

基于时空关联图模型的 视频监控目标跟踪

张诚 马华东 傅慧源

{zhangcheng, mhd, fhy}@bupt.edu.cn

智能通信软件与多媒体北京市重点实验室



提纲

- 研究背景
- 相关工作
- 算法
 - 时空关联图模型的建立
 - 数据关联与相似性度量
- 实验
- 结论及贡献



提纲

- 研究背景
- 相关工作
- 算法
 - 时空关联图模型的建立
 - 数据关联与相似性度量
- 实验
- 结论及贡献

多摄像机监控下的目标跟踪

- 在**非重叠视域**的摄像机网络中对特定目标进行持续跟踪
- 应用普遍：视频监控、目标轨迹预测、物理空间检索等
- 三个子问题：
 - 单摄像头目标跟踪（无起始帧的初始化）
 - 目标识别与匹配
 - 跨摄像头目标跟踪

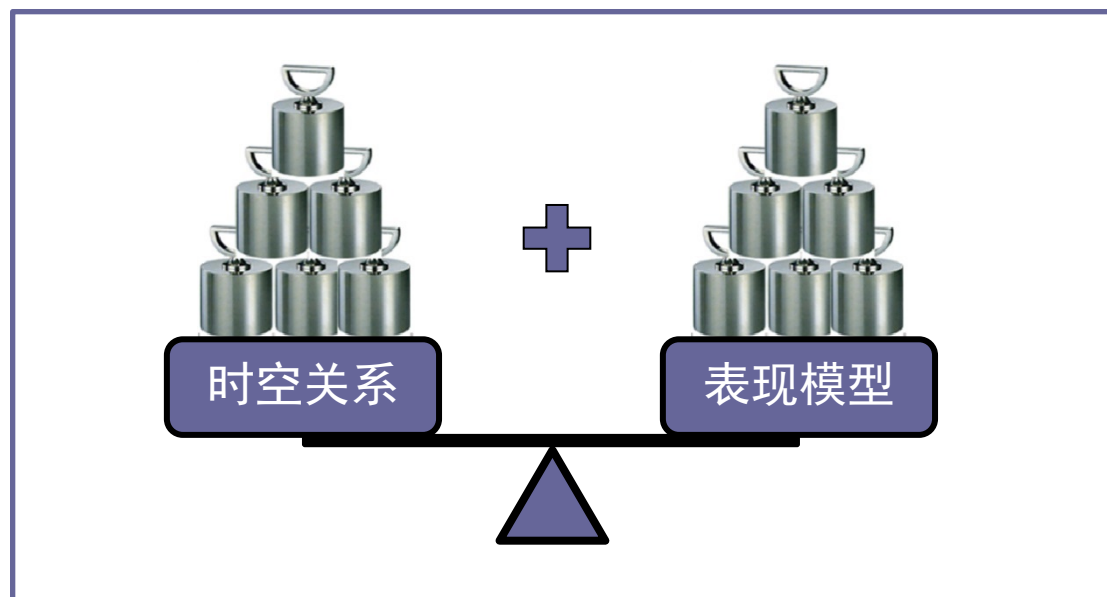
研究动机

- 跨头跟踪：解决两个问题
 - 目标关联：对不同摄像机的观测数据进行数据关联 (i.e., 监控区域的拓扑结构, 时空约束)
 - 对应关系：判断是否为同一目标 (i.e., 特征匹配)
- 传统的特征匹配注重于建立目标的表现模型
- 众多优秀特征：颜色特征、几何特征、区域特征等  ✓
- 目标关联 \neq 目标重识别  ✗
- 监控环境复杂，对鲁棒性的要求
- 目标运动和转移的规律依赖于摄像机网络的拓扑结构

时空相关性

研究动机

- 如何协同多个摄像头来可靠地跟踪目标?
- 需要将摄像机的时空拓扑关系与目标的外观表现模型匹配算法进行结合。



提纲

- 研究背景
- 相关工作
- 算法
 - 时空关联图模型的建立
 - 数据关联与相似性度量
- 实验
- 结论及贡献

相关工作

- 单摄像头目标跟踪
 - TLD算法 [Kalal et al 2012]
- 图模型
 - 利用摄像机网络的拓扑结构来分析视频监控活动 [Stauffer03, Makris04, Tieu05, Song07, Loy09, Wang10]
- 特征匹配
 - 颜色直方图、SURF纹理特征等
- 快速特征融合
 - 区域协方差描述子 Region covariance [Tuzel et al 2006]

提纲

- 研究背景
- 相关工作
- **算法**
 - 时空关联图模型的建立
 - 数据关联与相似性度量
- 实验
- 结论及贡献

相似性度量

- 相似性度量:

对于两个目标，采用如下方式进行度量：
目标A和目标B的相似度

$$Cor(A, B) = \alpha \cdot Sim(A, B) + (1 - \alpha) \cdot E[|T_A - T_B|]$$

表现特征项 (AM)

平衡系数

时空关联项
(GM)

基于网络拓扑的空间关联

- 图模型定义: $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$, $v_i \in \mathcal{V}$, $e(i, j) \in \mathcal{E}, 1 \leq i, j \leq N$
- 两个空间约束:
 - 如果跟踪目标在当前视域的节点处能够监测到, 则使用单摄像头跟踪算法对其进行持续跟踪
 - 如果跟踪目标在某一节点处消失, 则在其相邻一步可达的节点处进行目标检测, 再进行匹配与关联

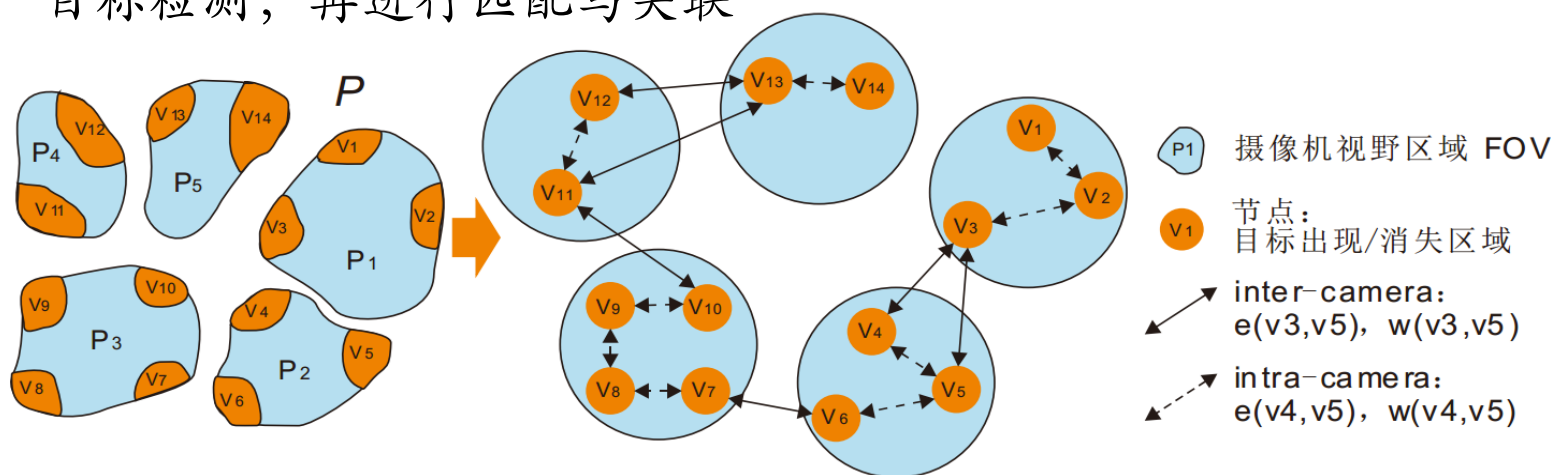


图1 摄像机网络与图模型建立示意图

Fig. 1 Example of camera network and its corresponding graph model

基于统计学习的时间关联

- 边权的计算：对通过相邻节点的观测时间序列 T_{ij} 进行聚类
- 期望权重相似度^[1]： (EWS)

$$w(e) \triangleq \exp(e) - 1$$

$$E[w(e)] = \exp\left(e' + \frac{\sigma^2}{2}\right) - 1$$

边 $e(i, j)$ 的权重根据高斯分布建模
 $p(e) \sim \mathcal{N}(e', \sigma)$

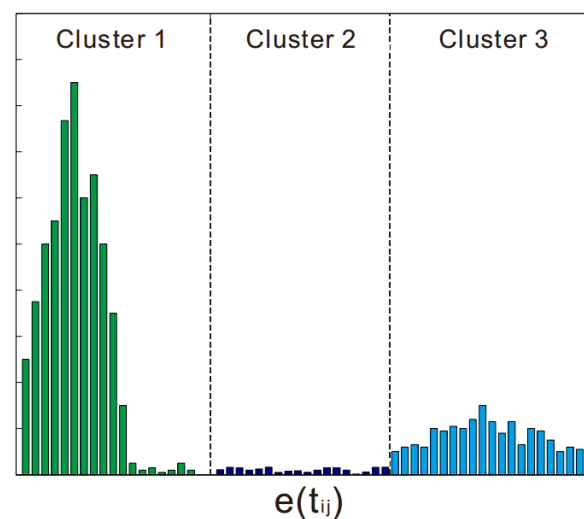
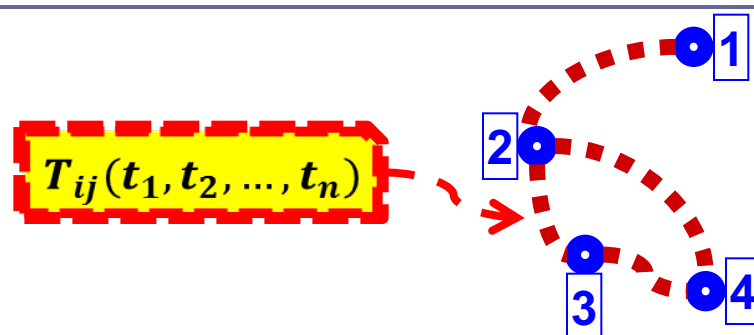


图 2 通过训练数据学习而来的节点 v_i 和 v_j 之间边权统计直方图

Fig. 2 Histogram of edge weights between v_i and v_j learned during the training phase

特征提取与相似性度量

- 相似性度量:

对于

节点i处观测到的目标n
节点j处观测到的目标m

目标

平衡系数

方式

时空关联项
(GM)

$$Cor(O_n^i, O_m^j) = \alpha \cdot Sim(O_n^i, O_m^j) + (1 - \alpha) \cdot E[w(e)]$$

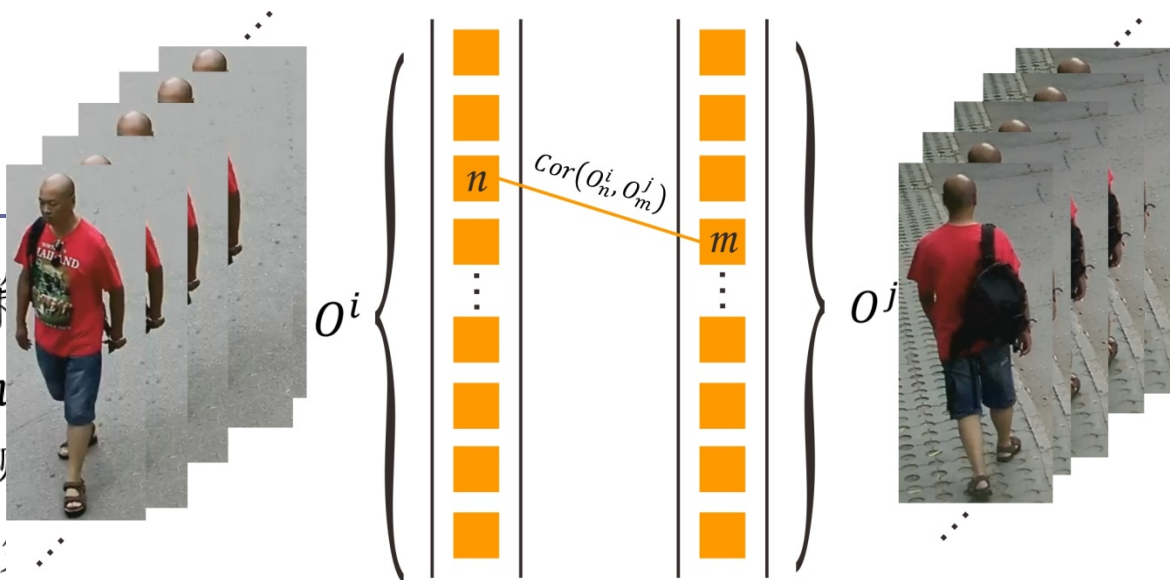
表现特征项 (AM) , 颜色 + SURF特征
采用区域协方差描述子进行特征融合^[2]

$$Sim(C_1, C_2) = \sqrt{\sum_{i=1}^s \ln^2 \lambda_i(C_1, C_2)} \quad C_R = \frac{1}{s-1} \sum_{k=1}^s (\mathbf{z}_k - \mu)(\mathbf{z}_k - \mu)^T$$

目标关联

- 基于二部图匹配第

- 对目标消失节点
- 的节点 v_j 处的观测
- 基于二部图匹配



目标函数: $\hat{\mathcal{M}} = \operatorname{argmax} \sum_{n=1}^N \sum_{m=1}^M \operatorname{Cor}(O_n^i, O_m^j) \cdot X(O_n^i, O_m^j)$

$\{X(O_n^i, O_m^j)\}_{N \times M}$ 为0-1二值矩阵, 如果目标 O_n^i 与目标 O_m^j 相匹配, 则 $X(O_n^i, O_m^j) = 1$, 否则为0

候选节点计算: $v_j = \operatorname{argmax}_{j \in \{j | \exists e(v_i, v_j), \text{ s.t. } v_i \rightarrow \text{next} = v_j\}} \mathcal{M}(v_i, v_j)$

对节点 V_i 所有一阶邻域节点的观测帧序列进行遍历

提纲

- 研究背景
- 相关工作
- 算法
 - 时空关联图模型的建立
 - 数据关联与相似性度量
- **实验**
- 结论及贡献

实验设计

- 数据集: PKU-SVD-B^[3]



图 4 5个目标跟踪实例

Fig. 4 Five examples of the image of people in the network

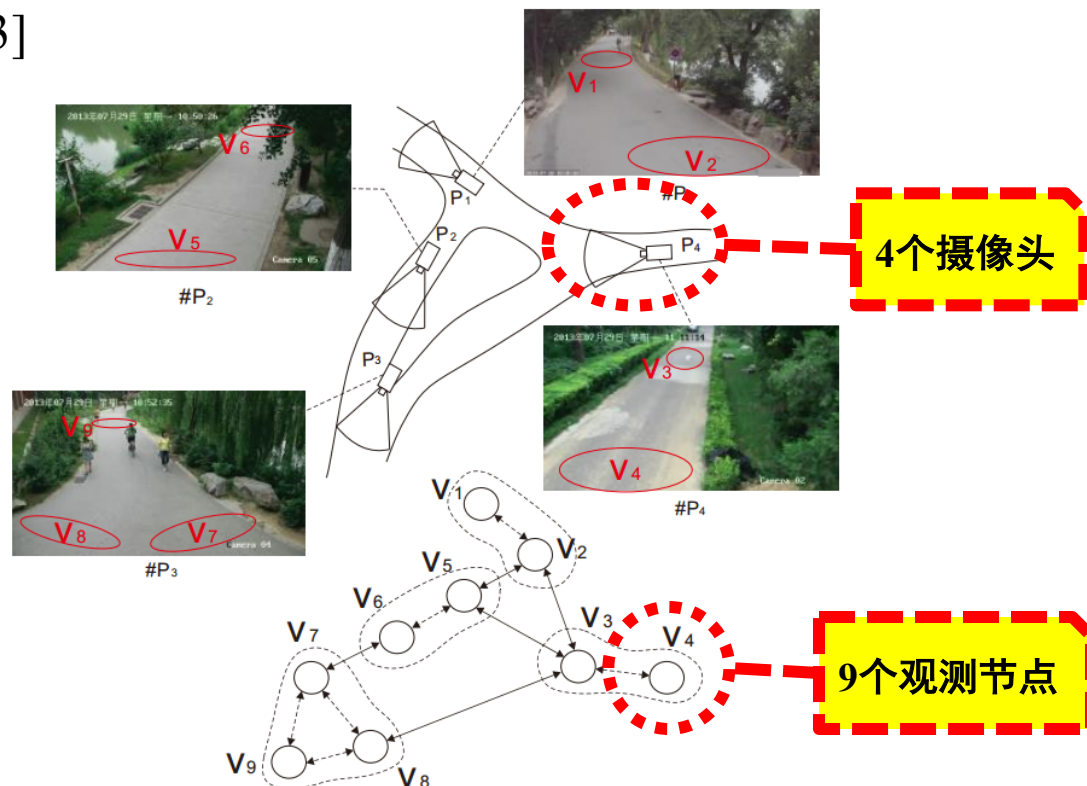
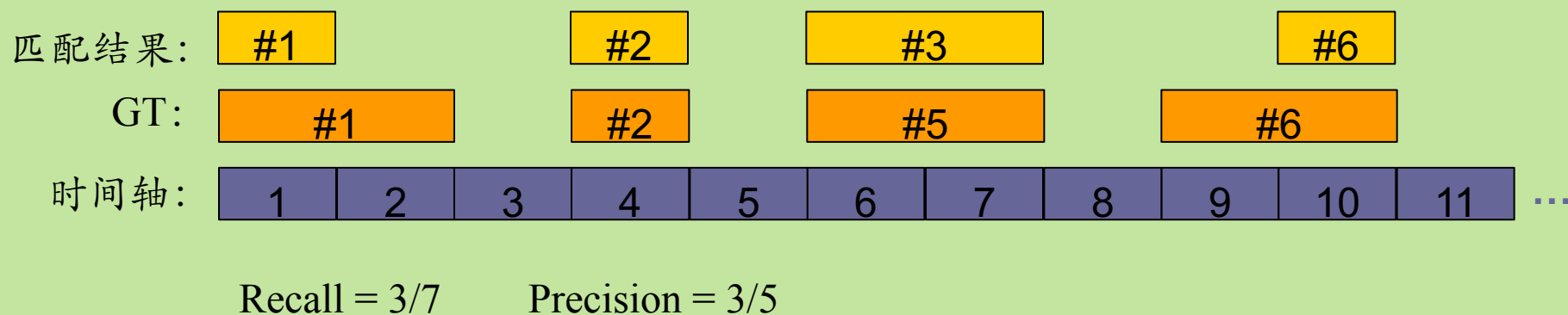


图 3 摄像机视域及网络拓扑

Fig. 3 The topology of camera views

性能评价

- 某个指定对象的跟踪结果:



- $$Recall = \frac{\sum_{i=1}^P C_i}{\sum_{i=1}^P G_i} \quad Precision = \frac{\sum_{i=1}^P C_i}{\sum_{i=1}^P D_i} \quad F_1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

- 每个节点在时间轴上设置 $F_i (i = 1, 2, \dots, N)$ 个观测帧序列
- P : 总的节点个数
- $G_i (G_i < F_i)$: 跟踪对象的Ground Truth轨迹序列在第 i 个节点处的数目
- C_i : 跟踪对象在节点 i 处正确匹配到的帧的数目。
- D_i : 跟踪对象在节点 i 处包含的所有匹配帧的数目。

比较结果

表 1 目标跟踪精度及对比

Table 1 Tracking accuracy and comparison

目标 ID	方法 对比	准确率 (Precision)	召回率 (Recall)	F 值 (F1-measure)
1	AM+GM	0.5956	0.6006	0.5980
	AM	0.4923	0.4950	0.4936
2	AM+GM	0.6559	0.6264	0.6408
	AM	0.5019	0.5996	0.5464
3	AM+GM	0.6519	0.6592	0.6555
	AM	0.5611	0.6040	0.5817
4	AM+GM	0.6537	0.6590	0.6563
	AM	0.5580	0.5976	0.5771
5	AM+GM	0.6021	0.6878	0.6421
	AM	0.5056	0.5590	0.5309

结合目标表现模型与图模型的方法在跟踪匹配的精度上能够得到约10%的提升

传统的基于特征的识别方法没有加入时空关联约束，会产生较多误匹配，因此准确率较低

随机参数实验

$$Cor(\cdot) = \alpha \cdot Sim(\cdot) + (1 - \alpha) \cdot E[\cdot]$$

平衡系数:

手工调整参数的最佳性能范围: 0.05-0.15

对每一个目标进行跟踪时, 在该范围内
随机选取10组参数进行实验

大部分数据集中在平均值附近, 说明在 α
的经验范围内, 算法性能比较稳定

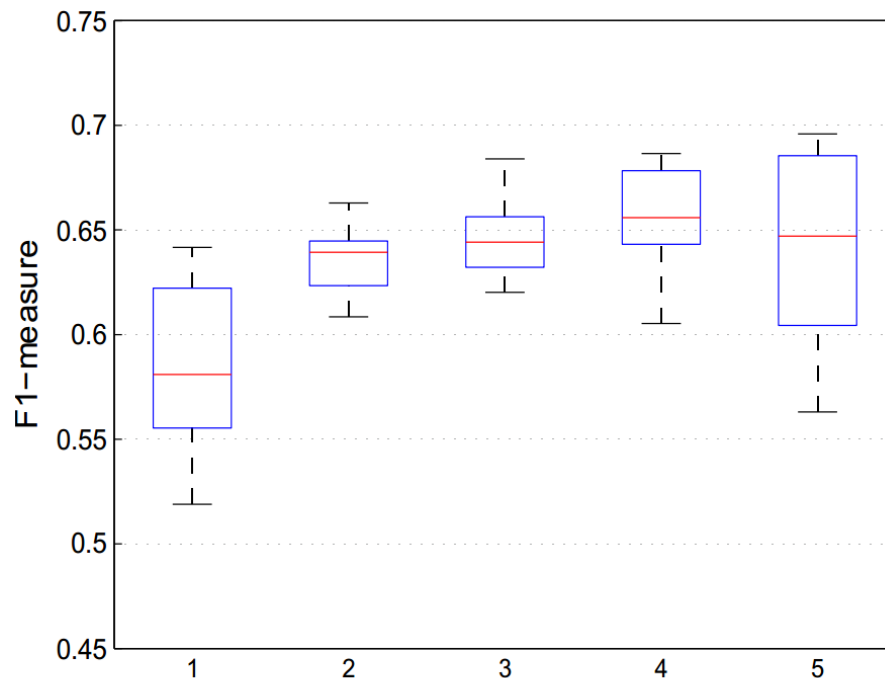


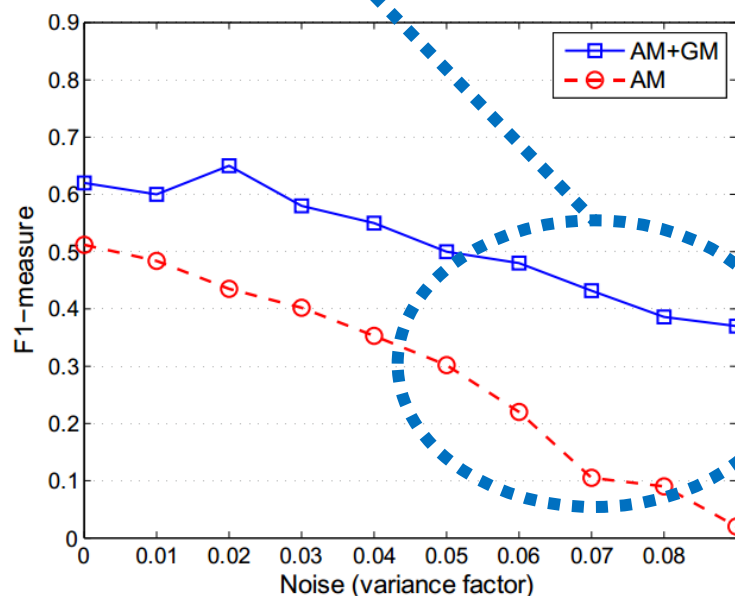
图 5 在不同跟踪目标上的随机参数实验

Fig. 5 The experiment of random parameters in different tracking case

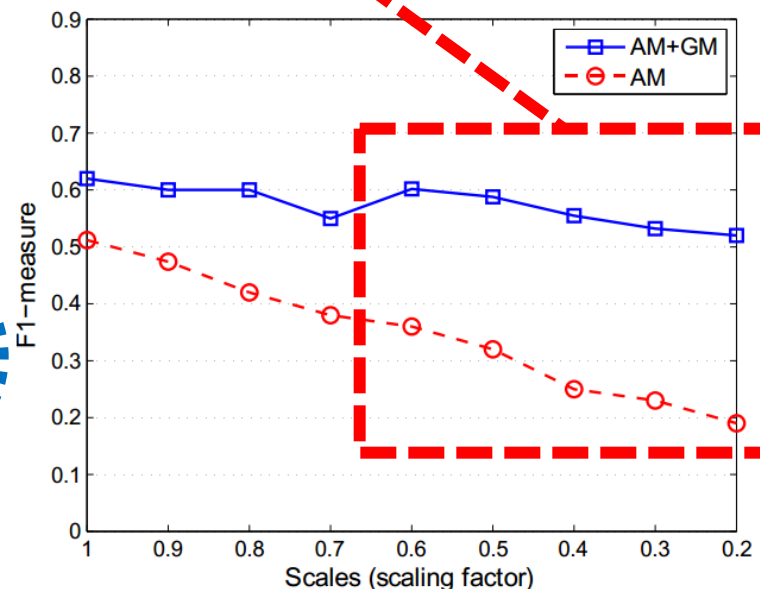
鲁棒性实验

使用高斯噪声对检测的帧序列进行不同程度模糊，AM方法会逐渐失效

监控视频中的运动目标由于远离或者靠近摄像机，往往会产生较大幅度的尺度改变。不同尺度之下我们的方法也具有较好的鲁棒性



a 噪声鲁棒性



b 尺度鲁棒性

图 6 跨摄像头跟踪时不同方法对噪声和尺度的鲁棒性

Fig. 6 The robustness of different approach to noise and scale variances in object tracking

提纲

- 基础问题
- 相关工作
- 算法
 - 时空关联图模型的建立
 - 数据关联与相似性度量
- 实验
- 结论及贡献

结论及贡献

- 将特征融合与时空关联相结合来进行多摄像机监控环境下的目标跟踪，能够充分利用现有的各种优秀的特征提取方法并将其融入提出的算法框架
- 采用区域协方差算子对目标外观特征进行融合，在噪声和尺度上具有较好的鲁棒性
- 采用了针对多个跟踪序列的二部图匹配策略来解决目标匹配问题，相比于传统的特征匹配方法，可以降低匹配的错误率，修正误匹配

谢谢



北京邮电大学

Beijing University of Posts and Telecommunications