*件2：论文范例（绿色斜体字均为说明，论文编辑完毕后请删除绿色斜体文字并保持格式不变 ）*

*（不空行，用学校提供的蓝色打印）*



**题目：横幅视频识别的关键技术**

**学 院 软件学院**

**专 业 软件工程**

**学生姓名 张爱文**

**学生学号 200730555364**

**指导教师 黄翰**

**提交日期 年 月 日**

毕 业 设 计 （论文） 任 务 书

兹发给 班学生 毕业设计（论文）任务书，

内容如下：

1. 毕业设计（论文）题目：

2、应完成的项目：

（1）

（2）

（3）

（4）

3、参考资料以及说明：

（1）

（2）

（3）

（4）

（5）

（6）

（7）

4、本毕业设计（论文）任务书于 年 月 日发出，应于 年 月 日前完成，然后提交毕业考试委员会进行答辩。

指导教师 签发， 年 月 日

教研组（系、研究所）负责人 审核， 年 月 日

# 摘 要

本文详细介绍了从视频中识别运动横幅的原理和具体实现方法等。

视频识别属于机器视觉的领域，它涉及多方面的学科，有着深入广泛的应用。

视频识别的首要任务是识别出其中的运动物体，然后对运动物体进行结构分析，提取特征向量，再应用人工智能，模式识别或者概率分布等原理进行特征分析，从而识别出目标物。传统的数字视频运动物体识别提取方法，一般是先将视频分解成一帧帧的数字图像，然后对当前帧与背景帧，进行背景差法，获取运动物体区域；或者利用相邻两帧的像素差，利用帧差法进行运动物体的识别提取。考虑到传统背景差法和帧差法的局限性，此次研究将采用结合背景差法，帧差法的数学形态学法提取出视频中的运动物体。

特征向量的直接影响到识别结果的准确性，考虑到所识别的目标物是横幅，属于具有类矩形几何特征的人造物体，另外考虑到视频图像的质量较低，无法获取清晰的图像块细节，而轮廓和角点是图像中最容易提取的结构元素，因此采用图像轮廓作为结构元素，而线段则为目标物的必备特征，轮廓线段的运动特征作为特征向量。

物体遮挡问题一直是机器视觉领域的难点。相对于单物体识别而言，多物体的视频识别难度要更大，因为物体的遮挡会令计算机产生歧义，所以在没有逻辑推理能力的情况下，计算机无法识别出物体遮挡，物体重叠等问题。考虑到轮廓提取算法，线段识别算法必然存在的误差，这里采用对连续K帧图像进行线段序列提取，统计，利用统计规律对图像进行两两比对，匹配，最后区分哪些线段轮廓属于同一个物体，从而解决物体区分，物体遮挡的问题。

**关键词**：视频横幅识别；特征向量；物体遮挡问题；物体区分问题

# Abstract

This paper describes the field of machine vision object recognition in video motion the principle and implementation methods. the main research work is the last link, the digital video image analysis and processing.   
Video identification is the field of machine vision, which involves a wide range of disciplines, with in-depth a wide range of applications.   
The primary task of the video identification is to identify the moving object which then undergo a structural analysis of moving objects, Feature Extraction, and then use artificial intelligence, pattern recognition, or the probability distribution characteristics of principle, in order to identify the target.   
The traditional identification of digital video moving object extraction method, is generally broken down into a first video frame of the digital image, and then the current frame and background frame, the background subtraction method for moving object region; or use the difference between two adjacent frames of the pixel using the frame difference method to extract moving objects recognition. Taking into account the traditional background subtraction and frame difference of the limitations of the study will use the combination of background subtraction, frame difference of the mathematical morphology method to extract moving objects in the video.   
Feature vector directly affects the accuracy of recognition results, taking into account the identified target is the banner, are a class of rectangular geometry of the man-made objects, while taking into account the lower quality video image, can not get a clear image block details The contour is the image most easily extracted structural elements, so a profile as a structural element of the image, while the line was the target of essential features, characteristics of linear motion as the feature vectors.   
Object occlusion field of machine vision has been difficult. Relative to the single-object recognition, the more difficult to identify video objects more, because the computer will block the object ambiguity, so there is no logical reasoning ability, the computer does not recognize the object occlusion, object overlap and other issues. Taking into account the contour extraction algorithm, linear recognition algorithm errors are bound to exist, where the use of continuous linear sequence of K extracted frame images, statistics, law on the use of statistical comparison between two images, matching, and finally to distinguish which objects belong to the same contour line to solve the distinction between objects, object occlusion problem.

**Keywords:** video banner recognition; feature vector; object occlusion

**目 录**

[摘 要 III](#_Toc294076616)

[Abstract IV](#_Toc294076617)

[第一章 绪 论 7](#_Toc294076618)

[1.1 视频识别技术的需求 7](#_Toc294076619)

[1.2 实验开发环境简介 7](#_Toc294076620)

[1.3 主要研究工作 8](#_Toc294076621)

[1.3.1 基于运动物体的视频识别技术当前研究现状及方法 8](#_Toc294076622)

[1.3.2 特征向量 8](#_Toc294076623)

[1.3.3 线段变换能量 8](#_Toc294076624)

[1.4 本文安排 9](#_Toc294076625)

[第二章 横幅视频识别技术概况 9](#_Toc294076626)

[2.1 视频目标识别相关技术介绍 9](#_Toc294076627)

[2.1.1视频目标识别主要问题和难点 10](#_Toc294076628)

[2.1.2视频目标检测常用的技术 10](#_Toc294076629)

[2.1.3视频目标检测中的图像分割技术 10](#_Toc294076630)

[2.1.4 模式识别与横幅识别 13](#_Toc294076631)

[2.2 横幅视频识别系统基本模型 14](#_Toc294076632)

[2.3 横幅识别基本特点以及难点问题浅析 15](#_Toc294076633)

[2.4 本章小结 15](#_Toc294076634)

[第三章 横幅识别技术原理 16](#_Toc294076635)

[3.1检测运动物体 16](#_Toc294076636)

[3.3.1帧差法 16](#_Toc294076637)

[3.3.2背景差法 16](#_Toc294076638)

[3.3.3光流法 16](#_Toc294076639)

[3.3.4数学形态学运动目标检测法 17](#_Toc294076640)

[3.2物体轮廓提取 18](#_Toc294076641)

[3.2.1Prewitt算法 18](#_Toc294076642)

[3.2.2Roberts算法 18](#_Toc294076643)

[3.2.3Canny算法 18](#_Toc294076644)

[3.3从边缘到轮廓 19](#_Toc294076645)

[3.3.1边缘锐化 19](#_Toc294076646)

[3.3.2USM锐化 20](#_Toc294076647)

[3.4边缘分析算法 20](#_Toc294076648)

[3.4.1Hough参数空间 20](#_Toc294076649)

[3.5线段轮廓特征分析 21](#_Toc294076650)

[3.5.1线段特征和线段相似度 22](#_Toc294076651)

[3.5.2线段变换 22](#_Toc294076652)

[3.5.3线段相似度 23](#_Toc294076653)

[3.5.4线段转换能量矩阵 24](#_Toc294076654)

[第四章 横幅识别系统设计 24](#_Toc294076655)

[4.1线段参数数据结构设计 24](#_Toc294076656)

[4.2 线段转换能量表设计 25](#_Toc294076657)

[第五章 横幅识别系统的实现与实验结果 26](#_Toc294076658)

[5.1系统实现： 26](#_Toc294076659)

[5.2实验结果 27](#_Toc294076660)

[结 论 28](#_Toc294076661)

[参考文献 28](#_Toc294076662)

[致谢 29](#_Toc294076663)

# 第一章 绪 论

## 1.1 视频识别技术的需求

随着信息时代的到来，用计算机处理各种信息的需求越来越多，多媒体信息处理技术已成为日常生活中各领域的迫切需要，而人们就更希望能用计算机技术来处理视觉问题，例如利用人脸，虹膜，指纹等识别技术来处理与个人有关的一切事务。利用自动识别技术帮助盲人，利用视觉自动监控系统监控监视环境中发生的非常事件，例如陌生人的入侵，老年人的异常行动等，本次研究的横幅识别技术正是此类应用。

我们现在的社会并不完美，人民内部的矛盾依旧存在。在国内，常常有不够理性的公民们因为他人煽动，不熟悉法律，或者情绪过激，采取非法游行的措施表达意愿。

根据我国的《中华人民共和国集会，游行，示威法》规定：集会、游行、示威的时间、地点、路线等都要按事先批准的内容进行。而非法集会往往是在未经相关政府部门的批准下，由一些不法分子秘密组织并发起；由于集会是表达公众意愿的强烈形式， 再加上不法分子的别有用心，非法集会往往在开阔场地，人流量大的地方，例如政府广场，交通枢纽，商业旺地等进行，一旦集会群众出现过激行为，甚至失去控制，其造成的公共安全危害将难以预料。

鉴于非法集会的突发性，偶然性，以及巨大的危害性，如果网络治安视频监控系统中能实现智能监控集会事件的功能，就可以及时提醒相关治安部门，迅速采取措施，避免意外的发生。

本次横幅视频识别技术研究正是为基于公共安全领域的急切需求而进行的。

## 1.2 实验开发环境简介

Opencv：

OpenCV的全称是：Open Source Computer Vision Library

OpenCV于1999年由Intel建立，现在由Willow Garage提供支持。OpenCV是一个基于BSD许可证授权（开源）发行的跨平台计算机视觉库，可以运行在Linux、Windows和Mac OS操作系统上。它轻量级而且高效——由一系列 C 函数和少量 C++ 类构成，同时提供了Python、Ruby、MATLAB等语言的接口，实现了图像处理和计算机视觉方面的很多通用算法。

Visual C++ express Edition

Visual c++ express edition 是微软提供的免费C++开发工具，它不自带MFC类库，小巧方便，是广大C++爱好者的实用工具。

操作系统：Window 7 32bit

内存：2G

中央处理器：AMD Athlon 2 P3600 2.3GHz

## 1.3 主要研究工作

### 1.3.1 基于运动物体的视频识别技术当前研究现状及方法

计算机视觉是在20世纪50年代从统计模式识别开始的，当时的工作主要集中在二维图像分析，识别和理解上。发展到今天，不仅出现了各种各样的图像识别技术，同时将应用扩展到了视频识别领域中。

视频识别问题最终会归一化为图像识别问题，其关键在于提取图像的特征向量。假如所提取的特征向量能够描述问题的所有解集，并且能成功实现特征向量的提取，则可以认为该图像识别问题在理论上已经实现。

### 1.3.2 特征向量

在模式识别过程中，特征向量的选取对图像识别效果极其关键。

因为特征向量常常作为图像识别问题的解集描述，所以特征向量必须与目标物体紧密联系。本次研究将整个识别过程分解为多个连续的识别问题，每个图像识别问题都有相对应的特征向量来描述识别目标物。其中比较关键的是线段变换能量特征向量（见4.13），用于识别两幅图的轮廓变换情况，从而识别出不同速度的物体，进行物体区分，解决物体遮挡问题。

### 1.3.3 线段变换能量

线段变换能量是本文提出的一个概念，用于描述两条线段之间的相似度。众所周知，在同一个二维空间中，如果两条线段起始点（x,y），方向相同，长度相同，则我们认为这两条线段相似度为100%。那么如何衡量两条不同线段之间的相似度呢？线段变换能量正是这样一个可衡量的标量，后续将详细描述。

## 1.4 本文安排

本文第一章介绍绪论，讲述论文内容、背景等基本情况。

第二章详细介绍视频识别技术原理的概况，其中包含视频识别的详细步骤，以及图像识别的具体原理，图像尺度空间，图像位置空间，图像结构元素等图像图形识别的基本概念，以及它们之间的联系。

第三章重点介绍本文选取的特征向量：线段参数特征向量，以及最重要的线段转换能量特征向量，并将详细阐述选取该特征的原因、好处、缺点等。

第四章重点介绍视频识别系统模型，运用前面介绍到的理论知识设计一个视频横幅识别的系统模型，并分析该模型的详细设计、优缺点等。

第五章是该软件系统模型的编码实现，以及相关的实验过程记录和实验结果分析，该过程进一步验证了系统模型的设计是否达到识别的功能、要求等。

第六七章，将会对编码出来的系统进行系统的测试，分析其识别率，并且进一步探讨特征向量的选定对于识别的影响。

第八章，本文提出了一种对系统的改进方法，弥补系统的一些不足之处，使得系统更加完善。在每一章之后都有该章的小结，在论文最后，是对整个论文题目的总结、参考文献目录。

# 第二章 横幅视频识别技术概况

## 2.1 视频目标识别相关技术介绍

　　在图像序列中进行运动目标的识别跟踪是计算机视觉中一个必不可少的关键技术。与视频序列识别技术有密切关系的应用主要有：智能视频监控，基于视频的人机交互，自动驾驶等等。此外，目标识别跟踪在基于视觉的控制、农业自动化、医学图像、视觉重构等领域均有应用。

　　目标检测和跟踪，就是在一段序列图像的每幅图像中找到所感兴趣的运动目标所处的位置，从而达到检测并跟踪的目的。目标检测跟踪的理论及其实现已经进行了大量、深入的研究工作，并取得了令人瞩目的成果。但是，已有的目标检测跟踪系统大多受限于特定的应用背景，目标检测跟踪算法还有待于进一步研究和深化，研究一种鲁棒性好、精确度高、性能稳定和适用性强的运动目标检测跟踪方法依然面临巨大挑战。

### 2.1.1视频目标识别主要问题和难点

运动目标检测和跟踪算法的主要问题和难点有：

1。各个目标之间的相互遮挡，造成目标信息的不完整，给目标识别带来困难。

2。不同视角和光照变化带来目标形状变化使得实时识别和跟踪非常困难，如何选择和提取具有不变性的特征具有挑战性。

3。在复杂背景下，如何对目标进行精确图像分割，如何提高算法鲁棒性和实时性具有实际意义。

4。由于图像序列有丰富的原始数据，相邻帧之间有很强的相关性，如何利用好时间域上的相关性，提高运动目标检测、识别和跟踪效率具有实际意义。

### 2.1.2视频目标检测常用的技术

　　检测的方法大致分为两类

　　1。先检测后跟踪（deteet一beofre一track，简称DBT），它对每帧图像都作出目标存在与否的判决，然后通过对所有目标运动轨迹进行分析来确认出真正的目标，该方法受图像信噪比的影响较大，只有在高信噪比时才能取得好的效果。这类方法需要被处理图像中目标的能量足够大，否则不能很好地检测出目标。

　　2。先跟踪后检测（tarck一beofer一detect，简称TBD），它开始并不在单帧图像内对目标的存在与否作出判断，而是通过对每条跟踪的轨迹进行分析后，再对目标的存在与否作出合理的判断，大量实验表明，该方法在低信噪比时也可以取得不错的效果。

动态场景下分割常采用DBT方法，因为在实时的检测、跟踪系统中都希望每帧都检测到目标，尤其是发现新目标。

本次研究采用第一种方法，即先检测后跟踪。

### 2.1.3视频目标检测中的图像分割技术

图像分割是目标检测、跟踪系统中的重中之重，序列图像分割结果的好坏直接影响到目标检测、特征提取及后续的跟踪环节。图像分割算法有基于区域的分割技术，基于边缘的分割技术以及基于特定理论的分割技术，除此之外，还有利用序列图像中运动信息的检测技术。

本次研究采用阈值分割法，属于基于区域的分割技术。

**基于区域的分割技术**

　　常用的区域的分割算法有：阈值分割法和区域生长法。阈值分割法阈值的选取在算法中占有非常重要的地位，它对图像分割的效果精度产生直接的影响。

　　80年代初，人们开始运用信息论中熵的概念进行阈值选取。1980年，Pun首先提出了最大后验熵上限法。1985年，Kpaar等人提出了一维最大熵与阈值法。随后，局部熵、条件熵、相关熵等相继被用来选取阈值。

　　从所选取阈值的作用来看，阈值分割算法又可以分为全局阈值和局部阈值。所谓全局阀值就是指在整个分割过程中，只选取一次阈值，并根据该阈值完成对整幅图像的两类分割，常用的全局阈值选取方法有图像灰度直方图峰谷法、最小误差法、最大类间方差法、最大熵自动阀值法等；局部阀值是指在图像的分割过程中，选取的阈值不止一个，可以根据实际情况选取多个阈值，并根据所选取的阈值完成对图像的多类分割。事实上，图像的阈值分割过程也就是像素点的分类过程。经过多年的研究，人们提出了多种基于全局阈值和局部阈值的图像分割算法。阈值分割的优点是计算简单，运算效率较高，速度快。全局阈值对于灰度相差很大的不同目标和背景能进行有效地分割。当图像的灰度差异不明显或不同目标的灰度值范围有重叠时，应采用局部阈值或动态阈值分割法。另一方面，这种方法只考虑像素本身的灰度值，一般不考虑空间特征，因而对噪声很敏感。在实际应用中，阈值法通常与其他方法结合使用。

区域生长法是利用图论的工具，通过把具有某种相似性的集连通而构成最终的分割区域。在这种分割算法中，区域的相似性和连通性是构成同一区域的基本属性。由于这一技术在区域生长过程中融入了像素邻域信息，因而它具有较强的抗噪声性能。使用这种分割技术的关键主要是区域之间相似性的确定。

**基于边缘的分割技术**

　　基于边缘的分割是在边缘特征提取的基础上进行的。首先通过边缘检测算子得到图像属性（通常是灰度属性）变化较大的边缘像素，然后，利用这些边缘像素对图像进行分割。边缘提取方法可以分为经典边缘提取方法和现代边缘提取方法。经典的边缘提取算法是考察图像中每个像素的某个邻域内灰度的变化，利用边缘邻近的一阶或二阶方向和变化规律，用简单的方法检测边缘。常用的边缘检测算子有：sobel算子，Lapxae算子，Robert算子，Prewitt算子等。但这些局部的边缘算子，由于微分运算固有的特性而对噪声干扰十分敏感。70年代出现了几种改进的边缘检测算子。1976年，Rosneefdl通过组合多个不同尺度算子的输出，提出了非线性边缘检测技术，随后Perwitt，Haralick等先后提出了基于曲面拟合的边缘检测算法。80年代，Marr和Hidreth通过将尺度概念融入边缘检测技术中，提出了著名的过零交叉理论及LoG（LaplaeeofGaussina）算子，即著名的Marr-Hidreths算子，从而拉开了现代边缘检测技术的序幕。1983年，witkin提出了多尺度滤波理论，Burt等人提出了基于Laplaee塔式分解的图像分解和重建理论。1986年，woods运用带通滤波器实现了信号的多分辨分解。基于边缘的图像分割技术中的关键问题是：能否得到具有精确定位的边缘特征图。对于所得到的边缘特征图，能否得到封闭的区域边界，也就是如何将杂乱的不封闭边界封闭化，并通过对这些封闭的区域进行分裂和合并，最终得到封闭的边界。由于分割的依据是图像的边缘，因而基于边缘的分割算法的抗噪声能力较差。

**基于特定理论的图像分割技术**

　　1基于小波变换的边缘检测方法

　　该方法是借助新出现的数学工具小波变换来分割图像的一种方法，也是非常新的一种方法。小波变换是一种多尺度多通道分析工具，比较适合对图像进行多尺度的边缘检测。90年代初，Mallat在图像分析的基础上，将小波变换技术运用于图像多尺度边缘检测中，提出了小波变换局部模极大值的边缘检测法。小波变换的计算复杂度较高，抗噪声能力强。理论证明，以零点为对称点的对称二进小波适应检测屋顶状边缘，而以零点为反对称点的反对称二进小波适合检测阶跃状边缘。近年来多进制（MultiBnad）小波也开始用于边缘检测。另外，利用正交小波基的小波变换也可提取多尺度边缘，并可通过对图像奇异度的计算和估计来区分一些边缘的类型。利用小波变换和其它方法结合起来的分割技术也是现在研究的热点。靳华等人提出用树型小波来提取纹理特征进行纹理图像分割的方法。贾天旭等人提出的基于小波包分解自适应Gbaor函数设计的纹理分割算法，首先用Shannon小波包检测纹理的主频，设计Gbaor函数，然后根据Gbaor函数与纹理图像的卷积，就可以在纹理的连接处产生良好的阶跃边缘。

　　2基于分形理论的图像分割方法

　　分形理论的发展为图像边缘检测和提取提供了一条崭新的思路。复杂自然景物具有自相似性，虽然这种自然景物的自相似性不能与人为构造出的纯数学分形相比，但与人造目标来比，自然物体的分形特征还是有别于人造目标的分形特征。近年来，人们对基于分形特征的目标图像边缘提取技术进行了初步的探讨。

　　3基于模糊技术的图像分割方法

　　由于传统的图像分割算法对噪声敏感，或者存在计算量大的缺陷。为了得到性能更为优异的图像分割算法，人们对基于人工智能原理的图像分割算法，尤其是基于模糊技术的图像分割进行了有益的尝试。1980年Pal和Knig首先提出了模糊边缘检测的概念。由于图像分割过程所存在的不确定性和模糊性，模糊边缘检测技术引起了人们的极大关注。在图像分割过程中所涉及到的模糊技术主要有：模糊阈值技术，模糊聚类技术以及模糊边缘检测技术等。

　　模糊阈值技术：所谓模糊阈值技术，一般来讲，就是通过对模糊度的优化计算来获得最佳阈值。具体过程就是通过选取合适的隶属函数来定义物体的模糊区域，进而通过对模糊度的最小化的优化计算，来进一步获取最佳的隶属度函数。这个最佳的隶属度函数给出了像素对模糊物体区域的隶属度，根据隶属度函数的交点来确定最佳阈值。

　　模糊聚类技术：根据前面的分析可知，图像的分割过程，实质就是根据某种标准对图像像素进行分类。由于图像分割过程中的不确定性往往是由模糊性造成的，故模糊技术应用于图像分割，是图像分割技术发展的必然。在图像分割领域中，应用最为广泛的模糊聚类技术是模糊C--均值聚类技术模糊C--均值算法也称为FCM算法，它是通过对目标函数门迭代优化来实现对数据集合的模糊分类的，其聚类的准则是使误差的平方和最小。

### 2.1.4 模式识别与横幅识别

模式是指人们在一定的条件环境下，根据某种特定的需求，对自然事物的一种抽象分类的概念，比如桌子、椅子等，人们把它归并为家具，铅笔、橡皮等，人们把它归并为文具，它们是不同的种分类。模式样本是人们按照一定的类别特性去区分自然界的具体事物，，它是抽象模式的具体体现。

在机器视觉中，模式识别是非常重要的。机器视觉系统通过照相机捕捉图像，然后通过分析，生成图像描述信息。以本次研究为例，计算机从数字视频文件中进行数字图像的采集，然后进行模式识别，识别系统必须将识别物分为“横幅”和“不是横幅”两种，同时还需完成“横幅A”“横幅B”的识别功能，然后根据分类结果采取相应的行动，比如发出预警声音或闪动预警灯等。

模式识别是对未知模式样本的分类属性进行辨别的过程。但是要进行模式样本类别属性的正确判定，必须根据对样本的各种可可以观察和测量得到的特征量度来进行。

说到模式识别，就离不开特征，特征向量和分类器。

特征，顾名思义，就是人们对识别物的抽象。比如一个多边形，它的边数，角点数，紧凑度和中心距等都是它稳定的特征，不受伸缩，平移，旋转变化的影响。特征向量则是识别物特征的数学抽象，将所选取的特征用对应标量标示成一个向量，则称之为特征向量。特征分类器用于判定特征标本属于哪个解集，通常表现为特征平面空间的决策线，或者决策面。下图给出了模式识别的各个阶段：

（图《模式识别》4）

从这些反馈箭头可以看出，每一步都不是独立的，相反它们之间互相关联，互相依赖。为提高整体性能，每一阶段都有可能返回到前一阶段重新设计。而且有一些阶段可以合并，例如，特征选择和分类器设计阶段处于同一优化任务中。

本文要做的事情，就是把模式的概念运用到横幅识别上面来。模式是一种抽象的分类，我们把横幅（类矩形）作为一个“类”，因此我们区分相应的物体的时候，必然有能表现该“类”并且辨别其他“类”的特征，那么我们人眼是怎么辨别两个图形是不是类似矩形呢？一个明显的特征是角点数，如果一个几何图形具有四个角点，另一个几何图形只有三个角点，那明显前者是横幅的可能性比后者更大。另一个特征是轮廓中包含的线段数量，如果一个物体拥有两条线段作为轮廓，另一个物体的轮廓全部为矩形，则明显前者是矩形的可能性更大。考虑到横幅会产生形变，而且其中央通常有文字，图案等不规则纹理，此次研究将采用多种特征组合向量的形式，力求取得稳定的特征分析结果。

## 2.2 横幅视频识别系统基本模型

横幅识别系统的系统框架图如图下图。它包含图像预处理、图像边缘处理、目标物识别、后处理器等几个主要步骤：

1. 预处理阶段：对图像（包括背景图像，当前帧）进行高斯平滑，去除一定的噪声，并进行数学形态法获取运动图像区域，然后用阈值进行二值化；

同时对原图像进行USM锐化，强化图像边缘

1. 图像边缘提取：将第一阶段获得的运动物体图像区域作为掩膜，从源图像的灰度图像中获取运动图像区域的边缘。
2. 识别运动物：对每帧图像的目标区域轮廓进行线段识别，统计每帧图像的线段轮廓信息，然后综合图像序列的线段特征值，判定每帧图像中的物体区域，并把它们分别识别和提取出来。
3. 横幅识别：统计每个识别区域的特征向量，然后再跟特征库进行特征比对，识别出其中的横幅。

图像预处理

图像边缘提取

横幅识别

识别运动物

视频图像字图形

识别器

图 2-1 横幅识别系统的系统框架图

2.3 横幅识别基本特点以及难点问题浅析

横幅是典型的人造物体，类似矩形的外轮廓特征是它的显著特征，而且它应该会具有一定的形变。

## 2.4 本章小结

本章主要介绍了视频目标识别与跟踪的相关技术，阐明了所选技术的原因，以及所选技术的优势与局限性。另外描述了模式识别与横幅识别的关系，以及整个横幅识别系统的基本模型和大概流程。

# 第三章 横幅识别技术原理

## 3.1检测运动物体

传统的运动物体检测方法：

运动目标检测是视频识别技术的第一步。常用的运动物体检测方法有帧差法,光流法和背景差法。

### 3.3.1帧差法

帧差法是在连续的图像序列中两个或三个相邻帧间采用基于像素的时间差分并且阈值化来提取图像中的运动区域。帧差法对环境的噪声污染和光照变化要求低，具有较好的抗干扰性，但是帧差法只适用于运动速度较快的物体，对于运动速度比较慢的物体，容易出现检测空洞。测试效果如下：

### 3.3.2背景差法

背景差法通过当前图像与背景图像之间的差分来检测运动区域。其优点是位置准确, 速度快, 只需要获取当前的一幅图像;不足之处是受环境光线变化的影响, 在非受控环境下, 需要加入背景图像更新机制。图背景差法能有效检测出各种速度的物体，但容易受环境变化的影响，比如光照条件。另外 ，背景差法要求背景环境灰度比较单一，对于灰度跳跃丰富的背景，容易在检测结果中出现背景图像的结构噪声影像。测试结果如下图：

### 3.3.3光流法

光流法[ 1] 是利用运动估计进行分割和跟踪运动目标的一种方法。光流法在不需要背景区域的任何先验知识条件下就能够实现对运动目标的检测和跟踪, 还可以应用于摄像机运动的情况, 但计算量非常大, 而且对噪声极为敏感。考虑到横幅识别环境的复杂多变，本次研究不采用光流法。

### 3.3.4数学形态学运动目标检测法

针对传统运动目标检测算法的不足，文献【基于数学形态学的运动目标检测算法】提出一种基于数学形态学得帧差法与背景差法相结合的运动物体目标检测算法，即先将帧差法得到的图像经过数学形态学处理，再将其与背景差法的图像相结合，实现表明，该算法能适应复杂变化的环境，准确地提取运动目标，对复杂干扰场景下的实时运动目标检测得到了较为满意的效果。终上所述 ，我们采用了第四种运动目标检测算法，来提取目标物体。

文献【基于数学形态学的运动目标检测算法】提出的数学形态学运动物体检测方法，其具体流程如下：

背景图像

当前帧

上一帧

背景差分并且二值化

帧间差分并二值化

数学形态学处理（先进行腐蚀，然后进行膨胀操作）

对应像素进行与操作

其识别结果如图：

（图）

## 3.2物体轮廓提取

对图像函数局部梯度的近似是很多经典的边缘检测算法的基础。实际上，不同边缘检测算法的差异，仅在于用于估算梯度分量的滤波器类型以及将这些梯度分量结合起来的方法。在血多情况下，我们不仅关心边缘强度，还比较关心边缘的方向，这两类信息都包含在梯度函数中，并可以简单地由方向分量算出。传统的边缘检测算法有很多，其中一些不一定有使用价值。

### 3.2.1Prewitt算法

Prewitt使用的线性滤波器包含相邻的三行三列，以抵消简单的（单行单列）梯度滤波器的噪声敏感性。Prewitt算子采用如下滤波器：

我们可以看到，Prewitt算法采用一个三行箱式滤波器，一个三列平滑滤波器，对八个方向进行边缘检测，把其中响应幅度最大的方向作为图像的边缘。其优点在于对噪声具有平滑作用。与prewitt算法类似的还有sobel算法

### 3.2.2Roberts算法

作为最简单，最古老的边缘检测器之一，Roberts算法使用了两个非常小的2\*2滤波器，沿图像的对角线方向来估算方向梯度。

Roberts算法对图像中的对角线边缘有着良好的响应，但却没有很好的方向选择性。

### 3.2.3Canny算法

比起prewitt算法，sobel算法，和Roberts算法，Canny算法更受欢迎。他在不同分解的图像上使用一些列的相对大尺寸的有向滤波器进行滤波，得到的边缘图比其他简单算子的到得结果更加清晰。

我们采用的工具包opencv中，有完整的canny算法库函数。所以下面的实验过程中，获取图像轮廓部分都是采用canny算法实现的。

### 3.3从边缘到轮廓

我们很容就能想到，循着已经检测到得边缘去找轮廓线，并且看上去也是很容易实现的。比如，从边缘强度较大的地方开始，沿着两个不同方向跟踪边缘点，直到这两条轨迹相遇并形成一条闭合的轮廓线。不过事实常常令人遗憾，因为总有那么几个难点让这项工作并非看起来那么简单，比如在灰度梯度很弱甚至消失的地方，边缘将会中断。或者比如，交叉的边缘会产生歧义，还有，轮廓线有时可能会存在多个分支，等等。所以轮廓跟踪通常不用于原始图像或者连续取值的边缘图像中，而是在二值图像中进行轮廓跟踪，这样会简单很多。

在实际情况下，边缘图中通常很少存在完美的轮廓线。更普遍的情况是包含需对细小的，不连续的轮廓片段，他们常常在边缘强度不足的地方中断。为此，我们需要对源图像进行边缘锐化，尽可能地保留图像的边缘信息。

### 3.3.1边缘锐化

拉普拉斯滤波器

二维函数的的拉普拉斯算法定义为沿x和y方向的二阶偏导数之和：

它可以由一些列简单的线性滤波器来估计。

这样我们就可以将它们结合起来成为一个二维拉普拉斯滤波器：

拉普拉斯边沿锐化算法的缺点在于对噪声相当敏感，通常情况下，需要通过采用预平滑的措施来减弱噪声的影响，比如高斯平滑等。

### 3.3.2USM锐化

USM是“Unsharp masking”的缩写。顾名思义，这种边缘锐化技术的原理是从原始图像中减去一个它平滑后的副本，这一结果称作掩膜。然后将此掩膜重新叠加到源图像上，这样图像中的边缘就得到了强化。USM的具体步骤如下：

1. 获取掩膜：通过源图像I减去经过平滑后的副本

其中，

1. 将掩膜M以因子a加权后叠加到原图像I上。权重因子a用于控制锐化后的强度

USM滤波器相对于拉普拉斯滤波器而言，其优势在于它降低了对噪声的敏感性，因为它包含了一个平滑的过程，并且通过参数a提高了滤波结果的可控性。

当然，USM不仅仅对真实的边缘有响应，它对任何灰度变化的地方都有响应，因此它也会放大图像中的一些噪声。由于其简单性和灵活性，USM滤波器在几乎所有的图像处理软件（比如adobe photoshop）中都可以找到。在下面的实验过程中，我们将利用OPENCV提供的CvSmooth()等库函数实现USM滤波器，锐化图像边缘，并从中提取出图像中运动物体的完整边缘。

## 3.4边缘分析算法

Hough算法是著名的几何图形识别算法之一，是由美国的Paul Hough提出的。它通常用来定位一些形状，比如许多几何形状，如直线，圆，椭圆等可以用简单的只有几个参数的公式来描述的几何形状。我们知道，通常情况下，人造物体往往以简单的几何形状出现。比如卡车的外轮廓近似矩形，飞机的外轮廓近似于五边形，而我们要识别的横幅则是其中之一，它的外轮廓也近似于矩形。

### 3.4.1Hough参数空间

我们知道，图像是个二维的平面空间。假设图像中有一个点，对于所有适当的，值，每条通过点的线段必须满足：

即：

这是直线的另一种表达方式，其中，是变量，，是常量。这样，上述式子的解集就可以描述通过点的所有肯能的直线。

现在假设有平面，以，为横纵轴，平面中的直线和直线相交，

这意味着有一条直线同时通过，.

参数空间中在一点处相交的直线越多，则图像空间中相应直线上的点数就越多。

Hough算法正是通过这个原理，用几何参数函数，创建参数空间，并为参数空间的k,d轴选择合适的步长，转连续空间为离散空间，然后对参数空间进行累加统计，从而获得图像空间的显著直线。

通常情况下，上述的直线表达式有其局限性。我们知道，垂直直线,这使得上述的直线表达式在实际应用中不被使用。为此，我们采用另一种更安全的表达式：

这样参数空间的横纵坐标则为：

转换为离散参数空间则为：

式中H，W分别为图像的高和宽

Hough参数空间变换算法对离散近似直线响应极好，而且具有很好的抗干扰能力。在下面的研究中，我们将采用opencv自带的cvHough()函数对边缘图像进行线段结构线的提取，为后续的进一步分析研究做好图像预处理工作。

## 3.5线段轮廓特征分析

从这里开始是整个图像识别流程的核心部分。

现在我们回顾一下前面的流程：源图像去噪，锐化边缘，结合帧差法，背景差法获取运动物体区域；然后提取运动物体的边缘轮廓；并提取运动物体轮廓中的线段部分。

经过一系列流程下来在对视频图片进行Hough变换提取线段之后，我们需要对线段序列进行分析。

### 3.5.1线段特征和线段相似度

线段作为基本的几何形状单元，在二维空间有其特定的特征值。我们知道，在二维图像空间中，线段可以用两个点来完全描述，如。但这组参数只包含了两个端点信息，其他的几何信息都被隐藏了。为了方便研究和计算，通常情况下，采用可读性更强，针对性更强的参数形式其中Θ是线段所在直线的角度，其范围是l是线段的长度，x，y是线段起点坐标

说起线段的特征向量，可能了解模式识别的人会觉得不可思议，或者认为这是对特征向量概念的一种误解。因为所有线段几乎都“一样”，大家研究的是每个汉字有什么不同，诸如结构，方向线，笔画数等等，或者研究每个指纹有什么不同，采用Gabor滤波提取一个个32维甚至更高维数的特征向量，然后进行比对。计算机能识别出甲和乙的指纹，或者识别出甲大拇指和食指的指纹，因为它们的特征向量近似相等；我们会说“人”字和“入”字很像，因为它们的特征很接近，但没人会说线段“--”和“-”很像，而且这样区分线段显得没有意义。在模式识别中，通常情况我们会选取那些具有平移不变性，旋转不变性，和缩放不变形的属性作为特征值。以简单图形的识别为例，假如我们要在一堆不规则的图像中找出一个最像矩形的物体出来，十有八九我们会选取角点数，紧凑度，中心矩等属性，而不是质心，对角线长度，面积或者周长。

事实上，参数向量 确实能描述线段“-”和“--”它们各自的特征，本质上而言就是它们的特征向量，而且与其他不规则几何物体一样，我们可以用它们的特征向量来描述它们之间的“相似”程度，虽然它并不是用于模式识别，所以，为避免混淆，这里称之为线段相似度。

### 3.5.2线段变换

假设同一平面空间中，有两条不同的线段a，b，当线段从a变换为b时，需要经过三个过程：

1. 平移
2. 旋转
3. 缩放

线段平移时，方向不变，长度不变，起点坐标改变，变换向量是;

线段旋转时，方向改变，长度不变，起点坐标不变，变换向量是;

线段缩放时，方向不变，长度改变，起点坐标不变，变换向量是

则线段a到线段b的变换可以描述为：

为方便描述，以及归一化，我们将平移变换转换为沿x坐标变换和沿y坐标变换，则任何二维空间内的线段变换都可以表示为四个单位变换向量 的线性累加：

（a）

事实上，不仅仅线段在二维空间中的线性变换如此，其他规则的几何图像结构元素，甚至不规则图像结构元素，比如区域，边缘碎片等在二维空间中的变换也都同样可以用四个变换向量的线性累积来描述。对于规则几何图像结构元素而言，获取变换向量并不是件很难的事情。

### 3.5.3线段相似度

考虑如下两幅图，假设它们是视频中连续的两帧图像，并且这两幅图像均为物体低速运动下的图像：

（图）

第一幅图中有一条线段a，位于（1,0）,长度为4，角度为π/4;第二幅图中出现了两条线段，其中一条是第一幅图中的a。第二幅图中一条是b,位于（1.5，0）长度为4.2， 角度为π/4；另一条是c，位于（5,0）长度为4.3，角度为π/3。有什么方法可以判断b，c到底谁是a的可能性大点？或者说谁“像”a多点？

其中一个办法是使用线段变换所需的能量来描述。

任意线段变换都可以表现为平移，缩放和旋转的线性累加。我们知道，物体在空间中变换位置是需要消耗能量的。假设点为单位质量，作用力为单位力，则线段a进行平移，旋转和缩放时所需的能量分别为：

Ep = l\*(dx + dy);

Et = l\*dΘ;

Ec = l\*dp;

在视频中检测物体时，考虑到视频设备获取图像的速度非常快，物体在每两帧相邻的图像中的位移差并不大，基本近似直线运动，再考虑进线段提取过程中产生的误差，我们采用线段变换能量的概念来衡量两帧图中两条直线边缘属于同一线段边缘的可能性。

假设第一帧图像中的线段a，变换为第二帧图像中的线段b，所需要的能量为 ,当

越小时，a,b属于视频中同一线段的可能性越大。

### 3.5.4线段转换能量矩阵

设想我们从相邻的两帧图像, 中分别检测到了n, m条线段，对它们进行两两变换，则矩阵描述了在我们检测范围内的两幅图像所有直线变换情况。考虑线段集合

；

假设绝对理想的情况，即两幅图的所有直线都被检测出来。则矩阵元素可以区分为三类：

1.矩阵描述的两条直线之间存在实际转换：即对应两条直线在S,S’

# 第四章 横幅识别系统设计

## 4.1线段参数数据结构设计

Opencv中有方便的结构

类型来存储图像数据。

CvRect

{

Int x;

Int y;

Int height;

Int width;

}

CvPoint

{

Int x;

Int y;

}

本次研究，采用Opencv函数库动态数据结构中的队列CvSeq进行线段参数的动态存储。每幅图像I都有一个线段序列，它们是构成线段转换能量矩阵的基础。

1）以CvRect方式存储，CvpPoint取出元素

Opencv中的队列是个强大的动态数据结构。它在创建时，必须指定元素的类型和大小。如下例：

CvSeq\* test = cvCreateSeq( 0, sizeof(CvSeq), sizeof(ElemName), cvCreateMemStorage(0) )；

Opencv中并没有cvline类型，只有cvRect{int,int,int,int}，cvpoint{int,int}。可以看到：

所以我们可以以cvRect类型作为元素创建队列，然后用cvPoint指针来获取队列中的元素。具体代码如下：

## 4.2 线段转换能量表设计

线段转换能量的具体表达方式是个矩阵,直接采用CvMat方式存储。矩阵的行，代表前一幅图像的线段索引；矩阵的列，代表后一幅图像的线段索引。

实现代码如下：

CvMat\* enChange = cvCreateMat( height, width , CV\_32FC1)

其中参数CV\_32FC1表示该矩阵是个32浮点性单通道矩阵。它可以转换为一幅相同大小的灰度图像，并显示在窗口上。

# 第五章 横幅识别系统的实现与实验结果

## 5.1系统实现：

本次系统采用面向过程的设计方式。

主要数据结构有：CvImage, CvMat,用于存储图像信息，CvRect，CvPoint，用于存储图元信息。

主要的函数接口如下：

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*函数声明\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

void findObj( CvMat\* );

CvMat\* showObjByColor( const CvMat\* );

void findContours( IplImage\* binaryImg ,CvSeq\* contours, CvMemStorage\* contoursMem ); //寻找二值图的轮廓

void findContours( const IplImage\* srcBinImg, IplImage\* dstBinImg);

int myFindObj\_FloodFill( const IplImage\* srcImg, IplImage\* dstImg, int\* objPoint); //寻找并标识图片中的不同物体，返回物体数量(洪泛法)

CvSeq\* myLine2Feature( const CvSeq\* ); //转换直线数组为特征量数组

CvSeq\* myFeaDiff( const CvSeq\* lineFea1, const CvSeq\* lineFea2 , int range );

CvSeq\* myMatchObjLines( const CvSeq\*, const CvSeq\* , int range );

static int myFeaDiffSeqSort( const void\* \_a, const void\* \_b, const void\* userData); //自定义排序算法

static int mySortLine( const void\* \_a, const void\* \_b, const void\* userData );//自定义排序算法

CvSeq\* feaDiff2Objs( const CvSeq\* feaDiff, const CvSeq\* linesA ); //转换特征差数组为对象数组

CvSeq\* objs2rects( const CvSeq\* objs );

int theaVec( const CvPoint\* vectorA, const CvPoint\* vectorB );

void getClosePoint( const CvPoint\* lineA, const CvPoint\* lineB, CvPoint\* recPoint );

void getRectFromPoint( CvPoint\* pointA, CvPoint\* pointB, CvPoint\* pointC, CvRect\* rect );

int vecDot( CvPoint\* pointA, CvPoint\* pointB ); //两个向量的点积

void printLineSeq( const CvSeq\* lines );

void drawPointSeqImg( IplImage\* img, const CvSeq\* points, const CvScalar color, int lineWidth);

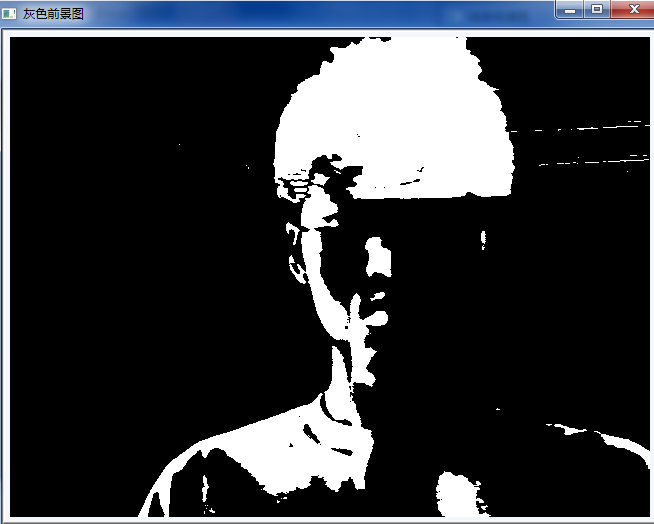
void drawLineSeqImg( IplImage\* img, const CvSeq\* lines, const CvScalar color, int lineWidth);

## 5.2实验结果

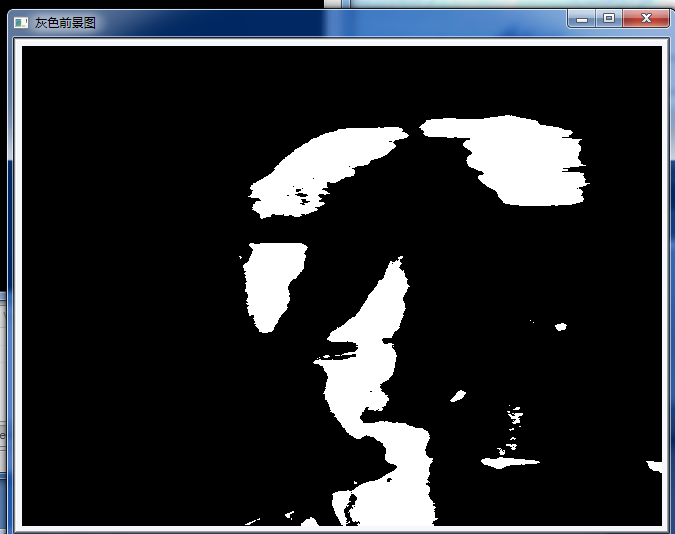
5.2.1 运动物体区域检测结果

实验表明，采用数学形态学提取运动物体区域，比单纯的背景差法和帧差法所取得的运动物体区域要完整，而且几乎没有空洞现象出现

背景差法所检测出的运动物体，可以明显看到背景图中显著结构元素的影响：



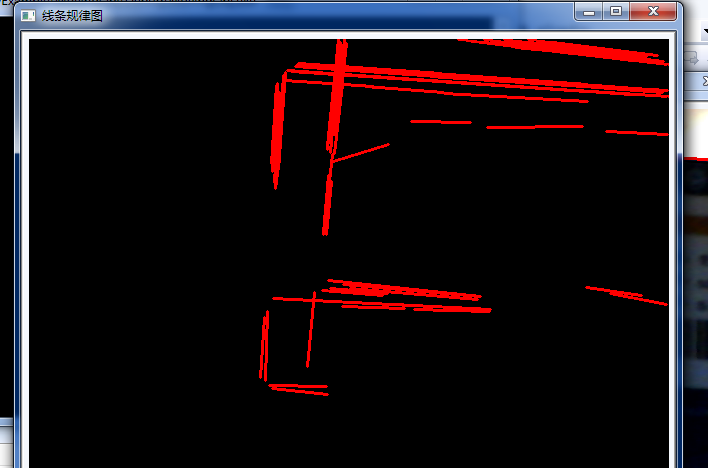
帧差法所检测出的运动物体，可以明显看到图中的空洞效应

数学形态学法检测出的运动物体

5.2.2 运动物体轮廓直线检测结果

实验表明，Hough算法对物体的直线轮廓或者类似直线的轮廓有很好的响应。

Canny 算法，Hough算法检测出的运动物体的直线轮廓

累计20帧的运动物体的直线轮廓

5.2.3 运动物体轮廓直线的匹配结果

实验表明，根据直线变换最小能量法在直线变换矩阵中获取的匹配结果相当准确，在多次实验中都能准确匹配出来。

|  |  |  |  |
| --- | --- | --- | --- |
| 实验次数 | 1 | 2 | 3 |
| 运动物体 | 学生卡 1张 | 白色书本1本 | 黑色书本 1本 |
| 识别结果  （直线轮廓边数） | 4 | 3 | 4 |

上述实验结果均在以宿舍墙壁作为背景，因为白色书本与背景颜色接近，所以在进行数学形态学提取运动物体区域时进行图像像素的减法，必然导致得出的运动区域像素值很低，很容易低于阈值，这样就造成了检测空洞，无法识别出运动物体区域。

# 结 论

该系统总体上能够识别出大部分矩形物体，大致实现了监控环境中的横幅识别功能，在论文的过程中，本次研究深入地研究了视频图像序列以及运动物体轮廓的微运动特征，同时深入学习当前主流的识别技术，结合我自身的知识水平、能力，设计了该方案，并且编码模拟实现了核心功能。

实践证明，直线轮廓确实可以很好地描述一个运动物体的外形特征，由于时间以及个人能力的关系，没有能够运用多种结构元素，例如角点与轮廓特征结合，这样导致一些细节的或者有歧义的物体轮廓无法正确识别，只能停留在大轮廓方向粗糙范围内的识别效果，对于精确度要求较高的识别还是力不从心。

# 参考文献

[1] (美)Rafael\_C\_Gonzalez,数字图像处理（第二版）［M］电子工业出版社 2003.3

[2] Wilhelm Burger Mark J.Burge ,数字图像处理——java语言算法描述［M］ 清华大学出版社 2010.2:130-142

[3] 徐慧,数字图像实用工程案例精选［M］人民邮电出版社 2004.4: 214-254

[4] 陆宗骐 ,C/C++图像处理编程［M］ 清华大学出版社 2005.1: 260-265

[5] 边肇祺 .张学工 ,模式识别（第二版）［M］清华大学出版社 2007.7: 89-128

[6] Welsley E.Snyde, 机器视觉教程［M］机械工业出版社 2005.1:130-165

[7] Steger C ,机器视觉算法与应用［M］清华大学出版社 2008.11:260-265

[8] 高隽，谢昭，图像理解理论与方法［M］科学出版社 2009

[9] 王熙照 模糊测度和模糊积分 [M] 科学出版社 2008

[10] JohnD.Woodward, Jr. Nicholas M. Orlans Peter T. Higgins 生物认证 [M] 清华大学出版社 2004

[11] 张鹏林 复杂场景视频序列图像运动物体提取方法研究 [D] 武汉：武汉大学 2006

[12] 何娜 视频监控中运动物体自动跟踪技术的研究 [D] 湖南：南华大学 2010

[13] 杨明明 基于DirectShow的数字视频监控报警系统的研究与设计

# 致谢

本次毕业论文的顺利完成，首先要感谢黄翰老师的精心指导，让我在方向上有了一个正确的指引，其次要感谢其他在我写论文的过程中给予我帮助的同学，最后是感谢软件学院，给予我们一个很好的机会去锻炼总结我们四年来所学习到的东西，让我们的水平进一步的提高。