



普通高等教育“十一五”国家级规划教材
高等院校通信与信息专业规划教材

数字视频 图像处理与通信

DIGITAL VIDEO PROCESSING
AND COMMUNICATION



刘富强 王新红 宋春林 等编著



机械工业出版社
CHINA MACHINE PRESS

普通高等教育“十一五”国家级规划教材
高等院校通信与信息专业规划教材

数字视频图像处理与通信

刘富强 王新红 宋春林 陈康力 李志鹏 王平 编著



机械工业出版社

本书共12章,第1~4章介绍了数字视频图像的基础知识,包括视频图像的表示、变换和预处理。第5、6章介绍了视频图像分割技术,包括图像分割的基本方法和视频分割的理论与方法。第7~9章介绍了视频图像分析,包括图像匹配、图像识别和运动目标跟踪。第10、11章介绍了视频图像传输,包括视频图像编码和多媒体通信。第12章介绍了视频图像在智能交通系统中的应用。

本书内容详尽,重点突出,注重理论和实践的结合,每章后有习题供读者思考或练习,可作为信息工程、计算机、通信、自动化等专业高年级本科生和研究生教材或教学参考书,也可作为工程技术人员的参考资料或培训教材。

图书在版编目(CIP)数据

数字视频图像处理与通信/刘富强等编著. —北京:机械工业出版社, 2009.12

普通高等教育“十一五”国家级规划教材. 高等院校通信与信息专业规划教材

ISBN 978-7-111-28834-3

I. 数… II. 刘… III. ①视频信号-数字图像处理-高等学校-教材②视频信号-数字通信-高等学校-教材 IV. TN941.3 TN914.3

中国版本图书馆CIP数据核字(2009)第231717号

机械工业出版社(北京市百万庄大街22号 邮政编码100037)

策划编辑:李馨馨 责任编辑:李馨馨 版式设计:张世琴

封面设计:鞠杨 责任校对:樊钟英 责任印制:杨曦

北京市朝阳区展望印刷厂印刷

2010年6月第1版第1次印刷

184mm×260mm·18.75印张·463千字

0001-3000册

标准书号:ISBN 978-7-111-28834-3

定价:35.00元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

社服务中心:(010)88361066

门户网:<http://www.cmpbook.com>

销售一部:(010)68326294

销售二部:(010)88379649

教材网:<http://www.cmpedu.com>

读者服务部:(010)68993821

封面无防伪标均为盗版

高等院校通信与信息专业教材

编委会名单

(按姓氏笔画排序)

编委会主任	乐光新	北京邮电大学
编委会副主任	张文军	上海交通大学
	张思东	北京交通大学
	杨海平	解放军理工大学
	徐澄圻	南京邮电大学
	吴镇扬	东南大学
	王金龙	解放军理工大学
	刘 陈	南京邮电大学
编委会委员	赵尔沅	北京邮电大学
	邹家禄	东南大学
	张邦宁	解放军理工大学
	张玲华	南京邮电大学
	徐惠民	北京邮电大学
	王成华	南京航空航天大学
	王建新	南京理工大学
	彭启琮	电子科技大学
	南利平	北京信息科技大学
	刘增基	西安电子科技大学
	刘富强	同济大学
	李少洪	北京航空航天大学
	冯正和	清华大学
秘 书 长	胡毓坚	机械工业出版社
副 秘 书 长	许晔峰	解放军理工大学

出版说明

为了培养 21 世纪国家和社会急需的通信与信息领域的高级科技人才，为了配合高等院校通信与信息专业的教学改革和教材建设，机械工业出版社会同全国在通信与信息领域具有雄厚师资和技术力量的高等院校，组成阵容强大的编委会，组织长期从事教学的骨干教师编写了这套面向普通高等院校的通信与信息专业规划教材，并且将陆续出版。

这套教材将力求做到：专业基础课教材概念清晰、理论准确、深度合理，并注意与专业课教学的衔接；专业课教材覆盖面广、深度适中，不仅体现相关领域的最新进展，而且注重理论联系实际。

这套教材的选题是开放式的。随着现代通信与信息技术日新月异的发展，我们将不断更新和补充选题，使这套教材及时反映通信与信息领域的新发展和新技术。我们也欢迎在教学第一线有丰富教学经验的教师及通信与信息领域的科技人员积极参与这项工作。

由于通信与信息技术发展迅速而且涉及领域非常宽，这套教材的选题和编审如有缺点和不足之处，诚恳希望各位老师和同学提出宝贵意见，以利于今后不断改进。

机械工业出版社
高等院校通信与信息专业规划教材编委会

前 言

随着数字视频技术的迅速发展,近年来各种视频应用层出不穷,特别是网络技术和无线通信技术的快速发展,不仅为视频技术的发展和应用提供了良好的平台,也使得视频应用有了更大的需求。为了适应目前高等院校本科生和研究生教育对于数字视频知识的需求,作者在长期从事有关数字视频技术教学、科研实践的基础上编写了本书。

本书共12章。其中第1~4章介绍了数字视频图像的基础知识,包括视频图像的表达、变换和预处理。第5、6章介绍了视频图像分割技术,包括图像分割的基本方法和视频分割的理论与方法。第7~9章介绍了视频图像分析,包括图像匹配、图像识别和运动目标跟踪。第10、11章介绍了视频图像传输,包括视频图像编码和多媒体通信。第12章介绍了视频图像在智能交通系统中的具体应用。

本书内容详尽,重点突出,注重理论和实践的结合,每章后有习题供读者思考或练习,智能可作为信息工程、计算机、通信、自动化等专业高年级本科生和研究生教材或教学参考书,也可作为工程技术人员的参考资料或培训教材。

本书由同济大学刘富强教授组织编写、统稿并定稿,王新红编写了第1~4章(不包括第3章的3.5节),陈康力编写了第5、6章,宋春林编写了第3章的3.5节和第7、8章,李志鹏编写了第9、12章,王平编写了第10、11章。博士研究生董楠、崔建竹、邵洁等以及硕士研究生张夏宜、张珊珊等,在本书的编写过程中也做了不少工作,在此表示感谢。此外,还要感谢同济大学教材、学术著作出版基金委员会的资助。

在写作过程中我们参考了很多文献,在此对这些文献资料的作者予以感谢。

由于作者水平有限,疏漏和错误之处在所难免,敬请有关专家和读者批评指正。

本书配有电子教案,读者可在机械工业出版社网站(www.cmpedu.com)下载。

作 者

2009年9月于同济大学

目 录

出版说明

前言

第1章 绪论 1

1.1 视频图像处理的概念 1

1.2 视频图像处理的应用 1

1.3 视频图像处理系统的组成 6

1.4 习题 8

参考文献 8

第2章 视频图像的表达 10

2.1 图像的视觉基础 10

2.1.1 图像的概念 10

2.1.2 分辨率 12

2.1.3 人眼的视觉特性 13

2.2 颜色模型 14

2.2.1 颜色基础知识 14

2.2.2 颜色模型的基本概念 15

2.3 图像数字化 19

2.3.1 图像的采样 19

2.3.2 图像的量化 22

2.3.3 图像的表达 23

2.4 图像格式 24

2.4.1 BMP 24

2.4.2 JPEG 25

2.4.3 TIFF 27

2.4.4 GIF 28

2.5 数字视频格式 29

2.5.1 MPEG 29

2.5.2 MOV 31

2.5.3 ASF 31

2.5.4 RA/RM 33

2.6 习题 34

参考文献 34

第3章 图像变换 35

3.1 傅里叶变换 35

3.1.1 傅里叶变换的基本概念 35

3.1.2 离散傅里叶变换 36

3.1.3 傅里叶变换的性质 37

3.1.4 快速傅里叶变换 40

3.2 离散余弦变换 43

3.2.1 离散余弦变换的定义 44

3.2.2 离散余弦变换的正交性 45

3.2.3 离散余弦变换的计算 45

3.3 沃尔什—哈达玛变换 46

3.3.1 沃尔什变换 46

3.3.2 哈达玛变换 47

3.4 K—L变换 48

3.5 小波变换 49

3.5.1 小波的基本概念 49

3.5.2 小波变换 50

3.5.3 多分辨率分析 52

3.5.4 Mallat 算法 52

3.5.5 图像的小波变换及算法 55

3.6 习题 58

参考文献 59

第4章 视频图像预处理 60

4.1 灰度变换 60

4.1.1 直接灰度变换 60

4.1.2 直方图变换 62

4.2 图像平滑 67

4.2.1 邻域平均法 68

4.2.2 中值滤波 71

4.2.3 多图像平均法 72

4.2.4 低通滤波法 73

4.3 图像锐化 74

4.3.1 图像锐化的基本思路 74

4.3.2 梯度算子 75

4.3.3 拉普拉斯算子 77

4.3.4 高通滤波法 77

4.4 图像校正	78	6.5 习题	136
4.4.1 图像校正的基本思路	79	参考文献	136
4.4.2 常用的图像几何变换	80	第7章 图像匹配	138
4.5 图像形态学处理	82	7.1 图像匹配的基本概念	138
4.5.1 形态学的基本概念	83	7.2 图像匹配算法分类	138
4.5.2 形态学的基本运算	83	7.3 模板匹配算法	139
4.6 习题	84	7.3.1 ABS 算法	139
参考文献	85	7.3.2 归一化互相关匹配算法	140
第5章 图像分割	86	7.3.3 图像矩匹配算法	141
5.1 阈值分割	87	7.3.4 基于图像特征点的匹配算法	143
5.1.1 阈值分割的原理与方法	87	7.4 改进算法	149
5.1.2 基于像素的阈值选取	88	7.4.1 序贯相似性检测算法	149
5.1.3 基于区域的阈值选取	91	7.4.2 基于排序的序贯相关算法	150
5.2 基于区域的图像分割	95	7.4.3 FFT 的相关算法	152
5.2.1 区域生长	95	7.4.4 分层搜索的序贯判决算法	153
5.2.2 区域分裂—合并技术	98	7.5 其他算法介绍	155
5.3 基于边缘提取的图像分割	101	7.5.1 基于对数极坐标变换的图像匹配 算法	155
5.3.1 边缘检测	101	7.5.2 不同分辨率图像的角点匹配 方法	157
5.3.2 用于边缘检测的算子	102	7.6 待研究的问题	159
5.3.3 边缘连接	108	7.7 环境匹配技术发展简介	160
5.3.4 哈夫变换	108	7.7.1 环境匹配算法研究的现状 分析	160
5.4 图像分割的量化评价	112	7.7.2 环境匹配算法存在的主要 问题	163
5.4.1 评价方法与分类	112	7.7.3 环境匹配技术的发展趋势	164
5.4.2 常用测度	113	7.8 习题	165
5.5 习题	114	参考文献	165
参考文献	114	第8章 图像识别	166
第6章 视频分割的理论与方法	116	8.1 图像识别的基本概念	166
6.1 视频分割的相关概念	116	8.2 基于匹配的识别技术	167
6.1.1 视频分割的定义	116	8.2.1 图像变换	167
6.1.2 视频分割的研究和应用 现状	117	8.2.2 相似性测度	168
6.2 视频分割的基本理论	121	8.2.3 插值算法	169
6.2.1 视频分割的框架	121	8.2.4 最小二乘法	169
6.2.2 视频对象分割方法的分类	121	8.3 统计识别方法	170
6.3 视频分割技术	122	8.3.1 线性决策函数	170
6.3.1 空间域分割技术	122	8.3.2 最小距离分类器	173
6.3.2 时域分割技术	129		
6.3.3 时空联合分割技术	134		
6.4 视频分割算法的性能评价标准	135		

8.4 神经网络识别法	174	10.3 视频编码的整体框架和关键技术	231
8.4.1 神经网络的概念	175	10.3.1 预测编码	232
8.4.2 BP神经网络	175	10.3.2 帧内预测	234
8.4.3 自组织网络	184	10.3.3 变换编码	237
8.5 支持向量机识别方法	192	10.3.4 变长编码	239
8.5.1 支持向量机	193	10.4 习题	240
8.5.2 支持向量机的学习算法	196	参考文献	240
8.6 习题	197	第11章 多媒体通信概述	241
参考文献	197	11.1 多媒体通信概述	241
第9章 运动目标跟踪	199	11.2 流媒体技术	242
9.1 运动目标跟踪技术简介	199	11.2.1 流式传输的基础	242
9.2 卡尔曼滤波跟踪	200	11.2.2 流媒体的实现原理	243
9.2.1 离散卡尔曼滤波器	200	11.2.3 流媒体的播放方式	243
9.2.2 扩展卡尔曼滤波器	201	11.2.4 流媒体的应用	245
9.2.3 卡尔曼滤波器在目标检测与跟踪中的应用	202	11.3 多媒体通信协议	246
9.3 光流分析法	204	11.3.1 会话初始化协议	246
9.3.1 运动场与光流场	204	11.3.2 实时传输协议	252
9.3.2 Horn-Schunck 经典光流计算方法	205	11.3.3 实时流协议	256
9.3.3 光流计算方法实验结果	208	11.3.4 QoS 保障	258
9.4 均值漂移跟踪	209	11.4 多媒体通信系统	263
9.4.1 均值漂移算法概述	209	11.4.1 基于 H.320 的多媒体通信系统	263
9.4.2 基于色彩直方图的均值漂移	211	11.4.2 基于 H.323 的多媒体通信系统	265
9.4.3 均值漂移目标跟踪效果	212	11.4.3 基于 SIP 的多媒体通信系统	269
9.5 蒙特卡洛随机采样	213	11.4.4 基于 P2P 的多媒体通信系统	272
9.5.1 贝叶斯滤波原理	213	11.5 习题	277
9.5.2 蒙特卡洛方法	214	参考文献	277
9.5.3 粒子滤波算法综述	215	第12章 视频图像应用	278
9.5.4 粒子滤波目标跟踪效果	215	12.1 基于视频图像的智能交通监控系统	278
9.6 习题	217	12.1.1 系统硬件框架设计	278
参考文献	217	12.1.2 系统软件总体结构	279
第10章 视频图像编码	218	12.1.3 系统主要模块	280
10.1 视频图像编码概述	218	12.1.4 系统软件界面	281
10.2 图像/视频编码标准	220	12.2 基于视频图像的客流量检测	
10.2.1 MPEG 系列	220		
10.2.2 H.26x 系列	225		
10.2.3 AVS	230		

系统.....	282	12.3 基于视频图像的车载安全预警	
12.2.1 客流密度估计	282	系统.....	286
12.2.2 行人检测与计数	284	12.3.1 交通标志识别系统.....	287
12.2.3 客流速度检测	285	12.3.2 驾驶员注意力判别系统	289

第 1 章 绪 论

据统计,人类从外界获取的信息中,75%来自于视觉,这些信息实际上就是图像。在图像的基础上再加上时间因素,就形成了视频,因此视频有时又称为时基媒体。从最早出现的模拟视频发展到数字视频,是视频技术的一大飞跃,由于数字视频具有易存储、易编辑等特性,因此获得了越来越广泛的应用。

1.1 视频图像处理的概念

视觉是人类最重要的感觉,也是人类获取信息的主要来源。视频信息与其他的信息形式相比,具有直观、具体、生动等特点,并且视频所包含的信息量很大,“百闻不如一见”、“一图值千言”等谚语都说明了这一特点。因此,各种视频技术的研究和应用一直吸引着国内外广大科技人员的关注。随着多媒体技术及 Internet 的迅速发展,各种信息采集和生产手段在不断投入使用,例如,数码照相机、摄像机、扫描仪、打印机等的使用,使视频信息的来源不断扩大。同时,人们对视频的应用也越来越广泛,视频点播、可视电话、视频会议等应用形式不断涌现。视频信息正在对人类的生活方式和社会发展起着越来越重要的作用。

视频是一组图像在时间轴上的有序排列,是二维图像在一维时间轴上构成的图像序列。又称为动态图像、活动图像或者运动图像,它不仅包含了静止图像所包含的内容,还包含了场景中目标运动的信息和客观世界随时间变化的信息。电影、电视等都属于视频的范畴。早期的视频主要指模拟视频信号,随着电子技术的发展以及全球数字化进程的推进,视频的采集设备和采集方式有了很大的进展,直接采集数字视频的设备得到了广泛的开发和应用。

所谓的视频图像处理,就是指用数字计算机及其他有关的数字技术,对图像施加某种运算和处理,从而达到某种预期的目的。

1.2 视频图像处理的应用

视频图像处理技术广泛应用于生物医学、航天遥感、军事和公安部门、工业生产、智能交通等众多的领域中。

1. 生物学、医学方面的应用

生物学、医学是图像处理技术的重要应用领域之一,也是较早进行图像处理研究的学科之一。其中一类应用是对生成医学显微图像的处理和分析,如红、白细胞和细菌、虫卵的分类计数,癌细胞的识别以及染色体分析等。此外,利用图像处理技术还可以得到常规方法无法得到的图像,如由 CT(计算机断层扫描成像)得到人体横截面的图像等。

在医疗诊断方面,血细胞自动分类技术是指应用图像处理和模式识别技术对医院常规验血实行自动化处理的技术。从玻片图像输入开始,对红、白血细胞自动分类及计数。对细胞图片的分析,是在细胞学的基础上,把细胞图像分割,求出细胞核和细胞浆区域,再计算它

们的几何参量和光密度参量，求出表征细胞特征的少数最重要的参数，进而区分正常细胞与癌细胞。还可以进一步把医生的经验研制成专家系统，指导计算机进行细胞分类。另外，对B型超声波图像进行图像增强，图像分析测量有关参数，其结果可以大大提高图像的清晰度和图像判断的准确度，也提高了整个设备的分辨率。肺部病人的透视照片，经过图像增强与分析后，能更清楚地显示病灶情况。肿瘤照片经处理后，可清楚地显示出肿瘤的位置与大小，并把照片中的血管等图像弄清楚。

在生物学研究方面，图像处理技术可用于染色体的自动分类。将医生制作的染色体玻片通过带摄像头的显微镜输入图像处理设备，经过单个染色体分割，提取轮廓，测量参数，实现自动分类，然后，再把每个染色体重新按类排序，得到人类染色体的自动排序结果，由此可区分出正常与异常染色体。随着医学技术的发展，可以制作分带的染色体，应用图像处理技术也可对它们进行自动分类。此外，多普勒血流计及血流图的检测设备已成功地应用于血流的检测与血流图的显示。

计算机断层扫描成像（CT）是利用投影数据重建来生成人体横截面图像的技术。现在流行的CT技术有X射线CT与超声CT两大类。其基本原理为：用X射线或超声波束横向通过人体某一截面，射线穿过人体时，由于各器官对它的吸收率不同，因而接收到的射线是经过沿途吸收衰减的结果。设这些吸收源是彼此无关的，则输出的结果将是每个吸收源衰减结果的线性叠加。由此可列出一系列线性方程，求解这个方程组可得到各吸收源的衰减因子，由于各器官的衰减因子不同，因此，可以得到人体某一截面上的器官分布图，通称CT图像。CT技术解决了通常投影式X线摄影所无法解决的问题。还有一种核磁共振CT，使人体免受各种硬射线的伤害，并且图像更为清晰。在处理技术进一步改进的情况下，由CT生成立体图像甚至活动图像也已成为可能，这为医疗诊断和医学研究带来了极大的便利。这是医学史上具有划时代意义的重大发明，其发明者因此获得了1979年诺贝尔医学奖。

2. 航天、遥感方面的应用

遥感是利用装载在飞机或人造卫星上的传感器，搜集由地球表面的物体反射或放射出的电磁波，并利用这些数据得到有关对象物和现象的信息的技术。遥感技术的特点是检测面积大，光谱谱段多，获取信息丰富，对物质有很强的识别能力，因此被成功地应用于国民经济的各个部门。它主要应用在空间探测、资源普查、地图测绘、军事侦察等方面。

在农业方面，可利用遥感技术进行农作物和森林资源分布情况的分析，制订土地利用规划及农作物收成估计等。

在水文、水资源工作方面，对于水库和较大河流中下游的湖泊，一般都有水面面积和库容的关系曲线资料，只要知道了水面面积就可以求得水体水量，从而通过卫星图片资料，随时掌握水库和湖泊的水量情况，尤其是洪水过后和干旱期的情况。水质污染问题是水资源工作中十分引人关心的问题。利用不同水质的辐射光谱特性的不同可以监测水的污染。例如，缺氧的水色调常呈黑色或暗色，水温增高在照片上表现为浅色调异常，硫酸排放对水的污染在全色黑白片加浅黄滤光片的照片上呈浅色调。在1972年10月10日的多光谱卫星图像上，十分明显地发现一个从纽约州新建的国际造纸公司排入恰普林湖的污染流，这污染流流到万佛蒙特州边，为此，万佛蒙特州第一次以卫星照片作为一个证据向该造纸厂提出法律诉讼。在我国的津勃环境遥感试验中对海河污染状况的观测分析表明，海河自天津解放桥以上河水水质清洁，在彩色红外图片上，色调浅蓝，结构均匀。

可利用遥感图像更新速度快的特点来观察自然现象的发展动态。例如，在天气预报方面，从气象雷达或气象卫星发回的图像中分析台风、暴雨以及天气变化等方面的信息，由此做出正确的预报；在灾情监察方面，监视洪水、森林火灾、病虫害等受灾区域的变化动态，以及环境污染的监测，据此做出预报并制订抗灾计划。

有些物质对某些波段的光线或微波的吸收率很低，可用来发现掩盖在这些物体下的东西。例如，在军事方面，可用于发现潜伏在海底的潜艇；在地质勘探方面，可用于寻找被植被和泥土覆盖的地质构造；在考古方面，可用于寻找史前遗址等。

可利用遥感技术进行资源调查，例如，在工业方面进行矿产、石油等分布情况的分析，用于指导找矿、石油勘探等。

遥感技术其他方面的应用还有空间探测、星际探险照片处理、水文海洋调查、环境污染监测、军事侦察、气象、地震后的损失统计等。

3. 军事和公安部门的应用

图像处理在军事上的应用主要是侦察照片的判读，雷达图像的处理，军用地图的存储与自动检索，小型图像地面接收站的装备和飞机、坦克、军舰等的模拟训练等。成像雷达的成像部分使用信号处理方法把扫描得来的数据重新构成图像，并显示出来，然后进行存储、分类、归档，以便自动检索，这可免去参谋人员大量的标图作业。装备了小型图像地面接收站的移动式战术侦察系统，可把传感器传来的图像记录下来，并进行目标分析和归档转发等，可将前线的情况及时传送到指挥部。训练模拟装置可提供接近实战条件的训练环境。

在公安方面的应用是图片的判读分析，如指纹鉴别、不完整图片的复原、跟踪、监视等。其中，指纹鉴别系统将指纹经过预测处理后提取特征，再用聚类分析方法进行分类，然后将指纹数据无失真压缩，存储在光盘上，就可以在几百万人的大城市中，实现居民指纹的存储和检索。

图像处理在公安方面的应用中，个人鉴别是广泛研究的一大类问题，包括指纹、掌纹、面部照片等的识别问题。这一领域的研究对于安全与侦破工作有重大意义。

其中指纹识别已实用化，实际的开发系统已经实现了在 1.3ms 中求出一对指纹的相似度。有人提出将来开门不用钥匙，而用指纹，这在保护财产、保守机密方面有着重大意义。据考古学家证实：公元前 7000 年到 6000 年以前，指纹作为身份鉴别的工具已经在古叙利亚和我国开始应用。在那个时代，一些粘土陶器上留有陶艺匠人的指纹，中国的一些文件上印有起草者的大拇指指纹，在耶利哥古城的房屋留有砖匠的指纹等。指纹的纹路在图案、断点和交叉点上是不相同的，在信息处理中将它们称做“特征”，医学上已经证明这些特征对于每个手指都是不同的，而且这些特征具有唯一性和永久性。指纹识别系统目前已经广泛应用于公安、银行、证券、各种考试的考生身份鉴别，及机关、企事业单位的考勤等领域。

人脸识别技术主要针对面部不易产生变化的部分进行图像处理，其中包括眼眶轮廓、颧骨的周围区域以及嘴的边缘区域等。计算机可在 0.05~0.1s 的时间内“捕获”人脸，并在 1s 内完成识别。但是目前的人脸识别方法主要集中在二维图像方面，由于受到光照、姿势、表情变化的影响，识别的准确度受到很大限制，迄今为止，建立一个鲁棒的人脸识别系统仍然是一个很难解决的问题。1993 年，美国国防部高级研究项目署（Advanced Research Projects Agency）和美国陆军研究实验室（Army Research Laboratory）成立了 FERET（Face

REcognition Technology) 项目组, 建立了 FERET 人脸数据库, 用于评价人脸识别算法的性能。到 1997 年底为止, 参加测试的算法中, 对 FB 图像 (图像与目标集中的图像是在同日、同光照条件下拍摄) 的识别率很高, 库藏 1196 人中首选率 (输出的候选图像集合中, 排在第一幅的图像与查询图像为同一人的概率) 达 96%; 对 FC 图像 (图像与目标集中的图像是在同日、不同光照的条件下拍摄) 的识别率也比较高, 库藏 1196 人中首选率达 81%。测试发现的主要问题如下: 识别算法对光照变化敏感; 查询图像和目标图像相隔一年半后, 识别算法的性能下降; 识别算法对人脸位置变化敏感, 人脸偏转角度大于 15° 后, 性能下降; 识别算法对相机敏感。美国陆军实验室在 13 周时间内, 用 270 人的图像测试一个人脸识别系统, 发现识别率只有 51%。这套系统在机场中进行测试时, 存储了 250 人的图像, 其中的 15 人在 1 个月内通过摄影机 958 次, 只有 455 次被正确辨认, 识别率只有 47%。在美国一个机场开展的一项为期 8 周的公开测试中, 使用一家公司的人脸识别系统, 在 4 周内出错率为 53%。在另一个机场开展的一项为期 90 天的测试中, 人脸识别系统发出的错误警报也太多。

针对人脸识别的难点, 一个较好的办法是利用三维信息进行人脸的识别。三维信息能够更精确地描述人的脸部特征, 提取的某些特征具有刚体变换不变性, 并且不易受化妆和光照的影响。但是由于三维数据获取方面存在困难, 现在利用三维信息进行识别的报道并不多见, 然而三维信息加入到现有的人脸识别算法中, 识别效果将会大大提高。

4. 工业生产方面的应用

由于数字图像的信息量大, 可作远距离、非接触测量, 在不破坏物体的情况下, 也可检测到物体的内部, 另外, 它还具备处理方法灵活, 便于自动化处理的特点。图像处理还广泛地应用于工业生产中, 如弹性力学照片的应力分析, 流体力学照片的阻力和升力分析, 金相照片的结构和成分分析, 流水线上机械零件的自动检测、分类, 信件、包裹的自动分拣, 高炉火焰的分析, 轧钢厂中轧制中的钢坯位置的确定和尺寸的测量, 工件内部结构的分析或裂缝的检测 (工业 CT) 等。

在现代工业自动化生产中, 涉及各种各样的检查、测量和零件识别应用, 例如汽车零配件尺寸检查和自动装配的完整性检查, 电子装配线的元器件自动定位, 饮料瓶盖的印刷质量检查, 产品包装上的条码和字符识别等。有些时候, 如微小尺寸的精确快速测量, 形状匹配, 颜色辨识等, 用人眼根本无法连续稳定地操作。这时, 人们开始考虑把计算机的快速性、可靠性、结果的可重复性, 与人类视觉的高度智能化和抽象能力相结合, 由此产生了机器视觉的概念。其最大的应用行业为: 汽车、制药、电子与电气、制造、包装/食品/饮料、医学。如对汽车仪表盘加工精度的检查, 高速贴片机上对电子元器件的快速定位, 对管脚数目的检查, 对 IC 表面印字符的辨识, 胶囊生产中对胶囊壁厚和外观缺陷的检查, 轴承生产中对滚珠数量和破损情况的检查, 对食品包装上的生产日期的辨识, 对标签贴放位置的检查。印制电路板的检测指标主要有: 缺件、错位、极性、颠倒、漏焊、短路、焊量、贴片脱落等。一般采用多摄像头技术并辅以灯光。

图像处理的重要分支“计算机视觉”为机器人提供了“视觉”。采用摄像机输入二维图像的机器人, 可以确定物体的位置、方向、属性以及其他状态等, 它不但可以完成普通的材料搬运、产品组装、部件组装、生产过程自动监控, 还可以在人不宜进入

的环境里进行喷漆、焊接、自动检测等。现在已发展到具备视觉、听觉和触觉反馈的智能机器人。

5. 智能交通方面的应用

交通事故的预防一直是世界各国学者们研究的热点，每年全世界都会发生数量惊人的交通事故，欧洲每年因交通事故造成的经济损失达 500 亿元之多，我国大城市的交通事故发生率更是高出西方发达国家数倍，其财产与生命损失无法估量。据统计，2006 年我国交通事故的死亡人数为 8.9 万人，约占全球总数的一半；2006 年，上海市交通事故的万车死亡率在 10 人以上，伦敦、巴黎和东京的万车死亡率则在 2 人以下。

另外，汽车的大量增加造成了多数城市的严重交通堵塞问题。据有关部门统计，全国因城市交通堵塞造成的损失每年高达上千亿元，并且交通问题未因环行路和立交桥的兴建而缓解。从大城市到小城市，交通堵塞成了普遍现象，这使得城市的效率大为降低，城市的功能没得到充分的发挥，同时大量的汽车尾气的排放严重污染了空气。显然，我国城市在发展过程中所面临的交通拥挤、环境污染等问题，已经成为影响城市交通环境的主要因素。于是，智能交通系统（Intelligent Transportation System, ITS）这一新的研究领域便应运而生，并迅速发展起来。

视频图像处理在道路交通状况监测中的应用主要包括：

- 1) 车速的识别，从移动时间和距离来判断车的速度。
- 2) 车型的识别，从车长来区别大车、小车。
- 3) 车牌的识别，通过视频图像处理的手段识别车牌号码。
- 4) 交通流量的监视，从车辆存在的位置来获得交通流量。

目前已投入运行的高速公路不停车自动收费系统中的车辆和车牌的自动识别都是图像处理技术成功应用的例子。

另外，一个在智能交通系统中应用广泛的方向是车载导航。通过在车辆前方装载的摄像头捕获交通环境视频信息，利用视频和图像处理技术对于捕获的视频信息进行分析和理解，结合机器视觉和模式识别技术实现为车辆驾驶中的道路环境识别、车辆和障碍物检测、危险环境预判、车辆和环境参数及和其他交通参数获取。以上信息最后转化成驾驶员可以直观理解的视觉或者听觉信号，给出驾驶提示和危险预警，为车辆和驾驶者提供详尽和直观的驾驶帮助和提示，起到减少驾驶员负担的作用，同时让车辆更加智能、更加安全。

6. 在其他方面的应用

在电影、电视和电子游戏等方面，现在也越来越多地采用数字图像技术，如电影、电视特殊场景和特技的制作，早期黑白电影片的自动着色，电视画面的数字编辑，电影动画片的制作等。《玩具总动员》就是第一部全部由计算机制作的动画片。电子游戏采用数字图像技术后，可使场景更加生动、逼真，使人产生身临其境的感觉，极大地提高了游戏的刺激性和吸引力。

动态序列图像的分析应用于运动员的动作分析，可提高运动员的训练水平，已有许多应用此项技术提高运动成绩和帮助运动员获得世界冠军的实例。此项技术还可帮助裁判员进行评分。

在艺术方面，目前较为成熟的应用有服装纺织工艺品的花纹设计、制作，文物资料照片

的存储、复制和修复等。广告和图书封面的设计现在已广泛使用图像编辑软件，不仅制作速度快，而且质量也很好。现在还在逐渐形成一门新的艺术——计算机美术。

采用了计算机技术后，出版印刷工业发生了彻底的变革，出版界采用计算机进行排版、分色；印刷界废弃了沿用百年的铅字，采用了激光照排技术，这些都极大地提高了印刷工业的自动化程度，缩短了出版周期，减轻了劳动强度，提高了印刷品的质量。这中间所用的技术有些也与数字图像有关。值得一提的是，我国这一技术处于国际领先地位，例如，王选教授长期致力于文字、图形和图像的计算机处理研究，应用自己的发明成果开发了汉字激光排版系统并形成产业，取代了沿用上百年的铅字印刷，推动了中国报业和出版业的跨越式发展，创造了巨大的经济和社会效益。世界上华文报纸大多采用北大方正的排版系统，除了传统的黑白排版外，它也可出彩色版。与卫星通信技术相结合，还实现了报纸版面的异地传输，这极大地缩短了《人民日报》等全国性大报的发行周期。

现在，大型数据库正在逐步取代图书馆的作用，从数据库中查阅资料既迅速又方便。现在在工程图样的设计已广泛使用 CAD 技术，它极大地提高了工程设计的效率和速度。但是要将现有的书面资料、工程图样输入计算机却需要花费巨大的人力物力，在这些方面采用图像处理技术可大大提高功效。工程图样矢量化软件可将扫描输入计算机的图样图像转换成 CAD 软件可以接受的图形数据，然后直接在 CAD 软件中修改、编辑；文字识别软件可将扫描输入计算机的文本图像转换成字处理软件可以接受的文本数据，然后直接在字处理软件中修改、编辑。在印刷体汉字的自动识别方面，我国已有产品问世。联机手写体识别已有产品问世，脱机手写体识别也在研究之中。

1.3 视频图像处理系统的组成

一个基本图像（处理和分析）系统的构成可用图 1-1 来表示。图中各模块都有特定的功能，分别是采集、显示、存储、通信、处理和分析。为完成各自的功能，每个模块都需要一些特定的设备。

1. 图像采集模块

为采集数字图像，需要两种装置（器件）：一种是对某个电磁能量谱波段（如 X 射线、紫外线、可见光、红外线等）敏感的物理器件，它能产生与所接收到的电磁能量成正比的（模拟）电信号；另一种称为数字化器，它能将上述（模拟）电信号转化为数字（离散）的形式。所有采集数字图像的设备都需要这两种装置。

以常见的 X 光透视成像仪为例，由 X 光源发出的射线穿越物体到达另一端对 X 光敏感的媒体。这个媒体能获得物体对 X 光不同吸收率的图像。它可以是胶片、一个带有将 X 光转化为光子的电视摄影机，或其他能将 X 光转换为数字图像的离散检测器。

电荷耦合器件（CCD）的成像机理是基于感光基元能产生与所接收的输入光强成正比

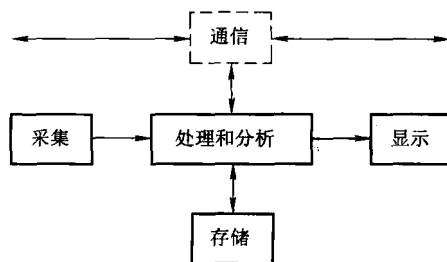


图 1-1 图像系统的构成示意图

的输出电压。电视摄像机一般由 CCD 阵组成，要得到数字图像需把摄像机的视频输出送到一个数字化器（称为图像捕获器）中。这个工作常通过在计算机中插入专门的硬件捕获卡实现。

图像卡仍是目前常用的图像数字化设备，目前低端的图像采集卡一般不具有图像帧存体，而是直接将图像采集到计算机的内存中以供处理，如加拿大 Matrox 公司的 Metro-II 采集卡，高端的图像卡是集采集和处理于一身的昂贵的非标准配件，如 Matrox 公司的 Genesis 图像卡，具有帧存体和数字信号处理器（DSP）及邻域处理加速器（NOA），用于开发高速或实时处理应用。此外，还有一类普及型的多媒体视频采集卡，如宝狮 Boser602，主要用于视频会议、视频邮件等。最后，还应提到的是一类多媒体应用中使用的压缩卡，如 AV8 MPEG 压缩卡，可以将视频压缩成 MPEG-I 格式，主要用于 VCD 制作和视觉保安系统中，当然也具有图像采集功能。后两种都支持微软公司的 VFW（Video For Windows）标准。此外，高档的压缩卡，如 RT2000 压缩卡可以将视频压缩成 MPEG-II 格式等，影视制作行业还有各种高性能的图像及视频编辑设备等。

扫描仪的原理是利用光电转换原理，通过扫描仪光码的移动或原稿的移动，把黑白或彩色的原稿信息数字化后输入到计算机中，所形成的是点阵图。与数码相机类似，在图像扫描仪中，也使用 CCD 作图像传感器。不同的是，数码相机使用的是二维平面传感器，成像时将光图像转换成电信号，而图像扫描仪的 CCD 是一种线性 CCD，即一维图像传感器。

扫描仪对图像画面进行扫描时，线性 CCD 将扫描图像分割成线状，每条线的宽度大约为 $10\mu\text{m}$ 。光源将光线照射到待扫描的图像原稿上，产生反射光（反射稿所产生的）或透射光（透射稿所产生的），然后经反光镜组反射到线性 CCD 中。CCD 图像传感器根据反射光线强弱的不同转换成不同大小的电流，经 A/D 转换处理，将电信号转换成数字信号，即产生一行图像数据。同时，机械传动机构在控制电路的控制下，使步进电动机旋转带动驱动带，从而驱动光学系统和 CCD 扫描装置在传动导轨上与待扫原稿作相对平行移动，将待扫图像原稿逐条线的扫入，最终完成全部原稿图像的扫描。

数码照相机不用胶片，而使用 CCD 阵列，把来自 CCD 阵列的电压信号送到 A/D 转换器后，转换成图像的像素值。

2. 图像显示模块

对图像处理来说，处理的结果主要用于显示给人看。对图像分析来说，分析的结果也可以借助计算机图形学技术转换为图像形式直观地展示。所以图像显示对图像处理和分析系统来说是非常重要的。

常用的显示设备是电视显示器。除了电视显示器，可以随机存取的阴极射线管（CRT）和各种打印设备也可用于图像输出和显示。

3. 图像存储模块

图像存储可采用磁带、磁盘、光盘和磁光盘等。

4. 图像通信模块

图像通信可借助综合业务网、计算机局域网（LAN），甚至普通电话网（PSTN）等。

5. 图像处理和分析模块

这是图像处理系统的主要和关键模块，主要包括：图像变换、图像编码压缩、图像增