



普通高等教育“十一五”国家级规划教材

“十二五”江苏省高等学校重点教材



数字图像处理与 图像通信

(第4版)

shuzi tuxiang chuli yu
tuxiang tongxin

朱秀昌 刘峰 胡栋 ©编著



北京邮电大学出版社
www.buptpress.com



普通高等教育“十一五”国家级规划教材

“十二五”江苏省高等学校重点教材



数字图像处理与图像通信

(第4版)

朱秀昌 刘峰 胡栋 编著



北京邮电大学出版社
www.buptpress.com

内 容 简 介

本书主要介绍数字图像处理和图像通信方面的基本原理、主要技术和典型应用。全书共分13章,系统叙述了数字图像信号的基本特征,数字图像处理的基本原理和方法,静止和活动图像的压缩编码、网络传输,图像处理 and 图像通信领域的部分新技术和新应用。

本书可作为信号处理、通信工程、计算机应用、广播电视、自动控制、生医工程、地理信息等领域的工程技术人员、大专院校相关专业的高年级学生或研究生学习、应用图像处理和图像通信技术的教材。

图书在版编目(CIP)数据

数字图像处理与图像通信 / 朱秀昌, 刘峰, 胡栋编著. -- 4版. -- 北京: 北京邮电大学出版社, 2016.8
ISBN 978-7-5635-4785-2

I. ①数… II. ①朱… ②刘… ③胡… III. ①数字图像处理 IV. ①TN911.73

中国版本图书馆 CIP 数据核字(2016)第 127308 号

书 名: 数字图像处理与图像通信(第4版)

著作责任者: 朱秀昌 刘 峰 胡 栋 编著

责任编辑: 刘 颖

出版发行: 北京邮电大学出版社

社 址: 北京市海淀区西土城路10号(邮编:100876)

发行部: 电话: 010-62282185 传真: 010-62283578

E-mail: publish@bupt.edu.cn

经 销: 各地新华书店

印 刷: 北京通州皇家印刷厂

开 本: 787 mm×1 092 mm 1/16

印 张: 21.5

字 数: 560 千字

印 数: 1—3 000 册

版 次: 2002年5月第1版 / 2016年8月第4版 2016年8月第1次印刷

ISBN 978-7-5635-4785-2

定 价: 43.00 元

· 如有印装质量问题,请与北京邮电大学出版社发行部联系 ·

第4版前言

本书初版、再版及三版以来,受到了不少热心的读者及使用本书作为教材的老师和学生的肯定。正因为如此,2015年本书被江苏省教育厅认定为“十二五”江苏省高等学校重点教材。出于对广大师生和读者的回报,也是为了更加完善本教材,我们对第3版进行了修订。参照读者意见和我们的教学体验,对全书的所有部分作了仔细的审视,改正了所发现的错误,修改了一些表述不妥的地方,删减了一些过时的或冗余的内容、话语,修正和改进了一些插图。

这次修订保持了本书结构大体不变,以适应多种教学环境和不同背景读者的需要。全书大致分为三个部分:第一部分,第1~3章,为数字图像的理论基础;第二部分,第4~8章,为数字图像处理的基本内容;第三部分,第9~13章,为数字图像通信的基本内容。

本书在编写中既注意到各章节之间的关联,又特别注意到使各章节之间保持一定的独立性。因此,在实际的教学应用中,大体上可以有三种选择方案:如果课时允许,可以全书讲授,形成较系统和全面的“数字图像处理与图像通信”课程;如果课时有一定的限制,可以讲授第1~8章,形成较系统的“数字图像处理”基础课程;还可以选择第1~3章,再加上第9~14章,形成较系统的“数字图像通信”基础课程。同时,在具体的讲授过程中可以根据实际情况,增减有关的内容,尤其是可灵活处理打“*”的章节。

我们虽然尽力作了修订,估计还难免存在差错之处,诚恳地希望广大读者一如既往地关注我们的教材,并给我们提出宝贵的意见、建议和希望,这是我们今天修订和今后改进的根本动力所在。

作 者

2016年5月于南京

目 录

第 1 章 绪论	1
1.1 图像信息和图像技术	1
1.1.1 图像信息的特点	1
1.1.2 图像信号	3
1.1.3 数字图像技术	5
1.2 人眼视觉特性和图像质量评价	6
1.2.1 人眼视觉特性	6
1.2.2 图像质量的评价方法	11
1.3 图像系统的构成	14
1.3.1 图像系统的线性模型	14
1.3.2 图像处理系统	15
1.3.3 图像通信系统	16
1.4 图像处理与图像通信的发展简况	17
习题与思考	20
第 2 章 数字图像基础	21
2.1 图像信号的数字化	21
2.1.1 图像信号的频谱	22
2.1.2 二维采样定理	23
2.1.3 量化和编码	26
2.2 数字图像的表达	27
2.2.1 数字图像文件格式	27
2.2.2 视频信号的数字化	29
2.2.3 数字视频国际标准	31
2.3 图像设备和器件	34
2.3.1 图像信号的采集	34
2.3.2 图像信号的接入	38
2.3.3 图像信号的显示	38
2.3.4 视频信号的转换	42
2.3.5 图像信号的处理	42
习题与思考	44
第 3 章 图像变换	46
3.1 离散傅里叶变换	46

3.1.1	一维离散傅里叶变换	46
3.1.2	二维离散傅里叶变换	51
3.2	离散余弦变换	53
3.2.1	一维离散余弦变换	54
3.2.2	二维离散余弦变换	55
3.3	图像信号的正交基表示	56
3.3.1	变换核的一般表示	56
3.3.2	变换的矩阵表示	57
3.3.3	基本图像和基本频谱	58
3.4	沃尔什和哈达玛变换	59
3.4.1	离散沃尔什变换	59
3.4.2	离散哈达玛变换	62
3.5	离散 K-L 变换	64
3.5.1	K-L 变换	64
3.5.2	最小均方误差重建	66
	习题与思考	66
第 4 章	图像增强	68
4.1	灰度级修正	68
4.1.1	灰度变换	68
4.1.2	直方图修正	70
4.1.3	直方图规定化	74
4.2	图像的同态增晰	74
4.3	图像的平滑和去噪	75
4.3.1	图像噪声的特性	76
4.3.2	空间域邻域平均	78
4.3.3	频率域低通滤波	79
4.3.4	多幅图像平均	80
4.3.5	中值滤波	81
4.4	图像的锐化	82
4.4.1	一阶微分算子法	83
4.4.2	拉普拉斯算子法	85
4.4.3	高通滤波法	86
4.5	图像的伪彩色处理	87
4.5.1	图像的彩色表示	87
4.5.2	伪彩色处理	90
4.6	图像的几何校正	92
4.6.1	空间几何变换	93
4.6.2	像素点插值	94
	习题与思考	96

第 5 章 图像复原	98
5.1 图像的降质模型	98
5.1.1 连续图像的降质模型	99
5.1.2 几个典型的降质模型	100
5.1.3 离散图像的降质模型	104
5.1.4 降质模型的矩阵表示	106
5.2 无约束图像复原	109
5.2.1 无约束最小二乘方复原	109
5.2.2 逆滤波复原	110
5.2.3 运动模糊的消除	111
5.3 有约束图像复原	113
5.3.1 有约束最小二乘方复原	114
5.3.2 维纳滤波复原	115
5.3.3 功率谱均衡复原	117
5.3.4 平滑约束复原	118
5.4 非线性图像复原*	119
5.4.1 最大后验复原	119
5.4.2 最大熵复原	121
习题与思考	123
第 6 章 图像分割	124
6.1 图像分割定义和分类	124
6.2 基于阈值的图像分割	125
6.2.1 直接阈值法	125
6.2.2 直方图阈值法	127
6.2.3 统计最优阈值法	128
6.2.4 最大类间方差法	129
6.3 基于边界的图像分割	130
6.3.1 高斯-拉普拉斯算子法	130
6.3.2 边界模板匹配法	131
6.3.3 边界跟踪法	133
6.3.4 边界拟合法	135
6.4 基于区域的图像分割	136
6.4.1 区域生长法	136
6.4.2 分裂合并法	137
6.5 基于二值数学形态学的图像分割*	139
6.5.1 数学形态学基本概念	139
6.5.2 二值数学形态学运算	141
6.5.3 二值图像的形态学处理	142
习题与思考	147

第7章 图像重建	149
7.1 投影重建基础	149
7.1.1 投影重建方式	150
7.1.2 投影定理	151
7.2 傅里叶投影重建	154
7.2.1 基本原理	154
7.2.2 离散化处理	154
7.3 卷积逆投影重建	155
7.3.1 重建过程	156
7.3.2 重建滤波器	157
7.4 代数法重建	157
7.4.1 代数法基本原理	158
7.4.2 迭代算法	159
7.5 计算机断层扫描技术	160
习题与思考.....	161
第8章 图像处理新技术	162
8.1 图像插值技术	162
8.1.1 最近邻插值	162
8.1.2 双线性插值	163
8.1.3 双立方插值	163
8.1.4 核回归插值	164
8.2 超分辨率图像重建	167
8.2.1 超分辨率的基本概念	167
8.2.2 基于重建的超分辨率方法	169
8.2.3 基于学习的超分辨率方法	171
8.3 图像水印技术	175
8.3.1 图像水印的分类	175
8.3.2 图像水印的嵌入和提取	176
8.3.3 图像水印的基本算法	178
8.3.4 图像水印系统的性能	181
8.4 基于模型的图像分割技术	183
8.4.1 分水岭分割算法	183
8.4.2 基于主动轮廓模型的分割	185
第9章 图像压缩	188
9.1 图像的统计特性	188
9.1.1 图像的信息熵	188
9.1.2 图像的自相关函数	190
9.2 压缩编码原理	192

9.2.1 无失真编码	192
9.2.2 有限失真编码	199
9.2.3 编码性能参数	203
9.3 预测编码和变换编码	204
9.3.1 预测编码原理	204
9.3.2 最佳线性预测	205
9.3.3 变换编码原理	206
9.3.4 分块 DCT 编码	208
9.4 量化	209
9.4.1 标量量化	209
9.4.2 矢量量化	210
9.4.3 量化压缩机理	212
9.4.4 结合量化的整数变换*	212
习题与思考	214
第 10 章 静止图像编码	216
10.1 静止图像的编码传输	216
10.2 方块编码	217
10.2.1 基本方法	217
10.2.2 参数选择	218
10.2.3 改进方法	220
10.3 比特面编码	221
10.3.1 次最佳编码	221
10.3.2 子块再划分编码	222
10.3.3 其他改进方法	222
10.4 多分辨率编码	223
10.4.1 四叉树编码	223
10.4.2 金字塔编码	225
10.5 二值图像编码	225
10.5.1 跳过白色块编码	226
10.5.2 游程长度编码	227
10.5.3 JBIG 标准	228
10.6 JPEG 与 JPEG 2000 标准	230
10.6.1 JPEG 标准	230
10.6.2 JPEG 2000 标准	232
习题与思考	236
第 11 章 活动图像编码	238
11.1 视频编码传输系统	238
11.1.1 系统结构	238
11.1.2 分辨率和编码方式	239

11.2 帧间预测编码	240
11.2.1 帧间预测的依据	240
11.2.2 简单的帧间预测	241
11.2.3 编码模式的选择	242
11.3 运动估计与运动补偿	242
11.3.1 运动补偿的帧间预测	242
11.3.2 基于块的运动补偿预测	243
11.3.3 像素递归运动估计	244
11.3.4 块匹配运动估计	245
11.4 混合编码	247
11.4.1 混合编码结构	247
11.4.2 H.261 混合编码方案	248
11.5 视频编码的国际标准	250
11.5.1 H.261/H.263 建议	250
11.5.2 MPEG-1/MPEG-2 标准	253
11.5.3 MPEG-4 标准	256
11.5.4 H.264/AVC 标准	257
11.5.5 HEVC 标准	260
11.6 压缩视频的差错控制技术	264
11.6.1 信道的传输特性	264
11.6.2 信道编码技术	264
11.6.3 差错掩盖技术	265
11.6.4 抗误码编码技术	266
习题与思考	267
第12章 图像编码新方法	270
12.1 小波变换与图像编码	270
12.1.1 连续小波变换	270
12.1.2 多分辨率分析*	274
12.1.3 离散小波变换	276
12.1.4 小波变换的提升算法	278
12.1.5 图像的小波变换	279
12.2 可分级编码和多描述编码	280
12.2.1 可分级视频编码	281
12.2.2 多描述视频编码	284
12.2.3 可分级编码与多描述编码的差异	287
12.3 分布式视频编码	287
12.3.1 分布式编码原理	287
12.3.2 分布式视频编码框架	289
12.3.3 解码和边信息估计	292
12.4 图像的压缩感知	293

12.4.1	压缩感知基础	293
12.4.2	信号的稀疏表示	294
12.4.3	随机测量矩阵	295
12.4.4	优化重建算法	296
12.4.5	视频的压缩感知编码	298
第 13 章 图像的网络传输		302
13.1	通信网基础	302
13.1.1	拓扑结构和服务质量	302
13.1.2	信息交换方式	303
13.1.3	三种通信网络	304
13.1.4	信息传输方式	306
13.1.5	下一代网络	307
13.1.6	软交换	309
13.2	通信网接入技术	310
13.2.1	电话网接入	310
13.2.2	光纤宽带接入	312
13.2.3	局域网接入	312
13.2.4	宽带无线接入	312
13.2.5	移动通信网接入	313
13.3	模拟视频基带信号	314
13.3.1	噪声影响	315
13.3.2	线性失真	315
13.3.3	非线性失真	316
13.3.4	清晰度和信号带宽	317
13.4	基带信号的数字调制	317
13.4.1	多相相移键控调制	318
13.4.2	多电平正交幅度调制	318
13.4.3	正交频分复用调制	319
13.4.4	残留边带调制	321
13.5	图像通信应用系统	323
13.5.1	会议电视系统	323
13.5.2	远程监控系统	325
13.5.3	网络视频	327
习题与思考		331

第1章 绪 论

图像信息具有直观、形象、易懂和信息量大等特点,因此它是人们日常生活、生产中接触最多、最受欢迎的信息种类之一。近年来,随着超大规模集成电路、个人计算机以及通信网络技术的发展,图像信息的处理和传输无论是在理论研究还是在实际应用等方面都取得了长足的进展。尤其是因特网的普及、物联网需求的增加、无线通信的兴起、云计算的问世,对数字图像处理与图像通信技术的发展起了关键性的推动作用,而数字图像技术的发展又反过来促进和加速了相关技术的进步。

本章在简单归纳图像信息特点和数字图像的基础上,简要介绍了人眼视觉系统特性和图像质量的评价方法,给出了图像处理系统和图像通信系统的基本模型,回顾了图像信息技术发展的简要历程。

1.1 图像信息和图像技术

1.1.1 图像信息的特点

人们经常接触的信息主要有三大类,它们分别是语音信息、文本信息和图像信息。这里所说的图像信息的特点,是和语音、文本信息相比较而言,关注这些信息的发出、接收、理解和表达等方面有何不同。

1. 三类信息的比较

(1) 理解和表达方式比较

对于语音信息,人们无论是说话(发出语音信息)或听话(接收语音信息),都必须通过大脑的思维,将不同的音节“转换”为不同的含义。很显然,人们都必须遵循同一“转换”规则,使语音的含义比较明确,才能进行有效的信息交流。

文本信息是语音信息的书面“符号”表达方式,所表示的含义和相应的语音一样。但文本信息和语音信息相比更加简洁和抽象,其产生和理解,需要遵循一套严密的“转换”法则,才能将符号及其组合赋予确切的含义。文本信息的规则往往较语音信息更为复杂和严格,它所表示的内容也更为精确。为此,人们需要努力学习和掌握这些规则。

和语音、文本不同,(自然)图像反映的是人眼看到的实际存在景物,“所见即所得”,直观性强。对于通常的自然景物图像,人们不必要经过特别的学习就能自主理解,从而对同一图像也可能出现各人的解读不尽相同的情况,由此带来了图像的含义不确切问题。

(2) 接收和发出方式比较

人们在发言或聆听的过程中,基本上是一种“串行”的输入、输出方式。以发言为例,说话者必须一个字一个字地说,由若干个字形成一句话,由若干句话形成一段发言,不可能同时将这段话的每个字的所有的语音一并发出。当然,听话也是如此。

文本信息的接收或发出方式也和语音差不多,文章必须一个字一个字、一句话一句话地写

出来,读书也必须逐字、逐句地读。显然,这也是一种串行的信息交流方式。

“百闻不如一见”中的“一见”表明人眼接收图像信息的方式是一种“并行”的方式,所看之处,所有的内容尽收眼底,而不是逐行逐点地观看。可见,由于图像信息的直观和便于“并行”接收,所以尽管图像的信息量庞大,但人们的接收速度却没有问题。

(3) 图像信息的特点

和语音、文本信息相比较,图像信息(这里主要指自然场景图像,不包括某些特殊的图形、标记或图标等)主要具有以下三方面特点:

一是图像的直观性强。在一般情况下图像是外界场景的直接反映,它的内容和由眼睛直接观察到的、呈现在人们脑海里的印象非常接近,或者说人们摄取图像的方法本身就是受人眼获得图像机理的启示。

二是图像的信息量特别大。俗话说“百闻不如一见”,它表明一幅图像带给人们的信息量是巨大的。例如,可以凭一张某人的照片在人群中识别出此人,但很难依据一篇描述此人的文章(尽管可以用成千上万的文字来描述)来识别他。

三是图像的表意不确切。尤其是自然场景图像,表意存在一定的模糊性,这是相对于语音和文本信息而言的。例如,面对同一幅图像,不同的观察者会有不同的理解和感受,甚至有可能给出不同的解释,如让他们写出各自观察的内容,则几乎是各不相同。

将以上叙述的内容归纳后形成表 1.1。图像信息的直观性强,易于为人们所接受,能表达语音或文本信息难以表达的内容,这些就是图像信息备受人们欢迎的根本原因之一,也是图像处理 and 通信近来得到迅速发展的根本动力之一。

表 1.1 三类信息的特点比较

信息类型	确切性	直观性	接收方式	信息量	易于理解程度
语音信息	中	中	串行	中	难(需转换)
文本信息	好	差	串行	小	较难(需转换)
图像信息	差	好	并行	特大	易(无须转换)

2. 图像的数据量

数字图像的数据量很大。例如一幅普通的 512×512 像素组成的灰度图像,其灰度级如用 8 bit 的二进制数表示,则有 $2^8 = 256$ 级灰度,那么该图像的数据量高达 $512 \times 512 = 256$ KB。对庞大数据量的图像进行处理,必须用高速信号处理器或计算机才能胜任。如果要快速传输这一类图像数据,则必然要占用较宽的频带。再如传输标准数字电视图像所需的速率约为 160 Mbit/s,而一路标准的数字语音只需 64 kbit/s 速率,两者相差几个数量级。所以在采集、传输、存储、处理、显示等各个环节的实现上,图像处理技术难度较大,成本也高,往往形成技术瓶颈。

虽然看起来数字图像的数据量很大,但是在图像画面上,经常有很多像素具有相同或接近的灰度,各个像素之间也不是完全独立的,具有一定的相关性。如在电视画面中,同一行中相邻两个像素或相邻两行间的像素,其相关系数可达 0.9 以上,而相邻两帧之间的相关性比帧内相关性一般说来还要高些。如后续章节所述,正是由于图像存在相当大的统计相关性,所以图像数据存在压缩的可能性。

1.1.2 图像信号

虽然我们对图像很熟悉,对图像的特点也很了解,但到底什么是图像却很难下一个严格的定义。一般而言,自然图像是当光辐射能量照在物体上,经过物体的反射或透射,或由发光物体本身发出的光能量,在图像采集设备中所呈现出的、能够为人的视觉所感知的物体的视觉信息。照片、电影、电视、图画等都属于图像的范围。

1. 图像信号的数学表示

图像按其亮度等级的不同,可以分成二值图像(只有黑白两种亮度等级)和灰度图像(有多种亮度等级)两种。按其色调不同,可分为无色调的灰度(黑白)图像和有色调的彩色图像两种。按其内容的变化性质不同,有静态图像和活动图像之分。而按其所占空间的维数不同,又可分为平面的二维图像和立体的三维图像等。

图像的亮度一般可以用多变量函数来表示:

$$I = f(x, y, z, \lambda, t) \quad (1.1)$$

其中, x, y, z 表示空间某点的坐标, t 为时间轴坐标, λ 为光的波长。

当取 $z = z_0$ 时,则表示二维图像;当取 $t = t_0$, 或 I 与 t 无关时,则表示静态图像;当 λ 取为定值时则表示单色图像。

一般地,由于 I 表示的是物体的反射、透射或辐射能量,因此它是正的、有界的,即

$$0 \leq I \leq I_{\max} \quad (1.2)$$

其中, I_{\max} 表示 I 的最大值, $I = 0$ 表示黑色。

式(1.1)是一个多维函数,它不易于分析处理,为此需要采用一些有效的方法进行降维。首先,根据三基色原理可知, I 可以表示为 3 个基色分量 I_R 、 I_G 和 I_B 之和,即

$$I = I_R + I_G + I_B \quad (1.3)$$

其中,

$$\begin{cases} I_R = f_R(x, y, z, \lambda_R, t) \\ I_G = f_G(x, y, z, \lambda_G, t) \\ I_B = f_B(x, y, z, \lambda_B, t) \end{cases} \quad (1.4)$$

其中, $\lambda_R, \lambda_G, \lambda_B$ 为三个基色的波长。

为了进一步降低维数,还可以采用扫描技术,如电视摄像机、扫描仪等都采用了这种方法,此时一幅画面不是一次性获得,而是采用线性扫描的方法以极快的速度逐行逐点地采集,这与我们阅读一页书非常类似。对于平面灰度图像,式(1.1)中 $z = 0$, λ 为常数, x 和 y 分别用以 t 为参数的扫描函数表示:

$$\begin{cases} x = \psi_1(t) \\ y = \psi_2(t) \end{cases} \quad (1.5)$$

于是,高维的图像被转换为一维信号。同样,利用式(1.3)、(1.4)、(1.5)也可以再恢复出原来的高维图像。

2. 模拟视频信号

根据三基色原理,利用 R(红)、G(绿)、B(蓝)三色不同比例的混合可以表示各种色彩。摄像机在拍摄时,通过光敏器件(如 CCD, 电荷耦合器件),将光信号转换为 RGB 三基色电信号。在电视机或监视器内部,最终也是使用 RGB 信号分别控制撞击荧光屏的电子流,或液晶显示屏的驱动电路使其发光产生影像。

在实际的传输和存储中,为了和黑白电视信号兼容、节约视频信号带宽,将三基色信号按一定比例组合成亮度(Y)和色度(U, V)信号,它们之间的关系如下:

$$\begin{cases} Y=0.3R+0.59G+0.11B \\ U=B-Y \\ V=R-Y \end{cases} \quad (1.6)$$

为了使 U, V 和 Y 能在一个频带内传输,到达黑白/彩色视频信号接收兼容的目的,还需将这两个色度信号进行正交幅度调制。设 $U(t), V(t)$ 为色度信号, $Y(t)$ 为亮度信号,则经调制后的两个色度信号分别为

$$\begin{cases} u(t)=U(t)\sin(\omega_{sc}t) \\ v(t)=V(t)\Phi(t)\cos(\omega_{sc}t) \end{cases} \quad (1.7)$$

其中, $\omega_{sc}=2\pi f_{sc}$ 为色度信号的副载波角频率, $\Phi(t)$ 是开关函数。由此产生的正交幅度调制的色度信号为

$$c(t)=u(t)+v(t)=C(t)\sin[\omega_{sc}t+\theta(t)] \quad (1.8)$$

其中, $\theta(t)=\Phi(t)\arctan[V(t)/U(t)], C(t)=\sqrt{U^2(t)+V^2(t)}$ 。

$\Phi(t)$ 为开关函数,如 $\Phi(t)=1$,可表示 NTSC 电视制式的色度信号;如 $\Phi(t)=+1$ (偶数行)或 -1 (奇数行),则可表示彩色副载波逐行倒相的 PAL 制色度信号。

在 PAL 制视频中,行频 $f_h=15.625$ kHz,帧频为 25 Hz,场频为 50 Hz,色度副载频 $f_{sc}=283.75f_h=4.43$ MHz。而在 NTSC 制视频中,行频 $f_h=15.75$ kHz,帧频为 30 Hz,场频为 60 Hz,色度副载频 $f_{sc}=227.50f_h=3.59$ MHz。两种制式都采用隔行扫描的方式,图像宽高比皆为 4:3。

从视频信号的频谱上看,色度信号的副载波位于亮度信号频谱的高频端,如图 1.1 所示。这样,在亮度信号的高频部分间插经过正交调制的两个色度分量,形成彩色电视的基带信号,又称为复合电视信号或全电视信号:

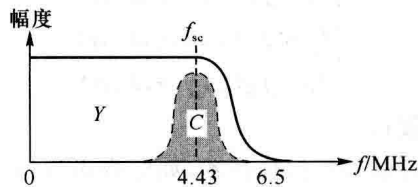


图 1.1 复合视频信号的频谱(PAL 制)

$$e(t)=Y(t)+c(t)=Y(t)+C(t)\sin[\omega_{sc}t+\theta(t)] \quad (1.9)$$

应用复合视频主要是为了方便传输以及发射。将 RGB 信号转换成 YUV 信号、Y/C 信号以及复合视频信号的过程,被称为(电视)编码,而其逆过程称为(电视)解码。另外,为了保证传送的图像能够稳定再现,实际的全电视信号还包括复合同步信号(包括行场同步、行场消隐)及色同步信号等。上面介绍的是彩色电视信号,黑白电视信号可以看作是彩色电视信号的特殊情况,此时的 $c(t)=0$ 。

除了复合视频输出外,几乎所有的视频设备都具有 S-Video 输出端子。S-Video 信号将亮度 $Y(t)$ 和色度信号 $c(t)$ 分两条线输出,避免了将 Y, C 复合起来输出,然后在应用设备中又要进行 Y, C 分离,这样的一个反复过程是有损于图像质量的。

和电影一样,视频图像也是由一系列单个静止画面组成的,这些静止画面被称为帧

(frame)。一般当帧频在每秒 20~30 帧之间时,视频图像的运动感觉就非常光滑连续,而低于每秒 15 帧时,连续运动图像就会有动画感。我国的电视标准是 PAL 制,它规定每秒 25 帧,每帧有水平方向的 625 扫描行。由于采用了隔行扫描(interlaced scan)方式,625 行扫描线分为奇数行和偶数行,这分别构成了每一帧的奇、偶两场(field)。

1.1.3 数字图像技术

图像是人类获取信息的一个重要来源,图像信息的采集、处理、存储和传输等环节也和其他信息数字化浪潮同步实现了全面的数字化。

1. 图像信息的获取、处理及应用

(1) 获取

最常见图像处理是在以计算机为中心的包括各种输入、输出、存储及显示设备在内的数字图像处理系统上进行的。为此,必须首先为计算机获取数字图像信息。早先是先获取模拟图像,再对模拟图像进行数字化之后交由计算机处理。在如今数字化时代,大多数情况是由数字设备直接获取数字图像,然后交由计算机处理。

(2) 处理

数字图像信号处理实际上是数字信号处理(Digital Signal Processing, DSP)学科的一个分支。简单地从技术角度来看,数字图像处理就是用现代数字信号处理的方法来处理数字图像信号。因此我们有必要关注数字信号处理领域的发展。

数字信号处理通常是指利用计算机或/和专用处理设备(包括器件),以数字的形式对信号进行采集、滤波、检测、均衡、变换、调制、压缩、去噪、估计等处理,以得到符合人们需要的信号形式。显然,这里所指的处理,就是对数字化后的信号施加某种数学运算。如对图像信号进行去噪,就是利用某种滤波运算,滤除或抑制混杂在图像数据中的噪声和干扰,提高图像质量,使图像看起来更加清晰。

从 1965 年快速傅里叶变换算法(FFT)提出以来,数字信号处理技术获得了重大突破。随着超大规模集成电路技术和计算机技术的发展,各种快速数字信号处理器件大量问世并得到广泛应用。目前,国际市场上涌现出多种多样的高速数字信号处理器(Digital Signal Processor, DSP)、大容量(千万门级)的现场可编程门阵列(Field Programmable Gate Array, FPGA)、高性能图像处理专用集成电路(Application Specific Integrated Circuits, ASIC)以及融合 DSP、ASIC 和 FPGA 于一体的片上系统(System On Chip, SOC)等,它们的高速运算能力解决了许多信号实时处理问题,使复杂的图像处理易于实时实现。例如,目前高性能 DSP 的处理能力已达数千 MIPS(Million Instructions Per Second)以上,集成度高达数千万门以上,而功耗却低至 0.001mW/MIPS 以下。同时,器件性能的更新周期还在逐步缩短。

(3) 应用

在近代科学研究、军事技术、工农业生产、气象、医学等领域中,人们越来越多地利用图像来认识和判断事物,解决实际问题。例如,在环境保护中,人们利用人造卫星所拍摄的地面照片,来分析获取地球资源、全球气象和污染情况;在航天探测中,利用宇宙飞船所拍摄的月球、火星等表面照片,来分析这些星球的形成;在医学上,通过 CT 断层扫描,医生可以观察和诊断人体内部是否有病变组织;在公安侦破中,采用对指纹图像提取和比对来进行侦破;在军事上,目标的自动识别和自动跟踪都需要进行高速图像处理。图像处理的应用范围愈来愈广,时至今日,已经很难找到哪个领域不需要图像技术了。

2. 图像处理的主要内容

(1) 图像变换

图像变换通常是利用正交变换,如傅里叶变换、余弦(正弦)变换、沃尔什变换、哈达码变换、小波变换等,将图像转换到变换域中进行处理。如由时间域或空间域的图像转换到频率域,在频率域进行去噪处理、压缩处理等。这些变换大多都有快速算法,从而大大提高了处理运算的速度,有利于实际应用。

(2) 图像增强

图像增强主要是指利用各种灰度函数变换方法,对图像中的像素值进行处理,加强人们感兴趣部分的图像内容,包括图像灰度修正、图像平滑、噪声去除、边缘增强等。

(3) 图像复原

在景物成像过程中,由于目标的运动、系统畸变、噪声干扰等因素,导致成像后的图像的降质(或退化)。图像复原就是把降质图像(场景)尽可能恢复成原来的图像(场景)。图像复原主要研究内容包括对图像降质因素的分析 and 降质模型的建立,以及针对降质模型的多种复原处理方法。

(4) 图像分割

图像分割的主要目标是按照具体应用要求将图像中有意义或感兴趣的部分分离或提取出来,这种分离或提取通常是根据图像的某些特征或属性来进行的。图像分割可以帮助人们进一步理解、分析或识别图像的内容;图像分割往往不是最终的目的,经常是模式识别和图像分析的预处理。

(5) 图像重建

图像重建是对一些三维物体,应用 X 射线、超声波等方法获取物体内部的结构数据,再将此数据进行重建处理而构成物体内部某些部位的图像。其中,最成功的实际应用之一就是 CT 成像技术。

(6) 图像压缩

尽管数字图像的数据量庞大,但图像信息存在较强的内在相关性,所以可以对数字图像进行数据量的压缩处理,减少其相关性,形成高效的表示方法,在满足一定图像质量要求的前提下,最大限度地减少图像的数据量,以存储更多的图像,或节省图像传输带宽。

由于图像信息处理包括的内容太多,而且随着技术的发展,学科本身的发展及学科之间的渗透、融合,有些内容已发展成为独立的学科领域,如图像融合、图像分析、图像理解等。因此,上面列举的几项内容仅是数字图像处理最为基础的部分。

1.2 人眼视觉特性和图像质量评价

一般说来,图像是给人看的,人眼是图像信息最终、最重要的接受者,也是图像质量最权威的判断者。人的视觉系统具有一些重要的特殊性能,图像处理只有尽可能地顺应或利用这些特点,才能够达到高效处理、获得高质量图像的目的。

1.2.1 人眼视觉特性

1. 人眼构造和视觉现象

图 1.2 为人眼的横截面的简单示意图。眼睛的前部为一圆球,其平均直径约为 20 mm,它