



信息通信领域产学研合作特色期刊

第三届全国期刊奖百种重点期刊 | 中国科技核心期刊

ISSN 1009-6868

CN 34-1228/TN

# 中兴通讯技术

## ZTE TECHNOLOGY JOURNAL

tech.zte.com.cn

2017年12月 • 第6期

**专题：移动VR/AR技术与应用**



# 《中兴通讯技术》第7届编辑委员会委员名单

**顾问** 侯为贵（中兴通讯股份有限公司创始人） 钟义信（北京邮电大学教授） 陈锡生（南京邮电大学教授）

**主任** 陆建华（中国科学院院士,清华大学教授）

**副主任** 赵先明（中兴通讯股份有限公司总裁） 糜正琨（南京邮电大学教授）

**副主任** 马建国（广东工业大学教授） 陈前斌（重庆邮电大学副校长）

## 编委（按姓名拼音排序）

- |                                       |  |
|---------------------------------------|--|
| <b>曹淑敏</b> 中国信息通信研究院前任院长              | <b>谈振辉</b> 北京交通大学教授                      |
| <b>陈建平</b> 上海交通大学教授                   | <b>唐雄燕</b> 中国联通网络技术研究院首席专家               |
| <b>陈杰</b> 中兴通讯股份有限公司高级副总裁             | <b>童晓渝</b> 中电科软件信息服务有限公司副总经理             |
| <b>陈前斌</b> 重庆邮电大学副校长                  | <b>王京</b> 清华大学教授                         |
| <b>葛建华</b> 西安电子科技大学通信工程学院副院长          | <b>王文东</b> 北京邮电大学软件学院副院长                 |
| <b>管海兵</b> 上海交通大学电子信息与电气工程学院副院长       | <b>王翔</b> 中兴通讯股份有限公司副总裁                  |
| <b>洪波</b> 中兴发展股份有限公司总裁                | <b>王喜瑜</b> 中兴通讯股份有限公司副总裁                 |
| <b>洪伟</b> 东南大学信息科学与工程学院院长             | <b>卫国</b> 中国科学技术大学教授                     |
| <b>纪越峰</b> 北京邮电大学信息光子学与光通信研究院<br>执行院长 | <b>吴春明</b> 浙江大学教授                        |
| <b>蒋林涛</b> 中国信息通信研究院科技委主任             | <b>邬贺铨</b> 中国工程院院士                       |
| <b>李尔平</b> 浙江大学信息学部副主任                | <b>徐安士</b> 北京大学教授                        |
| <b>李红滨</b> 北京大学教授                     | <b>续合元</b> 中国信息通信研究院副总工                  |
| <b>李建东</b> 西安电子科技大学副校长                | <b>徐慧俊</b> 中兴通讯股份有限公司执行副总裁               |
| <b>李军</b> 清华大学教授                      | <b>薛一波</b> 清华大学教授                        |
| <b>李乐民</b> 中国工程院院士,电子科技大学教授           | <b>杨义先</b> 北京邮电大学教授                      |
| <b>李融林</b> 华南理工大学教授                   | <b>杨震</b> 南京邮电大学校长                       |
| <b>李少谦</b> 电子科技大学通信与信息工程学院院长          | <b>尤肖虎</b> 东南大学教授                        |
| <b>李涛</b> 南京邮电大学计算机学院院长               | <b>张宏科</b> 北京交通大学教授                      |
| <b>李星</b> 清华大学教授                      | <b>张平</b> 北京邮电大学网络技术研究院执行院长              |
| <b>刘建伟</b> 北京航空航天大学教授                 | <b>张云勇</b> 中国联通研究院院长                     |
| <b>陆建华</b> 中国科学院院士,清华大学教授             | <b>赵慧玲</b> 中国通信标准化协会网络与业务能力技术工作<br>委员会主席 |
| <b>马建国</b> 广东工业大学教授                   | <b>赵先明</b> 中兴通讯股份有限公司总裁                  |
| <b>孟洛明</b> 北京邮电大学教授                   | <b>郑纬民</b> 清华大学教授                        |
| <b>糜正琨</b> 南京邮电大学教授                   | <b>钟章队</b> 北京交通大学教授                      |
| <b>庞胜清</b> 中兴通讯股份有限公司执行副总裁            | <b>周亮</b> 南京邮电大学通信与信息工程学院副院长             |
| <b>孙知信</b> 南京邮电大学现代邮政学院院长             | <b>朱近康</b> 中国科学技术大学教授                    |



信息通信领域产学研合作特色期刊  
第三届国家期刊奖百种重点期刊  
中国科技核心期刊  
工信部优秀科技期刊  
中国五大文献数据库收录期刊  
ISSN 1009-6868  
CN 34-1228/TN  
1995年创刊

## 办刊宗旨

以人为本,荟萃通信技术领域精英;  
迎接挑战,把握世界通信技术动态;  
立即行动,求解通信发展疑难课题;  
励精图治,促进民族信息产业崛起。

# Contents 目次

中兴通讯技术 总第137期 第23卷 第6期(卷终) 2017年12月

## 专题:移动VR/AR技术与应用

- 02 虚拟环境中人和虚拟角色互动的关键技术.....翁冬冬,薛雅琼  
06 支持大规模视频融合的混合现实技术.....周忠,孟明,周颐  
10 增强现实中的空间深度感知问题.....陈东义  
14 基于虚拟现实技术的无线自组网仿真.....罗训,王荣  
19 虚拟现实视频传输架构和关键技术.....郭宗明,班怡璇,谢澜  
24 大视频VR直播业务及其技术.....徐火顺,叶剑章,尹海生  
28 HEVC兼容的全景视频运动补偿预测算法.....胡强,马文广,石志儒  
32 移动增强现实应用及进展.....潘志庚,高翔,丁丹丹

## 专家论坛

- 37 关于增强现实技术潜在发展方向的思考.....乔秀全,任沛,商彦磊

## 专家视点

- 41 封装天线技术发展历程回顾.....张跃平  
50 面向沉浸式视频编码和传输的高效投影方法.....王悦名,王荣刚,高文

## 企业视界

- 56 智能工业装备操作系统技术及创新.....崔云峰,钟卫东,刘东

## 综合信息

- 《中兴通讯技术》2018年专题计划(55) 《中兴通讯技术》第23卷总目次(I)

期刊基本参数:CN 34-1228/TN\*1995\*b\*16\*64\*zh\*P\*¥ 20.00\*15000\*12\*2017-12

# Contents 目次

ZTE TECHNOLOGY JOURNAL Vol. 23 No. 6 Dec. 2017

## Special Topic: Technologies and Applications of Mobile VR/AR

- 02 The Key Technologies of Human-Virtual Character Interaction in Virtual Environment ..... WENG Dongdong, XUE Yaqiong
- 06 Massive Video Integrated Mixed Reality Technology ..... ZHOU Zhong, MENG Ming, ZHOU Yi
- 10 Spatial Depth Perception in Augmented Reality ..... CHEN Dongyi
- 14 Wireless ad hoc Network Simulation Based on Virtual Reality Technologies ..... LUO Xun, WANG Rong
- 19 Architecture and Key Technologies of Virtual Reality Video Transmission and Streaming ..... GUO Zongming, BAN Yixuan, XIE Lan
- 24 Service and Technology of VR Live on Big Video System ..... XU Huoshun, YE Jianzhang, YIN Haisheng
- 28 Motion Compensation and Prediction for Panoramic Video Coding on HEVC ..... HU Qiang, MA Wenguang, SHI Zhiru
- 32 Application and Development of Mobile Augmented Reality ..... PAN Zhigeng, GAO Xiang, DING Dandan

## Expert Forum

- 37 Thoughts on the Potential Development Direction of Augmented Reality Technology ..... QIAO Xiuquan, REN Pei, SHANG Yanlei

## Expert Perspective

- 41 A Review of the Development of Antenna-in-Package Technology ..... ZHANG Yueping
- 50 Efficient Projections for Immersive Video Coding and Transmission ..... WANG Yueming, WANG Ronggang, GAO Wen

## Enterprise View

- 56 Technologies and Innovations of Intelligent Industrial Equipment Operating System ..... CUI Yunfeng, ZHONG Weidong, LIU Dong

## 敬告读者

本刊享有所发表文章的版权,包括英文版、电子版、网络版和优先数字出版版权,所支付的稿酬已经包含上述各版本的费用。

未经本刊许可,不得以任何形式全文转载本刊内容;如部分引用本刊内容,须注明该内容出自本刊。

## 2017年第1—6期专题

### 1 NB-IoT 体系标准与应用

王喜瑜 中兴通讯股份有限公司副总裁  
孙知信 南京邮电大学教授

### 2 通信网络重构新技术

赵慧玲 中国通信标准化协会网络与业务能力技术工作委员会主席

### 3 5G 无线传输技术

卫国 中国科学技术大学教授

### 4 神经网络与深度学习

焦李成 西安电子科技大学教授  
杨淑媛 西安电子科技大学教授

### 5 硅基光电子集成技术与器件

陈建平 上海交通大学教授  
张佩华 中兴通讯股份有限公司技术总监

### 6 移动 VR/AR 技术与应用

潘志庚 杭州师范大学教授





潘志庚教授

杭州师范大学数字媒体与  
人机交互研究中心主任,教授、博导,广东虚拟现实产业研究院院长,中国图像图形学学会副理事长,中国虚拟现实与可视化技术战略联盟副理事长,中国仿真学会数字娱乐仿真专委会主任,中国计算机学会虚拟现实与可视化专委会副主任等,2006—2011年期间担任国际虚拟现实期刊《IJVR》主编,2013年担任IEEE VR国际会议程序委员会共同主席;1992年开始从事虚拟现实及相关技术的研究,承担或完成国家自然科学基金重点项目2项、国家科技支撑计划项目2项、“863”高科技计划1项等项目;2001年获中国高校科技进步奖二等奖,2008年获得国家科技进步二等奖,2011年获得上海市科技进步一等奖,2015年获得浙江省自然科学二等奖,2001年和2013年两次入选浙江省“钱江学者”特聘教授;在顶级期刊和会议上发表论文80余篇。

## 导读

虚拟现实(VR)与增强现实(AR)作为一种新的仿真与互动技术,已成为学术界及产业界的关注热点,同时也在改变着人们的生活,并且这种改变正不断地深入与扩大。移动VR/AR技术是VR/AR的一个分支,目前主要有传统移动VR/AR技术和基于移动终端的移动VR/AR技术两种。

本专题共有8篇文章:《虚拟环境中人和虚拟角色互动的关键技术》一文,从交互环路剖析虚拟环境中人-虚拟角色互动中存在的核心技术,包括空间定位、动作捕捉、眼动追踪和语音输入等;《支持大规模视频融合的混合现实技术》一文介绍了国际上的相关研究,及作者在虚拟现实视频融合技术方面的相关成果;《增强现实中的空间深度感知问题》一文对AR中感知问题进行分类,阐述并分析了影响AR中距离感知的4个因素,即增强叠加、显示装置、用户感知;《基于虚拟现实技术的无线自组网仿真》一文提出了一种网络仿真的新技术——基于虚拟现实的无线自组网仿真技术,认为游戏引擎中的场景管理、动画渲染、脚本控制、物理引擎等为自组网仿真提供了方便的接口,利用这些接口可以很轻松的实现自组网仿真工作;《虚拟现实视频传输架构和关键技术》一文介绍了VR视频传输架构以及全景传输、视点自适应传输等关键技术,并详细论述了均匀映射、非均匀映射和分块等各种传输方式的优缺点;《大视频VR直播业务及其技术》一文提出了一种在IPTV/OTT大视频系统上的VR直播业务的实现技术,可以让IPTV/OTT的用户通过机顶盒在电视上观看360°场景的VR直播视频,或通过手机加头盔、一体化集成的VR眼镜等移动设备来观看VR直播视频;《HEVC兼容的全景视频运动补偿预测算法》一文提出了一种全新的运动补偿预测算法,将当前像素投影回球面并寻找其匹配的参考像素块,在此基础上利用球面上运动相对一致的特点,提出了一种新的运动向量预测模式;《移动增强现实应用及进展》一文综述移动VR技术的应用,并对商业应用前景和商业模式进行了展望。

VR/AR作为一种新技术,将带动一大批新的产业发展模式。另外,改变人们生活和社会的还有4G与5G网络技术。5G网络将会带来更高的速率、更宽的带宽,能够充分满足消费者对虚拟现实、超高清视频等对高带宽、低延时要求更高的业务体验需求。可以预见:在未来的生活中,AR/VR将具有极其广阔的发展与应用前景,将会成为5G的一种“杀手级”应用。

本期专题我们将和大家一起讨论移动VR/AR的一些最新进展,这些论文凝聚了作者多年的研究成果和工作经验,希望能给读者有益的启示与参考。在此,对各位作者的积极支持和辛勤工作表示衷心的感谢。

潘志庚

2017年10月25日



# 虚拟环境中人和虚拟角色互动的关键技术

## The Key Technologies of Human-Virtual Character Interaction in Virtual Environment

翁冬冬/WENG Dongdong  
薛雅琼/XUE Yaqiong

(北京理工大学, 北京 100081)  
(Beijing Institute of Technology, Beijing  
100081, China)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2017) 06-0002-004

**摘要:** 剖析虚拟环境中人和虚拟角色互动的关键技术, 其中空间定位、动作捕捉、眼动追踪和语音输入实现虚拟角色对用户的多通道交互输入信号的接收, 反馈控制系统建立从交互输入信号到呈现给用户的多通道反馈输出信号的映射, 反馈信号渲染将多通道反馈输出信号以虚拟角色的表情、动作、语音等形式呈现给用户, 同时力触觉反馈技术支持虚拟角色提供触觉通道的交互, 使交互具有更多的物理性。

**关键词:** 人-虚拟角色互动; 虚拟现实; 多通道交互信号接收; 反馈控制; 多通道反馈信号渲染; 力触觉反馈

**Abstract:** In this paper, the key technologies of human-virtual character interaction in virtual environment are analyzed. The reception of multi-channel interactive input signal of user is achieved by space positioning, motion capture, eye tracking and voice input. The mapping from interactive input signal to response signal is constructed by response control system. The multi-channel response signal is shown to user by response render technology with lifelike facial appearance, motion and voice. Moreover, force feedback technology adds the haptic interaction to human-virtual character interaction, and enhances the physicality of interaction.

**Key words:** human-virtual character interaction; virtual reality; reception of multi-channel interactive signal; response control; rendering of multi-channel response signal; force feedback

### 1 人和虚拟角色的互动

虚拟现实(VR)技术通过以计算机技术、人机交互为核心的高新技术生成逼真的视觉、听觉、触觉等多通道的、一定范围内的虚拟环境, 用户可以借助必要的设备以自然的方式与虚拟环境中的物体交互, 获得亲临等同真实环境的感受和体验。VR具有“3I”特性, 即沉浸感(immersion)、交互性(interaction)和构想性(imagination), 并且随着同人工智能技术的不断结合, VR系统表现出更多的智能性(intelligence), 并逐渐向4I发展。

VR技术在发展的早期阶段多用于军事、航空航天、医学、培训、工业仿真、城市规划等严肃的专业领域, 充分利用3I特性对现实世界进行高精度的模拟与呈现, 辅助用户研究分析, 解决复杂问题。随着Oculus Rift、HTC Vive和PlayStation VR等消费级

VR产品的发布, VR技术逐渐从专业应用领域走向消费市场, 教育、娱乐等个人应用领域进一步蓬勃发展。

最简单的VR应用是虚拟电影院, 在用户的视野前方设置一块可以播放电影的巨大屏幕, 虽然播放的资源仍然是原先的传统电影, 但是可以让用户享受到在电影院观影的感觉。360°视频则更进一步地不再固定用户的视角, 给予用户更多自主选择观看视角的自由。严格来看, 这两者不能算作真正意义上的VR应用, 因为并没有充分体现出VR的3I特性, 单单通过虚拟环境的360°环绕给

人仿佛“身临其境”的感官错觉, 缺乏空间感知和运动支持, 实现的沉浸感是粗糙而脆弱的, 交互性和构想性更是不曾体现。它们最多只能作为当前阶段缺乏成熟的制作VR互动体验的经验, 为方便广大消费者了解VR技术而做出的折衷。

在真正的VR互动体验中, 用户不再以一个旁观者的角度观看, 而是走进故事中, 亲身存在于虚拟世界里, 以第一人称视角对虚拟世界进行感知与理解, 通过四处走动加深对虚拟世界的空间感知, 同时通过自身活动影响周围的虚拟环境和虚拟角

收稿日期: 2017-09-26

网络出版日期: 2017-11-09

基金项目: 国家高技术研究发展计划(“863”)计划(2015AA016303); 国家重点研发计划(2016YFB1001401)

色的发展,从而更深刻地感受到踏入一个全新世界的真实感。人和虚拟角色的互动就可以归入这一范畴。如果需要实现复杂的、接近真实世界中人-人交互的人和虚拟角色的互动体验,需要众多相关技术的支持。

## 2 人和虚拟角色互动的关键技术

如图1所示,一个完整的人-虚拟角色交互环路及各环节中需要解决的关键技术包括:(1)虚拟角色对用户的多通道交互输入信号进行接收,即通过对用户的头部和手部进行空间定位,赋予虚拟角色感知用户当前位置和视线方向,以及打招呼、递东西等手部动作的能力;通过动作捕捉,对用户全身的动作进行跟踪;通过眼动追踪,更准确地了解用户注视点的变化;通过麦克风采集用户的语音输入。(2)虚拟角色根据用户的多通道交互输入信号,确定将要呈现给用户的多通道反馈输出信号,例如面对用户的打招呼动作,确定是否做出反应,以及做出怎样的反应。(3)虚拟角色将多通道反馈输出信号通过自身的表情、姿态、动作、语音等形式呈现给用户。

此外,在纯虚拟的交互基础上,可以进一步借助触觉反馈技术为虚拟角色提供触觉支持,使交互具有

更多的物理性,即用户在虚拟环境中看到虚拟角色的同时,可以在真实世界的相同位置处触摸到虚拟角色。

### 2.1 空间定位

为了给用户提供真实的包含视、听、触多通道信息的虚拟环境,需要实时检测用户头部的位置和视线方向,计算机能根据这些信息确定所要呈现的虚拟感官通道信息,并通过各通道的输出设备实时呈现。为了使用户在虚拟环境中体验到更高的沉浸感,同时赋予用户一定的手部交互能力,需要实时检测用户手部的位置和姿态并在虚拟环境中渲染,同时通过将手部与可交互的虚拟目标进行碰撞检测做出交互判断。

应用于VR领域的跟踪技术,通常以传感器为核心构建跟踪系统,根据选用的传感器种类不同,跟踪系统分为机械跟踪、电磁跟踪、超声波跟踪、光学跟踪和惯性跟踪。机械跟踪会极大地限制用户自身的运动,不适合人和虚拟角色存在丰富互动的应用。电磁跟踪和超声波跟踪易受工作环境中磁场、金属物体与刺激性声波脉冲的干扰,抗干扰性较差,且跟踪精度会随着跟踪范围的增大而迅速衰减,也不适合包含空间定位,允许用户在较大范围内走动交互的应用。而光学跟踪(包括广义的光学跟

踪技术,例如激光扫描空间定位技术)和惯性跟踪能够实现较大的跟踪范围,且两者相结合的跟踪方案能够取长补短,光学跟踪提供高精度的空间定位,同时对惯性跟踪中随时间推移产生的较大累积误差进行校正,惯性跟踪克服光学跟踪在遮挡、画面模糊时定位失败的问题,同时高采样率确保能实时跟踪目标的快速运动,保证交互的实时性,是当前技术阶段中主流的空间定位解决方案。

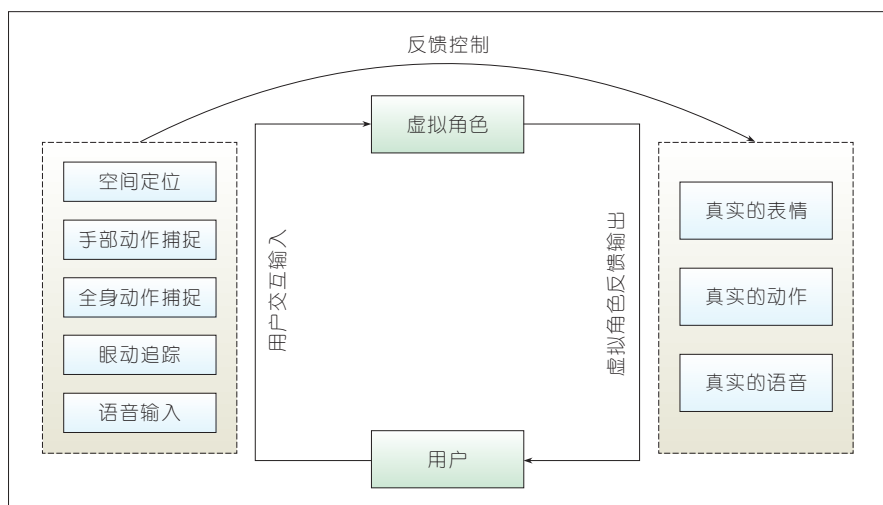
### 2.2 动作捕捉

通过对用户的手部进行空间定位,虚拟角色能够感知用户手部做出的简单动作,例如:招手、递送物体,并做出相应的回应,但是手柄实现的简单空间定位无法完全跟踪手部丰富自由度的运动,且在交互过程中持续地手持设备会干扰人-虚拟角色交互的自然性。为此可以采用手部动作捕捉技术,无需手持设备,裸手或者穿戴数据手套就能跟踪全手运动,与虚拟角色自然互动。此外还可以进一步通过全身动作捕捉来实现用户全身的运动跟踪。

动作捕捉技术从实现方式上分场景深度解析方案和可穿戴方案两种。场景深度解析方案通过光学传感器接收到的光信号来分析场景的深度信息,进而确定手部或者全身的位置和姿态,无需用户手持或者穿戴专用设备就可实现动作捕捉,最大程度地减少真实世界的干扰,使人和虚拟角色的互动更加自然;但是获取的人体骨骼运动较为粗糙,限制了交互的真实感。相比之下,可穿戴方案通过在用户身上多个关键点处固定传感器或者标志点,测量该点的位置变化或者弯曲程度,反算用户的身体运动,实现的动作捕捉更为精细。

### 2.3 眼动追踪

通过对用户的头部进行空间定位,虚拟角色能够利用屏幕中心粗略地估计用户的注视点,做出与用户视



▲图1 人-虚拟角色交互环路中的关键技术



线接触等交互。若是需要更准确地确定用户的注视点,以便更好地理解用户的交互意图,可以将眼动追踪技术<sup>[1]</sup>集成至头戴式显示器中,例如FOVE头戴式显示器通过嵌入两枚小型红外摄像头,采集被红外发光二极管照亮的人眼的图像,利用角膜反射法<sup>[2]</sup>计算用户的眼球位置。

眼动追踪技术还能捕捉、记录反映用户一定情绪和认知过程的眼部活动<sup>[3-4]</sup>,例如表征情绪状态变化的瞳孔缩放和与心理负荷息息相关的眨眼频率。综合分析这些反映情绪和认知的眼动数据和其他通道获取的用户输入,能更好地判断用户的情绪和交互意图,使虚拟角色做出更加合适的反应。

除此之外,利用眼动追踪技术获取的眼球运动数据,还可以实现模拟人眼视觉的视网膜中心凹渲染技术。该技术可以只对视域中央的画面进行高分辨率渲染,视域边缘采用逐渐降低的分辨率,大幅降低硬件计算负担,避免精细地渲染整幅画面耗费大量计算资源,导致渲染帧率下降,人-虚拟角色互动实时性变差<sup>[5-6]</sup>。

## 2.4 语音输入

人们倾向于将交互对象拟人化,面对一个虚拟角色,尤其是类人形时,会自然地期待它能表现出类似人类的行为,而语音交互作为人-人交互中最重要交互手段之一,非常有必要实现于人-虚拟角色互动中。

在交互过程中,虚拟角色将采集到的用户语音输入,先通过语音识别转化为相应的文本内容,再通过语义理解进行基于上下文的交互意图判断,同时虚拟角色还可以通过对音调、响度等声学特征和语音内容进行分析,判断用户在与之互动时的情绪变化<sup>[7-8]</sup>,综合分析用户的交互意图和情绪,做出更为合适的反应。

## 2.5 反馈控制系统

反馈控制系统的作用在于根据

获取到的多通道交互输入信号,确定将要呈现给用户的多通道反馈输出信号。传统方式是建立一个由交互输入到反馈输出的程式化的映射,一切按照预先写好的程序进行,一定的输入必然对应一定的输出,或者按照一定概率对应一系列输出中的一种,这种程式化的虚拟角色反馈极度单调、不自然,缺乏使人与之长期互动的吸引力。

为了建立丰富的、自然的虚拟角色反馈,感知-控制-行动模型(SCA)、并行转换网络模型(PaT-Nets)<sup>[9]</sup>、等多种虚拟角色行为控制模型被建立,且随着深度学习和大数据的不断发展,在不久的将来虚拟角色甚至能够以学习的方式自行建立起人-虚拟角色交互的反馈模型。

## 2.6 反馈信号渲染

反馈信号渲染的目标在于将包括表情、动作和语音在内的多通道反馈输出信号呈现给用户。为了使用户感到虚拟角色是真实的,吸引用户按照人-人交互的方式同虚拟角色互动,要求虚拟角色的表情、动作和语音都是接近真实的。

真实的表情和动作可以通过动作捕捉并录制获得,通过在动作录制者的脸上和身上粘贴或绘制标志点,并对标志点进行跟踪,即可利用跟踪数据来驱动虚拟角色做出同样的表情和动作,但是此方法获得动作和表情需预先录制,限制交互的丰富性。虚拟角色的表情和动作也可以通过深度学习进行训练,目前Google使用强化学习算法训练人工智能越过障碍物从起点跑至终点,已经成功地使人形模型自行学会了行走、跳跃等动作<sup>[10]</sup>,这种通过学习产生的动作和表情能够实时生成无需预先录制,同时有望做到非常接近真实的程度,但当前研究进展距离商业可应用还有一段不短的路要走。

传统的语音合成方案为参数化语音合成和拼接式语音合成,均利用

已有的声音进行重组来合成新的语段音频。该方式产生的语音能基本接近人类表达的流畅度,但是听起来不自然,且由于无法产生可以自适应变化的语调和语速来反映说话者的情绪,很难让人产生“我在跟一个人说话”的感觉。为了获得真实的语音,一方面可以针对交互场景预先录制,此方法同样存在交互丰富性受限的问题;另一方面可以引入学习的手段,例如Google的WaveNet<sup>[11]</sup>利用真实的人类声音和相应的语言、语音特征来训练卷积神经网络,使其掌握不同语音、语言的模式,能够实时合成出更加接近自然人声的语音音频,并且模拟一定的语调、情感和口音,但是距离让用户无法区分是机器合成还是真人讲话尚有很大差距。

## 2.7 力触觉反馈技术

上述的关键技术已涵盖人-虚拟角色交互的整个环路,但是无论是用户的交互输入还是虚拟角色的反馈输出都完全虚拟,看得见摸不着,可能会发生用户的虚拟化身穿过虚拟角色身体造成临场感中断的现象。为了避免出现此类视觉穿透现象破坏人-虚拟角色互动体验,一方面可以通过巧妙的方式拉开人和虚拟角色之间的距离,但是遏制了视觉穿透可能性的同时也可能给交互带来距离感;另一方面则可以在纯虚拟的交互基础上,借助力触觉反馈技术为虚拟角色提供触觉支持,避免视觉穿透的同时对视觉、听觉双通道交互进行触觉通道的扩展。

力触觉反馈技术从实现机制上分为主动式和被动式两种。主动式力触觉反馈设备包括场景/桌面式、手持式和可穿戴式3种;场景/桌面式设备固定放置于桌面上或者立于交互场景中,通过电机驱动操纵杆或者线绳的方式来输出三维空间中的虚拟作用力,由于需要用户一直与设备接触以感受其产生的作用力输出,会损伤用户与虚拟角色互动时的沉浸



感,并且有限的工作范围也严重限制了人-虚拟角色交互的自由性,故不适合用于人和虚拟角色的互动体验;手持式设备,顾名思义需要用户时刻持于手中,通常以手柄、手持道具类设备出现,通过振动触觉、气动等技术模拟作用于手部的力,同样存在持续接触干扰交互自然性的问题;可穿戴式设备通常以触觉衣、臂带和手套等形式出现,利用振动、气动、肌肉电刺激、挤压、力矩操纵等技术模拟力触觉作用于人体的感受,目前高精度的触觉分布模拟的计算难关还未突破,只能对和虚拟角色的握手、拥抱等触觉交互进行较为粗糙的实现。

被动式力触觉反馈则是通过跟踪真实世界中一个和虚拟角色近似1:1对应的实物,并在其上精确地叠加虚拟角色,使用户在虚拟环境中看到虚拟角色的同时,在相同位置处触摸到与虚拟角色对应的实物。该方案利用真实物体本身的属性提供力触觉反馈,真实感更高,并且不存在持续接触的问题,更容易实现和虚拟角色的握手、拥抱等触觉交互,但是需要在真实世界中存在一个类似于虚拟角色的实物限制了该类技术应用的灵活性。

当前阶段,选用被动方案为静态的虚拟角色提供力触觉反馈更为自然真实。而随着传感驱动装置的小型化集成技术更加成熟,随着对触觉这一感官通道的研究更加深入,对触觉的模拟更加真实,使用被动方案或可穿戴式方案为动态的虚拟角色赋予物理性变得可行,其中被动方案的实现更加自然,接近真实世界中的交互;而可穿戴式方案更加灵活,可以方便调整为不同的虚拟角色提供力触觉反馈。

### 3 结束语

在虚拟环境中实现同虚拟角色的互动需要解决3个核心问题:如何实时精确地采集用户的多通道交互输入信号;如何建立虚拟角色的由交

互刺激到反馈的“体现智能与情感的”映射;如何将虚拟角色的多通道反馈信号真实地呈现出来。就这3个核心问题,又需要解决一系列相关的关键技术。

对用户交互信号的采集是3个环节中最为依赖于硬件设备发展的一环。目前,空间定位和动作捕捉领域已有较多相对成熟的技术与产品,更多在于针对应用特性选择适合的方案,光学和惯性相结合的空间定位与动作捕捉能提供较大范围内高精度的空间定位和实时性高、无惧遮挡的动作捕捉,是当前适合于人-虚拟角色互动应用的成熟且优秀的方案。眼动追踪方面,目前已有FOVE、七鑫易维、Tobii等几家公司完成了眼动追踪技术到头戴式显示器的集成。近场语音识别更是在借助深度学习以后识别准确率有了实质性提高,已经达到了初期的商业可用的阶段。综上所述,目前已有技术已经能够比较完整地实现用户交互信号的采集环节;而虚拟角色的由交互刺激到反馈的“体现智能与情感的”映射建立和多通道反馈信号的真实呈现,目前还处于研究阶段,需依靠认知心理学、人机交互、人工智能等技术的进一步发展。

虚拟角色的触觉支持是对上述各环节实现的增强,避免发生视觉穿透现象影响交互沉浸感。目前触觉的发展相对视觉和听觉还有很大差距,并没有一个完善的触觉解决方案,使用和虚拟角色近似1:1的实物来提供被动力触觉反馈只是一个权宜之计。还需待传感驱动装置的小型化集成更加成熟,或者对触觉这一感官通道的研究更加深入之后,才能利用主动/被动方案为动态的虚拟角色提供更加灵活、真实的触觉支持。

#### 参考文献

- [1] DUCHOWSKI A T. Eye Tracking Methodology: Theory and Practice[M]. London:Springer, 2003
- [2] SIGUT J, SIDHA S A. Iris Center Corneal Reflection Method for Gaze Tracking Using

Visible Light[J]. IEEE Transactions on Bio-Medical Engineering, 2011, 58(2):411.DOI: 10.1109/TBME.2010.2087330

- [3] GAO Y, BARRETO A, ZHAI J, et al. Digital Filtering of Pupil Diameter Variations for the Detection of Stress in Computer Users[C]// Proceedings of the 11th World Multi-Conference on Systemics, Cybernetics and Informatics. USA:IEEE, 2007:30-35.DOI: 10.1109/TBME.2010.2087330
- [4] ISHIMARU S, KAI K, KISE K, et al. In the Blink of An Eye: Combining Head Motion and Eye Blink Frequency for Activity Recognition with Google Glass[C]// Augmented Human International Conference. USA:ACM, 2014: 15. DOI: 10.1145/2582051.2582066
- [5] PATNEY A, SALVI M, KIM J, et al. Towards Foveated Rendering for GazeTracked Virtual Reality[J]. ACM Transactions on Graphics, 2016, 35(6):179
- [6] GUENTER B, FINCH M, DRUCKER S, et al. Foveated 3D Graphics[J]. ACM Transactions on Graphics, 2012, 31(6):164
- [7] JUSLIN P N, SCHERER K R. Vocal Expression of Affect[J]. The New Handbook of Methods in Nonverbal Behavior Research, 2005: 65-135. DOI: 10.1093/acprof:oso/9780198529620.003.0003
- [8] SCHERER K R. Vocal Affect Expression: A Review and A Model for Future research[J]. Psychological Bulletin, 1986, 99(2):143
- [9] BADLER N I, WEBBER B L, BECKETT W, et al. Planning and Parallel Transition Networks: Animation's New Frontiers[J]. Center for Human Modeling and Simulation, 1995: 91
- [10] Google's DeepMind AI Just Taught Itself to Walk[EB/OL]. (2017-07-11)[2017-09-23]. <http://www.businessinsider.com/google-deepmind-ai-artificial-intelligence-taught-itself-walk-2017-7>
- [11] WaveNet: A Generative Model for Raw Audio[EB/OL]. [2017-09-23]. <https://deepmind.com/blog/wavenet-generative-model-raw-audio/>
- [12] SHOJI M, MIURA K, KONNO A. U-Tsu-Shi-O-Mi: the Virtual Humanoid You Can Reach[C]// ACM SIGGRAPH 2006 Emerging technologies. USA:ACM, 2006: 34

#### 作者简介



翁冬冬,北京理工大学光电学院副研究员;主要研究领域为虚拟现实、增强现实与新型人机交互技术;先后主持和参加国家级项目20余项;已发表被SCI/EI检索论文30余篇,获授权专利5项。



薛雅琼,北京理工大学在读硕士研究生;主要研究方向为虚拟现实与新型人机交互技术。

# 支持大规模视频融合的混合现实技术

## Massive Video Integrated Mixed Reality Technology

周忠/ZHOU Zhong  
孟明/MENG Ming  
周颐/ZHOU Yi

(北京航空航天大学, 北京 100191)  
(Beihang University, Beijing 100191, China)

**虚**拟现实(VR)是一个新兴的科学技术领域,该技术建立人工构造的三维虚拟环境,用户以自然的方式与虚拟环境中的物体进行交互,极大地扩展了人类认识、模拟和适应世界的能力。

虚实融合(MR)技术将虚拟环境与真实环境进行匹配合成,降低了三维建模的工作量,并借助真实场景及实物提高用户的体验感和可信度。随着当前视频图像的普及,MR技术的探讨与研究更是受到关注。

视频融合技术利用已有的视频图像,将它们融合到三维虚拟环境中,可以实现具有统一性的、深度的视频集成。该技术最早可追溯到1996年Paul Debevec<sup>[1]</sup>提出的一种视点相关的纹理混合方法,即实现了不在相机视点的真实感漫游效果,但仍在很多方面存在难点。

### 1 MR的技术特点

现在业内普遍认可从真实世界到虚拟环境中间经过了增强现实与增强虚拟环境这两类VR增强技术<sup>[2]</sup>,

收稿日期: 2017-09-23

网络出版日期: 2017-11-06

基金项目: 国家高技术研究发展(“863”)计划(2015AA016403); 国家自然科学基金(61572061、61472020)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2017) 06-0006-004

**摘要:** 视频融合是视频监控控制相关应用领域关注的虚拟现实(VR)热点问题之一。根据实现虚实融合的维度不同,将相关研究方法分为4类:视频标签地图、视频图像拼接、视频叠加到三维场景、视频融合到三维场景。介绍了VR视频融合技术方面的相关工作,认为虚实融合技术正在快速发展中,虚实信息的可视关联对于人类认知和人工智能都已表现出显著的提升作用,未来将可能作为一种基础的地理信息资源来提供,有着重要的应用价值。

**关键词:** VR; 混合现实; 虚实融合; 视频融合; 视频监控控制

**Abstract:** Video fusion is a hot topic in some virtual reality (VR) application areas such as video surveillance. According to dimension number, related works in virtual-reality integration are surveyed in this paper in four categories: map with video tag, video image stitching, video overlapping on 3D scenes, and the video fusion in 3D scenes. Then our work in VR video fusion is presented in this paper, especially the new idea of video model and mixed reality rendering. The virtual reality technology is developing rapidly, and the visual correlation of the virtual and real environments has shown significant improvement on both human cognition and artificial intelligence. In the future, the video model may be provided as a fundamental geographic layer resource, which has significant values for various applications.

**Keywords:** VR; mixed reality; VR integration; video fusion; video surveillance

混合现实则是包含这两类技术及其融合。

增强现实技术与增强虚拟环境技术,可分别形象地描述为“实中有虚”和“虚中有实”。增强现实技术通过运动相机或可穿戴显示装置的实时连续标定,将三维虚拟对象稳定一致地投影到用户视口。增强虚拟环境技术通过相机或投影装置的事先或实时标定,提取真实对象的二维动态图像或三维表面信息,实时将对象图像区域或三维表面融合到虚拟环境中。两项技术的MR方式如图1和图2所示,图中虚线对象代表虚拟环境对象,实线对象代表真实对象或其图像。

随着VR技术的发展,其与现实世界正趋向于深度融合,一些技术开始兼具“虚中有实”和“实中有虚”这两种模式。20世纪70年代电影《星球大战》中展示的全息甲板是科幻中的终极理想状态,但U.C.Berkeley提出的Tele-immersion远程沉浸系统、微软的Holoportation、Magic Leap所设计的光场头盔显示原型等已经具备了这种深度MR特点。

### 2 视频融合技术分类

早期的视频融合技术只能做到将图片向地面或简单立面映射,现在的新技术则已经能够快速将视频实时地映射到复杂的三维模型上。根

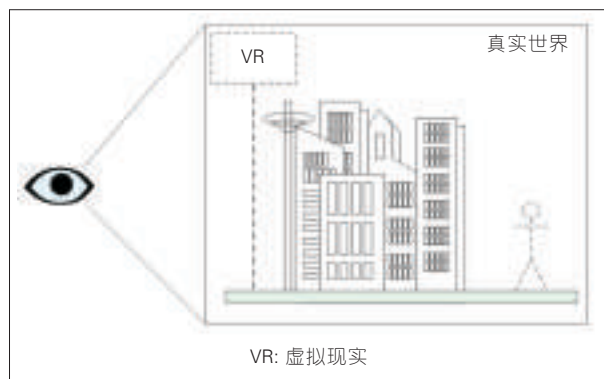


图1  
增强现实技术“实中有虚”

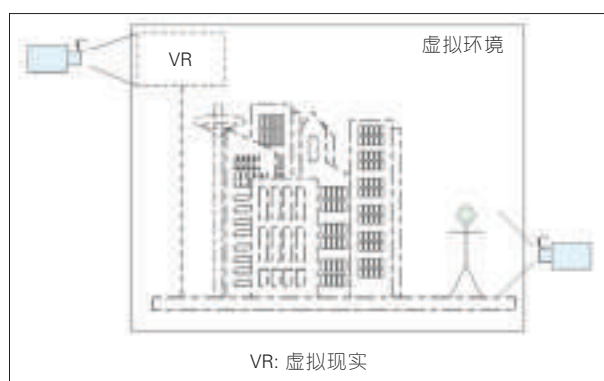


图2  
增强虚拟环境技术“虚中有实”

据实现 MR 的维度不同,可将相关方法分为 4 类:视频标签地图、视频图像拼接、视频叠加到三维场景,视频融合到三维场景。需要说明的是:这些技术针对的是易于获取的普通摄像头视频,FreeD 等基于多视图几何的三维重建方法不在讨论中。

### 2.1 视频标签地图

基于视频监控控制对于多地点视频有效组织的需求,采用视频标签与地图的索引集成,实现在地图上放置和观看视频。美国 FX Palo Alto 实验室、美国三菱电机研究所、韩国电子通信研究院和法国原子能署 CEA LIST 等机构在该方面开展了工作。

美国 FX Palo Alto 实验室提出了用于室内的多相机实时监测控制系统,动态物体跟踪系统(DOTS)<sup>[3]</sup>,该系统通过对平面设计图的特征分割,获得系统中视频与位置信息的关联,直接向用户提供一种视频分析结果的展示手段。美国三菱电机研究所 MERL 的 Ivanov 等也实现了用于室内

居住环境可视化的类似系统,通过加入运动传感数据等信息,为建筑系统的设计人员和管理人员提供整栋楼的设计人员和管理人员提供整栋楼的上下文信息<sup>[4]</sup>。视频标签索引的融合方法除了用于监测控制系统,还广泛应用于地理信息系统(GIS),它们通过建立提取的视频内容与数据库中 GIS 数据的对应关系,进行视频与 GIS 的融合。韩国电子通信研究院的 Tae-Hyun Hwang 等基于这个思想,将视频和虚拟场景中的 meta 元素提出,建立了基于 meta 元素的视频与 GIS 系统的关联<sup>[5]</sup>,通过简单的点击地图查询即可在手机上直接访问视频。此外,法国原子能署 CEA LIST 的 Gay-Bellile 等通过增强现实 AR 和相机追踪的方式建立了实时视频与 2D GIS 的关联<sup>[6]</sup>。

### 2.2 视频图像拼接

视频图像拼接是指将空间上可配准、相互之间具有足够重叠区域的图像序列经过特征对齐、空间变换、重采样和拼接合成之后形成宽视角

甚至全景图像的方法。经过数十年的发展,图像拼接算法目前已经比较成熟,全景相机出现了低成本、消费级的特点,利用鱼眼相机来降低对相机数量的要求,达到小型化成为新的“爆点”。图像拼接主要针对的是窄基线相机图像序列,宽基线相机不具备统一的单应性,特别是遮挡大大影响重叠区域的匹配,其图像拼接是目前研究的难点。

### 2.3 视频叠加到三维场景

视频叠加到三维场景的方法以 2D 和 3D 特征注册为基础进行虚实融合,允许用户在相机视点的转移路径上观看叠加的结果,其效果优于视频标签地图方法。

塞尔维亚利兹大学 CG&GIS 实验室中 Milosavljević 等提出基于 GIS 增强的视频监控控制系统,将视频窗口叠加到 3D 模型视图窗口的上方显示,在 3D GIS 环境中建立视频与空间信息的位置关联<sup>[7-8]</sup>。美国微软公司的 Snively 等于 2006 年提出了照片旅游系统<sup>[9]</sup>,该系统利用对互联网上照片集的匹配,重构一个稀疏的三维点云场景,通过图像变换的渲染方法进行照片浏览。荷兰代尔夫特理工大学的 Haan 等人受到 Snively 等共平面视点转移方法的启发,于 2009 年面向监测控制系统提出了第一人称式的场景导航方法<sup>[10]</sup>,通过交互地在场景中放置画布的方法完成视频的注册,进而通过动态视频嵌入实现导航,缓解了视点移动时视频间重叠区域的视觉差异现象。

### 2.4 视频融合到三维场景

视频与三维场景的融合方法,实质是将相机捕捉的视频图像,以纹理的方式实时注册到虚拟环境中,达到增强虚拟环境的效果,能够允许用户从非相机虚拟视点观察融合结果。这种方法比前几种融合方法在视点可选范围上进一步扩大,实际上是从另一个角度解决了宽基线相机以及



无重叠视域相机的图像拼接问题。但这类技术仍会存在一些难以克服的问题。

在 IEEE VR 03 上,南加州大学的 Ulrich Neumann 等人<sup>[11]</sup>系统阐述了增强虚拟环境的概念,实现了随着图像数据变化的动态三维模型效果,解决了非相机视点下贴图扭曲现象<sup>[12]</sup>。在 ACM MM 10 上,麻省理工学院的 DeCamp 等人<sup>[13]</sup>设计了一套用于智能家居的沉浸式系统 HouseFly,通过鱼眼相机的三维融合,让用户可以漫游于掀顶式楼宇。在 ISMAR 09 上,佐治亚理工学院 Kihwan Kim 等<sup>[14]</sup>提出基于动态信息增强 Google Earth 等航拍地球地图的方法,提出对视频进行分类处理和增强显示的方法。2012 年国立台湾大学的 Chen 等人<sup>[15]</sup>建立了 GIS 辅助的可视化框架,融入了多分辨率监测控制策略,以固定视角的相机提供低分辨率图像,球基相机根据用户交互提供兴趣区的高分辨率图像。

### 3 MR 技术进展

近几年,我们在这方面开展了一系列工作,主要特色是将图片建模技术用到 MR 中,以得到准确的虚实对齐效果。其中图片建模技术是利用二维图片恢复场景三维结构的数学过程和计算技术,这一技术能够很容易地达到虚实融合过程中对三维模型精度的高要求,克服了视频投影本身带来的二三维深度不匹配问题。

#### 3.1 基于图片建模的视频模型

该方法的核心部分是一种快速建模视频背景的交互式方法,使用体元和场景树来描述图像中各点之间的建模关系,首先针对单幅图像,在图像几何分析的预处理基础上,进行图像与场景模型的三维注册,然后提出了一种体元的定义,支持交互式的方式进行基本几何结构的恢复,实现单幅图像场景的视频模型生成。场景树结构示意如图 3 所示。

监测控制场景中存在大量相机

视频区域重叠度很小的情况,现有基于多视图的建模方法不能适用。进一步针对低重叠度图像序列,使用点线联合的匹配方法进行新图像与现有视频模型的注册,用户可以进一步进行新图像场景的结构建模,最终链式地匹配和注册更多的图像场景结构,如图 4 所示。

在此基础上,我们定义了一种基于单幅照片建模生成的视频模型,它描述了该照片对应的三维几何结构,

可以供二次开发使用。

#### 3.2 基于视频模型的 MR 方法

我们提出了基于视频模型的 MR 方法,针对每个视频创建对应的视频模型<sup>[16]</sup>,然后通过纹理投影<sup>[17]</sup>和阴影投影方法将视频与其模型进行融合。方法整体流程如图 5 所示,分为两个阶段:(1)预处理阶段。提取视频的背景帧进行交互式建模,得到简单的视频模型,然后与三维场景模型

图 3 场景树结构示意图

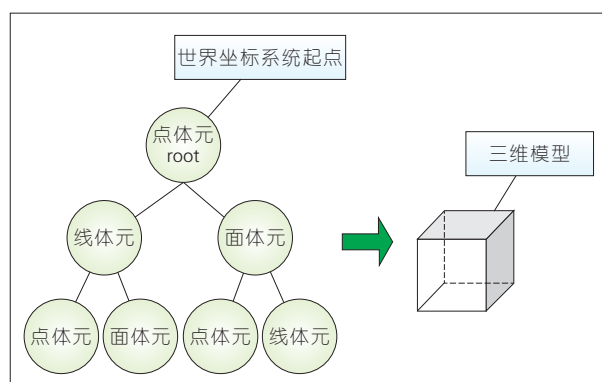


图 4 某火车站 6 幅相机图片的链式建模结果

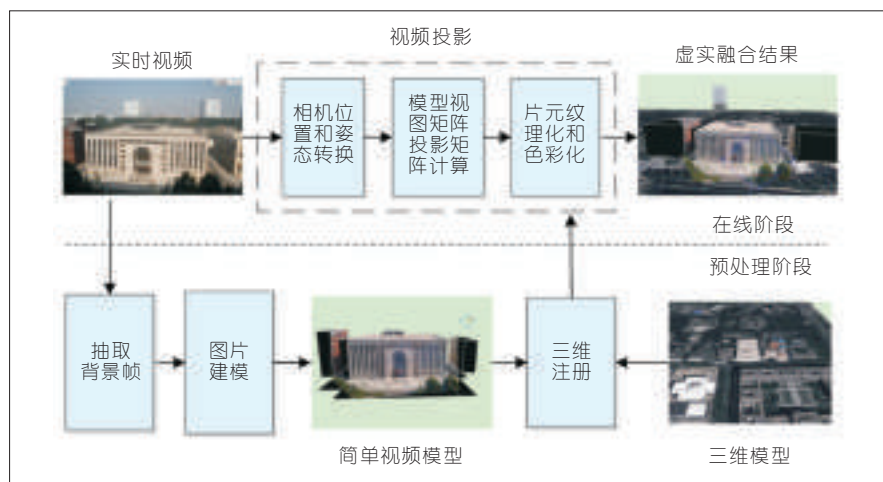
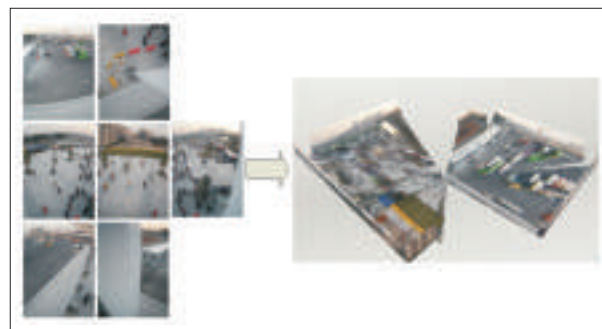


图 5 基于视频模型的 MR 流程



进行注册;(2)在线阶段。该阶段与直接投影的融合方法类似,但不再执行遮挡测试。

### 3.3 MR 场景中的自动路径规划方法

由于MR中视频中的动态物体并未三维化,这类技术不可避免存在偏离视点会出现画面畸变的现象。研究视点变化与画面畸变之间的关系,我们给出了一种基于视频投影中的画面畸变的视点质量评价方法,进一步提出了一种MR场景中的自动路径规划方法,来尽量减少畸变现象。

### 3.4 支持大规模视频融合的视频监测控制

以上技术被应用于支持大规模视频融合的视频监测控制,如图6所

示,各建筑模型是手工建模得到的精细模型,视锥区域是视频模型。通过空间划分和瓦片关联,可以很容易地扩展到大规模的视频模型场景。

## 4 展望

随着视频监控控制在公共安全、交通等领域的广泛应用,多相机监测控制网络中相机数量日益增多,MR技术将现实世界中大量的相机视频进行整合,提供与真实世界具有几何结构一致性的统一视图,解决的是人类“认知”的问题。另一方面,实际上不同位置的相机在内容上也很难关联分析,对于计算机的智能分析能力也带来了巨大的挑战,这同样需要MR的信息支持。

在MR技术的基础上,我们正在

研究多相机拓扑中的几何-语义联合理解与关联问题,研究并实现了一种基于几何-语义结构分析的多相机场景拓扑连通图的构建方法。以多相机拍摄的监测控制视频为输入,如图7所示,解析相机图像的语义信息与基本几何结构,通过分析目标在不同相机之间的转移状态,计算各相机区域之间的连通概率,从而建立多相机间的拓扑连通图。

图8中展示了使用我们的算法在Duke<sup>[18]</sup>数据集上恢复出的相机拓扑连通图,各视频图像和地图上对应的语义区域进行对齐,和原始场景中的拓扑连通关系相符。这种细粒度的相机视频融合方式可以很容易地作为一个地图图层推广到大规模GIS系统中,可以从根本上解决海量的视频碎片化问题。

MR技术正在快速发展中,这种虚实信息的可视关联对于人类认知和人工智能都已表现出显著的提升作用,未来作为一种基础的地理信息资源来提供,有着重要的发展意义。

### 参考文献

- [1] DEBEVEC P E, TAYLOR C J, MALIK J. Modeling and Rendering Architecture from Photographs: A Hybrid Geometry-and Image-Based Approach[C]//Proceedings of the Annual Conference on Computer Graphics and Interactive Techniques. USA: ACM, 1996:11-20. DOI: 10.1145/237170.237191
- [2] AZUMA R, BAILLOT Y, BEHRINGER R, et al. Recent Advances in Augmented Reality[J]. Computer Graphics and Applications, 2001, (21): 34-47
- [3] GIRGENSOHN F, SHIPMAN T, TURNER, et al. Wilcox, Effects of Presenting Geographic Context on Tracking Activity Between Cameras[C]//Conference on Human Factors in Computing Systems. USA: California, 2007: 1167-1176
- [4] IVANOV Y, WREN C, SOROKIN A, et al. Visualizing the History of Living Spaces[J]. IEEE Transactions on Visualization and Computer Graphics, 2007, 13(6):1153-1160. DOI: 10.1109/TVCG.2007.70621
- [5] HWANG T H, CHOI K H, JOOL I H, et al. MPEG-7 Metadata for Video-Based GIS Applications[C]//Proceedings of IEEE International Geoscience and Remote Sensing Symposium. USA: IEEE, 2003, (6): 3641-3643. DOI: 10.1109/IGARSS.2003.1294880
- [6] GAY-BELLILE V, LOTHE P, BOURGEOIS S,

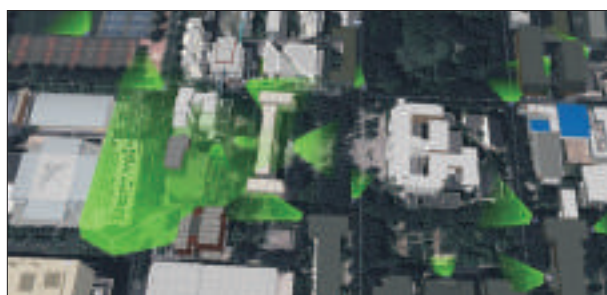


图6  
多路相机视频的融合概览



图7  
Duke数据集的相机关系



图8  
相机图片覆盖区域及拓扑连通

下转第36页

# 增强现实中的空间深度感知问题

## Spatial Depth Perception in Augmented Reality

陈东义/CHEN Dongyi

(电子科技大学, 四川 成都 611731)  
(University of Electronic Science and  
Technology of China, Chengdu  
611731, China)

多年来,对头戴式增强现实(AR)的研究得到了诸如手持式AR和投影摄像机系统等新平台研究的补充。随着手机应用的快速发展,AR应用概念已逐渐成为市场新兴增长点。然而,研究人员和从业者仍在尝试解决AR设计开发中的许多根本问题。目前,众多研究人员正在解决由跟踪注册算法引起的定位精度问题,但在AR应用中正确地感知仍然是AR应用中的一个关键挑战。这些障碍可以追溯到增强显示中相互关联的深度感知和照明感知问题,或是与环境感知有关的问题。这些问题均可能会导致场景感知失真和深度感知失真的情况。虽然其中的一些问题源于技术限制,然而许多是由于有限的理解或显示信息方法的不足所造成的。

感觉刺激的识别和解释是一个复杂的构造过程。我们经常从观察到的环境中获得不同的线索,并尝试匹配那些线索。线索提示可以是互相重叠,或者是互相冲突的。其中重要的是,不正确的感知通常是线索提示冲突的结果。在文中,我们仅关注与视觉感知相关的问题,而忽略与其

收稿日期: 2017-09-23  
网络出版日期: 2017-11-02

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2017) 06-0010-004

**摘要:** 在增强现实领域中,用户关注点已经转移到精确的虚实叠加感知和空间理解上,并且存在诸多导致用户对增强内容产生不正确感知的因素。提供了一种感知问题分类方法,将信息增强作为一个感知理解管道从而对各个环节分别阐述。主要分析了增强现实中影响空间深度感知的因素,并为改善感知问题提出了参考研究方向。

**关键词:** 场景感知; 空间深度感知; 增强现实

**Abstract:** For the users of augmented reality, concerns have shifted to precise sense of virtual superposition and spatial understanding. And there are many factors that lead to the incorrect perception of enhanced content. In this paper, a method of perceptual question classification is presented, and information enhancement is used as a perception and understanding pipeline to organize each link separately. It mainly analyzes the factors that influence the spatial depth perception in the augmented reality and puts forward the reference research directions for improving mentioned perception problems.

**Keywords:** scene perception; spatial depth perception; augmented reality

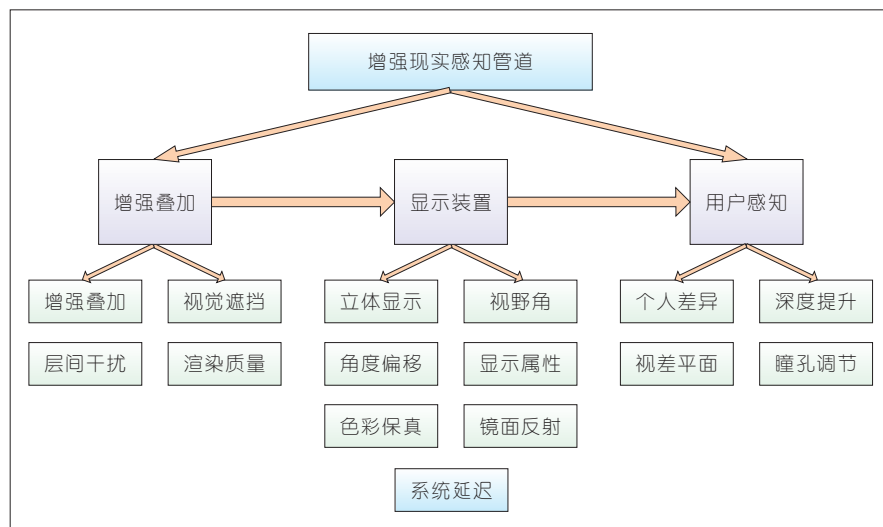
他感知的相互作用。感知问题涉及对生成的虚拟世界以及现实世界在观察和解释时出现的一些问题,不仅可能由真实信息和虚拟信息的组合引起,也可能源自于现实世界本身的表现。

针对深度感知失真的问题,目前的研究大多是针对深度相关线索进行修正以达到增强的目的。Christoph等人<sup>[1]</sup>通过增加景深和选择性的局部对比度提升来对单目深度线索进行增强。Athena等人<sup>[2]</sup>指出运动视差等动态透视线索能够提升深度感知。Chiuhsiang等人<sup>[3]</sup>调查了使用立体眼镜观看近场增强现实虚拟目标的中心距离感知精度,指出当目标是垂直放置,靠近观察者位置并且目标间隔较小时精度会有提高。Ernst等人<sup>[4]</sup>详细探讨了增强现实应用过程中与

感知相关的问题,并按各种感知类别给出了总结。Chen等人<sup>[5]</sup>给出了一种感知导向的深度线索增强方法,该方法展示了一种非线性视差映射,能增加观察者关注的深度范围。刘自强<sup>[6]</sup>研究了增强现实中深度一致性问题,并提出了一种基于双目摄像机的深度计算方法来解决深度一致性问题。郑毅<sup>[7]</sup>则介绍了增强现实应用中存在的虚实遮挡问题。相比前述工作,文中我们提供了一种感知问题分类方法,将信息增强作为一个感知理解管道从而对各个环节分别组织阐述,图1给出了感知理解管道中存在的因素的分层关系。

### 1 增强现实中感知问题的来源

我们可以从影响感知的因素入



▲图1 增强现实感知理解管道的影响因素分层关系

手,了解增强显示时用户对内容不正确感知的几个问题。对于某一些任务,不正确的增强现实感觉可能没有影响,而对于其他一些任务,这类问题是至关重要的。这些问题大致可分为3类:

(1)场景扭曲和抽象。场景增强可能会出现扭曲和部分抽象变形,使得对象识别、尺寸感知、对象分割或对象内部关系的感知变得困难。

(2)深度失真和物体排序。与上一个问题相关,不正确的深度解释是AR应用程序中最常见的感知问题。AR中的深度是指第一人称视角,视野中的物体与重叠的信息之间在空间关系上的解释和相互作用。深度失真和物体排序问题使用户无法正确地将重叠的信息匹配到现实的世界中。

(3)能见度。用户可能无法查看内容本身,主要是由屏幕问题引起的,例如:屏幕大小、屏幕反射和亮度,或环境干扰中颜色和纹理图案。

AR中的目标是将真实世界对象和数字内容之间在感知上正确连接,正确地解释真实和虚拟对象之间的空间关系。不正确的深度感知是AR应用程序中最常见的感知问题,妨碍了第一人称视角下视图中的真实对象与叠加(嵌入)信息之间的空间关

系理解。

### 1.1 增强叠加

增强是指在光学图像或视频图像上注册数字内容,并可能因技术水平限制而产生一系列相关的深度感知问题。

(1)注册错误。精确注册依赖于跟踪设备的正确定位和方向信息(姿势)。当前的手机具有相对不精确的位置和方位传感器,导致跟踪精度和方向测量存在显著漂移。所需的跟踪精度取决于所观察对象的环境和距离要求:较低精度跟踪对于距离较远的大型环境中的物体可能是可以接受的,即使产生偏移也不太明显,而对增强设备附近物体准确地增强则显得较为困难。

(2)视觉遮挡。物体视觉上被遮挡。与遮挡相关的主要问题是前景和背景的分割不正确:需要在特定对象后面呈现对象,而不是出现在特定对象的前面,否则会导致不正确的深度排序,使得对象看起来像它们不属于当前场景。自AR出现以来,研究人员试图使遮挡或隐形物体再次可见,主要是利用称为X射线视觉的叠加显示技术,允许用户通过前面的对象来查看遮挡对象。然而,X射线视觉也容易发生深度排序问题,因为重

叠的顺序是相反的。

(3)层间干扰。环境物体的排列组合模式会限制增强感知和识别。根据环境背景和增强后的特征,感知干扰可能发生在图形相交或者是视觉合并的地方,并因此导致前景背景的视觉解释问题。这些特征受到图案的方向、透明度、密度和规律性以及使用的颜色方案的影响。另外,前景背景图案问题与在多层AR系统中发生的问题有关,其中多层AR是指多个虚拟对象被渲染在彼此之上。这带来的问题是容易出现层间干扰,这取决于虚拟对象的数量及各自的不透明度。一旦层数太大,对象可能会重叠,这可能会降低环境增强信息的可读性。

(4)渲染质量。渲染质量定义了屏幕上显示数字对象的保真度。在保真度和数字再现图形中的深度判断之间没有发现直接的关系。除了渲染质量之外,照明可以影响增强物体的保真度及其正确的感知。在分辨率和清晰度之间的差异可能导致不正确的立体感知。

### 1.2 显示装置

显示装置向用户显示增强的环境,并且像其他阶段一样可以引起空间感知问题。大多数问题可能与屏幕相关联,但是还有一些问题来自于处理器和图形单元。

(1)立体显示。主要存在于头戴显示器(HWD),典型的问题包括实际和预设瞳孔间距的差异、视力和对比效应,对齐和校准问题以及与瞳孔汇聚有关的问题。然而,当使用HWD来查看完全人工合成的虚拟环境时出现的一些感知问题可以用AR来减轻。目前来说,立体显示问题对于手持设备和投影摄像机系统来说还不重要。

(2)视野角(FOV)。FOV是指可观察世界的角度。在视频透视显示中,FOV显然限制了现实世界中可以看到多少内容。尽管人类视网膜视



觉分辨率不到 $1^\circ$ ,但人类还依赖于非焦点的视力,而有限的FOV使许多视觉任务变得非常困难。目前研究来看有限的FOV不一定会导致深度估计错误。在光学透视和手持设备中,问题则变得更加复杂,因为信息空间不再是统一的。人类具有超过 $180^\circ$ 的水平FOV,而视频透视HWD通常支持 $30^\circ \sim 100^\circ$ 水平FOV(尽管有些达到近 $180^\circ$ )。光学透视显示器和手持设备使用相对较小的FOV。在一些光学透视显示器中,光学器件无框架或被非常薄的框架包围,用户可以观察现实世界中FOV的大部分,而不是重叠的图形:用户以正确的尺度看现实世界两部分。同样,大多数手持式视频浏览显示器允许用户查看显示器边框周围的真实世界。然而,在这些显示器中,相机使用的宽FOV镜头与显示器中心的偏移相结合,通常会在小且不正确缩放的增强视图中产生明显感知差异。

(3)角度偏移。HWD是直接放置在眼睛的前方,因此在被观察的现实世界和被看见的增强显示之间通常存在相对小的角度偏移。这种偏移还会通过由连接到显示器的相机引起可能的进一步偏移强化。根据设备重量,此角度偏移量可能会随着时间的推移动态增加。偏移影响了用户对增强后真实世界的间接观察。这种问题也可以称之为对齐问题,用户不能很容易地直接看到现实世界中内容之间的关系,以及在屏幕显示时两者的比较关系,这需要用户经历一个困难的心理旋转和缩放适应过程。另外,相机相对于显示器的角度和放置位置可能会使视角偏移进一步出现恶化。

(4)显示属性。当与环境光混合时,显示亮度和对比度会影响内容的可视性。显示亮度是指显示器的亮度,大约在 $250 \sim 500$ 烛光/平方米( $\text{cd}/\text{m}^2$ )之间变化。对比度可以由显示能够产生的最亮颜色(白色)的亮度与最暗颜色(黑色)的亮度的比率

来表示。特别是在户外应用中,由于环境光的影响,显示器对比度受到很大限制,环境光将降低显示屏的对比度,这导致人类视觉系统无法区分更细微的颜色。

(5)色彩保真。在AR中,色彩保真度是指现实世界与屏幕上显示的颜色相似度。而在打印介质中,不同设备的颜色表示之间存在标准转换,目前的做法通常不涉及采样的真实世界颜色之间的映射。色彩空间转换使用色域映射方法将颜色转换为可在给定设备上显示的范围。自然色彩的全系列不能忠实地表现在现有的显示器上,特别是不能再现高度饱和的色彩。这可能扭曲基于颜色的感知线索,并影响对颜色编码信息的解释。户外环境中的色彩保真度是一个非常复杂的问题。更改户外条件会影响光学透视显示器,因为在视频透视显示器中,现实世界和重叠式显示在相同的色域中。在投影机相机系统中,投影面上的纹理变化可能会干扰色彩表现。

(6)镜面反射。反射是扰乱AR内容感知最重要的影响因素。在HWD中,闪亮的物体可能会扰乱感知。在具有暴露屏幕的手持AR系统中,内容可能几乎看不见,这通常取决于环境条件,例如朝向太阳或人造灯的亮度和方向以及被反射的物体之间的相对观察位置。反射也引入了多个视差平面的问题,因为反射物体通常与屏幕内容不同。当内容投射在镜面反射面上时,投影机相机系统中的反射也是一个问题。

(7)系统延迟。延迟涉及捕获或显示内容的延迟,并且直接取决于显示设备能够产生的每秒帧数。大多数情况下取决于处理器和图形卡的性能,并与内容的复杂性直接相关。硬件性能可能会影响内容的捕获和渲染质量。延迟包括动态注册延迟、相机图像更新和渲染覆盖延迟。延迟影响用户与内容的直接互动。延迟对正在观看的内容的感知影响较

小,例如:许多AR应用程序涉及静态或至少缓慢变化的内容,渲染速度对这类内容影响程度较小。依赖于快速图形生成(如游戏)或依赖于重叠图形的灵巧运动类应用程序的可用性也存在受到延迟引起的感知限制的影响。

### 1.3 用户感知

用户是感知流水线的最后阶段,受到各种平台的不同影响。

(1)个人差异。在显示屏幕上呈现的数字内容的最终感知受到用户之间个体差异的高度影响。这些差异可能需要对我们表示信息的方式进行明显的修改,例如图标或文本。个体差异体现在用户感知细节的能力,包括眼睛视力、色觉能力和空间分辨能力等差异。

(2)深度提示。深度线索在解释增强内容中起着至关重要的作用。图形深度线索是图像中的特征,给人以物体的不同深度的印象。这些线索包括遮挡、视野中的高度、相对尺寸、空中视角、相对密度、相对亮度和阴影。动态深度提示可以通过视点位置动态变化而获得深度信息,例如相对运动视差和运动视角。生理深度提示来自眼睛的肌肉控制系统,包括会聚(相反方向的眼睛旋转以固定在一定深度)、调焦(通过改变眼睛的晶状体形状来抵消模糊)和瞳孔直径(通过改变眼睛的景深来抵消模糊,但是也受环境照明水平的影响)。最后,双目视差通过组合由眼睛提供的场景的两个水平偏移视图来提供深度线索。在所有这些深度线索中,遮挡是最主要的,这驱动了AR中最普遍的深度提示问题:空间距离上不正确的深度排序。当只有有限数量的深度线索可用时,这个问题变得更加成问题,这可能导致物体深度不足,甚至出现矛盾。

(3)视差平面。现实世界和虚拟物体都会有不同的双目差异,并导致与视差平面和视差区域相关的知觉



问题。在双视 AR 系统中,经常会出现深度差异:增强信息存在于一个区域,而现实世界存在另一个区域。由于这些区域处于不同的焦点深度,所以用户可能需要在这些区域之间持续地切换其聚焦(眼睛旋转)来比较内容,或者因为注意力被吸引到其他区域。此外,在 HWD 中,在视锥的前平面中呈现的用户界面元素与实际 AR 内容之间可能存在深度偏移。当用户经常需要使用界面时,会导致这些不同深度平面之间的定期切换,这将导致视觉疲劳。

(4)瞳孔调节。立体显示器的用户通常会遇到所谓的会聚调节冲突。当眼睛收敛于提供给左眼和右眼两个空间偏移视图中的观察物体时,就会发生这种冲突。人类视觉系统一般具有容忍这种不匹配的能力,但深度感知会被扭曲。在单视镜和投影摄像机系统中,所有内容都在单一深度平面上显示和观看,因此不存在这个问题(以损失会聚调节深度线索为代价)。投影机相机系统可能没有焦平面(差异)问题。然而,由于除激光投影机之外的所有焦点都具有固定的焦点深度,因此不同的多个非连续深度面将产生深度感知问题。

## 2 增强现实中感知问题的解决思路

研究人员提出了各种方法来解决增强现实中感知渠道的问题。在本节中,我们概述除用户之外的主要解决思路。

### 2.1 环境分析

物体在杂乱场景中的增强通常需要使用一种视觉辅助机制,来对物体进行独特的增强绑定。具体来说,当增强信息重叠在几个对象上时,正确的对象布局判断可以辅助正确的增强叠加,可以使用计算机视觉分析方法分离前景和背景层。增强也可能需要一些颜色上的拮抗,以避免增强信息与被覆盖物体的出现视觉

融合现象。然而,如果对象颜色发生改变,则对象和增强信息可能会出现分离的情况。

### 2.2 捕获设备

捕获受到镜头和摄像机参数两者的限制。然而,解决由镜头引起的问题通常很困难,目前只有少数解决方案存在。对于广角镜头,观察者不一定会注意到广角引入的扭曲,而双重视图条件下真实世界的视图可以纠正包括由低分辨率引起的线索提示之间的潜在冲突。然而,双重视图的情况会增加视差平面切换和认知负荷,并且当对象快速移动时无效。

### 2.3 增强提示

对于与遮挡对象相关的问题,大多数研究人员通过避免对象裁剪来纠正深度排序问题。此外,已经出现了许多技术能够改善被遮挡对象的 X 射线可视化。正确的光照渲染也有助于与遮挡物体相关的深度排序,例如使用阴影信息能够提供重要的深度提示。正确的光照渲染也可以使场景深度更加可信,防止增强的虚拟物体看起来像一个突兀的平面模型。此外,可以使用诸如网格或深度标签(距离指示器)的人造深度提示来直接向用户提供深度线索判别。

### 2.4 显示设备

目前的显示技术水平不断提高,预计将出现新的显示技术,以更好地应对亮度和对比度问题。尽管显示内容在阳光充足的条件下还难以清晰呈现,但目前显示技术经常使用背光和抗反射涂层使内容更加可见。可以通过增加涂层来最小化反射光,但这也可能降低屏幕的亮度。类似地,在 HWD 壳体的内部应避免设计反射表面。匹配环境光照的动态范围同样是一个问题。飞机中使用的平视显示器可以匹配该动态范围,并且基于激光的显示技术具有匹配该动态范围的应用潜力,但是目前还没

有被广泛使用。

## 3 结束语

文章中,我们概述了主流 AR 平台上影响用户正确感知增强信息的主要问题来源,特别是影响空间深度感知的一些因素。总而言之,除去对内容主体增强,如何使用户在感知上也获得正确的增强仍然是一个困难的问题,并对 AR 平台的用户体验构成了潜在困扰。从技术和方法看,还有很大的改进空间。由于涉及到人的主观感受,通过改进的硬件和软件来实现感知正确的增强是非常具有挑战性的研究方向。相关技术平台开发也是未来研究的基础方向。

### 参考文献

- [1] ROBIN C, HANIKA J, LENSCH H. Real-Time Disparity Map-Based Pictorial Depth Cue Enhancement[J]. Eurographics, 2012, 31(2):11-14. DOI:10.1111/j.1467-8659.2012.03006.x
- [2] BUCHTHOUGHT A, YOONESSI A, BAKER C L. Dynamic Perspective Cues Enhance Depth Perception from Motion Parallax[J]. Journal of Vision, 2017, 17(1):10,1-19. DOI: 10.1167/17.1.10
- [3] LIN C J, WOLDEGIORGIS B H, CAESARON D, et al. Distance Estimation with Mixed Real and Virtual Targets in Stereoscopic[J]. Displays, 2015, 36:41-48. DOI: 10.1016/j.displa.2014.11.006
- [4] KRUIJFF E, SWAN II J E, FEINER S. Perceptual Issues in Augmented Reality Revisited[C]// IEEE International Symposium on Mixed and Augmented Reality 2010. USA: IEEE, 2010: 13-16. DOI: 10.1109/ISMAR.2010.5643530
- [5] CHEN Y, CHANG T. Perceptual Oriented Depth Cue Enhancement For Stereoscopic View Synthesis, 3dTV-Conference[C]//the True Vision - Capture, Transmission and Display of 3d Video. USA:IEEE, 2016:1-4. DOI: 10.1109/3DTV.2016.7548884
- [6] 刘自强. 增强现实中深度一致性问题的研究[D]. 沈阳: 沈阳工业大学, 2017
- [7] 郑毅. 增强现实虚实遮挡方法评述与展望[J]. 系统仿真学报, 2014, 26(1):1-10

### 作者简介



陈东义,电子科技大学教授;主要研究方向为可穿戴计算、增强现实、人机交互等;作为负责人承担“863”、国家自然科学基金、国际合作等国家级科研项目10余项;已发表论文100余篇。

# 基于虚拟现实技术的无线自组网仿真

## Wireless Ad-Hoc Network Simulation Based on Virtual Reality Technologies

罗训/LUO Xun

王荣/WANG Rong

(天津理工大学, 天津 300384)

(Tianjin University of Technology, Tianjin 300384, China)

近几年,随着物联网的快速发展,无线技术也得到了飞速的发展,而在无线技术中,无线自组网技术<sup>[1]</sup>又是一个广受欢迎的技术,例如 Zigbee 技术和蓝牙技术等。当面对特定环境和特定需求需要开发一种新的无线组网技术或无线广播算法时,如何测试评估网络协议或算法的性能和可靠性是一个关键问题。

如果直接购买或者开发硬件平台进行实地部署测试,不仅成本高、费时费力,而且还需要考虑硬件故障、环境因素、实地部署时供电功耗、维护等很多问题,偏离了实验重心,不能聚焦在需要解决的关键问题上,而通过计算机仿真却可以避免以上问题,研究者只需要重点关注仿真的性能和结果来改进协议或算法即可。由此可见,网络仿真工作有着不可或缺的作用,占据重要的地位。

目前大部分的网络仿真都是通过 NS-2 或 OPNET 仿真器来实现,但是这些仿真器也存在一些不足,例如: NS-2 仿真器不便于构建复杂的仿真场景;调试功能受限;体系庞大,使用较为复杂。OPNET 仿真器则运行速度慢,仿真网络规模小,且对无

线网络的支持较弱<sup>[2]</sup>。而这其中的某些不足对于游戏引擎而言,却恰恰是优势所在。在文章中,我们介绍了一种基于虚拟现实的自组网仿真技术。

该方法大体流程是:首先利用 3D 建模软件建立实验所需的各种模型,之后导入游戏引擎,建立自组网仿真的虚拟现实场景,然后利用游戏引擎的接口,通过引擎脚本来控制仿真流程。通过实验证明:该仿真方法可以实现自组网络的仿真工作,同时也通过实验得到的数据论证了该仿真技术的有效性和可靠性。

### 1 游戏引擎

#### 1.1 游戏引擎简介

通过分析可知仿真是自组网必

**摘要:** 提出了一种网络仿真的新技术——基于虚拟现实(VR)的无线自组网仿真技术。游戏引擎中的场景管理、动画渲染、脚本控制、物理引擎等为自组网仿真提供了方便的接口,利用这些接口可以很轻松地实现自组网仿真工作。从游戏引擎的优势、仿真流程实现来说明利用游戏引擎实现在虚拟现实中的自组网络仿真技术,并与其他常用仿真技术的实验数据对比论证仿真方案的可行性和准确性。

**关键词:** VR; 无线自组网络; 仿真; 游戏引擎

**Abstract:** In this paper, the wireless ad-hoc network simulation based on virtual reality(VR) technology is proposed. The scene management, animation, rendering, scripting, physics engine in the game engine can provide convenient interfaces for ad-hoc network simulation. Using these interfaces, the simulation of ad-hoc networks can be easily implemented. The ad-hoc network simulation technology can be realized in VR by game engine, which is illustrated by advantages of the game engine, the simulation process in this paper. The feasibility and accuracy of the simulation scheme are demonstrated by comparing the experimental data with other commonly used simulation techniques.

**Keywords:** VR; wireless ad-hoc networks; simulation; game engine

不可少的一步,但是为了弥补传统仿真器的不足,提出了利用游戏引擎来进行仿真的新方法。

游戏引擎是指一些已经编写好的可编辑游戏系统或者一些交互实时图像应用程序的核心组件,游戏引擎包括以下系统:渲染引擎、动画制作、物理引擎、碰撞检测、系统脚本、引擎动画、粒子系统和场景管理等<sup>[3]</sup>。

对于自组网络仿真中需要的场景管理、组网动画效果和脚本控制等,游戏引擎正好提供了这些接口,因此使得利用游戏引擎来进行基于虚拟现实的自组网络仿真成为可能。

#### 1.2 游戏引擎优势

与传统的仿真器经过对比,游戏引擎具有以下优势:

- 简单易用。游戏引擎提供了整体的框架,根据游戏引擎所提供的接口,在进行自组网仿真时,只需要重点关注需要解决的问题,例如自组网协议研发或算法实现等,然后研发人员只需要通过脚本来控制节点之间的数据交换流程和动画展示等内容即可。

- 真实环境模拟。在之前的无线仿真技术中,都是通过生成随机位置,然后根据研究者配置进行模拟仿真。这种仿真的缺点是无法模拟出具体真实场景和地理位置对仿真实验的影响。在虚拟现实的游戏引擎中,研究者可以对仿真场景进行随意添加更改,包括丛林、河流、山峰和建筑物等,利用这些仿真真实的环境,增加了真实环境中各种障碍物对无线信号衰减的影响,从而提高了仿真的准确性。

- 可视化效果增强。在虚拟现实,用户可以很直观地看到仿真环境中无线信号的传播过程和组网过程,以及组网是如何受真实场景影响的,不仅对于研究者更加容易找出网络中的问题,对于初学者而言,这种直观真实的可视化效果也更加容易记忆和学习。

- 清晰的层次结构。游戏引擎中针对每个场景、每个状态和场景中的每个物体都有独立的脚本来控制,因此对于自组网仿真而言,这种清晰的层次结构可以很轻松地添加新功能或修改已有功能,例如网络仿真中各个节点的控制都是由其独立来控制实现,因此无论是修改功能还是更改流程在该节点的脚本中都可以很方便实现,整体控制可以通过关卡脚本来实现。

- 复用性高。在游戏引擎中,各个部件和物体的控制可以在脚本中使用函数和事件来实现,游戏引擎也提供了场景脚本调用所有物体内函数的功能,大大提高了程序的复用性;而且可以在场景中生成多个物体对象,由同一脚本去控制工作,例如

网络仿真中根据配置生成不同数量的节点进行无线组网,每个节点的工作流程完全一致,因此除了每个节点的ID不同外,其内部对收到的数据处理流程完全一样,因此共用一个程序文件来控制即可,完全不需要针对每一个节点单独编写。

- 消息广播机制。在组网过程中路由建立的基础就是通过路由广播以及路由回应来完成的,而在游戏引擎中也提供了全局消息广播的机制,通过该机制,研发者不需要考虑自己去实现复杂的广播消息处理。

我们选择了虚幻游戏引擎(简称UE4)来实现无线自组网仿真技术的研究。该游戏引擎除了其强大的各种子系统外,还提供了一个“所见即所得”的编辑器,在可视化的编辑窗口中,开发人员可以直接对游戏中的角色、物品道具、角色个体智能(AI)及光源进行自由摆放和属性控制,并且全部是实时渲染,虚幻引擎为了使开发者更好使用,还完全开源了其游戏引擎的代码,开发者可以按照自己的需求开发自己特色的内容。

### 1.3 游戏引擎相关工作

在此之前,张青和曲洲青<sup>[2]</sup>等人研究过基于游戏引擎的力学仿真方法,同样是由于力学实验耗费成本较高且不易运作,传统仿真方法又存在一些缺点,所以提出了利用游戏引擎来进行仿真的方法<sup>[4]</sup>;LUO和YU等人提出在网络仿真中利用游戏引擎来增强移动模型的研究<sup>[5-6]</sup>,GODSEY等人<sup>[7]</sup>则利用游戏引擎来实现仿真传感器网络应用,这些研究都为文中所提的基于虚拟现实的无线自组网仿真提供了启发思路。除此之外,YU-CHEE TSENG和SZE-YAO NI等人在研究关于移动自组网中的广播风暴问题时,利用C++设计了一种仿真器,并利用该仿真器仿真了在不同广播算法下广播风暴的缓解情况,对实验数据结果进行了分析和论证<sup>[8]</sup>。

在相关工作的基础上,我们利用

游戏引擎设计无线自组网仿真器,最后通过仿真YU-CHEE TSENG<sup>[8]</sup>论文中完全相同的实验场景进行实验,仿真得到的实验数据与YU-CHEE TSENG论文中的数据对比的结果,验证了文章中的仿真器是真实可靠的。

## 2 仿真流程实现方式

无线仿真遵循网络仿真的一般步骤,来确定无线自组网整体的仿真流程,包括问题定义、建模、模型确认、数据采集、程序编制、运行、结果展示和最后的验证等环节<sup>[9]</sup>。

我们采用了主从节点的方式进行无线组网仿真,主节点根据从节点的地址进行轮询广播查找,整个仿真场景的参数主要有场景大小、节点数量、信号传输半径、广播包大小、传输速率、广播时间间隔等。

自组网仿真路由由协议采用的是精简的按需矢量距离(AODV)<sup>[10-11]</sup>路由协议。在其基础上我们针对游戏引擎仿真进行了适当的优化,在游戏引擎中先对优化后的路由协议进行数据结构定义,主要包括以下几部分:路由广播包、路由回应包、路由错误包、路由消息包。除此之外,还定义了场景中节点物体与其对应地址的数据结构,用来在全局范围内查找某节点信号传输半径范围内的节点并将路由信息发送给该接收节点。当每个节点收到路由广播或路由回应后,节点需要将其存储至节点的路由表中,建立路由,因此还需要定义路由表结构。

在该仿真中,场景内主要包括以下几个部分:(1)范围盒子,在仿真时并不可见,主要用来根据用户配置的场景大小来自动变化范围,初始化后不可改变;(2)主节点,整个自组网仿真中路由广播的发起者,在轮询各个节点时,根据其内的路由表信息来判断是否进行路由广播;(3)从节点,根据配置信息来自动生成规定数量的从节点,同时自动分配IP。主节



点和从节点的位置是在范围盒子内部随机生成,当然用户也可以手动调整其位置。

根据虚幻 UE4 的设计,场景中的所有物体均可以获取到游戏状态和游戏模式,场景关卡可以获取到场景中的所有物体以及其内部的属性、方法函数和事件等。

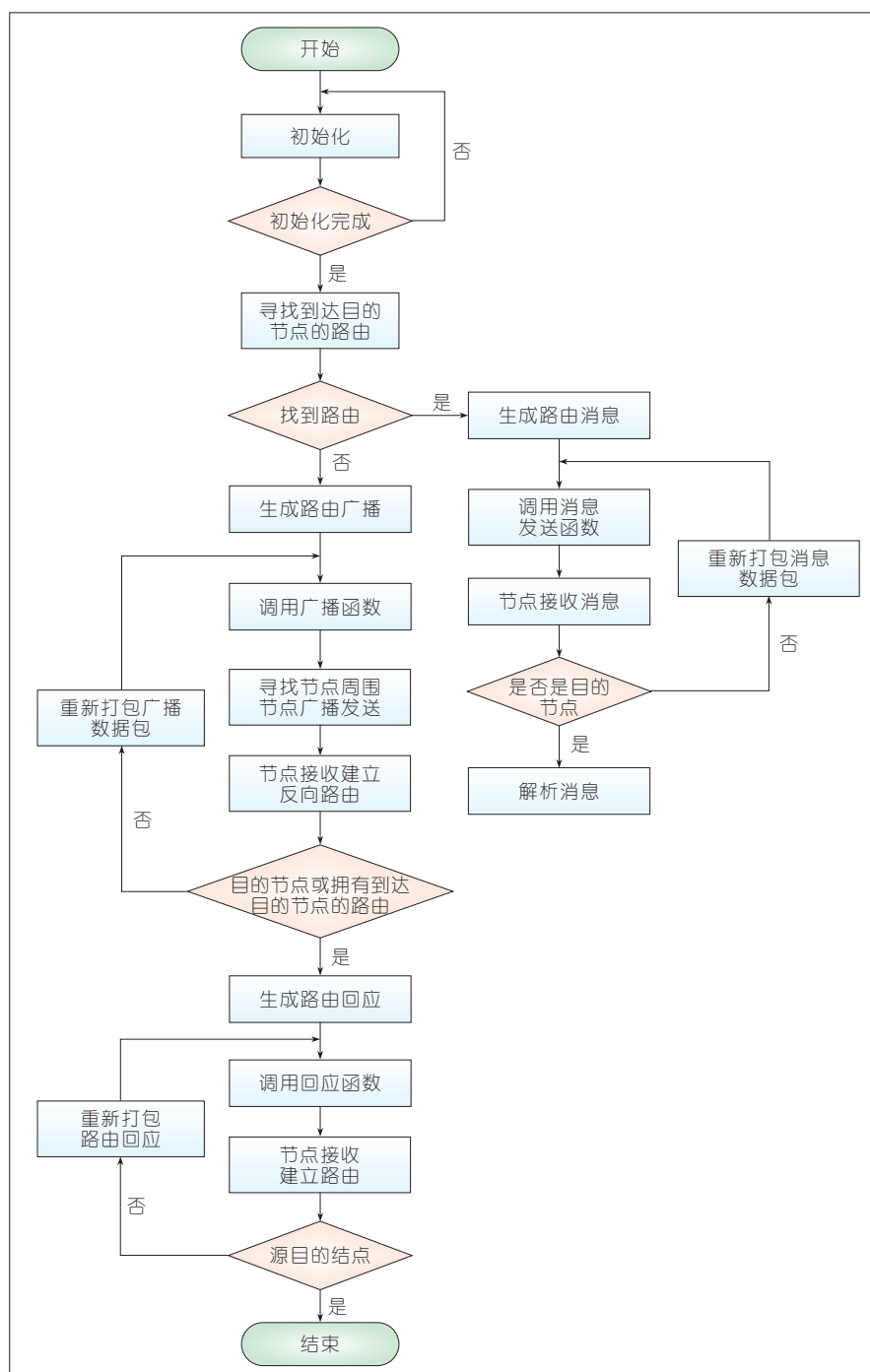
据此可知配置文件的读取由游戏状态来完成,并存入其中供其他物体来获取,所有的节点都具有发送消息的功能,因此在游戏状态中定义了路由广播发送函数接口、路由回应发送函数接口、路由错误发送函数接口和路由消息发送函数接口,当节点需要发送相应的消息时,直接通过调用游戏状态中的相应接口即可。相应消息的接收函数仍是由各自独立实现,各种发送函数接口的具体实现是在场景关卡中完成,由于场景关卡可以获取到场景内的所有物体,因此在场景关卡中通过对传入参数做解析之后,将需要发送的消息发送给各个对应的物体接收函数,从而完成了整个组网过程。图 1 是整个自组网系统的仿真流程。

最终仿真则效果如图 2 和图 3 所示,其中图 2 为主节点路由表中没有到达目的节点的路由信息,进行路由广播,从而引起其信号半径范围内的从节点依次向外中继转发广播,当广播到达目的节点或某中继节点中拥有到达目的节点的路由时开始回应。图 2 中有 3 个节点进行了路由回应,其中黑色箭头表示路由回应信号,右边红色的数字表示在某个节点信号半径范围内的可达节点。

图 3 是在路由建立之后,主节点发送路由消息给目的节点,同时目的节点回应路由消息给主节点的数据传输仿真过程。

### 3 仿真流程实验数据论证

在 YU-CHEE TSENG<sup>[8][12]</sup>等人的工作中,利用仿真器进行广播算法的仿真,得到的仿真数据通过平均时延、



▲图 1 自组网系统的仿真流程

节约转发率和可达率 3 方面来评估不同广播算法在不同阈值下的广播风暴缓解情况,然后通过横向对比,给出该广播算法的最优阈值范围。

评估观察的性能指标解释如下:

- 可达率(RE)。接收到广播消息的主机节点数与从源主机直接或

间接到达的所有主机节点数的比值。

- 节约转发率(SRB)。 $(r-t)/r$ ,  $r$  为接收到的广播消息的主机节点数,  $t$  为真正发送消息的主机节点数。

- 平均时延。从广播开始到最后主机节点结束转发的时间间隔。

我们利用本次设计的基于游戏





图2  
路由广播仿真场景



图3  
路由回应仿真场景

引擎的仿真器通过完全相同的仿真环境配置进行实验,将数据结果与其进行比较,仿真环境配置是:节点之间信号传输半径为500 m,然后生成100的主机节点进行仿真,范围大小有 $1 \times 1$ 、 $3 \times 3$ 、 $5 \times 5$ 、 $7 \times 7$ 、 $9 \times 9$ 、 $11 \times 11$ 单元大小,每个单元的长度是500 m(即信号传输半径);然后则通过仿真概率广播算法<sup>[4]</sup>和计数广播算法<sup>[4-5]</sup>获得数据作图对比,其中每次仿真执行10 000个广播请求。

### 3.1 概率广播算法仿真

概率广播算法是指当一个主机节点收到路由广播消息时,该主机会以概率 $P$ 进行转发。当概率 $P$ 是1时,该算法等同于泛洪广播<sup>[8][13]</sup>。

图4为通过本文中的仿真器仿真结果,统计在不同场景大小时,不同的转发概率 $P$ 条件下的节约转发率;图5是可达率;图6是平均时延。通过与YU-CHEE TSENG的论文中的数据对比,统计结果相差不超过1%。

### 3.2 计数广播算法仿真

计数广播算法是指当一个主机

节点收到路由广播时,在该主机转发广播之前会由于解析数据重新打包

广播包退避时间,忙碌或者处理其他队列消息等而等待一段时间。在该段时间内,该主机有可能再一次监测到相同的路由广播,因此通过一个计数器 $k$ 来统计该主机收到相同路由广播包的次数,当该次数 $k$ 超过计数阈值 $C$ 时,该主机取消转发广播。

图7为通过文中所提的仿真器仿真结果,统计在不同场景大小时,不同的计数阈值 $C$ 条件下的节约转发率;图8是可达率;图9是平均时延。通过与YU-CHEE TSENG的论文中的数据对比,统计结果相差不超过1%。

因此,通过以上数据统计分析和对比,可以证明本论文中的在游戏引擎中实现的基于虚拟现实的无线自组网仿真技术是有效可靠的。

## 4 结束语

文章中,我们通过分析自组网仿真的重要性总结了传统网络仿真中的一些不足之处,进而提出了利用游

图4  
概率广播算法仿真中不同场景大小下的节约转发率

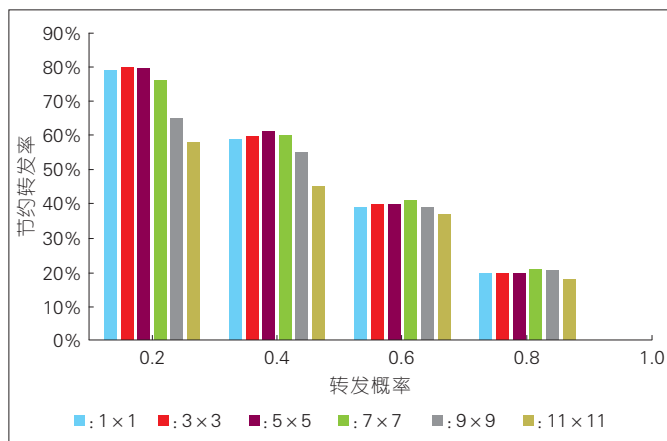
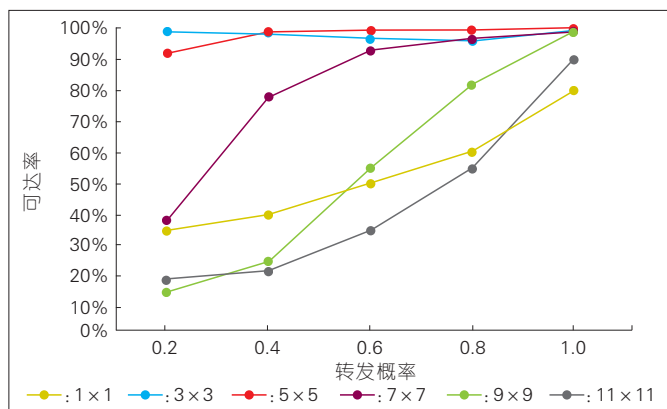
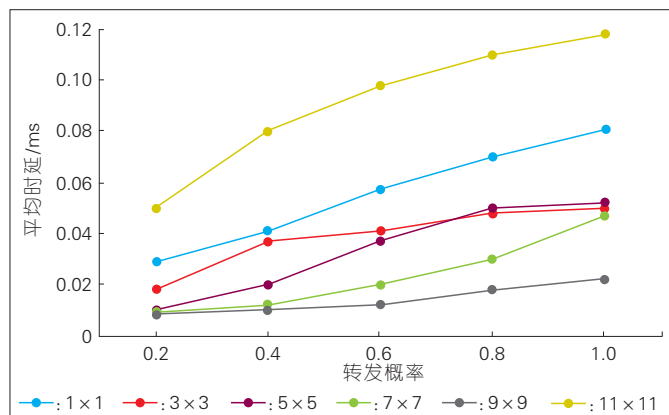
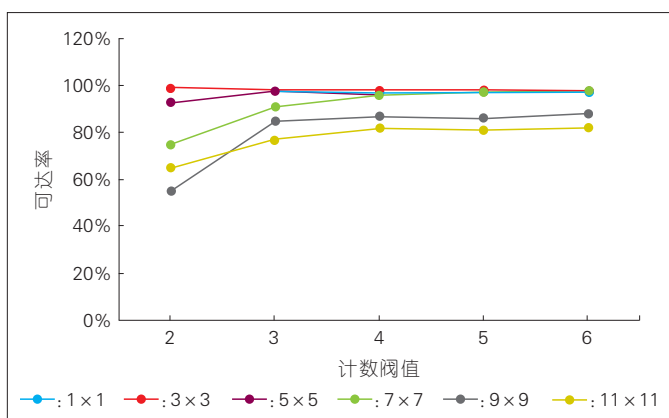


图5  
概率广播算法仿真中不同场景大小下的可达率





▲ 图6 概率广播算法仿真中不同场景大小下的平均时延



▲ 图8 计数广播算法仿真中不同场景大小下的可达率

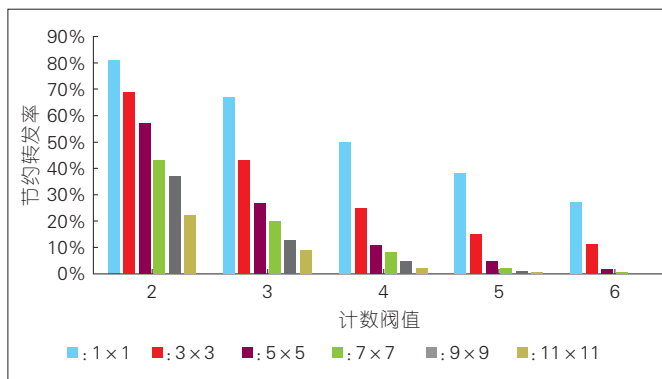
戏引擎来进行基于虚拟现实的自组网仿真的方法。之后通过UE4游戏引擎实现了该自组网仿真技术,实现了自组网仿真功能,同时利用该仿真器通过对两种广播算法进行仿真实验,与前人的实验数据进行比较和分析,验证了该基于虚拟现实的无线自组网仿真技术的可行性,研究也达到了预期的效果。

我们相信在之后的发展过程中,利用游戏引擎的虚拟现实技术会应用到更多的领域中去,并且更加简单和方便地为研究者提供便利。

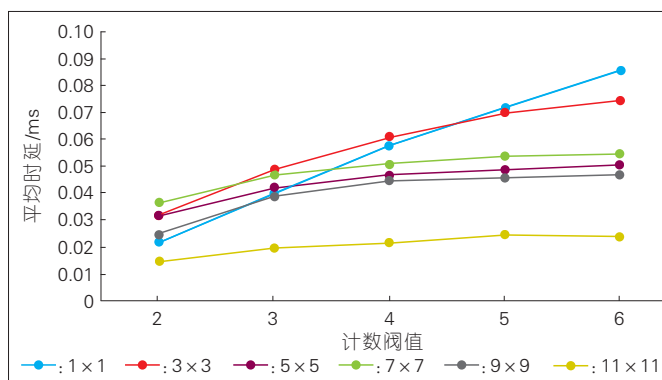
#### 参考文献

- [1] MACKER J P, CORSON M S. Mobile ad hoc Networking and the IETF[J]. Acm Sigmobile Mobile Computing & Communications Review, 1998, 2(2):9-12
- [2] 高振国, 温东新, 刘宏伟, 等. 构建无线自组网仿真场景[J]. 宇航学报, 2006, 27(4):806-813
- [3] 张青, 曲洲青. 基于游戏引擎的力学仿真新方法[J]. 中国传媒大学学报(自然科学版)自然科学版, 2013, 20(2):45-49

- [4] MOHAMMED A, OULD K M, MACKENZIE L M, et al. Improving the Performance of Counter-Based Broadcast Scheme for Mobile AD HOC Networks[C]// IEEE International Conference on Signal Processing and Communications. USA:IEEE, 2008:1403-1406.DOI: 10.1109/ICSPC.2007.4728591
- [5] LUO X. Using Game Engine to Enhance Mobility Modeling in Network Simulations [C]// Virtual Reality.USA: IEEE, 2013:177-178
- [6] LUO X, YU N. Fast Mobility Model Prototyping in Network Simulations using Game Engine[C]// 2013 中国虚拟现实大会暨 ICVRV'2013. USA:IEEE, 2013:145-152. DOI: 10.1109/ICVRV.2013.31
- [7] GODSEY C, SKUBIC M. Using Elements of Game Engine Architecture to Simulate Sensor Networks for Eldercare [J]. 2009: 6143-6146
- [8] TSENG Y C, NI S Y, CHEN Y S, et al. The Broadcast Storm Problem in a Mobile ad hoc Network[J]. Wireless Networks, 2002, 8(2-3): 153-167
- [9] 侯宗浩, 王秉康, 黄泳翔. 网络仿真的研究[J]. 计算机仿真, 2003, 20(10):89-91
- [10] PERKINS, BEDOMGRPUER. Ad hoc On-Demand Distance Vector (AODV) Routing [J]. RFC, 1997, 6(7):90
- [11] ROYER E M, DAS S R, PERKINS C E. A d Hoc On-Demand Distance Vector (AODV)



▲ 图7 计数广播算法仿真中不同场景大小下的节约转发率



▲ 图9 计数广播算法仿真中不同场景大小下的平均时延

Routing (Internet-Draft)[J]. RFC, 2000, 6(7): 90

- [12] TSENG Y C, NI S Y, SHIH E Y. Adaptive Approaches to Relieving Broadcast Storms in a Wireless Multihop Mobile ad hoc Network[J]. Computers IEEE Transactions on, 2001, 52(5):545-557
- [13] BERTSEKAS D, GALLAGER R. Data Networks (2nd ed.)[M]. British: Prentice-Hall, Inc., 1992

#### 作者简介



罗训, 天津理工大学计算机科学与工程学院教授、博士生导师, 天津市“千人计划”专家; 主要研究方向为虚拟现实与普适计算; 先后参与支持国家/省市级科研项目9项; 发表论文30余篇, 获发明专利27项。



王荣, 天津理工大学计算机科学与工程学院硕士研究生; 主要研究方向为无线自组网通信、虚拟现实仿真等。

# 虚拟现实视频传输架构和关键技术

## Architecture and Key Technologies of Virtual Reality Video Transmission and Streaming

郭宗明/GUO Zongming  
班怡璇/BAN Yixuan  
谢澜/XIE Lan

(北京大学, 北京 100871)  
(Peking University, Beijing 100871, China)

虚拟现实(VR)视频, 又称全景视频或360°视频, 以其充分的沉浸性、真实性、制作简单性等特点正受到人们的广泛关注。它通过全景摄像机记录三自由度或六自由度的自然图像, 降低了VR内容制作的门槛, 提供现有普通视频所不能达到的真实感。近几年, 随着VR设备、网络和终端技术的成熟, VR视频正越来越多地用于影视、游戏、医疗、教育等应用中。据思科的统计报告显示<sup>[1]</sup>: VR应用在2016年的平均月流量达到13.3 PB, 预计到2021年达到140 PB, 年增长率达到60%。如此广泛的VR视频给人们的生活和工作带来了新体验, 大量的网络基础设施和智能穿戴设备也推动了VR视频应用的发展。

然而, 高质量的VR视频服务仍面临若干挑战。其中, 网络传输是一个关键性问题。VR视频对网络带宽有较高的消耗, 例如: 一个分辨率为2 K×1 K的虚拟现实视频经过压缩编码, 需要4 Mbit/s的数据量; 若分辨率达到4 K×2 K, 则数据量将会上涨至

20 Mbit/s, 至少需要25 Mbit/s的网络带宽才可以承载<sup>[2]</sup>。现有互联网带宽难以满足如此高码率虚拟现实视频的流畅播放, 进而严重影响其应用。

为了解决VR视频的传输瓶颈, 工业界和学术界提出了众多解决方案。早期, Google、YouTube采用可兼容现有视频编解码器的全景传输模式, 但存在码率高、视频变形等问题。2015年, Facebook提出了视点自适应传输和金字塔映射模式, 可根据用户视点区域传输视频内容, 减少总体码率。2016年, 三星、高通等公司提出了多种非均匀映射方式, 动态图像专家组(MPEG)也专门成立一个I-AHG沉浸式视频专题组, 制订全景视频文件格式、映射、传输标准。在中

国, 数字音视频编解码技术标准工作组(AVS)标准化组织也于2016年开始制订AVS虚拟视频编码和系统标准, 中兴通讯、北京大学、上海交通大学等单位相继提出多种映射和传输方案。

文中, 我们将详细介绍VR视频传输的发展现状和关键技术, 重点从传输框架、全景传输、自适应传输等方面展开。

### 1 虚拟现实视频传输框架

完整的虚拟现实视频传输架构包括5个部分: 全景采集、拼接、映射、编码、传输, 如图1所示。

(1) 全景采集。VR视频是由多摄像头组合成的全景相机拍摄而

**摘要:** 介绍了虚拟现实(VR)视频传输架构, 提出了两种传输方式: 全景传输和视点传输。全景传输可提供多角度质量均匀的图像, 但存在高码率和带宽浪费问题; 视点传输仅传输视点范围内图像, 可减少码率浪费, 但存在头部运动延迟问题。提出的视点自适应传输技术、视点预测方法, 可以较好地实现低码率、高质量的VR视频传输。

**关键词:** VR视频; 视点自适应传输; 非均匀映射传输; 分块传输; 视点预测; 用户质量优化

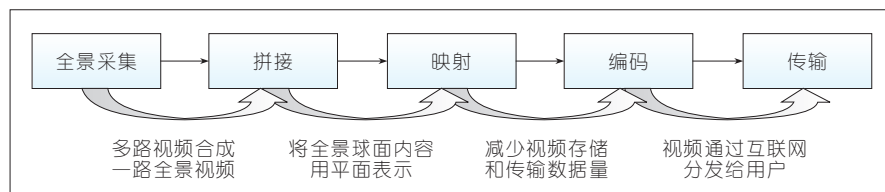
**Abstract:** In this paper, a main framework of virtual reality (VR) video transmission is introduced, and two schemes are also proposed: panorama transmission and viewport transmission. The panorama transmission can provide 360° symmetric image with high bit-rate and bandwidth wasting; the viewport transmission only transmits a part of images so as to reduce bit-rate wasting, but can cause head movement delay. In this paper, viewport-adaptive streaming and viewport prediction technologies are introduced, which enable high quality service of VR video streaming.

**Keywords:** VR video; viewport adaptation streaming; asymmetric projection streaming; tile-based streaming; viewport prediction; QoE optimization

收稿日期: 2017-09-26

网络出版日期: 2017-11-09

基金项目: 国家自然科学基金(61471009)



▲图1 虚拟现实视频传输架构

成。利用全景摄像机采集自然图像,可以极大简化虚拟现实内容的制作。近几年,全景相机硬件技术得到飞速发展。从早期由多个普通摄像头合成的全景相机,发展到各种专用相机,例如:Facebook Surround 360、三星 Gear 360 等。可以提供从视频采集到拼接等一系列功能,方便虚拟现实视频的创作。

(2)视频拼接。视频拼接是全景视频采集的后处理过程,它将若干个摄像头的视频合成一路全景视频,形成一个完整的虚拟现实视频提供给用户。目前已经有多种商业化或开源的视频拼接软件,例如:Video Stitch、Facebook Surround 360 SDK、Nuke 等。

(3)映射过程。为了便于存储和压缩编码,球面视频内容需要被几何映射到平面。映射影响了VR视频编码前的像素量,在一定程度上也决定了视频内容所包含的信息量。映射过程与传输紧密相关,分为两类:均匀映射和非均匀映射。均匀映射保留全部全景视频信息;非均匀映射仅在部分区域保留完整信息,其他区域保留部分信息,并用较少像素表示。

(4)视频编码。VR视频使用压缩编码减少视频中的冗余信息。为了兼容现有编码器和播放设备,目前虚拟现实视频主要采用H.264,高效视频编码(HEVC)等编码标准。

(5)网络传输。VR视频需要通过互联网分发给用户,网络传输是目前最大的挑战。为了提供清晰、流畅的虚拟现实视频体验,映射和传输方法起到了关键性作用。传输方法主要分为两类:

- 全景传输方法。将360°全景

视频以同等质量、完整的发送给用户。可以保证映射内容完整保留了原始球面的所有内容,保留信息量最大;但由于播放设备和视角的限制,在某一时刻,用户只关心视角范围内的部分内容,全景传输方法势必会造成巨大的资源浪费。

- 视点自适应传输方法。顾名思义,指的是客户端可根据用户视点,动态传输视角范围内的视频,避免了资源浪费。然而,若用户当前视点与下载视频的视点不符,会出现黑屏或视频质量降低等问题。

在虚拟现实视频传输框架中,映射与传输方式紧密相关。其中,全景传输对应均匀映射,完整的映射和传输360°视频;视点自适应传输对应非均匀映射和分块传输,按用户视角映射和传输局部视频。

下面将详细介绍这两种虚拟视频传输方式,以及各自优缺点。

## 2 全景传输

全景传输方法是将球面视频均

匀的映射到平面上,按照矩形布局编码和传输完整的360°视频内容。全景传输可直接使用现有编码器和传输方式,因为其简单和通用而成为当下虚拟现实视频使用最广泛的传输方法。

均匀映射是全景传输的主要映射方式,决定了编码效率和带宽消耗。代表性的均匀映射方法有矩形映射(ERP)、立方体映射(CMP)、八面体映射等,如图2所示。

ERP方法是使用最多的虚拟现实视频映射方法,这种方法在球的纬度上按照赤道的采样频率对内容进行采样并用平面进行表示,生成的虚拟现实视频的宽高比为2:1。但ERP方法在南北两极存在着严重的过采样现象,使得整体的冗余像素增加。

立方体映射是将球面内容投影到外接的一个立方体的6个面上,并将6个面拼接重组成一个完整的虚拟现实视频。

除此之外,更多的多面体被应用于VR视频的映射方法中,例如:正八面体(OHP)、正二十面体(ISP)等。理论上,使用面数越多的正多面体,可以使得映射后的数据量越小。但是,这些正多面体都需要通过“再拼接”形成矩形平面。拼接过程导致内容上的不连续现象,在一定程度上也会影响编码效率。

图2 矩形、六面体等均匀映射方法示例

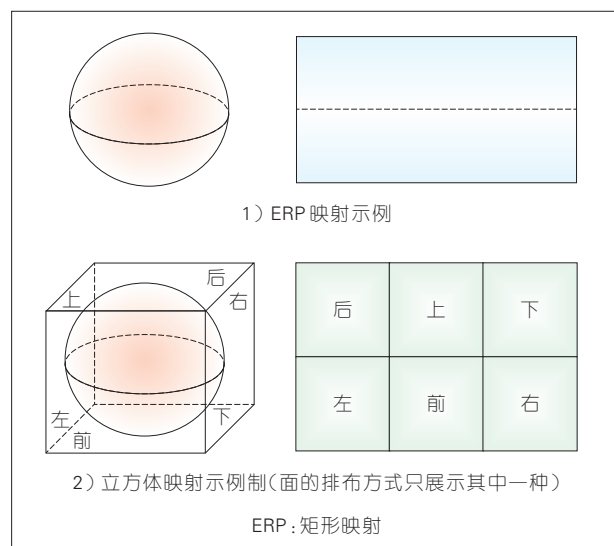




表1总结了目前主要的均匀映射方式的优缺点。其中,像素比例代表映射后面积和原始球面的比例关系,

▼表1 均匀映射方法比较

映射方法	像素比例	优点	缺点
矩形映射	1.57	映射方法简单	南北两极存在严重过采样
立方体映射	1.91	每个面失真小	6个面拼接存在不连续现象,影响编码效率
八面体映射	1.65	映射面积小	多面拼接存在不连续现象,影响编码效率
二十面体映射	1.21	映射面积小	多面拼接存在不连续现象,影响编码效率
柱状映射	1.20	映射面积小	存在空白区域

是影响视频编码码率的一个因素。

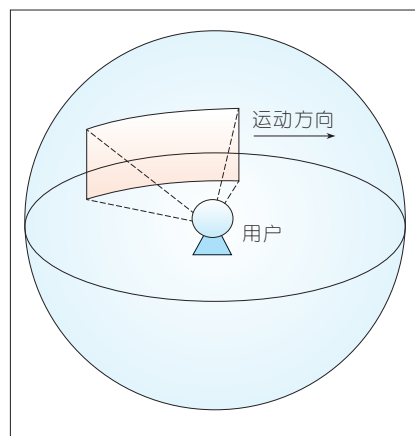
### 3 视点自适应传输

视点自适应传输技术是一种根据用户视点,按需下载视频内容的方法。如图3所示,对于用户视角内的内容可以选择下载高质量的版本,对于视角外的部分可下载低质量的版本,或者不进行下载。

视点自适应传输可有效减少虚拟现实视频的带宽浪费问题。在相同带宽条件下,它可分配更多带宽给视角区域,从而提高用户视窗内视频的质量,增强用户的观看体验。

按照传输内容的构成形式,视点自适应方法可以分为两类:非均匀映射和分块传输。

(1)非均匀映射<sup>[3]</sup>。传输一个质量不均匀的360°全景视频,用户视点范围内是高分辨率,其他区域是低分辨率,从而减少整体码率。



▲图3 视点自适应技术原理

(2)分块传输<sup>[4]</sup>。将全景视频从空间上分为不同视频块,根据视点范围传输部分视频块。

两种的主要差别为:非均匀映射需传输360°的全景视频,可防止头部快速运动导致的黑场;分块方法可传输局部或全部视频内容,灵活性高。

视点预测是视点自适应传输的另一个重要技术。VR视频要求显示延迟低于20 ms<sup>[5]</sup>。同时,由于存在网络传输延迟,在头部快速运动时,要避免视野范围内出现黑场,必须对视点区域进行预测和预取。因此视点预测的准确度极大影响虚拟现实视频的传输质量。

下面将详细介绍视点自适应的非均匀映射、分块传输和视点预测等关键技术。

#### 3.1 非均匀映射传输技术

非均匀映射传输技术是一种通过将原始球面视频进行非均匀映射处理后进行传输的技术。其在对球面内容进行采样时,令球面上的像素点有不同的权重,使得关键视频内容得到保留,而不重要的区域被下采,仅保留少部分关键信息,如图4所

示。其优点是在保证关键区域质量的同时,从整体上降低视频的大小,极大地减轻了网络传输和客户端解码的压力。

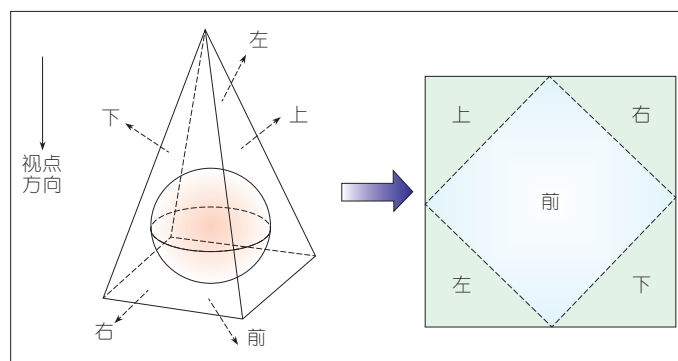
目前已有多种非均匀映射传输方式,例如:Facebook提出的多分辨率立方体、金字塔映射、偏移立方体等,高通提出的阶段金字塔等。它们共同目标是在保持视点区域质量的前提下,降低整体分辨率,但不同映射方式存在变形、失真或边界等问题,对编码效率有不同影响。

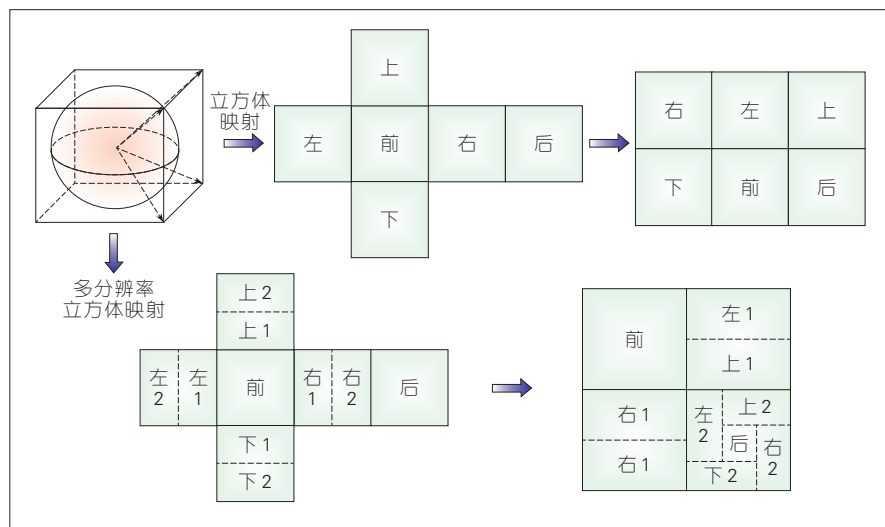
Facebook最早提出了一种多分辨率立方体映射方法,它是通过将立方体映射方式进行变换后得到的。保持正前面和与其相邻的上、下、左、右4个面的一半分辨率不变,对这4个面剩下的一半和后面的视频内容进行下采样,最终重新组合成一个二维平面视频帧,如图5所示。这种方法的缺陷是在上、下、左、右面的中间部位由于采样率突变,存在明显的边界,影响用户的观看体验。

除Facebook外,高通曾提出一种截断的金字塔映射方法<sup>[6]</sup>,并被MPEG OMAF标准采纳。这种方法与Facebook的金字塔映射方法类似,只不过不关注区域投影至四棱台的顶部,而不是投影为一个像素点,如图6所示。这种方法相较于金字塔方法,在一定程度上消除了由于映射造成的斜边效应,在内容表示上更加连续,有利于视频编码。

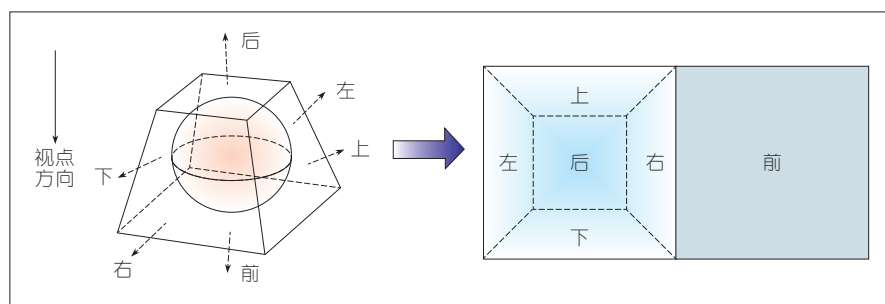
为了支持非均匀映射传输,需要针对每一个主视点都预先生成不同的视频版本。一旦用户发起请求,就

►图4 非均匀映射的原理





▲图5 多分辨率立方体映射



▲图6 截断金字塔映射

立即将对应视点的版本提取出来进行传输,在保证视频质量的同时减少了网络流量。

### 3.2 分块传输技术

分块传输技术是另一种视点自适应传输方式。它将全景视频按照空间划分为若干个子视频块,客户端可以根据网络状况和用户头部运动有针对性的向服务器端请求视频片段,具体过程如图7所示。

与非均匀映射传输方式不同,分

块传输仅传一部分内容,进一步减少了传输数据量。它可以自由地选择各个分块的质量,增加了传输的灵活性。同时,最新的HEVC编码标准支持tile方式编码,可用一个编解码器对整个视频流进行编解码,大大减少了客户端的解码复杂度。所以,分块传输是目前最实用的全景视频传输技术之一。

分块传输可有多种切分方式,例如:六面几何式<sup>[7]</sup>、18格矩形分割方式<sup>[8]</sup>、12格矩形分割方式<sup>[8]</sup>等。不同

的切分方式将会影响编码效率、传输块数量和边界失真等。

图8以18格和12格矩形分割方式为例。18格划分是将水平方向按照60°间隔分成6列,垂直方向按照45°-90°-45°方式分3行,共计18个视频块。12格划分是水平方向按照90°间隔分成四列,垂直方向同上。

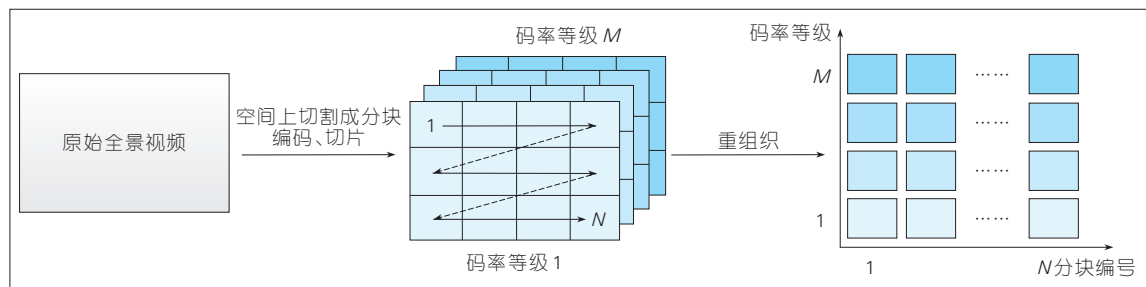
值得注意的是:在全景视频传输过程中,切分粒度越小,视频块的组合就越贴近用户视角,浪费的视频内容也就越少,也即自适应传输技术就越灵活。但是,一味减小视频块大小并不会使视频数据量有持续显著性下降。这是由于在切分后像素之间的相关性被破坏,使得切分后编码数据量之和略大于切分前的数据量,即便是在高压缩率的HEVC标准下,18格划分方式都会招致5.82%的压缩损失<sup>[9]</sup>,在极端条件下,浪费的视频内容大小甚至不足以抵消分块带来的损失。所以,如何决定分块大小也是一个值得关注的问题。

### 3.3 视点预测方法

视点预测是视点自适应传输的关键技术之一。由于网络传输存在延迟,为保证视点自适应播放的连续性,减少头部运动带来的卡顿甚至黑屏,如何准确预测视角位置是一个较大的挑战。

例如:线性预测方法中<sup>[9]</sup>,预测用户未来1s的头部位置时,准确率可达到90%以上,而将预测时间延长至3s时,准确率会大幅降至70%。一旦预测错误,由于本地缓存中缺少相应视角的视频片段,会导致用户视窗中出现黑窗,极大地影响用户的观看

图7  
分块传输示意



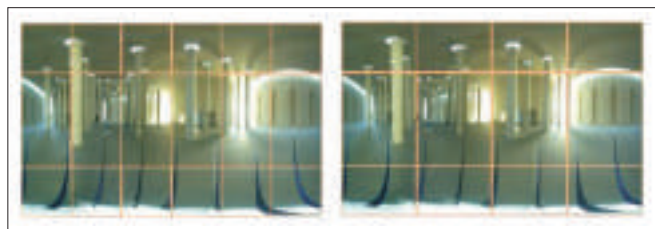


图8  
18格/12格划分示意

体验。

目前视点预测方式主要分为两类:运动预测和内容分析方法。运动预测是根据用户的历史浏览行为,预测未来视角位置,包括均值算法<sup>[9]</sup>、线性回归<sup>[9]</sup>、概率统计<sup>[10]</sup>、运动估计<sup>[11]</sup>、行为学习<sup>[12]</sup>等方法。

其中,线性回归算法尤为重要,它是一种利用数理统计中的回归分析来预测下一时刻用户位置的算法。具体来说,它先将用户在滑动窗口内不同时刻的运动状况进行记录,再通过记录下的头部数据来训练回归模型,得到回归曲线,预测下一时刻观看概率最大位置,如图9所示。

另外一类是内容分析方法。它是一种是基于视频中各个对象的显著性特征来进行预测。显著性特征代表着物体吸引用户的程度,显著性越强,表示用户越关注,同时也代表用户观看概率越高,显著性越弱,则代表观看概率越低。目前,学术界主要是通过机器学习的方法得到视频的显著性特征<sup>[12]</sup>,再进行预测。另外一种内容分析方法<sup>[13]</sup>是同时基于内容以及用户位置进行预测的神经网络。其中,视频内容包括显著性特征

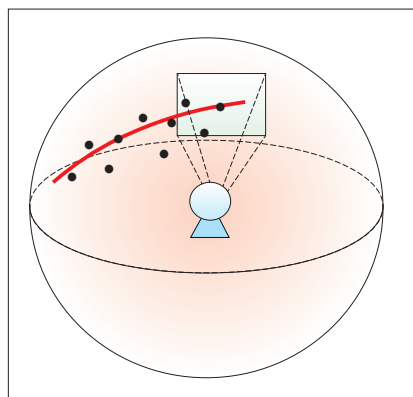


图9 线性回归预测视点

和运动特征。它需要在视频上架前对视频内容和测试头部数据集进行训练,学习出用户最可能出现的行为模式,进而在用户实际观看时根据学习结果预测下一时刻用户头部位置。

## 4 结束语

VR视频正得到越来越广泛的应用,但由于带宽消耗巨大,传输瓶颈已成为它面临的重大挑战之一。文章中,我们系统地介绍了VR视频传输框架,以及关键技术的发展现状,详细比较各种技术的优缺点,总结VR视频传输的发展方向。随着VR视频传输技术的发展,在有限带宽上传输高质量VR视频将成为现实。

### 参考文献

- [1] Cisco Visual Networking Index: Globe Mobile Data Traffic Forecast Update[R]. Technical report, 2017
- [2] Internet Connection Speed Recommendations[EB/OL]. [2017-10-12]. <https://help.netflix.com/en/node/306>
- [3] XU Z M, XIE L, ZHANG X G, et al. Optimal Viewport Adaptive Streaming for 360-Degree Videos[C]// IEEE International Conference on Multimedia&Expo(ICME '2017). USA:IEEE, 2017. DOI:10.1109/ICC.2017.7996611
- [4] BAN Y X, XIE L, ZHANG X G, et al. An Optimal Spatial-Temporal Smoothness Approach for Tile-Based 360-Degree Video Streaming[C]//IEEE Conference on Visual Communications and Image Processing (VCIP). USA:IEEE, 2017
- [5] YAO R, HEATH T, DAVIES A, et al. Oculus VR Best Practices Guide[EB/OL]. (2014-03-17)[2017-10-11]. <http://brianschrank.com/vrgames/resources/OculusBestPractices.pdf>
- [6] VR/360 Video Truncated Square Pyramid Geometry for OMAF: ISO/IEC JTC1/SC29/WG11/M[S]. US: MPEG, 2016
- [7] HOSSEINI, MOHAMMAD, and SWAMINATHAN V. Adaptive 360 VR Video Streaming: Divide and Conquer[C]//2016 IEEE International Symposium on Multimedia (ISM). USA:IEEE, 2016
- [8] ZARE A, AMINLOU A, HANNUKSELA M, and GABBOUJ M. HEVC-Compliant Tile-Based Streaming of Panoramic Video for Virtual

Reality Applications[C]// the ACM International Conference on Multimedia(ACM MM). USA:ACM,2016

- [9] QIAN F, JI L, HAN B, et al. Optimizing 360 Video Delivery over Cellular Networks[C]// 5th ACM Workshop on All Things Cellular, 2016. USA:ACM, 2016:1-6. DOI: 10.1145/2980055.2980056
- [10] XIE L, XU Z, ZHANG X, et al. 360ProbDASH: Improving QoE of 360 Video Streaming using Tile-based HTTP Adaptive Streaming[C]//ACM International Conference on Multimedia(ACM MM). USA: ACM, 2017
- [11] BATTLE L, CHANG R, and STONEBRAKER M. Dynamic Prefetching of Data Tiles for Interactive Visualization[C]//the ACM International Conference on Management of Data, 2016
- [12] LO W-C, FAN C L, LEE J, HUANG C Y, et al. 360° Video Viewing Dataset in Head-Mounted Virtual Reality[C]// 8th ACM on Multimedia Systems Conference. USA: ACM, 2017
- [13] FAN C L, LEE J, LO W-C, et al. Fixation Prediction for 360° Video Streaming in Head-Mounted Virtual Reality[C]//the 27th Workshop on Network and Operating Systems Support for Digital Audio and Video. USA:ACM, 2017:67-72. DOI: 10.1145/3083165.3083180

### 作者简介



**郭宗明**, 北京大学博导, 北京大学计算机科学技术研究所所长; 主要研究方向为数字视频处理及传输、彩色图像处理; 曾主持、领导研发团队攻克多项数字视频处理技术与工程难关, 研制成功多项工程应用系统并获得广泛应用, 促进了中国电视行业从模拟播出到数字播出的技术变革, 促进了中国网络视频行业的发展和普及; 获中国高等学校十大科技进展、国家科技进步二等奖、教育部科技进步一等奖和二等奖、教育部技术发明奖二等奖、广电总局科技创新二等奖等; 发表论文150余篇, 申请专利120余项。



**班怡璇**, 北京大学计算机科学技术研究所读本科生; 研究方向为网络视频传输、HTTP动态自适应流媒体; 发表论文1篇。



**谢澜**, 北京大学计算机科学技术研究所读硕士生; 研究方向为HTTP动态流传输、全景视频视点自适应传输; 已发表学术论文4篇, 拥有专利2项、标准提案3项。



# 大视频 VR 直播业务及其技术

## Service and Technology of VR Live on Big Video System

徐火顺/XU Huoshun  
叶剑章/YE Jianzhang  
尹海生/YIN Haisheng

(中兴通讯股份有限公司, 广东 深圳  
518057)  
(ZTE Corporation, Shenzhen 518057, China)

### 1 VR 直播业务简介

虚拟现实(VR)是一种“互动式的计算机模拟环境,能感知用户的状态和行为,替换或者加强对一种或多种感知系统的感官反馈信息,从而使用户获得一种沉浸在模拟环境虚拟环境中的感觉”。虚拟现实技术的特点是沉浸性高,当用户处于虚拟环境时,如身临其境。用户转变角度时,虚拟环境也会作出相应的改变<sup>[1]</sup>。

普通视频直播中受众往往只能从某一角度观看直播,而不能全方位地了解主播周围环境的状况;而虚拟现实技术满足了这个需求,使受众能从各个角度观看直播,增强了用户体验,参与感也大大增强。因此,虚拟现实与视频直播两者相互契合,可以给受众营造更好的观看效果<sup>[2-3]</sup>。

现阶段的VR直播主要是指360°全景直播,利用多路摄像机将视频信号拼接为360°全景视频,解决传统直播受众受镜头推移、视线角度的限制不能获得最佳视觉体验的问题。用户可以通过机顶盒(STB)、VR眼镜等

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868 (2017) 06-0024-004

**摘要:** 提出了一种在网路协定电视(IPTV)/互联网应用服务(OTT)大视频系统上的虚拟现实(VR)直播业务的实现技术,可以让IPTV/OTT的用户通过机顶盒在电视上观看360°场景的VR直播视频,或通过手机加头盔、一体化集成的VR眼镜等移动设备来观看VR直播视频。该VR直播业务的实现技术也实现了端到端低时延,提升了用户体验;在IPTV/OTT大视频系统上承载VR直播业务,需要对系统进行端到端的技术改造。

**关键词:** IPTV/OTT大视频系统;直播业务;机顶盒;手机;低时延

**Abstract:** The implementation technology of virtual reality (VR) live service on the Internet Protocol Television (IPTV)/Over The Top (OTT) big video system is proposed in this paper, which allows IPTV/OTT users to watch 360-degree scene VR live video on TV with set top box(STB) or mobile devices such as smart phone plus helmet, integrated VR client. This technology can realize the end-to-end low-latency and improve the user experience. End-to-end technological transformation for the IPTV/OTT system is also needed when bearing VR live service.

**Keywords:** IPTV/OTT big video system; live service; STB; smart phone; low-latency

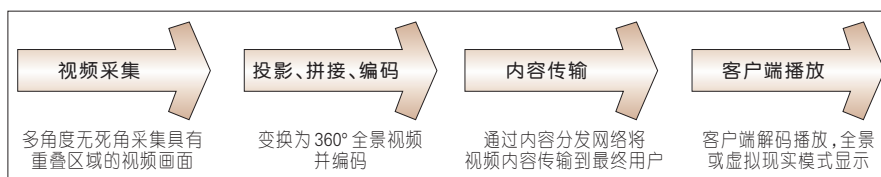
设备进行观看直播,体验沉浸感和现场感<sup>[4-5]</sup>。

VR直播,通过全景摄像机进行视频的实时采集,并对视频进行拼接、编码,通过内容分发网络进行传输,最终在终端上进行视频的解码播放。其中用到了多项视频处理的技术,和普通直播视频的要求有着很大不同。VR直播业务主要流程如图1所示。

VR直播可以通过多种方式呈现给最终用户,在家中可以通过家庭网关,利用机顶盒显示在电视上VR视

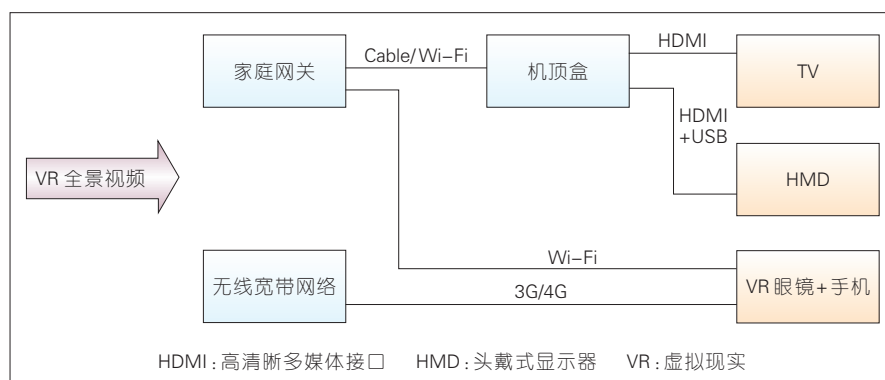
频内容,电视上显示的全景视频需要用遥控器控制转动;也可以使用机顶盒通过高清晰度多媒体接口(HDMI)和USB线连接到头戴式显示器(HMD),头部转动可以显示所面向方向的视频内容,显示的内容随着用户头部的转动而转换显示内容;另外也可以使用手机+VR眼镜的方式,手机可以通过Wi-Fi或移动网络获取VR全景视频,显示效果与HMD的方式类似。VR直播用户使用场景如图2所示。

用户不仅可以观看直播内容,也



▲图1 VR直播业务主要流程

收稿日期: 2017-09-28  
网络出版日期: 2017-11-02



▲图2 VR直播用户使用场景

可以通过同样的方法观看VR视频点播,VR直播也可以通过回放的方式播放。

## 2 VR直播业务中的相关技术

### 2.1 视频图像同步技术

与普通直播视频拍摄不同,VR直播视频的采集需要多镜头或摄像机同时完成。VR直播视频拍摄设备的取景范围为水平 $360^\circ$ 、垂直 $180^\circ$ ,如图3所示。

全景摄像机需要配置好参数,在拍摄过程中,还需要解决多相机的采集同步的问题。常见的同步方式有:闪光同步,即检测所有相机视频帧内的“闪光”,如明亮帧、白色帧,利用这个信号进行同步;运动同步,即检测所有相机视频帧内的运动信息,通过匹配各帧运动量进行同步;声音同步,即分析所有相机采集到的声音频谱进行同步;手动同步,即根据某一个时刻的所有相机采集的视频帧手动进行微调。

### 2.2 投影变换技术

多相机拍摄的图像是在不同方向下拍摄得到的,所以并不在同一投影平面上,在对重叠的图像进行无缝拼接前,如果不进行投影平面的变换,直接拼接会破坏实际景物的视觉一致性。投影变换的方式一般有平面投影、柱面投影、球面投影和鱼眼

投影等<sup>[6-7]</sup>。

### 2.3 视频缝合技术

完成投影变换后进行拼接,拼接过程主要有特征提取—特征匹配—配准—融合等步骤<sup>[8]</sup>。常用的特征提取方法有尺度不变特征变换(SIFT)、加速稳健特征(SURF)、定向快速特征点提取算法(ORB)、二值化描述算法(BRIEF)等<sup>[9-10]</sup>。特征点匹配可以采用SIFT的作者Lowe提出的比较最近邻距离与次近邻距离的SIFT匹配方式,或者KD-Tree算法等。为了将图像注册到同一个坐标系中,需要在多幅图像配准的过程中根据几何运动模型来配准,采用的几何运动模型主要有:平移模型、相似性模型、仿射模型和透视模型等。

完成图像配准后,便可以进行图像融合操作,即形成了 $360^\circ$ 全景图像。图像匹配拼接如图4所示。

### 2.4 图像增强技术

由于不同角度的画面是通过不

同的相机采集得到,融合后的 $360^\circ$ 全景图像会遇到各个区域的曝光不一致的情况,通过曝光补偿的技术可以使得拼接后的全景图像曝光一致。

此外,当场景中的物体存在运动的情形时,融合后的全景图像中会出现“鬼影”的情况。区域差分(ROD)算法都能够消除这种“鬼影”<sup>[11-12]</sup>。

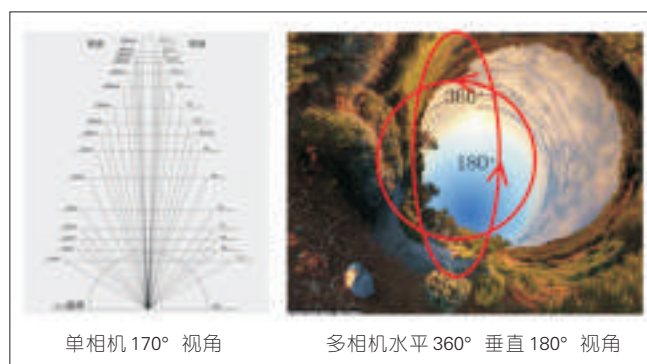
### 2.5 投影映射及编码技术

输入图像经过拼接后投影到一个三维投影结构上,例如一个单位球体或者正六面体。投影结构上的图像数据将被进一步地分布到一个二维平面的投影帧。等角投影(ERP)是目前VR视频主流格式。

ERP投影类似于地球球面展开成地图,将球面展开为平面矩形,所以也被称为经纬图展开。这种方式简单并已经普遍使用,但其缺点也是明显的:球面赤道部分投影展开后失真小,而两极部分纬度越高,失真越大。同时引入了过多的无效像素,球面展开成ERP后,面积即像素数增加了57%,从而降低了传输效率。

正多面体投影(PSP)是业界关注的新方向,具有失真小、压缩效率高的特点。正多面体投影将球体以球心为中心向外划分为多个球面区域,并投影到多面体某个面上,正多面体可以是四面体、立方体、金字塔、十二面体、二十面体等。由于每个球面区域单独投影,因此失真较小。对于面数多的方案,其总面积增加也较小。同时,正多面体投影格式有利于实现基于视点VR视频传输方案,传

图3 VR直播视频拍摄设备覆盖范围





▲图4 图像匹配拼接示例

输部分“面”的视频图像,或者对于不同“面”提供不同的分辨率。

对于单目虚拟现实视频,同一时刻的输入图像拼接后生成一个投影帧,用来代表一个视场。对于双目立体虚拟现实视频,同一时刻的输入图像拼接后生成的一个投影帧代表两个视场:左眼和右眼。两个视场可以映射到相同的封装帧,并基于传统2D视频编码器进行编码,或者投影数据帧的不同视场可以被映射到各自的封装帧。

## 2.6 内容分发网络

内容分发网络(CDN)的基本思路是尽可能避开互联网上有可能影响数据传输速度和稳定性的瓶颈和环节,使内容传输更快、更稳定。通过在网络各处放置节点服务器所构成的在现有的互联网基础之上的一层智能虚拟网络,CDN系统能够实时地根据网络流量和各节点的连接、负载状况以及到用户的距离和响应时间等综合信息将用户的请求重新导向离用户最近的服务节点上。

VR直播业务使用的是全景视屏,现阶段大部分采用的是全传输的方案,带宽占用较大,考虑到终端的解码能力,一般使用展开面积为2~4 K的分辨率,码率一般可以设置在10 Mbit/s以内的HEVC编码方式。

## 3 中兴通讯大视频VR业务的实现

### 3.1 VR直播视频实时采集编码系统

中兴通讯大视频VR全景视频采集系统,通过组合4个4 K单反照相机,每个单反配有鱼镜头,通过

HDMI将视频内容传输到视频编码服务器中,在服务器上实现视频的投影变换、视频拼接等操作。视频内容会编码为多码率格式,传输到远端CDN节点一份,并本地保存一份。

4个4 K单反相机进行视频录制充分保证了视频的清晰度,画面的表现力非常强,在暗光下效果很好,视频拼接后可以拼接出最大12 K分辨率,一般直播使用4 K分辨率。

### 3.2 VR直播无法看清细节问题的解决

为了解决VR 360°全景视频无法看清人脸的问题,中兴大视频项目组在直播中采用了全景视频拼接方法,增加主要演讲人员的高清视频,使用户在观看整体直播时可以在需要时适当转头朝向旁边可以看到宣讲人员的大画面展示,视频拼接的效果很好地解决了用户既想观看全景视频,又想看清演讲者细节的问题。

### 3.3 支持VR直播业务的大视频业务平台

中兴通讯大视频业务系统是一个开放的、先进的、标准的系统,能为全球的客户提供的个性化需求及服务。该系统支持网路协定电视(IPTV)&互联网应用服务(OTT)基础视频业务,包括:直播电视、视频点播(VOD)、直播回看(TVOD)、时移电视(TSTV)等;支持广告、卡拉OK、游戏、早教等增值业务;支持多屏互动体验,包括:TV屏、PC屏、手机屏和PAD屏。实现多屏内容、用户、产品的统一管理,用户可以在任何时间、任何地点,通过任何设备访问IPTV和OTT业务。

在大视频业务系统基础上,新增

加的VR子系统使大视频业务系统增加了VR业务的能力,包括VR直播、VR VOD、VR TVOD。

### 3.4 支持VR直播视频的内容分发网络

中兴通讯大视频融合CDN广泛应用于各类场景需求,以全球运营商视频服务为基础,还为TV游戏、电子阅读、应用商店、互联网内容提供商(IDC)业务、大文件/小文件缓存加速、终端管理和云盘等各种业务提供内容分发服务。

在视频加速方面,中兴大视频融合CDN系统集实时流传输协议(RTSP)和HTTP视频码流传输技术的两家之长,创造性地提出了OTT组播解决方案。不仅能对公众用户提供OTT直播大规模商用解决方案,也可用同一套系统对集团、政企等优质用户提供有保障的RTSP高服务质量(QoS)解决方案。

中兴通讯大视频融合CDN系统无缝集成了VR直播能力,采用分布式文件系统(DFS)实现文件更细分的内容分片聚合,充分保证VR直播业务中巨型分片文件细小化、媒体分发的大带宽和低延迟的产品需求。主动的安全防护,有效地提升了融合CDN产品的安全性,并从系统的主机、网络、应用3个层面对系统进行防护,把安全检查变成系统自身、例行工作,及时发现系统隐藏的隐患;提高系统的安全性,同时根据黑客攻击模型,变被动(事后处理)为主动(防预),做到事前、事中发现,及时处理;降低安全事件的危害。

### 3.5 大视频全流程质量监测控制系统

中兴通讯大视频智能运维系统,支持IPTV、OTT、VR直播等多种视频直播业务,能实现从头端—业务系统—网络侧—用户终端侧的服务质量监测控制与故障定位,如图5所示。

支持VR直播源服务质量的监测控制,大视频智能运维系统在VR直播摄像头入向的频道部署监测点,用



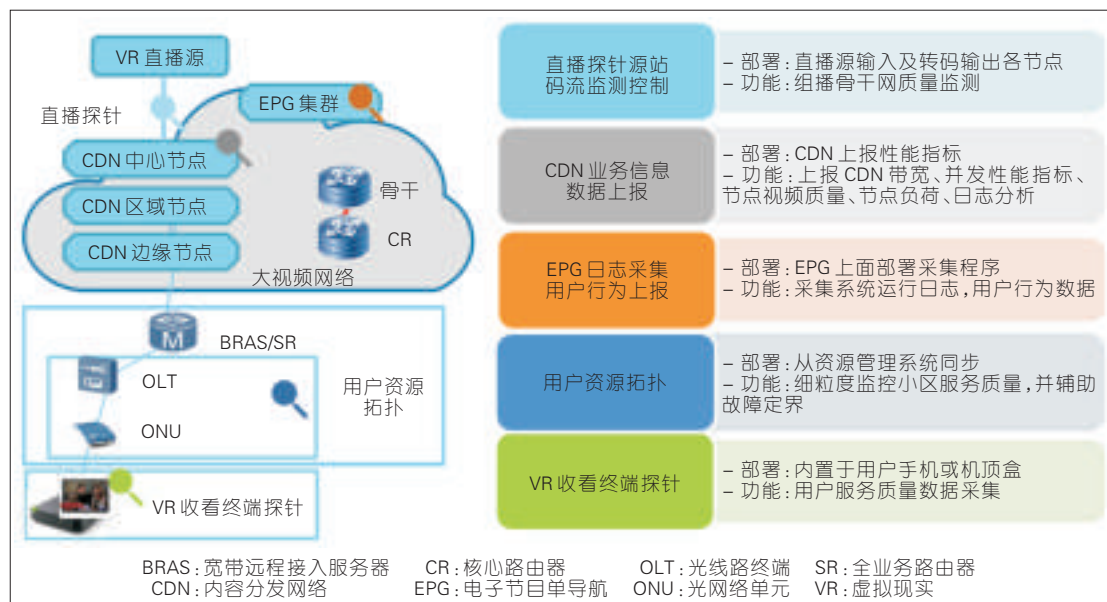


图5  
大视频 VR 直播全流程  
质量监测控制体系

于检测头端的编码质量以及网络 QoS。

(1) 支持 CDN 节点及业务系统质量监测控制。在 VR 直播 CDN 节点及系统侧, 按需部署码流监测点, 检测 OTT 平台本身媒体出口以及相关网络性能指标。监测点探针支持不同的厂家、平台以及不同厂家的终端设备, 帮助运维人员第一时间发现内容源或者平台的故障。

(2) 支持网络侧服务质量监测控制。大视频智能运维系统具备专家系统与推理机功能, 通过与运营商的 IT 系统接口同步网络拓扑数据, 结合先进的知识库与推理机。当用户出现服务质量异常时, 可迅速定位存在的网络故障, 支持全网 VR 直播用户的网络侧故障监测与定位。

## 4 结束语

中兴通讯已经在多次商用场景下进行了 VR 直播, 在这一领域已经积累了丰富的应用经验。目前, 基于 IPTV/OTT 大视频系统的 VR 直播解决方案已经步入商用阶段, 应用场景非常广泛, 适用于如 VR 娱乐直播、VR 教育直播、VR 医疗直播、VR 购物直播、VR 网红直播等, 为用户带来革命性的直播视频体验。

VR 直播业务的背后, 是中兴通讯在大视频领域的深厚沉淀。截至 2016 年底, 中兴通讯大视频方案在全球拥有 90 多个商用局, 系统容量达到 8 000 万, CDN 产品全球局点超过 100 个, 总并发能力超过 100 T, 位居全球领先地位。中兴通讯提出大视频 3.0 的概念, 让系统更加智能、更加弹性, 助力大视频发展迈向新高度。

## 致谢

本研究得到中兴通讯大视频实验室左罗总工和黄珂资深工程师, 以及大视频客户端研发产品涂成义经理的帮助, 谨致谢意!

## 参考文献

- [1] 陈静, 黄慧. 虚拟现实技术在视频直播中的应用分析[J]. 新闻研究导刊, 2016, 7(18): 8-9
- [2] 王跃华. 虚拟现实视频的制作和应用场景分析[J]. 视听, 2016(7): 11-12
- [3] 杨磊. 基于虚拟现实头戴式显示器用户体验研究——以摄像机运动模式为例[J]. 科技资讯, 2017, 15(1): 1-2
- [4] 夏明革, 唐小明, 夏仕昌. 图像融合的发展现状与展望[J]. 舰船电子工程, 2002, (6): 2-12
- [5] 杨恺. 智能手机 HMD 高沉浸感体验的研究与实现[D]. 北京: 北京工业大学, 2017
- [6] BOURKE P. Converting Dual Fisheye Images into a Spherical (Equirectangular) Projection[EB/OL]. (2016-08)[2017-10-23]. <http://paulbourke.net/dome/dualfish2sphere/>
- [7] BOURKE P. Classification of 3D to 2D Projections[EB/OL]. (1994-12)[2017-10-23]. <http://paulbourke.net/geometry/transformationprojection/>

- [8] 张洋, 李庆忠, 臧风妮. 一种多摄像机全景视频图像快速拼接算法[J]. 光电子·激光, 2012(9): 1821-1826
- [9] 林枝叶. 面向全景视频拼接的图像融合算法及其 GPU 实现[D]. 成都: 电子科技大学, 2017
- [10] 胥陈或. SIFT 算法的图像特征处理模块的芯片设计研究[D]. 南京: 东南大学, 2016
- [11] 宋振兴. 图像拼接融合技术中去鬼影及裂缝方法研究[D]. 大连: 大连海事大学, 2011
- [12] 袁晨鸿. 快速创建全景视频及去鬼影方法的研究与实现[D]. 哈尔滨: 哈尔滨工业大学, 2013

## 作者简介



徐火顺, 中兴通讯股份有限公司多媒体业务产品总工; 主要研究领域为电信智能业务、IPTV&OTT 视频业务及增值业务; 获得 4 项省市级科研成果奖; 已发表论文 10 余篇。



叶剑章, 中兴通讯股份有限公司技术规划部资深战略规划师; 主要研究领域为 5G 网络业务及应用、电信增值业务产品、人工智能等; 已发表专利 10 余篇。



尹海生, 中兴通讯股份有限公司多媒体业务产品产品经理; 长期从事电信运营业务相关产品的规划工作, 主要研究领域为 IPTV&OTT 业务运营、智能感知、开放型客户端等相关领域。

# HEVC兼容的全景视频运动补偿预测算法

## Motion Compensation and Prediction for Panoramic Video Coding on HEVC

胡强/HU Qiang<sup>1</sup>  
马文广/MA Wenguang<sup>2</sup>  
石志儒/SHI Zhiru<sup>2</sup>

(1. 上海交通大学, 上海 200240;  
2. 上海科技大学, 上海 201210)  
(1. Shanghai Jiao Tong University,  
Shanghai 200240, China;  
2. ShanghaiTech University, Shanghai  
201210, China)

### 1 三维球面投影及所提编码框架

随着消费级虚拟现实(VR)技术的不断完善, HTC VIVE、Oculus Rift、三星 Gear VR、Deepon 等 VR 设备得到井喷式发展。作为能够快速制作推广的全景视频也随之进入公众的视野, 在景观展示、视频会议、远程医疗和会诊、远程监测控制、机器人自动导航等领域得到广泛应用。目前 VR 视频主要有两种生产途径: 一种是通过渲染引擎创作的纯虚拟场景; 另一种是通过相机阵列拍摄拼接的 360° 视频<sup>[1]</sup>。

使用者在观看 360° 全景视频, 在同一位置通过改变观察方向, 能够浏览四周全部场景<sup>[2]</sup>。360° 视频包含了四周空间的全部运动信息, 利用虚拟现实头盔显示观看时, 只选取其中一部分视野进行渲染显示, 所以全景视频需要更清晰(至少空间分辨率 4 K 以上), 才能匹配现在人们对高清

收稿日期: 2017-10-20

网络出版日期: 2017-11-13

基金项目: 上海市科技人才计划项目(15XD1502900); 上海市科委项目(17JC1403800)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2017) 06-0028-004

**摘要:** 利用图像渲染技术, 全景视频能够为终端用户提供沉浸的虚拟现实体验。不同于传统二维视频, 360° 全景视频不仅包含平移运动还包括仿射运动, 后者主要由于球面映射所产生的畸变。特别是对于广泛使用的经纬球面投影, 其两极区域的扭曲失真特别严重。提出了一种全新的运动补偿预测算法, 其将当前像素投影回球面并寻找其匹配的参考像素块。在此基础上利用球面上运动相对一致, 提出了一种新的运动向量预测(MVP)模式。实验结果表明: 该算法与传统高效视频编码器(HEVC)相比, 最高提升 3.65% 平均码流差(BDBR)的编码性能。

**关键词:** 全景视频; 虚拟现实; 运动补偿; 视频编码

**Abstract:** Virtual reality can provide immersive panoramic video for the end user experience by image rendering technology. Different from the traditional 2D video, 360° panoramic video includes the translational motion and the affine motion. The affine motion is mainly due to the distortion generated by the spherical mapping. For the latitude and longitude spherical projection which is widely used, the polar area distortion is especially serious. In this paper, a novel motion compensated prediction scheme is proposed to deal with non-translational motions for panoramic video. Furthermore, a compensation of motion vector predictor (MVP) is employed on the sphere plane to further improve the coding performance. Experimental results show that the proposed method reaches a maximum 3.65% Bjontegaard Delta Bitrate (BDBR) reductions compared to high efficiency video coding (HEVC).

**Keywords:** panoramic video; virtual reality; motion compensation; video coding

分辨率的体验。全景视频的数据量相比于传统视频是成倍的增长趋势, VR 全景视频的点播和直播的传输面临带宽瓶颈<sup>[3]</sup>。传统视频编解码的方案不太适应全新的视频格式的需求, 需要根据全景视频的特性提出有针对性的编解码方案, 满足存储和传输的需求<sup>[4]</sup>。

360° 全景视频应用主要包括前端相机阵列的视频同步采集、拼接、编码及传输和播放。其中多相机阵列同步采集的各视角视频, 进行校正配准之后, 对重叠的图像无缝拼接, 需要将图像进行投影变换到三维球面模上。投影到球体表面模型进

行拼接处理, 然后再以一定的投影模式变换成二维图像, 进行压缩传输和储存。所谓的投影变换是指将全景视域的三维场景和二维平面相互映射过程, 作为全景视频内容的最广泛的转播载体, 投影后的二维视频, 最终通过渲染在虚拟现实终端设备进行观看。在投影变化和反变换过程中, 存在着图像失真, 运动一致被破坏等问题, 传统的视频编码技术在应对 360° 全景视频的效果并不理想<sup>[5-7]</sup>。

针对以上存在的问题, 由国际电信联盟电信标准分局(ITU-T)/视频编码专家组(VCEG)和国际标准化组织-国际电工委员会(ISO-IEC)/动态

图像专家组(MPEG)成立的联合视频研究组(JVET)将 $360^\circ$ 全景视频纳入到未来视频编码标准范围内<sup>[8]</sup>。与传统的高效视频编码器(HEVC)/H.265编码标准<sup>[9]</sup>相比,正在制订的联合视频探测组(JVET)标准,将通过优化投影变换避免过多的扭曲失真,作为一个重要的实验测试和研究方向,从而避免过多的扭曲失真产生的回放质量。目前使用最为广泛的是球面投影(ERP)和立方图投影(CMP)。

球面投影是把球型表面的内容映射到平面的矩形图中,如图1所示,其对自然图像有良好的兼容性,在转换过程中不同维度区域有不同的延展程度,赤道区域的失真比两级区域小很多,像素的数量在两级被放大,运动方向扭曲,产生大量的时间和空间的冗余信息<sup>[10]</sup>。而立方图投影通过外接立方体,将球体表面内容投影到立方体表面。展开后按照一定的布局格式进行布局,通常的布局方法有 $3\times 2$ 布局和 $4\times 3$ 布局,如图2所示。不用的布局方案对编码效率和内存管理都有不同程度的影响。相比球面投影,立方图投影节省了25%的像素面积,对比特率有较大的提升,但存在像素密度减少照成的模糊问题<sup>[11]</sup>。文章中,我们针对全景视频投影过程中产生畸变和运动扭曲,所带来的编码问题,提出了一种全新的运动补偿算法。

在传统的视频编码算法中,基于块匹配的运动估计和运动补偿得到广泛的应用。其建立在相邻时域帧内的二维物体所做的都是平移运动,大小形状不改变的基础上,对动态帧之间的时域冗余有非常高压缩效果。但是对于全景视频,由于从三维球面坐标投影到二维平面坐标转换的过程中,产生了一系列相机旋转、缩放、错切运动的形变,影响了运动的连贯性和一致性,从而降低了帧间预测的效率。

针对全景视频出现的大量非平

移运动,我们提出了一种全新的运动补偿和预测的模式,用于提高 $360^\circ$ 全景视频的时域编码效率。为避免几何失真对运动预测带来的影响,首先将当前编码块的像素投影到三维球面坐标,利用球面运动一致性的特点,在球面坐标系中找相对应的参考像素,并最终将三维球面坐标投影回二维坐标。更进一步,我们提出了基于球面坐标的运动向量预测方法,避免了投影变换带来非平移运动的影响,视频编码框架如图3所示。我们提出的一整套运动补偿预测算法都在块划分的基础上提出,可以作为HEVC编码的附加模式,对 $360^\circ$ 全景视频有着更高效的编码压缩性能。

## 2 基于三维球面投影的运动补偿算法

为了提升 $360^\circ$ 全景视频的编码压缩性能,我们针对全景视频中大量产生的非平移运动,进行了深入的分

析,其中绝大部分非平移运动,是由投影变换产生的畸变导致。针对这种情况,我们在HEVC编码框架下提出了一种全新的运动补偿模式;基于三维球面投影的运动补偿(SphMC)。整体算法流程如图3所示。本算法作为 $360^\circ$ 视频的一种补充,同时存在于传统的帧间预测的框架内。HEVC编码过程中,最优的预测模式仍然是遍历所有块划分和预测模式后,选择最小率失真值的模式,作为最优的编码模式。其中SphMC模式的加入,提供了一种针对全景视频的运动补偿模式,对整体编码性能有明显的增益。

基于前面章节分析,纹理的不连续性会严重影响运动预测的性能,针对全景投影视频的有效参考像素,需要从运动连续的三维球面坐标系中推导而来。文中涉及到的投影变换都以立方图投影为例,SphMC模式的参考块像素点的展示在图4中。假

图1 球面三维坐标与球面投影变换

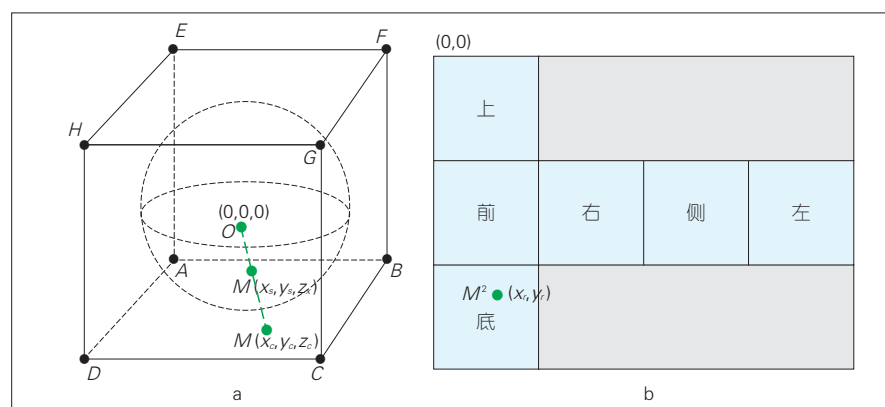
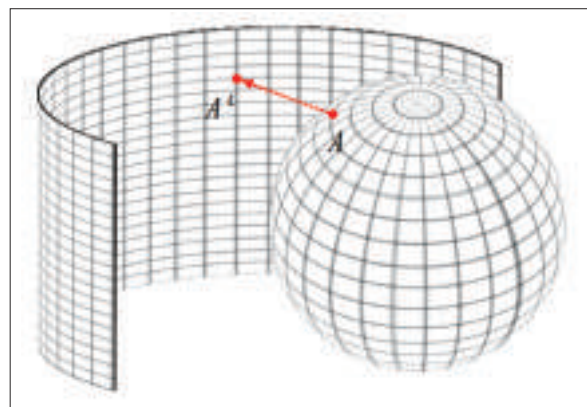
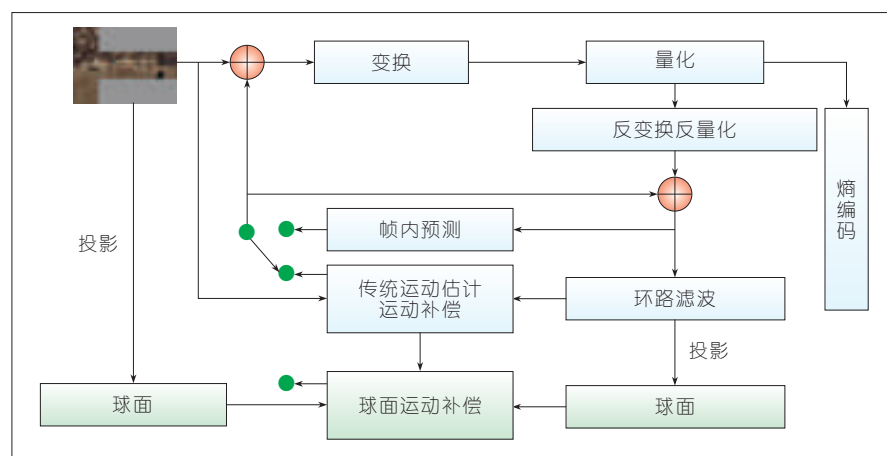


图2 球面三维坐标与立方图投影变换





▲ 图3 基于三维球面投影的运动补偿和预测模式

设当前编码块所在帧  $t$ ，该块中心位置  $n_0$  移动到帧  $t$  帧的  $n_2$  位置，基于三维球面的投影问题就简化为：在参考帧中利用球面投影变换，寻找当前编码块  $n_1$  的对应参考点  $n_3$ 。

为了在参考帧得到出对应的点  $n_3$  的位置，其在三维球面上的坐标位置首先会被推出来，然后再投影回立方图。在这个投影关系中，立方图和三维球面坐标是一一对应。 $s = P(n)$  代表了从立方图平面  $N$  到三维全面  $S$  的映射，反变换由  $n = P^{-1}(s)$  表示。所以点  $n_0$ 、 $n_1$  和  $n_2$  的坐标投影到三维球面分别为： $s_0 = P(n_0)$ 、 $s_1 = P(n_1)$  和  $s_2 = P(n_2)$ 。在三维球面的表达中，绝大多数的像素点和运动具有一致性，所以  $s_3$  的可通过以下式(1)得到：

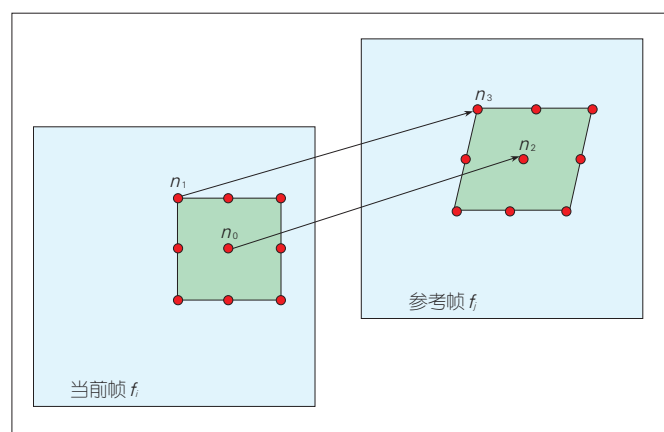
$$s_3 = P(n_2) + P(n_1) - P(n_0) \quad (1)$$

得到在三维球面坐标点  $s_3$  后，其

在立方图坐标系中的相对应的二维坐标  $n_3$  可以通过反投影  $n_3 = P^{-1}(s_3)$  得到。其余参考点可以通过相似的方法推演得到。对于非整像素的位置，可通过双线性差值得到。最后，基于三维球面投影得到的参考区域与编码块进行运动补偿的计算。

### 3 基于三维球面投影的运动向量预测算法

在 HEVC 帧间预测过程中，会针对空域和时域上已经确定的运动向量，来预测当前块运动估计的初始向量方位，该过程通过相邻预测单元 (PU) 的运动向量来进行预测。但对于 360° 全景视频，由于投影变换带来的空间不连续，传统的基于临块的运动预测已经不太适用。本章中，我们提供了一种基于三维球面投影的运动向量，其针对不同 PU 的相对距离，



▲ 图4 基于三维球面投影的运动补偿算法

如图5所示。

针对 360° 全景视频的帧间预测，假设  $n_4$  和  $n_5$  分别为分块 C 和分块  $B_1$  的中心点， $MV_1$  为分块  $B_1$  的运动向量，而  $MV_0$  代表了当前编码块 C 的预测向量。类似于我们提出的运动补偿方法， $n_4$  和  $n_5$  都利用二维立体图平面投影回三维球面坐标，从而寻找正确的运动预测向量。我们可以得到以下关系：

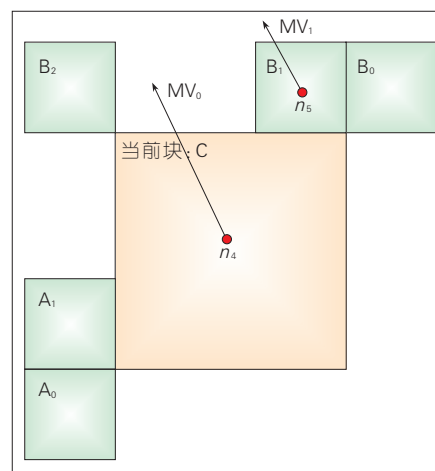
$$\begin{aligned} P(n_2 + MV_0) - P(n_4) &= \\ P(n_5 + MV_1) - P(n_5) \end{aligned} \quad (2)$$

所以，基于球面投影的运动向量  $MV_0$  可以由此得到：

$$MV_0 = P^{-1}(P(n_5 + MV_1) + P(n_4) - P(n_5)) - n_4 \quad (3)$$

## 4 实验结果

为了验证所提算法的有效性，基于投影变换的运动补偿算法已集成到了 HEVC 标准 HM16.15<sup>[12]</sup> 测试软件中，通过对比 HM16.15 与该算法的性能，证明了改算法的有效性。基于 HEVC 常用测试条件<sup>[12]</sup>，我们主要测试 3 种配置：randomaccess\_main (RA\_main)、lowdelay\_main (LD\_main)、lowdelay\_P\_main (LP\_main)。所采用的 QP 分别是 22、27、32、37，包括高比特率和低比特率的情况。鉴于该算法的目标在于更好地表征 360° 视频立



▲ 图5 针对球面投影模式的运动向量预测

体图投影中的无规则运动,我们选择了一些相对于立方体4×3格式有更大运动的序列来表明该算法的优势。测试序列信息如表1所示。另外,由于这些序列都具有极高的空间分辨率,因此我们只测试了1 s所对应的帧数来证明算法的有效性。

RA\_main、LD\_main、LP\_main的实验结果分别如表2、表3所示。从这些表格中,我们可以看到:通过该算法,在RA\_main、LD\_main两种配置下,视频Y通道分别可达到1.83%,1.21%,1.19%平均码流差(BDBR)<sup>[13]</sup>的减少。鉴于编码顺序相邻帧之间

▼表1 实验所用测试序列

测试序列	分辨率	帧数
SkateboardInLot	4 736 × 3 552	32
DrivingInCountry	3 840 × 2 880	24
Bicyclist	3 840 × 2 880	24
Glacier	3 840 × 2 880	24
ParamotorTraining	3 840 × 2 880	24
TimelapseBuilding	3 840 × 2 880	24

▼表2 RA-Main配置下所提算法实验结果

测试序列	BDBR/(%)Y	BDBR/(%)U	BDBR/(%)V
SkateboardInLot	-1.33	-0.98	-0.85
DrivingInCountry	-2.72	-0.73	-1.31
Bicyclist	-1.15	-1.02	-1.44
Glacier	-3.65	-2.72	-3.04
ParamotorTraining	-0.42	-0.33	-0.56
TimelapseBuilding	-1.72	-1.24	-1.44
平均	-1.83	-1.17	-1.44

BDBR: 平均码流差

▼表3 LD-Main配置下所提算法实验结果

测试序列	BDBR/(%)Y	BDBR/(%)U	BDBR/(%)V
SkateboardInLot	-1.24	-0.23	-0.65
DrivingInCountry	-1.83	-0.68	-1.12
Bicyclist	-1.12	-0.31	-1.08
Glacier	-1.53	-0.62	-0.93
ParamotorTraining	-0.34	-0.15	-0.21
TimelapseBuilding	-1.23	-0.75	-0.71
平均	-1.21	-0.45	-0.78

BDBR: 平均码流差

的平均距离要比LD\_main和LP\_main中的距离大一些,RA中的不规则运动会更加明显。这就是为什么我们能实现更好的R-D性能。特别是对于具有较大运动的序列Glacier,算法可将RA\_main中Y通道的平均码流差减少3.65%。因此,对于具有较大运动的序列,就可实现更显著的平均码流差的减少。

## 5 结束语

360°全景视频与传统的视频存在着许多差别,传统的HEVC编码不能达到最理想的压缩效果,我们分析了由于投影变换带来的运动失真的问题,并相应地提出了一种基于球面投影的运动补偿和预测算法,其主要利用三维球面的运动具有较强的一致性,通过立方体二维坐标与球面三维坐标的相互转换,寻找出新的参考像素和运动向量预测。经过大规模的实验,结果表明:该算法与传统HEVC相比,最高提升3.65%BDBR的编码性能。在后续的研究中,以下问题将会着重进行考虑:提出针对360°全景视频的运动估计算法,并且修改相应的率失真准则,以更好地适应360°全景视频的编码需求。

## 参考文献

- [1] ROUTHIER P P. Virtually Perfect: Factors Affecting the Quality of a VR Experience and the Need for a VR Content Quality Standard [C]// Processing of SMPTE Annual Technical Conference and Exhibition. USA:IEEE, 2016: 1-20. DOI: 10.5594/M001733
- [2] 张洋, 李庆忠, 臧风妮. 一种多摄像机全景视频图像快速拼接算法[J]. 光电子. 激光, 2012, 23(9): 1821-1826
- [3] NG K T, CHAN S C, SHUM H Y. Data Compression and Transmission Aspects of Panoramic Videos [J]. IEEE Transaction on Circuits and Systems for Video Technology, 2005, 15(1): 82-95. DOI: 10.1109/TCSVT.2004.839989
- [4] 张洪彬, 伏长虹, 苏卫民, 等. 3D-HEVC深度图像帧内编码单元划分快速算法[J]. 电子与信息学报, 2016, 38(10): 2523-2530
- [5] DE LA, FUENTE Y S, SKUPIN R, et al. Video Processing for Panoramic Streaming Using HEVC and its Scalable Extensions [J]. Multimedia Tools and Applications, 2016: 1-29
- [6] ZHU Z G, XU G Y, EDWARD M R, et al. Fast Construction of Dynamic and Multi-Resolution 360° Panoramas from Video

- Sequences [J]. Image and Vision Computing, 2006, 24(1): 15-17
- [7] ZHENG J L, SHEN Y F, ZHANG Y D, et al. Adaptive Selection of Motion Models for Panoramic Video Coding [C]// Process of IEEE International Conference on Multimedia and Expo. USA:IEEE, 2007: 1319-1320
- [8] NORKIN A, YANG H, OHM J R, et al. Call for Test Materials for Future Video Coding Standardization [C]// Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, ITU, 2016
- [9] SULLIVAN G J, OHM J, WOO J H, et al. Overview of the High Efficiency Video Coding (HEVC) Standard [C]//IEEE Transaction on Circuits System Video Technology. USA:IEEE, 2012,(22):1649-1668
- [10] LIN H, LIN J L, CHANG S K, et al. AHG8: An Alternative Arrangement for Cube Format (double-cross layout) [C]// Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11. ITU, 2016
- [11] KUZUYAKOV E, PIO D. Next-Generation Video Encoding Techniques for 360 Video and VR [EB/OL]. (2016-07-15)[2017-11-07]. <https://code.facebook.com/posts/1126354007399553/next-generation-video-encoding-techniques-for-360-video-and-vr/>. 2016. 7. 15
- [12] JCT-VC Subversion Respository for the HEVC Test Model Version HM16.15[EB/OL]. [2017-11-07]. [https://hevc.hhi.fraunhofer.de/svn/svn\\_HEVCSoftware/tags/HM-16.15](https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.15)
- [13] BOSSEN F. Common HM Test Conditions and Software Reference Configurations, Document JCTVC-K1100[R]. China: Shanghai, 2012

## 作者简介



胡强, 上海交通大学图像通信与网络工程研究所在读博士; 研究方向主要包括图像处理和视频编码; 已发表论文4篇, 获得专利4项。



马文方, 上海科技大学信息与工程学院研究助理; 研究方向主要包括光场摄影技术、计算成像技术及相关设备制作以及虚拟现实技术。



石志儒, 2012年在英国萨里大学电子工程系获得博士学位, 之后于上海交通大学电子信息与工程学院从事博士后研究, 现任上海科技大学副研究员; 从事计算机视觉、视频编解码、虚拟现实等方面的研究; 已发表学术论文10余篇。

# 移动增强现实应用及进展

## Application and Development of Mobile Augmented Reality

潘志庚/PAN Zhigeng<sup>1,2</sup>  
高翔/GAO Xiang<sup>2</sup>  
丁丹丹/DING Dandan<sup>2</sup>

(1. 广州玖的数码科技有限公司, 广东 广州 510000;

2. 杭州师范大学, 浙江 杭州 310036)

(1. Guangzhou nine Digital Technology Co., Ltd., Guangzhou 510000, China;

2. Hangzhou Normal University, Hangzhou 310036, China)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2017) 06-0032-005

**摘要:** 移动增强现实技术是增强现实(AR)的一个分支,其目前主要有传统移动增强现实技术和基于移动终端的移动增强现实技术两种。移动增强现实设备的普及和技术的高速发展使其可以应用到多个领域。对移动增强现实技术的应用进行综述,并对其商业应用前景和商业模式进行展望。

**关键词:** 虚拟现实(VR); AR; AR应用

**Abstract:** Mobile augmented reality is a branch of augmented reality (AR), and it has traditional mobile augmented reality technology and mobile augmented reality technology based on mobile terminal. With the popularity of mobile augmented reality devices and the rapid development of technology, mobile augmented reality can be applied to many fields. In this paper, the application of mobile augmented reality is summarized, and the commercial applications prospect and the business models are also discussed.

**Keywords:** virtual reality (VR); AR; AR application

### 1 移动AR简介

虚拟现实(VR)是一种对现实世界的数字化模拟技术,包括对于环境的视觉、听觉、触觉等方面的感觉模拟。用户通过使用必要的交互设备,最终获得身临其境的感受。VR涉及多个学科,包括心理学、计算机科学、电子学、多媒体等,具有极强的学科交叉性,目前已经成为了科研中一种常用的辅助手段。VR的典型特征,即沉浸感、交互性和构想性。沉浸感指的是用户在虚拟环境中,得到了视觉、听觉、嗅觉、味觉、触觉等多种感官刺激,仿佛身临其境的感受。交互性是指在VR环境中用户能够拥有人性化的人机交互方式和操作。构想性指的是通过前两个特性,用户能够根据环境变化和交互变化而激发出更大的创意和想法。图1是我们开发的虚拟健身系统。

增强现实(AR)是VR的技术拓展,它是一种将真实世界信息和虚拟世界信息“无缝”集成的新技术,将真

实环境和虚拟物体实时地叠加到了同一个画面或空间中。AR强调的是现实世界叠加虚拟世界,从而实现“虚实结合”。移动AR技术是AR的一个分支。目前主要有传统移动AR技术和基于移动终端的移动AR技术两种。图2是我们开发的AR乐器虚拟教学系统。

传统移动AR技术主要依靠PC机、头盔显示器或者全球定位系统(GPS)、磁传感器等一些外接设备来完成系统。该类系统具有设备昂贵,不利于长期携带,维护成本较高等局限性。美国哥伦比亚大学研制的户外AR导航系统,也是世界上第一个移动AR系统。它使用视频透视式头盔显示器和方向跟踪器,用户身后的背包里面装有计算机、GPS以及无线电,手上是一个具有触摸笔的可触屏的手持设备。

而基于移动终端的移动AR技术近几年发展迅速,为人们的生活娱乐等带来极大的便捷和乐趣。图3所示为谷歌眼镜,是谷歌公司2012年4月发布的一款“AR”眼镜。它同智能手机一样,主要采用音频交互的方式,可以控制拍照、视频通话和导航以及网上冲浪、处理文字信息和电子邮件等。2013年11月12日,谷歌公司发布一系列新功能,包括搜索歌曲、扫描已保存播放列表,以及收听



▲ 图1 虚拟健身系统

收稿日期: 2017-11-14  
网络出版日期: 2017-11-29





▲ 图2 AR乐器虚拟教学系统



▲ 图3 谷歌眼镜

高保真音乐等。在某种程度上,谷歌眼镜改变了人们的生活方式。

移动AR技术最重要的特征是移动,它的应用场景更加广大,不仅仅只是局限在桌面上。因此移动AR技术将会是未来发展的主流趋势,尤其是基于移动终端的移动AR技术。虽然AR技术的应用在硬件设备以及发展形式上发生了变化,但是AR技术的关键问题和关键技术依然没有改变。而且,随着移动设备的硬件与软件条件越来越高端,移动AR技术也得到了突飞猛进的发展。

## 2 移动AR的应用

移动AR的应用面很广,限于篇幅无法一一列举,重点介绍几类。

### 2.1 教育与娱乐

相对于传统的教学方式来说,移动AR能帮助学生更加高效学习,直观的演示和生动的场景更是提升了知识吸收率和留存时长。早在2001年,Billinghurst等人<sup>[1]</sup>发明了一种“魔法书(MagicBook)” (如图4所示),将可交互的数字内容覆盖到真实书本

上,为读者提供了一种在真实书本和沉浸式虚拟环境的过渡效果。Dünser在2007年<sup>[2]</sup>和2008年<sup>[3]</sup>分别做了实验证明魔法书确实能提升学生的阅读体验和复述效果。自此之后,AR图书开始大力发展,现在市面上已经有很多这样的产品。AR图书是将文本内容进行了增强,还有很多产品是将图形内容进行增强。Feng等人<sup>[4]</sup>发明的MagicToon可以将纸上的画转化成栩栩如生的卡通形象。通过简单便捷的点击拖拽交互即可编辑生成的3D模型。

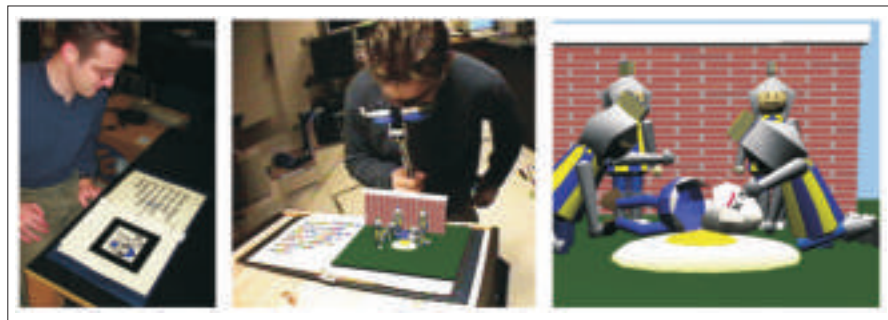
Fonseca等人<sup>[6]</sup>观察学生用传统和移动AR的方式去学习建筑学知识,发现学生用后者学习对“空间”的理解更深刻,直接导致他们的课程评估分数变高。可以看出:在需要立体模型或强交互的教学场景中,移动AR可以帮助人们更形象、直接地观察3D结构,从而使学习更加有效。因此在天文学、化学、生物学、建筑学、工程学、设计和几何学等学科,移动AR都已经有了类似的应用。Nguyen等

人<sup>[7]</sup>设计了一个用户可交互、可编辑的说明书,生动地教用户操作。

### 2.2 建筑工程

传统的建筑图或工程图通常都是2D视图,从不同角度绘制的视图或电脑辅助设计建模视图等。2D视图虽然利于工程计算,但是通常需要工程师有非常强的空间想象能力才能在脑海中构建出大致模型,3D视图则有可能导致失真、比例不对和丢失了真实建筑的细节等问题。此外,这些工程制图大部分都需要专业人士去解读,不同的水平也容易产生误解。如果使用移动AR,这些问题都可以充分解决。Velasco等人<sup>[7]</sup>提出了一种在移动AR中通过交互来规划建筑的设计方式。文献[8-9]评估了移动AR在建筑规划方面的优势,一致认为移动AR的直观性促进了决策制定和对规划的理解。不仅如此,通过结合传感器收集的数据、开发者提供的数据,可以将生成的各类3D模型叠加到现实场景中,使建筑规划、维护、勘探和装配都更加高效。

Schall等人<sup>[10]</sup>开发了一个手机应用,可以生成像X光一样的视图,用于展示建筑或设施内部的框架结构,帮助现场勘探人员之间观察建筑内部结构。Schall等人<sup>[11]</sup>还提供了一个增强现实的测绘工具,通过让有经验的工程师使用传统技术和完全没经验的工程师使用该系统进行对比,最后发现其在交互性方面胜过传统技术。Lee等人<sup>[12]</sup>开发了一个能快速搜索钻井日志并用增强现实可视化出



▲ 图4 用“魔法书”在虚拟与现实之间浏览书籍内容

来的系统。Kollatsch 等人<sup>[13]</sup>则用增强现实可视化出了装配流水线相关的信息。此外移动 AR 在灾后(如地震)建筑勘察和检测方面也有应用。Behzadan 等人<sup>[14]</sup>开发的视觉避撞系统可以帮助挖掘机操作员在工程实施时看到埋在建筑中的东西避免造成破坏或人员伤亡,如图 5 所示。如图 6 所示,可在 AR 帮助下将先前存储的建筑物基线叠加到现在的建筑物上,实现对建筑物损坏程度的评估。

### 2.3 地理与导航

导航一直是增强现实的一个重要应用。传统的导航都是在 2D 地图上,缺乏三维信息和对真实世界的感知。在移动 AR 中,可以直接在视野中添加方向注释(箭头、路径跟随等)或一些必要的提示(交规、停车场位置等),如图 7 所示。典型的应用有飞机飞行导航和汽车驾驶导航等。通过对视野内事物的识别和追踪,动态地将需要的信息图像叠加到真实场景上。除了交通方面,移动 AR 在旅游景点、博物馆、图书馆或商场的浏览和探索方面也有应用。这类导览式的应用都大同小异,如博物馆中,参观者可以通过移动 AR 和工艺品进行转动等交互,更好地观察工艺品的全貌,还可以通过 AR 展示出工艺品的数据信息等<sup>[4]</sup>。商场中可以直接通过移动 AR 展示商品的信息,不用亲自动手去货架上查看,甚至可以在移动 AR 中比货。

### 2.4 医疗

移动 AR 可以将电子计算机断层扫描(CT)或磁共振成像(MRI)等医学三维重建得到的虚拟模型图像融合到病人相应的部位、器官,使医生的视角得到增强<sup>[15]</sup>,在看到真实身体部位的同时,还能看到器官内部的具体信息。这对于医学手术计划制定、手术演练、手术教学和术前对话等都提供了巨大的便利。典型的工作有用移动 AR 帮助验尸<sup>[16]</sup>,通过将 CT 图

像叠加到尸体上,可以保证法医在不丢失真实伤口信息的情况下,观察分析肉眼不可见的内伤,而且不会对尸体造成破坏,从而保护了证据(如图 8 所示)。Soeiro 等人<sup>[17]</sup>开发了一款增强现实 Android 应用来指导医生进行经颅磁刺激技术(TMS)治疗。其中增强现实部分用来显示虚拟的病人脑模型,以便医生了解目前病人脑中的脉冲变化。Hansen 等人<sup>[18]</sup>开发了一种减少三维规划模型复杂性的视觉算法,将 3D 模型更好地可视化到

真实器官上,并在外科手术中运用。

## 3 移动 AR 的应用前景

### 3.1 移动 AR 的商业应用模式

根据网易科技报道<sup>[19]</sup>,移动 AR 技术的应用前景多种多样。虽然目前有很多顶尖的 AR 企业提供一系列的产品、服务和更多其他的东西,但归纳起来基本上属于 5 种基本类型的公司,他们“钱”景无限。提供的技术产品或服务(也称为商业模式)如

图 5  
增强现实视觉避撞助手

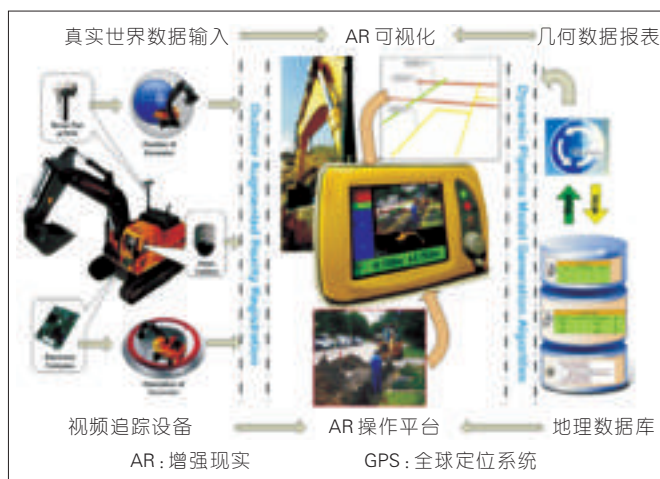


图 6  
建筑破坏检测



▲ 图 7 飞行员 AR 导航和 AR 驾驶训练





▲图8 AR 医学应用

表1所示。

### 3.2 移动AR的市场预测

根据VR网报道<sup>[20]</sup>, AR/VR已经成为一个双速前进的市场, 其中移动AR在2021年的安装基数将超过2017年总AR/VR市场的2倍。苹果ARKit、谷歌ARCore和Facebook Camera Effects平台的推出改变了整个市场的轨迹, 到2018年有可能获得9亿安装基数。Digi-Capital已经彻底修改了关于AR/VR市场的分析和预测报告。

- 由于移动AR的出现, VR的市场潜力减弱, 移动AR在未来将支配AR/VR市场;

- 智能手机仍然是AR/VR的未来, 但可能需要10年时间才能成为一个大众消费市场;

- 直到第2代VR一体机(无需连接PC或移动设备)在2019—2020年

期间爆发之前, 高端VR的发展可能不会提速;

- 由于头显厂商和开发者正在转向移动AR, 到2021年, 移动AR的安装基数将增长至30亿以上;

- 智能眼镜、高端VR和移动VR的组合在同一时间段内可能达到1亿(即每个平台数千万), 这使得移动AR的安装基数超出所有AR/VR头显的25倍以上。

在2021年, 主导AR/VR安装基数的移动AR将占据总市场营收的2/3。移动AR软件的体量跟总移动市场的体量接近, 尽管每用户平均收入(ARPU)较低, 但巨大的用户基数将能带来高增长和高盈利。得益于硬件销售, 智能眼镜、高端VR和移动VR的ARPU更高, 但较小的安装基数限制了其非硬件营收的潜力。在2018年, 移动AR的营收将达到智能眼镜的4.8倍, 高端VR的1.3倍, 以及

移动VR的1.5倍; 而2021年将增长至智能眼镜的3.9倍, 高端VR的4.8倍, 以及移动VR的9.6倍。

鉴于移动AR的安装基数和转换率的提高, 电子商务在2021年可能成为AR/VR的最大营收来源。Houzz已经通过移动AR实现了11倍的销售转换, 所以移动AR电子商务可以进一步巩固亚马逊、eBay和阿里巴巴的主导地位。硬件销售(特别是消费类智能眼镜和第二代高级VR一体机)为第二营收来源; 其次是广告、游戏和非游戏应用收入。AR将成为企业/B2B营收的主要驱动驱动, 而视频和线下娱乐(主题公园)营收则处于较低水平。

## 4 结束语

从前面的介绍可知, 移动AR的应用已经扩展到我们生活的多个方面。AR和VR的发展一开始受5G网络的限制。5G网络被视为解锁市场机会的关键, 拥有低延迟、高可靠和网络统一性的优势, 能够应付VR和AR对网络日益增长的速度和容量的需求。对于移动AR应用扩展的速度以及应用的广度, 让我们拭目以待。

### 参考文献

- [1] BILLINGHURST M, HIROKAZU K, POUPYREV I. The Magicbook: A Transitional AR Interface[J]. Computers & Graphics, 25(5): 745–753, 2001
- [2] DUNSER A, HORNECKER E. Lessons from an AR Book Study[C]//In Proceedings of the 1st International Conference on Tangible and Embedded interaction. USA: ACM, 2007: 179–182
- [3] DUNSER A. Supporting Low Ability Readers with Interactive Augmented Reality[J]. Annual Review of Cyber Therapy and Telemedicine, 2008, 12(1): 41–48, 2008
- [4] WHITE M, LIAROKAPIS F, DARCY J, et al. Augmented Reality for Museum Artefact Visualization[C]//Proceedings of Irish Workshop on Computer Graphics, Eurographics Ireland Chapter. UK: Northern Ireland, 2003: 75–80, 2003
- [5] FENG L, YANG X, XIAO S. MagicToon: A 2D-to-3D Creative Cartoon Modeling System with Mobile AR[C]//Virtual Reality. USA: IEEE, 2017: 195–204
- [6] FONSECA D, VILLAGRASA S, MARTI N, et al. Visualization Methods in Architecture

▼表1 商业模式

编号	服务形式	服务内容
1	AR开发平台 SDK	为经验丰富的软件开发人员提供基础开发平台和工具箱, 以开发先进的AR增强现实解决方案, 其中一些做AR平台的公司还建立和推广自己的产品, 或提供客户服务
2	AR产品和游戏	开发和销售自己的专属AR产品, 例如: 直接面向消费者零售的书籍或游戏, 这些公司使用移动平台的AR, 但在更大程度上它们依赖于专用的设备, 这类公司包括索尼、微软和任天堂
3	自助AR工具和通用AR浏览器	专为快速简单的AR体验或活动而设计, 提供内容管理工具和基本AR效果的简单菜单; 借助自助AR工具, 精通技术的个人可以创建简单的体验, 例如: 发布单个视频或简单的动画
4	定制的AR应用开发	直接与品牌营销人员和机构合作, 为大型广告活动、贸易展览和现场活动构建定制的增强现实解决方案; 允许营销人员结合个性化的服务, 通过项目管理打造独一无二的AR增强现实体验
5	特定垂直行业的AR解决方案	专门服务于某些专业领域, 如: 奢侈品零售、医疗服务、工业应用、制药公司和化妆品公司

AR: 增强现实 SDK: 软件开发工具包



- Education Using 3D Virtual Models and Augmented Reality in Mobile and Social Networks[J]. Procedia – Social and Behavioral Sciences, 2013, 93(1):1337–1343
- [7] VELASCO R, BRAKKE A P, CHAVARRO D. Computer-Aided Architectural Design Futures[M]. The Next City – New Technologies and the Future of the Built Environment. Germany: Springer-Verlag
- [8] OLSSON T D, SSAVISALO A T, HAKKARAINEN M, et al. User Evaluation of Mobile Augmented Reality in Architectural Planning[J]. eWork and eBusiness in Architecture, Engineering and Construction, Gudnason G. & Scherer R. (Eds.), 2001:203–211. DOI:10.1201/b12516-116
- [9] LEE G A, DUNSER A, KIM S, et al. Cityviewer: A Mobile Outdoor AR Application for City Visualization[C]// In Mixed and Augmented Reality (ISMAR-AMH), 2012 IEEE International Symposium on. USA: IEEE, 2012: 57–64
- [10] SCHALL G, MENDEZ E, KRUIJFF E, et al. Handheld Augmented Reality for Underground Infrastructure Visualization[J]. Personal & Ubiquitous Computing, 2009, 13(4):281–291
- [11] SCHALL G, ZOLLMANN S, REITMAYR G. Smart Vidente: Advances in Mobile Augmented Reality for Interactive Visualization of Under-Ground Infrastructure[J]. Personal & Ubiquitous Computing, 2013, 17(7):1533–1549
- [12] LEE S, SUH J, PARK H D. BoreholeAR: A Mobile Tablet Application for Effective Borehole Database Visualization Using an Augmented Reality Technology[J]. Computers & Geosciences, 2015, 76:41–49
- [13] KOLLATCH C, SCHUMANN M, KLIMANT P, et al. Mobile Augmented Reality Based Monitoring of Assembly Lines [J]. Procedia Cirp, 2014, 23:246–251
- [14] BEHZADAN A H, DONG S, KAMAT V R. Augmented Reality Visualization: A Review of Civil Infrastructure System Applications [J]. Advanced Engineering Informatics, 2015, 29(2):252–267
- [15] 张军毅, 罗述谦. 医学增强现实建模方法研究[J]. 系统仿真学报, 2009, 21(12):3658–3661
- [16] KILGUS T, HEIM E, HAASE S, et al. Mobile Markerless Augmented Reality and Its Application in Forensic Medicine [J]. International Journal of Computer Assisted Radiology & Surgery, 2015, 10(5):573–586
- [17] SOEIRO J, CLAUDIO A P, CARMO M B, et al. Mobile Solution for Brain Visualization Using Augmented and Virtual Reality[C]// International Conference Information Visualisation. USA: IEEE, 2016:124–129. DOI: 10.1109/IV.2016.18
- [18] HANSEN C, WIEFERICH J, RITTER F, et al. Illustrative Visualization of 3D Planning Models for Augmented Reality in Liver Surgery[J]. International Journal of Computer Assisted Radiology & Surgery, 2010, 5(2):133–141
- [19] 移动AR才是未来?细数AR公司的五种商业模式.(2016–11–01)[2017–11–10].http://tech.163.com/16/1101/13/C4PR76NP00097U7U.html
- [20] 移动AR未来将支配AR/VR市场.(2017–10–10)[2017–11–10].https://www.hiavr.com/

## ←上接第9页

- et al. Augmented Reality in Large Environments: Application to Aided Navigation in Urban Context[C]//IEEE International Symposium on Mixed and Augmented Reality. USA:IEEE, 2010:225–226
- [7] MILOSAVLJEVIC A, DIMITRIJEVIC A, RANCIC D. GIS-Augmented Video Surveillance[J]. International Journal of Geographical Information Science, 2010, 24(9): 1415–1433. DOI: 10.1080/13658811003792213
- [8] MILOSAVLJEVIC A, RANCIC D, DIMITRIJEVIC A, et al. Integration of GIS and Video Surveillance [J]. International Journal of Geographical Information Science, 2016, 350(10):198–9. DOI: 10.1080/13658816.2016.1161197
- [9] SNAVELY N, STEVEN M S, SZELISKI R et al. Photo Tourism: Exploring Photo Collections in 3D [J]. ACM transactions on graphics (TOG), 2006, 25(3):835–846
- [10] HAAN G D, SCHEUER J, VRIES R D, et al. Egocentric Navigation for Video Surveillance in 3D Virtual Environments[C]// IEEE Symposium on 3D User Interfaces, 2009. 3DUI 2009. USA, IEEE, 2009:103–110. DOI: 10.1109/3DUI.2009.4811214
- [11] NEUMANN U, YOU S, HU J, et al. Augmented Virtual Environments (ave): Dynamic Fusion of Imagery and 3d Models [C]//Proceedings of Virtual Reality. USA: IEEE, 2003: 61–67. DOI: 10.1109/3DUI.2009.4811214
- [12] SEBE I O, HU J, YOU S, et al. 3D Video Surveillance with Augmented Virtual Environments[C]//Proceedings of ACM SIGMM International Workshop on Video surveillance. USA:IEEE, 2003:107–112. DOI: 10.1109/3DUI.2009.4811214
- [13] DECAMP P, SHAW G, KUBAT R, et al. An Immersive System for Browsing and Visualizing Surveillance Video[C]// Proceedings of the International Conference on Multimedia. USA:ACM, 2010:371–380. DOI: 10.1145/1873951.1874002
- [14] KIM K, OH S, LEE J, et al. Augmenting Aerial Earth Maps with Dynamic Information [C]// Proceedings of the IEEE International Symposium on Mixed and Augmented Reality. IEEE:USA, 2009:35–38. DOI: 10.1109/ISMAR.2009.5336505
- [15] CHEN S C, LEE C Y, LIN C W, et al. 2D and 3D Visualization with Dual-Resolution for Surveillance[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. USA: IEEE, 2012:23–30
- [16] ZHOU Z, YOU J, YANG J, et al. Method for 3D Scene Structure Modeling and Camera Registration from Single Image: 20160249041[P].2016–08–25
- [17] SEGAL M, KOROBIKIN C, VAN W R, et al. Fast Shadows and Lighting Effects Using Texture Mapping[C]//Conference on Computer Graphics and Interactive Techniques. USA: ACM, 1992:249–252
- [18] Duke Multi-Target, Multi-Camera Tracking Project [EB/OL] [2017–10–23].http://vision.cs.duke.edu/DukeMTC/

## 作者简介



**潘志庚**, 杭州师范大学数字媒体与人机交互研究中心主任, 教授、博导, 广州玖的数码科技有限公司VR研究院院长; 长期从事虚拟现实等技术的研发; 承担或完成国家自然科学基金重点项目2项、国家科技支撑计划项目2项、“863”高科技计划1项等; 于2001年获中国高校科技进步二等奖, 2008年获国家科技进步二等奖; 发表论文80余篇。



**高翔**, 杭州师范大学数字媒体与人机交互研究中心硕士研究生; 从事数据可视化与增强现实等相关技术研究。



**丁丹丹**, 就职于杭州师范大学国际服务工程学院; 主要从事视频图像处理、视频编解码及优化技术、SoC系统芯片设计方法学、多媒体通信等研究; 在重要期刊与会议上发表相关论文20余篇, 申请发明专利6项。

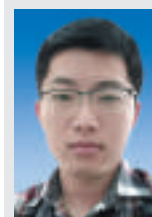
## 作者简介



**周忠**, 北京航空航天大学教授、博导, 现于虚拟现实/增强现实技术及应用国家重点实验室、北京航空航天大学虚拟现实技术与系统国家重点实验室工作; 研究方向为虚实融合等; 获国家技术发明二等奖1项、国家科技进步二等奖1项等; 发表论文70余篇, 获发明专利20余项, 出版专著1部。



**孟明**, 北京航空航天大学虚拟现实技术与系统国家重点实验室在读博士, 主要研究方向为虚实融合。



**周颀**, 北京航空航天大学虚拟现实技术与系统国家重点实验室科研人员; 主要研究方向为三维重建、路径规划和虚实融合技术; 近5年来, 在全球期刊及会议上发表论文4篇, 申请国家发明专利3项, 国际发明专利1项。

# 关于增强现实技术潜在发展方向的思考

## Thoughts on the Potential Development Direction of Augmented Reality Technology

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2017) 06-0037-04

**摘要:** 增强现实(AR)和虚拟现实(VR)技术近年来迅猛发展,引起了产业界和学术界的高度关注。AR/VR技术当前仍然是雷声大雨点小,并未广泛渗透到普通老百姓的日常生活中。针对现有AR/VR提供模式的局限性,提出了一种新的轻量级、普适化和跨平台的Web AR/VR技术,分析了其发展方向。同时,还探讨了Web AR/VR发展所面临的机遇和挑战。

**关键词:** Web AR; Web VR; 移动边缘计算; 5G

**Abstract:** The rapid development of augmented reality (AR) and virtual reality (VR) in recent years has attracted much attention from industry and academic. However, these two technologies have not been widely used in people's daily life as expected. Based on the existing AR/VR provisioning modes, a promising light-weight, pervasive and cross-platform Web AR/VR technology is proposed. Furthermore, the related development directions, challenges and opportunities of Web AR/VR are figured out.

**Key words:** Web AR; Web VR; mobile edge computing; 5G

乔秀全/QIAO Xiuquan  
任沛/REN Pei  
商彦磊/SHANG Yanlei

(北京邮电大学网络与交换技术国家重点实验室, 北京 100086)  
(State Key Laboratory of Networking and Switching Technology, Beijing University of Posts & Telecommunications, Beijing 100086, China)

- AR/VR具有极其广阔的发展与应用前景,将会成为5G的一种杀手级应用
- Web AR是普适化AR发展的转折点
- 在硬件设备、软件平台、内容制作工具、应用开发等多方面,AR/VR需要先形成产业联盟,推进协同合作

随着近两年支付宝“扫福集福卡”、基于位置的增强现实(AR)红包、QQ AR传递奥运火炬等活动的开展,AR与虚拟现实(VR)作为一种新型计算机视觉技术逐渐进入人们的视线,从学术界到产业界,AR/VR都成为了非常值得关注的热点。人们的生活由于AR/VR的出现而在慢慢发生着改变,并且这种改变还在不断地深入与扩大。我们可以预见的是,在未来的生活中,AR/VR将具有极其广阔的发展空间与应用前景,并且将会成为5G的一种杀手

级应用。

### 1 对AR/VR技术发展的思考

AR<sup>[1]</sup>是一种将真实世界信息和虚拟世界信息无缝集成的新技术,是把原本在现实世界的一定时间空间范围内很难体验到的实体信息,通过计算机视觉等科学技术,应用到真实世界,被人类感官所感知,从而达到超越现实的感官体验。通过AR技术,真实的环境和虚拟的物体实时叠加到了同一个画面或空间。

VR<sup>[2]</sup>则是综合利用计算机图形系统和各种现实及控制等接口设备,在计算机上生成的、可交互的三维环境中提供沉浸感觉的技术,通过提供使用者关于视觉、听觉、触觉等感官的

模拟,让使用者如同身临其境一般,可以及时、没有限制地观察三维空间内的事物。

#### 1.1 AR/VR具有广泛的应用场景

AR/VR作为一种新的视觉交互技术,将带动一大批新的产业发展模式。在教育领域,通过VR技术,可以实现沉浸式的教学方式,寓教于乐;通过AR技术,可以更加形象生动地展现书本中的内容,从而大大增加阅读的互动性与趣味性。在文化旅游及展览领域,通过VR技术可以打造数字化旅游景点,使用户足不出户就可以遍览天下美景;而AR技术在展览领域可以实现规划效果与真实场景的叠加,大大便利了市政建设规划任务。在娱乐领域,VR游戏、VR视

收稿日期: 2017-10-11

网络出版日期: 2017-11-07

基金项目: 国家自然科学基金重点国际合作项目(61720106007)、面上基金项目(61671081);北京市自然科学基金(4172042)

频、AR 游戏等也为人们提供了一种全新的娱乐休闲的方式。在工业制造领域,通过 AR 可以将多种辅助信息显示给用户;而 VR 则可以帮助企业更好地训练工人。

### 1.2 AR/VR 的历史尝试与发展

虽然 AR/VR 是在近两年内才得以井喷式的发展,但其起源甚至可以追溯到现代计算机技术的诞生之际。著名计算机科学家、图灵奖获得者伊万·萨瑟兰(Ivan Sutherland)在 1968 年发明的头戴式显示器能够将简单线框图转换为 3D 效果的图像,这种头戴式显示器(HMD)在结构上与现在的各种 AR/VR 产品也惊人的相似。1998 年,AR 技术第 1 次被成功应用于实时直播, Sportvision 作为当时体育转播图文包装和运动数据追踪领域的领先公司,开发了 1st & Ten 系统,从此开启了 AR 的应用与探索之路。

### 1.3 AR/VR 具有极强的发展潜力

AR/VR 技术离不开硬件设备的发展。VR 讲究的是沉浸感、交互性和构想性,构想性的关键在内容设计,而沉浸感和交互性的关键则在于硬件实现。AR 讲究的是现实世界叠加虚拟世界,从而实现“虚实结合”,显示技术与感知技术都是有待解决的重要问题。包括 Oculus、Google、SONY、HTC、Samsung、Microsoft、百度、阿里巴巴、腾讯、乐视、爱奇艺、优酷等在内的多家公司已经开始参与从事 VR/AR 的相关研发,仅在 2015 年内就涌现出了 100 多家 VR/AR 的硬件厂商,其中不乏各大手机终端厂商。尽管目前的 AR/VR 行业仍然处于起步阶段,但依据 Digi Capital 预测,至 2020 年,全球 AR 与 VR 市场规模将达到 1 500 亿美元;BI Intelligence 则统计,2020 年仅头戴式 VR 硬件设备市场规模将达到 28 亿美元。AR/VR 市场巨大的潜力吸引了越来越多的厂商、研究机构参与到了 AR/VR 软

件、硬件的研发浪潮之中。

### 1.4 目前提供 AR 的两种主要形式

(1) 配备专门的 AR/VR 设备(比如 AR/VR 头盔或者 AR/VR 眼镜):这是一种极其重量级的解决方案:一方面因为设备的费用相对较贵,如 Oculus Rift (VR 设备)的消费成本约为 1 500 美元,微软推出的全息眼镜 HoloLens、SONY 推出的 Morpheus 头盔、Samsung 推出的 Gear VR 头戴设备都由于其高额的成本,其推广范围大大受到了限制;而专用的 AR/VR 设备的另外一个缺陷在于其较差的便携性,用户对于 AR/VR 的使用范围受限,大大局限了 AR/VR 的应用范围与场景;因此,通过 AR/VR 近几年的发展历程,可以看出,基于专用设备的 AR/VR 应用推广遇到了很大的障碍,还无法大规模、普适化地应用起来,目前更多的是在非常垂直的领域和特定的场合提供此类 AR/VR 体验。事实上,给用户增添个额外的设备,是很有挑战性的事情,3D 电视的市场经验证明:哪怕只是多增加一副 3D 眼镜,对于内容消费来说都是巨大的障碍。因此,AR 设备的便捷化、轻量化,嵌入用户的日常生活用品是一个必然的趋势。

(2) 基于智能手机终端的 AR APP 应用:随着智能手机终端计算、存储等能力的不断提升,使得基于智能手机终端的 AR 应用成为可能。利用 APP 调用智能手机摄像头,从而对物理世界进行识别,并叠加虚拟信息,从而达到增强现实的目的。目前,基于 APP 的 AR 服务提供目前也有两种形式:

- 一种是专门的、独立的、离线的 APP 形式。所有 AR 的功能(包括开启摄像头、图像识别、目标跟踪和动画渲染等)都在手机终端来完成;在这种模式下,AR 服务中所涉及到的计算需求都要在本地进行,其特点在于实时跟踪快,用户体验好;但这种模式下,用户需要下载专门的

APP,应用推广的成本比较高;另外,单纯依靠智能终端进行计算,AR 服务的功能将大大受到受限(如识别图片数量、种类、精确度等能力)。

- 另一种则是采用“云端”的模式,如支付宝、聚划算、QQ AR 等服务,借助云端强大的计算能力,可以精确识别更多的图片,但与此同时,由于网络传输时延的增大,要做到实时的识别与跟踪难度很大,因而无法为用户提供优良的 AR 体验。

目前看来,基于专用 AR 设备的应用提供模式太过重量级,无法大规模、普适化推广与应用;而基于智能手机终端 APP 形式的 AR 应用,对于每个 APP 而言,又都是一个个独立封闭的孤岛,信息只局限于该 APP 的用户,难以跨平台提供普适化的 AR 服务,无法触及更多受众群体。另外一个致命问题在于 AR 应用的下载与安装。AR 服务的体验必须依托于特定的应用软件,从而导致了服务体验的门槛过高,并不方便用户的使用。总体来说,跨平台地提供普适化的 AR 应用是目前 AR 提供技术存在的最大的一个问题。因此,需要提供一种轻量级、跨平台的 AR 应用解决方案,使用户不需要下载专门的 APP 或者不局限于某个具体的 APP,就可获得 AR 体验,从而推动基于移动智能终端的 AR 大规模移动化、普适化应用。

## 2 Web AR 是普适化 AR 发展的转折点

目前,AR 技术出现了一种新的动向。这主要归功于伟大的 WWW 之父——蒂姆伯纳斯·李(2016 年图灵奖得主),他在 20 世纪 90 年代所发明的 Web 天然地具备了跨平台的特性。现有的很多 APP 是支持移动智能终端浏览器(HTML5)网页的,比如微信、QQ、微博、今日头条等,因此如果提供基于 Web 的 AR 应用,那么用户不用下载专门的 APP,只需要利用通用的 Web 浏览器或者这些内在支持 HTML5 网页的 APP 就可以体验 AR



服务。Web天然的普适特性将为AR/VR的应用提供一条更加便捷的传播方式。目前,越来越多的厂商开始关注原生支持AR/VR服务的智能终端硬件设备组件,这也给基于普适化的Mobile Web AR应用带来了很大的机遇。因此,基于Web的AR技术,特别是基于移动智能终端的Mobile Web AR技术,开始逐渐萌芽,为实现AR应用的大规模、跨平台的传播和分享奠定了基础,为解决AR的普适化应用难题提供了一种新的研究方向和技术可能。

然而,Mobile Web AR目前所面临的技术难题主要有以下3点。

(1)目前Mobile Web AR还处于技术的萌芽状态,现有的生态环境和技术条件、网络支撑等都还不是很成熟,无法完全发挥出它的全部潜力。比如:Web浏览器计算能力很弱,无法像现有的APP一样,在移动终端就完成全部AR的功能;参考现在基于独立APP的AR的实现方式,想要完全利用前端JavaScript技术来完成全部AR的工作,包括图像采集、特征点提取、图像匹配、动画渲染等,功能非常受限,目前只能识别简单的基于标识的图片(即基于Marker的AR技术),不足以支持自然图片和实物(即Markerless图像/实物)在前端的实时计算和识别;另外,由于JavaScript的弱计算能力,造成AR服务非常严重的卡顿现象,严重影响AR服务的用户体验。因此,只靠纯前端的方式,目前看来基本上还行不通,无法大规模商用。这也就意味着更多的计算功能需要放到云端服务器来完成,也就是所谓的“云+端”的实现模式。但这时候,网络的通信时延就成了很大的一个问题,现有的网络环境又很难满足AR技术在自动跟踪和识别方面的实时性和低延时的苛刻要求。

(2)浏览器的功能和性能差异比较大。现在手机终端上有多种不同的浏览器,如百度浏览器、360浏览器、QQ浏览器、Chrome浏览器、Safari

浏览器、火狐浏览器、UC浏览器等,除此之外,还有一些APP应用内嵌了简化版的浏览器功能模块(如微信、微博、QQ等),因此对于众多浏览器的适配工作也比较复杂。然而,与原生提供AR服务的APP相比,手机浏览器的功能要弱化很多,较大的3D动画模型和VR场景的加载和运行都会比较慢,直接影响了用户的业务体验;另外,Web网页不具备直接调用很多手机底层资源的权限,这也就给AR服务的提供带来了很大程度上的不便。

(3)内容制作和适配的问题也是Web AR普及所面临的一大难题。现有的大部分3D建模工具对基于Web的3D呈现兼容性较差,很多工具导出的模型无法直接加载使用。另外一个关键的问题就是用户与动画模型的交互技术现在也不成熟,使得用户在Web AR应用体验的过程中娱乐性和趣味性无法很好地得到体现。

### 3 Mobile Web AR:一种轻量级、普适化的AR技术

面对Web AR应用轻量化和普适化的拦路虎,我们团队经过无数次的技术攻关,在Mobile Web AR方面取得了技术性突破。首次成功研制了一套完整的、具有自主知识产权的、基于HTML5的、能够支持自然图片和实物识别功能的Web增强现实技术,超前性地采用了面向Web的“端计算+边缘计算+云计算”无缝协同的分布式AR实现方案,利用边缘服务器来提供即时的复杂计算能力,利用云服务器来保存实时性要求比较低的、与大量业务运营相关的数据和逻辑,解决当前Web增强现实面临的浏览器能力弱、图像计算量大、扩展性不好、实时性低的技术难题。基于Web AR技术,用户可以利用网页超链接机制将AR的功能嵌入到各种Web网页中,并且能够方便地进行转发、分享和传播,让用户不再受限于专门的APP就能获得AR体验,开启

了一个全新的、基于Web的AR应用新时代。

如图1所示,Mobile Web AR技术可以灵活地嵌入到微信公众号中,用户只需要关注特定的公众号,通过点选公众号中的菜单栏选项即可体验AR技术,程序将通过调取手机摄像头进行图像的采集、云端图像识别,以及在微信浏览器下进行3D模型的加载与渲染等一系列操作,方便快捷地为互联网用户带来轻量级、普适化的AR体验。

### 4 5G网络为AR/VR的大规模商用提供了机会

4G改变生活,5G<sup>[3-4]</sup>改变社会。5G网络将会带来更高的速率、更宽的带宽,能够充分满足消费者高带宽、低延时等更高的业务体验需求。同时,5G在网络的整体设计中也会采用一些全新的设计思路,比如软件定义网络(SDN)/网络功能虚拟化(NFV)/网络切片技术,从而可以根据业务需求来灵活动态地调整网络资源,进而满足不同业务类型的需求;设备到设备(D2D)的自组织通信使得终端与终端之间就可以直接进行数据传输和内容分享;移动边缘计算(MEC)<sup>[5]</sup>技术使得计算和存储等功能进一步下沉到网络的边缘,在更靠近用户的地方运行,进一步满足了低延时类的应用需求,避免了核心网络的传输时延,对时延敏感的AR、自动驾驶、工业物联网等应用将会在5G网络中得到更好的支持。

#### 4.1 5G网络下AR/VR的机遇与挑战

5G到来之后,很多新型应用将会出现并大规模普及。AR/VR具有高带宽、低延时、计算量大的特点;高清的VR视频每秒容量都高达百兆,现有4G网络下,移动用户很难进行流畅的体验,只能对视频进行压缩,或者采用将球形全景图转换成立方体切片逐步加载的方式来进行处理;而AR对实时性和计算的要求更高,需



▲图1 基于微信社交平台的Mobile Web AR实例展示

要对目标进行实时识别和自动跟踪,若所有计算都从云端服务器进行处理,则网络时延将会是一个很大的挑战。然而,5G网络具有的新特点,能够很好地满足消费者对超高清视频、AR/VR、自动驾驶等高带宽、低延时的业务体验需求,可以说AR/VR将是5G的一种杀手级应用。

#### 4.2 5G网络下Mobile Web AR的机遇与挑战

相比较于AR/VR专用设备或者基于APP的AR/VR应用,Mobile Web AR理论上更加依赖于通信网络的支持。由于Web前端计算能力弱,很多的工作需要在云端服务器来进行。在现有网络环境下,可以在各地计算中心通过部署边缘服务器来进行处理,但边缘服务器的部署和管理都将浪费巨大的人力和物力。5G时代,有了统一的移动边缘服务,应用提供商就能够很方便地进行全网的按需灵活部署和动态迁移,解决现在Mobile Web AR所面临的时延和带宽问题。

Mobile Web AR技术能真正提供一种移动化、跨平台、普适化的AR服务,可以应用到各行各业,包括:城市规划宣传、文化旅游行业、博物馆、房地产行业、大型企业、连锁酒店行业、连锁餐饮行业、大型商场超市、商贸城、大型展会、广告营销行业等。AR/

VR将不仅是一次视觉技术的革命,更是一次生活方式的巨大转变。

目前,AR/VR还在进一步发展中,生态链尚不成熟。因此,AR/VR相关产业链(硬件设备、软件平台、内容制作工具、应用开发商等多方面)需要先形成产业联盟,共同推进协同合作。在此基础上,还应制订相关标准,通过标准来进行互联互通,让产业链上的各个环节无缝协作,这样才能开展大规模的AR/VR产业化应用,形成一个良好的产业生态链。

## 5 结束语

5G的低延迟、高吞吐量、D2D、多输入多输出(MIMO),以及边缘计算等新的技术,将推动AR/VR的快速发展,而Web AR/VR轻量级和良好的跨平台性将极大地提升用户的使用体验,从而产生更多更丰富的AR/VR应用,并带来更多的业务和技术创新。目前Web AR的研究仍然处在初级阶段,存在诸多待解决的问题,这需要学术界和产业界共同努力,在算法、协议、标准、Web 3D内容、产业链等方面完善和创新。

## 致谢

Mobile Web AR的成功实现与应用,离不开团队老师和同学们的辛苦努力。感谢研究生张鹏、闫浩然、郭佳豪、徐言、宋丽敏以及任沛博士、南

国顺博士等同学的辛勤付出,对他们谨致谢意!

## 参考文献

- [1] KARHU A, HEIKKINEN A, KOSKELA T. Towards Augmented Reality Applications in A Mobile Web Context[C]// International Conference on Next Generation Mobile Apps, Services and Technologies. USA: IEEE, 2014:1-6. DOI: 10.1109/NGMAST.2014.36
- [2] VINCE J. Virtual Reality Systems[M]. New York: ACM Press/Addison-Wesley Publishing Co., 1995
- [3] AGIHAL M, ROY A, SAXENA N. Next Generation 5G Wireless Networks: A Comprehensive Survey[J]. IEEE Communications Surveys & Tutorials, 2016, 18(3): 1617-1655. DOI: 10.1109/COMST.2016.2532458
- [4] FOUKAS X, PATOUNAS G, ELMOKASHFI A, et al. Network Slicing in 5G: Survey and Challenges[J]. IEEE Communications Magazine, 2017, 55(5): 94-100. DOI: 10.1109/MCOM.2017.1600951
- [5] TRAN T X, HAJISAMI A, PANDEY P, et al. Collaborative Mobile Edge Computing in 5G Networks: New Paradigms, Scenarios, and Challenges[J]. IEEE Communications Magazine, 2017, 55(4): 54-61. DOI: 10.1109/MCOM.2017.1600863

## 作者简介



**乔秀全**, 北京邮电大学教授、博士生导师、网络与交换技术国家重点实验室网络服务基础研究中心副主任; 主要研究方向为Mobile Web AR/VR应用、未来网络服务提供、物联网服务提供、移动互联网应用创新相关的应用基础理论和技术研究; 曾获“北京市科技新星”, 作为负责人承担了国家“973”课题、国家自然科学基金面上项目、国家自然科学基金青年基金项目、移动通信网国家科技重大专项子课题等国家级项目; 近5年来, 以第一作者身份在《IEEE Communications Magazine》等权威SCI期刊上已经发表论文60余篇。



**任沛**, 北京邮电大学在读博士生; 主要研究方向为Mobile Web AR/VR应用、移动边缘计算。



**商彦磊**, 北京邮电大学副教授、硕士生导师; 主要研究方向为云计算、Web AR/VR应用; 承担过移动通信网国家科技重大专项子课题等国家级项目; 已发表SCI/EI检索论文30余篇。



# 封装天线技术发展历程回顾

## A Review of the Development of Antenna-in-Package Technology

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2017) 06-0041-09

**摘要:** 封装天线(AiP)是基于封装材料与工艺,将天线与芯片集成在封装内,实现系统级无线功能的一门技术。AiP技术顺应了硅基半导体工艺集成度提高的潮流,为系统级无线芯片提供了良好的天线解决方案,因而深受广大芯片及封装制造商的青睐。AiP技术很好地兼顾了天线性能、成本及体积,代表着近年来天线技术的重要成就。另外,AiP技术将天线触角伸向集成电路(IC)、封装、材料与工艺等领域,倡导多学科协同设计与系统级优化。AiP技术已逐渐趋于成熟,在技术方面有很多论文和专利可供参考,但还没有一篇专门回顾AiP技术发展历程及其背后的故事,文章旨在填补这一方面的空白。将以AiP技术发展历程中起到重要推动作用的经典设计为例,加上自己亲身经历的故事,为大家勾勒出AiP技术发展的来龙去脉。

**关键词:** 封装天线;射频芯片(RFIC);移动通信;物联网

**Abstract:** Antenna-in-Package (AiP) technology is an antenna solution technology that implements an antenna or antennas on (or in) an integrated circuit (IC) package that can carry a highly-integrated radio or radar transceiver die (or dies). Keeping with the trend of silicon semiconductor technologies, AiP technology provides elegant antenna solutions to radio-frequency system-on-chip. Hence, AiP technology has been widely adopted by chip makers for single-chip radios and radars. Having well balanced among performance, size, and cost, AiP technology represents an important antenna technology achievement in recent years. In addition, AiP knowledge is permeating from the antenna to packaging, IC, material, microwave, and wireless communication fields, AiP technology is getting matured. There are literally many papers and patents devoted to AiP technology. However, there is no paper about its development history. This paper aims to fill this void by reviewing those important design examples that have driven the development of AiP technology as we know today.

**Key words:** AiP; radio frequency IC (RFIC); mobile communications; Internet of things

无线通信发展迅速,4G的商用才刚刚铺开,5G研发的热潮已迎面扑来。在未来的几年里,5G旨在实现低时延、高速率、大容量万物互联,将会彻底改变我们同世界互动的方式。为了使5G的愿景变为现实,必须突破几个关键技术藩篱,其中一个核心技术的难题就与我们的领域

收稿日期: 2017-10-25  
网络出版日期: 2017-11-13

息息相关,即如何利用大规模多输入多输出(MIMO)天线阵列实现波束成形、扫描、追踪、锁定来有效对抗毫米波移动信道的路径损耗<sup>[1]</sup>。

汽车雷达在改善驾车安全的同时,也提升了全新的驾车体验。目前,汽车雷达主要工作在24 GHz和77 GHz的窄频带范围,仅起到预警及辅助驾驶的作用。未来汽车雷达将朝着工作在79 GHz宽频带发展,利

张跃平/ZHANG Yueping

(南洋理工大学,新加坡 639798)  
(Nanyang Technological University,  
Singapore 639798, Singapore)

用4 GHz带宽获得更高的空间分辨率,甚至实现无人驾驶<sup>[2]</sup>。2015年,谷歌手势雷达一经问世,便立刻造成全球轰动。手势雷达工作在60 GHz频带,跟踪人手移动及其变化,非常适合嵌入在可穿戴设备、手机和其他电子产品中作为用户界面。

消费类电子产品的硬件主要是通过系统级芯片(SoC)和系统级封装(SiP)技术来实现。SoC技术通过半导体工艺在同一个芯片上集成实现系统功能的各种电路。而SiP技术则是通过封装工艺将各个功能模块集成在一个封装内<sup>[3]</sup>。尽管SoC技术可以以更低系统成本来提高系统的可靠性和功能,但是由于使用相同的材料和工艺,没办法使每个类型的电路性能达到最优,进而导致系统性能降低和系统功耗增加等问题。相反,SiP技术可以提升系统性能,降低系统功耗,但是由于功能模块和封装制作采用不同的材料和工艺,会导致系统的可靠性降低和系统成本增加。

天线是无线系统中的重要部件,有分离和集成两种形式。分离天线司空见惯<sup>[4]</sup>,集成天线也已悄悄地进入到我们的视线。集成天线包括片上天线(AoC)和封装天线(AiP)两大类<sup>[5]</sup>。AoC技术通过半导体材料与



工艺将天线与其他电路集成在同一个芯片上<sup>[6-10]</sup>。考虑到成本和性能,AoC技术更适用于太赫兹频段<sup>[11-13]</sup>。AiP技术是通过封装材料与工艺将天线集成在携带芯片的封装内。AiP技术很好地兼顾了天线性能、成本及体积,代表着近年来天线技术重大成就,因而深受广大芯片及封装制造商的青睐。如今几乎所有的60 GHz无线通信和手势雷达芯片都采用了AiP技术<sup>[14-24]</sup>。除此之外,在79 GHz汽车雷达<sup>[2]</sup>,94 GHz相控阵天线,122 GHz、145 GHz和160 GHz的传感器,以及300 GHz无线链接芯片中都可以找到AiP技术的身影<sup>[25-30]</sup>。毋庸置疑,AiP技术也将会为5G毫米波移动通信系统提供很好的天线解决方案。

很显然AoC和AiP分别属于上述SoC和SiP概念的范畴,那么我们为什么要将它们从SoC和SiP技术中明确区分开来呢?原因其实很简单,就是为了强调它们独有的辐射特性。关于AoC技术,需另辟专文详述,本文仅拟论及AiP技术。尽管AiP技术方面的论文和专利很多,但还没有一篇用于回顾AiP技术发展历程及其背后故事,本文旨在填补这一空白。

## 1 早期与蓝牙无线技术一起发芽

AiP技术早在该术语被提出和普及之前就已经存在。AiP技术继承与发扬了微带天线、多芯片电路模块及瓦片式相控阵结构的集成概念。它的发展主要得益于市场的巨大需求,硅基半导体工艺集成度的提高,驱动了研究者自90年代末不断探索在芯片封装上集成单个或多个天线。

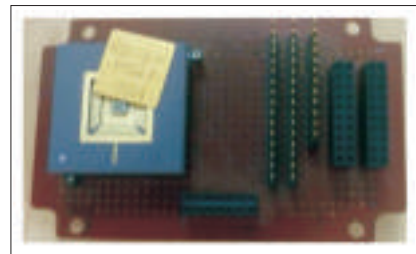
我于1993年初荣幸地成为香港中文大学国际知名卫星天线专家黄振峰博士课题组成员,有机会参与制造和测试多款微带天线。通过使用一种刚问世不久的低损耗高介电常数陶瓷材料,我们成功地将900 MHz微带天线小型化到只有手指甲大小,利用几个小型化天线就可以实现手

机天线辐射方向图成形,减少向人体侧辐射。研究成果不知怎样引起了时任香港中文大学校长高锟教授的关注,有一天召集我们到他办公室向他汇报。然而,我们关于实现天线小型化的研究似乎没有给高锟教授留下深刻印象。他打比喻说:将4条腿的长凳缩小到3条腿的板凳只是进化而已,大学应该尝试做一些革命性的研究。高锟教授获2009年度诺贝尔物理学奖,也许高锟教授这样伟大的科学家更关注研究的科学价值,而我们则更强调潜在的应用。幸运的是,我们关于天线小型化的工作在无线领域受到欢迎,并在1995年电气与电子工程师协会(IEEE)天线与传播国际研讨会上与摩托罗拉公司设计的类似天线在同一会场宣读<sup>[31]</sup>,直接促进了陶瓷贴片天线的发展。1996年,我加入了香港城市大学国际著名的天线实验室从事介质谐振器天线研究。偶然的机会我在香港城市大学遇到了材料科学家李国源博士,他热情地向介绍了他研究的低温共烧陶瓷(LTCC)材料与工艺,并用一块可以表贴集成电路内有埋置去耦电容的LTCC基板讲解了厚膜电路的优缺点,临别时还慷慨地向我赠送了好多块他烧好的LTCC基板用于天线研究。这些LTCC基板除了后来用于天线试验毁坏的以外,剩余的我还保留着。李国源博士现在是华南理工大学教授。1998年,我离开任教的香港大学前往新加坡南洋理工大学就职。令我惊讶的是,我被分派到电路与系统系而非通信工程系,后者有几位教授及先进的实验室从事天线与电磁波传播研究。在参观集成电路实验室时,我看到了图1所示的装置,就问实验室一个研究生那是微带天线吗?研究生回答:“不,那不是,那是一个集成电路芯片。”不久,电路与系统系启动了“片上软件无线电”的战略性研究项目,我的任务是为这个项目开发天线技术。因为对图1所示的集成电路芯片同微

带天线结构相似性的着迷,预测到未来有可能产生一种革命性的天线解决方案,我很快决定研究图1所示集成电路芯片作为天线的可行性<sup>[32]</sup>。

首先,我找来许多现成的陶瓷封装集成电路芯片来进行天线及电路实验,研究它们之间的相互影响。图2a)就是当时实验过的一个在双列直插式封装上实现的2.4 GHz天线。后来发现利用现成的陶瓷封装集成电路芯片来进行天线实验有很大的局限性,于是决定利用印刷电路板(PCB)工艺加工集成电路封装结构模型且印制有天线。图2b)所示的集成电路封装结构模型利用了3层电路板,天线印制在顶层板上,信号线及封装地在低层板上实现,中间层中空,夹在顶、低层之间形成腔体来携带裸芯片。顶层板印制了5.2 GHz微带天线,如果顶层板换成左下角所示的板,则模拟集成电路封装结构是一款集成有2.4 GHz及5.2 GHz双频微带天线。上述在现成的陶瓷封装集成电路芯片和PCB加工的模型上尝试,都获得了令人满意的实验结果。受其鼓舞,我和学生林伟、薛阳、王珺珺2003年利用LTCC工艺实现了多款真正工业意义上的封装天线<sup>[33]</sup>。图2c)是一款利用LTCC工艺为蓝牙芯片开发的差分封装天线。

与此同时,英国伯明翰大学C. T. SONG, P. S. HALL和H. GHAFOURI-SHIRAZ提出了两个有关天线封装的概念。第1个概念突出体现了将小天线埋入到芯片封装材料中,然后在埋入式天线近距离处放置一个寄生单元,从而改善封装天线的低增益并



▲图1 拥有密封金属环及盖的陶瓷封装集成电路芯片



▲图2 封装天线技术的进化

增加带宽。第2个概念建议在半导体芯片上实现射频前端电路及电小馈电天线,并在馈电天线上方增加寄生单元并充当封装顶盖,密封整个芯片<sup>[34]</sup>。HALL教授学识渊博、谦虚低调,是国际天线界一位德高望重的学者。为表彰他在微带天线方面所做出的杰出贡献,美国IEEE天线与传播学会授予他2012年the John Kraus Antenna Award,英国工程技术学会(IET)授予他2013年度the James R. James Lifetime Achievement Award。

几乎同时,封装工程师也在尝试解决相同的问题。D. J. MATHEWS等人<sup>[35]</sup>申报了一项内置电磁防护罩和天线的用于蓝牙芯片封装的发明专利。美国佐治亚理工学院K. T. LIM等人设法在封装系统(SoP)上集成射频无源器件、天线和有源芯片,以增强封装系统的整体性能和增加更多功能<sup>[36]</sup>。比利时校际微电子中心S. BREBELS等人也实现了集成有天线的SoP<sup>[37]</sup>。但是,由于已经有SiP的概念,SoP的概念未被广泛接受。

稍后香港城市大学梁国华教授及新加坡微电子研究所A. P. POPOV博士分别独立发明基于介质谐振器天线的AiP技术<sup>[38-39]</sup>。梁国华教授90年代初同我相识在香港中文大学微波实验室。当时他博士即将毕业。据说梁国华博士论文答辩时,答辩委员会主席认为他提交给中文大学的博士论文等于其他学校的两份博士论文。梁国华教授后来被任命为IEEE天线与传播汇刊的主编。梁国华教授的博士导师实际上是陆贵文

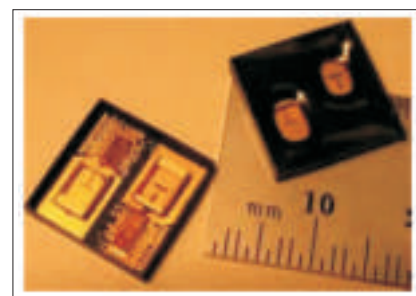
教授,陆贵文教授获得了2017年度IEEE天线与传播学会the John Kraus Antenna Award,另外曾长期在香港中文大学及城市大学工作过的李敏方教授也于2009年获此殊荣,微带天线的研究至少已产生了3位获奖者。

## 2 中期与60 GHz无线技术及毫米波雷达一起成长

2005年3月初,在新加坡举办的第1届小型天线国际研讨会上,我第一次见到了来自IBM Thomas J. Watson Research Center的Brian GAUCHER先生和刘兑现博士,并邀请他们访问了南洋理工大学。Brian就IBM的60 GHz SiGe芯片、天线、封装和测试设备做了学术报告。图3为IBM用于概念验证的60 GHz芯片照片。SiGe裸芯片通过倒装焊技术与天线连接、封装成为栅格阵列模块。由于需要在封装内加金属墙及封装上开天线窗口,因此该概念封装天线不易大批量生产。我向Brian GAUCHER先生和刘兑现博士简要介绍了几款基于LTCC工艺适合批量生产的2.4 GHz和5.2 GHz频段的封装天线。双方当场就达成了基于LTCC工艺合作开发用于IBM 60 GHz SiGe芯片组的封装天线的可行性研究计划。我和我的学生孙梅博士负责设计工作,邀请新加坡制造技术研究所(SIMTech)的一个研究小组负责LTCC加工,刘兑现博士负责评估并向我反馈测试结果。

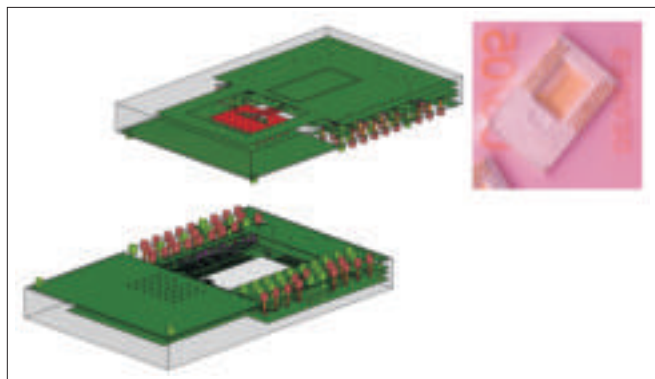
早期在封装上集成天线,所用英文名称五花八门。随着开发的深入

我意识到一个专门响亮的名称非常有利于去推广它。2006年起,我首先使用Antenna-in-Package的名称去推广这一新颖的天线解决方案<sup>[40]</sup>。采用Antenna-in-Package(AiP)而不是Antenna-on-Package(AoP)主要考量是前者更有可能使天线靠近芯片,减少互连损耗<sup>[40]</sup>。低插损的天线与芯片互连是毫米波AiP技术的一大挑战。2006年3月初,我参加了在美国纽约举办的第二届小型天线国际研讨会,并访问了IBM Thomas J. Watson Research Center,与刘兑现、U.R. PFEIFFER和Janusz GRZYB博士讨论了AiP技术问题。很遗憾,这次访问并没有见到已于2004年离开IBM的Thomas ZWICK博士。Thomas在开发探针式毫米波集成天线测试系统及AiP键合线互连方面做出了突出贡献。此次会面增强了双方合作,加速了AiP技术的发展。图4展示了设计阶段获取的截图和刘兑现博士评测的基于LTCC加工的AiP样品。这一样品集成了共面波导馈线、准腔体、定向保护环、基板材料调制的槽天线。天线输入阻抗故意设计成容性的与芯片通过感性的键合线互连,结果令人满意,并在2007年3月英国剑桥举办的第3届小型天线国际研讨会上被授予最佳论文奖<sup>[41]</sup>。事实上,在赴英参会之前我在一次内部会议上就对孙梅博士及新加坡制造技术研究所的合作者预测到60 GHz AiP技术论文会获奖。而且,还有一件有趣的事情,那就是在剑桥大学的演讲厅等待颁奖时,我在一张会议用纸上



▲图3 IBM公司用于概念验证的60 GHz SiGe芯片、天线、封装模块照片





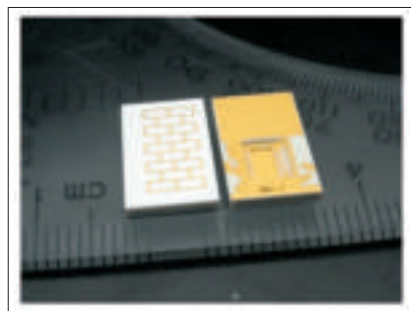
▲图4 LTCC 60 GHz AiP设计时截图及实物照片

写了个便条,再一次预测有关AiP的工作将会赢得更高奖项。我也请刘兑现博士在便条上签了字,便条至今由我保管。果然不出所料,2012年我、孙梅、刘兑现和陆亿浈博士荣获当年IEEE天线与传播学会谢昆诺夫论文奖<sup>[42]</sup>。这是该奖项自1957年设立以来,亚洲研究者首次及至今唯一获此殊荣。谢昆诺夫是国际著名的电磁理论学家。他于1920年代初期从前苏联经中国移居美国。他在工程电磁场、天线理论、波导理论、电磁屏蔽等方面提出了许多定理、原理、概念、方法,做出了重要的贡献。他使应用数学焕发出光彩,许多工作带有奠基性质。就经典电动力学方法(即量子理论以外领域)而言,中国著名物理学家黄志询先生认为可以把他比作20世纪的麦克斯韦<sup>[43]</sup>。

John KRAUS是对天线做出卓越贡献的老一辈天线专家。IEEE天线与传播学会的the John Kraus Antenna Award就是用他名字命名的。KRAUS发明的螺旋天线应用非常广泛,但他发明的栅格天线却鲜有应用<sup>[44]</sup>。孙梅博士在2008年发现栅格天线的网状结构非常适合于LTCC工艺,故将其应用于60 GHz AiP设计中<sup>[45]</sup>。紧接着,Wolfgang MENZEL等人将其应用于79 GHz<sup>[2]</sup>,陈梓浩等人将其应用于94 GHz<sup>[46]</sup>,THOMAS等人将其应用于122 GHz<sup>[47]</sup>,张冰等人将其应用在了145 GHz AiP设计中<sup>[48]</sup>。几乎快被人遗忘的栅格天线得以在AiP技术中发扬光大。MENZEL教授是微带漏波

天线的发明者,曾访问南洋理工大学并做了有关车载雷达中毫米波天线的邀请报告,期间与我们交流了关于栅格天线的设计方法。图5是孙梅等人利用LTCC工艺为IBM 60 GHz SiGe接收机裸芯片设计的AiP。它采用了键合线球栅阵列(BGA)封装结构集成了14个网格的栅格天线,尺寸为 $13.5 \times 8 \times 1.265 \text{ mm}^3$ 。刘兑现博士测试了AiP的天线性能并给我发送了电子邮件,邮件中只写了两个字“excellent results”。确实,结果表明栅格天线具有频带宽、辐射效率高的优点,且在57 ~ 64 GHz频率范围内主波束辐射都在天顶方向,60 GHz频段最大增益可达到14.5 dBi<sup>[49]</sup>,代表了当时最好的60 GHz AiP设计。

AiP技术的成功主要归功于人们重拾了对60 GHz无线系统的研究与开发兴趣。2007年,AiP技术发展进入新阶段,IEEE开始着手制订60 GHz频带标准,很多企业开始重视60GHz芯片及封装天线的研发。美国Sibeam公司第一个将60 GHz天线阵与互补金属氧化物半导体(CMOS)裸

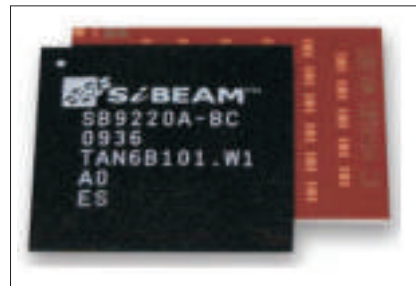


▲图5 LTCC 60 GHz AiP实物照片

芯片利用LTCC工艺集成在一起,引入消费市场,用于高清视频内容的无线传输。图6是Sibeam公司60 GHz芯片的照片,集成的微带天线阵清晰可见。值得一提的是,我在2005年9月美国加州圣克拉市举办的天线系统和短程无线会议上做主题演讲后,休息之余,与Sibeam的创始人之一、毫米波CMOS电路先驱者C. H. DOAN先生热情地讨论了一些有关基于LTCC封装天线集成的问题<sup>[50]</sup>。

2010年,美国IBM公司公布了用于60 GHz相控阵系统的完整AiP方案<sup>[51]</sup>。如图7所示,基于LTCC工艺,16个矩形微带天线被集成在BGA封装中,发射或接收裸芯片通过倒装焊技术与AiP相连。AiP尺寸为 $28 \times 28 \times 1.47 \text{ mm}^3$ ,在4个IEEE 802.15.3c通道中,天线单元增益均可达到5 dBi。2011年,IBM还展示了另外一个用于60 GHz相控阵系统的基于有机材料高密度互连工艺(HDI)的完整AiP方案<sup>[52]</sup>。值得一提的是,IBM与封装材料及工艺商通过努力实现了在AiP中嵌入空气腔体来改善微带天线阻抗及辐射特性。

2011年,韩国Samsung公司发表了用于60 GHz相控阵系统的完整AiP方案<sup>[53]</sup>。如图8所示,基于LTCC工艺,24个圆形微带天线被集成在BGA封装中,发射或接收裸芯片可以通过倒装焊技术与AiP相连。为了避免像IBM公司那样在AiP中嵌入空气腔体来改善微带天线阻抗及辐射特性可能带来的可靠性问题,Samsung公司AiP设计采用了圆形叠



▲图6 Sibeam公司60 GHz LTCC封装天线CMOS芯片实物照片





▲图7 IBM公司 60 GHz LTCC 封装天线 SiGe 芯片实物照片

层微带天线。AiP 的尺寸为  $20 \times 15 \times 1.02 \text{ mm}^3$ , 分别实现了 9 GHz 带宽及 14.5 dBi 增益。Samsung 公司还分别在 2012 和 2013 年提出了用于 60 GHz 相控阵系统的基于低成本 FR4 材料与 HDI 工艺的完整 AiP 方案<sup>[19,21]</sup>。叠层微带天线有助于满足 HDI 工艺对金属密度的要求。Samsung 公司 AiP 技术主要贡献者是一位名叫 Wonbin HONG 的年轻学者, 我们经常通过电子邮件及在国际学术会议上交流 AiP 技术方面的心得。后来 HONG 博士率先报道了 28 GHz 5G 手机天线方面的工作, 引起天线界的关注。

2012 年, 美国英特尔 (Intel) 公司发表了用于 60 GHz 相控阵系统的完整 AiP 方案<sup>[53]</sup>。如图 9 所示, 基于 LTCC 工艺, 36 个矩形微带天线 (含 4 个哑元) 被集成在 BGA 封装中, 收发裸芯片通过倒装焊技术与 AiP 相连。AiP 的尺寸为  $25 \times 25 \times 1.4 \text{ mm}^3$ , 在 60 GHz 频段,  $\pm 30^\circ$  扫描范围内增益达 19 dBi。针对 60 GHz 相控阵系统, 英特尔还分别在 2013 年、2014 年和 2015 年提出利用 PCB<sup>[57]</sup>、玻璃<sup>[55]</sup> 和液晶聚合物 (LCP)<sup>[56-57]</sup> 实现低成本低损耗 AiP 解决方案。

2015 年, 美国谷歌 (Google) 公司首次公开亮相的手势雷达名震四海。手势雷达使用 60 GHz 信号来快速追踪人手移动, 精度可以达到亚毫米级。也许对我而言, 最振奋人心的就是 AiP 技术被应用于手势雷达芯片, 如图 10 所示, 德国英飞凌 (Infineon) 公司利用嵌入式晶圆级封装 (eWLB) 技术, 在 AiP 中集成了 1 个

60 GHz SiGe 收发裸芯片、2 个用于发射的差分微带天线和 4 个用于接收的单端口微带天线<sup>[24]</sup>。AiP 的尺寸为  $14 \times 14 \times 0.8 \text{ mm}^3$ , 显而易见, 其尺寸已经足够小, 可用于穿戴设备。并且对于智能手表、手机和其他装置而言, 手势雷达作为用户界面潜力巨大。

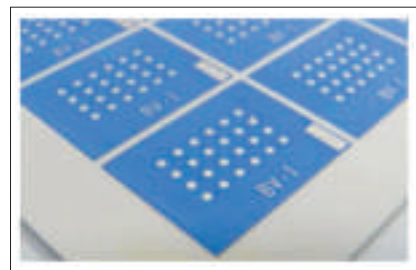
几乎所有主要的日本电子公司都开发出了适用于 60 GHz 应用的芯片组和 AiP 方案。图 11 所示的是日本 NEC 公司早期开发的 60 GHz 接收机模块, 以及 NTT 公司近期开发的 60 GHz 收发模块。两家公司分别用不同的 LTCC 工艺在模块中集成了缝隙天线及抛物面天线<sup>[58-59]</sup>。

像 AiP 技术用于谷歌手势雷达中一样, 英飞凌公司也为 77 GHz 车载雷达研制了 SiGe 芯片组及基于 eWLB 工艺的 AiP 技术, 并自 2016 年 6 月以来就同比利时校际微电子中心合作开发基于 28 nm CMOS 的芯片组和基于低成本低损耗 PCB 工艺的 AiP 技术, 用于 79 GHz 车载雷达<sup>[60]</sup>。比利时校际微电子中心负责 AiP 技术开发的是 Guy A. E. VANDENBOSCH 教授。VANDENBOSCH 教授每次来中国讲学, 都会在演讲前向学生们赠送著

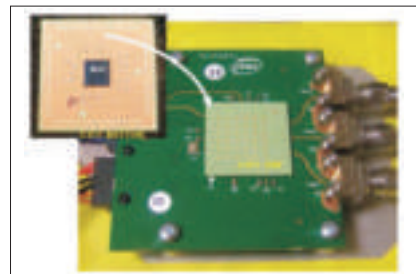
名的比利时巧克力, 很受学生们欢迎。IBM 公司将其 AiP 技术的工作频段推进到 94 GHz, 并在 2014 年实现了用于 W 波段的可扩展相控阵系统的 SiGe 芯片及完整 AiP 解决方案<sup>[61]</sup>。如图 12 所示, 该 AiP 设计采用多层有机基板及 HDI 集成了 64 个双极化叠层微带天线和 36 个哑元, 其尺寸为  $16.2 \times 16.2 \times 0.75 \text{ mm}^3$ 。

在欧盟科技委员会的赞助下, 2009 年 11 月—2013 年 5 月, SUCCESS 合作团体基于 SiGe 工艺开发了如图 13 所示的 122 GHz 及 145 GHz 雷达芯片, 且用键合线将它们分别与天线阵列集成在 8 mm 见方的扁平无引脚 (QFN) 封装内。

图 14 是奥德利 JKU 在欧盟科技

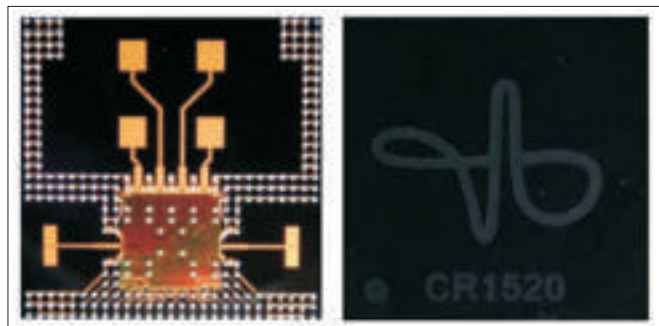


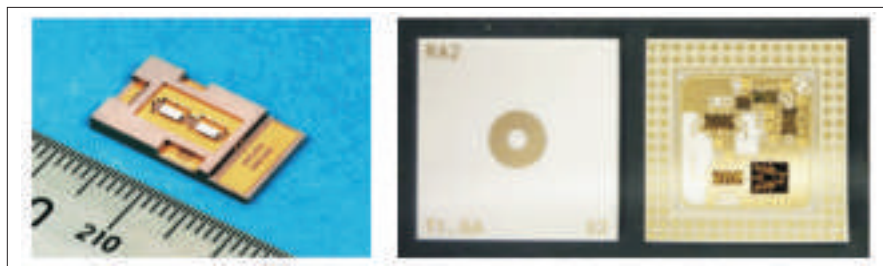
▲图8 Samsung公司集成了24个天线的AiP实物照片



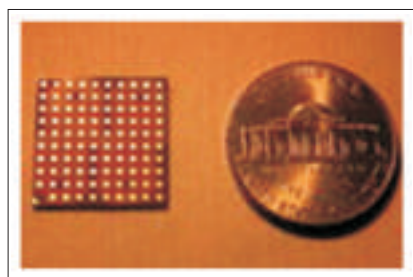
▲图9 Intel公司 60 GHz LTCC 封装天线 CMOS 芯片实物照片

图 10 ▶  
Google 公司 60 GHz  
eWLB 封装天线 SiGe  
芯片实物照片





▲图 11 NEC 公司和 NTT 公司 60 GHz LTCC 封装天线 GaAs 芯片实物照片



▲图 12 IBM 公司 94 GHz HDI 封装天线 SiGe 芯片实物照片

委员会、英飞凌等公司赞助下于 2013 年 10 月推出的基于 SiGe 工艺开发的 160 GHz 雷达芯片, 基于 eWLB 工艺将芯片与天线阵集成在 BGA 封装内<sup>[62]</sup>。

### 3 近期助力太赫兹、物联网和 5G 移动通信发展

太赫兹技术是改变未来世界的重要技术, 已引起各国政府的重视。在日本政府的支持下, NTT、NICT 和 FUJITSU 都参与到世界上第一个使用 300 GHz 无线链接的收发信机研发工作中。NTT 成功研发了如图 15 所示用于 300 GHz 发射机芯片的 AiP 结构。该 AiP 设计采用 LTCC 工艺, 其中喇叭天线尺寸为  $5 \times 5 \times 2.7 \text{ mm}^3$ , 最大增益为 18 dBi, 带宽达 100 GHz<sup>[63-65]</sup>。

物联网 (IoT) 是互联网发展的新阶段, 它通过智能感知、识别技术与普适计算等手段实现万物互联。最近, 美国 Silicon Labs 公司发布了如图 16 所示的世界上最小的蓝牙无线系统, 它的封装内集成有天线, 尺寸只有  $6.5 \times 6.5 \times 1.5 \text{ mm}^3$ , 这使得设计真正紧凑的物联

网设备变得可行。

AiP 技术是近期国际上 5G 移动通信研发的一个重要课题, 难点是如何实现高辐射效率及低成本量产。图 17 为 IBM 应用于未来 5G 基站 28 GHz AiP 照片<sup>[66]</sup>。该 AiP 包含 4 个单片 SiGe 裸芯片和 64 个双极化天线, 尺寸约为  $7.1 \times 7.1 \text{ cm}^2$ 。刘兑现博士是 IBM 公司所有 AiP 设计背后的灵魂人物, 他指出相控阵列天线的并行双极化运作方式能够形成两个波束支持低于  $1.4^\circ$  的波束扫描精度, 同时保持发送和接收模式, 进而使服务的用户量增加一倍。

图 18 为 Qualcomm 近日发布的用于 5G NR 首款智能手机参考设计中采用的 28 GHz 毫米波芯片<sup>[67]</sup>。参考设计旨在于手机的功耗和尺寸要求下, 对 5G 技术进行测试和优化。该芯片天线方案采用 AiP 技术, 尺寸约为 5 美分大小。Qualcomm 希望能在一年内将尺寸缩小一倍。

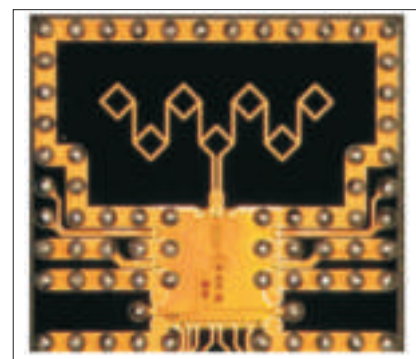
### 4 结束语

不知不觉 AiP 技术已走过了多年发展历程。早期 AiP 技术的研究主

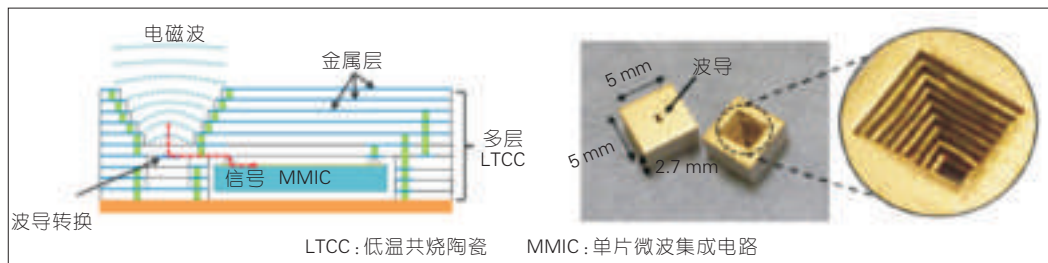
要集中在了大学的实验室, 围绕着 2.4 GHz 蓝牙芯片展开。如何实现天线小型化是当时 AiP 研究者所面临的技术难题。中期 AiP 技术的开发主要集中在大公司, 围绕着 60 GHz 芯片及毫米波雷达展开。如何实现宽频带、高增益天线及芯片与天线低损耗互连是中期 AiP 开发者所面临的挑战。中期也是 AiP 技术茁壮发展的阶段, 很多大公司投入大量人力物力开发适合于 AiP 设计的新材料和新工艺, 实属罕见。据我所知, 也只有在 1970 年代微带天线曾获得过如此瞩目与投入。近期 AiP 技术的研发一方面向更高的频率扩展, 另一方面正围绕着 IoT 及毫米波移动通信 5G 芯片如火如荼展开。更高频率



▲图 13 122 及 145 GHz 封装天线 SiGe 芯片实物照片



▲图 14 160 GHz eWLB 封装天线 SiGe 芯片实物照片



▲图 15 NTT 公司 300 GHz LTCC 封装天线实物照片





▲ 图 16 Silicon Labs 公司的蓝牙 CMOS 芯片内置封装天线实物照片



▲ 图 17 IBM 公司 28 GHz HDI 封装天线 SiGe 芯片实物照片



▲ 图 18 Qualcomm 公司用于 5G NR 首款智能手机参考设计中采用封装天线技术的 28 GHz 毫米波芯片实物照片

AiP 技术的关键在于材料损耗及工艺精度,5G AiP 技术的难点是如何实现高辐射效率及低成本量产。

如今 AiP 技术不仅仅被工业界广泛采用,也已从学术界天线领域扩散到集成电路、封装、材料与工艺、微波、雷达及通信等领域。这一点既可以从发表 AiP 技术相关文章的刊物看出,也可以从不同领域作者出版的书籍中窥到。比如国际著名的无线通信专家、美国纽约大学 T. S. RAPPAPORT 教授不仅在他发表的新书毫米波无线通信中专门详细介绍 AiP 技术<sup>[68]</sup>,也在很多无线通信类国际学术会议的主题演讲中用我们的 AiP 设计作为例子阐述封装天线的优点<sup>[69]</sup>。再比如以前在射频集成电路

工程师眼中,天线只不过就是一片金属,现在他们意识到没有好的天线解决方案,设计再好的射频集成电路也就是一块石英。另外,IEEE 微波理论与技术学会比天线与传播学会对推广 AiP 技术更加积极,几年前几乎在同一时段选择任命了两位杰出讲师讲授 AiP 技术。一位是德国 Karlsruhe Institute of Technology 的 Thomas ZWICK 教授,另一位是奥地利 Johannes Kepler University 的 Andreas STELZER 教授。同时在一个题目上任命了两位杰出讲师,在 IEEE 微波理论与技术学会历史上是前所未有的。德国 Karlsruhe Institute of Technology 是 Heinrich Rudolf HERTZ 1887 发现电磁波的地方,Thomas ZWICK 教授是 IBM Thomas J. Watson

Research Center 前雇员,在 AiP 设计、制造及测试方面做出过突出贡献。Andreas STELZER 教授由于在 SiGe 毫米波雷达芯片设计方面的贡献获 2011 年度 IEEE 微波理论与技术学会微波奖,在基于 eWLB 工艺开发差分 AiP 技术方面的贡献获首届 IEEE 亚太天线与传播年会最佳论文奖。

自 20 世纪 90 年代末,我有幸参与并推动了 AiP 技术的发展。早在 2001 年我就同上海交通大学毛军发教授团队就 AiP 技术进行学术交流。毛军发教授团队在三维系统级集成及多物理场仿真方面经验丰富、硕果累累。自行开发的热仿真软件对分析 AiP 热效应及散热设计非常有用。近期毛军发教授团队与中电

科 41 所进行合作,建成了中国第 1 套从 50 GHz (为适应 5G 高频段天线测试已向扩展到 18 GHz)—325 GHz (为适应太赫兹频段天线测试可扩展到 500 GHz 或更高)集成天线远场自动测试平台。该测试平台达到世界先进水平,支持探针及波导馈电,110 GHz 以下也可用同轴馈电。该测试平台已为中国多家科研院所的研究项目及公司产品开发提供了测试服务,极大地助进了中国在片上天线及封装天线方面的研究与发展。在各种科研计划的支持下,清华大学冯正和教授团队,东南大学洪伟教授团队、崔铁军教授团队,香港城市大学薛泉教授团队,香港城市大学梁国华教授团队,浙江大学尹文言教授团队,山西大学张文梅教授团队都对 AiP 技术发展做出了积极贡献。张文梅教授曾两次应邀在新加坡南洋理工大学进行长期学术访问与讲学。张文梅教授 2008 年回国后率先在国际上开展了用滤波器综合方法设计滤波天线。滤波天线目前是国际上微波与天线领域的一个研究热点,华南理工大学褚庆听教授团队、章秀银教授团队分别在滤波天线的设计方面做出了突出贡献。中国公司尽管在 AiP 技术开发方面起步比较晚,但得益于后发优势,60 GHz AiP 技术与相关芯片研发已取得重大突破,在交大测试平台多次所做的相控阵、大规模 MIMO 辐射测试获得令人满意的效果。毫米波频段 5G 移动通信 AiP 技术也已取得进展。

最后,让我将 AiP 与基片集成波导(SIW)联系起来结束这篇文章。我在 2016 年南京举办的华人微波论坛上讲过,吴柯教授及洪伟教授的合作将 SIW 技术做成微波领域的国际主流,我同刘兑现博士一起努力让 AiP 技术在国际天线领域引起人们的重视。我俩在相距很近的黄河东西岸边的乡村出生长大,相识却在远隔万里的南洋,珠联璧合,开创出封装天线一片天地,并且 3 次携手登上国



际天线领域的颁奖舞台,成就了一个小小奇迹,一段佳话。另外,我们几位都是77、78级大学生,我们的名字有着鲜明的时代特征,伟大的跃进,可否兑现?我想我们没忘初心,兑现了父辈的期望及我们自己的选择!

## 致谢

太原理工大学盛剑桓教授,香港中文大学黄振峰博士,香港中文大学程伯中教授,以及南洋理工大学杜茂安教授。

## 参考文献

- ANDREWS J G, BUZZI S, WAN C, et al. What Will 5G Be? [J]. IEEE Journal on Selected Areas in Communications, 2014, 32 (6):1065–1082. DOI: 10.1109/JSAC.2014.2328098
- BAUER F, WANG X, MENZEL W, et al. A 79-GHz Radar Sensor in LTCC Technology Using Grid Array Antennas [J]. IEEE Transactions on Microwave Theory & Techniques, 2013, 61 (6):2514–2521. DOI: 10.1109/TMTT.2013.2260766
- MAURELLI A, BELOT D, CAMPARDO G. SoC and SiP, the Yin and Yang of the Tao for the New Electronic Era [J]. Proceedings of the IEEE, 2009, 97(1):9–17. DOI: 10.1109/JPROC.2008.2007453
- HEININGER N, KIVIKOSKI M, LEE Y H. Advanced Antenna for Mobile Phone [C]// Processing of. Int. Symp. Antennas Propaga., USA:IEEE, 2005
- ZHANG Y P, LIU D. Antenna-on-Chip and Antenna-in-Package Solutions to Highly Integrated Millimeter-Wave Devices for Wireless Communications [J]. IEEE Transactions on Antennas & Propagation, 2009, 57(10):2830–2841. DOI: 10.1109/TAP.2009.2029295
- O KK, KIM K, FLOYD BA, et al. On-Chip Antennas in Silicon ICs and Their Application [J]. IEEE Transactions on Electron Devices, 2005, 52(7):1312–1323. DOI: 10.1109/TED.2005.850668
- ZHANG Y P, SUN M, GUO L H. On-Chip Antennas for 60-GHz Radios in Silicon Technology [J]. IEEE Transactions on Electron Devices, 2005, 52(7):1664–1668. DOI: 10.1109/TED.2005.850628
- ZHANG Y P, GUO L H, SUN M. High Transmission Gain Inverted-F Antenna on Low-Resistivity Si for Wireless Interconnect [J]. IEEE Electron Device Letters, 2006, 27(5): 374–376. DOI: 10.1109/LED.2006.872351
- BABAKHANI A, GUAN X, KOMIJANI A, et al. A 77-GHz Phased-Array Transceiver With On-Chip Antennas in Silicon: Receiver and Antennas [J]. IEEE Journal of Solid-State Circuits, 2006, 41(12):2795–2806. DOI: 10.1109/JSSC.2006.884811
- SUN M, ZHANG Y P. 100-GHz Quasi-Yagi Antenna in Silicon Technology [J]. IEEE Electron Device Letters, 2007, 28(5):455–457. DOI: 10.1109/LED.2007.895447
- OJEFORS E, PFEIFFER U R, LISAUSKAS A, et al. A 0.65 THz Focal-Plane Array in A Quarter-Micron CMOS Process Technology [J]. IEEE Journal of Solid-State Circuits, 2009, 44(7):1968–1976. DOI: 10.1109/JSSC.2009.2021911
- GRZYB J, STATNIKOV K, PFEIFFER U R. A 2 × 2 Lens-Integrated On-Chip Antenna System for A 820 GHz Multiplier-Chain Source in SiGe Technology [C]// Antennas and Propagation Society International Symposium. USA:IEEE, 2014:737–738. DOI: 10.1109/APS.2014.6904698
- HAN R, JIANG C, MOSTAJERAN A, et al. A SiGe Terahertz Heterodyne Imaging Transmitter With 3.3 mW Radiated Power and Fully-Integrated Phase-Locked Loop [J]. IEEE Journal of Solid-State Circuits, 2015, 50(12):2935–2947. DOI: 10.1109/JSSC.2015.2471847
- PFEIFFER U R, GRZYB J, LIU D, et al. A Chip-Scale Packaging Technology for 60-GHz Wireless Chipsets [J]. IEEE Transactions on Microwave Theory & Techniques, 2006, 54(8):3387–3397. DOI: 10.1109/TMTT.2006.877832
- ZHANG Y P, SUN M, CHUA K M, et al. Antenna-in-Package Design for Wirebond Interconnection to Highly Integrated 60-GHz Radios [J]. IEEE Transactions on Antennas & Propagation, 2009, 57(10): 2842–2852. DOI: 10.1109/TAP.2009.2029290
- DONG G K, LIU D, NATARAJAN A, et al. LTCC Packages With Embedded Phased-Array Antennas for 60 GHz Communications [J]. IEEE Microwave & Wireless Components Letters, 2011, 21(3):142–144. DOI: 10.1109/LMWVC.2010.2103932
- KAM D G, LIU D, NATARAJAN A, et al. Organic Packages with Embedded Phased-Array Antennas for 60-GHz Wireless Chipsets [J]. IEEE Transactions on Components Packaging & Manufacturing Technology, 2011, 1(11):1806–1814. DOI: 10.1109/TCPMT.2011.2169064
- HONG W B, GOUDELEV A, BAEK K H, et al. 24-Element Antenna-in-Package for Stationary 60-GHz Communication Scenarios [J]. IEEE Antennas & Wireless Propagation Letters, 2011, 10(1897):738–741. DOI: 10.1109/LAWP.2011.2162640
- HONG W B, BAEK K H, GOUDELEV A. Multilayer Antenna Package for IEEE 802.11ad Employing Ultralow-Cost FR4 [J]. IEEE Transactions on Antennas & Propagation, 2012, 60(12):5932–5938. DOI: 10.1109/TAP.2012.2214196
- COHEN E, RUBERTO M, COHEN M, et al. A CMOS Bidirectional 32-Element Phased-Array Transceiver at 60 GHz With LTCC Antenna [J]. IEEE Transactions on Microwave Theory & Techniques, 2013, 61 (3):1359–1375. DOI: 10.1109/TMTT.2013.2243749
- HONG W B, BAEK K H, GOUDELEV A. Grid Assembly-Free 60-GHz Antenna Module Embedded in FR-4 Transceiver Carrier Board [J]. IEEE Transactions on Antennas & Propagation, 2013, 61(4):1573–1580. DOI: 10.1109/TAP.2012.2232635
- TSUTSUMI Y, ITO T, HASHIMOTO K, et al. Bonding Wire Loop Antenna in Standard Ball Grid Array Package for 60-GHz Short-Range Wireless Communication [J]. IEEE Transactions on Antennas & Propagation, 2013, 61(4):1557–1563. DOI: 10.1109/TAP.2012.2232262
- BOERS M, AFSHAR B, VASSILIOU I, et al. A 16TX/16RX 60 GHz 802.11ad Chipset With Single Coaxial Interface and Polarization Diversity [J]. IEEE Journal of Solid-State Circuits, 2014, 49(12):3031–3045. DOI: 10.1109/JSSC.2014.2356462
- NASR I, JUNGMAIER R, BAHETI A, et al. A Highly Integrated 60 GHz 6-Channel Transceiver With Antenna in Package for Smart Sensing and Short-Range Communications [J]. IEEE Journal of Solid-State Circuits, 2016, 51(9):2066–2076. DOI: 10.1109/JSSC.2016.2585621
- GU XX, VALDES-GARCIA A, NATARAJAN A, et al. W-Band Scalable Phased Arrays for Imaging and Communications [J]. Communications Magazine IEEE, 2015, 53 (4):196–204. DOI: 10.1109/MCOM.2015.7081095
- BHUTANI A, GULAN H, GOETTEL B, et al. 122 GHz Aperture-Coupled Stacked Patch Microstrip Antenna in LTCC Technology [C]// European Conference on Antennas and Propagation. USA:IEEE, 2016:1–5. DOI: 10.1109/EuCAP.2016.7481147
- GOTTEL B, BEER S, PAULI M, et al. Ultra Wideband D-Band Antenna Integrated in A LTCC Based QFN Package Using AFIP-Chip Interconnect [C]// Microwave Conference (EuMC). USA:IEEE, 2013:227–230. DOI: 10.23919/EuMC.2013.6686632
- BEER S, RUSCH C, GOTTEL B, et al. D-Band Grid-Array Antenna Integrated in The Lid of A Surface-Mountable Chip-Package [C]// European Conference on Antennas and Propagation. USA:IEEE, 2013:1318–1322
- HAMIDIPOUR A, FISHER A, JAHN M, et al. 160-GHz SiGe-Based Transmitter and Receiver with Highly Directional Antennas in Package [C]// Microwave Integrated Circuits Conference. USA:IEEE, 2013:81–84
- TAJIMA T, SONG H J, YAITA M, et al. 300-GHz LTCC Horn Antennas Based on Antenna-in-Package Technology [C]// Microwave Conference (EuMC). USA: IEEE, 2013:231–234. DOI: 10.23919/EuMC.2013.6686633
- ZHANG Y P, LO K C, HWANG Y M. A Dielectric-Loaded Miniature Antenna for Microcellular and Personal Communications [C]// Antennas and Propagation Society International Symposium. USA:IEEE, 1995: 1152–1155. DOI: 10.1109/APS.1995.530223
- ZHANG Y P, DO M A. Integrated-Circuit Pressed-Ceramic Package Antenna for The Single-Chip Solution of A Wireless Transceiver [J]. Microwave & Optical Technology Letters, 2001, 30(5):330–332. DOI: 10.1002/mop.1305
- Nanyang Technological University. Integrated Circuit Package Antennas, Internal Reports [R]. Singapore:NTU, 2003.
- SONG P C T, HALL P S, GHAFOURI-SHIRAZ H. Novel RF Front End Antenna Package [J]. IEEE Proceedings-Microwaves Antennas and Propagation, 2003, 150(4): 290–294. DOI: 10.1049/ip-map:20030414

- [35] MATHEWS D J, HILL R J, GAYNOR M P, et al. Multi-Chip Semiconductor Package with Integral Shield and Antenna:US6686649 B1 [P].2001-05-14
- [36] LIM K, OBATOYINBO A, SUTONO A, et al. A Highly Integrated Transceiver Module for 5.8 GHz OFDM Communication System Using Multi-Layer Packaging Technology [C]// Microwave Symposium Digest, 2001 IEEE MTT-S International. USA:IEEE, 2002: 1739-1742. DOI: 10.1109/MWSYM.2001.967242
- [37] BREBELS S, RYCKAERT J, COME B, et al. SOP Integration and Codesign of Antennas [J]. Advanced Packaging IEEE Transactions on, 2004, 27(2):341-351. DOI: 10.1109/TADVP.2004.828822
- [38] LEUNG K W. The Hollow DRA and Its Novel Application as APackagingcover[R]. Hong Kong: Hong Kong Research Grants Council, 2004
- [39] POPOV A P, ROTARU M D. A Novel Integrated Dielectric Resonator Antenna for Circular Polarization[C]// Electronic Components and Technology Conference. USA:IEEE, 2003:470-473. DOI: 10.1109/ECTC.2003.1216320
- [40] ZHANG Y P. Antenna-in-Package Technology for Modern Radio Systems[C]// IEEE International Workshop on Antenna Technology Small Antennas and Novel Metamaterials. USA:IEEE, 2006:37-40. DOI: 10.1109/IWAT.2006.1608969
- [41] ZHANG Y P, SUN M, CHUA K M, et al. Antenna-in-Package in LTCC for 60-GHz Radiol[C]//IEEE International Workshop Antenna Technology. USA:IEEE, 2007.
- [42] ZHANG Y P, SUN M, LIU D, et al. Dual Grid Array Antennas in AThin-Profile Package for Flip-Chip Interconnection to Highly-Integrated 60-GHz Radios[J]. IEEE Trans. Antennas Propaga., 2011, 59(4):1191-1199
- [43] 黄志洵. S.A. 谢昆诺夫及其对波导理论的贡献[J]. 物理, 1988, 17(1):53-56
- [44] KRAUS J. A Backward Angle-Fire Array Antenna[J]. IEEE Transactions on Antennas & Propagation, 1964, 12(1):48-50. DOI: 10.1109/TAP.1964.1138171
- [45] SUN M, ZHANG Y P. Design and Integration of 60-GHz Grid Array Antenna in Chip Package[C]// Microwave Conference. USA: IEEE, 2008:1-4. DOI: 10.1109/APMC.2008.4958260
- [46] CHEN Z H, ZHANG Y P, BISOGNIN A, et al. An LTCC Microstrip Grid Array Antenna for 94-GHz Applications[J]. IEEE Antennas & Wireless Propagation Letters, 2015, 14(4): 1279-1281. DOI: 10.1109/LAWP.2015.2402159
- [47] BEER S, RUSCH C, GOTTEL B, et al. D-Band Grid-Array Antenna Integrated in The Lid of ASurface-Mountable Chip-Package [C]// European Conference on Antennas and Propagation. USA:IEEE, 2013:1318-1322
- [48] ZHANG B, GULAN H, ZWICK T, et al. Integration of a 140 GHz Packaged LTCC Grid Array Antenna with An InPDetector[J]. IEEE Transactions on Components Packaging & Manufacturing Technology, 2015, 5(8):1060-1068. DOI: 10.1109/TCPMT.2015.2453407
- [49] SUN M, ZHANG Y P, GUO Y X, et al. Integration of Grid Array Antenna in Chip Package for Highly Integrated 60-GHz Radios[J]. IEEE Antennas & Wireless Propagation Letters, 2010, 8:1364-1366. DOI: 10.1109/LAWP.2009.2039031
- [50] ZHANG Y P. Integrated Circuit Package Antenna: An Elegant Antenna Solution for Single-Chip RF Transceiver[R]. USA:IEEE, 2005
- [51] DONG G K, LIU D, NATARAJAN A, et al. Low-Cost Antenna-in-Package Solutions for 60-GHz Phased-Array Systems[C]// Electrical Performance of Electronic Packaging and Systems. USA:IEEE, 2010: 93-96. DOI: 10.1109/EPEPS.2010.5642554
- [52] LIU D, AKKERMANS J A G, CHEN H C, et al. Packages With Integrated 60-GHz Aperture-Coupled Patch Antennas[J]. IEEE Transactions on Antennas & Propagation, 2011, 59(10):3607-3616. DOI:10.1109/TAP.2011.2163760
- [53] COHEN E, RUBERTO M, COHEN M, et al. A CMOS Bidirectional 32-Element Phased-Array Transceiver at 60 GHz With LTCC Antenna[J]. IEEE Transactions on Microwave Theory & Techniques, 2013, 61(3):1359-1375. DOI: 10.1109/TMTT.2013.2243749
- [54] COHEN E, RUBERTO M, COHEN M, et al. Antenna Packaging of A 32 Element TX/RX Phased Array RFIC for 60 GHz Communications[C]// IEEE International Conference on Microwaves, Communications, Antennas and Electronics Systems. USA:IEEE, 2013:1-5. DOI: 10.1109/COMCAS.2013.6685306
- [55] KAMGAING T, ELSHERBINI A A, FRANK T W, et al. Investigation of APhotodefinableGlass Substrate for Millimeter-Wave Radios on Package[C]// Electronic Components and Technology Conference. USA:IEEE, 2014:1610-1615. DOI: 10.1109/ECTC.2014.6897510
- [56] KAMGAING T, ELSHERBINI A A, OSTER S N, et al. Low-Profile Fully Integrated 60 GHz 18 Element Phased Array on Multilayer Liquid Crystal Polymer Flip Chip Package [C]// Electronic Components and Technology Conference. USA:IEEE, 2015: 994-998. DOI: 10.1109/ECTC.2015.7159716
- [57] KAMGAING T, ELSHERBINI A A, OSTER S N, et al. Ultra-Thin Dual Polarized Millimeter-Wave Phased Array System-in-Package with Embedded Transceiver Chip [C]// Microwave Symposium. USA:IEEE, 2015:1-4. DOI: 10.1109/MWSYM.2015.7166967
- [58] OHATA K, MARUHASHI K, ITO M, et al. Millimeter-Wave Broadband Transceivers [EB/OL].(2005-07-30)[2017-10-25]. <http://www.nec.com/en/global/techrep/journal/g05/n03/pdf/a211.pdf>
- [59] NTT. Near-Field High-Speed Transmission Technology Using 60-GHz Band[EB/OL]. [2017-10-25]. <http://www.ntt.com.jp/mirai/organization/organization03/02.html>
- [60] Infineon, Imec. Infineon and Imec Cooperate On 79 GHz CMOS Radar Sensor Chips For The Automotive Industry[EB/OL]. (2016-05-24)[2017-10-25]. <http://www.infineon.com/cms/en/about-infineon/press/market-news/2016/INFATV201605-062.html>
- [61] GU X, LIU D, BAKS C, et al. A Compact 4-Chip Package with 64 Embedded Dual-Polarization Antennas for W-Band Phased-Array Transceivers[C]// Electronic Components and Technology Conference. USA:IEEE, 2014:1272-1277. DOI:10.1109/ECTC.2014.6897455
- [62] STELZER A, FEGGER R. Integrated Microwave Sensors in SiGe with Antenna in Package: From Concepts to Solutions[C]// International Conference on Applied Electromagnetics and Communications. USA:IEEE, 2016:1-7. DOI: 10.1109/ICECom.2016.7843873
- [63] TAJIMA T, SONG H J, YAITA M, et al. 300-GHz LTCC Horn Antennas Based on Antenna-in-Package Technology[C]// Microwave Conference. USA:IEEE, 2013: 231-234. DOI: 10.23919/EuMC.2013.6686633
- [64] TAJIMA T, SONG H J, AJITO K, et al. 300-GHz Step-Profiled Corrugated Horn Antennas Integrated in LTCC[J]. IEEE Transactions on Antennas & Propagation, 2014, 62(11):5437-5444. DOI: 10.1109/TAP.2014.2350520
- [65] Fujitsu. Fujitsu Develops World's First Compact 300 GHz Receiver for Wireless Communications of Tens of Gigabits per Second s[EB/OL]. (2015-09-08)[2017-10-25]. <http://www.fujitsu.com/global/about/resources/news/press-releases/2015/0908-02.html>
- [66] IBM. IBM & Ericsson Announce Research Advance for 5G Communications Networks [EB/OL]. (2017-02-07)[2017-10-20]. <https://www-03.ibm.com/press/us/en/pressrelease/51542.wss>
- [67] Qualcomm. Qualcomm Achieves World's First Announced 5G Data Connection on a 5G Modem Chipset for Mobile Devices[EB/OL]. (2017-10-18)[2017-10-25]. <https://www.qualcomm.com/news/releases/2017/10/17/qualcomm-achieves-worlds-first-announced-5g-data-connection-5g-modem>
- [68] RAPPAPORT T S, HEATH R W, DANIELS R C, et al. Millimeter-Wave Wireless Communications[M]. New York: Prentice Hall, 2015
- [69] RAPPAPORT T S, GUTIERREZ F. The Emerging World of Massively Broadband Devices: 60 GHz and Above[EB/OL]. (2009-06-04)[2017-10-25]. [http://wireless.engineering.nyu.edu/presentations/Keynote\\_Rappaport.pdf](http://wireless.engineering.nyu.edu/presentations/Keynote_Rappaport.pdf)

## 作者简介



张跃平, 新加坡南洋理工大学电子工程学讲座教授、IEEE Fellow、IEEE 天线与传播学会杰出讲师、上海交通大学“千人计划”国家特聘专家、《ZTE COMMUNICATIONS》杂志编委; 目前主要研究方向为无线电电子学; 曾荣获 IEEE 天线与传播学会谢昆诺夫论文奖; 已发表论文 260 篇, 拥有授权美国专利 7 项。

# 面向沉浸式视频编码和传输的高效投影方法

## Efficient Projections for Immersive Video Coding and Transmission

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2017) 06-0050-06

**摘要:** 全景视频投影技术是全景视频应用中的一项关键技术,为此成立了 IEEE 1857.9 沉浸视频内容编码工作组,致力于制订高效的沉浸视频投影和编码方法。IEEE 1857.9 沉浸视频内容编码工作组设计了一系列的映射格式,其中双极方形的投影格式相较于球面投影技术(ERP)取得了 11.56% 的编码性能增益。此外,IEEE 1857.9 工作组还针对流切换的场景设计了基于主视点的投影模型,在保证主视点区域视频质量不变的情况下,可获得 70% 的编码性能增益。

**关键词:** 虚拟现实;沉浸视频;投影;编码

**Abstract:** Since the panoramic video projection technology is a key technology in panoramic video applications, the IEEE 1857.9 immersive video content coding working group was established so as to develop efficient immersive video projection and coding methods. A series of projection schemes are designed by the IEEE 1857.9 immersion video content coding working group, in which the bipolar square projection format achieves 11.56% coding performance gain compared to the equirectangular projection (ERP). In addition, for the flow switching scenario, a projection model based on the main view point is designed by the IEEE 1857.9 working group, which can achieve 70% coding performance gain while guaranteeing the video quality of the main view area.

**Key words:** virtual reality; immersive video; projection; encoding

王悦名/WANG Yueming  
王荣刚/WANG Ronggang  
高文/GAO Wen

(北京大学深圳研究生院,广东 深圳 518055)  
(Shenzhen Graduate School, Peking University, Shenzhen 518055, China)

虚拟现实(VR)是一种计算机仿真技术,它使用头戴式设备(有时与物理空间或多投影环境相结合)产生逼真的图像、声音和其他感觉,给用户带来沉浸式的体验。沉浸视频(或球面视频,全景视频)是通过全景摄像机在同一时刻捕获 360° 范围内的视频来获得,在观看过程中,观看者能看到全方位的视频,感受到身临其境的感觉。然而现有的编码和存储技术不支持球形视频的处理,因

此球形全景视频需要投影到二维平面上进行存储和编码。

将球面全景视频映射到二维平面上方法多种多样<sup>[1]</sup>,其中最常见的是球面投影技术(ERP)<sup>[2]</sup>,但是 ERP 技术在高纬度区域存在很严重的过采样,因此会浪费很多传输带宽。随后,等面积圆柱投影(与 ERP 类似,但其通过降低两极在纬度方向的采样密度是球面总体的采样密度一致)<sup>[3]</sup>、立方体投影(通过透视投影的方法,将球面投影到立方体平面上)<sup>[4]</sup>、自适应条带投影(将 ERP 格式的全景图像根据纬度分割成多个

条带,然后根据图像内容以及条带的纬度来进行自适应的下采样)<sup>[5]</sup>、条带分割投影(将球面视频根据纬度分割成多个环带,两极投影为平面,中间环带投影为矩形)<sup>[6]</sup>等投影方案被纷纷提出。这些投影方法可以减少全景视频的传输带宽,但是由于全景视频需要提供 360° 的视频内容,因此全景视频的传输带宽和播放复杂度仍然比传统视频大得多。此外,随着虚拟现实技术的视频采集设备的发展,全景视频的分辨率也提高到 8K 甚至更高。为了解决上述问题,IEEE 1857.9 沉浸视频内容编码工作组于 2015 年 12 月 5 日成立,该工作组致力于制订高效的沉浸视频投影和编码方法。

IEEE1857.9 工作组致力于推动用于压缩、解压缩和重建沉浸式视觉内容的高效编码工具的标准化。该标准针对的应用场景和服务对象包括但不限于 VR,例如:基于无人机的 VR、增强现实、全景视频和其他视频/音频驱动的服务,以及诸如沉浸式视频流、广播、存储和通信之类的

收稿日期: 2017-10-20  
网络出版日期: 2017-11-08



应用。最近,IEEE 1857.9 沉浸视频内容编码工作组拟定的标准草案即将定稿。本文对 IEEE1857.9 工作组采纳或研究的投影技术进行了综述。

## 1 全景投影技术

在众多的投影方法中,ERP 投影是最常用的投影格式,但是其编码效率却非常低。为了提高 VR 视频的编码效率,IEEE1857.9 工作组研发了一系列针对 VR 视频的高效投影格式,并且在标准中采纳了其中的部分投影格式。

### 1.1 ERP 投影

最常见的全景视频的投影格式是 ERP 投影<sup>[2]</sup>。ERP 投影根据等间隔的经度和纬度将球面投影到二维平面上,如图 1 所示。

ERP 在 VR 视频中使用很广泛,但是 ERP 存在很多问题,比如这种投影方式会造成两极区域的过采样和失真,导致使用 ERP 投影得到的平面视频的编码效率很低。

### 1.2 多面体投影

多面体投影<sup>[6]</sup>通常通过透视投影将球面投影到外切多面体上。在 IEEE1857.9 工作组会议中,多种多面体投影模型被提出,包括:立方体投影、八面体投影、二十面体投影。

多面体的面越多,越接近球面,采样密度也就越均匀,多面体投影的多种投影模型如表 1 所示。

### 1.3 双极方形投影

分析发现 ERP 投影格式在两极区域存在较严重的过采样,而在中间区域采样密度比较均匀(但不完全均匀),针对以上特点,设计出了一种双极方形的投影格式<sup>[7]</sup>。如图 2 所示,双极方形投影将球面根据纬度(南北纬 45°)划分为 3 个区域,根据等间距的经度和纬度将球的区域 II 投影到二维平面上(投影方法与 ERP 相同),并将区域 I 和区域 III 投影到由多个

图 1 ERP 投影示意

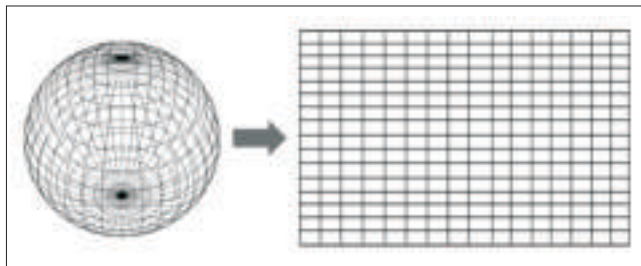


表 1 多种多面体投影示意

投影方式	3D 模型	投影得到的 2D 平面
立方体投影 (6 面)		
八面体投影 (8 面)		
二十面体投影 (20 面)		

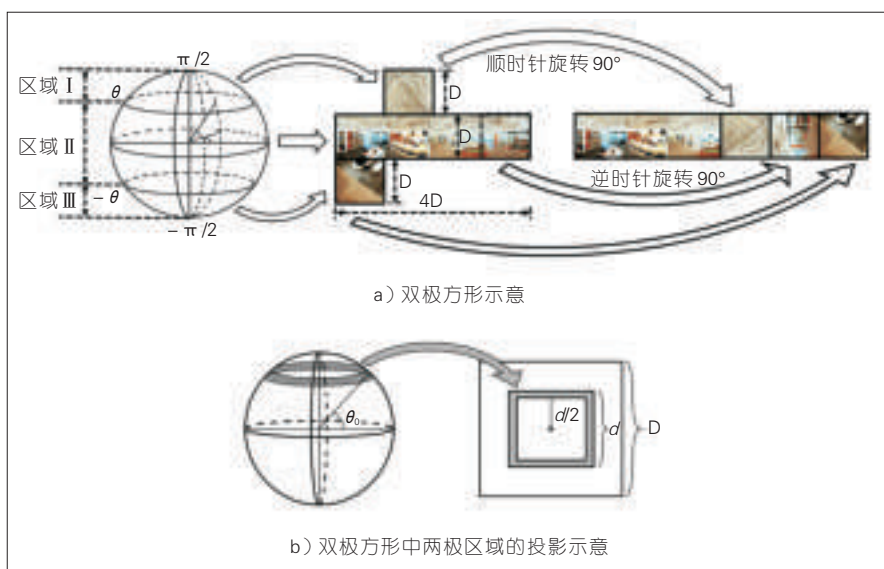


图 2 双极方形投影

同心方形环组成的矩形平面中。

### 1.4 双环带投影

双环带投影<sup>[8]</sup>通过北纬 30° 和南纬 30° 的两条纬线将球面分成 3 部分。中间区域是一个环形区域,称为“环区域”;另外的两部分分别是顶区域和底区域。然后,双环带投影进一步将环区域分成 6 个均匀的子区域,将顶和底区域分别分成 4 个子区域,球面的区域划分如图 3 所示。双环

带投影进一步将这 14 个子区域投影为平面上的 14 个方向平面,并进一步将 14 个方向平面拼接成一个矩形平面。

## 2 非对称投影技术

在具有一对多信道和反馈信道(如直播广播系统、视频点播系统等)的应用场景中,动态流切换技术能在很大程度上减少传输带宽和播放的复杂度。在动态流切换技术中,全景

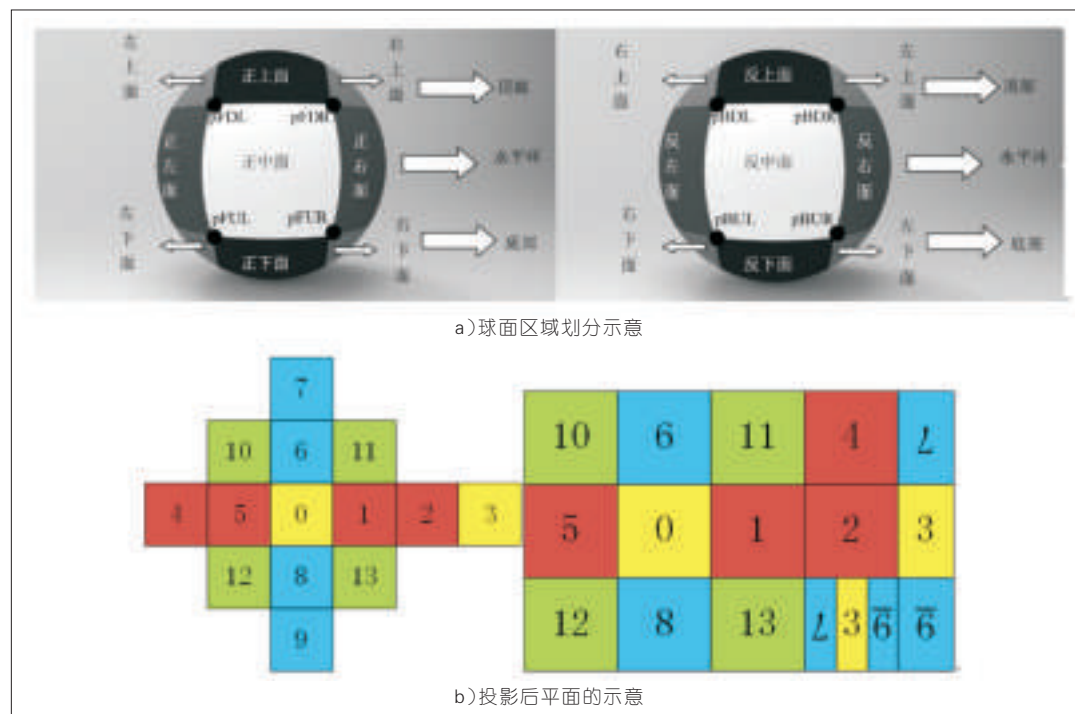


图3  
双环带投影示意

视频由覆盖整个全景视频的多个主视点不同的非对称投影序列表示。根据人观看VR视频时头部转动的方向,每一时刻只传输一个主视点与当前头部朝向最接近的非对称投影序列的码流。下一时刻,如果观看者头部的转动幅度超过现有的主视点范围,服务端将传输对应主视点区域的非对称投影序列的码流到客户端。通常,全景视频的非对称投影格式的主视点区域的采样密度较高,而非主视点区域的采样密度很低,并且整个非对称投影格式的全景视频的分辨率低于全分辨率的全景视频。IEEE1857.9工作组设计了几种非对称投影格式,有效地降低了传输带宽和播放复杂度。

### 2.1 非对称圆形投影

非对称圆形投影(ASC)<sup>[9]</sup>根据每个区域与主视点中心的角度将VR视频的球面划分成3个区域。如下图所示,C是主视点的中心,区域I是主视点区域,区域I与主视点中心之间的角度小于 $Z_1$ ,区域II是非主视点区域,区域和主视点中心之间的角度

从 $Z_1$ 到 $Z_2$ ,区域III也是非主视点区域,区域和主视点中心之间的角度大于 $Z_2$ 。ASC将图4 a)球面上的区域I、区域II和区域III分别投影到图4 b)中的二维平面上的区域I(半径为 $0 \sim \rho_1$ )、区域II(半径为 $\rho_1 \sim \rho_2$ )和区域III(半径为 $\rho_3 \sim 0$ )。

ASC在主视点区域使用等积投影,以确保主视点的采样密度高且均匀;在非主视点区域,采样密度随着区域与主视点中心角度的增加而减小。另外,主视点区域的大小是可变的, $Z_1$ 、 $Z_2$ 和 $\rho_1$ 可以根据应用场景、网络状况或其他因素自定义,因

此ASC投影具有很高的灵活性。

### 2.2 等角金字塔投影

等角金字塔投影(EAP)<sup>[10]</sup>首先通过透视投影将球面投影到金字塔,然后将金字塔投影到二维平面。图5 a)是主视点中心为D时,通过透视投影将球面投影到金字塔的示意图。金字塔的底面对应于主视点区域(主视点区域的角度为 $\theta$ ),而金字塔的侧面对应于非主视点区域。图5 b)是展开后金字塔的示意图。图5 c)是通过将金字塔投影到二维平面的示意图。

金字塔底面对应主视点区域,

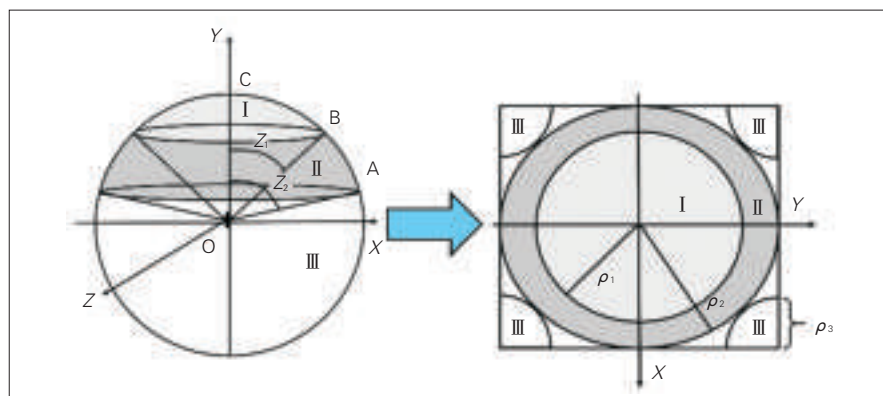


图4 非对称圆形投影示意

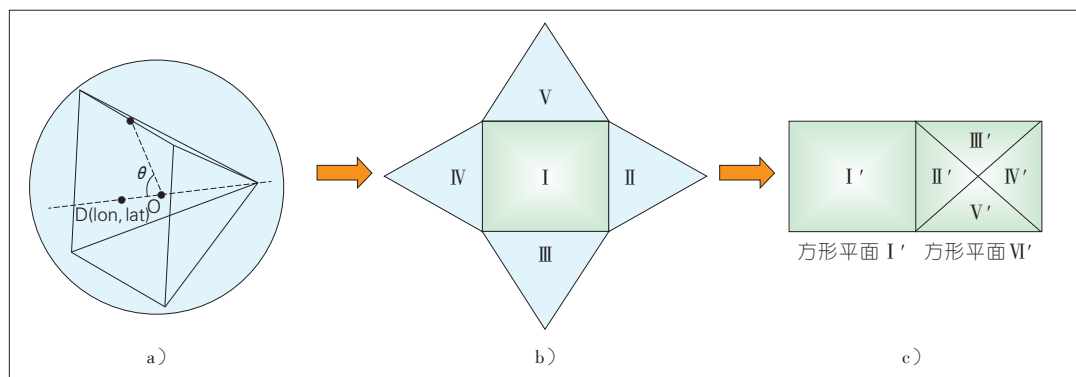


图5  
等角金字塔投影的示意

EAP对金字塔的主视点区域使用等角投影;金字塔侧面对应非主视点区域,使用采样密度按梯度下降的方式投影,离主视点越远,采样密度越低。

### 2.3 非对称投影的自适应滤波方法

在非对称投影中,由于非主视点区域的采样密度较低,该区域产生锯齿现象。为了消除非主视点区域的锯齿现象,提出了一种用于非对称投影的自适应滤波方法<sup>[11]</sup>。该方法将非主视点区域划分为多个子区域,并对不同的子区域使用不同强度的滤波器进行滤波。

如图6所示,对于EAP,自适应滤波方法将非主视点区域划分为4个子区域(A、B、C、D),分别对子区域A、B、C和D分别进行强滤波、中等强度滤波、弱滤波和无滤波。自适应滤波可以消除锯齿效应,降低码率。

## 3 不同映射格式编码效率的测试

为了评估不同投影格式的编码

效率,IEEE1857.9工作组进行了一系列的测试。

IEEE1857.9工作组使用LETIN VR<sup>[12]</sup>提供的测试序列集,其中包含8个分辨率为4 096×2 048的ERP格式的VR视频序列,4个分辨率为8 192×4 096的ERP格式的VR视频序列。测试序列的长度为300帧,其帧率为30 f/s。

实验流程如图7所述。将分辨率更高的ERP格式的原始序列作为参考基准,然后使用不同的投影格式将原始序列投影到不同的平面格式。

本实验使用RD-VR16.1编码和解码不同投影格式生成的序列。最后,计算解码后的序列和原始序列之间的球面峰值信噪比(S-PSNR)值。

在实验中,将ERP格式作为实验基准。实验结果如表2所示,立方体投影、八面体投影、二十面体投影、双极方形投影和双环带投影的增益分别为:3.76%、2.93%、8.39%、11.56%和8.61%。

对非对称投影格式的测试流程如图8所示,使用RD-VR16.1(RA的默认配置)对ERP格式的序列进行编

图6  
自适应滤波示意

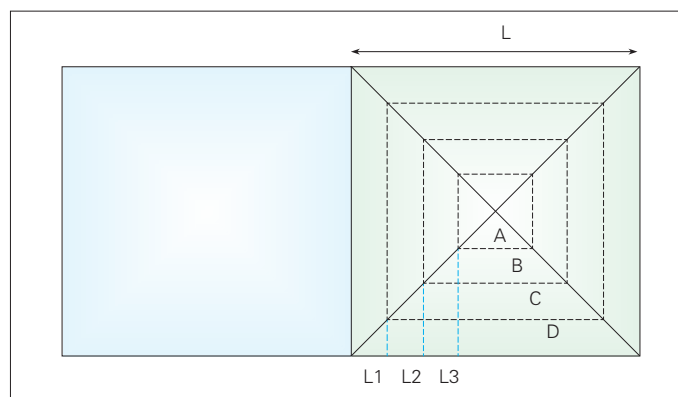
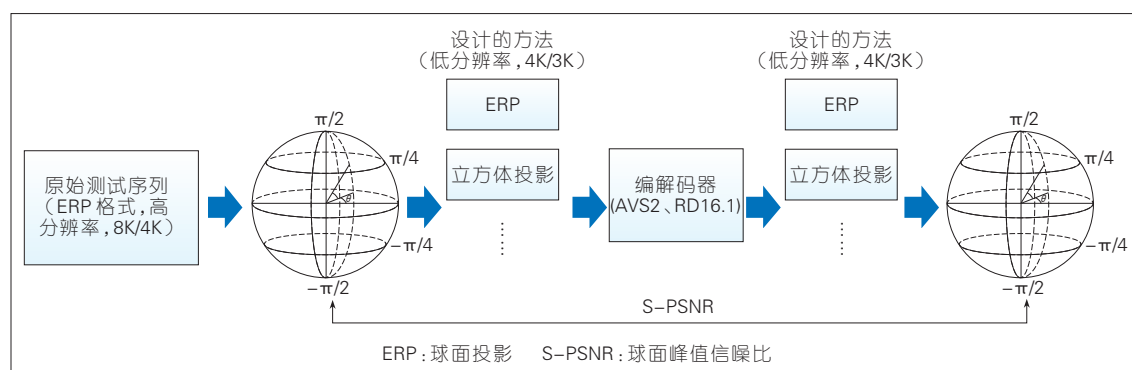


图7  
实验流程





▼表2 不同格式的编码效率的实验结果

原始分辨率	序列	立方体投影	八面体投影	二十面体投影	双极方形投影	双环带投影
4 096x2 048	Fengjing_1	-19.81%	-22.16%	-25.40%	-29.27%	-25.17%
	Fengjing_3	-3.35%	-5.55%	-14.34%	-23.18%	-17.35%
	Hangpai_1	-8.09%	-3.52%	-7.16%	-10.89%	-9.28%
	Hangpai_2	-14.73%	-10.33%	-12.45%	-13.85%	-14.88%
	Hangpai_3	-7.08%	-3.46%	-7.87%	-9.04%	-9.30%
	Xinwen_1	-0.83%	6.38%	3.07%	-1.88%	0.25%
	Xinwen_2	-0.37%	10.74%	3.88%	-0.24%	3.62%
	Yanchanghui_2	3.73%	8.85%	0.64%	-7.31%	-1.59%
8 192x4 096	THEPLACE	16.14%	11.87%	0.29%	1.9%	5.74%
	Natatorium	-1.94%	-0.23%	-9.86%	-11.43%	-10.26%
	Highway	-13.75%	-19.43%	-26.87%	-27.25%	-25.00%
	Canolafield	6.02%	-1.83%	-4.58%	-6.24%	-0.15%
4K 均值		-6.31%	-2.38%	-7.45%	-11.96%	-9.21%
8K 均值		1.62%	-2.40%	-10.25%	-10.75%	-7.42%
所有序列的均值		-3.67%	-2.39%	-8.39%	-11.56%	-8.61%

码和解码,然后计算解码序列与原始序列之间的加窗 S-PSNR。

非对称投影的实验过程如图9所示:首先将ERP格式的序列分别投影为ASC和EAP格式;然后,使用RD-VR16.1(RA的默认配置)对ASC和EAP格式的序列进行编码和解码;最

后,计算解码序列与原始序列之间的加窗 S-PSNR。实验参数的设置如下:ASC的分辨率为 $1\,448 \times 1\,448$ , $Z_1$ 、 $Z_2$ 和 $\rho_1$ 分别设置为 $45^\circ$ 、 $90^\circ$ 和 $512$ 。EAP分辨率为 $2\,048 \times 1\,024$ , $\theta$ 为 $45^\circ$ 。

实验结果如表3所示。由于ASC的主视点区域是圆形的,因此添加了

使用循环窗口计算加窗 S-PSNR 的一组数据。实验结果表明:ASC和EAP能在主视点区域实现超过50%的增益,并且EAP的BD-rate增益比ASC高出4%~10%,这主要是由于ASC存在被浪费的区域。但是ASC中的所有参数( $Z_1$ 、 $Z_2$ 和 $\theta$ )都是可调的,因此ASC比EAP更灵活。例如:当主视点区域设置为 $120^\circ$ 时,ASC可以实现比EAP更好的性能。另外,自适应滤波可以在EAP的基础上进一步将码率降低约11%,并且消除非主视点区域的锯齿现象。

#### 4 结束语

IEEE1857.9工作组致力于标准化用于压缩、解压缩和重建沉浸式视觉内容的高效编码工具。该工作组已经设计了一系列针对VR视频的投影格式。被采纳的全景视频投影格式最多能节省30%的码率。此外,工作组为动态流切换应用设计了各种投影格式和自适应滤波方法,实现了70%的编码性能的增益。未来IEEE1857.9工作组还将继续致力于沉浸

图8  
基准实验流程

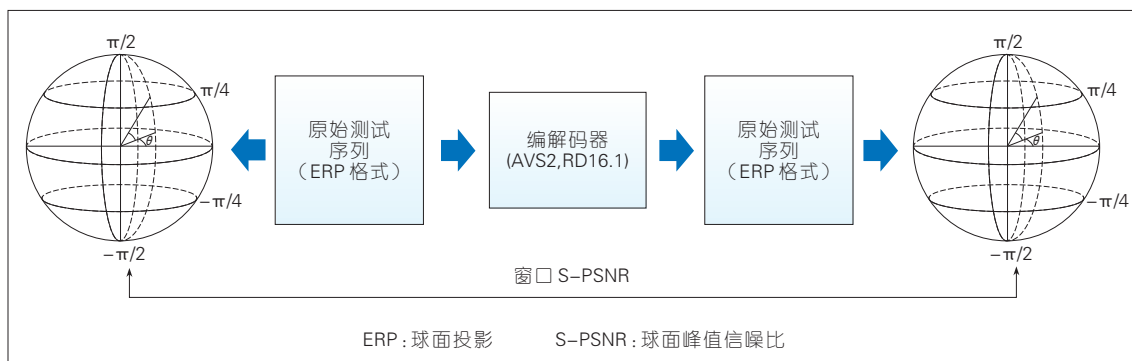
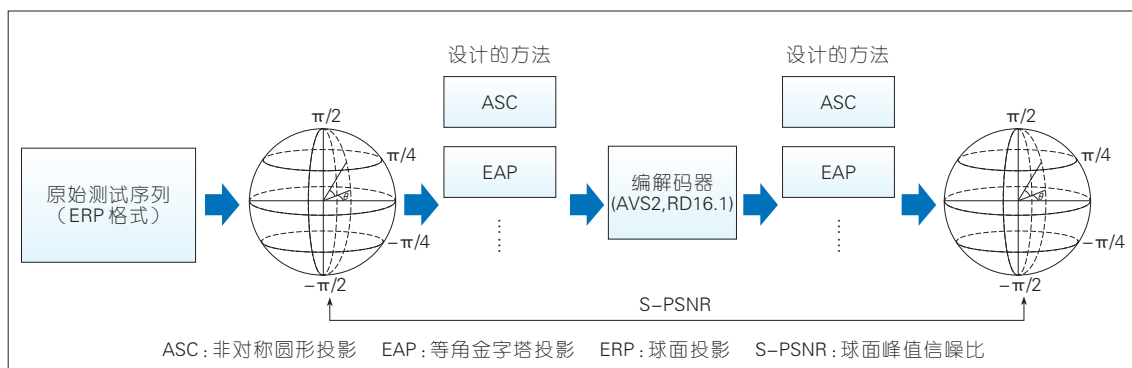


图9  
非对称投影实验流程



▼表3 实验结果

窗口大小	方形窗口计算的 S-PSNR			圆形窗口计算的 S-PSNR
	ASC	EAP	EAP + 自适应滤波	ASC
70°	-54.71%	-59.28%	-70.19%	-55.08%
80°	-52.98%	-59.45%	-70.31%	-54.87%
90°	-49.07%	-58.45%	-69.61%	-54.36%
100°	-42.47%	-54.63%	-66.96%	-46.82%
110°	-36.56%	-50.70%	-63.89%	-40.94%
120°	-30.91%	-46.30%	-58.83%	-35.56%
130°	-24.84%	-41.20%	-50.71%	-30.25%
140°	-18.59%	-35.40%	-33.39%	-24.28%
150°	-11.32%	-28.21%	-11.65%	-18.02%

ASC: 非对称圆形投影 EAP: 等角金字塔投影 S-PSNR: 球面峰值信噪比

式视觉内容的高效编码工具的研究, 寻求新的高效的映射格式, 此外还将在 3D 沉浸式视觉内容的映射和编码工具上投入更多的研究。

## 参考文献

- [1] ZORIN D, BARR A H. Correction of Geometric Perceptual Distortions in Pictures [C]// Conference on Computer Graphics and Interactive Techniques. USA: ACM, 1995: 257-264. DOI: 10.1145/218380.218449
- [2] SMOLIC A, MCCUTCHEN D. 3DAV Exploration of Video-Based Rendering Technology in MPEG[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2004, 14(3):348-356. DOI: 10.1109/TCSVT.2004.823395
- [3] NG K T, CHAN S C, SHUM H Y. Data Compression and Transmission Aspects of Panoramic Videos[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2005, 15(1):82-95. DOI: 10.1109/TCSVT.2004.839989
- [4] YU M, LAKSHMAN H, GIROD B. Content Adaptive Representations of Omnidirectional Videos for Cinematic Virtual Reality[C]// The 3rd International Workshop on Immersive Media Experiences. USA: ACM, 2015:1-6. DOI: 10.1145/2814347.2814348
- [5] LI J, WEN Z, LI S, et al. Novel Tile Segmentation Scheme for Omnidirectional Video[C]// IEEE International Conference on Image Processing. USA: IEEE, 2016:370-374.
- [6] CHOIK P, VLADYSLAVZ, CHOIM, et al. On 2D Representation Format of Panoramic Video: IEEE1857.9-04-M1027[S]. Guiyang: IEEE, 2016:6
- [7] WANG Y M, WANG R G, WANG Z Y, et al. A New Panoramic Video Projection Scheme: IEEE1857.9-04-M1028[S]. Guiyang: IEEE, 2016:6
- [8] LIN C, GU X, WU C J, et al. A New Projection Method of Omnidirectional Video: IEEE1857.9-04-M1025[S]. Guiyang: IEEE, 2016:6
- [9] WANG Y M, WANG R G, WANG Z Y, et al. A New Asymmetric Projection Scheme Based on ROI: IEEE1857.9\_M1060/AVSVR\_M1060 [S]. Haikou: IEEE, 2016:12
- [10] WANG Y M, WANG R G, WANG P, et al. Improved Pyramid Projection: IEEE1857.9\_M1107/AVSVR\_M1107[S]. Dalian: IEEE, 2017:8
- [11] SHENG X J, WU Y X, YIN H B, et al. An Improved Pyramid Projection Based on Adaptive Filtering: IEEE1857.9\_M1111/AVSVR\_M1111[S]. Dalian: IEEE, 2017:8
- [12] CHEN J, ZHANG J. Projection CE: Summary Report: IEEE1857.9\_M1072/AVSVR\_M1072 [S]. Haikou: IEEE, 2016:12

## 作者简介



王悦名, 北京大学信息工程学院数字媒体研究中心在读硕士研究生; 主要研究方向为视频编解码技术和虚拟现实视频编码技术; 已发表论文 2 篇, 申请发明专利 3 项。



王荣刚, 北京大学信息工程学院教授、院长助理, 国际 MPEG 互联网视频压缩标准 (IVC) 专题组联合组长; IEEE1857.9 专题组组长; 主要研究方向为新一代视频编解码技术、视频增强和虚拟现实视频技术等; 已发表论文 100 余篇, 申请国际/国内发明专利 70 余项。



高文, 北京大学教授、国家自然科学基金委员会副主任、中国工程院院士、ACM/IEEE Fellow; 主要研究方向为智能人机交互与多媒体技术、计算机视觉、多媒体数据压缩等; 曾获国家技术发明二等奖、科技进步二等奖和自然科学二等奖; 已出版著作 6 部, 在 IEEE T-CSVT、IEEE T-MM、IEEE T-LP 等顶级国际期刊发表论文近百篇。

## 《中兴通讯技术》2018 年专题计划

期次	专题名	专题名
1	5G 承载网技术和优化组网	中国联通研究院 张云勇院长
2	大数据智能化无线网络技术	重庆邮电大学 陈前斌教授
3	毫米波与太赫兹通信技术	东南大学 洪伟教授 东南大学 王海明教授
4	5G 回传网络光电子器件技术	中兴光电子技术有限公司 孙笑晨技术总监 中兴光电子技术有限公司 徐积勇总经理
5	可再生能源供电的无线通信	清华大学 牛志升教授
6	区块链技术及其物联网应用	北京航空航天大学 刘建伟教授

# 智能工业装备操作系统技术及创新

## Technologies and Innovations of Intelligent Industrial Equipment Operating System

崔云峰/CUI Yunfeng  
钟卫东/ZHONG Weidong  
刘东/LIU Dong

(中兴通讯股份有限公司, 广东 深圳  
518057)  
(ZTE Corporation, Shenzhen 518057, China)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2017) 06-0056-005

**摘要:** 针对工业技术的发展和《中国制造 2025》发展规划, 探讨了智能工业装备操作系统的重要性的意义。围绕智能工业装备的 4 种关键需求(关键性能、智能化、信息安全、功能安全), 中兴通讯提出了智能工业装备操作系统及其创新方案, 通过 5 种创新技术(安全可信的微内核架构、嵌入式虚拟化技术、健康管理技术、智能互联技术、安全防护技术)解决智能装备的关键需求, 并将在通信、汽车、工业控制等诸多领域促进智能工业装备发展。

**关键词:** 智能工业装备; 实时操作系统; 微内核; 嵌入式虚拟化

**Abstract:** In this paper, the importance and significance of the operating system of intelligent industrial equipment are discussed. Focusing on four key requirements of intelligent industrial equipment including key performance, intelligence, security and safety, the intelligent industrial equipment operating system and its innovative scheme are proposed by ZTE Corporation. In this way, key requirements of intelligent equipment are solved by five innovation technologies, such as safety credible micro-kernel architecture, embedded hypervisor technology, health management technology, intelligent ecosystem and interconnection technology, and security technology. Furthermore, the development of intelligent industrial equipments including communications, automotive, and industrial control should be promoted too.

**Key words:** intelligent equipment; real-time operating system; micro-kernel; embedded hypervisor

### 1 智能工业发展趋势

信息化和工业化的发展以及深度融合, 将在世界范围内带来深刻的变革。以物联网、大数据、人工智能为代表的新一代信息技术与现有工业进行深度融合, 并推动工业技术变革。全面和持续的技术变革必将改变产业格局, 因此部分发达国家相继颁布再工业化战略, 国际竞争日趋激烈。

中国连续多年制造业总产值世界第一, 成为世界的制造大国。但中国与先进国家相比还有较大差距: 创新能力弱, 关键核心技术与高端装备对外依赖度高; 产业结构不合理, 高端装备制造业和生产性服务业发展滞后; 信息化水平不高, 与工业化融合深度不够; 产业国际化程度不高, 企业全球化经营能力不足等。

为了推进中国从制造大国向制造强国的转变, 国务院组织编制并正式发布了《中国制造 2025》, 对中国制造业转型升级和跨越发展作了整体部署, 总体分三步走的战略规划。为

了力争到 2025 年达到国际领先地位或国际先进水平, 《中国制造 2025》选择了十大优势和战略产业作为突破点, 包括: 新一代信息技术产业(含芯片、操作系统、信息通信等基础产业)、高档数控机床和机器人、航空航天装备、海洋工程装备及高技术船舶、先进轨道交通装备、节能与新能源汽车、电力装备、农业装备、新材料、生物医药及高性能医疗器械。

新一代信息技术产业是其他行业发展的基础, 而操作系统是整个智能工业发展的核心基础。在《中国制造 2025 重点领域技术路线图》中也明确指出: “新一代科技革命与产业变革是以数字化网络化智能化为特

征, 操作系统是工业数字化网络化智能化的基石, 是新一轮工业革命的核心要素。”《中国制造 2025 重点领域技术路线图》中对智能工业装备操作系统领域有明确的规划: “到 2025 年, 绝大部分核心技术取得突破, 形成自主可控的操作系统与工业软件及其标准体系, 自主工业软件市场占有率超过 50%”。

### 2 智能工业装备操作系统的关键需求

智能工业装备操作系统作为底层的基础软件, 存在两个方面的需求来源。首先需要满足传统嵌入式系统的功能和性能需求; 其次是由物联

收稿日期: 2017-10-11  
网络出版日期: 2017-11-08



网、大数据、人工智能等新技术引入的智能化的功能和性能需求。总体来说包括4方面需求,如图1所示。

(1)关键性能需求。首先,智能工业装备由传统的嵌入式实时系统发展而来,系统具有明确的硬实时需求。例如:高速行驶汽车上自动驾驶系统属于硬实时系统,需系统及时处理外部事件,否则可能造成不可预期的后果。其次,嵌入式实时系统存在严格的确定性需求。系统的某些关键业务必须在确定的时间内完成。Windows、Linux、Android等桌面操作系统或手机操作系统,采用相同的软硬件,有时业务运行快,有时业务运行慢。而嵌入式系统中的关键任务需要明确固定的执行节奏,从而保证系统行为的确定性。

(2)智能化需求。首先,智能工业装备引入互联网、大数据、人工智能等新技术,这些新技术需要操作系统提供开放的智能软件生态。新技术多数从IT行业发展而来,依赖开放的智能软件生态,例如:Android的应用生态、深度学习软件生态等。其次,智能工业装备可能由多个系统相互协同完成工作任务。因此需要底层操作系统提供互联互通技术,以及进一步提供互操作互调用的机制。

(3)功能安全需求。部分嵌入式系统涉及到行业以及个人生命安全,例如:工业控制设备、轨道交通控制设备、汽车自动驾驶系统等。对于这类系统需要进行系统全面的失效分析,从失效概率、危害大小、危害可控

性等维度评估出业务模块的功能安全等级。需要底层的操作系统符合响应功能安全等级的要求,并且提供故障监测控制、故障隔离以及故障恢复的功能。

(4)信息安全需求。开放的软件生态环境以及互联网技术引入到嵌入式实时操作系统上,也给现有封闭的嵌入式系统带来了安全隐患。大规模的开放软件生态环境可能含有安全漏洞、未知后门;高速的以太网接入方式也给黑客提供了便捷的攻击路径。但是嵌入式系统的网络安全问题不能直接照搬IT领域的安全解决方案,需要考虑到现有嵌入式系统的实时性、确定性要求;同时频繁升级、打补丁会影响系统可用性,因此同样不适用于嵌入式系统。

### 3 中兴通讯智能工业装备操作系统及技术创新方案

中兴通讯立足于自主创新,从2003年开始研发自主可控的嵌入式操作系统,经过10余年的积累掌握底层操作系统的核心技术。中兴通讯所研制的电信级嵌入式实时操作系统于2016年荣获了第4届中国工业大奖(工业领域最高荣誉),并于2017年荣获了国际软件博览会金奖。目前中兴嵌入式操作系统广泛应用于通信、电力、轨道交通、汽车、航空等众多工业领域,累计发货量达到2亿,稳定运行于全球160个国家和地区。

针对智能工业装备的关键需求,

中兴通讯自主研发了智能工业装备操作系统ZEOS。如图2所示,中兴智能工业装备操作系统以安全可信的微内核架构和嵌入式虚拟化技术为支撑,同时具备实时操作系统和嵌入式Hypervisor的功能。通过健康管理技术实现故障隔离、运行状态监测控制、故障恢复,保障系统持续稳定运行;通过面向服务的互联网络架构支持多种异构系统互联互通,并借助虚拟机支撑智能应用生态;通过内建的安全可信体系架构保障基础系统安全可信,引入移动目标防御技术防御未知威胁。

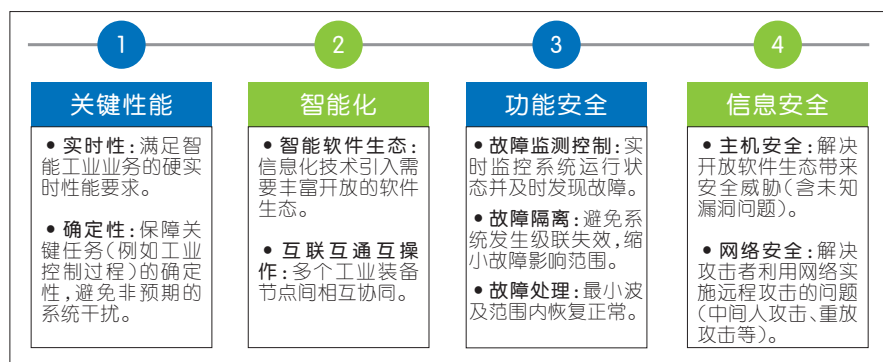
#### 3.1 安全可信的微内核架构

中兴通讯从关键需求出发,研究系统实时性、确定性的关键影响因素以及内核各个模块的代码变更和质量特征,得出智能工业装备操作系统的微内核软件架构。依据安全系统的最小特权原则(POLP),建立智能工业装备操作系统的可信计算基(TCB)。如图3所示,在处理器最高特权模式下仅保留少量的核心功能模块(空间管理、线程管理、进程间通信机制、中断管理部分)作为系统的可信计算基;其余的操作系统功能模块运行于普通模式(普通模式)。以此基础架构为支撑,中兴智能装备操作系统提供实时操作系统功能和嵌入式虚拟化功能。

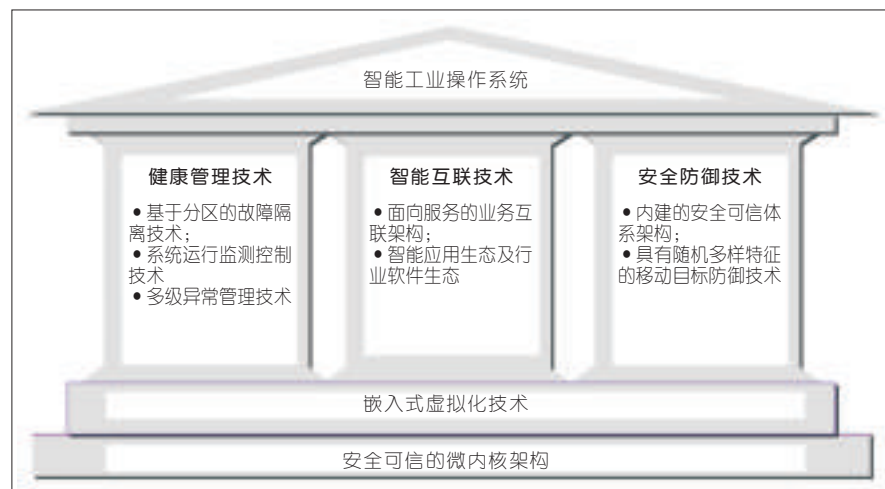
智能工业装备操作系统的基础架构有利于保障智能工业装备的关键需求,例如实时性、确定性、功能安全、信息安全。

(1)实时性:利用微内核的软件架构可以保障硬实时性能。通过减少内核功能模块,减少系统关中断、关抢占的频度和时长,从而有效地提升系统的实时性指标(中断延时、调度延时)。

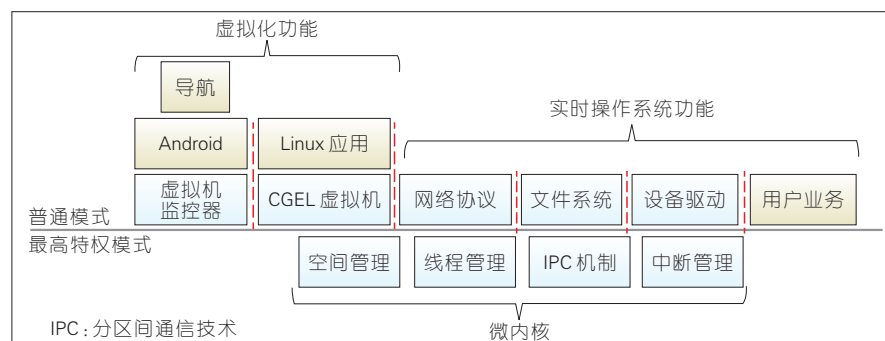
(2)确定性:利用微内核架构可减少后台干扰,保障业务的确定性。从架构设计上减少内核功能模块,减少后台任务和定时器,降低系统自身



▲ 图1 智能工业装备的关键需求



▲ 图2 中兴通讯智能装备操作系统创新技术



▲ 图3 安全可信的微内核架构

开销,从而保障用户业务的确定性。

(3)功能安全和信息安全:利用微内核架构可隔离核心功能模块和业务功能模块,避免宏内核架构下单个模块故障或安全威胁波及影响整个操作系统。采用微内核架构可将少数稳定的核心模块置于处理器的特权模式下,其他业务功能模块放置于用户态模式下,隔离了高缺陷率的功能模块并缩小了核心模块的对外攻击表面。

### 3.2 嵌入式虚拟化技术

虚拟化技术已在IT领域广泛商用,但应用到嵌入式设备上有两个方面的挑战:首先是虚拟机的性能问题。嵌入式系统对实时性以及关键业务性能有比较高的要求,引入Hypervisor仍需要保证虚拟机的关键性能。其次,嵌入式处理器发展相对

滞后,部分处理器不具备硬件辅助虚拟化特性,需要在传统硬件上提供虚拟化解决方案。

在IT领域广泛使用I型和II型Hypervisor架构:II型架构中虚拟机管理监测控制程序(VMM)作为应用程序运行在主机操作系统之上,虚拟机的性能受主机操作系统和虚拟机监测控制程序的影响;I型架构将虚拟机直接运行在Hypervisor之上,相对而言虚拟机的性能更高。

为了保障虚拟机的性能,中兴智能装备操作系统采用更加轻量级的0型Hypervisor架构。将Hypervisor做的更“薄”:去除虚拟机迁移、容灾热备份等IT虚拟化功能特性;仅保留CPU虚拟化、内存虚拟化以及部分外设虚拟化等核心功能;对于性能要求较高的虚拟机或实时业务通过分区隔离机制保障其物理资源分配,从而

达到与物理机相近的性能指标。

针对处理器相对滞后(不支持嵌入式虚拟化特性)的问题,中兴智能装备操作系统引入半虚拟化技术。通过对客户机操作系统和底层Hypervisor的深度定制和融合,实现在普通硬件环境下支持多个半虚拟化客户机操作系统(例如Linux、Android)。具体包括了以下技术。

(1)处理器虚拟化:在定制的客户机操作系统中实现了虚拟机处理器架构,并在非特权模式下客户机操作系统与Hypervisor配合完成特权指令的模拟执行。

(2)内存虚拟化:针对半虚拟化虚拟机和全虚拟化虚拟机,分别提供基于影子页表技术和基于硬件辅助虚拟化的内存虚拟化功能,从而保障客户机虚拟机的内存访问性能。

(3)外设虚拟化:对于不同需求的场景提供外设透传技术(高性能使用场景)和外设模拟共享技术(共享使用场景)。外设模拟技术通过Hypervisor模拟多个虚拟外设,并将虚拟外设请求转交给实际物理外设,实现多个虚拟机共享物理外设功能。

### 3.3 健康管理技术

在通信、汽车、工控、航空、轨道交通等高可靠领域,对于系统的功能安全要求比较高。例如:汽车电子的ISO26262 ASIL D级要求每小时失效概率小于 $10^{-9}$ ,需要系统底层操作系统提供多种功能安全保障措施。中兴智能装备操作系统从故障隔离、运行监测控制、异常恢复等角度保障系统的持续稳定运行。

#### 3.3.1 基于分区的故障隔离技术

在高可靠领域故障隔离是解决功能安全的核心关键。汽车电子领域ISO26262标准指出避免波及影响的重要性,系统必须提供两个或多个单元发生级联失效的手段。航空航天领域ARNIC653标准提出分区隔离方式支持多种电子应用,通过分区隔

离保障单点失效不影响其他业务。

中兴通讯智能装备操作系统采用微内核架构,对于用户态的不同功能安全等级业务提供分区隔离机制。当某个业务发生故障并失效后,其他分区业务不受影响。如图4所示,具体技术包括:分区间的时间隔离机制、空间隔离机制、权限隔离、硬件隔离机制。

### 3.3.2 系统运行监测控制及多级异常处理技术

中兴智能装备操作系统提供多种不同粒度的系统运行状态监测控制功能,包括:任务级、分区级、系统级、硬件级的运行监测控制。通过不同粒度的健康监测控制功能可以及时发现不同范围的异常或告警事件,并实时监测控制关键任务的运行状态,确保其运行状态与预期一致。

当发现系统不同类型的故障后,中兴智能装备操作系统提供多级的异常处理机制。通过任务级、分区级、系统级、硬件级逐级处理异常,确保每个异常问题可以在系统最小波及、影响范围内恢复,避免恢复异常造成过大的影响。

中兴智能装备操作系统基于高可靠行业标准设计功能安全特性,及时监测控制系统的运行状态并细粒度地恢复系统运行状态,从而保障高可靠性业务持续、稳定运行。

### 3.4 智能互联技术

随着工业化和信息化的深度融合,智能化和网络化是智能装备的主要发展趋势。在智能化方面,中兴智能装备操作系统通过嵌入式虚拟化技术引入成熟的开放软件生态,并提供POSIX等基础库支撑硬实时的智能任务;在网络化方面,中兴智能装备操作系统通过分区间通信技术(IPC)打通不同智能节点间的通信通道,通过支持各个行业的互联管理协议支撑多种异构系统的互联互通互操作。嵌入式虚拟化技术已在前述

章节描述,在此重点介绍网络互联方面相关技术。

中兴智能装备操作系统采用微内核架构,所有内核功能和业务功能都以服务方式对外提供,并支持本地和远程的连接和访问,具体过程如图5所示。首先由智能应用服务业务向本地服务管理器注册,并由服务管理器发布此服务;再由本地业务通过IPC机制访问服务管理器,并由服务管理器建立服务连接实现本地访问服务功能。远程业务也同样访问服务管理器并查找到服务节点,由服务管理器建立远程服务通信通道。通过此技术实现跨节点的智能互联互通互操作功能,应用业务可透明使用本地服务或远端服务。

中兴智能装备操作系统针对不同行业需求,提供应用互联互通互操作的业务协议和安全管控协议。例如:在物联网领域的约束应用协议(CoAP)和消息队列遥测传输协议(MQTT)等应用互联协议;在汽车电子领域的AUTOSAR SOME/IP协议;在工业控制领域的IEC62351(传输层安

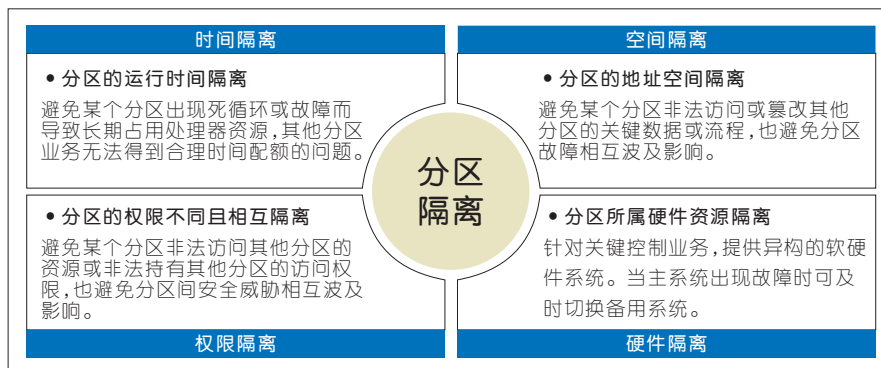
全协议(TLS))。通过上述应用互联和管控协议,实现多种异构系统的相互协同,形成了一个完整的智能工业系统。

### 3.5 安全防御技术

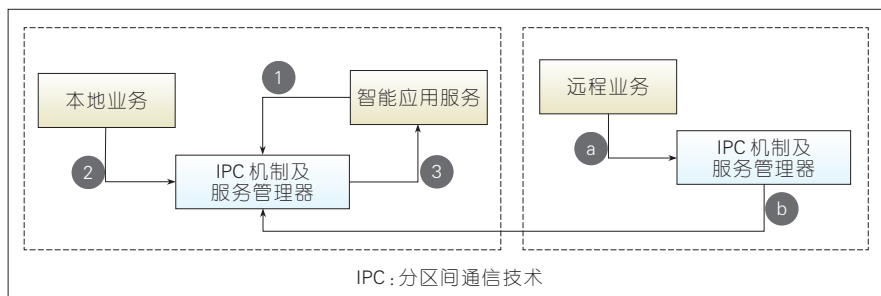
智能工业装备引入开放的软件生态以及互联网技术,也引入了安全威胁:开放的智能软件生态可能包含各种安全隐患(甚至包括未知安全威胁),互联网技术也给攻击者提供了便利的攻击途径。中兴智能装备操作系统基于可信计算的理念,设计安全可信的操作系统体系架构;同时从攻击链入手,利用底层基础软件随机多样的变化改变攻击路径,实现主动防御。利用基础软件建立智能工业装备内生的信息安全机制,无需叠加外部信息安全功能,保障了智能工业装备的实时性、确定性以及关键业务性能。

#### 3.5.1 内建的安全可信体系架构

首先,基于前述安全可信的微内核架构,建立智能工业装备操作系统



▲ 图4 分区隔离技术功能



▲ 图5 IPC过程示意



的可信计算基(TCB)。可信计算基从两个方面实现内建的安全可信:一是控制可信计算基的规模和复杂度,杜绝其自身的安全漏洞;二是缩小及加固可信计算基的对外攻击面,采用基于权限映射表的强制访问控制技术实现外部接口的安全控制功能。

其次,中兴智能装备操作系统支持内核分离保护轮廓(SKPP),通过分区隔离机制保障单个分区的安全威胁不扩散。操作系统非核心功能及智能节点业务之间都处于时间隔离、空间隔离、访问权限隔离、多核和外设隔离的运行环境中。单个分区业务故障不会影响到其他分区业务功能。

中兴通讯通过上述技术途径搭建了一个安全可信的体系架构,实现内建的可信体系架构,解决了工业装备系统的核心问题。

### 3.5.2 具有随机多样特征的移动目标防御技术

高级持续威胁(APT)是智能工业装备面临的最大的安全威胁形式,例如:伊朗震网事件、乌克兰电网事件、康明斯发动机事件等。如图6所示,高级持续威胁攻击过程包括如下4个步骤:系统探测、漏洞挖掘、系统突破、系统控制。通过对此攻击模型分析可以看出,攻击者在攻击过程依赖系统中固定的、一致性的、可预期的系统规律。例如:全局符号布局相对相对关系、关键数据结构的布局结构、关键函数的地址信息等。

中兴智能装备操作系统基于移动目标防御(MTD)的思想,打破攻击者所依赖的系统固定的、一致性的规

律。通过编译器和操作系统等底层基础软件,实现全局符号随机化、关键数据结构随机化、运行地址空间随机化、异构发布等技术。通过上述技术对整个系统引入随机、多样的变化,使攻击者无法找到系统或漏洞的规律,有效阻断攻击链。

## 4 结束语

中兴智能装备操作系统通过技术创新,解决了智能工业装备的4个核心需求:关键性能(实时性、确定性)、智能化、功能安全、信息安全。结合智能工业行业特点,提供全面的解决方案。

在电信行业:电信业务中数据面业务可直接运行于中兴智能装备操作系统,由操作系统保障业务的实时性、确定性和吞吐性能;控制面业务和云化网络功能可通过虚拟化方式运行在电信级嵌入式Linux操作系统之上,满足业务对软件生态的要求。

在工业控制行业:首先,中兴智能装备操作系统通过硬实时、高确定性,保障工控组态业务的实时性和确定性;其次,通过虚拟化技术和电信级嵌入式操作系统支持多个控制器集约化发展,降低产品成本;最后,通过安全可信的体系架构和移动目标防御技术解决工业控制系统的安全问题。

在汽车电子行业:针对车载电子,中兴智能装备操作系统利用虚拟化技术支持“一机多屏”,在一个硬件环境上同时支持仪表业务(Linux+QT)和中控台业务(Android),并满足仪表快速启动和功能安全和信息安全要求。针对车控电子,中兴智能装

备操作系统利用虚拟化引入自动驾驶软件生态,提供感知、决策的底层中间件;同时多分区的软件架构支持不同功能安全等级的业务,并兼容AUTOSAR行业基础软件标准。

综上所述,中兴智能装备操作系统满足不同行业的智能装备发展的关键需求,为中国制造2025的规划奠定坚实的软件基础。

### 参考文献

- [1] 中华人民共和国国务院. 中国制造2025[R]. 北京: 中华人民共和国国务院, 2015
- [2] 国家制造强国建设战略咨询委员会. 中国制造2025重点领域技术路线图[R]. 北京: 国家制造强国建设战略咨询委员会, 2015
- [3] Lynx Software Technologies. The Rise of the Type Zero Hypervisor[EB/OL]. (2012-09-18) [2017-10-08]. <http://www.lynx.com/the-rise-of-the-type-zero-hypervisor/>
- [4] ISO. ISO26262 Part 1[S]. Geneva: ISO, 2011: 11
- [5] Aeronautical Radio. ARINC653 Standards[EB/OL]. (2013-10-20) [2017-10-08]. [http://store.aviation-ia.com/cf/store/catalog\\_detail.cfm?item\\_id=496](http://store.aviation-ia.com/cf/store/catalog_detail.cfm?item_id=496)
- [6] AUTOSAR. AUTOSAR Standards[EB/OL]. (2017-03-31) [2017-10-08]. <https://www.autosar.org/standards/>

### 作者简介



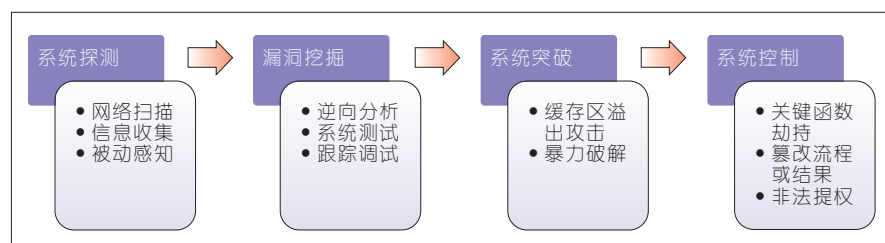
**崔云峰**, 中兴通讯股份有限公司操作系统产品部技术总工; 主要研究方向为嵌入式操作系统、网络安全等; 参与国家重大专项2项, 获得国家级、省部级奖励3项; 已申请专利10余项。



**钟卫东**, 中兴通讯股份有限公司操作系统产品部部长; 主要研究方向为程控交换机、数据通信、操作系统等; 承担多项“863”、核高基等项目; 获得国家级、省部级科技奖7项; 已申请专利20余项。



**刘东**, 中兴通讯股份有限公司操作系统产品部资深研发经理; 主要负责嵌入式操作系统产品经营与研发管理; 先后参加2项重大专项课题, 获得3项国家级科学技术与产品大奖。



▲ 图6 高级持续威胁攻击模型

# 《中兴通讯技术》第23卷总目次

卷·期·页

卷·期·页

## 卷首特稿

无线通信若干问题的思考 ..... 陆建华 23-1-02

## 专题

### 专题:NB-IoT 体系标准与应用

NB-IoT 系统物理随机接入信道设计 ..... 刘锐,戴博,杨维维 23-1-06

NB-IoT 窄带物联网系统现状与发展 ..... 张万春,陆婷,高音 23-1-10

NB-IoT 系统资源调度研究 ..... 邱刚,陈宪明,戴博 23-1-15

NB-IoT 运营策略 ..... 贾雪琴,张云勇 23-1-21

一种基于 NB-IoT 节点的多蜂窝网络最优选择方法 ..... 潘甦,陶帅,陈宇青 23-1-25

NB-IoT 运营商面临的机遇与挑战 ..... 鲁娜,朱雪田,张成良 23-1-29

基于雾计算的 NB-IoT 框架、关键技术及应用 ..... 张红,王玉峰 23-1-32

智慧能源物联网应用研究与分析 ..... 陈永波,刘建业,陈继军 23-1-37

NB-IoT 关键技术及应用前景 ..... 邹玉龙,丁晓进,王全全 23-1-43

### 专题:网络重构新技术

中国电信网络重构及关键技术分析 ..... 史凡,赵慧玲 23-2-02

中国联通网络重构与新技术应用实践 ..... 唐雄燕,曹畅 23-2-06

网络功能虚拟化的关键技术 ..... 王路,赵鹏,付乔 23-2-12

M-ICT 时代随选网络架构与实施策略 ..... 王延松,王伟忠,卢华 23-2-17

vBRAS 应用场景及评估方法探讨 ..... 雷波,解云鹏,王波 23-2-23

SDN/NFV 测试方法的研究 ..... 穆域博,马军锋,徐晓麟 23-2-27

构建网络创新试验环境的实践 ..... 毕军,胡虹雨,王扬扬 23-2-33

运营商开源策略研究 ..... 王峰,赵慧玲,杨明川 23-2-37

### 专题:5G 无线传输技术

类有线的无线接入 ..... 尹华锐,陈晓辉,卫国 23-3-02

非正交多址接入的关键技术研究 ..... 肖可鑫,夏斌,陈智勇 23-3-06

毫米波大规模 MIMO 无线传输关键技术 ..... 尤力,高西奇 23-3-11

宽带毫米波数模混合波束赋形 ..... 朱宇,李先驰 23-3-14

5G 大规模机器类通信中的传输技术 ..... 宫诗寻,陶小峰 23-3-20

基于滤波器组多载波的多点协作通信系统 ..... 宋国超,陈达,江涛 23-3-24

Filtered-OFDM 系统的次优接收机设计与验证 ..... 梁潇,武刚 23-3-29

干扰对齐技术的发展与应用 ..... 刘伟,李建东,盛敏 23-3-35

基于加权类分数傅立叶变换的变换域通信系统 ..... 梅林,房宵杰,沙学军 23-3-38

### 专题:神经网络与深度学习

机器学习在大视频运维中的应用 ..... 屠要峰,吉锋,文韬 23-4-02

基于 3D CNN 的大规模视频手势识别研究 ..... 苗启广,李宇楠,徐昕 23-4-09

基于深度学习的多目标跟踪算法研究 ..... 陆平,邓硕,李伟华 23-4-14

车辆属性识别及跨场景标注方案 ..... 董振江,高燕,吴文熙 23-4-20

领域自适应目标识别综述 ..... 唐宋,叶茂,李旭冬 23-4-25

深度神经网络的结构基础:自动编码器与限制玻尔兹曼机 ..... 康文斌,彭菁,唐乾元 23-4-32

深度学习进展及其在图像处理领域的应用 ..... 刘涵,贺霖,李军 23-4-36

### 专题:硅基光电子集成技术与器件

硅光子通信产品技术和商业化进程 ..... 孙笑晨,张琦 23-5-02

硅光子芯片工艺与设计的发展与挑战 ..... 郭进,冯俊波,曹国威 23-5-07

基于复模式匹配的半矢量硅基光波导模式求解方法 ..... 赵佳,韩林,黄卫平 23-5-11

宽带电光调制器的研究现状与新型硅基混合集成调制器的发展趋势·····李金野,于丽娟,刘建国	23-5-15
硅基马赫曾德电光调制设计优化与实现·····周林杰,周砚扬,陆梁军	23-5-21
多维复用硅基集成光子器件·····储涛,郭德汾,吴维轲	23-5-30
基于高Q值氮化硅微环反射镜的窄线宽单模光纤激光器·····章元珏,李渔,陈明华	23-5-34
25 Gbit/s CMOS 直接调制 DFB 激光器驱动电路·····石泾波,祁楠,姜培	23-5-38
硅基片上光电传感及相关器件·····周治平,邓清中	23-5-43

### 专题:移动 VR/AR 技术与应用

虚拟环境中人和虚拟角色互动的关键技术·····翁冬冬,薛雅琼	23-6-02
支持大规模视频融合的混合现实技术·····周忠,孟明,周颐	23-6-06
增强现实中的空间深度感知问题·····陈东义	23-6-10
基于虚拟现实技术的无线自组网仿真·····罗训,王荣	23-6-14
虚拟现实视频传输架构和关键技术·····郭宗明,班怡璇,谢澜	23-6-19
大视频 VR 直播业务及其技术·····徐火顺,叶剑章,尹海生	23-6-24
HEVC 兼容的全景视频运动补偿预测算法·····胡强,马文广,石志儒	23-6-28
移动增强现实应用及进展·····潘志庚,高翔,丁丹丹	23-6-32

## 专家论坛

NB-IoT 中安全问题的若干思考·····孙知信,洪汉舒	23-1-47
云计算驱动了雾计算的发展 薛育红·····	23-1-51
5G 新需求下无线网络重构的若干思考·····冯建元,冯志勇,张奇勋	23-2-41
工业互联网——重构网络架构的起点·····张恒升	23-2-45
着力突破与创新,实现超越与引领·····钟义信	23-2-47
关于移动通信终极形态的部分思考·····唐友喜,李晨兴,邵士海	23-3-45
5G 大规模接入技术:如何应对差异化服务的挑战·····张朝阳	23-3-50
深度学习的能与不能·····于剑	23-4-41

当深度学习遇到大视频数据·····曾文军,罗翀	23-4-44
关于人机对话系统的思考·····王小捷	23-4-47
硅基光子技术发展的特点、机遇与挑战·····杨建义,王根成	23-5-47
对硅基光电子技术发展的思考·····郝然	23-5-52
关于增强现实技术潜在发展方向的思考·····乔秀全,任沛,商彦磊	23-6-37

## 企业视界

M-ICT 2.0 战略——探索数字化转型之道·····苏永钦	23-1-53
M-ICT 应用发展趋势及其关键技术分析·····吕达,董振江,杨勇	23-2-50
无线网络与产品演进·····张万春,崔丽,段晓伟	23-3-53
移动通信 Small Cell 基站供电解决方案与趋势分析·····胡先红,刘明明	23-4-51
5G 承载网的需求、架构和解决方案·····李光,赵福川,王延松	23-5-56
智能工业装备操作系统技术及创新·····崔云峰,钟卫东,刘东	23-6-56

## 专家视点

封装天线技术发展历程回顾·····张跃平	23-6-41
面向沉浸式视频编码和传输的高效投影方法·····王悦名,王荣刚,高文	23-6-50

## 技术广角

一种基于安全标记的多租户访问控制方法·····彭勇,侯超平,童遥,申光	23-1-58
新一代视频业务安全解决方案·····华新海,贺镇海,刘志军	23-2-56
一种基于 Docker 的数据中心云平台实现方法及系统·····彭勇,谢剑,童遥,申光	23-2-60
一种可信身份网络架构及在互联网安全中的应用·····毛玉欣,郝振武,江家仁	23-3-58
泛在信息化智能制造系统·····于海斌,王鹏,曾鹏	23-3-62
基于运行大数据学习的复杂装备故障诊断技术及其典型应用·····刘达新,裘乐淼,王志平	23-4-56
BRAS SDN&NFV 的演进思路·····王怀滨,王延松	23-4-60
智能工厂的感知、通信与控制·····关新平,吕玲,杨博	23-5-61



# 《中兴通讯技术》杂志(双月刊)投稿须知

## 一、杂志定位

《中兴通讯技术》杂志为通信技术类学术期刊。通过介绍、探讨通信热点技术,以展现通信技术最新发展动态,并促进产学研合作,发掘和培养优秀人才,为振兴民族通信产业做贡献。

## 二、稿件基本要求

### 1. 投稿约定

- (1)作者需登录《中兴通讯技术》投稿平台:tech.zte.com.cn/submission,并上传稿件。第一次投稿需完成新用户注册。
- (2)编辑部将按照审稿流程聘请专家审稿,并根据审稿意见,公平、公正地录用稿件。审稿过程需要1个月左右。

### 2. 内容和格式要求

- (1)稿件须具有创新性、学术性、规范性和可读性。
- (2)稿件需采用WORD文档格式。
- (3)稿件篇幅一般不超过6000字(包括文、图),内容包括:中、英文题名,作者姓名及汉语拼音,作者中、英文单位,中文摘要、关键词(3~8个),英文摘要、关键词,正文,参考文献,作者简介。
- (4)中文题名一般不超过20个汉字,中、英文题名含义应一致。
- (5)摘要尽量写成报道性摘要,包括研究的目的、方法、结果/结论,以150~200字为宜。摘要应具有独立性和自明性。中英文摘要应一致。
- (6)文稿中的量和单位应符合国家标准。外文字母的正斜体、大小写等须写清楚,上下角的字母、数据和符号的位置皆应明显区别。
- (7)图、表力求少而精(以8幅为上限),应随文出现,切忌与文字重复。图、表应保持自明性,图中缩略词和英文均要在图中加中文解释。表应采用三线表,表中缩略词和英文均要在表内加中文解释。
- (8)参考文献以20条左右为宜,不允许公开发表的资料不应列入。所有文献必须在正文中引用,文献序号按其在文中出现的先后次序编排。常用参考文献的书写格式为:
  - 期刊[序号]作者.题名[J].刊名,出版年,卷号(期号):引文页码.数字对象唯一标识符
  - 书籍[序号]作者.书名[M].出版地:出版者,出版年:引文页码.数字对象唯一标识符
  - 论文集中析出文献[序号]作者.题名[C]/论文集编者.论文集名(会议名).出版地:出版者,出版年(开会年):引文页码.数字对象唯一标识符
  - 学位论文[序号]作者.题名[D].学位授予单位所在城市名:学位授予单位,授予年份.数字对象唯一标识符
  - 专利[序号]专利所有者.专利题名:专利号[P].出版日期.数字对象唯一标识符
  - 国际、国家标准[序号]标准名称:标准编号[S].出版地:出版者,出版年.数字对象唯一标识符
- (9)作者超过3人时,可以感谢形式在文中提及。作者简介包括:姓名、工作单位、职务或职称、学历、毕业于何校、现从事的工作、专业特长、科研成果、已发表的论文数量等。
- (10)提供正面、免冠、彩色标准照片一张,最好采用JPG格式(文件大小超过100kB)。
- (11)应标注出研究课题的资助基金或资助项目名称及编号。
- (12)提供联系方式,如:通信地址、电话(含手机)、Email等。

### 3. 其他事项

- (1)请勿一稿两投。凡在2个月(自来稿之日算起)以内未接到录用通知者,可致电编辑部询问。
- (2)为了促进信息传播,加强学术交流,在论文发表后,本刊享有文章的转摘权(包括英文版、电子版、网络版)。作者获得的稿费包括转摘酬金。如作者不同意转摘,请在投稿时说明。

编辑部地址:安徽省合肥市金寨路329号国轩凯旋大厦1201室,邮政编码:230061

联系电话:0551-65533356,联系邮箱:magazine@zte.com.cn

本刊只接受在线投稿,欢迎访问本刊投稿平台:tech.zte.com.cn/submission

办刊宗旨

以人为本,荟萃通信技术领域精英;  
迎接挑战,把握世界通信技术动态;  
立即行动,求解通信发展疑难课题;  
励精图治,促进民族信息产业崛起。

---

双月刊 1995 年创刊 总第 137 期  
2017 年 12 月 第 23 卷 6 期

主管:安徽省科学技术厅  
主办:安徽省科学技术情报研究所  
中兴通讯股份有限公司  
编辑:《中兴通讯技术》编辑部

总编:陈杰  
常务副总编:黄新明  
责任编辑:徐烨  
编辑:卢丹,朱莉,赵陆  
排版制作:余刚  
发行:王萍萍  
编务:王坤

---

《中兴通讯技术》编辑部  
地址:合肥市金寨路 329 号凯旋大厦 12 楼  
邮编:230061  
网址:tech.zte.com.cn  
投稿平台:tech.zte.com.cn/submission  
电子信箱:magazine@zte.com.cn  
电话:(0551)65533356  
传真:(0551)65850139

出版、发行:中兴通讯技术杂志社  
发行范围:全球发行  
印刷:合肥添彩包装有限公司  
出版日期:2017 年 12 月 10 日  
中国标准连续出版物号:ISSN 1009-6868  
CN 34-1228/TN  
定价:每册 20.00 元,全年 120.00 元