

分类号\_\_\_\_\_

密级\_\_\_\_\_

UDC \_\_\_\_\_

编号\_\_\_\_\_

# 中国科学院研究生院

## 硕士学位论文

视频目标跟踪中的特征评估算法研究

韩 振 军

指导教师\_\_\_\_\_焦建彬教授 博导\_\_\_\_\_

\_\_\_\_\_中国科学院研究生院

申请学位级别工学硕士 学科专业名称计算机科学与技术

论文提交日期2009 年 4 月 论文答辩日期2009 年 5 月

培养单位\_\_\_\_\_中国科学院研究生院

学位授予单位\_\_\_\_\_中国科学院研究生院

答辩委员会主席\_\_\_\_\_

# 独 创 性 声 明

本人郑重声明：所提交的学位论文，是本人在指导教师的指导下，独立进行研究工作所取得的研究成果。尽我所知，文中除特别标注和致谢的地方外，学位论文中不包含其他人或集体已经发表或撰写过的研究成果，也不包含为获得中国科学院研究生院或其它教育机构的学位或证书所使用过的材料。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

签名：\_\_\_\_\_日期：\_\_\_\_\_

## 关于学位论文使用授权的说明

本人完全了解中国科学院研究生院有关保管、使用学位论文的规定，其中包括：①学校有权保管、并向有关部门送交学位论文的原件与复印件；②学校可以采用影印、缩印或其它复制手段复制并保存学位论文；③学校可允许学位论文被查阅或借阅；④学校可以学术交流为目的,复制赠送和交换学位论文；⑤学校可以公布学位论文的全部或部分内容。

（涉密的学位论文在解密后应遵守此规定）

签名：\_\_\_\_\_导师签名：\_\_\_\_\_日期：\_\_\_\_\_

## 摘要

视频图像中的运动目标跟踪作为计算机视觉领域的核心研究课题之一,经过近 50 多年的发展得到了广泛而深入的研究。它融合了图像处理、模式识别、人工智能、自动控制等相关领域的研究成果,在视频监控、机器人导航、视频传输、视频检索、医学图像分析、气象分析等许多领域有着广泛的应用。本文以此为研究内容,具有重要的理论意义和广阔的应用价值。

本文提出一种通过实时调整目标特征权值来进行背景自适应跟踪的算法。首先,定义了一种综合特征集合用以描述目标的颜色和局部轮廓。其次,提出了在滤波框架中对目标特征进行评估的算法,从而使得具有强区分能力的特征占有较大的权值,进而使其能够在跟踪过程起到较大的作用。采用传统的 Kalman 滤波和粒子滤波对所提出的算法进行了验证。

本文的主要贡献和研究内容如下:

1) 一种面向跟踪的目标表征综合特征集。基于颜色直方图 (Histogram of Color, HC) 和梯度方向直方图 (Histogram of Oriented Gradient, HOG) 的综合直方图融合了目标的颜色、局部轮廓、梯度方向等特征,固定尺度的综合直方图和 Scale Invariant Feature Transform(SIFT)特征点具有相近的描述能力。同时,利用 SIFT 特征点的主方向方法改进了 HOG 特征的提取,在一定程度上解决了 HOG 对目标旋转敏感的缺点。

2) 基于时间连续性的特征权值评估方法。在每帧视频中,通过当前视频帧中的跟踪目标特征区分目标本身与相关背景的能力以及前一视频帧中跟踪目标特征的权值对特征进行重新评估。将这种特征权值评估的方法嵌入到传统的滤波框架中,从而可以保证特征权值的变化是个连续的过程,即具有时间连续性。这也是本文的主要贡献。

3) 对传统滤波算法应用的扩展。基于滤波框架的特征权值评估有效地将滤波算法从对目标运动状态的建模扩展到了对目标特征重要性的评估,拓展了传统滤波框架的应用范畴。

**关键词:** 目标跟踪、特征评估、Kalman 滤波、粒子滤波

## Feature Evaluation for Adaptive Visual Object Tracking

**Han Zhen-jun**(Computer Application)

Directed by: **Jiao Jian-bin**(Professor)

Moving objects tracking is one of the most important branches in the computer vision, which combines advanced technologies and research achievements in image processing, pattern recognition, artificial intelligence, automatic control and other relative fields. It has widely applied in video surveillance, robot navigation, video transmission, video retrieval, medical image analysis, meteorological analysis, and other fields. Therefore, this subject has important theoretical significance and wide practical value.

In this paper, we propose a new adaptive visual object tracking method based on online feature evaluation approach. First, a feature set is built by combining color histogram (HC) with gradient orientation histogram (HOG), which emphasizes both color and contour representation. Then a feature confidence evaluation approach is proposed to make features with higher confidences play more important roles in the instantaneous tracking ensuring that the tracking can adapt to the appearance change of both the object and its background. The feature evaluation approach is fused with filter frameworks, e.g. Kalman and Particle filter, to keep the temporal consistency of feature confidence evolution.

The contributions of this paper are as follows.

**1) A combined feature set for adaptive object tracking** The basic idea of using the combined feature set is that an object can often be well characterized by both the color (HC) and the contour (HOG) representation or either of them. The combined feature set is the evolvement of color, edge orientation histograms and Scale Invariant Feature Transform (SIFT) descriptors. The improvement is that we propose an approach to reduce the orientation sensitivity of the HOG features by calculating the dominant orientation of the object.

**2) A novel feature evaluation approach of temporal consistency** In each frame, the proposed evaluation approach adjusts the confidence of each feature based on the discriminative abilities of the feature bins in current video frame and the confidence of each feature in the previous video frames. The evaluation is formulated in the traditional filter frameworks so that the discriminative abilities of current frame and the temporal consistency from previous video frames are well comprised. This is also the main advantage of our method compared with existing feature evaluation

methods.

**3) Function extension of traditional tracking filters** The proposed feature evaluation approach extends the function of traditional filter framework from modeling motion states to modeling feature evaluation problem. To our knowledge, this is the first research of such function extension.

**Key Words:** Moving Object Tracking, Feature Evaluation, Kalman Filter, Particle Filter

# 目录

摘要 .....	I
Abstract .....	I
目录 .....	IV
图目录 .....	VI
表目录 .....	VII
<b>第一章 绪论</b> .....	<b>1</b>
1.1 课题来源、背景和意义 .....	1
1.1.1 课题来源 .....	1
1.1.2 课题背景及意义 .....	1
1.2 视频跟踪的国内外研究现状 .....	2
1.3 视频跟踪的主要研究内容和分类 .....	5
1.4 本文的研究内容 .....	7
1.5 本文的组织结构 .....	7
<b>第二章 视频跟踪特征评估算法基础</b> .....	<b>9</b>
2.1 视频跟踪的主要算法 .....	9
2.1.1 基于表达和相似性度量的运动目标跟踪算法 .....	9
2.1.2 基于滤波预测方法的运动目标跟踪算法 .....	12
2.1.3 基于优化匹配搜索的运动目标跟踪算法 .....	14
2.2 跟踪特征评估算法 .....	18
<b>第三章 基于 Kalman 滤波的特征评估算法</b> .....	<b>21</b>
3.1 Kalman 滤波算法及其应用 .....	21
3.1.1 Kalman 滤波算法基本原理 .....	21
3.1.2 Kalman 滤波算法在目标跟踪中的应用 .....	23
3.2 基于 Kalman 滤波的特征评估算法研究 .....	24
3.2.1 综合特征的提取 .....	24
3.2.2 基于 Kalman 滤波的特征评估 .....	27
3.3 实验结果比较与算法性能分析 .....	29
3.3.1 视频测试集合 .....	29
3.3.2 实验结果分析及比较 .....	29
3.4 跟踪算法在测试系统中的应用 .....	35
<b>第四章 基于粒子滤波的特征评估算法</b> .....	<b>37</b>
4.1 粒子滤波算法及应用 .....	37
4.1.1 粒子滤波算法原理 .....	37
4.1.2 粒子滤波算法在目标跟踪中的主要应用 .....	38
4.2 基于粒子滤波特征评估 .....	38
4.3 实验结果比较与算法性能分析 .....	40

总结与展望.....	47
参考文献.....	49
发表文章目录.....	53
致谢.....	55

## 图目录

图 1-1 本文研究的目标跟踪框图.....	6
图 2-1 Mean Shift 示意图.....	15
图 3-1 跟踪目标及其背景区域.....	25
图 3-2 HOG 特征提取.....	26
图 3-3 利用 HOG 主方向方法对目标进行旋转的示意图.....	27
图 3-4 基于 Kalman 滤波的特征权值调整示意图.....	29
图 3-5 相对简单背景的跟踪结果图.....	30
图 3-6 复杂背景的跟踪结果图.....	32
图 3-7 四种跟踪算法的 DER 比较结果图.....	33
图 3-8 本章所提出算法的跟踪结果图.....	34
图 3-9 本章所提出算法在基于 PTZ 摄像头系统的应用.....	35
图 4-1 基于粒子滤波的特征评估方法示意图.....	40
图 4-2 相似背景中的目标跟踪结果.....	41
图 4-3 存在遮挡和相似目标的跟踪结果.....	42
图 4-4 四种跟踪算法的 DER 比较结果图.....	43
图 4-5 本章所提出跟踪算法的跟踪实验结果.....	44



## 表目录

表 3-1 基于 kalman 滤波的特征权值评估算法.....	28
表 3-2 用于计算 DER 的视频文件列表.....	33
表 4-1 基于粒子滤波的特征权值评估算法.....	39



## 第一章 绪论

### 1.1 课题来源、背景和意义

#### 1.1.1 课题来源

本课题来源于国家自然科学基金（课题编号：60672147、60872143）以及中国科学院“百人计划”资助课题。

作者在论文工作期间，参与了以下课题：

- 1) 百人计划“复杂环境下动态目标检测及跟踪技术研究”；
- 2) 国家自然科学基金“复杂环境下新型多目标动态检测、识别及跟踪算法研究”；
- 3) 国家自然科学基金“视频中的人体目标检测和跟踪关键技术研究”；
- 4) 南京军区徐州某部“弹药库智能监控系统”；
- 5) 航天科工集团哈尔滨风华机器厂“‘天目’全景摄像预警控制系统开发”。

#### 1.1.2 课题背景及意义

计算机视觉与专家系统、自然语言理解一起成为人工智能最活跃的三大领域。基于视频的运动目标跟踪技术（以下简称为：视频跟踪或者目标跟踪）作为计算机视觉领域的核心研究课题之一，主要目的是模仿生理视觉系统的运动感知功能，通过对摄像头捕捉到的图像序列进行分析，计算出运动目标在每一帧图像中的二维坐标位置；然后，根据运动目标相关的特征值，将图像序列中连续帧间的同一运动目标关联起来，得到每帧图像中目标的运动参数以及相邻帧图像间运动目标的对应关系，从而得到各个运动目标完整的运动轨迹，即在连续的视频序列中建立运动目标的对应关系。简单来说，目标跟踪技术就是在下一帧图像中找到目标的确切位置并反馈给跟踪系统进行跟踪，为视频序列分析和理解提供重要的基础运动信息和数据。

视频跟踪技术融合了图像处理、模式识别、人工智能、自动控制以及计算机等众多领域中的先进技术和核心思想。与传统的雷达系统相比，视频跟踪系统采用摄像机或者其它光学系统等比较廉价的设备，采用被动式工作模式（即工作时不向外辐射无线电波，不易被敌方的电子侦察设备发现），因此具有更高的性价

比、一定的隐蔽性和抗电子干扰的能力。“眼见为实”的视频跟踪系统，可以从视频监视器上直接看到运动目标图像，获得关于运动目标更丰富的信息，因此具有很好的直观性和可靠性。

经过近 50 多年的深入研究和发 展，视频跟踪技术已在军事制导、视觉导航、安全监控、智能交通、医疗诊断以及气象分析等方面广泛应用。例如，可见光跟踪和红外跟踪技术已应用于导弹制导、火炮控制、空间飞行体和靶场测量等。视频监控技术已在大多社区、大型公共场所及重要设施使用；进行车辆的实时检测和跟踪，可以实时监控车流量、车速、车流密度、交通事故、违章逃逸车辆等交通状况，用于实时的智能交通调度。另外，视频跟踪技术在视频会议、视频分析、视频检索、基于视频的运动分析和合成、基于运动信息的身份识别、图像检索、水文观测、港口管理、医学图像分析、远距离测量、零部件质量检测等许多领域也有大量应用。

然而，就目前来说，一般意义上的视频跟踪技术还远未成熟，要开发出真正可靠、实用的视觉跟踪应用系统还需要开发更为鲁棒的核心算法。在运动目标通过摄像机的检测区域过程中，运动目标的图像可能发生明显的变化，这些变化主要来自于四个方面：1) 运动目标的尺度伸缩的变化；2) 运动目标姿态和形状的变化；3) 运动目标被其他目标或者背景部分或者完全遮挡；4) 环境光线亮度的变化。这些情况的存在使得对运动目标的跟踪变得更加困难。现有的目标跟踪算法虽然能够在一定程度上完成对运动目标的跟踪，但这些算法大多是基于目标某一种特征集合的跟踪，其主要缺陷是对目标特征描述不完备，并且无法自动适应目标和环境的变化，因此当背景发生较大变化时，往往会失效。在本研究中，长时间复杂背景中的鲁棒性跟踪是关键问题，也是难点问题。欲解决此问题需要解决：1) 如何更好地提取目标的完备特征集合，从而使得对目标的描述更加准确与鲁棒；2) 如何使目标的特征描述更好地适应目标和环境的变化。目前我们的研究主要集中在这两个方面。

## 1.2 视频跟踪的国内外研究现状

国外对于视频跟踪理论研究起步较早。美国自然科学基金委员会及美国军方对复杂环境下多目标的动态检测、识别及跟踪的算法的研究及其应用非常重视。美国自然科学基金多次资助有关公司进行相关算法研究，美国国防部

(Department of Defense, DoD)、美国国防部高级研究项目署 (Defense Advanced Research Project Agency, DARPA)、美国陆军夜视及电子探测设备委员会(Night Vision and Electronic Sensors Directorate, NVESD)、美国陆军 SBIR(Small Business Innovation Research)项目, 美国海军水下战事中心(Navy Underwater Warfare Center, NUWC)等部门也先后投巨资资助该类算法的开发及其相关应用研究并取得了可观的成果。早在 20 世纪 50 年代初期, GAC 公司就为美国海军研制开发了自动地形识别跟踪系统 (Automatic Terrain Recognition and Navigation System, ATRAN)。近几年, 自适应跟踪和智能跟踪的思想被相继提出, DARPA 成立了自动目标识别工作组 ATRWG (Automatic Target Recognition Work Group) 进行目标智能识别和跟踪等相关研究。

许多国际学术会议和期刊也包含了这一领域的最新研究成果, 如: Int. Conf. on Computer Vision and Pattern Recognition (CVPR)、Int. Conf. on Computer Vision (ICCV)、European Conf. on Computer Vision (ECCV)、Asian Conf. on Computer Vision (ACCV)、Int. Conf. on Pattern Recognition (ICPR)、Int. Journal of Computer Vision (IJCV)、IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)、IEEE Trans. on Image Processing、Pattern Recognition、Image and Vision Computing 等。

随着图像处理和分析技术、计算机技术和各种图像传感器技术的飞速发展, 国内一些高校和科研院所已经逐渐开展这方面的研究, 并且已经取得了喜人的成绩。如: 中科院计算所先进人机通信技术联合实验室对于体育视频中的目标分割和跟踪、精彩片断分析的研究[9][13]; 中科院自动化所模式识别国家重点实验室图像和视频分析研究组开发了人体运动的视觉分析系统、交通行为事件分析系统、交通场景监控系统和智能轮椅视觉导航系统[16]; 清华大学开发的适用于自然环境中的视觉侦查系统具有运动检测、跟踪和全景图生成等功能; 西安交通大学的人工智能与机器人研究所使用光流和帧差两种算法对自适应巡航控制中的车辆跟踪进行了深入的研究等。

对于目标进行跟踪, 可以获取目标的移动速度、移动方位甚至更高的语义特征。近年来研究人员开发了大量新的跟踪算法[1-31], 同时, 一些传统的跟踪方法也不断得到改进[16]。这些算法大致分为确定性方法和随机方法两大类。

典型的确定性跟踪方法在对感兴趣目标进行跟踪时,根据事先设定的相似性度量函数,在当前视频帧局部区域内,通过搜索或者迭代算法找到与目标模板或者目标表观最为相似的区域。在该类算法中,人们广泛使用误差平方和(Sum of Squared Difference, SSD)相似性度量函数与全局搜索算法实现目标模板匹配。最近的研究中, Mean Shift 等优化算法被广泛用于寻找局部最优区域 [17-22]。但是, Mean Shift 算法对带宽选择和跟踪过程初始化的苛刻要求限制了其应用。Bouttefroy 等人采用了一种新的视频跟踪技术,使用投影 Kalman 滤波的 Mean Shift 改进了传统 Mean Shift 的不足[22]。对于特定的目标,人们将目标先验知识,如表观、颜色分布、轮廓信息等用于目标模板的建立,从而使跟踪算法更为鲁棒、有效。其中基于表观(Appearance-based)的方法可以统计学习跟踪目标在特征空间(eigenspace)或核空间(kernel space)内的分布规律 [17-20]。

随机跟踪方法利用状态空间对当前跟踪系统的运动进行模型化。当系统的运动模型呈线性或高斯分布且观察模型也为线性时,其后验概率密度函数的模式比较简单,此时可以采用 Kalman 滤波方法对后验概率分布进行时序上的传递与更新。然而,对于非线性、非高斯的问题,由于目前尚未找到准确的解析描述,因此,人们提出一些近似解决方法。粒子滤波(Particle Filter),也被称为时序蒙特卡罗(sequential Monte Carlo) [20, 27],是目前应用最为广泛的算法 [22]。随机跟踪方法引入了概率统计的思想,将不确定性观察(如概率密度函数)与不同的状态相联系,从而不再假设运动系统输入与输出的完全确定性。对于运动系统而言,由于目标自身及相互之间存在遮挡、高速运动等不确定性因素,因此,随机跟踪方法更适用于大多数基于视觉特征的目标跟踪。

确定性的跟踪方法是基于假设观察模型可以很方便地建立且具有较好的精度来实现的,因此,这种方法经常是运用简单的规则或推理来得到目标的当前状态。但这种方法却忽略了最终得到的假设或状态存在不确定因素的事实。而基于随机的方法引入了概率统计的思想,将不确定性观察(如概率密度函数)与不同的状态相联系,从而不再假设运动系统输入与输出的完全确定性。对于运动系统而言,由于目标自身及相互之间的遮挡、高速运动等不确定性因素的存在,因此随机跟踪方法更适用于大多数的基于视觉特征的目标跟踪。目前,基于粒子滤波的目标随机跟踪方法已被大量使用,所采用的模型有轮廓模型、颜色模型和表面

模型等。

### 1.3 视频跟踪的主要研究内容和分类

现有的视频运动目标跟踪算法，虽然在一定程度上能够完成对运动目标的跟踪，但都存在着一定的问题，例如，一些算法的跟踪的精度不高，或者是跟踪精度有所提高，但特征提取和匹配的计算复杂度也随之增大，难以满足实时处理的要求。如何提高运动目标跟踪算法的实时性、鲁棒性和精确性，是目前视频跟踪算法研究的热点和难点。

以下一些问题也是视频跟踪中需要重点解决的难题：

- 1). 运动目标图像的三维重建。三维空间的运动目标变换为二维图像，在这个过程中会丢失重要的深度信息；
- 2). 图像的外界影响。现实中图像成像模型受到目标自身的材质、外界光源、摄像机的视角和距离、透明度和遮挡等诸多因素的影响，但图像像素值与这些因素之间的依赖关系往往是非线性的；
- 3). 运动目标的部分及完全遮挡。背景对目标的遮挡以及目标之间的相互遮挡是视频跟踪系统中的又一个难题，遮挡是造成目标的图像表达突然变化的重要原因之一，而且这种变化具有突然性和不连续性，从而很容易引起跟踪算法的失效；
- 4). 运动目标的尺度伸缩变化和形变。在跟踪的过程中，目标的运动会引起与摄像机的距离和视角变化，造成摄像机所捕捉到的目标图像的尺度伸缩以及目标图像表达等产生非线性的变化，从而导致跟踪的失效；
- 5). 场景光线亮度变化。同一目标在不同的光线亮度条件下会产生有明显差异的目标图像，该影响是很多室外实际应用系统中阻碍性能提升的瓶颈；
- 6). 目标复杂的运动模式。跟踪算法主要采用滤波预测算法，在运动目标可能出现的区域内进行全局或者局部搜索，但当目标具有复杂的运动模式时，如目标的运动速度或者方向发生突然改变，使得运动目标不在预测的搜索区域内，这样简单滤波预测算法就会失效，进而造成跟踪目标丢失。

视频跟踪算法可以按照不同的标准进行分类。根据运动目标和背景的相对运动关系不同，视频跟踪算法可以分为：动态背景的运动目标跟踪算法、动态背景的静止目标跟踪算法、静止背景的运动目标跟踪算法、静止背景的静止目标跟踪

算法。本文以前三类算法为研究重点，因此在没有特别说明的情况下，本文所提到的跟踪算法均为基于运动目标或者运动背景的跟踪算法。根据跟踪目标的个数不同，视频跟踪算法可以分为单目标跟踪算法和多目标跟踪算法。本文主要研究单目标跟踪算法，但该跟踪框架很容易扩展到多目标跟踪中。根据跟踪目标的表达方法和相似性度量方法不同，视频跟踪算法可以分为基于轮廓的跟踪算法、基于特征的跟踪算法、基于区域统计特性的跟踪算法和基于模型的跟踪算法等。根据跟踪目标属性的不同，视频跟踪算法可以分为：刚体目标跟踪算法和非刚体目标跟踪算法。

目前一些跟踪算法部分是针对提高跟踪匹配的准确性而提出的，这类算法的主要特点是通过选取稳定的、可区分性好的特征来提高目标的搜索匹配速度和匹配的正确程度；部分算法是针对缩小目标搜索范围提出的，这类算法的主要特点是通过滤波方法预测目标下一时刻可能出现的位置，通过缩小目标的搜索范围来缩短目标搜索时间。

为了开发一种鲁棒实时的单目标跟踪系统来解决运动目标的尺度伸缩、部分及完全遮挡、运动目标的形变以及光线亮度变化等视频跟踪中的开放性难题，进而达到对目标长时间的鲁棒性跟踪，本文提出了基于综合颜色直方图和梯度方向直方图的目标跟踪综合特征并研究了基于传统经典滤波方法和目标跟踪综合特征的特征权值调整算法，对环境噪声具有一定的自适应性，并对目标的尺度伸缩、遮挡以及环境的光线变化等都具有比较好的鲁棒性。本文跟踪算法的研究框图如下：

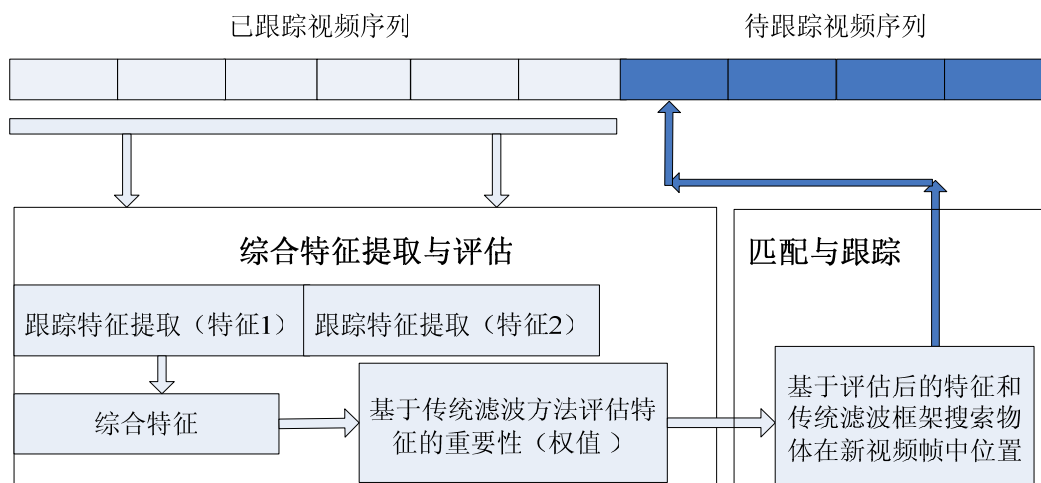


图 1-1 本文研究的目标跟踪框图



## 1.4 本文的研究内容

本文主要提出了基于综合颜色直方图和梯度方向直方图的目标跟踪特征,并在传统滤波算法的框架内,对目标跟踪的综合特征的特征权值进行自适应的调整,旨在提高目标特征的表达能力,使其在目标的尺度伸缩、遮挡以及环境的光线变化等的复杂环境中,仍然具有比较好的鲁棒性。主要研究内容如下:

- **综合直方图的目标描述方法:**以目标本身及其背景为研究对象,提出了一种基于颜色直方图(Histogram of Color, HC)和梯度方向直方图(Histogram of Oriented Gradient, HOG)[34]的综合直方图。同时我们利用 Scale Invariant Feature Transform (SIFT)的主方向方法很好地解决了 HOG 特征对目标旋转比较敏感的弱点。多尺度的综合直方图包含了目标颜色,边缘方向,局部轮廓以及类似于 SIFT 等特征的描述,从而可以很好的在目标尺度伸缩变化、部分遮挡以及形变等情况下对目标进行有效地表达。实验结果表明,基于综合直方图的目标描述方法对视频运动目标的跟踪是非常有效的。
- **基于 Kalman 滤波的目标特征评估方法:**基于目标特征变化与目标运动变化在时间连续性方面的相似性,我们提出了一种基于 Kalman[37][38]滤波方法的特征权值调整新方法。这种特征评估与权值调整方法,通过融合跟踪的历史帧特征权值信息以及当前帧的特征判别力信息,可以保证特征评估在跟踪过程中是一个连续平滑的过程,从而有效地增加了特征权值的连续性和稳定性,提高了跟踪的稳定性和鲁棒性。通过大量的实验验证了该方法的有效性。
- **基于粒子滤波的目标特征评估方法:**由于 Kalman 的局限性(线性、高斯的假设),我们提出了基于一种非线性、非高斯假设的特征评估算法,即基于 Particle[21][39]滤波的特征评估方法。这种特征评估方法即保留了基于 Kalman 滤波特征评估方法的优点,又克服了 Kalman 滤波高斯和线性假设的缺点,因此能够更好地预测和修正目标特征权值的变化,使得对目标的特征评估更加准确。本文通过大量实验验证了算法的有效性。

## 1.5 本文的组织结构

第一章,绪论。主要论述了视频跟踪算法的研究背景和意义,分析了国内外的研究现状以及视频跟踪算法中存在的难题,列出了本文的主要研究目的和研究

内容，总结了本文的主要贡献。

第二章，视频跟踪特征评估算法基础。主要综述了视频跟踪算法的主要研究内容和方法，同时也对视频跟踪过程中的特征权值评估进行了简单的综述。

第三章，基于 Kalman 滤波的特征权值评估算法。主要论述了所提出的目标跟踪的综合特征集以及其提取方法，详细介绍了基于此综合特征集和 Kalman 滤波对特征权值进行评估的算法。最后，给出了该跟踪算法在各种复杂场景下的运动目标跟踪实验，从而对算法在运动目标尺度伸缩变化、部分及完全遮挡和部分形变等情况下的鲁棒性跟踪进行验证，同时将该算法与传统的跟踪算法的性能进行了比较并对比较结果进行了分析。

第四章，基于粒子滤波的特征权值评估算法。主要论述了基于综合特征集和粒子滤波对特征权值进行评估的目标跟踪算法。该算法基于特征权值非高斯、非线性的假设，使得评估之后的特征权值的后验概率更符合当时场景变化的真实情况。最后，给出了该跟踪算法在各种复杂场景下的运动目标跟踪实验，从而对算法在运动目标尺度伸缩变化、部分及完全遮挡和部分形变等情况下的鲁棒性跟踪进行了验证，同时将该算法与传统的跟踪算法的性能进行了比较与分析。

最后是对现有工作的总结和对未来工作的展望，以及对如何进一步提高跟踪算法的鲁棒性、准确性和实时性等研究热点和难点的探讨。

## 第二章 视频跟踪特征评估算法基础

视频跟踪是计算机视觉领域的热点研究问题之一。目前,针对如何提高运动目标跟踪算法的实时性、鲁棒性和准确性等方面的研究逐步成为视觉跟踪研究的热点。本章将主要介绍视频跟踪的一些传统方法,并对基于目标特征评估的跟踪算法进行介绍和综述。

### 2.1 视频跟踪的主要算法

视频跟踪的算法思想通常有两种:通过识别目标来进行跟踪和通过运动目标检测来进行跟踪。前一种算法思想通过识别每一帧的图像中跟踪目标,来确定该运动目标的位置和实现跟踪,该算法思想包含了目标识别和目标匹配两部分;而后一种跟踪算法思想通过检测和发现运动目标并确定运动目标的位置进行跟踪,该方法不需要考虑目标的形状、尺度,可以检测任何目标。

#### 2.1.1 基于表达和相似性度量的运动目标跟踪算法

视频跟踪算法根据跟踪目标的表达方法和相似性度量方法可以大致分为:基于轮廓的跟踪算法[43]、基于特征的跟踪算法[30]、基于区域统计特征的跟踪算法[32][33]和基于模型匹配的跟踪算法[17][42]。

基于轮廓的跟踪算法通过对目标的分割和边缘提取实现对运动目标的跟踪。然而,当目标被部分及全部遮挡时,跟踪算法将受到很大的限制。基于模型的跟踪算法采用目标的物理知识而不仅仅是目标的图像信息建立模型,因此基于 3D 模型跟踪算法是比较完善的。该方法的局限在于无先验知识的实际跟踪系统中很难得到目标的模型。基于特征表达的跟踪算法通过提取具有尺度伸缩、旋转等不变性的特征点(如 SIFT、KLT、Harris Corners、SUSAN Corners 等),解决尺度伸缩和部分遮挡等问题。SIFT[30]特征点主要是通过图像金字塔的多尺度极值点检测得到的,用梯度方向进行描述,因此这种特征点具有尺度和旋转不变性等优点,在图像配准和跟踪方面具有很高的匹配精度和鲁棒性。然而,特征自身的检测和提取往往计算复杂难以满足跟踪系统实时处理的要求。基于区域统计特性的跟踪算法利用的是目标图像的全局描述信息,很难精确定位目标的位置。同时,当背景或者其他目标的区域统计特性相似的时候,跟踪算法也会失效(如基于颜

色或者灰度直方图的跟踪算法)。

基于特征匹配的跟踪方法不考虑运动目标的整体特征,即不关心具体的运动目标,而只通过其特征来进行跟踪。由于图像采样时间间隔通常很小,可以认为这些特征在运动形式上是平滑的,因此可以利用其完成目标的整个跟踪过程。

无论是刚体运动目标还是非刚体运动目标,利用基于特征匹配的方法进行目标跟踪时主要包括特征提取和特征匹配两个过程。在特征提取中要选择适当的跟踪特征,并且在后续的视频帧提取相应的特征;在特征匹配中将提取的当前帧图像中目标的特征与特征模板相比较,根据比较的结果来确定目标,从而实现目标的跟踪。例如使用特征点对人体进行运动跟踪时,首先把需要跟踪的每一个人用一个矩形框封闭起来,在封闭框内进行跟踪特征的提取并建立特征模板,然后在后续的视频帧中提取相应的特征并与特征模板进行匹配,从而通过找到最佳的匹配位置完成跟踪。除了用单一的特征来实现跟踪外,还可以将多个特征信息进行综合作为跟踪特征,来提高跟踪的稳定性。另外,如果运动目标简单,可以将整个目标作为特征模板来进行跟踪,这种方法也被称为模板匹配。

基于区域统计特征的跟踪方法是把图像中运动目标连通区域的共有特征信息作为跟踪特征的一种方法。在图像中常用的共有特征信息如颜色特征、纹理特征等。这种方法不需要在视频序列中找到完全相同的特征信息,而是通过计算候选区域与原始目标之间的区域统计特征的相关性来确定跟踪目标的最佳位置。例如,[57]利用小区域特征进行室内单目标人体的跟踪。该方法将人体看作由头、躯干、四肢等部分所对应的小区域块的联合体,通过分别跟踪各个小区域块最终完成对整个人体的跟踪。同时,基于区域统计特征的跟踪方法常常利用滤波技术来估计和预测区域的几何形状和运动速度。

基于模型匹配的跟踪方法是通过建立模型来表示目标,也就是对运动目标进行建模,然后在图像序列中匹配这个建立好的模型来实现目标的跟踪。然而,在实际应用中跟踪的目标大部分都是非刚体的,其形状在不断发生变化。即使是刚体目标,由于拍摄视角及拍摄距离的不断变化也会导致其形态不停地发生变化,很难得到准确的几何模型。因此,一些学者提出通过变形轮廓模板来进行目标跟踪。目前有两种较为流行的可变形模型:自由式的可变形模型和参数可变的模型。前者主要通过满足一些简单的约束条件(连续性、平滑性等)来跟踪任意形状的运动

动目标；后者通过使用参数公式或变形公式来描述目标的形状。

图像处理与分析理论中，常用的相似性度量方法有欧氏距离、街区距离、棋盘距离、加权距离、巴特查理亚系数（Bhattacharyya）、Hausdorff 距离等，其中欧氏距离由于其简便性而得到最为广泛的应用。部分相似性度量方法介绍如下：

### ● 欧氏距离（Euclidean Distance）

首先，定义二维平面的两点  $(x_1, y_1)$  和  $(x_2, y_2)$ ，可以被看作是两个二维向量。

其定义如下：

$$d[(x_1, y_1), (x_2, y_2)] = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (2-1)$$

### ● 加权距离（Weighted Distance）

首先，定义二维平面的两点  $(x_1, y_1)$  和  $(x_2, y_2)$ ，可以被看作是两个二维向量。

加权距离可以看作是街区距离、棋盘距离以及类欧氏距离的扩展距离，其定义如下：

$$d[(x_1, y_1), (x_2, y_2)] = \begin{cases} \omega_0 |x_1 - x_2| + \omega_1 |y_1 - y_2| & \text{if } |x_1 - x_2| > |y_1 - y_2| \\ \omega_0 |y_1 - y_2| + \omega_1 |x_1 - x_2| & \text{others} \end{cases} \quad (2-2)$$

其中， $\omega_i > 0, i = 1, 2$ 。当  $\omega_0 = 1, \omega_1 = 1$  时，加权距离就是街区距离（City-Block Distance）；当  $\omega_0 = 1, \omega_1 = 0$  时，加权距离就变成了棋盘距离（Chessboard Distance）；当  $\omega_0 = 1, \omega_1 = \sqrt{2} - 1$  时，加权距离就变成了类欧式距离（Quasi-Euclidean Distance）。

### ● 巴特查理亚系数（Bhattacharyya）

Bhattacharyya 系数，其定义如下：

$$\rho(y) \equiv \rho[\hat{p}(y), \hat{q}(y_0)] = \sum_{b=1}^m \sqrt{\hat{p}_b(y) \hat{q}_b(y_0)} \quad (2-3)$$

其几何意义为两个  $m$  维向量  $(\sqrt{\hat{p}_1}, \dots, \sqrt{\hat{p}_m})$  和  $(\sqrt{\hat{q}_1}, \dots, \sqrt{\hat{q}_m})$  夹角的余弦。

### ● Hausdorff 距离

Hausdorff 距离（Hausdorff Distance, HD）是对两个特征点集的之间的相似度进行度量的方法，即有限像素点集合之间的一种距离定义形式。

已知两组有限像素点集合  $A = \{a_1, a_2, \dots, a_p\}$  和  $B = \{b_1, b_2, \dots, b_q\}$ ，则 HD 距离

定义如下：

$$H(A, B) = \max\{h(A, B), h(B, A)\} \quad (2-4)$$

其中  $h(A, B) = \max_{a_i \in A} \min_{b_j \in B} \|b_j - a_i\|$ ，称为  $A$  到  $B$  的有向 HD，表示点集  $A$  中所有点到点集  $B$  的距离  $d_B(a_i)$  的最大值； $d_B(a_i) = \min_{b_j \in B} \|a_i - b_j\|$ ，表示点到集合的距离，是该点到集合  $B$  中每一点的距离的最小值。

### ● 改进的 Hausdorff 距离

HD 的改进算法都是对有向 HD 的定义做了修正，例如部分 HD 距离和基于平均距离值的 MHD。

部分 HD，先求出点集  $A$  中所有点到点集  $B$  的距离，然后将这些距离由小到大排序，其中序号  $k$  的距离即为  $h_k(A, B)$ ；部分 HD 的有向 HD 定义如下：

$$h_k(A, B) = K_{a_i \in A}^{th} d_B(a_i) \quad (2-5)$$

基于平均距离值的 MHD，有向 MHD 的改进如下：

$$h_{MHD}(A, B) = \frac{1}{N_A} \sum_{a_i \in A} d_B(a_i) \quad (2-6)$$

部分 HD 和平均距离值的 MHD 的融合距离定义如下：

$$h_{LTS}(A, B) = \frac{1}{k \times N_A} d_B(a_i)_{(m)} \quad (2-7)$$

### 2.1.2 基于滤波预测方法的运动目标跟踪算法

滤波方法的优点是：①采用递归滤波的方法，可以将任意一点作为初始状态开始递归；②计算量小，可实时计算；③预测具有无偏、稳定和最优等特点。它的提出为现代控制理论的发展做出了巨大的贡献。滤波方法首先被应用于航天领域，后来在机器人导航、控制、传感器数据融合以及很多军事领域等都得到了广泛的应用。

运动目标跟踪常用的预测算法有 Kalman 滤波、扩展 Kalman 滤波以及粒子滤波等算法，在这里我们将主要介绍扩展 Kalman 滤波算法，在后续的章节中我们将对 Kalman 滤波算法和粒子滤波算法进行简单的介绍。

标准的 Kalman 滤波算法解决了线性过程中的状态估计问题，然而，实际应

用中状态方程、观测方程一般是非线性非高斯的。扩展 Kalman 滤波的提出就是为了解决这类系统的状态估计问题，其详细介绍如下：

非线性的状态方程和观测方程可以表示为：

$$\begin{aligned} x_k &= f(x_{k-1}, u_{k-1}, w_{k-1}) \\ z_k &= h(x_k, v_k) \end{aligned} \quad (2-8)$$

忽略状态噪声和观测噪声可以得到状态和观测值的近似估计，即：

$$\begin{aligned} \hat{x}(k|k-1) &= f(\hat{x}(k-1|k-1), u_{k-1}, 0) \\ \hat{z}(k|k-1) &= h(\hat{x}(k|k-1), 0) \end{aligned} \quad (2-9)$$

为了估计非线性过程，必须对公式(2-9)进行线性化，将函数  $f$  分别对  $x$  和  $w$  求偏导，得到雅克比矩阵  $A$  和  $W$ ；将函数  $h$  分别对  $x$  和  $v$  求偏导，并且得到雅克比矩阵  $H$  和  $V$ 。线性化过程为：

$$\begin{aligned} x_k &= \hat{x}(k|k-1) + A(x_{k-1} - \hat{x}(k-1|k-1)) + Ww_{k-1} \\ z_k &= \hat{z}(k|k-1) + H(x_k - \hat{x}(k|k-1)) + Vv_k \end{aligned} \quad (2-10)$$

$$\begin{aligned} A_{[i,j]} &= \frac{\partial f_{[i]}}{\partial x_{[j]}}(\hat{x}(k-1|k-1), u_{k-1}, 0) \\ W_{[i,j]} &= \frac{\partial f_{[i]}}{\partial w_{[j]}}(\hat{x}(k-1|k-1), u_k, 0) \\ H_{[i,j]} &= \frac{\partial h_{[i]}}{\partial x_{[j]}}(\hat{x}(k|k-1), 0) \\ V_{[i,j]} &= \frac{\partial h_{[i]}}{\partial v_{[j]}}(\hat{x}(k|k-1), 0) \end{aligned} \quad (2-11)$$

我们利用线性化后的状态方程和观测方程，在后验估计误差协方差最小的条件下，进行扩展 Kalman 滤波的两个过程，即预测部分和修正部分，如下所示：

**预测部分：**

状态预测：

$$\hat{x}(k|k-1) = f(\hat{x}(k-1|k-1), u_{k-1}, 0) \quad (2-12)$$

先验估计误差协方差预测：

$$P(k|k-1) = A_k P(k-1|k-1) A_k^T + W_k Q_{k-1} W_k^T \quad (2-13)$$

**修正部分：**

滤波增益：

$$K_k = P(k | k-1)H_k^T (H_k^T P(k | k-1)H_k^T + V_k R_k V_k^T)^{-1} \quad (2-14)$$

修正过的状态向量的后验估计：

$$\hat{x}(k | k) = \hat{x}(k | k-1) + K_k (z_k - h(\hat{x}(k | k-1), 0)) \quad (2-15)$$

修正后的后验估计误差协方差：

$$P(k | k) = (I - K_k H_k) P(k | k-1) \quad (2-16)$$

利用上述的滤波器来估计目标运动，当系统对目标的运动位置和速度有了可靠估计后，可以在相对较小的区域内进行搜索，完成对目标的跟踪过程。当目标被遮挡时，利用滤波器对目标的运动轨迹进行可靠预测，可以方便地在特定区域内搜索目标，等待目标的重新出现。一般意义上，定义滤波器的系统状态为  $x_k$ ，其是一个四维向量  $(s_x, s_y, v_x, v_y)^T$ ，各分量分别表示运动目标在坐标轴上的位置和坐标轴方向上的速度。然后在基于滤波算法的跟踪过程中，初始化目标的状态参数（如：速度、位置等），即使在不准确的情况下，经过有限次的迭代运算，上述滤波算法仍然可以准确地估计出目标的真实运动状态参数，具有预测的稳定性。

### 2.1.3 基于优化匹配搜索的运动目标跟踪算法

Mean Shift 最早是由 Fukunaga 等人于 1975 年提出来的，其最初含义是偏移的均值向量。但随着 Mean Shift 理论的发展，Mean Shift 的含义也逐步地发生着变化，目前所说的 Mean Shift 算法一般是指一个迭代的过程，即先算出当前点的偏移均值，移动该点到其偏移均值，然后以此点为新的起始继续移动，直到满足一定的条件结束。

#### Mean Shift 理论基础和原理

##### ● 基本 Mean Shift

给定  $d$  维空间  $R^d$  中的  $n$  个样本点  $x_i, i=1, \dots, n$ ，在  $x$  点的 Mean Shift 向量的基本形式定义为：

$$M_h(x) \equiv \frac{1}{k} \sum_{x_i \in S_h} (x_i - x) \quad (2-17)$$



其中， $S_h$  是一个半径为  $h$  的高维球区域，满足以下关系的  $y$  点的集合，

$$S_h(x) \equiv \left\{ y : (y-x)^T (y-x) \leq h^2 \right\} \quad (2-18)$$

$k$  表示在这  $n$  个样本点  $x_i$  中，有  $k$  个点落入  $S_h$  区域中。

我们可以看到  $(x_i - x)$  是样本点  $x_i$  相对于点  $x$  的偏移向量，公式 (2-17) 定义的 Mean Shift 向量  $M_h(x)$  就是对落入区域  $S_h$  中的  $k$  个样本点相对于点  $x$  的偏移向量求和然后再平均。直观上讲，如果样本点  $x_i$  从一个概率密度函数中采样得到，由于非零的概率密度梯度指向概率密度增加最大的方向，因此从平均上来说， $S_h$  区域内的样本点更多的落在沿着概率密度梯度的方向。因此，对应的 Mean Shift 向量  $M_h(x)$  应该指向概率密度梯度的方向。

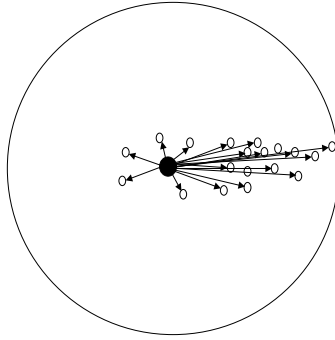


图 2-1 Mean Shift 示意图

如上图所示，大圆圈所圈定的范围就是  $S_h$ ，小圆圈代表落入  $S_h$  区域内的样本点  $x_i \in S_h$ ，黑点就是 Mean Shift 的基准点  $x$ ，箭头表示样本点相对于基准点  $x$  的偏移向量，因此我们可以得到，平均的偏移向量  $M_h(x)$  会指向样本分布最多的区域，也就是概率密度函数的梯度方向。

### ● 扩展 Mean Shift

从 (2-16) 式我们可以看出，只要落入  $S_h$  区域内的采样点，无论其离  $x$  远近，对最终的  $M_h(x)$  计算的贡献是一样的。然而，一般来说离  $x$  越近的采样点对估计  $x$  周围的统计特性越有效，因此我们引进核函数的概念，即在计算  $M_h(x)$  时可以

考虑距离的影响,认为在所有的样本点  $x_i$  中,重要性并不一样,因此我们对每个样本都引入一个权重系数。

因此,基本的 Mean Shift 形式扩展如下:

$$M(x) \equiv \frac{\sum_{i=1}^n G_H(x_i - x)w(x_i)(x_i - x)}{\sum_{i=1}^n G_H(x_i - x)w(x_i)} \quad (2-19)$$

其中:

$$G_H(x_i - x) = |H|^{-1/2} G(H^{-1/2}(x_i - x))$$

$G(x)$  是一个单位核函数

$H$  是一个正定的对称  $d \times d$  矩阵,我们一般称之为带宽矩阵

$w(x_i) \geq 0$  是一个赋给采样点  $x_i$  的权重

在实际应用的过程中,带宽矩阵  $H$  一般被限定为一个对角矩阵  $H = \text{diag}[h_1^2, \dots, h_d^2]$ , 甚至可以更简单的被定义为正比于单位矩阵, 即  $H = h^2 I$ 。

由于后一形式只需要确定一个系数  $h$ , 因此常常被采用, 因此 (2-19) 式又可以被改写为:

$$M_h(x) \equiv \frac{\sum_{i=1}^n G(\frac{x_i - x}{h})w(x_i)(x_i - x)}{\sum_{i=1}^n G(\frac{x_i - x}{h})w(x_i)} \quad (2-20)$$

可以看到, 如果对所有的采样点  $x_i$  满足:

$$(1) \quad w(x_i) = 1$$

$$(2) \quad G(x) = \begin{cases} 1 & \text{if } \|x\| < 1 \\ 0 & \text{if } \|x\| \geq 1 \end{cases}$$

则 (2-20) 式完全退化为 (2-17) 式, 也就是说扩展的 Mean Shift 在某些情况下会退化为最基本的 Mean Shift。

### Mean Shift 在目标跟踪中的应用

我们采用一个目标的灰度或颜色分布来描述这个目标, 假设目标中心位于  $x_0$ , 则该目标可以表示为:

$$\hat{q}_u = C \sum_{i=1}^n k \left( \left\| \frac{x_i^s - x_0}{h} \right\|^2 \right) \delta [b(x_i^s) - u] \quad (2-21)$$

加权的灰度或颜色直方图，使距离目标中心的像素点具有较大的权值，具有抗部分遮挡和边缘噪声影响的优点。

候选的位于  $y$  的目标可以描述为：

$$\hat{p}_u(y) = C_h \sum_{i=1}^{n_h} k \left( \left\| \frac{x_i^s - y}{h} \right\|^2 \right) \delta [b(x_i^s) - u] \quad (2-22)$$

因此目标跟踪可以简化为寻找最优的  $y$ ，使得  $\hat{p}_u(y)$  与  $\hat{q}_u$  最相似。

$\hat{p}_u(y)$  与  $\hat{q}_u$  的相似性用 Bhattacharyya 系数  $\hat{\rho}(y)$  来度量，即

$$\hat{\rho}(y) \equiv \rho[p(y), q] = \sum_{u=1}^m \sqrt{p_u(y) \hat{q}_u} \quad (2-23)$$

式(2-23)在  $\hat{p}_u(\hat{y}_0)$  点泰勒展开可得，

$$\rho[p(y), q] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p(y_0) q_u} + \frac{1}{2} \sum_{u=1}^m p_u(y) \sqrt{\frac{q_u}{p_u(y_0)}} \quad (2-24)$$

把式(2-22)带入式，整理可得：

$$\rho[p(y), q] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p(y_0) q_u} + \frac{C_h}{2} \sum_{i=1}^n w_i k \left( \left\| \frac{y - x_i}{h} \right\|^2 \right) \quad (2-25)$$

其中，

$$w_i = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{q_u}{p_u(y_0)}} \quad (2-26)$$

对式(2-25)右边的第二项，可以利用 Mean Shift 算法对其进行最优化。

Mean Shift 算法的形式简洁，收敛条件宽松。只要确定了跟踪目标，整个跟踪过程就不需额外的参数输入。由于无需对概率密度分布进行估计，就可以直接沿着梯度方向搜索局部最大值，大大减少了运算量，所以这种算法具有良好的实时性。

Mean Shift 算法用于视频跟踪有以下几个优势：首先，算法计算量不大，在目标区域已知的情况下完全可以做到实时跟踪；其次，作为一个无参数概率密度

估计算法，很容易和其他算法结合使用；然后，采用加权直方图模型，对部分遮挡、目标旋转、形变和背景运动不敏感。但是，其也有不足之处，比如：缺乏必要的模板更新算法，整个跟踪过程中窗宽的大小保持不变，因此当目标有尺度变化时，可能跟踪失败；颜色直方图是一种比较通用的目标特征描述子，当背景和目标的颜色直方图相似时，跟踪效果往往不好；另外，**Mean Shift** 算法是局部最优的优化算法，当出现多个局部峰值时，算法可能会收敛于局部最优值而造成跟踪算法失效。

## 2.2 跟踪特征评估算法

目前对于目标跟踪的研究越来越多的集中到对跟踪目标的表达及复杂背景的处理上。其中对于目标表达的特征越来越复杂，**Haar** 小波、轮廓特征、纹理特征、颜色特征及其组合受到了研究人员的重视。另外，在近年的目标跟踪研究中，为了使得目标特征对背景的改变具有适应性，人们提出了通过在跟踪过程中实时更新目标特征权值的方法来评估特征有效性。然而，如何进行实时的跟踪特征评估还有待深入的研究。

近些年来，人们更多地将视觉跟踪作为模式识别中一个分类问题进行研究，即如何将跟踪目标（前景）与跟踪背景（背景）进行更好的区分，这样使得属于前景或者背景的像素更容易进行聚类，从而提高跟踪的精度和鲁棒性。其中，特征权值评估的研究问题便是将视觉跟踪作为分类问题的一种典型方法。该算法主要是基于前景与背景特征相似度的变化而进行特征权值的调整。卡内基-梅隆大学的 **Collins**[11]等人根据颜色特征对目标的视角、目标之间的部分遮挡以及目标形状的不敏感性提出了一种基于颜色的目标特征描述的新方法。他们所提出的跟踪特征集合是颜色特征 **R**、**G** 和 **B** 的线性组合，定义如下：

$$F_1 = \{w_1 R + w_2 G + w_3 B \mid w_* \in [-2, -1, 0, 1, 2]\} \quad (2-27)$$

其中，每个颜色分量参数的取值范围为 $[-2, 2]$ 之间的整数，因此原则上便会得到 125 种基于三种颜色分量的组合特征。消除特征之间的冗余，例如三个颜色分量的参数相等等情况之后，最终得到 49 种组合特征，从而构成了他们所定义的颜色特征集合。然后将每种颜色特征都归一化到 0-255 之间，通过遍历视频帧，将得到的统计结果离散化为直方图，这样就得到特征集合中每个特征的直方图描

述形式。这种特征集合有如下三个优点：1) 容易计算，由于特征都是基于视频帧中每个像素颜色分量的统计，所以比较直观、快速；2) 这种特征集合类似于三维颜色空间在一维空间中的投影，基本上覆盖了原有的颜色特征空间；3) 相比于原有的颜色特征定义，这种特征集合还包含了比原有颜色特征更多的描述形式，例如： $R-B$ 、 $2G-R-B$  等。

当提取得到前景和背景的跟踪特征之后，Collins 等人首先计算了每种特征在前景与背景之间的相似性，定义如下：

$$L(i) = \log \frac{\max\{p(i), \delta\}}{\max\{q(i), \delta\}} \quad (2-28)$$

其中， $p(i)$  定义为目标中的第  $i$  个特征，而  $q(i)$  为背景中的第  $i$  个特征。 $L(i)$  为目标与背景基于第  $i$  个特征的 log 似然比，包含了前景与背景在某一个特征之间的相似度信息。然后通过统计第  $i$  个特征在视频帧中得到的每一个  $L(i)$ ，并计算其方差便得到了目标与背景在第  $i$  个特征之间的判别力，公式表达如下：

$$VR(L; p, q) = \frac{\text{var}(L; (p+q)/2)}{[\text{var}(L; p) + \text{var}(L; q)]} \quad (2-29)$$

其中  $\text{var}$  定义为某一特征相似度的方差。

根据目标和背景特征之间的特征判别力，直接选取特征判别力最大的前  $N$  个特征作为下一帧视频跟踪的特征集。

Liang 等人[9]提出了一种类似的方法，但是特征的判别力是通过贝叶斯误差率定义的，即通过目标和前景颜色特征直方图的交集进行计算得到的。Wang 等人[8]提出了一种通过 Adaboost 学习选取 Harr 特征子集作为跟踪特征集的方法。尽管这些方法具有一定程度的前景/背景变化的自适应性，但是它们很少考虑到跟踪过程的特征变化的时间连续性，即跟踪的特征集合变化是个连续的过程。

由于 Collins 等都采用单一的特征来描述（颜色或者纹理）跟踪目标，并且他们在特征评估方面都有一定的局限性，因此，在他们研究的基础上，本文将主要研究综合直方图的提取和评估。综合直方图将目标的颜色和局部轮廓特征有效地融合起来，这样描述目标更加准确，所包含关于目标的信息量更多，便于计算机处理和进行有效的跟踪。同时，本文提出的目标特征的评估方法通过利用滤波

方法有效地解决了跟踪过程中特征权值的帧间预测和帧内校验的问题，使得到的特征权值是真实权值的最优估计，并达到均方误差最小，从而使得目标的综合特征能够切实有效地达到背景自适应，很好地解决复杂动态背景中的单目标跟踪，使得跟踪算法更加稳定、鲁棒。

### 第三章 基于 Kalman 滤波的特征评估算法

Kalman 滤波方法是 Kalman 和 Bucy 于 1960 年提出的一种递推最优线性均方误差估计的方法，他们把状态空间的概念引入到随机估计理论中，把信号过程视为白噪声作用下的一个线性系统的输出，用状态方程来描述这种输入输出关系，估计过程中利用系统状态方程、观测方程和白噪声激励(系统噪声和观测噪声)的统计特性形成滤波算法，由于所用的信息都是时域内的量，所以不但可以对平稳的一维的随机过程进行估计，也可以对非平稳的多维随即过程进行估计。

#### 3.1 Kalman 滤波算法及其应用

Kalman 滤波方法的优点是：采用线性递归滤波的方法，可以任意一点作为初始化开始递归；计算量小，可实时计算；预测具有无偏、稳定和最优的特点。它的提出为现代控制理论的发展做出了巨大的贡献，首先被应用于航天领域，后来在机器人导航、控制、传感器数据融合以及很多军事领域都得到了广泛的应用。目前，Kalman 滤波方法在视频跟踪算法中也得到了非常广泛的应用，主要用来预测目标下一时刻目标的位置，以此作为量测算法匹配搜索的起点，从而提高运算速度。

##### 3.1.1 Kalman 滤波算法基本原理

系统的状态方程：

$$X_k = A \times X_{k-1} + B \times u_{k-1} + w_{k-1} \quad (3-1)$$

观测方程：

$$z_k = H \times X_k + v_k \quad (3-2)$$

$w$  和  $v$  分别为独立不相关的状态噪声和观测噪声，并且分布满足，

$$\begin{aligned} p(w) &\sim N(0, Q) \\ p(v) &\sim N(0, R) \end{aligned} \quad (3-3)$$

状态预测方程（先验估计）：

$$\hat{X}(k|k-1) = A \times \hat{X}(k-1|k-1) + B \times u_{k-1} \quad (3-4)$$

经观测值修正过的后验估计：

$$\widehat{X}(k|k) = \widehat{X}(k|k-1) + K_k(z_k - H\widehat{X}(k|k-1)) \quad (3-5)$$

其中,  $\widehat{X}(k|k-1)$  是  $k$  时刻的先验估计, 即由  $k-1$  时刻的状态向量预测得到  $k$  时刻的状态向量;

$\widehat{X}(k|k)$  是  $k$  时刻的后验估计, 即得到  $k$  时刻的观测值  $z_k$  后, 经过观测值修正后的状态向量;

$z_k - H\widehat{X}(k|k-1)$  称为差值, 是观测估计和实际观测之间的误差;

$K_k$  是  $k$  时刻的滤波增益, 用来最小化  $k$  时刻的后验估计误差协方差。

那么,  $k$  时刻的先验估计误差和后验估计误差分别为:

$$\begin{aligned} \widetilde{X}(k|k-1) &= X(k) - \widehat{X}(k|k-1) \\ \widetilde{X}(k|k) &= X(k) - \widehat{X}(k|k) \end{aligned} \quad (3-6)$$

先验估计误差和后验估计误差的协方差分别为:

$$P(k|k-1) = E[\widetilde{X}(k|k-1) \widetilde{X}^T(k|k-1)] \quad (3-7)$$

$$\begin{aligned} P(k|k) &= E[\widetilde{X}(k|k) \widetilde{X}^T(k|k)] \\ &= E[(X(k) - \widehat{X}(k|k)) (X(k) - \widehat{X}(k|k))^T] \\ &= E[(X(k) - \widehat{X}(k|k-1) - K_k(z_k - H\widehat{X}(k|k-1))) \\ &\quad (X(k) - \widehat{X}(k|k-1) - K_k(z_k - H\widehat{X}(k|k-1)))^T] \\ &= P(k|k-1) - P(k|k-1)H^TK^T - KHP(k|k-1)^T \\ &\quad + K(HP(k|k-1)H^T + R)K^T \end{aligned} \quad (3-8)$$

得到后验估计误差协方差  $P(k|k)$  关于滤波增益  $K_k$  的表达式, 为了最小化后验估计误差协方差, 经  $P(k|k)$  对  $K_k$  求偏导, 令其结果等于零, 解出滤波增益  $K_k$  的表达式:

$$K_k = P(k|k-1)H^T(HP(k|k-1)H^T + R)^{-1} \quad (3-9)$$

考虑两个极端的情况, 滤波增益是由先验估计误差协方差  $P(k|k-1)$  和观测噪声协方差  $R$  共同决定的。



$$\begin{aligned}
 \lim_{R \rightarrow 0} \hat{X}(k|k) &= \hat{X}(k|k-1) + H^{-1}(z_k - H\hat{X}(k|k-1)) = z_k \\
 \lim_{P(k|k-1) \rightarrow 0} \hat{X}(k|k) &= \hat{X}(k|k-1) + 0 \times (z_k - H\hat{X}(k|k-1)) \\
 &= \hat{X}(k|k-1)
 \end{aligned} \tag{3-10}$$

Kalman 滤波使每一时刻  $k$  的目标状态向量的后验估计值的误差协方差最小，由预测和修正两步骤组成；预测部分包括状态预测和先验估计误差协方差预测；修正部分包括滤波增益的计算和利用滤波增益对状态预测值和后验估计误差协方差进行修正，计算得到状态向量的后验估计和最小化的后验估计误差协方差，具体如下：

预测部分：

$$\begin{aligned}
 \hat{X}(k|k-1) &= A\hat{X}(k-1|k-1) + Bu_{k-1} \\
 P(k|k-1) &= AP(k-1|k-1)A^T + Q
 \end{aligned} \tag{3-11}$$

修正部分：

$$\begin{aligned}
 K_k &= P(k|k-1)H^T(HP(k|k-1)H^T + R)^{-1} \\
 \hat{X}(k|k) &= \hat{X}(k|k-1) + K_k(z_k - H\hat{X}(k|k-1)) \\
 P(k|k) &= (I - K_kH)P(k|k-1)
 \end{aligned} \tag{3-12}$$

滤波增益  $K_k$  是预测与量测的权衡。当量测精确时应该增大 Kalman 增益，反之则减少 Kalman 增益。针对传统 Kalman 滤波过多依赖于历史量测，一段时间后会导 Kalman 增益减小而失去滤波的作用，因此改进的算法增添了消渐因子  $\gamma$  ( $\gamma > 1$ )，降低历史量测的影响，增加新近量测数据在滤波中的作用。改进 Kalman 滤波的协方差预测方程为：

$$P(k|k-1) = \gamma A(k-1)P(k-1|k-1)A^T(k-1) + Q(k-1) \tag{3-13}$$

### 3.1.2 Kalman 滤波算法在目标跟踪中的应用

Kalman 滤波器对于提高跟踪系统的处理速度和性能有着极其重要的作用。假设目标的运动状态参数为某一时刻目标的位置和速度。在跟踪过程中，由于相邻两帧图像的时间间隔比较短，目标在这么短的时间间隔内状态变化比较小，因此可以假设目标在此时间间隔内是匀速运动，所以匀速运动模型在一般情况下可以很好地反映目标的运动趋势。

定义 Kalman 滤波器的系统状态为  $x_k$ ，其是一个四维向量  $(s_x, s_y, v_x, v_y)^T$ ，分别代表运动目标在坐标轴上的位置和坐标轴方向的速度。根据运动目标在单位时间间隔内是匀速运动的，定义状态转移矩阵  $A$  和观测矩阵  $H$ ：

$$\begin{aligned} A &= \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ H &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \end{aligned} \quad (3-14)$$

### 3.2 基于 Kalman 滤波的特征评估算法研究

#### 3.2.1 综合特征的提取

传统的颜色直方图是一个对图像区域有效的特征表达。近些年出现了一些经典的基于颜色直方图的跟踪系统，它们均利用了直方图计算简单、能够满足实时处理的需求，对运动目标的旋转和尺度伸缩不敏感等特性。但是，由于颜色直方图信息隐藏了图像的空间位置信息并且易受光照等外界环境因素的影响，因此，导致了传统的直方图跟踪算法跟踪精度不高和效果不稳。针对上述问题，有学者提出了梯度方向直方图的概念并且用于人体检测[34]，还有学者提出将改进的直方图方法用于运动目标跟踪[11]，也有学者提出了直方图相似度的改进方法。

本章受到 SIFT 特征提取、匹配以及 HOG 特征在人体检测中的启发，提出了综合直方图的概念，即融合了颜色和梯度方向信息的直方图。多尺度的综合直方图的目标描述形式可以有效地解决动态背景等困难场景下运动目标的鲁棒跟踪。在目标跟踪中，需要着重研究两个方面的问题：1) 目标的特征表达；2) 跟踪过程中的特征对于复杂背景的自适应问题。

#### ● 前景和背景区域的定义

为了使特征具有背景自适应性，我们需要评测综合特征集合中的每个特征对于当前背景的判别能力，对于当前帧中的目标，按照下图定义目标区域和背景区域：

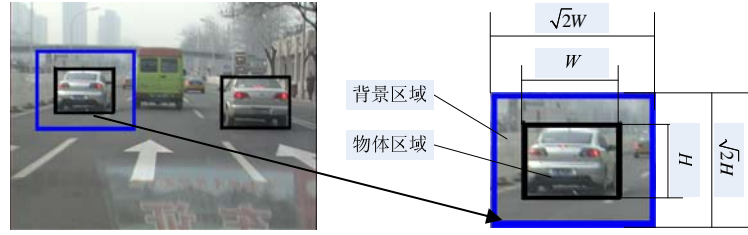


图 3-1 跟踪目标及其背景区域

如上图所示，目标区域定义为  $h \times w$  的矩形框，背景区域定义为  $\sqrt{2}h \times \sqrt{2}w$  的矩形框与目标区域之间的环形区域。目标的特征  $\{F_i(x)\}, i = 0, 1, \dots, N$  从以  $x$  位置为中心的目标区域的像素中提取，背景的特征  $\{B_i(x)\}, i = 0, 1, \dots, N$  从以  $x$  位置为中心的背景区域的像素中提取。

### ● 跟踪综合特征的提取

目前对于目标跟踪的研究越来越集中到对于跟踪目标的表达及复杂背景的处理上。在过去的几年，颜色直方图（HC）被大量的应用于目标跟踪，并且取得了不错的效果。但是大量的实验表明，当目标和背景颜色比较相似的时候，基于颜色的跟踪算法往往不能取得很好的效果。近几年，梯度方向直方图（HOG）开始被广泛地用来进行目标的检测和跟踪，并且取得了很好的效果。HOG 提取的是目标的边缘和轮廓特征，而这些特征对目标的颜色变化不敏感。Dalal 等人 [34] 已经证明了固定尺度的 HOG 和 SIFT 对目标的描述能力几乎一样强。Nguyen 等人 [12] 利用 Gabor Wavelet 描述目标。Wang 等人 [13] 利用 Harr-like 特征子集描述跟踪目标。这些特征同样是为了取得目标的纹理或者轮廓特征。但是，HOG 同样有一些缺点，比如在目标跟踪方面 HOG 的性能要差于 HC，并且当目标出现大面积的平滑区域时，HOG 的性能会大大地降低。虽然每种单独的特征都有一些缺陷，但 HC 和 HOG 两种特征之间的互补性使得我们可以有效地把它们融合在一起。

### 颜色直方图

在 RGB 颜色空间中，每个颜色分量（R、G 和 B）定义一个 16 维的颜色直方图，从而为运动目标构建了一个 48 维的颜色直方图。以 R 颜色分量为例，采用窗宽为  $h_1$  的剖面函数  $k(\bullet): [0, \infty) \rightarrow R$  进行平滑后的颜色直方图为：

$$\hat{q}_b(y_0) = C_1 \sum_{i=1}^n k\left(\left\|\frac{y_0 - X_i}{h_1}\right\|^2\right) \delta[I(X_i) - b_j], \quad \text{if } 1 \leq j \leq m \quad (3-15)$$

其中,  $I(X_i)$  是视频帧中  $X_i$  位置处像素的颜色值 (R、G 或者 B)。其中,  $y_0$  是目标区域的中心位置;  $n$  是目标区域像素的数目;  $\hat{q}_b(y_0)$  是平滑后直方图分量  $b$  的值;  $\delta(\cdot)$  是 Kronecker Delta 函数;  $m$  是颜色直方图级的个数 ( $m=16$ );  $X_i$  是第  $i$  个像素点的位置且  $X_i = [h_i, w_i]$ , 常数  $C_1$  由约束条件  $\sum_{b_j=1}^M \hat{q}_{b_j}(y_0) = 1$  得到, 即

$$C_1 = \frac{1}{\sum_{i=1}^n k\left(\left\|\frac{y_0 - X_i}{h_1}\right\|^2\right) (1 + d(X_i))} \quad (3-16)$$

### 梯度方向直方图

根据 HOG 梯度方向提取的思想来定义和计算灰度图像窗口的梯度方向直方图(HOG)。当给定特定的图像大小和特征尺度, 梯度直方图可以很好地表达出跟踪目标的局部轮廓。在研究中, 我们使用了 72 维的 HOG 特征对目标的局部轮廓进行描述。

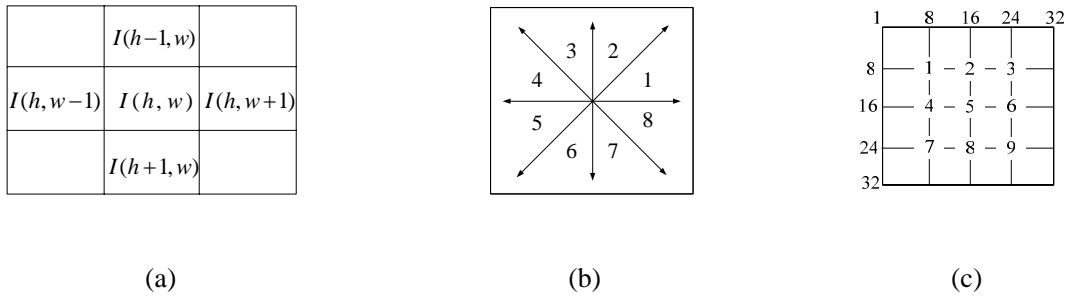


图 3-2 HOG 特征提取, (a) 计算梯度方向的模板, (b) 8 个方向投影区间, (c) 图像窗口中的 9 个块。

下面是 HOG 特征提取的详细过程: 将跟踪目标转化为灰度图像; 将跟踪目标窗口归一化到固定大小的图像窗口: 32x32 像素; 然后把图像划分为 16 个 8x8 的小区域, 并且每个 2x2 的小区域组成一个图像块, 因此每个相邻的图像块有 2 个小区域彼此重合; 本文中将每个图像块的梯度方向统计投影到一个 8 维的梯度方向直方图中。从图 3-2 (c) 中可以看到, 由于每个图像窗口包含 9 个图像块, 因此每个目标得到 72 的梯度方向直方图来描述目标的局部轮廓。每个像素的梯

度方向计算如下：

$$\begin{aligned}
 I &= G(\sigma, 0) * I_0 \\
 dy &= I(h+1, w) - I(h-1, w) \\
 dx &= I(h, w+1) - I(h, w-1) \\
 ori(h, w) &= a \tan 2(dy, dx) \quad ori \in [0, 2\pi]
 \end{aligned}
 \tag{3-17}$$

上式是计算一个像素在特定尺度下一个像素点方向的过程。 $G(\sigma, 0)$  表示一个高斯函数， $I_0$  表示原图像， $I$  表示经过平滑滤波后的图像。 $ori(h, w)$  表示像素梯度方向（目标轮廓方向）。对目标区域/背景区域的所有像素进行累计，可以获得 72（9 个图像块，每个块 8 个方向，如图 3-2（b）和（c）所示）维 HOG 特征。此特征因为只保留方向特征，不保留强度特征，因此对于亮度、光照等变化具有一定的适应性。

由于 HOG 对目标的旋转敏感，因此，为了使得 HOG 特征能够很好地处理目标旋转运动，本文采用了 SIFT 算法中的主方向思想，即在计算的过程中，在图像窗口中心的邻域范围内进行采样，然后用直方图统计该邻域像素的梯度方向，直方图的峰值则代表了该图像窗口的主方向。在跟踪的过程中，将候选目标图像窗口的主方向都归一化到初始跟踪目标的主方向上，这样使得目标的 HOG 特征和目标的旋转运动无关。如下图所示：

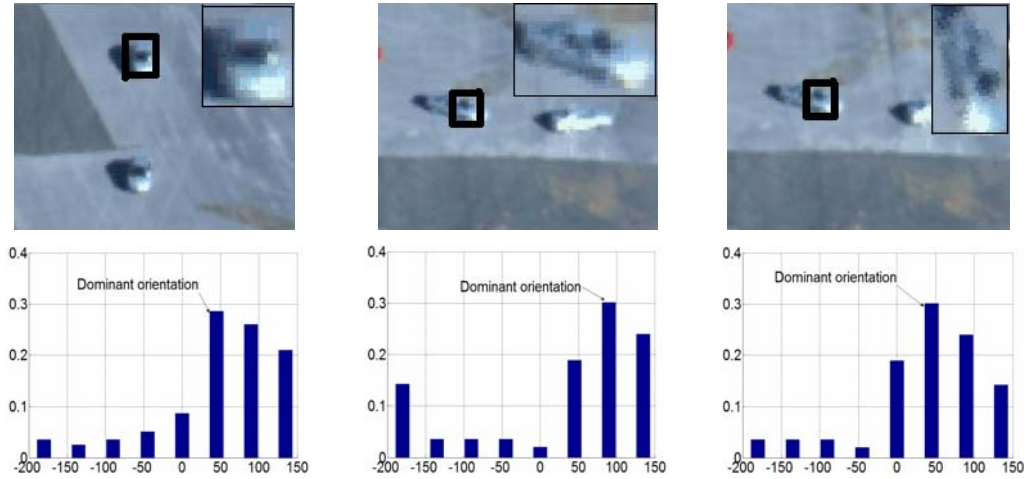


图 3-3 利用 HOG 主方向方法对目标进行旋转的示意图

### 3.2.2 基于 Kalman 滤波的特征评估

提取出来的综合特征（Histogram of Oriented Gradient and Color, HOGC）在

跟踪的过程中将会根据背景和目标本身的变化进行实时的自适应评估。不失一般性，在此我们做以下两个假设：（1）特征权值和特征判别力都是在（0.0,1.0）之间的浮点数；（2）权值调整之后，背景/前景判别力大的特征的权值也应该大，反之亦然。

在跟踪过程中，跟踪目标和背景的表现时刻发生着变化。在近年的目标跟踪研究中，为了使得特征对于背景的改变具有适应性，一些学者提出在跟踪过程中实时更新目标特征的权值来评估特征有效性。在本章中，我们提出了一种全新的基于 Kalman 滤波的特征权值调整策略。假定目标特征的变化是满足线性、高斯的假设，Kalman 滤波中的状态定义为特征的权值，Kalman 滤波中的观察定义为特征的判别力。则我们所提出的 Kalman 滤波的状态方程和观察方程可表示如下：

$$\begin{cases} \begin{pmatrix} W_{t+1} \\ \Delta W_{t+1} \end{pmatrix} = \begin{pmatrix} I_{N \times N} & I_{N \times N} \\ 0 & I_{N \times N} \end{pmatrix} \begin{pmatrix} W_t \\ \Delta W_t \end{pmatrix} + u_t \\ (S_t) = (I_{N \times N} \ 0) \begin{pmatrix} W_t \\ \Delta W_t \end{pmatrix} + v_t \end{cases} \quad (3-18)$$

其中， $W_t = \{w_t(1), w_t(2), \dots, w_t(N)\}$ ，表示第  $t$  视频帧中特征集所对应的特征权值集合， $w_t(i)$  是第  $t$  视频帧中第  $i$  个特征所对应的特征权值，并且  $\Delta W_t = W_t - W_{t-1}$ ； $S_t = \{S_t^1, S_t^2, \dots, S_t^N\}$  定义为第  $t$  视频帧中特征集所对应的特征判别力集合， $S_t^i$  是第  $t$  视频帧中第  $i$  个特征所对应的特征判别力。 $u_t$  和  $v_t$  为高斯白噪声。

表 3-1 基于 Kalman 滤波的特征权值评估算法

1. 初始化 ( $t=0$ ). 在这个阶段，首先初始化目标的特征权值为：

$$(w_0(i), \Delta w_0(i))^T = \left( \frac{1}{N}, 0 \right)^T;$$

2. 预测 ( $t>0$ ). 对于特征集合中的每一个特征，进行如下操作：

- (a) 利用 Kalman 滤波预测该特征的先验权值；
- (b) 利用先验的加权 HOGC 特征来指导目标跟踪，找到目标所在的位置。

3. 修正 ( $t>0$ ). 当搜索到目标在下一帧的最佳位置以后，进行如下操作：

- (a) 提取目标在下一帧中的 HOGC 特征；
- (b) 计算 HOGC 特征的判别力；

(c) 通过 Kalman 滤波来修正 HOGC 特征的权值, 得到当前帧中后验特征权值。

4.  $t = t + 1$ ; 跳转到步骤 2 或者结束循环。

上面描述的利用 Kalman 滤波调整特征权值的过程可以用下图表示:

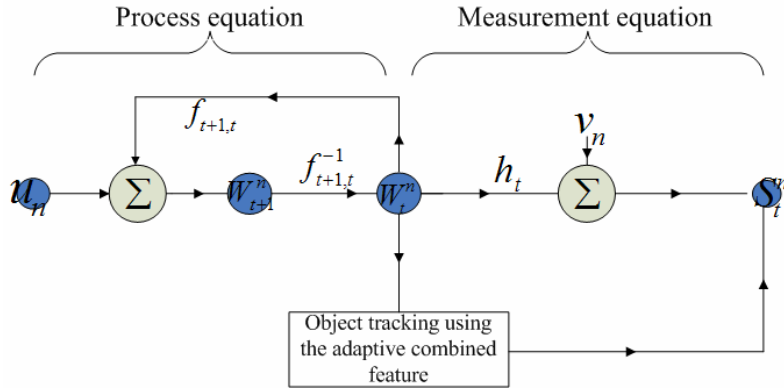


图 3-4 基于 Kalman 滤波的特征权值调整示意图

### 3.3 实验结果比较与算法性能分析

本节将本章所提出的跟踪算法的实验结果与传统算法的结果进行比较。同时, 还把本章的跟踪算法嵌入到了 PTZ (pan-tilt-zoom) 摄像机中, 从应用的角度分析所提出的跟踪算法在实际应用中的效果。

#### 3.3.1 视频测试集合

本章所采用的视频集合主要来自目前比较流行的公共测试集合和我们实验室自己的跟踪测试集合, 包括: VIVID data set [47]、CAVIAR data set [48] 和 SDL data set [54]。这些测试视频都是针对运动背景和前景的, 其中背景和目标的相对变化主要包括了跟踪目标之间的相互遮挡、目标的尺度变化、目标的旋转、背景光照变化以及相似背景等。在测试视频中不仅有对人体的跟踪, 也有对运动车辆的跟踪。

#### 3.3.2 实验结果分析及比较

把本章所提到的跟踪算法与目前比较流行的其他三种跟踪算法进行比较, 其中包括: 基于特征选择的跟踪算法(Collins method)[11]、基于 SIFT 特征点的跟



跟踪算法（SIFT-Based method）[30]和传统的 Kalmam 滤波跟踪算法（traditional Kalman filter method）[39]。下面的实验测试视频来自于 VIVID data set，跟踪目标是在动态背景中的一辆小吉普车，视频是通过直升飞机从空中拍摄得到的。

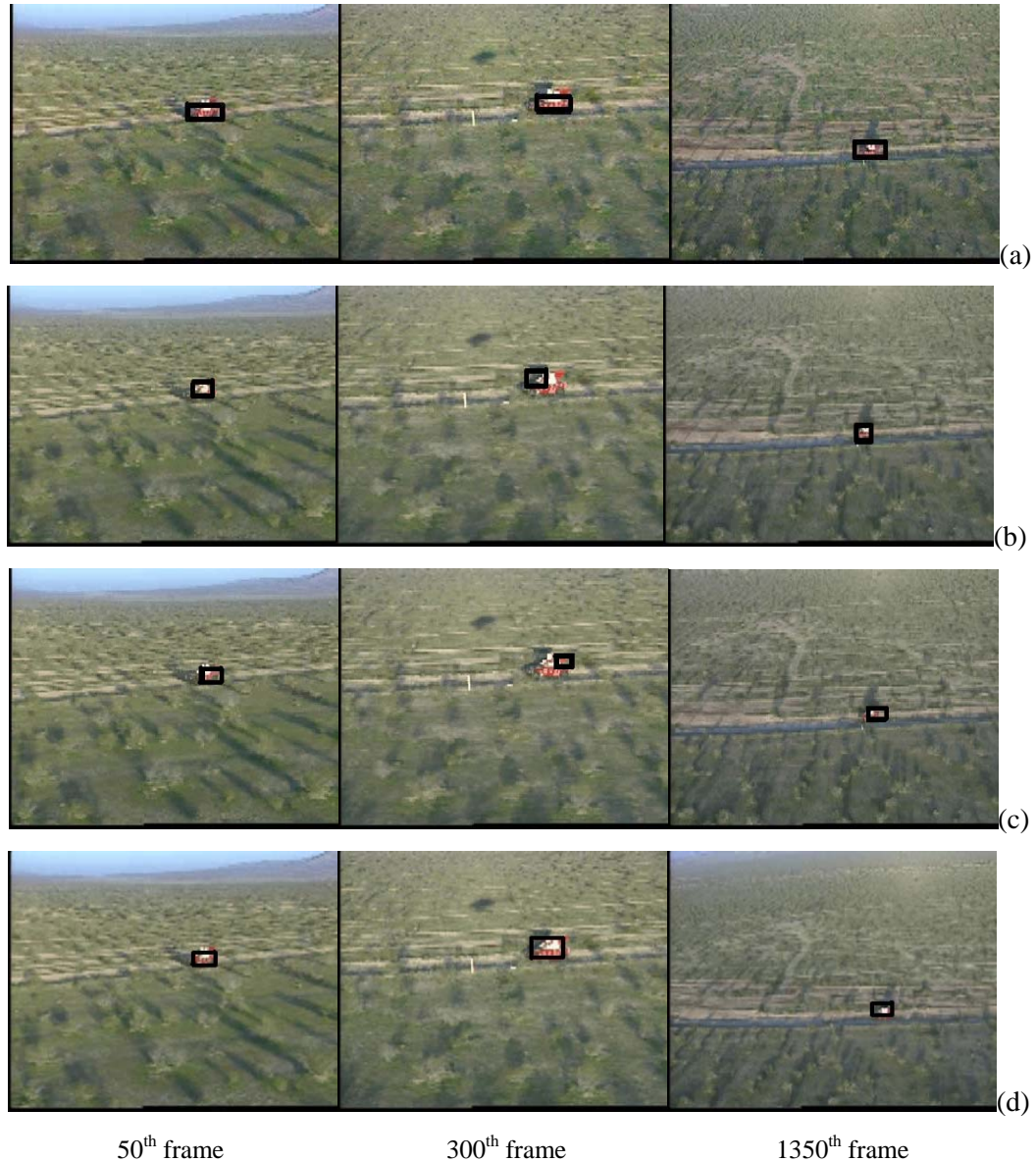


图 3-5 相对简单背景的跟踪结果图. (a) Collins method 的跟踪结果, (b) SIFT-based method 的跟踪结果, (c) traditional Kalman filter method 的跟踪结果, (d)我们所提出算法的跟踪结果.

在上图所示的跟踪实验结果中，由于跟踪目标的背景的变化、目标本身表观的变化以及相互之间的运动模式都相对比较简单，因此四种跟踪算法都取得了比较好的实验结果。与其他三种算法相比，SIFT-based method 算法取得的结果相对较差，这是因为虽然 SIFT 特征点对目标有很强的描述能力，但是由于背景中的灌木丛在目标特征点的计算过程中同样会产生大量的 SIFT 特征点，从而导致有



很多背景的噪声点掺杂在跟踪目标 SIFT 特征点的描述中, 使得对跟踪目标描述的准确度降低, 从而导致跟踪实验结果不是很稳定。Collins method 和 traditional Kalman Filter method 在这段测试视频的跟踪结果都比较好, 其中 Collins method 算法通过特征选择可以使对跟踪目标的描述自适应背景的变化, traditional Kalman Filter method 通过预测和修正目标的运动模型可以很好地估计目标在视频帧中的位置。尽管四种跟踪算法都取得了比较好的实验结果, 从实验结果图中还是可以明显看出本文提出的跟踪算法所取得的实验结果最为鲁棒和稳定。

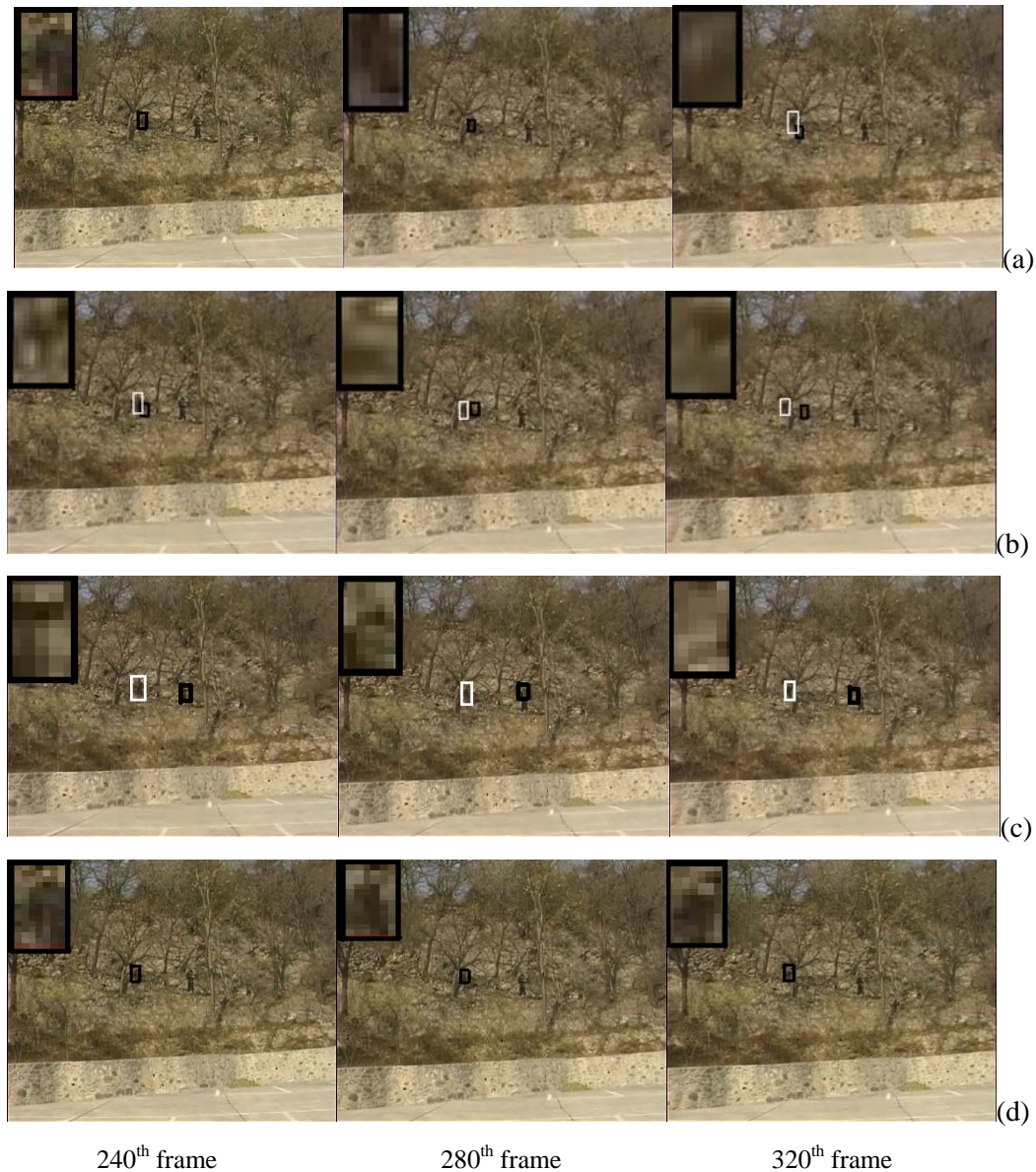


图 3-6 复杂背景的跟踪结果图. 跟踪结果用黑框表示, 真实的结果用白框表示. (a) Collins method 的跟踪结果, (b) SIFT-based tracking method 的跟踪结果, (c) traditional Kalman filter method 的跟踪结果, (d) 我们所提出方法的跟踪结果.

第二组实验比较结果如上图所示。此视频是我们所采用的视频集中难度最大的一段，跟踪目标是一个尺度很小的人体，背景对跟踪目标的干扰很大。不仅背景的颜色和目标的比较类似，而且背景中树的结构与跟踪目标的结构也极其相似，这些都给此运动目标的跟踪带来极大的困难。我们仍对上述提到的四种跟踪算法进行比较。为了更清楚地描述每种跟踪算法的结果，将跟踪结果在每个图片的左上角进行了放大显示。从各个跟踪算法的结果可以清楚地看到，本章所提出的跟踪算法在如此复杂的背景下仍能取得比较稳定的跟踪结果。Traditional Kalman filter method (图 3-6c)在跟踪的初始阶段便丢掉了跟踪目标，同样，SIFT-based method (图 3-6b)同样在跟踪的初始阶段产生了跟踪错误。相比上述两种跟踪算法，基于特征选择的 Collins method 算法 (图 3-6a)取得了相对比较好的跟踪结果，尽管部分视频帧产生了跟踪错误（如 320<sup>th</sup> 视频帧），但是其在大部分的视频帧中都能比较好地对目标进行跟踪。其中跟踪错误的产生在一定程度上是由于 Collins 等所提到的特征选择算法没有考虑到特征变化的时间连续性，而是直接通过特征判别力及其相关的变化对特征进行排序和筛选，这就使得背景噪声很容易对跟踪目标的特征描述产生干扰，从而导致部分跟踪结果不正确。

为了更加形象并且量化地评价本章所描述的跟踪算法，我们对跟踪的视频进行了人工标注，然后分别对上述四种跟踪算法的跟踪结果与手工标注的结果进行比较。跟踪的效果采用跟踪错误率（displacement error rates, DER)来衡量，其定义如下：

$$DER = \frac{\text{Displacement error between tracked object position and groundtruth}}{\text{Size of the object}}.$$

从 DER 的定义可知：跟踪算法的性能越好，其 DER 便会越小；跟踪算法越稳定，其 DER 的在图像上表现的越平稳，这样便可以更加直观地比较上面提到的四种跟踪算法。在比较实验中，我们共计算了 10 段视频（如下表所示）的平均 DER 来评价上述四种跟踪算法。DER 的比较结果如下图所示。从图中可以看出，我们的跟踪算法的 DER 不仅平均值比其他三种跟踪算法的都要小，而且 DER 的变化也比较平坦(大约在 0.05 到 0.2 之间)，表明该算法更为鲁棒。

表 3-2 用于计算 DER 的视频文件列表

视频测试集 (Video test set)	视频名称 (Video name)
------------------------	-------------------

VIVID Tracking video set	redteam.avi
	egtest01.avi
	egtest02.avi
	egtest04.avi
SDL Tracking video set	xiangshan_0032.avi
	xiangshan_0043.avi
CARVIA Tracking video set	Browse1.avi
	Fight_Chase.avi
	OneStopMoveEnter1cor.avi
	EnterExitCrossingPaths2front.avi

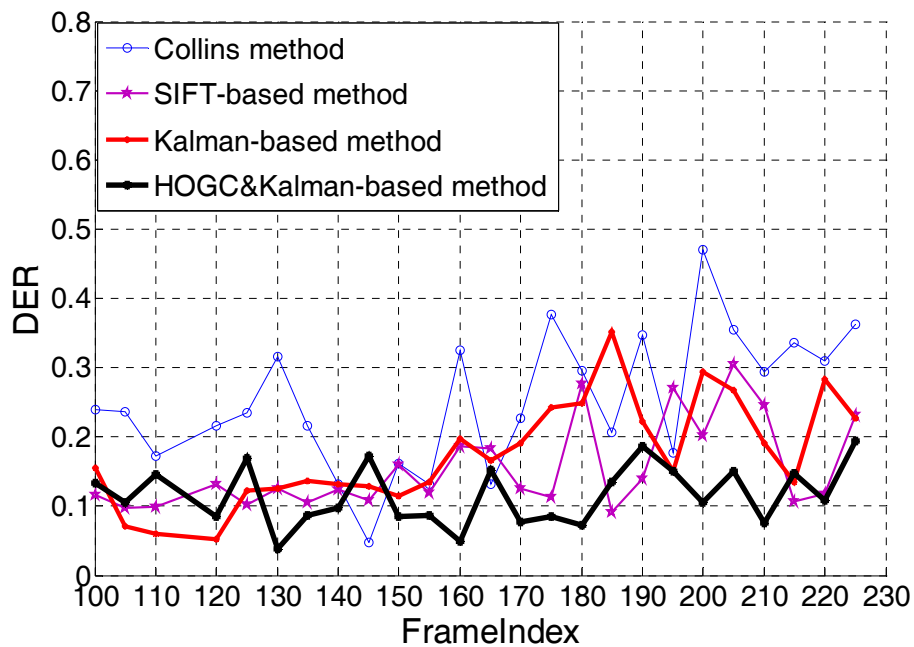
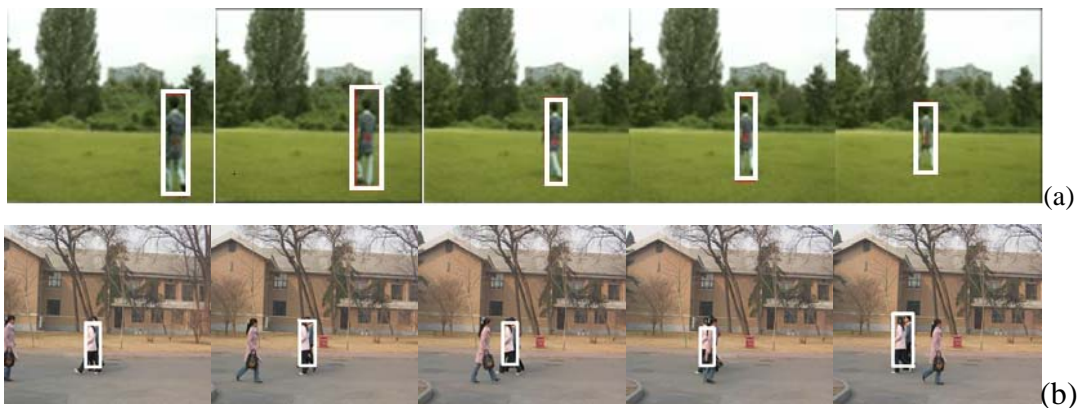


图 3-7 四种跟踪算法的 DER 比较结果图.

本章所提到算法在其他测试视频上的跟踪结果如下图所示：





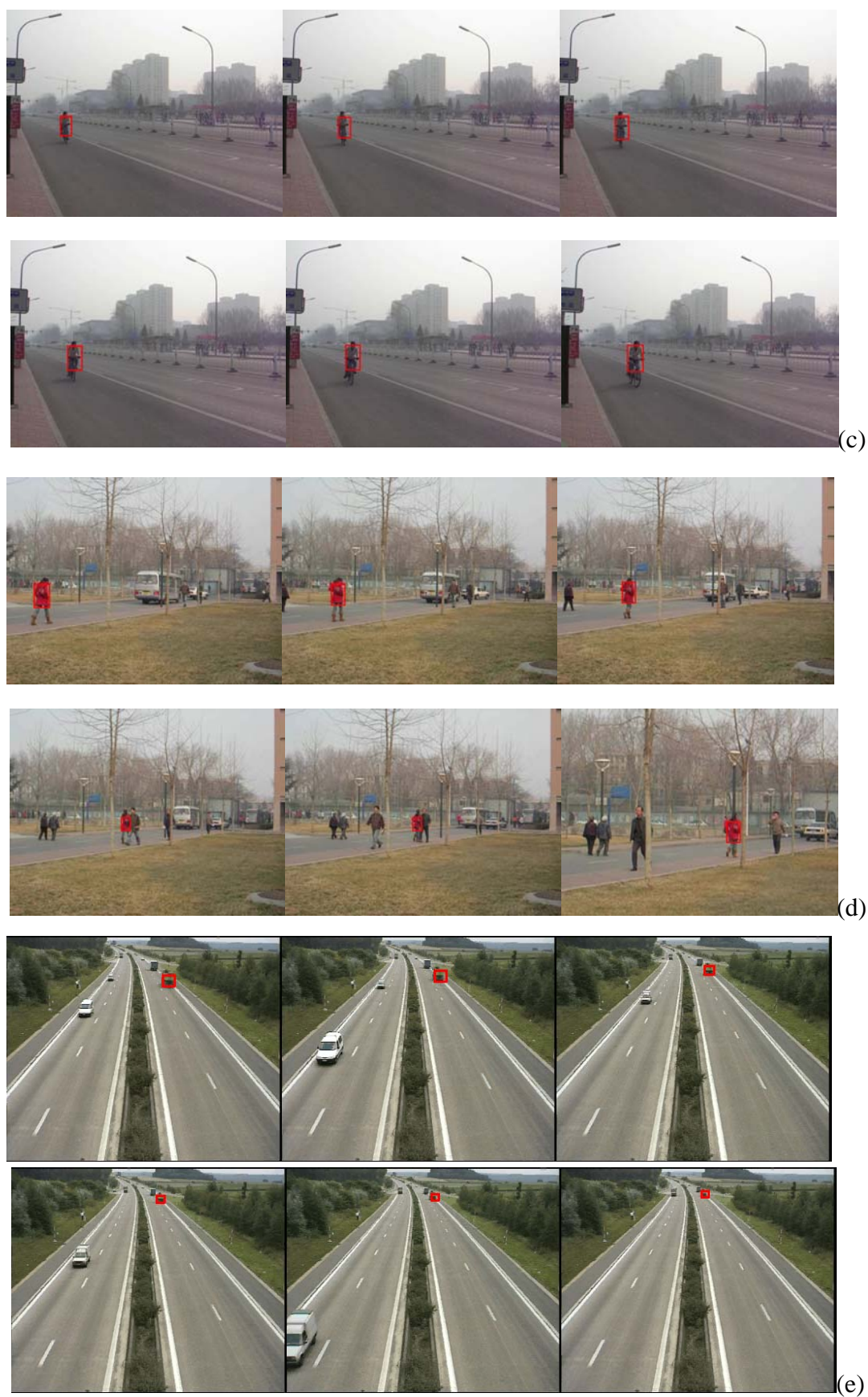


图 3-8 本章所提出算法的跟踪结果图

## 3.4 跟踪算法在测试系统中的应用

我们将本章所提到的跟踪算法嵌入到一个实际的应用测试系统中,其中包括一个 PTZ 和一个视频服务器。首先通过背景建模[49][51][65]得到运动的目标区域并对其进行初始化,然后利用 PTZ 随动地对目标进行跟踪以使得 PTZ 能够在目标的整个运动过程中都能观察到被跟踪的目标。系统结构图和跟踪结果如下所示:

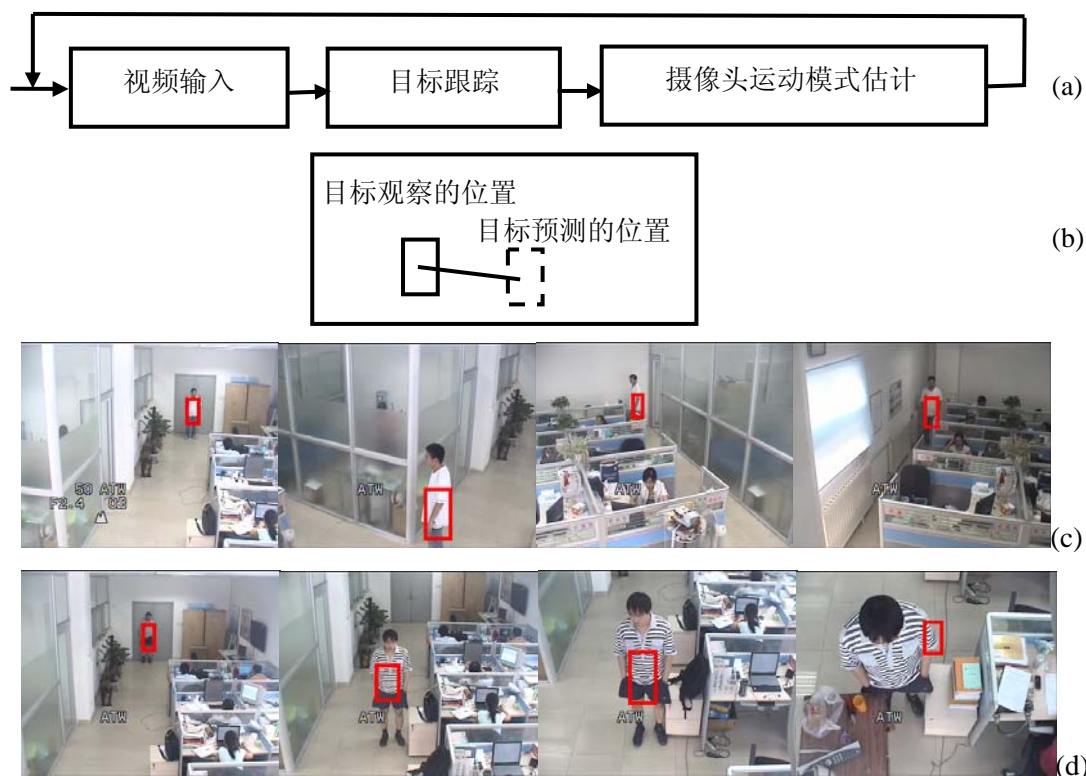


图 3-9 本章所提出算法在基于 PTZ 摄像头系统的应用. (a)基于 PTZ 的目标跟踪流程图, (b)跟踪目标示意图, (c)正确的跟踪结果和(d)错误的跟踪结果.



## 第四章 基于粒子滤波的特征评估算法

Kalman 滤波通常基于线性和高斯的假设，因此在一些情况下不能很好的预测和修正目标的特征权值，本文又提出了一种基于非高斯非线性假设的特征权值评估方法，即基于粒子滤波的特征权值评估。

### 4.1 粒子滤波算法及应用

粒子滤波 (Particle Filter)，亦称 Bootstrap 滤波、Condensation 算法、序列蒙特卡罗 (Sequential Monte Carlo) 滤波等，是一种目前研究较多的非线性非高斯环境下的状态估计算法。近年来引入视觉跟踪研究领域的序列蒙特卡罗滤波算法与 Kalman 滤波算法一样根植于贝叶斯时序滤波框架。在某种意义上，Kalman 滤波算法可以看作序列蒙特卡罗滤波算法的特例。序列蒙特卡罗滤波算法由于下述优点而得到了越来越多研究人员的重视：

- 1) 相对于 Kalman 滤波要求后验分布满足高斯形式，算法可以求解后验分布为任意函数形式的视觉跟踪问题；
- 2) 算法在理论上能够解决带有任意非线性特性的复杂视觉跟踪问题；
- 3) 算法能够在时间轴上传递条件分布的变化，从而系统地处理不确定性问题；
- 4) 当粒子数目足够多时，算法总能够得到收敛的跟踪结果；
- 5) 由于在贝叶斯的框架下进行推断，算法能够有效融合多通道信息。

#### 4.1.1 粒子滤波算法原理

作为蒙特卡罗方法 (Monte Carlo) 的一种，序列蒙特卡罗滤波器在很多研究领域中都有应用，如 Bootstrap filtering、Condensation、Particle filtering、Interacting Particle approximation 和 Survival of Fittest 都是序列蒙特卡罗滤波算法在不同领域中的名称。序列蒙特卡罗滤波器是通过蒙特卡罗模拟对贝叶斯时序滤波器的实现，该算法的基本思想是通过一组带有相关权值的随机样本，以及基于这些样本的估计来表示状态的后验概率分布  $p(x_k | z_{1:k})$ 。令  $\{x_{0:k}^i, \omega_k^i\}_{i=1}^{N_s}$  表示后验概率分布  $p(x_k | z_{1:k})$  的随机观测值，其中  $N_s$  是采样点数， $\omega_k^i$  是  $k$  时刻第  $i$  个采样点的权值，

并且满足  $\sum_i \omega_k^i = 1$ ，则  $k$  时刻后验概率分布  $p(x_k | z_{1:k})$  可近似表示为：

$$p(x_{0:k} | z_{1:k}) \approx \sum_{i=1}^{N_k} \omega_k^i \delta(x_{0:k} - x_{0k}^i) \quad (4-1)$$

上式中  $\omega_k^i$  可以通过重要采样 (Importance Sampling) 方法得到。若采样点集合  $x_{0:k}^i$  可以由某重要度函数  $q(x_k | z_{1:k})$  取得，则第  $i$  个粒子的归一化权值为

$$\omega_k^i \propto \frac{p(x_{0:k}^i | z_{1:k})}{q(x_{0:k}^i | z_{1:k})} \quad (4-2)$$

根据 Markov 过程基本概念，后验概率分布  $p(x_k | z_{1:k})$  亦可以表示为

$$\begin{aligned} p(x_{0:k} | z_{1:k}) &= \frac{p(z_k | x_k) p(x_k | x_{k-1})}{p(z_k | z_{1:k-1})} p(x_{0:k-1} | z_{1:k-1}) \\ &\propto p(z_k | x_k) p(x_k | x_{k-1}) p(x_{0:k-1} | z_{1:k-1}) \end{aligned} \quad (4-3)$$

#### 4.1.2 粒子滤波算法在目标跟踪中的主要应用

当系统对目标的运动位置和速度有了可靠估计，可以在相对较小的区域内进行搜索，完成对目标的模板匹配，而且当目标被遮挡时，利用粒子滤波可以对目标的运动轨迹进行可靠预测，方便在特定区域内搜索目标，等待目标的重新出现。

粒子滤波对于提高跟踪系统的处理速度和性能有着极其重要的作用。假设目标的运动状态参数为某一时刻目标的位置和速度，在跟踪过程中，由于相邻两帧图像的时间间隔比较短，因此我们可以利用粒子滤波器对目标的运动模型进行很好地估计，速度和速度的变化率足以反映目标的运动趋势。通常定义粒子滤波器的系统状态为  $x_k$ ，并且是一个六维向量  $(s_x, s_y, v_x, v_y, \nabla v_x, \nabla v_y)^T$ ，分别代表运动目标在坐标轴上的坐标、在坐标轴方向的速度以及速度的变化率。

粒子滤波算法中，即使在初始化目标的状态参数（如：位置）不准确的情况下，经过有限次的迭代运算，仍然可以准确地估计出目标的真实运动状态参数（如：位置），因此，具有预测的稳定性。

#### 4.2 基于粒子滤波特征评估

首先将目标的 HOGC 特征集合分为 12 个特征子集，其中包括 3 个颜色子集 (R、G 和 B) 和 9 个 HOG 子集。然后为每个特征子集定义一个粒子滤波，用



来独立地调整目标特征子集中特征的权值。具体定义如下：在时刻  $t$ ， $\{p_t(j)\} = \{(j^i, w_t(j^i))\}$ ,  $j = 1, 2, \dots, 12$ ，代表每个特征子集的 12 个粒子滤波，其中  $j$  表示 R, G, B 颜色子集或者一个 HOG 子集， $j^i$  表示第  $j$  个特征子集（粒子滤波）的第  $i$  维特征， $w_t(j^i)$  表示第  $j^i$  特征所对应的特征权值。

**表 4-1 基于粒子滤波的特征权值评估算法**

1. 初始化第  $j$  个特征子集（粒子滤波）中每个特征（Particle）相同的权值，其中  $M$  为该特征子集中特征的个数；
2. 对于第  $j$  个特征子集  $\{p_t(j)\}$ ,  $j = 1, 2, \dots, 12$ ，我们进行如下的操作：
  - 2.1 计算 每个 Particle 的累计权值，如下：
 
$$\begin{cases} c_t^0 = 0 \\ c_t^i = c_t^{i-1} + w_t(j^i) \end{cases} \quad i = 1, 2, \dots, M$$

得到新的集合  $Set_t(j) = (j^i, w_t(j^i), c_t^i | i = 1, 2, \dots, M)$ ；
  - 2.2 重采样 从  $Set_t(j)$  重采样选择  $k$  个 Particles，对于每个重采样得到的粒子  $i'$  按照如下的方法进行选择：
    - (a) 从 0 到 1 的均匀分布中，产生一个随机数  $r$ ， $r \in [0, 1]$ ；
    - (b) 找到最小的  $i$  使得  $c_t^i \geq r$ ；
    - (c) 然后  $i' = i$ ；
  - 2.3 预测 通过  $p(i|i')$  预测重采样得到的 Particle  $\{i' | i' = 1, 2, \dots, M\}$  的权值。在我们的实验中  $p(i|i')$  被定义为  $p(i|i') \sim N(0, \delta)$ ，其中  $\delta$  被设定为  $2M$ ；
  - 2.4 更新 通过  $p(S_{t+1}^i | i)$  来更新第  $i$  个特征（Particle）的权值  $w_{t+1}(j^i)$ ；在我们的实验中， $p(S_{t+1}^i | i)$  定义为相应特征的判别力；
  - 2.5 归一化  $\sum_i w_{t+1}(j^i) = 1$ 。
3. 在跟踪过程中，搜索目标的最适位置；
4.  $t = t + 1$ ；跳转到步骤 2 或者结束循环。

以某一个 HOG 特征子集为例，特征的权值更新过程如下图所示：

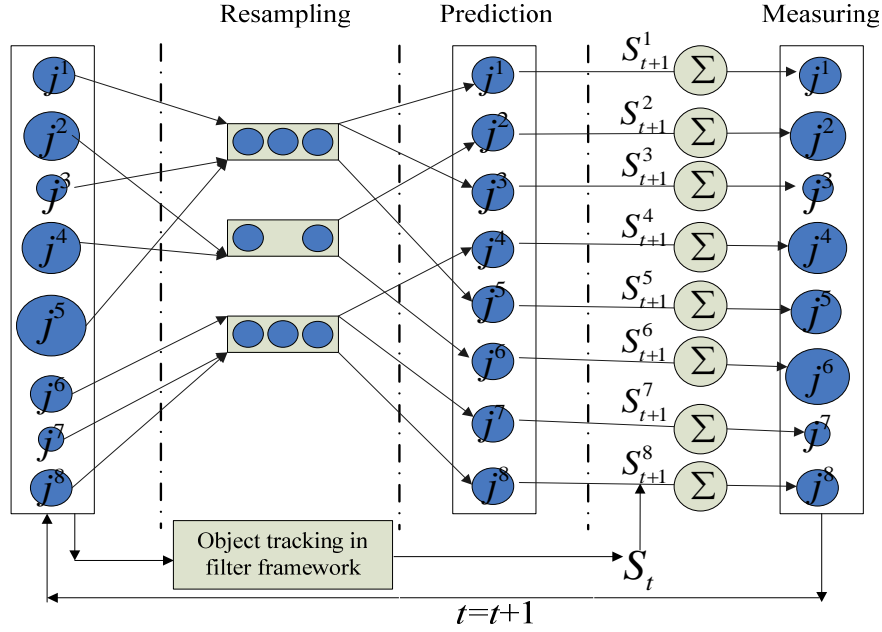


图 4-1 基于粒子滤波的特征评估方法示意图.

### 4.3 实验结果比较与算法性能分析

第一段测试视频来自于 VIVID data set, 其中跟踪目标为一辆和背景颜色极为相似的车辆, 视频同样是由直升机从空中拍摄所得。比较的算法包括本文提出的算法与另外三种传统的跟踪算法: 基于特征选择的跟踪方法 (Collins method) [11]、基于 SIFT 特征点的跟踪方法 (SIFT-Based method) [30] 和传统的粒子滤波跟踪方法 (traditional Particle Filter method) [39]。实验结果如下图所示, 其中 Collins method 在第 700 和 1350 帧中产生了跟踪错误, 由于 Collins method 只采用了颜色特征对目标进行跟踪, 所以当有背景颜色对其干扰时, 其往往会产生较差的跟踪结果; 由于背景比较的平滑, 所以不会产生大量的 SIFT 噪声点对目标的描述进行干扰, 因此 SIFT-Based method 取得了比较令人满意的跟踪结果; 同时, 由于跟踪目标与背景之间的相对运动比较简单, 所以 traditional Particle Filter method 能够得到很好的预测和修正目标的运动模型, 准确地得到目标在视频帧中的位置。

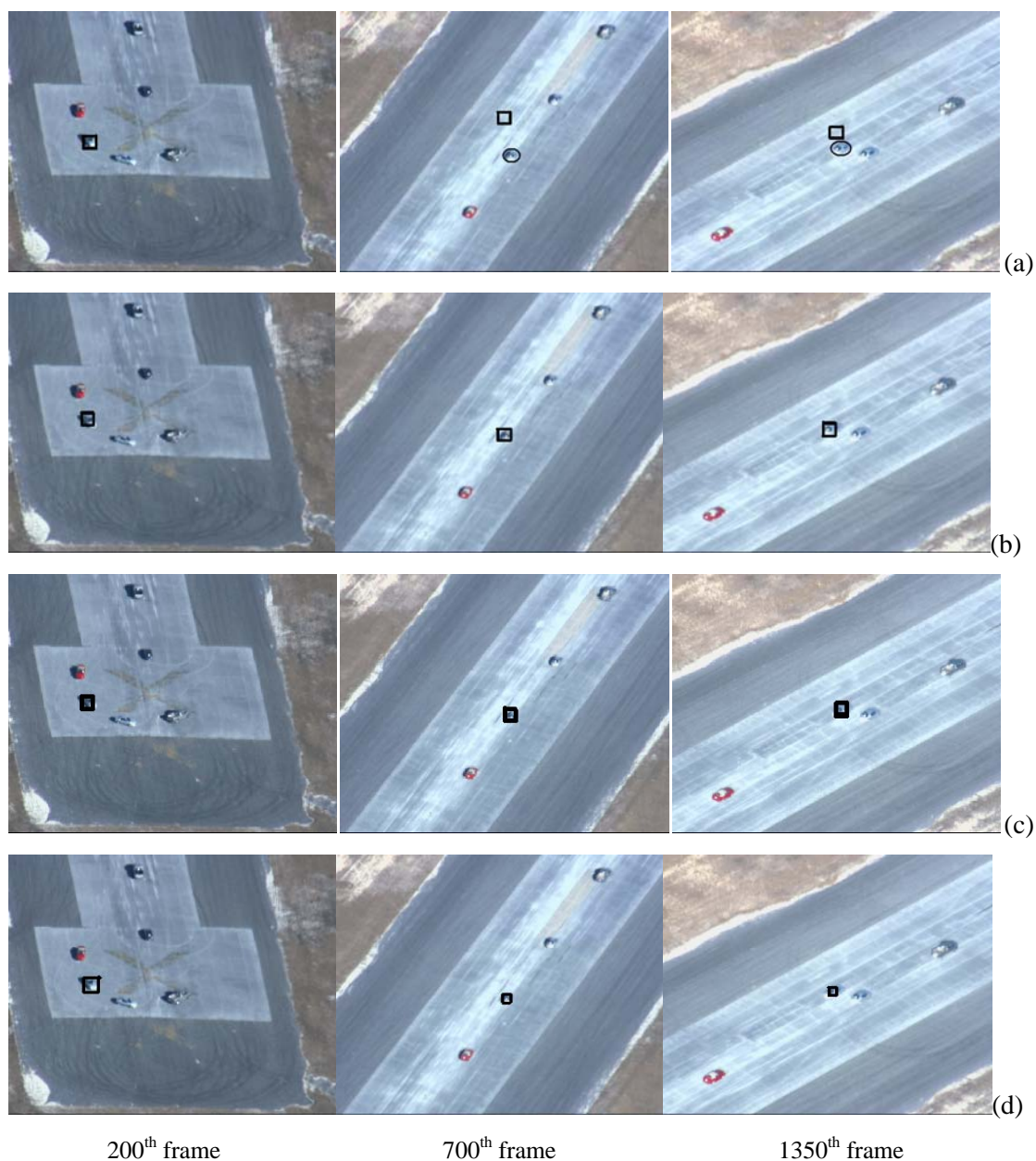


图 4-2 相似背景中的目标跟踪结果. 跟踪结果用黑框表示, 真实结果用黑色的椭圆表示. (a) Collins tracking method 的跟踪结果, (b) SIFT-based tracking method 的跟踪结果, (c) traditional Particle filter method 的跟踪结果和(d)我们所提出跟踪算法的跟踪结果.

在下图所示的实验结果比较中, 跟踪目标为人体, 其中跟踪目标与其他目标之间存在着遮挡且目标本身也具有尺度变化。从实验结果可以看出, 我们的跟踪算法仍然取得了最好的跟踪结果。traditional Particle Filter method 方法也取得了较为令人满意的跟踪结果, 但是基于 SIFT 的跟踪算法(图 4-2b)产生了一些跟踪错误, 这是由于在背景中存在着相似的人体, 对跟踪目标 SIFT 特征点的提取会产生很大的干扰, 从而使得跟踪结果不是特别稳定。Collins method 虽然也会根

据背景的变化来自动进行特征选择,但是由于自适应特征选择机制的缺陷使得当背景存在噪声干扰时,不能很好地跟踪目标,因此在第 829 帧中产生了跟踪错误。



图 4-3 存在遮挡和相似目标的跟踪结果. (a)Collins method 的跟踪结果, (b) SIFT-based method 的跟踪结果, (c) traditional Particle filter method 的跟踪结果和(d)本文所提出算法的跟踪结果.

下面仍然沿用上一章中所定义的 DER 来定量地比较本章中所提到的四种跟踪算法的整体性能,测试视频仍然选用为上一章中的 10 段视频,对比结果如下:



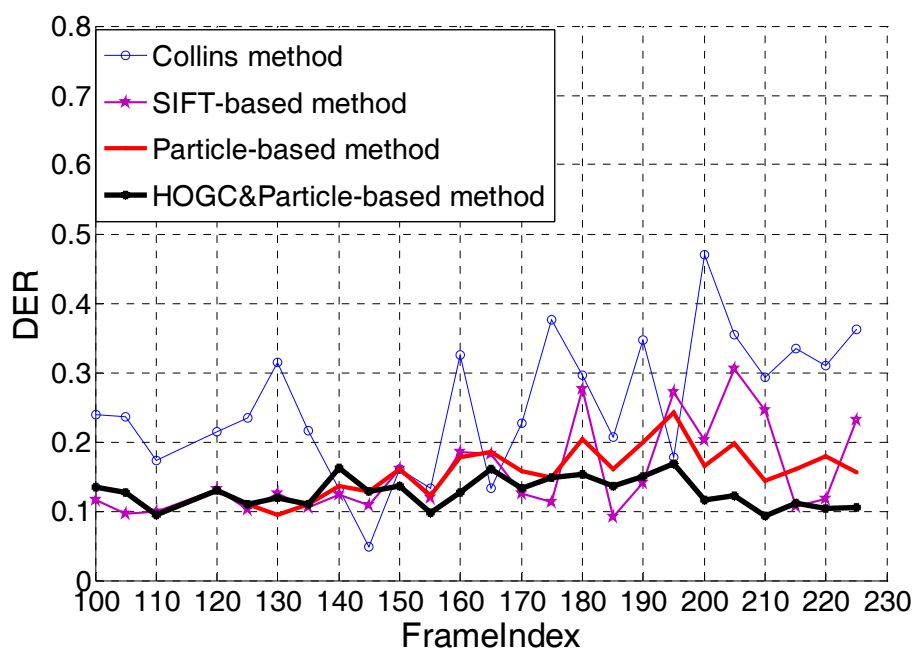
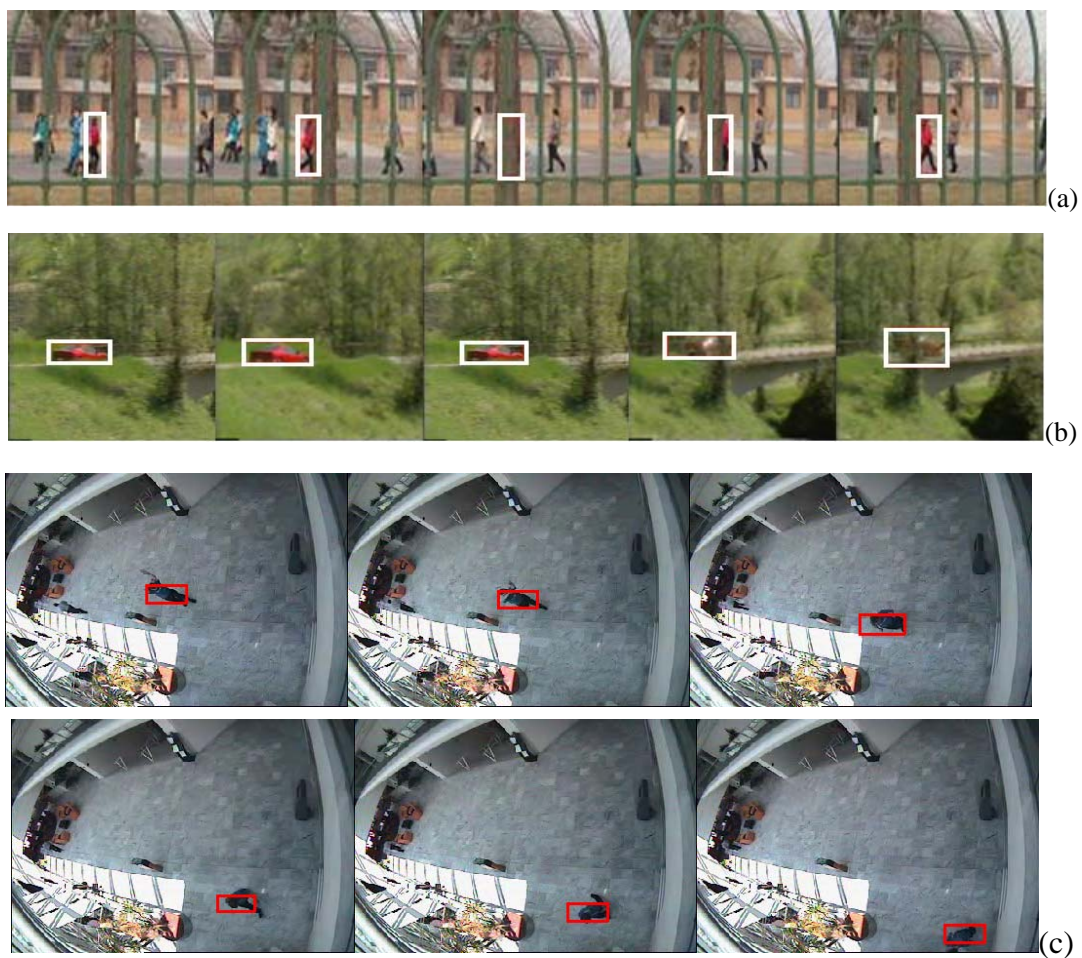


图 4-4 四种跟踪算法的 Displacement Error Rates (DER)的结果.

本章所提到的算法在其他测试视频上的跟踪结果如下所示:



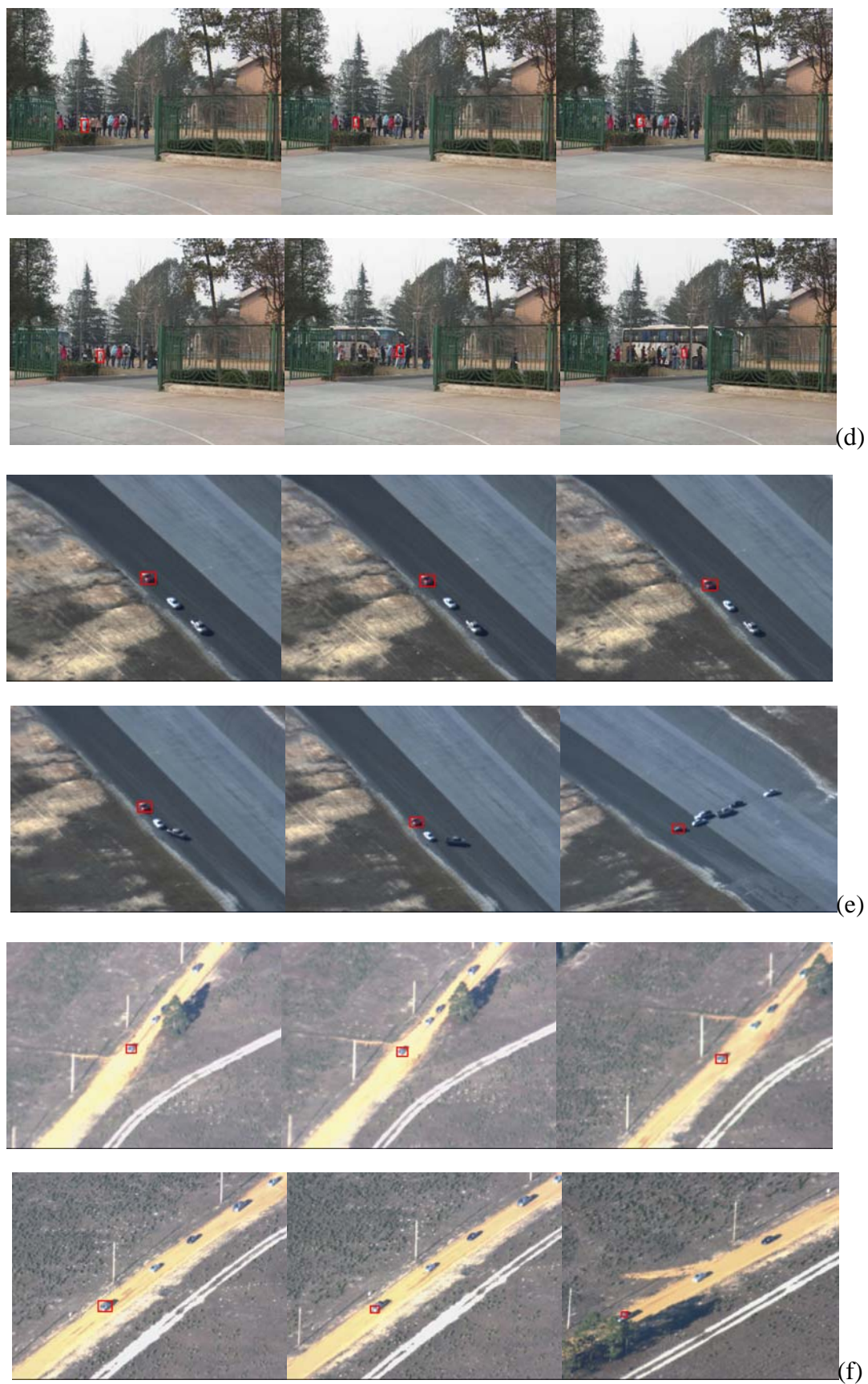


图 4-5 本章所提出跟踪算法的跟踪实验结果

跟踪效率是跟踪算法的另一个重要衡量标准。在实验比较中，在 **Pentium IV CPU (2.4G)** 的测试平台上，我们的两种跟踪方法基本可以达到近似实时处理的程度（20 帧/每秒），远快于 **SIFT-based tracking method** 的处理速度。





## 总结与展望

目前,跟踪特征的评估是视频跟踪中一个非常重要的研究课题。本文提出了一种新的在滤波框架中对跟踪特征进行评估的算法,并且分别通过 Kalman 滤波与粒子滤波对该算法进行了验证。实验结果证明,与其他跟踪算法相比本文所提出的基于滤波框架的特征权值调整策略和跟踪算法的有效性。定量实验对比表明我们的跟踪算法在动态复杂背景下对目标进行跟踪的优势更加明显。同时,我们也将基于 Kalman 滤波权值调整算法嵌入到 PTZ 系统中,对所提出的跟踪算法在实际应用中进行了验证。

本文提出的新概念和新技术主要包括:基于颜色和梯度方向的综合直方图、基于跟踪特征判别力和滤波框架的跟踪特征权值评估机制。其中,综合直方图可以很容易地融合其他一些特征,例如纹理特征等。同时,所提到的滤波框架也不仅仅局限于 Kalman 或者粒子滤波算法,其他类似的滤波算法也可以在此框架中进行跟踪特征权值评估。

由于时间和经历的关系,我们没有对 Kalman 滤波或者粒子滤波的理论进行深入的研究并对其进行改进以使其对目标运动的描述更加精确,因此,在某些情况下,我们得到的跟踪结果并没有某些基于运动预测的跟踪算法好。另一个不足在于,我们提出的综合特征集在跟踪目标的尺度发生很大变化的时候,往往容易产生跟踪错误。在将来的工作中,我们拟提出更多的对光照和尺度变化不敏感的特征并将其融合到所提出的综合特征中,同时,将对跟踪目标的模板更新问题进行深入地研究,使得所提出的跟踪方法能够在目标具有大尺度变化的情况下仍然能对其进行鲁棒的跟踪。



## 参考文献

- [1] C. Stauffer, and W. Grimson. Adaptive background mixture models for real-time tracking. *Proceedings of IEEE Conference on CVPR*, 1999:246–252.
- [2] G. Bradski. Real time face and object tracking as a component of a perceptual user interface. *Proceedings of IEEE Workshop on Applications of Computer Vision*, 1998:214–219.
- [3] N. Papanikolopoulos, P. Khosla, and T. Kanade. Visual tracking of a moving target by a camera mounted on a robot: a combination of control and vision. *IEEE Transactions on Robotics and Automation*, 1993:14–35.
- [4] S. Avidan. Ensemble tracking. *Proceedings of IEEE Conference on CVPR*, 2005:494–501,.
- [5] J. Shi, and C. Tomasi. Good features to track. *Proceedings of IEEE Conference on CVPR*, 1994:593–600,.
- [6] M. Isard, and A. Blake. CONDENSATION-conditional density propagation for visual tracking. *International Journal on Computer Vision*, 1998:5–28.
- [7] D. Comaniciu, V. Ramesh, and P. Meer. Kernel-based object tracking. *IEEE Trans. PAMI*, 2003:564–577.
- [8] S. Baker, and I. Matthews. Lucas-Kanade 20 years on: A unifying framework. *International Journal on Computer Vision*, 2004:221–255.
- [9] D.W. Liang, Q.M. Huang, W. Gao, and H.X. Yao. Online selection of discriminative features using Bayes error rate for visual tracking. *7th Pacific-Rim Conference on Multimedia*, 2006:547–555.
- [10] D. Chen, and J. Yang. Robust object tracking via online spatial bias appearance model learning. *IEEE Trans. PAMI*, 2007:2157–2169.
- [11] R. Collins, and Y. Liu. Online selection of discriminative tracking features. *Proceedings of Ninth IEEE ICCV*, 2003:346–352.
- [12] H. Nguyen, and A. Smeulders. Tracking aspects of the foreground against the background. *Proceedings of ECCV*, 2004:446–456.
- [13] J. Wang, X. Chen, and W. Gao. Online selecting discriminative tracking features using Particle Filter. *Proceedings of IEEE Conference on CVPR*, 2005:1037–1042.
- [14] J.Q.Wang, and Y.S.Yagi. Integrating color and shape-texture features for adaptive real-time object Tracking. *IEEE Trans. Image Processing*, 2008:235–240.
- [15] H. Bohyung, and L. Davis. Object tracking by adaptive feature extraction. *IEEE International Conference on Image Processing*, 2004:1501–1504.

- [16] Weiming Hu, Dan Xie and Tieniu Tan. A hierarchical self-organizing approach for learning the patterns of motion trajectories. *IEEE Transactions on Neural Networks*, 2004: 135-144.
- [17] 李美花. 三维人体跟踪. 大连理工大学硕士学位论文, 2004.
- [18] T. Yang, Q. Pan, J. Li, and S.Z. Li. Real-time multiple objects tracking with occlusion handling in dynamic scenes. *IEEE Conference on CVPR*, 2005:970-975.
- [19] 彭宁嵩, 杨杰, 刘志, 张凤超. Mean-Shift 跟踪算法中核函数窗宽的自动选取. *软件学报*, 2005:1542-1550.
- [20] D. Comaniciu, V. Ramesh, and P. Meer. Kernel-based object tracking. *IEEE Transaction on Pattern Analysis and Machine Learning*, 2003:564-577.
- [21] A. Elgammal, R. Duraiswami, and L. Davis. Probabilistic tracking in joint feature-spatial spaces. *IEEE Conference on Computer Vision and Pattern Recognition*, 2003:781-788.
- [22] R. Collins. Mean-shift blob tracking through scale space. *IEEE Conference on Computer Vision and Pattern Recognition*, 2003:234-240.
- [23] M. Isard, and A. Blake. Condensation - conditional density propagation for visual tracking. *International Journal of Computer Vision*, 1998:5-28.
- [24] T.X. Han, V. Ramesh, Y. Zhu, and T.S. Huang. On optimizing template matching via performance characterization. *International Conference on Computer Vision*, 2005:182-189.
- [25] B.Han, Y. Zhu, D. Comaniciu, and L. Davis. Kernel-based Bayesian filtering for object tracking. *IEEE Conference on Computer Vision and Pattern Recognition*, 2005:227-234.
- [26] Y. Huang, and I Essa. Tracking multiple objects through occlusions. *IEEE Conference on Computer Vision and Pattern Recognition*, 2005:1051-1058.
- [27] Y. Wu. Robust visual tracking by integrating multiple cues based on co-inference learning. *International Journal of Computer Vision*, 2004:55-71.
- [28] Z. Khan, T. Balch, and F. Dellaert. A rao-blackwellized particle filter for eigentracking. *IEEE Conference on Computer Vision and Pattern Recognition*, 2004:980-986.
- [29] P. Perez, C. Hue, J. Vermaak, and M. Gangnet. Color-based probabilistic tracking. *European Conference on Computer Vision*, 2002:661-675.
- [30] David G. Lowe. Distinctive image features from scale-invariant key points. *International Journal of Computer Vision*, 2004:91-110.
- [31] K. Okuma, A. Taleghani, N. de Freitas etc. A boosted particle filter: multi target detection and tracking. *European Conference on Computer Vision*, 2004:28-39.
- [32] Z.j. Han, Q.x. Ye, and J.b. Jiao. Online Feature Evaluation for Object Tracking using Kalman Filter. *International Conference on Pattern Recognition*, 2008.
- [33] Han, Z., Ye, Q., Liu, Y., et al. Feature evaluation by particle filter for adaptive object tracking. *Proceedings of SPIE*, 2009.

- [34] N. Dalal, and B. Triggs. Histograms of oriented gradients for human detection. Proceedings of IEEE Conference on CVPR, 2005:886–893.
- [35] W.T. Freeman, and M. Roth. Orientation histograms for hand gesture recognition. IEEE Intl. Wkshp. on Automatic Face and Gesture Recognition, 1995:296–301.
- [36] Q. Zhu, S. Avidan, M.C. Yeh, and K.T. Cheng. Fast human detection using a cascade of histograms of oriented gradients. Proceedings of IEEE Conference on CVPR, 2006.
- [37] R. G. Brown, and P. Y. C. Hwang. Introduction to random signals and applied Kalman Filtering. New York:John Wiley & Sons,Inc., 1992.
- [38] E. Cuevas, D. Zaldivar, and R. Rojas. Kalman filter for vision tracking. Technical Report B, Fachbereich Mathematikund Informatik, Freie Universität Berlin, 2005.
- [39] E. Cuevas, D. Zaldivar, and R. Rojas. Particle filter for vision tracking. Technical Report B, Fachbereich Mathematikund Informatik, Freie Universität Berlin, 2005.
- [40] D. Comaniciu, V. Ramesh, and P. Meer. Real-time tracking of non-rigid objects using mean shift. Proceedings of IEEE Conference on CVPR, 2000:142–149.
- [41] S. Birchfield. Elliptical head tracking using intensity gradients and color histograms. Proceedings of IEEE Conference on CVPR, 1998:232–237.
- [42] G. Hager, and P. Belhumeur. Efficient region tracking with parametric models of geometry and illumination. IEEE Trans. PAMI, 1998:1025–1039.
- [43] A. Yilmaz, X. Li, and B. Shah. Contour based object tracking with occlusion handling in video acquired using mobile cameras. IEEE Trans. PAMI, 2004:1531–1536.
- [44] J.y. Pan, B. Hu, and J.Q. Zhang. Robust and accurate object tracking under various types of occlusions. IEEE Trans. CSVT, 2008:223–236.
- [45] K. Hariharakrishnan and D. Schonfeld. Fast object tracking using adaptive block matching. IEEE Trans. Multimedia, 2005:853–859.
- [46] I. Matthews, T. Ishikawa, and S. Baker. The template update problem. IEEE Trans. PAMI, 2004:810–815.
- [47] VIVID Tracking Evaluation Web Site at: <http://www.vividevaluation.ri.cmu.edu/datasets/datasets.html>.
- [48] CAVIAR Test Case Scenarios at: <http://homepages.inf.ed.ac.uk/rbf/CAVIAR>.
- [49] A. Monnet, A. Mittal, and N. Paragios. Background modeling and subtraction of dynamic scenes. Ninth IEEE International Conference on Computer Vision, 2003:1305–1312.
- [50] B. Lucas, and T. Kanade. An iterative image registration technique with an application to stereo vision. Proceeding of DARPA Image Understanding Workshop, 1981:121–130.
- [51] T.F. Cootes, G.J. Edwards, and C.J. Taylor. Active appearance models. IEEE Trans. PAMI, 2001:681–685.
- [52] Catalin. G, and Nedevschi. S. Object tracking from stereo sequences using particle filter. Intelligent Computer Communication and Processing, 2008:279-282.

- [53] Huaping Liu and Fuchun Sun. Particle Filter with Improved Proposal Distribution for Vehicle Tracking. *Advances in Neural Networks* ISSN 2008, 2008:422-431.
- [54] SDL data set at: <http://cear.gucas.ac.cn/SDL-HomePage/Dataset.htm>.
- [55] Y. Qian, G. Medioni, and I. Cohen. Multiple target tracking using spatial-temporal Markov chain Monte Carlo data association. *Proceedings of IEEE Conference on CVPR*, 2007:1–8.
- [56] Y. Alper, J. Omar, and S. Mubarak. Object tracking: A survey. *ACM Computing Surveys*, 2006.
- [57] N. Pérez de la Blanca, J.M. Fuertes, and M. Lucena. Deformable Object Matching Based on Multi-scale Local Histograms. *AMDO 2004*, 2004:154-162.
- [58] Tony Lindeberg. Feature Detection with Automatic Scale Selection. *International Journal of Computer Version*, 1998.
- [59] Bastian Leibe, and Bernt Schiele. Scale-Invariant Object Categorization using a Scale-Adaptive Mean-Shift Search. *DAGM'04 Pattern Recognition Symposium*, 2004.
- [60] 赵建伟, 刘重庆. 适用于遮挡问题的目标跟踪算法. *计算机辅助设计与图形学学报*, 2003.
- [61] 薛雪, 刘泽平, 丁艳. 区域增长的图像跟踪算法的研究. *光学技术*, 2005.
- [62] 胡明昊, 任明武, 杨静宇. 一种基于直方图模式的运动目标实时跟踪算法. *计算机应用与研究*, 2004.
- [63] 常发亮. 彩色图像分割与复杂场景下视觉目标跟踪方法研究. 山东大学博士学位论文, 2006.
- [64] Tao Zhang, and Daniel Freedman. Tracking Objects Using Density Matching and Shape Priors. *International Conference on Computer Vision*, 2003.
- [65] A. Elgammal, D. Harwood, and L. Davis. Non-Parametric Model for Background Subtraction. *Proc. IEEE Frame Rate Workshop*, 1999.

## 发表文章目录

### 已录取文章目录

- **Z.j. Han**, Q.x. Ye, and J.b. Jiao, “Online Feature Evaluation for Object Tracking using Kalman Filter,” International Conference on Pattern Recognition, 2008.
- **Han, Z.**, Ye, Q., Liu, Y., et al. Feature Evaluation by Particle Filter for Adaptive Object Tracking. Proceedings of SPIE, 2009.
- Zhihua Liu, **Zhenjun Han**, Qixiang Ye, Jianbin Jiao. An improved segment-based algorithm in stereo matching. IEEE international conference on Mechatronics and Automatic, 2009.
- 常向魁, 叶齐祥, 刘先省, 焦建彬, **韩振军**. 基于综合色度和梯度方向直方图的运动目标跟踪算法.

### 审稿文章目录

- **Z.j. Han**, Q.x. Ye, and J.b. Jiao+, “Feature Evaluation in Filter Framework for Adaptive Object Tracking,” submitted to IEEE Trans on Circuits and Systems for Video Technology, 2008.
- 张宇, **韩振军**, 焦建彬, 一种基于综合特征评估的运动目标跟踪算法, 中国科技大学学报.

### 国家发明专利

- 焦建彬, 叶齐祥, **韩振军**, 基于 Internet 和 GRPS 的智能视频监控系统, 中国发明专利, 批准日期: 2008 年.





## 致谢

在攻读硕士学位三年的学习生活中，我经历了诸多坎坷，也付出了艰辛的努力，同时也得到了很大的收获。在毕业论文完成之际，由衷地感谢这三年来曾经给予我无数帮助的老师、同学、朋友和家人。

首先，本课题的研究工作是在导师焦建彬教授的悉心指导下完成的。焦建彬教授在我的学习中，从理论和实验方面给予了大量的、极其有益的建议和具体的指导，并在论文的撰写和审稿中倾注了大量的心血，他诲人不倦的精神和对我的谆谆教导，使我从恩师身上学到了不少可贵的东西，让我受益匪浅。此外，他还在生活和思想上关心我、照顾我，再次表示万分的感谢！尤其是恩师严谨的治学态度、积极的进取精神、平易近人的工作作风将使我终身受益。

其次，特别要感谢叶齐祥老师在三年的学习与生活中对我的极大帮助和支持，他渊博的专业知识和精益求精的动手能力，在理论学习和项目开发过程中给我提供了耐心的引导和帮助。

再次，感谢同实验室的常向魁、周锐、李志远、刘志花、徐冉等同学，他们在我的论文工作和撰写中，给予了我巨大的支持和帮助。

我还要衷心感谢我敬爱的亲人们，特别是我的未婚妻崔少娟，她对论文的校对与修改给予了我许多的帮助。亲人们对我工作和学业的热情鼓励和殷切期望，都给我注入了巨大的动力，他们无私的爱和无条件的支持一直指引我奋发拼搏、积极向上，我将对此终生铭记。