

引用格式:李卓,刘洁瑜,周伟.基于几何约束的视觉闭环检测和位姿优化[J].电光与控制,2018,25(5):79-83. LI Z, LIU J Y, ZHOU W. Visual loop closure detection and pose optimization based on geometric constraints[J]. Electronics Optics & Control, 2018, 25(5):79-83.

基于几何约束的视觉闭环检测和位姿优化

李卓, 刘洁瑜, 周伟
(火箭军工程大学导弹学院, 西安 710025)

摘要:为实现视觉里程计在大规模复杂环境下的准确导航,消除载体位姿估计的累计误差,提出了一种基于几何约束的视觉闭环检测和位姿优化方法。首先建立基于 LATCH 二进制描述子的视觉词袋(BoVW),并采用视觉单词向量描述图像进而归一化相似度的闭环检测方法,然后对闭环候选关键帧和当前关键帧进行基于 RANSAC-HORN 运动估计的闭环验证,最后将闭环关键帧的局部地图点投影到当前帧并最小化重投影误差以使得位姿得到优化。数据集实验表明,提出的闭环检测和位姿优化算法能够有效地对闭环进行准确检测和验证,并对视觉里程计的误差积累进行闭环位姿优化,提高了视觉导航精度。

关键词:视觉里程计;几何约束;闭环检测;位姿优化

中图分类号: TP242 **文献标志码:** A **doi:**10.3969/j.issn.1671-637X.2018.05.016

Visual Loop Closure Detection and Pose Optimization Based on Geometric Constraints

LI Zhuo, LIU Jie-yu, ZHOU Wei

(Missile Academy, Rocket Force University of Engineering, Xi'an 710025, China)

Abstract: In order to realize the accurate navigation of visual odometers in the large-scale complex environment and eliminate the cumulative error of pose estimation, this paper presents a method of visual loop closure detection and pose optimization based on geometric constraints. Firstly, a Bag of Visual Words (BoVW) based on LATCH descriptor is set up, and the loop closure detection method is adopted with the images described by the visual word vector and the similarity degree normalized. Then, the loop closure between the candidate keyframe and the current keyframe is checked based on RANSAC-HORN motion estimation. Finally, the local map points of the loop closure keyframe are projected to the current frame, and the re-projection error is minimized to optimize the pose. The experimental results show that the loop closure detection and pose optimization algorithm proposed in this paper can effectively and accurately detect and verify the loop closure, carry out the loop closure pose optimization of the error accumulation of the visual odometer and improve the accuracy of visual navigation.

Key words: visual odometer; geometric constraints; loop closure detection; pose optimization

0 引言

视觉里程计(Visual Odometer, VO)仅利用摄像机的输入图像对载体的位姿进行估计^[1],近十几年来已广泛应用于各类机器人自主导航^[2-3]。VO一般将局部运动估计级联得出摄像机的全局导航信息,从而使得误差不断积累繁殖,在长时间大范围情况下导航精度不高。

为实现大规模复杂环境下的准确导航,消除载体位姿估计的累计误差,闭环检测成为了视觉导航领域的基础问题和研究热点。所谓闭环检测就是为了使相机识别出曾经到达过的场景,进而校正累积误差。

由于闭环检测应用场景越来越复杂,环境规模越来越庞大,其对图像信息的压缩要求也越来越高,因此满足此要求的基于图像特征相似性的闭环检测在实际应用中占据了主导地位^[4],且大多采用视觉词袋(BoVW)^[5-9]技术。其中:ANGELI等^[5]研究了增量式视觉词典构建问题;CUMMINS等^[6-7]采用Chow-Liu树继承关系的闭环概率方法;LABBE等^[8-9]设计了内存管理机制,提

收稿日期:2017-06-14 修回日期:2018-01-04

基金项目:国家自然科学基金(61503390,61503392)

作者简介:李卓(1992—),男,河北保定人,硕士生,研究方向为视觉导航。

高了实时性。由于采用不同的特征提取和描述算子,基于SURF算子的FAB-MAP算法^[10]无法保证实时性; GALVEZ-LOPEZ提出了基于FAST检测算子和BRIEF特征算子的二进制视觉词袋技术,计算效率得到了显著提升^[11]; MUR-ARTAL等^[12]采用具有旋转不变性的ORB特征构建词袋,保证实时性的同时提高了稳定性; 2016年LEVI^[13]提出了LATCH二进制描述子,将以往二进制描述子的像素点对的比对转变为三元组图像块的F范数比对,在保证二进制描述子实时性的同时提高了鲁棒性。另外在闭环验证方面,基于文献[11]的DBoW2库只考虑图像中的特征描述,缺乏特征的几何约束。

基于以上问题与分析,本文建立基于图像关键帧数据库,并融合特征跟踪率与RGB直方图匹配的关键帧选取方法,降低信息冗余。采用视觉词袋查找当前帧的候选闭环关键帧,利用RANSAC-HORN运动估计方法来验证闭环,最终进行当前帧位姿优化。在KITTI数据集进行仿真实验,验证了该闭环方法的有效性,进一步提高了VO精度。

1 LATCH二进制描述子与视觉词袋

1.1 LATCH二进制描述子

通常二进制描述子由抽样模式、方向补偿以及抽样比对3部分组成,目前的几种二进制描述子在抽样的策略和辨别能力上均有所不同。

先前的二进制描述子以检测到的特征点为中心选取一定大小的图像块 W ,一个二进制描述子 b_w 由 T 对抽样坐标序列 $S = \{s_t\}_{t=1, \dots, n} = \{[p_{t,1}, p_{t,2}]\}_{t=1, \dots, n}$ 组成,其中, $p_{t,1} = (x_{t,1}, y_{t,1})$ 和 $p_{t,2} = (x_{t,2}, y_{t,2})$ 定义在 W 坐标系。索引 t 既与 W 中的一对坐标关联,又与高斯光滑核 $\sigma_t = (\sigma_{t,1}, \sigma_{t,2})_{t=1, \dots, n}$ 关联。对于每一抽样对 s_t ,比较 $p_{t,1}$ 和 $p_{t,2}$ 经过光滑后的灰度,从而设置二进制中的相应位的值,即

$$f(W, s_t, \sigma_t) = \begin{cases} 1 & W(p_{t,1}, \sigma_{t,1}) > W(p_{t,2}, \sigma_{t,2}) \\ 0 & \text{其他} \end{cases} \quad (1)$$

式中, $(W(p_{t,1}, \sigma_{t,1}), (W(p_{t,2}, \sigma_{t,2}))$ 是图像块 W 中坐标 $p_{t,1}(p_{t,2})$ 经标准差 $\sigma_{t,1}(\sigma_{t,2})$ 高斯滤波后的值。

最终的二进制串 b_w 定义为

$$b_w = \sum_{1 \leq t \leq n} 2^t f(W, s_t, \sigma_t) \quad (2)$$

这种对像素直接比对的方法易受噪声和局部表现变化的影响,进行预高斯光滑可以在一定程度上使其得到缓解,但是这也意味着在存在特征较多的高频区域内信息的缺失。LATCH二进制描述子引进了三元组像素块的比对方法,从而确定二进制中相应位的取值。考虑 $t = 1, \dots, n$ 个三元组序列, S 转化为 $\hat{S} = \{\hat{s}_t\}_{t=1, \dots, n} = \{[p_{t,\alpha}, p_{t,1}, p_{t,2}]\}_{t=1, \dots, n}$, $p_{t,\alpha}, p_{t,1}$ 以及 $p_{t,2}$ 给出了大小为 $k \times k$ 像

素块中心像素的定位。将 $P_{t,\alpha}$ 定义为主像素块, $P_{t,1}$ 和 $P_{t,2}$ 是其伴随像素块,通过计算之间的F范数来评估主像素块和伴随像素块的相似程度,由式(3)确定二进制串中的对应位的值。将 g 替换式(2)中函数 f 就确定了基于三元组方法的二进制串 b_w 。

$$g(W, \hat{s}_t) = \begin{cases} 1 & \|P_{t,\alpha} - P_{t,1}\|_F^2 > \|P_{t,\alpha} - P_{t,2}\|_F^2 \\ 0 & \text{其他} \end{cases} \quad (3)$$

在运行时间方面,LATCH二进制描述子保持了二进制描述子的优势,比基于直方图描述子快一个数量级;在鲁棒性方面,LATCH二进制描述子在大多数数据集上的效果优于其他二进制描述子,缩小了与基于直方图描述子的差距^[13]。

1.2 建立视觉词袋

在计算机视觉领域,图像通常以特征点及其描述来表达,以特征描述看作单词构建出相应的词袋模型,图像便可以转化为低维稀疏的数字向量,便于大量图像信息的处理,图像间的相似度转化为向量间的相似度,其本质是一个信息压缩的过程。

本文采用LATCH二进制描述子构建视觉词袋,对训练集进行特征提取并将所有特征通过K-means聚类,从而所有特征被分为 K 类,构成层级词袋的第一层,以此类推,采用同样方法对每层聚类空间分为 K 类,构建出其对应的子层级,直至第 L_q 层为止,最终所有的 Q 个叶节点在 L_q 层形成 K 的 L_q 次方类,对每一类用其中的所有特征的平均特征作为代表(称为单词),从而构建了一个如图1所示的具有 Q 个单词和 $\frac{K^{L_q} - 1}{K - 1}$ 个节点的树形视觉词典。树结构中的中间节点供快速查找给定特征对应单词时使用,只需将其与中间节点的聚类中心比较(一共 L_q 次),即可查找到对应单词,保证了查找效率。

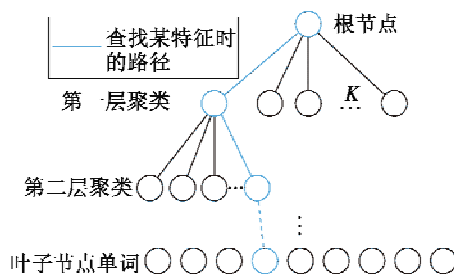


图1 K 叉树视觉字典示意图

Fig. 1 Sketch map of the K -ary tree visual dictionary

根据文献[14]所采用的频率-逆文档频率(TF-IDF)方法为每个叶节点赋予权重。频率-逆文档频率为

$$\begin{cases} \eta^i = tf(i, I_t) \times idf(i) \\ tf(i, I_t) = n_{it} / n_t \\ idf(i) = \ln(N / n_i) \end{cases} \quad (4)$$

式中: n_{i_i} 为单词 w_i 在图像中 I_i 出现的次数; n_{I_i} 为 I_i 中的单词总数; N 为训练数据库的图像总量; n_i 为单词 i 在整个训练图像库中出现的次数。

TF 部分的思想是某单词在一幅图像中的出现次数越高, 其区分度就越高。IDF 思想是某单词在字典中出现的频率越低, 则分类图像时区分度越高。TF-IDF 越大, 表明该单词在某图像中出现的频率高, 并且在其他图像中较少出现, 因此该单词具有很好的区别分类性。

2 关键帧闭环检测与验证

2.1 关键帧的选取

与传统的基于运动变化的关键帧选取方法不同, 本文选取方法采用当前帧与上一关键帧之间的相似度。为提高相似度的可靠性, 融合特征跟踪率与 RGB 直方图匹配评判相似度, 相似度函数定义为

$$\gamma = \alpha_{\text{color}} |h_k - h_c| + \alpha_{\text{keypoint}} \cdot \left[\left(1 - \frac{m_M}{\max(m_k, m_c)} \right) + \frac{1}{\max(m_k, m_c)} \sum_{i=1}^{m_M} \|x_i^k - x_i^c\| \right] \quad (5)$$

式中: α_{color} 为 RGB 直方图匹配的权重, α_{keypoint} 为 LATCH 特征匹配的权重, 且 $\alpha_{\text{color}} + \alpha_{\text{keypoint}} = 1$; m_M 为两幅图像匹配的特征点数目; m_k, m_c 分别为第 k 个关键帧与当前帧的特征点数目; x_i^k 和 x_i^c 为匹配的 LATCH 特征描述向量; h_k, h_c 分别表示两幅图像的 RGB 统计直方图, 采用的特征跟踪方法参见文献[15]。

在式(5)中可根据 α_{color} 和 α_{keypoint} 来调整特征匹配和色彩直方图匹配的权重, 鉴于特征数目较少时随机因素容易对关键帧的选取产生影响, 因此本文根据特征数目适当调整 α_{keypoint} , 即

$$\alpha_{\text{keypoint}} = \frac{1}{2} - \frac{1}{2} \exp\left(-\frac{\min(m_k, m_c)}{m_e}\right) \quad (6)$$

式中, m_e 为期望的特征提取数量。

2.2 关键帧闭环检测

根据式(4)可计算图像 I_i 中特征点对应的单词及其权重, 从而获得该图像的 BoW, 即

$$A = \{(w_1, \eta_1), (w_2, \eta_2), \dots, (w_N, \eta_N)\} \triangleq \mathbf{v}_A \quad (7)$$

通过词袋可用向量 \mathbf{v}_A 描述图像 A , 而向量 \mathbf{v}_A 是稀疏向量, 非零部分表达了图像中含有的单词种类, 且这部分的值为 TF-IDF 的值。对于给定 \mathbf{v}_A 和 \mathbf{v}_B , 采用 1 范数形式来表示向量间的差异(图像间的距离), 即

$$s(\mathbf{v}_A, \mathbf{v}_B) = 1 - \frac{1}{2} \left| \frac{\mathbf{v}_A}{\|\mathbf{v}_A\|} - \frac{\mathbf{v}_B}{\|\mathbf{v}_B\|} \right| \quad (8)$$

从式(8)可以得出: 两图像越相似, $s(\mathbf{v}_A, \mathbf{v}_B)$ 得分越高(距离越小), 而当两图像无差别时, 距离 $s(\mathbf{v}_A, \mathbf{v}_B) = 1$ 。为进一步提高评分的可靠性, 在先验相似度 $s(v_i, v_{i-\Delta t})$ 基础上进行归一化, 即

$$s(v_i, v_{i'})' = s(v_i, v_{i'}) / s(v_i, v_{i-\Delta t}) \quad (9)$$

先验相似度表示某时刻关键帧图像与前一时刻关键帧的相似性, 从而 $s(v_i, v_{i'})'$ 反映了当前帧与之前的关键帧间在先验相似度基础上归一化的相似度。当归一化相似度达到一定阈值条件时, 对应的关键帧构成了闭环候选帧。

基于词袋的闭环检测依赖于外观而未利用任何几何信息, 从而易造成外观相似的图像被检测为闭环, 为了保证闭环检测的可靠性, 需要对闭环进行验证, 以及进行时间和空间一致性的检验后, 闭环候选帧才会被认为是正确的。时间一致性需要设立闭环缓存机制, 在一段时间内持续检测到闭环; 空间一致性是对检测到的闭环两帧进行特征匹配, 估计摄像机帧间运动, 最后将该运动级联检查是否满足闭环要求。

2.3 基于 RANSAC-HORN 运动估计的闭环验证

与此同时也看到, 词袋只考虑了图像中的特征描述, 丢失了特征的几何约束。文献[11]采用随机抽样一致性(RANSAC)方法计算两幅图像之间的基本矩阵, 判断其内点数是否足够, 从而对当前帧和闭环候选帧进行几何一致性校验。由于双目相机不存在尺度问题, 本文采用基于 RANSAC-HORN 运动估计的闭环验证方法。

HORN 方法最少可以使用 3 对匹配的三维点求取图像间的运动信息, 具体实现为:

1) 设当前关键帧的 3 个三维点质心坐标为 P_{oc} , 则将三维点减去质心坐标得到质心坐标系下的新坐标 $P_c \{P_{c1}, P_{c2}, P_{c3}\}$, 同法可得到闭环候选帧 3 个三维点在其质心 P_{ol} 坐标系下的新坐标 $P_l \{P_{l1}, P_{l2}, P_{l3}\}$;

2) 根据点集 P_c 和 P_l 分别确定的平面法向量 $\mathbf{n}_c = P_{c1} \times P_{c2}$ 和 $\mathbf{n}_l = P_{l1} \times P_{l2}$, 以及单位法向量 $\hat{\mathbf{n}}_c$ 和 $\hat{\mathbf{n}}_l$;

3) 令 $\mathbf{n}_a = \hat{\mathbf{n}}_c \times \hat{\mathbf{n}}_l$, 两平面间夹角 φ 满足

$$\begin{cases} \sin \varphi = \|\mathbf{n}_a\| \\ \cos \varphi = \hat{\mathbf{n}}_c \cdot \hat{\mathbf{n}}_l \end{cases} \quad (10)$$

则此夹角对应的四元数向量为

$$\mathbf{q}_a = \left(\cos \frac{\varphi}{2}, \sin \frac{\varphi}{2} \cdot \hat{\mathbf{n}}_a \right); \quad (11)$$

4) 定义变量 $C = \sum_{i=1}^3 (P_{ci} \times P_{li})$ 和 $S = \mathbf{n}_l \sum (P_{ci} \times P_{li})$, 可得平面内两组点集间的旋转角分别为

$$\begin{cases} \sin \theta = S / \sqrt{S^2 + C^2} \\ \cos \varphi = C / \sqrt{S^2 + C^2} \end{cases} \quad (12)$$

对应的四元数为 $\mathbf{q}_l = \left(\cos \frac{\varphi}{2}, \sin \frac{\varphi}{2} \cdot \hat{\mathbf{n}}_l \right)$;

5) 由 \mathbf{q}_a 和 \mathbf{q}_l 得到旋转矩阵 \mathbf{R}_a 和 \mathbf{R}_l , 则总旋转矩阵为 $\mathbf{R}_z = \mathbf{R}_l \times \mathbf{R}_a$, 平移 $T = P_{ol} - \mathbf{R} P_{oc}$ 。

RANSAC 算法可实现对外点的滤除, 对输入数据进

行多次随机采样并确定模型,根据确定的模型对所有数据进行划分。经过多次随机抽样试验后会找到符合误差范围内最多点的集合,利用该内点集合确定最终的模型。考虑 RANSAC 算法的抽样点数尽量小的原则,根据 HORN 方法确定抽样点数的最小集为 3 对三维点。当将 R_z 和 T 确定后,可引导当前关键帧特征点在闭环候选帧中的大致区域,以及闭环候选帧中特征点在当前关键帧中的大致区域,查找更多匹配以弥补在词袋匹配环节的漏匹配,当匹配特征数目达到足够点数时,该闭环候选帧才被接收为成功闭环,从而实现了闭环验证。

3 闭环位姿优化

经过 2.3 节基于 RANSAC-HORN 运动估计的闭环验证后,已知当前关键帧与闭环关键帧之间的旋转矩阵 R_z 和平移 T 以及由此引导查找到满足一定数目的匹配点,则此闭环候选帧被接收为闭环关键帧,实现对闭环验证。为对当前关键帧位姿进一步优化,将闭环关键帧的局部地图点投影到当前帧并最小化重投影误差。

通过最小化世界坐标系下三维路标点 $X^i \in \mathbb{R}^3$ 和当前关键帧的特征点 x^i 之间匹配的重投影误差来实现对相机的旋转 $R \in SO(3)$ 和位置 $l \in \mathbb{R}^3$ 的优化,即

$$\{R, l\} = \underset{R, l}{\operatorname{argmin}} \sum_i \rho(\|x^i - \pi(RX^i + l)\|_{\Sigma}^2) \quad (13)$$

式中: ρ 为具备鲁棒性的 Huber 代价函数; Σ 为与特征点尺度相关的协方差矩阵,目标函数为

$$F(x) = \sum_i \rho(\|x^i - \pi(RX^i + l)\|_{\Sigma}^2) \quad (14)$$

投射函数 π 定义为

$$\pi \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{bmatrix} f_x \cdot x/z + c_x \\ f_y \cdot y/z + c_y \\ f_x \cdot (x - a)/z + c_x \end{bmatrix} \quad (15)$$

式中: (f_x, f_y) 为摄像机焦距; (c_x, c_y) 为投影中心; a 为基线长度。以上参数均能通过相机的标定来获取。

在相机的初始位姿 \tilde{x} 已知情况下,利用一阶泰勒展开线性化目标函数 $F(x)$,最终建立线性化系统,需求解的方程为

$$H\Delta x = -b \quad (16)$$

式中: b 为系数向量; H 为系数矩阵(信息矩阵)。通过求解式(16)线性方程组可得最小值 Δx^* ,新的状态向量估计值可依据

$$x^* = \tilde{x} + \Delta x^* \quad (17)$$

进行更新,进而迭代直到满足预定要求。

常用的非线性最小平方算法有 Gauss-Newton 算法(GN),列文伯格-马夸尔特算法(LM)等,均可以直接用来求最后的最优解。由于系数向量和系数矩阵在形式上都是稀疏的,一些开源的稀疏矩阵运算算法结合非

线性最小二乘算法可提高计算效率。本文利用在图优化通用框架(g2o)^[16]中的 LM 算法来实现非线性优化。

4 实验与分析

本文实验数据采用 KITTI 数据集^[17-18],该数据集是具备组合测量系统的车辆进行跑车实验而采集得到的,包括搭载有 Flea2 Firewire 摄像机采集的高质量双目图像序列(已预先矫正)和 OXTS RT 3000 IMU/GPS 测量系统采集的地理数据(作为真值)。运行计算机为 DELL OPTIPLEX 7010 台式电脑(Intel Core i5-3470 CPU,主频 3.20 GHz,4.00 GB 内存,64 位的 Windows7 操作系统)。

在 KITTI 数据集中,00-10 号图像序列是测试序列,包括了城市道路、小镇、乡村、高速多种环境,00,02,05,06,07,09 号图像序列含有闭环,包含较多的相似场景。

图 2 所示为本文算法在 KITTI 数据集上的重建轨迹,图中,红色线段为闭环检测到的路段,可以看出,在含有闭环的 00,02,05,06,07,09 号图像序列中全部正确识别闭环。

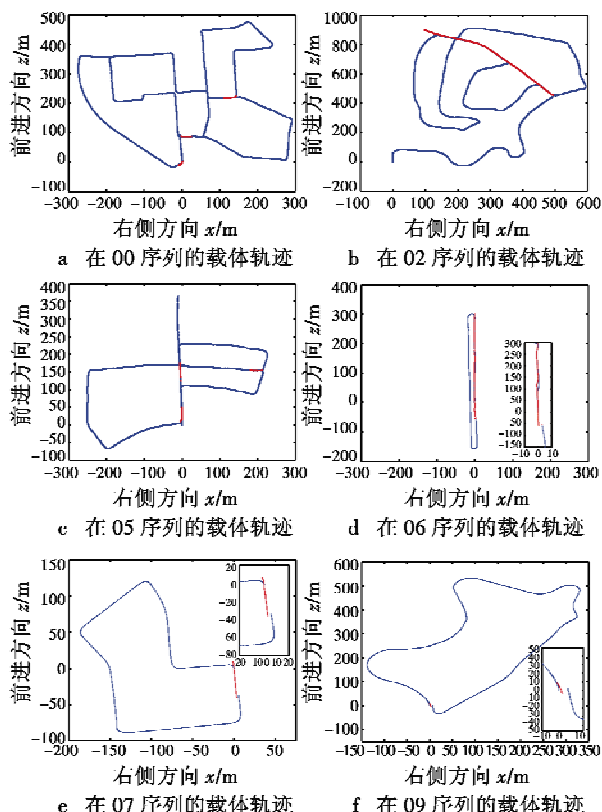


图 2 本文算法在 KITTI 数据集的重建轨迹

Fig. 2 Reconstructed trajectory of the proposed algorithm on KITTI data set

经本文算法闭环优化后的重建轨迹,平移和旋转误差均显著减小。在含闭环的 00,02,05,06,07,09 序列下,双目 VO 平均平移误差 1.0714%,平均旋转误差为

0.003 050 ($^{\circ}$)/m;而本文提出基于闭环检测的VO平均平移误差0.835 8%,平均旋转误差为0.002 712 ($^{\circ}$)/m。

在实时性方面,由于闭环检测采用关键帧和词袋技术,同时优化只是当前帧位姿优化而非全局优化,因此在实时性上略有下降,但并不影响VO的实时导航。

综上实验结果对比分析,本文提出的闭环检测和位姿优化算法能够有效地对闭环进行准确检测和验证,并对VO导航过程的误差积累进行闭环位姿优化,提高了视觉导航精度。

5 结束语

本文提出了一种基于几何约束的视觉闭环检测和位姿优化方法,旨在提高视觉导航精度。首先建立基于LATCH二进制描述子的视觉词袋,并采用视觉单词向量描述图像进而归一化相似度的闭环检测方法,然后对闭环候选关键帧和当前关键帧进行基于RANSAC-HORN运动估计的闭环验证,最后将闭环关键帧的局部地图点投影到当前帧并最小化重投影误差以使得位姿得到优化。数据集实验表明,本文提出的闭环检测和位姿优化算法能够有效、准确地对闭环检测和验证,并对VO导航过程的误差积累进行闭环位姿优化,提高了视觉导航精度。下一步将在本文基础上探索闭环验证后位姿的全局优化,以实现闭环前路径的优化。

参考文献

- [1] YOUSIF K, BAB-HADIASHAR A, HOSEINNEZHAD R. An overview to visual odometry and visual SLAM: applications to mobile robotics[J]. *Intelligent Industrial Systems*, 2015(1):289-311.
- [2] KONOLIGE K, AGRAWAL M, SOLÀ J. Large-scale visual odometry for rough terrain[J]. *International Journal of Robotics Research*, 2011(66):201-212.
- [3] ANNAIYAN A, YADAV M, OLIVARES-MENDEZ M A, et al. Visual odometry based absolute target geo-location from micro aerial vehicle[C]//*IEEE International Conference on Robotics, Automation, Control and Embedded Systems*, 2015:1-7.
- [4] WILLIAMS B, CUMMINS M, NEIRA J, et al. A comparison of loop closing techniques in monocular SLAM[J]. *Robotics and Autonomous Systems*, 2009, 57(12):1188-1197.
- [5] ANGELI A, FILLIAT D, DONCIEUX S, et al. A fast and incremental method for loop-closure detection using bags of visual words[J]. *IEEE Transactions on Robotics*, 2008, 24(5):1027-1037.
- [6] CUMMINS M, NEWMAN P. Probabilistic appearance based navigation and loop closing[C]//*IEEE International Conference on Robotics and Automation*, 2007:2042-2048.
- [7] CUMMINS M, NEWMAN P. Highly scalable appearance-only SLAM-FAB-MAP 2.0[C]//*Robotics: Science and Systems*, 2009. doi:10.15607/RSS.2009.V.039.
- [8] LABBE M, MICHAUD F. Memory management for real-time appearance-based loop closure detection[C]//*IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011:1271-1276.
- [9] LABBE M, MICHAUD F. Appearance-based loop closure detection for online large-scale and long-term operation[J]. *IEEE Transactions on Robotics*, 2013, 29(3):734-745.
- [10] CUMMINS M, NEWMAN P. Appearance-only SLAM at large scale with FAB-MAP 2.0[J]. *International Journal of Robotics Research*, 2011, 30(9):1100-1123.
- [11] GALVEZ-LOPEZ D, TRADOS J D. Bags of binary words for fast place recognition in image sequences[J]. *IEEE Transactions on Robotics*, 2012, 28(5):1188-1197.
- [12] MUR-ARTAL R, TARDOS J D. Fast relocalisation and loop closing in keyframe-based SLAM[C]//*IEEE International Conference on Robotics and Automation(ICRA)*, 2014:846-853.
- [13] LEVI G, HASSNER T. LATCH: learned arrangements of three patch codes[C]//*IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2016:202-210.
- [14] CUMMINS M, NEWMAN P. Accelerating FAB-MAP with concentration inequalities[J]. *IEEE Transactions on Robotics*, 2010, 26(6):1042-1050.
- [15] MUR-ARTAL R, MONTIEL J M M, TARDOS J D. ORB-SLAM: a versatile and accurate monocular slam system[J]. *IEEE Transactions on Robotics*, 2015, 31(5):1147-1163.
- [16] KUMMERLE R, GRISETTI G, STRASDAT H, et al. G2o: a general framework for graph optimization[C]//*IEEE International Conference on Robotics and Automation (ICRA)*, 2011:3607-3613.
- [17] PAZ L M, PINIES P, TARDOS J D, et al. Large-scale 6-DOF SLAM with stereo-in-hand[J]. *IEEE Transactions on Robotics*, 2008, 24(5):946-957.
- [18] SCARAMUZZA D, FRAUNDORFER F. Visual odometry part I: the first 30 years and fundamentals[J]. *IEEE Robotics & Automation Magazine*, 2011, 18(4):80-92.