



山东建筑大学

本科毕业论文

题 目：计算机视觉跟踪运动人体
目标方法研究

院（部）：理学院

专 业：应用物理学

班 级：光电 043

姓 名：曹海涛

学 号：2004121097

指导教师：贺长伟

完成日期：2008 年 6 月 20 日

目 录

摘 要.....	III
ABSTRACT.....	IV
1 前 言	1
1.1 课题研究背景和意义.....	1
1.2 国内外研究现状.....	2
1.3 图像的预处理.....	4
1.4 目标检测.....	4
1.5 目标跟踪.....	4
1.6 技术难点.....	4
1.7 面临的问题.....	5
2 图像预处理	7
2.1 图像增强.....	7
2.1.1 对比度增强	7
2.1.2 直方图均衡化	7
2.1.3 实时中值滤波快速算法的并行实现	8
2.2 算法介绍.....	8
2.3 预处理结果与分析.....	10
2.4 噪声消除.....	11
2.5 形态滤波.....	11
2.5.1 腐蚀	11
2.5.2 膨胀	12
2.5.3 开运算	13
2.6 阴影消除.....	14
3 目标物的检测	16
3.1 运动目标检测.....	16

3.2 目标分类.....	17
4 目标物的跟踪.....	19
4.1 运动目标的跟踪技术.....	19
4.1.1 基于 Hausdoff 距离的运动跟踪.....	19
4.1.2 基于区域的运动跟踪.....	19
4.1.3 相关跟踪.....	20
4.1.4 光流跟踪.....	20
4.1.5 基于网格的运动跟踪.....	20
4.2 人体运动状态估计.....	21
4.3 卡尔曼滤波理论.....	22
4.3.1 滤波理论基础.....	22
4.3.2 卡尔曼滤波器的特点.....	22
4.3.3 卡尔曼滤波发散的原因和抑制.....	23
5 结 论.....	24
谢 辞.....	25
参考文献.....	26

摘要

视觉跟踪问题是当今计算机视觉领域的研究热点。随着社会公共安全体系的逐步完善，公共场所中对安全智能监控系统的要求越来越高，这使得许多知名公司和科研机构在此课题中投入大量人力财力。人体运动目标的检测与跟踪是视觉跟踪的重要部分，它在未来研究领域中将有广阔的前景并能带来巨大的社会效益。

运动目标检测与跟踪的研究对象是视频序列，或者说图象序列。运动目标检测就是判断视频序列中是否存在运动目标，并确定运动目标的位置;运动目标跟踪是指在整个视频序列中监控运动目标的时空变化，如目标的出现与消失，目标的位置、尺寸和形状等。运动目标的检测与跟踪是紧密关联的两个过程，跟踪始于检测，而目标在后续图象序列中的重复检测也有助于目标的跟踪。由于光照变化、背景干扰、阴影、摄像机的抖动以及运动目标之间遮挡等现象的存在，都给运动目标的正确检测与跟踪带来了极大的挑战。

关键词： 运动目标跟踪；图像处理；卡尔曼滤波

ABSTRACT

Visual tracking is currently one of the most active topics computer vision. As U.N. attaching the importance to the social public safety, the safe intelligent detection system needs to be improved and become more and more effective in public. Then the research agency of many countries and some world famous companies are now paying much attention to this field. As an essential part of the human motion analysis system, the research in people detecting and tracking can provide broad potential use in the future and economic benefits to the society.

The research interest of moving objects detection and tracking is video sequence or image sequence .Object detection in videos involves verifying the presence of an object in image sequences and possibly locating it precisely. Object tracking is to monitor an objects spatial and temporal changes during a video sequence,including presence,position,size, shape,etc. These two processes are closely related because tracking usually starts with detecting objects, while detecting an object repeatedly in subsequent image sequence is often necessary to verify tracking. Due to changes in illumination of the scene background perturbations shadows, vibration of the camera and occlusions between moving objects, precise detection and tracking of moving objects are still a challenge field of research.

Keyword: Moving object tracking ;image processing ;Kalman filter

1 前言

1.1 课题研究背景和意义

计算机视觉是计算机科学和人工智能的一个重要分支。它研究的主要内容是怎样利用各种成像系统代替视觉器官来获取信息，由计算机代替大脑完成对信息的处理和解释，使计算机具有人一样的视觉功能。智能视频监控是计算机视觉的一个重要应用领域，有着广阔的应用前景，尤其是那些对安全要求比较敏感的场所，如飞机场、地铁站、银行、商店、停车场等；另外，还有工业生产现场的监控以及交通系统的监控等等。目前监控摄像机的应用已经比较普遍，但大多没有充分发挥实时主动监督作用，没有智能化，摄像机只是一个电子眼而已，其具体监视过程，仍由人来完成。显然依靠人来执行这种长期的枯燥无味的例行监视工作其效果是不理想的，不但费用高，而且易受人为因素的影响，例如责任心，情感因素等等。因此在这些场合引入智能监控是非常合适的。要求监控系统每天 24 小时连续监控，能自动分析摄像机捕捉到的图像信息，当监视区内有违法活动或有异常状况出现时，能采取必要的安全措施，同时向保安人员及时准确地发出报警信号，从而防治犯罪的发生，避免不必要的损失。2001 年美国 9.11 事件。2005 年英国 7.7 伦敦地铁爆炸案以及去年的美国校园枪击案等恐怖暴力事件的发生更使得世界各国都高度重视如何对国家重要安全部门和敏感的公共场合进行全天候、自动的、实时的监控，而智能视频监控系统为解决这一问题提供了一种有效的途径。

近年来，各行各业对视频监控需求不断升温。特别是美国 9.11 事件以来，一些人群比较密集的公共场所或比较容易受袭击的公共场所，如机场、体育馆、外国使馆、地铁和银行等，都纷纷安装视频监控系统以保障人民生命和财产安全。目前在建造智能大厦和选购住房时，安全防范系统越来越受到人们的重视。如果对现有的视频监控系统加以改进，实现对被监控目标的自动识别功能，就能够大大地降低犯罪率，节省人力物力资源，节约投资。有效的交通管理是一些大都市面临的难题。智能视频交通控制系统能及时提供各路段的车辆流量和路况信息，记录违章车辆，以便实现准确快速的交通指挥调度，达到充分利用现有的道路资源，提高突发交通事故的处理能力，从而为人们的出行提供快捷舒适的交通服务。一些工业生产线上，也利用无人监控系统检测产品质量。视频监控在军事上也有广阔的应用前景。准确及时地掌握海防区域的军事情况，对于有效保卫祖国的领海和领土，在未来战争中做出快速反应、掌握战

争主动权有着极其重要的意义。建立海防远程视频监控系统，对关键口岸、哨所和敏感地区实施监控，就能使情报部门直观、及时地监视边海防前线情况，提高情报获取的实时性和综合处理能力，也能有效防止偷渡、出逃、走私和贩毒等非法行为。毋庸置疑，在 2008 年的北京奥运会中，计算机视频监控系统将大放异彩，它将广泛用于智能化的交通调度、现代化的体育场馆和优雅舒适的奥运村，为参加奥运会的各国朋友在北京的比赛、游览提供安全舒适的服务。

1.2 国内外研究现状

目标的目标跟踪是基于对一个图像序列的研究，从序列图像中监测到运动的目标，并对目标运动的规律加以预测，实现对指定的目标进行准确且连续的跟踪。以下讨论目标的分割与跟踪的简单分类^{[1][2]}：

按照检测与跟踪的先后可分为：先跟踪后检测（Track-before-Detect）和先检测后跟踪（Detect-before-Track）两种方法。

在远距离搜索与监视过程中，多半是使用先跟踪后检测的方法。如 CCD(Charge Coupled Device)宽场望远镜搜索或跟踪天空中的流星、卫星或其他运动目标，用记载或地面红外（电视）搜索跟踪系统搜索远距离目标等。目标距离传感器较远，它们在传感器中的成像为点目标，相对于背景和传感器噪声而言，目标强度较弱，基本上被噪声所淹没。由于没有形状、大小、纹理等特征，无法用传统的图像处理技术从单帧图像中检测到目标，必须采用基于目标运动特征的序列图像处理办法，边检测边跟踪，设法使目标航迹积累能量，提高信噪比，以达到抑制噪声、目标跟踪的目的。反之，可采用先检测后跟踪的策略。先检测识别目标，再利用目标特征建立模式，然后对目标状态进行预测，搜索匹配实现跟踪。

按照目标在视野中的大小比例可分为：点探测跟踪和成像探测跟踪两种探测跟踪方式。点探测跟踪能提供的信息有点辐射能量，点的位置及其变量，通常应用于短距离目标；成像探测跟踪能提供的信息有目标的辐射亮度及其分布，并进一步感知目标的形状特征及矩描绘特征，进而确定目标的位置及其变量。因此，成像跟踪方式较点跟踪方式所能感知的信息要丰富的多，精确的多。在自然干扰及人工干扰情况下，成像跟踪可以根据其丰富的信息量去除干扰的影响以及提高探测跟踪的精度。

一般来说，运动目标的分割与跟踪就是通过计算有传感器所获得的图像与包含的参考图之间的相似性，确定目标当前的位置、运动参数、空间结构，并给出相应的跟

踪过程。因此，运动目标的分割与跟踪是以图像处理技术为核心、有机融合了计算机技术、传感器技术、模式识别、人工智能等多种理论和技术的新型的目标分割跟踪技术。它是图像处理技术的直接应用。上世纪 50 年代初，GAC 公司就为美国海军开发研制了自动地形识别跟踪系统 (ATRAN) [3]。图像信息的获得和处理是图像跟踪的基础。进入 70 年代以来，随着相关理论与技术的不断发展，图像跟踪技术无论从理论研究、还是从应用研究上都取得了巨大的进展。在学术研究方面：自 W.Meger 和 G.Drius 报道了有关二维相关计算以来 [4]，Anuta、Blom、Blatnea、Maurer 等人对图像跟踪研究领域的新思想、新方法、新进展作了系统而全面的论述 [5]。到上世纪 70 年代初期，自适应跟踪、智能跟踪的思想方法相继提出。1980 年，由美国国防预研局 (DARPA) 和联合军种委员会 (JSG&CC) 联合发起并成立了自动跟踪识别工作组 (ATRWG) [6]。它主要负责图像数据的标准化，为自动目标识别跟踪制定统一的规格，每年定期举行三次会议。许多国际性刊物，如：IEEE on AES, IEEE on PAMI, IEEE on AC, Pattern recognition, Proceedings of SPIE 等成了许多专家、学者学术交流的重要园地。在国际上已经为图像跟踪的研究提供了良好的环境，并取得了可喜的成就。我国对图像跟踪技术的研究起步较晚，但是近年来，随着图像处理技术、计算机技术、特别是大规模集成电路技术 (VLSI)、超大规模集成电路技术 (HVLSI) (TMS320 系列芯片, Transputer 芯片等)、以及各种传感器技术 (如红外、雷达、激光、毫米波、微波等) 等的迅速发展，再加上国际上的影响、国内政府的支持，为图像跟踪技术的研究创造了优越的环境、提供了必要的理论基础和技术条件。目前，一些高等院校、科研院所都已经开展了此项研究工作。

目前，国外的视频跟踪技术相对成熟，已经有了一些雏形系统，比如卡内基梅隆大学领导的视频安全与控制 (VSAM) [7] 研究计划。根据这个计划，研究人员研制了一个端到端的测试系统，集成了很多高级视频安全监控技术，比如静止背景和运动背景下的实时目标探测与跟踪，普通目标 (比如人、轿车、卡车) 的分类识别，特殊物体 (比如校车等具有特殊标记的物体) 的分类识别，目标姿势估计，摄像机的自主控制，多摄像机协同跟踪，人体步法分析等等。

此外，马里兰大学的实时监控系 [8] W4 可以利用单摄像头对人体以及人体的各个部分进行实时地跟踪。所谓 W4，是指 Who, When, Where, What，也就是说这个系

统可以确定目标是谁，什么时间、什么地点，他在干什么。而 IBM 等大公司也在资助这个领域内的相关研究，期望能将研究成果应用到商业领域中。

视频跟踪对运动人体的检测与跟踪主要包括，图像的预处理，运动目标的检测，运动目标跟踪三个方面。

1.3 图像的预处理

对于视频采集到的图像，由于其受到环境等因素的干扰，会产生噪声信号，需要对采集的原始图像先进行预处理，再实现目标的提取。图像的预处理主要指图像的去噪，传统的去噪方法有中值滤波，均值滤波，Donoho 软门限去噪，Lee 滤波等等^[9]，经过去噪处理后图像更加平滑，对于后续的处理工作起到至关重要的作用。

1.4 目标检测

运动检测的目的是从序列图像中将变化区域从背景图像中提取出来。运动区域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，因为以后的处理过程仅仅考虑图像中对应于运动区域的像素。然而，由于背景图像的动态变化，如天气、光照、影子及混乱干扰等的影响，使得运动检测成为一项相当困难的工作。

1.5 目标跟踪

跟踪等价于在连续的图像帧间创建基于位置、速度、形状、纹理、色彩等有关特征的对应匹配问题，常用的数学工具有卡尔曼滤波 (Kalman Filtering)、Condensation 算法及动态贝叶斯网络 (Dynamic Bayesian Network)等。其中 Kalman 滤波是基于高斯分布的状态预测方法，不能有效地处理多峰模式(Multi-mode)的分布情况;Condensation 算法是以因子抽样为基础的条件密度传播方法，结合可学习的动态模型，可完成鲁棒的运动跟踪。目前，就跟踪对象而言，有跟踪如手、脸、头、腿等身体部分与跟踪整个人体的，就跟踪视角而言，有对应于单摄像机的单一视角、对应于多摄像机的多视角和全方位视角，当然还可以通过跟踪空间(二维或三维)、跟踪环境(室内或户外)、跟踪人数(单人、多人、人群)、摄像机状态(运动或固定)等方面进行分类。

1.6 技术难点

尽管视觉监控已经取得了一定的成果，但下面几个方面仍是技术上的难点，快速准确的运动分割是个相当重要但又是比较困难的问题。这是由于动态环境中捕捉的图像受到多方面的影响，比如天气的变化、光照条件的变化、背景的混乱干扰、运动目标的影

子、物体与环境之间或者物体与物体之间的遮挡、甚至摄像机的运动等，这些都给准确有效的运动分割带来了困难。就以运动目标的影子为例，它可能与被检测的目标相连，也可能与目标分离。在前者情况下，影子扭曲了目标的形状，从而使得以后基于形状的认识方法不再可靠；在后者情况下，影子有可能被误认为场景中一个完全错误的目标。尽管目前图像运动分割主要利用背景减除方法，但如何建立对于任何复杂环境的动态变化均具有自适应性的背景模型仍是相当困难的问题。一个可喜的发展是，一些研究者们正利用时空统计的方法构建自适应的背景模型，这也许对于不受限环境中的运动分割而言是个更好的选择。目前，大部分人的运动分析系统都不能很好地解决目标之间互遮挡和人体自遮挡问题，尤其是在拥挤状态下，多人的检测和跟踪问题更是难于处理。遮挡时，人体只有部分是可见的，而且这个过程一般是不可训练的，简单依赖于背景减除进行运动分割的技术此时将不再可靠，为了减少遮挡或深度所带来的歧义性问题，必须开发更好的模型来处理遮挡时特征与身体各部分之间的准确对应问题。另外，一般系统也不能完成何时停止和重新开始身体部分的跟踪。当然，可喜的进步是利用统计方法从可获得的图像信息中进行人体姿势、位置等的预测；不过，对于解决遮挡问题最有实际意义的潜在方法应该是基于多摄像机的跟踪系统。

1.7 面临的问题

虽然对目标检测的援救已经经历了几十年的时间,但是到目前为止面临很多不能很好解决的问题:

1、在运动目标分割过程中，背景的提取与更新是比较关键的步骤。如果背景静止不变,则运动目标分割相对容易实现。然而，多数情况下，背景是动态变化的，常存在物体的移入或移出，如果不能很好的处理背景中的干扰现象，则作为背景的运动物体的监测必然将受到影响。

2、视频场景中光线变化问题(渐变、突变、反光)同样会影响背景的提取与更新。必须不断地对背景进行实时估计，解决亮度变化问题，才能更好的完成运动目标的检测。

3、光照条件下，目标存在阴影。与背景不同，运动目标的阴影将随运动目标移动。检测分割的结果可能是阴影与目标的出现。

4、遮挡也是运动目标检测过程中一个难以解决的问题，在运动目标前方的遮挡物很可能会作为目标的一部分被提取出来，从而造成监测目标形变.严重时会造成运动目

标检测的失败。

5、背景物体的干扰运动。背景中存在的如风中树叶的摇动,水面波光的闪动,车窗玻璃的反光以及天气的变化等许多细微活动,同样会影响到运动目标的检测。

6 另外目标运动过程是否连续,速度如何,是否存在往返运动的现象等问题也关系到运动目标的检测。

2 图像预处理

在实际应用中,由于图像在微光 CCD 摄像系统成像的过程中不可避免地受到各种因素的干扰,以及周围环境的影响,使得所摄取的图像质量较低,对这样的图像直接进行目标的提取和跟踪是比较困难的。因此,我们需要对图像进行预处理,减少或滤除各种噪声和随机干扰,增强有用信息,提高后续处理的有效性和可靠性,为图像分割创造良好的条件。在对运动图像序列进行处理之前,应对其进行预处理,目的在于提高图像质量、增强有用信息。

2.1 图像增强

由于 CCD 摄像系统的影响,或者是摄像机周围环境的影响等因素,使得摄像系统摄取的图像存在一些问题,如图像的灰度分布不均匀,整帧图像的灰度范围较小,这样图像显得对比度不够、模糊不清,目标与周围景物之间的灰度差别很小,这样的图像直接进行目标的提取和跟踪将比较困难,而且使得目标的轮廓不清,这样我们需要预先对图像进行图像增强处理,图像增强技术是用于改善图像视感质量所采取的一种较通用的方法。通常要完成的工作是除去图像中的噪声,使边缘清晰以及突出图像中的某些性质等。图像增强技术发展比较迅速,图像增强处理的目的是改善图像的视觉效果,将图像转换为一种更适于人或机器分析的形式。

图像增强技术可使图像灰度级的动态范围增大,使得轮廓比较清晰,对比度增大。使被跟踪目标和干扰与背景的灰度区别增大,易于将被跟踪目标和干扰从背景图像中分离出来进行处理。下面介绍几种图像增强的方法:

2.1.1 对比度增强

照片或电子图像常见的毛病是对比度不理想,这由于图像亮度不够或非线性而引起的。对每一灰度值重新分度,改善图像对比度,从而改善图像的视觉效果。若令 Γ 代表原图像灰度值, S 代表增强后图像的灰度值 $S=T(\Gamma)$ 。利用这种变换技术可使原图像中灰度小于 M 或大于 N 的部分对比度减小,而在 $M-N$ 区域内的对比度增加,从而使得增强后的图像视觉效果更好。

2.1.2 直方图均衡化

直方图表示数字图像中各灰度级与其出现的频数间的统计关系。一幅自然图像的线性量化图,各灰度值的概率通常是不一致的,有些分布集中在较暗的区域,图像具有较暗的特性,因此暗区的细节常常看不出来;有些反之。这类图像增强的一种方法

是所谓的直方图修正技术，直方图修正是图像增强的最常用、最重要的方法之一。利用图像灰度分布（直方图）信息，对灰度分布形式作校正来修正图像灰度，最终达到图像增强的目的。因此它是基于度量空间（灰度分布）的灰度修正方法。采用直方图修正可使原图像灰度级集中的区域拉开或使灰度分布均匀，从而增大反差，使图像的细节清晰，达到增强的目的。增强后的图像量化误差增大，但视觉效果却得到了改善。

直方图修正中具有最优性质的是直方图均衡。图像直方图均衡处理就是用灰度变换函数对输入图像直方图进行修正，修正后的图像直方图趋向均匀分布。

设 r_k 为原始图像的第 k 级灰度， s_k 为变换后的图像的第 k 级灰度，那么图像直方图均衡的具体方法如下：

首先计算原始图像的灰度级直方图，用 $n(i)(i=0,1,\dots,255)$ 表示。

计算灰度级变换函数 $T(r_k)$

$$T(r_k) = \sum_{i=0}^k \frac{n(i)}{N} \quad k = 0, 1, \dots, 255 \quad (2.1)$$

其中 N 是图像的像素总数。

然后根据灰度级变换函数 $T(r_k)$ 完成图像的灰度级变换。

由于 $\frac{n(i)}{N}$ 只是图像灰度级概率密度函数的近似，因此经过直方图均衡后，只能得到接近平坦的直方图。

2.1.3 实时中值滤波快速算法的并行实现

滤波能增强有用信号，提高后续处理的有效性和可靠性，为图像分割创造良好的条件。目前最常用的图像去噪工具是均值滤波器和中值滤波器。它们分别是线性滤波器和非线性滤波器的代表。其中均值滤波器主要用来抑制高斯噪声，而中值滤波器则主要用来抑制脉冲噪声。均值滤波器，算法简单，易于实时处理，且对高斯噪声有较好的平滑能力。但均值算法会破坏图像边缘，且对脉冲噪声十分敏感，没有充分利用图像像素间的相关性和像素的位置信息。中值滤波器具有能够彻底滤除尖波干扰噪声，同时又有能较好保护目标图像边缘的特点。

2.2 算法介绍

中值滤波是由 Tukey 首先提出的一种典型的非线性滤波技术^[10]。由于它较好地解

决了消除脉冲干扰和保持信号边缘问题，在图像滤波中得到广泛应用。

标准一维滤波器的定义为：

$$y_k = med \{x_{k-N}, x_{k-N+1}, \dots, x_k, x_{k+N-1}, x_{k+N}\} \quad (2.2)$$

其中：*med* 表示取中值操作中值滤波的滤波方法是对滑动滤波窗口 $(2N+1)$ 内的像素作大小排序，滤波结果的输出像素值规定为该序列的中值。例如取 3×3 滑动窗口，中值为窗口内第五个最大的像素值。

中值的计算在于对滑动窗口内像素的排序操作。要进行排序，就必须对序列中的数据像素作比较和交换，数据元素之间的比较次数是影响排序速度的一个重要因素。传统的排序串行算法是基于冒泡排序法，若窗口内像素为 m 个，则每个窗口排序需要作 $m(m-2)/2$ 次像素的比较操作，时间复杂度为 $o(m^2)$ 。此外，常规的滤波算法使窗口每移动一次，就要进行一次排序，这种做法实际上包含了大量重复比较的过程。若一幅图像的大小为 $M \times N$ ，则整个计算需要 $o(m^2 N^2)$ 时间，当窗口较大时计算量很大，较费时。

针对 3×3 中值滤波，采用了一种快速的并行滤波方法，通过巧妙设计，避免了大量的重复比较操作，每一窗口排序需要 $o(m)$ 时间，整个计算需要 $o(mN^2)$ 时间，易于实时处理。

快速算法及其实现：为便于说明，将 3×3 窗口内的 9 个像素分别定义为

$$P_{00}, P_{01}, P_{02}, P_{10}, P_{11}, P_{12}, P_{20}, P_{21}, P_{22}$$

表 1 3×3 窗口内像素排列

	第 0 列	第 1 列	第 2 列
第 0 行	P_{00}	P_{01}	P_{02}
第 1 行	P_{10}	P_{11}	P_{12}
第 2 行	P_{20}	P_{21}	P_{22}

首先对窗口内的每一列分别计算最大值,中值和最小值.这样就得到三组数据,分别为最大值组,中值组和最小值组.计算过程表示如下：

$$\begin{aligned}
 \text{最大值组: } & P_{\max 1} = M \max [P_{01}, P_{11}, P_{21}] \\
 & M \max = \{P_{\max 0}, P_{\max 1}, P_{\max 2}\} \\
 \text{中值组: } & P_{\max 1} = Med [P_{01}, P_{11}, P_{21}] \\
 & Med = \{P_{med 0}, P_{med 1}, P_{med 2}\} \\
 \text{最小值组: } & M \min_1 = M \min \{P_{\min 0}, P_{\min 1}, P_{\min 2}\}
 \end{aligned} \tag{2.3}$$

(2.3) 公式中 \max 表示取最大值操作， med 表示取中值操作， \min 表示取最小值操作。由此可以看到，最大值组中的最大值与最小值组中的最小值一定是 9 个像素中的最大值和最小值。除此，中值组中的最大值至少大于 5 个像素：本列中的最小值和其它两列的中值和最小值；中值组中的最小值至少小于 5 个像素：本列中的最大值和其它两列的最大值和中值。同样，最大值组中的中值至少大于 5 个像素，最小值组中的中值至少小于 5 个像素。即最大值组中的最小值为 $p_{M \max \min}$ ，中值组中的中值为 $p_{Med \text{ med}}$ ，最小值组中的最大值为 $p_{M \max \max}$ 中的中值。这一计算过程表示如下：

$$\begin{aligned}
 p_{M \max \min} &= \min [P_{\max 0}, P_{\max 1}, P_{\max 2}] \\
 p_{Med \text{ med}} &= \text{med} [P_{med 0}, P_{med 1}, P_{med 2}] \\
 p_{M \min \max} &= \max [P_{\min 0}, P_{\min 1}, P_{\min 2}] \\
 p_{M \min \text{ med}} &= \text{med} [P_{\max \min}, P_{Med \text{ med}}, P_{\min \max}]
 \end{aligned} \tag{2.4}$$

采用该方法中值的计算仅需作 19 次比较，与传统算法相比，比较次数减少了近 2 倍，且该算法十分适用于在 FPGA 上作并行处理。

2.3 预处理结果与分析

我们取 3×3 正方形中值滤波的窗口，效果如下图 1 所示为原始图像，经快速中值滤波和均值滤波处理后的结果。显而易见，快速中值滤波平滑掉了原图中的点状噪声，而且边缘清晰度比均值滤波的效果好。



(a) 原始图



(b) 均值滤波效果图



(c) 快速中值滤波效果图

图 1

2.4 噪声消除

在背景区域中被错分的像素表现为零散的噪声点或小的噪声块，而在前景区域中会使人体的区域出现空洞现象，这些对于判断前景物体是否为人类和对人体的颜色特征进行提取都会产生不利影响。所以在分割步骤以后还要对二值化映射图进行去噪声和补空洞的处理步骤。这里采取数字图像处理中的数学形态学处理的开、闭运算来完成这个工作。我们可以通过调节膨胀、腐蚀模板的尺寸大小来控制去噪声的力度，它的值的大小可根据场景和噪声的具体情况确定。

2.5 形态滤波

数学形态学是一门建立在严格数学理论上的学科，其基本运算包括腐蚀(Erosion)，膨胀(Dilation)，开(opening)运算和闭(Closing)运算。用这些算子及组合来进行图像结构和形状的分析 and 处理，包括图像分割，特征抽取，边缘检测，图像滤波，图像增强和恢复等方面的内容。形态学是从数学理论的集合角度去定义分析的，有关数学的理论这里不再累述，只把形态学在图像滤波中的应用总结如下。

2.5.1 腐蚀

腐蚀算子在对二值图像处理时可以采用不同的腐蚀元，3x3 窗口或 5x5 窗口或更大的窗口，下面表 2 以 3x3 窗口为例说明该算子的算法。从二值图像的起始位置开始，取一 3x3 的窗口，只有当窗口内的 9 个像素点的灰度值都为 0 时，中间点的像素值才为 0，如图(d)、(e)所示，如果 9 个像素点中有一个以上的点的值为 255，则中心点的像素值为 255，如图(b)、(c)所示。该算法可以用公式描述如下：

$$X(i, j) = \begin{cases} 0 & a = b = \dots = i = 0 \\ 255 & else \end{cases} \quad (2.5)$$

其中， $X(i, j)$ 为窗口中心点的像素。

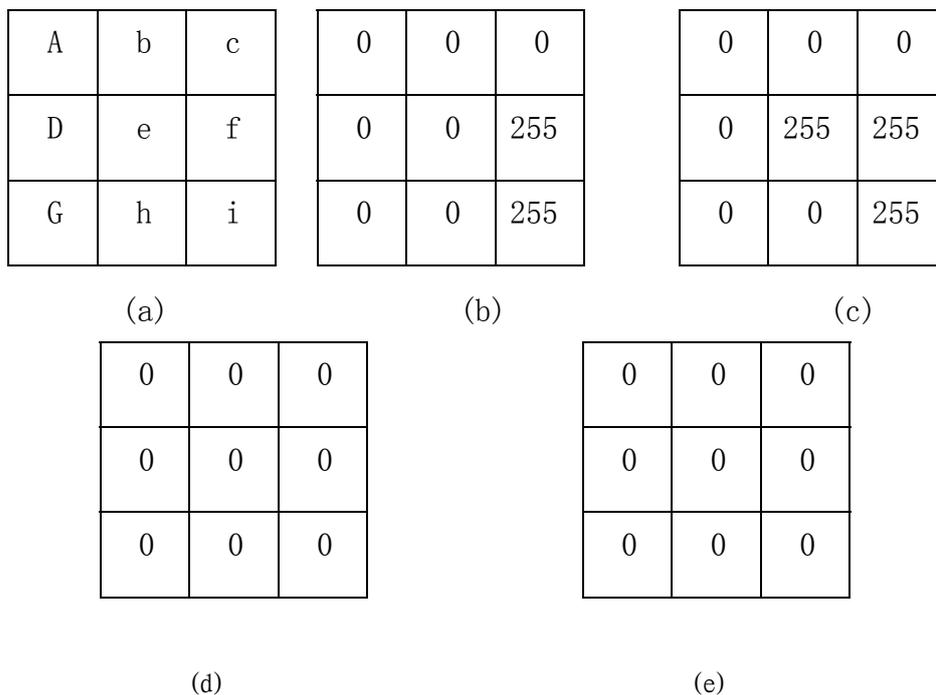


表 2 腐蚀算子滤波演算图

(a) 为图像某位置处像素的表示 (b) 为该位置处像素的实际灰度值 (c) 为腐蚀处理后 c 点的像素值 (d) 为另一位置处像素的实际灰度 (e) 为该点腐蚀处理后 c 点的像素值

腐蚀的一种最简单的用途是从二值图像中消除不相关的细节，从上面的分析可以看出，经过腐蚀算法后的图像中的小的黑色区域被去掉，只保留连续的较大的目标区域，从而达到滤波的作用。不同窗口的腐蚀元具有不同的滤波效果，虽然较大的窗口可以得到更好的滤波效果，但是太大的窗口会使图像损失过大，实验表明，采用 3x3 窗口或 5x5 窗口比较合适。

2.5.2 膨胀

膨胀算子在对二值图像处理时也可以采用不同的膨胀元，通常采用 3x3 窗口或 5x5 窗口，下表 3 以 3x3 窗口为例说明该算子的算法。从二值图像的起始位置开始，取一 3x3 的窗口，只有当窗口内的 9 个像素点的灰度值都为 255 时，中间点的像素值才为 255，如图 (d) (e) 所示，如果 9 个像素点中有一个以上的点的值为 0，则中心点的像素值为 0，如

图 (b) (c) 所示。该算法可以用公式描述如下：

$$X(i, j) = \begin{cases} 255 & a = b = \dots = i = 0 \\ 0 & else \end{cases} \quad (2.6)$$

其中， $X(i, j)$ 为窗口中心的像素值。

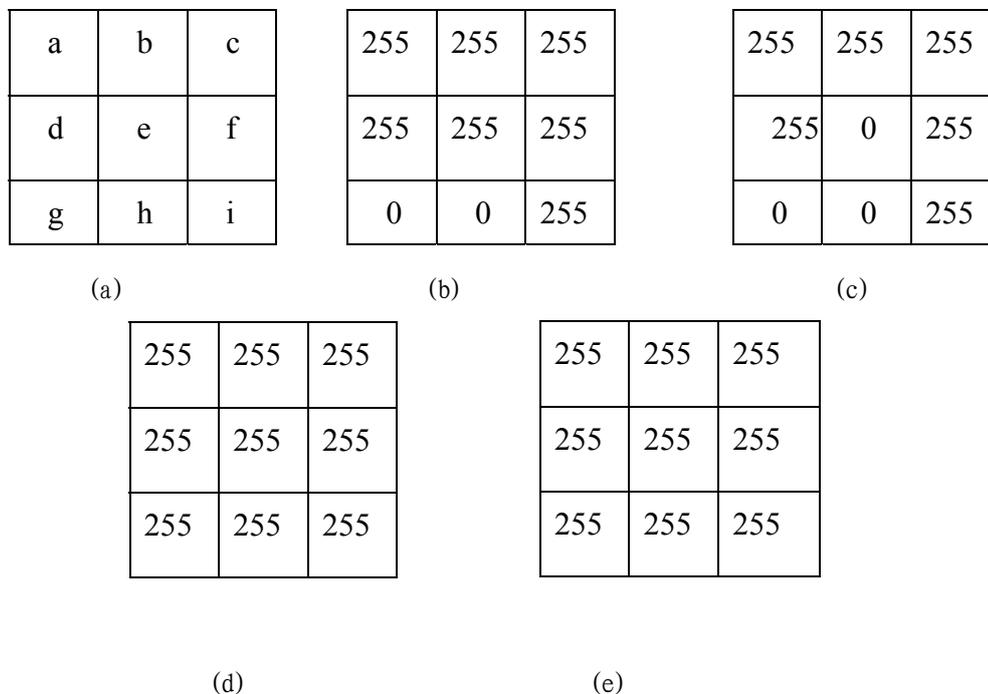


表 3 膨胀算子滤波演算图

(a) 为图像某位置处像素的表示 (b) 为该位置处像素的实际灰度值 (c) 为膨胀处理后 c 点的像素值 (d) 为另一位置处像素的实际灰度 (e) 为该点膨胀处理后 c 点的像素值

膨胀最简单的应用之一是将裂缝桥接起来。从上图的分析可以看出，二值图像经膨胀算法后可以把原来断开的部分连起来，这种算法虽然可以增大目标检测的区域，但是也把噪声增强了。实验表明，采用的窗口越大，目标的断开点越少，但是窗口选择过大时引入的噪声也大大增强，实际中采用 3x3 窗口或 5x5 窗口比较理想。

2.5.3 开运算

前面已经分析过，腐蚀使图像缩小而膨胀使图像扩大，而有时单纯的腐蚀或膨胀不能满足我们的要求，如我们希望图像在去噪后再进行断点桥连等，这时需要引入开操作和闭操作。开操作使对象的轮廓变得光滑，断开狭窄的间断和消除长细的突出物。闭操作同样使轮廓更为光滑，但与开操作相反的是，它通常用于填充狭窄的间断和长细的鸿沟，消除小的孔洞，并填补轮廓线中的断裂。

开操作的基本过程是先对图像进行腐蚀处理，然后再对腐蚀的结果进行膨胀，而闭操作正好是一个相反的过程，先对图像进行膨胀，然后对膨胀的结果腐蚀。

图 2 给出了形态学滤波的结果，其中，第一列表示原始图像，第二列为差分图像，第三列为腐蚀处理后的差分图像，第四列为开运算后的结果。

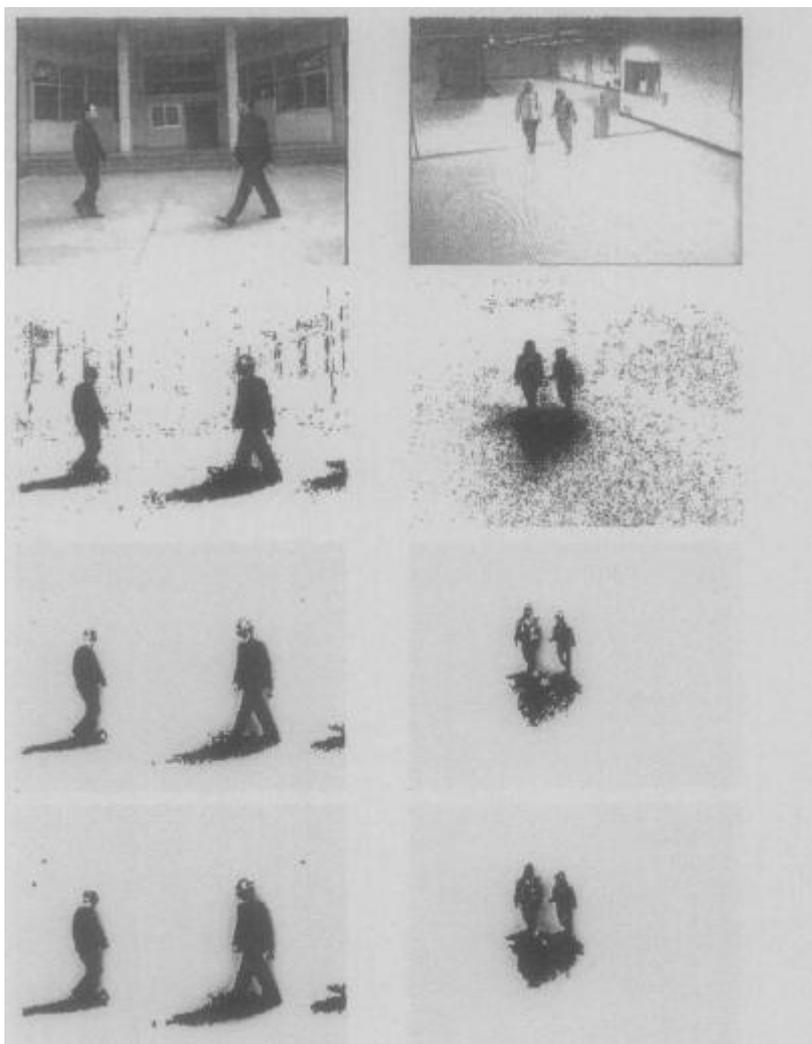


图 2 采用形态学方法对差分图像进行处理的结果

从上图中可以看出，采用开操作后噪声被减小或完全被抑制，同时对目标区域出现的断点也基本得到桥连。

2.6 阴影消除

阴影消除是运动目标精确检测的主要挑战之一。当分割和提取运动目标时，因为阴影具有两个重要的视觉特征：阴影显著地不同于背景而被理解为前景和阴影与产生阴影的目标具有相同的运动属性。而且视频中检测前景目标大多数使用帧间差分 and 背景差分方法，所以阴影点将被误检测为目标点。在检测运动目标时，如不考虑阴影的影响，将可能导致如下问题：

- (1) 场景内目标的融合(阴影可能连接若干独立目标)；

(2) 目标几何特征的扭曲(光照等外部条件引起的阴影变化将影响目标形状);

(3) 目标的丢失(阴影遮住了其他目标);

(4) 虚假目标的出现(目标与其阴影分离时, 阴影被检测为目标);

这些都将直接影响到目标分类、识别、跟踪等后续处理。目前阴影检测方法主要有基于模型和基于颜色特征的方法。基于模型的方法假设目标的三维形状和光源属性已知, 通过模型可以精确地计算出阴影的形状和位置, 但这在实际应用中是不现实的。基于颜色特征的方法主要根据目标和阴影不同的光学和颜色特征来分离目标和阴影, 是目前常用的检测阴影的方法。

3 目标物的检测

3.1 运动目标检测

运动检测的目的是从序列图像中将变化区域从背景图像中提取出来。运动区域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，因为以后的处理过程仅仅考虑图像中对应于运动区域的像素。然而，由于背景图像的动态变化，如天气、光照、影子及混乱干扰等的影响，使得运动检测成为一项相当困难的工作。下面介绍目前几种常用的方法。

(1) 背景减除法

背景减除方法是目前最常见的一种运动目标检测的方法。它的思想是通过当前图像与背景图像之间的差分来检测图像中的变化区域实现对运动目标的检测。这种方法实现起来比较简单，但一般只适用于背景静止的场景中，而对动态背景以及一些外来无关事件的干扰较为敏感。另外即便是背景保持静止，由于环境光照强度的变化，影子的干扰以及环境噪声的影响等使得这种方法在实际运用中面临着许多鲁棒性问题。背景减除法的核心是如何获得一个较为准确的背景图像，目前最简单的背景模型是帧平均图像，另外现在许多研究人员也开发出其他的背景模型，例如^{[11] [12] [13]}

(2) 时间差分法

时间差分法指的是在动态图像序列中，利用相邻的两帧或三帧间的差分来实现运动目标的检测，例如^[14]。这种方法一般也能快速的从背景中分割出运动区域，并且对动态环境有较强的自适应性，对环境噪声以及光线强度等干扰相对与背景减除法有着较强的抗干扰性，但一般不能获得运动变化区域，在运动实体内部容易产生空洞现象，需要通过进一步的分析处理实现运动目标的完全检测。

(3) 光流法

光流法的历史见较长，是早期计算机视觉研究的比较多的方法之一，他的基本思想是构造光流方程，求解各个像素的光源，根据运动目标与背景随时间变化的光流人、特性的不同，将运动目标与背景进行区分，达到目标检测的效果，如^[15]。这种方法的优点是在摄像机运动时仍然可以有效的实现运动目标的检测。然而光流法的运算一般都比较复杂，即便是现在的计算机处理水平也难以满足实际应用中的实时性的要求，所以光流法在没有相关硬件设备的保证支持下很难被应用于实时性要求较强的实际问

题中。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性，如 Meyer 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，从而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，大多数的光流计算方法相当复杂，且抗噪性能差，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。

(4) 运动能量法

运动能量法把连续图像看作由二维空域加上时间维构成的三维空间，计算空时梯度，运动对象经过位置的像素空时梯度一致性高，梯度能量大，这种特性可以用于运动对象检测。运动能量法适合于复杂变化的环境，能消除背景中振动的像素，使按某一方向运动的对象更加突出的显示出来，但运动能量法分割出对象不够精确，而且方法过于复杂，不利于实现。在运动变化检测中还有一些其他的方法，如 Friedman 与 Russe 利用扩展的 PEMIZ(Peetition Maximization)算法，为每个像素建立了混合高斯分类模型，该模型可以自动更新，并能自适应地将每个像素分类为背景、影子或者运动前景，在目标运动速度缓慢的情况下亦能较好地完成运动区域的分割，并可以有效地消除影子的影响；另外，Stringa 也提出了一种新颖的基于数学形态学的场景变化检测算法，在变化的环境条件下获得了相对稳定的分割效果。

3.2 目标分类

目标分类的目的是从检测到的运动区域中将对应于人的运动区域提取出来。不同的运动区域可能对应于不同的运动目标，比如交通道路上监控摄像机所捕捉的序列图像中可能包含行人、车辆及其他诸如飞鸟、流云、摇动的树枝等运动物体，为了便于进一步对行人进行跟踪和行为分析，运动目标的正确分类是完全必要的。注意，这个步骤在一些情况下可能是不必要的(比如已经知道场景中仅仅存在人的运动时)。下面介绍两种常用的目标分类方法。

(1) 基于形状信息的分类(Shape-based Classification)

基于形状信息的分类是利用检测出的运动区域的形状特征进行目标分类的方法。例如 VSAMF 采用区域的分散度、面积、宽高比等作为特征，利用三层神经网络方法将运动目标划分为人、人群、车和背景干扰；Lipton 等^[16]利用分散度和面积信息对二维运动区域进行分类，主要是区分人、车及混乱扰动，时间一致性约束使其分类更加准确；Watanabe^[17]使用简单的人体轮廓模式的形状参数从图像中检测运动的人。

(2)基于运动特性的分类(Motion-base Classification)

基于运动特性的分类是利用人体运动的周期性进行目标分类的方法。例如 Cutler 与 Davis^[18]通过跟踪感兴趣的运动目标,计算出目标随着时间变化的自相关特性,而人的周期性运动使得其自相关也是周期性的,因此通过视频化方法分析目标是否存在周期性的运动特性而将人识别出来:Lipton^[19]通过计算运动区域的残余光流(Residual Flow)来分析运动实体的刚性和周期性,非刚性的人的运动相比于刚性的车辆运动具有较高的平均残余光流.同时它呈现了周期性的运动特征,据此可以将人区分出来。

4 目标物的跟踪

4.1 运动目标的跟踪技术

运动目标跟踪的目的就是通过对传感器拍摄到的图像序列进行分析，计算出目标在每帧图像上的位置，给出目标速度的估计。可靠性和精度是跟踪过程中的两个重要指标。对运动对象的跟踪较复杂，需要考虑每帧图像背景和对象的变化。可能有四种情况：从当前帧到下一帧，映射的目标完全被包含在实际运动目标的区域内；映射的目标大部分被包含在实际运动目标的区域内，小部分超出相应的运动区；出现了一个新的对象覆盖了映射的运动目标；实际的运动目标从屏幕上消失，而映射的运动目标还存在。跟踪算法的思想有：把在当前帧中分割出的目标轮廓根据运动信息投影到下一帧，作为下一帧中运动目标分割的初始轮廓；或者把当前帧中的目标轮廓与下一帧中提取的粗边沿图像进行匹配，以便确定跟踪的目标；或者根据当前帧中目标的特征合并下一帧中分割的各个区域，从而得到跟踪的目标。已有的跟踪算法各有千秋，不同的跟踪算法适用于不同运动状态的目标。目前主要从以下几个方面解决复杂背景下的目标跟踪问题。

4.1.1 基于 Hausdoff 距离的运动跟踪

使用 Hausdoff 距离跟踪视频运动目标是个很有效的方法^[20]。这种方法使用二进制模板进行比较，即使当前目标的形状有较大的变化或背景移动时，也能很精确的检测目标新的位置。使用 Hausdoff 距离方法来跟踪视频运动目标的四个主要优点：混乱的图像背景对其影响不大；计算简单且速度快；允许仅适用图像的部分轮廓进行比较；允许运动对象的不连续运动。使用 Hausdoff 距离跟踪器关键是在初始时如何提取二进制运动目标模板，即如何分割出运动目标和背景。把得到的视频对象的二值化模板用于在后记帧中寻找相匹配的模板。为了适应模板的形状改变，每帧都对模板进行更新。Hausdoff 跟踪器对刚体目标跟踪较好，而对非刚体目标跟踪较差，而且只能跟踪一个运动对象。

4.1.2 基于区域的运动跟踪

在进行刚体运动目标跟踪时采用的算法分为四步：首先根据运动轨迹的光滑性进行运动预测；其次采用把参数透视模型进行运动估计；再用前一帧运动估计的参数对当前帧图像进行反向变换，把对应前一帧目标边界的像素作为当前帧中对象的边界；最后根据估计的边界产生内、外两个初始界，并采用形态学分水岭算法得到真实的边

界。文献^[21]采用修改的形态学分水岭的区域跟踪算法，通过对上一帧区域进行运动估计和投影，在下一帧重新提取标记，然后再使用分水岭算法进行分割，该方法也可以检测新出现的目标。

采用交互的方式，按照跟踪算法的性质将跟踪分为基于区域的刚体跟踪和基于区域的非刚体跟踪。基于区域的非刚体跟踪算法为：得到上一帧中目标的边界；估计边界像素的运动；重新提取当前帧的准确边界；生成目标的模板。这种方法简单，但要求对象和背景间的对比度较强。基于区域的刚体跟踪算法为：首先进行参数的运动估计；提取目标模板；再进行参数的修正；最后进行目标边界分割。这种方法可以处理目标和背景间的对比度弱的情况，以及对象之间出现局部遮挡的情况，但运动目标必须是刚体运动。

4.1.3 相关跟踪

在众多的成像跟踪算法中，相关跟踪算法具有对场景图像质量要求不高，可在信噪比低的条件下稳定工作，能适应较复杂场景结构的目标和背景条件，具有较强的局部抗干扰能力等特点。因此，相关跟踪算法在成像跟踪中有着重要地位，特别适用于地面背景等复杂场景条件下的目标跟踪。针对相关跟踪算法运算量大的不足，出现了许多改进算法，如基于塔型结构的匹配跟踪、多子模板匹配、Kalman 滤波器。

4.1.4 光流跟踪

光流跟踪运用目标的运动信息，避免了灰度变化对目标跟踪的影响，因而具有较好的抗噪能力。光流分析可以分为连续光流法和特征光流法。特征光流法是通过特征匹配求得特征点出的光流，通过对前后帧图像中的角点匹配来估计运动目标的位移并进行跟踪。特征流法的主要优点在于：对目标在帧间的运动的限制较小，可以处理大的帧间位移；对噪声的敏感降低，只处理图像中很少数的特征点，计算量较小。主要缺点是：得到的是稀疏光流场，难以提取运动目标的精确状态；特征匹配问题尚未得到较好的解决。全局光流的计算方法有 Horn-Schunck 算法、Nagel 算法等^[21]，得到全局光流场后通过比较运动目标与背景之间的运动差异对运动目标进行光流分割，从而达到跟踪的目的。运用全局光流场来跟踪运动目标的关键是要得到精确的光流估计，这就依赖于施加合理的平滑约束条件。

4.1.5 基于网格的运动跟踪

提出基于 2-D 网格的运动跟踪算法分为五步：上一帧的边界作为下一帧的粗边界；

按照最大对比度约束条件查找新的准确的目标边界；新的未覆盖区域的边沿用 2-D 网格和搜索算法重新提取，然后根据运动矢量、彩色信息和边沿信息指定为背景、原对象一部分或新出现的目标，对新出现的目标也进行跟踪；每帧需要更新网格结构。当遮挡区域超过一个给定比值，分割过程重新初始化。

这种方法的优点为：可跟踪刚体或非刚体目标；可跟踪多个目标；可处理目标间的遮挡、目标的消失和新目标的出现。缺点有：边界对比度不能太弱；分割出的目标边界不够准确。

4.2 人体运动状态估计

在图像序列中跟踪运动物体，实际上就是在帧间把已知的运动物体与在新一帧中检测到的运动区域匹配起来。所以跟踪主要是处理每一帧中已知的运动物体与下一帧的运动区域相匹配的问题。在匹配的过程中可能出现下列几种情况：

1、在新一帧中出现了一个前景运动区域，但是该运动区域不能与上一帧中任何已知物体相对应。在这种情况下，把这个前景运动区域作为新出现的运动物体来处理。

2、一个已知的运动物体不能与新一帧中任何前景运动区域相对应，那么这个物体可能已经出了观测区域，或者是被背景区域遮挡住，所以不能被观察到。这时认为该运动物体已经丢失。

3、一个已知的运动物体与唯一的前景运动区域相对应。这种情况在跟踪中是最好的情况，这时就认为新一帧中检测到的前景运动区域就是上一帧中已知的运动物体。这样我们就可以利用运动物体在新一帧中得到的观测值来更新描述该运动物体的状态向量。

4、多个已知运动物体与一个运动前景区域相对应。这种情况可能是由多个单独人体相互遮挡形成人群造成的。人体在场景中的运动状态是不断变化的，为了跟踪人体的运动就要对人体的运动状态进行估计，卡尔曼滤波器是一种递推线性最小方差估计，具有估计快速准确的优点。本文使用卡尔曼滤波器对人体的运动状态进行估计，并且根据估计出的运动状态对人体运动进行跟踪。本文针对卡尔曼滤波器发散的原因进行了分析，并且根据不同的发散原因采用了不同的抑制措施。在本文中采取组合的方法抑制卡尔曼滤波器的发散，在实际应用中取得了较好的效果。

4.3 卡尔曼滤波理论

4.3.1 滤波理论基础

在数字信号处理中，滤波器的作用是把有用信号从噪声中提取出来。它的信号差分模型一般为： $x(n) = s(n) + v(n)$

其中， $x(n)$ 是输入信号， $s(n)$ 是有用信号， $v(n)$ 是噪声。设滤波器的单位函数响应为 $h(n)$ ，则输出为：

$$y(n) = \sum_m h(m)x(n-m) \quad n=1,2,3,\dots n \quad (4.1)$$

这里希望 $y(n)$ 尽量接近于 $s(n)$ ，所以称 $y(n)$ 为 $s(n)$ 的估计值。实际上，(4.1)式的卷积形式可以理解为从当前和过去的观测值 $x(n), x(n-1), x(n-2)\dots$ 来估计有用信号的当前值 $s(n)$ ，用 $\hat{s}(n)$ 表示。因此，用滤波器进行滤波的问题，可以看成是一个估计问题。

一般的估计方法有：

- 1) 从当前和过去的观测值来估计信号的当前值，即 $y(n) = s(n)$ ，称为滤波；
- 2) 从过去的观测值来估计信号的现在值或将来值，即 $y(n) = s(n+N)$ $N > 0$ ，称为预测或外推；
- 3) 从过去的观测值来估计过去的信号值，即 $y(n) = s(n-N)$ $N > 1$ ，称为平滑或内插。

4.3.2 卡尔曼滤波器的特点

滤波器的种类有多种，上述是典型的维纳数字滤波器。它的优点是物理概念清楚，但缺点是不能递推计算，也不适于时变系统。设计维纳滤波器需做功率谱分解，只有当被处理信号为平稳的、干扰信号和有用信号均为一维、且功率谱为有理分式时，维纳滤波器的传递函数才可用伯特—香农设计法较容易地求解出。

卡尔曼滤波器不仅同时克服了上述两个缺点，而且只用信号的前一个估计值和最近一个观测值就可以在线性无偏最小方差估计准则下估计信号的当前值，而不必像维纳滤波器需要全部过去的观测值。

卡尔曼滤波从与被提取信号有关的量测中通过算法估计出所需信号。其中被估计信号是由白噪声激励引起的随机响应，激励源与响应之间的传递结构(系统方程)已知，

量测量与被估计量之间的函数关系(量测方程)也已知。在估计过程中,利用如下已知信息:系统方程、量测方程、白噪声激励的统计特性、量测误差的统计特性。卡尔曼滤波是在时域内设计的,所用的信息也都是时域内的量。并且卡尔曼滤波器适用于多维。因而卡尔曼滤波的适用范围无维纳滤波的范围广,卡尔曼滤波有如下特点:

- 1、卡尔曼滤波处理的是随机信号;
- 2、被处理信号无有用和干扰之分,滤波的目的是要估计出所有被处理信号;
- 3、系统的白噪声激励和量测噪声不是要滤除的对象,它们的统计特性正是估计过程中需要利用的信息。

卡尔曼滤波一般根据物理系统的系统方程和量测方程的性质可分为线性的和非线性的,每一种又可分为连续的和离散的。非线性卡尔曼滤波又称为扩展卡尔曼滤波,这时,系统方程是非线性的,或者系统方程和量测方程都是非线性的时候,本文中建立的系统方程和量测方程都是线性的并且是离散的。

4.3.3 卡尔曼滤波发散的原因和抑制

在卡尔曼滤波计算中,经常会发生这样一种现象:当量测值数目 k 不断增大时,按滤波方程计算的估计均方误差阵趋于零或趋于某一稳态值,但估计值相对实际的被估计值的偏差却越来越大,使滤波器逐渐失去估计作用,这种现象称为滤波器的发散。导致滤波器发散的主要原因有两点:

- 1、描述系统动力学特性的数学模型和噪声的统计模型不准确,不能真实的反映物理过程,使模型与获得的量测值不匹配,导致滤波器发散。这种由模型过于粗糙或失真引起的发散称为滤波发散。

- 2、卡尔曼滤波是递推过程,随着滤波步数的增加,舍入误差逐渐积累。这种积累误差可能使估计的均方误差阵失去非负定性甚至失去对称性,使增益阵的计算值逐渐失去合适的加权作用而导致发散。这种由计算的舍入误差积累引起的滤波发散称为计算发散。

在实际应用中,系统方程可能不够准确,这样就能够造成卡尔曼滤波器发散。并且,如果进行长时间监控,随着时间的推移,以前得到的量测值对现在状态估计的影响越来越小,但是卡尔曼滤波器对过去的量测值的影响不断积累,也将导致卡尔曼滤波器的发散。由于上述原因造成的滤波器发散,可以通过加大当前量测值的加权系数,从而相应地降低早期量测值的影响来抑制。本文通过衰减记忆法来抑制这种滤波器的发散。对于舍入误差积累造成的卡尔曼滤波发散,可以通过平方根滤波来抑制。

5 结 论

本文主要阐述了视觉跟踪领域中运动人体目标检测与跟踪的相关知识背景及技术流程，在三个方面作了细致的分析，并且在图像预处理及动态背景目标检测方面提出了新的算法，通过大量的实验验证了新方法的可行性。

计算机视觉跟踪技术是个热门的研究方向，不过它的发展到现在还不是很完善，还有很多技术难点等待我们去解决：

1、阴影的处理。在人体目标检测中我们可以看到，由于光照影响，在目标周围还有大面积的阴影存在，如果不给予去除将会给后面对个体的位置及行为的判断上带来很大障碍。不少研究者在这方面也做了一些工作，比如文献[22]中基于灰度投影来判断影子，并利用人体的对称原理加以消除。利用色度来判别阴影的方法，这些方法的计算量较大导致实时性较差。因此在这方面还需要做更加细致的工作。

2、跟踪过程中，多目标的出现常常会相互遮挡，导致跟踪的错误。现在的方法也有很多，比如基于颜色的差别、基于高度的不同等等。但实际应用中还要考虑它的实时性，普遍性。在利用 kalman 对目标进行行为的估计时，没有考虑到行为忽然变化的情况，因此类似的问题需要更深入的研究。

总之，运动人体跟踪的应用领域是非常广泛的，适用性非常强，他对我们智能安全控制、人机交互管理的实现起到非常重要的作用，希望自己有机会在这方面做更多的工作。

谢 辞

衷心的感谢导师贺长伟对我的热情指导和精心帮助。半年来的大四学习期间，她渊博的知识、敏捷的思维、平易近人的态度使我终生受益。本论文无论从理论方面还是实验方面都得到了贺老师的细心指导和大力支持，所以本论文无不凝聚着贺老师心血和关怀。在此衷心地感谢贺老师！同时也感谢山东建筑大学图书馆给我提供了很多的参考资料以及网络资源。

本文的顺利完成与那些所有关心和帮助过我的老师和同学是分不开的，在此一并表示我衷心地感谢！感谢父母多年来对我的支持！

参考文献

- [1] 李智勇,沈振康,杨卫平等.动态图像分析.北京:国防工业出版社,1999年.
- [2] 高文,陈熙霖.计算机视觉.北京:清华大学出版社,1999.
- [3] R.F.Roch,D.C.Evans.Atr an terrain sensing guidanceThe grand-daddy system.Pro Of the society of Photo-optical Instrumentation Engineers,Vol.238,Image Processinfor Missile Guidance,Ssan Siego California,1950,July:2-9.
- [4] W.Meger-Eppler and G.Drius.Two dimensional photographic autocorrelation opictures and alphabet letters.In Proc.3rd London symposium on information theory C.cherry,Ed.,Academic Process,New York,1956:34-36.
- [5] chen h.,shen z.,liz.New target tracking method for optical/IR Image sequencConference Proc.Of SPIE,1997,Apr.
- [6] Narendra P.M.,Westover B.L.Advanced target tracker concepts.1980,AD-A101845.
- [7] Collins R,Lipton A,Kanade T.Introduction to the Special Section on Video Surveillance.IEEE Tans.Pattern Analysis and Machine Intelligence.2000,22(8):745-746.
- [8] Haritaoglu I,Harwood D,Davis L.W4:Real-time surveillance of people and the iractivities.IEEE Transations on Pattern Analysis and Machine Intelligence.2000,22(8):809-803.
- [9] Julier S,Uhlmann J.Unscented Filtering and Nonlinear Estimation.PROCEEDINGS OF THE IEEE.2004,3,92(3):401-422.
- [10] Gordon N,Salmond D,Smith A.Novel approach to nolinear/non-Gaussian Bayesian state estimation.IEE PROCEEDINGS-F.1993,4,140(2):107-113.
- [11] Isard M,Blake A.Condenstation-Conditional Density Propagation for Visual Tracking.International Journal of Computer Vision.1998,290(2):5-28.
- [12] 章毓晋. 图象分割. 科学出版社.
- [13] J.Barron, n.Fleet, s.Beauchemin.“Peorfimrnaeeofopticalnowteehniques Intenrational Journal of ComPuter ComPuterVision, 12(1), 1994:20 — 77.
- [14] Lipton A,Fujiyoshi H,Patil R.Moving Targrt Classifivation and Travking from Real-time-video,Proc IEEE workshop on Application of Computer Vision,Vision,Princeton,NJ,1984; 8-4.
- [15] Meyer D,Denzler J,Niemann H,Model Based Extraction of Articulated Objects in Image Sequences for Gait Analysis,Proc IEEE International Conference on Image Processing Santa Barbara,California,1997:78-81.
- [16] Huang P,Harris C,Nixon M,Human Gait Recongnition in Canonical Space Using Templates,Image Signal Process,1999,146(2):93-100.
- [17] Mao Xia, et al. Analysis of Affective Characteristics and Evaluation of Harmonious Feeling of Image Based on Fluctuation Theory[A] . International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems (IEA/AIE) 2002:17-19.

- [18] 贾云得.机器视觉.北京:科学出版社,2004 年.
- [19] Zhou O,Aggarwal J.Tracking and classifying moving objects from videos.In:Proceedings of IEEE Workshop on Performance Evaluation of Tracking and Surveillance,Hawaii,USA,2001.
- [20] Friedman N,Russell S.Image segmentation in video sequences:A probabilistic approach In of the Conference on Uncertainty in Artificial Intelligence,Rhode Island,USA,1997:175-181.
- [21] Gao X,Boult T,Coetzee F,et al.Error analysis of background adaption.In:Proceedings ofIEEE International Conference on Computer Vision and Pattern Recognition,Hilton Head Isand ,SC,USA, 2000:503-510.
- [22] 侯志强,韩崇昭.基于像素灰度归类的背景重构算法.软件学报,2005,16(9):1568-1576.