

# 基于三维视频融合的 监控分析系统

» 与传统的单纯获取监控目标的实时音频、视频信息的视频监控系统相比,集成了空间位置信息的三维视频监控系统在增强用户空间位置意识、辅助用户应急决策等方面将发挥更大的作用。

■ 陈宝权

随着GIS技术的广泛应用,集成二维地图等空间位置信息的视频监控系统解决方案在视频监控领域正逐步走向成熟。与传统的单纯获取监控目标的实时音频、视频信息的视频监控系统相比,集成了空间位置信息的三维视频监控系统在增强用户空间位置意识、辅助用户应急决策等方面将发挥更大的作用。如何把三维GIS与广泛分布于城市各个角落的视频监控系统集成起来,已成为基于位置的视频监控系统的研究热点。

## 三维视频融合分析系统简介

在三维视频监控方面,目前比较普遍的做法是从各地实时监控影像中采集出图片,以标注的形式在三维场景中进行显示。当用户想知道某地的监控影像时,点击该标注即可弹出实时监控影像截图。这种做法只是视频与三维平台的初步结合,并没有发挥出三维场景的直观特点,也远没有达到视频位置与周边三

维信息的融合目标。与直接标注的方式不同,我们采用将视频投影到三维地图的方法,研发出一套基于三维视频融合的监控分析系统,使监控视频具有了空间位置感,实现了三维场景与动态视频的融合。

近年来公安部正在大力推行创建“平安城市”试点工作,在城市中安装了大量的视频监控设备以强化治安管理,防范和追述重大事件的发生。“平安城市”作为一个特大型且综合性很强的管理系统,主要通过视频监控系统,数字化城市管理系统以及道路交通等多个系统的紧密配合来实现城市的平安和谐。在“平安城市”建设的任务中,维护公共安全,建立高效可靠的城市安防体系与应急响应平台是整个系统实现对城市的有效管理和打击违法犯罪的关键。该平台所涉及的主要关键技术包括智能视频分析,生物识别以及射频识别技术等,并整合了环境监测,通路监测,物品

安全审查和出入口控制等多项功能的整体安全防护系统。

在刑事犯罪等重大事件发生后,为了分析犯罪现场和提取犯罪线索,公安局刑侦部门通常会调取在案发地点附近的多个摄像头相关时段的视频并对嫌疑人进行甄别和行为分析,通常需要众多有经验的刑侦人员对不同的视频记录前后反复对比观看,获取有价值的犯罪线索。然而,由于案件数目的不断增加,以及数以万计的摄像头所产生的海量视频数据已经对公安刑侦部门的案件处理能力造成巨大的压力,如何从这些视频数据中快速、自动和智能地提取,处理与理解场景中的关键信息,从而有效地应用于社会安全事件的解决已成为目前公安系统和城市管理部门最为关心的问题之一。

基于三维视频融合的监控分析系统结合了业界领先的计算机视觉技术,能够对在不同摄像头中出现的人员进行

有效检测、跟踪和关联并快速地甄别犯罪嫌疑人及车辆，从而大幅度减少刑侦人员查看视频的时间，显著提高案件的取证速度和侦破效率。同时，该系统采用了三维沉浸式显示，可给出监控区域和视频更为直观和立体的视觉效果，并增强了用户与系统交互的友好性和便捷性。通过对应用层的扩展，该系统还可以用于日常监控视频的在线处理，从而实时地实现对视频中人员和车辆等目标的智能识别，跟踪和行为分析，为日常的城市管理和生产生活服务。目前该系统已在福建泉州、浙江宁波等地的公安机关进行试点并进行产业化项目实施。

### 三维视频融合分析系统的组成

基于三维视频融合的监控分析系统以自底向上的形式可分为数据层、应用层和用户交互层上述三个层次，如图1左侧的结构图所示。图1右侧的关联图则给出了不同模块之间的相互关联性：应用层中的目标跟踪模块是系统的核心模块，其结合了应用层的追寻策略模块和数据层的地理信息系统(GIS)模块给予的信息来获取被跟踪目标的详细情报，并通过用户交互界面将上述情报传递给监控人员，从而便于其做出下一步的决策。

系统的数据层由视频管理和地理信

息系统(GIS)两个模块组成，用于给上层模块提供底层的原始待处理数据：

**视频管理：**对监控视频进行图像去噪/增强等预处理，通过H.264标准对视频进行压缩并存储以及为了满足其他层的需求实现视频的调度和筛选。

**地理信息系统：**通过诸如Google Earth等GIS系统提供城市的地理分布信息以及各个摄像机所处的位置信息，如图2所示，系统的应用层是整个智能视频分析系统的核心，其主要由增强现实(Augmented Reality)，追寻策略和目标跟踪上述三个模块组成。

**增强现实：**根据地理信息系统提供的信息实现城市的三维模型构建，将各个摄像头对应的二维监控视频投影并融合到三维城市模型中以呈现直观而立体的视觉效果，如图3所示。此外，可根据用户需要对摄像机进行设置和切换。基于三维沉浸式的场景显示更加便于刑侦人员将视频内容与实际的周边环境相联系，从而增强监控的现场感，是目前国内市场里绝大多数智能视频分析系统所不具备的。

**追寻策略：**在地理信息系统输出的地图信息基础上根据到被跟踪目标所处位置的距离来对摄像机进行优先级定位，并重点关注被跟踪目标附近一定范围内的摄像机在案件发生时间段内拍摄的监控视频，从而提高系统整体的智能程度，其大致流程如图4左侧下方的流程图(Camera Locator)所示。

**目标跟踪：**选择通过追寻策略得到的最高优先级摄像机并对其视频内容进行分析，进而实现基于跨相机的单目标跟踪，其基本过程如图4中部的流程图所示。基于单个相机的目标跟踪流

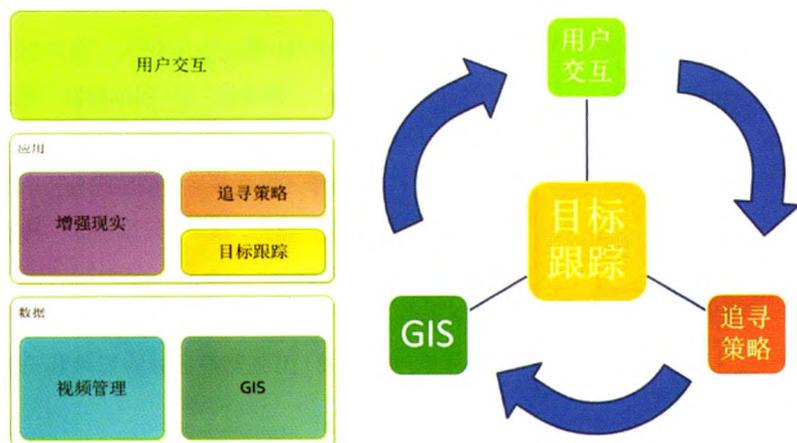


图1: 智能视频分析系统示意图



图2: Google Earth地理信息系统



(a) 城市三维建模



(b) 视频投影1



(c) 视频投影2



(d) 视频投影3

图3: 基于增强现实的三维城市模型

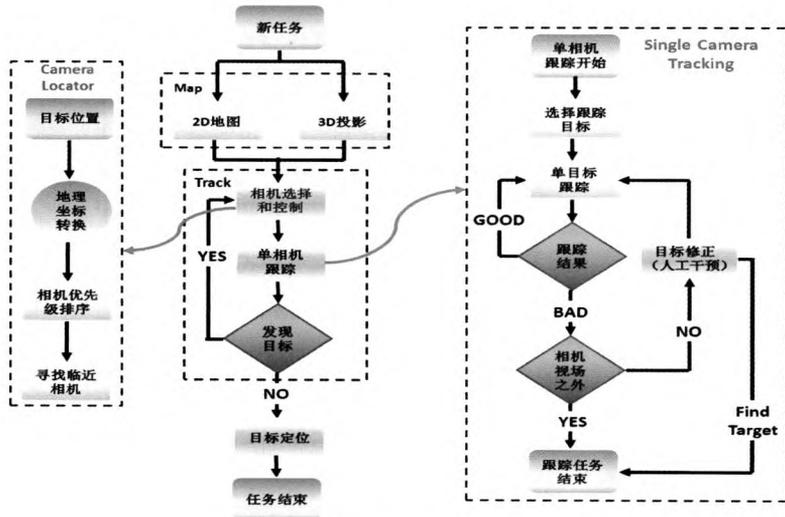


图4: 基于跨相机的单目标跟踪流程图



图5: 基于三维视频融合的监控分析系统用户界面

程如图4右侧的流程图 (Single Camera Tracking) 所示: (1) 选定一个目标并进行自动跟踪; (2) 在一定的人工干预基础上判断跟踪结果是否正确, 如果出错问题则暂停视频并人工对跟踪结果进行修正; (3) 重复上述过程直到完成整个跟踪流程。如果目标成功地在某个相机中被跟踪, 则在跟踪结束时输出目标在地图中的地理坐标, 否则将换优先级更低的相机重复上述单相机跟踪过程直到找到被跟踪的目标为止。

### 三维视频融合分析系统功能

系统的用户交互层将数据层输出的信息以清晰而直观的方式呈现给用户, 并允许用户方便地进行信息查询和系统设置, 其用户界面如图5所示。用户交互层具有“三维显示”和“目标跟踪”两大功能。

基于三维视频融合的监控分析系统具备如下功能, 其中“三维显示”功能主要包含:

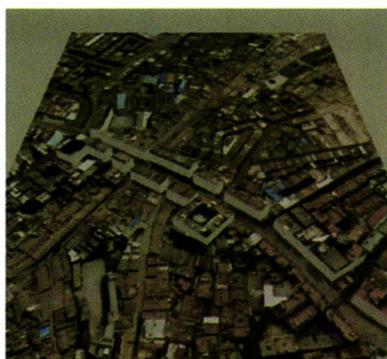
- (1) 查询案件相关信息;
- (2) 浏览监控区域的二维和三维地图;
- (3) 任意点击播放监控区域中各个摄像机记录的原始视频以及其在三维地图中的投影视频 (如图6所示);

(4) 可自由选择任意时间段摄像机拍摄的监控视频;

(5) 每一帧跟踪得到的目标区域采样列表;

“目标跟踪”功能主要包含:

(1) 单目标跟踪 (如图7所示): 在人工设定初始目标位置的情况下能够有效实现单目标的长时间持续跟踪, 尤其在被跟踪目标从一个摄像机视场中消失



(a) 三维街区模型



(b) 三维沉浸式监控视频播放  
图6: 基于增强现实的浏览效果

后在具有一定公共视场的另一个摄像机中出现时依然具有较为稳定的跟踪效果。

(2) 目标检索: 在单个摄像机跟踪的基础上, 通过检索类似目标物体的特征, 实现多个摄像机的同一目标分析以及检索, 并以图片的形式保存有效信息, 提高视频分析和存储效率。

(3) 多目标跟踪: 在单目标跟踪的基础上实现多种同一类别目标以及多种不同类别目标的跟踪, 如图9所示。多目标跟踪的优势在于不仅能够受关注人员乘坐不同交通工具时依然实现有效



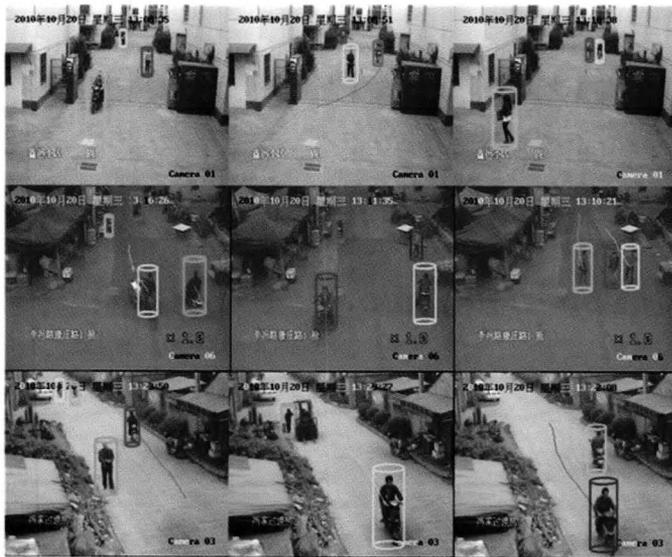
(a) 单目标跟踪结果

(b) 无公共视场的跨相机目标关联

图7: 单目标跟踪



图8: 目标检索



(a) 多个同一类别目标的同步跟踪



(b) 多种不同类别目标的同步跟踪

图9: 多目标跟踪

跟踪,而且还可以对多个不同的受关注人员实施同步监控,尤其针对街道,商场和小区门口等人员出入密集的场所。

(4) 异常行为分析:对监控视场中行为举止明显异于常人的目标进行筛选和跟踪,同时对其潜在的意图进行判断,从而提高系统监控的智能化程度。

### 结束语

基于三维视频融合的监控分析系统在不同的场所有不同的应用,逐渐发展为视频监控预警系统,在突发事件的应急指挥过程中,可以提供有效的决策支持,是视频监控系统“深度应用”的重要组成部分。CPS

作者单位:中科院深圳先进技术研究院