

计算机视觉——目标跟踪

2025年春季
桑农 王岳环



目标跟踪任务

Image 1

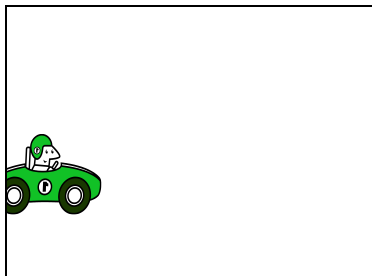


Image 2

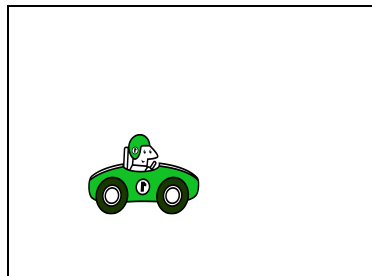


Image 3

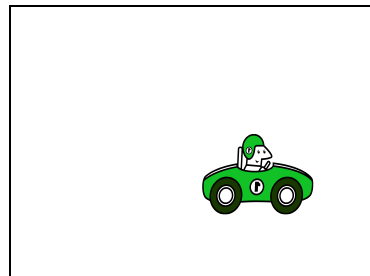
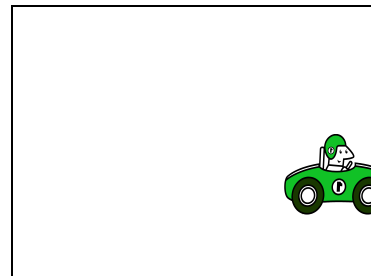


Image 4

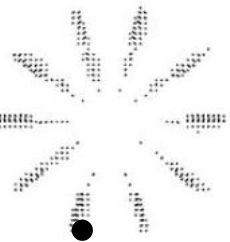


- 可以确定目标的位置吗？
- 可以确定目标的速度吗？
- 可以预测目标的位置吗？

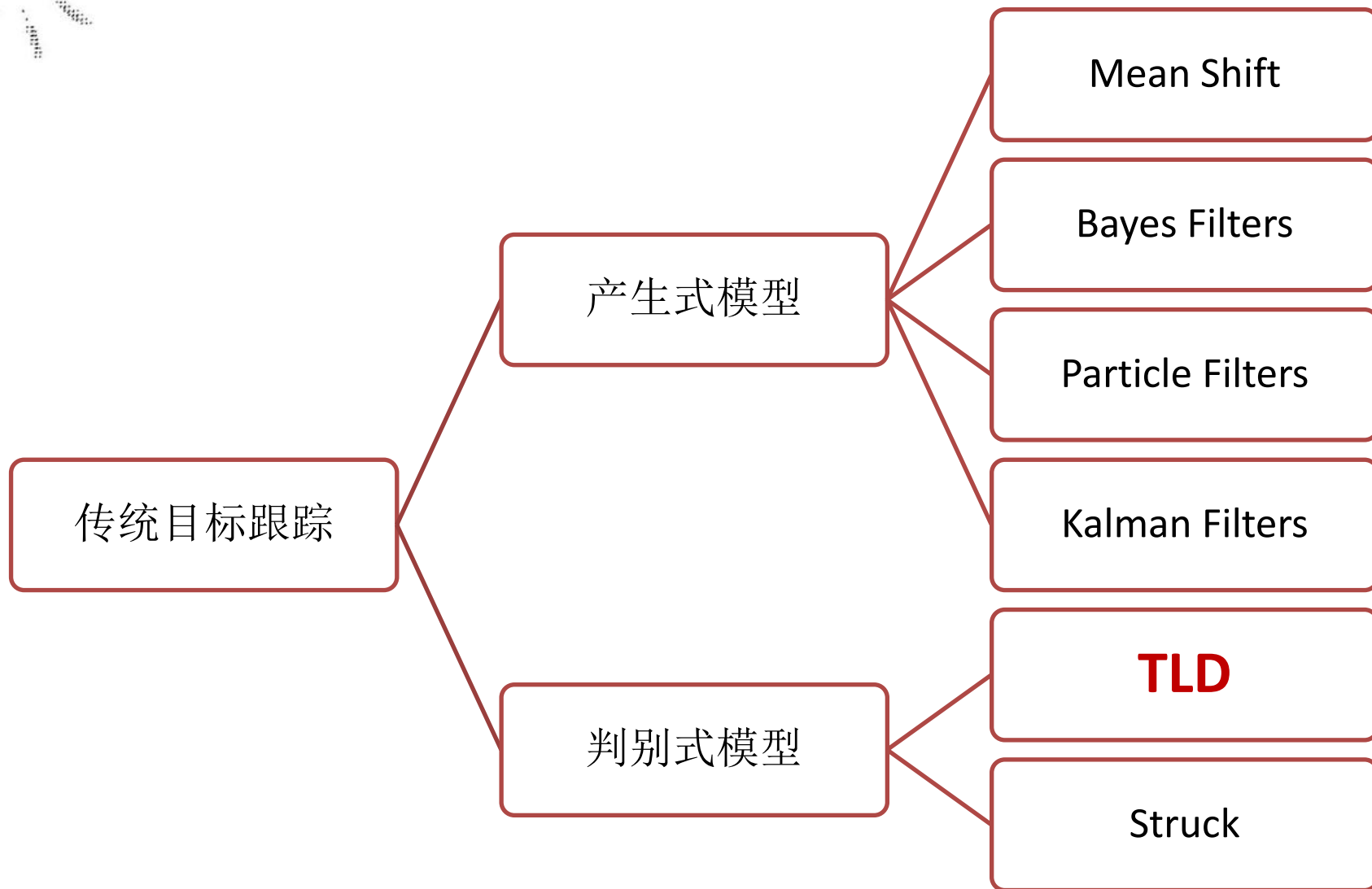


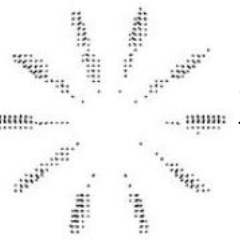
目标跟踪问题

- 给定图像序列
- 找到运动目标的位置
- 摄像头可以是禁止或者运动
- 假定：每一幅图像中都能找到目标
- 问题：在序列图像上进行跟踪
- 目标跟踪是计算机视觉的基本问题之一



目标跟踪方法





基于检测的目标跟踪方法

- 检测 VS 跟踪



t=1



t=2

...



t=20



t=21

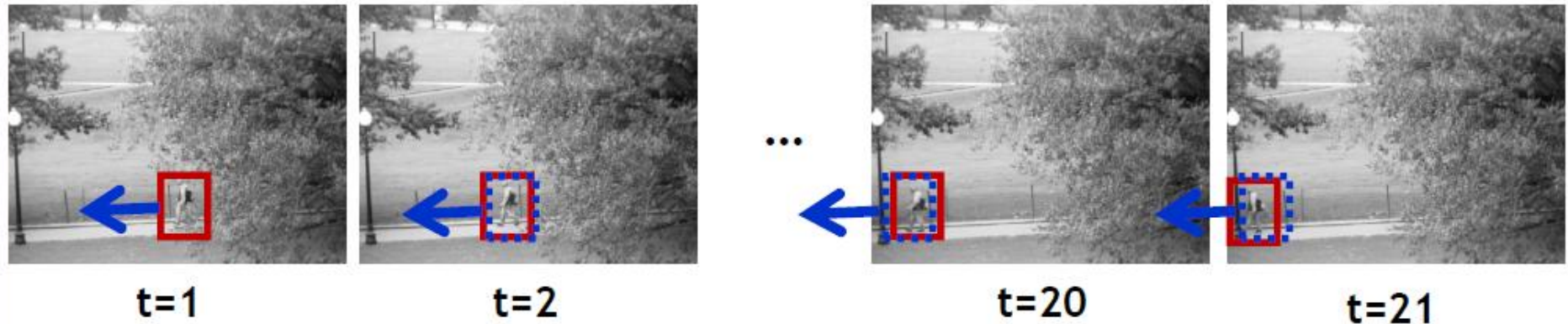
Detection vs. Tracking



- **Detection**

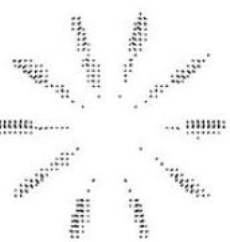
- We detect the object independently in each frame and can record its position over time, e.g., based on blob's centroid or detection window coordinates.

Detection vs. Tracking



- Tracking with *dynamics*:

- We use image measurements to estimate the object position, but also incorporate the position predicted by dynamics, i.e., our expectation of the object's motion pattern.

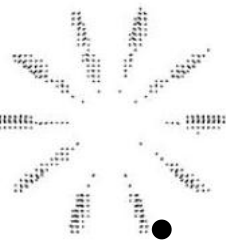


TLD 目标跟踪方法

TLD: Tracking-Learning-Detection

作者相关文章:

- Z. Kalal, K. Mikolajczyk, and J. Matas, "[Tracking-Learning-Detection](#)," Pattern Analysis and Machine Intelligence 2011.
- Z. Kalal, K. Mikolajczyk, and J. Matas, "[Face-TLD: Tracking-Learning-Detection Applied to Faces](#)," International Conference on Image Processing, 2010.
- Z. Kalal, K. Mikolajczyk, and J. Matas, "[Forward-Backward Error: Automatic Detection of Tracking Failures](#)," International Conference on Pattern Recognition, 2010, pp. 23-26.
- Z. Kalal, J. Matas, and K. Mikolajczyk, "[P-N Learning: Bootstrapping Binary Classifiers by Structural Constraints](#)," Conference on Computer Vision and Pattern Recognition, 2010.



TLD 目标跟踪方法

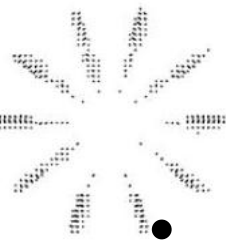
- 目标

长期, 实时 目标跟踪 (未知类别)

- 挑战

- 表观变化
- 部分或全部遮挡
- 目标可能短期消失
- 相似物体
- 实时性能
-

模型更新:
逐帧更新
选择更新



TLD 目标跟踪方法

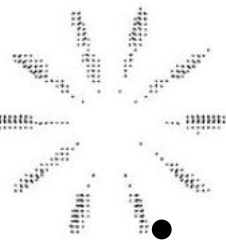
- 目标

长期, 实时 目标跟踪 (未知类别)

- 挑战

- 表观变化
- 部分或全部遮挡
- 目标可能短期消失
- 相似物体
- 实时性能
-

通过检测器初始化跟踪器



TLD 目标跟踪方法

- 目标

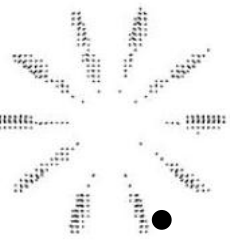
长期, 实时 目标跟踪 (未知类别)

- 挑战

- 表观变化
- 部分或全部遮挡
- 目标可能短期消失
- 相似物体
- 实时性能
-

重识别

A light orange rectangular callout box with a blue border. A blue line extends from the bottom-left corner of the box, pointing towards the text '相似物体' in the list of challenges.



TLD 目标跟踪方法

- 目标

长期, 实时 目标跟踪 (未知类别)

- 挑战

- 表观变化
- 部分或全部遮挡
- 目标可能短期消失
- 相似物体
- 实时性能
-

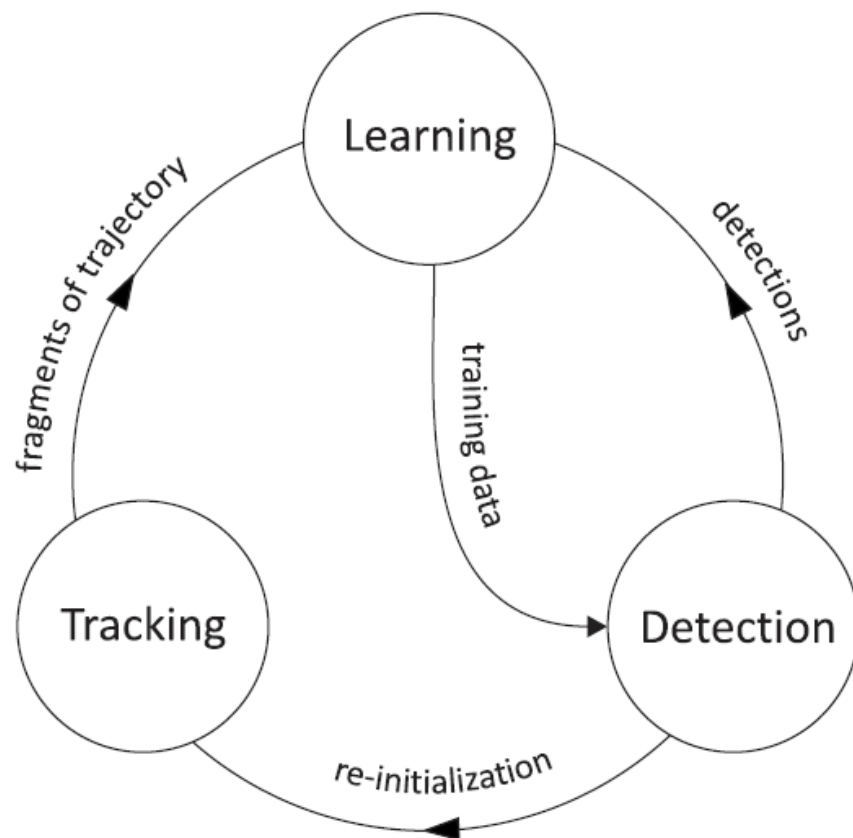
- 方法

- **TLD**: a Framework that integrates adaptive **T**racking with online **L**earning of object-specific **D**etector

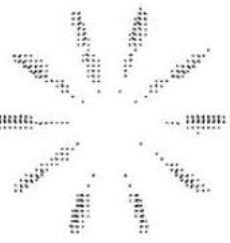


TLD目标跟踪方法

- TLD 框架

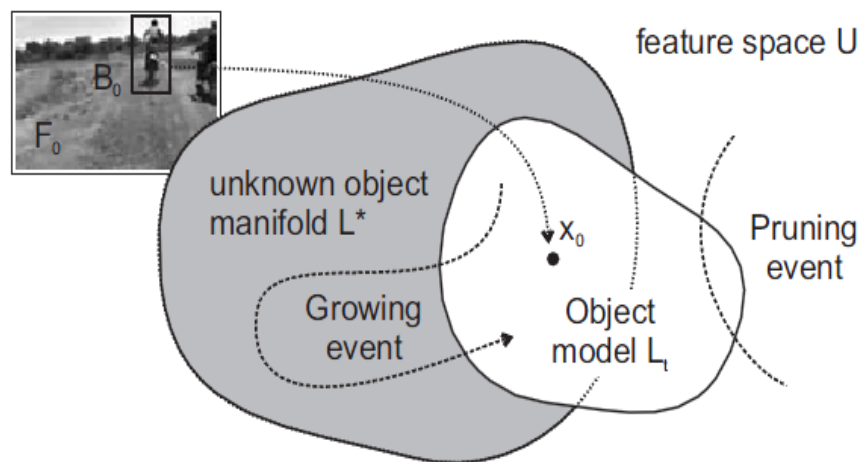


- **跟踪器**是观察帧与帧之间的目标运动
- **检测器**将每帧图像当成独立的，然后去定位。检测器可以定位之前所有出现的目标
- **学习器**评估检测器的误差，并更新检测器防止后续错误



TLD 目标跟踪方法

- TLD 框架



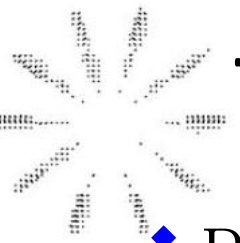
- **Growing events** 通过在线模型扩展空间
- **Pruning events** 去除错误样本



TLD目标跟踪方法

❖ P-N learning

- **目的**：提高检测器性能。采用半监督学习方法进行在线检测器更新，每一帧评估当前检测器，针对错误情况进行更新避免将来出现类似错误
- **两类检测器错误**：假阴 && 假阳
- P-N learning主要思想是利用两类“专家（expert）”解决这两类错误
 - **P-expert**：识别假阴
 - **N-expert**：识别假阳



TLD目标跟踪方法

❖ P-N learning

- P-N learning**主要思想是利用两类“专家（expert）”解决这两类错误**
 - **P-expert**: 识别假阴
 - **N-expert**: 识别假阳
- **专家的特点**
 - 采用独立的信息
 - 可以反驳对方
 - 可能发生错误



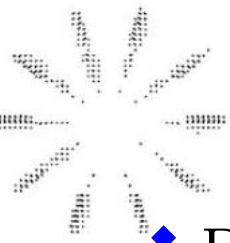
TLD 目标跟踪方法

❖ P-N learning

■ 数据

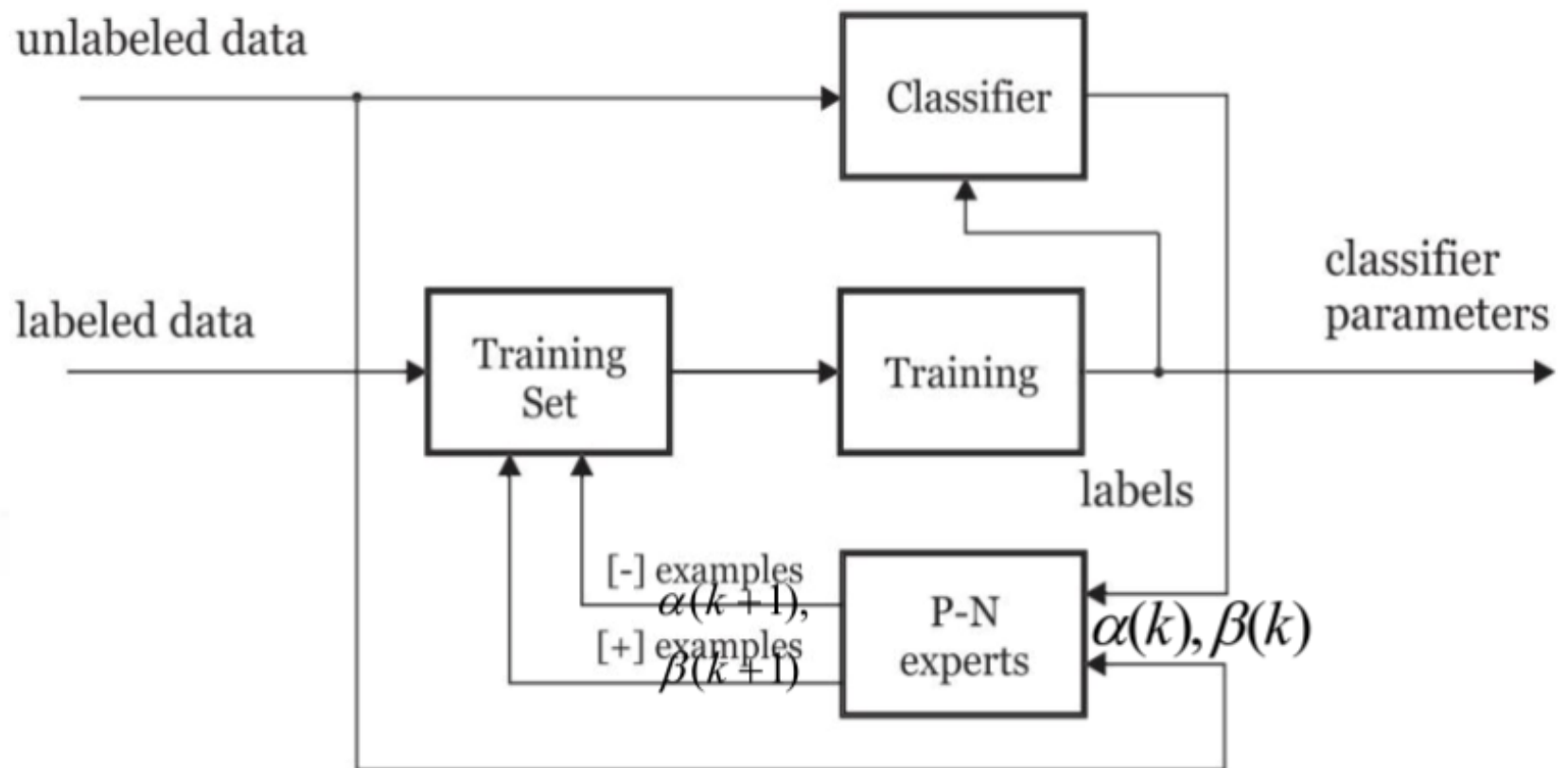
- 标记数据集: $L_l = (\mathcal{X}, \mathcal{Y})$
- 未标记数据集: $X_u = (X)$

- P-N学习的任务是根据标记数据集学习一个分类器 $f: \mathcal{X} \rightarrow \mathcal{Y}$ ，并利用未标记数据集提升其性能
- 分类器 $f: \mathcal{X} \rightarrow \mathcal{Y}$ 为一个来自由参数 θ 参数化的函数族 F 的函数，训练过程和参数 θ 的估计有关



TLD 目标跟踪方法

❖ P-N learning



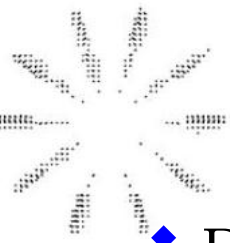


TLD目标跟踪方法

❖ P-N learning

■ Procedures of P-N Learning

- 学习过程的初始化：将标记数据集 L 嵌入训练集
- 训练集用于监督学习，训练一个分类器，即估计初始参数 Θ^0
- 在第 k 次迭代中，利用之前训练好的分类器对未标记样本进行分类， $y_u^k = \hat{f}(x_u | \Theta^{k-1})$, for all $x_u \in X_u$.
- P-N experts对分类进行性能评估，找到错误分类的样本，并加入训练集
- 重新训练分类器，即估计参数 Θ^k
- 迭代，直到收敛或者达到迭代次数



TLD 目标跟踪方法

❖ P-N learning

■ P-N Learning 的关键要素

- 估计分类器的错误：将两种错误分别处理

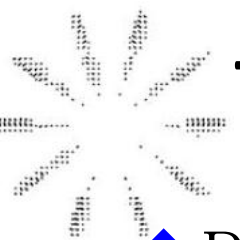


TLD目标跟踪方法

❖ P-N learning

■ P-N Learning的关键要素

- 估计分类器的错误：将两种错误分别处理
- 无监督的数据集根据当前分类器分为两部分，每部分都用不同的expert处理
 - P-expert 分析错分的负样本，赋予正的标签并加入训练集。在第 k 次迭代，P experts输出 $n^+(k)$ 个正样本
 - N-expert 分析错分的正样本，赋予负的标签并加入训练集。在第 k 次迭代，N experts输出 $n^-(k)$ 个负样本

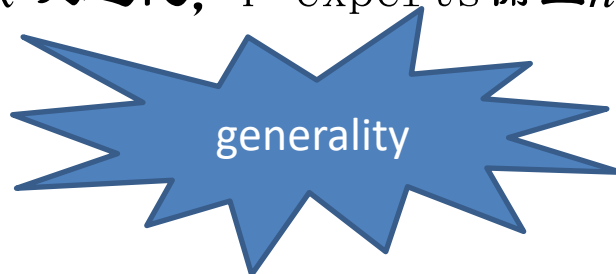


TLD目标跟踪方法

❖ P-N learning

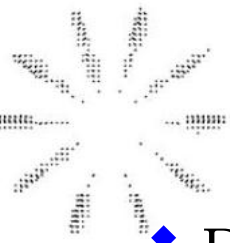
■ P-N Learning的关键要素

- 无监督的数据集根据当前分类器分为两部分，每部分都用不同的expert处理
 - P-expert 分析错分的负样本，赋予正的标签并加入训练集。在第 k 次迭代，P experts输出 $n^+(k)$ 个正样本



- N-expert 分析错分的正样本，赋予负的标签并加入训练集。在第 k 次迭代，N experts输出 $n^-(k)$ 个负样本

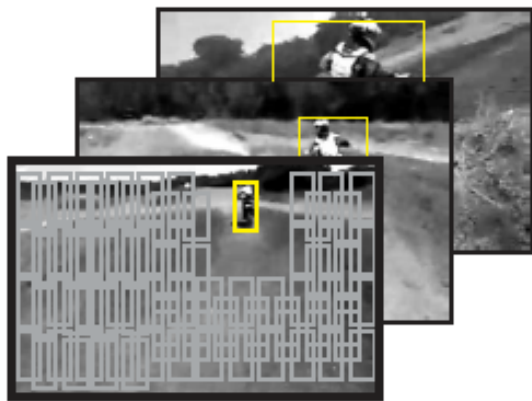




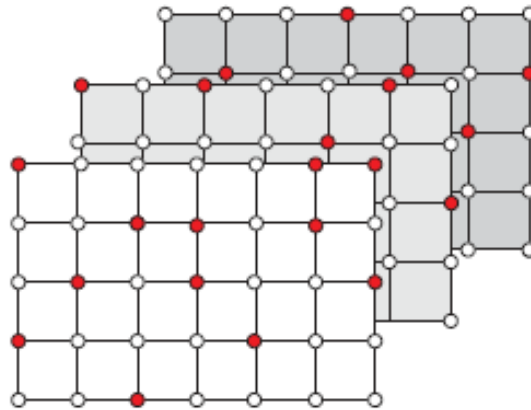
TLD 目标跟踪方法

❖ P-N learning

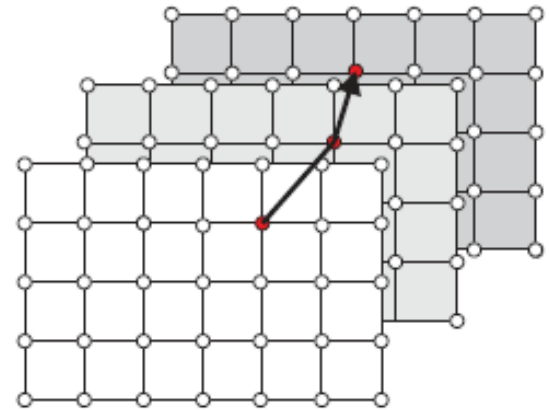
■ Design of experts



a) scanning grid



b) unacceptable labeling



c) acceptable labeling

- 每个图像块的标签是**独立的**
- 这种唯一连续的性质称为**结构性**
- P-N 学习的关键就是找到这种结构性的数据，来判别检测器所产生的错误标签



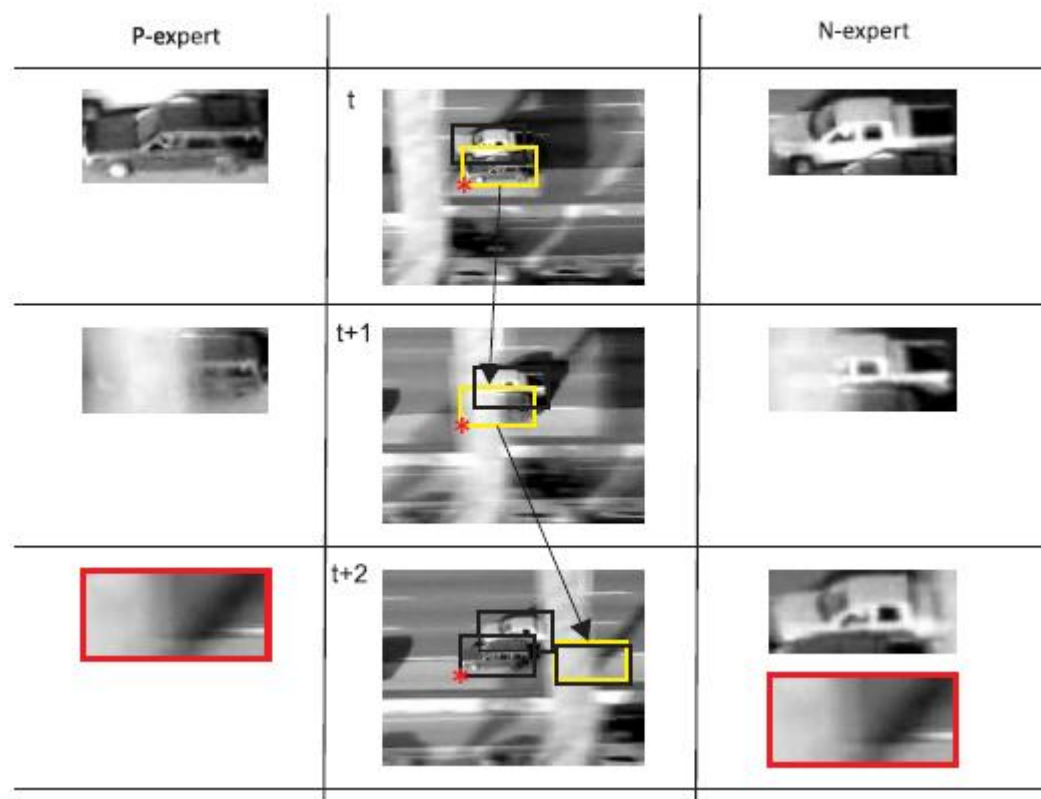
TLD目标跟踪方法

❖ P-N learning

■ Design of experts

P-experts寻找视频序列中的**时域**上的结构性特征，并且假设目标是沿着轨迹线移动的。P-experts记录目标在上一帧中的位置，并根据帧与帧之间的跟踪算法来预测目标在当前帧中的位置。如果**检测模块**将目标在**当前帧中的位置**标记为**负标签**，那么P-experts就产生一个**正样本**。

N-experts寻找视频序列中的**空域**上的结构性特征，并且假设目标在一个视频帧中只可能出现在一个位置。N-experts对**检测模块**在当前帧中的所有输出结果以及**跟踪模块**的输出结果进行分析，并找到**置信度最高**的结果。同**置信度最高**区域之间没有重叠的图相片被认定为**负样本**。另外，具有最大可能性的那个区域，被用于**重新初始化**跟踪模块。



- 目标是下面深色的车
- 黑色框是检测器检测到的正样本
- 黄色框是跟踪器产生的正样本
- 红星标记的是每一帧最后的跟踪结果
- 跟踪器代表产生正样本的P专家

- 在第t帧，检测器没有发现深色车，P专家认为深色车是正样本，N专家经过比较，认为深色车的样本更可信，所以把浅色车输出为负样本。
- 在第t+1帧的过程与之类似。
- 在第t+2帧时，P专家产生了错误的结果，但经过N专家的比较，又把这个结果排除了，算法仍然可以追踪到正确的车辆。



TLD 目标跟踪方法

- 局限性
 - 不能很好解决发生平面外旋转问题
 - 不能很好跟踪铰链式的物体
 - 仅适合单目标跟踪
 -



The end !