

基于光束平差法的双目视觉里程计研究

罗杨宇¹, 刘宏林^{1,2}

(1. 中国科学院自动化研究所, 北京 100190; 2. 上海奥的斯电梯有限公司, 上海 201206)

摘要: 机器人自定位是实现机器人自动导航及其他智能行为的前提, 一种基于光束平差法的移动机器人双目视觉里程计可以有效地实现机器人自定位. 为此, 首先采用点模式匹配方法建立相邻图像之间的特征匹配关系, 根据立体视觉算法得到匹配点对的三维对应关系; 然后, 计算摄像机的相对运动参数, 并采用光束平差分段优化算法对其进行优化. 所提出的双目视觉里程计能够避免车轮半径变化、空转、打滑等对里程计测量精度的影响, 相对定位精度较高.

关键词: 移动机器人; 双目视觉里程计; 图像匹配

中图分类号: TP24

文献标志码: A

Research on binocular vision odometer based on bundle adjustment method

LUO Yang-yu¹, LIU Hong-lin^{1,2}

(1. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China; 2. Otis Elevator Co. Ltd., Shanghai 201206, China. Correspondent: LUO Yang-yu, E-mail: yangyu.luo@ia.ac.cn)

Abstract: Self-localization is the premise of navigation and other intelligent behavior for mobile robot. This paper proposes a kind of binocular vision odometer based on the bundle adjustment method. The feature point set is matched based on the point pattern matching method, and the corresponding 3D coordinate is calculated through 3D reconstruction. Then, the relative motion parameter is optimized through segmenting the image sequence into several subsequences according to overlapped regions. The performance of the binocular vision odometer does not depend on the count of rotating shaft, so that the effect of wheel radius, idle running and wheel slipping on the measurement precision can be avoided, and the relative positioning accuracy is higher.

Keywords: mobile robot; binocular vision odometer; image matching

0 引言

自定位的准确性是移动机器人实现导航任务的关键. 通常机器人自定位采用航迹推算的方法: 依靠车轮里程计和惯导信息, 通过累积测量给出机器人的大概位置. 这种自定位方法的最大问题是车轮打滑及地面不平整情况下的累积误差增大. 此外, 传感器精度降低、惯导漂移等情况下也会引起误差增大. 为降低自定位误差并提高定位对环境的鲁棒性, 人们将计算机视觉应用于机器人自定位, 利用摄像机采集的序列图像, 通过特征跟踪和相对运动估计出机器人自身的运动参数. 这种位姿估计方法类似于里程计的航迹推算, 也称为视觉里程计定位技术. 与传统定位方式相比, 视觉里程计采用被动视觉传感器, 结构简单、信息量大, 只依靠视觉输入信息即可完成定位, 无需场

景及运动的先验知识, 而且不存在传感器精度降低导致的误差, 适用于各种环境.

视觉里程计属于无地图定位, 介于相对定位与绝对定位之间, 实现方法主要有基于特征的离散处理方法和基于光流的连续处理方法. 基于光流的视觉里程计, 利用摄像机拍摄的时间序列图像来估计图像上各像素点的速度, 即光流场, 进而根据相对运动原理得到摄像机本身的运动参数. 常用的估计光流的方法有块相关^[1]和图像差分方法(如 Horn-Schunck 方法^[2]、Lucas-Kanade 方法^[3])等. 文献[4]以单目摄像机采集图像, 在室内及室外多种环境下对 Lucas-Kanade 光流方法计算机人位姿信息进行了实验验证. 文献[5]采用块相关方法计算光流, 实现全向移动机器人视觉里程计, 并通过实验验证了视觉里程计的位置精度高

收稿日期: 2015-03-30; 修回日期: 2015-11-27.

基金项目: 国家 863 计划项目(2007AA04Z227).

作者简介: 罗杨宇(1975—), 女, 副研究员, 博士, 从事智能服务机器人的研究; 刘宏林(1980—), 男, 高级工程师, 博士, 从事智能控制、智能机器人的研究.

于常用的轮轴编码器里程计. 文献[6]将卡尔曼滤波(KF)与差分光流计算方法相结合,用以实现移动机器人的速度测量. 光流场的计算基于图像灰度守恒这一假设条件,在许多实际环境中,由于物体遮挡、多光源和噪声等原因,这一条件难以满足,不能求解出精确的光流场. 同时,大多数光流场计算方法的复杂度较高,无法满足实时性要求,因此,基于特征跟踪的视觉里程计以其对环境条件变化鲁棒的优点受到了更广泛的关注^[7-13].

基于特征跟踪的视觉里程计,其思路是从相邻时刻的两帧或多帧图像中提取诸如角点、边缘、稳定多边形等显著性特征并建立特征对应关系,根据这些特征对应估计机器人的运动参数.

早在 1976 年, Hans Moravec 研制的“斯坦福推车”^[14]在导航过程中用相机采集图像,并使用特征跟踪算法测定机器人的移动距离,这可以被看作是最早的视觉里程计. 随后, Matthies 在卡耐基梅隆大学对这种自定位方式进行研究,提出了视觉里程计的概念并设计了从特征提取、特征匹配与跟踪到运动估计的视觉里程计框架,至今仍为大多数视觉里程计系统所遵循. 基于这一框架,文献[15]设计了基于特征跟踪的视觉里程计,该系统可以采用单目摄像机或立体视觉来实现,而且对于运动模式或场景结构都没有要求,利用视频序列来实现里程的测量. 文献[16]提出了一种立体视觉里程计,采用 Shi-Tomasi 方法^[17]提取图像角点特征,用归一化的互相关算法在连续的图像间进行特征跟踪,最后,利用最小化均方和函数在 3D 空间计算运动参数,并采用迭代最邻近点算法(ICP)^[18]得到最终的运动参数估计. 文献[19]设计了针对室内环境的视觉里程计,以 SUSAN 角点^[20]为特征,由仿射变换关系求解移动机器人运动参数. 文献[13]设计了一种同时适用于室内及室外的立体视觉里程计,以 CensurE 算法^[21]提取图像中的尺度不变特征,通过三维重建得到三维特征点,并以此为基础求解运动参数. 此外,随着同时定位与导航(SLAM)算法的飞速发展^[22-23],视觉里程计作为其重要组成部分得到了更多研究者的关注^[24-31]. 如文献[27]将视觉里程计与地图创建相结合,创建了基于机器人视角的全局地图,采用里程计信息提高地图中路标的定位精度,同时利用全局地图降低视觉里程计的累计误差,取得了良好的地图创建及定位结果.

摄像机在静态环境中移动,得到不同时刻拍摄的多幅图像. 假设这些图像是同一刚性物体的投影,则可由图像特征对应关系估计出摄像机的运动参数. 在计算机视觉中,这一过程称为运动分析或由运动重建物体结构(SFM). 当空间物体结构参数未知时,

SFM 问题可分为两类:单目视觉下的二维特征对应和多目视觉下的三维特征对应.

采用二维特征对应关系估计相对运动需要给定先验的空间尺度信息,这为单目视觉里程计的实现带来一定的不便.

三维特征对应关系下求解运动估计问题的一般方法为:首先采用双目或多目摄像机三维重建得到空间物体的三维数据;然后由二维图像特征对应关系建立空间物体的三维特征对应,进而进行三维运动问题求解. 三维数据的信息量远高于二维图像,因此三维运动估计问题的求解大为简化. 但是,由于立体视觉中三维重建过程对像素误差有放大作用,三维运动估计的结果对图像点误差非常敏感,需要采取一定的措施对三维重建结果进行优化以提高运动估计精度.

本文首先对视觉里程计进行综述;然后介绍点模式匹配方法以及基于光束平差法的双目视觉里程计. 实验结果表明,光束平差法双目视觉里程计具有较高的定位精度.

1 点模式匹配

1.1 点模式匹配迭代算法

本文中图像检测对象为天花板,这种场景包含丰富的直线边缘,而直线的交点是比较明显的角点特征. 本文采用由直线边缘的交点构成的角点作为特征点,这类特征点具有显著性高、易于识别、定位准确及在视觉条件变化下鲁棒性较强的优势,适合于视觉定位任务中对特征提取精度和效率的要求. 基于边缘的角点特征提取过程分为 3 步:首先,在图像中提取反映灰度变化的边缘;然后,将这些边缘连接为具有意义的线类特征;最后,通过检测直线的交点提取出连接点式角点特征.

室内天花板图像纹理简单且结构重复性高,在这一情况下, SIFT 描述子与 KLT 特征跟踪方法等都不能取得满意的效果. 为得到稳定、准确的特征点匹配关系,本文提出一种基于点模式的匹配方法. 点模式匹配是从点特征匹配问题中抽象出来的更基本的问题,其思路是将图像匹配转化为从两个满足一定映射关系的点集中找出最优匹配点对问题. 对于已知具有一定变换关系的两幅图像,基于空间变换关系求解的算法,按照某种度量准则对点集间空间变换参数进行估计,通过迭代获得最优的变换参数,从而得到点集之间的匹配关系. 本文中,摄像头采集的图像为天花板,当天花板上各物体高度远小于摄像机与天花板的距离时,可将天花板视为平面场景,于是两幅图像之间的变换参数满足射影变换.

假设从两幅图像提取的特征点集分别为

$$P = \{p_1, p_2, \dots, p_n\}, Q = \{q_1, q_2, \dots, q_m\},$$

其中 p_i 和 q_j 表示特征点的二维像素坐标. 已知 P 与 Q 之间满足射影变换关系, 即对于其中某一对对应点 (p_i, q_i) , 下式成立:

$$s_i \tilde{q}_i = H \tilde{p}_i. \quad (1)$$

其中: \tilde{p}_i, \tilde{q}_i 分别表示 p_i, q_i 的齐次坐标; $H \in R^{3 \times 3}$ 表示射影变换关系, 称为单应矩阵; s_i 为常系数.

由式 (1), 特征点匹配问题便转化为在匹配关系未知情况下变换矩阵 H 的估计问题. 由于匹配关系未知, 点集匹配的思路是首先找到可能的对应点, 估计 H 矩阵的初值, 然后建立优化函数, 通过迭代的方法调整 H 矩阵各参数, 直至收敛条件满足, 完成匹配.

初始对应关系的建立可采用最近点原则: 设点集 P 中点的数目小于点集 Q , 定义点集 P 中任意一点 p_i 到点集 Q 的最近距离为

$$d(p_i, Q) = \min_{q_j} \|p_i - q_j\|. \quad (2)$$

由最近距离决定点集 P 中每一点在 Q 中的对应点, 若 P 中有两个点对应 Q 中同一点, 则选择距离最小的点为对应点. 考虑 P 与 Q 为不完全匹配情况, 且两者都存在出格点, 于是当 P 中可匹配点在最近距离准则下都找到相应的对应点时, P 中出格点由于不存在可匹配点而会寻找到 Q 中出格点, 其对应距离大于可匹配点的最近距离. 因此, 可设定距离阈值 σ , 当最近点之间距离大于 σ 时, 将对应点作为出格点处理. 设 $n \leq m$, P 与 Q 有 k 对匹配点, 通过上述划分出格点处理, 点匹配问题可转化为求解如下最小化问题:

$$\varepsilon = \min_{H, n} \left[\sum_{i=1}^k \|\tilde{q}_i - H \tilde{p}_i\|^2 + (n - k) \sigma^2 \right]. \quad (3)$$

匹配过程中, 随着图像之间畸变关系的变化, 可匹配对应点之间的最近距离也会随之发生变化, 因此阈值 σ 的选择非常关键: 较大的 σ 值会导致将出格点也视为真实对应点, 使式 (3) 无法收敛; 较小的 σ 值会将可匹配点错误地认为是出格点, 使式 (3) 收敛缓慢. 因此, 本文采取以下方法估计 σ 值:

借鉴高斯尺度空间方法, 采用零均值规格化方法^[32]对点集的中心位置和尺度进行配准: 对两幅图像的坐标分别进行平移, 使得点集的质心移至坐标原点; 然后选择合适的缩放因子使点集中各点到原点的平均距离为 1. 由规格化后的点集求得 P 中各点相对于 Q 的最近距离, 选择合适的 σ 值, 使 P 中最少有 k 个点满足最小距离小于 σ . $k \geq 4$, 因为 H 矩阵的估计最少需要 4 对对应点.

估计 σ 后, 点集被分割为出格点和对应点两个子集. 对对应点子集再次使用零均值规格化, 一般经过规格化后点集在坐标系的 4 个象限都有点分布, 于

是可以选择每个象限子点集质心作为初始匹配点估计 H 矩阵的初值, 然后按照式 (3) 迭代求解 H 并得到匹配对.

图 1 为实际图片迭代过程, 过程中 $\sigma = 0.24$, 经过 3 次迭代, 算法取得了较好的匹配效果.

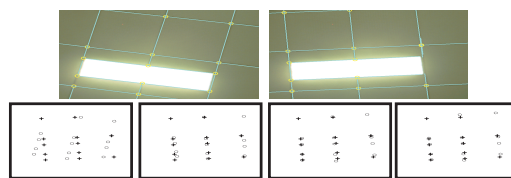


图 1 点模式匹配迭代过程 ($\sigma = 0.24$)

图 1 点模式匹配迭代过程 ($\sigma = 0.24$)

基于点模式的匹配方式只利用特征的位置信息, 而不考虑特征点及其邻域的灰度信息, 在图像光照变化下匹配稳定、准确. 与基于特征及灰度相关的匹配方法相比, 点模式匹配方法的另一显著优点是算法简单、匹配速度快. 其主要缺点在于要求可匹配点之间只存在较小的位置偏移. 本文采用的摄像机取样速度为 30 帧/s, 图像序列中相邻帧之间位置变化较小, 满足点模式匹配的要求.

1.2 特征点跟踪过程中定位及匹配精度

特征点跟踪过程中的定位误差及匹配准确性是影响视觉自定位精度的主要因素. 本文以边缘连接点作为特征, 采用点模式方法实现匹配跟踪, 并与基于 Harris 角点的 KLT 特征跟踪方法进行比较.

特征点定位及匹配精度可用重投影误差表征. 同一平面场景投影形成的两幅平面图像之间满足射影关系, 图像上任意一对匹配点满足下式:

$$s_i \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = H \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}, H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}. \quad (4)$$

其中: H 表示平面图像之间的射影关系, s_i 为常系数. 于是可定义匹配点之间重投影误差为

$$\varepsilon_i = \left(\left(x'_i - \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 + \left(y'_i - \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 \right)^{1/2}. \quad (5)$$

当两幅平面图像之间的射影关系矩阵 H 已知时, 重投影误差 ε_i 由特征点定位误差及匹配准确性决定, 特征点定位不准确或错误的匹配关系都会引起 ε_i 增大. 因此, 可用 ε_i 的均值及方差表征不同特征点定位及匹配方法的优良程度.

为得到图像平面之间的射影关系矩阵 H , 本文实验条件设置如图 2 所示 (采用标定块顶点匹配计算平面图像之间的射影关系, 并计算匹配对重投影误差, “□”标记标定块顶点, “○”标记 KLT 角点, “+”标记连接点): 在天花板粘贴棋盘式标定图案, 并采用

Harris 特征检测方法提取标定图案上各块的顶点. 由于标定块结构的特殊性, 顶点的精确位置可以求得; 在图像特征点位移变化较小的前提下, 采用点模式匹配方法可避免图像点灰度变化对匹配精度的影响, 并求出匹配点集之间的射影关系矩阵 H .

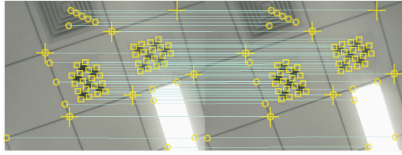
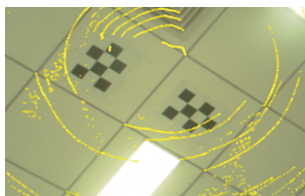
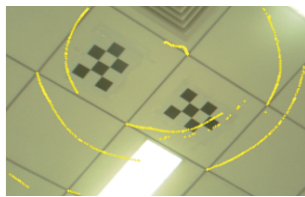


图 2 实验条件设置

求得图像之间的射影关系 H 后, 在每幅图像标定图案之外区域进行特征点提取与匹配. 为便于比较, 摄像机旋转 90° , 在此过程中采集多帧图像, 对同一图像序列分别进行 KLT 跟踪和由点模式匹配实现的连接点特征跟踪, 并计算各帧之间的重投影误差. 图 3 为特征跟踪效果, 图 4 为特征跟踪过程中各帧之间重投影误差均值与均方差数据. 由于日光灯附近灰度比较均匀, KLT 方法跟踪到较多的离散点(图 3 左图中日光灯附近离散轨迹), 这些点定位精度差, 使重投影误差均值较大, 且误差比较分散, 说明存在错误的匹配关系. 与此相比, 连接点特征与点模式匹配受灰度对比度变化影响小, 重投影误差较小且分布稳定, 更适于对特征点定位精度要求较高的视觉定位任务.



(a) KLT 跟踪



(b) 连接点特征跟踪

图 3 特征点跟踪

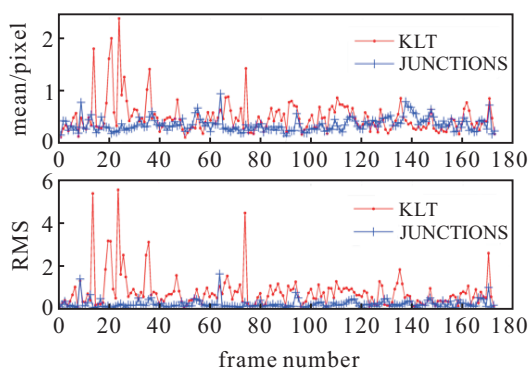


图 4 KLT 和连接点跟踪过程中重投影误差均值及均方差

2 光束平差法优化双目运动估计

单目视觉运动参数估计中需要给定先验的尺度信息(如摄像机光心到平面的距离 d), 而这一先验知识与场景信息相关, 在摄像机运动过程中可能会发生漂移, 这也会引入一定的误差. 随着立体视觉系统的研究深入和应用范围的扩大, 通过双目摄像机可直接重建物体的三维结构, 获得场景的深度信息, 从而使摄像机运动估计问题大为简化.

令 $P = [X \ Y \ Z]^T$ 为空间一点, $p_i = [x_i \ y_i \ z_i]^T$ 为运动前该点在摄像机坐标系下的坐标, 运动后的坐标为 $p'_i = [x'_i \ y'_i \ z'_i]$. 摄像机运动用旋转矩阵 R 和平移矢量 t 描述, 由空间点对应关系可得

$$p_i = R p'_i + t. \tag{6}$$

当有多组点对应(最少 3 对非共线点)时, 可用最小二乘法由式 (6) 解出摄像机运动参数.

上述采用最小二乘法求解摄像机的相对运动参数时, 假设数据不包含噪声. 实际上, 由于图像点定位误差的存在, 三维重建数据也存在一定的误差. 在具有误差的情况下, 一般以最小二乘法的结果为初值, 采用一定的优化准则对结果进行迭代优化, 以提高其精确度. 本文采用图像点和三维点之间的重投影误差对摄像机的相对运动参数进行优化.

2.1 投影关系优化运动估计

根据摄像机针孔成像模型, 空间点与对应图像点之间投影关系定义为

$$s_i \tilde{p}_i = M \tilde{q}_i. \tag{7}$$

其中: \tilde{q}_i 和 \tilde{p}_i 分别表示空间点 q_i 和图像点 p_i 对应的齐次坐标; M 为投影矩阵, 由摄像机内参数矩阵 K 、机器人运动中心坐标系与摄像机的相对位置数 R 、 t 构成, 即

$$M = [K \ 0] \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix}. \tag{8}$$

当参数 R 、 t 的估计存在误差时, 三维点经投影矩阵计算得到的重投影点与图像点不能完全重合. 定义重投影误差

$$d^2(\tilde{p}_i, M \tilde{q}_i) = \|\tilde{p}_i - M \tilde{q}_i\|^2. \tag{9}$$

采用投影关系优化双目运动估计时, 首先对运动前后采集的两幅图像分别进行三维重建, 得到三维点对应关系, 采用奇异值分解方法求得运动参数的初始估计; 然后, 将运动前图像重建的三维点视作空间点, 以其与运动后图像上对应点的重投影距离作为优化准则求解运动参数的最佳解.

当运动过程中摄像机采样速度较快时, 序列图像中相邻多帧由同一三维物体投影而成, 形成如图 5 所示的图像重叠. 将序列图像以重叠区域分段, 在各段

中分别对重叠区域图像特征进行跟踪,并以段首帧图像特征的三维重建数据作为空间点坐标的初始估计,后续各帧依次相对于此空间点求解运动参数的初值,并采用投影关系对运动参数和空间点坐标同时进行优化.这一过程称为序列图像的分段优化,具体优化实施流程如下.

Step 1: 在第 0 帧进行三维重建,得到三维点 q_{0i} , $i = 1, 2, \dots, n$;

Step 2: 在第 1 帧进行三维重建,得到与 q_{0i} 对应的三维点 q_{1i} 及其匹配图像点 p_{1i} , 计算运动参数 R_{10} 和 t_{10} , 以重投影误差 $\sum d^2(\tilde{p}_{1i}, M_{10}\tilde{q}_{0i})$ 最小为准则优化 $(R_{10}, t_{10}q_{0i})$;

Step 3: 获取第 2 帧图像,通过与第 1 帧各点匹配关系得到与 q_{0i} 对应的 q_{1i} 和 p_{1i} , 计算运动参数 R_{10} 和 t_{10} , 以重投影误差的和 $\sum (d^2(\tilde{p}_{2i}, M_{20}\tilde{q}_{0i}) + d^2(\tilde{p}_{1i}, M_{10}\tilde{q}_{0i}))$ 最小为准则优化 $(R_{20}, t_{20}, R_{10}, t_{10}, q_{0i})$.

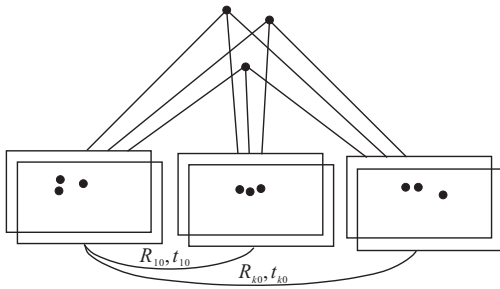


图 5 序列图像重叠情况

上述分段优化方法中,在采用空间点与图像点得到的投影关系对运动参数进行优化的同时,利用在不同摄像机位置观察得到的多帧图像数据对空间点坐标的估计值进行优化,从而降低了运动估计中三维重建误差的影响,提高了运动参数的估计精度.

2.2 光束平差法

以上优化流程中在每一步同时对运动参数及第 0 帧三维点坐标进行优化,当图像点较多及参与优化的帧数较多时,可采用光束平差法计算,以提高计算效率和运动估计精度.

在解析摄影测量中,将外方位元素和模型点坐标的计算放在一个整体内进行,称其为光束法.光束平差(BA)是以共线方程式作为数学模型,像点的像平面坐标测量值是未知数的非线性函数,经线性化方法后按照最小二乘原理进行计算.假设给定一系列在图像中观测到的一组对应点集相应的三维坐标的初始估计,以及关于每张图像对应摄像机姿态的初始估计,则光束法平差是一个大的最优化问题,包括同时优化三维结构和摄像机姿态.为了获得一个在特定的假设下最优化的重建,考虑与观测的图像特征有关的噪声:如果图像误差满足均值为零的正态分布,

则 BA 是最大似然法估计.目前,人们常用 Levenberg-Marquardt 非线性最小二乘算法的常规实现来解决与 BA 相联系的稀疏的大规模的优化问题.

记测量数据向量为

$$X =$$

$$(X_{11}^T, \dots, X_{1n}^T, X_{21}^T, \dots, X_{2n}^T, \dots, X_{m1}^T, \dots, X_{mn}^T),$$

其中 X_{ij} 表示第 i 幅图像上第 j 个点.

记参数向量为

$$P = (p_1, p_2, \dots, p_n, Rt_1, Rt_2, \dots, Rt_m).$$

其中: p_1, p_2, \dots, p_n 表示第 0 幅图像对应的三维点; Rt_j 表示第 j 幅图像相对于第 0 幅图像的对应摄像机姿态.

由投影关系,可得到

$$X_{ij} = f(p_j, Rt_i), \quad (10)$$

即测量向量的估计值 $\hat{X} = f(P)$.

记测量值残差为 $\varepsilon = X - \hat{X}$, 根据非线性最小二乘法,有

$$J^T \Sigma_X^{-1} J \delta = J^T \Sigma_X^{-1} \varepsilon. \quad (11)$$

其中: J 为 f 的雅克比矩阵, δ 为每次迭代中参数向量 P 的更新值, $\Sigma_X^{-1} \text{diag}(\Sigma_{X_{11}}^{-1}, \Sigma_{X_{12}}^{-1}, \dots, \Sigma_{X_{mn}}^{-1})$ 表示测量向量的协方差矩阵.

注意到 $\frac{\partial \hat{X}_{ij}}{\partial p_k} = 0, \forall j \neq k$, 且 $\frac{\partial \hat{X}_{ij}}{\partial Rt_k} = 0, \forall i \neq k$, 并记 $A_{ij} = \frac{\partial X_{ij}}{\partial p_j}$, $B_{ij} = \frac{\partial X_{ij}}{\partial Rt_i}$. 雅克比矩阵 J 可表示为

$$J = \begin{bmatrix} A_{11} & 0 & \cdots & 0 & B_{11} & 0 & \cdots & 0 \\ 0 & A_{12} & & & B_{12} & 0 & & \\ \vdots & & \ddots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{1n} & B_{1n} & 0 & \cdots & 0 \\ A_{21} & 0 & \cdots & 0 & 0 & B_{21} & \cdots & 0 \\ 0 & A_{22} & & & 0 & B_{22} & & \\ \vdots & & \ddots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{2n} & 0 & B_{2n} & \cdots & 0 \\ A_{m1} & 0 & \cdots & 0 & 0 & 0 & \cdots & B_{m1} \\ 0 & A_{m2} & & & 0 & 0 & & B_{m2} \\ \vdots & & \ddots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{mn} & 0 & 0 & \cdots & B_{mn} \end{bmatrix}. \quad (12)$$

记

$$U_j = \sum_{i=1}^m A_{ij}^T \Sigma_{X_{ij}}^{-1} A_{ij}, \quad V_i = \sum_{j=1}^n B_{ij}^T \Sigma_{X_{ij}}^{-1} B_{ij},$$

$$W_{ij} = A_{ij}^T \Sigma_{X_{ij}}^{-1} B_{ij},$$

并记

$$\delta = [\delta_p \quad \delta_{Rt}]^T,$$

表 1 机器人近似直线运动位置测量数据

序号	实际位置 $(x, y)/\text{cm}$	双目测量值 $(x, y, \text{error})/\text{cm}$	BA 双目测量值 $(x, y, \text{error})/\text{cm}$
1	-0.41, 0.78	-0.42, 1.78, 1.00	-0.34, 0.70, 0.10
5	-3.02, 16.77	-3.59, 19.48, 2.76	-2.76, 16.71, 0.27
9	-5.00, 39.45	-5.74, 41.38, 2.07	-4.61, 39.51, 0.40
13	-6.19, 61.67	-6.68, 64.25, 2.62	-5.66, 61.69, 0.54
17	-6.93, 83.24	-6.69, 84.89, 1.67	-6.51, 83.47, 0.48
21	-7.31, 105.80	-7.29, 109.76, 3.96	-7.12, 105.81, 0.19
25	-7.14, 127.39	-8.81, 131.46, 4.40	-6.97, 127.25, 0.23

表 2 机器人近似直线运动姿态测量数据

序号	实际姿态 $\theta/(\text{°})$	双目测量值 $(\theta, \text{error})/(\text{°})$	BA 双目测量值 $(\theta, \text{error})/(\text{°})$
1	1.18	1.16, -0.02	1.17, -0.02
5	5.32	5.48, 0.16	5.48, 0.16
9	4.47	4.67, 0.20	4.67, 0.20
13	3.25	3.21, -0.03	3.33, 0.09
17	2.17	2.26, 0.09	2.36, 0.19
21	1.13	1.18, 0.05	1.32, 0.19
25	0.39	0.44, 0.04	0.71, 0.31

令机器人做半径为 58 cm 的圆周运动 2 圈, 机器人累计运动 729 cm, 共获取 509 组轨迹点数据. 图 8 和图 9 分别为 3 种里程计的位置和姿态及对应误差. 为清晰展示视觉里程计累积误差, 表 3 和表 4 给出了圆周运动第 2 圈等间隔抽取的部分位姿测量数据 (仅列出常规双目视觉里程计和基于光束平差法双目视觉里程计的实验数据). 圆周运动中, 单目视觉里程计位置测量误差均值为 12.82 cm, 最大值为 22.81 cm, 姿态测量误差均值为 0.29°, 最大值为 0.94°; 双目视觉里程计位置测量误差均值为 12.73 cm, 最大值为 25.81 cm, 姿态测量误差均值为 0.69°, 最大值为 1.96°; 光束平差法优化双目视觉里程计位置测量误差均值为 5.51 cm, 最大误差为 10.88 cm, 姿态测量误差均值为 1.22°, 最大值为 2.39°.

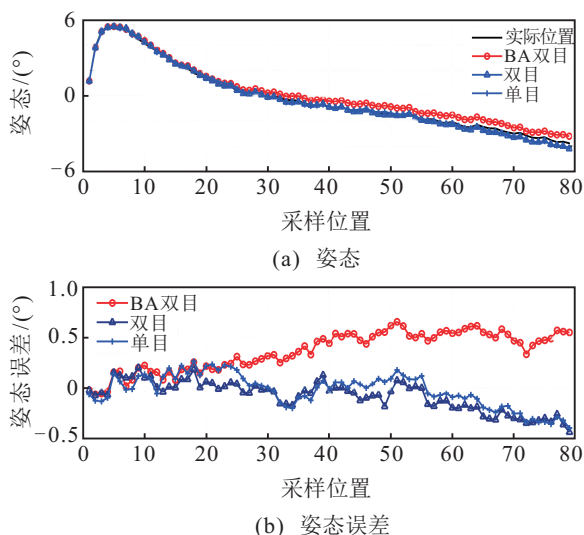


图 8 机器人近似直线运动姿态测量值和姿态误差

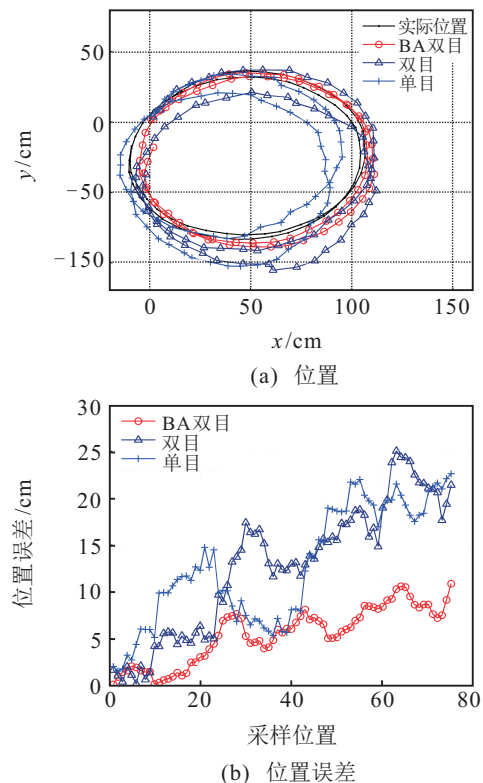


图 9 机器人圆周运动位置轨迹和位置误差

从实验结果可以看出: 在位置测量方面, 光束平差法优化里程计测量结果与实际位置基本吻合, 误差较小; 而在姿态的测量中, 光束平差法优化双目视觉里程计的测量误差在可接受范围内. 综合考虑位置与姿态的测量结果, 光束平差法优化双目视觉里程计优于其他两种方法. 3 种视觉里程计测量结果的差异性缘于实验中运动估计采用的特征点数目. 图 10 为圆周运动旋转过程中特征点跟踪效果, 由于天花板纹理

表 3 机器人圆周运动第 2 圈位置测量数据

序号	实际位置 $(x, y)/\text{cm}$	双目测量值 $(x, y, \text{error})/\text{cm}$	BA 双目测量值 $(x, y, \text{error})/\text{cm}$
39	-3.33, -0.72	2.64, -11.46, 12.29	2.12, 2.01, 6.10
41	7.82, 15.04	14.18, 3.56, 13.13	14.49, 16.11, 6.75
43	23.12, 25.69	27.65, 13.61, 12.90	31.25, 26.13, 8.14
45	42.13, 31.65	50.62, 21.12, 13.52	49.35, 32.55, 7.27
47	60.15, 30.91	68.47, 17.65, 15.66	66.70, 31.34, 6.57
49	76.27, 24.44	81.66, 9.52, 15.87	81.34, 24.96, 5.10
51	91.65, 12.23	99.61, -3.14, 17.31	97.41, 12.72, 5.78

表 4 机器人圆周运动第 2 圈姿态测量数据

序号	实际姿态 $\theta/(\circ)$	双目测量值 $(\theta, \text{error})/(\circ)$	BA 双目测量值 $(\theta, \text{error})/(\circ)$
39	-9.93	-9.77, 0.16	10.80, 0.87
41	-29.01	-28.59, 0.42	30.18, 1.17
43	-46.92	-46.53, 0.40	48.13, 1.20
45	-66.56	-65.89, 0.67	67.87, 1.31
47	-86.09	-85.62, 0.47	87.53, 1.44
49	-104.31	-103.62, 0.69	105.45, 1.14
51	-123.63	-122.78, 0.85	125.01, 1.38

简单, 可用于运动估计的特征点为 12 个或 13 个. 特征点数目对运动估计结果的影响可简单解释如下: 以 4 个点为例, 单目视觉运动估计中用 8 个约束优化单应矩阵分解产生的 8 个参数, 其中旋转参数与其他参数的耦合性不强, 而平移参数估计要受到平面法向量估计结果的影响; 双目视觉运动估计中 4 对对应点可提供 12 个约束, 用以优化 6 个参数; 采用光束平差法分段优化运动估计中, 多帧图像同时参与优化, 相当于增加了特征点数. 由此可见, 进一步提高里程计精度的有效方法是增加图像提取特征点的数量.

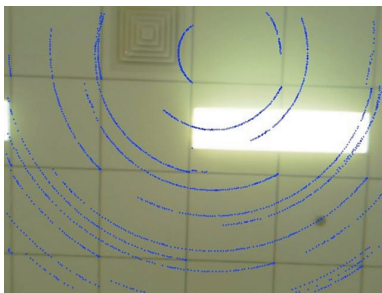


图 10 旋转过程特征点跟踪

4 结 论

本文讨论了基于图像匹配的运动估计问题与视觉里程计实现. 采用点模式匹配方法建立图像对的特征匹配关系; 利用序列图像多帧之间存在重叠区域这一有利情况, 采用投影距离对运动参数和三维点坐标同时进行优化, 以进一步降低三维重建误差对估计结果的影响, 并给出了优化过程的光束平差法实现; 通过手眼标定将摄像机坐标系下的相对运动转化为移动机器人运动中心坐标系的运动, 通过运动累计构成

了视觉里程计. 最后, 对单目里程计、采用三维点对应运动估计的双目里程计, 以及光束平差法优化的双目里程计定位精度进行了比较, 实验表明, 光束平差法双目视觉里程计具有较高的定位精度. 对于非平面天花板假设下的室内机器人视觉里程计尚待深入研究.

参考文献(References)

- [1] Gyaourova A, Kamath C, Cheung S C. Block matching for object tracking[R]. Livermore CA: Lawrence Livermore National Laboratory, 2003.
- [2] Horn B, Schunck B. Determining optical flow[J]. Artificial Intelligence, 1981, 17(1/2/3): 185-204.
- [3] Lucas B, Kanade T. An iterative image registration technique with an application to stereo vision[C]. Proc of Imaging Understanding Workshop. San Francisco, 1981: 121-130.
- [4] Campbell J, Sukthankar R, Nourbakhsh I. Techniques for evaluating optical flow for visual odometry in extreme terrain[C]. IEEE/RSJ Int Conf on Intelligent Robots and Systems. Pittsburgh, 2004, 4: 3704-3711.
- [5] Sanada A, Ishii K, Yagi T. Self-localization of omnidirectional mobile robot using silicon retina based optical flow sensor[C]. World Automation Congress(WAC). Hukuoka, 2010: 1-6.
- [6] Song X, Seneviratne L, Althoefer K. A Kalman filter-integrated optical flow method for velocity sensing of mobile robots[J]. IEEE/ASME Trans on Mechatronics, 2011, 16 (3): 551-563.

- [7] Agrawal M, Konolige K. Rough terrain visual odometry[DB/OL]. https://www.researchgate.net/publication/228958153_Rough_terrain_visual_odometry.
- [8] Garcia-Garcia R, Sotelo M A, Parra I, et al. 3D visual odometry for road vehicles[J]. *J of Intelligent & Robotic Systems*, 2008, 51(1): 113-134.
- [9] Fang H, Yang M, Yang R Q, et al. Ground-texture-based localization for intelligent vehicles[J]. *IEEE Trans on Intelligent Transportation Systems*, 2009, 10(3): 463-468.
- [10] Comport A I, Malis E, Rives P. Real-time quadrifocal visual odometry[J]. *Int J of Robotics Research*, 2010, 29(2/3): 245-266.
- [11] De X, Liwei H, Min T, et al. Ceiling-based visual positioning for an indoor mobile robot with monocular vision[J]. *IEEE Trans on Industrial Electronics*, 2009, 56(5): 1617-1628.
- [12] Jiaolong Y, Lei C, Wei L. Monocular vision based robot self-localization[C]. *IEEE Int Conf on Robotics and Biomimetics*. Beijing, 2010: 1189-1193.
- [13] Konolige K, Agrawal M, Sola J. Large-scale visual odometry for rough terrain[J]. *Robotics Research*, 2011, 66: 201-212.
- [14] Moravec H. The cmu rover[C]. *Proc of the National Conf of Artificial Intelligence*. Pittsburgh, 1982: 377-380.
- [15] Nister D, Naroditsky O, Bergen J. Visual odometry[C]. *IEEE Conf on Computer Vision and Pattern Recognition*. Washington, 2004, 1: 652-659.
- [16] Milella A, Siegwart R. Stereo-based ego-motion estimation using pixel tracking and iterative closest point[C]. *IEEE Int Conf on Computer Vision Systems*. New York, 2006: 21-22.
- [17] Shi J, Tomasi C. Good features to track[C]. *IEEE Conf on Computer Vision and Pattern Recognition*. Seattle, 1994: 593-600.
- [18] Besl P, McKay H. A method for registration of 3-D shapes[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1992, 14(2): 239-256.
- [19] Nguyen V, Jeong M, Ahn S, et al. A robust localization method for mobile robots based on ceiling landmarks[C]. *Int Conf on Modeling Decisions for Artificial Intelligence*. Kitakyushu, 2007: 422-430.
- [20] Smith S, Brady J. SUSAN-A new approach to low level image processing[J]. *Int J of Computer Vision*, 1997, 23(1): 45-78.
- [21] Agrawal M, Konolige K, Blas M. Censure: Center surround extremas for realtime feature detection and matching[C]. *European Conf on Computer Vision*. Marseille, 2008: 102-115.
- [22] Won D H, Chun S, Sung S, et al. Improving mobile robot navigation performance using vision based SLAM and distributed filters[C]. *Int Conf on Control, Automation and Systems*. Seoul, 2008: 169-174.
- [23] Munguia R, Grau A. Monocular SLAM for visual odometry[C]. *IEEE Int Symposium on Intelligent Signal Proc*. Alcala de Henares, 2007: 443-448.
- [24] Pink O, Moosmann F, Bachmann A. Visual features for vehicle localization and ego-motion estimation[C]. *IEEE Intelligent Vehicles Symposium*. Xi'an, 2009: 254-260.
- [25] Williams B, Reid I. On combining visual SLAM and visual odometry[C]. *IEEE Int Conf on Robotics and Automation Anchorage*. Alaska, 2010: 3494-3500.
- [26] Hwang S Y, Song J B. Monocular vision-based SLAM in indoor environment using corner, lamp, and door features from upward-looking camera[J]. *IEEE Trans on Industrial Electronics*, 2011, 10: 4804-4812.
- [27] Lee Y-J, Yim B-D, Song J-B. Mobile robot localization based on effective combination of vision and range sensors[J]. *Int J of Control, Automation and Systems*, 2009, 7(1): 97-104.
- [28] Jwu-Sheng Hu, Ming-Yuan Chen. A sliding-window visual-IMU odometer based on tri-focal tensor geometry[C]. *2014 IEEE Int Conf on Robotics and Automation(ICRA)*. Hong Kong, 2014: 3963-3968.
- [29] Guo C X, Mirzaei F M, Roumeliotis S I. An analytical least-squares solution to the odometer-camera extrinsic calibration problem[C]. *2012 IEEE Int Conf on Robotics and Automation(ICRA)*. St Paul, 2012: 3962-3968.
- [30] Guyue Zhou, Jiaxin Ye, Wei Ren, et al. On-board inertial-assisted visual odometer on an embedded system[C]. *2014 IEEE Int Conf on Robotics and Automation(ICRA)*. Hong Kong, 2014: 2602-2608.
- [31] Yu-Cheol Lee, Seung-Hwan Park. 3D map building method with mobile mapping system in indoor environments[C]. *2013 16th Int Conf on Advanced Robotics(ICAR)*. Montevideo, 2013: 1-7.
- [32] Han J, Kamber M. Data mining: Concepts and techniques[M]. Burlington: Morgan Kaufmann, 2006: 191-257.