

# 图像识别 与项目实践

——VC++、MATLAB  
技术实现



• 杨淑莹 著



电子工业出版社  
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

# 图像识别与项目实践

## ——VC++、MATLAB 技术实现

杨淑莹 著

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

## 内 容 简 介

本书是一本有关数字图像处理应用项目开发与实践指导类的教材, 主要介绍图像处理应用项目开发的基本流程、图像识别处理应用项目关键技术。本书直击当今研究热点, 选择有代表性的专题项目, 详细介绍了手写数字识别、邮政编码识别、汽车牌照号码识别、印刷体汉字识别、一维条形码识别、人脸识别、虹膜识别、指纹识别 8 个应用项目的实现方法。同时, 针对每一个项目介绍项目的应用及意义, 该项目的数据特征分析、识别系统设计、图像预处理技术、特征提取技术, 以及识别方法等。书中实例程序的框架结构简单, 代码简洁, 读者可在数字图像处理技术的基础上进一步深化学习内容, 提高实践应用能力和项目开发能力。

本书可作为高等院校计算机工程、信息工程、生物医学工程、智能机器人学、工业自动化、数字图像处理、模式识别及相关学科的教材或参考书, 也可供有关工程技术人员参考。

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有, 侵权必究。

## 图书在版编目(CIP)数据

图像识别与项目实践: VC++、MATLAB 技术实现/杨淑莹著. —北京: 电子工业出版社, 2014.5  
ISBN 978-7-121-22966-4

I. ①图… II. ①杨… III. ①数字图象处理—C 语言—程序设计②数字图象处理—Matlab 软件  
IV. ①TN911.73②TP312

中国版本图书馆 CIP 数据核字 (2014) 第 075055 号

策划编辑: 张 榕

责任编辑: 周宏敏 文字编辑: 张 迪

印 刷: 北京季峰印刷有限公司

装 订: 北京季峰印刷有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 787×1 092 1/16 印张: 18.5 字数: 473.6 千字

印 次: 2014 年 5 月第 1 次印刷

印 数: 3 000 册 定价: 48.00 元



凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 [zltz@phei.com.cn](mailto:zltz@phei.com.cn), 盗版侵权举报请发邮件至 [dbqq@phei.com.cn](mailto:dbqq@phei.com.cn)。

服务热线: (010) 88258888。

# 前言

## <<<<< PREFACE

图像识别技术涉及数字图像处理、模式识别、人工智能、计算机等学科，是一门综合性的技术。自从电子计算机诞生以来，用机器来模拟实现人类所具有的能力便成为一项非常热门且颇具挑战性的研究领域。由于图像信息易于采集，信息相关性多，抗干扰能力强，因此获得了越来越广泛的应用。当前，人类社会已经进入了信息化和数字化时代，随着国家信息化进程的加速，对于图像识别的需求量越来越大，应用也将越来越广泛。

目前，数字图像处理与模式识别的教材以纯学术性为主，十分侧重于理论推导和分析，与实际编程实现和具体工程脱节。因此，带来如下问题。

### 1. 不能胜任高端业务

虽然掌握图像处理技术可应用于图像预处理环节，但是人们对图像数据更多的是应用于模式识别。在学习完图像处理后，还缺乏高端的图像识别技能。

### 2. 实践能力欠缺

传统教学以理论介绍为主，强调理论的体系和概念，忽视技术的联系和灵活应用，对理论的理解仅限于表面认识，很难看到理论的实际应用效果。由于缺乏实践性教材指导，读者面对应用领域，往往不知所措，无从下手，更谈不上创新应用。

### 3. 人才培养和市场需求之间脱节

目前，市场上对人才的需求越来越倾向于具有实践能力和项目开发经验的人才，而这方面需要积累和引导。

本书从在人们日常生活中产生重大影响的项目入手，综合了作者多年从事数字图像处理和模式识别领域研究的成果，结合实际应用项目，阐述了图像识别的相关知识，介绍了开发实例的实现方法。这里，打破单方面介绍数字图像处理和模式识别的理论知识体系，而是将两者与项目实践有机地结合起来，由易到难，循序渐进，介绍项目实现技术。每个项目实例介绍研究意义、背景和要求，并从数据特征分析、图像预处理、特征提取、模式识别 4 个主要方面入手，提供了项目开发技术原理、实现流程、算法的实现步骤，将实现技术落实到编程，提供 VC++ 或 MATLAB 的编程代码。

本书介绍了手写数字识别、邮政编码识别、汽车牌照号码识别、印刷体汉字识别、一维条形码识别、人脸识别、虹膜识别、指纹识别等项目的开发技术。这些项目在社会中广泛应用，对人们的生活产生重大的影响。项目例程框架结构简单，每个项目的具体实现方法都提供编程实例，代码简洁，使读者能够掌握图像识别工程的核心技术，掌握图像识别的方法，为本领域的研究打下坚实的基础，提高独立承担科研项目开发的能力。

本书在天津理工大学教材建设基金的资助下，得以顺利完成，作者表示感谢。参加编写的还有刘婷婷、刘旭鹏、陶冲、申婷婷、叶诚、祁稳杰、谷林、陈品、王博凯、王光彪、冯帆、贾紫娟、牛廷伟、邓飞。最后本书作者对书中所引论文和参考书籍的作者表示感谢。

由于作者水平有限，书中难免会有缺陷和不足之处，敬请广大读者批评和指正！

著 者

# 目录

## <<<< CONTENTS

|                       |      |
|-----------------------|------|
| <b>第 1 章 图像识别概述</b>   | (1)  |
| 1.1 图像识别意义            | (1)  |
| 1.2 图像识别技术            | (3)  |
| 1.3 手写数字识别            | (6)  |
| 1.4 邮政编码识别            | (8)  |
| 1.5 车牌识别简介            | (10) |
| 1.6 印刷体汉字识别           | (13) |
| 1.7 条形码识别             | (16) |
| 1.8 人脸识别              | (18) |
| 1.9 虹膜识别              | (20) |
| 1.10 指纹识别             | (22) |
| 1.11 图像识别系统性能评价       | (24) |
| <b>第 2 章 图像识别关键技术</b> | (27) |
| 2.1 图像识别开发基本流程        | (27) |
| 2.2 图像预处理             | (28) |
| 2.2.1 灰度化             | (29) |
| 2.2.2 灰度变换            | (30) |
| 2.2.3 灰度分布均衡化         | (32) |
| 2.2.4 几何变换            | (33) |
| 2.2.5 去噪              | (34) |
| 2.2.6 锐化              | (36) |
| 2.2.7 形态学处理           | (38) |
| 2.2.8 细化              | (43) |
| 2.2.9 目标物体的轮廓提取       | (45) |
| 2.2.10 分割             | (48) |
| 2.2.11 基于投影法定位        | (51) |
| 2.2.12 测量             | (51) |
| 2.3 变换域处理             | (53) |
| 2.3.1 傅里叶变换           | (54) |
| 2.3.2 Gabor 变换        | (55) |

|       |                   |      |
|-------|-------------------|------|
| 2.3.3 | 小波变换 .....        | (56) |
| 2.4   | 特征提取 .....        | (59) |
| 2.4.1 | 基于颜色的特征提取 .....   | (60) |
| 2.4.2 | 基于纹理的特征提取 .....   | (60) |
| 2.4.3 | 基于形状的特征提取 .....   | (62) |
| 2.4.4 | 基于空间关系的特征提取 ..... | (63) |
| 2.5   | 模式识别 .....        | (63) |
| 2.5.1 | 模式识别简介 .....      | (63) |
| 2.5.2 | 模式识别方法 .....      | (65) |
| 2.5.3 | 模板匹配法 .....       | (67) |

## 第3章 手写数字识别 .....

|     |                    |      |
|-----|--------------------|------|
| 3.1 | 手写数字图像数据特征分析 ..... | (70) |
| 3.2 | 手写数字识别系统设计 .....   | (72) |
| 3.3 | 特征提取 .....         | (73) |
| 3.4 | 手写数字识别 .....       | (77) |

## 第4章 邮政编码识别 .....

|       |                      |      |
|-------|----------------------|------|
| 4.1   | 邮政编码图像数据特征分析 .....   | (81) |
| 4.2   | 邮政编码识别系统设计 .....     | (82) |
| 4.3   | 邮政编码预处理 .....        | (83) |
| 4.3.1 | 去除红色边框 .....         | (84) |
| 4.3.2 | 灰度化与二值化 .....        | (87) |
| 4.3.3 | 基于投影法的编码定位 .....     | (89) |
| 4.3.4 | 数字切割 .....           | (90) |
| 4.4   | 邮政编码样本特征提取与特征库 ..... | (96) |
| 4.4.1 | 邮政编码样本特征提取 .....     | (96) |
| 4.4.2 | 构建邮政编码样本特征库 .....    | (97) |
| 4.5   | 邮政编码识别 .....         | (99) |

## 第5章 汽车牌照号码识别 .....

|       |                    |       |
|-------|--------------------|-------|
| 5.1   | 汽车牌照图像数据特征分析 ..... | (104) |
| 5.2   | 汽车牌照号码识别系统设计 ..... | (105) |
| 5.3   | 图像预处理 .....        | (106) |
| 5.3.1 | 二值化 .....          | (106) |
| 5.3.2 | 去噪 .....           | (111) |
| 5.3.3 | 车牌定位 .....         | (113) |
| 5.3.4 | 车牌图像标准化 .....      | (118) |
| 5.3.5 | 字符分割 .....         | (120) |

|                           |       |
|---------------------------|-------|
| 5.3.6 字符细化 .....          | (124) |
| 5.4 车牌号码识别 .....          | (128) |
| <b>第6章 印刷体汉字识别</b> .....  | (141) |
| 6.1 印刷体汉字图像数据特征分析 .....   | (141) |
| 6.2 汉字识别系统设计 .....        | (142) |
| 6.3 图像预处理 .....           | (142) |
| 6.3.1 二值化 .....           | (142) |
| 6.3.2 消除噪声 .....          | (146) |
| 6.3.3 汉字行切分与字切分 .....     | (148) |
| 6.4 特征提取 .....            | (155) |
| 6.5 汉字识别 .....            | (161) |
| <b>第7章 一维条形码识别</b> .....  | (167) |
| 7.1 一维条形码图像数据特征分析 .....   | (167) |
| 7.2 一维条形码识别系统设计 .....     | (170) |
| 7.3 一维条形码图像预处理 .....      | (171) |
| 7.3.1 灰度化 .....           | (171) |
| 7.3.2 二值化 .....           | (173) |
| 7.3.3 图像校正处理 .....        | (175) |
| 7.3.4 噪声处理 .....          | (178) |
| 7.4 一维条形码识别 .....         | (180) |
| <b>第8章 人脸识别</b> .....     | (189) |
| 8.1 人脸图像数据特征分析 .....      | (189) |
| 8.2 人脸识别系统设计 .....        | (190) |
| 8.3 人脸图像预处理 .....         | (192) |
| 8.3.1 去除背景 .....          | (192) |
| 8.3.2 二值化 .....           | (194) |
| 8.3.3 噪声消除 .....          | (195) |
| 8.4 基于复合多重投影检测的人脸定位 ..... | (197) |
| 8.4.1 复合多重投影检测方法 .....    | (197) |
| 8.4.2 脸部区域定位 .....        | (198) |
| 8.4.3 眼部区域定位 .....        | (202) |
| 8.4.4 嘴部区域定位 .....        | (208) |
| 8.5 特征提取 .....            | (211) |
| 8.6 人脸识别 .....            | (226) |



|                        |       |
|------------------------|-------|
| <b>第9章 虹膜识别</b>        | (229) |
| 9.1 虹膜图像数据特征分析         | (229) |
| 9.2 虹膜识别系统设计           | (230) |
| 9.3 虹膜定位               | (231) |
| 9.3.1 基于感兴趣区域的虹膜快速定位   | (232) |
| 9.3.2 虹膜外圆定位           | (232) |
| 9.3.3 虹膜内圆定位           | (234) |
| 9.4 虹膜区域处理             | (238) |
| 9.4.1 提取虹膜区域           | (238) |
| 9.4.2 虹膜区域极坐标变换        | (240) |
| 9.4.3 虹膜图像规范化          | (243) |
| 9.5 虹膜特征提取             | (244) |
| 9.5.1 二维 Gabor 滤波器     | (244) |
| 9.5.2 虹膜特征提取           | (247) |
| 9.6 虹膜特征降维             | (250) |
| 9.7 虹膜识别               | (254) |
| <b>第10章 指纹识别</b>       | (262) |
| 10.1 指纹识别图像数据特征分析      | (262) |
| 10.2 指纹识别系统设计          | (263) |
| 10.3 指纹图像预处理           | (264) |
| 10.4 指纹图像 Gabor 滤波     | (272) |
| 10.4.1 Gabor 滤波        | (272) |
| 10.4.2 指纹图像 Gabor 滤波方法 | (274) |
| 10.5 指纹特征降维            | (276) |
| 10.6 指纹识别              | (279) |
| <b>参考文献</b>            | (284) |

# 第1章

## 图像识别概述

### 1.1 图像识别意义

图像是指物体的描述信息，数字图像是一个物体的数字表示。视觉是人类感知外部世界最重要的手段，据统计，在人类获取的信息中，视觉信息占 60%，而图像正是人类获取信息的主要途径。因此，和视觉紧密相关的数字图像处理技术的项目开发越来越受到人们的关注，逐渐形成图像识别技术。

随着数字图像处理技术的发展和实际应用的需求，许多问题不要求其输出结果是一幅完整图像的本身，而是将经过一定处理后的图像再分割和描述，提取有效的特征，进而加以判决分类，这种技术就是图像的模式识别。

图像识别技术是利用计算机视觉采集物理对象，以图像数据为基础，让机器模仿人类视觉，自动完成某些信息的处理功能，达到人类所具有的对视觉采集图像进行识别的能力，以代替人去完成图像分类及辨识的任务。对图像识别来说，面对的是二维数据信号或平面图形，除掉与它们各不相同的物理内容，考虑对样品数据分类这一共性来研究的，把同一种共性者归为一类，另一种共性归为另一类。要求在最小的错误概率条件下，使识别的结果尽量与客观物体相符合，具备人所具有的对各种事物与现象进行分析、描述与判断的能力。

图像的识别属于当代计算机科学研究的重要领域，已经发展成为一门独立的学科。这一学科在近几年里，发展十分迅速，应用范围相当广泛，几乎遍及各个领域，从宇航领域扩展到生物医学、信息科学、资源环境科学、天文学、物理学、工业、农业、国防、教育、艺术等各个领域与行业，在国民经济、国防建设、社会治安和社会发展等方面得到广泛应用，对整个社会都产生了深远的影响。目前，光学字符识别（如手写数字识别、邮政编码识别、汽车牌照识别、汉字识别、条形码识别等），以及生物特征识别（如人脸识别、指纹识别、虹膜识别等）已经在人们的日常生活中广泛应用，对经济、军事、文化及人们的日常生活产生重大的影响。



## 1. 光学字符识别 ( Optical Character Recognition, OCR )

光学字符识别使用 OCR 读取设备和智能视觉系统软件, 识别可同时被机器和肉眼读取的文本。OCR 所使用的输入设备可以是任何一种图像采集设备, 如 CCD、扫描仪、数码相机等。通过使用这类采集设备, OCR 系统将书写者自己写好的文字作为图像输入到计算机中, 然后由计算机去识别。光学字符识别技术已经广泛地应用于各种商业应用, 现在又开始运用到自动化任务中。字符识别处理的信息可分为 3 大类: 文字信息识别、数据信息识别和条形码识别。

(1) 文字信息识别对各民族文字书写的或印刷的文本信息进行识别。如汉字识别, 目前已趋向成熟, 并推出了很多应用系统。

(2) 数据信息识别对阿拉伯数字和少量特殊符号组成的各种编号和统计数据识别。如邮政编码、汽车牌照、统计报表、财务报表、银行票据等, 处理这类信息的核心就是数字识别。常见的应用包括对邮局邮件分拣、车牌号码读取、通行证处理、安全文件处理 (支票、财务文件、账单)、消费品包装 (单号、批号、有效期) 及临床应用等。

(3) 条形码识别对由一组按特定编码规则排列的宽度不等的多个黑条和空白组成的信息进行识别。根据条形码的维度, 通常将条形码分类为一维条形码和二维条形码。条形码具有成本低、扫描速度快、识别可靠性高等优点, 同时为了克服条形码不能被人识别的缺点, 又在条形码下面印上了相应的数字和字符, 就成为条形码识别和光学字符识别的双重形式。条形码可以标出物品的生产国、制造厂家、商品名称、生产日期、图书分类号、邮件起止地点、类别、日期等许多信息, 因而在商品流通、图书管理、邮政管理、银行系统等许多领域都得到广泛的应用。

## 2. 生物特征识别

生物特征识别就是采用某种技术和手段对人的身份进行标识, 从而依据该标识对人进行身份识别, 以达到监督、管理和控制目的的一种技术。用于身份识别和个人信息管理的技术和手段层出不穷, 传统的身份鉴定方法包括个人持证。如身份证、工作证、学生证、人密码、磁卡、智能卡、口令密码等, 这些身份验证方法普遍存在易丢失、易破解、易伪造、不易携带等缺点, 而且在安全性和鉴定速度方面也已经不能满足人们的要求, 这些技术虽然方便、快捷, 但其致命的缺点是安全性差、易伪造、易窃取等。近年来, 计算机技术的广泛应用使得使用生物特征识别进行身份识别成为可能。

生物特征识别的方法越来越多地被应用于身份识别领域。生物识别技术 (Biometric Identification Technology) 是指以人体固有的特征为判别标准, 达到精确鉴定人身份的技术。这些固有特征包括人脸、虹膜、指纹、掌纹等, 也被称为生物模态。这些特征除了外伤等特殊情况下一般会伴随人的一生, 而不会改变或者变换很小。生物识别技术对每个个体都具有随身携带性和持久性; 对不同个体具有普遍性和唯一性等优于传统身份识别的特点。基于人类生物特征的识别技术具有安全可靠、特征唯一、不易伪造、不可窃取等优点。

结合计算机技术, 发展起来了众多的基于人类生物特征的人类身份识别技术, 如人脸识别技术、指纹识别技术、虹膜识别技术。这些识别技术具有特征录入较为方便、信息丰富、适用范围广等优点。因此有着广阔的应用前景。

(1) 人脸识别主要通过分析人脸特征进行识别,也是人们最早使用的生物特征识别技术之一,是一种比较直观、友好、更容易被人接受的识别方式。在实际应用中,人脸识别易于使用,无须使用者的主动参与,尤其适于视频监控等应用。但人脸识别的缺点在于稳定性较差,很容易受周围环境、饰物、年龄、表情等干扰,造成错误的识别。另外,对双/多胞胎的鉴别仍然无能为力。

(2) 虹膜识别主要基于虹膜的生理结构,利用虹膜中存在的细丝、斑点、凹点、射线、皱纹和条纹等特征进行识别。据称,没有任何两个虹膜是一样的。虹膜身份认证的可靠性高,其错误接受率和错误拒绝率很低。

(3) 指纹识别主要通过分析指纹的全局特征和局部特征进行识别,常用的特征如指纹中的嵴、谷、终点、分叉点和分歧点等。随着指纹识别技术的发展及指纹采集设备的价格降低,指纹识别不仅广泛应用于司法和商务活动中,也越来越多地用在笔记本电脑、手机、存储器等终端设备中使用。但采集指纹时要求手指保持洁净和光滑,污垢或疤痕都会给识别带来困难。老年人和手工劳动者的指纹由于磨损严重而不易识别。另外,在实际采集中发现,由于在犯罪记录中常使用指纹,导致很多人害怕将指纹记录在案,从心理上不愿意接受这种识别方式。

生物特征识别可应用于社会生活的很多方面,帮助人们更快捷和更方便地解决潜在的安全问题。总结起来,生物特征识别可以在以下几个主要领域得到应用。

(1) 职员或会员管理:内部授权管制、考勤、薪资计算、俱乐部会员确认。

(2) 重要区域的门禁管制:军事基地、枪械库、核能设施、物料放置库房、电脑机房、政府办公室、保密资料室等。

(3) 金融、证券、保险、社会福利机构的身份确认:如柜台提款、自动提款机、保险箱、金库、大额取款客户身份确认、公司提现确认、交易终端客户身份确认、远程交易身份确认、保险受益人等各种社会福利受益人身份确认等。

(4) 个人财产使用管制:如移动电话、笔记本电脑、汽车等。

(5) 门禁安全:包括社区人员进出及访客出入记录。

(6) 社会安全:包括全民证照系统(身份证、暂住证、驾驶证等)、海关出入境管理、刑侦罪犯查缉过滤、法律上罪犯认定等。

(7) 信息安全:电子交易、网络安全、网上银行及电子商务的安全交易等。

(8) 卫生保健:血液管理、公费医疗确认、个人医疗档案管理。

目前,无论是字符识别(如手写数字识别、邮政编码识别、汽车牌照识别、文字识别等)还是人类生物特征识别(如人脸识别、指纹识别、虹膜识别等)的项目开发技术,它们涉及数字图像处理、模式识别、人工智能、智能计算等多个学科领域。随着高科技的发展,这些项目应用已经成为衡量当代高科技水平的重要手段。

## 1.2 图像识别技术

图像识别技术是数字图像处理和模式识别技术相结合的产物,是一门新兴的应用学科。数字图像处理则是指利用计算机或其他数字设备对图像信息进行各种加工和处理,以满足目



标识需求的基础行为。模式识别研究如何用机器来实现人（及某些动物）对事物的学习、识别和判断能力，因而是以满足目标识别的判断行为。在实际应用中需要将这两个学科技术结合应用，相辅相成、相互促进和发展。

为了模拟人类图像识别活动，人们提出了不同的图像识别模型。例如，模板匹配模型。这种模型认为，识别图像中的某个物体，必须在过去的经验中有这个图像对物体的记忆模式，又叫模板，当前的刺激如果能与大脑中的模板相匹配，这个物体也就被识别了。

图像识别的基本过程是抽取代表未知样本模式的本质表达形式（如各种特征）和预先存储在机器中的标准模式表达形式的集合（称为字典）逐一匹配，用一定的准则进行判别，在机器存储的标准模式表达形式的集合中，找出最接近输入样本子模式的表达形式，该表达模式对应的类别就是识别结果。因此，图像识别技术是一种从大量信息和数据出发，在已有经验和认识的基础上，利用计算机和数学推理的方法自动完成图像中物体的识别和评价的过程。

这种采用模板匹配模型的方法就是模式识别。模式识别方法可以大致分为结构模式识别、统计模式识别及两者的结合。结构模式识别利用结构信息的方法进行识别，类似一个逻辑推理器。其主要优点在于适应性强，区分相似能力强。但是，在实际应用中，面临着抗干扰能力差、描述结构复杂、匹配过程的复杂度高等问题。在识别领域中，纯结构模式识别方法已经逐渐衰落。统计模式识别提取待识别模式的一组统计特征，然后按照一定准则所确定的决策函数进行分类判决。统计模式识别是将物体点阵看作一个整体，其所用的特征是从这个整体上经过大量的统计而得到的。统计特征的特点是抗干扰性强，匹配与分类的算法简单，易于实现。

结构模式识别与统计模式识别各有优缺点，随着对于两种方法认识的深入，统计识别与结构识别两种方法正在逐渐融合。网格化特征就是这种结合的产物。图像被均匀地或非均匀地划分为若干区域，称之为“网格”。在每一个网格内寻找各种特征，如笔画点与背景点的比例；交叉点和笔画端点的个数；细化后的笔画的长度、网格部分的笔画密度等。特征的统计以网格为单位，即使个别点的统计有误差也不会造成大的影响，增强了特征的抗干扰性。这种方法正得到日益广泛的应用。

图像识别可简单地分为两个过程：学习（训练）过程和识别过程。学习过程就是让计算机通过样本学习或训练，提取出每个已知类别的模式特征并存储起来，作为标准特征库，即模板库。识别过程中，计算机首先按学习过程中的特征提取方法，提取出输入模式的特征，然后再与标准特征库中的特征进行匹配，匹配程度最大的模式类别即为识别结果。前者是对样本进行特征选择，构建特征库，寻找分类的规律；后者是根据分类规律对未知样本集进行分类和识别。

图像识别过程包括图像采集、图像预处理、特征提取、模式匹配 4 个环节。一个典型的识别过程如图 1-1 所示，下面简要论述其工作流程。

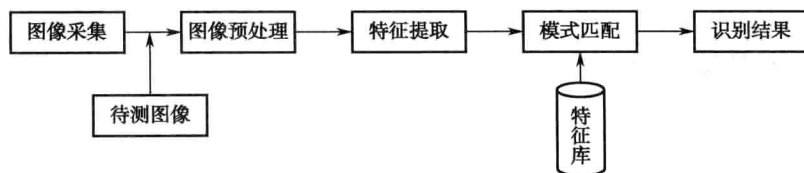


图 1-1 单模态生物识别过程

首先,通过高清摄像机、扫描仪或其他图像采集仪器采集图像的原始信息。图像的采集过程中,由于设备的机械原因或是其他人为因素造成的图像的尺寸、角度、格式、光照强度等的不同,会对以后的操作产生较大的影响,所以要对采集来的原始图像进行预处理操作。图像预处理的作用可以总结为:采用某种手段将图像信息进行归一化,以便于后续处理工作。图像特征提取部分的作用是提取出最能表征一个物体的特征信息,并将其转变成特征向量或矩阵的形式。模式匹配是指系统用待测图像的特征与特征库中的信息进行比对,通过选择合适的分类器达到识别的目的。

## 1. 光学字符识别

如果需要计算机去认识这些已经成为文字的东西,就需要 OCR 技术。比起联机字符识别来,OCR 不要求书写者在特定输入设备上书写,它可以与平常一样书写,所以 OCR 的应用更为广泛。一个典型字符识别过程包括图像采集、图像预处理、字符特征提取和字符识别4个环节。

### 1) 图像采集

OCR 所使用的输入设备可以是任何一种图像采集设备,如 CCD、扫描仪、数字相机等。通过使用这类采集设备,OCR 系统将书写者已写好的文字作为图像输入到计算机中,然后由计算机去识别。由于 OCR 的输入只是简单的一幅图像,它就不能像联机输入那样比较容易地从物理特性上获得字符笔画的顺序信息。因此,OCR 是一个更具挑战性的问题。

### 2) 图像预处理

图像预处理主要包括二值化、噪声去除、倾斜校正、版面分析、字符切割等操作。

(1) 二值化:对于摄像头拍摄的图片,大多数是彩色图像。彩色图像所含的信息量巨大,对于图片的内容,我们可以简单地分为前景与背景。为了让计算机更快和更好地识别文字,需要先对彩色图像进行处理,使图片只有前景信息与背景信息。可以简单地定义前景信息为黑色、背景信息为白色,这就是图像二值化。

(2) 噪声去除:对于不同的文档,对噪声的定义可以不同。根据噪声的特征进行去噪,就叫作噪声去除。

(3) 倾斜校正:由于一般用户,在拍照文档时,都比较随意。因此,拍照出来的图片不可避免地产生倾斜,这就需要文字识别软件进行校正。

(4) 版面分析:将文档图片分段落和分行的过程就叫作版面分析。由于实际文档的多样性和复杂性。因此目前还没有一个固定的和最优的切割模型。

(5) 字符切割:由于拍照条件的限制,经常造成字符粘连和断笔,因此需要文字识别软件有字符切割的功能。

### 3) 字符特征提取

由于文字笔画的粗细、断笔、粘连、旋转等因素的影响,极大地影响了特征提取的难度。

### 4) 字符识别

采用模式识别方法进行字符识别,最常用的方法是模板匹配法。

## 2. 基于视觉的生物识别

基于视觉的生物识别系统利用视觉传感器采集生物图像,如人脸、指纹、虹膜、步态





等，是一种根据人的生理特征或行为特征来识别人身份的模式识别系统。基于视觉的生物识别系统仍然包括图像采集、图像预处理、特征提取和身份识别 4 个环节。

#### 1) 图像采集

用于测定和量化生物特征。

#### 2) 图像预处理

负责对原始数据进行处理，包括图像增强、去背景、目标定位、分割、归一化等措施，为完成特征提取和模式匹配打下基础。

#### 3) 特征提取

如何有效地描述丰富的人体生物特征，并采用合适的方法来提取特征是生物识别的关键部分。

#### 4) 身份识别

决策部分则根据模式匹配的结果做出最终的判决，即确定使用者的身份。所采用的策略往往取决于系统在安全性和实用性等方面的要求。

## 1.3 手写数字识别

### 1. 研究意义

数字是人们日常生活进行信息交流中的不可缺少的重要工具之一，人们经常运用的数据都是阿拉伯数字。随着社会的发展，各国的经济交流日益加深，人们每天要处理大量的票据，都要与大量数据打交道。手写数字识别应用在邮政编码、财务报表和银行票据等方面。

#### 1) 手写数字识别在邮件分拣中的应用

随着我们生活水平的提高和经济活动的发展，通信联系的需求使信函的互换量大幅度增加，我国函件业务量也在不断增长，一些大城市的中心邮局每天的处理量高达几百万件。业务量的急剧上升使得邮件的分拣自动化成为大势所趋。

#### 2) 手写数字识别在财务、税务和金融领域中的应用

财务、税务和金融是手写数字识别大有可为的又一领域。随着我国经济的迅速发展，每天等待处理的财务、税务报表、支票和付款单等越来越多。如果能把它们用计算机自动处理，无疑可以节约大量的时间、金钱和劳力。

#### 3) 手写数字识别在大规模数据统计中的应用

在行业年检和人口普查等需要处理大量字符信息录入的场合有很大的实用价值。这些行业往往需要对手写数字进行大量录入操作，手写数字识别的研究对优化业务流程和提高工作效率有巨大的推动作用。近年来，在这类工作中采用 OCR 技术已成为一种趋势。

手写数字识别作为模式识别领域的一个重要问题，也有着重要的理论价值。

(1) 阿拉伯数字是唯一的被世界各国通用的符号，为各国和各地区的研究工作者提供了一个施展才智的大舞台。对手写数字识别的研究基本上与文化背景无关，在这一领域大家可以探讨和比较各种研究方法。

(2) 由于数字识别的类别数较小，有助于做深入分析及验证一些新的理论。这方面最明

显的例子就是人工神经网络算法以手写数字识别作为具体的实验平台,验证理论的有效性和评价各种方法的优缺点。

(3) 尽管人们对手写数字识别技术已从事了很长时间的研究,并已取得了很多成果,但到目前为止,机器的识别本领还无法与人的认知能力相比,这仍是一个有难度的问题。

(4) 手写数字识别方法很容易推广到其他一些相关问题,一个直接的应用是对英文这样的拼音文字的识别。事实上,很多学者就是把数字和英文字母的识别放在一块儿研究的。

## 2. 手写数字识别系统的构成

手写数字识别的研究对象就是利用计算机系统智能地和准确地识别出手写数字。手写数字识别是图像识别学科下的一个分支,它是图像处理和模式识别领域研究的课题之一,由于其具有很强的实用性,一直是多年来的研究热点。手写数字识别是光学字符识别技术的一个重要分支,包括两类。一类为联机手写数字识别,另一类为脱机手写数字识别。在整个OCR领域中,最为困难的就是脱机手写数字识别。目前,脱机手写数字识别的研究已经取得了较大的进展,有相对完善的识别技术与流程。

脱机手写数字识别系统的构成如下。

### 1) 获取数字图像

手写数字识别时,首先将印(写)在纸上的字符,经光电扫描产生模拟电信号,再通过模数转换把带灰度值的数字信号输入计算机中,再通过所建立的模式识别系统,自动识别出所对应的数字。在此过程中,纸张厚度、洁白度、光洁度、油墨深浅、印刷或书写质量都会造成字形畸变、产生污点、飞白、断笔、交连等干扰。输入设备的鉴别率、线形度、光学畸变、量化过程也要产生噪声。

### 2) 数字图像预处理

在单个字符识别之前,要对带有随机噪声的字符灰度值数字信号进行预处理。对手写数字进行必要的预处理,包括去噪、二值化、细化、图像倾斜校正、图像归一化、提取数字的外界矩形框等操作,将图像带来的外界干扰物排除,得到清晰的图像,这样能够更好地进行手写数字识别操作,所以图像预处理至关重要,直接影响数字识别效果。

### 3) 数字特征提取

提高手写数字识别率的重要方法是提取数字的可靠特征,手写数字特征的提取和选择是决定识别率的关键。当特征值过少时,由于决定性的分类特征太少,使得分类器无法发挥学习分类的功能,造成系统无法辨识。当特征值过多时,除了使系统存储量变大之外,也会因特征值的某些部分与其他特征值冲突,从而造成系统辨识的误差。

### 4) 数字进行识别

采用模式识别技术,设计对应的识别算法,达到识别的目的。

## 3. 识别难点

数字的类别只有10种,笔画又简单,其识别问题表面上是一个较简单的分类问题。但实际上,由于不同的人所写的数字体形态各异,千差万别,手写数字随意性大,书写不规范,经常出现连笔和断笔等现象,甚至同一个人写出的数字也不一定相同。所以,手写数字识别是极其复杂的。而在实际的应用系统中,对数字识别单字识别正确率的要求要比文字苛





刻得多，这是因为数字没有上下文关系，每个单字的识别都至关重要，而且数字识别经常涉及财会和金融领域，特别是有关金额的数字识别时，如支票中填写的金额部分，其严格性更是不言而喻的。因此，用户的要求不是单纯的高正确率，更重要的是极低的、千分之一甚至万分之一以下的误识率。

阿拉伯数字只包括 10 种符号，笔画书写也不复杂，但它的识别难度较大。造成其识别困难的若干原因如下。

- (1) 笔画简单，某些数字的字形相差不大，使得这些手写数字的识别变得相当困难。
- (2) 书写随意，单个数字的变化可能很大。
- (3) 数字之间没有语义相关性，要通过后续处理进行识别结果纠正的可能性很低。

到目前为止，尽管人们在脱机手写英文和汉字识别的研究中已取得很多可喜成就，但距实用还有一定距离。因此，研究高性能的手写数字识别算法是一个有相当挑战性的任务。随着信息化的发展，实现计算机手写数字识别是加快社会信息化进程的关键所在。手写数字识别的应用需求将会更加广泛。一旦研究成功并投入应用，将产生巨大的社会和经济效益，有着重大的现实意义。

## 1.4 邮政编码识别

### 1. 研究意义

世界通信技术飞速发展，有线通信与无线通信形式种类繁多，然而一般信函的通信联系量还是巨大的，而且随着世界人口的增加和经济活动的发展，人们需要信函通信联系的互换量不断增加。目前，世界上已有 40 多个国家先后实行了邮政编码制度，每天都有无数的信件在流通。对于成千上万的信件，依靠手工分拣，不仅速度慢，而且会浪费大量的人力和物力。利用机器自动分拣信件，成为了一个非常有前景的课题，并以此作为衡量一个国家通信技术和邮政服务水平标准之一。

为了实现邮件分拣自动化和邮政网络数字化，我国从 20 世纪 70 年代末开始对字符识别进行研究，主要用于邮政信函分拣的数字识别和计算机输入用的英文、数字和符号的识别。目前，我国城市与农村现代化建设的加速，楼宇建筑林立，居民住房分布比较复杂。所以，对于信函上书写或打印的 6 位邮政编码的扫描识别，在邮件分拣处理中采用先进的自动分拣系统，准确有效地自动识别信封上的邮政编码，可以提高邮件的分组投递速度，加快邮件传递速度，大大节省时间和人力，提高邮政服务质量。因此，进行这方面的研究有着广泛的社会意义和现实意义。具有较高的可靠性。

随着计算机技术日益飞速的发展，邮政编码识别技术在计算机通信、多媒体计算机系统开发、图像的远程通信、电视会议传输、可视电话、视频管理控制、指纹和唇纹识别等研究项目中得到大量应用。图像处理、模式识别、图像识别等技术还在其他多个领域应用广泛，已经渗透到人们的日常生活中，在机器制造、冶金、地质、农林、渔业、天文气象、医疗、教育、邮政、运输、公安、财会、工矿等部门得到广泛应用，并受到重视，在世界各地的国民经济中发挥越来越大的作用。