

摘 要

随着农业生产的飞速发展和农业劳动力成本的迅速上升,农业劳动力的匮乏将成为许多发达国家和发展中国家共同面对的问题。目前机器人技术在农业上应用研究越来越成为热门课题,与工业机器人特定的工作环境不同,农业机器人主要在自然场景下工作,要面对更为复杂多变的情况,有更多的问题需要解决。

本研究作为柑橘采摘机器人研究的一部分,利用双目立体视觉技术,研究自然场景下的成熟柑橘的识别和定位方法、障碍物(树枝)的检测方法,为未来开发收获机器人采摘柑橘进行前期视觉系统方面的准备工作,这部分相关的研究也是水果收获机器人实用化的关键。

本研究的主要内容和方法如下:

1. 成熟果实定位

果实定位主要通过识别和匹配等步骤完成。识别的目的是把成熟柑橘从背景中识别出来,为空间定位做准备。本研究利用对 RGB 颜色系统中的色差分量 $2R-G-B$ 值进行迭代,自动寻找阈值的方法分割柑橘图像;将区域分割后的彩色图像转化为二值图像;经过形态学运算消除噪声;对图像进行区域标记,根据区域面积和区域最小外接矩形长宽比设定阈值,去除小块及非类圆形干扰区域;进行区域填充、轮廓提取,并采用优化圆形 Hough 变换拟合出成熟果实图像中的圆心坐标、半径等特征值;然后以单个柑橘为匹配对象,采用基于特征的立体匹配方法,加入柑橘的区域重心、外接矩形尺寸等特征值作为

匹配约束来实现左右两幅图像中对应果实的唯一匹配, 研究结果显示正确匹配率达 80% 以上。最后在对摄像机内外部参数进行了标定的情况下, 求得柑橘的空间三维坐标, 并利用激光测距仪进行了对比验证, 当测量距离小于等于 1.5m 时, 平均误差不超过 1%。

2. 障碍物检测

为保持算法的可执行性, 障碍物检测的步骤和果实定位基本相同, 采用对图像 2R-G-B 和 2G-R-B 色差分量值进行迭代, 并结合图像灰度阈值法能快速有效的分割出图像树枝区域, 通过图像二值化、形态学运算、区域标记、空洞填充提取出图像树枝区域; 通过区域细线化提取树枝骨架, 并进行骨架修剪、恢复遮挡骨架等处理; 然后找出骨架中端点、分支点等特征点并记录它们的连接关系; 最后通过对特征点的立体匹配恢复障碍物的三维信息。试验表明障碍物的正确识别率为 67.3%, 当障碍物实际距离大于 1.5m 时, 识别误差增大。

通过研究, 在成熟水果识别、匹配和定位方面取得了较大进展, 提出了一种果实收获机器人的障碍物检测方法。本文的研究内容对我国开展农业收获机器人视觉识别技术领域的研究具有参考价值, 为进一步的研究打下了基础, 对提高我国农业的国际竞争力有重要的经济意义。

关键词: 识别, 匹配, 圆形 Hough 变换, 空间定位, 障碍物

ABSTRACT

With the rapid development of agricultural production, the cost of agriculture labor force will become more and more costly. In recent years, the agricultural application of robot technique have already become popular issue, because of the shortage of the agriculture labor force both in developed countries and developing countries. Different from industrial robot, which works in particular environment, the agriculture robot mainly works in the natural environment, and the agriculture robot has to face more complicated and uncertain circumstance, thus there are more problems to be resolved.

As a part of research on citrus picking robots, this research used binocular stereo vision to researching on recognition and location mature citrus, obstacle (branches) detection under natural environment. The main contents and methods are as follows:

1. Mature fruit location

The main steps to complete that are image recognition and stereo matching. The purpose of image recognition is recognizing mature citrus region from image to make preparation for location. This research used iterate on 2R-G-B chromatism component in RGB color system to finding threshold automatically to segment original image. Then the segmented images were converted to two-value images and eliminated noise by morphological operation. Region labeling was done and eliminated region which was small or had great differences from circle by defining threshold which based on area and length to width ratio of the smallest circumscribed rectangles of each region. Then region filling and contour extraction operation were been done and used improved Circular Hough Transformation (CHT) to found out circle centre and radius of each region's approaching round. Then used feature-based match and added extra characteristics such as barycenter of orange region in the image, sizes of each region's approaching round to get correct match results of each image pair. Experimental results show that the matching accuracy can reach over 80%. Finally, after calibrated the camera interior and exterior parameters, the 3-d space coordinate of each orange

was been obtained and used the laser range finder to verification and comparison. Results show that the average error ratio is below 1% when the measuring distance is no more than 1.5m.

2. Obstacle detection

To ensure executable of algorithms, the steps of obstacle detection are the same as that of mature fruit location. Used iterate on 2R-G-B and 2G-R-B chromatism component, combined with gray threshold method to segment image quickly and effectively. Got the branch regions by image binaryzation, morphological processing, region labeling and filling. Extracted skeleton of obstacle by thinning and did some processes so as to pruning the skeleton and recovering the occluded skeleton. Then obtained the feature points such as endpoints and branch points of the skeleton, recorded their connecting relationship. Finally the 3D information of obstacle was restored by stereo matching on feature points. Experimental results show that the identification accuracy of obstacle can reach 67.3%, the identification error ratio was increased when the actual distance of obstacle is more than 1.5m.

Through the research, some achievements have been made. Such as mature fruit recognition, match and location. This research also provides a method for fruit harvest robot to detect obstacle. The research results of this research have reference value for the study on visual recognition in the field of harvest robot in our country. They also provide a basis for further study and have important economic significance to enhance international competitive power of our country's agricultural.

KEY WORDS: recognition, matching, circular hough transformation (CHT), spatial location, obstacle

学位论文版权使用授权书

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。本人授权江苏大学可以将本学位论文的全部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  在 年我解密后适用本授权书。

学位论文作者签名：周明军
2009年6月6日

指导教师签名：李健康
2009年6月6日

独创性申明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，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以外，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。

学位论文作者签名：
2009年6月6日

第一章 绪 论

1.1 研究目的和意义

进入二十一世纪,我国的农业和农村经济发展正面临新的机遇和挑战。一方面,由于加入 WTO,农业和农村经济发展正面临着世界各地的冲击和挑战;另一方面也给我国带来了一次加快农业和农村经济发展以及对农村经济结构调整的绝好的机遇。随着农业生产正朝着规模化、多样化、精确化方向发展,农业劳动力的成本必然迅速上升,劳动力不足的现象也会日趋明显,因而作为高科技的机器人技术进入农业领域变得越来越现实。

柑橘作为世界第一大类水果,2002 全世界年贸易额达到 82 亿多美元,是仅次于小麦和玉米的第三大贸易农产品。过去 30 年,世界柑橘产量增长了一倍,据联合国粮农组织预测,到 2010 年,全球柑橘产量将达 6640 万吨,中国具有发展柑橘产业得天独厚的自然条件,适宜栽培柑橘的地域广阔,2007 年栽培面积达 191 万公顷,产量达到 2059 万吨,中国柑橘种植面积和产量均跃居世界第一。目前,中国人均消费柑橘鲜果 10.5kg,比 1978 年的 0.3kg 增加了 35 倍^[1]。在柑橘的生产作业中,收获采摘约占整个作业量的 40~50%,由于采摘作业的复杂性,采摘自动化程度很低。由于本世纪我国农业劳动力将逐渐向社会其它产业转移,而且我国面临着严重的人口老龄化问题,劳动力资源不足将逐渐变为现实,我国柑橘产量大,季节性强,大量的成熟柑橘如不及时收获和保存,由此造成的腐烂和掉落将损失巨大,柑橘自动采摘机器人能够降低人劳动强度,提高劳动生产率和产品质量,保证果实适时采收,因此,研究开发柑橘采摘机器人,具有重要的意义。

机器人技术的研究和发展分几个时代。第一代机器人很多是示教再现型工业用机器人,即按预先规定好的程序反复动作,这类机器人不具备识别外界状况的功能。近几年,利用各种传感器技术检测环境条件并进行工作的智能型第二代机器人研究迅速展开,在环境条件受到严格控制的作业现场实现应用。在此基础上,研究人员进一步研究开发具有计算机人工智能和推理功能的第三代机器人,并不断走向成熟。由于工作要求不同、使用行业不同,目前机器人研究领域形成了几

代并存、共同发展的局面。即使是这样，将现在的工业机器人系统直接应用于农业，面对田间作业现场的不同环境和多变条件，还是无法满足应用要求。与工业应用的人工控制环境截然不同，农业用机器人主要在自然环境下工作，要识别的作业对象形状、颜色等特征会因生长环境的不同而变化。如收获水果时，机器人需要从枝叶、地面、空间等杂乱背景中分辨出水果，并根据其成熟度有选择地实施操作，在不同的光照条件下，原本就很复杂的问题，又变得更为复杂。因此，目前收获机器人存在的问题主要有果实的识别率和采摘率不高、果实的平均采摘周期较长等，解决收获机器人果实的识别和定位问题的关键是视觉功能的智能化。作为机器人的核心系统：即立体视觉系统的研究，成为机器人能否在农业上成功应用的关键，使机器人能像人一样，准确识别和定位目标，对实现机器人的智能化以及机器人升级换代起决定性作用，因此这方面的研究成为焦点。

收获机器人主要包括行走机构、控制系统和机械手等几大部分，其中机械手是机器人的主要执行部件，其构造复杂，造价昂贵，是机器人的最重要组成部分。在采摘苹果、柑橘等高大果树的果实时，机械手需要较大的工作空间^[2]，但实际生长的果树形态各异、参差不齐、许多果实生长在树枝后面，或者在两个树枝之间，而且这些果树的树枝较粗大，若机械手与这些树枝相碰撞将造成机械手的损坏，要想使机械手采摘到这些果实，必须避开树枝，所以对于开发柑橘、苹果等收获机器人，避障研究是必要的。要使机械手自动避障，必须先使机器人准确的识别障碍物，得到障碍物空间三维信息，然后通过空间建模和机械手路径规划实现自动避障。目前在机器人导航以及汽车行业中，对机器人避障已进行了较广泛的研究，但在收获机器人方面，对机械手避障的研究较少，目前开发出的收获机器人，如日本的茄子采摘机器人^[3]，荷兰的黄瓜采摘机器人^[4]，美国的柑橘采摘机器人等都没有机械手避障功能，所采摘的都是表层的无障碍的果实，因此目前果实收获机器人的应用范围较窄，在有坚硬树枝等障碍物存在的情况下，机械手的采摘较危险。为了增加收获机器人的实用性，提高采摘的安全性，机械手必须能自动避障，而如何准确的识别和恢复障碍物三维信息则是机器人避障必须首先解决的问题。

综上所述，本研究的两大主要内容：柑橘采摘机器人果实定位和障碍物检测，是实现柑橘采摘机器人自动、快速、准确、安全采摘果实的关键，是柑橘收获机

机器人发展必须首先解决的问题。

1.2 国外研究现状

在果实识别和定位方面,国外对计算机视觉在农产品收获中应用的研究起步很早,尤其是在树上果实的检测方面做了很多工作,为我们的研究提供了许多宝贵的经验。

机器视觉技术在果蔬收获机器人中的应用开始于1968年,美国学者 Schertz 和 Brown 最早提出果实自动化采摘的想法。利用果实和叶子在电磁光谱的可见光和红外区域反射率的不同来检测水果^[5,6],在他们的文章中,建议果实的位置可以通过光度信息测定,尤其是利用叶子和果实对可见光或者电磁光谱红外部分反光度的差异。同时文章中也指出了在果实自动化检测中必须考虑的问题:(1)不均匀照明;(2)枝叶遮挡使得只有70%到100%的果实可以看到。

受 Schertz 和 Brown 想法的启发,各种有关果实自动化采摘的研究在世界各地陆续开展起来。Parrish 和 Goksel 在1977年据上述理论,使用一台 B/W 照相机和一个用于增强红色苹果和绿色叶子对比度的红色光学滤波器,构建了第一个较为实用的苹果识别视觉系统^[7],这项研究拓宽了收获机器人研究的思路。他们利用人造苹果树进行了一些试验,但是没有报道具体的检测结果。

D'Esnon 等人报道了用于苹果检测的第一套彩色视觉系统^[8]。通过对二值图像的处理能找到各水果的区域的中心。但是这种系统误识别率较高,检测时为了避免视觉系统看到未被树叶挡住的天空,要求在果树后面实施遮挡,防止产生大量虚假目标^[9]。

1987年,Whittaker^[10]首先提出基于形状信息检测果实的方法。他认为基于局部像素颜色值的分析方法不适用于建立对颜色不敏感的处理系统,建议基于形状信息同时考虑局部像素颜色值和整体颜色值的内在关系,在处理的过程中用到了优化圆形 Hough 变换。该方法虽然仅限于理论上的探讨,但是开辟了果实识别的一个新的途径。

Illingworth 等人^[11]用一台黑白摄像机拍摄得到灰度为256级的灰度图像,经过 Sobel 算子处理得到每幅图像的梯度矢量图,通过阈值化得到边界图像和方向图像。方向图像包括每个边界的矢量方向和最大亮度变化方向,运用优化的圆形

Hough 变换 (CHT) 检测图像中符合西红柿轮廓的弧度。该研究采用人为给定阈值的方法, 不能适用光线变化。给定阈值对结果影响很大, 最好的检测准确率为 68%。本文采用的迭代算法能根据采集图像的亮度变化自动选择分割阈值, 很好的解决光线变化问题。

Harrell 等人对柑橘类采摘机器人 (CPR) 进行了研究^[12~14]。采用色调、饱和度以及 RGB 颜色信息进行分类, 利用经典的贝叶斯分类器, 这种分类方法将成熟的柑橘或其它水果像素从背景中分离比较有效。但是需要人工照明, 并且当背景颜色和水果相似时, 如大片枯叶、地面等杂物容易被误识别成目标, 并且不能检测有遮挡的果实。

1990 年, 匈牙利和美国科研联合基金会^[15]联合研制开发出苹果收获机器人。该机器人用立体视觉系统自动检测苹果, 给出检测到的苹果的空间位置。试验结果表明: 可以找到 41% 的可视果实, 同时出现了一些误识别。他们认为出现这样的结果原因有两方面: 一是图像分割时出现了一些错误的判断, 图像上有些非苹果部分被判定是苹果; 二是利用立体视觉寻找果实上同一部分在两幅图像上的对应区域时, 由于遮挡、重叠等缘故, 当果实在果树深处时, 没有充分利用匹配算法中的约束条件, 如极线约束、唯一性约束、保序性约束等, 在遍历搜索过程中导致多重匹配, 出现虚假目标, 采用基于区域的匹配算法, 在灰度分布相似区域匹配不佳是该算法的缺陷。为此, 本文采用基于特征的立体匹配算法, 运用各种约束条件, 减少了定位误差和虚假目标的出现。

1995 年, 意大利的 Buemi^[16]开发了用于温室柑橘采摘的自动化系统。该视觉系统使用彩色相机获取图像, 并将颜色转换到 HIS 颜色系统下, 利用色调和饱和度对图像进行分割, 通过对同一场景两幅不同的图像进行空间匹配获得三维信息, 试验结果认为可以检测到大约 90% 的成熟柑橘, 出现的错误检测, 基本上都是由遮挡问题造成的。

西班牙的 Ceres^[17]等认为使用光学相机系统, 造成错误检测的原因是光照的变化和场景中同果实和背景具有相同颜色或形状属性。1998 年, 他们在西班牙 AGRIBOT 工程中提出使用激光区域探测传感器获得图像, 利用形状识别果实。他们把检测的目标近似看作球体, 把轮廓线、凸度、凸面、反射率四种不同的点阵用于估计球体的参数: 空间位置、半径、反射率。分别利用人工树和自然树进

行了测试, 研究表明 80% 的可视果实被探测到, 没有错误检测。根据他们的视觉识别系统, 果实的空间位置可精确到 10mm。

高桥照夫等人(2000)利用双目立体视觉对果园的苹果进行定位研究。以红富士苹果为采摘对象进行实验, 结果表明, 目标距离为 1~4m 的有效范围内测量误差较小; 距离在 1.2~3.5m 时, 误差在-4~2%之间; 大于 3.5m 时误差在 -7~0% 之间。系统要求的光线强度控制在 1~90klx, 果实的重叠度应小于 0.5。

荷兰农业环境工程研究所(IMAG B.V.) (2002)研究开发了一种移动式黄瓜收获机器人样机^[18], 该研究在荷兰2hm²的温室里进行。实验结果表明, 高峰期需要4台机器人。每台机器人每日工作18h, 作业速度为10秒/根, 相当于12个工人每日6h的工作量。蔬菜果实收获自动化可以节约果实收获人工费用, 但要满足商用产品的各种要求, 还需对样机加以改进和完善。

日本 Shimane 大学(2000)构建了一种收获机器人视觉系统^[19]。Bulanon^[20] (2002, 日本)根据颜色和灰度的组合特征对不同光照条件下的 Fuji 苹果进行识别, 成功率 88%, 他们认为影响识别结果的主要因素是光照条件。Limsiraratana^[21]等(2002, 日本)以木瓜为对象, 利用椭圆傅立叶描绘子(Elliptic Fourier Descriptors)根据形状在图像上识别果实, 为了优化识别过程, 采用了快速傅立叶变换(FFT)和遗传算法(GA), 最后结果认为这种方法是可行的, 但仍需进一步深入研究。目前, 日本、荷兰、美国、以色列等国都开展了果蔬收获机器人方面的研究工作, 但是研究大都停留在理论或实验阶段, 还没真正实现商业化^[22]。

在机器人障碍物检测方面, 目前在机器人导航以及汽车行业中, 对机器人避障已进行了较广泛的研究, 但在收获机器人方面, 对机械手避障的研究较少, 其原因主要是农业环境中由于背景的复杂性和光照的不稳定性, 障碍物的检测较复杂, 因此这方面的研究任重而道远。本文提出了一种障碍物检测的方法, 在收获机器人障碍物检测方面做了一些基础性的研究。

1.3 国内研究现状

国内利用机器视觉技术在农业收获机器人方面的研究起步较晚。

1997年, 吉林工业大学的杨秀坤等^[23]提出利用计算机视觉对果实表面缺陷

进行检测的方法。该方法首先对果实图像进行滤波、增强等预处理，然后根据像素颜色值，利用自适应特征聚类神经网络和模糊加权决策树相结合的方法对果实表面缺陷进行检测。实验结果认为该方法优于传统的检测方法。

2001年，南京农业大学的张瑞合等^[24]对自然环境下番茄的识别与定位进行了研究，他们根据颜色特征识别番茄，利用双目立体视觉从两幅二维图像中恢复出番茄的三维空间坐标，实验结果认为当目标距离为300~400mm时，深度误差可以控制在3%~4%。

2002年，中国农业大学的孙明等^[25]提出利用计算机视觉识别萝卜幼苗。该方法首先突出图像中各像素的绿色颜色分量，然后结合亮度信息，采用最大方差自动取阈值等方法识别目标。

2005年，江苏大学的蔡健荣等^[26]将RGB颜色空间转换到球形HIS颜色系统，并利用Otsu算法自动获取分割阈值，以成熟的西红柿为对象，根据水果的颜色信息和形态信息能够识别出成熟水果，但对有遮挡或果实重叠的情况研究较少。

综上所述，国内对机器视觉在农产品收获和加工中的应用进行了大量的研究，但和国外仍存在一定差距，特别是在农产品收获应用方面，国内的研究很有限。为了赶超国外先进水平，早日实现我国农业生产自动化，必须进一步开展广泛和深入的研究。

1.4 研究内容和方法

根据国内外研究状况，基于双目立体视觉对自然环境下植株上的果实和障碍物进行识别并确定其位置的构想是可行的，本研究以柑橘为研究对象，以两个光轴平行的同型号CCD摄像头构成双目立体视觉系统，对大棚和田间生长的成熟柑橘分别在不同时间、不同天气下进行图像采集，基于VC++6.0平台自行开发了处理程序进行图像处理，运用相关的图像处理算法及空间测距理论，对处理结果进行分析和计算，从而达到识别和定位的目的。本研究的主要内容如下：

1、柑橘果实定位

选择合适的颜色系统，对图像中目标和背景进行合理分割，研究受环境影响较小的分割方法；对分割出的图像目标区域，采用有效的特征识别方法，力求准确快速提取目标区域特征，为立体匹配和空间定位做准备；使用正确的立体匹配

方法对左右图像的果实区域进行匹配,根据匹配的结果计算果实的空间位置,将计算的结果同实际位置相比较,从而调整相应算法,以达到准确定位。

2、障碍物检测

主要步骤和果实定位相同,根据柑橘收获机器人机械手自动采摘过程中遇到的树枝等主要障碍物的特点,研究对于障碍物的分割和提取方法,针对树枝的多样性的特点,研究有效的特征提取、立体匹配和三维重建方法恢复树枝三维信息,本文提出了相应的解决方案。

1.5 本章小结

本章首先介绍了研究的目的是和意义,然后对国内外研究情况做了介绍,最后简要的提出了本研究的主要内容。

第二章 双目立体视觉系统简介

众所周知，人们从外界环境获取的信息中，80%来自于视觉，其他来自于触觉、听觉、嗅觉等感觉器官。当人的眼睛从自己周围的环境获取大量信息，并传入大脑后，由大脑根据知识或经验，对信息进行加工、推理等处理工作，最后识别、理解周围环境，包括环境内的对象物，如物体间的相对位置、形状、大小、颜色、纹理、运动还是静止等。计算机视觉就是用计算机模拟人眼的视觉功能，从图像或图像序列中提取信息，对客观世界的三维景物和物体进行形态和运动识别。计算机视觉研究的目的之一就是要寻找人类视觉规律，从而开发出从图像输入到自然景物分析的图像理解系统。

目前许多计算机视觉专家都是在马尔(Marr)创立的视觉计算理论框架下探索。美国麻省理工学院人工智能实验室的马尔教授认为，视觉可分为三个阶段，第一阶段是早期视觉，其目的是抽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、光流计算等；第二阶段是二维半简图或本征图像，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；第三阶段是二维模型，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。

尽管马尔视觉计算理论在解决具体问题时遇到了困难，但可以说，计算机视觉有了长足的发展，其主流是沿着马尔所指出的方向取得的。

根据以上的马尔视觉理论，可以将计算机视觉分为二维视觉和三维也就是立体视觉。二维视觉是把输入图像转换成具有所希望特性的另一幅图像，例如滤波、锐化、图像相加等；或者提取图像的某些特征，例如边缘提取、图像分割等。立体视觉研究最多的是马尔提出的双目立体视觉，由于立体视觉很多处理要以二维

视觉为基础,如图象分割、特征提取等。因此,可以认为立体视觉的研究包括了对二维视觉信息的研究,这里主要介绍双目立体视觉。双目立体视觉是计算机视觉的一个重要分支,它由不同位置的两台或者一台摄像机经过移动或旋转拍摄同一幅场景,通过计算空间点在两幅图像中视差,获得该点的三维坐标值。马尔提出了一种视觉计算理论并应用在双眼匹配上,使两张有视差的平面图产生有深度的立体图形,奠定了双目立体视觉发展的理论基础。相比其他类的体视方法,如透镜板三维成像、投影式三维显示、全息照相术等,双目体视直接模拟人类双眼处理景物的方式,可靠简便,在许多领域均极具应用价值,如微操作系统的位姿检测与控制、机器人导航与航测、三维测量学及虚拟现实等^[27]。

双目体视技术的实现可分为以下步骤:图像获取、摄像机标定、特征提取、图像匹配和三维重建,下面对各个步骤作简要介绍。

图像获取,双目体视的图像获取是由不同位置的两台或者一台摄像机(CCD)经过移动或旋转拍摄同一幅场景,获取立体图像对,两 CCD 平行放置或者成 50° 到 80° 之间的夹角。

摄像机的标定,对双目体视而言,CCD 摄像机、数码相机是利用计算机技术对物理世界进行重建前的基本测量工具,对它们的标定是实现立体视觉基本而又关键的一步。通常先采用单摄像机的标定方法,分别得到两个摄像机的内、外参数,再通过同一世界坐标中的一组定标点来建立两个摄像机之间的位置关系。

特征点提取,立体像对中需要提取的特征点应满足以下要求:与传感器类型及抽取特征所用技术等相适应,有足够的鲁棒性和一致性。在进行特征点像的坐标提取前,需对获取的图像进行预处理,因为在图像获取过程中,存在一系列的噪声源,通过此处理可显著改进图像质量,使图像中特征点更加突出。

立体匹配,立体匹配是双目体视中最关键、困难的一步。与普通的图像配准不同,立体像对之间的差异是由摄像时观察点的不同引起的,而不是由其它如景物本身的变化、运动所引起的。根据匹配基元的不同,立体匹配可分为区域匹配、特征匹配和相位匹配三大类。

三维重建,在得到空间任一点在两个图像中的对应坐标和两摄像机参数矩阵的条件下,即可进行空间点的重建。通过建立以该点的世界坐标为未知数的 4 个线性方程,可以用最小二乘法求解得该点的世界坐标。实际重建通常采用外极

线约束法。空间点、两摄像机的光心这三点组成的平面分别与两个成像平面的交线称为该空间点在这两个成像平面中的极线。一旦两摄像机的内外参数确定，就可通过两个成像平面上的极线的约束关系建立对应点之间的关系，并由此联立方程，求得图像点的世界坐标值。对图像的全像素的三维重建目前仅能针对某一具体目标，计算量大且效果不明显，因此通常使用提取图像中目标区域的特征，通过对图像对中区域特征的处理达到目标三维重建的目的。

本研究选用双目立体视觉系统主要基于以下考虑：

1) 计算结果精度较高。用双目立体视觉技术计算的结果其绝对误差可达到 0.03mm 以下^[28]，可以达到要求；

2) 计算相对简单。由于双目立体视觉系统中双摄像机的参数彼此一致，光轴相互平行且垂直于基线 (baseline 通过两摄像机光心的直线)，构成一共极性 (epipolar) 结构，这样可以缩小对应的搜索空间，只在水平方向存在视差 ($X_l - X_r$)，从而简化了立体视觉中的匹配过程；

3) 价格比较便宜。双目立体视觉系统不需要对光源做特殊控制，相比其它主动的测距方法，它不需要专门的主动光照射装置，通常在自然光或一定环境照明的条件下，由相距一定距离的两个摄像机各自摄取目标的图像，找出空间物点在两个图像中的对应点，就能得到点的距离信息。且系统的硬件组成比较简单，不需要昂贵的设备。

2.1 硬件组成

根据以上原理，本研究将两个同型号的 CCD 摄像机平行安装在摄像机架上，两摄像机通过 USB 接口与同一台电脑相连接构成双目立体视觉系统，如图 2-1 所示，该系统的硬件部分主要由两个相同型号的 CCD 摄像机、镜头、云台（支架）、激光测距仪和计算机等设备构成。CCD 摄像头用于图像采集，激光测距仪用于精确测距进行标定和验证，计算机用于图像处理。在本研究中，图像采集主要在自然环境下进行，没有采用人工照明。

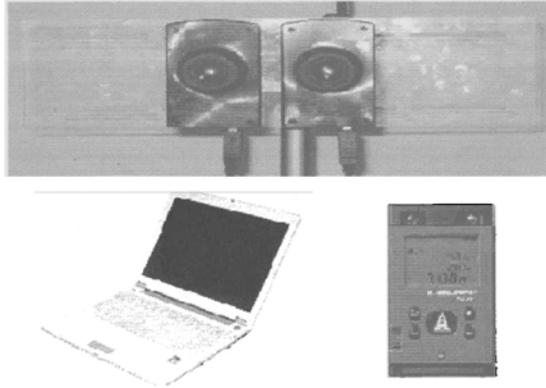


图 2-1 机器视觉系统装置示意图

Fig.2-1 Image of devices of computer vision system

在前期图像识别研究中,为了图像采集方便,使用数码相机代替 CCD 摄像头,在不同的天气情况(晴天、阴天)和拍摄条件下(顺光、逆光)对田间和盆景柑桔进行图像采集,后期的果实匹配和空间定位研究,利用双目立体视觉采集图像进行匹配实验,确定柑橘和障碍物的空间位置。本研究硬件系统各部分介绍如下:

数码相机: nikon-990 数码相机(最大分辨率: 2048×1536 、位深: 真彩 24 位、图像格式: BMP);

CCD: 美国 Lumenera 公司 Lu075C 型摄像头(所采集图像的有效尺寸: 640×480 、最大帧频率: 60fps、与电脑的连接方式: USB 接口);

镜头: 日本 Computar 公司 M0814-MP 和 M1214-MP 两种型号摄像头(每种型号各两个,焦距分别为 8mm 和 12mm);

计算机: CPU 为 AMD 公司 Athlon(tm)64 Processor 3000+(主频为 1.8GHz) 型兼容机;

激光测距仪: 德国喜利得(Hilti)公司的 PD 30 laser range meter;

采集地点: 镇江市江心洲果林场(大田)、计算机视觉研究室(室内);

采集时间: 2007 年 10-11 月份, 2008 年 10-11 月份。

摄像头支架为普通的摄像机支架,可调节摄像头高度和拍摄角度,调节方式为手动调节(自动的调节支架正在开发中),摄像头安装在有机玻璃板上,玻璃板上开有两条窄槽可使两摄像头的基线长度在 6cm-25cm 之间调整。

2.2 相关图像处理软件

果实定位及障碍物检测所用算法在 Microsoft VC++6.0 平台下开发，开发过程中还使用了 MATLAB7.0.4 和 OpenCV1.0 等图像处理工具。

2.3 本章小结

本章主要介绍了双目立体视觉系统的原理、组成，同时介绍了本研究所用的硬件组成、软件平台及图像采集的时间、地点、环境及方式。

第三章 成熟柑橘果实定位

利用机器视觉进行图像识别,首先要选择合适的特征,把要识别的目标与背景区分开,然后运用一些合适的分割方法,达到去除背景,识别目标的目的。基于机器视觉识别果实可划分两类:一是根据颜色或者灰度特征分割图像、寻找果实;二是在图像采集时使用滤波片或其它设备获取较为简单的图像,再根据形状特征寻找果实^[29,30]。

图像识别中所用的图像是在自然光下采集得到,采集的地点分为田间和室内,采集的时间段有上午、中午、下午,采集时的天气情况有晴天、多云、阴天,采集的方式有顺光和逆光等。图像中的成熟柑橘为识别目标,其余的都视为背景。图像采集时,没有对自然生长的植物做特殊处理,因此背景极为复杂,背景中有柑橘植株上的叶子、茎、透过枝叶的天空和土壤、用于支持植株的木杆或者绳线等,其中叶子又有绿色和枯黄两种,而图像采集主要是在自然光下进行,光照对采集到的图像质量有很大影响。光照太强会导致图像失真,图像上柑橘与背景区别不明显,从而导致错误的识别结果;光照太弱也会造成采集到的图像上信息区分不明显,同样会造成错误识别或降低识别精度。为了减少错误识别和提高识别精度,本研究中在特征选择和图像处理方法上给以足够重视,充分考虑各种情况,力求达到最佳效果。在目标特征提取时,对于柑橘图像中容易出现的遮挡和重叠等情况,本文对比了多种不同的特征识别方法,力求达到满意的识别效果。

3.1 计算机视觉中的颜色表示

人们可以根据颜色识别不同的物体或者物体的不同部分,但是颜色究竟是什么、如何准确的鉴定、表示或者显示颜色,却不是一个简单的问题。大约 60 多年前,科学家们开始研究这些问题,并逐渐形成一门新的学科——色度学。色度学发展到现在,已经出现了十多种用于表示颜色的模型,如 RGB、HIS、HSV、YCbCr(亦称 YUV)、L*a*b*等,依据这些颜色模型,可以在不同领域方便的鉴定、使用颜色。可见光的波长在 400nm-700nm,当光谱采样限制到三个人类视觉系统敏感的红、绿、蓝光波段时,对这三个光谱带的光能量进行采样,就可以

得到一幅彩色图像。国际照明委员会(CIE)规定以 700nm(红)、546.1nm(绿)、435.8nm(蓝)三个色光为三基色(如图 3-1 所示)。RGB 模型是根据三基色原理建立起来的,是基本的颜色表示模型,其他的颜色表示模型都可通过 RGB 模型转化得到。本研究中使用了目前数字图像通常使用的 RGB 和 HIS 颜色系统。

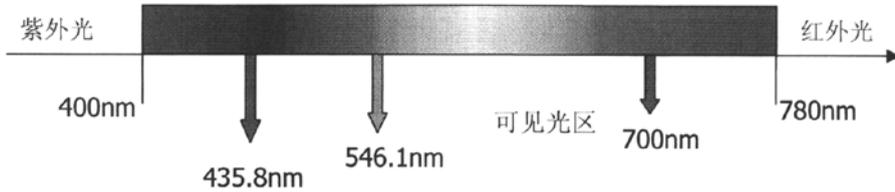


图 3-1 三基色及其波长

Fig.3-1 Image of tricolor and their wavelength

RGB颜色模型是最常见的一种颜色空间(如图3-2所示)。自然界的各种颜色都可以由红(R)、绿(G)和蓝(B)三基色加权混合而成,绝大多数彩色成像设备也采用该格式。RGB颜色系统是在三基色学说下建立起来的颜色系统。它通过改变三基色的数量,混合出其他各种颜色,可以用颜色方程(3.1)表示混合后的颜色:

$$C \equiv R(R) + G(G) + B(B) \quad (3.1)$$

式中, C 表示混合后的颜色; \equiv 表示视觉上的相等, (R)、(G)、(B) 表示红、绿、蓝三基色, R、G、B 分别表示三基色的数量。所有的颜色,包括黑白系列的各种灰度、各种色调与饱和度的彩色,都可以用红、绿、蓝三原色混合而成,三原色的数量为三刺激值,这样的色度学系统称为 CIE1931-RGB 彩色系统^[31]。

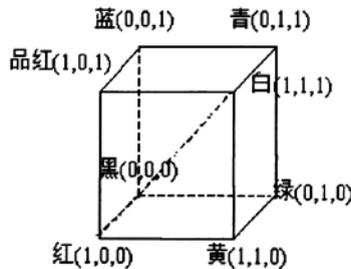


图 3-2 RGB 彩色系统

Fig.3-2 RGB color system

本研究所处理的图像均为真彩色数字图像,也就是图像中的三原色 R、G、B 各有 256 个彩色级(图 3-2 中的“1”表示最高级别,即 255)。

HIS颜色模型是由孟塞尔(Munsell)最早提出的一种颜色系统格式,因此也有

人称其为孟塞尔空间^[32]。HIS颜色空间使用非常近似于人对彩色感知的方法来定义彩色，是一种很直观的方法，符合人对颜色的描述习惯。它由三个量来度量，即色调、饱和度和亮度，它们共同产生视觉效果。其中，色调（Hue）主要由可见光光谱中各分量成分的波长来确定，是颜色光的基本特性；饱和度（Saturation）反映了彩色的浓淡，它取决于彩色光中白光的含量，掺入白光越多，彩色越淡，当白光占主要成分时，彩色淡化为白色；亮度（Intensity）指彩色光对人眼引起的光刺激强度，即该彩色光的明亮程度，它和光的颜色无关。HIS模型与RGB模型的转化关系如式3.2~3.4所示：

$$H_0 = \text{Arc cos} \left\{ \frac{2R - G - B}{2\sqrt{(R - G)^2 + (R - B)(G - B)}} \right\} \times 180 / \pi$$

$$H = 360 - H_0 \quad B > G \quad (3.2)$$

$$H = H_0 \quad B \leq G$$

$$S = 1 - \frac{3\text{Min}(R, G, B)}{R + G + B} \quad (3.3)$$

$$I = \frac{R + G + B}{3} \quad (3.4)$$

HIS 彩色系统模型与理想的双锥体模型吻合（如图 3-3 所示）。一个颜色样本的特征矢量在垂直于 I 的平面上的投影与水平方向的夹角为样本的色调 H，H 的取值范围为 0 ~ 360；投影线的长度表示颜色样本的饱和度 S，取值在 [0,1]；I 轴与明度线重合，以明度值最低的黑点表示原点，如果 I 经过归一化后的 I' 的值域在 [0,1]，0 表示明度最低的黑点，1 表示明度最高的白点。当 I' = 0.5 时，饱和度和度方向可分辨的颜色数最多，如下图 3-4 所示。

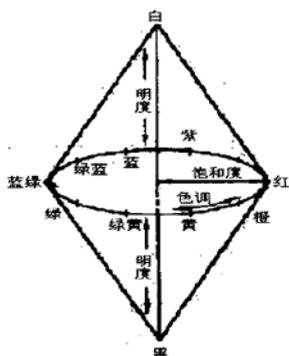


图 3-3 HIS 的理想双锥体

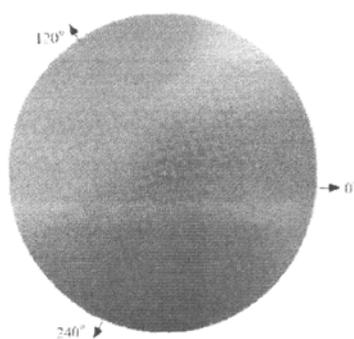


图 3-4 当 I = 0.5 时饱和度方向上的颜色

Fig.3-3 Ideal double vertebral of HIS Fig.3-4 Color distribution when I=0.5

3.2 成熟柑橘区域提取

3.2.1 图像分割原理

机器视觉系统中图像分割是目标提取的关键和难点, 图像分割的准确程度将直接影响后续处理。图像分割多年来一直得到人们的高度重视, 至今已提出了上千种不同的分割算法^[33]。以前提出的大多数分割算法是基于灰度图的, 近年来由于电子技术的发展和彩色 CCD 的普及, 同时由于彩色图像本身包含大量有利于图像分割的信息, 人们对彩色图像分割越来越重视。彩色图像分割主要分为以下几类: 特征空间聚类方法、基于随机模型的方法、基于区域的方法和基于直方图的方法^[34]。由于本研究的对象是具有生命的离散个体, 生长环境(光照、营养、水分等)不同, 会引起个体的形状或颜色不同, 而且水果在树上的位置是不确定的, 光照条件时时在变。自然光照强度变化和阴影等都会对图像获取和图像分析带来困难, 影响识别的正确率。

本研究在提取和理解成熟柑橘、树叶和树枝等颜色信息的基础上, 对不同条件下拍摄的图像在其相应颜色模型中各个分量的直方图进行分析比较(如图 3-5 所示)。

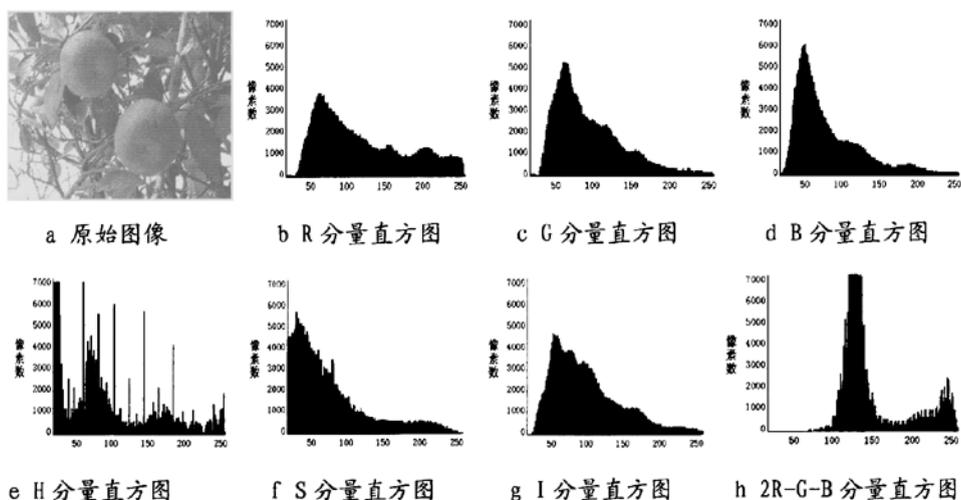


图 3-5 图像各分量直方图

Fig.3-5 Histogram of different image component

实验发现, 在原始 RGB 和 HIS 等颜色模型的分量中, 只有 H 分量有比较明显的目标区域波峰(由于柑橘颜色与背景有明显的差异, 其色调介于红色与黄色之间, 其 H 分量值在 0~30 附近, 故图 3-5e 中 0 附近的波峰区域为柑橘区域),

其余分量均无法区分目标区域与背景区域，在 HSV、L*a*b*等颜色系统下二者的区分更困难。因此可考虑使用 H 分量分割成熟柑橘，但柑橘图像的 H 分量往往呈现多峰状态，较难确定分割阈值，若人为设定阈值，则鲁棒性低，适应性差，且图像中成熟柑橘的 H 分量值随着柑橘种类、拍摄环境、相机型号等的变化而变化。如图 3-6b 所示，由于图像中柑橘区域占整个图像的比例较小，直方图中柑橘的 H 分量峰值低，图 3-5e 中的峰值 0 处成了波谷，在图像分割时不容易确定分割阈值。

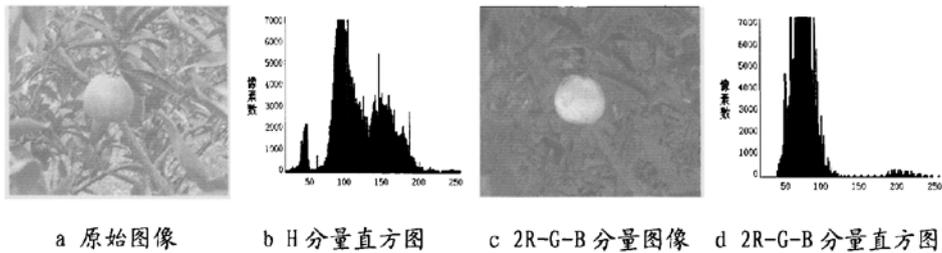


图 3-6 柑橘图像 H 和 2R-G-B 分量比较

Fig.3-6 Comparison of H and 2R-G-B component of citrus image

参考文献[35]利用 RGB 系统中的色差分量 2R-G-B，结合 Ostu(最大类间方差法)自适应阈值分割法分割成熟柑橘区域，使用 Ostu 的前提是图像的直方图存在明显的双峰，但很多柑橘图像的 2R-G-B 分量直方图双峰状态不明显，如图 3-6d 所示，在这种情况下，使用 Ostu 分割不准确。从柑橘图像的 2R-G-B 分量图像(图像每个像素的 RGB 值都等于原图像中该像素的 2R-G-B 值)可以看出，图像中成熟柑橘与背景的 2R-G-B 值存在较大差别，且在直方图中图像的背景区域波峰明显，如图 3-5h 和 3-6d 所示。正确的图像分割阈值应该在这一波峰右边的第一个波谷处，若能正确找到这一波谷，则能实现准确的自动阈值分割。根据这一要求，本研究参考了文献[36]的迭代自动求阈值方法，并将该方法运用于本研究中，取得了满意的效果。

3.2.2 阈值分割

迭代法使用图像直方图求取图像分割阈值，迭代法的实现原理如下：

设图像的灰度值使用 256 级，即灰度值从 0~255 变化。设 S_i 表示图像内灰度值为 i 的像素点的总数， i 的取值范围为 0~255； T_i 表示阈值，有如下公式：

$$Tmp = \sum_{i=0}^{T_i} i \times S_i \quad (3.5)$$

$$Tmp1 = \sum_{i=0}^{T_i} S_i \quad (3.6)$$

$$Tmp2 = \sum_{i=T_i+1}^{255} i \times S_i \quad (3.7)$$

$$Tmp3 = \sum_{i=T_i+1}^{255} S_i \quad (3.8)$$

$$T_{i+1} = (Tmp / Tmp1 + Tmp2 / Tmp3) / 2$$

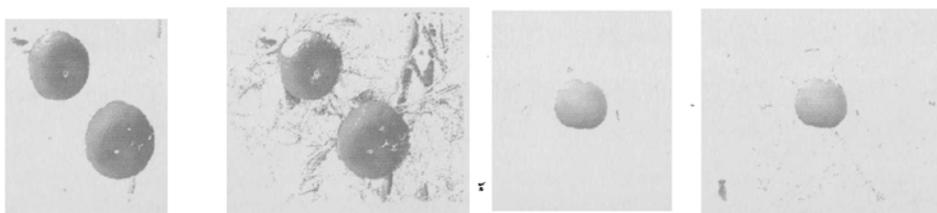
$$= \left(\sum_{i=0}^{T_i} i \times S_i / \sum_{i=0}^{T_i} S_i + \sum_{i=T_i+1}^{255} i \times S_i / \sum_{i=T_i+1}^{255} S_i \right) / 2 \quad (3.9)$$

若指定一个极小值 ε ，有：

$$|T_{i+1} - T_i| \leq \varepsilon \quad (3.10)$$

T_{i+1} 即为最后的迭代结果，否则令 $T_i = T_{i+1}$ 重新执行上面的计算过程，直到满足(3.10)式的条件^[36]。上述的迭代法用计算机实现很简单，一般只需四五次迭代即可找到阈值。迭代中，初始值 T_0 的选择决定迭代的收敛速度，初始值越靠近目标值，收敛速度越快，本研究设初始值 T_0 为 150。

运用迭代法对图 3-5 和 3-6 的原始图像 2R-G-B 分量直方图进行迭代求阈值，将图像中 2R-G-B 值小于阈值的像素变为白色，得到的结果如图 3-7a 和 c 所示。



a 2R-G-B 迭代阈值分割 b H 分量阈值分割 c 2R-G-B 迭代阈值分割 d H 分量阈值分割

图 3-7 图像分割结果

Fig.3-7 Result of image segmentation

为了与对 H 分量设定阈值的图像分割方法相比较，图 3-7b 和 d 给出了对 H 分量人为设定阈值的方法分割的结果，由图对比可知，2R-G-B 迭代阈值分割法

能更好的分割出目标区域，去除背景区域，且自动寻找分割阈值，无需人为干预。

3.2.3 二值化处理

将经过阈值分割后的彩色图像转化为二值图像，为方便后续处理，将背景区域像素的 RGB 值设为 255，目标区域设为 0，效果如图 3-8a 所示。经过以上分割处理后，可以去除绝大部分背景区域，但由于自然场景下拍摄的图像受外界干扰较大，图像中与成熟柑橘颜色接近的区域、随机噪声等仍无法完全去除，可采用形态学处理和区域标记法将其除去。

3.2.4 形态学运算

数学形态学是以形态结构元素为基础对图像进行分析的数学工具。它的基本思想是用具有一定形态的结构元素去度量和提取图像中的对应形状以达到对图像分析和识别的目的。数学形态学的应用可以简化图像数据，保持它们基本的形状特征，并除去不相干的结构。图像形态学基本运算包括膨胀（Dilation）、腐蚀（Erosion）、开启（Open）和闭合（Close）运算。它们在二值图像中和灰度图像中各有特点。基于这些基本运算还可以推导和组合成各种数学形态学实用的算法 [37]。

设 F 为灰度图像， B 为结构元素（Structure element）， B 可以是圆、三角形、正方形等简单的几何基元，可以是对称或不对称。但不对称的结构元素会使图像产生位移，因此要用 B 的对称集 B^V 配对运算。设膨胀运算符为 \oplus ，腐蚀运算符为 \otimes ，开启运算符为 \circ ，闭合运算符为 \bullet 。二值图像形态运算中：

腐蚀是一种消除边界点，使边界向内部收缩的过程，同时使区域内的小孔变大，可以用来消除颗粒噪声，定义为

$$E(x, y) = (F \otimes B)(x, y) = \underset{i, j=0}{\overset{m}{AND}} [F(x+i, y+j) \& B(i, j)] \quad (3.11)$$

膨胀是将与目标接触的所有背景点合并到该物体中，使边界向外部扩张的过程，同时使区域内的小孔变小，可以用来消除内部空洞，定义为

$$D(x, y) = (F \oplus B)(x, y) = \underset{i, j=0}{\overset{m}{OR}} [F(x+i, y+j) \& B(i, j)] \quad (3.12)$$

$$\text{开启: } F \circ B = (F \otimes B) \oplus B \quad (3.13)$$

$$\text{闭合: } F \bullet B = (F \oplus B) \otimes B \quad (3.14)$$

开启和闭合运算都是膨胀和腐蚀的组合。开启运算能够消除尺寸比结构元素小的结构，因此可消除图中的孤立点，从内部平滑物体的边缘并能断开相连的物体。闭合从外部平滑物体的边缘并能闭合物体内部的空洞。

本文对图像进行的形态滤波处理方式如公式3.15所示，表示先进行 n 次腐蚀，再进行 n 次膨胀，其中 $n = 2 \sim 3$ （一般取2次），结构元素是 3×3 的正方形。

$$G = (F \otimes nB) \oplus nB \quad (3.15)$$

经过滤波消除阈值分割后遗留的随机噪声，减少后续处理运算量，对图3-8a进行形态滤波的结果如图3-8b所示。

3.2.5 区域标记

区域标记是一种当图像中有多个区域时，将它们区分开来，分别进行描述的方法。在二值图像中，传统的标记方法有顺序法、蚕食法等^[38]。这些方法通常需要对图像执行二次扫描。第一次通过逐行逐列扫描像素，判断像素之间的邻域关系，对属于同一连通区域的像素赋予相同的连通标号，实现连通标识。这种逐行逐列的次序扫描的结果，会产生同一像素点被重复标记的现象，同一连通区域的不同子区域被赋予了不同的标记号。因此，需要执行第二次扫描消除重复性的标记，合并属于同一连通区域的但标记不同的子区域，效率较低。

本文采用类似区域增长的思想，避免了重复标记的问题。不对二值图像作逐行或逐列的次序扫描，而是一次性地标记整个连通区域，然后再标记下一个区域，直到所有的连通区域都被标记。具体算法操作步骤为：首先对输入的二值图像进行逐行扫描，找到一个未标记区域的第一点，标记该点；检查该点的八邻域点并标记满足连通性要求且尚未被标记的点。将新增的标记点记录下来作为“区域增长”的种子点。在后续的标记的过程中，不断地从记录种子点的数组中取出一个种子，进行上述的操作，如此循环，直到记录种子点的数组为空，一个连通区域标记结束。接着再标记下一个未标记区域，直到输入二值图像的所有连通区域都被标记。然后统计每个连通区域的像素总数，将像素总数小于 500 的小区域去除，这里的阈值 500 是统计的结果，随着柑橘品种、镜头焦距和物距的不同而改变，在应用中可根据情况调整。运用该法对图 3-8b 进行标记的结果如图 3-8c

所示。由图可以看出，较小的背景区域被去除。

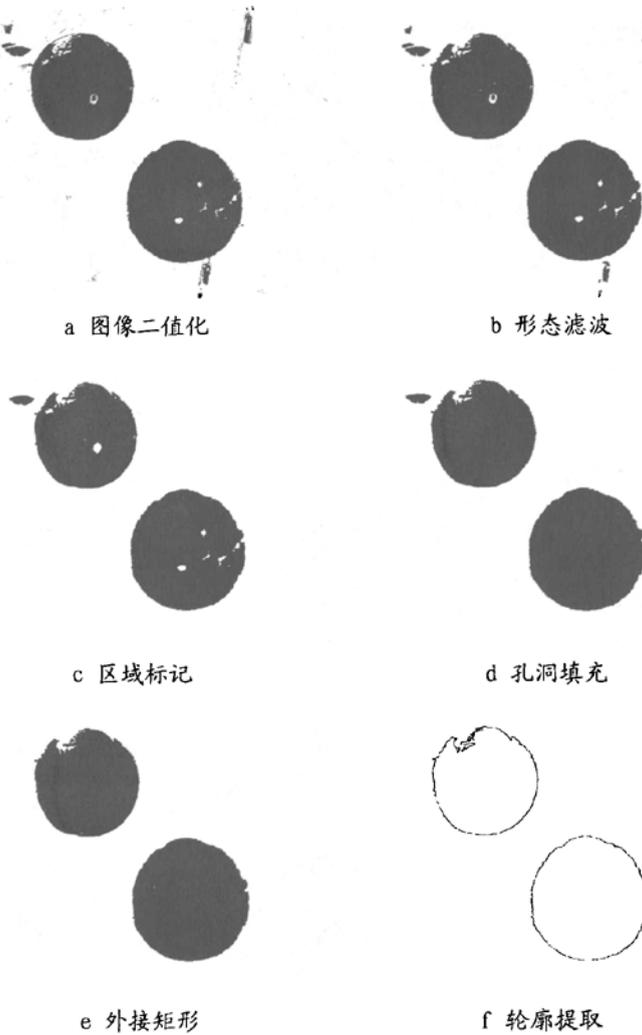


图 3-8 二值图像处理

Fig.3-8 Binary image process

3.2.6 区域填充

由于柑橘呈球形，会产生耀斑，通常情况下经过分割后的目标区域都存在一些孔洞，这些孔洞将对后续的处理产生较大影响，必须进行填充。常用的区域填充算法有多边形有序边表填充算法、种子填充算法及多边形边标志填充算法^[39-41]。有序边表填充算法的主要缺点是对各种表的维持和排序开销太大；种子填充算法的缺点是需要栈结构，因而需要较大的存储空间以实现栈结构，而且在有多个对象需要填充时种子点的逐个选取会降低效率，在有些情况下甚至是不现实

的;多边形边标志填充算法速度虽然较快,但对任意形状区域的填充却存在缺陷,这对于在图像处理时需要完整地保持原形态的情况下尤为突出。本研究使用参考文献[42]提出的利用图像外接矩形进行搜索的孔洞填充算法,该算法可用于任意形状区域中若干孔洞的快速填充。在该算法中,为了方便填充区域的计算、减少辅助内存,申请了两个数组,分别用于保存扫描过程中所搜索到的所有点的横坐标和纵坐标,称为坐标数组。坐标数组的存储大小和存储方式与原图像相同。该孔洞填充算法的描述如下:

1) 首先标记图像中目标区域,标记方法使用3.2.5节的标记法,将不同的黑色区域标记为不同的值,假设图像中 k 个区域标记的值分别为 $1\sim k$ 。

2) 通过已知区域内(即图中每个黑色区域)所有点的坐标找出包含该区域的最小矩形区域,并且将每个矩形内的白色区域(包括孔洞区域和柑橘目标区域与外接矩形之间的白色区域)均标记为 $k+1$ 。这样,只要能把多边形区域内孔洞外的标记为 $k+1$ 的部分去掉,剩下的就是孔洞区域,将孔洞区域用柑橘区域的颜色(此处为黑色)填充即可完成孔洞填充。

3) 找出矩形内标记为 $k+1$ 的区域,判断各个区域是否为孔洞。判断准则是该区域是否有像素与矩形边缘相连,如果有,说明不是孔洞,否则,就是孔洞。将非孔洞区域标记为另一个值 $k+2$,而多边形区域内的孔洞由于不与矩形边界相连,所以保留原来的标记值 $k+1$,由此可见,该算法可适用于任意形状区域的孔洞填充。

4) 用柑橘区域的颜色(黑色)填充矩形区域中标记为 $k+1$ 的孔洞像素。这样,多边形区域中的孔洞就一次性地全部被填充了^[42]。

用上述方法对图3-8c的目标图像进行填充后的结果如图3-8d所示。

3.2.7 基于面积和最小外接矩形的目标区域提取

柑橘成熟时呈现橘黄色,在田间拍摄的图像中,虽然背景中大部分是绿色的枝叶,但也有少量枯萎的黄色枝叶、黄色土地等与柑橘颜色接近的背景区域,而且当图像中成熟柑橘被遮挡部分超过60%时,如图3-8d中左上方的小区域,由于识别误差太大,也应该视作背景区域去除。前面所述的方法无法去除这些区域,但这些区域形状往往与圆形差别较大,而柑橘区域呈现圆形或者近圆形,根据这个特点,本研究使用外接矩形法去除非目标区域。

首先求出每个连通区域的面积(像素总数)和最小外接矩形(四条边分别在水

平和垂直方向), 根据外接矩形的长宽比以及连通区域面积和外接矩形面积之比去除图像中背景区域, 综合考虑图像面积与外接矩形面积之比的范围, 去除不在此范围内的区域, 考虑到柑橘的具体形状特征, 设定范围如式 3.16, 满足此式者为目标区域, 否则为背景区域, 将其去除。记录下得到的目标区域外接矩形边长值, 方便后面 Hough 变换设定搜索范围。

$$0.7 \leq a/b \leq 1.4 \text{ 且 } S_0/S_i \geq 0.6 \quad (3.16)$$

其中 a , b 为外接矩形的两条边的长度, S_0 为目标区域的面积, S_i 为其外接矩形的面积, 经过处理后一般只留下柑橘图像区域, 其它干扰物体被除去, 成熟柑橘被有效地分割出来。实践证明该分割方法分割准确率较高、适用性较强, 能在复杂的自然场景下准确地识别柑橘。

使用上述方法对 3-8d 处理的结果如图 3-8e 所示, 图像中左上方的小区域被去除。通过上面几步的处理, 图像中成熟柑橘区域被提取出来, 背景区域被完全去除, 为对柑橘的特征提取和定位处理创造了条件。

3.3 果实区域特征提取

根据柑橘果实类圆形的特点, 使用同样大小的圆代表单个成熟柑橘。对于求取柑橘区域的拟合圆, 本研究使用圆形 Hough 变换 (CHT) 法。首先提取柑橘区域轮廓, 然后使用优化的圆形 Hough 变换拟合柑橘轮廓, 最后提取圆心坐标、半径等特征值。

3.3.1 果实区域的轮廓提取

目标物体的边缘对图像识别和计算机分析十分有用。边缘蕴含了丰富的内在信息(如方向、阶跃性质、形状等), 是图像识别中抽取图像特征的重要属性。图像边缘是图像局部特性不连续性(灰度突变、颜色突变等)的反映, 它标志着一个区域的终结和另一个区域的开始。轮廓提取是边界分割中非常重要的一种处理, 图像轮廓提取的方法有边界跟踪法、区域增长法、区域分裂合并法等, 而二值图像轮廓提取的原理非常简单, 就是掏空内部点: 如果原图像中有一点为黑, 且它的 8 个相邻点皆为黑, 则证明它是区域内部点, 将该点删除。具体地说, 在进行轮廓提取的时候, 需使用一个一维数组, 用来记录处理的像素点周围 8 个邻域的

信息,如果像素点的灰度值为 0 且 8 个邻域的像素点的灰度值和中心像素点的灰度值相同,就认为该点是在图像的内部,可以删除,否则,认为该点是在图像的边缘需要保留。依次处理图像中的每一个像素,剩下的就是图像的轮廓^[43,44]。运用该方法处理图 3-8e 得到的轮廓图像如图 3-8f 所示。

3.3.2 利用优化圆形 Hough 变换提取果实特征

在自然环境下,圆形目标检测大多采用基于 Hough 变换的方法。如球场内足球位置的跟踪研究,就是采用 Hough 变换来确定足球位置。在农业方面,如柑橘类水果、卷心菜等的自动定位研究,也采用 Hough 变换算法。

Hough 变换是通过坐标空间的变换,将识别问题转化为空间坐标系中点的聚类问题,即将图像从位置空间转换到描述曲线的参数空间。其优点是:当发生遮挡、光线变化、视点变化和图像采集误差等引起的噪声,使图像边缘不连续或变形时,通过变换仍有效的检测出目标,有较强的抗噪声能力。所以当识别物体含噪声、有数据丢失或冗余时,Hough 变换被认为是最佳识别方法。实践也已证明基于 Hough 变换技术的圆形目标检测方法鲁棒性最好,已在科学各领域实现应用。但其复杂的运算和巨大的数据量严重限制了它的应用和发展。

3.3.2.1 Hough 变换的基本原理

Hough 变换是对图像进行某种形式的坐标变换。它将原始图像中给定形状的曲线或直线变换成 Hough 空间中的一个点,即原始图像中给定形状的曲线或直线上的所有点都集中到变换空间的某个位置,形成峰点。这样,就把原始图像中给定形状的曲线或直线的检测问题,变成在变换空间中寻找峰点的问题,也即把检测整体特性(给定曲线的点集)变成检测局部特性的问题。

Hough 变换的基本原理详见参考文献[45],从 Hough 变换的原理可见,算法不要求边界线连续,如果待检测线上有扰动或者断裂甚至是虚线,进行 Hough 变换后,根据概率大小,在变换空间仍能得到明显的峰点。抗干扰能力强是 Hough 变换的一个突出优点,因此适用于处理来自自然场景中存在失真、遮挡和视点变化引起变形的图像。虽然 Hough 变换具有很突出的特点,但复杂的运算、巨大的数据量制约着它的应用,噪声的沉积也会使输出失去意义,因此必须处理好几个关键问题:

- 1) 合理的图像分割,提取有效目标轮廓,减少 Hough 变换的数据量,以降

低噪声干扰，提高变换速度；

- 2) 在变换域中选取多大的区间作为单元设置累加器，累加器单元的大小就是检测算法的置信区间的大小，将直接影响检测结果；
- 3) 选择合理的搜索半径范围，以减少不必要的无用变换，提高变换速度；
- 4) 使用相应算法搜索峰值，以准确找到目标圆。

本研究运用 OpenCV 的 Hough 变换程序，为了解决以上问题，根据柑橘区域的外接矩形尺寸设定搜索半径范围、累加器尺寸等，提高了变换的速度和准确度。OpenCV 是 Intel 公司开发的开源机器视觉类，里面包含了常用的机器视觉相关处理，如图像处理与识别、视频处理、运动物体跟踪等。OpenCV 的圆形 Hough 变换函数原型如下：

```
CvSeq* cvHoughCircles(CvArr* image, void* circle_storage, int method,
                      double dp, double min_dist, double param1=100,
                      double param2=100, int min_radius=0, int
                      max_radius=0);
```

各参数的意义如下：

image 输入的 8 位单通道灰度图像，本实验输入为 8 位的二值轮廓图像；
circle_storage 检测到的圆存储于此。圆的参数被记录在此变量中。如果圆的存储是一个矩阵且实际检测到的圆超出了矩阵的存储能力，那么超出的可能性最大的一些圆被返回。每个圆使用三个浮点数表示：中心坐标 (x, y) 和半径。

method Hough 变换方法。目前，唯一实现的方法是 CV_HOUGH_GRADIENT，它就是在 H.K.Yuen 于 2003 年发表的文章中介绍的 2IHT 方法。所以此参数是固定的，为 CV_HOUGH_GRADIENT。

dp 用来检测圆心累加器的分辨率。如果它的取值为 1，则累加器将与输入图像具有相同的分辨率；如果它的取值为 2，则累加器的长和宽只有输入图像的一半，本研究为节约时间提高速度，此处取值为 2。

min_dist 检测到的圆心之间的最小距离。如果此参数的取值太小，则在只有一个圆的情况下会误检测到多个相邻的圆；如果太大，则有些圆

(如重叠比较大或者两个距离很近的目标)不能被检出。

param1 第一个与方法相关的参数。在 **CV_HOUGH_GRADIENT** 情况下，此参数是传递给 **Canny** 边缘检测算子的两个阈值中较高的那一个(低阈值为高阈值的一半)。本实验设为 100。

param2 第二个与方法相关的参数。在 **CV_HOUGH_GRADIENT** 情况下，此参数是在圆心检测阶段的累加器阈值。其取值越小，将会误检到越多本不存在的圆；其取值越大，对应于累加器中最大累加值的圆将首先被返回。

min_radius 被搜索的圆的最小半径，小于此半径的圆被忽略。

max_radius 被搜索的圆的最大半径，默认情况下，最大半径被设置为 $\max(\text{image_width}, \text{image_height})$ 。

本研究根据柑橘区域的外接矩形尺寸设定 **min_dist**, **param2**, **min_radius**, **max_radius** 四个参数的具体值，以提高 Hough 变换的速度和准确度^[46]。

3.3.2.2 累加器阈值设置

Hough 变换中，累加器阈值实际是圆形目标检测的容错区间，其大小影响 Hough 变换的精度。累加器阈值取值越小，圆形目标检测越全面，但对噪声敏感，导致一个轮廓检测出很多个圆，同时累加器阈值过小，由于返回的圆形目标参数太多，会消耗过多的存储空间；反之如果累加器阈值设置越大，可以去除由噪声引起的虚假目标，但会导致圆度稍低的目标会作为噪声被去除。图 3-9 所示为 Hough 变换时累加器阈值设置过大和过小时的情况。因此，应根据需要选择合理的累加器单元尺寸。



a 累加器阈值过大时的 Hough 变换结果 b 累加器阈值过小时的 Hough 变换结果

图 3-9 累加器阈值过小和过大时的图像 Hough 变换结果

Fig.3-9 Result of Hough Transform when the threshold of accumulator is too large or too small

一定大小的柑橘，摄像机拍摄的位置不同，图像上的柑橘区域的大小也不相同。例如近距离拍摄得到的图像与远距离所拍图像相比较，相同柑桔的区域会变大，其轮廓线也会相应的变长，且轮廓线的长短与区域外接矩形的边长呈线性正相关。本文根据这一关系，将目标区域的尺寸作为累加器阈值设置依据，根据柑橘外接矩形尺寸设置累加器阈值。公式如下：

$$\text{param2} = \frac{\sum_{i=1}^k (a_i + b_i)}{8k} \quad (3.17)$$

其中 k 为图像中柑橘区域的个数； a_i ， b_i 为 3.2.7 节中求得的每个区域外接矩形的边长，所有数字的单位都为像素数。根据上述公式求得的 param2 值作为 Hough 变换函数中参数 param2 的值。这样，累加器阈值会随着区域的尺寸变化而变化，提高了检测的鲁棒性。

3.3.2.3 搜索半径范围和圆心距离设定

图像中某一个边界点，可能有不同半径和不同圆心的圆通过，因此在进行 Hough 变换时，要对所有的可能性进行“投票”，这种投票在时间和空间上的消耗相当大。因为圆心位置可以是图像上所有的像素点，半径的取值范围为 1 至图像的宽度或高度。因此，在 OpenCV 的圆形 Hough 变换中半径的默认搜索范围为 0 到图像的宽度或者高度，这种“地毯式”遍历搜索方法耗时并且输出结果也无意义。为缩短处理时间，可根据目标区域的具体情况设置半径和圆心位置的搜索范围，半径的搜索范围和圆心距离可通过柑橘区域外接矩形的边长来确定，确定方法如式(3.18)~(3.20)所示：

$$\text{min_radius} = \min\{a_i, b_i\}/4 \quad (3.18)$$

$$\text{max_radius} = \max\{a_i, b_i\} * 1.5 \quad (3.19)$$

$$\text{min_dist} = \min\{a_i, b_i\}/2 \quad (3.20)$$

其中， a_i ， b_i 分别为图像中各区域外接矩形的长或宽（单位：像素），数字 4，1.5 和 2 是由实验统计得出的参数。经过优化，搜索范围得到有效控制，提高了变换速度，并且阈值选择和参数设定均由计算机自动完成，无需人工干预，实现了智能化的要求。根据以上设定的参数对图像进行圆形 Hough 变换的结果如图

3-10 所示。由图可知，算法能正确找到果实中心，拟合圆半径和果实半径基本吻合。



图 3-10 圆形 Hough 变换法提取区域拟合圆

Fig.3-10 Use Circular Hough Transformation (CHT) to extract regional approaching round

3.3.3 实验结果及分析

本章前面表述了成熟柑橘区域提取及使用 Hough 变换法提取区域特征，本研究使用上述方法分别对 67 幅图像进行了处理，67 幅图像总共包含 363 个柑橘区域，其中无遮挡的区域 204 个，正确识别 194 个，识别率为 95.1%；部分遮挡的区域为 159 个，正确识别 133 个，识别率 83.6%。识别结果如表 3-1 所示。

表 3-1 成熟柑橘识别情况

Tab.3-1 Result of mature citrus recognition

类别	综合情况	无遮挡	部分遮挡
区域总数	363	204	159
正确识别	327	194	133
识别率%	90.1	95.1	83.6

从表 3-1 可知，本研究所用方法识别率高，由于 Hough 变换鲁棒性较强，当果实区域存在重叠或遮挡时识别率仍然达到了 83.6%，错误识别(不能够准确且完整表述其特征的图像均认为是错误识别)的原因分析：

1) 成熟度不够、反光等因素导致识别出的目标残缺部分较多，使算法无法识别完整果实，图 3-11a 表示由于成熟度不够造成的拟合半径变大、圆心位置上移；图 3-11b 示意了因反光造成拟合半径变大，圆心位置上移。实验发现，基于 2R-G-B 色差图的迭代阈值分割算法，对顺光、晴阴天等不同条件下的图像识别

率较高,对逆光和阳光直射下的识别率相对较差,尤其是果实表面有较大反光时,特征提取准确率下降明显;

2) 优化的圆形 Hough 变换算法,比较准确提取出类球形果实的圆心坐标和半径等特征,但是当果实区域残缺大于 $1/2$ 或果实发生扭曲、变形等形状与圆形相差较大情况严重时,提取的轮廓不规则,准确率会明显下降;

3) 遮挡、重叠过大或者存在多重重叠(遮挡或重叠部分超出果实整体的 $1/2$),使较小的目标区域经外接矩形算法操作后,误当成背景去除没有被识别,图 3-11c 示意了遮挡过大造成的拟合半径变小及识别不到的情况,图 3-11d 表示因果实之间重叠过大,使较小目标区域在经过外接矩形运算后作为背景被误去除。

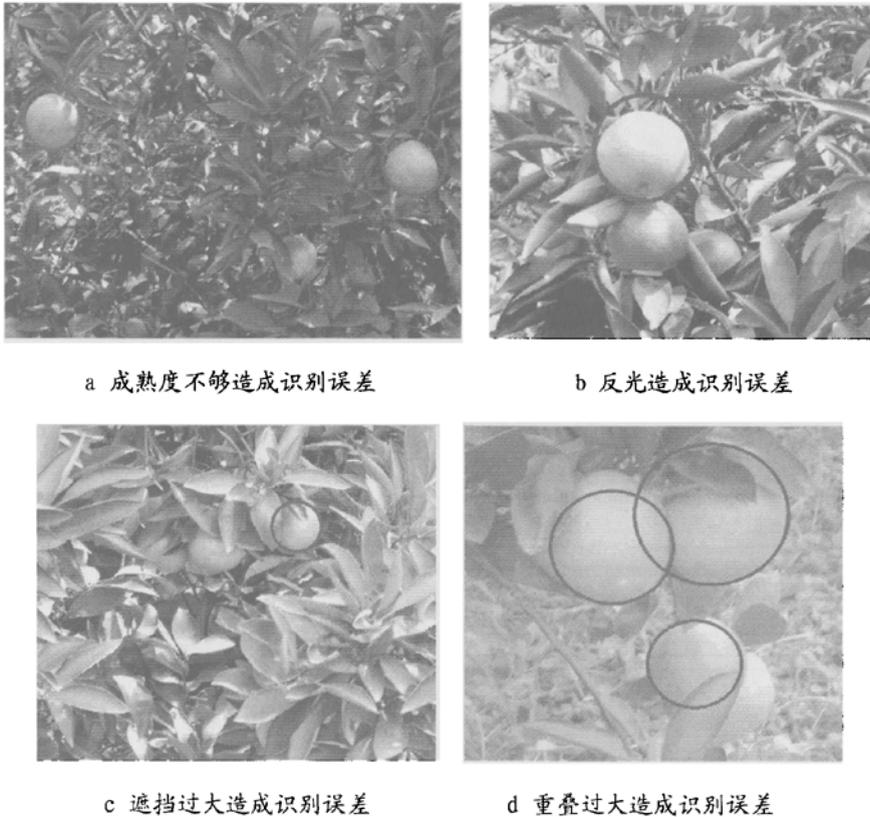


图 3-11 错误识别的原因分析

Fig.3-11 Cause analysis of error recognition

3.4 成熟柑橘区域立体匹配和空间定位

把柑橘从图像中识别出来之后，只是得到了柑橘在图像平面上的二维信息，柑橘在三维空间的信息还未知，要想实现机器人对柑橘的自动采摘，必须对柑橘进行空间定位。早期对水果的研究主要集中在识别方法和提高识别率方面，未涉及定位和采摘方法研究，近些年来才开始有较多涉及。定位算法主要采用双目立体视觉系统或“手眼”系统，有些研究还采用激光、超声、红外测距仪、结构光等。本研究采用双目立体视觉技术来实现成熟柑橘的空间定位。根据双目立体视觉的原理，结合图像中柑橘区域的特点，采用基于特征的立体匹配方法匹配立体图像对，得到目标在立体图像对中的视差，根据视差值得目标三维空间坐标。

3.4.1 双目立体视觉成像模型

本文采用相同焦距的镜头和相同型号的 CCD 摄像机(详见第二章)，两个摄像机的光轴互相平行，因此可以用类同求取简单的空间点深度的计算方法。由于光轴与图像平面垂直，两个摄像机的图像坐标系的 x 轴重合， y 轴互相平行，如图 3-12 所示。从理论上讲，如果将第一个摄像机沿其 x 轴平移一段距离后，会与第二个摄像机完全重合。

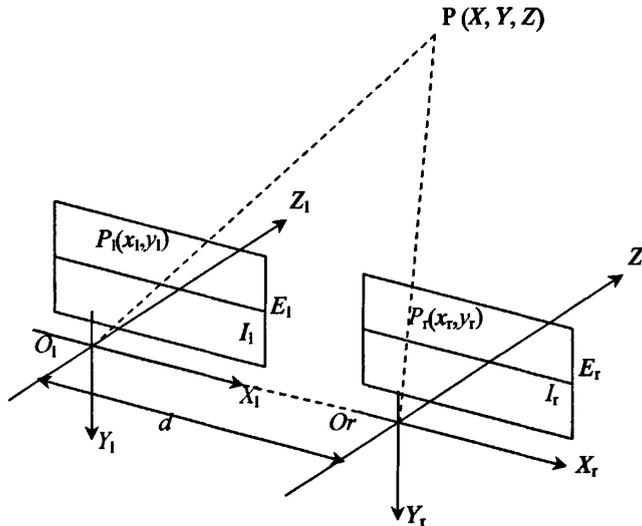


图 3-12 双摄像机配置图

Fig.3-12 Configuration diagram of two-camera

在如图 3-12 所示的 $O_l P O_r$ 平面上, P 为空间任意点, O_l 和 O_r 分别为左、右摄像机坐标系中心, 则中心连线 $O_l O_r$ 与 X 轴重合。若 $O_l P O_r$ 平面与两个图像平面的交线分别为 E_l 和 E_r , 由于两个图像平面位于同一平面, 则 E_l 和 E_r 为该平面上的同一直线, 又由于图像平面平行于 x 轴, 故 E_l 和 E_r 与 x 轴平行。设 P_l 为 $O_l P$ 与 I_l 的交点, P_r 为 $O_r P$ 与 I_r 的交点, 则 P_l 与 P_r 应分别位于 E_l 和 E_r 上, 此时 E_l 和 E_r 称之为极线, 这种约束关系就是所谓的极线约束^[7]。

3.4.2 双目立体视觉的深度计算

由于特征提取后的特征点的坐标值是在图像坐标系下, 以像素为单位表示, 因此要将坐标值转换到成像平面坐标系下, 以 mm 形式表示, 这是就成像平面坐标系中的公式的逆过程, 转换公式为:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & s & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (3.21)$$

其中: f_x , f_y 为摄像机镜头沿 x 和 y 方向的焦距, 以像素点表示, (u_0, v_0) 为主点坐标, s 为歪斜系数, 这些值均为通过对摄像机标定得到。利用上式分别对左右图像中的对应点进行计算, 可得到 (x_l, y_l) 和 (x_r, y_r) 。

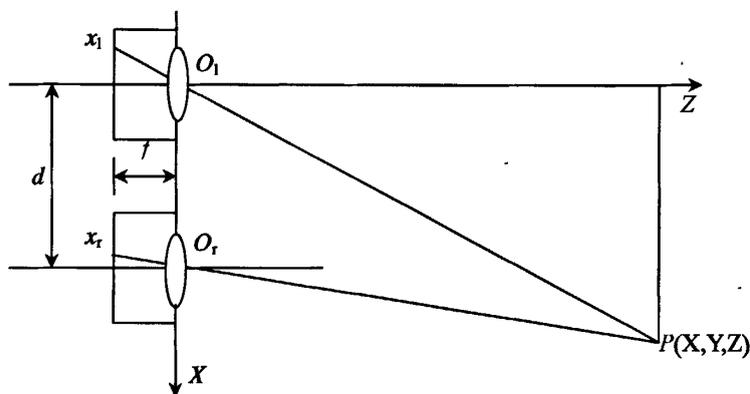


图 3-13 三角形深度计算示意图

Fig.3-13 Schematic diagram of triangular depth calculation

经标定本研究的摄像机间距 $d \approx 100.8 \text{ mm}$ ，焦距 $f_1 = 12.269 \text{ mm}$ ， $f_2 = 12.245 \text{ mm}$ ，取二者的平均值，焦距为 $f = 12.26 \text{ mm}$ 。为便于计算，设世界坐标系和左摄像机坐标系重合(如图 3-13 所示)，空间任一点在世界坐标系下的坐标为 $P(X,Y,Z)$ ，在左、右摄像机坐标系下的坐标分别为 (x_l, y_l) 和 (x_r, y_r) ，则其对应关系可用两对相似三角形计算得到，公式如下：

$$\frac{X}{x_l} = \frac{X - d}{x_r} \Rightarrow X = \frac{x_l \cdot d}{x_l - x_r} \quad (3.22)$$

由于摄像机置于同一高度，根据共极线约束 $y_l = y_r$ ，则：

$$\frac{Y}{y_l} = \frac{d}{(x_l - x_r)} \Rightarrow Y = \frac{y_l \cdot d}{x_l - x_r} \quad (3.23)$$

$$\frac{Z}{f} = \frac{d}{(x_l - x_r)} \Rightarrow Z = \frac{f \cdot d}{x_l - x_r} \quad (3.24)$$

在上式中， d 称为基线 (baseline) 长度， $x_l - x_r$ 称为视差 (disparity)。

分析表明，基线长度越长，空间位置计算相对误差越小，但基线长度不可太长，否则由于物体各部分的互相遮挡等原因，两个摄像机可能不能同时观察到 P 点。视差是由于双摄像机位置不同使 P 点在图像中的投影点的位置不同所引起的，由式(3.24)可见， P 点的距离越远，视差就会越小。当 P 点趋于无穷远时， $O_l P$ 和 $O_r P$ 趋于平行，视差趋于零。

3.4.3 空间定位的实现

从以上介绍可以看出，如果知道了某一点在图像平面上的像的坐标，只要求出该点的 Z 坐标，该点在空间中的位置即可找到。所以要想根据图像求某一点在空间的位置，最重要的是求该点的 Z 坐标，也就是求该点到观察者的距离。在本研究中，研究目标成熟柑橘近似于一个球体。要实现对柑橘的自动采摘，只需求出柑橘上具有代表性的某一点的空间位置坐标即可。对于球体，在几何意义上质心应该是最具有代表性的点。但是质心位于柑橘内部，而柑橘大小有差异，在实际应用中，不能选择质心作为代表柑橘的点，若选择成熟柑橘图像轮廓线经过优化后的圆形 Hough 变换拟合出的圆心，用该点的坐标代替柑橘在空间的位

置, 这样既简化了计算, 又利于实现自动化采摘。根据前面的分析可知, 求取特征点的坐标主要是求取特征点到观察者的距离, 而要得到此距离, 必须首先得到特征点在立体图像对中的视差, 要得到视差则先要对特征点进行立体匹配。

3.4.4 立体匹配

3.4.4.1 立体匹配概述

立体匹配是双目视图中最重要、最关键也是最困难的一步。与普通的图像配准不同, 立体像对之间的差异是由拍摄时观察点的不同引起的, 而不是由其它原因, 如景物本身的变化、运动所引起的。当空间三维场景被投影为二维图像时, 同一景物在不同视点下的图像会有很大不同, 而且场景中的诸多因素, 如光照条件、景物几何形状、物理特性、噪声干扰、畸变以及摄像机特性等, 都被综合成单一的图像中的灰度值。因此, 要准确地对包含了如此之多不利因素的图像进行无歧义的匹配, 显然是十分困难的。立体视觉匹配的研究基本分为两个方向: 从理解人类视觉的立体融合机制出发, 建立一种通用的双眼视觉模型; 从实际应用和要求出发, 建立一种专用的面向对象的立体视觉系统。对于任何一种立体匹配方法, 其有效性有赖于 3 个问题的解决, 即: 选择正确的匹配特征、寻找特征间的本质属性和建立能正确匹配所选特征的稳定算法。立体匹配的研究都是围绕着这三方面开展工作, 并已提出了大量各具特色的匹配方法。

根据匹配基元的不同, 立体视觉匹配算法通常分为三大类: 基于区域匹配, 基于特征匹配和基于相位匹配^[48]。这三类算法因匹配基元的不同, 判断对应点匹配的理论依据也有所不同。虽然各类匹配算法基元不同, 但他们之间还是存在有一些共有的约束条件, 例如:

- 1) 极线约束: 在这种约束条件下, 两幅图像的匹配点位于同一条外极线上;
- 2) 唯一性约束: 两幅图像中的匹配点必须唯一, 即一一对应;
- 3) 几何学约束: 特征点在左右图像中的横坐标值满足 $x_l > x_r$;
- 4) 保序性约束: 即匹配点在两幅图像中的位置关系必须一致。

上述三类匹配方法中: 基于特征的方法主要是通过提取特征点、线或边缘等特征, 然后进行匹配。因为这些特征对噪声不太敏感, 所以可以得到比较精确的匹配, 但基于特征方法只能得到稀疏的视差图。如今很多研究人员利用插值或表

面重建等方法来得到高密度的深度图，但这在大多数情况下都得不到很好的效果；基于区域方法虽可以得到高密度深度图，但在遮挡区域、无纹理区域以及深度不连续处也得不到正确的结果，而且基于区域方法的出发点是，在相关窗口内的像素点视差相同，但当视差变化比较大时，相关算法误匹配的可能性很大；基于相位匹配方法，首先假设图像对中的对应点相位相等，然后进行基元匹配，其匹配基元是相位，当假设条件不成立时匹配就会失去意义，所以它一般只能得到景物的粗糙结构，而且在一些傅立叶相位区域尚需进行特殊处理。总之，这些匹配算法各有优劣，良好的匹配特征应具有可区分性、不变性、稳定性、唯一性以及有效解决歧义匹配的能力。立体视觉的最终目的是为了恢复景物可视表面的完整信息，目前，随着数学算法的不断发展，以及计算机技术的飞速进步，新的更加完善的匹配算法也将会越来越多，如基于模拟退火技术的图像匹配、基于遗传算法的图像匹配、基于概率松弛迭代算法的图像匹配等^[49~51]，但是无论是哪种匹配方法都还不能恢复出所有图像点的视差。

3.4.4.2 本研究所采用的匹配算法

基于区域和特征的匹配方法已有多年的研究，算法也相对成熟，而基于相位的匹配方法是近十多年才发展起来的，相应的配准算法还在不断摸索和完善当中，本文进行的是应用性研究，因此主要对基于区域和基于特征的匹配算法进行了实验。结合成熟柑橘实际情况，采用了基于特征的立体匹配方法，为避免误差传递和扩大，本文将各单个原始柑橘区域(原始图像识别出的成熟柑橘区域)作为整体，提取区域的重心坐标、外接矩形尺寸、面积等特征作为匹配约束条件，使用竞争机制，根据特征创建竞争函数，通过候选区域之间的竞争达到准确匹配。在成熟柑橘匹配过程中，因为要匹配的特征是图像中各个柑橘区域的，因此匹配的图像只是一个稀疏的矩阵，相对而言降低了匹配难度提高了匹配准确度。区域的外接矩形尺寸已在 3.2.7 节中得到，这里还需得到柑橘区域的面积和重心坐标。

1) 区域面积

图像区域一个最基本的特性就是区域面积，它描述区域的大小。如果区域用 R 表示，则二值图像区域 R 的面积 A 计算公式如 3.25 所示：

$$A = \sum_{(x,y) \in R} 1 \quad (3.25)$$

2) 区域重心

为了进一步描述区域特征, 先定义矩的概念。对一个二维连续图像 $f(x, y)$, 可以定义其 $(p+q)$ 阶矩为:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad p, q = 0, 1, 2, \dots \quad (3.26)$$

矩描述也是一种区域描述方式。其特点是不需要知道边界, 只需要知道分割出来的区域内部像素情况即可。对一个区域而言, 矩描述不受平移、旋转、变化等的影响。由式 3-25 描述的区域面积实际上是一个零阶矩 $A = m_{00}$ ($f(x, y) = 1$)。而一阶矩可表示为:

$$m_{10} = \int_{-\infty}^{\infty} x dx, \quad m_{01} = \int_{-\infty}^{\infty} y dy \quad (3.27)$$

区域重心是一种区域全局特性的描述符, 其坐标是根据所有属于区域的点计算出来的, 其计算公式如下:

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}} \quad (3.28)$$

对于实际的图像处理应用而言, 都应该用 $\bar{x} = \frac{m_{10}}{m_{00}}$, $\bar{y} = \frac{m_{01}}{m_{00}}$ 的离散形式。将 p 、 q 或为 1 或为 0 代入, $f(x, y) = 1$ 代入, 并改写为离散形式, 则重心的坐标为:

$$\bar{x} = \frac{1}{A} \sum_{(x,y) \in R} x \quad (3.29)$$

$$\bar{y} = \frac{1}{A} \sum_{(x,y) \in R} y \quad (3.30)$$

得到区域的面积和重心坐标后, 结合 3.2.7 节得到的外接矩形尺寸, 即可通过这些特征值设置匹配约束条件和竞争函数进行立体匹配。

基本约束条件:

$$\text{几何学约束: } x_l > x_r \quad (3.31)$$

$$\text{共极线约束: } |y_l - y_r| \leq Y_{\max} \quad (3.32)$$

其中 x_l , x_r 分别为区域重心在左右图像中的横坐标, y_l , y_r 为区域重心在左右图像中的纵坐标, 根据立体匹配的几何学约束原理, 特征点在左图像中的横坐标必须大于在右图像中的横坐标, 且由于双目立体视觉是基于共极线约束, 即摄像机光轴平行, 并在同一水平面上, 因此搜索空间被限制在沿 y 方向的允差带上, Y_{\max} 是由于镜头及安装引起的垂直方向的误差范围。这些值均需要从摄像机标定时得到, 根据测量对象的深度范围, 设定一定的置信区间。满足以上约束条件的区域才能作为匹配的候选区域。

竞争函数:

$$G = 2\Delta a + 2\Delta b + 2\Delta A + \Delta y \quad (3.33)$$

其中, Δa , Δb 分别表示左右图像中区域外接矩形横坐标和纵坐标方向上的边长之差, ΔA 为区域面积之差, Δy 为区域重心纵坐标之差, 根据竞争函数的定义可知, 两个区域的差别越小, 函数值越小, 因此, 可根据候选匹配区域与待匹配区域的竞争函数值, 得到待匹配区域的最佳匹配。

具体实现过程如下:

1) 用 3.2.5 节介绍的区域标记法将左右图像中的目标区域编号, 并分别求出各目标区域的面积(像素总数)及其重心坐标, 单位为像素;

2) 选择左图像中一个未匹配目标区域, 在右图像中, 逐个考察未被匹配的各个区域, 考察步骤为: 首先考察两个区域之间是否满足基本约束条件(即考察是否满足式 3.31 和 3.32), 若不满足, 则本区域考察完毕, 选择下一区域, 若右图中没有一个区域满足基本约束条件, 则右图中没有区域与左图中待匹配区域匹配, 则将左图中待匹配区域标记为已匹配状态; 若满足, 则计算该区域的竞争函数 G 值, 然后选择右图中下一个未匹配区域进行考察。若右图中所有未匹配区域考察完毕, 则转到步骤(3);

3) 当右图中各个未被匹配的区域考察结束之后, 得到各区域的竞争函数 G 值, 比较各区域的 G 值大小, G 值最小的区域作为左图待匹配区域的匹配结果, 并将这两个区域标记为已匹配状态, 然后转到步骤(2), 继续左图下一区域的匹配, 直到左图中所有区域匹配完毕;

4) 输出两幅图像的匹配结果。

输出结果的形式如图 3-14c 所示, 其中, ΔX 为两区域拟合圆心的横坐标之差, ΔY 为纵坐标之差, 有了这两个值之后, 即可根据公式得到果实目标的三维坐标。

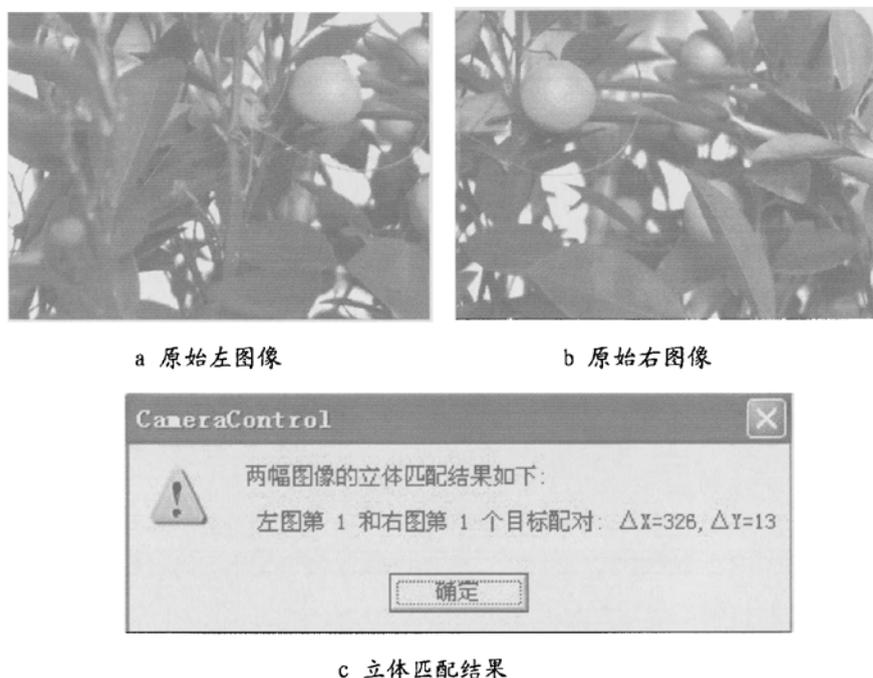


图 3-14 区域立体匹配

Fig.3-14 Regional stereo matching

3.4.4.3 实验结果及分析

成熟柑橘的空间匹配与定位是在识别的基础上进行的, 识别的准确率和特征参数提取的精度对目标匹配的准确性有着直接的影响。从原始立体图像对到最终的立体匹配结果, 经历了识别和匹配两个过程, 故最终的匹配误差实际上是识别误差和匹配误差的叠加(最终识别和匹配的准确率近似等于两个步骤误差的乘积)。本研究使用上述匹配方法分别对 67 对图像进行了识别和匹配, 67 对图像总共包含 304 个柑橘区域(左右图像中都存在的区域), 其中无遮挡的区域 173 个, 正确识别和匹配 152 个, 识别率为 87.9%; 部分遮挡的区域为 131 个, 正确识别和匹配 94 个, 识别率 71.8%; 总体识别率为 80.9%。识别结果如表 3-2 所示。错误匹配的主要原因分析:

1) 果实的成熟度不够、反光等因素使左右图像中识别出的成熟柑橘区域差别较大, 导致不符合匹配条件而无法匹配;

2) 因为叶子、枝干的遮挡、果实之间重叠过大或者存在多重重叠（遮挡重叠部分超出果实整体的 1/2）等原因，破坏了果实的轮廓，误把目标当成背景去除导致匹配错误，或左右图像中同一目标重叠或被遮挡的部位不同，导致左右图像中识别出的目标区域差别较大，从而导致匹配错误；

3) 拍摄的角度不同、基线过宽（基线长度的取值范围 8-20cm），使左右图像中果实个数不同，导致了图像无法匹配；

4) 摄像机校正的方法和精度等引起的误差。

表 3-2 成熟柑橘识别和匹配结果

Tab.3-2 Results of mature citrus recognition and stereo matching

类别	综合情况	无遮挡	部分遮挡
区域总数	304	173	131
正确匹配	246	152	94
识别率%	80.9	87.9	71.8

3.4.5 成熟柑橘的三维坐标恢复

在得到空间任一点在两个图像中的对应坐标、摄像机参数矩阵和摄像机的成像模型条件下，即可进行空间点三维坐标的求解。一旦建立精确的对应关系，距离的计算相对而言只是一个简单的三角计算。本研究中，根据识别出来的成熟柑橘拟合圆心点的坐标及半径值，可以判断出目标在空间的三维位置。

在求柑橘的 X、Y 和 Z 三个坐标中，X、Y 坐标根据图像识别结果容易求出，为简化实验，仅考虑求 Z 坐标。公式 (3.24) 仅适用于当研究对象是理想球体时的情况。在进行实际采摘的过程中，还要考虑到果实半径 R 的影响，在成熟柑橘定位研究中，研究对象是不规则的半球体，R 值通过圆形 Hough 变换得到。为准确起见，Z 的坐标用如下公式表示：

$$Z = \frac{f \cdot d}{(x_l - x_r)} + R \tag{3.34}$$

式中 f 和 R 是常数，Z 取决于基线长度 d 和视差 $x_l - x_r$ 。从理论上来说，基线长度 d 值大一些，当 $x_l - x_r$ 有较小误差时求出的 Z 值偏差不大。但是实际应用中，d 不能太大，因为在复杂的自然场景条件下，存在枝叶遮挡、果实重叠等问

题, 如果 d 太大, 会造成图像采集时一个摄像头可以拍摄到柑橘, 另一个摄像头无法拍摄到柑橘或者无法获取相同柑橘信息的现象, 这样对利用双目立体视觉进行定位造成很大影响, 所以基线长度 d 的取值应该慎重考虑。

在具体的研究过程中, 是以像素为单位进行求解的, 像素宽度可以通过图像宽度除以图像每行像素的个数 (或以图像的高度除以图像每列像素个数) 计算得到。建立一个固定的比例关系 k , 摄像头 f 是一个经过标定得到的数值。则公式 (3.22)、(3.23) 和式 (3.34) 可以转化为 (3.35)~(3.37):

$$X = \frac{x_l \cdot d}{x_l - x_r} \quad (3.35)$$

$$Y = \frac{y_l \cdot d}{x_l - x_r} \quad (3.36)$$

$$Z = \frac{f \cdot d}{k(x_l - x_r)} + R \quad (3.37)$$

3.4.6 实验方案

实验在室内进行, 白天利用太阳光作为光源, 晚上利用室内日光灯作为照明设备, 图像采集无特殊照明设备。摄像头和目标间距设定为 900mm~1500mm。为保证摄像头物理参数的一致, 使用同一型号经过标定后的两个摄像头 (摄像头预先经过标定, 标定工作由本课题组另一成员负责), 两个摄像头使用 USB 接口同时连接到电脑上, 实验过程中保持摄像头基线距离、焦距等参数不变。两个摄像头对同一场景同时拍摄, 得到的图像储存在计算机硬盘中, 同时用激光测距仪测量目标到基准点的距离。用计算得到的数据和测距仪实际测得的数据比较, 即可得到测量准确度。

3.4.7 实验所得数据

为了验证基于立体视觉进行成熟柑橘识别定位理论的正确性, 进行了多次实验。以下是一次实验结果 (在距离摄像机镜头不同的 9 个位置拍摄 9 组柑橘图像, 同时用激光测距仪记录不同位置的测量距离), 数据记录与处理结果如表 3-3 和图 3.15 所示:

表 3-3 检测结果数据记录 (单位: mm)

Tab.3-3 Data of detection results

测量距离	计算距离	误差	百分比误差
1000	1004	-4	-0.4%
1100	1104	4	0.36%
1200	1202	2	0.17%
1300	1298	-2	-0.15%
1400	1396	-4	-0.29%
1500	1508	8	0.53%
1600	1575	-25	-1.56%
1700	1711	11	0.65%
1800	1836	36	2.00%

表 3-3 中计算距离是把每组图像识别的结果代入 Z 坐标的标定方程得到的, 误差是激光测距仪测量距离与计算距离之差的绝对值, 百分比误差是误差与测量距离之比, 测量距离与误差的关系如图 3-15 所示。

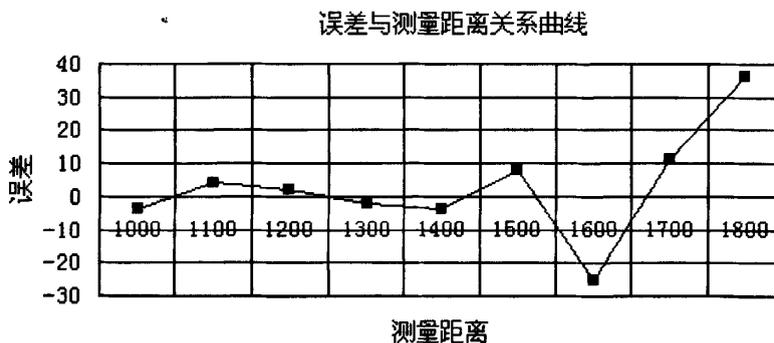


图 3-15 误差与测量距离关系曲线

Fig.3-15 The curve of relationship of error and measuring distance

结合图 3-15 和表 3-3 可以看出: 当测量距离小于 1500mm 时, 误差基本上都在 $\pm 4\text{mm}$ 以内。当测量距离大于(或等于)1500mm 时, 误差变大, 这是由于当摄像机镜头与柑橘之间的实际距离超过 1.5m 时, 采集到的图像上柑橘变小, 极小的误差就会造成最终计算结果极大的偏差, 所以在研究中, 一般设定摄像机镜

头与柑橘之间的实际距离在 1.5m 以内。

3.4.8 结论及分析

多次实验验证的结果表明,本研究方法对于柑橘空间位置的确定是可行的。当测量距离介于 1.0~1.5m 时,距离误差基本上都在 $\pm 4\text{mm}$ 以内。误差可能由以下原因引起:

1) 成熟度不够、反光等因素导致识别出的目标残缺较多,使算法无法识别完整果实,导致拟合圆心、半径值差异较大;

2) 柑桔为类球形,不是绝对意义上的球体,使用 Hough 变换算法得到的圆心、半径值只是近似值;

3) 当测量距离大于 1.5 米时,由于摄像机(摄像机镜头的焦距所限)并不是严格遵守小孔成像模型,所以误差可能较大。

4) 由于基准点是空间一点,无法满足基准点和测距仪测距起始点绝对重合;

5) 系统标定的精度不够,由于摄像机镜头存在多种畸变,镜头各个部分光学参数不一致。

3.5 本章小结

本章首先介绍了成熟柑橘的图像识别方法,详细介绍了识别各步骤的处理过程、实验结果及分析讨论;然后对常用的双目立体视觉技术进行了介绍和分析,根据常用双目立体视觉技术的难点,使用了一种新的匹配算法,并且对空间坐标公式进行了校正,通过实验验证了该理论的正确性及可行性,最后对该理论在实际应用中遇到的问题做了分析和总结。

第四章 障碍物检测

柑橘类水果的果树比较高，具有比较粗大坚硬的树枝，由于树上果实位置的不确定性，当采摘位于粗大树枝后面的柑橘时，若没有检测到果实前面的树枝，则机械手路径规划时将视果实为无障碍状态，机械手按照这一状态规划出的路径进行采摘时，极有可能碰上树枝，从而对机械手造成损害。本研究针对这一问题，进行了树枝障碍物的检测研究，通过提取和处理障碍物图像的骨架恢复障碍物三维信息。采用图像分割、形态学处理、区域细化等步骤提取障碍物骨架；然后找出骨架中端点、分支点等特征点并记录它们的连接关系；最后通过对特征点的立体匹配恢复障碍物的三维信息，试验表明障碍物的正确识别率为 67.3%。

4.1 图像树枝区域提取

在障碍物检测中，图像分割是难点。能否准确地将障碍物区域从背景中分割出来，将直接影响到后续处理。根据图像中柑橘树枝的颜色特点，参考第三章的图像分割方法，进行了不同颜色系统的对比试验，发现利用对RGB颜色系统的色差分量 $2R-G-B$ 进行迭代自适应阈值分割法去除图像成熟果实、棕色土地等区域，并利用对 $2G-R-B$ 进行迭代分割法去除图像未成熟果实、树叶、嫩树枝等区域，再采用灰度阈值分割法去除图像较亮和较暗区域，将以上处理相结合能快速有效地分割出图像的树枝区域。

4.1.1 颜色分割

从采集的图像分析可知，柑橘图像主要包含树枝、树叶、果实、天空、土地等区域(如图4-1a所示)。由于树枝区域颜色特征不明显，要想直接从图像中提取出树枝区域比较困难，但图像中其他区域的颜色特征却非常明显，如树叶和未成熟柑橘为绿色，成熟柑橘为黄色，天空为白色或蓝色等。利用这一特点，本研究使用去除背景留下目标的方法提取树枝区域。参考第三章的图像分割方法，由于迭代法能自动快速准确地找到图像分割阈值，因此考虑使用迭代阈值分割法去除背景，图4-1b中高亮度区域为成熟柑橘区域，图4-1c中高亮度区域为绿叶区域。

使用迭代法对图4-1b和c的色差图分别进行自动阈值分割，分割结果如图4-1d所示。由图可知，除了一些反光较大的树叶之外，图像中成熟果实和树叶部分基本被去除，而绿色的树枝由于比较娇嫩和柔软，对机械手采摘不会产生危险，也视同非障碍物一并去除。

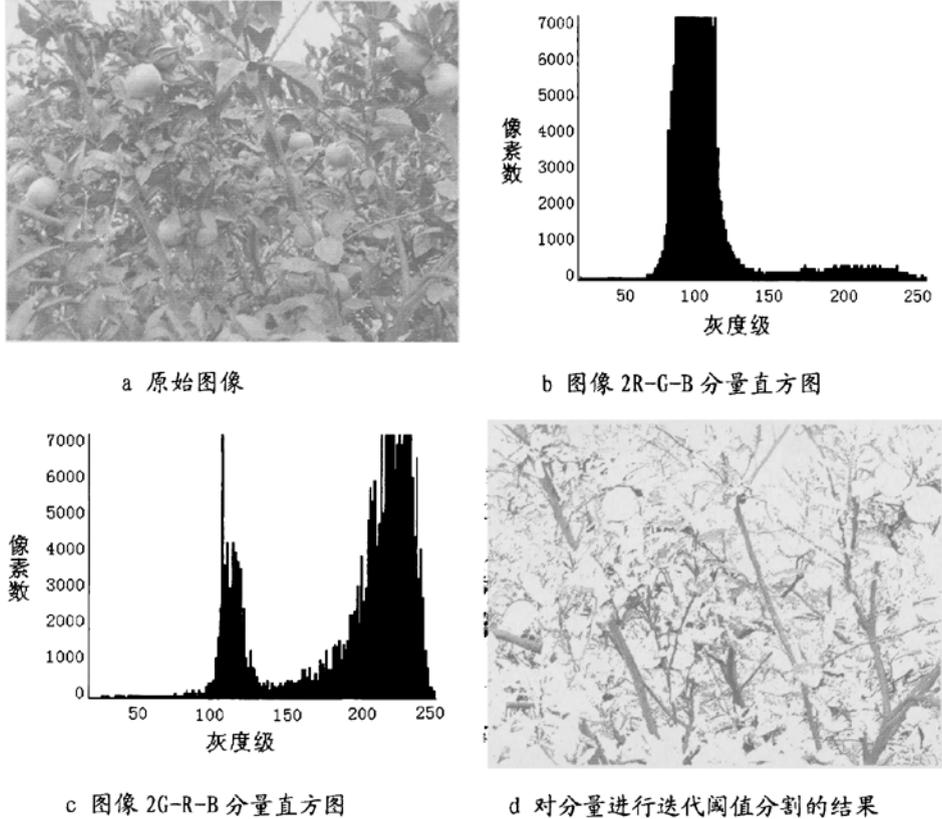


图 4-1 对图像色差的迭代阈值分割

Fig.4-1 Use iteration on chromatism component to scgment image

4.1.2 灰度阈值分割

以上分割利用了像素的颜色分量间的差异，但图像中高亮度和低亮度的区域，如天空和果树深处的阴影由于颜色分量差异较小而无法去除。从采集的图像分析可知柑橘的树枝比较粗糙、不反光，在灰度直方图的中间区域，如图4-2a所示，阴影下的树枝一般在树丛中，不会对机械手采摘构成障碍。因此，将真彩色图像转化为灰度图像，就可以反映图像各像素的明暗程度，转换公式如式(4.1)所示，可将图像中亮度较大和较小的区域去除从而保留有用的树枝区域。

为了提高算法的鲁棒性,经实验统计,以公式(4.2)作为判据能有效分割出障碍物(树枝)。对图4-1d的灰度阈值分割结果如图4-2b所示。

$$f(x, y) = 0.299R + 0.597G + 0.114B \quad (4.1)$$

$$\left. \begin{array}{l} f(x, y) > 1.6 fa \\ f(x, y) < 0.4 fa \end{array} \right\} \text{背景} \quad (4.2)$$

其中 R, G, B 为图像 (x, y) 像素点的颜色分量, fa 为图像平均灰度(即图像各像素点 $f(x, y)$ 的均值)。

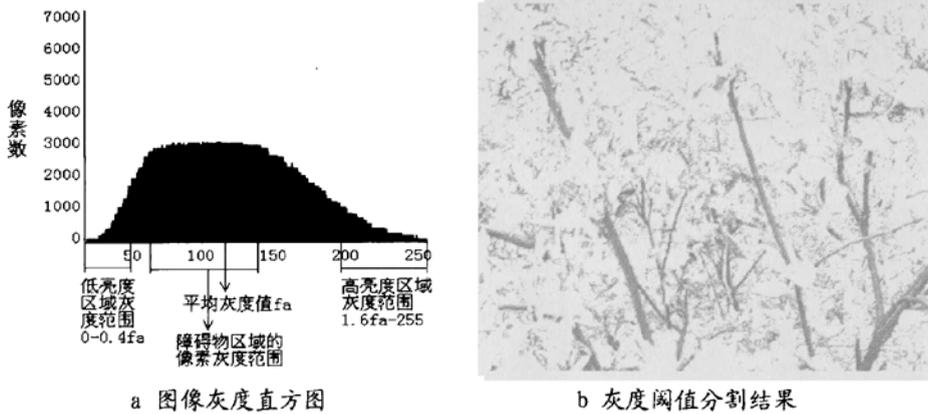


图 4-2 图像灰度阈值分割

Fig.4-2 Image segmentation use gray threshold

经过以上分割处理后,可以提取大部分树枝区域,但由于自然场景下拍摄的图像受外界干扰较大,图像中阴影区域、随机噪声等仍无法完全去除。可采用形态学处理和区域标记法将其除去。

4.1.3 提取树枝区域

分割后的图像,经过二值化、形态学运算、区域标记及小区域去除、区域填充处理提取出图像树枝区域,以上几步的处理方法和第三章中果实区域的相应处理方法相同,故此处不再赘述。最终的树枝区域提取结果如图4-3所示,背景区域被完全去除,绝大部分障碍物(树枝)区域被提取出来。



图 4-3 柑橘树枝区域提取结果

Fig.4-3 Result of citrus branch region extraction

4.2 树枝特征提取

从识别出的树枝区域可以看出，树枝区域形状较复杂，利用此区域来恢复树枝的三维信息显然非常困难，为了准确快速地获取障碍物三维信息，需要提取树枝区域特征，由于正常树枝形状为细长的类圆柱体，因此本研究使用相应的圆柱体代表树枝，由几何知识可知，要想得到一个圆柱体的位置和尺寸，只需得到圆柱上下底面的圆心位置以及半径即可。根据这一思想，本研究首先提取树枝区域骨架，并检测骨架上的特征点(端点、分支点等)，然后利用特征点将骨架分段，将每段树枝看成一段圆柱体，恢复出每段圆柱体的三维信息，也就恢复出了每段树枝的三维信息。

4.2.1 树枝骨架提取

为了对二值图像中各个图形分量(对象物)的形状进行分析，需要有各种能表示图形的特征，骨架就是其中之一。骨架是对象物的核心部分，不同形状的对象物就有不同的骨架^[31]。一般来说，骨架具有三个主要特征：连续性、最小宽度为1和中心性^[52]。从1967年Blum等^[53]首先用中轴表示连续平面上图形以来，有许多骨架求解算法产生，可归纳为骨架子化算法和细化算法两种类型。目前区域骨架的提取方法主要有形态学方法、距离变换法、细线化法等。

4.2.1.1 形态学方法

形态学法是利用一个正方形的结构元素腐蚀待处理区域,将各个腐蚀到空集之前的最后一次腐蚀的结果相加,即得骨架。例如: $S(A)$ 代表 A 的骨架,数字骨架可以从形态学的角度给定义,对于 $n=0,1,2,\dots$, 定义骨架子集 $S_n(A)$ 为图像 A 内所有最大内切圆盘 nB 的圆心构成的集合,从骨架定义可知,骨架是所有骨架子集的并^[54]。它可以表示为:

$$S(A) = \bigcup_{n=0}^K S_n(A) \quad (4.3)$$

式中的 $S_n(A)$ 称为骨架子集,可写成:

$$S_n(A) = (A \otimes nB) - [(A \otimes nB) \cdot B] \quad (4.4)$$

其中 B 为正方形结构元素, (4.3) 式中的 K 代表将 A 腐蚀成空集前的最后一次迭代次数, 即

$$K = \text{MAX}\{n \mid (A \otimes nB) \neq \Phi\} \quad (4.5)$$

形态学骨架提取实例图如图 4-4 所示, 由图可知, 形态学提取法实际上是将各个部分的骨架相加, 这种方法不能保证骨架的连接性, 当骨架区域宽度发生突变或区域边缘不光滑时会造成对连续区域识别出的骨架不连续。

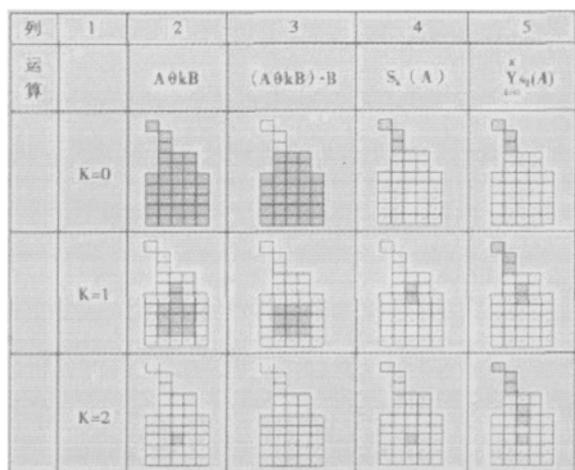


图 4-4 形态学骨架提取示例图

Fig.4-4 Morphological skeleton extraction

4.2.1.2 距离变换法

距离变换法提取区域骨骼通过距离变换和骨骼化两步实现, 该方法首先通过距离变换得到距离图像, 然后比较图像中区域像素的距离值, 将所有距离值

大于或等于邻域中最大距离值的像素的集合作为区域的骨骼。

通过距离变换得到用距离值来表示的图像称为距离图像。所谓距离变换是将二值图像中的1-像素变换为它和0-像素间的最短距离值。两点 $(f_{i,j}$ 与 $f_{m,n})$ 之间的四邻域距离值 d_4 如式(4.6)所示。

$$d_4(f_{i,j}, f_{m,n}) = |i-m| + |j-n| \quad (4.6)$$

获得距离图像需用两次扫描完成，第一次采用顺向扫描，设 $f_{i,j}$ 为原图像像素值， $S_{i,j}$ 为第一次顺向扫描时的运算结果。顺向扫描时，像素点的4个邻点中只有两个已完成距离运算，故只能进行以局部邻域为基础的距离运算。 $S_{i,j}$ 的计算方法如式(4.7)所示。

$$S_{i,j} = \begin{cases} 0 & f_{i,j} = 0 \\ \min\{(S_{i,j-1} + 1), (S_{i-1,j} + 1)\} & f_{i,j} = 1 \end{cases} \quad (4.7)$$

顺向扫描时由于只考虑了局部邻域，因而右下部像素的距离值不是与0-像素的最短距离。第2次采用逆向扫描进行运算，设 $S_{i,j}$ 为第2次扫描前的像素值， $d_{i,j}$ 为第2次运算结果。计算方法如式(4.8)所示。由于 $S_{i,j}$ 是从左上部邻域中求得的局部最短距离值，再加入右下部局部邻域的运算，最终完成距离变换，得到距离图像^[31]。

$$d_{i,j} = \begin{cases} 0 & S_{i,j} = 0 \\ \min\{(d_{i+1,j} + 1), (d_{i,j+1} + 1), S_{i,j}\} & S_{i,j} \neq 0 \end{cases} \quad (4.8)$$

距离图像中各像素距离值的大小直接反映了该像素离开对象物边缘的远近，因此，作为对象物核心部分的骨架提取，只要通过邻域比较找到距离值大的像素部分即可完成。计算方法如下式：

$$d_{i,j} = \begin{cases} d_{i,j} & d_{i,j} \geq \max\{d_{i-1,j}, d_{i+1,j}, d_{i,j-1}, d_{i,j+1}\} \\ 0 & \text{其他} \end{cases} \quad (4.9)$$

距离变换法根据距离图像得到区域骨架点，由于每行(或每列)中距离最大的点仅与其边缘有关，故相邻两行(或两列)的骨架点位置关系并没有必然的联系，当边缘不光滑或者区域宽度发生突变时，会造成连续区域检测到的骨架点不连续的情况。使用距离变换法提取图 4-3 的区域骨架，提取结果如图 4-5 所

示，由图可以看出，提取出的骨架区域不连续，在边缘不光滑处提取的骨架不连续或者产生噪声点，这将会对进一步处理造成困难。虽然距离变换法在本研究中不能很好的得到区域骨架，但得到的距离图像却能很好的反映树枝的粗细（即树枝半径的大小），可用于本研究后面确定树枝的半径。

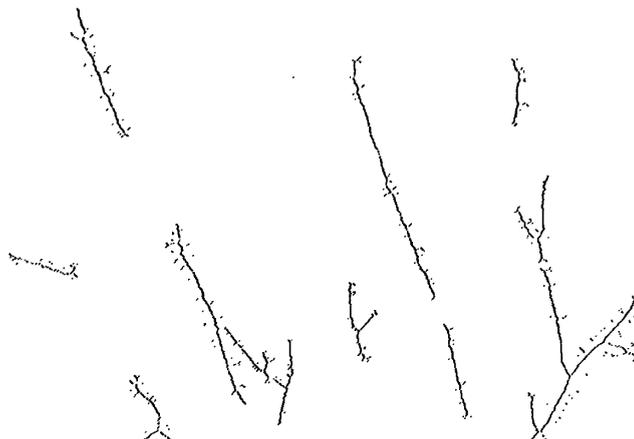


图 4-5 距离变换法提取树枝骨架

Fig.4-5 Extract branch region's skeleton use distance transformation

4.2.1.3 细线化法

区域细线化将区域细化为由线条构成的线图形，线图形有存储量小，便于识别等优点。

为使获得的细线能够准确代表对象物的形状，细线化处理必须满足以下要求：

- 1) 线宽为一个像素；
- 2) 细线位置基本处于原线宽的中心；
- 3) 保持图形的连接性不变，并且不能出现孔和点的新生或消失现象；
- 4) 图形端部基本不缩短。

由此可以看出，区域细线化过程实质上是一个在保持连接性和图形长度不变的前提下求出图形中心线的过程^[31]。

这里在介绍具体的方法之前，首先介绍连接数 N_c 的概念。细线化处理中，保持连接性不变是一个至关重要的要求。所谓连接数不仅用来表示像素之间的连接或非连接关系，还可用于反映各像素与其邻域像素保持连接关系的数量。

连接数 N_c 定义为:在中心为1-像素 x_0 的 3×3 邻域中, 与 x_0 有连接关系的图形分量(不包括 x_0)的个数。采用4-连接和8-连接时连接数分别用 $N_c^{(4)}$ 和 $N_c^{(8)}$ 表示。为了方便, 3×3 邻域看成是一个小图像, 因此也存在有多个图形分量的可能性, 对于4-连接来说(图4-6c), $x_1 - x_2 - x_3$ 和 $x_5 - x_6$ 是两个与 x_0 相连的图形分量, 即 $N_c^{(4)}=2$, 而对于8-连接来说(图4-6d), 由于 x_3 和 x_5 间的八连接关系, 因此, 与 x_0 相连的图形分量仅一个, 即 $N_c^{(8)}=1$ 。连接数可用以下公式计算。

4-连接数:

$$N_c^{(4)} = \sum_{k \in S} (x_k - x_k x_{k+1} x_{k+2}) \quad (4.10)$$

式中 $S=\{1,3,5,7\}$, 并规定 $x_9 = x_1$, 式(4.10)表示了当 k 分别取值为1, 3, 5, 7时的和。以图4-6b为例的计算过程为: $N_c^{(4)} = (1-1) + (1-0) + (1-0) + (0-0) = 2$ 。

8-连接数:

$$N_c^{(8)} = \sum_{k \in S} (\bar{x}_k - \bar{x}_k \bar{x}_{k+1} \bar{x}_{k+2}) \quad (4.11)$$

式中 \bar{x}_k 为 x_k 的取反值, 即 $\bar{x}_k = 1 - x_k$, 其他符号同上。以同图为例的计算过程为: $N_c^{(8)} = (0-0) + (0-0) + (0-0) + (1-0) = 1$ 。

不论是4-连接还是8-连接, 它们的连接数范围都是0~4。

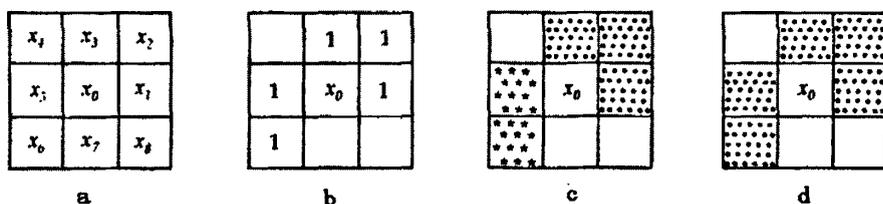


图 4-6 连接数 N_c 的概念

Fig.4-6 Concept of N_c

a 3×3 邻域中像素的代号; b x_0 为 1-像素的 3×3 邻域;

c 4 邻域连接数 $N_c^{(4)}=2$; d 8 邻域连接数 $N_c^{(8)}=1$

为得到图形中心线而进行去除表面像素是有条件的, 其中保持连接性条件

是指该1-像素的删除必须不破坏原图形的连接关系。当 $N_c=2, 3, 4$ 时,说明该1-像素与2个、3个、4个图形分量相连,换句话说该1-像素是这些图形分量的连接点,通过它将这此图形分量连成为一体。因此,如果删除该1-像素,就使原来连接一体的关系遭到分离。当 $N_c=0$ 时,说明该1-像素 x_0 是一个孤立像素点,若将它删除,则必然产生原图像中点的消失。而只有当 $N_c=1$ 时,说明中心1-像素 x_0 不是多个图形分量的连接点,删除它不会破坏原有的连接关系。由此得出保持连接性条件或者称可删除性条件为: $N_c=1$ 。

细线化过程即去除1-像素的过程,被去除的1-像素应满足以下条件:

- 1) 去除的1-像素必须是表面1-像素;
- 2) 满足保持连接性条件,即 $N_c=1$;
- 3) 去除的1-像素不是线图形端点。为了保证图形长度基本不缩短,当该1-像素 x_0 已经成为线图形端点时,尽管它满足 $N_c=1$ 的保持连接性条件,但仍然不能去除。即当中心1-像素 x_0 的邻域中0-像素的个数为7时,该1-像素是线图形的端点。因此,设 K 为中心1-像素的8-邻域中0-像素的个数,非线图形端点的条件为: $K < 7$ 。

实现细线化的具体方法有多种,本研究采用8-连接的横井法细线化实现区域的细线化。

横井法在反复去除图形表面像素时,将图形表面像素分为上下表面像素和左右表面像素两类,然后交替进行顺向扫描去除上下表面像素和逆向扫描去除左右表面像素的处理,直至获得中心线为止。具体步骤如下:

- 1) 按光栅扫描顺序,搜索1-像素(本研究为黑色像素);
- 2) 判别图形表面像素,并加上标记。图形表面像素定义为4邻点表面像素,并规定按上→左→下→右的顺序检查1-像素的4邻域中0-像素的位置,从而确定图形表面像素的种类。设上下表面像素和左右表面像素的标记分别为21和31。反复执行步骤1)和2),直至扫描结束。
- 3) 再次顺向扫描,搜索标记为21的上下表面像素。
- 4) 判别线图形端点,当8邻域中0-像素的个数 $K=7$,则为端点,不满足

去除条件, 并加标记 41; 当 $K < 7$, 则为非端点, 可以继续。

- 5) 计算连接数 $N_c^{(8)}$ 。当 $N_c^{(8)}=0, 2, 3, 4$ 时, 因不满足保持连接性条件, 故不能去除, 同样加标记 41; 当 $N_c^{(8)}=1$, 满足可去除条件, 则此 1-像素变为 0-像素。

反复执行 3), 4), 5) 直至所有标记为 21 的像素全部处理完毕。采用逆向扫描, 搜索标记为 31 的全部左右表面像素, 并按同样步骤执行 4), 5)。至此, 第一层图形表面像素处理完毕。回到步骤 1), 重复上述处理, 直至图像上只有标记为 41 的不可去除的表面像素为止。

- 6) 最后一次扫描, 将标记为 41 的像素全部改为 1-像素。

运用上述细线化方法提取图 4-3 的区域骨架, 提取结果如图 4-7 所示。提取的骨架保持了原区域的连接性, 无中断现象, 效果较好, 本研究选用此方法提取区域骨架。

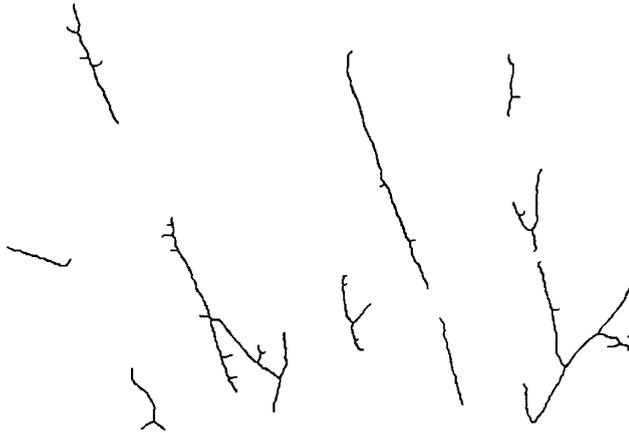


图 4-7 细线化法提取区域骨架

Fig.4-7 Extract branch region's skeleton use graph thinning

4.2.2 骨架特征点提取

树枝骨架图形特征点包括端点和分支点, 分支点为树枝分叉的开始点。特征点可通过八邻域分支点数 $N_b^{(8)}$ 来判别^[31]。特征点与连接数的关系为:

$$N_b^{(8)}=1 \quad \text{端点} \quad (4.12)$$

$$N_b^{(8)}>2 \quad \text{分支点} \quad (4.13)$$

非特征点的 $N_b^{(8)}=2$ 。骨架特征点提取中，分支点数 $N_b^{(8)}$ 的计算是关键，本研究采用对八邻域点逐点检测的方式得到每个骨架点的分支点数。算法具体实现过程如下：

- 1) 从图像左上角开始按光栅扫描顺序，搜索图像骨架点；
- 2) 按顺时针方向逐点检测骨架点的八邻点，检测方式为：首先检测中心点左上方邻点，判断其是否为骨架点，若不是，则跳过；若是，则检查其前面一点是否为骨架点(即其反时针方向的邻点，例如：左上方邻点的前面一点即中心点左边的邻点)，若前面一点也为骨架点，说明此骨架点和前面一点是一条分支上的点，中心点的分支点数不变；若前面一点不是骨架点，说明此点是另一条分支上的点，中心点的分支点数加1，左上点检测完毕，然后用同样的做法按顺时针方向依次检测中心点正上方点、右上方点、右边点……，直到中心点的左边邻点。则整个八邻域点检测完毕，得到中心点的分支点数；
- 3) 继续扫描图像，直到所有骨架点处理完毕。

得到每个骨架点的 $N_b^{(8)}$ 之后，即可根据式(4.12)和(4.13)判断各个骨架点，从而提取出骨架的特征点。

4.2.3 骨架修剪

通过细线化处理得到的树枝骨架，虽然很好的保持了连接性，但当树枝区域边缘不平滑时会产生“假分支”。如图 4-8 所示，从矩形框中的部分可以明显的看出，某些较短的分支并非真正的树枝，而是图像区域在此处的边缘向外突起造成的，这些“假分支”将影响本研究后续处理，需将其去除，目前比较常用的骨架修剪方法是采用形态学方法^[54]。但形态学方法只能修剪较短的分支(不超过三个像素，即毛刺)，本研究中树枝骨架的假分支大部分均超过 3 个像素，因此使用形态学修剪法不能去除“假分支”。本研究使用直接统计分支骨架长度，然后对长度设定阈值去除较短骨架分支的方法去除“假分支”。具体处理过程如下：

- 1) 定义一个与原始图像同样尺寸的矩阵，将原始骨架图像复制一份；
- 2) 在复制的图像上进行处理，将特征点中分支点(即 $N_b^{(8)} > 2$ 的骨架点)的八邻点全部变为背景点，由于分支点是几个分支的连接点，将分支点的八邻点变成背景点之后，这些分支将断开成一条条独立的树枝段；
- 3) 使用八邻域标记法标记每段树枝，并统计每段树枝的像素总数(即该段树枝总长度)；
- 4) 在原始图像上依次判断每段树枝骨架，若骨架的两端点均为分支点，则该段骨架位于树枝主干上，不能去除；若两个端点均为端点，则该段骨架为独立的骨架段，不为分支段，也不能去除；若骨架的两端点中只有一点为端点，则该段骨架为非主干，再判断其长度，若其像素总数小于 15 个，表明该段骨架为“假分支”，将其去除。

树枝骨架经过以上处理后，骨架中毛刺和“假分支”被有效去除，如图 4-9 所示，特征点和分支数大大减少，简化了后续处理。

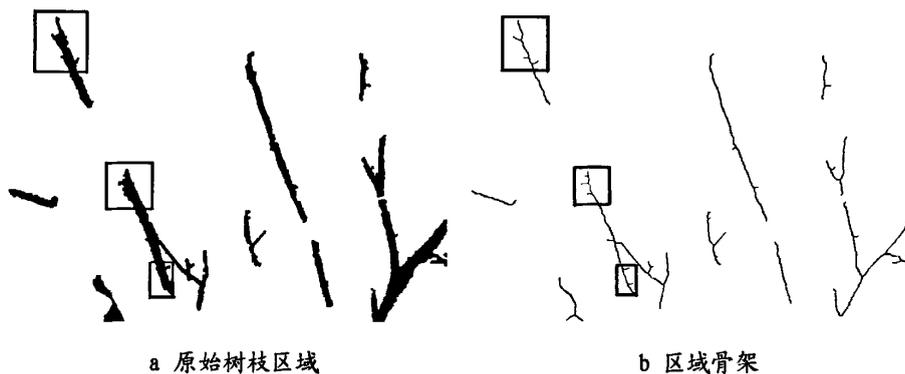


图 4-8 细线化后包含“假分支”的骨架(矩形部分更明显)

Fig.4-8 Skeletons which containing “false branch” after graph thinning (It is more obvious in rectangles)



图 4-9 修剪后的区域骨架

Fig.4-9 Regional skeletons after being pruning

4.2.4 恢复遮挡树枝

在树枝识别时由于遮挡等原因，有时会造成识别出的树枝发生中断的情况，例如图 4-9 中向右下方倾斜约 30 度方向的几条树枝，由于前面有树叶遮挡，从而使识别出的树枝发生中断，但该区域的树枝是实际存在的，机械手路径规划时将视此处为无树枝状态，若机械手从此处通过将树枝相碰，造成机械手的损坏，因此，必须恢复被遮挡的树枝。

本研究通过实验统计发现，当图像中某段树枝有一部分被遮挡而发生中断时，余下的两段骨架斜率基本相同，它们基本在一条直线上，因此，本研究采用直线 Hough 变换法检测图像中的直线段，用检测出的直线使中断的树枝相连，从而恢复遮挡树枝。具体处理过程如下：

- 1) 使用八邻域标记法标记各个联通的骨架，并检测和保存各个骨架的特征点；
- 2) 对骨架图像进行 Hough 变换，为了使检测出的直线能使中断的树枝相连，本研究将检测出的直线段用三个像素粗度划出，对图 4-9 进行 Hough 变换的结果如图 4-10 所示，由图可以看到，中断的树枝被检测出的直线段连接了起来；
- 3) 依次检测各条线段的连接情况，即检测其和哪些编号的骨架相连，若一条线段和两个不同编号的骨架相连，则表明这两个编号的骨架之间有待恢复的骨架(即有骨架中断现象)，记下这些骨架对的编号及将它们连接

在一起的直线段的斜率；

- 4) 分别处理步骤 3) 得到的所有需要连接的骨架对，处理方法为：首先统计出两个骨架中所有的端点 ($N_b^{(8)}=1$ 的点)，然后任意两两组合两个骨架中的端点。例如：A 骨架中有 1、2 两个端点；B 骨架中有 1、2、3 三个端点，组合方式可以为 A 骨架中 1 分别和 B 骨架中 1、2、3 组合，A 骨架中 2 分别和 B 骨架中 1、2、3 组合，一共有 6 种组合方式。每种组合方式中，需要求出两个端点的距离 $D_{i,j}$ 和此两点连线的角度 $\theta_{i,j}$ 。得到两点的 $D_{i,j}$ 和 $\theta_{i,j}$ 值之后，即可计算此两点之间的判别函数值 $T_{i,j}$ ，计算公式如下：

$$D_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (4.14)$$

$$\theta_{i,j} = \arctan\left(\frac{y_j - y_i}{x_j - x_i}\right) \quad (4.15)$$

$$T_{i,j} = D_{i,j} + 100 \times |\theta_i - \theta_{i,j}| \quad (4.16)$$

其中， x_i ， x_j 分别为两端点的横坐标； y_i ， y_j 分别为两端点的纵坐标，在计算 $\theta_{i,j}$ 时，必须使 $y_j > y_i$ ，即图像下方的端点坐标值减去上方的坐标值。为避免分母为零的情况，规定当 $x_i = x_j$ 时， $\theta_{i,j} = \pi/2$ ，式(4.16)中 θ_i 为将这两个骨架连接在一起的直线段的倾斜角。将两个骨架中各种组合的判别函数求出之后，比较他们的大小，判别函数值最小的组合为待连接的两个端点，使用粗细为一个像素的直线段将这两个点相连，即完成了该骨架对的处理。

使用上述方法处理图 4-9 的骨架，处理结果如图 4-11 所示，由图可知，被遮挡而中断的骨架得到了恢复。经过以上处理后，某些骨架端点会因为骨架连接而成为非特征点，因此，此处需要再进行一次特征点检测，并记录下特征点之间的连接情况，两个特征点之间的连线代表一段树枝。



图 4-10 骨架的直线 Hough 变换结果

Fig.4-10 Result of Linear Hough Transform on skeletons



图 4-11 遮挡树枝恢复结果

Fig.4-11 Result of the shielded branches recovery

4.2.5 树枝半径提取

机械手避障过程中，树枝的粗细也是避障的重要信息。根据4.2.1.2节得到的区域的距离图像可得到骨架点与区域边缘的距离，从而得到树枝以像素为单位的半径，经过摄像机标定之后便可得到树枝的实际半径。在距离图像中，骨架点的距离值就是树枝各截面的半径(以像素为单位)，求出两特征点连线上所有骨架点距离值的均值 R ，如式(4.17)所示，用 R 代表该段树枝的半径。

$$R = \frac{\sum_{P \in S} d_1(P)}{N} \quad (4.17)$$

其中， S 代表两特征点连线上的所有骨架点， $d_1(P)$ 为 P 骨架点的距离值， N

为S中骨架点的总数。由于进行了遮挡骨架的恢复处理，恢复出来的骨架点由于在原始树枝区域中不存在，因此没有距离值，本研究在处理含有此种骨架点的骨架段时，实行以部分代表整体的办法，即用有距离值的部分得到的半径代表整个骨架半径。

4.3 立体匹配

为减少匹配运算量，降低误匹配数，以特征点为匹配对象，采用基于特征的立体匹配方法。在遵守匹配三约束的条件下，加入分支点数 $N_b^{(8)}$ 约束，即 $N_b^{(8)}$ 相同的点才能进行匹配。匹配的约束条件如下：

$$1) |y_l - y_r| \leq Y_{\max} \quad (\text{共极线约束})$$

$$2) x_l > x_r \quad (\text{几何约束})$$

$$3) N \leq 1 \quad (\text{唯一性约束})$$

$$4) N_b^{(8)(L)} = N_b^{(8)(R)} \quad (\text{分支点数相等})$$

其中， x_l ， x_r 分别为区域重心在左右图像中的横坐标， y_l ， y_r 为区域重心在左右图像中的纵坐标， N 为一幅图像中的特征点在另一图像中匹配的点的个数。

4.4 障碍物三维信息恢复

匹配后的特征点，首先根据其在左右图像的视差，计算其在世界坐标系中的X，Y，Z坐标，计算公式如式(3.22)~(3.24)所示。得到特征点的实际位置后，结合4.2.4中得到的特征点之间的连接关系和4.2.5节中求得的每段树枝的半径，用两特征点的连线为轴，以该段树枝半径为半径的圆柱体代表该段树枝，即可得到树枝的三维信息。为了确保机械手不碰到树枝，将此处得到的半径乘以1.2作为安全系数。最终的树枝三维信息识别结果如图4-12所示，假设该段圆柱体代表一段树枝，则最终识别出上下底面的圆心坐标在三维空间的位置和圆柱体半径。有了这些信息，机器人即可进行三维建模和路径规划，实现自动采摘。

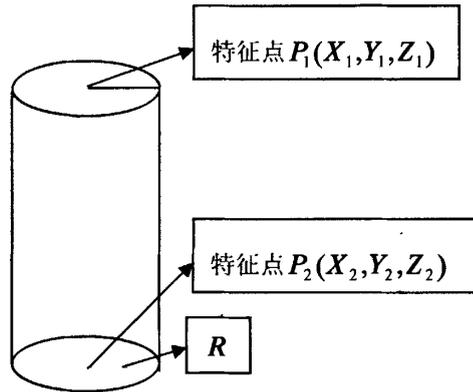


图 4-12 树枝三维信息识别结果

Fig.4-12 Result of recognition of branches' 3-D information

4.5 算法流程及结果分析

障碍物识别的执行流程如图4-13所示。

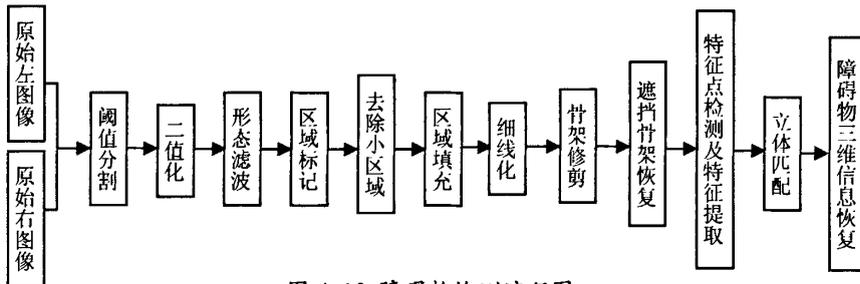


图 4-13 障碍物检测流程图

Fig.4-13 Flowchart of obstacle detection

本研究采用2个美国Lumenera公司的Lu075系列摄像头构成平行立体视觉系统，视觉系统首先通过标定得到相关参数，实验过程中保持摄像头基线距离和焦距等不变。实验地点为镇江市江心洲果林场，果树平均高度约2.5m，试验过程中，程序首先控制左右摄像头拍摄两幅图像；然后对左右图像按图4-13所示的流程进行处理，得到树枝三维信息值；最后实验人员通过激光测距仪测定树枝的实际位置（主要测量树枝特征点如分叉点等处的位置），并与计算得到的数值进行比较，得到识别误差。对于树枝半径，本研究利用游标卡尺检测每段树枝中较粗部分的直径，直径的一半则为半径，并与计算得到的该段树枝半径值比较。本研究按以上方法共处理了58对不同条件下拍摄的树枝图像，共包括211个树枝区域（左右图

像中都存在的树枝区域), 识别结果如表4-1所示。表4-1中的正确识别是指程序得出的树枝三维信息与树枝的实际位置差别在5%以内, 其余为误识别, 误识别的主要原因有:

1) 图像背景过于复杂导致图像区域分割不准确, 如图4-14所示, 图像中褐色土地, 灰色杂草等区域使得图像分割不准确;

2) 光照强度变化太大导致障碍物区域的像素灰度值偏离正常范围, 导致图像分割不准确;

3) 互不相连的两个树枝在图像中有时会产生重叠, 在图像识别时重叠处会产生假特征点, 造成识别不准确;

4) 系统标定误差及实际位置测量过程中的人为误差。

表4-2给出了不同距离(障碍物与两摄像机基线中点的距离)下障碍物的识别平均误差, 由表可知, 随着距离的增大, 由于障碍物在摄像机中的图像变小, 识别误差增大。因此, 实际应用中摄像头与障碍物间的距离应控制在1.5m以内。

表 4-1 柑橘树枝识别结果

Tab.4-1 Result of citrus branch recognition

类别	树枝总数	正确识别	识别率/%
晴天	122	88	72.1
阴天或多云	89	54	60.6
综合情况	211	142	67.3

表 4-2 不同距离下的障碍物识别平均误差

Tab.4-2 Average error of obstacle identification in different distances

距离范围/m	0.5-1.0	1.0-1.5	1.5-2.0	2.0-2.5
识别平均误差/%	3.53	3.77	5.48	6.84

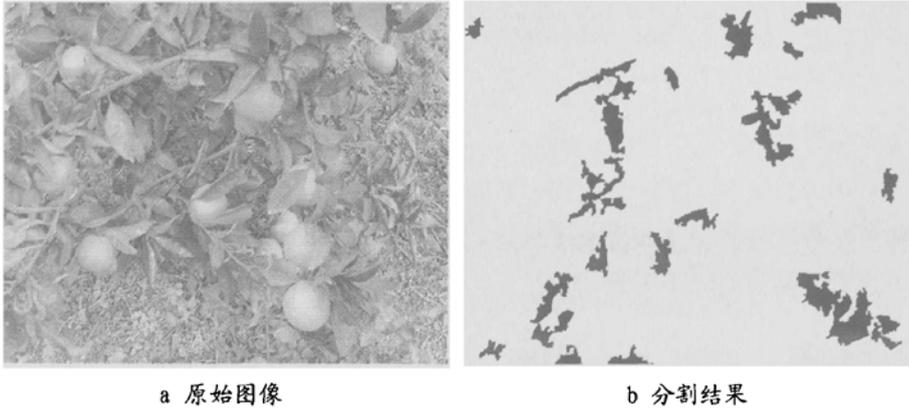


图 4-14 带有复杂背景的图片分割

Fig.4-14 Segmentation on image with complex background

4.6 本章小结

本章介绍了本研究的障碍物检测方法,首先介绍了图像分割及树枝区域识别方法,根据树枝区域的特点,提取树枝骨架,区域骨架提取方法有多种,介绍了三种方法,即形态学方法、距离变换法、细线化法。本研究利用了细线化法提取区域骨架,然后对树枝骨架进行了骨架修剪和遮挡树枝恢复操作,并利用距离变换法求树枝半径。最后通过对树枝特征点的立体匹配恢复树枝三维信息,本章最后部分给出了障碍物检测的算法流程及结果分析。

第五章 结论与展望

国际上,果实收获机器人的研究已有 30 多年,主要有苹果、柑橘、甜瓜、黄瓜、葡萄等收获机器人的研究报道,有些还研制出实验样机,但总体上来讲仍处于实验室研究阶段。其关键原因是机器人在自然场景下工作,光线变化、果实遮挡或粘连等原因导致成熟果实的识别、定位困难,障碍物检测、机械手避障采摘果实的问题也没有得到满意的解决,因此至今还没有研究出实用的果实收获机器人。

本研究基于双目立体视觉技术对自然场景下的成熟柑橘进行识别、匹配与定位,并对收获机器人障碍物(主要是树枝)检测做了初步研究。主要研究成果总结如下:

1) 基于 2R-G-B 色差图的迭代自动阈值分割算法,对光线影响不敏感,能够在顺光、晴阴天等不同条件下比较准确、快速的分割识别出成熟柑桔。

2) 利用优化圆形 Hough 变换算法,比较准确提取出果实的圆心坐标和半径等特征。针对 Hough 变换高耗时问题,提出根据目标区域大小自动设置搜索半径区间和累加器单元尺寸,使变换时间与图像大小基本无关,同时提高了变换的准确度,实验证明算法鲁棒性较强。

3) 以单个成熟柑橘区域为匹配对象,采用原始目标区域的重心、面积以及外接矩形尺寸等特征建立匹配约束条件,同时建立竞争函数,实现图像对中目标区域的最优匹配。利用双目立体视觉技术和三角形原理进行柑橘的空间定位,同时对传统的求取目标空间坐标公式进行了校正处理,提高了准确性。

4) 提出了一种树枝障碍物检测方法,采用区域细线化方法提取树枝骨架,并通过对区域骨架的处理得到树枝三维信息。并通过实验得到了检测准确率,达到了预期的效果,实验证明该方法是可行的。

水果收获机器人的研究是一项任重道远工程。本文针对视觉系统进行了相关研究,在果实定位和障碍物检测方面提出了相应的解决方案,取得了一定的进展,但仍需要改进和提高。

1) 在处理采集图像时,实验发现逆光和阳光直射下的识别率相对较差,尤

其是果实表面有较大反光时,特征提取准确率下降明显,所以对采集方式要注意。

2) 柑橘的尺寸不大,当实际距离大于 1500mm 时,利用本研究中使用的图像采集装置拍摄到的图像上柑橘变得很小,在定位时很小的误差就会导致错误的结果。如果采用更好的图像采集和定位装置(如使用长焦镜头等),应该可以扩大定位的范围和精度。

3) 遮挡、重叠过大或者存在多重重叠(遮挡或重叠部分超出果实整体的 1/2),使较小的目标区域当成背景没有识别,对于这处情况,可以尝试用红外热成像技术或多光谱、高光谱等新技术处理^[55-58],如果采集装置可以透过枝叶采集到它们的图像,根据本研究中的定位方法,也可以确定它们的空间位置。

4) 树枝区域提取时,图像分割是难点,本研究曾考虑使用小波变换,基于纹理的分割方法等,但由于算法处理太慢无法实现实时性而没有采用。如何从图像中准确的分割出树枝区域,将是进一步研究的难题。

5) 探索新的定位方法,除本研究中采用改进的双目立体视觉技术进行定位,还有单目定位、三目定位、超声波定位、激光定位、红外定位、结构光定位等定位的方法。在这些方法中,是否存在更好的定位方法值得探讨。

6) 机器人视觉的实时性问题是一个关键因素,建议改进系统的软硬件设备,采用并行的数据处理方式。如硬件设备可采用高速的 DSP 处理器。

参考文献

- [1] <http://www.china-citrus.cn/> 中国柑橘网.
- [2] 姜丽萍, 陈树人.果实采摘机器人的研究综述.农业装备技术.2006, 32(1): 8-10.
- [3] Shigehiko Hayashi, Katsunobu Ganno, Yukitsugu Ishii, et al. Robotic harvesting system for eggplants[J]. JARQ ,2002, 36 (3) : 163-168.
- [4] MurakamiN, Inoue K, Otsuka K. Selective harvesting robot of cabbage[J]. JSAM , 1995, 2: 24-31.
- [5] Harrell R C ,Adsit P D ,Munilla R D , et al . Robotic picking of citrus [J] . Robotica , 1990 , 8(4) . 269 - 278.
- [6] Jimenez A R ,Ceres R ,Pons J L. A survey of computer vision methods for locating fruit on trees[J] . Transactions of the ASAE ,2000 ,43 (6) :1911- 1920.
- [7] Parrish, E., and A. K. Goksel, 1977. Pictorial pattern recognition applied to fruit harvesting. Transactions of the ASAE 20(5): 822-827;
- [8] D'Esnon,A G, Rabatel G, Pellenc R. A self-propelled robot to pick apples[J]. ASAE paper No.87-1037, 1987.
- [9] Rabatel, G. A vision system for magali, the fruit picking robot[C]. Agricultural Engineering, Paris, London, U.K. Paper 88293, AGENG88, Int, Conf. Ministry of Agriculture, Fisheries and Food 1988.
- [10] A. Dale Whittaker, G. E. Miles, 1987. Fruit location in a partially occluded image. Transactions of the ASAE 30(3): 591-596;
- [11] Illingworth, J., and J.Kittlen. A survey of the hough transform[J]. Computer Vision, Graph. & Image Proc.,1988, 44:87-116
- [12] Slaughter, D., and R. C. Harrel. Color vision in robotic fruit harvesting. Transactions of the ASAE 1987, 30(4): 1144-1148.
- [13] Harrell, R. C., D. C. Slaughter etc. A fruit-tracking system for robotic harvesting. Machine Vision & Appli. 1989, 2: 69-80.
- [14] Harrell, R. C., P. D. Adsit, R. D. Mumlla, and D. C. Slaughten. Robotic picking of citrus. Robotica 1990, 8: 269-278.
- [15] Kassay, L., 1992. Hungarian robotic apple harvester. ASAE Paper No. 92-7042, 1-14. St. Joseph, Mich.: ASAE.
- [16] Buemi, F., M.. Massa, and G. Sandini, 1995. Agrobot: A robotic system for greenhouse operations. In 4th Workshop on Robotics in Agriculture, IARP, Tolouse, 172-184;
- [17] Ceres, R., J. L. Pons, A. R. Jiménez, J. M. Martin, and L. Calderón, 1998. Design and implementation of an aided fruit harvesting robot(Agribot). Industrial Robot 25(5): 337-346;
- [18] Van Henten, E. J., B. A. J. Van Tuijl, J. Hemming, J. G. Kornet, J. Bontsema and E. A. Van

- Os. 2003. Field Test of an Autonomous Cucumber Picking Robot. *Biosystems Engineering* 2003,86(3): 305-313.
- [19] Fujiura T, Ueda K, Hyun Ch S, et al. Vision system for cucumber-harvesting robot [A]. *IFACBio-Robotics, Information Technology and Intelligent Control for Bio-Production Systems*[C]. Sakai, Osaka, Japan:2000.61 - 65.
- [20] D. M. Bulanon, T. Kataoka, Y. Ota, T. Hiroma, 2002. Automatic recognition of Fuji apples at harvest. *Biosystems Engineering* 83(4): 405-412;
- [21] S. Limsiraratana, Y. Lkeda, 2002. On image analysis for harvesting tropical fruits. *SICE02-0824*: 1336-1341;
- [22] 汤修映, 张铁中. 果蔬收获机器人研究综述. *机器人*. 2005, 27 (1): 90-91.
- [23] 杨秀坤, 陈晓光等. 水果表面缺陷自动检测系统中的人工智能方法研究, *农业工程学报*, 1997, 13 (增刊): 242-247;
- [24] 张瑞合, 姬长英等. 计算机视觉技术在番茄收获中的应用, *农业机械学报*, 2001, 32 (5): 50-53;
- [25] 孙明, 凌云, 基于计算机视觉的萝卜幼苗自动识别技术, *农业机械学报*, 2002, 33 (5): 75-77;
- [26] 蔡健荣, 赵杰文. 自然场景下成熟水果的计算机视觉识别. *农业机械学报*. 2005, 36(2):62-64.
- [27] 隋婧, 金伟其. 双目立体视觉技术的实现及其进展. *电子技术应用*. 2004, 10: 4-6.
- [28] 孙明, 凌云等. 基于计算机视觉的萝卜幼苗自动识别技术. *农业机械学报*, 2002, 33 (5): 75-77
- [29] Jimenez A R, Ceres R, Pons J L. A Survey of Computer Vision Methods for Locating Fruit on Trees. *Transaction of the ASAE*, 2000, 43 (6): 1911-1920.
- [30] 应义斌, 章文英, 蒋亦元等. 机器视觉技术在农产品收获和加工自动化中的应用. *农业机械学报*, 2000, 5 (3): 112-115.
- [31] 方如明, 蔡健荣, 许俐. 计算机图像处理技术及其在农业工程中的应用. 清华大学出版社, 1999, 103-104.
- [32] Kenneth R, Castleman 著, 朱志刚等译. 数字图像处理. 电子工业出版社, 1998, 473-483
- [33] 章毓晋. 图像工程[M]. 北京: 清华大学出版社, 1999 179~182.
- [34] 黄敦, 游志胜. 对彩色和亮度通道进行各向异性扩散的彩色图像分割[J]. *计算机工程*, 2002, 28(6): 166~170.
- [35] 蔡建荣, 周小军, 李玉良, 等. 基于机器视觉自然场景下成熟柑橘识别[J]. *农业工程学报*, 2008, 24(1): 175-178.
- [36] 黎妹红, 张其善. 用迭代法求指纹图像中的阈值. *电子技术应用*. 2004, 3: 12-13.
- [37] 章筑晋. 图像处理和分析. 北京: 清华大学出版社, 1999, 254-278.
- [38] Slaughter D C, Harrell R C. Discriminating Fruit for Robotic Harvest Using color in Natural Outdoor Scenes[J]. *Trans of the ASAE*, 1989, 32 (2): 757-763.

- [39] Burger P. Interactive Computer Graphics Functional, Procedural and Device level Methods[M]. Addison Wesley, Massachusetts,1989.
- [40] Pavlidis T. Algorithms for Graphics and Image Processing[M]. Computer Science Press Inc,1982.65-73.
- [41] 孙家广, 杨长贵.计算机图形学.北京: 清华大学出版社, 1998, 178-190.
- [42] 张燕, 曾立波, 吴琼水, 等.一种适用于任意形状区域的快速孔洞填充算法.计算机应用研究.2004, 12: 155-156.
- [43] 杨淑莹.VC++图像处理程序设计[M].第二版.北京: 清华大学出版社; 北京交通大学出版社, 2005.
- [44] 高文, 陈熙霖.计算机视觉——算法与系统原理[M].第一版, 北京: 清华大学出版社, 1999.
- [45] [美]冈萨雷斯, 温茨.数字图像处理[M].第二版.北京: 电子工业出版社, 2003.
- [46] 刘瑞祯, 于仕琪.OpenCV 教程(基础篇).北京: 北京航空航天大学出版社, 2007.
- [47] 马颂德, 张正友.计算机视觉-----计算理论与算法基础[M]. 北京: 科学技术出版社, 1998: 72~78
- [48] 徐奕, 周军, 周源华. 立体匹配技术[J]. 计算机工程与应用, 2003(15):1~5.
- [49] 吕朝辉, 张兆阳, 安平. 一种基于遗传算法的立体匹配. [J] 计算机工程, 2003, 29(20): 24-25, 30.
- [50] 赵丽萍, 王建华, 黄国建, 曾芬芳.模拟退火及其改进算法在图像匹配问题中的应用[J], 电子科学学刊.1996, 18(增刊). 70-75.
- [51] 蔡健荣.基于立体视觉的成熟水果识别定位及机器人路径规划[D],PhD 江苏大学博士学位论文.2005.
- [52] 杨承磊.基于无向图的图像整体骨架表示模型及其算法.计算机学报.2000, 23(3).
- [53] Blum H. A Transformation for Extracting New Description of Shape. Model for the Perception of Speech and Visual. Cambridge, Massachusetts:MITPress,1967:362-380.
- [54] 程志君, 杨德强.基于数学形态学的汉字骨架提取算法.山西电子技术.2008, 3: 41-42.
- [55] Plebe, Alessio and Giorgio Grasso, Localization of spherical fruits for robotic harvesting[J]. Machine Vision and Applications, 2001, 13(2):70-79.
- [56] 周增产, J. Bontsema等. 荷兰黄瓜收获机器人的研究开发.农业工程学报.2001, 17(6).
- [57] 何东健, 杨青等. 农产品分光反射特性及近红外图像处理在农业中的应用. 农业工程学报, 1996, 12(4).
- [58] A.R. Jimenez, R.Ceres, J.L.Pons. A vision system based on a laser range-finder applied to robotic fruit harvesting. Machine Vision and Applications, 2000, 11:321-329.

致 谢

本文是在蔡健荣教授的关心和指导下完成的。值此论文完稿之际，向所有关心和支持我的老师、同学和朋友们表示衷心的感谢！

首先，我向导师蔡健荣教授致以最诚挚的谢意！在三年硕士的学习期间，无论是在学习上还是在生活上，导师都给予了极大的关心和帮助。特别是在该论文的选题、研究方案的设计以及最终的文字修订，导师都给予了精心指导和亲切关怀，为作者提供了大量前沿的研究资料，指明了研究方向，创造了良好的研究条件；学生取得的每一点进步都凝聚了导师的教诲和心血。导师严谨求实的治学精神、精益求精的科学态度和优良高尚的道德品质，是我学习的榜样和终生受用的宝贵财富。

其次，在课题研究过程中，得到了黄星奕教授、陈斌教授、姜松教授、刘伟民教授、邹小波老师、陈全胜老师、孙宗保等老师的悉心指导，在此表示衷心的感谢。另外，感谢计算机视觉工作室的周平、王梁、张庆磊、王慧、辛君伟、郭志明、王峰、李永平、徐占利等师兄们对我的帮助。

最后，感谢我的父母和其他亲人在我求学生涯中给予的一贯支持和鼓励！

硕士期间发表的论文及参与的课题

硕士期间作者发表的论文:

- [1] 蔡健荣, 周小军, 李玉良, 等. 基于机器视觉自然场景下成熟柑橘识别. 农业工程学报. 2008, 24(1):175-178(EI收录)
- [2] 蔡健荣, 周小军, 王锋. 柑橘采摘机器人障碍物识别研究, 农业机械学报(录用)

硕士期间参与的科研项目:

- [1] 国家863项目: 基于立体视觉的果实信息快速提取、定位及机械手避障研究 (2006AA10Z263)
- [2] 国家自然科学基金资助项目: 果实收获机器人的多信息感知方法研究 (30771243)
- [3] 江苏省高校自然科学基金项目: 自然场景下遮挡果实的识别定位方法研究 (05KJB210019)