

GY

中华人民共和国广播电视和网络视听行业标准

GY/T 424—2025

制作域超高清视频近无损编码图像质量 主观评价方法

Subjective assessment methods for image quality of nearly lossless coding of ultra
high definition video in production domain

2025 - 08 - 18 发布

2025 - 08 - 18 实施

国家广播电视总局 发布

目 次

前言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	1
5 超高清视频近无损编码图像质量主观评价方法.....	1
5.1 主观评价条件、显示器技术要求及参数值.....	1
5.2 测试图像	2
5.3 观看员	2
5.4 评价周期	2
5.5 测试图像编排和显示.....	2
5.6 评价打分	3
5.7 打分表	3
5.8 评价数据的分析计算.....	4
参考文献	5

前 言

本文件按照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专利的责任。

本文件由全国广播电视和网络视听标准化技术委员会（SAC/TC 239）归口。

本文件起草单位：国家广播电视总局广播电视规划院、中央广播电视总台、鹏城实验室、北京广播电视台、上海广播电视台、广东广播电视台、四川广播电视台、湖南广播影视集团有限公司、江苏省广播电视总台、浙江广播电视集团、山东广播电视台、江西广播电视台、宁波广播电视集团、广州柯维新数码科技有限公司、四川新视创伟超高清科技有限公司、成都索贝数码科技股份有限公司、中移（杭州）信息技术有限公司、工业和信息化部电子第五研究所。

本文件主要起草人员：许帅、宁金辉、李岩、葛涛、王惠明、张乾、周立、潘晓菲、王立冬、尚峰、杨勇强、魏忠书、覃益明、朱光荣、陈晓东、韩国栋、温劲楠、朱红天、曾志华、王向炯、邹双泽、谢超平、缪云海、汪芮、刘斌、吴成志、刘汉源、朱易、潘波、赵新生、张宜静、王龙威、王刚、赵超、梁枫、周建海、管志鹏、余佳豪、罗天、张金沙、沙源、韦胜钰。

制作域超高清视频近无损编码图像质量 主观评价方法

1 范围

本文件规定了对超高清视频近无损编码图像质量进行主观评价的方法。
本文件适用于对制作域超高清视频的近无损编解码系统和设备的图像质量进行主观评价。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；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。

GB/T 41808—2022 高动态范围电视节目制作和交换图像参数值

GB/T 41809—2022 超高清清晰度电视系统节目制作和交换参数值

3 术语和定义

本文件没有需要界定的术语和定义。

4 缩略语

下列缩略语适用于本文件。

EOTF 电光转换函数 (Electro-Optical Transfer Function)

HDR 高动态范围 (High Dynamic Range)

SDR 标准动态范围 (Standard Dynamic Range)

5 超高清视频近无损编码图像质量主观评价方法

5.1 主观评价条件、显示器技术要求及参数值

主观评价条件应符合表1的规定，评价用显示器技术要求及参数值应符合表2的规定。

表1 主观评价条件

序号	项目		技术要求
1	环境亮度		$\leq 5 \text{cd/m}^2$
2	背景色温		D ₆₅
3	观看距离	3840×2160 图像	1.6 倍图像高度
		7680×4320 图像	0.8 倍图像高度

表2 显示器技术要求及参数值

序号	项目		技术要求及参数值
1	显示器尺寸	3840×2160 图像	对角线的尺寸应不小于 1.40m (55in)，宜大于 1.78m (70in)
		7680×4320 图像	对角线的尺寸宜大于等于 1.78m (70in)
2	显示器物理分辨率	3840×2160 图像	$\geq 3840 \times 2160$
		7680×4320 图像	$\geq 7680 \times 4320$

表 2 (续)

序号	项目	技术要求及参数值	
3	显示器色域	支持 BT. 2020 色域	
4	显示器峰值亮度 ^a (cd/m ²)	SDR	150~300
		HDR	宜大于等于 1000
5	显示器对比度 ^b	SDR	≤0.02
		HDR	≤0.000005
6	显示器显示比特深度	≥10bits	
7	显示处理	应支持被测视频 EOTF 曲线、色域和元数据 (如有) 的正确处理	
^a 峰值亮度是指 1%面积白窗在 100%峰值视频电平下的亮度。 ^b 该值为显示器黑场亮度与峰值亮度之比, 会受到环境光的影响。			

5.2 测试图像

测试图像宜为苛刻度较高的图像, 4K超高清视频优先选用GY/T 329—2020中适合评估清晰度、层次、彩色还原、运动特性的图像。测试图像的视频参数应符合GB/T 41808—2022或GB/T 41809—2022的要求。

一套测试图像应至少包含4个图像, 内容场景宜包含人物或动物、自然风景、不同字体及大小的滚动字幕、快速切换镜头、大型会议全景和主席台、行进的队列等, 甚至可采用计算机生成的特别复杂的难以压缩编码处理的测试图像。每个测试图像时长应为1s~10s。

5.3 观看员

观看员不应包括评价的组织者, 应为受过专业训练的专业观看员; 应具有正常的视力(含校正视力)和色觉; 应具有一定分析判断能力; 年龄宜在18岁~30岁之间。

每条视频主观评价所需观看员的人数应不少于15人。

5.4 评价周期

在每个评价周期开始时, 应向观看员详细、准确地介绍评价方法、质量要素或可能出现的损伤类型、评分标度、测试图像和评价时间长度, 并进行评价示范显示。

一个评价周期包括示范说明在内应不超过30min。在正式评价开始前, 应引入3个~5个示范图像来稳定观看员的判别力, 其评价数据不纳入测试结果统计。示范图像应不同于正式测试图像, 且能代表被测系统的损伤类型和损伤程度。为了评估相关性, 部分评价内容可以重复进行。

为便于评价数据筛选, 每次评价应引入若干控制图像对 (每个控制图像对包含源图像和损伤图像, 损伤图像应具有相对明显的质量损伤, 可使观看员很容易地区分源图像和损伤图像), 数量不少于测试图像总数的5%, 随机编排插入到测试图像中。

评价周期的显示流程图1。

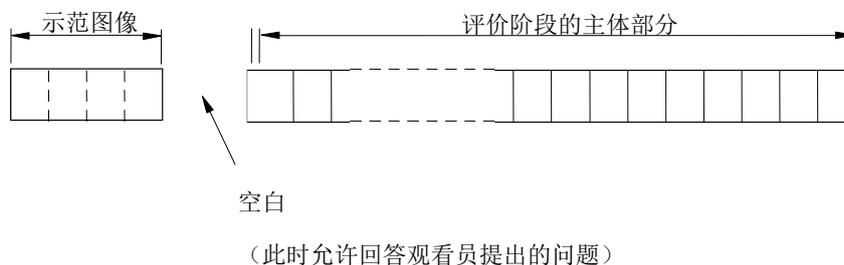


图1 评价周期的显示流程

5.5 测试图像编排和显示

一个评价周期由多次播放显示组成。图2为主观评价图像同屏拼接显示方式的示例, “A” “B” 分别表示图像的左半部分和右半部分。在每一次播放显示时, 显示器分为左右两侧, 一侧为源图像, 另一侧为处理图像 (由源图像和被测图像逐帧交替形成), 源图像和处理图像的显示位置 (屏幕左侧或右侧)

应以伪随机方式编排。评价组织人员需提前对经过近无损编码处理的图像进行质量分析，并标记出图像中损伤相对明显的区域，辅助观看员的评价。



a) 图像A部分的评价显示



b) 图像B部分的评价显示

图2 源图像和处理图像显示方式

评价时，视频的播放帧率可为实际帧率的1/5~1/10。

观看员可以反复观看图像，必要时可静帧观看，直到完成评价。在一组图像不同码率点的评价间隙，应允许观看员休息。

5.6 评价打分

评价时，观看员应在源图像和处理图像二者中进行选择，选出其认为的处理图像，如果不能准确判断，进行猜测也要给出判断结果。

一组评价过程中包含多个测试图像，应等待观看员完成对当前图像的选择确认后，再开始播放下一个图像。

5.7 打分表

评价时采用的打分表见表3。

表3 打分表

序号	源图像名称	打分			
		A (图像左半部分)		B (图像右半部分)	
		显示器左侧	显示器右侧	显示器左侧	显示器右侧
1	图像1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	图像2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	图像3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...	图像...	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
N	图像N	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
注：如果您认为显示器左侧或者右侧为处理图像，请在相应位置处画√。					
				观看员姓名： 年龄： 性别：	

5.8 评价数据的分析计算

5.8.1 评价数据收集

对于每个观看员，所有评价选择的数据都应收集并与所评价的图像严格对应。
对于每个测试图像，应收集所有观看员的评价结果。

5.8.2 纳入数据统计的观看员有效数据筛选

对于某个控制图像，观看员只要正确选择出图像左半部分（或者图像右半部分）的源图像和处理图像，则认为观看员对该控制图像的判断准确。在此规则下，观看员对所有控制图像的判断准确率需大于95%，其评价数据才可视为有效数据纳入统计分析。

5.8.3 评价数据的统计分析

评价数据分析时，对于每个测试图像，先分析每个观看员对图像左半部分的判断结果，如果判断准确，则认为是1，否则认为是0；图像右半部分同样按照此种方法判断。对于第j个测试图像，分别分析有效数据中图像左半部分的平均准确率 S_{j1} 和右半部分的平均准确率 S_{j2} ，计算见公式（1）和公式（2）。

$$S_{j1} = \frac{\sum_{i=1}^B x_{ij1}}{B} \dots\dots\dots (1)$$

$$S_{j2} = \frac{\sum_{i=1}^B x_{ij2}}{B} \dots\dots\dots (2)$$

式中：

B ——评价数据有效的观看员的个数；

x_{ij1} ——第*i*个观看员对于第*j*个图像左半部分的判断结果；

x_{ij2} ——第*i*个观看员对于第*j*个图像右半部分的判断结果。

S_{j1} 和 S_{j2} 较大的一个作为第*j*个测试图像的最终准确率 S_j ，计算见公式（3）。

$$S_j = \max[S_{j1}, S_{j2}] \dots\dots\dots (3)$$

S_j 接近0.5时，表示观看员为随机判断。 S_j 为(0.5, 0.75)时，表示观看员未察觉到源图像和被测图像之间的明显差异。 S_j 为0.75时，表示观看员刚好可察觉到源图像和被测图像之间的差异。 S_j 高于0.75时，表示观看员可明显察觉到源图像和被测图像之间的差异。

参 考 文 献

- [1] GY/T 329—2020 4K超高清视频图像质量主观评价用测试图像
 - [2] GY/T 340—2020 超高清清晰度电视图像质量主观评价方法 双刺激连续质量标度法
 - [3] ISO/IEC 29170-2:2015 Information technology — Advanced image coding and evaluation — Part 2: Evaluation procedure for nearly lossless coding
 - [4] ITU-R BT.500-15 Methodologies for the subjective assessment of the quality of television images
-