

学校编码: 10384  
学号: 200228031

分类号 \_\_\_\_\_ 密级 \_\_\_\_\_  
UDC \_\_\_\_\_

厦 门 大 学  
硕 士 学 位 论 文

视频监控中的运动目标检测技术研究

Moving Object Detection Research in video  
surveillance system

王 华 伟

指导教师: 李翠华 教授/博士生导师  
专业名称: 计算机应用技术  
论文提交日期: 2005年5月 日  
论文答辩日期: 2005年6月 日  
学位授予日期: 2004年 月 日

答辩委员会主席: \_\_\_\_\_

评 阅 人: \_\_\_\_\_

二〇〇五年四月

厦门大学博硕士学位论文摘要库

## 摘要

随着网络、通信和微电子技术的迅速发展和人民物质生活水平的不断提高,视频监控以其直观、方便和内容丰富等特点,日益受到整个社会的重视。新一代视频监控系统以网络为依托,以数字视频的压缩、传输、存储和播放为核心,以智能图像处理和分析为特色。近年来视频监控广泛的应用前景及巨大的潜在经济价值,受到了学术界、产业界和管理部门的高度重视。对视频监控系统及其关键技术的研究具有重要的学术价值和实用价值。

视频监控最基础、也是最重要的工作就是检测场景中的运动目标。运动目标的有效分割对于目标分类、跟踪和行为理解等后期处理是非常重要的。然而,由于背景图像的动态变化,天气、光照、影子及混乱干扰等的影响,运动目标检测成为一项有相当难度的工作。本文针对视频监控系统中的关键技术——运动目标检测进行了研究。

对静止平台,即摄像机固定情况,本文实现了一个能有效提取视频场景中运动目标的检测系统。研究了两种新颖的 RGB 空间内的背景建模方法:掩码图像建模法和类 Hausdorff 建模法,经过性能比较我们选取后者作为本系统的背景建模方法。然后利用背景差分法获得了输入帧的前景区域,并将之分类为运动目标、静止目标、虚假目标、背景、噪声,最后利用基于目标的方法对背景进行更新。该系统建立了可靠的背景模型,背景更新速度快,而且解决了“虚假目标”带来的背景更新“死锁”问题。

检测到的运动目标往往带有阴影。阴影与目标可能相连,也可能分离。两者相连时,影子扭曲了目标的轮廓,使得以后基于形状的认识方法不再可靠;两者分离时,影子有可能被误认为目标。因此,阴影的存在对视频监控系统的有效处理产生严重威胁。本文从阴影的物理原理出发,提出了一种基于 HSV 颜色信息和一阶梯度信息的混合阴影剪除算法。HSV 颜色空间真实地反映了人类观察彩色的方式,更精确的反映一些灰度信息和色彩信息;由于阴影内部灰度值分布的连续性,因此可以用梯度的方法去除阴影,但不能处理阴影内部存在纹理的情况,且求梯度运算也会破坏目标的完整性。结合两者的优缺点,提出了一种新颖的混合阴影剪除算法。与文献[49]中阴影剪除算法相比,该方法对阴影区分率略有改进,对阴影检测率有较大提高。

基于视频的交通监控是视频监控的一个重要内容。针对移动平台,本文研究了目前现有的车辆检测算法,提出了一个基于阴影检测的实时车辆检测系统。首先在场景中检测道路,将其颜色分布拟合为高斯分布,然后利用该分布在 HSV 颜色空间中检测阴影区,确定候选目标区域。最后利用纹理信息剔除虚假目标。实验结果表明,本系统检测性能可靠,能够实现实

时检测。

**关键词：**视频监控；运动检测；阴影剪除

厦门大学博硕士论文摘要库

## Abstract

With the rapid development of the technology of network, communication and micro-electronics, video surveillance system has been laid store by the whole society in recent years for its frank, convenience and abundant content. The new generation of video surveillance system is supported by network, while its kernel is the compression, transmission, storage and play of digital video, and its feature is image procession and analyse. More and more academe, industry and administrators has paid attention to video surveillance system for its extensive appliance feature and huge potential economic value.

Detecting moving object in scenes is one of the most basic and important task in video surveillance system. The effective segmentation of moving object is very important to the following process including object tracking and behavior understanding. However, due to dynamic change of background and the influence of weather, illumination, shadow, moving object detection is a quite difficult task. This paper does some research in the field of moving object detection which is the key technology of video surveillance system.

For static camera situation, this paper implements a moving object diction system. First, two novel background modeling methods in RGB color space are studied: the one is Mask image background modeling method; the other is like-Hausdorff background modeling method. Comparing the two methods, this system chooses the latter which is more robust and effective. Then the foreground area is obtained by background subtraction, and detection system classifies it to moving object, still object, false object, background and noise. Finally, updating background model using an object based algorithm. This system builds robust background model, which updates background rapidly and resolves the "lock" problem caused by false object.

Moving object always has shadow with it. Shadow sometimes connects with object, sometimes separates with object. When shadow is joined with object, the contour of object will be twisted by the shadow. While shadow is apart with object, the shadow will be likely be detected as moving object incorrectly. Thereby, effectiveness of video surveillance system is ruined by shadow. This paper gives a hybrid shadow extermination algorithm based on HSV color space and first order grads information. HSV color space truly reflects observing mode of human beings. Due to continuity of gray distribution in shadow, grads information is used to detect shadow. But grads method cannot handle the texture inside shadow, and will destroy the integrality of object. Unite this two methods, this paper gives a hybrid shadow extermination algorithm. Comparing with other common algorithm, this method improves shadow distinction slightly, and enhances detection obviously.

Another important content of video surveillance system is video based traffic monitor. To moving camera, this paper studied vehicle detection algorithm at the present time, and gives a shadow based detection system. First,

detecting road zone in the scenes, and put its color distribution into Gauss Model. Then, system detects shadow zone using the distribution, and determinates candidate ROI. Finally, system eliminates false objects using texture information. Experiment shows this system detect vehicles robustly, and can reach real time detection.

**Keywords:** Video Surveillance System; Moving Object Detection; Shadow Extermination.

厦门大学博硕士论文摘要库

摘要 .....	i
Abstract .....	3
目 录 .....	5
第一章 绪论 .....	1
1. 1 引言 .....	1
1. 2 视频监控研究的发展历程 .....	2
1. 3 视频监控系统的现状 .....	3
1. 4 小结及本文结构 .....	5
第二章 运动目标检测技术 .....	7
2. 1 引言 .....	7
2. 2 光流法 (Optical Flow) .....	7
2. 3 时间差分法 (Temporal Difference) .....	8
2. 4 背景差分法 (Background Substraction) .....	11
2. 5 其它方法 .....	14
2. 6 小结 .....	17
第三章 静止背景下的运动检测系统 .....	19
3. 1 引言 .....	19
3. 2 运动目标检测处理流程 .....	20
3. 3 背景建模 .....	21
3. 3. 1 掩码图像建模方法 .....	21
3. 3. 2 类 Hausdorff 距离建模方法 .....	23
3. 4 前景目标分类 .....	26
3. 5 阴影剪除 .....	29

3. 6 背景的更新 .....	34
3. 7 实验结果及分析 .....	35
3. 8 小结 .....	38
<b>第四章 移动平台上的车辆检测系统</b> .....	<b>39</b>
4. 1 引言 .....	39
4. 2 检测处理流程 .....	40
4. 3 道路检测 .....	41
4. 4 阴影区检测 .....	43
4. 5 生成候选 ROI .....	44
4. 6 去除虚假目标 .....	47
4. 7 实验结果分析 .....	49
4. 7. 1 算法性能分析 .....	49
4. 7. 2 系统效率分析 .....	59
4. 8 小结 .....	59
<b>第五章 总结</b> .....	<b>61</b>
5. 1 本文工作总结 .....	61
5. 2 后续研究方向 .....	62
<b>参考文献</b> .....	<b>67</b>
<b>研究生期间的研究成果</b> .....	<b>75</b>
<b>致谢</b> .....	<b>77</b>

## 第一章 绪论

### 1.1 引言

常言道“百闻不如一见”，在人类感官接受的各种信息中约有80%来自视觉。图象、视频是对客观事物形象、生动的描述，是直观而具体的信息表达形式，是人类最重要的信息载体，特别是在今天的信息社会，随着网络、通信和微电子技术的快速发展和人民物质生活水平的不断提高，视频以其直观、方便和内容丰富等特点，日益受到人们的青睐。当前，视频监控系统由传统的模拟方式步入了数字式的网络时代。与模拟系统相比，数字监控系统借助于计算机网络突破了模拟信号传输距离上的限制，监控可以无所不在；更重要的是，数字监控系统以计算机为处理核心，除了能够实现多媒体信息处理如压缩、传输、存储和播放等基本功能之外，还能够实现自动异常报警、智能存储和快速检索等高级功能，实现模拟监控无法实现的真正意义上的“主动监控”功能。

视频监控近年来得到了非常迅速的发展，应用范围不断地扩展。民用系统主要集中在安全、保卫、交通等方面，如银行、宿舍楼、公路、停车场的监控系统。还包括一些企事业单位的智能无卡考勤，某些旅游景点的人流统计和楼宇进出的安全控制等等。

应用先进的监控技术，建立完善的道路交通监控系统，实现城市道路的现代化监控与管理，从而有效地抑制交通事故的发生，保证战时交通道路的畅通，已成为军事部门与各地交通管理部门最为关注的问题之一。通过对主要道路上车辆的监控，可以实时、直观地对可疑目标实现有针对性的检测、跟踪，了解各主要交通要道和交叉路口的车流量、车辆通行状况，还可根据各监控点反馈的信息，预测某些交通要道和交叉路口可能会出现的情况，以便及时调整和疏导交通流，减少交通堵塞，最大限度地求得道路系统的利用率，确保道路的畅通。由此可见，对道路实行交通监控，有着

重要的军事意义与应用价值。

## 1. 2 视频监控研究的发展历程

视频监控系统大致经历了以下三个发展阶段：

(1) 在 20 世纪 90 年代初及其以前，主要是以模拟设备为主的闭路系统，称为第一代视频监控系统，即模拟视频监控系统。这种监控系统多是以摄像机、分割器、录像机为核心，辅以其它传感器的模拟系列。每一监视点的视频信号通过模拟线路传输到中心控制室；采用手动方式进行切换，在电视机构成的监视器上难以同时观察到所有监视点的情况；其存储不但耗费大量的存储介质，查询取证也很烦琐。这种模拟监控系统大多采用模拟电路与中、小规模数字集成电路制成的控制器，功能简单、可靠性差，而且易受干扰，系统寿命也短。

(2) 20 世纪 90 年代中期，随着计算机处理能力的提高和数字视频技术的发展，监控系统也逐渐从模拟方式转向数字处理方式。通过视频采集卡采集视频信号，利用计算机的高速数据处理能力进行视频处理。利用显示器的高分辨率和高清晰度，实现图像的多画面显示，从而大大提高了图像质量。随着数字视频压缩编/解码技术的日益成熟和存储媒介的发展，为视频信号的传输、存储提供了极大的方便，提高了信息的可利用率。多媒体技术的不断发展使得监控不仅能提供各类数据、文本、图形信号，还有视频图像、声音等更加丰富的信息支持生产和管理活动，提供监视、控制、存储、报警、查询、回放等功能，监控的系统性能得到了极大的改善和提高。

(3) 到了 20 世纪 90 年代末，随着网络带宽、计算机处理能力和存储容量的迅速提高，以及各种实用视频信息处理技术的出现，视频监控进入了全数字化的网络时代，称为第三代视频监控系统，即全数字视频监控系统或网络数字视频监控系统。第三代视频监控系统以网络为依

托，以数字视频的压缩、传输、存储和播放为核心，以实用的智能图像分析为特色，引发了视频监控行业的技术革命。

(4) 虽然监控系统功能日益强大，但是依然需要工作人员不间断地分析监视场景内的活动，日夜值守工作量繁重，同时远程视频传输的实时性、流畅性也还有待于进一步的研究，因此计算机视觉和应用研究学者适时提出了新一代监控——视觉监控(Visual Surveillance)的概念<sup>[1]</sup>。视觉监控在不需要人为干预的情况下，利用计算机视觉和视频分析的方法对摄像机拍摄的图像序列进行自动分析，实现对动态场景中目标的定位、识别和跟踪，并在此基础上分析和判断目标的行为，从而做到既能完成日常管理又能在异常情况发生时及时做出反应，同时在远程传输中提供更加丰富、流畅的视频监控效果。视觉监控的概念逐渐为人们所关注，成为视频监控未来的发展方向。

### 1.3 视频监控系统的现状

视频监控以其强大的功能、低廉的成本、灵活的使用方式、广泛的应用前景及巨大的潜在经济价值，受到了学术界、产业界和管理部门的高度重视。

目前，国外比较有代表性的实时系统有：

#### (1) MIT 的 Pfunder 系统<sup>[2]</sup>

这个系统能用来实时地跟踪人，并能分析理解人的行为。通过固定的摄像头，它可以适应于不同类型的场景，但只能处理单个没有被遮挡的人体，并且要求人是站立的。

#### (2) Haritaoglu 的 W<sup>4</sup> 系统<sup>[3]</sup>

Maryland 大学的 Haritaolu 等人开发出了 W<sup>4</sup>(what, where, when, who) 系统，这个系统是一个室外的实时监控系统，是一个比较全面的人体动作监控系统，主要用于处理高质量的非压缩的灰度图像数据。它使用双峰分

布模型，即通过最小、最大强度值和最大时间差分值为场景中每个像素进行统计建模，并且进行周期性的背景更新。

在目标分类方面，它通过分析前景区域的主轴上的水平和垂直投影直方图，从而计算它与数据库中标准的直方图的匹配，并结合人的周期性运动分析来区分人与其它的物体(如车)。在人的跟踪方面，它能同时对多人进行检测与跟踪，当某个特定的人与运动人群分开时，这个系统使用运动模型来跟踪这个特定的人，当这个人融入人群后，只要它的头部是可见的，这个系统仍能准确地跟踪他。它也能通过建立外观模型实现多人的跟踪，并可以检测和分析人是否携带物体、放置物体、交换物体等简单行为。在人的行为理解上，它利用人的拓扑结构及对人的运动分析(如速度、加速度、人的周期性运动等)，定义人的头、躯干、两手、两脚这六个主要部分，从而进行人的标准姿势，如正面图、左视图、右视图上的站、坐、躺、卧的行为识别。

### (3) CMU 的 VASM 系统<sup>[4]</sup>

在 1996 年至 1999 年间，美国国防高级研究计划局(DARPA)资助卡内基梅隆大学、戴维 SARNOFF 研究中心等著名大学和公司合作，联合研制了视频监视与监控系统 VSAM。目标是开发自动视频理解技术，用于实现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等场合的监控。

VASM 系统能根据监控区域的特征建立一个虚拟的场景，在虚拟场景中每个固定目标都有一个对应的虚拟物体。然后系统检测并识别实时输入的视频中的运动物体，再将这些物体放入虚拟场景中，得到场景的真实再现。物体的信息如运动速度、人体动作也同时输出。这个系统能处理多场景视频，摄像头还可以运动、调焦，以得到目标的清晰图像。整体说来，VASM 系统是一个比较成熟的监控系统，但缺陷也不少，主要工作集中在低层次处理运动检测和跟踪、物体识别，很少涉及高层次如轨迹分析、交互事件识别、异常事件处理等。

另外，DARPA在2000年又资助了重大项目HID 计划(Human Identification at a Distance)，研究开发多模式的监控技术以实现远距离情况下的人的检测、分类和识别，以增强国防、民用等场合免受恐怖袭击的保护能力。

#### (4) Dockstader 的物体跟踪和人脸检测系统<sup>[5]</sup>

这个系统可以跟踪中等复杂的场景中的多个人，并确定这些人头部的位置。它能获得所有的物体轮廓，能比较好地处理碰撞问题，但不能得到物体的组成像素点，这使得进一步的处理如物体识别难以实现。这个系统还假设人的头部肯定在检测前景的上方，这也限制了它的用途。

国内也有一些视频监控方面的产品，如黄金眼、行者猫王等，应用于交通控制、监狱管理等方面。相比之下，国外的技术相对来说比较成熟。国内的研究则起步较晚，技术还不够完善，开发出的产品距离智能化还有一定差距，在实际的应用中，受到很多限制，还有待于进一步的完善。

在学术上，IEEE从1998年起资助了国际视觉监控系列会议，至今已分别在印度、美国、爱尔兰召开了三届。国际权威期刊International Journal of Computer Vision也在2000年6月出版了视觉监控专题。国际著名期刊IEEE Trans. On Pattern Analysis and Machine Intelligence已于2000年8月出版了有关视觉监控算法的专集。

### 1. 4 小结及本文结构

本章介绍了视频监控系统在军事、民用方面的主要应用领域，讨论了其发展历程及未来发展趋势，研究了视频监控系统的现状，重点介绍了目前国际上著名的几个系统。

本文共五章，具体结构如下：

第一章：绪论；

第二章：运动目标检测技术，主要介绍了视频监控中目前常用的检测运

动目标的方法，着重介绍了最为常用的背景差分法；

第三章：在静止背景下，提出了一个能有效提取视频场景中运动目标的检测系统。给出一种基于前景目标的背景更新方法，并提出了一种新颖的基于 HSV 颜色信息和一阶梯度信息的混合阴影剪除算法。最后，给出了实验结果，并对阴影剪除算法效率进行了量化评估；

第四章：移动平台上的车辆检测。回顾了现阶段车辆检测的主要技术，提出了一个实时车辆检测系统。最后对实验结果进行了分析；

第五章是结束语，总结了本文的主要工作，并提出了下一步工作的设想。

本课题主要来源于国家自然科学基金项目和福建省自然科学基金项目。

## 第二章 运动目标检测技术

### 2.1 引言

视频序列中的运动目标检测是指通过运动及图像特征分析(如灰度、边缘、纹理等)检测视频序列中是否存在与背景有相对运动的前景目标,如果有则将其从原始图像序列中提取出来。运动目标检测是视频监控的第一步工作,是运动目标跟踪、分类和识别的基础。

如何从图像序列中检测和选择出感兴趣运动目标,长期以来一直是人们研究的热点问题。常用的运动目标检测方法有时间差分法(temporal differencing)、背景差分法(background subtraction)、光流法(optical flow)等。本章将对这些方法逐一介绍。

### 2.2 光流法 (Optical Flow)

当人的眼睛观察运动物体时,物体的景象在人眼的视网膜上形成一系列变化的图像,这一系列连续变化的信息不断“流过”视网膜(即图像平面),好像一种光的“流”,故称之为光流。

光流计算技术是Gibson于1950年提出的。所谓光流是指图像中模式运动的速度,它是一种二维(2D)瞬时速度场,其中2D速度矢量是可见的三维速度矢量在成像平面上的投影。光流法检测运动目标的基本原理是:给图像中的每一个像素点赋予一个速度矢量,这就形成了一个图像运动场,在运动的一个特定时刻,图像上的点与三维物体上的点一一对应,这种对应关系可由投影关系得到,根据各个像素点的速度矢量特征,可以对图像进行动态分析。如果图像中没有运动目标,则光流矢量在整个图像区域是连续变化的,当物体和图像背景存在相对运动时,运动物体所形成的速度矢

量必然和邻域背景速度矢量不同，从而检测出运动物体的位置。

检测示例如图 2.1:

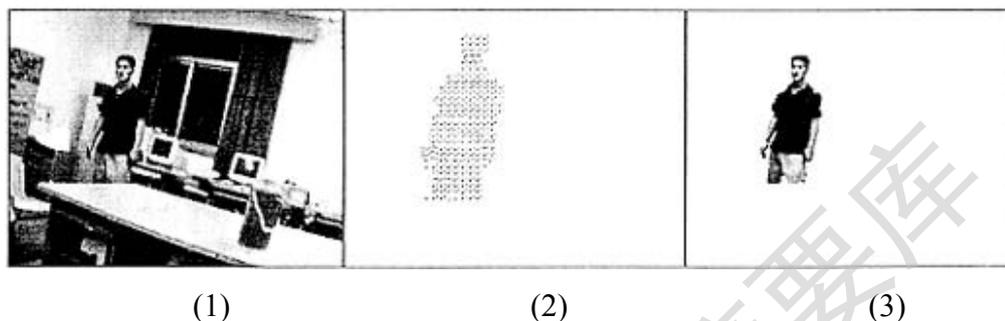


图 2.1 光流法 (1)为当前帧；(2)为光流分析；(3)为检测结果

基于光流方法的运动检测采用了运动目标随时间变化的光流特征，如 Meyer<sup>[6]</sup>等通过计算位移向量光流场来初始化基于轮廓的跟踪算法，有效地提取和跟踪运动目标。该方法的优点是在摄像机运动存在的前提下也能检测出独立的运动目标，然而，大多数光流计算方法相当复杂，计算量大，易受噪声影响，不利于实时处理。并且具有不确定性问题，无法解决孔径问题，抗噪性差，如果没有特别的硬件装置则不能被应用于全帧视频流的实时处理。关于光流法的详细内容可以参考 Barron<sup>[7]</sup>的文章。日本大阪大学的 Ryuzo Okada, Shinya Yamamoto 等人对此方法作了深入的研究，并已研制出了比较成熟的系统<sup>[8-11]</sup>。借助于多个数字信号处理器（DSP），这些系统都实现了实时目标检测与跟踪，处理速度可以达到 15 帧/s。

### 2.3 时间差分法（Temporal Difference）

这种方法的突出特点是实现简单、运算速度快，在大多数情况下检测效果良好<sup>[12]</sup>。基本原理就是将前后两帧图像对应像素点的灰度值相减，在环境亮度变化不大的情况下，如果对应像素灰度相差很小，可以认为此处景物是静止的，如果图像区域某处的灰度变化很大，可以认为这是由于图像

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to [etd@xmu.edu.cn](mailto:etd@xmu.edu.cn) for delivery details.

厦门大学博硕士论文摘要库