

图1 智能摄像机功能结构图

2.1 传感器单元

图像传感器是传感器单元的核心，一般采用 CCD 或 CMOS 感光器将光信号转换成电信号，通过接口将电信号传输给处理器单元。嵌入式智能摄像机安装的图像传感器按解析率、帧率、色彩以及镜头可以分为多种类型。通常输出高分辨率和高帧率的图像传感器需要较高的处理能力，而特殊镜头(比如全景镜头)则需要软件上的特殊处理方法。多图像传感器也常常出现在嵌入式智能摄像机设计中，用来完成立体视觉和协作跟踪等特殊功能。

2.2 嵌入式处理器

嵌入式处理器是智能摄像机的“大脑”，对视频数据进行初步的分析和处理，控制、辅助系统运行的硬件单元，它通常把通用计算机中许多由板卡完成的任务集成在芯片内部。

目前多数智能摄像机的硬件设计方法是通过系统总线(比如 PCI 总线)将负责视频分析计算的一个或者多个 DSP 处理器和负责网络通讯的控制处理器进行连接作为嵌入式智能摄像机的核心。智能摄像机中嵌入式处理器的选择主要有三种方案：

视频采集芯片+DSP 处理器结构。由视频采集芯片完成图像的预处理，DSP 处理器完成图像的编码压缩、处理。如美国德州仪器公司(TI)和美国模拟器件公司(ADI)的 DSP 芯片。TI 生产的 TMS320 系列 DSP 芯片由于价格低廉、简单易用、功能强大等特点，得到广泛应用，DM642 是 TMS320 系列中为视频应用而优化的高速处理器^[4,5]。独立 DSP 处理器的缺点是控制不够

灵活，不合作系统控制，因为 DSP 通常没有强大的操作系统。

多 DSP 处理器结构。有的 DSP 处理器完成视频处理算法，如视频压缩、视频分析、数据处理，其它 DSP 处理器完成前端摄像头、云台的控制以及网络控制等功能^[6]。如文献[8]中采用多个高性能 TMS320C64X(美国 TI 公司)DSP 处理器和 Xscale 450(美国 Intel 公司)网络处理器，两种处理器通过 PCI 总线进行连接，在该平台上能够进行实时运动检查、姿态识别等算法。多 DSP 结构满足智能摄像机的实时视频分析和压缩的要求，但是缺乏较强的事务管理功能，不能很好的进行系统整体的控制。

DSP 处理器+嵌入式处理器的双核结构。TI 公司近年推出双核的达芬奇处理器，如 TMS320DM644x 系列^[7]是基于 ARM926 和 TMS320C64x+DSP 双核、高度集成的数字媒体处理器。文献[7]中采用 TI 的 DSP+ARM 双核处理器 TMS320DM6446，主要由 DSP 完成图像压缩、分析，ARM 负责控制芯片各个模块的工作以及运行操作系统、网络协议、应用软件等。由于将图像的处理和控制两个芯片上分工完成，DSP 集中完成图像的处理分析，对大量视频数据的处理效率大大提高，ARM 处理器作为控制器优于 DSP 处理器。充分发挥了 ARM 的事件处理控制能力和 DSP 对数字视频的处理能力，双核结构是智能摄像机发展的方向。

2.3 网络通信

智能摄像机是智能化的网络摄像机，可借助网络实现远程监控，在远程不同地点的分控中心或同一个分控中心的不同主机上同时调看某一个或几个监控现场的视频数据，或远程控制前端摄像机。

智能摄像机通过内部网络服务器将音视频数据以及视频分析结果传输到监控中心主机上，传输方式主要有三种：(1) 以太网(Ethernet)；(2) 无线局域网(WLAN 等)；(3) 移动网络(GPRS/3G)。

智能摄像机上配有 RJ-45 以太网接口，可以与现有 IP 网络直接相连，基本支持现阶段主流的通信格式，比如支持 PPPOE、DNS、UDP、PTP、PTCP、HTTP、FTP 等^[8]。智能摄像机拥有独立的 IP 地址，经过处理的视频信息通过网络控制器转换成符合网络传输的数据流。用户可以通过标准的浏览器在 PC 机上对智能摄像机进行访问，观看实时画面，控制摄像机的镜头

和云台。

智能摄像机内部可以增加无线网卡模块或提供无线网卡的插槽,支持无线功能的智能摄像机可以通过无线局域网技术(WLAN)将采集、处理后的视频信息传送到近距离的监控中心。无线局域网技术标准 802.11,可实现十几兆至几十兆的无线接入。文献[10]设计了一种在局域网下传输实时视频的方案,采用 IEEE802.11 协议,但是无线网络及其协议无法给视频的实时传输提供 QoS 保障。

智能摄像机与 3G 无线宽带技术配合,使得移动监控的实现成为可能,采集压缩后的视频图像通过移动的 3G 无线网络模块传送到监控中心。3G 标准主要有欧洲主导的 WCDMA,它可支持 384Kbps 到 2Mbps 不等的数据传输速率;美国主导 CDMA2000 无线通讯标准,能够满足市场对移动通信容量和品质的高要求;中国国产 3G 标准 TD-SCDMA^[9,10]。3G 网络摄像机适用于野外、海上等无法接宽带网的场所。

3 智能摄像机的关键技术

从上述智能摄像机的功能结构可以看出,智能摄像机硬件技术已经较为成熟。在处理器平台上完成智能视频监控的任务,要用到智能摄像机的软件技术,视频编码压缩和视觉分析算法是实现嵌入式智能摄像机智能化的关键和核心技术。

3.1 视频编码压缩

新一代智能视频监控系统中应用较为广泛的视频压缩标准主要是 MPEG-4 和 H.264^[11]。MPEG-4 标准是超低数码率的视频压缩标准,其发展目的是基于 IP 的视频数据传送,最大优点是考虑了网络的传输问题,可以利用很窄的带宽,通过帧重建技术来压缩和传输数据,以求利用最少的数据获得最佳的图像质量。H.264 充分考虑了多媒体通信对视频编解码的各种要求,并借鉴了以往视频压缩标准的研究成果,具有比 MPEG-4 算法简单的特点,因而在网络视频监控领域具有明显的优势。H.264 获得优越性能的代价是计算复杂度的增加,从而增加了软硬件技术的难度。

目前的视频编码系统主要实现方法有三种:纯软件方式、通用 DSP 方式、专业处理器与专业芯片方式。以下对这几种方式作概括性介绍:

(1) 纯软件方式:可以采用直接修改移植开源代码的方式以求缩短开发周期。如开源 MPEG-4 编码器

XviD 是最新的 MPEG-4 codec,而且是第一个真正开放源代码的编码器。JM、X264 和 T264 是目前常见的 H.264 开源软件编解码器。软件压缩的缺点:对 CPU 处理能力、算法速度要求比较高,一般很难达到实时。

(2) 专业处理器与专用芯片方式:将压缩算法中的大量简单重复运算独立出来,用专门的硬件引擎实现,用一个嵌入式的 RISC 或 DSP 完成算法的其余任务。这种实现方式提高了运算效率和执行速度,开发周期短,但其扩展性差,价格通常比较昂贵。

(3) 通用 DSP 方式:基于专用图像处理 DSP 芯片实现,DSP 芯片采用改进的哈佛结构,具有对数字信号的高速处理能力。基于 DSP 的软件实现方式具有灵活性和可移植性等优点,成本也较低,虽然运算效率和执行速度不太理想,但基于 DSP 方式的实时视频压缩是智能摄像机实现视频压缩的有效方法。

3.2 智能视频分析

智能摄像机在 DSP 上执行智能视频分析算法,实现自动视频分析的功能。智能视频分析算法主要包括运动目标检测、目标分类、目标跟踪、行为分析等四个方面^[12],如图 2 所示。其中,运动目标检测、目标跟踪属于视觉底层处理部分,而目标分类、行为分析则属于中层和高层处理。

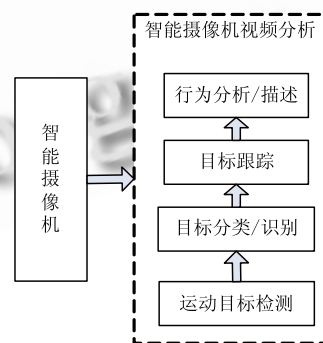


图 2 智能视频分析研究内容

3.2.1 运动目标检测

运动目标检测是从图像序列中将变化区域从背景图像中提取出来。运动目标的准确检测和分割对于目标分类和行为识别等后期处理非常重要。目前检测运动目标最常用的方法主要有背景减除法、帧差法和光流^[13,14]。

背景减除法(Back Ground Subtraction)是目前运动检测中的主流方法,利用当前图象与背景图象的差分

来检测运动区域。背景减除法的基本思想是:首先,用事先获取背景图像序列为每个象素进行统计建模,得到背景模型;其次,将当前每一帧图像和背景模型相减,计算出在一定阈值 T 限制下当前图像中出现的偏离背景模型值较大的那些象素,超出阈值的象素被认为出现在运动目标上。

帧差法是将前后两帧图像对应像素值相减,在环境亮度变化不大的情况下,如果对应像素值相差很小,认为此处物体是静止的,如果某处的像素值变化很大,则认为这是由于图像中运动物体引起的,将这些区域标记下来,利用这些标记的像素区域,就可以求出运动目标在图像中的位置。另外,当目标有阴影干扰时也要进行特殊处理,文献[21]对此进行了详细的讨论。

光流场的计算最初是由美国学者 Horn 和 Schunch 提出的^[15]。1981年 Horn 等人在相邻图像之间的间隔时间很小且图像中灰度变化也很小的前提下,推导出灰度图像光流场计算的基本等式,奠定了光流场计算的基础。此外还出现了基于块区域的光流场计算方法、基于随机平滑度约束条件的贝叶斯光流场分析法、傅立叶自适应平滑约束方法等。

3.2.2 目标分类

目标分类的目的是从检测到的运动区域中将特定类型物体的运动区域提取出来。实现运动目标分类方法主要依据所使用运动目标图像特征来区分,这些常用的特征包括外观特征和运动特征。如利用 Haar 小波函数进行行人正面图像的探测,利用包络椭圆的方向和轴长度的比值等形状信息进行机动车和行人的分类,利用周长和面积的比值判断行人和机动车,利用轮廓上各点到目标物体重心的距离统计区别复杂轮廓目标(行人)和简单轮廓目标(机动车)^[16,17]。除了选择合适的图像特征之外,选择合适的分类方法也是实现运动目标分类的关键,目前主要使用的是基于样本训练的方法包括神经网络、Adaboost 算法和支撑向量机(SVM)^[18,19]等。

3.2.3 目标跟踪

目标跟踪主要用于确定某个特定目标在观测图像上的位置坐标(包括 2D 模型和 3D 模型上的),根据是否使用运动目标检测结果可以分为两大类,即在运动目标检测的结果上的目标定位和直接在图像帧上的目标跟踪。直接图像上的目标跟踪方法主要基于相关性的统计方法,包括卡尔曼滤波(Kalman filtering)、Condensation 算法,均值漂移算法(Mean shift)^[20-22]。另外基于区域匹配跟踪、基于轮廓匹配跟踪、特征匹配跟踪和基于运动特性的跟踪是目标跟踪的常见方法。

3.2.4 行为分析

前面了解了目标检测、目标分类和目标跟踪,但计算机视觉的最终目标是为了行为分析。它是在对视频图像序列进行低级处理的基础上,进一步研究图像中各目标的性质和它们之间的相互联系,并得出对图像内容含义的理解以及对客观场景的解释,从而指导和规划行动。在过去的十年中,大量的行为理解的方法被开发和研究。在文献[22]贝叶斯网络用于识别行为"sitting",文献[17]采用有限状态机来表达和识别行为。

智能视频分析技术的核心是由各种算法组成,由于应用环境的不同,对算法的复杂性要求也不同,这就需要具有开放度高、兼容性好及操作方便的嵌入式平台与之适应。文献[12]中 Nikhil Gagvani 给出了智能视频分析在嵌入式处理器上的实现。随着视频分析方案和算法的成熟,视频分析已经逐渐从 PC 平台转移到实时嵌入式处理器上实现。

4 展望

智能摄像机平台能够有效地完成计算机视觉实时分析任务,并且能够通过网络进行协作和数据融合,是实现智能视频监控的理想平台,具有安装容易、实施简单以及可扩充性好等优点。根据很多行业的差异化需求,如银行、机场、零售业等,针对各行业的特点,推出基于场景内容分析的智能摄像机,因其能适应复杂、多变的场景而得到了一致的认可。毋庸置疑,智能摄像机具有广阔的市场前景。智能摄像机有以下值得关注的趋势:

(1) 多摄像机协同工作(Multi-Camera)

智能摄像机的协同工作是指使用多个视频输入源及使用不同的摄像机参数配置和安装方式,使每个摄像机在最合适的光源和物理安装条件下,执行最有利的一个或者多个任务,并且和同一场景或者跨越多个场景的摄像机群联动,完成智能化的系统要求。这样符合需要多摄像机的广阔场景的应用,能实现实时多任务系统,并实现多角度智能配合。

(2) 视频传感器网络(Video Sensor Network)^[23]

智能摄像机可以作为视频传感器构成智能视频传感器网络,作为传感器网络的高级形式,可以扩展人们对大信息量媒体的获取和处理能力,将环境中的视觉、声音等多种媒体信息有效的引入环境检测中来,实现全面而有效的环境监测,采集的数据通过 Internet 网络或通信卫星到达控制中心。

(3) 移动式监控系统

随着国家大力推进 3G 网络建设,智能摄像机可

以通过无线的方式直接接入无线网络,将视频图像传送到远端的智能手机上,用户可以通过手机直接监视画面。为了适应方便、灵活的智能监控,将智能摄像头(Smart Cam)植入智能手机成为可能,让智能手机的摄像头变为智能摄像机的应用程序。智能手机可以通过 WiFi 或者蓝牙与电脑连接,同样可以完成智能视频监控的功能。

参考文献

- 1 Hampapur A, Connell J, Pankanti S, Senior A. Smart Surveillance: Applications, Technologies and Implications. IEEE Pacific-Rim Conference on Multimedia, Singapore, 2003.1133-1138.
- 2 Wolf W, Ozer B, Lv T. Smart Cameras as Embedded Systems. IEEE Computer, 2002,35(9):48-53.
- 3 Bramberger M, Brunner J, Rinner B, Schwabach H. Real-Time Video Analysis on an Embedded Smart Camera for Traffic Surveillance. Proc. of the 10th IEEE Real-Time and Embedded Technology and Applications Symposium. 2004. 174-181.
- 4 Maier A, Rinner B, Schwabach H, Trathnigg T. Combined Dynamic Power- and QoS- Management in Embedded Video Surveillance Systems. Proc. of the Second Workshop on Intelligent Solutions in Embedded Systems, 2004.63-77.
- 5 Kahar NF, Ahmad RB, Hussin Z, Rosli ANC. Embedded Smart Camera Performance Analysis. International Conference on Computer Engineering and Technology. Singapore, 2009.79-83.
- 6 Quaritsch M, Kreuzthaler M, Rinner B, Bischof H, Strobl B. Autonomous Multi-camera Tracking on Embedded Smart Cameras. EURASIP Journal on Embedded Systems, 2007. 1-10.
- 7 Liu Y, Li RF, Xu C, Yu F. Design and Implementation of Embedded Multimedia Surveillance System. 1st International Workshop on Knowledge Discovery and Data Mining. Adelaide. 2008.570-573.
- 8 Bramberger M, Pflugfelder RP, Maier A, Rinner B, Strobl B, Schwabach H. A smart camera for traffic surveillance. Proc. of the 1st Workshop on Intelligent Solutions in Embedded Systems. 2003.153-164.
- 9 金然.基于 3G 标准、技术体制的比较.电子科技大学学报, 2005,7(2):85-87.
- 10 孙中伟,张福炎.自动视频分析综述.计算机科学,2002, 29(5):80-84.
- 11 Bramberger M, Doblander A, Maier A, Rinner B. Distributed Embedded Smart Cameras for Surveillance Applications. IEEE Computer Society, 2006,39(2):68-75.
- 12 Kisananin B, Shuvra S, Chai B. Embedded Computer Vision. London: Springer-Verlag, 2009.219-255.
- 13 Shotton J, Blake A, Cipolla R. Contour-Based Learning for Object Detection. Proc. of the 10th IEEE International Conference on Computer Vision. 2005.503-510.
- 14 王亮,胡卫明,谭铁牛.人运动的视觉分析综述.计算机学报,2002,25(3):225-237.
- 15 Horn B, Schunch B. Determining optical flow. Artificial Intelligence, 1981,17(1-3):185-203.
- 16 Rivlin E, Rudzsky M, Goldenberg R. A real-time system for classification of moving objects. Proc of the 16th International Conference on Pattern Recognition, Quebec City: IEEE Computer Society, 2002.688-691.
- 17 Ayers D, Shah M. Monitoring human behavior from video take in an office environment. Image and Vision Computing, 2001,19(12):833-846.
- 18 Fan RE, Chen PH, Lin CJ. Working set selection using second order information for training SVM. Journal of Machine Learning Research, 2005,6(12):1890-1918.
- 19 Xu LQ, Landabaso JL, Lei B. Segmentation and Tracking of Multiple Moving Objects for Intelligent Video Analysis. BT Technology Journal, 2004,22(3):140-150.
- 20 Comaniciu D, Ramesh V, Meer P. Real-Time Tracking of Non-Rigid Objects using Mean Shift. IEEE Conference on Computer Vision and Pattern Recognition. Hilton Head Island, 2000.142-149.
- 21 Paul L. Rosin, T. Ellis. Image difference threshold strategies and shadow detection. Proc. of the British Conference on Machine Vision. 1995.347-356.
- 22 Madabhushi A, Aggarwal J. A Bayesian approach to human activity recognition. 2nd IEEE International Workshop on Visual Surveillance, 1999.25-30.
- 23 Liu Y, Das K. Information-Intensive Wireless Sensor Networks: Potential and Challenges. IEEE Communications Magazine, 2006,44(11):142-147.