团体标准

《信息技术 媒体质量评价 第1部分：增强视频的编解码重建质量客观评价》（征求意见稿）编制说明

**一、工作简况**

1. **任务来源**

《信息技术 媒体质量评价 第1部分：增强视频的编解码重建质量客观评价》（以下简称增强视频客观质量评价）标准规定了视频质量客观评价技术，适用于评价由视频增强及编解码过程所重建的视频质量。其项目计划号为：2025041101。

1. **编制组成**

本标准起草组成员单位有：腾讯科技（深圳）有限公司、浙江大学、华为技术有限公司、天津科技大学、武汉大学、北京大学、国家广播电视总局广播电视科学研究院，北京邮电大学。

1. **编制背景**

数字音视频编解码技术标准化工作组（简称AVS工作组）由国家信息产业部科学技术司于2002年6月批准成立。工作组的任务是：面向我国的信息产业需求，联合国内企业和科研机构，制（修）订数字音视频的压缩、解压缩、处理和表示等共性技术标准，为数字音视频设备与系统提供高效经济的编解码技术，服务于高分辨率数字广播、高密度激光数字存储媒体、无线宽带多媒体通讯、互联网宽带流媒体等重大信息产业应用。自AVS工作组2002年成立以来，至今，已制订了多代视频/图像压缩相关的国家标准。增强视频客观质量评价标准针对原始视频由视频增强及编解码过程所重建的视频质量评价问题，通过相关算法和数学模型来模拟人类视觉系统对视频内容质量的感知。增强视频客观质量评价标准为视频编解码提供准确高效的客观质量评价准则，进一步为视频处理算法设计、编解码参数优化、面向显示设备的渲染优化等决策环节提供科学依据。

1. **编制过程**

工作组自2023年8月第86次会议开始，讨论启动增强视频客观质量评价标准的制定工作。基于综合应用需求分析，要求视频质量客观评价算法应支持1）从360P到4K多种分辨率，2）包括传统视频编解码器、感知优化的视频编解码器、端到端主观优化的AI编解码器所产生的多种失真类型，并输出N3631《视频编码质量评价标准的技术需求》（征求意见稿）。经过对潜在技术方案的详细梳理，以及国际最新视频质量客观评价标准发展态势的分析，和对已有视频质量客观评价算法性能分析之后，工作组于2024年6月第89次会议发布视频客观质量评价技术提案征集书（N3856），启动增强视频客观质量评价项目技术的征集工作。随后于2024年8月第90次会议输出了第一份视频客观质量评价技术工作组草案（N3913）。工作组共计召开了7次工作组全体会议、4次专题组加会、收到并讨论了58项增强视频客观质量评价标准相关提案，最终采纳2项视频质量客观评价方案并进行充分的迭代和性能验证，形成了此次的征求意见稿。

1. **主要起草人所做的工作**

腾讯科技（深圳）有限公司和浙江大学主要负责项目的组织与实施、关键技术的研发、标准文本的起草与编写、标准意见的征集与修改、定稿与报批工作。华为技术有限公司、天津科技大学、武汉大学、北京大学、国家广播电视总局广播电视科学研究院，北京邮电大学负责标准技术内容的讨论、意见的征集与修改、视频主观质量数据集构建、客观算法性能测试等工作。新一代人工智能产业技术创新战略联盟/中关村视听产业技术创新联盟主要负责标准技术讨论会议组织、编制和标准流程管理。

主要起草人：王海强、张家琪牵头标准制定工作，承担技术方案的设计实现、标准文本编辑和参考软件维护、组织标准文本修订等工作；主要起草人：张滢雪、陈震中、贾川民负责大规模主观质量数据集的构建、算法准确度与性能评测、标准文本审阅修订等工作；主要起草人：张立春、许晓中、刘杉、虞露、庄晓琪、郑月、赵寅、马思伟、石惠颖、雷正艺、谢嘉仪、宋莘鹏、王昊、郑蕤荻、闫石参与了标准制定过程中的讨论、算法性能测试和标准文本审阅修订等工作。

**二、标准编制原则和确定主要内容的论据及解决的主要问题**

1. 本标准的编制原则

本标准是为了适应视频处理链路中由视频处理与编码技术智能化对视频客观质量评价技术的更高需求而制定的，目标是为多媒体视频业务提供准确高效的视频客观质量评价准则。本标准的编制符合《GB/T 1.1-2020标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定。

本标准是以我国重点企业和科研机构为首、引领全球信息技术领域有影响力企业共同集体创新而成。为了形成优化的技术方案，工作组质量评价专题组根据标准的实际应用设定了详细的评估条件和测试方法，收集了大量反映各种典型情况的测试序列，技术评估的基本依据是综合考虑提案对质量评价准确度的贡献、方案实现复杂度和知识产权情况，这些原则与手段为本标准的技术先进性和妥善解决知识产权问题奠定了坚实基础。

1. 确定主要内容的论据

本标准是经过多次内部研讨，以科学严谨的态度面对征集意见，对内容进行反复商议和修改，从而逐渐优化形成的。在2024年3月召开的专题组会议上，多家单位共同参与，研究了标准的初步制定计划，详细论证了相关现有技术的不足，讨论了国际有关标准的实施现状，计划了本标准的核心需求与关键技术，经过一系列的讨论和研究，达成了本标准的研究技术路线。在随后召开的数十次专家组内部研讨会上，研究了标准编制的核心技术框架和性能指标，研究了标准制定的若干关键技术与核心算法，研究了用于技术验证的大批量测试序列构建方法，讨论了标准存在的技术和协调性问题，经过各单位专家的交叉检查与讨论，最终达成了标准在内容与技术路线上的一致性。

1. 解决的主要问题

本标准采用了基于深度学习的算法框架，面向不同业务场景形成无参考客观质量评价技术一项和全参考客观质量评价技术各一项。与国内外现有视频质量客观评价标准的不同之处在于，通过构建包含生成式视频编码、智能预处理与先进视频编码技术结合等多种增强视频编码范式的大规模视频质量评价数据集作为训练集，并针对性设计神经网络结构和训练方法，本标准实现了对增强视频重建质量的深度特征提取、更符合人眼感知的质量评价，解决了智能画质提升处理与视频编码技术深度融合带来的视频质量客观评价失准问题。

1. 有关技术问题说明

增强视频的编解码重建质量客观评价标准包括以下特色技术。

* **更高效的视频质量深度特征提取技术**
	+ **基于视觉语言多模态大模型（VLM）的无参考质量评价技术**

本标准中的无参考视频质量客观评价技术利用强大的预训练多模态大模型技术，通过多阶段的训练策略来模拟人类视觉感知特性，以实现高效的视频质量评估。通过在海量图文对数据上训练，多模态模型已经学习到了丰富的视觉概念、语言语义以及它们之间的对齐关系，为无参考评估提供了坚实的基础。一方面，基于VLM对视觉图像和自然语言指令的理解能力，通过精心设计的提示，可以更灵活、更符合人类直觉地定义和完成质量评估任务。另一方面，通过使用工作组自建的大规模视频主观质量数据集作为训练集，本方案在指令微调阶段进一步提升模型遵循指令进行质量评估的能力，实现了在无需参考视频的情况下，对视频质量进行自动化、高效且符合人类感知的评估。该标准所定义的无参考视频质量客观评价技术的架构见图1。



图 1 无参考视频客观质量评价技术方案系统架构图

* + **基于可变形卷积的多尺度特征融合技术**

本标准中的全参考视频质量客观评价技术利用ResNet-18作为主干网络，分别提取原始视频和重建视频帧级空域质量的深度特征，并使用可形变卷积网络对相同尺度的特征进行局部内容自适应融合。特别地，可形变卷积通过在感受野中引入偏置，可应对如移动、缩放、旋转等物体复杂形变的丰富场景，进一步提升深度特征的泛化能力。相比传统算法通过度量原始视频和重建视频深度特征的距离等来估计重建视频质量，该技术通过神经网络学习特征差异度量的最佳方式。通过上述多尺度特征融合技术获得视频帧级空域质量后，结合基于视频帧差特征的时域质量预测，实现了全参考条件下的增强视频编解码重建质量高效评估。该标准所定义的全参考视频质量客观评价技术的架构见图2。



图 2 全参考视频客观质量评价技术方案系统架构图

* **更多样性的质量评价训练集与测试集**

 本标准通过自建大规模视频主观质量数据集的方式，来支持以上深度学习算法模型的训练、验证与测试。相比于已有视频主观质量数据集，新建数据集有以下特色。

* + **测试视频内容多样性**

本标准测试所使用的数据集原始序列均为高质量拍摄的商业级高清视频，内容类型丰富且数量巨大，结合实际视频应用场景，涵盖了电视剧、电影、综艺、纪录片、动画等品类，涉及多种风格的室内室外、人物近景群像、古代现代建筑、自然风光、城市街景、舞台表演、体育赛事等场景。

* + **增强视频编码器多样性**

为了充分验证在多个编解码器场景的泛化性，在构建增强视频编码重建质量评价数据集的过程中，编码器类型的选择兼顾视频编码标准前沿、最新发展技术路线和实际业务应用，包括了2种先进的国内（AVS3）和国际视频编码标准（HEVC），3种端到端智能视频编码、1种面向主观质量优化的生成式智能视频编码以及1种智能预处理与先进视频编码技术结合的编码器。通过全方位、多层次、前瞻性、差异化的编码器选择策略，为训练高鲁棒性的质量评价算法提供了坚实的数据支持。

**三、主要试验或验证情况分析**

2024年12月，工作组发布了增强视频的编解码重建质量客观评价参考软件平台AVSVQA。相应地，根据不同的应用需求，工作组构建了大规模视频主观质量评价数据集用于客观质量评价技术的训练、验证与测试。

数据集包含1070段分辨率为1920×1080、帧率为25fps或30fps、色彩格式为YUV420P的典型视频序列作为参考序列，使用表1所示7个编解码器共计25个码率点对参考序列进行编码，生成26750段测试序列；参照国际标准ITU-T P.910建议的双刺激损伤量表评分方法，使用表2所示评分量表开展主观实验，对每段测试序列平均采集24个主观评分并计算得到平均意见分数（MOS）。

表 1 视频主观质量评价数据集构建所用编解码器配置

|  |  |  |  |
| --- | --- | --- | --- |
| **序号** | **编解码器** | **码率点数量** | **编解码配置** |
| 1 | X265 | 4 | CRF [20, 25, 31, 38] |
| 2 | AVS3 | 4 | CRF [27, 32, 38, 45] |
| 3 | X265Enh | 4 | CRF [23, 28, 34, 41] |
| 4 | EEM | 4 | Pre-trained |
| 5 | EEM-P-L | 2 | Pre-trained |
| 6 | EEM-P-H | 3 | Pre-trained |
| 7 | alphavc | 4 | Pre-trained |

表 2 评分量表

|  |  |
| --- | --- |
| 1 | 非常令人讨厌 |
| 2 | 令人讨厌 |
| 3 | 轻微令人讨厌 |
| 4 | 可感知但不令人讨厌 |
| 5 | 不可感知 |
| 6 | 轻微偏好 |
| 7 | 明显偏好 |

以参考序列为基准，对数据集所包含的测试序列和主观评分进行划分，其中800段参考序列及其对应的测试序列作为训练集，70段参考序列及其对应的测试序列作为验证集，200段参考序列及其对应的5000段测试序列和主观评分作为测试集。AVSVQA在测试集上的测试结果如表3所示。

表 3 算法准确度评价结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **客观质量评价方法** | **PLCC** | **SROCC** | **KROCC** | **RMSE** |
| Baseline | PSNR | 0.2700 | 0.3510 | 0.2998 | 0.9162 |
| FSIM-Y | 0.7654 | 0.6643 | 0.5466 | 0.5946 |
| VSI | 0.6648 | 0.5993 | 0.5003 | 0.6806 |
| VMAF | 0.8858 | 0.8319 | 0.6767 | 0.4351 |
| AVSVQA | 无参考视频客观质量评价技术方案 | 0.9450 | 0.9135 | 0.7845 | 0.2983 |
| 全参考视频客观质量评价技术方案 | 0.9091 | 0.8603 | 0.7039 | 0.3901 |

其中3个相关性指标，PLCC、SROCC、KROCC 的取值区间为[0,1]，其值越大代表客观分数与主观分数的相关性越强，客观算法的准确度越高。类似的，RMSE的值越小代表客观算法的准确度越高。对应指标的计算方式为先对每一个参考序列生成的所有测试序列计算上述指标，再对所有参考序列的同一指标取平均值，计算过程中对各技术方案的原始估计分数均做一次线性映射。

由表3所示测试结果可见，AVSVQA包含的无参考和全参考视频客观质量评价技术方案在PLCC、SROCC、KROCC、RMSE等4个指标上的表现均优于现有客观评价算法。

以上测试结果表明，本标准所提供的2项技术在性能上处于国际先进水平，尽快颁布增强视频的编解码重建质量客观评价标准，并在视频处理算法设计、编解码参数优化，面向显示设备的渲染优化等应用领域全面推广使用，将有利于促进数字音视频产业的跨越发展，为国家赢得巨大的经济利益。

**四、知识产权情况说明**

《信息技术 媒体质量评价 第1部分：增强视频的编解码重建质量客观评价》标准部分参与单位声明专利X项，包括技术提案单位或工作组成员正在申请或拥有的自主专利。这些专利或潜在专利的权利人均承诺同意将专利纳入AVS专利池统一进行实施许可。

**五、产业化情况、推广应用论证和预期达到的经济效果**

本标准旨在定义适用于视频经增强处理和编码后重建视频质量的客观质量评价方法。

本部分定义和使用增强视频客观质量评价技术标准的主要目的是：面向我国的超高清信息产业需求，联合产业链上下游企业与科研机构，制订覆盖视频采集、传输、处理、显示全链路的客观质量评价方法、提供评价依据及测试规范，为视频内容的生产、平台处理与编码、网络分发、终端播放等环节提供科学高效的评估工具，服务于超高清内容的制作、流媒体服务优化、智能监控系统、远程协作平台等核心信息产业应用。

**六、采用国际标准和国外先进标准情况**

本部分所规范的技术方案未采用国际标准，而是采用自主制定的技术路线，其基本原因在于有关国际标准背后不确定的专利费。同时，在本部分的制定过程中，不断吸收国内本领域更多研究机构成为会员单位，积极参与增强视频客观质量评价标准的制定。

目前国外暂无相应的标准发布。

**七、与现行相关法律、法规、规章及相关标准的协调性**

本标准符合相关的现行法律、法规和规章。

**八、重大分歧意见的处理经过和依据**

无。

**九、标准性质的建议**

鉴于本标准的内容，建议将该标准作为推荐性团体标准发布。

**十、贯彻标准的要求和措施建议**

建议按照专家意见修改送审稿并尽快完成编制以及标准的最终发布。同步开展标准宣贯培训与应用示范工作以便更好地引导产业发展，规范市场，从而促进超高清视频内容的落地和推广，为公众和行业提供更高效更高质量的媒体服务，让普通大众分享科技发展的红利。

**十一、替代或废止现行相关标准的建议**

无

**十二、其他应予说明的事项**

无。

《信息技术 媒体质量评价 第1部分：增强视频的编解码重建质量客观评价》标准编制组

 2025年7月1日