

DOI: 10.13973/j.cnki.robot.2017.0036

基于 SURF 和 ORB 全局特征的快速闭环检测

刘国忠¹, 胡钊政^{1,2}

(1. 河北工业大学电子信息工程学院, 天津 300401; 2. 武汉理工大学智能交通系统研究中心, 湖北 武汉 430063)

摘要: 针对移动机器人 SLAM (同时定位与地图创建) 中的闭环检测问题, 提出了一种基于 SURF (加速鲁棒特征) 和 ORB (oriented FAST and rotated BRIEF) 全局特征的快速闭环检测算法. 首先利用 SURF 与 ORB 分别提取查询图像的全局特征, 实现对图像的特征表征. 在特征提取过程中, 对查询图像进行归一化操作, 并将归一化的图像中心直接作为 SURF 与 ORB 的特征点位置, 从而避免了耗时的特征点定位过程. 然后将归一化的图像直接作为特征点的邻域区域, 把计算的 SURF 与 ORB 局部特征描述符作为图像的全局特征. 为了融合 SURF 与 ORB 全局特征实现查询图像的位置识别, 提出了 H-KNN (混合 K 最近邻) 的改进算法: WH-KNN (加权混合 K 最近邻). 最后通过跟踪模型实现闭环检测, 其核心思想是利用之前闭环检测的结果预测查询图像在地图图像中的位置范围. 实验中分别使用采集数据集和牛津大学公开数据集对本文算法进行了验证, 同时与传统的 BOW (词袋) 算法进行了对比. 本文算法在两种数据集上分别达到了 94.3% 和 94.5% 的准确率, 并且查询图像位置识别与全局特征提取的平均时间不到 3 ms. 其准确性及计算速度都超过了 BOW 算法, 可以准确快速地实现实时闭环检测.

关键词: 同时定位与地图创建; 闭环检测; 机器人; 全局特征; 跟踪模型

中图分类号: TP242

文献标识码: A

文章编号: 1002-0446(2017)-01-0036-10

Fast Loop Closure Detection Based on Holistic Features from SURF and ORB

LIU Guozhong¹, HU Zhaozheng^{1,2}

(1. School of Information Engineering, Hebei University of Technology, Tianjin 300401, China;

2. ITS Research Center, Wuhan University of Technology, Wuhan 430063, China)

Abstract: A fast loop closure detection algorithm based on holistic features from SURF (speeded-up robust feature) and ORB (oriented FAST and rotated BRIEF) is proposed for the loop closure detection problem in mobile robot SLAM (simultaneous localization and mapping). Firstly, holistic features of a query image are extracted for feature representation by using SURF and ORB. The query image is normalized and its center is directly set as the feature point position of SURF and ORB in the process of extracting holistic features, so that the time-consuming step of feature point localization is avoided. Afterwards, the normalized image is directly used as the patch of the feature point. The local feature descriptors computed from SURF and ORB are then used as the holistic features of the image. In order to fuse the holistic features from SURF and ORB for image matching, an improved method of H-KNN (hybrid K-nearest neighbor) called WH-KNN (weighted hybrid K-nearest neighbor) is also proposed. Finally, loop closure detection is finished by a tracking model. The core idea behind the model is to utilize the previous loop closure detection results to predict the position of the query image in the map. In the experiments, the proposed algorithm is not only tested with the collected datasets and the public datasets of Oxford University, but also compared with the classic BOW (Bag of Words) algorithm. It can achieve 94.3% and 94.5% detection accuracy on these two datasets, respectively. In average, it takes less than 3 ms to match each query image and extract holistic features. The test results demonstrate that the proposed algorithm outperforms the BOW method both in term of detection accuracy and computation efficiency. Therefore, the algorithm is accurate and fast enough for real-time loop closure detection.

Keywords: simultaneous localization and mapping; loop closure detection; robot; holistic feature; tracking model

1 引言 (Introduction)

随着机器人技术的快速发展, 移动机器人正在逐步走向完全自主化, 而 SLAM (同时定位与地图

创建) 是机器人能否真正实现完全自主化的关键技术之一^[1-3].

闭环检测^[4]是 SLAM 研究中的一个基础问题, 准确有效地判断机器人是否位于之前已经去过的区

基金项目: 国家自然科学基金 (51679181, 51208168); 河北省普通高等院校青年拔尖人才计划 (BJ2014-013); 湖北省技术创新专项重大项目 (2016AAA007); 湖北省自然科学基金 (2015CFB252).

通信作者: 胡钊政, zhaozheng.hu@gmail.com 收稿/录用/修回: 2016-07-04/2016-11-17/2016-11-21

域, 对减少机器人运动累积误差以及后端优化等具有至关重要的作用. 闭环检测也是视觉定位^[5]系统的最核心技术之一. 闭环检测以其重要性及挑战性吸引了广大科研学者参与研究. 在闭环检测的方法研究中, 大多数采用基于视觉的方法. 根据不同的数据关联方法, 将基于视觉的闭环检测方法大致分为 3 类: (1) 地图与地图间匹配的闭环检测方法. 例如, Clemente 等人^[6]提出寻找子图特征间的相似性来实现闭环检测, 通过场景特征间的相似度, 获得子图间的公共特征. 该方法的缺点是需要稀疏地图上执行, 包含信息较少, 难以准确地实现闭环检测. (2) 图像与地图间匹配的重定位闭环检测方法. 例如, Williams 等人^[7]提出通过寻找当前图像与地图之间的相似性, 来获得当前图像在地图中的位置. 该方法的缺点是, 利用分类器进行分类时占用的内存过多, 因此不适合用于大规模环境. (3) 图像与图像间匹配的闭环检测方法. 例如 Cummins 等人^[8]提出提取已去过地方的场景图像特征, 将特征形成视觉表像实现闭环检测.

近几年出现了越来越多的基于 BOW (词袋) 模型的闭环检测算法^[9-10]. 例如, 李博等人^[11]提出了一种基于视觉词典树的金字塔 TF-IDF (term frequency-inverse document frequency) 得分匹配方法, 李永峰等人^[12]在此方法基础上进行了改进, 通过构建历史模型集提高闭环检测的准确率和实时性. 李维鹏等人^[13]提出了一种基于空间位置不确定性约束的改进闭环检测算法, 通过空间约束减小闭环检测的范围, 还排除了大部分的感知歧义. 另外, BOW 算法在特征分析、模式识别、图像分类等领域均得到了广泛的应用^[14-15].

在基于 BOW 闭环检测的方法中, 一张图像被当作字典中词汇出现频率的直方图. 通常情况下, 先把地图图像中提取的特征经过聚类来产生词典. 然后通过比较查询图像与地图图像的直方图之间的相似性, 来判断查询图像对应的机器人拍摄位置是否为已去过的区域, 这是目前闭环检测最主流的方法之一. 该方法的缺点是在构建词典前需要提取局部特征进行训练, 随着机器人地图的扩建, 其计算量会急剧增大, 不具有实时性和可靠性^[16].

计算机视觉研究领域的进步, 有效地促进了局部特征算子的发展. 其中, SIFT (尺度不变特征变换)^[17]和 SURF^[18]都是尺度和旋转不变的检测子和描述子, 而 ORB^[19]是旋转不变描述子, 对尺度不具有不变性, 但是速度比 SIFT 快 2 个数量级. 在 SLAM 过程中如何对图像特征进行选择 and 处

理对 SLAM 的实时性和高效性非常关键. 龙超等人^[20]采用 GPU-SURF (GPU accelerating SURF) 进一步提高了特征提取的速度. Mur-Artal 等人^[21]利用 ORB 描述子完成了所有的 SLAM 工作, 包括目标追踪、构建地图、定位和闭环检测, 并达到了良好的执行效果.

综上所述, 在 SLAM 闭环检测领域尽管取得了不少研究成果, 但是目前主流的基于 BOW 的闭环检测算法一般需要在一张图像上提取成千上万个特征点以形成稳定的直方图表征. 而无论是 SIFT、SURF 还是 ORB 算法, 在特征点定位上均需要消耗大量的时间, 而在特征描述符计算上则速度较快. 此外, 特征聚类生成视觉词典的计算量也非常庞大, 系统的实用性和准确率不能令人满意. 因此, 提出了一种准确快速的闭环检测算法——基于 SURF 和 ORB 全局特征的闭环检测. 本文算法首先对查询图像进行灰度化、均衡化、归一化处理, 并将归一化的图像中心直接作为 SURF 与 ORB 的特征点位置, 从而避免了繁琐的特征点定位过程, 极大提高了计算速度. 在此基础上将归一化的图像直接作为 SURF 与 ORB 特征点的邻域区域, 把计算的 SURF 和 ORB 局部特征描述符作为图像的全局特征. 本文还提出基于 H-KNN^[22]的改进算法 WH-KNN, 对 SURF 与 ORB 全局特征进行有效融合, 实现查询图像与地图图像的特征匹配. 最后, 通过构建跟踪模型获取候选集, 减小了闭环检测的范围, 从而显著提高了闭环检测的准确性和实时性. 在所采集的室内数据集和牛津大学的公开数据集^[23]上验证了文中提出的算法, 并且与基于 BOW 的闭环检测方法进行了对比, 结果表明本文算法在实时性和准确率方面更适合用于闭环检测.

2 本文算法 (The proposed method)

本文算法的流程图如图 1 所示. 先对载入的查询图像进行灰度化及均衡化处理, 然后对其进行归一化操作. 再将图像中心直接作为特征点, 提取图像的 SURF 与 ORB 全局特征. 在此基础上, 构建跟踪模型来获取查询图像位置所属的候选集, 然后采用 WH-KNN 算法在候选集中计算与查询图像最相近的 M 张图像 (SURF 全局空间) 和 N 张图像 (ORB 全局空间), 最终把 $M+N$ 张图像组成新的判决空间, 实现闭环检测. 当闭环检测结束后, 利用闭环检测的结果对跟踪模型进行更新, 从而产生新的候选集.

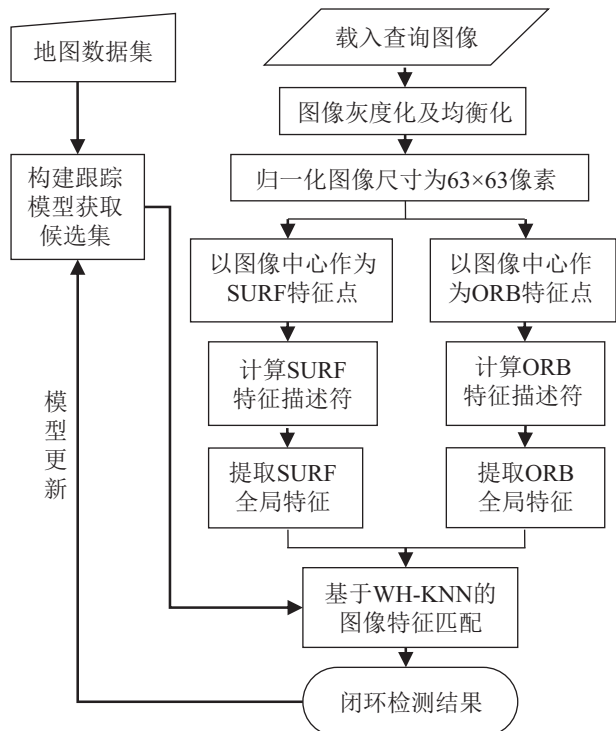


图1 本文算法流程图

Fig.1 Flow chart of the proposed method

2.1 基于 SURF 的全局特征计算

SURF 算子是由 SIFT 算子改进而来, SURF 不仅在运行速度上优于 SIFT, 并且在视角、光照、尺度变化等情形下具有更好的鲁棒性^[17-18]. SURF 算法主要包含两部分: 第一部分是特征点定位, 第二部分是特征描述符计算. 本文把归一化后的图像的中心点设定为 SURF 特征点位置, 从而避免了耗时的特征点定位过程, 进而直接把计算出的局部特征描述符当作图像的全局特征.

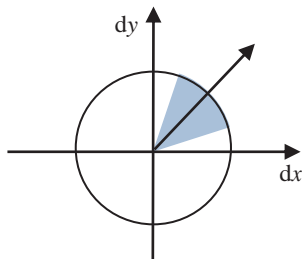


图2 SURF 特征点主方向

Fig.2 Main orientation of the SURF feature point

为了计算 SURF 描述符, 需要计算特征点的主方向信息, 如图 2 所示. 首先计算特征点附近圆形区域内的像素在 x 与 y 方向的 Haar 小波响应, 并对这些响应赋予高斯权重系数, 使靠近特征点的响应贡献大, 远离特征点的响应贡献小. 之后小波响应 dx 和 dy 被表示成一个向量, 该向量包括小波响应在水平和垂直两个方向的信息. 再在圆形区域中计

算 60° 滑动窗内像素的 x 与 y 方向的小波响应之和 m_w :

$$m_w = \sum_w dx + \sum_w dy \quad (1)$$

滑动窗以每步 5° 的步幅转动, 计算合成向量角度:

$$\theta_w = \arctan\left(\frac{\sum_w dx}{\sum_w dy}\right) \quad (2)$$

求出各方向扇形的合成向量模长最大值:

$$l_{\max} = \sqrt{\left(\sum_w dx\right)^2 + \left(\sum_w dy\right)^2} \quad (3)$$

其中向量模长最大值对应的角度即特征点主方向. 获取特征点方向后, 把特征点附近的正方形邻域划分为 4×4 个子区域, 对每个子区域计算 Haar 小波响应, 并对这些响应赋予高斯权重系数. 然后将每一个子区域内的 dx 、 dy 分别求和. 为了使特征描述符具有强度变化的极性信息, 同时也对 dx 和 dy 的绝对值分别求和. 因此, 每一个子区域的特征描述符 v 计算如下^[18]:

$$v = \left(\sum dx, \sum dy, \sum |dx|, \sum |dy|\right) \quad (4)$$

所有的 16 个子区域的描述符合起来形成此特征点处的特征描述符. 最终的 SURF 描述符是一个 64 维的向量^[18]:

$$V = [v_1, v_2, \dots, v_{15}, v_{16}] \quad (5)$$

其中 v_i ($i = 1, 2, \dots, 15, 16$) 是第 i 个子区域的特征描述符.

计算 SURF 全局特征的主要思想是把预处理图像的中心直接作为特征点的位置, 并把图像整体作为特征点的邻域区域, 最后把计算得到的 SURF 局部描述符作为图像的全局特征描述符. 由于不需要构建图像金字塔, 进一步地加快了 SURF 特征提取的速度. 图 3 是对室外场景图像提取 SURF 全局特征描述符的过程, 左边为预处理后的图像, 红色圆代表图像中心, 右边为 SURF 特征全局描述符.

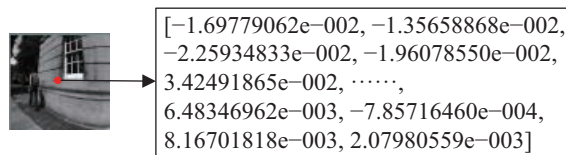


图3 SURF 全局特征提取 (1x64 向量)

Fig.3 Extraction of SURF holistic features (1x64 vectors)

2.2 基于 ORB 的全局特征计算

ORB 算子包含 FAST (feature from accelerated segment test) 特征点的检测方法与 BRIEF (binary

robust independent elementary feature) 特征描述子两部分, 见文 [19]. 同样, 由于直接将归一化的图像中心作为 ORB 特征点的位置, 因此可以直接忽略 FAST 特征点检测过程, 从而极大的提高计算速度.

为了使特征具有旋转不变性, 在计算特征点方向时利用图像矩来计算. 由特征点邻域图像 $I(x,y)$ 定义如下图像矩:

$$m_{pq} = \sum_{x,y} x^p y^q I(x,y) \quad (6)$$

其中 $I(x,y)$ 为点 (x,y) 处的灰度值. 通过图像矩计算该邻域图像的质心:

$$\mathbf{C} = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (7)$$

在此基础上, 定义特征点的方向:

$$\theta = \text{atan2}(m_{01}, m_{10}) \quad (8)$$

其中, atan2 表示根据 x 和 y 所在象限的符号来计算反正切, 见文 [19]. 获取特征点方向后, 计算 BRIEF 描述符. 其核心思想是在特征点周围选取 n 个点, 并把这 n 个点对的比较结果组合起来作为描述符. 定义点对比较准则 $\tau^{[19]}$:

$$\tau(P;x,y) := \begin{cases} 1 : P(x) < P(y) \\ 0 : P(x) \geq P(y) \end{cases} \quad (9)$$

$P(x)$ 表示点 x 处的灰度值. BRIEF 描述符的结果是一个 n 维向量, 通过 n 个点对比较得到^[19]:

$$f_n(P;x,y) := \sum_{i=1}^n 2^{i-1} \tau(P;x,y) \quad (10)$$

为了使描述符具有旋转不变性, 根据特征点方向设置 BRIEF 的方向. 对 n 个点对 (x_i, y_i) 组成的特征集合, 定义了一个 $2 \times n$ 的矩阵:

$$\mathbf{S} = \begin{pmatrix} x_1, \dots, x_n \\ y_1, \dots, y_n \end{pmatrix} \quad (11)$$

由图像像素邻域的方向, 可以计算相应的旋转矩阵 \mathbf{R}_θ :

$$\mathbf{R}_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (12)$$

经过旋转, 得到一个新的特征集合, 定义为矩阵 \mathbf{S}_θ :

$$\mathbf{S}_\theta = \mathbf{R}_\theta \mathbf{S} \quad (13)$$

由点对矩阵 \mathbf{S}_θ 可以获取集合 T , 其中 T 中的元素为引入方向后的新点对, 最终得到如下旋转不变的 ORB 特征描述符^[19]:

$$g_n(P, \theta) := f_n(P) | (x_i, y_i) \in T \quad (14)$$

本文采用了性能最佳的高斯分布, 向量长度 n 设置为 256. 因此, ORB 特征描述符是一个 256 位的二进制串.

计算 ORB 全局特征的主要思想是把预处理图像的中心直接作为特征点的位置, 同时把图像整体作为特征点的邻域区域, 最后把计算得到的 ORB 局部描述符作为图像的全局特征描述符. 如图 4 所示, 左边是经过预处理的室内场景图像, 红色圆表示图像中心, 右边为 ORB 特征全局描述符.

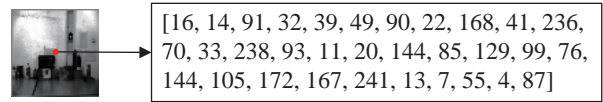


图 4 ORB 全局特征提取 (32×8 bits)

Fig.4 Extraction of ORB holistic features (32×8 bits)

2.3 基于 WH-KNN 的图像特征匹配

在计算查询图像对应的最近图像时, 本文使用 H-KNN (混合 K 最近邻)^[22] 的改进算法: WH-KNN (加权混合 K 最近邻). 图 5 描述了在 SURF 和 ORB 全局空间使用 WH-KNN 算法计算图中圆形是属于正方形类还是三角形类的过程:

(1) 在 SURF 全局空间采用 KNN 算法计算圆形的 M 个近邻 ($M=2$), 即最近的正方形 A 与次近的三角形 B;

(2) 在 ORB 全局空间采用 KNN 算法计算圆形的 N 个近邻 ($N=2$), 即最近的正方形 C 与次近的三角形 D;

(3) 然后将 K ($K=M+N=4$) 个近邻 (2 个正方形和 2 个三角形) 融合, 形成新的判决空间.

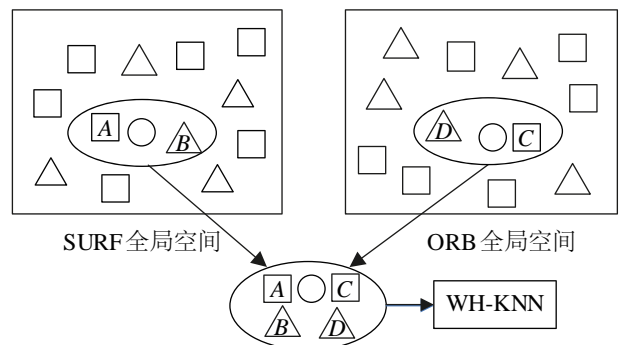


图 5 WH-KNN 示意图

Fig.5 Illustration of WH-KNN

由图 5 可知, 在新的判决空间中正方形和三角形的个数一样多, 对于 H-KNN 无法通过图形的个数判断圆形属于正方形类还是三角形类. 而在 WH-KNN 中, 距离圆形越近的图形, 其权值越小, 因此正方形的权值小于三角形. 通过 WH-KNN 算法可以判断圆形属于正方形类, 从而解决了 H-KNN 算法中的分类模糊性问题. WH-KNN 还可以用于多类别的判别中, 而不仅局限于两种类别. 在选择 M 和 N 的具体值时, 需要多次测试选择出最佳组合, 使最终判别的准确率最高.

对图像进行匹配时, 在 SURF 全局空间采用欧氏距离作为度量, 如:

$$D_S(\mathbf{L}^1, \mathbf{L}^2) = \|\mathbf{L}^1 - \mathbf{L}^2\| = \sqrt{\sum_i (L_i^1 - L_i^2)^2} \quad (15)$$

其中 \mathbf{L}^1 和 \mathbf{L}^2 是两个输入的 SURF 全局特征, 下标 S 代表 SURF 全局空间. L_i^j 表示第 j 个特征 \mathbf{L}^j 的第 i 个元素, 即 64 维特征向量中的第 i 维.

在 ORB 全局空间采用汉明距离作为度量, 汉明距离是通过对两个 (256 位) 字符串 \mathbf{H}^1 和 \mathbf{H}^2 进行位异或运算得来, 如:

$$D_R(\mathbf{H}^1, \mathbf{H}^2) = \sum_{i=1}^{256} \text{xor}(H_i^1, H_i^2) \quad (16)$$

其中 \mathbf{H}^1 和 \mathbf{H}^2 是两个输入的 ORB 全局特征, H_i^j 是 \mathbf{H}^j 的第 i 位, 下标 R 代表 ORB 全局空间. 首先提取查询图像的 SURF 与 ORB 全局特征, 然后使用 KNN 算法在 SURF 全局空间计算 M 个近邻, 同理, 在 ORB 全局空间计算 N 个近邻. 在此基础上对每个近邻赋予权值, 其中最近邻权值最小, 剩下近邻权值依次增大. 在 SURF 全局空间中, 对计算的 M 个近邻赋予权值, 如:

$$W_S(\mathbf{L}^i) = \alpha \frac{D_S(\mathbf{L}, \mathbf{L}^i)}{\sum_{i=1}^M D_S(\mathbf{L}, \mathbf{L}^i)} \quad (17)$$

$D_S(\mathbf{L}, \mathbf{L}^i)$ 表示查询图像与 SURF 全局空间第 i 张图像的欧氏距离. 在 ORB 全局空间中, 对计算的 N 个近邻赋予权值, 如:

$$W_R(\mathbf{H}^j) = \beta \frac{D_R(\mathbf{H}, \mathbf{H}^j)}{\sum_{j=1}^N D_R(\mathbf{H}, \mathbf{H}^j)} \quad (18)$$

$D_R(\mathbf{H}, \mathbf{H}^j)$ 表示查询图像与 ORB 全局空间第 j 张图像的汉明距离. 式 (17) 与 (18) 中的加权系数

$\alpha = \frac{M}{M+N}$, $\beta = \frac{N}{M+N}$, 分别表示两种不同的全局特征对应的权重.

对于从 SURF 全局空间计算得到的 M 个近邻, 在此采用 P_m ($1 \leq m \leq M$) 来表示这些近邻在地图中对应的图像标签, 这些标签构成集合 P . 同样的, 针对从 ORB 全局空间计算的 N 个近邻, 利用 Q_n ($1 \leq n \leq N$) 来表示这些近邻在地图中对应的图像标签, 这些标签构成集合 Q . 在此基础上, 定义一个新的图像标签集合 G :

$$G = P \cup Q = \{G_1, G_2, \dots, G_i, \dots\} \quad (19)$$

闭环检测的目标就是在集合 G 中找到与查询图像“最近”的一张图像, 即在集合 G 中找到权值最小的图像标签 (“最近”意味着查询图像对应的拍摄位置与地图中某一张图像对应的拍摄位置最近). 其中, 每个图像标签 G_i 对应的权值可利用式 (17) ~ (19) 计算如下:

$$W(G_i) = \prod_{P_m=G_i} W_S(\mathbf{L}^m) \times \prod_{Q_n=G_i} W_R(\mathbf{H}^n) \quad (20)$$

对于 SURF 全局空间和 ORB 全局空间计算得到的同一个近邻, 它为最近图像的概率更大, 因此它的权值为 SURF 全局空间和 ORB 全局空间分别计算得到的权值的乘积. 除同一个近邻以外, 剩下的近邻只属于 SURF 全局空间或只属于 ORB 全局空间, 它们的权值计算只由自身空间决定. 最后, 选择权值最小的近邻作为闭环检测最近的图像, 如:

$$G^* = \underset{G}{\text{argmin}} W(G_i) \quad (21)$$

WH-KNN 通过引入权值来代替 H-KNN 或 KNN 算法中的类标签的数目, 有效解决了 H-KNN 或 KNN 算法的识别模糊性问题, 提高了闭环检测的准确率.

2.4 基于跟踪模型的闭环检测

当地图的规模在不断扩大时, 采集的图像数量会急剧增加. 因此, 在如此多的地图图像中进行闭环检测, 计算量将会显著增大, 不能满足实时性要求. 同时, 闭环检测的精度也会随之下降. 为了提高闭环检测的实时性和准确性, 本文提出一种基于跟踪模型的闭环检测方法. 其核心思想是利用之前闭环检测的结果来预测查询图像在地图图像中的位置.

如图 6 所示, 三角形表示查询图像, 其中绿色三角形表示当前查询图像, 其序号为 i . 圆形表示

地图图像, 其中绿色圆形表示当前查询图像在地图中对应的最近图像, 其序号为 X_i . 我们需要对当前查询图像进行闭环检测, 即计算其在地图中所对应的最近图像, 若最近图像的序号为 X_i , 则识别正确.

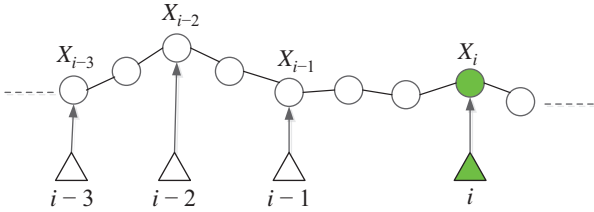


图 6 闭环检测跟踪模型示意图

Fig.6 Schematic of the tracking model for loop closure detection

在本文中查询图像序列号和地图图像序列号都是公差为 1 的等差数列. $i-1, i-2, \dots, i-n$ 为当前查询图像之前的图像序列号, 它们在地图中对应的最近图像的序列号为 $X_{i-1}, X_{i-2}, \dots, X_{i-n}$. 假设摄像机在很短的时间内做近似匀速运动, 在考虑速度干扰的情况下, 对 $X_{i-1}, X_{i-2}, \dots, X_{i-n}$ 序列中每相邻两序号做差, 也就是对其相应的地图图像序号做差, 它们的差值近似服从高斯分布, 如:

$$D_n = X_{i-n} - X_{i-n-1} \sim N(\mu, \sigma^2) \quad (22)$$

其中高斯分布的均值 μ 和方差 σ^2 可利用前面闭环检测的结果计算:

$$\mu = \frac{\sum_{i=1}^{n-1} D_i}{n-1} \quad (23)$$

$$\sigma^2 = \frac{\sum_{i=1}^{n-1} (D_i - \mu)^2}{n-1} \quad (24)$$

通过之前闭环检测的结果得到的期望值 μ 和方差 σ^2 来预测当前查询图像在地图图像中的位置. 把序列号在 $[X_{i-1} + \mu - 10\sigma, X_{i-1} + \mu + 10\sigma]$ 区间的地图图像作为当前查询图像的最近图像候选集. 当前查询图像只需与候选集中的图像进行匹配, 从而避免了与地图中每一张图像都进行匹配. 在此基础上, 采用 WH-KNN 算法计算候选集中与当前查询图像最近的 M 张图像 (SURF 全局空间) 和 N 张图像 (ORB 全局空间), 然后在这 $M+N$ 张图像形成的新判决空间中判断当前查询图像对应的最近图像. 当闭环检测结束后, 在此候选集的末尾添加当前查询图像的最近图像, 同时删除候选集的首

幅图像, 从而产生新的候选集, 待下一次闭环检测使用. 通过跟踪模型进行闭环检测不仅减少了计算量, 同时也提高了闭环检测的准确性.

3 实验结果 (Experimental results)

本文实验所用计算机为华硕 K55DR, 其配置为: CPU 为 A8-4500M 处理器, 主频 1.9 GHz, 内存 4 G. 分别使用室内和室外数据集对本文算法进行了验证.

室内实验环境为工作室和楼道, 如图 7 所示, 图 7(a) 为 $6\text{ m} \times 7\text{ m}$ 的实验室, 包含桌椅和书柜等物品, 环境比较复杂. 图 7(b) 为长度为 9 m 的楼道环境, 由于墙面特征较少而且图像间相似度高, 因此是一个闭环检测难度较大的场景. 室内数据集由普通摄像机采集得到, 图像尺寸为 640×480 像素. 实验室和楼道地图数据集为摄像机每隔约 10 cm 连续拍摄的场景图像. 其中摄像机在实验室移动距离约为 15.6 m , 共拍摄 157 张图像. 在楼道移动距离约为 8 m , 共拍摄 81 张图像. 用来作闭环检测的数据集是在地图数据集的基础上发生前后平移以及左右旋转所拍摄的场景图.



(a) 实验室场景

(b) 楼道场景

图 7 室内环境

Fig.7 Indoor environment



(a) 城市中心采集路线图

(b) 牛津大学新学院采集路线图

图 8 室外环境

Fig.8 Outdoor environment

室外数据集采用了牛津大学移动机器人研究小组公开的数据集, 即 City Center (城市中心) 数据集和 New College (新学院) 数据集, 见文 [8,23]. 航拍场景如图 8 所示, 黄点表示车载相机运动轨迹, 红点表示闭环检测相对位置. 地图数据集为车载相机每隔约 1.5 m 连续拍摄的场景图像, 包括

2474 张城市中心采集图像和 2146 张新学院采集图像. 用来作闭环检测的查询图像为机器人第 2 次回到相同地点拍摄的场景图. 在这 2 组数据集上验证了本文提出的算法, 并且与 BOW 算法进行了对比.

3.1 室内数据集测试结果

室内地图数据集部分图像如图 9 所示, 闭环检测数据集部分图像如图 10 所示.



图 9 室内地图数据集部分图像

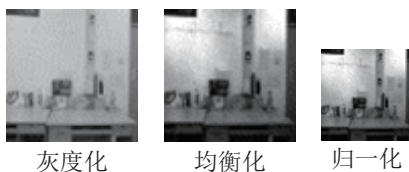
Fig.9 Partial images of indoor map datasets



图 10 室内闭环检测数据集部分图像

Fig.10 Partial images of indoor loop closure detection datasets

首先, 对输入的查询图像进行灰度化处理, 同时由于 ORB 直接使用灰度信息在太亮或太暗的条件下对效果有较大影响, 因此需要在提取 ORB 全局特征前对灰度图进行直方图均衡化. 然后把图像尺寸归一化为 63×63 像素, 这样还可以有效解决 ORB 不具有的尺度不变性问题.



灰度化 均衡化 归一化

图 11 室内预处理后的图像

Fig.11 Indoor pre-processed images

图 11 是某一查询图像灰度化、均衡化及归一化的结果. 对于预处理后的图像, 直接把图像中心作为特征点位置, 并且把整个图像作为 SURF 与 ORB 特征点的邻域区域, 进而把计算的局部描述符作为 SURF 与 ORB 的全局描述符. 在提取了查询图像的 SURF 与 ORB 全局特征后, 通过构建跟踪

模型来获取候选集. 然后使用 WH-KNN 算法对查询图像的位置进行识别.

在使用 WH-KNN 时, 分析了 SURF 与 ORB 全局空间中近邻个数的选取对闭环检测识别率和效率的影响, 如表 1 所示. 从室内闭环检测数据集中选取了连续的 40 张场景图像作为查询图像, 然后与地图数据集进行匹配来计算查询图像的最近图像.

表 1 WH-KNN 中不同近邻个数的室内闭环检测结果
Tab.1 Indoor loop closure detection results with different nearest neighbors of WH-KNN

M	N	平均时间 /ms		识别率
		特征提取	识别	
1	1	1.78	0.10	87.5%
1	2	1.94	0.18	87.5%
2	1	1.83	0.17	92.5%
1	3	1.92	0.18	90.5%
3	1	1.86	0.19	92.5%
2	2	1.98	0.18	95.0%
3	2	1.93	0.20	95.0%
2	3	2.01	0.22	92.5%
3	3	1.97	0.27	95.0%

在表 1 中, M 与 N 分别表示从 SURF 和 ORB 全局空间选取近邻的个数. 随着近邻个数逐渐增多, 闭环检测的识别率也在提高. 而当 $M > 2$ 且 $N > 2$ 时, 识别率并没有提高, 变化较小. 此外, 查询图像的全局特征平均提取时间趋于 2 ms, 识别时间却随着全局空间近邻数的增多在逐渐增大, 这是由于当近邻总数增多时, 需要耗费更多的时间去计算近邻的权值. 因此, 为了平衡计算时间和识别率, 在实验中分别设置 M 、 N 的值为 2. 然后对 $M+N$ 个近邻赋予权值作最后的判断, 判断查询图像的最近图像是否和设定的标签一致, 若一致则认为闭环检测正确, 否则错误. 另外, 对于 BOW 闭环检测方法而言, 每幅查询图像由 SURF 提取局部特征和计算图像直方图的平均时间为 246.52 ms, 识别的平均时间为 38.47 ms, 所需的总时间约为 285 ms.

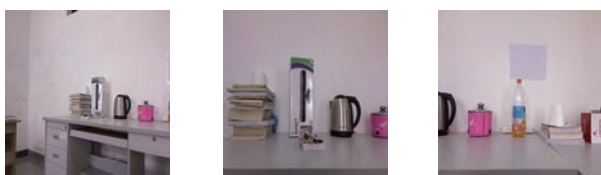
对于室内数据集, 一共做了 10 组实验, 每组实验从闭环检测数据集选取连续的 40 张图像作为查询图像, 然后与地图数据集进行匹配来计算查询图像的最近图像. 由表 2 可知, 400 幅查询图像正确判别了 377 幅, 错判了 23 幅, 平均识别率为 94.3%. 而采用 BOW 算法得到的平均识别率为 84.8%. 因此, 通过提取 SURF 与 ORB 全局特征,

采用 WH-KNN 和构建跟踪模型的方法提高了查询图像的识别准确率.

查询图像是在地图数据集的基础上随意拍摄的, 对于前后左右平移和小角度旋转的查询图像, 计算得到的最近图像全部正确. 然而对于大角度的旋转, 判别的正确率比较低. 我们分析了判决错误的图像, 如图 12, 查询图像对应的摄像机位置本应该与实际最近图像对应的相机位置最近, 但是由于摄像机发生了较大角度的旋转, 导致拍摄到的查询图像与后面的场景图像相似度更大, 因此由算法计算得到的最近图像与我们人为标记的实际最近图像不符, 从而判决错误.

表 2 室内闭环检测结果
Tab.2 Indoor loop closure detection results

编号	40 张连续图像	正确识别图像个数		识别率	
		本文方法	BOW	本文方法	BOW
1	40	38	35	95.0%	87.5%
2	40	40	36	100.0%	90.0%
3	40	33	26	82.5%	65.0%
4	40	39	38	97.5%	95.0%
5	40	37	35	92.5%	87.5%
6	40	40	37	100.0%	92.5%
7	40	39	33	97.5%	82.5%
8	40	37	32	92.5%	80.0%
9	40	36	31	90.0%	77.5%
10	40	38	36	95.0%	90.0%
统计	400	377	339	94.3%	84.8%



查询图像 计算的最近图像 实际最近图像

图 12 室内错误闭环检测结果展示

Fig.12 Illustration of indoor inaccurate loop closure detection results

3.2 室外数据集测试结果

室外数据集包括 City Center (城市中心) 数据集与 New College (新学院) 数据集. 地图数据集部分图像如图 13 所示, 闭环检测数据集部分图像如图 14 所示.

同理, 将查询图像灰度化及均衡化处理, 归一化图像尺寸为 63×63 像素, 如图 15 所示.

在提取了图像的 SURF 和 ORB 全局特征后, 通过构建高斯模型选取候选集然后使用 WH-KNN

算法获得查询图像对应的最近图像. 在使用 WH-KNN 时, 分别设置 M 、 N 的值为 2. 该测试集每幅查询图像提取 SURF 和 ORB 全局特征的平均时间为 1.95 ms, 识别的平均时间为 0.19 ms, 总时间少于 3 ms. 当采用基于 BOW 的闭环检测算法时, 每幅图像由 SURF 提取特征并计算图像直方图的平均时间为 254.33 ms, 识别的平均时间为 42.38 ms, 总时间为 296.71 ms.



图 13 室外地图数据集部分图像

Fig.13 Partial images of outdoor map datasets



图 14 室外闭环检测数据集部分图像

Fig.14 Partial images of outdoor loop closure detection datasets



灰度化 均衡化 归一化

图 15 室外预处理后的图像

Fig.15 Outdoor pre-processed images

同样, 对于室外数据集, 做了 15 组实验, 每组实验从闭环检测数据集选取连续的 40 张图像作为查询图像, 然后与第一次生成的地图数据集进行图像匹配, 计算查询图像的最近图像. 从表 3 中可以看出本文算法的平均识别率为 94.5%, 600 张查询图像正确判别了 567 幅, 判错了 33 幅. 而采用 BOW 算法的平均识别率为 83.2%, 并且局部特征提取和识别所耗费的总时间较本文算法更长. 因此, 通过牛津大学标准数据集进一步验证了本文算法精度高, 实时性强.

另外, 分析了室外图像错判的原因, 从图 16 可以看出, 由于室外场景比较复杂, 以及光照强度的影响, 许多场景具有很高的相似性. 其中查询图

像与计算的最近图像有很多相似的地方,包括城墙和草丛等,从而把计算得到的最近图像误判为实际最近图像。

表3 室外闭环检测结果
Tab.3 Outdoor loop closure detection results

编号	40张连续图像	正确识别图像个数		识别率	
		本文方法	BOW	本文方法	BOW
1	40	40	34	100.0%	85.0%
2	40	38	31	95.0%	77.5%
3	40	40	32	100.0%	80.0%
4	40	35	33	87.5%	82.5%
5	40	39	35	97.5%	87.5%
6	40	34	28	85.0%	70.0%
7	40	37	34	92.5%	85.0%
8	40	38	35	95.0%	87.5%
9	40	40	37	100.0%	92.5%
10	40	39	32	97.5%	80.0%
11	40	38	34	95.0%	85.0%
12	40	36	33	90.0%	82.5%
13	40	39	36	97.5%	90.0%
14	40	35	31	87.5%	77.5%
15	40	39	34	97.5%	85.0%
统计	600	567	499	94.5%	83.2%



图16 室外错误闭环检测结果展示

Fig.16 Illustration of outdoor inaccurate loop closure detection results

4 结论 (Conclusions)

在未知环境下,SLAM是机器人实现自主定位和地图构建的关键性问题。高效准确的定位和地图构建可以给人们生活带来极大的帮助和服务。闭环检测是机器人研究SLAM的一个基础问题,通过闭环检测可以提高地图构建的准确性,同时使机器人的定位更加精确。本文提出了一种全新的快速且准确的闭环检测算法。本算法中场景图像的特征由两部分组成,一部分为SURF全局特征,另一部分为ORB全局特征。对每幅图像提取两种全局特征后,通过构建跟踪模型来获取查询图像所属的候选集,然后采用WH-KNN算法进行图像特征匹

配,从而实现闭环检测。利用牛津大学数据集和所采集的数据集验证了提出的算法。在一台1.9 GHz的CPU和4G RAM的华硕K55DR计算机上进行全局特征提取的时间少于2 ms,场景位置识别的时间少于1 ms,室内和室外判决最近位置正确率分别为94.3%和94.5%。从特征提取时间和正确率方面可以看出,本文提出的算法可以准确快速地实现闭环检测。

参考文献 (References)

- [1] Thrun S, Leonard J J. Simultaneous localization and mapping[M]//Springer Handbook of Robotics. Berlin, Germany: Springer-Verlag, 2008: 871-889.
- [2] Xu H X, Zhou W, Zhu J. 3D visual SLAM with a time-of-flight camera[C]//IEEE International Workshop on Signal Processing Systems. Piscataway, USA: IEEE, 2015.
- [3] 付梦印, 吕宪伟, 刘彤, 等. 基于RGB-D数据的实时SLAM算法[J]. 机器人, 2015, 37(6): 683-692.
Fu M Y, Lü X W, Liu T, et al. Real-time SLAM algorithm based on RGB-D data[J]. Robot, 2015, 37(6): 683-692.
- [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] Latégahn H, Stiller C. Vision-only localization[J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 15(3): 1246-1257.
- [6] Clemente L, Davison A J, Reid I D, et al. Mapping large loops with a single hand-held camera[C]//Robotics: Science and Systems III. Cambridge, USA: MIT Press, 2008: 27-30.
- [7] Williams B, Cummins M, Neira J, et al. An image-to-map loop closing method for monocular SLAM[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, USA: IEEE, 2008: 2053-2059.
- [8] Cummins M, Newman P. FAB-MAP: Probabilistic localization and mapping in the space of appearance[J]. International Journal of Robotics Research, 2008, 27(6): 647-665.
- [9] Paul R, Newman P. FAB-MAP 3D: Topological mapping with spatial and visual appearance[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2010: 2649-2656.
- [10] Kejriwal N, Kumar S, Shibata T. High performance loop closure detection using bag of word pairs[J]. Robotics and Autonomous Systems, 2016, 77: 55-65.
- [11] 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 vocabulary tree with pyramid TF-IDF scoring match scheme for loop closure detection[J]. Acta Automatica Sinica, 2011, 37(6): 665-673.
- [12] 李永峰, 张国良, 王峰, 等. 一种基于历史模型集的改进闭环检测算法[J]. 机器人, 2015, 37(6): 663-673.
Li Y F, Zhang G L, Wang F, et al. An improved loop closure detection algorithm based on historical model set[J]. Robot, 2015, 37(6): 663-673.

- [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, 321.
Li W P, Zhang G L, Yao E L, et al. An improved loop closure detection algorithm based on the constraint from space position uncertainty[J]. Robot, 2016, 38(3): 301-310, 321.
- [14] 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.
- [15] Li F F, Fergus R, Perona P. Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories[J]. Computer Vision and Image Understanding, 2007, 106(1): 59-70.
- [16] 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.
- [17] Lowe D. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [18] Bay H, Tuytelaars T, van Gool L. SURF: Speeded up robust features[M]//Lecture Notes in Computer Science, vol.3951. Berlin, Germany: Springer, 2006: 404-417.
- [19] Rublee E, Rabaud V, Konolige K, et al. ORB: An efficient alternative to SIFT or SURF[C]//IEEE International Conference on Computer Vision. Piscataway, USA: IEEE, 2011: 2564-2571.
- [20] 龙超, 韩波, 张宇. 基于 Kinect 和视觉词典的三维 SLAM [J]. 计算机应用, 2016, 36(3): 774-778.
Long C, Han B, Zhang Y. Three-dimensional SLAM using Kinect and visual dictionary[J]. Journal of Computer Applications, 2016, 36(3): 774-778.
- [21] 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.
- [22] Hu Z Z, Hu Y Z, Li Y C. Fast road sign recognition based on ORB-encoded holistic and local features[C]//95th Annual Meeting of the Transportation Research Board on Artificial Intelligence and Advanced Computing Applications. 2016.
- [23] Oxford Mobile Robotics Group. New college and center city dataset [DB/OL]. [2016-06-28]. http://www.robots.ox.ac.uk/~mobile/IJRR_2008_Dataset/data.html.

作者简介:

刘国忠 (1990 -), 男, 硕士生. 研究领域: 移动机器人 SLAM, 模式识别, 计算机视觉.

胡钊政 (1979 -), 男, 博士, 教授. 研究领域: 3D 计算机视觉理论与应用, 视觉与激光 SLAM 定位, 智能交通系统等.