



北京 2022 年冬奥会官方合作伙伴
Official Partner of the Olympic Winter Games Beijing 2022

中国联通 5G 新通信 技术白皮书

中国联通研究院

2022 年 5 月

版权声明

本报告版权属于中国联合网络通信有限公司研究院，并受法律保护。转载、摘编或利用其他方式使用本报告文字或者观点的，应注明“来源：中国联通研究院”。违反上述声明者，本院将追究其相关法律责任。



中国联通研究院

目录

前言.....	1
一、背景分析.....	2
（一）市场需求驱动	2
（二）业务演进趋势	3
二、总体构架.....	4
（一）总体构架	4
（二）网络要求	6
（三）终端要求	7
三、关键技术.....	9
（一）可信交互网络	9
1.交互式数据通道.....	9
2.企业/第三方可信接入.....	10
（二）增强内生能力	11
1.下一代编解码技术	11
2.音视频增强技术.....	12
3.统一媒体面	12
4.AR/VR 技术.....	13
5.AI 技术.....	14
（三）灵活泛在端侧	14
四、业务场景.....	15

（一）主叫名片	15
（二）智能客服	17
（三）智能定位	18
（四）远程协助	19
（五）实时通话翻译	20
（六）AR/XR 应用	21
（七）手势翻译	22
（八）量子安全通话	22
五、总结和展望	23
六 附录：缩略语	25
七 参考文献	26



前言

5G VoNR 技术成熟和大规模商用部署在即，5G 创新业务具备了更加强大的网络基础。伴随通信网络和多媒体技术的持续演进，以及人们多元化的通信需求，更高清、交互式、全业务的增强实时通信业务成为越来越多的用户和产业的刚性需求。

本白皮书阐述了 5G 新通信的产业环境，提出网络及业务总体架构，分析涉及的关键能力和技术及典型新通信场景，并提出发展 5G 新通信业务的策略和展望。

本次发布的白皮书通过向业界介绍 5G 新通信最新技术方向和应用场景、演进思路等，推动对现有技术路线可持续发展，促进产业和生态构建。

本白皮书由中国联通研究院牵头编制，联合编制单位（排名不分先后）：华为技术有限公司、中兴通讯股份有限公司。

编写组成员（排名不分先后）：

马瑞涛、魏群、龚子丹、穆佳、马泽芳、陈婉琨、任驰、罗佩华、吴琼、李延斌、程粹茹、邢天齐、李晨仪、刘强、夏肖、孙新国，朱俊、缪永生、夏沫、虞建成。

一、背景分析

（一）市场需求驱动

传统音视频是运营商最为基础和关键的实时通信功能，满足着全球人类迫切的社交需求，在应急场景时为维系生命提供了重要手段。中国联通在 2020 年即完成了全球首个 5G NSA/SA 双模组网，并将全网升级支持 3GPP R16 的 eMBB（增强移动宽带）、uRLLC（低时延高可靠）和 mMTC（海量机器类通信）的最新国际标准，高起点建设 5G 精品网络。更高的速率和更低的时延为 5G 新通信奠定了坚实的网络基础。

从 VoLTE 伊始，通过近十年培育和通信发展，视频与多媒体业务已不仅是社会生活的刚性需求，数字通信的实时互动性、趣味性、能力叠加的趋势明显。AI、区块链、语音语义识别、区块链、AR/VR 技术逐步成熟，以及数据通道技术快速创新，引发了交互和增强通信业务的新模式。“5G 新通信”是中国联通为用户提供极致体验的奋斗目标，为此，中国联通基于 GSMA NG.114，以 VoLTE+VoNR 为 5G 语音解决方案启动和完成了 10 个城市 16 种组合的核心网与无线配套测试，此后从终端、业务、无线三个维度持续拉网调优，致力于全球最大 5G 共建共享商用网络，携手合作伙伴们在全国 125 个城市正式开通 5G VoNR（超高清音视频通话）服务。随着中国联通数字化转型的步伐，我们还将突破传统，为用户提供更为丰富的客户

直达、多元化通信等混合业务体。“5G 新通信”将构建新型的多媒体业务系统,催生更多新的应用,同时极大拓展产业链的完善和发展。

(二) 业务演进趋势

相比拨号盘入口的传统通话,5G 新通信业务特点是高清、可视、互动,使用交互式通信提升运营商原生通话体验,从话音到多媒体,从双向到实时交互,从平面到全感知。通过富媒体能力,提升通话互动和趣味性,打造通话新入口,升级基础通话体验,重塑通话入口优势。

5G 新通信更希望提升业务感知。利用 VoLTE/VoNR 基础网的数据通道,将原生视频通话,发展到更高清、视频交互式 and 沉浸式通信,未来还将进入全息通信,包括听觉、视觉、嗅觉、味觉体验等。而它将从以下方面提出需求:

交互式: 通信参与方在实时通信时增加了多方面的互动,可以发送/接收图片、文字、表情,文件传输,位置和可视菜单等信息,交互的内容可以是扩展的和多变的。

沉浸式: VR/AR 沉浸式通信、协作,构建身临其境与虚实融合的沉浸式体验。

安全性: 增强的通信安全方式将提升终端、企业和网络品牌好感度,包括面向企业的第三方 ID 接入,第三方呼入和呼出场景等,企业与运营商网络之间建立可信关系,帮助用户识别并防范电信网络诈骗

智能化：运营商系统智能等级提升，包括语音/语义识别、多语言翻译、图像识别；例如手语识别等 AI 技术，能满足商务场景和无障碍等更丰富的通信业务需求。

图 1 业务演进趋势

二、总体构架

（一）总体构架

5G 新通信业务以“端网协同、融合开放、使能创新”为设计理念，基于 IMS 网络升级演进。

引擎、原生支持 IMS 数据通道 (Data Channel)，按需从网络侧下载各类应用；通过音频、视频、数据三通道传递各种多媒体信息，与网络侧高效协同，为 ToC 用户提供高清、强交互、沉浸体验的通信业务，并帮助 ToB 行业实现高效、低成本的生产活动。

2.媒体能力层：不仅提供放音、收号、编解码、混音/混屏、转码等传统媒体处理能力，还支持媒体渲染、视频特效、语音识别(ASR)、语音合成(TTS)等新的媒体处理能力。媒体能力层采用插件化框架，支持媒体能力的编排，可以动态加载新增媒体能力库和媒体资源库，实现实时通信的创新业务快速上线。

3.网络能力层：为上层应用提供业务签约、接入鉴权、路由控制、呼叫控制、设备管理、计费管理等基本通信能力，并提供与其他运营商网络的互联互通能力。

4.业务使能层：为了充分发挥运营商通信网络的价值，使能 ToB 和 ToC 实时通信业务创新，业务使能层为各类应用提供了全生命周期管理平台，可以实现应用的开发、联调、部署、发布、更新、下载等；并通过提供业务编排引擎，为运营商自身业务创新和应用生态建设提供支撑。



图 2 5G 新通信业务总体架构

(二) 网络要求

VoLTE 作为不同于传统 CS 电路域提供的音视频解决方案，基于 IP 承载，叠加在 PS（数据分组域）之上，由 IMS（IP 多媒体子系统）为用户提供文本、语音、视频、图片等不同的多媒体信息通信服务。与 VoLTE 类似，VoNR 语音由 5G 网络端到端承载，并由 IMS 系统控制提供基础音视频服务。

为保障用户的良好业务体验，VoNR 与 VoLTE 之间应支持互操作，当 5G NR 不支持语音业务时，网络应支持 EPS Fallback 方式回落到 VoLTE，即 UE 在 5G NR 中发起或接收语音呼叫时，通过重定向或切换的方式回落到 4G 网络，由 VoLTE 来提供语音业务，当语音通话结束后，UE 再返回到 5G 网络。

Wi-Fi 接入作为一种接入速度快、建设成本低的无线接入手段，已经被社会广泛应用。VoWiFi (Voice over Wi-Fi) 是用户使用具有 VoWiFi 能力的智能终端，利用 Wi-Fi 接入，经 ePDG 连接到运营商的 EPC 网络，并通过 EPC 网络连接到 IMS 网络中，实现语音、短信等 IMS 业务。VoWiFi 技术作为一种典型的非 3GPP 接入方式，促进了通信终端向着多元化方向发展，从单一的手机，扩展到了各种可穿戴设备、家庭设备以及行业设备，如智能手表、Pad、电脑、智能音箱、机顶盒、智慧屏、智能汽车、摄像头、路灯等。

原生的新通信业务要求 IMS 网络支持交互式数据通道 (Data Channel, 简称 DC)，称为 IMS DC。IMS DC 在传统音视频通道上叠加新的数据通道，支持传送图片、视频、菜单、表情、位置、AR/VR 等多媒体信息，并提供带宽控制和 QoS 策略，保证实时可靠地传送交互数据信息。同时，通过网络能力开放平台为传统电信业务增加交互式业务特征，同时提供全新的沉浸式业务体验。

支持 IMS DC 的网络应兼顾各种应用对底层通道的多样化诉求，但不关注所传递的内容和格式，只要通信双方达成一致即可，以实现适用于多种场景的新业务。

(三) 终端要求

5G 新通信不仅为新的不同种类的 5G 终端提供崭新的体验，同时还要为现有终端提供一些穿透性的功能，也就是不需要终端任何的

改造也能享受新通信红利。这其中包括：

1.新通信要求现有 MTSI 终端升级为支持 DC 的 DCMTSI 终端。DCMTSI 终端应支持 VoLTE 和 VoNR 音视频协议栈和新的 DC 协议栈，支持 DC 创建、删除和维护等相关管理功能，同时还支持下载/集成 H5 小程序的能力，保证新业务的快速分发和升级。

2.采用 APP+SDK 业务方式实现 5G 新通信时，APP 通过 Data APN 接入，建立新通信平台和媒体面，并与 VoLTE 系统协同。此种方式下，无需终端配合，只需安装 APP 或者 SDK。

3.除此之外，新通信尽可能为更多领域客户服务，是兼容各种标准 SIP 范围内的 IMS 接入网络，具备与企业 PBX、IP 关口局、CS 传统关口局平台互联的能力，能够为包括普通 POS 终端、IP 多媒体终端在内的固定终端，和 2/3G 终端、VoLTE、VoNR 终端的移动终端提供一致性业务服务。



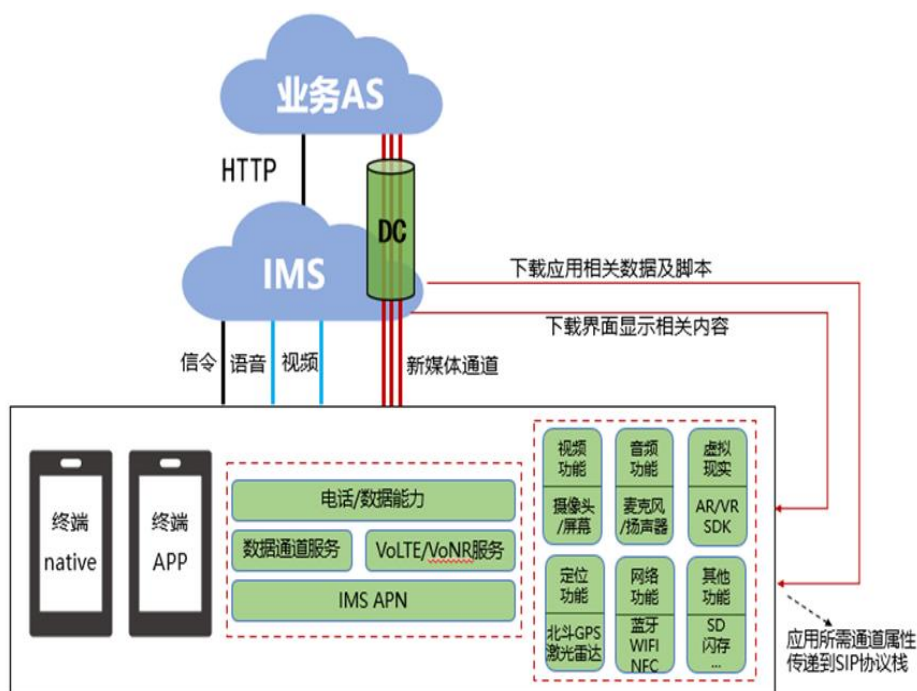


图 3 终端功能框架

三、关键技术

（一）可信交互网络

1.交互式数据通道

音频和视频是 IMS 网络已经支持的两种媒体通道，3GPP 标准在音视频通道基础上引入了新的数据通道，可以用于传输任何类型的数据，并由 IMS 网络为之提供 QoS 保障，以满足应用对时延、带宽和可靠性方面的特殊要求。

终端通过数据通道，在呼叫接通前从网络侧下载各种最新应用，应用可以在通话前、通话中或者通话后传递文本、图片、位置、文件、视频等各种媒体，与音视频媒体流结合，实现通话双方或多方的交互，

显著丰富了用户的业务体验。

数据通道技术具有如下方面的优势：

1.前向兼容：基于 IMS 架构增强，VoLTE/VoNR 均可支持数据通道能力。

2.灵活定制：数据通道与音视频通话可以同步关联独立建立。

3.全量类型：数据通道可传输全量多样化数据类型，并于业务和终端协同感知，为用户提供更优质的用户体验。

2.企业/第三方可信接入

企业外呼是行业用户的刚需，目前通过商业中继或者点击回呼方式实现的企业外呼，存在呼叫拒绝率高、部署周期长、成本过高等诸多问题。

基于企业和 IMS 网络的双向认证以及企业对自己员工的认证鉴权实现的企业可信接入，使得企业员工能够以其企业员工账号接入 IMS 网络，呼叫大网用户展开业务活动。借助企业可信接入技术，能够使企业充分享受到运营商网络通信随处可达、免 App 安装的优势，有效降低企业获客成本，同时能够将企业认证的企业员工身份信息通过 IMS 网络安全、可信地传递到被叫侧，有效提高企业外呼的接通率，从而提高业务成功率。

（二）增强内生能力

1. 下一代编解码技术

随着通信业务的视频化及高清化，移动通信网络从 2G、3G、4G 发展到 5G 时代，其语音编解码技术也从 AMR（Adaptive Multi-Rate）、AMR-WB（Adaptive Multi-Rate Wideband）演进到 EVS（Enhanced Voice Service），朝着沉浸式语音编解码 IVAS（Immersive Voice and Audio Services，沉浸式语音及音频服务）持续演进；而移动通信网络的视频编解码技术，也从 H.264/AVC（Advanced Video Coding，高级视频编码）逐步演进到 H.265/HEVC（High Efficiency Video Coding，高效视频编码）。

EVS 是继 AMR-WB 之后语音编解码技术的又一次改进，它提高了编码的灵活性和效率，主要应用于 VoLTE 高清语音。EVS 的频率范围扩展到 14~17KHz，已接近人耳可感知的 20KHz 极限，更加逼真。IVAS 是正在进行标准化工作的下一代音频编解码器，基于 EVS 编解码，具有更好的抗丢包能力。IVAS 编解码可以更轻松地在通话、会议场景跟踪多方的对话，并区分人声与背景声，为用户带来沉浸式通话体验。

H.265/HEVC 视频编解码标准在 H.264/AVC 的基础上，从提高压缩效率、提高鲁棒性和错误恢复能力、减少实时的时延、减少信道获取时间和随机接入时延、降低复杂度等方面进行了大幅优化，以达成在有限带宽下传输更高质量的网络视频，仅需原先的一半带宽即

可播放相同质量视频的目标，满足超高清时代的视频编解码需求。

2. 音视频增强技术

为了应对丢包、卡顿、时延等弱网场景，通常会采用带宽自适应技术，通过发送端动态调整比特率、数据包速率，减少发送的数据量，以适应网络资源的动态变化，最大程度保证通话质量。同时，在接收端可以应用超分（Super-Resolution，超分辨率）技术，通过 AI 深度学习将低分辨率图像/视频重建成高分辨率，使模糊的图像/视频瞬间变高清，为移动终端用户带来极致视频体验。

这一技术，利用了无线网络的上下行资源不对称部署模型和当前视频电话上下行资源对称模型，在移动终端支持发送和接收不对称分辨率的前提下，在终端和网络媒体处理节点上部署超分能力，保证移动终端能在视频通话过程中减轻对网络上行资源的依赖，并借助网络媒体处理节点调整分辨率，从而实现通话路径中资源和效果的双优。

3. 统一媒体面

随着实时通信业务向着高清、强交互、沉浸体验发展，对通信网络提出了更大带宽、更低时延、更丰富媒体处理能力的要求。

传统的媒体处理网元，存在功能分散、部署重复、接口不灵活等问题，造成媒体资源浪费、媒体路径迂回、新功能部署困难。而通过构建统一媒体面，可以有效地解决这些问题。

1. 采用服务化接口（Service based Interface, SBI），为业务网元提供了丰富的媒体处理能力，不仅有放音、收号、编解码、混音/混

屏、转码等传统媒体处理能力，还支持媒体渲染、视频特效、语音识别、语音合成等新的媒体处理能力。

2.采用插件化框架，支持媒体能力编排，可以动态加载新增的媒体能力库和媒体资源库，支撑实时通信创新业务快速上线。

3.实现了媒体能力的按需部署、最优调度，可以更好地支撑未来 5G 新通信业务的创新和发展。

4.AR/VR 技术

增强现实（Augmented Reality，AR）技术，将虚拟信息与真实世界巧妙融合，运用多媒体、三维建模、实时跟踪、智能交互、传感等多种技术手段，将计算机生成的文字、图像、三维模型、音乐、视频等虚拟信息叠加到真实世界中，并被人所感知，虚实结合，实现超越现实的感官体验。

虚拟现实（Virtual Reality，VR）技术，利用建模、实时三维图形生成技术、立体现实、传感等技术，利用现实生活中的数据，通过计算机技术产生电子信号，并将其与各种输出设备结合使其转化为能够让用户在虚拟世界感受到的现象，其模拟环境的真实性与现实世界难辨真假，让人有种身临其境的感觉。

基于 AR/VR 技术，在 IMS 网络中引入 AR/VR 会话管理和 AR/VR 标记、媒体渲染、跟踪定位等媒体处理能力，实现 AR 通话、远程协作等新业务，为 5G 新通信带来更丰富的业务体验。

5.AI 技术

为残障人士、老年人等提供无障碍的通信服务,消除“数字鸿沟”,是 5G 新通信的努力方向。

在这些通信场景,需要实现音频、视频、文本不同媒体间的实时互转,这就需要在 5G 新通信中引入 AI 技术,主要包括:

语音识别:提取语音的特征向量,将其编码成数字向量;然后利用基于神经网络的声学模型转换成字母,再利用基于神经网络的语言模型组装成单词。

自然语言处理 (Natural Language Processing, NLP):对文本进行去噪声、词汇归一化等预处理后,再进行语法分析、句法分析、语义理解、文本分类、文本相似度处理、情感倾向分析等处理。

语音合成(Text-To-Speech, TTS):将文本转化成音素序列后,通过拼接法或参数法,根据音素序列生成语音。

基于这些 AI 技术,可以实现语音和文本之间的实时转换,帮助听力下降的老年人更轻松地与他人通信;还可以实现不同方言、语种之间的实时互转,让不同国家和地区的人们直接交流。此外,基于手语识别和手语合成 AI 技术,还可以实现语音和手语之间的实时互转,帮助聋哑人更好地与正常人无障碍通信。

(三) 灵活泛在端侧

当前的通信业务创新严重依赖于终端的配合,导致新业务部署周

期长且缺乏灵活性。为了解除业务创新对终端的依赖，初期可以采用端侧小程序架构，基于 Web 框架将终端能力解耦，在保持终端基础 SDK/Native App 稳定的同时，将变化频繁的业务逻辑等部分以 H5 小程序方式实现，并在呼叫过程中通过 IMS 数据通道从网络侧下载到终端。

小程序方式可以在终端无感知的情况下从网络侧下发新业务，无需终端升级，有助于新业务的快速部署。网络侧可以基于一定的分发策略下发新业务，比如面向友好用户进行小规模试点，有助于业务的快速迭代和创新，再如面向特定用户群推送新业务，提高业务成功率。

四、业务场景

（一）主叫名片

相比于传统来电业务只能显示电话号码或通讯录姓名，在语音/视频通话过程中根本无法进行任何信息交互，主叫名片业务支持主叫用户在通话前设置认证名称/通话主题/位置/图片/视频/动图/表情等内容，拨打电话后被叫终端能收到主叫用户设置的系列内容，并展示在来电显示界面。可应用于包括向经运营商认证的用户/企业服务商提供通话前身份展示的名片、向移动用户提供诈骗/骚扰电话提示的防诈骗提醒等场景。

通话主题/位置/图片/视频/动图/表情等来电显示内容需要根据

业务平台要求提前上传并进行审核,或根据业务平台提供的模板进行选择,内容审核通过后方可向被叫展示。

1.在 To C 应用场景下,5G 新通信可以提供“更好玩”和“更好用”的业务,使用户获得身心愉悦感和展示个性的机会,吸引大中学生、职场精英、年轻情侣和“时尚达人”等新潮用户。

主叫名片业务允许通过丰富的多媒体信息展示个性化形象,并在通话前表明来意,以提升接通率,给喜欢尝鲜的年轻用户带来新鲜感,提升通话的炫酷感和娱乐性。

2.在 To B 应用场景下,客服中心(如银行、保险公司、电信运营商等)或企业员工(如理财经理、保险顾问、网约车司机、快递小哥、房地产顾问等)呼叫用户时,往往被认为是骚扰电话,遭到用户的拒绝,需要多次呼叫才能接通,5G 新通信可以显著改善这种尴尬状况。

在呼叫接通前,主叫身份认证、设置的呼叫主题和呼叫的重要性等信息通过图片/视频/文字等方式在被叫终端呈现,明确呼叫的目的性和指向性,大大提升呼叫接通率。同时,还可以向用户发送企业预置的多媒体宣传信息,提升品牌宣传推广效果。

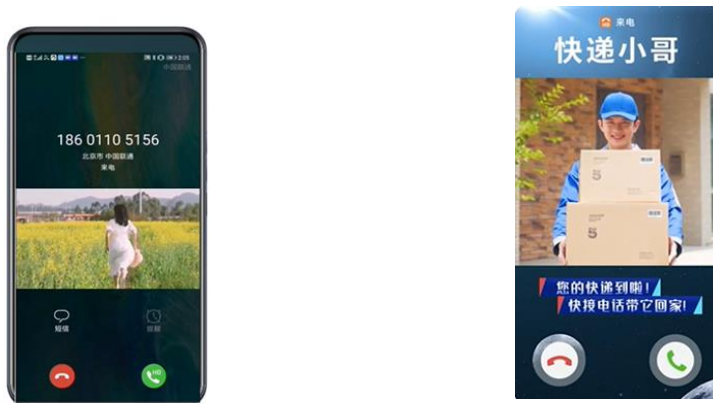


图 4 主叫名片

(二) 智能客服

在 To B 应用场景下，5G 新通信可以提供“更经济”和“更高效”的业务，让用户从中获得更大的经济效益和社会效益，更好地服务于垂直行业、企业和党政机关等用户。

智能客服业务是 To B 业务的典型代表，通常可与企业名片、内容分享等业务进行集成。

目前的客服 IVR 导航冗长，引导能力差；服务完全依赖语音，信息呈现层次叠加复杂；服务全过程碎片化，客户满意度不高。

智能客服业务保留传统 IVR 导航，同时提供交互式菜单，用户可根据自身习惯选择使用。在服务过程中，向用户推送文字、图片、视频、位置和 HTML5 页面等，并根据用户反馈收集所需信息。用户只需拨打一个号码，即可享受“一条龙”服务，显著提升用户体验，帮助企业降低运营成本，改进服务质量，提高工作效率。

当用户与客服平台进行音频或视频通话时，平台可向用户推送可

可视化菜单，用户点击菜单后，平台根据用户菜单点击结果，完成业务逻辑处理，支持发送图片/音视频/文件。菜单页面支持文本、图片、视频、交互按钮等内容的显示。



图 5 智能客服

(三) 智能定位

在 To H 应用场景下，5G 新通信可以提供“更贴心”和“更智能”的业务，让用户从中获得温馨的亲情/友情和集体/社会融入感，适用于家庭、班级、办公室和微型企业等用户。智能定位业务是 To H 业务的典型代表。

对于家庭中的特殊成员（老人、儿童或行动不便者等），由于自理能力比较弱，容易出现意外情况，需要其他家庭成员的协助，因此随时定位特殊成员的位置显得尤为重要。现在手机虽然很普及，但只能进行音视频通话，往往难以描述确切位置。

智能定位业务可以在电话接通前显示特殊成员的当前位置，通话交流只作为辅助手段，无论是报警、赶往现场，还是下达指令，都可以第一时间完成，避免双方的低效沟通，及时采取救助措施，提高特

殊成员的安全系数。

用户拨打特殊成员的手机号码，电话接续过程中，实时采集特殊成员的网络位置信息（如 IP 地址、基站小区等），并结合 GPS 定位信号形成物理位置信息。用户接收该物理位置信息，被叫方的 IP 可以在通话前显示在主叫终端上，所在位置信息用文字显示或直接在地图上标识。



图 6 智能定位

（四）远程协助

远程协助是一种综合业务，兼具交互式和沉浸式两种业务特征，既可应用于 To B 和 To H 场景，也可应用于 To C 场景。该业务通常会和屏幕共享、内容分享或电子白板等业务协同工作。

对于很多老人/儿童或普通用户来说，电脑、电视机、路由器和手持式设备都比较复杂，往往无法独立完成其连线或设置，需要得到专业人士的指导。但传统电话只能进行语音讲解，不够直观形象，难以达成目标。

在远程协助场景中，用户发起 AR 远程协作请求，发起方（求助者）将本地摄像头信息实时共享给接收方（专家），专家在求助者共

享的图像或视频中进行标记，指导求助者进行相应操作。远程协助方式简化了专家的指导过程，避免用户的无效操作，提高解决问题的效率，增进生活的便利和家庭的和谐。

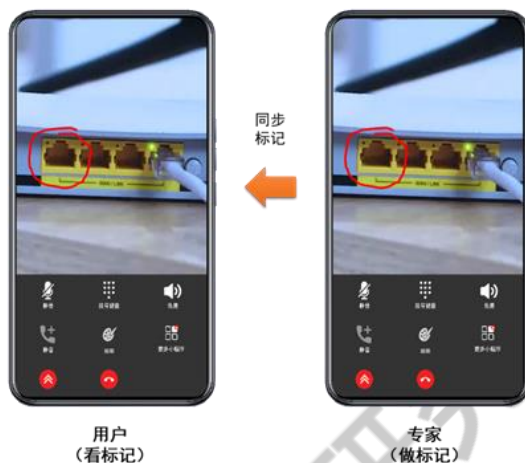


图 7 远程协助

（五）实时通话翻译

实时通话翻译即在通话中通过实时翻译，实现不同语音的一方或者多方自由通话。5G 新通信的实时通话翻译业务，是在不同语种的点对点的音视频通话中，网络将语音转换成文字，在通话中将本端语音翻译后的文字显示在对端的终端侧。比如正在进行的 IMS 通话会话中，外国游客用户 A 与本地用户 B 交谈，用户 A 不理解用户 B 的语言，可以在 IMS 网络中为用户 A 和 B 提供实时翻译服务。

网络侧在保持原通话的前提下：1、根据终端用户指示，执行翻译功能；2、实现将原语言的原媒体流，翻译成目标语言的文字；3、通话过程中，将双方语音转换成文字，并同步显示在主被叫用户终端。



图 8 实时通话翻译

(六) AR/XR 应用

继 AR 游戏、VR 视频直播/点播和 360 度视频直播/点播等应用后, AR 已率先在实时音视频领域开始商用, 5G 新通信可以提供 AR 社交、远程购物协助和 AR 远程维修等业务, 提高工作效率, 改善用户满意度。

在 XR 沉浸式应用场景下, 5G 新通信可以提供“更逼真”和“更直观”的业务, 通过个人穿戴设备(娱乐、游戏、运动)和行业应用网络(医疗、教育、中介等)结合, 让用户从中获得身临其境的感受, 适用于年轻个人用户和行业用户。

在线看房业务是 XR 业务的典型代表。就传统的看房方式而言, 由于受到各方时间和交通因素的限制, 每次看房都要兴师动众, 而且基本上无法一次完成, 导致成交率下降。

在线看房业务随时随地都可以进行, 不受时间和空间的制约, 无需见面, 只需通过事先录制或现场拍摄的 XR 影像资料就可以 360 度感受目标房屋, 并听取中介的语音介绍, 获得身临其境的看房体验, 提高交易的成功率。



图 9 在线看房业务

（七）手势翻译

通过 AI 技术，手势翻译可以实现语音和手语之间的实时互转，手势识别可以通过安装 API，集成 SDK 检测手部动作的视频或者图片，将手势信息上传到平台，翻译成语音或者文字进行交互。

业务平台完成实时翻译后基于手机呼叫界面展现翻译内容，实现视听全方位翻译同步，帮助聋哑人进行无障碍沟通。

手势翻译通过隔空手势，采用视觉识别技术，无需语音或文字输入，创造了一种更加多元化的交互方式。手势识别也可以应用于在线直播、在线教育互动、自动驾驶手势引导等场景。

（八）量子安全通话

传统的密码体制一般以数学原理为基础，在计算机信息处理能力越来越强的今天，有很大概率可以破解这些密码。量子密码是以量子物理原理为基础，采用量子力学原理，利用卫星实现量子密钥的分配。其优势主要有可证明安全性和对抗动的可检测性。

量子安全通话是基于量子信息技术的加密通话产品，使用终端、量子安全 SIM 卡和加密算法三重保护。用户每次发起“量子密话”都会随机抽取卡内密钥校验用户身份，认证通过后再由量子网络实时生成一个新密钥加密会话数据。密钥真随机，难破译，会话结束后密钥立刻被废弃，一话一密，无法破译。

在不降低原有业务话音清晰度的情况下，用户通过原生终端拨号盘即可直接拨打加密通话，操作便捷，接通迅速，无顿挫感，在保障通话高清品质的同时，极大增强通话的安全保密性，通话高清实时，切换便捷。



图 10 量子安全通话

五、总结和展望

中国联通从 2020 年开展 5G 新通信技术研究，实现了“5.17” VoNR 业务商用。2022 年，中国联通将继续开展终端集成 APP/SDK 新通信业务创新，试点业务迭代，联合产业计划发布支持 DC 的芯片原型机，芯片集成及实验室验证，预计 2022 年底部署新通信方案试点环境。2023 年开展规模业务落地和商用工作。下一阶

段 5G 新通信将至少攻坚三个关键问题：

1.简化网络：5G 新通信网络和基础承载网络相对独立，研发通过 5G 新通信媒体面和 IMS 媒体面融合，以及网络能力开放平台和智能业务能力平台融合的方式，简化网络架构，降低网络复杂度，提高运行效率。

2.开放网络：在 IMS 网络广泛使用 SIP 协议的基础上，5G 新通信网络引入 HTTP 协议和 RESTful 接口，同时进一步强化网络音视频和数据能力的封装，提升网络开放性，降低第三方业务商的进入门槛。

3.安全网络：5G 新通信网络安全方案综合考虑物理安全、数据安全、组网安全和供应链安全，针对不同客户的安全需求，提供差异化的分层安全保障机制，特别增加对客户关键数据的保护，消除客户的后顾之忧。

中国联通紧紧围绕“1+8+2”新战略，聚焦大联接、大应用等主责主业，协同产业伙伴发布 5G 新通信技术白皮书，为 5G 新通信业务发展提供参考技术方向和应用场景。未来中国联通还将探索媒体处理能力与芯片/终端解耦，以及 IMS Data Channel 技术向下一代演进。基础网络将使能更多的业务场景，随着 AR/VR、AI 等技术的引入，有望创造新的行业应用，进一步通过差异化的 5G 新通信业务体验，提升品牌知名度和市场，以促进产业发展和繁荣，进一步提升 5G 技术发展，为社会经济创造价值。

六 附录：缩略语

缩略语	全称	中文名称
AR	Augmented Reality	增强现实
AS	Application Server	应用服务器
APN	Access Point Name	接入点名称
ASR	Automatic Speech Recognition	自动语音识别
DC	Data Channel	数据通道
DCEP	Data Channel Establishment Protocol	数据通道建立协议
EPS	Evolved Packet System	演进分组系统
NLP	Natural Language Processing	自然语言处理
QCI	QoS Class Identifier	服务质量分类识别码
QoS	Quality of Service	服务质量
S-CSCF	Serving-Call Session Control Function	服务类型的会话控制功能
SIP	Session Initiation Protocol	会话发起协议
SR	Super Resolution	超分辨率
TCP	Transmission Control Protocol	传输控制协议
TLS	Transport Layer Security	传输层安全协议
TTS	Text-to-Speech	语音合成
UE	User Equipment	用户设备
UDP	User Datagram Protocol	用户数据报协议
UMF	Unified Media Function	统一媒体功能
VR	Virtual Reality	虚拟现实
VoLTE	Voice over Long-Term Evolution	长期演进语音承载
VoNR	Voice over New Radio	新空口语音
XR	eXtended Reality	增强/虚拟/混合现实的总称

七 参考文献

- 1 3GPP TS26.114, IP Multimedia Subsystem (IMS);
Multimedia Telephony; Media Handling and interaction
(Release 17);
- 2 CCSA TC3 WG1, 增强实时通信网络技术研究;



中国联通研究院

中国联通研究院是根植于联通集团（中国联通直属二级机构），服务于国家战略、行业发展、企业生产的战略决策参谋者、技术发展引领者、产业发展助推者，是原创技术策源地主力军和数字技术融合创新排头兵。联通研究院以做深大联接、做强大计算、做活大数据、做优大应用、做精大安全为己任，按照 4+1+X 研发布局，开展面向 CUBE-Net 3.0 新一代网络、大数据赋能运营、端网边业协同创新、网络与信息安全等方向的前沿技术研发，承担高质量决策报告研究和专精特新核心技术攻关，致力于成为服务国家发展的高端智库、代表行业产业的发言人、助推数字化转型的参谋部，多方位参与网络强国、数字中国、智慧社会建设。联通研究院现有员工近 700 人，平均年龄 36 岁，85% 以上为硕士、博士研究生，以“三度三有”企业文化为根基，发展成为一支高素质、高活力、专业化、具有行业影响力的人才队伍。

战略决策的参谋者
技术发展的引领者
产业发展的助推者

态度、速度、气度

有情怀、有格局、有担当

中国联合网络通信有限公司研究院

地址：北京市亦庄经济技术开发区北环东路 1 号

电话：010-87926100

邮编：100176



中国联通研究院



中国联通泛终端技术