

硕士学位论文

基于机器视觉的工业机器人分拣系统设计

**DESIGN OF INDUSTRIAL ROBOT SORTING
SYSTEM BASED ON MACHINE VISION**

何泽强

哈尔滨工业大学
2016 年 6 月

国内图书分类号：TP391
国际图书分类号：004.93

学校代码：10213
密级：公开

工程硕士学位论文

基于机器视觉的工业机器人分拣系统设计

硕 士 研 究 生：何泽强

导 师：马家辰教授

申 请 学 位：工程硕士

学 科：控制工程

所 在 单 位：信息与电气工程学院

答 辩 日 期：2016 年 6 月

授予学位单位：哈尔滨工业大学

Classified Index: TP391
U.D.C: 004.93

Dissertation for the Master Degree in Engineering

DESIGN OF INDUSTRIAL ROBOT SORTING SYSTEM BASED ON MACHINE VISION

Candidate:	He Zeqiang
Supervisor:	Professor Ma Jiachen
Academic Degree Applied for:	Master of Engineering
Speciality:	Control Engineering
Affiliation:	School of Information and Electrical Engineering
Date of Defence:	June, 2016
Degree-Conferring-Institution:	Harbin Institute of Technology

摘 要

随着工业 4.0 时代的到来，机器人在工业智能化中发挥着越来越重要的作用。在机械零件分拣的生产线上会存在繁重的任务，如果用人工进行分拣，不但作业效率较低而且成本较大。把机器视觉技术应用到机器人上，实现自动分拣机械零件，是工业流水线智能化的趋势。利用机器视觉算法来识别和跟踪目标并且实现快速分拣作业具有重要的理论意义和实用价值。

本文针对基于机器视觉的分拣机器人系统开展工作，进行理论研究和实验验证。分拣系统涉及到两个核心问题，其一是基于机器视觉的图像识别和目标跟踪问题，其二是基于 SCARA 机器人的抓取分拣控制策略问题。待识别的目标是常见的机械零件，包括大小螺栓和螺帽。对智能分拣作业进行了总体方案的设计，详细说明了系统工作原理，通过实验对分拣系统的可行性进行验证。主要内容如下：

首先，对视觉分拣系统的构成进行简单介绍，其中对重要的组成部分即机械手和控制器的各项指标进行说明；对图像采集系统的相机和镜头选取原则做了详细介绍。完成了实验平台的搭建并且给出摄像机标定的具体步骤以及摄像机标定的实验结果。

其次，研究了传送带上机械工件的跟踪问题。分拣的前提是可以识别和跟踪目标，目标跟踪的方法很多，本课题中采用的是基于 CamShift 目标跟踪算法。详细地介绍了该算法的基本原理和主要实现步骤，并且分析了该算法连续跟踪目标工件的基本过程，通过实验证实了跟踪的有效性。

再次，研究了针对三种不同工件的识别方法。由于目标图像的识别是依据其某些特征进行的，本文以图像的矩特征作为识别的重要依据。针对某一工件多角度的图片进行特征提取并结合该工件的标号组成实验样本，依此训练 BP 神经网络。为了提高神经网络的识别准确率，引入 Adaboost 算法对单一神经网络构成的弱分类器进行性能提升。通过对实验数据分析，验证了改进的分类器可较准确的识别目标。

最后，针对机械臂的控制策略问题进行研究。首先利用 D-H 参数法对 SCARA 机器人进行建模，然后分析了机器人运动学的逆解问题，在此基础上，从最优控制角度分析了机械手分拣抓取点的位置和机械手转角之间的关系，得出选取最优抓取点的基本原则，最后利用视觉分拣平台完成分拣作业实验。

关键词：目标工件识别；目标工件跟踪；最优分拣点；SCARA 机器人

Abstract

With the advent of the industrial era 4.0, the robot is playing an increasingly important role in the industrial intelligence. There will be a heavy task in work-piece sorting production line. If manually sorting operations be applied, it will be not only low work efficiency but also higher costs. It is a new trend in the industrial pipeline that the machine vision technology be applied to robotics in order to execute intelligent sorting task. So, using machine vision algorithm to identify and track the target and to achieve fast sorting operations are of great theoretical significance and practical value.

In this paper, the theoretical research and experimental verification based on machine vision sorting robot system will be carried out. The sorting system involves two core issues: one is image recognition and object tracking subsystem based on machine vision, the other is sorting strategy based on SCARA robot. Targets to be identified are the common mechanical work-piece, including bolts and nuts. Overall program of intelligent sorting operations be designed, elaborating the basic operating principle and verifying the feasibility of the sorting system through experiments. The main contents are shown as follow:

Firstly, there is a brief introduction about the components of the visual sorting system including that describing the indicators of SCARA and the controller, analyzing principles of selecting the industrial cameras and optical lens, introducing the specific steps of camera calibration and eventually showing the experimental results of the camera calibration.

Secondly, the tracking problem of the mechanical parts on the conveyor belt is analyzed. Identifying and tracking targets is the premise of the sorting operation. There are many ways to target tracking at present. However, the work-piece tracking algorithm based on CamShift is applied to visual sorting system. The basic principle and main implementation steps of the algorithm are introduced in detail. The basic process of continuous tracking target work-piece is analyzed. Then, the experiment proves that the tracking is effective.

Thirdly, the identification method for three different parts is studied. Because the recognition of the target images is based on some of the features of work-pieces, so the moment feature of image combined with the label of the work-piece is used as an important sample to train BP neural network. Adaboost algorithm is introduced to improve the performance and the recognition accuracy of neural network. The results show that the improved classifier is effective in target recognition based on the experimental data analysis.

Finally, the control strategy of the manipulator is studied. The SCARA robot is modeled by the D-H parameter method and the inverse solution of robot kinematics is analyzed. After that, the relationship between the picking positions and the rotation angles of the manipulator is analyzed based on the optimal control. The basic principles of selecting the optimal grasping points are obtained. Finally, the sorting operation experiment is finished by using the visual sorting platform.

Keywords: object recognition; object tracking; optimal sorting point; SCARA robot

目 录

摘 要	I
ABSTRACT.....	II
第 1 章 绪 论	1
1.1 视觉分拣机器人的研究目的及意义	1
1.2 视觉分拣机器人的国内外研究现状	2
1.3 视觉分拣系统的关键问题概述	5
1.3.1 目标跟踪相关理论概述.....	5
1.3.2 目标识别相关理论概述.....	6
1.3.3 视觉伺服相关理论概述.....	7
1.4 本课题主要研究内容	9
第 2 章 视觉分拣系统构建	10
2.1 视觉分拣系统方案设计	10
2.2 SCARA 机器人系统介绍.....	11
2.2.1 SCARA 机器人	11
2.2.2 SCARA 控制器	12
2.3 图像采集系统部件选型	14
2.3.1 工业相机的选择.....	14
2.3.2 光学镜头的选择.....	16
2.4 摄像机标定实验.....	18
2.5 本章小结	20
第 3 章 基于均值偏移的目标工件跟踪	21
3.1 颜色特征.....	21
3.1.1 颜色模型.....	21
3.1.2 颜色直方图.....	23
3.1.3 直方图后向投影.....	24
3.2 均值偏移算法基本原理	25
3.2.1 核密度估计法.....	25
3.2.2 均值偏移迭代算法.....	26
3.3 基于均值偏移迭代算法的跟踪框架	29
3.3.1 基于 Mean-Shift 目标跟踪算法	30

3.3.2 基于 CamShift 目标跟踪算法	33
3.4 目标工件跟踪实验	35
3.5 本章小结	36
第 4 章 基于神经网络的目标工件识别	37
4.1 不变矩特征	37
4.2 基于 BP 神经网络的目标识别	38
4.2.1 人工神经网络概述	38
4.2.2 BP 神经网络结构	40
4.2.3 BP 神经网络学习算法	41
4.2.4 基于 BP 神经网络的目标识别流程	43
4.3 BP 神经网络识别算法优化	45
4.3.1 Adaboost 基本原理	45
4.3.2 Adaboost 算法实现及算例分析	45
4.3.3 基于 Adaboost 的 BP 神经网络识别算法优化	46
4.4 目标工件识别实验	48
4.5 本章小结	50
第 5 章 基于 SCARA 机器人的目标工件分拣	51
5.1 SCARA 机器人的 D-H 参数建模	51
5.2 SCARA 机器人的运动学逆解分析	55
5.2.1 运动学逆解的求取方法	55
5.2.2 运动学逆解的选取标准	57
5.3 基于最优抓取点的分拣策略	58
5.4 目标工件分拣实验	61
5.5 本章小结	64
结 论	65
参考文献	66
哈尔滨工业大学学位论文原创性声明和使用权限	71
致 谢	72

第1章 绪 论

1.1 视觉分拣机器人的研究目的及意义

本文是研究基于机器视觉的工业机器人分拣系统，所要实现的功能是机器人对流水线上的工件进行识别和分拣作业。因为机器人分拣是视觉伺服的具体应用，所以从机器分拣在工业生产中的应用出发进行研究目的说明，最终落实到对视觉伺服研究意义的阐述。

工业机器人是一种应用范围非常广并且技术附加值较高的数字控制装备。为了避免工人在工业化生产流水线上进行不断地机械式重复劳动，或是期望工人可以从危险的工作环境下解放出来，于是掀起了研究工业机器人的热潮。随着机器人研究领域的不断发展，机器人在工业化生产中的相关技术已经越来越成熟，可以严格保证产品质量，极大地提高了工业化生产效率。此外，机器人行业的竞争激烈，促使其成本正在逐步降低，然而人工劳动力的成本却不同程度地上涨，这导致高性价比的机器人需求量迅速增加，从而进一步推动着该行业的发展^[1,2]。

如今，在工业领域中自动化生产越来越普及，机器人被广泛地应用到各种各样的工业化流水线上，完成种类繁多的工作。图 1-1 是分拣机器人在进行分拣作业。

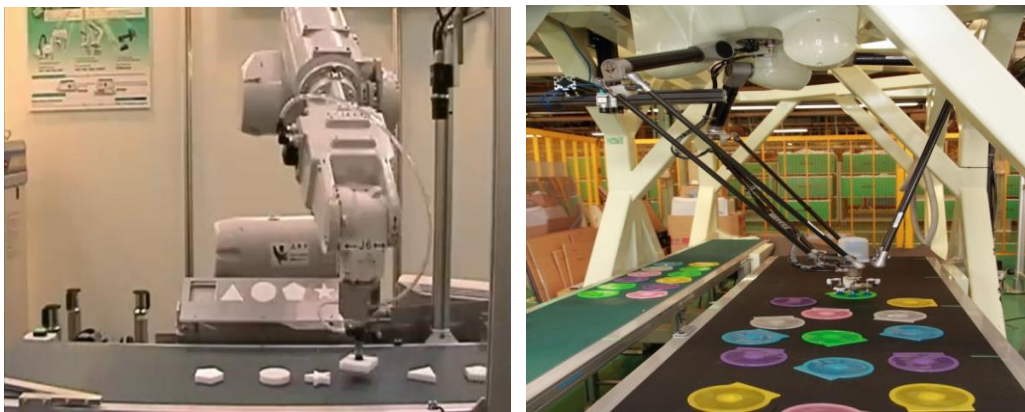


图 1-1 流水线上的视觉分拣机器人

针对工业生产过程中常见的作业方式，已经开发出各种各样的特种机器人来完成相应工作，例如在电子产品领域中的元器件焊接操作，在工业集装箱领域中的搬运操作，在汽车制造业领域中的喷漆操作，在机械零部件生产领域中

的分拣操作等。分工明确化、功能专一化的各类专用工业机器人在减轻工人工作量的同时，也在不断地提高社会生产力水平。

课题研究的机械工件分拣作业是工业生产流程中一项常见的工作任务，其目的是可以把不同种类的目标物体进行正确的分类处理，也可以用于分拣出有质量缺陷的机械工件^[3]。如果用人工操作的方式来完成工业流水线上的分拣作业任务，不论是在分拣速度上还是在分拣成本上，相比于机器人分拣来说并没有明显优势。

虽然工业机器人的特点是可以长时间、高强度的不间断工作，但是采用机器人分拣作业需要克服较多困难。在目前的实际应用中，很多机器人是通过示教方式或是离线编程方式完成一些固定的操作，虽然提高了工业化的总体水平，但是仍然无法适应人们期待的更加智能化的工业生产工作。为了能够实现分拣操作，机器人需要知道待分拣的工件类型以及工件在传送带上运动时的实时位置信息，所以需要给机器人安装图像传感器设备，使其具有获取外部图像信息的能力，增强其自适应性以满足分拣种类繁多的机械零部件的要求^[4,5]。

人类通过视觉可以获取众多的外部信息，若将这一特点应用到机器人上，可以使机器人感知外部信息的变化，并反馈给控制器，其应用领域将更加广泛。可见，视觉伺服对于扩大工业机器人的应用范围起到了重要作用^[6]，视觉分拣操作中涉及到的相关技术正在改变我们传统的工业，解放人类劳动力，提高社会生产力水平，这正是我国实现工业化道路过程中的一次重要机遇。

世界一些工业强国的发展战略报告指出，“机器人技术是提高社会生产力和提升工业化水平的重要保障，该技术和信息技术具有同样重要的地位，不断发展更加智能化的工业机器人技术，不断提高工业信息化水平，是新世纪工业化强国的内在要求，具有极其重要的战略意义”。我国的“863 计划”以及中科院等科研院所一直在做关于机器人的基础理论研究和相应实践研究，其意义在于该技术是衡量一个国家工业和信息化水平的重要标志，是不断提高我国工业自动化领域核心竞争力的重要工具，同样也是有助于加快我国产业转型升级的务实之选^[7]。

1.2 视觉分拣机器人的国内外研究现状

视觉分拣机器人利用机器视觉技术，对传送带上的目标进行识别和跟踪并将相应信息反馈给机器人控制器，导引机械手臂完成自动分拣任务。在 1959 年，美国人约瑟夫发明了第一台工业机器人，从此人类进入了工业机器人时代；

后来生产出利用示教方式仅能机械式工作的机器人，该类型机器人具有可编程功能，所以机器人得到了较为广泛的应用；为了让机器人适应更为复杂的工作环境，进而发展到配备传感器，具有感知外界信息的机器人，这极大地提高了机器人的自主工作能力，其中视觉机器人就是属于此类；目前机器人各项技术更加成熟，被应用到工业、农业、医学、航天等众多领域。

工业生产流水线上目标分拣工作是视觉分拣机器人的主要应用之一，分拣方式主要分为两种，一种是相对静止分拣，另一种是动态分拣，两者的主要区别在于传送带在机器人分拣过程中是连续运动状态还是间歇性运动状态^[8]。在将机器视觉应用到机器人实现分拣工作初期阶段，主要以前一种方式为主，这种分拣方式适用于传送带上物件密度小范围集中分布的情况，在目标物件进入相机视野范围内时，传送带停止运动，机械手臂快速对目标进行分拣。该方式对于电机的磨损是较严重的。对于动态分拣作业来说，传送带保持匀速运动，目标工件源源不断的进入分拣区域，这对整个视觉分拣系统的实时性要求较高，对于动态目标的准确定位提出更高要求。目前视觉分拣机器人工作方式主要是连续动态分拣。

总体来说，视觉分拣机器人在国外的研究已经成熟，应用场合非常广泛；然而国内的相关研究虽然起步较晚，但是经过多年的研究积累，正在缩小和国外著名机器人公司的差距。

近年来，国际上著名的机器人生产商主要有日本的 FANUC、EPSON、瑞典的 ABB 以及德国的 KUKA 等企业，本文通过介绍上述企业的典型机器人来阐述当前本领域的国外研究现状。

FANUC 公司在 1974 年生产了首台机器人，经过 40 年的发展，现在已有多达 240 种各种类型的机器人，适用于分拣作业的典型机器人是 M-1iA 型。它具有重量轻、结构紧凑的特点可以安装的狭小的空间或任意斜面；可在 0.3 秒内完成 25-200-25mm 搬送动作，末端执行器转速可达 $3000^{\circ}/s$ ，可迅速调整工件位姿；内置视觉功能，利用软件开发包 iRVision 可以实现机器人的视觉分拣功能。

EPSON 公司生产的机器人在小型部件组装作业方面具有较强竞争力，机器人的种类很多，其中一款典型机器人是 S5L，该款机器人的机械结构刚性强，功能强大的 RC620+控制器可以使得机器人运动快速且平稳，具有多种配件可供选择，具有功能模块化的趋势。利用系统 VB Guide 5.0 软件可以实现 .Net 连接，具有支持多种编译器功能。基于 Vision Guide 系列视觉软件可以最多控制

16 个相机且支持多种分辨率目标图像。

ABB FlexPicker 机器人擅长拾取操作，运行速度快且抓取精度高，在工件分拣方面具有较大优势。该款机器人支持外部图像采集设备，通过 Cognex 公司生产的配套视觉软件，可以实现机器人的准确拾取操作，特别适合流水线上的对工件的分拣或包装任务。

KUKA 是国际著名的机器人生产供应商，机器人不论在结构设计上还是电机控制上都已经非常成熟，该公司提供大量的应用软件可供用户选择，每一种软件包都对对应着实现不同的功能，例如 KUKA.ConveyorTech 是针对工业流水线而设计的一款应用软件，根据传送带运动情况，实时调整机械手臂的运动状态，具有自适应性的操作特点；KUKA.VisionTech 是另一款跟分拣关联紧密的应用软件，利用该软件可以分析摄像头捕捉到的运动目标信息，识别出目标的位置或是工件上的文字内容，从而引导机器人进行相应的分拣操作。

通过对国际上顶级的机器人公司的产品特点分析，可以得出国外针对分拣机器人的研究现状。目前国外对机器人本体的研究已经相当成熟，具有运动速度快和控制精度高的特点。现在机器人的外围设备是研究的热点，将视觉传感器或触觉传感器和机器人本体相结合，配备以功能强大的应用软件程序，可实现机器人的智能化控制。对于分拣机器人来说，现有的配套视觉软件可以引导机器人进行准确快速的分拣作业，但是关于应用软件涉及到的核心技术属于商业机密，仍然处于严格保密状态。

对于国内研究现状主要概括为以下几个方面^[9-11]：

首先，国内对于各类的机器人需求很大，但是国内知名机器人生产商却并不多，包含周边配件的机器人系统核心技术专利持有量还是比较有限，有些重要零部件需要依赖进口，机器人的控制精度往往难以达到国际领先水平。很多微小企业往往通过购买国外核心零部件，进行组装生产机器人。

其次，为了引进机器人重要技术，企业或是科研院所通常会与国外机器人公司进行技术合作，优势互补，共同研发。其中，安川首钢机器人有限公司就是一个典型例子，该公司通过技术合作研发的 MOTOMANMPK2 型机器人具有较高的性能，分拣速度可达 2 次/s，控制精度可达 0.2mm。

再次，国内很多学者针对分拣作业中的具体问题进行深入研究。例如，针对分拣工作环境恶劣的情况下，如何避免粉尘对摄像机的干扰；针对分拣中图像准确识别问题进行细致研究，从而提高分拣的准确率；针对传送带上目标坐标点的预测及获取进行分析，为能够快速分拣提供必要前提；针对特定分拣目标进行专一化的分拣设计。

最后，攻坚克难，自主研发核心技术。例如以技术转让、技术咨询和技术服务为特点的哈工大机器人集团，集国内众多专家、学者以及科研攻关人员，着力研究机器人领域的重要问题；以中科院沈阳自动化研究所为代表的科研机构一直专注于机器人研究领域，已积累了较多的核心技术，研发的 CR-01 型水下机器人获得国家科技进步一等奖。

1.3 视觉分拣系统的关键问题概述

1.3.1 目标跟踪相关理论概述

本文中涉及到传送上动态目标的分拣工作，对动态目标实时跟踪是实现分拣的前提。由于摄像机获取的是连续的目标图像，可以用视频目标跟踪的方法实现待分拣物体的跟踪。目标跟踪的方法可总结为三个主要类别，包括基于图像特征匹配的跟踪方法，基于运动检测的跟踪方法和基于对比度分析的跟踪方法^[12]。

基于图像特征匹配跟踪方法的基本思想是利用图像中的特征来表示图像中的待跟踪目标，然后不断地寻找相邻两帧图像之间的同一目标特征，从而实现目标的跟踪。可见，跟踪的质量很大程度上取决于对图像特征的选取。Marr 提出过系统地描述图像特征的方法，但是实际特征匹配中，应用最多的是图像的特征点，关于提取特征点的方法也较为成熟，例如 KLT, SIFT, SURT 以及 Harris 等方法^[13-15]。

被跟踪目标的运动往往具有不确定性，基于贝叶斯框架下的跟踪方法适用于对随机运动目标进行跟踪。在该框架下的卡尔曼滤波可以对目标位置进行预测，可以更有利于提高特征匹配的效率，但是该方法仅适用于线性高斯模型。为了扩展跟踪应用范围，在卡尔曼滤波 KF 的基础上，相继提出 EKF 算法、UKF 算法。除此之外，粒子滤波 PF 算法，隐马尔卡夫模型 HMMs 以及动态贝叶斯模型 DBNs 等方法均在目标跟踪方面有所应用^[16-18]。

基于运动检测跟踪方法的基本思想是根据目标运动方向和背景运动方向的不同，突显出目标进而实现对其跟踪。典型的跟踪算法是光流法，光流是描述图像各个像素点的瞬间运动状态，通过计算光流可以实现对目标运动方向的检测。

基于对比度分析的目标跟踪方法是对比目标和背景之间的差异，进而实现目标的识别与跟踪。对视频中每一帧图像进行分块处理，将目标图像和图像块进行对比，确定目标所在位置，依次类推，获取连续位置信息，完成目标跟踪。

本文中涉及到的跟踪算法包括 Mean-Shift 算法和 CamShift 算法。Mean-Shift 跟踪算法具有计算量小、跟踪效果明显以及易于编程实现等特点。但是该算法也有自身的不足，由于搜索窗口的大小在初始阶段被固定下来，这导致该算法的应用受到一定限制。CamShift 算法有效地避免了因搜索框大小无法调整导致跟踪失败的缺点。需要说明的是很多跟踪算法已经不局限于传统的跟踪理论，而是将目标检测等视觉处理算法融合到跟踪算法中来，以提高跟踪性能。

1.3.2 目标识别相关理论概述

图像识别的方法比较多，但是几乎每一种识别方法都是基于图像的某些特征进行识别的。从待识别的目标图像特征来说，主要可以分为颜色特征、形状特征、纹理特征和空间层次关系特征等。颜色特征是根据颜色直方图具有对图像转动等变换不敏感的特点来识别图像，一般应用于对彩色图像的识别。针对形状特征的识别方法，关键是找到图像中目标的形状特征，往往需要对图像中目标的边缘轮廓、重心等参数进行提取来完成对目标的描述及识别。基于纹理特征的图像识别方式，主要依据对图像的颜色信息分布情况以及利用图像中特殊的结构特征来完成对图像的识别。基于空间层次的图像识别方式可以依据目标图像中的层次关系来估计出目标的位置以及方向^[19]。

在图像模式识别中，除了图像特征这个关键识别因素之外，对于图像特征的判别决策方法也是影响着图像识别准确率的关键因素。根据图像的模式特征以及判别决策方法的不同，可以将图像模式识别进行分类。对常见的识别方法简单介绍如下：

句法模式识别算法可以有效地识别较为复杂的图像模式，基本思想是将一种模式进行不断地细分，形成一种由众多子模式构成的树状分枝结构。其中最底层的模式被称作模式基元，支配一定量的基元构成完整模式的法则被称为文法，用一组模式基元以及特定的文法来对某一模式进行描述的语言，被称之为模式描述语言。识别的过程是通过自下而上的识别过程即首先对基元的识别进而利用文法的定义来判定是否属于某一类^[20]。

句法模式识别系统主要是由以下几个部分构成，包括图像预处理、图像基元提取、针对句法的分析以及相应的文法推断。根据预处理完成对模式的分割工作，通过基元提取获得描述模式的一系列的有序基元串。句法分析再依据文法推理得到的信息，来完成对给定模式的识别任务，获得最终的判决结果。为了得到与模式类相适应的文法，句法模式识别需要事先采集足够多的训练模式

样本，经基元提取后把相应的文法推断出来。

统计模式识别算法是目前应用最为广泛也是最成熟的图像识别方法之一，其核心思想是建立决策边界函数以区别不同种类的模式^[21,22]。需要说明的是图像的特征选择及其提取对基于统计方法的模式识别非常重要。因为识别分类的对象并不是目标图像本身，而是通过特征提取之后获得的数据。特征提取后可以用决策函数对目标进行识别分类。整个过程就是由对象空间到特征空间再到决策空间的映射。

近年来模糊数学和神经网络的理论也逐渐开始在图像识别领域中得到广泛应用。模糊模式识别中较为成熟的理论方法是模糊聚类，已经被应用到染色体和白血球的识别，手写文字识别等领域。神经网络是模拟动物大脑的一种数学模型。该网络由大量神经元构成，神经元之间相互联接，最终形成拓扑结构。神经网络具有分布式存储信息的特点，同时具有自组织、自学习的能力。利用图像特征数据作为样本，训练神经网络，利用训练好的网络可以识别目标。需要说明的是，基于神经网络的识别并不能保证识别完全准确^[23]。

1.3.3 视觉伺服相关理论概述

机器人视觉伺服是指对摄像头等图像采集设备获取的图像进行分析处理，依据得到的外界信息作为控制机器人运动的基本依据。机器人视觉伺服系统的分类方法很多，最常见的分类方式是从反馈的信息类型角度，把系统分类为基于位置的视觉伺服系统，基于图像的视觉伺服系统和混合型的视觉伺服系统。这三种视觉伺服系统又分别叫做 3D、2D 和 2.5D 视觉伺服系统^[24]。

基于位置的视觉伺服系统^[25]是对由图像采集设备得到的实时目标图像进行三维空间建模并结合摄像机模型，计算出目标物体在传动带上的空间位置坐标，将这一位置信息和机械手臂末端执行器的位置信息进行对比，得到相应的误差信号，利用控制器使误差减小，机械手末端逐渐靠近目标工件，完成分拣作业。基本控制流程如图 1-2 所示：

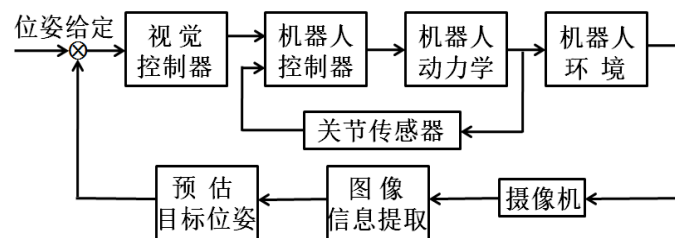


图 1-2 基于位置的视觉伺服系统框图

与上述基于位置的视觉伺服系统不同，基于图像的视觉伺服方法并不需要对目标进行 3D 建模估计坐标位置，但是需要求解用来描述图像特征位置点和机械手末端执行器关系的雅克比矩阵。2D 视觉伺服系统将检测的图像信号与给定图像信号进行对比分析计算图像的误差，产生相应的控制信号进行闭环反馈控制^[26]。基本控制流程如图 1-3 所示：

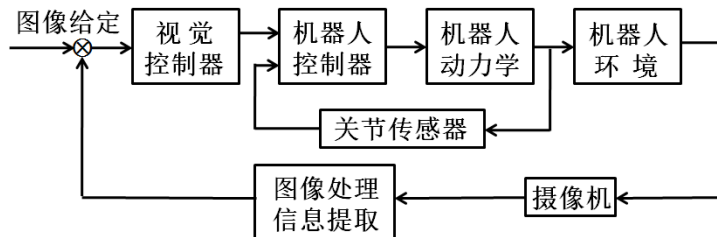


图 1-3 基于图像的视觉伺服系统框图

由于纯粹的三维空间位置信息反馈或是纯粹的二维平面图像信息反馈，都具有一定的局限性。Malis 等综合上述两种视觉伺服系统的各自优点，提出了一种混合型的视觉伺服方式，该方法融合了以三维空间坐标表示的反馈信号和以二维图像空间坐标表示的反馈信号，利用综合的误差信号对系统进行控制，可适度提高系统的稳定性和收敛性^[27]。基本控制流程如图 1-4 所示：

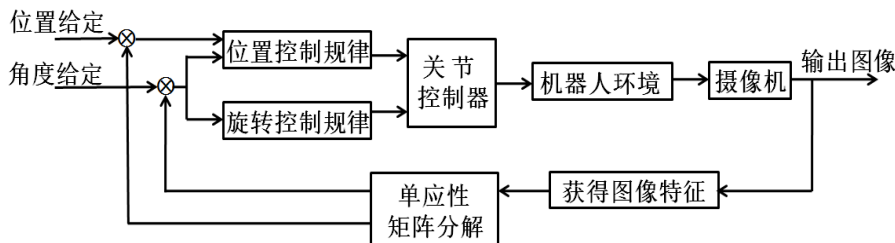


图 1-4 2.5D视觉伺服系统框图

视觉伺服系统实现的是视觉和运动的映射关系，但是往往这种关系较为复杂且一般具有较强的非线性特征，无法用解析式对其明确表示，然而神经网络通过大量的样本训练，就能实现对上述 Visual - Motor关系的描述。

Hashimoto 把提取到的图像特征和给定的对应机器人关节运动信息进行组合，形成一个训练样本，利用大量样本来训练 BP 神经网络，最终可以利用该网络实现对机器人关节角度的控制^[28]；Wells 仍然利用神经网络建立图像特征和机械手运动之间的关系，但是他将重点放在了图像特征的选取上。利用更具抗干扰的全局特征构成样本集合对神经网络进行训练^[29]。

本文中的机器人分拣运动实质上是根据摄像头采集的连续图像，识别待分拣的目标类型并且跟踪待分拣的目标位置，将上述信息传递给机器人控制器，完成对机械手的关节控制，进而实现分拣操作。

1.4 本课题主要研究内容

本课题研究的是基于机器视觉的工业机器人分拣系统设计。首先给出分拣系统的总体设计方案以及各个子系统的实验平台组成；然后构建神经网络分类器对目标机械工件进行识别并且利用均值偏移迭代算法对运动目标进行跟踪；最后阐述了 SCARA 机器人分拣策略并且完成分拣实验。全文分为五章，内容简介如下：

第 1 章是绪论。简述了视觉分拣机器人的研究目的及意义，介绍了本课题的国内外研究现状，概括了本课题所涉及到的相关技术，包括目标识别、目标跟踪和视觉伺服等。

第 2 章是视觉分拣系统构建。设计了基于机器视觉的工业机器人分拣系统的总体方案并说明其基本工作原理。详细地介绍了实验平台的主要组成部分，包括基于 SCARA 机器人的分拣抓取子系统和图像获取分析子系统。说明了工业相机和光学镜头的选取原则并且完成了基于 HALCON 的摄像机标定实验。

第 3 章是基于均值偏移的目标工件跟踪。首先介绍了目标跟踪算法中涉及到的颜色特征，着重阐述了直方图后向投影的过程。然后详细地推导了均值偏移迭代算法并结合图示进行迭代过程说明，进而给出 Mean-Shift 算法以及 CamShift 算法的基本原理以及实现步骤。最后利用 CamShift 算法完成工件跟踪实验并能够获取目标的实时坐标。

第 4 章是基于神经网络的目标工件识别。首先对目标工件进行图像预处理和特征提取，然后利用图像矩特征完成对神经网络的训练，最后引入 Adaboost 算法对基于神经网络的分类器进行性能优化。详细地阐述了识别算法的基本原理以及具体实现步骤，最后通过工件识别实验证实了改进后的神经网络分类器具有较高的识别准确率。

第 5 章是基于 SCARA 机器人的目标工件分拣。通过标准 D-H 参数法建立了 SCARA 机器人的运动学模型，阐述了运动学逆解的求取方法和选取标准。详细分析了机器人关节转角和末端执行器位置之间的关系，得到基于最优抓取点的分拣策略。最后在视觉分拣实验平台上完成了工件分拣作业，详细描述了实验现象并分析了原因。

第 2 章 视觉分拣系统构建

2.1 视觉分拣系统方案设计

视觉分拣机器系统如图 2-1 所示。不同种类的工件被随机放置在传送带上，图像采集设备获取图像序列，由计算机识别目标类型并提取实时坐标，将上述信息传递到控制器，进而导引机械手完成工件分拣工作。

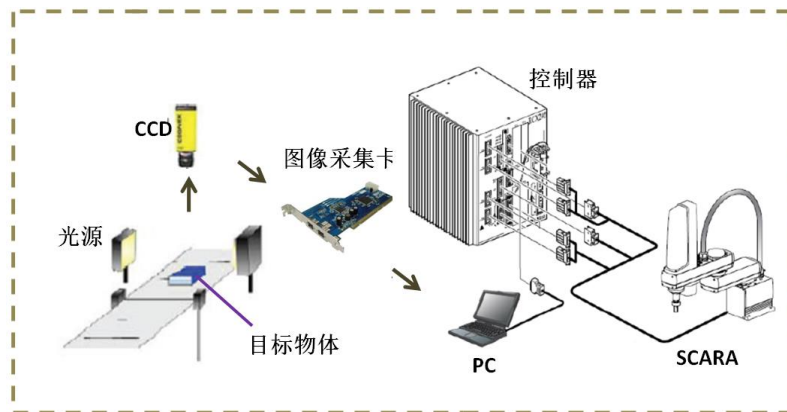


图 2-1 视觉分拣机器人系统示意图

该系统主要由工件传送单元、图像采集单元、信息处理单元以及机械手与控制器单元等部分组成。

工件传送单元由工件传送履带以及三相交流电机构成。本课题中分拣的机械零件是大、小螺栓和螺帽，他们都具有金属光泽，与工件传送履带形成了较大的颜色反差，这有助于实现对目标工件的准确识别与跟踪，为更好的完成分拣抓取任务提供了必要基础。

图像采集单元由工业相机与光学镜头、图像采集卡以及照明设备构成。光源作为分拣平台的照明设备，被固定在传送履带的两侧，用以消除工件自身的阴影。工业相机连续拍摄传送带上的目标工件，并将获取到的信息传递到图像采集卡中，为后续图像处理做好必要准备。

信息处理单元主要由计机构成。该单元对视频序列进行一系列处理，其中包括采用相应视觉算法对目标工件进行识别和跟踪，进而获取到目标种类及其位置坐标，将该信息传递给机械手控制器，进而控制机械手完成分拣任务。该部分是整个系统的核心，具有分析图像信息以及给出控制信息的重要作用，是连接自动控制系统中传感器和执行器的桥梁。

机械手与控制器单元主要是由基于 SCARA 的四自由度机械手臂和相对应的 RC40 控制器构成。机械手与控制器通过专用的电缆进行连接，控制器配有标准的 RS-232 接口，使用该接口可与主机进行数据通信。主机将控制信号传递给控制器，进而控制机械臂的关节运动，在末端执行器上具有电磁吸盘，可实现对待分拣工件的抓取作业。

2.2 SCARA 机器人系统介绍

2.2.1 SCARA 机器人

SCARA (Selective Compliance Assembly Robot Arm) 即选择顺应性装配机器手臂，是一种内置 4 轴伺服电机、谐波减速器和制动器系统的具有四自由度的圆柱坐标型工业机器人。

SCARA 机器人具有四个自由度，包括两个水平关节以及一个既可以垂直移动又可以旋转的连杆。其中第 1、2 自由度是由大机械臂和小机械臂构成的旋转关节，可以使得水平连杆在水平面内进行旋转运动完成平面内快速且精确的定位；第 3 个自由度是具有竖直升降功能的移动关节，可以完成垂直于平面的运动；第 4 个自由度为旋转关节，可以使得末端执行器方便抓取目标。由于 SCARA 机器人在垂直方向上既可以保证较强的刚度又可以保证较高的精度，同时在水平面内可以自由快速的转动，所以非常适合进行平面定位以及完成分拣抓取工件任务，其实物如图 2-2 所示。



图 2-2 SCARA 型四自由度机械手

SCARA 机器人的 4 个关节的执行器部分均使用的是交流伺服电机，可以提供较大的转矩，在机械臂的运动速度和精度上具有良好的执行效果。机器人的关节 1、2 基于谐波减速，而关节 3、4 是基于同步带进行减速。旋转关节

1、2 通过各自转动，可以实现末端执行器在水平面内的移动，关节 3 将电机与滚珠丝杠连接，从而可以实现机械连杆的竖直移动，关节 4 是通过花键将电机和丝杠连接起来，使得丝杠可以完成相应的转动动作。这样通过 4 个关节的相互配合运动，就能够完成物体垂直升降和水平旋转的功能^[30]。该机械手的相关参数见表 2-1。

表 2-1 SCARA 机械手基本规格参数

SCARA 规格	X 轴	Y 轴	Z 轴	R 轴
臂长 (mm)	225	175	150	/
转角范围	$\pm 115^\circ$	$\pm 140^\circ$	/	$\pm 360^\circ$
最大速度	4.0m/s	4.0m/s	1.0m/s	1080°/s
可重复精度	$\pm 0.01\text{mm}$	$\pm 0.01\text{mm}$	$\pm 0.01\text{mm}$	$\pm 0.005^\circ$
电机	200W	100W	100W	100W

在上位机对 SCARA 机器人的控制方面，既有基于计算机的运动控制系统，可以提供准确且灵活的控制效果，也有专供复杂现场使用的且具有抗干扰能力强的运动控制系统，可编程逻辑控制器可以在外部干扰大、工作环境较为恶劣且要求不间断生产的场合下工作。在整个系统的动力方面，该款机器人采用的是交流式伺服电机，可以提供较高运动速度。采用高精密的谐波减速器以及制动器，可以实现运动过程中机器臂的高精度定位和低震动运行。

因为该款机器人的结构非常紧凑，所需安装的空间小，定位精确度高，机械的刚性好，相对同类产品效率较高。故本文中采用的是具有四自由度的 SCARA 机器人作为视觉分拣系统的执行器。

2.2.2 SCARA 控制器

在机器视觉分拣系统中，SCARA 机器人所配套的控制设备是 RCX 系列控制器，本文使用的是 RCX40，该款控制器功能强大，允许外接多种类的执行器，最多支持 8 个不同优先级的任务；具有端口信息并行处理以及中断处理能力，有利于实现多执行器工作效率的最大化；断电后再启动不需要做原点复归，可以有效保存机械手臂的点位数据；有多种类的省配线架构可供选择，支持 CC-Link、DeviceNet、Profibus 和 Ethernet；RCX 系列的控制器具有类似 PLC 的控制效果，在没有 PLC 控制器的情况下，同样可以实现对周围器件的控制功能；可判断 SCARA 型四自由度机器人在移动中所经过的轨迹位置，并输出相应的信息。

该控制器的外观如图 2-3 所示，对应的基本规格参数见表 2-2。



图 2-3 SCARA 机械手控制器

表 2-2 RCX40 控制器基本规格参数

RCX40 基本规格参数	规格说明
外形尺寸	W180 H250 D235 (mm)
控制、马达电源	AC200~230V 10% 以内(50/60Hz)
最大耗电	2500VA
驱动方式	AC 全数字伺服
位置显示单位	脉冲、毫米、度
原点复位方式	增量式、绝对式
程序语言	YAMAHA BASIC
内存容量	364KB (最大 10000 点)
运转方式	PTP 动作、弯曲动作、直线插补、圆弧插补

本文采用 VIP+上位机软件对 RCX40 控制器进行编程，该软件是 YAMAHA RCX 系列控制器的专用支持软件，可利用以太网连接，控制多台控制器，具有强大的图形用户界面，可实现丰富的功能，如对控制器数据进行实时备份；利用 GUI 直接将控制信号传递给 RCX40，从而实现对机械手的实时控制；可以获取机器人的位姿信息，利用自身的函数命令完成机器人的自动控制；支持多种运动方式，包括点对点运动、圆弧式插补运动、直线式插补运动等。

VIP+可以给予 RCX40 控制信号，实现对机器人的三种控制形式，即人工手动控制模式、自动运行模式以及系统模式。VIP+上位机的控制界面见图 2-4。

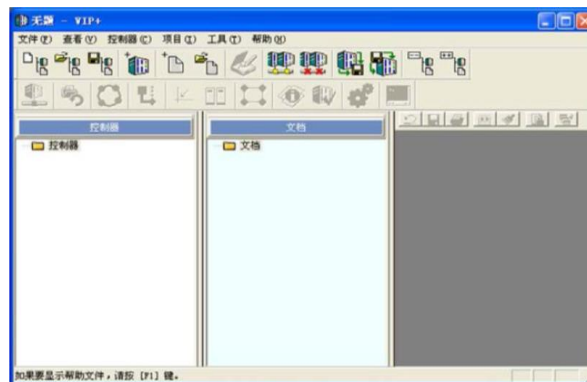


图 2-4 VIP+控制界面

利用该软件对机器人的控制方法如下：首先添加并设置控制器，此时窗口左侧会显示包括程序、坐标系、参数和监控等信息的控制器树状图；其次，点击运行该软件的返回原点按钮，实现位置初始化，机械手后续的运动将以此为参考点进行运动；最后，由于 RCX40 控制器支持点对点运动控制，可基于坐标点完成程序的编写工作，切换到自动控制模式并运行程序，机械手末端将会运行到给定的每一个坐标位置，进而实现对其控制操作。

2.3 图像采集系统部件选型

2.3.1 工业相机的选择

工业摄像机是基于视觉的机器人分拣系统的重要传感器，配合光学镜头共同构成了图像获取单元，完成对分拣工件的图像序列获取工作。在机械手分拣过程中，对于所获得的目标图像质量要求较高，图像的分辨率不但影响着计算机对目标信息的分析处理，而且决定着对目标跟踪的准确性以及定位精度。所以应该选择具有分辨率高、噪声低、抗干扰能力强、对比度高、集速度快等特点的工业摄像机。主要从以下几个角度作为选择摄像机的标准：摄像机的芯片类型、分辨率、帧率等^[31-32]。

基于相机不同类型的感光芯片，摄像机通常被划分为两种类型，即 CCD 摄像机以及 CMOS 摄像机。基于 CCD 的感光芯片是一种半导体成像器件，由大量排列整齐的光敏二极管构成，可将光信号转换为电荷信号，再经外部的采样放大电路以及模数转换电路得到数字型的图像信号。CCD 摄像机具有较多优势，如受强光、磁场和震动的影响较小、成像过程中的畸变较小、成像速度较快。CMOS 摄像机的感光芯片是一种基于互补金属氧化物的半导体成像器件，基于 CMOS 工艺，可将图像采集模块和处理模块集成化，具有体积小、价格低

的特点。虽然 CMOS 摄像机的拍摄图像效果可以满足大部分科研需要，但是本文中需要利用工业相机对传动带上运动的工件进行跟踪并分拣，由于是针对动态目标图像的拍摄，所以选择 CCD 型的摄像机。

依据待分拣工件的精度需要来对摄像机分辨率进行选择。摄像机的像素精度是单方向视野范围大小和摄像机单方向上分辨率的比值，而摄像机单方向的分辨率是单方向视野范围与理论精度的比值。为了提高整个分拣系统的稳定性，一般会采用多倍像素单位对应一个观察精度值，进行目标分辨率的选择。

由于被分拣的工件是处于运动状态，这就对摄像机帧率提出较高要求，以便可以获得更多的目标物体图像，一般来说，帧率不应小于对工件的检测跟踪速度，从而达到对目标的实时跟踪效果，但是分辨率和帧率是呈负相关的，在追求高帧率的同时，也要考虑到摄像机拍摄图片的分辨率大小。

本课题根据分拣螺栓及螺帽的具体要求，选择日本 ARTRAY 公司生产的 ARTCAM-150P5-WOM 型 150 万像素相机，实物如图 2-5 所示，该相机的主要规格参数见表 2-3。



图 2-5 ARTCAM 型工业相机

表 2-3 ARTCAM-150P5-WOM 型工业相机的特性参数

相机指标	特征参数
图像传感器	CCD
分辨率(有效像素)	1360(H)×1024(V)
成像区域	6.47(H)×4.83(V)mm1/2型
像素尺寸	4.65(H)×4.65(H) (μm)
扫描方式	逐行扫描
帧率	20fps
感应度	1.0V/Lux-sec(550nm)
输出方式	USB2.0
镜头卡口	C卡口
接口	USB2.0批量传送
供电方式	DV 5V
功耗	约2W

ARTCAM-150P5-WOM 相机是一款像素约为 150 万的 CCD 型相机，可以满足对目标机械工件的图像获取要求。该相机除了适合于运动目标的图像采集外，经常被应用到静态工件的伤痕检测，医学显微镜成像研究，以及计算机视觉等领域。

2.3.2 光学镜头的选择

镜头最主要的功能就是实现光束调制。在基于机器视觉系统的实际应用中，镜头将待成像目标的光束变换到图像传感器的光敏面上，通过相机将光信号转换为最终的图像信号。镜头质量对于基于机器视觉的分拣系统整体性能会产生重要影响，所以要合理地选择镜头，以便获得最佳效果。

在选择工业相机镜头时，选择的依据为镜头的各项参数，所以先介绍一下镜头的主要性能参数。

1) 焦距 f

在光学系统中，衡量光的聚集或发散的度量方式被称为焦距 f ，通常情况下，指透镜的光心到焦点之间的距离。在相机中，指的是从镜头的中心到相机成像平面之间的距离。镜头的焦距在一定程度上决定了该镜头所拍摄的目标在成像平面上所能够形成的影像大小。假设以同样的距离对某一个目标进行拍摄，那么如果光学镜头的焦距越长，则被摄目标在相机的光电传感器上所成像的放大倍数就越大。

2) 视场角

简单来说视场角是指摄像机镜头可以捕捉到目标图像的最大角度范围。光学仪器的视野范围取决于镜头的视场角的大小，如果镜头的视场角越大，那么能够拍摄的视野就越大，相应的光学倍率就会越小。一般来说，被拍摄的目标若超过这个角度，那么图像信息便不会完全被镜头捕获到。

3) 景深

简单来说，在相机完成聚焦之后，在焦点前后的范围内都可以形成较为清晰的影像，那么就称这个距离范围为景深。利用景深可以拍摄出效果完全不同的图像，例如对于同一场景，若目标的影像清晰而背景的影像模糊，则称景深浅；相反，背景和目标的影像同样清晰，则称为景深深。

调节景深最简单的方法就是改变光圈大小。若果将光圈调大，那么景深就会变浅；若果调小光圈，那么景深就会相应地变深。当然光圈变大，快门速度也相应变快，光圈变小，快门速度则变慢。拍摄传送带上运动的物体时，为了

避免被拍摄影像出现模糊的情况，应选择快速度并且是大光圈。

4) F 数

F 数是描述可以进入镜头内光线的多少的度量， F 数越大，那么可进入镜头内的光线就越多，反之，就越少。 F 数与镜头的有效孔径成正比，与镜头的焦距成反比。由于镜头的焦距一般是固定的，所以可以通过调节镜头的有效孔径，来实现对 F 数的调节。

在基于机器视觉的工业机器人分拣系统中，需要对传送带上运动的工件进行实时的图像获取，根据实验要求，选择的镜头是 OPT-C2514-5M 镜头，实物如图 2-6 所示，该镜头的各项参数如表 2-4 所示。



图 2-6 OPT-C2514-5M 镜头

表 2-4 OPT-C2514-5M 镜头的主要参数

镜头指标	相应参数
放大倍数(\times)	0.068
图像圈 ϕ (mm)	8.0
工作距离(mm)	2786
$F/\#$	8
景深(mm)	145
接口方式	C
长度/直径(mm)	317.0/14.3
重量(g)	2500
1/3"探测器(4.8 \times 3.6mm)	70.6 \times 52.9
1/2.5"探测器(5.70 \times 4.28mm)	83.8 \times 62.9
1/2"探测器(6.4 \times 4.8mm)	94.1 \times 70.6

该镜头有如下特点：成像畸变较小，畸变率可以控制在 0.3% 以内；清晰成像最近距离可达 0.2m 以内；在各种光源条件下均可成像清晰均匀；结构紧凑，外形尺寸小型化，适合安装。

2.4 摄像机标定实验

通过以上小节，介绍了基于机器视觉的工业机器人分拣系统的组成以及各主要模块的器件选择，其中图像采集模块中的摄像机将传送带上的三维空间目标转换为二维的图像信息，依据所获得的目标图像序列实现三维空间上的机械手分拣抓取任务，就需要通过摄像机的内参数和外参数进行坐标转换，从而得到相关的位置坐标信息^[33-34]。为了获得摄像机的内参数和外参数，本文给出了基于 HALCON 软件的摄像机的标定方法。

利用 HALCON 软件中的标定助手对设备进行标定，方法简单且标定准确度高^[35]。为了尽量减少摄像机标定过程中的误差，该软件对标定板有一定要求，例如需要至少 10 张清晰整洁的图像标定板进行摄像机标定。在利用摄像机对标定板进行拍摄时，要保证标定板均匀的出现在摄像机的视场角中。在拍摄标定板时需要获得多幅不同位姿的图片，这就要求相机和标定板之间的拍摄角度进行不断地调整。还要注意光源对标定的影响，保证标定板光照均匀，防止光线过亮或是过暗影响标定结果。

从不同的角度拍摄标定板，一共获得 16 幅标定板图像，如图 2-7 所示，利用以上图像作为对摄像机进行标定的样本。

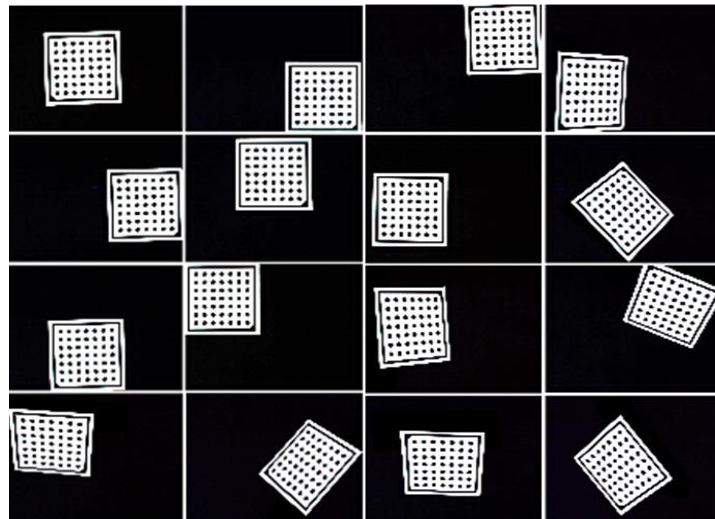


图 2-7 相机标定图像样本

利用上述图像样本进行摄像机标定的步骤如下：

- 1) 使用函数 `gen_caltab()` 生成文件集用以描述标定板信息；
- 2) 使用函数 `read_image()` 获得标定板图像信息；
- 3) 使用函数 `find_caltab()` 将标定板从背景中分离出来；

4) 使用函数 `find_marks_and_pose()` 获取标志点的坐标值;

5) 使用函数 `camera_calibration()` 得到摄像机的标定结果。

基于 HALCON 实现摄像机标定的过程如图 2-8 所示。

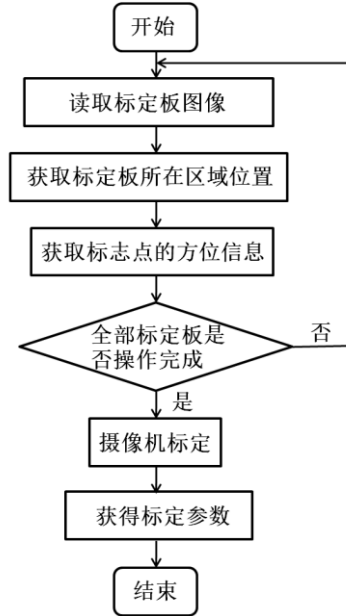


图 2-8 基于 HALCON 摄像机标定流程图

得到相机的内部参数为:

$\text{CamParam} = [4.64, -2342.28, 8.31, 8.30, 276.84, 263.52, 2480, 1852]$

其中, 标定结果中的元素依次代表: 焦距 f 、畸变系数 k 、像元的宽度 S_x 、像元的高度 S_y 、图像中心点的横坐标 C_x 和纵坐标 C_y 和图像的宽与高。

外部参数为:

$\text{CamPose} = [0.112532, -0.1594584, 0.9645815, 342.518, 364.298, 87.6142]$ 。

其中, 标定结果中的元素依次表示为: X、Y、Z 轴上的平移分量 t_x, t_y, t_z 以及旋转角度 α, β, γ 。标定结果如表 2-5 所示。

表 2-5 摄像机标定数据结果

各项参数	针对不同方位图像标定结果
f/mm	4.64
k/mm	-2342.28
C_x/mm	276.84
C_y/mm	263.52
$S_x/\mu\text{m}$	8.31
$S_y/\mu\text{m}$	8.30

2.5 本章小结

本章主要对视觉分拣系统的构建以及实验平台的重要器件选取问题进行阐述。给出视觉分拣系统示意图并简要介绍该系统的基本工作过程。着重介绍了机器人及其控制器的基本信息，详细说明了图像采集设备中关键器件的选取原则。最后给出基于 HALCON 的摄像机标定方法和步骤，最后通过实验获得摄像机标定参数。

第 3 章 基于均值偏移的目标工件跟踪

3.1 颜色特征

3.1.1 颜色模型

1. RGB颜色模型

RGB颜色模型最早出现于彩色电视的颜色模型配置中，通过将红、绿、蓝三种色光按一定比例相互叠加可以产生种类繁多的颜色，而其他颜色混合却不能得到上述三种颜色，故将红、绿、蓝称作三基色，式（3-1）表示用RGB颜色模型构成任意颜色的原理：

$$F = r(R) + g(G) + b(B) \quad (3-1)$$

其中， r, g, b 表示比例系数， F 表示任意一种由红、绿、蓝三基色按一定比例构成的颜色。

为了直观理解RGB颜色模型，可将其形象化为空间立方体模型，定义坐标轴原点代表黑色，红绿蓝颜色轴分别作为笛卡尔坐标系的三个方向轴，在各颜色轴上截取单位距离作为边长，构成的可表示颜色的立方体如图 3-1 所示。

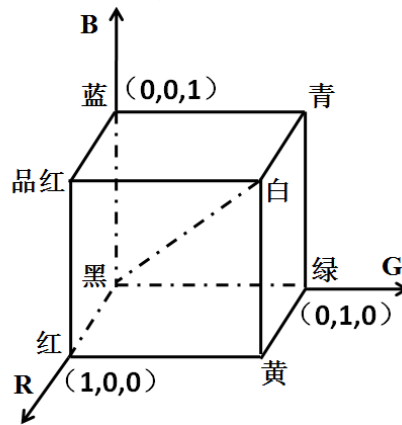


图 3-1 RGB 颜色模型

由于RGB颜色模型结构简单，易于理解，被广泛地应用到数字摄像机，扫描仪等各类图像获取设备中，但是该模型有其局限性。由于各种颜色都存在着一定的亮度信息，这使得该模型易受光照影响，导致后续的颜色处理较为困难，所以在很多关于基于机器视觉的领域都对所获图像的RGB颜色进行转换。若将上述RGB颜色模型进行线性变化可以得到YUV颜色模型，若经非线性变换可得到HSV颜色模型。

2. YUV 颜色模型

YUV 颜色模型是一种颜色描述方法，在欧洲电视系统中使用较多。通过彩色摄像机等图像采集设备获取到颜色信息，对颜色信号进行分离和放大，再经过线性变换电路处理，可以得到由亮度信号 Y 和色差信号 U 和 V 构成的 YUV 颜色模型。该模型最大的特点是颜色的亮度和色度是分离的，这有效地避免了 RGB 颜色模型易受光照强度干扰等问题。本质上，YUV 颜色模型是根据 RGB 颜色模型转换而来，它们之间的变换关系见式 (3-2)。

$$\begin{cases} R = Y + 1.402 \times V \\ G = Y - 0.344 \times U - 0.714 \times V \\ B = Y + 1.722 \times U \end{cases} \quad (3-2)$$

3. HSV 颜色模型

HSV(hue, saturation, value) 颜色模型将图像中的亮度和色彩进行有效的区分，将亮度信息用明度值度量，颜色信息由色度值，饱和度值进行度量。该模型符合人类对色彩的感知形式，因而被广泛使用。

色度是反应了图像中目标颜色的本质属性，描述的是物体反射的光谱波长大小，色度值的范围是 [0, 360]；饱和度可以反应颜色的纯正程度，饱和度值大，表示图像中目标颜色越鲜艳，取值范围是 [0, 1]；明度是描述图像中物体亮暗程度的物理量，与目标物体表面的反射系数有关，反射系数越大，明度值越高，反之越低，取值范围是 [0, 1]。HSV 颜色模型如图 3-2 所示。

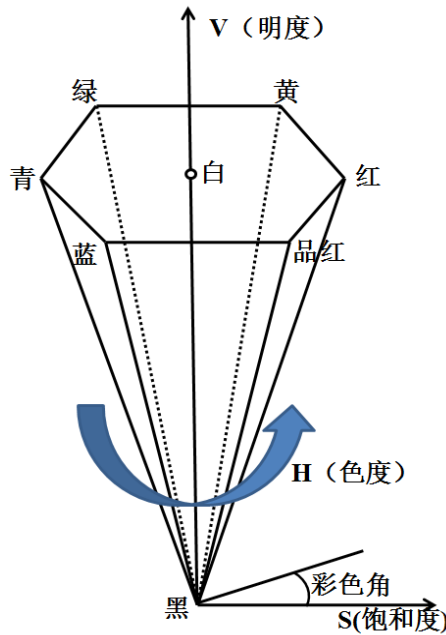


图 3-2 HSV 颜色模型

该模型呈倒锥形，中轴线表示明度，与中轴线垂直的方向表示饱和度，锥面表示色度。当明度和饱和度取最大值时，沿着锥面边缘的顶角上每隔 120° 分别表示三基色红、绿、蓝。同样地，三补色青、品红、黄之间的相对位置也间隔 120° 。

HSV 颜色模型和 RGB 颜色模型之间的转换关系见式 (3-3)、式 (3-4) 和式 (3-5)。

$$H = \begin{cases} \theta & B < G \\ 2\pi - \theta & B > G \end{cases}, \theta = \arccos\left(\frac{\frac{1}{2}[(R-G) + (R-B)]}{[(R-G)^2 + (R-G)(G-B)]^{1/2}}\right) \quad (3-3)$$

$$V = \frac{\max(R, G, B)}{255} \quad (3-4)$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R + G + B)} \quad (3-5)$$

上述两个颜色模型之间的另一种转换方式见式 (3-6)。

$$\begin{bmatrix} V \\ S \times \sin H \\ S \times \cos H \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3-6)$$

3.1.2 颜色直方图

颜色直方图是一种常见的图像特征，表征的是整幅图形中颜色的种类和数量的分布情况。该图像特征通常以颜色分布作为横坐标，以颜色出现的频数作为纵坐标值，构成直方图形式。颜色直方图的计算如式 (3-7) 所示。

$$H(k) = \frac{n_k}{N}, \quad (k = 0, \dots, L-1) \quad (3-7)$$

其中， k 表示颜色直方图的某一个维度， L 表示颜色直方图的 bin 值， n_k 表示第 k 个颜色子空间中所包含的像素点的数量， N 表示图像的所有像素点之和。

以如图 3-3 (a) 所示的通用 lena 测试图片为例，该图片的颜色信息分布如图 3-3 (b) 所示，在颜色直方图的两侧颜色分布较多，中间部分没有颜色分布。颜色直方图可以描述图形的整体颜色特征，这适用于跟踪某些难以分割的图像目标，但并不会得到关于图像目标的具体形状等外部特征信息。颜色直方图与实际图像之间是一对多的对应关系，故该颜色直方图并不能唯一表征 lena 图像。

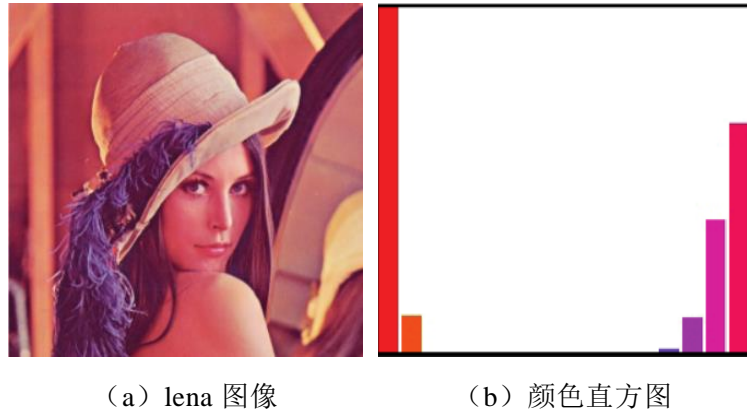


图 3-3 lena 图像的颜色直方图

在对传送带上目标的跟踪过程中,传送带作为背景并没有明显的颜色变化,由于目标工件与背景的颜色反差很大,通过提取目标的颜色直方图,可以很准确的将目标从背景中区分出来,所以本课题的跟踪部分采用颜色直方图作为目标跟踪特征。

3.1.3 直方图后向投影

基于连续自适应均值偏移的目标跟踪算法是以图像的颜色直方图特征作为基础实现跟踪,然而跟踪的对象并不是摄像机获取的基于 **RGB** 颜色空间的目标图像本身,而是跟踪经过一系列图像处理后的基于概率分布的目标工件灰度图。该灰度图可以表征目标的可能位置。

由于很多图像采集设备是基于 **RGB** 颜色空间来获取目标的颜色,为了避免光照强度对目标跟踪的影响,需要将 **RGB** 颜色空间转换为 **HSV** 颜色空间,提取图像色度值作为待分拣目标的颜色直方图模型。

在对目标工件进行跟踪时,会根据颜色直方图模型得到待处理图像区域中每个像素点的概率值,将概率值替换原图像的像素点对应的颜色灰度值,得到跟踪区域的概率分布图,这个过程被称为直方图后向投影。在概率分布图中,概率值越大表示该像素点是目标的可能性越高,反之越低。

为了便于显示,将经过直方图后向投影的概率分布图进行如下处理:把概率分布图中概率为 0 的像素点灰度值赋为 0,同理将概率为 1 的像素点灰度值赋值 255。将概率分布区间和灰度值分布区间进行线性对应,按照上述线性关系可以将每个像素点的概率信息转换为灰度值信息。需要说明的是,最终得到的灰度图像具有概率含义,此时的灰度图像可较为直观地反应出目标可能出现的位置^[36,37]。

目标图像的直方图后向投影如图 3-4 所示。

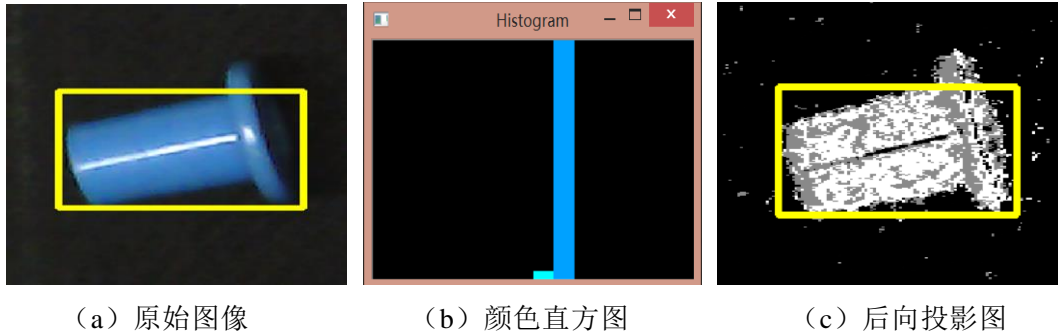


图 3-4 目标图像的后向投影

3.2 均值偏移算法基本原理

3.2.1 核密度估计法

基于均值偏移算法的数学基础是数理统计中的无参密度估计理论。该理论相对于参数密度估计理论具有相对宽泛的使用条件。参数密度估计理论要求被估计的随机变量服从某一个已知的概率密度函数，实际应用中往往难以满足这个条件。然而，无参密度估计法几乎不依赖先验知识，而是根据采集到的数据完成对密度函数值的估计^[38-40]。

常见的无参密度估计法包括最近邻域法、直方图法以及核密度估计法。其中直方图法的基本思想是将数据的取值范围分成多个小区间，将每个区间包含的数据量和总的数据量的比值作为该区间的因变量值，最终形成直方图。核密度估计法在直方图法的基础上，添加具有平滑数据功能的核函数，使其能够对服从任意分布的随机数据进行密度估计。

Mean-Shift 算法是一种基于核密度估计的非参数特征空间分析方法。为此，介绍核密度估计中的重要概念，为后续推导 Mean-Shift 迭代算法做必要准备。

1. 核函数定义

设 X 是一个 d 维空间，向量 x 是该空间中的任意一点，该向量的模记为 $\|x\|^2 = x^T x$ ，实数域记为 R 。若对函数 $K: X \rightarrow R$ ，总存在映射 $k: [0, \infty] \rightarrow R$ ，使

$$K(x) = k(\|x\|^2) \quad (3-8)$$

成立，且满足如下条件：

- (1) k 是非负的，即 $k(x) \geq 0, \forall x \in R^d$
- (2) k 是非增的，即若 $a < b$ ，则 $k(a) \geq k(b)$ ；
- (3) k 分段连续且满足 $\int_0^\infty k(r) dr < \infty$ 。

函数 $K(x)$ 被称为核函数，映射 k 被称为核函数 $K(x)$ 的剖面函数。若将剖面函数 k 的导数取负值后赋给函数 $g(x)$ ，即

$$g(x) = -k'(x) \quad (3-9)$$

则函数 $g(x)$ 被称作核函数 $K(x)$ 的影子核函数。

2. 常见核函数举例

均值偏移算法中常见的核函数是 Epanechnikov 核函数和高斯核函数。

Epanechnikov 核函数的数学表达式为：

$$K(x) = k(\|x\|^2) = \begin{cases} 1 - \|x\|^2, & \|x\| \leq 1 \\ 0, & \|x\| > 1 \end{cases} \quad (3-10)$$

对应的 $k(x)$ 表达式为：

$$k(x) = \begin{cases} 1 - x, & 0 < x \leq 1 \\ 0, & x > 1 \end{cases} \quad (3-11)$$

高斯核函数的数学表达式为：

$$K(x) = k(\|x\|^2) = \begin{cases} \exp(-\frac{\|x\|^2}{2}), & \|x\| \leq 1 \\ 0, & \|x\| > 1 \end{cases} \quad (3-12)$$

对应的 $k(x)$ 表达式为：

$$k(x) = \begin{cases} \exp(-\frac{x}{2}), & 0 < x \leq 1 \\ 0, & x > 1 \end{cases} \quad (3-13)$$

3. Mean-Shift 向量基本形式

设 d 维空间 R^d 中存在着 n 个样本点 $x_i (i=1, 2, \dots, n)$ ，选取其中的任意一点 x_i ，那么定义 Mean-Shift 向量的基本形式为：

$$M_h = \frac{1}{K} \sum_{x_i \in S_k} (x_i - x) \quad (3-14)$$

其中， s_k 为某一个高维球空间， K 为落入 s_k 中的样本点个数。

式 (3-14) 仅描述了 Mean-Shift 向量的形式，其物理含义可以理解为：在某个高维空间中，以一个初始点 x 为球心，作一个高维度球体，在球体中包围的其他样本点 x_i 与 x 之间形成的向量相加的平均值。

3.2.2 均值偏移迭代算法

均值偏移迭代算法是基于 Mean-Shift 跟踪算法基于 CamShift 跟踪算法的理论基础。算法迭代的最终目的是寻找到数据分布密集区域，也就是跟踪并得到图像中最有可能是目标像素的位置。所应用的迭代数据具有随机性，属于数理

统计中的随机变量，往往用概率密度函数描述其分布特征。

对于 d 维空间 R^d 中任意一点 x 的核密度函数定义为：

$$\hat{f}(x) = \frac{1}{nh^d} \sum_i^n K\left(\frac{x-x_i}{h}\right) \quad (3-15)$$

其中， $K(x)$ 表示核函数， h 表示核函数的窗宽。

核函数 $K(x)$ 和剖面函数 $k(x)$ 的关系式为：

$$K(x) = c_{k,d} k(\|x\|^2) \quad (3-16)$$

其中，系数 $c_{k,d}$ 是归一化系数。

将式 (3-16) 代入到核密度函数表达式 (3-15) 中，经过简化处理后，得到的核密度估计公式为：

$$\hat{f}(x) = \frac{c_{k,d}}{nh^d} \sum_i^n k\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \quad (3-17)$$

假设剖面函数 $k(x)$ 的一阶导数在区间 $x \in [0, \infty)$ 上除了少数的有限点外均存在，那么其影子核函数 $g(x)$ 为：

$$g(x) = -k'(x) \quad (3-18)$$

定义核函数 $G(x)$ 为：

$$G(x) = c_{g,d} g(\|x\|^2) \quad (3-19)$$

其中， $c_{g,d}$ 表示归一化系数。

核密度函数的梯度指向随机数据样本分布稠密变化最快的方向，该梯度估计的推导过程为：

$$\begin{aligned} \hat{\nabla} f_K(x) &\equiv \nabla \hat{f}(x) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (x-x_k) k'\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \\ &= \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (x-x_k) g\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \\ &= \frac{2c_{k,d}}{nh^{d+2}} \left[g\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \left[\frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \right] \right] \end{aligned} \quad (3-20)$$

式 (3-20) 中的最后一项因式是两个变量之差，其中一个变量可以理解为所有落入给定高维球中随机样本的坐标平均值，另一个变量可以理解为给定的初始位置坐标点。这种形式可以看成是空间中的某一向量表达式，事实上，该向量即为样本点的 Mean-Shift 向量。

若将 Mean-Shift 向量记为 $M_{h,G}(x)$ ，那么该向量的表达式为：

$$M_{h,G}(x) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \quad (3-21)$$

若利用式 (3-19) 定义的核函数 $G(x)$ 对某一数据样本点 x 进行密度估计，其表达式为：

$$\hat{f}_G(x) = \frac{c_{g,d}}{nh^d} \sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \quad (3-22)$$

将 Mean-Shift 向量表达式 (3-21) 和基于核函数 $G(x)$ 的概率密度估计表达式 (3-22) 代入到核密度梯度估计表达式 (3-20) 中，化简可得：

$$\hat{\nabla} f_K(x) = \hat{f}_G(x) \frac{2}{h^2} \frac{c_{k,d}}{c_{g,d}} M_{h,G}(x) \quad (3-23)$$

整理式 (3-24)，可得 Mean-Shift 表达式如下：

$$M_{h,G}(x) = \frac{h^2 c}{2} \frac{\hat{\nabla} f_K(x)}{\hat{f}_G(x)} \quad (3-24)$$

其中，系数 $c = c_{g,d}/c_{k,d}$ 是核函数 $G(x)$ 以及核函数 $K(x)$ 的归一化系数之商。

式(3-24)说明基于核函数 G 的均值偏移矢量 $M_{h,G}(x)$ 正比于基于核函数 $K(x)$ 的密度梯度估计 $\hat{\nabla} f_{h,k}(x)$ ；基于核函数 $G(x)$ 所推导的样本均值偏移矢量是基于核函数 $K(x)$ 计算的归一化密度函数的梯度；均值偏移向量总是指向随机变量概率密度变化最大的方向^[41-43]。

利用核密度函数梯度估计，可得概率分布的局部极值点。由式 (3-24) 可知，若要使得 $\hat{\nabla} f_K(x) = 0$ 成立，则必有 $M_{h,G}(x) = 0$ 。根据公式 (3-21) 可以得到新的高维球心坐标为：

$$x = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} \quad (3-25)$$

均值偏移算法就是不断地迭代计算 $M_{h,G}(x)$ 值，直到该值小于给定的误差 ε 时停止迭代，否则利用式(3-25)求取高维球心坐标，以该坐标为新的起始点，重新作一个半径为 h 的高维球，继续计算 Mean-Shift 向量，最终运动到概率密度最大的区域。

基于均值偏移算法的迭代过程如图 3-5 所示。

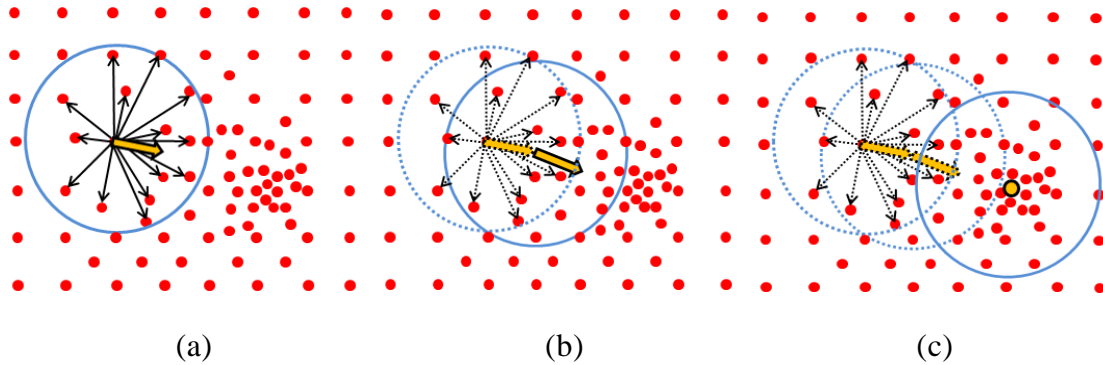


图 3-5 均值偏移算法迭代过程

在 d 维空间 R^d 中任选一个点，以该点为球心， h 为半径作一个高维球。落在这个球内的所有样本点和球心点之间都会产生一个向量，再把这些向量进行矢量相加，得到的结果就是 Mean-Shift 向量，如图 3-5 (a) 所示。

再以 Mean-Shift 向量的终点为球心，继续作高维球。不断地重复以上步骤，高维球就会随着 Mean-Shift 向量不断移动，如图 3-5 (b) 所示。

如此重复下去，Mean-Shift 算法可以收敛到概率密度最大得地方，也就是样本点最稠密的地方，如图 3-5 (c) 所示。

3.3 基于均值偏移迭代算法的跟踪框架

目标跟踪检测是计算机视觉领域的一项基础的核心技术，在机器人分拣系统组成中至关重要。利用目标跟踪算法对摄像机拍摄到的视频序列进行处理，提取目标的位置信息和运动速度信息，将这些信息传递给机械手控制器，才能够引导机械手臂进行分拣作业。

由于摄像机获取图像的过程是一种三维空间对二维空间的映射过程，期间目标信息存在一定程度的缺失，并且传送带上目标工件本身的位置及运动信息是不确定的。除此之外，光照变化、目标被遮挡等环境中的干扰因素都会增加运动目标跟踪的复杂性。在这样背景下，要求实时准确地完成图像跟踪是一项具有挑战性的任务。

1975 年，Fukunaga 和 Hosteler 在研究核函数的密度估计时，首次提出均值偏移迭代算法的原型，用来进行概率密度函数的梯度估计，但是当时并没有得到足够重视。直到二十年后，由 Yizong Cheng 重新定义了核函数和权重系数后，并且成功地将该迭代算法扩展到计算机视觉领域才引起许多学者的关注^[44,45]。Comaniciu 证明了均值偏移算法的收敛性并且将 Mean-Shift 算法成功地应用到

视频跟踪算法中。Bradski 在此基础上提出一种连续自适应均值偏移算法，CamShift 算法具有自适应更新搜索窗口的特点，具有更加广泛地应用。在此之后，很多学者对基于 Mean-Shift 算法和基于 CamShift 算法的跟踪框架进行不断改进和完善。

3.3.1 基于 Mean-Shift 目标跟踪算法

Mean-Shift 算法是一种高效且准确的跟踪算法，且具有很多优势，并不需要对整体区间的概率密度进行估计，也不需要区间中的样本分布情况进行预估假设。该跟踪算法利用核函数加权颜色直方图描述目标，用巴式系数度量目标和候选目标之间的相似性，将搜索相似度量的局部极值问题转换为核密度估计问题，通过均值偏移迭代算法来寻找其局部最大值。迭代的过程就是目标跟踪窗口沿着较高相似度的位置不断地移动的过程，直至算法收敛完成目标跟踪任务^[46,47]。

将从建立目标模型和候选模型、描述模型之间的相似程度、推导目标跟踪算法以及实现目标跟踪具体步骤等四个方面详细介绍基于 Mean-Shift 目标跟踪算法。

1. 直方图模型

图像特征是实现跟踪目标工件的重要依据，由于不同的特征空间对应着不同的目标模型，建立描述待分拣工件的有效目标模型是进行目标跟踪的首要任务。对于 Mean-Shift 跟踪算法而言，目标模型和候选模型本质上是一种离散概率密度函数，该函数是通过统计被跟踪区域像素点的颜色值得到的，可由颜色直方图表示。

给定一幅图像，设被跟踪的目标区域中包含 n 个像素点 $\{x_i\}, i=1,2,\dots,n$ ，若直方图 bin 值取为 m ，那么目标直方图模型可由下式表示：

$$q = \{q_u\}_{u=1,2,\dots,m} \quad (3-26)$$

直方图中数据的分布情况为：

$$q_u = C_h \sum_{i=1}^n k\left(\left\|\frac{x_i}{h}\right\|^2\right) \delta(b(x_i) - u) \quad (3-27)$$

其中， h 表示窗口带宽矩阵，其作用是限定给定图像中被跟踪目标的像素个数； $b(x_i)$ 表示一种量化函数，建立了像素点 x_i 的特征值和相应的 bin 值之间的映射

关系； δ 表示 Kronecker delta 函数； C_h 表示归一化系数，可以保证 $\sum_{u=1}^m q_u = 1$ 。

需要说明的是，基于直方图描述的目标模型往往具有一定的离散特点，由此引入核函数对统计的数据进行平滑处理，使得直方图模型趋于连续且光滑，进而满足均值偏移迭代算法的条件。

同样地，候选模型定义如下：

$$p(y) = \{p_u(y)\}_{u=1,2,\dots,m} \quad (3-28)$$

$$p_u(y) = C_h \sum_{i=1}^n k\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \delta[b(x_i) - u] \quad (3-29)$$

其中， y 表示被跟踪目标图像的中心坐标。其他符号含义可参见目标模型中的相关说明。

2. 相似性度量

建立了描述目标的模型后，需要将目标模型和候选模型进行对比分析。描述模型之间相似度的方法较多，在均值偏移算法中常采用的是 Bhattacharyya 系数法，巴氏系数的定义如下：

$$d(y) = \sqrt{1 - \rho[p(y), q]} \quad (3-30)$$

$$\rho(y) \equiv \rho[p(y), q] = \sum_{u=1}^m p_u(y) q_u \quad (3-31)$$

Bhattacharyya 系数的几何意义是指单位向量之间的内积值。向量夹角的大小表征着相似程度。若夹角为 0 ，表明两个向量指向同一方向，此时最相似；若夹角为 90° ，表明两个向量相互垂直，此时最不相似。如果图像的描述模型之间相似度较低，那么对应的巴氏系数值将会较小，反之，则对应的巴氏系数值将会较大。

3. 目标跟踪算法

图像目标跟踪就是在不断地寻找每一帧图像中和给定目标模型相似度最高的区域，这等价于寻找一个图像块的中心坐标点 y ，使得目标模型和候选模型之间巴氏系数最大。事实上，寻找坐标位置 y 的过程正是基于均值偏移迭代算法实现的。

由于传送带上待分拣目标的运动大多具有平稳连续性，在不考虑速度突变的情况下，初始位置 y_0 处的候选模型 $p(y_0)$ 和下一帧图像区域的候选模型 $p(y)$ 之间的变化不大。利用泰勒公式将函数 $\rho[p(y), q]$ 在初始位置 y_0 处展开，那么 $\rho[p(y), q]$ 近似为：

$$\rho(y) \approx \rho[p(y), q] \approx \rho[p(y_0), q] + [p(y) - p(y_0)]^T \frac{\partial \rho}{\partial p}(y_0) \quad (3-32)$$

将目标模型和候选目标的表达式代入到式 (3-32)，得到如下结果：

$$\rho(y) \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(y_0)q_u} + \frac{1}{2} \sum_{u=1}^m p_u(y) \sqrt{\frac{q_u}{p_u(y_0)}} \quad (3-33)$$

定义中间变量 w_i ，见下式：

$$w_i = \sum_{u=1}^m \sqrt{\frac{q_u}{p_u(y_0)}} \delta(b(x_i) - u) \quad (3-34)$$

联立式 (3-29)、式 (3-33) 和式 (3-34)，可以得到：

$$\rho(y) \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(y_0)q_u} + \frac{C_h}{2} \sum_{u=1}^n w_i k\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \quad (3-35)$$

式(3-35)描述的是目标模型和候选模型之间的相似程度。该式的第一项为常数项，要使得 $\rho(y)$ 取最大值，等价于式 (3-35) 中的第二项取最大值。其中，使函数 $\rho(y)$ 取最大值的 \hat{y} 表示某一图像块的中心位置坐标，并且该坐标点 \hat{y} 是被跟踪目标坐标点的可能性最大。由于梯度 $\nabla \rho(y)$ 描述的是函数 $\rho(y)$ 变化情况，若令梯度 $\nabla \rho(y) = 0$ ，即可求出 \hat{y} 。

梯度 $\nabla \rho(y)$ 的计算过程如下：

$$\begin{aligned} \nabla \rho(y) &\approx C_h \sum_{i=1}^n w_i (y - x_i) k'\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \\ &= C_h \sum_{i=1}^n w_i (x_i - y) g\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \\ &= \left[\frac{\sum_{i=1}^n x_i w_i g\left(\left\|\frac{y-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n w_i g\left(\left\|\frac{y-x_i}{h}\right\|^2\right)} - y \right] \left[C_h \sum_{i=1}^n w_i g\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \right] \end{aligned} \quad (3-36)$$

其中， $g(x) = -k'(x)$ 。式 (3-36) 中的第一个因式被称为 **Mean-Shift** 向量。

为了使 $\nabla \rho(y) = 0$ ，只要令式(3-36)的第一个因式为0即可。利用 **Mean-Shift** 算法得到的最终停止位置为：

$$\hat{y} = \frac{\sum_{i=1}^n x_i w_i g\left(\left\|\frac{y-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n w_i g\left(\left\|\frac{y-x_i}{h}\right\|^2\right)} \quad (3-37)$$

其中， \hat{y} 表示图像序列中目标最可能出现的位置点。

4. 跟踪算法实现

根据选定图像的初始点，计算出相应的 Mean-Shift 向量，待目标图像运动到新的位置时，以 Mean-Shift 向量的终点为新的起点，重复上述过程直到跟踪算法收敛。具体算法如下所示：

- 1) 建立基于颜色特征的目标模型并计算候选直方图向量 $p(y_0)$ ；
- 2) 计算偏移向量，找到目标下一次迭代的新位置 \hat{y} ；
- 3) 计算候选目标在新位置上与目标模板的相似度 $\rho[p(\hat{y}), q]$ ；
- 4) 当 $\rho[p(\hat{y}), q] < \rho[p(y_0), q]$ 成立时，将 $1/2(\hat{y} + y_0)$ 赋值给 \hat{y} ，计算 $\rho[p(\hat{y}), q]$ ；
- 5) 如果 $\|\hat{y} - y_0\| < \varepsilon$ ，停止本帧的跟踪，否则设定 $y_0 \leftarrow \hat{y}$ ，返回步骤 2)。

3.3.2 基于 CamShift 目标跟踪算法

CamShift(Continuously adaptive Mean Shift)算法是在 Mean-Shift 算法基础上改进的，是一种具有连续自适应特点的目标跟踪算法。该算法以工件的颜色作为目标跟踪特征，通过自适应地更改搜索目标窗口的大小，寻找并发现目标，解决了 Mean-Shift 算法无法适应尺寸变化的不足。

Camshift 将获取到的目标图像进行颜色模型转换，构建颜色直方图，并进行直方图后向投影，在得到的基于概率分布的灰度图上，用 Mean-Shift 算法进行迭代计算，将上一帧中的目标窗口的大小和位置作为下一帧 Mean-Shift 算法的初始值，通过多阶矩进行迭代计算，不断地调整目标搜索窗口的大小和位置，直到收敛于跟踪的目标位置。

将从颜色直方图的后向投影、目标搜索窗口的更新以及目标跟踪的具体步骤等三个方面详细介绍基于 CamShift 的目标跟踪算法。

1. 直方图后向投影

直方图后向投影是 CamShift 算法的基础，其基本概念已在前面章节有所叙述。后向投影应用于基于 CamShift 跟踪算法的具体步骤如下：

- 1) 读取图像采集设备获取的传送带目标工件图像，将该图像由 RGB 颜色空间转换为 HSV 颜色空间；
- 2) 给定搜索窗口初始值，在该范围内提取 HSV 颜色空间中的 H 通道颜色数据分量，计算基于色度分量的直方图 $\{\hat{q}_u\}_{u=1, \dots, n}$ ；
- 3) 计算像素点的概率值，并将其线性变换到区间 $[0, 255]$ ，完成后向投影。

$$p_u = \min\left\{\frac{255\hat{q}_u}{\max(\hat{q}_u)}, 255\right\}_{u=1, \dots, m} \quad (3-38)$$

2. 搜索窗口的更新

CamShift 算法相比于 Mean-Shift 算法较大的改进在于搜索窗口的自适应更新，结合跟踪工件几何外形的特征，将跟踪框定义为矩形。通过计算矩形框的长度 l 和宽度 w 及其质心位置 (x_c, y_c) ，实现搜索窗口的不断更新。相关计算公式总结如下：

计算 0 阶矩：

$$M_{00} = \sum_i \sum_j f(i, j) \quad (3-39)$$

计算 1 阶矩：

$$\begin{cases} M_{01} = \sum_i \sum_j j f(i, j) \\ M_{10} = \sum_i \sum_j i f(i, j) \end{cases} \quad (3-40)$$

计算 2 阶矩：

$$\begin{cases} M_{20} = \sum_i \sum_j i^2 f(i, j) \\ M_{02} = \sum_i \sum_j j^2 f(i, j) \\ M_{11} = \sum_i \sum_j ij f(i, j) \end{cases} \quad (3-41)$$

利用图像的上述多阶矩，可以计算关于矩形搜索框的基本参数。搜索窗口的质心坐标为：

$$\begin{cases} x_c = \frac{M_{10}}{M_{00}} \\ y_c = \frac{M_{01}}{M_{00}} \end{cases} \quad (3-42)$$

搜索窗口的质心坐标为：

$$\begin{cases} l = \sqrt{\frac{(a+b) + \sqrt{b^2 + (a-c)^2}}{2}} \\ w = \sqrt{\frac{(a+b) - \sqrt{b^2 + (a-c)^2}}{2}} \end{cases} \quad (3-43)$$

其中， $a = \frac{M_{20}}{M_{00}} - x_c^2$ ， $b = 2\left(\frac{M_{11}}{M_{00}} - x_c y_c\right)$ ， $c = \frac{M_{02}}{M_{00}} - y_c^2$ 。

3. CamShift 算法实现

CamShift 跟踪算法的主要步骤如下：

- 1) 对所获取图像的给定区域进行直方图后向投影；
- 2) 利用 Mean-Shift 迭代算法寻找目标区域的中心坐标，同时计算搜索框位置和尺寸；
- 3) 若跟踪算法收敛，那么将目前搜索框的信息作为下一帧跟踪的初始值，在此基础上继续跟踪，否则返回步骤 2)。

CamShift 跟踪算法的实现流程如图 3-6 所示。

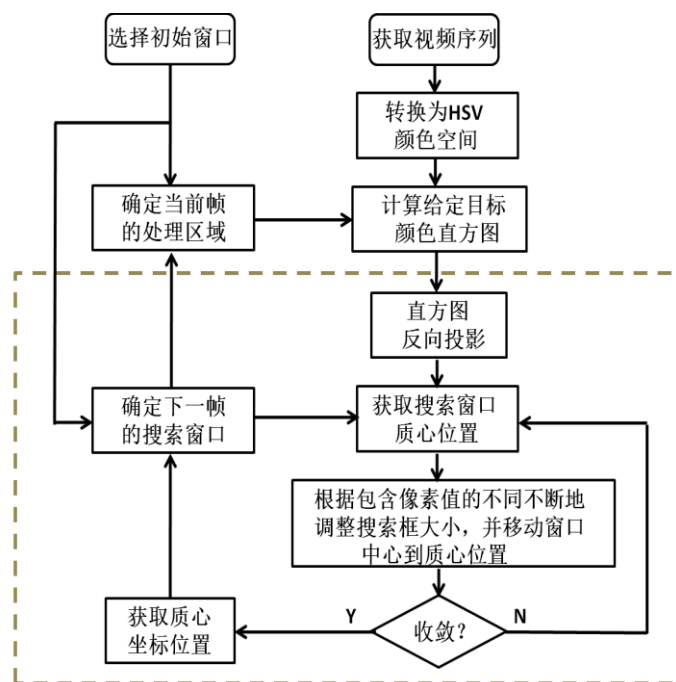


图 3-6 CamShift 算法流程图

3.4 目标工件跟踪实验

在跟踪实验中，传送带上目标工件为小螺栓、大螺栓以及螺帽，通过图像采集设备获取传送带上运动工件的视频序列 2750 帧，采用基于 CamShift 的算法对其进行跟踪，被跟踪的目标用黄色框图进行标注，截取部分跟踪视频片段如图 3-7 所示。

在采用基于 Mean-Shift 算法进行目标跟踪时，由于搜索窗口大小相对固定，这对连续地跟踪大小和形状不同的工件有较大影响。实验发现，若初始窗口包含的是尺寸较小的工件，那么基于 Mean-Shift 的跟踪算法就难以识别并连续地跟踪尺寸较大的工件，常常导致跟踪失败。

采用 CamShift 算法进行目标跟踪可以有效地解决上述问题，虽然每个工件的形状大小不同，但是他们的颜色特征非常相似，当机械手分拣完成前一个工件时，搜索框会在原来位置的基础上，自适应地寻找下一个颜色特征相近的目标，并很快收敛到新的目标位置处，正是由于 CamShift 跟踪算法的自身特点，才能实现对目标的连续跟踪以及获取相对准确的坐标信息。

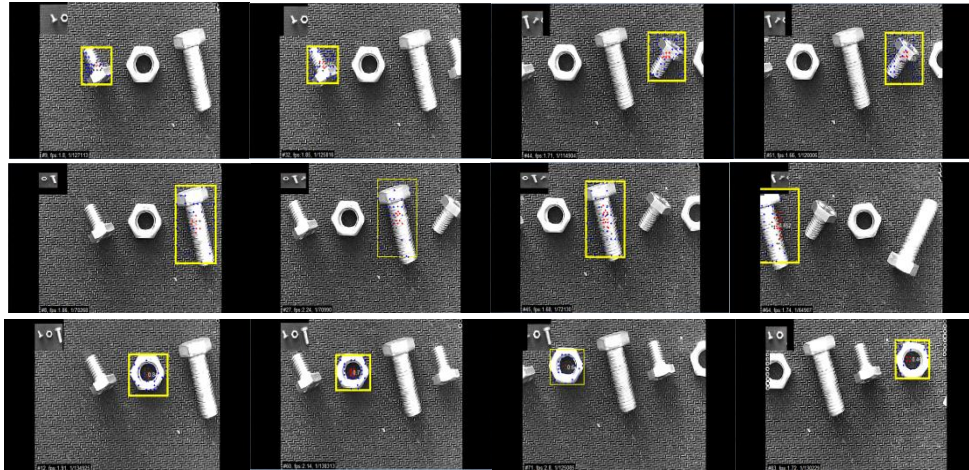


图 3-7 传送带上目标工件的跟踪

由于工件和传送带之间的颜色反差较明显，摄像机视野内的图像相对来说并不复杂，基于 CamShift 算法总体上可以较准确地跟踪目标工件，跟踪效果好，能够满足本课题对跟踪的基本要求。需要说明的是，通过实验发现，当机械手没有及时分拣工件，使其运动到摄像机视野边缘时，容易使 CamShift 跟踪算法发散。

3.5 本章小结

本章介绍了视觉分拣系统中的目标跟踪算法，目的是将待分拣目标的实时坐标传递给控制器，以便完成对机械手的导引分拣任务。首先对目前常见的颜色模型特点及其相互转换方法，表征目标区域颜色分布特征的颜色直方图以及直方图后向投影基本过程等三个方面做了基本介绍。对均值偏移算法的基本原理进行了详细推导。详细阐述了基于均值偏移的两种典型跟踪算法基本原理，并对 Mean-Shift 和 CamShift 跟踪算法的实现方法进行了说明，结合本课题具体情况，分析了上述两种方法在视觉分拣平台上的优缺点。最终采用基于 CamShift 跟踪算法完成对目标工件的跟踪，为后续的视觉分拣实验奠定了基础。

第 4 章 基于神经网络的目标工件识别

4.1 不变矩特征

为实现对目标工件的分拣作业，除了完成跟踪目标并获取实时坐标的任务外，还应识别出目标的具体类型。本文中涉及到的待分拣目标是大小螺栓和螺母。摄像机采集图像的过程是将三维空间中的目标转换到二维平面中，期间会产生目标信息的部分丢失，又由于随传送带运动的工件与摄像机之间存在着拍摄角度的原因以及工件被随机地放置在传送带上导致其位姿具有多样性，以至于同一个工件在图像序列中具有不同的形态表现。难以用单一的图像特征准确地识别目标，故本文中采用神经网络来对上述三种工件进行目标识别。

由于 CamShift 跟踪算法可以用矩形框标示出被跟踪目标，基于这个特点将包含目标的矩形框进行归一化处理为大小 10×10 的图像块，并作为神经网络学习的一类训练样本；CamShift 跟踪过程中已将直方图进行后向投影，得到基于概率分布的灰度图，以此求取跟踪框内的目标不变矩特征，作为神经网络学习的另一类训练数据。最终将两类训练数据组合起来训练神经网络，以获取更为准确的识别效果，为此先介绍图像的矩特征概念。

在统计学中常使用矩来描述随机变量的分布情况。因为一幅灰度图像可以用二维灰度密度函数来表示，将矩特征引入到图像识别中，以完成对灰度图像的描述。

一幅 $M \times N$ 的数字图像 $f(i, j)$ ，其 $p+q$ 阶几何矩 m_{pq} 为：

$$m_{pq} = \sum_{i=1}^M \sum_{j=1}^N i^p j^q f(i, j) \quad (4-1)$$

数字图像 $f(i, j)$ 的 $p+q$ 阶中心矩 μ_{pq} 为：

$$\mu_{pq} = \sum_{i=1}^M \sum_{j=1}^N (i - \bar{i})^p (j - \bar{j})^q f(i, j) \quad (4-2)$$

其中， $\bar{i} = \frac{m_{10}}{m_{00}}$ ， $\bar{j} = \frac{m_{01}}{m_{00}}$ 。

为了避免比例变化的影响，规格化中心矩为：

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad (4-3)$$

其中， $\gamma = \frac{p+q}{2} + 1$ 。

将二阶以及三阶规格化的中心矩进行非线性组合，可以推导出七个量值，它们对于对图像平移变化、旋转变换以及按比例缩放变化不敏感，故将其称为图像的不变矩。计算公式如下：

$$\begin{cases} \Phi_1 = \eta_{20} + \eta_{02} \\ \Phi_2 = (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \\ \Phi_3 = (\eta_{30} + 3\eta_{12})^2 + 3(\eta_{21} - \eta_{03})^2 \\ \Phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \Phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ \quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \Phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \Phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ \quad + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{cases} \quad (4-4)$$

4.2 基于 BP 神经网络的目标识别

人工神经网络是一种模仿大脑神经联系而进行信息处理的数学模型，由大量的神经元节点和它们之间带有权重的连接构成。每一个节点都相当于一种特定的输出函数，被称作激励函数；任意节点间的连接代表着神经网络的记忆，具有不同的加权值，被称作权重；神经网络的输出则由激励函数、网络权重以及连接方式共同决定。

4.2.1 人工神经网络概述

1. 人工神经元模型

生物神经元细胞的种类很多，但是就其结构组成的角度来看，都是由细胞体以及大量轴突、树突组成，基本结构具有极大的相似性。其中，细胞体主要是由外部的细胞膜和内部的细胞核构成，功能上类似于初等处理器。在细胞体外侧存在着大量的树状分枝，可以感受到周围的抑制或是兴奋信号，这种具有感知能力的部分相当于细胞的信息输入端，被称作树突。与细胞的接受信息端相对应的是信息输出端，即由细胞体向外延伸出较长距离的分支，被称为轴突。其末端具有丰富的神经末梢，释放神经递质进而传递输出信号。不同神经细胞的轴突和树突通过突触相连接，突触作为信息传递的中介，代表了神经元之间所传递信息的性质以及连接强度^[48]，生物神经元模型如图 4-1 所示。

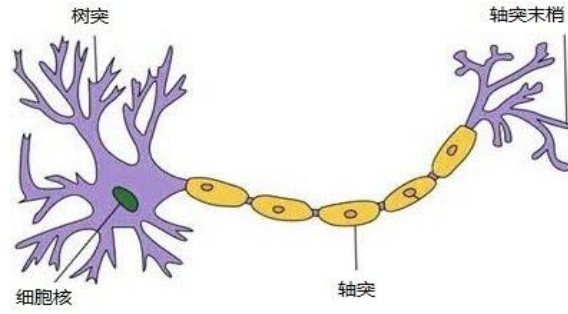


图 4-1 神经元模型示意图

类似地，人工神经网络由大量的人工神经元构成。其中，信息的处理主要是由网络各个节点之间连接的权值决定。人工神经网络在使用之前需要进行训练，训练的过程实际上就是不断地修正和调整各个权值的过程。

人工神经元是模拟生物神经元的一种数学模型。如图 4-2 所示，人工神经元是多输入单输出的非线性处理单元。

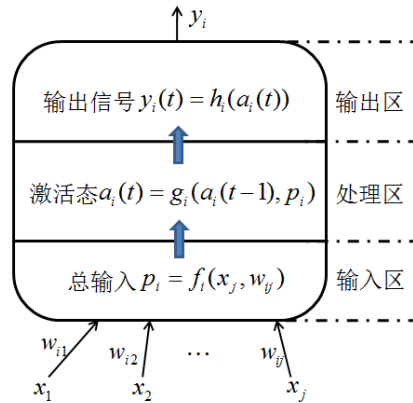


图 4-2 人工神经元非线性结构

人工神经元中的每个输入量 x_j 都有与其相关的权重 w_{ij} 。在处理区中，计算输入及其对应的权重值得到加权之和，计算出唯一的输出量。这个输出量 y 和权重和之间的映射关系被称为神经元的传递函数。这个过程可以用下式表示。

$$y = f\left(\sum_j w_{ij}x_j + b\right) \quad (4-5)$$

其中 f 代表着神经网络中该神经元所采用的传递函数类型。传递函数的不同也使得各种神经网络在结构和功能上有所差异。一般情况下，神经网络的传递函数在训练过程保持不变，改变的是网络的权重，而权重用以存储最重要的信息。

2. 人工神经网络的互联结构

由于神经元的连接方式不同，导致神经网络的结构差异较大。神经网络常被分为前馈模型和后馈模型^[49]。可进一步细分为无反馈的前向网络、有反馈的前向网络、层内有相互结合的前向网络以及相互结合型网络。

人工神经网络的互联结构如图 4-3 所示。

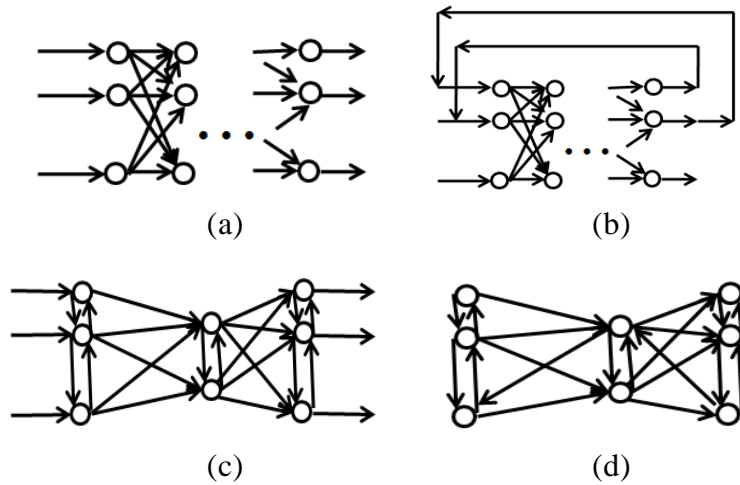


图 4-3 网络结构的各种类型

1) 无反馈的前向网络, 如图 4-3 (a)所示。神经元按形成不同分层有序排列, 通常由输入层、隐蔽层以及输出层构成。层间的神经元信号是单向传播, 从输入层开始向后逐层传递直至到输出层, 最终获得输出信号。

2) 有反馈的前向网络, 如图 4-3 (b)所示。该网络的特点是引入了反馈, 使得神经元之间的信号传递关系更为丰富。

3) 具有层内互联的前向网络, 如图 4-3 (c)所示。该网络的基本结构与上述网络相比并没有较大变化, 但是同层之间的神经元存在着连接关系, 这样可以实现同层之间神经元信号的互相传递。

4) 相互结合型网络, 如图 4-3 (d)所示。该网络神经元之间的连接较为复杂, 既有前馈又有反馈, 层内的神经元也互相连接。神经元间连接得越多就意味着信号传递的可能种类越多, 在信号传递过程中, 网络中的连接权值在不断变化, 最终可能达到动态平衡也可能是周期振荡状态。

4.2.2 BP 神经网络结构

神经网络中应用最为成熟的 BP 神经网络属于多层前向型网络, 其信号前向传播, 误差反向传播。标准的 BP 神经网络是采用梯度下降法, 网络权值沿着性能函数的梯度反向调整。

图 4-4 所示模型即为基本 BP 神经元模型, 其 S 个输入均被分别赋予相应的权值, 求这些输入的加权和之后与偏差共同组成该神经元传递函数的输入。

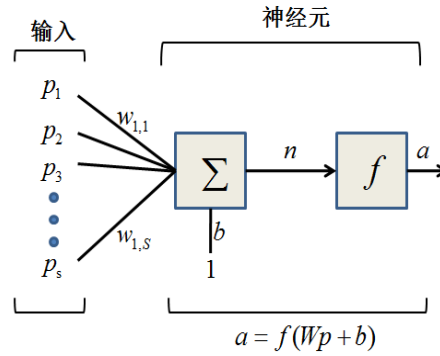


图 4-4 BP 神经元模型示意图

相邻层之间的节点为全连接，连接强度采用权重参数表示。网络输出层的具体形式由所要进行的任务决定，如在多分类问题中，输出层的每一维可以表示对应类别的预测置信度。在前馈网络中，网络层数、节点连接权重共同确定具体的前馈网络实例。在给定神经网络的结构之后，神经网络的训练目标是根据训练样本学习网络节点之间的连接权重，使得神经网络在未知的测试数据集上上进行相同的任务时，能够取得更好的测试性能。

4.2.3 BP 神经网络学习算法

使用 BP 学习算法训练前馈网络的基本思想是：首先通过一个前向过程计算训练样本中的输入特征经过前向网络后的实际输出值；然后将输出值与期望值进行比较，得到误差信号，再根据误差信号从后向前调节各个神经元之间的连接强度；再次重复进行前向过程计算，使得误差减小，再将新的输出值与期望值进行比较，得到新的比原来小的误差信号；接着根据小的误差信号，从后向前调整各神经网络层神经元之间的连接强度，这个前向过程和后向过程不断地多次进行，直到误差满足要求为止。

为了确定网络的所有连接权重，一般是通过训练样本集合来学习出这些权重参数^[50-52]。一个包含 N 个样本的训练集可以表示为：

$$X = \{x_n, t_n\}_{n=1}^N \quad (4-6)$$

其中 x_n 表示该样本对应的 n_i 维输入特征， t_n 为 n_0 维类别输出变量。

定义损失函数：

$$l(\Theta; x_n, t_n) \quad (4-7)$$

来计算一组特定的模型参数在一个样本上的损失值。其中， Θ 是一个参数集合，表示所有的连接权重，即

$$\Theta = \{W_{ij}^l, b_j^l\} \quad (4-8)$$

其中, w_{ij}^l 表示连接第 l 层第 j 个节点和第 $l+1$ 层第 i 个节点之间的权重, b_j^l 为第 $l+1$ 层第 i 个节点的偏置项。

为了得到最优的连接权值, 需要求解如下的最优化问题:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \frac{1}{N} \sum_{n=1}^N l(\Theta; x_n, t_n) \quad (4-9)$$

该优化问题的目标是求取特定的参数值, 使得损失函数在训练集上的总损失和最小。由于损失函数的具体定义不同, 上述最优化问题的求解过程也会相应不同。当损失函数的导数存在时, BP 学习算法给出一种通用过程来求解上述最优化问题。

若以平方和误差作为优化目标, 则损失函数 $l(\Theta; x_n, t_n)$ 可以表示为:

$$l(\Theta; x_n, t_n) = \frac{1}{2} \|t_n - y_n\|^2 \quad (4-10)$$

其中, y_n 表示前馈网络的输入为 x_n 时所得到的输出函数值。相应地, 整个训练集上的目标损失函数可以表示为:

$$L(\Theta) = \frac{1}{2} \sum_{n=1}^N l(\Theta; x_n, t_n) \quad (4-11)$$

该目标函数相对于参数 Θ 的梯度方向为:

$$\nabla L(\Theta) = \frac{\partial L(\Theta)}{\partial \Theta} = \frac{1}{N} \sum_{n=1}^N (h_{\Theta}(x_n) - t_n) \frac{\partial h_{\Theta}(x_n)}{\partial \Theta} \quad (4-12)$$

假设参数 Θ 的初始值为 Θ_0 , 根据梯度下降法, 不断修正参数 Θ :

$$\Theta_{k+1} = \Theta_k + \Delta \Theta_k = \Theta_k - \eta \nabla L(\Theta_k) \quad (4-13)$$

其中 $\Delta \Theta_k = -\eta \nabla L(\Theta_k)$ 表示第 k 次迭代参数的更新量, $\eta \in (0, 1)$ 为学习速率。

BP 学习算法巧妙地使用链式求导法则, 给出一种从后向前分层更新参数的过程。首先通过前向过程计算得到训练样本在当前网络的输出; 然后根据网络损失函数比较输出值与期望值, 计算网络的误差; 接着通过将网络误差从后向前逐层传播来更新网络参数, 根据更新后的网络参数, 不断重复上述过程, 直到训练过程收敛。

假设网络中计算节点使用的非线性函数为 sigmoid 函数, 为了方便后面推导反向传播算法, 将前馈网络第 l 层的第 i 个节点的总输入记为 z_i^l , 其输出记为 a_i^l , 对于网络的隐层有:

$$\begin{cases} a_i^l = \sigma(z_i^l) = \frac{1}{1 + \exp(-z_i^l)} \\ z_i^{l+1} = \sum_{j=1} W_{ij}^{l+1} a_j^{l+1} + b_i^{l+1} \end{cases} \quad (4-14)$$

假设当前网络的错误为 $E = L(\Theta)$ ，对于连接权重 ΔW_{ij}^l 和偏置 Δb_i^l ，在一次更新过程的更新量分别为：

$$\begin{cases} \Delta W_{ij}^{n_l} = \frac{\partial E}{\partial W_{ij}^{n_l}} \\ \Delta b_i^l = \frac{\partial E}{\partial b_i^l} \end{cases} \quad (4-15)$$

反向传播过程从最后一层开始，不断应用链式求导法则，由后向前逐层地计算 ΔW_{ij}^l 和 Δb_i^l ，进而更新各层参数。

对于输出层来说，其连接偏置 $\Delta b_i^{n_l-1}$ 的更新量为：

$$\Delta b_i^{n_l-1} = \frac{\partial E}{\partial a_i^{n_l}} \frac{\partial a_i^{n_l}}{\partial z_i^{n_l}} \frac{\partial z_i^{n_l}}{\partial b_i^{n_l-1}} = -(t_i - a_i^{n_l}) \sigma'(z_i^{n_l}) \quad (4-16)$$

其连接权重 $\Delta W_{ij}^{n_l-1}$ 的更新量为：

$$\Delta W_{ij}^{n_l-1} = \frac{\partial E}{\partial z_i^{n_l}} \frac{\partial z_i^{n_l}}{\partial W_{ij}^{n_l-1}} = \Delta b_i^{n_l-1} a_j^{n_l-1} \quad (4-17)$$

对于隐层，其连接偏置 $\Delta b_i^{(l-1)}$ 的更新为：

$$\Delta b_i^{(l-1)} = \sum_{j=1}^{s_{l+1}} \frac{\partial E}{\partial z_j^{l+1}} \frac{\partial z_j^{l+1}}{\partial a_i^l} \frac{\partial a_i^l}{\partial z_i^l} \frac{\partial z_i^l}{\partial b_i^{l-1}} = \left(\sum_{j=1}^{s_{l+1}} W_{ji}^{l+1} \Delta b_j^l \right) \sigma'(z_i^l) \quad (4-18)$$

其连接权重 $\Delta W_{ij}^{(l-1)}$ 的更新量为：

$$\Delta W_{ij}^{(l-1)} = \sum_{j=1}^{s_{l+1}} \frac{\partial E}{\partial z_j^{l+1}} \frac{\partial z_j^{l+1}}{\partial a_i^l} \frac{\partial a_i^l}{\partial z_i^l} \frac{\partial z_i^l}{\partial W_{ij}^{l-1}} = \Delta b_i^{l+1} a_j^{l-1} \quad (4-19)$$

4.2.4 基于 BP 神经网络的目标识别流程

用大量数据样本对神经网络进行不断地训练是应用 BP 神经网络对不同目标进行识别的前提，其中的样本包括训练输入和期望输出两个部分。训练输入数据往往是图像的某些特征量，而期望输出则可设置为图像对应的代号。

在用样本对神经网络进行训练的过程中，图像中存在很多干扰信息，这增加了网络训练的难度，因此本文采用基于不变矩的策略滤除干扰项。基于不变矩的方法将目标图像中的不变矩和图像中分拣目标的代号共同组成训练样本，

可以将目标图像变为含有 7 个数字的一组不变矩，该策略可以简化训练样本滤除干扰信息，使训练效率得到大幅提高的同时还可以提升识别准确率。为了最大限度的滤除干扰信息，还需对目标图像进行二值化处理，为提取不变矩特征做必要的准备。

本文采用 BP 神经网络对工件进行识别，应用基于不变矩特征的样本对网络进行训练，识别算法实现流程步骤如下：

1) 获取目标图像信息

获得所要识别的对象各个角度的图像。

2) 图像预处理

灰度化并且二值化为适合求不变矩的图像。

3) 求不变矩

用高度浓缩的图像特征作为识别图像的依据。

4) 生成神经网络的训练样本和识别样本

将多角度拍摄的目标图像进行预处理后，求出每幅图像的不变矩，把形成的向量和目标代号组合起来，就形成了一个训练样本，每一幅图像都可以得到一个训练样本。若去掉训练样本中的目标代号，余下的不变矩序列就可以作为待识别的样本。

5) 建立神经网络

本文采用单隐含层结构的 BP 神经网络对工件进行识别，其结构如图 4-5 所示，其隐含层中 20 个神经元的传递函数均为 S 函数。由于待识别的工件有三种，故将输出层设置成 3 个神经元且传递函数均为线性函数。

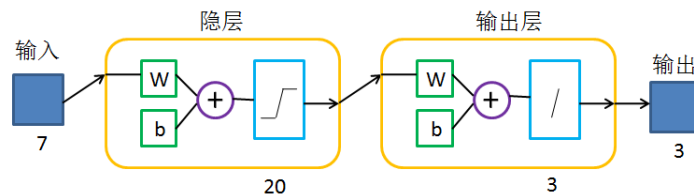


图 4-5 单隐含层 BP 神经网络结构

6) 训练神经网络

本文为了使训练算法同时具有高斯-牛顿法的局部收敛性还具有梯度下降法的全局特性，采用了性能稳定的 Levenberg-Marquardt 算法。

7) 识别目标

经大量不变矩特征样本训练后的神经网络权重值趋于稳定，当给定待识别的图像不变矩时，神经网络通过映射给出识别结果。

4.3 BP 神经网络识别算法优化

4.3.1 Adaboost 基本原理

1988 年, Kearns 等^[53]在研究 PAC 学习模型时提出 Boosting 问题, 即“弱可学习是否等价于强可学习”。Schapire 认为只要获得比随机猜测略好的弱分类器, 就可以通过这些弱分类器的组合, 获得强分类器的识别效果。

Freund 根据在线分配算法, 改进 Boosting 方法, 提出 Adaboost 算法^[54]。由于训练样本中有些样本点是较难分类或是分类错误率较高, 可以通过改变样本的权重, 而不是随机地抽取样本点, 将分类重点放在那些较难区分的样本点上。在多个弱分类器的共同决策方法上也引入了权重的思想, 对那些分类效果较好错误率较低的弱分类器给予较大的权值, 反之给予较小的权重。通过对样本点和弱分类器采用加权方式, 可以获得更好的识别效果。

Adaboost 通过维护一个定于训练样本集上的样本权重分布, 改变数据分布实现强分类器。判断当前轮迭代是训练集中的每一样本是否正确分类, 结合上轮迭代时的总体分类正确率, 确定下一轮迭代时训练集样本的权值。将修改过权值的新数据集送给下轮迭代中的分类器进行训练, 最后将每轮训练得到的分类器赋予不同的权重系数并叠加融合, 作为最后的决策分类器。

为了获得不同类别的训练集合训练每一个弱分类器, Addboost 算法不断地调节着样本点所对应的权重值。每组样本初始权重相同, 通过 Addboost 算法在对弱分类器进行训练过程中实现对错误分类样本和正确分类样本的权重值调整, 即可得到凸显分错样本比重的新样本分布, 据此不断对弱分类器进行训练, 逐次迭代经过 T 轮后可以得到 T 个新的弱分类器。最终, 赋予这 T 个弱分类器一定的权重并且进行叠加, 得到最终的强分类器^[55,56]。

4.3.2 Adaboost 算法实现及算例分析

1. Adaboost 算法实现流程

给定训练样本以及分类 $(x_1, y_1), \dots, (x_n, y_n)$, 其中 $x_i \in X, y_i \in Y = \{-1, +1\}$ 。初始化样本权重为 $D_1(i) = 1/n$, 即为训练样本的初试权值分布, 考虑最大循环次数为 T 的迭代训练, 实现流程为:

- 1) 对于第 $t = 1, \dots, T$ 次迭代, 使用样本的权值分布 D_t 训练弱分类器;
- 2) 获得弱分类假设 $h_t : X \rightarrow \{-1, +1\}$, 计算弱分类器所对应的错误率:

$$\varepsilon_t = \Pr_{i \sim D_t} [h_t(x_i) \neq y_i] \quad (4-20)$$

3) 选取 $\alpha_t = 1/2\ln[(1-\varepsilon_t)/\varepsilon_t]$, B_t 为归一化参数, 更新第 $t+1$ 次迭代中的权值分布:

$$D_{t+1}(i) = \frac{D_t(i)}{B_t} \exp\{-\alpha_t y_i h_t(x_i)\} \quad (4-21)$$

4) 在经过 T 次迭代训练之后, 得到 h_1, h_2, \dots, h_T , T 个弱分类器;

5) 最后通过加权投票, 得到强分类器, $\text{sign}(\cdot)$ 表示符号运算:

$$H(x) = \text{sign}\left\{\sum_{t=1}^T \alpha_t h_t(x)\right\} \quad (4-22)$$

2. Adaboost 算法算例分析

记号 "+", "-" 分别代表训练集 $X = \{x_1, x_2, \dots, x_n\}$, $n=10$ 中的两种类别, 正确分类集 $Y = \{+1, +1, +1, -1, -1, +1, -1, +1, -1, -1\}$, 按列从上至下排列。在对 X 进行分类中, 使用水平或者垂直直线作为分类器。

1) 初始化样本 x_i 的权值 $D_1 = \{0.1, 0.1, \dots, 0.1\}$;

2) 采用分类器 h_1 对 X 分类, 假设分类错误的样本为 3 个, 在图示中用已标出错误样本, 则 h_1 的分类错误率 $\varepsilon_1 = 0.30$; 根据公式 $\alpha_t = 1/2\ln[(1-\varepsilon_t)/\varepsilon_t]$, 得到权系数 $\alpha_1 = 0.42$;

3) 根据权值更新迭代公式 $D_{t+1}(i) = D_t(i)/Z_t \exp\{-\alpha_t y_i h_t(x_i)\}$, 更新权值分布, 可得 $D_2 = \{0.0714, 0.0714, 0.1667, 0.0714, 0.1667, 0.0714, 0.1667, 0.0714, 0.0714\}$ 。显然在第 $t=1$ 轮迭代训练中, 被 h_1 错分样本的权系数在 D_2 中增大;

4) 对于第 $t=2$ 轮迭代训练, 计算可得分类器 h_2 的权系数 $\alpha_2 = 0.21$ 和错误率 $\varepsilon_2 = 0.65$ 继而更新样本权系数分布 D_3 。同样, 被 h_2 错分样本的权系数在 D_3 中增大;

5) 对于第 $t=3$ 轮迭代训练, 计算可得分类器 h_3 的权系数 $\alpha_3 = 0.14$ 和错误率 $\varepsilon_3 = 0.92$;

6) 将上述三轮迭代训练中得到的 3 个弱分类器 h_1, h_2, h_3 加权投票, 组合最终强分类器 H :

$$H = \text{sign}(0.42h_1 + 0.65h_2 + 0.92h_3) \quad (4-23)$$

4.3.3 基于 Adaboost 的 BP 神经网络识别算法优化

基于神经网络的方法对目标机械工件的识别不会绝对准确, 为了尽可能地提高分类网络的识别性能, 引入 Adaboost 算法将多个弱分类器进行加权组合, 构成识别性能更佳的强分类器。该分类器可以通过根据训练时的目标分类是否正确调整下一次网络训练的权重, 综合多个神经网络的表现给出最终识别结果。

利用 Adaboost 算法可以提高 BP 神经网络识别目标工件准确率，具体实现流程如图 4-6 所示：

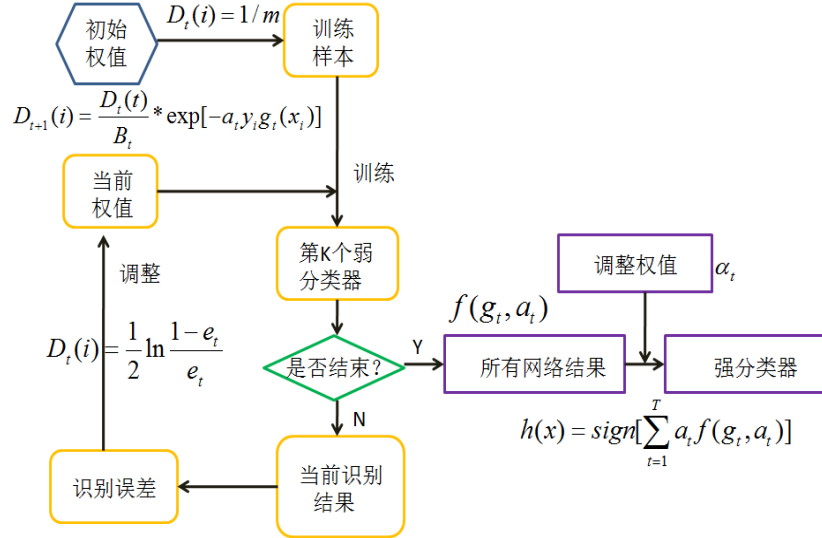


图 4-6 基于 Adaboost 算法的权值调整流程

提升神经网络识别率的具体步骤如下：

1) 神经网络初始化

在样本空间中随机选取 m 组数据用于多层神经网络结构，其中每组数据分初始布权值为 $D_t(i) = 1/m$ ，初始化每层神经网络的权值和相应阈值。

2) 弱分类器处理

通过由单层神经网络训练得到的弱分类器预测训练数据的输出，其中预测分类序列 $g(t)$ 的分类错误和 e_t 的计算公式如下：

$$e_t = \sum_i D_t(i) \quad i = 1, 2, \dots, m \quad (g(t) \neq y) \quad (4-24)$$

式中， $g(t)$ 和 y 分别代表预测分类序列和期望分类序列

3) 分类权重序列计算

分类权重序列 a_t 的权重计算公式如下：

$$a_t = \frac{1}{2} \ln\left(\frac{1-e_t}{e_t}\right) \quad (4-25)$$

4) 数据权值实时调整

依据式 (4-25) 的权重不断调整下一时刻训练数据的权重，实时调整公式如下：

$$D_{t+1}(i) = D_t(i) / B_t * \exp[-a_t y_i g_t(x_i)] \quad i = 1, 2, \dots, m \quad (4-26)$$

其中，的 B_t 表示归一化系数。

5) 经过 T 轮的训练后可以得到其弱分类函数 $f(g_i, a_i)$ ，再经实时权重组合调整后可以得到强分类函数 $h(x)$

$$h(x) = \text{sign}\left(\sum_{i=1}^T \alpha_i \cdot f(g_i, a_i)\right) \quad (4-27)$$

4.4 目标工件识别实验

本节利用 Matlab 软件进行目标识别，通过调用神经网络工具箱中的函数进行 BP 神经网络的构建、训练以及仿真。

1) 建立 BP 神经网络的函数原型为：

```
newff(minmax(inputdata),[m,n],{'tansig','pureline'},'trainlm');
```

该函数需要提供四个参数。其中，参数 `minmax(inputdata)` 表示输入数据的取值范围；参数 `[m,n]` 表示神经网络隐层的神经元个数为 m ，输出神经元个数为 n ；参数 `{'tansig','pureline'}` 表示神经网络中采用的传递函数类型；参数 `'trainlm'` 表示的是训练神经网络的方法。

2) 训练 BP 神经网络的函数原型为：

```
train(netname,inputdata,outputdata);
```

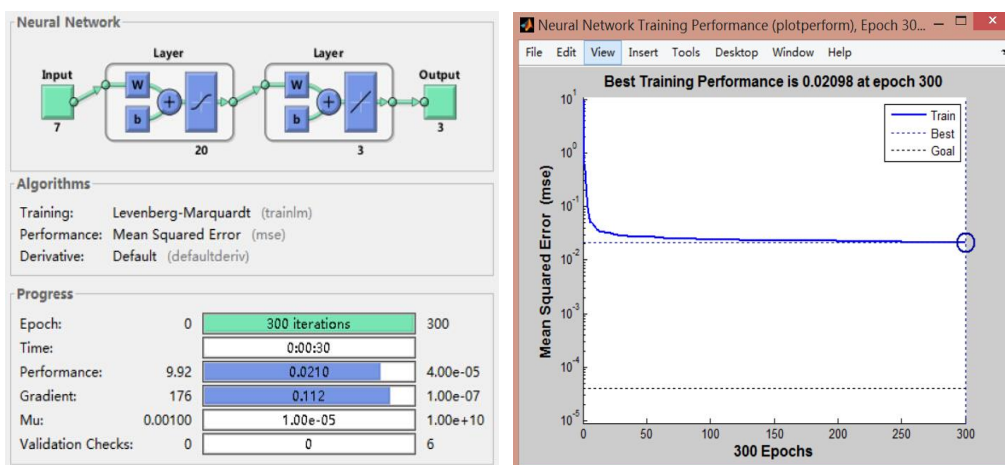
在利用数据 `inputdata,outputdata` 对神经网络进行训练之前，需要设定训练的相关参数，其中包括训练周期、训练误差和训练步长。

3) 仿真 BP 神经网络的函数原型为：

```
sim(netname,inputdata_test);
```

利用训练好的神经网络 `netname` 对数据 `inputdata_test` 进行识别。

神经网络的训练信息如图 4-7 所示。



(a) 网络训练参数

(b) 均方误差变化曲线

图 4-7 BP 神经网络训练信息

图 4-7 (a) 表示的是 BP 神经网络的训练参数。该 BP 神经网络隐层的神经元个数设定为 20，由于要识别的是三种不同种类的工件，输出层神经元个数设定为 3；训练方法采用的是 Levenberg-Marquardt 算法；在训练过程中 Gradient 表示误差曲面的梯度，当其达到设定值时，会终止网络训练。变量 μ 是随着迭代过程中误差的改变而改变的，若 μ 数值过大，同样会停止训练进程。

图 4-7 (b) 表示的是均方误差随着训练周期变化的曲线，本实验中设定的误差限为 0.00004，训练周期为 300，训练步长为 0.1，随着训练周期数的增加，误差值迅速降低，由于达到训练周期而停止训练，此时的训练误差为 0.02098，若希望进一步减小误差直至所设定的目标误差，可以适当更改训练样本数量以及训练次数，目前训练误差可以满足实验识别要求。

从不同角度对每一个工件拍 400 幅图片，每幅图片均可以提取一组矩特征，并结合工件标号作为一组训练样本，共获取 1200 组训练样本对神经网络进行训练；在识别部分，采集每个工件 33 幅图片，共 99 组不变矩作为测试样本。

通过上述训练样本实现神经网络中神经元权值的确定，完成整个神经网络的学习。再将待识别图像的不变矩特征作为识别数据，经过神经网络的分类得到的识别结果绘制表格如下，其中基于 BP 神经网络的分类器识别结果见表 4-1。

表 4-1 单个弱分类器识别结果统计

错误类型	小螺栓	大螺栓	螺帽	总错误	错误率
错误统计	3	3	1	7	7.07%

构成强分类器的 10 个弱分类器总的识别错误数以及相应的分类误差率见表 4-2，强分类器识别错误数和相应误差率见表 4-3。

表 4-2 多个弱分类器识别结果统计

弱分类器	1	2	3	4	5	6	7	8	9	10	总错误	弱分类器总体分类错误率
错误数	6	9	7	8	7	6	9	7	5	9	73	7.37%

表 4-3 强分类器识别结果统计

错误类型	小螺栓	大螺栓	螺帽	总错误	错误率
错误统计	2	1	1	4	4.04%

将表 4-2 和表 4-3 中的数据进行对比和分析，可发现强分类器的分类错误率低于 10 个弱分类器的平均分类错误率，这表明多个神经网络通过加权投票方式获得识别效果要好于单一的神经网络。之所以强分类器的识别效果更好，是因为强分类器可以根据每个网络的训练情况调节权重，不断地削弱无用特征，最终使得整体识别率较高。

将强弱分类器识别工件的正确率进行对比，如表 4-4 所示。

表 4-4 强、弱分类器识别效果对比

分类方式	基于 BP 神经网络的弱分类器	基于 Adaboost 算法的强分类器
正确率	91.93%	94.82%

通过大量实验可以发现基于 BP 神经网络的分类器在识别本文中的机械工件时，识别率是比较高的。但是如果拍摄目标工件图样的角度特殊时，会对弱分类器的识别效果产生较大影响，相对基于 Adaboost 算法的强分类器而言，不论在识别的准确率还是在识别的抗干扰程度上都要更加可靠。

4.5 本章小结

本章目的是实现对分拣平台上静态工件的识别，为机器人分拣不同目标工件做好前期准备。首先由绪论中介绍的图像模式识别基本理论可知，在图像识别过程中，关键是找到图像的特征，其中图像的矩特征具有平移旋转不变性，可以表征目标工件的本质属性并且提取相对容易，故作为识别的依据；其次从神经网络基本单元到网络互联方式以及学习过程等几个方面详细介绍了神经网络相关理论，给出基于 BP 神经网络识别目标工件的具体步骤；再次，为了提高 BP 神经网络的识别准确率，本文将基于单个神经网络的弱分类器进行性能提升，引入 Adaboost 算法将多个神经网络组成强分类器，并且从原理上分析了提高 BP 网络识别准确率的原因；最后通过实验给出弱分类器和强分类器对于工件图像的识别结果，验证了改进算法在目标识别上的优势。

第 5 章 基于 SCARA 机器人的目标工件分拣

5.1 SCARA 机器人的 D-H 参数建模

SCARA 机器人是基于视觉机器人分拣系统中的重要组成部分，完成对传送带上的目标工件的分拣抓取任务，关于该类型的机器人的外观及基本参数已经在前面章节中叙述，本章主要对基于 SCARA 机器人的分拣控制策略进行研究。首先介绍 D-H 参数法及机器人坐标系表示法，在此基础上建立 SCARA 机器人运动学模型；其次针对机器人运动学方程的求解进行分析并简要介绍运动学方程的逆解选取原则；最终结合视觉分拣机器人在分拣过程中的实际情况，详细分析了机械手末端执行器在三维空间中的位姿和机械手臂各个主要关节转角之间的关系，进而得到基于最优抓取点的分拣策略。

Denavit 和 Hartenberg 提出一种系统地描述串行链路式机械手臂的方法，后来被广泛地采用，该方法即为 D-H 参数表示法^[57]。虽然针对机械手的 D-H 描述法有很多种改进，但是由于本文中的机器人具有四个自由度，使用经典的 D-H 法就可以准确地得到 SCARA 机器人模型，下面对 D-H 参数以及机械手的坐标系确定方法进行介绍。

连杆是机器人的重要组成部件，通过关节将一组有序的杆件连接在一起构成机械手。除了首末两个特殊连杆之外，其他的连杆两端的轴线均存在着一法线，这两条法线的距离记为 d_i ，相邻的法线夹角记为 θ_i ，公共法线的距离记为 a_i ，垂直于 a_i 所在平面内两轴的夹角记为 α_i ，相关参数如图 5-1 所示。

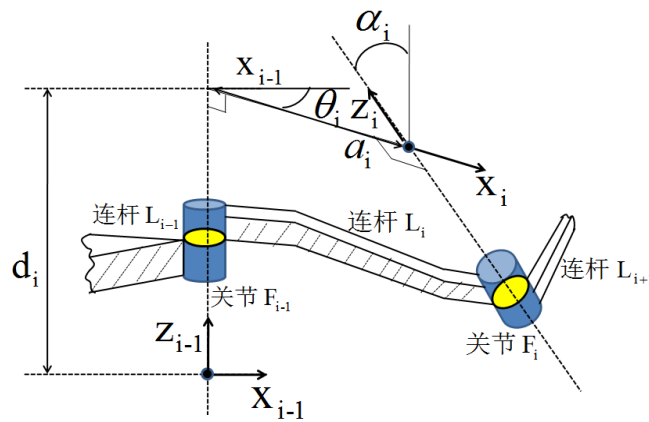


图 5-1 机械手 D-H 参数描述

一般地，称 a_i 为连杆长度， α_i 为连杆扭角， d_i 为两个连杆的距离， θ_i 为两连杆的夹角。需要采用两个参数可对连杆进行描述，分别为连杆的长度 a_i 和连杆的扭曲度 α_i 。同样需要另外两个参数来描述相邻两杆之间的关系，分别为两连杆距离 d_i 和两杆件的夹角 θ_i 。标准 D-H 法参数含义如表 5-1 所示。

表 5-1 标准 D-H 参数描述

名称	符号	含义
杆件长度(mm)	a_i	z_{j-1} 轴与 z_j 轴之间的公法线长度
扭曲度($^\circ$)	α_i	z_{j-1} 轴与 z_j 轴之间绕 x_i 旋转的角度
杆件偏移(mm)	d_i	x_{j-1} 轴与 x_i 轴之间沿 z_{j-1} 轴的距离
关节转角($^\circ$)	θ_i	x_{j-1} 轴与 x_j 轴之间绕 z_{j-1} 轴旋转的角度

通常，对机械手坐标系的确定是按照如下顺序：

- 1) 将与转轴轴线平行的方向定义为 z_{j-1} 轴；
- 2) 选取 z_{j-1} 轴和 z_j 轴的公共法线上的点作为坐标轴原点；
- 3) 将垂直两连杆公垂线由 z_{j-1} 指向 z_j 方向定义为 x_{j-1} 轴；
- 4) 通过右手定则，最终确定 y_{j-1} 轴。

明确了机器人的 D-H 参数以及坐标系含义后，根据以上规则，可以得到本实验中 SCARA 机械手的 D-H 模型，如图 5-2 所示。

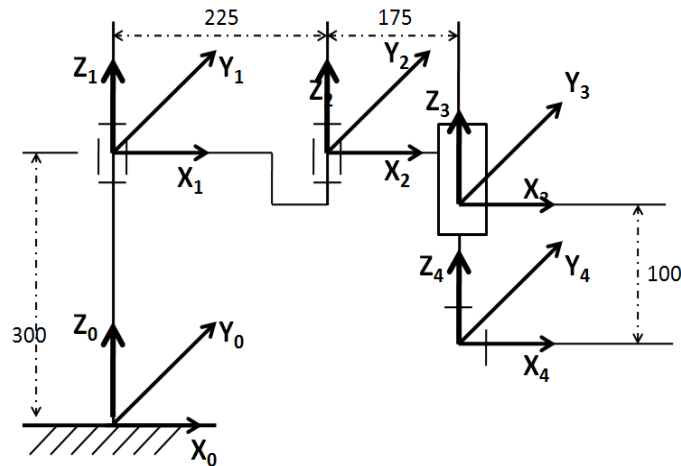


图 5-2 机械手 D-H 模型

描述 SCARA 机器人关节坐标系之间相互转换的矩阵 ${}^{j-1}_jT$ 表达式如下：

$${}^{j-1}_jT(\theta_j, d_j, a_j, \alpha_j) = T_{rz}(\theta_j)T_z(d_j)T_x(a_j)T_{rx}(\alpha_j) \quad (5-1)$$

其中， $\theta_j, d_j, a_j, \alpha_j$ 描述的是关于串行链路式机械手臂的相关 D-H 参数。根据上可以得到关节的齐次变换矩阵。

基于旋转关节的齐次变换矩阵是：

$${}^{j-1}_jT = \begin{bmatrix} \cos \theta_j & -\sin \theta_j \cos \alpha_j & \sin \theta_j \cos \alpha_j & a_j \cos \theta_j \\ \sin \theta_j & \cos \theta_j \cos \alpha_j & -\cos \theta_j \sin \alpha_j & a_j \sin \theta_j \\ 0 & \sin \alpha_j & \cos \alpha_j & d_j \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5-2)$$

基于移动关节的齐次变换矩阵是：

$${}^{j-1}_jT = \begin{bmatrix} \cos \theta_j & -\sin \theta_j \cos \alpha_j & \sin \theta_j \sin \alpha_j & 0 \\ \sin \theta_j & \cos \theta_j \cos \alpha_j & -\cos \theta_j \sin \alpha_j & 0 \\ 0 & \sin \alpha_j & \cos \alpha_j & d_j \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5-3)$$

本文是应用平面四自由度机械手臂来实现分拣作业的，该机械手的基本参数见表 5-2：

表 5-2 SCARA 机器人参数表

编号	关节变量	d_j	a_j	α_j	变量范围
1	θ_1	0	$l_1 = 225 \text{ mm}$	0	$+/-115^\circ$
2	θ_2	0	$l_2 = 175 \text{ mm}$	0	$+/-140^\circ$
3	d_3	$-d_3$	0	0	$0 \sim 150 \text{ mm}$
4	θ_1	0	0	0	$+/-360^\circ$

其中，描述机械臂旋转关节 1 和 2 的角度变量分别是 θ_1 和 θ_2 ；描述移动关节 3 的变量是 d_3 ；描述机械手末端的旋转关节 4 的角度变量是 θ_4 。

机械手的基座与关节 1 之间的齐次变换矩阵为：

$$\begin{aligned} {}^0_1T &= \text{Trans}(0,0,0) \cdot \text{Rot}(z, \theta_1) \cdot \text{Trans}(225,0,0) \\ &= \begin{bmatrix} \cos \theta_1 & -\sin \theta_1 \cos \alpha_1 & \sin \theta_1 \sin \alpha_1 & l_1 \cos \theta_1 \\ \sin \theta_1 & \cos \theta_1 \cos \alpha_1 & -\cos \theta_1 \sin \alpha_1 & l_1 \sin \theta_1 \\ 0 & \sin \alpha_1 & \cos \alpha_1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad (5-4)$$

机械手关节 1 与关节 2 之间的齐次变换矩阵为：

$$\begin{aligned}
 {}^1_2T &= Rot(z, \theta_2) \cdot Trans(175, 0, 0) \\
 &= \begin{bmatrix} \cos \theta_2 & -\sin \theta_2 \cos \alpha_2 & \sin \theta_2 \sin \alpha_2 & l_2 \cos \theta_2 \\ \sin \theta_2 & \cos \theta_2 \cos \alpha_2 & -\cos \theta_2 \sin \alpha_2 & l_2 \sin \theta_2 \\ 0 & \sin \alpha_2 & \cos \alpha_2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5-5)
 \end{aligned}$$

机械手关节 2 与关节 3 之间的齐次变换矩阵为：

$$\begin{aligned}
 {}^2_3T &= Trans(0, 0, -d_3) \\
 &= \begin{bmatrix} \cos \theta_3 & -\sin \theta_3 \cos \alpha_3 & \sin \theta_3 \sin \alpha_3 & 0 \\ \sin \theta_3 & \cos \theta_3 \cos \alpha_3 & -\cos \theta_3 \sin \alpha_3 & 0 \\ 0 & \sin \alpha_3 & \cos \alpha_3 & -d_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5-6)
 \end{aligned}$$

机械手关节 3 与关节 4 之间的齐次变换矩阵为：

$${}^3_4T = \begin{bmatrix} \cos \theta_4 & -\sin \theta_4 \cos \alpha_4 & \sin \theta_4 \sin \alpha_4 & 0 \\ \sin \theta_4 & \cos \theta_4 \cos \alpha_4 & -\cos \theta_4 \sin \alpha_4 & 0 \\ 0 & \sin \alpha_4 & \cos \alpha_4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5-7)$$

将描述各个关节之间位姿变化的矩阵依次相乘，得到如下方程：

$$\begin{aligned}
 {}^0_4T &= {}^0_1T(\theta_1) \cdot {}^1_2T(\theta_2) \cdot {}^2_3T(d_3) \cdot {}^3_4T(\theta_4) \\
 &= \begin{bmatrix} n_x & o_x & a_x & p_x \\ n_y & o_y & a_y & p_y \\ n_z & o_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5-8) \\
 &= \begin{bmatrix} C(\theta_1 + \theta_2 + \theta_4) & -S(\theta_1 + \theta_2 + \theta_4) & 0 & l_1 C(\theta_1) + l_2 C(\theta_1 + \theta_2) \\ S(\theta_1 + \theta_2 + \theta_4) & C(\theta_1 + \theta_2 + \theta_4) & 0 & l_1 S(\theta_1) + l_2 S(\theta_1 + \theta_2) \\ 0 & 0 & -1 & d_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}
 \end{aligned}$$

式 (5-8) 是本实验中平面四自由度机器人的运动学方程。其中， 0_4T 表示的是 SCARA 机器人末端执行器的位姿矩阵，矩阵中两两垂直的向量 $\vec{n}, \vec{o}, \vec{a}$ 描述 SCARA 末端执行器的方向信息，向量 \vec{p} 描述 SCARA 机器人末端执行器的位置信息。符号 S 和 C 表示正弦函数和余弦函数， $\theta_1, \theta_2, d_3, \theta_4$ 是描述 SCARA 机器人旋转关节和移动关节的变量值。

5.2 SCARA 机器人的运动学逆解分析

机械手的运动学方程描述的是旋转或移动关节变量和末端执行器之间关系。如果已知的是机械手关节变量，则可以通过运动学方程得到机械手在三维空间的末端位置坐标，这个过程即为机械手的运动学正解。同样的道理，如果已知的是机械手的在三维空间中的末端位置坐标，依据运动学可以得到机械手臂的各个关节变量的值，这个过程即为机械手的运动学逆解。机械手运动学的正逆解关系如图 5-3 所示。

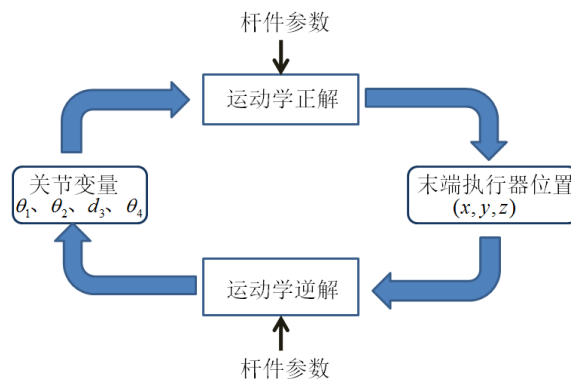


图 5-3 机械手的正逆运动学解

机械手可以通过改变其关节变量值使得末端执行器到达传送带上的待分拣工件的空间坐标点处，但是有时候存在机械臂运动死区以及机械臂关节变量存在多种解情况，为了确定分拣抓取的位置，需要研究 SCARA 机械手末端位置和关节转角之间的关系。

5.2.1 运动学逆解的求取方法

一般地，若机械手的关节较少且关节轴相交于一点时，则该机械手的运动学方程存在着逆解。本文讨论的平面四自由度机器人满足运动学逆解存在的必要条件，可以利用逆矩阵进行不断左乘，逐渐递推的方法得到逆解。

1) 求解关节变量 θ_1

将可逆矩阵 $({}^0_1T)^{-1}$ 左乘运动学方程 (5-8) 两边，可得：

$$({}^0_1T)^{-1} \cdot {}^0_4T = {}^1_2T \cdot {}^2_3T \cdot {}^3_4T$$

$$\begin{bmatrix} \cos \theta_1 & \sin \theta_1 & 0 & 0 \\ -\sin \theta_1 & \cos \theta_1 & 0 & 0 \\ 0 & 0 & 1 & -d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} n_x & o_x & a_x & p_x \\ n_y & o_y & a_y & p_y \\ n_z & o_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos(\theta_2 - \theta_4) & \sin(\theta_2 - \theta_4) & 0 & a_1 + a_2 \cos \theta_2 \\ \sin(\theta_2 - \theta_4) & -\cos(\theta_2 - \theta_4) & 0 & a_2 \sin \theta_2 \\ 0 & 0 & -1 & -d_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5-9)$$

根据式 (5-9)，由于矩阵中对应位置上的元素相等，可得下式：

$$\begin{cases} p_x \cdot \cos \theta_1 + p_y \cdot \sin \theta_1 = a_1 + a_2 \cdot \cos \theta_2 \\ -p_x \cdot \sin \theta_1 + p_y \cdot \cos \theta_1 = a_2 \cdot \sin \theta_2 \end{cases} \quad (5-10)$$

进而可以化简得到：

$$\theta_1 = \arctan\left(\frac{A}{\pm \sqrt{1-A^2}}\right) - \varphi \quad (5-11)$$

$$\text{式中 } A = \frac{p_x^2 + p_y^2 + a_1^2 - a_2^2}{2a_1 \sqrt{p_x^2 + p_y^2}}, \quad \varphi = \arctan \frac{p_x}{p_y}。$$

2) 求解关节变量 θ_2

将求解的关节变量值 θ_1 代入到方程组 (5-10)，可得下式：

$$\theta_2 = \arctan\left(\frac{r \cdot \cos(\theta_1 + \varphi)}{r \cdot \sin(\theta_1 + \varphi) - a_1}\right) \quad (5-12)$$

$$\text{式中, } r = \sqrt{p_x^2 + p_y^2}, \quad \varphi = \arctan \frac{p_x}{p_y}。$$

3) 求解关节变量 d_3

根据式 (5-9)，由于矩阵方程中的元素对应相等，可得下式：

$$d_3 = d_1 - p_z \quad (5-13)$$

4) 求解关节变量 θ_4

同理，根据矩阵方程 (5-9) 可得：

$$\begin{cases} n_x \cdot \cos \theta_1 - n_y \cdot \sin \theta_1 = \cos(\theta_2 - \theta_4) \\ -n_x \cdot \sin \theta_1 + n_y \cdot \cos \theta_1 = \sin(\theta_2 - \theta_4) \end{cases} \quad (5-14)$$

解上式 (5-14)，可得下式：

$$\theta_4 = \arctan(\cos \theta_1 \cdot n_y - \sin \theta_1 \cdot n_x) + \theta_2 \quad (5-15)$$

机械手臂的所有关节变量均已求出，但是上述方法的计算量非常大，而且对于本课题来说，并不需要知道 SCARA 四个自由度的全部变量值，事实上，实际分拣作业时，移动关节 3 的运动距离是固定的，就是垂直于传送带的距离，而关节 4 是吸盘式的末端执行器，不需要获得具体转角来实现吸附抓取分拣作

业。重要的是关节 1 和 2 的变量值，它们决定了机械手是否可以运动到工件的正上方。

针对本课题，在求取机械手的运动学逆解问题中，可以做适当简化，SCARA 机械手的俯视图如图 5-4 所示：

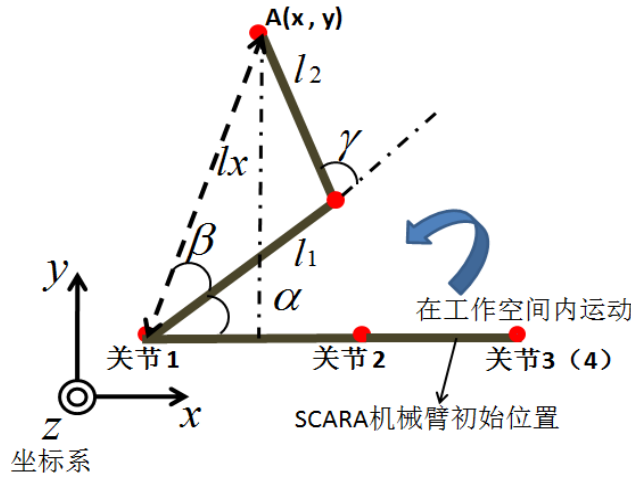


图 5-4 SCARA 机械手俯视图

利用余弦定理可得到抓取位置点 A 和关节 1 和 2 之间的关系：

$$l_x^2 = l_1^2 + l_2^2 - 2l_1l_2 \cos(\pi - \gamma) \quad (5-16)$$

化简上式，可得：

$$\gamma = \arccos \frac{l_x^2 - l_1^2 - l_2^2}{2l_1l_2} \quad (5-17)$$

继续使用余弦定理，可得：

$$\begin{cases} l_2^2 = l_x^2 + l_1^2 - 2l_xl_1 \cos \beta \\ \sin(\alpha + \beta) = \frac{y}{l_x} \end{cases} \quad (5-18)$$

化简上式，可得：

$$\alpha = \arcsin \frac{y}{l_x} - \arccos \left(\frac{l_x^2 + l_1^2 - l_2^2}{2l_xl_1} \right) \quad (5-19)$$

其中， α 和 γ 分别代表着关节 1 和关节 2 的转角。

5.2.2 运动学逆解的选取标准

对机械手关节转角变量值的研究是实现准确分拣抓取作业的前提。通过运动学逆解的求解过程分析可知运动学方程往往存在多解的情况，在逆解的选取问题上需要有一定的标准，通常选取逆解需考虑如下几点^[60]：

1) 机械手的工作空间范围。由于存在着机械结构上的制约,通过关节相互的机械手臂存在着工作死区,即机械手末端是无法到达某些位置点的。为了保证各个关节可以在其工作状态下正常转动,需要满足如下约束条件:

$$q_i^L \leq q_i \leq q_i^U, \quad i=1,2,\dots,n \quad (5-20)$$

式中, q_i^L 代表着第 i 个关节所能转动的最小角度; q_i^U 代表着第 i 个关节所能转动的最大角度。

2) 位置误差最小原则。机械手臂在分拣作业过程中,会存在着一定的位置运动误差,并且随着传送带的运动,导致实际机械手末端抓取位置和理论值有误差,为了保证抓取的准确性,要求误差大小 ε 在一定范围之内。

$$\|{}^oP_w - {}^oP_{w,des}\| < \varepsilon \quad (5-21)$$

式中, oP_w 代表着机械手末端实际运动到的坐标点; ${}^oP_{w,des}$ 代表着机械手末端理论上运动到的坐标点。

3) 柔顺度原则。柔顺度指的是各个关节角度变化具有连续性,表现为机械手臂在运动过程中相对平稳,波动较小。为了尽量减少因关节转动所带来的机械磨损,需要满足下面的条件:

$$\Delta q = \|\{q\} - \{q_{cur}\}\| \leq k \quad (5-22)$$

式中, $\{q\}$ 代表的是机械手的关节在下一位姿的转角; $\{q_{cur}\}$ 代表的是机械手的关节在当前位姿中的转角。 k 是常数,代表着阈值,对关节在两个位置之间转角差值进行约束。

4) 小关节运动优先原则。SCARA 机器人关节所连接的杆件质量不同,靠近基座的杆件一般质量较大,惯性也较大,在运动过程中不但容易产生振动,而且需要耗费更多的能量。为避免上述不利现象,所以在机械手可以实现分拣作业的前提下,尽量选择小关节运动。

5.3 基于最优抓取点的分拣策略

SCARA 型机器人进行分拣抓取作业时,依靠着旋转关节 1、2 转动的角度,到达目标工件的正上方,然后移动关节 3 伸出机械连杆,使机械末端配备的电磁吸盘靠近目标物体,对相应的工件进行抓取,放到分类容器中,实现分拣任务。由于关节 3 与传送带的垂直距离是固定的,只要机械手末端到达了目标上方,后续的移动关节 3 和旋转关节 4 的配合运动所需的时间必将也是相对固定

的，所以我们将研究的重点放在机械手从初始位置运动到目标上方的过程中，怎样实现更好的分拣策略。

事实上，在讨论 SCARA 机械臂的分拣策略时候，存在两个关于机械手运动学的问题。其一，由于 SCARA 机器人的手臂存在运动盲区，即对于某些特殊抓取位置点，机械手臂末端是无法到达的，也就是运动学方程的逆解不存在。其二，SCARA 机器人对于同一个抓取位置，如图 5-5 位置点 A，具有多种实现方法，即关节 1、2 可以具有至少两种的组合形式，也就是说该机器人运动学方程的逆解存在且不唯一。本文主要讨论的是运动学存在逆解的情况下，机械臂抓取位置的最佳选择问题。

不难想象，对于机器人来说，在机械臂角速度 ω 一定的条件下，如果机械臂的关节转角越小，机械臂到达抓取目标位置的时间就越少，所需要的能耗就越少，我们把 SCARA 机械臂的运动时间和能耗均最小的抓取位置称为最优抓取点。对于在 SCARA 机器人作业范围的某一个位置点，总有着一种对应关系，即传送带上的位置点和 SCARA 机器人关节 1、2 转角之间的映射关系。那么最优抓取点的选择就转换为关节 1、2 最优转角的选择。

首先分析传送带上只放置一个目标工件的情况，如图 5-5 所示。不妨设工件被随机地放置在距机械臂初始位置的垂直距离为 h 处，在随着传送带运动过程中，工件会形成一条直线运动轨迹 l ，其中 A、B、C 三点可视为任意一个工件在传送带上经过的三个不同位置。通过 SCARA 机械手的运动学逆解，我们可以得到 l 上任意一点 (x, y) 和关节 1 的转角 α 和关节的 2 的转角 γ 之间的关系，即 $\alpha(x, y)$ 和 $\gamma(x, y)$ 。很显然，对于上述三个不同的可能抓取位置点，机械手关节 1、2 的组合运动角度 (α, γ) 是不同的。确定最优抓取点的原则是使得 SCARA 机械手的组合转角 (α, γ) 最优。

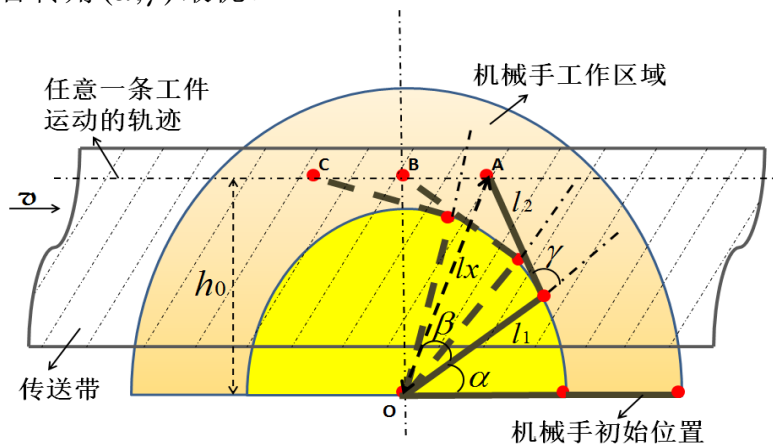


图 5-5 SCARA 机械手分拣作业示意图

假设 SCARA 型机器人不同关节具有相同的角速度 ω ，那么对于机械臂到达目标工件正上方的时间和能耗就取决于 α 和 γ 之间较大的一个。由此我们定义最优目标函数如下：

$$\phi = \max\{\alpha(x, y), \gamma(x, y)\} \quad (5-23)$$

其中，角度值 ϕ 表示机械手运动到点 (x, y) 所需要的必要关节转角。

在直线轨迹 l 上，除了机械臂无法到达的奇异点之外的所有抓取 (x, y) 中，如果它们对应的 SCARA 机械臂必要关节转角 ϕ 最小，那么相应的位置就是最优抓取点。在直线轨迹 l 上的最优抓取位置所对应的角度为 $\phi = \min\{\phi\}$ 。

如果对于某一点 (x, y) ，SCARA 运动学方程有两个逆解，所对应的两个关节转角分别为 $\{\alpha_1, \alpha_3\}$ 和 $\{\alpha_2, \alpha_4\}$ ，如图 5-6 所示。

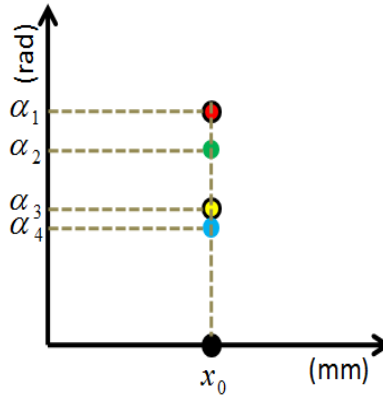


图 5-6 运动学逆解选取分析

对于图 5-6 设定的机械手末端位置与转角关系，那么

$$\alpha_2 = \min(\max(\alpha_1, \alpha_3), \max(\alpha_2, \alpha_4)) \quad (5-24)$$

所以 (α_2, α_4) 所对应的机械臂末端执行器位置就是最优抓取点。

对于更一般的情况如图 5-7 所示。

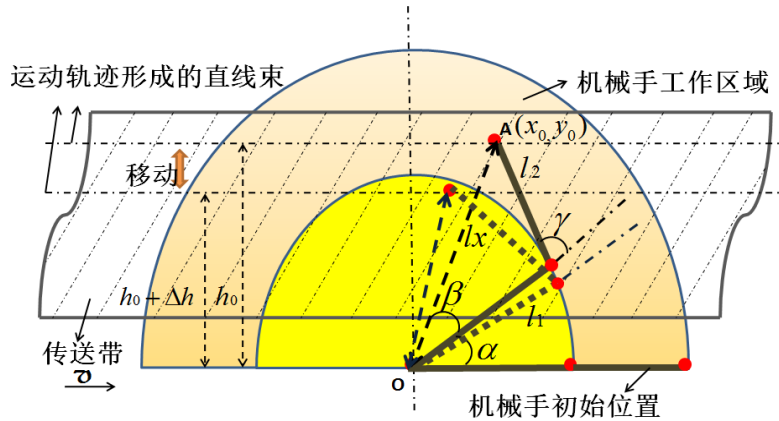


图 5-7 最优抓取位置分析

某一个工件被随机的放置在传送带上，工件初始位置距 SCARA 机械臂初始位置的垂直距离为 h_0 ，那么如前所述，在该工件运动的直线轨迹中存在一点使得机械臂关节 1、2 的转角相对最小。同样的道理，如果工件初始位置距 SCARA 机械臂初始位置的垂直距离为 $h_0 + \Delta h$ ，那么对于新生成的直线轨迹中也存在着最优抓取位置，在工件随传送带运动形成的直线束可以找到一系列的最优抓取点，即形成了最优抓取轨迹。

5.4 目标工件分拣实验

本实验中的视觉分拣系统平台的构建以及重要的组成部件选取问题已经在第二章中有所叙述，此处不再赘述。通过第四章的目标跟踪可以获取实时的图像目标位置点坐标，将该坐标转换到空间三维坐标系中，利用 RC40 控制器实现对机械手臂的控制，最终完成分拣任务。

下面介绍视觉分拣系统中的常见坐标系及其相互转换。

1. 图像坐标系

通过摄像机拍摄目标物体，成像系统的底片上会形成一个 $M \times N$ 的网格结构，每个小空格所在的位置是像素点，所拍摄的图像由这些像素点构成。可以将该网络中的某一点定位为原点 O_o ，利用目标物体像素点相对于原点 O_o 的位置定义该像素点的坐标为 (u, v) ，表示像素点在横坐标上距离原点 O_o 有 u 个像素，同理在纵坐标上距离原点 O_o 有 v 个像素。但是该坐标描述法并没有量纲单位，为了进行精确的描述，引入度量单位。一般地，在图像的中心位置选择一个原点 O_i ，该原点正是摄像机的光心轴和图像平面相交点，并且平行于 u 轴和 v 轴建立 X 轴和 Y 轴，其中图像坐标系如图 5-8 所示。

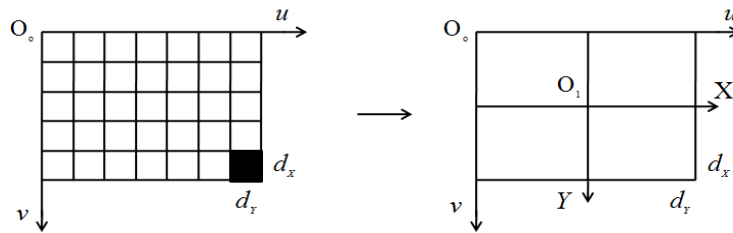


图 5-8 图像坐标系

如果坐标原点 O_i 位于 $u-v$ 坐标系第 u_0 行和第 v_0 列，该坐标系的像素点之间的距离用 d_x 和 d_y 表示，那么 $u-v$ 坐标系和 $X-Y$ 坐标系中的任意一像素点的坐标之间转换关系为：

$$\begin{cases} u = \frac{X}{d_x} + u_0 \\ v = \frac{Y}{d_y} + v_0 \end{cases} \quad (5-25)$$

将上述用矩阵简化表示为：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/d_x & 0 & u_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (5-26)$$

2. 摄像机坐标系

图 5-9 为摄像机坐标系示意图，不妨设点 P 是三维空间中任一点，经过摄像机透镜投影到相机底片上形成图像点 p ，点 P 和点 p 之前的连线与摄像机光轴相较于一点，该点 o_c 定义为摄像机坐标系原点，沿着光轴通过原点 o_c 的线称为 Z_c 轴，与图像坐标系的 X 轴和 Y 轴分别平行的轴定义为摄像机的 X_c 轴和 Y_c 轴。

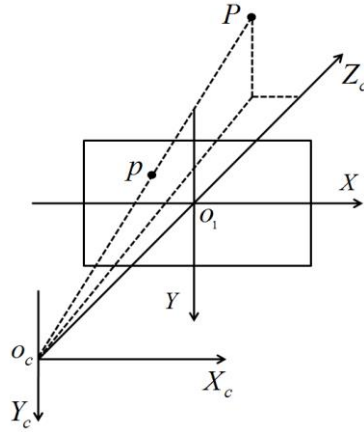


图 5-9 摄像机坐标系

3. 世界坐标系

世界坐标系就是三维空间坐标系，描述摄像机在现实空间中的位置，往往作为基准坐标系，用字母 $O_w X_w Y_w Z_w$ 表示。若三维空间中存在一点 P ，该点在世界坐标系下的坐标表示为 $(X_w, Y_w, Z_w, 1)^T$ ，在摄像机坐标系下的坐标表示为 $(X_c, Y_c, Z_c, 1)^T$ ，那么世界坐标系和摄像机坐标系之间的转换关系如下：

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (5-27)$$

下面分析视觉分拣实现的基本过程以及分拣存在的问题及原因。

利用 Camshift 算法获取传送带上运动工件的实时坐标，如图 5-10（a）所示。利用图 5-10（b）所示的串口助手并通过 RS232 可将 PC 机上的二进制程序代码传递给控制器 RCX40。将图像坐标转化为三维空间坐标后，发送给控制器，由控制器发送点对点的控制信号，导引 SCARA 机器人运行到工件的上方，通过机械手末端的电磁吸盘对目标工件进行抓取。



(a) 工件实时坐标的获取

(b) 工件实时坐标的传送

图 5-10 位置信息的获取与传递

SCARA 机器人末端执行器采用电磁吸盘结构，控制器 RC40 导引机器人运动到工件附近对其进行吸附，分拣过程如图 5-11 所示。



图 5-11 SCARA 机器人分拣实验

总体来说，视觉分拣系统可以完成工件分拣任务，但是仍然存在如下两个有待进一步解决的问题。

1) 搭建好视觉分拣实验平台，在调试过程中发现工件的分布情况对视觉分拣系统有较大影响，若待分拣的目标在传送带上的分布密度较大，或者是目标之间发生重叠，基于 CamShift 的目标跟踪算法会将工件视为一个整体，难以给出每个工件的坐标，这样的结果是机械手末端的电磁吸盘将多个工件共同吸附，无法实现分拣任务。

2) 在传送带上相距一定间隔地放置工件,测得机器人对螺栓和螺帽的实际分拣速度最快约为 1 次/2s,分拣速度与目标工件位置坐标点及其运行快慢有关,若提高传送带的运行速度,跟踪算法容易发散,控制器将无法获得目标物体的坐标而导致分拣失败。由于传送带速度过快,会对基于 CamShift 的跟踪算法产生较大的影响。一方面该算法对被跟踪目标的运动速度有一定限制,特别是无法跟踪速度突变的情况;另一方面,当目标被分拣后,该算法不断地更新搜索框尺寸寻找下一个工件,这个过程需要一定的时间,这一特点同样限制了传送带的运行速度。

5.5 本章小结

本章主要介绍了目标分拣策略和视觉分拣实现。利用 D-H 参数法对 SCARA 机器人进行建模,分析关于机械手臂运动学逆解的求取问题,进而讨论逆解的选取标准。在此基础上,详细研究了机械手的末端执行器位置和机械臂关节转角之间的关系,从最优控制角度出发,得到分拣抓取点的选取原则,使得机械手的抓取时间和能量损耗达到最优。计算机将跟踪算法中提取到的目标坐标信息传递给机器人控制器,进而导引机械手运动,实现了视觉分拣作业。最后,针对实验中视觉分拣的不足进行说明并分析原因。

结 论

在工业生产线上，工件的分拣作业是一种常见工序。机器视觉可以让机器人更具智能化，满足复杂环境下的生产要求，自动地完成繁重流水线上的分拣任务，可以显著提高生产力。本文讨论了基于机器视觉的工业机器人分拣系统相关技术，所做的研究和成果有：

(1) 设计了基于机器视觉的工业机器人分拣系统的构建方案并介绍了系统各个组成部分。其中，针对图像采集部分用到的相机和镜头进行了详细地说明，并且给出摄像机标定的步骤以及标定实验结果；

(2) 机器人分拣过程中，目标工件是随着传送带进行运动的，为了实现动态目标抓取任务，研究了图像跟踪相关技术。本文基于 CamShift 算法进行目标跟踪，从算法的原理上进行了详细地分析并且给出了算法实现的具体步骤，最后通过实验得到所跟踪目标的实时坐标；

(3) 针对分拣过程中的必要步骤，研究了图像的识别方法。通过提取待分拣的螺栓和螺帽特征来训练 BP 神经网络，为了提高识别的准确率，应用 Adaboost 算法对 BP 神经网络进行性能提升。实验结果表明基于多个神经网络有机组合的强分类器识别分类效果更好；

(4) 对于 SCARA 机器人的抓取策略进行研究。利用 D-H 参数法对机器人进行建模，分析运动学逆解相关问题，从最优控制角度出发，分析了抓取位置和机械臂转角之间的关系，得到基于最优抓取点的分拣策略。

本文是针对视觉分拣系统进行研究，取得一定成果，并且得出以上结论。但是仍然存在一些问题，可以从如下两个方面进行深入研究。

(1) 分拣工件任务对实时性的要求是非常严格的，虽然本文中采用的 CamShift 算法可以实现对目标的跟踪并能够获取目标的实时坐标，但为使分拣系统可以在工业流水线上更高效工作，需要在更快速且稳定的目标跟踪算法上进行深入研究；

(2) 分拣过程中，传送带上的工件分布是随机的，密度大时要防止漏拣，密度小时要减少等待分拣时间，所以可以尝试着从传送带的运动速度控制角度进行研究，得到一种变速度的传送带控制方法，进而提高工件分拣效率。

参考文献

- [1] 刘振宇, 李中生, 赵雪, 邹凤山. 基于机器视觉的工业机器人分拣技术研究[J]. 制造业自动化, 2013, 35(9): 25-30.
- [2] 栗俊艳, 米月琴, 弓靖, 刘鹏飞. 一种基于机器视觉的工业机器人分拣系统[J]. 电子科技, 2016, 29(1): 105-110.
- [3] Jyoti Jhavar. Orange Sorting by Applying Pattern Recognition on Colour Image[J]. Procedia Computer Science, 2016, 78: 691-697.
- [4] Kevin C. Deegan, Laura Koivisto, Juha Näkkilä, Lea Hyvönen, Hely Tuorila. Application of a sorting procedure to greenhouse-grown cucumbers and tomatoes[J]. LWT - Food Science and Technology, 2010, 43(3): 393-400.
- [5] Muhammad Ali Ashraf, Naoshi Kondo, Tomoo Shiigi. Use of Machine Vision to Sort Tomato Seedlings for Grafting Robot[J]. Engineering in Agriculture, Environment and Food, 2011, 4(4): 119-125.
- [6] Gangqi Dong. Autonomous robotic capture of non-cooperative target by adaptive extended Kalman filter based visual servo[J]. Acta Astronautica, 2016, 122: 209-218.
- [7] 蔡自兴, 郭璠. 中国工业机器人发展的若干问题[J]. 机器人技术与应用, 2013, 3: 9-12.
- [8] 张朝阳. 基于视觉的机器人废金属分拣系统研究[D]. 北京: 中国农业大学学位论文, 2015: 8-9.
- [9] 计时鸣, 黄希欢. 工业机器人技术的发展与应用综述[J]. 机电工程, 2015, 32(1): 1-13.
- [10] 孙英飞, 罗爱华. 我国工业机器人发展研究[J]. 科学技术与工程, 2012, 12(12): 2912-2918.
- [11] 赵杰. 我国工业机器人发展现状与面临挑战[J]. 航空制造技术, 2012, 12: 26-29.
- [12] 蔡荣太, 吴元昊, 王明佳, 吴庆祥. 视频目标跟踪算法综述[J]. 视频应用与工程, 2010, 34(12): 135-138.
- [13] Mohamed F.A., Rama C., Zheng Q. Integrated motion detection and tracking for visual surveillance[C]. Proc. of the 4th IEEE international conference on

- computer vision system, New York, 2006: 28.
- [14]Paul Viola, Michael J. Jones, Daniel Snow. Detecting pedestrians using pattern of motion and appearance[C]. Proceeding of the ninth IEEE international conference on computer vision, 2003, 6: 13-16.
- [15]高文, 朱明, 贺柏根, 吴笑开. 目标跟踪技术综述[J]. 中国光学, 2014, 7(3): 365-375.
- [16]闫庆森, 李临生, 徐晓峰, 王灿. 视频跟踪算法研究综述[J]. 计算机科学. 2013, 40(6A): 204-209.
- [17]Long Jin, Yunong Zhang, Tianjian Qiao, Manchun Tan, Yinyan Zhang. Tracking control of modified Lorenz nonlinear system using ZG neural dynamics with additive input or mixed inputs[J]. Neurocomputing, 2016, 196: 82-94.
- [18]Yilmaz A., Li X., Shah M. Contour-based object tracking with occlusion handling in video acquired using mobile cameras[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2004, 26(11): 1531-1536.
- [19]张家怡. 图像识别的技术现状和发展趋势[J]. 电脑知识与技术, 2010, 6(21): 6045-6046.
- [20]戴汝为. 语义、句法模式识别方法及其应用[J]. 模式识别与人工智能, 1995, 8(2): 89-93.
- [21]Jing Li, Nong Sang, Changxin Gao. Completed local similarity pattern for color image recognition[J]. Neurocomputing, 2016, 182: 111-117.
- [22]Mariusz Młynarczuk, Andrzej Górszczyk, Bartłomiej Ślipek. The application of pattern recognition in the automatic classification of microscopic rock images[J]. Computers and Geosciences, 2013, 60: 126-133.
- [23]王强, 张小溪, 韩一红. 基于神经网络的图像识别[J]. 电子设计工程, 2012, 20(9): 187-189.
- [24]赵清杰, 连广宇, 孙增圻. 机器人视觉伺服综述[J]. 控制与决策, 2001, 16(6): 849-853.
- [25]倪受东, 刘洋, 袁祖强. 机器人视觉伺服综述[J]. 机床与液压, 2007, 35(9): 227-230.
- [26]王麟琨, 徐德, 谭民. 机器人视觉伺服研究进展[J]. 机器人, 2004, 6(3): 277-282.
- [27]方勇纯. 机器人视觉伺服研究综述[J]. 智能系统学报, 2008, 3(2): 109-114.
- [28]H Hashimoto, M Sato, F Harashima. Visual control of robot ics manipulator

- based on neural networks[J]. IEEE Trans on Industrial Electronics, 1992, 39(6): 490-496.
- [29]G Wells, C Venaille, C Torras. Promising research vision-based robot positioning using neural networks[J]. Image and Vision Computing, 1996, 14(10): 715-732.
- [30]卢军, 郑国穗, 马金锋, 刘杰. SCARA 机器人结构优化设计与运动分析[J]. 陕西科技大学学报, 2014, 32(1): 137-141.
- [31]沈中城. 检测技术与仪器[M]. 北京: 高等教育出版社, 2003.
- [32]刘燕. H 型钢截面尺寸线结构光测量[D]. 天津: 天津大学学位论文, 2012: 16-20.
- [33]张铨伟, 王彪, 徐贵力. 摄像机标定方法研究[J]. 计算机技术与发展, 2010, 11(20): 174-179.
- [34]Qian Sun, Xiaoyi Wang, Jiping Xu, Li Wang, Huiyan Zhang, Jiabin Yu, Tingli Su, Xun Zhang. Camera self-calibration with lens distortion[J]. Optik International Journal for Light and Electron Optics, 2016, 127(10): 4506-4513.
- [35]张芝贤, 赵远方, 武旭娟. 基于 Halcon 的多目摄像机标定技术[J]. 2015,23(38): 13-140.
- [36]赵小川, 何灏, 缪远诚. MATLAB 数字图像处理实战[M]. 北京: 机械工业出版社, 2012: 159-161.
- [37]Zongjie Cao, Liyuan Xu, Jilan Feng. Automatic target recognition with joint sparse representation of heterogeneous multi-view SAR images over a locally adaptive dictionary[J]. Signal Processing, 2016, 126: 27-34.
- [38]Haichao Zheng, Xia Mao, Lijiang Chen, Xiaogeng Liang. Adaptive edge-based mean-shift for drastic change gray target tracking[J]. International Journal for Light and Electron Optics, 2015, 126(23): 3859-3867.
- [39]Youness Aliyari Ghassabeh, Frank Rudzicz. The mean-shift algorithm and its relation to kernel regression[J]. Information Sciences, 2016, 348: 198-208.
- [40]Jianfang Dou, Jianxun Li. Robust visual tracking based on joint multi-feature histogram by integrating particle filter and mean-shift[J]. Optik, 2015, 126: 1449-1456.
- [41]Tomas Vojir, Jana Noskova, Jiri Matas. Robust scale-adaptive mean-shift for tracking[J]. Pattern Recognition Letters, 2014, 49: 250-258.
- [42]X. Mei, H.B. Ling, Y. Wu, E.P. Blasch, L. Bai. Efficient minimum error

- bounded particle resampling L1 tracker with occlusion detection[J]. IEEE Trans. Image Process, 2013, 22 (7): 2661-2675.
- [43]Zhang Shuang, Qin Yu-Ping. Mean-shift algorithm apply for infrared imaging tracking[C]. AASRI Procedia, 2012, 1: 52-57.
- [44]杨杰, 张翔. 视频目标检测和跟踪及其应用[M]. 上海: 上海交通大学出版社, 2012: 161-185.
- [45]徐磊. 基于 Mean Shift 算法的 Camshift 跟踪技术研究[J]. 价值工程, 2015, 29: 202-203.
- [46]查宇飞, 毕笃彦, 杨源, 鞠明. 视频目标跟踪方法[M]. 北京: 国防工业出版社, 2015: 265-296.
- [47]E.G. Karakasis, A. Amanatiadis, A. Gasteratos, S.A. Chatzichristofis. Image moment invariants as local features for content based image retrieval using the Bag-of-Visual-Words model[J]. Pattern Recognition Letters, 2015, 55: 22-27.
- [48]韩丽. 神经网络结构优化方法及应用[M]. 北京: 机械工业出版社, 2012: 98-105.
- [49]范佳妮, 王振雷, 钱锋. BP 人工神经网络隐层结构设计的研究进展[J]. 控制工程, 2005, 12: 105-109.
- [50]Liao GM, Zheng X. An improved BP neural network based on IPSO and its application[J]. Comput 2013, 8(5): 1267-1272.
- [51]Xuefeng Zhao, Qin Ba, Lei Zhou, Weijie Li, Jinping Ou. BP neural network recognition algorithm for scour monitoring of subsea pipelines based on active thermometry[J]. Optik, 2014, 9(7): 5426-5431.
- [52]S.P. Adam, D.A. Karras, G.D. Magoulas, M.N. Vrahatis Solving the linear interval tolerance problem for weight initialization of neural networks[J]. Neural Netw, 2014: 27-37.
- [53]Kearns M., Valiant L. Cryptographic limitations on learning Boolean formulae and finite automata[J]. Journal of the ACM, 1994, 41(1): 67-95.
- [54]Freund Y., Schapire R.E. A decision-theoretic generalization of on-line learning and an application to Boosting[J]. Journal of Computer and System Sciences, 1997, 55(1): 119-139.
- [55]谢剑斌. 视觉机器学习[M]. 北京: 清华大学出版社, 2015: 81-82.
- [56]曹莹, 苗启广, 刘家辰, 高琳. Adaboost 算法研究进展与展望[J]. 自动化学报, 2013, 39(6): 745-758.

- [57]郭发勇, 梅涛, 赵江海. D-H 法建立连杆坐标系存在的问题及改进[J]. 中国机械工程, 2014,25(10): 2710-2714.
- [58]Niphun Surapong, Chowarit Mitsantisuk. Position and Force Control of the SCARA Robot Based on Disturbance Observer[C]. Procedia Computer Science, 2016, 86: 116-119.
- [59]Wenchang Zhang, Jiangping Mei, Yabin Ding. Design and Development of a High Speed Sorting System Based on Machine Vision Guiding[C]. Physics Procedia, 2012, 25: 1955-1965.
- [60]闫昊. SCARA 机器人动力学分析及鲁棒性控制研究[D]. 哈尔滨: 哈尔滨工业大学学位论文, 2013: 14-15.

哈尔滨工业大学学位论文原创性声明和使用权限

学位论文原创性声明

本人郑重声明：此处所提交的学位论文《基于机器视觉的工业机器人分拣系统设计》，是本人在导师指导下，在哈尔滨工业大学攻读学位期间独立进行研究工作所取得的成果，且学位论文中除已标注引用文献的部分外不包含他人完成或已发表的研究成果。对本文的研究工作做出重要贡献的个人和集体，均已在文中以明确方式注明。

作者签名：何译强 日期：2016年6月20日

学位论文使用权限

学位论文是研究生在哈尔滨工业大学攻读学位期间完成的成果，知识产权归属哈尔滨工业大学。学位论文的使用权限如下：

(1) 学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文，并向国家图书馆报送学位论文；(2) 学校可以将学位论文部分或全部内容编入有关数据库进行检索和提供相应阅览服务；(3) 研究生毕业后发表与此学位论文研究成果相关的学术论文和其他成果时，应征得导师同意，且第一署名单位为哈尔滨工业大学。

保密论文在保密期内遵守有关保密规定，解密后适用于此使用权限规定。
本人知悉学位论文的使用权限，并将遵守有关规定。

作者签名：何译强 日期：2016年6月20日

导师签名： 日期：2016年6月20日

致 谢

光阴似箭，岁月如歌。两年的研究生生活即将结束，时光短暂，我却收获良多。回首求学过程中的点滴，仍历历在目。对那些引导我、帮助我、激励我的人，我心中充满了感激。

感谢导师马家辰教授的悉心指导。从论文定题到写作定稿，倾注了马老师大量的心血，在课题的研究中，给了我很多宝贵的建议和启发，当我遇到实验瓶颈的时候，总是耐心地帮我答疑解惑，并监督我课题的进展。同时，在生活中，教给我做人做事的方法，使我受益终身。

感谢实验室的各位师兄师姐、同届的伙伴们和师弟们，能在实验室遇到你们，很幸运。实验室氛围很好，希望实验室不断壮大。

感谢寝室的同学们，我们来自不同的专业，朝夕相处两年，不论在学习上还是在生活中遇到怎样挫折，我们都互相鼓励。积极面对困难，共同分享喜悦，一路走来，建立了深厚的友谊。

感谢母校，感谢控制专业的老师们，是你们给了我接触更多专业知识的平台，拓宽了我的视野。

需要特别感谢的是我的家人。在我面临人生选择而迷茫之际，为我排忧解难，是我十多年求学路上的坚强后盾；在我做出成绩之时，流露出的喜悦是对我最大的鼓励和肯定。父母的养育之恩无以为报，对我无私的爱与照顾是我不断前进的动力。