

目 次

前言

1 范围	(1)
2 引用标准	(1)
3 术语	(1)
4 主观评价的一般要求	(1)
5 评分方法和评分标度	(4)
附录 A(标准的附录) 数字电视图像质量主观评价数据统计方法	(7)
数字电视图像质量主观评价方法 编制说明.....	(8)

前 言

本标准等效采用国际电信联盟 ITU-R BT. 500-7(1997 年)和 Rec. 710-1(1992 年)建议书。

使用本标准中规定的数字电视图像质量主观评价方法,与国际推荐的方法一致,便于进行国际间的技术交流,促进广播电视技术的发展。

本标准依据 ITU-R BT. 500-7(1997 年)和 Rec. 710-1(1992 年)建议书制订。由于双刺激连续质量标度法具有灵敏度高、稳定性好的特点,适合用于高质量电视系统图像质量的主观评价,因此,本标准规定其作为数字电视图像质量的主观评价方法。具体内容包括 ITU-R BT. 500-7(1997 年)附录 1 中 1, 2.1.1.1, 2.2, 2.3, 2.5, 2.6, 2.7, 2.8, 4.1, 4.3, 4.4, 5.1, 5.2, 5.3, 5.4和 5.5;附录 2 中 1, 2, 2.1, 2.2;Rec. 710-1(1992 年)中的附录 1。只对观看员评分的一致性检验参数作了修改。

本标准的附录 A 是标准的附录。

本标准由国家广播电影电视总局提出。

本标准由国家广播电影电视总局标准化规划研究所归口。

本标准起草单位:国家广播电影电视总局广播科学研究院电视研究所。

本标准主要起草人:侯 洁、李若霜。

数字电视图像质量主观评价方法

GY/T 134—1998

The method of the subjective assessment of
the quality of digital television picture

1 范围

本标准规定了数字电视图像质量的主观评价方法。
本标准适用于对数字电视系统的性能进行主观评价。

2 引用标准

下列标准所包含的条文,通过在本标准中引用而构成为本标准的条文。本标准出版时,所示版本均为有效。所有标准都会被修订,使用本标准的各方应探讨使用下列标准最新版本的可能性。

GB/T 7401—1987 彩色电视图像质量主观评价方法

3 术语

本标准采用下列定义。

3.1 主观评价 subjective assessment

直接利用观察者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 质量评价 quality assessment

在最佳条件(即理想收看条件)下,确定电视系统性能的主观评价。

3.3 损伤评价 impairment assessment

在非最佳条件(如发送和传输等条件)下,确定系统保持质量性能的主观评价。

3.4 测试图像 test picture

用于电视图像质量主观评价的、在图像素材上有特定要求的图像或图像序列。

3.5 评分标度 scale

在主观评价中使用的评分尺度。

3.6 标准清晰度电视 standard definition television(SDTV)

其等效质量是由源于 ITU—R 601 建议书中 4:2:2 等级的图像,经过码率压缩编解码处理后,而其主观评价质量优于 PAL—D 常规清晰度电视的数字电视。

3.7 高清晰度电视 high definition television(HDTV)

高清晰度电视在水平和垂直两个方向上,分辨力约为标准清晰度电视的两倍,并具有 16:9 的幅型比。

4 主观评价的一般要求

4.1 观看条件

数字电视图像质量的主观评价应该在表 1 中给出的观看条件下进行。

表 1 数字电视图像质量主观评价观看条件

项 目	SDTV 参数值	HDTV 参数值
(1)观看距离	4 倍到 6 倍图像高度	3 倍图像高度
(2)显示屏幕的峰值亮度	70 cd/ m ²	150~250 cd/m ²
(3)束流截止时,屏幕亮度与峰值亮度之比	≤0.02	同左
(4)暗室中,黑电平亮度与峰值亮度之比	约 0.01	同左
(5)显示器背景亮度与峰值亮度之比	约 0.15	同左
(6)室内环境光照明	宜低	同左
(7)背景光和照明光光源的色温	D ₆₅	同左
(8)背景光部分对观看员的张角	高≥ 43°,宽≥57°	高 53°,宽 83°
(9)观看员的座位布局	水平方向在中垂线±30°内 (垂直方向待定)	同左
(10)显示图像(对角线)尺寸	至少 50 cm	1.4m

注:

- 1 条件(2),屏幕的峰值亮度对应于 100%的视频信号幅值。在 HDTV 的峰值亮度达不到规定值之前,应满足 ≥70 cd/m²。
- 2 条件(3)受室内照明和显示对比度范围的影响。
- 3 条件(4)黑电平对应于 0%的视频信号幅值。
- 4 为了满足条件(3)和(5),应该设置室内照明。
- 5 若 HDTV 达不到条件(8)规定值之前,应该满足高≥28°,宽≥48°。
- 6 若 HDTV 达不到条件(10)规定的显示尺寸之前,应该满足≥76.2cm。

4.2 评价系统

主观评价系统设备应是高质量设备,不应影响被测系统或设备的质量,系统的原理方框图如图 1 所示。

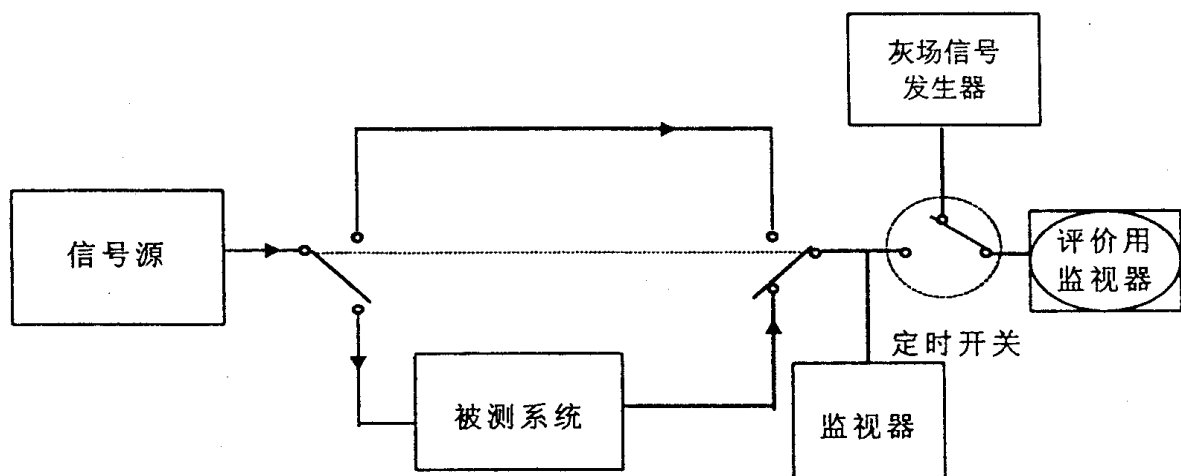


图 1 评价系统方框图

观看员从评价用监视器上看到的是经过定时开关给出的图像。定时开关的输出信号或直接来自信号源,或已经过被测系统,或来自灰场信号发生器。灰场信号发生器的输出调定为 28.6% 的峰值视频电平。

4.3 信号源

信号源直接提供基准数字图像信号,并作为被测系统的输入,经被测系统则成为被测图像。

4.4 测试素材的选择

一套测试图像通常包含一定数目的静止图像和含运动物体的图像序列,每个序列大约持续 10~15 秒钟。

测试素材应该是“严格的,但又不过分”,即包含对各种评价因素,如静止空间分辨率、动态空间分辨率、亮度再现、彩色再现和运动再现等具有比较敏感特点的那些内容。而且,测试素材同时又应包含能代表电视节目的典型内容。

4.5 观看员

观看员即应邀参加主观评价的评分人员。观看员通常有两种类型,即专业观看员和非专业观看员。一般都由非专业观看员来进行主观评价,当需要精确判断时,可由受过专业训练的专业观看员来进行评价和分析。

观看员应该具有代表性,即应包括不同性别、年龄、文化层次的观众;他们应具有正常的视力(含校正视力)和色觉;有一定分析判断能力;对于评价方法和要求应能较快地接受和掌握。

主观评价所需观看员的数目一般不少于 15 人。

4.6 评价测试阶段

在每个评价测试阶段开始时,应该向观看员详细、正确地介绍评价方法和存在的质量损伤类型,并进行评价示范显示。示范显示应该使用正式测试的图像或序列以外的图像或序列,但应与正式测试中使用的图像或序列具有可比性。

一个评价测试阶段包括示范说明在内一般不超过半个小时。在正式测试开始前,需要引入 3 至 5 个测试周期来稳定观看员的判别力,其结果的数据不纳入测试结果的统计中。不同测试图像的显示顺序采用伪随机方式。为了检测相关性,有些测试可以重复进行,但要避免相同测试图像在相继的测试周期出现。

测试阶段的显示流程如图 2 所示

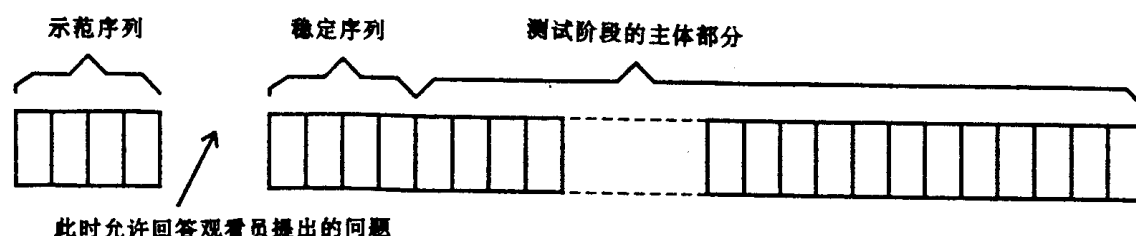


图 2 测试阶段的显示流程图

4.7 结果表达

对于每个测试图像,应该给出评分值统计分布的平均值、标准偏差和 95% 置信度区间。

结果还必须包含以下的信息：

- a) 测试情况的说明
- b) 测试素材的情况
- c) 图像源的类型和显示监视器型号
- d) 观看员的人数和类型
- e) 使用的基准系统
- f) 测试总的平均分
- g) 如果淘汰了某些观看员,还要说明原始的和调整后的平均分和 95% 置信度区间

5 评分方法和评分标度

5.1 双刺激连续质量标度法

5.1.1 评分方法

在双刺激连续质量标度法中,要求观看员只简单地对每个测试图像的两种状态,即基准状态(直接来自信号源)和被测状态(经被测系统)的图像总体质量进行评价,并在评分表上作出标记。

5.1.2 评分标度

评分表由若干对纵向标度线组成,纵向标度线成对地排列,以适应对每个测试图像两种状态的评价。纵向标度线提供连续标度,且被分成 5 个等分,相当于标准的 ITU-R 5 级质量标度范围,并在评分表的左边标有与不同等级相对应的质量术语,即优、良、中、差、劣,可作为观看员评分的一般指导,如图 3 所示。观看员记分笔的颜色应与标度印刷线的颜色有区别。

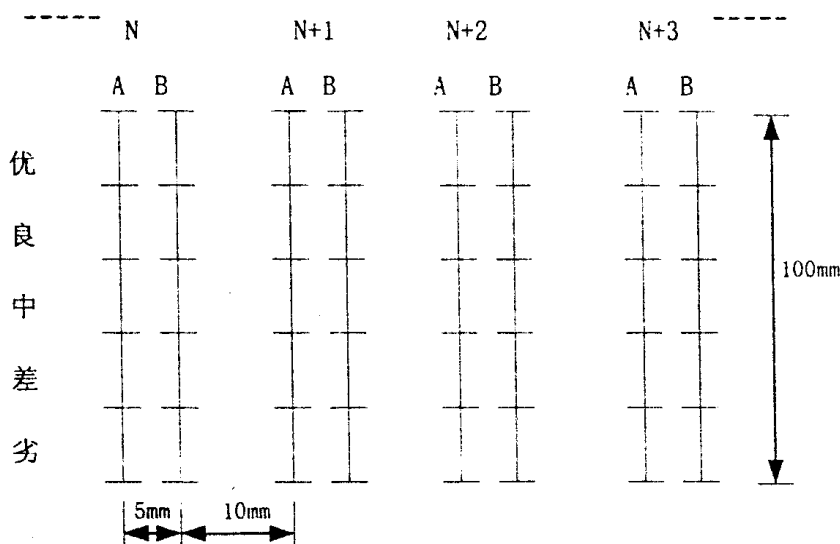


图 3 双刺激连续质量评分表

5.1.3 测试程序

5.1.3.1 示范说明阶段

在每个测试阶段开始时,要向观看员详细、正确地介绍评价方法和评分标度以及图像显示的顺序和时间,但不应出现影响评分结果的暗示。然后进行评分过程的示范,静止图像和图像序列各示范一次。

示范说明应该事先录音,每次测试开始时播放,以免每次口头说明有出入或有暗示。

5.1.3.2 图像显示和评分阶段

一个测试周期是对一个测试图像(或图像序列)进行评分操作的过程,是组成评价测试阶段的基本单元。它包括显示图像和评分两部分。对于静止图像,其基准图像和被测图像交替显示5次,每次各持续3—4秒钟,最后两次为评分期;对于运动图像,其基准图像与被测图像交替显示2次,每次各持续约10秒钟,然后评分。图像序列显示评分的顺序如图4所示。在不同测试图像的一连串显示评分过程中,基准图像与被测图像的先后次序以伪随机方式变动。要求观看员将T1显示图像的评分和T3显示图像的评分分别记录在评分表相应序号的某对标度线A和B上。

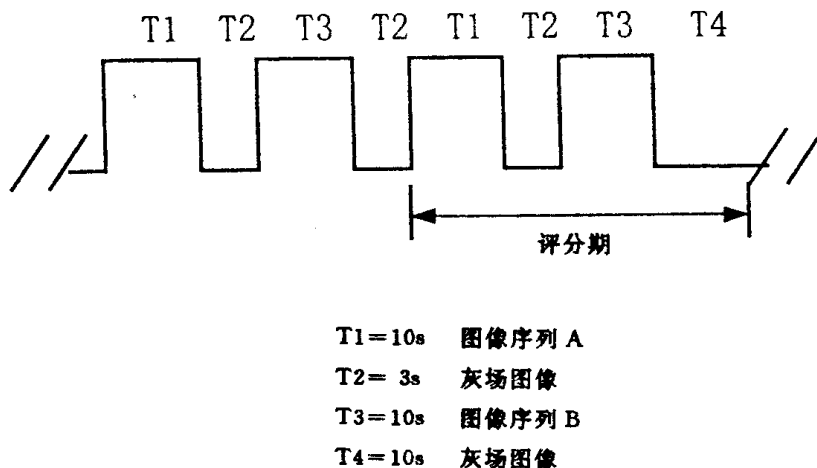


图4 双刺激连续质量标度法图像序列显示和评分的顺序

5.1.4 数据处理

具体的数据统计方法见附录A。

5.2 双刺激损伤标度法

5.2.1 评分方法

在双刺激损伤标度法中,要求观看员根据每个测试图像无损伤的基准图像,对被测图像质量的受损情况作出评价,并在评分表上作出标记。在被测的图像或序列中,也可包括无损伤的图像或序列。

5.2.2 评分标度

双刺激损伤标度法使用5级损伤评分标度。其评分等级和相应的损伤术语在表2中给出,要求观看员根据被测图像质量的受损情况,以规定的评分标度适当的等级评分。

表2 五级损伤标度

评分等级	损伤术语
5	不觉察
4	可觉察、但不讨厌
3	稍有讨厌
2	讨厌
1	非常讨厌
注:“讨厌”不是指对图像内容的讨厌	

5.2.3 测试程序

5.2.3.1 示范说明阶段

同 5.1.3.1。

5.2.3.2 图像显示和评分阶段

图像显示和评分的顺序如图 5 所示,在每个测试周期,均为基准图像显示在先,被测图像显示在后,基准图像和被测图像交替显示两次之后评分。

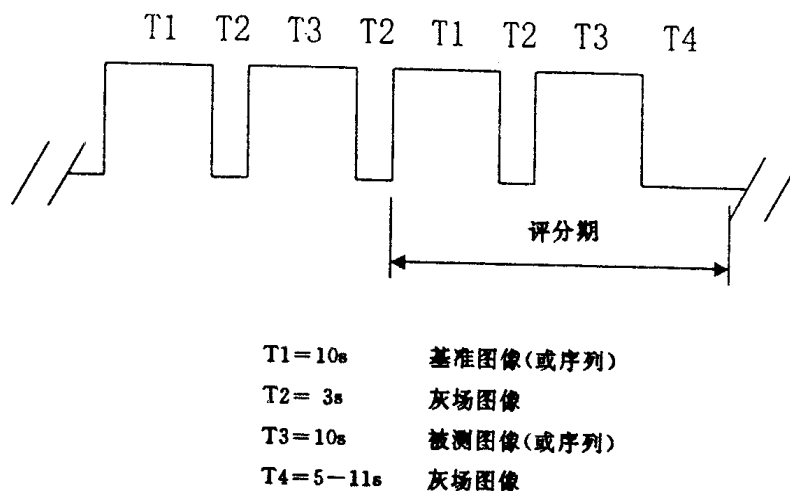


图 5 双刺激损伤标度法图像显示和评分顺序

5.2.4 数据处理

具体的数据统计方法见附录 A。

5.2.5 结果表达

结果表达方式见 4.7 节。

附录 A

(标准的附录)

数字电视图像质量主观评价数据统计方法

A1 双刺激连续质量标度法的数据统计

首先要确认评价过程的正确性和数据的有效性。先将评分表上所有基准和被测状态的测量值转换成标称分 0 到 100 的整数,并区分整理出每个测试图像基准图像和被测图像的评分值。然后进行评分的一致性检验,即用每位观看员对同一图像在同一测试阶段中的不同测试周期给出的两个评分值来核对,如果评分值相差 20 分(含 20 分)以上,则被视作无效而舍去,若某位观看员在同一测试阶段给出的有效评分次数低于其总评分次数的 85% 时,取消其全部评分。

最后计算出每个测试图像两种状态的平均分、标准偏差和 95% 置信度区间。

$$\text{平均分 } U = \frac{1}{N} \sum_{i=1}^N u_i$$

$$\text{标准偏差 } S = \sqrt{\frac{\sum_{i=1}^N (U - u_i)^2}{N - 1}}$$

95% 置信度区间为 $[U - \delta, U + \delta]$

$$\text{其中 } \delta = 1.96 \frac{S}{\sqrt{N}}$$

以上各式中: u_i ——观看员 i 的评分;

N ——观看员的数目

A2 双刺激损伤标度法的数据统计

首先要确认评价过程的正确性和数据的有效性,即进行评分的一致性检验,用每位观看员在同一测试阶段中对同一图像的不同测试周期给出的两个评分值来核对,如果评分值相差 2 级或 2 级以上,则被视作无效而舍去,若某位观看员在同一测试阶段给出的有效评分次数低于其总评分次数的 85% 时,取消其全部评分。

最后计算出每个被测图像的平均分、标准偏差和 95% 置信度区间,计算公式见 A1。

A3 一个测试阶段中,全体观看员给出的累加有效评分次数总和低于应有总评分次数的 85% 时,取消该测试阶段的全部评分。

数字电视图像质量主观评价方法

编制说明

1 任务来源

根据原广播电影电视部广技监字[1996]590号文“关于下达广电部1996年制、修订行业标准项目计划”的通知中第13项,制定《数字电视图像质量主观评价方法》行业标准,由广播科学研究院电视研究所负责起草,由国家广播电影电视总局标准化规划研究所归口管理。

2 编制原则及目的

本标准编制原则是在适合我国实际情况的前提下,充分向国际标准靠拢,以便与国际标准接轨。

编制本标准的目的是统一我国数字电视图像质量主观评价方法,使主观评价方法标准化,为数字电视图像质量评审提供重要依据,为数字电视发展建立基础;另外,通过使用本标准中规定的数字电视图像质量主观评价方法,能与国际推荐的方法一致,便于进行国际间的技术交流,促进广播电视技术的发展。

3 主要工作过程

在以ITU-R BT. 500-7(1997年)和Rec. 710-1(1992年)建议书为依据的前提下,在国家“八五”科技攻关项目《高清晰度电视技术研究》(HDTV图像质量的主观评价专题)研究成果的基础上,我们对较有代表性的双刺激法和单刺激法做了多种主观评价方法的主观评价试验,并对试验结果进行了深入分析、研究和比较。实践证明双刺激连续质量标度法最适合于对数字电视图像质量进行主观评价。

在理论分析和多次实践的基础上,我们写出了《数字电视图像质量主观评价方法》标准的初稿,并在技术内容和文字修饰上,经过了仔细推敲和10余次的修改之后,形成征求意见稿。将征求意见稿送交19个有关单位和专家征求意见,根据有关专家的意见,进行补充修改,形成送审稿。

4 技术内容的有关说明

本标准主要内容包括主观评价的一般要求和两种评价方法。主观评价的一般要求包括七个方面:观看条件、评价系统、信号源、测试素材的选择、观看员、评价测试阶段和结果表达。两种主观评价方法为双刺激连续质量标度法和双刺激损伤标度法。具体有以下几点说明:

1)关于双刺激连续质量标度法。在最佳条件(即理想收看条件)下确定电视系统性能时,使用这种主观评价方法。在双刺激连续质量标度法中,需要对每个测试图像的两种状态,即基准状态(直接来自信号源的图像)和被测状态(经被测系统处理输出的图像)进行评分。在不同测试图像的一连串评价测试中,基准和被测状态呈现的先后次序以伪随机方式变动,要求观看员只简单地对每个图像的总质量进行评价,并在评分表上作出标记。双刺激连续质量标度法使用连续质量标度。

2)关于双刺激损伤标度法。在非最佳条件(即如发送和传输等条件)下确定系统保持质量性能时,使用这种主观评价方法。双刺激损伤标度法是在观看员中每次首先显示没有损伤的基准图像,然后显示可能有损伤的同样图像,要求观看员根据第一个基准图像对第二个被测图像作出评价。双

刺激损伤标度法使用 5 级损伤标度。

5 本标准编写格式按照 GB/T1.1—1993。

6 建议本标准作为推荐性行业标准。