



“十三五”江苏省高等学校重点教材 (编号: 2017-2-029)
高等院校通信与信息专业规划教材

数字图像 与视频处理

DIGITAL IMAGE AND VIDEO PROCESSING



卢官明 唐贵进 崔子冠 编著



机械工业出版社
CHINA MACHINE PRESS



教学资源下载网址
<http://www.cmpedu.com>

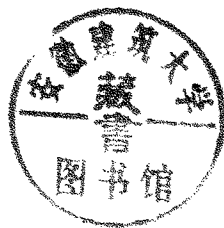




“十三五”江苏省高等学校重点教材(编号: 2017-2-029)
高等院校通信与信息专业规划教材

数字图像与视频处理

卢官明 唐贵进 崔子冠 编著



机械工业出版社

本书深入浅出地介绍了数字图像与视频处理的基本概念、基本原理、关键技术和典型应用。全书共 11 章, 主要包括图像与视频处理基础、图像增强、形态学图像处理、图像分割、数字图像与视频压缩编码原理及相关标准、图像和视频文件格式、数字水印技术、图像与视频的质量评价、基于内容的图像和视频检索、图像识别等内容。每章都附有小结与习题, 以指导读者加深对本书主要内容的理解。

本书注重选材, 内容丰富, 条理清晰, 通俗易懂, 重点突出。在强调基本概念、基本原理的同时, 注重理论与实际应用相结合, 介绍了相关领域的最新研究成果及发展新动向。

本书可作为高等院校电子信息工程、通信工程、电子科学与技术、计算机应用、广播电视工程等专业的低年级本科生或研究生的教材或教学参考书, 也可供相关专业的工程技术人员和技术管理人员阅读。

本书配套授课用电子课件等教学资源, 需要的教师可登录 www.cmpedu.com 免费注册、审核通过后下载, 或联系编辑索取 (QQ: 6142415, 电话: 010-88379753)。

图书在版编目 (CIP) 数据

数字图像与视频处理/卢官明, 唐贵进, 崔子冠编著. —北京: 机械工业出版社, 2018. 7

高等院校通信与信息专业规划教材

ISBN 978-7-111-60177-7

I. ①数… II. ①卢… ②唐… ③崔… III. ①数字图像处理-高等学校-教材②数字视频系统-数字信号处理-高等学校-教材 IV. ①TN911. 73
②TN941. 3

中国版本图书馆 CIP 数据核字 (2018) 第 146487 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

责任校对: 陈越 封面设计: 鞠杨

责任印制: 李昂

北京宝昌彩色印刷有限公司印刷

2018 年 8 月第 1 版第 1 次印刷

184mm × 260mm · 21.5 印张 · 541 千字

0001—3000册

标准书号: ISBN 978-7-111-60177-7

定价: 59.80 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

服务咨询热线: 010-88379833

机工官网: www.cmpbook.com

读者购书热线: 010-88379649

机工官博: weibo.com/cmp1952

教育服务网: www.cmpedu.com

封面无防伪标均为盗版

金书网: www.golden-book.com

前 言

当前,以数字图像与视频处理技术为核心的网络视频、智能视频分析、图像识别等领域正在积极创新,为产业结构的调整和升级带来新的机会。国内许多高校的电子信息类、数字媒体、教育技术等专业纷纷开设数字图像与视频处理相关课程,以满足社会对相关专业人才的需求。尽管国内外出版的《数字图像处理》相关教材不少,但缺少将数字图像处理技术与数字视频处理技术有机地整合在一起的教材。近年来,作者一直关注着数字图像与视频处理技术的发展,并致力于该领域的教学与研究,深感出版一本《数字图像与视频处理》教材实有必要。

编写本教材的指导思想是:将图像与视频信息的处理技术有机地整合在一起,揭示其内在的关联,以便让学生在有限的学时内掌握更系统、更全面的知识。本书的特色主要体现在以下几方面。

(1) 取材先进,内容新颖。本书充分吸收了相关领域的新技术、新标准和新成果。例如,在第1.6节视频信号的数字化中,介绍了针对4K与8K超高清显示的国际标准ITU-R BT.2020;在第2.3节,介绍了基于稀疏表示的图像去噪技术;在第2.6节,介绍了基于Retinex理论的图像增强技术;在第4.5节,介绍了基于主动轮廓模型的图像分割新方法;在第6章,介绍了H.265/HEVC、AVS+视频编码标准;在第8章,介绍了数字图像与视频水印技术;在第9章,介绍了图像和视频质量的评价方法;在第10章,介绍了基于内容的图像和视频检索;在第11章,介绍了图像识别。

(2) 结构合理,条理清晰。本书突出定性分析和系统原理框图流程分析,科学系统地归纳本学科知识点的相互联系与发展规律,符合认知规律和教学规律,富有启发性,适合教学与自学,有利于激发学生的学习兴趣及创新能力培养。

(3) 重点突出,注重实用。本书以掌握基本原理、强化应用为重点,在强调基本概念、基本原理的同时,注重理论与实际应用相结合,列举了大量具有实际应用价值的MATLAB编程实例,使学生能较快地掌握图像与视频处理的基本理论、方法、实用技术及一些典型应用,学以致用,有利于培养学生解决实际问题的能力。

本书共11章,比较系统地介绍了数字图像与视频处理的基本概念、基本原理、关键技术和典型应用,知识体系完整,结构合理,各章内容既相互独立,又兼顾其内在关联及系统性。在对不同专业或不同层次的教学进行安排时,教师可根据学生已有的知识基础和专业方向等情况,有针对性地选择其中的部分内容。对于不作为重点的教学内容,如果学生感兴趣,也可以自学。

本书的编写得到江苏省重点研发计划(BE2016775)以及“十三五”江苏省高等学校重点教材立项建设项目资助。在编写过程中,作者参考和引用了一些学者的研究成果、著作和论文,具体出处见参考文献。在此,作者向这些文献的著作者表示敬意和感谢!

本书的第9章由崔子冠老师编写,第11章及第2.3.4节、第2.3.6节、第2.6节、



第4.5节由唐贵进老师编写，其余内容由卢官明编写，全书由卢官明统审、定稿。鉴于作者水平所限，加之相关技术发展迅速，书中难免存在不妥之处，敬请同行专家和广大读者批评指正，提出宝贵意见和建议。

作者
2018年3月

目 录

前 言

第 1 章 图像与视频处理基础 1

- 1.1 光的特性与光源 1
 - 1.1.1 光的特性 1
 - 1.1.2 光通量和发光强度 2
 - 1.1.3 照度和亮度 2
- 1.2 彩色三要素与三基色原理 3
 - 1.2.1 光的颜色与彩色三要素 3
 - 1.2.2 三基色原理及应用 4
 - 1.2.3 几种典型的颜色空间模型及转换关系 7
- 1.3 人眼的视觉特性 11
 - 1.3.1 视觉光谱光视效率曲线 11
 - 1.3.2 人眼的亮度感觉特性 12
 - 1.3.3 人眼的分辨力与视觉惰性 13
- 1.4 图像信号的数字化 15
- 1.5 彩色模拟电视制式 19
 - 1.5.1 NTSC 制 20
 - 1.5.2 PAL 制 21
 - 1.5.3 SECAM 制 22
- 1.6 视频信号的数字化 23
 - 1.6.1 ITU-R BT. 601 建议 24
 - 1.6.2 ITU-R BT. 709 建议 25
 - 1.6.3 ITU-R BT. 2020 建议 27
 - 1.6.4 我国数字电视节目制作及交换用视频参数 27
- 1.7 MATLAB 在数字图像与视频处理中的应用 30
 - 1.7.1 MATLAB 简介 30
 - 1.7.2 MATLAB 中图像与视频文件的基本操作 30
 - 1.7.3 MATLAB 编程实例 34
- 1.8 小结 35
- 1.9 习题 35

第 2 章 图像增强 36

- 2.1 引言 36
- 2.2 图像的灰度变换 37
 - 2.2.1 灰度的线性变换 37

- 2.2.2 灰度的非线性变换 39
- 2.2.3 直方图修正 40
- 2.2.4 直方图规定化 45
- 2.3 图像平滑与去噪 48
 - 2.3.1 模板操作和卷积运算 48
 - 2.3.2 邻域平均法 49
 - 2.3.3 中值滤波 51
 - 2.3.4 基于非局部相似性的图像去噪 54
 - 2.3.5 频率域低通滤波 57
 - 2.3.6 基于稀疏表示的图像去噪 60
- 2.4 图像锐化 62
 - 2.4.1 梯度运算 (算子) 62
 - 2.4.2 索贝尔 (Sobel) 算子 64
 - 2.4.3 拉普拉斯 (Laplacian) 算子 64
 - 2.4.4 频率域高通滤波 66
- 2.5 图像的同态滤波 69
- 2.6 基于 Retinex 理论的图像增强 70
- 2.7 彩色增强 72
 - 2.7.1 伪彩色增强 72
 - 2.7.2 假彩色增强 74
- 2.8 MATLAB 编程实例 75
- 2.9 小结 76
- 2.10 习题 78

第 3 章 形态学图像处理 79

- 3.1 引言 79
 - 3.1.1 数学形态学的发展简史和基本思想 79
 - 3.1.2 集合论基础 80
 - 3.1.3 数学形态学中的几个基本概念 82
- 3.2 二值形态学基本运算 83
 - 3.2.1 腐蚀 83
 - 3.2.2 膨胀 84
 - 3.2.3 腐蚀运算与膨胀运算的对偶性 85
 - 3.2.4 开运算 86
 - 3.2.5 闭运算 87
- 3.3 二值图像的形态学处理 88
 - 3.3.1 边缘提取 88



3.3.2	区域填充	88	第5章 数字图像与视频压缩编码	
3.3.3	骨架抽取	89	原理	130
3.3.4	细化	90	5.1 数字图像与视频压缩编码概述	130
3.3.5	粗化	92	5.1.1 数字图像与视频压缩的必要性和可能性	130
3.3.6	形态滤波	92	5.1.2 数字图像与视频压缩编码的主要方法及其分类	132
3.4	灰度形态学基本运算	93	5.2 熵编码	133
3.4.1	灰度腐蚀	93	5.2.1 图像的信源熵	133
3.4.2	灰度膨胀	94	5.2.2 游程编码	135
3.4.3	灰度开运算与闭运算	96	5.2.3 哈夫曼编码	135
3.5	灰度图像的形态学处理	97	5.2.4 算术编码	137
3.5.1	形态学梯度	97	5.3 预测编码	142
3.5.2	形态学平滑滤波	98	5.3.1 图像差值信号的统计特性	143
3.5.3	高帽变换	98	5.3.2 帧内预测编码	144
3.6	MATLAB 编程实例	98	5.3.3 帧间预测编码	145
3.6.1	MATLAB 中形态学基本运算函数	98	5.4 变换编码	149
3.6.2	编程实例	101	5.4.1 图像的频率域统计特性	149
3.7	小结	102	5.4.2 变换编码的基本原理	149
3.8	习题	103	5.4.3 正交变换基的选择	150
			5.4.4 基于 DCT 的图像编码	151
第4章 图像分割		104	5.5 MATLAB 编程实例	156
4.1	图像分割的概念及分类	104	5.6 小结	158
4.1.1	图像分割的概念	104	5.7 习题	159
4.1.2	图像分割的依据和方法分类	105	第6章 数字图像与视频压缩编码	
4.2	基于灰度阈值化的图像分割	105	标准	160
4.2.1	阈值化分割的原理	105	6.1 静止图像编码标准	160
4.2.2	全局阈值化分割法	106	6.1.1 JPEG 标准概述	160
4.2.3	局部阈值化分割法	108	6.1.2 JPEG 基本编码系统	161
4.3	基于边缘检测的图像分割	108	6.1.3 基于 DCT 的渐进编码	162
4.3.1	边缘检测的基本原理和步骤	109	6.1.4 分级编码	162
4.3.2	梯度算子	110	6.1.5 JPEG2000 标准概述	163
4.3.3	拉普拉斯算子和 LoG 算子	111	6.1.6 JPEG2000 标准的基本框架	163
4.3.4	Canny 算子	114	6.1.7 JPEG2000 的主要特点	164
4.3.5	边缘跟踪	116	6.2 数字视频编码标准概述	165
4.4	基于区域的图像分割	118	6.2.1 H.26x 系列标准	166
4.4.1	区域生长法	118	6.2.2 MPEG-x 系列标准	168
4.4.2	区域分裂与合并法	119	6.2.3 AVS 和 AVS+ 标准	171
4.5	基于主动轮廓模型的图像分割	120	6.3 H.264/AVC 视频编码标准	172
4.5.1	参数主动轮廓模型	121	6.3.1 H.264/AVC 视频编码器的分层结构	173
4.5.2	几何主动轮廓模型	122	6.3.2 H.264/AVC 中的预测编码	174
4.6	MATLAB 编程实例	124		
4.7	小结	128		
4.8	习题	129		

6.3.3 整数变换与量化	178	7.5.1 RealMedia 文件格式	240
6.3.4 基于上下文的自适应熵编码	181	7.5.2 ASF 文件格式	242
6.3.5 H.264/AVC 中的 SI/SP 帧	182	7.5.3 QuickTime (MOV) 文件 格式	244
6.3.6 H.264/AVC 的其余特征	185	7.5.4 FLV 文件格式	246
6.3.7 H.264/AVC 的类和 FRExt 增加的 关键算法	188	7.5.5 其他流媒体文件格式	247
6.4 H.265/HEVC 视频编码标准	190	7.6 小结	248
6.4.1 H.265/HEVC 视频编码原理	190	7.7 习题	249
6.4.2 基于二叉树结构的编码单元 划分	191	第 8 章 数字水印技术	250
6.4.3 帧内预测	194	8.1 数字水印概述	250
6.4.4 帧间预测	196	8.1.1 数字水印技术的产生背景和 应用	250
6.4.5 变换与量化	199	8.1.2 数字水印的基本特征	251
6.4.6 环路滤波	200	8.1.3 数字水印系统的组成	253
6.4.7 上下文自适应的熵编码	201	8.1.4 数字水印的分类	254
6.4.8 并行化处理	202	8.2 数字图像水印算法	256
6.4.9 HEVC 的语法和语义	203	8.2.1 最低有效位方法	256
6.4.10 HEVC 的类、级和层	203	8.2.2 基于 DCT 域的方法	257
6.5 AVS 与 AVS+ 视频编码标准	204	8.3 数字视频水印的嵌入和提取方案	258
6.5.1 AVS1-P2	204	8.3.1 基于未压缩的原始视频的 水印方案	259
6.5.2 AVS1-P2 与 H.264 的比较	211	8.3.2 基于视频编码的水印方案	259
6.5.3 AVS+ 标准	212	8.3.3 基于压缩视频码流的 水印方案	259
6.6 MATLAB 编程实例	214	8.4 水印的攻击方法和对策	260
6.7 小结	216	8.4.1 简单攻击及对策	260
6.8 习题	216	8.4.2 同步攻击及对策	261
第 7 章 图像和视频文件格式	217	8.4.3 排除攻击及对策	261
7.1 资源交换文件格式	217	8.4.4 混淆攻击及对策	262
7.2 数字图像文件格式	219	8.5 MATLAB 编程实例	262
7.2.1 位图和调色板的概念	219	8.6 小结	267
7.2.2 图像文件的一般结构	220	8.7 习题	267
7.2.3 BMP 文件格式	221	第 9 章 图像与视频的质量评价	268
7.2.4 GIF 文件格式	224	9.1 常见的图像与视频失真类型	268
7.2.5 JPEG 文件交换格式	227	9.2 图像与视频质量的主观评价	271
7.2.6 其他图像文件格式	230	9.2.1 对比敏感度与视觉掩盖效应	271
7.3 常见的动画文件格式	232	9.2.2 电视图像质量的主观评价 方法	273
7.3.1 FLI/FLC 文件格式	232	9.3 图像质量的客观评价	274
7.3.2 SWF 文件格式	233	9.3.1 图像质量客观评价方法的 分类	274
7.4 数字视频文件格式	235	9.3.2 全参考图像质量评价	275
7.4.1 AVI 文件格式	235		
7.4.2 MPEG/MPG/DAT/DivX/ XviD	238		
7.5 流媒体文件格式	240		



9.3.3 半参考图像质量评价	281	10.3.1 基于内容的视频检索概述	305
9.3.4 无参考图像质量评价	283	10.3.2 视频内容的结构化	306
9.4 视频质量的客观评价	285	10.3.3 基于内容的视频检索工作 流程	307
9.4.1 全参考视频质量评价	286	10.3.4 基于内容的视频检索系统 结构	308
9.4.2 半参考视频质量评价	288	10.3.5 镜头切换的基本概念	309
9.4.3 无参考视频质量评价	289	10.3.6 镜头边界检测	311
9.5 MATLAB 编程实例	290	10.3.7 关键帧的提取	311
9.6 小结	291	10.3.8 镜头聚类(场景检测)	313
9.7 习题	291	10.4 小结	314
第 10 章 基于内容的图像和 视频检索	292	10.5 习题	314
10.1 基于内容检索技术概述	292	第 11 章 图像识别	315
10.1.1 多媒体信息的内容	292	11.1 图像识别概述	315
10.1.2 内容处理技术	292	11.2 统计学习方法	316
10.1.3 基于内容检索的查询方式	294	11.2.1 经验风险最小化	316
10.1.4 基于内容检索系统的 一般结构	294	11.2.2 结构风险最小化	317
10.1.5 基于内容的检索过程	295	11.2.3 支持向量机	318
10.1.6 基于内容检索的特点	295	11.3 人工神经网络	321
10.2 基于内容的图像检索	296	11.3.1 人工神经元模型	321
10.2.1 基于内容的图像检索概述	296	11.3.2 前馈神经网络	323
10.2.2 图像颜色特征的提取与表示	299	11.3.3 Hopfield 网络	324
10.2.3 图像纹理特征的提取与表示	301	11.3.4 卷积神经网络	325
10.2.4 图像形状特征的提取与 表示	302	11.4 基于 LeNet 网络的手写数字识别	328
10.2.5 图像空间关系特征的提取与 表示	302	11.5 MATLAB 编程实例	329
10.2.6 图像的相似性度量	303	11.6 小结	330
10.2.7 图像检索中的相关反馈 机制	304	11.7 习题	330
10.3 基于内容的视频检索	305	附录 缩略语英汉对照	331
		参考文献	336

第1章 图像与视频处理基础

本章学习目标:

- 掌握光的特性与度量的基本知识,包括光通量、发光强度、照度、亮度等主要光度学参量。
- 掌握彩色三要素、三基色原理及混色方法等色度学知识。
- 理解 RGB、YUV、YIQ、 $YCbCr$ 、HSL/HSV 等颜色空间的表示及转换。
- 掌握人眼视觉特性的知识,包括亮度感觉特性以及人眼的分辨力与视觉惰性。
- 熟悉图像、视频信号数字化的过程,掌握均匀量化的原理。
- 了解 NTSC、PAL 和 SECAM 三种兼容制彩色电视制式。
- 熟悉 ITU-R BT. 601、ITU-R BT. 709、ITU-R BT. 2020 建议和我国数字电视节目制作及交换用视频参数。
- 了解 MATLAB 中图像与视频文件的基本操作。

1.1 光的特性与光源

1.1.1 光的特性

光是一种电磁波,它具有波粒二象性——波动特性和微粒特性。电磁波包括无线电波、红外线、可见光、紫外线、X射线和宇宙射线等,它们分别占据的频率范围如图 1-1 所示。其中人眼能看见的可见光谱只集中在 $(3.85 \sim 7.89) \times 10^{14}$ Hz 的频段内,其波长范围在 380 ~ 780nm 之间。不同波长的光作用于人眼后引起的颜色感觉各不相同,可见光谱的波长由 780nm 向 380nm 变化时,人眼产生的颜色感觉依次是红、橙、黄、绿、青、蓝、紫色。

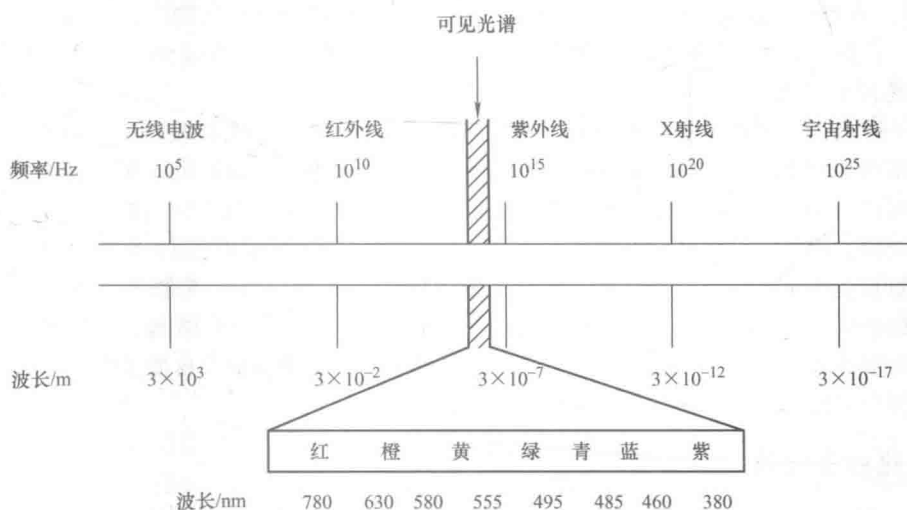


图 1-1 电磁辐射波谱



1.1.2 光通量和发光强度

通量这个术语在光辐射领域是常用的。光源辐射通量就是指其辐射功率，而光源对某面积的辐射通量是指单位时间内通过该面积的辐射能量；光源总的辐射功率（或总辐射通量）是指单位时间内通过包含光源的任一球面的辐射能量。通量与功率的意义是相同的，其单位是瓦（W）或焦 [耳] /秒（J/s）。

通常光源发出的光是由各种波长组成的，每种波长都具有各自的辐射通量。光源总的辐射通量应该是各个波长辐射通量之和。

由于在相同的亮度环境条件下，辐射功率相同、波长不同的光所引起的亮度感觉不同；辐射功率不同、波长也不相同的光可能引起相同的亮度感觉。为了按人眼的光感觉去度量辐射功率，特引入光通量的概念。

在光度学中，光通量（Luminous Flux）明确地定义为能够被人的视觉系统所感受到的那部分辐射功率的大小的度量，单位是流 [明]（lm）。

因此，只要用到光通量这个术语，首先想到它把看不见的红外线和紫外线排除在外了，而且在数量上也并不等于看得见的那部分光辐射功率值。那么，光通量的大小是怎样度量的呢？按照国际上最新的概念，它表示用标准人眼来评价的光辐射通量，其数学表达式为

$$\Phi_v = K \int_{380}^{780} \Phi_e(\lambda) V(\lambda) d\lambda \quad (1-1)$$

式中， $V(\lambda)$ 是明视觉光谱光视效率函数，人眼的视觉特性，就是从这里开始被引入到对光的定量评价中来的； $\Phi_e(\lambda)$ 是光源的辐射功率波谱； K 是一个转换常数，过去也曾称为光功当量，现在称为最大光谱效能，它的数值是一个国际协议值，规定 $K = 683 \text{ lm/W}$ ，即表示在人眼视觉系统最敏感的波长（555nm）上，辐射功率为 1W 相对应的光通量，有时称这个数为 1 光瓦。

因为人眼只对 380 ~ 780nm 的波长成分有光感觉，因此式中的积分限与此二数值相对应。由此可见，光通量的大小反映了一个光源所发出的光辐射能量所引起的人眼光亮感觉的能力。

一个 40W 的钨丝灯泡所能输出的光通量为 468lm，一个 40W 荧光灯可以输出的光通量为 2100lm。通常用每瓦流明数来表示一个光源或一个显示器的发光效率，如钨丝灯泡的发光效率为 11.7lm/W；荧光灯的发光效率为 52.5lm/W；用于电视照明的金属卤化物灯，发光效率可达 80 ~ 100lm/W。目前许多国家都在努力研制新型人工光源，并已取得不少成果，不仅提高了发光效率，而且延长了光源的使用寿命。

对于一个光源，可以说这个光源发出的光通量是多少；对于一个接收面，可以说它接收到的光通量有多少；对于一束光，可以说这束光传播的光通量是多少。从时间上讲，光通量可以是变化的，也可以是恒定的；从空间上来分析，可以导出光度学中其他几个常用的量。

一个光源，例如一个电灯泡，在它发光的时候，可以向四面八方照射，但它向各个方向所发出的光通量可能是不一样的，于是定义发光强度（Luminous Intensity）来描述在某指定方向上发出光通量的能力。发光强度的单位是坎 [德拉]（cd）。1979 年第十六届国际计量大会决定：坎德拉是一光源在指定方向上的发光强度，该光源发出频率为 $540 \times 10^{12} \text{ Hz}$ 的单色辐射，而且在此方向上的辐射强度为 $1/683 \text{ W/sr}$ （瓦 [特] 每球面度）。

1.1.3 照度和亮度

当有一定数量的光通量到达一个接收面上时，人们称这个面被照亮了，照明程度的大小可以用照度（Illuminance）来描述。照度是物体单位面积上所得到的光通量，其单位是勒 [克斯]（lx）。1lx 等于 1lm 的光通量均匀地分布在 1 m^2 面积上的光照度，即 $1 \text{ lx} = 1 \text{ lm/m}^2$ 。

下面举几个实际生活中的照度值。

教室中的标准照明是指在课桌面上的照度不低于 50lx；白天无阳光直射自然景物上的照度为 $(1 \sim 2) \times 10^4 \text{lx}$ ；晴天室内的照度为 100 ~ 1000lx；阴天自然景物上的照度约为 10^3lx ；阴天室内的照度为 5 ~ 50lx；夜间满月下为 10^{-1}lx 。

发光强度只描述了光源在某一方向上的发光能力，并未涉及光源的面积，采用单位面积上的发光强度更能反映各种光源的“优劣”，这就要用到亮度这个概念。

亮度 (Luminance) 是一个表示发光面发光强弱的物理量，表示单位面积上的发光强度，其单位是坎 [德拉] 每平方米 (cd/m^2)。

1.2 彩色三要素与三基色原理

1.2.1 光的颜色与彩色三要素

光的种类繁多，下面仅从颜色、频率成分和发光方式等方面将其分类。

- 按颜色可分为彩色光和非彩色光。非彩色光包括白色光、各种深浅不一样的灰色光和黑色光。
- 按频率成分可分为单色光和复合光。单色光是指只含单一波长成分的光或者所占波谱宽度小于 5nm 的色光；包含有两种或两种以上波长成分的光称为复合光。
- 按频率和颜色综合考虑可分为谱色光和非谱色光。谱色光主要是指波长在 780 ~ 380nm 之间，颜色按红、橙、黄、绿、青、蓝、紫顺序排列的各种光；把两个或者两个以上的单色光混合所得，但又不能作为谱色出现在光谱上的色光称为非谱色光。白光是非谱色光。

单色光一定是谱色光，非谱色光一定是复合光，而复合光也可能是谱色光。例如，红单色光和绿单色光合成的复合光为黄色，它属于谱色光。

- 按发光方式可分为直射光，反射光和透射光。发光体 (光源) 直接发出的光称为直射光；物体对光源发出的光，能够进行反射所形成的光称为反射光；能进行透射所形成的光称为透射光。若设光源的功率波谱为 $\Phi(\lambda)$ ，物体反射或透射特性分别为 $\rho(\lambda)$ 和 $\tau(\lambda)$ ，则直射光、反射光和透射光的功率波谱将分别为 $\Phi(\lambda)$ 、 $\Phi(\lambda) \rho(\lambda)$ 和 $\Phi(\lambda) \tau(\lambda)$ 。

无论是什么光，它的颜色都是取决于客观与主观两方面的因素。

客观因素是它的功率波谱分布。光源的颜色直接取决于它的辐射功率波谱 $\Phi(\lambda)$ ；而彩色物体的颜色不仅取决于它的反射特性 $\rho(\lambda)$ 和透射特性 $\tau(\lambda)$ ，而且还与照射光源的功率波谱 $\Phi(\lambda)$ 有密切关系。因此，在色度学和彩色电视中，对标准光源的辐射功率波谱，必须做出明确而严格的规定。

主观因素是人眼的视觉特性。不同的人对于同一 $\Phi(\lambda)$ 的光的颜色感觉可能是不相同的。例如，对于用红砖建造的房子，视觉正常的人看是红色，而有红色盲的人看是土黄色。

在色度学中，任一彩色光可用亮度 (Lightness，也称为明度)、色调 (Hue) 和饱和度 (Saturation) 这三个基本参量来表示，称为彩色三要素。

1. 亮度 (明度)

亮度也称明度或明亮度，是光作用于人眼时所引起的明亮程度的感觉，用于表示颜色明暗的程度。一般来说，彩色光的光功率大则感觉亮，反之则暗。就非发光物体而言，其亮度决定于由其反射 (或透射) 的光功率的大小。若照射物体的光功率为定值，则物体反射 (或透射) 系数越大，物体越明亮，反之，则越暗。对同一物体来说，照射光越强 (即光功率越大)，越明亮，反之则越暗。



亮度是非彩色的属性，用于描述亮还是暗，彩色图像中的亮度对应于黑白图像中的灰度。

2. 色调

色调是指颜色的类别，通常所说的红色、绿色、蓝色等，就是指色调。色调是决定色彩本质的基本参量，是色彩的重要属性之一，彩色物体的色调由物体本身的属性——吸收特性和反射或透射特性所决定。但是，当人们观看物体色彩时，还与照明光源的特性——光谱分布有关。色调与光的波长有关，改变光的波谱成分，就会使光的色调发生变化。例如在日光照射下的蓝布因反射蓝光而吸收其他成分而呈现蓝色，而在绿光照射下的蓝布则因无反射光而呈现黑色。对于透光物体（例如玻璃），其色调由透射光的波长所决定。例如红玻璃被白光照射后，吸收了白光中大部分光谱成分，而只透射过红光分量，于是人眼感觉到这块玻璃是红色的。

3. 饱和度（彩度）

饱和度是指彩色光所呈现色彩的深浅程度，也称为彩度。对于同一色调的彩色光，其饱和度越高，说明它的颜色越深，如深红、深绿等；饱和度越低，则说明它呈现的颜色越浅，如浅红、浅绿等。高饱和度的彩色光可以通过掺入白光而被冲淡，变成低饱和度的彩色光。各种单色光饱和度最高，单色光中掺入的白光愈多，饱和度愈低。当白光占绝大部分时，饱和度接近于零，白光的饱和度等于零。物体色调的饱和度决定于该物体表面反射光谱辐射的选择性程度，物体对光谱某一较窄波段的反射率很高，而对其他波长的反射率很低或不反射，表明它有很高的光谱选择性，物体这一颜色的饱和度就高。

色调与色饱和度合称为色度，它既说明彩色光的颜色类别，又说明颜色的深浅程度。色度再加上亮度，就能对颜色做完整的说明。

非彩色只有亮度的差别，而没有色调和饱和度这两种特性。

1.2.2 三基色原理及应用

在自然界中呈现的万紫千红的颜色，是人眼所感觉的颜色。在人眼的视觉理论研究中，眼睛视网膜的中心部分布满了锥体视觉细胞，它既有区别亮度的能力，又有区别颜色的能力。因此人们能看到自然界中的五颜六色，尤其是雨后的彩虹，黄、青、绿、紫、红、蓝的颜色给人以美的感觉。

三基色原理是指自然界中常见的大部分彩色都可由三种相互独立的基色按不同的比例混合得到。所谓独立，是指其中任何一种基色都不能由另外两种基色混合得到。三基色原理包括如下内容。

- 1) 选择三种相互独立的颜色基色，将这三基色按不同比例进行组合，可获得自然界各种彩色感觉。
- 2) 任意两种非基色的彩色相混合也可以得到一种新的彩色，但它应该等于把两种彩色各自分解为三基色，然后将基色分量分别相加后再相混合而得到的颜色。
- 3) 三基色的大小决定彩色光的亮度，混合色的亮度等于各基色分量亮度之和。
- 4) 三基色的比例决定混合色的色调，当三基色混合比例相同时，色调相同。

按照1931年国际照明委员会所作统一规定，选水银光谱中波长为700nm的红光为红基色光；波长为546.1nm的绿光为绿基色光；波长为435.8nm的蓝光为蓝基色光。常分别用R、G、B表示。当红、绿、蓝三束光比例合适时，就可以合成出自然界中常见的大多数彩色。

利用三基色原理，将彩色分解和重现，最终实现在视觉上的各种不同彩色，是彩色图像显示和表达的基本方法。

不同颜色混合在一起，能产生新的颜色，这种方法称为混色法。混色分为相加混色和相减混色。

1. 相加混色

相加混色是各分色的光谱成分相加，混色所得彩色光的亮度等于三种基色的亮度之和。彩色电视系统就是利用红、绿、蓝三种基色以适当的比例混合产生各种不同的彩色。经过对人眼识别颜色的研究表明：人的视觉对于单色的红、绿、蓝三种形式的色刺激具有相加的混合能力，例如：用适当比例的红光和绿光相加混合后，可产生与黄色光相同的彩色视觉效果；同样用适当比例的红光和蓝光相加混合后，可产生与品红色光（或称紫色光，严格地说，品红色与色谱中的紫色不同）相同的彩色视觉效果；用适当比例的蓝光和绿光相加混合后，可产生与青色光相同的彩色视觉效果。自然界中所有的万紫千红都可以用红、绿、蓝这三种颜色以适当的比例相加混合而成。相加混色的结果如图1-2所示。

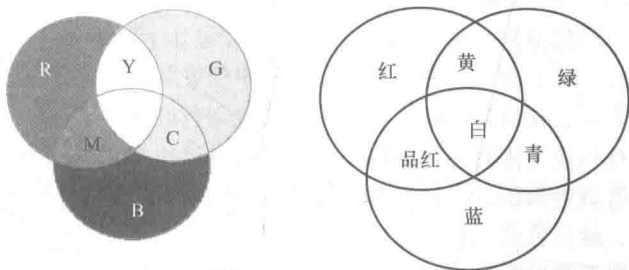


图1-2 相加混色

用等式表示为

$$\begin{aligned}
 & \text{红色} + \text{绿色} = \text{黄色} & \text{绿色} + \text{蓝色} & = \text{青色} \\
 & \text{红色} + \text{蓝色} = \text{品红色} & \text{蓝色} + \text{黄色} & = \text{白色} \\
 & \text{红色} + \text{绿色} + \text{蓝色} & = \text{白色} & \text{红色} + \text{青色} & = \text{白色} \\
 & & & \text{绿色} + \text{品红色} & = \text{白色}
 \end{aligned}$$

因为“蓝色 + 黄色 = 白色”，所以在色度学中称蓝色为黄色的补色，黄色为蓝色的补色。同样，红色和青色互为补色，绿色和紫色互为补色。也就是说三基色红、绿、蓝相对应的补色分别是青色、品红色、黄色。在彩色电视中，常用的彩条信号，即黄色、青色、绿色、品红色、红色、蓝色彩条，就是由红、绿、蓝三基色和它们对应的补色组成的。

三基色原理是彩色电视的基础，人眼的彩色感觉和彩色光的光谱成分有密切关系，但不是决定性的，只要引起的彩色感觉相同，都可以认为颜色是相同的，而与它们的光谱成分无关紧要。例如，单色青光可以由绿色与蓝色组合而成，尽管它们的光谱成分不同，但人眼的彩色感觉却是相同的。因此，在彩色视觉重现的过程中，并不一定要求重现原景象的光谱成分，而重要的是应获得与原景象相同的彩色感觉。千变万化的彩色景象，无须按其光谱成分及强度的真实分布情况来传送，只要传送其中能合成它们的三种基色就可以完全等效，并能获得与原景象相同的彩色视觉。利用三基色原理就可以大大简化彩色电视信号的传输。

实现相加混色的方法通常有以下4种。

(1) 时间混色法

时间混色法将三种基色光按一定的时间顺序轮流投射到同一平面上，只要轮换速度足够快，由于人眼的视觉惰性，分辨不出三种基色，而只能看到混合彩色的效果。如单片 DLP (Digital Light Processing, 数字光处理) 色轮技术就是利用了时间混色法。

(2) 空间混色法

空间混色法是将三种基色光分别投射到同一表面的相邻三点上，只要三点相隔足够近，由于人眼的分辨力有限，故看到的不是三种基色光而是它们的混色光。空间混色法是同时制彩色电视的基础，CRT (Cathode Ray Tube, 阴极射线管)、PDP (Plasma Display Panel, 等离子体显示器)、LCD (Liquid Crystal Display, 液晶显示器) 的显像就是利用了空间混色法。

(3) 生理混色法

当两只眼睛同时分别观看不同的颜色（例如，左眼观看红光，右眼观看绿光），人们所感



觉到的彩色不是两种单色，而是它们的混合色。立体彩色电视的显像方法就利用这种生理混色法。

(4) 全反射法

全反射法是将三种基色光以不同比例同时投射到一块全反射的平面上。由此构成了投影彩电。例如，多媒体教室中的前投影彩电、家电中的背投影彩电的显像就是利用了这种方法。

利用空间和时间混色效应，就可以对彩色图像进行空间和时间上的分割，将其分解为像素，采用顺序扫描的方式，来处理 and 传送彩色电视信号。

彩色电视从 20 世纪初到现在，经过几十年的研究和发 展，从摄像、传输到显示技术都是利用红、绿、蓝三基色原理把自然界中的五颜六色的景物显示到电视机屏幕上，供观众欣赏。就目前而言，在世界范围内，无论是模拟彩色电视机还是数字电视接收机，无论是扫描型阴极射线管电视机还是固有分辨力电视机（例如液晶电视、等离子体电视机），无论是直视型电视机还是投影型电视机，都是利用三基色原理工作的。阴极射线管电视机、等离子体电视机，选用红、绿、蓝三色荧光粉作为三基色，利用荧光粉发出的三基色光进行混合而成；LCD 电视机（包括直视型和投影型）、LCoS（Liquid Crystal on Silicon，硅基液晶）投影机都是通过光学系统滤光分色，分出红、绿、蓝三基色信号后经信号调制再相加混合而形成彩色图像。

但是，目前出现了各种不同成像原理的成像器件，有的成像器件重现还原的色域范围较小，限制了在电视中的应用，液晶面板就是其中的一种。为了提升液晶电视的彩色重现范围，生产液晶面板的一些公司研究不同的方法，改进和提高彩色的还原能力。采用四色、五色或六色滤色器面板，以提高液晶电视的彩色重现范围。对单片 DLP 投影机，为了增加亮度和彩色鲜艳度，将由过去的 R、G、B 三段色轮改造成 R、G、B、C（青）、Y（黄）、M（品红）六段色轮，并在驱动和显示电路上，实现单独地对 R、G、B、C、Y、M 进行补偿，以提高投影机的亮度和彩色鲜艳度，同时也可以根据用户的需要进行修正。

随着数字化处理技术的发展，近几年对显示器的色度处理方法也越来越多，可以根据显示器内部电子装置的需要，将一些信号从一种形式变换成另一种形式，以便完成各种处理任务。例如，首先将这些信号实时地、一个像素一个像素地变换成亮度和色度坐标，以这种形式对其进行独立处理，最后变换成电子信号，传送给显示设备进行显示。这样做的最大优点就是将信号源信号的校正与参数设置和显示器的标准和设置隔离开来，可以独立地对某种颜色进行修改和校正，可以消除灰度、色调和饱和度之间的相互作用而产生的误差，可以允许因观众喜好不同而和信号源有一定的误差等优点。还有通过对电路的设计，可以单独对红、绿、蓝和它们对应的补色分别进行修正，获得更明亮、更鲜艳的彩色，以符合某些观众对颜色的喜好。

但无论采用哪种彩色的补偿修正方法，以红、绿、蓝作为彩色电视的三基色原理是不会改变的。因为彩色电视系统到目前为止，在前端摄像机采集景物图像的颜色、演播室的节目制作和中间的节目传输都是采用红、绿、蓝三基色；而在终端显示部分，只是有些企业为渲染彩色重现效果，在电视机的信号处理电路部分分别采用“六色”或“五色”或“四色”的处理技术，但在终端显示还是以 R、G、B 三基色相加混合重现彩色图像，重现的彩色范围不会超过三基色相加混色限定的范围。

2. 相减混色

在彩色印刷、彩色胶片和绘画中的混色采用相减混色法。相减混色是利用颜料、染料的吸色性质来实现的。例如，黄色颜料能吸收蓝色（黄色的补色）光，于是在白光照射下，反射光中因缺少蓝色光成分而呈现黄色。青色染料因吸收红光成分，在白光照射下呈现青色。若将黄、青两色颜料相混，则在白光照射下，因蓝、红光均被吸收而呈现绿色。混合颜料时，每增加一种颜

料,都要从白光中减去更多的光谱成分,因此,颜料混合过程称为相减混色。在相减混色法中,通常选用青色(C)、品红(M)、黄色(Y)为三基色,它们能分别吸收各自的补色光,即红、绿、蓝光。因此,在相减混色法中,当将三基色按不同比例相混时,在白光照射下,红、绿、蓝光也将按相应的比例被吸收,从而呈现出各种不同的彩色。相减混色的结果如图1-3所示。

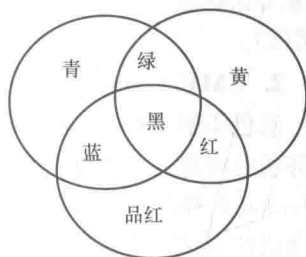


图1-3 相减混色

用等式表示为

$$\text{青} = \text{白} - \text{红} \quad \text{黄} + \text{品红} = \text{白} - \text{蓝} - \text{绿} = \text{红}$$

$$\text{品红} = \text{白} - \text{绿} \quad \text{黄} + \text{青} = \text{白} - \text{蓝} - \text{红} = \text{绿}$$

$$\text{黄} = \text{白} - \text{蓝} \quad \text{品红} - \text{青} = \text{白} - \text{绿} - \text{红} = \text{蓝}$$

$$\text{黄} + \text{青} + \text{品红} = \text{白} - \text{蓝} - \text{红} - \text{绿} = \text{黑色}$$

这种以青色(C)、品红(M)、黄色(Y)为三基色的彩色空间模型称为CMY模型。

1.2.3 几种典型的颜色空间模型及转换关系

在多媒体系统中通常用几种不同的颜色空间模型表示图形和图像的颜色,如计算机显示时采用RGB颜色空间模型;在彩色全电视信号数字化时使用 $YCbCr$ 颜色空间;彩色印刷时采用CMYK颜色空间模型等。不同的颜色空间对应不同的应用场合,在图像的生成、存储、处理及显示时对应不同的颜色空间,需要做不同的处理和转换,下面简单介绍几种典型的颜色空间模型及转换关系。

1. RGB 颜色空间模型

在多媒体计算机中,用得最多的是RGB颜色空间模型,因为计算机和彩色电视机的彩色显示器的输入需要RGB的彩色分量,通过3个分量的不同比例,在显示器屏幕上合成所需要的任一颜色。不管其中采用什么形式的颜色空间表示方法,多媒体系统最终的输出一定要转换成RGB空间表示。

在RGB颜色空间,对任意彩色光 F ,其配色方程可写为

$$F = r[R] + g[G] + b[B] \quad (1-2)$$

式中, r 、 g 、 b 为三色系数; $r[R]$ 、 $g[G]$ 、 $b[B]$ 为 F 色光的三色分量。

RGB颜色空间模型可以用笛卡尔坐标系(Cartesian coordinates)中的立方体来形象表示,3个坐标轴的正方向分别是 R 、 G 、 B 三基色,用三维空间中的一个点来表示一种颜色,如图1-4所示。每个点有3个分量,分别代表该点颜色的红(R)、绿(G)、蓝(B)三基色的值。为了方便描述,将各基色的取值范围从0~255归一化到0~1。

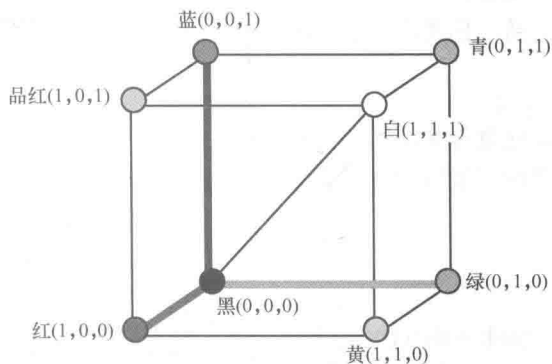


图1-4 RGB颜色空间模型

在RGB模型立方体中,原点所对应的颜色为黑色,它的3个分量值都为0。距离原点最远的顶点对应的颜色为白色,它的3个分量值都为1。从黑到白的灰度值分布在这两个点的连线上,该线称为灰色线。立方体内其余各点对应不同的颜色。彩色立方体中有3个角



对应于三基色——红、绿、蓝。剩下的三个角对应于三基色的3个补色——黄色、青色、品红色(紫色)。

2. CMY/CMYK 颜色空间模型

彩色印刷或彩色打印的纸张是不能发射光线的,因而印刷机或彩色打印机就只能使用一些能够吸收特定的光波而反射其他光波的油墨或颜料。油墨或颜料的三基色是青(Cyan)、品红(Magenta)和黄(Yellow),简称为CMY。理论上说,任何一种由颜料呈现的颜色都可以用这三种基色按不同的比例混合而成,人们称这种颜色表示方法为CMY颜色空间表示法。彩色打印机和彩色印刷系统都采用CMY颜色空间模型。

CMY颜色空间正好与RGB颜色空间互补,即用白色减去RGB颜色空间中的某一颜色值就等于这种颜色在CMY颜色空间中的值。

根据这个原理,很容易把RGB颜色空间转换成CMY颜色空间。由于彩色墨水和颜料的化学特性,用等量的CMY三基色得到的黑色不是真正的黑色,因此在彩色印刷技术中常加一种真正的黑色墨水(Black Ink)。由于B已经用来表示蓝色,因此黑色用K表示,于是CMY颜色空间也称为CMYK颜色空间。

3. YUV 和 YIQ 颜色空间模型

在现代彩色电视系统中,通常采用三管彩色摄像机或彩色CCD(Charge Coupled Device,电荷耦合器件)摄像机,它把得到的彩色图像信号,经分色,分别放大校正得到R、G、B,再经过矩阵变换电路得到亮度信号Y和2个色差信号R-Y、B-Y,最后发送端将亮度和2个色差信号分别进行编码,用同一信道发送出去。这就是PAL彩色电视制式中使用的YUV颜色空间模型和NTSC彩色电视制式中使用的YIQ颜色空间模型。其中Y表示亮度信号,U和V(I和Q)构成彩色的2个分量。

采用YUV颜色空间模型的重要性在于它的亮度信号Y和色差信号U、V是分离的。如果只有Y信号分量而没有U、V分量,那么表示的图就是黑白灰度图。彩色电视采用YUV空间模型正是为了用亮度信号Y解决彩色电视机与黑白电视机的兼容问题,使黑白电视机也能接收彩色信号。

另外,人眼对彩色图像细节的分辨能力比对黑白图像低,因此,对色度信号U和V可以采用“大面积着色原理”,即用亮度信号Y传送细节,用色度信号U、V进行大面积涂色。

根据美国国家电视制式委员会(NTSC)制式的标准,当白光的亮度用Y来表示时,它和红、绿、蓝三色光的关系可用式(1-3)描述为

$$Y = 0.299R + 0.587G + 0.114B \quad (1-3)$$

这就是常用的亮度公式。色差信号U、V是由B-Y、R-Y按不同比例压缩而成的。YUV颜色空间模型与RGB颜色空间模型的转换关系为

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1-4)$$

如果要由YUV转换成RGB,只要进行相应的逆运算即可

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.140 \\ 1 & -0.395 & -0.581 \\ 1 & 2.032 & 0 \end{bmatrix} \cdot \begin{bmatrix} Y \\ U \\ V \end{bmatrix} \quad (1-5)$$

美国、日本等国采用了NTSC制式,选用的是YIQ颜色空间模型。Y仍为亮度信号,I、Q仍