

浙 江 大 学

硕士学位论文文献综述

(专业学位)

文献综述题目： 智能家居监控系统

姓 名： 刘双妹

学 号： 21451196

专 业： 软件工程

院 别： 软件学院

导 师： 张泉方

日期：二零一五年 11月

目录

一、智能视频监控概述.....	1
二、视频监控的发展.....	2
三、智能视频监控的关键技术.....	4
3.1 高效大容量的视频数据存储.....	4
3.2 监控场景运动检测技术.....	4
3.3 监控场景中人脸识别方案.....	5
参考文献.....	7

一、智能视频监控概述

智能视频监控（IVS: Intelligent Video Surveillance）是计算机视觉领域中近年来新兴起的一个应用方向，是融合了图像处理、计算机视觉、自动控制和计算机等若干领域先进技术，利用计算机视觉技术对视频信号进行处理、分析和理解，并对视频监控系统进行控制，从而提高视频监控系统智能化水平。智能视频监控系统在民用和军事领域中都有着极大的应用前景。目前虽然在银行、商店、车站、港口等一些重要的公共场所普遍架设了监控摄像机，但实际的监控任务仍需要较多的人工工作来完成。在很多情况下，目前的视频监控系统所提供的信息是没有经过任何分析的视频裸数据，这就不能充分发挥监控系统应有的实时主动的监督作用。另一方面，为了防止和阻止犯罪，对无人值守的视频监控系统的需求量日益上升，这类系统的主要目标是减少对繁琐人工的依赖，自动完成对复杂环境中人和车辆等进行实时观测以及对感兴趣的对象的行为进行分析和描述^[1]。

智能视频监控技术主要包括对视频图像序列自动地进行运动对象的提取、描述、跟踪、识别和行为分析等方面的内容。通俗来讲，智能视频监控技术就是借助计算机强大的数据处理功能，对视频画面中的海量数据进行高速分析，过滤掉监控者不关心的信息，仅仅为监控者提供有用的关键信息。智能视频监控以数字化、网络化视频监控为基础，但又有别于一般的网络化视频监控，它是一种更高端的视频监控应用。智能视频监控系统能够识别不同的物体，发现监控画面中的异常情况，并能够以最快和最佳的方式发出警报和提供有用信息，从而能够更加有效的协助安全人员处理危机，并最大限度的降低误报和漏报现象。

智能视频监控系统总结起来主要有如下优势：

1. 全天候可靠监控

智能视频监控系统彻底改变了完全由工作人员对监控画面进行监视和分析的模式，它通过智能视频分析模块由计算机自动对监控的画面进行分析，一旦发现安全威胁立即报警。

2. 响应速度快

智能视频监控系统比普通视频监控具有更强大的智能特性，往往在安全威胁发生的同时能够预警，使得安全部门有足够的时间为所面对的威胁做好相应的准备工作。

3. 提高报警精确度

智能视频监控系统能够有效提高报警精确度，随着监控系统前端设备性能的提高，使得设备可以具有强大的图像处理能力，并可以运行较复杂的智能算法，这样用户可以更加精确的定义安全威胁的特征，从而大大降低误报和漏报。

正是由于智能监控系统优势如此明显，使得其应用范围很广，作为一种高端的视频监控技术，智能视频监控不仅能够成为应对恐怖主义袭击和处理突发事件的有力辅助工具，而且还可以应用在交通管理、客户行为分析、客户服务等多种非安全相关的场景。

目前根据智能算法的研究进展，智能视频监控潜在的应用有^[2]：

1. 多环境下运动目标检测：在复杂的天气环境中(例如：雨雪、大雾、大风等)精确的检测和识别单个物体或多个物体的运动情况。

2. 目标跟踪：检测到运动目标后，根据运动目标的运动情况，系统可以跟踪目标，较难的情况是多摄像头下的目标跟踪，这样可以对可疑人物进行自动跟踪，监控其行为。

3. 目标分类：分类可以有多种，可以进行人车分类，进行颜色分类，多人和单人分类等。

4. 人脸识别：自动识别人物的脸部特征，并通过与已存的人脸模型进行比较来识别或验证人物身份。

5. 密度估计：可以根据场景进行人群或者车辆密度估计来估计场景中人群或车辆拥挤程度，在广场或道路高峰期有很好的拥挤预警作用。

6. 注意力控制：统计人们在某一物体前面的停留时间，可以评估新产品或者新促销策略的吸引力。

二、视频监控的发展

智能视频监控是在传统的视频监控基础上发展起来的。传统的视频监控系统投入实际应用已经数十年，其技术发展也随之经历了三个时代^[3]。

第一代：模拟时代

在 20 世纪 90 年代初以前，视频监控主要是以模拟式磁带录像机(VCR: Video Cassette Recorder)为代表，该系统主要由模拟摄像机、专用电缆、视频切换矩阵、模拟监视器、模拟录像设备和盒式录像带等构成。其缺点是无法进行远程访问、无法与其他安防系统(如门禁、边界防护等)有效集成，信息存储方式的臃肿也给检索和查询带来众多不便。

第二代：数字时代

在上世纪 90 年代中期，得益于数字视频压缩编码技术的发展，数字式视频录像机(DVR: Digital Video Recorder)开始出现。DVR 的使用让用户可以将模拟的视频信号数字化，并存储在电脑硬盘而不是盒式录像带上。数字化的存储大大提高了用户对视频信息的处理能力。此外对于报警事件，以及事前和事后报警信息的搜索也变得异常简单。

第三代：网络时代

进入 21 世纪后，随着网络技术的发展，DVR 系统又进一步发展成为网络数字视频录像机(NVR: Network DVR)系统。与 DVR 系统相比，NVR 系统不但实现了视频信息的数字化存储，还实现了视频档案信息的数字化传播，即 NVR 可以直接接入到 IP 网络中，从而使存储下来的视频信息可以通过网络方便的进行共享。进一步，网络化视频监视系统，又称为 IP 监视系统(IPVS: IP Video Surveillance)开始出现。网络化视频监视系统从一开始就是针对在网络环境下使用而设计的，因此它克服了 DVR/NVR 无法通过网络获取视频信息的缺点，用户可以通过网络中的任何一台电脑来观看、录制和管理实时的视频信息。第三代视频监控系统是完全数字化的系统，它基于标准的 TCP/IP 协议，能够通过局域网/无线网/互联网传输，布控区域大大超过了前两代系统；它采用开放式架构，可与门禁、报警、巡更、语音、管理信息(MIS: Management Information System)等系统无缝集成；它基于嵌入式技术，性能稳定，无需专人管理；其灵活性大大提高，监控场景可以实现任意组合，任意调用。

摄像头的增加带来了大规模防范的可能，即可以获取海量的视频数据用于实

时报警和事后查询。但是对以人为主的使用对象而言,大规模视频数据也带来巨大的挑战。美国圣地亚国家实验室专门做了一项研究,结果表明,人在盯着视频画面仅仅 22min 之后,人眼将对视频画面里 95% 以上的活动信息视而不见。基于以上需求,智能视频监控系统应运而生,其中最核心的部分是基于计算机视觉的视频内容理解技术,通过对原始视频图像经过背景建模、目标检测与识别、目标跟踪等一系列算法,进而分析其中的目标行为以及事件,然后按照预先设定的安全规则,及时发出预警信号。

国际上,计算机视觉领域中的权威期刊“国际计算机视觉期刊(IJCV, International Journal of Computer Vision)”和“IEEE 模式分析和机器智能汇刊(PAMI, IEEE Transaction on Pattern Analysis and Machine Intelligence)”相继在 2000 年 6 月和 2000 年 8 月出版了关于视频监控的专刊^[4]。美国计算机学会(ACM, Association for Computing Machinery)连续几年都有大量的文章在视频监控的智能化、网络化和系统化方面做出研究。英国雷丁大学计算机系的 VIEWS 项目组开展了对车辆和行人的跟踪及其交互作用识别的相关研究,这一技术可用于视频监视系统中的危险识别^[5]。IBM 公司于 2006 年 11 月 7 日宣布已开发并销售一款用于分析视频实时监控系统的软件——智能监控系统(S3:Smart Surveillance System),它能够将视频摄像头捕捉到的信息通过计算机网络传递到整个系统软件上,发现监控环境中的潜在安全隐患并自动报警^[6]。由 Steve J.Maybank 和谭铁牛组织的 IEEE 视觉监控专题讨论会(VS,IEEE International Workshop on Visual Surveillance)也已经成功地举办了三届。美国国防高级研究计划局(DARPA)资助卡内基梅隆、戴维 SARNOFF 研究中心等著名大学和科研机构,联合研制出了智能场景监视与监控系统 VSAM(Visual Surveillance and Monitoring),用于在未来战争中人力监控费用昂贵、非常危险或者人力无法实现的场合的监控^[7];美国 ISS 公司的研制出的 AUTOSCOPE2004 是一种大区域视频监控系统,已经作为北美铁路运输监视系统实现了应用。

在国内,智能视频监控的研究也已经有了长足的进步,2002 年 5 月,第一届“全国智能视觉监控学术会议”在北京召开,迄今为止,已经举行了三届,得到了广泛的关注。此外,中国科学院自动化研究所模式识别国家重点实验室已经成立了智能视频监控研究小组,正在开展智能视频监控方面的研究,研究内容包括:快速准确的运动检测,实时性、鲁棒性的基于三维模型的车辆与行人的定位、识别和跟踪,基于移动摄像机的视觉监控技术,多摄像机的协作监控,事件的机器学习方法,异常现象的检测、报警与目标的行为预测,对目标运动情况给出语义上的解释的方法以及远距离的身份识别等等^[8]。目前视觉监控研究组针对目前交通管制中大量人力和物力资源的浪费,提出了一种基于三维模型的交通监控系统,采用了基于三维线框模型的车辆定位算法和基于改进的扩展卡尔曼滤波的车辆跟踪算法;研究组在自主研发的交通场景模拟平台上最终实现了交通监控原型系统等。北京大学视觉与听觉信息处理国家重点实验室主要是实现高度智能化的机器感知系统,并在言语听觉技术、三维视觉信息处理、智能机器人等研究领域取得许多研究成果^[9]。清华大学智能技术与系统国家重点实验室主要从事人工智能基本原理、基本方法的研究,包括智能信息处理、机器学习、智能控制,以及神经网络理论等,还从事与人工智能有关的应用技术和系统集成技术的研究,主要有智能机器人、声音、图像、文字及语言处理等。另外,浙江大学信电系智能交通研发中心致力于智能交通监控系统的研究工作。

三、智能视频监控的关键技术

3.1 高效大容量的视频数据存储

视频监控系统的重要功能之一就是可以保存大量的视频记录,供日后查询浏览使用,这就需要有一套完整的视频数据存储方案来对监控场景的视频数据进行高效快速的保存。在视频监控系统中,目前有两种方法可以存储大量视频数据,一种是文件存储系统,另一种是数据库技术,而在实际应用中一般是通过文件存储系统来存储的^[10]。

视频监控系统中的存储方式主要有三种方式^[11-14]:

1、完全存储

将所有的连续视频数据都保存下来,适用于易出现突发事情的场合,这种应用最为广泛。

2、流水存储

按事先定义好的方式存储,丢掉了部分帧,这种方式存储的图像一般不连续,空间占用比完全存储小得多,这种存储方式适用于不太重要的场合或者不太重要的时段。

3、报警存储

存储动作由报警启动,并且需要保存报警前一段时间内的图像,当报警解除后则停止存盘,适用于相对比较静止的场合。

此外还有定时存储和手动存储等方式,对于这些不同的存储方式,采用不同的存储机制来实现,以达到最好的性能要求,具体使用的存储方式则根据用户设置来选取。

3.2 监控场景运动检测技术

由于视频监控的监控场景基本已经知道,于是其关注的焦点集中在场景本身的各种异常行为的发生上,这些异常行为多以运动的形式表现出来。运动检测技术即是利用计算机自动检测场景中活动的人和物体,并实时产生可靠的报警信息,从而提高监控系统工作效率^[15]。

现有的运动检测算法适用于不同的范围,通常根据不同需要来确定检测方法。目前几种常用的方法有:帧间差分法、背景差分法、光流法和块匹配法^[16-23]。

帧间差分法其基本原理就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,可以认为此处景物是静止的,如果图像区域某处的像素值变化很大,可以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。帧间差分法对动态变化场景有较强的适应性,但一般不能很好的提取出所有的特征像素点。

背景差分法是通过背景建模,利用相邻序列图像估计视频中不变的背景,进而提取当前帧中的非背景像素集合(即运动目标)。其基本思想是输入图像与背景

模型进行比较,通过判定灰度等特征的变化,或用直方图等统计信息的变化来判断异常情况的发生和分割运动目标。背景差分法一般能得到比较全面的特征数据,但对动态场景变化和光线变化等外界干扰非常敏感。

光流法可以分为连续光流法和特征光流法。连续光流法一般采用基于帧间图像强度守恒的梯度算法,其中最为经典的算法是 L- K(Lucas-Kanade)法和 H-S(Horn- Schunck)法。特征光流法是通过特征匹配求得特征点处的光流,可以采用图像边缘和角点作为特征点。光流法不需要预先知道场景的信息,并在摄像头运动时能有效的检测目标,但是该算法计算方法又过于复杂,且抗干扰性差,不适合用于实时性要求很强的场合。

块匹配方法的基本思想为:图像序列的每一帧被划分成若干个相同的子块,为了在 $k+1$ 帧中找出第 k 帧中第 j 个子块的最佳匹配块,在 $k+1$ 帧中某一给定区域内,根据一定的匹配准则进行搜索,第 j 块与其匹配块之间的相对位移即为该块的匹配运动向量。这种方法与光流计算不同,无需计算每一个像素的运动,而只是计算由若干像素组成的像素块的运动,这一点尤其适合应用在视频监控系统中。

3.3 监控场景中人脸识别方案

随着视频监控系统的普及和广泛应用,用户对其有了更多及更高的要求,近年来国内外也开始了新一代视频监控系统的研究及开发,而新一代的视频系统最主要的发展方向是在传统视频监控系统中加入带有人工智能的自动影像捕捉及识别功能,使固有的视频监控系统从被动的监控录像系统提升至实时自动监视及主动报警系统,可有效的解决现今简易传感器的误报警问题,大大提高了视频监控系统的应用领域^[24-27]。在各种自动影像捕捉及识别方法中,人脸识别是最重要的技术。人脸识别技术是使用人类面部特征作为识别个人身份的一种方式,它可将某特定个人的面部特征资料储存在数据库里,作为识别的依据^[28]。

人脸识别包括两个技术环节:人脸检测、人脸识别^[29-30]。人脸检测的目的是确定图像中人脸的位置、大小和数量;人脸识别是对检测到的人脸进行特征提取、模式匹配与识别。人脸识别系统主要有三个步骤:人脸的检测和定位、人脸的特征提取和人脸的识别,在这些步骤之前还应有预处理这一步,即对采集到的图像先进行预处理,以达到位置校准和灰度归一的目的,然后寻找人脸,如果有则确定人脸的位置并提取人脸,然后提取人脸特征,最后根据提取的特征进行识别^[31-32]。

1.人脸检测

人脸检测是指在动态的场景与复杂的背景中判断是否存在人脸,并分离出这种人脸。人脸检测首先检测确定人脸五官和人脸外轮廓,然后精确定位出眼睛位置,通过识别算法对人脸生成模板,通过分析同一个人在不同条件下人脸图像的差异(类内差异)以及不同的人脸之间的差异(类间差异),进而生成一个变换矩阵。在这个过程中提取人脸特征时需要注意的是减少类内差异、扩大类间差异,让提取的人脸特征利于分类。然后根据提取的人脸特征,采用识别算法计算两幅人脸图像的相似度。一般有下列几种方法:

(1) 参考模板法

首先设计一个或数个标准人脸的模板,然后计算测试采集的样品与标准模板之间的匹配程度,并通过阈值来判断是否存在人脸。

(2) 人脸规则法

由于人脸具有一定的结构分布特征,所谓人脸规则的方法即提取这些特征生成相应的规则以判断测试样品是否包含人脸。

(3) 样品学习法

这种方法即采用模式识别中人工神经网络的方法,即通过对人脸样品集和非人脸样品集的学习产生分类器。

(4) 肤色模型法

这种方法是依据面貌肤色在色彩空间中分布相对集中的规律来进行检测。

(5) 特征子脸法

这种方法是将所有人脸集合视为一个人脸子空间,并基于检测样品与其在子空间的投影之间的距离判断是否存在人脸。

值得提出的是,上述 5 种方法在实际检测系统中也可综合采用。

2、人脸比对

人脸比对是对被检测到的人脸进行身份确认或在人脸库中进行目标搜索。这实际上就是说,将采样到的人脸与库存的人脸依次进行比对,并找出最佳的匹配对象。所以,面像的描述决定了面像识别的具体方法与性能。目前主要采用特征向量与面纹模板两种描述方法。

(1) 特征向量法

该方法是先确定眼虹膜、鼻翼、嘴角等面像五官轮廓的大小、位置、距离等属性,然后再计算出它们的几何特征量,而这些特征量形成一个描述该面像的特征向量。

(2) 面纹模板法

该方法是在库中存贮若干标准面像模板或面像器官模板,在进行比对时,将采样面像所有像素与库中所有模板采用归一化的相关量度量进行匹配。此外,还有采用模式识别的自相关网络或特征与模板相结合的方法。

参考文献

- [1] 黄凯奇,陈晓棠,康运锋,谭铁牛.智能视频监控技术综述[J].计算机学报,2015,06:1093-1118.
- [2] 李立仁,李少军,刘忠领.智能视频监控技术综述[J].中国安防,2009,10:90-95.
- [3] 杨建全,梁华,王成友.视频监控技术的发展与现状[J].现代电子技术,2006,21:84-87.
- [4] 王晨.智能视频监控设计[D].南京:南京理工大学,2007.
- [5] 张海清.智能视频监控中的运动目标检测与跟踪[D].合肥:中国科学技术大学,2007.
- [6] 陈轶博.智能视频监控的设计与实现[D].大连:大连海事大学,2008.
- [7] Collins R,et al.A system for video surveillance and monitoring[J].Carnegie Mellon University's Technical Report,2000.
- [8] 谭铁牛.智能视觉监控技术概述[J].第一届全国智能视觉监控学术会议,2002.
- [9] 朱守刚.智能视频监控技术研究及应用[J].数字技术与应用,2012,01:26+28.
- [10] 肖沁雨.智能视频监控关键技术分析[J].制造业自动化,2012,12:21-23.
- [11] 李立仁,李少军,刘忠领.智能视频监控技术综述[J].中国安防,2009,10:90-95.
- [12] 韩国强.浅谈智能视频监控技术及其主要应用[J].计算机与网络,2014,02:62-65.
- [13] 范金刚.智能视频监控系统若干关键技术的研究与实现[D].成都:电子科技大学,2006.
- [14] 程向娇,黄言恣.无线视频监控关键技术研究及开发[J].中国科技信息,2010,16:155-156.
- [15] 林结斌.智能视频监控中的运动目标检测研究[D].广州:华南师范大学,2009,35-39.
- [16] 姚敏.数字图像处理[M].北京:机械工业出版社,2006,225-264.
- [17] 李毅,孙正兴,远博等.一种改进的帧差和背景减相结合的运动检测方法[J].中国图像图形学报,2009,14(6),1162-1168.
- [18] 鱼亚锋.运动目标检测和智能视频监控系统设计[D].北京:北京邮电大学,2008,20-27.
- [19] 耿淼.视频监控系统中的运动目标检测研究[D].无锡:江南大学,2008,7-10.
- [20] 李德禄.异常运动目标的智能视频监控技术研究[D].广州:华南理工大学,2008,22-27.
- [21] 李敏.视频监控系统中运动目标检测方法研究[D].西安:西北大学,2008,45-58.
- [22] 钟庆.智能视频监控中的运动检测算法研究[D].合肥:中国科学技术大学,2004,25-30.
- [23] 方颖.运动目标的检测、定位与跟踪研究[D].济南,山东大学,2008,12-27.
- [24] 黄家荣.一种基于人脸识别技术的智能监控系统设计[J].煤炭技术,2010,03:64-67.
- [25] 程向娇,黄言恣.无线视频监控关键技术研究及开发[J].中国科技信息,2010,16:155-156.
- [26] 李彬,曲寒冰,靳薇.浅谈人脸识别技术在智能视频监控中的应用与发展趋势[J].中国安防,2011,03:50-53.
- [27] 马海兵,白洁.人脸识别技术在智能视频监控系统中的应用[J].现代电子技术

术,2007,20:125-128.

[28]李建勇.人脸识别技术在视频监控系统中的应用[J].中国安防,2009,04:48-50.

[29]李尚国.基于人脸识别技术的智能监控系统设计方案[J].成都大学学报(自然科学版),2009,01:56-58+62.

[30]桑海峰,吴丹阳,王会.视频监控下的人脸跟踪与识别系统[J].计算机工程与应用,2014,12:175-179.

[31]邓楠.视频监控系统与人脸识别技术的结合应用[J].计算机光盘软件与应用,2012,19:149-150.

[32]谢仪,鲍可进.智能视频监控中人脸检测的研究与实现[J].计算机测量与控制,2013,11:2921-2923.