

Sparse Optimization Based LoopClosing

分享人：艾攀 何叶

- I. 论文简介
- II. 已有方法
- III. 本文方法
- IV. 实验及结论

I. 论文简介

Name: Sparse Optimization for Robust and Efficient Loop Closing

Source: Robotics and Autonomous Systems 2017 (SCI 水刊)

Key words: loop-closure, sparse optimization, robust

Key idea: loop-closure detection 问题建模为sparse convex
L1-minimization optimization 问题

Highlight points: Global unique Any representation Robust and
efficient

II. 已有方法

- Appearance-based/ FAB-MAP[13] —— visual BoW text-retrieval, offline learn
- BBoW-based —— Fast keypoint and BRIEF descriptors text-retrival, offline learn

针对极端环境（天气，光线）的方法：

SeqSLAM[33] —— 匹配序列图像

Experience-based maps[12] ——学习同一个场景各种情形下的特征（比较笨的）

深度学习：

CNN-based[44, 45] ——用CNN去学习图像特征（不需手动设计或选择），来构建词典

总结：都需要根据特定的特征离线构建词典，然后查询匹配词典计算当前图像得分

III. 本文方法

关键思路：在线构建一个冗余的词典，以便找到一个稀疏的解来表示检测到了回环这件事

推导：

假设当前图像用一个 n 维向量 \mathbf{b} 表示，表示是任意的， \mathbf{b} 可以是像素值，也可以是描述子，甚至是CNN提取出的特征，通过在线不断的构建一个词典 $\mathbf{B}=[\mathbf{b}_1 \mathbf{b}_2 \cdots \mathbf{b}_m]$ 得到一个过去场景的总的表示，那么通过求解

$$\mathbf{B}\mathbf{x} = \mathbf{b}$$

得到的解 \mathbf{x} 就可以看做是当前图像和过去 m 帧图像的一个关联程度，用数学语言讲就是 \mathbf{b} 可以由基 \mathbf{B} 表示，当 \mathbf{x} 稀疏时，就说明当前图像只跟 \mathbf{x} 中非0部分对应的基 \mathbf{b}_i 相关性大，也即检测到回环

问题1：并不是所有 \mathbf{b} 可以由基 \mathbf{B} 描述，怎么办？

理论：当前仅当 \mathbf{b} 在 \mathbf{B} 的range space（列向量张成的线性空间），才能表示，当 \mathbf{B} 秩亏时，可能不能表示

问题2：怎么确保解是稀疏的？

III. 本文方法

实际情况：我们可以通过确保构建一个冗余的满秩的B来解决，而且在实际求解过程中，由于信号就是稀疏的，得到的解一般也是很稀疏的
建模：

$$\min_x \|x\|_0 \quad \text{subject to} \quad \mathbf{B}x = b$$

Combinatorial optimization --- NP-hard（枚举所有满足约束的解），放松条件求解最小L1范数（L1范数是最接近L0范数的凸优化解，且也满足稀疏特性）

$$\min_x \|x\|_1 \quad \text{subject to} \quad \mathbf{B}x = b$$

考虑图像噪声（如动态物，模糊等），引入误差向量 e

$$\begin{aligned} & \min_{x,e} \|x\|_1 + \|e\|_1 \quad \text{subject to} \quad \mathbf{B}x + e = b \\ \Rightarrow & \min_{\alpha} \|\alpha\|_1 \quad \text{subject to} \quad \mathbf{D}\alpha = b \quad : \mathbf{D} := \begin{bmatrix} \mathbf{I}_n & \mathbf{B} \end{bmatrix} \text{ and } \alpha := \begin{bmatrix} e \\ x \end{bmatrix}. \end{aligned}$$

III. 本文方法

求解上式，引入噪声等级参数ksai，进一步放松约束条件

$$\min_{\alpha} \|\alpha\|_1 \quad \text{subject to} \quad \|\mathbf{D}\alpha - \mathbf{b}\|_2 \leq \epsilon$$

转化为无约束的优化问题，lamda为权重因子

$$\min_{\alpha} \lambda \|\alpha\|_1 + \frac{1}{2} \|\mathbf{D}\alpha - \mathbf{b}\|_2^2$$

求解方法：primal-dual method 或 **homotopy method**

III. 本文方法

算法步骤:

Algorithm 1 Closing Loops via ℓ_1 -Minimization

Input: Dictionary \mathbf{D}_{i-1} , Current image \mathbf{i}_i , Threshold τ , Weight λ , Ignoring-time window t_g

Output: Loop-closure hypotheses \mathbf{H} , Updated dictionary \mathbf{D}_i

1: $\mathbf{b}_i := f(\mathbf{i}_i)$

2: *Hypothesis generation:*

3: Solve $\min_{\boldsymbol{\alpha}_i} \lambda \|\boldsymbol{\alpha}_i\|_1 + \frac{1}{2} \|\mathbf{D}_{i-1} \boldsymbol{\alpha}_i - \mathbf{b}_i\|_2$ using the homotopy approach (see Section 4.2)

4: Normalize $\hat{\boldsymbol{\alpha}}_i := \frac{\boldsymbol{\alpha}_i}{\|\boldsymbol{\alpha}_i\|_2}$

5: Find hypotheses $\mathbf{H} := \{j \mid \hat{\alpha}_{i,j} > \tau, \|i - j\|_1 > t_g\}$

6: *Dictionary update:*

7: $\mathbf{D}_i := \begin{bmatrix} \mathbf{D}_{i-1} & \mathbf{b}_i \end{bmatrix}$

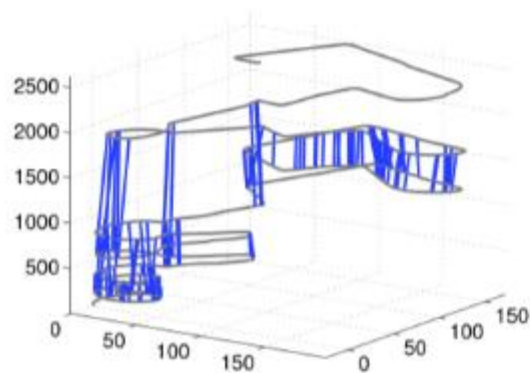
优点: global unique—只有一个解j, any representation——bi不限形式
Robust——动态物体, 运动模糊

IV. 实验及结论

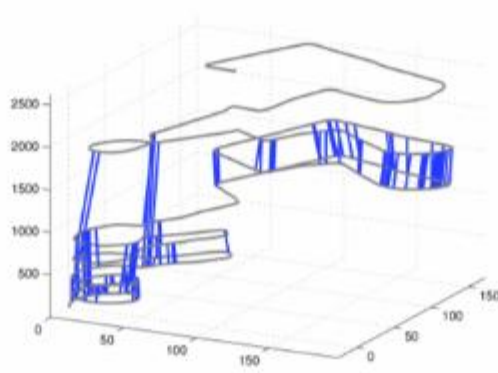
硬件：Laptop i5 2.5GHz, 16G RAM

软件：Matlab

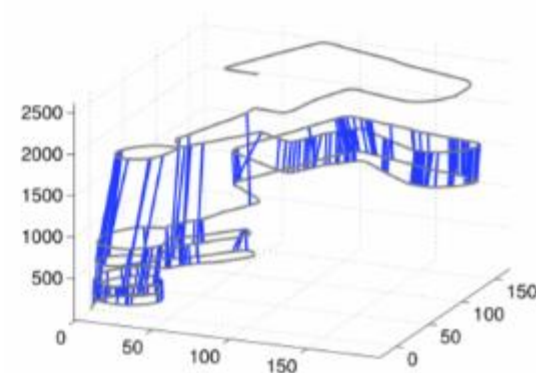
1.不同类型的Basis——数据集 New College dataset[41] 1Hz 共2624张



(a) Basis used as raw 64×48 images



(b) Basis used as raw 8×6 images

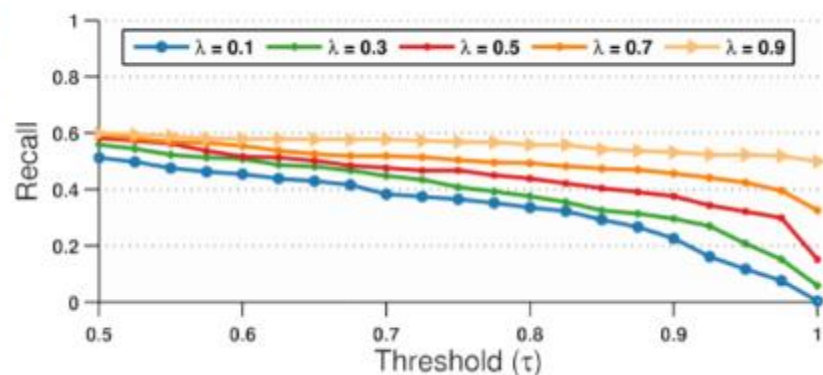
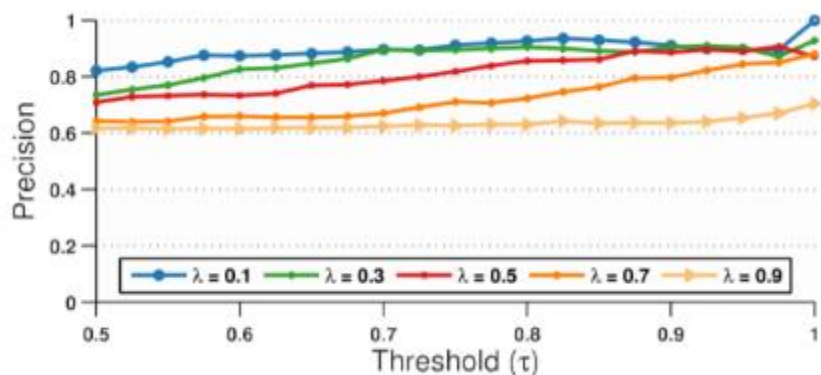


(c) Basis used as GIST descriptors

基本都能正确检测回环

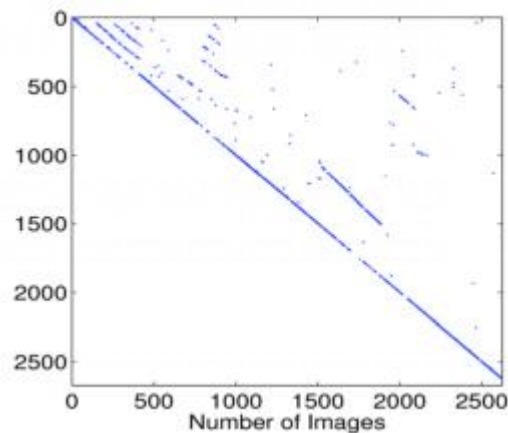
IV. 实验及结论

2.调参——数据集Bicocca 25b dataset 5Hz 共8358张
阈值 τ 和权重 λ , 评估指标: 准确率和召回率



求解出的 α 的各分量超过阈值 τ ,
说明对应图像有贡献, 可能存在关联, 阈值越大, 限制越严格,
对应准确率就越高, 相应的召回率越低, 权重则相反

右图横坐标为当前图像的序号, 在y上有对应的
稀疏点, 表明和对应历史图有关联, 对角线表明
当前图像和自己最相似



IV. 实验及结论

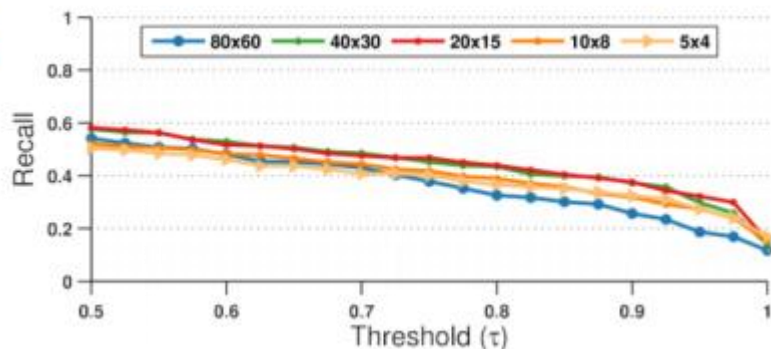
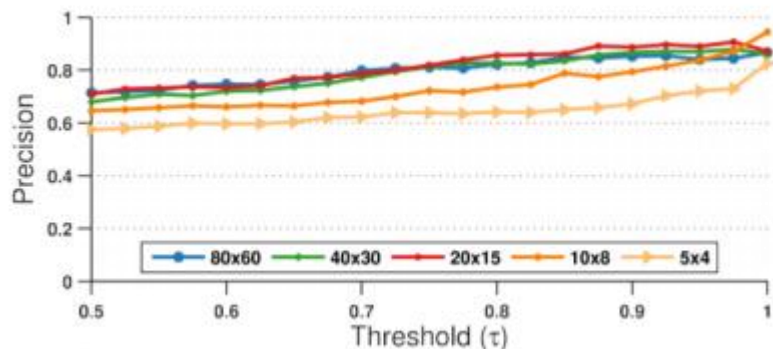
2.调参——数据集Bicocca 25b dataset 5Hz 共8358张

Image Size影响

a.运行时间——可满足实时检测

b.算法性能

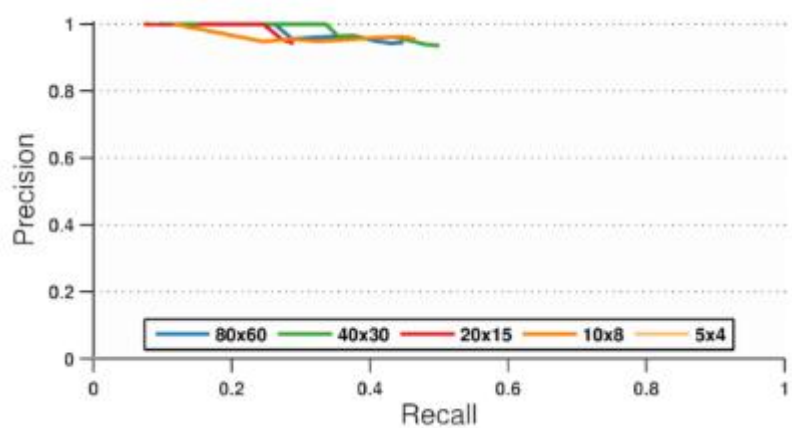
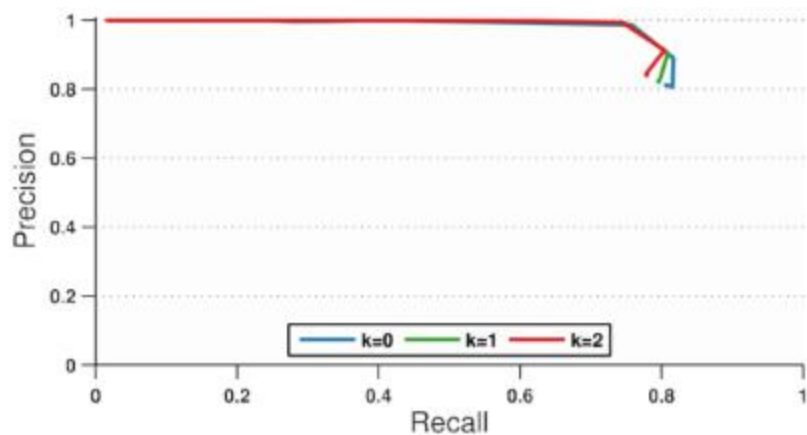
size (feature dimension)	min (ms)	mean (ms)	max (ms)	std (ms)
80×60 (4800)	45.762	116.45	417.290	43.229
40×30 (1200)	2.689	17.334	70.940	9.500
20×15 (300)	0.111	3.662	20.828	2.904
10×8 (80)	0.029	0.707	4.448	0.552
5×4 (20)	0.021	0.410	2.541	0.3163



前三种分辨率的准确率召回率曲线几乎相同，当分辨率降到10*8时才有显著的变化

IV. 实验及结论

3.和DBoW算法对比



和DBoW相比效果还是差了很多，准确率-召回率曲线， DBoW明显占优（左图）

IV. 实验及结论

4.其他实验（感觉不是和很重要）

CNN提取的特征

多种特征混合（论文中只是简单的并列起来）

多次拜访问题

剧烈环境变化（比如冬季和夏季）

这里就不贴实验结果了，按论文的说法，结果都很好，多种特征混合可以提升算法性能——准确率和召回率都有一定程度的提升

5.稀疏性和参数lamda的关系

纵轴是求解出的alpha中non-zero的比例，也就是稀疏性，可以看出是反比关系，因为这个权重代表的L1范数的部分的影响

