



普通高等教育“十一五”国家级规划教材

高等院校电子、信息类教材

数字图像处理及模式识别

Digital Image Processing and Pattern Recognition

(第二版)

沈庭芝 王卫江 闫雪梅 编著

 北京理工大学出版社
BEIJING INSTITUTE OF TECHNOLOGY PRESS

责任编辑：王艳丽
封面设计：庚辰年代

内 容 简 介

本书是高等院校本科生、研究生数字图像处理及模式识别技术的教材。内容包括图像的变换、增强、复原、编码、分析与识别等，还增加了现代先进图像处理技术，如三维可视化、盲信号处理、随机场模拟、遗传及进化算法、智能模式识别等。每章配有习题。

定价：26.00 元

7-81045-403-2
数字图像处理及模式识别（第二版）



781045403260

RMB:26.00

2007

TN919.8/4=2

2007

普通高等教育“十一五”国家级规划教材
高等院校电子、信息类教材

数字图像处理及模式识别

(第二版)

沈庭芝 王卫江 闫雪梅 编著

 北京理工大学出版社

BEIJING INSTITUTE OF TECHNOLOGY PRESS

图书在版编目 (CIP) 数据

数字图像处理及模式识别/沈庭芝, 王卫江, 闫雪梅编著. —2 版.

—北京: 北京理工大学出版社, 2007. 7

普通高等教育“十一五”国家级规划教材. 高等院校电子、信息类教材

ISBN 978-7-81045-403-2

I. 数… II. ①沈…②王…③闫… III. ①数字图像处理-高等学校-教材②数字图像-模式识别-高等学校-教材 IV. TN911.73

中国版本图书馆 CIP 数据核字 (2007) 第 109548 号

出版发行 / 北京理工大学出版社

社 址 / 北京市海淀区中关村南大街 5 号

邮 编 / 100081

电 话 / (010)68914775(办公室) 68944990(批销中心) 68911084(读者服务部)

网 址 / <http://www.bitpress.com.cn>

经 销 / 全国各地新华书店

印 刷 / 北京圣瑞伦印刷厂

开 本 / 787 毫米 × 1092 毫米 1/16

印 张 / 17.75

字 数 / 411 千字

版 次 / 2007 年 7 月第 2 版 2007 年 7 月第 5 次印刷

印 数 / 11001 ~ 16000 册

定 价 / 26.00 元

责任校对 / 张 宏

责任印制 / 吴皓云

前 言

自从 20 世纪 20 年代人类第一次实现在纽约和伦敦之间通过海底电缆传送图片至今, 仅仅 80 余年, 数字图像处理技术已经得到了突飞猛进的发展。随着计算机的普遍使用, 人类已经进入了一个高速发展的信息化时代, 图像处理技术愈来愈成为科学技术各领域必不可少的手段。它的应用是如此之广, 从空间技术到显微图像, 从军事领域到工业生产, 从天文地质到医学诊断, 从电视广告到少儿游戏, 图像处理技术无处不在, 它与每个人的工作、学习和生活密切相关。

广大青年学生对图像处理技术的学习怀有浓厚的兴趣, 本书就是为高等院校的本科生、研究生编写的教材。它包含了图像处理的主要技术: 如图像的变换、增强、复原、编码、分析与识别等, 阐述了它们的主要基础理论知识和实现的途径, 提供了习题及实验框架。

反映本学科的新思想、新技术、新方法是本书编写的一大宗旨, 为此除基本内容之外, 增加了最近不断讨论的三维成像、遗传算法、随机场模拟、神经网络的应用、盲信号处理以及先进的视频编解码技术。加强理论与实践的联系也是本书编写过程中考虑的问题, 除理论分析之外, 本书还强调了实现的途径, 提供了几个主要实验的框架, 以期本书对工程技术人员有参考价值。

本书还有配套的计算机辅助教学软件, 供学生自学与自查, 软件中包括了大部分图像处理技术, 并提供了许多精美的图片。凡具有中文 Windows 3.1 以上的计算机皆可使用, 欲购买此软件者请与北京理工大学电子工程系方子文同志联系, 电话 68912612 - 802。

本书的诞生是在多位教师多年教学与科研的基础上编写的, 本次修订, 有几位年轻教师参加编写, 为本教材的编写增加了新生力量。在此特向在百忙之中为本书进行审阅并提出许多宝贵意见的著名信号处理专家柯有安教授、张宝俊教授致以崇高的敬意。对曾参与计算机辅助教学软件制作的林海、陈建军、顾建军、朱少娟等同志表示感谢。

本书第 1、第 4、第 5、第 6 章由沈庭芝编写, 第 2、第 3 章由闫雪梅编写, 第 7、第 8、第 9、第 10 章由王卫江编写。本次修订版主要作以下修改: 删减相对落后的内容, 保留基础理论, 增加在此领域最新出现的一些新理论、新方法以及应用实例。由于篇幅有限, 增加的内容不能太多, 但力求精练实用, 既有全面系统的基础知识, 又有实际可操作的步骤, 还有部分应用实例, 适合于各类学校、研究所参考使用。由于作者水平有限, 书中难免有不当之处, 恳请各位读者批评指正。

目 录

第 1 章 绪论	(1)
§ 1.1 数字图像处理技术的概况	(1)
§ 1.2 视觉的原理与模型	(3)
§ 1.3 图像的数字化的	(9)
§ 1.4 彩色图像	(11)
习题	(13)
参考文献	(13)
第 2 章 图像变换	(14)
§ 2.1 傅里叶变换	(14)
§ 2.2 离散余弦变换	(22)
§ 2.3 离散沃尔什变换/离散哈达玛变换	(24)
§ 2.4 小波变换	(30)
§ 2.5 离散 K-L 变换	(35)
习题	(37)
参考文献	(38)
第 3 章 图像增强	(39)
§ 3.1 基于点操作的增强	(39)
§ 3.2 图像平滑	(50)
§ 3.3 图像锐化	(61)
§ 3.4 彩色增强	(69)
习题	(73)
参考文献	(74)
第 4 章 图像的复原	(75)
§ 4.1 退化的数学模型	(75)
§ 4.2 图像中的噪声	(76)
§ 4.3 连续系统的图像复原	(78)
§ 4.4 离散情况下的退化模型	(80)
§ 4.5 离散情况下的复原	(83)
§ 4.6 运动模糊图像的复原	(85)
§ 4.7 非线性图像复原	(87)
§ 4.8 同态滤波复原	(91)
§ 4.9 图像的盲复原	(92)

§ 4.10 三维图像可视化处理及在医学上的应用	(94)
习题	(96)
参考文献	(97)
第 5 章 数字图像的压缩编码	(98)
§ 5.1 概述	(98)
§ 5.2 基础知识	(99)
§ 5.3 熵编码方法	(104)
§ 5.4 轮廓编码	(108)
§ 5.5 变换编码与小波编码	(113)
§ 5.6 分形编码	(118)
§ 5.7 图像压缩标准	(120)
习题	(127)
参考文献	(127)
第 6 章 图像随机场模拟及处理	(129)
§ 6.1 图像的随机场模型	(129)
§ 6.2 图像模拟的实现	(133)
§ 6.3 图像参数估计方法的研究	(137)
§ 6.4 遗传算法及其应用	(141)
习题	(145)
参考文献	(146)
第 7 章 图像的分析	(147)
§ 7.1 图像的特征	(147)
§ 7.2 图像的分割	(151)
§ 7.3 图像的纹理分析	(161)
§ 7.4 图像的符号描述	(164)
§ 7.5 多维信息及运动图像的分析 and 利用	(171)
习题	(172)
参考文献	(173)
第 8 章 图像的匹配	(174)
§ 8.1 模板匹配法	(174)
§ 8.2 其他快速计算法	(179)
§ 8.3 受几何失真等影响小的匹配算法	(184)
§ 8.4 几种实用的图像匹配算法	(185)
习题	(191)
参考文献	(192)
第 9 章 模式识别技术	(193)
§ 9.1 统计模式识别法	(194)
§ 9.2 遥感图像的几种监督分类方法	(202)

§ 9.3 模糊模式识别	(212)
§ 9.4 结构模式识别	(225)
§ 9.5 神经网络及其在模式识别中的应用	(236)
§ 9.6 支持向量机及其在模式识别中的应用	(245)
习题	(249)
参考文献	(251)
第 10 章 数字图像处理与模式识别技术的应用举例	(252)
§ 10.1 牌照图像的预处理	(252)
§ 10.2 基于综合特征的牌照定位技术	(254)
§ 10.3 牌照字符的切分	(257)
§ 10.4 字符识别算法	(259)
参考文献	(262)
附录 实用图像处理程序	(263)

第 1 章 绪 论

§ 1.1 数字图像处理技术的概况

1.1.1 图像处理技术的发展与应用

早在 20 世纪 20 年代,人们利用巴特兰 (Bartlane) 电缆图片传输系统,经过大西洋传送了第一幅数字图像,它使传输的时间从一个多星期减少到了三小时,使人们感受到数字图像传输的威力。它的传输方法,首先是对图像进行编码,然后在接收端用一台电报打印机利用字符模拟中间色调把图像还原出来,这是个初步尝试。为了对图像的灰度、色调和清晰度进行改善,人们采用各种方法对图像的传输、打印和复原等技术进行改进,这种努力一直延续到此后的 40 年。直到大型计算机出现后,人们才开始用计算机来改善图像。在 1964 年,美国喷射推进实验室 (JPL) 进行了太空探测工作,当时用计算机来处理测距器 7 号发回的月球图片,以校正飞船上电视摄像机中各种不同形式的固有的图像畸变,这些技术都是图像增强和复原的基础。同时他们成功地用计算机绘制出月球表面的地图。随后在 1965 年又对徘徊者 8 号发回的几万张照片进行较为复杂的数字图像处理,使图像的质量进一步提高。JPL 的工作引起了世界许多有关方面的注意, JPL 也更加重视数字图像技术的研究,投入了更大的力量,并取得了许多非凡的成果。与此同时, JPL 以及世界各国有关部门已把数字图像处理技术从空间技术开发到生物医学、X 射线图像增强、光学显微图像分析、遥感图像分析、粒子物理、地质勘探、人工智能和工业检测等方面。

数字图像处理技术在近 30 多年的时间里,迅速地发展成为一门独立的有强大生命力的学科,下面仅就几个方面的某些应用举些例子。

1. 遥感技术

遥感技术可以是飞机遥感和卫星遥感技术,从前,许多国家每天派出很多侦察飞机对地球上感兴趣的地区进行大量的空中摄影。对由此得来的照片进行判读分析需要雇用几千人,而现在改用配备有高级计算机的图像处理系统来判读分析,既节省人力,又加快速度,还可以从照片中提取人工所不能发现的大量的有用情报。由于各种原因,从遥感卫星所获得的地球资源图片图像质量总不是很好,如果仍采用简单的直观判读如此昂贵代价所获取的图像是不合算的,因此必须采用图像处理技术。如 LANDSAT 系列陆地卫星,采用多波段扫描器 (MSS),在 900 km 高空,对地球每一个地区以 18 天为一周期进行扫描成像,其图像分辨率大致相当地面上十几米或 100 m。这些图像无论在成像、存储、传输过程中,还是在判读分析中,都必须采用很多的数字图像处理方法。目前遥感技术,尤其是卫星遥感,已经在资源调查、灾害监测、农业规划、城市规划、环境保护等方面有很大的应用效果。我国也在以上诸方面的实际应用中取得了良好的成果,对我国国民经济的发展起到了相当大的作用。

2. 医用图像处理

医学上不管是基础科学还是临床应用，都是图像处理种类极多的领域，例如对生物医学的显微图像的处理分析方面，如红白细胞和细菌，染色体分析，另外像胸部 X 线照片的鉴别，眼底照片的分析以及超声波图像的分析等都是医疗辅助诊断的有力工具，目前这类应用已经发展到专用的软件和硬件设备，最普遍使用的是计算机层析成像，亦称为 CT 技术。它是由英国的 Hounsfield 和美国的 Cormack 发明的。通过 CT，可以获取人体剖面图，使得肌体病变特别是肿瘤诊断发生革命性的变化，两位发明者因此获得 1979 年诺贝尔医学奖。近年来又出现的核磁共振 CT，使人体免受各种硬射线的伤害，并且图像更为清晰。图像处理技术在医学上的应用正在进一步地发展。

3. 工业领域中的应用

在工业领域中的应用一般有以下几方面：工业产品的无损探伤，表面和外观的自动检查和识别，装配和生产线的自动化，弹性力学照片的应力分析，流体力学图片的阻力和升力分析。最值得注意的是“计算机视觉”，采用摄影和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，产品集装，部件装配，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。

4. 军事公安方面

主要应用是：各种侦察照片的判读，运动目标的图像自动跟踪技术，例如目前电视跟踪技术已经装备到导弹和军舰上，并在实践和演习中取得很好的效果。另外还有公安业务图片的判读分析，如指纹识别，不完整图片的复原等，在公安的跟踪、窃视、交通监控、事故分析中都用到了图像处理技术。

5. 文化艺术方面

在文化艺术方面有电视画面的数字编辑，动画片的制作，服装的花纹设计、制作，文物资料照片的复制和修复。在体育方面有运动员的训练，动作分析和评分等。

随着计算机技术的日益发展，图像处理技术的日益完备，图像处理的应用范围将越加深入和广泛。

1.1.2 数字图像处理技术的特点

在计算机处理出现以前，图像处理都是光学照相处理和视频信号处理等模拟处理，把数字图像与其他模拟方法在以下各方面比较：

方式	处理速度	灵活性	精度	调整	再现性	其他
光学	√	○	○	×	○	× 现象
照片	√	×	×	×	×	× 现象
录像	√	○	×	○	○	
数字	×	√	√	√	√	× 内存

从上面的列表可以看出，除了处理速度和内存要求高以外，数字图像处理技术在灵活性、精度、调整和再现性方面都是卓越的，它用程序能自由地进行各种处理，并且能达到较

高的精度。这与模拟处理中,要提高一个数量级的精度,就必须对装置进行大幅度改进相比确实为一优点。另外,由于半导体技术的不断进步,开发出普遍适用的微处理器的图像处理专用高速处理器,以 IC 存储器为基础的图像显示也达到可行的程度,这些都进一步加快了数字图像处理技术的发展和实用化。

为了用计算机处理图像,必须把图像作为数值来表示,数字图像就是二维平面上的灰度分布。数字图像信息有以下的特点。

① 信息量很大。例如一帧电视图像取 512×512 个像素组成,如其灰度级用 8 bit 的二进制制来表示,则有 $2^8 = 256$ 个灰度级,那么一帧图像的信息量即为 $512 \times 512 \times 8 = 2\,097\,152$ bit。对这样大信息量的图像进行处理,必须要有计算机才能胜任,而且计算机的内存量要大。

② 数字图像占用的频带较宽。与语言信息相比,占用的频带要大几个数量级。如电视图像的带宽为 5~6 MHz,而语言带宽仅为 4 kHz 左右。频带愈宽,技术实现的难度就大,成本亦高,为此对频带压缩技术提出了较高的要求。

③ 数字图像中各个像素是不独立的,其相关性很大。例如在电视画面中,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,而相邻两帧之间的相关性比帧内相关性还要大一些。因此图像信息压缩的潜力很大。

④ 处理后的数字图像是需要给人观察和评价的,因此受人的因素影响较大。由于人的视觉系统很复杂,受环境条件、视觉性能、人的主观意识的影响很大,因此要求系统与人的良好的配合,这还是一个很大的研究课题。

§ 1.2 视觉的原理与模型

1.2.1 视觉的原理

光线照在物体上其透射或反射光的分布就是“图”。而人的视觉系统对图的接收在大脑中形成的印象或认识就是“像”。前者是客观存在,后者为人的感觉,图像就是二者的结合。下面首先谈谈人眼的构造。

在图 1-1 中画出了人眼的横断面,眼睛的形状为一圆球,即平均直径约为 20 mm,它由三层薄膜包着,即角膜和巩膜外壳,脉络膜和视网膜。角膜是一种硬而透明的组织,它盖着眼睛的前表面,巩膜与角膜连在一起,巩膜是一层包围着眼球剩余部分的不透明的膜。脉络膜位于巩膜的里边,这层膜包含有血管网,它是眼睛的重要滋养源,脉络膜外壳着色很重,因此有助于减少进入眼内的外来光和眼球内的回射。在脉络膜的最前面被分为睫状体和虹膜。虹膜的收缩和扩张控制着允许进入眼内的光量。虹膜的中间开口处(瞳孔)的直径是可变的,约由 2 mm 变到

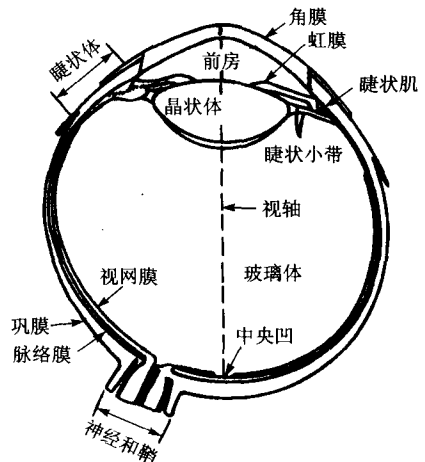


图 1-1 人眼截面的简化图

8 mm。虹膜的前部含有眼睛的明显的色素，而后部则含有黑色素。

眼睛最里层的膜是视网膜，它布满在整个后部的内壁上，当眼球被适当地聚集，从眼睛的外部物体来的光就在视网膜上成像。整个视网膜表面上分布的分离的光接收器造成了图案视觉。这种光接收器可分为两类：锥状体和杆状体。每只眼睛中锥状体的数目在 600 万 ~ 700 万之间。它们主要位于视网膜的中间部分，叫做中央凹，它对颜色很敏感，人们用这些锥状体能充分地识别图像的细节，因为每个锥状体都被接到其本身的神经的一端。控制眼睛的肌肉使眼球转动，从而使人所感兴趣的物体的像落在视网膜的中央凹上，锥状视觉又叫白昼视觉。

杆状体数目更多，约有 7 500 万 ~ 15 000 万个，分布在视网膜表面上，因为分布面积较大并且几个杆状体接到同一根神经的末端上，因而使接收器能够识别的细节的量减少了，杆状体用来给出视野中一般的总的图像。它没有色彩的感觉而对低照明度的景物则较敏感。例如，在白天呈现鲜明颜色的物体，在月光之下却没有颜色，这是因为只有杆状体受到了刺激，而杆状体没有色彩的感觉。杆状视觉又叫夜视觉。

眼睛中的晶状体与普通的光学透镜之间的主要差别在于前者的适应性强，如图 1-1 所示，晶状体前面的曲率半径大于后表面的曲率半径。晶状体的形状由睫状体的韧带的张力来控制。为了对远方的物体聚焦，控制用的肌肉就使晶状体变得较扁平。同样，为使眼睛近处的物体得到聚集，肌肉就使晶状体变得较厚。

当晶状体的折射能力由最小变到最大时，晶状体的聚集中心与视网膜之间的距离约由

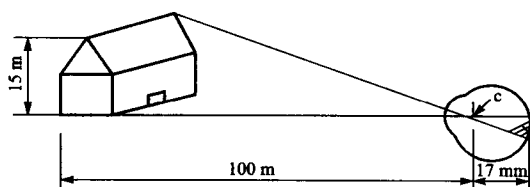


图 1-2 用眼睛观察房子的光学表示法
c 点—晶状体的光学中心

17 mm 缩小到 14 mm。当眼睛聚焦到远于 3 m 的物体时，晶状体的折射能力最弱，当聚集到非常近的物体时，其折射能力最强。利用这一数据，将易于计算出任何物体在视网膜上形成图像的大小。例如图 1-2 中，观测者看一个距离为 100 m，高为 15 m 的房子，设 x 为视网膜上形成的图像的大小，单位为 mm，由图 1-2 的几何形状来看，则有

$15/100 = x/17$ ， $x = 2.55$ mm。正如前面所指出的，视网膜图像主要反映在中央凹的面积上。然后由光接收器的相对刺激作用产生感觉。这样，感觉把辐射来的能量转变为电脉冲，最后由脑子判别出来。

由于数字图像是作为许多分离的亮点显示出来的，因此眼睛对于不同亮度之间的鉴别能力，是图像处理结果中所要考虑的一个重要方面，人的视觉系统能够适应光强度的级别的范围是很宽的，由夜视阈值到强闪光之间光强度的级别约为 10^{10} 级，相当多的实验证据表明主观亮度是进入眼内的光强度的对数函数。这一特性示于图 1-3，它表示光强度与主观亮度之间的关系。长的黑线代表人的视觉系统所能适应的光强度的范围。昼视觉范围为 10^6 ，由夜视觉到昼视觉是逐渐过渡的，过渡的范围大致由 0.001 ~ 0.1 毫朗伯^①在图中画出了这一范围内的适应曲线。

解释在图 1-3 中的特殊动态范围的关键在于人的视觉绝对不能同时在这一范围工作，

注：① 1 朗伯 = 1 lm/cm^2

说得更确切一点,它是利用改变其整个灵敏度来完成这一大变动的,这就是所谓亮度适应现象。当与整个适应范围相比时,能同时鉴别光强度级的总范围是很小的。对于任何一组给定的条件,视觉系统的当前的灵敏度级叫做亮度适应级,例如它相当于图1-3中的亮度 B_a ,短的交叉曲线表示当眼睛适应于这一强度级时,人眼能感觉的主观亮度的范围。必须注意这一范围是有一定限制的,在 B_b 处和 B_b 以下时,所有的刺激都是作为不可分辨的黑来理解的。

眼睛的对比灵敏度可用下述方法测量出来,使观测者处于亮度为 B 的均匀光场中,中部有一个具有鲜明边界的圆形靶,其亮度为 $B + \Delta B$,如图1-4(a)所示。 ΔB 由零值一直升到刚好被觉察到的值为止。这个刚能觉察到的亮度差值 ΔB 是作为 B 的函数而被量测的。 $\Delta B/B$ 称为韦伯比,在非常宽的亮度范围内,它近似为一常数,约为2%,如图1-4(b)所示。这一现象导致下述看法,即人眼具有比人造图像系统宽得多的动态范围,如果利用图1-5(a)的图案可得到更加适用的结果,再一次测出 $\Delta B/B$,但这里的 B_0 (周围亮度)是一个参数,图1-5(b)中画出了结果。动态范围在以周围亮度为中心的2.2单位内变动,如果背景亮度调节正确的话,那么这一结果可以与用电学图像系统所能达到的结果相比较。真正引人注意的特性是眼睛自身适应的难易强度和快慢(在视网膜的不同部分是不同的),而不是它的整个动态范围。

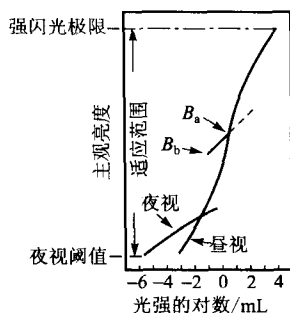


图1-3 表示特殊适应能力的
主观亮度感觉的范围

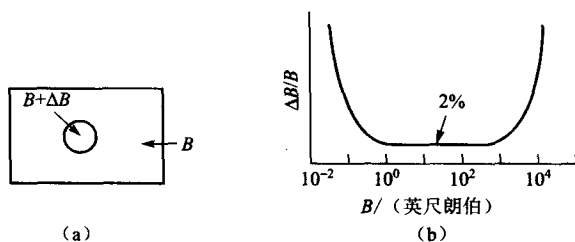


图1-4 具有恒定背景的对比灵敏度

在复杂图像的情况下,视觉系统不会适应于单一光强度级。相反,它适应于一平均光强度级,这一平均光强度级取决于图像的性质。当眼睛在景物周围徘徊时,瞬时适应能力围绕其平均值上下波动。在图像中任何一点或小面积上,韦伯比一般要比在实验环境中得到的值大得多,这是因为在背景中缺乏鲜明的限制边界和强度变化。其结果,眼睛只能在一个复杂的图像中的任何一点上的十多个或二十多个亮度级附近进行检测。但这并不意味着图像只需要用二十多个光强度级进行显示就能获得令人满意的视觉效果。上面的窄的

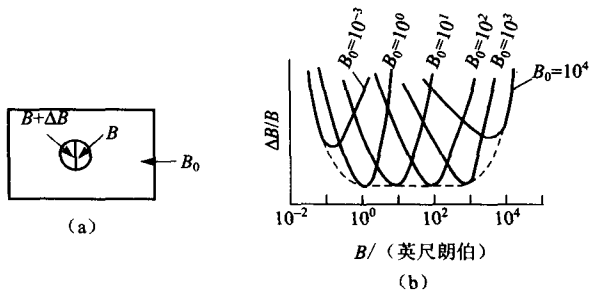


图1-5 具有可变背景的对比灵敏度

鉴别范围跟随着适应力而变化,因为适应能力需要变化,以便当眼睛沿着景物移动时,适应不同的光强度级。这允许有大得多的全光强度鉴别的范围。对于一大类图像类型来说,为了使眼睛得到适当舒服的显示,一般需要大于 100 光强度级的范围。

人眼能感觉到的一个区域的高密度取决于许多因素,而不是简单地取决于由该区辐射出的光,根据图像处理的用途,有关亮度感觉的最有价值的现象是人的视觉系统的响应在不同光强度区的边界周围引起的“过量调整”。这一过量调整的结果会造成某个恒定光强度区域的出现,如同这些区域原来就具有可变的亮度一样。例如,在图 1-6 (a) 中所画的图像是

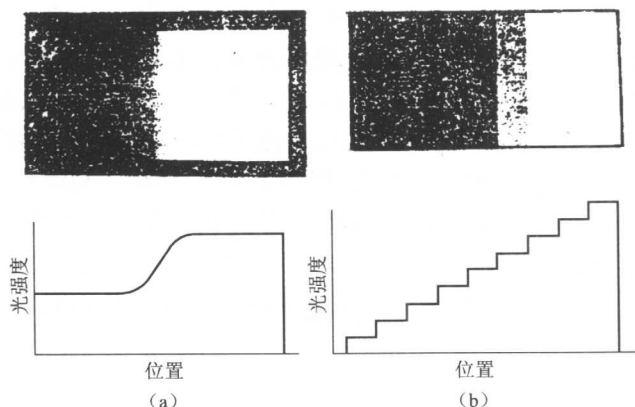


图 1-6 马赫带效应的例子

根据照片下面的光强度分布而形成的。虽然光强度变化是完全圆滑的,但人的眼睛所感受到的是 B 区中较亮的带和 D 区中较暗的带。这些带叫做马赫带。马赫带效应的更加明显的例子示于图 1-6 (b) 中。正如光强度分布图所指示的那样,照片中的每一条带都是由一固定光强度造成的,然而对于眼睛来说,图像中的亮度图案,特别在边界的周围,出现了强度的变化。

1.2.2 视觉模型

将视觉系统的功能抽象化为简单的模型。并以此模型为基础来对视觉系统的功能进行研究或加以模仿是很吸引人的。建立模型首先要提出假说,根据假说建立模型,再根据试验修改模型,反复进行使之逐渐完善合理。建立模型可以有电子线路模型以及化学模型等,对我们来说,一般是对电子线路模型感兴趣,这样可以把视觉系统的一些优越性引入图像通信和信息处理系统加以研究和应用。

下面简单地介绍一下神经元模型,黑白视觉模型和彩色视觉模型。

1. 神经元模型

从信息处理观点出发,在神经元所具有的各种机能中最重要的是,在突触处许多输入在空间和时间上进行加权的性质以及细胞的阈值作用。有两种神经元模型,一种是针对研究空间加权特性的不计时间特性的神经元模型,另一种是考虑到时间特性的空间加权特性的模型。现对后一种略加介绍。

(1) 神经元信息的产生及传递特性。

神经元由细胞体及其轴突(神经纤维)和树突(细胞体粗短突起)组成,轴突末端称为末梢,一个神经元末梢与另一个神经元细胞体或树突相接触形成突触,神经元之间的信息传递都是通过突触进行的,见图 1-7^[1-1]。平常细胞膜内外保持一个约 70 mV 的电位差,当细胞兴奋时就会发出幅度为 100 mV,宽度为 1 ms 左右的脉冲电位,其脉冲的频率随细胞兴奋程度而变。这一电位经突触传给另一个神经元,产生所谓突触后电位简称 PSP (Post Synaptic Potential)。把能够产生的 PSP 的突触称为兴奋性突触,它产生的兴奋性突触后电位,简称 EPSP,与此相反地还有抑制性突触产生的抑制性后电位 IPSP。

在神经元内部之间的传输信息（兴奋）是有速度的，约几米每秒到几十米每秒，因此神经元有时间效应。

轴突有保持某一固有脉冲的性质，比较小的波形在传递中可以增幅，比较大的波形在传递中可以衰弱。对脉冲宽度有同样的性质，因此有相似于自动增益控制作用。不管输入波形有什么变化，输出波形是恒定的，而且对同类波形可以恒定速度传送。

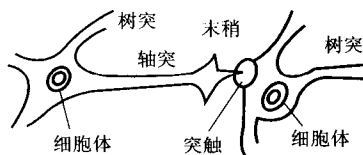


图 1-7 神经元之间突触联系

但轴突传送的信号，对于它的脉冲幅度要有一定的阈值，比阈值小的信号在传送过程中就消失了，这种阈值作用相似于静化噪声的作用。神经元的阈值变化是这样的：在发放之后有一段时间约为 1 ms，无论给予多么强的刺激都不再发放，称为绝对不应期。稍后阈值慢慢下降并恢复到原来状态，这段慢慢下降期，即阈值比正常值高又难于引起反应的期间称为相对不应期。一般来讲，信息在轴突中的传送是单方向的，而突触传送既有单向也有双向的。

(2) 考虑时间特性的神经元脉冲模型。

在考虑时间特性的模型中，有些是和实际神经元同样产生脉冲输出的模型，也有电子线路模型^[1-2]。其主要功能框图如图 1-8 所示。

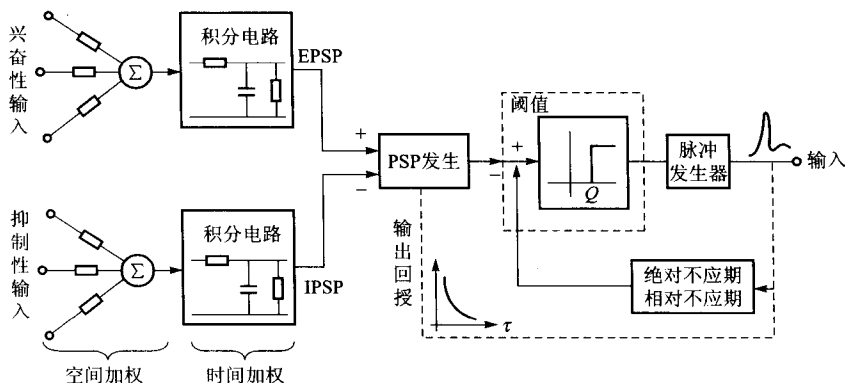


图 1-8 脉冲输出神经元模型功能框图

模型的输入输出都是脉冲序列，由 EPSP 和 IPSP 的干涉（如两者的差或和）来决定 PSP，当 PSP 超过一定阈值 Q 时，产生脉冲输出，脉冲发生后进入绝对不应期，脉冲放电暂时被抑制，接着进入了相对不应期，随着阈值恢复正常，下一个脉冲立即发生，PSP 越大，脉冲重复频率越高。

2. 黑白视觉模型

研究下列线性光学系统如图 1-9 所示。

设 $\varphi_i(\omega_x, \omega_y)$ 为输入图像 $I_i(x, y)$ 的傅里叶变换， $\varphi_o(\omega_x, \omega_y)$ 为输出图像 $I_o(x, y)$ 的傅里叶变换。根据线性系统理论可得

$$\varphi_o(\omega_x, \omega_y) = H(\omega_x, \omega_y) \cdot \varphi_i(\omega_x, \omega_y)$$

在许多情况下，人们只对相对输入图像强度和振幅变化的输出图像强度感兴趣，因此由下式描述

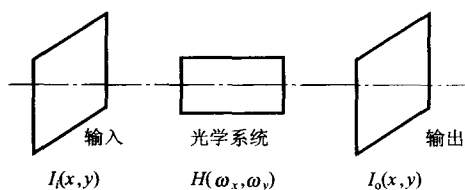


图 1-9 线性光学系统

$$H(\omega_x, \omega_y) = \frac{\varphi_o(\omega_x, \omega_y)}{\varphi_i(\omega_x, \omega_y)}$$

式中, $H(\omega_x, \omega_y)$ 为该光学系统的调制传递函数 MTF。

把光学系统的概念用到人的视觉系统方面已经作出了很多努力, 从实验结果可以看出, MTF 测试结果与输入对比度的大小有关, 而且如果当输入正弦光栅相对于眼球光轴进行旋转时, 则测得的 MTF 形状也有某些改变。于是可以推出以下结论: 人的视觉系统由实验测出的 MTF 是非线性的和各向异性的 (即旋转可变的)。另外, 根据一些人的实际结果^[1-3] 可以认为, 眼睛对光强度的非线性响应呈对数型, 并且发生在视觉系统的开始附近 (亦就是视觉信号在锥状及杆状细胞空间上发生相互作用之前)。由此得出, 人眼黑白视觉的简单对数模型如图 1-10 所示。

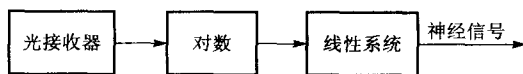


图 1-10 黑白视觉对数模型

下面介绍一个表示黑白视觉扩展的模型, 如图 1-11 所示。在模型中把对不同波长有不同灵敏度的感受器输出馈送到代表眼睛光学性能的低通线性系统 $H_1(\omega_x, \omega_y)$, 接着呈代表杆状和锥状视细胞非线性强度响应的一般黑白非线性网络, 下面就是代表侧抑制过程的具有带通性能的线性系统 $H_2(\omega_x, \omega_y)$, 最后是具有时间滤波效应的线性系统 $H_3(\omega_x, \omega_y)$ 。

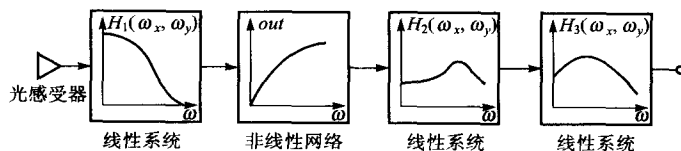


图 1-11 黑白视觉扩展模型

3. 彩色视觉模型

根据三色假说为基础来研究彩色视觉模型, 图 1-12 是 Frei 建议的彩色模型。在该模型中, e_1, e_2, e_3 代表视网膜三个具有 $S_1(\lambda), S_2(\lambda), S_3(\lambda)$ 谱灵敏度的感受器, 其输出分别为

$$e_1 = c(\lambda) S_1(\lambda)$$

$$e_2 = c(\lambda) S_2(\lambda)$$

$$e_3 = c(\lambda) S_3(\lambda)$$

式中, $c(\lambda)$ 为入射光源的谱线分布函数。 e_1, e_2, e_3 经对数传递后并组合为 d_1, d_2, d_3 输出

$$d_1 = \lg(e_1)$$

$$d_2 = \lg(e_2) - \lg(e_1) = \lg \frac{e_2}{e_1}$$

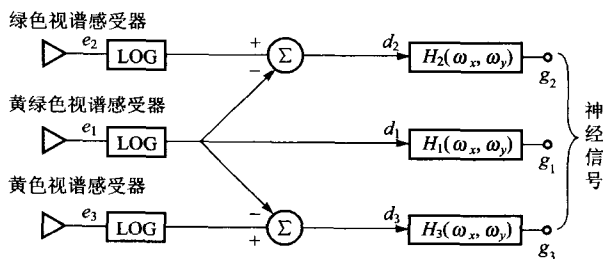


图 1-12 彩色视觉模型

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

最后, 信号 d_1 , d_2 , d_3 分别经传递函数 $H_1(\omega_x, \omega_y)$, $H_2(\omega_x, \omega_y)$, $H_3(\omega_x, \omega_y)$ 的线性系统输出 g_1 , g_2 , g_3 , 由这些信号提供大脑感受彩色的基础。

信号 d_2 和 d_3 与彩色光的色度有关, 而 d_1 则正比于它的亮度。这个模型相当准确地预测许多彩色现象, 也能满足色度学的基本定律。例如, 彩色光的谱能量被改变为一个常数, 则光线的色调和饱和度正如色度坐标定量描述那样, 在一个宽的动态范围内是不会变的, 从以上的公式可以看出 d_2 和 d_3 彩色信号是不会变的, 而亮度信号 d_1 则按对数规律增加。

与图 1-11 所示的黑白视觉扩展模型类似, 对数的彩色视觉模型很容易被扩展, 考虑到眼睛的响应, 只要在每一种锥状视细胞感受器后插入一个线性传递函数即可。同样, 可用一般的非线性网络代替对数传递函数网络。应当注意, 感受器的相加和传递函数的运算次序可以互换而不影响其输出, 因为两者都是线性运算, 图 1-13 为彩色视觉扩展模型。

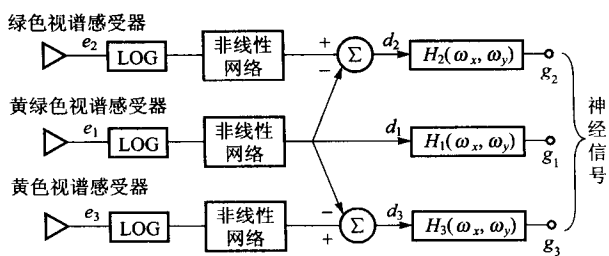


图 1-13 彩色视觉扩展模型

由 Vonkrio^[1-4] 提出的关于彩色适应最简单的模型, 就是在图 1-13 中的锥状细胞感受器和第一个线性系统之间插入自动增益控制, 其增益为

$$a_i = [w(\lambda) S_i(\lambda) d \lambda]^{-1} \quad i = 1, 2, 3$$

被这样调整后, 当观察参考白色及谱能分布为 $w(\lambda)$ 图像时, 修正后的锥细胞响应为 1。Vonkrio 的模型是很吸引人的, 因为它在定性上是合理的, 而且简单。但实际彩色测试证明, 该模型并不能完全预测彩色适应效应。Wallio 提出彩色适应部分是由网膜在神经抑制机构中所产生的, 该机构使图像中变化分量慢慢地线性衰减。这可将图 1-13 中网膜后传递函数 $H_1(w_x, w_y)$ 的低空间频率衰减用来模型化。毫无疑问, 视神经和网膜后机构都可能产生彩色适应效应, 为此更精确的模型尚需进一步研究。

§ 1.3 图像的数字化

我们已经知道, 计算机所能处理的信息必须是数字信号, 而我们得到的照片、图纸或景物等信息都是连续信号, 为此必须将此连续信号进行抽样和量化, 即进行数字化处理。设连续图像 $f(x, y)$ 经过等间隔抽样以后, 可以用一个离散量组成的矩阵来表示

$$f(x, y) \approx \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,n-1) \\ f(1,0) & f(1,1) & \cdots & f(1,n-1) \\ \vdots & \vdots & & \vdots \\ f(n-1,0) & f(n-1,1) & \cdots & f(n-1,n-1) \end{bmatrix}$$

矩阵中的每一个元素称作像元、像素或图像元素。而 $f(x, y)$ 代表 (x, y) 点的灰度值, 即亮度值。以上数字化有以下几点补充说明一下。