

GY

中华人民共和国广播电影电视行业标准

GY/T 262—2012

节目响度和真峰值音频电平测量算法

Algorithms to measure audio programme loudness and true-peak audio level

(ITU BS.1770-2: 2011, IDT)

2012 - 08 - 07 发布

2012 - 08 - 07 实施

国家广播电影电视总局 发布

目 次

前言	II
引言	III
1 范围	1
2 术语和定义	1
3 多声道响度客观测量算法	1
4 真峰值电平测量算法	5
附录 A (资料性附录) 多声道响度测量算法描述	8
附录 B (资料性附录) 关于准确测量数字音频信号峰值电平的考虑事项	16

前 言

本标准按照GB/T 1.1-2009给出的规则起草。

本标准使用翻译法等同采用 ITU-R BS.1770-2:2011《节目响度和真峰值音频电平测量算法》。

本标准由全国广播电影电视标准化技术委员会（SAC/TC239）归口。

本标准起草单位：国家广播电影电视总局广播电视规划院。

本标准主要起草人：张建东、邓向冬、崔俊生、覃毅力、宁金辉、韦安明、董文辉。

引 言

考虑到：

- a) 现代数字音频技术支持传输动态范围非常大的音频节目；
- b) 现代数字音频节目制作和传输技术支持单声道、立体声和多声道格式的混合应用，因此以这三种格式制作的声音节目并存；
- c) 受众希望不同来源和不同类型的节目声音在主观响度上保持一致；
- d) 音频电平测量方法有很多，但在节目制作中，现有测量方法不能提供主观响度指示；
- e) 为在节目交换中实施响度控制，提高受众满意度，有必要规定一个唯一的节目主观响度的客观测量算法；
- f) 将来基于心理声学模型的复杂算法，可能为多种多样的音频节目提供与主观感觉更一致的响度客观测量；
- g) 应避免数字媒体在记录中突然出现的过载，即便是瞬时的过载。

进一步考虑到：

- a) 通用的音频信号处理，如滤波处理或者比特率压缩，可能导致信号峰值电平的增加；
- b) 由于信号的真峰值可能出现在两个采样点之间，现有的测量技术没有反映出包含在数字信号中的真峰值电平；
- c) 目前的数字信号处理技术，使得实现一种算法来近似地估算信号的真峰值电平成为可能；
- d) 使用真峰值算法可以准确地指示出数字音频信号的峰值电平和削波电平之间的余量；
- e) 广播电视有使用指示节目响度和真峰值电平的仪表的需求。

建议：

- a) 为利于节目传送和交换而有必要对音频通道或节目的响度进行客观测量时，应采用本标准第 3 章规定的算法；
- b) 在节目制作和后期制作中用来测量节目响度的方法，可以本标准第 3 章规定的算法为基础；
- c) 当需要指示数字音频信号的真峰值电平时，测量算法应基于本标准第 4 章的规定，或者其他能给出类似或更准确测量结果的方法。

进一步建议：

- a) 进一步扩展第 3 章定义的算法，以适应对短期响度的测量；
- b) 如果证实新的响度测量算法的性能明显优于第 3 章规定的算法性能，可考虑更新本标准。

节目响度和真峰值音频电平测量算法

1 范围

本标准规定了一至五个声道节目音频的响度客观测量算法以及音频信号真峰值电平的测量算法。

本标准适用于在电视节目或广播节目的收录、制作、交换、传送和监测环节，对节目响度和真峰值电平的测量，也适用于响度和真峰值测量仪表的研究和开发。

2 术语和定义

下列术语和定义适用于本标准。

2.1

绝对响度单位 LKFS absolute loudness unit (LKFS)

使用符合本标准响度算法计算出的响度值的单位，以LKFS标示。其中，“K”表示测量中使用的频率加权，具体频率加权算法见第3章。LKFS表示“K加权下相对于标称满刻度的响度”。

2.2

真峰值电平 true-peak level

信号在连续时域中的最大峰值电平，是相对于离散采样点的最大峰值电平而言的。信号的真峰值电平可能大于在采样时刻获得的最大离散采样点的峰值电平。

2.3

RLB 加权 revised low-frequency B-curve weighting

对音频噪声电平测量中B加权曲线进行修改形成的加权曲线，具体频率响应见第3章。

3 多声道响度客观测量算法

本章定义了多声道响度客观测量模型化算法，该算法包含四个阶段：

- “K”频率加权；
- 各声道能量均方值计算；
- 声道加权求和（环绕声道的权重较高，且不包括LFE声道）；
- 在一个响度测量周期内，以400ms采样块组成的滑动窗口为单位（相邻窗口75%时间重叠）对响度测量值进行两步阈值处理：

- 第一步处理的阈值为-70LKFS；
- 第二步处理的阈值为经过第一步阈值处理后测得的响度值减去10dB。

多声道响度客观测量算法处理流程见图1。为便于算法描述，对图1中的流程节点设置了标签。图1显示了五个主要声道的输入（左、中、右、左环绕和右环绕），算法允许监测包含一至五个音频声道，对于少于五个声道的节目，一些输入可以不使用。本算法不对低频效果（LFE）声道进行测量。

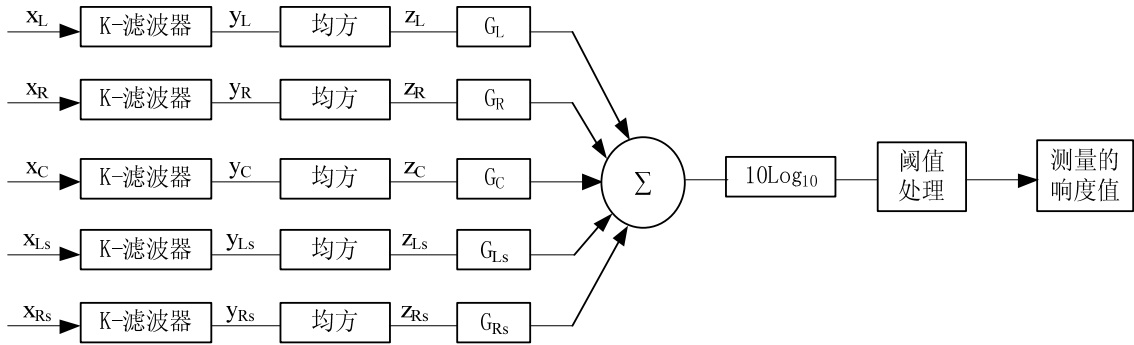


图1 多声道响度客观测量算法处理流程框图

算法的第一阶段是对信号进行两级预滤波，第一级用于考虑头部的声学效果，这里头部被建模为一个严格的球体。用于考虑头部声学效果的第一级预滤波滤波器的频率响应见图2。

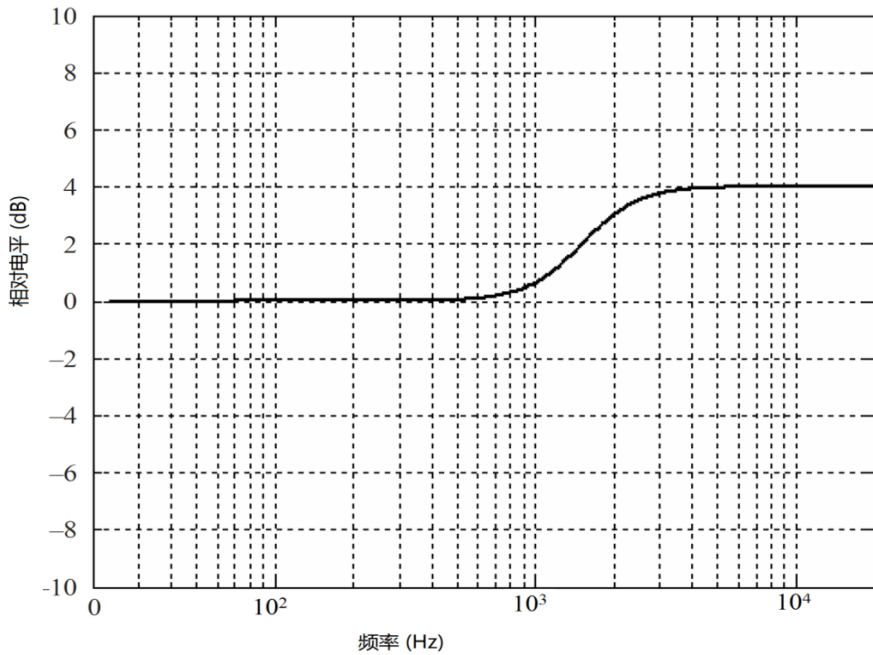


图2 用于考虑头部声学效果的第一级预滤波滤波器的频率响应

考虑头部声学效果的第一级预滤波滤波器的信号流程见图3。滤波器系数见表1。

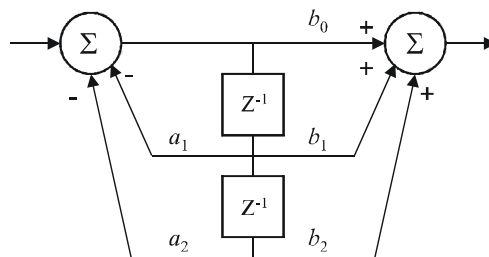


图3 一个二阶滤波器的信号处理流程

表1 第一级预滤波滤波器系数

a_1	a_2	b_0	b_1	b_2
-1.69065929318241	0.73248077421585	1.53512485958697	-2.69169618940638	1.19839281085285

表1规定的滤波器系数是针对采样率为48kHz音频信号的,针对其他采样率的音频信号则需要不同的滤波器系数,以使滤波器的频率响应符合图2所示的频率响应曲线。在算法的实际实现中,可能由于硬件支持的精度不够,需要对表1中的系数值进行量化。经过测试表明,本算法的性能对系数的微小变化不敏感。

第二级预滤波是应用RLB加权曲线,通过一个简单的高通滤波器实现。RLB加权二阶滤波器的信号流程见图3,频率响应曲线见图4,RLB加权滤波器系数见表2。

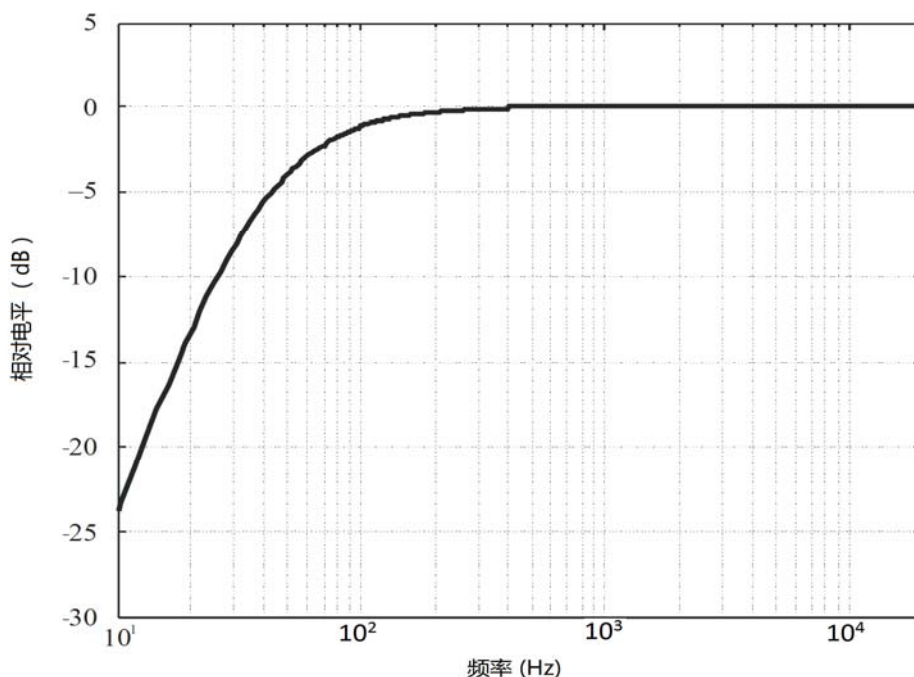


图4 第二级预滤波—RLB 加权滤波器的频率响应曲线

表2 第二级预滤波—RLB 加权滤波器系数

a_1	a_2	b_0	b_1	b_2
-1.99004745483398	0.99007225036621	1.0	-2.0	1.0

表2规定的滤波器系数是针对采样率为48kHz音频信号的,针对其他采样率的音频信号则需要不同的滤波器系数,以使滤波器的频率响应符合图4所示的频率响应曲线。

对预滤波后的信号计算测量周期T内的能量均方值 Z_i ,计算公式见式(1)。

$$Z_i = \frac{1}{T} \int_0^T y_i^2 dt \dots\dots\dots (1)$$

式中:

y_i ——对输入信号预滤波后的值;

$i \in I, I = \{L, R, C, Ls, Rs\}$, I 为输入声道的集合。

测量周期T内的响度 L_k 的定义见式(2)。

$$L_k = 0.691 + 10 \log_{10} \sum_i G_i \cdot z_i \quad \text{LKFS} \dots\dots\dots (2)$$

式中:

G_i ——各个声道的加权系数。

为使测量响度值与主观响度更加接近,需对按照式(2)测量出的响度值进行阈值处理。将测量周期T分为一系列交互重叠的滑动窗口,窗口长度 $T_g=400\text{ms}$,窗口内的一组连续音频采样值组成采样块,各个采样块应有75%的时间重叠。

测量周期T的末端应限制在最后一个完整采样块的结尾,即测量周期结束端对应的不完整的采样块不予使用。

在测量周期T内,第*i*个输入声道的第*j*个采样块能量均方值 z_{ij} 的计算见式(3)。

$$z_{ij} = \frac{1}{T_g} \int_{T_g \cdot j \cdot \text{step}}^{T_g \cdot (j \cdot \text{step} + 1)} y_i^2 dt \quad \dots\dots\dots (3)$$

式中:

$\text{step} = 1 - \text{重叠率}$; (重叠率=75%);

$$j \in \left\{ 0, 1, 2, \dots, \frac{T - T_g}{T_g \cdot \text{step}} \right\}.$$

第*j*个采样块的响度 l_j 的定义见式(4)。

$$l_j = -0.691 + 10 \log_{10} \sum_i G_i \cdot z_{ij} \quad \dots\dots\dots (4)$$

对于阈值 Γ ,存在一个集合 $J_g = \{j: l_j > \Gamma\}$,集合中的元素*j*为采样块响度值高于阈值的采样块在测量周期T内的索引, $|J_g|$ 表示集合 J_g 中元素的个数。

响度测量周期T内,进行阈值处理后的响度 L_{KG} 的定义见式(5)。

$$L_{KG} = -0.691 + 10 \log_{10} \sum_i G_i \cdot \left(\frac{1}{|J_g|} \cdot \sum_{j \in J_g} z_{ij} \right) \quad \text{LKFS} \dots\dots\dots (5)$$

测量算法的阈值处理分为两步,第一步为绝对阈值处理,第二步为相对阈值处理。相对阈值 Γ_r 通过式(6)计算,其中, Γ_a 为绝对阈值, $\Gamma_a = -70$ LKFS。

$$\Gamma_r = -10 \log_{10} \left(-0.691 + 10 \log_{10} \sum_i G_i \cdot \left(\frac{1}{|J_g|} \cdot \sum_{j \in J_g} z_{ij} \right) \right) \quad \text{LKFS} \dots\dots\dots (6)$$

式中:

$$J_g = \{j: l_j > \Gamma_a\};$$

用 Γ_r 按公式(5)计算阈值处理后的响度值。

式中:

$$J_g = \{j: l_j > \Gamma_r\}。$$

测量中使用的频率加权，由预滤波器（补偿头部声学效果的第一级滤波器和RLB加权第二级滤波器的级联）产生，命名为“K”加权。由式（2）计算出的响度值应以“LKFS”为单位，“LKFS”表示“K加权下相对于标称满刻度的响度”。信号电平每增加1dB，将导致响度值读数增加1LKFS，从这个意义上讲，LKFS单位与dB相当。

如果将一个0dB FS的1kHz正弦波信号分别输入到左声道、中置声道或右声道，则对应的响度指示值均为-3.01LKFS。

音频信号各个声道的加权系数见表3。

表3 音频信号各个声道加权系数

声道	加权, G_i
左 (G_L)	1.0 (0dB)
右 (G_R)	1.0 (0dB)
中置 (G_C)	1.0 (0dB)
左环绕 (G_{Ls})	1.41 (~+1.5dB)
右环绕 (G_{Rs})	1.41 (~+1.5dB)

有关多声道响度测量算法更详细地描述，参见附录A。

应该注意到：本标准规定的算法对于典型的广播内容的音频节目的响度估算有效，但是通常不适用于估算纯音信号的主观响度。

4 真峰值电平测量算法

本章描述了一种单声道线性PCM数字音频信号真峰值电平的估算算法，以下算法描述以48kHz采样率为例。真峰值电平是指信号波形在连续时域中的最大值（正值或负值），该数值可能大于以48kHz采样率采样获得的最大采样值。

考虑到实际应用中，测量点下游的一些处理设备有可能对原信号滤除直流分量，因此，本算法不仅可以估算原信号的真峰值电平，而且还包含了一个可选项，用来估算原信号滤除直流分量后的真峰值电平；本算法的另一个可选项是轻微的高频预加重，这样做是考虑下游信号处理环节（例如奈奎斯特滤波器）的相移可能引起高频信号峰值加大。预加重选项使算法估算出的高频信号峰值电平高于实际值，从而进一步防止在下游的处理环节中产生削波。

4.1 概述

处理步骤如下：

- a) 衰减：12.04dB的衰减；
- b) 4倍过采样；
- c) 加重（可选）：预加重搁架式滤波器，零点为14.1kHz，极点为20kHz；
- d) 隔直流（可选）；
- e) 取绝对值；
- f) 取最大值：最大值检测，若包含隔直流处理步骤则包含本步骤。

隔直流前后信号的绝对值检测结果可用于估算信号在当前测量点的真峰值电平，也可用于估算在下游设备滤除直流分量的情况下信号的真峰值电平。

4.2 框图

信号真峰值电平测量算法框图见图5。

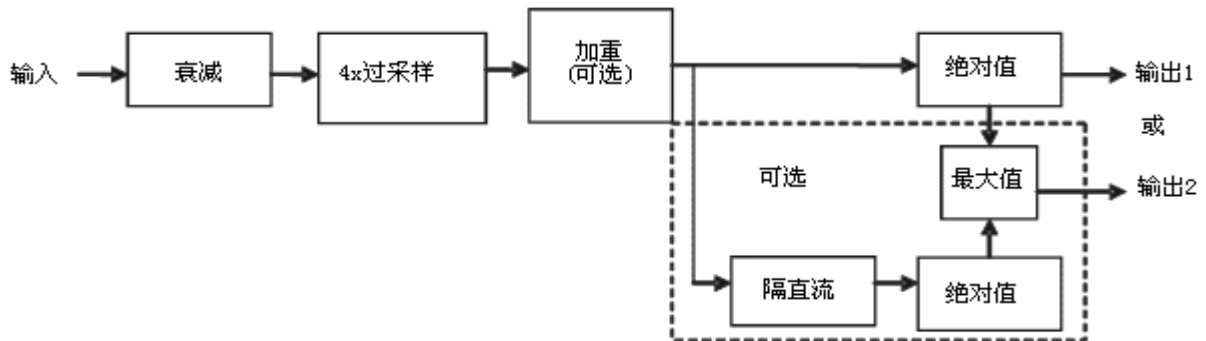


图 5 信号真峰值电平测量算法框图

4.3 说明

真峰值电平测量算法的第一步为对输入信号进行12.04dB的衰减（采样值右移2位）。信号衰减的目的是为采用整型运算的后续信号处理提供裕量。如果采用浮点计算，则无需本步。

4倍过采样滤波器将信号采样率从48kHz增加到192kHz。应用更高采样率得到的信号样本能够更准确地表示信号的真实波形。高采样率和高过采样比有助于准确估算信号的真峰值电平，具体参见附录B。如果输入信号本身是高采样，则可相应降低过采样比（例如，对于采样率为96kHz的输入信号，2倍过采样比则可达到较高的测量精度）。

考虑到音频信号处理中经常应用奈奎斯特滤波器，而滤波器产生的离散（相移）效应使测量和控制信号高频分量的峰值更加困难，在算法中选用预加重滤波器可提升信号的高频分量，使测量出的真峰值稍高于信号的实际峰值。

考虑到存在信号高度不对称或者信号中包含直流偏置的情况，算法除了可以测量当前测量点信号（包括非对称和/或含有直流偏置的信号）的真峰值电平，在算法中选用隔直流滤波器，还可以测量下游设备滤除直流分量后信号的真峰值电平。

取绝对值处理步骤对负的采样值取反得到采样的绝对值，即负值用相同幅度的正值代替。该步骤的输出信号是单极性的。如果不采用隔直流选件，输出数据流来自输出1；如果采用了隔直流选件，“取最大值”模块针对同一个采样点选择来自两个信号路径的较大采样值，输出数据流取自输出2。

后续的处理模块可将输出采样值与标称100%峰值信号电平（如果在输入端已经采用了12dB的衰减，则为满刻度的1/4）进行比较，产生相对于数字满刻度的真峰值电平测量值，本标准不对这些后续处理的实现进行规定。

遵循本真峰值电平算法且至少采用192kHz过采样率的真峰值表，应以“dB TP”为单位标识测量结果。该标识表示相对于100%满刻度的真峰值电平值。

附录 A

(资料性附录)

多声道响度测量算法描述

本附录描述了一个用于测量音频信号主观响度的客观算法。该算法能够准确测量单声道、立体声和多声道音频信号的响度。本算法的一个主要优势在于其简单性，可以很低的成本实现。本附录还描述了正式主观测试的结果，由此形成的数据库可用于评估客观响度测量算法的性能。

A.1 概述

在许多应用中需要测量和控制音频信号的响度。此类应用包括：在电视和无线电广播中，音频节目的类型和内容经常发生变化，在音乐、语音和声效或者这些类型的某些组合之间频繁地切换，这种节目内容的变化可能导致主观响度的显著改变；此外，对信号的各种动态处理也可能导致主观响度的明显改变。当然，主观响度对音乐产业也有重要意义，音乐录音中通常采用动态处理使录音节目的响度最大化。

近年来，无线电通信工作小组6P一直致力于为广播应用中的典型节目确定一种主观响度的客观测量算法。在工作的第一阶段，ITU-R专门对单声道信号的客观响度算法进行了评价，结果发现：一种加权的均方测量算法 Leq (RLB)对单声道信号的测量性能最佳。

进一步而言，广播应用需要能够测量单声道、立体声和多声道信号响度的仪表。本标准建议了一种新的响度测量算法，成功应用于单声道、立体声和多声道音频信号。新的多声道算法以 Leq (RLB)算法的直接扩展作为基础，并保留了单声道 Leq (RLB)算法的极低计算复杂度。

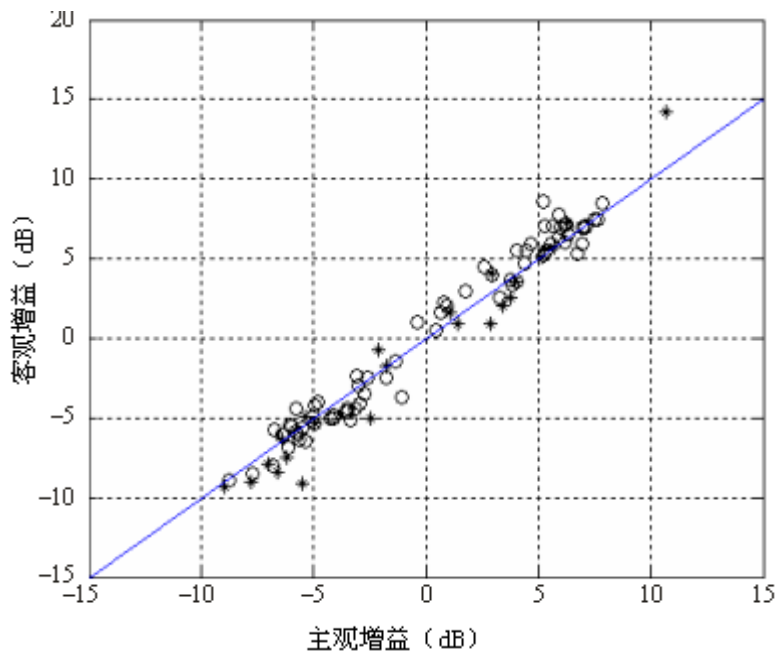
A.2 背景

在研究的第一阶段，ITU-R开发了一种主观测试方法，用于测试典型单声道节目素材的响度。并在全世界五个地点进行了主观测试，建立了主观测试结果数据库，可用于评估未来响度客观测量算法的性能。主观测试方法为：评价员通过调整重放增益将各种单声道音序列的响度和一个参考序列的响度相匹配。音序列取自实际的电视和广播素材。

同时，七家支持单位提交了总计十个单声道响度仪表/算法，在加拿大通信研究中心的音频感知实验室进行评价。

此外，还有另外两个基本响度算法作为性能对比的基础。一种测量算法 Leq (RLB)，包括一个简单的频率加权函数和其后的一个均方测量单元，加权函数是被称为RLB(修正低频B曲线)的一种高通频率加权曲线；另一种测量算法 Leq ，仅为无加权的均方值测量。

图A.1所示的是ITU-R对单声道序列，使用 Leq (RLB)响度表的测量结果和主观测试结果的对比。图中，横轴表示的是来自主观数据库的相对主观响度，纵轴表示的是由 Leq (RLB)算法客观测量的响度。图中每一个点代表一个音序列的测试结果。圆圈代表基于语音的音序列，星号代表非语音序列。可以看出数据点紧紧地聚集在对角线周围，表明 Leq (RLB)算法性能良好。



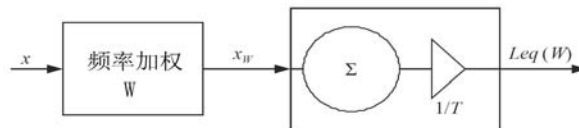
图A.1 单声道 Leq (RLB) 响度表客观测量结果对比主观结果 ($r=0.982$)

评估结果发现：在所有被评估的响度仪表中， Leq (RLB) 的性能最好（尽管从统计意义上，某些基于心理声学的响度测量算法同样性能良好）； Leq 表现的性能几乎和RLB一样好。评估结果说明对于典型的单声道广播素材，基于能量的简单算法，与更加复杂的可能包含详细感知模型的算法相比，效果同样不错。

A.3 Leq (RLB) 算法设计

Leq (RLB) 响度算法被特意设计得非常简单。 Leq (RLB) 算法框图见图A.2，它由一个高通滤波器和一个求一段时间内能量均值的处理单元组成。即信号先经过滤波器，滤波器的输出再进入处理单元，由处理单元进行求和并计算一段时间内的平均值。

滤波器的目的是根据人耳的感觉特性对信号频谱进行加权。响度测量算法采用这种基本结构的优势为：所有的处理均在简单的时域处理单元完成，计算量低。



图A.2 Leq (RLB) 算法框图

图A.2所示的 Leq (RLB) 算法只不过是等效声级 (Leq) 测量的一个频率加权形式， Leq 定义见式(A.1)。

$$Leq(W) = 10 \log_{10} \left[\frac{1}{T} \int_0^T \frac{x_W^2}{x_{Ref}^2} dt \right] \quad \text{dB} \dots\dots\dots (A.1)$$

式中：

x_W ——加权滤波器的输出信号；

x_{Ref} ——参考信号幅度；

T ——音序列长度；

$Leq(W)$ 中的符号 W 代表频率加权，此处采用的是修正低频B曲线(RLB)。

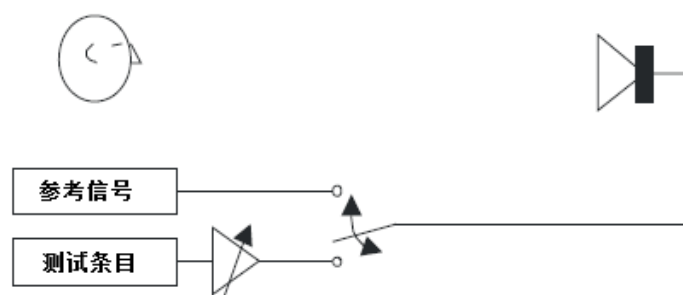
A.4 主观测试

为了评估多声道响度测量算法的优劣，有必要进行正式主观测试并建立响度主观测试数据库。通过主观数据库就可以评估客观响度测量算法预测主观测试结果的能力。数据库为各种单声道、立体声和多声道节目序列提供了主观响度评价结果。测试序列取自世界范围内真实的电视节目和广播节目，以及CD和DVD。序列类型包括音乐、电视剧和电影、体育赛事、新闻广播、音效和广告，序列中含有多种语言的语音片段。

A.4.1 主观测试安排

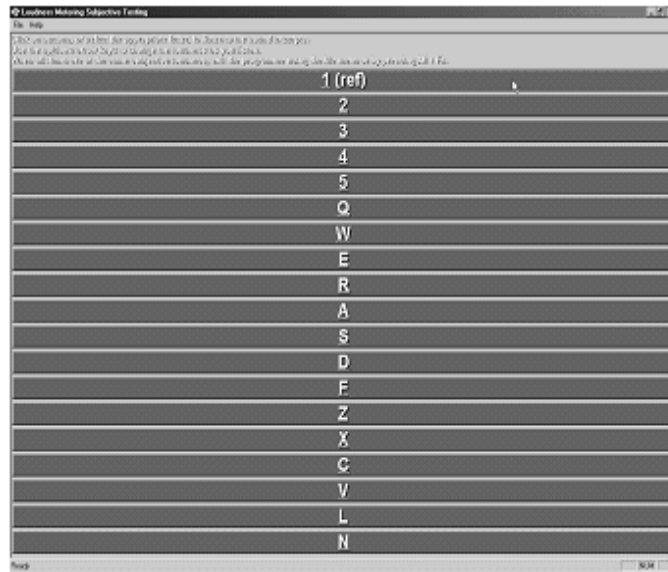
主观测试包含了一个响度匹配的过程。评价员收听各种类型的典型节目序列，调整每个测试条目的电平增益直至其感知响度与参考信号的响度相当，响度主观测试方法示意图见图A.3。

参考信号以60dBa声压级播放，该声压级数值为Benjamin所发现的在实际家庭中收看电视的典型值。



图A.3 主观测试方法示意图

测试采用澳大利亚广播公司开发并提交的基于软件的多声道主观测试系统，该系统允许评价员在测试条目之间来回瞬时切换，并调整各个条目的电平（响度）。测试软件的截屏图见图A.4，测试条目的电平能够以0.25dB为步长进行调整。选择标记为“1”的按钮播放参考信号，参考信号的电平保持固定不变。



图A.4 主观测试系统的用户界面截屏图

评价员使用计算机键盘选择一个给定的测试条目，调整增益，直至感觉其响度与参考信号的响度相当。通过按键选择，评价员可以在任意的测试条目之间瞬时切换。测试序列在测试期间连续（循环）播放。软件记录下评价员对各个测试条目的增益设置。因此，主观测试产生了一组使各个测试序列的响度与参考序列的响度相匹配所需的增益值，从而可以直接确定各个测试条目的相对响度。

正式盲测前，所有评价员均接受过训练，已熟悉测试软件和测试任务。由于很多测试条目混合了语音和其他声音（如音乐，背景噪声，等等），评价员已被专门告知，在评价时需将整个信号的响度，而不仅仅是信号中语音部分的响度与参考信号的响度相匹配。

正式盲测中，给每个评价员的测试序列的顺序是随机的。因此没有两个评价员的测试序列的顺序相同，从而消除了因顺序效应引起的任何可能的偏差。

A.4.2 主观数据库

用于评估响度客观算法性能的主观数据库实际上由三个单独的数据集组成，这三个数据集通过数年来开展的三个独立的主观测试建立。

第一个数据集由ITU-R的最初研究结果组成。在该项研究中，97名评价员在全世界五个不同的地点进行了主观测试，匹配了96个单声道音频序列的响度。由无线电通信WP 6P SRG3成员组成的一个三人专门小组挑选测试序列和参考序列。本次测试中的参考信号由英语女性语音组成。序列由直接放置在评价员前面的单个扬声器播放。

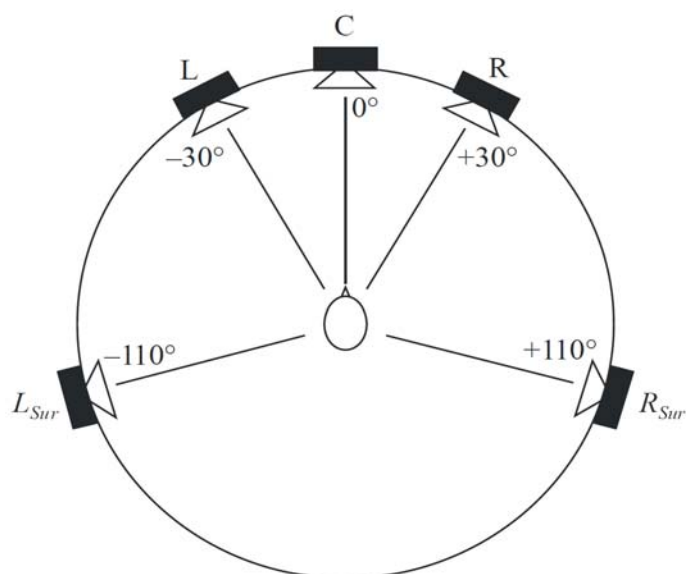
在最初的ITU-R单声道研究之后，一些算法的提交方推测主观测试中采用的序列的范围和类型还不够广泛，并进一步推测正是这个原因导致简单的基于能量的 Leq (RLB)算法优于其他算法。

为此，响度算法的提交方被要求提交自认为对 Leq (RLB)算法更具挑战性的新的单声道音频序列用于下一轮主观测试。只有两个响度算法的提交方提供了新的测试序列。

应用新序列，在加拿大通信研究中心的音频感知实验室进行了正式的主观测试。共计20名评价员对96个单声道序列进行了响度评价，测试采用与建立第一个数据集时相同的主观测试方法和参考信号。测试结果形成了主观测试数据库的第二个数据集。

第三个数据集包括144个音频序列的相对响度值，包括48个单声道序列、48个立体声和48个多声道序列。而且一半的单声道条目由中置声道（单声道）播放，另一半的单声道条目由左声道和右声道扬声器（两个单声道）播放。这样做是出于考虑收听单声道信号时可能采用的两种不同方式。本次测试中的

参考信号由具有立体声氛围和低电平背景音乐的英语女性语音组成。共计20名评价员参与了这次测试，扬声器配置采用ITU-R BS. 775建议书的规定，用于第三个数据集的扬声器配置见图A. 5。



图A. 5 用于第三个数据集的扬声器配置

前两个数据集限于单声道序列，因此无需考虑声像因素。对包含了立体声和多声道序列的第三个数据集，声像是一个重点考虑因素。一个序列的声像和环境感可能对序列的响度感觉产生显著影响，因此，第三个数据集中选择了包含广泛声像类型（例如，声像在中间与声像极左/极右，声源在前方与声源在四面八方）和环境类型（例如混响小与混响大）的立体声和多声道序列。

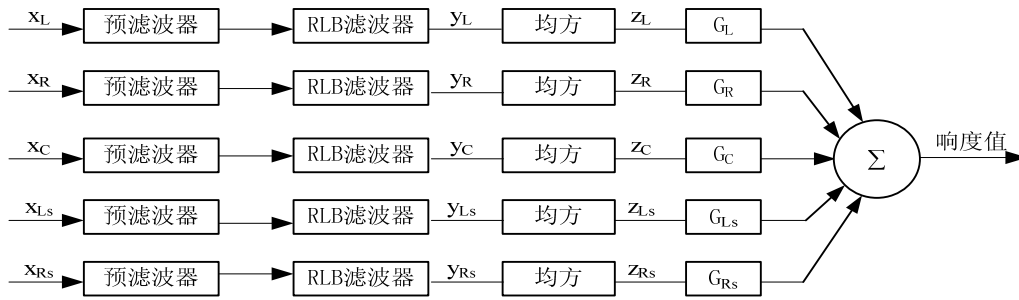
对第三个数据集，评价员需同时将单声道、立体声和多声道序列的响度与参考信号相匹配，这比前两个数据集时只对单声道序列进行匹配更为困难，而且，序列中包含的各种声像类型和环境感觉加剧了这种困难。因此，有人认为评价员会对测试任务不知所措。幸运的是，初步的测试结果表明该项测试是可控的，事实上，20名评价员能够提供一致的结果。

A. 5 多声道响度算法设计

如前所述，*Leq*(RLB)算法旨在用于单声道信号，而且，ITU的研究结果表明，该算法成功地应用于单声道。多声道响度算法的设计带来了新的挑战，一个成功的多声道算法的关键要求在于它必须也能很好地适用于单声道、双声道和立体声信号。也就是说，这些声道格式必须被看作是多声道信号的特例（虽然很常用）。

在当前的研究中，假设多声道信号符合ITU-R BS. 775建议书中5.1声道配置要求，不考虑LFE声道。

在多声道响度算法中，每个音频声道的响度测量独立地采用*Leq*(RLB)算法，推建的多声道响度算法框图见图A. 6。不过，在采用*Leq*(RLB)之前应对各个声道进行预滤波。



图A.6 推荐的多声道响度算法框图

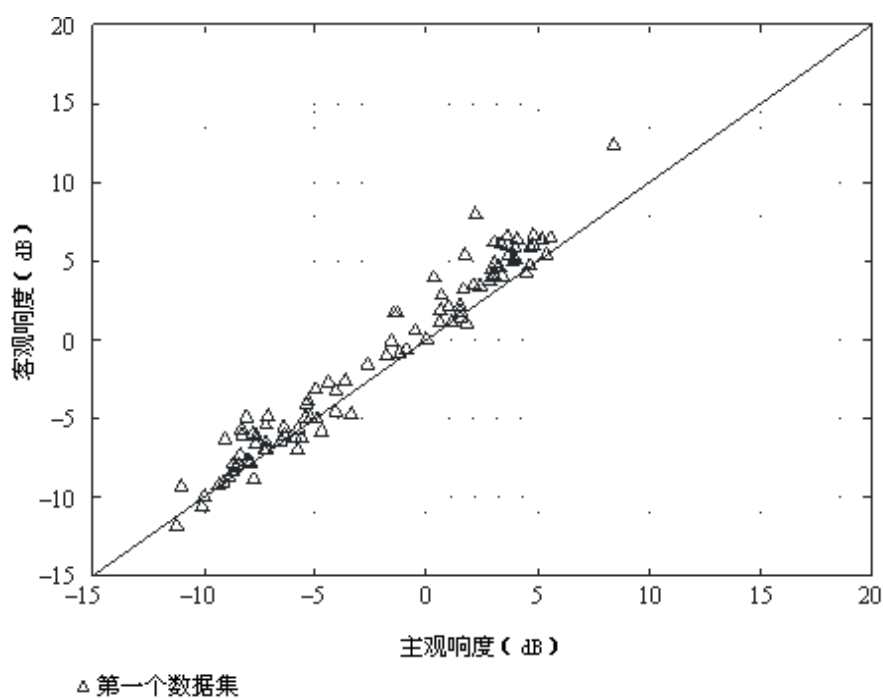
引入预滤波器的目的是考虑人体头部对到达信号的声学效应。这里，头部被模型化为一个严格的球体。各路声道采用相同的预滤波器。根据信号到达人耳的角度对产生的各路响度值进行加权(G_i)，然后求和（在线性域），从而提供一个合成的响度测量值。使用加权是为了反映这样一个事实：从收听者后面到达的声音比从前面到达的声音在感觉上响度更高。

该多声道响度算法的一个重要优势在于其简单性，该算法完全由非常基本的信号处理单元组成，能够在时域以廉价的硬件轻松地实现。该算法的另一个重要优势在于它的可伸缩性，由于各个声道采用的处理相同，所以很容易实现一种能够容纳从1到 N 任意声道数的响度测量仪表。此外，由于以各个声道响度作用的总和而非信号电平作用的总和作为响度值，算法不依赖于声道之间的相位或相关性，这使得本标准算法更加通用和强大。

A.6 多声道算法评估

应用多声道响度算法测量三个数据集中的336个音频序列，并记录测量值。某响度算法的总体性能可通过比较测量值与上述主观测试结果之间的一致性来评估。

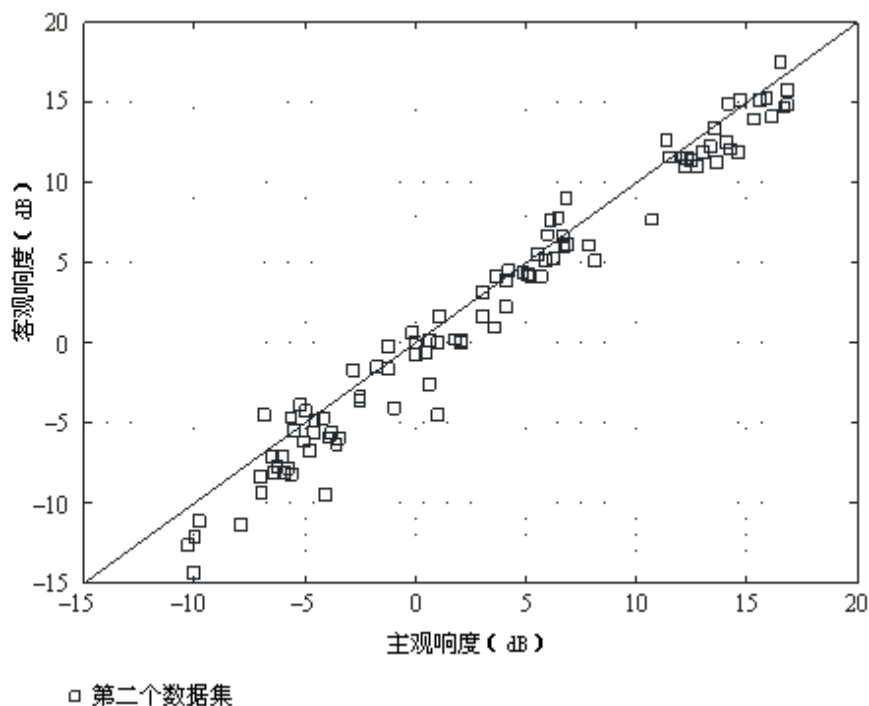
应用符合本标准算法的响度仪表测量三个数据集中各序列的响度，图A.7、图A.8和图A.9分别给出了三个数据集客观测量结果与主观测试结果的对比统计。图中，横轴表示数据集中每个音频序列的主观响度值，纵轴表示采用符合本标准响度仪表测量产生的客观响度值。图中每一点代表一个音频序列的结果。理想的客观算法将使所有的数据点均落在斜率为1且经过原点的对角线上。



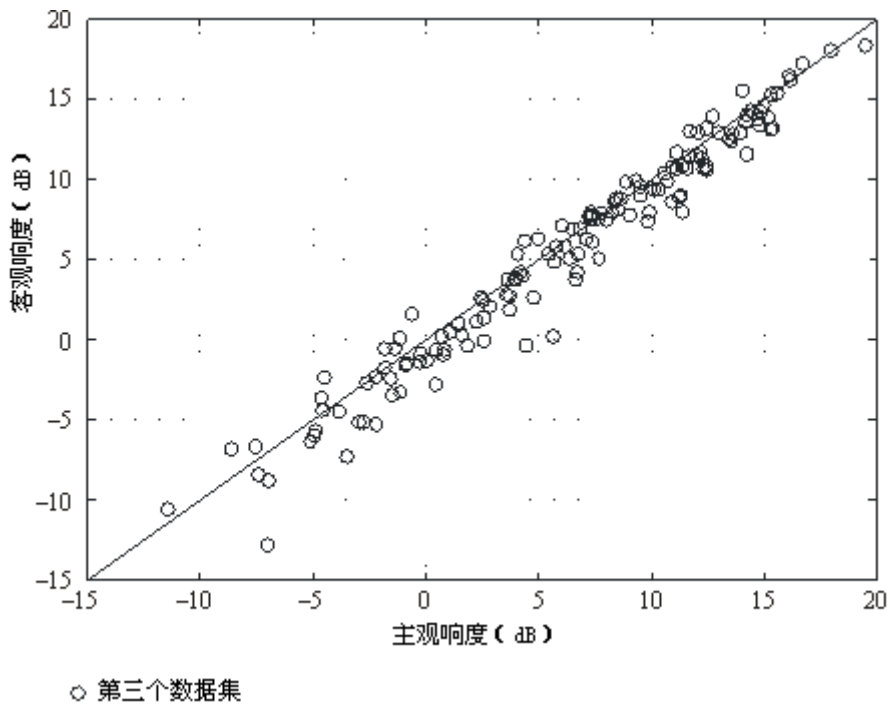
图A.7 第一个（单声道）数据集客观测量与主观测试结果的对比（ $r = 0.979$ ）

如图A.7所示，本标准规定的多声道响度算法在测量第一个（单声道）数据集时性能良好，主观响度测试与客观响度测量结果之间的相关性很好（ $r=0.979$ ）。

如图A.8所示，对于第二个数据集，主观响度测试与客观响度测量结果之间的相关性也非常好（ $r=0.985$ ）。该数据集中差不多一半的序列为音乐。



图A.8 第二个（单声道）数据集客观测量与主观测试结果的对比（ $r = 0.985$ ）

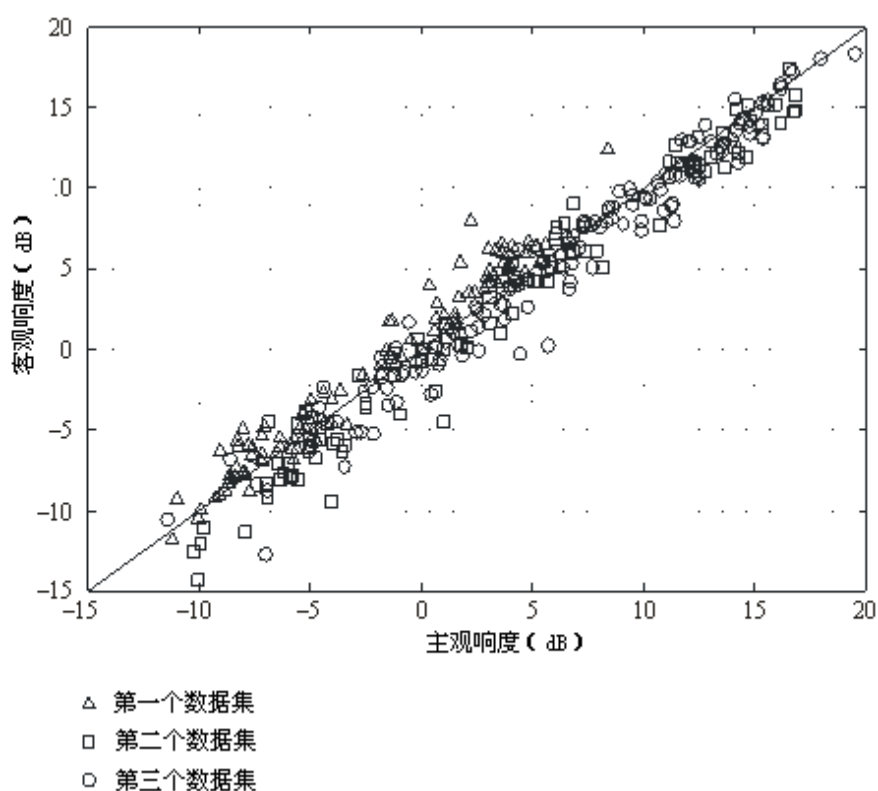


图A.9 第三个（单声道、双单声道、立体声和多声道）数据集客观测量与主观测试结果的对比（ $r = 0.980$ ）

第三个数据集的结果见图A.9，该数据集包括单声道、双声道、立体声和多声道信号。同样地，主客观吻合性非常好（相关性 $r=0.980$ ）。

图A.9和图A.10中包含的多声道结果是应用本标准算法产生的，但环绕声道的加权设为4dB（最初的提议）而不是1.5dB（最后的规范）。事实证实加权值从4.0dB到1.5dB的变化不会对结果产生任何显著的影响。

图A.10为对主观测试数据库三个数据集中所有336个音序列的测试结果，总相关性 r 为0.977。



图A.10 全部三个数据集客观测量与主观测试结果的对比 ($r = 0.977$)

评估结果表明基于 $Leq(RLB)$ 的多声道响度测量算法，应用于主观数据库中所有336个序列，得到的测量结果与主观测试结果的吻合性都非常好。由此证明，符合本标准的响度仪表适用于单声道、双声道、立体声以及多声道信号。评估用主观数据库中测试序列范围广泛，包括音乐、电视剧和电影、体育赛事、新闻广播、音效和广告。序列中也包含了多种语言的语音。

附录 B (资料性附录)

关于准确测量数字音频信号峰值电平的考虑事项

B.1 存在的问题

数字音频系统中的峰值电平表通常指示“采样峰值”，而不是“真正峰值”。

采样峰值表通常是各个输入采样的绝对值（经整流）与峰值表的当前读数进行比较，如果新采样值较大，则取代当前读数，否则，当前读数乘以一个略小于1的常数，产生一个对数衰减。这种峰值表易于实现，因而得到普遍应用，但不一定能够指示音频信号的真正峰值电平，根源在于真正的信号峰值通常出现在前后两个采样点之间而不是精确地出现在采样瞬间，造成采样峰值表错失指示信号的真正峰值。因而，在需要准确测量显示节目峰值的应用场合，采用采样峰值表可能会带来一些问题。

一些常见的采样峰值表应用异常包括：

- a) 峰值读数不一致。重复播放一段模拟录音并输入至带采样峰值表的数字系统，每次播放产生的节目峰值读数完全不同，这样的情景经常发生。同样地，重复播放一段数字录音，经采样率转换器后进行测量，每次播放显示的峰值也不同。这是因为每次播放时，采样时刻可能落在实际信号的不同位置。
- b) 意外过载。采用采样峰值表监测信号过载时，即使没有信号采样值达到或接近数字满刻度电平，实际信号也可能包含过载电平，因此，采样峰值表的过载指示并不可靠，可能在后续的D/A或重采样处理环节中导致削波。
- c) 纯音信号的读数偏低和差拍。采样峰值表的输入信号为频率接近于采样率整数分之一的纯音（例如校准音）时，可能出现读数偏低，或者即使纯音信号的幅度固定也可能出现不断变化的读数。

B.2 问题的严重性

通常，输入信号频率越高，采样峰值表出现以上问题的可能性越大。

对于连续纯音信号，读数偏低很容易证明。例如，频率为四分之一采样率、相位不巧的单音，采样峰值表读数比信号的真正峰值低3dB；频率为二分之一采样率的单音，读数低很多，但绝大多数的数字音频信号在该频点能量很弱（原因是，一方面在D/A转换期间，该频点能量基本上被抗混叠滤波器滤除了，另一方面是因为“真实的”声音中连续高频信息并不占主导地位）。

连续单音信号频率不接近小的整数分之一倍采样率时，由于差频（ $n \cdot f_{tone}$ 和 f_s 之间的差）高于峰值表衰减速率的倒数，故在采样峰值表中不会出现读数偏低。换言之，采样时刻越接近单音的真正峰值时刻，采样峰值表越不容易出现读数偏低。

不过，对于信号中的个别瞬态，仍可能出现读数偏低的情况，而且出现瞬态的信号频率越高，读数可能越低。在“真实的”声音中出现具有显著高频成分的瞬态是正常的，由此引起的读数偏低通常为数分贝。

由于真实声音信号的频谱能量通常随频率升高而降低，而且信号频谱不会因采样率的增加而改变，采用较高采样率时，采样峰值表读数偏低会明显降低。

B.3 解决方案

为了测量被采样信号的真正峰值，有必要对信号进行“过采样”（或“上采样”），从本质上重建已有采样之间的原始信号，从而提高信号的采样率。

为确定所需要的过采样比，首先需要确定：

- a) 最大可接受的读数偏低误差；
- b) 被测量的最高频率与采样率的比（最大“归一化频率”）。

如果知道这些参数，就可以通过一种直接的方法计算出需要的过采样比（甚至不用考虑过采样的实现细节）。简单起见，可以将最大归一化频率点，对称地出现在正弦曲线波峰两侧的一对过采样值产生的最大读数偏低当作“最坏情况”下的读数偏低。

因而对于：

过采样比： n ；

最大的归一化频率： f_{norm} ；

采样率： f_s ；

可以得到：

在过采样率下的采样周期为： $\frac{1}{n \cdot f_s}$ ；

最大归一化频率的信号周期为： $\frac{1}{f_{norm} \cdot f_s}$ ；

则：

最大的读数偏低（dB）为 $20 \cdot \log(\cos(\frac{2\pi \cdot f_{norm} \cdot f_s}{n \cdot f_s \cdot 2}))$ ，（分母中的 2 是因为错过波峰的最大间隔为

半个过采样周期）；

简化为：

最大的读数偏低（dB）= $20 \cdot \log(\cos(\frac{\pi \cdot f_{norm}}{n}))$ 。

由以上公式可得到表 B. 1。

表 B. 1 不同过采样比下的真峰值表读数偏低的最大值

过采样比	读数偏低（dB）最大值	
	$f_{norm} = 0.45$	$f_{norm} = 0.5$
4	0.554	0.688
8	0.136	0.169
10	0.087	0.108
12	0.060	0.075
14	0.044	0.055
16	0.034	0.042
32	0.008	0.010

B.4 真峰值表的实现

为实现过采样，在原始数据流采样之间插入0值采样，得到对应过采样率速率的数据流，然后应用低通“内插”滤波器滤除所要求的最大 f_{norm} 以上的频率成分。此时再对过采样信号运用采样峰值算法，将得到满足最大读数偏低要求的真峰值表。

下面研究过采样器的实现。这种低通滤波通常被当作一个对称的FIR来实现。在采用低通滤波器传送高质量音频的应用中，例如在（老式的）过采样D/A转换器或者采样率转换器中，为了保持极低的通带纹波、获得最大的阻带衰减和狭窄的过渡带，有必要计算大量的“抽头”。为了保持动态范围且使失真最小化，也必须保留较长的字长。

然而，真峰值表中过采样器的输出不用于收听，只是用于显示读数或者条形图，因此不需要同样的精度要求。只要通带纹波，加上来自阻带的伪分量的耦合造成的读数准确度的降低不超出目标允许的范围即可。因此，尽管仍然需要一个由最大归一化目标频率决定的狭窄过渡带，所需的抽头数量已大大减少。同样，除了对于较低的电平幅度要求准确的数值输出的情况，只需足够的字长保证目标精度即可。

因此，适当的过采样器（可能适用于很多声道）有可能用普通的低成本DSP或者FPGA或者更普通的处理器轻松实现。另一方面，采用原用于D/A转换器的高精度过采样芯片已经实现了过采样峰值表，虽然这相当浪费硅片和能量，但是元器件比较廉价和易于获得。

对于一种特定的峰值表规范，确定必需的抽头数量和抽头系数的最简单方法是使用递归FIR滤波器设计程序，例如Remez或者Meteor。

按惯例，音频表采用隔直流设计，因此在峰值表中也要求去除输入信号中直流分量的影响。另一方面，如果关注真峰值电平值的目的是为了消除过载，则必须保留并测量直流分量。如果需要，可以在峰值表的输入端引入一个低阶IIR高通滤波器，从而以较低的计算强度去除直流对峰值表的影响。

为了强调某一部分频带的影响，有时需要在采用一些加权滤波器之后再计量峰值信号幅度。实现方式取决于具体加权滤波器的特性。

参 考 文 献

- [1] ITU-R BS.775-1 Multichannel stereophonic sound system with and without accompanying picture



中 华 人 民 共 和 国
广 播 电 影 电 视 行 业 标 准
节 目 响 度 和 真 峰 值 音 频 电 平 测 量 算 法
GY/T 262—2012

*

国家广播电影电视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.abp.gov.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