

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/343280756>

# 基于SLAM的视觉定位

**Article** in *Beijing Hangkong Hangtian Daxue Xuebao/Journal of Beijing University of Aeronautics and Astronautics* · January 2021

DOI: 10.13700/j.bh.1001-5965.2019.0642

CITATIONS

0

READS

521

5 authors, including:



Ping An

Shanghai University

288 PUBLICATIONS 2,757 CITATIONS

SEE PROFILE



Yilei Chen

Shanghai University

21 PUBLICATIONS 126 CITATIONS

SEE PROFILE

http://bhxb.buaa.edu.cn      jbuaa@buaa.edu.cn  
DOI: 10.13700/j.bh.1001-5965.2019.0642

# 一种高效准确的视觉 SLAM 闭环检测算法

安平\*, 王国平, 余佳东, 陈亦雷, 尤志翔

(上海大学 通信与信息工程学院, 上海 200444)

**摘 要:** 同步定位与地图构建(SLAM)是视觉导航领域的关键技术之一,闭环检测是 SLAM 的基础问题。针对视觉 SLAM 闭环检测准确率不高的问题,提出一种高效准确的闭环检测算法。该算法由词袋模型、图像结构校验、跟踪预测模型 3 个模块构成。首先,将局部特征与全局特征相结合,设计了词袋模型与图像结构校验模块。词袋模型通过视觉单词比较图像之间的相似性,得到闭环候选帧。然后,图像结构校验模块灰度化、归一化当前图像与闭环候选图像。归一化之后的图像被直接作为局部特征的邻域,计算得到全局描述符,通过全局描述符判断闭环候选帧是否为有效的闭环。最后,针对传统闭环检测算法耗时随图像数量增加而显著增加的问题,设计了跟踪预测模块,以提高计算效率。实验中,与主流的 DBoW 算法相比,提出的闭环检测算法的准确率提升了 20% 以上,实时性也有更好的表现。

**关 键 词:** 同步定位与地图构建(SLAM); 闭环检测; 局部特征; 全局特征; 跟踪预测  
**中图分类号:** TP242

**文献标志码:** A      **文章编号:** 1001-5965(2021)01-0024-07

视觉同步定位与地图构建(Simultaneous Localization and Mapping, SLAM)技术是视觉导航的关键技术之一<sup>[1]</sup>,目前主要应用于机器人领域。主流的视觉 SLAM 技术主要通过相机获得连续图像,利用几何方法估计相邻图像的运动并在后端对其进行优化<sup>[2]</sup>。闭环检测<sup>[3]</sup>作为后端优化的关键技术,通过检测出相机是否经过同一位置区域,以此作为后端优化的重要依据。闭环检测算法的准确性主要包括消除误检和漏检 2 部分,正确高效的检测判定是闭环检测算法的关键所在。

基于不同的实现方法,目前主流的闭环检测算法大致可以分为 3 类:①地图到地图(Map-to-Map)类方法。Clemente 和 Davison<sup>[4]</sup>于 2007 年提出一种“分段式”地图概念,极大地提高了其在动态和复杂环境中的稳健性,但是因为空间特征需要构建稀疏地图,最终的闭环检测结果往往因为信息量不够而欠准确。②图像到地图(Image-to-

Map)类方法。Williams 等<sup>[5]</sup>在设计重定位模块时提出了一种通过查找当前图像帧与地图特征对应关系来确定寻找两者的匹配关系,但是这种方法需要把每一次的闭环信息都用来训练分类器,浪费大量的计算资源。③目前主流的图像到图像(Image-to-Image)类方法。这种匹配算法主要通过特征算子评估图像间的相似性确定闭环。其中,特征算子大致可以分为局部算子和全局算子 2 类<sup>[6]</sup>。全局算子主要通过单个全局描述子去描述整张图像<sup>[7]</sup>。早期的研究成果主要是全局描述子的设计,Krose 等<sup>[8]</sup>通过主成分分析(Principal Component Analysis, PCA)降维生成图像的全局描述子。后来研究人员开始考虑图像本身的性质。Lowry 等<sup>[9]</sup>通过在线学习改进了 PCA 描述子,指出 PCA 的前半部分的维度代表了图像序列相似的信息,对环境变换比较敏感,而后半部分维度的 PCA 特征则相反,更能应对环境的变换。相

收稿日期: 2019-12-23; 录用日期: 2020-01-17; 网络出版时间: 2020-03-13 10:37  
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20200313.0820.001.html  
基金项目: 国家自然科学基金(61828105); 上海市科委项目(17DZ2292400,18XD1423900)  
\* 通信作者。E-mail: anping@shu.edu.cn

**引用格式:** 安平, 王国平, 余佳东, 等. 一种高效准确的视觉 SLAM 闭环检测算法[J]. 北京航空航天大学学报, 2021, 47(1): 24-30. AN P, WANG G P, YU J D, et al. An efficient and accurate visual SLAM loop closure detection algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(1): 24-30 (in Chinese).

比全局描述子,局部描述子的方法目前更为通用。局部描述子方法的流行主要因为视觉词袋(Bag-of-Words,BoW)技术。Cummins 和 Newman<sup>[10]</sup>较早地把词袋技术引入到闭环检测中,通过 Chow-Liu 树计算视觉单词的联合概率分布,从而获得闭环的后验概率,以此判断闭环是否可能发生。Galvez-Lopez 和 Tardos<sup>[11]</sup>创造性地提出 DBoW 算法。DBoW 利用词汇树向量化图像,并通过正向(Direct Index)和反向(Inverse Index)2 种索引结构提高计算效率。ORB-SLAM2 系统<sup>[12]</sup>作为近几年视觉 SLAM 系统的代表性工作,通过 DBoW 算法获得了较强的闭环检测和重定位能力。

综上所述,闭环检测已经取得了一些进展。但即使是应用程度最广的闭环检测 DBoW 算法,仍然存在诸多问题:①词袋模型通过聚类生成视觉单词,通过向量比较图像序列之间的相似性,随着场景范围的扩大,尤其是大范围的场景重建,计算量也会随之增加,实时性将会面对挑战<sup>[13]</sup>。②在闭环检测中,一个真阳性闭环(预测为闭环的真闭环)能够显著优化视觉里程计的误差,但一个假阳性闭环(预测为闭环的假闭环)可能导致整个后端优化模块收敛到错误的方向。正确地检测出闭环对整个 SLAM 系统至关重要。目前,一些闭环检测算法的准确率在复杂场景下还远远不够,尽管已经有一些研究希望能够降低闭环检测的错误率<sup>[14-15]</sup>,但往往会加重整个 SLAM 后端的负担,并且提高也不明显。

针对上述问题,本文提出了一种高效准确的闭环检测算法 TPSV-DBoW (Tracking Prediction and Structural Verification-Distributed Bag of Words)。

为了提高闭环检测算法的准确率与鲁棒性,将局部算子与全局描述符结合。利用局部算子(视觉单词)比较图像的相似性,得到闭环候选帧;通过全局描述符进行结构校验,验证闭环候选帧是否是真闭环,以此设计了词袋模型和图像结构校验模块。针对传统闭环检测算法耗时随图像数量增加而显著增加的问题,进而设计了跟踪预测模型,基于相邻前几帧图像的闭环检测结果预测当前帧的状态。

## 1 闭环检测算法

如图 1 所示,TPSV-DBoW 算法由词袋模型、图像结构校验、跟踪预测模型 3 个模块组成。算法通过词袋模型模块和图像结构校验模块将局部特征与全局特征相结合,以提高闭环检测算法在复杂场景下的准确率与鲁棒性。词袋模型通过二元鲁棒独立基本特征(Binary robust independent elementary feature,Brief)离线词袋树得到图像的视觉单词向量,实现图像的数字化。之所以选用 Brief 而不是其他更加鲁棒的特征算子是为了提高算法运行速度,且本模块的设计是提供闭环候选帧而不是直接得到真闭环,Brief 算子的性能能够满足本模块的需要。图像结构校验模块通过全局算子进行结构校验。出于节省计算资源的考虑,没有设计专门用于校验的全局算子,而是通过对图像灰度化、归一化处理,将归一化之后的图像平均分成 4 份,每份图像的中心像素点直接作为定向快速旋转 Brief ORB 的特征点,避免了提取特征点的复杂操作。之后,计算特征点的局部描

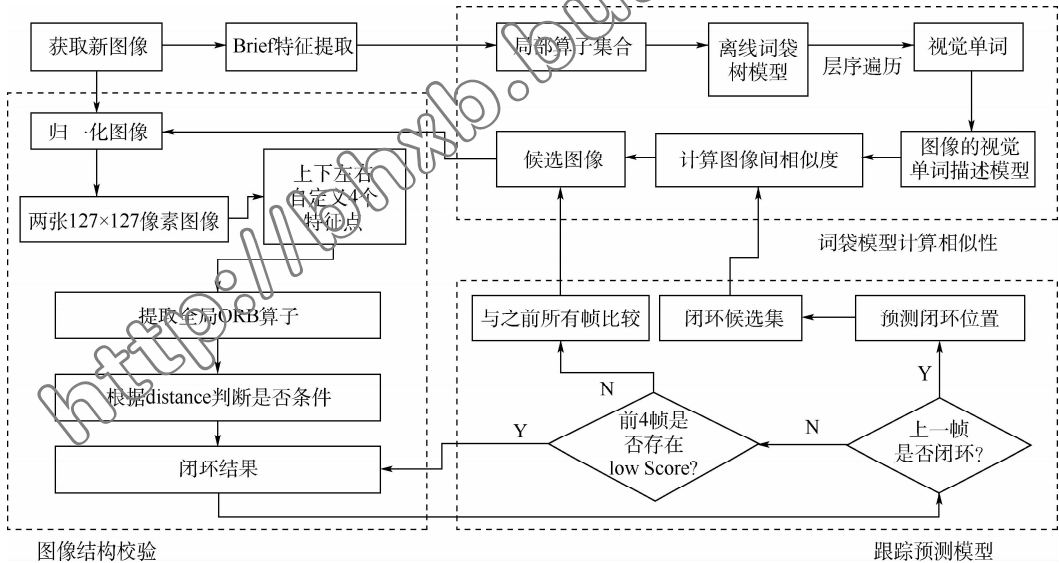


图 1 闭环检测流程

Fig. 1 Flowchart of loop closure detection

述符。每个 ORB 局部描述符恰好覆盖每份图像 (即归一化图像的四分之一), 可作为归一化图像的全局描述符。词袋模型提供候选闭环, 图像结构校验模块验证是否为真闭环。这种局部描述符与全局描述符相结合的方法可以进一步提高闭环检测的准确率。在跟踪预测模块中, 前几帧的闭环检测结果可以作为后一帧闭环判断的先验条件。在传统的闭环检测算法中, 当前图像都需要与之前所有已经出现过的图像比较相似性得分才能判断闭环是否存在, 这会导致计算时间随着图像数量的增加而急剧增加。本模块的设计主要是根据前几帧的闭环状态判断当前图像的闭环状态。主要有 2 点, 其一先判断当前图像是否存在闭环, 其二如果存在则预测出闭环候选集, 借此提高闭环检测的实时性。

2 核心算法模块

2.1 词袋模型计算相似性

对每一帧图像计算 Brief 算子。Brief 是典型的二进制描述符, 可降低计算复杂度, 提高计算速度。Brief 只是计算描述符的方法, 不包含提取特征点的过程, 所以需要先利用 FAST (Features from Accelerated Segment Test) 算法检测特征点, 再计算 Brief 描述符。之后通过 Brief 词袋树完成图像的向量化。视觉词袋有多种组织方式, 对应于不同的搜索复杂度。词袋模型采用的结构如图 2 所示, 其是一个度为  $K=10$ 、层数  $L=6$  的搜索树, 时间复杂度为  $\lg 6$ 。

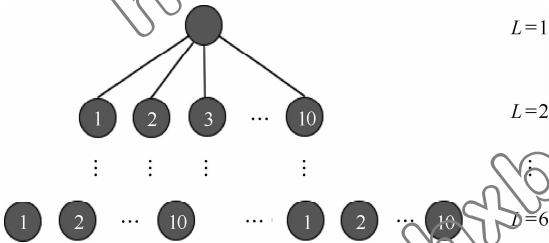


图 2 Brief 词袋树  
Fig. 2 Vocabulary tree of Brief

词袋树每个叶节点的单词都被赋予一个 TF-IDF 权重, 如下:

$$w_i = \frac{n_{id}}{n_d} \lg \frac{N}{N_i} = TF \cdot IDF \tag{1}$$

式中:  $n_{id}$  为图像  $d$  中视觉单词  $i$  出现的频数;  $n_d$  为图像  $d$  中视觉单词的总数;  $N$  为图像序列总数;  $N_i$  为所有图像中视觉单词  $i$  出现的总数。TF 越高, 说明单词在这幅图像中出现的越多, 单词的 IDF 越高, 说明其他图像中很少出现, 二者结合起来, 则认为此单词具有很好的区分能力, 权重也就越高。

对图像中的每个特征算子通过层序遍历池袋树, 即从根节点开始搜索, 首先确定根节点的 10 个子节点中与其最匹配的节点, 再从该子节点开始搜索。这样一层层往下遍历, 从而得到其对应的视觉单词。最终得到整张图的视觉单词描述模型:

$$I_u \rightarrow \{d_1, d_2, \dots, d_n\} \tag{2}$$

式中:  $I_u$  为图像;  $d_u$  为视觉单词。这些值构成图像的描述向量  $\mathbf{v}$ 。对 2 幅图像比较计算两者的相似度, 计算公式如下:

$$s(\mathbf{v}_1, \mathbf{v}_2) = 1 - \frac{1}{2} \left| \frac{\mathbf{v}_1}{|\mathbf{v}_1|} - \frac{\mathbf{v}_2}{|\mathbf{v}_2|} \right| \tag{3}$$

2.2 图像结构校验模块

本模块是基于全局描述符的设计, 通过全局描述符校验闭环候选帧。词袋模型通过局部描述符寻找候选闭环, 在此基础上, 通过本模块进行图像结构校验。这种设计可大大提高闭环检测的鲁棒性与准确率。

将候选闭环图像与当前图像灰度化、归一化, 如图 3 所示, 左边为未处理之前的  $1241 \times 376$  的彩色图像, 右边为归一化之后的  $127 \times 127$  的灰度图像。与仅用中心特征作为全局特征<sup>[7]</sup>的方法不同, 将归一化之后的图像平均分成 4 块 (见图 4), 每块为  $64 \times 64$  像素。之后把每份图像的中心像素点直接作为 ORB 特征点, 这样可以避免特征点定位的复杂操作, 从而提高计算效率。在每个特征点的邻域直接计算 ORB 局部描述符, 图 4 左边的数字即为第一份图像的 ORB 描述符。每个局部描述符恰好覆盖每份图像, 即归一化图像的四分之一。4 个局部描述符恰好覆盖整张图像, 可作为图像的全局特征。这种方式不需要设计特别复杂的全局描述符即可完成结构校验。

ORB 拥有旋转不变性, 对噪声不敏感, 相对鲁



图 3 图像归一化  
Fig. 3 Image normalization

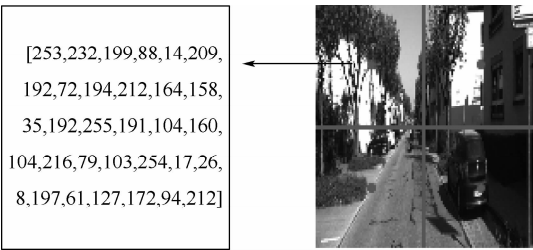


图 4 ORB 全局特征  
Fig. 4 ORB holistic features

棒,并且计算速度也较快,完全可以满足结构校验的需要。ORB 使用矩来计算特征点以  $r$  为半径范围内的质心。矩的定义如下:

$$m_{pq} = \sum_{x,y} x^p y^q I(x,y) \quad (4)$$

式中:  $I(x,y)$  为  $(x,y)$  处的灰度值。通过该矩计算邻域质心为

$$C = \left( \frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (5)$$

特征点坐标到质心的向量作为该特征点的方向。假设角点坐标为  $O$ ,此向量的角度即为特征点方向<sup>[7]</sup>。

$$\theta = \arctan\left(\frac{m_{10}/m_{00}}{m_{01}/m_{00}}\right) = \arctan\left(\frac{m_{10}}{m_{01}}\right) \quad (6)$$

在进行图像结构校验时,将归一化图像各区域的全局算子两两比较,记录汉明距离大于 100 的数目,如下<sup>[7]</sup>:

$$\sum_{x \in \{1,2,3,4\}} \text{distance}(N_x, I_x) \quad (7)$$

式中:  $N_x$  为候选闭环的第  $x$  份区域;  $I_x$  为当前图像的第  $x$  份区域,  $x \in \{1,2,3,4\}$ 。

闭环候选图像与当前图像归一化之后,每张图像各划分为 4 份区域,匹配之后,若其中有 2 份及以上区域间的汉明距离大于 100,则认为当前的候选闭环不是正确闭环,应当舍弃。

### 2.3 跟踪预测模型

随着场景的扩大,尤其是大范围场景的闭环检测,当前图像需要与之前所有已经出现过的图像逐帧比较,以此寻找闭环。因此,计算量会随着图像序列的增加而增加。由于词袋模型和图像结构校验模块将局部特征与全局特征相结合,在一定程度上也提高了计算的时间复杂度,因而设计了跟踪预测模型来处理以上问题。

跟踪预测模型能够有效地节约整个算法的耗时。跟踪预测模型的设计原则为:根据前几帧的检测结果判断当前图像是否可能有闭环。如果预测不存在闭环,则可以跳过与已经出现过的图像逐帧比较的阶段<sup>[7]</sup>。此外,认为如果预测当前图像存在闭环并且连续前几帧图像有闭环,则可以预测当前图像的闭环候选集,即把闭环图像限定在少量连续的图像集合内,减少与当前图像逐帧比较的数量,提高闭环检测的速度。

本模块的核心思想是:依据前几帧图像的闭环检测结果预测当前图像的闭环检测结果。因为相邻图像具有相似性,即使相机以一个相对较快的速度在移动,由于曝光速度的原因,相邻图像描述的场景仍然没有太大差别。所以,如果某张图

像不存在闭环且其与之前所有图像的相似度得分都很低,那么其下一帧图像也应该不存在闭环。如图 5 所示,如果当前图像与之前所有的图像相似度得分都小于阈值  $S_{\text{low}}$ , 将其定义为 low Score 的图像。low Score 图像的相邻图像得分应该相对较低,而存在闭环的图像的相邻图像的相似度得分应该较高。

如图 6(a)所示,若  $i+n-1$  为当前图像,  $i+n$ 、 $i+n+1$ 、 $i+n+2$ 、 $i+n+3$  为其后 4 帧图像。如果当前图像为 low Score 的图像,认为其后 4 帧图像之内不存在 high Score 的图像。这 4 帧图像都可跳过比较相似性的阶段。4 帧之后则需要重新通过词袋模型与之前所有的图像逐帧比较相似性。另一种可预测的情况如图 6(b)所示,  $i$  为当前图像的序号,  $i-3$ 、 $i-2$ 、 $i-1$  为当前图像之前的图像序号,其对应的闭环图像为  $x_{i-3}$ 、 $x_{i-2}$ 、 $x_{i-1}$ 。因为相机的速度在很短的时间内近似于匀速,在考虑干扰的情况下,闭环图像的差值近似于高斯分布,如下<sup>[7]</sup>:

$$D_i = x_{i-n} - x_{i-n-1} \sim N(\mu, \sigma^2) \quad (8)$$

式中:  $\mu$  为高斯均值;  $\sigma^2$  为高斯方差。

$$\sigma^2 = \frac{\sum_{i=1}^{n-1} D_i^2}{n-1} \quad (9)$$

$$\mu = \frac{\sum_{i=1}^{n-1} (D_i - \mu)^2}{n-1} \quad (10)$$

计算可得闭环候选集的范围为  $[x_{i-1} + \mu - 10\sigma, x_{i-1} + \mu + 10\sigma]$ 。

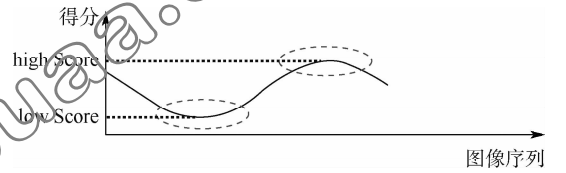


图 5 相似得分的连续性

Fig. 5 Continuity of similarity score

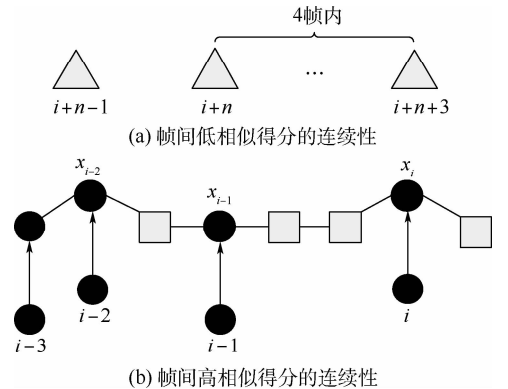


图 6 跟踪预测模型

Fig. 6 Tracking prediction model

3 实验结果分析

实验所用主机的配置为:CPU 为 I7-6700,主频 3.4 GHz,内存 32 GB。对比算法为 DBoW。实验所用数据集为 KITTI Visual Odometry 数据集。

KITTI Visual Odometry 数据集部分图像如图 7 所示,其属于城市公路数据集,包含较多相似场景,并且其中含有不少假闭环,是一个比较具有挑战性的实验数据集。



图 7 实验场景  
Fig. 7 Experimental scene

3.1 准确率数据集实验

KITTI Visual Odometry 数据集中含有 ground truth 真实地图标号为 00~10 的数据集,图 7 中的实验图像便是标号为 00 数据集的场景。标号 00~10 数据集含有连续闭环的为 00、02、05、06 数据集,可以作为用来检测闭环检测算法性能的数据集。本文在这几组数据集上与 DBoW 算法比较,验证了提出的闭环检测算法的准确率。

图 8(a)、(c)、(e)、(g) 为 DBoW 算法在 00、02、05、06 数据集上的实验结果,图 8(b)、(d)、(f)、(h) 为 TPSV-DBoW 算法在 00、02、05、06 数据集上的实验结果。图 8 中连续的线条为相机的运动轨迹。正确的闭环(即相机确实曾经到过此处)旁边有黑色实线标记,没有黑色实线标记的路径表示相机轨迹未发生闭环。如果算法在没有黑色实线标记的路径中检测出闭环,则认为出现误检情况。误检越多,闭环检测算法的准确率越低。图 8(b)、(d)、(f)、(h) 的错误闭环分别少于图 8(a)、(c)、(e)、(g),TPSV-DBoW 的准确率表现更优。

具体闭环的准确率如表 1 所示,可以看出 TPSV-DBoW 算法的优越性。在表 1 中,除 02 数据集外,闭环检测算法的准确率都达到了 80% 以上。由于 02 数据集相似场景很多(道路两旁几乎都是相同的树木),本文所提算法和 DBoW 算法的准确率都相对较低,但相比 DBoW 依旧有 25.35% 的提升。总体上,除了 06 数据集,TPSV-DBoW 算法基本都有 25% 左右的显著提升。由于 06 数据集本身比较简单,故性能普遍较好。DBoW

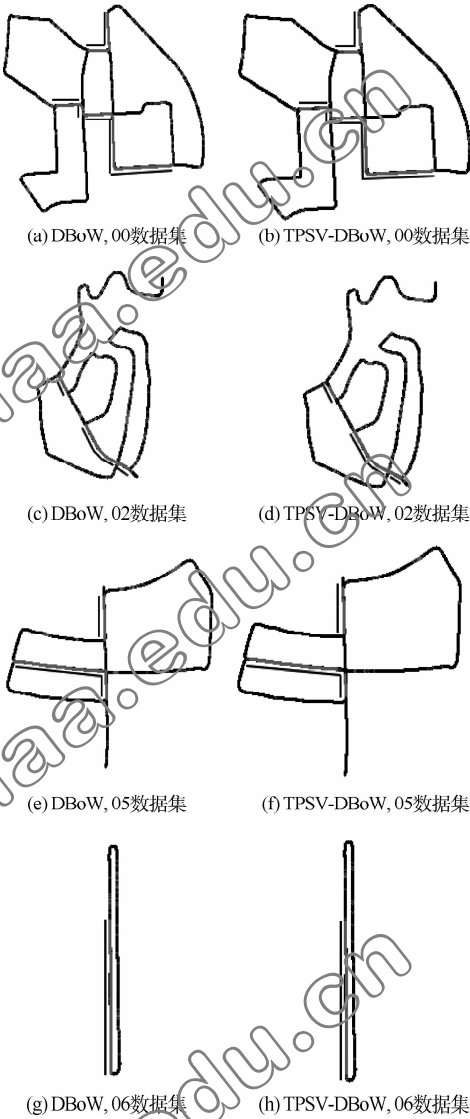


图 8 实验结果  
Fig. 8 Experimental results

表 1 闭环检测准确率  
Table 1 Loop closure detection accuracy

| 数据集标号 | 准确率/%   |              | 准确率<br>提升/% |
|-------|---------|--------------|-------------|
|       | DBoW 算法 | TPSV-DBoW 算法 |             |
| 00    | 67.00   | 83.07        | 23.99       |
| 02    | 26.67   | 33.43        | 25.35       |
| 05    | 62.78   | 84.89        | 35.22       |
| 06    | 78.86   | 89.93        | 14.04       |

算法已经达到了 78.86% 的准确率,所提出的算法在此基础上仅提升了 14.04%。

3.2 实时性检测数据集实验

为了验证 TPSV-DBoW 算法的实时性,实时性检测实验使用 01~10 号 KITTI Visual Odometry 数据集(其中 01、06、07 数据集都为 1 100 张图像)。图 9 展示了闭环检测耗时与图像数量的关系,表 2 记录了闭环检测耗时的具体数据。可以看到,随着图像数量的增加,闭环检测算法耗时急

剧增加,因为要检测出当前图像是否存在闭环,需要与之前所有的图像序列比较。由图 9 可见,随着图像数目的增加,算法在实时性上具有更好的表现,实时性的提升也更加显著。主要是因为跟踪预测模型能够预测当前图像是否可能存在闭环,以及存在闭环的情况下,预测当前的闭环候选集来减少与当前图像比较相似性的图像数量,从而提高算法的计算效率。

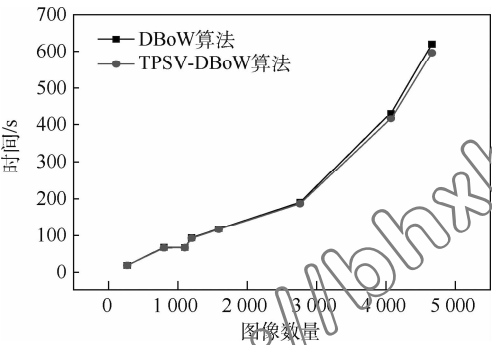


图 9 实时性实验

Fig. 9 Real-time performance experiment

表 2 实时性对比

Table 2 Real-time performance comparison

| 数据集<br>标号 | 图像<br>数量 | 耗时/s    |              | 实时性<br>提升/% |
|-----------|----------|---------|--------------|-------------|
|           |          | DboW 算法 | TPSV-DboW 算法 |             |
| 04        | 271      | 18.55   | 18.53        | 0.10        |
| 03        | 801      | 67.44   | 66.41        | 1.96        |
| 01,06,07  | 1100     | 67.30   | 66.70        | 0.89        |
| 10        | 1200     | 93.56   | 92.17        | 1.48        |
| 09        | 1590     | 118.76  | 116.99       | 1.48        |
| 05        | 2761     | 190.71  | 186.87       | 2.01        |
| 08        | 4070     | 432.77  | 416.78       | 3.56        |
| 02        | 4661     | 649.58  | 595.71       | 3.85        |

4 结 束 语

SLAM 是移动机器人视觉导航领域的关键技术,闭环检测在其中至关重要。为了提高闭环检测的准确率和实时性,提出了一种高效准确的闭环检测算法。①词袋模型与图像结构校验模块通过将全部算子与局部算子结合的方法大大提高了闭环检测的准确率。②考虑到算法复杂度及大场景下闭环检测等因素,跟踪预测模型用于提高计算效率。③相比于 DBoW 算法,TPSV-DBoW 算法的准确率基本提升了 20% 以上,实时性也更加优越。

虽然 TPSV-DBoW 算法的闭环检测准确率有着较大的提升,但精度还存在提升空间,对于复杂场景准确率也有所降低。算法检测时间较 DBoW 算法有所减少,但当图像序列数量增加时,耗时也呈指数级增长。除此以外,动态场景等复杂环境

下算法的鲁棒性也遭受挑战。下一步工作将着重研究提高复杂场景下闭环检测的性能及闭环检测算法对各种场景的适应能力。

参考文献 (References)

[ 1 ] GUI J J, GU D B, WANG S, et al. A review of visual inertial odometry from filtering and optimisation perspectives[ J ]. Advanced Robotics, 2015, 29 ( 20 ) : 1289-1301.

[ 2 ] 何俊华,李战明. 基于视觉的同时定位与地图构建方法综述[ J ]. 计算机应用研究, 2010, 27 ( 8 ) : 2839-2844.

HE J X, LI Z M. Survey of vision-based approach to simultaneous localization and mapping[ J ]. Application Research of Computers, 2010, 27 ( 8 ) : 2839-2844 ( in Chinese ).

[ 3 ] HESS W, KOHLER D, RAPP H H. Systems and methods of detecting loop closure in simultaneous localization and mapping (SLAM) applications. U. S. 14/972,938[ P ]. 2019-06-11.

[ 4 ] CLEMENTE L, DAVIDSON A. Mapping large loops with a single hand-held camera[ C ] // Robotics Science and Systems, 2007, 2 ( 2 ) : 297-304.

[ 5 ] WILLIAMS B, CUMMINS M, NEIRA J, et al. An image-to-map loop-closing method for monocular SLAM[ C ] // IEEE International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2008: 2053-2059.

[ 6 ] 刘强,段富海. 复杂环境下视觉 SLAM 闭环检测方法综述[ J ]. 机器人, 2018, 40 ( 6 ) : 123-136.

LIU Q, DUAN F H. A survey of loop-closure detection method of visual SLAM in complex environments[ J ]. Robot, 2018, 40 ( 6 ) : 123-136 ( in Chinese ).

[ 7 ] 刘国忠,胡钊政. 基于 SURF 和 ORB 全局特征的快速闭环检测[ J ]. 机器人, 2017, 39 ( 1 ) : 36-45.

LIU G Z, HU Z Z. Fast loop closure detection based on holistic features from SURF and ORB[ J ]. Robot, 2017, 39 ( 1 ) : 36-45 ( in Chinese ).

[ 8 ] KROSE E J A, VLASSIS N, BUNSCHOTEN R, et al. A probabilistic model for appearance-based robot localization[ J ]. Image & Vision Computing, 2001, 19 ( 6 ) : 381-391.

[ 9 ] LOWRY S M, WYETH G F, MILFORD M J. Unsupervised on-line learning of condition-invariant images for place recognition[ J ]. Procedia-Social and Behavioral Sciences, 2014, 106: 1418-1427.

[ 10 ] CUMMINS M, NEWMAN P. Probabilistic appearance based navigation and loop closing[ C ] // Proceedings 2007 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2007: 2042-2048.

[ 11 ] GALVEZ-LOPEZ D, TARDOS 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. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras[ J ]. IEEE Transactions on Robotics, 2017, 33 ( 5 ) : 1255-1262.

[ 13 ] 梁志伟,陈燕燕,朱松豪,等. 基于视觉词典的单目视觉闭环检测算法[ J ]. 模式识别与人工智能, 2013, 26 ( 6 ) : 561-570.

LIANG Z W, CHEN Y Y, ZHU S H, et al. Loop closure detection algorithm based on monocular vision using visual dictionary

[J]. Pattern Recognition and Artificial Intelligence, 2013, 26 (6):561-570 (in Chinese).

[14] GUCLU O, CAN A B. Fast and effective loop closure detection to improve SLAM performance [J]. Journal of Intelligent and Robotic Systems, 2019, 93 (3-4):495-517.

[15] ZHANG G, YAN X, YE Y. Loop closure detection via maximization of mutual information [J]. IEEE Access, 2019, 7: 124217-124232.

**作者简介:**  
安平 女, 博士, 教授, 博士生导师。主要研究方向: 图像与视频处理、计算机视觉。  
王国平 男, 硕士研究生。主要研究方向: 视觉 SLAM。  
余佳东 男, 硕士研究生。主要研究方向: 视觉 SLAM。

An efficient and accurate visual SLAM loop closure detection algorithm

AN Ping\*, WANG Guoping, YU Jiadong, CHEN Yilei, YOU Zhixiang

(School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China)

**Abstract:** Simultaneous Localization and Mapping (SLAM) is one of the key technologies in visual navigation, and loop closure detection is a basis of SLAM. An efficient and accurate loop closure detection algorithm is proposed to solve the problem of low accuracy rate of SLAM loop closure detection. The loop closure detection algorithm consists of bag of words module, structure checking module, and tracking module. First, we design the bag of words model and structure checking module, combining local features with holistic features. The bag of words model compares the image similarities using visual words to obtain the closed-loop candidate frame. Then, structure checking module grayscales and normalizes the current image and the closed-loop candidate image. The normalized images are directly used as the patch of local feature to obtain holistic feature. Whether the closed-loop candidate frame is a valid closed loop is determined by the holistic descriptor. To address the problem that time consumption increases rapidly with the increase of image numbers, we design the tracking module to improve the computational efficiency. The comparative experiments with DBoW algorithm show that the proposed algorithm improves the accuracy by more than 20% and also has better real-time performance.

**Keywords:** Simultaneous Localization and Mapping (SLAM); loop closure detection; local features; holistic features; track and predict