

## 视频体验评估的需求和场景

Application scenarios and demand of assessment for QoE of video services

2017年9月29日发布

2017年10月1日实施

---

中关村现代信息消费应用产业技术联盟 发布

## 目 次

目次.....	I
前言.....	II
视频体验评估的需求和场景.....	1
1 范围.....	1
2 规范性引用文件.....	1
3 缩略语.....	2
4 视频业务用户体验评估技术.....	4
5 视频业务类型划分方法.....	5
6 各类视频服务应用场景与用户体验评估需求.....	8
6.1 分发类视频服务.....	8
6.1.1 直播式分发类视频服务.....	8
6.1.2 点播式分发类视频服务.....	11
6.2 会议类视频服务.....	12
6.2.1 应用场景.....	12
6.2.2 视频体验评估需求.....	14
6.3 会话类视频服务.....	14
6.3.1 应用场景.....	14
6.3.2 视频体验评估需求.....	15
6.4 采集类视频服务.....	15
6.4.1 应用场景.....	15
6.4.2 视频体验评估需求.....	15
6.5 检索类视频服务.....	15
6.5.1 应用场景.....	16
6.5.2 视频体验评估需求.....	16
6.6 信息类视频服务.....	16
6.6.1 应用场景.....	16
6.6.2 视频体验评估需求.....	17
6.7 其他视频业务.....	17
6.7.1 远程教育.....	17
6.7.2 远程医疗.....	17
7 用户体验评估技术的扩展应用场景.....	18
7.1 用户自助评测业务质量.....	18
7.2 运营商采集用户体验评估协助视频质量监测.....	18

## 前 言

本标准是针对视频服务用户体验评估系列标准之一，本标准预计的架构如下：

- 视频体验评估的需求和场景
- 视频服务用户体验评估的系统架构
- 视频服务用户评估算法和参数-直播
- 视频服务用户评估算法和参数-点播
- 视音频业务主观评价技术专家组章程

本标准按照GB/T 1.1-2009给出的规则起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本标准由中关村现代信息消费应用产业技术联盟提出并归口。

本标准起草单位：中国联通集团公司

本标准主要起草人：中国信息通信技术研究院、国家新闻出版广电总局广播电视规划院、中国电信集团公司、中国移动集团公司、华为技术有限公司、腾讯公司、爱奇艺公司、优酷土豆股份有限公司、网宿科技股份有限公司、上海交大、博汇科技有限公司

## 视频体验评估的需求和场景

### 1 范围

本标准规定了视频服务用户体验评估的需求和场景。

本标准适用于公用电信网、广播电视网、公用互联网、企业专用网等环境下的视频服务用户体验评估需求和场景的分析和分类。

### 2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅所注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

RFC 4445 A Proposed Media Delivery Index (MDI)

RFC 3550 Real-time Transport Protocol

ETSI TR101 290 V1.2.1, Digital Video Broadcasting (DVB); Measurement Guidelines for DVB Systems

ITU-R BT.500-11, Methodology for the subjective assessment of the quality of television pictures

ITU-T F.700 (2000) Framework Recommendation for multimedia services

ITU-T F.701 (2000) Guideline Recommendation for identifying multimedia service requirements

ITU-T F.702 (1996) Multimedia conference services

ITU-T F.703 (2000) Multimedia conversational services

ITU-T F.741 (2005) Service description and requirements for audiovisual on-demand services

ITU-T F.742 (2005) Service description and requirements for distance learning services

ITU-T G.1011 (2015) Reference guide to quality of experience assessment methodologies

ITU-T G.1080 (2009) Quality of experience requirements for IPTV services

ITU-T P.10/G.100, Amd 1 (2007) Definition of Quality of Experience (QoE)

ITU-T P.800.1 (2006) Mean Opinion Score (MOS) terminology

ITU-T P.910 (2008) Subjective video quality assessment methods for multimedia applications

ITU-T P.1201 (2012) Parametric non-intrusive assessment of audiovisual media streaming quality

ITU-T P.1201.1 (2012) Parametric non-intrusive assessment of audiovisual media streaming quality - lower resolution application area

ITU-T P.1201.2 (2012) Parametric non-intrusive assessment of audiovisual media streaming quality - higher resolution application area

ITU-T P.1202.1 (2012) Parametric non-intrusive bitstream assessment of video media streaming quality - lower resolution application area

ITU-T P.1202.2 (2013) Parametric non-intrusive bitstream assessment of video media streaming quality - Higher resolution application area

ITU-T P.1401 (2012) Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models

ITU-T J.144 (2004) Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference

ITU-T J.341 (2011) Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference

ITU-T J.343 series (2014) Hybrid perceptual/bitstream models for objective video quality measurements

### 3 缩略语

下列缩略语适用于本标准。

1080P	1920x1080 Progressive Scanning	一种在逐行扫描下达到 1920×1080 的分辨率的显示格式
4CIF		一种 CIF 图像格式，它的主要参数规定如下：亮度取样的象素个数(dx) ×亮度取样的行数 (dy)、色度取样的象素个数(dx/2) 和色度取样的行数(dy/2)分别为 704×576、352、288
720P	1280x720 Progressive Scanning	一种在逐行扫描下达到 1280×720 的分辨率的显示格式
ABR	Available Bit Rate	自适应码率
AP	(Wireless)Access Point	(无线)访问接入点
APP	Application	手机软件
AVoD	Audiovisual on-Demand	视听点播
BBS	Bulletin Board System	电子公告牌系统
BRAS	Broadband Remote Access Server	宽带远程接入服务器

BSC	Base Station Controller	基站控制器
CDN	Content Delivery Network	内容分发网络
CIF	Common Intermediate Format	标准化图像格式
CM	Cable Modem	线缆调制解调器
CMTS	Cable Modem Terminal Systems	线缆调制解调器终端系统
CR	Core Router	核心路由器
DHCP	Dynamic Host Configuration Protocol	动态主机配置协议
DL	Distance Learning	远程教育
DLSP	Distance Learning Services Platform	远程教育服务平台
DNS	Domain Name System	域名系统
DRM	Digital Right Management	数字版权管理
DSLAM	Digital Subscriber Line Access Multiplexer	数字用户线路接入复用器
DTS	Decode Time Stamp	解码时间标签
DTV	Digital Television	数字电视
EPG	Electronic Program Guide	电子节目单
FTP	File Transfer Protocol	文件传输协议
FTTH	Fiber To The Home	光纤到户
FW	FireWall	防火墙
GGSN	Gateway GPRS Support Node	网关 GPRS 支持节点
HAS	HTTP Adaptive Streamin	HTTP 自适应流媒体
HG	Home Gateway	家庭网关
HLS	HTTP Live Streaming	Apple 的动态码率自适应技术
HPD	HTTP Progressive Download	HTTP 渐进下载
HTTP	HyperText Transfer Protocol	超文本传输协议
IGMP	Internet Group Management Protocol	互联网组管理协议
IPTV	InternetProtocolTelevision	交互式网络电视
KPI	Key Performance Indicator	关键功能、性能指标
KQI	Key Quality Indicators	关键质量指标
LAN	Local Area Network	局域网
MCU	Multipoint Control Unit	多点控制单元
MDI	Media Delivery Index	媒体传输指标
MDU	Multiple Dwelling Unit	多用户居住单元
MOS	Mean Of Score	主观感知评分
MSE	Middleware Service Element	中间件服务元
MTU	Multi-Tenant Unit	多商户单元
OLT	Optical Line Terminal	光线路终端
OTT	Over The Top	基于开放互联网的服务
PAD	Portable Android Device	平板电脑
PC	Personal Computer	个人电脑
PDA	Personal Digital Assistance	掌上电脑
PE	Provider Edge	边缘路由器
PGW	PDN(Packet Data Network) Gateway	PDN(公用数据网)网关
PIM	Protocol Independent Multicast	协议无关组播

PSTN	Public Switched Telephone Network	公用电话交换网
PTS	Presentation Time Stamp	显示时间标签
QAM	Quadrature Amplitude Modulation	正交调幅
QoE	Quality of Experience	用户体验质量
QoS	Quality of Service	服务质量
RNC	Radio Network Controller	无线网络控制器
RTSP	Real Time Streaming Protocol	实时流传输协议
SGSN	Service GPRS Support Node	服务 GPRS 支持节点
SGW	Signaling GateWay	信令网关
SMTP	Simple Mail Transfer Protocol	简单邮件传输协议
SR	Service Router	全业务路由器
STB	Set Top Box	机顶盒
STU	Set Top Unit	机顶盒
TCP	Transfer Control Protocol	传输控制协议
TS	Transport Stream	传送流
UC	User Center	用户中心
UDP	User Datagram Protocol	用户数据报协议
VCR	Video Cassette Recorder	盒式磁带录像机
VoD	Video on Demand	视频点播
VR	Virtual Reality	虚拟现实
Web Cache		网页快照
Wi-Fi	Wireless Fidelity	一种允许电子设备连接到一个无线局域网（WLAN）的技术
YUV		一种颜色编码方法（属于 PAL），是 PAL 和 SECAM 模拟彩色电视制式采用的颜色空间

#### 4 视频业务用户体验评估技术

和语音、数据业务相比，视频业务占用带宽资源较多、实时性要求较高，并且对数据丢失、时延、抖动等网络特性非常敏感。影响用户视频体验的关键因素如下图所示，主要包括视频源质量、音频源质量、用户的交互体验和观看体验。

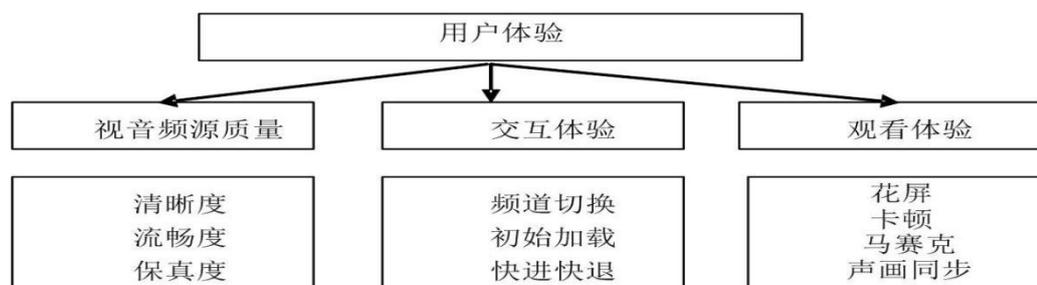


图 1 影响用户体验的因素

视频源质量取决于视频源的清晰度、流畅度、保真度（色调、对比度）等因素，涵盖了视频源的码率、帧率、分辨率、编码方式、编码参数、信号层参数和编码格式等客观指标。

音频源质量取决于音频源的清晰度、流畅度、保真度等因素，涵盖了音频源的码率、采样率、声道数、编码方式、编码参数、信号层参数和编码格式等客观指标。

用户的交互体验取决于系统对用户交互操作的响应速度，涵盖了平台、网络、终端性能指标，可以通过用户一次观看行为中各类请求（直播/点播/EPG/索引请求）的总次数和成功次数、各类请求（直播/点播/EPG）的时延、快进、快退、暂停响应时延，以及异常中断的次数等衡量。

观看体验取决于观看视频过程中出现的画质劣化情况，包括花屏、马赛克、卡顿、声画不同步等因素，可以通过视音频信息的传输性能和损伤等客观指标衡量。目前主要包括缓存卡顿（次数/时长/发生时间）和画面花屏（次数/总时长/平均花屏面积百分比）等参数。

客观指标对用户视频体验质量的影响权重随着时间的持续而弱化，当花屏/卡顿事件出现时，视频的质量体验立即下降，当花屏/卡顿结束，恢复正常播放时，用户对于视频的质量体验就会逐渐地缓慢恢复。假如后续一直能持续正常播放，实时的质量体验会逐渐恢复到正常值。

用户的视频体验质量得分，是综合视频源质量、交互体验质量、观看体验质量，尽量真实反映用户体验感知的平均主观评分（MOS）。

如果能够实时准确地对网络视频服务质量进行监控，获得反映用户感受的视频体验质量，就可以基于实际的用户体验来调节编解码器或传输网络的参数，有效改善视频业务的服务质量。

ITU-T P.10/G.100 Amd 1定义用户体验（QoE）为终端用户对应用或者服务整体的主观可接受程度，它从用户角度来体验一个系统的整体性能。它由终端用户的主观感知决定。体验质量包含完整的端到端系统的影响（客户、终端、网络、服务基础设施等等）。整体可接受性可能受用户期望或所处环境的影响。QoE的影响因素主要可分为三个层面：服务层面、环境层面及用户层面。

基于视频业务运营的诉求，需要建立一套以QoE为中心的评价体系，采用统一的衡量标准，来定量地评价不同网络、不同屏幕、不同场景应用下视频体验的好坏，同时该视频体验标准体系需要具备可实施性、可扩展性并支持演进。用户的视频体验应采用客观评估算法评价，客观评估算法根据其输入信息的不同可划分为4个层级，如下图所示。

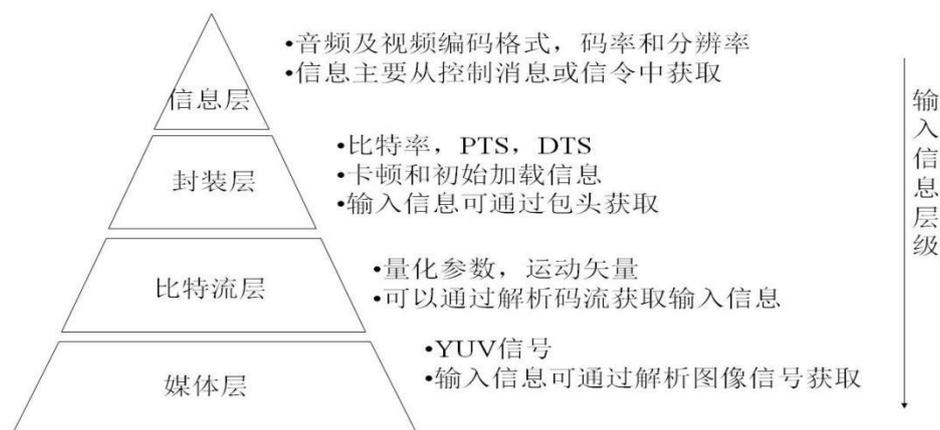
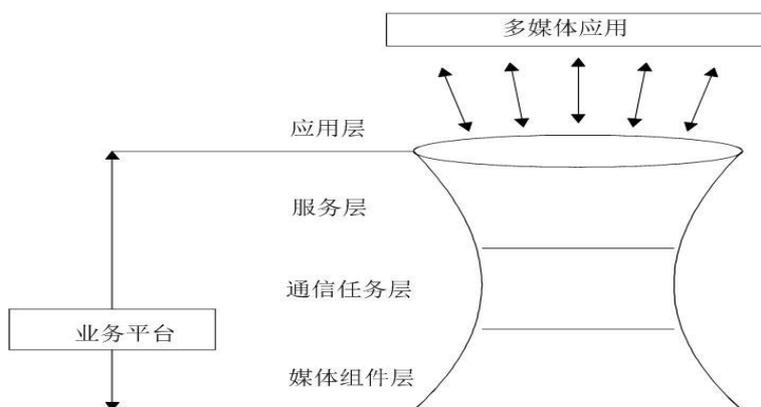


图2 视频评估信息模型

视频传输流程包括采集、传输和接收，视频质量评估也可以在这三个层面进行视频质量分析。

## 5 视频业务类型划分方法

以视频为主要内容的多媒体业务可以用模块化的四个层次的参考模型来表示。如下图所示，最顶层的为应用层，向下依次为服务层、通信任务层、媒体组件层。



应用层从终端用户的角度描述了各种应用的必要功能特性，不关联底层技术或特定的网络解决方案。多媒体应用需要同时使用两个（含）以上的表示媒体（对应不同信息类型），它们将构成一个共同的信息空间。应用实例有合作编辑文档、远程会议、远程监控、医疗文件远程分析和远程培训等。

服务层定义了用来满足应用层功能要求的通信服务或支持工具的基本集合，这里定义和描述的整体服务原则包括：性能、服务质量、安全、计费、互通。通过联合通信任务可建立服务。服务水平可以包含如何定位最终用户和终端位置。从用户的角度来看，多媒体服务是同步处理两个（含）以上类型的电信服务。多媒体服务可以涉及多方、多个连接，以及在单个通信会话内的增减资源和用户。

通信任务层定义和描述了构建的服务所需的特定的通信任务。通信任务是执行多媒体服务的通信功能的功能实体。它以同步方式运用一组媒体组件，以传达和控制各种复杂的信息类型。

在媒体组件层，将各类服务的各个单一媒体组件关联到单一的信息类型（例如音频、视频等），然后进行鉴定和描述，这样多媒体在各方面的特征将变得明显。

上述4个层中，服务层、通信任务层、媒体组件层可合在一起称为通信平面。为了运营这些服务需要的控制和处理功能被称为控制和处理平面，它由被称为中间件服务元（MSE）的模块化实体构成。

从用户请求开始，由上向下分析，多媒体服务可以首先被分解成一组通信任务，其中每一个任务由用户和/或服务提供商单独或配合完成。相反，在自下而上的分析中，通信任务可以看作是将为了同一个服务而彼此相关的各媒体组件进行汇集的一种手段。

在对通信任务的划分中，可以参考以下特性。

表1 通信任务特性

属性	可能值
通信配置	点到点、点到多点、多点到点、多点到多点
信息流对称性	单向、双向对称、双向非对称
传送控制实体	源、宿、源和宿、第三方
通信延迟	实时、近实时、非实时、特定时间
必选媒体组件	音频/视频/文本/静态图片/图形/数据
可选媒体组件	音频/视频/文本/静态图片/图形/数据/无

属性	可能值
媒体组件相互关系	1) 相互同步： <ul style="list-style-type: none"> <li>● 音频和视频（唇同步，位置相关同步）；</li> <li>● 音频和文本（语音合成）；</li> <li>● 文本和视频/静态图片/图形（图像带有同步字幕）；</li> <li>● 图形和音频。</li> </ul> 2) 相同类型的媒体组件之间的对称性，是否允许双向。           3) 信息类型（或媒体组件）之间的转换。
时间连续性	同步/非同步

通信延迟有四个可能值：

——实时，即在1秒钟以内，这适用于会话或一些视频游戏中，这时游戏用户的操作必须立即显示在屏幕上；

——近实时，即在几秒钟内，这适用于用户等待系统应答的检索服务，这时可接受的延迟取决于请求的复杂程度和应答的长度；

——非实时，这适用于存储信息，例如用于存储或获取消息，这时的行动可以允许在任何合理的但不严重的延迟内完成；

——特定时间，即在将来精确选定的时间点或未来的一个时间周期，这可以适用于某些类型的点播视频服务，用户可以提前要求在未来的某个特定的时间将选中的电影视频发送给他，或者是一个检索服务，由于检索请求较复杂因此需要采取一些时间研究和处理数据，该系统将在给定时间后回拨；它也可以适用于在夜间网络流量少期间发送信息。类似地，需要在接收端人工介入的非紧急消息则可以在办公时间的任何时候被发送。

时间连续性具有的属性值为“同步”或“非同步”。如果该用户终端设备无缓冲能力则必须进行同步传输，或者如果数据捕捉设备不能容忍传输速度的中断或变动，或者它不具备足够大的存储容量。

通过结合不同的属性，可以从不同角度对通信任务进行分类。例如，将“通信配置”、“信息流对称性”和“传送控制实体”三个参数相结合的划分方法如下：

表2 通信任务分类示例

通信配置	三类业务流的通信任务		
	双向	单向-源端控制	单向-宿端控制
点到点	会话	发送	接收
点到多点		分发	
多点到点			采集
多点到多点	会议		

而“通信延迟”和“时间连续性”两个时间特性参数相结合的划分方法如下表。在下文中，对分发类视频服务的进一步分类中将会用到该分类方法。

表3 “通信延迟”和“时间连续性”结合分类

通信延迟	服务示例	
	同步	非同步

实时	会话	无
近实时	在线音视频检索 (注)	在线数据检索 (注)
非实时	近视频点播	数据信息
特定时间	音视频信息 (注)	音视频信息 (注)
注：同步特性取决于终端设备的存储和捕捉能力。		

服务描述有两个阶段：一种是独立于网络的一般说明，另一种是对于承载服务的网络的一般说明。

相关服务的定义与相关实现技术的建议可以通过两种方式进行关联：第一种是利用在诸如协议和编码算法等技术的建议中所定义的技术构建模块来构建通信能力，即通信任务和媒体组件；第二种方式则相反，它先定义了服务对于设备、属性、功能和性能的要求，相应的实现技术建议必须能够反映和满足上述要求。一方面为用户期望的服务质量和功能，另一方面是技术的限制和实施成本，需要统筹权衡应对这两方面要求，以得到最佳综合效果。

以表2的“通信任务分类”为基础，并结合用户角度，可将常见视频服务分为六大类：分发、会议、会话、采集、检索、信息。本标准将分别阐述。

## 6 各类视频服务应用场景与用户体验评估需求

本章节将简要陈述视频服务的主要应用场景和各场景对于用户体验评估的需求。

在用户体验评估技术成熟后，与网络质量相关的部分指标（如时延）可进一步分解为对链路上各段网络的要求，对于网络的规划、建设、运维将具有指导意义。

目前以4K高清视频、VR等为代表的视频新技术发展较快，需要研究明确相关的用户体验评估技术。与此同时，随着视频业务对于业务质量要求的提升、应用的推广发展及对网络要求的提高，也将对视频业务用户体验提出新的要求。

### 6.1 分发类视频服务

参考前文表3中“通信延迟和时间连续性结合分类”方法，可以将分发类视频服务分为直播式和点播式两大类。前者通信延迟要求为“实时”，时间连续性通常为“同步”，用户不能控制观看内容的临时次序。后者通信延迟要求为“特定时间”，时间连续性通常为“非同步”。

#### 6.1.1 直播式分发类视频服务

##### 6.1.1.1 应用场景

直播视频是娱乐性视频的一种重要业务，可以通过网络信号，在线收看体育赛事、重大活动和新闻等，此时用户不能控制观看内容的临时次序。

直播按技术实现手段可分为两大类：广播电视直播和IP网络直播。广播电视直播示意图如下：

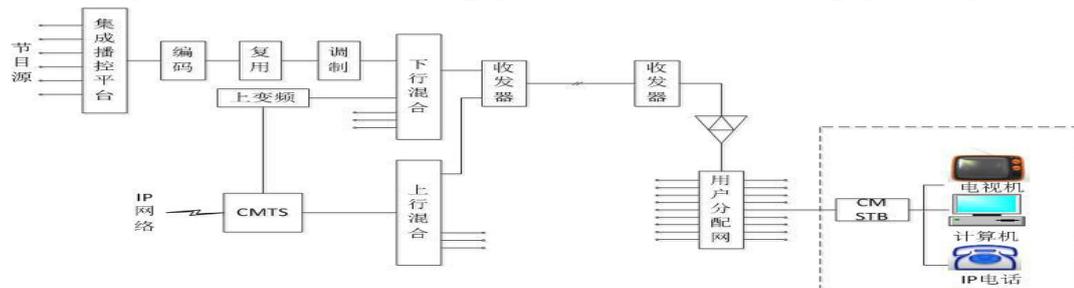


图4 广电视直播架构示意图

在广电的有线数字电视网络中，数字前端将多个编码后的电视频道节目信号复用到 TS 码流中。复用了节目的 TS 码流被调制（QAM 方式）到同轴电缆的频点上，通过信号分配网络送到机顶盒。上图 4 就是一个典型的双向数字电视网络示意图。双向有线数字电视网络由多个环节构成，在下行方向，集成播控平台负责视频节目的收录、播出和信号调度，待传送的多路节目信号经编码、复用、调制后与 CMTS 送来的 IP 调制信号混合，经收发器传送到用户分配网络。在用户家庭中，CM 完成 IP 信号的解调后送到计算机或 IP 电话机，STB 完成有线电视 QAM 信号的解调后送到电视机。在上行方向，IP 信号经 CM 调制后经传送网回传到 CMTS，解调后可接入 IP 网络。上行通道也是有线电视网络的回传通道。

IP 网络直播首先可以按配置的用户数量来分为组播和单播两大类。

在组播模式下，网络通过组播路由协议（如 PIM 等）建立好组播树。头端系统将流媒体编码封装后，利用该组播树将视频流数据分发至网络边缘的组播复制点。用户设备（如机顶盒）在请求收看某一频道时，发送 IGMP 加入消息到组播复制点。用户加入组播后，组播复制点立刻复制流媒体数据到用户设备中。

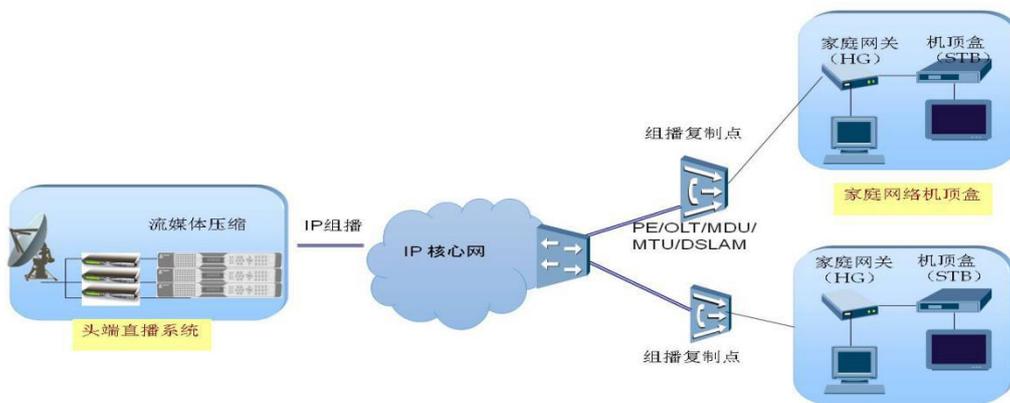


图5 直播视频组播架构示意图

直播视频一般通过频道的形式呈现给用户。在视频传输时，根据不同网络结构和用户规模，可以采用不同的接入封装方式以及承载方式。大多数运营商自建的IPTV直播都采用组播方案和UDP承载传输，为实时流媒体，没有业务层交互控制，报文在传输过程中出现差错或丢包时不会重传。在媒体服务器到网络边缘节点之间，一般采用组播方式传送如上图5，可以节省城域网的带宽资源；而网络边缘节点到用户之间可能采用单播或者组播模式。组播模式下，基本采用UDP来推送内容码流。而单播模式下，可能使用TCP或UDP进行传输。本文的直播视频场景主要考虑最常见的组播UDP这种直播场景。TCP传输直播视频的场景与点播场景类似。

如图6，目前国内的IPTV直播频道有一部分是通过单播向用户提供服务的，在这类场景下可能使用TCP或UDP进行传输。在单播之前，内容提供商通过组播把直播流推送到运营商中心节点，从中心节点到区域或边缘节点可能是组播，也可能是单播中继（Relay）的。

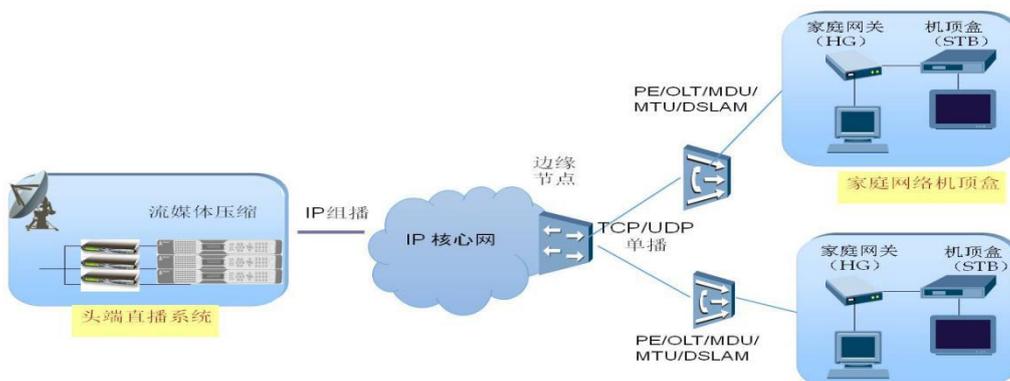


图6 直播视频单播架构示意图

单播模式下，头端系统先通过组播把直播流推送到网络边缘节点上。用户设备（如机顶盒）通过 TCP 或 UDP 单播向边缘节点请求内容。

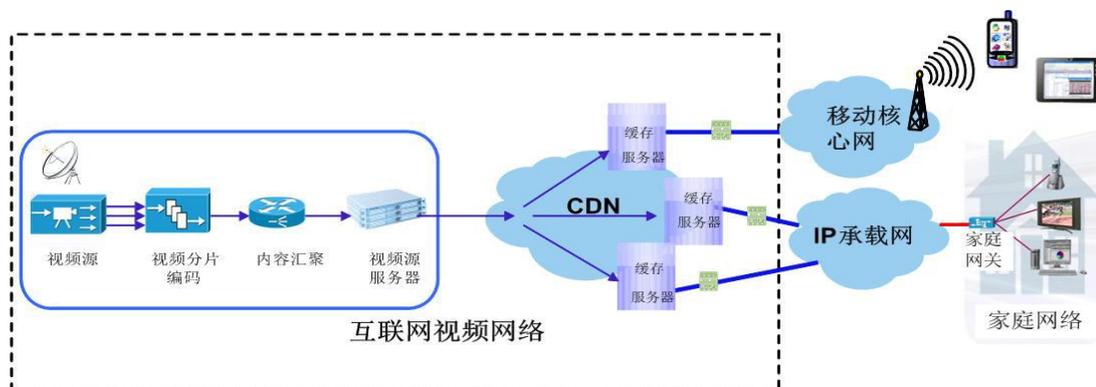


图7 互联网视频直播架构示意图

现在流行的互联网直播属于 IP 网络直播的一种应用，其系统结构示意图如图 7，它通过公共互联网提供面向电视传输的 IP 视频和互联网应用相融合的服务。其接收终端为互联网电视一体机或机顶盒+电视机及移动终端。

互联网视频主要有两种实现方式：HTTP 渐进下载（HTTP Progressive Download，简称 HPD）和 HTTP 自适应流媒体（HTTP Adaptive Streaming，简称 HAS）。传统的 OTT 视频一般采用 HPD 技术。基于 HPD 的客户端在开始播放之前仅需等待一段较短的时间用于下载和缓冲媒体文件最前面的一部分数据，之后便可以一边下载一边播放。HPD OTT 视频存在诸多的局限性，例如：不适合对实时性要求较高的直播节目的传输；初始播放的等待时延一般较长；当网络带宽不稳定时比较容易出现卡顿现象；由于客户端会持续下载视频文件，当用户中途放弃节目观看，会造成已下载文件（消耗带宽）的浪费。

为了克服 HPD OTT 视频技术的局限性，HAS OTT 采用视频分片和自适应码率（ABR）技术，可向不同屏幕大小的终端提供适合分辨率的视频分片文件，并在不同网络带宽情况下实现视频的流畅播放。

移动视频服务也是 IP 网络直播一种应用场景，其系统结构如下图所示，主要涉及网络包括互联网骨干网、IP 承载网、IP 城域网、无线网络、有线接入网等。它与固网的业务流程差异主要体现在承载网内。如图有 4 种视频业务流程：一是由本地 CDN 提供视频内容；二是通过 DNS 解析，由本地 Web Cache（网页快照）提供视频内容；三是由外省 CDN 或 Web Cache 提供视频内容；四是由外网 CDN 或源站提供视频内容。

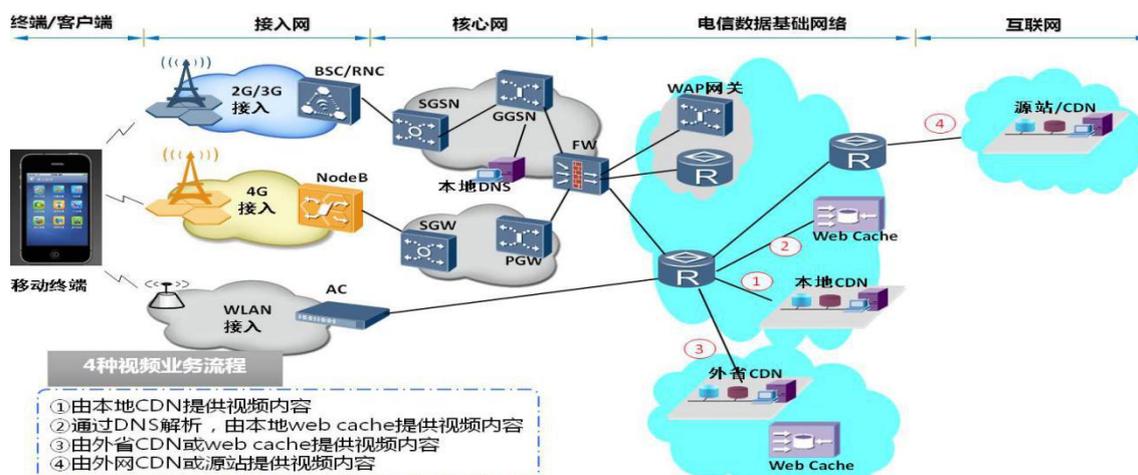


图8 移动视频直播架构示意图

### 6.1.1.2 视频体验评估需求

目前的直播视频传输方式下，整条传输链上可能同时包含TCP和UDP传输环节，即同时存在重传和丢包带来的质量劣化（如卡顿和花屏），该情况对其他视频业务如回看等也是同样存在的。在这种混合网络视频传输中，需要同时考虑卡顿和花屏的影响。而且为了反映前面几个环节的问题，视频体验评估需要支持云管端部署，以实现故障定界定位。

用户对直播视频最直观的体验是画面质量和频道切换响应：直播由于是实时流媒体，故当出现丢包时，如被丢弃的报文包含重要帧如I帧，则会出现短暂黑屏现象，而当被丢弃的报文仅有B帧或P帧，则会出现面积和范围不等的花屏（马赛克）现象，这是影响直播体验的关键因素；另一方面，由于直播视频一般以频道的方式呈现，基于广播电视换台操作经验，用户对换台时延的期望值较高，因此在对直播视频的体验评估进行研究时需要考虑以上因素。

用户在使用直播视频业务时，其心理预期会对用户体验产生影响，例如在不同的网络接入条件下（公共场合无线连接、家中Wi-Fi等），用户对视频业务的预期质量是有差别的，因此需要将用户心理预期纳入评估范围。

视频用户对视频体验优劣进行评价时将跨越视频业务的全流程，并站在最终消费者体验的角度去横向比较不同的分辨率等各类参数带来的不同体验，经过研究可指导运营商网络特别是直播视频网络的设计和优化。

移动视频相关技术正在快速发展中。根据融合CDN的技术发展思路，移动视频与固网视频在业务平台和CDN的层次上可以通用多项质量监测设备和技术。对于影响到用户体验的移动视频业务质量，需要重点研究无线网络（3G/4G网或家庭内部的Wi-Fi）质量监测、移动终端（手机、PAD）特性、跨省分漫游等新特性带来的新挑战。为了更好地评估移动视频的用户体验，可进行以下研究：

一是需要加强移动视频节目源检测标准研究，明确节目源的技术参数要求。移动视频节目源质量监测技术与IPTV相似，但是由于移动终端（手机、PAD）屏幕等小于电视屏幕，因此节目的帧率、分辨率通常略低于IPTV，需要通过研究明确技术要求。

二是需要推进移动视频的用户体验评估算法走向成熟，争取达到在一定网络条件（如带宽）和终端软硬件条件下能够基本确定提供给客户的业务质量。以提高和改善用户的观看感知和体验为目标，研究基于用户感知的移动视频节目指标体系，如综合权衡流畅度和清晰度要求后来配置各项参数。

三是需要加强移动通信链路末端的无线网络段（3G/4G网或家庭内部的Wi-Fi）的网络质量监测研究。由于移动视频与通常的上网相比，对于网络流量要求更高，需要研究确认移动视频业务对于无线网络的要求。Wi-Fi环境下，开放空间AP频段竞争调度对于用户上网的影响需要研究解决。

四是需要加强对于手机、PAD等移动终端研究，研究明确在终端上采取业务质量和用户体验的监测手段的技术可行性和相关的软硬件要求。监测移动视频质量时即要保证对于视频质量的用户体验，又要确保监测信号不能占用过多的带宽、内存、CPU，以免影响用户观看体验和互动体验。

### 6.1.2 点播式分发类视频服务

对于普通家庭用户的娱乐类视频服务，通常称为视听点播服务。它是由内容提供商和服务提供商向多个普通用户提供视频节目观看服务，用户可在提供的节目中选择观看。根据本标准表2中对于通信任务的分类方法，它的“通信配置”是“点到多点”结构而“传送控制实体”是“单向-源端控制”。它是用户和机器（视听资源）之间不对称的互动多媒体通信，通过电信网络在任何时间为用户提供视听信息。用户可以在导航子系统协助下查找所需信息。用户选择确定服务模式后，视听点播服务平台发送具有一些QoS保证等级的内容信息。用户可使用类似VCR功能，以查看接收到的信息。

由视听点播业务平台提供的内容可以是文本、音频、图形、图片或视频信息。媒体资源可被存储在网络的任何位置上。视听点播业务平台整合媒体资源，进行数字版权管理（DRM），并通过电信网络提供导航服务和内容交付。

用户可以在固定或移动的位置如家庭、办公室，或在移动的火车上来使用视听点播服务。用户应该能够使用PC、机顶盒+电视机、手机、PDA或其他设备，通过流媒体通信能力接入宽带网络获得服务。

### 6.1.2.1 应用场景

点播视频允许用户根据自己的需求选择相应的视频进行收看，而不是被动地收看某个特定的广播视频。视频点播的实现过程如下：当用户发出点播请求时，流媒体服务系统就会根据点播信息，将存放在片源库中的节目信息检索出来，以视频和音频流文件，通过高速传输网络传送到用户终端。点播业务一般采用单播网络来实现。视频点播架构如下图9。

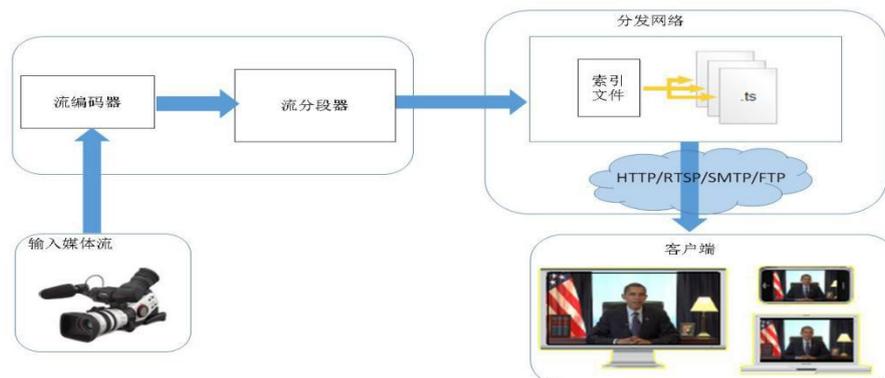


图9 视频点播架构示意图

网络结构由VoD前端处理系统、服务器、分发网络、客户端（机顶盒/移动终端）组成。从业务流来看，用户在客户端启动播放请求，这个请求通过网络发出，到达并由VoD服务器的网卡接收，传送给VoD服务器。经过请求验证后，服务器把存储子系统中可访问的节目名准备好，使用户可以浏览到所喜爱的节目菜单。用户选择节目后，服务器从存储子系统中取出节目内容，并传送到客户端进行播放。

点播视频可以是运营商提供的IPTV业务，也可以是内容提供商提供的视频业务。

### 6.1.2.2 视频体验评估需求

影响点播视频体验的因素包括视频源质量、操作体验和播放体验三个模块的主要因素。点播视频评价模型主要分为三个部分，即视频质量、交互体验和播放体验。视频源质量主要取决于编码和码率等因素。交互体验即用户的互动过程体验，例如对于点播视频而言，从用户点击“开始”按键到视频开始播放的时延对于用于体验影响较大。播放体验即观看体验，例如在视频播放过程中由于网络性能的波动导致视频卡顿，会对用户体验产生较差的影响。

影响点播视频的主要因素有初始缓冲时延和卡顿。初始缓冲时延即从终端客户点击“开始”按钮到视频开始播放的时间差。卡顿即由于网络状态的抖动使得视频出现暂停，经过一段时间的缓冲后才能继续播放。点播视频的播放环境也会影响终端用户的感受，例如手持移动终端观看视频或在TV或者电脑前观看视频，用户对于视频质量如卡顿的忍耐程度不同。终端的大小，例如是手机、PAD或电视，用户对视频清晰度的要求也不同，因此研究点播视频的体验评估时需要考虑以上因素。

点播视频在进行体验评估时，还应同时支持加密和非加密两种场景。对于非标准分辨率的视频质量，也能够准确评估；还须考虑用户操作响应对视频体验造成的影响，如快进快退、拖动操作、EPG操作响应等。用户在使用点播视频业务时，与直播视频一样，应该考虑心理预期对用户体验产生的影响。

## 6.2 会议类视频服务

### 6.2.1 应用场景

一个多媒体会议服务提供不同位置的多个用户之间的实时通信，它具备良好的音频设施，并结合了参与者运动影像和/或传输多媒体信息。这类服务适用于公司的专有会议室以及出租使用的公共接入会议室。它适用于多种类型的多媒体会议终端。

该服务是基于电信网络的双向服务，它提供两个或多个多媒体会议终端的互连。其他类型的终端可以加入到会议中，诸如可视电话或甚至普通电话；虽然通常多媒体会议中各类信息的性能在发送和接收上会受到一些限制，但是至少能够交换语音，以允许会议用户参加讨论。

当会议包括两个以上的终端时，通常需要一个多点控制单元（MCU）。所有地点分别连接到MCU上，由它来对各地点来的信号进行选择或适当组合，并且管理信令和可选信道。

会议服务的视频通信允许用户使用自己的设备进行点对点视频通话，或者是参加多人的视频会议。视频通信架构如下图10，从功能上可分为三层：用户接入、媒体处理和业务平台。

用户接入层包括各种类型的终端和会议室，是用户接入通信系统的途径。媒体处理层包括MCU、企业UC (User Center) 和媒体网关等，负责媒体的适配、转发等处理。业务平台负责业务调度、资源分配等功能。视频通信的流程如下：用户发出视频通信请求，对端选择接受，双方建立通信连接，通过高速传输网络收发音视频等数据，进行视频通信。

视频通信业务可以采用单播或组播网络来实现，目前多采用单播模式。



图10 视频通信架构示意图

商企用户视讯会议的典型组图如图11所示。各类型的终端通过MCU接入视频通信系统，并且在业务平台的管理下，合理、高效地使用视频通信服务。管理员可通过Web Portal登陆业务平台，进行适当的配置和管理。



图11 商企用户视频会议服务组网示意图

## 6.2.2 视频体验评估需求

视频通信业务可以由运营商自建和运营的网络提供，也可以由服务提供商提供的应用来承担，如各类网络通信工具等OTT。影响视频通信的主要因素除音视频质量之外，还包括通话建立时延和通话体验。通话建立时延，即从点击连接通话，到连接成功的接听和画面展示的等待时间。通话体验包括：卡顿，即由于网络状态的抖动使得视频出现暂停，经过一段时间的缓冲后才能继续播放；花屏，即数据包丢失会起视频的失真；端到端时延以及音视频同步。其次，进行视频通信时，终端所处的环境也会影响用户的感受：例如在移动环境手持移动终端进行视频通信，与在固网环境TV或者电脑前进行视频通信相比，用户对于视频质量（如卡顿）的忍耐程度会不同。另外，终端的大小不同，例如手机/PAD与大屏电视相比，用户对视频清晰度的要求也不一样，因此需要考虑以上因素。

## 6.3 会话类视频服务

### 6.3.1 应用场景

多媒体会话业务通过电信网络给两个不同位置的用户之间提供了实时双向通信。它通常同时传输用户的音频和动态图像，和/或多媒体信息。特定的应用下可能不存在音频。该服务适用于专用终端设备，或基于微机的终端。

多媒体会话业务基本上是围绕会话通信任务来构建，也可选使用接收和发送等其他通信任务。控制功能是中间件服务组件会话控制。

在多媒体会话中，两个终端通过电信网络交换多媒体信息。基本配置是指两个终端通过双向连接点对点通信。这种连接通常是对称的，但在特殊情况下每个方向上的媒体组件可能不同，或者它们可以是相同的，但速率和服务质量不同。

除特殊情况外视频服务均需要音频，因此为了实现多媒体会话服务的基本功能，终端设备应包括下列必要的音频通信部分：麦克风、扩音器、音频编解码器、一些音频相关的控制（可选）。终端还必须包括一个网络接口单元。

处理多媒体文件的设备包括以下若干个功能部分：有屏幕和调制解调器的微机；配有相机或扫描仪、屏幕和调制解调器的静态图像设备；打印机（可选）。视频的基本装备包括摄像机、屏幕、视频编解码器。当传送视频时需要有展示输出图象的手段，也可提供屏幕来显示图片。

会话服务还可能有以下应用实例：

——远距离双方之间各种类型的谈话。它类似于采用电话进行语音会话，但相比之下增强配置了运动的视频、静止图片或文本。

——远距离双方，其中至少一人有听力困难需使用另一个媒体来协助进行沟通，这可能是由于残疾或环境嘈杂，例如使用可视电话、唇读以帮助听力障碍者、文本会话；

——两方共同处理一个文件，可能伴随有文件合作处理操作。

家庭用户主要进行点对点视频通话如图12，图中不同地点的家庭用户通过两个由互联网连接的终端在进行视频的沟通。

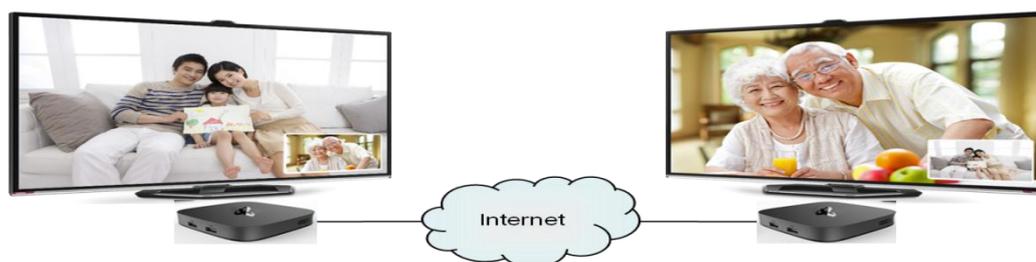


图12 家庭用户视频会话服务组网示意图

### 6.3.2 视频体验评估需求

会话服务与会议服务相近，均属于视频通信业务，二者的视频体验评估需要也相近。两种业务的参与方数量和通信任务的“通信配置”特性不同，对于视频源质量、音频源质量、互动体验、观看体验等带来的差异也需要进一步研究。

## 6.4 采集类视频服务

### 6.4.1 应用场景

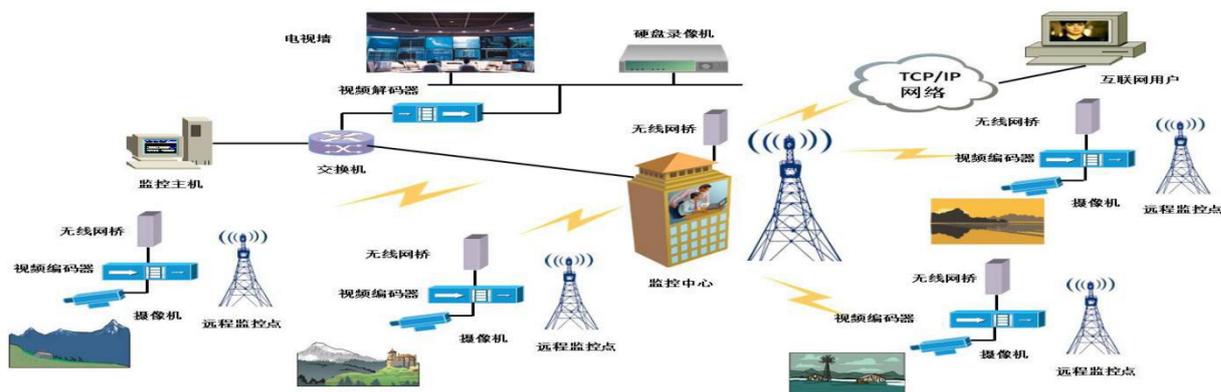


图13 视频采集场景示意图

随着网络技术、编解码技术及存储技术的不断发展，采集类视频业务的应用越来越广泛。公共场所的监控业务是采集类视频业务最常见的形式之一。以景点游客流量监控为例的视频采集场景如上图13。在每一个远程监控点，视频编码器将摄像机设备采集到的音视频信号经过压缩编码后再通过网络传输到监控中心。监控主机和互联网用户均可以通过网络从监控中心获取各个远程监控点的音视频信号。这些信号可以经过视频编解码器进行解码并在电视墙上进行播放，或者通过硬盘录像机保存。该业务可以帮助景点管理者应对各种突发状况，也会对互联网用户的相关出行决策起到一定的参考作用。

采集类视频业务主要是点对点的上行业务，当前终端上传视频的主流格式包括CIF、D1/4CIF、720P、1080P。单台终端（如：摄像头）的上行吞吐量主要取决于码率。影响码率的因素很多，主要包括分辨率、帧率、编码方式、压缩比、画面动态等。

### 6.4.2 视频体验评估需求

采集类业务可以由运营商自建，也可以由服务提供商提供云侧、端侧，运营商提供管道来合作建设。影响采集类视频的主要因素除音视频质量之外，还包括会话建立时延和业务体验。会话建立时延即从点击连接通话，到连接成功的画面展示的等待时间。影响业务体验的主要因素包括花屏、卡顿、视频降质和音视频失步等。各因素含义如下：花屏，即数据包丢失引起的视频失真；卡顿，即由于网络状态的抖动使得视频出现暂停，经过一段时间的缓冲后才能继续播放；视频降质，指带宽不足的情况下，视频编码的分辨率/码率主动下降以减少花屏/卡顿影响；端到端时延指视频数据包从离开源点时算起一直到抵达终点时一共经历了多长时间的时延，主要由排队延迟、发送延迟、传输延迟和传播延迟组成。音视频同步一般情况下是指在视频设备协同工作时需要视频与音频动作与口型要一致，没有交错的发生。

另外，采集终端的视频编码能力直接决定视频源质量，所以采集类视频业务体验受终端的影响很大。在进行体验评估的研究时需要一并考虑。

## 6.5 检索类视频服务

### 6.5.1 应用场景

随着计算机技术和网络技术的发展，以及多媒体的推广应用，出现了数字图书馆、数字博物馆等许多新服务形式和信息交流手段。

在传统的数据库系统中，信息的检索一般以数值和字符型为主，而在多媒体数据库中集成图像、视频、音频等类型信息，它们具有数据量大、信息不定长、结构复杂等特点。每一种媒体数据都有一些难以用字符数字符号描述的内容线索，如图像中某一对象的形状、颜色和纹理，视频中的运动，声音的音调等。当用户要利用这些线索对数据进行检索时，首先要将其人工转化为文本或关键词形式，这种转换带有一定的主观性，且极其费时，因而仅仅基于关键词的检索已不能满足用户的检索要求。数据库和其他信息系统不仅要能对图像、视频和声音等媒体进行存储以及基于关键字的检索，而且要对多媒体数据内容进行自动语义分析、表达和检索。

视频检索需要从大量的视频数据中找到所需的视频片段。根据所给出的例子或是特征描述，系统能够自动地找到所需的视频片段。根据提交视频内容的不同，视频检索一般分为镜头检索和片段检索。从用户的角度分析，他们对视频数据库的查询通常会是一个视频片段而很少会是单个的物理镜头。从信息量的角度分析，由几个镜头组成的视频片段有比单个镜头更多的语义，它可以表示用户感兴趣的事件。因此查询的结果也比较有意义。

由于视频拍摄的多样性和后期编辑的复杂性，片段的相似性有多种可能。把片段检索分为这样两种类型：精确和相似性检索。一个完整的视频检索系统的关键技术主要有：关键帧提取、图像特征提取、图像特征的相似性度量、查询方式、以及视频片段匹配等方法。

视频检索是一门交叉学科，以图像处理、模式识别、计算机视觉、图像理解等领域的知识为基础，涵盖了认知科学、人工智能、数据库管理系统及人机交互、信息检索等领域，引入媒体数据表示和数据模型，从而设计出可靠、有效的检索算法、系统结构以及友好的人机界面。在片源检索中，用户通过终端选择检索视频内容主题，获取需要的节目信息，如VoD内容检索，或是DTV检索。在信息检索中，用户通过终端检索信息类主题，如新闻检索等。

视频检索方法完全不像全文检索，在很大程度上也不同于图像检索。视频本身的层次化结构则要求视频检索必须层次化进行，因此视频的特征决定了视频检索必须是层次化的，且用户接口是多表现模式的。下面是几种常用的检索方法：

一是基于框架的方法：该方法通过知识辅助对视频内容建立框架，并进行层次化检索。

二是基于浏览的方法：基于浏览的方法始终是视频检索中一个不可缺少的方法。如果用户没有明确的查询主题或用户的主题在框架中没有被定义等，用户可以通过浏览来确定其大概目的。

三是基于描述特征的检索：该检索针对视频的局部特征检索，描述特征包括说明性特征和手绘特征。

四是视频的检索反馈在检索的实现中除利用图像的视觉特征进行检索外，还应根据用户的反馈信息不断学习改变阈值重新检索，实现人机交互，直到达到用户的检索要求。

随着网络技术的不断的发展，视频数据会有更多的应用，因此基于内容的视频检索技术有着很大的发展前途。

### 6.5.2 视频体验评估需求

该类业务的视频体验评估有待进一步研究。

## 6.6 信息类视频服务

### 6.6.1 应用场景

信息类视频服务目前主要包括以下几个方面：

一是信息提供，即通过视频系统向用户提供各种信息，例如：天气预报、股票交易信息、租房/售房信息、政府公告信息、电影预报、旅游信息、教育课程、用户的帐单信息、电话号码本和联系方式、网页黄页、免费电话等等；

二是广告，即针对特定区域用户定向投放的广告，用户通过EPG选择观看；

三是视频插播，在用户节目观看的中间提供其他视频内容，如广告或是紧急通告；

四是滚动字幕，在用户观看的节目上提供简短的消息提示或是广告。

### 6.6.2 视频体验评估需求

该类业务的视频体验评估有待进一步研究。

## 6.7 其他视频业务

随着视频业务飞速发展已产生众多的新型应用业务，下面以远程教育和远程医疗为例进行说明。

### 6.7.1 远程教育

远程教育即位于不同地点的学习者和学习资源之间的交互式和非交互式多媒体通信。学习者使用远程教育服务的目的可能是获得水平认证、参加业务培训，或独立学习专业知识。远程教育服务也可以有教学中心，它类似于传统的面对面的课堂学习、个别辅导、自学、多角色学习、团队学习等。

在远程教育的过程中，信息可能源于含有学习资源的远端数据库或者直播讲座。该材料可以是文本的、听觉的、图形或视频，也可以存储为一个多媒体格式。该信息可以点对点、点对多点或多点对多点方式传送。

在远程教育中学习者可能位于配备有相关设施的教室，也可能在办公室、家中或火车上等其他地方，在那里他们能够访问到远程教育服务平台（DLSP）。学习者在学习中可能伴随有与他人的实时交互，可能按照课程表学习，也可能按需学习非实时课程。学习者可能使用PC、PDA、移动电话，甚至一台配备有机顶盒的电视机。学习者可以在DLSP的协助下更换设备而不会中断学习。

远程教育服务系统由学习资源、DLSP和学习者组成。学习者通过远程教育服务平台从学习资源获取的学习材料。

学习资源包括由教师通过视听教室播出的直播课程、存储在服务器的直播课程数字视频记录、基于教师给出的课程讲义经特殊的软件制作、数字化材料（可视化图书馆，可视化实验室）、数字新闻等。学习者和DLSP之间、学习资源和DLSP之间可以是IP网、卫星通信网、移动通信网和PSTN等。

在VoD形式的多媒体课件点播中，教学内容是预先制作好的，用户通过点播的形式选择。

在在线课堂授课中，用户可通过遥控器，选择观看直播课堂内容的角度，并通过摄像头等参与课堂问答交互。

远程教育与视频会议、音视频点播、IPTV、基于Web的应用等密切相关。不同之处在于远程教育的目标是教学和学习，使用的信息是相关学习资源和学习记录，并且参与者是教师和学生。

### 6.7.2 远程医疗

远程医疗及顾问服务涉及位于多个地点的医学专家之间的交互式多媒体通信。这种通信通常由需要讨论特定病人的医生发起，并与相关专家进行通信。它可能发生在医生和一名顾问之间，也可能需要医生安排同时和几个顾问之间的互动会议。

在咨询的过程中，可能从包含患者医疗文件的远程数据库提取所需的信息：如从若干个诊断检测中心调取X光、超声波扫描、心电图或类似的医疗图像；或从参考图书馆提取技术信息、医疗影像说明或其他协助资料。这些资料可以是文本、音频、图形或图像，也可存储为多媒体格式。

参与咨询顾问的各方可能位于办公室或医疗设施中，这些地点具备全方位的宽带多媒体通信设施；也可能位于移动的车辆或者在一些其他偏远位置，只具备有限通信接入手段。为了应对所有的可能性，要求支持动态资源分配的规定，无论是在“呼叫”起始和“呼叫”进行中，均需要确保更重要的交互需要也能完全满足。

## 7 用户体验评估技术的扩展应用场景

### 7.1 用户自助评测业务质量

视频业务质量监测涉及宽带用户、网络、业务质量的综合监测管理，需要全程监测网络、视频业务质量，并聚合客户、网络、业务数据等信息，以可视化的形式呈现，实现对业务质量的评估与预警、故障的定位与分析，能够有效支撑运营商和智能终端厂家等的业务质量管理及运营管理需求。

类似家庭用户进行宽带测速以确认所购买的宽带服务质量，本标准系列涉及的用户体验评估技术在发展成熟后也可供终端用户对自身接收到的IPTV、OTT等固移视频服务进行量化测试评价。宽带测速场景是用户在测速网站的网页上进行操作，移动终端视频可通过安装APP完成。IPTV视频用户的实现方式可进一步研究。

目前在视频节目的质量监测中主要是单设备视角，对于用户体验的关注度偏少。当用户感知存在不足时，用户主要采用电话投诉等方式反馈问题，通常描述不够专业。如果能给用户提供更直观的评估界面，将有利于双方良好沟通，快速处理存在的问题，增加用户粘度。

### 7.2 运营商采集用户体验评估协助视频质量监测

用户体验评估技术所用的算法和参数在业务质量监测应用中将以软件模块形式嵌入监测工具（如STB内置软探针）中，作为延伸到宽带用户侧的网络信息采集、网络及业务质量拨测工具。根据实际需要，业务质量监测工具也可以与硬件设备集成，部署到网络相关节点，对网络质量、业务质量进行采集。

业务质量监测工具获取的信息能够涵盖网络、业务、客户三个层面数据。

业务质量监测工具与中心平台联动，实现数据采集、数据管理、数据分析与数据应用。

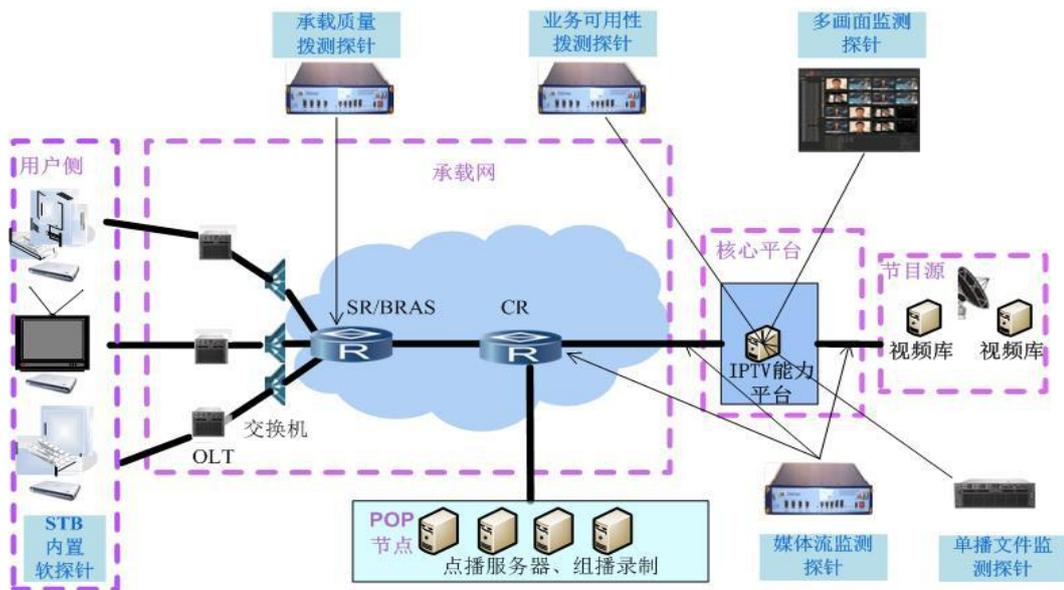


图14 运营商视频质量监测系统架构示意图

以固网IPTV为例，运营商视频质量监测系统架构示意如图14。IPTV质量监测系统由集中监测平台和IPTV质量监测工具构成。IPTV质量监测工具通过探针的部署，获取IPTV端到端各节点的关键指标，并将关键指标发送给集中监测平台。集中监测平台对收集到的关键指标进行统计、分析和关联，实现端到端的质量监测、故障告警、故障诊断、质量分析评估等应用功能。

IPTV质量监测工具包括几种：硬探针、软探针、手持测试仪。其中硬探针又分为两类：监测类探针、拨测类探针。监测类探针采用被动抓包为主，拨测类探针采用主动拨测为主。监测类探针包括：单播文件监测探针、媒体流监测探针（主要为直播频道媒体流监测，可包括组播和单播等传输方式，点播节目的媒体流监测应用场景较少）、多画面监测探针。拨测类探针包括：业务可用性拨测探针、承载网络业务和质量拨测探针（简称为承载质量拨测探针）。拨测类探针产品也称为拨测服务器。

STB内置软探针能够对机顶盒状态进行7\*24小时的实时监测。在用户使用业务过程中处于实时监测状态，用户不感知。监测模块可记录下用户的操作行为、视频播放的质量参数，并定期上报。

本标准所涉及的视频用户体验评估的算法和参数在软探针中有广泛的应用，是监测上报的重要参数之一，同时也是媒体流监测探针的主要上报参数。

IPTV质量监测系统与外部网管平台配合能够实现故障诊断定位、详细原因分析并满足客户服务、网络及平台维护、装机修障等场景下的各类需求。

在视频业务发展中，运营商面临着运维方面的诸多挑战，从用户侧到网络侧端到端的业务质量监测一直是运营商和第三方厂商头疼的问题，业务质量直接影响运营商业务的发展。然而当前网络指标无法反应业务质量、用户业务感知无法量化评估、业务质量指标不完善、缺少有效监控手段，急需端到端的全流程服务质量监控整体解决方案，并且需要从运营层面进行分析，以提升业务运营水平。

对于IPTV业务质量监测，要实现以下目的：一是系统性能强，确保质量监测的效率和准确率；二是在业务链路上多点部署，实现全视角运维；三是预设运行基线，强化告警管理，实现主动运维，四是以业务全流程数据采集分析，实现综合数据分析。