

Visual C++

数字图像模式识别 技术详解

Digital Image Pattern Recognition Programming Using
Visual C++

冯伟兴 唐墨 贺波◎等编著



CD-ROM



机械工业出版社
China Machine Press

Visual C++

数字图像模式识别 技术详解

Digital Image Pattern Recognition Programming Using
Visual C++

冯伟兴 唐墨 贺波◎等编著



机械工业出版社
China Machine Press

本书内容包括11章：绪论、Visual C++数字图像处理基础、图像特征、统计模式识别、模式识别决策方法及实现，以及人脸检测与特征点定位、汽车牌照识别、脑部医学影像诊断、印刷体汉字识别、手写体数字识别、运动图像分析共6个数字图像模式识别应用实例。系统地介绍了数字图像模式识别技术的基本概念和理论、基本方法和算法，并将图像模式识别的基础理论与Visual C++软件实践相结合。

本书条理清晰、系统全面、由浅入深、实例引导、贴近实用。全书基于数字图像模式识别技术的基本流程，结合实例详细讲解相关基本概念、理论以及几种典型实践和应用，内容翔实、实践性强，其特色在于重点强调了怎样通过开发新代码来增强这些软件工具。另外，本书还配有完整的实例代码光盘，便于学习。

本书可作为普通高等院校计算机、自动化等相关专业数字图像模式识别技术方面的本科生教材，也可作为相关工程技术人员的学习参考用书。

封底无防伪标均为盗版

版权所有，侵权必究

本书法律顾问 北京市展达律师事务所

图书在版编目 (CIP) 数据

Visual C++数字图像模式识别技术详解/冯伟兴等编著. —北京：机械工业出版社，2010.7

ISBN 978-7-111-30949-9

I. V… II. 冯… III. C语言—数字图像处理—程序设计 IV. ①TP391.41 ②TP312

中国版本图书馆CIP数据核字 (2010) 第107538号

机械工业出版社 (北京市西城区百万庄大街22号 邮政编码 100037)

责任编辑：张少波

北京诚信伟业印刷有限公司印刷

2010年9月第1版第1次印刷

184mm×260mm · 21.25印张

标准书号：ISBN 978-7-111-30949-9

ISBN 978-7-89451-561-2 (光盘)

定价：45.00元 (附光盘)

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：(010) 88378991；88361066

购书热线：(010) 68326294；88379649；68995259

投稿热线：(010) 88379604

读者信箱：hzjsj@hzbook.com

作为一门实践性强的综合性边缘学科,数字图像模式识别技术研究的内容主要包括图像变换、图像增强、图像特征提取、图像识别以及运动图像分析等。本书将以实践为导向,以实用为目标来介绍这些重要的数字图像模式识别技术,在介绍数字图像模式识别技术基础理论及算法原理的同时,重点详细地介绍如何用Visual C++编程实现这些典型及常用算法,并结合实际应用,介绍作者所在实验室近年来在工程实践和课题研究中总结出来的一些经典案例,力求理论、应用与实际编程的紧密结合,使读者真正掌握用Visual C++进行图像模式识别编程的基本方法和技巧。

本书内容丰富、层次清晰、力求较强的实践性和可用性。在学习完本书之后,相信读者能够深入掌握数字图像模式识别技术的基础理论和经典算法,并能顺利进行实际项目的开发。

本书特点

本书主要有以下特点。

1. 循序渐进,由浅入深

为了方便读者学习,本书全部实例程序均采用同一个应用程序界面。从基于Visual C++构建应用程序界面,到增加图像处理功能,再到最终的数字图像模式识别经典实例的实现,全书内容前后连贯,互相依托,构成一个整体。使读者高效地掌握基于Visual C++实现数字图像模式识别技术的基本方法。

2. 技术全面,内容充实

本书以理论和编程实践相结合的方式介绍了数字图像模式识别技术的常用算法。按照数字图像模式识别技术的基本体系结构,全面地从数字图像模式识别技术基础知识、数字图像处理应用、数字模式识别技术实例三个层次组织内容,并有机结合了数字图像模式识别技术、软件开发方面的专业知识。

3. 对比讲解,理解深刻

本书针对不是非常熟悉Visual C++编程和数字图像模式识别技术的初学者,采用图文并茂、对比讲解的方式,详细介绍算法实现的每一个步骤。希望通过这种讲解方式,帮助读者加深、加快对Visual C++数字图像模式识别技术的理解和掌握。

4. 代码完整，讲解详尽

书中的每个知识点都有相应的实例代码，并对关键的代码部分进行了注释说明。每段代码的后面都有详细的分析，并给出了代码运行后的结果。读者可以参照运行结果阅读源程序，以便于加深理解。

主要内容

本书共11章，各章的主要内容如下。

第1章：本章介绍了计算机获取、显示、存储数字图像的方法。重点解释了计算机中数字图像的存储格式以及数字图像处理的核心内容。结合模式空间、特征空间和类别空间介绍了模式识别原理，以及数字图像模式识别的工作原理和系统组成。

第2章：系统地介绍了Visual C++作为应用程序编译器的编程思路、编程方法以及如何基于Visual C++ 6.0进行应用程序开发。着重讲述数字图像的特点及其在Windows中的表示方式。在此基础上，介绍了在数字图像模式识别中常用的图像处理基本算法，包括图像增强、形态学运算和图像分割。

第3章：介绍了图像特征的定义及其提取方法。包括图像的统计特征、幅值特征、几何特征、形状特征、纹理特征等基本图像特征的定义及其提取方法。通过实例展示了这几种图像特征提取方法的效果和目的。

第4章：介绍了统计模式识别的主要研究内容，即特征提取与选择、模式分类和模式聚类的研究目的和研究方法。详细介绍了分支界定和基于K-L变换等两种特征提取方法，基于贝叶斯决策、线性分类器和非线性分类等三种模式分类方法，以及模式聚类需要解决的两个问题，即衡量两个样本相似程度的方法和聚类准则。

第5章：介绍了常用的模式识别决策方法。包括人工神经网络的原理及基本实现方法、隐马尔可夫模型的概念及基本算法、决策树的基本概念及设计方法、模板匹配的概念及基于Hausdorff距离的匹配实现方法。在详细介绍这几种决策方法的同时，提供了详尽的实现代码。

第6章：人脸检测与特征点定位应用。介绍了基于Visual C++利用数字图像模式识别技术实现对面脸的自动检测与特征点定位。包括人脸相似度计算、人脸轮廓提取、人脸定位、脸内轮廓提取、眼睛定位、鼻子定位、嘴定位等内容。

第7章：汽车牌照识别应用。按照模式识别系统组成，分车牌预处理、车牌特征提取和车牌识别等三个环节介绍了汽车牌照的自动识别过程。

第8章：脑部医学影像自动诊断应用。介绍了利用灰度共生矩阵进行脑部医学图像纹理特征提取的技术方法，以及基于BP神经网络的分类器实现方法。

第9章：印刷体汉字识别应用。分别介绍了基于统计模式、结构模式和人工神经网络的分类器设计方法及其在汉字识别中的应用，并基于Visual C++利用数字图像模式识别技术实现了印刷体汉字识别。

第10章：手写体数字识别应用。介绍了利用数字图像技术对获得的手写体数字图像进行二值化和反色处理，在已经定位的数字区域上进行特征提取，以及采用模板匹配法对手写体数字进行识别。

第11章：运动图像分析应用。介绍了运动图像分析的主要研究内容及其分析方法，并设计了基于Visual C++在视频中进行动态目标检测和跟踪的应用实例。其中，目标检测采用了帧间差分法，目标跟踪采用了Mean Shift法。

读者对象

- 高等院校的学生
- 社会培训班的学生
- Visual C++ 开发人员
- 数字图像处理技术研究人员
- 模式识别技术研究人员

本书光盘

- 各章实例程序的源代码。
- 与本书内容相关，但由于篇幅所限，未写入本书的内容。

本书由冯伟兴（第1、2、3、4、6、7章）、唐墨（第2、5、9、10、11章）、贺波（第2、8章）编著。冯伟兴、唐墨共同负责全书程序代码的编程和调试。其他参与编著和资料整理的人员有黄晓辉、王宝玉、邹国峰、林天威、马慧、刘靖宇、李阳、阎涛、杨晓飞、宋一兵、管殿柱、付本国、赵景波、王臣业、张忠林等，在此对他们的辛勤工作表示感谢！

感谢您选择了本书，希望我们的努力对您的工作和学习有所帮助，也希望您把对本书的意见和建议告诉我们。

作者联系方式：gdz_zero@126.com

编辑联系方式：sdl@hzbook.com

作者

2010年6月

目 录 CONTENTS

前言

第1章 绪论1

1.1 数字图像处理概述1

1.1.1 数字图像获取1

1.1.2 图像显示与存储1

1.1.3 数字图像文件2

1.1.4 数字图像处理4

1.2 模式识别基本概念5

1.2.1 模式和模式识别的概念5

1.2.2 模式空间、特征空间和类别空间5

1.2.3 模式识别系统的组成5

1.2.4 数字图像模式识别7

1.3 实践拓展7

第2章 Visual C++数字图像处理基础9

2.1 Visual C++编程方法9

2.1.1 面向对象编程9

2.1.2 MFC类12

2.1.3 程序框架14

2.1.4 集成开发环境16

2.1.5 生成多文档应用程序19

2.2 Visual C++数字图像处理23

2.2.1 BMP图像文件23

2.2.2 位图文件读取25

2.2.3 图像增强33

2.2.4 图像形态学处理43

2.2.5 图像分割48

2.3 实践拓展51

第3章 图像特征53

3.1 统计特征53

3.2 幅值特征55

3.3 几何特征56

3.3.1 位置与方向56

3.3.2 周长57

3.3.3 面积57

3.3.4 长轴与短轴58

3.3.5 距离58

3.4 形状特征59

3.4.1 多边形描述59

3.4.2 曲线描述59

3.4.3 标记60

3.4.4 矩形度60

3.4.5 圆形度60

3.4.6 不变矩61

3.4.7 偏心率62

3.5 纹理特征62

3.5.1 纹理63

3.5.2 纹理分析63

3.6 实践拓展64

第4章 统计模式识别65

4.1 统计模式识别的研究内容65

4.2 特征的提取与选择66

4.2.1 特征评判标准——类别可分性判据66

4.2.2 特征选择及分支界定法67

4.2.3 特征提取及主分量分析68

4.3 模式分类71

4.3.1 最小错误率的贝叶斯决策71

4.3.2	感知器分类器	73
4.3.3	近邻分类器	74
4.4	模式聚类	75
4.4.1	模式相似性测度和聚类准则	76
4.4.2	层次聚类法	77
4.4.3	c-均值算法	77
4.5	实践拓展	79

第5章 模式识别决策方法及实现 ···80

5.1	人工神经网络	80
5.1.1	神经网络基本原理	80
5.1.2	误差反向传播算法	83
5.1.3	BP网络的设计	85
5.1.4	BP算法的C语言实现	86
5.2	隐马尔可夫模型	89
5.2.1	隐马尔可夫概念	89
5.2.2	隐马尔可夫模型基本算法	90
5.2.3	隐马尔可夫模型的C语言实现	94
5.3	决策树	104
5.3.1	决策树的基本概念	104
5.3.2	决策树的设计	106
5.3.3	决策树的C语言实现	107
5.4	模板匹配	112
5.4.1	模板匹配概念	113
5.4.2	Hausdorff距离	113
5.4.3	基于改进的Hausdorff距离的模板匹配算法	113
5.4.4	模板匹配的C语言实现	114
5.5	实践拓展	115

第6章 人脸检测与特征点定位 ···117

6.1	人脸检测方法	117
6.1.1	基于肤色的检测方法	117
6.1.2	其他人脸检测方法	119
6.2	人脸检测实例	120
6.2.1	系统设计	120
6.2.2	肤色相似度计算	122
6.2.3	人脸轮廓提取	125
6.2.4	人脸定位	127

6.2.5	脸内轮廓提取	131
6.2.6	眼睛定位	134
6.2.7	鼻子定位	142
6.2.8	嘴部定位	145
6.3	实践拓展	148

第7章 汽车牌照识别 ···151

7.1	系统概述	151
7.2	车牌定位	152
7.2.1	车牌颜色识别	153
7.2.2	车牌形状识别	159
7.2.3	车牌纹理识别	165
7.2.4	车牌倾斜校正	165
7.2.5	车牌定位及提取	166
7.3	车牌字符分割	172
7.3.1	车牌二值化	172
7.3.2	去除边框	178
7.3.3	字符分割	178
7.4	车牌字符识别	178
7.4.1	字符归一化	179
7.4.2	字符细化	179
7.4.3	除噪	180
7.4.4	字符模板匹配	180
7.5	实践拓展	184

第8章 脑部医学影像诊断 ···185

8.1	医学影像自动诊断	185
8.2	脑部医学影像的特征提取	187
8.2.1	灰度共生矩阵	187
8.2.2	脑CT图像纹理特征提取实例	188
8.3	脑部医学影像分类器设计	195
8.3.1	神经网络分类器的训练	195
8.3.2	脑CT图像分类器训练实例	196
8.3.3	分类器评估	201
8.4	实践拓展	201

第9章 印刷体汉字识别 ···203

9.1	印刷体汉字的特征提取	203
9.1.1	汉字特征的分类	203
9.1.2	常用的汉字特征	204

9.2 印刷体汉字的分类器设计	205
9.2.1 统计模式识别	205
9.2.2 结构模式识别	206
9.2.3 统计模式识别与结构模式识别的结合	207
9.2.4 人工神经网络	207
9.3 印刷体汉字识别实例	207
9.3.1 系统设计	207
9.3.2 图像预处理	210
9.3.3 文本区域处理	212
9.3.4 多特征提取	220
9.3.5 多分类器集成	221
9.4 实践拓展	237
第10章 手写体数字识别	239
10.1 系统概述	239
10.2 手写体数字图像的预处理	241
10.2.1 图像的二值化	241
10.2.2 图像反色	242
10.3 手写体数字的特征提取	244
10.4 手写体数字的识别	251
10.5 实践拓展	255
第11章 运动图像分析	257
11.1 运动图像分析概述	257
11.1.1 运动的分类	257

11.1.2 运动图像分析内容	258
11.2 运动目标检测与跟踪实例	258
11.2.1 系统设计	258
11.2.2 运动目标检测	264
11.2.3 运动目标跟踪	268
11.3 实践拓展	273

附录A 图像处理子函数代码——灰度变换	276
----------------------------	-----

附录B 图像处理子函数代码——几何变换	282
----------------------------	-----

附录C 图像处理子函数代码——空域增强	290
----------------------------	-----

附录D 图像处理子函数代码——频域增强	298
----------------------------	-----

附录E 图像处理子函数代码——形态学	310
---------------------------	-----

附录F 图像处理子函数代码——图像分割	317
----------------------------	-----

参考文献	331
-------------	-----

第1章 绪 论

可以说模式识别 (pattern recognition) 伴随人们生活的每一天。听到某个人说话, 就能分辨出这个人是谁; 当某种东西出现在视觉范围之内时, 就能认出这种东西是什么, 如此等等。随着计算机技术的发展, 人们希望利用计算机模仿人的识别能力, 从而开始了模式识别技术研究。

作为模式识别技术的一个重要分支, 数字图像模式识别技术已在各行各业中得到广泛应用, 并为提高有关技术的自动化、智能化水平发挥着重要作用。本章将着重阐述数字图像模式识别技术的一些基本概念。

1.1 数字图像处理概述

数字图像处理也称为计算机图像处理, 是指将图像信号转换成数字格式并利用计算机对其进行处理的过程。数字图像处理起始于20世纪50年代, 并于20世纪60年代初期形成一门独立学科。目前, 数字图像处理技术作为自动化、信息化的一个重要手段广泛应用于各行各业的工作环节中, 为提高其工作效率和工作精度作出了不可忽视的贡献。正因为如此, 图像处理理论和技术受到了越来越广泛的重视。

1.1.1 数字图像获取

图像是指自然景物存储在各种介质上的视觉信息。这类原始图像也称为模拟图像。计算机无法直接处理模拟图像, 而需要将其转换成数字图像并采集到计算机中再进行处理。这一过程称为数字图像获取。

数字图像获取一般包括图像摄取、转换及数字化等几个步骤, 主要由硬件实现。例如扫描仪可以将印有文字的书页转换为数字图像并采集到计算机中, 它的核心元件包括光学成像部件和光电转换部件。光学成像部件由灯管、反光镜、镜头以及电荷耦合器件组成。扫描仪工作时, 灯管发出强光照射书页; 镜头在反光镜辅助下对书页图像的反光进行聚焦, 并由电荷耦合器件进行光采集, 生成关于书页的图像。这种图像还需由光电转换部件将其转换成计算机可以识别的二进制数字格式, 即生成最终的数字图像。最后, 扫描仪通过通信接口将数字图像传送至计算机中完成整个书页图像的获取。

1.1.2 图像显示与存储

图像采集到计算机中, 是以图像文件的形式存储的。在需要使用时, 由应用程序对这些文件进行读取及信息处理。

1. 图像显示

图像在计算机屏幕上的显示,实际上是从存有图像数据的图像文件中取出数据,再按计算机显示电路的要求送到显示存储器和调色板中的过程。其中,显示存储器中的数据是关于在计算机屏幕不同位置上显示的颜色值,而颜色的定义则由调色板控制。同时,从图像文件中可取得图像的各种参数,这些参数用于控制图像的显示。如果文件中数据的存储形式与显示模式不一致,还需进行类型转换。

2. 图像存储

显示在屏幕上的图像,其调色板寄存器和显示存储器中的数据可用专门编制的程序读出。图像参数可以在显示过程中得到。屏幕图像的存储过程按显示过程的反方向进行。

1.1.3 数字图像文件

为了对数字图像文件进行正常存取,这里介绍数字图像文件的文件结构。尽管图像文件种类繁多,但其结构是相似的。

1. 图像基本类型

为了更清楚地解释数字图像文件结构,先简单介绍数字图像按颜色分类的基本类型。

(1) 二值图像

一幅二值图像的像素值仅由0、1两个值构成,“0”代表黑色,“1”代表白色。由于每一像素取值仅有0、1两种可能,所以计算机中二值图像像素的数据类型通常为1个二进制位。

(2) 16色图像

一幅16色图像像素值的取值范围为[0,15]。由4个二进制位表示。实际上,0~15并不代表具体颜色,而是调色板中16种颜色的索引值。由于这类图像只能显示16种颜色,因此显示效果不是很精细,目前只在图标等较少的场合中有应用。

(3) 256色图像

一幅256色图像像素值的取值范围为[0,255]。由8个二进制位表示。和16色图像一样,0~255也不代表具体颜色,而是调色板中256种颜色的索引值。和16色图像相比,由于可显示颜色更丰富了,显示效果得到明显改善。

256色图像更重要的一种应用格式是256灰度图像。256灰度图像实际上是256色图像的一个特例。图像像素值的取值范围仍然是[0,255],但每一个数值代表一种灰度,“0”表示纯黑色,“255”表示纯白色,中间的数字从小到大表示由黑到白的过渡色。就文件格式而言,256灰度图像仍然采用256色图像的格式,只不过其调色板中256种颜色已明确由小到大定义成不同灰度颜色。由于256灰度图像去除了颜色信息,因此,在很多图像分析场合,尤其是数字图像模式识别应用中较常用。

(4) 真彩色图像

和上述图像不同,真彩色图像的像素值直接表示具体颜色,因此没有调色板。一般真彩色图像的每个像素值由三个字节组成,分别表示红(R)、绿(G)、蓝(B)三原色的亮度值。组合在一起,真彩色图像共可以表示 $256 \times 256 \times 256 = 16\,777\,216$,或16M种颜色。由于真彩色图像所能表示的颜色数量已经超过人眼所能分辨的颜色数量,因此,实际应用中认为真彩色图像的显示效果和数字化前的原始图像是一样的,这也是真彩色图像名称的由来。

2. 图像文件结构

在计算机中,数据以文件的形式存储在外存储器中,图像数据也不例外。图像文件就是以数字形式存储的图像。为了便于读写,图像数据一般以一定的格式存放。目前已知有几十种图像文件格式,它们虽各不相同,但都具有相似的特征。图像文件一般由文件头、像素数据和调色板数据3部分组成。

(1) 文件头

文件头用于存放图像基本信息,包括各种特征参数、像素数据与调色板数据在文件中的存放位置,以及文字注解等。它分固定格式和灵活格式两大类。前者如BMP、PCX文件等,它们存储规定的参数,且数据在文件中的存储位置是固定的。后者如TIF文件,文件中除了规定的参数外,还可自定义特征参数,数据在文件中的存储位置也不固定,且同一文件中还可存放多幅图像。还有一些格式介于两者之间,如GIF和TGA文件格式。

(2) 像素数据

像素数据以位图的形式存放,每个像素数据对应图像相应位置上的颜色值,具体颜色值的定义由调色板数据提供。像素数据有压缩和不压缩之分。压缩数据可以节省存储容量,但在存取时需进行压缩和解压缩处理,处理速度稍慢。不压缩的像素数据在不同格式的图像文件中的存储结构基本相同。

(3) 调色板数据

调色板数据是指二值、16色和256色彩色图像的色调数据,真彩色图像的像素数据中已有颜色分量,故不再需要调色板数据。有些图像文件格式中调色板数据放在文件头中。有的文件中还有输入、输出设备及存储介质的光学特性曲线等,据使用场合的不同而不同。图像显示时需要调色板数据。

3. 图像文件常用参数

在图像的显示控制中最为重要的参数包括图像的宽度、高度、每像素所占位数、位平面数、图像类型、像素数据存储位置、调色板数据存储位置、压缩类型和图像扫描方向等。它们可分为图幅参数、类型参数、位置参数和其他参数等4类。

(1) 图幅参数

图像的宽度和高度是图像的基本参数。由图像宽度可以计算出每行字节数,即由图像宽度乘上每像素字节数直接求得,但有的图像格式要求每行字节数是4的倍数(如BMP文件)或偶数(如PCX文件),计算时需作调整。

(2) 类型参数

图像的每像素所占位数、位平面数决定了图像的存储结构。每像素所占位数与位平面数相乘可得每像素所占总位数,由它可以计算出图像中用到的色彩数,后者决定了图像所需调色板中色彩的数量。总位数1、4、8分别对应于色彩数2、16、256。总位数为24时图像是真彩色图像,它不需要调色板数据,故色彩数为0。得到每像素所占位数和位平面数后,就可以确定图像的类型。

有的图像类型还特别注明图像是黑白还是彩色。有了这个参数,黑白图像文件中可以省略调色板数据(如TIF、TGA文件)。

(3) 位置参数

像素数据和调色板数据的存储位置用于对图像数据的存数、取数指针进行定位。它们的

存放位置可以是固定的，也可以是不固定的，因不同文件而不同。

这里，需注意的是图像扫描方向或数据存储顺序。大多数图像文件采用的扫描方向为自上而下（如PCX、GIF、TIF与JPEG文件），少数为自下而上（如BMP、TGA文件）。因此，在对读取图像数据进行显示或处理时，需按照文件类型进行相应调整。

(4) 其他参数

压缩类型用于注明图像文件是否压缩，采用何种压缩等，供存数、取数程序确定是否需要压缩和解压缩。

综上所述，尽管计算机中存储的图像文件类型繁多，但不同类型的图像文件均可由图像中每像素所占位数和位平面数等图像类型参数来区分，如表1-1所示。图像文件按每像素所占总位数可分为4类，即黑白二值、16色彩色、256色彩色和24位真彩色。有些图像文件中用图像类型这个参数来区别黑白图像和彩色图像。表1-1中，pal表示该文件类型采用调色板，需要调色板数据；rgb表示彩色分量排列顺序为红、绿、蓝，bgr则相反，为蓝、绿、红；bw、wb则省略调色板，前者表示黑白图像中零值为黑色，高数值为白色，后则相反，零值为白色，高数值为黑色；*表示该类图像支持的数据类型。

此外，16色彩色和真彩色图像可分为第一位平面结构与多位平面结构两类。单一位平面结构下，图像数据以像素为单位集中存放，与显示存储器中的存放形式相同。而图像文件中的多位平面存储结构与显示存储器中的有所不同，它是按行为单位存放的。在同一行中图像数据以颜色分量为单位集中在一起，每个位平面相当于一行黑白图像，几个位平面的数据依次存放。

表1-1 图像文件类型参数

位数	1			4		8			24		
类型	黑白二值			16色彩色		黑白灰阶		256彩色	真彩色		
位平面	1	1	1	1	4	1	1	1	1	1	3
	bw	wb	pal	pal	pal	bw	wb	pal	rgb	bgr	rgb
BMP			*	*				*		*	
TGA						*		*		*	
GIF			*	*				*			
PCX			*		*			*			*
TIF	*	*		*		*	*	*	*		
JPEG									*		

1.1.4 数字图像处理

数字图像处理即针对在空间上离散的，在幅值上量化分层的数字图像，采用一些特定数理模式进行加工处理，以达到有利于人眼视觉或某种接收系统所需要的图像的过程。数字图像处理的方法主要包括：图像变换、图像增强、图像复原、图像压缩编码等。这部分处理也称作数字图像的预处理。其基本特点是输入和输出均为图像。在此基础上，可以进行图像分析，模式识别等内容研究。一般认为图像分析和模式识别是和图像预处理相对独立的另两个部分。其基本特点是输入的是图像，输出不为图像，而是对图像的分析 and 特征分类，或对图像的描述和解释。

1.2 模式识别基本概念

模式识别是20世纪60年代迅速发展起来的一门前沿学科,在许多领域得到了广泛应用。本节简要介绍模式识别的基本概念和系统组成,以及在图像识别中的应用方法。

1.2.1 模式和模式识别的概念

模式识别可以说是伴随人们生活的每一天。听到某个人说话,就能分辨出这个人是谁;当某种东西出现在视觉范围之内时,就能认出这一东西是什么,如此等等。在模式识别技术出现之前,人们从来没有过多地关注自己所具备的识别能力,而认为是理所当然的。

模式识别的目的就是利用计算机模仿人的识别能力,它往往包含有推理过程,需要有专家系统、知识工程等相关学科的支持。当前的模式识别理论主要是对人的低级识别能力的模拟,即是实现“观察对象是什么”的判断。

模式是指具有某种特定性质的观察对象。特定性质是指可以用来区分观察对象是否相同或相似而选择的特性。观察对象是指凡是人类感官直接或间接接受的外界信息。例如,一个数字、一句话、一张照片等都是观察对象。而把具有某些共同特性的模式的集合称为模式类。

模式识别是研究一些自动技术,依靠这些技术,计算机自动地(或者人进行少量干涉)把待识别的模式分到各自模式类中。即模式识别技术就是根据模式的特性,将其判定为某一模式类的技术。

1.2.2 模式空间、特征空间和类别空间

从技术途径来说,模式识别实质上完成的是从模式空间经过特征空间到类别空间的映射过程。

在物理上可以觉察到的世界里,适当地选择某些物体和事件,把它们称为样本。对它们分别进行观察所得到的观测数据的集合就构成了模式,所有的观察样本数据则构成模式空间。显然,模式空间的维数与所选择的样本和观测方法有关,也与特定的应用有关,一般是一个很大的有限值。在模式空间中,每个样本都是一个点,点的位置由该模式在各维上的数据来确定。有物理上可以觉察到的世界到模式空间所经历的过程称为模式采集。

模式空间的维数虽多,但有些并不能揭示样本的实质。对模式空间里的各坐标元素进行综合分析,获取最能揭示样本属性的观测量作为主要特征,这些主要特征就构成了特征空间。显然,特征的维数大大压缩了。由模式空间到特征空间所需要的综合分析,往往包含适当的变换和选择,称之为特征提取和选择。

利用某些知识和经验可以确定分类原则,称之为判别规则。根据适当的判别规则,将特征空间里的样本区分成不同的类型,从而将特征空间转换成了类别空间。类别空间中不同类别的分界面,常称为决策面。类别空间的维数与类别的数目相等,通常小于特征空间的维数。由特征空间到类别空间所需要的操作是分类判别。

1.2.3 模式识别系统的组成

一个典型的模式识别系统如图1-1所示,由信息获取、预处理、特征提取和选择、分类决策和分类器设计五部分组成。可以分为上下两部分:上半部分完成未知类别模式的分类;下

半部分完成分类器的设计训练过程。

模式识别系统各组成单元的功能如下。

(1) 信息获取

利用计算机可以运算的符号来表示所研究的对象，对应于外界物理空间向模式空间的转换。一般获取的信息类型有以下几种。

- 一维波形：心电图、脑电波、声波、震动波形等。
- 二维图像：文字、地图、照片等。
- 物理参量：体温、化验数据、温度、压力、电流、电压等。

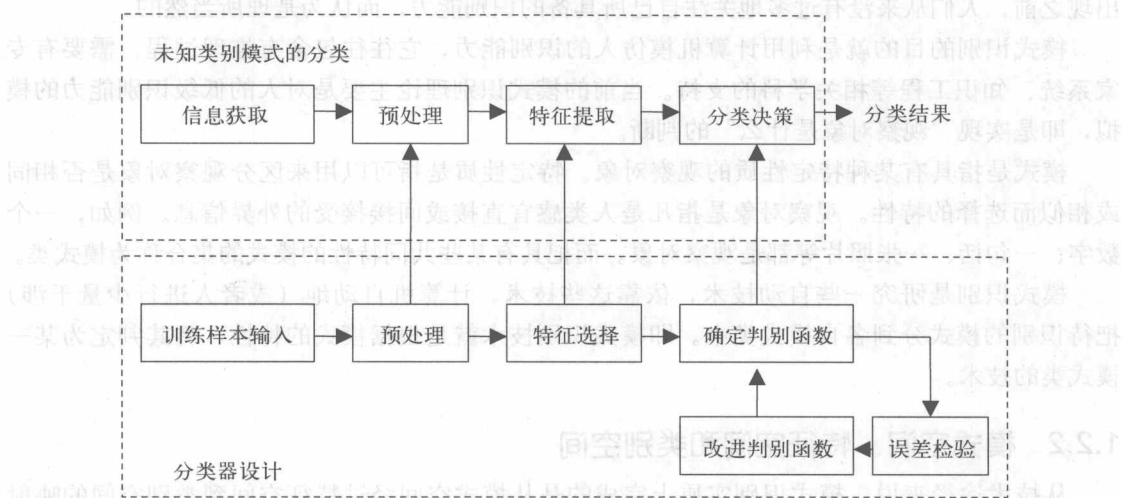


图1-1 模式识别系统的组成

(2) 预处理

对由于信息获取装置或其他因素所造成的信息退化现象进行复原、去噪，加强有用信息。

(3) 特征提取

由信息获取部分获得的原始信息，其数据量一般相当大。为了有效地实现分类识别，应对经过预处理的信息进行选择或变换，得到最能反映分类本质的特征，构成特征向量。其目的是将维数较高的模式空间转换为维数较低的特征空间。

(4) 分类决策

在特征空间中用模式识别方法（由分类器设计确定的分类判别规则）对待识模式进行分类判别，将其归为某一类别，输出分类结果。这一过程对应于特征空间向类别空间的转换。

(5) 分类器设计

为了把待识模式分配到各自的模式类中，必须设计出一套分类判别规则。基本做法是收集一定数量的样本作为训练集，在此基础上确定判别函数，改进判别函数和误差检验。

模式识别的关键是解决如何利用计算机进行模式识别，并对样本进行分类。执行模式识别的基于计算机系统（可以是台式机、笔记本电脑或基于单片机、DSP和ARM等有计算能力的系统）称为模式识别系统。

1.2.4 数字图像模式识别

在各种信息中,图像含有的信息最多,也是人类视觉的基础,因此基于视觉的图像识别具有特别重大的意义。图像可以是各种物体的黑白或彩色图画、手写字符、遥感图片、X射线透视胶片等等。图像识别的目的在于用计算机自动处理图像上的信息,以代替人去完成图像分类及辨识的任务。

图像识别作为模式识别的一类,其识别过程与模式识别的过程类似。图像识别的四个主要步骤如图1-2所示。

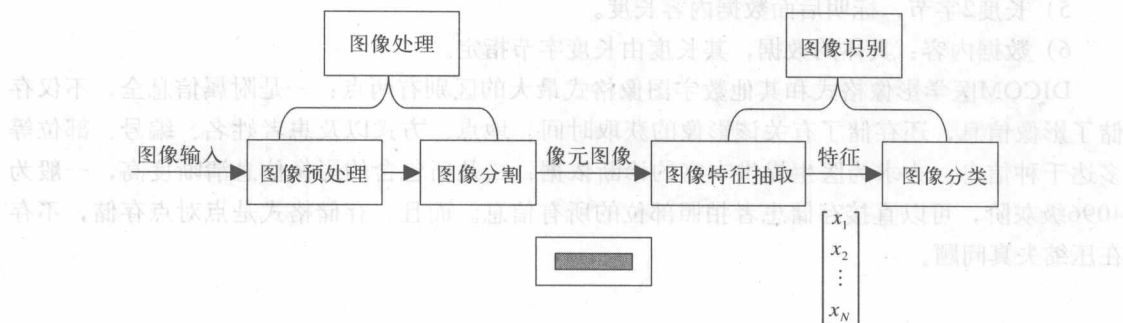


图1-2 图像识别的四个主要步骤

(1) 图像预处理

为了研究图像内容的识别,首先要对获得图像进行预处理,滤去干扰、噪声,当图像中的信息微弱无法辨识时,还需对图像进行增强处理,几何调整,颜色校正等,以便进一步进行人机分析。

(2) 图像分割

为了从图像中找到需要识别的物体,还需要对图像进行分割,也就是定位和分离出不同的待识别物体。这一过程输入的是整幅图像,输出的是像元图像。

(3) 图像特征抽取

在图像中需要识别的物体被分割出来的基础上,提取需要的特征,并对某些参数进行计算、测量,根据测量的结果进行分类。

(4) 图像分类

根据提取的特征值,利用模式识别方法进行分类,确定类别名称,以便对图像的重要信息得到一种理解和解释。这一过程的输入是特征信息,输出是类别名称。

1.3 实践拓展

随着现代计算机技术的飞速发展,在医学领域,各种最新的计算机技术的引进与应用已成为进一步促进医疗诊断水平提高必不可少的手段。以前,各种医学影像诊断设备如超声、CT、核磁共振等探测到的人体信息均采用普通数字图像格式,如BMP、JPG等进行存储。但由于这类数字图像存储格式存储的医学影像信息存在信息不全、清晰度低等问题,因而,由美国国家电子产品制造商协会和美国放射医学学院联合制定了DICOM医学数字影像存储格式。该格式已为全世界各大医疗器械厂家所接受,并最终成为世界共同遵守的协定,从而极

