Bags of Binary Words for Fast Place Recognition in Image Sequences

Huang Yao

2015 - 9 - 7

内容

1. 简介
2. 二进制特征
3. 图片数据库的构建
4. 闭环检测
5 .实验结果以及结论

Why?

when areas non-observed for long are re-observed, standard matching algorithms fail. When they are robustly detected, loop closures provide correct data association to obtain consistent maps.

The same methods used for loop detection can be used for robot relocation after track lost, due for example to sudden motions, severe occlusions or motion blur.

1. 位置识别、闭环检测可以纠正长时间SLAM积累的误差
2. 当机器人track lost，可以relocation

the state-of-the-art FAB- MAP 2.0 algorithm

Our approach is based on bag of words and geometrical check, with several important novelties that make it much faster than current approaches.

1. 方法最大改进之处在于使用FAST特征点和BRIEF描述子。
2. 使用二进制的词袋（The fist time a binary vocabulary for loop ）
3. 采用一定的策略阻止短时间大量图片数据库查询的冲突
4. 加入geometrical check

优点：

1. 计算速度快，适合实时的项目，而且占用内存少
2. 鲁棒性好

缺点：

1. BRIEF无法对scale,rotation保持不变性

二进制特征

FAST 关键点 + BRIEF 描述子

1. The patches are previously smoothed with a Gaussian kernel to reduce noise
2. 17.3μs per keypoint when Lb = 256 bits) and to compare.
3. 选择Hamming distance快速计算二进制串的距离

图片数据库IMAGE DATABASE

hierarchical bag of words

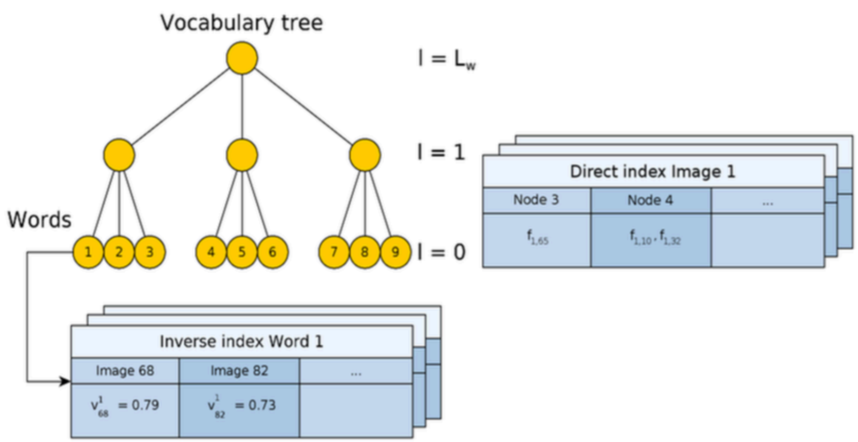

Fig. 1. Example of vocabulary tree and direct and inverse indexes that compose the image database. The vocabulary words are the leaf nodes of the tree. The inverse index stores the weight of the words in the images in which they appear. The direct index stores the features of the images and their associated nodes at a certain level of the vocabulary tree.

1. 视觉词袋创建offline
2. 通过k-means聚类形成层次结构

1. 将图片转换为向量v
2. 从root遍历到树叶
3. 每一层最小化Hamming距离

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

inverse index

1. 对于每个word，存储其所有出现过的图片索引，以及在图片中得权重
2. 当数据库插入新图片时被更新；
3. 主要用来加速数据库的查询

store pairs < It, vti >

direct index

store the features of each image

1. 每张图片需要建立一个direct Index
2. store in the direct index the nodes at level l that are ancestors of the words present in It, the list of local features ftj associated to each node
3. 用来加速geometrical verification
4. 新图片加入时被更新

LOOP DETECTION ALGORITHM

Database query

1. 获取图片之后将其转换为向量v之后在数据库中查询，会得到一系列匹配的结果
2. 通过normalized similarity score 对每一个candidate评估
3. 设定最小的阈值，reject一些匹配的结果

$$\eta(\mathbf{v}_t, \mathbf{v}_{t_j}) = \frac{s(\mathbf{v}_t, \mathbf{v}_{t_j})}{s(\mathbf{v}_t, \mathbf{v}_{t-\Delta t})}$$

Match grouping

1. 图片采集速度快，在数据库查询中发生竞争
2. 把较小时间区间tni , . . . , tmi 内的匹配< vt, vtni >, . . . , < vt, vtmi >看成同一个< vt, VTi >，作为一个island
3. The island with the highest score is selected as matching group and continue to the temporal consistency step.

Temporal consistency

The match < vt, VT′ > must be consistent with k previous matches < vt−Δt, VT1 >, . . . , < vt−kΔt, VTk >, such that the intervals Tj and Tj+1 are close to overlap. If an island passes the temporal constraint, we keep only the match < vt, vt′ >, for the t′ ∈ T′ that maximizes the score η, and consider it a loop closing candidate, which has to be finally accepted by the geometrical verification stage.

几何一致性检验 geometrical consistency

This check consists in finding with RANSAC a fundamental matrix between It and It’ supported by at least 12 correspondences.

暴力搜索 K-D树

To obtain correspondences between It and It’ , we look up It’ in the direct index and perform the comparison only between those features that are associated to the same nodes at level l in the vocabulary tree.

This condition speeds up the correspondence computation.

实验结果分析

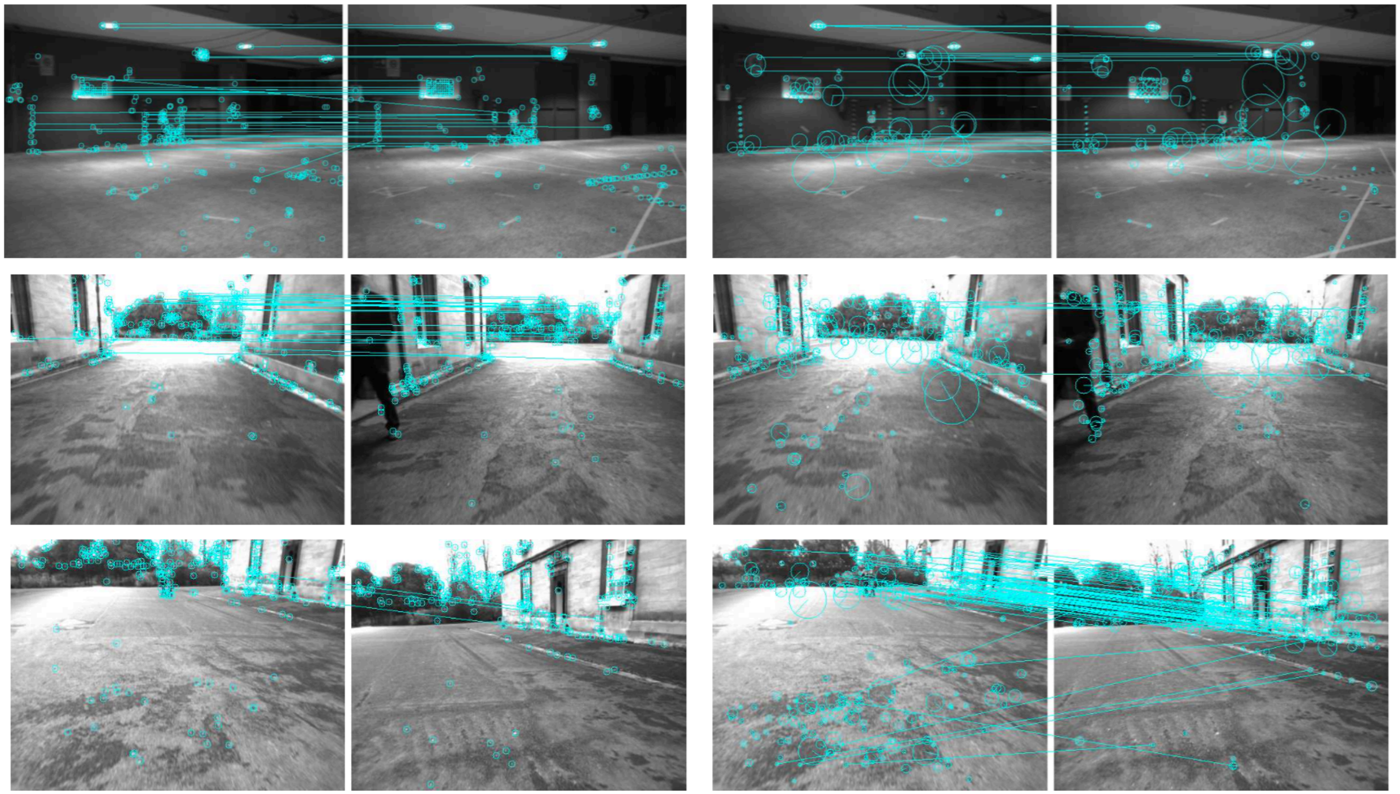

Fig. 3. Examples of words matched by using BRIEF (pair on the left) and SURF64 descriptors (pair on the right).

1. 背景特征丰富时，BRIEF和SURF方法都比较可靠；
2. 第三幅图，存在尺度或旋转变化时，BRIEF效果不如SURF

13ms per image instead of 100–400ms)

32MB instead of 256MB to store a 1M word vocabulary

		Execution time (ms / query)			
		Mean	Std	Min	Max
Features	FAST	11.67	4.15	1.74	30.16
	Smoothing	0.96	0.37	0.79	2.51
	BRIEF	1.72	0.49	1.51	4.62
Bag of words	Conversion	3.59	0.35	3.27	8.81
	Query	3.08	1.91	0.01	9.19
	Islands	0.12	0.04	0.08	0.97
	Insertion	0.11	0.02	0.06	0.28
Verification	Correspondences and RANSAC	1.60	2.64	0.61	18.55
Whole system		21.60	4.82	8.22	51.68

结论

1. binary fea- tures are very effective and extremely efficient in the bag-of-words approach.

2. The main limitation of our technique is the use of features that lack rotation and scale invariance. It is enough for place recognition in indoor and urban robots, but surely not for all-terrain or aerial vehicles, humanoid robots, wearable cameras, or object recognition

参考资源

1. FABMAP

http://www.robots.ox.ac.uk/~mobile/wikisite/pmwiki/pmwiki.php?n=Software.FABMAP

2. DBoW2: Enhanced hierarchical bag-of-word library for C++

http://webdiis.unizar.es/~dorian/index.php?p=3#GalvezIROS11

3. Gálvez-López D, Tardós J D. Bags of binary words for fast place recognition in image sequences[J]. Robotics, IEEE Transactions on, 2012, 28(5): 1188-1197.

4. Nister D, Stewenius H. Scalable recognition with a vocabulary tree[C]//Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. IEEE, 2006, 2: 2161-2168.