

中国视频云场景应用洞察白皮书

2021年





发展环境：1) 我国通信能力不断提升，以视频为代表的富媒体信息成为主流。短视频不仅成为人们日常生活中的“时间杀手”，同时带动了营销模式的改变。2) 需求端：碎片化的消费模式、极致的观感需求、创新应用场景，让**短视频、超高清视频和实时音视频**成为需求风口。3) 供给端：视频云为行业提供技术支撑，AI赋能全链条，实现全民生产、全域应用，行业步入视频服务社会的新时代。



关键技术：视频云进一步延伸服务链条，对视频全流程进行全方位支撑。1) 生产环节：智能媒资、智能生产快速找出相关创作素材，提升后期**创作效率**。2) 处理环节：超清转码和窄带高清大幅提升**画质和降低成本**。3) 传输环节：CDN向边缘云发展，云边协同降低**视频时延**、节省**带宽成本**。4) 消费环节：视频特效丰富**视频体验**，智能推荐算法实现**精准营销**。



应用场景：视频云渗透率最高的场景在**互联网泛娱乐**以及**在线教育**。资源和技术门槛的存在加之应用玩法的快速迭代均凸显了视频云的价值。云服务与产业互联网的结合不断加深，视频云厂商场景化、私有化能力加强，能更好地匹配传统行业的部署需求，在传媒、教育、平台电商化、协同办公等领域的应用也正变得越来越多。



发展趋势：1) 视频云技术将在视频处理和传输环节发力，实现**技术极致化**。除了精研软件算法，同时积极与硬件厂商合作，实现性能优化和成本下降。2) **低代码**开放模式持续在音视频行业渗透，向企业提供面向全场景的音视频互动一站式解决方案。3) **高新视频**新业态，让视频赋能千行百业的能力得到充分的发挥。

新时代下的视频云

1

云上视频新场景

2

视频云发展趋势展望

3



新时代下的大视频产业：

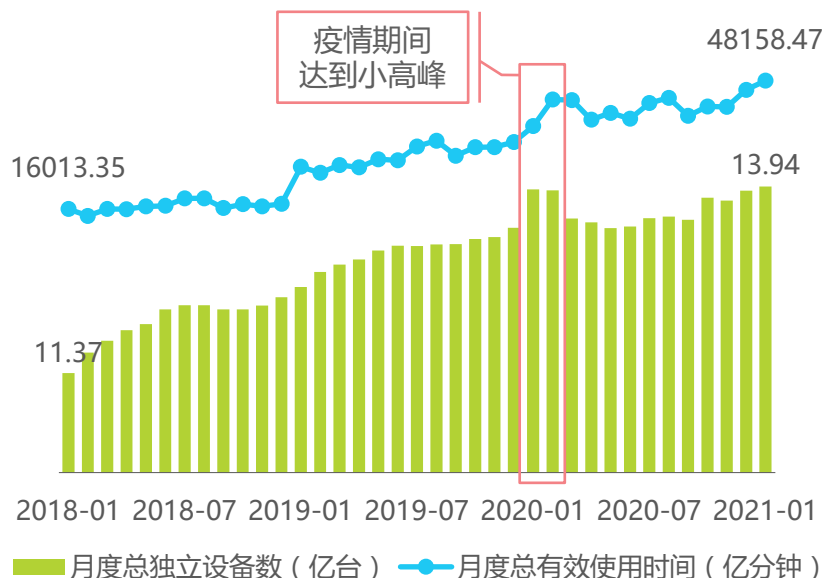
“内容视频化、交互线上化”

大视频产业发展环境

通信基础设施成熟，以视频为代表的富媒体信息成为主流

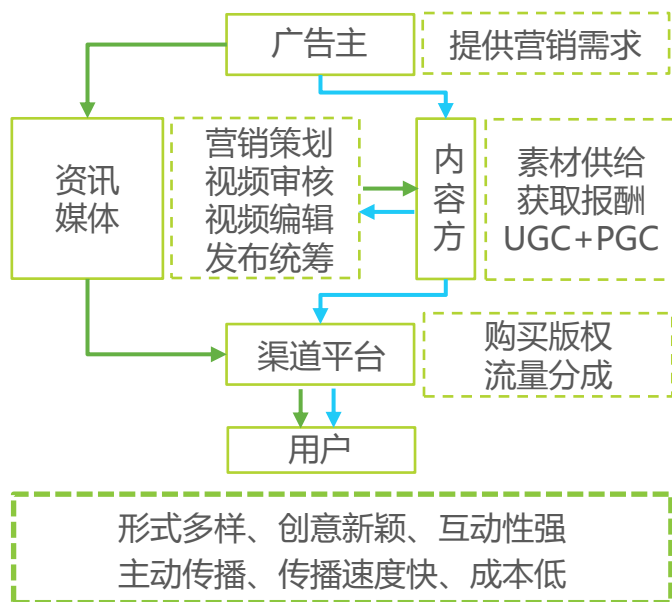
近年来，我国持续推进网络基础设施建设，以5G、数据中心、人工智能、物联网等为代表的信息数字化能力不断提升。根据工信部数据，截至2020底，我国新建5G基站超60万个，全部已开通5G基站超过71.8万个，5G基础设施网络覆盖全国地级以上城市及重点县市。移动通信技术的发展不仅带来网速的提升，同时推动流量单位成本下降，“提速降费”让视频服务类App的用户量和活跃度不断增加。艾瑞UserTracker监测数据显示，自2018年以来，视频服务App月总独立设备数和月使用时间虽偶有波动，但整体呈现稳步增长趋势。在各细分领域中，短视频以承载量大、内容多元兼趣味化、传播力强为优势，市场规模不断扩大，同时带动了营销模式的改变，短视频逐渐成为新型的营销手段与分发渠道。

mUserTracker-2018年1月-2021年1月
视频服务类App月独立设备数及总有效时间



来源：UserTracker 多平台网民行为监测数据库（桌面及智能终端）。

短视频营销模式及优势



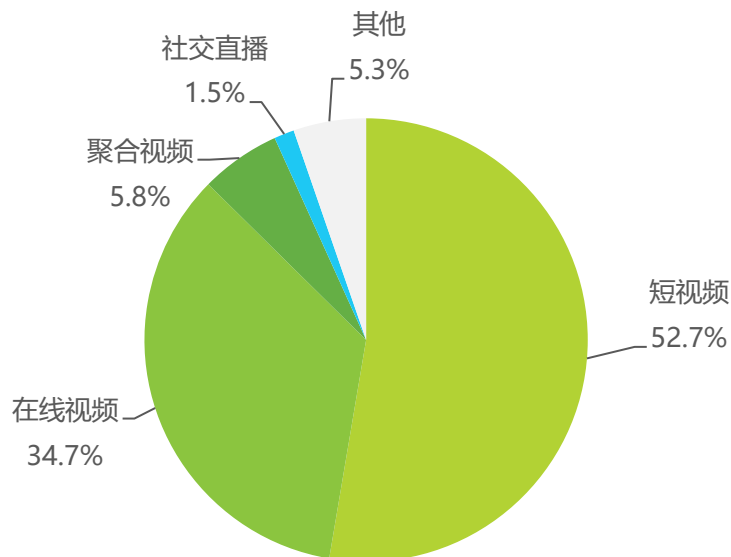
来源：艾瑞咨询研究院自主研究及绘制。

大视频产业需求端分析（1/3）

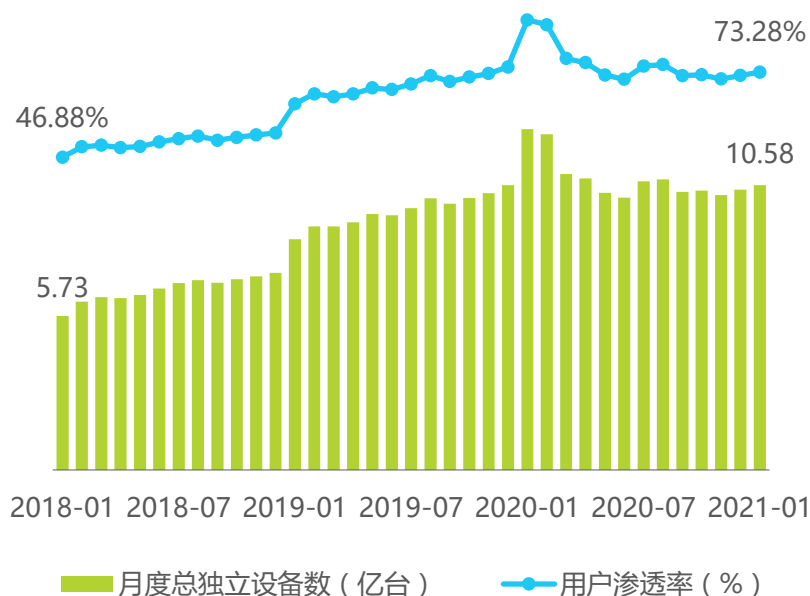
短视频强势崛起，视频消费趋于“短、频、快”

受消费习惯的影响，短视频市场需求旺盛。艾瑞UserTracker监测数据显示，视频服务类App用户使用的内容形式中，短视频位列首位。伴随5G网络的普及，短视频用户规模持续增长，移动网民端短视频渗透率达到73.28%。短视频带来了全新的节目形态和传播场域，小屏幕、碎片化、快节奏成为行业潮流。在这种趋势下，视频消费习惯的改变对短视频生产端提出新的要求：更短的制作周期确保更新频率、画质清晰、剪辑节奏流畅、内容策划紧跟热点热度等。

mUserTracker-2021年4月视频服务类App
月总有效使用时间比例



mUserTracker-2018年1月-2021年1月
短视频App月独立设备数及用户渗透率



来源：UserTracker 多平台网民行为监测数据库（桌面及智能终端）。

来源：UserTracker 多平台网民行为监测数据库（桌面及智能终端）。

大视频产业需求端分析 (2/3)

高清迈向超高清，视频功能创新、体验升级

2019年2月28日,工业和信息化部、国家广播电视总局和中央广播电视总台部门联合发布了我国的超高清产业发展计划——《超高清视频产业发展行动计划(2019-2022年)》，明确将以“4K先行、兼顾8K”的总体技术路线，大力推进超高清视频产业发展和相关领域的应用。政策利好推动超高清视频产业发展，同时也对技术和服务提出了更高的要求。除了要满足用户还原真实的观感需求，还需不断探索并服务于新的应用场景：远程医疗、安防监控、工业制造等，应用端利用超高清技术更快、更准确、更高效的完成专业作业。超高清在各领域应用的加速推进，将使其步入发展的黄金时期。

超高清视频需求分析及产业链各层次市场规模

2020年，我国超高清视频产业总规模达到1.8万亿，其中：服务层和设备层的市场规模达8100亿元；应用层市场规模9800亿元。预计到2022年超高清视频产业总规模将达4万亿元。

政策推动

- 国家政策要求大力推进超高清视频产业链各环节共同发展。
- 地方政策加速跟进，北京、上海等地陆续发布地方超高清视频产业发展行动计划，明确主要任务。

用户观感需求提升

- 图像清晰度
- 图像流畅度
- 颜色渐变平滑度
- 颜色丰富度
- 明暗对比度
- 立体临场感
- 场景空间感
- 更贴近真实

视频应用场景多样化

- 自由视角、360度全景、沉浸式视频需要超高清视频技术支撑。
- 场景需求愈加广泛，目标用户除了电视观众，已延伸至专业领域。

➤ **设备层**：包括视频生产设备、网络传输、终端呈现设备、核心元器件，超高清视频带动设备生产数量的提高以及传输和芯片技术的提升。

➤ **服务层**：超高清对内容制作、内容版权提出更高的要求，建立良好的平台服务和版权保护生态，形成内容制作盈利闭环。



应用层：市场规模占比大，各领域需求旺盛

- 文教娱乐：直播、教育、云游戏要求还原画面真实感、带来沉浸式体验。
- 工业制造：自动化涉及精细识别、精密测量。
- 安防监控：前端高清成像准确还原监控场景。
- 智能交通：交通管控、行为和步态精确化识别。
- 医疗健康：精准治疗、精确诊断需要视频呈现微小图像细节。

来源：超高清市场规模来自中国电子信息产业发展研究院、中国超高清视频产业联盟政策工作组《超高清视频产业发展白皮书（2021年）》，艾瑞咨询研究院自主研究绘制。

大视频产业需求端分析（3/3）

强调视频的实时交互，视频通话、视频会议兴起

疫情爆发推动了实时交互视频的飞速发展，高带宽消耗的视频通讯从高端消费层下沉，视频通话、视频会议、在线教育等成为人们生活、工作和学习中的“刚性需求”。疫情期间，企业通过协作办公确保业务连续性，视频通讯成为企业和员工间主要的沟通方式；在教育领域，“停课不停学”的政策鼓励学生利用线上平台教学，而教学质量取决于师生间互动质量的优劣，此时在线互动教育平台顺应了主动化、多元化、开放性的需求；在医疗领域，为了缓解资源缺乏和分布不均的问题，在线医疗迎来爆发，通过远程问诊、线上复诊等分散线下压力，开发新用户的同时提升用户粘性。此外，实时音视频向社交娱乐、金融服务、公用事业等领域渗透，应用场景日渐丰富的同时对画面质量、延时、抖动提出更高要求。

实时音视频应用场景

社交娱乐

单人私聊、多人群聊
游戏社交、荷尔蒙社交



会议协作

内部沟通、外部沟通
员工培训、售后服务



在线教育

双师课堂、小班教学
互动教学、K12教育



良好的用户体验提升用户粘性和支付意愿。



在线医疗

医患沟通、远程问诊
病例研讨、手术观摩示教



金融服务

远程定损、视频双录
线上开户、客户服务



公用事业

民生政务、视频报警
远程接访、可视化执法

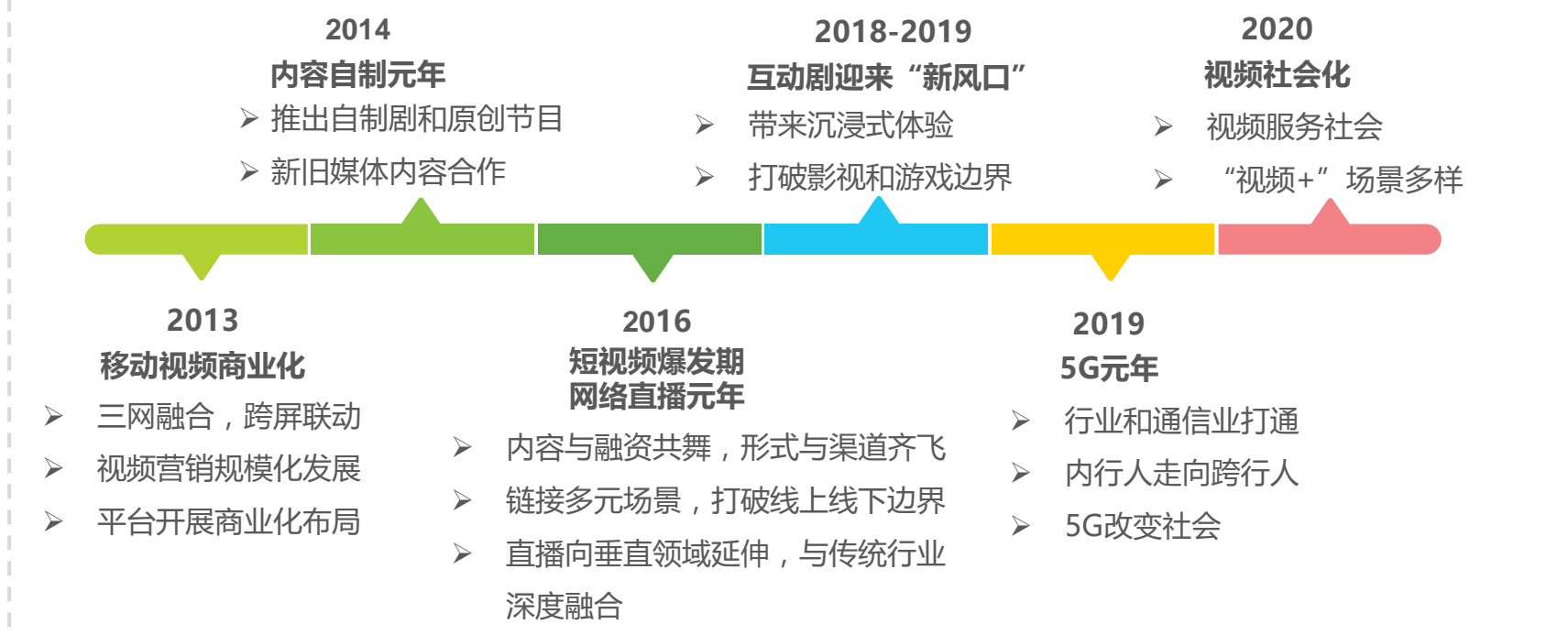
用户深度参与和紧密协作促进创新，推动办公运转的敏捷性。

大视频产业供给端分析（1/3）

视频社会化时代，产业边界消融、应用场景拓宽

从2006年网络视频的兴起，到现在步入视频社会化时代，视频不再局限于电影、电视、广告等传统媒体领域，音视频通话、互动直播等新型应用使产业边界逐渐消融。内容上，网络自制剧、短视频的出现打破传统媒体的垄断地位，迫使其加速转型。形式上，音视频通话、直播以互动性强的特点，向电商、游戏、办公等垂直领域延伸。技术上，“5G+AI+4K”的战略布局推进了新旧媒体一体化发展，借助电视媒体在音视频制作方面的优势，提升高品质新媒体制作能力。同时，技术进步也降低了视频生产的成本和技术门槛，为加速推进传统行业的智能化裂变做出贡献，使应用场景进一步拓宽。

视频行业迈向“视频社会化时代”



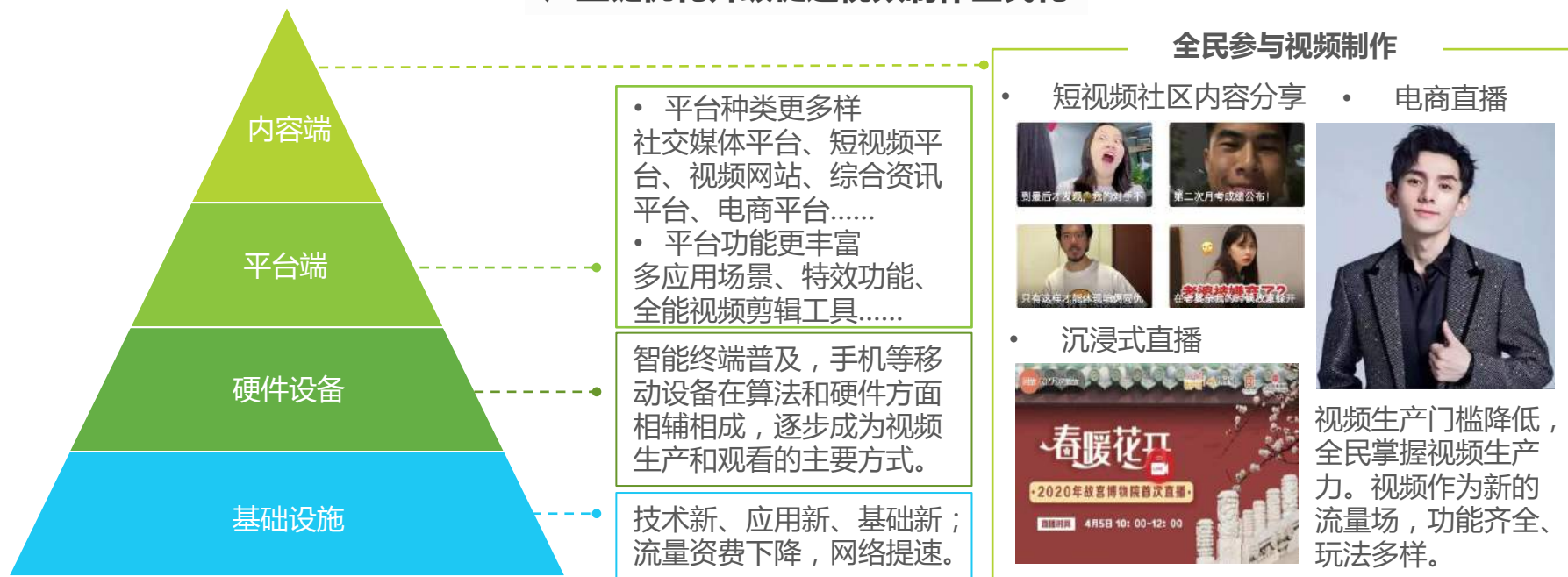
来源：艾瑞咨询研究院自主研究绘制。

大视频产业供给端分析（2/3）

大视频产业链持续优化升级，全民参与视频制作

智能终端的普及和流量资费的下降，为大视频产业提供良好的发展环境，人们逐渐习惯以直播、短视频作为信息交流的媒介。视频拍摄无需依托专业的硬件设备，手机的拍摄功能即可满足需要，硬件和软件相辅相成，操作简单、成像效果清晰。手机摄影方便创作者记录日常生活中的每一个瞬间，题材还原生活本质，增强纪实性，也让内容更加丰富。与此同时，各类型视频平台为了争夺市场，不断更新平台功能、拓展应用场景、降低生产门槛，从而吸引更多的流量和内容供应。网络基础设施、拍摄设备和平台端的优化升级为视频内容创作者提供了肥沃的土壤，万物皆可拍、万物皆可播、全民可参与。

产业链优化升级促进视频制作全民化



大视频产业供给端分析（3/3）

视频云提供底层技术支持，助力各行业向视频领域跨界

视频云在智能生产、智能分发以及通过音视频处理保障流畅度等方面提供强大的技术支持，对视频生产供应、传播和跨领域应用起到重要作用。首先，视频智能生产赋能视频生产环节，帮助创作者高效处理和制作视频内容，加上视频AI的多种能力，能够在大型直播场景下实现内容识别、定位、剪辑、合理聚合，高效产生精彩素材。其次，在视频传播方面，按照内容分发特点可分为强中心化和去中心化两种模式。去中心化将内容选择权留给用户，智能标签实现内容快速检索，提升资源检索效率和准确性。在强中心的传播途径中，视频内容分析和智能标签实现精准、个性化的内容推荐，达到千人千面的效果。最后，实时音视频能力保障稳定流畅的观看体验，推动视频行业跨界融合，在传统媒体、电商、医疗、教育等领域广泛应用。

视频云厂商赋能各行业向视频的跨界发展



来源：艾瑞咨询研究院自主研究绘制。



新时代大视频产业的底座：

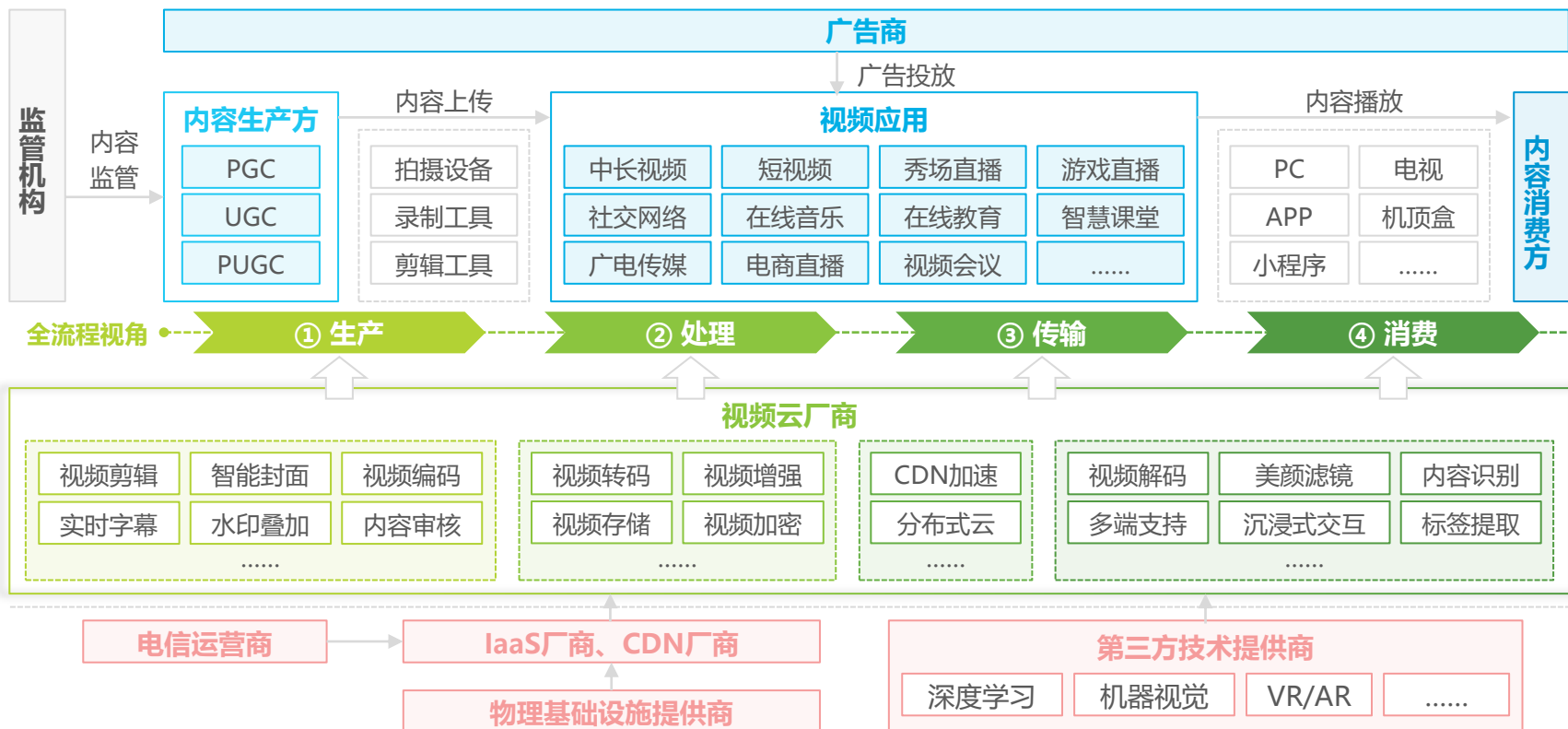
视频云

视频云产业链拆解

降低准入门槛，全方位支撑视频的生产、处理、传输及消费

对于竞争激烈、快速迭代的大视频产业而言，视频云已经逐渐发展为一项关键的基础设施。经过多年发展，视频云由最早期的聚焦CDN加速到以PaaS服务助力视频应用快速开发，进一步延伸服务链条，对视频从生产到处理、传输以及最终消费的全流程进行全方位支撑。视频云不仅大幅降低了视频应用的准入门槛，更通过促进产业效率提升不断推动大视频产业的繁荣。

中国视频云产业链拆解



来源：艾瑞咨询研究院自主研究及绘制。

视频云产业图谱

2021年中国视频云产业图谱



注释：行业解决方案商分类依据厂商所聚焦的核心场景进行划分。
来源：艾瑞咨询研究院自主研究及绘制。

视频云关键技术解析：生产环节

内容管理与搜索识别审核降本，智能化处理提升后期创作效率

视频生产是视频产业链的内容产出环节，主要基于视觉与深度学习技术，在素材采集后提供视频内容的识别、编辑、生成、增强与摘要等能力。视频生产可广泛应用于互联网媒体、短视频、娱乐直播、在线教育、广电传媒等行业应用。近年来，通过智能媒资系统，使用多模态搜索、智能拆条、智能审核等AI处理技术对各类采集源的视频素材进行处理，客户可以轻松高效地从海量视频中找出自己所关注的目标，快速找出相关创作素材，提升后期创作效率。

视频云关键技术解析：生产环节



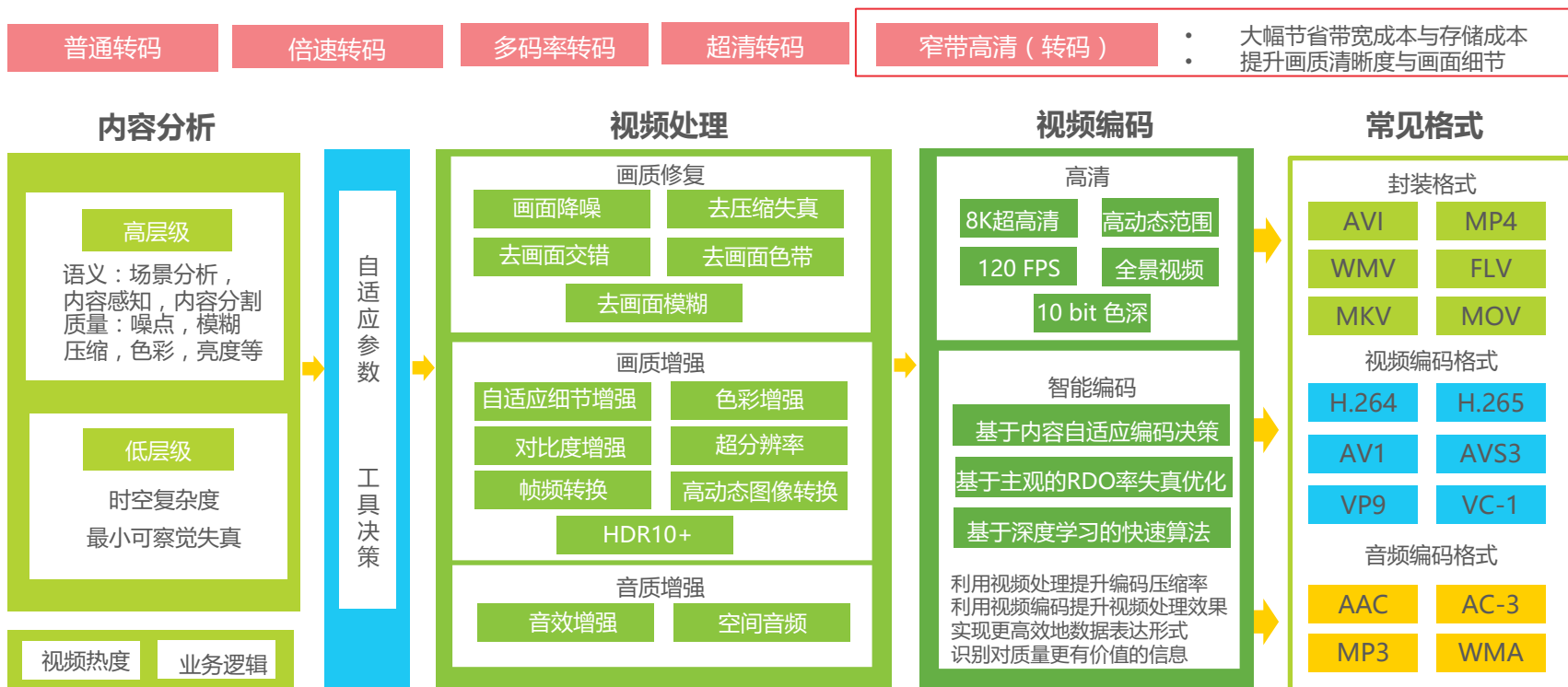
来源：艾瑞咨询研究院自主研究及绘制。

视频云关键技术解析：处理环节

超清转码和窄带高清转码发展，大幅提升画质和降低成本

视频编码的主要作用是对视频像素数据压缩，从而降低视频的数据量。高效率的视频编码在同等的码率下，可以获得更高的视频质量。转码处理主要是针对海量媒体资源提供高效、智能的服务，包括普通转码、倍速转码、超清转码、窄带高清转码等。其中窄带高清技术在编码前先对内容进行分析，在不降低甚至增强源视频画质的基础上，对视频分辨率、流畅度、码率、帧率等进行算法优化，大幅提升画面清晰度和转码速度，同时节省带宽和存储成本。

视频云关键技术解析：处理环节



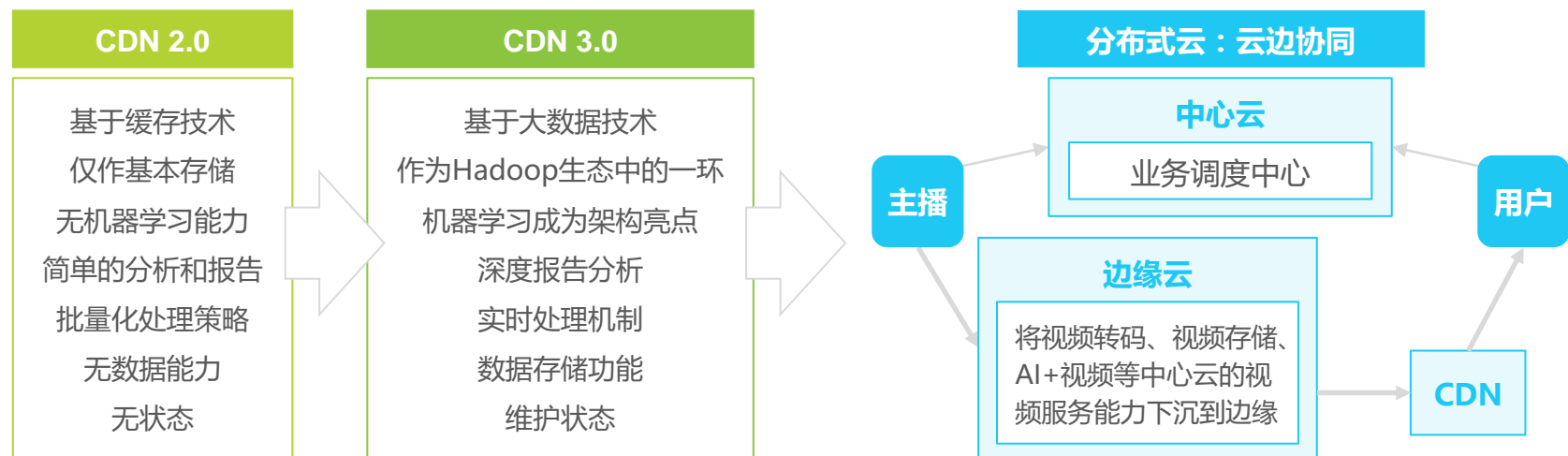
来源：艾瑞咨询研究院自主研究及绘制。

视频云关键技术解析：传输环节

CDN智能加速，云边协同降低传输时延、节省带宽成本

CDN即内容分发网络，通过将视频内容发布到最接近用户的网络边缘，提高用户访问时的响应速度。发展至今，CDN已由2.0阶段过渡到3.0阶段，从单纯以缓存边缘内容为目标步入到运用机器学习实现智能调度、智能分发，从而有效降低运营成本、提升传输质量。但传统的CDN通常只下沉到地市的城域网，距离终端用户依然比较远，在流量高峰时段，CDN资源紧张会影响用户使用体验。随着5G和边缘计算技术的逐步成熟，传统CDN开始向边缘云发展，通过将中心云的视频服务能力下沉到边缘，在边缘云上进行转码、分发，可以进一步降低视频时延、节省带宽成本，满足大规模、实时互动的需求。

视频云关键技术解析：传输环节



更精准的调度：由地理就近性调度转变为基于现网带宽、负载、成本等多种算法策略的综合调度

更智能的分发：对现网热点内容数据特征进行机器学习，预判内容热度，指导CDN智能预分发

更先进的运维：根据告警数据特征进行故障定位和原因分析，系统自动或人工障排除

关键技术参数：

时延、抖动率、丢包率、卡顿率、并发量

视频云关键技术解析：消费环节

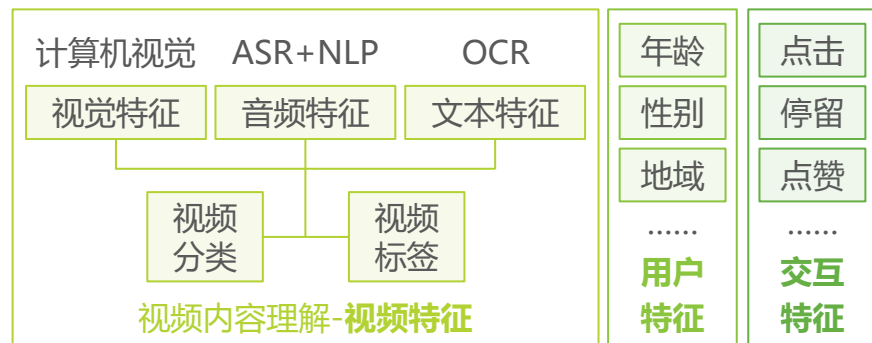
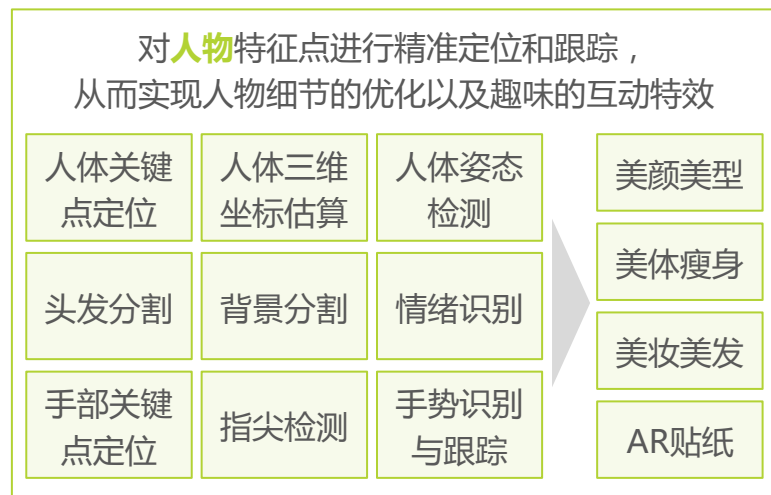
多样化玩法丰富客户端体验，智能化加持视频精准营销

在视频消费环节，视频云可以提供包括美颜特效、美声变声特效、沉浸式交互等在内的多样化玩法，以丰富终端用户的使用体验。视频特效多以人为中心，关键点定位、识别和跟踪是其技术核心，在此基础上可以对人物细节的优化处理，或添加AR贴纸等互动特效。基于视觉特征、音频特征、文本特征进行视频内容理解，并与用户特征和交互特征结合建模，可用于预测用户在另一视频数据集上的消费行为，从而实现视频内容的精准推荐与营销。

丰富视频体验

视频云关键技术解析：消费环节

智能视频营销



常用的视频推荐算法

内容推荐算法：将与用户喜欢和自己关注的item在内容上类似的item给用户，可避免冷启动问题，但可能会推荐重复的item

协同过滤算法：挖掘与用户具有相似兴趣的用户喜欢过的item进行推荐，随交互记录增加将更加准确，但初期存在冷启动问题

混合推荐算法：融合以上方法，通过加权、串联、并联等建模

变调算法中时域法较常见，采用变速不变调处理与重采样相结合的方式实现变调不变速的**人声**效果

语聊美声

歌唱变声

音色转换

变声特效

新时代下的视频云

1

云上视频新场景

2

视频云发展趋势展望

3

视频云应用场景总述

由互联网泛娱乐出发，逐步向更广泛的行业领域渗透

目前来看，视频云渗透率最高的场景依然是在互联网泛娱乐以及在线教育。视频属于这两类场景的核心业务，资源和技术门槛的存在加之应用玩法的快速迭代均凸显了视频云的价值。随着网络基础设施的优化，人们越来越习惯利用视频进行日常沟通、信息分享。在短视频风口的助推下，视频演变为各类互联网应用的标配，视频云自然而然地向游戏、社交、K歌、长音频等非视频服务应用渗透。现如今，云服务与产业互联网的结合不断加深，视频云厂商场景化、私有化能力加强，能更好地匹配传统行业的部署需求，在传媒、教育、平台电商化、协同办公等领域的应用也正变得越来越多。

中国视频云应用场景渗透率路径



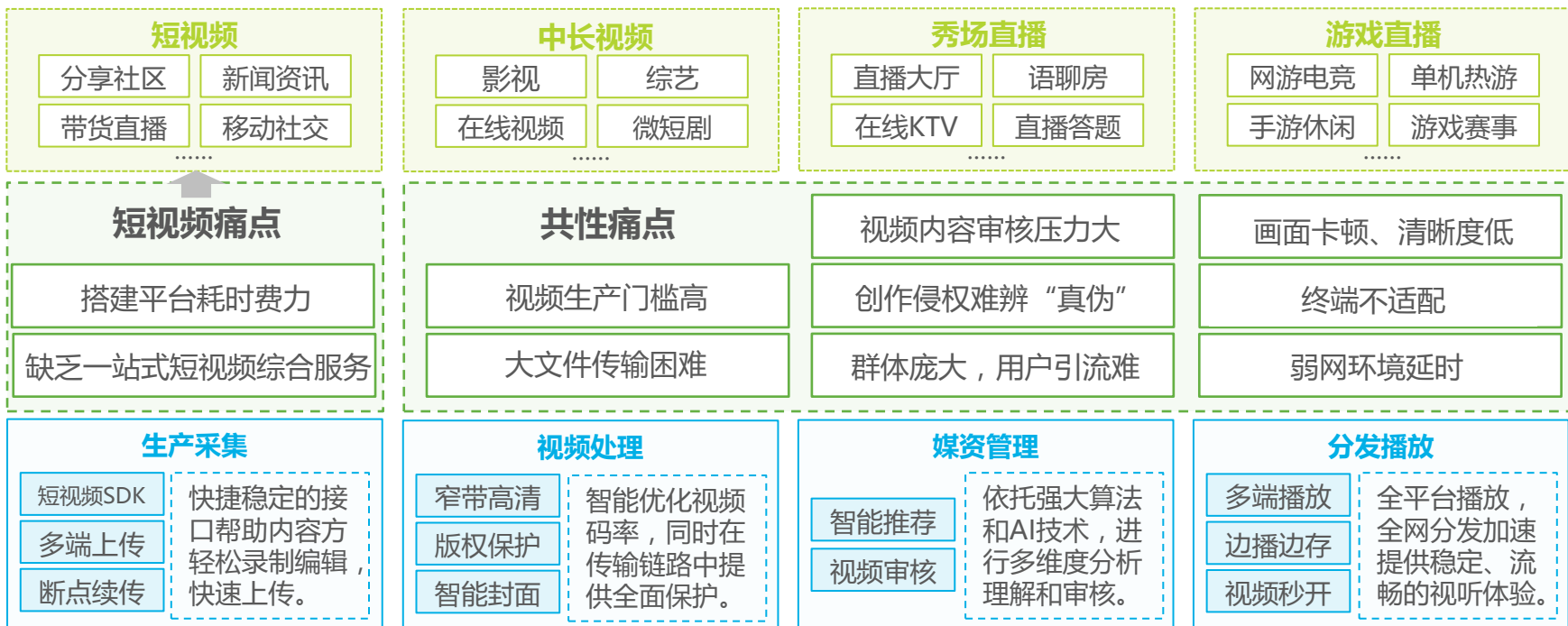
来源：艾瑞咨询研究院自主研究及绘制。

互联网泛娱乐场景需求分析

一站式综合服务打造全民化、智能化互联网泛娱乐平台

互联网泛娱乐主要以中长视频、短视频、秀场直播以及游戏直播应用场景为细分，具备用户群体大、涵盖内容丰富、覆盖场景多样的特点。面对内容生产方自身水平和所处网络环境的差异性，视频云能够提供拍摄剪辑、特效编辑和高速上传的解决方案，并基于智能化编转码技术和媒体资源管理能力，将清晰流畅的视频内容准确推送给目标人群，提升用户观看体验。在短视频领域，除了在产业链各环节存在与其他应用场景共同的应用痛点，还面临着平台搭建和运营所涉及的成本及效率问题。视频云能够提供一站式的短视频综合服务，快速搭建平台、满足基础功能需求，让客户能够专注核心业务创新。

互联网泛娱乐场景：视频云需求痛点与解决方案



来源：艾瑞咨询研究院自主研究绘制。

互联网泛娱乐场景应用案例：Kumu

菲律宾移动社交平台Kumu利用视频云快速打开短视频市场





Kumu是菲律宾领先的移动社交平台，自2018年成立以来积累了大量的注册用户，运营方式主要以直播为主。2020年，疫情限制了人们的出行自由，但是却让Kumu捕捉到用户对短视频服务的强烈需求。为了增加活跃度以及用户粘性，Kumu希望为平台增加UGC短视频的板块，在不花费大量开发成本的前提下，简单和快速的搭建短视频业务。为解决搭建耗时长和网络稳定性差等问题，阿里云为Kumu提供客户端短视频SDK和CDN分发的解决方案，帮助客户快速在自有平台上搭建短视频业务，并保证流畅的上传和播放体验。

Kumu×阿里云：一站式短视频社交



背景介绍

- 构建 SDK 任务艰巨、昂贵且耗时，需要各种开发资源，需要专家团队来创建、部署和维护。
- 需要领先于用户需求，抢占市场先机。

客户需求

	快	快速无缝部署短视频应用程序
	易	提供简单的界面、可定制的UI以满足业务需求。
	惠	开发、维护、部署成本最小化
	稳	视频的起播要求高，在东南亚网络不稳定的情况下提升播放体验。

解决方案

	SDK	实时滤镜	动画贴纸	MV效果
		实时美颜	断点续录	气泡字幕
	CDN	网络优化	✓ 快速满足视频内容分发播放需求。	
		播放端处理	✓ 视频达到秒开效果，提升播放体验。	

应用效果

- 一站式的接入能力，基础设施投资少，降低开发门槛，帮助客户快速验证新业务。
- 在生产端提供市面主流短视频功能，帮助用户快速生产个性化内容，增强社区粘性以及丰富社区内容。

互联网泛娱乐场景应用案例：快手

推出云剪云直播，赋能创作者高效产出优质内容

随着短视频在互联网泛娱乐场景中的广泛应用，用户对视频质量和直播技术支持提出更高的要求。对于创作者而言，原始的视频生产工具已不能满足专业化、多元化的需求，加之生产耗时长、素材处理困难等问题，亟需量身定制的新工具协助生产。快手推出两款全新的免费视频工具：云剪和云直播。通过深入视频制作每一环的“云”，向专业自媒体、MCN机构、媒体机构等专业和准专业创作者提供易用、高效、便协作的云端产品。

快手：短视频云剪云直播



来源：艾瑞咨询研究院自主研究绘制。

教育场景需求分析

实时音视频RTC提升互动性，助力在线课堂提升教学质量

教育行业覆盖的用户群体广泛，从低龄儿童到中老年，从学生到上班族，用户需求多样化。线下教育机构向线上转型不仅需要投入巨大的花销，而且会遇到开发难度层面、集成对接以及扩容速度等方面的困难。视频云从基础服务出发，通过实时音视频RTC技术在流畅度、成本、时延等需求间实现平衡，实现用户体验的整体提升。此外，视频云厂商还通过设计多种课堂功能，运用AI能力和版权保护等附加功能为在线教育提供更多的可延展性。

应用场景

1V1

1V多小班课

1V多大课堂

大型公开课

双师课堂

AI互动课

客户类型

低幼及素质教育
K12学科培训
高等学历教育
职业和成人教育

教育场景：视频云需求痛点与解决方案

痛点

成本不可靠控

场景不齐全

集成不便捷

扩展不灵活

功能不易用

互动不及时

底层不稳定

数据不安全

视频泄露

版权侵权

解决方案

基础服务

RTC

点播

白板

直播

媒体处理

实时消息

课堂功能

互动课堂 互动白板

互动签到 学生连麦

随堂考试 录制回看

问答功能 课堂管理

数据统计

AI功能

注意力管理 人像悬浮

虚拟老师 发音识别

音准节拍识别 体态识别

手势捕捉 无障碍处理

智能字幕 随堂报告

其他辅助功能

防盗链 播放鉴权

视频加密保障视频安全

教育场景应用案例：学信网

学信网开通远程复试系统，在线考试效能获提升

学信网全称“中国高等教育学生信息网”，由全国高等学校学生信息咨询与就业指导中心主办。学信网在疫情期间响应教育部关于远程复试的相关政策，快速上线了研究生招生远程面试系统，每年服务考生约150万人次。在构建平台的过程中，由于距离研究生复试时间短，且团队无视频应用开发经验，因此解决方案需简单易操作并支持快速开发集成。同时，防作弊、强监管以及支持高并发的视频通信也是在线考试系统的核心诉求。阿里云依靠过硬的企业级音视频通信服务保障能力和丰富的视频会议服务经验，保障学信网研究生视频面试稳定运行，提升复试工作效率。

学信网×阿里云：在线考试



视频云应用效果

3199

覆盖考场数量

9758

最高并发连接数

43%

收到待录取时间同比缩短

来源：艾瑞咨询研究院自主研究绘制。

教育场景应用案例：浙江大学

智云课堂助力“网上浙大”全面开课，智慧教育走向常态化

浙江大学“智云课堂”以智能媒体处理平台为基础，综合应用视频媒体能力与AI能力，实施“地对空”在线课堂新模式，即利用新建的录播教室向智云课堂推流进行在线直播，并同步存储课程视频为点播文件。过程中同时实现对所有教室信号的AI智能分析，并提供同声传译、AI课堂知识解构、AI协同笔记等多个功能模块。阿里视频云强大的底层技术支撑，提升课程视频质量及观看体验，联合华栖云提出的智能化解决方案进一步提高教学精准性。2020年2月24日，浙江大学全面网上开课，线下资源快速上线，实现教育常态化。

直播教学过程中的AI特色应用

浙江大学×阿里云&华栖云：智云课堂

资源分析、智能探测及制作



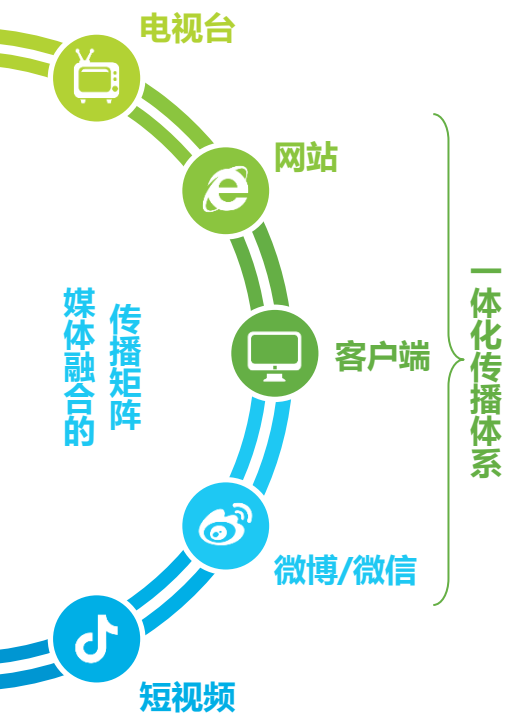
来源：艾瑞咨询研究院自主研究绘制。

传媒场景需求分析

媒体融合纵深发展，媒介形式升级倒逼生产流程重塑

移动互联网冲击传统媒体业态，传媒行业以媒体融合为关键词进行转型。经历了初期的产品融合和平台打造之后，各层级媒体开始构建打通台、网、端、微、视等多种资源的一体化、全媒体传播体系。传统媒体行业拥有一套成熟且模式化的策、采、编、发流程，严谨有余但效率不足，为适应新媒体的媒介形式与消费习惯，需要向更加高效智能的生产模式演进。视频云不仅可以提供分布式的内容分发网络，通过超清转码、窄带高清、画质重生等满足传媒行业高质量视频更低时延、更低成本的传播需求，更能够以云端一体化的解决方案帮助传媒行业重塑生产流程，提高产业效率。

传媒场景：视频云需求痛点与解决方案



时政新闻
体育赛事
剧集综艺
晚会直播
4K/8K超高清
互动直播
慢直播
.....

- 传统媒体的IT架构难以适应媒体融合背景下的全网内容传播
- 直播、录制、生产、发布的链路隔离，缺乏云端一体化解决方案
- 高清视频传输效率低，海量素材无法实时共享，多人多地协同编辑、审查、修改困难
- 传媒行业注重内容时效性，人工进行策、采、编、发的压力大、效率低
- 新媒体时代，短视频创作者大多缺乏版权意识，侵权方式多样且隐蔽

视频云解决方案关键点

公有云CDN + 全域边缘节点

面向大型活动、体育赛事、晚会演出，提供高可靠、低延时、高并发的直播服务

超清转码 窄带高清 画质重生

满足传媒行业超高清内容的传播需求，优化视频画质，降低视频码率，节约带宽成本

内容库 视频拆条 智能剪辑

云导播台 语音识别 文字校对

虚拟主播 智能审核 版权保护

构建云端一体化的视频采编平台，运用AI技术提升剪辑、编辑、审核的工作效率

来源：艾瑞咨询研究院自主研究及绘制。

传媒场景应用案例：人民日报

云上AI编辑部集合策采编发，赋能两会高时效、全媒体报道

为构建“AI+5G+4K/8K+VR”的全媒体采编生产新格局，人民日报技术社积极同头部科技公司展开交流合作，在语音转文字、新闻拆条、虚拟主播、视频打标签等方面进行探索。为保障两会报道的时效性与首发率，人民日报与阿里云共创AI编辑部，直击传统媒体生产效率低、传播渠道窄的痛点。AI编辑部集“看”“听”“悟”“审”“查”能力于一身，深度贴合前方记者“多栖作战”的采编需求，现场即拍即编、快速发布，全面赋能人民日报社“报、网、端、微、屏”全媒体两会报道，并持续助力报社“服贸会”“进博会”等重大报道活动。

人民日报×阿里云：云上AI编辑部



来源：艾瑞咨询研究院自主研究及绘制。

传媒场景应用案例：中央广播电视总台 iResearch 艾瑞咨询

基于超高清在线虚拟包装系统，打造欧洲杯数据视觉盛宴

为创造高品质观赛体验，广播电视和体育赛事积极融合创新，探索VR、AR、5G+4K/8K超高清的新形态、新模式。现下2020年欧洲杯激战正酣，中央广播电视总台在欧洲杯专题栏目演播室中引入新奥特全程提供超高清图文演播的技术服务，打造了一场欧洲杯数据的视觉盛宴。通过应用“石墨”超高清在线虚拟包装系统，中央广播电视总台实现了超高清虚拟植入和超高清大屏展示，一方面利用传感定位、虚拟分层渲染合成、GPU色键等技术，将生成的虚拟AR场景精准植入到转播现场实景之中，一方面通过对接内外部数据源进行项目播控，实现对赛事数据的可视化、直观化呈现。

中央广播电视总台×新奥特：虚拟包装

新奥特·石墨——超高清在线虚拟包装系统

① 超高清虚拟植入

虚拟AR+在线包装无缝结合，虚拟场景与包装场景可同时存在、叠加播出。

传感定位

图像识别技术结合传感器数据，自动定位摄像机在虚拟场景中的位置和姿态，搭建时间短，跟踪精度高，确保了虚拟植入场景的准确定位

虚拟分层渲染合成

输入视频、抠像视频和多个虚拟图层渲染共同组成虚拟演播、虚拟植入的场景，并可与其它包装图层叠加播出，融为一体，实现丰富的演播效果

GPU色键

包括亮度和阴影处理的高质量抠像，保证真实人物在虚拟场景中的“落地感”；支持三级色键级联、分工逐级抠像及对多区域、多颜色的抠像



② 超高清大屏展示

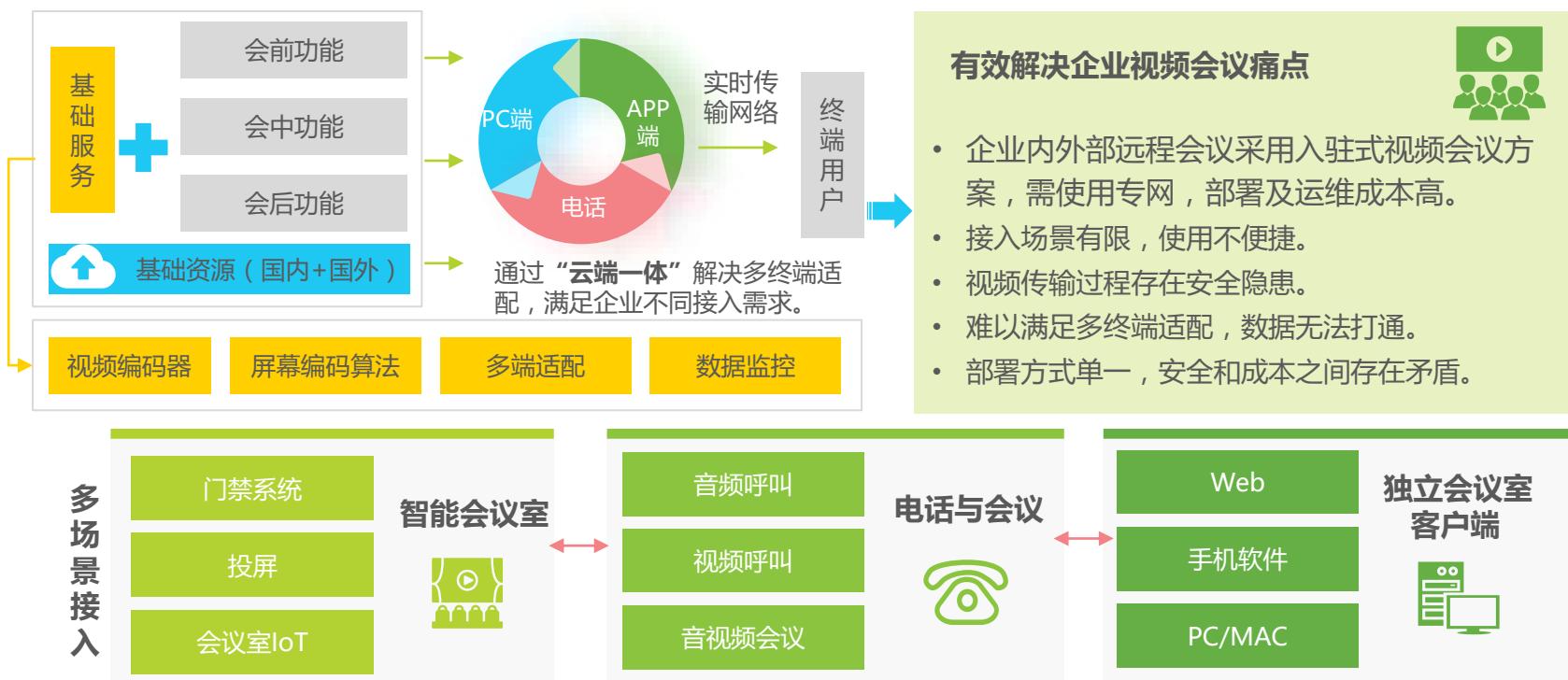
提供大量基于数据的图文场景，支持外部实时赛事数据输入和处理；将赛事场馆、赛程预告、赛事积分榜、球员情况、球队阵型等数据内容可视化、直观化，为演播室节目拍摄和制作提供灵活多变的制作手段。

协同办公场景需求分析

“云端一体” 满足企业不同接入需求，实现随时随地入会

视频云在协同办公应用中通过稳定可靠的IM及音视频通信底层能力，为客户提供开放、可靠、智能的会议服务。视频云厂商基于视频会议技术，提供公有云、私有云、混合云的多模式部署，帮助企业快速构建多端互通与简单易用的会议协作应用。此外，覆盖Windows、Mac、Android、iOS、Web等多端的SDK可以满足不同终端用户的接入需求，在各端APP内可直接发起会议，通过“云端一体化”解决方案帮助客户实现随时随地入会，提升沟通效率。

协同办公场景：视频云需求痛点与解决方案



来源：艾瑞咨询研究院自主研究及绘制。

协同办公场景应用案例：越秀地产

打通视频会议系统，满足企业复杂多样的会务场景需求

越秀地产成立于1983年，1992年于香港上市（股票代码：00123HK），是全国第一批成立的综合性房地产开发企业之一、中国第一代商品房的缔造者，也是中国第一家拥有香港房地产投资信托基金的内地房企。截止2020年底，公司总资产超过2600亿元。

2019年4月起，越秀地产同钉钉、阿里云、蓝凌开展四方共创，打造了全连接、数字化和智慧化的移动集中业务处理平台—悦工作（钉钉大企业专属版）。视频会议方面，越秀地产日常有经营分析会、方案研讨会、条线工作会、培训学习等需求，通过将悦工作与华为视频会议系统打通，在沿用既有成熟视频会议系统的同时，提供了更加简单便捷的入会体验，实现移动端与传统会议室设备的融合互通，解决了以往开会只能二选一的问题，有效利用了现有资源，提高开会效率。

越秀地产×钉钉：视频会议



来源：艾瑞咨询研究院自主研究及绘制。

协同办公场景应用案例：金蝶云之家

产品深度整合，利用会议平板实现异地员工的远程协作

云之家是金蝶旗下的移动办公平台，定位新一代智能协同云。金蝶云之家团队经常需要针对产品创意进行头脑风暴，以白板开会不仅经常会出现书写面积不够、会议内容难以保存的问题，也不利于各分公司之间员工的异地协同。为此，金蝶云之家引入MAXHUB高效会议平台，以一体化、轻部署的方式打造流畅、清晰、稳定的视频会议体验，通过文档、应用、屏幕内容等的异地同步，全面提升会议效率。在金蝶云之家园区内，云之家移动办公平台与MAXHUB会议平板已经可以实现互联互通，金蝶云之家与MAXHUB将通过在产品层面的深度整合，共同打造面向未来的办公场景。

金蝶云之家×MAXHUB：基于会议平板的远程协作

应用场景1：展厅

金蝶云之家将MAXHUB作为展厅内的**电子看板**使用

- ✓ 清晰展现各产品线最新产品和技术应用，以及财务、供应链等流程和数据
- ✓ 通过展厅负责人的触控或语音控制，将产品的价值直观地传递给用户



应用场景2：会议室

金蝶云之家全国各个分公司利用MAXHUB会议平板进行**异地协同合作**

- ✓ **一体化、轻部署**：集成高性能摄像、拾音、扩音模块，提供流畅、清晰、稳定的视频会议体验
- ✓ **远程办公、身临其境**：搭配MindLinker等视频会议软件，异地员工在MAXHUB超大屏幕上可以同时清晰地看到视频画面及共享内容，仿佛在同一会议室面对面协作

屏幕共享

书写同步



双方产品深度整合

云之家移动APP与MAXHUB会议平板深度融合，用户可将云之家的文件一键投射到MAXHUB的会议平板上



会议结束后，MAXHUB会议平板上的会议纪要可一键回传到云之家移动办公平台，有效提升会议结果的执行力

平台电商化场景需求分析

融合视频直播技术，全面提升客户体验和个性化需求

在电商直播中，视频云广泛用于美妆、母婴、3C、快消等各个垂直分类直播场景。如何解决高并发、直播延迟卡顿等痛点，保证直播画质、提高观众参与度和积极性是视频云在电商直播中核心关注的问题。基于视频云低时延解决方案，融合直播管理、内容审核等视频处理技术，通过PC端和手机端触达观众，可以更好的保证观众互动体验，提升直播转换效果。除此之外，视频云厂商针对企业直播提供丰富多样的产品功能模块，如共享直播、弹幕打赏、直播抽奖等，满足企业在直播过程中的个性化需求，并通过视频加密、监播等技术，全面保障直播安全，实现高质量直播解决方案落地。

平台电商化场景：视频云需求痛点与解决方案



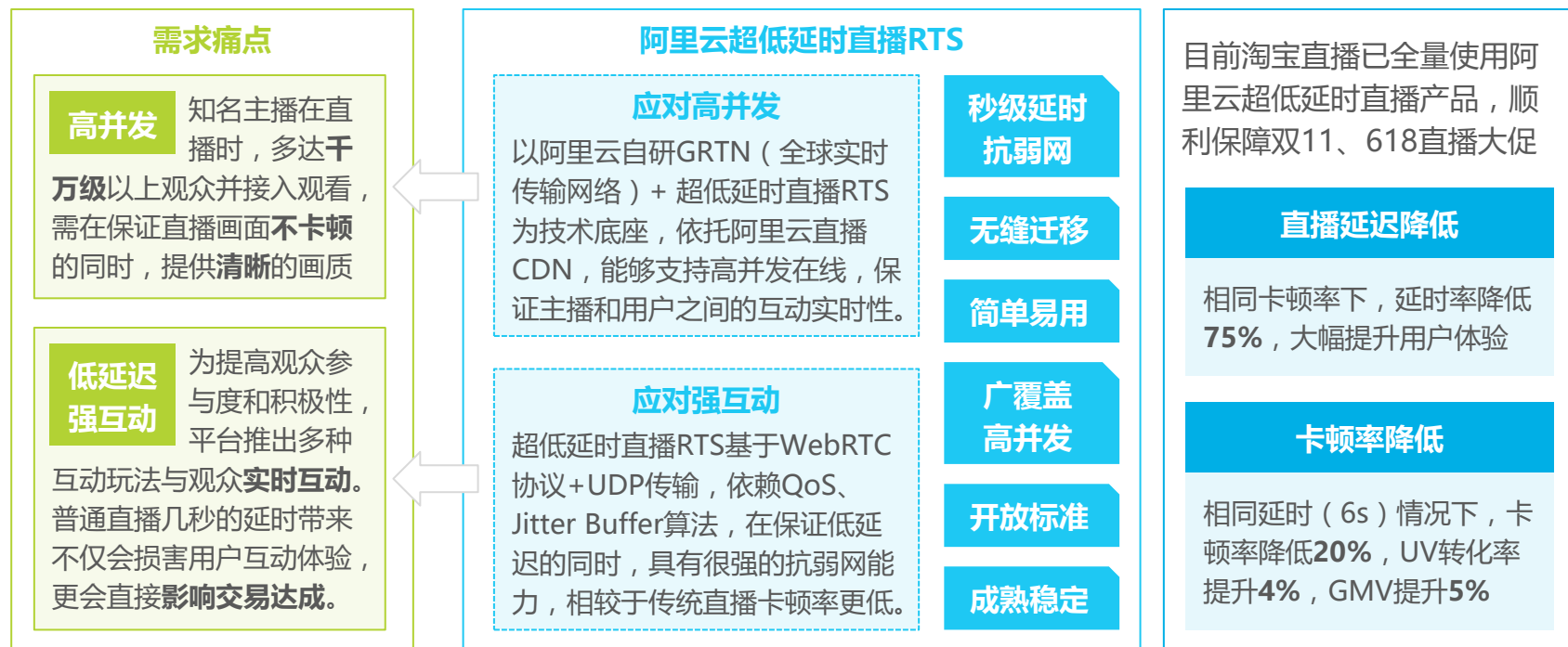
来源：艾瑞咨询研究院自主研究及绘制。

平台电商化场景应用案例：淘宝直播

超低延时电商直播保障618、双11大促的高并发、强互动

2019年直播电商行业爆发，然而传统的直播技术难以满足直播电商对实时互动的高要求。阿里云与淘宝直播共同攻坚并推出了超低延时直播服务RTS（Real-time Streaming），通过WebRTC协议优化、流媒体大脑动态选路及达摩院XGlab Qos专项提升，在承载千万级大规模并发的同时，实现将端到端延时降低到1秒以内。由于RTS服务部署于阿里云直播CDN节点，能够复用阿里云直播CDN的节点和网络资源，从而有效平衡了直播电商对接入成本、节点覆盖与承载能力的需求。阿里云RTS经历了2020年淘宝直播618和双11大促的双重考验，目前已经在淘宝直播全量上线。

淘宝直播×阿里云：超低延时电商直播



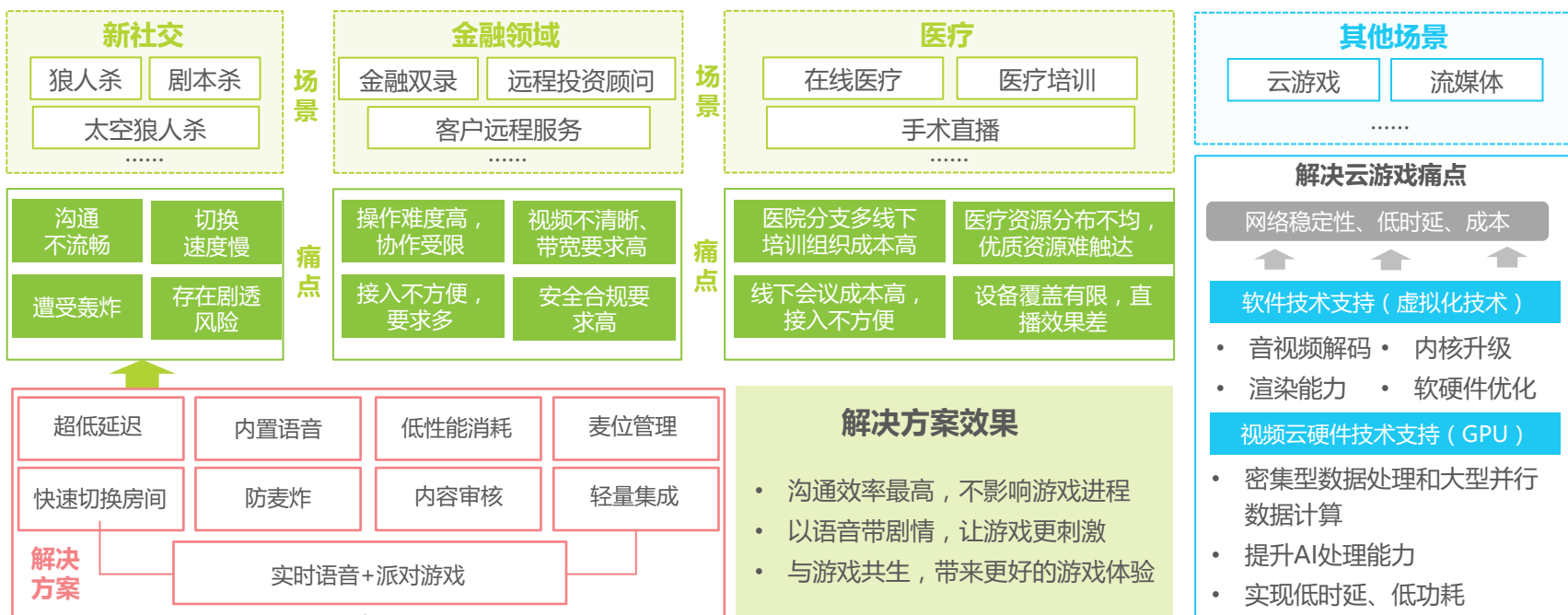
来源：艾瑞咨询研究院自主研究及绘制。

创新应用场景需求分析

剧本杀、云游戏等新兴场景叠出，视频云应用更具想象空间

除了互联网泛娱乐、教育、传播、平台电商化、协同办公等典型场景外，视频云也开始尝试触达新社交、金融业务、医疗业务、云游戏等创新型场景。尤其是社交游戏领域，无论是狼人杀、剧本杀还是太空狼人杀，游戏进程的推进都依赖参与者的线上实时沟通。语音互动中，发言者的语气、声调变化可以成为玩家判断的依据，实时语音不仅充当了主要沟通工具，也是影响玩家体验的重要功能，应用视频云可以帮助客户快速将更多创意派对游戏落地。云游戏的落地面临网络稳定性、低时延和成本等痛点，对云端编码优化和客户端解码有更高的要求。Intel SG1 GPU高密度、低时延、低功耗、低TCO的特性，能够满足《王者荣耀》等安卓云游戏对GPU的硬件需求，为视频云带来更大的市场价值和想象空间。

创新应用场景：需求痛点及解决方案



来源：艾瑞咨询研究院自主研究及绘制。

视频云行业产学研情况

音视频产学研结合不紧密，通过竞赛填补创新实践空白

产业端，推动视频云创新的计算机视觉产业投融资市场活跃度下降，市场趋于理性。但图像识别、人脸识别、物体与场景识别、视频识别等细分领域的投融资热度不减。专业人才供应方面，阳光高考信息平台发布的数据显示，计算机科学与技术为高考最热门专业，开设院校数量高达969个，课程设置包括图像处理与模式识别、三维计算机视觉等音视频专业课程。虽然企业和学校的供需资源丰富，但是音视频产学研专项项目、企业院校联合研究中心、校园创新大赛较少，影响科研成果向市场转化。放眼国外市场，全球音视频赛事种类繁多，在业界享有盛誉的国外音视频类比赛有CVPR、ICCV、ICML等计算机视觉机构协会主办的竞赛，每年全球参赛人数众多、奖金池丰富、难度含金量高，这三个要素使国外的比赛在全球业界占技术主导地位。而国内的音视频比赛主要由视频平台以及云服务厂商主导举办，影响力略有不足。推动产学研结合，需要发挥国内机构协会和头部企业的引领作用，组织开展创新型竞赛，提升大学生实践能力，补足音视频人才缺口。

2016-2020年中国计算机视觉融资案例数量



国内外音视频类竞赛主办方类型



国内外头部的视频平台和音视频服务厂商通过举办竞赛将音视频技术、计算机视觉的产学研相结合，实现科研、教育、生产在功能与资源优势的协同与集成化，将技术创新在上、中、下游对接与耦合。

产学研案例：全球视频云创新挑战赛

聚焦视频云，以顶级赛事驱动技术与行业创新

全球视频云创新挑战赛是由阿里云联手英特尔主办，与优酷战略技术合作，面向企业以及个人开发者的音视频领域的数据算法及创新应用类挑战。大赛依托阿里云天池平台丰富资源，其中算法赛道聚焦计算机视觉领域的热点领域——视频分割，应用创新赛道鼓励参赛选手从行业痛点出发，挖掘视频云技术在各个行业场景的应用。大赛累计吸引了全球23个国家和地区，超过4000名选手参赛，成为颇具影响力的专业领域赛事。

视频云顶级赛事驱动技术与行业创新

算法挑战赛

视频分割将传统图像分割问题延伸到视频领域，可服务于视频理解处理和编辑等任务。该问题的目标是在视频帧中分割目标物体的区域，需精确到像素级别。



视频类数据集对计算机视觉算法至关重要，但受制于视频版权、用户隐私、视频标注等因素影响，**优质的视频数据集稀缺**。

本次全球视频云创新挑战赛使用的数据集依托于**阿里巴巴的数据资源**，形成了目前国内最大规模的视频类数据集。

创新应用挑战赛

创新应用赛道参赛选手应用视频云技术，解决了视频领域和相关行业的痛点问题，实现应用场景的创新，**创造出下一代音视频新场景**。参赛选手让视频云技术切实融入、普惠到我们的生活当中，同时赋能社会各行业的革新，**标志着视频云全场景时代的开启**。

视频+互动娱乐

基于Reaction Video视频网站

“一起看视频”应用里观众可以邀请好友在观影房中一起看娱乐视频、赛事直播或者发布会，互相分享彼此的感受。



深度挖掘视频网站评论和弹幕的背后观众群体所需的认同感和分享欲，应用视频云技术搭建互动式观影房。



交通安全监控

智慧停车场系统

对监控视频进行智能处理，获得停车场实时停车情况以及危险报警系统。充分利用监控视频资源，无需增加额外的设备。



遍布城市的公共摄像头



对摄像头采集视频中的人及车辆进行分割、比对，对行人危险情况预警

孤独人群陪伴

AI钟生伴侣

提供可自定义表情+声音的AI伴侣。



基于多源信息融合交互式的动态仿真系统，提供集观、听、感于一体的智能伴侣，服务面向单身、二次元、LGBT、孤寡等任何需要陪伴的人群。

闹钟+视频陪伴+人工智能

新时代下的视频云

1

云上视频新场景

2

视频云发展趋势展望

3

视频云行业发展趋势：技术极致化

软件优化×硬件适配，不断追求高清化、实时性和互动性

视频处理和传输环节是视频云技术攻坚的重点方向，以更高的压缩比呈现更高质量的视频，保障在高并发、弱网环境下的低时延、去抖动、抗丢包，始终是视频云厂商技术实力对比的关键指标。一方面，视频云厂商将进一步精研软件算法，运用人工智能技术降低视频处理的计算量、提高网络传输效率，另一方面，视频云厂商也将积极同英特尔等硬件厂商合作，利用硬件厂商提供的高性能芯片、AI硬件加速产品、边缘智能产品等实现性能优化和成本下降。

视频云行业发展趋势：技术极致化



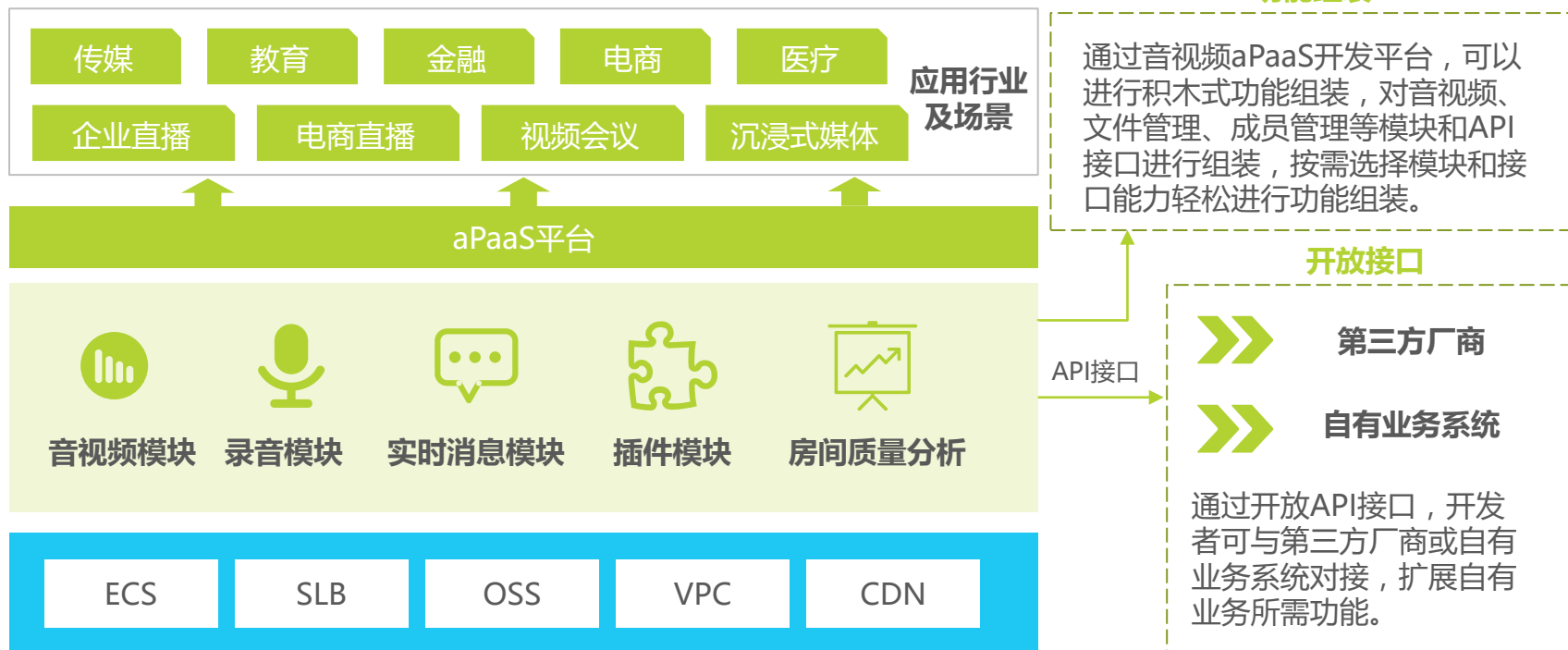
来源：艾瑞咨询研究院自主研究及绘制。

视频云行业发展趋势：低代码开发

通过功能组装和开放接口，灵活满足应用场景音视频需求

相对传统软件开放，低代码开发平台具有开发效率高、可配置性强、易用性强等特点，近年来不少厂商开始将音视频解决方案aPaaS平台化。一方面，在通用能力模块化封装的基础上，aPaaS平台针对不同场景进行了大量业务层面的功能整合，无需开发人员写代码，非技术人员也可按需完成应用开发。另一方面，通过开放各类标准API接口，扩展自有业务所需功能，满足个性化开放需求。未来，低代码开发模式将持续在音视频行业渗透，和视频云其他层级结合向企业提供面向全场景的音视频互动一站式解决方案。

视频云行业发展趋势：低代码开发



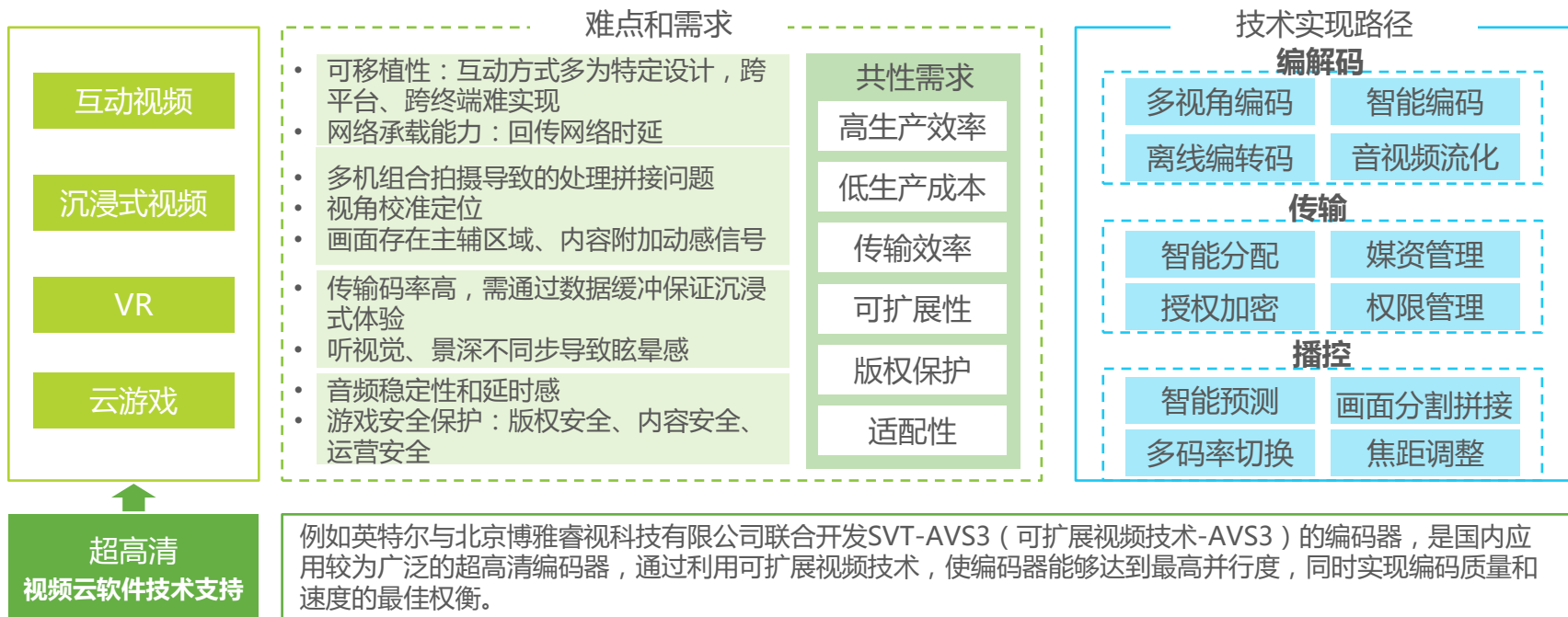
来源：艾瑞咨询研究院自主研究及绘制。

视频云行业发展趋势：交互新体验

超高清和5G高新视频：强交互、重体验、深融合的未来之路

5G时代，用户除了关注视频内容之外，更加看重体验感和应用能力。5G基础设施建设为高新视频发展提供了基础条件，同时，我国自主研发了拥有自主知识产权的AVS3编码标准，以英特尔为代表的8K超高清工程化厂商在编解码技术开发上持续发力。再加上后疫情时代人们的娱乐和消费习惯从线下转向线上，高新视频的发展可谓占据了“天时、地利、人和”。互动视频、沉浸式视频、VR、云游戏作为5G基础设施和超高清视频共同叠加产生的高新视频新业态，让视频赋能千行百业的能力得到充分的发挥。超高清、沉浸式、互动性视频带来的海量数据流对算法、算力提出更高的要求，这为视频云的技术迭代和场景应用带来新的机遇和挑战。

视频云行业发展趋势：交互新体验



来源：艾瑞咨询研究院自主研究及绘制。

艾瑞新经济产业研究解决方案



行业咨询

- 市 场 进 入 为企业提供市场进入机会扫描，可行性分析及路径规划
- 竞 争 策 略 为企业提供竞争策略制定，帮助企业构建长期竞争壁垒



投资研究

- IPO行业顾问 为企业提供上市招股书编撰及相关工作流程中的行业顾问服务
- 募 投 为企业提供融资、上市中的募投报告撰写及咨询服务
- 商业尽职调查 为投资机构提供拟投标的所在行业的基本面研究、标的项目的机会收益风险等方面的深度调查
- 投后战略咨询 为投资机构提供投后项目的跟踪评估，包括盈利能力、风险情况、行业竞对表现、未来战略等方向。协助投资机构为投后项目公司的长期经营增长提供咨询服务

关于艾瑞




艾瑞咨询是中国新经济与产业数字化洞察研究咨询服务领域的领导品牌，为客户提供专业的行业分析、数据洞察、市场研究、战略咨询及数字化解决方案，助力客户提升认知水平、盈利能力和综合竞争力。

自2002年成立至今，累计发布超过3000份行业研究报告，在互联网、新经济领域的研究覆盖能力处于行业领先水平。

如今，艾瑞咨询一直致力于通过科技与数据手段，并结合外部数据、客户反馈数据、内部运营数据等全域数据的收集与分析，提升客户的商业决策效率。并通过系统的数字产业、产业数据化研究及全面的供应商选择，帮助客户制定数字化战略以及落地数字化解决方案，提升客户运营效率。

未来，艾瑞咨询将持续深耕商业决策服务领域，致力于成为解决商业决策问题的顶级服务机构。

联系我们 Contact Us

 400 - 026 - 2099

 ask@iresearch.com.cn



企 业 微 信



微 信 公 众 号

法律声明

版权声明

本报告为艾瑞咨询制作，其版权归属艾瑞咨询，没有经过艾瑞咨询的书面许可，任何组织和个人不得以任何形式复制、传播或输出中华人民共和国境外。任何未经授权使用本报告的相关商业行为都将违反《中华人民共和国著作权法》和其他法律法规以及有关国际公约的规定。

免责条款

本报告中行业数据及相关市场预测主要为公司研究员采用桌面研究、行业访谈、市场调查及其他研究方法，部分文字和数据采集于公开信息，并且结合艾瑞监测产品数据，通过艾瑞统计预测模型估算获得；企业数据主要为访谈获得，艾瑞咨询对该等信息的准确性、完整性或可靠性作尽最大努力的追求，但不作任何保证。在任何情况下，本报告中的信息或所表述的观点均不构成任何建议。

本报告中发布的调研数据采用样本调研方法，其数据结果受到样本的影响。由于调研方法及样本的限制，调查资料收集范围的限制，该数据仅代表调研时间和人群的基本状况，仅服务于当前的调研目的，为市场和客户提供基本参考。受研究方法和数据获取资源的限制，本报告只提供给用户作为市场参考资料，本公司对该报告的数据和观点不承担法律责任。

为商业决策赋能

EMPOWER BUSINESS DECISIONS

