

ICS 33.160
CCS M74

团 体 标 准

T/CVIA 93-2022

代替 T/CVIA 93-2021

彩电产品质量评价指标体系及测试方法

(数字电视音频质量主观评价规范)

Evaluation Index System and test method of color TV product quality

Specification for audio subjective assessment of digital television

CVIA

2022 - 4 - 21 发布

2022 - 4 - 21 实施

中国电子视像行业协会 发布

目 次

前言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 主观评价的一般要求	2
4.1 听音室	2
4.2 测试基本条件	2
4.3 信号源	2
4.4 测试素材选择	2
4.5 评价人员	2
4.6 评分方法	3
5 测试与评价	5
5.1 平衡度	5
5.2 力度	5
5.3 清晰度	5
5.4 明亮度	5
5.5 丰满度	5
5.6 柔和度	5
5.7 真实感	6
5.8 立体感	6
5.9 空间感	6
6 评价过程以及数据处理	6
6.1 评价的一般流程	6
6.2 数据处理	7
6.3 结果的表达	8
附录 A（资料性）听音室设计的一般原则	9
附录 B（规范性）评价结果的筛选	10
附录 C（规范性）评价项目权重及分值表	11

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这些专利的责任。

本文件由中国电子视像行业协会提出并归口。

本文件起草单位：海信视像科技股份有限公司、TCL 实业控股股份有限公司、深圳创维-RGB 电子有限公司、四川长虹电器股份有限公司、深圳康佳电子科技有限公司、青岛海尔多媒体有限公司、北京小米电子产品有限公司。

本文件主要起草人：王伟、孙婧、郝亚斌、冯晓曦、彭健锋、魏于凡、韩秋峰、朱生林、王麟、李鹏、王艳、田秀明、王富裕、王玉乾、张利利。

本文件为首次发布。



CVIA

彩电产品质量评价指标体系及测试方法

（数字电视音频质量主观评价规范）

1 范围

本文件规定了数字电视声音质量的主观评价条件和主观评价方法。

本文件适用于各种接收方式（如卫星、地面和有线等）和显示方式（液晶、OLED和投影等）的数字电视终端，其他类似用途的设备可参照执行。

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 12060 声系统设备_第13部分：扬声器听音试验

GB/T 22123 数字电视接收设备图像和声音主观评价方法

SJ/T 11324 数字电视接收设备术语

IEC 61672-1 电声学 声级计 第1部分：规范（Electroacoustics - Sound level meters - Part 1: Specifications）

IEC 61672-2 电声学 声级计 第2部分：型式评定试验（Electroacoustics - Sound level meters - Part 1: Pattern evaluation tests）

ITU-R BT.2245-8 HDTV and UHDTV including HDR-TV test materials for assessment of picture quality

3 术语和定义

SJ/T 11324 界定的以及下列术语和定义适用于本文件。

3.1

主观评价 subjective assessment

通过评价人员对被测系统质量的主观反应，来确定电视系统性能的一种方法。

3.2

测试素材 test fodder

用于评价电视声音质量，并且对声音内容有特定要求的音源或音源序列。

3.3

评级标度 grading scale

主观评价中用于评定声音质量的评定尺度。

4 主观评价的一般要求

4.1 听音室

主观评价实验室环境应类似或优于一般家庭试听环境。良好的听音室设计的一般原则参考附录A。如房间大小可自由选择，标准听音室可以参考如下尺寸：

$$L = 7.0 \text{ m};$$

$$W = 5.3 \text{ m};$$

$$H = 2.7 \text{ m}.$$

4.2 测试基本条件

听音室环境以及观看条件的确定参考GB/T 22123和GB/T 12060中的规定，应满足表1规定的要求。

表 1 观看条件

项目	参考值
环境条件	环境温度：18° ~27° 相对湿度：25%~75% 大气压强：86 kPa~106 kPa
电源电压	220(1±10%) V, 频率振动 ≤±2 %， 谐波分量≤±5 %
背景噪声声级（A计权）	≤35dB
混响时间	在250 Hz~4 kHz之间：0.3s~0.6s，最好0.4±0.05 s
听音距离	对于1920*1080格式：3.2倍图像高度 对于3840*2160格式：1.6倍图像高度 对于7680*4320格式：0.8倍图像高度
评价人员的座位安排	2 K：水平方向在中垂线±15度以内 4 K：水平方向在中垂线±30度以内 8 K：水平方向在中垂线±45度以内

4.3 信号源

信号源产生的测试信号输出接口推荐采用数字音频接口，如采用其他接口测试，应记录所采用的接口类型。

4.4 测试素材选择

主观评价测试素材参考ITU-R BT.2245-8选取，并应满足以下要求。

- 测试音源的音色、形式、细节等不应使评价人员分心；
- 测试音源序列应包括人语声、声乐、乐器和效果声四大类节目素材；
- 测试音源序列应包含新闻类、音乐类、电影类、综艺类、动画类、游戏类等节目类型的节目素材；

4.5 评价人员

评价人员又称听音员，听音员应具有正常的听力，应熟练的掌握评价方法，具有独立的分析判断能力。

根据不同的评价目的评价小组可分为专家评价小组和非专家评价小组。

4.5.1 专家评价小组

专家是指那些受过听音训练，具有高保真听音经验的专业人士，专家对于测试系统引入的声音具有专长，最好对于音乐具有专业知识。为得到较高置信度的评价结果，专家评价小组建议由3至5名专家人员组成。

4.5.2 非专家评价小组

非专家对于测试系统引入的声音不具有专长，所需评价人员的数目取决于测试程序的感受度和信度，建议应至少由15位评价人员组成。

4.6 评分方法

本标准采用刺激比较法和单刺激法两种方法作为主观评价方法，在对于有对比样机的主观评价中，使用刺激比较法；对于无对比的情况，使用单刺激法。

4.6.1 刺激比较法

4.6.1.1 系统框图

测试系统的系统框图如图1所示。

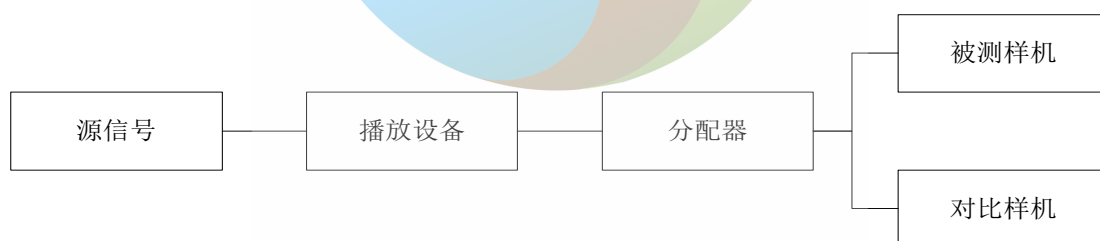


图 1 测试系统框图——刺激比较法

4.6.1.2 评分标度

刺激比较法使用连续量表法判断，评价人员在垂直标尺上插入一个标记，对每次演示的总体声音质量进行评价。为了避免量化误差，标度提供了连续的等级系统，并将垂直标尺分成了相等的七段长度，与ITU-R比较量表相对应，见图2和表3。

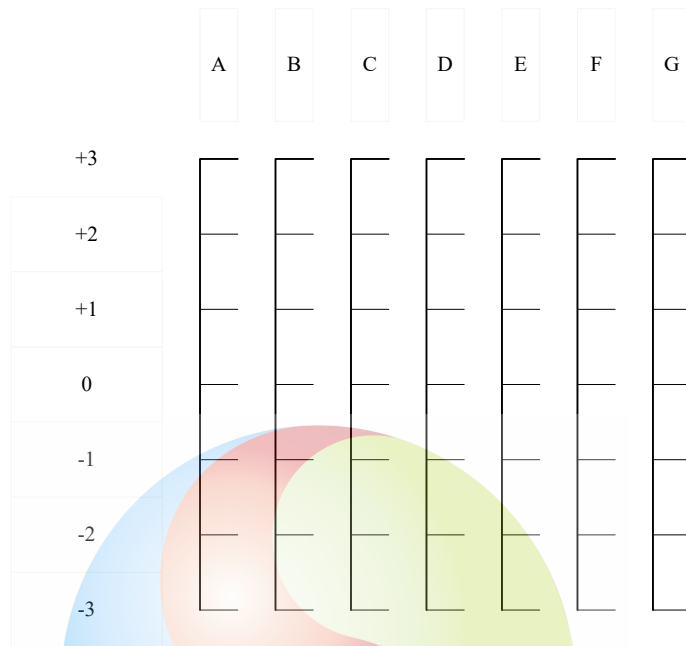


图 2 刺激比较连续质量标度

表 3 评级标度描述

等级	术语描述
-3	差得多
-2	差
-1	稍差
0	相同
1	稍好
2	好
3	好得多

4.6.2 单刺激法

4.6.2.1 系统框图

测试系统的系统框图如图3所示。

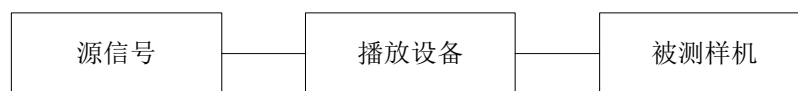


图 3 测试系统框图——单刺激法

4.6.2.2 评分标度

单刺激法使用数值量表法判断，评价人员为每一个音源或音源序列打分，分值范围为0-100。分值可与正常的ITU-R五级质量标度相对应，见表4。

表 4 分值与五级质量量表对应关系

分值	等级	质量术语描述
100	5	优
80	4	良
60	3	中
40	2	差
20	1	劣

5 测试与评价

5.1 平衡度

选择高、中、低频段均有的音源序列进行评价，根据声音各频率段的比例是否协调，高、中、低音搭配是否合理，频率特性是否好，是否存在某一频段过于提升或衰减的情况进行综合评价。

5.2 力度

选择低音频率成分且带有动态冲击的音源序列进行评价，根据音源序列中声音是否有劲、中低频成分动态范围是否较宽进行综合评价。音源序列可包含鼓声、手雷爆炸声等场景。声音表现应强劲有力，收放自如；不应有明显的乏力感。

5.3 清晰度

根据音源序列中各频率的信号是否都能很好的表现出来进行综合评价。音源序列可包含人声、多乐器演奏等场景。整体声音应易懂度高，有清澈见底之感；不应存在声音含糊不清、单色混成一片，低音过多的情况。

5.4 明亮度

选择高、中音丰富的音源进行评价，根据音源序列中、高音能量是否充分，混响比例是否得当、是否出现啸叫等进行综合评价。音源序列可选取音乐会的场景。声音应有明快、活跃感；不应有低音松弛，高音缺乏的现象。

5.5 丰满度

选择中、低音丰富的音源进行评价，根据音源中的中低频是否丰富，高频是否适度，混响是否充足、响度是否合宜等情况进行综合评价。音源序列可包含人声、小号、吉他、萨克斯等场景。声音听起来应圆润、甜美；不应有单薄、干瘪之感。

5.6 柔和度

根据声音是否松弛不紧，高音比例是否适度，混响时间是否得当等情况进行综合的评价。音源序列声音混响时间应稍长，失真小，高频部分应有延伸但不刺耳，整体听音感觉悦耳舒服。

5.7 真实感

根据声音还原过程中是否保持声音原来的所有特点、是否出现失真或染色的情况进行综合评价。声音整体应逼真；不应因增加过多的修饰使声音的特点变形。

5.8 立体感

选择景深层次富有变化的音源序列进行评价，根据声像定位是否准确，声像群分布是否正确，是否有宽度和纵深感进行综合评价。音源序列可选取音乐会片段或演唱会片段，首先声音方位要清楚，有层次，宽度感好；其次重放器乐或人声应真实。

5.9 空间感

选择能展现宽广声场气氛的音源序列进行评价，根据音源序列是否能很好展示原来声源位置和现有播放位置的空间真实感进行综合评价。音源序列可选取交响乐、大合唱等，评价人员应产生强烈的临场感，距离不失真，且不应有明显的压抑感和侵略性。

6 评价过程以及数据处理

6.1 评价的一般流程

6.1.1 标准工作状态调整

数字电视进行测试时请将电视恢复出厂默认状态，对声音设置状态做如下规定：

- a) 将数字电视的声音设置恢复到出厂设置；如无出厂设置，将声音模式调整到“标准”或与之相对应的模式，其它菜单设置为开机后的设置；
- b) 数字电视如有声音输出选项，将电视机的选项设置到内置扬声器的状态。
- c) 数字电视的音量设置：信号使用连续的粉红噪声，调整数字电视音量的大小，直到使放在参考试听点的声级计读数为 75 dBA（50 寸及以上）或 70 dBA（50 寸以下）来获得，且需读数相差小于 0.5 dB。声级计符合 IEC 61672-1 和 IEC 61672-2 的标准。

6.1.2 稳定时间

为了确保数字电视被测终端状态的稳定性，应在标准工作状态下工作30 min以上，方可进行测试。

6.1.3 评价过程

6.1.3.1 评价训练

在每一评价训练开始时，应向评价人员详细介绍评价方法、容易产生的质量损伤类型、声音音源播放顺序和时间、评分制、评分等级及使用设备等。介绍人员语气中不应带有影响评定的暗示。

评价训练阶段应持续0.5 h以内。第一段开始时，应播放5个左右“模拟演示”，以稳定评价人员的评分。这几个序列中给出的数据不能在测试结果中考虑。如果需要多个测试阶段，在后续阶段开始时仅需要3个左右的“模拟演示”。

训练序列播放顺序采用随机顺序，训练序列不应使用正式的测试序列，但应与正式的测试序列具有可比性。

6.1.3.2 正式评价

- a) 测试序列应按照评价表中的维度顺序播放，测试阶段每一个测试序列演示 3 次。第一次演示用于稳定观察者的意见，此次演示不评分；对于测试序列的评分是从第二个和第三个演示中得到的。序列演示以及评分顺序如图 4 所示：

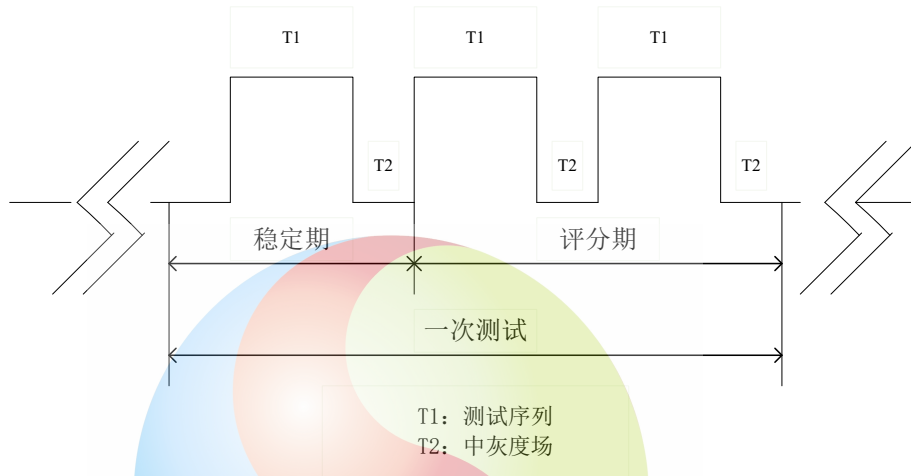


图 4 序列演示和评分顺序

- b) 评价人员应按有关评分等级的说明对被测音源或音源序列独立地进行评价打分，试验过程中评价人员应集中注意力，防止环境干扰对评价人员的听觉产生影响。

6.2 数据处理

刺激比较法需要先将结果归一化为0和100之间的整数。

评价人员评价结果的筛选见附录B。

刺激比较法以及单刺激法都需要计算出每个被测序列的平均分 \bar{u}_{jkr} ，公式如下：

$$\bar{u}_{jkr} = \frac{1}{N} \sum_{i=1}^N u_{ijkr} \dots \dots \dots (1)$$

式中：

u_{ijkr} ——评价人员*i*在测试条件*j*、序列/音源*k*、重复*r*次情况下的评分；

N ——评价人员的数目。

每一个被测序列的标准偏差 S_{jkr} 由下式给出：

$$S_{jkr} = \sqrt{\sum_{i=1}^N \frac{(\bar{u}_{jkr} - u_{ijkr})^2}{(N-1)} \dots \dots \dots (2)}$$

δ_{jkr} 代表置信区间的半径微变量，由下式给出：

$$\delta_{jkr} = 1.96 \times \frac{S_{jkr}}{\sqrt{N}} \dots \dots \dots (3)$$

式中：

S_{jkr} ——每一个被测序列的标准偏差；

N ——评价人员的数目，单位为个。

95%置信度区间如下：

$$[\bar{u}_{jkr} - \delta_{jkr}, \bar{u}_{jkr} + \delta_{jkr}] \dots \dots \dots (4)$$

6.3 结果的表达

针对每个被测序列的平均分 \bar{u}_{jkr} ，计算测试条件 j 时的平均分 \bar{u}_j ，最后按照附录C进行评分汇总，给出被测数字电视终端终评得分 U 。报告中还应包含以下内容：

- a) 被测数字电视终端型号、版本等信息；
- b) 评价使用的设备以及测试素材的情况；
- c) 评价时数字电视终端陈设的描述，如：使用的是木质电视柜、被测终端距墙 50cm 等
- d) 主观评价采用的方法；
- e) 评价人员人数；
- f) 数字电视终端终评得分，其中单刺激法终评得分为数据处理后的结果；刺激比较法中对比样机按照 50 分终评分参与计算，若对比样机声音质量实得分为 U_{ds} ，则被测样机加权后的终评得分 U_z 可按照如下公式得出：

$$U_z = U \times \frac{U_{ds}}{50} \dots\dots\dots (5)$$

- g) 总声音质量的描述。



附录 A
(资料性)
听音室设计的一般原则

A.1 房间大小

任何大小的房间的低频响应都会受到室内本征音频分布的影响。

为了确保一个适度均衡的低频本征音的分布，要求听音室宜采用以下尺寸比例：

$$\frac{W}{H} \leq \frac{L}{H} \leq \left(4.5 \frac{W}{H} - 4\right)$$

式中：

W——宽度；

H——高度；

L——长度。

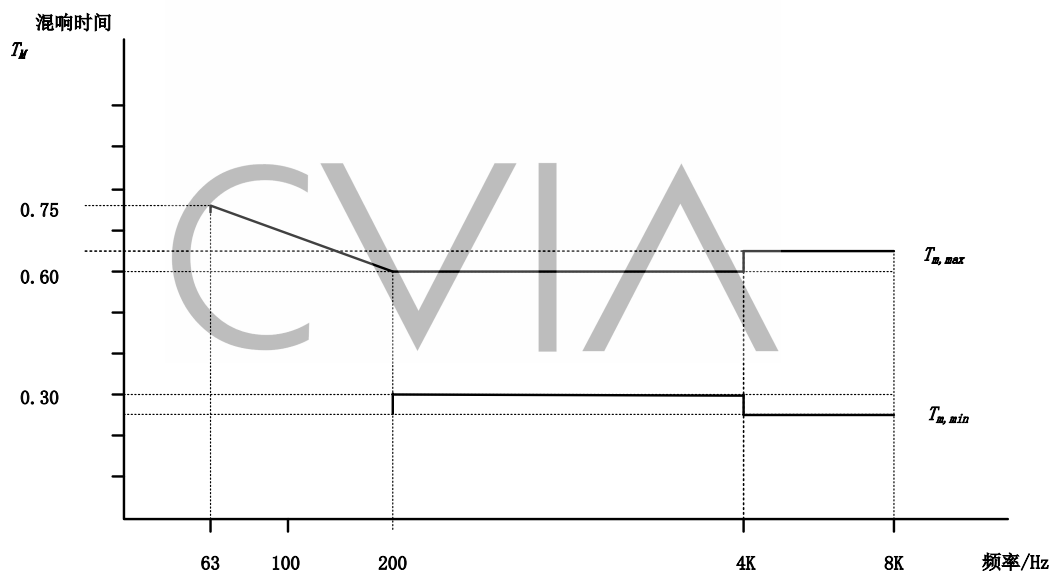
同时满足 $\frac{L}{H} < 3$ 和 $\frac{W}{H} < 3$ 的条件。

在设计试听室时，宜首先规定高度。然后根据设计标准计算符合条件的长度和宽度。单声道或双通道立体声的重放，要求地面面积宜在25m²—40m²，而多通道立体声的重放则要求30m²—45m²。

A.2 混响时间

根据ISO 3382的要求，应该在有家具布置但无试听人员在场的条件下按1/3 oct对混响时间T进行测量。

从200 Hz到4000 Hz测试的平均混响时间T_m应控制在0.3 s—0.6 s之间。当实验结果并不局限于某一特定地区或某一特定目的时，建议T_m控制在图A.1所示的容差范围内。



图A.1 评价混响时间的容差范围

试听室的天花板要求几乎全部反射，而地面则要求几乎全部吸声。为了达到与一般家居环境相近的效果，要求加入普通的家具布置。而为了达到进一步的声吸收，需要附加吸声材料均匀的分布在墙上已达到指定的混响时间，同时抑制颤动回声。

对椅子的要求：使评价人员舒适，并且椅子靠背应低于坐时肩部。

附录 B
(规范性)
评价结果的筛选

首先,利用 β_2 (通过计算函数的峰态系数,即四阶中心距与二阶中心距平方的比值)确定每个测试显示的得分是否是正态分布。对于每个测试显示,计算平均分 \bar{u}_{jkr} 、标准偏差 S_{jkr} 和峰度系数 β_{2jkr} , β_{2jkr} 由下式给出:

$$\beta_{2jkr} = \frac{m_4}{(m_2)^2} \quad m_x = \frac{\sum_{i=1}^N (u_{ijkr} - \bar{u}_{jkr})^x}{N} \dots\dots\dots (B.1)$$

式中:

- m——中间过程量;
- x——取值为2或4;
- N——评价人员的数目;
- j——测试条件的数目,包括基准在内;
- k——测试音源或音源序列的数目;
- r——重复次数。

对于每一位评价人员*i*,找出每一 P_i 和 Q_i ,即:

若 $2 \leq \beta_{2jkr} \leq 4$,这个分布可以取做正态(即视为正常)。则:

$$\text{当 } u_{ijkr} \geq \bar{u}_{jkr} + 2 S_{jkr} \text{ 时 } P_i = P_i + 1; \text{ 当 } u_{ijkr} \leq \bar{u}_{jkr} - 2 S_{jkr} \text{ 时 } Q_i = Q_i + 1$$

若 $\beta_{2jkr} < 2$ 或者 $\beta_{2jkr} > 4$,则:

$$\text{当 } u_{ijkr} \geq \bar{u}_{jkr} + \sqrt{20} S_{jkr} \text{ 时 } P_i = P_i + 1; \text{ 当 } u_{ijkr} \leq \bar{u}_{jkr} - \sqrt{20} S_{jkr} \text{ 时 } Q_i = Q_i + 1$$

最后,如果满足下面的计算结果,则剔除评价人员*i*。

$$\frac{(P_i+Q_i)}{j \cdot k \cdot r} > 0.05, \text{ 且 } \left| \frac{P_i-Q_i}{P_i+Q_i} \right| < 0.3$$

式中:

u_{ijkr} ——评价人员*i*的评分。



附 录 C
(规范性)
评价项目权重及分值表

表C.1规定了各评价项目对应的权重及分值。

表 C.1 评价项目权重及分值表

序号	测试项目	权重系数	加权满分值
1	平衡度	15%	15
2	力度	10%	10
3	清晰度	20%	20
4	明亮度	12%	12
5	丰满度	10%	10
6	柔和度	8%	8
7	真实感	8%	8
8	立体感	7%	7
9	空间感	10%	10
	总评分	100 %	100