

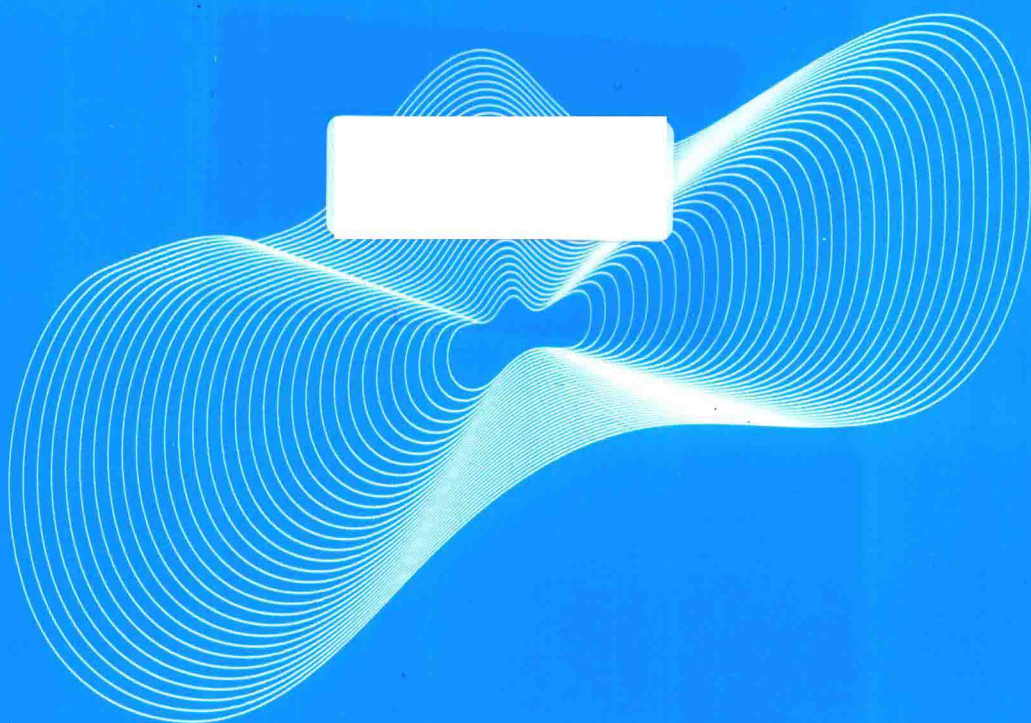


大数据和人工智能技术丛书

AUGMENTED REALITY:
TECHNICAL PRINCIPLES AND
APPLICATION PRACTICE

增强现实： 技术原理与应用实践

方维◎著



北京邮电大学出版社
www.buptpress.com



大数据和人工智能技术丛书

增强现实：技术原理与应用实践

方 维 著



北京邮电大学出版社
www.buptpress.com

内 容 简 介

本书围绕增强现实的技术原理和典型应用进行介绍,以帮助读者了解增强现实的理论方法,践行相关行业应用。本书以增强现实的关键技术为主线,对其涉及的计算机视觉、跟踪注册、3D建模、显示和自然交互等技术原理进行了介绍,并展示了增强现实在典型行业中的实际应用案例,最后对增强现实的发展趋势进行了展望。

本书适合对增强现实技术感兴趣的学生或技术人员阅读,可作为他们入门增强现实领域的指引之书,也适合在增强现实领域已有一定实践经验的从业者阅读,可帮助他们梳理知识体系和开拓创新思路。

图书在版编目 (CIP) 数据

增强现实:技术原理与应用实践 / 方维著. -- 北京:北京邮电大学出版社, 2022. 4
ISBN 978-7-5635-6608-2

I. ①增… II. ①方… III. ①虚拟现实 IV. ①TP391.98

中国版本图书馆 CIP 数据核字 (2022) 第 038255 号

策划编辑:姚 顺 刘纳新 责任编辑:王小莹 封面设计:七星博纳

出版发行:北京邮电大学出版社

社 址:北京市海淀区西土城路 10 号

邮政编码:100876

发行部:电话:010-62282185 传真:010-62283578

E-mail: publish@bupt.edu.cn

经 销:各地新华书店

印 刷:保定市中国画美凯印刷有限公司

开 本:787 mm×1 092 mm 1/16

印 张:13.5

字 数:336 千字

版 次:2022 年 4 月第 1 版

印 次:2022 年 4 月第 1 次印刷

ISBN 978-7-5635-6608-2

定 价:42.00 元

· 如有印装质量问题,请与北京邮电大学出版社发行部联系 ·

前 言

信息技术已从固定的桌面计算转变到以手机、平板电脑为代表的移动计算，这将传统在办公室或书房进行的计算工作转变为能随时随地进行的活动，因此将不同情景体验融入我们的真实世界中变得越来越重要。

增强现实（Augmented Reality, AR）技术的出现为上述情景计算提供了可能，其可将真实世界和与之相关的虚拟信息直接关联，通过模拟的视觉、听觉、触觉等感受，给人提供超越真实世界认知的体验。增强现实技术有望改变人们的生活和生产方式，提供实时且无限的信息来源，构建彼此相互交流、沟通和理解的新范式。

增强现实技术是一门涉及多门学科的技术，其知识体系仍在不断完善中，很难用逻辑概括的方式来说明它包含的所有内容。本书旨在描述增强现实的技术原理与应用实践。作者的目的是让这本书既能服务于科学研究人员，又能服务于对增强现实应用感兴趣的从业者，特别是工程师。读者提前了解计算机视觉以及计算机图形学领域的相关知识，会对理解本书内容有所帮助。由于篇幅的限制，作者无法进一步提供背景技术的特定细节，读者可参考书中文献。

全书分为 4 部分。

第 1 部分由第 1 章组成，介绍了增强现实相关的基本概念，如增强现实技术的特点，其与虚拟现实的区别，以及增强现实的发展历史和增强现实系统的构成，使得读者对增强现实技术有更全面的认识。

第 2 部分由第 2~6 章组成，共 5 章，主要围绕增强现实的技术原理进行了介绍。其中：第 2 章简要介绍增强现实中的计算机视觉基础，如相机成像模型、相机标定、图像特征检测算子、空间位姿解算等；第 3 章介绍了增强现实中的跟踪注册技术，主要围绕基于传感器、基于视觉和基于视觉-惯性传感器融合的跟踪注册技术进行描述；第 4 章介绍了基于双目结构光的 3D 建模和基于多视图的 3D 场景重建；第 5 章介绍了现有增强现实中不同显示装置的原理和技术特点；第 6 章介绍了增强现实的自然交互技术。

第 3 部分由第 7~9 章组成，共 3 章，结合现有的典型软/硬件系统，对增强现实的应用实践部分进行了介绍。其中：第 7 章介绍了增强现实在工业制造中的典型应用，如在航空航天、汽车和智慧物流等工程领域的应用；第 8 章介绍了增强现实在国防军事、医疗健康和教育等行业的应用实践；第 9 章对现有的典型增强现实的软/硬件系统进行了介绍，

并给出了简要的开发实践案例，但由于篇幅有限，本书不能一一描述所有优秀的增强现实软/硬件系统。

第4部分由第10章组成，展望了增强现实技术的未来发展趋势。该部分主要结合当前信息科技的发展，对增强现实与AI、5G、空间计算和数字孪生等技术的融合前景进行了展望，其将有助于进一步推进增强现实的普适性应用。

当前，虽然距离制造出一款可被不同行业广泛接受并认可的增强现实设备，我们还需要时间来应对并解决所面临的问题，如移动AR应用中跟踪注册技术的完善、即时交互稳定性的保障和5G基础设施的建设等，但这条道路是清晰且明确的。可以预见，在不远的将来，增强现实将会融入人们工作和生活的方方面面，并促使人们变换新的方式来思考问题，丢弃那些自认为宝贵的想法。

作者多年来一直从事于视觉感知、增强现实技术及其相关应用的研究。本书是作者在多年研究的基础上，结合当前增强现实技术的行业发展而写成的。本书获得了国家自然科学基金（编号：52105505）和北京市自然科学基金（编号：3204050）的资助，在此表示感谢。

本书的出版得益于许多友人的支持与鼓励，在此感谢以下各位为本书提供的见解和建议：杨奎、李根、安泽武、陈黎茜、胥小宇、王泽瑜、李张文驰。

感谢北京邮电大学出版社为本书提供了很多修改和建议。

最后，感谢我的家人在书稿撰写过程中给予的巨大支持和耐心。

由于作者水平有限，书中难免有疏漏，恳请广大读者批评指正，如蒙读者不吝告知（作者邮箱：fangweichn@163.com）将不胜感激。

方 维

目 录

第 1 章 增强现实概论	1
1.1 增强现实技术概述	1
1.1.1 增强现实的定义	1
1.1.2 增强现实技术的特点	2
1.1.3 增强现实与虚拟现实的区别	2
1.2 增强现实的发展历程	3
1.2.1 起源阶段	3
1.2.2 初步发展阶段	4
1.2.3 快速发展阶段	7
1.3 增强现实系统的构成	10
1.3.1 增强现实系统的基本组成	10
1.3.2 增强现实系统的关键技术	11
本章小结	12
本章参考文献	13
第 2 章 增强现实中的计算机视觉基础	15
2.1 相机成像模型	15
2.2 相机标定	17
2.2.1 基于标定物的标定方法	17
2.2.2 基于主动视觉和自标定的标定方法	20
2.2.3 相机畸变矫正	20
2.3 图像特征检测算子	21
2.3.1 图像特征点提取	21
2.3.2 图像特征描述	26
2.3.3 线特征检测	29

2.3.4	特征匹配及外点剔除	31
2.4	空间位姿解算	33
2.4.1	三维空间变换	33
2.4.2	2D-2D 位姿估计	34
2.4.3	3D-2D 位姿估计	36
2.4.4	3D-3D 位姿估计	37
本章小结		38
本章参考文献		38
第3章	增强现实中的跟踪注册技术	41
3.1	增强现实中的跟踪注册要求	41
3.2	基于传感器的跟踪注册技术	43
3.2.1	磁场跟踪	43
3.2.2	机械式跟踪	43
3.2.3	声学跟踪	44
3.2.4	惯性跟踪	44
3.3	基于视觉的跟踪注册技术	45
3.3.1	基于人工标识的跟踪注册	45
3.3.2	基于模型的跟踪注册	48
3.3.3	基于自然特征的跟踪注册	50
3.4	基于视觉-惯性传感器融合的跟踪注册技术	52
3.4.1	基于滤波的视觉-惯性传感器融合的跟踪注册	53
3.4.2	基于自适应滤波的实时平滑跟踪注册	62
3.5	实例验证	66
本章小结		68
本章参考文献		68
第4章	增强现实的视觉3D建模技术	72
4.1	3D建模技术概述	72
4.2	基于双目结构光的3D建模技术	73
4.2.1	自动3D模型获取方法概述	74
4.2.2	双目结构光视觉测量系统参数构建	76

4.2.3	SFS 被动视觉预测	81
4.2.4	增强现实中的自动 3D 模型重建策略	83
4.2.5	实例验证	88
4.3	基于多视图的 3D 场景重建	91
4.3.1	基于多视图的 3D 场景重建概述	91
4.3.2	逆深度自适应加权的三角化模型	92
4.3.3	实例验证	95
	本章小结	99
	本章参考文献	99
第 5 章	增强现实的显示技术	101
5.1	视觉感知的基础知识	101
5.1.1	视觉方式	101
5.1.2	视觉结构的基本概念	102
5.2	显示装置的基本特点	104
5.2.1	成像显示技术	104
5.2.2	视场	106
5.2.3	分辨率与刷新率	107
5.2.4	亮度和透光率	108
5.2.5	延迟	108
5.2.6	畸变	109
5.2.7	人因工程学	109
5.3	增强现实显示器	110
5.3.1	头戴式显示器	110
5.3.2	手持式显示器	111
5.3.3	固定式显示器	112
5.3.4	投影式显示器	113
5.4	光学透视式显示	114
5.4.1	棱镜方案	114
5.4.2	Birdbath 方案	115
5.4.3	自由曲面方案	115
5.4.4	光波导方案	116

本章小结·····	120
本章参考文献·····	120
第 6 章 增强现实的自然交互技术·····	122
6.1 人机交互概述·····	122
6.2 触控交互·····	123
6.2.1 电阻式触摸屏·····	124
6.2.2 电容式触摸屏·····	125
6.2.3 红外触摸屏·····	126
6.2.4 表面声波式触摸屏·····	126
6.2.5 触控交互总结·····	127
6.3 语音交互·····	128
6.3.1 语音交互概述·····	128
6.3.2 语音交互原理·····	129
6.3.3 语音交互局限·····	131
6.4 手势交互·····	132
6.4.1 手势交互概述·····	132
6.4.2 手势识别方法·····	132
6.4.3 典型的手势识别方法·····	135
6.4.4 手势交互总结·····	137
6.5 眼动交互·····	138
6.5.1 眼动交互概述·····	138
6.5.2 眼动交互的输入方式·····	140
6.5.3 眼动交互的眼动指标·····	141
6.5.4 眼动仪·····	142
6.6 脑机接口交互·····	142
6.7 多模态交互·····	143
6.8 多用户协同交互·····	145
6.8.1 多用户协同交互特征·····	145
6.8.2 共享空间协同交互设计·····	147
本章小结·····	148
本章参考文献·····	149

第 7 章 增强现实与工业制造	152
7.1 航空航天工程	152
7.2 汽车工程	155
7.3 智慧物流工程	158
7.4 其他工业领域	161
本章小结	161
本章参考文献	162
第 8 章 增强现实在其他典型行业中的应用	163
8.1 国防军事	163
8.1.1 增强现实模拟演练	163
8.1.2 单兵作战态势感知	164
8.2 医疗健康	165
8.2.1 医学教育和手术培训	166
8.2.2 手术规划	167
8.2.3 手术导航	169
8.3 教育	169
8.4 生活导航	171
8.5 游戏娱乐	172
8.6 广告商务	173
8.7 文化旅游	173
本章小结	174
本章参考文献	174
第 9 章 典型增强现实软/硬件系统	176
9.1 典型增强现实眼镜	176
9.1.1 Google Glass	176
9.1.2 Epson Moverio	177
9.1.3 Vuzix M4000	178
9.1.4 Hololens	179
9.1.5 Magic Leap One	181

9.1.6 晨星 DaystAR	182
9.1.7 Nreal Light	183
9.1.8 影创 Hong Hu	184
9.2 移动增强现实软件开发平台	185
9.2.1 ARKit	186
9.2.2 ARCore	187
9.2.3 Vuforia	188
本章小结	195
本章参考文献	195
第 10 章 增强现实的趋势与未来	196
10.1 硬件平台的未来	196
10.1.1 传感器	196
10.1.2 显示设备	197
10.2 AR 与 AI	198
10.3 AR 与 5G	199
10.4 AR 云	200
10.5 多人 AR	201
10.6 AR 与数字孪生	202
10.7 AR 空间计算	203
本章小结	205
本章参考文献	205

第 1 章

增强现实概论

增强现实(Augmented Reality, AR)是一项革命性的信息技术,常与虚拟现实(Virtual Reality, VR)相提并论,但二者却有显著区别。前者可将计算机渲染生成的虚拟场景与真实物理世界进行无缝融合,通过显示设备将虚实融合的场景呈现给用户,而后者要截断人类感知真实物理世界的信息,将其替换为计算机生成的数字信息,是一个纯虚拟的世界。在本章中,我们将探索增强现实和虚拟现实的区别,分析增强现实系统的关键技术,并由远及近地梳理增强现实的发展历程,为后续增强现实技术原理的学习和应用实践打下基础。

1.1 增强现实技术概述

1.1.1 增强现实的定义

增强现实是一种可以把计算机生成的信息(如三维模型、图片、音乐等)叠加到真实环境中,以增强人们对真实场景感官和认知的技术,该技术是在虚拟现实基础上发展而来的。虽然增强现实是近年来才“火”起来的一个名词,但早在 1992 年,波音公司的科学家 Caudell 和 Mizell 就首次提出这一概念^[1],并尝试在飞机装配布线过程中引入这一技术。1994 年多伦多大学的 Milgram 等人首次提出了“现实-虚拟连续体”的概念^[2],认为从虚拟世界到真实环境是一个连续渐变的过程,如图 1-1 所示。在虚拟环境和真实环境的交界部分存在着一种“混合现实”的可能,在交界地带里,相对靠近真实环境的部分为增强现实。其可利用计算机技术生成虚拟对象来“增强”用户的真实体验。

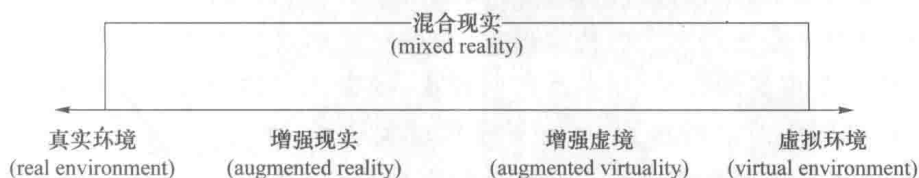


图 1-1 虚拟现实连续体

与上述定义类似,1997年 Azuma 对增强现实技术提出了一个更为广泛的定义^[3],认为增强现实技术是指将虚拟环境准确注册到物理世界,使虚拟与现实相融合,并实现实时交互的一种技术。增强现实技术能有效突破现实场景中时间和空间等因素的限制,使得传统上难以直接被用户体验到的实体信息通过增强现实技术与现实场景叠加并被人感知到,从而加强用户的操作体验,在工业制造、军事、医疗和教育等诸多方面有了越来越广泛的应用。

1.1.2 增强现实技术的特点

根据 Azuma 对增强现实的定义,增强现实是一个具有虚实融合、实时交互和三维沉浸特征的系统,具体表现在以下方面。

(1) 虚实融合

虚实融合可将虚拟的物体信息叠加或合成到真实世界中,是增强现实的具体表现形式。它允许观测者看见虚拟和现实融合的世界,并强化真实物理场景(并不是直接替换),因此通过虚拟信息与真实环境的融合,用户可以“身临其境”地对虚拟内容进行观测和理解。

(2) 实时交互

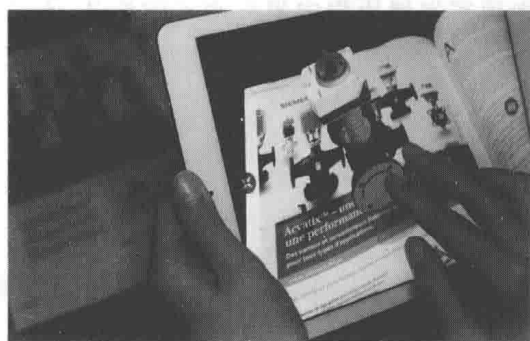
在增强现实为用户提供场景虚实融合的基础上,为使融合的视觉效果变得更为逼真,需要突破增强现实虚物理的界限,构建虚拟物体与用户和真实环境间的互动方式。通过利用语音、手势等实现实时多模态信息的操作交互,提升增强现实过程中用户认知体验和知识互动的能力,从而使用户获取更多的认知体验。

(3) 三维沉浸

三维沉浸是根据用户在三维空间的运动,实时调整计算机产生的增强信息,为用户构建出具有“身临其境”效果的体验环境,保证用户沉浸在其中的认知体验与真实物理世界的认知体验相似或者相同。这种“身临其境”的体验为用户构建了三维沉浸式环境,可让用户更容易融入增强现实构建的虚实一体化系统中。

1.1.3 增强现实与虚拟现实的区别

虽然增强现实技术是在虚拟现实技术的基础上发展起来的,两者有一些共同点,但是随着增强现实技术探索的深入,研究者发现两者有着明显的区别。增强现实技术更侧重于将虚拟物体叠加到真实的物理场景,在用户接触真实环境的基础上融合一些虚拟场景,构建一个虚实结合、三维沉浸的新环境,使得用户仍然保持对真实场景的感知和交互能力〔如图 1-2(a)所示〕。而对于虚拟现实技术,在用户与虚拟物体交互时,所有的交互场景都是计算机通过特定的程序仿真生成的,用户的感受完全是基于虚拟场景产生的〔如图 1-2(b)所示〕。



(a) 增强现实示意图



(b) 虚拟现实示意图

图 1-2 增强现实与虚拟现实的差别

1.2 增强现实的发展历程

1.2.1 起源阶段

利用计算机生成的信息在物理世界中进行注释首次出现在 20 世纪 60 年代。计算机图形学和增强现实之父 Sutherland 于 1965 年在“*The Ultimate Display*”(终极显示)^[4]一文中给出了关于 VR 和 AR 的著名论断：“终极显示应该是有这么一个房间，在这个房间中计算机可以控制物体的存在，显示的椅子可以坐下，显示的手铐可以将你束缚，显示的子弹是可致命的，通过适当的编程，这样的显示可以真正地被称为爱丽丝梦游的仙境。”

此后不久，Sutherland 于 1968 年开发出了第一套增强现实系统^[5]，它被命名为“达摩克利斯之剑”(sword of damocles)(如图 1-3 所示)。这套系统使用一个光学透视头戴式显示器，同时配有两个 6DoF(degree of freedom)追踪仪(一个是机械式，另一个是超声波式)，头戴式显示器由其中之一进行追踪。受制于当时计算机的处理能力，这套系统较重。其显示设备悬吊在用户头顶的天花板上，并通过连接杆和头戴式显示器相连，它能够简单线框图转换为 3D 效果的图像。

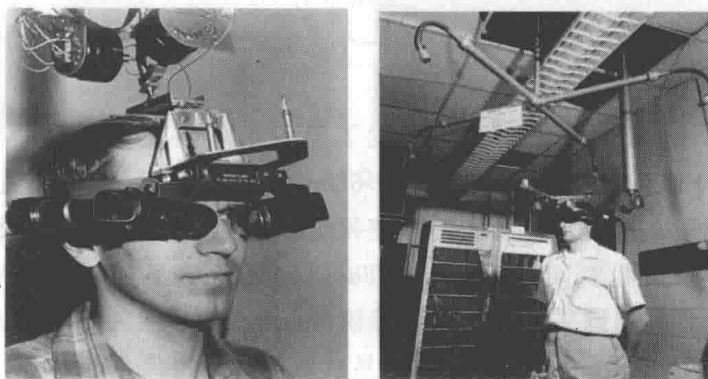


图 1-3 “达摩克利斯之剑”的头戴式显示器系统

虽然这套系统被业界认为是虚拟现实和增强现实发展历程中里程碑式的作品，不过在当时除了得到大量科幻迷的热捧外，并没有引起很大轰动，笨重的外表和粗糙的图像系统都大大限制了产品在普通消费者群体里的应用。

1.2.2 初步发展阶段

20 世纪八九十年代计算机性能的提高促使增强现实发展成为一个独立的研究领域。1992 年波音公司 Caudell 和 Mizell 在辅助布线系统中最早提出增强现实的概念^[1]，他们通过开发头戴式透视显示系统，使工程师能够使用叠加在电路板上的数字化增强现实图解来组装这个电路板上的复杂电线束。该增强现实系统虚拟化了布线图，可指导工程师们完成布线操作，有效提高了工作效率。

1993 年，哥伦比亚大学的 Feiner 等人提出了基于知识的 AR 系统 KARMA^[6]，其全称为基于知识的增强现实维修助手 (knowledge-based augmented reality for maintenance assistance)。通过在头戴式显示设备中叠加虚拟的作业辅助信息，指导人们如何安装和维护打印机。这或许是增强现实技术最早在消费者端的应用产品。该 AR 系统也可应用于激光打印机维修和汽车门锁的安装，能够指导缺乏维修经验的人员高效完成维修任务，如图 1-4 所示。



图 1-4 KARMA 增强现实系统

1997 年，在前人研究的基础上，哥伦比亚大学的 Feiner 等人又开发了第一个户外 AR 游览系统 (touring machine)^[7]。这套系统使用带有 GPS 和姿态跟踪的透视式头戴显示器，为了在移动过程中输出三维图形，系统包括一个内装计算机和传感器的背包，并配有光笔和触控界面的手持平板电脑。

1999 年，Kato 和 Billinghurst 共同开发了第一个增强现实开源框架 ARToolKit^[8]。ARToolkit 基于 GPL 开源协议发布，是一个 6DoF 姿势追踪库，其使用直角基准和基于模板的方法来进行识别。ARToolKit 的出现使得增强现实技术不仅仅局限在专业的研究机构中，许多普通程序员也都可以利用 ARToolKit 开发自己的增强现实应用。早期的 ARToolKit 可以识别和追踪一个黑白人工标识 (marker)，并在黑白的标识上显示 3D 模型 (如图 1-5 所示)。这个巧妙的软件设计与容易获取的网络摄像头相结合，使 ARToolKit 得到了广泛的应用。

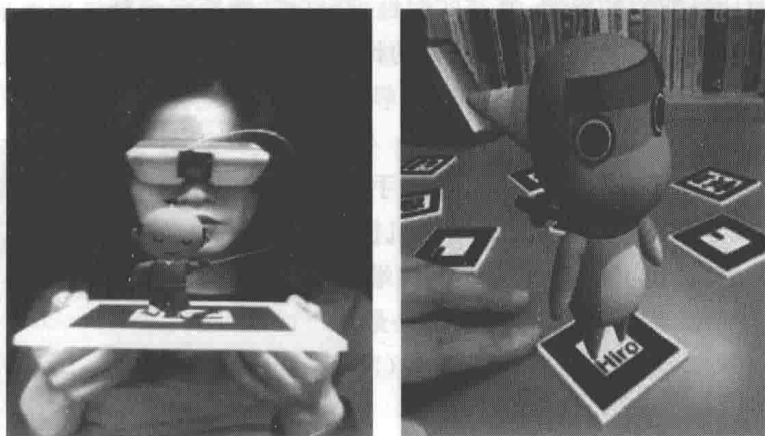


图 1-5 基于 ARToolKit 的增强现实应用

与此同时,德国联邦教育和研究部于 1999 年启动了一项 2 100 万欧元的 AR 项目——ARVIKA(augmented reality for development, production and servicing)^[9],以将 AR 应用在产品开发、生产和服务领域。来自工业界和学术界的 20 多人组成一个研究小组,针对工业应用,特别是德国汽车工业应用的增强现实系统开展研究工作,该计划有效推进了全球对 AR 在专业领域应用的认识。

2000 年,来自南澳大利亚大学的 Bruce Thomas 等人开发了一个项目 ARQuake^[10],将 AR 带到了室外的真实场景。ARQuake 是一个基于 6DoF 追踪系统的第一人称应用。这个追踪系统使用了 GPS、数字罗盘和基于标记的视觉追踪系统。使用者背着一个可穿戴式计算机的背包、一台头戴式显示器和一个只有两个按钮的输入器(如图 1-6 所示)。这款游戏在室内或室外都能进行,一般游戏中的鼠标和键盘操作由使用者在实际环境中的活动和简单输入界面代替。

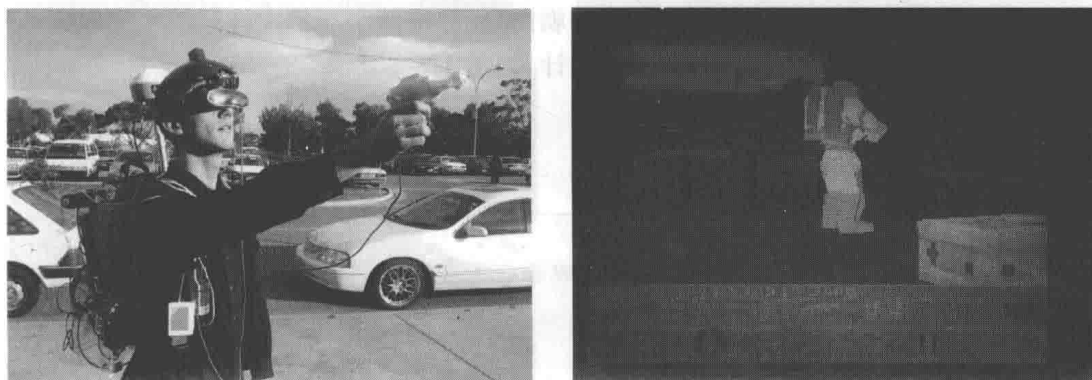


图 1-6 ARQuake 项目

2003 年,Kooper 和 MacIntyre 开发出第一个基于移动增强现实的浏览器(Real-World Wide Web, RWW^W)^[11],其可基于万维网(World Wide Web, WWW)实现三维空间信息与物理世界的交互融合。这套系统受限于当时笨重的增强现实硬件设备,需要一个头戴式显示器和一套复杂的跟踪定位设备。

此外,增强现实技术所需要的设备比较昂贵,对应的软件算法对 CPU 和 GPU 性能的

要求高,受到早期计算能力的限制,其在实时性、交互性、跟踪定位精度等方面都难以得到有效保证。信息技术的发展,尤其是移动平台的计算能力和传感器性能的快速提升,为更具舒适性的移动增强现实应用提供了基础。

2005年, Henrysson 等人将 ARToolKit 与软件开发工具包 (Software Development Kit, SDK) 相结合,可为早期的塞班智能手机提供服务^[12]。开发者通过 SDK 启用 ARToolKit 的视频跟踪功能,实时计算出手机摄像头与真实环境中特定标志之间的相对方位,基于该技术框架,实现了第一个移动手机端的多人协同的 AR 应用——AR 网球游戏,如图 1-7 所示。直到今天,ARToolKit 依然是最流行的增强现实开源框架之一,支持几乎所有的主流平台,并且已经实现自然特征追踪 (Nature Feature Tracking, NFT) 等更高级的功能。



图 1-7 AR 网球游戏

在 2007 年, Klein 和 Murray 提出了一种实时的同步定位和建图 (Parallel Tracking and Mapping, PTAM) 方法^[13],该方法将 SLAM (Simultaneous Localization and Mapping) 中的定位和建图分别在两个独立的线程中运行,通过对场景中自然纹理特征的提取,在进行三维稀疏场景重建的同时解算相机的空间位置和朝向,无须依赖于场景中先验的人工标识,即可为未知场景下增强现实应用提供实时位姿估计,如图 1-8 所示。



图 1-8 基于 PTAM 方法跟踪定位的增强现实

在此基础上,为拓展增强现实的应用场景, Wagner 等人于 2008 年首次提出一种能在手机上进行平面自然特征跟踪的方法^[14]。其对特征描述算子 SIFT (Scale Invariant Feature Transform) 进行修改,以适应于手机有限的计算和存储能力,实验结果表明,其能