

图像处理与模式识别

范立南 韩晓微 张广渊 著

科学出版社

北京

内 容 简 介

本书结合作者对图像处理与模式识别进行的研究和工作,对图像处理与模式识别方面的几个关键问题进行了论述与探讨。全书共分10章,内容包括颜色视觉特性与颜色相似性度量、基于脉冲噪声检测的图像滤波方法、灰度图像多结构元形态学边缘检测、基于形态变换的彩色图像边缘检测、彩色图像区域分割及应用、基于不变矩方法的图像特征提取、基于BP神经网络的图像模糊分类、满文文字识别及处理等。

本书可供信息科学、数理科学、生命科学和遥感科学的科研人员和工程技术人员阅读,还可作为计算机科学与技术、自动化、电子工程、信号与信号处理等专业研究生和高年级本科生的参考书。

图书在版编目(CIP)数据

图像处理与模式识别/范立南,韩晓微,张广渊著.—北京:科学出版社,2007

ISBN 978-7-03-018728-4

I .图… II .①范… ②韩… ③张… III .①图像处理 ②模式识别
IV .TN911.73 0235

中国版本图书馆CIP数据核字(2007)第034072号

责任编辑:刘宝莉 / 责任校对:陈玉凤
责任印制:刘士平 / 封面设计:王浩

科 学 出 版 社 出 版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

中国科学院印刷厂印刷

科学出版社发行 各地新华书店经销

*

2007年3月第一版 开本:B5(720×1000)

2007年3月第一次印刷 印张:15 1/2 插页:1

印数:1—3 000 字数:296 000

定价:36.00元

(如有印装质量问题,我社负责调换(环伟))

前 言

图像是用各种观测系统以不同形式和手段观测客观世界而获得的、可以直接或间接作用于人眼并进而产生视知觉的实体。人的视觉系统就是一个观测系统，通过它得到的图像就是客观景物在人脑中形成的影像。图像是人类最重要的信息源，其直观性和易理解性是显而易见的，是其他类信息无法比拟的，因而是现代信息化社会的重要支柱。科学研究和统计表明，人所获取的信息大约 75% 来自视觉系统。

随着计算机的发展，数字图像技术近年来受到人们极大的重视取得了长足的进展，出现了许多有关的新理论、新方法、新算法、新手段和新设备，并已在科学研究、工业生产、医疗卫生、教育、娱乐、管理和通信等方面得到了广泛的应用，对推动社会发展、改善人们生活水平都起到了重要的作用。人们对这方面的研究集中在计算机视觉、图像工程、机器人视觉等研究领域，它们有个共同之处就是都基于对图像的处理与理解。由于图像本身的特殊性和复杂性，有很多问题需要研究，而图像处理与图像模式识别又是其中的研究热点。本书就图像处理与模式识别方面的几个关键问题进行阐述与研究，包括颜色相似性度量、图像滤波、图像分割及边缘检测、图像特征提取、图像分类、文字识别等。

本书基本上从专著和科研的角度出发，理论性、实用性、系统性和实时性相结合，既参考了许多他人的有关文献，更结合了作者近年来在图像处理与模式识别方面的研究成果，不仅有较系统的原理介绍，还结合科研成果给出了许多实例。

全书共分 10 章。第 1 章对人眼视觉基本原理、颜色视觉特性等方面的国内外研究成果进行了阐述，对主要的颜色空间进行了归类分析，并对各种颜色空间的转换关系进行了论述，定义了颜色相似系数，实现了颜色之间的相似性度量。第 2 章根据脉冲噪声的幅值特征，分别提出了脉冲噪声多颜色通道标量检测方法和基于颜色相似性的矢量检测方法，采用改进的自适应中值滤波方法对脉冲噪声图像有选择地滤除。第 3 章提出了多结构元的灰度形态学边缘检测梯度算法，既能提取精细的边缘，又能很好地抑制噪声。第 4 章建立了基于 HSI 颜色空间彩色形态学边缘检测梯度算法，基于该算法实现彩色图像边缘检测。第 5 章基于所提出的颜色相似性度量方法，在颜色聚类中实现图像区域分割过程，给出了一种新的彩色图像区域分割方法。第 6 章采用不变矩特征作为图像模式识别的特征矢量，讨论了利用边缘图像计算的线矩特征矢量。第 7 章基于 BP 神经网络的分类器，分别对医学图像和颜色等进行分类，采用特征矢量标准化以及模糊化预处理

理、竞争选择等一系列措施，提高了识别效果。第8章介绍了满文的特点和满文编码系统，提出了对扫描文档的倾斜检测和校正算法以及版面分析算法，并对满文单词图像进行去噪、细化和剪枝操作。第9章建立了满文笔划基元的分割、提取、特征提取和识别，以及笔划基元的组合识别方法。第10章为了优化满文识别正确率，建立了一个基于语料和规则相结合的满文知识基，实现了对脱机手写满文识别的后处理测试。

本书第1、2、5章由韩晓微编写，第3、4、6章由范立南编写，第7章由范立南和韩晓微共同编写，第8、9、10章由张广渊编写，全书由范立南统稿。

本书的研究课题得到了国家自然科学基金、中国博士后科学基金、辽宁省自然科学基金的资助。另外，在撰写本书的过程中，还参考和引用了一些论文和资料。在此一并表示衷心的感谢。

由于图像处理与模式识别技术正在迅速发展之中，加之作者水平所限，本书的内容一定有不足之处，恳请广大读者批评指正。

目 录

前言

第 1 章 颜色视觉特性与颜色相似性度量	1
1.1 引言	1
1.2 光度学基本概念	2
1.2.1 可见光范围	2
1.2.2 光强	3
1.2.3 光通量	3
1.2.4 亮度	4
1.2.5 照度	5
1.3 视觉基本原理	5
1.3.1 人眼构造和工作原理	5
1.3.2 人眼视觉模型	7
1.3.3 视觉现象	8
1.4 颜色基本特性	10
1.4.1 颜色属性	10
1.4.2 颜色的定量描述	11
1.4.3 混色原理	12
1.4.4 亮度方程	13
1.5 颜色视觉原理	13
1.5.1 视觉系统颜色感知	14
1.5.2 颜色视觉模型	15
1.6 颜色空间分析	16
1.6.1 面向设备的颜色空间	16
1.6.2 面向视觉感知的颜色空间	19
1.6.3 均匀颜色空间	23
1.6.4 色差度量	25
1.7 颜色相似性度量	26
1.7.1 RGB 与 HSI 颜色空间变换特征	27
1.7.2 颜色相似系数	27
1.8 实验结果与分析	29

1.9	本章小结	30
	参考文献	31
第2章	基于脉冲噪声检测的图像滤波方法	33
2.1	引言	33
2.2	图像噪声	34
2.2.1	噪声的来源	34
2.2.2	噪声的模型	34
2.2.3	脉冲噪声	34
2.3	灰度图像脉冲噪声滤除	35
2.3.1	脉冲噪声的图像特征分析	36
2.3.2	灰度图像脉冲噪声检测	37
2.3.3	脉冲噪声滤除	39
2.4	彩色图像脉冲噪声滤除	39
2.4.1	彩色图像脉冲噪声检测	41
2.4.2	脉冲噪声的自适应滤除	44
2.5	实验结果与分析	46
2.5.1	灰度图像	46
2.5.2	彩色图像	47
2.6	本章小结	52
	参考文献	53
第3章	灰度图像多结构元形态学边缘检测	55
3.1	引言	55
3.2	数学形态学的特点	55
3.3	二值形态学	56
3.3.1	腐蚀	56
3.3.2	膨胀	58
3.3.3	开运算	60
3.3.4	闭运算	61
3.4	灰度数学形态学的基本运算	62
3.4.1	灰度腐蚀	62
3.4.2	灰度膨胀	63
3.4.3	灰度开运算	64
3.4.4	灰度闭运算	65
3.4.5	复合形态学滤波	65
3.5	灰度图像形态学的边缘检测梯度	68

3.6	抗噪型形态学边缘检测算法	69
3.7	多结构元形态学边缘检测算法	70
3.8	结构元的选择	71
3.9	实验结果与分析	72
3.10	本章小结	78
	参考文献	78
第4章	基于形态学变换的彩色图像边缘检测	80
4.1	引言	80
4.2	图像阴影过渡区的检测研究	81
4.3	HSI 颜色空间形态学变换	82
4.4	彩色形态学开运算和闭运算	85
4.5	多结构元彩色形态学边缘检测算法	85
4.5.1	彩色边缘	85
4.5.2	彩色形态学边缘检测算法	86
4.6	实验结果与分析	87
4.7	本章小结	93
	参考文献	94
第5章	彩色图像区域分割及应用	96
5.1	引言	96
5.2	彩色图像区域分割	98
5.2.1	图像区域分割定义	98
5.2.2	彩色图像区域分割	98
5.3	足球机器人视觉系统概述	100
5.3.1	足球机器人系统	100
5.3.2	视觉子系统工作原理	100
5.3.3	足球机器人色标	101
5.3.4	色标颜色特征分析	102
5.4	足球机器人色标搜索与分割	103
5.4.1	排异聚类色标主色提取	103
5.4.2	网格搜索策略	105
5.4.3	色标区域分割	106
5.4.4	分割结果后处理	107
5.5	实验结果与分析	108
5.5.1	彩色图像区域分割	108
5.5.2	足球机器人色标搜索与分割	109

5.6	本章小结	110
	参考文献	111
第6章	基于不变矩方法的图像特征提取	113
6.1	引言	113
6.2	不变矩理论	114
6.3	线矩特征的计算	116
6.3.1	基于线积分的线矩计算	116
6.3.2	基于二值边缘图像的线矩计算	117
6.4	实验结果与分析	118
6.5	本章小结	122
	参考文献	123
第7章	基于BP神经网络的图像模糊分类	124
7.1	引言	124
7.2	多层前向神经网络	127
7.2.1	前向神经网络模型	127
7.2.2	反向传播的训练	129
7.2.3	反向传播BP网络有关问题	132
7.2.4	BP网络学习步骤	133
7.3	神经网络训练与学习的预处理	134
7.3.1	不变矩矢量标准化	134
7.3.2	以BP网络做模糊分类	135
7.3.3	竞争选择	135
7.4	基于BP神经网络的颜色分类	136
7.4.1	颜色类的选择	136
7.4.2	颜色样本的获取	137
7.4.3	颜色特征提取	138
7.4.4	网络节点数量的确定	138
7.4.5	网络的训练与分类	139
7.5	实验结果与分析	139
7.5.1	医学图像分类	139
7.5.2	颜色分类	147
7.6	本章小结	151
	参考文献	152
第8章	满文文字识别预处理	153
8.1	引言	153

8.2	满文编码系统	155
8.2.1	满文编码现状	155
8.2.2	满文音码编码方案	156
8.2.3	满文形码编码方案	158
8.3	二值化处理	161
8.3.1	整体阈值二值化	162
8.3.2	局部阈值二值化	163
8.3.3	动态阈值二值化	163
8.4	倾斜校正	163
8.4.1	图像的倾斜检测	163
8.4.2	图像的旋转	167
8.5	版面分析	169
8.5.1	版面分析的对象与任务	169
8.5.2	版面分析的方法	170
8.5.3	基于投影的版面分析	171
8.6	满文单字识别预处理	174
8.6.1	噪声过滤	175
8.6.2	细化	176
8.6.3	剪枝	179
8.7	本章小结	180
	参考文献	180
第9章	满文文字识别	183
9.1	引言	183
9.2	文字识别的原理和方法	184
9.2.1	统计模式识别	184
9.2.2	结构模式识别	186
9.2.3	统计识别与结构识别的结合	187
9.2.4	人工神经网络	187
9.3	文字识别现状	188
9.3.1	蒙古文特点	189
9.3.2	蒙文的切分和识别	190
9.4	满文文字识别系统结构	190
9.5	基于笔划基元的满文单字识别	192
9.5.1	笔划基元的划分与提取	192
9.5.2	笔划基元的特征提取与识别	196

9.5.3	游离笔划的识别	199
9.5.4	满文基元的组合识别	200
9.6	基于二级子结构引导的满文识别	201
9.7	基于 WNN 的满文连接笔划识别	202
9.7.1	小波神经网络	202
9.7.2	网络的具体训练算法	204
9.7.3	基于 WNN 的笔划识别	205
9.8	实验结果与分析	205
9.9	本章小结	207
	参考文献	208
第 10 章	满文文字识别后处理	211
10.1	引言	211
10.2	后处理的原理和方法	211
10.2.1	后处理的重要性	211
10.2.2	对单字识别器输出的要求及其描述	212
10.2.3	基于词匹配的文字识别后处理	214
10.2.4	基于句法-语义分析的后处理	218
10.2.5	基于人工神经网络的文字识别后处理	220
10.2.6	基于统计方法的文字识别后处理	222
10.3	基于知识基的满文识别后处理	226
10.3.1	系统结构	226
10.3.2	词典	227
10.3.3	规则描述语言	228
10.3.4	语料标注	229
10.3.5	后处理策略	232
10.4	实验结果与分析	233
10.5	本章小结	233
	参考文献	234

第 1 章 颜色视觉特性与颜色相似性度量

1.1 引 言

颜色是光作用于人眼所引起的视觉感知现象，是计算机视觉、多媒体和虚拟现实领域基础研究的一部分，与认知学、生理学和信息科学等领域的研究密切相关，共同发展。在生物学领域，人们希望通过对颜色信息处理的研究来了解人脑信息表象和处理过程。在计算机视觉领域，人们试图借助对人脑颜色信息处理模型的研究成果，设计颜色信息的表象和处理方法，实现景物理解。颜色的应用有两个主要原因：第一，颜色是一个强有力的描绘子，它常常可简化目标物的区分及从场景中抽取目标。第二，人可以辨别 30 多万种不同的颜色。与灰度图像相比，彩色图像由于具有了颜色，因此携带了更多的可视信息。颜色在客观上是光的一种特性，在主观上是一种视知觉现象，是人的感觉器官对外界刺激的反映。对颜色的感觉决定于光的物理本质，也决定于光在人的眼-脑视觉系统中引起的反应。与一般的物理量不同，颜色包含着复杂的生理和心理过程。

不同波长的光作用于人眼后产生了颜色的主观感觉。光作为一种能量进入人眼，汇聚在视网膜上，刺激感光细胞，最后以电脉冲形式传入大脑皮层，产生颜色感觉。光本身并没有颜色，光是通过眼睛的光感机制而被感知。人类神经系统获得了外界环境的不同波长光的信息，并对光信息进行转换才产生了颜色。颜色是人类视觉系统内在的特性，依赖于视觉而存在，颜色和其他感觉的存在一样是人们以一种特殊的方式理解这个世界。

颜色分析是彩色图像处理的重要基础。颜色分析首先要对光谱分布和光的度量等进行研究，建立以光的照射、反射现象为基础的物理模型，从而解释光的辐射、传播、照射、反射和吸收等方面现象。由于颜色是光作用于人眼的结果，因此颜色分析还需要对人眼视觉系统进行研究，建立一个符合客观实际的人眼视觉模型。

颜色空间是用来表示颜色的三个参数所构成的三维空间，是颜色抽象表示和描述的方法，是在某些标准下用通常可接受的方式来简化颜色规范。因此，颜色空间是进行颜色信息研究的理论基础。表达颜色的空间可以是非均匀、非线性的，用于彩色图像处理的空间必须同时具备独立性和均匀性的要求。独立性的要求是指颜色空间的三个分量不互相影响，对其中某个分量的处理不会导致其他分

量相对于人眼感觉发生变化。均匀性的要求是指对颜色空间中的每一个分量来说,在不同取值处的相同改变量引起的视觉变化是相同的。可以说,彩色图像处理与灰度图像处理的区别主要在于建立一个能满足上述条件的颜色空间,而具体对各个分量的处理则完全可以参照成熟的灰度图像处理技术。

颜色空间要求能形象、直观地反映人眼的颜色视觉特性,使得人们可以有目的地进行各种处理,同时又能保证在各种处理过程中都能达到预期的处理效果而不产生负面影响。由于人眼对颜色的感觉是一个复杂的非线性关系,绝对满足上述条件的颜色空间是不存在的,但是可以实验为基础,寻找一个能在较大范围内符合上述条件的颜色表达空间。

为了定量地表达自然界丰富多样的色彩,人们借助于数学坐标,已先后提出了各种各样的颜色空间。在颜色空间中,颜色由三维坐标表示,这个坐标描述了颜色在特定颜色空间中的位置。有很多种颜色空间可以精确地表达任意一种颜色,但由于颜色感知所固有的主观性和非线性,它们大多不能直接运用于彩色图像处理,必须经过必要的非线性变换和可视化处理,这个过程需要花费大量的计算时间,是一个要在软硬件设计中综合考虑的因素。

颜色空间把抽象的、主观的视觉感受具体化为三维空间中的位置、矢量,使各种彩色图像、彩色设备的颜色特性都能实现可视化,成为颜色研究的重要工具。各种混色系统都有其相应的颜色空间,不同的颜色空间具有不同的特性,分别具有各自的优缺点和局限性。在进行彩色图像处理之前,首先应选择合适的颜色空间。

本章内容是后续章节必要的理论基础。

1.2 光度学基本概念

光度学是光学中研究光的辐射、吸收、照射、反射、散射、漫射等有关光的度量的学科。光度学研究是从辐射的角度对图像进行分析和度量,用光度学对于人眼所见的图像进行度量,结合人眼的视觉特征来确定光的度量及所使用的单位。使用光电器件接受图像信息时,亮度刺激由辐射强度和接受器件的敏感特性共同决定,此时往往把电磁辐射作为纯物理现象来研究,应采用辐射量值系统;通过人眼观察图像或景物时,人们往往只关心辐射强度和人眼光敏感特性共同作用的效果,即视觉系统受刺激的程度,需要采用光学量值系统。

1.2.1 可见光范围

作为能量和信息的载体,电磁辐射的波谱范围很广,可见光是电磁光谱中非常小的一部分。要严格定义可见光的波长范围是困难的,一般说来,可以取 380~

780nm 波长之间, 其中红色光线有最长的波长, 而紫外光线的波长最短, 如图 1.1 所示。然而, 当光很强时, 人眼又在暗适应的情况下, 可感受的波长至少可扩大到 350~900nm。有时在色度学计算中, 取 400~700nm 也可以得到一定的精度。不同波长的可见光作用于视觉系统时, 不仅会产生不同的颜色感觉, 还会产生不同的亮度感觉。

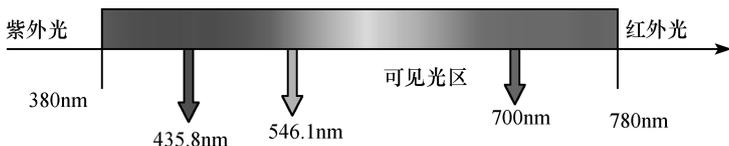


图 1.1 可见光

实际上, 见到单色的机会很少, 所见到的光都是混合光。一定成分的混合光, 有一种确定的颜色与之对应; 但反过来, 一种颜色并不只对应一种光谱组合。两种光谱成分完全不同的混合光, 有可能引起同样的颜色感觉。这种现象在色度学中被称为“同色异谱”。

1.2.2 光强

设某个点光源向各个方向都均匀辐射, 则光强可定义为发射到单位立体角 $d\Omega$ 的光通量 $d\Phi$ 的比值, 即 $I = d\Phi/d\Omega$ 。光强 I 的单位为坎 (cd); Ω 的单位为球面度 (sr); Φ 为光通量, 单位为流明 (lm)。

光强表示光源在给定方向上的发光强度, 单位为坎, 坎定义为光源发出频率为 540×10^{12} Hz 的单色辐射光, 且在此指定方向上的辐射强度为 $1/683 \text{ W/sr}$, W 为能量单位, 由于点光源向四周均匀辐射, 而球心对球面总立体角有 4π 球面度, 所以 $I = \Phi/(4\pi)$ 。1cd 光强的点光源发出的总光通量为 $4\pi \text{ lm}$ 。

1.2.3 光通量

光源以电磁波的形式辐射出的光功率称为光通量。光通量是用于度量可见光对人眼的视觉刺激程度的量, 与辐射度量中的辐通量对应, 单位为流明 (lm)。单位流明定义为发光强度为 1cd 的均匀点光源在一球面立体角内发射的光通量, 即 $1 \text{ lm} = 1 \text{ cd} \cdot \text{sr}$ 。流明的测量常以人眼的光感觉来度量其辐射功率。人眼对不同波长、相同辐通量的可见光的敏感程度不同, 这表明人的视觉系统对不同波长可见光有不同的相应灵敏度, 称为光谱光效率函数。观察场明暗程度不同时, 人眼光谱敏感特性稍有不同。国际照明委员会 (CIE) 根据实验确定两种光谱光效率函数: 明视觉光谱光效率函数 $V(\lambda)$ 和暗视觉光谱光效率函数 $V'(\lambda)$, 函数曲线如图 1.2 所示, 这里函数值已经归一化。

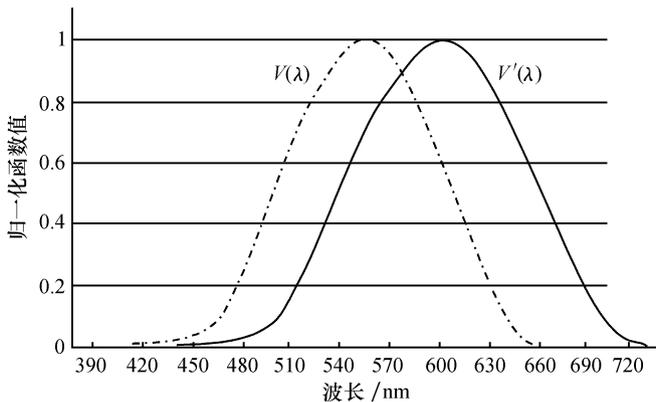


图 1.2 光谱光效率函数

图 1.2 中， $V(\lambda)$ 是在亮视场观察时的光谱光效率函数，最敏感波长为 555nm； $V'(\lambda)$ 是在暗视场观察时的光谱光效率函数，最敏感波长为 607nm。

对于单一波长的光源来讲，光通量 Φ 与辐射功率的关系为

$$\Phi = kP(\lambda)V(\lambda) \tag{1.1}$$

式中， $k=683\text{lm/W}$ ， $P(\lambda)$ 为光源辐射功率的波谱分布。

在可见光范围内对所有波长求和，即为光源的光通量：

$$\Phi = k \int_{\Delta} P(\lambda)V(\lambda)d\lambda \tag{1.2}$$

式中，积分的 Δ 范围为可见光范围，即 $\Delta=380\sim780\text{nm}$ 。

1.2.4 亮度

亮度 B 是发光面的明亮程度的度量，决定于单位面积的光强。光可以从一个面光源直接辐射，也可以由入射光照射下从单一表面反射出来或者是某光源经介质透射出来。从某一角度 α 去看发光面 S ，如图 1.3 所示，亮度与该角度的光强 I_{α} 成正比，与所见的面积 $S\cos\alpha$ 成反比。其亮度定义为 $B = I_{\alpha} / (S\cos\alpha)$ 。 I_n 为发光面法线方向的发光强度， I_{α} 为 S 发光面在指定方向的发光强度， α 为该方向与法线的夹角。在任何方向看到的 I_{α} 越大或 $S\cos\alpha$ 越小，则亮度 B 越大，即明亮程度越高。按前述的理想漫射的面光源，即光强按余弦规律分布，则

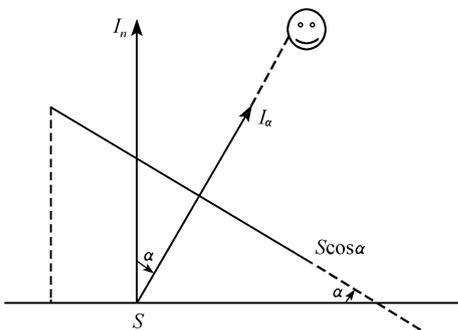


图 1.3 发光面的视在面积

$$B = \frac{I_{\alpha}}{S \cos \alpha} = \frac{I_n \cos \alpha}{S \cos \alpha} = \frac{I_n}{S} \quad (1.3)$$

也就是说,理想漫射的面光源的亮度与方向角度 α 无关。亮度的单位是坎每平方米 (cd/m^2)。

1.2.5 照度

照射在单位面积上的光通量就是照度 E , 即 $E = \Phi/S$, 单位为勒 (lx), lx 定义为在 1m^2 的面积上照有 1lm 的光通量。照度与成像的好坏有很大关系, 为了得到良好的图像, 常用强光来照射。由于光强 $I = \Phi/\Omega$, 而立立体角 Ω 又和球半径 r 有关, 这样照度 E 定义为 $E = \Phi/S = I/r^2$, 即当光强 I 一定时, 照度与距离平方成反比。光源为平行光束时, 照度和距离无关。

1.3 视觉基本原理

人类视觉反映了客观刺激与人脑神经系统的相互作用, 是我们观察和理解周围环境的一个人体组成部分。视觉现象中的绝大部分是由物理和生理因素引起的, 只有极少部分是由心理因素引起。人类视觉的产生是由眼、视神经和视觉中枢的共同活动完成的, 因此研究人眼的生理结构及其物理现象十分重要。生理学家和物理学家对人眼结构进行了精确的解剖分析, 进行了有关人眼的光度、黑白感觉、彩色感觉、色度等实验, 并提出了假说。虽然还有许多机理没有完全搞清楚, 但黑白及彩色电视以及图像处理的实验结果都从实践证明这些机理是十分有效的。视觉是人类最高级的感知器官, 人类视觉系统特性的研究对于彩色图像处理是非常重要的。数字图像处理领域是建立在数学和概率统计表示法的基础上的, 但相比之下人类视觉在技术选择上起到核心作用。

1.3.1 人眼构造和工作原理

人眼作为一种较理想的图像传感器, 是人体接收外界信息最多的信息信道。人眼径向横截面如图 1.4 所示。

眼睛的形状近似于一个圆球, 其平均直径大约为 20mm 。有多层薄膜包围着眼睛, 即角膜、巩膜、脉络膜、虹膜和视网膜。角膜是一种硬而透明的组织, 它覆盖着眼睛的前表面。与角膜相连的巩膜是一层包围着眼球剩余部分的不透明膜。脉络膜有助于减少进入眼内的外来光和眼球内反向散射光的数量, 在脉络膜的最前面分为睫状体和虹膜。虹膜的收缩和扩张控制着进入眼睛的光量, 虹膜中间开口处(瞳孔)的直径是可变的, 大约在 $2\sim 8\text{mm}$ 的范围。当眼球适当地聚焦时, 来自眼睛外部的光在视网膜上成像。肌肉控制眼球转动, 直到感兴趣的物体图像落到中央凹上。在眼睛的后部是视网膜, 布满了整个后部的内壁, 这里是

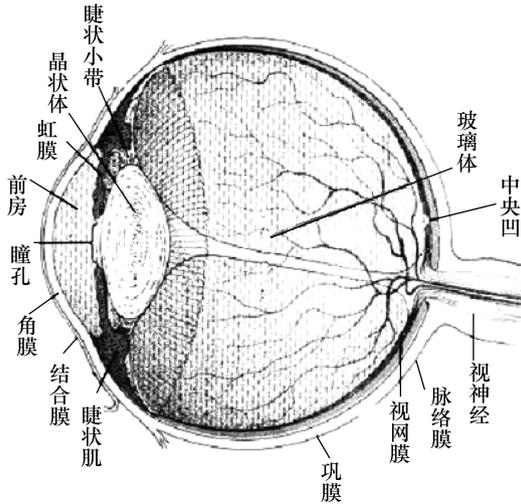


图 1.4 人眼横截面简图

眼睛的光线接受器。这些接受器由两类不同的细胞组成，分别被称为杆状体和锥状体，它们都含有被称为视紫质的色素将进入的光转变为神经脉冲。约有 7500 万~15 000 万个杆状体分布在视网膜表面，由于分布面积较大而且几个杆状体连接到一个神经末端，因此，减少了这些接受器感知细节的数量。杆状体用来给出视野内一般的总体图像。它们没有彩色感觉，而在低照明度下对图像较敏感，主要在暗光下使用以提供黑白视觉。在黑夜或其他低照明的时候世界看上去是黑白的，这是因为这些时候的视觉几乎只是由杆状体所提供的。在人的眼睛中大约有 600 万~700 万的锥状体，它们主要位于视网膜的中间部分，称之为中央凹，且对颜色灵敏度很高。用这些锥状体，人们可以充分地分辨图像细节，因为每一个锥状体都连接到自身的神经末端，在亮光下发挥作用。中央凹是视网膜中间部分的一个小底槽，锥状体几乎集中在视网膜中心中央凹的中心区，而杆状体存在于非中央凹的区域。有三种不同类型的锥状体，它们分别对不同的视觉光谱部分产生反应。正是这些锥状体使我们感知颜色。锥状体与杆状体细胞特性归纳为表 1.1 所示。

表 1.1 锥状体与杆状体细胞特性

	锥状体细胞	杆状体细胞
分布位置	视网膜中央凹区域	视网膜非中央凹区域
细胞数量	约 600 万~700 万	约 7500 万~15 000 万
颜色辨别	可细分为三种类型，分别对红、绿、蓝光敏感	只对亮度敏感
光暗辨别	只对中至强光敏感	能感受微弱光线变化
动作辨别	不敏锐	敏锐

1.3.2 人眼视觉模型

1) 调制传递函数 (MTF)

人眼是一种可见光探测器,其输入为光的辐射量,输出则为用光学量表示的光感觉,即视觉系统受刺激的程度。可以将人的眼睛看成是一个精密的光学成像系统,可以使用光学系统的概念来解释某些视觉特性。该系统可以被视为二维线性系统,其特征可以用二维光学传递函数来描述,代表系统的空间频率响应。

一个线性光学系统可以考虑为如图 1.5 所示的线性光学系统。假设该系统输入的图像是一幅亮度按空间正弦规律分布的图案,由于系统是线性的,所以输出的图像也应是一幅按同样规律分布的图案。但是,实际上两者的分布在亮度大小及空间相位上都可能会不同。若保持输入正弦规律亮度大小不变而改变其频率,则可以得到一系列对应的输出亮度与相位,从而测得光学系统的传递函数。

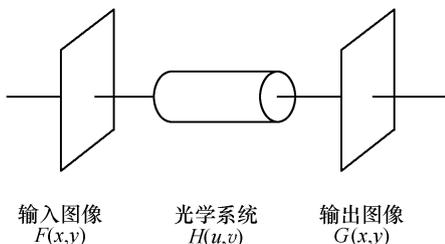


图 1.5 线性光学系统

该系统为位移不变的线性系统,其传递函数 $H(u,v)$ 、 $F(u,v)$ 和 $G(u,v)$ 分别代表系统的输入与输出图像的傅里叶变换,那么输出与输入图像的关系为

$$G(u,v) = H(u,v)F(u,v) \quad (1.4)$$

式中, u 、 v 分别代表 x 、 y 轴的空间频率分量。

若 $F(u,v)$ 和 $G(u,v)$ 已知,则光学传递函数 $H(u,v)$ 为

$$H(u,v) = \frac{G(u,v)}{F(u,v)} \quad (1.5)$$

在许多场合,只对输入和输出图像空间频率的绝对值大小的比值感兴趣,即

$$|H(u,v)| = \frac{|G(u,v)|}{|F(u,v)|} \quad (1.6)$$

称此比值为光学系统的调制传递函数(modulation transfer function, MTF)。

$H(u,v)$ 可表征线性位移不变光学系统的作用。测试人的视觉系统的调制传递函数的实验是测量人眼分辨力的一种方法。通过实验可以发现,视觉响应对较低与较高的空间频率来说,灵敏度均较差;对于空间中频而言,则灵敏度较高。同时也发现人的视觉系统是各向异性的。

2) 对数空间响应模型

从人的视觉感受到的效果上来看,人眼适应亮度的范围很大,在两幅图像中目标相对于背景的亮度的跳跃经过对数运算后是相等的。对于亮度恒定现象来说,有如下的等式成立:

$$\ln I_1 - \ln \Delta I_1 = \ln I_2 - \ln \Delta I_2 \quad (1.7)$$

另一方面，从视觉系统调制传递函数形状测试结果来看，视觉的空间频率响应具有低通特性。这是由于人眼晶状体透镜的光学孔径是有限大小的，而且具有一定的像差；并且锥状体和杆状体也都有一定的尺寸，因而人眼不能分辨非常细小的物体，所以视觉系统在空间频率上是有限带宽的。另外，从生理学上看，视神经传递光信息时具有一定的延时和残留现象，即当光闪现时，视觉系统要滞后一定时间才感觉到。而当闪光结束后，视觉对光的感受仍能维持一定的时间。这也是低通滤波器所具有的典型响应特征。

调制传递函数是线性系统的一个参量，但视觉系统对光强度的响应是非线性的，因此要线性化，才能求得真正的调制传递函数。假设视觉系统是圆对称的，则调制传递函数的一种可能的数学表达式为

$$|H(u,v)|^2 = (1 + 0.05 D^2) \exp\left[-\left(\frac{D}{50}\right)^2\right] \quad (1.8)$$

式中， $D^2 = u^2 + v^2$ 。



图 1.6 视觉系统模型

综上所述，视觉系统模型示于图 1.6 中，实验已经证明这种模型能在很宽亮度范围内作出视觉响应的合理预测。但是在高空间频率下，以及在很弱或很强的亮度下，所观察到的响应与模型预测的数值不符。尽管如此，这种模型由于其简单性，对大多数应用来说是适合的。

1.3.3 视觉现象

1) 亮度适应力

人在从亮处进入暗处时，视力首先会很差，但会慢慢地变得敏锐，可以使人在暗环境下看到物体，这一过程被称为暗适应。相反的过程，当人从暗处进入亮处也是一样，这一过程被称为亮适应。人类视觉系统可以在一个比较宽的亮度变化区间内，自动调节对亮度的适应过程，保持一定的灵敏状态正常工作。人眼对所感光强度的感觉范围称为视觉的动态范围约在 $10^{-2} \sim 10^6 \text{ cd/m}^2$ ，这样大的动态范围是许多光敏传感器所不及的。实验表明，人的视觉系统感觉到的亮度即主观亮度是进入眼内光亮度的对数函数。如图 1.7 所示，图中曲线

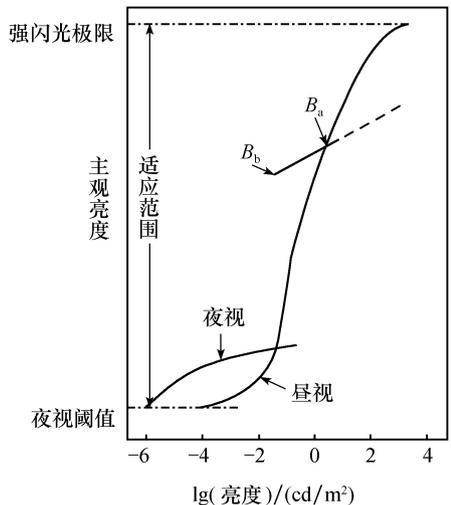


图 1.7 人的视觉系统对光亮度的适应力

表示出了光亮度与主观亮度之间的关系，即人眼的视觉动态范围。图中箭头所示为人的视觉系统所能适应的光亮度范围，即夜视与昼视动态范围。在这个动态范围内，人眼并不能同时工作，它是利用改变其整个灵敏度来完成如此大的变化的，这就是所谓的亮度适应现象。图中亮度 B_a 处的短交叉曲线表示当眼睛适应这一亮度级时，人眼所能感觉到的主观亮度范围，但这一范围是有限制的，在 B_b 处和 B_b 以下，眼睛响应为黑，曲线上部虚线部分无限制。在某一个亮度背景下，人眼感光动态范围却并不大，例如，在背景亮度约为 1 cd/m^2 条件下，人眼仅有 $10^{-2} \sim 10^1 \text{ cd/m}^2$ 的动态范围；环境背景亮度为 10^2 cd/m^2 时，其动态范围总体偏移到从 $10^1 \sim 10^3 \text{ cd/m}^2$ 。人眼总的动态范围相对大，但这个动态范围不是一成不变的，而是能够适应外界光的强度而自动调节的。

2) 对比灵敏度

眼睛对亮度变化的影响是非线性的。设有一个强度为 $I + \Delta I$ 的光斑，其周围的背景亮度为 I ，如图 1.8 所示。在不同的亮度范围，人眼所能感受到的亮度差也发生变化。实验表明，人眼睛刚能分辨的亮度差 ΔI 是 I 的函数，而且韦伯分数 (Weber ratio) $\Delta I/I$ 在相当宽的亮度范围内近似为常数，约等于 0.02。但在照度很高或很低时，所得结果与此数不同，如图 1.8 (a) 中曲线所示，从图中可以看出人眼系统有比人工图像系统更宽的动态范围。另外在图 1.8 (b) 中，

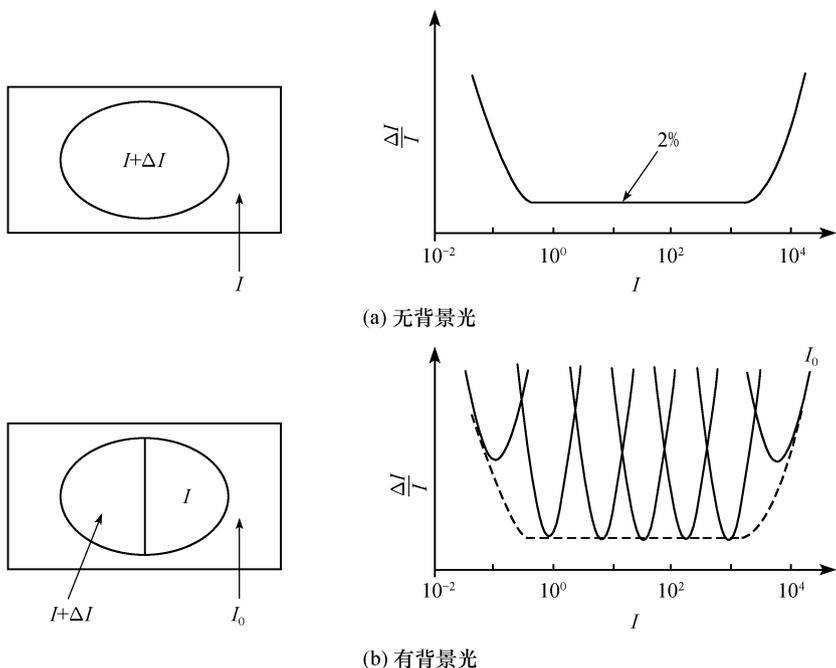


图 1.8 对比灵敏度测定

有两个相邻的光斑，一个强度为 I ，另一个强度为 $I + \Delta I$ ，而它们的背景强度为 I_0 ，实验结果表明，韦伯分数 $\Delta I/I$ 为背景强度的函数。从图 1.8 (b) 的韦伯分数与强度 I 的关系曲线可以看出，背景强度 I_0 的存在和变化，使得韦伯分数为常数的强度范围减小了很多。但是曲线簇的包络线（图中的虚线）相当于图 1.8 (a) 中的关系曲线。尽管如此，当背景光强度为 I_0 时， $\Delta I/I$ 近似为常数的强度变化范围，仍然高于多数电子成像系统的动态范围。

3) 亮度感觉

由于对数强度的微分为

$$d(\ln I) = \frac{dI}{I} \quad (1.9)$$

即在韦伯分数为常数的强度范围内，相等的对数强度变化，有相等的刚刚可辨认的强度变化。如图 1.9 所示，图 1.9 (a) 中 $\Delta I_1 = 50$ ， $I_1 = 100$ ；图 1.9 (b) 中 $\Delta I_2 = 100$ ， $I_2 = 200$ 。 $\Delta I_2 > \Delta I_1$ ，而 $I_2 > I_1$ ，但人眼的亮度感觉却是一样的，这说明人眼对亮度的感觉，并非单独与 ΔI 和 I 有关，而且还与亮度的对数成线性关系。由式 (1.9) 可知：只要两种情况的 $dI/I \approx \Delta I/I$ 相等，亮度感觉是一样的，称这种现象为亮度恒定现象。因此，在许多图像处理系统中，不是运算像素点的强度，而是强度的对数。

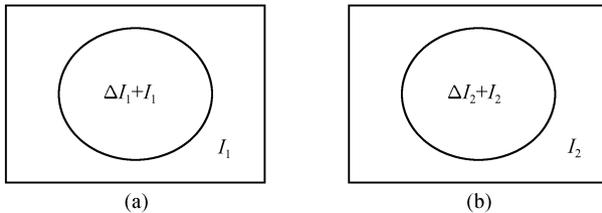


图 1.9 人眼亮度感觉

1.4 颜色基本特性

颜色是不同波长的可见光辐射作用于人的视觉器官后所产生的心理感受，对人眼颜色视觉进行定性和定量研究是彩色图像处理的基础。这些研究包括人眼颜色视觉的生理和心理现象、颜色的合成分解以及相应理论的建立。

1.4.1 颜色属性

颜色具有三种属性：色调、饱和度和亮度，它们共同引起视觉的总效果。灰度图像只有亮度特征，而彩色图像还具有色调和饱和度两个颜色特征。

色调 (hue) 反映颜色的类别，是区分不同颜色的特征属性。色调是以光波长为基础的，是以到达人眼的各波长分量综合形成的。人感觉到的颜色光可有不可

同组合，不是唯一的，这是人颜色视觉的一个特点。可见光谱范围内，光源辐射波长不同，在视觉上呈现不同色调，如表 1.2 所示。

表 1.2 颜色波长范围

颜色	紫	蓝	青	绿	黄绿	黄	橙	红
波长范围/nm	380~430	430~470	470~500	500~530	530~560	560~590	590~620	620~780

饱和度 (saturation) 表示颜色接近光谱色的程度，反映颜色的纯度。任意一种颜色都可以看作是某种光谱色与白色混合的结果。光谱色所占的比例越大，颜色接近光谱色的程度就越高，颜色的饱和度就越高。

亮度 (lightness) 指描述颜色是亮或暗的一种属性，表示颜色明暗的程度。亮度是彩色光对人眼引起的光刺激强度，是一种光的强度测量方法，与光的能量有关。如果两种颜色有相同的相关光谱组成，那么它们的不同仅仅在于它们亮度上的区别。

1.4.2 颜色的定量描述

颜色是人类视觉系统对光波刺激的反应，但它不是光波本身的物理性质，而是由人眼的生理结构和心理活动所产生的。一方面，任何视觉正常的人，其眼睛的生理结构和颜色的心理活动是基本一致的，所以人们对颜色的感知具有确定性的一面，人眼的颜色视觉特性可以定量描述出来；另一方面，不同的人，眼睛的生理结构和颜色的心理活动都有或多或少的差别，所以对颜色的定量描述应尽量避免主观因素的影响，应以大量统计实验为基础。

颜色光学的大量实验表明，人类对色光刺激的反应是三变数的，主要表现为：①人眼只能感受颜色的三种变化，如亮度、色调、饱和度，也可以用其他类似的三种特性表示；②每一种颜色都能用三个选定的原色按适当的比例混合而成，即三基色原理。

颜色可以由其亮度、色度和饱和度来定量描述，如图 1.10 所示。垂直方向表示亮度变化，圆周方向表示色度变化，径向表示饱和度变化。

用亮度、色调和饱和度来表达颜色具有形象、直观的优点，但主观性强。常用的颜色表达多是根据三基色原理而构成。

三基色原理：适当选取三种基色（如红、绿、蓝），将它们按不同的比例进行合成，就可以引起不同的颜色感觉，合成彩色光的亮度由三个基色的亮度之和决定，色度由三基色分量的比例决定。三基色彼此独立，任一种基色不能由其他两种基色配出。

随着三基色选取的不同，可以构成任意多个颜色空间。国际照明委员会 (CIE) 推荐使用波长分别为 700nm (红)、546.1nm (绿)、435.8nm (蓝) 的光

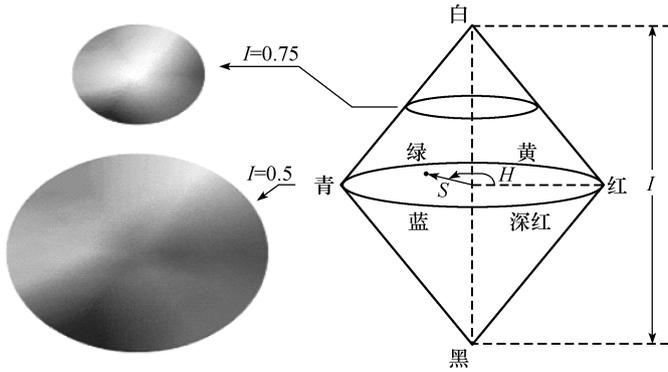


图 1.10 颜色的立体表示

谱色为三原色，并用 (R)、(G)、(B) 表示。选定不同的三基色，则该三基色系统所能展现的颜色范围也随之确定。

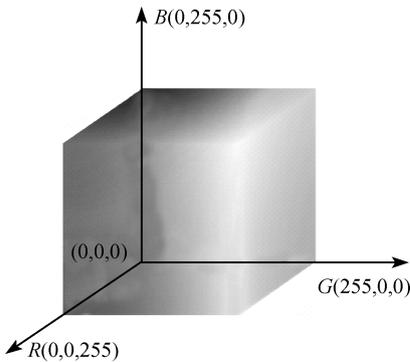


图 1.11 使用三基色表示颜色

选定三基色 R、G、B，如果一种颜色 S 可以由 r 份 R、g 份 G、b 份 B 构成，则点 (r, g, b) 可唯一确定该颜色，三基色 R、G、B 构成了一个颜色空间，用一个立方体来表示，如图 1.11 所示。

红色、绿色和蓝色组成了三个轴。在三基色 RGB 彩色立方体上，每一种颜色都有其几何坐标。黑色位于立方体后方下部。(0, 0, 0) 表示完全没有光线，或者是加色色彩模型中的黑色。立方体前方的顶部是白色，在这一点上，所有色彩的光线都集中在此。

光线增加可以生成新的色彩，光线越多，色彩就越呈浅色。白色的光线是由红色、绿色和蓝色组成的。彩色电视正是利用这一原理，用三基色分量值来传输和显示丰富多彩的颜色。

1.4.3 混色原理

任何一种颜色都可以通过混合所选择的适当的主颜色来形成，任何的主颜色都可以被使用，条件是没有其中的一种颜色可由另外两种颜色的混合来形成。混色原理包括加混色和减混色两种方式，这里主要论述加混色原理。

几种色光按不同比例混合满足加混色原理。设色光 $S_1 = (r_1, g_1, b_1)$ 与色光 $S_2 = (r_2, g_2, b_2)$ 混合，将产生新的颜色 $S = (r_1 + r_2, g_1 + g_2, b_1 + b_2)$ 。混合色的亮度由原始色光的亮度之和决定，色度由原始色光的比例决定。如图 1.12 所示，

红、绿、蓝三种颜色的光线叠加在一起，就形成了白色的光线。实践中，我们所见到的都是混合光。一定成分的混合光，有一种确定的颜色与之对应；但反过来，一种颜色并不只对应一种光谱。也就是说，两种光谱成分完全不同的混合光，有可能引起的颜色感觉完全一样。例如，6W的540nm（绿光）与25W的650nm（红光）的单色光相混合，产生的颜色与10W的580nm（黄光）的单色光相同。这种现象在色度学中称为“同色异谱”。

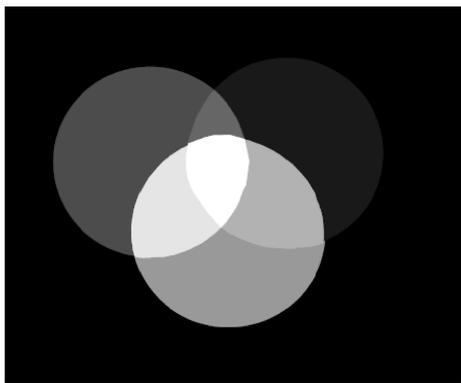


图 1.12 加色三原色

格拉斯曼定律说明了视觉对颜色的反应取决于红、绿、蓝三输入量的代数和这一事实，格拉斯曼定律包括如下内容：

- (1) 所有颜色都可以用互相独立的三基色混合得到。
- (2) 假如三基色的混合比相等，则色调和饱和度也相等。
- (3) 任意两种颜色相混合产生的新颜色与采用三基色分别合成这两种颜色的各自成分混合起来得到的结果相同。
- (4) 混合色的光亮度是原来各分量光亮度的总和。

1.4.4 亮度方程

辐通量相同而波长不同的色光，对人眼的刺激程度——光通量不同。因此，由三基色构成的颜色，每种颜色对亮度的贡献不同。由三基色分量计算该颜色亮度的公式称为亮度方程，在提取彩色图像的灰度特征时经常用到。以下给出三个最常用的方程：

PAL 电视制式亮度方程：

$$L = 0.222r + 0.707g + 0.071b \quad (1.10)$$

NTSC 电视制式亮度方程：

$$L = 0.299r + 0.587g + 0.114b \quad (1.11)$$

CIE 推荐三基色相对亮度方程：

$$L = r + 4.5907g + 0.0601b \quad (1.12)$$

1.5 颜色视觉原理

由于颜色具有物理和心理的双重结构，因此无论对于颜色刺激的物理性质的

客观描述还是对于颜色感知的主观描述都是极其重要的。颜色的感知特性依赖于光源的相对光谱功率分布、物体的选择吸收特性以及人眼的视觉响应特点，而颜色的刺激特性是前两者的函数。

1.5.1 视觉系统颜色感知

当眼睛接受到的光包含所有波长的可见光且其强度大致相等时，则发出光线的光源或所看到的物体是非彩色的。非彩色的光源为白光，而从物体反射或透射的非彩色光可能呈白色、黑色或不同程度的灰色。在白光光源照射下，若某物体能反射 80% 以上的入射光，则看上去是白色的。若反射光小于 30%，则物体看上去是黑色的。中间的反射率形成了各种深浅不同的灰色。

如果眼睛感受到的光信号中各波长的光占任意比例且均不相同，则形成彩色光。若光线集中在可见光谱的上端，即光信号主波长位于可见光谱的红色区段内，则所产生的光将呈红色。若光线集中在可见光谱的下端，即光信号主波长位于可见光谱的蓝色区段内，则形成蓝色光。

在眼睛的中央部位有三种类型的对色敏感的锥状体细胞。其中一种类型的锥状细胞对位于可见光谱中间位置的光波敏感，这种光波经过人的视觉系统转换产生绿色感。其他两种锥状体细胞对位于可见光谱的上、下端即较长和较短波长的光敏感，它们分别识别红色和蓝色。在明视觉条件下，人眼由感红、感绿、感蓝三种不同的感光细胞对光刺激进行响应，其光谱响应特征如图 1.13 所示。图中横坐标为波长，纵坐标为响应敏感程度，响应曲线已作归一化处理。

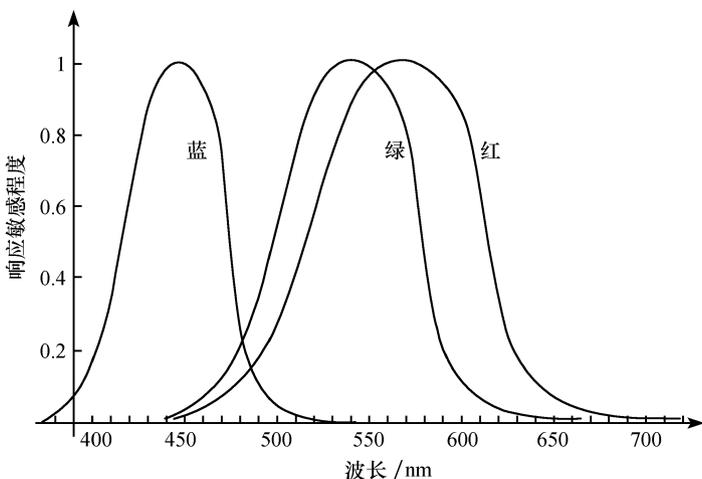


图 1.13 人眼颜色光谱敏感特性曲线

对中心波长为 λ 、微小波长间隔 $d\lambda$ 波长范围内的色光，其三刺激值 $dR(\lambda)$ 、

$dG(\lambda)$ 、 $dB(\lambda)$ 和颜色刺激 $\varphi(\lambda)$ 、三种感光细胞的敏感特性 $V_r(\lambda)$ 、 $V_g(\lambda)$ 、 $V_b(\lambda)$ ，以及波长间隔 $d\lambda$ 成比例，可得以下关系：

$$\begin{aligned} dR(\lambda) &= k\varphi(\lambda)V_r(\lambda)d\lambda \\ dG(\lambda) &= k\varphi(\lambda)V_g(\lambda)d\lambda \\ dB(\lambda) &= k\varphi(\lambda)V_b(\lambda)d\lambda \end{aligned} \quad (1.13)$$

对于整个可见光谱范围，积分式 (1.13) 得

$$\begin{aligned} R &= \int_{\lambda} k\varphi(\lambda)V_r(\lambda)d\lambda \\ G &= \int_{\lambda} k\varphi(\lambda)V_g(\lambda)d\lambda \\ B &= \int_{\lambda} k\varphi(\lambda)V_b(\lambda)d\lambda \end{aligned} \quad (1.14)$$

R 、 G 、 B 值即构成了对该混合色的描述，由人的视觉系统解释为人所熟悉的颜色。对于光源色，颜色刺激函数 $\varphi(\lambda) = S(\lambda)$ ，其中 $S(\lambda)$ 为光源的光谱功率分布。对于透射物体色，颜色刺激函数 $\varphi(\lambda) = S(\lambda)\tau(\lambda)$ ，其中 $\tau(\lambda)$ 为物体的光谱透射率。对于漫反射物体，颜色刺激函数 $\varphi(\lambda) = S(\lambda)\beta(\lambda)$ 或 $\varphi(\lambda) = S(\lambda)\rho(\lambda)$ 。 $\beta(\lambda)$ 为光谱反射率因数或光谱辐亮度因数； $\rho(\lambda)$ 为物体的反射率。对于后两种情况，当已知光源光谱分布后，可以推导出实际的物面光学参数，即图像的色度或饱和度可以不受照明条件的影响。实际上，人眼判断景物色度时，能够在较大程度上不受光源的影响，这一特性称为颜色视觉不变性。

若这三种锥状体细胞都感受到相同水平的辐射（单位时间内的能量），则眼睛看到的是白光。当然自然界的白光包含可见光谱中的所有波长的光的辐射。然而，从生理学的角度看，由于眼睛仅包含三种不同类型的锥状细胞，因而对任意三种颜色适当混合就可以产生白光视觉，条件是这三种颜色中任意两种的组合都不能生成第三种颜色。

1.5.2 颜色视觉模型

颜色刺激是视网膜上三种锥状细胞分别感受不同波长的光刺激后由大脑合成的一种主观知觉，是波长的函数，同时又受到人眼生理结构及感知颜色的心理活动等主观因素的影响。颜色视觉模型是一个研究中的问题，三基色合成颜色的原理是基于实验得到的。目前有许多学说企图解释这个实验，因而人视觉可用三通道的模型来建模。这种假设仍然有待于生理学、解剖学的严格证明。但由于实验的成功，这种建模方法必然和实验符合。设 $C(\lambda)$ 为照射光源的波谱能量分布， $v_i(\lambda)$ 为第 i 个锥状体视敏度响应，则三种锥状体产生的信号表示为

$$e_i = \int C(\lambda)v_i(\lambda)d\lambda \quad i = 1, 2, 3 \quad (1.15)$$

由 Frei 提出的彩色视觉模型，如图 1.14 所示。

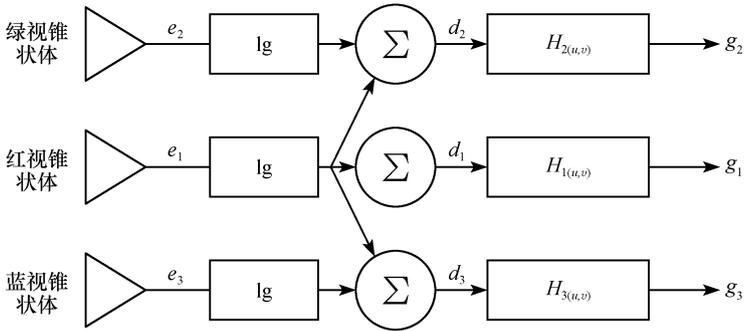


图 1.14 颜色视觉模型

图 1.14 中 e_1 、 e_2 、 e_3 表示神经膜的色素吸收作用后三个感光锥状体所产生的信号。图中的加权系数 d_i 即线性组合系数：

$$d_1 = \lg e_1$$

$$d_2 = \lg e_2 - \lg e_1 = \lg \frac{e_2}{e_1} \tag{1.16}$$

$$d_3 = \lg e_3 - \lg e_1 = \lg \frac{e_3}{e_1}$$

图中 $H_1(u,v)$ 、 $H_2(u,v)$ 、 $H_3(u,v)$ 为三个通道的转移函数，最后产生输出信号 g_i ($i=1, 2, 3$)。这个模型对彩色图像的分解、合成均有意义。

1.6 颜色空间分析

1.6.1 面向设备的颜色空间

面向设备的颜色空间适合在输出显示场合下使用。计算机彩色输入输出设备使用与设备相关的颜色空间，如 CRT 显示器和扫描仪使用 RGB 颜色空间，打印机使用 CMYK 颜色空间，但各种外部设备使用的颜色空间并不一致。因此，除了分析研究它们各自所在的颜色空间外，还必须清楚它们之间的相互转换关系，这在计算机系统彩色管理工作中十分重要。

1) RGB 颜色空间

光谱中最重要的颜色是红 (R)、绿 (G)、蓝 (B) 三基色，所有颜色都可以由三基色相加而产生。RGB 颜色空间是一个立方体三维坐标空间结构，分别用红、绿、蓝表示三个坐标轴，如图 1.15 所示。立方体的底部 $R=G=B=0$ 处为黑色，顶部与其相对角 $R=G=B=255$ 处为白色。由于图像采集和显示设备

使用的是 RGB 颜色空间，所以，RGB 颜色空间是彩色图像处理中最基础、最常用的颜色空间。

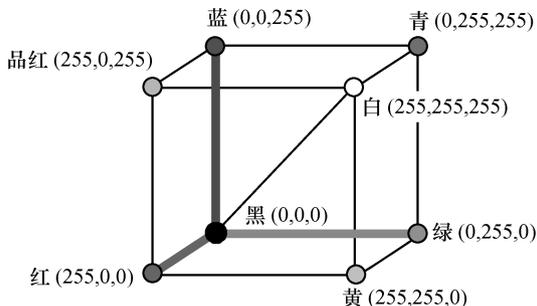


图 1.15 RGB 颜色空间

RGB 颜色空间的主要缺点：①不直观，从 RGB 值中很难看出其所表示的颜色的认知属性；②不均匀，两个色点之间的距离不等于两个颜色之间的知觉差异；③对硬件设备具有依赖性。因此，RGB 颜色空间是一个与设备相关的、颜色描述不完全直观的空间。

为了克服 RGB 颜色空间的不均匀和不直观的缺点，在彩色图像处理中大多采用更加符合颜色视觉特性的颜色空间。RGB 颜色空间能被转变成所需要的其他任何颜色空间。由于任何三个基色都能表示一种颜色，根据三基色理论的实现要求，就可以使用不同颜色空间的不同基色来表达同一种颜色。颜色空间变换提供了一种三基色颜色空间向另一种三基色颜色空间的映射方法，实现从一组原色向另一组原色转换，这是由于任何原色刺激都可以由其他组原色刺激的混合来生成。例如，从 RGB 到 $R'G'B'$ ，变换可以写成以下形式：

$$R' = a_{11}R + a_{12}G + a_{13}B \quad (1.17)$$

$$G' = a_{21}R + a_{22}G + a_{23}B \quad (1.18)$$

$$B' = a_{31}R + a_{32}G + a_{33}B \quad (1.19)$$

或者写为

$$P' = TP \quad (1.20)$$

式中， T 为变换矩阵， P 和 P' 是不同的颜色空间。

$$T = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad (1.21)$$

2) CMY (CMYK) 颜色空间

CMY 颜色空间是一个矩形直角空间结构，其三基色分别为深蓝 (C)、品红

(M) 和黄 (Y)。它利用减色法获得各种颜色，常用于非发射显示，如喷墨式打印机等。等量的颜料三基色可以产生黑色，但实际上这样产生的黑色是不纯的。通常为了产生真正的黑色，加入了第四种颜色——黑色 (K)，提出了 CMYK 颜色空间。CMY 颜色空间不是很直观的，而且是非线性感知的。CMY 颜色空间的三基色与 RGB 颜色空间的三基色互为互补色。虽然这两个三基色集合产生的空间并不完全真正对应，但有简单的近似公式可以相互变换，从 RGB 颜色空间到 CMY 颜色空间的变换公式如下：

$$\begin{aligned} C &= 1 - R \\ M &= 1 - G \\ Y &= 1 - B \end{aligned} \quad (1.22)$$

从 CMY 到 CMYK 空间的变换公式如下：

$$\begin{aligned} K &= \min(C, M, Y) \\ C &= \frac{C - K}{1 - K} \\ M &= \frac{M - K}{1 - K} \\ Y &= \frac{Y - K}{1 - K} \end{aligned} \quad (1.23)$$

3) YIQ 与 YUV 颜色空间

YIQ 和 YUV 颜色空间都产生一种亮度分量信号和两种色度分量信号，而每一种变换使用的参数都是为了适应某种类型的显示设备。其中，YIQ 适用于 NTSC 彩色电视制式，YUV 适用于 PAL 和 SECAM 彩色电视制式。

在彩色电视制式中，使用 YIQ 和 YUV 模型来表示彩色图像。YIQ 颜色空间是美国国家电视标准委员会 (NTSC) 定义的电视广播标准。在 YIQ 颜色空间中，Y 为亮度信息，I、Q 为色度值，是两个彩色分量。其中，I 为橙色分量，Q 为品红色分量，各分量近似正交。正交颜色空间主要应用在彩色电视系统，可以减少电视信号传输的带宽，也是彩色图像编码和压缩的基础。欧洲定义了相交替制式 (PAL)，使用 YUV 格式。在这种颜色空间中，Y 为亮度信息，U、V 为色差信号，U、V 是构成彩色的两个分量。YUV 颜色空间与 YIQ 颜色空间类似，差别仅在于空间多了一个 33° 的旋转。目的是为了可以与黑白电视兼容。电视信号在发射时，转换成 YUV 形式，接收时再还原成 RGB 三基色信号，由显像管显示。YUV 表示法中的亮度信号 (Y) 和色度信号 (U、V) 是相互独立的，也就是 Y 信号分量构成的黑白灰度图与用 U、V 信号构成的另外两幅单色图是相互独立的。此外，黑白电视能接收彩色电视信号也就是利用了 YUV 分量之间的独立性。YUV 表示法的另一个优点是可以利用人眼的特性来降低数字彩色

图像所需要的存储容量。

(1) YUV 与 RGB 的空间变换。

在考虑人的视觉系统和阴极射线管 (CRT) 的非线性特性之后, RGB 和 YUV 的对应关系可以近似地用下面的方程式表示:

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ U &= -0.147R - 0.289G + 0.436B \\ V &= 0.615R - 0.515G - 0.100B \end{aligned} \quad (1.24)$$

(2) YIQ 与 RGB 的空间变换。

RGB 和 YIQ 的对应关系可用下面的方程式表示:

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ I &= 0.596R - 0.275G - 0.321B \\ Q &= 0.212R - 0.523G + 0.311B \end{aligned} \quad (1.25)$$

1.6.2 面向视觉感知的颜色空间

面向设备的颜色空间与人眼的视觉感知有一定的距离。例如, 给定一个颜色, 人很难判定其中的 RGB 各个分量, 这时使用面向视觉感知的颜色空间比较方便。

在面向视觉感知的颜色空间中, 人们通过将人类对颜色的观察情况组合而建立的孟塞尔 (Munsell) 颜色空间是一个基本的颜色空间, 它主要基于主观观察, 用于艺术领域。在此基础上, 人们又提出了许多对孟塞尔系统近似的比较客观的颜色空间, 如 HSI、HSB 等。这些颜色空间与人类关于颜色的视觉感知比较接近, 同时独立于设备, 是非线性的。

1) 由色卡系列构成的颜色空间

主要有孟塞尔颜色空间以及奥斯特瓦尔德 (Ostwalds) 颜色空间, 是经过大量实验得到的一套实验数据或比色卡, 不可能直接应用于图像的表达和处理, 但可以作为判断其他颜色空间独立性和均匀性的标准。孟塞尔和奥斯特瓦尔德颜色空间, 都是色调、饱和度和亮度的三维表达, 它们都有沿着圆周的光谱色调。沿着垂直中线到颜色环平面亮度发生变化, 主要的亮度轴被称为非彩色轴, 沿着这个轴, 颜色从白到灰到黑变化。饱和度被表示为与非彩色轴之间的距离。其中孟塞尔颜色空间是目前使用的最重要的表色系统之一, 美国和日本的颜色标准都以它为基础, 并得到了世界公认。

孟塞尔颜色空间于 1905 年被提出, 以后, 又经过几百万人次的实验和修正, 其明度值, 以及色调、彩色的编排均接近视觉上等距的原则, 因此常常作为判断其他颜色空间均匀性和独立性的标准。孟塞尔颜色空间包括 1500 种系统排列的颜色样本。这是一种均匀颜色空间的三维表示, 一个色调环被分为 5 个主要色

调,即红、黄、绿、蓝和紫。这些颜色又被进一步地分为 5 个中间色调,即红黄、黄绿、绿蓝、蓝紫和红紫,而这些被分出来的颜色中的每一个又再被划分为 10 种颜色,从而得到 100 种分离的色调。孟塞尔颜色空间的一个重要的属性是饱和度和亮度的数量级别是按着相等的知觉间隔来确定的。在三维空间中,孟塞尔颜色空间为一立体,如图 1.16 (a) 所示。它的奇怪外型是由于饱和度和亮度是随着相等的知觉间隔而变化的。

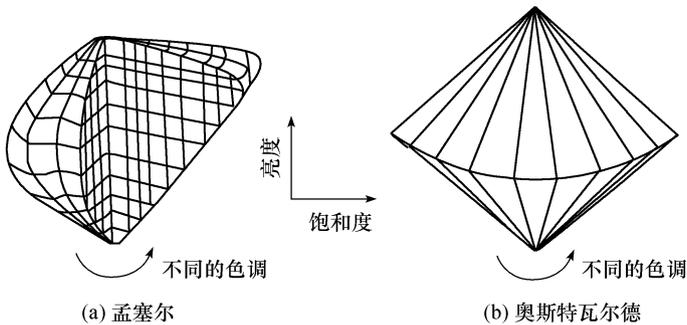


图 1.16 孟塞尔和奥斯特瓦尔德颜色空间

奥斯特瓦尔德颜色系统与孟塞尔颜色系统是相似的,但在奥斯特瓦尔德颜色系统中,定义每一种色调都是按着固定比例数量混合得到的,8 个相等空间分布的颜色从白到黑变化。如图 1.16 (b) 所示,图中的这种规则外型是由于奥斯特瓦尔德颜色空间是不依赖于相等的感知差别的。

2) XYZ 颜色空间

由于用 RGB 颜色空间比配等能光谱时存在负比配,为了用三基色定义出所有的颜色,国际照明委员会 CIE 定义了三种标准基色 X、Y、Z,这三种基色是虚拟的,使颜色比配全部为正值,称为 XYZ 颜色空间,是一种设备独立的颜色空间。在这个空间内,监视器的颜色范围成为一个变形的六面体,它与 RGB 颜色空间之间的关系是线性关系。其形状大约为一个顶点在坐标原点的圆锥体,锥体的外表面可以看成由一些从原点出发而终止于一条边的光线组成,这些光线可以看成是某些色彩的各种颜色的集合,而 Y 值则表示某个颜色的亮度。

XYZ 颜色空间包含了所有人类能够发觉的颜色,而且它是基于由实验测定的颜色匹配函数的,因此它不同于 RGB 颜色空间只是表示监视器所能显示的颜色范围,而是显示所有的颜色。在 XYZ 颜色空间内,可以通过改变三个分量的数值来得到所需要的颜色。

XYZ 颜色空间的一个重要属性是它的设备非依赖性,每一个颜色空间有一个来自 CIE XYZ 颜色空间转变量。这个 XYZ 颜色空间通常作为一个参考颜色空