

GY

中华人民共和国广播电影电视行业标准

GY/T 314—2017

立体电视图像质量主观评价方法

Subjective methods for the assessment of stereoscopic television image quality

(ITU-R BT. 2021-1, Subjective methods for the assessment of stereoscopic 3DTV systems, MOD)

2017-12-26 发布

2017-12-26 实施

国家新闻出版广电总局 发布

目 次

前言	II
1 范围	1
2 规范性引用文件	1
3 术语、定义和缩略语	1
4 评价维度	2
5 主观评价方法	3
6 观看条件	15
7 图像序列	15
8 评价用设备	16
9 评价员	16
10 评价员须知	16
11 评价时长	16
12 参考图像序列	17
13 评分差异	17
14 评价员的筛选	17
15 统计分析	17
附录 A (资料性附录) 本标准与 ITU-R BT. 2021-1 相比的结构变化情况	18

前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

本标准使用重新起草法修改采用 ITU-R BT. 2021-1 《立体 3DTV 系统主观评价方法》。

本标准与 ITU-R BT. 2021-1 相比，在结构上有较多的调整，附录 A 中列出了本标准与 ITU-R BT. 2021-1 章条编号变化对照一览表。

本标准与 ITU-R BT. 2021-1 的技术性差异及其原因如下：

- 为符合 GB/T 1.1—2009 的要求，增加了第 1 章“范围”、第 2 章“规范性引用文件”、第 3 章“术语、定义和缩略语”；
- 删除了附件 1 “其他评价维度”相关内容，因 ITU-R BT. 2021-1 缺乏对“其他评价维度”的评价方法；
- 修改了附件 1 中图 3、图 5、图 6，在图像序列上叠加文字标识，以方便评价员依序观看和打分；
- 删除了附件 1 中 2.6.1 “如果可能，两个图像序列可在同一监视器上并排播放”及图 8，因实际中一般采用两个并排摆放的监视器的实现方式，此方式极少采用；
- 删除了附件 1 中表 4 中 1280×720 显示设备的最佳观看距离，因 GY/T 155—2000 在参照 ITU-R BT. 709-3 时删除了分辨率 1280×720。

本标准还做了下列编辑性修改：

- 删除了附件 1 中第 1 章关于立体视觉深度感知的论述性语句；
- 删除了附件 1 中第 1 章关于 ITU-R BT. 500 图像质量评价方法和评价维度的论述性语句；
- 删除了附件 1 中 2.2.1 关于图像序列时长选取的论述性语句；
- 删除了附件 1 中 2.2.3 关于双刺激损伤标度法的重复的语句；
- 删除了附件 1 中 2.3.3 关于双刺激连续质量标度法实现方式的重复的语句；
- 删除了附件 1 中 2.5 关于单刺激连续质量评价法的论述性语句；
- 删除了附件 1 中 2.6.1 关于同时双刺激连续质量评价法的论述性语句；
- 删除了附件 1 中第 3 章关于观看环境和观看距离的论述性语句；
- 删除了附件 1 中 4.1 关于视觉舒适限值的论述性语句；
- 删除了附件 1 中 4.2 关于左右眼图像的误差的论述性语句；
- 删除了附件 1 中第 5 章关于评价设备选取的论述性语句；
- 在附件 1 表 4 中增加了以毫米为单位的屏幕尺寸的数值；
- 删除了附件 1 中 6.1 关于样本容量的论述性语句；
- 删除了附件 1 中 6.2 关于视觉检查的论述性语句；
- 删除了附件 1 中第 7 章关于伦理准则的论述性语句；
- 删除了附件 1 中第 9 章关于参考序列的论述性语句；
- 删除了附录 1 “视觉检查用测试素材”。

请注意本标准的某些内容可能涉及专利。本标准发布机构不承担识别这些专利的责任。

本标准由全国广播电影电视标准化技术委员会（SAC/TC 239）归口。

本标准起草单位：国家新闻出版广电总局广播电视规划院、中央电视台、中国传媒大学、清华大学。

本标准主要起草人员：董文辉、路晓俐、戴琼海、邓向冬、郑冠雯、张乾、姜秀华、史萍、王珮、范静涛、郝涛、汪芮。

立体电视图像质量主观评价方法

1 范围

本标准规定了立体电视图像质量的主观评价方法。
本标准适用于立体电视系统和设备图像质量的主观评价。

2 规范性引用文件

下列文件对于本标准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标准。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本标准。

ITU-R BT. 500 电视图像质量主观评价方法 (Methodology for the subjective assessment of the quality of television pictures)

ITU-R BT. 2022 平板监视器的SDTV和HDTV图像质量主观评价的通用观看条件 (General viewing conditions for subjective assessment of quality of SDTV and HDTV television pictures on flat panel displays)

3 术语、定义和缩略语

3.1 术语和定义

下列术语和定义适用于本标准。

3.1.1

立体电视 stereoscopic television

利用人眼的视觉特性进行拍摄、制作，展现景物的实际空间关系和立体感的电视技术。在显示终端上观众能感受到观看的影像是具有深度特性的三维立体场景，增强观看的临场感和沉浸感，也被称为3D电视。本标准中所指立体电视均为双画面立体电视。

3.1.2

双画面立体电视 binocular stereoscopic television

用双摄像单元模拟人的双眼同时拍摄的、或计算机生成的左右眼两路图像信号，经过制作、播出、传输，最后显示在立体电视屏幕上，使观众左右眼分别观看到具有适当视差的左右两幅图像，从而在大脑中形成三维立体效果的电视。

3.1.3

视差 disparity/parallax

被摄物体上的同一点在左右眼图像上对应像素的水平位置差。观看者感觉到被摄物体的影像在显示屏幕后方的是正视差，在显示屏幕前方的是负视差，在显示屏幕上的是零视差。

3.1.4

视差角 disparity angle/parallax angle/angular disparity

双眼看到的屏幕上的实像会聚角与屏幕前、后方的虚像会聚角之差。

3.1.5

评价单元 judgement trial

图像质量主观评价中，图像序列播放和打分的最小单元，也是能够体现主观评价方法和获得主观评价结果的一个最小单元。

3.1.6

评价段 test session

一个评价段由若干个评价单元组成，评价段之间安排休息时间。

3.1.7

保真度 fidelity

被测系统相对于图像质量无损伤的理想系统的性能。它可通过对比源图像序列和被测图像序列的主观评价打分来计算。

3.2 缩略语

下列缩略语适用于本标准。

DMOS 平均评价分数差异 (Difference Mean Opinion Score)

DSCQS 双刺激连续质量标度法 (Double-Stimulus Continuous Quality Scale method)

DSIS 双刺激损伤标度法 (Double-Stimulus Impairment Scale method)

MOS 平均意见分 (Mean Opinion Score)

SC 刺激比较法 (Stimulus-Comparison method)

SDSCE 同时双刺激连续质量评价法 (Simultaneous Double Stimulus for Continuous Evaluation method)

SS 单刺激法 (Single-Stimulus method)

SSCQE 单刺激连续质量评价法 (Single-Stimulus Continuous Quality Evaluation method)

4 评价维度

4.1 概述

立体电视图像质量的主观评价维度可包括图像质量、深度质量、观看舒适度三个基本评价维度。本标准给出了采用图像质量、深度质量、观看舒适度三个基本维度进行评价的方法。

4.2 基本评价维度

4.2.1 图像质量

图像质量是指评价员对立体电视图像优劣的主观感知。图像质量主要受到技术参数和引入误差（如编码和（或）传输过程）的影响。

2D电视图像质量的评价要素，如清晰度、色彩重现、运动表现、整体质量、锐度等也可用于立体电视图像质量的评价。

4.2.2 深度质量

深度质量是指评价员对立体电视图像深度表现能力的主观感知。

评价要素包括：

- 深度分辨率，即深度方向上的分辨率；
- 深度运动，即深度方向上的运动或移动是否平滑流畅；
- 空间变形失真，比如木偶效应，即深度方向上物体的视觉比例出现失真；卡片效应，即物体是立体的，但深度方向上物体变薄了，产生厚度失真。

4.2.3 观看舒适度

观看舒适度是指评价员对立体电视图像观看舒适程度的主观感受。不恰当的拍摄或显示等可能导致观看不舒适。

5 主观评价方法

5.1 主观评价方法概述

ITU-R BT. 500建议书针对各种被测对象（包括同一算法不同压缩参数、同一编码方式不同码率或不同传输场景等）给出了多种图像质量评价方法，本标准采用了其中六种评价方法：

- a) 单刺激（SS）法；
- b) 双刺激损伤标度（DSIS）法；
- c) 双刺激连续质量标度（DSCQS）法；
- d) 刺激比较（SC）法；
- e) 单刺激连续质量评价（SSCQE）法；
- f) 同时双刺激连续质量评价（SDSCE）法。

各种评价方法的图像序列的显示时长和评分标度汇总见表1、表2和表3。

具体评价时，可对表中的评价方法进行一些细微的调整。例如，单刺激（SS）法的离散标度也可采用七级标度。

表1 图像质量主观评价方法

评价方法	图像序列时长	二选一标度	离散标度	连续标度										
单刺激（SS）法	约 10s	——	5 优 4 良 3 中 2 差 1 劣	<table style="border: none; margin-left: auto; margin-right: auto;"> <tr><td style="border: none;"> </td><td style="border: none;">优</td></tr> <tr><td style="border: none;"> </td><td style="border: none;">良</td></tr> <tr><td style="border: none;"> </td><td style="border: none;">中</td></tr> <tr><td style="border: none;"> </td><td style="border: none;">差</td></tr> <tr><td style="border: none;"> </td><td style="border: none;">劣</td></tr> </table>		优		良		中		差		劣
	优													
	良													
	中													
	差													
	劣													

表 1 (续)

评价方法	图像序列时长	二选一标度	离散标度	连续标度										
双刺激损伤标度 (DSIS) 法	约 10s	——	5 损伤不可察觉 4 损伤可察觉, 但不令人讨厌 3 损伤令人有些讨厌 2 损伤令人讨厌 1 损伤令人很讨厌	——										
双刺激连续质量标度 (DSCQS) 法	约 10s	——	——	<table style="border: none; margin: auto;"> <tr><td style="border: none;"> </td><td style="border: none;">优</td></tr> <tr><td style="border: none;"> </td><td style="border: none;">良</td></tr> <tr><td style="border: none;"> </td><td style="border: none;">中</td></tr> <tr><td style="border: none;"> </td><td style="border: none;">差</td></tr> <tr><td style="border: none;"> </td><td style="border: none;">劣</td></tr> </table>		优		良		中		差		劣
	优													
	良													
	中													
	差													
	劣													
成对比较 (PC) 法	约 10s	A 或 B	-3 差很多 -2 差 -1 稍差一些 0 相同 1 稍好一些 2 好 3 好很多	——										
单刺激连续质量评价 (SSCQE) 法	约 3min~5min	——	——	<table style="border: none; margin: auto;"> <tr><td style="border: none;"> </td><td style="border: none;">优</td></tr> <tr><td style="border: none;"> </td><td style="border: none;">良</td></tr> <tr><td style="border: none;"> </td><td style="border: none;">中</td></tr> <tr><td style="border: none;"> </td><td style="border: none;">差</td></tr> <tr><td style="border: none;"> </td><td style="border: none;">劣</td></tr> </table>		优		良		中		差		劣
	优													
	良													
	中													
	差													
	劣													
同时双刺激连续质量评价 (SDSCE) 法	——	——	——	<table style="border: none; margin: auto;"> <tr><td style="border: none;"> </td><td style="border: none;">保真度最佳 (打分100)</td></tr> <tr><td style="border: none;"> </td><td style="border: none;">保真度最差 (打分0)</td></tr> </table>		保真度最佳 (打分100)		保真度最差 (打分0)						
	保真度最佳 (打分100)													
	保真度最差 (打分0)													

表2 深度质量主观评价方法

评价方法	图像序列时长	二选一标度	离散标度	连续标度
单刺激 (SS) 法	约 10s	——	5 优 4 良 3 中 2 差 1 劣	优 良 中 差 劣
双刺激损伤标度 (DSIS) 法	约 10s	——	5 损伤不可察觉 4 损伤可察觉, 但不令人讨厌 3 损伤令人有些讨厌 2 损伤令人讨厌 1 损伤令人很讨厌	——
双刺激连续质量标度 (DSCQS) 法	约 10s	——	——	优 良 中 差 劣
成对比较 (PC) 法	约 10s	A 或 B	-3 差很多 -2 差 -1 稍差一些 0 相同 1 稍好一些 2 好 3 好很多	——
单刺激连续质量评价 (SSCQE) 法	约 3min~5min	——	——	优 良 中 差 劣
同时双刺激连续质量评价 (SDSCE) 法	——	——	——	保真度最佳 (打分100) 保真度最差 (打分0)

表3 观看舒适度主观评价方法

评价方法	图像序列时长	二选一标度	离散标度	连续标度
单刺激 (SS) 法	约 10s	——	5 非常舒适 4 舒适 3 轻微不适 2 不舒适 1 非常不适	非常舒适 舒适 轻微不适 不舒适 极度不适
双刺激损伤标度 (DSIS) 法	约 10s	——	5 损伤不可察觉 4 损伤可察觉, 但不令人讨厌 3 损伤令人有些讨厌 2 损伤令人讨厌 1 损伤令人很讨厌	——
双刺激连续质量标度 (DSCQS) 法	约 10s	——	——	非常舒适 舒适 轻微不适 不舒适 极度不适
成对比较 (PC) 法	约 10s	A 或 B	-3 差很多 -2 差 -1 稍差一些 0 相同 1 稍好一些 2 好 3 好很多	——
单刺激连续质量评价 (SSCQE) 法	约 3min~5min	——	——	非常舒适 舒适 轻微不适 不舒适 极度不适
同时双刺激连续质量评价 (SDSCE) 法	——	——	——	保真度最佳 (打分100) 保真度最差 (打分0)

5.2 单刺激 (SS) 法

5.2.1 概述

该方法包含一系列的评价单元，每个评价单元只包含一个图像序列，根据相应评分标度对图像序列单独进行打分。在适当的情况下，可将若干个评价单元组成一个评价段，评价段之间安排休息时间。

5.2.2 单刺激法评价单元结构

评价单元中图像序列的前后均应有灰场显示。图像序列前的灰场应为零视差且持续时间小于等于3s，可包含序号等标识。图像序列后的灰场可包含提示语，如“打分”，该灰场应持续足够长的时间（如小于等于10s）以便评价员进行打分。图像序列的时长应为10s左右。单刺激法评价单元结构见图1。

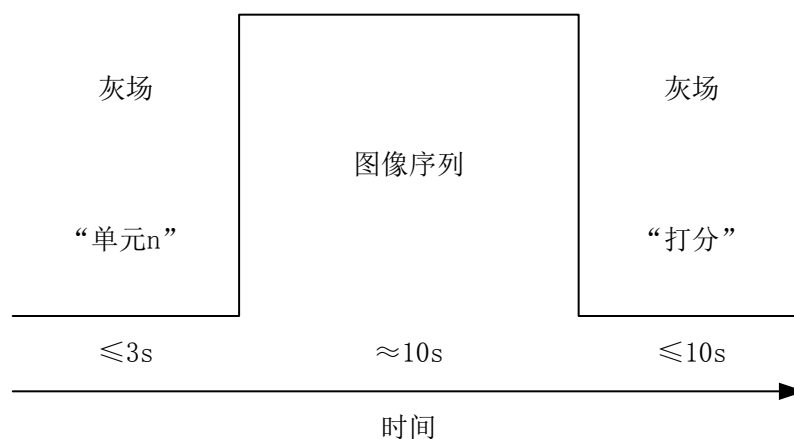


图1 单刺激法评价单元结构

5.2.3 单刺激法评分标度

图像质量和深度质量的评价可用两种评分标度：离散五级标度和标准的ITU连续质量标度，见表1和表2。评分等级划分为“优”、“良”、“中”、“差”、“劣”。

观看舒适度的评价可用两种评分标度：离散五级标度和连续舒适度标度，见表3。评分等级划分为“非常舒适”、“舒适”、“轻微不适”、“不舒适”、“极度不适”。

5.2.4 单刺激法评价分数

对每个图像序列的一次打分为一个“评价分数”。对所有评价分数进行平均，得到MOS值。

评价中可使用参考图像序列，即源图像序列。参考图像序列可用于计算“评价分数下降值”，即参考图像序列和被测图像序列的评价分数的差值。将所有评价分数下降值进行平均，得到DMOS值。

5.3 双刺激损伤标度 (DSIS) 法

5.3.1 概述

双刺激损伤标度法先给评价员随机播放一个源图像序列，再播放相应的一个被测图像序列。评价员对比源图像序列，对被测图像序列进行打分。被测对象的顺序应随机，各被测对象和各图像序列应无遗漏，每个评价段的持续时长不应超过0.5h。双刺激损伤标度法系统框图见图2。

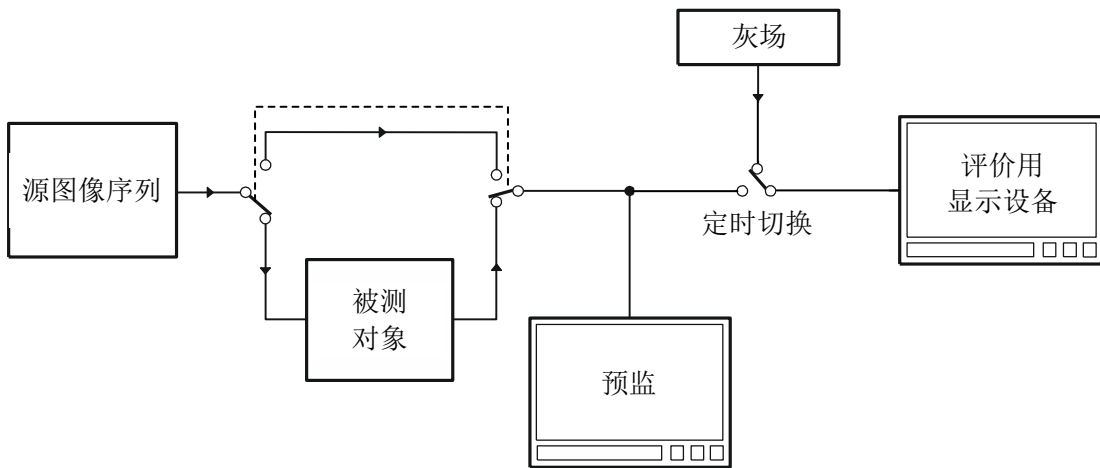
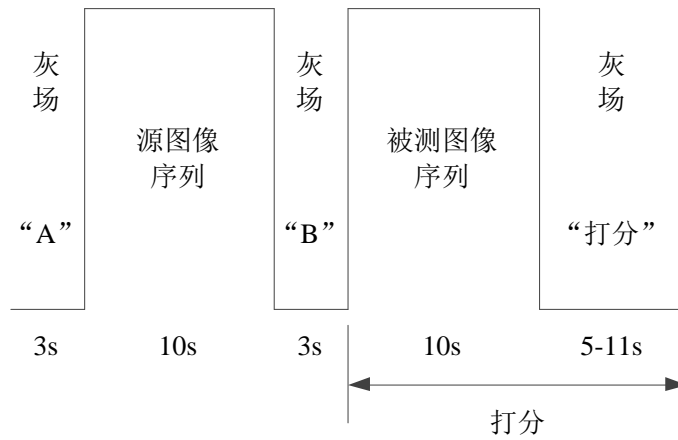


图2 双刺激损伤标度法系统框图

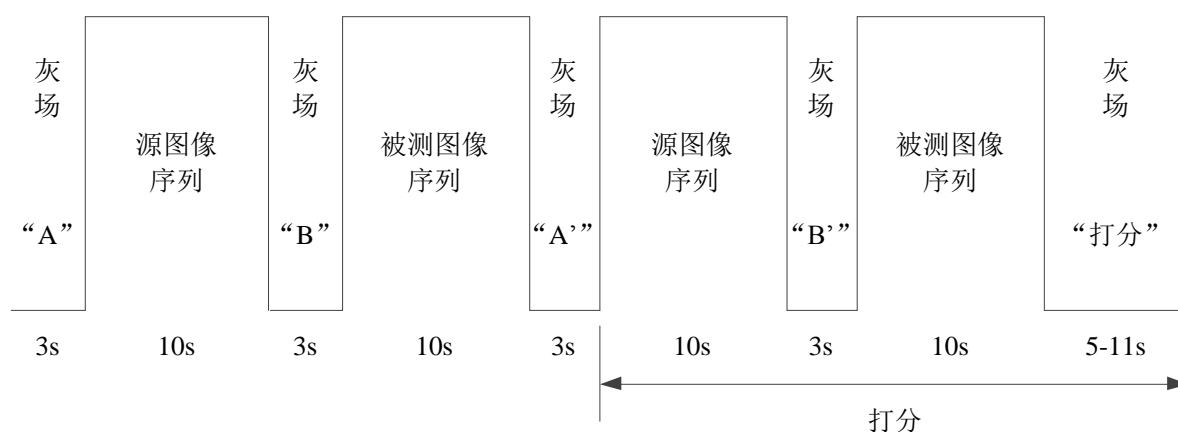
5.3.2 双刺激损伤标度法评价单元结构

评价单元结构有两种：

- a) 方式一：源图像序列和被测图像序列仅各播放一次，见图3中a)；
 - b) 方式二：源图像序列和被测图像序列各播放两次，见图3中b)。
- 当需要分辨很小的损伤或评价动态图像序列时，可采用方式二。



a) 方式一



b) 方式二

图3 双刺激损伤标度法的评价单元结构

5.3.3 双刺激损伤标度法评分标度

应使用五级损伤标度，见表1、表2和表3。评价员使用的打分记录表应有非常清晰的标度，并且表格中应含有带编号的方框（或采用其他形式）以记录打分结果。

5.3.4 双刺激损伤标度法评价分数

评价结束后，给出每个被测对象的各被测图像序列的平均分和总平均分作为评价结果。

本方法通常对损伤较小的评价比对损伤较大的评价在评价结果的稳定性上更好。本方法宜对图像整体损伤进行评价，不宜对图像某方面的损伤进行评价。

5.4 双刺激连续质量标度（DSCQS）法

5.4.1 概述

该方法包含一系列的评价单元，在每个评价单元中，同一视频的源图像序列与被测图像序列均播放2次，共播放4次。双刺激连续质量标度法系统框图见图4。

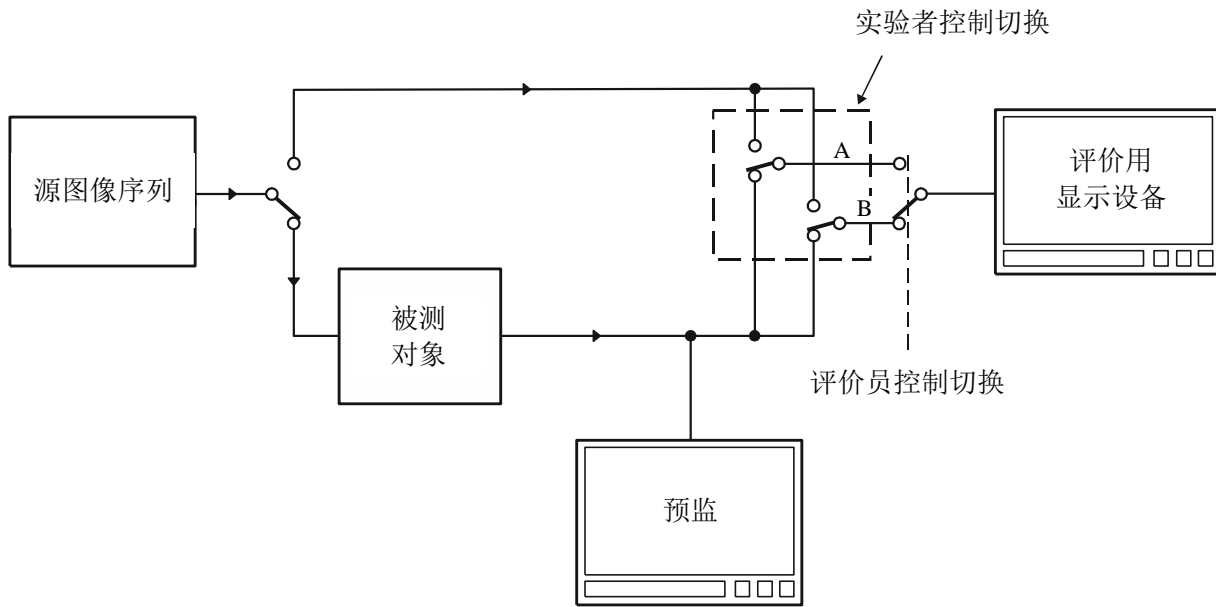


图4 双刺激连续质量标度法系统框图

5.4.2 双刺激连续质量标度法的实现方式

双刺激连续质量标度法的实现方式有两种：

- a) 方式一：每次由一名评价员单独观看，该评价员可自由切换 A、B 信号，直至形成对每个信号质量的打分意见；
- b) 方式二：每次由多名评价员同时观看。对于静态图像序列，应每次播放 3s~4s，A、B 信号重复播放 5 次，在最后 2 次进行打分。对于动态图像序列，应播放 10s，A、B 信号重复播放 2 次，在最后 1 次进行打分，其评价单元结构见图 5。

对于以上两种实现方式，实验者应随机地安排 A 为源图像序列、B 为被测图像序列或 A 为被测图像序列、B 为源图像序列，并不将这种安排告知评价员。

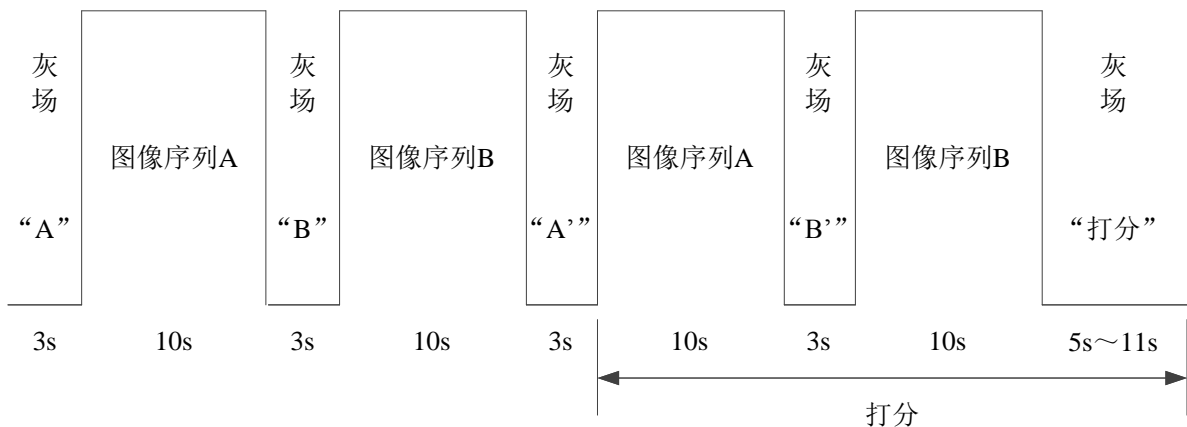


图5 双刺激连续质量标度法的评价单元结构

5.4.3 双刺激连续质量标度法评分标度

评价员应对A、B图像序列分别打分。在评价图像质量和深度质量时，应使用连续质量标度，见表1和表2。在评价观看舒适度时，应使用连续舒适度标度，见表3。

5.4.4 双刺激连续质量标度法评价分数

使用源图像序列与被测图像序列的主观评价得分，计算得到被测对象的DMOS值。

5.5 成对比较（PC）法

5.5.1 概述

该方法将被测对象的图像序列进行两两的配对显示，评价员需判断两个中哪一个更好。当被测对象数目为 n 时，应将被测对象（如X、Y、Z等）组合成所有可能的 $n(n-1)$ 种序列对（如XY、YZ、ZX等）进行播放，且所有的序列对均应以不同顺序分别播放一次（例如XY、YX）。

5.5.2 成对比较法评价单元结构

评价单元的开始处应有小于等于3s的灰场显示，可叠加标识文字，如评价单元的序号。该灰场应为零视差。然后播放要比较的图像序列对，每个图像序列应持续10s左右。应在评价单元结尾处的灰场上叠加标识文字，如“打分”。该灰场应持续足够长时间（比如小于等于10s）以便评价员打分。被测图像序列可在两个显示设备上同时播放（或者在同一显示设备上并排播放），也可在同一个显示设备上先后播放。当在同一显示设备上先后播放时，图像序列对中间应有3s的灰场，其评价单元结构见图6。

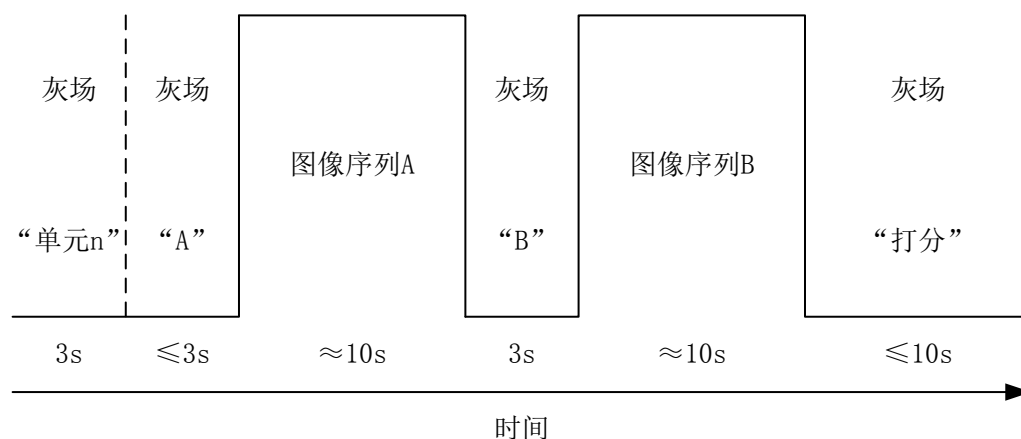


图6 成对比较法评价的评价单元结构

5.5.3 成对比较法评分标度

应让评价员从图像序列对中选出相对较好的（采用二选一标度），或者让评价员对质量差异进行更细致的评价（采用离散标度）。评价图像质量、深度质量和观看舒适度时可采用相同的标度方式，见表1、表2和表3。

5.5.4 成对比较法评价分数

汇总评价员的打分，给出成对比较法的判定结果。

5.6 单刺激连续质量评价（SSCQE）法

5.6.1 概述

在该方法中，评价员对图像质量、深度质量和观看舒适度进行连续不断地（在时间上）评价。

5.6.2 单刺激连续质量评价法的实现方式

应按如下方式组织单刺激连续质量评价法：

- a) 节目片段：是指经过某个被测对象处理过的某种类型（如体育、新闻、戏剧等）的一段节目，节目时长至少 5min。
- b) 评价段：由节目片段组成，节目片段之间没有间歇，且以伪随机顺序排列。每个评价段应含有所有被测对象和各种类型的节目片段，但不必含有所有节目片段。每个评价段持续时间不应超过 60min。
- c) 评价场：评价场包含所有被测对象的所有节目片段，可分成多个评价段完成。

为了保证质量评价效果，在评价中也可播放伴音。应注意，在评价之前应选取不干扰正常评价的伴音。

5.6.3 单刺激连续质量评价法评分标度

对于图像质量和深度质量评价，应使用标准的ITU连续质量标度，见表1和表2；对于观看舒适度评价，应使用连续舒适度标度，见表3。图7是单刺激连续质量评价法输出结果的一个示例。

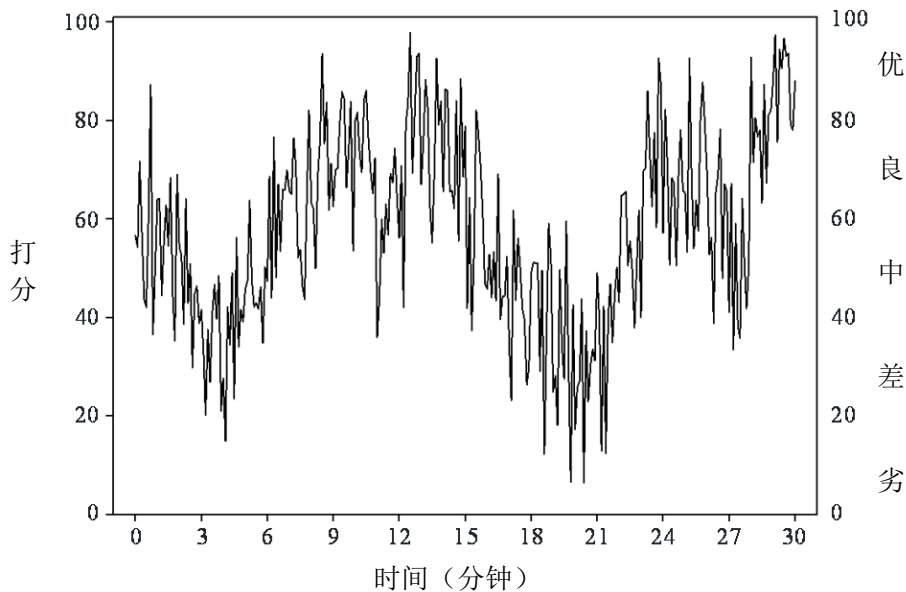


图7 单刺激连续质量评价法输出结果的一个示例

5.6.4 单刺激连续质量评价法评价分数

通过计算在全部评价段中获得的数据，可得到平均质量打分值和时间的函数关系 $q(t)$ 。报告应记录各个节目片段、视频内容、评价段下所有评价员的平均打分结果。

5.7 同时双刺激连续质量评价（SDSCE）法

5.7.1 同时双刺激连续质量评价法图像序列的实现方式

评价员同时观看两个图像序列：左边是源图像序列，右边是被测图像序列。可使用两个并排摆放的监视器播放两个图像序列。

5.7.2 同时双刺激连续质量评价法评分标度

评价员比较两个图像序列，并通过手持打分装置的推杆为被测图像序列的保真度进行打分。当保真度最佳时，应将推杆推到顶端（即打100分），当保真度最差时，应将推杆拉到底端（即打0分）。应使用标准ITU连续质量标度，见表1、表2和表3。

应让评价员提前知道哪个是源图像序列，并要求评价员在观看过程中连续不断地对被测图像序列的保真度进行打分。

5.7.3 同时双刺激连续质量评价法评价分数

下列定义适用于对评价过程的描述：

视频片段：即图像序列。

被测条件：即被测对象。应将参考图像序列添加到被测条件列表中。

评价段：是指一系列某被测条件下的视频片段，视频片段之间没有间歇，且以伪随机顺序排列。每个评价段应含有各个被测条件和所有视频片段，但不要求显示所有被测条件下的所有视频片段。

评价场：评价场是一系列的评价段，包含了所有被测条件下的所有视频片段。所有被测条件下的所有视频片段必须被相同数量的评价员打分（但不必是相同的评价员）。

打分段：每个评价员应在一个评价段中连续不停地打分。

打分片段：10s打分的片段，每个打分片段中应有20个连续的、不重叠的、分布均匀的打分。

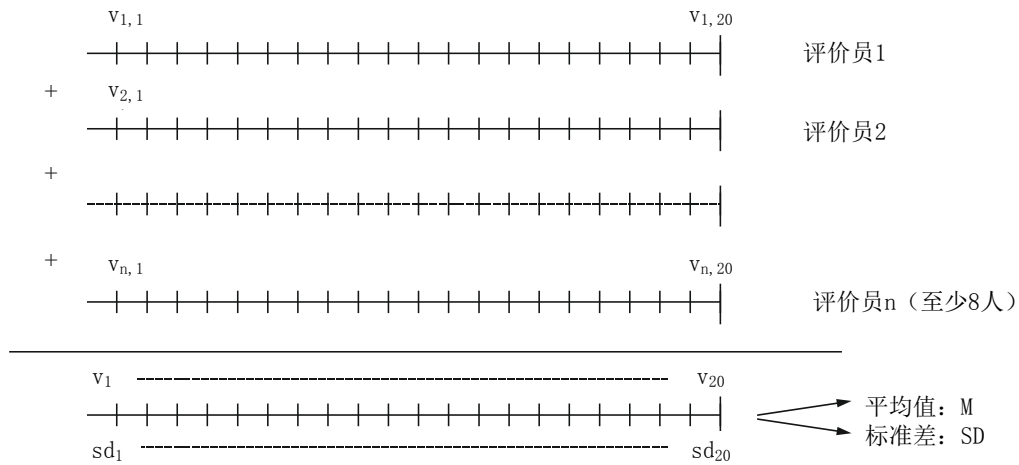
打分后，通过检查每个被测条件下的视频片段是否有相同数量的打分，可检查评价结果的有效性。

可使用以下三种不同的方式处理评价中得到的打分数据：

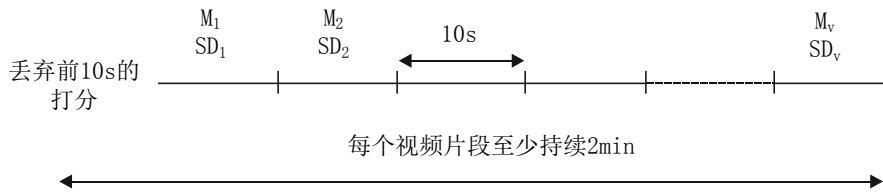
- a) 每个视频片段分别统计分析；
- b) 每个被测条件分别统计分析；
- c) 所有被测条件下的所有视频片段整体统计分析。

每种方式可采用四步分析：

- a) 针对打分片段的每个打分点，累计所有评价员的打分，计算其均值和标准差。
- b) 计算每个打分片段的均值和标准差，见图9。这一步的计算结果可表示为一个时序图，见图10。
- c) 给出每个打分片段的均值的统计分布，分析各均值出现的频率。为了避免之前视频片段带来的近因效应，将每个视频片段中前10s的打分丢弃。
- d) 通过累计各均值的频率计算整体损伤特性。计算中应考虑置信区间，见图11。通过每个打分片段的均值和它们的累加频率之间的函数关系，得到整体损伤特性。



a) 计算视频片段的各打分片段的每个打分点所有评价员的均值 v 和标准差 sd



b) 计算视频片段的各打分片段的均值 M 和标准差 SD

图8 双刺激连续质量评价法的数据处理

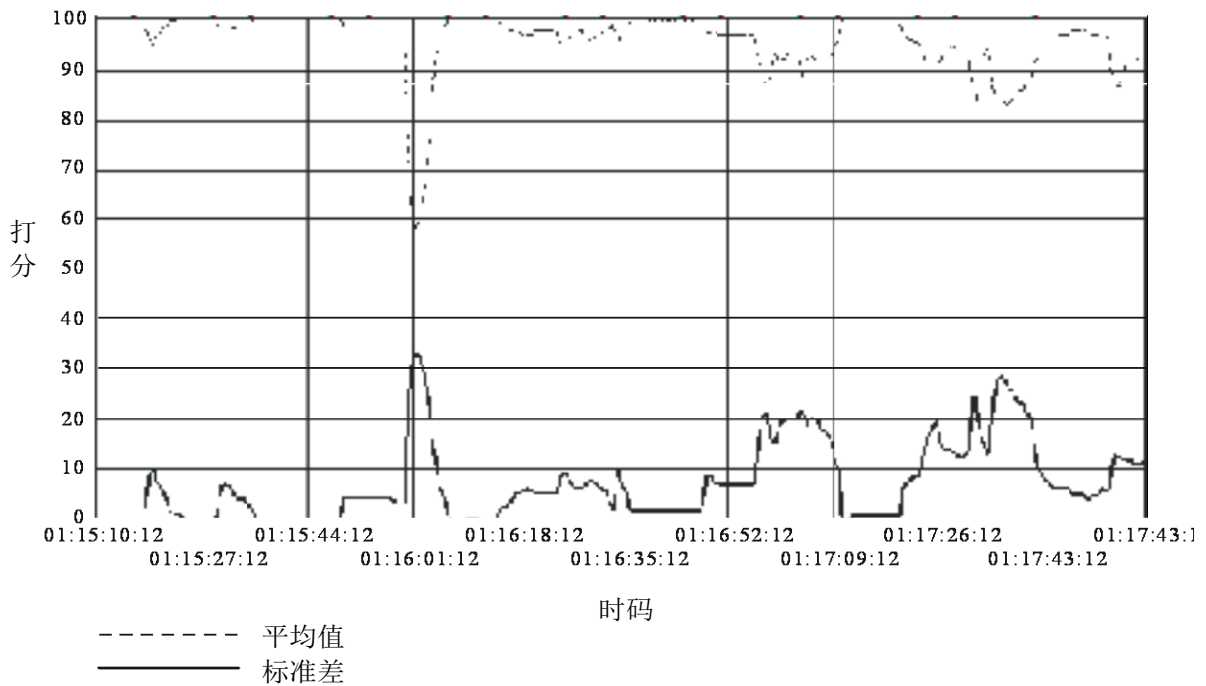


图9 双刺激连续质量评价法的原始数据时序图

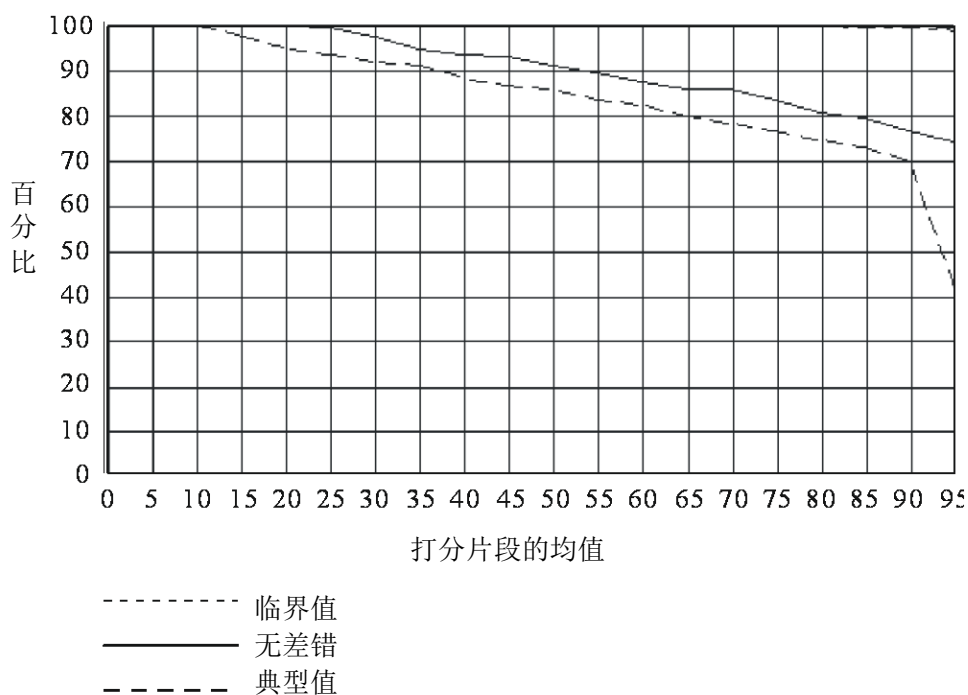


图10 双刺激连续质量评价法的整体损伤特性和置信区间

6 观看条件

屏幕亮度、对比度、背景照明度、观看距离等观看条件应与ITU-R BT. 2022建议书中高清电视平板显示设备的主观质量评价观看条件保持一致。

评价员宜在最佳观看距离处（立体电视相邻两个像素在观看点处形成的夹角恰好为 $1'$ ）进行观看和评价。

对于图像分辨率为 1920×1080 的家用高清显示设备，最佳观看距离为 $3.1H$ （ H 为显示设备画面高度）。表4为一些典型屏幕尺寸的最佳观看距离。

表4 一些典型屏幕尺寸的最佳观看距离

屏幕尺寸 毫米（英寸）	最佳观看距离 米
812.8（32）	1.24
1066.8（42）	1.62
1320.8（52）	2.01
1574.8（62）	2.39
1828.8（72）	2.78
2082.8（82）	3.17
2336.8（92）	3.55
2590.8（102）	3.94

7 图像序列的基本要求

7.1 概述

应根据测试目的，有针对性地选择图像序列。应根据被测对象所应用的节目场景，选择具有相应内容（如体育、戏剧、电影等）和时空特性的图像序列。

所选图像序列应具有正常的观看舒适度。因此，除了特殊目的之外，所选图像序列的视差范围不应超出7.2所规定的限值。另外，应测量并记录图像序列视差分布的平均值、标准差和最大最小值。

当选择适宜观看的立体图像序列时，应考虑左右眼图像之间的视差、视差分布和视差变化。

7.2 视觉舒适限值

以屏幕水平尺寸的百分比表示，负视差的限值宜为1%，正视差的限值宜为2%（总和为3%）。

7.3 左右眼图像的误差

在选取图像序列时，应尽量避免左眼图像与右眼图像间存在如下误差：

- a) 几何误差，包括尺寸误差、垂直误差、旋转误差；
- b) 亮度误差，包括白电平误差、黑电平误差；
- c) 串扰。

7.4 视差范围、分布及变化

宜按如下方式设计图像序列：

- a) 在画面顶部的物体离观看者较远、在画面底部的物体离观看者较近；
- b) 视差角范围在 $\pm 1^\circ$ 之内；
- c) 字幕范围在所遮挡画面物体前 $1/6^\circ \sim 1/4^\circ$ （视差角）；
- d) 场景切换时的视差角改变不超过 1° 。

8 评价用设备

评价用设备（视频服务器、显示设备等）应能够播放全分辨率高清的图像序列。

应在报告中提供关于显示设备的尽可能详细的参数信息。

9 评价员

9.1 评价员数量

评价员数量至少为30人。

9.2 视觉检查

应对评价员的视敏度、色盲、立体视三方面进行检查，比如用视力表衡量视敏度、用色盲检测图或其他方式衡量色觉、用随机点立体图或其他方式衡量立体视，并将检查结果记录在报告中。

10 评价员须知

应根据评价要素（如深度质量、观看舒适度等）给出相应的评价员须知。应细致全面地将评价目的以及评价产生的任何可能的负面影响都告知给评价员。

11 评价时长

如果图像序列的视觉舒适限值符合7.2的规定，则评价段时长可为20min~40min。如果图像序列包含过大视差，有可能给评价员带来不适，则评价时长应有所限制。

12 参考图像序列

在评价中可使用参考图像序列。参考图像序列可以是立体参考图像序列，也可以是立体参考图像序列的左眼或右眼2D图像序列。

13 评分差异

允许评价员之间存在评分差异，可通过增加评价员数量以增强评分结果的稳定性。

为了使评价员在评价过程中的打分尺度稳定，应提供足够的训练过程。

为了使评价结果稳定，可采用多重随机播放（即对不同的评价员以不同的随机顺序播放测试序列）以及重复打分。

14 评价员的筛选

打分结束后，对评价员的筛选方法见ITU-R BT. 500建议书附件2。

15 统计分析

结果数据的统计分析与ITU-R BT. 500相同。

附 录 A
(资料性附录)

本标准与 ITU-R BT. 2021-1 相比的结构变化情况

本标准与 ITU-R BT. 2021-1 相比，章条编号发生了变化，具体对照见表 A. 1。

表A.1 本标准与 ITU-R BT. 2021-1 的章条编号对照情况

本标准章条编号	对应 ITU 标准章条编号
1	---
2	---
3	---
4	附件 1 中 1
5	附件 1 中 2
5.1	---
5.2	附件 1 中 2.1
5.3	附件 1 中 2.2
5.4	附件 1 中 2.3
5.5	附件 1 中 2.4
5.6	附件 1 中 2.5
5.7	附件 1 中 2.6
6	附件 1 中 3
7	附件 1 中 4
8	附件 1 中 5
9	附件 1 中 6
10	附件 1 中 7
11	附件 1 中 8
12	附件 1 中 9
13	附件 1 中 10
14	附件 1 中 11
15	附件 1 中 12
附录 A	---

中 华 人 民 共 和 国
广 播 电 影 电 视 行 业 标 准
立 体 电 视 图 像 质 量 主 观 评 价 方 法
GY/T 314—2017

*

国家新闻出版广电总局广播电视规划院出版发行

责任编辑：王佳梅

查询网址：www.abp2003.cn

北京复兴门外大街二号

联系电话：(010) 86093424 86092923

邮政编码：100866

版权专有 不得翻印