|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | |  |
|  | | |
|  | | |
| 方圆ok1 | | | |
|  | 视频体验衡量体系评价标准  U-vMOS白皮书 | | 附件1-16K |
|  | |
| 文档版本 | 01 |
| 发布日期 | 2015-09-15 |
|  | |
| 华为技术有限公司 | |
|  | | |

|  |
| --- |
| 版权所有 © 华为技术有限公司 2013。保留一切权利。  非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。  商标声明  附件3-版权声明页图和其他华为商标均为华为技术有限公司的商标。  本文档提及的其他所有商标或注册商标，由各自的所有人拥有。  注意  您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。  由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。 |

|  |  |
| --- | --- |
| 华为技术有限公司 | |
| 地址： | 深圳市龙岗区坂田华为总部办公楼 邮编：518129 |
| 网址： | [http://www.huawei.com](http://www.huawei.com/) |
| 客户服务邮箱： | [support@huawei.com](mailto:Support@huawei.com) |
| 客户服务电话： | 4008302118 |

前 言

概述

本文档阐述了华为定义的视频体验指标体系评价标准U-vMOS，该标准体系通过对主观视频体验的样本调研和技术研究分析，定义了一套量化的指标体系，从而形成完整的视频体验客观评价标准体系。该体系的建立，为视频业务体验提供了一个完善和统一的评估标准，为运营商和视频产业链提升用户体验给出了切实的指导建议，解决了一直困扰视频产业高速发展的一大难题。

读者对象

本文档只针对华为内部发布，不能直接发送给客户。主要适用于以下人群：：

* 客户经理
* 行销产品经理
* MKT营销经理
* SR

修订记录

|  |  |  |  |
| --- | --- | --- | --- |
| 日期 | 修订版本 | 描述 | 作者 |
| 2015年9月 | V1.0 | 视频体验衡量体系评价标准  U-vMOS白皮书 | 胡湘慧，黎静，孙丽娜，贺勇，宋新峰 |

目 录

[前 言 ii](#_Toc430359208)

[1 视频正在成为评价网络优劣的唯一标准 1](#_Toc430359209)

[1.1 视频是人类和社会信息发展的必然选择 1](#_Toc430359210)

[1.2 视频成为网络上的最主要流量 1](#_Toc430359211)

[2 U-vMOS提出的背景 3](#_Toc430359212)

[3 U-vMOS的设计思路 4](#_Toc430359213)

[3.1 主观测试方法和分制 4](#_Toc430359214)

[3.2 总体设计思想：跨屏幕、跨网络、跨业务 4](#_Toc430359215)

[3.3 视频体验的三大因素 5](#_Toc430359216)

[3.4 U-vMOS评价体系模型全景图 6](#_Toc430359217)

[3.5 视频质量（sQuality） 7](#_Toc430359218)

[3.5.1 人因工程定义视频源的体验极致 7](#_Toc430359219)

[3.5.2 满足人眼视觉特性的“身临其境”的关键视频指标 7](#_Toc430359220)

[3.5.3 最佳分辨率取决于人眼最小分辨角 7](#_Toc430359221)

[3.5.4 系统的色度空间应尽可能覆盖人眼可分辨的颜色 8](#_Toc430359222)

[3.5.5 动态范围与Bit深度对视觉感受的影响 10](#_Toc430359223)

[3.5.6 视频的帧率与景物运动相匹配才能获得最佳体验 10](#_Toc430359224)

[3.5.7 视频编码对SQuality的影响 11](#_Toc430359225)

[3.5.8 视频内容复杂度——画面清晰度 13](#_Toc430359226)

[3.5.9 屏幕显示尺寸——视频清晰度 14](#_Toc430359227)

[3.5.10 视频质量计算公式 15](#_Toc430359228)

[3.6 操作体验（sInteraction） 17](#_Toc430359229)

[3.6.1 操作体验影响因子介绍 17](#_Toc430359230)

[3.6.2 操作体验计算公式及典型值描述 18](#_Toc430359231)

[3.7 播放体验（sView） 19](#_Toc430359232)

[3.7.1 直播丢包花屏 19](#_Toc430359233)

[3.7.2 点播卡顿 26](#_Toc430359234)

[3.8 U-vMOS评价体系模型总公式 28](#_Toc430359235)

[4 U-vMOS的应用 30](#_Toc430359236)

[4.1 测量方法 30](#_Toc430359237)

[4.1.1 确定测量样本空间 30](#_Toc430359238)

[4.1.2 U-vMOS的具体测量方案 30](#_Toc430359239)

[4.2 应用案例 31](#_Toc430359240)

[5 华为对于U-vMOS的展望 32](#_Toc430359241)

# 视频正在成为评价网络优劣的唯一标准

## 视频是人类和社会信息发展的必然选择

视频是否是人们的刚需，在全球是否是普遍的大规模业务？

从生物学上分析，耳朵有约10万个神经细胞，眼睛有超过1.2亿个神经细胞，是耳朵的1200倍，神经细胞数量上的巨大差距体现在对外部信息量获取的差距。难怪大家会常常感叹“说一席话，不如画一张图，一张图表达不清的，最好用一段视频来示意”。眼睛是人类感知世界的最主要途径，视频才是人类生活的第一刚性需求。

从通信学上分析，耳朵是听觉的入口，连接语音时代；眼睛是视觉的入口，连接视频时代。进入数字时代后，只需要64kbps网络带宽就可以传递清晰的话音，而仅仅是高清视频，都需要不止10Mbps网络带宽。如今，音频已经发展到极致，22.2Hz沉浸式音频已经达到人耳听觉的极限。但视频从黑白到彩色，从标清到高清、4K，未来还有8K、增强现实AR、虚拟现实VR等发展。视频需求还远远没有得到满足。在信息爆炸的大时代中，视频是信息的主要载体，视频需求是一个全球性的普遍需求。

## 视频成为网络上的最主要流量

下面一组数据充分证明了视频将成为互联网主流业务的趋势。视频已经成为了人们的基本需求，是将成为继语音、短信、数据之后的运营商的基础业务，并将成为运营商的新增长引擎。

**全球IP视频流量**

2019年，每月134.8EB，从2014–2019年，复合年增长率为27%。

从2014到2019年，IP视频流量将增长3倍。

到2019年，消费类IP视频流量将占消费类IP流量的84%，相比2014年的75%有明显增长。

到2019年，商业IP视频流量将占商业IP流量的63%，相比2014年的36%有增长显著。

**全球互联网视频流量**

2019年，每月105EB，从2014–2019年，复合年增长率为33%

从2014到2019年，互联网视频流量将增长4倍。

到2019年，互联网视频流量将达到每月105EB，相比2014年的每月25EB增长显著。

2019年，互联网视频流量（商业和消费类）将占全部互联网流量的77%，相比2014年的59%有明显增长。

到2019年，消费类互联网视频流量将占消费类互联网流量的80%，相比2014年的64%有明显增长。

到2019年，商业互联网视频流量将占商业互联网流量的65%，相比2014年的38%有明显增长。

**高清和超高清（4K）互联网视频流量**

2019年，每月66EB，从2014–2019年，复合年增长率为53%

到2019年，高清和超高清互联网视频将占互联网视频流量的63%。

从2014年到2019年，高级互联网视频（高清和超高清）将增长8.5倍。

已安装/在用的4K电视将从2014年的990万台增长到2019年的3.715亿台（2014年占到平板电视总量的2.7%；到2019年将占到平板电视总量的31%，复合年增长率高达106%）

全球消费者视频点播（VoD）流量

到2019年，每月26.8EB，从2014–2019年，复合年增长率为14%

从2014到2019年，视频点播流量将增长近2倍。

高清视频将在运营商网络中无处不在，消费者对运营商网络质量的评价从原来的话音质量好坏和上网速率快慢逐步向视频观看体验优劣迁移， 构建以视频体验为核心的网络是运营商的提升品牌影响力和市场竞争力，从而取得商业上的持续成功所不得不面对的问题。

# U-vMOS提出的背景

目前，电信传统的语音、数据业务都有一套完整的质量指标评估体系。ITU-T给出了衡量通信系统语音质量的重要指标MOS（Mean Opinion Score），根据这个指标可以对运营商的语音网络质量进行评估。影响MOS的主要因素有以下几个方面：语音编码方案（AMR、HR 、FR or EFR）、Abis传输、Abis压缩、不连续发射（DTX）、C/I、切换频次及质量(RxQual) 对MOS的影响等，运营商以此来优化自己的网络从而提供MOS水平。而IPTV等视频业务还缺乏有效的评估手段，因此，当前在视频业务运营过程中很难评估用户对业务质量的体验，很难快速判断故障点。IP承载网是一个“尽力而为”的网络，网络视频业务占用带宽资源较多、实时性要求较高，并且对分组丢失、时延、抖动等网络特性非常敏感，尤其是时变的网络特性严重影响网络视频业务的质量。通过对网络视频质量的监控和反馈，可以调节编解码器或信道的参数，改善传输视频的服务质量。因此，需要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。有效的网络视频质量评估已经成为网络视频业通过对网络视频质量的监控和反馈，可以调节编解码器或信道的参数，改善传输视频的服务质量。因此，需要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。视频业务是一个需要高带宽、低传输时延和低时延抖动的业务，任何一个环节出现问题都可能导致用户观看视频时出现马赛克、停顿等现象，甚至无法观看，因此，视频业务的质量评估更具迫切性。

早在2009年ITU-T就启动了针对视频业务的vMOS标准研究项目，并于2012年参考语音MOS指标体系发布了第一个基于视频体验的VMoS指标，用于监控视频经过网络传输后的质量损失，关注点在于视频QoE的检测和问题定位。这套指标完全参考了语音MOS的定义，先定义影响因素Compression、packet-loss、rebuffering，自下而上地计算vMOS。出发点是为了发现问题，用于视频质量监控，只站在技术视角看问题，没有考虑消费者对视频体验优劣的评价是跨越视频业务的全流程，也没有站在最终消费者体验的角度去横向比较不同的分辨率带给用户的不同体验。因此也无法完整的指导运营商网络的设计和优化。

华为认为在原有vMOS的基础上，需要根据以用户体验为中心的评价体系标准，用统一的衡量标准，来评价不同网络，不同屏幕，不同场景应用下的视频体验的好坏。基于以上出发点，华为视频研究团队结合人体工程学实验，样本调研和深入技术研究，提炼出适配全场景的视频体验TOP3影响因子，即视频质量（sQaultiy），互动体验(sInteraction)和观看体验(sView)。华为基于三大核心思想，设计了视频体验衡量体系评价标准U-vMOS，使TOP3视频体验影响因子得以量化，使得视频体验标准体系实现可采集、可评估、可演进。

本次发布U-vMOS 1.0标准，范围包括平面屏幕上基于UDP的直播，以及基于HTTP STREAMING的点播。范围锁定在逼近人眼极限的各种屏幕大小下的视频观看体验。

# U-vMOS的设计思路

## 主观测试方法和分制

本文的主观测试采用单激励方法（single stimulus methods，SSM），设置打分制为五分制（1-5）。选取拟合人眼“沉浸式”感受的视频源，以完全无压缩的原始采集视频序列为当前观看的最佳体验，定为5分。主观测试方法遵循标准ITU-R BT. 500，ITU-T P.910，ITU-T P.911。具体主观质量评分标准如所示。

表格 1 视频主观质量评分标准

|  |  |  |
| --- | --- | --- |
| 分值 | 评分尺度 | 质量等级 |
| 5 | 质量很好，没有察觉失真 | 很好 Excellent |
| 4 | 觉察失真但不反感 | 好 Good |
| 3 | 觉察失真且稍微反感 | 一般 Fair |
| 2 | 觉察失真且令人反感 | 差 Poor |
| 1 | 觉察失真且令人很反感 | 很差 Bad |

## 总体设计思想：跨屏幕、跨网络、跨业务

U-vMOS 旨在构建一个统一跨屏、跨网、多业务的客观视频体验标准体系，通过客观的视频体验评估标准感知来反应视频业务体验的好坏。

在不同场景下根据用户对不同体验维度的重要性进行综合评价，得到用户整体视频体验vMOS评估准则和指导基线。该基线将成为未来评价视频网络性能的基本衡量标准。U-vMOS覆盖的场景将会不断扩展。从传统的娱乐类视频业务：点播和直播，到视频通讯和视频监控等实时通讯视频业务。



图 1 U-vMOS评价体系设计思想全景图

## 视频体验的三大因素

根据人因工程学实验6000+样本测试，以及3000+样本的用户调研，得到用户对于视频体验最关心的三大因素，分别是视频源的质量，播放启动时的操作体验和播放过程中视频是否有损伤。因此，U-vMOS评价体系主要涵盖这三个方面。

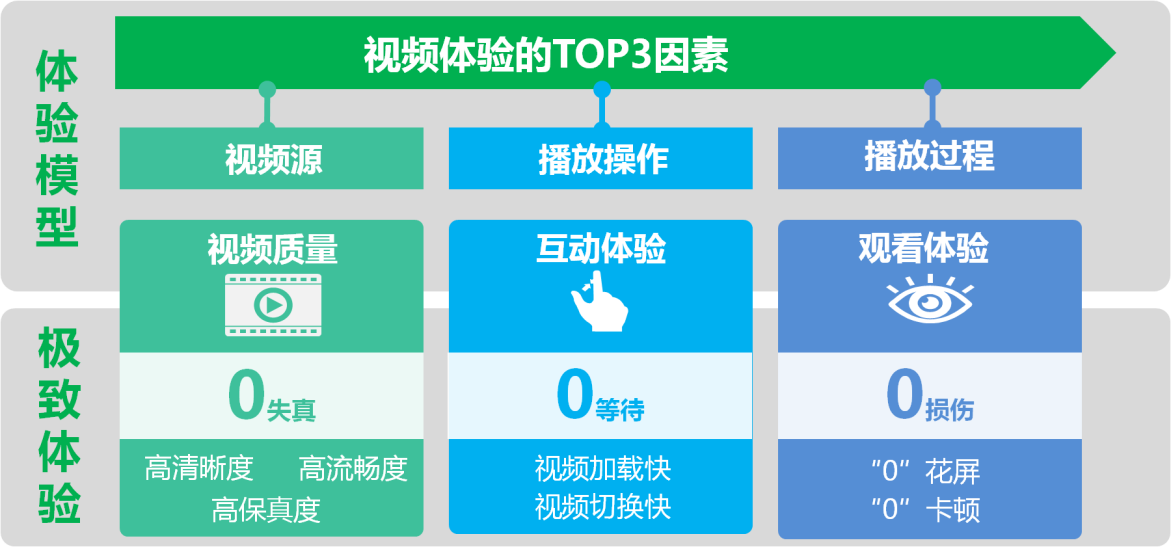


图 2 视频体验的TOP3因素

## U-vMOS评价体系模型全景图

影响视频质量，操作体验和播放体验三个模块的主要因素如图2所示U-vMOS的评价模型主要分为三个部分，即视频质量（sQuality），操作体验（sInteraction）和播放体验（sView），拟合出如下公式。

|  |  |
| --- | --- |
| 公式 1    图 3 统一视频体验指标体系U-vMOS建模方法    图4 sQuality的影响因子    图 5 sInteraction的影响因子    图 6 sView 的影响因子 |  |

下面将分别对三个关键因素分别进行详述。

## 视频质量（sQuality）

视视频质量影响因子如图4所示，包括视频分辨率，视频编码帧率，编码类型，编码码率，视频内容复杂度和最终显示器的尺寸。本章将详细描述这些因子如何合理取值。

### 人因工程定义视频源的体验极致

核心思想是通过人因工程学，从人眼体验极限的角度出发进行了深入研究，以人眼的“沉浸式体验”为满分设计，突破了原有vMOS仅能在同一分辨率下进行比较的局限，实现了在同一个标准下对不同分辨率和不同屏幕进行横向和纵向的比较。因此，计算得出的分值可以完全真实的反映客户对不同屏幕不同分辨率下观看视频体验的好坏。

### 满足人眼视觉特性的“身临其境”的关键视频指标

图 7 满足人眼视觉特效的浸入式体验

人眼是一个复杂而精密的系统，要想满足人眼视觉的“身临其境”体验，是视频内容源质量达到最高分，需要从分辨率，帧率，动态范围，色度等几个关键因素去最大限度拟合人眼的视觉最佳感受。

### 最佳分辨率取决于人眼最小分辨角

当观看距离越近时，对视频分辨率的要求越大。一般，视频分辨率越高，画质越好。其实对视频分辨率的要求取决于人眼的分辨力，即人眼分辨图象的细节能力称为分辨力，可用分辨角来衡量。它也反映了人眼的视力。

在量值上，分辨角用θ表示，因为实际的θ很小，它大致和可分辨的紧邻的两点间距成正比(视觉分辨率与视网膜retina上的视椎细胞大小为6um有关，这也是视网膜屏叫法的来源)，和观看距离L成反比， 即

θ ≈ d/L（弧度）＝ 360\*d/(2π\*L)(度) = 3438d/L (分)

如图所示。分辨角的倒数为分辨力。分辨力还和照度及景物相对对比度有关。正常视力的人，在中等亮度和对比度下，θ=1’-1.5’。这个值对电视系统扫描行数的选择提出了基本要求。在观看距离一定的情况下，若两相邻扫描线的视角小于分辨角，人眼已不能分辨，造成购机的浪费；若大于分辨角，则会有颗粒状的感觉，影响视觉效果。

因此，θ=1’之内，是分辨黑白两条扫描线的极限距离，那么，可以按照θ=1’之内(也即是下图d距离内)容纳1个像素来计算分辨率极限。

图 8 视角与分辨力关系

根据上述方法，可以得出不同屏幕在典型观看距离下的最佳分辨率要求（按2.0视力取值）



图 9 不同屏幕在典型观看距离下的最佳分辨率要求

### 系统的色度空间应尽可能覆盖人眼可分辨的颜色

色度（chromaticity）：指不包括亮度在内的颜色的性质，它反映的是颜色的色调和饱和度。包括色域，色深。系统的色度空间需要尽可能覆盖人眼可分辨的颜色。

色域(ColorGamut)：是对一种颜色进行编码的方法，也指一个技术系统能够产生的颜色的总和。在现实世界中，自然界中可见光谱的颜色组成了最大的色域空间，该色域空间中包含了人眼所能见到的所有颜色。 采用ITU-R Recommendation BT.2020 标准可以覆盖75.8%的人眼可见色彩空间，是目前拟合度最高的标准。

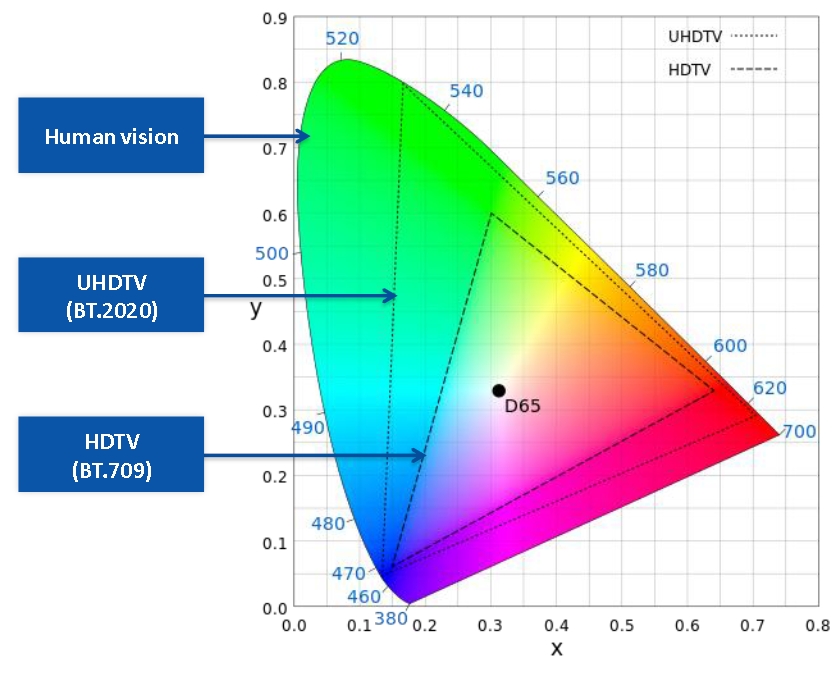


图 10色域

色深（Color Depth）**：**是一种主要取决于物体色的明度并与色调和彩度有关的对物体色的综合颜色感觉。色深亦可称为色位深度，是用 bit 数来表示数码影像色彩数目的单位；色深用 2 的幂指数来表示，bit 数愈高，色深值便愈高，影像所能表现的色彩也愈多。

如下图，8bits的影像出现了伪边界，而12bits的影像能够比较真实的还原物体的真实。



图 11 色深

### 动态范围与Bit深度对视觉感受的影响

人眼的亮度感知范围远大于目前显示器的显示范围，现实真正存在的亮度差，即最亮的物体亮度，和最小的物体亮度之比为 ， 而人类的眼睛所能看到的范围是 左右，但是一般的显示器，照相机能表示的只有256种不同的亮度。通过HDR（ High-Dynamic Range）技术可以提供更多的动态范围和图像细节，能够更好的反映人真实环境中的视觉效果。

图 12 动态范围

随着动态范围的增加，需要增加量化位数（Bit深度），否则可能产生产生伪轮廓；

图 13 不同Bit深度产生的效果

### 视频的帧率与景物运动相匹配才能获得最佳体验

现实世界是连续的，人眼是以连续的方式观察到现实中连续、清晰度运动；视频系统以离散度方式记录现实世界的运动，采集帧率与景物运动相匹配才能获得最佳体验。

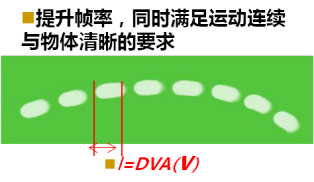
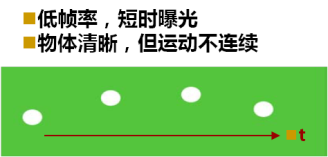


图 14 不同帧率对图像清晰度的影响

在观察运动物体时，人眼的分辨率（DVA）会随物体运动角速度变化。运动物体清晰度需要满足人眼动态分辨率（DVA）的要求，根据David M.H， Yasuhiro Y等的研究1，运动失真与运动速度及帧率等的关系如下：

公式 2

参数定义：f为帧率， Os是物体运动速度（deg/sec）,DVa为动态视觉精度（cpd），cff为闪烁频率，取40；

根据上述原理，测算出不同屏幕尺寸下不同的运动序列对帧率的最高要求

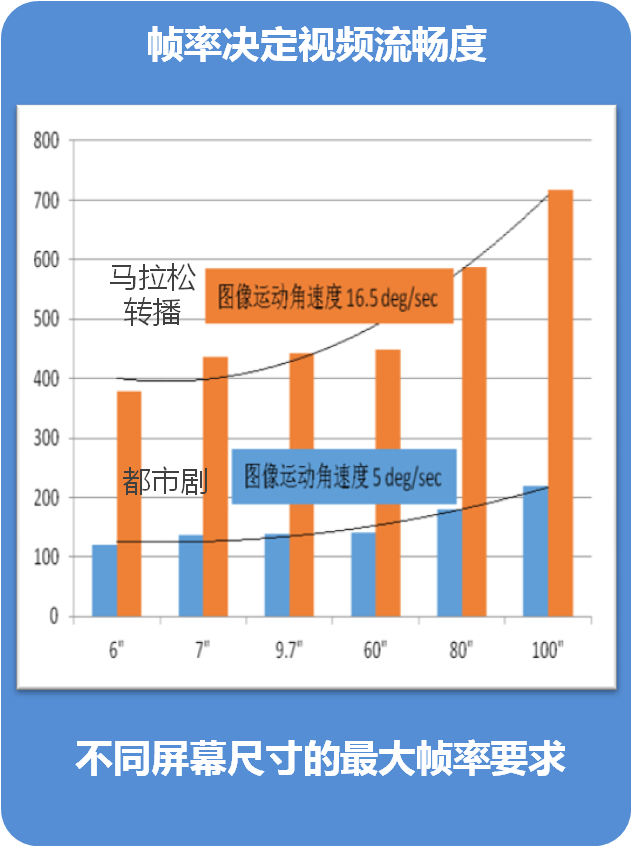


图 15 不同屏幕尺寸下不同的运动序列对帧率的最高要求

### 视频编码对SQuality的影响

前面我们定义了影响原始视频内容质量的关键因素，由于视频要经过IP网络传输，在传输之前要先经过编码，视频的编码类型（MPEG4，H.264，H.265，MPEG2，H.263等）及码率（Bitrate）大小是影响画面清晰度最重要的因素。当用某一种编码类型时，编码码率越高时，画面清晰度越高。

图16给出了码率与sQuality的关系图（以H.264/4K/60fps为例），可以看到，当码率越大时，质量越高，当码率大到达一定值时，质量趋于平稳。

图17，18给出了不同分辨率在相同编码类型和帧率时，不同码率时的质量（以H.264/30fps为例）。可以看出，在码率低的部分，相同码率下，高分辨率体验差于低分辨率；高分辨率要达到高体验，码率必须高于某个阈值。

图 16 编码码率与视频质量的关系（以H.264/4K/60fps为例）



图17 分辨率、码率与视频质量的关系图（以H.264/30fps为例）

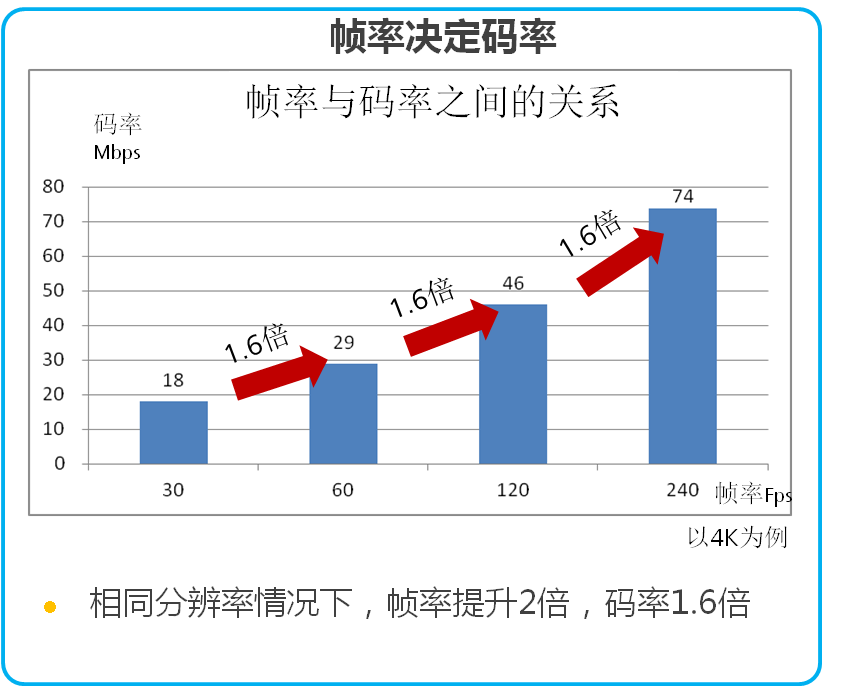


图 18 帧率对码率要求

### 视频内容复杂度——画面清晰度

在用某个固定的编码类型编码时，当视频的帧率，码率以及分辨率都固定时，视频源的内容复杂程度会影响画面的清晰度。视频内容复杂度一般包括时间复杂度（画面内运动剧烈程度）和空间复杂度（画面细节丰富程度）。

一般，在帧率，码率以及分辨率都相同时，时间复杂度较大的视频比时间复杂度较小的视频质量更差，空间复杂度较大的视频比空间复杂度较小的视频质量更差。如图9所示，当码率与帧率相同时，各个视频序列的质量分数仍有一定的差异，这是由于视频内容复杂度不同导致的。

图19 不同内容复杂度在相同码率下质量不同的示意图

### 屏幕显示尺寸——视频清晰度

随着各种智能终端的迅速发展，智能终端的显示器种类也越来越多。为了适应不同的显示终端，播放不同格式的视频时，需要对视频进行缩放，经过缩放后的视频在不同的屏幕上显示时，视频的质量存在差异。当播放视频源的分辨率小于屏幕分辨率时，播放器全屏播放会放大视频源，在原有视频源质量的基础上产生马赛克等视频损伤；当播放视频源的分辨率大于屏幕分辨率时，播放器全屏播放会缩小视频源，在原有视频源质量的基础上丢失视频细节；当播放视频源的分辨率等于屏幕分辨率时，播放器全屏播放，显示设备会将原始的视频源质量输出。

其中视频缩小产生的细节丢失通常很难被人察觉，但是视频扩大产生的马赛克现象相对比较明显。同一个视频源扩大播放，在点距相同的情况下，屏幕分辨率越大，视频源扩大的倍数越大，马赛克现象越严重，人眼的视频体验越差；在屏幕分辨率相同（放大倍数一定）的情况下，点距（屏幕尺寸）越大，马赛克现象越严重，人眼的视频体验越差。

图20给出了各个分辨率的无损视频在不同尺寸的屏幕上观看时的极限质量。可见屏幕越大，分辨率要求越高，不同分辨率的体验差异越明显。每种分辨率都存在sQuality极限，如果要提升用户体验，必须提升分辨率。

图20 不同显示屏幕尺寸对于sQuality的影响

### 视频质量计算公式

#### 公式因子描述

视频的显示屏幕尺寸被记为DisplaySize，视频内容复杂度被记为VideoComplexity，视频分辨率被记为Resolution，码率被记为BitRate，编码类型被记为CodecType，帧率被记为VideoFrameRate。则视频质量为

|  |
| --- |
| 公式 3 |

#### 公式及典型值具体描述

当帧率等于30fps（认为此时帧率为视频流畅播放的基准帧率）时，

|  |  |
| --- | --- |
| 公式 4 |  |

当帧率大于30fps时，

|  |  |
| --- | --- |
| 公式 5 |  |

当帧率小于30fps时，

|  |  |
| --- | --- |
| 公式 6 |  |

上述*v1-v10*为训练的参数值，其中，。当视频内容复杂度可以提取到时，*v3*和*v4*的取值与视频内容复杂度呈线性关系。

在不同尺寸的屏幕显示视频时，需要用有效显示PPI及屏幕尺寸（*ScreenSize*）来评估在特定设备上播放视频时用户的体验质量。有效显示PPI由原视频的分辨率，屏幕分辨率以及屏幕尺寸计算得到。

|  |  |
| --- | --- |
| 公式 7 |  |

各个分辨率的无损质量视频在不同尺寸的屏幕上显示的最高质量如表格 2所示。

表格 2 各分辨率在不同尺寸屏幕显示时的最高质量

## 操作体验（sInteraction）

### 操作体验影响因子介绍

##### 直播操作

直播操作时，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。因此频道切换的时延是影响直播操作体验的最重要因素。频道切换时延越长，体验越差。

|  |  |
| --- | --- |
| 公式 8 |  |

##### 点播操作

基于互联网点播操作经验，用户对点播初始载入时间的期望值较高。点播初始载入时间越长，体验越差。

|  |  |
| --- | --- |
| 公式 9 |  |

### 操作体验计算公式及典型值描述

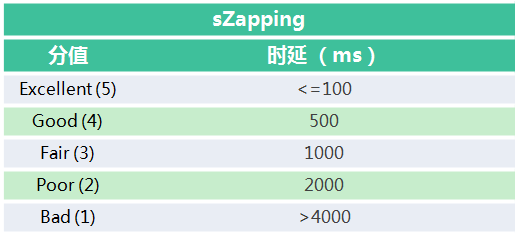
##### 直播操作

直播时频道切换时延被记为，则*sZapping*的计算公式如下

|  |  |
| --- | --- |
| 公式 10 |  |

上述**为训练的参数值。具体的典型分值如表格 5所示。

表格 3 sZapping与频道切换时延的关系



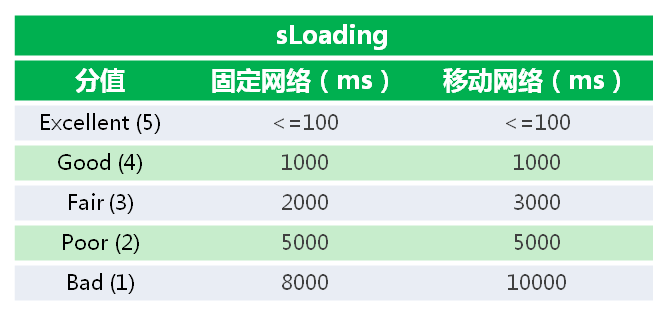
##### 点播操作

点播时初始缓冲时延被记为，则*sLoading*的计算公式如下

|  |  |
| --- | --- |
| 公式 11 |  |

上述**为训练的参数值。具体的典型分值如所示。

表格 4 sLoading与初始加载时长的关系



## 播放体验（sView）

视频在播放的过程中会由于丢包花屏（sBlocking）和卡顿（sStalling）而带来视频劣化现象。一般，直播时会发生丢包花屏现象，点播时一般会发生数据包到达延迟而引起视频卡顿现象。

### 直播丢包花屏

##### 丢包花屏影响因子介绍

（一）帧类型简介

在MPEG/H.26x视频编码标准中，有三种帧类型，分别是I帧(Intra-frame)，P帧(Predictive Frame)，B帧(Bi-directional predictive frame)。

I帧是帧内编码模式帧，表示关键帧，可以理解为这一帧画面的完整保留，解码时只需要本帧数据就可以完成（因为包含完整画面）；

P帧是帧间编码模式帧，表示的是这一帧跟之前的一个关键帧（或P帧）的差别，解码时需要用之前缓存的画面叠加上本帧定义的差别，生成最终画面。（也就是差别帧，P帧没有完整画面数据，只有与前一帧的画面差别的数据）；

B帧是双向差别帧，也就是B帧记录的是本帧与前后帧的差别，换言之，要解码B帧，不仅要取得之前的缓存画面，还要解码之后的画面，通过前后画面的与本帧数据的叠加取得最终的画面。

因此数据包丢失发生在不同的帧类型里，会引起的花屏效果是不同的。按照影响程度的比重，I帧大于P帧，P帧大于B帧。

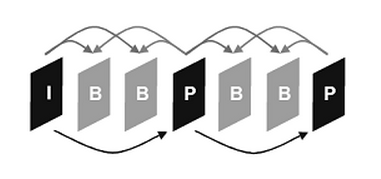


图21 帧类型示意图

（二）GOP简介

一般每个视频序列都包含有多个图像组（Group of pictures）。每个GOP从一个I帧开始，直到下一个I帧前，GOP长度是不固定的，可以几十帧一个GOP，也可以几百帧一个GOP，具体由视频的编码参数设定。GOP的示意图如图12所示。

图22 视频的GOP示意图

帧内编码模式帧（I 帧）和帧间编码模式帧（P 帧）数据的丢失会导致后续多个视频帧质量下降，时域错误传播通常会持续到下一个I 帧才结束。I帧编码时不参考其他任何帧，如果I帧的帧头信息损坏或丢弃，等同于丢弃整个I帧，就会影响整个GOP的视频。同样I帧中包丢失并没有损坏帧头信息，将会引起片（slice）损伤，并且也会持续到整个GOP结束；同I帧一样，如果P帧的头信息被损坏或丢弃，对视频的影响则等同于整个P帧被丢弃，P帧后面的所有帧都会受到影响直到GOP结束。如果丢包发生在P帧内，则会引起片（slice）损伤，影响持续到GOP结束；而双向帧间编码模式帧（B 帧）数据丢失只影响当前帧的质量。因此，评价数据丢失对视频质量的影响还需要考虑数据分组丢失的具体位置和所处的帧类型等特性。具体示意图如图13所示。



图23 分组丢失发生在不同类型帧造成的错误传播

（三）视频内容复杂度

当丢包位置和丢包个数等情况相同时，内容复杂度的差异会引起不同的观看体验效果。观看者对视频质量下降感知主要取决于I帧或P帧损伤的程度和这段视频的时间复杂度。在一个GOP内，如果P帧损伤得越早，整个GOP内的视频帧被损伤得就越多，直到接收到没有损坏的I帧结束。当P帧受损时，解码端依然要对受损视频进行误码掩盖。由于误码掩盖效果的好坏很大程度上取决于这段视频的时间复杂度，包丢失对视频质量的影响很大程度上也取决于视频内容和复杂度。

一般，运动复杂度越大，误码传播越严重，质量越差；空间复杂度越大，误码掩盖效果越好，质量稍好。主观体验如图24所示。



（a）运动复杂度大，花屏体验明显



（b）运动复杂度小，空间复杂度大，花屏体验不明显

图24 视频内容复杂度不同带来的主观体验差别

##### 丢包花屏计算公式参数介绍

数据包丢失会引起视频的误码失真（本帧数据丢失对本帧的影响）和误码传播失真（前面或后面的参考帧的误码失真对本帧造成的传播影响）。因此，只考虑整体的丢包率来计算丢包的影响并不准确，需要从视频帧的损伤程度方面进行考虑。视频帧的空间损伤程度和时间损伤程度，以及损伤的频率对于视频的失真程度有直接的影响。

简单来说，花屏sBlocking主要由花屏时长占比（TimeRatio，TT），花屏面积占比（BlockAreaRatio，BAR）和花屏次数（Frequency）共同影响。

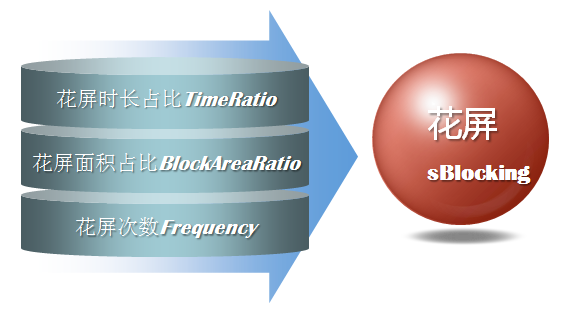


图25 花屏sBlocking的计算输入参数

对应于从视频里提取的参数，花屏时长占比对应于视频受损帧的比例，花屏面积占比对应于受损帧的平均损伤程度，花屏次数对应于一段时间内的损伤频率。具体描述如下：

(1) 受损帧的比例：在待评分序列中受到损伤的帧(包括有丢包的帧和受到误码传播的帧)占全部视频帧数的比例；

(2) 受损帧的平均损伤程度：受到损伤的帧的平均损伤比例；

(3) 一段时间内的损伤频率：损伤频率也称为丢包事件频率。一次丢包事件的失真将截止在I/IDR帧或场景切换帧处。当在某些应用场景下无法得知失真截止边界时，损伤频率可直接统计一段时间内的实际有丢包的帧的损伤次数，此时临近的受损的帧算在一次损伤里。

##### 丢包花屏计算公式

（一）公式因子描述

受损帧的比例（花屏时长占比）被记为*TR*，受损帧的平均损伤程度（花屏面积占比）被记为*BAR*，一段时间内的损伤频率（花屏次数）被记为*Frequency*，视频内容复杂度被记为*VC*。则丢包花屏为

|  |  |
| --- | --- |
| 公式 12 |  |

（二）公式及典型值具体描述

丢包花屏的失真因子被记为*sBlocking\_Dn*，当能提取到内容复杂度时，它的公式描述如下

|  |  |
| --- | --- |
| 公式 13 |  |

当提取不到内容复杂度时，可用经验值给*VC*赋值，此时公式描述如下

|  |  |
| --- | --- |
| 公式 14 |  |

sBlocking（sView）的具体公式描述如下

|  |  |
| --- | --- |
| 公式 15 |  |
|  |  |

上述*v11-v16*为训练的参数值。

特别的，在不考虑操作体验时，视频的体验*vMOS*由视频质量*sQuality*和只有丢包时的视频劣化体验质量*sView*共同得到，计算公式描述如下

|  |  |
| --- | --- |
| 公式 16 |  |

表格 7 花屏典型取值及质量

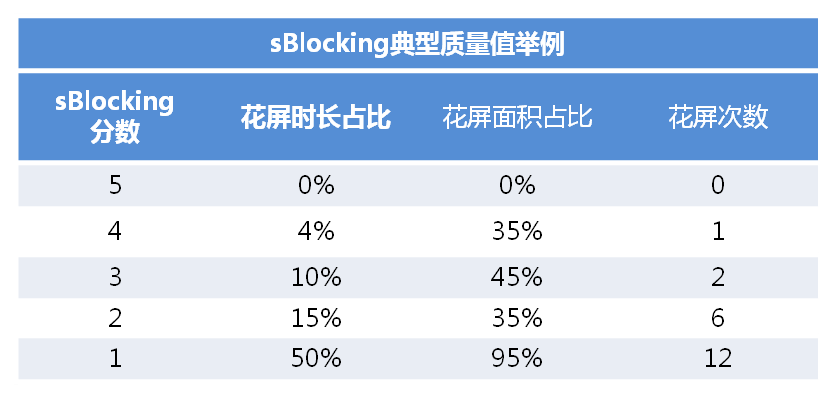




图26 固定花屏时长占比及面积占比时sBlocking与花屏次数的关系示意图

### 点播卡顿

##### 视频卡顿影响因子介绍

在视频播放的中间，视频由于数据包未能按时到达而进行缓冲，造成卡顿效果，也对视频体验影响较大。视频质量变差主要跟停顿时间的长短及多次停顿之间的间隔有关，在一个长视频观看的过程中，视频质量变差主要跟卡顿次数有关。当发生视频停顿现象时，停顿的时间越长，体验越差。当停顿结束恢复正常播放时，对于视频的质量体验就逐渐的缓慢恢复。假如后续一直能持续正常播放，实时的质量体验会恢复到正常值；但是如果在质量恢复过程中又出现一次停顿，则质量体验不仅跟本次停顿时长有关，还跟前一次停顿的时间距离有关。一个具体的实时质量分数变化如图27所示。



图27 视频中间停顿的实时质量分数图

##### 卡顿计算公式

（一）公式因子描述

多次停顿的平均卡顿时长被记为*Duration*，相邻两次卡顿间隔的平均值被记为*Interval*，卡顿事件的次数被记为*Frequency*。则卡顿表达式为

|  |  |
| --- | --- |
| 公式 17 |  |

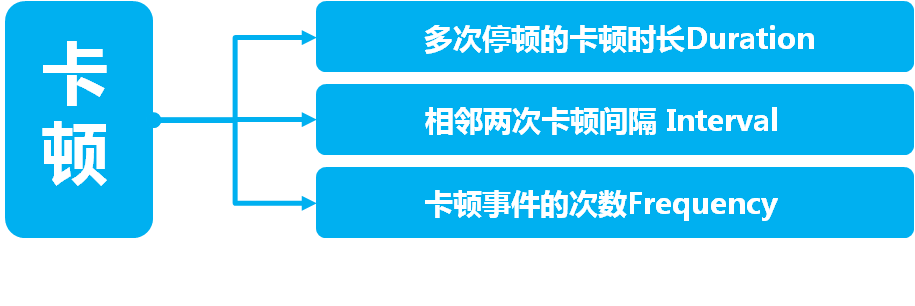


图28 卡顿的影响因子

（二）公式及典型值具体描述

卡顿的失真因子被记为*sStalling\_Dn*，它的公式描述如下

公式 18



sStalling（sImpariment）的具体公式描述如下

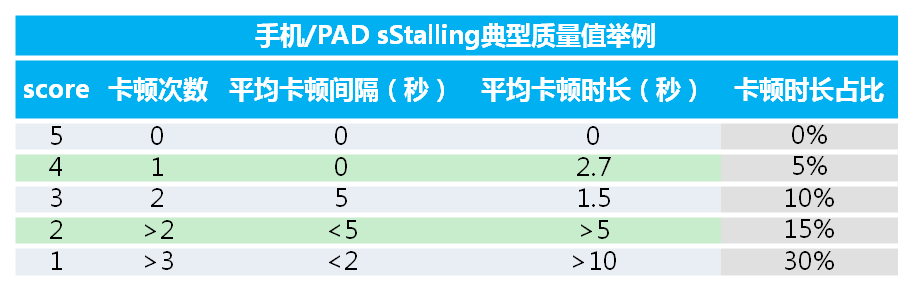
|  |  |
| --- | --- |
| 公式 19 |  |
| 公式 20 |  |

上述*v17-v22*为训练的参数值。

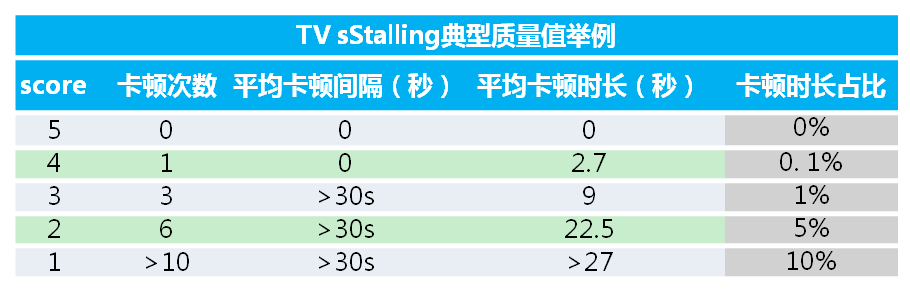
特别的，在不考虑操作体验时，视频的体验*vMOS*由视频质量*sQuality*和只有卡顿时的视频劣化体验质量*sView*共同得到，计算公式描述如下

|  |  |
| --- | --- |
| 公式 21 |  |

表格 6 卡顿典型取值及质量



（a）统计周期1分钟



（b）统计周期45分钟

## U-vMOS评价体系模型总公式

由上述三小节的计算公式分别得到视频质量（*sQuality*），操作体验（*sInteraction*）和播放体验（*sView*）后，U-vMOS的计算公式为

|  |  |
| --- | --- |
| 公式 22 |  |

其中，和分别是操作体验和播放体验的加权系数。具体的权重系数值由大数据调研的结果得到，详见图29。

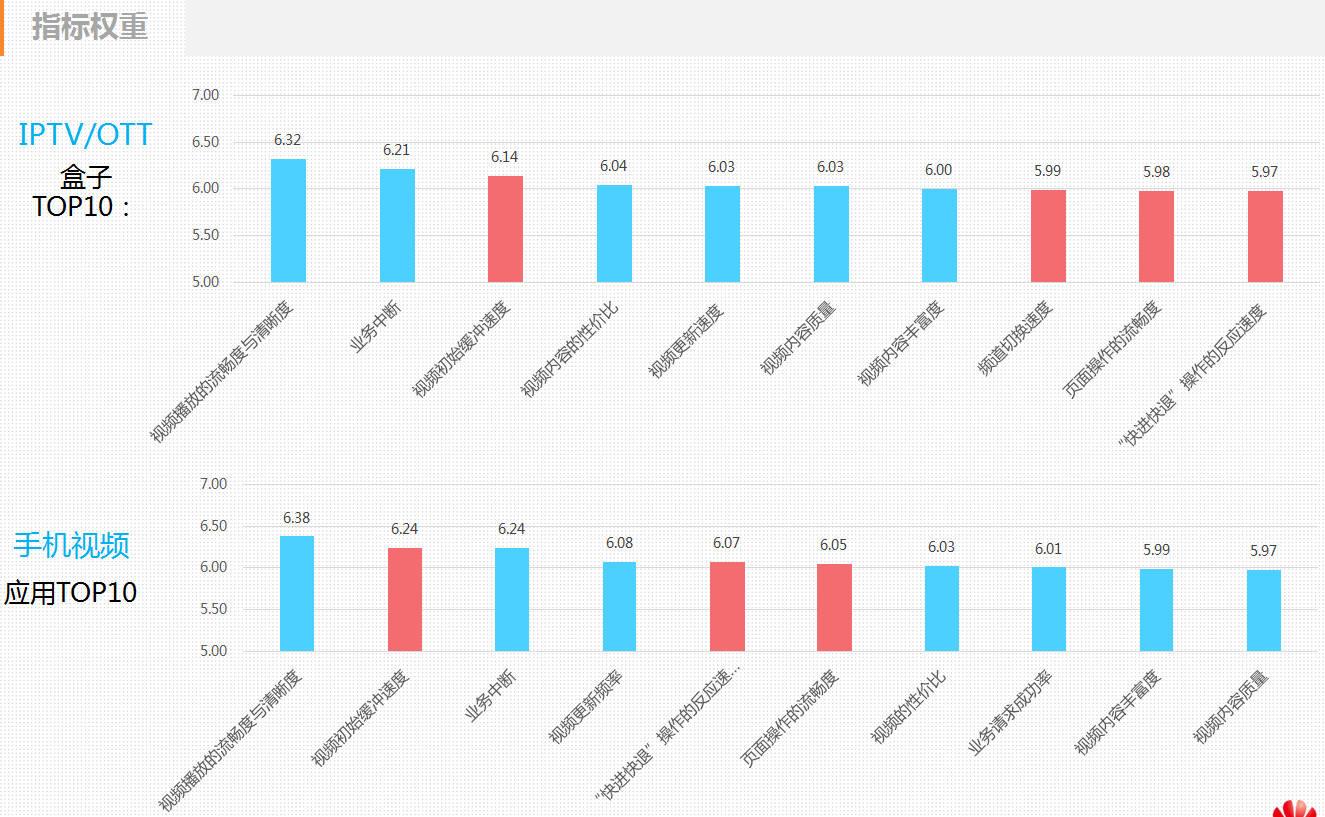


图29 观看视频业务的各项指标权重

表格 9 各项指标的权重值

|  |  |  |  |
| --- | --- | --- | --- |
|  | 参数 | 类别 | 权重 |
| 直播 |  | sZapping | 0.66 |
|  | sBlocking | 0.77 |
| 点播 |  | sLoading | 0.71 |
|  | sStalling | 0.77 |

# U-vMOS的应用

## 测量方法

### 确定测量样本空间

测量样本的选择决定了测量结果能否完整全面地反映被测网络的整体体验情况。样本包含三个维度的要求：

* 云样本：即选取哪些视频内容作为被测对象。云样本的选取要考虑覆盖不同的视频提供商的TOP热点节目的不同清晰度视频源。
* 管（网络）样本：即选择何种网络场景下的用户进行测试。管（网络）样本的选择需要考虑用户接入类型、用户套餐、不同区域、不同的网络层次、不同网络负载、不同线路质量（如：与基站距离的远近、光纤接入/铜线接入等）等因素。
* 测量时间段：不同的时间段，网络负载不同，测试时间段需要覆盖网络的忙时和闲时。忙时、闲时可由用户数或者流量大小来决定。宽带发展联盟认为通用场景下19：00～23：00为网络忙时，但不同运营商的用户行为可能不同，精确的忙时、闲时定义建议提前对运营商网络进行用户数及流量的评估，从而准确判断忙时、闲时时间段。

### U-vMOS的具体测量方案

U-vMOS的测量手段，主要有如下三类：

#### 视频播放器集成U-vMOS算法引擎

机顶盒、PC、手机/PAD等终端上集成华为U-vMOS算法引擎，通过视频播放器获取加载时长、卡顿、花屏等体验参数，调用U-vMOS引擎实时计算出结果并上报服务器。

优势：测量结果精确，覆盖范围全，实时获取结果。

限制：播放器客户端需要升级

#### 网络设备打点分析

中间网络设备可以基于播放器的关键行为特征针，对每用户的视频交互流程进行分析打点，判断加载时长、频道切换时长等指标进行测量。同时还能够基于CDN与终端的实时数据传输速率和传输的数据量判断卡顿指标。

优势：两端无感知，不需要客户主动发起测试。

限制：测量精度较差，部分指标无法测量，对于HTTPS加密传输的视频实施难度较大。

#### 专用的U-vMOS测试终端

提供一种基于X86/Android通用硬件盒子，集成各种视频的客户端播放器，同时集成U-vMOS算法引擎，有选择性地抽样部署在网络的不同位置（基站、接入机房或者部分用户家中），对视频质量进行例行监控。

优势：测试结果精确，实时获取结果。

限制：终端用户抽样有限。

## 应用案例

基于U-vMOS对东亚、中东、东南亚和拉美12个运营商无线网络的路测数据，统计发现：韩国代表全球视频体验标，目前平均U-vMOS：~3.7

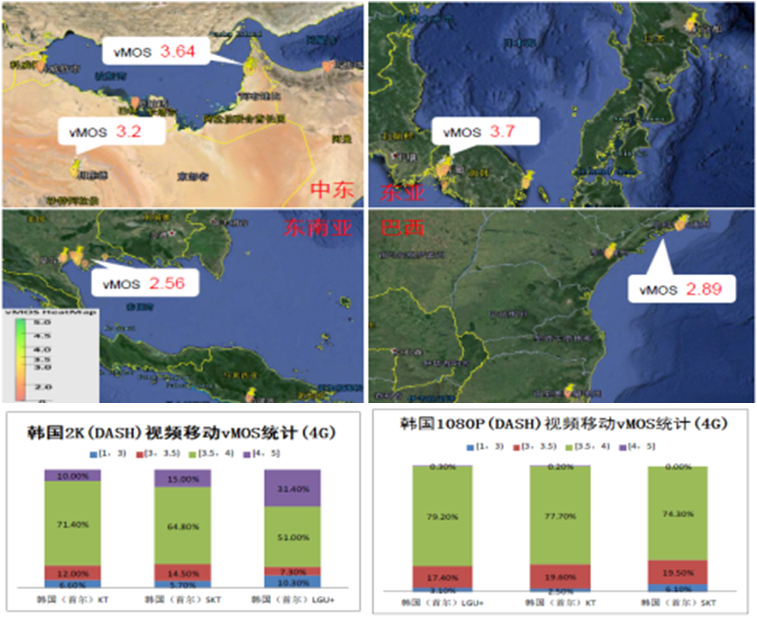


图 30全球视频体验指标评价案例

基于U-vMOS对中国大陆的3大运营商固定网络终端用户的抽样测试数据，统计发现：中国大陆1080P及以下的视频体验良好，但4K视频非常少，北京地区平均U-vMOS得分最高：~3.19

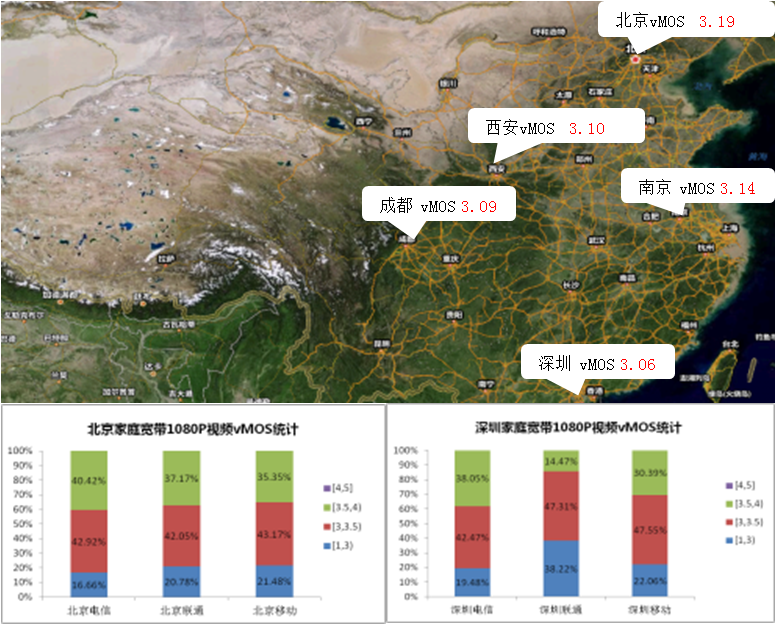


图 31中国大陆各运营商视频体验指标评价案例

# 华为对于U-vMOS的展望

华为在视频领域精心耕耘二十余年，致力于通过视频联接数十亿用户和数百亿物体的全新世界。华为将用户的体验总结为“ROADS”，即实时（Real-time）、按需（On-demand）、全在线（All-online）、服务自助（DIY）和社交化（Social）。华为vMOS视频评估体系可以帮助运营商构建并评估以ROADS为中心的最佳体验网络。

华为将与产业链合作，持续优化视频U-vMOS评估体系，与业界权威机构共同构建指标体系。并适配新业务，随着业务发展不扩充和更新。

华为将开发相关算法给产业链各方使用，从APP/SDK/算法/标准等多个层面向第三方开放。华为将在个产品线产品中集成U-vMOS计算引擎中间件，并开源发布提供给OTT/P3等第三方使用。并将推出U-vMOS体验测试工具，可安装在手机、STB等终端，成为普通用户视频体验感知评价手段。

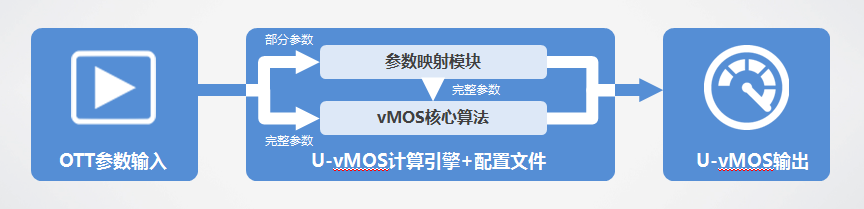


图 32 U-vMOS体验测试工具

华为还将发布基于U-vMOS体系，面向视频体验的网络架构，构建以U-vMOS为标准的评估－建网－优化－验收的闭环体系。

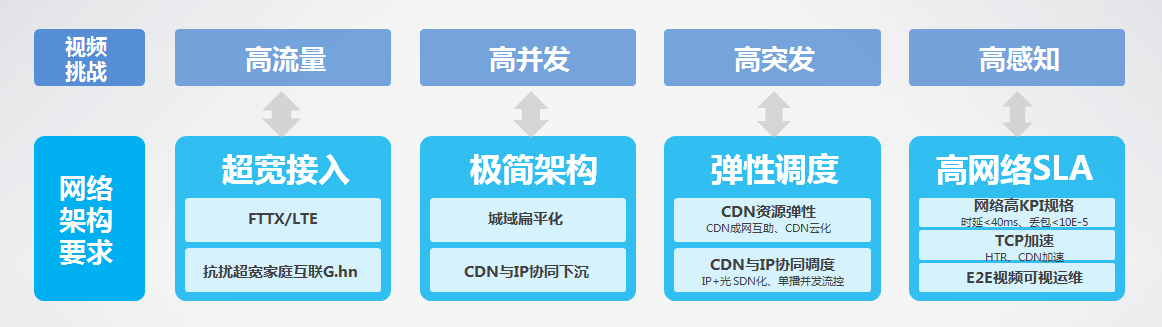


图 33 构建评估-建网-验收的闭环体系

华为将与产业链合作，持续优化视频U\_vMOS评估体系。华为希望通过这一指标体系的完善和发布，为视频行业提供统一的跨屏，跨网，多业务的视频体验指标评价体系，使得视频业务能够在统一的评价标准下，实现更加快速稳健的发展。