

视线追踪

迟健男 王志良 张闯◎著



机械工业出版社
CHINA MACHINE PRESS

视 线 追 踪

迟健男 王志良 张 闯 著



机械工业出版社

本书以作者多年的创新性研究成果为主要内容,对视线追踪的各项关键技术进行了详细的阐述。全书共分九章,主要介绍了与视线追踪技术相关的人眼生理学和心理学知识及其发展历史,以及一些知名的视线追踪产品;各种视线追踪技术的原理和相关知识;图像预处理技术;视线参数检测;瞳孔跟踪;视线追踪系统标定技术;视线方向估计模型;视线追踪技术在各领域的应用。本书集合了先进性、系统性、理论性与实用性,是第一部视线追踪技术方面的专业著作。

本书适合大专院校、研究机构中从事相关领域研究的教师、研究人员使用,也可作为相关技术公司的技术参考用书。

图书在版编目 (CIP) 数据

视线追踪/迟健男,王志良,张闯著. —北京:机械工业出版社, 2011.6

ISBN 978-7-111-34689-0

I. ①视… II. ①迟…②王…③张… III. ①计算机视觉 IV. ①TP302.7

中国版本图书馆 CIP 数据核字 (2011) 第 088503 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:王欢 责任编辑:王欢

版式设计:霍永明 责任校对:李秋荣

封面设计:陈沛 责任印制:乔宇

北京机工印刷厂印刷 (三河市南杨庄国丰装订厂装订)

2011 年 6 月第 1 版第 1 次印刷

184mm×260mm · 16.25 印张 · 399 千字

0 001—3 000 册

标准书号:ISBN 978-7-111-34689-0

定价:49.80 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

社服务中心:(010)88361066

门户网:<http://www.cmpbook.com>

销售一部:(010)68326294

教材网:<http://www.cmpedu.com>

销售二部:(010)88379649

读者购书热线:(010)88379203 封面无防伪标均为盗版

序 言

视线追踪技术广泛地应用于人机交互、车辆辅助驾驶、心理与生理学研究、人因分析、虚拟现实等领域。视线追踪技术在国外开展较早，有着长期和深厚的技术积累，并形成了很多实际应用产品，但价格比较昂贵。我国在这一领域起步较晚、投入相对较少，研究基础相对薄弱，但目前我国视线追踪技术的研究越来越受到重视。

北京科技大学视线追踪技术课题组长期从事视线追踪技术的研究，其研究工作得到了国家自然科学基金和国家高技术研究发展计划（863 计划）的支持。本书的主要内容来源于课题组多年的研究工作及国内外相关文献（主要为国外文献），主要介绍了单摄像机系统和双摄像机系统的相关关键技术，具有一定的学术价值和实际应用价值。据作者检索，目前（成书前）本书为国内第一部有关视线追踪技术的著作。参考检索到的国外两部专著是《Eye Tracking Methodology Theory and Practice》（Andrew T Duchowski 著，2002 年 Springer 出版社出版）和《Passive eye monitoring: algorithms, applications and experiments》（Riad I. Hammoud 著，2008 年 Springer 出版）。通过作者的不懈努力，本书在内容的组织上更加专业，内容也更丰富。

本书的写作目的在于尽快将视线追踪相关先进技术介绍给国内的研究人员，推动视线追踪技术在国内的发展和应用，开发属于我国知识产权的视线追踪产品。

王志良

2011 年 5 月 28 日

前 言

视线追踪技术是利用机械、电子、光学等技术手段获取受试者“视觉注意”方向的技术，在认知障碍诊断和人机交互领域有着广泛的应用，如人机交互、车辆辅助驾驶、心理与生理学研究、人因分析、虚拟现实等。视线追踪技术的研究范畴涉及图像处理与分析、计算机视觉、光学、心理学与解剖学等众多领域的专业知识，技术手段需融合多种生物特征识别技术，是典型的多学科交叉性研究课题。按照系统构成和采用的检测方法不同，视线追踪技术可以粗略地划分为侵入式和非侵入式。随着相关技术的发展，基于数字视频分析的非侵入式视线追踪系统以其对人的干扰小、操作简便、精度较高等优点成为当前研究的热点方向。

视线追踪主要包括两部分关键技术，即视线参数检测和视线估计模型建立。作者长期从事视线追踪技术的研究，本书是以作者多年的创新性研究成果为主要内容，对视线追踪的各项关键技术进行了详细阐述。全书共分9章。第1章绪论，主要介绍了与视线追踪技术相关的人眼生理学和心理学知识。第2章视线追踪方法，回顾了视线追踪技术的发展历史，介绍了各种视线追踪技术的原理和相关知识。第3章典型视线追踪系统及构成，介绍了国外一些知名的视线追踪产品。第4章图像预处理，主要包括两部分内容，即基于反对称双正交小波的图像增强方法和基于边缘检测和双边滤波的彩色图像滤噪方法。这两种方法分别在VOG视线追踪系统采集的图像中有所应用，可提高图像质量。第5章视线参数检测，主要讨论主动红外光源系统的人眼视线特征提取方法，介绍了窄视野摄像机系统视线参数检测方法和宽视野摄像机系统视线参数检测方法。第6章瞳孔跟踪，是以瞳孔目标检测的准确性和快速性为目标，阐述了粒子滤波和卡尔曼滤波方法在瞳孔跟踪中的应用。第7章视线追踪系统标定，主要是针对双摄像机系统的光源和屏幕位置标定，介绍了基于平面镜反射的视线追踪系统标定方法和基于多摄像机全局标定的视线追踪系统标定方法。第8章视线方向估计模型，是视线追踪最为重要的关键技术，分别介绍了单摄像机平面视线参数到视线落点的映射模型和双摄像机系统视线估计几何模型。第9章视线追踪技术应用，详细讨论了视线追踪技术在各领域的应用。

本书由迟健男副教授统稿和主要编写工作。第1章由北京科技大学王志良教授编写。第2章、第3章、第7章、第9章由迟健男副教授编写，硕士研究生史光远、王翠娟、谢秀珍、胡涛、覃艳君、黄荣辉和博士研究生张琼参与了部分工作。第4章、第5章、第6章、第8章由迟健男和博士研究生张闯共同编写，博士研究生张鹏翼参与第8章的部分编写工作。

本书是国内第一部关于视线追踪技术的专业书籍，适用于大专院校、研究机构中从事相关领域研究的教师、研究人员和研究生参考使用。本书的相关工作得到了国家自然科学基金、国家高技术研究发展计划（863计划）的支持，在此表示感谢。由于作者的水平有限，书中难免有错误和不足，请读者批评指正。

迟健男

于北京科技大学

目 录

序言

前言

第 1 章 绪论	1
1.1 视觉注意	1
1.1.1 早期注意理论	1
1.1.2 近代注意理论	1
1.1.3 当前视觉注意理论	3
1.1.4 未来注意的研究	3
1.2 人眼系统的视觉机制	3
1.2.1 眼睛	3
1.2.2 视网膜	5
1.2.3 视束和视觉通道	6
1.3 视觉	6
1.3.1 颜色视觉	6
1.3.2 立体视觉	7
1.3.3 临时视觉	8
1.4 人眼运动	9
1.4.1 控制眼球运动的肌肉	9
1.4.2 平稳跟随运动	9
1.4.3 扫视	10
1.4.4 固视微动	10
1.4.5 眼球震颤	11
1.5 双眼视觉	11
1.6 视野与视敏度	11
参考文献	12
第 2 章 视线追踪方法	14
2.1 直接观察法	15
2.2 后像法	15
2.3 机械记录法	15
2.3.1 杠杆直接记录法	16
2.3.2 气动法	16
2.3.3 角膜吸附环状物法	16
2.4 电流记录法	16
2.5 电磁感应法	17
2.6 光学记录法	18
2.6.1 反光记录法	18
2.6.2 影视法	18

2.6.3 角膜反光法	19
2.6.4 双普金野法	20
2.6.5 虹膜-巩膜边缘法	21
2.7 本章小结	21
参考文献	22
第 3 章 典型视线追踪系统及构成	23
3.1 典型视线追踪系统	23
3.1.1 穿戴式(头盔式)视线追踪系统	23
3.1.2 非穿戴式(遥测式)视线追踪系统	29
3.2 典型的视线追踪系统构成	35
3.2.1 单摄像机无光源系统	35
3.2.2 单摄像机单光源或双光源系统	36
3.2.3 单摄像机多光源系统	38
3.2.4 多摄像机多光源系统	38
3.3 本章小结	41
参考文献	41
第 4 章 图像预处理	43
4.1 基于反对称双正交小波重构的图像增强方法	43
4.1.1 双正交小波	45
4.1.2 基于反对称双正交小波重构的图像多尺度边缘提取	48
4.1.3 基于反对称双正交小波变换的图像锐化增强	54
4.1.4 图像滤噪增强	55
4.1.5 实验结果与分析	58
4.2 基于边缘检测和双边滤波的彩色图像滤噪方法	69
4.2.1 细胞神经网络(CNN)	71
4.2.2 基于 CNN 的彩色图像分块自适应边缘检测	76
4.2.3 彩色图像滤噪	83
4.2.4 实验结果及分析	85
4.3 本章小结	87
参考文献	88

第5章 视线参数检测	90
5.1 基于主动红外光源的窄视野摄像 机系统视线参数检测	92
5.1.1 亮瞳和暗瞳图像判别及图像 差分	92
5.1.2 视线特征参数提取	93
5.1.3 实验结果	105
5.2 基于主动红外光源的宽视野摄像 机视线参数检测	107
5.2.1 暗瞳图像上基于 AdaBoost 算法的 人脸检测和人眼区域定位	108
5.2.2 投影法缩小人眼检测区域	113
5.2.3 瞳孔检测及瞳孔椭圆拟合	115
5.2.4 角膜反射（普尔钦斑）检测与中 心定位	117
5.2.5 实验结果与分析	119
5.3 本章小结	122
参考文献	122
第6章 瞳孔跟踪	124
6.1 基于卡尔曼滤波和均值漂移的瞳孔 跟踪方法	125
6.1.1 卡尔曼滤波	125
6.1.2 基于卡尔曼滤波和均值漂移的 瞳孔跟踪方法	128
6.2 基于粒子滤波的瞳孔跟踪方法	132
6.2.1 粒子滤波	132
6.2.2 基于伪彩色图的粒子滤波瞳孔 跟踪方法	140
6.2.3 实验结果与分析	149
6.3 粒子滤波和卡尔曼滤波组合的瞳孔 跟踪方法	154
6.3.1 粒子滤波跟踪	155
6.3.2 卡尔曼滤波瞳孔参数估计	159
6.4 本章小结	161
参考文献	162
第7章 视线追踪系统标定	164
7.1 摄像机模型	164
7.1.1 成像变换	164
7.1.2 摄像机的镜头畸变	165
7.1.3 摄像机模型	166
7.2 摄像机标定方法	170
7.3 双摄像机双光源视线追踪系统构成 ..	174
7.4 基于平面镜反射的视线追踪系统 标定	177
7.4.1 三维空间坐标测量	177
7.4.2 平面估计	178
7.4.3 确定屏幕平面及光源中心的 空间位置	179
7.4.4 实验结果	180
7.5 基于多摄像机全局标定的视线追踪 系统标定	180
7.5.1 单摄像机标定	182
7.5.2 双摄像机立体视觉系统标定	182
7.5.3 视线追踪系统的多摄像机全局 标定	182
7.5.4 系统标定结果与分析	187
7.6 本章小结	191
参考文献	192
第8章 视线方向计算模型	193
8.1 单摄像机主动红外光源系统视线方 向计算模型	194
8.1.1 基于非线性多项式和广义回归 神经网络的视线估计方法	194
8.1.2 基于非线性多项式和支持矢量 回归的视线落点估计方法	200
8.1.3 基于解析头动补偿的视线估计 方法	205
8.2 双摄像机主动红外光源系统的视线 方向计算模型	220
8.2.1 基于眼球成像模型的视线估计 方法 ^[6]	220
8.2.2 视线方向估计的 five-spot 模型 ..	223
8.3 本章小结	228
参考文献	229
第9章 视线追踪技术应用	231
9.1 计算机领域	232
9.1.1 选择性眼动系统	232
9.1.2 注视跟随系统	234
9.2 心理学与生理学检测	238
9.3 市场营销与广告	239
9.3.1 市场营销	239
9.3.2 网页可用性	239
9.3.3 电视广告的改进	240
9.3.4 印刷广告	241
9.3.5 商品包装设计	242

9.4 工业工程与人因分析	243	9.5 本章小结	248
9.4.1 航空	243	参考文献	249
9.4.2 车辆辅助驾驶	245		

第 1 章 绪 论

研究有关视线追踪问题，首先应该了解有关视觉注意、人眼运动、人眼的生理结构及人眼视觉特性等一些相关知识。视觉注意主要从心理学的角度研究人眼对视觉信息的加工过程，已有较长时间的研究历史。视线追踪技术是研究视觉注意的工具，视觉注意是指导视线追踪技术应用的基础。眼睛运动的目的是为了注视目标和提高视场的分辨率，无论人眼视线的方向如何，人们都努力看清视场中的细节并将注意力集中在目标区域或感兴趣区域。因此，如果能够追踪和记录人眼的运动，便可以跟踪观察者的视觉路径，进而了解到观察者感兴趣的内容，以及观察者是如何感知观察内容的^[1]。人眼的生理结构和人眼运动的特点在进行视线追踪技术研究必要的理论基础。

1.1 视觉注意

视觉注意的研究开始于 19 世纪末期，20 世纪 50 年代中期认知心理学的兴起带动了注意理论及视觉注意研究的广泛开展。目前，视觉注意依然是一个心理研究的热点课题。下面内容根据本章参考文献 [1, 2]，概述注意理论及视觉注意理论的发展过程。

1.1.1 早期注意理论

认知心理学不仅将心理过程看做是信息加工过程，而且在整体上强调了人的心理活动的主动性，所以注意研究成为认知心理学的一个重要领域。注意具有选择性、集中性等特征，由于认知心理学主要强调注意的选择性，即注意的作用过程，从这个角度出发，心理学家们提出了一些注意模型。

1958 年，英国著名心理学家布罗德班特提出了过滤器理论。他认为所有需要加工的信息首先要进入过滤器过滤掉无用的信息。然而，美国普林斯顿大学心理学系教授特瑞斯曼认为，未被注意的信息并未全部丢弃，只是信号被衰减，即部分信息得到进一步加工。基于这种观点，特瑞斯曼于 1964 年提出了注意的衰减理论。以上两种理论都认为注意过滤早于刺激再认，但 1963 年加利福尼亚大学圣地亚哥分校神经科学教授德尤奇等人提出了刺激再认先于注意过滤的理论，他们认为所有信息都被加工，只是在输出结果时才对重要刺激做出反应^[1]。

由此可见，过滤器理论和衰减理论属于早期选择观点，认为注意处于信息加工的知觉阶段，在识别前对信息进行选择。

1.1.2 近代注意理论

20 世纪中期心理学家提出的注意理论只是比较粗略地描述了注意的过程。直到 20 世纪 80 年代以后，一些心理学家提出新的注意模型才更为详细地解释了注意的过程。

1. 特征整合理论

特瑞斯曼等人在 1980 年提出了注意的特征整合理论。特征整合理论假设,对知觉对象的加工首先在特征水平上进行,然后以注意“胶水”将游离的视觉特征粘合在一起继续加工,最后获得关于对象的知觉。在前注意加工阶段,人们以平行的、自动的方式对于整个视域中各对象的所有特征同时进行加工,无需注意参与,更无容量限制;在注意加工阶段,当聚焦注意辨别目标时,聚焦注意就像“胶水”一样把分离的各个特征整合成一个完整的知觉对象。这时人们对刺激特征的加工是系列的,且有加工容量的限制。

视觉搜索在注意特征整合理论中占有重要地位。视觉搜索包括特征搜索和联合搜索。特征搜索是指当视觉搜索目标只有一个单一的、不被任何干扰子共享的物理特征时对目标的平行搜索。联合搜索是指如果干扰子多于两类且目标与每类干扰子共享一个特征时的搜索。联合搜索是在特征搜索的基础上进行的^[3]。

特征整合理论首次把刺激特征纳入目标刺激的检测过程,即刺激特征影响视觉搜索的结果,因此该理论对视觉注意理论的进展产生了巨大的影响。

2. 指导搜索模型

特征整合理论认为,特征搜索属平行加工,联合搜索属系列加工。然而,有些研究人员认为实际情况并不完全是这样,特征搜索也有系列加工,而联合搜索中也有平行加工。针对特征整合理论的不足,乌尔夫等人提出指导搜索模型。该模型认为在平行加工阶段,首先对有限基本视觉特征进行平行加工,每个特征维度都有一个分离特征地图,由特征地图产生一组激活值,综合每一特征维度的激活值,便可得到激活构图。系列加工阶段根据该激活构图中每个位置的激活值检测目标是否存在,低激活值的部分被排斥并不再重新检测,直到发现目标刺激或者发现无目标刺激为止^[4]。指导搜索模型描述了视觉注意的神经机制。

3. 相似性理论

乌尔夫的指导搜索模型详细地描述了视觉搜索的脑机制,还有一些理论宏观地描述视觉注意过程。

英国剑桥大学的认知脑科学学者邓肯等人于 1989 年提出了刺激相似性理论。该理论认为目标搜索是由目标与干扰子之间的相似性和干扰子彼此之间相似性所决定。目标-干扰子之间的相似性越高,干扰子彼此之间的相似性越低,目标搜索越困难,目标搜索越慢,反之亦然。刺激相似性理论简单地概括了视觉注意的规律,对后来的视觉注意研究产生了较大的影响^[5]。

4. 基于空间和基于物体的视觉注意理论

认知心理学关注的问题是选择注意是“基于空间”的,还是“基于物体”的。

“基于空间”的选择注意理论认为注意优先选择位置,然后对被选择位置上的刺激进行知觉加工。研究人员用基于空间的聚光灯式的注意模型,作为“基于空间”选择注意理论的基本框架。

“基于物体”理论则认为,选择注意基于物体本身进行。相关研究显示,被测试者对来自同一物体的两个靶子的反应并不比只对一个靶子的反应难,此时注意是将被注意的客体作为一个整体对待。相反被测试者对来自不同客体的靶子的反应则比较困难,这说明不同客体对集中注意产生了竞争。上述结果是在两个客体占有相同空间的基础上进行的,难以用基于空间的选择注意进行解释^[6]。

尽管有证据表明,视觉注意受组合因素的影响,但其中空间因素仍起特殊作用,组合因

素通过影响注意空间分配而产生特定影响^[7]。研究显示,空间是视觉注意特有的性质,当个体选择注意的客体时,可能忽略其位置。而在注意刺激位置时,却不会忽视该客体^[8]。

因此,视觉认知系统既可选择客体所在空间,也可选择被注意的客体本身。“基于物体”和“基于空间”的注意选择机制并非互不相容,而是相互补充的。到目前为止,很多学者已从神经生理学或认知神经科学研究中获取许多与认知心理学一致的结果,这些结论将会合理地解释注意选择性机制问题^[9]。

1.1.3 当前视觉注意理论

对于特征整合理论、指导搜索模型和刺激相似性理论等模型化视觉注意的规律理论,由于这些理论仅定性地描述了目标搜索的决定因素,所以这些结果在实际中难以操作。因此有必要量化视觉注意规律,于是最近有一些心理学研究人员建立了视觉注意的数学模型。例如,丹麦哥本哈根大学心理学系视觉认知中心的认知心理学教授邦德森试图把单个刺激再认、整体报告、检测和视觉搜索整合成一个一般视觉注意理论。邦德森等人认为视觉分类属平行加工。在视觉搜索时,尤其在联合搜索时,被测试者首先对呈现的刺激进行分类。邦德森等人的模型是有实际意义的^[1]。

1.1.4 未来注意的研究

近年来,选择注意的研究,例如注意的认知心理和神经生理效应研究、注意机制的解剖定位和功能研究等,取得了很大进展。在认知领域,关于不同种类注意之间的关系,如是否听觉空间注意影响视觉空间注意等问题也受到关注。现在的研究人员逐步认识到,选择注意加工所面临的主要问题是挑选特定刺激来控制特定行为。这是因为大脑在一定时间内只能完成一件任务,如支配眼睛或一只手朝向某个物体。上述认知心理学及神经心理学的各种问题,都是今后有待深入研究的。显然选择注意研究已从认知心理学角度逐步转向认知神经科学领域。未来的研究趋势是运用复杂的功能成像方法,促进心理学与神经科学的逐步整合^[2]。

1.2 人眼系统的视觉机制

人类约80%的信息是通过视觉系统获得的,在所有的感觉通道中,视觉起着最重要的作用。视觉系统的各种功能使人们能够感知外界环境物体的大小、形状、颜色等属性。视觉系统的主要功能是由视网膜和视觉中枢共同完成的。为了有效地获得视觉信息,眼睛的屈光系统把外界物体的像清晰地成像在视网膜上,视网膜上的光感受器把光信号转变为电信号,该信号通过视网膜上的神经回路逐级传递和处理,再由视神经传送至视觉中枢,最后分析形成视知觉。

1.2.1 眼睛

人眼可见光线只占全部电磁波波长中的一小段(即约400~700nm)。图1.1所示为眼球基本结构及视线方向。由图可见眼睛近似于球状,半径约为12mm。

眼睛的外层,依次可见角膜、巩膜、虹膜和瞳孔。角膜是一层透明的保护膜,它保护着

眼睛前端凸出的血管，覆在虹膜外面。虹膜中心有一个环状体，称作瞳孔，它的作用是通过持续改变大小来调节进入眼睛的光量。虹膜后面是晶状体，是眼球中的一种透明的双面凸体，位于虹膜与玻璃体之间，它可使穿过瞳孔进入眼球的光线聚集在视网膜上来形成图像。视网膜位于眼球的后面，它蕴含大量感光细胞。在角膜和晶状体之间是水样体，在晶状体和视网膜之间是玻璃体。光线在进入眼睛后穿过这一系列的光介质并在各个层面上经反射和折射后才最后到达视网膜。

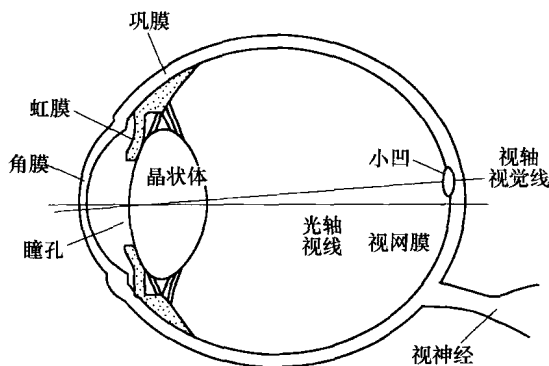
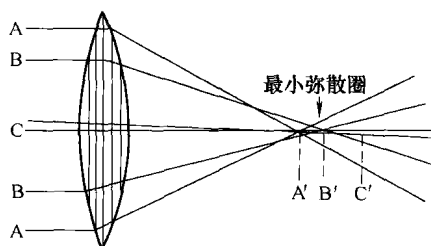


图 1.1 眼球结构图

在视网膜上有一个非常特殊的小区域，称之为小凹。上面集中了绝大部分对颜色敏感的细胞。并且，人对景物细节的感知也全依赖于它。小凹的位置并不是准确地位于由眼球和瞳孔中心所确定的光轴上。眼睛的光轴被称为视线 (Line of Gaze, LoG)，而把自小凹穿过瞳孔中心的线称为视觉线 (Line of Sight, LoS)。正是视觉线而非视线决定了人的视觉关注。如果可以估计视线或视觉线的方向并且掌握景物的信息，那么关注点就是由离视线和视觉线交叉点最近的景物来确定。

(1) 球面像差

光学透镜周边的屈光力量要比中央部强，因此经过光学透镜周边部分的光线比中央部分的光线形成的焦点更靠近透镜，这种现象称为球面像差。如图 1.2 所示，光 A 的焦点 A' 要比光 B 的焦点 B' 更靠近透镜。因此，光学成像系统后的任何焦点都不是一个锐利的点，而是一个小的光斑。



A'、B'、C' 分别为 A、B、C 的像；C' 为由接近光轴的 C 所成的像，离透镜最远；A' 为由透镜边缘的光所成的像，离透镜最近。

图 1.2 球面像差

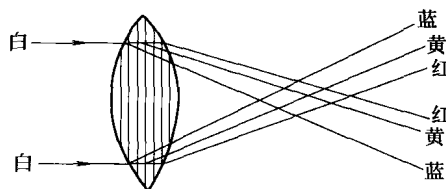
(2) 色像差

如图 1.3 所示，白光是由很多不同波长的光所组成。白光通过三棱镜可产生分光现象，成为红、橙、黄、绿、青、蓝、紫七色光谱。不同波长的光经过屈光介质时，短波光传播比较慢，因此在透镜内行程中的弯曲程度要比长波的光大，所以短波中的蓝光要在长波中的红光之前形成焦点。这种现象称为色像差，色像差可以降低视网膜成像的清晰度。

(3) 周边像差

由于眼睛受到某些光学因素的影响，使视网膜周边部的像总是不如中心凹成像清楚，这种现象称为周边像差。

然而，眼睛也通过各种机制来减少这些缺陷的影响。为了减少球面像差，虹膜被用来遮挡周边部位的大部分光线，所以球面像差影响并不明显。为了克服色像差，眼睛致力于制造中间波长的高质量



短波的蓝光先聚焦，长波的红光后聚焦

图 1.3 色像差

图像。为了匹配周边像差,视网膜的表面呈弯曲状,使得眼睛的光学系统非常接近于理想光学系统^[10]。

1.2.2 视网膜

视网膜的作用相当于照相机的感光底片或摄像机的光电转换器件,主要用来感光成像。当人们看东西时,物体的光线通过眼球光学系统,成像在视网膜上。

视网膜居于眼球壁的内层,是一层透明的薄膜,厚度约 200 ~ 300 μm ,是大脑的延伸部分,也是视觉形成的起点。视网膜由色素上皮层和神经感觉层组成。色素上皮层与脉络膜紧密相连,由色素上皮细胞组成。它们具有支持和营养光感受器细胞,并起到遮光、散热及再生和修复等作用。神经感觉层共分为 9 层,又称为视网膜的内 9 层,由外向内依次分别是视锥,视杆细胞层,外界膜,外颗粒层,外丛状层,内颗粒层,内丛状层,神经节细胞层,神经纤维层,内界膜。

(1) 视网膜色素上皮层 (Retina Pigment Epithelium, RPE)

由单层排列整齐的六面柱形细胞组成。每只眼睛约有 4 ~ 6 百万个 RPE 细胞。在眼底的不同区域, RPE 细胞的形态不同。后极部特别是黄斑区, RPE 细胞细长而均匀,色素较多。靠近锯齿边缘处的 REP 细胞较为短肥,色素亦逐渐减少。

(2) 椎体与杆体层

脊椎动物的光感受器在结构上由内段、外段、连接绒毛、体部和突触五部分组成。锥体与杆体层由光感受器的外段和部分内段组成。根据外段的形态不同,光感受器分为视锥细胞和视杆细胞。成年人每只眼睛的视锥细胞约有 6 百万个,视杆细胞约有 1.2 亿个。

(3) 外界膜

由光感受器之间、光感受器与 Muller 细胞,以及 Muller 细胞之间的粘连小带构成。

(4) 外核层

由光感受器的细胞体组成,含细胞核和细胞质。视盘鼻侧外核层较厚,有八九层细胞核,越靠近周边,细胞核层数越少。视盘颞侧周边的视网膜,外核层最薄,只有 4 层。黄斑中心凹处有 10 层,其他部位只有 5 层。

(5) 外丛状层

由光感受器伸出的轴突与双极细胞和水平细胞的突起相互连接的突触构成的疏松网状结构。视网膜的毛细血管至此为止,不再伸向外核层。黄斑的外丛状层较厚,该处的视锥细胞的轴突最长且走行方向倾斜,基本与外界膜平行,呈放射状的排列,称为 Henle 纤维层。黄斑以外的外丛状层逐渐变薄。

(6) 内核层

内核层有四种细胞,从外至内依次为水平细胞、双极细胞、Muller 细胞及无长突细胞。这些细胞的树突或轴突分别向上、向下伸展至内、外丛状层。而 Muller 细胞的凸起则分布于视网膜各层,起支持作用。

(7) 内丛状层

内丛状层较外丛状层厚,为内核层的双极细胞和无长突细胞与神经节细胞的树突连接而成的丛状结构。

(8) 神经节细胞层

主要由神经节细胞组成, 还有 Muller 细胞、神经胶质细胞及视网膜血管分支。在视网膜的绝大部分区域, 神经节细胞仅为一层, 而在视盘 (也叫视乳头) 的颞侧则为两层, 在黄斑区为 8~10 层。向黄斑中心凹方向, 神经节细胞明显减少, 中心凹处则几乎无神经节细胞。

(9) 神经纤维层

神经节细胞的轴突不分支, 其轴突沿视网膜平行行走而构成神经纤维层。除神经纤维外, 还有 Muller 纤维、神经胶质细胞和丰富的视网膜血管。

(10) 内界膜

由 Muller 细胞的基层膜及胶质细胞的突起组成, 厚约 $1 \sim 2 \mu\text{m}$, 随着年龄增大逐渐增厚。它靠近玻璃体的内表面, 十分光滑, 近视网膜一侧因 Muller 细胞凸起的延伸不同而起伏不平^[10]。

1.2.3 视束和视觉通道

视觉通道是由视网膜的视细胞接受外界的光刺激后, 将视信息传递到神经节细胞, 然后汇集到视神经, 由视神经进入颅内形成视交叉, 再由视交叉左右分开成为视束, 再由视束到视放射, 最后终止于大脑皮层枕叶的视中枢^[10]。

1.3 视觉

1.3.1 颜色视觉

视网膜位于眼睛的后方, 是构成视觉的重要部分。视网膜的主要功能是接收并向大脑传输图像信号, 而执行这些功能的细胞主要有两种: 视杆细胞和视锥细胞。在黄斑中央凹处只有视锥细胞, 无视杆细胞。在中央凹的边缘才开始有视杆细胞。再向外, 视杆细胞逐渐增多, 视锥细胞则逐渐减少。视杆细胞和视锥细胞主要完成感光功能, 其中视杆细胞的数量是视锥细胞的 18 倍。视杆细胞对弱光敏感 (可感受到一个光子), 主要负责在昏暗环境中产生暗视觉, 但只能辨别明暗, 不能分辨物体的细节和颜色。视锥细胞感受强光和颜色, 产生明视觉, 对物体细节和颜色分辨能力强。

视杆细胞和视锥细胞接收到的信息随后被传送给视网膜上的近 1 百万个节细胞。这些节细胞来自视杆细胞和视锥细胞的信息通过视神经发送到大脑。

人类每个眼球的视网膜内约有 6~7 百万个视锥细胞, 视锥细胞有三种, 它们分别含有对红、绿、蓝三种光敏感的感光色素。三种感光色素均由视黄醛和视蛋白构成, 其中视黄醛基本相同, 主要不同在于视蛋白结构中存在微小差异, 因而它们对色光的敏感性也存在差别。色光引起色觉, 这是一种复杂的物理和心理现象。根据三原色原理解释色觉的机制认为: 不同的色光作用于视网膜时, 三种视锥细胞产生了不同程度的兴奋, 这种兴奋信息经处理后转化为不同组合的视神经冲动, 传到大脑皮层就产生不同的色觉。例如, 红、绿、蓝三种视锥细胞兴奋程度的比例为 4:1:0 时, 产生红色色觉; 比例为 2:8:1 时, 产生绿色色觉。人眼可分辨波长在 380~760nm 的约 150 种颜色^[10]。

1.3.2 立体视觉^[11,12]

立体视觉是人类对三维空间各种物体的远近、凸凹和深浅的感知能力，是双眼视觉的高级形式^[11]。人眼为保证立体视觉的建立有三大机能和三级视功能。三大机能为感觉机能、运动机能和整合机能。三级视功能为同时知觉、融合功能和运动立体视觉^[12]。其中三大机能的具体功能如下：

(1) 感觉机能

感觉机能的主要功能是接收外界刺激，包括物体形状、颜色及影像刺激方位。感觉机能是人眼最原始、最基本的机能，也就是人眼具有看的机能。

(2) 运动机能

运动机能的主要功能是外眼肌运动协同作用能力，除此之外调节作用和瞳孔反应也属于运动机能的范畴，它们之间有着密切的关系。运动机能的作用是跟踪目标，使由双眼视轴焦点形成的注视方向始终停留在目标上，维持正常的视差。外眼肌的运动包括同向运动和异向运动两种。同向运动作用是跟踪远目标，此时双眼的运动是同时向相同的方向运动。异向运动的作用是跟踪不同距离的目标，双眼的运动方向是向相反的方向运动。由远向近注视的运动过程称为集合，而由近向远注视的运动过程称为散开。异向运动与调节和瞳孔反应有着密切的关系，它们被称为三联动。此外瞳孔反应还与光线的强度有关。

(3) 整合机能

整合机能的主要功能是在正常的运动机能作用下，将双眼所收集到的正确视觉信号，通过大脑的作用组合成完整的视觉信息^[12]。

为保证立体视觉，人眼的三级视功能分别如下：

(1) 同时知觉

视觉中枢能同时感受落在视网膜及双眼黄斑上的像刺激。这种同时知觉需要正常的感觉机能来保证。

(2) 融合功能

双眼看到的像能够准确重叠，即使在一定程度上改变双眼注视角度，在眼肌参与运动反应的条件下仍能准确重叠。这种功能需要正常的运动机能和整合机能来保证。融合功能分为三级融合：第一级重叠视，不相似的物体同时投射于两眼视觉相等的位置上，通过大脑的整合形成单一的完整影像；第二级融合视，单一影像同步反映于视觉知觉上，成为视轴目标，形成单一影像；第三级立体视，融合部分相似但不相同或部分完全不同的目标而产生立体视，形成三维空间。

(3) 运动立体视觉

在时间的参与下产生运动立体视觉^[12]。

人类用两只眼睛观察事物，双眼的共同活动提供了对方向和深度进行信息加工这一特殊功能的生理基础。在观察外界事物时，每只眼都形成单独的影像，双眼视像经大脑的信息加工处理。形成双眼立体影像。人类对物体方向的判断，是以双眼睛连线的中点处为依据。此点的正前方方向线是人类判断方向的依据，其他视网膜上相应点的视觉方向都以这个方向为基准来判定偏移方向和距离。形成立体视觉的基础是双眼单视，双眼单视与两眼看物体有着本质的不同。当两眼注视空间的某一点时，两眼视轴将集合于该点，同时该点作用于双眼视

网膜的相应点上, 并引起视觉兴奋, 产生双眼的单一视觉, 简称双眼单视。形成双眼单视必须符合以下条件:

- 1) 每一只眼睛必须有优良的单眼成像能力;
- 2) 两眼视场必须在一个大的范围内重叠;
- 3) 每个眼睛的视网膜上须形成大致相似的图像;
- 4) 视网膜上必须有生理相应点;
- 5) 神经与肌肉组织控制着眼睛运动及产生融合运动的反射活动。

眼睛时刻处于复杂机构的协调下, 保证具有正常视觉方向的视网膜总是接收同样的图像。立体视觉是双眼单视三级功能 (同时感觉、融合感觉、立体感觉) 中的最高级视功能, 它的形成是由于两眼对空间某一物体具有相对性双眼视差。在观察三维物体时, 由于两眼之间存在 65mm 左右的距离, 所以左右眼以各自的角度观察物体, 双眼视网膜上感受的刺激是不完全相同的, 这种视觉差异称为双眼视差。两眼视网膜不完全相应部位感受刺激并产生兴奋, 经视觉通路传到大脑皮层, 就产生了立体知觉。

如图 1.4 所示, 人眼注视平面上的一点 F , 在左右眼视网膜上的投影分别为 F_L 和 F_R , 它们分别位于两眼视网膜的对称点上。比点 F 近的一个点 B 在左右眼视网膜上的投影分别为 B_L 和 B_R , 它们分别位于两眼视网膜的非对称点上。 F 和 B 在左眼视网膜影像之间的距离 (B_L 和 F_L 之间的距离) 大于它们在右眼视网膜影像之间的距离 (B_R 和 F_R 之间的距离)。这就是 F 和 B 的相对双眼视差, 通常表示为两物体所形成的辐合角的差值, 即 $\angle B_L B R - \angle F_L F R$ 。点 F 的视差定义为零视差。比点 F (位于注视面上) 离观察者近的点 (如点 B) 相对于点 F 之间的视差称为交叉视差。反之, 如果一个点比注视平面上的物体离观察者远 (如点 A), 这时的视差称为非交叉视差。

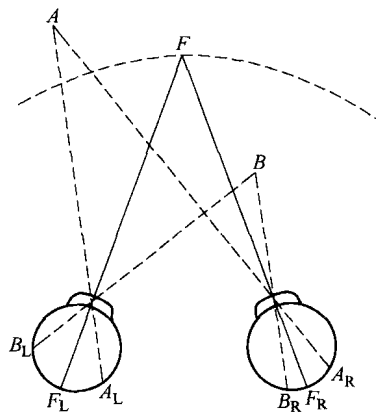


图 1.4 双眼视差示意图

由上述分析可知, 双眼视差不但有大小之分, 而且还有不同的种类, 即零视差、交叉视差、非交叉视差。双眼视差提供了物体之间的相对深度信息, 是产生立体感知的一种重要线索^[12]。

1.3.3 临时视觉

视觉的持续性说明了人类视觉运动不同的响应特征。视觉的持续性描述了人类视觉系统的临时采样频率。

视觉的持续性说明视网膜不能识别高频率采样的图像, 人眼对频率大约为 50 ~ 60Hz 的刺激闪光感觉是稳定的, 这个频率称为临界融合频率 (Critical Fusion Frequency, CFF)。CFF 解释了人眼对高频率显示图像响应的局限性。当图像以 24 帧/s 的速度播放时, 三叶快门提供的闪烁频率能达到 72Hz, 因此人们能够看清电影。电视以 60 张/s 的频率播放画面达到了临界融合频率, 类似采用三叶快门的电影。电视采取了隔行扫描技术, 典型的电视帧速率是 30 帧/s, 但是在每一轮扫描过程中只有奇数行或者偶数行能够在画面中显示出来。

人的眼睛并不是一种完善的记录工具，因为视网膜反应并不能随闪光开始而立即开始，也不能随闪光停止而立即停止。事实上，在刺激的开始和终止时，都存在着视网膜的时延。在断续的照明下，视网膜时延的存在，对人类视觉是有利的，使人们在闪烁照明下能够稳定地看清物体。如果人眼没有视网膜延时，视网膜的反应完全是实时的，那么在日常生活中交流电灯的照明下，人们会感觉任何物体都是闪烁的^[13]。

人们在电影银幕上看到的物体运动，其实是由影片上一系列略有区别的静止画面产生的，这种运动叫动景运动。在黑暗中，如果注视一个细小的光点，人们会感觉到它来回飘动，这叫自主运动。在皓月当空的夜晚，人们感觉月亮在“静止”的云朵后徐徐移动，这种运动感觉是由于云朵的飘动诱发产生的，因而叫诱发运动。在注视了倾泻而下的瀑布以后，如果将目光转向周围的田野，人们会觉得田野上的景物都在向上飞升，这叫运动后效。上述这些现象，人们看到的运动都不是物体的真正位移^[14]。

1.4 人眼运动

眼球的运动主要是由外眼肌动作支配眼球成像对焦机构来完成的，使目标能够正确落在视网膜的中心区域，再经过视传导路径在大脑皮质形成影像。眼球的运动可分为以下几个基本形态：平稳跟随运动、扫视运动、固视微动和眼球震颤。

1.4.1 控制眼球运动的肌肉

眼球上下左右方向的运动是由三对肌肉控制的，三对肌肉的协同活动能使眼球以角膜顶端后方 13.5mm 处为中心转动。每对眼肌控制眼球在一个平面上转动。这三对肌肉是，内直肌和外直肌、上直肌和下直肌、上斜肌和下斜肌。内直肌和外直肌收缩，眼球向内外方向转动；上直肌收缩，眼球向上内方向转动；下直肌收缩，眼球向下内方向转动；下斜肌收缩，眼球向上外方向转动；上斜肌收缩，眼球向下外方向转动。这三对肌肉的活动紧密配合，当眼球向一个方向移动时，除有关肌肉收缩外，其他肌肉，特别是相对称的肌肉必须同时放松，这样才能保证眼球的自由活动。例如，眼球向上方转动时，上直肌和下斜肌同时收缩，下直肌和上斜肌要相应地放松，同时内直肌和外直肌亦须做适当的放松配合。

1.4.2 平稳跟随运动

平稳跟随运动也称跟随运动，或眼追迹运动。这是一种慢速的眼动，所以也可叫慢速跟踪运动。跟随一个缓慢而平稳运动目标时的眼动，叫做平稳跟踪运动。此时眼球运动与目标物的运动之间，保持一种固定关系。

眼球跟踪运动是因注视对象运动的刺激而产生的，至于扫视运动则因注视对象与视线方向之间的背离而引起的。这两种运动成分产生的原因不同，其产生的生理结构也有区别。在观察对象静止的情况下，不出现跟踪运动。例如，读书或看静止的画面，都是由扫视运动实现的。

跟随运动的速度最高可达 30m/s，比之扫视运动是较慢的速度。如运动的物体超过此值，在眼睛追不上的情况下，为了补偿这种误差，则出现扫视运动。

跟踪运动是由前枕叶控制的。右侧枕叶负责向左跟踪，左侧枕叶负责向右跟踪，双侧枕