deepsort:使用深度关联度量的简单的在线实时跟踪

1. 前言

deepsort是sort的升级版，进一步优化了sort的效果。

1. 要解决什么问题

sort的关系度量使用的是矩形框的长宽，导致使用预测时，矩形框变动和目标遮挡时，身份切换次数太多。

1. 用了什么方法

关系度量替换未运动和外观信息。克服sort缺点:身份切换次数过多。

采用传统的单假设跟踪方法，递归卡尔曼滤波和逐帧数据关联。下面是跟踪系统的核心组件。

**跟踪轨迹处理和状态估计**

假设一个非常普遍的跟踪场景，相机是不加标记的，没有自我运动信息可用。

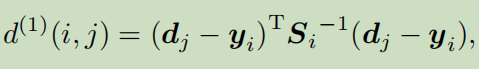
在跟踪场景上定义了八维空间状态(u,v,r,h,x`,y`r`,h`)。其中包围框中心位置(u,v)，纵横比r，高h，和它们各自在图片坐标中的速度。使用一个匀速运动的标准卡尔曼滤波和线性观测模型。以边界坐标(u,v,r,h)作为直接观察对象。

对每一个跟踪器k，计算自上一次成功关联到ak以来的帧数，此计数器在卡尔曼滤波预测的时候递增，并在轨迹和测量关联时候重置为0。超过预先设定最大范围Amax的轨迹被认为已经离开了场景并在轨迹集里面删除。当检测目标不能关联到现存的轨迹，则初始化一个新的轨迹。在前三帧时新轨迹被分类为暂定的，在此期间，期望每个时间步骤中都有一个

成功的度量关联，否则删除暂定的新轨迹。

**分配问题**

传统的解决卡尔曼预测状态和新的测量之间的关联的方法是建立一个可以用匈牙利算法求解的分配问题。在本问题公式中，通过两个合适的度量组合来集成运动和外观信息。为了结合运动信息，使用马氏距离在卡尔曼预测状态和新的测量之间：



使用(yi,Si)来表示第i条跟踪轨迹到度量空间的预测，用dj表示第j个检测框。通过计算检测位置和平均跟踪位置之间的标准差，马氏距离考虑了状态估计的不确定性。使用该度量能排除不可能的关联。用一个指示器来表示这个决策



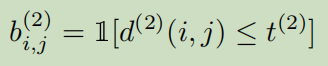
如果第i个跟踪和第j个检测允许关联时，指示器值为1。对于我们的四维测量空间，相应的Mahalanobis阈值为t (1) = 9.4877

当运动的不确定性低时，马氏距离是一种合适的关联度量，但在图像空间问题构成中，从卡尔曼滤波框架获得的运动状态分布仅提供了粗糙的对象位置估计。尤其是，未明确的相机运动可能会在图像平面引入快速位移，使马氏距离失效。因此，在分配问题中引入第二种度量。

对每个检测框dj，计算外观描述符rj，且||rj||=1.对每个跟踪器k，保留最后Lk=100个相关外观描述符集合Rk。因此，第二度量是计算第i个跟踪器和第j个检测在外观空间最小的余弦距离：



同样引入一个二进制变量来指示关联是否允许。

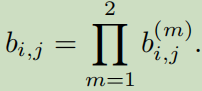


在一个独立的训练数据集上为此指标找到合适的阈值。实际上，使用预训练的CNN来计算边界框的外观描述符。

通过结合应对分配问题不同的方面，两个度量互补。一方面，马氏距离基于运动提供物体可能的位置信息，这在短期预测中十分有用。另一方面，当运动缺乏辨别力时，余弦距离考虑的外观信息对长时间遮挡后恢复身份十分有用。为了建立关联问题，使用加权和将两个度量结合



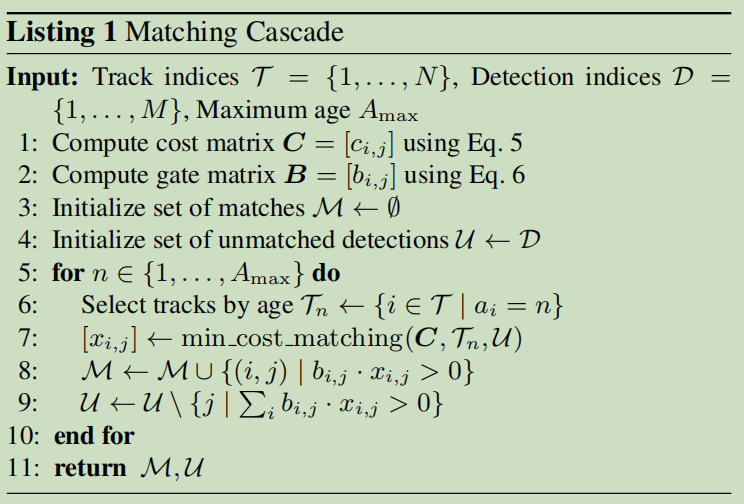
如果一个关联在两个度量的控制区域，则称为是允许的



可以通过超参数λ来控制每个度量对合并的关联成本的影响。在实验中，当摄像机运动较大时，设为0时合理的选择，此时，仅外观信息用于关联度量。然而，马氏距离仍然用来忽略基于卡尔曼滤波推断的目标可能位置的不可行的分配。

**级联匹配**

并不在全局匹配问题上解决测量和跟踪的关联，而是引入一个级联来解决一系列子问题。为了激发该方法，考虑下面的情况：当一个目标长时间遮挡后，后续的卡尔曼滤波预测增加了目标位置的不确定性。因此，概率质量在状态空间中扩散，观测的可能性变小。直觉上，关联度量通过增加测量和跟踪的距离来应对概率质量的扩散。反直觉的，当两个跟踪器竞争同一个检测结果，马氏距离支持更大的不确定性，因为它有效的减少了任何对于预计跟踪均值的检测的标准差的距离。这是个不被期望的行为，会导致跟踪破碎和不稳定的跟踪。因此，引入了一个级联匹配，将更频繁见到的目标给予优先级，编码进我们关联可能性中概率扩散的概念中。



将跟踪集合T和检测集合D，最大范围Amax作为输入。

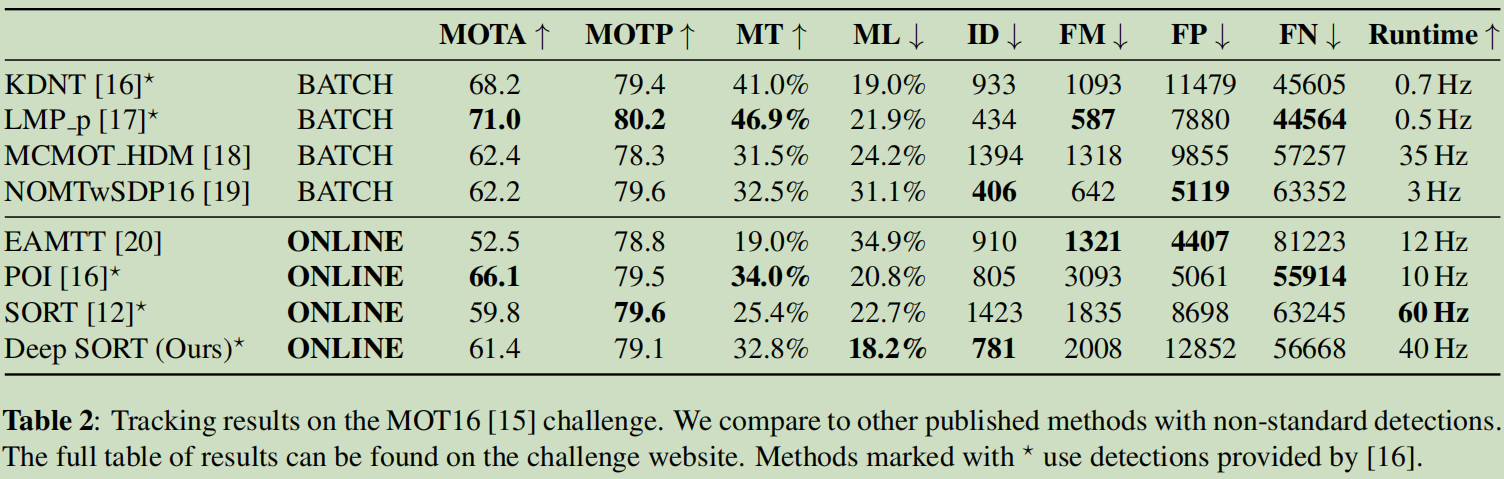
第一行和第二行计算关联代价矩阵和允许关联的矩阵。之后迭代跟踪范围n来解决跟踪范围增加的线性分配问题。第6行选择出在最后n帧中未被关联到检测目标的跟踪器kn的子集。第7行解决Tn和未匹配的检测目标U的线性匹配。在第8和第9行，更新匹配和未匹配的检测集合，在11行完成后返回。注意到此级联匹配优先范围较小的跟踪器，即最近匹配成功的匹配器。

在最后匹配阶段，在原始的sort算法上提出iou关联，用于未确定的和未匹配的范围n=1的跟踪器。这有助于解释突然的变化，如由于静态场景中的部分遮挡，提升错误初始化的鲁棒性。

**深度网络外观描述符**

通过使用简单的最近邻查询而不需要额外的度量学习，我们的方法成功运用需要辨别力强的特征，且在实际在线跟踪应用之前就离线训练好。通过使用CNN来提取这样的特征。

1. 效果如何



相比sort，ID switch从1423减少到781，减少了45%。由于Amax=30，允许30帧的丢失，造成了fp的提升。

1. 还存在什么问题和可借鉴的地方

deepsort是sort的扩展版，添加了外观特征来找回跟踪目标，减少了身份转换。