

摘 要

机器视觉是智能机器人发展的一个重要方向,虚拟仪器是计算机应用技术领域里一个新兴的、富有生命力的软件平台。本文基于虚拟仪器技术,在研究近年来机器人视觉技术的发展状况基础上,开发了气动机器人的视觉定位系统,研究了基于 BP 神经网络的摄像机标定及畸变校正技术,针对图像的平滑滤波、阈值分割、特征提取和视觉跟踪轨迹插补等问题进行了详细的讨论。

在摄像机的标定和调整上,通过对摄像机成像的物理模型的分析,利用 BP 神经网络对镜头进行畸变校正,并根据机械手实验台的特点,设计了一种简单易行的网格平行线平面模板标定方法,很好地实现了摄像机标定;在图像预处理方面,本文以实际零件图像为例,利用分段线性变换方法进行了对比度增强,分析了图像检测中 CCD 摄像机的噪声影响要素,利用差值滤波器对图像进行平滑滤波;在图像分割方面,设计了一种改进的自适应阈值分割算法,可以根据图像特征自动分割图像,提高了检测的精度和快速性;在图像特征提取方面,针对数字图像的特征,讨论了基本的特征提取方法,利用获得的图像形心坐标描述目标物体的位置信息;在视觉跟踪和路径规划方面,研究并比较了常用的跟踪理论,采用图像分割求形心和直线轨迹插补的方法实现机械手快速跟踪定位,并对跟踪算法进行了仿真和试验验证;在视觉伺服系统的实现方面,将虚拟仪器技术和机器视觉技术的优点结合起来,利用 LabVIEW 和 NI IMAQ 开发了一套完整的视觉伺服定位系统。

实践证明,将虚拟仪器与机器视觉结合起来开发机器人视觉系统,缩短了系统的开发周期,同时也使系统的灵活性、可靠性和性价比大为提高,为视觉系统的二次开发提供了很好的试验平台。

关键词: 机器视觉; 气动机器人; 虚拟仪器; 路径规划

ABSTRACT

Machine vision is regarded as the most important branch of the research of intelligent robots , virtual instrument is a fresh software flat in the field of the application of computer . Based on the technology of virtual instrument and the study of the status of robot vision in recent years , this thesis design the pneumatic robot vision orientation system . The thesis firstly analyzes the camera calibration and its adjustment for the robot vision . Further we particularly discuss the noise filtering and feature extraction .

The thesis analyses the physical model of camera and the geometrical model of grille imaging in camera plane. According to the feature of 2-D plane image, we presented a method for the camera adjusting that can reduce the camera distort and improve the precision of orientation based on BP Neural Networks . For the image process , the study is illustrated with a hexangular nut experiments relating to the method development and analysis . For the feature extraction, based on the basic shape feature inspection in the work piece inspection, the thesis presents a method to get the center coordinate of the image with the area.

On the aspect of the realization of robot vision servo control system, this thesis based the advantage of the visual programming , combined with the software IMAQ Vision , developed a completed vision servo control system . And the route of robot line interpolation also programmed .

Practice have proved that combined the virtual instrument with the machine vision to design the robot vision system shortened the designing period , and also improved the reliability and the ratio to capability and cost . Moreover , combined the virtual instrument with the machine vision , in term of the logic of the machine vision system to analyse the whole system , explore the way to develop the machine vision . Therefore , it has profundity significance in both project and theoretics .

Key words: machine vision; pneumatic robot; virtual instrument; route programming

湖南大学

学位论文原创性声明

本人郑重声明：所呈交的论文是本人在导师的指导下独立进行研究所取得的研究成果。除了文中特别加以标注引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写的成果作品。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律后果由本人承担。

作者签名： 

日期：2004年 5月 15日

学位论文授权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权湖南大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于

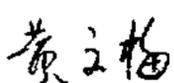
1、保密 ，在 _____ 年解密后适用本授权书。

2、不保密 。

(请在以上相应方框内打“√”)

作者签名： 

日期：2004年 5月 15日

导师签名： 

日期：2004年 5月 15日

第 1 章 绪 论

1.1 气动机器人

1.1.1 工业机器人技术概述

工业机器人由操作机（机械本体）、控制器、伺服驱动系统和检测传感装置构成，是一种仿人操作、自动控制、可重复编程、能在三维空间完成各种作业的机电一体化自动化生产设备^[1]。特别适合于多品种、变批量的柔性生产。它对稳定、提高产品质量，提高生产效率，改善劳动条件和产品的快速更新换代起着十分重要的作用。

机器人技术是综合了计算机、控制论、机构学、信息和传感技术、人工智能、仿生学等多学科而形成的高新技术，是当代研究十分活跃，应用日益广泛的领域。机器人应用情况，是一个国家工业自动化水平的重要标志。

机器人并不是在简单意义上代替人工的劳动，而是综合了人的特长和机器特长的一种拟人的电子机械装置，既有人对环境状态的快速反应和分析判断能力，又有机器可长时间持续工作、精确度高、抗恶劣环境的能力，从某种意义上说它也是机器的进化过程产物，它是工业以及非产业界的重要生产和服务性设备，也是先进制造技术领域不可缺少的自动化设备。

1.1.2 气动技术发展状况及优缺点

气动技术是一门正在蓬勃发展的新技术，用气动元件组成的传动和控制系统已广泛用于国民经济各部门的成套设备和自动化生产线上。气动技术是以压缩气体为工作介质进行能量和信号的传递，从而实现生产过程自动化的一门技术。它包含气压传动和气动控制两方面的内容。

目前，气动和液压是两种较为普遍应用的传动和控制方式，两者有许多相同点，也有许多不同点，气动技术真正为全世界各个工业部门所接受并广泛应用，是由于日益迫切的生产自动化和操作程序合理化的需要。气动技术具有以下许多优点：

- 1、气动技术以空气为工作介质，空气随处可取，且粘性小，在管内流动阻力小，便于集中供气和远距离输送。大多数工厂有方便的压缩空气气源。作为工作介质的压缩空气的物理性质，使气动技术在广泛的各种应用中具有安全、方便和费用低的优点。压缩空气没有产生电火花的安全。因此，适用于易燃、易爆等危险的工况。

2、气动元件结构简单，价格低廉，并可直接向大气排放，处理方便，不必使用回收管道。

3、气动系统清洁，即使有泄漏，也不会像液压系统那样污染产品和环境，也不像电子系统受到电磁干扰。

4、气动系统维护不复杂，也不需要特殊的培训和试验设备。

5、适应性强，现有的机器可方便地改为气动传动，气缸可以直接安装在要求出力的地方。

6、便于进行能量存储，可以进行应急或系统需要用。

7、气压传动本身有过载保护性能。气动执行元件能长期在满负荷下工作，在过载时会自动停止。

8、气动元件运动速度快。普通气缸的运动速度一般为 $0.05\sim 0.7\text{m/s}$ ，有的高达 $1\sim 3\text{m/s}$ ，高速气缸可达 15m/s 。

当然，气动技术也有其缺点：

1、由于空气的可压缩性大，使得工作部件运动速度稳定性差，要获得精确的进给运动是困难的。

2、由于气压传动的工作压力远低于液压传动的工作压力，气动执行元件输出的力小。

3、气动工作介质无润滑性能。气体排放时噪声大，需加消声器。

4、气动信号的传递速度远比电信号低，而且有较大的延迟和失真，因而气动控制技术不宜用于高速传递和处理信息的复杂系统，而且气动信号的传送距离也受限制。

1.1.3 气动机器人（机械手）应用现状及发展趋势

尽管气动技术上有一些缺点，但它的优点还是主要的。所以气动技术能在各个工业部门中得到日益广泛的应用。而气动元件更是一种经济实用的机械化、自动化的理想元件。现在，气动技术和电子电器、液压技术一样，都成为自动化生产过程的有效技术之一，在国民经济建设中起着越来越大的作用。气动技术由风动技术及液压技术演变，发展而成为独立的技术门类不到 50 年，却已经充分显示出它在自动化领域中强大的生命力，成为二十世纪应用最广，发展最快，也最容易接受及重视的技术之一。气动技术已成为各个行业不可缺少的一部分。在国外，气动被成为“廉价的自动化技术”。微电子加气动技术的应用已成为现代化的一个重要标志。

据统计，在工业发达国家中，全部自动化流程中约有 30% 装有气动系统^[2]。90% 的包装机、70% 的铸造和焊接设备、50% 的自动操作机、40% 的锻压设备和洗衣设备、30% 的采煤机械、20% 的纺织机械、制鞋业、木材加工、食品机械、

43%的工业机器人有气压系统。美、日、德等国的气动元件销售平均每年增长超过10%~15%。许多工业发达国家的气动元件产值已接近液压元件的产值，且仍以较大的速度发展。

为了适应这一形势的发展，近年来在气动控制技术上特别引人注目的发展趋势之一是通过气动技术与电子技术以及机械技术的有机结合，以及周边技术的成熟，在工业自动化领域里，气动机械手、气动机器人的实用性已经充分体现出来。因为气动伺服定位技术一出现，就受到工业界和学术界的高度重视，为气动机器人、气动机械手大规模进入工业自动化领域开辟了十分宽广的前景^[3]。由于大量采用传感器，使气动元件智能化，国外发展的气动定位系统，采用反馈控制，速度在10m/s，行程为300mm时，定位精度可达±0.01mm，国外还开发了新型智能型电磁阀，可直接接受传感器信号，并自行完成动作，达到控制的目的，以应用于物件的传输，并能识别搬运物的大小。广泛采用总线连接集成系统，提高气动控制系统的可靠性，简化安装和维修。

1.2 虚拟仪器技术

在纷繁的测试产品中，虚拟仪器（Virtual Instrument，简称VI）已经成为越来越多测试人员的最佳选择，这是因为VI系统能更迅捷、更经济、更灵活地解决测试问题。随着VI驱动程序标准化及软件开发环境的发展，代码复用成为仪器编程中的基础，这意味着可以避免仪器编程过程中的大量重复性劳动，从而大大缩短复杂程序的开发时间，可以用不同的模块构造自己的VI系统，这不仅会节省大量人力物力，而且已有的测试投资在未来仍能得到可靠保护。

1.2.1 虚拟仪器的概念

虚拟仪器通过应用程序将通用计算机与仪器硬件结合起来，用户可以通过友好的图形界面（通常叫做虚拟前面板）操作这台计算机，就像在操作自己定义、自己设计的一台单个传统仪器一样^[4]。VI以透明的方式把计算机资源（如微处理器、内存、显示器等）和仪器硬件（如A/D、D/A、数字I/O、定时器、信号调理等）的测量、控制能力结合在一起，通过软件实现对数据的分析处理、表达以及图形化用户接口。

应用程序将可选硬件（如GPIB、VXI、RS-232、DAQ板）和可重复用原码库函数等软件结合在一起，实现了仪器模块间的通信、定时与触发等。原码库函数为用户构造自己的VI系统提供了基本的软件模块。由于VI的模块化、开放性和灵活性，以及软件是关键的特点，当用户的测试要求变化时，可以方便地由用户自己来增减硬、软件模块，或重新配置现有系统以满足新的测试要求。这样，当用户从一个项目转向另一个项目时，就能简单地构造出新的VI系统而不丢弃已

有的硬件和软件资源。

1.2.2 虚拟仪器的特点

VI 与传统仪器的比较见表 1.1，其最主要的区别是 VI 的功能由用户使用时自己定义，而传统仪器的功能是由厂商事先定义好的。有人把功能确定的计算机插卡式仪器也笼统地叫做 VI，这有悖于 VI 与传统仪器本质的区别。由此可见，没有面向科学家与工程师的图形化编程平台就很难谈得上广泛普及 VI。

表 1.1 VI 与传统仪器的比较

VI	传统仪器
软件使得开发与维护费用降至最低	开发维护开销高
技术更新周期短（1~2 年）	技术更新周期长（5~10 年）
关键是软件	关键是硬件
价格低、可复用与可重配置性强	价格昂贵
用户定义仪器功能	厂商定义仪器功能
开放、灵活，可与计算机技术保持同步	封闭、固定
与网络及其它周边设备方便互联的面向应用的仪器系统	功能单一，互联有限的独立设备

1.3 机器人视觉系统与虚拟仪器技术

1.3.1 智能机器人与机器视觉

智能机器人的“智能”特征在于它具有通过各种传感器（视觉、触觉、接近觉、测距传感、力/力矩传感、滑觉、声觉等）与外部世界（包括工作对象、工作环境和人）相协调工作的机能。“让机器人学会感知它们所触到的部件”已成为机器人发展的趋势。因此，机器人视觉系统的研究一直是各类智能机器人研究的一个热点^[5]。

如果想要赋予机器人较为高级的智能，那么离开视觉系统是无法做到的。第一代工业机器人只能按照预先规定的动作往返操作，一旦工作环境变化，机器人就不能胜任工作。这是因为第一代机器人没有视觉系统，无法感知周围环境和对象的情况，因此对于智能机器人来说，视觉系统是不可缺少的。

从六十年代开始，人们着手研究机器视觉系统。一开始，视觉系统只能够识别平面上的类似积木的物体。到了七十年代，已经可以认识某些加工部件，也能认识室内的桌子，电话等物品了。当时的研究工作虽然进展很快，但却无法用于实际。这是因为视觉系统的信息量极大，处理这些信息的硬件系统十分庞大，花费的时间也很长。

机器人视觉是一门新兴的发展迅速的学科，八十年代以来，机器人视觉的研究已经历了从实验室走向实际应用的发展阶段。从简单的二值图象处理到高分辨率多灰度的图象处理，从一般的二维信息处理到三维视觉机理以及模型和算法的研究都取得了很大的进展。而计算机工业水平的飞速提高以及人工智能、并行处理和神经网络等学科的发展，更促进了机器人视觉系统的实用化和涉足许多复杂视觉过程的研究。目前，机器人视觉系统正在广泛地应用于视觉检测、机器人的视觉引导和自动化装配领域中。

1.3.2 机器视觉与虚拟仪器的结合

机器视觉与虚拟仪器结合的必要性是显而易见的。传统的基于专用硬件来实现的机器视觉系统由于硬件的价格昂贵、开发方法难度大（一般用汇编语言来编程）因此开发周期和开发的成本非常高，这与市场竞争的日益加剧格格不入。而机器视觉和虚拟仪器强大的开发工具和低廉的成本使大多数传统机器视觉系统必将为其所替代。机器视觉注重对图像进行分析并作出控制外部控制机构动作的决定，但是任何所做的决定，最终都用于指导外部的运动控制模块工作。虚拟仪器除了具备较强的分析功能外还增强了控制功能，将二者结合就可以使机器视觉的分析功能和虚拟仪器的控制功能同时为系统所用，而机器视觉和虚拟仪器各自的灵活性、可靠性、低成本必将给予系统很强的灵活性和可靠性，同时也使系统有很高的性价比。

机器视觉软件与虚拟仪器开发平台的结合使机器视觉传统的实现模式（如基于 PLC 和专用硬件实现方式）向基于 PC（PC Base Machine vision）的实现模式逐步转化。这种转化的可能性和原因有以下几个方面：

1、计算机技术、多媒体技术使基于 PC 实现机器视觉系统称为可能性，也使得机器视觉系统的价格进入平民化时代。过去，由于机器视觉系统需要对大量的数据进行处理，而计算机的处理能力远远不能达到要求，因此大部分机器视觉

表 1.2 PCI 总线与其它传输介质的传输速率

总线类型	传输速率 (MB/S)
ISA	3-5
EISA	33
PCI	95-132
Full Montion Video	2-11/window
SVGA	30-40
Hard Disk(SCSI)	4-20
10-100Mbps Ethernte	2/20

系统都采用专用的硬件来实现，这些硬件往往由专用 DSP 芯片构成，价格非常

昂贵。而到今天，计算机的处理能力已经达到了很高的水平，如 Pentium 系列处理器的处理能力已经达到 1G 以上。这使基于多媒体计算机来实现机器视觉系统成为可能，而计算机高的性价比也使得机器视觉系统的价格让用户很容易接受。

2、局部总线技术使基于 PC 的视觉系统有更强的数据处理能力。表 1.2 列出 PCI (Peripheral Component Interconnect) 总线与其他几种传输介质传输速率的比较。由表 1.2 可知 PCI 总线的带宽最高，因此，PCI 总线日趋代替了 ISA (Industry Standard Architecture) 总线。不仅如此，PCI 总线还允许设备以主模式 (Master) 和从模式 (Slave) 两种方式与之连接，以主模式与之连接的设备可以不经 CPU 的干涉直接将大量的数据转入内存，这样系统就可以为 CPU 留出足够的时间来处理其他工作，机器视觉系统的速度也会因此得到提高。机器视觉中很多设备都可以主模式工作在 PCI 总线上，如 NI 的图像采集卡 PCI-1407 和 Data Translation 公司的 DT3133 系列图像采集卡。

3、32 位的 Windows NT/2000 操作系统和全方位的机器视觉软件，使机器视觉系统的开发更快速可靠。过去，开发机器视觉系统时，往往要为某一处理过程编写专用程序，使系统的开发要花费大量的人力和财力，从而系统的价格也很昂贵。近几年，Windows 操作系统，以其稳定性、可靠性和安全性赢得了用户的欢迎，而基于 Windows 平台的机器视觉软件近几年日趋完善，这些软件包括了几乎所有机器视觉开发应用到的算法，使用时可以直接拿来使用，或者将自己的算法和这些已有的算法集成到一起。这样开发者就可以将更多的时间集中到系统功能的研究上而不必拘泥于某一算法或过程，开发周期大大缩短，系统的价格也会因此降低。NI IMAQ vision6.0 包括图像滤波、识别、几何变换、形状匹配、斑点分析和测量等近 400 种完整的算法，同时还利用了多媒体技术对这些算法进行优化，其性能要比不具有 MMX 技术的奔腾处理器提高 4 倍。

综上所述，将虚拟仪器和机器视觉结合起来不仅是必要的同时也是可行的。同时由于虚拟仪器技术得到众多商家的支持，这两种技术的结合不仅可以使技术各自的优势得到充分发挥，同时也为缩短项目的开发周期、提高项目的可靠性和性价比作出了保证。因此将虚拟仪器与机器视觉结合起来开发机器人视觉系统，必将使系统的开发周期缩短，同时也可以使系统的灵活性、可靠性和性价比大为提高。

1.4 本文主要研究内容

本文提出了一种基于虚拟仪器的机器人实时视觉伺服控制系统，该系统由气动机械手，CCD 摄像机，以及控制目标运动系统组成。该系统把机器人视觉和机器人控制紧密地结合起来，基于图象实时生成机器人执行机构的运动轨迹，以达到实时控制机器人跟踪目标的目的。

本文将从以下几个方面进行论述：

1、气动机械手总体方案及其视觉平台

结合气动机械手的系统结构和组成原理，对其进行了运动学建模分析，研究电气比例控制的性能和特点，利用虚拟仪器技术构建机械手视觉平台，通过视觉技术识别不同形状和尺寸的目标物体，反馈机械手与目标物体的位置偏差，引导机械手朝缩小偏差的方向前进，实现自动准确抓取物体。

2、摄像机标定与畸变校正

通过建立摄像机成像和畸变模型，研究摄像机标定和畸变校正技术，利用神经网络进行畸变校正，采用网格式平行线模板标定了像素的水平当量和垂直当量。

3、物体识别理论与实验

主要研究待识别物体图像的预处理技术、分割方法和识别特征等，包括对比度增强，噪声滤波，图像分割，并通过实验比较图像处理的效果，针对处理后的二值化图像，设计了用于识别目标物体位置的特征参量，经过坐标变换，将这些特征量转化为空间坐标中的实际位置，以控制机器人完成相应的操作。

4、视觉跟踪理论与路径规划

根据图像处理反馈的物体位置信息，通过跟踪算法规划机械手的运动轨迹，通过逆运动学分析，转化成运动气缸的目标位移，实现机械手跟踪控制。

5、机械手视觉系统软件实现策略

将虚拟仪器技术和机器视觉技术的优点结合起来，利用 LabVIEW 和 NI IMAQ 开发机械手视觉系统软件，由软件实现图像处理和跟踪控制功能，建立起功能完善、操作简单的机械手视觉平台。

第 2 章 气动机械手总体方案及视觉平台

机械手是自动化生产线最常见的设备。本章介绍抓取气动机械手的系统结构、控制原理和运动学模型，在原有气动机械手系统的基础上增加摄像机和采集卡，配以识别与控制软件，构建机械手视觉系统，自动对准并抓取目标物体。

2.1 气动机械手系统原理

2.1.1 气动机械手基本结构及其工作原理

机械手是机电系统典型的设备之一，可在空间抓放物体，动作灵活多样，适用于可变换生产品种的中、小批量自动化生产，广泛应用于柔性自动线。本文研究的自动抓取气动机械手，它以压缩气体为动力源，通过编程控制，可以实现简单形状的目标物体抓取动作。气动机械手系统原理如图 2.1 所示。

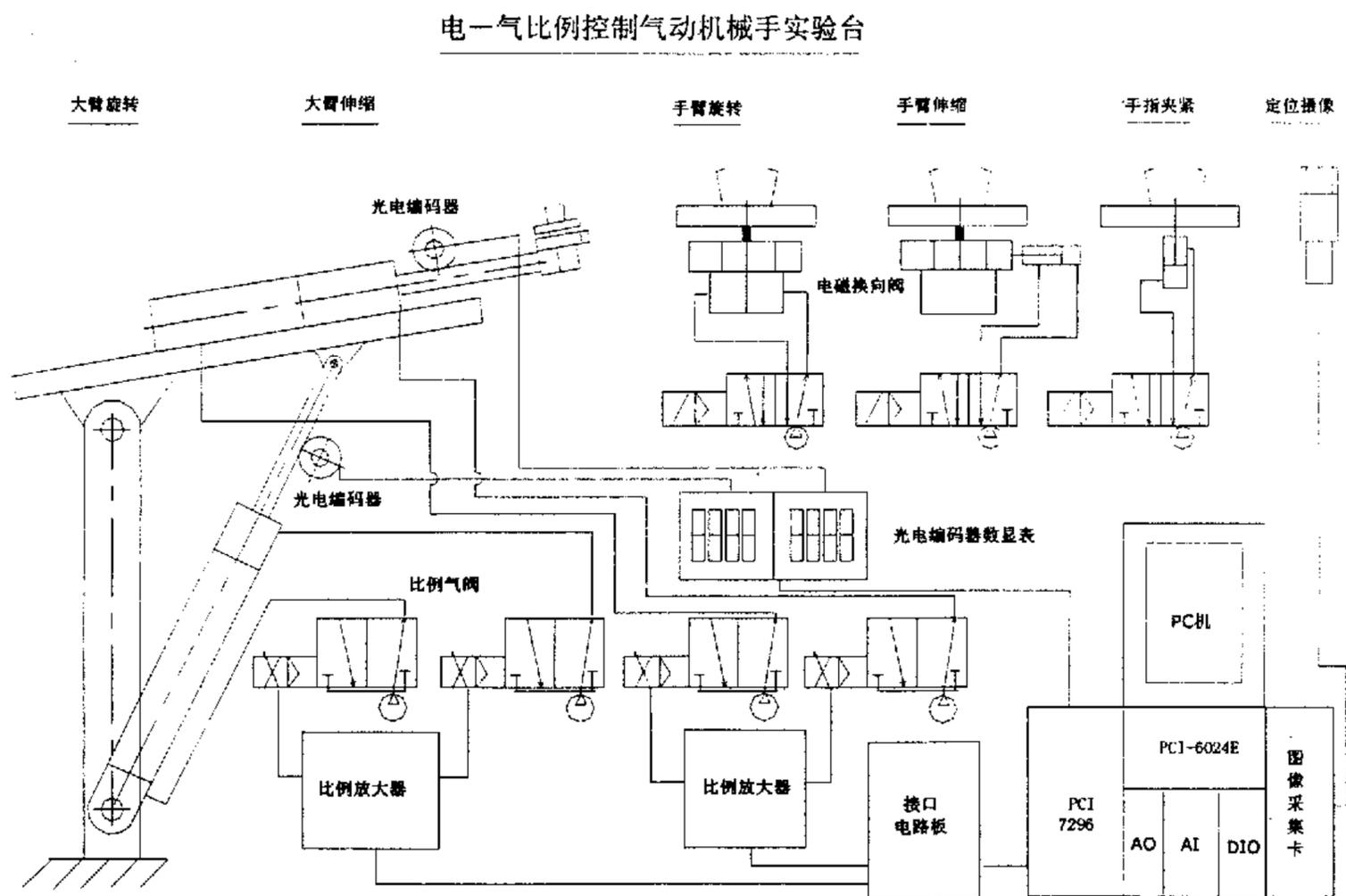


图 2.1 气动机械手总体方案

气动机械手机械系统由机架、大臂、小臂、手腕和手指等部分组成，图 2.2 为气动机械手的组成原理图。大臂 2 包含两个运动：绕铰接中心的摆动和沿大臂方向的伸缩，摆动由气缸 2 的伸缩实现，气缸 2 的缸体铰接在机架上，活塞端与

大臂铰接，气缸伸缩带动大臂摆动，实现机械手上下摆动；大臂伸缩运动由气缸 4 实现，气缸 4 的缸体固定在大臂上，活塞端与小臂连接，缸体伸缩带动机械手左右移动；小臂是由一个伸缩旋转组合缸驱动，可以前后伸缩和旋转，使机械手接近物体，保持正确的夹持姿态；手指在夹紧气缸的驱动下，夹紧放开物体。

图 2.3 为机械手的运动简图，抓取物体的工作原理为：首先大臂摆动气缸 2 动作，带动大臂 3 绕铰接点摆动，然后大臂伸缩气缸 4 动作，带动机械手臂作直线移动，对准二维平面内任意位置的物体，最后小臂 5 气缸动作，手腕转动，手指 6 夹紧，完成夹取目标物体的动作。

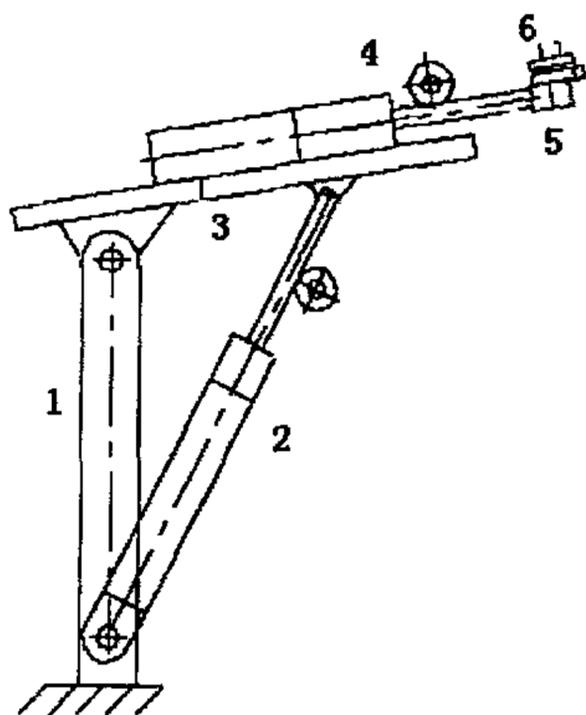


图 2.2 气动机械手机械原理图

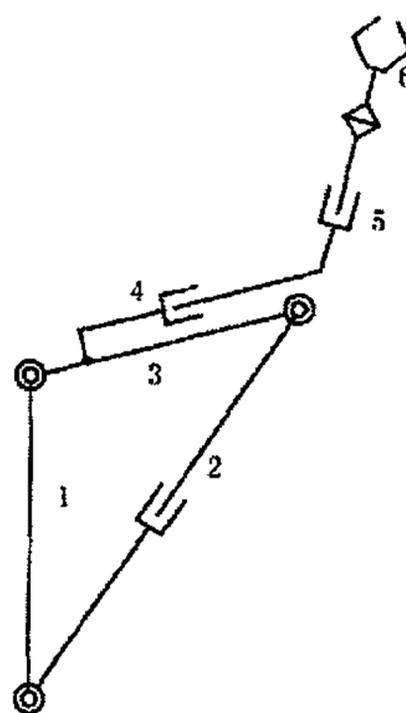


图 2.3 气动机械手运动简图

2.1.2 气动机械手控制原理

本文研究的机械手的驱动采用了气动控制技术，气动控制具有性能稳定，可靠性高，具有本质安全防爆功能，不受电磁场干扰，结构简单，维护方便，清洁无污染等优点，在工业机械手和物料输送系统有较多应用。

气动控制系统一般由控制器、电—气控制元件、气动执行元件、传感器和接口电路等组成^[6]。控制器是气动系统的大脑，负责信号采集和控制指令的生成，通常采用计算机、单片机或可编程控制器等。电—气控制元件负责将控制器生成的指令信号转换成气动信号，对气流的压力、流量和方向进行调节，一般由电—气伺服阀、电—气比例阀和电—气开关阀等组成。气动执行元件将气动能转换成机械能以驱动工作机构的元件，包括气缸、气爪或者气动马达等。气动控制系统传感器负责将执行元件的状态反馈给控制器，常用的一般有位置传感器、压力传感器或速度传感器等。接口是指控制器与控制元件、控制器与传感器之间的信号匹配电路，包括放大、滤波、调制解调和功率驱动等，可以实现不同电信号之间

的相互转换。气动控制系统的组成原理如图 2.4 所示。

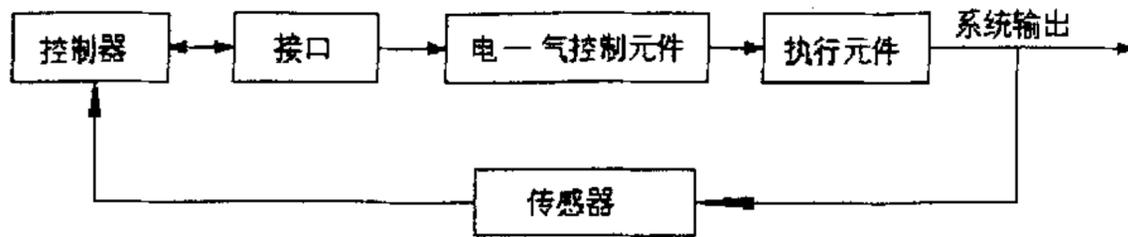


图 2.4 气动控制系统组成示意图

气动控制系统可按控制元件来分类^[7]，分为伺服控制系统、比例控制系统和开关控制系统。电-气伺服系统以气动伺服阀为控制元件，工作频宽较高，响应速度快，性能最好，但是成本昂贵，在一般的工业场合较少应用。电-气比例控制系统是以电-气比例阀为主要控制元件的气动控制系统。电-气比例阀采用了比例电磁铁技术，相对于伺服阀要低廉很多，频宽一般为 10Hz~25Hz 左右。具有高精度电反馈的电气比例控制系统，在控制精度上已经很接近电气伺服控制系统。同时电-气比例控制系统控制简单，系统成本低、有利于简化机械结构、提高控制品质和智能化程度，可以达到精密控制、自动化和节能的目的，在气动领域得到了广泛的应用。开关阀系统以高速开关阀作为控制元件，与计算机接口方便，成本最低，但开关阀系统存在噪音大、开关阀寿命不足等问题，近年来发展较慢，一直处于实验研究阶段。

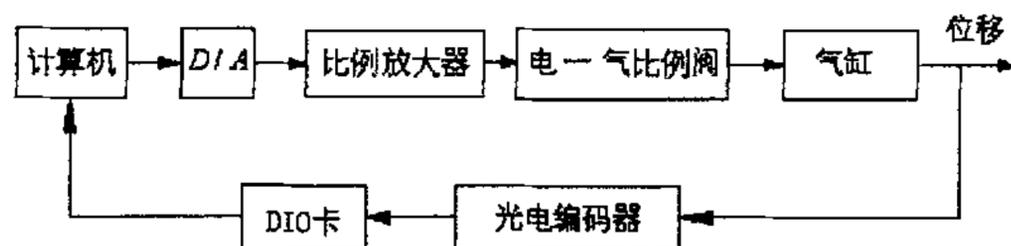


图 2.5 电-气比例阀系统示意图

本文研究的机械手采用了电-气比例控制技术实现大臂的摆动和伸缩，其控制系统原理如图 2.5 所示。机械手动作时，计算机根据目标位移和传感器反馈的气缸当前位置，通过相应的算法生成位移控制量，经 D/A 转换和比例放大后，控制电-气比例阀的阀口开度，高压气流驱动气缸产生位移，光电编码器时刻检测气缸位移并反馈给计算机，达到目标位移控制气缸停止动作。

机械手的电-气比例控制回路如图 2.6 所示。气缸 2 完成大臂摆动运动，其型号为 L-CDG1BN50-750，气缸 4 完成大臂伸缩运动，其型号为 L-MDBB63-550，两个气缸的控制阀采用日本 SMC 公司的比例流量型控制阀 VEF3121，模数转换卡选用 NI 公司的 PCI-6024E，位移传感器使用光电编码器，控制方法为连续无级控制。

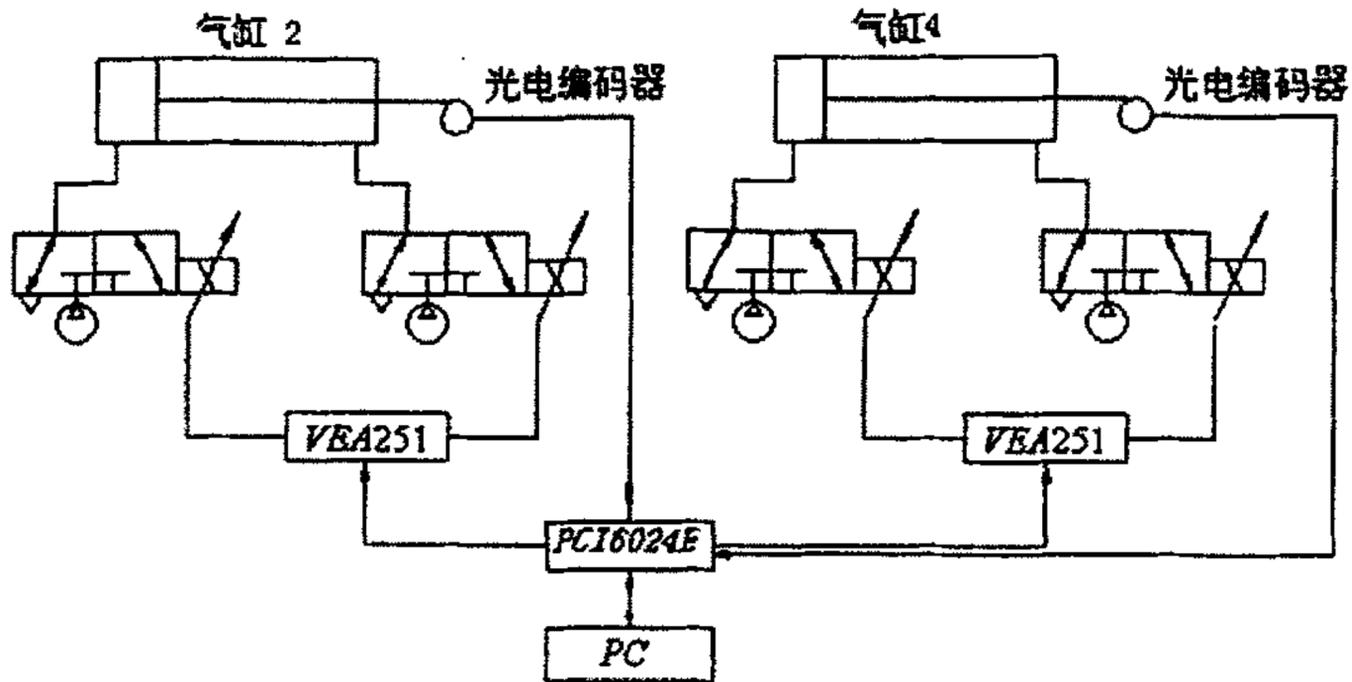


图 2.6 机械手大臂气动控制回路

2.1.3 电—气比例控制系统建模分析

机械手电—气比例位置控制的模型如图 2.7 所示，系统由数模转换卡、比例放大器、电—气比例流量阀、气缸、光电编码器、DIO 卡等组成，构成了闭环控制回路。

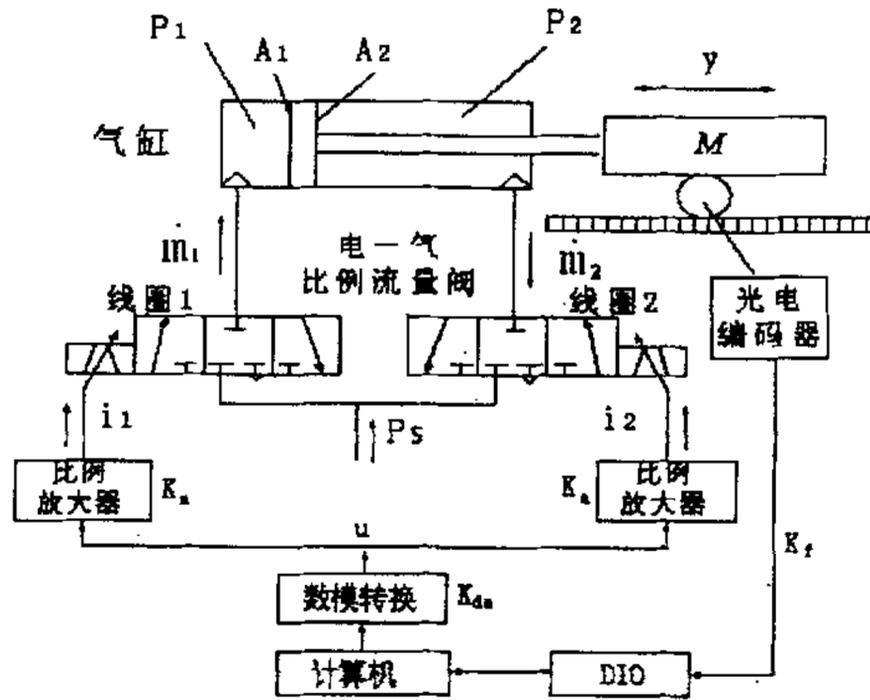


图 2.7 机械手电—气比例位置控制模型

(1) 数模转换卡 PCI-6024E

PCI-6024E 的功能是将数字控制量信号转换成模拟电压信号，相当于一个比例环节，其模型为：

$$K_{da} = 5 \text{ v} / 2048$$

(2) 比例放大器

比例放大器的作用是将计算机发出的电压信号经过比例放大,转换成电流信号,用来控制比例流量阀的开度。比例系数为 K_a 。

(3) 电气比例流量阀

电气比例流量阀采用电磁比例电磁铁控制阀口开度,使通过阀的气体质量流量和阀的开度成一定的关系。当线圈电流越大,阀的开度越大,通过阀的气体质量流量也越大。其数学模型为:

进气腔:

$$\dot{m}_1 = \sqrt{K_{x1}(X_v)P_1 + K_{x2}(X_v)} \quad (2.1)$$

排气腔:

$$\dot{m}_2 = \sqrt{K_{x1}(-X_v)P_2 + K_{x2}(-X_v)} \quad (2.2)$$

其中, \dot{m}_1 和 \dot{m}_2 为质量流量, X_v 为阀芯位移, K_{x1} 、 K_{x2} 是与阀芯位移 X_v 的系数,或者说是阀工作电流的函数,可以写成

$$\begin{aligned} K_{x1} &= f_1(i) \\ K_{x2} &= f_2(i) \end{aligned} \quad (2.3)$$

在实际应用中, K_{x1} 、 K_{x2} 由实验测得。

(4) 气缸

气缸的运动方程为

$$M\ddot{y} = P_1A_1 - P_2A_2 - F_R - F_L \quad (2.4)$$

式中,

M ——运动部分质量;

F_L ——负载力;

F_f ——摩擦力;

P_1, P_2 ——分别为气缸进、排气腔绝对压力;

y ——活塞位移;

A_1, A_2 ——分别为气缸进、排气腔有效面积。

气缸进气腔

$$\dot{P}_1 = \frac{k}{V_{10} + A_1y} (RT\dot{m}_1 - P_1A_1\dot{y}) \quad (2.5)$$

气缸排气腔

$$\dot{P}_2 = \frac{k}{V_{20} + A_2(L-y)} (P_2A_2\dot{y} - RT\dot{m}_2) \quad (2.6)$$

式中, V_{10}, V_{20} ——分别为气缸进、排气腔余隙空间;

k ——绝热指数;

R ——气体常数；

T ——气体的热力学温度；

L ——气缸总有效长度；

(5) 光电编码器和数显表

光电编码器的功能是将角位移转换为电脉冲，每转产生 1024 个脉冲，这也相当于一个比例环节，模型参数为

$$K_f = 1024 \text{ 个脉冲/滚轮周长}$$

(6) 电—气比例位置控制系统线性化数学模型

机械手电—气比例位置控制模型是上述环节的串联，相关参数可以通过实验或者计算得到，机械手电—气比例位置控制模型的线性化数学模型为：

$$G(s) = \frac{1054}{s^3 + 20.56s^2 + 732.4s} \quad (2.7)$$

可见，本文研究的气动机械手具有积分环节，对阶跃输入的理论稳态误差为零，但实际系统存在死区、延迟等非线性因素，因此机械手还需采用 PID 及其变形方式进行运动控制，才能达到预期的动态性能。

2.1.4 气动机械手运动分析

气动机械手的运动包括大臂摆动、大臂伸缩、小臂伸缩、手腕转动和手指夹持等，其中大臂运动通过电—气比例控制实现，小臂、手腕和手指的运动比较简单，只有两个状态，通过开关量控制，这里仅对机械手大臂的运动过程进行分析。

机械手大臂的运动简图如图 2.8 所示。AO 段表示机架，AB 段表示机械手大臂，线段 AO、AB 的长度为安装尺寸，是定值。AC 为大臂伸缩气缸，OB 为大臂摆动气缸。

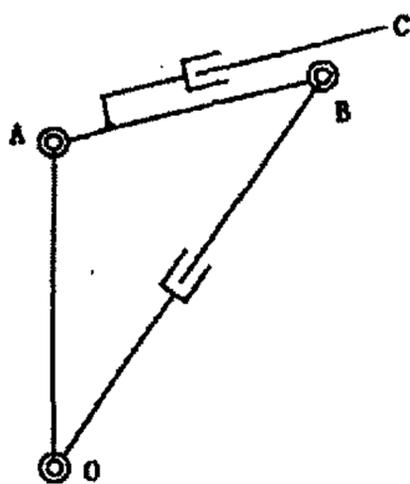


图 2.8 气动机械手大臂运动简图

机械手的末端位置 C 的运动是两个气缸运动的合成，要让气动机械手在其活动范围内快速定位，需要研究其运动合成。由于机械手立柱不是垂直安装，建立两个直角坐标系来研究 C 点的运动：一是物件台架坐标系 X'—Y'，另一个是

机械手臂运动坐标系 $X-Y$ 。两个坐标系的夹角为 $\alpha=10^\circ$ ，如图 2.9 所示。设 $OA=l_1$ ， $OB=l_2$ ， $AB=l_3$ ， $AC=l_4$ ， θ 、 β 分别为气缸 AC 、 OB 的水平倾角，端点 C 在坐标系 $X'-Y'$ 中的坐标为 (x',y') ，则 C 在 $X-Y$ 坐标系中的坐标 (x, y) 为

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

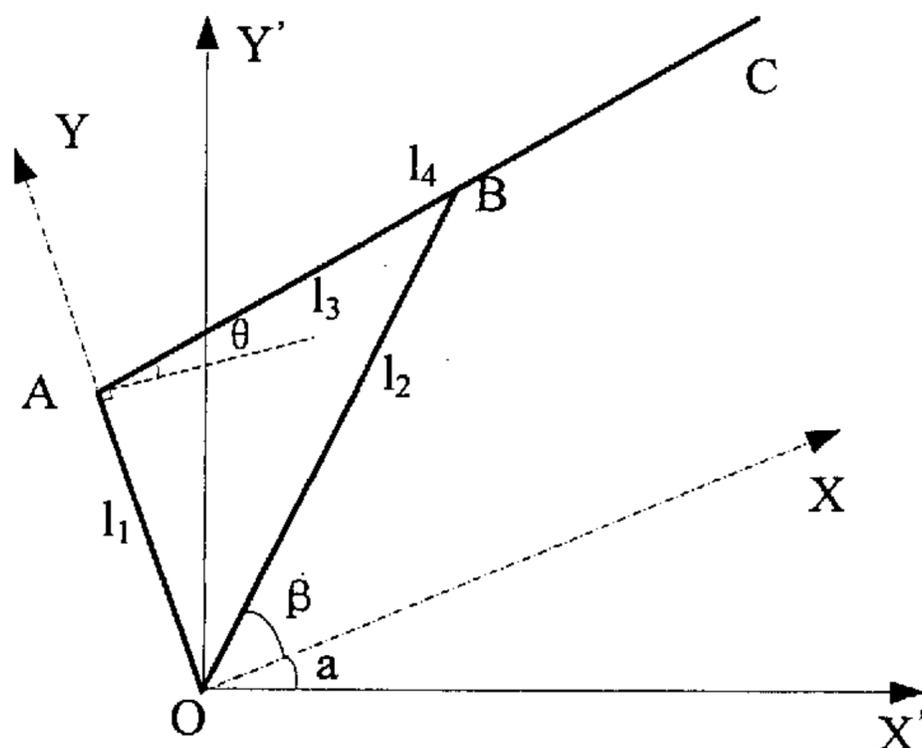


图 2.9 气动机械手的运动分析图

则有

$$\begin{cases} x = l_4 \cos \theta \\ y = l_4 \sin \theta + l_1 \end{cases} \quad (2.8)$$

对三角形 AOB ，由余弦定理得到

$$\cos(90^\circ + \theta) = \frac{l_1^2 + l_3^2 - l_2^2}{2 \cdot l_1 \cdot l_3} \quad (2.9)$$

$$\cos(90^\circ - \beta) = \frac{l_1^2 + l_2^2 - l_3^2}{2 \cdot l_1 \cdot l_2} \quad (2.10)$$

则

$$\sin \theta = -\frac{l_1^2 + l_3^2 - l_2^2}{2 \cdot l_1 \cdot l_3}, \cos \theta = \sqrt{1 - \left(\frac{l_1^2 + l_3^2 - l_2^2}{2 \cdot l_1 \cdot l_3} \right)^2} \quad (2.11)$$

代入(2.8)得

$$\begin{cases} x = \frac{l_4}{2 \cdot l_1 \cdot l_3} \sqrt{4 \cdot l_1^2 \cdot l_3^2 - (l_1^2 + l_3^2 - l_2^2)^2} \\ y = l_1 - \frac{l_4}{2 \cdot l_1 \cdot l_3} (l_1^2 + l_3^2 - l_2^2) \end{cases} \quad (2.12)$$

由上式可知，若已知气缸 OB 和 AB 的位移 l_2 和 l_4 ，就可求得机械手的运动轨迹。

反之，在知道了机械手的目标位置以后，需要进行逆运动学建模分析，计算出为了达到目标位置气缸 OB 和 AB 的期望位移。

由式(2.8)，得

$$\theta = \arctan\left(\frac{y-l_1}{x}\right) \quad (2.13)$$

$$l_4 = \frac{x}{\cos\theta} \quad (2.14)$$

由式(2.9)得

$$l_2 = \sqrt{l_1^2 + l_3^2 + 2 \cdot l_1 \cdot l_3 \sin\theta} \quad (2.15)$$

已知气缸 AC 的原始长度为 700mm，OB 的原始长度为 1200mm，将 $l_1=1370\text{mm}$ ， $l_3=320\text{mm}$ 代入(2.5)和(2.6)，得

$$\theta = \arctan\left(\frac{y-1370}{x}\right) \quad (2.16)$$

$$l_1' = \frac{x}{\cos\theta} - 700 \quad (2.17)$$

$$l_2' = \sqrt{1979300 + 876800 \sin\theta} - 1200 \quad (2.18)$$

式中 l_1' 、 l_2' 为气缸所需要的行程，单位为毫米。

将 l_1 、 l_3 代入 (2.10)，已知 $1200 \leq l_2 \leq 1500$ ，计算得

$$76.5^\circ \leq \beta \leq 78.5^\circ$$

在机械手运动过程中，可以简化计算过程，对三角形 AOB，由余弦定理得

$$\frac{l_2}{\sin(90^\circ + \theta)} = \frac{AB}{\sin(90^\circ - \beta)} \quad (2.19)$$

则有

$$l_2 = \frac{AB}{\cos\beta} \cos\theta \quad (2.20)$$

式(2.18)可以替换为

$$l_2' = \frac{320}{\cos\beta} \cos\theta = 331.263 \cos\theta \quad (2.21)$$

2.2 机械手视觉平台设计

2.2.1 目标物体位置自动识别

气动机械手抓取目标物体过程中,机械手的末端坐标与姿态通常根据机械手运动学关系求出,由于机械手各关节存在制造时引入的加工误差,传动机构存在间隙,各部件的变形以及计算误差等,这样计算的坐标与姿态难免含有误差,机械手末端一般不能十分精确地对准目标物体,需要根据目标物体位置进行微调,因此正确识别目标物体位置是准确抓取的关键。在开环控制中,目测起到反馈作用,观察并估计目标物体的位置,从而确定每个气缸的位移,在接近物体时调整机械手末端的姿态,实现准确抓取。在闭环控制中,为了实现自动抓取,需要采用目标物体位置自动识别技术,在机械手运动过程中,能够根据目标物体位置信息调整运动方向,实现精确定位。

目标物体位置自动识别系统可分为接触式与非接触式两大类,非接触式应用较多。非接触式识别方法有电磁反射、激光、红外线、超声波、光学和视觉传感器等。早期的目标物体位置自动识别系统多用电容式或电感式邻近效应测距仪,这些传感器精度高但测距短,主要用于位置固定、形状确定的目标物体识别;超声波设备也是一个选择,但是精度不高,且易受环境噪音影响;光电、磁电接近开关能提供是否对准的信息,但不能反映位置偏差大小,一般需要在目标物体附近安装传感器或者利用目标物体的某些物理特性;另外还有将发光二极管装在目标上以便于识别,或者利用激光扫描以确定目标的方位,还有些光学系统采用激光雷达或激光测距仪来测量目标的距离与方位^[6~8],这些方法又都过于昂贵。

随着工业摄像技术和图像处理技术的发展,计算机视觉识别技术得到了越来越多的应用。视觉识别技术模仿人的视觉原理,使用摄像机和图像采集卡等设备,利用图像识别技术进行测量定位^[8],不但能够识别不同形状和尺寸的目标物体,而且可以反馈目标物体位置与机械手位置的偏差,引导机械手朝缩小偏差的方向前进,直至对准目标物体。因此本文的气动机械手采用了视觉技术识别目标物体位置。

2.2.2 气动机械手视觉系统原理

在原有气动机械手控制系统的基础上增加了摄像机和采集卡,配以识别与控制软件,就构成了机械手视觉控制系统,其组成原理如图 2.10 所示。

摄像机安装在手指上方,摄取目标物体信息,采集卡采集目标物体图像后送入计算机进行处理,计算机通过图像处理获得目标物体的空间位置和形状信息,并通过机械手运动轨迹控制算法计算出机械手各个气缸的运动量,送入 PID 位置控制器,计算出目标控制量,由 D/A 变换成模拟控制电压送出,经比例放大

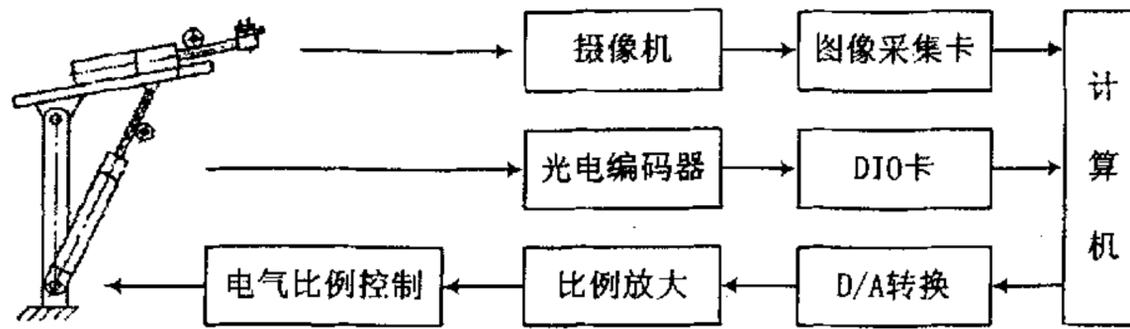


图 2.10 气动机械手视觉系统结构框图

器转换为控制电流，在控制电流的作用下，电气比例阀阀芯移动，控制流向气缸的流量和方向，推动气缸动作，使机械手接近目标位置，光电编码器检测气缸的实际位移并反馈给计算机实现负反馈的位置闭环控制，当计算机判断机械手对准目标物体后，抓取目标物体并且把目标物体移走。

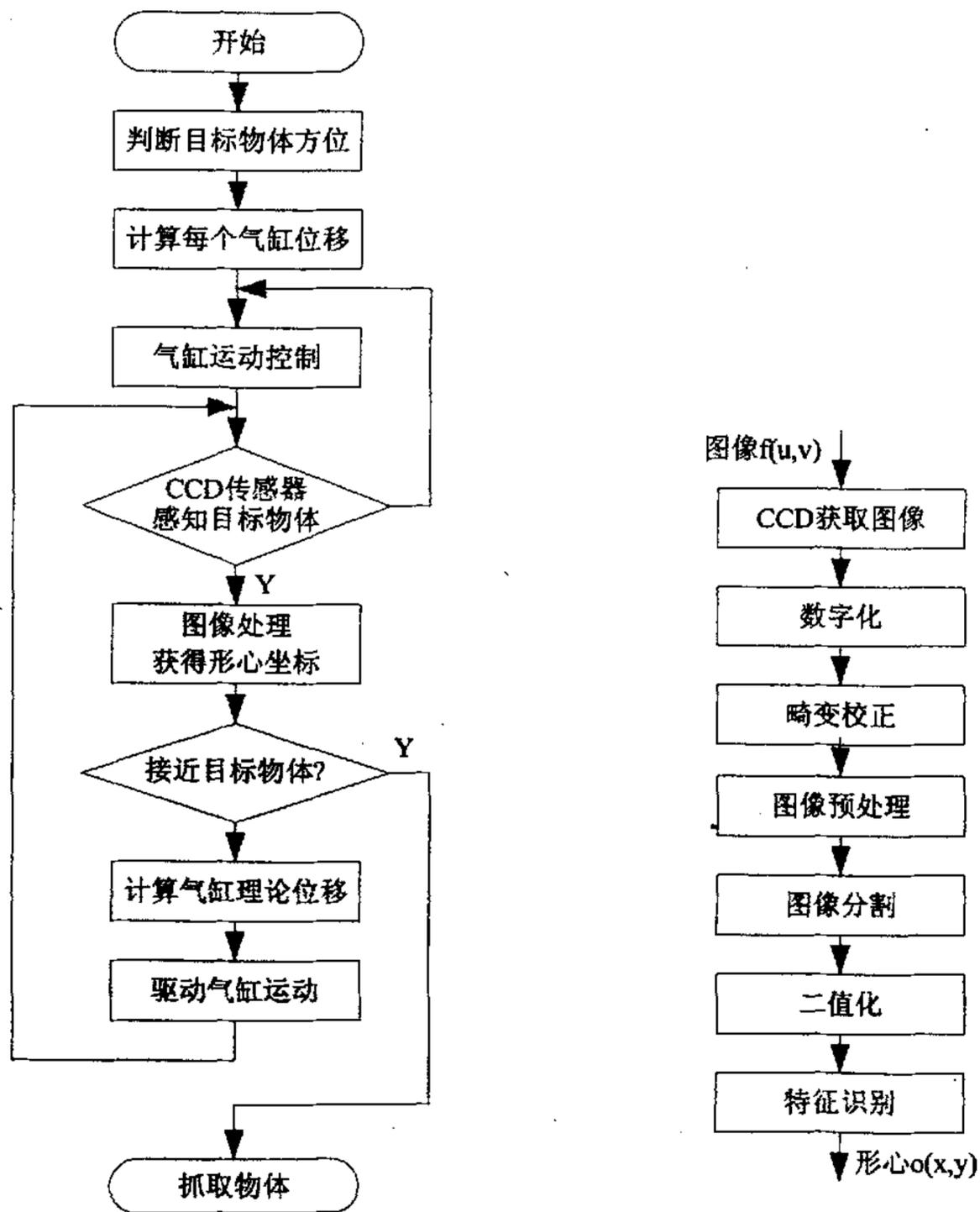


图 2.11 机械手视觉伺服系统流程图

2.12 目标物体图像识别流程

视觉处理和控制在用一台 586 微型计算机实现，主频 166MHz，它从图像输入

接口接收图像采集卡传来的数字信号,经过二值化处理和分割后,判断目标物体位置并将机械手的位置偏差转换成实际位移,然后通过逆运动学分析,求出机械手两个气缸需要运动的距离并控制气缸运动,机械手系统工作流程和物体识别流程如图 2.11 和图 2.12。

识别和控制软件采用 LabVIEW 和 IMAQ vision 开发,主要包括图像预处理,目标物体分割、识别和视觉跟踪等,后面章节将详细介绍。

2.2.3 气动机械手视觉伺服工作流程

带有视觉功能的气动机械手抓取物体的过程分为两个阶段:首先根据物体所在的方位接近物体(人工根据物体柜判断);接近物体后,视觉传感器(CCD 摄像机)获得物体的精确位置,通过相应的控制算法,驱动气缸运动,接近并抓取物体。气动机械手视觉伺服系统的工作流程如图 2.11 所示。首先根据物体方位,估计每个气缸的位移,系统控制气缸运动;接近物体时,摄像机采集目标物体图像,数字化后送入计算机进行处理,将目标物体和背景分离,计算目标物体的形心,得到机械手当前位置相对目标位置的偏差;然后利用逆运动学分析,换算成大臂的旋转角度和平移距离,作为控制目标,送入 PID 控制器生成每个气缸的控制量;经 D/A 转换和比例放大后驱动气缸动作,引导机械手逐渐对准目标物体;目标物体对准后,根据目标物体形状调整手腕姿态,手指张开,包含目标物体后合拢,夹紧并取出目标物体。

2.2.4 系统主要硬件组成

2.2.4.1 图像采集卡

图像采集卡的基本任务是将 CCD 摄像机输出的模拟信号(RS170/CCIR 黑白电视信号, PAL/NTSC 彩色电视信号, S-Video 视频信号等)转换成方便计算机处理的数字信号。图像采集卡的工作过程可以描述为:实时采集 CCD 输出的视频信号,通过内部时序控制 A/D 转换后,存放在采集卡的帧存储器通道中,可以直接送监视器显示,计算机可以通过指令读取帧存储器的图像进行各种变换和处理分析。

图像采集卡的种类很多,按照处理图像颜色的不同,通常可以分为黑白系统、伪彩色系统和真彩色系统,他们分别基于不同的颜色模型。为了方便和图像处理系统连接,采集卡一般做成标准接口形式,例如在通用计算机系统中常用 ISA 和 PCI 接口,直接插在 PC 扩展槽中。根据系统要求,图像采集卡选用 NI 公司的 PCI-1407,它的主要指标如下:

- ◆ 8-bit 的垂直精度(采集的图像具有 256 级灰度);
- ◆ 空间分辨率: 576×768;

此采集卡的驱动程序相当高效、可靠,与开发环境结合紧密,选用此卡可以

在提高系统可靠性的同时缩短开发周期。

2.2.4.2 CCD 摄像头

在气动机械手视觉伺服系统中,摄像机起到相当于人眼睛的作用,其基本任务是摄取目标物体图像,计算机采集和识别后引导机械手抓取目标物体。

摄像机由光学系统、CCD 图像传感器、信号处理电路和机壳等部分组成,其中 CCD 图像传感器是它的核心部件。CCD 即光电荷耦合器件(Charge Coupled Devices),它是一种新型全固体自扫描摄像器件,既可以把光信号转换成电信号,又具有信号电荷的存储、转移和读出功能,能把一幅空间域分布的图象,变换为一系列按时间域离散分布的电信号,在图像传感、信号处理、数字存储等领域内得到了广泛应用。作为摄像器件,CCD 摄像机与其他摄像机相比,有如下优点:

- (1) 体积小、重量轻、功耗低;
- (2) 图像畸变小,测量重复性好;
- (3) 尺寸精度高,具有较高的空间分辨率,可获得较高的测量精度;
- (4) 动态范围大,具有较高的光电灵敏度和较宽的光谱响应频率;
- (5) 抗振动和潮湿,可靠性高,工作寿命长。
- (6) 输出信号易于与计算机连接,进行数字化处理。

根据成像单元的不同结构,CCD 摄像器件可以分为线阵和面阵两种类型,线阵 CCD 价格便宜,行分辨率比较高,一次只能得到一行图像信息,信息量少,可以通过驱动装置循环扫描的得到面阵图像信息;面阵 CCD 价格昂贵,行分辨率比线阵 CCD 低,但是成像信息量大,随着微电子技术的进步,分辨率越来越高,价格也不断下降,因此工业视觉领域常用面阵 CCD 摄像器件。

对于面阵 CCD 摄像机,主要有黑白和彩色两大类。相对于彩色 CCD 摄像机来说,黑白 CCD 摄像机具有价格便宜、高分辨率、低照度等许多彩色 CCD 摄像机所不具有的优点。同时在数字图像处理领域,灰度图像处理技术已经相当成熟,较容易在通用计算机系统上实现,而彩色图像由于信息量庞大,无法满足工业控制的实时性要求,因此在图像处理领域常用黑白面阵 CCD 摄像机。

在选择黑白摄像机的时候,需要考虑象素数、分辨率、最低照度、信噪比、电源等参数,气动机械手的摄像头选用台湾 Life View 出品的 FlyVideoEZ 98,具有以下性能指标和特点:

- ◆ 安装简易、PCI 总线即插即用
- ◆ 视频输入格式:模拟信号/数字信号
- ◆ 提供两路 AV 端子和一路 S 端子输入
- ◆ 高性能,支持 30 帧/秒连续画面捕获和单幅图像抓拍
- ◆ 视频输出格式: AVI/MPEG-1/MPEG-2
- ◆ 图像分辨率(dpi): 576×768

2.3 本章小结

本章首先介绍了气动机械手的系统结构和视觉平台的工作原理,分析了电一气比例控制的性能及其在气动机械手上的实现。在此基础上,对气动机械手进行了运动学分析,得到了机械手的运动学方程和用于控制的逆运动学方程。

为了实现自动抓取,建立了气动机械手视觉平台,通过视觉技术识别不同形状和尺寸的目标物体,反馈机械手与目标物体的位置偏差,引导机械手朝缩小偏差的方向前进,对准目标物体实现准确抓取。本章最后介绍了视觉平台主要硬件的选型。

第 3 章 图像畸变校正与摄像机标定

3.1 概述

本文研究的气动机械手利用视觉原理识别物体位置,为了将图像表示的位置信息转换成物理空间的平面位置信息,首先需要建立图像坐标和空间坐标的对应关系,这就用到了摄像机标定技术。

摄像机标定技术最早应用在摄影测量学中^[9]。摄影测量学中所使用的方法是数学解析分析的方法,在标定过程中利用数学方法对从数字图像中获得的数据进行处理,得到摄像机参数。

常用的摄像机标定技术大体可以归结为两类:传统的摄像机标定方法和摄像机自标定方法。传统的摄像机标定方法是在一定的摄像机模型下,基于特定的实验条件,如形状、尺寸已知的标定参照物,经过对其进行图像处理,利用一系列数学变换和计算方法,求取摄像机模型的内部参数和外部参数。基于主动视觉的自标定技术是通过控制摄像机的运动获取多幅图像来标定摄像机内参数,不需要标定块,但是一般要解非线性方程组,运算量大。

摄像机标定都是基于一定的摄像机模型进行的。摄像机模型是光学成像几何关系的简化,最简单的模型是小孔摄像机模型 (pin-hole model),它是摄像机标定研究的基本模型。然而当图像精度要求较高,或者在近距、广角的情形时,这种线性模型不能准确地描述摄像机的成像几何关系,须引入线性或非线形畸变补偿,利用校正后的模型进行三维重建才能得到更高的精度。

3.2 摄像机模型与图像几何畸变

3.2.1 小孔成像模型

摄像机模型是光学成像几何的简化,小孔透视模型 (pin-hole model) 是最简单最常用的摄像机模型,这是一种理想状态模型,它简单实用而不失准确性。

如图 3.1 所示,视点为 O ,空间点 $P(X_w, Y_w, Z_w)$,在摄像机坐标系下的坐标为 (X, Y, Z) ,它在摄像机成像平面上的投影点为 $p(x, y)$,它们的透

视投影几何关系为: $x = f \frac{X}{Z}, y = f \frac{Y}{Z}$,应用齐次坐标表示为:

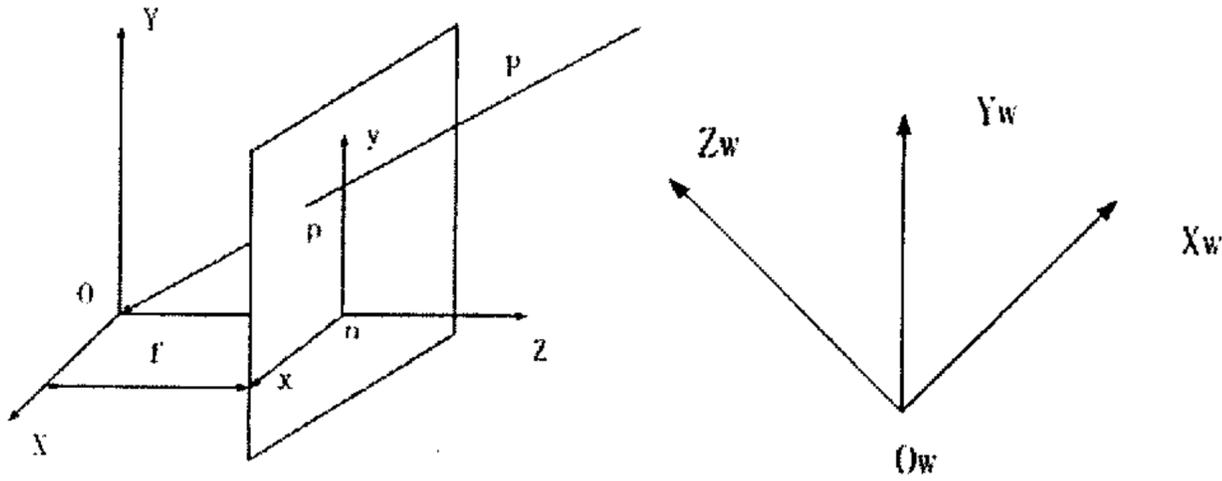


图 3.1 小孔摄像机模型

$$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

设图像的像素坐标是 (u, v) ，则图像点的像素坐标与以毫米为单位的图像

坐标的关系为：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & u_0 \\ 0 & s_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

其中， s_x, s_y 为图像平面单位距离上的像素数 (pixels/mm)，称为比例系数， (u_0, v_0) 为摄像机光轴与图像平面的交点，即计算机帧存图像中心的坐标。

摄像机坐标系与世界坐标系的转换关系为：

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

其中， R 和 T 分别为从世界坐标系到摄像机坐标系的旋转和平移变换， R 是一个 3×3 的正交矩阵， T 是 3×1 的平移向量。

于是，我们得到以世界坐标系表示的 P 点坐标 (X_w, Y_w, Z_w) 与其投影点 p 的坐标 (u, v) 的关系如下：

$$Z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} a_x & 0 & u_0 & 0 \\ 0 & a_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M_1 M_2 X = MX \quad (3.1)$$

其中， $a_x = f \cdot s_x$ ， $a_y = f \cdot s_y$ ； M 为 3×4 矩阵，称为投影矩阵； M_1 完全由内部

参数决定； M_2 完全由摄像机相对于世界坐标系的方位即外部参数决定。

可见，如果已知摄像机的内外参数，就已知投影矩阵 M ，对任何空间点 P ，如果已知其空间三维坐标 (X_{wj}, Y_{wj}, Z_w) ，就可以求出其图像坐标点的位置 (u, v) 。但是，反过来，如果已知空间某点的图像点的位置 (u, v) ，即使已知投影矩阵，其空间坐标也不是唯一确定的，它对应的是空间的一条直线。

3.2.2 图像的几何畸变

由于摄像机光学系统并不是精确地按理想化的小孔成像原理工作，存在有透镜畸变，物体点在摄像机成像面上实际所成的像与理想成像之间存在有光学畸变误差。主要的畸变误差分为三类：径向畸变，偏心畸变和薄棱镜畸变。第一类只产生径向位置的偏差，后两类则既产生径向偏差，又产生切向偏差^[10]。图 3.2 所示为无畸变理想图像点位置与有畸变实际图像点位置之间的关系。

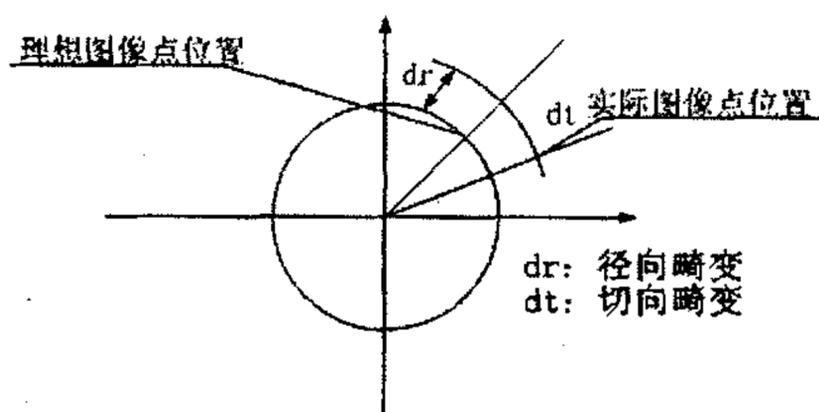


图 3.2 理想图像点和实际图像点

1. 径向畸变 (Radial Distortion)

径向畸变主要是由镜头形状引起的，是关于摄像机镜头的主光轴对称的。正向畸变称为枕形畸变，负向畸变称为桶形畸变，如图 3.3 所示。其数学模型为：

$$\begin{aligned} \delta_x &= X_d \cdot (k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots) \\ \delta_y &= Y_d \cdot (k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots) \end{aligned} \quad (3.2)$$

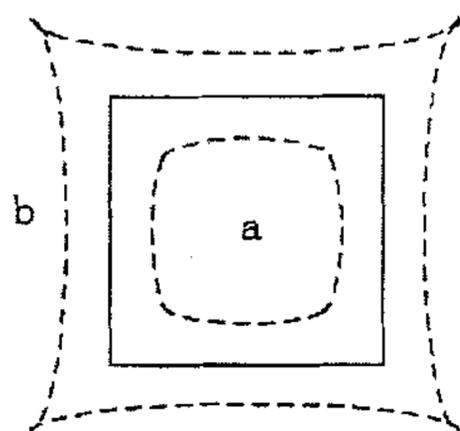
其中， k_1, k_2, k_3, \dots 为径向畸变系数。

2. 偏心畸变 (Decentering Distortion)

偏心畸变主要是由光学系统光心与几何中心不一致造成的，即镜头器件的光学中心不能严格共线。这类畸变既含有径向畸变，又含有切向畸变(图 3.4)。切向畸变的数学模型为：

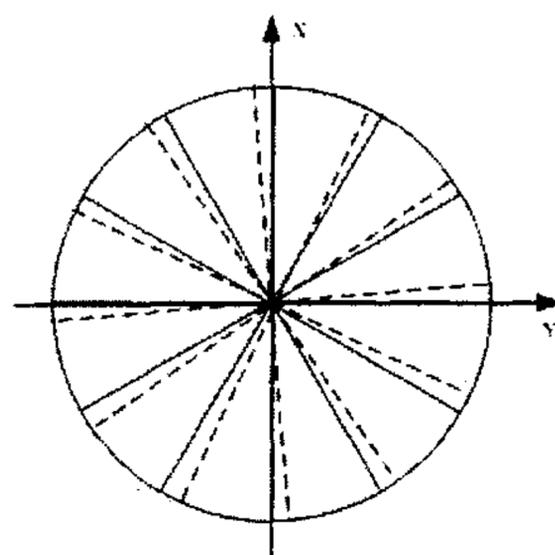
$$\begin{aligned} \delta_{xd} &= p_1 \cdot (3X_d^2 + Y_d^2) + 2p_2 \cdot X_d \cdot Y_d + \dots \\ \delta_{yd} &= p_2 \cdot (X_d^2 + 3Y_d^2) + 2p_1 \cdot X_d \cdot Y_d + \dots \end{aligned} \quad (3.3)$$

其中, p_1, p_2, \dots 为切向畸变系数。



a 桶形畸变 b 枕形畸变

图 3.3 切向畸变



实线: 无畸变 虚线: 有畸变

图 3.4 偏心畸变

3. 薄棱镜畸变

薄棱镜畸变是由于镜头设计缺陷和加工安装误差所造成的,如镜头与摄像机像面有很小的倾角等。这类畸变相当于在光学系统中附加了一个薄棱镜,不仅会引起径向偏差,而且引起切向误差。其数学模型为:

$$\begin{aligned} \delta_{xp} &= s_1 r^2 + \dots \\ \delta_{yp} &= s_2 r^2 + \dots \end{aligned} \quad (3.4)$$

其中, s_1, s_2 为薄棱镜畸变系数。

图像平面理想图像点坐标 (X_u, Y_u) 为实际图像点坐标 (X_d, Y_d) 与畸变误差之和, 即:

$$\begin{aligned} X_u &= X_d + \delta_{xr} + \delta_{xd} + \delta_{xp} \\ Y_u &= Y_d + \delta_{yr} + \delta_{yd} + \delta_{yp} \end{aligned} \quad (3.5)$$

薄棱镜畸变很小, 可以忽略。

一般情况下, 在图像边缘处存在较大的畸变误差, 采用普通 CCD 摄像机捕获大小为 500×500 像素的图像时, 其边缘的畸变大约为 5 个像素。径向畸变关于摄像机镜头的主光轴对称, 是对边缘处存在较大畸变的一种模型化。在工业视觉中, 一般只需要考虑径向畸变, 因为引入过多的非线性参数, 往往不仅不能提高解的精度, 反而会引起解的不稳定性。

图 3.6 显示了利用机械手视觉系统拍摄的一幅标定模板(图 3.5)图像, 可以看到原来的直线畸变成了弧状, 明显发生了桶形失真, 这是由摄像机系统的径向畸变引起的。如果不进行畸变矫正, 将影响到目标物体的正确识别。

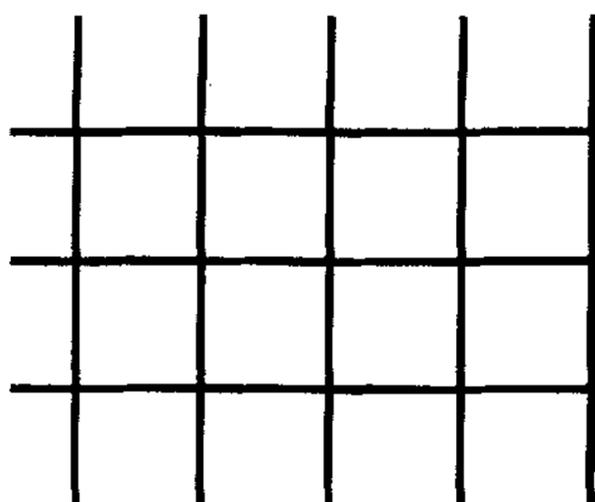


图 3.5 标定模板



图 3.6 径向畸变引起的桶形失真

3.3 基于人工神经网络的畸变校正

图像几何畸变校正的思想,就是通过一些已知的正确象素点和畸变点间的对应关系,从畸变图像恢复正常的图像。图像畸变校正经典的方法是通过线性投影变换求得摄像机的参数矩阵,然后通过线性变换进行像差纠正。由于摄像机畸变是一个复杂的非线性过程,人工神经网络在以下几个方面与摄像机标定具有相似的特点:1)神经网络由许多具有非线性特性的神经元组成,因此它有学习非线性关系的能力;2)神经网络和摄像机标定的工作机制是相近的,都是从一些已知数据去获得系统参数,并由此来计算未知的数据;3)使用神经网络来确定镜头畸变模型,该镜头畸变可以是任意模型。

人工神经网络是近年来发展起来的一门新兴技术,它可以处理那些难以用数学模型描述的系统,具有很强的自适应、自学习能力。已经证明如果多层前馈神经网络隐含层节点可自由设定,则它可以实现任意精度近似任何连续函数[11]。

误差反传播(Back Propagation,简称BP)神经网络是当今科学技术领域中广泛应用的一种人工神经网络,属多层前馈神经网络,其突出的优点是具有很强的非线性映射能力以及根据具体问题灵活处理的柔性网络结构,在语言综合、信息处理和智能控制等领域发挥着重要作用。

下面首先简单介绍一下BP神经网络模型及用于畸变图像校正的方法。

3.3.1 BP神经网络模型

1986年,以Rumelhart和McClelland为首的科学家小组在《Parallel Distributed Processing》一书里,对具有非线性连续转移函数的多维多层前馈网络的误差反传播算法进行了详尽地探讨,从而解决了感知器不能解决的多层网络学习算法的问题,人们也常把按误差反传播算法训练的多层前馈网络,直接称为BP网络。

BP算法的基本思想是,根据样本的期望输出与实际输出之间的平方误差,利用梯度下降法,从输出层开始,逐层修正权系数。每个修正周期分两个阶段:前向传播阶段和反向传播阶段。

图 3.7 给出了一个典型的 BP 网络结构，分为输入层、隐含层和输出层。假设输入层节点数为 n ，隐含层节点数为 h ，输出层节点数为 c ，输入样本为 x_1, x_2, \dots, x_n ，样本的期望输出为 d_1, d_2, \dots, d_c ，则有

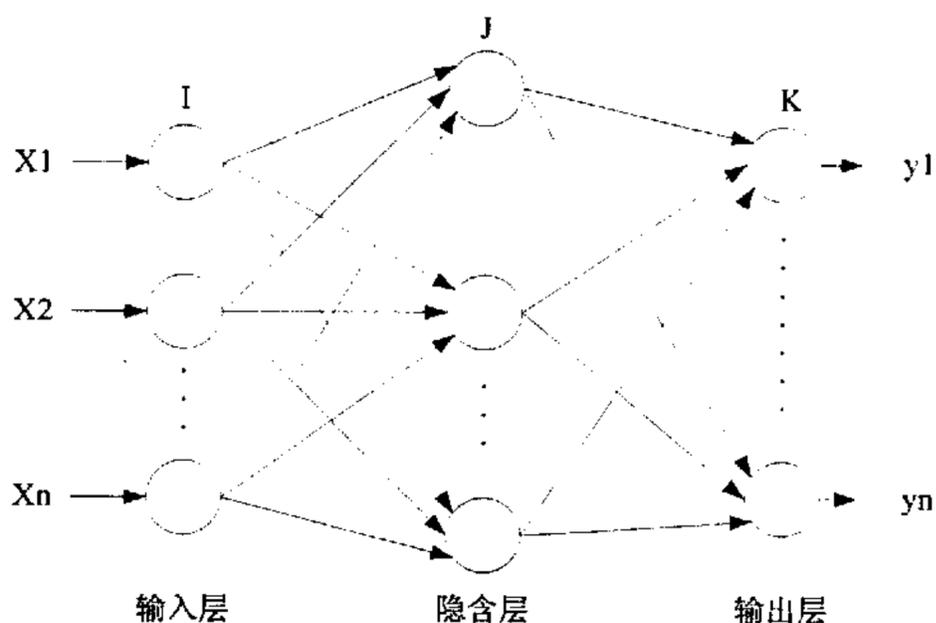


图 3.7 BP 网络结构

隐含层节点 j 的输入为

$$net_j = \sum_{i=1}^{n+1} w_{ji} o_i, \text{ 其中 } x_{n+1} = 1, o_i = x_i, w_{j,n+1} = -\theta_j \quad (3.6)$$

隐含层节点的输出为

$$o_j = f(net_j), j = 1, 2, \dots, h \quad (3.7)$$

输出层节点的输入为

$$net_k = \sum_{j=1}^{h+1} w_{kj} o_j, o_{h+1} = 1, w_{j,k+1} = -\theta_j \quad (3.8)$$

输出层节点的输出为

$$y_k = f(net_k), k = 1, 2, \dots, c \quad (3.9)$$

如果任意设置网络初始权系数，那么对于每个输入样本，网络输出与期望输出一般总有误差，定义平方误差为

$$E = \frac{1}{2} \sum_{k=1}^c (d_k - y_k)^2 \quad (3.10)$$

如果采用单极型 S 形函数作为激励函数，则输出层节点的权系数修正公式为：

$$\begin{aligned} w_{kj}(t+1) &= w_{kj}(t) + \eta \delta_k y_j \\ \delta_k &= (d_k - y_k) y_k (1 - y_k) \end{aligned} \quad (3.11)$$

式中， $k = 1, 2, \dots, c; j = 1, 2, \dots, h+1, \eta$ 为学习率， t 表示第 t 次迭代。

隐含层节点的权系数修正公式为：

$$w_{ji}(t+1) = w_{ji}(t) + \eta \delta_j y_i$$

$$\delta_j = y_j(1 - y_j) \sum_{k=1}^c \delta_k w_{kj} \quad (3.12)$$

式中， $j=1,2,\dots,h; i=1,2,\dots,n+1$ 。

在实际应用时，考虑到学习过程的收敛性，通常在权系数修正公式中右侧再加上一个“惯性”项 $\alpha \Delta w(t)$ ， α 为动量系数， $\Delta w(t)$ 为第 t 次学习权系数的变化。一般地，BP 学习算法描述为如下步骤：

初始化网络及学习参数，如设置网络初始权系数，学习率 η ，动量系数 α 等：

- (1) 给定输入样本和期望输出；
- (2) 计算实际输出；
- (3) 修正权系数：从输出层开始，将误差信号沿连接通路反向传输方向传播，通过修正各权系数，使误差最小；
- (4) 达到误差精度或循环次数要求，则输出结果，否则回到 (2)。

3.3.2 图像几何畸变校正的神经网络方法

我们知道，摄像机线性标定方法具有计算简单、速度快的优点，但由于没有考虑摄像机的镜头畸变，标定精度受到影响。如果能够求得每一像素点处的镜头畸变校正，问题就可迎刃而解，因此我们提出了镜头畸变校正的多层前馈神经网络方法，步骤为：

- (1) 首先利用线性标定方法确定摄像机的针孔模型，求得摄像机参数；
- (2) 利用得到的针孔模型，分别计算用于标定的控制点的理想图像点；
- (3) 求取用于标定的控制点的理想图像点与对应的实际图像点的偏差；
- (4) 摄像机需要一个校正网络，将其实际图像坐标作为神经网络的输入，由第三步得到的各自偏差值作为神经网络的输出，通过网络的训练建立起摄像机镜头畸变的校正模型。

3.3.3 实验方法

3.3.3.1 样本的选取

我们选择模板图像与标准模板的数据作为摄像机标定的样本数据集，在图像上选取 15 个点作为特征点，其图像坐标值 (U, V) 与所对应的世界坐标值 (X, Y) 见表 3.1。

3.3.3.2 神经网络训练策略

BP 网络具有很强的非线性映射能力以及根据具体问题灵活处理的柔性网络

结构^[12]，但是，在实际应用中，目前尚未找到较好的网络构造方法，确定神经网络的结构和权系数，来描述给定的映射或逼近一个未知的映射，只能通过学习方式来得到满足要求的网络模型。神经网络的学习可以理解为：对确定的网络结构，寻找一组满足要求的权系数，使给定的误差函数最小。下面给出具体的实验步骤：

表 3.1 样本数据集 单位: pixel

	U	V	X	Y
1	70	129	67	117
2	218	125	220	117
3	368	120	373	117
4	523	120	526	117
5	678	117	679	117
6	71	218	67	269
7	219	277	220	269
8	373	276	373	269
9	527	273	526	269
10	681	269	679	269
11	75	434	67	421
12	223	429	220	421
13	374	428	373	421
14	529	427	526	421
15	682	424	679	421

(1) 确定 BP 网络输入层和输出层节点的个数，本文采用的输入层为 2 个节点，输出层为 2 个节点；

(2) 获取样本集并进行处理；

(3) 确定 BP 网络的拓扑结构。确定网络有多少个隐含层，中间的隐含层有多少个神经元数。对隐含层数和每个隐含层的神经元数的选择，虽然有一定的指导原则，但更多的是靠经验和试凑。我们的做法是，初始选用比较大的隐单元数和比较多的隐含层数，逐步减小，直到网络通过学习仍然能达到所要求的精度为止。经过试验，比较合适的隐含层数为 1 个，最佳节点数为 5。

(4) 确定学习率 η 和动量因子 α ，也需要试验。如果 η 太小，收敛的速度可能很慢；如果太大，可能出现麻痹现象或连续不稳定。本文最终选取的 η 为 0.5， α 为 0.7。

(5) 初始化权系数。初始化值一般取随机数，而且比较小。初始值对于学习过程能否达到局部最小和能否收敛有很大的关系。

(6) 将训练数据输入网络，通过对网络的连接系数做调整的迭代运算，对网络的权系数反复调节，直到误差满足要求为止。

(7) 训练完毕后，就需要对这个网络模型进行测试。与训练样本类似，任意收集若干数据对作为测试样本，计算网络的输出值与期望输出值之间的误差，依次检验网络的有效性。

(8) 如果测试结果满意，即可认为该神经网络可以用来实现这个关系映射。否则，可能需要修改前面(3)、(4)和(5)步骤中的参数，直到满意为止。

利用 BP 神经网络，训练函数分别采用 trainbp, trainbpx, trainlm 函数，对网络进行训练。经过大量的实验和测试，取误差指标为 0.01 时，得出其性能如表 3.2，图 3.8，3.9 为神经网络校正前后效果比较。

表 3.2 性能对比结果

训练函数	网络类型	时间(秒)	训练步数
trainbp	BP 网络	259.1	4123
trainbpx	采用快速 BP 算法的前向网络	42.4	570
trainlm	采用 L—M 算法的前向网络	3.3	5

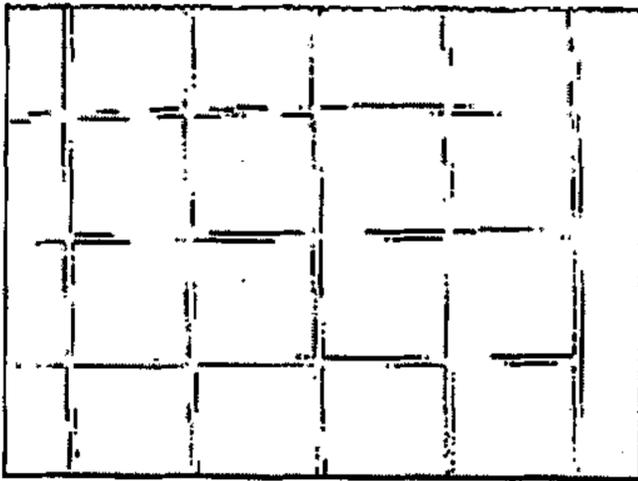


图 3.8 神经网络校正前

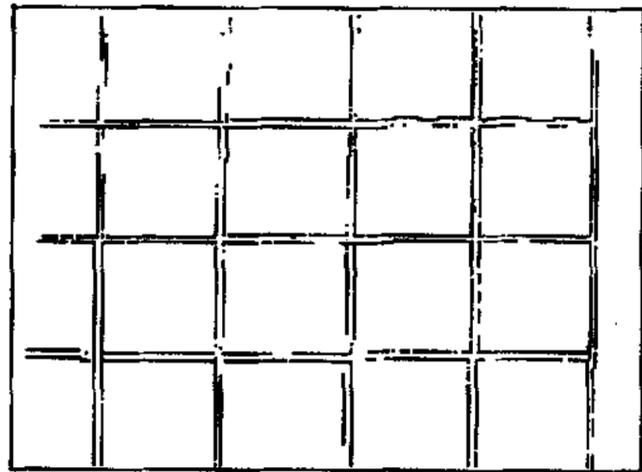


图 3.9 神经网络校正后

3.3.4 校正算法精度和评价

3.3.4.1 算法精度分析

利用上述方法得到如图结果。将每一个点的均方根 (Root Mean Square,

简称 RMS) 误差和整个样本测试点集的均方根误差作为评价指标。计算公式

$$\begin{aligned}
 P_e &= \sqrt{\frac{(Xa - Xn)^2 + (Ya - Yn)^2}{2}} \\
 R_e &= \sqrt{\frac{P_{e1}^2 + P_{e2}^2 + \dots + P_{ei}^2}{I}}
 \end{aligned}
 \tag{3.13}$$

式如下：

式中， P_e ——任一点的均方根误差；

R_c ——整个测试点集的均方根误差;

X_a, Y_a ——测试点的真实坐标值

X_n, Y_n ——计算出的坐标值

P_{cl} ——第 I 点的均方根误差

I——测试点总数

表 3.3 给出了计算结果, 表中最后一行为整个测试点集的均方根误差。

表 3.3 计算结果 单位: pixel

测试点序号	计算结果
1	1.392956
2	2.815582
3	2.831605
4	2.479630
5	3.316151
6	2.886430
7	1.963392
8	2.578660
9	3.307720
10	0.890320
11	2.772072
12	1.950103
13	2.150103
14	2.557703
15	2.984065
Re	2.4371

为了更直观看出误差, 我们还给出了测试点在 X, Y 方向上的误差结果, 如图 3.10, 图 3.11 所示, 其计算公式为

$$\begin{aligned} E_x &= X_n - X_a \\ E_y &= Y_n - Y_a \end{aligned} \tag{3.14}$$

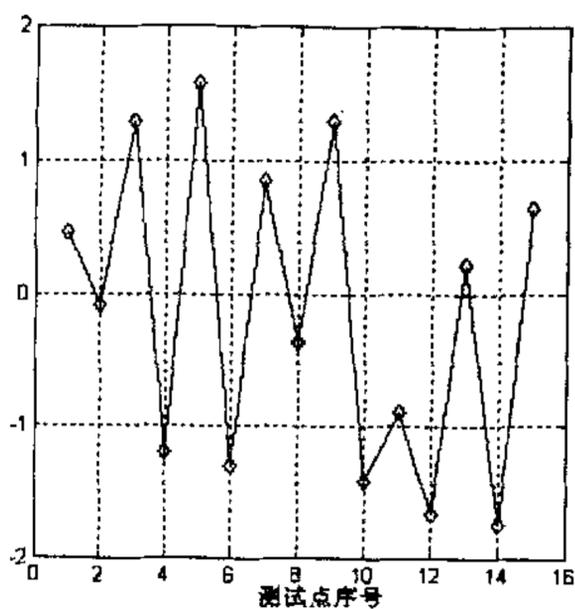


图 3.10 X 方向误差

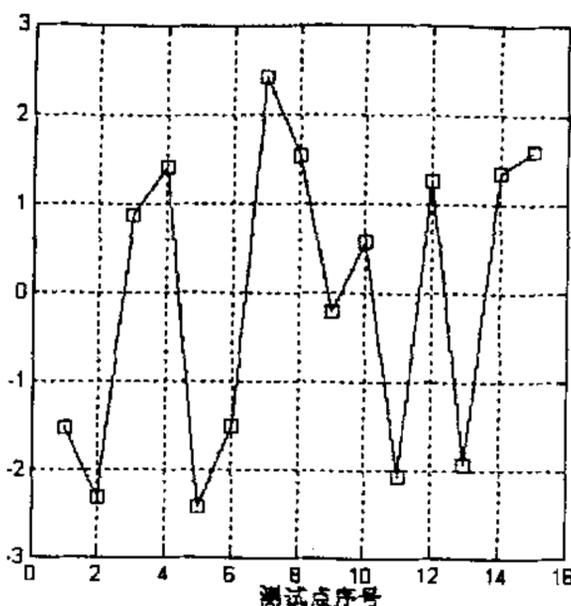


图 3.11 Y 方向误差

3.3.4.2 影响校正精度的主要误差分析

影响图像畸变校正精度的因素由许多，正确分析这些因素，是评价校正精度的前提，也是改善算法，提高精度的基础。包括光学中心估计误差、原点中心定位误差、截取误差、插值误差和灰度校正误差，其中误差的产生主要在以下几个方面。

摄影头的分辨率不高，这样，每一像素点表示的距离值较大，这样产生的误差也就较大。我们可以采用分辨率更高的摄像头来获得更高的精度。

由于摄像头光轴与目标物体平面很难做到完全垂直，从而造成拍摄图像的不对称变形，由此可能造成较大误差。

另外，精确测量控制点的世界坐标和图像坐标是一项严格的工作，需要精密的装置和测量方法。因此，训练样本集的获得不太精确也是产生误差的一个原因。

3.4 平面标定技术

一般二维图像测量系统主要由被测物、光学成像系统、CCD 摄像机、信号处理电路及计算机组成，CCD 摄像机把带有边缘位置信息或尺寸信息的光学信号变为视频信号，经图像处理电路送入计算机处理，因此图像处理的数据是数字信号，得出的边缘位置是以像素为单位来表示的，如果要给出物体的实际尺寸，必须建立数字图像像素与实际尺寸的对应关系。不妨把一个图像像素在水平方向和垂直方向相当于实际尺寸的大小分别简称为像素的水平当量和垂直当量。确定像素的水平当量和垂直当量的过程，就是摄像机标定的过程。

传统的摄像机标定技术需要一个标定参照物，即需要在摄像机前放置一个已知物体。在标定过程中，利用已知物体上一些点的三维坐标和它们相应的图像点坐标，计算摄像机的内外参数。

从计算方法的角度上看，传统的摄像机标定方法可分为四类，即利用最优化

算法的标定方法,利用摄像机透视变换矩阵的标定方法,进一步考虑畸变补偿的两步法和双平面标定方法。文献[13]中按照求解算法的特点分为直接非线性最小化方法、闭式求解法和两步法,这也是一种好的划分方法。

传统的摄像机定标方法是通过结构已知的物体(如定标块)在图像中的投影来计算摄像机的内参数。这种方法的缺点是需要一个定标块,这在很多实际应用中难以实现。基于主动视觉的自定标技术可以克服这个缺点,其特点是通过控制摄像机的运动获取多幅图像来标定内参数。Maybank 和 Faugeras[14]首次提出基于 Kruppa 方程的摄像机自定标方法,继之很多研究人员提出了一些其它的类似方法[15-23]。在这些方法中都需要解一个非线性方程组,一般计算量都很大。针对这一问题,马颂德[24]提出了一种基于主动视觉系统的摄像机自定标的线性方法。这种方法通过控制摄像机在三维空间作两组平移运动,其中每组运动包括三次两两正交的平移运动,简称三正交平移运动,利用 FOE(Focus of Expansion)点建立一个线性方程组来求解摄像机内参数。杨长江等人对此提出了一种改进方法,要求摄像机作四组平移运动[25],其中每组包括两次相互正交的平移运动,利用 FOE 点仍可线性求解摄像机内参数。文[24, 25]均是在四参数即畸变因子为零的摄像机模型下,实现线性自定标,在五参数的模型下,利用他们的方法,不可能线性求解。由于线性方法计算简单,在数值上比非线性方法稳定,一直是摄像机定标所追求的目标。

最近, Zhang 提出利用平面模板来标定摄像机[19]。在 Zhang 的方法中,要求摄像机在两个以上不同的方位拍摄一个平面模板的图像,模板可以用普通的激光打印机打印,贴在一个平面上(如硬皮书面),摄像机和模板都可以自由地移动,不需要知道运动参数。这种方法本质上是一种传统的摄像机标定方法,只是用一个平面模板代替了传统摄像机标定中的标定块,其特点是方法简单方便,成本低,标定稳定性和精度相对于自标定一般来说要高。本文就是基于这样的思想设计了一种简单易行的方法进行标定——网格式平行线标定方法。

3.4.1 网格式平行线标定方法的基本原理及实现的关键技术

网格式平行线标定方法的标准件设计成如下形式:图 3.12 为激光打印的平面模板。设网格交叉点分别为: O_1, O_2, \dots, O_{16} ;这些点的坐标可统记为 (X, Y) , 坐标平面记为 $X-Y$, CCD 感光面的坐标平面为 X_s-Y_s , 点的坐标为 (X_s, Y_s) 。由物面到 CCD 感光面的成像关系如图 3.13 所示。下面的分析不考虑光学系统像差、CCD 像机及图像卡的影响。

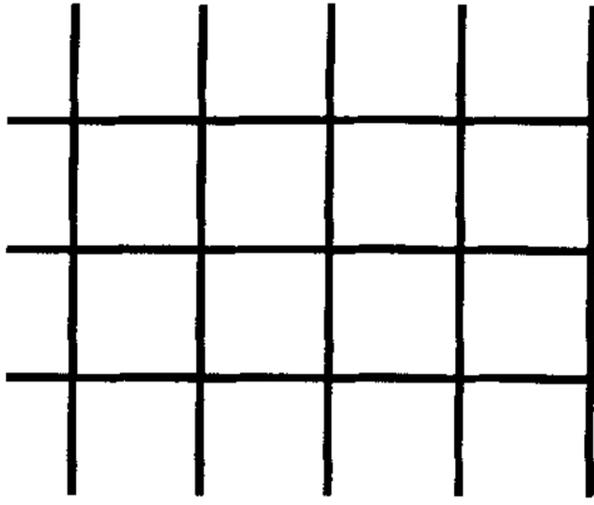


图 3.12 模板原图

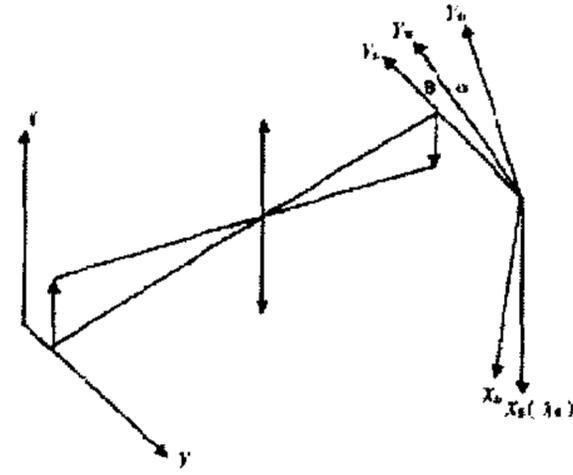


图 3.13 物体成像图

令光学系统的放大倍数为 M ，根据成像关系，可得出

$$\begin{aligned} X_s &= MX \\ Y_s &= MY \end{aligned} \quad (3.15)$$

由于 CCD 摄像机的转移频率与图像卡的采样频率并不相同，所以， X_s-Y_s 坐标系与图像坐标系 X_D-Y_D 之间存在一个几何变换关系[26]；另外，CCD 摄像机阵列水平方向与垂直方向不一定垂直，存在一个轴偏变换。设变换后的坐标系为 X_w-Y_w ，网格交叉点图像坐标为 (X_D, Y_D) ，水平像素当量及垂直像素当量分别为 CX, CY ，可得出如下两个关系式：

$$\begin{aligned} X_w &= X_s - Y_s \sin \theta \\ Y_w &= Y_s \cos \theta \end{aligned} \quad (3.16)$$

$$\begin{aligned} CX &= MXD = X_w \cos \alpha - Y_w \sin \alpha \\ CY &= MYD = X_w \sin \alpha + Y_w \cos \alpha \end{aligned} \quad (3.17)$$

式(3.16)、(3.17)中 α 为旋转角， θ 为轴偏角。把式(3.16)、(3.17)代入式(3.15)得

$$\begin{aligned} X_D &= a_0 + a_1 X + a_2 Y \\ Y_D &= b_0 + b_1 X + b_2 Y \end{aligned} \quad (3.18)$$

式中：

$$\begin{aligned} a_1 &= \cos \alpha CX, a_2 = -\sin(\alpha + \theta)CX \\ b_1 &= \sin \alpha CY, b_2 = -\cos(\alpha + \theta)CY \end{aligned} \quad (3.19)$$

a_0, b_0 为考虑数字图象坐标系相对于物面坐标系平移的量[27]。以标准件交叉点的坐标 (X, Y) 为自变量，交叉点的对应的图像坐标 (X_D, Y_D) 为因变量进行最小二乘拟合，求出 a_1, a_2, b_1, b_2 其公式为 $E = [a_0 \ a_1 \ a_2 \ b_0 \ b_1 \ b_2]^T$

式中：

$$\begin{aligned}
 a_1 &= \frac{\cos \alpha}{C_x} & a_2 &= -\frac{\sin(\alpha + \theta)}{C_x} \\
 b_1 &= \frac{\sin \alpha}{C_y} & b_2 &= \frac{\cos(\alpha + \theta)}{C_y}
 \end{aligned} \tag{3.20}$$

a_0, b_0 为考虑数字图像坐标系相对于物面坐标系平移的量。以标准件交叉点的坐标 (X, Y) 为自变量, 交叉点的对应的图像坐标 (XD, YD) 为因变量进行最小二乘拟合, 求出 a_1, a_2, b_1, b_2 其公式为 $E = [a_0 \ a_1 \ a_2 \ b_0 \ b_1 \ b_2]^T$

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \dots \\ A_n \end{bmatrix} \quad A_i = \begin{bmatrix} 1 & X & Y & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & X & Y \end{bmatrix}$$

$$B = \begin{bmatrix} B_1 \\ B_2 \\ \dots \\ B_n \end{bmatrix} \quad B_i = [X_D \ Y_D]^T$$

$$E = (A^T A)^{-1} A^T B \tag{3.21}$$

式中, n 为参与最小二乘的交叉点数。利用下面的公式即可计算出 C_x, C_y :

$$\begin{aligned}
 C_x &= \frac{\cos \alpha}{a_1} \\
 C_y &= \frac{\sin \alpha}{b_1}
 \end{aligned} \tag{3.22}$$

式中, $\sin \alpha = \pm \sqrt{\frac{a_1^2 b_2^2 - a_2^2 b_1^2}{a_1^2 b_2^2 + a_2^2 b_1^2}}$, $\sin \alpha$ 的符号选择与 b_1 相同的符号。

标准模板交叉点的坐标 (X, Y) , 通过精确测量网格水平和垂直两个方向的间距来获得。数字图像交叉点的坐标获取方法是: 先通过图像处理的边缘检测技术获得各条水平线和垂直线的双边线上各点坐标, 然后应用最小二乘法来进行直线拟合, 求出交叉各点的坐标。由于标准件线条的边缘不可能很完美, 这就造成边缘检测得到的位置坐标值有一定分散性, 给标定带来误差。为减少这种影响, 本文采用平行线拟合的方法, 也就是说, 把彼此互相平行的水平直线或垂直线一起参与最小二乘拟合, 得到各条直线的方程, 然后再求交叉点的坐标, 这样由于参与拟合的点数大大增加, 即使某处边缘有个别缺陷, 也不影响求出的交叉点的坐标。设参与拟合的直线数为 m , 每条直线参与拟合的点数为 n_1, n_2, \dots, n_m , 每条直线的点集记为 $P_j, j=1, 2, \dots, m$, 各条直线的方程为

$$\begin{aligned} \sin \beta x - \cos \beta y + d_1 &= 0 \\ \sin \beta x - \cos \beta y + d_2 &= 0 \\ \dots\dots\dots \\ \sin \beta x - \cos \beta y + d_m &= 0 \end{aligned} \tag{3.23}$$

利用最小二乘法，构造如下目标函数：

$$\begin{aligned} W(\beta, d_1, d_2, \dots, d_m) &= \sum_{(x_i, y_i) \in P_1} [\sin \beta x_i - \cos \beta y_i + d_1]^2 + \sum_{(x_i, y_i) \in P_2} [\sin \beta x_i - \cos \beta y_i + d_2]^2 \\ &+ \dots + \sum_{(x_i, y_i) \in P_m} [\sin \beta x_i - \cos \beta y_i + d_m]^2 \end{aligned} \tag{3.24}$$

式中， (x_i, y_i) 为拟合直线上的点的坐标。

可以得出

$$\begin{aligned} \bar{x}_j &= \frac{1}{n_j} \sum_{(x_i, y_i) \in P_j} x_i & j=1, 2, \dots, m; i=1, 2, \dots, n_j \\ \bar{y}_j &= \frac{1}{n_j} \sum_{(x_i, y_i) \in P_j} y_i & j=1, 2, \dots, m; i=1, 2, \dots, n_j \\ \gamma_j &= \frac{1}{n_j} \sum_{(x_i, y_i) \in P_j} x_i^2 - \bar{x}_j^2 - \frac{1}{n_j} \sum_{(x_i, y_i) \in P_j} y_i^2 - \bar{y}_j^2 & j=1, 2, \dots, m; i=1, 2, \dots, n_j \\ \bar{\xi}_j &= \frac{1}{n_j} \sum_{(x_i, y_i) \in P_j} x_i y_i - \bar{x}_j \bar{y}_j & j=1, 2, \dots, m; i=1, 2, \dots, n_j \end{aligned} \tag{3.25}$$

令

$$\gamma = \sum_{j=1}^m \gamma_j, \quad \xi = \sum_{j=1}^m \xi_j \quad j=1, 2, \dots, m$$

则有

$$\gamma \sin 2\beta - 2\xi \cos 2\beta = 0 \tag{3.26}$$

当 $\xi \neq 0$ 时

$$\tan \beta = \frac{-\gamma + \sqrt{\gamma^2 + 4\xi^2}}{2\xi} \tag{3.27}$$

再利用：

$$d_j = \cos \beta \bar{y}_j - \sin \beta \bar{x}_j \quad j=1, 2, \dots, m \tag{3.28}$$

求出 d_j 。

为了提高精度，可将目标放置在工作平面的各个区域，求得标定结果后取平

均值。由于本文所述方法简单，多做几组数据很容易实现。通过实验，本文研究的气动机械手视觉系统的标定结果如表 3.4 所示：

表 3.4 像素当量值 单位： μ/mm

Cx	Cy
3.79	3.76

3.5 本章小结

由于摄像机观察目标是相对于自身坐标系的，得到的视觉反馈也是相对与摄像机坐标系描述的，而机械手控制及轨迹规划是在机械手坐标系中进行的。因此要把相对于摄像机坐标系的观察描述转换到机械手坐标系中。这就需要标定摄像机坐标系与机械手坐标系之间的关系。摄像机参数的标定是实现机械手视觉伺服控制的基础。本章主要研究摄像机标定和畸变校正，分析了摄像机模型和摄像机畸变原理，利用神经网络进行畸变校正，采用网格式平行线方法标定了像素的水平当量和垂直当量。

第 4 章 图像预处理及物体特征的识别

4.1 引言

一个图像识别系统可分为三个主要部分，其框图如图 4.1 所示。第一部分是图像信息获取，它相当于对被研究对象的调查和了解，从中得到数据和材料，对图像识别来说就是把图片、底片、文字图形等用光电扫描设备变换为电信号以备后续处理。第二部分是信息的加工与处理。它的作用在于把调查了解到的数据材料进行加工、整理、分析、归纳以去伪存真、去粗取精，抽出能反映事物本质的特征。当然，抽取什么特征，保留多少特征与采用何种判决有很大关系。第三部分是判决或分类。这相当于人们从感性认识升到理性认识而做出结论的过程，与特征抽取的方式密切相关。它的复杂程度也依赖与特征的抽取方式。例如，类似度、相关度、最小距离等。

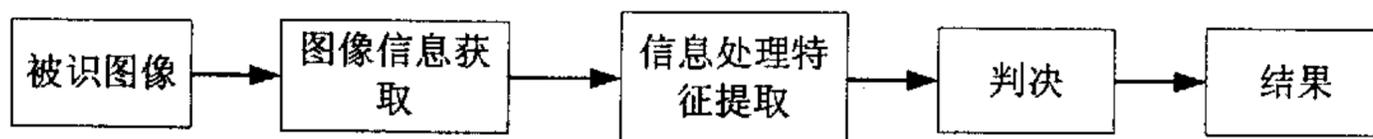


图 4.1 图像识别系统

模式识别已初步形成三大类：统计模式识别、结构模式识别、模糊集识别等三大类^[28]，本文主要用到统计模式识别法。

统计模式识别认为图像可能包含一个或多个物体，并且每个物体属于若干事先定义的类型、范畴或模式之一。在给定一幅含有多个物体的数字图像的条件下，模式识别过程由三个主要阶段组成（见图 4.2）。

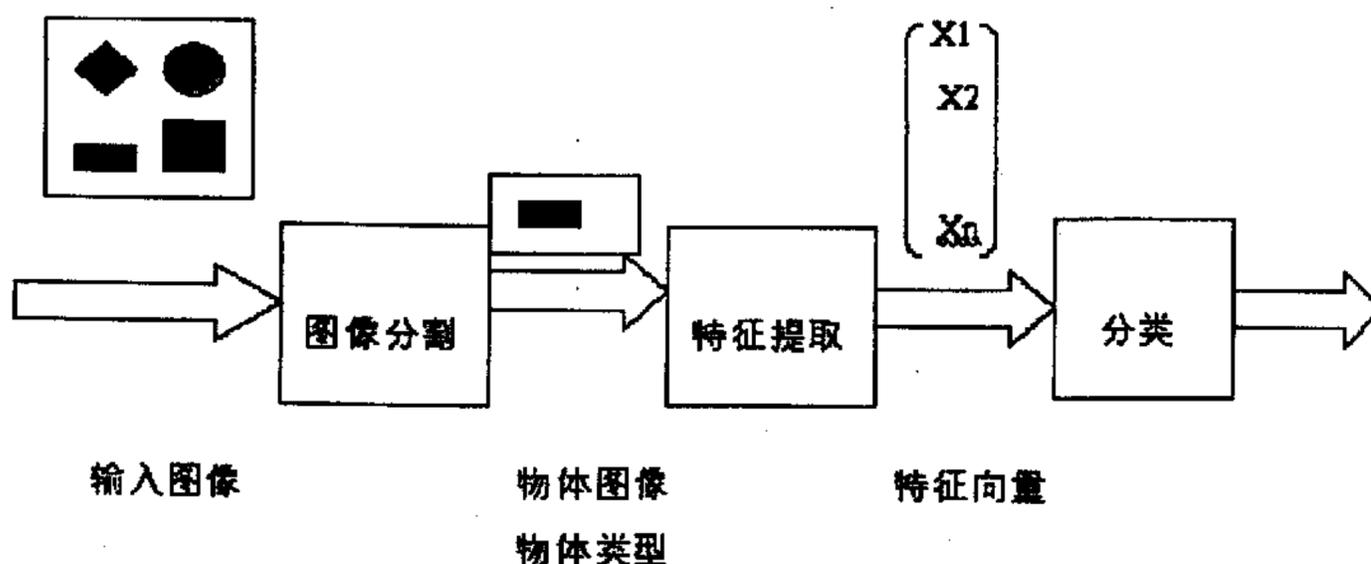


图 4.2 图像分割原理图

第一个阶段称为图像分割或物体分离阶段。在该阶段中检测出各个物体，并把它们的图像和其余景物分离。

第二个阶段称为特征提取阶段。在该阶段中对物体进行度量。一个度量是指一个物体某个可度量的度量值，而特征是一个或多个度量的函数。通过计算可以对物体的一些重要特性进行定量化表示。特征抽取过程产生了一组特征，把他们组合在一起，就形成了特征向量。这种被大大减少了的信息（与原始图像相比）代表了后续分类决策必须依靠的全部知识。引进一个 n 维空间的概念是很有益的，这个空间包含了所有可能的 n 维特征向量。因此任意一个特定物体都对应于特征空间中的一点。

模式识别的第三个阶段是分类。它的输出仅仅是一种决策，确定每个物体应该归属的类别。每个物体被识别为某一特定类型，它是通过一个分类过程加以实现的。每个物体被指定属于若干预先定义好的组（类）中的某一个组。这些组代表了预期存在于图像中物体的所有可能类别。如果把指派物体归到一个不正确的类，就产生一个分类错误。错分类发生的概率称为误判率。

4.2 图像预处理

对图像的预处理是为了去掉图像中的噪声和畸变等无关的信息，并且还能够恢复有用的信息，把图像具有的信息变得更容易观察，增强有关信息的可测性和最大限度的简化数据，使边缘检测、特征提取和识别分析更容易进行，更有可靠性。图像预处理包括对比度的增强、几何畸变、噪声的去除等操作[29]。

图像的对比度增强一般是利用线性灰度变换、非线性灰度变换或直方图平坦化等操作来进行对比度增强，但在灰度的绝对值具有意义的图像或地图、图纸一类的二值图像中，信息将被歪曲，有时会得到相反的结果。

在远离图像中心的图像不可避免存在几何畸变，本文第三章采用 BP 神经网络进行畸变校正，因此在图像的预处理中不再对图像进行几何畸变校正。

图像中包含着多种噪声，所以在从图像提取各种特征之前，有必要除去这样的噪声。

4.2.1 对比度增强

图像在成像、采集、转换和传输中，由于光照条件、成像器设备和外部环境噪声等各种因素的影响，会产生降质，对比度很低，图 4.3 显示了四种类型的物体图像：暗、亮、低对比度和高对比度。因此，在识别之前需要对图像进行对比度增强处理，扩大图像的灰度范围。图像增强技术是不考虑图像降质的原因，只将图像中感兴趣的特征有选择地突出，而衰减其不需要的特征，故改善后的图像

不一定要去逼近原图像。如突出目标物轮廓、增强对比度、去除各类噪声等，图像增强的主要目的是提高图像的可懂度。

对比度增强最简单的方法是灰度映射，对图像中每一个像素的灰度进行标度变换，改变图像的灰度范围^[30]。图像对比度增强的算法公式为

$$g(x,y) = T[f(x,y)] \quad (4.3)$$

式中， $f(x,y)$ 表示处理前的图像， $g(x,y)$ 表示处理后的图像。

对比度增强的方法有很多，常用的有分段线性灰度变换和直方图均衡化。这里采用的是分段线性灰度变换。

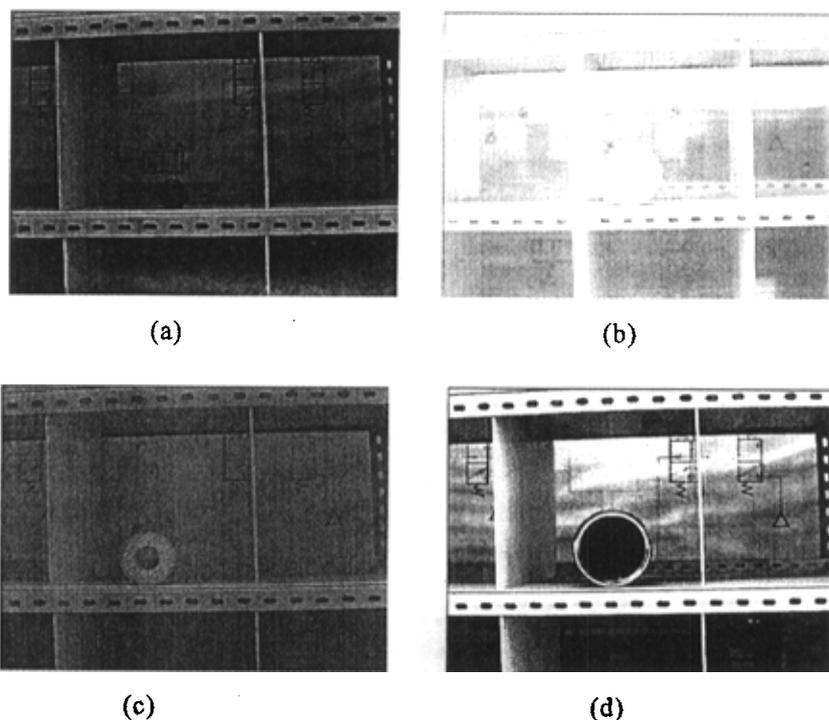


图 4.3 四种基本类型的图像

(a)暗，(b)明，(c)低对比度，(d)高对比度

分段线性变换是一种简单的灰度变换函数，既可以用于对比度增强，也可以用于灰度压缩，分段线性变换函数的图形表示为：

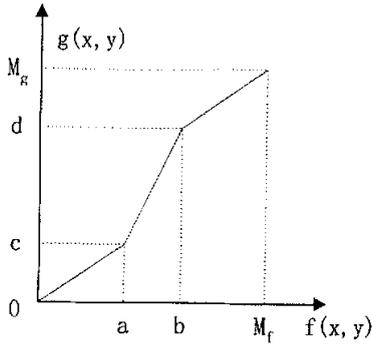
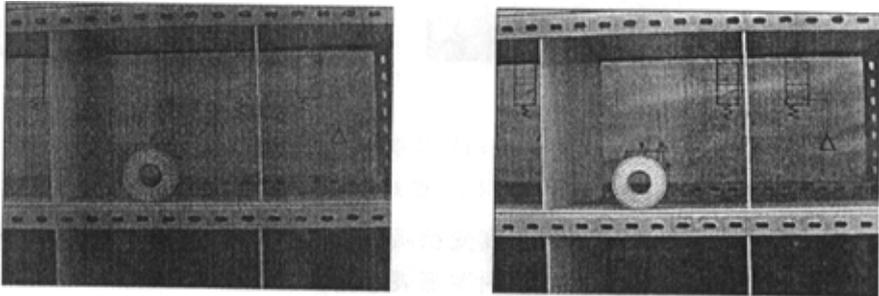


图 4.4 分段线性变换

数学表达式为：

$$g(x, y) = \begin{cases} \frac{c}{a} * f(x, y) & 0 \leq f(x, y) < a \\ \frac{d-c}{b-a} * [f(x, y) - a] + c & a \leq f(x, y) < b \\ \frac{M_g - d}{M_f - b} * [f(x, y) - b] + d & b \leq f(x, y) < M_f \end{cases} \quad (4.4)$$

在应用中，一般将原图像的灰度级范围（fmin, fmax）拉伸到（0, 255），使图像可以显示所有的灰度级，增加了对比度。此时 $c=0$, $d=255$, $a=fmin$, $b=fmax$ ，图 4.5 显示了利用分段线性灰度变换对引导线图像进行对比度增强的效果。一般来讲，对于灰度分布集中在一个位置的情况，这种方法有较好的效果，如果灰度分布有两处以上的集中位置，这种方法的效果就不很明显，需要用到另一种增强方法——直方图均衡化。



(a)

(b)

图 4.3(c) 图像

(b) 分段线性拉伸结果

图 4.5 分段线性拉伸结果

4.2.2 图像滤波

图像在生成和传输的过程中受到各种噪声的影响，使图像质量下降，为了抑制噪声改善图像质量，必须对图像进行平滑滤波处理^[31]。

影响图像质量的噪声一般来说主要有三种：电子噪声，光电子噪声和感光颗粒噪声。电子噪声是由于阻性器件的电子随机热运动而造成的，这类噪声很早就被电路设计人员成功地建模并研究，一般用零均值高斯白噪声作为其模型，这种模型具有一个高斯函数形状的直方图分布以及平坦的功率谱；光电子噪声是由光的统计本质和图像传感器中光电转换过程引起的，在弱光照的情况下，光电子噪声的影响就更为严重，通常用具有泊松密度分布的随机变量作为光电噪声的模型，在光照较强时，泊松型分布趋向高斯分布，而标准差仍等于均值的平方根，这意味着噪声的幅度是与信号有关的；感光片颗粒噪声是由感光片上感光乳胶中悬浮的卤化银颗粒因感光效果而引起的，在 CCD 摄像机中不存在这种噪声，但是由于 CCD 摄像机中 CCD 感光器件光电转换和一个光电感受单元而产生图像的噪声，对这种噪声目前还不能有效地进行数学上的模型化，但可采用根据噪声所具有的一般性质进行噪声消除。综上所述，图像噪声的产生机理是多种多样的，针对不同的噪声采用不同的处理办法来消除噪声可以得到不错的效果，但这样会加长图像处理的时间，影响图像处理的效率。目前，一般采用中值滤波加权平均法，但这些方法只能在一定程度上保持了图像的边缘细节，相当多的细节被抑制。这样，在处理过程中，既要尽可能的滤除噪声，又要更多的保留图像的细节就变得很困难^[30-33]。本文采用差值滤波法来对图像进行平滑图像滤波除噪声^[33]。

差值滤波法的原理是：定义了一个以某点为中心的 3×3 的图像窗口，将其分解成 45° 、 135° 和垂直水平三个方向上的 2×2 “矩形”，见图 (4.6)，通过计算“矩形”对角线顶点之间的梯度绝对值的和值并比较其大小，检测出该点周围的边缘细节分布状况，然后用边缘细节中和值最小的四个点的平均值来代替该点的灰度值，更好地平滑了图像的噪声和保护图像的细节，这样既不同于中值滤波法，忽视了周围的边缘细节，也不同于加权平均滤波法，衰弱了图像噪声的平滑。若 $f(x, y)$ 的数值代表原图像的灰度值， $g(x, y)$ 是滤波后的图像，则差值滤波用数学式子表示为三个方向的五个“矩形”的梯度绝对值的和值 t_0, t_1, t_2, t_3, t_4 ：

$$\begin{aligned} t_0 &= |f(x-1, y-1) - f(x, y)| + |f(x-1, y) - f(x, y-1)| \\ t_1 &= |f(x-1, y) - f(x, y+1)| + |f(x-1, y+1) - f(x, y)| \\ t_2 &= |f(x, y-1) - f(x+1, y)| + |f(x+1, y-1) - f(x, y)| \\ t_3 &= |f(x, y) - f(x+1, y+1)| + |f(x, y+1) - f(x+1, y)| \\ t_4 &= |f(x-1, y) - f(x+1, y)| + |f(x, y-1) - f(x, y+1)| \end{aligned}$$

判断 t_0, t_1, t_2, t_3, t_4 中最小值 $t_{\min} = \min(t_0, t_1, t_2, t_3, t_4)$ ，那么 $g(x, y)$ 的值为与 t_{\min} 相关的四个点灰度值的平均值。运用差值滤波法虽然解决了平滑滤波和细节保护的矛盾，但经过差值滤波法处理后的图像存在有微小的锯齿。

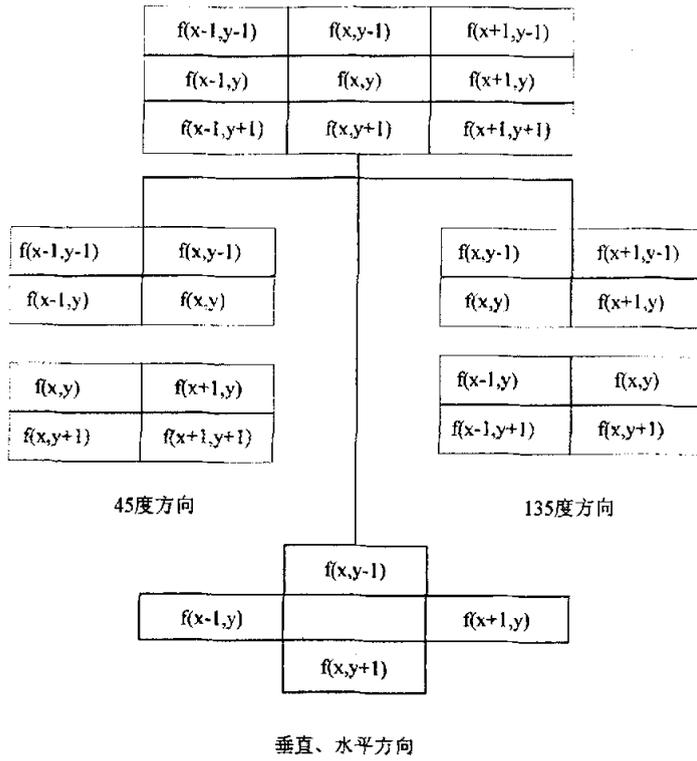
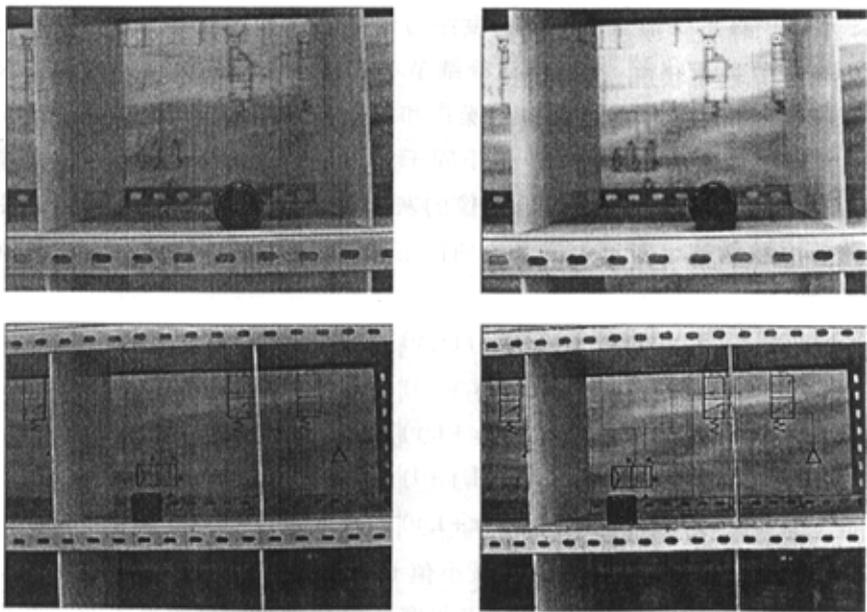


图 4.6 将图像窗口分解成 5 个 2×2 的像素窗口



(a)

(b)

图 4.7 (a) 滤波前图像, (b) 滤波后图像

4.3 图像分割技术与方法

图像分割一直是图像处理领域的一个比较活跃的分支^[34, 35], 是图像处理中的一项关键技术, 图像分割的任务就是合理地将图像中的目标与背景分离。虽然文献中包含的分割方法很多, 但是在具体应用中都存在各种各样的问题, 主要是这些分割方法对具有某种特定特征的图像具有较好的分割结果, 而对实际工作中遇到的其他类型的图像分割效果却较差。

4.3.1 图像分割的方法

分割图像的简单方法之一, 便是基于图像灰度直方图的整体阈值法, 它利用灰度直方图的统计结果, 求出呈现“双峰”的直方图的“谷点”作为合理的阈值来分割图像。这种整体阈值法虽然具有简单、高效的优点, 但却难以实用。因为现实中很难出现图像的灰度直方图恰好呈现“双峰”的情形。Chow 与 Kaneko 于 1972 年提出将图像划分为若干区块, 对各区块分别用基于直方图的阈值分割方法进行分割, Chow 的这种“分而割之”的方法, 对一类灰度虽然分布均匀但“缓慢”变化(随区块不同)的图像, 其分割效果较整体阈值法有明显的改善, 这使图像分割方法在实用性方面迈进了一步。Yasuo Nakagava 与 Rosenfeld 在 Chow 的基础上对算法作了进一步完善, 并提出了相应的改进算法, 以增强其实用性。此后, 人们又陆续提出了各种基于不同学科理论的分割算法, 如基于模糊论的模糊 C 均值(FCM 法), 基于区域相关的区域生长算法, 以及近年来日益受到关注的基于人工神经网络的 Hopfield 网络法等^[36~39]。

根据图像分割机理, 图像分割的方法主要有: 基于像素分类的分割, 将像素归属于不同的集合, 利用分类特征实现分割; 基于区域的方法, 这种方法假设图像的子区域有相同的性质, 不同区域没有共同性质; 基于边缘的方法, 这种方法假设图像中子区域之间一定会有边缘存在, 先确定边缘像素并把它们连接在一起以构成所需的边界; 基于纹理特征的分割, 这种方法利用纹理的统计分析或者结构分析的方法。图像分割的基本方法如图 4.8 所示。

不同的分割方法有不同的适用范围, 应该根据图像的特点选择。根据气动机械手图像的特点, 应该采用基于区域的分割方法。

4.3.2 阈值分割

阈值处理是一种区域分割技术, 它对物体与背景有较强对比景物的分割特别有用。它计算简单, 而且总能用封闭而且连通的边界定义不交叠的区域。

阈值分割的基本思想是确定一个阈值, 然后把每个像素点的像素值和阈值比较, 根据结果将该像素确定为前景还是背景, 所有灰度值大于或等于某阈值的像素都被判属于物体, 所有灰度值小于该阈值的像素被排除在物体之外。于是, 边

界就成为这样一些内部点的集合，这些点都至少有一个邻点不属于该物体^[36]。

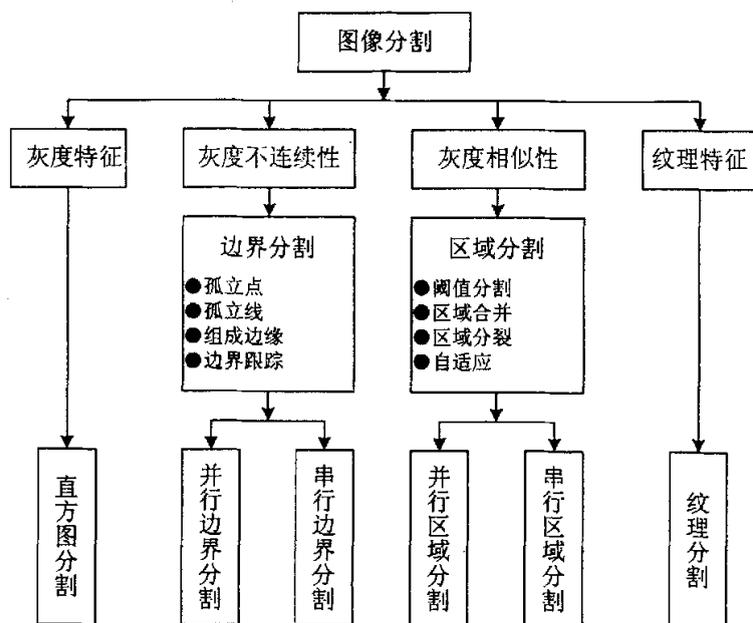


图 4.8 图像分割分类

如果感兴趣的物体在其内部具有均匀一致的灰度值并分布在一个具有另一个灰度值的均匀背景上，使用阈值方法效果就很好。如果物体同背景的差别在于某些性质而不是灰度值（如纹理），那么，可以首先把那个性质转化为灰度，然后，利用灰度阈值化技术分割待处理的图像。

采用阈值确定边界的最简单做法是在整个图像中将灰度阈值的值设置为常数。如果背景的灰度值在整个图像中可合理地看作为恒定，而且所有物体与背景都具有几乎相同的对比度，那么，只要选择了正确的阈值，使用一个固定的全局阈值一般会有比较好的分割效果。

阈值分割分为全局阈值和局部阈值两种，如果分割过程中每个像素的阈值都相等，称为全局阈值。

$$T = T(f) \tag{4.6}$$

如果每个像素所使用的阈值可能不同，称为局部阈值。

$$T = T(f, x, y) \tag{4.7}$$

除非图像中的物体有陡峭的边沿，否则灰度阈值的取值对所抽取物体的边界的定位和整体的尺寸有很大的影响。这意味着后续的尺寸（特别是面积）的测量对于灰度阈值的选择很敏感。由于这个原因，我们需要一个最佳的，或至少是具有一致性的方法确定阈值。最佳分割阈值的确定方法有很多，常用的是：试验法、直方图法和最小误差方法。

- (1) 试验法。根据图像先验特征，试验不同的阈值，直到满意为止。
- (2) 直方图法。如果图像前景和背景灰度比较均匀，直方图会出现明显的双峰，可以利用这个特性用峰值间的谷点作为门限值。
- (3) 最小误差法。这种方法假设前景和背景的分布都是正态分布，从概率上找到一个最优阈值。

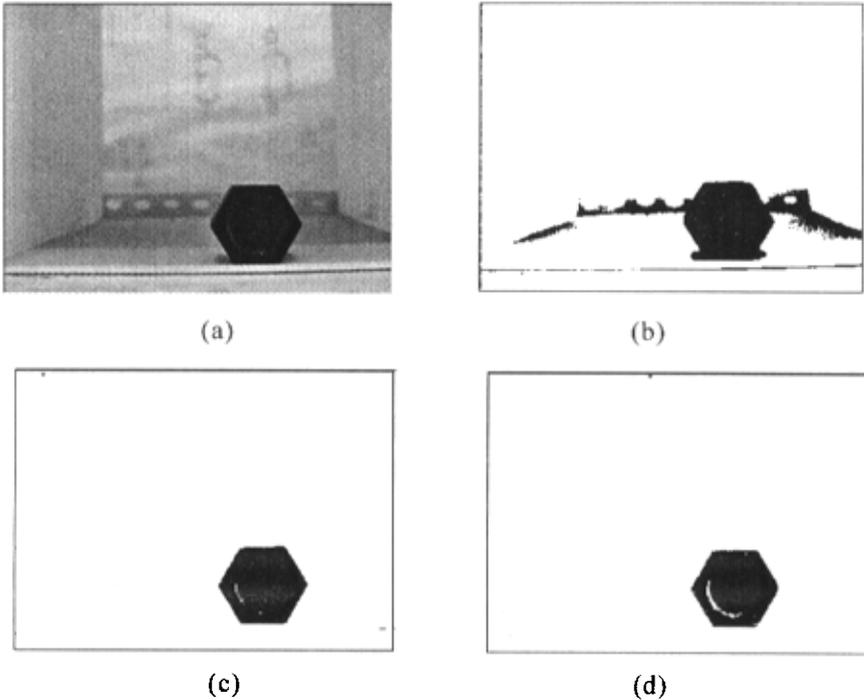


图 4.9 阈值分割结果

(a)原图 (b)阈值为 60 的分割 (c)阈值为 50 的分割 (d)阈值为 40 的分割

通过物体图像可以看出，物体总是占据直方图灰度最小的部分。通过统计试验，物体主要集中在图像灰度值最小的 2%上，即物体图像的分割阈值在 40~60 之间，图 4.9 为利用试验法确定阈值的分割结果。

在许多情况下，背景的灰度值并不是常数，物体和背景的对比度在图像中也有变化，图像的最佳分割阈值也会发生变化，因此利用试验法得到的分割阈值没有普遍适应性。当照明不均匀、有突发噪声，或者背景灰度变化较大时，整幅图像分割将没有合适的单一门限，可对图像按照坐标分块，对每一块分别选定阈值进行分割，称为动态阈值分割或者自适应阈值分割^[40]。

自适应阈值的选取一种方法是对每个像素确定以他为中心的一个邻阈窗口，计算窗口内的最大值和最小值，然后取他们的均值作为阈值。图 4.10 表示当前像素的 8 邻域窗口（实际可以根据需要选取）。

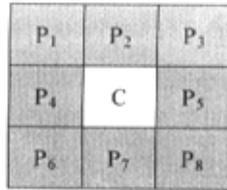


图 4.10 图像分割分类

P_i 表示当前像素 C 的 8 邻域像素灰度值，则阈值可以设定为：

$$T = \frac{\max(P_i) + \min(P_i)}{2} \quad (4.8)$$

另一种方法采用均值作为阈值：

$$T = \frac{\sum_{i=1}^8 P_i}{8} \quad (4.9)$$

对图像分块以后的每一个子块可以采用直方图分析，如果某个子块内有目标和背景，则直方图呈双峰^[41]。如果块内只有目标或者背景，则直方图没有双峰，可以根据邻域各块分割得到的参数插值进行分割。

针对本项目中图像的特点，笔者设计了一种简单的自适应阈值分割算法，其算法步骤为：

- (1) 对图像进行灰度拉伸；
- (2) 搜索图像灰度最大值 f_{max} ；
- (3) 将灰度值 $f_{max} - 150$ 作为该行的阈值对物体图像进行分割；
- (4) 对分割后的图像进行孤点去除。

一般情况下，由于物体的灰度值明显低于背景的灰度值，通过这种简单的方法，可以很方便地将物体和背景分割开来，图 4.11 显示了利用该算法对六角螺母图像进行处理的结果，(a)为原始图像，(b)为分割结果。这种算法原理简单，效果较佳。

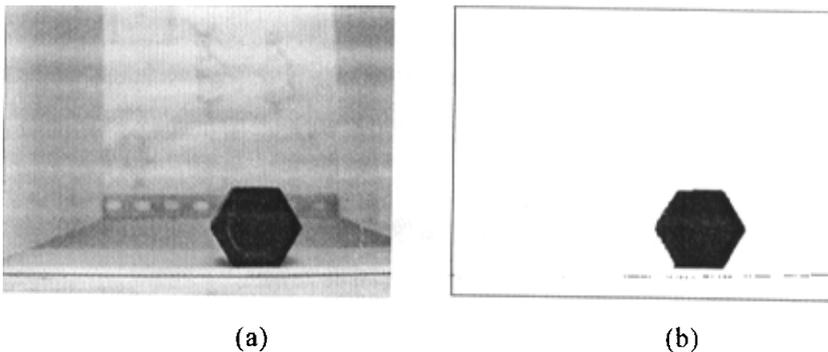


图 4.11 阈值分割结果

(a)原始图像，(b)自适应阈值分割结果

4.3.3 图像二值化

为了提高图像处理的速度,工业上一般采用处理过程相对比较简单二值化图像,因此如何合理选取二值化阈值,使之同时满足速度和精度的要求,就成了一个关键的问题。通过对多种图像处理算法进行比较,发现使用直方图同高斯滤波相结合的方法来求取二值化阈值,对图像进行二值处理,速度特别快^[42],精度也完全能够满足测量的要求。

4.3.3.1 直方图技术

一幅含有一个与背景明显对比的物体的图像具有包含双峰的灰度直方图。两个尖峰对应于物体内部和外部较多数目的点。两峰间的谷对应于物体边缘附近相对较少数目的点。在类似这样的情况下,通常使用直方图来确定灰度阈值的值。

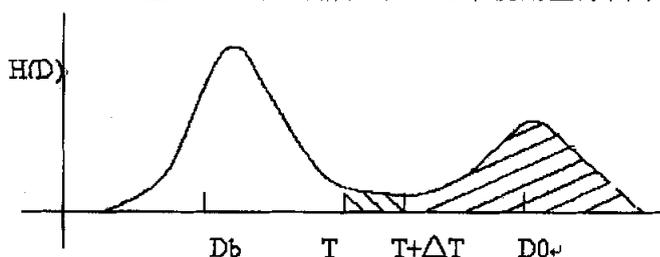


图 4.12 双峰直方图

利用灰度阈值 T 对物体面积进行计算的定义是:

$$A = \int_{T}^{D_0} H(D) dD \quad (4.10)$$

显然,如果阈值对应于直方图的谷,阈值从 T 增加到 $T + \Delta T$ 只会引起面积略微减少。因此把阈值设在直方图的谷点,可以把阈值选择对面积测量的影响降到最低。

如果图像或包含物体图像的区域面积不大且有噪声,那么,直方图本身就会有噪声。除了凹谷特别尖锐的情况外,噪声会使谷的定位难以辨认,或至少是通过不同幅图像得到的结果不稳定可靠。这个问题在一定程度上可以通过用卷积或曲线拟合过程对直方图进行平滑加以克服。如果两峰大小不一样,那么,平滑化可能会导致最小值的位置发生移动。但是,在平滑程度适当的情况下,峰值还是容易定位并且也是相对稳定的。一种更可靠的方法是把阈值设在相对于两峰的某个固定位置,如中间位置上,这两个峰分别代表物体内部和外部点典型(出现最频繁)的灰度值。

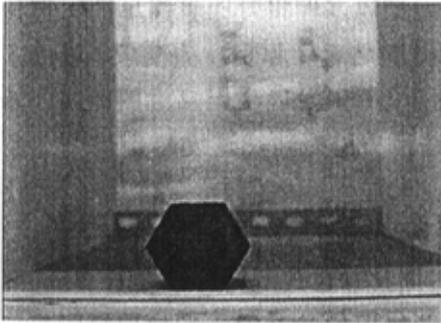


图 4.13 零件原始图

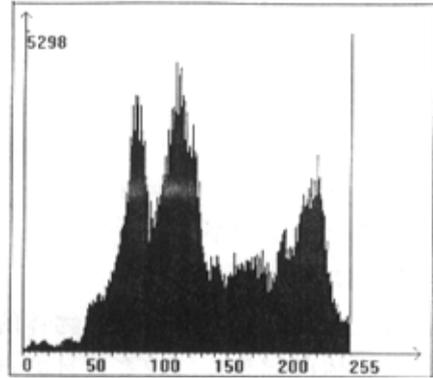


图 4.14 零件直方图

观察图 4.13 的原始图像及图 4.14 所示的相应直方图，可以发现，原始图像中，背景较亮，直方图上相应的像素多集中在直方图的高段，零件较暗，其相应的像素集中在直方图的低段。因此，根据直方图可推断图像的某些特征。例如，直方图偏低，意味着图像偏暗；直方图偏高，意味着图像偏亮；直方图偏窄，意味着图像对比度偏低。事实上，直方图可以作为图像质量的一种主观评价。

4.3.3.2 二值化阈值的确定

从直方图中可以看出，背景和零件在直方图上都出现了峰值，如果把二值化阈值取为双峰之间的谷底，就可以将零件和背景最大限度的分割开来。从图 4.5 中可以看出直方图并不是一个规则的双峰形，直接求取双峰之间的谷点有些困难。一般情况下，直方图中的不规则可以认为是由白噪声引起的。因为高斯函数能较好地滤除白噪声，所以可采用加窗高斯函数对直方图进行滤波^[43-45]。加窗高斯函数表示如下：

$$G(x) = (1/2\pi\delta) \exp\left\{-\frac{(x-\delta)^2}{\delta}\right\} \quad (4.11)$$

将加窗高斯函数的离散值同直方图函数的离散值进行卷积运算，就完成了滤波操作。设加窗高斯函数的窗口大小为 δ ，直方图的灰度等级为 h ，卷积后的结果为 $R(x)$ ，则离散卷积的公式如下：

$$R(x) = \sum_{k=0}^{20} P'(x-k)G(k) \quad (4.12)$$

其中， $P'(x)$ 是由直方图的离散值 $P(x)$ 扩展所得到的：

$$P'(x) = \begin{cases} P(x) & x \leq h \\ 0 & x > h \end{cases} \quad (4.13)$$

$\delta=15$ 时的高斯曲线如图 4.6，滤波后的结果如图 4.7。

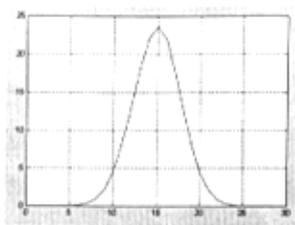
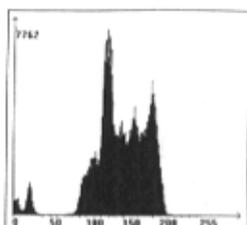
图 4.15 $\delta=15$ 高斯曲线

图 4.16 滤波后的直方图



4.17 零件二值化图像

从图 4.16 可以看出, 直方图经窗口高斯函数滤波呈现规则的双峰特征, 这是就可以用一阶差分算得到直方图各个极值点对应的灰度值, 但是, 因为直方图的峰值和谷点都是极限点, 所以, 在这些地方都会出现过零点, 而且当高斯函数的窗口 δ 较小时, 高斯函数对直方图细节的影响较小, 对灰度变化比较敏感, 从而导致过零点较多, 不利于得到阈值。实验表明对于图 4.13 所示的图像, 当 $\delta=8$ 时, 极小值点数为 5, 但是, 当 δ 增大时, 极小值点数大部分情况下都是 3, 于是可以简单地判定第二个极小值点就是所求的二值化阈值。用这个阈值对原始图像分割, 就得到了图 4.17 所示的二值化图像。

4.4 图像特征识别

4.4.1 图像特征

前面介绍了模式识别并且讨论了怎样从一个复杂的景物中分割和抽取物体。下面将着重论述物体特征识别方面的内容。

如果要建立一个识别不同种类对象的系统, 必须首先确定应测量对象的那些特征以产生描述参数。被测量的这些特殊属性被称为特征, 而所得的参数值组成了每个对象的特征^[46-47]。

根据本课题的要求, 图像处理的最终结果是找出物体的形心坐标。物体的特征中反映其尺寸的有面积、周长等。在本研究中尽量做到形心的计算准确, 这样在抓取物体时就不会因为物体的重心偏离而导致物体滑脱, 从而使物体抓取的准确性得到提高。

面积是物体总尺寸的一个方便的度量。面积只与该物体的边界有关, 而与其内部灰度级的变化无关。面积可以很容易地从已分割的图像抽取物体的过程中计算出来。

边界定义 在给出一个计算物体面积的算法之前, 必须明确关于物体边界的定义。尤其应该注意的是必须确认是在测量同一个多边形的面积。

像素计数面积 最简单的面积计算方法是统计边界内部(也包括边界上)的像素的数目。

4.4.2 面积与形心定位

对图像进行相应的处理后，得到一幅效果很好的二值化图像，其目的是为了计算图像中零件的特征参量。我们所使用的特征参量主要有面积、零件的形心坐标等。有了这几个特征量，再经过坐标变换，将这些特征量转化为零件在空间坐标中的实际位置量，就可以控制机器人完成相应的操作。

计算平面图形的形心公式如下：

$$\bar{x} = \frac{\int x dA}{A}, \bar{y} = \frac{\int y dA}{A} \quad (4.14)$$

式中：A 为平面图形的面积，其它符号的含义见图 4.18。

形心公式不难理解，将它应用到图像处理中进行形心计算也很容易。我们将一个像素看成一个 dA，且为一个单位，则图像的面积即为像素点的个数，分子上的积分则变为像素坐标的和。以求和代替积分，则式 (4.14) 可变为

$$\bar{x} = \frac{\sum_{x=1}^M x \cdot f(x, y)}{\sum_{x=1}^M \sum_{y=1}^N f(x, y)}, \quad \bar{y} = \frac{\sum_{y=1}^N y \cdot f(x, y)}{\sum_{x=1}^M \sum_{y=1}^N f(x, y)} \quad (4.15)$$

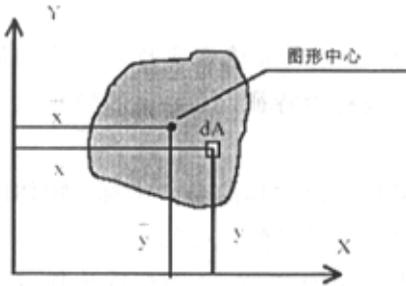


图 4.18 形心的计算

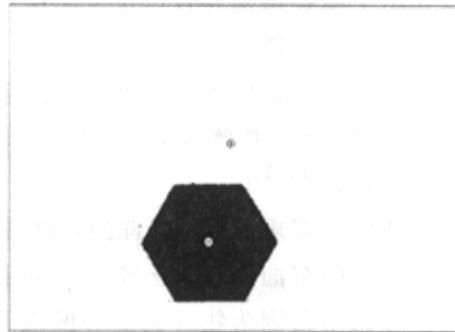


图 4.19 二值图像及形心点

按照上述的计算方法，算出图 4.15 所示的零件的形心坐标为：

	X 方向 (pixel)	Y 方向 (pixel)
图像中心	386.33	280.55
零件形心	341.54	442.80

4.5 本章小结

本章主要对图像处理及其特征识别进行研究。首先讨论了检测图像的预处理

理,进而对图像分割方法进行了研究。利用分段线性变换方法进行了对比度增强,分析了图像检测中 CCD 摄像机的噪声影响要素,并给出了相应的滤波方法。针对图像的特点,笔者设计了一种简单的自适应阈值分割算法,经过实验比较,改善了处理效果。对图像进行相应的处理后,得到一幅效果很好的二值化图像,其目的是为了计算图像中零件的特征参量。本章利用的特征参量主要有面积、目标物体的形心坐标等。有了这几个特征量,再经过坐标变换,将这些特征量转化为目标物体在空间坐标中的实际位置量,就可以控制机械手完成相应的操作。

第 5 章 视觉跟踪理论与路径规划

5.1 引言

本文研究的气动机械手是采用视觉传感技术实现抓取跟踪的。视觉跟踪 (Visual Tracking), 在机器人视觉伺服控制文献中也常被称作“图像特征跟踪” (Image Feature Tracking)^[48], 主要用于跟踪机器人末端工具和移动工件, 以实时把它们的位置信息反馈到控制器。

视觉跟踪是计算机视觉中的一个目前相当活跃的研究领域, 并且可用于多种场合: 例如跟踪飞机、导弹、车辆、人和动物等。在机器人视觉伺服里视觉跟踪用于跟踪机器人末端工具和移动工件, 以实时把它们位置信息反馈到控制器。

视觉跟踪可一般性地定义为^[48]:

给定

- 1) 图像 I_0 中的一个特定目标;
- 2) 目标的状态 S_0 ;
- 3) 一系列图像, I_1, I_2, \dots 。

则产生对目标相应状态一系列估计 S_1, S_2, \dots 。在机器人视觉伺服应用中, S_1 一般代表目标的位置向量。

视觉跟踪可分为物体识别和跟踪两个过程。根据物体的识别方法不同, 可以将视觉跟踪大致分为三类^[48]: 基于模板的跟踪 (Template-based Visual Tracking)、基于活动轮廓的跟踪 (Snake Tracking) 和基于分割的跟踪 (Segmentation-based Tracking)。

(一) 基于模板的区域跟踪^[49]

基于模板区域跟踪的基本思想是按像素对感兴趣的区域进行一帧帧地匹配, 基本的方法为^[48]:

- 1) 从初始图像 I_0 中选择一感兴趣的区域 R_0 , R_0 在 I_0 的位置记为 u_0 。
- 2) 对于系列图像 I_1, I_2, \dots 。
- 3) 采用某种方法搜索 u_i , 使得 R_{i-1} 在图像 I_i 获得最好的匹配。
- 4) 在 u_i 位置, 从图像 I_i 中采样 R_i 。

这一系列的区域位置坐标 u_i 就构成了被跟踪目标的在这一系列图像中的位置。

采用基于模板的跟踪算法主要应用于目标外形和光照条件等环境因素基本保持不变的情况。因为需要进行图像相关性分析, 模板的尺寸和物体在相邻两帧

图像间的位移对运算速度影响较大,例如文献 50 中采用的模板窗口尺寸为 16×16 ,每帧图像中物体最大位移为 $\Delta x = \Delta y = \pm 32$ 像素,跟踪时间大约为 116 毫秒。在应用中为了减少处理时间,一般通过重采样技术动态改变图像分辨率。从算法原理上可以看出,这种方法无法分辨低于一个像素的运动(指相邻两帧之间)。

(二) 活动轮廓跟踪^[51,52]

活动轮廓(snake)是一个能量最小化样条,其跟踪过程可以描述如下:

- 1) 从一幅初始图像 I_0 中确定一个初始活动轮廓 C_0 。
- 2) 对于系列图像 I_1, I_2, \dots
- 3) 沿着当前的活动轮廓的离散控制点 s_1, s_2, \dots, s_n ,沿活动轮廓正交方向按某种策略寻找一个边沿,并将 s_j 到边沿的垂直距离 d_j 保存下来。
- 4) 根据所有 (s_j, d_j) 信息,用某种拟合方法,比如最小二乘法,估计在 I_i 中新的活动轮廓 C_i 。

这类方法最大特点是从整体上识别物体,在物体具有变形、背景复杂和局部被遮挡的情况下,仍可以得到较为满意的跟踪效果,具有较强的鲁棒性。但是这种方法要求被跟踪物体具有清晰的轮廓,尤其当局部遮挡现象发生时更是要求这样。为了保证轮廓的参数具有良好的行为(well-behaved),一般都限制搜索范围,例如文献 51 将搜索范围限制到从物体中心出发的一系列射线上,避免了轮廓发生重迭交叉的可能。

(三) 基于分割的跟踪^[53, 54]

基于分割的方法一直是视觉跟踪的主流方法,其基本原理是从包含有目标物体的图像中将目标和背景分割出来,用相应的特征参数对其位置进行描述,分割的线索包括了图像强度、颜色、运动、纹理、灰度直方图等。“灯泡”位置跟踪,是这种方法的一种典型应用,它的基本步骤如下:

- 1) 任何一个图像 I_0, I_1, I_2, \dots

在图像 I_i 位置 u_{i-1} 获得一个区域 R_i , 该区域含有目标物体

- 2) “灯泡” $B_i = \pi(R_i)$
- 3) 计算 $u_i = u_{i-1} + \text{centroid}(B_i)$

“灯泡”位置跟踪假设一个区域仅有一个被跟踪的目标,目标的尺寸基本上不变,并且不能脱离被选择的区域。由于分割本质上是基于二值化图像的,因此可以容易地将多种线索组合起来,以克服光照条件变化、遮挡等不利因素,从而构建一个更加鲁棒的系统。基于分割的方法采用不同的分割策略,计算量的差别将很大,文献[51]报道的处理速度为 60Hz,而文献[52]报道的处理速度为 30Hz。另外,这种方法的跟踪精度受环境因素影响较大。

以上介绍了三种常见的视觉跟踪方法,它们都各自有优缺点,例如基于模板的视觉跟踪的方法简单,精度高,但计算量大,鲁棒性差;而活动轮廓跟踪的鲁

棒性好，计算量也不大，但是要求跟踪物体具有清晰的轮廓，精度也不是很高；基于分割的跟踪方法，在确定的环境下可以同时使用多个特征来进行跟踪，具有较高的精度，也具有较强的鲁棒性，但在非确定环境下，这种方法的性能就会下降很多。

本文所研究的气动机械手视觉跟踪系统中，被抓取的目标物体是静止的，物体轮廓和形状以及光照等环境条件基本不变，因此采用了基于分割的跟踪方法，对目标物体的图像进行分割和二值化后，利用形心坐标来描述目标物体的位置特征。

5.2 机械手抓取物体及轨迹计算

5.2.1 物体抓取路径规划

通过分割获得目标物体的位置描述以后，下一步的工作就是根据物体的位置与机械手当前位置的偏差，确定机械手的移动轨迹和速度，以便快速准确地靠近目标物体，这就涉及到机械手路径规划问题。

路径规划技术是机器人研究领域中的一个重要分支。路径规划可以用 Durrant-Whyte H F^[55]提出的三个问题（1）“我现在何处？”，（2）“我要往何处去？”，（3）“要如何到该到处去？”来描述。即给定一个移动机器人所处的环境（环境可以通过移动机器人视觉系统或其它途径获得），一个起始点和一个期望的终止点，机器人路径规划根据一定的任务要求，寻求一条连接起始点到终止点且能避开环境中障碍物的运动轨迹，即最优或次优有效路径。

运动轨迹是机器人系统工作的依据，它决定了系统的工作方式和效率，机器人系统要完成某种操作作业，就必须对其运动轨迹进行规划，因此研究机器人系统运动轨迹的规划尤为重要。轨迹规划可分为关节坐标空间的轨迹规划和笛卡尔坐标空间的轨迹规划^[56~59]。

关节空间进行轨迹规划，需要给定路径点 A、B 的位姿，进行位姿反解，即可求出相应的关节变量 q_A 和 q_B （对于转动关节即 $\theta_{A1}, \theta_{A2} \dots; \theta_{B1}, \theta_{A2} \dots$ ）。然后对关节进行插值，得到中间点的关节值。关节坐标空间轨迹规划不必考虑两路径中间点的位姿要求，如图 5.1 所示，由 A 到 B 可由 1 或 2 或 3……到达。关节空间轨迹规划的方法简单，不会产生奇异位置。

在本课题中，机械手的定位跟踪是通过两个气缸的伸缩来实现的，计算机通过图像处理得到目标物体的位置描述，并将其转换成空间坐标关系，然后按照一定的路径插补算法计算各个气缸的位移量，完成跟踪并抓取物体。

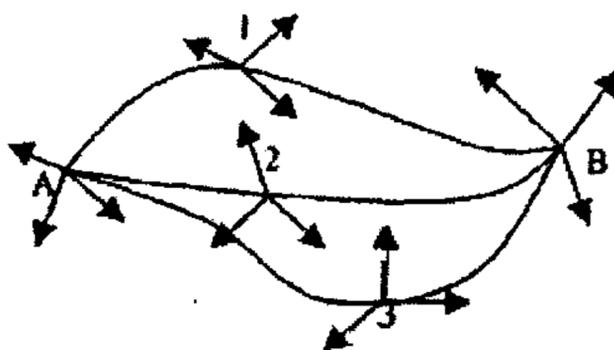


图 5.1 关节空间轨迹规划

5.2.1.1 直线轨迹插补

对于一般的机器人轨迹跟踪系统来说，起始点和目标点的位置是已知的（本文可由视觉系统得到），机械手从起始点到目标点的运动通过轨迹插补实现。轨迹插补可以获得中间点的坐标，通过机器人逆运动学分解算法，转变为对应的气缸位移，然后由控制系统实现气缸运动控制，其计算流程如图 5.2 所示。

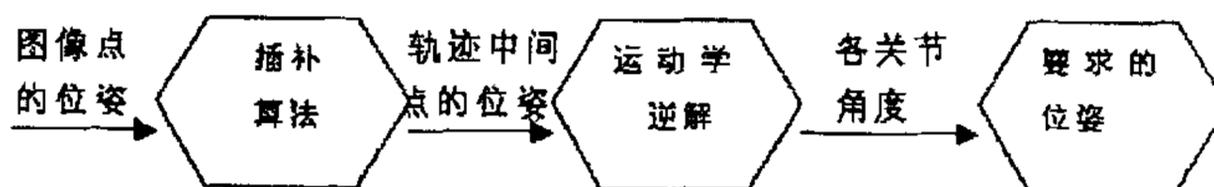


图 5.2 轨迹插补流程

从上面介绍的内容可以知道，机器人实现一个空间轨迹过程，是实现轨迹离散点的过程，如果这些离散点间隔很大，机器人运动轨迹就与要求轨迹有较大误差。只有这些离散点（插补得到的）彼此很近，才有可能使机器人以足够精度逼近要求的轨迹。

实际上，机器人运动是从一点到一点的过程，如果始末两点距离很大，机器人只保证运动经过这两点，但不能保证这两点中间路径，也就是说其两点中间路径不确定。如果插补中间点足够密集，则能逼近要求的曲线。

直线插补和圆弧插补是机器人系统中两种基本的插补算法。对于非直线和圆弧插补轨迹，可以采用直线或圆弧逼近，以实现这些轨迹。对于本课题的气动机械手来说，我们关心的只是目标点的位置，和具体的跟踪路径无关，因此采用简单的平面直线插补。

平面直线插补是根据已知直线始末两点的坐标，求取轨迹中间点（插补点）的坐标。本文用来运动合成的两个气缸运动不是成正交关系，也就是说每个气缸的运动对机械手末端的两个坐标都有影响，不能用传统的插补算法，这里需要做一些简化和假设。从第二章的机械手运动分析知道，大臂摆动气缸主要影响机械手 Y 坐标，大臂伸缩气缸主要影响 X 坐标，因此可以分别控制两个气缸的运动实现直线轨迹插补。

已知机械手起始点坐标 $P1(x1, y1)$ ，目标物体的坐标为 $P2(x2, y2)$ ，这些

坐标值是通过图像处理及坐标变换得到的，其他参数含义见第二章气缸运动分析，可以通过以下步骤进行直线轨迹的插补：

1. 给定步长 L_s ，在程序实现中，步长可以由操作者输入；

2. 求轨迹直线方程： $(y - y_1) = \frac{(y_2 - y_1)}{(x_2 - x_1)}(x - x_1)$

3. 由于机械手结构限制，只需要考虑第一象限插补，计算插补点与轨迹直线的距离 L

$$L = (y - y_1) - \frac{(y_2 - y_1)}{(x_2 - x_1)}(x - x_1)$$

4. 计算插补方向标志 a ：

如果 $(x_2 + y_2) \geq (x_1 + y_1)$ ， $a=1$ ；反之， $a=0$

5. 插补运算

当 $a=1$ 时，正向插补，如果 $L \geq 0$ ， $x_2 = x_2 + L_s$ ；反之， $x_1 = x_1 + L_s$

当 $a=0$ 时，反向插补，如果 $L \geq 0$ ， $x_1 = x_1 - L_s$ ；反之， $x_2 = x_2 - L_s$

6. 反复进行 (3)、(4)、(5)，直到 $(y_2 - y_1)^2 + (x_2 - x_1)^2 < \varepsilon$ ， ε 为给定的正数，表示抓取定位精度。

5.2.1.2 直线插补算法的仿真

根据以上算法，对机械手直线轨迹插补进行了方针分析，初始条件如下：

机械手初始坐标 (800,1300)，目标物体形心坐标(850,1450)，这些都由图像处理得到；插补步长为 0.5mm，抓取定位精度为 1mm，其他条件见第二章机械手运动分析。

用 Matlab 进行仿真分析的结果如图 5.3 和 5.4 所示。

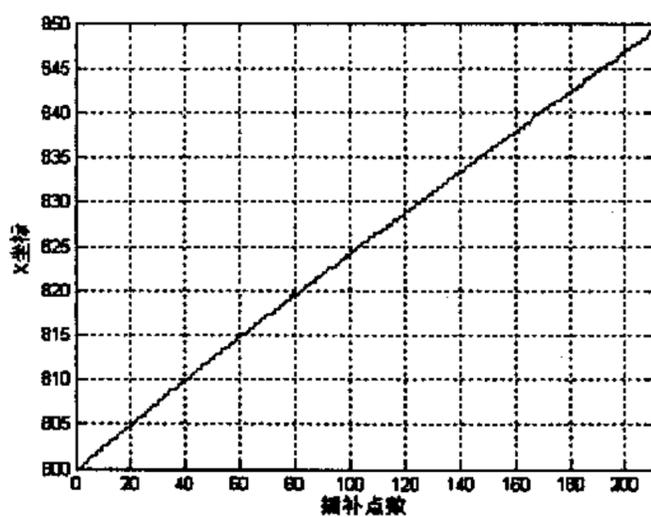


图 5.3 x 轴插补仿真结果

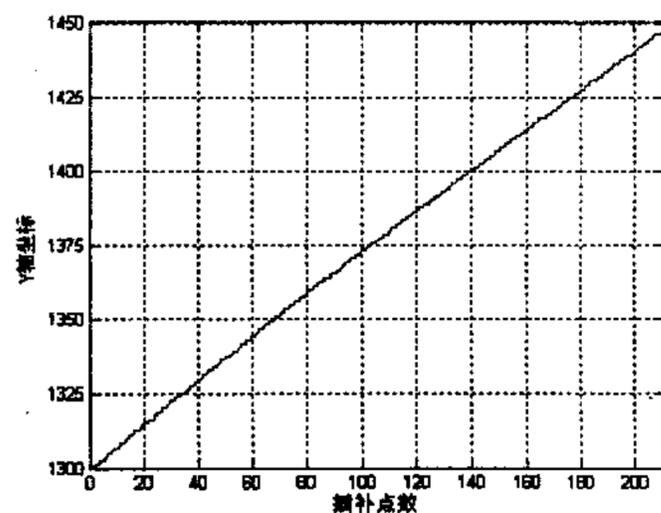


图 5.4 y 轴插补仿真结果

5.3 视觉跟踪系统的实现

在气动机械手系统中，系统的准确定位依靠图像检测来实现。当机械手接近目标物体时，视觉系统根据采集到的图像，自动计算当前的相对位置偏差，并转

换成空间坐标,通过计算机控制程序发出跟踪控制命令,驱动气缸完成位置跟踪,对准后进行抓取。在机器人控制中,跟踪一般靠插补完成,下面介绍视觉系统的跟踪插补算法。

5.3.1 跟踪算法实现

本课题中的跟踪系统采用基于目标形心的跟踪算法。通过寻找图像中物体的形心,以它偏离图像中心的方向为控制量,通过直线插补可以实现目标跟踪。整个过程可以分三步完成。

- 1) 图像分割与特征提取
- 2) 计算目标形心坐标
- 3) 跟踪控制

前两步的实现方法在第四章中已经做了大量讨论。这里就第三步跟踪控制的实现方法作以讨论。

求出的目标形心坐标 (l, k) 与图像中心坐标 $(384, 288)$ 进行比较,并通过尺度变换计算出机械手的初始坐标和目标点坐标,利用直线轨迹插补控制机械手的运动并抓取物体。

5.3.2 实验结果

为了验证跟踪子系统的可行性和有效性,本文做了分别以 x 轴、 y 轴做了跟踪实验。其结果如下图:

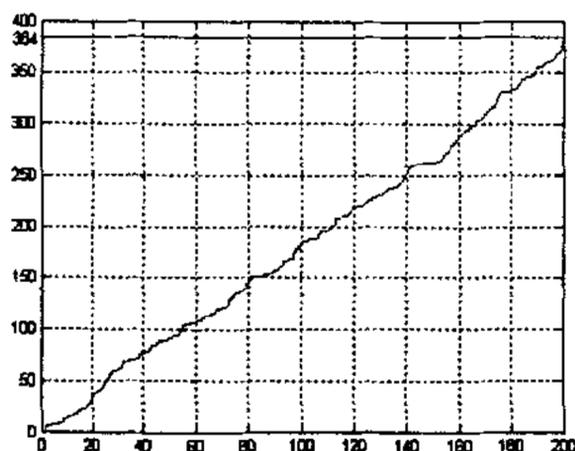


图 5.5 x 轴跟踪结果

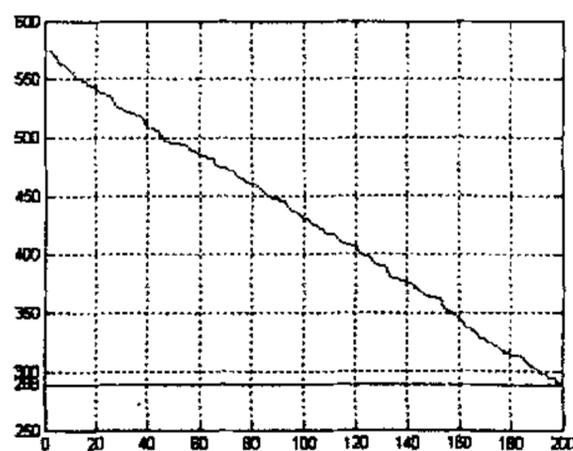


图 5.6 y 轴跟踪结果

通过和仿真结果进行比较,和实际情况比较接近(注:曲线 Y 坐标为像素表示的相对坐标, X 轴表示采样点数)。

5.4 本章小结

本章介绍了三种常用的视觉跟踪方法,比较了他们的优劣,根据本项目的特点选用了基于分割的跟踪算法,利用目标物体图像的形心坐标描述物体位置,设

计了直线轨迹插补算法,通过逆运动学分析规划机械手的运动轨迹,转化成运动气缸的目标位移,实现机械手跟踪控制,对跟踪结果进行了仿真分析和试验验证。

第 6 章 气动机械手视觉系统设计

6.1 软件平台概述

气动机械手为实验室建设项目,为降低成本,计算机系统仅采用普通 PC 机,各种执行机构的控制、图像的处理以及各种信号的监测均在此计算机上实时完成,这无疑大大增加了本文算法的程序设计的难度。为此,控制软件除了要在优化算法的基础上之外,还必须在计算机软件平台、程序结构等方面进行深入研究,以实现在较低的硬件条件下,完成一系列的功能。需要指出的是,在软件设计过程中,尽量减少人工干预,设计的软件还应操作简单,人机交互方便,软件界面美观,并注意与硬件的配合,尽量发挥出硬件的潜力。

本文选用 LabVIEW 作为软件开发平台,其最大的特色是采用编译型图形化编程语言——G 语言(Graph Programming),即用户设计好程序的大体框架后,如同画流程图一般,只需将系统提供的各种图形化功能模块连接起来,就可得到所需的应用软件,大大节约了开发时间。

另外,鉴于 LabVIEW 不适于进行大量数据处理,又不能完全提供用户所需要的驱动子程序,对底层操作不易实现,可以借助其他语言(C 语言)或利用其它软件开发环境(如 Visual C++、MATLAB 等)开发满足特殊功能的动态链接库文件成为可以考虑的解决途径。本课题中就是通过接口调用其他语言弥补其不足之处,从而达到扩充 LabVIEW 功能,满足系统要求的目的。

6.1.1 虚拟仪器开发平台 LabVIEW

LabVIEW 是 NI 公司推出的图形化虚拟仪器开发平台,各个专业领域的工程师、科学家通过定义和连接代表各种功能模块的图标,可方便迅速地建立起通常只有高超编程技巧的程序员才能编制的高水平应用程序。因此,它又被称为“面向工程师和科学家的编程平台”^[60]。

LabVIEW 具有以下特点:

1. 图形化和数据驱动式的开发环境。LabVIEW 使用可视化技术建立人机界面。针对测试测量的过程控制领域,LabVIEW 提供了虚拟仪器面板上所必须的大量显示或控制对象,如表头、旋钮、图表等。用户还可以方便地将现有控制对象修改成适合自己工作领域的控制对象。LabVIEW 用图标表示功能模块,使用图标间的连线表示各种功能模块间传递的数据流,使用数据流程图式的图形化语言编写代码。开发时开发者在开发环境的前面板定制界面,后面板会自动生成与前面板相应的功能图标,开发者只要按照自己的意图将这些功能图标连接起来即

可完成某种特定的功能（图 6.1）。

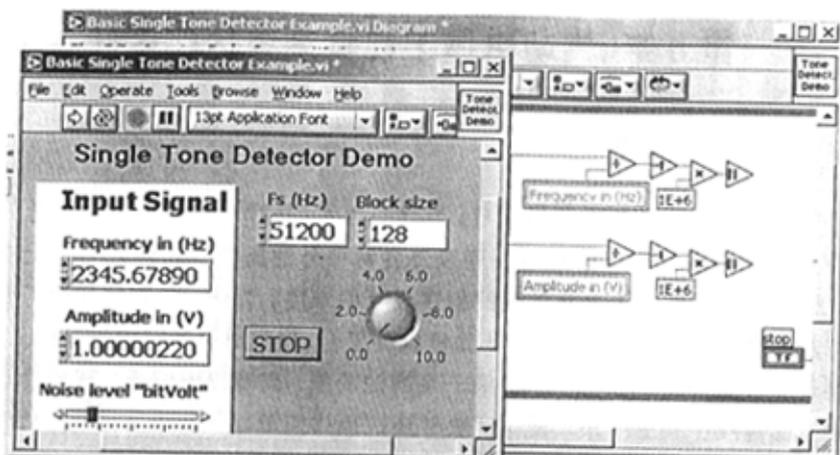


图 6.1 LabVIEW 的前面板与后面板

2. 灵活的程序调试手段和高速编程效率。用户可在源代码中设置断点单步执行源代码,在源代码的数据流上设置探针,还可在程序运行过程中观察数据变化。LabVIEW 采用编译方式运行 32 位应用程序,这就解决了其他按解释方式工作的图形化编程平台速度慢的问题,其运行速度相当于 C 语言。

3. 功能强大的函数库^[61]。LabVIEW 提供了大量的基本函数库供用户直接调用。从底层的 VXI、GPIB、串口及数据采集板的硬件控制子程序 600 多个仪器驱动程序,从基本的数学函数、字符串处理函数、数据运算函数、文件 I/O 函数到高级分析库(包括信号处理、窗函数、滤波器设计、线性代数、数理统计与概率论、曲线拟合等),涵盖了虚拟仪器设计中几乎所有需要的函数。此外,LabVIEW 还支持 NI 各种专用函数库,来实现特殊功能,只要安装相应的软件,用户就可以像调用基本函数库中的函数一样来调用库中的功能模块。如 NI SQL 封装了复杂的数据库操作程序,用户可以通过 NI SQL 很容易的进行数据库操作;NI IMAQ VISION 包含了各种机器视觉开发用的算法,基于它用户可以快速的开发自己的机器视觉应用。

4. 支持多种系统操作平台。在 Windows2000/NT/9X、Power Macintosh、HP-UX、SUN SPARC、Concurrent Computer Corporation 的实时 Unix 系统平台上,NI 公司都提供了相应版本的 LabVIEW,并且在任何一个平台上开发的 LabVIEW 应用程序可直接移植到其他平台上。

5. 开放式的开发平台^[62]。LabVIEW 提供 DLL 接口,使用户能够在 LabVIEW 平台上调用其他软件平台编译的模块,从而在 LabVIEW 环境下可以控制用户自己开发的专用仪器硬件。LabVIEW 提供了 DLL 接口,可以使用户将自己用 C 语言编写的程序集成到整个软件系统中来,用户可以根据项目的需要,在 LabVIEW

现有的功能模块基础上定制自己的算法和功能。本课题在实现读取光电编码器时,就将自己编写的 C 语言程序通过 DLL 接口嵌入到整个软件系统中。LabVIEW 提供了与 LabWindows/CVI 源代码相互调用的接口,提供对 MATLAB 的接口,可与其他应用软件一起构成功能更为强大的应用程序开发环境。

LabVIEW 如今已被认为是虚拟仪器开发的通用平台,它也越来越被广大的开发人员所接受。

6.1.2 IMAQ vision

NI 公司的 IMAQ Vision 软件包为图像处理提供了完整的功能。它将 400 多种功能整合到应用软件中,以实现功能强大的图像处理解决方案。现在的图像采集用户可以利用 LabVIEW 图形化编程的优异性能和高效率,利用 CVI 对 C 代码的控制能力的流畅性,开发出基于通用编程语言的解决方案。

◆ IMAQ Vision 的主要功能^[63]

NI 公司的 IMAQ Vision 软件包在 LabVIEW、BridgeVIEW、LabWindows/CVI、ComponentWorks 以及其他 ActiveX 容器应用程序中加入了机器视觉和图像处理的功能。IMAQ Vision 中包含一套丰富的为 MMX 而优化的函数,可用来完成灰度、彩色以及二值图像的显示、处理(统计、滤波和几何变换)、形状匹配、斑点分析、计算和测量等。最终用户、系统集成商和原始设备制造商都可以使用 IMAQ Vision 以加快工业视觉和科学图像应用的开发。IMAQ Vision 可用于工厂和实验室里那些需要高可靠性、高速的视觉系统的自动化操作中。

◆ IMAQ Vision 的主要特点

1) 加快了应用程序开发速度

IMAQ Vision 考虑了使用的简便性,尤其能满足图像应用开发人员的需求,以减轻他们在降低成本与缩短上市时间方面的压力。开放的内存管理以及符合命名逻辑的 VI、函数和参数使得 IMAQ Vision 易于学习;其内建的各种高级函数可以直观地相互配合工作,因此可以使用很少的函数而得到更快的开发速度。

2) 使用 Intel MMX 技术以提高性能

Windows NT/98/95 版的 IMAQ Vision 利用了 Intel MMX 技术。在图像处理应用中处理 8 位图像时,通常会用到整数或浮点运算函数,而 MMX 技术可以加快这些运算的速度。使用具有 MMX 技术的奔腾处理器执行很多 IMAQ Vision 的函数时,其性能要比不具有 MMX 技术的奔腾处理器提高最大 4 倍。

3) ActiveX 控件—ComponentWorks IMAQ Vision

IMAQ Vision ActiveX 控件是 ComponentWorks 产品家族的一员,它为 Microsoft Visual Basic、Microsoft Visual C++和其它开发环境提供了三个等级的控件:浏览器控件可以 30 帧/秒或更高的速率显示图像,并且提供了感兴趣

区域 (ROI) 选择工具; NI-IMAQ 图像采集控件使用属性页的形式简化了图像捕捉; 机器视觉和图像处理控件提供了上百种处理函数。

4) DAQ 配合工作进行图像采集^[64]

IMAQ 软、硬件的设计目标之一就是能够很容易的与 NI DAQ 产品集成。IMAQ 硬件上的 RTSI 总线可以使不同板卡使用共同的定时信号, 从而实现图像和数字量的定时、同步。

◆ IMAQ Vision 的使用环境

IMAQ Vision 针对不同开发环境有不同的使用方式, 开发人员可以根据需要进行灵活选择。利用 LabVIEW、BridgeVIEW 等图形化开发环境可以加快开发速度, 同时提高系统的可靠性。这种环境下 IMAQ Vision 的各种处理功能以 VI 的形式出现: 若开发者习惯于常规语言的开发环境, 可选用基于 C 语言的开发环境 LabWindows/CVI, IMAQ Vision 提供了丰富的 C 函数库以供调用; 若开发人员想利用现有的通用开发环境, 可以使用 IMAQ Vision ActiveX 控件, 它可以在 Microsoft Visual Basic、Microsoft Visual C++和其它 ActiveX 开发环境中以可视化控件的形式提供图像处理功能。

◆ 图像采集驱动软件 NI-IMAQ

不论使用 LabVIEW、LabWindows/CVI、C 或是 Visual Basic, NI IMAQ 都能提供对 NI IMAQ 采集设备的高层控制。NI-IMAQ 是针对图像采集的一个完整而可靠的 API。NI-IMAQ 可以完成图像采集所需的与计算机和板卡有关的工作而无须进行寄存器级编程。

NI-IMAQ 与 NI-DAQ 和其它的 NI 驱动软件完全兼容, 以便将图像集成到任何基于 NI 产品的解决方案中。

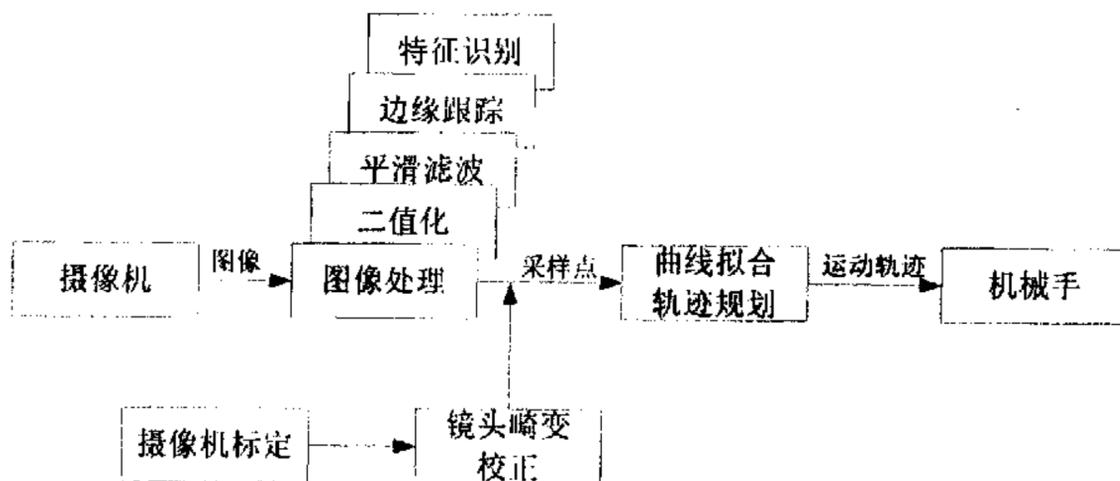


图 6.2 系统结构示意图

6.2 系统功能

- ◆ 本系统的关键在于软件设计。整个系统的优劣在很大程度上取决于软件设计的质量。系统软件要求可靠性高、有升级空间、易于使用、有

一定的智能化。因此本系统软件采用模块化设计思想，将系统功能进行分割。每个功能都用一个相对独立的软件模块实现，可以加快软件的编制、调试工作，并很容易加入新的功能模块。根据气动机械手系统的功能要求，软件系统总体流程如图 6.3。

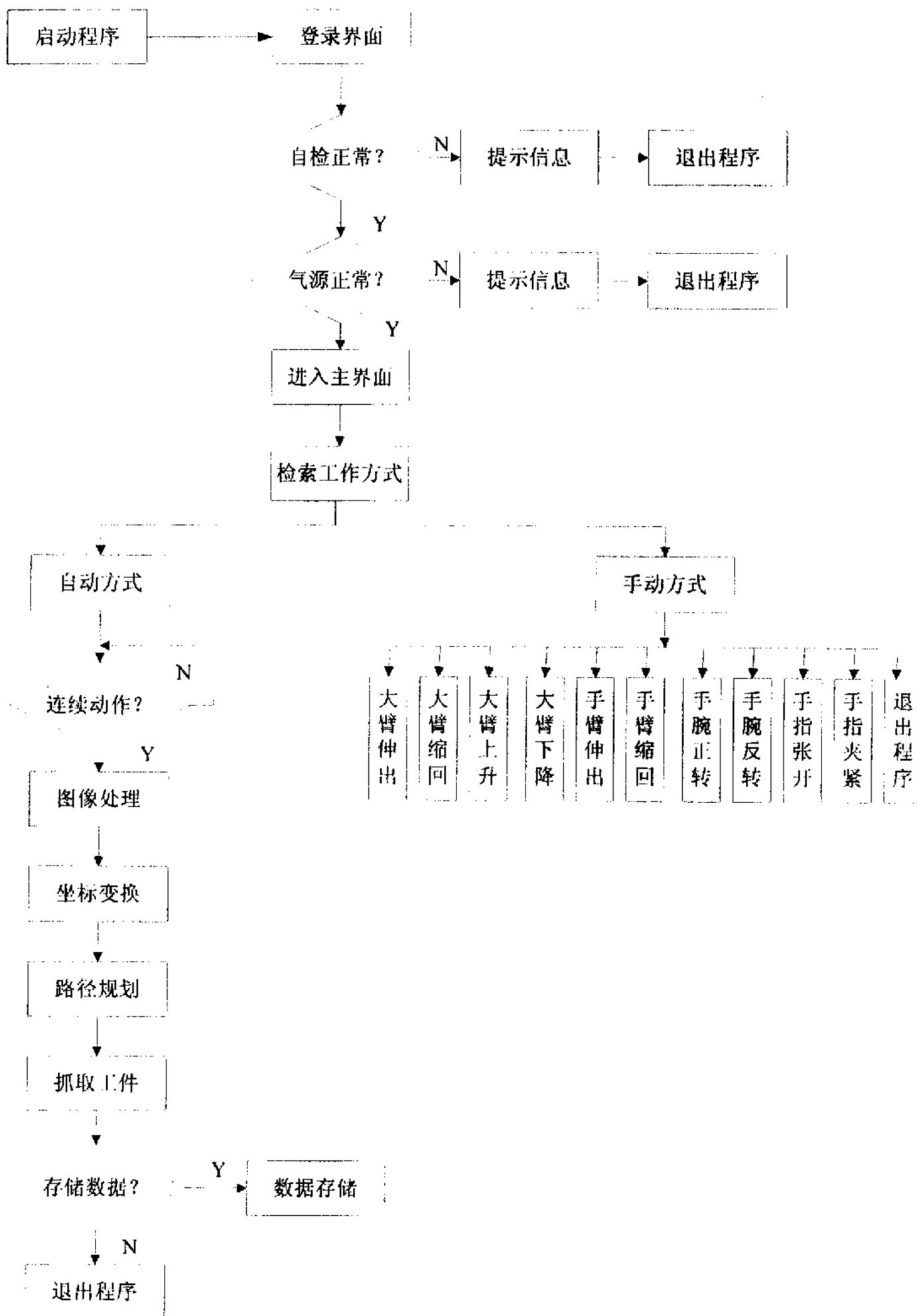


图 6.3 程序总流程图

在流程图的指导下，程序可以分为以下几个功能模块：

- ◆ 登录及自检模块。
- ◆ 图像采集和显示模块。
- ◆ 图像处理及形心跟踪模块。
- ◆ 机械手动作控制模块。

数据存储及调用模块。为了提高软件的可读性，本文采用模块化的方法来设计整个系统的软件部分。除了上述的软件模块之外，在每个模块的内部也尽量用模块设计，将一个比较独立的功能封装在一个模块内实现。

6.3 程序实现

6.3.1 登录模块

6.3.1.1 信息录入

为便于管理，在实验开始前要输入有关信息，包括实验操作人员姓名、学号、组别、实验开始时间等。为了增加软件的可靠性，在程序中加入了一些判断功能，若出现有一项未填或者数据类型不正确，则认为登录信息有误，需要重新输入。登录完成后登录信息填入数据库。

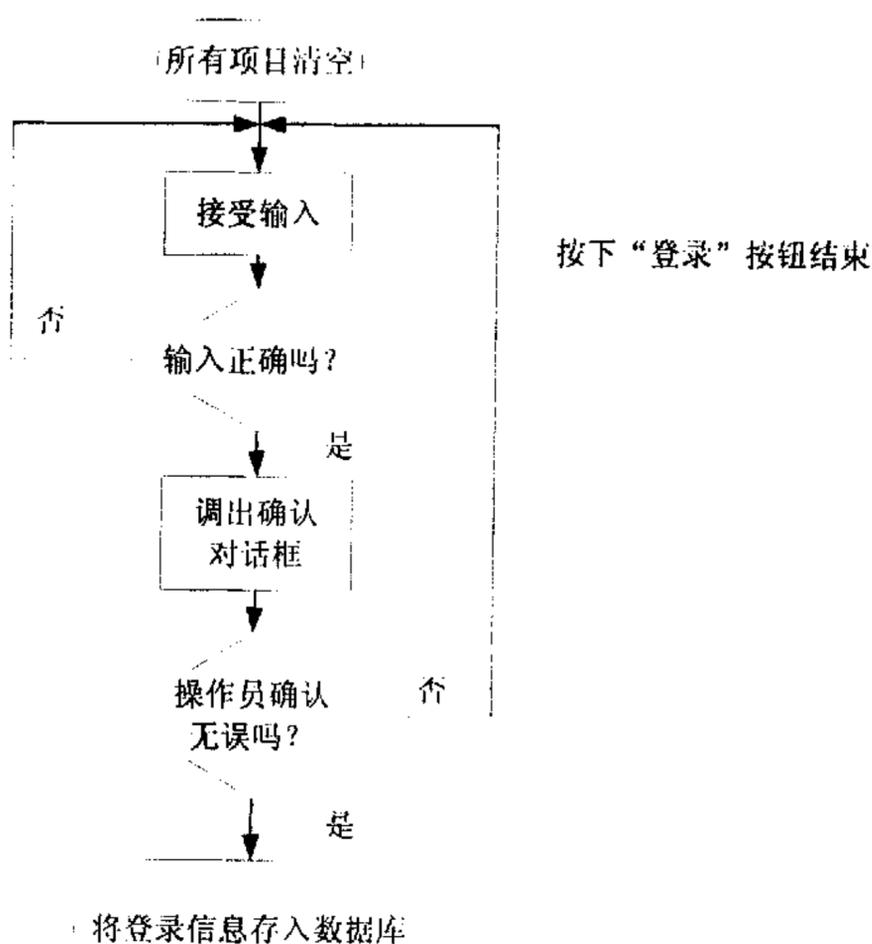


图 6.4 登录模块软件流程图

根据系统要求，在最终的实验报告中要包括整个实验的条件在内，如实验时间、操作人员姓名学号等，所以在系统运行的最开始设立了登录程序。操作人员可在产生的登录窗口中输入登录信息，程序内部调用 LabVIEW 的“获取当前时间”子 VI 来得到实验开始时间。录入的信息利用 LabVIEW SQL Toolkit 存入数据库。

其软件流程如图 6.4:

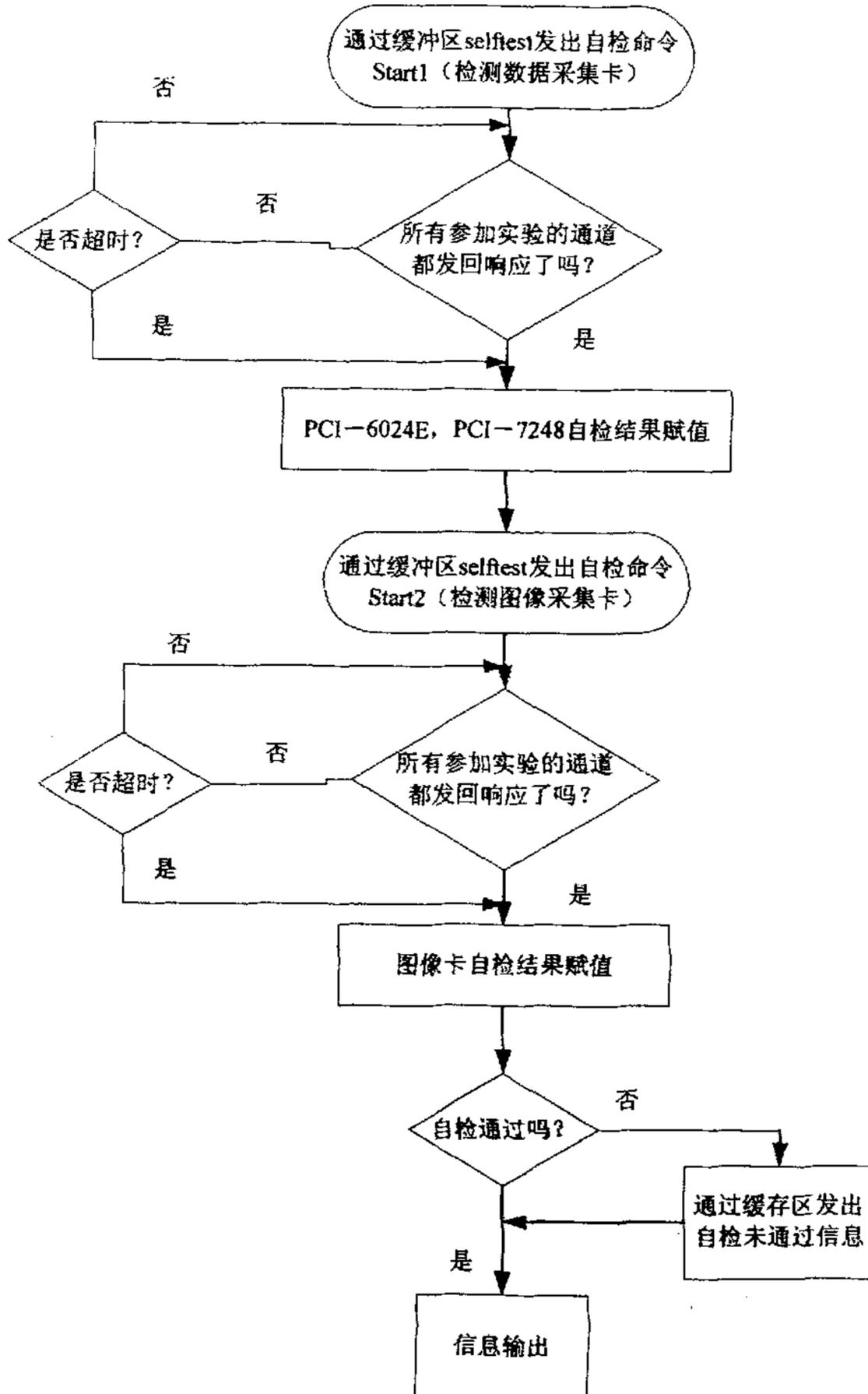


图 6.5 系统自检流程

6.3.1.2 系统自检

自检包括判断各个板卡是否插好，CCD 摄像头是否向图像采集卡传送了视

频图像，气源是否打开，并判断气压值是否在工作状态。

对于系统状态，物体图像的采集程序依赖于相应的采集卡，若采集卡不存在或未正确安装、图像采集卡未接收到视频图像等，程序会给出相应的出错信息，在本模块中对出错信息进行截取判断，可给出相应的自检结果信息。软件流程如图 6.5 所示。

6.3.2 图像采集及处理模块

系统图像采集及处理程序流程如图 6.6 所示。

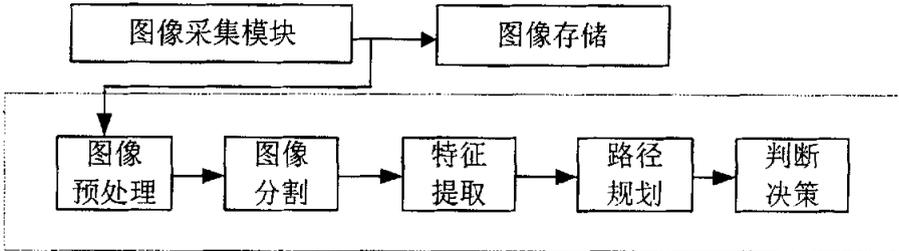


图 6.6 图像采集及处理流程

系统图像采集及处理程序流程利用 LabVIEW 及图像采集卡 PCI-1407 的驱动程序，设计了实时图像采集及存储模块。该程序主要利用了 NI-IMAQ 所提供的对图像采集卡的控制函数来进行图像采集的，并使用 LabVIEW 自带的图像子虚拟仪器把此图像存储为可供选择的多种文件格式：BMP、JPEG 和 PNG 等。

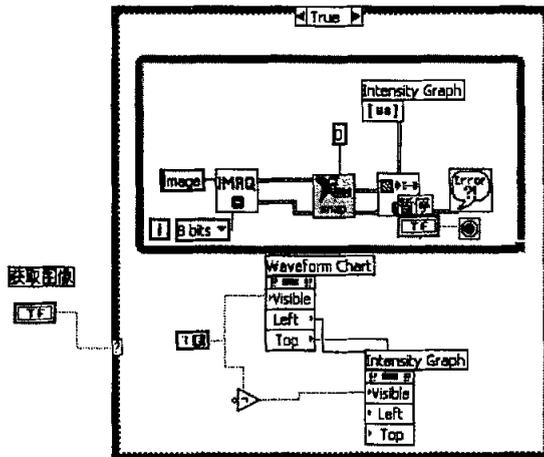


图 6.7 图像采集子 VI

针对目标物体图像，设计了图像处理模块，以获取目标物体的准确信息。系统的图像预处理包括图像增强和噪声过滤。图像分割是图像识别的关键步骤，本系统的图像分割主要用到了图像处理技术中的阈值分析技术。用自适应阈值进行图像二值化处理，把灰度图像转换为二值图像。具体的阈值大小则应根据不同的

情况选取不同的值。本系统还采用了面积形心法找出图像及物体的形心坐标。特征提取及判断决策模块则要针对目标物体的位置进行实时判别。

基于以上算法，LabVIEW 的库函数相应地提供了子 VI，如：图像采集子 VI（图 6.7）、平滑滤波子 VI、边缘检测子 VI、求取阈值子 VI、二值化子 VI、物体形心坐标子 VI 等等，用户只要用导线将必要的其他子 VI 连接起来，就可以完成对获得图像的预处理、理解，从而实现机器人的视觉。

6.3.3 控制模块

此模块中，由 PCI-7248 读得当前的坐标点位置，由图像采集及处理模块得出目标物体的形心位置及路径规划的轨迹，根据控制算法，由 PCI-6024E 发出指令，控制机械手的动作。其流程如图 6.8 所示：

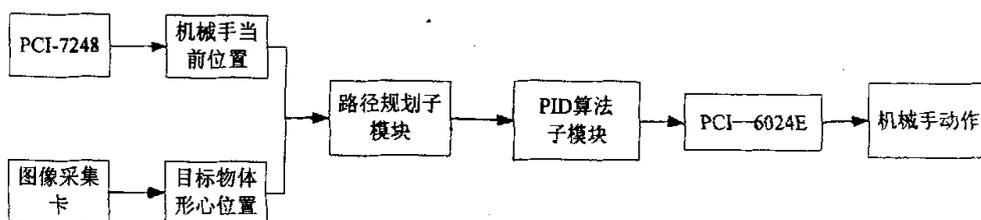


图 6.8 控制模块流程图

6.3.4 数据管理模块

LabVIEW 直接操作数据库的方法都显得过于繁琐，为了实现更强大的功能和做出合乎要求的报告，本系统中的数据管理软件采用 Excel 表格文件作为中介，通过动态数据交换调用数据库的宏操作实现数据库访问，从而避免了 LabVIEW 直接和数据库的数据打交道^[65]。



图 6.9 登录界面

为了使用户在同一个软件中就能对数据库进行备份、恢复、清空等本应在

SQL Server 中进行的操作，使用动态 SQL 语句调用了一些数据库管理的高级功能，用户无需数据库知识就能对本系统中形成的数据库进行操作。

这里的工作主要集中在数据存储与调出、检索窗口、报表的格式设计上，而代码实现比较简单，加之基于消息、事件驱动的程序流程图较为复杂，为节省篇幅，这里不给出流程图。图 6.9 和图 6.10 为程序用户界面。

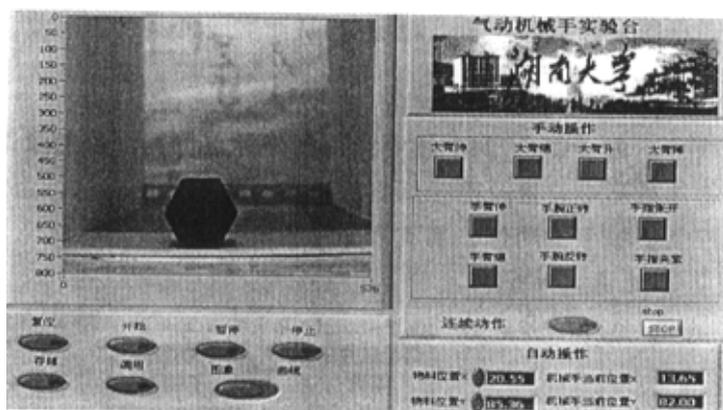


图 6.10 气动机械手总体界面

6.4 本章小结

本章对气动机械手视觉系统的程序设计作了详细的论述，由于采用了基于虚拟仪器的机器视觉技术，使得绝大部分的处理功能都由软件实现，因此本系统的软件设计质量的高低直接关系到整个系统性能甚至功能的优劣。

本系统功能众多，软件体系采用模块化设计，每一个功能模块都定义其输入输出接口，可以在其它场合方便调用。在论述系统中的功能模块时，首先说明功能要求，然后论述设计方案，并给出程序设计流程图。通过详细介绍视觉系统 LabVIEW 程序的设计，完成了一个初步的视觉软件，实现了目标物体的自动识别和路径跟踪，达到了设计目的。

总 结

本论文围绕机器人视觉伺服跟踪平面运动目标这个任务,以气动机械手为研究对象,以虚拟仪器为开发平台,研究手和眼的协调工作过程。从系统标定、视觉信息处理、跟踪等几个关键技术方面,对整个系统的实现进行了讨论。在前人工作的基础上,本文主要完成了以下几个部分:

由于基于虚拟仪器的机器视觉技术是近年才发展起来的,本文首先对这一技术的有关概念、原理、特点做了论述。在对机器视觉技术的概念与原理做了论述后,将其与人类视觉相比较,得到了机器视觉技术的应用范围,根据最近机器视觉技术的动态总结出了其发展的技术趋势,指出将虚拟仪器技术和机器视觉技术相结合得到的基于虚拟仪器的机器视觉系统,完全体现了机器视觉技术的发展趋势,成为机器视觉系统的重要分支。

对实验室原有的气动机械手进行了恢复和改造,并在机械手的末端执行器上安装一 CCD 摄像机,建立了一个“Eye-in-hand”构成的手眼系统。引入了虚拟仪器的机器视觉系统概念之后,对这类系统的主流开发平台 LabVIEW 和图像处理工具包 IMAQ Vision 在功能及特点等方面做了介绍。

本文第三章完成了摄像机标架到末端手架之间的坐标转换,即标定。给出了一种视觉平面标定方法,该方法实现简单,算法简捷。但其标定结果只能用于要求不太高的场合。事实上,如何从本质上提高标定精度是许多研究者始终关注的课题。一种方法是把上述所有的标定参数用一个标定模型统一起来,同时估计所有参数的值,但这使得问题的复杂程度大大加重了。即使能给出一个标定模型,工程上的可行性,是否有好的算法来求解,还有待于进一步的研究。

针对基于虚拟仪器的机器视觉系统的特点,主要使用软件进行图像的处理工作。本文重点研究了图像增强,滤波,二值化,图像分割等方面的理论并用于实验。最后,根据实验要求,进行特征提取,求出物体图像的面积,并利用面积算出形心坐标。

本文的第五章分析了视觉跟踪理论及实现方法,并对各种方法作以比较,比较其精度和计算量。根据本项目的特点选用了基于分割的跟踪算法,利用目标物体图像的形心坐标描述物体位置,设计了直线轨迹插补算法,通过逆运动学分析规划机械手的运动轨迹,转化为运动气缸的目标位移,实现机械手跟踪控制,对跟踪结果进行了仿真分析和试验验证。

最后,对气动机械手视觉系统的程序设计作了详细的论述,由于采用了基于虚拟仪器的机器视觉技术,使得绝大部分的处理功能都由软件实现,因此本系统

的软件设计质量的高低直接关系到整个系统性能甚至功能的优劣。系统功能众多，软件体系采用模块化设计，每一个功能模块都定义其输入输出接口，可以在其它场合方便调用。在论述系统中的功能模块时，首先说明功能要求，然后论述设计方案，并给出程序设计流程图。通过详细介绍视觉系统 LabVIEW 程序的设计，完成了一个初步的视觉软件，实现了目标物体的自动识别和路径跟踪，达到了设计目的。

论文的研究工作以“气动机械手的视觉系统”为主线，查阅国内外大量的参考资料，通过仔细分析系统研究的现状、可行性和系统的需求，最终根据系统性价比最高的原则，将虚拟仪器技术和机器视觉技术结合起来，完成了整个系统的设计。所完成的系统完全满足要求，不仅具有自动化程度高、可靠性和精确性好等特点，还节省了人力和财力。

另外，限于本人水平和时间，以及基础条件的不足，本文的研究工作也存在一些缺陷。随着相关技术的研究和进一步发展，以下方面的工作有待进一步研究和完善^[66]：

- 1) 图像特征的选择问题。视觉伺服的性能密切依赖于所用的图像特征，特征的选择不仅要考虑识别的指标，还要考虑控制指标。
- 2) 结合计算机视觉及图像处理的研究成果，建立机器人视觉系统的专用软件库。在视觉伺服中，需要进行图像采集、图像处理、特征抽取及由二维信息重构三维信息等，要处理的数据量较大，算法复杂多样。
- 3) 加强系统的动态性能研究。目前的研究主要是根据图像信息确定机器人的期望运动轨迹，没有对整个视觉伺服系统的动态性能进行过多的研究。
- 4) 研究智能控制理论和技术在机器人中的应用。虽然神经网络在机器人视觉伺服中已得到应用，但多数都是针对具体物体的具体特征，或只进行了仿真实验，还有待于进一步的研究。随着智能理论和技术的发展，很多机器人难题将得到解决。

参考文献

- [1] 吴广玉, 姜复兴. 机器人工程导论. 哈尔滨: 哈尔滨工业大学出版社. 1988, 1—9
- [2] 李建藩. 气压传动系统动力学. 广东: 华南理工大学出版社. 1991, 1—3, 53—59
- [3] 周洪. 气动比例控制技术及其应用. 液压与气动, 1999. (3): 1—3
- [4] James Kimery, "Choosing the right test platform", EE-EVALUATION ENGINEERING, 1996, 9—10
- [5] 陈熙霖. 视觉计算——人类感知能力的延伸. 测控技术, 2000, 19(5): 19—20
- [6] 王雄耀. 近代气动机器人的发展及应用, 液压与气动密封, 1999(5): 17—19
- [7] 周洪, 电—气比例/伺服系统及其控制策略研究: [浙江大学博士论文]. 杭州: 浙江大学图书馆, 1988, 10—20
- [8] 周洪, 气动伺服定位技术及其应用. 液压与气动, 1999(1): 18—21
- [9] 阮秋琦. 数字图像处理学. 北京: 电子工业出版社. 2001, 13—18
- [10] 章毓晋. 图像理解与计算机视觉. 北京: 清华大学出版社. 2000, 21—29
- [11] 胡守仁, 神经网络导论. 长沙: 国防科技大学出版社, 1993, 42—48
- [12] 楼顺天, 施阳. 基于 MATLAB 的系统分析与设计——神经网络. 西安: 西安电子科技大学出版社, 1998, 114—126
- [13] Eshed Robotec. ACL——Reference Guide and ATS. 11—17
- [14] Maybank S J, Faugeras O. A theory of self-calibration of a moving camera. Int. J. Computer Vision, 1992, 8(2):123—151
- [15] Faig W. Calibration of close-range photogrammetry system: Mathematical formulation. Photogrammetric Engineering and Remote Sensing, 1975, 41(12):1479-1486
- [16] Faugeras O. D, Luong Q T, Maybank S J. Camera self-calibration: Theory and experiments. In: Proceedings of the European Conference on Computer Vision, Santa-Margherita, Italy, 1992. 321-334
- [17] Maybank S J, Faugeras O D. A theory of self-calibration of a moving camera. International Journal of Computer Vision, 1992, 8(2):123-151
- [18] Tsai RY. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. IEEE

- Journal of Robotics and Automation, 1987, 3(4):323-344
- [19] Zhang Z . A flexible new technique for camera calibration . Microsoft Corporation : Technical Report MSR-TR-98-71, 1998
- [20] Triggs B . Autocalibration from planar scenes . In : Proceedings of the European Conference on Computer Vision, Freiburg, Germany, 1998. 89-105
- [21] Sturm P , Maybank S . On plane-based camera calibration : A general algorithm , singularities , applications . In : Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition , Fort Collins, Colorado, USA, 1999, 432—437
- [22] Ma S . D . Conics-based stereo , motion estimation , and pose determination . International Journal of Computer Vision, 1993, 10(1):7—25
- [23] Quan L . Conic reconstruction and correspondence from two views . IEEE Transaction on Pattern Analysis and Machine Intelligence, 1996, 18(2):151—160
- [24] Mundy J , Zisserman A . Geometric Invariance in Computer Vision . Cambridge , MA : MIT Press, 1992, 130—135
- [25] Golub G , VanLoan C . Matrix Computations . 3rd edition . The John Hopkins University Press, 1996
- [26] 陶闯. 摄像系统标定中尺度因子的鲁棒测定. 模式识别与人工智能, 1995, 8(3):248—254.
- [27] 卢宏等. 摄像机校准的精度分析. 见: 第三届全国机器人学术讨论(下). 1991 (1), 378—382
- [28] 郑南宁. 计算机视觉与模式识别. 北京: 国防工业出版社, 1998, 10—23
- [29] 朱志刚等. 数字图像处理. 北京: 电子工业出版社, 1998, 375—406
- [30] 王晖, 张基宏. 图像边界检测的区域对比度模糊增强算法. 电子学报, 2000, 28 (1), 45—47
- [31] 张永平, 郑南宁, 赵荣椿. 基于视觉模型的图像边缘检测算法. 电子学报, 2000, 28 (1), 101—103
- [32] 华长发, 范建平等. 基于二维熵阈值的图像分割及其快速算法. 模式识别与人工智能. 2000, 13 (1), 42—45
- [33] 王思贤, 曾发龙. 平滑图像噪声的差值滤波法. 电子科学学刊. 2000, 22 (3), 412—415
- [34] T. I. Hsu, J. L. Kuo and R. Wilson, A multi-resolution texture gradient method for unsupervised segmentation, Pattern Recognition 2000. 33(11):1819—1834
- [35] N. P. Pal and S. K. Pal, A review on image segmentation techniques , Pattern

- recognition, 1992, 26(9):1277-1294
- [36]刘宏兵, 杨万海, 马剑虹. 图像小波边缘提取中阈值选取的一种自适应方法. 西安电子科技大学学报(自然科学版). 2000, 27(3): 294-296
- [37]孙龙祥. 程义民, 王以孝. 深度图像分析. 北京: 电子工业出版社. 1996
- [38]贾云得. 机器视觉. 北京: 科学出版社, 2000, 69-89
- [39]A . Hammoude . An empirial parameter selection method for border identification algorithms Computers in Cardiology . 1998:445-448
- [40] 吴晓波. 图像边缘特征分析. 光学精密分析, 1997, 7(3): 59-63
- [41] 王博, 潘泉, 张洪才, 戴冠中. 图像平滑与边缘检测的模糊向量描述. 小型微型计算机系统, 1999, 20(3): 218-221
- [42]张凯丽, 刘辉边. 边缘检测技术的发展研究. 昆明理工大学学报. 2000, 25(5): 36-39
- [43] 贺忠海, 王宝光, 廖怡白, 张晓芳. 图像处理中的边缘跟踪法研究. 仪器仪表学报. 2001, 22(3s), 261-262
- [44]周德龙, 潘泉, 张洪才. 自动图像边缘检测方法的研究. 计算机工程与科学. 2001, 23(4), 11-14
- [45]张天序. 一种新的边缘检测计算模型和算法. 自动化学报. 1994, 20(4), 436-444
- [46]戚飞虎, 模式识别与图像处理. 上海: 上海交通大学出版社, 1989
- [47]傅京孙, 模式识别及其应用. 北京: 科学出版社, 1983
- [48]Joon H . Hana , tae Y . Kimb . Ambiguity distance : an edge evaluation measure using fuzziness of edges . Fuzzy Sets and Systems . 2002, 126:311-324
- [49]Hutchinson, S., Hager , G. D. and Corke , P. I:A Tutorial on Visual Servo Control. IEEE Trans. on Robotics and Automation , 12(5):651-670
- [50]Hager, G., D., A Brief Reading Guide to Dynamic Vision, The Johns Hopkins University <http://www.cs.jhu.edu/~hager>, 2001, 1-8
- [51]Ferrier , N . J ., Active Control of Resolution for Stable Visual Tracking , In Workshop on Modelling and Planning for Intelligent Systems, Christensen, Vunke&Bolles, Editors, LLNCS, Spinger-Verlag, 1999, 37-46
- [52]Nelson , B . , Papanikolopoulos , N. , and Khosla , P. , Visual Servoing for Robotic Assembly . Visual Servoing-Real-Time Control of Robot Manipulators Based on Visual Sensory Feedback , Ed . K . Hashimoto , World Publishing Co . Pte . Ltd . River Edge , New Jersey , 1993, 1: 39-64

- [53]Denzler, J. and Niemann, H. Active rays:Polar-transformed active contours for real-time contour tracking. Journal on RealTime Imaging, 1997, 45-47
- [54]shen, D., Wong W. and Horace, H. Affine Invariant Images Retrieval by Correspondence Matching of Shapes, Image And Vision Computing, 1999, 17(7): 489-499
- [55]Leonard J J, Durrant-Whyte H F. Directed sonar sensing for mobile robot navigation. Kluwer Academic Publishers, 1992, 29-38
- [56]刘万春, 贾云得等. 基于肤色的人脸实时跟踪算法. 北京理工大学学报, 2000, 8, 20(4), 461-465
- [57]Rasmussen, C., Toyama, K., and Hager, G. D., Tracking Objects By Color Alone. Technical Report TR1114, Dept. of Computer Science Yale University, June1996
- [58]S. Sarkar, T Sanocki, K. Bowyer, M. Heath Comparison of edge detectors: a methodology and initial study. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Proceedings CVPR '96, 1996:143-148
- [59]M. C. Shin, D. Goldgof, K. W. Bowyer. Comparison of edge detectors using an object recognition task. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999, 1:365
- [60]National Instruments, 《LabVIEW User Manual》, 1999. 1-223
- [61]National Instruments, 《LabVIEW Measurement Manual》, 1999. 5-20
- [62]National Instruments, 《LabVIEW Developments Manual》, 1999. 40-49
- [63]National Instruments, 《Using External Code in LabVIEW》, 1999. 50-55
- [64]National Instruments, 《IMAQ Vision for G Reference Manual》, 1999. 2-20
- [65]王宁, 黄文梅. LabVIEW 在智能切割机控制软件开发中的应用. 湖南大学学报(自然科学版), 2003, (3) 53-56
- [66]孙增圻. 智能控制理论与技术[M]. 北京:清华大学出版社,1996:1-10

致谢

首先，我要感谢我的导师黄文梅教授和钟志华教授。

黄文梅教授和钟志华教授以其渊博的知识、严谨的治学态度、敏锐的洞察力、诲人不倦的教育风范和高度的工作热情，给我留下深刻的印象，将使我终身受益。

钟志华教授使我有机会参加“十五”攻关项目的实际开发与研究，有幸接触到许多机电控制领域发展的前沿。在此过程中，得到了黄文梅教授的言传身教，其严谨的治学态度、扎实的理论功底、高尚的思想修养，如涓涓细流，似绵绵春雨，润物无声而让我受益匪浅。

在读研究生的3年里，得到了唐亚利老师的关心和帮助，毕业设计中，液压实验室的邹宪军老师为我们提供了方便，向他们表示感谢。在和王秋菊长期的相处中，始终是在坦诚和互相促进中相得益彰的，向她表示由衷感谢！王局柏和左飞雁在项目完成的过程中，给过我无私的帮助，一并表示感谢！

感谢父母及家人在我成长的过程中，给我的关怀和帮助！

感谢男友徐祖龙在我学习和生活上的关心与呵护！

特别感谢在百忙之中抽出时间，对本论文进行评阅与指点的专家与学者！

最后，谨以此文，献给所有关心我成长的老师、亲人和朋友们！

附录 A 攻读学位期间发表的学术论文

1. 王宁, 黄文梅. LabVIEW 在智能切割机控制软件开发中的应用. 湖南大学学报 (自然科学版) 2003 年第三期;
2. 徐祖龙, 傅勤毅, 王宁. 基于组件技术的励调器智能检测软件设计. 微型电脑应用, 2003 年第 8 期;
3. 牛刚, 傅勤毅, 王宁等. 货车轮对滚动轴承压装机微机监测系统设计. 长沙电力学院学报. 2004 年 5 月