

虚拟现实与数字媒体技术前沿

增强现实： 原理、算法与应用

鲍虎军 章国锋 秦学英 著

AUGMENTED REALITY:
PRINCIPLES, ALGORITHMS
AND APPLICATIONS



科学出版社

虚拟现实与数字媒体技术前沿

增强现实：原理、算法与应用

Augmented Reality: Principles, Algorithms and Applications

鲍虎军 章国锋 秦学英 著

科学出版社

北京

内 容 简 介

本书是关于增强现实核心技术原理、算法和应用的一本专著。全书共8章,包括增强现实概论、虚实视觉融合的相机模型、实景的三维结构恢复和重建、虚实环境的实时三维注册、虚实环境的视觉融合、空间增强现实技术、增强现实环境中的交互、移动增强现实系统的设计与应用案例。本书不仅全面介绍增强现实的基本概念和领域知识,更专注于增强现实的最新研究进展和算法阐释,力求涵盖各相关领域的前沿进展。

本书既可作为高等院校计算机类专业的本科生或研究生的参考书,也适合广大从事增强现实技术的研发人员与工程技术人员阅读。

图书在版编目(CIP)数据

增强现实:原理、算法与应用 / 鲍虎军, 章国锋, 秦学英著. —北京: 科学出版社, 2019.5

ISBN 978-7-03-056736-9

I. ①增… II. ①鲍… ②章… ③秦… III. ①虚拟现实—研究 IV. ①TP391.98

中国版本图书馆 CIP 数据核字(2018)第 043854 号

责任编辑: 任 静 / 责任校对: 张凤琴

责任印制: 张克忠 / 封面设计: 迷底书装

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>

北京画中画印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2019 年 5 月第 一 版 开本: 720×1 000 1/16

2019 年 5 月第一次印刷 印张: 22 1/4

字数: 427 000

定价: 145.00 元

(如有印装质量问题, 我社负责调换)

作者简介



鲍虎军，男，博士，浙江大学计算机科学与技术学院教授，博士生导师。国家杰出青年科学基金获得者，教育部长江学者特聘教授。主要从事计算机图形学和混合现实等方面的研究。所领导的团队曾获国家创新研究群体科学基金的资助，并作为首席科学家，先后承担了国家 973 计划项目“虚拟现实的基础理论、算法及其实现”和“混合现实的理论和方法”的研究。在虚拟环境的几何表示和高效建模、虚拟环境的实时高保真绘制、虚实混合环境的实时三维注册和融合呈现等方面取得了重要突破，自主研发了混合现实基础支撑软件平台，实现了成功应用。部分成果荣获高等学校优秀成果奖一等奖和国家自然科学奖二等奖。所指导的多位博士生的学位论文被评为全国百篇优秀博士学位论文和计算机学会优秀博士学位论文。



章国锋，男，博士，浙江大学计算机科学与技术学院教授，博士生导师，国家优秀青年科学基金获得者。主要从事三维视觉与增强现实方面的研究，尤其在同时定位与地图构建和三维重建方面取得了一系列重要成果，研制了一系列相关软件，如 ACTS、LS-ACTS、RDSLAM、RKSLAM 等(<http://www.zjucvg.net>)，并开源了基于非连续特征跟踪的大尺度运动恢复结构系统 ENFT-SfM、分段集束调整 SegmentBA 和高效的增量式集束调整 EIBA、ICE-BA 等算法的源代码(<https://github.com/zju3dv/>)。获全国百篇优秀博士学位论文奖、计算机学会优秀博士学位论文奖以及教育部高等学校科学研究优秀成果奖科学技术进步奖一等奖(排名第 4)。



秦学英，女，博士，山东大学软件学院教授，博士生导师。主要从事增强现实、计算机视觉、计算机图形学等方面的研究。2009年获山东省自然科学基金杰出青年基金资助。参与了国家973计划项目“虚拟现实的基础理论、算法及其实现”和“混合现实的理论和方法”的研究，并主持多项虚实融合相关的863计划项目以及自然科学基金项目。近年从事的主要研究工作包括：虚实融合与交互、光照环境重建、目标跟踪、三维模板跟踪等。

前 言

拥有诸如千里眼、顺风耳、先知先觉等超自然能力，增强感知、探索和改造世界的的能力，一直是人类孜孜追求的梦想和目标。随着信息的获取、传输、存储、处理、分析、模拟和呈现等技术的迅猛发展，信息空间和现实物理空间日益融合，人类的这一梦想逐渐成为现实。作为用户的感知体验终端，增强现实是其中极为关键的一环。它是一种集定位、呈现、交互等软硬件于一体的人机界面技术，通过虚拟物体与现实环境的自动注册融合，营造出虚拟与现实共享同一空间的沉浸式体验，以增强用户对复杂现实环境的临场感知。著名的好莱坞影片《阿凡达》和《黑客帝国》形象阐释了这种虚拟与现实无缝融合的科幻景象。但是，技术的实现远不像壮丽的影视特效制作那么简单，这也是自其诞生之始，增强现实技术一直是信息技术的前沿研究方向的原因。

在增强现实环境中，一切由计算机产生或者存储的信息，不再简单地显示在计算机屏幕上，而是以时空一致的方式注册在相应的实际环境中，一起融合呈现在用户面前，以使用户用感官去自然感知。增强现实技术的理想是伴随人类左右，去感知和探究复杂世界。谷歌眼镜和微软 HoloLens 系统展现了增强现实的预想世界，例如，看向天空时显示天气情况；走向地铁口，则显示地理信息及相關属性信息；远程的人们可以坐在一起，进行面对面的交流；工人们可以临场了解装备内部的运行状态，并在有关信息的导引下，进行精准的系统运维操作；等等。而苹果 ARKit、谷歌 ARCore 和商汤 SenseAR 等移动增强现实软件平台的推出，则催生了增强现实的广泛应用研发，从高端的工业生产到简单的孩子学习和家庭应用，无不体现出其强大的威力。增强现实技术正在以迅雷不及掩耳之势影响现代社会。

我国虚拟现实和增强现实技术的研究，始于 20 世纪 90 年代末，是计算机图形学和计算机视觉研究的自然拓展。2002 年底，浙江大学鲍虎军教授任首席科学家的 973 计划项目“虚拟现实的基础理论、算法及其实现”立项，重点研究虚拟环境表达和呈现、人机交互以及虚实融合等关键问题，开启了虚拟现实和增强

现实技术的系统性研究。2009年,该项目得到持续支持,并重点转向增强现实和混合现实技术的研究。这两个项目的研究持续十年,浙江大学、中国科学院软件研究所和自动化研究所、北京理工大学、北京航空航天大学、清华大学、东南大学等单位深度参与了相关研究,并取得了重要进展,培养了一批优秀人才,建立起了一支高水平的虚拟现实和增强现实技术研究队伍。在虚拟环境的高效建模和实时高保真绘制、多通道人机交互、头盔显示和多投影融合显示、虚实环境的实时三维注册和融合等方面取得了一批国际先进水平的研究成果,成功研发了虚拟现实和增强现实驱动引擎和支撑软件平台,并实现了成功应用,极大地促进了我国虚拟现实和增强现实技术的研究。作为973计划项目的主持单位,浙江大学在虚拟现实和增强现实的建模和实时绘制、虚实环境的三维注册和融合呈现方面取得了重要研究成果,并发表在IEEE TPAMI、ACM TOG、IEEE TVCG、IEEE TIP、ACM SIGGRAPH、ICCV、CVPR、ECCV、EUROGRAPHICS和ISMAR等重要学术期刊与会议上。本书围绕虚实融合的沉浸视觉感知这一主题,集成我们过去十余年的研究成果和相关国内外前沿进展,根据增强现实系统的核心架构组织撰写而成。

本书汇聚了增强现实的前沿核心算法,不仅介绍了相关技术的基本概念和算法,而且较为详细地阐述了虚实视觉融合最为精妙的算法原理及其实现,力求为广大增强现实技术和系统的研发者提供所需的实现途径。本书共8章,第1章介绍了增强现实技术的基本原理、发展历史、显示装置、核心技术和典型应用,展现了国内外的的发展态势;第2章介绍了相机的基本概念、成像模型及畸变校正模型;第3章详细阐述了基于影像的现实环境三维结构恢复和重建,这是三维计算机视觉的核心;第4章专门介绍了增强现实系统必需的实时三维注册技术,包括同时定位与地图构建(SLAM)、二维标志跟踪定位和三维模板跟踪定位等算法;第5章介绍了虚实环境的视觉融合技术,包括虚拟景物的实时高保真绘制、现实环境光照恢复、虚实景物的相互映照、信息可视化等算法;第6章介绍了空间增强现实技术,重点阐述了空间几何校正和颜色校准方法;第7章简单介绍了增强现实系统所需的人机交互技术,包括触控、手势、语音、实物等交互方法;第8章介绍了我们的移动增强现实系统的框架设计,并分析了若干增强现实应用案例。

本书由鲍虎军制定编写大纲并撰写了第1章、第5章和第6章的大部分,章国锋撰写了第3章和第8章,秦学英和章国锋共同撰写了第2章和第4章,黄进和田丰撰写了第7章的绝大部分。王锐、华炜、巫英才、王进参与了部分章节的撰写。李津羽、赵长飞、杨镑镑、叶智超、王儒、任浩城、李佰余、徐超、

王斌、黄昭阳、曹健、柯锦乐、杨腾达、王懿芳、李洋、张浩、李念龙、钱权浩、褚冠宜、丁奇超、唐庆等研究生参与了数据、材料的收集和整理；全书由鲍虎军、章国锋和秦学英定稿。

由于作者水平有限，书中疏漏之处在所难免，恳请读者批评指正。

作 者

2018年8月

目 录

前言

▶ 第 1 章 增强现实概论 / 1

- 1.1 增强现实的基本原理 / 2
- 1.2 增强现实的发展简史 / 4
- 1.3 增强现实的显示装置 / 6
- 1.4 增强现实的关键技术 / 9
 - 1.4.1 三维注册与几何一致性 / 10
 - 1.4.2 真实感绘制与光照一致性 / 12
 - 1.4.3 人机交互 / 14
- 1.5 增强现实应用 / 16
 - 1.5.1 数字娱乐应用 / 16
 - 1.5.2 文化教育应用 / 16
 - 1.5.3 医学诊疗应用 / 17
 - 1.5.4 工业生产应用 / 18
 - 1.5.5 城市运维应用 / 20
 - 1.5.6 电子商务应用 / 21
 - 1.5.7 其他应用 / 22
- 1.6 小结 / 22

▶ 第 2 章 虚实视觉融合的相机模型 / 23

- 2.1 相机成像原理及其数学模型 / 23
 - 2.1.1 相机成像原理 / 23
 - 2.1.2 相机的数学模型 / 24
 - 2.1.3 透视投影矩阵 / 28

2.2 镜头的非线性畸变 / 30

2.3 其他相机模型 / 32

2.4 小结 / 33

第3章 实景的三维结构恢复和重建 / 34

3.1 多视图几何原理 / 35

3.1.1 双视图几何 / 35

3.1.2 多视图几何 / 43

3.2 特征与匹配 / 45

3.2.1 特征提取 / 46

3.2.2 特征匹配的主流算法 / 50

3.2.3 连续特征跟踪和非连续特征匹配 / 53

3.2.4 特征点匹配的发展趋势 / 57

3.3 运动恢复结构 / 57

3.3.1 初始化 / 58

3.3.2 相机位姿估计 / 61

3.3.3 集束调整 / 62

3.3.4 自定标 / 73

3.3.5 代表性 SfM 方法 / 80

3.4 稠密深度估计 / 83

3.4.1 双视图立体匹配 / 84

3.4.2 多视图立体匹配 / 86

3.4.3 深度学习方法 / 89

3.4.4 数据集和测评网站 / 90

3.5 三维几何重建 / 91

3.5.1 基于深度图融合的三维重建 / 91

3.5.2 基于点云的三维重建 / 93

3.5.3 基于体素的三维重建 / 94

3.5.4 代表性的三维重建系统 / 95

3.6 表面纹理重建 / 96

- 3.6.1 纹理映射框架 / 98
- 3.6.2 纹理拼接 / 98
- 3.6.3 颜色融合 / 99
- 3.6.4 配准误差 / 101
- 3.7 小结 / 101

▶ 第 4 章 虚实环境的实时三维注册 / 103

- 4.1 同时定位与地图构建 / 103
 - 4.1.1 视觉 SLAM / 105
 - 4.1.2 视觉惯性 SLAM / 119
 - 4.1.3 RGB-D SLAM / 129
 - 4.1.4 问题与发展趋势 / 134
- 4.2 平面标志物的检测、识别与跟踪 / 138
 - 4.2.1 人工标志的识别与跟踪 / 139
 - 4.2.2 自然标志的识别与跟踪 / 143
 - 4.2.3 问题与讨论 / 145
- 4.3 三维实物的检测、识别与跟踪 / 146
 - 4.3.1 概述 / 146
 - 4.3.2 基于特征点的三维物体跟踪 / 149
 - 4.3.3 基于边缘的三维物体跟踪 / 150
 - 4.3.4 基于区域的三维模板跟踪 / 156
 - 4.3.5 问题与讨论 / 161
- 4.4 小结 / 162

▶ 第 5 章 虚实环境的视觉融合 / 163

- 5.1 虚拟物体的绘制 / 163
 - 5.1.1 绘制方程 / 164
 - 5.1.2 BRDF 函数 / 165
 - 5.1.3 局部光照计算与绘制流水线 / 171
 - 5.1.4 全局光照计算 / 174

- 5.2 现实环境的光照估计 / 177
 - 5.2.1 高动态范围图像 / 178
 - 5.2.2 室内光照估计 / 179
 - 5.2.3 室外光照估计 / 184
- 5.3 虚实视觉融合 / 188
 - 5.3.1 虚实景物间的阴影生成 / 188
 - 5.3.2 屏幕空间环境光遮蔽技术 / 192
 - 5.3.3 虚实景物间的间接光照计算 / 196
- 5.4 信息与现实环境的融合呈现 / 206
 - 5.4.1 用户视觉感知机制 / 207
 - 5.4.2 信息的可视化和可视分析 / 209
 - 5.4.3 信息的时空注册和呈现 / 211
- 5.5 小结 / 218

▶ 第6章 空间增强现实技术 / 220

- 6.1 空间增强现实的原理和方法 / 220
 - 6.1.1 光学穿透式空间呈现 / 221
 - 6.1.2 光场式空间呈现 / 222
 - 6.1.3 多投影空间呈现 / 224
- 6.2 投影画面的几何校正 / 226
 - 6.2.1 几何校正的基本原理 / 226
 - 6.2.2 图像的几何形变映射 / 228
 - 6.2.3 交互式几何校正 / 229
 - 6.2.4 自动几何校正 / 230
- 6.3 投影画面的颜色校正 / 234
 - 6.3.1 投影光照度特征函数 / 235
 - 6.3.2 颜色校正原理 / 237
 - 6.3.3 非均质投影系统的颜色校正 / 238
- 6.4 多投影空间增强现实呈现系统 / 240
 - 6.4.1 系统架构和实现 / 241

6.4.2 应用案例 / 243

6.5 小结 / 245

▶ 第 7 章 增强现实环境中的交互 / 246

7.1 增强现实系统的人机交互概述 / 246

7.1.1 增强现实系统的交互设计原则与评估方法 / 247

7.1.2 增强现实系统的主流交互技术 / 250

7.1.3 问题与发展趋势 / 254

7.2 触控交互 / 256

7.2.1 触控交互设备 / 256

7.2.2 触控交互技术 / 257

7.2.3 具触控交互的典型增强现实系统 / 261

7.3 手势交互 / 261

7.3.1 手势输入设备 / 262

7.3.2 手势识别技术 / 264

7.3.3 应用案例 / 273

7.4 语音交互 / 274

7.4.1 语音交互的基本原理 / 274

7.4.2 语音交互的典型应用 / 277

7.5 实物交互 / 279

7.5.1 实物交互模型 / 279

7.5.2 实物感知与驱动技术 / 283

7.5.3 实物交互界面 / 286

7.6 小结 / 288

▶ 第 8 章 移动增强现实系统的设计与应用案例 / 289

8.1 系统架构设计 / 289

8.1.1 输入模块 / 290

8.1.2 三维注册和重建模块 / 291

8.1.3 虚实融合模块 / 295

8.2 应用案例解析 / 298

8.2.1 教育类产品 / 298

8.2.2 AR 游戏 / 300

8.2.3 AR 旅游 / 302

8.2.4 工业应用 / 304

8.3 小结 / 306

► 参考文献 / 307

第 1 章

增强现实概论

复杂现实环境往往拥有大尺度的空间结构、高动态的运行演化以及多关联的社会活动，这些特点使得人类难以直接凭借自身有限的感知能力来全面理解掌握其中的运行演化进程，从而影响人类对现实环境所发生的复杂对象(包括事物、人物)和事件的准确认知和有效管控。增强对复杂现实环境的感知能力一直是人类孜孜以求的重大科学目标。半个多世纪以来，广大科研人员为此付出了巨大的努力，在现实物理世界的信息获取、智能处理和分析、信息传输和呈现等方面取得了重大突破，诞生了以互联网和物联网、云计算和移动计算、机器学习和数据智能、虚拟现实和增强现实(augmented reality, AR)等为代表的新一代信息技术，赋予人类千里眼、顺风耳、先知先觉等超自然的感知能力，极大提升了人类探索和改造世界的能力。

增强现实技术通过建立虚拟环境和现实物理世界的映射关系，智能实时地将相关信息融合呈现在位于现实环境的用户面前，达到增强用户临场感知能力的目的。这里的“增强”主要体现在以下三个方面：①空间穿透。通过三维结构信息与相应现实环境的在线、精准的空间注册和融合呈现来增强用户的环境智能理解能力，使得用户可以在视觉上穿透现实空间环境，让“不可见的”变为“可见的”，让“可见的”变为“不可见的”，让“遥远的”变为“眼前的”。②时间折叠。通过所关联的历史和实时信息来扩展现实环境在时域上的维度，使得用户既可临场

回溯再现历史、分析预测未来，也可自由地组合呈现不同时刻的信息，让用户拥有先知先觉的能力。③知识叠加。自动识别环境中的实物对象，将其所关联的信息、知识或不同条件下的演化模拟结果实时地叠加到该对象上，使得用户即时了解现实环境和事件所需的关键信息和知识。增强现实的这三个特点为用户对复杂现实环境的准确感知和分析决策提供了一种重要的技术手段，得到了国际学术界和产业界的高度重视和广泛关注。

近年来，增强现实技术发展迅猛，日益成熟。国际 IT 巨头苹果、谷歌、微软等分别推出了移动增强现实软件开发平台 ARKit 和 ARCore 以及增强现实头盔显示器 HoloLens，掀起了增强现实应用系统的研发热潮。尤其是，随着移动通信和智能终端的普及，增强现实技术逐渐由国防安全、工业生产、医疗康复、城市管理等高端应用扩展到电子商务、文化教育、数字娱乐等大众应用，成为人们认知世界和改造世界的基础性工具。

1.1 增强现实的基本原理

增强现实是一种将虚拟景物或信息与现实物理环境叠加融合起来，交互呈现在用户面前，从而营造出虚拟与现实共享同一空间的技术。本质上，增强现实是一种集定位、呈现、交互等软硬件技术于一体的新型界面技术，其目标是让用户在感官上感觉到虚实空间的时空关联和融合，来增强用户对现实环境的感知和认知。因此，它拥有三个基本要素(Azuma, 1997)，即虚实空间的融合呈现、实时在线交互以及虚实空间的三维注册。虚实空间的融合呈现，强调虚拟元素与真实元素的并存，这是用户对现实环境的感知得以增强的关键；实时在线的交互，强调用户和虚实物体之间互动响应计算的实时性，以满足用户感官对时间维度的响应需求；而虚实空间的三维注册，强调用户对空间感知的精确性和智能性，体现了虚实融合呈现的时空一致性。显然，这三个要素是实现现实环境增强感知的关键所在。由于这种增强感知是空间方位依赖的，因此，增强现实系统通常借助头盔等特制设备来呈现虚实融合的效果。

增强现实技术是虚拟现实技术的进一步发展。虚拟现实是完全由计算机营造的世界，为用户提供沉浸式的视觉、听觉、触觉等互动感知体验，这种沉浸式系统能够逼真再现现实世界，也可以模拟虚构的、现实世界不存在的环境。虚拟现实的这种模拟仿真特性使得我们可以用它来增强用户对现实环境的感知，从而催生了将虚拟景物或信息与现实世界融为一体的增强现实技术。Milgram 等(1994)

将介于虚拟与现实之间的世界统称为混合现实连续统(mixed reality continuum),并将这种虚拟世界和现实世界的融合技术称为混合现实技术(如图 1.1 所示),其中用虚拟信息来增强用户对现实世界的感知称为增强现实技术,而用现实信息来增强用户对虚拟世界的感知则称为增强虚拟技术。从技术上来说,增强现实和增强虚拟的实现方法是类似的,它们均要求用户对虚实融合环境的感知体验等同于现实物理世界的知觉。

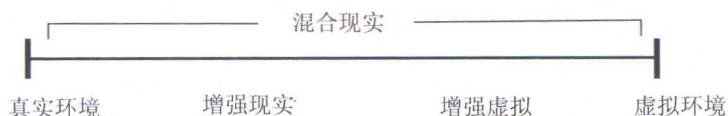


图 1.1 混合现实连续统

根据人类知觉的特点,虚实环境的融合无疑应该涵盖人类知觉的视、听、触等多个感知通道。例如,一个虚拟的球掉落在现实物理世界的地面上,这个虚拟球不仅应该反弹跳起,而且在反弹的时候,还要发出球击地面的声音,产生一致的听觉体验。三十年来,虚实环境的多通道感知融合呈现技术得到了快速发展,尤其是视觉感知融合呈现技术日益成熟,极大地促进了增强现实技术的普及应用。

为了实现虚实环境的沉浸式视觉融合呈现,满足增强现实三要素的要求,需要在算法上涉及三大核心技术,即虚实空间的实时三维注册、虚实景物的高保真融合、高效的人机交互。人类的眼睛经历了亿万年的进化而高度精密,使其对场景空间关系和景物细节非常敏感,由于视觉感知与视点和观察方向强相关,因此当虚实环境融为一体时,需要实时地跟踪景物和用户的观察方位,实现高精度的虚实空间注册。其次,由于虚拟物体要“假装”真实地存在于现实世界之中,增强现实系统需要实时逼真地呈现虚实环境间光能传递和动力学的相互作用,才能产生视觉感知的沉浸性。再次,虚实融合场景中的景物往往复杂而动态变化,因此需要实时地理解用户的交互意图,以提高任务的临场完成效率。可以说,增强现实技术将在几何、物理、智能等层面实现虚拟世界与物理世界的视觉融合,以增强人类感知和改造客观世界的能力。例如,在复杂装配过程中,可将设计方案叠加在现实安装组件上,以迅速判断装配的正确与否,避免错误装配的发生(如图 1.2 所示)。增强现实系统以符合人类感知习惯、直观的信息传递方式,赋予用户额外的感知客观世界的能力,从而成为未来信息系统的重要使能技术。