

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^[1]。

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^[2-3]。

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

收稿日期: 2017-09-28
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) 06-0024-004

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

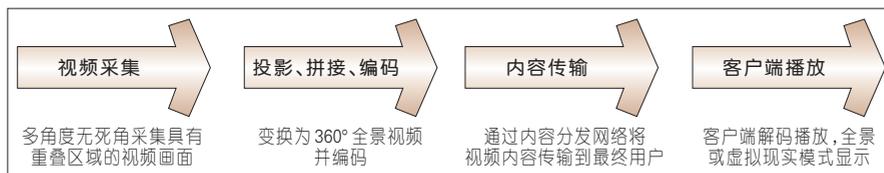
设备进行观看直播, 体验沉浸感和现场感^[4-5]。

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。

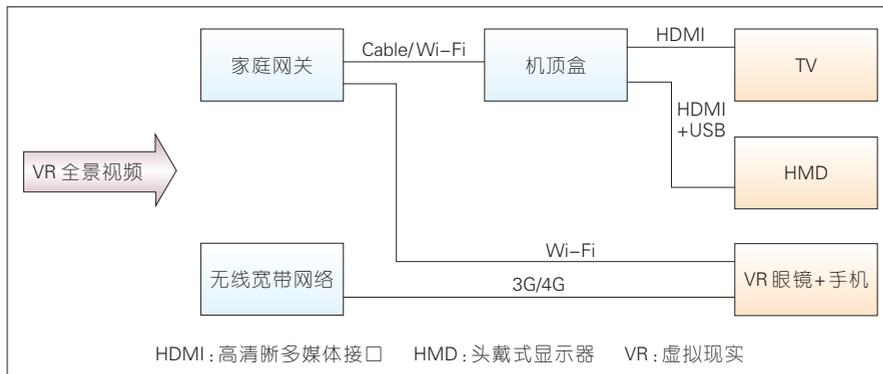
VR 直播可以通过多种方式呈现给最终用户, 在家中可以通过家庭网关, 利用机顶盒显示在电视上 VR 视

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



▲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° 、垂直 180° ,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^[6-7]。

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^[8]。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^[9-10]。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° 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° 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”^[11-12]。

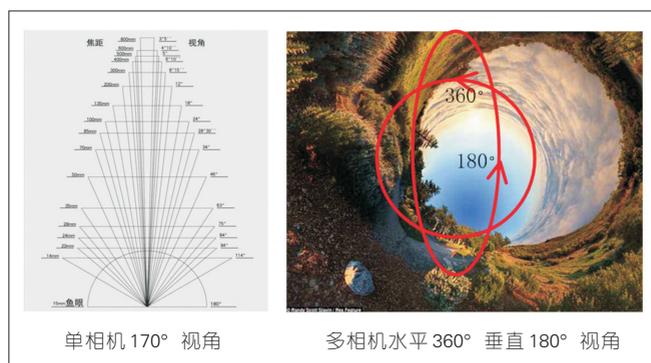
2.5 投影映射及编码技术

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2~4K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4K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4K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12K分辨率,一般直播使用4K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满意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系统集成实时流传输协议(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、例行的工作,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

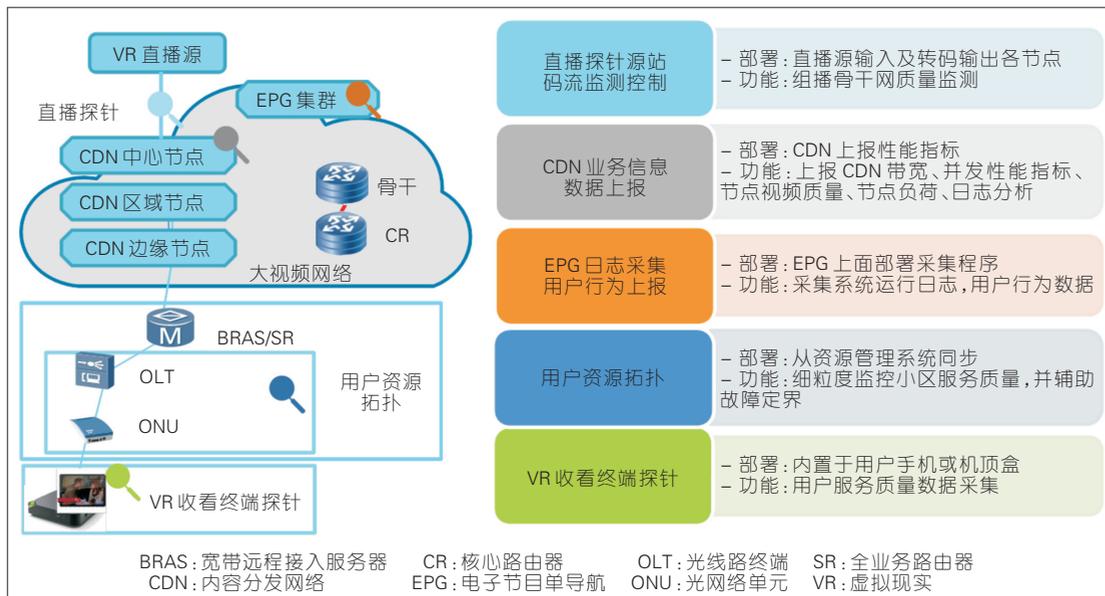


图5
大视频 VR 直播全流程
质量监测控制体系

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程师和黄珂资深工程师, 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。