

同时定位与地图构建(SLAM)概述 II

Paul Robertson

回顾

- 定位
 - 跟踪，全局定位，绑架问题
- 卡尔曼滤波
 - 二次型
 - 跟踪，全局定位，绑架问题
- 同时定位与地图构建
 - 闭环回路
 - 尺度：
 - 分割空间至重叠区域，采用重路由算法
- 未提及的问题
 - 特征
 - 探索

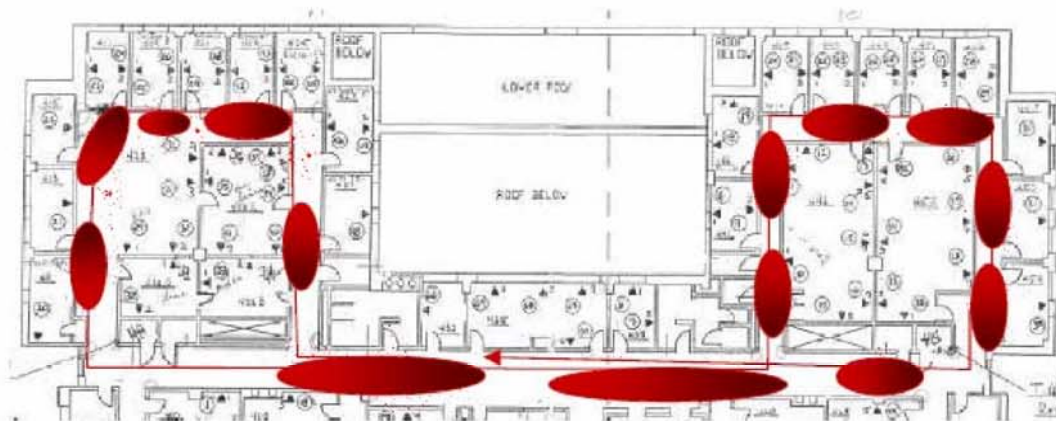
提 纲

- 拓扑地图
- 隐马尔可夫模型
- 尺度不变特征变换
- 基于视觉的定位

拓扑地图

思想：

构建一个量化地图，该地图中的节点代表具有相似性质的传感器数据标识，节点之间的变换表示这期间的机器人控制行为



拓扑的优越性

- 可以解决全局定位问题
- 可以解决绑架问题
- 类似于人所构建地图
- 支持度量性的定位
- 可以表述为隐马尔可夫模型(HMM)

1

隐马尔可夫模型(HMM)

已知条件

- 作业域可以表示为一系列的变量
- 状态定义了任意给定状态之后什么后继状态可达
- 状态的改变包括行为
- 行为可观，状态不可观
- 你想尽量搞清楚一系列机器人行为

f

例子

部分语音的标识，自然语言解析，语音识别，场景分析，定位/路径估计

隐马尔可夫模型概述

什么是隐马尔可夫模型

找到最可能状态序列的算法

一个动作序列的概率似然算法
(计算所有可能的状态路径)

训练马尔可夫模型的算法

只能工作在状态结构能够被定义为FSM的问题，
此时某一时刻单独行为能够作为状态之间的变换
本身是一种非常流行的算法，原因在于该算法与
行为序列长度成线性关系。

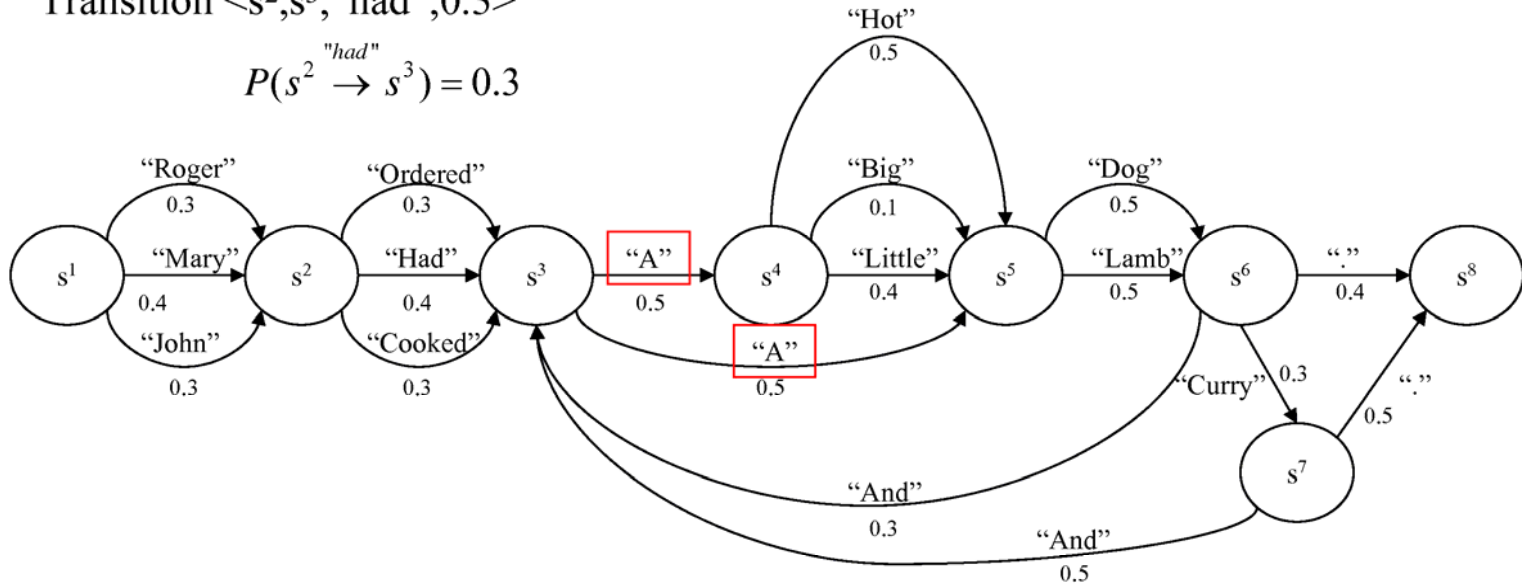
隐马尔可夫模型

用概率表示节点链接弧的有限状态机

$\langle s^1, S, W, E \rangle$, 其中 $S = \{s^1, s^2, s^3, s^4, s^5, s^6, s^7, s^8\}$; $W = \{\text{"Roger"}, \dots\}$; $E = \{\langle \text{transition} \rangle \dots\}$

Transition $\langle s^2, s^3, \text{"had"}, 0.3 \rangle$

$$P(s^2 \xrightarrow{\text{"had"}} s^3) = 0.3$$



S_1 : 玛丽有一个小台灯和一只大狗

S_2 : 罗杰点了一份羔羊咖喱和一个热狗

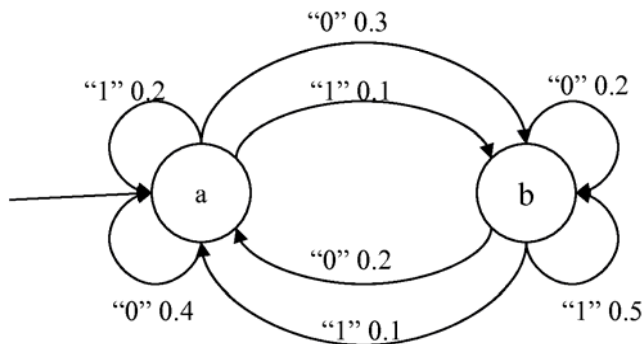
S_3 : 约翰做了一份热狗咖喱

$$P(S_3) = 0.3 * 0.3 * 0.5 * 0.5 * 0.3 * 0.5 = 0.003375$$

寻找极大似然路径

Viterbi 算法：对于一组t-1长度的动作序列，在线性时间内找到：

$$\sigma(t) = \arg \max_{s_{1,t}} P(s_{1,t} | w_{1,t-1})$$



“1110”

Viterbi 算法：

对于每一个状态，找到终止于该状态的最有可能的状态序列

| States | | ϵ | 1 | 11 | 111 | 1110 |
|--------|-------------|------------|-----|------|-------|-------|
| a | Sequence | a | aa | aaa | aaaa | abbba |
| | Probability | 1.0 | 0.2 | 0.04 | 0.008 | 0.005 |
| b | Sequence | b | ab | abb | abbb | abbbb |
| | Probability | 0.0 | 0.1 | 0.05 | 0.025 | 0.005 |

行为序列的概率

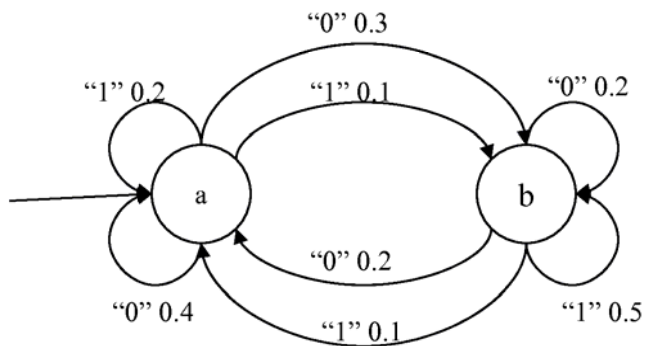
$$P(w_{1,n}) = \sum_{i=1}^{\sigma} P(w_{1,n}, S_{n+1} = s^i)$$

令 $\alpha_i(t)$ 为概率 $P(w_{1,t-1}, S_t = s^i)$, 因此 $P(w_{1,n}) = \sum_{i=1}^{\sigma} \alpha_i(n+1)$

$$\alpha_i(1) = \begin{cases} i=1 \rightarrow 1.0 \\ \text{否则} \rightarrow 0 \end{cases} \quad (\text{必须从初始状态开始}).$$

$$\alpha_j(t+1) = \sum_{i=1}^{\sigma} \alpha_i(t) P(s^i \xrightarrow{w_t} s^j)$$

隐马尔可夫的前向概率



"1110"

| t | 1 | 2 | 3 | 4 | 5 |
|---------------|-----|-----|------|-------|--------|
| ϵ | 1 | 1 | 1 | 1 | 0 |
| $\alpha_a(t)$ | 1.0 | 0.2 | 0.05 | 0.017 | 0.0148 |
| $\alpha_b(t)$ | 0.0 | 0.1 | 0.07 | 0.04 | 0.0131 |
| $P(w_{1,t})$ | 1.0 | 0.3 | 0.12 | 0.057 | 0.0279 |

$$0.2 * 0.1 = 0.02$$

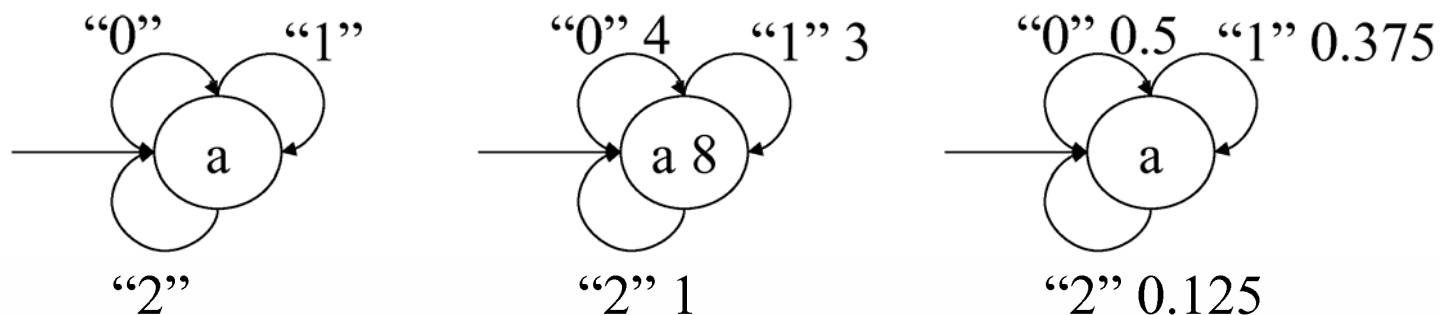
+

$$0.1 * 0.5 = 0.05$$

隐马尔可夫模型的训练

(Baum-Welch 算法)

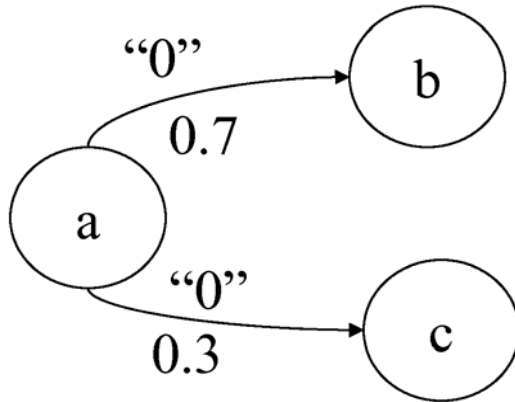
给定训练序列，调整HMM的状态转移概率，使得这组行为序列的可能性最大



训练序列：01010210

隐状态

直觉上.....



当计算状态转移情况时，
利用他们的概率按比例分
配状态转移情况

？问题在于我们并不知道
转移的概率是多少！

1. 猜想一组状态转移概率

2. (while (improving)

(propagate-training-sequences))

通过比较每次迭代的交叉熵来计算这种“提升”。当交叉熵在某一次迭代中降低至小于某一先验阈值参数时，计算结束。

交叉熵定义为：

$$-\frac{1}{n} \sum_{w_{1,n}} P_{M-1}(w_{1,n}) \log_2 P_M(w_{1,n})$$

尺度不变特征变换

David Lowe ‘Distinctive Image Features from Scale-Invariant Keypoints’ IJCV 2004.

内容:

- 尺度空间(Witkin ‘83)极值提取
- 关键点删剪及定位
- 方向指派
- 关键点描述符

SIFT中的尺度空间

目的：

- 物体可以在许多级的细节中被识别出来
- 长距离对应于低的探测限制
- 不同类型的信息在每一个级别中都可能获得

思想：在一副图像中每一探测级别提取信息内容。细节的减少通过高斯模糊来实现：

- $I(x, y)$ 为输入图像. $L(x, y, \sigma)$ 为该图像在 σ 尺度下的形式
- $G(x, y, \sigma)$ 为2D高斯函数，其方差为 σ^2
- $L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$
- $D(x, y, \sigma) = L(x, y, k \sigma) - L(x, y, \sigma)$

尺度不变特征

不变性质：

尺度

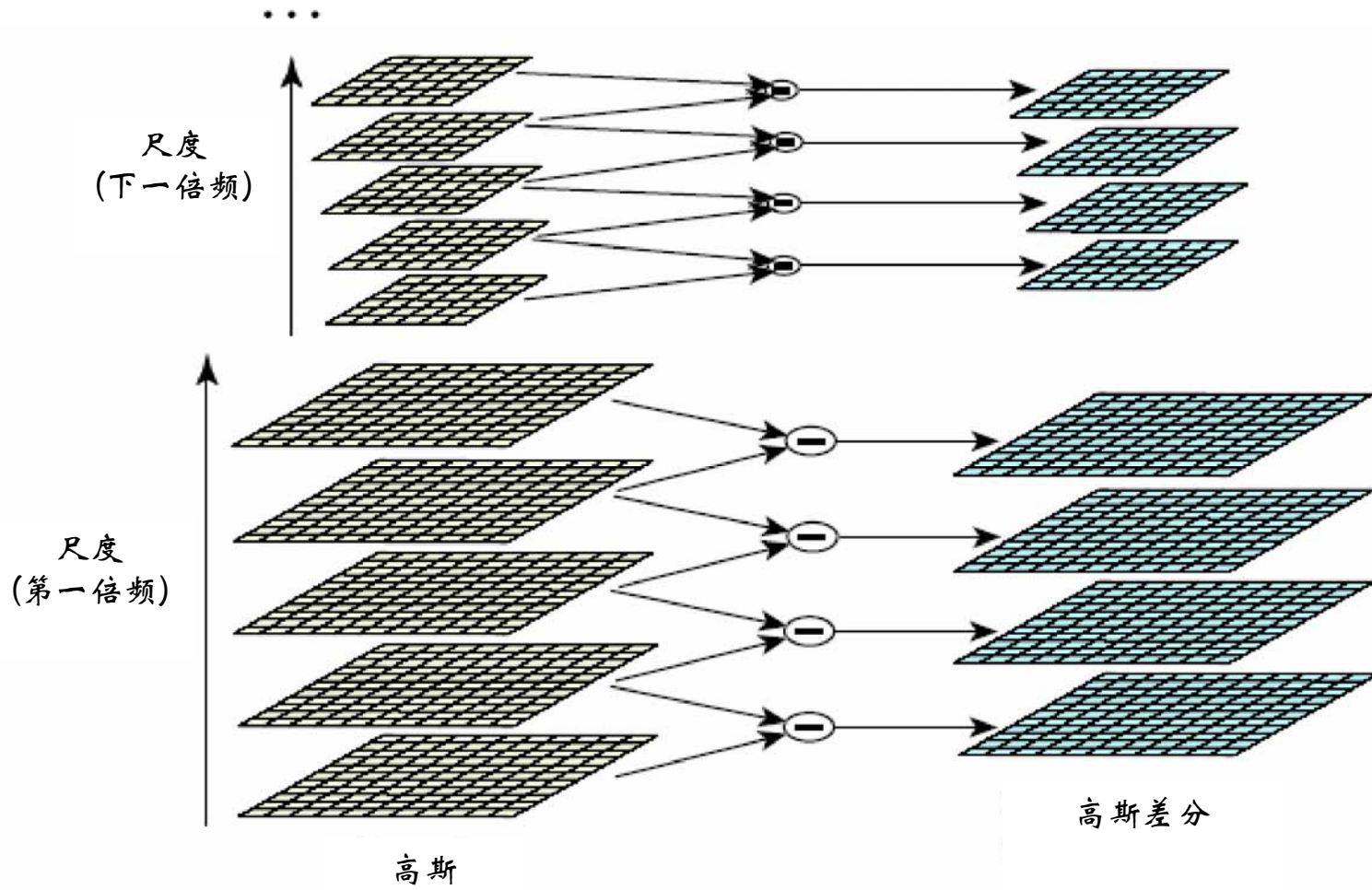
平面旋转

对比度

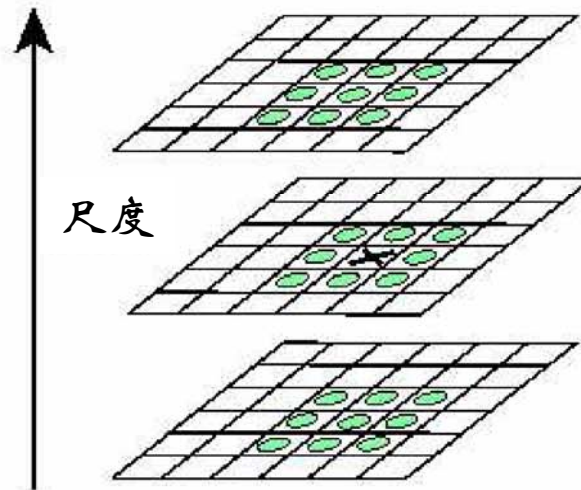
光照

大量特征

高斯差分空间



尺度空间

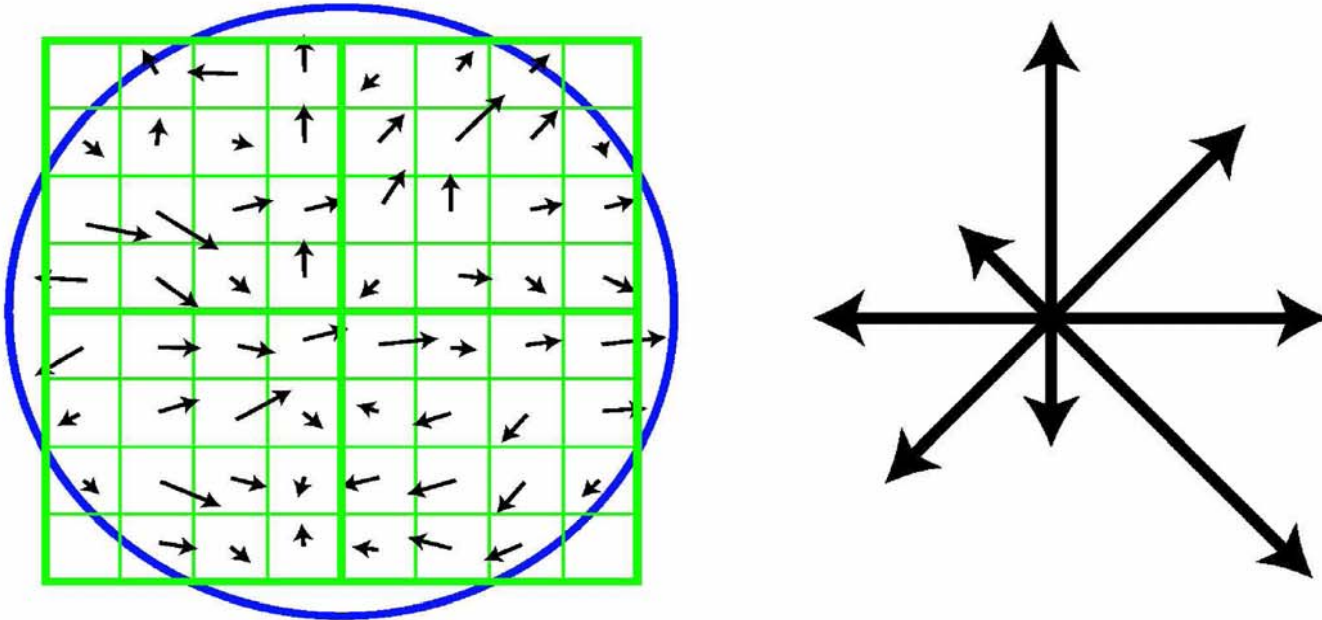


- 计算差分空间中的局部极值
- 每个 (x, y, σ) 为一个特征
- (x, y) 尺度和平面旋转的不变性

针对稳定性的剪裁

- 去除候选特征
 - 低对比度
 - 不稳定的边界响应

方向指派



对于每一个特征 (x, y, σ) :

- 在 $L(x, y, \sigma)$ 中找到围绕 (x, y) 的某一固定像素区域
- 计算梯度直方图；并命名为 b_i
- 对于极大值80%内的 b_i ，定义特征 (x, y, σ, b_i)

基于视觉的SLAM

参考读物：

Se, S., D. Lowe and J. Little, ‘Mobile Robot Localization and Mapping with Uncertainty using Scale-Invariant Visual Landmarks’, The International Journal of Robotics Research, Volume 21 Issue 08.

Kosecka, J. Zhou, L. Barber, P. Duric, Z. ‘Qualitative Image Based Localization in Indoor Environments’ CVPR 2003.

可预测的视觉自定位

1. 在当前位置计算SIFT特征
2. 利用立体视觉来定位3D特征
3. 运动
4. 基于里程计和卡尔曼滤波来预测新的位置
5. 基于机器人的运动来预测SIFT特征的位置
6. 找到SIFT特征，并且找到每一个特征的3D位置
7. 根据每一个匹配的特征来计算机器人当前的位置估计

基于视觉的定位

- 在机器人探索新环境时获得视频序列
- 以位置和空间中他们的联系来构建环境的模型
- 拓扑定位-通过位置的识别实现
- 度量性定位-通过计算当前场景和最可能位置的表述的相对位姿

同一个位置？



全局拓扑，局部几何

问题：

1. 独立位置的表述
2. 学习代表性位置的特征
3. 学习这些位置之间的关系
4. 每一个观测场景由一系列SIFT特征表述
5. 这些位置对应于子序列，使得特征能够成功匹配
6. 位置上的空间关系是通过位置图来捕捉的

图像匹配

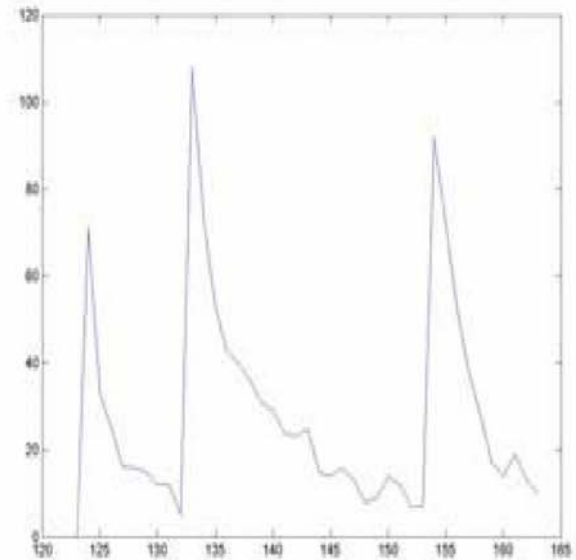


每幅图像包含
10-150个特征

- 每一个特征找到具有差别性的最近邻特征
- 图像距离对应于成功匹配特征数的比例

分割视频序列

- 状态的过渡取决于探索的过程
- 位置子序列中的特征能够被成功的匹配
- 位置表述：
具有代表意义的图像集合及其关联特征





图像匹配

- 给定一副图像，它究竟是从哪一位置获得的呢？

认知—投票机制

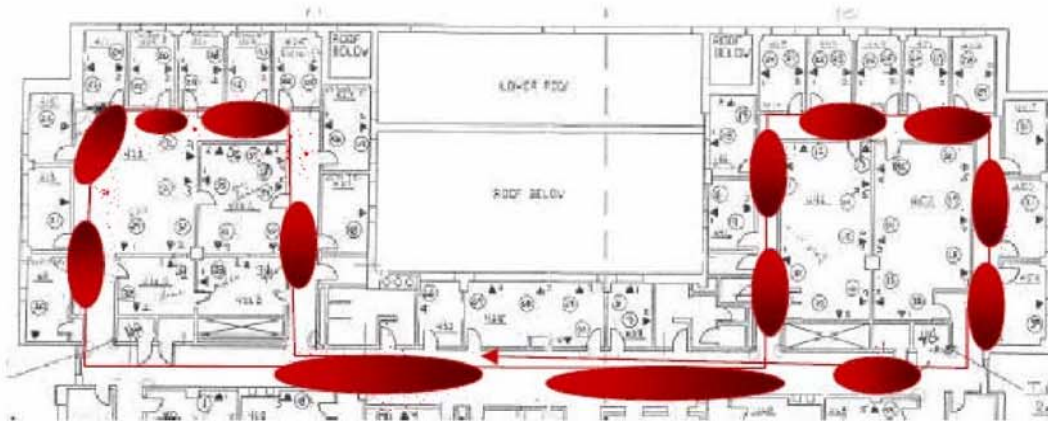
对于每一个在探索过程中获得的代表性图像

1. 计算匹配的特征数目
2. 具有最多匹配数目的位置被定义为最可能的位置

拓扑模型中的马尔可夫定位

探索未知之间的空间关系

- S – 离散状态集 $L \times \{N, W, S, E\}$ 位置和方向
- A – 离散行为集 (N, W, S, E)
- $T(S, A, S')$ – 转移方程, 离散马尔可夫模型



马尔可夫定位

$$P(L_t=l_i|o_{1:t}) \propto P(o_t|L_t=l_i) P(L_t=l_i|o_{1:t-1})$$

位置的后验概率

观测概率

$P(\text{location}|\text{observations})$

$P(\text{image}|\text{location})$

观测概率

$$P(o_t|L_t=l_i) = \frac{C(i)}{\sum_j C(j)}$$

$P(\text{image}|\text{location})$

$$P(L_t=l_i|o_{1:t-1}) = \sum A(i,j)P(L_{t-1}=l_j|o_{1:t-1})$$

位置变换概率矩阵

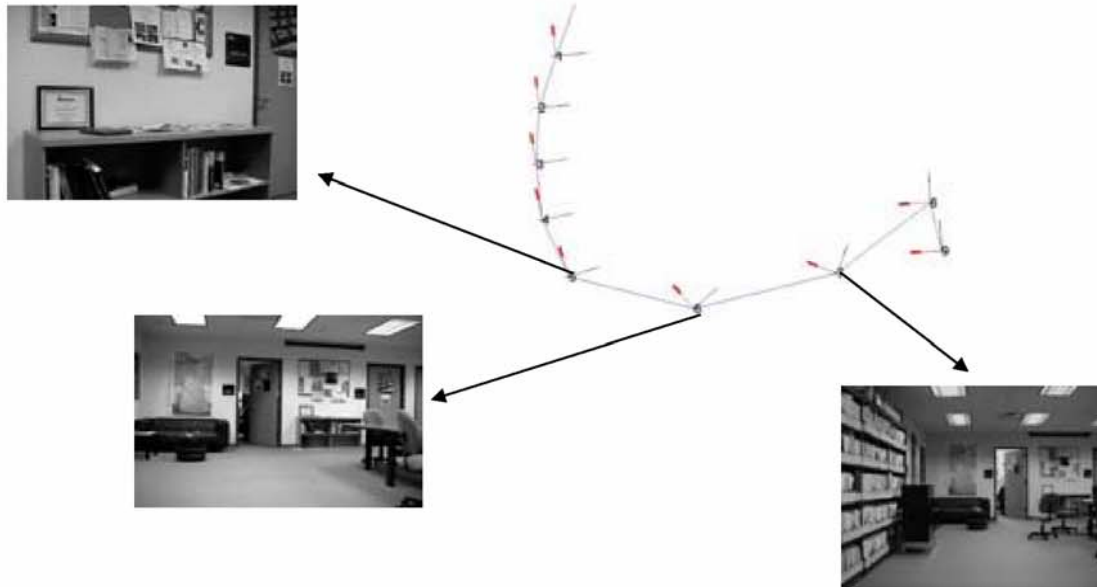
隐马尔可夫识别

- 在实验测试中的识别效率提高到82%至96%

位置内的度量性定位

1. 给定位置对应的最近视图
2. 建立关键点直接按的对应关系
3. 概率匹配，结合几何，关键点描述符和内在的尺度

计算关于参考视图的相对位置



总 结

- 我们介绍了什么？
 - 支撑方法
 - 卡尔曼滤波
 - 隐马尔可夫模型
 - 尺度不变特征
 - 定位与地图构建
 - 基本SLAM
 - 大尺度SLAM (Leonard)
 - 拓扑地图
 - 基于视觉的定位及SLAM