



普通高等教育
电气工程与自动化类
“十一五”规划教材

DIGITAL IMAGE PROCESSING TECHNOLOGY AND APPLICATION

数字图像技术与应用

朱 虹 主编



数字图像技术与应用

朱虹 主编



机械工业出版社

本书采用了崭新的思路来构建编写体系,将静态的单帧图像的基本处理方法作为基础知识部分;然后对动态图像,即数字视频方面的知识进行介绍;最后介绍了三维重建技术。这样,就构成了从二维平面信息(静态图像)到时空三维信息(数字视频),以及空间三维信息(三维成像技术)的理论体系。在介绍了相关理论之后,配备了应用案例,给读者提供对应的解决方法,以达到启发、拓展思维的目的。

本书适合工科院校计算机应用、电子信息工程等专业本科生使用,也可供模式识别与智能系统、信号与信息处理、计算机应用等学科方向的硕士研究生参考。

本书配有免费电子课件,欢迎选用本书作教材的老师发邮件到 Jinacmp@163.com 索取,或登录 www.cmpedu.com 注册下载。

图书在版编目(CIP)数据

数字图像技术与应用/朱虹主编. —北京:机械工业出版社, 2011.3

普通高等教育电气工程与自动化类“十一五”规划教材
ISBN 978-7-111-33291-6

I. ①数… II. ①朱… III. ①数字图像处理—高等学校—教材 IV. ①TN911.73

中国版本图书馆CIP数据核字(2011)第017301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:王保家 责任编辑:吉玲

责任校对:刘怡丹 责任印制:杨曦

北京市朝阳展望印刷厂印刷

2011年5月第1版第1次印刷

184mm×260mm·15.5印张·384千字

标准书号:ISBN 978-7-111-33291-6

定价:32.00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>

读者购书热线:(010) 88379203

封面无防伪标均为盗版

前 言

随着科技的进步，数字图像技术得到了长足的发展，在各个领域获得了广泛的应用。对科技的进一步发展起着举足轻重的作用。

编者协同其研究团队在数字图像技术与应用的教学与研究中，总结了相关核心问题的解决方法，形成了本书的基本架构。

本书致力于为读者提供在数字图像工程中若干研究方向上的研究方法，其中包括了多个研究项目的积累，以及多名研究生的研究课题的成果。

本书内容分为 10 章，包括了数字图像、数字视频处理等相关内容，既有基本方法理论方面的阐述，又有对研究问题的剖析。每章内容后面均附有一至两个应用示例及其分析，为读者灵活运用所介绍的方法起到提示作用。

本书是在编者所在研究室的全体人员的共同努力下完成的，全书由朱虹统稿，并任主编，其余参与编写的人员有：王栋，邓颖娜，杨坤，杨向波，张然，金欢，史静，刘薇，邢楠，贾明桥，王文静，赵朝杰，张军，陈飞雄，雷林华，季瑞瑞，辛力，徐骁斐，钱慧芳，张瞳，沈懂，楚勇，闫思彤，周鹏。

在编写过程中，参考并引用了大量的参考文献，在此，向原作者表示感谢。由于编者水平有限，书中难免出现不当之处，敬请各位读者批评指正。

作 者

目 录

前言

第1章 引言	1
1.1 图像的基本概念	1
1.2 数字图像处理系统	3
1.2.1 静态数字图像处理	3
1.2.2 三维成像	4
1.2.3 数字视频处理	5
1.3 数字图像处理的主要研究内容	6
1.4 本书的结构安排	9
第2章 图像的增强与去噪	10
2.1 图像的灰度增强	10
2.1.1 线性对比度展宽	10
2.1.2 直方图均衡化	11
2.1.3 伪彩色增强	14
2.1.4 Retinex 图像增强方法	16
2.2 图像的噪声抑制	18
2.2.1 图像质量的客观评价指标	18
2.2.2 均值滤波	18
2.2.3 中值滤波	20
2.2.4 基于小波变换的阈值去噪	21
2.2.5 基于维纳滤波的图像去噪	23
2.2.6 基于偏微分方程的图像去噪	24
2.3 图像的锐化	26
2.3.1 一阶锐化算法	27
2.3.2 二阶锐化算法	32
2.4 应用示例1: 大雾天气下的图像清晰化	34
2.4.1 雾况的建模	34
2.4.2 光学厚度的确定与分割	35
2.4.3 清晰化处理	36
2.5 应用示例2: 具有高光抑制的夜间图像清晰化	37
2.5.1 亮度分量的估计与处理	37
2.5.2 反射分量的处理	38
2.5.3 高光抑制	39

第3章 图像的几何变换	41
3.1 齐次坐标与图像的仿射变换	41
3.2 三维图像的投影变换	42
3.2.1 投影与投影变换	43
3.2.2 透视投影	43
3.2.3 平行投影	46
3.3 图像几何畸变的校正	51
3.3.1 几何畸变的描述	51
3.3.2 畸变关系已知情况下的畸变校正	51
3.3.3 畸变关系未知情况下的畸变校正	52
3.4 应用示例1: 基于内容感知的图像非等比例缩放	53
3.4.1 能量函数的定义	53
3.4.2 低能量线的选取	53
3.4.3 敏感区域的能量提升	54
3.5 应用示例2: QR 码图像的几何畸变校正	55
第4章 图像的分割	57
4.1 阈值法	57
4.1.1 类间类内最大方差比法	57
4.1.2 最大熵法	58
4.1.3 聚类法	59
4.2 区域提取法	60
4.2.1 区域生长法	60
4.2.2 分裂-合并法	61
4.3 基于色彩特征的图像分割方法	62
4.3.1 RGB 色系下的图像分割	63
4.3.2 YCbCr 色系下的图像分割	63
4.3.3 HSV 色系下的图像分割	64
4.4 神经网络在图像分割中的应用	65
4.4.1 基于 BP 神经网络的图像分割	65
4.4.2 基于 SOFM 神经网络的图像分割	70
4.5 应用示例1: 肾小球区域的分割	72
4.5.1 基于遗传算法的肾小球区域定位	72
4.5.2 基于分水岭算法的肾小球边缘提取	73
4.5.3 肾小球区域的分割	74

4.6 应用示例2: 具有黑白混叠的二维条码的条空分割	74	6.2.2 信号的小波变换	106
4.6.1 明显白区域的提取与分割	75	6.2.3 离散小波变换与二进小波变换	107
4.6.2 明显黑区域的提取与分割	76	6.2.4 二维小波变换	108
4.6.3 边缘过渡区域与小骨架区域的确定	76	6.3 图像的多分辨率分析与 Mallat 算法	108
4.6.4 边缘过渡区域的处理	76	6.3.1 信号的多分辨率分析	108
4.6.5 小骨架区域的处理	77	6.3.2 一维 Mallat 算法	110
第5章 图像的数学形态学处理	79	6.3.3 二维 Mallat 算法	111
5.1 二值图像中的基本概念	79	6.4 应用示例: 基于小波变换的图像融合	112
5.1.1 连接关系	79	6.4.1 图像融合的概念	112
5.1.2 内部点与边界点	80	6.4.2 基于小波变换的图像融合	113
5.1.3 连接数与交叉数	80	第7章 数字视频	115
5.1.4 几何特征	81	7.1 数字视频的基本概念	115
5.2 贴标签	82	7.1.1 数字视频的采样格式	115
5.3 细线化	83	7.1.2 数字视频的格式与应用	116
5.3.1 距离变换	84	7.1.3 数字视频的数据形式	118
5.3.2 骨架线的提取	85	7.1.4 数字视频的表现特征	120
5.4 二值图像的腐蚀与膨胀	86	7.2 数字视频的常用处理技术	120
5.4.1 腐蚀	86	7.3 视频的镜头分割与关键帧的提取	121
5.4.2 膨胀	88	7.3.1 镜头切变检测	123
5.5 灰度图像的腐蚀与膨胀	90	7.3.2 镜头渐变检测	126
5.5.1 灰度值腐蚀	90	7.3.3 镜头分割方法性能的评价指标	127
5.5.2 灰度值膨胀	91	7.3.4 关键帧的提取	128
5.6 二值图像的开运算与闭运算	92	7.4 视频运动目标的检测	128
5.6.1 开运算	92	7.4.1 帧间差分法	129
5.6.2 闭运算	93	7.4.2 背景差分法	130
5.7 灰度图像的开运算与闭运算	94	7.4.3 光流法	130
5.8 应用示例1: 基于灰度腐蚀、膨胀的断层插值图像	96	7.5 视频的背景建模方法	131
5.9 应用示例2: 印章图像的预处理	98	7.5.1 单高斯背景建模	132
5.9.1 采用贴、删标签的方法去除噪声与空穴	98	7.5.2 混合高斯背景建模	133
5.9.2 采用闭运算弥补印鉴中的断裂	99	7.5.3 灰度归并法的背景建模	134
5.9.3 利用圆弧度对印鉴进行分类	99	7.5.4 基于码书的背景建模	135
第6章 图像的傅里叶变换与小波变换	101	7.6 单摄像头固定监视模式下的运动目标跟踪	138
6.1 图像的傅里叶变换	101	7.6.1 目标特征的提取	139
6.1.1 二维离散傅里叶变换	101	7.6.2 运动目标的跟踪	141
6.1.2 快速傅里叶变换	102	7.7 多摄像头联合有重叠视场模式下的运动目标跟踪	142
6.1.3 傅里叶变换在图像处理中的应用	103	7.7.1 基于单应性映射的全景图生成	143
6.2 图像的小波变换	105	7.7.2 运动目标的跟踪	145
6.2.1 小波分析的基本概念	105	7.8 多摄像头联合无重叠视场模式下的运动目标跟踪	145
		7.8.1 颜色特征的描述与匹配	146

7.8.2 速度特征的描述与匹配	147	9.2.2 哈夫曼编码	180
7.9 应用示例: 基于视频序列的手语识别 ..	149	9.2.3 静态二进制算术编码	181
7.9.1 关键帧的提取	150	9.2.4 自适应二进制算术编码	182
7.9.2 手势区域的提取	150	9.2.5 概率自适应二进制算术编码	183
7.9.3 静态手势特征的提取	151	9.3 有损压缩编码	184
7.9.4 手语识别	152	9.3.1 离散余弦变换编码	184
第8章 图像的三维重建	154	9.3.2 小波变换编码	185
8.1 可视化与三维重建的概念	154	9.3.3 标量量化编码	187
8.1.1 可视化技术	154	9.3.4 压扩量化编码	188
8.1.2 三维重建技术	155	9.3.5 矢量量化编码	189
8.2 面绘制方法	155	9.3.6 预测编码	192
8.2.1 Marching Cubes 算法	156	9.3.7 基于分形的图像压缩编码	194
8.2.2 Marching Tetrahedra 算法	161	9.4 静态图像的压缩编码标准	200
8.2.3 连接轮廓线法	162	9.4.1 JPEG 压缩编码标准	200
8.2.4 网格简化技术简介	163	9.4.2 JPEG2000 压缩编码标准简介	206
8.3 体绘制方法	164	9.5 MPEG 数字视频压缩编码标准	209
8.3.1 体绘制的基本原理	164	9.5.1 MPEG-1 视频压缩编码标准	211
8.3.2 体绘制的算法步骤	165	9.5.2 MPEG-2 视频压缩编码标准	215
8.3.3 体绘制方法的分类	169	9.5.3 MPEG-4 视频压缩编码标准	216
8.3.4 体绘制算法的加速	170	9.6 H.26X 系列数字视频压缩编码标准	219
8.4 应用示例1: 三维虚拟人脸表情的生成 (面绘制方法的应用)	170	9.6.1 H.261 视频压缩编码标准	219
8.4.1 通用中性人脸三维模型的确定	171	9.6.2 H.263 视频压缩编码标准	220
8.4.2 人脸特征点的确定及模型的调整	172	9.6.3 H.264 视频压缩编码标准	221
8.4.3 相位无关纹理图的生成	173	第10章 图像的匹配与跟踪	224
8.4.4 生成三维虚拟人脸表情	173	10.1 图像的匹配	224
8.5 应用示例2: 结肠的三维重建与虚拟内窥 (体绘制方法的应用)	174	10.1.1 基于区域特征的匹配方法	224
8.5.1 虚拟相机	174	10.1.2 基于特征点的匹配方法	225
8.5.2 路径规划	175	10.1.3 基于抽象特征的匹配方法	226
第9章 图像及视频数据压缩	177	10.2 图像的跟踪	229
9.1 图像数据的冗余	177	10.2.1 基于块运动估计的运动目标 跟踪	229
9.1.1 数据冗余的概念	177	10.2.2 运动目标的快速搜索	232
9.1.2 图像及视频数据的冗余	177	10.3 应用示例1: 车载摄像视频的稳像	235
9.2 无损压缩编码	178	10.4 应用示例2: 变背景下的运动目标 跟踪	237
9.2.1 行程编码	179	参考文献	240

第1章 引言

“一图千字”的古语，道出了人类从视觉所获得的信息量。图像带给人的信息具有很强的直观性，例如，我们能够从一张照片中感受到照片所记录时刻的年代信息。电视剧通过视频再现了曾经有过的历史，以及讲述大家都听得懂的故事，把观众带入特定的故事情节之中。这种图像信息的直观性与生动模仿性，就带给人们许多精神上的享受。同时，采用可视化技术，能够为人们提供可观察的画面，例如，现代医学影像技术等为后续的信息利用与处理提供了丰富的信息源。此外，作为无损检测手段，图像在工业生产中也起着很重要的作用。正是因为图像具有能带给人们信息直观性的特点，使得图像处理技术随着计算机技术、多媒体技术的飞速发展，得到了长足的进步；并且图像具有可反映人类第一感觉下的思维魅力，这些年来，图像技术快速地向各个领域渗透。

由于计算机处理能力的不断增强，使得图像处理的广泛应用有了可实现的硬件基础。本章首先就数字图像处理的基本概念进行介绍。

1.1 图像的基本概念

图像 是对客观存在物体的一种相似性的生动模仿与描述，是对物体的一种不完全的、不精确的，但在某种意义下是适当的表示^{[1][2]}。

图 1-1 所示是某个校园中的一景，从图中可以感受到校园的氛围与美丽，这在反映校园文化时，是一个贴切的表示，尽管这幅图能够展现的内容很有限。

按照上面对图像的定义，又可将图像按照其表述方式分为物理图像、数字图像、三维图像和数字视频等。

物理图像 是指物质或能量的实际分布。例如，光学图像是光强度的空间分布，就能够被人的肉眼看见，因此也称为可见图像。可见图像是我们所接触的、与人类的视觉特

性相吻合的通常意义下的图像。图 1-1 即为可见图像的例子。不可见的物理图像，如温度、压力、高度等的分布图，在医学诊断中所使用的以超声、放射线手段成像得到的医学影像等。这类图像是将不可见的物理量通过可视化的手段将其转换成人眼可非常方便进行识别的图像形式。

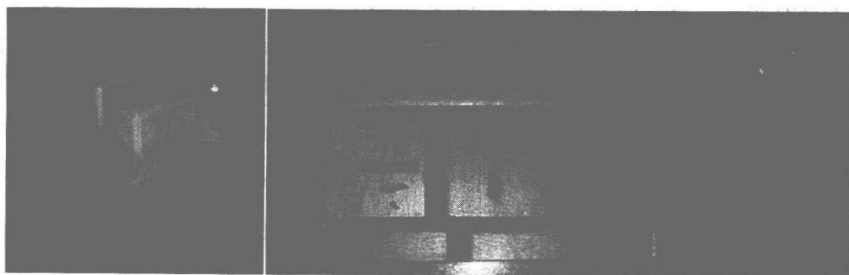
物理图像信号的好坏，在很大程度上依赖于物理信号检测设备的性能。以光学图像为例，光感应特性好的设备可以得到效果好的图像，同时，光感应器件的适应范围（即可以感知的最大、最小发光强度的范围）不同，所使用的目的也不同。



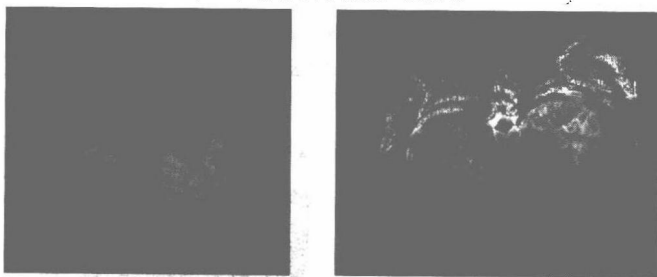
图 1-1 校园

数字图像 是用一个数字阵列来表示的图像。因此，从数学模型上来说，数字图像可以用矩阵来描述。数字阵列中的每个数字，表示数字图像的一个最小单位，称为像素。通过对每个像素点的颜色，或者是亮度等进行数字化的描述，就可以得到在计算机上进行处理数字图像。显然，数字图像可以是物理图像，也可以是虚拟图像。

三维图像 是用一个三维的数据集合来表示的图像。由于三维图像不能进行整体的直观可视化表示，因此，通常是在一个二维的平面图上显示其某个投影面，以达到表现三维物体的目的，如图 1-2 所示。图 1-2a 是一个立体实物在不同角度拍摄得到的画面，表现出人从不同视角观察到的物体信息；图 1-2b 是根据 CT 断层数据三维重建之后得到的在不同角度下的图像。可以看到，对于三维图像来说，看似平面却隐藏了奇妙的三维世界。



a) 不同角度实际拍摄的立体实物



b) 根据体数据生成的不同视角下的三维医学CT图像

图 1-2 三维图像示例

数字视频 是连续播放的数字图像序列。与单幅图像相比，又多了在一个时间轴上的图像间的变化。图 1-3 所示是从国际标准测试视频序列 BUS 中取出的三帧图像。可以看到，为了使所播放的图像序列在视觉上保持连续性，视频帧之间的差异不是很大，其视觉内容常有较大的冗余。

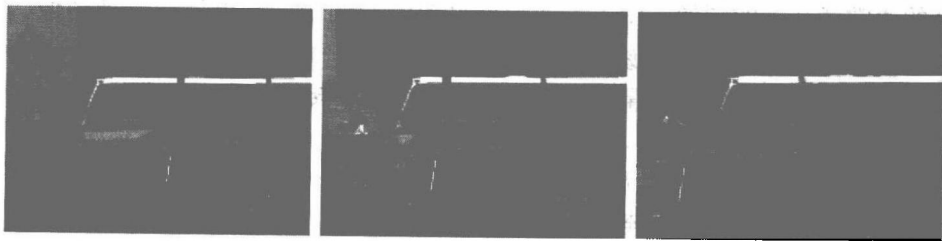


图 1-3 视频帧序列图像示例

从视频的组织结构上来讲,数字视频数据的最基本的构成元素是帧,一系列的帧可以归纳为镜头,多个镜头可以归纳为情节。一个镜头是指由一个摄像机连续拍摄得到的、在时间上连续的若干帧图像的集合,当视频的故事情节发生变化时,会出现镜头的切换。

1.2 数字图像处理系统

数字图像处理系统包括图像的采集、处理两个环节。

对于可视物理图像的采集,要获得比较好的图像质量,在很大程度上取决于光照环境的设置。如果能够人为确定光照条件,则可以充分考虑光源与对象物之间的位置关系,以及光照强度的设置等。

光源包括各种人造光源,以及白昼自然光,而光源与对象物的位置关系则基本上可以分为背光光照、正面光光照、斜射光光照几种情况。

在背光光照下,由于背景光的强度大于前景(对象物),这时,目标物的细节部分在图像中所呈现的效果就不是很好。但是,在某些工业自动化生产线上,为了快速获得目标物的定位,而不需要对目标物的细节进行观察时,常常将光源设置为背光照明方式。

在正面光光照下,如果目标物有非常光滑的反射表面,如金属表面,并且其表面是凸面的话,会在画面上产生高光区,高光区部分的颜色等细节均会退化。但是对一般的非特殊光滑表面的物体,正面照射可以获得反映目标物细节的图像。

在斜射光光照下,画面会产生光照不均的效果,如果是进行景物渲染,这是一种非常好的方法,但是当需要从画面提取相应目标物时,光照不均是阻碍正确获取目标物的一个非常严重的障碍。

显然,在构造一个数字图像处理系统时,如果允许设置光源,则需要综合考虑系统的功能目标来进行合理的设置。

与数字图像相关的研究领域,包括了计算机视觉、数字图像处理等。其所研究的对象都是图像,所使用的硬件手段都是计算机(或者是DSP处理器),并且这些研究领域所研究的内容有一定的交叉和覆盖。但是这三个研究领域所研究的内容有不同的侧重点。

计算机视觉 是指能够理解自然景物的系统。我们知道,计算机视觉是为一个设备或者是机器人提供眼睛的功能。因此,在计算机视觉的处理中,包括了三维景物信息的处理与识别,对景物中所包含目标的内容及信息进行理解,以及最终得到一个决策。例如,在一个生产线上,机械手是由三个装有吸盘的手爪构成的,当需要平稳地抓起工件时,就需要计算机视觉处理,给出机械手的三个手爪可以抓到的最大、最平稳的面。

数字图像处理 一般特指静态图像的处理,本文中,将数字图像处理的范畴扩展到三维图像,以及数字视频的处理,为此,数字图像处理的概念可以从以下几个方面来理解。

1.2.1 静态数字图像处理

静态数字图像的处理,从其获得的结果形式来分类,可以分为从图像到图像的处理以及从图像到非图像的决策处理两个方面的工作。

1. 从图像到图像的处理

这是将一幅效果不好的图像处理为视觉效果好的图像的处理方法。图1-4所示是一个智能交通信息处理系统的显示界面,图中的主画面的右下角视窗中的画面,是对大雾天气下的

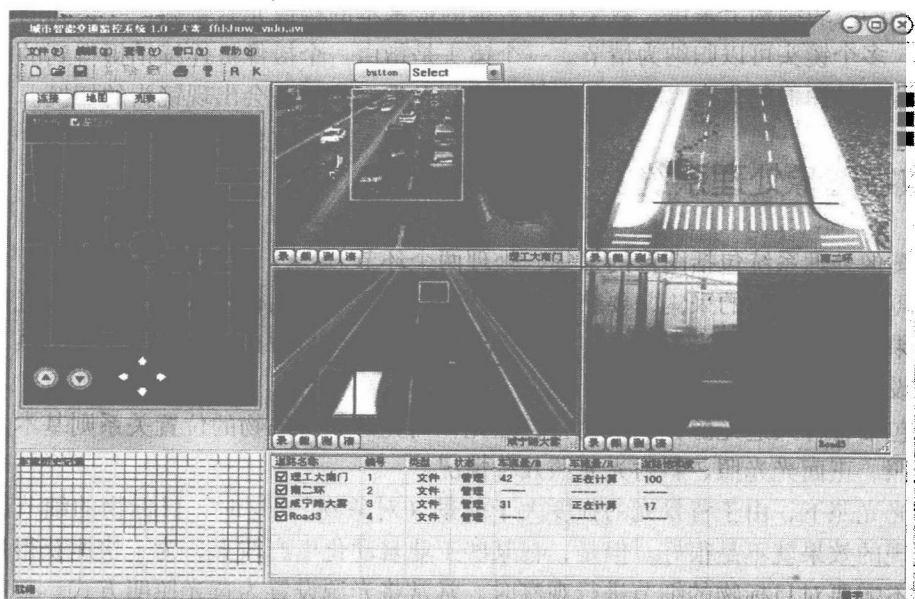


图 1-4 智能交通信息处理系统示例

监视图像进行清晰化处理的比较效果。画面的左侧是清晰化后的效果，画面的右侧是未清晰化处理之前的原始图像。经过分析可知，原图之所以能见度低，是因为在空气中悬浮着许多微小的水颗粒，这些水颗粒在光线的散射下，使景物与镜头（或人眼）之间形成了一个半透明层。这样，如果通过适当的图像处理方法，消除或减弱这层遮住人视线的大雾层，就可以得到一幅清晰的图像^[3]。

2. 从图像到非图像的决策处理

这类处理通常又称为数字图像分析。通常是对一幅图像中的若干个目标物进行识别分类后，给出其特性测度。如图 1-4 所示的智能交通信息处理系统，主画面中的左上、右上、左下三个视图均为交通状况分析画面。左上图给出了当前道路上塞车的分析状态，右上图给出了车辆异常行驶状态的轨迹分析，左下图则给出了通过虚拟检测线的车辆数目、行驶速度等状态的分析结果。这类方法在图像检测、图像测量等领域中，有着非常广泛的应用^[4]。

1.2.2 三维成像

三维成像技术采用了成像几何原理，将一幅或多幅二维图像与三维图像模型进行匹配映射，获得三维图像在某个投影面上成像的过程，称为三维成像。按照获得的平面图像信息的多少以及三维成像的原理，又可分为三维体绘制方法和三维面绘制方法。

三维体绘制方法是指在三维成像之前，从一个序列图像中获得了三维目标体的所有信息，称为体数据，根据该数据完成三维成像的方法。一个典型的例子就是医学影像中的三维 CT 图像重建技术。CT 数据是一组断层数据序列，该序列给出了三维立体目标物的所有体数据，这些体数据为三维成像提供了原始数据。

体绘制方法是对在三维空间上定义的三维物体，从任意给定的视点来跟踪体素，最终在投影面上获得三维物体的投影图像的方法。体绘制不仅能够构造出物体的表面，对其内部的信息也可以通过定义透明度的方法来显示。

图 1-5 所示是一个 CT 三维成像系统的界面。左侧主视窗中显示的就是胸片中得到的骨骼与肺的三维体绘制方法的显示结果。因为在三维成像时,有充分的三维数据作为依据,所以所得到的三维图像的细节信息比较清晰,这在医学影像辅助诊断中是非常重要的。



图 1-5 CT 三维图像重建系统示例

面绘制方法则是利用三维物体的表面数据,通过光线投影以及光线跟踪的方法,获得三维目标物在给定视角下的二维投影图像。

图 1-5 的右下侧视窗中的图像为面绘制方法得到的骨骼部分的三维立体目标物在某个视角下的投影。可以看到,因为只能给出表面信息,因此在细节表述方面不及体绘制方法,但其运算速度快是一个突出的优点。

1.2.3 数字视频处理

数字视频处理与静态图像处理的最大不同点,就是充分利用了数字视频中的图像序列在时间轴上的变化信息。

例如,图 1-6 所示是将三段国际标准测试视频组合后,进行视频镜头检测与关键帧聚类的示例。在视频节目的镜头检测与视频关键帧聚类中,通过视频帧间内容差异的分析,获得对镜头的分割,以及对分割后的镜头中的关键帧,根据其表示内容的相似性进行聚类。这类问题的研究,是数字视频处理的一类典型的研究。



图 1-6 视频镜头检测与聚类示例

数字视频处理的另一类典型的研究是对智能监控系统的人物行为自动分析方面的研究。图 1-7 所示是一个智能监控系统的人物行为的自动分析的几个示例，通过分析视频帧间内容的变化，对人物的跨越敏感检测线（见图 1-7a）、进入敏感区域（见图 1-7b），以及在监视场中聚众（见图 1-7c）等行为进行自动分析，这些方面的处理，是智能监控系统中的核心技术。

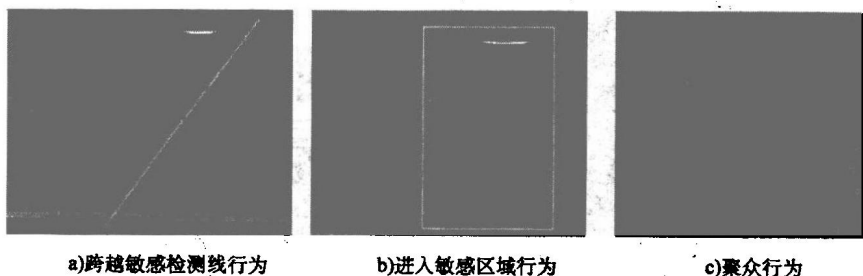


图 1-7 人物行为的自动分析示例

1.3 数字图像处理的主要研究内容

数字图像处理的主要研究内容，根据其主要的处理流程与处理目标大致可以分为图像信息的描述、图像信息的处理、图像信息的分析、图像信息的编码以及图像信息的显示、图像的三维重建、视频检索与镜头分割、视频运动目标检测与跟踪、运动目标行为自动分析等几个方面。

1. 图像的数字化

图像数字化的目的是将一幅图像以数字的形式进行表示，并且要做到既不失真又便于计算机进行处理。换句话说，图像数字化要达到以最小的数据量来不失真地描述图像信息。

图像数字化包括了采样与量化。所谓的采样是指将在空间上连续分布的图像转换成离散的图像点，量化是指将图像的亮暗信息以离散的数字来表示。

数字化技术既包括将模拟信号转换成数字信号时的采样，也包括在图像处理之前，对数字信号的重采样、重量化处理。

值得重点提到的是，受图像采集以及存储设备的物理特性的限制，当图像的分辨率达不到获取景物细节的程度时，目前采用的超分辨率技术提高原始景物的物理分辨率的方法成了研究热点之一，通过该技术完成对图像中细节的恢复^[5]。

虽然超分辨率技术属于图像恢复领域的研究内容，但其达到的效果是提高了图像的采样频率。

2. 图像的增强

图像增强的目的是将一幅图像中有用的信息（即感兴趣的信息）进行增强，同时将无用的信息（即干扰信息或噪声）进行抑制，提高图像的可观察性。

3. 图像的几何变换

图像几何变换的目的是改变一幅图像的大小或形状。例如，通过平移、旋转、放大、缩小镜像等，可以进行两幅以上图像内容的配准，以便进行图像之间内容的对比检测。例如，在印章的真伪识别，以及相似商标检测中，通常都会采用这类处理。

另外,对图像中景物的几何畸变进行校正,对图像中的目标物大小进行测量等,大多也需要有图像几何变换的处理环节。

4. 图像的恢复

图像恢复的目的是将退化了的以及模糊了的图像的原有信息进行恢复,从而使图像清晰化。图像的退化是指图像经过长时间的保存之后,因发生化学反应而使画面的颜色以及对比度发生了退化改变的现象;或者是因噪声污染等导致画面退化的现象;或者是因为现场的亮暗范围太大,导致暗区或高光区信息退化的现象。图像的模糊则常常是因为运动,以及拍摄时镜头的散焦等原因所导致的。无论是图像的退化还是图像的模糊,本质上都是原始信息部分丢失,或者是原始信息间相互混叠,或者是原始信息与外来信息的相互混叠所造成的。因此,根据退化、模糊产生原因的不同,采用不同的图像恢复方法即可达到图像清晰化的目的。

5. 图像的隐藏

图像隐藏的目的是将一幅图像或某些可数字化的媒体信息隐藏在另一幅图像中。例如,在保密通信中,将需要保密的图像在不增加数据量的前提下,隐藏在一幅可公开的图像之中,同时要求达到不可见性及抗干扰性。

所谓的抗干扰性是指当承载着保密图像的公开图像受到各种恶意攻击,或者是经过压缩等处理之后,通过解密方法获得的保密图像损失不大或基本上没有损失。

图像隐藏技术目前还有一个非常重要的拓展应用,就是数字水印技术。数字水印在维护数字媒体版权方面起着非常重要的作用。数字水印有时允许是可见的,但必须具有抗干扰性,特别是可以抵抗二次水印的添加等。同时,数字水印技术已经不仅限于图像的隐藏,而且可以在数字化的多媒体信息之间进行隐藏,如语音中隐藏图像、图像中同时隐藏语音和文字说明等。

6. 图像的变换

图像变换是指通过一种数学映射的方法,将空间域中的图像信息转换到如频域、时频域等变换域上进行分析的数学手段。最常采用的变换有傅里叶变换、小波变换等。通过二维傅里叶变换可进行图像的频率特性分析;通过小波变换,则可以将图像进行多频段分解;通过对不同频段的不同处理,可以达到良好的效果。

除了以上所述的狭义上的图像变换之外,二维条码技术也是一种图像变换的方法。二维条码相比于一维条码最大的不同就是其承载的信息量比较大,因此,可将数字化的、包括图像在内的多媒体信息按照条码的标准转换成二维条码的形式。二维条码是可视(黑白相间的条和块)但不可直接理解的信息(不能直接看出这些条块的含义),因此,作为防伪手段,将可视并可理解的图像信息转换成可视但不可理解的条码信息,同时印刷在产品之上,则通过对条码的逆变换获得可视可理解的图像信息,将该信息与印刷在产品上的可视且可理解的信息进行比较,即可获得真伪的鉴别。

7. 图像的编码

图像编码的目的是简化图像的表达方式,压缩表示图像的数据,以便于存储和传输。图像编码主要是完成对图像数据的压缩。因为图像信息具有较强的相关特性,因此通过改变图像数据的表示方法,可对图像的数据冗余进行压缩。另外,利用人类的视觉特性,可对图像的视觉冗余进行压缩,由此达到减少描述图像的数据量的目的。

8. 图像的识别与理解

所谓的图像识别与理解是指通过对图像中各种不同的物体特征进行定量化描述之后,将所期望获得的目标物进行提取,并且对所提取的目标物进行一定的定量分析。要达到这个目的,实际上就是要实现对图像内容的理解,以及对特定目标的一个识别。因此,其核心是要完成依据目标物的特征对图像进行区域分割,获得期望目标所在的局部区域。

图像中目标物的特征描述一般包括形状特征、纹理特征以及颜色特征等。形状特征包括长宽比、圆度、面积、周长等;纹理特征则包括如砖墙、布纹这类规则纹理,以及如沙滩、草坪这类非规则纹理。

9. 图像的三维重建

图像三维重建的目的是根据二维平面图像数据构造出三维物体的图像。例如,在医学影像技术中的CT成像技术,就是将多幅断层二维平面数据重建成可描述人体组织器官三维结构的图像。有关三维图像的重建方法,在计算机图形学中有非常详细的说明。三维重建技术也成为目前虚拟现实技术以及科学可视化技术的重要基础。

10. 数字视频内容分析与理解

数字视频资源以爆发性的速度增加,使得对这些资源的管理以及在这海量数据中有效地挖掘出所需要的信息,就需要对视频内容进行分析与理解。对视频进行基于语义的分析和自动标注,是进行视频数据库管理、便于用户查询与检索的一类方法。

图1-6所给出的示例,是基于视频中的低级特征的内容描述来完成对视频的分类、组织和管理。低级特征主要包括颜色、纹理和形状特征。

基于语义的视频分析则是根据电视节目的制作特点来考虑的高级视频特征。例如,在视频节目制作过程中,首先确定视频节目的主题,然后用故事单元来表述,最后是镜头的调度与运动。因此,通过镜头与故事单元的相互关系,进行镜头分割,以及关键帧的聚类,获得故事单元,由此完成基于语义的视频分析,是基于高级视频特征的视频分类、组织与管理,这类方法更加贴近人的思维方式^[6]。

11. 视频中的运动目标检测与跟踪

运动目标检测的目的是从序列图像中将运动目标区域从背景中提取出来。运动目标区域的提取,根据摄像头是否固定拍摄,可分为背景景物固定不变情况以及背景景物随摄像头的转动而变化的情况。

对于前一种情况,虽然背景景物是相对固定不变的,但是背景受到如天气、光照、阴影等干扰的影响,使得背景随着时间的推移会发生缓慢的变化,需要对背景进行更新才能够采用背景差分法获取运动目标区域。

对于后一种情况,在运动目标运动的同时,背景也在运动,但是两者的运动规律是不相同的,为此,这种情况下的运动目标检测可采用对运动目标进行建模的方法,也可采用描述两者运动差异的方法来获得运动目标区域。

运动目标的跟踪,可以通过运动目标区域的检测直接完成其跟踪。然而,运动目标的检测算法通常比较复杂,为了保证处理的实时性,一般情况下,利用视频中运动目标在各帧间的相关性,通过匹配与运动估计的方法来进行跟踪,则是有效的办法。

12. 视频中的运动目标行为自动分析

视频中的运动目标行为自动分析是指根据运动目标的运动规律找出其行为特征的处理过

程。在智能交通信息处理中,判断车辆的违章行为就是一个典型的示例。例如,判断车辆是否压黄实线行驶,只要判断在黄实线上是否有运动目标覆盖、覆盖的运动目标是否是车辆即可。图1-7所示是对人物行为自动分析的例子,对于聚众行为的判别,如果聚集在一起的目标随着时间的增多而增多,到了一定的计数范围之后,就可判断其存在着聚众的行为。

由于视频中运动目标行为的不一致性以及环境的复杂性,通常采用模式识别的方法,以及智能方法综合来完成对目标的行为分析。

1.4 本书的结构安排

前面我们简单地介绍了数字图像处理领域所包含的主要的研究方面,以此,本书安排了以下内容。

第1章为引言,给出本书所要讨论内容的一个概貌。

第2章为图像的增强与去噪,主要介绍图像的增强、图像噪声的抑制以及图像的锐化方法,并对这章的内容给出两个实际项目的应用示例供读者参考。

第3章为图像的几何变换,主要针对图像的放大、缩小、旋转和三维投影变换以及图像的畸变校正等方法进行讨论。

第4章为图像的分割,图像分割是进行图像分析与图像识别的基础,在本章中,主要介绍图像分割中常用的阈值方法和区域提取方法,并对这章的内容给出两个实际项目的应用示例供读者参考。

第5章为图像的数学形态学处理,主要介绍二值图像,以及灰度图像的腐蚀、膨胀,开、闭运算;介绍二值图像的细线化方法以及贴标签方法;并对这章的内容给出两个实际项目的应用示例供读者参考。

第6章为图像的傅里叶变换与小波变换,在介绍傅里叶变换的基础上,主要讨论图像中常用的小波变换方法,并对这章的内容给出两个实际项目的应用示例供读者参考。

第7章为数字视频,主要介绍数字视频的数据形式,数字视频图像的表现特征,并对这章的内容给出两个实际项目的应用示例供读者参考。

第8章为图像的三维重建,在介绍可视化与三维重建概念的基础上,对三维重建方法中的体绘制与面绘制两类典型方法进行详细的介绍,并对这章的内容给出两个实际项目的应用示例供读者参考。

第9章为图像及视频数据压缩,在介绍数据冗余的基础上,讨论数字图像的无损压缩和有损压缩方法,以及数字视频的压缩标准。

第10章为图像的匹配与跟踪,重点介绍目前常用的几类运动目标的跟踪方法,并对这章的内容给出了两个实际项目的应用示例供读者参考。

由于篇幅所限,本书主要介绍了最基本的数字图像处理方法,这十章的内容相对比较独立,有一定基础的读者可以根据需要选择使用相关的章节,也可以按照本书的内容安排进行系统的学习。相信经过本书的学习,在进行相关文献阅读以及解决实际问题时,会有一个较为清晰的物理概念,并通过物理概念的建立,对问题进行正确分析,最终达到解决问题的目的。

第2章 图像的增强与去噪

图像的预处理是指在完成一个目标任务时,如果采集到的初始图像信息无法直接使用,则需对其进行前期处理,以加强目标特征、减少图像中的其他非目标特征的影响。例如,图像中会出现过暗或过亮的部分,还可能会出现噪声等,导致很难完成对特定目标的识别。图像的预处理就是在对图像进行特征提取、分割和匹配等操作之前进行的处理。预处理的目的是改善图像数据中所承载的信息,清除图像中的无关信息,去除噪声,恢复有用信息,抑制不需要的变形或者增强某些对于后续处理来说比较重要的图像特征,从而改进特征提取,增加图像分割、匹配和识别的可靠性。所以,预处理在图像处理的整个过程中占据了很重要的位置。通常根据预处理的目不同,预处理技术一般有图像增强、图像平滑、图像锐化等。

2.1 图像的灰度增强

图像的灰度增强是指通过一定的处理方法,提高图像中的亮暗对比度,由此加大亮暗差异的目标特征。由于描述一幅图像的灰度级有限,因此图像的灰度增强处理的核心思路是通过抑制非重要目标信息来增强重要目标信息。图像增强应用范围很广,涉及的处理对象的复杂度一般也很高,其难点是需要对场景中的景物进行有选择性的增强,为此,这些年来在这个方向上的研究成果非常多。例如,拟人类视觉特性增强法^[7],基于视觉感知特性增强法^[8],基于小波变换^[9]、曲波变换^[10]或其他变换的增强法^[11],基于微分^[12]或同态滤波增强法^[13],光照补偿增强法^[14],遗传优化增强法^[15]和神经网络增强法等^[16]。这里,将介绍几种典型的图像灰度增强法。

2.1.1 线性对比度展宽

对比度作为分析图像质量的依据,简单地讲它是亮与暗的对比程度。对比度表示了从黑到白的渐变层次。对比度越大,从黑到白的渐变层次就越多,灰度的表现能力越强,图像越清晰醒目;而对比度越小,画面清晰度越低,层次感越差。

有些情况下,因为某些客观原因,采集到的图像对比度不够,画面效果不够好。为了使画面中期望观察的对象能够便于识别,可以采用对比度展宽的方法调节对比度,达到改善画面效果的目的。

线性对比度展宽处理,实质上是通过降低不重要信息对比度的降低,来留出多余的空间对重要信息的对比度进行展宽。

图2-1所示是线性对比度展宽的灰度映射关系图例。假设处理前后图像的量化等级相同,并且都为8位位图,即处理前后图像的灰度等级都为 $[0, 255]$ 。

设处理前原图像的灰度为 $f(i, j)$,处理后图像的灰度为 $g(i, j)$;设原图像的重要目标区域的灰度分布在 $[f_a, f_b]$ 的范围内,对比度展宽的目的是使处理之后图像的重要目标区域的灰度分布在 $[g_a, g_b]$ 的范围内。

设 $\Delta f = (f_b - f_a)$,表示原图像重要目标区域的对比度特性;设 $\Delta g = (g_b - g_a)$,表示处