



专题: 5G

## 面向通信与计算融合的 5G 移动增强/虚拟现实

周一青<sup>1,2,3</sup>, 孙布勒<sup>1,2,3</sup>, 齐彦丽<sup>1,2,3</sup>, 彭燕<sup>1,2,3</sup>, 刘玲<sup>1,2,3</sup>,

张志龙<sup>4</sup>, 刘奕彤<sup>4</sup>, 刘丹谱<sup>4</sup>, 李兆歆<sup>1,2,3</sup>, 田霖<sup>1,2,3</sup>

(1. 中国科学院大学, 北京 100049; 2. 中国科学院计算技术研究所, 北京 100080;  
3. 移动计算与新型终端北京市重点实验室, 北京 100190; 4. 北京邮电大学, 北京 100876)

**摘要:** 面向通信与计算融合的 5G 移动通信网络, 剖析移动 AR/VR 信息处理和传输的特征, 提出融合通信与计算的能力, 在未来 5G 移动通信网络中采用多级计算, 通过协同网络多级计算节点的能力来解决移动终端计算能力有限的问题; 采用智能传输机制, 通过高效频谱感知、分层编码、空口自适应传输等来克服移动通信信道传输能力不稳定的问题; 采用时延保障机制, 通过通信与计算资源协同管理来确保移动 AR/VR 服务时延。对移动 AR/VR 多级计算模型、智能传输机制和服务时延保障的研究现状和发展趋势进行综述, 总结现有研究存在的问题, 并提出下一步研究的方向。

**关键词:** 通信与计算融合; 5G; 增强/虚拟现实; 多级计算

**中图分类号:** TN911

**文献标识码:** A

**doi:** 10.11959/j.issn.1000-0801.2018241

## Mobile AR/VR in 5G based on convergence of communication and computing

ZHOU Yiqing<sup>1,2,3</sup>, SUN Bule<sup>1,2,3</sup>, QI Yanli<sup>1,2,3</sup>, PENG Yan<sup>1,2,3</sup>, LIU Ling<sup>1,2,3</sup>,

ZHANG Zhilong<sup>4</sup>, LIU Yitong<sup>4</sup>, LIU Danpu<sup>4</sup>, LI Zhaoxin<sup>1,2,3</sup>, TIAN Lin<sup>1,2,3</sup>

1. University of Chinese Academy of Sciences, Beijing 100049, China

2. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China

3. Beijing Key Laboratory of Mobile Computing and Pervasive Device, Beijing 100190, China

4. Beijing University of Posts and Telecommunications, Beijing 100876, China

**Abstract:** Oriented to 5G mobile communication network combining with communication and computation, the characteristics of mobile AR/VR information processing and transmission were analyzed, and the ability to integrate communication and computation was proposed, and the capability of integrating communication and computing was proposed. In the future 5G mobile communication network, multi-level computing will be used to solve the problem of limited computing capacity of mobile terminals by the capability of multi-level computing nodes in collaborative network. Intelligent transmission mechanism could overcome the problem of unstable transmission capacity of mobile channel by efficient spectrum sensing, layered coding and adaptive space port transmission, and the mobile AR/VR service delay was ensured by using delay guarantee mechanism and cooperative management of communication and computing resources. The research status and development trend of mobile AR/VR multi-level computing model, intelligent transmission mechanism and service delay guarantee was reviewed, the existing problems were summarized and the next research direction was put forward.

**Key words:** convergence of communication and computing, 5G, AR/VR, hierarchical computing



## 1 引言

移动通信的发展日新月异,目前 5G 的研发已拉开大幕。相比 4G, 5G 将在数据传输速率、传输时延、网络容量等多个方面实现飞跃性突破: 峰值速率将达到 10 Gbit/s; 支持超低时延超高可靠的服务, 业务时延小于 5 ms; 联网移动设备数量增加到现在的 100 倍, 网络容量将提升 1 000 倍<sup>[1]</sup>。这些性能的突破增强了 5G 的智能业务服务能力。高通、ABI 等公司指出移动增强现实/虚拟现实 (augmented reality/virtual reality, AR/VR) 将成为 5G 的第一波杀手级应用<sup>[2]</sup>。

AR/VR 以计算技术为核心,生成逼真的视觉、听觉等,构成一定范围内的虚拟环境,用户可以与虚拟环境中的物体交互,获得身临其境的感受和体验。以 360°全景 VR 视频为例,其处理和传输流程如图 1 所示。为了增强用户的体验,首先用摄像机拍摄超高清分辨率的画面,然后拼接成 360°全景画面,给用户选择观看视角的自由,让用户具有身临其境的视觉体验。生成的 AR/VR 可缓存于服务器端,当用户端发出请求时,可通过有线或者无线传输将相应的视频提供给用户。显然,由于超高清分辨率和全景画面需要的多角度信息,生成的 AR/VR 比普通视频,信息量可能高出几十倍,从而对系统的处理和传输能力提出

很高的需求。另一方面,在用户端,当朝向、视角等状态发生变化时,系统可基于传感器等,运用动作捕捉技术,追踪用户行为并完成与虚拟场景的实时交互,实现更极致的互动体验。即用户发出新的画面需求,系统进行实时响应,在 20 ms 内将相应的画面提供给用户。由于用户的状态是随机改变的,所对应的画面通常是难以预存的,需要根据已有画面和变化的用户请求,渲染绘制出新内容,再提供给用户。渲染是一个复杂的综合性任务,通常需要借助深度信息计算、图像语义理解(也称作图像语义分割)等计算密集型处理来合成以假乱真的内容。可见,AR/VR 具有海量信息、密集计算的特点,它们对移动网络提出了大带宽、低时延的服务需求,为 5G 及未来移动网络的发展带来新的挑战 and 机遇。

AR/VR 在发展的早期多用于军事、航空航天、工业仿真等行业领域。近几年来,随着 Oculus Rift 等消费级 VR 产品的推出,AR/VR 逐渐向个人应用领域蓬勃发展。目前的 AR/VR 主要基于个人电脑 (personal computer, PC),交互设备如头戴显示设备(以下简称头显)采用有线的方式与 PC 直连,计算处理任务由 PC 完成并通过连接线传回头显。受限于有线连接,用户不能自由活动,极大地影响了 AR/VR 的体验效果。因此,通过无线方式连接的移动 AR/VR 成为近年的发展焦点。2017 年 HTC

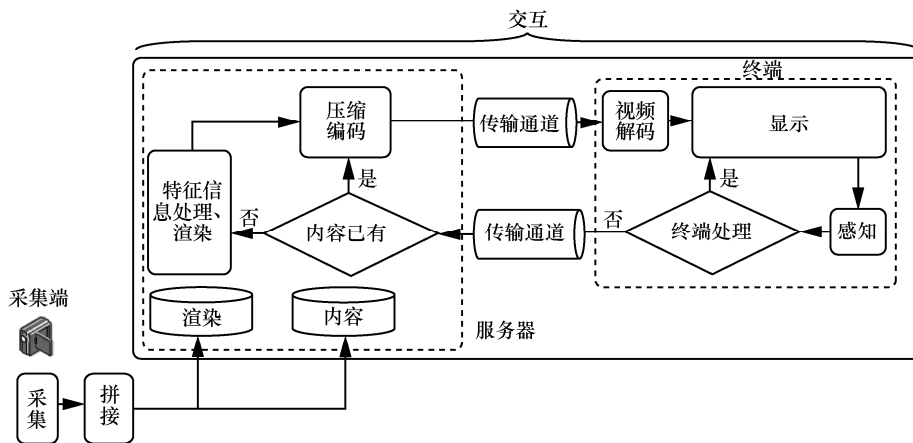


图 1 360°全景 VR 视频的处理和传输流程

发布了 VIVE 无线一体机, 将头盔与 PC 之间的多根数据线升级为无线方式连接, 但仍然需要配套一台高性能的 PC, 成本高昂。而以谷歌 Daydream 和三星 Gear VR 为代表的联合手机终端的头显则用手机终端提供显示、通信和简单的计算功能, 头显提供封闭的环境、镜头和交互功能。通过手机终端进行联网互动的同时将复杂的 AR/VR 处理任务传输到云端处理并将结果传回手机终端显示给用户。这种服务模式不需要配套性能强劲的 PC 来实现复杂任务的处理, 能以低廉的价格提供移动 AR/VR 服务, 正在逐渐成为移动 AR/VR 业务的主流实现方式。根据德意志银行的报告, 未来几年移动 AR/VR 产品将赶超基于 PC 的 AR/VR 产品<sup>[3]</sup>。

## 2 移动 AR/VR 面临的挑战

目前基于手机终端的移动 AR/VR 仅能提供简单有限的体验, 整体效果差强人意。如前所述, 移动 AR/VR 信息的处理过程中, 涉及渲染等高复杂度的密集型计算任务, 而手机终端自身的计算能力有限, 因此在终端处理任务的比例非常低, 大部分任务都需要通过移动互联网传送到云端服务器计算并传回给用户, 时延大。另一方面, 在移动 AR/VR 信息的传输过程中, 受限于移动通信的传输带宽以及信道质量不稳定等问题, 导致用户接收到的视频画面质量不稳定、不流畅, 体验效果差。因此, 移动 AR/VR

的信息处理和传输面临众多挑战, 需要解决图 2 所示的三大矛盾。

(1) 移动 AR/VR 处理的密集计算需求与有限的移动终端计算能力之间的矛盾

移动 AR/VR 具有的海量信息导致处理复杂度高, 对系统的密集计算能力提出了很强的需求。以典型的多视角三维场景重建为例, 在较低分辨率的摄像头下支持 320 像素×240 像素, 在约 1 m×1 m×1 m 真实环境空间中使用 256 体素×256 体素×256 体素的体素解析度进行计算, 通过图形处理器 (graphics processing unit, GPU) 并行处理也只能达到 15 帧/s 左右的交互式帧率 (电脑配置: 计算机处理器 i7, 32 GB RAM, NVIDIA Titan black 显卡, 6 GB 显存); 若提升摄像头分辨率, 且在更大的真实环境中进行应用, 那么需要采用更高密度的体素辅助计算, 如果体素解析度提高至 1 024 体素×1 024 体素×1 024 体素, 计算量又将提高 64 倍左右。相比 PC 配置, 移动终端虽然逐步配备了多核处理器和图形处理核等, 其计算能力还存在明显差距, 难以满足移动 AR/VR 密集计算的需求。若将计算迁移至远端云服务器进行, 则其时延性能难以保障。值得注意的是, 虽然移动终端本身计算能力有限, 如图 2 所示, 已有研究<sup>[4]</sup>指出, 未来移动网络将在基站、核心网等网络的不同层次融合移动边缘计算 (mobile edge computing, MEC), 成为一个通信与

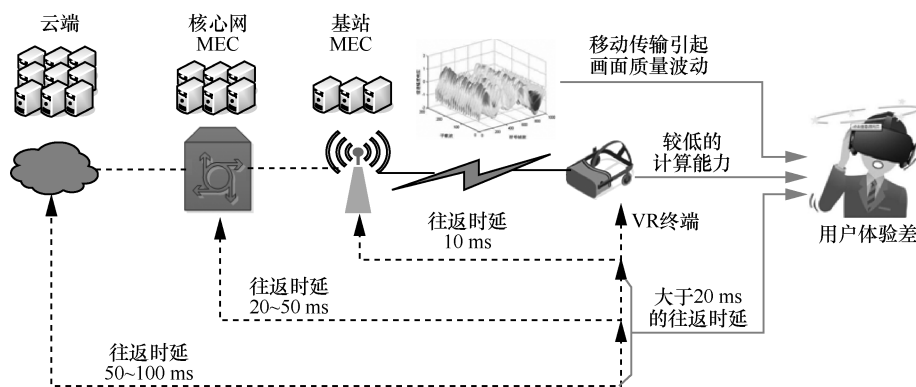


图2 移动 AR/VR 处理与传输问题分析



多级计算协同融合的网络。若能有效协同移动终端与未来 5G 移动网络的多级计算能力,有可能完成 AR/VR 的密集计算需求。

(2)移动 AR/VR 的高画质需求与随机时变的移动传输能力带来的画质不稳定之间的矛盾

为了增强用户的体验,需要提升移动 AR/VR 的清晰性、流畅性,因此对传输能力也提出了大带宽、高质量的需求。相比有线通信,采用无线传输的移动通信可以在任何时间、任何地点为任何人提供服务,这种服务的自由性提升了用户的体验,但无线传输信道不光带宽有限,而且其信道质量随着时间和地点不断动态变化,这种随机的时变传输能力为 AR/VR 传输带来了很大的问题。一方面,AR/VR 视频流具有大象流特征,对带宽需求特别高;另一方面,鉴于 AR/VR 的交互特性,用户对视频的需求存在强烈的随机性和突发性,即便是看同一个 AR/VR 视频,不同的用户视角不同,也会发出不同的画面请求,进一步增大传输带宽的需求。当用户发起新的画面需求时,如果当时信道质量较差,只能传输低质量画面,引发的视觉疲劳会带来眩晕感,如果以低速传输高质量画面,可能无法及时响应用户需求,导致时延过大或者画面停顿,同样严重影响用户体验。这个问题在通信与多级计算融合的未来无线网络中有可能得到解决。未来无线网络引入的 MEC 服务器,具有大数据计算和存储的能力。一方面可以通过无线大数据挖掘,获得频谱地图,当移动信道质量出现下降时,快速提供可切换或可扩展的频谱,提升传输能力。另一方面,正如已有研究<sup>[5]</sup>指出的,协同利用通信与计算资源,有可能设计智能高效的移动传输机制,完成传统移动通信网络中难以传输的移动 AR/VR。因此,解决第二个矛盾需要针对移动 AR/VR 和随机时变的移动无线信道的特征,研究通信与计算协同的智能传输机制,确保移动 AR/VR 传输画面的质量和流畅性。

(3)移动 AR/VR 的低时延需求与现有移动网络服务时延过大之间的矛盾

AR/VR 业务一般需要低至 20 ms 的服务时延来保障其自然、流畅的体验,而极致的体验,需要进一步降低服务时延。服务时延包括业务的传输时延和处理时延。如图 2 所示,传输时延部分,现有 4G 网络的移动终端与基站空口传输往返时延(round trip time, RTT)近似为 10 ms;如果需要云端的计算能力,由于终端到云端的数据传输需要经过网关等设备经过多重转发,往返时延一般为 50~100 ms,远大于移动 AR/VR 要求的 20 ms 的服务时延要求,严重影响用户的体验。另一方面,移动 AR/VR 的处理时延由计算节点的计算能力及业务的计算量决定。目前移动终端计算能力较弱,而云服务器计算能力则非常强大,故相同计算量的业务在移动终端和云服务器的处理时延差异较大;同时,移动 AR/VR 的海量数据也带来了非常大的计算量,以 4K RGB (red-green-blue) 360°全景视频为例,其全视角分辨率需达到 7 680×3 840 像素,是普通 720P 视频的 32 倍,处理和传输时延都将大幅增加。可见,目前移动网络以移动终端协同远端云服务的方式提供 AR/VR 业务,很多时候仅传输时延就超出了 AR/VR 的时延限制。但如果考虑未来通信与多级计算融合的移动网络,一方面 5G 的空口时延将降低到 1 ms 左右,另一方面有可能通过协同 MEC 边缘计算节点完成 AR/VR 的密集计算任务,将大幅降低传输时延。而 MEC 服务器相比云服务器,计算资源相对有限,其处理时延不可忽略。因此,针对移动 AR/VR 的需求,基于未来移动网络中的传输和处理时延模型,对通信及计算异质资源的分配进行优化,有可能降低总体服务时延,满足移动 AR/VR 服务的时延需求。

综上所述,为了解决当前移动 AR/VR 面临的三大矛盾,面向未来 5G 移动网络,以通信与计算融合为核心思想,应紧密结合移动 AR/VR 业务特

征与需求,研究未来 5G 移动网络的多级计算网络模型及移动 AR/VR 信息处理的多级协同计算方法,研究多级计算架构下移动 AR/VR 的智能传输、服务时延模型及时延优化与保障机制,提出基础理论和核心机制,解决移动 AR/VR 信息处理与传输的关键问题。下文将对移动 AR/VR 的“计算模型”“智能传输”和“时延保障”3 个方面的研究现状及发展动态进行综述。

### 3 AR/VR 信息处理计算模型与方法

AR/VR 信息处理过程中,需要实时完成场景深度估计、图像语义理解、三维场景重建、高真实感渲染等密集计算型任务,以保障用户可以获得自然、流畅的体验。目前 AR/VR 的信息处理研究主要是通过并行计算、分布式计算完成密集计算型任务。此外,针对移动终端有限的计算能力难以支持移动 AR/VR 密集计算任务的问题,研究人员也提出了终端—服务器的两级计算方法和面向移动终端的一系列低复杂度算法。

#### (1) AR/VR 的并行计算与分布式计算

在 AR/VR 应用中,考虑到 PC 的 CPU 计算资源有限,计算领域提出了 GPU 并行计算、多 PC 分布式计算的方法,通过增加计算资源的方式完成海量信息处理需求。在 GPU 并行计算方面,以 AR/VR 三维场景重建为例,Dame 等<sup>[6]</sup>提出了在具有两个 GPU 的 PC 上进行基于并行加速的三维场景重建,首先将三维场景重建拆分为深度估计与目标位姿估计、目标分割及关键帧提取等多个子任务,然后将它们分配到不同 GPU 上并行执行。Pradeep 等<sup>[7]</sup>提出了在具有一个 GPU 的 PC 上进行三维场景重建的并行处理方法,首先将场景重建按像素拆分为若干子任务,然后在 GPU 的多个线程上并行处理各子任务,实现快速计算。上述研究表明,AR/VR 业务的海量信息处理包含多个任务,如三维场景重建、高真实感渲染,部分任务之间可以进行并行处理,同时部分任务还可进一

步分割为多个子任务,如三维场景重建可拆分为深度估计、关键帧提取等子任务实现并行加速;但上述任务或子任务并行处理方法均是基于单一 PC 上的 GPU 并行计算的,故其任务分割方案具有局限性。

区别于参考文献[6-7]只基于一个 PC 计算节点进行并行计算,Lan 等<sup>[8]</sup>提出了一种在分布式计算机系统中实现三维场景重建的并行处理机制,其主要思想是利用不同计算机并行完成视频帧的去噪滤波及深度传感器校验,然后通过吉比特局域网将视频帧传输给主服务器进行数据融合。Pawel 等<sup>[9]</sup>研究了在具有不同计算能力的计算节点上进行并行处理的机制:首先在具有 NVIDIA Tesla S1070 与 NVIDIA Tesla C2070 两种共 4 个 GPU 的 Hal 机上进行 3D 图像重构,利用电容层析成像技术,选择输入矩阵的不同行对应的某一电极并将其分配给特定的 GPU 实现并行计算;其次,增加具有 2x GTX 570 GPU 的 Dave 机,仍采用上述任务分配方法,研究不同计算节点的并行计算。上述在分布式计算机系统进行并行计算的模型具有可扩展性,即可以通过增加计算节点的方式来完成更大规模的计算任务。但模型中各节点计算能力虽然不同,却基本相当,且计算任务对时延没有严苛要求。而通信与计算融合的未来 5G 移动网络将包含手机移动终端、PC、服务器等计算能力存在数量级差异的节点,同时传统基站、集中式接入网(centralized radio access network, C-RAN)设备<sup>[10]</sup>、终端直通(device to device, D2D)<sup>[11-12]</sup>等共存的异构通信网络环境时,单个性能差的计算节点或传输环境都可能导致整个业务出现不可预知的延迟。

#### (2) 移动 AR/VR 终端—服务器两级计算方法

在移动 AR/VR 方面,由于手机终端配备的计算能力逐步提升,如多核处理器和图形处理核,同时配备微电子传感器,能够处理角速度和线性加速度,为移动 AR/VR 信息处理提供了新的可



能。Alex 等<sup>[13]</sup>提出了一种移动终端与云服务器联合进行三维场景重建的方法,其中移动终端进行实时跟踪、关键帧选择及低复杂度关键帧更新,而具有深度信息的关键帧则合并到位姿图中传输给云服务器进行三维场景重建。该模型中云服务器作为移动终端协作网络的中心节点,能够同时支持大量移动终端的密集计算需求,同时用户的三维模型可以存储在云端,被其他用户复用或在此基础进行完善,有效节约计算资源。但参考文献[13]将计算任务卸载到云端进行计算,需要经过路由器、网关等层层关键设备,其交互时延难以得到保证。

针对移动终端计算能力难以满足移动 AR/VR 密集计算需求,而云端计算难以保证时延的矛盾,雾计算(fog computing)<sup>[14-16]</sup>、朵云(Cloudlet)计算<sup>[17]</sup>、MEC<sup>[18]</sup>等概念相继提出,旨在移动网络边缘提供 IT 服务环境和云计算能力,使业务处理更靠近终端,在满足密集计算要求的同时,有效降低业务延迟。Hou 等<sup>[19]</sup>提出将便携式 VR 眼镜与云/边缘计算设备通过无线进行连接,分别将服务器部署在云端、网络边缘(移动网络网关、基站或无线接入点)、终端设备边缘,利用服务器的计算能力来减轻 VR 眼镜重量,分析了不同部署策略适应的业务类型。但该研究主要关注利用视频编码、传输及压缩方法来降低传输视频流比特率,并未解决云/边缘计算服务器如何辅助终端设备进行密集计算的问题。

### (3) 面向移动终端的低复杂度 AR/VR 处理方法

Tanskanen 等<sup>[20]</sup>提出了一种在手机终端上进行稀疏三维重建的算法,利用手机终端的惯性传感器进行实时跟踪与建图,通过选择合适的关键帧、估计场景重建尺度,实现了一组基于稀疏点云的高效三维重建。Kolev 等<sup>[21]</sup>提出了一种在手机终端进行基于深度图的三维场景重建方法,通过评估局部几何方向、底层相机设置和光度等来确定每个深度估计的权重,进而实现深度信息融合。由于手机终端计算能力有限,只能实现稀疏三维

场景重建,与稠密重建质量仍有较大差距,影响用户体验。此外,为了在移动终端获得混合现实的体验,需要算法能够实时稳健地感知场景的三维信息。Ondruska 等<sup>[22]</sup>提出一种可在主流手机上运行的实时场景估计方法,但是由于需要基于体素的深度融合,限制了该方法只能在较小的场景下使用。Mur-Artal 等<sup>[23]</sup>提出基于 ORB (oriented FAST and rotated BRIEF) 特征的相机定位和稀疏场景重建方法,该方法可移植到移动端运行,但是由于场景重建仅包含一些稀疏的三维点,无法满足混合现实的需要。此外,利用 AR/VR 视频视野范围很大的特点,Lai 等<sup>[24]</sup>提出将全景视频划分为规则的区块以实现在移动端并行的低复杂度视频解码,能显著减少时延,然而这种简单的图像分块没有考虑图像本身的内容,也没有考虑时序上的冗余,例如一些图像的内容要比另外一些具有更高的视觉显著度,因此处理它们的优先级应是不同的。

总结上述研究现状可知,现有 AR/VR 信息处理研究利用并行计算和分布式计算来应对其海量信息的密集计算需求,说明 AR/VR 的密集计算任务是可分割的,但这些计算模型和方法都是针对具有相同或相当计算能力的节点设计的,在通信与计算融合的未来 5G 移动网络中,计算节点能力与其部署的通信网络位置有关,可能存在数量级的差异,现有计算模型和方法无法直接应用,必须面向通信与计算融合的未来 5G 移动网络研究新的计算模型,并提出相应的协同计算方法。此外,目前虽然面向移动终端提出了一些低复杂度的 AR/VR 视频信息处理算法,但这些算法尚未充分利用 AR/VR 视频的时空特性,以牺牲精度为代价降低计算量,会给用户带来很大的不适。因此亟待解决的难点问题是如何在移动终端实现轻量级移动 AR/VR 处理算法。

## 4 移动 AR/VR 的智能传输

为了解决移动网络时变的传输能力带来的移

动 AR/VR 服务质量不稳定的问题,主要从两个方面进行探讨:一是从系统的角度适配业务传输的需求,比如改变传输频带、增加传输带宽;二是从业务处理和传输的角度适应移动信道的变化,比如面向移动传输的视频分层编码、空口的自适应传输等。

### (1) 无线频谱感知与动态适配

提供更多更好的可用频谱是保障移动 AR/VR 传输需求的一个有效技术。从统计上看,移动 AR/VR 业务在不同的时刻数据传输量可能会发生剧烈的抖动,带宽需求具有快速时变特点。为了获得频谱,首先要进行频谱感知,主要包括节点频谱感知和频谱地图两种方式。节点频谱感知方面,已有研究针对未来移动网络,提出了基于认知导频信道的快速频谱感知方法<sup>[25]</sup>,但此类方法普遍存在实现简单但正确性低或者感知性能好但时延的问题<sup>[26-27]</sup>。而频谱地图基于大量的计算和存储,能快速提供较为准确的频谱信息。参考文献[28]采用频谱地图,记录用户使用频谱的频段、时间、用户位置以及服务质量等多维信息,为精确、安全、快速的频谱感知提供参考。参考文献[29]论述了频谱地图的应用,将事件学习算法和知识学习算法融入频谱地图,在更好地实现各项功能的同时,有效保护授权用户。参考文献[30]中介绍了一种多级无线电环境地图的架构,在核心网侧、基站侧、终端侧均放置了计算处理单元,分别用于生成各层级的无线电环境地图,实现整个覆盖区域频谱资源的高效协调。移动 AR/VR 业务作为时延敏感的高带宽传输业务,对频谱质量、时延等性能提出全面的要求。但上述研究重点关注可用频谱信息,未考虑频谱质量问题,不能满足 AR/VR 获取高质量传输频带的需求。在频谱适配方面,现有研究主要考虑用户公平性、系统吞吐量等方面性能。参考文献[31]提出了跨成员载波的比例公平调度算法,可以保障载波聚合下多用户调度的公平性。参考文献[32-33]提出了基于贪

婪算法的联合载波选择与资源块分配的资源调度方法以提高系统吞吐量。参考文献[34]提出了多载波共用一个缓存队列的联合队列调度方式,相比于独立队列调度方式可以充分利用分集增益。参考文献[35]分析了用户调度服从泊松分布时,基于轮询与移动散列算法的频谱适配的性能。值得注意的是,上述研究都未能与 AR/VR 结合。在 AR/VR 场景中,一个用户请求在处理时通常被分割成多个业务,如三维场景重建、高真实感渲染等,这些业务的数据传输量有很大差异。因此,在 AR/VR 应用时,需要考虑这些不同业务的用频需求差异,此外,还需要考虑同一业务不同编码层对服务质量的需求差异。因此,现有频谱适配无法满足 AR/VR 多种差异化业务的需求。

### (2) 面向移动传输的视频分层编码

在传统视频信息处理研究中,目前已经提出了使用不同的编码方式适应移动传输环境,为用户提供不同等级质量的视频服务。两种主流的编解码方案为可分级视频编码(scalable video coding, SVC)和感知视频编码。SVC 方面,主要是基于 H.264/AVC(advanced video coding)混合编码框架,在时间、空间、质量上实现分层编码,以实现不同帧率、图像分辨率和图像质量等级的自适应调整<sup>[36]</sup>。H.265/HEVC(high efficiency video coding)与 SHVC(scalable extension of HEVC)作为 H.264/AVC 的继任者,同样继承了分层编码的特性<sup>[37-38]</sup>。SHVC 支持不同空间分辨率或重建信噪比(signal noise ratio, SNR)的多层视频序列的编码,支持高达 8 个分层,包括一个基本层和多个增强层。通过引入上采样滤波、层间纹理预测、层间运动预测等技术,SHVC 还支持混合编解码可分级、比特深度可分级和色彩域可分级。感知视频编码方面,主要是利用了视频信号中存在的视觉冗余,即人眼不能察觉的图像中的某些信息。人类视觉系统的研究表明,人眼对图像和视频的感知是有选择性的,不同的对象或者区域



具有不同的视觉重要性,并且对视觉信号的各种失真具有不同的敏感和容忍程度。基于视觉特性的视频编码主要思路是如何根据视觉感知特性选择优化的编码参数,实现码率的优化分配,例如为视觉重要区域分配更多的码率资源来提高其主观质量;减少非视觉重要区域的码率资源来减少视觉冗余,从而提高编码性能<sup>[39]</sup>。此外,多视点视频能提供立体感和交互性,结合多视点视觉感知模型的特点,基于视觉特性的多视点视频编码可分为基于立体视觉注意的多视点视频编码和基于立体视觉可见度的多视点视频编码<sup>[40-41]</sup>。基于下一代编码标准 H.265/HEVC 的视频感知编码可能成为未来 AR/VR 编解码的重要工具<sup>[42-43]</sup>。由上述研究可见,当前视频编码大多面向传统视频,尚未有针对 AR/VR 的编解码机制。在 AR/VR 编解码研究中,应结合其独有的特征如深度信息、六自由度和用户的感兴趣区域的差异性等,一方面提高编码压缩比,轻量化业务数据量;另一方面需要降低编码复杂度,降低对移动设备的计算压力和续航压力,为智能地在时变的移动无线环境中提供不同的体验提供编码方案的基础。

### (3) 空口感知的 AR/VR 自适应传输

由于无线信道的时变特性,很难满足 AR/VR 持续大带宽和低时延的需求,移动 AR/VR 传输面临很大挑战。此外,考虑多用户场景,由于用户状态不同,即使观看同一个移动 AR/VR,所需的画面也不完全相同,无法像普通视频那样,采用多播的传输方式来提升频谱效率。这进一步加大了移动 AR/VR 的带宽需求。如何设计能够感知空口的变化、符合业务和无线网络特点的自适应机制是亟待解决的问题。现有 AR/VR 传输的成果大多基于有线网络环境,只有少量的相关工作基于无线环境。例如,在参考文献[44]中,运用博弈论等方法,提出了一种异构蜂窝网络下适合 AR/VR 传输的资源管理策略,但时延仍然难以达到 AR/VR 业务要求。参考文献[45]提出了一种时延

导向的基于无线局域网的 AR/VR 多用户接入策略,分析了传输时延的组成部分,制定出了一套 AR/VR 多用户接入方案。另外一方面,云服务的引入为移动视频传输提供了新的思路。首先,可利用集中式特点来进行更好的资源管理。已有研究面向集中式移动网络架构,提出负载感知的资源管理,可提高资源利用率达 70%<sup>[11]</sup>。其次,可利用集中式架构中心处理单元的计算资源进行视频转码,根据用户需求将原始视频转码成较低版本,在空口传输时节省传输内容,提高空口资源利用率。在参考文献[46]中,提出基于集中式云服务的转码机制,利用云计算,以用户的信道条件所需的视频质量为参数来计算转码版本,提升视频质量和频带利用率。参考文献[47-48]针对一对多实时转码的视频直播服务,提出了一种云计算资源分配的方案,在提供用户所需的视频质量的前提下结合地域性差异最小化计算资源的开销。参考文献[49]则是在参考文献[47-48]的基础上,针对流式直播服务,为视频接收者和发布者设计了一套云服务器联合选择策略,降低了系统开销和地域性差异的影响。根据直播视频发布者受欢迎程度的不同,参考文献[50]借助计算资源,通过建立多目标优化问题,最小化计算资源开销和最大化用户体验,在时延约束的条件下利用李雅普诺夫优化的方法求解并提出了一套计算和带宽资源联合分配方案,使网络性能和用户体验得到了权衡。参考文献[51]将云计算与内容中心移动网络结合,为多媒体服务提出了一种较为完整的云网络架构,包括计算资源的部署和网络的拓扑结构。但以上研究尚未与移动 AR/VR 特征相结合。可见,目前针对移动 AR/VR 业务的传输策略研究尚处于起步阶段,在通信与计算融合的移动网络架构下对 AR/VR 传输策略的研究国内外均属空白,尚未见任何成果公布,其中蕴含着大量研究机会。针对移动 AR/VR 传输的资源分配、版本选择、多播等技术,都有必要探索新的解决思



路和方法,从而更为有效地改善移动 AR/VR 业务的用户体验。

总结上述研究现状可知,作为一种崭新的移动多媒体服务,面向 AR/VR 的移动传输机制研究目前尚处于初级阶段,在频谱感知、信源分层编码、空口自适应传输等都少有针对性的研究,亟需根据移动 AR/VR 的需求,剖析其与移动传输相关的特征,面向通信与计算融合的未来 5G 移动网络,设计智能的传输机制。

## 5 移动 AR/VR 服务时延模型及优化保障

用户交互得到即时响应是移动 AR/VR 良好用户体验的主要来源。当服务时延大于图像刷新时间间隔时,用户将产生晕眩感。因此,服务时延是影响体验的重要因素。目前在这个方向上已有大量的研究,提出了多种时延优化保障机制,研究的主要思路是基于给定的服务时延模型,从移动网络的资源管理、移动切换、AR/VR 用户位姿预测、边缘缓存等方向展开服务时延的优化或保障。

### (1) 移动 AR/VR 的服务时延模型

目前移动 AR/VR 服务时延模型相对比较简单,仅考虑单服务器节点情况。Deng 等<sup>[52]</sup>研究了在小区基站部署云服务器进行密集计算任务的卸载,其中假设应用的任务已在云服务器虚拟机上进行备份,故其服务时延模型不考虑任务的传输时延,仅考虑任务的处理时延以及各任务之间彼此依赖且上一任务与当前任务在不同节点进行处理时上一任务结果的传输时延;Chen 等<sup>[53]</sup>研究了将密集计算任务卸载到移动边缘云计算的时延模型,其主要包括任务在 3G/4G 无线信道的上行传输时延及任务在移动边缘云服务器的处理时延,而未考虑任务处理结果的下行回传时延;Lai 等<sup>[24]</sup>研究了手机终端 VR 业务卸载到台式机进行处理的时延模型,其中手机终端与台式机通过基于 IEEE 802.11ac 协议的无线局域网(wireless local

area network, WLAN) 进行相连,服务时延则主要包括请求等待时延、视频流原始帧的传输时延及在台式机计算时延,其中视频流原始帧的 WLAN 中的传输时延为主要时延;Ahn 等<sup>[45]</sup>研究了 WLAN 中基于 IEEE 802.11 协议的多用户 VR 服务时延模型,其服务时延主要包括用户感知时延、终端决策时延、上行传输时延、PC 处理时延、下行传输时延及终端整合时延,该模型中考虑了多用户场景下信道接入时延问题,重点分析了多用户业务上行数据分组传输时延。上述模型均是基于单服务器节点的服务时延模型,其结构简单,很难直接拓展到包含 D2D 通信、传统小区基站通信与 C-RAN 集中式架构等并存的异构通信网络,且异构通信网络中包含终端、MEC、云服务器等多级不同计算能力的服务节点,存在多种计算方式,同样影响移动 AR/VR 服务时延建模。

### (2) 基于资源管理的移动 AR/VR 的服务时延优化

在给定服务时延模型基础上,针对移动 AR/VR 服务时延的优化主要集中在 3 个方面,即研究时延约束下用户终端功率损耗问题、时延与功耗的联合优化以及直接以最小化服务时延为目标。在时延约束下最小化用户终端功耗方面:Cao 等<sup>[54]</sup>提出了单用户场景中基于组合优化算法的最优自适应算法和基于贪婪算法的次优算法,将终端能耗分别降低 48%和 47%;Zhao 等<sup>[55]</sup>假设所有用户具有相同的信道质量和计算能力,其提出的最优算法和次优算法能分别节约 40%、30%的能耗;Vondra 等<sup>[56]</sup>探讨了通信与计算负载的平衡,在满足时延要求的前提下提出一种应用考虑算法,根据当前计算和通信负载来选择合适的接入基站。在时延与功耗的联合优化方面:Muñoz 等<sup>[57]</sup>提出一种联合分配无线通信和计算资源的通用架构以对用户终端能耗和时延做折中处理;对于多用户场景,Mao 等<sup>[58]</sup>基于 Lyapunov (李雅普诺夫)优化提出一种在线算法用于决策计算任务的分配,



在每个时隙, 利用 Gauss-Seidel (高斯-塞德尔) 方法来决定计算卸载的优化传输功率及带宽分配。在最小化服务时延方面: Liu 等<sup>[59]</sup>采用马尔可夫决策过程方法对基与任务缓存排队状态、本地处理单元执行状态以及传输单元状态的计算任务进行调度, 通过分析每个任务的平均时延及移动终端的平均功率损耗, 研究功率约束下的最小化时延问题, 并通过一维搜索算法寻找优化调度策略; Mao 等<sup>[60]</sup>研究了加入能量收集设备的 MEC 系统中单用户计算卸载策略, 综合考虑了卸载决策、CPU 周期频率以及计算卸载的传输功率, 以最小化时延为目标, 通过 Lyapunov 动态优化算法实现优化目标。上述时延优化问题均假设待处理的任务是可被任意分割的, 但是实际上, 移动 AR/VR 任务的分割比例是与业务本身息息相关的; 同时上述问题中仅针对终端及 MEC 或云端两级计算节点, 且都假设网络中通信、计算与存储资源是有限的, 但是并没有探究如何进行资源部署以及资源部署对服务时延的影响。

### (3) 计算节点移动切换机制

在通信与计算融合的未来 5G 移动通信网络中, 通信与计算节点切换是必不可少的核心技术, 切换时延对移动 AR/VR 的体验有重大影响<sup>[61]</sup>。计算节点的切换主要分为按需切换和主动推送两大类。按需切换方面, 参考文献[62]提出分布式 MEC 服务器总是跟随服务基站无线切换而变化, 可以实现业务传输时延最小化。参考文献[63]考虑协作小区通信方式, 根据终端时延需求等确定服务小区, 通过无线链路和终端相连, 协作组内的其他小区和服务小区通过回传链路连接。通过马尔可夫决策过程 (Markov decision process, MDP) 算法选择最优切换节点, 在满足能耗需求的条件下, 最小化业务总时延。参考文献[64]考虑云、分布式 MEC 和 D2D 协作多级计算架构, 在终端与基站连接时间等限制下, 进行计算节点切换决策, 最大化系统接入能力。已有计算节点按需切换机

制普遍存在切换时延延长的弊端, 切换时延至少是 60 ms<sup>[65]</sup>, 甚至可达数秒<sup>[62]</sup>, 远不能满足移动 AR/VR 提出的 RTT 为 20 ms 的要求。主动推送切换方面, 其核心思想是利用移动终端具备定位能力, 其移动具有规律性和可预测性的特点, 主动将计算任务提前推送到将要切换的目的节点, 改善切换时延性能。参考文献[66]基于终端、MEC 和云端三层架构, 通过挖掘终端历史数据, 采用多项式非线性回归的方法, 预测终端位置, 提前进行 MEC 切换, 保证服务的连续性。参考文献[67]假设可以采用一定的方法预测出未来  $T$  时间窗口内, 每一个 MEC 执行任务的传输成本及业务在 MEC 间切换的成本上限, 采用 MDP 方法, 确定服务 MEC 序列, 最小化终端在  $T$  时间的平均成本。参考文献[68]在参考文献[63]的基础上, 加入终端位置预测, 进行最优切换选择机制的设计, 进一步降低业务处理时延和能耗。现有的主动推送切换机制大都是基于位置预测进行推送, 没有考虑切换时机和新入网终端等因素。而移动 AR/VR 的主动推送切换需要切换的时间、地点、内容 3 方面的预测与实际需求精确匹配, 才能保证内容推送的实时性和精确性。因此, 仅基于位置预测的传统主动推送机制性能无法保证移动 AR/VR 业务切换的性能需求。

### (4) 基于用户位姿预测的时延优化机制

AR/VR 业务良好体验的重要来源之一就是自由的用户位姿, 主要分为用户在虚拟现实中的运动任务、用户在空间中的自由移动以及基于空间位置的行为交互。无论何种用户行为, 都需要大量的计算资源开销。基于用户行为的位姿降低 AR/VR 服务时延, 提升实时计算性能, 是目前业界研究的热点。Fang 等<sup>[69]</sup>提出了一种基于视觉惯性的实时用户位姿跟踪算法, 结果表明该方法能够实时、稳定地为移动虚拟现实提供一种平滑、稳健的六自由度运动跟踪。Kim 等<sup>[70]</sup>使用事件相机 (event camera) 实现了基于用户位姿实时重建

立体场景的方法, 基于 3 个解耦的概率过滤器, 分别追踪不同自由度的用户位姿, 计算场景强度的梯度图和场景的深度图, 并合成为关键帧, 通过时域和空域的超分辨率重建从低比特率的事件流中恢复实时三维场景视频序列。Dobbins<sup>[71]</sup>提供了一个运动捕捉和虚拟现实的协同可视化系统, 确保了六自由度 AR/VR 业务与用户交互的可能。位姿捕获系统捕获一个或多个用户的头部旋转信息, 从而控制真实世界视频的平移、倾斜和缩放。当用户头部的位置改变时, 其所视的虚拟视频内容需要实时相应改变。以上方法分别从视觉惯性、运动捕捉等方面进行了用户位姿的跟踪及相应视频内容的播放, 并没有考虑视频内容的兴趣区域。对于移动 AR/VR 而言, 用户对视频内容兴趣区域的变化同样可以概括为用户位姿的范畴。此外, 当前没有针对用户位姿预测的研究, 由于移动 AR/VR 业务计算量大, 会产生较大的业务处理时延, 如果能够对下一时刻用户位姿进行预测, 并根据预测结果进行计算资源的提前分配实现任务的提前计算, 可有效降低处理时延。

#### (5) 低时延缓存策略

如前所述, 未来移动网络将融合 MEC 节点, 提供计算与存储功能, 可以使移动 AR/VR 内容更加靠近用户。如果用户请求的视频内容被缓存在其可接受服务范围内的缓存点中, 那么用户可以直接从这些缓存点获取视频内容, 不需要再经过核心网或从远端视频服务器获取内容。缓存命中时, 服务时延可以得到有效降低<sup>[72]</sup>。目前, 研究人员针对不同无线网络场景下的高效缓存策略进行了大量研究, 通过分析视频业务的特点、用户历史观看行为和具体的网络架构, 使用户所请求的视频内容能以更高的效率被缓存<sup>[73-74]</sup>。随着研究的深入, 针对具体业务特点的缓存机制研究逐渐受到关注。例如, 为降低业务时延, 参考文献[75]基于内容请求变化率和文件流行度设计了缓存策略, 可满足严格的时延要求; 参考文献[76]

参考内容分发网架构提出了分布式的自适应流媒体服务, 能提高视频服务效率降低时延, 但是文中缓存策略针对基于内容分发网络的超文本传输协议 (hypertext transfer protocol, HTTP) 动态自适应流媒体 (dynamic adaptive streaming over HTTP, DASH) 业务设计, 不能直接用于 AR/VR 业务。参考文献[77]提出了预取与缓存整合策略, 降低了比特率抖动问题, 有效提升了视频流命中概率。为适应视频流媒体业务的可按照多种速率进行分发的特点, 参考文献[78]基于最近最少使用内存清理提出通过最高速率视频段缓存和转码的结合, 有效提升了缓存命中率。此外, 参考文献[79]通过采用分层视频编码的视频传输, 提高了缓存命中率。参考文献[80]提出了一种基于云服务的多级视频缓存架构, 在边缘缓存的基础上, 还在核心网部署云服务器和云缓存, 在降低网络传输时延的同时提升了用户体验质量 (quality of experience, QoE)。在缓存策略设计方面, 已有以提高缓存命中率为目标, 设计了面向 DASH 的云接入网缓存策略<sup>[81]</sup>。根据上述调研, 目前针对业务特点的缓存研究已经逐步深入, 但是尚未有研究充分结合未来移动网络通信与计算融合的架构特点以及移动 AR/VR 业务的多视角特点进行缓存策略设计。

总结上述研究现状可知, 鉴于服务时延对移动 AR/VR 的重要性, 目前已经展开了大量相关研究, 但仍存在一系列挑战。在服务时延模型方面, 现有基于单服务器节点的服务时延模型过于简单, 必须面向通信与计算融合的未来 5G 移动网络架构, 研究匹配的移动 AR/VR 服务时延模型; 在基于资源管理的时延优化与保障方面, 未能与移动 AR/VR 特征紧密联系, 也缺乏从整个系统层面的资源保障, 必须基于移动 AR/VR 处理与传输的特征, 分析通信与计算资源对服务时延的影响并提出相应的优化保障机制, 例如可以研究网络切片机制, 从系统层面确保移动 AR/VR 的资源, 从



而保障时延性能;在低时延缓存和基于用户位姿预测的时延优化机制方面,也有很大的研究空间。

## 6 结束语

虽然移动 AR/VR 有望成为未来 5G 移动网络的杀手级应用,但目前由于移动终端计算能力低、移动传输能力不稳定、移动网络服务时延大等原因,难以满足移动 AR/VR 的信息处理与传输需求,用户体验差强人意。未来 5G 移动网络将是一个异构通信与多级计算融合的网络,协同通信与计算、存储资源,有可能大幅提升网络性能,满足移动 AR/VR 的信息处理与传输需求。面向通信与计算融合,未来 5G 的移动 AR/VR 信息处理与传输应在基于多级计算的移动 AR/VR 海量信息处理、融合计算的移动 AR/VR 智能传输机制、保障 AR/VR 时延的移动网络异质资源协同与优化等方面展开研究,深入探讨其中的机理,抽象出理论模型,并提出高效的机制,大幅改善移动 AR/VR 的用户体验,促进 5G 发展。

## 参考文献:

- [1] 3GPP. System architecture for the 5G system, version 1.2.0: TS23.501[S], 2017.
- [2] ETSI. Mobile edge computing (MEC) ETSI Industry Specification Group (ISG) version 1.1.1[R], 2016.
- [3] 德意志银行. 了解关于 VR 的一切[R], 2016. Deutsche Bank. Learn all about VR[R], 2016.
- [4] 齐彦丽, 周一青, 刘玲, 等. 融合 MEC 的未来 5G 移动通信网络[J]. 计算机研究与发展, 2018, 55(3): 478-486. QI Y L, ZHOU Y Q, LIU L, et al. MEC coordinated future 5G mobile wireless networks[J]. Journal of Computer Research and Development, 2018, 55(3): 478-486.
- [5] 周一青, 李国杰. 未来移动通信系统中的通信与计算融合[J]. 电信科学, 2018, 34(3): 1-7. ZHOU Y Q, LI G J. Convergence of communication and computing infuture mobile communication systems[J]. Telecommunications Science, 2018, 34(3): 1-7.
- [6] DAME A, PRISACARIU V A, REN C, et al. Dense reconstruction using 3D object shape priors[C]//IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2013, Portland, OR, USA. Washington DC: IEEE Computer Society, 2013: 1288-1295.
- [7] PRADEEP V, RHEMANN C, IZADI S, et al. MonoFusion: real-time 3D reconstruction of small scenes with a single web camera[C]//IEEE International Symposium on Mixed and Augmented Reality, Oct 1-4, 2013, Adelaide, Australia. Washington DC: IEEE Computer Society, 2013: 83-88.
- [8] LAN G, LUO Z, HAO, et al. Development of a virtual reality teleconference system using distributed depth sensor[C]//IEEE International Conference on Computer and Communications, December 13-16, 2017, Chengdu, China. Washington DC: IEEE Computer Society, 2017: 975-978.
- [9] KAPUSTA P, MAJCHROWICZ M, SANKOWSKI D, et al. Acceleration of image reconstruction in 3D electrical capacitance tomography in heterogeneous, multi-GPU system using sparse matrix computations and finite element method[C]//The Federated Conference on Computer Science and Information Systems, September 11-14, 2016, Gdansk, Poland. Washington DC: IEEE Computer Society, 2016: 679-683.
- [10] QIAN M, WANG Y, ZHOU Y, et al. A super base station based centralized network architecture for 5G mobile communication systems[J]. Digital Communications and Networks, 2015, 1(2): 152-159.
- [11] ZHOU Y, LIU H, PAN Z, et al. Energy efficient two-stage cooperative multicast based on device to device transmissions: effect of user density[J]. IEEE Transactions on Vehicular Technology, 2016, 65(9): 7297-7307.
- [12] SUN Q, TIAN L, ZHOU Y, et al. Energy efficient incentive resource allocation in D2D cooperative communications[C]//IEEE International Conference on Communications (ICC), June 8-12, 2015, London, UK. Washington DC: IEEE Computer Society, 2015: 2632-2637.
- [13] LOCHER A, PERDOCH M, RIEMENSCHNEIDER H, et al. Mobile phone and cloud-a dream team for 3D reconstruction[J]. IEEE Applications of Computer Vision, 2016: 1-8.
- [14] MASIP-BRUIN X, MARÍN-TORDERA E, TASHAKOR G, et al. Foggy clouds and cloudy fogs: a real need for coordinated management of fog-to-cloud computing systems[J]. IEEE Wireless Communications, 2016, 23(5): 120-128.
- [15] ZHENG K, MENG H, CHATZIMISIOS P, et al. An SMDP-based resource allocation in vehicular cloud computing systems[J]. IEEE Transactions on Industrial Electronics, 2015, 62(12): 7920-7928.
- [16] BONOMI F, MILIT RO, ZHU J, et al. Fog computing and its role in the internet of things[C]//ACM Workshop on Mobile Cloud Computing, August 17, 2012, Helsinki, Finland. New York: ACM Press, 2012: 13-16.
- [17] LIU Y, LEE M J, ZHENG Y, et al. Adaptive multi-resource allocation for cloudlet-based mobile cloud computing system[J]. IEEE Transactions on Mobile Computing, 2016, 15(10): 2398-2410.
- [18] ETSI. Mobile edge computing (MEC): framework and reference architecture, version 1.1.1[R], 2016.
- [19] HOU X, LU Y, DEY S, et al. Wireless VR/AR with edge/cloud computing[C]//International Conference on Computer Communication and Networks, September 18, 2017, Vancouver, Canada. Washington DC: IEEE Computer Society, 2017: 1-8.
- [20] TANSKANEN P, KOLEV K, MEIER L, et al. Live metric 3D reconstruction on mobile phones[C]//IEEE International Con-

- ference on Computer Vision, March 3, 2013, Sydney, Australia. Washington DC: IEEE Computer Society, 2013: 65-72.
- [21] KOLEV K, TANSKANEN P, SPECIALE P, et al. Turning mobile phones into 3D scanners[C]//IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, USA. Washington DC: IEEE Computer Society, 2014: 3946-3953.
- [22] ONDRUŠKA P, KOHLI P, IZADI S, et al. Mobile fusion: real-time volumetric surface reconstruction and dense tracking on mobile phones[J]. IEEE Transactions on Visualization & Computer Graphics, 2015, 21(11): 1251-1258.
- [23] MUR-ARTAL R, MONTIEL J M M, TARDÓ S J D, et al. ORB-SLAM: a versatile and accurate monocular SLAM system[J]. IEEE Transactions on Robotics, 2015, 31(5): 1147-1163.
- [24] LAI Z, HU Y C, CUI Y, et al. Furion: engineering high-quality immersive virtual reality on today's mobile devices[C]//ACM 23rd Annual International Conference on Mobile Computing and Networking, October 16-20, 2017, Snowbird, Utah, USA. New York: ACM Press, 2017: 409-421.
- [25] LIU L, ZHOU Y, TIAN L, et al. CPC-based backward compatible network access for LTE cognitive radio cellular networks[J]. IEEE Communication Magazine, 2015, 53(7): 93-99.
- [26] SIMON N. Cognitive radio: brain-empowered wireless communications[J]. IEEE Journal on Selected Areas in Communications, 2005, 23(2): 201-220.
- [27] YIN W S, REN P Y, DU Q H, et al. Delay and throughput oriented continuous spectrum sensing schemes in cognitive radio networks[J]. IEEE Transactions on Wireless Communications, 2012, 11(6): 2148-2159.
- [28] QUAN Z, CUI S, SAYED A H, et al. Optimal linear cooperation for spectrum sensing in cognitive radio[C]//IEEE Military Communications Conference, Oct 29-31, 2007, Orlando, FL, USA. Washington DC: IEEE Computer Society, 2007: 1-6.
- [29] ZHAO Y, MORALES L, GAEDDERT J, et al. Applying radio environment maps to cognitive wireless regional area networks[C]//IEEE 2nd International Symposium on New Frontiers in Dynamic Spectrum Access Networks, Apr 17-20, 2007, Dublin, Ireland. Washington DC: IEEE Computer Society, 2007: 115-118.
- [30] PEREZ-ROMERO J, ZALONIS A, BOUKHATEM L, et al. On the use of radio environment maps for interference management in heterogeneous networks[J]. IEEE Communications Magazine, 2015, 53(8): 184-191.
- [31] WANG Y, PEDERSEN K I, SORESENSEN T B, MOGENSEN P E, et al. Utility maximization in LTE-advanced systems with carrier aggregation[C]//IEEE Vehicular Technology Conference, Sept 5-8, 2011, Yokohama, Japan. Washington DC: IEEE Computer Society, 2011: 1-5.
- [32] LIAO H S, CHEN P Y, CHEN W T, et al. An efficient downlink radio resource allocation with carrier aggregation in LTE-advanced networks[J]. IEEE Transactions on Mobile Computing, 2014, 13(13): 2229-2239.
- [33] SUNDARESAN K, RANGARAJAN S. Energy efficient carrier aggregation algorithms for next generation cellular networks[C]//IEEE 21st International Conference on Network Protocols, Oct 7-10, 2013, Goettingen, Germany. Washington DC: IEEE Computer Society, 2013: 1-10.
- [34] CHUNG Y L, JANG L J, TSAI Z, et al. An efficient downlink packet scheduling algorithm in LTE-advanced systems with carrier aggregation[C]//IEEE Consumer Communications and Networking Conference, Jan 10-13, 2011, Las Vegas, NV, USA. Washington DC: IEEE Computer Society, 2011: 632-636.
- [35] WANG Y, PEDSRSEN K I, SORESENSEN T B, MOGENSEN P E, et al. Carrier load balancing and packet scheduling for multi-carrier systems[J]. IEEE Transactions on Wireless Communications, 2010, 9(5): 1780-1789.
- [36] SCHWARZ H, MARPE D, WIEGAND T, et al. Overview of the scalable video coding extension of the H.264/AVC standard[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2007, 17(9): 1103-1120.
- [37] BOYCE J M, Ye Y, CHEN J, et al. Overview of SHVC: scalable extensions of the high efficiency video coding standard[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2016, 26(1): 20-34.
- [38] SULLIVAN G J, BOYCE J M, CHEN Y, et al. Extensions of high efficiency video coding (HEVC)[J]. IEEE Journal of Selected Topics in Signal Processing, 2013, 7(6): 1001-1016.
- [39] ZHAO Y, CHEN Z, ZHU C, et al. Binocular just-noticeable-difference model for stereoscopic images[J]. IEEE Signal Processing Letters, 2010, 18(1): 19-22.
- [40] 周俊明, 郁梅, 蒋刚毅, 等. 面向 ROI 编码的立体图像比特分配策略分析[J]. 高技术通讯, 2011, 21(10): 1048-1055.
- ZHOU J M, YU M, JIANG G Y, et al. Analysis of bit allocation for ROI based coding of stereoscopic images[J]. Chinese High Technology Letters, 2011, 21(10): 1048-1055.
- [41] SHAO F, JIANG G, YU M, et al. A novel rate control technique for asymmetric-quality stereoscopic video[J]. IEEE Transactions on Consumer Electronics, 2012, 57(4): 1823-1829.
- [42] WEI H, ZHOU X, ZHOU W, et al. Visual saliency based perceptual video coding in HEVC[C]//IEEE International Symposium on Circuits and Systems, May 22-25, 2016, Florence, Italy. Piscataway: IEEE Press, 2016: 2547-2550.
- [43] KIM J, BAE S H, KIM M, et al. An HEVC-compliant perceptual video coding scheme based on JND models for variable block-sized transform kernels[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2015, 25(11): 1786-1800.
- [44] CHEN M, SAAD W, YIN C, et al. Virtual reality over wireless networks: quality-of-service model and learning-based resource management[J]. IEEE Transactions on Communications, 2018 (99): 1.
- [45] AHN J, KIM Y Y, KIM R Y, et al. Delay oriented VR mode WLAN for efficient wireless multi-user virtual reality device[C]//IEEE International Conference on Consumer Electronics, Jan 8-10, 2017, Las Vegas, NV, USA. Washington DC: IEEE Computer Society, 2017: 122-123.
- [46] LAI C F, CHAO H C, LAI Y X, et al. Cloud-assisted real-time transrating for HTTP live streaming[J]. IEEE Wireless Communications, 2013, 20(3): 62-70.
- [47] WANG F, LIU J, CHEN M, et al. Calms: cloud-assisted live media streaming for globalized demands with time/region diversities[C]//IEEE INFOCOM, Mar 25-30, 2012, Orlando, FL, USA. Washington DC: IEEE Computer Society, 2012: 199-207.



- [48] WANG F, LIU J, CHEN M, et al. Migration towards cloud-assisted live media streaming[J]. *IEEE/ACM Transactions on Networking*, 2016, 24(1): 272-282.
- [49] HE J, XUE Z, WU D, et al. CBM: online strategies on cost-aware buffer management for mobile video streaming[J]. *IEEE Transactions on Multimedia*, 2014, 16(1): 242-252.
- [50] ZHENG Y, WU D, KE Y, et al. Online cloud transcoding and distribution for crowd sourced live game video streaming[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2016: 1.
- [51] TANG J, QUEK T Q S. The role of cloud computing in content-centric mobile networking[J]. *IEEE Communications Magazine*, 2016, 54(8): 52-59.
- [52] DENG M, TIAN H, FAN B. Fine-granularity based application offloading policy in small cell cloud-enhanced networks[C]//*IEEE International Conference on Communications Workshops*, May 23-27, 2016, Kuala Lumpur, Malaysia. Washington DC: IEEE Computer Society, 2016: 638-643.
- [53] CHEN X, JIAO L, LI W, et al. Efficient multi-user computation offloading for mobile-edge cloud computing[J]. *IEEE/ACM Transactions on Networking*, 2016, 24(5): 2795-2808.
- [54] CAO S, TAO X, HOU Y. An energy-optimal offloading algorithm of mobile computing based on HetNets[C]//*International Conference on Connected Vehicles and Expo*, Oct 19-23, 2015, Shenzhen, China. Washington DC: IEEE Computer Society, 2015: 254-258.
- [55] ZHAO Y, ZHOU S, ZHAO T, et al. Energy-efficient task offloading for multiuser mobile cloud computing[C]//*IEEE/CIC International Conference on Communications in China*, Oct 14-17, 2016, Shenzhen, China. Washington DC: IEEE Computer Society, 2016: 1-5.
- [56] VONDRA M, BECVAR Z. QoS-ensuring distribution of computation load among cloud-enabled small cells[C]//*IEEE International Conference on Cloud Networking*, Oct 8-10, 2014, Luxembourg. Washington DC: IEEE Computer Society, 2014: 197-203.
- [57] MUNOZ O, PASCUAL-ISERTE A, VIDAL J. Joint allocation of radio and computational resources in wireless application offloading[C]//*IEEE Future Network and Mobile Summit*, Jul 3-5, 2013, Lisboa, Portugal. Washington DC: IEEE Computer Society, 2013: 1-10.
- [58] MAO Y, ZHANG J, SONG S H, et al. Power-delay trade-off in multi-user mobile-edge computing systems[C]//*IEEE Global Communications Conference*, Dec 4-8, 2016, Washington DC, USA. Washington DC: IEEE Computer Society, 2016: 1-6.
- [59] LIU J, MAO Y, ZHANG J, et al. Delay-optimal computation task scheduling for mobile-edge computing systems[C]//*IEEE International Symposium on Information Theory*, Jul 10-15, 2016, Barcelona, Spain. Washington DC: IEEE Computer Society, 2016: 1451-1455.
- [60] MAO Y, ZHANG J, LETAIEF K B. Dynamic computation offloading for mobile-edge computing with energy harvesting devices[J]. *IEEE Journal on Selected Areas in Communications*, 2016, 34(12): 3590-3605.
- [61] SATYANARAYANAN M, BAHL P, DAVIES N. The case for VM-based cloudlets in mobile computing[J]. *IEEE Pervasive Computing*, 2009, 8(4): 14-23.
- [62] WANG K, SHEN M, CHO J, et al. MobiScud: a fast moving personal cloud in the mobile network[C]//*ACM Workshop on All Things Cellular: Operations, Applications and Challenges*, Aug 17, 2015, New York, USA. New York: ACM Press, 2015: 19-24.
- [63] BECVAR Z, PLACHY J, MACH P. Path selection using hand-over in mobile networks with cloud-enabled small cells[C]//*IEEE International Symposium on Personal, Indoor, and Mobile Radio Communication*, Aug 30-Sept 2, 2015, Washington DC, USA. Washington DC: IEEE Computer Society, 2015: 1480-1485.
- [64] RAVI A, PEDDOJU S K. Handoff strategy for improving energy efficiency and cloud service availability for mobile devices[J]. *Wireless Personal Communications*, 2015, 81(1): 101-132.
- [65] CLARK C, FRASER K, HAND S, et al. Live migration of virtual machines[C]//*2nd Conference on Symposium on Networked Systems Design and Implementation*, May 2-4, 2005, Lombard, IL, USA. New York: ACM Press, 2005: 273-286.
- [66] BELLAVISTA P, ZANNI A, SOLIMANDO M. A migration-enhanced edge computing support for mobile devices in hostile environments[C]//*IEEE Wireless Communications and Mobile Computing Conference*, June 26-30, 2017, Valencia, Spain. Washington DC: IEEE Computer Society, 2017: 957-962.
- [67] WANG S, URGAKONKAR R, CHAN K, et al. Dynamic service placement for mobile micro-clouds with predicted future costs[C]//*IEEE International Conference on Communications*, June 8-12, 2015, London, UK. Washington DC: IEEE Computer Society, 2015: 5504-5510.
- [68] PLACHY J, BECVAR Z, STRINATI E C. Dynamic resource allocation exploiting mobility prediction in mobile edge computing[C]//*IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications*, Sept 4-8, 2016, Valencia, Spain. Washington DC: IEEE Computer Society, 2016: 1-6.
- [69] FANG W, ZHENG L, DENG H, et al. Real-time motion tracking for mobile augmented/virtual reality using adaptive visual-inertial fusion[J]. *Sensors*, 2017, 17(5): 1-22.
- [70] KIM H, LEUTENEGGER S, DAVISON A J. Real-time 3D reconstruction and 6-DoF tracking with an event camera[C]//*European Conference on Computer Vision*, October 8-16, 2016, Amsterdam, Netherlands. New York: Springer International Publishing, 2016: 349-364.
- [71] DOBBINS M K, RONDOT P, SCHWARTZ K, et al. Providing a collaborative immersive environment using a spherical camera and motion capture: US 8217995 B2[P]. 2012-07-10.
- [72] AHLEHAGH H, DEY S. Hierarchical video caching in wireless cloud: approaches and algorithms[C]//*IEEE International Conference on Communications*, June 10-15, 2012, Ottawa, ON, Canada. Washington DC: IEEE Computer Society, 2012: 7082-7087.
- [73] WANG X, CHEN M, TALEB T, et al. Cache in the air: exploiting content caching and delivery techniques for 5G systems[J]. *IEEE Communications Magazine*, 2016, 52(2): 131-139.
- [74] MOLISCH A F, CAIRE G, OTT D, et al. Caching eliminates the wireless bottleneck in video aware wireless networks[J]. *Advances in Electrical Engineering*, 2014(9): 74-80.
- [75] LI W, OTEAFY S M A, HASSANEIN H S. Dynamic adaptive

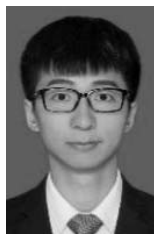
streaming over popularity-driven caching in information-centric networks[C]//IEEE International Conference on Communications, June 8-12, 2015, London, UK. Washington DC: IEEE Computer Society, 2015: 5747-5752.

- [76] LIU Y, GEURTS J, POINT J C, et al. Dynamic adaptive streaming over CCN: a caching and overhead analysis[C]//IEEE International Conference on Communications, June 9-13, 2013, Budapest, Hungary. Washington DC: IEEE Computer Society, 2013: 2222-2226.
- [77] LIANG K, HAO J, ZIMMERMANN R, et al. Integrated prefetching and caching for adaptive video streaming over HTTP: an online approach[C]//ACM Multimedia Systems Conference, March 18-20, 2015, Portland, OR, USA. New York: ACM Press, 2015: 142-152.
- [78] GRANDL R, SU K, WESTPHAL C. On the interaction of adaptive video streaming with content-centric networking[J]. Journal of Sports Sciences, 2013, 23(9): 977-989.
- [79] FUENTE Y S D L, SCHIERL T, HELLE C, et al. iDASH: improved dynamic adaptive streaming over HTTP using scalable video coding[C]//ACM SIGMM Conference on Multimedia Systems, Feb 23-25, 2011, San Jose, CA, USA. New York: ACM Press, 2011: 257-264.
- [80] TRAN T X, POMPILI D. Octopus: a cooperative hierarchical caching strategy for radio access networks[C]//IEEE International Conference on Mobile Ad Hoc and Sensor Systems, Oct 10-13, 2016, Brasilia, Brazil. Washington DC: IEEE Computer Society, 2016: 154-162.
- [81] ZHANG Z, LIU D, YUAN Y. Layered hierarchical caching for SVC-based HTTP adaptive streaming over C-RAN[C]//IEEE Wireless Communications and Networking Conference, March 19-22, 2017, San Francisco, CA, USA. Washington DC: IEEE Computer Society, 2017: 1-6.

#### [作者简介]



周一青(1975-),女,中国科学院大学岗位教授,中国科学院计算技术研究所“百人计划”研究员、博士生导师,无线通信技术研究室副主任,移动计算与新型终端北京市重点实验室研究员,主要研究方向为移动通信、通信与计算融合等。



孙布勒(1992-),男,中国科学院计算技术研究所博士生,主要研究方向为通信与计算融合、大规模多天线、毫米波通信等。

齐彦丽(1991-),女,中国科学院计算技术研究所博士生,主要研究方向为移动边缘计算、通信与计算融合、无线资源管理等。

彭燕(1993-),女,中国科学院计算技术研究所博士生,主要研究方向为通信与计算融合、超密集网络等。

刘玲(1990-),女,中国科学院计算技术研究所助理研究员,主要研究方向为5G无线通信、通信与计算融合、超密集网络、无线资源管理等。

张志龙(1985-),男,北京邮电大学讲师,主要从事多媒体通信、无线视频传输策略优化等方面的研究工作。

刘奕彤(1982-),女,北京邮电大学讲师,主要从事VR视频和三维模型的无线传输等方面的研究工作。

刘丹谱(1972-),女,北京邮电大学国际学院电信工程及管理系主任、博士生导师,IEEE高级会员,中国通信学会和中国电子学会高级会员,一直致力于无线视频传输策略优化领域的基础理论与关键技术研究。

李兆歆(1983-),男,中国科学院计算技术研究所助理研究员,主要从事三维重建领域的研究工作。

田霖(1980-),女,中国科学院计算技术研究所副研究员,主要研究方向为绿色无线通信系统与无线资源管理技术。