

VR的技术发展趋势和行业应用

Trends and Industry Applications of VR Technology

尹芹/YIN Qin

吕达/LYU Da

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

1 VR的概念

1.1 VR的作用

1965年,虚拟现实(VR)的概念被首次提出。2016年,Oculus、HTC VIVE、PS VR等各种VR硬件层出不穷地面世,该年也被称为VR元年。受限于制作精良、吸引用户眼球并可持续使用的VR内容,受限于带宽以及更高清分辨率的显示技术,VR仍然处在爆发前夜。

VR可用于视频直播、视频点播、视频游戏,以及垂直行业的多个应用场景,拉近人与人、人与物之间的距离,带给用户沉浸、互动的感受。

2017年,中兴通讯推出端到端的VR直播解决方案,并成功应用于南艺520直播。

1.2 VR的挑战

(1)带宽挑战。运营商开展VR视频业务之后,对带宽的需求约为300 Mbit/s~1.2 Gbit/s,最高可能将会超过1 Gbit/s,因此千兆带宽及至5G网络将会给用户带来更佳服务体验,

收稿日期:2018-06-18

网络出版日期:2018-07-31

基金项目:国家高技术研究发展(“863”)计划(2014AA015200)、国家科技支撑计划(2012BAH02B03)

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868(2018)03-0054-004

摘要: 提出虚拟现实(VR)面临的2种技术发展挑战:高带宽、低时延。围绕VR的发展,中兴通讯创新性地提出端到端VR系统方案的技术理念,研发出一系列创新技术用于实现VR业务,例如:低码高清技术、基于视野(FOV)的自适应传输技术、VR电子节目指南(EPG)、VR机顶盒、VR播放器,并基于网络功能虚拟化(NFV)技术构建出云化VR网络架构。这些技术和方案进一步地促进了VR的发展。

关键词: 低码高清; FOV; 基于FOV的自适应传输

Abstract: In this paper, two kinds of technical development challenges faced by virtual reality (VR) are proposed: high bandwidth and low delay. Focusing on the development of VR, ZTE innovatively proposed the technical concept of an end-to-end VR system solution. A series of innovative technologies were developed to implement VR services, such as low-code high-definition technology and angle of view (FOV)-based adaptation transmission technology, VR electronic program guide (EPG), VR set-top box, VR player, and cloud-based VR network architecture based on network function virtualization (NFV) technology. These technologies and solutions further promote the development of VR.

Keywords: low-code high-definition; FOV; FOV-based adaptive transmission

内容分发网络(CDN)也将为运营商节约更多带宽消耗。

(2)时延和丢包挑战。VR要求运动到图像的最大时延在20 ms,运动到声音的最大时延在20 ms,并要求音视频保持同步。

(3)完美拼接。通常需要在一组摄像机设备上采集,然后再进行拼接处理,将来自不同摄像机的视图合并到一个视图中。为保障完美的VR体验,不应引入任何拼接错误,不应看到任何拼接线。

1.3 VR的演进过程

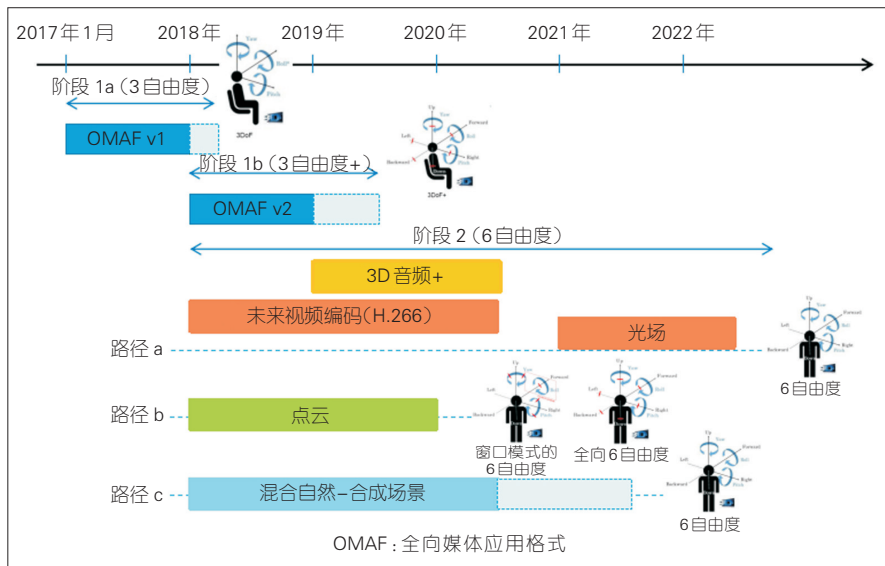
动态图像专家组(MPEG)、第3代合作伙伴项目(3GPP)、数字音视频编解码技术标准(AVS)、中国通信标准化协会(CCSA)等多个全球标准组织正在积极推进和制订VR相关标准。其中,MPEG国际标准组在2015

年10月启动全向媒体应用格式(OMAF)项目,主要针对360视频以及对应的音频,研究相应的文件封装格式,以及在基于HTTP的动态自适应流(DASH)方式下的传输,同时还包括编码的配置、视点的投射等。MPEG OMAF之后,MPEG标准组织认为有必要成立MPEG-I,并包含OMAF。MPEG-I标准的工作时间轴如图1所示。

MPEG-I的标准制定,又分为3个阶段^[1]:阶段1a、阶段1b、阶段2,具体如下表1所示。

阶段1a(3自由度):在特定观察位置,当头部左右旋转、俯仰旋转、摇摆旋转时,VR头显能正确显示相应VR内容,需要VR内容、VR采集和VR显示设备的支持。

阶段2(全向6自由度):在特定观察位置,当头部左右旋转、俯仰旋



▲图1 MPEG-I标准工作的不同阶段

▼表1 MPEG-I的不同阶段及说明

阶段	自由度	说明
1a	3自由度	围绕X轴、Y轴、Z轴的3次旋转和非限制运动(分别为俯仰、偏航、翻滚)
1b	3自由度+	在3自由度的基础上,沿X轴、Y轴、Z轴额外、受限的平移运动(典型情况下,头部左右、上下、前后移动)
2	(内容)窗口模式的6自由度	围绕X轴、Y轴(分别为俯仰、偏航)进行受限的旋转运动,并沿Z轴进行受限的平移运动
	全向6自由度	沿着X轴、Y轴、Z轴进行受限的平移运动(典型情况下,步行距离为几步)
	6自由度	在3自由度的基础上,沿着X轴、Y轴、Z轴进行完全平移(左右、上下、前后自由移动)

转、摇摆旋转,以及一定范围内向前后、左右、上下3个方向平移时,VR头显能正确显示相应VR内容,需要VR内容、VR采集和VR显示设备的支持。

阶段2(6自由度):典型的使用案例是用户自由穿过头戴式显示器(HMD)上显示的3D 360 VR内容(物理地或通过专用的用户输入装置)。

1.4 VR端到端架构

3GPP定义的VR视频架构^[2]由采集、球面拼接(可选)、投影(可选)、封包(可选)、编码、封装、传输、解封装、解码、渲染、显示各环节组成,如图2所示。

各细分环节的详细描述,参见如

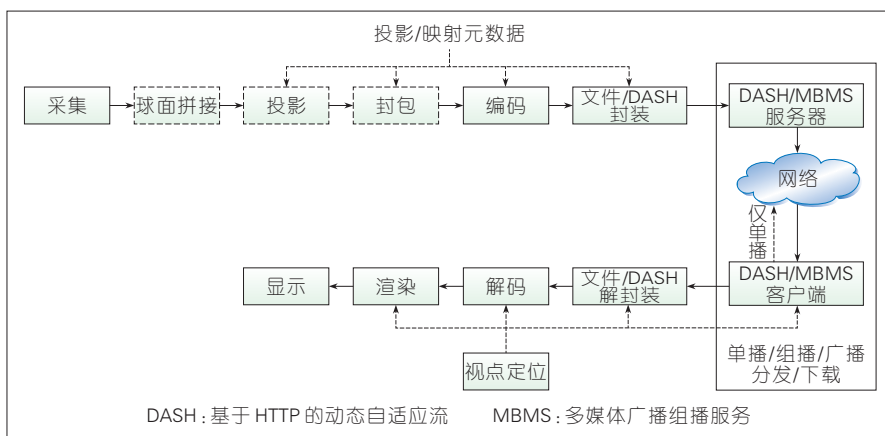
下各章节。

1.4.1 采集

取决于采集系统的功能,VR内

1.4.2 球面拼接

球面拼接分为3种:基于镜像的



▲图2 3GPP虚拟现实视频架构

容以不同的格式表示,例如:全景图或球体。许多系统采集覆盖整个360°×180°球体的球形视频。通常需要配置多个相机来采集这样的内容。配置的各种相机可用于记录2D和3D内容。采集环节可以使用立体相机组、鱼眼、广角镜头、相机阵列、光场相机阵列,产生立体效果或光场渲染效果的VR内容。

(1)立体效果

(a) 分片立体效果:立体相机组采集3D内容,并以相对较小的重叠排列成星形配置;但这样的照相机系统一方面可能会有视差错误,另一方面基于镜像的系统可以使用减少视差错误的相机组采集3D图像。

(b) 极度重叠的立体效果:立体内容由鱼眼或广角镜头、相机阵列采集的重叠图像创建。在处理过程中,每个图像传感器分成左右2部分,并拼接成左右全景图。

(2)光场渲染

(a) 基于深度的光场渲染:是创建3D内容的有前途的方法;但是,光场渲染需要密集的相机阵列。

(b) 使用基于深度的渲染来生成中间相机视图:这是现有方法,可以减少所需相机的数量;但该方法需要非常精确的深度图和复杂的基于深度的处理流程,增加了所需算力。

系统直接拼接、深度感知的拼接(分片立体效果,极度重叠的立体效果)、深度使能的光场渲染。

1.4.3 投影

最常用的VR投影方法包括2种:经纬度展开投影(ERP)、立方体投影(CMP)。

(1)ERP:水平、垂直坐标分别对应经度、纬度,不变换、不缩放。该方法的特点是:赤道上的像素拉伸最小(或一点都没有拉伸),而越向两极拉伸越严重,因此失真越严重。这就产生了大量冗余信息,不适合使用高效视频编码(HEVC即H.265)对其进行压缩。

(2)CMP:将球面全景图像映射到了立方体的6个面上,中间的4个面为前后左右的图像,上下3个面为顶部和底部的图像,每个面都有 $90^\circ \times 90^\circ$ 的视野(FOV)。在立方体投影中,直线保持笔直,便于对图像进行处理,相比于圆柱映射,失真要小很多。好处在于:减少了两极的冗余,减少了数据量;立方体投影中直线保持笔直,没有发生弯曲失真(这对于视频编码来说十分重要,因为运动矢量是直线);立方体投影对像素进行了分配,两极和侧面都是一样的;立方体投影的映射更加简单,只要将每一个面贴到对应的立方体面上即可。

投影方法除了以上这2种以外,根据用于渲染的几何类型,还有一些其他投影类型,如:球、压扁的球、圆柱体、柏拉图固体(正多面体)、立方体(6面)、八面体(8面)、二十面体(20面)、截断的金字塔、分段球体、直接鱼眼(用于视频监控控制等)。

1.4.4 区域映射(封装)

在投影之后,所获得的二维矩形图像可被分割成可重新排列以产生“封装”帧的区域。从投影帧产生封装帧的操作(表示为“封装”或“区域映射”)可能包括平移、缩放、旋转、填充、仿射变换等。执行区域映射的

原因包括:提高编码效率或依赖视点的流管理(详见多流方法)。

区域映射是可选过程,如果未使用区域映射,则封装的VR帧与投影帧相同。

1.4.5 编码&解码

目前的360视频服务提供了非常有限的用户体验,因为用户视点的分辨率、视觉质量与传统视频服务差不多。需要多倍于现有超高清(UHD)分辨率的分辨率,才能以足够清晰的分辨率覆盖完整的 360° 环境。这对现有已建立的视频处理流程链、现有的终端设备都构成了重大挑战。360视频传输主要考虑3种解决方案:单流方式、多流方式、分片式流方法。

1.4.6 文件/DASH封装/解封装

如果DASH用于传送360视频,则可能需要额外的信令,例如:投影和映射格式可能需要在媒体呈现描述(MPD)信令上体现,以便用户可以请求合适的不同码率、不同码率的码流,然后根据考虑的解决方案类型(单流、多流、分片式流)执行不同的文件/DASH封装。

接收器可以根据当前的视点位置、设备能力(例如:视频解码器能力),选择仅解封装所接收的视频流的子集。

2 VR行业应用

为解决用户体验、VR成本等问题,Cloud VR通过云端渲染为VR发展提供最佳的解决方案。

Cloud VR场景^[9]发展分为近、中、远期3个阶段,Cloud VR巨幕影院、Cloud VR直播、Cloud VR 360视频、Cloud VR游戏4个场景处于近期阶段;Cloud VR教育、医疗、营销、大空间竞技、健身、音乐、K歌场景处于中期阶段;Cloud VR旅游、社交、购物、军事、工程、房地产等场景处于远期阶段。十七大场景的商业潜力空间巨大,但离不开运营商的牵头。电信

运营商具备规模发展Cloud VR业务的必需条件,通过大管道、平台基础、用户、接入光纤资源等优势吸引优质的VR服务提供商、终端厂商、内容商等加入生态链,不仅可以赢得VR行业浪潮的商业先机,还可以带动整个VR产业的发展。运营商可以先发展Cloud VR 2C市场,再聚合行业应用,扩展到2B市场。

3 中兴通讯VR技术创新

3.1 低码高清技术创新

中兴通讯创新地提出低码高清、低码超高清技术,对H.264和H.265核心算法提出了5个方面的改进,在保证主观质量的情况下编码码率降为原来的40%左右,为高清和4K超高清视频业务的发展扫清了接入带宽方面的障碍,具体包括:

- 提出帧级码率控制方法,有效控制瞬时码率过高;
- 提出基于宏块距离的空间滤波方法,提高图像质量;
- 提出恰可察觉失真(JND)和人类视觉的感兴趣区域(ROI)相结合的分级量化编码方法,对在相同或降低码率的情况下,大幅提升视频编码的主观质量;
- 提出色度分量策略性编码方法,有效提升视频压缩效率,且视频主观质量不会有明显下降;
- 提出非ROI宏块采用变换域JND进行预测残差自适应调整编码方法,从而既保证了编码的主观质量,又降低了编码码率。

低码高清、低码超高清算法是一种复杂的综合算法,经过不同片源的综合评测,成为有效的图像算法,既可以应用在H.264的框架下,也可以融合在H.265编码标准下。

3.2 基于视点FOV自适应传输技术创新

中兴通讯融合CDN支持基于视点FOV自适应传输技术。

该技术的原理为:人眼视野范围内不同区域的敏感度不同,35°范围外不敏感;而FOV显示模式可以针对不同区域给予不同分辨率的图像。当视点发生变动时及时切换到对应的内容频道。

中兴通讯融合CDN能支持对各种自适应码率格式,扩展偏航角、俯仰角、主视场的水平视角,主视场的垂直视角等信息,根据用户头部运动,传输用户感兴趣的、基于视点的内容,有效降低带宽,缩短时延。

3.3 多场景的VR视频播放技术创新

(1) 机顶盒+头盔

机顶盒作为计算中心,提供图形计算、渲染等能力;头显提供VR呈现,具备九轴传感器,提供位置等信息给机顶盒;机顶盒和头显之间通过高清晰多媒体接口线(HDMI)和USB连接,HDMI传输VR视频,USB传输传感器信号。

(2) 手机+头盔

手机作为计算中心,提供图形计算、渲染等能力;手机屏幕提供VR呈现,提供位置、角度等信息给手机;通过VR眼镜的自带触控板或遥控器,便捷用户操作。中兴通讯提供VR页面、VR EPG能力、VR播放器能力支持VR点播业务。

(3) 裸眼收看VR业务

手机作为计算中心,提供图形计算、渲染等能力;手机屏幕提供VR呈现,提供位置、角度等信息给手机;通过VR眼镜的自带的触控板或遥控器,便捷用户操作。

4 结束语

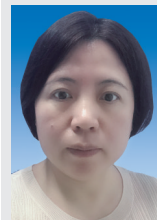
VR的发展过程是视频领域的一个技术创新的过程,中兴通讯创新地提出端到端的VR架构理念,形成了FOV视点自适应传输、低码高清、VR EPG、VR机顶盒、VR播放器等一批新技术并成功应用在VR系统中,从

而推动VR的全面发展。

参考文献

- [1] MPEG.MPEG-I Part 1 Technical Report on Architectures for Immersive Media[S].2017
- [2] 3GPP.Virtual Reality (VR) Media Services over 3GPP: 3GPP TR26.918[S].2018
- [3] 中国信息通信研究院.“Cloud VR+”场景白皮书[R].2018

作者简介



尹芹,中兴通讯股份有限公司产品总经理;主要研究领域为大视频、大数据、人工智能、云计算、固网、5G、网络安全等;先后主持和参加国家级基金项目5项,荣获省部级以上科学技术奖4项;已发表多篇论文,拥有发明专利30余项。



吕达,中兴通讯股份有限公司有线研究院副院长;先后主持和从事程控交换机、软交换、电信增值业务、数字家庭和基于IP的大视频解决方案的研发和管理工作,尤其是对于多媒体技术发展趋势和业务应用有深入了解;多次获得省部级奖;已发表论文5篇。