

doi:10.16055/j.issn.1672-058X.2020.0003.015

# 一种基于 TF-IDF 和 ORB 的改进闭环检测算法研究\*

荣桂兰, 许 钢, 邢广鑫

(安徽工程大学 检测技术与节能装置安徽省重点实验室, 安徽 芜湖 241000)

**摘 要:**针对视觉 SLAM 闭环检测过程中由于感知歧义导致的闭环不准确问题,基于 TF-IDF 方法提出了一种带权重的计算两幅图像相似性得分的算法,用于视觉 SLAM 的闭环检测研究;首先在特征点检测时,为了得到均匀分布且重叠点较少的特征点,使用 FAST 角点检测方法得到关键点,而后对关键点进行非极大值抑制;其次使用改进的算法计算两幅图像间带权重的相似性得分;最后根据场景图像的特征,进行闭环确认,进一步剔除错误闭环。实验通过搭建平台和使用标准数据集进行测试,证明了改进的闭环检测方法能够有效提高闭环的识别率与准确率。

**关键词:**闭环检测;ORB;TF-IDF 熵;相似性得分函数

**中图分类号:**TP391.4

**文献标志码:**A

**文章编号:**1672-058X(2020)03-0094-06

## 0 引 言

闭环检测对于准确判断机器人位于之前已到达的场景至关重要,能有效减少地图的不准确性<sup>[1-3]</sup>。正确的闭环检测能够消除位姿的累计误差并得到一致性地图,构建精确有效的场景地图;错误的闭环检测则可能导致地图创建的失败,造成导航的失误。

闭环检测主要基于外观和几何两个方面。基于几何的闭环检测准确率完全依赖于位姿估计的准确性,而因为存在观测噪声,所以全局位姿估计存在着巨大的累计误差<sup>[4]</sup>。基于外观的闭环检测则是直接使用当前数据与之前的数据进行相似性计算,满足相似性的一定条件便认为检测到了闭环<sup>[5-6]</sup>。近年来 BOW(词袋)<sup>[7]</sup>模型越来越多的应用于闭环检测,BOW 算法还可以应用于许多领域,

如图像分类等<sup>[8]</sup>。文献[3,9]在 BOW 算法基础上提出了一些用于闭环检测的改进方法,如 Chow-Liu 树和增量式闭环检测算法。近年来有许多学者将深度学习与闭环检测算法联系在了一起,且得到了令人满意的实验结果。Hou 和 Nate 等<sup>[10-11]</sup>在此方面有许多成果。如 Hou 使用 Places-CNN 模型进行闭环检测,证明了深度学习用于闭环检测时对光照变化的鲁棒性优于 BOW 算法;李维鹏等<sup>[12-13]</sup>认为将场景显著区域和空间不确定性约束方法加入闭环检测中能够有效减少闭环检测中的感知歧义;Murartal 等<sup>[14-15]</sup>提出了一种基于 ORB 算法的闭环检测方法,但因为 ORB 算法的尺度一致性较差,所以该方法的闭环检测准确率较低。因此,为了解决感知歧义,本文提出一种改进的方法用以闭环检测的研究。在特征点检测时,使用 FAST 角点检测并用非极大值抑制方法得到均匀分布的且重叠点较少的特征点,防止出现特征点聚集在一起的现象,

收稿日期:2019-08-26;修回日期:2019-09-28.

\* 基金项目:检测技术与节能装置安徽省重点实验室开放基金(2017070503B026-A01);安徽高校自然科学研究重点项目(KJ2018A0111).

作者简介:荣桂兰(1994—),女,安徽淮北人,硕士研究生,从事智能信息处理及应用研究.

并剔除重叠点。在文献[16]的基础上,改进计算两图像间相似性的得分算法,有效减少感知歧义,并在最后根据场景图像的特性进行闭环确认,剔除错误闭环,有效检测出正确的闭环。

## 1 图像特征检测与描述

### 1.1 提取特征点

#### 1.1.1 具有方向的 FAST 特征点检测

FAST 算法的定义:如果在一幅灰度图像中随机取一个像素点的灰度值与其周围邻域内一圈像素点的灰度值进行比较,结果显示该像素点的灰度值比其中连续的 9~12 个像素点的灰度值都大或小,则认为该像素点便是满足算法需要的角点,如图 1 所示。

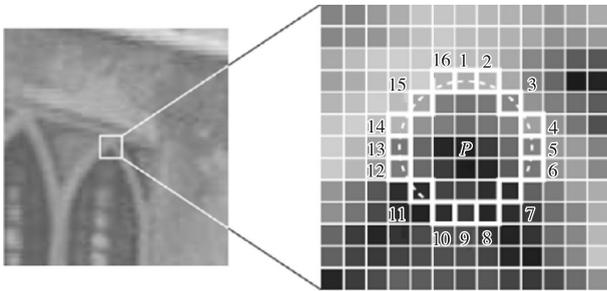


图 1 FAST 角点检测示意图

Fig. 1 Schematic diagram of FAST corner detection

具体操作如下:

- (1) 从灰度图像中随机选取一个像素点  $P$  并且假设该点  $P$  的像素值为  $I_p$ 。
- (2) 以像素点  $P$  为圆心,在该点  $P$  周围一圈内选取像素点,一般选取 16 个像素点。
- (3) 设定一个阈值  $T$ 。
- (4) 当选取的这些像素点中有连续 9~12 个的像素值大于  $I_p+T$  或均小  $I_p-T$  时,那么该点  $P$  是算法需要的特征点。

为了使图像在旋转时也能有精确的结果,所有特征点需要一个主方向。具体方法如下:

- (1) 随机选择一个特征点  $P$  为中心并且为原点。
- (2) 在特征点的领域  $S$  内计算该领域的质心位置。
- (3) 构造一个以点  $P$  为起点,领域  $S$  的质心为

终点的向量,该向量的方向便为此特征点  $P$  的方向。

计算过程如下所示:

$$M_{p,q} = \sum_{(x,y) \in S} x^p y^q f(x,y) \quad (1)$$

式(1)中: $f(x,y)$ 表示图像的灰度值,则此区域质心位置为

$$C = \left( \frac{M_{1,0}}{M_{0,0}}, \frac{M_{0,1}}{M_{0,0}} \right) \quad (2)$$

式(2)中: $M_{0,0}$ 为零阶矩; $M_{1,0}$ 、 $M_{0,1}$ 为一阶矩。因此,特征点的方向为

$$\theta = \arctan \left( \frac{M_{0,1}}{M_{1,0}} \right)$$

#### 1.1.2 非极大值抑制算法

上述算法得到的特征点容易出现聚集、重叠的现象,导致图像上有的地方特征点多,有的地方特征点少,容易出现错误匹配等问题。为了解决这一问题,本文使用非极大值抑制<sup>[17]</sup>方法。该方法的基本思想是能够在—幅图像中检测并提取到目标的局部极大值,并且抑制其他非极大值元素的输出。该方法能够在特征点检测时,剔除掉重叠点,让特征点均匀分布,减少特征点聚集在一起的现象,防止特征点的重叠输出。

### 1.2 计算特征描述子

根据上面求得的关键点,利用 ORB 算法求取描述子,即使用 rBRIEF 描述检测到的关键点。rBRIEF 的思想是从上述步骤得到的关键点中随机选取像素点,而后使用高斯分布采样方式将选取的像素点组成若干像素点对(比如  $k$  组点对,通常设置  $k=256$ ),之后分别比较每一组像素点对中两个像素点的灰度值,而后将比较的结果存储,进而得到一个二进制串,该二进制串便是 BRIEF 的特征描述算子。计算方法如下:

定义一个二值比较函数  $\tau$ :

$$\tau(B;a,b) = \begin{cases} 1, I_a < I_b \\ 0, \text{其他} \end{cases} \quad (3)$$

式(3)中: $I_a$ 、 $I_b$ 为随机像素点对中两个像素点的像素灰度值,选择  $k$  对像素点对进行比较,从而得到二进制编码位串:

$$f_k(I) = \sum_{1 < i < k} 2^{i-1} \tau(B;a,b) \quad (4)$$

式(4)得到的描述子的旋转不变性较差,且该算子对噪声敏感。

针对上述方法不具有旋转不变性的问题,定义一个  $2 \times n$  的矩阵:

$$Q = \begin{pmatrix} x_1, x_2, \dots, x_n \\ y_1, y_2, \dots, y_n \end{pmatrix}$$

其中  $(x_i, y_i)$  为对应的测试点对,特征点的方向  $\theta$  对应的旋转矩阵为  $R_\theta$ , 便能够得到矫正的矩阵  $Q_\theta = R_\theta Q$ , 如此得到的描述子为

$$g_n(p, \theta) = f_n(p) | (x_i, y_i) \in Q_\theta$$

此时得到的特征描述子便会具有较好的旋转不变性。

## 2 改进的闭环检测算法

现实场景中的图像有很多是相似的,相似的场景图像会引起感知歧义,闭环检测时会因为感知歧义而损坏闭环的精度,所以为了将闭环检测的准确率提高,提出一种改进的闭环检测算法,该算法主要包括以下 3 方面内容:创建字典、改进计算两幅图像之间的相似性得分算法、利用场景图像的特性进行闭环确认。

### 2.1 创建字典

BOW 算法应用于很多的闭环检测中,实际场景图像大多需要占据大量的数据,所以创建的视觉字典树也需要满足大规模的需求。为了满足大规模需求,在构建视觉词典树时,采用 K 均值聚类算法。视觉词典树的生成方法如下:

(1) 创建一颗  $k$  个分支数、 $L$  层的树。

(2) 对每一层的每一分支递归调用 K 均值聚类算法,得到  $k$  个子分支,一共调用到  $L$  层。该词典树的单词容量为

$$\sum_{i=1}^L k^i = k^L$$

根据文献[16]定义图像在每个树节点的得分权重可以用各个节点的 TF-IDF 熵表示,定义 TF-IDF 熵为

$$w_i^l(X) = \frac{n_i}{n} \log \frac{N}{N_i} \quad (5)$$

式(5):中  $w_i^l(X)$  表示图像  $X$  在视觉词典树的

第  $l$  层的第  $i$  个节点  $O_i^l$  处的得分权重,  $n$  和  $n_i$  分别表示特征点总数和投影到节点  $i$  上的特征数,  $N$  和  $N_i$  分别表示待处理图像总数和至少有一个特征投影到节点  $i$  上的图像数。

### 2.2 改进的相似性算法

在现实场景中有很多相似的场景难以分辨,往往会引起感知歧义。针对感知歧义问题,在文献[16]的基础上,提出一种改进的计算图像间相似性得分的算法,可以更为精准地表示图像的相似性得分函数。

根据文献[16]可知,通过判断两幅图像间得分权重的最小值表示两幅图像单个节点  $O_i^l$  的相似性得分:

$$S_i^l(X, Y) = \min \{ w_i^l(X), w_i^l(Y) \} \quad (6)$$

根据式(6)的表示,如果存在三幅图像,分别为图像  $X$ 、图像  $Y$ 、图像  $Z$ ,且它们在单个节点  $O_i^l$  处的得分权重为  $w_i^l(Z) > w_i^l(Y) > w_i^l(X)$ ,在节点  $O_i^l$  处的图像  $X$  与图像  $Z$  的相似性得分和图像  $X$  与图像  $Y$  的相似性得分一样,也就是说图像  $X$  与图像  $Z$  和图像  $Y$  的相似性得分相同,即  $S_i^l(Z, X) = S_i^l(Y, X)$ ,可是这是与事实相悖的,因为图像  $X$  与图像  $Y$  的相似性更高。所以为了避免出现这种问题,本文提出了一个改进的算法以求相似性得分函数。

$$S_i^l(X, Y) = \min \{ w_i^l(X), w_i^l(Y) \} + \frac{1}{1 + (w_i^l(X) - w_i^l(Y))^2} \quad (7)$$

由式(7)可以看出:公式的第一项仍旧可以满足大部分图像相似性的判定,而公式的第二项为了不影响整体的判断结果,所以将其值范围控制在  $(0, 1]$  中。公式的第二项为反比例函数,图像  $X$  与图像  $Y$  在节点  $O_i^l$  处的相似性得分与二者 TF-IDF 熵的差值成反比,即二者的 TF-IDF 熵越接近,相似性越高。该方法能够有效地排除感知歧义。

### 2.3 闭环确认

机器人在运动过程中往往会因为摄像机的旋转、伸缩以及数据传输过程中的一些干扰和场景的混乱、尺度的变换及其他多方面因素,导致得到的候选闭环中存在着一些错误闭环,影响建图效果,难以构建精确的地图。所以本文使用场景图像的特性,将时间和空间上的约束作为后验处理的约束

条件用以提高闭环检测的准确性。

机器人的运行是在一段时间内的,所以机器人在运行的过程中得到的图像往往是关于场景的连续变换,所得到的场景图像在时间上也应该是连续的,不是断开的。所以当每一时刻出现闭环时,往往后面几个时刻均发生闭环。因此可以通过此特征剔除错误闭环。比如在  $s$  时刻图像与之前所到达的场景图像发生闭环,那么接下来几个时刻的图像均会发生闭环,如果候选闭环不满足该条件,则删除。

当发生闭环现象时,两幅场景图像对应的应该是同一个场景,只是两幅图像所拍摄的角度及方向是有区别的,所以那两幅图像应该满足对极几何约束。通过计算两幅图像间的基础矩阵,当两幅图像间基础矩阵的内点超过设定的阈值时,则认为这两幅图像存在闭环,否则,图像从候选闭环中删除。

### 3 实验结果及分析

本文方法所用的系统是 Ubuntu 系统,实验所采用的数据集为 TUM 数据集。

#### 3.1 算法检测

在室内做 5 组实验,将本文改进的闭环检测方法方法与基于 SIFT 算法的方法进行比较。在室内获取 100 幅图像,并选取其中 20 幅图像作为检测的图像,将这 20 幅图像与原 100 幅图像进行匹配,计算识别率。实验的部分图像如图 2 所示。

由表 1 可知:在 100 幅图像中,改进算法能够正确检测出 98 幅,平均准确率约达到 98%,而基于 SIFT 算法的平均准确率约为 95%,因此通过本文改进的算法能够较好地达到闭环效果。

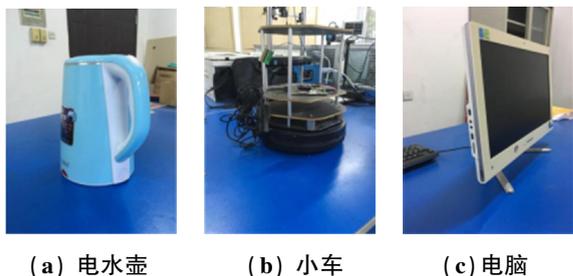


图 2 部分实验场景

Fig. 2 Part of the experimental scene

表 1 闭环检测结果

Table 1 Closed loop test results

编号	改进算法的 正确检测 图像/幅	SIFT 的 正确检测 图像/幅	改进算法的 准确率/%	SIFT 的 准确率/%
1	19	19	95	95
2	20	20	100	100
3	20	19	100	95
4	19	18	95	90
5	20	19	100	95

#### 3.2 算法评估

为了更好地验证本文算法,从 TUM 数据集选取两组 RGBD 数据集进行实验,通过均方根误差 (RMSE),比较改进算法与传统 RGB-D SLAM 算法的不同,表 2 为数据集的基本信息。通过表 3 可知:本文算法的均方根误差 (RMSE) 比 RGB-D SLAM 的均方根误差 (RMSE) 在不同数据集下所得的数据较小。图 3 和图 4 为本文改进算法在两组数据集下的绝对轨迹误差图,通过图 3 和图 4 可以看到估计轨迹与真实轨迹相比还是稍有不同,但是估计轨迹接近于真实轨迹。其中,虚线表示真实的运动轨迹,实线表示算法的估计轨迹。由图 3 可知:改进后的算法可在应用于场景构建时有着量化的优势,且能生成完整性较好的点云地图。

表 2 数据集的基本信息

Table 2 Basic information of the data set

数据集	轨迹	时长/s	平均线	平均角
	长度/m		速度/m/s	速度/deg/s
fre1_desk2	10.161	24.86	0.426	29.308
large-with-loop	39.111	173.19	0.231	17.211

表 3 不同算法的 RMSE 对比

Table 3 RMSE comparison of different algorithms

数据集	轨迹长度/m	RGB-D SLAM	本文方法
fre1_desk2	10.161	0.043	0.026
large-with-loop	39.111	0.456	0.122

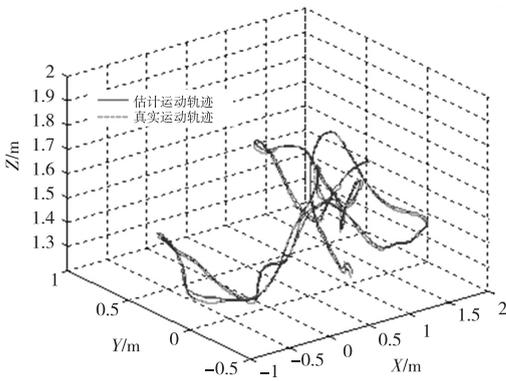


图 3 Fre\_desk

Fig. 3 Fre\_desk

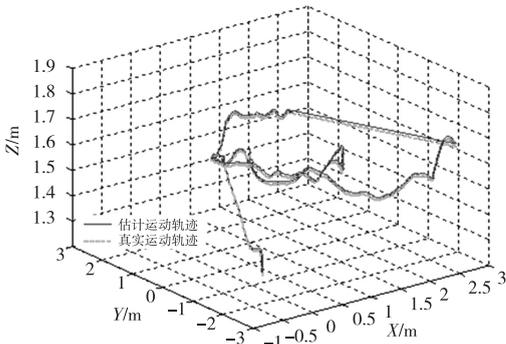


图 4 Large-with-loop

Fig. 4 Large-with-loop

## 4 结 论

针对因感知歧义导致闭环检测的不准确,提出一种改进的 TF-IDF 方法用于视觉 SLAM 的闭环检测研究,改进了计算两幅图像间性相似性得分的算法。首先在特征点检测时,通过改进的方法能够得到均匀分布且重叠点较少的特征点;其次使用改进的算法计算两幅图像的相似性得分;最后根据场景图像的特性,进行闭环确认,进一步剔除错误闭环。对比实验表明:本文提出的改进闭环检测算法相较于基于 SIFT 算法的闭环检测方法在准确率方面有略微的优势,在 TUM 数据集中使用改进的算法能够得到较好的估计轨迹,证明了该改进算法的有效性。

## 参考文献 (References):

[1] BAZEILLE S, FILLIAT D. Combining Odometry and Visual Loop-Closure Detection for Consistent Topometrical Mapping[J]. *RAIRO - Operations Research*, 2010, 44(4):365—377

[2] XIANG G, TAO Z. Unsupervised Learning to Detect Loops Using Deep Neural Networks for Visual SLAM System[J]. *Autonomous Robots*, 2017, 41(1):1—18

[3] WILLIAMS B, CUMMINS M. An Image-to-map Loop Closing Method for Monocular SLAM[C]// *IEEE/RSJ International Conference on Intelligent Robots & Systems*. IEEE, 2008

[4] BEESON P, MODAYIL J, KUIPERS B. Factoring the Mapping Problem: Mobile Robot Map-Building in the Hybrid Spatial Semantic Hierarchy[J]. *The International Journal of Robotics Research*, 2010, 29(4):428—459

[5] 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

[6] SIAM S M, ZHANG H. Fast-Seq SLAM: A Fast Appearance Based Place Recognition Algorithm[C]// *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017

[7] BAEZAYATES R, RIBEIRONETO B. *Modern Information Retrieval*[M]. New York, USA: ACM, 1999

[8] MA H, CHAN J C, SAHA T K, et al. Pattern Recognition Techniques and their Applications for Automatic Classification of Artificial Partial Discharge Sources[J]. *IEEE Transactions on Dielectrics and Electrical Insulation*, 2013, 20(2):468—478

[9] ANGELI A, FILLIAT D, DONCIEUX S. Fast and Incremental Method for Loop-Closure Detection Using Bags of Visual Words[J]. *IEEE Transactions on Robotics*, 2008, 24(5):1027—1037

[10] HOU Y, ZHANG H, ZHOU S. Convolutional Neural Network-Based Image Representation for Visual Loop Closure Detection[C]// *IEEE International Conference on Information and Automation*. Piscataway, USA: IEEE, 2015:2238—2245

[11] MERRIL N, HUANG G. Lightweight Unsupervised Deep Loop Closure[J/OL]. (2018-05-24) [2018-12-29]. <https://arxiv.org/abs/180707703>

[12] 李维鹏, 张国良, 姚二亮, 等. 基于场景显著区域的改进闭环检测算法[J]. *机器人*, 2017, 39(1):23—35

LI W P, ZHANG G L, YAO E L, et al. Improved Closed-loop Detection Algorithm Based on Significant Regions of Scenes[J]. *Robot*, 2017, 39(1):23—35 (in Chinese)

- [13] 李维鹏, 张国良, 姚二亮, 等. 基于空间位置不确定性约束的改进闭环检测算法[J]. 机器人, 2016, 38(3):301—310  
LI W P, ZHANG G L, YAO E L, et al. Improved Closed-loop Detection Algorithm Based on Spatial Position Uncertainty Constraints[J]. Robot, 2016, 38(3):301—310(in Chinese)
- [14] MURARTAL R, TARDOS J D. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras[J]. IEEE Transactions on Robotics, 2017, 37(5):1439—1444
- [15] MURARTAL 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] 李博, 杨丹, 邓林. 移动机器人闭环检测的视觉字典树金字塔 TF-IDF 得分匹配方法[J]. 自动化学报, 2011, 37(6):665—673  
LI B, YANG D, DENG L. Visual Dictionary Tree Pyramid TF-IDF Score Matching Method for Closed-Loop Detection of Mobile Robots[J]. Acta Automatica Sinica, 2011, 37(6):665—673(in Chinese)
- [17] NEUBECK A, GOOL L J V. Efficient Non-Maximum Suppression[C]// International Conference on Pattern Recognition. 2006

## An Improved Closed Loop Detection Algorithm Based on TF-IDF and ORB

**RONG Gui-lan, XU Gang, XING Guang-xin**

(Key Laboratory of Detection Technology and Energy Saving Devices of Anhui Province, Anhui University of Technology, Anhui Wuhu 241000, China)

**Abstract:** Aiming at the closed-loop inaccuracy caused by perceptual ambiguity in the closed-loop detection process of visual SLAM, a weighted algorithm for calculating the similarity scores of two images is proposed based on TF-IDF method, which is used for closed-loop detection of visual SLAM. First, in the feature point detection, in order to obtain feature points that are evenly distributed and have fewer overlapping points, the FAST corner point detection method is used to obtain the key points, and then the key points are subjected to non-maximum suppression. Secondly, an improved algorithm is used to calculate the weighted similarity score between the two images. Finally, according to the characteristics of the scene image, closed loop confirmation is performed to further eliminate the error closed loop. The experiment tests the platform and uses the standard data set to prove that the improved closed-loop detection method can effectively improve the recognition rate and accuracy of the closed loop.

**Key words:** closed-loop detection; ORB; TF-IDF entropy; similarity score function

责任编辑:李翠薇

引用本文/Cite this paper:

荣桂兰, 许钢, 邢广鑫. 一种基于 TF-IDF 和 ORB 的改进闭环检测算法研究[J]. 重庆工商大学学报(自然科学版), 2020, 37(3):94—99

RONG G L, XU G, XING G X. An Improved Closed Loop Detection Algorithm Based on TF-IDF and ORB[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2020, 37(3):94—99