

智能制造系列教材

虚拟现实 与增强现实

FUNDAMENTAL
OF VIRTUAL AND AUGMENTED REALITY

鲍劲松 武殿梁 编著



清华大学出版社
北京

版权所有,侵权必究。举报:010-62782989, beiqinquan@tup.tsinghua.edu.cn。

图书在版编目(CIP)数据

虚拟现实与增强现实/鲍劲松,武殿梁编著. —北京:清华大学出版社,2023.10
智能制造系列教材
ISBN 978-7-302-63284-9

I. ①虚… II. ①鲍… ②武… III. ①虚拟现实—教材 IV. ①TP391.98

中国国家版本馆 CIP 数据核字(2023)第 059392 号

责任编辑:刘 杨

封面设计:李召霞

责任校对:赵丽敏

责任印制:丛怀宇

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-83470000 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:三河市春园印刷有限公司

经 销:全国新华书店

开 本:170mm×240mm 印 张:10.5

字 数:210 千字

版 次:2023 年 10 月第 1 版

印 次:2023 年 10 月第 1 次印刷

定 价:32.00 元

产品编号:088940-01

智能制造系列教材编审委员会

主任委员

李培根 雒建斌

副主任委员

吴玉厚 吴 波 赵海燕

编审委员会委员(按姓氏首字母排列)

陈雪峰	邓朝晖	董大伟	高 亮
葛文庆	巩亚东	胡继云	黄洪钟
刘德顺	刘志峰	罗学科	史金飞
唐水源	王成勇	轩福贞	尹周平
袁军堂	张 洁	张智海	赵德宏
郑清春	庄红权		

秘书

刘 杨

清华大学出版社

多年前人们就感叹,人类已进入互联网时代;近些年人们又惊叹,社会步入物联网时代。牛津大学教授舍恩伯格(Viktor Mayer-Schönberger)心目中大数据时代最大的转变,就是放弃对因果关系的渴求,转而关注相关关系。人工智能则像一个幽灵徘徊在各个领域,兴奋、疑惑、不安等情绪分别蔓延在不同的业界人士中间。今天,5G 的出现使得作为整个社会神经系统的互联网和物联网更加敏捷,使得宛如社会血液的数据更富有生命力,自然也使得人工智能未来能在某些局部领域扮演超级脑力的作用。于是,人们惊呼数字经济的来临,憧憬智慧城市、智慧社会的到来,人们还想象着虚拟世界与现实世界、数字世界与物理世界的融合。这真是一个令人咋舌的时代!

但如果真以为未来经济就“数字”了,以为传统工业就“夕阳”了,那可以说我们就真正迷失在“数字”里了。人类的生命及其社会活动更多地依赖物质需求,除非未来人类生命形态真的变成“数字生命”了,不用说维系生命的食物之类的物质,就连“互联”“数据”“智能”等这些满足人类高级需求的功能也得依赖物理装备。所以,人类最基本的活动便是把物质变成有用的东西——制造!无论是互联网、物联网、大数据、人工智能,还是数字经济、数字社会,都应该落脚在制造上,而且制造是其应用的最大领域。

前些年,我国把智能制造作为制造强国战略的主攻方向,即便从世界上看,也是有先见之明的。在强国战略的推动下,少数推行智能制造的企业取得了明显效益,更多企业对智能制造的需求日盛。在这样的背景下,很多学校成立了智能制造等新专业(其中有教育部的推动作用)。尽管一窝蜂地开办智能制造专业未必是一个好现象,但智能制造的相关教材对于高等院校与制造关联的专业(如机械、材料、能源动力、工业工程、计算机、控制、管理……)都是刚性需求,只是侧重点不一。

教育部高等学校机械类专业教学指导委员会(以下简称“机械教指委”)不失时机地发起编著这套智能制造系列教材。在机械教指委的推动和清华大学出版社的组织下,系列教材编委会认真思考,在 2020 年新型冠状病毒感染疫情正盛之时进行视频讨论,其后教材的编写和出版工作有序进行。

编写本系列教材的目的是为智能制造专业以及与制造相关的专业提供有关智能制造的学习教材,当然教材也可以作为企业相关的工程师和管理人员学习和培

训之用。系列教材包括主干教材和模块单元教材,可满足智能制造相关专业的基础课和专业课的需求。

主干教材,即《智能制造概论》《智能制造装备基础》《工业互联网基础》《数据技术基础》《制造智能技术基础》,可以使学生或工程师对智能制造有基本的认识。其中,《智能制造概论》教材给读者一个智能制造的概貌,不仅概述智能制造系统的构成,而且还详细介绍智能制造的理念、意识和思维,有利于读者领悟智能制造的真谛。其他几本教材分别论及智能制造系统的“躯干”“神经”“血液”“大脑”。对于智能制造专业的学生而言,应该尽可能必修主干课程。如此配置的主干课程教材应该是本系列教材的特点之一。

本系列教材的特点之二是配合“微课程”设计了模块单元教材。智能制造的知识体系极为庞杂,几乎所有的数字-智能技术和制造领域的新技术都和智能制造有关,不仅涉及人工智能、大数据、物联网、5G、VR/AR、机器人、增材制造(3D 打印)等热门技术,而且像区块链、边缘计算、知识工程、数字孪生等前沿技术都有相应的模块单元介绍。本系列教材中的模块单元差不多成了智能制造的知识百科。学校可以基于模块单元教材开出微课程(1 学分),供学生选修。

本系列教材的特点之三是模块单元教材可以根据各所学校或者专业的需要拼合成不同的课程教材,列举如下。

课程例 1——“智能产品开发”(3 学分),内容选自模块:

- 优化设计
- 智能工艺设计
- 绿色设计
- 可重用设计
- 多领域物理建模
- 知识工程
- 群体智能
- 工业互联网平台

课程例 2——“服务制造”(3 学分),内容选自模块:

- 传感与测量技术
- 工业物联网
- 移动通信
- 大数据基础
- 工业互联网平台
- 智能运维与健康管理

课程例 3——“智能车间与工厂”(3 学分),内容选自模块:

- 智能工艺设计
- 智能装配工艺

- 传感与测量技术
- 智能数控
- 工业机器人
- 协作机器人
- 智能调度
- 制造执行系统(MES)
- 制造质量控制

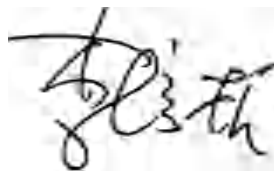
总之,模块单元教材可以组成诸多可能的课程教材,还有如“机器人及智能制造应用”“大批量定制生产”等。

此外,编委会还强调应突出知识的节点及其关联,这也是此系列教材的特点。关联不仅体现在某一课程的知识节点之间,也表现在不同课程的知识节点之间。这对于读者掌握知识要点且从整体联系上把握智能制造无疑是非常重要的。

本系列教材的编著者多为中青年教授,教材内容体现了他们对前沿技术的敏感和在一线的研发实践的经验。无论在与部分作者交流讨论的过程中,还是通过对部分文稿的浏览,笔者都感受到他们较好的理论功底和工程能力。感谢他们对这套系列教材的贡献。

衷心感谢机械教指委和清华大学出版社对此系列教材编写工作的组织和指导。感谢庄红权先生和张秋玲女士,他们卓越的组织能力、在教材出版方面的经验、对智能制造的敏锐性是这套系列教材得以顺利出版的最重要因素。

希望本系列教材在推进智能制造的过程中能够发挥“系列”的作用!



2021 年 1 月

清华大学出版社

制造业是立国之本,是打造国家竞争能力和竞争优势的主要支撑,历来受到各国政府的高度重视。而新一代人工智能与先进制造深度融合形成的智能制造技术,正在成为新一轮工业革命的核心驱动力。为抢占国际竞争的制高点,在全球产业链和价值链中占据有利位置,世界各国纷纷将智能制造的发展上升为国家战略,全球新一轮产业升级和竞争就此拉开序幕。

近年来,美国、德国、日本等制造强国纷纷提出新的国家制造业发展计划。无论是美国的“工业互联网”、德国的“工业 4.0”,还是日本的“智能制造系统”,都是根据各自国情为本国工业制定的系统性规划。作为世界制造大国,我国也把智能制造作为推进制造强国战略的主攻方向,并于 2015 年发布了《中国制造 2025》。《中国制造 2025》是我国全面推进建设制造强国的引领性文件,也是我国实施制造强国战略的第一个十年的行动纲领。推进建设制造强国,加快发展先进制造业,促进产业迈向全球价值链中高端,培育若干世界级先进制造业集群,已经成为全国上下的广泛共识。可以预见,随着智能制造在全球范围内的孕育兴起,全球产业分工格局将受到新的洗礼和重塑,中国制造业也将迎来千载难逢的历史性机遇。

无论是开拓智能制造领域的科技创新,还是推动智能制造产业的持续发展,都需要高素质人才作为保障,创新人才是支撑智能制造技术发展的第一资源。高等工程教育如何在这场技术变革乃至工业革命中履行新的使命和担当,为我国制造企业转型升级培养一大批高素质专门人才,是摆在我们面前的一项重大任务和课题。我们高兴地看到,我国智能制造工程人才培养日益受到高度重视,各高校都纷纷把智能制造工程教育作为制造工程乃至机械工程教育创新发展的突破口,全面更新教育教学观念,深化知识体系和教学内容改革,推动教学方法创新,我国智能制造工程教育正在步入一个新的发展时期。

当今世界正处于以数字化、网络化、智能化为主要特征的第四次工业革命的起点,正面临百年未有之大变局。工程教育需要适应科技、产业和社会快速发展的步伐,需要有新的思维、理解和变革。新一代智能技术的发展和全球产业分工合作的新变化,必将影响几乎所有学科领域的研究工作、技术解决方案和模式创新。人工智能与学科专业的深度融合、跨学科网络以及合作模式的扁平化,甚至可能会消除某些工程领域学科专业的划分。科学、技术、经济和社会文化的深度交融,使人们

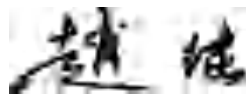
可以充分使用便捷的软件、工具、设备和系统,彻底改变或颠覆设计、制造、销售、服务和消费方式。因此,工程教育特别是机械工程教育应当更加具有前瞻性、创新性、开放性和多样性,应当更加注重与世界、社会和产业的联系,为服务我国新的“两步走”宏伟愿景做出更大贡献,为实现联合国可持续发展目标发挥关键性引领作用。

需要指出的是,关于智能制造工程人才培养模式和知识体系,社会和学界存在多种看法,许多高校都在进行积极探索,最终的共识将会在改革实践中逐步形成。我们认为,智能制造的主体是制造,赋能是靠智能,要借助数字化、网络化和智能化的力量,通过制造这一载体把物质转化成具有特定形态的产品(或服务),关键在于智能技术与制造技术的深度融合。正如李培根院士在丛书序1中所强调的,对于智能制造而言,“无论是互联网、物联网、大数据、人工智能,还是数字经济、数字社会,都应该落脚在制造上”。

经过前期大量的准备工作,经李培根院士倡议,教育部高等学校机械类专业教学指导委员会(以下简称“机械教指委”)课程建设与师资培训工作组联合清华大学出版社,策划和组织了这套面向智能制造工程教育及其他相关领域人才培养的本科教材。由李培根院士和雒建斌院士、部分机械教指委委员及主干教材主编,组成了智能制造系列教材编审委员会,协同推进系列教材的编写。

考虑到智能制造技术的特点、学科专业特色以及不同类别高校的培养需求,本套教材开创性地构建了一个“柔性”培养框架:在顶层架构上,采用“主干教材+模块单元教材”的方式,既强调了智能制造工程人才必须掌握的核心内容(以主干教材的形式呈现),又给不同高校最大程度的灵活选用空间(不同模块教材可以组合);在内容安排上,注重培养学生有关智能制造的理念、能力和思维方式,不局限于技术细节的讲述和理论知识的推导;在出版形式上,采用“纸质内容+数字内容”的方式,“数字内容”通过纸质图书中列出的二维码予以链接,扩充和强化纸质图书中的内容,给读者提供更多的知识和选择。同时,在机械教指委课程建设与师资培训工作组的指导下,本系列书编审委员会具体实施了新工科研究与实践项目,梳理了智能制造方向的知识体系和课程设计,作为规划设计整套系列教材的基础。

本系列教材凝聚了李培根院士、雒建斌院士以及所有作者的心血和智慧,是我国智能制造工程本科教育知识体系的一次系统梳理和全面总结,我谨代表机械教指委向他们致以崇高的敬意!



2021年3月

前言

PREFACE

当前,计算机、人工智能、新一代通信与传感器等技术加快提升了制造业数字化、智能化、网络化的水平,从根本上提高了工业知识产生和利用的效率。虚拟现实(VR)和增强现实(AR)技术作为新一代的人机交互界面,成为先进的制造技术,像机器人、3D 打印和物联网一样,正以创新的方式被使用,是智能制造重要的使能技术。随着智能制造的快速推进,VR/AR 技术正在重塑制造领域,在制造全生命周期发挥不可替代的作用,可以提高设计功能和性能,洞察制造过程使之透明化,使得人机协同更加高效。

VR/AR 的主要技术基础是计算机图形学,同时还涉及计算机其他学科、认知科学、电磁、机械等多学科。在智能制造工程类专业开设 VR/AR 课程是非常必要的。本书共分为 8 章,首先介绍 VR/AR 的发展历史和相关概念,然后按照制造系统中虚实融合的步骤展开:从场景的几何建模、虚拟场景搭建,到真实感渲染的基本流程、场景中的动画;再接着介绍人机交互技术,利用 VR/AR 和科学计算可视化进行集成等,最后给出常见的 VR/AR 开发方法。

全书主要部分由东华大学鲍劲松教授负责撰写,第 3 章由上海交通大学武殿梁副研究员撰写。本书涉及的图形学相关内容参考了美国斯坦福大学的课程 cs148、cs248、cs468、ee267,加利福尼亚大学圣塔芭芭拉分校的课程 cs184,加利福尼亚大学圣迭戈分校的课程 cse167,乔治梅森大学的课程 15462,北卡罗来纳大学的课程 comp575,布朗大学的课程 cs123,以及英国剑桥大学的课程 1819 等内容。另外,东华大学的刘天元、丁志昆、刘世民、卢山雨、江亚南、吕其彬、胡富琴等博士、硕士研究生参与了本书文献检索、公式校核和图表等编辑工作,魏静庵、丁志昆、卢山雨提供了本书的案例代码,在此表示感谢。同时,感谢上海交通大学的杨旭波教授,蔡鸿明教授,华中科技大学的王俊峰教授,同济大学的贾金原教授对本书提出的宝贵意见。

本书中涉及的相关随书代码、例程,可通过右侧二维码扫描获取。

由于时间仓促,作者知识浅陋,不足之处敬请广大读者指正。

作 者

2022 年 12 月于沪上

随书代
码网址

清华大学出版社

目录

CONTENTS

第 1 章 虚拟/增强现实技术基础	1
1.1 概念与发展历史	1
1.1.1 定义	2
1.1.2 VR/AR 的特点	3
1.1.3 VR/AR 的发展	5
1.2 VR/AR 的沉浸原理	5
1.2.1 深度暗示	7
1.2.2 自然人机交互	13
1.2.3 虚实融合	14
1.3 VR/AR 系统与组成	16
1.3.1 VR 系统	16
1.3.2 AR 系统	18
1.4 面向制造的 VR/AR 典型应用	21
习题	21
参考文献	21
第 2 章 三维几何表达与处理	23
2.1 三维模型数据及处理	23
2.1.1 概述	23
2.1.2 三维多边形网格	26
2.1.3 三维点云	34
2.2 常用的三维网格模型文件	36
习题	37
参考文献	37
第 3 章 三维虚拟场景	39
3.1 坐标系与坐标变换	39

3.1.1	坐标系统	39
3.1.2	坐标变换	41
3.2	三维场景图	42
3.2.1	一个简单三维场景	42
3.2.2	场景图数据结构	42
3.2.3	场景图的基本要素	43
3.2.4	场景图累积变换矩阵	49
3.3	AR 场景的虚实融合	50
3.3.1	虚实融合概述	50
3.3.2	相机标定技术	51
3.3.3	虚实场景注册技术与跟踪技术	58
3.4	场景优化	61
3.4.1	概述	61
3.4.2	模型优化	63
3.4.3	场景结构优化	65
3.4.4	场景加载优化	67
	习题	68
	参考文献	69
第 4 章	真实感渲染	70
4.1	图形渲染流程概述	70
4.2	模型、视图与投影变换	71
4.3	光栅化	71
4.4	着色	72
4.4.1	光源	73
4.4.2	光线表示	75
4.4.3	光照模型	76
4.5	纹理	83
4.5.1	纹理定义	83
4.5.2	纹理映射	84
4.6	使用 Unity 进行真实感渲染	84
	习题	85
	参考文献	85
第 5 章	运动仿真与动画	86
5.1	计算机动画概述	86

5.2	基于关键帧的仿真动画	87
5.2.1	关键帧动画基本方法	87
5.2.2	关键帧动画应用	88
5.3	基于物理的仿真动画	89
5.3.1	粒子	89
5.3.2	刚体运动	90
5.3.3	柔性体仿真	91
5.4	基于运动学的机构仿真动画	92
5.4.1	正向运动学计算	92
5.4.2	逆运动学计算	94
5.4.3	虚拟人运动仿真方法	94
	习题	95
	参考文献	96
第 6 章	工程数据可视化	97
6.1	数据可视化概述与基本流程	97
6.1.1	概述	97
6.1.2	可视化流程	99
6.2	科学计算可视化	100
6.2.1	有限元分析数据表示	100
6.2.2	标量场可视化	101
6.2.3	矢量场可视化	105
6.2.4	多物理场可视化	109
6.2.5	基于 VTK 的可视化	110
6.3	信息可视化	113
6.3.1	信息可视化概述	113
6.3.2	信息可视化方法	114
6.4	数据可视化与虚拟场景融合	123
	习题	125
	参考文献	126
第 7 章	人机交互技术	128
7.1	人机交互概述	128
7.1.1	概念	128
7.1.2	流程	129
7.1.3	技术发展	131

7.2	人机交互输入设备与原理	132
7.3	人机交互输出界面	132
7.4	人机交互关键技术	132
7.4.1	选择与操作	132
7.4.2	漫游	137
7.4.3	用户界面设计	138
7.5	案例：人机交互式虚拟维修训练	139
习题		141
参考文献		141
第 8 章	VR/AR 开发方法	143
8.1	VR/AR 开发概述	143
8.1.1	VR/AR 应用类型	143
8.1.2	系统工程方法	144
8.2	利用 Unity 进行系统开发	146
8.2.1	使用 PTC Vuforia 软件包	146
8.2.2	使用苹果 ARKit 软件包	147
8.2.3	使用谷歌 ARCore 软件包	148
8.2.4	使用微软 MRTK 软件包	148
8.3	利用 OpenSceneGraph 进行开发	149
8.4	基于 WebXR 的开发	149
习题		149
参考文献		149
附录	术语简表	150

第1章

虚拟/增强现实技术基础

虚拟现实和增强现实(VR/AR)是一种沉浸式、互动式体验,既包括人的感官体验,又包括人的认知体验,是基于三维数字仿真的可视化人机交互接口技术。VR/AR 技术的发展已逾 50 年,但直到最近几年才得到突飞猛进的发展。它们不仅在娱乐业领域得到了非常广泛的应用,在工业领域也有着广阔的应用前景。周济院士认为,智能制造系统是由相关的人、信息系统以及物理系统(HCPS)有机组成的综合智能系统,其中物理系统是主体,信息系统是主导,人是主宰^[1]。制造系统中人的作用不可或缺,VR/AR 被认为是智能制造的重要使能技术,将人与信息系统(HCS)、信息物理系统(CPS)和人与物理系统(HPS)有机融合起来,如图 1-1 所示。本章就按照这种融合关系,从 VR/AR 技术出发,介绍构建信息系统、人与信息系统的感知和认知,以及人与 HCS、HPS 和 CPS 的交互,展开基础理论、实现方法和应用开发介绍。

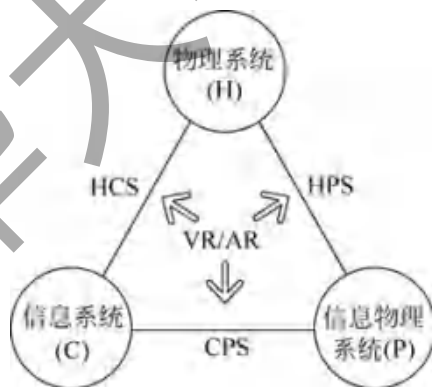


图 1-1 VR/AR 与 HCPS 系统的关系

1.1 概念与发展历史

VR/AR 在今天来说并不是新概念,两者在游戏、电影等领域已被广泛应用,被普遍认为是一种先进的可视化人机接口。但 VR/AR 到底是什么?

1.1.1 定义

虚拟现实(virtual reality,VR),也称虚拟技术、虚拟环境,早期译为“灵境技术”。VR的术语起源可追溯到德国哲学家康德(Immanuel Kant)提到的“Reality”,现代意义的VR术语则是由杰伦·拉尼尔(Jaron Lanier)在20世纪80年代提出的。从字面上看,“虚拟现实”一词本身就是矛盾的,奥卢大学(University of Oulu)的Steven M. LaValle在其2019年出版的专著*Virtual Reality*中指出,VR系统会使人保持一种知觉上的错觉,应该从人的心理、感知和认知角度来定义。在本书中,作者倾向于将虚拟现实定义为:完全利用数字化技术模拟产生三维虚拟世界,并使用立体显示设备和三维交互装置,为用户提供视觉、力觉等多感官的模拟,让用户沉浸在所营造的虚拟场景中,并与之互动的技术。

增强现实(augmented reality,AR)的术语最早由波音公司的Tom Caudell在1990年使用。增强现实可定义为融合了现实世界场景和虚拟仿真模型或信息,且现实世界中的物体被计算机生成的信息所增强,可实现与现实世界环境进行互动的一种体验技术。

值得注意的是,对VR/AR目前还没有被业界普遍接受的定义,并且目前虚实融合的人机交互接口领域还出现了新的概念,比如混合现实(mixed reality,MR)、扩展现实(eXtended Reality,XR)等。VR、AR和MR技术的界限可以通过1994年P. Milgram等给出的虚拟-现实界限图进行区分,如图1-2所示。

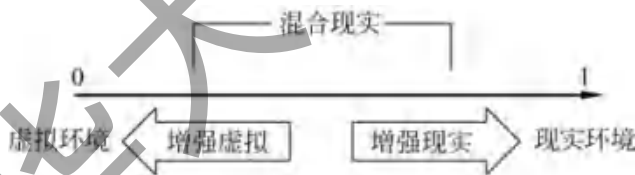


图 1-2 虚拟与现实的界限^[2]

在图1-2中添加一个数轴,虚拟环境和现实环境分别位于一条数轴的两端,越靠近1,越趋向于物理世界(现实环境),场景中物理世界的要素越多,虚拟要素越少。反之,越靠近0,则越趋向于虚拟世界(虚拟环境),场景中虚拟对象的比例就越多。

(1) 虚拟现实:看到的一切都是虚拟的,所有的要素都是数字化营造的假象。它是一种可以创建和体验虚拟世界的计算机仿真系统,即利用计算机生成一种模拟环境,场景是完全由计算机生成的三维虚拟环境,通过多源信息融合的/交互式的三维动态视景和实体行为的仿真,使用户沉浸到该环境中,让人的意识完全地进入虚拟世界,与现实环境相隔绝。VR位于图1-2中数轴的最左侧。

(2) 增强现实:在现实环境中增添由计算机生成的、可以交互的虚拟物体或信息,其混合度是由里面虚拟对象和物理对象的比例来决定的。可以看出,AR系统

在图 1-2 中靠近右侧。AR 将虚拟信息加入实际生活场景,就是将现实环境扩大了,是在现实场景中加入虚拟信息。例如,在汽车维修时,将需要维修的部件标注出来;汽车抬头显示器(HUD)可将车速、导航等信息投影(或反射)在挡风玻璃上,让驾驶员可以避免低头,这些都是典型的 AR 应用。

因此,通过虚拟环境和现实环境的距离,可以定性地度量虚拟和现实的混合方式,比如某系统更加“虚拟化”些。人类通常更多感知的是物理世界,在体验数字化过程中,“增强”根据接近虚拟世界和物理世界的程度,分为增强虚拟(augmented virtuality, AV)和增强现实。AV 将真实信息加入虚拟环境里,例如电玩游戏时可通过游戏手把感应重力,并且将现实中才有的重力特性加入游戏中,用来调整、控制赛车的方向。AR 则是将虚拟对象和信息添加到物理世界中,使物理世界得到“扩展”。将物理世界和虚拟世界混合在一起,统称为 MR,这是最近几年提出的名词,尽管某些文献论述中 MR 和 AR 有细微不同,在本书中作者不特意强调 AR 和 MR 的区别,都统一归为 AR。

1.1.2 VR/AR 的特点

1. VR 的特点——“3I”

虚拟现实环境是完全由计算机生成的三维虚拟环境,人如果融入到系统中,就非常强调系统具有沉浸感、逼真性,既要求有高的真实感、自然的交互方式,又要满足实时性的交互要求。因此,虚拟现实可以总结为有“3I”特点(分别为沉浸(Immersion)、交互(Interaction)、想象(Imagination)的三个首字母),如图 1-3 所示。

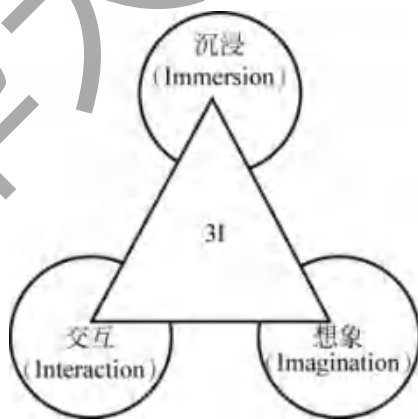


图 1-3 虚拟现实三大特点(3I)

(1) 沉浸: 沉浸是指用户感觉到完全置身于虚拟环境中,被虚拟世界所包围,觉得自己是虚拟世界中的一部分,使用户由被动的观察者变成主动的参与者,沉浸于虚拟世界之中,参与虚拟世界的各种活动。“沉浸”包括身体沉浸和精神沉浸两方面的含义。虚拟现实的沉浸性来源于对虚拟世界的多感知性,包括视觉感知、听觉感知、触觉感知,以及运动感知、味觉感知、力觉感知、嗅觉感知、身体感觉等。

(2) 交互：交互是指用户可与虚拟世界中的各种对象进行交互。在传统的多媒体技术中,人机之间主要是通过键盘与鼠标进行一维、二维的交互,而 VR 系统中人与虚拟世界之间则以自然的方式进行交互,人借助各种交互硬件设备,以自然的方式,与虚拟世界进行交互,实时产生如在现实世界中一样的感知。比如用户可以用手直接抓取虚拟世界中的物体,并可以感觉到物体的重量、软硬等,这种自然的人机交互极大地加强了用户的沉浸感。

(3) 想象：VR 为人更深入地认识世界提供了一种全新的接口和手段,使人突破时间与空间,去体验世界上早已发生或尚未发生的事情,“进入”宏观或微观世界进行研究和探索,从而完成某些因为条件限制而难以完成的事情。沉浸在虚拟世界中会激发人的想象力,尤其当多人参与在同一个虚拟场景中时。因此,VR 系统在汽车设计中得到了广泛应用。

2. AR 的特点——“3R”

AR 也要实现 VR 的“3I”,但是 AR 更强调虚拟世界和物理世界的融合,因此,其特点也围绕虚实融合来形成。对应“3I”,本书将 AR 的特点总结为“3R”,分别是虚实共融(Reaction)、增强(Reinforcement)和三维注册(Registration),如图 1-4 所示。

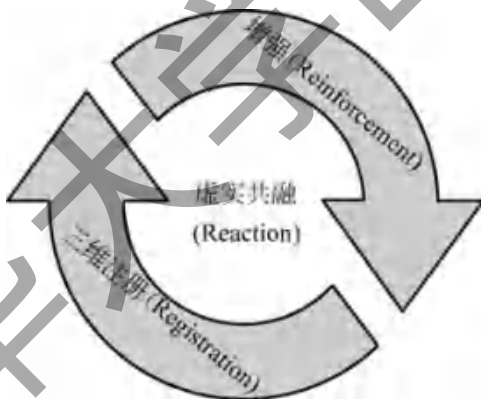


图 1-4 增强现实的三大特点(3R)

(1) 虚实共融：利用光学反射原理,将信息投射在镜片上,并经过平衡反射将影像投射入用户的眼睛,这样就可以将虚拟对象和真实环境对象融合在一起。为了获得以假乱真的虚实共融场景,三维虚拟模型需要是高度逼真的,同时能够对物理行为进行表达和对环境进行响应。通过人机交互从精确的位置扩展到整个环境,从简单的人面对屏幕交流发展到将自己融合于周围的空间与对象中。运用信息系统不再是自觉而有意义的独立行动,而是和人们的当前活动自然而然地成为一体。

(2) 增强：AR 系统通过融入虚拟对象或对物理世界添加虚拟的数字化标签,实现物理世界的增强。这种增强不仅可以增加物理世界的三维场景对象,还通过信息标注增强了对真实场景中对象的理解,将隐式的信息显现出来。

(3) 三维注册：实现虚实共融和环境增强，要让待增强的对象和虚拟对象之间精确匹配和精准融合，根据用户在三维空间的运动调整计算机产生的增强信息。

1.1.3 VR/AR 的发展

VR 术语的起源很早，可以追溯到德国哲学家康德。1938 年法国剧作家安托南·阿尔托(Antonin Artaud)将剧院描述为虚拟现实。而和现在意义一致的 VR 术语则是由贾龙·拉尼尔(Jaron Lanier)在 20 世纪 80 年代提出，他创建了 VPL Research 公司，对推广 VR 概念起到重要作用。截至目前，VR/AR 的发展已经超过 40 年，如图 1-5 所示，经历了启蒙、成长发展阶段，目前可认为基本成熟。



VR/AR 的发展历史

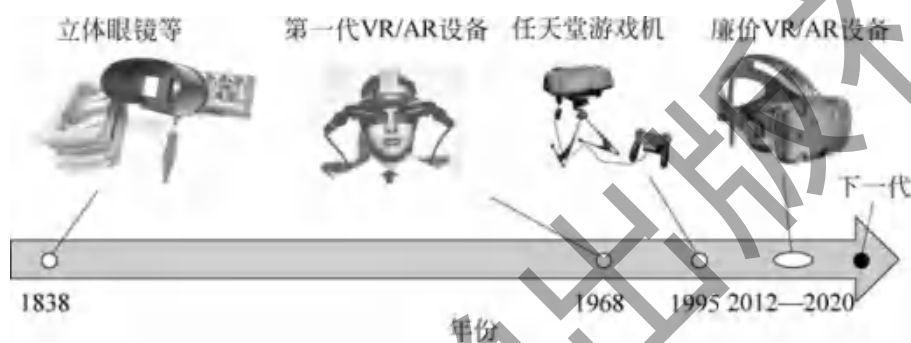


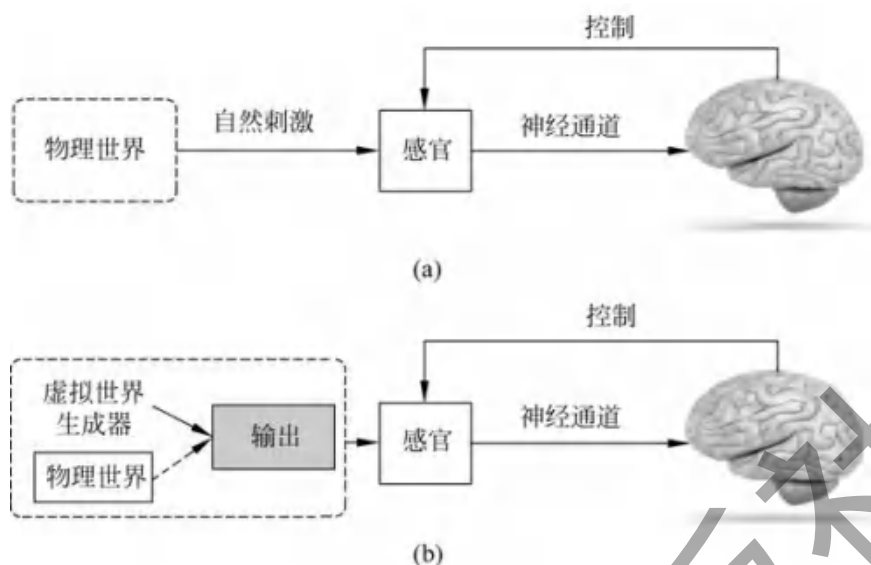
图 1-5 VR/AR 发展的重要时间节点

1.2 VR/AR 的沉浸原理

人类观察世界是立体的，真实世界和人完全融合在一起。但是人所看到的所谓三维世界都是三维世界在其视网膜上的二维投影而已，只不过这些二维投影包含了大量的三维信息，大脑通过二维投影来重建并理解三维世界。正常情况下大脑(和身体部位)控制着感觉器官(眼睛、耳朵、手指)，因为它们接受来自周围物理世界的自然刺激，如图 1-6(a)所示。

VR/AR 系统建立逼真的三维数字化虚拟环境，当人融入到增强的场景中时，强烈的沉浸感来源于真实环境下与虚拟对象自然的人机交互过程。如图 1-6(b)所示。计算机生成的虚拟世界如果足够真实，就会填充虚拟和真实世界之间的间隔，人的大脑就会被“欺骗”，认为虚拟世界其实就是周围的物理世界，电影黑客帝国中就描述了这种情形。因此 VR 和 AR 的首要问题就是要实现“沉浸”，尽管沉浸的定义是广泛且可变的，但是在此仅假设用户感觉自己像是模拟世界的一部分。虚拟环境能否真正地使用户身临其境取决于许多因素，如生成可信逼真的三维环境、产生深度感知暗示，其他通道(包括声音、力觉和触觉)感知、显示，自然的人机交互等。整个 VR/AR 系统的结构和层次，如图 1-7 所示。本书章节依其主线安排。

沉浸感知和幻觉并不限于视觉通道，表 1-1 显示了人类感官的分类。不同的

图 1-6 通常的感知过程和 VR/AR 感知过程^[3]

(a) 通常感知过程；(b) VR/AR 感知过程

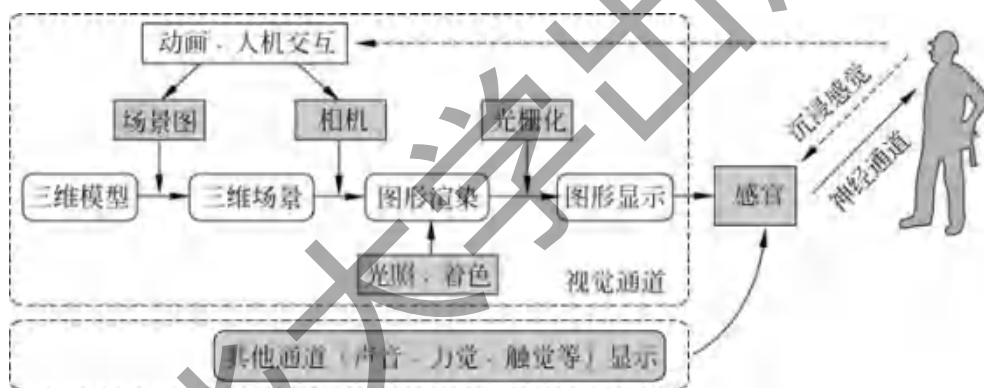


图 1-7 VR/AR 系统体系

刺激在不同的感官中将能量源转化为信号，对于人体意味着刺激被转化为神经脉冲。在人眼里，有超过 1 亿个感光体精确地感受可见光频率范围内的电磁能量。这些不同种类的光感受器可感知不同的颜色和光照度。其中近眼显示技术以沉浸感提升与眩晕控制为主要发展趋势。

表 1-1 感觉、刺激和接受分类

感 觉	刺 激	感 受 器	感 官
视觉	光电磁能量	图像传感器	眼
听觉	空气压力波	机械传感器	耳
触觉	组织扭曲	机械和热传感器	皮肤和肌肉
平衡感	重力和加速度	机械传感器	前庭
嗅觉/味觉	化学分解	化学传感器	鼻/舌头

听觉、触觉和平衡感涉及运动、振动或重力，是由人体机械感受器感受到的。

人的平衡感在前庭感知器中生成,可以帮助人体感知头部的方向,包括感知“向上”的方向。味觉和嗅觉被归为一类,称为化学感觉,它依赖于人体中的化学感受器。化学感受器根据人舌头上或鼻腔中出现的物质的化学成分提供信号。

1.2.1 深度暗示

让人感觉沉浸在计算机营造的环境中,首先要让人的心理确定是身处与真实世界一样的三维空间。感知心理学就是理解大脑如何将感觉刺激转化为感知现象的科学,需要研究物体看起来有多远,每秒多少帧才足以让人感知到运动是连续的,以及什么是存在感等问题。

人眼的视野很宽,水平方向约 220° ,垂直方向约 130° ,呈椭圆形,如图 1-8 所

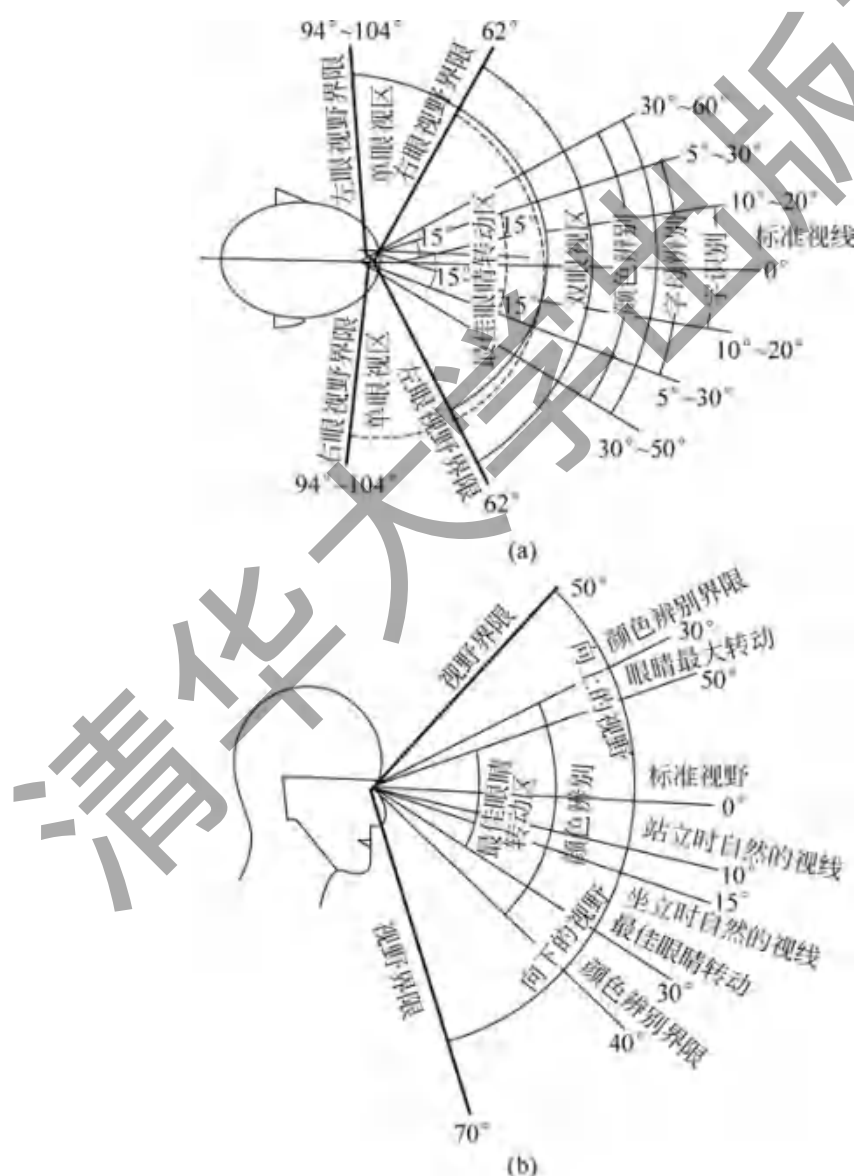


图 1-8 人眼的水平和垂视角

(a) 水平面内视野; (b) 垂直面内视野

示。但在通常的显示方式中,显示器均在视野之内,因此缺乏立体视觉的身临其境感。为此,增大显示器可以增强立体感。例如,宽银幕电影的立体感就比窄银幕的强,而全景电影由于没有画框,立体感更强。

现实世界中当人的头部旋转运动时,可以实现 360° 视野观察。在沉浸显示中,通过追踪人的头部旋转方向,来实时更新对应的显示画面,模拟人眼所看到的景物的变化。人具有深度感知的生理机能是对物理世界进行三维感知最重要的依据。VR/AR 采用的沉浸显示技术主要通过模拟人眼的立体视差、运动视差、视野范围来提供基本的视觉沉浸感,此外还可进一步通过模拟人眼聚焦、动态范围等方式来提高视觉沉浸感。

人类判断深度的方法可以分为两大类:一类是单眼深度暗示(monocular depth cues),指的是只依赖于单只眼睛作出判断;另一类是双眼深度暗示(binocular depth cues),或者称为立体深度暗示,指的是需要同时依赖于两只眼睛作出判断。

1. 单眼深度暗示

1) 线性透视

线性透视(linear perspective)是在平面上表现立体感的最有效的方法,在绘画艺术中被广泛采用。人通过线性透视可以在没有障碍物的情况下看到很远的距离。地平线是一条线,它将视野一分为二,上半部分是天空,下半部分是地面。由于透视投影,物体与地平线的距离直接对应于它们的距离。离地平线越近,感知的距离越远,如图 1-9(a)所示。而两根等长的线,远处的线看起来要长一点,如图 1-9(b)所示。

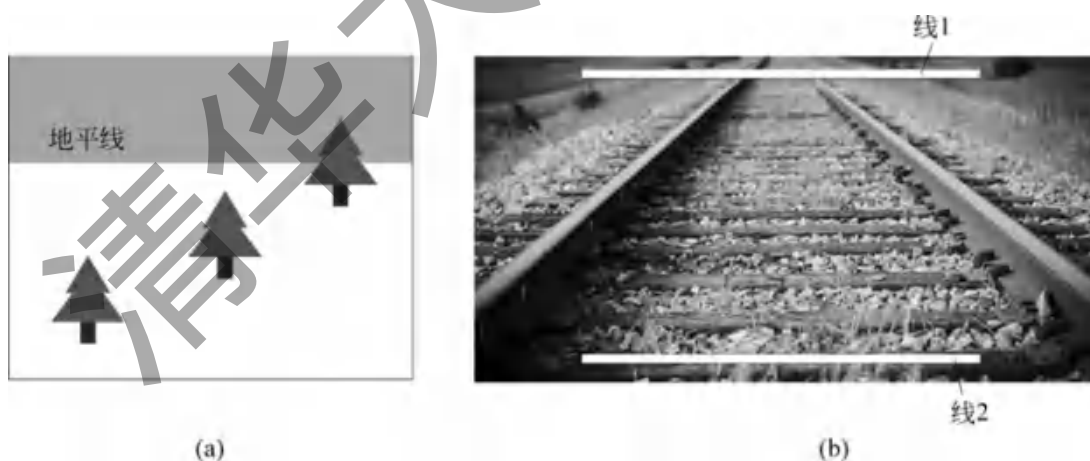


图 1-9 线性透视体现深度信息

2) 相对大小

当看到两个大小相似的物体时,你会判断它们的外观,感觉大的物体比小的物体更接近你,如图 1-10(a)所示。同样大小的物体,当观看距离不同时,在视网膜上成像的大小也不相同,距离越远,视网膜像越小,如图 1-10(b)所示。视线方向上平行线上对应的两点随着视距的增大,在视网膜上所成像点的距离线性减小。由此,

可通过比较视网膜像的大小来判断物体的前后关系。

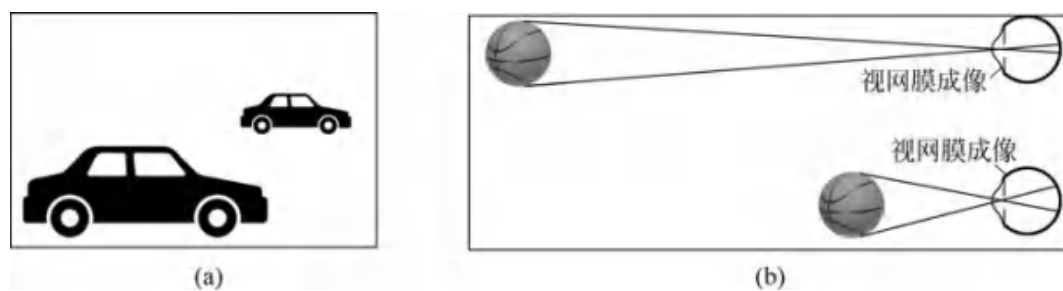


图 1-10 相对大小体现深度信息

3) 遮挡关系

当物体部分重叠时,与实际距离相比,后面的物体会显得最远。这种对深度的感知,使你有可能对相对的距离有一个直观的处置。当景物有相互遮挡时,也会产生深度暗示。如图 1-11 所示,球体、柱体和立方体在不同遮挡情况下将产生不同的立体视觉。

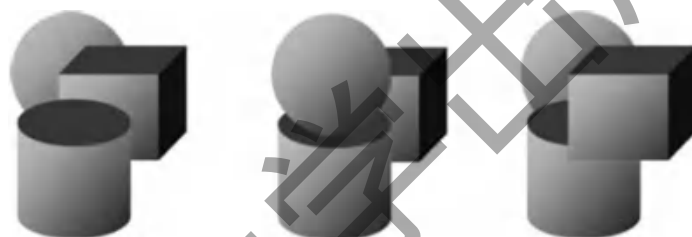


图 1-11 遮挡关系体现深度信息

4) 运动视差

运动视差(motion parallax)是人眼获得三维立体视觉感知的重要线索。如图 1-12(a)所示,当人在现实场景中左右移动时,所看到的景物会随之发生变化。当人和周围环境中的物体做相对平行运动时,远近不同的物体在运动速度和运动

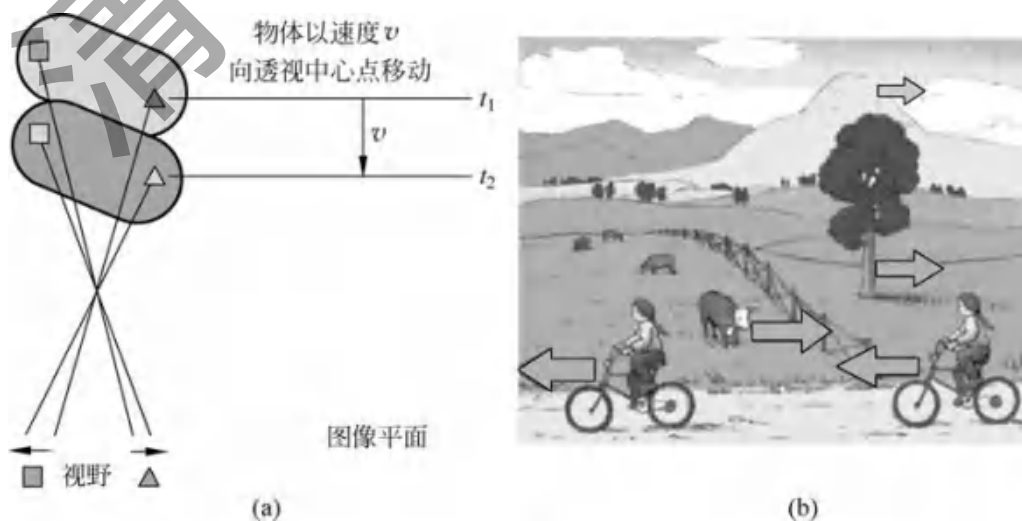


图 1-12 运动视差体现深度信息

方向上会出现差异,近处的物体看上去移动得快且方向相反,远处的物体移动得慢且方向相同,如图 1-12(b)所示。这是由于在同一时间内距离不同的物体在视网膜上运动的范围不同,近处物体的视角大,在视网膜上运动的范围大,而远处物体的视角小,在视网膜上运动的范围小,因而产生不同的速度和方向印象。在运动场景中根据对象的不同速度可以判断物体的远近,如图 1-12 所示。

5) 光和阴影

物体上光亮部分和阴影部分的适当分配可以改变或增强立体感,如图 1-13(a)所示。因此阴影所产生的深度感也是心理学上的重要暗示。例如,图 1-13(b)所示,在人行道上绘制的阴影,就有非常逼真的三维深度效果。



图 1-13 阴影体现深度信息

6) 纹理梯度

视野中物体在视网膜上的投影大小及投影密度上的递增和递减,称为纹理梯度(texture gradient)。当你站在一条用砖块铺成的路上向远处观察时,就会看到越远的砖块显得越小,这是因为远处部分在每单位面积上的砖块数量在网膜上的像较多。

7) 调整焦距

看不同远近的对象时,眼睛会一起移动,以便聚焦在近处的物体上,但又与远处的物体相距较远。当会聚发生时,眼睛必须先旋转,才能聚焦在一个物体上。这种聚焦的提示也可以帮助你确定物体有多远。看见近处一物体,有些模糊,所以睫状肌拉动晶状体调整焦距将眼睛聚焦上去,而改变晶状体的过程就叫调焦(accommodation)。如图 1-14 所示,看远处时晶状体压扁,看近处时晶状体拉长。为此,可以用景深(depth-of-field)来模拟眼睛调焦的效果。

8) 环境影响

由于大气层的影响,与较近的物体相比,远处的物体看起来不清晰或有些模糊。对于同一场景,远处的景物比近处的景物或多或少会更模糊,这样会产生深度暗示。景物越远,其发出的光线被空气中的微粒(如尘埃、烟、水气)散射越多,因



图 1-14 焦距调整体现深度信息

而显得越模糊。

2. 双眼深度暗示

双眼深度暗示主要包括两类：聚焦与视差。人眼注视物体时会进行调焦 (accommodative) 和聚散 (vergence)。如图 1-15 所示, 调焦就是调整晶状体大小把焦点对准当前的深度平面, 聚散则是眼睛内部转动让视线聚焦在某个点, 它们相互影响。收敛眼睛的聚散度会影响眼睛调节晶状体, 这时产生的深度暗示称为聚散暗示。

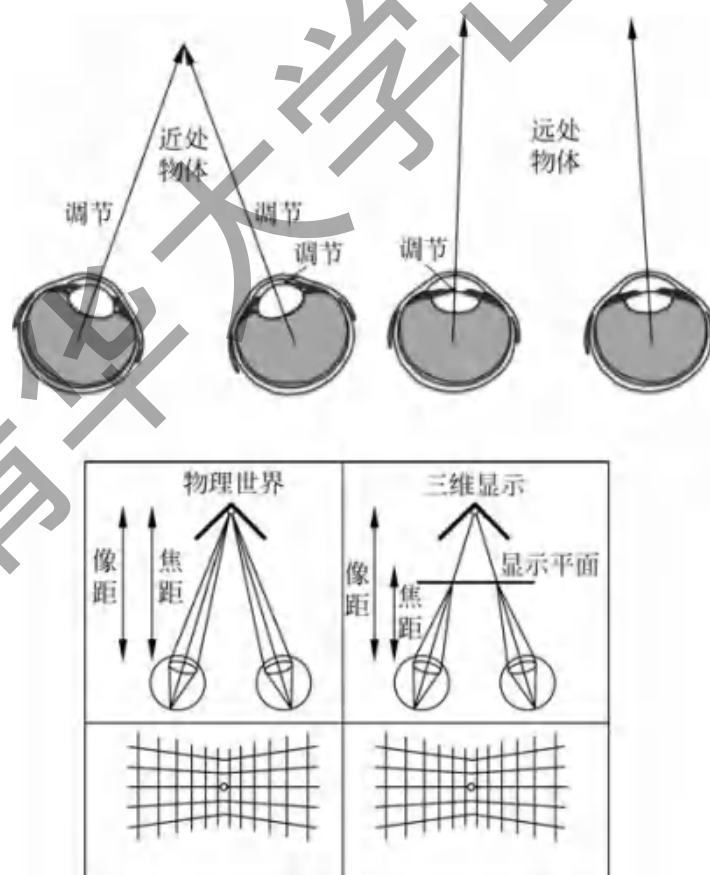


图 1-15 眼睛调焦的深度暗示

早在 1839 年,英国著名科学家温特斯顿(Winterston)就在思考一个问题——“人类观察到的世界为什么是立体的?”经过一系列研究发现:因为人长着两只眼睛。人的双眼大约相隔 6.5cm,观察物体(如一排重叠的保龄球瓶)时,两只眼睛从不同的位置和角度注视着物体,左眼看到左侧,右眼看到右侧。这排球瓶同时在视网膜上成像,而人的大脑可以通过对比这两幅不同的“影像”自动区分出物体的距离远近,从而产生强烈的立体感。引起这种立体感觉的效应叫作“视觉位移”。用两只眼睛同时观察一个物体时,物体上每一点对两只眼睛都有一个张角。物体离双眼越近,其上每一点对双眼的张角越大,视差位移也越大。

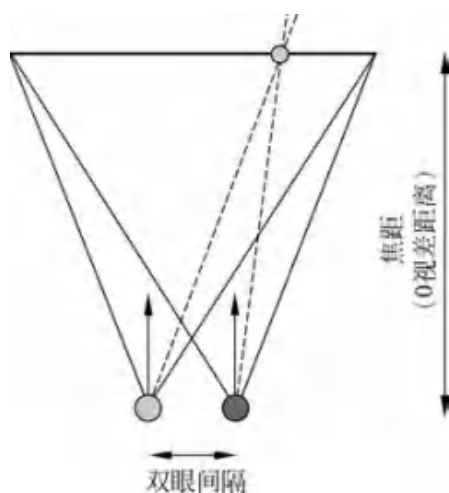


图 1-16 立体视差示意图

立体视差是人眼获得三维立体视觉感知的最重要线索。人眼在观察现实世界时,现实世界的光线在景物间产生反射、折射等现象,最终所形成的光线投射到眼底视网膜上成像,视神经将信号传输到大脑皮层的视觉处理区域,从而获得对景物的视觉感知。如图 1-16 所示,由于人的左右眼位置不同,景物在左右眼的视网膜上所投射的像也会有所不同。例如,在眼前竖起食指,并先闭上左眼,用右眼观察,然后闭上右眼,用左眼观察,你会发现食指和远处背景的相对位置,在左右眼看来会有明显不同,形成双目立体视差,由此产生不同深度的感觉。

立体眼镜发展已有近 200 年的历史。当前的沉浸技术是通过计算机图形图像技术来生成左右眼不同的画面,并通过立体显示技术,分别给左右眼同步展示不同画面,从而模拟人眼的立体视差效果。可通过直接给左右眼分屏来实现立体显示,通过左右眼的眼罩来保证左眼只看到左边屏幕的画面,右眼只看到右边屏幕的画面,由此带来立体深度的感觉。

立体视差是两眼图像的视差,根据聚焦点在投影屏幕的前后位置不同,可分为正视差、零视差和负视差,如图 1-17 所示。

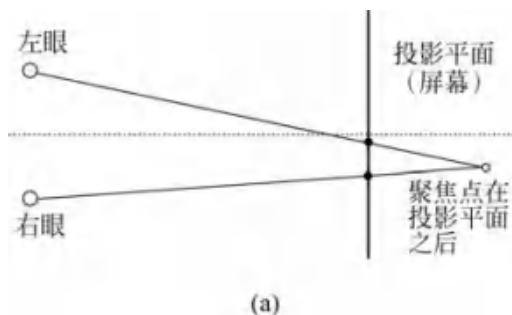


图 1-17 正视差、零视差和负视差

(a) 正视差; (b) 零视差; (c) 负视差

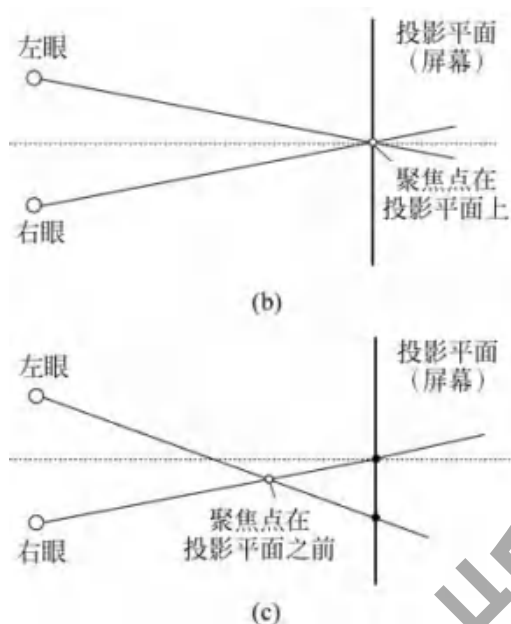


图 1-17 (续)

在视觉系统中常使用水平视差,不使用垂直视差。采用 toe-in(旋转)方法,会产生垂直视差,从而造成视觉上的不适感,如图 1-18(a)所示;而采用 off-axis 方法则可以避免垂直视差所导致的视觉不适感,如图 1-18(b)所示。

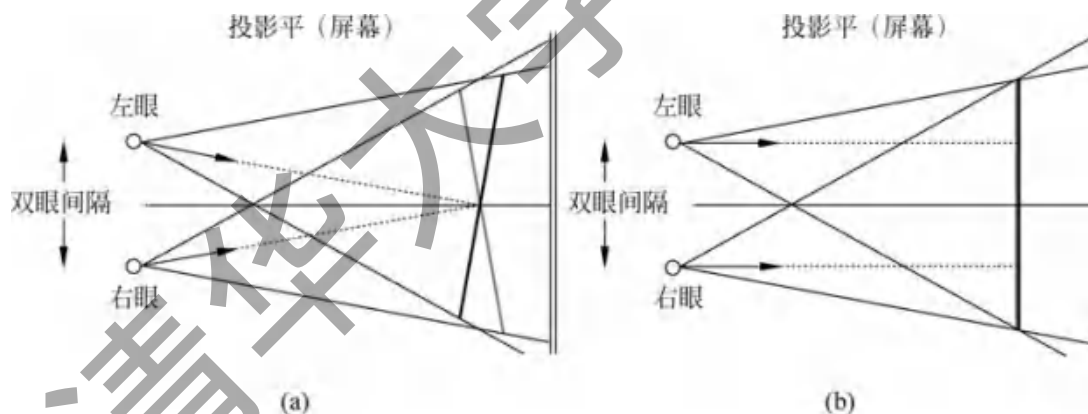


图 1-18 水平视差和垂直视差

(a) 垂直视差; (b) 水平视差

根据认知心理学分析,人们对不同的深度暗示,包括单眼和双眼的深度暗示,其敏感性不尽相同。

1.2.2 自然人机交互

对自然人机交互的衡量,并非取决于界面的交互模式,而是取决于用户自身的体验。一旦人充分相信数字环境是真实的,就能够以自然、直觉的方式与环境进行交互。人的手和身体本身就是极为灵活的工具,不需辅助就可以胜任许多通用型

的任务。同时,各式各样的工具造就了人类从事各种专门活动的能力,而每一种工具的设计和使用都尽最大可能为其支持的活动做了专门的优化,本身就是“自然人机交互”的典范。用笔写字实际上更好地利用了人多个手指灵活配合和控制的能力,大大提高了书写的准确性和丰富性。同样的道理,在今天的计算机使用中,特别是对于许多专门性的活动,一套设计合理的输入设备(例如数字笔)也可以延展人的能力,从而提供更加自然的界面。因此,“自然”并不是一个绝对的概念,无法抛开情境来下结论。自然人机交互使用手势、语音、触控,甚至是脑电波,这些交互的共性是不需要专门的输入工具,而是用身体的某一部分直接进行操作。

1.2.3 虚实融合

虚实场景融合是增强现实中的关键技术难点。虚拟物体和真实场景之间的融合,主要是解决虚实之间的几何一致性、时间一致性、遮挡一致性和光照一致性问题。不同于 VR 系统,在 AR 应用系统开发中面临的关键问题是如何解决真实场景和虚拟物体在几何、光照、时间和遮挡方面的一致性问题。

1. 几何一致性

场景图中的几何变换节点解决几何一致性问题。如图 1-19 所示。利用 AR 的三维注册技术,将虚拟物体与真实场景置于一个统一的坐标系中,并从同一个视点去观察,使虚拟物体和真实物体保持一致的透视变换和正确的遮挡关系。

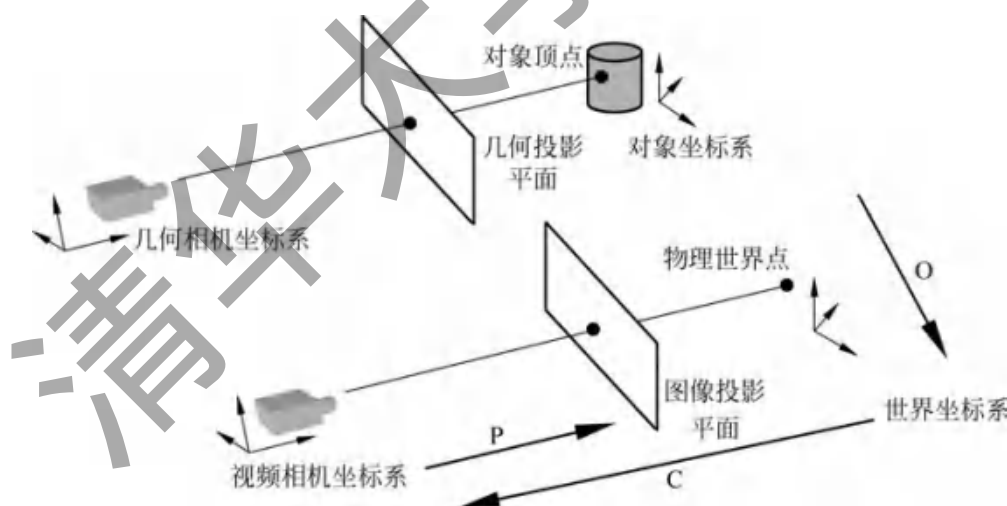


图 1-19 增强现实的坐标系统保证几何一致性^[5]

AR 系统须较好地解决增强现实注册定位问题,当用户转动和移动头部的时候,用户所感觉到的场景信息变化应是协调一致的。AR 系统不需要显示完整的虚拟场景,但需要分析大量的定位数据和场景信息,以保证由计算机生成的虚拟物体可以精确定位在真实场景中。