

分类号: \_\_\_\_\_  
UDC: \_\_\_\_\_

密级: \_\_\_\_\_  
编号: \_\_\_\_\_

# 基于双目立体视觉的距离测量

## RANGE MEASUREMENT BASED ON BINOCULAR STEREO VISION

学位授予单位用代码: 长春理工大学 (10186)

学科专业名称用代码: 光学工程 (080300)

研究方向: 光电测控技术及仪器

申请学位级别: 硕 士

指导教师: 杨进华

答辩委员会主席: 田地

研 究 生: 崔岩

论文评阅人: 李春槐

论文起止时间: 2005.8-2006.12

张秉仁

## 摘 要

立体视觉技术是机器人技术研究中最为活跃的一个分支，是智能机器人的重要标志。双目立体视觉是通过同一目标的两幅图像提取、识别、匹配和解释，重建三维环境信息的过程。主要包括：视频捕获、摄像机定标、图像预处理和特征提取、立体匹配以及三维重建。通过对双目立体视觉技术的进一步研究，对多目视觉具有重要的启发，为解决智能移动机器人、工业装配机器人、家用机器人、公共服务机器人的视觉问题奠定良好的技术基础。

计算机立体视觉的主要研究内容是由多幅二维的平面图像恢复出被摄物体的三维坐标，而其中基于两幅图像的双目视觉技术则是一个研究热点。双目立体视觉的基本原理是模仿人眼与人类视觉的立体感知过程，从两个视点观察同一景物，以获取不同视角下的感知图像，通过三角测量原理计算图像像素间的位置偏差，以获取景物的三维信息。

一个完整的双目视觉系统通常可分为图像获取、摄像机标定、特征提取、立体匹配、目标定位和深度图生成等六大部分。本论文主要是图像特征提取、立体匹配这两方面的内容展开研究。首先根据选择的立体视觉成像系统模型阐述了立体视觉测距的基本原理。其次，对图像特征提取等预处理方法进行详尽地分析，通过实例对算法进行实验。同时，对图像匹配问题涉及的内容包括图像匹配方式、质量控制策略等进行了总结。最后，以特征提取与图像匹配研究为基础，针对不同类型的立体图像提出了基于图像边缘特征和图像兴趣点的匹配方法。

**关键词：**立体视觉 摄像机定标 特征提取 立体匹配 深度图

## ABSTRACT

Stereovision, an active branch of robotic technology, is one important attribute of intelligent robot. Binocular stereovision can be defined as the process of tri-dimensional reconstruction through the abstraction, identification, matching and interpretation of twain images gathered from identical objective, mostly include: video frequency trapping, camera calibration, image preprocessing, feature extraction, stereo matching and tri-dimensional reconstruction as well. It is through further research of binocular stereovision that we can bring important enlightenment to multi-eye vision system and present supporting technology for visual problem of intelligent locomotive robot, industry assembling robot, domestic robot and community servicing robot.

Stereo vision is to reconstruct the three-dimensional coordinate of an object from some of its two-dimensional images. Generally, one would like to reconstruct the object in three-dimensional space with only two different view points. The fundamental principle of computer vision with dual-cameras have a strong resemblance to the visual perception procedure of mankind. We can get different images with two different view points and then calculate the position deviation of images pixels based on triangular measure. Finally we can reconstruct the three-dimensional coordinate of the object.

The integrated binocular stereo vision system generally includes image acquisition, camera calibration, features extraction, stereo matching, object location and depth map construction, etc. The features extraction and the stereo matching is mainly discussed in this paper. First, we elaborate the fundamental principle of measure distance in stereo vision based on camera modeling. Second, the paper gives the elaborate analysis on preprocessing methods such as features extraction, and verifies the algorithms by example. Furthermore, it summarizes the content of stereo matching such as match method, the strategy of match quality control. Finally, we submit the match algorithm based on image edge features and match algorithm based on the image interesting dot by features extraction and image matching.

**Key words: Stereovision Camera calibration Feature Extraction Stereo matching Depth map**

## 长春理工大学硕士学位论文原创性声明

本人郑重声明：所呈交的硕士学位论文，《基于双目立体视觉的距离测量》是本人在指导教师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

作者签名： 崔岩 2007年3月28日

## 长春理工大学学位论文授权使用授权书

本学位论文作者及指导教师完全了解“长春理工大学硕士、博士学位论文授权使用规定”，同意长春理工大学保留并向国家有关部门或机构送交学位论文的复印件和电子版，允许论文被查阅和借阅。本人授权长春理工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，也可采用影印、缩印或扫描等复制手段保存和汇编学位论文。

作者签名： 崔岩 2007年3月28日

指导导师签名： 杨建华 2007年3月28日

# 第一章 绪论

## § 1.1 引言

视觉是人类观察世界、认知世界的重要功能手段，人类感知外部世界主要通过视觉、触觉、听觉和嗅觉等感觉器官，其中 80% 的信息是由视觉获取的，机器视觉是人类利用计算机实现人的视觉功能——对客观世界三维场景的感知、识别和理解，机器视觉是一个相当新而且发展迅速的研究领域，机器视觉技术正广泛地应用于各个方面，许多人类无法感知的场合，如危险场景感知等，机器视觉更突显其优越性。机器视觉属于影像的重构技术，具有非接触性测量、实施简易的特点，具有广阔的应用前景，是有待于进一步开发的领域。

计算机视觉被认为是计算机科学和人工智能的一个分支，计算机视觉就是用机器代替人的眼睛和大脑，对客观世界进行感知和解释的技术。该系统的首要目标是使计算机具有通过一幅或者多幅二维图像认知周围的三维环境信息的能力。这种能力不仅可以使计算机能够感知三维环境中物体的几何信息，并且能够对它们进行描述、存储并使用图像来创建和恢复现实世界模型，从而达到认知现实世界的目的。因此，自从计算机视觉这门科学出现以来就成为计算机科学的重要研究领域之一并且发展十分迅速。

现实世界中的物体都是三维的，人眼所获得的景物图像却是二维的，但是人类的视觉系统能够很容易地从二维图像中感知三维世界，获得三维世界的信息。机器人视觉的研究目标就是使机器人具有通过一幅或多幅图像认知周围环境信息的能力。这种能力不仅使机器人能感知环境中物体的几何信息，如其形状、位置、姿态、运动等，而且能对它们进行描述、存储、识别与理解。机器人视觉系统的基本结构如图 1.1 所示。



图 1.1 机器人视觉系统基本结构

## § 1.2 国内外研究现状、发展趋势

### § 1.2.1 国内外研究现状

立体视觉的开创性工作是从 60 年代中期开始的。70 年代末, Marr 创立的视觉计算理论对立体视觉的发展产生了巨大影响, 现已形成了从图像获取到最终的景物可视表面重建的完整体系, 在整个计算机视觉中已占有越来越重要的地位。

现在, 国际上以计算机视觉为主题, 或主要以计算机视觉为主题的国际会议有国际计算机视觉会议 (ICCV), 国际模式识别会议 (ICPR), 国际计算机视觉与模式识别会议 (CVPR), 欧洲计算机视觉会议 (ECCV) 与亚洲计算机视觉会议 (ACCV)。以计算机视觉为主要内容之一的国际刊物也很多, 有 International Journal of Computer Vision, IEEE Trans. On PAMI, CVGIP (Computer Vision, Graphics and Image Processing), Visual Image Computing, IJRAI (International Journal of Pattern Recognition and Artificial Intelligence), Pattern Recognition 等, 每年的研究论文不下数千篇, 发表的论文也数以百计。在 Marr 理论的框架下, 取得了一大批科研成果, 也包括数据结构、算法层次上的各种算法和在硬件实施方面的一些实验系统。但这些理论、方法、算法在实际应用中并没有产生能完全取代人的视觉系统。

双目体视目前主要应用于四个领域: 机器人导航、微操作系统的参数检测、三维测量和虚拟现实。在国外, 日本大阪大学自适应机械学院研制了一种自适应双目视觉伺服系统, 实现了对运动方式未知的目标的自适应跟踪。日本东京大学将实时双目立体视觉和机器人整体姿态信息集成, 开发了仿真机器人动态行走导航系统。日本冈山大学使用立体显微镜、两个 CCD 摄像头、微操作器等研制了使用立体显微镜控制微操作器的视觉反馈系统, 用于对细胞进行操作, 对钟子进行基因注射和微装配等。麻省理工学院计算机系提出了一种新的用于智能交通工具的传感器融合方式, 由雷达系统提供目标深度的大致范围, 利用双目立体视觉提供粗略的目标深度信息, 结合改进的图像分割算法, 能够在高速环境下对视频图像中的目标位置进行分割。

在国内, 浙江大学机械系完全利用透视成像原理, 采用双目体视方法实现了对多自由度机械装置的动态、精确位姿检测。东南大学电子工程系基于双目立体视觉, 提出了一种灰度相关多峰值视差绝对值极小化立体匹配新方法, 可对三维不规则物体 (偏转线圈) 的三维空间坐标进行非接触精密测量。哈工大采用异构双目活动视觉系统实现了全自主足球机器人导航。火星 863 计划课题 “人体三维尺寸的非接触测量”, 采

用“双视点投影光栅三维测量”原理，由双摄像机获取图像对，通过计算机进行图像数据处理，不仅可以获取服装设计所需的特征尺寸，还可根据需要获取人体图像上任意一点的三维坐标。

### § 1.2.2 发展趋势

20 世纪 70 年代中期以 Marr、Barrow 和 Tenenbaum 等人为代表的一些研究者提出了一整套视觉计算的理论来描述视觉过程，其核心是从图像恢复物体的三维形状。在视觉研究的理论下，以 Marr 的理论影响最为深远，并成为这一领域的主导思想<sup>[1]</sup>。马尔教授认为：视觉可分为三个阶段。第一阶段是早期视觉(early vision)，其目的是拍取观察者周围景物表面的物理特性，如距离、表面方向、材料特性(反射、颜色、纹理)等，具体来说包括边缘检测、双目立体匹配、由阴影确定形状、由纹理确定形状等。第二阶段是二维半简图(2.5D sketch)或本征图像(Intrinsic Image)。它是在以观察者为中心的坐标系中描述表面的各种特性。根据这些描述，可以重建物体边界，按表面和体积分割景物。但在以观察者为中心的坐标系中只能得到可见表面的描述，得不到遮挡表面的描述，故称二维半简图。第三阶段是三维模型—视觉信息处理的最后一个层次，是用二维半简图中得到的表面信息建立适用于视觉识别的三维形状描述。这个描述应该与观察者的视角无关，也就是在以物体为中心的坐标系中，以各种符号关系和几何结构描述物体的三维结构和空间关系。

经过近三十年的发展，在马尔的理论框架下，机器人立体视觉取得了一大批成果。但随着机器人立体视觉研究的不断深入，人们发现了马尔视觉理论的局限性，即从景物图像或系列图像求出景物精确的三维几何场景并定量的确定景物中物体的性质时遇到了困难<sup>[2]</sup>。在 Marr 的理论中，输入是被动的，给什么图像系统就处理什么图像，整个处理过程基本上是“自底向上”的，没有反馈。处理的目的是不变的，总是要求出场景中物体的形状和它们的空间位置和关系。其理论强调表示的重要性以及从不同层次上去研究信息处理问题，在计算理论和算法实现上又特别强调计算理论的重要性。这些理论多数是建立在对人类感知三维信息分析的基础上。尽管这些方法在数学上是可行的，但由于各种干扰的存在以及逆成像的问题，使得问题本身是病态的。

进入 80 年代中后期，随着移动式机器人立体视觉研究的发展，大量运用空间几何的方法以及物理知识来研究双目立体视觉，其主要是完成对道路和障碍的识别处理。这一时期引入主动视觉的研究方法，并采

用了距离传感器以及视觉融合技术等,由于这种研究方法可直接取得深度图或通过移动获取深度图,因而使很多病态问题变成良态的。此外在视觉的研究中重视了对定性视觉、有目的的视觉等的研究。上述这些理论和方法的研究有力地促进了相关应用的发展,这些应用领域包括对照片特别是航空照片和卫星照片的解释、精确制导、移动机器人立体视觉导航、医学辅助诊断、工业机器人的手眼系统、地图绘制、物体三维形状分析与识别以及智能人机接口等<sup>[3][4]</sup>。在立体视觉的应用方面不乏成功的例子。如 1996 年日本成功研制出利用双目视觉来指导机械手动作的樱桃番茄收获机器人等<sup>[5]</sup>。

立体视觉经过二三十年的研究,已经有了很大的发展。从当前的发展来看,计算机视觉的发展动向可以归纳如下:

(1) 以全面的观点将立体视觉系统的各个模块联系起来,充分挖掘内在信息。

(2) 由传统的两视点立体视觉向多视点立体视觉并从完全静态的立体视觉向动态立体视觉发展,通过增加信息输入降低视觉计算的难度。

(3) 立体视觉系统逐渐向智能化方向发展,研究基于知识的、模型的和规则的立体视觉方法。

(4) 算法从以前的串行处理向并行处理发展,采用并行流水线机制和专用的信号处理器件,增强立体视觉系统的实用性。

(5) 由专用的视觉系统向通用的系统发展。强调场景与任务的约束,针对不同的应用目的,建立有目的的和面向任务的立体视觉系统。

随着计算机科学的飞速发展,VLSI 技术的提高,并行处理机的深入研究,立体视觉研究的前景逐渐显现出来。可以预见,计算机视觉、生物物理等多学科专家通力合作,必将在立体视觉的研究中取得丰硕成果。

与人类自身的视觉器官相比,目前的立体视觉系统还很不健全,还处在一个十分不成熟的阶段,计算机立体视觉系统的研究面临着一系列技术难点<sup>[6]</sup>:

(1) 立体视觉系统在分析和理解图像时都需要立体匹配,而这一技术还很不完善。在具有灰度失真、几何失真(透视、旋转、缩放等失真)和噪声的情况下匹配图像是十分困难的。目前还没有一个可靠的通用的匹配方法。尤其在失真和噪声比较严重的情况下,匹配更为困难。而现实世界中物体间的相互遮挡,物体本身各部分间的遮挡更增加了图像匹配的难度。



(2)人类视觉系统具有惊人的分析理解力,人是如何精确获取视觉知识的,至今还未研究清楚。由于视觉信息具有很高的冗余度,选取哪些特征可以解决立体视觉问题,目前尚无一种通用的方法以尽量减少视觉信息的冗余度,提高算法效率。

(3)立体视觉系统所需的计算量是非常庞大的,对于一幅标准的航空摄影照片(3000×3000),只用一个 $3\times 3$ 的算子对其进行一次卷积运算就相当于 $10^8$ 次乘法运算。然而实用的系统对时间的要求是有一定限制的。在现阶段集成电路和特殊功能部件的发展还不能满足实时性的要求,这为立体视觉系统完全变为实用的通用系统设置了一大障碍。

### § 1.3 双目立体视觉系统概述

生物功能的模拟是现代机器人视觉的最大挑战<sup>[7]</sup>。从二维灰度图像中获得三维深度信息、是机器人视觉中发展最迅速的领域。基于深度识别目的具有代表性的方法有:立体对计算深度、运动中计算深度、光流中恢复深度等。立体对深度计算就是利用双目立体视觉来完成深度识别的。双目立体视觉(也称双目视觉)是用两台性能相同、位置固定的 CCD 摄像机,获取同一景物的两幅图像,计算空间点在两幅图像中的“视差(disparity)”,以此确定场景的深度信息,进而构建场景的三维结构。一套完整的双目立体视觉算法可分为六个主要部分:图像获取、摄像机定标、特征提取、立体匹配、深度计算及深度信息内插<sup>[8]</sup>。

在一个典型的双目立体视觉系统中,两摄像机沿基线放置,其光轴相互平行,空间点的像分别在左右两个像平面上的投影位置的间距称为视差。由视差计算深度是非常容易的,但视差本身的计算是立体视觉中最困难的部分,它涉及到模型分析、摄像机定标、图像预处理、特征选取及特征匹配。

在机器人双目立体视觉系统的硬件结构中,目前通常采用两个 CCD 摄像机作为视频信号的采集设备,通过专用图像采集卡与计算机连接,把摄像机采集到的模拟信号经过采样、滤波、量化,最终提供给计算机图像数据。而当置于移动机器人上的摄像机光心平行放于一条直线上时,构成平行双目立体视觉系统,这是最简单最典型的一种情况,通过对它各个部分的几何关系的分析,我们可以获取目标场景的深度信息。

机器人双目立体视觉系统的软件结构包括计算机操作系统及其应用软件、视觉处理算法、机器人控制软件。其中视觉处理算法的范围十分广阔,根据应用目的的不同,可包括图像输入处理、摄像机定标算法、

图像预处理(滤波、边缘提取等)、特征选取、图像匹配、深度识别、深度图生成等。

目前,在机器人立体视觉中摄像机定标、立体匹配是最为活跃的研究分支。这两部分关系到移动机器人定位的准确性和目标跟踪的精度。另外,深度图生成技术的快速发展,为三维场景的生成及深度可视化提供了技术支持。

## § 1.4 数字图像处理及关键技术

图像处理是一个发展比较成熟的领域,图像处理通常是把一幅图像变换成另一幅图像,也就是说,图像处理系统的输入是图像,输出仍然是图像,信息恢复任务则留给人来完成。图像处理包括图像增强、图像压缩和模糊校正与非聚焦图像等内容。机器视觉系统把图像作为输入,产生的输出为另一种形式。比如图像中物体轮廓的表示。因此机器视觉的重点是在人的最小干预下,由计算机自动恢复和识别场景信息。图像处理算法在机器视觉早期阶段起着很大的作用,它们通常被用来增强图像特定信息并抑制噪声。

人工智能(artificial intelligent, AI)涉及到智能系统的设计和智能计算的研究。在经过图像处理和图像特征提取过程后,接下来要用人工智能方法对场景特征进行表示,并分析和理解场景。人工智能有三个过程:感知、认知和行动。感知是把反映现实世界的信息转换成信号,并表示成符号,人只是对符号进行各种操作,行动则把符号转换成影响周围环境的信号。人工智能的许多技术在机器视觉的各个方面起着重要的作用。事实上机器视觉常被视为人工智能的一个分支。

人工神经网络(artificial neural networks,ANNs)是一种信息处理系统,它是由大量简单的处理单元(称为神经元)通过具有强度的连接相互联系起来,实现并行分布式处理。人工神经网络的最大特点是可以通过改变连接强度来调整系统,使之适应复杂的环境,实现类似人的学习、归纳和分类等功能。人工神经网络已经在许多工程技术领域得到了广泛的应用。神经网络作为一种方法和机制将用于解决机器视觉中的许多问题。

现实世界是三维的空间,经摄像机成像后,成为丢失了距离信息的二维图象。然而,在计算机技术高速发展的今天,许多智能系统需要感知周围环境的三维信息。至今,研究人员提出多种从二维图象中恢复三维信息的方法<sup>[9]</sup>。以Marr<sup>[10]</sup>的计算视觉理论为基础的双目体视方法就是其中的一种重要的方法。双目立体视觉系统有多个部分组成。立体匹配是最为关键的技术,匹配结果的优劣直接影响系统的最后输出。

### § 1.4.1 双目体视的几何关系<sup>[11]</sup>

双目体视系统的成像几何关系如图1所示。系统采用双相机平行光轴成像，图

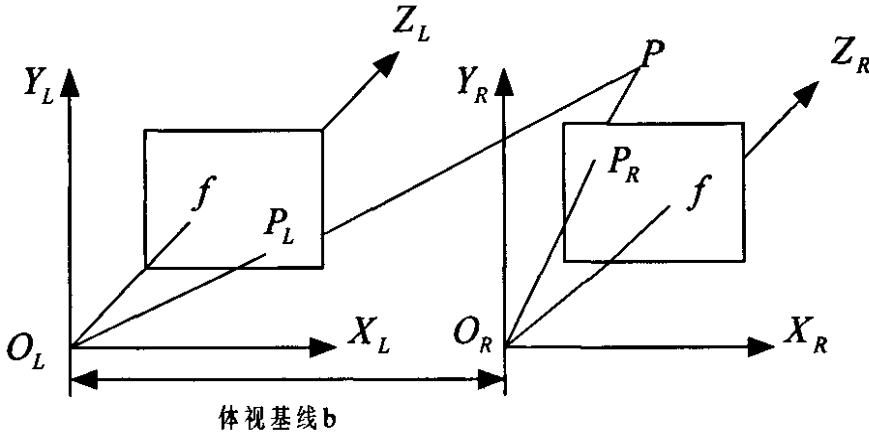


图1.2 双目体视几何关系图

中， $I_L$ 、 $I_R$ 分别表示左右象平面， $f$ 代表焦距， $b$ 是左右图象坐标系的原点 $O_L$ 、 $O_R$ 间的距离，称为体视基线。设世界坐标系 $W(X, Y, Z)$ 的3个坐标轴分别与左图象所在坐标系 $C_L(X_L, Y_L, Z_L)$ 的3个坐标轴重合。 $P_L(x_l, y_l, z_l)$ 、 $P_R(x_r, y_r, z_r)$ 分别表示三维空间中的一点 $P(x, y, z)$ 在左右象平面上的投影象点，投影线 $PO_L$ 和 $PO_R$ 所确定的平面被称为核心平面。该平面与左右象平面的交线为核心线。所以，左图中核心线上的一点，其右图中的对应点也在核心线上。利用核心线的这种几何关系，可以使得匹配过程极大的简化。

### § 1.4.2 立体匹配

立体匹配是寻求同一空间景物在不同视点下投影图像的像间的一一对应关系。与普通的图像模板匹配不同，立体匹配是在两幅存在视点差异，几何、灰度畸变和噪声干扰的图像间进行，不存在任何的标准模板。立体图像匹配的主要困难在于图像之间不可避免地存在着差异。造成图像之间存在差异的原因有很多，概括起来包括以下几种：

灰度差异：由于成像角度的不同及照明条件的变动，使景物表面上

某点在不同成像平面上的投影点的灰度值产生较大的差别。

遮挡效应引起的差异:现实景物之间,不可避免地存在着视线的遮挡效应,景物被遮挡的部位和遮挡的程度会随着视角的不同而改变。

特征抽取结果的差异:发生在立体像对特征之间的匹配是以图像特征提取为基础的,所以当两幅图的不同部位的反差存在差异时则会在反差弱的区域产生特征丢失的现象。丢失特征部位和丢失特征程度的差异给图像匹配带来了很大的困难。

噪声引起的差异:由于成像系统和图像预处理技术远非理想,在每个环节都会引进虚假噪声,并且对两幅图像的伤害程度可能不同,所以噪声也是引起图像匹配出错的重要原因。

由于这些难点,至今并无通用的匹配算法,各种算法都只是针对匹配技术的特定问题进行探讨,并且都有自己的限定条件。但是,在图像匹配过程中,各种算法都采用了各种匹配约束条件和匹配质量控制策略,以提高匹配质量,确保获得好的匹配结果。

### § 1.4.3 基本约束

为了帮助建立像素间的对应,已提出了许多约束,来减少搜索的范围和确定正确的对应。下面是一些最重要的约束。

#### (1) 外极线约束

如图 1.2 所示,投影线  $PO_L$  和  $PO_R$  所确定的核心平面与右象平面的交线称为左图像平面上那点  $P_L$  的外极线。容易看出,无论与  $P_L$  对应的点  $P$  是远是近,它总在那条外极线上,所以,对任一点  $P_L$ ,只要在它的外极线上去找它的对应点  $P_R$ ,这就是外极线约束。它大大限制了搜索对应点的范围,是立体视觉中最基本的约束。特别,如图 1.2 所示的成像系统模型,左右摄像机的光轴平行, X 轴一致,并且摄像机的水平扫描线位于同一平面时的简单情形时,左图像上一点  $P_L(x_l, y_l, z_l)$  外极线就是  $y_r = y_l$ 。

#### (2) 相容性约束

对应的特征,应有相同的属性,这一约束的具体含义随选用的特征及其属性不同而有所不同。

#### (3) 唯一性约束

除了极个别的情况，一幅图像(左或右)上的每个特征只能与另一幅图像上的唯一的一个特征对应。

#### (4) 连续性约束

由于物体表面一般是平滑的，因而它们的偏差一般也是平滑的，在物体边界处的深度不连续的地方除外。

#### (5) 灰度匹配约束

同名像点灰度应相同或相近。

#### (6) 顺序约束

景物在左右影像点的投影顺序应一致。

### § 1.4.4 匹配的实现

从立体匹配实现的技术上考虑，立体匹配可以分为基于区域的匹配和基于特征的匹配，由于基于特征的匹配不直接利用灰度值进行匹配，而是利用由灰度信息抽象得到的图像特征进行匹配。因此这种匹配技术具有对外界的变化不敏感、稳定性好、精度高、匹配速度快的优点，所以获得了较广泛的应用。基于图像特征的匹配方法的匹配基元为从灰度图像中获取的景物的抽象特征，通过强调空间景物的结构信息来解决匹配歧义性问题。该类匹配方法的匹配基元一般是经过处理后的图像边界点、直线、曲线或线段的组合、轮廓、区域等。这些特征都对应着图像灰度变化剧烈的像素位置，通常与三维景物的几何结构相对应，因而它们具有较强的抗干扰能力和较精确的空间几何关系。

零交叉匹配基元以及多通道协同处理方式对立体视觉的发展起了巨大的推动作用。1976年Marr基于其理论建立了模拟人类立视机制的匹配算法<sup>[12]</sup>，该算法的核心是首次采用多分辨率 $\nabla^2 G$  零交叉点作为匹配基元，利用零交叉分布特性和由粗到细引导的多通道协同匹配技术来解决匹配的歧义性和搜索空间大的问题。Baker<sup>[13]</sup>和Arnold<sup>[14]</sup>等提出以边缘点为匹配特征的匹配算法，并在匹配时首次采用了动态规划技术，将左右图像同一扫描线上的匹配特征，分别等效为动态规划的阶段与状态，通过在二维规划平面上搜索最佳路径得到最优的匹配。

已知立体成象模型和匹配视差后，三维距离的恢复是很容易的。影响距离测量精度的因素主要有摄象机标定误差、数字量化效应、特征检测与匹配定位精度等。一般来讲，距离的测量精度与匹配定位精度成正比，与摄象机基线长度成反比。增大基线长度可改善距离测量精度，但同时会增大图像间的差异，增加匹配的困难。因此，要设计一个精确的立视系统，必须综合考虑各方面的因素，保证各个环节都具有较高

的精度。

## § 1.5 课题主要研究内容

立体视觉的基本原理是从两个(或多个)视点观察同一景物,以获取在不同视角下的感知图像,通过三角测量原理计算图像像素间的位置偏差(即视差)来获取景物的三维信息,这一过程与人类视觉的立体感知过程是类似的。一个完整的立体视觉系统通常可分为图像获取、摄像机定标、特征提取、立体匹配、深度确定及内插等6个大部分。本课题的研究涉及到计算机视觉与图像理解,属于模式识别与人工智能领域。利用低层图像处理技术对双目图像进行分析,选择图像对中的目标特征并求解特征间的对应关系,通过图像匹配技术得到目标视差,从而转化为主体所需的深度信息。本课题沿着“图像获取—图像分割—特征提取—图像匹配—距离确定”这一主线,分层次地总结、分析了各个子阶段的相应问题和处理方法,并将本课题的重点集中于特征提取与图像匹配这两部分。

### (1) 图像预处理

立体视觉研究的对象为二维图像。由于各类图像系统中图像的传送和转换常造成图像的某些降质,如图像不清晰或发生畸变,不利于图像的特征提取,从而影响目标的定位和分析。应通过合适的图像预处理方法对图像进行滤波处理,以便于特征提取。

### (2) 图像特征提取

图像特征是景物的物理与几何特性在影像中的反映,理论上是灰度曲面的不连续点。图像特征可分为点状特征、线特征、边缘特征和面特征,为自动量测感兴趣的特征,必须借助于相应的图像特征提取方法,即利于一定算法提取构成明显目标的图像特征。图像特征的提取是图像分析和图像匹配的基础,也是单目图像处理的最重要的任务之一。

### (3) 双目视觉图像分析

#### ① 双目视觉中的成像系统模型的选择

双目视觉分析中的成像系统模型大致可分为两种:双目横模型结构和双目轴模型结构。模仿人眼的双目线索感知距离的方法,本课题确定了双目横模型结构,并分析了双目横模型结构的基本几何及其固有的问题。

#### ② 图像匹配

在图像特征提取与描述的基础上,即可进行图像的配准工作,即根据选取的匹配特征,在左、右视图中匹配这些特征。图像配准的主要研究内容包括特征类型的选择、匹配准则、选取有效的匹配控制策略等关键问题。选取合适的特征并准确匹配这些特征,才能得到比较准确的目

标视差图，从而恢复目标的三维信息，实现信息重建。

③求取目标点的距离

通过特征选取、图像匹配，利用视点几何即可求取目标点的距离。

## 第二章 图像获取

三维计算机视觉系统应能从摄像机获取的图像信息出发,计算三维环境物体的位置、形状等几何信息,并由此识别环境中的物体。图像上每一点的亮度反映了空间物体表面某点反射光的强度,而该点在图像上的位置则与空间物体表面相应点的几何位置有关。这些位置的相互关系,由摄像机成像几何模型决定<sup>[5]</sup>。该几何模型的参数称为摄像机参数,这些参数必须由实验与计算来确定,实验与计算的过程称为摄像机标定。

本章主要介绍摄像机工作原理、摄像机模型、立体视觉的测量原理。

### § 2.1 摄像机工作原理

摄像机是最常用的图像输入装置,它面向对象广泛,有输入速度快、灵敏度高、使用方便等特点。摄像机通常由摄像镜头、摄像器件、同步信号发生电路、偏转电路、放大电路、电源等部分组成。

按照摄像器件的组成,可将摄像机分为很多种。目前计算机视觉系统中常用的是电荷耦合器件(Charge Couple Device)摄像机,简称 CCD 摄像机,其工作原理是:

在硅单晶片上,生成很薄的一层二氧化硅,再于其上蒸发一层间距很小的铝电极条。加电压于电极上后,电场穿透二氧化硅薄层,并排斥硅中多数载流子,从而在电极下的硅单晶片上形成电荷耗尽层,又称势阱。当硅改变电极电压,使势阱内的电荷从一个电极传送到另一个电极,取出这移动的电荷就形成输出信号。

扫描系统顺序扫描各像素,在一个扫描周期内,用较长的时间感光、积累电荷,用极短的时间读取图像信号。由于图像各处的光亮不同,耗尽层内“捕捉”的电荷量形成差异,这就完成了图像的光电转换。

### § 2.2 摄像机模型

在计算机视觉的处理中,常用到一些坐标系,下面先对坐标系进行介绍。

#### (1) 图像坐标系

这是固定在图像上的二维坐标系,一般采用直角坐标系。通常将原点置于图像的左上角,其 X 轴向右, Y 轴向下,此时的 X 坐标变为列数  $j$ , Y 坐标变为行数  $i$ 。如图 2.1 所示



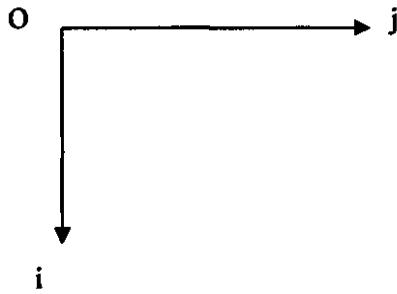


图 2.1 图像坐标系

(2) 观察者中心、坐标系

又简称摄像机坐标系，是固定在摄像机上的三维直角坐标系。

(3) 图像平面坐标系

在图像平面上的二维坐标系，一般采用直角坐标系。它与上面的图像坐标系之间存在一个线性变换的关系。

(4) 物体中心坐标系

固定于场景中某物体上的坐标系，一般采用三维直角坐标系。而且通常将该物体的三个主轴选为坐标轴，将该物体的中心(或某个顶点)选为坐标系原点。

这些坐标系之间的关系，特别是观察者中心坐标系与图像坐标系之间的关系的确就是摄像机的标定问题。下面讨论摄像机的模型。

我们要通过多个摄像机从不同方向拍摄两幅或两幅以上的二维图像，通常假定多个摄像机的参数一致，而且每个摄像机都被假设为一个理想的小孔摄像机，也就是采用透视投影成像模型。

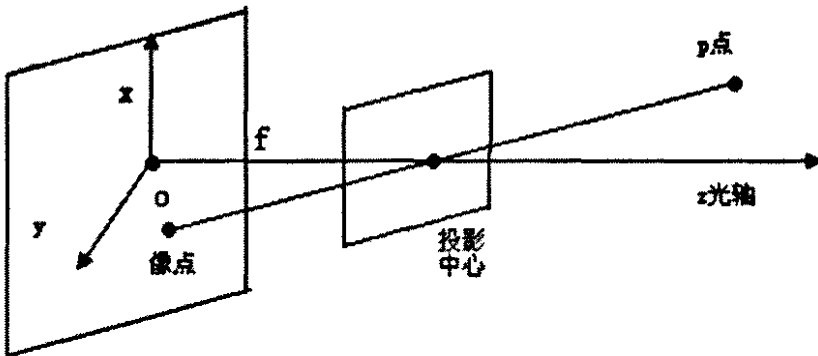


图 2.2 小孔摄像机模型

小孔成像模型如图 2.2 所示，图中位于图像平面前距离为  $f$  的点被

称为投影中心，由于图像平面位于投影中心的后面，这种小孔摄像机模型又称为后投影模型。小孔摄像机模型不考虑透镜的畸变，在大多数应用场合，这种模型可以满足精度要求。

## § 2.3 立体视觉测距原理

### § 2.3.1 视差原理

立体视觉是通过多幅图像获取物体三维几何信息的方法，对生物视觉系统，人们早就注意到，几乎所有具有视觉的生物都有两个眼睛，用两个眼睛同时观察一个物体时，会有深度或远近的感觉。双目立体视觉就是仿照这个原理，利用两台摄像机从不同角度同时获取同一景物的两幅图像，通过计算空间点在两幅图像中的视差来获取景物的三维坐标值。其视差测距原理<sup>[6]</sup>，如图 2.3 所示，设  $C_l$ 、 $C_r$  分别为左、右两个相机的光学中心位置(透镜中心)， $C_l$  与  $C_r$  之间的距离为  $b$ ，相机焦距为  $f$ 。设物体上的点  $P$  在左、右相机图像平面上的投影点分别为  $P_l$ 、 $P_r$ ， $P$  与  $C_l C_r$  连线间的距离为  $d$ ，过  $C_l$ 、 $C_r$  分别向图像面(即视平面)作垂线，垂足分别为  $A_l$ 、 $A_r$ ，过  $P$  向图像面作垂线，垂足为  $B$ 。令  $|A_l P_l| = l_a$ ， $|A_r P_r| = l_b$ ， $|P_r B| = a$ 。

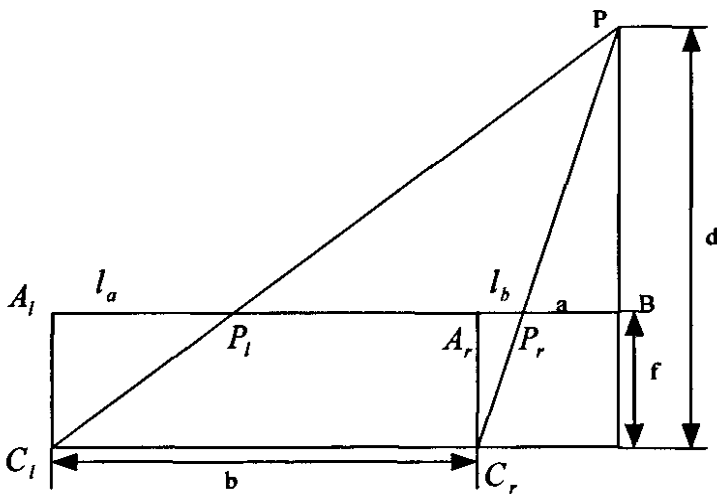


图 2.3 视差测距原理图

由相似三角形得

$$\frac{d-f}{d} = \frac{a}{a+l_b} \quad (2.1)$$

$$\frac{d-f}{d} = \frac{b-l_a+l_b+a}{b+l_b+a} \quad (2.2)$$

由式(2.1)、(2.2)可得

$$a = \frac{b \times l_a}{l_a - l_b} - l_b \quad (2.3)$$

代入式(2.1)得

$$d = f \frac{a+l_b}{l_b} = \frac{bf}{l_a - l_b} \quad (2.4)$$

由此可见距离  $d$  与  $b$ 、 $f$  和  $l_a - l_b$  有关,  $l_a - l_b$  称为点  $P$  在左右两个图像面上的视差, 两个图像相应的对应点只在水平方向存在视差, 而  $Y$  方向的坐标值是相等的。对于固定的成像系统模型而言, 参数  $b$ 、 $f$  取值已经确定, 仅需要对立体图像对求取对应像素视差即可获得景物的深度信息。

### § 2.3.2 双目视觉相机几何分析

在双目视觉中的成像系统模型分析中, 常见的有两种: 双目横模型结构和双目轴模型结构。在双目视觉分析中, 两相机不同的几何配置方法, 将直接影响双目的共同视野、搜索的范围以及定位的精度等。双目横模型结构是指两个相机平行放置, 而在双目轴模型结构中, 两个相机与目标在同一条直线上。常用的成像系统模型为仿照人眼的双目横模型结构。

用  $C_L$  和  $C_R$  两个相机同时观察  $P$  点, 并且如果我们能确定在  $C_L$  相机图像上的  $P_l$  点与在  $C_R$  相机图像上的  $P_r$  是空间同一点  $P$  的图像点(称  $P_l$ 、 $P_r$  为对应点), 则空间点  $P$  既位于  $O_l P_l$  上, 又位于  $O_r P_r$  上, 由此  $P$  点是  $O_l P_l$  和  $O_r P_r$  两条直线的交点, 即它的三维位置是唯一确定的。其关系如

图 2.4 所示。

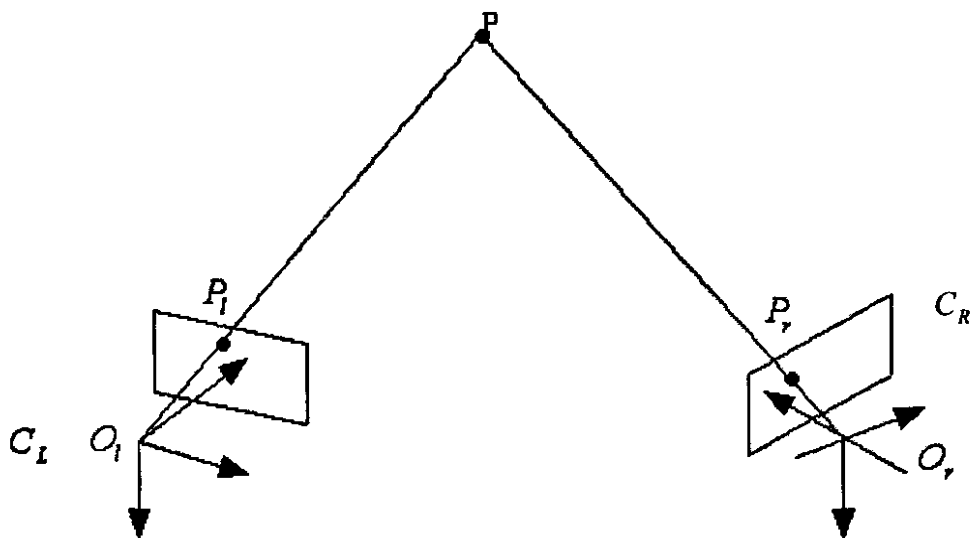


图 2.4 用摄像机观察空间点

关系式为<sup>[7]</sup>：

$$Zc_L \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11}^1 & m_{12}^1 & m_{13}^1 & m_{14}^1 \\ m_{21}^1 & m_{22}^1 & m_{23}^1 & m_{24}^1 \\ m_{31}^1 & m_{32}^1 & m_{33}^1 & m_{34}^1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.5)$$

$$Zc_R \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11}^2 & m_{12}^2 & m_{13}^2 & m_{14}^2 \\ m_{21}^2 & m_{22}^2 & m_{23}^2 & m_{24}^2 \\ m_{31}^2 & m_{32}^2 & m_{33}^2 & m_{34}^2 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.6)$$

其中：  $(u_1, v_1, 1)$  与  $(u_2, v_2, 1)$  为  $P_L$ 、 $P_R$  点在各自图像中的图像齐次坐标；  $m_{ij}^k$  ( $k=1,2; i=1,2,3; j=1,2,3,4$ ) 分别为投影矩阵的第  $i$  行第  $j$  列的元

素。 $(X, Y, Z, 1)$  为 P 点在世界坐标系下的齐次坐标。可在上面两个公式中消去  $Zc_L$ 、 $Zc_R$ ，得到关于 X, Y, Z 的四个线性方程<sup>[17]</sup>，如下所示：

$$(u_1 m_{31}^1 - m_{11}^1)X + (u_1 m_{32}^1 - m_{12}^1)Y + (u_1 m_{33}^1 - m_{13}^1)Z = m_{14}^1 - u_1 m_{34}^1$$

$$(v_1 m_{31}^1 - m_{11}^1)X + (v_1 m_{32}^1 - m_{12}^1)Y + (v_1 m_{33}^1 - m_{13}^1)Z = m_{14}^1 - v_1 m_{34}^1 \quad (2.7)$$

$$(u_1 m_{31}^2 - m_{11}^2)X + (u_1 m_{32}^2 - m_{12}^2)Y + (u_1 m_{33}^2 - m_{13}^2)Z = m_{14}^2 - u_1 m_{34}^2$$

$$(v_1 m_{31}^2 - m_{11}^2)X + (v_1 m_{32}^2 - m_{12}^2)Y + (v_1 m_{33}^2 - m_{13}^2)Z = m_{14}^2 - v_1 m_{34}^2 \quad (2.8)$$

式(2.7)的几何意义为过  $O_L P_L$  的直线，式(2.8)为过  $O_R P_R$  的直线。联立解得到的  $(X, Y, Z)$  即为点 P 的空间坐标。

在相机的双目横模型结构设置中，由于两相机在水平方向相隔一定距离，所以存在以下问题：

(1) 利用横模型系统的一个固有问题是无法匹配点所引起的不正确对应性。产生无法匹配点的原因有两种：一是目标元素不处于共同视野，只有一个像机的成像平面中有其投影；二是即使目标元素处于共同视野，一个像机平面中有其投影，而另一个像机的成像平面中，因为由别的目标元素所遮挡，不存在投影。在这两种情况下，对一幅图像选择这些投影点，而在另一幅图像中则找不到它的对应性。显然，减小像机之间的位移，可使其共同视野增加，出现上述情况的概率也随之减小。

(2) 在相机的双目横模型结构设置的另一个突出的问题是提高深度测量的精度与降低不正确对应性，对像机间的间隔  $\Delta X$  有不同的要求。为了扩大共同视野，就要减少像机间的距离，这样距离测量的准确度就变差。当像机间隔  $\Delta X = 0$  时，问题不存在了，但是双目视觉也就随之消失了。反之，若增加像机之间的距离，则将增加深度测量的精度，而共同视野及不含糊匹配的概率也随之降低。

## § 2.4 小结

本章介绍了计算机视觉系统中常用的 CCD 摄像机的工作原理，介绍了一些常用的坐标系并建立了摄像机的透视投影成像模型，以立体视觉中的平行轴立体成像模型为例，分析了立体视觉测距原理，引出了立体匹配问题。

## 第三章 基于机器视觉的图像处理技术的研究

图像处理技术涉及到数学、计算机科学、模式识别、人工智能、信息论、生物医学等学科，是一门多学科交叉应用技术。本章主要研究 Windows 环境下图像显示、像素灰度提取、图像滤波和图像增强等技术。

### § 3.1 Windows 位图

图像在计算机中是用文件的形式存放的，常用的图像文件格式有 BMP(Windows Bitmap)格式、PCX 格式、GIF 格式、TIFF 格式、JPEG 格式等<sup>[18]</sup>。Visual C++主要针对位图 BMP 进行处理，位图也是本文的主要研究图像，对于其它格式的图像，我们可以用图像处理软件如 ACDSee 把它转换成 BMP 格式。

BMP 格式的文件内容包括文件头(一个 BITMAP FILEHEADER 数据结构)、位图信息数据块(位图信息头 BITMAP INFOHEADER 和一个调色板 Palette)和图像数据<sup>[19]</sup>。对于用到调色板的位图，图像数据就是该像素在调色板中的索引值，对于真彩色图，图像数据就是实际的 R, G, B 值。

### § 3.2 像素灰度提取

图像数字化时，物理图像被划分为称作图像元素的小区域，图像元素简称像素，最常见的划分是采用方形采样网格。在每个像素位置，图像的亮度被采样和量化，从而得到图像对应点上表示其亮暗程度的一个整数值。对所有的像素都完成上述转化后，图像就被表示成一个整数矩阵。每个像素具有两个属性：位置和灰度。位置由采样点的两个坐标决定，又称为行和列；表示该像素位置上亮暗程度的整数称为灰度。此数字矩阵就作为计算机处理的对象。

图像处理的过程中可对像素灰度进行修改，处理后的结果由一个与数字化相反的逆过程显示出来，即用每个像素的灰度决定对应点在显示屏上的亮度。这样就转化成可视的、供人们解释的图像。

在计算机中，常用一个字节(Byte)存储灰度信息，即灰度级为 256 级。如果图像的灰度级只有黑白两种，用一位(bit)存储，则称为二值图像。一般常用三原色(红色、绿色、蓝色)来产生彩色图像，即用(r, g, b)表示，其中 r, g, b 为颜色分量。还有一种用亮度、饱和度和色度来表示

颜色的方法，这里不介绍了。

灰度图像的 RGB 三个分量的值相同，处理很方便。在 256 级灰度图中，图像数据一个字节就代表一个像素，如果灰度级是 2 级或 16 级，还要拼凑字节，比较麻烦。所以本文主要讨论 256 级灰度图。

对于灰度图像，因为其 RGB 值相同，所以只要定义一个二维数组  $p[x][y]$  即可，数组元素的值，也就是图像中每一点的灰度值，这个二维数组就是我们用计算机处理的对象，它和图像是一一对应的。

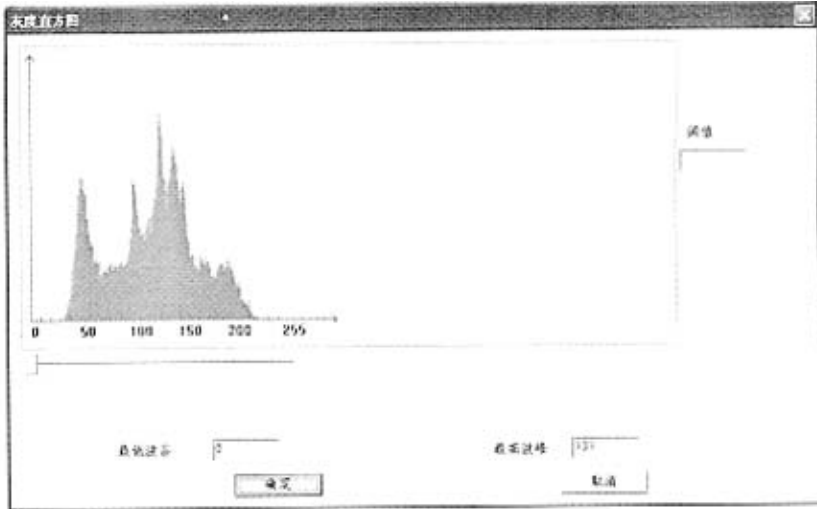
一幅图像所有像素的灰度信息可以用灰度直方图来描述，灰度直方图是灰度级的函数，描述的是图像中具有该灰度级的像素的个数：其横坐标是灰度级，纵坐标是该灰度出现的频率（像素的个数）。

计算灰度直方图的算法如下：

```
for(i=0;i<height;i++)
    for(j=0;j<height;j++)
    {
        k=p[i][j];
        gray[k]++;
    }
```

下图是实验中的一幅图像及其灰度直方图：





直方图给出了一个简单的指示,用来判断一幅图像是否合理的利用了全部被允许的灰度级范围,这对分析数字化的效果很有用。通过直方图还可以选择合适的灰度阈值,进行图像分割和边界提取。

### § 3.3 图像预处理

由于各种噪声的干扰和影响,在对图像进行特征提取前需先进行图像的平滑滤波等操作,以消除噪声的影响。为了抑制数字图像中的噪声,需采用空间低通滤波器。低通滤波器不影响低频分量,但由于图像边缘轮廓包含多种高频信息,所以在抑制噪声的同时会使图像边缘变得模糊。若采用高频滤波器提取图像特征又会增强图像噪声。图像处理中解决该矛盾的手段多采用中值滤波法或复合型中值滤波法。

图像预处理过程通常包括噪声消除、图像增强、边缘提取等过程。

图像增强是指对图像的某些特征,如边缘、轮廓、对比度等进行强调或锐化,以便于显示、观察或进一步分析与处理<sup>[20]</sup>。图像增强不增加图像数据中的相关信息,但它增加所选择特征的动态范围,从而使这些特征检测或识别更加容易。

图像增强的主要方法有:基于点运算的方法、基于空间运算的方法、基于频域运算的方法和彩色增强方法。

基于点运算的图像增强方法属于时域处理方法,实际上是对图像的灰度级进行变换,使图像对比度得到调整,是一种点对点的变换。主要有以下几种:

- (1) 灰度线性变换
- (2) 灰度非线性变换



(3) 灰度切片变换

(4) 直方图修整

基于空间运算的图像增强方法属于时域处理方法，是一种利用各像素及其邻近各点的像素值来判断该点是否含有噪声，并用适当的方法消除所发现的噪声的图像增强方法。主要有以下几种：

(1) 图像平滑

(2) 图像锐化

(3) 反对比度映射和统计比例尺度变换

### § 3.3.1 灰度变换

(1) 灰度线性变换

假定原图像的灰度范围为 $[a, b]$ ，希望变换后的图像 $g(x, y)$ 的灰度范围为 $[c, d]$ ，可用下述变换来实现：

$$g(x, y) = \frac{(d-c)[f(x, y) - a]}{b-a} + c \quad (3.1)$$

如果已知图像的大部分像素的灰度级分布在区间 $[a, b]$ ，小部分像素的灰度级不在此区间内，下列变换也可改善图像的效果。

$$g(x, y) = \begin{cases} \frac{(d-c)[f(x, y) - a]}{b-a} + c & a \leq f(x, y) \leq b \\ c & f(x, y) < a \\ d & f(x, y) > b \end{cases} \quad (3.2)$$

下面是实验中的一幅图像及对其进行灰度线性变换后的图像：



图 3.2 原图像及灰度线性变化图像

## (2) 灰度非线性变换

当用某些非线性函数，如对数、指数函数等作为映射函数时，可实现灰度的非线性变换。

对数变换一般为：

$$g(x, y) = a + \frac{\ln[f(x, y) + 1]/b}{\ln c} \quad (3.3)$$

这里  $a, b, c$  是可调参数。当希望对图像的低灰度区有较大的扩展而对高灰度区压缩时，可采用此变换。它能使图像灰度分布均匀，与人的视觉特性相匹配。

指数变换的形式为：

$$g(x, y) = b^{c[f(x, y) - a]} - 1 \quad (3.4)$$

这种变换能对图像的高灰度区给予较大的扩展。

## (3) 灰度分层切片

此变换把输入图像中的某一段灰度范围值抽取出来，转换成最大灰度值。

$$g(x, y) = \begin{cases} MAX & a < f(x, y) < b \\ 0 & \text{其它} \end{cases} \quad (3.5)$$

### § 3.3.2 直方图修整

直方图修整是指通过构造灰度级变换，改造原图像的直方图，使变换后的图像的直方图达到一定的要求。采用直方图修整可使原图像灰度级集中的区域拉开或使灰度分布均匀，从而增大反差，使图像的细节清晰，达到增强的目的<sup>[21]</sup>。

设变量  $r$  代表要增强图像中像素的灰度级，假定  $r$  已经归一化，即  $0 \leq r \leq 1, r=0$  代表黑， $r=1$  代表白。设新图像的灰度级为  $s$ ， $s$  与  $r$  的关系为  $s=T(r)$ 。假定  $T(r)$  满足 ①  $0 \leq r \leq 1, T(r)$  单调增加；②  $0 \leq r \leq 1, 0 \leq T(r) \leq 1$ 。

一幅图像中，每一像素点的灰度级  $r$  可看作是  $[0, 1]$  区间上的随机变量。假定  $r$  是连续变量，则可用概率密度函数  $P_r(r)$  表示图像的灰度级分布，即  $F_r(r) = \int P_r(w)dw$  表示的是图像中灰度级小于  $r$  的像素点的

数目。设  $P_s(s)$  表示变换后的图像的灰度级概念密度函数。根据概率论的知识,  $P_r(r)$  与  $P_s(s)$  间应有下列关系:

$$P_s(s) = [P_r(r) \frac{dr}{ds}]_{r=T^{-1}(s)} \quad (3.6)$$

通过改变  $T(r)$  就可控制  $P_s(s)$  的形状。由于  $P_r(r)$  与  $P_s(s)$  刻划了图像的灰度级相对分布, 因而选择适当的  $T(r)$  就可改变图像的外貌, 这就是直方图修整法的基础。

取  $T(r)$  为

$$T(r) = \int_0^r P_r(w) dw, \quad 0 \leq r \leq 1 \quad (3.7)$$

则易知  $P_s(s) = 1$ 。这说明, 在变换后的变量  $s$  的定义域上,  $P_s(s)$  是一均匀分布的概率密度函数。在增强的意义下, 这意味着像素灰度级的动态范围增加, 对比或反差增大。

在数字图像的情况下, 灰度级  $r$  取离散值  $\{r_k; k=0, 1, \dots, L-1\}$ ,  $L$  是灰度级的数目, 则连续函数  $P_r(r)$  成为  $\{P_r(r_k); k=0, 1, \dots, L-1\}$ 。其中  $P_r(r_k) = n_k / n$ ;  $n_k$  是图像中出现  $r_k$  灰度级的像素的数目,  $n$  是图像中像素总数。变换  $s=T(r)$  成为:

$$s_k = T(r_k) = \sum_j^k \frac{n_j}{n} = \sum_j^k P_r(r_j); \quad k=0, 1, \dots, L-1 \quad (3.8)$$

在上述变换下, 数字图像的直方图将成为均匀分布形状, 这一过程称为直方图的均衡化。其实质是减小图像的灰度等级以换取对比度的扩大, 但由于在均衡化的过程中, 原直方图上频次较小的灰度级被并入少数几个或一个灰度级, 对应的图像部分将得不到增强。

下面是实验中的一幅图像及对其进行直方图均衡化后的图像:



图 3.3 原图像及其直方图均衡化图像

### § 3.3.3 图像平滑

图像在生成和传输过程中常受到各种噪声源的干扰和影响而使图像质量变差，抑制或消除这类噪声而改善图像质量的过程称为对图像的平滑过程。图像平滑的方法主要有：

#### (1) 邻域平均法

在邻域平均法中，假定图像是由许多灰度恒定的小块组成，相邻像素间有很高的空间相关性，而噪声是统计独立的叠加在图像上的。因此，可用像素领域内的各像素灰度值的平均代替该像素原来的灰度值，实现

图像的平滑。根据求平均方式的不同,邻域平均法有如下几种主要形式:

①简单平均法

设图中某像素的灰度值为  $f(x, y)$ , 它的邻域  $S$  为  $M \times N$  的矩形窗口, 则平滑后该点的灰度值为:

$$\overline{f(x, y)} = \frac{1}{MN} \sum_{(u, v) \in S} f(u, v) = \frac{1}{MN} \sum_{i=-\frac{M-1}{2}}^{\frac{M-1}{2}} \sum_{j=-\frac{N-1}{2}}^{\frac{N-1}{2}} f(x+i, y+j) \quad (3.9)$$

如果图像中的噪声是随机不相关的加性噪声, 窗口内各点的噪声是独立分布的, 经过上述平滑后, 信噪比可提高  $\sqrt{MN}$  倍。

②阈值平均法

设原像素点的灰度值为  $f(x, y)$ , 取以它为中心点  $M \times N$  的大小的窗口, 则平滑后的灰度置为:

$$g(x, y) = \begin{cases} \overline{f(x, y)} & \text{若 } |f(x, y) - \overline{f(x, y)}| > T \\ f(x, y) & \text{其它} \end{cases} \quad (3.10)$$

$T$  为给定的阈值,  $\overline{f(x, y)}$  的计算方法同简单平均法。这种算法对抑制椒盐(salt&pepper)噪声比较有效, 同时也能较好的保护仅有微小灰度差的图像细节。

③K-近邻平均法

在一个  $M \times N$  的窗口中, 属于同一个物体的像素, 它们的灰度值将高度相关。因此, 窗口中心像素的灰度值可用窗口内与中心点灰度最接近的  $K$  个邻点的平均灰度来代替。一般而言,  $K$  值越小, 噪声方差降低越小, 但保持细节较好。较大的  $K$  值平滑噪声效果好, 但也使图像模糊。

④梯度倒数加权平均法

通常图像在一个区域内的灰度变化要比区域之间的变化小, 在边沿处的梯度绝对值要比区域内部的绝对值高。基于这样的假设, 在一个  $M \times N$  的窗口内, 若把中心像素点与其它各邻点之间梯度绝对值的倒数定义为各邻点灰度平均时的加权值, 则在区域内部的点, 加权值大; 而在一条边沿近旁的和位于区域外的那些邻点, 加权值小。这种加权平均既可以使图像得到平滑, 又不致使边沿和细节有明显模糊。具体算法如下:

设点  $(x, y)$  的灰度值为  $f(x, y)$ , 在它的一个  $N \times N$  邻域内, 定义梯度倒数为<sup>[22]</sup>:

$$g(x, y; i, j) = \begin{cases} \frac{1}{|f(x+i, y+j) - f(x, y)|} & \text{如果 } f(x+i, y+j) \neq f(x, y) \\ k & \text{如果 } f(x+i, y+j) = f(x, y) \end{cases} \quad (3.11)$$

上式中  $i, j = -\frac{N-1}{2}, \dots, \frac{N-1}{2}$ 。k 为一个大于 1 的数。

定义一个归一化的权重矩阵 W 如下：

$$W = \begin{bmatrix} w(x, y; -\frac{N-1}{2}, -\frac{N-1}{2}) & \dots & w(x, y; 0, -\frac{N-1}{2}) & \dots & w(x, y; \frac{N-1}{2}, -\frac{N-1}{2}) \\ \dots & \dots & \dots & \dots & \dots \\ w(x, y; -\frac{N-1}{2}, 0) & \dots & w(x, y; 0, 0) & \dots & w(x, y; \frac{N-1}{2}, 0) \\ \dots & \dots & \dots & \dots & \dots \\ w(x, y; -\frac{N-1}{2}, \frac{N-1}{2}) & \dots & w(x, y; 0, \frac{N-1}{2}) & \dots & w(x, y; \frac{N-1}{2}, \frac{N-1}{2}) \end{bmatrix}$$

规定  $w(x, y; 0, 0) = \frac{1}{k} < 1$ ；对  $i, j$  不同时为 0 时，有：

$$w(x, y; i, j) = \left(1 - \frac{1}{k}\right) \frac{g(x, y; i, j)}{\sum_{i, j \text{ 不同时为 } 0} g(x, y; i, j)} \quad (3.12)$$

平滑后的灰度值为：

$$\overline{f(x, y)} = \sum_i \sum_j w(x, y; i, j) f(x+i, y+j) \quad (3.13)$$

## (2) 低通滤波法

从频谱上看，噪声特别是随机噪声是一种具有较高频谱分量的信号。平滑的目的就是要通过一定的手段滤去这类信号。一个很自然的想法就是使图像经过一个二维的低通数字滤波器，让高频信息得到较大的衰减。在空间域上进行的这种滤波实际上就是对图像和滤波器的冲击响应函数进行卷积。

设图像为  $f(x, y)$ ，滤波器的冲击响应函数为  $H(x, y)$ ，则卷积表达式为：

$$g(u, v) = \sum_x \sum_y f(x, y) H(u-x+1, v-y+1) \quad (3.14)$$

在进行卷积之前， $f(x, y)$  和  $H(x, y)$  都要进行延拓。

常用的冲击响应函数有：

$$H_1 = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$H_2 = \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$H_3 = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

采用  $H_1$  作为滤波器，其效果和  $3 \times 3$  窗口下的简单邻域平均法得到的结果类似。

### (3) 中值滤波法

中值滤波的思想是对一个窗口内的所有像素的灰度值进行排序，取排序结果的中间值作为原窗口中心点处的像素灰度值。这种方法对脉冲干扰的抑制效果较好。中值滤波的关键在于选择合适的窗口大小和窗口形状。一般很难事先确定窗口的尺寸，通常的做法是从小到大进行多次尝试。窗口的形状可选为正方形，也可选为十字形。另一个重要的因素是算法的执行速度，由于每次都需要对窗口内的像素进行排序，因而必须有有效的快速算法。

中值滤波器是一种非线性的空间滤波器。为了对一幅图像上的某个像素点作中值滤波处理，必须先将掩模内欲求的像素及其邻域的像素值排序，确定出中值，并将中值赋予该像素点。掩模的大小可根据情况而定，一般情况下可选用  $5 \times 5$  的方形窗口，有时为了更好地保留图像的边角特征也采用交叉的十字型窗口。中值滤波器的主要功能是使拥有不同灰度的点看起来更接近于它的邻近值。事实上，是用  $n \times n$  的中值滤波器去除那些相对于其邻域像素更亮或更暗，并且，其区域小于  $n^2/2$  (滤波器区域的一半) 的孤立像素集。在这种情况下，将其像素值强制为邻域的中间亮度。

下面是实验中一幅图像及其采用  $3 \times 3$  矩阵窗口进行中值滤波后的图像。(图像含有椒盐(salt&pepper)噪声。)



图 3.3 原图像及其中值滤波法平滑后的图像

如上图所示，噪声被明显得消除了，图像质量得到了改善。

#### § 3.3.4 图像锐化

图像锐化的目的是增强图像中景物的边缘和轮廓。常用的方法有微分法、高通滤波法等<sup>[23]</sup>。

##### (1) 微分法

在图像中，边缘是由灰度级和邻域点不同的像素点构成的。因而，若想增强边缘，就应该突出相邻点的灰度级的变化。微分运算可用来求信号的变化率，因而具有加强高频分量的作用。如果将其应用在图像上，可使图像的轮廓清晰。由于无法事先确定轮廓的取向，因而挑选用于轮廓增强的微分算子时，必须选择那些不具备空间方向性的和具有旋转不变性的线性微分算子。



图像处理中最常用的微分方法是求梯度。对于图像  $f(x, y)$ ，它在点  $(x, y)$  处的梯度是一个矢量，定义为：

$$G(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (3.15)$$

梯度的两个重要性质是：

- (1) 矢量  $G(x, y)$  指向函数  $f(x, y)$  最大增加率的方向。
- (2)  $G(x, y)$  的幅度表示在  $G$  的方向上每单位距离  $f(x, y)$  的最大增加率。

$G(x, y)$  的幅度定义为：

$$GM(x, y) = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad (3.16)$$

对数字图像而言，微分运算一般用差分来代替。常用的差分形式有两种：

$$(1) GM(x, y) \approx |f(x, y) - f(x+1, y)| + |f(x, y) - f(x, y+1)| \quad (3.17)$$

$$(2) GM(x, y) \approx |f(x, y) - f(x+1, y+1)| + |f(x+1, y) - f(x, y+1)| \quad (3.18)$$

利用差分运算时，图像的最后一行和最后一列的像素的梯度无法求得，一般用前一行和前一列的梯度值近似代替。

梯度算出后，有多种方法可产生不同的梯度增强图像。

① 让梯度图像的灰度值  $g(x, y)$  就等于该点的梯度幅度， $g(x, y) = GM(x, y)$ ，这种方法的特点是增强的图像仅显示变化较陡的边缘轮廓，灰度变化较小的区域则呈黑色。

$$\textcircled{2} \text{ 使 } g(x, y) = \begin{cases} GM(x, y) & \text{若 } GM(x, y) \geq T \\ f(x, y) & \text{其它} \end{cases}, \text{ 式中 } T \text{ 是一个非负的}$$

阈值，适当选取  $T$ ，有可能既不破坏平滑背景的特征，又强调了有效边缘。

$$\textcircled{3} \text{ 是上一种的变化，它使 } g(x, y) = \begin{cases} L_0 & \text{若 } GM(x, y) \geq T \\ f(x, y) & \text{其它} \end{cases}, \text{ 式}$$

中  $L_0$  是根据需要指定的一个灰度级，在这种图像中，有效边缘是用一个

固定的灰度级表征的。

$$\textcircled{4} \text{使 } g(x, y) = \begin{cases} GM(x, y) & \text{若 } GM(x, y) \geq T \\ L_B & \text{其它} \end{cases}, \text{ 式中 } L_B \text{ 是对背景规定的}$$

的灰度值。

$$\textcircled{5} \text{使 } g(x, y) = \begin{cases} L_0 & \text{若 } GM(x, y) \geq T \\ L_B & \text{其它} \end{cases}, \text{ 得到二值图像, 容易观察边}$$

缘。

## (2) 高通滤波法

如前所述, 边缘是由灰度级跳变点构成, 因此具有较高的空间频率。采用高通滤波的方法让高频分量顺利通过, 使低频分量得到抑制, 就可以增强高频分量, 使图像的边缘变得清晰, 实现图像的锐化。这一思想反映在空间域处理中, 就是让图像和高通滤波器的冲击响应函数进行卷积。所用的表达式与低通滤波法中所用的相同, 只是冲击响应函数  $H$  不同, 常用于高通滤波的  $H$  有:

$$H_1 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$H_2 = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

$$H_3 = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

下面是实验中一幅图像及对其采用高通滤波法锐化后的图像, 采

用的冲击响应函数为  $H_2 = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$ 。



图 3.4 原图像及其高通滤波法锐化后的图像

### § 3.4 小结

对灰度图像进行特征提取等预处理过程是进行图像配准的前提条件。像素灰度提取和灰度直方图计算，可以得到图像中每个像素点的灰度值和所有点的灰度统计，以便进行灰度运算和为图像分割、边界提取选择合适的灰度阈值。图像增强的目的是使图像特征检测或识别更加容易，本章研究了灰度变换、直方图修整、图像平滑、图像锐化等图像增强方法，用 Visual C++编程实现了其中一些算法，并给出了处理后的效果图。

## 第四章 图像特征点的提取

图像的特征是指图像中可用作标志的属性。它有多种形式，可大致分为图像的统计特征和图像的视觉特征两类。图像的统计特征是一些人为特征，需通过变换才能得到，如图像的直方图、矩、频谱等。图像的视觉特征是指人的视觉可直接感受到的自然特征，如区域的亮度、纹理或轮廓等。利用这两类特征将图像分解成一系列有意义的目标或区域的过程称为图像分割。

本章主要讨论图像的边缘检测技术和图像特征点的提取。

### § 4.1 边缘检测

边缘是图像最基本的特征之一，包含了丰富的信息。边缘检测也是一种图像分割方法，在计算机视觉、图像分析等应用中起着非常重要的作用，是图像分析与识别的重要环节，同时，也是图像分析和模式识别的主要特征提取手段。所谓的边缘检测是指通过检测图像不同特性区域的边界来解决图像分割问题。与其它图像分割方法相比，边缘检测可以说是被研究的最多的一种分割方法。一般情况下图像边缘为图像灰度变化比较剧烈的地方，很自然地，我们可以联想到数学中的导数就反映了函数值的突变情况，在突变的地方求导数，可以得到绝对值较大的数值。很多传统的图像边缘检测算法都直接或间接的利用了这一思路<sup>[24]</sup>。

经典的边缘提取方法是考察图像的每个像素的某个邻域内灰度的变化，利用边缘邻近一阶或二阶方向导数变化规律进行边缘检测。在一幅图像中，沿边缘走向的灰度变化平缓，而垂直于边缘走向的灰度变化剧烈。边缘点对应于一阶微分幅度最大的点，也对应于二阶微分的零交叉点。因此，利用一阶微分(梯度)最大值或一阶导数过零点提取边缘点就成为一种有力的手段，这种方法也称边缘检测局部算子法。最基本的边缘检测算子有 Roberts 算子、Sobel 算子、Prewitt 算子、Laplacian 算子等。微分算子运算简单，但对噪声敏感、抗干扰性能差。

#### § 4.1.1 传统的边缘检测算子

边缘发生在灰度函数不连续的位置，可以适应灰度函数的一阶或二阶导数求得。直接使用导数运算判别边缘是不可取的，因为导数运算具有方向性。因此，我们经常需要的一种导数运算，能够保证各向同性。可以证明，偏导数的平方是各向同性的，即梯度幅度的平方。实际上，在图像处理中，梯度是经常被使用的一种微分运算。设图像为  $f(x, y)$ ,

则在点  $(x, y)$  的梯度矢量  $G[f(x, y)]$  为<sup>[25]</sup>：

$$G[f(x, y)] = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (4.1)$$

梯度具有两个重要的性质：

(1) 梯度的方向是函数  $f(x, y)$  最大变化率的方向。

(2) 梯度的幅度用  $G[f(x, y)]$  表示，则

$$G[f(x, y)] = \sqrt{[\Delta_x f(x, y)]^2 + [\Delta_y f(x, y)]^2} \quad (4.2)$$

根据  $\Delta_x f(x, y)$ 、 $\Delta_y f(x, y)$  的不同表示法，我们可以得到常用的几种边缘检测算子。

(1) 一阶微分算子

① Roberts 算子

图像函数  $f(x, y)$  在点  $(x, y)$  处梯度值的大小为

$$g(x, y) = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad (4.3)$$

Roberts 提出的算子是在  $2 \times 2$  邻域上计算对角导数，图像点  $(x, y)$  的梯度  $g(x, y)$  的幅度是用方向差分的均方值来近似，即

$$g(x, y) \approx R(x, y) = \sqrt{[f(x, y) - f(x+1, y+1)]^2 + [f(x+1, y) - f(x, y+1)]^2} \quad (4.4)$$

实际应用中，也可用

$$g(x, y) \approx R(x, y) = |f(x, y) - f(x+1, y+1)| + |f(x+1, y) - f(x, y+1)| \quad (4.5)$$

来近似。

Roberts 算子的卷积核为

1	0
0	-1

0	1
-1	0

② Sobel 算子

Sobel 算子是在以点  $(x, y)$  为中心的  $3 \times 3$  邻域上计算  $x$  和  $y$  方向的

偏导数，即

$$S_x = \{f(x+1, y+1) + 2f(x+1, y) + f(x+1, y+1)\} \\ - \{f(x-1, y-1) + 2f(x-1, y) + f(x-1, y+1)\} \quad (4.6)$$

$$S_y = \{f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1)\} \\ - \{f(x-1, y-1) + 2f(x, y-1) + f(x+1, y-1)\} \quad (4.7)$$

其梯度大小为

$$g(x, y) \approx S = \sqrt{S_x^2 + S_y^2} \text{ 或 } g(x, y) \approx S = |S_x| + |S_y| \quad (4.8)$$

Sobel 算子的卷积核为

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

这两种算子简单易用，边缘检测的效果较好<sup>[26]</sup>。但要注意的是，由于 Sobel 算子先做加权平均，然后再微分，具有一定的噪声抑制能力，但检测到的边缘宽度至少为两个像素，所以在实际应用中还需涉及到边缘细化技术。

### ③Kirsch 算子

对数字图像  $f(x, y)$  的每一像素  $(x, y)$ ，考察它的八个邻点的灰度变化，以其中三个相邻邻点的加权和减去剩下五个邻点的加权和，取其中差值的最大值作为 Kirsch 算子值。

$$K(x, y) = \max(5S_i - 3T_i) \quad (i=0, 1, 2, \dots, 7) \quad (4.9)$$

式中  $S_i$  = 三邻点之和， $T_i$  = 五邻点之和。

Kirsch 算子的卷积核为

-3	-3	5
-3	0	5
-3	-3	5

-3	-3	-3
-3	0	5
-3	5	5

-3	-3	-3
-3	0	-3
-3	-3	5

-3	-3	-3
5	0	-3
5	5	-3

5	-3	-3
5	0	-3
5	-3	-3

5	5	-3
5	0	-3
-3	-3	-3

5	5	5
-3	0	-3
-3	-3	-3

-3	5	5
-3	0	5
-3	-3	-3

用这 8 个卷积核对图像中的每一个点进行卷积，每个卷积核对某个特定的边缘方向做出最大响应，所有 8 个方向中的最大值作为边缘幅度图像的输出。

用一阶微分算子估计出梯度值后，将梯度值与给定的阈值比较，如果点(x,y)处的梯度值大于阈值，则(x,y)为边缘点。显然，阈值的选择是很重要的，在含有噪声的图像中，阈值选择必须在丢失有效边缘和噪声造成的错误边缘之间折中考虑。

## (2) 二阶微分算子

### Laplacian 算子

它是一种不依赖于边缘方向的二阶微分算子，它是一个标量而不是一个向量，具有旋转不变性即各向同性，在图像处理中常用来提取边缘。其表达式为：

$$\nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} \quad (4.10)$$

其离散形式定义为

$$\begin{aligned} \nabla^2 f(x, y) = L(x, y) = & \{ [f(x+1, y) - f(x, y)] - [f(x, y) - f(x-1, y)] \} \\ & + \{ [f(x, y+1) - f(x, y)] - [f(x, y) - f(x, y-1)] \} \end{aligned} \quad (4.11)$$

也可写成

$$L(x, y) = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y) \quad (4.12)$$

Laplacian 算子的卷积核为

0	1	0
1	-4	1
0	1	0

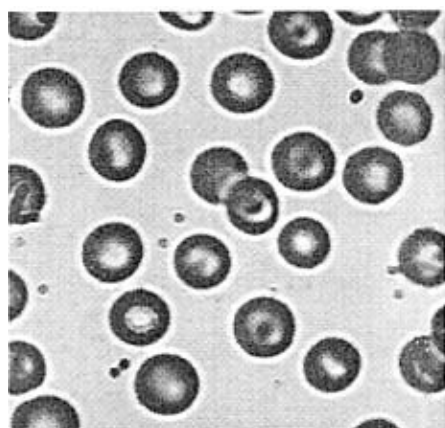
也可用

1	1	1
1	-8	1
1	1	1

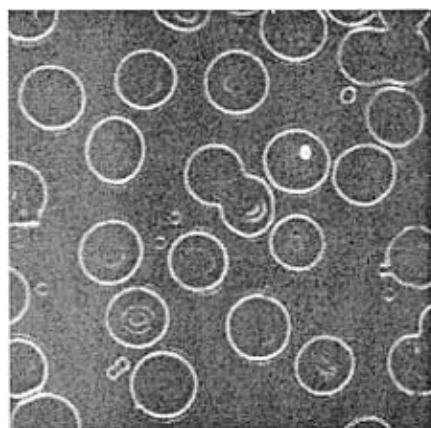
与前两种一阶微分算子相比较，二阶微分 Laplacian 算子对噪声更敏感，它使噪声成分加强，因此在实际使用中，常常先对图像进行平滑，再进行二阶微分。

对 blood 图像进行实验，几种简单边缘提取算法提取效果对比如图 4.1 所示，其中，(a)为 blood 原图，(b)为 Sobel 算子边缘检测结果图，(c)为 Robert 算子边缘检测结果图，(d)为 Kirsch 算子边缘检测结果图，(e)为 Laplacian 算子边缘检测结果图。

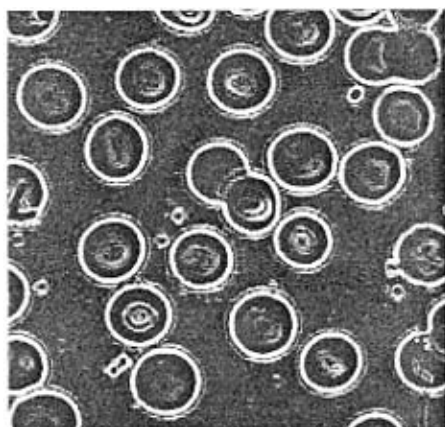




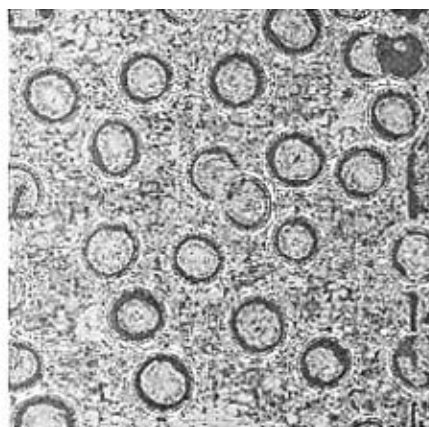
(a) 原图像



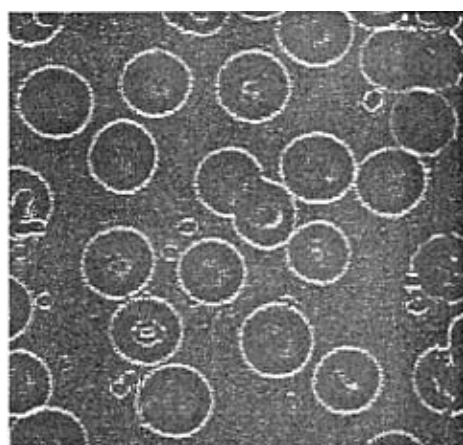
(b) Sobel 算子



(c) Robert 算子



(d) Kirsch 算子



(e) Laplacian 算子

图 4.1 原图像及其边缘检测效果图

### § 4.1.2 Marr 的边缘检测算法

图像的边缘是像素灰度值突变点,而图像的噪声常常也是这样的一类点。所以在使用上节的边缘检测算子时,图像的噪声点与边缘点同样被提取出来,这是我们不希望看到的。克服这一缺陷的一种可行的方法是首先对图像进行滤波,过滤掉那些孤立的噪声点,然后再进行边缘提取。Marr 算法进行边缘检测的思路就是如此<sup>[27]</sup>。

设平滑滤波器的冲击响应函数为  $h(x)$ ,待处理的信号为  $f(x)$ ,则滤波后的信号为  $g(x) = f(x) * h(x)$ 。然后,再对  $g(x)$  求一阶或二阶导数,就可以检测边缘点。注意到,卷积运算有以下性质:

$$\begin{aligned} g'(x) &= \frac{df(x) * h(x)}{dx} = \frac{d}{dx} \int_{-\infty}^{\infty} f(s)h(x-s)ds \\ &= \int_{-\infty}^{\infty} f(s)h'(x-s)ds = f(x) * h'(x) \end{aligned} \quad (4.13)$$

这样,就可以将卷积运算与微分运算合并,用平滑滤波器的一阶或二阶导数直接与图像卷积。检测  $f(x) * h'(x)$  的局部最大值或  $f(x) * h'(x)$  的过零点即可检测出图像的边缘点。这样,问题就转换到设计平滑滤波器上。平滑滤波器  $h(x)$  应满足以下条件:

- (1) 当  $|x| \rightarrow \infty$  时,  $h(x) \rightarrow 0$  为偶函数。
- (2)  $\int_{-\infty}^{\infty} h(x)dx = 1$ 。
- (3)  $h(x)$  一阶及二阶可微。

常用的平滑滤波器为 Gauss 函数,即  $h(x) = \frac{1}{\sqrt{2\pi}\sigma^3} e^{-\frac{x^2}{2\sigma^2}}$ , 求解它的一阶和二阶导数都很方便,分别为:

$$h'(x) = \frac{-x}{\sqrt{2\pi}\sigma^3} e^{-\frac{x^2}{2\sigma^2}} \quad (4.14)$$

$$h''(x) = \frac{1}{\sqrt{2\pi}\sigma^3} e^{-\frac{x^2}{2\sigma^2}} \left( \frac{x^2}{\sigma^2} - 1 \right) \quad (4.15)$$

其中,  $\sigma$  为 Gauss 函数的方差, 称为 Gauss 分布的空间尺度因子。 $\sigma$  越小, 则函数越“集中”仅在一个很小的范围内进行平滑滤波; 反之,  $\sigma$  越大, 则函数的滤波范围越大。但是随之又带来新的问题: 滤波范围太大, 很多我们感兴趣的信号突变点也被平滑掉了。因此, 选择合适的  $\sigma$ , 对边缘提取的质量有着很大的作用。

对于二维图像信号, Marr 采用的平滑函数为二维 Gauss 函数, 如下式所示:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (4.16)$$

它的一阶和二阶偏导数都是很容易求得的。但是, 在进行图像边缘检测时, 不能直接使用一阶和二阶的偏导数与图像进行卷积运算。这是因为, 偏导数都带有方向性, 只能处理某特定方向的边缘, 而对其它情况无能为力。回顾我们在前面使用的方法, 可能的一种处理手段是求沿着梯度方向的二维导数, 但其为非线性, 计算较为复杂。Marr 提出用 Laplacian 算子来代替, 这样得到边缘检测算子如下:

$$\nabla g(x, y) = \nabla^2(G(x, y, \sigma) * f(x, y)) = (\nabla^2 G(x, y, \sigma)) * f(x, y) \quad (4.17)$$

求  $\nabla^2 g(x, y)$  的零交叉点就可得到图像边缘点。在上式中

$\nabla^2 G(x, y, \sigma)$  称为 LOG (Laplacian of Gaussian) 滤波器<sup>[28]</sup>。由下式给出:

$$\nabla^2 G(x, y, \sigma) = \frac{\partial G}{\partial x^2} + \frac{\partial G}{\partial y^2} = \frac{1}{\pi\sigma^4} \left(\frac{x^2 + y^2 - 1}{2\sigma^2}\right) \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (4.18)$$

Marr 算法检测图像边缘的步骤即是用 LOG 滤波器与图像卷积, 然后根据零交叉点位置来确定图像中边缘的位置。在具体实现图像  $f(x, y)$  与  $\nabla^2 G$  之间的卷积时应取一个  $N \times N$  的窗口, 由于  $\nabla^2 G$  有无限尾长, 因此所用的窗口不能太小, 以免过分截尾。

利用 Marr 算法对 blood 图像检测边缘的效果如图 4.2 所示:

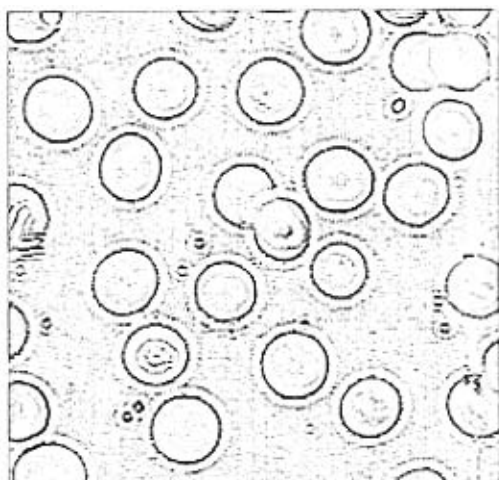


图 4.2 Marr 算法边缘检测效果图

## § 4.2 边缘细化

边缘细化的目的是将有一定宽度的边缘变窄,并保持其形状的拓扑结构不变,可以大大减少边缘的数据量。这里主要介绍 Hilditch 细化算法和 Sobel 细化算法。

Hilditch 细化算法适用于输入图像为 0 和 1 的二值图像。像素值为 1 的区域是需要细化的部分,像素值为 0 的区域是背景。Hilditch 细化算法可描述如下:

设  $p$  为被检测的像素,  $f(p)$  为像素  $p$  的灰度值,  $n_i (i=1,2,\dots,8)$  为  $p$  的 8 邻域像素,  $n_i$  的位置如图所示。

$n_4$	$n_3$	$n_2$
$n_5$	$p$	$n_1$
$n_6$	$n_7$	$n_8$

图 4.3 像素  $p$  的 8 邻域

设集合  $I=\{1\}$  表示需要细化的像素子集,集合  $N=\{g \mid g-m \leq 0\}$  表示背景像素子集,集合  $R=\{-m\}$  表示在第  $m$  次减薄时,  $I$  中被减掉的像素。

图像细化的减薄条件为：

$$(1) f(p) \in I$$

$$(2) U(p) \geq 1, \text{ 其中 } U(p) = a_1 + a_3 + a_5 + a_7, \text{ 这里 } a_i \text{ 为}$$

$$a_i = \begin{cases} 1 & f(n_i) \in N \\ 0 & \text{其它} \end{cases}$$

$$(3) V(p) \geq 2, \text{ 其中 } V(p) = \sum_{i=1}^8 (1 - a_i)$$

$$(4) W(p) \geq 1, \text{ 其中 } W(p) = \sum_{i=1}^8 c_i, \text{ 这里 } c_i \text{ 为}$$

$$c_i = \begin{cases} 1 & f(n_i) \in I \\ 0 & \text{其它} \end{cases}$$

$$(5) x(p) = 1, \text{ 其中 } x(p) = \sum_{i=1}^4 b_i, \text{ 这里 } b_i \text{ 为}$$

$$b_i = \begin{cases} 1 & f(n_{2i-1}) \in N \text{ 并且 } f(n_{2i}) \in I \cup R \text{ 或 } f(n_{2i+1}) \in I \cup R \\ 0 & \text{其它} \end{cases}$$

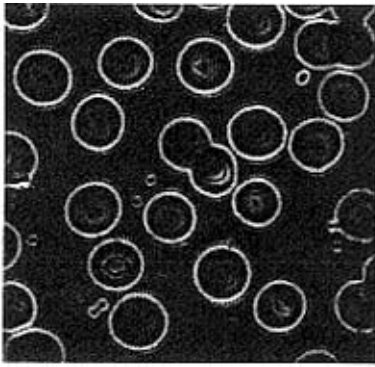
(6)  $f(n_i) \notin R$  或  $x_i(p) = 1, (i = 3, 5)$ , 其中  $x_i(p)$  表示对  $p$  的第  $i$  个邻域像素的  $x(p)$ 。

Sobel 细化算法适用于边缘检测后的边缘图像，Sobel 细化算法可描述如下：

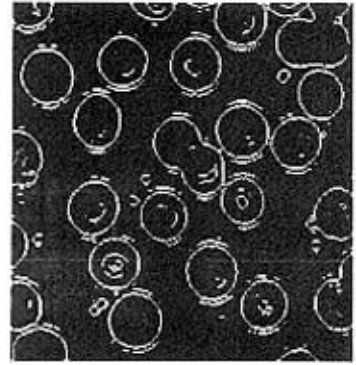
对边缘图作 Sobel 边缘检测可得到两侧的边界，边缘图像与对其进行 Sobel 处理后的图像相减可使波形更加尖锐，使边缘得到增强。这样即提高了定位的精度，又得到用其他方法难以得到的模糊边缘和微弱边缘。

有些图像边缘比较模糊，边界区域的过渡极为平坦，用一次 Sobel 细化还不能达到要求，这时可重复 2~3 次。由于细化处理会使信号强度减弱，已达到细化要求的部分在后面的细化过程中使其保留原来结果，不需要再做处理。

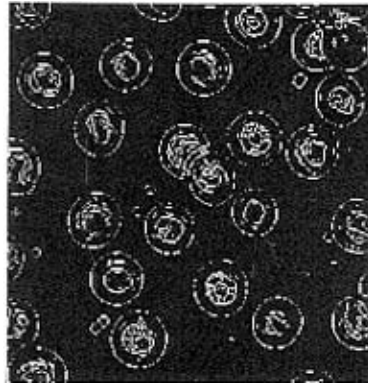
下面是 blood 边缘图及对其进行边缘细化后的结果。



(a)原边缘图



(b)Hilditch 算法细化后边缘图



(c)Sobel 算法细化后边缘图

图 4.3 边缘细化效果图

### § 4.3 特征点的提取

两眼立体视觉中左右眼图象的对应特征计算是立体视觉中很重要的问题。为图像上的每个象素找出对应点通常是不实际的。例如一幅图像上灰度相同的一个区域中的点可能与另一个图像上相应区域中的许多点对应。提取特征点就是为了解决这一问题从而实现唯一的匹配。

如果对左右立体图象对的所有点都进行相关计算，其计算量将会同基于面积的匹配一样很大。为了提高处理速度，人们进行了大量研究。

其中，Ohta<sup>[29]</sup>和 Llogd<sup>[30]</sup>提出利用边界方向的对应特征性质减少对应候选的数目，然后通过加局部约束和重复使用概率松弛法提高对应计算的可信度。但这种方法在选择递推计算次数的评价函数时遇到困难。与此同时奥富和三浦分别提出利用不同观察点的两眼立体视数据的融合消除 3D 视觉信号中的模糊性。

特征点的提取与它周围区域灰度变化有关。如果一个点的周围区域

灰度级有显著变化, 则有可能将此点作为特征点。找特征点实际上是对图像进行微分运算, 对数字图像而言, 微分运算可用差分来近似, 设原图像为  $f(i,j)$ , 特征图像为  $g(i,j)$ , 则可选择下列三个定义式之一来确定微分值  $g(i,j)$ 。

$$g(i,j) = \sqrt{\Delta_i^2 + \Delta_j^2} \quad (4.19)$$

$$g(i,j) = |\Delta_i| + |\Delta_j| \quad (4.20)$$

$$g(i,j) = \max\{|\Delta_i|, |\Delta_j|\} \quad (4.21)$$

$$\text{其中: } \Delta_i = \frac{\partial f}{\partial i} = g_x, \Delta_j = \frac{\partial f}{\partial j} = g_y$$

第(4.19)式为原定义, 由于要计算开方比较费时, (4.20), (4.21)式是其简化形式, 一般考虑计算量的原因常采用(4.20)或(4.21)。通常方法是用一个对灰度变化敏感的算子作用到图像上, 来提取特征点。常用的三种微分算子: Prewitt 算子, Sobel 算子, Roberts 算子。

特征图像  $g(i,j)$ , 可通过求算子  $g_x$  来求, 即:

$$g(i,j) = \begin{cases} f(i,j) & |g_x| \geq G \\ 0 & |g_x| \leq G \end{cases} \quad (4.22)$$

其中  $G$  为固定的阈值。通过改变  $G$  的大小, 可以决定提取特征点的多少。

提取特征点的另一方法是: 先求得  $g_x, g_y$ , 按(4.20)或(4.21)式得到  $g(i,j)$  的特征图像 (梯度大小图像), 而从  $h(i,j) = g_x / g_y$  得到另一特征图像 (梯度方向图像), 将左图像与右图像相匹配, 把满足(4.23)、(4.24)条件的点作为候选匹配点。

$$|g_l(i,j) - g_r(i,j)| \leq T_1 \quad (4.23)$$

$$|h_l(i,j) - h_r(i,j)| \leq T_2 \quad (4.24)$$

其中  $T_1, T_2$  为固定的阈值。通过改变  $T_1, T_2$  的大小, 可以决定提取特征点的多少。

在进行特征点提取时,一般方法是对二维图像中的每个像素作水平、垂直及对角4个方向的相关运算来确定特征点,计算量大<sup>[31]</sup>。为此,文献[31]提出了只需对二维图像中的像素点作水平(或垂直)方向上的灰度差值运算的二维图像特征点快速提取算法。采用文献[31]对图像进行特征点提取,存在合理选择阈值问题,但在文献[31]中,对阈值的选取没有作详细的讨论。在机器视觉系统中,由于视觉系统的自主性能要求,因此必须进行自动阈值选择。

图像中的特征点是某区域中具有较大灰度变化的点,特征点所有方向上的灰度方差通常比较大。典型的特征点检测算法有Barnard特征点选择算法和Harris角点检测算法。算法首先将大小为 $M \times N$ 的图像分为 $p \times q$ 个 $m \times n$ 的窗口 $W_{m \times n}$ ,从每个小窗口中按要求选取一个特征点,最后对 $p \times q$ 个特征点进行选择,剔除特征不明显的点,或从中留下指定的 $K$ 个特征点<sup>[32]</sup>。

在每个 $m \times n$ 窗口中,求出每个像素四个不同方向上与相邻像素灰度差的平方和。例如水平方向:

$$S_h(i, j) = [g(i, j) - g(i, j - 1)]^2 + [g(i, j) - g(i, j + 1)]^2 \quad (4.25)$$

其他还有:垂直方向 $S_v(i, j)$ ,左对角方向 $S_{ld}(i, j)$ ,右对角方向 $S_{rd}(i, j)$ 。

该点处方向平方和最小值为:

$$O(i, j) = \min\{S_h(i, j), S_v(i, j), S_{ld}(i, j), S_{rd}(i, j)\} \quad (4.26)$$

在 $W_{m \times n}$ 小窗口内查找 $O(i, j)$ 最大值点: $O(I, J) = \max_{(i, j) \in W_{m \times n}} \{O(i, j)\}$ ,得到的特征点 $(I, J)$ 在四个方向上的灰度方差最小值是窗口内最大的,即该点的各个方向上的灰度变化最显著。移动 $W_{m \times n}$ 的位置,不重复地对未覆盖区域继续处理,直到把整幅图像全部覆盖处理完毕,全部得到 $p \times q$ 个特征点。将所有的特征点按大小顺序排列,保留前 $K$ 个特征点,得到特征点图。该特征点检测算法的优点是可根据原始图像的特点通过改变窗口 $W_{m \times n}$ 的大小控制特征点的选择。



下面是blood边缘细化图进行特征点提取后的结果。

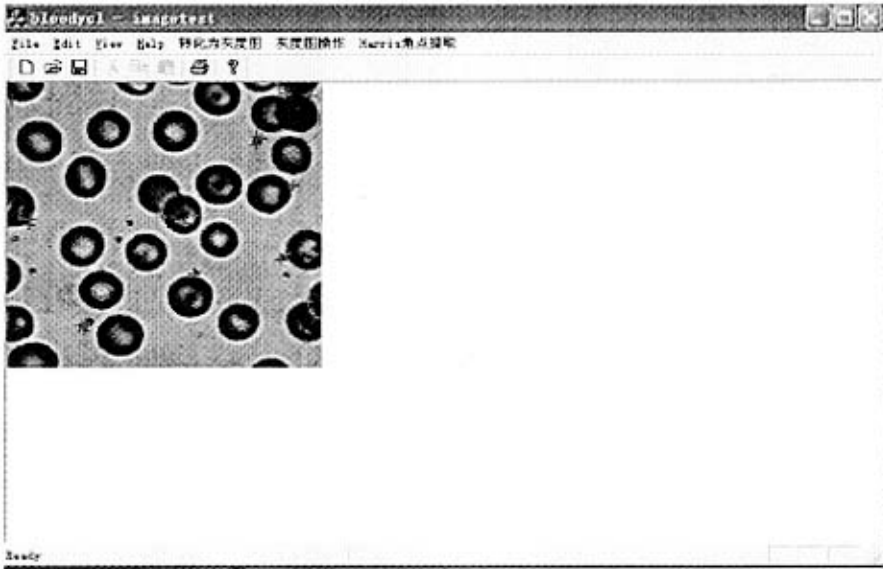


图 4.3 特征点提取效果图

#### § 4.4 小结

对灰度图像进行特征提取等预处理过程是进行图像配准的前提条件。图像的特征是多种多样的，这就决定了特征提取方式的多样性。对于各种特定条件下的应用而言，没有普遍适用的特征提取方法。特征从小到人主要有：点状特征、线状特征、区域特征等，根据特征的不同，对不同类型的特征特性和提取方法进行了全面的研究和比较。特征提取是为了得到匹配所需的图像特征。本章研究了基于微分算子的边缘特征检测方法，用Visual C++编程实现了多种微分算子检测算法，并给出了边缘检测结果，可根据实际情况选用不同的微分算子。

## 第五章 立体匹配

机器人立体视觉是指通过分析和综合从不同视点捕获的两幅或多幅图像推得三维景物的几何形体结构。从信息处理的角度看,一个机器人双目立体视觉系统必须顺序解决如下两个问题:立体匹配和三维重建。三维重建的先决条件是建立两幅图像中图元的对应关系,即立体匹配。对于实际的立体图像对,求解其对应问题极具挑战性,可以说是立体视觉中最困难的一步。

### § 5.1 立体匹配

立体匹配是寻找同一空间景物在不同视点下投影图像中像素间的一一对应关系。与普通的图像模板匹配不同,立体匹配是在两幅或多幅存在视点差异、几何和灰度畸变以及噪声干扰的图像对之间进行的,不存在任何标准模板。当空间三维场景被投影为二维图像时,受场景中诸多因素,如光照条件、景物几何形状和物理特性、噪声干扰和畸变以及摄像机特性等的影响,同一景物在不同视点下的图像会有很大不同,要准确地对包含了如此多不利因素的图像进行无歧义的匹配。是颇为困难的。

目前,对立体匹配的研究可归纳为两个方向:一个方向是从理解人类视觉的立体融合机制出发,试图建立通用的人类双目视觉计算模型。尽管已取得一些研究成果,但是就目前的研究而言,要构造出与人类视觉一样灵活的机器人视觉系统还有待于进一步的研究。另一个方向是从实际应用和要求出发,致力于建立适用于专门领域的和面向对象的实用立体视觉系统,通过强调场景和任务的约束等手段来降低视觉处理问题的难度,从而增加系统的实用性。

立体匹配方法的有效性有赖于三个问题的解决,即选择正确的匹配特征、寻找相应的匹配准则以及建立能够正确匹配所选特征的稳定算法。立体匹配的研究基本上都是围绕这三个方面的工作进行的,并已出现了大量各具特色的匹配方法。但是由于立体匹配涉及的问题太多,至今仍未得到很好的解决,特别是在复杂场景中,如何提高算法的去歧义匹配和抗干扰能力,如何降低计算的运算量,仍需要进行更深入的探索和研究。

### § 5.2 立体匹配的内容

对于任何一种立体匹配方法的设计,都必须解决以下三个方面问题:

(1)基元选择：选择适当的图像特征作为匹配基元，如点、直线、区域、相位等；

(2)匹配准则：将关于物理世界的某些固有特性表示为匹配所必须遵循的若干规则，使匹配结果能真实反映景物的本来面目；

(3)算法结构：利用适当的数学方法设计能正确匹配所选基元的稳定算法。

### § 5.2.1 选择匹配基元

匹配基元是指用以进行匹配的图像特征。由于立体成像的视点差异以及噪声干扰等因素的影响，要对图像中所有的像素都进行无歧义的匹配是很困难的，为此，应该选择能表征景物属性的特征作为匹配基元。

目前，常用的匹配基元主要有：点状特征、线状特征和区域特征等。其中，点状特征是最基本、最简单的特征基元，区域基元具有最好的全局属性，线状基元(包括直线和曲线)介于二者之间。一般来讲，点状特征具有定位准确、检测和描述容易以及重建精度高的优点，但它所含图像信息较少，在图像中的数目较多，因而在匹配时需要较强的约束准则和匹配策略，以克服歧义匹配和提高运算效率。线状基元和区域基元则含有更丰富的图像信息，在图像中的数目较少，易于实现快速匹配。但它们在特征提取和描述时比较困难，需要进行复杂的预处理，而且定位精度较差。

匹配基元类型的选择取决于匹配图像对本身的属性特点和应用要求。一般地，对于包含有大量非规则形状和高度突变的场景，如形状不规则、高度起伏变化大的山区、材料端口等，应选用点状特征作为匹配基元，因为在这样的场景中，要精确提取和描述诸如线段、区域等特征是非常困难的。由于景物结构的不规则性，其特征点的分布十分离散，此时若提取线状特征，用离散分布的特征点去近似连续线段，会引入大量的形状误差，而且不正确的连接还会造成重建表面的结构错误。对于具有规则结构的场景，如果直线或区域的特征提取和描述比较容易，且误差较小时，应该选用直线或区域特征，以便实现快速匹配。因此，在选择匹配特征时，应综合考虑各种因素，根据不同的景物特点和应用要求选择最合理的匹配特征。

除这些直观的匹配基元外，在不同的匹配算法中，还有其它的匹配基元，如Marr提出的零交叉匹配基元以及在此基础上Maynew<sup>[33]</sup>提出的卷积峰匹配基元，图像的相位基元等。

无论从视觉生理的角度还是在实际应用中，立体视觉匹配问题的解决很大程度上都取决于合理地选择匹配基元。良好的匹配基元应具有性

能稳定、抗噪声能力强、易于检测和描述的能力<sup>[34]</sup>。

### § 5.2.2 匹配准则

匹配准则是根据所选匹配基元将现实物理世界的某些固有属性表示成匹配所必须遵循的若干规则,用以提高系统的去歧义匹配能力和计算效率。Marr提出的唯一性约束、相容性约束和连续性约束被认为是匹配控制中最一般、最基本的物理约束<sup>[35]</sup>。在这三个基本约束的基础上,针对处理景物的特点和应用要求,通过附加某些先验知识能够引伸出一些特定的匹配控制准则。下面简单描述一些常用的匹配准则。

**唯一性约束:**一般情况下,一幅图像上的每一个特征点最多只能与另一幅图像上的一个特征点对应。

**连续性约束:**物体表面一般都是光滑的,因此物体表面上各点在图像上的投影是连续的,其视差也是连续的。但是,在物体边界处,比如边界两侧的两个点,连续性约束并不成立。

**相容性约束:**物体表面上一点在两幅或多幅图像上的投影在某些物体度量上(如灰度、灰度梯度变化等)或几何形状上具有相似性。比如空间某一个多面体的顶点,它在图像中的投影应是某一多边形的顶点。

**极线约束:**一幅图像上的任一点,在另一幅图像的对应点只可能位于一条特定的被称为极线的直线上。这个约束极大地降低了待验证的可能匹配点的数量,把一个点在另一幅图像上可能匹配点的分布从二维降到了一维。

**顺序一致性约束:**一幅图像的一条极线对应着另一幅图像中的一条极线,而且它们上面的对应点的排列顺序是不变的。但是,如果视点的方位变化很大,这个约束条件可能不被满足。

### § 5.2.3 算法结构

实现匹配的算法结构是和匹配基元的选择以及匹配准则紧密相关的,一般应该兼顾有效性和计算量<sup>[36] [37]</sup>。立体匹配实质上是在匹配准则下的最佳搜索问题,许多数学中的最优化技术都可应用于立体匹配,如动态规划方法、松弛法以及遗传算法等<sup>[38]</sup>。

## § 5.3 立体匹配方法

根据匹配基元和方式的不同,目前的立体匹配算法基本上可分为两

类：即基于区域相关 (Area-based) 的匹配和基于特征 (Feature-based) 的匹配<sup>[39]</sup>。

### § 5.3.1 基于区域相关的匹配

自四十年代相关技术被应用于航空立体测量以来，相关匹配方法一直在立体视觉的研究和应用中起着非常重要的作用，这是目前最简单、最易于用硬件实现的匹配方法。其基本思想是以统计的观点将图像看作二维信号，采用统计相关的方法，利用图像上局部区域的相关性来确定两个对应点间的匹配<sup>[40]</sup>。相关法的优点是：直接对图像像素进行匹配，匹配结果不受特征检测精度和密度的影响，可得到较高的定位精度。但是，这种方法过分依赖图像灰度统计特性，使得匹配结果对景物表面结构、光照反射和成像几何十分敏感，而且区域灰度相关的假设意味着相关窗口内的像素都具有相同的视差，这对含有深度不连续的区域是不成立的。因此，当空间景物表面缺乏足够的纹理细节及成像失真较大(如视点间距离过大)时，相关匹配法难以应用。

在基于区域相关的算法中，要匹配的元素是固定尺寸的图像窗口，相似准则是两幅图像中窗口间的相关性度量。当搜索区域中的元素使相似性准则最大化时，则认为元素是匹配的。具体描述如下：

输入两幅立体图像对  $I_l$  和  $I_r$ ，设  $P_l$  和  $P_r$  分别为这两幅图像中的像素点， $(2W+1)$  为相关窗口的宽度， $R(P_l)$  是  $I_l$  中与  $P_l$  相关的搜索区域，则  $\psi(u, v)$  是两个像素值  $u, v$  的相关函数<sup>[41]</sup>。对于  $I_l$  中的每一个像素， $P_l = [i, j]$ ，则：

(1) 对于每个区域  $d = [d_1, d_2]^T$ ，计算

$$C(d) = \sum_{k=-W}^W \sum_{l=-W}^W \psi(I_l(i+k, j+l), I_r(i+k-d_1, j+l-d_2)) \quad (5.1)$$

(2)  $P_l$  的视差就是在  $R(P_l)$  中使  $C(d)$  为最大值的矢量  $d$

$$\bar{d} = \arg \max_{d \in R} \{C(d)\} \quad (5.2)$$

输出结果是对应  $I_i$  中每一个像素点的视差的数组，即视差图。

常用的相关函数  $\psi = \psi(u, v)$  有两种形式： $\psi(u, v) = uv$  和  $\psi(u, v) = (u - v)^2$ ，前者称为交叉相关，后者称为块匹配。后者比前者的匹配效果要好，因为它不象交叉相关那样受区域内含有很大或者很小的灰度值的影响。Rosenholm<sup>[42]</sup>等提出在基于区域相关的匹配算法中，将函数  $\psi$  的值做成LUT(Look up table)，可大大加快算法的速度。

基于区域相关的匹配中有两个主要问题：一是适当地选取  $W$  和  $R$ ；一是恰当地寻找相关准则。窗口宽度  $(2W+1)$  的选择取决于从要处理图像中提取最重要的空间属性的能力。如果两个摄像机的几何位置未知，则搜索区域  $R(P_1)$  的初始位置和大小选择是很困难，但是根据极线约束，可以简化为在一条线段即极线上的搜索<sup>[43]</sup>。

### § 5.3.2 基于特征的匹配

与区域相关法不同，基于特征的匹配方法是有选择地匹配能表示景物自身特性的特征，通过更多地强调空间景物的结构信息来解决匹配歧义性问题。基于特征的匹配方法将匹配的搜索范围限制在一系列稀疏的特征上。利用特征间的距离作为度量手段，具有最小距离的特征对就是最相近的特征对，也就是匹配元素。特征间的距离度量有最大最小距离、欧氏距离等。大多数方法都通过增加约束来减少对每一个特征可能的对应特征数量。常用的有极线约束、唯一性约束及连续性约束等。

关于特征的选取，首先应使选取的特征对应可视景物的一定特征，并且这种景物特征能够在两幅图像中产生相似的结果；其次，所选取的特征不能太稀疏，否则，会给三维内插重建造成困难；另外，对所处理的图像来说，所选取的特征的分布还应该能够尽可能地避免误匹配的产生。常用的匹配特征有图像中的边缘点、角点等灰度不连续点和边缘直线等<sup>[44]</sup>。

在实际应用中，完成两幅立体图像对的匹配计算是相当耗时的。与人类能够瞬时完成双眼看到的两幅图像的对应匹配相比，机器人双目立体视觉仍显得相当笨拙，不仅速度太慢，而且匹配精度也不太高，这些

都需要进一步研究。一种思路是通过减小匹配的搜索空间来减少计算量；另一种是减少匹配中存在的歧义。已出现一些改进的方法，如基于带有全局性质的基元匹配和利用三个摄像机减少匹配的二义性<sup>[45]</sup>等。

立体匹配作为机器人立体视觉的核心，在理论上和技术上仍存在很多问题。例如，如何选择合理的匹配特征，以克服匹配准确性与恢复视差全面性之间的矛盾；如何选择有效的匹配准则和算法结构，以解决存在严重灰度失真、几何畸变(透视、旋转、缩放等)、噪声干扰、特殊结构(平坦区域、重复相似结构等)及遮挡景物对匹配的影响；如何建立更有效的图像表达形式和立体视觉模型，以便更充分地反映景物的本质属性，为匹配提供更多的约束信息，降低立体匹配的难度等<sup>[46]</sup>。

## § 5.4 基于边缘检测的分阶段立体匹配方法

边缘特征是最常见也是最重要的特征。这里，我们提出一种以边缘点为匹配基元，采用双阈值判决的灰度相关的分阶段立体匹配方法。该方法适用于双目立体视觉系统或者视点位置相差不太大的立体视觉系统<sup>[47]</sup>。在第一阶段，匹配的目的在于选出最佳匹配点或候选匹配点，在这一阶段中，如满足最佳匹配条件，则不须进行第二阶段匹配；如不满足最佳匹配条件，则选出若干候选匹配点，再进行第二阶段匹配测试，以选出其中的最佳匹配点。采用米字形窗口代替矩形窗口作为相关窗口，大大减少了计算量；采用双阈值判决，不仅提高了匹配的可信度，而且为第二阶段的匹配提供了候选匹配点。在第二阶段，基于视差相容性原则，通过检测其支持邻域的匹配相容以解决候选匹配的歧义性问题，输出唯一匹配，并且尽可能地消除误差。

设同一场景的两幅图像对分别记为  $I_1$  和  $I_2$ ，其中边缘特征点的集合分别为  $\{I_1^1\}$  和  $\{I_2^2\}$ ，相应的灰度值记为  $g_1^1$  和  $g_2^2$ ，以特征点为中心的相关邻域窗口记为  $W$ ，搜索范围记为  $W_1$ ，图像  $I_1$  和  $I_2$  中的边缘特征点的坐标分别记为  $(x_1, y_1)$  和  $(x_2, y_2)$ ，对  $I_1$  中坐标为  $(x_1, y_1)$  的边缘点  $I_1^1$ ，在  $I_2$  中搜索其匹配点的算法如下：

(1) 第一阶段：通过双阈值相关匹配选择最佳匹配点和候选匹配点。

①确定搜索范围 $W_1$ 。 $W_1$ 是以匹配特征点坐标 $(x_1, y_1)$ 为中心，宽度为 $(2H_{\max} + 1)$ ，高度为 $(2V_{\max} + 1)$ 的矩形区域，其中， $H_{\max}$ 为最大水平视差， $V_{\max}$ 为最大垂直视差。则可能匹配点 $I_j^2$ 的坐标 $(x_2, y_2)$ 应满足<sup>[48]</sup>

$$\{(x_2, y_2) | x_1 - H_{\max} \leq x_2 \leq x_1 + H_{\max}, |y_1 - V_{\max} \leq y_2 \leq y_1 + V_{\max}\} \quad (5.3)$$

②对于每组可能匹配，采用如下归一化的相关函数计算其灰度相关值 $R(I_i^1, J_j^2)$ 。取灰度相关函数为：

$$R(I_i^1, J_j^2) = \frac{\sum_{K \in W} (g_k^1 - \bar{g}^1)(g_k^2 - \bar{g}^2)}{\sigma_1 \sigma_2} \quad (5.4)$$

其中， $g_k^1$ 和 $g_k^2$ 为相关窗口中像素的灰度值， $\bar{g}^1$ 和 $\bar{g}^2$ 为相关窗口中的灰度平均值， $\sigma_1, \sigma_2$ 分别为局部灰度方差的平方根，

$$\bar{g}^1 = \sum_{K \in W} g_k^1 / N, \quad \bar{g}^2 = \sum_{K \in W} g_k^2 / N \quad (5.5)$$

$$\sigma_1^2 = \sum_{K \in W} (g_k^1 - \bar{g}^1)^2 / N, \quad \sigma_2^2 = \sum_{K \in W} (g_k^2 - \bar{g}^2)^2 / N \quad (5.6)$$

相关函数的匹配性能和计算效率与相关窗口的大小有关。本文选取米字条形窗口代替常用的矩形窗口，作为相关窗口，如图5.1所示。在窗口内的像素坐标如下<sup>[49]</sup>：

- 水平方向：(i-k, j)到(i+k, j)；
- 垂直方向：(i, j-k)到(i, j+k)；
- 正对角线方向：(i-k, j-k)到(i+k, j+k)；
- 对角线方向：(i-k, j+k)到(i+k, j-k)。



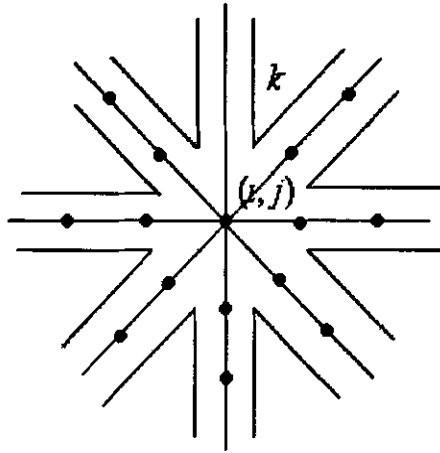


图5.1 米字型窗口

可见，米字条形窗口比矩形窗口极大地减少了运算量，运算速度更快，有利于建立正确匹配。对于  $(2n+1) \times (2n+1)$  的矩形窗口，需要计算的窗门内的像素个数为  $(2n+1)^2$ ，而对于相同大小的米字条形窗口，则只需计算  $(8n+1)$  个像素在具有相同计算点数的情况下，米字条形窗口将比矩形窗口覆盖更大的范围，包含更多的邻域信息，如当窗口大小为  $13 \times 13$  时，米字条形窗口的像素点数为 49，而具有相同点数的矩形窗口大小则为  $7 \times 7$ 。

③对相关匹配值  $R(I, J)$  做阈值处理，得到最佳匹配点和候选匹配点集。

在相关匹配中，一般取具有最大相关值的点作为最佳匹配点，但是往往存在因某些细微结构非常相似，并不存在明显峰值的情形，此时容易出现歧义匹配，两级阈值判断的方法能够很好地解决这个问题。

$$R_{\max}(I_i^1, I_p^2) = \max\left\{R(I_i^1, I_j^2)\right\} \quad (5.7)$$

$R_{\text{sec}}(I_i^1, I_q^2)$  为特征点与其可能匹配点集中相关值的次大值。

设最大值的阈值为  $T_1$ ，次大值的阈值为  $T_2$ ，若  $R_{\max}(I_i^1, I_p^2) > T_1$ ，且  $R_{\text{sec}}(I_i^1, I_q^2) < T_2$ ，则  $I_p^2$  为  $I_i^1$  的最佳匹配，不必再在第二阶段进行唯一匹配的操作；否则，若  $R_{\text{sec}}(I_i^1, I_q^2) > T_2$ ，则  $I_j^2$  为  $I_i^1$  的候选匹配点，所有满

足此条件的点  $I_j^2$  构成了  $I_i^1$  的候选匹配点集  $L_1(I_i^1)$ ，须在下一阶段进一步匹配。这种两种阈值处理的方法不仅提高了匹配的可信度，而且为第二阶段的处理提供了候选点。

匹配性能与阈值  $T_1$  和  $T_2$  的选择有直接关系。阈值  $T_1$  取值越低，意味着相关条件越弱，匹配的可信度就越低。如果  $T_1$  取值太高，则会丢失一些本可以在第一阶段就能够匹配的点，从而加重第二阶段匹配的负担；但如果  $T_1$  取值太低，会降低匹配的可靠度。阈值  $T_2$  的作用主要是为第二阶段的处理提供候选匹配点，其取值应适宜。若取值太高，就失去了设置该阈值的意义；但若取值太小，又会增加下一阶段匹配的搜索空间，降低计算效率。

④按照上面的步骤，对图像  $I_2$  中的特征点  $I_j^2$  在图像  $I_1$  中求得其最佳匹配和候选匹配点集  $L_2(I_j^2)$ 。若  $I_i^1$  和  $I_j^2$  互为最佳匹配，则认为  $I_j^2$  为  $I_i^1$  的最佳匹配。

因此，第一阶段的输出结果为一些最佳匹配点和候选匹配点集  $L_1(I_i^1)$  和  $L_2(I_j^2)$ 。

(2) 第二阶段：利用视差相容性准则对候选匹配点及其邻域进行最佳匹配。

这里，采用邻域视差支持的迭代算法。通过计算邻域内各对匹配间的相容程度，赋予候选匹配一个匹配的评价值，通过该值可得到在视差相容性约束下的最佳匹配。

① 计算候选匹配与其邻域匹配间的支持。

设  $(I_i^1, I_j^2)$  为一候选匹配， $W(I_i^1)$  为  $I_i^1$  一个边缘邻域，称为视差支持邻域，设  $I_m^1 \in W(I_i^1)$ ， $I_m^1$  在  $I_2$  中的匹配点为  $I_n^2$ 。则匹配  $(I_i^1, I_j^2)$  与匹配  $(I_m^1, I_n^2)$  的相互支持为

$$C(d_y, d_{mn}) = \frac{c_1}{D(d_y, d_{mn}) + c_2} \quad (5.8)$$

其中， $d_y$  和  $d_{mn}$  分别为匹配  $(I_i^1, I_j^2)$  与  $(I_m^1, I_n^2)$  的视差， $D(d_y, d_{mn}) = |d_y - d_{mn}|$  为视差之差， $c_1, c_2$  为正常数，满足  $c_1 + c_2 = 1$ 。

② 计算候选匹配的评价值。定义两对匹配的相容系数为<sup>[50]</sup>

$$r[(I_i^1, I_j^2), (I_m^1, I_n^2)] = \begin{cases} 2 & \text{当}(I_m^1, I_n^2)\text{为最佳匹配} \\ 1 & \text{当}(I_m^1, I_n^2)\text{为候选匹配} \\ 0 & \text{当}(I_m^1, I_n^2)\text{不匹配} \end{cases} \quad (5.9)$$

则每对候选匹配的匹配评价值为：

$$S(I_i^1, I_j^2) = \sum_{I_m^1 \in W(I_i^1)} \left\{ \max_{I_n^2 \in L_1(I_m^1)} (r[(I_i^1, I_j^2), (I_m^1, I_n^2)] c(d_y, d_{mn})) \right\} + \sum_{I_n^2 \in W(I_j^2)} \left\{ \max_{I_m^1 \in L_2(I_n^2)} (r[(I_i^1, I_j^2), (I_m^1, I_n^2)] c(d_y, d_{mn})) \right\} \quad (5.10)$$

若一对匹配满足条件：

$$S(I_i^1, I_j^2) > S(I_i^1, I_n^2), \quad \text{对于任意 } I_n^2 \in L_1(I_i^1);$$

$$S(I_i^1, I_j^2) > S(I_m^1, I_j^2), \quad \text{对于任意 } I_m^1 \in L_2(I_j^2);$$

则认为  $(I_i^1, I_j^2)$  为最佳匹配。

③ 对两幅图像的候选匹配点都进行上述两步处理后，将含有被唯一匹配的点的所有可能匹配点去除，再进行上述两步的处理，直至最佳匹配点的数目不变或达到设定的最大迭代次数为止。

## § 5.5 小结

立体匹配是机器人双目立体视觉系统必须解决的问题，是立体视觉的核心问题。立体匹配的有效性有赖于三个问题的解决，即：匹配基元的选取、匹配准则的制定、匹配算法实施。

本章提出了基于区域相关和基于特征两种匹配方法。最后又提出一种以边缘点为匹配基元，采用双阈值判决的分阶段立体匹配方法。第一阶段旨在寻找部分最佳匹配点以及候选匹配集。取米字形窗口代替矩形窗口作为相关窗口，提高了相关运算的速度；采用双阈值判决，不仅解决了灰度相关峰不明显时匹配判决困难的问题，而且能够得到大量

候选匹配点，大大缩小了下一阶段匹配的搜索空间。在第二阶段，利用视差相容性约束，基于邻域视差支持，对第一阶段中得到的候选匹配集进一步匹配，最后得到最佳匹配。

## 第六章 双目图像测距系统

### § 6.1 测距系统结构框图

图6.1为系统硬件配置图，它是由CCD摄像机、图像采集卡、微机系统和监视器组成。其工作原理为：

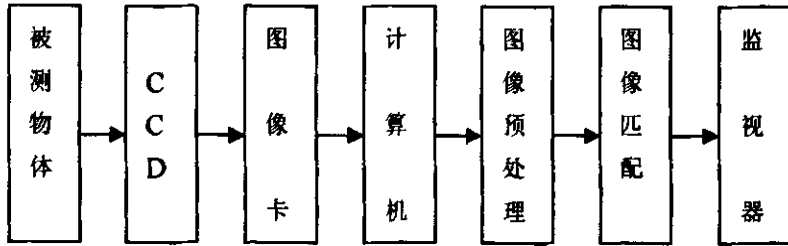


图6.1 测距系统结构框图

CCD摄像机对被测物体表面进行采集，图像卡用于接收摄像机送出的视频信号并将其转变为数字信号送入计算机，在计算机内完成图像处理过程并得出测量结果，最后在监视器(显示器)中将结果显示出来。

### § 6.2 实验方案设计

实验平台为一水平放置的玻璃板，在玻璃板上平铺一张坐标纸，两台摄像机水平放置在实验平台上，下面在坐标纸上选取一点 $O$ 作为空间坐标系的参考原点，建立空间坐标系 $OXYZ$ ，其中坐标纸平面被设定为 $XOZ$ 面，摄像机在拍摄过程中始终保持与 $XOZ$ 面平行。拍摄位置用与 $O$ 的距离 $L$ 表示。被测量的物体是形状规则的立方体方块。理想状态下两台摄像机应当完全平行放置，为了使实验更接近于实际工程测量，将两台摄像机呈一定角度放置，构成距离测量的传感器。如图6.2所示，两台摄像机的光心分别为 $O_1, O_2$ 。A和B分别为两个CCD相机光轴与其成像面的交点，其光轴与 $X$ 轴的夹角均为 $\theta$ 。P为被测点，其在两个CCD相机上的成像点分别为 $P_1$ 和 $P_2$ 。设 $O_1A = O_2B = f$ ， $AP_1 = x_1$  ( $P_1$ 位于A点左面为负，右面为正)， $BP_2 = x_2$  ( $P_2$ 位于B点左面为负，右面为正)， $OO_1 = OO_2 = L$ ，则直线 $O_1P_1$ 在 $XOZ$ 坐标系下的方程为

$$\frac{x+L}{-f \cos \theta + x_1 \sin \theta} = \frac{z}{-f \sin \theta - x_1 \cos \theta} \quad (6.1)$$

直线  $O_2P_2$  在  $XOZ$  坐标系下的方程为

$$\frac{x-L}{f \cos \theta + x_2 \sin \theta} = \frac{z}{-f \sin \theta + x_2 \cos \theta} \quad (6.2)$$

联立式 (6.1) 和 (6.2) 可以得到  $P$  点的  $Z$  坐标为

$$z = \frac{2(f \sin \theta - x_2 \cos \theta)(f \sin \theta + x_1 \cos \theta)L}{2 \sin \theta \cos \theta (f^2 + x_1 x_2) + f(x_2 - x_1)(\sin^2 \theta - \cos^2 \theta)} \quad (6.3)$$

$Z$  即为所求距离。

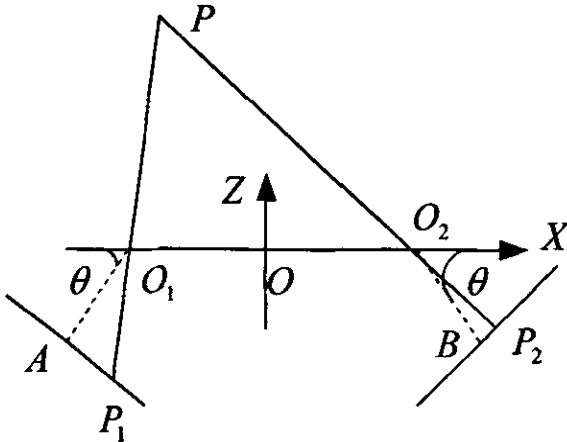


图6.2 测距系统实验方案图

### § 6.3 实验误差分析

在实际测量中, 为了保证两幅图像匹配的精度, 要求两相机光轴与纵轴夹角  $u (u = 90 - \theta)$  相等且值足够小,  $OO_1$  与  $OO_2$  相等, 根据误差合成原理可以得到  $Z$  的系统误差

$$\Delta z = \Delta \theta \frac{\partial Z}{\partial \theta} + \Delta L \frac{\partial Z}{\partial L} + \Delta u \frac{\partial Z}{\partial u} \quad (6.4)$$

由以上的结果可以看出当  $\tan u$  较小时,  $Z$  对  $u$  的导数值比对  $L$  和  $\theta$  的值大, 因此较小的偏差就可以引起较大的误差, 因此在光学系统校正中, 必须

保证 $u$ 能达到足够的精度。

根据误差理论得到测量的总误差是系统误差和对空间物体进行准确的识别。焦距对测量结果的影响主要反应在同名点的匹配上，两个焦距相差较大将使同名点的匹配精度降低，当两个光学系统的焦距相差超出一定范围以后，由于左右摄像机的放大倍数相差很大，这样使像差过大无法进行同名点的识别，从而使整个测量系统无法工作，因此在测量系统中为了提高匹配精度同时降低匹配难度，必须保证两个相机焦距一致。

## § 6. 4小结

本章首先从立体视觉的一般原理推导出了双目图像测距系统中空间点和空间直线的重建公式，接着设计了一种简易的双目图像测距系统的模型。

## 第七章 结论与展望

计算机立体视觉是被动式测距方法中最重要的距离感知技术,它直接模拟了人类视觉处理景物的方式,可以在多种条件下灵活地测量景物的立体信息。其作用是其它计算机视觉方法所不能取代的,对立体视觉的研究,无论从视觉生理的角度还是在工程应用中都具有十分重要的意义。

自计算机立体视觉技术诞生到现在,已经形成了从图像获取到最终的景物可视表面重建的比较完整的体系。本课题尝试从双目立体视觉的距离感知过程中的特征提取和图像匹配这两方面入手,重点对立体视觉中的若干问题如特征提取、立体匹配以及立体视觉测距原理进行了学习和研究。

模仿人眼,采用双目横模型结构的成像系统模型。景物在左右视网膜上成像位置的差异(即视差)反映了客观景物的深度。人类有深度感知,根据获得的视差并进行加工,进而获得景物的深度信息。基于此详细介绍了成像变换原理、成像系统模型和双目立体视觉中的视差计算原理。由于实际条件的约束,未对摄像机标定进行深入的研究。

目前尚没有一种普遍适用的理论可运用于匹配特征的提取,导致了立体视觉中匹配特征的多样性。良好的匹配特征应具有可区分性、不变性、稳定性、唯一性和有效解决歧义匹配的能力。由于特征和特征提取方式的多样性,所以在图像预处理与特征提取部分有重点地对特征提取技术(包括点特征、边缘特征等)进行了论述,并对比性的给出了相应实验结果。希望本阶段的研究和实验工作为下一步的立体图像匹配提供比较坚实的理论基础,对图像匹配起到指导作用。

虽然立体匹配技术经过数十年的研究,已经有了很大的发展。但无论是从视觉生理的角度,还是从实际应用方面来看,现有的立体匹配技术还处在十分不成熟的阶段。这不仅仅涉及到技术上的原因,而是更多地在于人类对自身机理还不十分了解,因此,立体匹配问题的彻底解决还有待少对人类自身视觉机理的深入研究。尽管如此,科技工作者仍坚持不懈地努力对立体匹配技术进行研究,并针对各自的应用目的和要求提出了很多尝试解决该方法的方法。本论文针对图像匹配详细论述了特征基元选择、匹配策略、约束条件以及相似性测度等立体匹配的主要组成部分,并针对不同特点的立体图像对分别提出了基于图像边缘特征的分阶段立体匹配方法和基于区域相关的匹配方法。

双目立体视觉测距系统需要用到摄像机标定所获取的摄像机内、外参数。由于课题研究各种条件的限制,没有涉及到图像匹配后的三维重



建技术的研究和实验工作。由于条件、时间和能力有限，本课题仅对基于双目机器视觉的被动测距的若干问题进行了探讨。希望通过对本课题的研究和学习，能够为将来的学习和工作确立一个明确的研究方向，并为其打下坚实的基础。

## 致谢

本文是在杨进华教授的精心指导、热情帮助下完成的。在两年半的硕士研究生学习期间，杨老师严谨的治学态度、求实的工作作风和高尚的为人品质使我深受教育和启发，并将使我终身受益。在我毕业课题的设计中，导师富于创造性的想法给予我很大帮助，使我能顺利完成课题研究和论文写作，也顺利完成研究生阶段的学业。在本论文即将结稿时，我要向导师表示衷心的感谢。

感谢父母和家人给予我的无私支持和帮助，使我能够专心于课题的研究。

最后，感谢审阅本文并提出指正意见的专家学者！

## 参 考 文 献

- [1] 吴立得. 计算机视觉. 上海: 复旦大学出版社, 1993
- [2] 项海兵, 计算机视觉发展中存在的问题. 新浪潮, 1996 (4): 29~30
- [3] 张福学. 机器人学——智能机器人传感技术. 北京: 电子出版社, 1996
- [4] 吴祥瑞. 机器人技术及应用. 北京: 北京航空航大出版社, 1994
- [5] 周国清, 唐晓芳. 计算机视觉及应用. 计算机用户, 1996(8):3~6
- [6] 邓文怡, 李乃光, 董明利, 张靖渝, 燕必希, 数字摄影测量技术在三维测量中的应用, 光电子激光, 2001, 12(7):697~700
- [7] 程存学, 朱晓昆著. 计算机视觉——低层处理技术. 电子工业出版社, 1993
- [8] Barnard, et al. Computational stereo. ACM Computing Servcegs. 1982, 14:552~572
- [9] Besl P J, Jain R C. Segmentation through variable-order surface fitting. IEEE Trans Pattern Anal Machine Intell, 1988, 10(2):167~192
- [10] Marr D. Vision San Francisco: Freeman, 1982
- [11] Dhond U R, Aggarwal J K. Structure from stereo: A review. IEEE Trans Pattern Anal Machine Intell, 1989, 11(12):113~120
- [12] Marr D, et al. Cooperative computation of stereo disparity. Sci, 1976, 194:283~287
- [13] Baker H H, et al. Depth from edge and intensity based stereo. Proc. 7th IJCAI, 1981:631~636
- [14] Aronid R D. Automated Stereo Perception. Stanford University Tech. Rep., STAN-CS-83-961, 1983
- [15] 郑南宁, 计算机视觉与模式识别, 国防工业出版社, 1998年3月
- [16] 王润生, 图像理解, 国防科技大学出版社, 1995:292~322
- [17] 高文, 陈熙霖, 计算机视觉——算法与系统原理, 清华大学出版社, 广西科学技术出版社, 1999年2月
- [18] 晶辰工作室, 最流行图像格式使用参考手册, 电子工业出版社, 1998年10月
- [19] 吕凤军, 数字图像处理编程入门, 清华大学出版社, 1999年9月
- [20] Kenneth. R. Castleman, 数字图像处理, 电子工业出版社, 1998年9月
- [21] 田捷等, 实用图像分析与处理技术, 电子工业出版社, 1995年1月
- [22] Catte F, Lions P L, Morel J M. Coll T. Image Selective Smoothing and Edge Detection by Nonlinear Diffusion II. Society for Industrial and Applied Mathematics Journal of Numerical Analysis, 1992, 29(3):845~866
- [23] 阮秋琦, 数字图像处理学. 北京: 电子工业出版社, 2001
- [24] 张凯丽, 刘辉, 边缘检测技术的发展研究, 昆明理工大学学报, 2000, 25(5):36~39
- [25] Lung P, wang Y F. Local Scale Controlled Anisotropic Diffusion with Local NoiseEstimate for Image Smoothing and Edge Detection. In: Proc of the 6th

- International Conference on Computer Vision, Bombay, India, 1998, 193~200
- [26] 马艳,张治辉,几种边缘检测算子的比较,工矿自动化,2004年2月,第一期
- [27] Marr D. Vision: A computational investigation into the human representation and processing of visual information, W. H. Freeman and Company, San Francisco. 1982
- [28] Eric W, Grimson L. Computational Experiments with a Feature Based Stereo Algorithm. IEEE Transactions on Pattern Analysis And Machine Intelligence 1998, PAMI-7(1): 17~24
- [29] Ohta Y, et al. Stereo by intra- and inter- scanline search using dynamic programming. IEEE Trans. on PAM I, 1985, 7: 139~154.
- [30] Lloyd S A , et al. A Paraller Binocular Stereo A lgorithm Utilized Dynamic Programm- ing and Relaxation Labelling. CV GIP, 1987, 39:202~225.
- [31] 管业鹏, 童林凤. 二维图像特征点快速提取算法[J]. 中国图像图形学报, 2002, 7A(12): 1296~1301
- [32] 潘晓露, 李一民, 刘辉, 平行视双目测量体系中特征检测方法, 云南工业大学学报, 1999, 15(4): 19~21
- [33] Mayfew J E W, et al. Psychophysical and computational study towards a theory of human stereopsis. A I, 1981, 17: 349~385.
- [34] 李作新, 胡广宁, 立体视觉与立体匹配小型微型计算机系统, 1990, 11(2): 43~50
- [35] Marr D. 著. 姚国正, 等译. 视觉计算理论. 北京: 科学出版社. 1988
- [36] 马颂德, 张正友, 计算机视觉—计算理论与算法基础, 科学出版社, 1998年1月
- [37] D. H. 巴拉德 C. M. 布朗, 计算机视觉, 科学出版社, 1987年9月
- [38] 吴立德, 计算机视觉, 复旦大学出版社, 1993年12月
- [39] 游素亚, 徐光佑. 立体视觉研究的现状与进展. 中国图像图形学报, 1997, 2(1): 15~24
- [40] Takeo Kanade et al, A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment, IEEE Trans. On Pattern Analysis and Machine Intelligence, 1994, 16(9): 920~932
- [41] Suresh B. Marapane et al, Multi-Primitive Hierarchical (MPH) Stereo Analysis, IEEE Trans. On Pattern Analysis and Machine Intelligence, 1994, 16(3): 227~240
- [42] D Rosenholm. Multi-point matching using the least-squares technique for an evaluation of three-dimensional models[J]. photogrammatic Engineering and remote Sensing, 1987, 53(6): 1214~1218
- [43] 丁震, 胡钟山, 唐振民, 杨静宇, 从二维图像中恢复距离信息的体视匹配方法, 南京理工大学学报, 1996, 20(5): 449~452
- [44] 任重, 邵军力, 双目立体视觉的匹配方法研究, 军事通信技术, 2001, 22(2): 45~48
- [45] 吕宏静, 贾云德. 用多基线立体视觉系统恢复稠密深度信息. 北京: 北京理工大学学报, 2000, 20(1): 70~72
- [46] 周东翔, 蔡宣平, 孙茂印, 一种基于特征约束的立体匹配方法, 中国图像图形学报, 2001,

6(7):653~656

- [47] 丁震, 唐振民等, 一种渐近的立体匹配方法, 计算机研究与发展, 1997, 34(10):42~47
- [48] 周骥, 石教英, 赵友兵, 图像特征点匹配的强壮算法, 2002, 14(8):754~757
- [49] 朱晓瑾, 候文永, 识别库函数的特征点匹配法, 上海交通大学学报, 1994, 28(9):34~38
- [50] 边肇祺, 张学工等. 模式识别. 清华大学出版社, 2002. 10

作者: [崔岩](#)  
学位授予单位: [长春理工大学](#)

## 相似文献(10条)

### 1. 学位论文 [王宁玲 立体视觉在非接触三维测量中的应用与研究](#) 2004

计算机立体视觉三维测量是一门新兴技术,在现代工业领域尤其是三维轮廓和体积测量等方面有广阔应用前景。本文系统研究了立体视觉三维测量中摄像机定标、特征点的提取与匹配、图像处理和三维计算等关键技术,提出了三维测量实验系统硬件及软件解决方案,在MATLAB环境下编制、调试了实验系统软件,建成了基于双目立体视觉技术的三维测量实验研究平台。在该平台上开展了摄像机定标及回转体、类回转体的三维测量实验,成功实现了空间立体三维信息重建和体积等关键参数的测取。

### 2. 学位论文 [张家林 基于立体视觉的皮肤表面检测技术的研究](#) 2002

如今,评价皮肤表面的光滑度越来越受到人们的重视,尤其是在化妆品研究领域,科学家们已经建立了几种检测皮肤皱纹、评价皮肤表面光滑度的方法,该文比较详细地介绍了各种皮肤皱纹检测的技术,并且对它们各自的优缺点以及适用范围进行了比较。另外,该文提出了基于立体视觉的皮肤表面检测技术,然后,分别对摄像机定标和立体匹配技术进行了研究。在摄像机定标部分,该文介绍了摄像机定标的基本原理以及定标的具体步骤。在立体分析部分,我们着重介绍了区域匹配技术以及用于全局控制的极线约束方法,并且在实验的基础上,提出了一种基于标准点约束的全局控制策略。与极线约束的控制策略相比,它不受定标精度的影响,而且处理简便。最后,总结全文,对自摄的图像进行了实验,并且分析了实验结果和系统的性能,指出今后需进一步研究和探索的方向。

### 3. 学位论文 [王军 双目立体视觉匹配改进算法及三维重建技术](#) 2005

本文主要就摄像机定标和立体匹配这两方面的内容展开研究。首先根据选择的立体视觉成像系统模型阐述了摄像机定标的基本原理。

其次,对图像匹配问题涉及的内容进行了详细的叙述,其中匹配特征的选取、匹配准则的确定、匹配算法的实现是立体匹配的三个重要步骤。立体视觉中,对于视差不连续场景,例如遮挡区域、物体边缘等,如何获得高性能的图像匹配是一个重要的研究课题。本文提出了一种移植性能很好的改进匹配算法,在对左右图像进行LOG窗口预处理后,用传统的SAD窗口匹配结合改进的子窗口技术,在保持实时性的同时能够有效地提高匹配的质量。它的突出优点是可以通过任何现有匹配算法的补充以增加匹配的精度,并且简单高效。

最后通过双目视觉模型分析,利用空间几何关系,提出了目标点的三维坐标计算表达式。

### 4. 期刊论文 [何春燕,周付根 双目立体视觉系统定标误差分析—中国体视学与图像分析](#)2003,8(4)

摄像机定标是提高双目立体视觉系统三维测量精度的关键之一,如何提高定标的精度对于提高整个系统的测量精度至关重要。本文在深入研究单目摄像机定标的基础上,利用一种基于双目立体视觉的三维测量系统定标模型,对摄像机定标给系统带来的误差进行了评价与分析。实验结果表明,在同一模型下,定标方式的选取对定标精度有较大影响,合理的进行定标可以有效地消除摄像机球面误差对测量精度的影响。

### 5. 学位论文 [石继雨 机器人双目立体视觉技术研究](#) 2003

该论文基于CCD摄像机和个人计算机组成的视觉系统,重点讨论了摄像机参数定标、立体匹配、目标定位、深度图生成等。在摄像机参数定标中,根据坐标关系,分别讨论了不同的摄像机模型,并对传统的定标方法和摄像机自定标做了详细阐述。双视觉立体匹配是寻找同一目标点在两幅图像中的成像位置,匹配特征的选取、匹配准则的确定、匹配算法的实现是立体匹配的三个重要步骤。目标定位是通过双目视觉模型分析,利用空间几何关系,最后提出了目标点的三维坐标计算表达式。深度图生成是深度信息的可视化过程,该文讨论了不同深度图表达形式和深度插值方法。至此构建一个完整的机器人双目立体视觉系统结构。

### 6. 学位论文 [雍超 基于避障的移动机器人立体视觉算法研究](#) 2004

自主移动机器人可以被广泛地应用到军事、生产、生活等许多领域,因此受到了全世界各国的普遍关注。然而,作为机器人技术中的一个关键部分,现有机器人视觉技术还存在很多问题需要改进。作者所在实验室的多个移动机器人研究项目都涉及机器人双目立体视觉的研究。依托这些项目提供的硬件支持,该文对摄像机定标、图像匹配和障碍探测等方面进行了深入细致的基础性研究,主要包括如下几方面的工作:1)开发了一种新的简单而有效的摄像机定标技术,它的精度高,容易操作,速度快,并具有很好的通用性,它可以到任意的计算机视觉系统中;2)实现了一种新的基于多重假定视差的匹配算法,很大程度解决了其他匹配算法中匹配某些像素精度不高的问题,它的另外一个重要特点是可以被移植到许多现有的匹配算法中去,有效提高这些现有算法的匹配性能;3)基于我们开发的摄像机定标技术,利用外极线约束条件对图像进行校正,一方面消除径向失真,另一方面使得立体图像对中的所有对应点都位于同一扫描行上,为图像匹配提供了良好基础;4)该系统不进行精确的三维场景重建,而是提出一种在较为粗糙的水平上恢复场景的策略,把摄像机定标、图像匹配等若干环节带来的累积误差的影响降到最低。同时它又可以节约三维重建的时间,提高系统效率。实验证明这样的策略对于超越机器人来说尤为合适。实验表明,该文介绍的基于立体视觉的障碍探测系统具有良好的表现。该文描述的双目立体视觉算法是一个相对通用的算法,其中的摄像机定标,匹配算法等技术可以移植到任何其他一种移动机器人双目视觉系统中去,这为后续研究提供了良好的基础。

### 7. 学位论文 [肖高逾 计算机立体视觉中三维重建技术的研究](#) 2000

计算机立体视觉的主要研究内容是由多幅二维的平面图像恢复出被摄物体的三维坐标,而其中基于两幅图像的双目立体视觉技术则是一个研究热点。该文即是对计算机双目立体视觉中的三维重建技术进行了探讨,着重研究了在图像匹配结果已知的情况下如何由像对重建物体三维坐标的问题,主要的研究内容包括摄像机定标、被摄物体三维坐标的获取、物体表面插值以及表面数据的表示。该论文的研究受国家“九五”项目的资助。

### 8. 学位论文 [陈华华 视觉导航关键技术研究:立体视觉和路径规划](#) 2005

陆地自主车(Autonomous Land Vehicle,简称ALV)是一种能够在道路和野外连续地、实时地自主运动的智能移动机器人,其研究涉及多个学科的理论与技术,体现了信息科学与人工智能技术的最新成果,具有重大的研究价值和应用价值,受到了世界各国的重视。在移动机器人的各项关键技术中,视觉导航的主要功能是对各种道路场景进行快速识别和理解,从而确定移动机器人的可行驶道路区域,而视觉导航的关键是障碍物的快速检测与识别。路径规划则根据视觉处理获得的信息,在可行区域内规划出一条安全、高效的路线。陆地自主车的行驶过程即立体视觉获取前方的路面环境信息,然后通过智能控制技术使自主车沿规划的路径行走。因此本文旨在研究陆地自主车的立体视觉系统设计和实现,以及在视觉处理获得的信息基础上进行路径规划。摄像机的精确定标是立体视觉研究的基础,本文第二章围绕摄像机定标对各种定标模板及其控制点、摄像机模型的优缺点进行了比较和分析,结合应用环境确定了适合于自主车的定标模板和控制点以及摄像机模型,并讨论了相应的摄像机定标方法,该方法只需要用多个任意角度的平面模板进行定标,操作过程简单,而且定标精度较高,能够较好的适合自主车导航现场环境的要求。此外,立体图像对的外极线校正是降低匹配算法复杂度必不可少的工作,在这一章中首先研究了双目立体视觉校正算法,但是其有一定的局限性,不适用于多基线立体视觉系统,由此局限性出发对此算法进行修改,将其扩展到平行三目立体视觉系统。实验结果表明,该算法正确有效并且可以推广到摄像机数目超过三个的平行多基线立体视觉系统。第三章对立体匹配算法作了深入的研究,研究中突出了算法实时性和鲁棒性要求。对于局域匹配,重点研究了区域法,对其各种相似性测度的准确性以及匹配窗口的大小进行了定性和定量的评估,对影响匹配结果的各种预处理和后处理作了深入的研究,并提出了多分辨率、盒滤波、并行指令、超线程、OpenMP等多种提高匹配实时性的加速方法。对全局匹配方法,研究了比较流行的动态规划方法和Graph cuts方法。对这两种全局匹配方法的准确性作了定性和定性的评估,并对这两种方法存在实时性较差的不足,提出了基于多分辨率的匹配方法,大大减少了匹配时间。立体匹配的结果最终通过三维重建来表现,在本文第四章提出了具有实时性较高的三维重建算法,对该算法结合摄像机定标过程中的模型误差和图像噪声误差、匹配过程中的透视投影变形误差、非校准误差、系统误差、量化误差以及双目视觉系统的误差模型对重建的误差进行了分析,并提出了误差修正算法,通过实例验证了该修正算法可以减少重建误差。在重建算法的基础上,结合自主车GPS/INS提供的车体位置和姿态,对越野环境的地貌进行了全局地图构建,通过实例验证表明构建的全局地图基本符合了真实三维场景。

### 9. 学位论文 [李华明 计算机视觉在城市建筑物测量中的应用](#) 2003

该文针对利用数码相机对城市建筑物进行测量这一崭新课题,运用计算机视觉、数字图像处理、图形图像学及射影几何学的知识,提出了一套可行性方案,并给出了初步的实验结果。首先,简要介绍了项目背景,指出了该文的主要工作,并根据项目及所使用的测量工具——数码相机的特点给出了测量工具的选择方案。其次,针对项目中的摄像机定标的需要,介绍了摄像机定标中的坐标系、线性摄像机模型以及摄像机定标的特点及分类,并给出了一种基于二维棋盘状定标参照物的定标方法和试验结果。再次,对计算机视觉中对获取场景深度信息的方法进行了分类和比较,讨论了双目立体视觉的方案及立体视觉中的极线约束。最后,推荐了利用消隐点和消隐线隐含的三维信息对建筑物进行三维信息恢复的方法,并利用Canny变换和Hough变换的方法给出了建筑物边缘提取的结果,同时介绍了高斯球面法提取消隐点的方法和对全文进行了小结。

### 10. 期刊论文 [周波,聂楠楠,李东辉,ZHOU Bo,NIE Nan-nan,LI Dong-hui 双目摄像机定标的神经网络方法—自动化与](#)

## 仪器仪表2009(1)

研究基于反向传播神经网络的摄像机双目立体视觉定标新方法.传统方法基于三角测量原理技术,会带入成像畸变非线性误差,而这种新方法可以消除非线性因素的影响.该方法利用了BP网络良好的非线性映射能力以及学习、泛化能力,通过采用高精度样本数据训练BP网络,最终建立起立体视觉定标的网络模型.由于不需要考虑视觉模型误差、光学调整误差、广角畸变等因素对视觉检测系统测量精度的影响,因而能够有效地克服常规建模方法的不足,保证了检测系统具有较高的精度.

本文链接: [http://d.g.wanfangdata.com.cn/Thesis\\_Y1041553.aspx](http://d.g.wanfangdata.com.cn/Thesis_Y1041553.aspx)

下载时间: 2010年5月9日