

• 光电子学工程与应用 •



基于原图-光照不变图视觉词典改进的 闭环检测方法

胡章芳, 曾念文*, 罗元, 肖雨婷, 钟征源

(重庆邮电大学光电工程学院 重庆南岸区 400065)

【摘要】当机器人所处环境光照发生变化时, 基于传统视觉词典的闭环检测算法性能会降低, 容易出现感知混叠和感知变异, 从而判断出假闭环。该文首先通过原彩色图像生成只与光源有关的光照不变图, 然后生成原图-光照不变图的视觉词典, 对每帧图像计算两个直方图和相似性得分, 通过最终的得分矩阵来判断是否闭环。实验结果表明, 与传统的视觉词典法相比, 该文提出的闭环检测算法对环境的光照变化具有较好的鲁棒性。

关键词 视觉词袋; 光照不变图; 闭环检测; 同步定位与建图

中图分类号 TP391 文献标志码 A doi:10.12178/1001-0548.2020272

A Method of Loop Closure Detection Improved by Bag-of-Visual Words Based on Original-Illumination Invariant Image

HU Zhang-fang, ZENG Nian-wen*, LUO Yuan, XIAO Yu-ting, and ZHONG Zheng-yuan

(School of Optoelectronic Engineering, Chongqing University of Posts and Telecommunications Nan'an Chongqing 400065)

Abstract When the ambient light of the robot changes, the performance of the loop closure detection algorithm based on the traditional visual word bag will decrease, and it is prone to perceptual aliasing and perceptual variation, thus judging the false closed-loop. In this paper, the original color image is used to generate an illumination invariant image related only to the light source, and then a visual dictionary of the original illumination invariant image is generated. For each image, two histograms and similarity scores are calculated to determine whether it is a closed loop. Finally, it is tested on the data set. The experimental results show that compared with the bag-of-words (BoW), the loop closure detection algorithm proposed in this paper has better robustness to the changes in the environment.

Key words bag-of-visual words; illumination invariant image; loop closure detection; simultaneous localization and mapping

近年来, 同步定位与建图 (SLAM)^[1] 是机器人自动驾驶技术方面的研究热点。在 SLAM 中, 通过后端优化可以不断地修正误差, 但是微小的误差会不断累积。消除累计误差的代表性方法是判断机器人是否“到访”过某个区域, 如果是则视为同一位置, 这就是闭环检测^[2]。闭环检测可以极大地校正漂移误差, 是视觉 SLAM 的重要组成部分。目前大多数研究是基于外观的方法, 利用提取的图像特征和描述符来估计是否到达过某个位置。如 3D-MAP^[3]、卷积神经网络 (convolutional neural networks, CNN)^[4]、视觉词袋法 (bag-of-words, BoW)^[5] 等。视觉词袋法目前在闭环检测中较为常用。文献 [6] 提

出了一种基于外观和颜色组合的双视觉词典方法, 并添加了贝叶斯滤波来判断是否闭环。文献 [7] 结合了词袋模型和 ORB-SLAM 算法, 使用场景分割的方法略微提高了准确率。最富标志性的是 FAB-MAP^[8] 和 DBoW2^[9], 它们使用离线的预训练词汇进行在线测试, 分为离线预处理和在线处理两个阶段。前者从采集到的整个图像数据库中提取视觉单词, 并对其量化构造成一个视觉词典; 后者查找的图像由基于视觉词典的直方图表示, 并与数据库直方图进行比较来判断是否闭环。然而, 移动机器人不可避免地会处于复杂多变的环境中, 包括光照、时间、天气等, 闭环检测是视觉 SLAM 的最重要

收稿日期: 2020-06-30; 修回日期: 2021-03-05

基金项目: 国家自然科学基金青年基金 (67103067); 重庆市教育委员会科学技术研究项目 (KJ1704072)

作者简介: 胡章芳 (1969-), 女, 博士, 教授, 主要从事光电信息处理方面的研究。

通信作者: 曾念文, E-mail: znianwen@qq.com

组成部分, 易受到上述因素的影响, 改善移动机器人在各种条件下 (尤其是在光照变化下) 的鲁棒性, 是视觉 SLAM 不可或缺的工作。

为了减少光照变化对闭环检测的影响, 本文改进了基于视觉词袋法的闭环检测方法。除了将原始的彩色图像作为视觉词典, 还将彩色图转为光照不变图, 同时并行生成光照不变图的视觉词典, 并对其直方图进行比较, 最后计算两者的最终得分矩阵来判断是否为闭环。实验结果表明, 本文算法能稳健应对光照变化较大的场景。

1 图像处理和词典构建

1.1 原图特征提取

考虑到系统的性能和运行时间要求, 本文使用 ORB^[10] 特征点, 选择描述符长度为 256 bit, 每位描述符的计算公式如下:

$$v_i(p) = \begin{cases} 0 & p + x_i < p + y_i \\ 1 & \text{otherwise} \end{cases} \quad \forall i \in [1, 256] \quad (1)$$

式中, v_i 是第 i 个位描述符的值; x_i 、 y_i 是通过高斯分布在该点周围预先建立的随机选择位置的灰度值; p 代表正在计算的点。本文使用汉明距离来计算描述符之间的距离:

$$d_{\text{Ham}}[v(p), v(q)] = \sum_i^{256} \text{XOR}[v(p), v(q)] \quad (2)$$

式中, $v(p)$ 、 $v(q)$ 是描述符; XOR 表示异或操作。获得特征和描述符后, 将它们转换为视觉单词, 并将图像转换为稀疏向量。

1.2 光照不变图的生成

文献 [11] 提出了一种算法, 使得图像外观变化只与光源有关。根据文献 [11], 光谱灵敏度 $F(\lambda)$, 图像传感器的光照响应 R , 物体的表面反射率 $S(\lambda)$ 和物体上的发射光谱功率分布 $E(\lambda)$ 之间存在以下关系:

$$R^{x,E} = \mathbf{a}^x \cdot \mathbf{n}^x I^x \int S^x(\lambda) E^x(\lambda) F(\lambda) d\lambda \quad (3)$$

式中, 单位向量 \mathbf{a}^x 和 \mathbf{n}^x 表示物体反射光源的方向和表面法线的方向, 取决于物体的材料特性; I^x 表示物体点 x 上反射光的强度。将光谱敏感度函数 $F(\lambda)$ 假设为以波长 λ_i 为中心的狄拉克增量函数^[12], 从而产生以下响应函数:

$$R^{x,E} = \mathbf{a}^x \cdot \mathbf{n}^x I^x S^x(\lambda_i) E^x(\lambda_i) \quad (4)$$

为减少光照强度 I^x 的影响, 得到取决于表面反射率 $S(\lambda_i)$ 的光照不变图 ζ , 将式 (4) 取对数:

$$\log(R^{x,E}) = \log\{\mathbf{G}^x I^x\} + \log\{S^x(\lambda_i)\} + \log\{E^x(\lambda_i)\} \quad (5)$$

将光照近似为普朗克光源^[13], 再将维恩常数近

似代入普朗克光源中:

$$\log(R^{x,E}) = \log\{\mathbf{G}^x I^x\} + \log\{2hc^2 \lambda_i^{-5} S_i\} - \frac{hc}{k_B T \lambda_i} \quad (6)$$

式中, $\mathbf{G}^x = \mathbf{a}^x \cdot \mathbf{n}^x$; h 为普朗克常数; c 为光速; k 为玻尔兹曼常数; T 是黑体源的相关色温。

文献 [11] 使用一维色彩空间 ζ , 该色彩空间由相机传感器在有序波长 $\lambda_1 < \lambda_2 < \lambda_3$ 下的峰值灵敏度 R_1 、 R_2 、 R_3 决定:

$$\zeta = \log(R_2) - \alpha \log(R_1) - (1 - \alpha) \log(R_3) \quad (7)$$

将式 (4) 代入式 (5), 当参数 α 满足以下约束时, 一维色彩空间 ζ 与相关色温 T 无关:

$$\frac{hc}{k_B T \lambda_2} - \frac{hc}{k_B T \lambda_1} - \frac{(1 - \alpha)hc}{k_B T \lambda_3} = 0 \quad (8)$$

简化为:

$$\frac{1}{\lambda_2} = \frac{\alpha}{\lambda_1} + \frac{1 - \alpha}{\lambda_3} \quad (9)$$

因此, 可根据相机传感器取适当的 α 值, 将原始图像转化为光照不变图像 ζ 。

1.3 视觉词典的构建

先对整个图像序列进行光照不变变换, 再生成原图-光照不变图视觉词典。

将采集到的彩色图像设为集合 S_{RGB} , 将光照不变图设为集合 S_{II} , 两者存在以下关系:

$$S_{\text{II}} = \zeta(S_{\text{RGB}})$$

从每个图像集中提取特征集 F_{RGB} 和 F_{II} , 该特征集由位置 x_i 、 s_i 和特征描述符 d_i 组成。

最后, 使用如图 1 所示的树数据结构^[5] 来构建词典, 在每一层使用 K-means 聚类方法对描述符进行分类。其中只有叶节点存储可视单词, 而中间节点仅用于查找单词。该词典的总容量为 K_d 。搜索特定的视觉词汇时, 只需要与聚类中心进行 d 次比较即可完成查找, 时间复杂度为 $O(\log N)$, 保证了搜索效率。同时, 为了区分每个单词的重要性, 使用直接索引法^[5], 该方法将单词的父节点存储在目录中以加快比较速度。

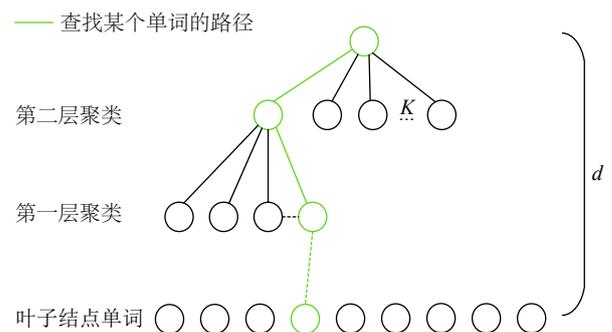


图 1 词袋树结构模型图

综上可以生成基于原图-光照不变图视觉词典, 图2总结了该算法。

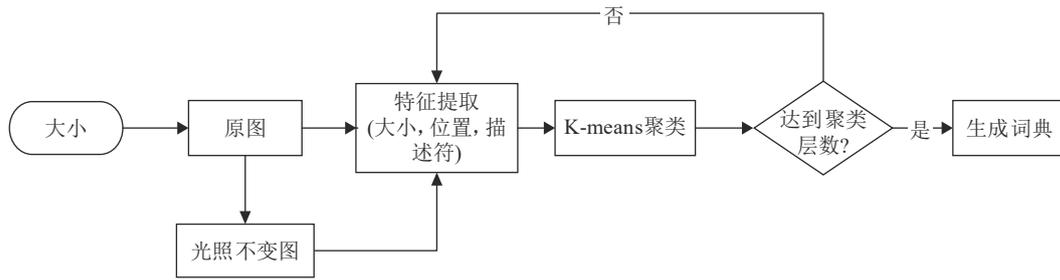


图2 原图-光照不变图词典生成框图

2 闭环检测算法

2.1 单词筛选

在提取图像特征和生成词典后, 再利用这些数据完成闭环检测。尽管视觉描述符已经转换为视觉单词, 但并非所有视觉单词都具有相同的识别和区分度。如某些单词很常见, 在许多帧中都可以找到, 这种情况下词类视觉单词在闭环检测中基本无作用。因此本文采用词频-逆文本频率指数 (term frequency-inverse document frequency, TF-IDF)^[14] 方法来区分不同单词的重要性。在词袋模型中, 可以在处理图像之前计算 IDF, 即在构建字典时确定 IDF 属性。将视觉单词 w_i 总数用 n 表示, 含有 w_i 的图像数量用 n_i 表示, 则 w_i 的逆文档频率为:

$$\text{IDF}_{(i)} = \log \frac{n}{n_i}$$

文档频率为:

$$\text{TF}_{(i)} = \frac{n_i}{n}$$

视觉单词 w_i 的权重为:

$$\beta_i = \text{TF}_{(i)} \text{IDF}_{(i)}$$

在考虑单词的权重之后, 对于特定的图像 A , 将原图特征和光照不变图特征添加到词袋中, 完善原图-光照不变图词典:

$$A = \{(F_{\text{RGB}1}, F_{\text{I}11}, \beta_1)(F_{\text{RGB}2}, F_{\text{I}12}, \beta_2) \cdots (F_{\text{RGB}n}, F_{\text{I}1n}, \beta_n)\}$$

2.2 候选帧的筛选和回环帧确定

根据前文生成的视觉词典, 对每帧图像计算两个直方图。为了快速进行两幅图像间的直方图比较, 本文采用直方图交叉核 (histogram intersection kernel)^[15] 方法来测量两个矩阵的相似性得分。其原理是先将图像特征在多分辨率的超平面上进行映射, 进而生成多层次的直方图, 最后进行相似度的

加权叠加。其交集函数 (核函数) 的定义如下:

$$\Gamma(H(h_1), H(h_2)) = \sum_{j=1}^r \min((H(h_1))_j, (H(h_2))_j)$$

式中, h_1, h_2 为欲比较的两个直方图; $H(h_1)_j$ 为 h_1 直方图中的第 j 个 bin; r 为直方图中 bin 的个数。两个 bin 的最小值为两个直方图每个 bin 的重叠数, 所有 bin 重叠数之和为该层次的交集函数值。直方图的相似性函数定义为:

$$K_{\Delta}(\psi(h_1), \psi(h_2)) = \sum_{i=0}^L \frac{1}{2^i} N_i$$

其值可以用来比较相似度。式中, L 表示直方图的层数; N_i 表示连续量的层次之间的交集函数值的差: $N_i = \Gamma(H_i(h_1), H_i(h_2)) - \Gamma(H_{i-1}(h_1), H_{i-1}(h_2))$ 。

定义两个矩阵间的相似性得分 $S = K_{\Delta}$, 完成两个矩阵的相似性得分计算以后, 对得分进行归一化:

$$S(h_i, h_{i'}) = \frac{S(h_i, h_{i'})}{S(h_i, h_{i-1})}$$

生成最终的得分为:

$$S_{\text{final}} = S_{\text{RGB}} + \eta \cdot S_{ii}$$

式中, S_{RGB} 是原始彩色图像的视觉词袋归一化得分矩阵; S_{ii} 是得到的光照不变图像的视觉词袋归一化得分矩阵; S_{final} 是最终的得分矩阵; η 是两个得分矩阵间的平衡因子。

由于相邻两幅图像十分相似, 容易误判为闭环。因此, 本文将图像序列根据规模进行分组。每个组通过累加组内每个候选帧的相似性得分来获得累加的得分。累加分最高的组视为闭环组, 闭环组中得分最高的图像将作为最佳候选帧。选择分数大于得分阈值的最佳候选帧作为判定正确的闭环对,

具体的算法伪代码如下:

输入: D 为前文生成的视觉词典

F 为提取的图像特征

T 为得分矩阵的阈值

输出: 闭环对 P

initialize P to empty

for all i

$h_{RGB,i}$ = nearest neighbors($F_{RGB,i}, D_{RGB}$)

$h_{II,i}$ = nearest neighbors($F_{II,i}, D_{II}$)

end for

for all i, j

$h_{RGB,i}$ = nearest neighbors($F_{RGB,i}, D_{RGB}$)

$h_{II,i}$ = nearest neighbors($F_{II,i}, D_{II}$)

end for

S_{final} = normalize(S_{RGB}) + η · normalize(S_{II})

for all i, j

if $S_{final}(i, j) > T$ and i, j are not adjacent frames

add(i, j) to P

end if

end for

3 实验结果

3.1 数据集选择

为了验证本文算法对光照的鲁棒性, 需要捕获同一场景在不同照明条件下的图像。因此, 本文选择了 Nordland 数据集, 该数据集采集自挪威北部 4 个不同季节, 分为春夏秋冬 4 个部分, 其相同场景在不同时间的比较如图 3 所示。同一场景随着时间的改变发生了很大的变化。本文首先选择光照变化最为显著的春季和冬季序列作为实验序列。



图3 同一场景在不同时刻的表现形式

3.2 评价指标

第一个评价指标是真(假)阴(阳)性, 当机器人经过同一位置时, 闭环检测算法应给出“是闭环”(真阳性)的结果, 反之则应该给出“不是闭环”(真阴性)的结果。

另外一个评价指标是准确率(precision)和召回率(recall)。准确率是指某个算法中检测到的闭环是真实闭环的概率, 即 $Precision = TP / (TP + FP)$, 召回率是指在所有真实闭环中被正确检测出来的概率, 即 $Recall = TP / (TP + FN)$ 。准确率-召回率是一对矛盾, 当算法设计得比较“严格”时, 准确率可以达到 100%, 但必然会造成召回率下降。因此, 在确定一个指标后, 若某算法得到的另一个指标比其他算法高, 则说明该算法更优。

3.3 图像匹配结果

为了直观地分析图像匹配的相似性规律, 本文与闭环检测中最常用的 BoW^[5] 算法和文献 [16] 提出的一种基于改进 TF-IDF 的 SLAM 回环检测算法进行了比较。首先在 Nordland 数据集的冬季序列中选取了 800 张在视觉上有明显差异的图像, 接着

在春季序列中选择了一张图像, 然后逐一匹配冬季序列中的图像并计算相似度, 获得的相似度曲线如图 4 所示。横轴表示冬季序列索引, 纵轴是与选取的春季图像的相似度值。曲线的最大值表示算法匹配最为相似的图像, BoW 算法、文献 [16] 算法、本文算法计算出的最大相似度图像分别位于第 29、200、482 帧处。

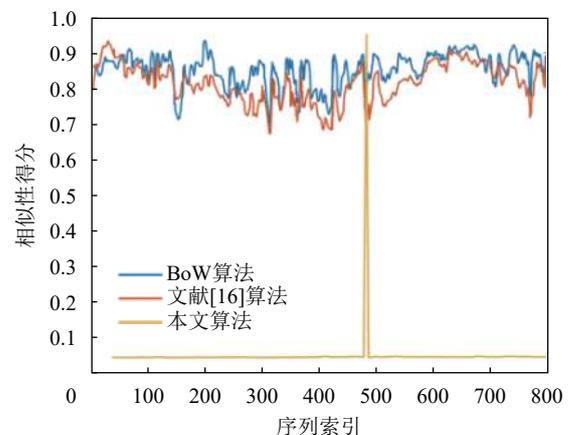


图4 相似性得分曲线

选取的春季序列图像与 3 种算法匹配的图像如

图5所示。由图5可知,当同一场景的光照发生变化时,本文算法匹配到了同一场景的图像,而其他算法都匹配到了错误的图像。

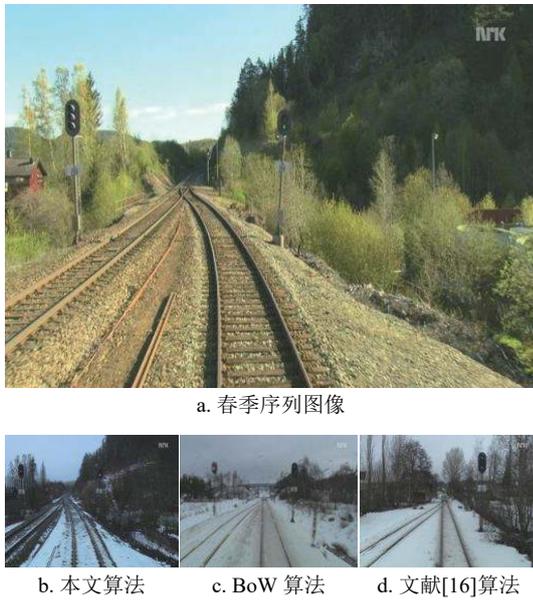


图5 选取的图像和算法匹配的图像

3.4 闭环检测性能

为了进一步验证本文算法对光照变化的鲁棒性,继续从春季选择了800个序列图像,从冬季选择了800个序列图像。将序列中的一个图像与另一序列中的所有图像进行比较。每两个序列计算一次相似度得分,然后根据编号排列得分,得到相似度矩阵 S 。显然, S 是一个对称矩阵。离对角线越近意味着图像与自身和连续帧最为相似,得到的相似度矩阵结果如图6所示。本文算法基于原图和光照不变图,而不是简单地进行一次图像的特征匹配,获得的相似度矩阵基本上沿对角线分布,且噪点较少;而BoW算法出现了多处对角线较为稀疏的地方,如图中圆圈标识;文献[16]方法获得的相似度矩阵有很多噪声点,且对角线较为稀疏,表示这是错误的闭环。因此,本文算法在光照变化明显的环境中表现良好。

3.5 召回率测试

本部分使用Northland数据集的夏季和冬季序列进行词典构建,然后选择春季序列进行召回率测试,在100%准确率的前提下实现较高的召回率,夏季序列和冬季序列的结果分别如图7a、图7b所示。在冬季序列中,部分场景被积雪覆盖,可提取的图像特征减少,本文提出的光照不变图算法表现更佳。

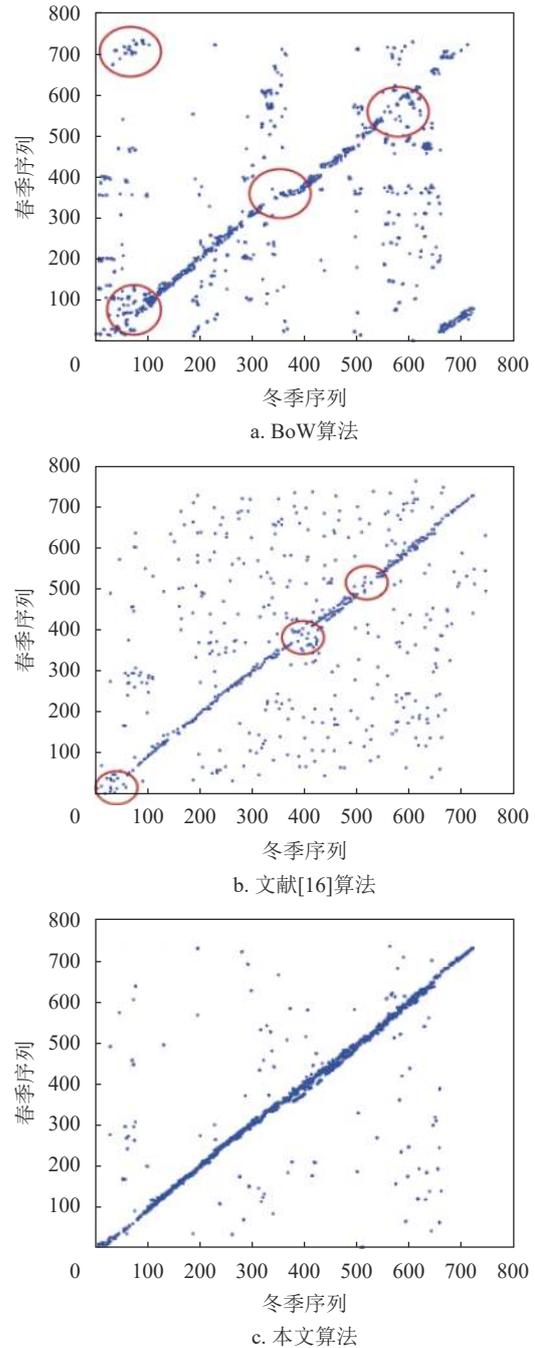
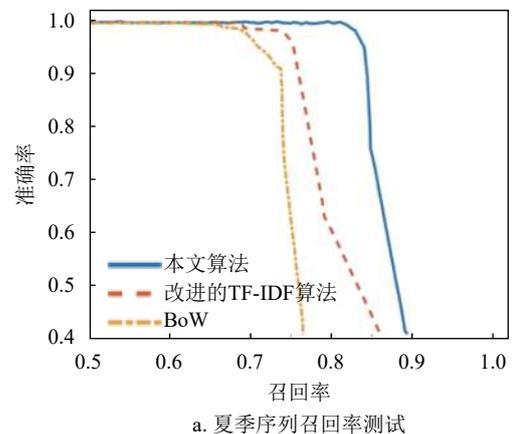
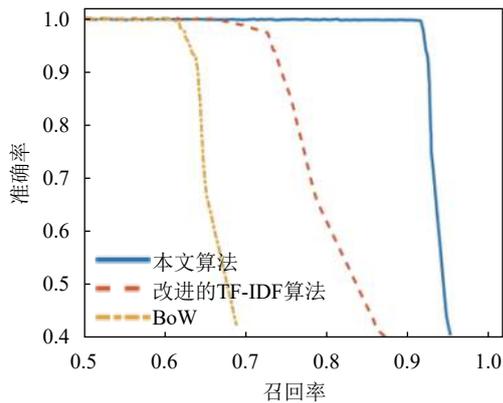


图6 3种算法的相似度矩阵



a. 夏季序列召回率测试



b. 冬季序列召回率测试

图7 召回率测试结果

4 结束语

视觉词袋模型因为其计算速度快且图像表示形式简单, 在闭环检测中得到了广泛的应用, 但是光照的变化会降低其鲁棒性。为了克服这个缺点, 本文在传统的词袋模型上进行了改进, 通过生成原图-光照不变图的视觉词典, 归一化计算最终矩阵相似性得分来提高对光照变化的鲁棒性。通过在数据集上和其他算法进行比较, 证明了本算法对光照变化具有更好的鲁棒性。最近热门的室外自动驾驶, 在长时间运作时, 可能会由于阳光的变化而产生错误闭环检测, 本文算法可适用于此类情况。然而, 在实验中发现由于另一个视觉词典的加入, 使得本文算法的实时性有所下降, 如何在光照鲁棒性的前提下保证速率的问题将是下一步研究的重点。

参 考 文 献

- [1] SAPUTRA M R U, MARKHAM A, TRIGONI N. Visual SLAM and structure from motion in dynamic environments: A survey[J]. *ACM Computing Surveys*, 2018, 51(2): 1-36.
- [2] 刘强, 段富海, 桑勇, 等. 复杂环境下视觉 SLAM 闭环检测方法综述[J]. *机器人*, 2019, 41(1): 112-123, 136.
LIU Qiang, DUAN Fu-hai, SANG Yong, et al. A survey of loop-closure detection method of visual SLAM in complex environments[J]. *Robot*, 2019, 41(1): 112-123, 136.
- [3] DING Jie, FANG Yong-chun. Multi-strategy based exploration for 3D mapping in unknown environments using a mobile robot[C]//2019 Chinese Control Conference (CCC). Guangzhou: IEEE, 2019: 4732-4738.
- [4] GAO Xiang, ZHANG Tao. Unsupervised learning to detect loops using deep neural networks for visual SLAM system[J]. *Autonomous Robots*, 2017, 41(1): 1-18.
- [5] GALVEZ-LO PEZ 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.
- [6] ANGELI A, FILLIAT D, DONCIEUX S, et al. Fast and incremental method for loop-closure detection using bags of visual words[J]. *IEEE Transactions on Robotics*, 2008, 24(5): 1027-1037.
- [7] CHEN J, LI J, XU Y, et al. A compact loop closure detection based on spatial partitioning[C]//International Conference on Image Vision and Computing. [S.l.]: IEEE, 2017: 371-375.
- [8] NOWICKI M, WIETRZYKOWSKI J, SKRZYPCZYNSKI P. Experimental evaluation of visual place recognition algorithms for personal indoor localization[C]//2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN). [S.l.]: IEEE, 2016, DOI: 10.1109/IPIN.2016.7743649.
- [9] JIE W, ALPER Y, LEI Y. DCF-BoW: Build match graph using bag of deep convolutional features for structure from motion[J]. *IEEE Geoscience and Remote Sensing Letters*, 2018(99): 1-5.
- [10] KYI P, YUTTANA K. Biomedical images stitching using ORB feature based approach[C]//2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS). [S.l.]: IEEE, 2018, DOI: 10.1109/ICIIBMS.2018.8549931.
- [11] MADDERN W, STEWART A D, MCMANUS C, et al. Illumination invariant imaging: Applications in robust vision-based localisation, mapping and classification for autonomous vehicles[C]//IEEE International Conference on Robotics and Automation (IEEE ICRA). Hong Kong, China: IEEE, 2004: 5-12.
- [12] FINLAYSON G D, HORDLEY S D. Color constancy at a pixel[J]. *J Opt Soc Am A*, 2001, 18(2): 253-264.
- [13] RATNASINGAM S, COLLINS S. Study of the photodetector characteristics of a camera for color constancy in natural scenes[J]. *J Opt Soc Am A*, 2010, 27(2): 286-294.
- [14] RABIE K, ADEBISI B, GACANIN H, et al. Energy-per-bit performance analysis of relay-assisted power line communication systems[C]//IEEE Transaction on Green Communications and Networking. [S.l.]: IEEE, 2018: 360-368.
- [15] LI Pan, LIU Yang, LIU Guo-jun, et al. A robust local sparse coding method for image classification with histogram intersection kernel[J]. *Neurocomputing*, 2016, 184: 36-42.
- [16] 董蕊芳, 柳长安, 杨国田. 一种基于改进 TF-IDF 的 SLAM 回环检测算法[J]. *东南大学学报(自然科学版)*, 2019, 49(2): 251-258.
DONG Rui-fang, LIU Chang-an, YANG Guo-tian. TF-IDF based loop closure detection algorithm for SLAM[J]. *Journal of Southeast University (Natural Science Edition)*, 2019, 49(2): 251-258.

编辑 税红