

随着公共安全对人类社会稳定发展的重要性愈发显著,城市中各类智能视频监控系统的规模也在变得越来越大,当需要针对特定目标(行人或者车辆)进行搜索时,传统图片匹配与以图搜图技术也渐渐无法满足现实的需要。针对表观相似对象在不同摄像头下难以区分这一难题,我们可以通过对特定目标在不同摄像头下的表观特征进行建模,从而实现目标精准搜索,因此,本问题具有重要的理论与实际应用价值。本文针对视频监控场景下目标精准搜索问题及其应用进行研究,主要贡献包括以下三点:

第一,为了实现高性能的快速运动对象检测,本文提出了一种能同时结合视频压缩域和像素域信息的快速运动对象检测框架。该方法可以在不进行视频解码的情况下利用压缩域运动向量信息提取出视频帧中运动前景区域,然后利用这些前景区域和运动向量的区域连通性分析结果来从检测速度和性能两方面优化像素域对象检测结果。在两个公开的监控视频数据集上的测试结果表明,在利用本快速运动对象检测方法之后,最终检测准确率与原始像素域对象检测算法相比可以得到 2%-6% 的性能提升。

第二,针对需要精准区分表观相似对象的目标精准搜索这一问题,本文提出了一种深度相对距离学习方法(DRDL)。该方法首先引入了一个簇聚类损失函数来作为类内与类间相对距离约束指导网络参数的学习,然后通过定义多任务混合差分深度网络(MDN)来同时学习对象的类别属性及高判别性深度特征,整个网络采用底层共享、中层分支及高层再融合的特殊结构来实现多任务融合。本文构建了一个超过 22 万张图片的车辆数据集“VehicleID”,该数据集中每张图片都有一个对应于汽车真实身份的 ID,且每辆汽车均包含两张及以上的不同图片样本。最终在该数据集上的实验结果表明,在不依赖于车牌的情况下,本文所提 DRDL 方法比同类方法在 MAP 指标上能有 5% 以上的提高。

第三,跨摄像头指定目标跟踪是目标精准搜索技术的一个重要应用场景,本文结合对象检测、单摄像头目标跟踪和跨摄像头目标精准搜索三种不同的视觉算法构建了一个完整的跨摄像头目标跟踪系统。该系统能够从多路摄像头设备获取实时拍摄的视频流数据,然后接收用户的目标跟踪指令对一个特定目标进行长时间的跨摄像头跟踪。但由于在实际场景中进行跨摄像头目标跟踪任务时还需要考虑视觉搜索过程中所产生的误检、漏检对象以及因为监控网络规模过大导致实际计算资源不足的情况,本文提出了一种同时结合目标对象视觉相似度与其时间空间历史信息的运动目标路径估计算法,并在此基础上使用一个基于计时器的摄像头搜索范围切换算法来避免每次搜索都要遍历整个摄像头监控网络的问题,有效降低了整个跟踪系统在计算资源上的需求。