



北京工业大学"211工程"资助出版

信息科学技术学术著作丛书

基于H.264的视频编码 处理技术与应用

贾克斌 刘鹏宇 吕卓逸 邓智玘 编著



科学出版社



信息科学技术学术著作丛书
北京工业大学“211 工程”资助出版

基于 H. 264 的视频编码 处理技术与应用

贾克斌 刘鹏宇 编著
吕卓逸 邓智玘

科学出版社

北 京

内 容 简 介

新一代的视频编码标准——H. 264/AVC, 与之前的国际标准相比, 采用了多种新编码技术, 因而应用前景更加广泛。本书是围绕 H. 264/AVC 标准的视频编码相关技术的一本学术著作, 在叙述数字视频概念、经典的视频压缩编码原理的基础上, 介绍了国际视频编码标准, 详细描述了 H. 264/AVC 的关键技术。针对近年来 H. 264/AVC 研究的最新进展, 详细介绍了关于 H. 264/AVC 的快速编码、三维视频编码与视频转码的概念和实现技术, 这是本书的特色与核心内容。同时结合具体的 DSP 芯片, 描述基于硬件的 H. 264/AVC 的编码实现技术与应用系统构建技术。此外, 本书也涉及并介绍了正在制定的 HEVC 视频编码的国际标准及相关内容。

本书可作为高等院校信息与通信工程、计算机科学与技术等相关专业研究生或高年级本科生的教材或参考书, 也可以作为多媒体信息处理领域工程技术人员的参考书。

图书在版编目(CIP)数据

基于 H. 264 的视频编码处理技术与应用/贾克斌等编著. —北京: 科学出版社, 2013

(信息科学技术学术著作丛书)

ISBN 978-7-03-036044-1

I. 基… II. 贾… III. 视频编码-研究 IV. TN762

中国版本图书馆 CIP 数据核字(2012)第 273508 号

责任编辑: 魏英杰 张艳芬 杨向萍 / 责任校对: 桂伟利

责任印制: 张 倩 / 封面设计: 陈 敬

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>

北京通州皇家印刷厂印刷

科学出版社发行 各地新华书店经销

*

2013 年 1 月第 一 版 开本: B5(720×1000)

2013 年 1 月第一次印刷 印张: 22 3/4

字数: 439 000

定价: 88.00 元

(如有印装质量问题, 我社负责调换)

《信息科学技术学术著作丛书》序

21 世纪是信息科学技术发生深刻变革的时代,一场以网络科学、高性能计算和仿真、智能科学、计算思维为特征的信息科学革命正在兴起。信息科学技术正在逐步融入各个应用领域并与生物、纳米、认知等交织在一起,悄然改变着我们的生活方式。信息科学技术已经成为人类社会进步过程中发展最快、交叉渗透性最强、应用面最广的关键技术。

如何进一步推动我国信息科学技术的研究与发展;如何将信息技术发展的新理论、新方法与研究成果转化为社会发展的新动力;如何抓住信息技术深刻发展变革的机遇,提升我国自主创新和可持续发展的能力? 这些问题的解答都离不开我国科技工作者和工程技术人员的求索和艰辛付出。为这些科技工作者和工程技术人员提供一个良好的出版环境和平台,将这些科技成就迅速转化为智力成果,将对我国信息科学技术的发展起到重要的推动作用。

《信息科学技术学术著作丛书》是科学出版社在广泛征求专家意见的基础上,经过长期考察、反复论证之后组织出版的。这套丛书旨在传播网络科学和未来网络技术,微电子、光电子和量子信息技术、超级计算机、软件和信息存储技术,数据知识化和基于知识处理的未来信息服务业,低成本信息化和用信息技术提升传统产业,智能与认知科学、生物信息学、社会信息学等前沿交叉科学,信息科学基础理论,信息安全等几个未来信息科学技术重点发展领域的优秀科研成果。丛书力争起点高、内容新、导向性强,具有一定的原创性;体现出科学出版社“高层次、高质量、高水平”的特色和“严肃、严密、严格”的优良作风。

希望这套丛书的出版,能为我国信息科学技术的发展、创新和突破带来一些启迪和帮助。同时,欢迎广大读者提出好的建议,以促进和完善丛书的出版工作。

中国工程院院士

原中国科学院计算技术研究所所长



序

“211工程”是新中国成立以来教育领域的国家重点建设工程,面向21世纪重点建设100所高水平大学,使其成为我国培养高层次人才,解决经济建设、社会发展和科技进步重大问题的基地,形成我国高等学校重点学科的整体优势,增强和完善国家科技创新体系,跟上和占领世界高层次人才培养的脚步和科技发展的制高点。

中国高等教育发展迅猛,尤其是1400所地方高校已经占全国高校总数的90%,成为我国高等教育实现大众化的重要力量,成为区域经济和社会发展服务的重要生力军。

在北京市委、市政府的高度重视和大力支持下,1996年12月北京工业大学通过了“211工程”部门预审,成为北京市属高校唯一进入国家“211工程”重点建设的百所大学之一。我校紧紧抓住“211工程”建设和举办奥运会的重要机遇,实现了两个历史性的转变:一是实现了从单科性大学向以工科为主,理、工、经、管、文、法相结合的多科性大学的转变;二是实现了从教学型大学向教学研究型大学的转变。“211工程”建设对于我校实现跨越式发展、增强服务北京的能力起到了重大的推动作用,学校在学科建设、人才培养、科学研究和服务北京等方面均取得了显著的成绩,综合实力和办学水平得到了大幅度的提升。

至2010年年底,我校的学科门类已经覆盖了8个:工学、理学、经济学、管理学、文学、法学、哲学和教育学。现拥有8个一级学科博士学位授权点、37个二级学科博士学位授权点和15个博士后科研流动站;15个一级学科硕士学位授权点和81个二级学科硕士学位授权点;拥有6种类型硕士研究生专业学位授权资格,工程硕士培养领域19个;拥有3个国家重点学科、16个北京市重点学科和18个北京市重点建设学科。

目前,学校有专任教师1536人,全职两院院士5人,博士生导师220人,有正高职称294人和副高职称580人,专任教师中具有博士学位教师的比例达到54.6%。有教育部“长江学者”特聘教授4人,国家杰出青年基金获得者6人,入选中组部“千人计划”1人,北京市“海聚工程”3人,教育部新(跨)世纪优秀人才支持计划15人。

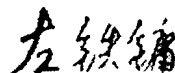
2010年学校的到校科研经费为6.2亿元。“十一五”期间,学校承担了国家科技重大专项28项,“973计划”项目16项,“863计划”项目74项,国家杰出青年基金2项,国家自然科学基金重点项目8项,科学仪器专项2项,重大国际合作项目

1 项,面上项目和青年基金项目 347 项,北京市自然科学基金项目 180 项,获国家级奖励 14 项。现有 1 个共建国家工程研究中心,7 个部级或省部共建科研基地,11 个北京市重点实验室和 3 个行业重点实验室。

为了总结和交流北京工业大学“211 工程”建设的科研成果,学校设立了“211 工程”专项资金,用于资助出版系列学术专著。这些专著从侧面代表了我校教授、学者的学科方向、研究领域、学术成果和教学经验。

展望北工大未来,任重而道远。我坚信,只要珍惜“211 工程”建设的重要机遇,构建高层次学科体系,营造优美的校园氛围,我校在建设国际知名、有特色、高水平大学的进程中就一定能够为国家,特别是为北京市的经济建设和社会发展作出更大的贡献。

中国工程院院士
北京工业大学原校长



2011 年 6 月

前 言

当今时代,随着多媒体信息技术和网络通信技术的迅猛发展,作为信息处理领域的重要分支,数字视频技术的应用领域不断拓宽,IPTV、PDA、立体电影和自由视点视频等多样化视频服务业务不断涌现,视频编码已成为数字电视、网络视频和移动多媒体等信息产业的基础核心技术之一,在国防、医疗、教育和传媒等领域具有广阔的应用前景。由于未经压缩的视频源常伴随海量数据的产生,因此压缩编码后的视频图像质量、编码效率以及编码实时性能是决定视频编码技术在多媒体通信等领域获得大规模应用的关键因素。过去的十几年里,ISO/IEC MPEG 和 ITU-T VCEG 两大国际标准化组织相继成功制定了面向多种应用的视频编码标准,对视频编码技术的发展起到了巨大的推动作用,对多媒体通信等信息产业的发展也产生了深远的影响。因此,对视频编码技术的研究已成为信息技术领域的研究热点和重点。

H. 264/AVC 是由 ISO/IEC 和 ITU-T 联合制定的新一代视频编码标准,它在继承预测变换混合视频编码框架的基础上,融入了视频编码的最新技术和成果,在预测、变换量化、熵编码和容错等技术细节上作出优化与改进,具有更高的压缩效率和稳定的网络友好性,应用前景十分广阔。与此同时,具有临场感和交互性的三维立体视频逐渐成为多媒体信息产业的热门话题。为了确定一种国际通用的三维视频格式,使其支持高质量的自由视点视频应用,2011 年 MPEG 成立最新的 3DG 工作组,着手制定国际三维视频编码标准——3DV/FTV,并且面向全球各大科研机构广泛征集三维视频编码技术的实现方案。然而,H. 264/AVC 在获得高效编码性能的同时也带来了巨大编码复杂度,阻碍了该标准的实时应用与推广。因此,如何有效降低编码复杂度并保持编码优势成为 H. 264/AVC 及三维视频编码标准获得成功的关键因素,也是数字视频技术大规模实用化的迫切需要。

本书是围绕 H. 264/AVC 标准的视频编码相关技术的一本学术专著,主要论述数字视频基础与视频编码的有关国际标准、H. 264/AVC 编码标准和最新进展、三维视频编码、视频转码的概念与关键技术以及基于硬件的实现和应用技术等。

全书共 10 章。第 1~4 章是基础知识部分,介绍人类视觉系统、数字视频和视频编码的原理和相关标准。第 5~9 章是本书的核心内容,主要描述 H. 264/AVC 及其相关的视频编码技术,其中第 5~6 章介绍 H. 264/AVC 编码原理及其主要技术;第 7 章介绍三维视频编码技术;第 8~9 章介绍视频转码概念与技术。第 10 章是应用部分,介绍基于 DSP 的 H. 264/AVC 编码的实现与应用方法。

本书是在北京工业大学数字多媒体信息处理与通信技术研究室近十年研究成果的基础上完成的,这些研究工作得到了国家自然科学基金、北京市自然科学基金和北京市教委重点科技发展基金的大力资助,直接参加这些研究工作的有刘鹏宇博士、吕卓逸博士、邓智玘博士、谢晶硕士、李可硕士、高海辉硕士、何絮硕士、李巍硕士、高俊巍硕士、田洪宁硕士等。目前,这些已经毕业的同学都已成为奋战在国内外著名研究所、高校和 IT 公司的业务骨干。在本书即将出版之际,衷心地感谢这些同学,他们当年在研究室辛勤工作的身影仿佛就在眼前。没有这些同学的勤奋工作,不可能有本书的研究成果。

值此书稿完成之际,深深感谢香港理工大学多媒体信号处理研究室的萧允治教授、陈锐霖博士,感谢美国纽约州立大学布法罗分校(UB)计算机系主任张爱冬教授。萧允治教授是国际著名的数字视频编码专家,他和陈锐霖博士多次到研究室来指导工作和参加博士生毕业答辩,并且曾多次邀请作者到香港理工大学进行合作研究和联合指导博士研究生,研究室取得的成果离不开他们的无私帮助。张爱冬教授在指导研究室科研工作和联合培养博士研究生方面也付出了辛勤的劳动。

贾克斌负责本书编著内容的统筹和大纲制定,刘鹏宇负责完成了本书第 5、6 章的初稿,邓智玘完成了第 7 章的初稿,吕卓逸完成了第 8、9 章的初稿,赵丽娅、周作成、蒋斌、窦环、王丽梅等帮助整理了书稿的部分素材,贾克斌负责完成了第 1~4 章、第 10 章的书稿,同时负责全书的定稿工作。杨明洁、李蕊审阅了主要部分的书稿,冯金超博士、孙中华博士等对本书的出版也作出了贡献,在此对他们辛勤的劳动一并表示衷心的感谢。

感谢北京工业大学“211 工程”学术专著出版基金对本书出版的大力资助。

由于作者水平和时间有限,书中难免有不妥之处,敬请批评指正。

贾克斌

于北京工业大学科学楼

2012 年 8 月 1 日

目 录

《信息科学技术学术著作丛书》序

序

前言

第 1 章 人类视觉与计算机成像	1
1.1 人类视觉系统	1
1.1.1 人眼结构	2
1.1.2 视神经系统	4
1.1.3 处理	4
1.2 颜色描述	5
1.2.1 颜色的感知	5
1.2.2 颜色模型	7
1.3 计算机视觉系统	10
1.3.1 摄像	11
1.3.2 计算机界面	14
1.3.3 处理与成像	15
1.4 小结	18
参考文献	18
第 2 章 数字视频	19
2.1 数字视频概念	19
2.2 视频质量	21
2.3 小结	26
参考文献	27
第 3 章 视频压缩编码基本原理	28
3.1 视频编码基础	28
3.1.1 视频编码的分类	29
3.1.2 图像的信息熵	30
3.2 预测编码	30
3.2.1 预测编码基本概念	30
3.2.2 帧内预测编码	32
3.2.3 帧间预测编码	33
3.2.4 运动估计	35
3.3 变换编码	43

3.3.1 变换编码基本概念	43
3.3.2 K-L 变换	44
3.3.3 离散余弦变换	44
3.3.4 zig-zag 扫描和游程编码	46
3.4 熵编码	47
3.4.1 变长编码	47
3.4.2 算术编码	49
3.5 现代视频编码技术	50
3.6 小结	51
参考文献	51
第 4 章 视频编码标准	53
4.1 视频编码标准的发展过程	53
4.2 H.26X 系列标准	54
4.2.1 H.261 标准	54
4.2.2 H.263 标准	55
4.2.3 H.264/AVC 标准	56
4.3 MPEG-1 标准	60
4.3.1 MPEG-1 标准内容	60
4.3.2 MPEG-1 标准视频数据流结构	60
4.4 MPEG-2 标准	61
4.4.1 MPEG-2 标准概述	61
4.4.2 MPEG-2 标准中的概念与原理	62
4.4.3 MPEG-2 标准的特点	63
4.5 MPEG-4 标准	64
4.5.1 MPEG-4 标准概述	64
4.5.2 MPEG-4 标准中的概念与原理	65
4.5.3 MPEG-4 编码标准的特点	66
4.6 MPEG-7 及 MPEG-21 标准	67
4.6.1 MPEG-7 标准	67
4.6.2 MPEG-21 标准	70
4.7 小结	73
参考文献	73
第 5 章 H.264/AVC 编码原理	75
5.1 H.264 视频编码标准简述及其关键技术	75
5.1.1 H.264 的编码特点	75
5.1.2 H.264 的帧内预测编码	76

5.1.3 H. 264 帧间预测编码	80
5.1.4 SP 帧和 SI 帧的码流切换	88
5.1.5 整数 DCT 变换与量化技术	88
5.1.6 熵编码技术	88
5.1.7 多参考帧预测技术	89
5.1.8 码率控制技术(率失真优化技术)	90
5.1.9 去块效应滤波器	92
5.2 H. 264 视频编码标准的码流结构和分层结构	92
5.2.1 H. 264 的档次和级	92
5.2.2 H. 264 标准的码流结构	93
5.2.3 H. 264 标准的分层结构	93
5.3 H. 264 视频编码标准的编码结构和解码结构	94
5.3.1 H. 264 标准的编码结构	94
5.3.2 H. 264 标准的解码结构	95
5.4 小结	95
参考文献	96
第 6 章 H. 264/AVC 编码技术新进展	98
6.1 H. 264/AVC 实时应用的瓶颈	98
6.2 快速模式选择技术	100
6.2.1 快速帧内预测技术	100
6.2.2 快速帧间模式选择技术	115
6.3 快速运动估计补偿技术	135
6.3.1 UMHexagonS 算法特点与计算冗余	135
6.3.2 基于运动矢量相关性的快速运动估计算法	141
6.3.3 仿真实验结果与分析	149
6.3.4 小结	154
6.4 基于 H. 264 的视频感兴趣区域的获取技术	155
6.4.1 视觉选择性注意机制在视频编码中的应用	155
6.4.2 基于编码信息的视觉感知特征分析	159
6.5 视频编码技术的最新进展	172
6.5.1 下一代视频编码标准 HEVC	172
6.5.2 基于 H. 264/AVC 的三维视频编码技术	179
6.5.3 基于 H. 264/AVC 的视频转码技术	180
6.6 小结	180
参考文献	183

第 7 章 三维视频编码技术	189
7.1 三维视频概述	189
7.1.1 深度感觉	189
7.1.2 立体成像原理	190
7.1.3 双目立体视频	191
7.1.4 多视点视频	191
7.2 三维视频采集方式	192
7.2.1 三维摄像机	192
7.2.2 二维转三维技术	193
7.3 三维视频显示技术	193
7.3.1 眼镜式三维立体显示技术	193
7.3.2 裸眼三维立体显示技术	196
7.4 多视点视频应用系统	198
7.4.1 基于多视点视频的三维电视	199
7.4.2 自由视点电视	201
7.4.3 基于深度图的三维视频系统	202
7.5 三维视频编码技术	204
7.5.1 纯视频格式三维视频	204
7.5.2 深度增强格式三维视频	226
7.6 国际三维视频的标准化进程	229
7.7 小结	231
参考文献.....	231
第 8 章 视频转码技术	234
8.1 视频转码技术发展现状	234
8.1.1 空间分辨率转码	234
8.1.2 时间分辨率转码	235
8.1.3 H.264 的视频转码	236
8.2 视频转码原理	237
8.3 视频转码技术的分类	239
8.3.1 比特率变换	239
8.3.2 分辨率转换编码	242
8.3.3 异类比特流间的转换编码	244
8.4 视频转码的关键技术	246
8.4.1 编码模式映射	246
8.4.2 运动矢量的映射与优化	246
8.5 小结	247

参考文献	247
第 9 章 基于 H. 264 同类转码技术	249
9.1 概述	249
9.2 基于 H. 264 的同类转码框架	249
9.3 转换分辨率转码	251
9.3.1 空间分辨率转码	251
9.3.2 时间分辨率转码	251
9.4 宏块模式选择	253
9.4.1 帧间宏块模式选择	253
9.4.2 帧内宏块模式选择	268
9.5 运动矢量重构	275
9.5.1 降尺寸转码运动矢量合成算法	275
9.5.2 跳帧转码运动矢量合成算法	278
9.6 实验结果	280
9.6.1 降尺寸视频转码	280
9.6.2 跳帧转码	291
9.7 小结	293
参考文献	294
第 10 章 基于 DSP 的 H. 264 编码技术的实现与应用	296
10.1 高速 DSP 应用现状	296
10.2 基于 DSP 的视频编码器实例	298
10.2.1 DM642 硬件平台和开发环境	299
10.2.2 代码移植	303
10.2.3 存储器优化	305
10.2.4 Ping-Pong 机制下的 EDMA 数据传输	309
10.2.5 代码优化	312
10.2.6 编码器性能测试	315
10.3 基于嵌入式 DSP 的智能视频监控系统	319
10.3.1 系统的整体架构	319
10.3.2 视频采集处理端功能设计与开发	321
10.3.3 视频集中监控业务平台端设计	330
10.3.4 用户端控制功能的实现	337
10.4 小结	338
参考文献	339
附录 英汉对照词汇表	341

第 1 章 人类视觉与计算机成像

视觉是人类感知外界信息的重要通道,外界信息的 80% 以上是通过视觉得到的。视觉系统通过眼睛获得图片,继而传送至大脑进行视觉信息的处理与理解。周围环境中的物体在太阳光的照射下,进入眼睛经视网膜成像,视细胞将其转换成神经脉冲信号,由神经纤维传送至大脑皮层进行处理与理解。

随着信号处理理论、计算机技术以及摄像机电子业发展逐渐成熟,通过电子摄像机获得图像并对其进行模数转换,用计算机对视觉信息进行处理,逐渐形成了计算机视觉这一新的学科。计算机视觉研究目标之一是利用二维图像认知三维景物世界。通过分析二维图像获得三维环境中物体的如形状、大小、位置和姿态等特征,从而达到对物体的描述、理解和认知^[1]。如今,计算机视觉应用于各种领域,像工业产品零件大小窥视、食品质量检测、宇宙卫星摄像、自动人脸识别、生物医学、视觉导航与虚拟现实等^[1]。与此同时,生物学家和心理学家也一直致力于人类视觉系统的研究,所形成的神经心理学、心理学与认知科学等学科对计算机视觉起到了很好的启发与指导作用。但计算机视觉并不完全局限于人类视觉系统,它拥有自己独立的计算方法和理论体系。

本章将在简单介绍人类视觉系统的同时,致力于介绍计算机视觉系统是如何展开工作的。首先介绍人类视觉系统的关键部分,展现人眼的工作模式。对于计算机视觉系统,作为基础知识,本章介绍了几种流行的图像颜色模型,视频采集设备以及处理和成像。

1.1 人类视觉系统

复杂的人类视觉系统能够感知视觉刺激并做出相应的反应。它经历了数百万年的发展,主要是为了防御和人类的生存^[2]。人类视觉系统分为眼睛、视神经通路、外膝体和视觉皮层 4 个部分。在人类视觉中,图像进入人眼这一感知器官,通过视神经传送到大脑作进一步处理。视神经没有足够的带宽传输由人眼感知的所有信息,因此,在将图像传送到视神经之前必须对图像进行预处理。根据功能,将人类视觉系统分为如下 3 种模型。

- (1) 人眼:一个物理模型,它的大部分功能可以通过病理学来呈现。
- (2) 处理系统:一个实验模型,它的功能可以被模仿,但是不能精确定位。
- (3) 大脑分析:一个心理学上的模型,不能直接访问或模拟它的处理过程,只

能通过实验和资料了解它的行为^[2]。

1.1.1 人眼结构

人类眼球近似一个球体,直径约 23mm,在肌肉的作用下,它可以转动而跟踪运动的对象,并产生运动的图像。

如图 1-1 所示,眼球由 3 层薄膜包围着,从外到内分别为蛋白质膜、虹膜和脉络膜、视网膜。其中位于最外层蛋白质膜正前方的大约 1/6 部分为有弹性的透明组织,称为眼角膜,光线经角膜产生屈折进入人眼内。其余 5/6 为白色不透明组织,称为巩膜,它主要起巩固和保护整个眼球的作用。

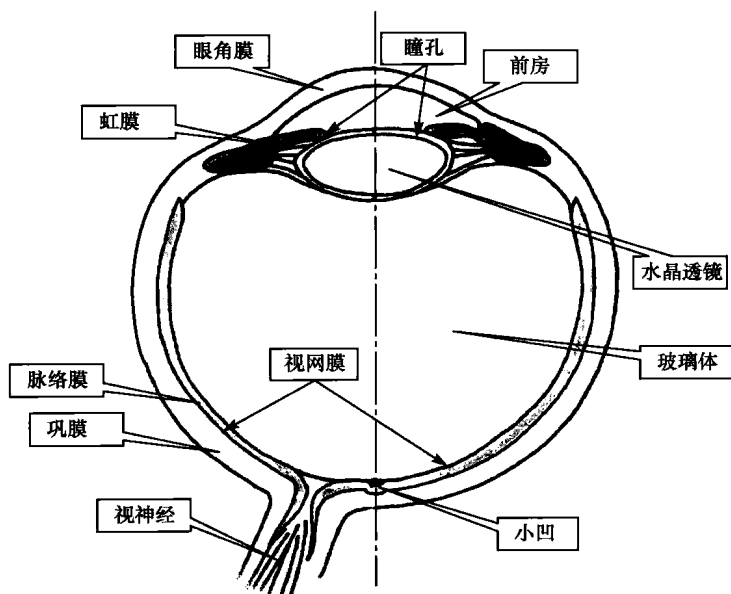


图 1-1 人眼横截面简图

中间一层由虹膜和脉络膜组成。巩膜的下面是脉络膜,脉络膜含有丰富的色素细胞,呈黑色,起着吸收外来散光的作用,消除光线在眼球内部的乱反射。在脉络膜的最前面分布着睫状体和虹膜。虹膜在眼角膜的后面,位于晶状体的前面,角膜中央有一个圆孔称为瞳孔,它的大小由两侧连接虹膜的睫状肌来调节,直径范围在 2~8mm,以达到控制进入眼睛内部光通量大小的目的。虹膜细胞中所含色素量的多少决定了虹膜的颜色,东方人是有色人种,虹膜中色素含量多,所以眼珠看上去呈黑色;西方人是白色人种,虹膜中色素含量少,基质层中分布有血管,所以眼珠看上去呈浅蓝色。在眼角膜层与虹膜层之间的空间充满了透明的液状体,称为前房。瞳孔正下方就是由水晶体构成的水晶透镜,称为晶状体。它是一种胶质透

明体,由纤维细胞层组成,主要吸收短波长光。

最内层为视网膜,布满眼球后壁,它是眼球的感光部分,与视神经相连。视网膜是一种透明薄膜,包含3类不同生理特性的细胞。最外层由锥体细胞和杆状细胞组成,统称视细胞^[3];中间一层是双极性细胞,连接视细胞与神经节细胞;内层是神经节细胞,组成视神经与大脑相连。

视网膜分为以视轴为中心直径约6mm的中央区和周边区。在中央区有一个直径为2mm的区域呈黄色,称为黄斑区。位于黄斑区的中央有一个直径为1mm的小凹,称为中央凹。视细胞中的锥体细胞和杆状细胞的分布与作用都不同。中央凹没有杆状细胞,只含有锥体细胞,锥体细胞总数有600万~700万个,细胞密度高达15万/mm²。离开中央凹,锥体细胞的数目急剧减少,而杆状细胞急剧增加,杆状细胞最多的地方位于离开中央凹20°角区域,杆状细胞总数达1亿个。图1-2显示了右眼中视网膜上锥体细胞和杆状细胞的密度分布。视细胞中,锥体细胞既可以辨别光的强弱,也可以辨别色彩,因而具有较高的视觉分辨率,能够充分识别图像的细节。锥体细胞需要较强的光源,主要完成的是白天的视觉过程,所以又称为明视觉细胞。杆状细胞产生的视觉是单色的,分辨率较低,只能辨识视野中景物总的形象,无法辨别图像中的细节。它主要在黑暗的条件下起作用,对于景物只有黑白、浓淡之分,没有彩色差别,所以又称为暗视觉细胞。

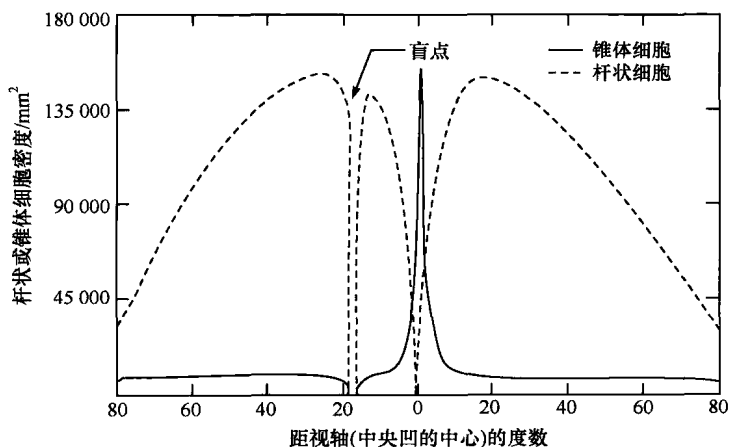


图 1-2 锥体细胞和杆状细胞在视网膜上的分布

整体来说,人眼的构造除了新陈代谢必需的一些功能部分,其余的部分均可以与现代社会中的照相机相对应。眼球外侧的六块肌肉通过移动使两只眼睛的视线指向某一特定位置,扫描更大的区域,获取更清晰的成像信息。照相机的自动调焦就是模仿了人眼的这一运动。

人眼工作时,物体反射的光线通过眼角膜,折射入前房,最后在水晶透镜的

聚焦作用下经过玻璃体成像在视网膜上的黄斑区。视网膜是响应光线感应的主体,其表层的锥体细胞和杆状细胞受到不同强度的光刺激,产生不同强度的电脉冲,通过与它们相连的双极性细胞经视神经系统传递到大脑的视觉皮层,大脑神经网络依据成像的位置和电脉冲强度,完成视觉信息的处理、加工与理解,感知到图像。

1.1.2 视神经系统

视神经系统的作用是将视网膜上得到的脉冲信号传送至大脑。信号主要传输路径如图 1-3 所示。视觉信息由人眼经视神经送出后,在视神经交叉处分为两路传到两个外侧膝状体,外膝体中继细胞所形成的视辐射与视皮质相连。大脑左右两个半球各自将信息通过视神经传送到外侧膝状体,然后再由上丘经视放线传送给大脑皮层的视觉区。

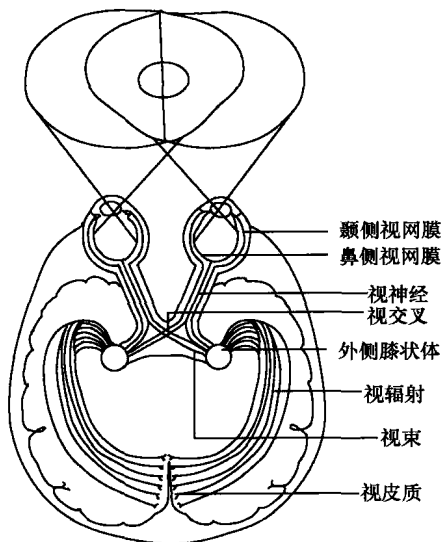


图 1-3 视神经通路

人眼视网膜的视细胞执行光—化学—电过程形成视觉信息后,将其传给双极细胞,继而再传给神经细胞,水平细胞和无足细胞存在于这两种细胞之间,支持双向输入输出,起反馈调节作用。视细胞输出连续电位,与光照相关,双极细胞和水平细胞的输入输出是与之形式不同的连续电位。无足细胞输入输出为瞬变性的连续电位,视神经细胞输出的是具有 0 和 1 两种状态的横幅度脉冲电位,其中频率由光强度决定。由此看来,视神经细胞以后的视觉信息都变成了电脉冲的形式。在大脑中枢神经中,以积分作用再将传来的密度调制脉冲信号变为模拟信号^[4]。

人类的视神经系统通过神经元的作用能够在 70~200ms 时间内完成感知和理解一幅复杂的图像。初级生物处理视觉感知问题也仅需要 18~46 步变换步骤,而在传统的计算机上执行算法达到认知图像,则需要进行上亿次变换。所以了解生物视觉的神经元机理对设计快速有效的计算机视觉系统很有指导意义。

1.1.3 处理

人类视觉具有对视觉信息获取、传输、处理和分析功能。视觉信息处理遵循自