论文编码：TP39

首都师范大学
Capital Normal University

首都师范大学本科毕业设计论文

基于多视角视频采集与分析的智能鼠标算法研究

——目标跟踪

An algorithm research of smart mouse based on multi-view video capture and analysis

—— Target tracking

论文作者_____
院　　系　信息工程学院
专　　业　信息工程
学　　号_____
指导老师_____

完成日期　2013 年 4 月 22 日

首都师范大学信息工程学院

摘　　要

随着计算机技术的发展，人机交互越来越成熟，图形人机交互具有简单、易懂、直观等优点，因此吸引了越来越多的人们去研究。

本文主要研究目标跟踪算法的设计与实现。Mean-Shift 算法是基于颜色信息的跟踪算法中应用最广泛的算法，能够对具有特定颜色信息的运动目标进行跟踪，在简单的跟踪背景下，该算法计算量小，对局部遮挡、目标形变具有一定的适应性。本文将以 Mean-Shift 算法为基础进行继续研究。

本文对传统 Mean-Shift 算法进行了两处改进：第一，Mean-Shift 算法跟踪快速移动目标时，经常会出现跟踪丢失，本文提出了基于肤色信息的位置修正算法，可以解决因运动目标运动速度过快导致的目标跟踪丢失问题。第二，当目标进入相似颜色信息干扰的复杂背景时，Mean-Shift 算法经常会出现跟踪丢失。Kalman 滤波具有预测功能，将传统 Mean-Shift 算法和 Kalman 滤波融合，可以提高 Mean-Shift 算法执行效率和跟踪准确性。

最后本文提出了基于 Mean-Shift 算法的 KSMS 跟踪算法(Kalman Skin Mean-Shift)，将基于肤色信息的 Kalman 滤波和 Mean-Shift 算法进行融合，解决因运动目标移动速度过快或者相似颜色信息干扰情况下的跟踪丢失问题。实验结果表明，KSMS(Kalman Skin Mean-Shift)跟踪算法对因目标运动速度过快、相似颜色信息干扰导致的跟踪丢失问题具有良好的鲁棒性。

关键词： 目标跟踪；Mean-Shift 算法；Kalman 滤波；肤色信息

首都师范大学信息工程学院

Abstract

With the development of computer technology, human-computer interaction is coming more and more mature, interactive graphics has the advantages of simple, intuitive and so on. So it has attracted more and more people to study.

This paper mainly studies the design and realization of the target tracking algorithm. Mean-Shift algorithm based on color information and most widely used in tracking. Mean-Shift algorithm can track moving targets with specific color information with small calculation when in simple background, with a certain degree of adaptability when partial occlusion or target deformation. This paper will be further research based on Mean-Shift algorithm.

We have two improvements to the traditional Mean-Shift algorithm in this paper. Firstly, Mean-Shift algorithm often lost target when the target moving fast. The position correction algorithm based on skin color information, can solve the problems of tracking lost caused by the target moving fast. Secondly, when the target runs into the complex background of similar color information interference, Mean-Shift often lost target. Kalman filter has the function of prediction. The integration of Mean-Shift and Kalman can improve efficiency and tracking accuracy.

Finally, we proposed KSMS tracking algorithm, the integration of Kalman based on the constraints of complexion and Mean-Shift algorithm to solve the problem of tracking loss due to moving fast or similar color interference. Experimental results show that the KSMS tracking algorithm has good robustness to tracking lost cased by moving fast or similar color information interference.

Key Words: target tracking; Kalman filter; mean-shift algorithm; complexion

首都师范大学信息工程学院

目　录

第一章 绪论 1
1.1 论文的选题背景 1
1.2 目标跟踪发展现状 2
1.3 本文的研究内容以及目的 5
1.4 论文章节安排 6
第二章 Mean-Shift 算法 8
2.1 目标跟踪技术 8
2.2 总体分布的非参数估计以及 Parzen 窗估计 9
2.2.1 总体分布的非参数估计 9
2.2.2 Parzen 窗估计 10
2.3 Mean-Shift 算法 10
2.3.1 Mean-Shift 基础 11
2.3.2 Mean-Shift 算法 11
2.3.3 算法实现步骤 13
2.4 位置修正算法 14
2.5 Mean-Shift 算法跟踪结果 19
第三章 KSMS 跟踪结果 21
3.1 卡尔曼滤波 21
3.2 Kalman 滤波算法 22
3.2.1 建立系统数学模型 22
3.2.2 滤波器计算原型 23
3.2.3 Kalman 滤波预测阶段 24
3.2.4 Kalman 滤波的更新 24
3.2.5 Kalman 滤波参数设置 25
3.3 融合的 KSMS 跟踪算法 25
3.4 KSMS 跟踪结果 27
第四章 实验平台与结果评估 28
4.1 实验平台介绍 28
4.2 跟踪结果对比 29
4.3 结论与讨论 30
第五章 总结与设想 32
参考文献 33
本科期间的科研情况 35
致　谢 36

首都师范大学信息工程学院

第一章 绪论

1.1 论文的选题背景

随着计算机技术的飞速发展，图像信息处理、人机交互技术得到了飞速推广。目标检测、目标跟踪、模式识别等高新技术的应用呈现出日新月异的发展趋势。研究者们提出了很多高效的算法，使得计算机能够实现实习检测与目标跟踪。

随着科技的进步，信息采集技术越来越成熟，可以采集的信息种类也越来越丰富，并对我们生活产生了积极的影响。在我们生活中随处可见信息采集，例如街道两旁的监控器，它可以采集过往车辆的各种信息。当车辆超速时，它可以记录车辆的车牌号，并将数据传给交通部的数据处理中心，最后由交通部决定对违规车辆的处罚。当出现入室抢劫等违法案件时，警察会对案发现场进行数据采集，指纹数据是最好的破案信息。通过对指纹数据的分析，可以直接知道作案人的姓名等基本信息，因为国家已经对每位公民进行了指纹数据存档，通过比对指纹信息就可以直接确定作案人选。

采集了足够的数据后，对数据的分析也是非常重要的。只有充分利用好数据，才能达到我们的目的。摄像头采集的视频信息是我们生活中最常见的信息。视频信息里面包含大量的数据，我们只需对我们感兴趣的部分进行分析即可得出我们需要的结果。近年来迅速发展的人机交互技术就利用了视频采集技术，通过对视频信息的分析完成人机交互。

人机交互技术发展迅猛，越来越高端的交互技术融入到了人们的日常生活。传统的鼠标已经不能完全满足人们的应用需求，同时伴随着更加高端的视觉、语音、行为交互技术，一种新型的视觉交互鼠标应运而生。使得人们可以轻松地在电脑摄像头前通过手势来完成本应由传统鼠标来完成的任务。然而该设计需要提取并分析手部所做出的动作来进行相应的鼠标操作，必须要跟踪手部以获取所需的手部信息。因此，需要更加精准的目标跟踪算法来实现。 运动目标跟踪是计算机视觉研究的核心内容之一，是一项融合图像处理、模式识别、人工智能、自动控制等多种不同领域先进成果的高技术课题。

研究并不断完善精准的目标跟踪算法并实现该设计，其意义在于更加深入理解目标跟踪的基本算法以及加深将算法理论与算法实践相结合的能力。

在我们日常生活中，目标跟踪也很常见。目前目标跟踪的应用领域很多，有军用也有民用。在军事上，战斗机上面安装的有源相控阵雷达就完成目标检查和目标跟踪的任务。在战争中，雷达扫描到敌方战机，指引导弹对目标进行轰炸。雷达将时刻调整导弹航向，

首都师范大学信息工程学院

不断修正跟踪轨迹，以应对敌方目标位置的变化。在这个过程中，就要求目标跟踪的准确性和实时性要非常好。很容易理解，一旦目标跟踪延迟，导弹就一定不会击中目标，从而留下后患，将自己置于危险之中。因此目标跟踪的实时性和准确性就对一个国家的安全就起到了举足轻重的作用。军用导航卫星是一个国家军事实力的具体表现，有了自己的导航系统，就可以在危机时刻，把命运掌握在自己手中。要想实现对全球任何地点进行精确打击，就必须有自己的导航系统。我国发展北斗导航系统就是为了保护我国的国家安全，保护我国领土不受侵犯。导航卫星对敌人进行实时跟踪，为精确打击做好准备。

在智能交通系统中，目标跟踪同样很常见。管理部门通过设置在道路附近的监控摄像头收集获取过往车辆的运动信息及车辆信息。对过往车辆进行跟踪分析，获取车辆的运动速度，一旦车辆超速行驶，就会记录车辆的信息，例如车牌号等，并对车主进行罚款等处罚。当出现交通事故时，如果肇事车主逃跑，可以调取监控录像分析车辆的信息，对车辆的运动路线进行跟踪，从而抓获肇事人。目标检测、目标跟踪、目标分析起到了决定性作用。

在人机交互领域，目标跟踪同样很重要，也是研究热点。近年来，随着机器视觉等技术的发展，人民已经不能满足与传统的以计算机鼠标或者键盘为主的人机交互方式，而更期待计算机能够以更人性化、更自由的方式与人交流。要实现这一目的，计算机就必须要获得操作者的运动信息，并识别运动方式并转换为控制信息从而和计算机进行交流。《基于多视角视频采集与分析的智能鼠标算法研究》就是利用摄像头进行视频采集，通过手势的不同方式表示不同含义与计算机进行交流并控制计算机实现一定的操作。在整个系统中，首先要检测出手，并对手势进行实时跟踪，并分析手势的含义从而操控计算机执行相应命令。

本文将主要研究《基于多视角视频采集与分析的智能鼠标算法研究》中的目标跟踪模块，通过不同视角的摄像头对计算机前的手势进行目标检测和目标跟踪，时刻记录手势的位置坐标，为后续目标分析、操作识别做准备。

1.2 目标跟踪发展现状

目标跟踪是视觉领域研究的热点问题，随着计算机技术的发展而得到迅猛的发展，目标跟踪技术也由此取得了长足的进步。上世纪对图像的处理主要是集中于单幅图像的处理，即使在动态图像序列中跟踪运动目标，也带有浓厚的静态图像处理的特点。直到上世纪 80 年代 HomBK 等人提出了光流法，目标跟踪研究才真正意义上步入了动态图像序列的

首都师范大学信息工程学院

研究领域。光流法对于现阶段的计算机处理速度提出了极大的挑战，在实际应用领域很难满足实时性的要求。另外，视频序列存在的噪声会对光流法跟踪产生极大的干扰，因此光流法现阶段很难应用到实际场合。

基于视觉跟踪的巨大实用价值和计算机技术的迅猛发展，视觉跟踪技术的强烈需求，欧美等发达国家对视频跟踪技术进行了大量深入的研究。1997 年，由美国国防高级研究项目署牵头，以美国卡耐基梅隆大学为首，麻省理工学院等高校参与了视觉监控项目 VSAM 即（Visual surveillance and Monitorings）的研究开发，研究应用于战场及民用场景的监控，卡内基梅隆大学由此建立了第一个校园监控系统。在智能监控方面，Maryland 大学研发的实时监控系统 W 实现了人的跟踪，可用于对人的行为监视，并可以判断人是否携带物体等简单行为。英国雷丁大学的车辆交通监控系统则是对车辆和行人的跟踪及交互作用识别相关方面的研究。在国内，视觉跟踪技术自 1986 年开始立项研究，视觉跟踪领域的理念和技术也有了长足的发展，中国科学院、浙江大学、武汉大学等高校和研究所在目标跟踪、运动行人的检测识别方面取得了一定的成果。

以上介绍的主要是视觉跟踪在民用方面的应用，军事上视觉跟踪同样有着广泛的应用。现代军事战争中，利用飞机在空中对特定目标进行打击是行之有效的军事手段，Kumar R 等人针对 Avs(Aerial video surveillance)系统进行了深入的研究，并取得了一定的成果。Tsao 和 wen 针对军事战场上的敌人运动目标跟踪进行研究，以期在不同的战场环境中，对敌人运动目标进行准确快速的搜索和跟踪，稳定而又可靠的跟踪为军事决策提供了可靠的信息屏障。

在视觉跟踪领域的跟踪算法层出不穷，在跟踪效果上都能满足某些应用背景的要求，但缺乏通用性。目前目标跟踪算法主要有基于对比度分析、基于匹配、核方法、基于运动检测等，如图 1.1。

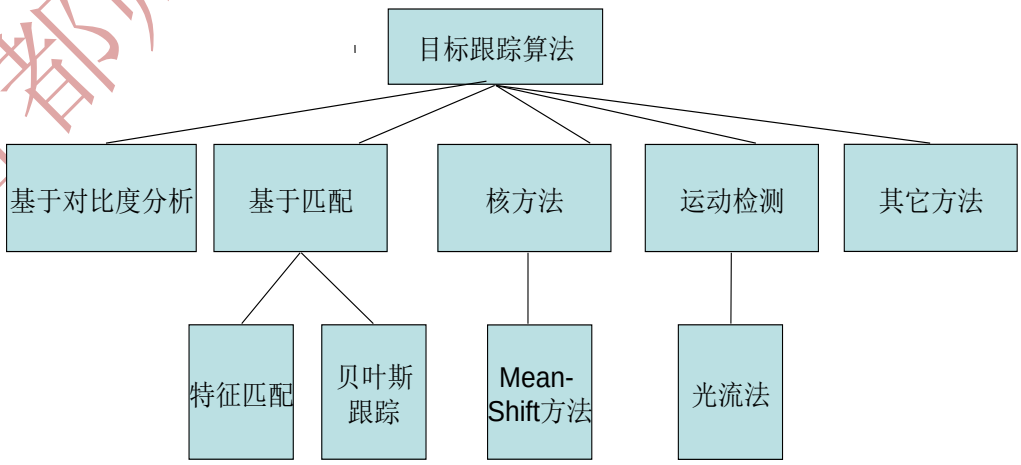


图 1.1 目标跟踪算法

首都师范大学信息工程学院

基于对比度分析的目标跟踪算法利用目标与背景在对比度上的差异来提取、识别和跟踪目标。主要有边缘跟踪、型心跟踪和质心跟踪三类。基于对比度分析的方法不适合复杂背景中的目标跟踪，但在空中背景下的目标跟踪中非常有效。基于匹配的目标跟踪算法需要提取目标的特征，并在每一帧中寻找该特征。寻找的过程就是特征匹配过程。目标跟踪中用到的特征主要有几何形状、子空间特征、外形轮廓和特征点等。其中，特征点是匹配算法中常用的特征。特征点的提取算法很多，如 Kanade Lucas Tomasi（KLT）算法、Harris 算法、SIFT 算法以及 SURF 算法等。特征点一般是稀疏的，携带的信息较少，可以通过集成前几帧的信息进行补偿。目标在运动过程中，其特征（如姿态、几何形状、灰度或颜色分布等）也随之变化。目标特征的变化具有随机性，这种随机变化可以采用统计数学的方法来描述。直方图是图像处理中天然的统计量，因此彩色和边缘方向直方图在跟踪算法中被广泛采用。基于运动检测的目标跟踪算法通过检测序列图像中目标和背景的不同运动来发现目标存在的区域，实现跟踪。光流算法是基于运动检测的目标跟踪的代表性算法。光流是空间运动物体在成像面上的像素运动的瞬时速度，光流矢量是图像平面坐标点上的灰度瞬时变化率。光流的计算利用图像序列中的像素灰度分布的时域变化和相关性来确定各自像素位置的运动，研究图像灰度在时间上的变化与景象中物体结构及其运动的关系。将二维速度场与灰度相联系，引入光流约束方程，得到光流计算的基本算法。光流法能够很好的用于二维运动估计，它也可以同时给出全局点的运动估计，但其本身还存在着一些问题：需要多次迭代，运算速度慢，不利于实时应用等。核方法对相似度概率函数或者后验概率密度函数采用直接的连续估计。其中最常用的是 Mean-Shift 算法。Mean-Shift 跟踪算法采用彩色直方图作为匹配特征。Mean-Shift 跟踪算法反复不断地把数据点朝向 Mean-Shift 矢量方向进行移动，最终收敛到某个概率密度函数的极值点。在 Mean-Shift 跟踪算法中，相似度函数用于刻画目标模板和候选区域所对应的两个核函数直方图的相似性，采用的是 Bhattacharyya 系数。因此，这种方法将跟踪问题转化为 Mean-Shift 模式匹配问题。核函数是 Mean-Shift 算法的核心，可以通过尺度空间差的局部最大化来选择核尺度，若采用高斯差分计算尺度空间差，则得到高斯差分 Mean-Shift 算法。Mean-Shift 算法假设特征直方图足够确定目标的位置，并且足够稳健，对其他运动不敏感。该方法可以避免目标形状、外观或运动的复杂建模，建立相似度的统计测量和连续优化之间的联系。但是，Mean-Shift 算法不能用于旋转和尺度运动的估计。为克服以上问题，人们提出了许多改进算法，如多核跟踪算法、多核协作跟踪算法和有效的最优核平移算法等。

首都师范大学信息工程学院

1.3 本文的研究内容以及目的

人机交互技术已经越来越成熟，发展的方向也日新月异。在人机交互过程中，目标跟踪技术起到了至关重要的作用。因此目标跟踪技术具有很高的研究价值。鼠标是我们与计算机进行交流的主要交流工具，是计算机输入设备的简称。鼠标的使用是为了使计算机的操作更加简便，来代替键盘那繁琐的指令。从原始鼠标、机械鼠标、光电鼠标、光机鼠标再到如今的光学鼠标，鼠标技术经历了漫漫征途终于修成正果。毫无疑问，光学鼠标是我们所追求的终极类型，诸多优点使它成为光机鼠标无可争议的接替者。而在光学鼠标发展的近几年中，我们亲眼目睹它的飞速进步，光学引擎的更新换代带来更高的精度、更快的速度以及更经得起推敲的性能。而鼠标相关的其他技术进展也不容小觑，纵横滚轮技术蔚为潮流，给我们带来更便捷的操作体验。蓝牙技术的引入让我们尽享无线操作的自由，皮革材料和丝绸表面处理工艺让鼠标成为艺术品的同时提供了绝佳的握感。然而人们仍然希望继续简化与计算机的交流方式，希望与计算机交流方式变得更加直接。因此人机交互技术变得更加盛行，成为研究热点。

智能鼠标的研究可以简化与计算机进行交流的方式，且更加直接。本课题研究的基于多视角视频采集与分析的智能鼠标算法研究，通过不同方向的多个摄像头采集手势信息，并分析手势被赋予的含义，来执行相应的操作，从而达到控制计算机的目的。首先进行目标检测，在视频采集区域找到进行操作的手，然后进行目标跟踪，在手移动的过程中，完成对手势的跟踪任务。并将手所在区域信息传给目标分析模块。分析手势的含义，并控制计算机执行相应操作。从而达到与计算机进行交流的目的。

本课题也存在一些限制因素值得注意。

1)只适合对单个手进行跟踪，不能对多个手进行同时跟踪。

2)由于光照等因素严重影响实验结果，因此在设计场景时应注意以下几个方面。如：影子抑制、低对比度的检测目标以及光照突然改变等情况。

3)大面积相似颜色信息干扰的问题。在目标跟踪过程中，当手势进入相似颜色信息干扰的复杂背景时，容易出现目标跟踪丢失。虽然本课题能够解决小面积肤色信息干扰问题，但当运动目标长时间进入复杂背景区域时，会出现跟踪丢失。

4)运动手势移动速度的问题。在手势跟踪过程中，进行操作的手的移动速度不能过快。虽然对手势进行重新检测能够重新定位到手的位置坐标，但运动速度过快会造成视频采集信息出现虚影，对后期特征提取和识别造成不必要的困难。

5)多视角采集过程中，手的动作幅度不能太小。例如：进行鼠标移动操作，手移动很

小的距离，很可能会出现识别失败，计算机无任何反应。因此在进行操控时，应使手的变化幅度尽量大。

本文的研究重点包含两方面，一方面是目标跟踪的设计以及获取目标位置信息；而另一方面也是重中之重进行算法的创新，解决传统跟踪算法因相似颜色信息干扰导致的跟踪丢失问题。

作者对课题的主要研究工作具体如下：

(1)对基于人手的目标跟踪与识别技术的现实可行性做了详细的分析。

(2)阅读了国内外大量有关目标跟踪、模式分析以及模式识等资料，对相关的学术论文和硕士论文进行了深入的研究和学习，为下面的实践工作做出充分的理论准备。

(3)具体分析了传统 Mean-Shift 跟踪算法的缺点，并根据 Mean-Shift 的缺点进行了部分改进，使其可以在相似颜色信息等复杂背景干扰下进行准确跟踪。

(4)本文提出了改进的 KSMS 算法（Kalman Skin Mean-Shift），即基于肤色信息约束的 Kalman 滤波和 Mean-Shift 融合的跟踪算法。利用 Kalman 滤波具有预测功能，对搜索窗口提前进行预测可以提高跟踪准确性。对大面积肤色信息干扰时具有良好的鲁棒性。

(5)通过实验验证了前期改进的 Mean-Shift 目标跟踪算法的有效性。

1.4 论文章节安排

本文由五个章节组成。

第 1 章首先论述了目标跟踪与识别技术研究的背景和意义，随后介绍了目标跟踪的发展现状，并简单介绍了本文的研究内容以及目的，并说明了前期的主要工作重点，最后提出了解决问题的具体策略。

第 2 章围绕目前主流的 Mean-Shift 跟踪算法进行了论述，在首节介绍了近期国内外的研究热点以及发展前景，并论文了目前众多方法的优点以及不足；其次在次小节介绍了 Mean-Shift 跟踪算法的工作原理以及需要的相关技术作为铺垫，对一些具体技术进行了详细的透彻的分析；然后介绍了本文所采用的算法以及框架的设计思路以及设计流程，并最终通过实验验证了该算法以及框架选择的优越性。

第 3 章提出了 KSMS 跟踪算法。首节论述 Kalman 滤波的简介，近期国内外的发展前景，以及在目标跟踪中引入 Kalman 滤波的原因。然后论述 Kalman 滤波的原理。将 Kalman 滤波与 Mean-Shift 算法融合来实现目标跟踪。最后介绍算法的框架设计思路以及设计流程，并最终通过实验验证该算法的优越性。

首都师范大学信息工程学院

第 4 章对全文进行总结，总结前期目标跟踪以及后期改进算法的优点，并最对目前系统框架的缺陷提出未来的研究的展望。

首都师范大学信息工程学院

第二章 Mean-Shift 算法

2.1 目标跟踪技术

目标跟踪技术在现代军事和民用中均占有非常重要的地位，提供可靠而准确的高质量目标信息始终是目标跟踪技术的主要任务。为此，国内外许多研究人员一直致力于该项目的研究。目前研究的目标跟踪主要分为三个步骤：对感兴趣移动目标的检测，从每一帧中跟踪这些目标，以及最后目标跟踪后的分析与识别[1]。

目标跟踪系统一般是基于对图像序列的处理，力图从复杂的背景中识别出目标，并对目标的运动规律加以预测，实现对目标的连续、准确的跟踪。系统的算法涉及图像数据的预处理、目标检测、图像分割、特征提取、运动分析和目标跟踪等。其中图像分割、特征提取和目标跟踪算法是目标跟踪系统的关键。

运动目标跟踪一般先进行目标检测，这是一个较简单的运动目标分割过程，提取出待跟踪的目标，将其表示为一定形式，如特征点、目标轮廓、与目标相关的几何形状等。目标检测后是目标跟踪的过程。目标跟踪的目的是提供一个目标随时间运动的轨迹，也可以提供目标在每一时刻所覆盖的区域，并建立不同帧之间目标实例的对应关系。

运动目标跟踪的方法根据目标描述的方式划分可分为三类：即点跟踪方法、基于核的跟踪方法和基于轮廓的跟踪方法(Object tracking survey)。

目前目标跟踪算法主要包含两个重要部分：1）目标的表示与定位；2）滤波与数据关联。其中，目标表示与定位是一个自底向上的运行过程，典型的应用为光流法、背景模板法、Mean-Shift 算法。而滤波与数据关联是一个自顶而下的过程，典型的应用是 Kalman filter、扩展 Kalman 滤波器、无极 Kamlam 滤波器以及粒子滤波器和它的变种等[2]。

虽然跟踪目标看似简单，但是它却是十分复杂的工作。这主要是由于一下原因造成：

(1)图像中存在大量的噪声。

(2)目标运动的复杂性。

(3)局部或者整个目标的遮挡。

(4)目标形状的复杂性。

(5)光照改变以及实时过程的需求。

首都师范大学信息工程学院

2.2 总体分布的非参数估计以及 Parzen 窗估计

该小节我们首先介绍总体分布的非参数估计技术，然后再介绍该技术的特例 Parzen 窗估计技术，通过其为下一小节介绍目标跟踪技术中的 Mean-Shift 即：均值漂移算法做铺垫。

2.2.1 总体分布的非参数估计

由于在大多数情况下，基于参数估计的方法需要知道样本的总体分布形式以及统计参数如：均值[1]、方差等，但这些条件在实际应用中却往往是无法准确获得的，因此，人们开始研究某些直接通过样本来估计总体分布的方法，即非参数估计[3]。

该方法的基本思想是，首先，我们给出概率密度的估计值 $\hat{p}(x)$ 如式(2.1)所示：

$$\hat{p}(x)=\frac{k}{NV} \tag{2.1}$$

显而易见，要想得到 $p(x)$ 的估计而不是 p(x)在区间内的平均就必须将体积 V 趋向于零。但是在把样本数目固定的情况下，令体积 V 趋于零则会出现区间不断缩小以至于无法包含任何样本的情况，即便是偶然有样本碰巧掉入区间内也会使 p(x)的估计趋近于无穷大而使统计变得毫无意义。

其实造成这种情况的原因是：样本数 N 总是有限的，所以体积 V 不允许任意的小。因此采用该估计，则 $\frac{k}{N}$ 和 $\hat{P}(x)$ 必然会存在一定的方差。

假设采用如下方法，即：为了估计 x 点的密度我们让第一个区间只采用一个样本来进行估计，而第二个区间则采用两个样本来进行估计,…，以此类推，则第 n 个区间采用 n 个样本，具体如(2.2)式所示。

$$\hat{p}_N(X)=\frac{k_N/N}{V_N}=\frac{k_N}{NV_N} \tag{2.2}$$

其中，式(2.2)应该必须满足三个条件，如式(2.3)、(2.4)和(2.5)所示：

$$\lim_{N\to\infty} V_N=0 \tag{2.3}$$

$$\lim_{N\to\infty} K_N=0 \tag{2.4}$$

$$\lim_{N\to\infty} \frac{K_N}{N}=0 \tag{2.5}$$

显而易见，其中，式(2.3)是为了使空间平均 $\frac{p}{v}$ 收敛于 p (x)；式(2.4)是为了当样本趋于无穷时，使获得的频率在概率意义上收敛于概率 P；式(2.5)表明尽管在小区域 R 内落入大量的样本，但相对于总样本来说却是小巫见大巫，少得可以忽略不计。

首都师范大学信息工程学院

非参数估计有两个主要应用：

(1) Parzen 窗体，通过固定 K_N，建立 N 与 V_N 之间的函数关系，使 V_N 不断减少，从而达到收敛，当然为了让 $\hat{P}_N(x)$ 收敛于 p (x)还需要对 k_N 以及 $\frac{K_N}{N}$ 加以限制。

(2) k_N 近邻估计，通过固定 V_N，建立 N 与 K_N 之间的函数关系，从而使包含 x 的 K_N 个近邻正好落入 V_N 中。

其中，我们将在下一小节中重点介绍 Parzen 窗估计，因为 Mean-Shift 的主要思想就是 Parzen 窗估计模型。

2.2.2 Parzen 窗估计

Parzen 窗估计[4]是总体分布中非参数估计的一个特例，它主要是将落入区间的样本数用一个函数进行描述如下：

$$K_N = \sum_{i=1}^{N} \varphi\left(\frac{x - x_i}{h_N}\right) \tag{2.6}$$

其中 $\varphi\left(\frac{x - x_i}{h_N}\right)$ 是一个服从其他分布的函数，比如：正态分布、均匀分布等。

此通过式(2.6)，我们可以转化式(2.2)如下：

$$\hat{P}_N(x) = \frac{1}{NV} \sum_{i=1}^{N} \varphi\left(\frac{x - x_i}{h_N}\right) \tag{2.7}$$

通过式(2.7)我们可以看出 p(x)的估计其实就是变量 x 和 x_i 函数的一种平均，实际上，Parzen 窗估计过程是一个内插过程，样本 x_i 距离 x 越近，对概率密度估计的贡献越大，越远贡献越小。当 h_N 较大时，x 和中心 x_i 距离大小的影响程度变弱，估计的 p(x)较为平滑，分辨率较差。当 h_N 较小时，x 和中心 x_i 距离大小的影响程度变强，估计的 p(x)较为尖锐，分辨率较好。

2.3 Mean-Shift 算法

目标跟踪是计算机视觉方向的一个重要研究领域，基于目标颜色信息和目标运动信息的跟踪算法是目前最常用的跟踪方法。基于颜色信息的跟踪算法中，Mean-Shift 算法是应用最广泛的算法。Mean-Shift 算法能够对具有特定颜色信息的运动目标进行跟踪，在背景简单的条件下，该算法计算量小，对局部遮挡、目标形变具有一定的适应性[5][6]。

Mean-Shift 算法是一种半自动跟踪算法。在跟踪的初始帧，它通过人工或其它检测算

首都师范大学信息工程学院

法找到并确定目标窗以及需要构建的目标模型；然后，在跟踪序列的第 N 帧所对应的位置计算候选目标模型；比较两个模型的相似度，并以相似度最大化为判断的原则来移动跟踪窗，从而定位目标的真实位置。

2.3.1 Mean-Shift 基础

Mean-Shift 这个概念最早是由 Fukunaga 等人于 1975 年在一篇关于概念密度梯度函数的估计中提出来的，其最初含义就是偏移的均值向量。Mean-Shift 指代一个向量，但随着 Mean-Shift 理论的发展，Mean-Shift 的含义也发生了变化。现今，Mean-Shift 算法指一个迭代的步骤，即先算出当前点的偏移均值，移动该点到其偏移均值，然后以此为新的起始点，继续移动，直到满足一定的条件结束[7]。

在很长的一段时间内 Mean-Shift 并没有引起人们的注意，直到 1995 年，另一篇关于 Mean-Shift 的重要文献发表。在这篇重要的文献中，Yizong Cheng 对基本的 Mean-Shift 算法在以下两个方面做了扩展。首先 Yizong Cheng 定义了一族核函数，使得随着样本与被偏移点的距离不同，其偏移量对均值偏移向量的贡献也不同，其次 Yizong Cheng 还设定了一个权重系数，使得不同的样本点重要性不一样，这大大扩大了 Mean-Shift 的试用范围。另外 Yizong Cheng 指出了 Mean-Shift 可能应用的领域，并给出了具体的例子。

Comaniciu 等人把 Mean-Shift 成功的运用在特征空间的分析，在图像平滑和图像分割中都得到了很好的应用。Comaniciu 等在文章中证明了 Mean-Shift 算法在满足一定条件下，一定可以收敛到最近的一个概率密度函数的稳态点，因此 Mean-Shift 算法可以用来检测概率密度函数中存在的模态。

Comaniciu 等人还把非刚体的跟踪问题近似为一个 Mean-Shift 最优化问题，使得跟踪可以实时的进行。Mean-Shift 算法计算量小，对刚性物体和具有特定颜色信息物体的跟踪准确性高，可实时性强。Mean-Shift 算法已经成为目标跟踪领域的常用算法。

2.3.2 Mean-Shift 算法

我们首先简要的介绍一下传统的 Mean-Shift 算法，传统的 Mean-Shift 算法主要分为四个模块，分别为目标模块、候选目标模块、分类度量与匹配模块、以及目标定位[8]。

给定 d 维空间 R^d 的 n 个样本点，i=1,2,…,n,在空间中任选一点 x，那么 Mean-Shift 向量的基本形式定义为：

$$M_h(x)=\frac{1}{k}\sum_{x_i\in S_h}(x_i-x) \tag{2.8}$$

首都师范大学信息工程学院

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：

$$S_h(x)=\{y:(y-x)^T(y-x)\le h^2\} \tag{2.9}$$

k 表示在这 n 个样本点 x_i 中，有 k 个点落入 S_h 区域中。

(x_i-x) 是样本点x_i相对于点x的偏移向量,(2-8)式定义的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。从直观上看,如果样本点x_i从一个概率密度函数$f(x)$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。

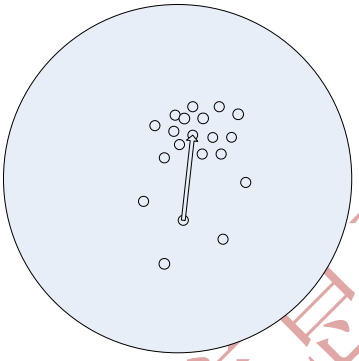


图2.1 Mean-Shift示意图

如上图所示，大圆圈区域表示S_h，小圆圈代表落入S_H区域内的样本点$x_i\in S_h$,箭头表示平均偏移向量$M_h(x)$,指向样本点分布最集中的区域，也就是概率密度函数的梯度方向。

从（2.9）式可以看出，只要是落入S_h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周围的统计特征越有效，因此我们引进核函数的概念，在计算$M_h(x)$时可以考虑距离的影响；同时我们可以认为在这所有的样本点x_i中，重要性不一样，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形式扩展为[9]：

$$M(x)=\frac{\sum_{i=1}^{n}G_H(x_i-x)w(x_i)(x_i-x)}{\sum_{i=1}^{n}G_H(x_i-x)w(x_i)} \tag{2.10}$$

其中：

$G_H(x_i-x)=|H|^{-\frac{1}{2}}G(H^{-\frac{1}{2}}(x_i-x))$，G(x)是一个单位核函数，H是一个正定的对称d*d矩阵，我们一般称之为带宽矩阵，$w(x_i)\ge 0$是一个赋给采样点x_i的权重。

实际应用中，带宽矩阵H一般被限定为一个对角矩阵 $H=diag[h_1^2,....,h_d^2]$，甚至更简单的被取为正比于单位矩阵，即$H=h^2I$。由于后一形式只需要确定一个系数h，在Mean-Shift

首都师范大学信息工程学院

中常常被采用，因此(2.10)式又可以被写为：

$$M_h(x)=\frac{\sum_{i=1}^{n}G(\frac{x_i-x}{h})w(x_i)(x_i-x)}{\sum_{i=1}^{n}G(\frac{x_i-x}{h})w(x_i)} \tag{2.11}$$

2.3.3 算法实现步骤

Mean-Shift 算法是基于颜色信息的目标跟踪算法，因此算法的首要任务就是计算运动目标的颜色信息。首先手动截取选定跟踪目标，得到目标矩阵 rect，计算出目标跟踪框内的加权直方图 hist。随后开始读取后续视频序列，先随机取一块与 rect 等大的矩阵，计算加权直方图 hist2，最后计算两者比重函数，如果超过设定的理想数值，则更新新的矩阵中心 Y，进行迭代，直到满足条件后停止。

Mean-Shift向量引入核函数的形式[10]：

$$M(x)\equiv\frac{\sum_{i=1}^{n}G_H(x_i-x)w(x_i)(x_i-x)}{\sum_{i=1}^{n}G_H(x_i-x)w(x_i)} \tag{2.12}$$

为了简化算法，在Mean-Shift中H经常正比于单位矩阵，即 $H=h^2I$,此时（2.12）式又可以写为：

$$M_h(x)\equiv\frac{\sum_{i=1}^{n}G\,(\frac{x_i-x}{h})w(x_i)(x_i-x)}{\sum_{i=1}^{n}G\,(\frac{x_i-x}{h})w(x_i)} \tag{2.13}$$

将（2）式的x提到求和号的外面来，可以得到：

$$M_h(x)=\frac{\sum_{i=1}^{n}G\,(\frac{x_i-x}{h})w(x_i)x_i}{\sum_{i=1}^{n}G\,(\frac{x_i-x}{h})w(x_i)}-x \tag{2.14}$$

令：

$$m_h(x)=\frac{\sum_{i=1}^{n}G\,(\frac{x_i-x}{h})w(x_i)x_i}{\sum_{i=1}^{n}G\,(\frac{x_i-x}{h})w(x_i)} \tag{2.15}$$

给定一个初始点x,核函数G(X),容许误差 ε，x为跟踪目标选择框的中心。

Mean-Shift算法循环执行下面步骤：

首都师范大学信息工程学院

(1)选择搜索窗口 ；

(2)计算 $m_h(x)$，即窗口（可能带权重）的重心；

(3)把 $m_h(x)$ 赋给x，即将窗口的中心移动到重心处；

(4)如果 $\|m_h(x)-x\|<\varepsilon$，结束循环；否则返回（2）步继续执行；

由（2.14）式可知， $m_h(x)=x+M_h(x)$,因此上面的操作就是不断的沿着概率密度的梯度方向移动，同时步长不仅与梯度的大小有关，也与该点的概率密度有关，在密度大的地方，更接近要找的概率密度的峰值，Mean-Shift算法使得移动的步长小一些，相反，在密度小的地方，移动的步长大一些。最终，Mean-Shift算法一定会收敛到该点附件的峰值，即窗口不能再移动。

2.4 位置修正算法

视频手势跟踪过程中，当手移动速度过快时，传统 Mean-Shift 算法会出现跟踪丢失问题。Mean-Shift 算法要求相邻两祯视频序列的运动目标有重叠区域，当运动速度过快时，相邻两祯图像序列中运动目标没有重叠区域，Mean-Shift 算法在周围区域搜索目标时无法找到跟踪目标，最终造成跟踪丢失问题。

由于手具有特定颜色信息，因此基于颜色信息的位置定位是可以实现的。在视频画面中，脸部和手具有相似颜色信息，因此只需分辨出手和脸部就可以实现手的位置重定位。因此要实现肤色检测。

在彩色空间中，皮肤颜色的分布较集中，与其他景物颜色间有较好的可区分性，可以用模型描述或对样本进行学习的方法加以判别。图像中皮肤颜色的差异主要由光照引起，在检测中只考虑色度信息，就可以减少光照的影响，使肤色的分布更趋集中。通过检测皮肤区域，可以缩小人脸的搜索范围，提高检测速度[11]。

根据计算机色彩理论，对每一种颜色在计算机中有不同的表达方式，这样就有了各种不同的颜色显示格式。这些格式只不过是颜色在计算机中不同的表示方法而已。在图像处理中最常用的模型是RGB、HSV、YUV等模型。由美国物理学家T.Yong和Helmholtz提出的著名的Yong-Helmholtz三色学说可以知道，光谱上的颜色能用红（Red）、绿（Green）、蓝（Blue）三种单色混合而成。研究的色彩图像是光谱图像的一种情况，所以用三基色来表征颜色是很自然的。多数的图像采集设备都是直接感知色彩的*R*、*G*、*B*三个分量，所以RGB格式在图像成像、显示、传输等工业上应用广泛。

首都师范大学信息工程学院

在RGB模型中(见图2.2)，光的色度取决于R、G、B 的比例关系：

$$r=\frac{R}{R+G+B}$$
$$g=\frac{G}{R+G+B} \tag{2.16}$$
$$b=\frac{B}{R+G+B}$$

其中r，g，b 成为色度坐标。因为$r+g+b=1$，所以只有两个色度坐标是相互独立的，也就是说是一个二维空间。

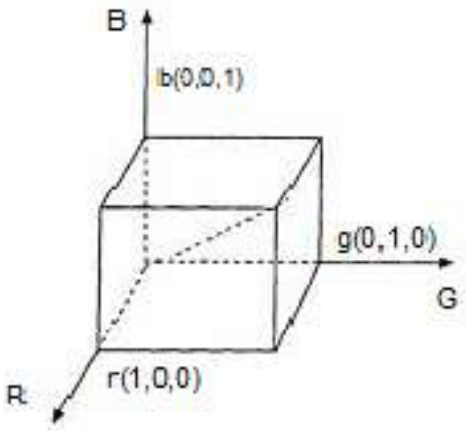


图2.2 RGB模型图

YUV是另外一种颜色模型，是以演播室质量标准的目标的CCIR601编码方案中采用的色彩表示格式。主要是用于多媒体计算技术和电视的色彩显示领域中。其基本特征是将亮度信号和色度信号分离表示。Y表示亮度；U、V是两个彩色分量，表示色差，一般是蓝和红色的相对值。由于人眼对亮度的变换比颜色的变化更敏感，因此YUV模型中Y 的值所占的带宽大于等于色彩分量所占的带宽。主要的YUV格式有4∶1∶1、4∶2∶2、4∶4∶4 三种[12]。

RGB颜色模型和YUV模型之间的转换公式:

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix}=\begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix}\begin{pmatrix} R \\ G \\ B \end{pmatrix} \tag{2.17}$$

HSV模型，这种色彩系统格式是由Munseu提出，常常为艺术家使用。这种格式反应了人类观察色彩的方法，它在色彩图像处理和计算机视觉的研究中常常被使用。这种格式是按照色度（hue）、色度饱和（saturation）和明亮值（value）建立起来的。有些也用光强

（Intensity）来表示HIS模型[13]。图4中可以看到：圆锥的顶面对应于V=1，它包含RGB模型中的R=1，G=1，B=1三个面，所代表的颜色较亮。在圆锥的顶点处，V=0，H和S无定义，代表黑色。圆锥的顶面中心处S=0，V=1，H无定义，代表白色。从该点到原点代表亮度渐暗的灰色，即具有同灰度的灰色。对于这些点，S=0，H的取值无定义。可以说，HSV模型中*V*轴对应于RGB颜色空间中的主对角线。在圆锥顶面的圆周上的颜色，V=1，S=1，这种颜色是纯色。

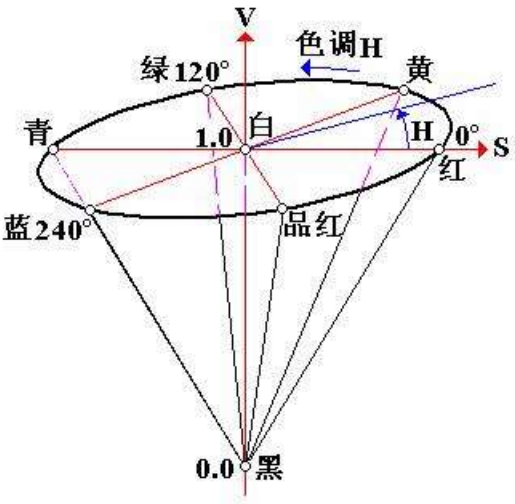


图2.3 HSV模型图

彩色图像常用R,G,B三分量的值来表示。但R,G,B三分量之间常有很高的相关性，直接利用这些分量常常不能得到所需的效果，所以需将RGB图像变换到其它的彩色特征空间中去，如YCbCr彩色空间。

利用 HSV 空间的 H（色调）分量进行肤色区域的检测[14]。与 HSI 中空间相比，从 RGB 空间到 HSV 空间的变换更容易计算。

在一些图像处理中如强亮度值处理时RGB模型将不再实用，即对于很多图像视频处理时，若背景亮度发生变化在目标跟踪时就会出现丢失目标现象。根据实验测试结果分析，

RGB 颜色模型和 HSV 模型之间的转换公式[15]：

$$\begin{cases} H = \arccos\left\{ \dfrac{[(R-G)+(R-B)]/2}{[(R-G)^2+(R-B)(G-B)]^{1/2}} \right\} \\ S = \dfrac{\max(R,G,B)-\min(R,G,B)}{\max(R+G+B)} \\ V = \dfrac{\max(R,G,B)}{255} \end{cases} \tag{2.18}$$

.

变换后的HSV空间是一个圆锥体，圆锥底面半径与0^0线之间的夹角代表色调(Hue)，圆锥底面半径大小代表饱和度（Saturation），圆锥的高代表亮度（Value）。

图2.3中可以看到：圆锥的顶面对应于V=1，它包含RGB模型中的R=1，G=1，B=1 三个面，所代表的颜色较亮。在圆锥的顶点处，V=0，H和S无定义，代表黑色。圆锥的顶面中心处S=0，V=1，H无定义，代表白色。从该点到原点代表亮度渐暗的灰色，即具有不同灰度的灰色。对于这些点，S=0，H的取值无定义。可以说，HSV模型中*V*轴对应于RGB颜色空间中的主对角线。在圆锥顶面的圆周上的颜色，V=1，S=1，这种颜色是纯色。

当色调（Hue）从0变化到1.0时，相应的颜色从红色变化到黄色、绿色、蓝绿色、蓝色、紫红色，然后回到红色，实际上0和1.0都代表红色值。当饱和度（Saturation）从0变化到1.0时，相应的颜色或者说色调（Hue）从不饱和（灰色阴影）变到全饱和（没有白色成分）。亮度从0变到1.0，相应的颜色变得越来越亮。因此，在圆锥的顶点，亮度（Value）值为0，所有的颜色均为黑色，饱和度S为0，此时的H值没有意义。

HSV空间是RGB空间的非线性变换，它将原本相关性很强的R、G、B值转换为相关性较弱的H、S、V值，H和S分量与人感受色彩的方式相一致。在HSV空间中彩色图像的每一个均匀性色彩区域都对应一个相对一致的色调（Hue），使得色调（Hue）能够单独用来进行彩色区域的分割。HSV变换比较容易计算，并且是可逆的，同时，HSV空间也能最好地满足颜色空间的均匀性、紧致性、完整性和自然性等属性[16]。

通常情况下，彩色图像的V值不会为0而使H无意义，也不会小于某个阈值，使得H无法分辨，并且在肤色检测中使用肤色条件是为了进行粗检测，使得肤色区域能够被迅速提取出来，肤色模型越简单有效越好，因此该文只使用H值来迅速检测肤色区域。

对于色彩信息的利用中，HSV格式的优点在于它将明度*V*与反映色彩本质特性的两个参数色度H和饱和度S一一分开，三个分量独立不相关。而RGB中根据公式（2.16）可知三个分量存在关系。采用HSV模型，将光照明暗给物体颜色带来直接影响，考虑用反映色彩本质特性的色度分量来进行聚类分析，由于HSV三个分量是独立的，可以考虑计算量比RGB模型的三个分量同时考虑要小。这也正是HSV格式在彩色图像处理和计算机视觉的研究中经常被使用的原因。

利用肤色信息修正跟踪目标的位置，首先要提取出肤色区域。将 RGB 颜色模型转换成 HSV 模型，通过设置合适的阈值实现肤色提取。一般情况下，连通域面积最大的为头部，面积小的为手。因此在手和脸部未形成连通域时，可以将面积较小的区域视为被跟踪

首都师范大学信息工程学院

目标，解决了因手移动速度过快所导致的跟踪丢失问题[17]。



图 2.4 肤色提取

从图中 2.4 可以看出，经过肤色检测后，成功将脸部和手从图像中分割出来。此时图像中仅剩下脸部和手区域，只需再确定手和脸部即可实现对手的重新定位。由图可知，脸部区域面积大于手区域面积，因此一般情况下，只需比较两个区域面积大小即可区分出脸部和手。图 5 第二帧图像可知，脸部连通域有未连通现象，区域内有细小空洞，因此对第二帧图像进行再次处理，使得两个区域没有未连续现象。

实验中需要对二值图像进行减噪处理，图像形态学中的腐蚀和膨胀能很好的解决此问题。如果在腐蚀和膨胀操作前，对灰度图像做一次滤波，减噪效果将更明显。腐蚀的具体操作是：用一个结构元素(一般是 3×3 的大小)扫描图像中的每一个像素，用结构元素中的每一个像素与其覆盖的像素做“与”操作，如果都为 1，则该像素为 1，否则为 0。膨胀的具体操作是：用一个结构元素(一般是 3×3 的大小)扫描图像中的每一个像素，用结构元素中的每一个像素与其覆盖的像素做“与”操作，如果都为 0，则 0 该像素为 0，否则为 1。腐蚀的作用是消除物体边界点，使目标缩小，可以消除小于结构元素的噪声点；膨胀的作用是将与物体接触的所有背景点合并到物体中，使目标增大，可添补目标中的空洞。开运算是先腐蚀后膨胀的过程，可以消除图像上细小的噪声，并平滑物体边界。闭运算时先膨胀后腐蚀的过程，可以填充物体内细小的空洞，并平滑物体边界。

膨胀和腐蚀

因此对二值图像进行闭运算，先膨胀后腐蚀，将脸部和手区域内的细小空洞填补，以使整个画面为两个完整的连通域，从而实现手的定位。处理结果如图：

首都师范大学信息工程学院

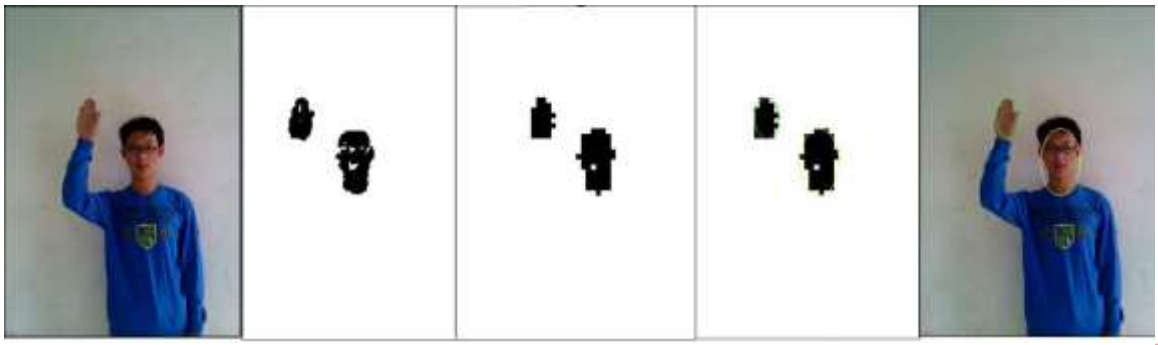


图2.5 基于肤色信息的手势定位

图2.5第一帧为原始图片，图5第二帧为经过二值化后的图片，图2.5第三帧是在图2.5第二帧图像的基础上，经过边缘和块处理函数处理后得到的二值图像。处理过程中，可以按需求来设置进行块操作的块大小。图2.5第四帧确定了两个目标，并分别用黄色和绿色的椭圆将连通域圈定。图2.5第五帧为找到手和脸部后，圈定肤色区域的图片。对两个连通域的面积进行排序，通常情况下最大的区域为脸部，其次是手。对面积小的连通域进行处理，求出连通域重心，即得到了跟踪目标的重心。

2.5 Mean-Shift 算法跟踪结果

在视频手势跟踪过程中，对跟踪结果有影响的因素主要有两个：光照条件和跟踪背景。当光照条件很理想时，背景的好坏对跟踪结果会产生巨大影响。Mean-Shift 跟踪算法是基于颜色信息的跟踪算法，当背景和肤色相近时，Mean-Shift 算法会因为无法识别跟踪目标导致跟踪丢失。本文对视频样本进行处理，处理结果如下图。



图 2.6 Mean-Shift 跟踪结果

图 2.6 序列图片是利用 Mean-Shift 算法对视频序列进行处理生成的跟踪结果图像，350 帧图像我们可以看出，整个跟踪背景里，只有脸部和手的颜色相近，因此脸部区域属于复杂背景区域，Mean-Shift 跟踪算法可能会跟踪丢失。350 帧前，手和脸部未形成连通域，此时跟踪准确。视频序列的 410 帧开始手和脸部形成了连通域，从 452 帧图像的跟踪结果可以看出，跟踪框始终停留在脸部，手的位置坐标停止更新。Mean-Shift 跟踪算法由于当脸部和手形成连通域后无法确定手是否在移动，导致跟踪框无法继续更新下去，从而造成

跟踪丢失。478 帧图像序列手开始和脸部分开，但因为传统 Mean-Shift 算法已经跟踪丢失，算法本身不能重新定位手的位置坐标，因此跟踪框将仍然在脸部区域晃动。不能完成跟踪任务，直到跟踪视频结束。Mean-Shift 跟踪算法在脸部和手形成连通域后，无法合理处理复杂背景下的跟踪问题，最终导致目标跟踪丢失。

Mean-Shift 跟踪算法是基于颜色信息的跟踪算法，当出现相似颜色信息的复杂背景时，Mean-Shift 跟踪算法不能很好的解决跟踪丢失问题。因此引入约束条件是必要的。

首都师范大学信息工程学院

第三章 KSMS 跟踪结果

在复杂背景下，Mean-Shift 算法的跟踪效果会减弱。例如手势跟踪过程中，手经过和肤色相近的背景时（例如脸部），会出现跟踪不准确，甚至跟踪目标丢失，一旦目标跟踪失败，后面的一系列操作就都无法完成了。Mean-Shift 算法只是利用颜色信息进行跟踪，当出现和目标颜色相近的背景时，系统无法确认哪一个是运动目标，引入约束条件是必要的。可以根据运动目标的运动信息（方向、速度等），估计运动目标下一时刻的最有可能位置。Kalman 滤波适用于解决随机线性离散系统的状态或参数估计问题，是一种实时递推算法[18]。用于估计参数的下一时刻的状态[19]。本文提出 Kalman 滤波与 Mean-Shift 算法相结合的跟踪算法，利用 Kalman 滤波预测手势下一时刻位置，并不断纠正跟踪目标位置信息，帮助 Mean-Shift 算法进行实时跟踪。实验结果表明此算法可以解决背景复杂情况下的丢失跟踪目标的问题。

3.1 卡尔曼滤波

卡尔曼匈牙利数学家， 1960 年, R.E. Kalman 出版了他的著名论文，描述了一个离散线性系统滤波问题的递归的解决方案，从那以后，由于计算机科学的发展，Kalman 滤波器被广泛的研究和应用于自治系统或导航系统。Kalman 滤波器是一组数学等式，它提供了一个有效的计算（回归）方法，来估计过程的状态，在这种方法中，将均方差最小化。该滤波器功能非常强大，它支持过去、现在、甚至将来状态的估计，并当模型系统的精确特性未知的情况下进行估计。

Kalman滤波器是根据最小均方误差准则建立起来的估计方法。Kalman滤波器是一个高效的递归滤波器，他可以实现从一系列的噪声测量中，估计动态系统的状态。它用线性递推的方法，对多个测量数据和多个信号参数进行处理，给出无偏的最小均方误差的估计。它只用信号的前一个估计值和最近一个观察值就可以估计信号的当前值，而不必像维纳滤波器那样需要全部过去的观察值。它根据受噪声干扰的观测输入对未知系统状态和参数进行估计，观测噪声和扰动输入都看作随机过程，Kalman滤波一般都假设它们是高斯白噪声过程，它们的概率密度函数是均值为零的高斯函数。Kalman滤波器是一个纯粹的时域滤波器，而不像低通滤波器等频域滤波器那样，需要在频域中设计，然后转换到时域中应用。

首都师范大学信息工程学院

Kalman滤波技术广泛应用于包含雷达、计算机视觉在内的工程应用领域。在控制理论和控制系统工程中也是一个非常重要的课题。

Kalman 滤波是一种实时递推算法，它所处理的是随机信号，利用系统噪声和观测噪声的统计特性，以系统的观测量作为滤波器的输入，以所要估计值（状态或参数）作为滤波器的输出，滤波器输入与输出是由时间更新和观测更新算法联系在一起的，根据系统方程和观测方程估计出所需要处理的信号——实质是一种最优估计方法。Kalman 滤波具有预测功能，常用于目标跟踪中。卡尔曼滤波器利用目标的动态信息，去掉噪声影响，获取目标此刻好的位置估计(滤波)，将来位置估计(预测)，也可以是过去位置估计的(插值或平滑)。

目前,卡尔曼滤波已经有很多不同的实现，有施密特扩展滤波器、信息滤波器以及一系列的 Bierman 和 Thornton 发明的平方根滤波器等，而卡尔曼最初提出的形式现在称为简单卡尔曼滤波器。最常见的卡尔曼滤波器应用是锁相环，它在收音机、计算机和几乎全部视频或通讯设备中广泛存在。

3.2 Kalman 滤波算法

手势跟踪时，只要手和脸部未形成连通域时，即使存在噪声干扰，传统Mean-Shift算法也能很好的对手进行实时跟踪。但当手和脸部形成连通域时，传统Mean-Shift算法无法识别并确认手的位置信息，从而造成跟偏甚至跟踪丢失，因此需要引入估计器预测运动目标的运动参数。Kalman滤波具有预测功能，将Kalman滤波与传统Mean-Shift算法结合可以增强抗干扰能力。Kalman滤波包括两个阶段：预测和更新。Kalman滤波器只要获知上一时刻的状态估计和当前状态的观测就可以计算出当前状态的估计，具有计算量小、可实时计算的特点[20]。

3.2.1 建立系统数学模型

在实际应用中，我们可以将物理系统的运行过程看作是一个状态转换过程，卡尔曼滤波将状态空间理论引入到对物理系统的数学建模过程中来，其假设系统状态可以用n 维空间的一个向量$X \in R^n$ 来表示。为了描述方便，我们作以下假设：①物理系统的状态转换过程可以描述为一个离散时间的随机过程；②系统状态受控制输入的影响；③系统状态及观测过程都不可避免受噪声影响；④对系统状态是非直接可观测的。

在以上假设前提下，定义系统状态变量为$X_k \in Rn$ ，系统控制输入为Uk ，系统过程激励噪声为W_k ，可得出系统的状态随机差分方程为：

首都师范大学信息工程学院

$$X_k = AX_{k-1} + BU_k + W_k \tag{3.1}$$

定义观测变量$Z_k \in R^m$，观测噪声为V_k，得到量测方程：

$$Z_k = HX_k + V_k \tag{3.2}$$

假设W_k，V_k为相互独立，正态分布的白色噪声，过程激励噪声协方差矩阵为Q，观测噪声协方差矩阵为R，即：

$$W_k \sim N(0, Q) \tag{3.3}$$

$$V_k \sim N(0, R) \tag{3.4}$$

A，B，H我们统称为状态变换矩阵， 是状态变换过程中的调整系数，是从建立的系统数学模型中导出来的，这儿我们假设它们是常数。

3.2.2 滤波器计算原型

从建立的系统数学模型出发，可以导出卡尔曼滤波的计算原型，包括：时间更新方程和测量更新方程。为了便于描述，做以下说明：

（1）$\hat{X}_k \in R^n$，第k步之前的状态已知的情况下第k步的先验状态估计值（－代表先验，^代表估计）；

（2）$\hat{X}_k \in R^n$，测量变量Z_k已知情况下第k步的后验状态估计值。由此定义先验估计误差和后验估计误差：

$$E_k^- = X_k - \hat{X}_{k^-} \tag{3.5}$$

$$E_k = X_k - \hat{X}_k \tag{3.6}$$

先验估计误差的协方差矩阵为：

$$P_k^- = E(E_k^- E_k^{-T}) \tag{3.7}$$

后验估计误差的协方差矩阵为：

$$P_k = E(E_k E_k^T) \tag{3-8}$$

式（3.8）构造了卡尔曼滤波器的表达式[21]：先验估计$\hat{X}_k^-$和加权的测量变量Z_k及其预测$H\hat{X}_k^-$之差的线性组合构成了后验状态估计$\hat{X}_k$：

$$\hat{X}_k = \hat{X}_k^- + K(Z_k - \hat{X}_k^-) \tag{3.9}$$

式中测量变量及其预测值之差$(Z_k - \hat{X}_k^-)$反映了预测值和实际值之间的不一致程度，称为测量过程的残余。n×m 阶矩阵K 叫做残余的增益。将P_k对K求导，使一阶导数为零，可以求出K_K 的一种形式为：

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \tag{3.10}$$

对卡尔曼增益K的确定是建立滤波模型的关键步骤之一，它能显著影响模型的效率[22]。

首都师范大学信息工程学院

卡尔曼滤波器包括两个主要过程：预估与校正。预估过程主要是利用时间更新方程建立对当前状态的先验估计，及时向前推算当前状态变量和误差协方差估计的值，以便为下一个时间状态构造先验估计值；校正过程负责反馈，利用测量更新方程在预估过程的先验估计值及当前测量变量的基础上建立起对当前状态的改进的后验估计。这样的一个过程，我们称之为预估－校正过程，对应的这种估计算法称为预估－校正算法。

3.2.3 Kalman 滤波预测阶段

Kalman滤波器具有递推算法结构，进行递推计算时，只保留上一次的计算结果即可。在估计阶段，滤波器利用上一次的计算结果和系统前一时刻方差预测系统当前时刻状态的预测值和系统当前状态的预测方差[23]。

状态预测：

$$\hat{X}_{k|k-1} = A\hat{X}_{k-1|k-1} + BU_{k-1} \tag{3.11}$$

式（3.11）中，$\hat{X}_{k|k-1}$是利用上一状态预测的结果，$\hat{X}_{k-1|k-1}$是上一状态最优的结果。U_{k-1}为现在状态的控制量，可以为0，A是作用在前一状态的状态转移矩阵，B为作用在控制向量上的控制输入输出矩阵。

估计协方差阵预测：

$$P_{k|k-1} = AP_{k-1|k-1}A^T + Q \tag{3.12}$$

式（3.12）中，$P_{k|k-1}$是$\hat{X}_{k|k-1}$对应的协方差，$P_{k-1|k-1}$是$\hat{X}_{k-1|k-1}$对应的协方差，Q是系统过程的协方差。

3.2.4 Kalman 滤波的更新

Kalman 滤波对目标当前状态进行预测后，利用在当前状态的观测值优化预测阶段的预测值，以获得一个更准确的当前状态的估计[24]。

状态估计更新：

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k[Z_k - H_k\hat{X}_{k|k-1}] \tag{3.13}$$

式（3.13）中K_k为Kalman增益矩阵，Z_k是系统的观测序列，H_k是观测矩阵。

滤波增益矩阵：

$$K_k = P_{k|k-1}H_k^T[H_kP_{k|k-1}H_k^T + P_k]^{-1} \tag{3.14}$$

式（3.13）中我们已经得到了K状态下最优的估算值$\hat{X}_{k|k}$，但为了令Kalman滤波器不断的运行下去直到系统过程结束，我们还要更新K状态下$\hat{X}_{k|k}$的估计误差方差阵$P_{k|k}$：

$$P_{k|k} = [I - K_kH_k]P_{k|k-1}[I - K_kH_k]^T + K_kR_kK_k^T \tag{3.15}$$

式（3.15）中I为1的矩阵，当系统进入k+1状态时，$P_{k|k}$就是式子（3.12）的$P_{k|k-1}$。这

首都师范大学信息工程学院

样算法就可以自回归的运算下去。

3.2.5 Kalman 滤波参数设置

跟踪过程中，相邻两祯图像中的运动目标位移较小，可近似认为跟踪目标在两祯图像的时间间隔内做匀速运动，所以采用等速度运动模型[25]，利用 Kalman 滤波器预测手心位置。

状态预测矩阵 $\hat{X}_{k-1|k-1}$ =（x_{k-1}，$y_{k-1,}Vx_{k-1},Vy_{k-1}$）^{T},观测序列向量 Z_k =（x_{k-1}，y_{k-1}）^{T},其中 x_{k-1}，$y_{k-1,}Vx_{k-1},Vy_{k-1}$分别为运动目标在图像中的x坐标，y坐标，水平方向速度和垂直方向速度[26]。根据方程（3-11）得状态转移矩阵A和观测矩阵H为：

$$A=\begin{bmatrix}1 & 0 & dt & 0\\ 0 & 1 & 0 & dt\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{3.16}$$

$$H=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\end{bmatrix} \tag{3.17}$$

Kalman滤波器首先初始化状态噪声协方差矩阵Q、观测噪声协方差矩阵R、初始状态预测矩阵X_0和初始误差协方差矩阵P_0。用状态预测矩阵中的x_k,y_k分量设定为Mean-Shift算法搜索窗口的中心位置，并把Mean-Shift算法得出的运动目标质心位置坐标作为观测值Z_k校正预测状态向量，得到 $\hat{X}_{k|k}$，实验中，

式中dt为时刻t_k与时刻t_{k-1}的时间间隔。观测噪声协方差矩阵R：

$$R=\begin{bmatrix}0.2845 & 0.0045\\ 0.0045 & 0.0455\end{bmatrix} \tag{3.18}$$

状态噪声协方差矩阵Q=0.01*eye(4),误差协方差矩阵P=100*eye(4),初始状态矩阵X_0中x_0，y_0分量分别取初始搜索窗口的中心位置坐标的x,y分量。Vx_0,Vy_0分量均取零。

3.3 融合的 KSMS 跟踪算法

算法流程如图8所示，每一祯视频图像都先转化为HSV图像，通过H分量的颜色直方图反投影建立肤色信息概率分布图像，然后用Mean-Shift算法对所选跟踪目标进行跟踪，算

首都师范大学信息工程学院

出运动目标的质心，质心的位置坐标作为下一祯图像跟踪目标的起始位置不断迭代更新下去。当手和脸部区域未形成连通域时，若出现跟踪丢失，则利用肤色信息重定位来重新计算手的位置，将结果作为下一祯图像搜索窗口的中心位置。当手和脸部区域形成连通域时，则将Mean-Shift算法算出的运动目标的质心坐标传给Kalman滤波器，作为Kalman滤波的观测序列[27]，利用Kalman滤波的预测功能，校正更新手心位置信息。Kalman滤波器的预测值作为下一帧图像搜索窗口的中心位置。

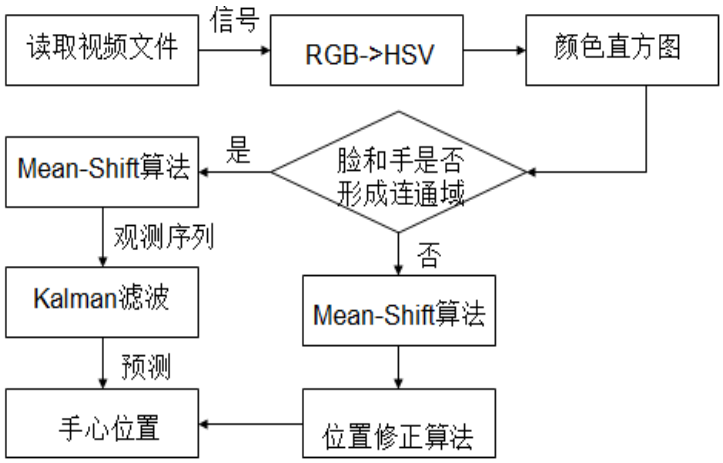


图 3.1 融合算法流程图

肤色信息修正的方法只能解决手和脸部未形成连通域时的跟踪丢失问题，一旦手和脸部形成连通域，算法将无法判断手心的位置坐标，误差不断的积累下去，最终导致目标跟踪丢失。因此引入Kalman滤波解决肤色信息干扰情况下的跟踪问题。

融合算法实现步骤：

(1)利用Mean-Shift算法跟踪运动目标，得到运动目标的中心坐标Y；

(2)利用肤色信息判断手和脸部是否已经形成连通域；

(3)如果已经形成连通域，将Y作为Kalman滤波的观测序列Z_k，作为输入参数，求出Kalman滤波对跟踪目标的预测位置Y_C。Y_C作为跟踪目标的中心，执行步骤(6)；

(4)若未形成连通域，找到手的重心G；

(5)比较G和Y，若相差较大，即为Mean-Shift算法跟踪目标丢失的情况，此时将G作为跟踪目标的中心；

(6)将跟踪目标的中心坐标Y_C或G传给Mean-Shift算法继续执行下去，直到满足一定的条件；

Mean-Shift算法融合Kalman滤波，能够解决运动目标经过复杂背景时跟踪失败的问题。在手和脸部未形成连通域时，基于肤色的位置修正能够解决运动目标速度过快时跟踪失败

的问题。

3.4 KSMS 跟踪结果

传统Mean-Shift跟踪算法在处理相似颜色信息干扰的复杂背景下的跟踪问题时，经常出现跟踪丢失现象，KSMS融合算法利用Kalman滤波的预测功能辅助Mean-Shift跟踪算法对手进行跟踪。我们对跟踪样本进行处理，跟踪结果如下图。



图3.2 KSMS跟踪结果

图3.2序列图片是利用Kalman滤波和Mean-Shift算法相结合的算法对视频序列进行处理得到的跟踪结果图像。视频序列的350帧图像显示跟踪准确，此时手和脸部未形成连通域。视频序列的410帧之后手和脸部形成了连通区域，432帧图像和452帧图像可以看出，KSMS融合新算法准确预测到了运动目标的位置，并通过实时更新运动目标的位置，来完成精确跟踪。系统判断手和脸部是否形成了连通域，但形成了连通域时，传统Mean-Shift算法将前几帧图像得到的手的位置坐标作为Kalman滤波的观测预测，充当输入参数，Kalman滤波利用观测序列对手的运动方向及速度进行预测，从而辅助Mean-Shift算法在相似颜色信息干扰的情况下进行继续跟踪。实验结果表明跟踪及时且准确。478帧图像开始，手和脸部脱离了连通区域，此时KSMS融合算法仍然准确跟踪到了手。486帧图像到视频序列结束，融合算法都成功跟踪到了手。跟踪结果理想，解决了相似颜色信息干扰情况下的跟踪丢失问题。

首都师范大学信息工程学院

第四章 实验平台与结果评估

4.1 实验平台介绍

MATLAB（矩阵实验室）是一款由美国 The MathWorks 公司出品的商业数学软件。MATLAB 是一种用于算法开发、数据可视化、数据分析以及数值计算的高级技术计算语言和交互式环境。除了矩阵运算、绘制函数/数据图像等常用功能外，MATLAB 还可以用来创建用户界面及与调用其它语言（包括 C，C++和 FORTRAN）编写的程序。

尽管 MATLAB 主要用于数值运算，但利用为数众多的附加工具箱（Toolbox）它也适合不同领域的应用，例如控制系统设计与分析、图像处理、信号处理与通讯、金融建模和分析等。另外还有一个配套软件包 Simulink，提供了一个可视化开发环境，常用于系统模拟、动态/嵌入式系统开发等方面。

MATLAB 和 Mathematica、Maple 并称为三大数学软件。它在数学类科技应用软件中在数值计算方面首屈一指。MATLAB 可以进行矩阵运算、绘制函数和数据、实现算法、创建用户界面、连接其他编程语言的程序等，主要应用于工程计算、控制设计、信号处理与通讯、图像处理、信号检测、金融建模设计与分析等领域。

MATLAB 的基本数据单位是矩阵，它的指令表达式与数学、工程中常用的形式十分相似，故用 MATLAB 来解算问题要比用 C，FORTRAN 等语言完成相同的事情简捷得多，并且 MATLAB 也吸收了像 Maple 等软件的优点，使 MATLAB 成为一个强大的数学软件。在新的版本中也加入了对 C，FORTRAN，C++，JAVA 的支持。可以直接调用,用户也可以将自己编写的实用程序导入到 MATLAB 函数库中方便自己以后调用，此外许多的 MATLAB 爱好者都编写了一些经典的程序，用户可以直接进行下载就可以用。

MATLAB 平台具有如下特点：

1) 高效的数值计算及符号计算功能，能使用户从繁杂的数学运算分析中解脱出来；

2) 具有完备的图形处理功能,实现计算结果和编程的可视化;

3) 友好的用户界面及接近数学表达式的自然化语言，使学者易于学习和掌握；

4) 功能丰富的应用工具箱(如信号处理工具箱、通信工具箱等) ,为用户提供了大量方便实用的处理工具。

4.2 跟踪结果对比

Mean-Shift 跟踪算法是基于颜色信息的跟踪算法，因此背景颜色、光照条件都会影响到跟踪效果。本实验采集的视频材料是在实验室的白色墙壁侧录制的。录制当天实验室光线理想，没有虚影，背景颜色对肤色影响很小，因此我们的跟踪的主要任务是解决当手和脸部形成连通域时的跟踪丢失问题。

本文基于 windows xp 操作系统，利用 MATLAB 2011a 软件平台，对 genzong.avi 视频文件进行试验研究，视频序列每帧大小为 480*640，视频序列长度为 498 帧。实验主要研究解决手和脸部形成连通域时，传统 Mean-Shift 算法出现跟踪丢失的情况。视频序列主要分为三个部分。第一部分为手从头部附近向脸部靠拢，此时未形成连通域；第二部分为手从头部上方向下移动，和脸部形成连通域，并继续向下移动，直到离开脸部，形成非连通域。第三部分为手离开脸部后，在胸前横向移动。此时脸部和手未形成连通域。

我们对相同的视频序列用不同的跟踪算法进行跟踪处理，首先利用传统 Mean-Shift 算法对视频序列进行跟踪，随后用 KSMS(Kalman Skin Mean-Shift)跟踪算法对视频序列进行跟踪。跟踪结果如下图：



(a)Mean-Shift跟踪结果



(b)KFMS算法跟踪结果

图4.1 Mean-Shift算法和融合算法的跟踪效果

通过图4.1对比可以发现，传统Mean-Shift算法在视频手势跟踪中，无法解决大范围肤色因素干扰情况下的跟踪问题，而本文提出的KSMS融合算法成功克服了脸和手重叠时，无法准确跟踪的问题。本文对视频中手势进行跟踪的过程中，记录了手心的位置坐标变化，

跟踪轨迹如图3.1所示。

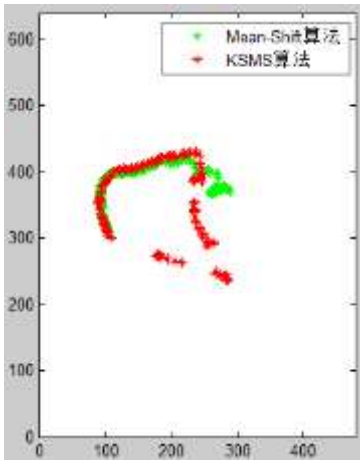


图4.2跟踪轨迹图

图4.2中绿色轨迹为利用Mean-Shift算法得到的跟踪结果，从图中可以看出，在跟踪序列的起始部分，Mean-Shift算法和KSMS融合算法的跟踪结果接近，都准确跟踪到了手的位置。当手和脸部形成连通域时，Mean-Shift算法开始出现波动，无法确定手和脸部的信息，跟踪框的中心坐标左右摆动，无法继续进行跟踪，手心的位置坐标停止更新。红色轨迹为利用KSMS算法得到的跟踪结果，从图中可以看出，当手和脸部形成连通域时，KSMS算法仍然能够对手进行准确跟踪，并实时更新手心的位置坐标。在简单背景下，Mean-Shift跟踪算法满足一般的跟踪任务，并且实时性好，但当运动目标进入复杂背景时，传统Mean-Shift算法开始力不足心，出现跟踪丢失问题。本文提出的KSMS(Kalman Skin Mean-Shift)融合算法成功解决了相似颜色信息干扰情况下跟踪丢失的问题。

4.3 结论与讨论

本文在研究了传统 Mean-Shift 跟踪算法的基础上，根据 Mean-Shift 算法的不足，提了改进的 KSMS(Kalman Skin Mean-Shift)跟踪算法。通过实验分析，可知 KSMS 融合算法相较传统 Mean-Shift 算法有了很有意义的提高。在简单背景下，Mean-Shift 算法和 KSMS 融合算法都能够很好的对运动目标进行准确跟踪；但当运动目标进入复杂背景时，例如相似颜色信息干扰的背景下，传统 Mean-Shift 算法经常出现目标跟踪丢失问题。本文提出的 KSMS 融合算法，利用 Kalman 滤波的预测功能，结合传统 Mean-Shift 算法实现了在复杂背景下的跟踪，实验结果表明，KSMS 融合算法跟踪准确，具有实用性。

该算法的不足之处在于 Kalman 滤波的复杂性，使得整个融合算法的执行效率降低。KSMS 融合算法虽然实现了很高的跟踪，但当运动目标进入相似颜色信息干扰的复杂背景

首都师范大学信息工程学院

时，会使得跟踪效率降低。近期，有文章指出快速 Kalman 滤波同样具有预测功能，算法复杂度降低，执行效率提高。如能将 Mean-Shift 算法和快速 Kalman 滤波相结合，应能提高整个算法的执行效率。令KSMS融合算法能够解决一些相似颜色信息干扰下的跟踪问题，但当大面积相似颜色信息干扰，并且运动目标长时间处于复杂背景中时，本算法对运动目标的跟踪情况并不理想。Kalman 滤波具有预测功能，但当长时间处于预测中时，误差会不断积累，最终也会导致跟踪失败。

目标跟踪应用领域广泛，而在日常应用中，背景不会是理想条件下，因此传统 Mean-Shift 跟踪算法在实际应用中效果并不理想，对 Mean-Shift 算法进行改进是必要的。本文着眼于相似颜色信息干扰的复杂背景下的目标跟踪，提出了改进的 KSMS(Kalman Skin Mean-Shift)跟踪算法，实验结果表明，在相似颜色信息干扰情况下，例如手和脸部形成连通域时，本算法能够很好的对运动目标进行跟踪。

首都师范大学信息工程学院

第五章 总结与设想

本文首先分析了目前跟踪领域的现状和发展趋势，并分析了目前主要的目标跟踪算法。对传统 Mean-Shift 算法进行了详细介绍，该方法计算量小，对局部遮挡、目标变形有一定适应性，因此广泛应用于目标跟踪领域。然而在实际应用中，该算法曝露出很多缺点，Mean-Shift 算法是基于颜色信息的跟踪算法，因此背景或者天气情况都会对跟踪结果产生巨大影响，本文主要解决相似颜色信息干扰情况下的跟踪问题。Kalman 滤波具有预测功能，本文研究将 Mean-Shift 跟踪算法和 Kalman 滤波融合，并结合肤色信息的位置修正算法解决相似颜色信息干扰下的手势跟踪丢失问题。

本文对 Kalman 滤波的发展历史和目前发展现状进行了详细介绍，对 Kalman 原理进行了详细分析。为 Mean-Shift 算法和 Kalman 滤波进行融合做好了理论知识准备。在简单背景下，当因手运动速度过快而导致的目标跟踪丢失时，可以通过基于肤色信息的位置修正算法能够重新检测并定位手的位置坐标；在复杂背景下，利用 Kalman 滤波的预测功能，将 Mean-Shift 算法和 Kalman 滤波进行融合，实现了相似颜色信息干扰下的目标跟踪。实验结果表明，本文提出的 KSMS 融合算法能够很好的解决相似颜色信息干扰情况下的跟踪丢失问题。

最后，本文分析了 KSMS 融合算法的缺点，在跟踪准确性提高的前提下，牺牲了一定的实时性。本文并对改进融合算法提出了自己的设想，将 Kalman 滤波改为快速 Kalman 滤波算法，可能会提高算法的执行效率。有利于实现实时跟踪任务。

首都师范大学信息工程学院

参考文献

[1]宋家东. 基于多媒体传感器目标跟踪与识别的研究[D]. 首都师范大学: 2011, 11-20.
[2]薛陈, 朱明, 陈爱华. 鲁棒的基于改进 Mean-shift 的目标跟踪[J]. 光学精密工程, 2010, 18(1): 234-239.
[3]安国成, 陈建军, 张凤军, 戴国忠. 改进权值计算的均值移动目标跟踪[J]. 中国图像图形学报, 2011, 16(7): 1283-1288.
[4]王明佳, 武治国, 韩广良, 王延杰, 张叶. 自适应窗口选取的 Mean-Shift 目标跟踪[J]. 光子学报, 2012, 42(1), 67-70.
[5]Vladimir Nedovic. Tracking moving video objects using mean-shift algorithm[J]. Artificial Intelligence Department Faculty of Science University of Amsterdam.
[6]王建华, 冯帆, 梁伟. 基于 Mean-shift 的运动目标跟踪[J]. 科学技术与工程, 2012, 12(3): 564-566.
[7]胡波, 陈恳, 徐建瑜, 张云. 基于kalman预测和Mean-shift算法的视频目标跟踪[J]. 光电子·激光, 2009, 20(11): 1517-1519.
[8]CHENG Y. Mean shift, mode seeking, and clustering[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1995, 17(8): 790-799.
[9]Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using mean shift[C]. IEEE Conference on Computer Vision and Pattern recognition, 2000, 142-149.
[10]Xinhong Zhang, Yanbin Cui, Duoyi Li, Xiangxing liu, Fan zhang. An adaptive mean shift clustering algorithm based on locality-sensitive hashing[J]. Optik International Journal for Light and Electron Optics, 2012, 123(20): 1891-1894.
[11]Yizong Cheng. Mean-shift, Mode Seeking, and Clustering[J]. IEEE TRANSACTI ONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 1995, 17(8): 790-793.
[12]J.Bernal, F.Vilarino, J.Sanchez. Feature Detectors and Feature Descriptors: Where We Are Now[D]. Computer Vision Center and Computer Science Department: 2010, 5-30.
[13]王新琭. 基于传感网的目标跟踪调度算法研究及业务服务子系统实现[D]. 首都师范大学: 2010, 38-39.
[14]章毓晋. 图象工程上册图象处理和分析[M]. 北京:清华大学出版社, 1999, 94-117.
[15]唐文平, 胡庆龙. 基于多目标 Camshift 手势识别[J]. 电子科技, 2012, 25(2): 71-74.
[16]范五东, 周尚波, 辛培宸. HSV 颜色空间特征与 Kalman 滤波融合的目标跟[J]. 计算机工程与应用, 2011, 47(13): 169-173.
[17]李文生, 解梅, 邓春健. 基于机器视觉的动态多点手势识别方法[J]. 计算机工程与设计, 2012, 33(5): 1988-1992.
[18]唐勇, 孙磊, 孙宾. 一种针对视频对象快速移动和遮挡的改进 mean-shift 跟踪算法[J]. 燕山大学学报, 2010, 34(1): 52-55.
[19] Dorin Comaniciu, Visvanathan Ramesh, Peter Meer. Kernel-Based Object Tracking[J]. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 2003, 25(5): 564-575.
[20]彭丁聪. 卡尔曼滤波的基本原理及应用[J]. 软件导刊, 2009, 8(11): 33-34.

首都师范大学信息工程学院

[21]刘玉进，蔡勇，武汇岳，张凤军，戴国忠．一种肤色干扰下的变形手势跟踪方法[J]．计算机工程与应用，45(35)：164-167.
[22]Sarah C. Rutan. Adaptive Kalman Filtering[J]. ANALYTICAL CHEMIST RY, 1991, 63(22): 1103-1109.
[23]乔坤，郭朝勇，史进伟．基于卡尔曼滤波的运动人体跟踪算法研究[J]．计算机与数字工程，2012，40(1)：1-4.
[24]ZHOU Dao-bing, XIAO Guo-qiang, ZHANG Bei-bei. Tracking Moving Objects in MPEG Videos via Kalman Filtering[J]. 西南师范大学学报(自然科学版)，2009，34(6)：113-118.
[25]柏菁，刘建业，袁信．模糊自适应卡尔曼滤波技术研究[J]．信息与控制，2002，31(3)：193-197.
[26]李庆瀛，褚金奎，李荣华，王洪青．基于卡尔曼滤波的移动机器人运动目标跟踪[J]．传感器与微系统，2008，27(11)：66-71.
[27]刘惟锦，章毓晋．基于 Kalman 滤波和边缘直方图的实时目标跟踪[J]．清华大学学报(自然科学版)，2008，48(7)：1104-1107.
[28]罗元，李玲，张百胜，杨红梅．基于 CamShift 和 Kalman 滤波混合的视频手势跟踪算法[J]．计算机应用研究，2009，26(3)，1163-1165.

首都师范大学信息工程学院

本科期间的科研情况

2012-2013 学年度首都师范大学实验室开放基金（开放性项目）支持。

纪成双，高金，解悦，丁辉. 改进的 KSMS 手势跟踪算法研究[J]. 光学技术(核心期刊，已录用)。

首都师范大学信息工程学院

致　谢

今天，我完成了本科生毕业论文的写作，我的大学生涯即将结束。

本研究及学位论文是在我的导师丁辉老师的亲切关怀和悉心指导下完成的。她严肃的科学态度，严谨的治学精神，精益求精的工作作风，深深地感染和激励着我。丁辉老师不仅在学业上给我以精心指导，同时还在思想、生活上给我以无微不至的关怀，在此谨向丁辉老师致以诚挚的谢意和崇高的敬意。

感谢毕业设计小组的同学们，他们是：吴宇航、高金、解悦、陈萌、刘舒荷。正是由于你们的帮助和支持，我才能克服一个又一个的困难和疑惑，直至本文的顺利完成。

感谢其他老师四年来的谆谆教诲，是你们丰富了我的计算机专业知识。使我从一个对计算机毫无了解的情况下，成长为一个具有一定计算机专业知识的本科毕业生。

感谢大学四年来的同学们，你们给我带来了无数欢乐，在这里请接受我诚挚的谢意!

同时也感谢学院为我提供良好的做毕业设计的环境。

感谢我的爸爸妈妈，养育之恩，无以回报，你们永远健康快乐是我最大的心愿。

最后再一次感谢所有在毕业设计中曾经帮助过我的良师益友和同学，以及在论文中被我引用或参考的论著的作者。

首都师范大学信息工程学院