

论文题目：基于目标识别的跟踪导航系统软件设计

专 业：测控技术与仪器

本 科 生：向莹光

(签名) _____

指导教师：秦学斌

(签名) _____

摘要

视觉跟踪是计算机视觉中的一个基础性和普遍存在的问题，具有众多的应用价值。任务是在第一帧中选择一个任意目标，通常是一个未知对象和一个图像的矩形区域，并在下面的连续帧中检测它的位置，称为跟踪检测。在过去几十年中作出了努力，并在结果的准确性和鲁棒性方面取得了重大进展。由于任务的性质，跟踪器被要求是通用的，并用于任何类型的对象类，这是不允许应用任何特定类的先验知识的。

本文利用图像处理技术，实现对目标的识别，目标的跟踪。在运动的背景下，由于背景的变化和目标运动产生形变，对目标的检测、跟踪产生很大的阻碍问题。跟踪器的核心组件是一种判别分类器，任务是区分目标和周围环境。为了应对自然图像的变化，此分类器通常使用转换和缩放的样本修补程序进行训练。通过控制最大样本采集量，来防止训练分类器太过单一，并遵循高斯分布，这不仅减少了分类器的错误检测率，而且对变异具有鲁棒性。

本设计在 Linux 系统下，移动小车为平台，搭建摄像头，实现对视频的实时采集。同时利用内核相关性滤波算法、目标的识别检测，并实现小车对目标的跟踪导航功能。

关键词：视觉跟踪，图像处理，分类器

Subject: Software design of tracking and navigation system based on target recognition

Specialty: Technology of Measure & Control and Instrument

Name:Xiangyingguang

(Signature) _____

Instructor: Qin Xue-bin

(Signature) _____

Abstract

Visual tracking is a basic and ubiquitous problem in computer vision, which has many application values. The task is to select an arbitrary target in the first frame, usually an unknown object and a rectangular area of an image, and detect its location in the following successive frames, called trace detection. Efforts have been made over the past few decades and significant progress has been made in the accuracy and robustness of the results. Because of the nature of the task, the tracker is required to be generic and used for any type of object class, which is not allowed to apply prior knowledge of any particular class.

This paper uses image processing technology to realize target recognition and target tracking. Under the background of the movement, because of the change of background and the deformation of the target motion, the detection and tracking of the target have a great obstacle. The core component of the tracker is a discriminant classifier, the task is to distinguish between the target and the surrounding environment. In order to respond to changes in natural images, this classifier is typically trained with a sample patch of transformations and scaling. To prevent the training classifier from being too single and to follow the Gaussian distribution by controlling the maximal sampling quantity, this not only reduces the error detection rate of the classifier, but also has robustness to the mutation.

This design utilizes the embedded development environment, the mobile trolley as the platform, constructs the camera, realizes the video real-time collection. At the same time, the kernel correlation filtering algorithm is used to realize the target recognition and detection, and the tracking and navigation function is realized.

Keywords: Visual tracking , image processing technology , discriminant classifier

目录

第一章 绪论	5
1.1 研究背景和意义	5
1.2 本课题国内外研究现状和研究成果	5
1.3 本课题基本任务及工作安排	6
第二章 目标识别跟踪导航技术	7
2.1 目标检测技术	7
2.1.1 最小二乘法拟合直线	7
2.1.2 SVM 超平面法	7
2.2 目标跟踪算法	8
2.2.1 光流法	8
2.2.2 背景差分法	9
2.2.2 TLD 算法	10
2.2.3 粒子滤波算法	11
2.2.4 相关性滤波算法	12
2.3 本章小结	13
第三章 系统的硬件设计	13
3.1 硬件功能简介	13
3.2 硬件模块	14
3.2.1 电源模块选取	14
3.2.2 电机选取	14
3.2.3 电机驱动	14
3.2.4 图像传输模块	15
3.3 本章小结	15
第四章 系统的软件设计	15
4.1 软件系统简介	15
4.1.1 OpenCV 图像处理	15
4.1.2 Linux 环境程序开发	16
4.1 软件系统框图	16
4.1.1 下位机程序流程图	16
4.2 子模块介绍	17

4.2.1 目标跟踪导航算法	17
4.2.2 网络图像传输	17
4.2.3 小车控制算法	18
4.3 本章小结	19
第五章 系统运行测试	20
5.1 系统测试方案	20
5.1.1 静止下运动目标物的跟踪	20
5.1.2 运动下运动目标物的跟踪	20
5.2 系统测试结果	21
5.3 本章小结	21
第六章 总结与展望	22
6.1 工作总结	22
6.2 展望	22

第一章 绪论

1.1 研究背景和意义

视觉作为人类获取外界信息重要的手段，是我们日常生活，工作中身体不可缺少的。研究人类视觉信息处理，是十分必要的。从信息的处理机制搞清楚它的原理，研究其中的信息处理算法，以便计算机实现人类的视觉、听觉等模式识别的能力。物体在可见光的照耀下，通过瞳孔折射在视网膜上形成图像，再由感光细胞转换成神经脉冲信号，经神经纤维传输大脑皮层进行信息处理。人类的视觉不仅是对光信号的感受，它包括了对视觉信息的获取、传输、处理、存储与解析。信号处理理论与计算机科学技术的发展，人们尝试用摄像机获取环境图像，通过图像处理技术将图像信息转换数字信号量，用计算机实现对视觉信息处理的全过程，被人类称之为机器视觉。

近些年来，得益于硬件电子和芯片技术的发展，计算机数字处理能力大幅提升，使运动物体的检测和跟踪在图像处理、计算机视觉、模式识别、人工智能、多媒体技术等领域越来越受到人们的关注。运动跟踪和检测的应用广泛，在智能监控和人机交互中，如：银行、交通、超市等场合常常使用运动跟踪分析技术，通过定位物体并对其行为进行分析，一旦发现物体有异常行为，监控系统就发出警报，提醒人们注意并即时的处理，改善了人工监督注意力分散、反应时间较慢、人力资源浪费等问题。运动目标的跟踪在虚拟现实、工业控制、军事设备、医学研究、视频监控、交通流量观测监控等很多领域也有重要的实用价值。特别在军事上，先进的武器导航、军事侦察和监控中都成功运用了自动跟踪技术。而跟踪的难点在于如何快速而准确的在每一帧图像中实现目标定位。正因如此，对运动目标的跟踪和检测的研究很有价值。

1.2 本课题国内外研究现状和研究成果

由运动目标所形成的图像序列可分为两种情况：一种是静态场景，一种是动态场景。前一种情况通常发生在摄像机相对静止状态（如监视某一路口车流量的固定摄像机），后一种情况通常发生在摄像机也在相对运动状态（如装在卫星或飞机上的监视系统）。

近几年来，研究人员提出了许多方法用于检测和跟踪序列图像中的运动目标，其中绝大多数方法都假设用于获取序列图像的背景是静止的，而大多数实际情况也满足这一假设，所以人们主要研究的是背景静止—物体运动这种情况下获取的动态序列图像。针对这种情况，现有的运动目标检测方法主要有三类：帧差法（Temporal difference）、背景减法（Background subtraction）、光流场法（Optical flow）。

（1）帧差法即帧间差分法，是在连续的图像序列中，提取两个或三个相邻帧，采用基于像素的时间差分来获取图像中的运动信息。帧间差分法计算简单，对动态环境具有较强的适应性，但不能提取出目标的所有相关点，容易在目标内部形成“空洞”，而且要求背景绝对静止或基本无变化(噪声较小)，因而适用场合有限。

(2) 背景减法也称背景差分法,基本思想是将当前帧图像与事先存储或者实时得到的背景图像相减,即当前图像与已知背景图像之间的差分。较早期的背景差分法有均值、中值、运动平均、高斯模型方法等,随着研究的深入,逐渐出现了一些比较成熟的背景差分方法。1999年 MIT 的 C.Stauffer 提出了基于混合高斯模型的背景差分方法,针对高斯噪声的复杂背景,采用多个高斯分布描述像素过程;2000年 MIT 的 Oliver 等人提出了背景特征化建模的方法。主要研究的焦点都落在了背景更新与建模方法上。从实际应用角度来看,帧差和背景减相结合的运动检测是使用最广泛的一种方法。Spagnol 等人提出了一种运用邻域的相关系数结合帧差和背景减来进行运动检测的方法,该方法不仅有效地抑制了光照变化对检测结果的影响,并解决了阴影、重影和伪前景等问题。但国内有人指出了 Spagnol 的方法的三个不足之处:①会将背景颜色一致的区域误判为前景;②当运动目标速度缓慢或静止时,存在漏检;③在背景复杂的情况下,背景更新策略将使背景模型退化。并且,提出了首先用灰度拉伸变换,并结合灰度值信息来改进邻域相关系数的计算方法,以排除背景的误判;然后通过帧差和背景减相结合的策略中加入运动分析,用来改善运动缓慢目标的检测结果;最后在背景复杂的情况下用运行期更新法进行背景更新,以防止背景模型的退化。

(3) 光流场法是分析序列图像中运动目标的重要方法,它既可以用于运动目标的检测,也可以用于运动目标的跟踪。所谓光流是指图像中灰度模式运动的速度;它是景物中可见点的三维速度矢量在成像平面上的投影,它表示了景物表面点在图像中位置的瞬时变化;同时光流场携带了有关运动和结构的丰富信息。光流场的计算最初是由国外两名学者提出的,它是一种以灰度梯度基本不变或亮度恒定的约束假设为基础的运动目标探测的有效方法。光流法能够较好的处理运动背景的情况,但计算量巨大,难以应用到实时系统,同时对噪声比较敏感,精度较低,难以得到运动目标的精确边界。而张泽旭等将 Canny 边缘提取融入光流场分割技术,对单运动目标和多运动目标均取得了比较满意的效果,实时性也大大提高。但如何进一步提高其抗噪性、实时性和运算速度,有待于深入研究。

总的来说,过去的二三十年中,国内外学者在运动目标检测理论及其实现方面做了大量、深入的研究工作,取得了令人瞩目的成果,并将研究的成果应用到很多领域,如智能监控系统、军事制导等等。但是由于运动目标检测问题本身的复杂性和目前视频技术发展水平的限制,运动目标的检测技术至今仍然处于研究和讨论阶段,到目前为止,还没有出现能够适用于各种场合、各种情况的通用算法。现今的各种算法在稳健性、准确性、可靠性等方面还有着这样、那样的不足。这些算法有的简单、易于实现,实时性好,但可靠性不高,在复杂场景和运动下算法容易失效;有的虽然可靠性较高,但算法过于复杂,不利于进行实时检测。而且,静态场景下的运动目标检测的研究较多;而动态场景下的检测研究较少,算法还不成熟。已有的运动目标检测系统大多受限于特定的应用场景,目标检测算法还有待于进一步研究和优化,研究一种鲁棒性好、精确度高、性能稳定和适用性强的运动目标检测算法依然任重道远,面临巨大挑战。

1.3 本课题基本任务及工作安排

本课题主要任务通过对摄像头输入图像信息,采用目标识别分类算法,检测目标物在视频流中的位置,在四轮车平台上,实现对实时视频中运动目标物的跟踪。

第二章 目标识别跟踪导航技术

2.1 目标检测技术

视频目标检测技术，是根据目标的表征信息，目标的运动轨迹，在视频中标出目标的位置。运动目标检测是实现目标跟踪的前提条件，在后期处理过程中，仅仅对检测出的运动区域的像素进行处理，因此选择好的检测方法非常重要。当前目标检测算法的发展目标是能够抵御环境变化的影响(如背景与目标相似性，光照的变化)，不依赖于摄像头的安装位置。

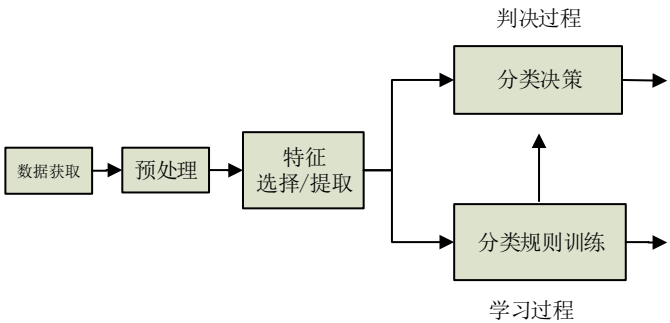


图 1 基于模式识别的目标特征提取技术

2.1.1 最小二乘法拟合直线

在目标识别中，常用到分类检测技术。在求采集样本和输出响应的函数关系的时候，常用最小二乘法，拟合直线描述输入输出效应。在模式识别中，如果预测的变量是离散的，我们称其为分类（如决策树，支持向量机等），如果预测的变量是连续的，我们称其为回归。回归分析中，如果只包括一个自变量和一个因变量，且二者的关系可用一条直线近似表示，这种回归分析称为一元线性回归分析。如果回归分析中包括两个或两个以上的自变量，且因变量和自变量之间是线性关系，则称为多元线性回归分析。对于二维空间线性是一条直线；对于三维空间线性是一个平面，对于多维空间线性是一个超平面。

对于一元线性回归模型，参数估计采用最小二乘法，最小二乘法（又称最小平方方法）是一种数学优化技术。它通过最小化误差的平方和寻找数据的最佳函数匹配。利用最小二乘法可以简便地求得未知的数据，并使得这些求得的数据与实际数据之间误差的平方和为最小。最小二乘法还可用于曲线拟合。

曲线拟合分为线性拟合与非线性拟合，其他一些优化问题也可通过最小化能量或最大化熵用最小二乘法来表达。

2.1.2 SVM 超平面法

Cortes 和 Vapnik 于 1995 年首先提出的 SVM 超平面法(Support Vector Machine)，它在解决小样本、非线性及高维模式识别中表现出许多特有的优势，并能够推广应用到函数拟合等其他机器学习问题中。

支持向量机方法是建立在统计学习理论的 VC 维理论和结构风险最小原理基础上的, 根据有限的样本信息在模型的复杂性 (即对特定训练样本的学习精度, **Accuracy**) 和学习能力 (即无错误地识别任意样本的能力) 之间寻求最佳折衷, 以期获得最好的推广能力[14] (或称泛化能力)。

统计学习因此而引入了泛化误差界的概念, 就是指真实风险应该由两部分内容刻画, 一是经验风险, 代表了分类器在给定样本上的误差; 二是置信风险, 代表了我们在多大程度上可以信任分类器在未知文本上分类的结果。很显然, 第二部分是没办法精确计算的, 因此只能给出一个估计的区间, 也使得整个误差只能计算上界, 而无法计算准确的值 (所以叫做泛化误差界, 而不叫泛化误差)。

置信风险与两个量有关, 一是样本数量, 显然给定的样本数量越大, 我们的学习结果越有可能正确, 此时置信风险越小; 二是分类函数的 VC 维, 显然 VC 维越大, 推广能力越差, 置信风险会变大

2.2 目标跟踪算法

2.2.1 光流法

1950 年, Gibson 首先提出了光流的概念, 所谓光流就是指图像表现运动的速度。物体在运动的时候之所以能被人眼发现, 就是因为当物理运动时, 会在人的视网膜上形成一系列的连续变化的图像, 这些变化信息在不同时间, 不断的流过眼睛视网膜, 就好像一种光流过一样, 故称之为光流。

光流场是一种二维 (2D) 的瞬时速度场, 在景物可见点中, 三维 (3D) 速度矢量在成像表面上形成的投影就是 2D 速度场矢量。光流包含了观测物体的运动变化信息, 同时还含有景物三维结构的信息。在许多问题的研究中, 光流都起着重要作用, 在计算机视觉的较高层视觉算法中, 光流是一个比较重要的输入, 可以实现诸如目标分割、识别、跟踪、机器人导航以及形状信息恢复等重要的任务; 通过二维物体表面上的光流来研究三维结构和运动, 是计算机视觉研究领域内一项很有意义和挑战的任务; 在其它领域的应用中, 光流信息的作用也是很明显的, 例如心脏手术后, 对病人心脏波动的监测、海洋及大气过程的研究和预测等。所以看出, 光流的研究不仅仅计算机视觉中, 也在其他有关研究领域是一个重要部分。光流法是空间运动物体在观察成像平面上的像素运动的瞬时速度, 是利用图像序列中像素在时间域上的变化以及相邻帧之间的相关性来找到上一帧跟当前帧之间存在的对应关系, 从而计算出相邻帧之间物体的运动信息的一种方法。一般而言, 光流是由于场景中前景目标本身的移动、相机的运动, 或者两者的共同运动所产生的。其计算方法可以分为三类:

- (1) 基于区域或者基于特征的匹配方法;
- (2) 基于频域的方法;
- (3) 基于梯度的方法;

简单来说, 光流是空间运动物体在观测成像平面上的像素运动的“瞬时速度”。光流的研究是利用图像序列中的像素强度数据的时域变化和相关性来确定各自像素位置的“运动”。研究光流场的目的是为了从图片序列中近似得到不能直接得到的运动场。

光流法用于目标检测的原理: 给图像中的每个像素点赋予了一个速度矢量, 这样就形成了一个运动矢量场。在某一特定时刻, 图像上的点与三维物体上的点一一对应, 这种对应关系可以通过投影来计算得到。根据各个像素点的速度矢量特征, 可以对图像进行动态分析。如果图像中没有运动目标, 则光流矢量在整个图像区域是连续变化的。当图像中有运动物体

时，目标和背景存在着相对运动。运动物体所形成的速度矢量必然和背景的速度矢量有所不同，如此便可以计算出运动物体的位置。需要提醒的是，利用光流法进行物体检测时，计算量较大，无法保证实时性和实用性。

光流法用于目标跟踪的原理：

- 1) 对一个连续的视频帧序列进行处理；
- 2) 针对每一个视频序列，利用一定的目标检测方法，检测可能出现的前景目标；
- 3) 如果某一帧出现了前景目标，找到其具有代表性的关键特征点；
- 4) 对之后的任意两个相邻视频帧而言，寻找上一帧中出现的特征点在当前帧中的最佳位置，从而得到前景目标在当前帧中的位置坐标；

如图 3 所示，假设已有 $H(x,y)$ ， $I(x,y)$ 两个图片，如何计算 H 到 I 像素点之间的运动？显然，对于 H 中特定的像素点，我们应该在 I 图片中对应位置的周围来寻找像素值一致或者接近一致的像素点。如此一来，不难发现，解决此类问题，一般有两个关键的假设：

- (1)颜色一致
- (2)微小运动；即，每个像素点，都不会产生较大的运动偏移。

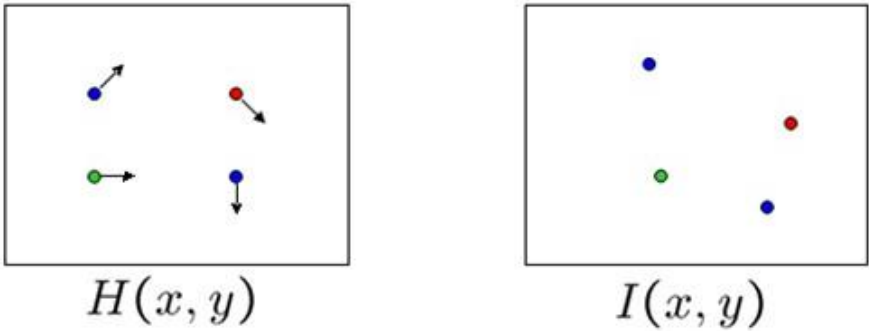


图 3 偏移过程

再如下图 4 所示， H 中的像素点 (x,y) 在 I 中的移动到了 $(x+u,y+v)$ 的位置，偏移量为 (u,v) 。

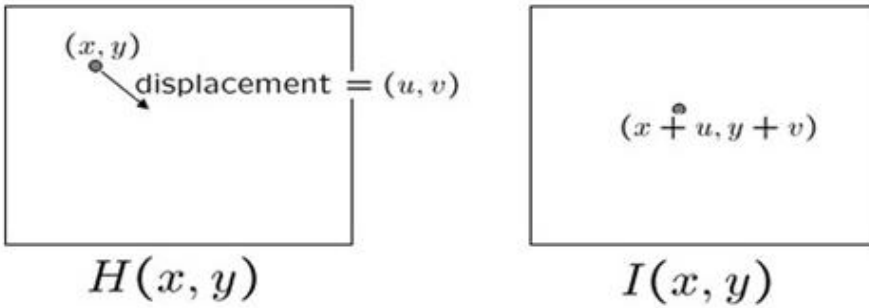


图 4 偏移量

2.2.2 背景差分法

背景差分法的基本原理是：先确定一个作为参考的背景图像，然后将当前帧与背景帧相减，得到差分图像，接着用阈值 T 判断得到当前时刻图像中与背景

模型灰度值相差较大的点，若差分值大于 T 则认为是目标，并令该点为 1；反之，认为是背景点，使该点为 0，从而分割出运动目标。函数表达式如下：

$$d_i(x, y) = \begin{cases} 1 & \text{if } |f(x, y, t) - b(x, y)| > T \\ 0 & \text{else} \end{cases} \quad (3.2)$$

背景差分法可以完整准确的分割出运动目标，提取运动物体的信息，这是它的一大优点，但是同时背景差分法也有诸多的限制，比如对环境的变化非常敏感，另外背景图像的选取也十分关键，直接关系到差分结果的准确性。背景差分法中有两个问题需要解决：

(1) 背景图像的获取：背景图像获取应当是在完全静止，没有任何运动物体的情况下进行的，而在有些实际情况中想得到一副完全静止的背景图像是很难办到的。环境条件的不断变化是一个不容忽视的因素，比如光照的变化，摄像机得抖动等等。

(2) 背景图像的更新：背景差分法只适用于室内等外界条件比较理想的场合，而在长时间的监控中监控点的背景不可能一成不变，所以需要在一段时间内更新背景图像，以适应环境的变化。

一般来说，实现背景差分法大致分为三个步骤：

第一步，获取监控点的理想背景图像；

第二步，将当前图像和背景图像作差分，然后设置合理的阈值，对差分后的结果进行二值化处理，从而得到运动的目标。

第三步，周期性的对背景图像进行更新以适应动态背景的变化。

2.2.2 TLD 算法

TLD (Tracking-Learning-Detection)算法是 Zdenek Kalal 在 2008 年到 2010 年间提出并不断完善的一种视觉跟踪算法，此算法主要由三个模块组成：跟踪模块，在线学习模块和检测模块。首先，TLD 算法训练一个目标检测器，通过第一帧图像获得运动目标特征，再通过仿射变换这些特征初始化检测器。跟踪器和检测器同时运行，估计目标的位置。在线学习模型不断的更新检测器和跟踪器的数据，同时跟踪器和检测器也将得到的数据反馈给在线学习模型，验证数据的精确性，使得到的目标特征更加精确，其三者相互作用，将得到的数据整合，最终确定下一帧中目标的位置。算法流程示意图如图 1 所示。

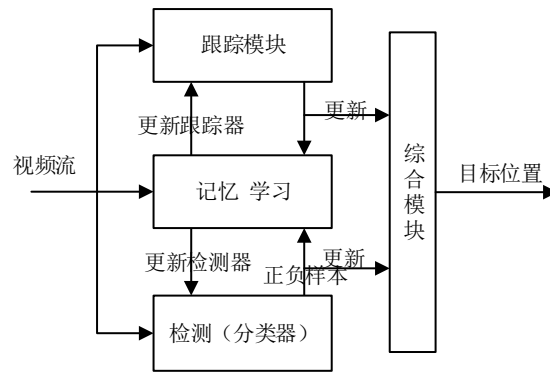


图 1 TLD 算法流程图示意图

由于 TLD 算法的跟踪器采用 Lucas-Kanade 光流法(LK 光流法) 进行跟踪，LK 光流法是基于运动一致连贯性假设的，算法不适用于大幅度运动目标的跟踪。

2.2.3 粒子滤波算法

粒子滤波算法即粒子滤波器，其核心思想为最优贝叶斯估计，由当前和以前的观测值来估计目标的当前状态。在非线性、非高斯的视频运动目标状态估计的中，滤波器表现出了非常高的优越性。



图 2.2 粒子跟踪图示意图

粒子滤波[12,14]最初是为了在混乱的背景下最终目标物而开发的。跟踪对象的状态由向量 X_t 描述，而向量 Z_t 表示到时间 t 的所有观察值 $\{z_1, \dots, z_t\}$ 。粒子是经常用于后验密度 $p(X_t | Z_t)$ 和观察密度 $p(Z_t | X_t)$ 是非高斯的。关键技术是用带权样本集合 $S = \{(s(n), \pi(n)) | n = 1, \dots, N\}$ 近似概率分布。样本 S 代表对目标状态的预测，并对应离散采样概率 π 。

通过根据系统模型传播每个样本来描述样本集的演变。然后，根据观察值对集合中的每个元素进行加权，并通过选择具有概率 $\pi^{(n)} = p(z_t | X_t = s_t^{(n)})$ 的特定样本来绘制 N 个样

本。在每个时间阶段 $E[S] = \sum_{n=1}^N \pi^{(n)} s^{(n)}$ 估计对象的平均状态。

粒子滤波提供了一个强大的跟踪框架，因为它模拟了不确定性。它可以保持其选项的打开，同时考虑多个状态假设。由于不太可能的对象状态有机会暂时保持在跟踪过程中，粒子滤波器可以处理短暂的遮挡。

2.2.4 相关性滤波算法

相关性滤波算法，基于判别方法的单目标运动跟踪算法，利用离散傅里叶变换降低分类器训练和检测的过程中计算量，使跟踪速度能到达 172 帧/秒。算法通过循环位移矩阵，对目标区域进行采样，因此形成的大量样本训练分类器，通过核函数计算样本区域和目标区域的相似度，输出相似度最大的作为跟踪目标的位置。

2.2.4.1 线性回归求解分类器。

在采集样本训练分类器中，为了找到函数 $f(z) = w^T z$ 描述样本特征与目标输出关系，其中 $w = [w^{(1)} w^{(2)} w^{(3)} \dots w^{(n)}]^T$ 为样本权重， $z = [z^{(1)} z^{(2)} z^{(3)} \dots z^{(n)}]^T$ 为样本的特征。令 x_i 为第 i 个样本， $x_i = [x_i^{(1)} x_i^{(2)} x_i^{(3)} \dots x_i^{(n)}]^T$ ， y_i 为 x_i 对应的回归值，通过训练找到最小化平方误差下的 w ，即

$$\min_w \sum_i (f(x_i) - y_i)^2 + \lambda \|w\|^2 \quad (2-1)$$

其中 λ 为防过拟合的参数，通过正则化稀疏模型最优闭式解为

$$w = (X^T X + \lambda I)^{-1} X^T y \quad (2-2)$$

其中矩阵 X 的每行为 x_i ，即 $X = [x_1 x_2 x_3 \dots x_n]^T$ ， I 为单位矩阵， $y = [y_1 y_2 y_3 \dots y_n]^T$ 。

2.2.4.2 循环矩阵采集样本

这里以单通道，一维图像为例。一个基样本 x_i 表示为 $n \times 1$ 的向量，通过置换矩阵循环左右位移样本产生正负样本。

$$P = \begin{bmatrix} 0 & 0 & 0 & & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & & 0 & 0 \\ & \vdots & & \ddots & \vdots & \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} \quad (2-3)$$

基样本乘一次 P ， x_i 中的元素就向右循环偏移一位， $P \bullet x_i = [x_n x_1 x_2 \dots x_{n-1}]^T$ 。如图 2.4 所示

$$X = C(x) = \begin{vmatrix} x_1 & x_2 & x_3 & \dots & x_{n-1} & x_n \\ x_n & x_1 & x_2 & \dots & x_{n-2} & x_{n-1} \\ x_{n-1} & x_n & x_1 & \dots & x_{n-3} & x_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_2 & x_3 & x_4 & \dots & x_n & x_1 \end{vmatrix} \quad (2-4)$$



图 2.4 循环位移采样

2.3 本章小结

通过对以上四种方法的比较，可以看出背景差分法能够直接得到运动对象的位置、大小、形状等信息，并且背景差分法计算复杂度适中，相对来说可以准确完整的提取运动目标的信息。因此本文采用背景差分法来实现运动目标的识别。

第三章 系统的硬件设计

3.1 硬件功能简介

本设计为目标识别跟踪导航系统，初始设计运动目标物选取以人为主。在四轮两驱小车平台上，使用摄像头传感器采集视频，利用图像处理算法，实现对特定的运动目标物的识别，并进行运动跟踪导航。

在设计中，四轮两驱小车搭载嵌入式开发板友善之臂的八核 A53 处理器，100WUSB 摄像头采集图像帧。首先选定需要跟踪的目标物的初始位置信息，基于目标识别的算法，输出选定特征物在接下来的摄像头采集上来的图像帧中的位置信息。利用 PID 控制器，输入为目标位置信息的变化，输出为小车的驱动控制，包括速度和方向的调节，实现小车的对运动目标物的跟踪。并利用 TCP/IP 网络传输协议，在 windows 系统的电脑上，利用集成开发环境 VisualStudio2013 搭建和图像处理库搭建 Socket 服务器，实现图像的接收与现实。在下位机 Linux 系统下利用图像处理库搭建 Socket 客户端实现视频的无线传输。图像的采集，图像的

处理，传输与显示，利用开源机器视觉 Opencv 图像处理函数库，实现了设计需求。本设计，经近期验证，具有一定的可行性。

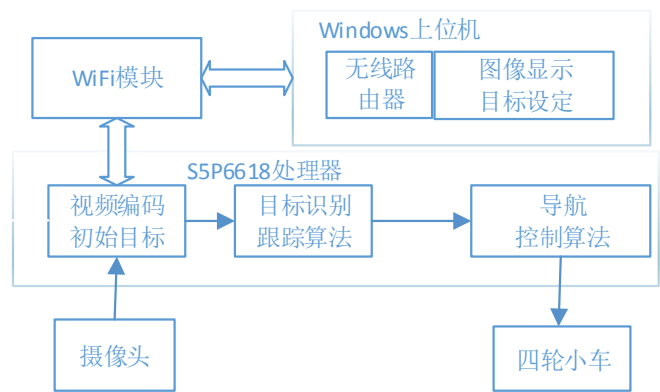


图 1 系统框架图

如图 1 所示，系统的硬件平台整体小车平台，控制器采用 S5P6618 处理器，基于 A53 的八核架构，运行主频 0.4~1.4GHz，满足设计中的目标识别算法，图像处理算法等运行要求。图像的采集采用 100W 像素的 USB 摄像头，摄像头输出为 YUV420 格式，满足系统设计的目标识别要求，以及在视频压缩编码过程中格式转换要求。图像的传输采用无线传输，Windows 上位机通过无线路由器实现与下位机的通讯。

3.2 硬件模块

3.2.1 电源模块选取

在本系统中，系统供电为 5V，L298N 芯片的电源 5V 和电机的电源 7-15V。所以需要对电源的提供必须正确和稳定可靠。用采用 11.1V 的 30C2200MA 航模电池给后轮电机供电，然后使用 LM2596 可调降压模块来把高电压稳成 5V 分别给单片机和电机驱动芯片供电。LM2596 输入直流 3V 到 40V,输出为 1.5V 到 35V，高效率最大输出电流 3A。

3.2.2 电机选取

采用普通直流电机。直流电机运转平稳，精度有一定的保证。直流电机控制的精确度虽然没有步进电机那样高，但完全可以满足本题目的要求。通过处理器的 GPIO 口的 PWM 输出同样可以控制直流电机的旋转速度，实现电动车的速度控制。并且直流电机相对于步进电机价格经济。

3.2.3 电机驱动

采用电机驱动芯片 L298N。L298N 为单块集成电路，高电压，高电流，四通道驱动，可直接的对电机进行控制，无须隔离电路。通过 GPIO 的输出改变芯片控制端的电平，即可以对电机进行正反转，停止的操作，非常方便，亦能满足直流减速电机的大电流要求,同时支持输

入可调 PWM 调节车速的改变。调试时在依照芯片手册，用程序输入对应的码值，能够实现对应的动作。

3.2.4 图像传输模块

图像传输采用无线 USB WIFI 模块 RTL8188CUS，通过路由器和上位机处于同一网段内。当 WiFi 模块在外加 13DB 增益的天线下，可以实现 150M 的发射功能，实现图像的稳定传输，网络传输标准为 802.11n。

3.3 本章小结

本系统硬件平台主要为四轮小车，摄像头实时采集视频。嵌入式控制系统现实对图像的处理，目标的检测，并利用目标检测的位置实现小车的对目标跟踪导航。

第四章 系统的软件设计

4.1 软件系统简介

目标识别跟踪导航系统，基于 Linux 开发平台，在 Windows 下开发上位机，实现上位机视频的显示，同时作为设备遥控端，实现对目标的选取与下发。

在 Linux 下，使用 Source Insight 编辑修改程序，Cmake 作为编译工具，生成 Makefile 文件。

4.1.1 OpenCV 图像处理

OpenCV 是 Open Source Computer Vision Library 的缩写，OpenCV 是因特尔公司于一九九九年创立发展的，现在的技术支持是由 Willow Garage 在进行。OpenCV 的主体是一个 C 的函数库，里面包含了众多的子函数。OpenCV 一个基于 C 语言发行的跨平台的计算机视觉函数库，可以在 XP、Windows、Linux 和 MacOS 多个操作系统上运行。它方便而且可利用性极高。它由一系列的 C 语言函数库和少量 C++ 函数库共同组成，同时提供了 Libview、MATLAB、Ruby、Python、等许多语言的接口，实现了深度数据图像处理和计算机视觉方面的很多通用算法。

OpenCV 函数库所提供的深度图像处理算法非常多，并且其中一部分函数以 C++ 语言来编写，加上它的开源的特性，更加大了它自身的可移植性，有时候不需要添加新的外部函

数，*OpenCV* 自身也可以完整的编译、加载、生成、执行程序，所以它有很大的应用范围。

作为一个开源项目函数处理库里，*OpenCV* 拥有大量的数字图像处理和 PC 摄影视觉等方面的基本算法及其函数源代码。这不仅使实际开发变得更加便利，同时其本身也成为各界需求者研究和学习这一方向理论知识的教科书——*OpenCV* 的生命活力体现不外如是。

OpenCV 不仅仅拥有 Intel 公司给予的多方面技术资源支持，更拥有万千专业人士及业余爱好者的自觉维护和即时更新，诸多国际领先的技术研究成果可以在它的技术基础上得以展现，使得各国技术人员能够便捷的吸收并运用到这些先进的技术理论和科研成果。

4.1.2 Linux 环境程序开发

4.1 软件系统框图

4.1.1 下位机程序流程图

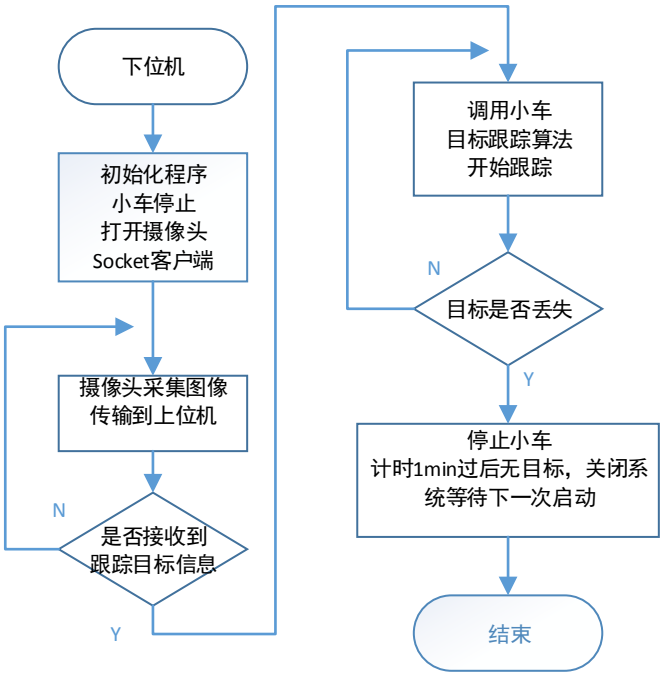


图 2 下位机软件流程图

如图 2 所示，为下位机软件流程图。在目标识别跟踪导航系统中，首先通过摄像头采集

一帧 640*480 大小的图像，通过 H264 视频编码，传输到上位机。同时接收上位机传输回来的跟踪目标信息的数据包。目标信息包括目标坐标与高和宽。下位机在接收到需要跟踪的目标信息过后，利用目标识别算法，在接下来的视频帧中寻找目标新的位置，控制小车移动。没有目标物的位置信息时候，小车停止。

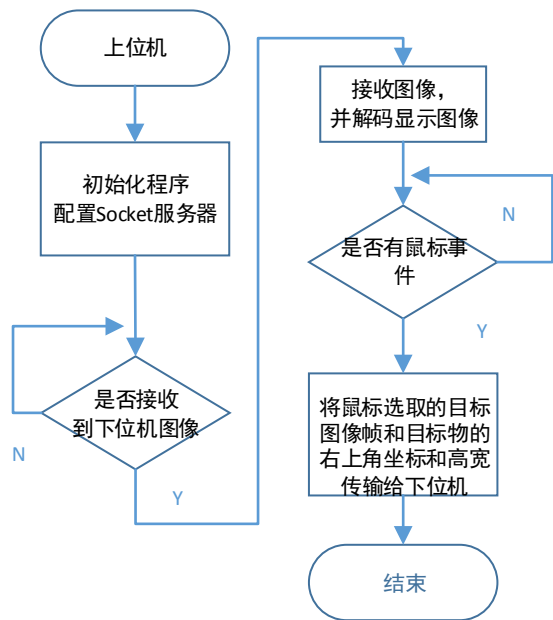


图 3 上位机软件流程图

如图 3 所示，为上位机的软件流程图。在目标识别跟踪导航系统中，上位机的搭建为 ViusalStudio2013 加上图像处理库 Opencv，接收传输回来的 H264 编码的视频流，解码为 Opencv 的图像帧格式，实现视频的显示，和选取需要识别跟踪的目标信息传输给下位机。目标信息包括目标坐标与高和宽。

4.2 子模块介绍

4.2.1 目标跟踪导航算法

算法模块的输入为目标物的一帧图像，和目标物在图像中的坐标，高和宽。输出为目标物在目标物在接下来的视频帧中的坐标，高和宽，在前期的验证中，基于学习，检测，跟踪的运动跟踪算法具有可靠性。

4.2.2 网络图像传输

搭建 Socket 服务器实现 TCP/IP 网络图像传输。Socket 为网络进程间的通信函数，在网络传输层的“协议+端口”标识唯一主机，利用(IP 地址，协议，端口)为标志网络进程，实现网络间的不同进程间的通信。

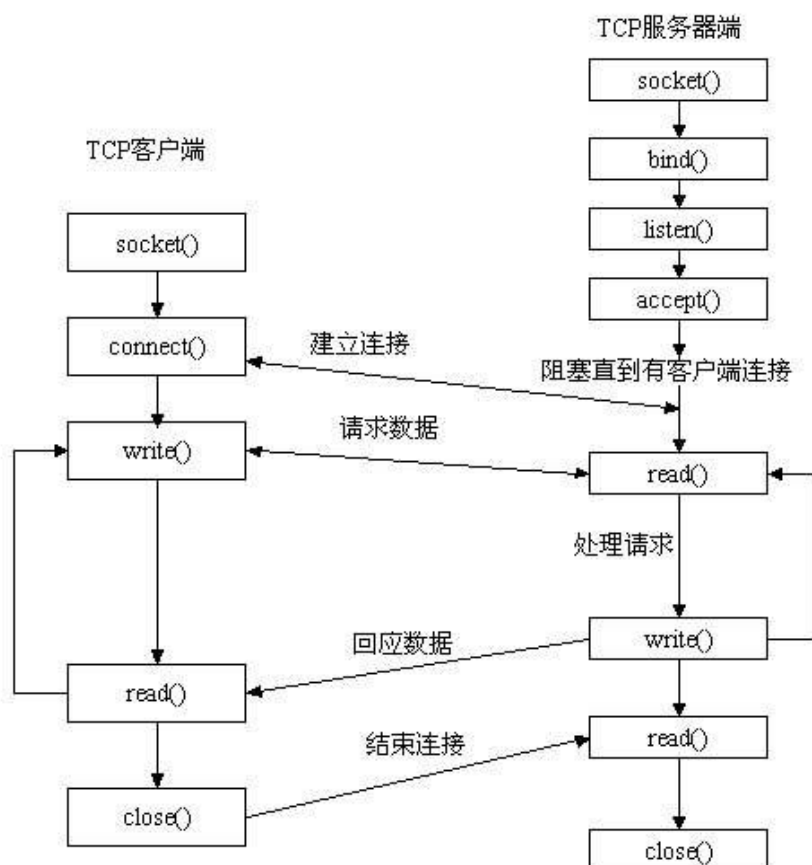


图 5 TCP 数据交互流程

服务器端先初始化 Socket，然后与端口绑定(bind)，对端口进行监听(listen)，调用 accept 阻塞，等待客户端连接。在这时如果有个客户端初始化一个 Socket，然后连接服务器(connect)，如果连接成功，这时客户端与服务器端的连接就建立了。客户端发送数据请求，服务器端接收请求并处理请求，然后把回应数据发送给客户端，客户端读取数据，最后关闭连接，一次交互结束。

4.2.3 小车控制算法

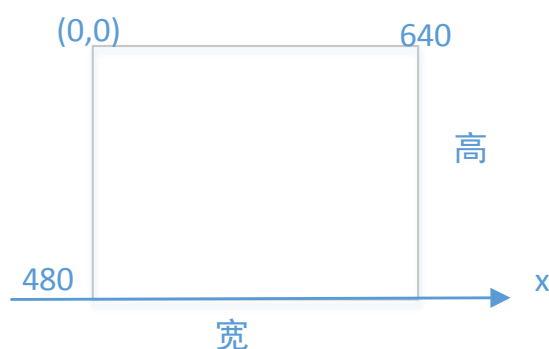


图 6 图像坐标

如图 6 所示，采集的一帧图像大小为 640*480 个像素点。以图像左上角为坐标原点，宽

为 X 轴，作为小车水平方向的控制，当 t 时刻，目标物坐标(X1,Y1) 在初始坐标(X0,Y0)的右边，则向右运动，反之小车向左运动。



图 7 小车速度控制

如图 7 所示,为小车的的速度控制器,输入 $\Delta x(t)$ 为跟踪目标在 $x(t_0)$ 时刻的面积和 $x(t)$ 的目标面积差值。采用开环单输入 PID 控制器，输出 $y(t)$ 为小车的速度。当面积为正值，小车开始加速，当面积为负值时，小车开始加速。

4.3 本章小结

第五章 系统运行测试

5.1 系统测试方案

基于目标识别的跟踪导航系统，分为目标检测，目标跟踪。目标检测过程中，由于目标在运动过程中，存在形变，光照，角度等因素的影响，使检测率大大下降。

5.1.1 静止下运动目标物的跟踪

(1) 目标物的形变

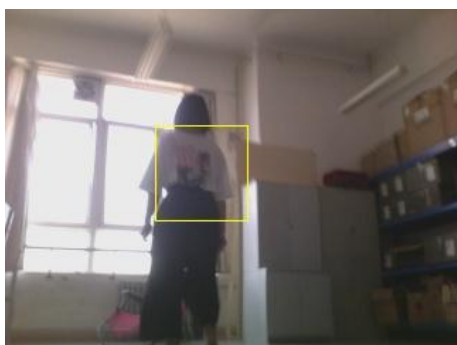


(a) 目标物



(b) 目标物侧身

(2) 目标物背光



(b) 背光 1



(c) 背光 2

5.1.2 运动下运动目标物的跟踪

在实时跟踪运动目标式，搭载摄像头的小车在移动，同时目标也是移动的，在跟踪过程中，由于环境背景的剧烈变化，对跟踪算法具有极大的挑战性。

(1) 室内跟踪



(2) 室外跟踪



5.2 系统测试结果

5.3 本章小结

第六章 总结与展望

6.1 工作总结

智能视频监控技术作为数字图像处理与计算机视觉领域中一个涉及多学科的研究前沿，具有非常重要的科学意义和广阔的应用前景。本文对智能视频监控系统中的运动目标检测进行了初步的探究。首先阐述了智能视频监控系统的发展历史与现状，分析了智能监控系统应用前景。之后针对运动目标检测，说明了视频图像预处理的必要性，图像的预处理可以有效地消除噪声，改善图像质量，使图像清晰且易于计算机识别，因此对运动目标识别的正确性提供了一定的保证。噪声处理算法采用了中值滤波的方法，中值滤波是一种非线性滤波方法，它不但对图像噪声消除非常有效，而且能够较好地保护图像边缘信息，因此本文用它来对视频图像进行预处理，并且取得了理想的效果，为目标的检测提供了良好的基础。在分析和比较了几种常用的运动目标检测的方法之后，选用了背景差分法，因为这个方法可以直接得到运动对象的位置、大小、形状等信息，能够完整准确地提取运动目标的信息，简单实用，易于实现，因此本文采用背景差分法来实现运动目标的识别，并取得了比较理想的效果。

6.2 展望

本文对摄像头目标检测与跟踪方法进行研究和分析，并深入探讨了摄像头之间的目标确认技术和目标跟踪技术。由于时间和实验条件有限，视频跟踪涉及领域广泛，系统的复杂度较高，因此仍然有很多影响系统跟踪效果的关键技术需要进一步的深入研究，这些关键的技术点包括：

(1) 如何得到一个准确单应性矩阵，单应性矩阵越准确，摄像头目标之间的确认精度越高。现有的单应性矩阵的计算方法是匹配两幅图像的特征点，但是由于存在噪声干扰，以及监控环境的不断变化，很难得到一个准确的单应性矩阵，因此，提高单应性矩阵的计算精度是一

个需要重点研究的问题。

（2）设计一个适应动态和异构的摄像头跟踪系统的跟踪算法。由于不同的跟踪任务需要不同类型的摄像头，而且在不同的时刻，可能有摄像头退出或加入，所以整个跟踪系统可能由不同类型的摄像头动态的组成。因此，需要有一个跟踪算法能够适应这种多摄像头网络

（3）在异构的多摄像头网络中的数据融合也是一个非常具有挑战性的课题。由于不同类型的摄像头得到的数据类型不一样，每个摄像头得到的目标特征也不相同，所以需要设计一种新的数据融合算法来有效地融合来自不同摄像头的的数据。

致谢

参考文献

- [1] N. Friedman, S. Russell. Image segmentation in video sequences: a probabilistic approach. In: Proc. 13th Conf. Uncertainty in Artificial Intelligence, 1997. 1–3
- [2] D. Koller, J. Weber, T. Huang, et al. Toward robust automatic traffic scene analysis in real-time. In: Proc. Int. Conf. Pattern Recognition. Israel: 1994. 126–131
- [3] M. Köhle, D. Merkl, and J. Kastner. Clinical gait analysis by neural networks: Issues and experiences. In: Proc. IEEE Symp. Computer-Based Medical Systems, 1997. 138–143
- [4] H. Z. Sun, T. Feng, and T. N. Tan. Robust extraction of moving objects from image sequences. In: Proc. Asian Conf. Computer Vision. Taiwan: R.O.C., 2000. 961–964
- [5] W. E. L. Grimson, C. Stauffer, R. Romano, et al. Using adaptive tracking to classify and monitor activities in a site. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition. Santa Barbara, CA, 1998. 22–31
- [6] I. Haritaoglu, D. Harwood, and L. S. Davis. W 4 : Real-time surveillance of people and their activities. IEEE Trans. Pattern Anal. Machine Intell, Aug. 2000, vol. 22: 809–830
- [7] S. McKenna, S. Jabri, Z. Duric, et al. Tracking groups of people. Comput. Vis. Image Understanding, 2000, vol. 80: 42–56
- [8] C. Stauffer, W. Grimson. Adaptive background mixture models for real-time tracking. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1999. 246–252
- [9] A. J. Lipton, H. Fujiyoshi, and R. S. Patil. Moving target classification and tracking from real-time video. In: Proc. IEEE Workshop Applications of Computer Vision, 1998. 8–14
- [10] D. Meyer, J. Denzler, and H. Niemann. Model based extraction of articulated objects in image sequences for gait analysis. In: Proc. IEEE Int. Conf. Image Processing, 1998. 78–81
- [11] D. Meyer, J. Psl, and H. Niemann. Gait classification with HMM's for trajectories of body parts extracted by mixture densities. In: Proc. British Machine Vision Conf., 1998. 459–468
- [12] O. Javed, M. Shah. Tracking and object classification for automated surveillance. In: Proc. European Conf. Computer Vision, 2002. 343–357
- [13] S. Funiak, C.E. Guestrin, M. Paskin, et al. Distributed inference in dynamical systems. Advances in Neural Information Processing Systems, MIT Press, 2006, vol. 19
- [14] S.J. Krotosky, M.M.Trivedi. Mutual information based registration of multimodal stereo videos for person tracking. Computer Vision and Image Understanding, 2007 ,106 (2–3) : 270–287.
- [15] I. Amundson, M. Kushwaha, B. Kusy, et al. Time synchronization for multi-modal target tracking in heterogeneous sensor networks. In: Workshop on Networked Distributed Systems for Sensing and Control, 2007

- [16] Y. Bar-Shalom. Extension of the probabilistic data association filter to multitarget tracking. In: Proceeding of the Fifth Symposium on Nonlinear Estimation, 1974
- [17] A. Dore, A. Calbi, L. Marcenaro, et al. Multimodal cognitive system for immersive user interaction. In: ACM/ICST First International Conference on Immersive Telecommunications, 2007
- [18] Omar Javed, Khurram Shafique, Mubarak Shah. Appearance Modeling for Tracking in Multiple Non-Overlapping Cameras. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. 26-33
- [19] S. Khan, O. Javed, and M. Shah. Tracking in Uncalibrated Cameras with Overlapping Field of View. In: Proc. Performance Evaluation of Tracking and Surveillance PETS, Dec. 2001.
- [20] R. Hartley, A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, 2000. 125-126
- [21] Z. Zhang. Determining the epipolar geometry and its uncertainty—a review. The International Journal of Computer Vision, 1998, 27 (2): 161–195.
- [22] K. Mikolajczyk, C. Schmid. Indexing based on scale invariant interest points. In: Proceedings of the IEEE International Conference on Computer Vision, 2001.
- [23] Thaler. M., Mö rzinger. R. Automatic Inter-image Homography Estimation from Person Detections. In: The Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2010.
- [24] P.T. Baker, Y. Aloimonos. Calibration of a multicamera network, In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. 7–72
- [25] S. Funiak, C. Guestrin, M. Paskin, et al. Distributed localization of networked cameras. In: Proceedings of the Fifth International Conference on Information Processing in Sensor Networks, 2006
- [26] R. Ghrist, A. Muhammad. Coverage and hole-detection in sensor networks via homology. In: Proceedings of the Fourth International Symposium on Information Processing in Sensor Networks, 2005.
- [27] E. Horster, R. Lienhart. Calibrating and optimizing poses of visual sensors in distributed platforms. ACM Multimedia Systems, 2006, 12 (3): 195–210
- [28] S. Ram, K.R. Ramakrishnan, P.K. Atrey, et al. A design methodology for selection and placement of sensors in multimedia surveillance systems. In: Proceedings of the 4th ACM International Workshop on Video Surveillance and Sensor Networks, 2006
- [29] B. Leibe, A. Leonardis, B. Schiele. Robust object detection with interleaved categorization and segmentation. International Journal of Computer Vision, 2008, 77 (1–3): 259–289
- [30] P. Viola, M.J. Jones, D. Snow. Detecting pedestrians using patterns of motion and appearance. In: Proceedings of the IEEE International Conference on Computer Vision, 2003
- [31] M.V. den Bergh, E. Koller-Meier, L.V. Gool. Fast body posture estimation using volumetric features. In: IEEE Visual Motion Computing, 2008
- [32] R. Kehl, M. Bray, L.V. Gool. Full body tracking from multiple views using stochastic sampling. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005
- [33] K. Bernardin, T. Gehrig, R. Stiefelhausen. Multi- and single view multiperson tracking for smart room environments. In: Proceedings of the Workshop on Classification of Events, Actions and Relations, 2006
- [34] K. Bernardin, R. Stiefelhausen. Audio-visual multi-person tracking and identification for smart

- environments. In: Proceedings of the fifteenth International Conference on Multimedia, 2007
- [35] A. Prati, S. Calderara, R. Cucchiara. Using circular statistics for trajectory shape analysis. In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2008
- [36] R.Vezzani, R. Cucchiara. ViSOR: Video surveillance on-line repository for annotation retrieval. In: Proceedings of the IEEE International Conference on Multimedia and Expo, 2008
- [37] S. Dockstader, A.Tekalp. Multiple camera fusion for multi-object tracking. In: Proceedings of the IEEE International Workshop on Multi-Object Tracking, 2001
- [38] J. Krumm, S. Harris, B. Meyers, et al. Multi-camera multi-person tracking for easy living. In: Proceedings of the IEEE Workshop on Visual Surveillance, 2000
- [39] P. Kumar, A. Mittal, P. Kumar. Study of robust and intelligent surveillance in visible and multi-modal framework. In: Informatica , 2008. 63–77
- [40] W. Hu, T. Tan, L. Wang, et al. A Survey on Visual Surveillance of Object Motion and Behaviors. In: IEEE Transactions Systems, Man and Cybernetics, 2004, 34: 334–352
- [41] Cutler R, Davis L. View-based detection. In: Proceedings of International Conference on Pattern Recognition, 1998: 495-500
- [42] Cucchiara R, PiccardiM, PratiA. Detecting moving objects, ghosts, and shadows in video streams. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(10): 1337-1342
- [43] Toyama K, Krumm J, Brumitt B, et al. Wallflower: Principles and practice of background maintenance. In: Proceedings of International Conference on Computer Vision, 1999. 255-261
- [44] C.Staffer, W.E.L.Grimson. Learning patterns of activity using real-time tracking. IEEE Trans. PAMI, 2000, 22(8): 747-757
- [46] A.Murat Tekalp. Digital video processing. Tsinghua University Press, 1998
- [47] Horn B, Schunch B. Detemining optical flow. Artificial Intelligence, 1981, 17: 185-203
- [48] C. Ridder, O. Munkelt, and H. Kirchner. Adaptive background estimation and foreground detection using Kalman-filtering. In: Proc. Int. Conf. Recent Advances in Mechatronics, 1995. 193–199
- [49] D. Comaniciu, V. Ramesh, and P. Meer. A Real-time tracking of non-rigid objects using mean-shift. In: Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit, 2000. 142–149
- [50] ChengY. Mean Shift, mode seeking, and clustering. IEEE Transactions on Patern Analysis and Machine Intelligence, 1995, 17(8)
- [51] Comaniciu D, Meer P. MeanShfit analysis and application. In: Proceedings of the Seventh IEEE Interactional Conference, Computer Vision, 1999
- [52] A. Doucet, N. D. Freitas, and N. Gordon. Sequential Monte Carlo Methods in Practice. New York: Springer-Verlag, 2001
- [53] M. Isard, A. Blake. BIcondensation: Unifying low-level and high-level tracking in a stochastic framework. In Proc. Eur. Conf. Comput. Vision, 1998. 767–781
- [54] J. S. Liu, R. Chen. BSequential monte carlo for dynamical systems. Amer. Statist. Assoc., 1998, vol. 93: 1031–1041
- [55] S. K. Zhou, R. Chellappa, and B. Moghaddam. Visual tracking and recognition using appearance-adaptive models in particle filters. IEEE Trans. Image Process., 2004, vol. 11: 1434–1456
- [57] S.M. Khan, M. Shah. A multiview approach to tracking people in crowded scenes using a

planar homography constraint. In: ECCV, 2006. 133–146

[58] Du. W., Piater. J. Multi-camera people tracking by collaborative particle filters and principal axis-based integration. In: ACCV. Heidelberg: Springer, 2007. 365–374

[59] Pham. N., Huang. W., Ong. S. Probability hypothesis density approach for multi-camera multi-object tracking. In: ACCV. Heidelberg: Springer, 2007. 875–884

[60] Khan. S., Shah. M. Consistent labeling of tracked objects in multiple cameras with overlapping fields of view. *Transaction on Pattern Analysis and Machine Intelligence*, 2003, 25: 1355–1360

[61] Alahi, A., Boursier, Y., Jacquesy, et al. Sport players detection and tracking with a mixed network of planar and omnidirectional cameras. In: *Proceedings of the International Conference on Distributed Smart Cameras*, 2009

[62] Morariu, V., Camps, O. Modeling correspondences for multi-camera tracking using nonlinear manifold learning and target dynamics. In: *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, 2006

[63] Khan, S., Shah, M. Tracking multiple occluding people by localizing on multiple scene planes. *Trans. on Pattern Analysis and Machine Intelligence*, 2009, 31: 505–519

[64] Kayumbi, G., Cavallaro, A. Multiview trajectory mapping using homography with lens distortion correction. *EURASIP Journal on Image and Video Processing*, 2008

[65] Kim, K., Davis, L.S. Multi-camera tracking and segmentation of occluded people on ground plane using search-guided particle filtering. In: ECCV. Heidelberg: Springer, 2006. 98–109

[66] Del Rincon, J.M., Herrero-Jaraba, et al. Automatic left luggage detection and tracking using multi-camera ukf. In: *Proceedings of the International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, 2006. 51–58

[67] Focken, D., Stiefelhagen, R. Towards vision-based 3-d people tracking in a smart room. In: *Proceedings of the International Conference on Multimodal Interfaces*, 2002

[68] Mittal, A., Davis, L.S. A multi-view approach to segmenting and tracking people in a cluttered scene using region-based stereo. In: ECCV. Heidelberg: Springer, 2002. 18–33

[69] Eshel, R., Moses, Y. Homography based multiple camera detection and tracking of people in a dense crowd. In: *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, 2008

[70] Taj, M., Cavallaro, A. Multi-camera track-before-detect. In: *Proceedings of the International Conference on Distributed Smart Cameras*, 2009

[71] Delannay, D., Danhier, N., Vleeschouwer, C.D. Detection and recognition of sports (wo)man from multiple views. In: *Proceedings of the International Conference on Distributed Smart Cameras*, 2009

[72] Khan, S., Shah, M. A multiview approach to tracking people in crowded scenes using a planar homography constraint. In: ECCV. Heidelberg: Springer, 2006. 133–146

[73] Khan, S., Shah, M. Tracking multiple occluding people by localizing on multiple scene planes. *Trans. on Pattern Analysis and Machine Intelligence*, 2009, 31: 505–519

[74] Black, J., Ellis, T., Rosin, P. Multi view image surveillance and tracking. In: *Proceedings of the International Workshop on Motion and Video Computing*, 2002

[75] Anjum, N., Cavallaro, A. Trajectory association and fusion across partially overlapping cameras. In: *Proceedings of the International Conference on Advanced Video and Signal Based Surveillance*, 2009

- [76] J. Kang, I. Cohen, and G. Medioni. Multi-views tracking within and across uncalibrated camera streams. In ACM International workshop on Video surveillance, 2003
- [77] C. Soto, B. Song, and A. K. Roy-Chowdhury. Distributed multi-target tracking in a self-configuring camera network. In: IEEE Int. Conf. on Computer Vision and Pattern Recognition, 2009

附录

