

GY

中华人民共和国广播电视和网络视听行业标准

GY/T 262—XXXX

代替 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-R BS. 1770-5:2023, MOD)

(报批稿)

XXXX—XX—XX 发布

XXXX—XX—XX 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 多声道（多至 3/2 声道）响度客观测量算法	2
6 先进声音系统扬声器配置的扩展响度测量算法	6
7 基于对象的音频信号或基于声道与对象混合的音频信号的响度测量算法	10
8 真峰值电平测量算法	10
8.1 通则	10
8.2 框图	10
8.3 详细说明	11
附录 A（资料性） 本文件与 ITU-R BS.1770-5:2023 结构编号对照一览表	12
附录 B（资料性） 多声道响度测量算法描述	13
B.1 概述	13
B.2 Leq(RLB) 算法设计	13
B.3 主观测试	13
B.4 多声道响度算法设计	15
B.5 多声道算法评估	16
附录 C（资料性） 基于对象的音频信号或基于声道与对象混合的音频信号的响度测量算法与渲染条件的关系	20
C.1 概述	20
C.2 客观响度测量结果因渲染条件不同而产生的差异	20
C.3 主观测试	20
附录 D（资料性） 数字音频信号准确峰值电平测量的技术要点	24
D.1 存在的问题	24
D.2 对测量准确性的影响	24
D.3 解决方案	24
D.4 真峰值表的实现	25
参考文献	27

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。

本文件代替GY/T 262—2012《节目响度和真峰值音频电平测量算法》，与GY/T 262—2012相比，除结构调整和编辑性改动外，主要技术变化如下：

- a) 增加了“先进声音系统扬声器配置的扩展响度测量算法”（见第6章）；
- b) 增加了“基于对象的音频信号或基于声道与对象混合的音频信号的响度测量算法”（见第7章）；
- c) 增加了“基于对象的音频信号或基于声道与对象混合的音频信号的响度测量算法与渲染条件的关系”（见附录C）。

本文件修改采用ITU-R BS. 1770-5:2023《节目响度和真峰值音频电平测量算法》。文件类型由ITU的建议书调整为我国的推荐性行业标准。

本文件与ITU-R BS. 1770-5:2023相比，在结构上有较多调整。两个文件之间的结构编号变化对照一览表见附录A。

本文件与ITU-R BS. 1770-5:2023的技术差异及其原因如下：

- 更改了“范围”的表述，并补充了文件的适用范围（见第1章），以符合我国标准编写要求；
- 增加了“术语和定义”一章（见第3章），以符合我国标准编写要求；
- 增加了“缩略语”一章（见第4章），以符合我国标准编写要求；
- 将图2、图4、表3、表4的提及方式更改为要求型条款（见图2、图4、表3、表4）；
- 将响度读数（见第5章）和真峰值表读数（见8.3）单位标识的推荐性条款修改为要求型条款，以适应我国实际应用需求；
- 将引用的ITU-R BS. 2051替换为我国对应的标准GB/T 42758；
- 将客观响度测量所用渲染条件的记录修改为一般性陈述（见C.2），以符合我国标准编写要求。

本文件做了下列编辑性改动：

- 删除了ITU-R BS. 1770-5:2023的“关键词”；
- 将引用的IEC 61606替换为我国对应的标准GY/T 285；
- 将ITU-R BS. 1770-5:2023中的两处“参考文献”合并，以符合我国标准编写要求；
- 删除了附录B“多声道响度测量算法描述”中的背景信息。

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、苏州大学、华为技术有限公司、腾讯科技（深圳）有限公司、西藏广播电视台、音王电声股份有限公司。

本文件主要起草人：张建东、邓向冬、宁金辉、陈晨、朱梦尧、夏丙寅、李大龙、晋美、王素君、崔俊生、覃毅力、韦安明、董文辉、汪芮、刘汉源、王倩男。

本文件及其所代替文件的历次版本发布情况为：

- 2012年首次发布为GY/T 262—2012，本次为第一次修订。

节目响度和真峰值音频电平测量算法

1 范围

本文件规定了节目音频响度的客观测量算法以及音频信号真峰值电平的测量算法。

本文件适用于广播电视和网络视听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 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 42758 用于节目制作的先进声音系统

3 术语和定义

下列术语和定义适用于本文件。

3.1

绝对响度单位 absolute loudness unit

使用符合本文件响度算法计算出的响度值的单位，以LKFS标示。

注：“K”表示测量中使用的频率加权，具体频率加权算法见第5章。LKFS表示“K加权下相对于标称满刻度的响度”。

3.2

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。

注：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。

3.3

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。

注：具体频率响应见第5章。

4 缩略语

下列缩略语适用于本文件。

DSP 数字信号处理器 (Digital Signal Processor)

dB TP 真峰值分贝 (decibels True Peak)

FIR 有限脉冲响应 (Finite Impulse Response)

FPGA 现场可编程门阵列 (Field-Programmable Gate Array)

- LFE 低频效果 (Low Frequency Effects)
- LU 响度单元 (Loudness Unit)
- PCM 脉冲编码调制 (Pulse Code Modulation)
- RLB 修正低频B曲线加权 (Revised Low frequency B curve weighting)

5 多声道（多至 3/2 声道）响度客观测量算法

多声道响度客观测量算法包含四个步骤。

- a) “K” 频率加权。
- b) 各声道计算均方功率。
- c) 声道加权求和（环绕声道的权重较高，且不包括 LFE 声道）。
- d) 对 400ms 数据块（75%样本重叠）的门限处理，其中设置双阈值：
 - 1) 第一阈值：-70LKFS；
 - 2) 第二阈值：经过第一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。为便于描述，图1中各流程节点设置了标签。图1显示了五个主要声道的输入（左、右、中、左环绕和右环绕），即测量算法允许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LFE声道进行测量。

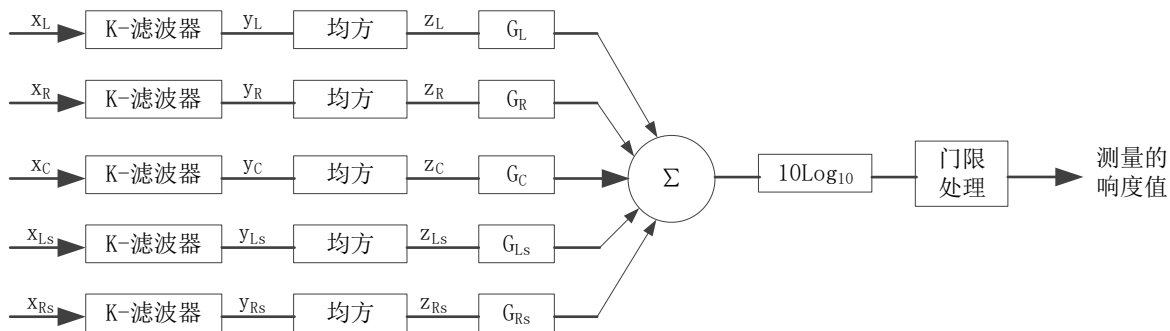


图 1 多声道响度客观测量算法简化框图

算法的第一步是对信号进行两级预滤波，该两级预滤波统称为K加权滤波。第一级预滤波用于考虑头部的声学效果，头部被建模为刚性球体。第一级预滤波器的频率响应应与图2相符合。

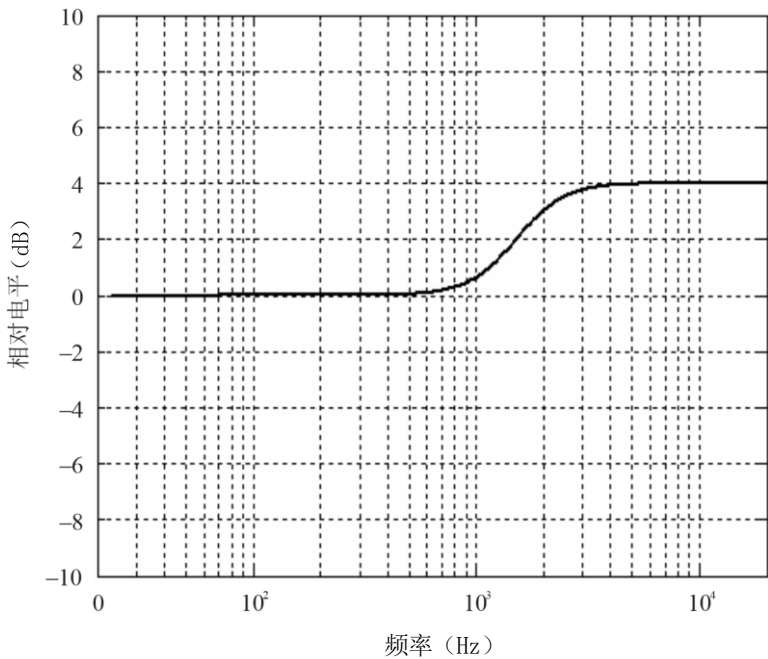


图 2 用于考虑头部声学效果的第一级预滤波器的频率响应

考虑头部声学效果的第一级预滤波器的信号流程见图3。滤波器系数见表1。

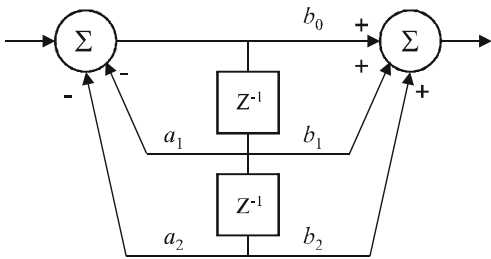


图 3 一个二阶滤波器的信号处理流程

表 1 第一级预滤波器系数

系数	系数值	系数	系数值
—	—	b_0	1.53512485958697
a_1	-1.69065929318241	b_1	-2.69169618940638
a_2	0.73248077421585	b_2	1.19839281085285

表1规定的滤波器系数适用于48kHz音频信号，其他采样率则需要不同的滤波器系数，以使滤波器的频率响应与图2相符。在算法的实际实现中，由于硬件的精度限制，可能需要对表1中的系数值进行量化。测试表明，算法性能对系数的微小变化不敏感。

第二级预滤波通过一个简单的高通滤波器实现，滤波器的频率响应应与图4相符。滤波器系数见表2。

表2规定的滤波器系数适用于48kHz音频信号，其他采样率则需要不同的滤波器系数，以使滤波器的频率响应与图4相符。

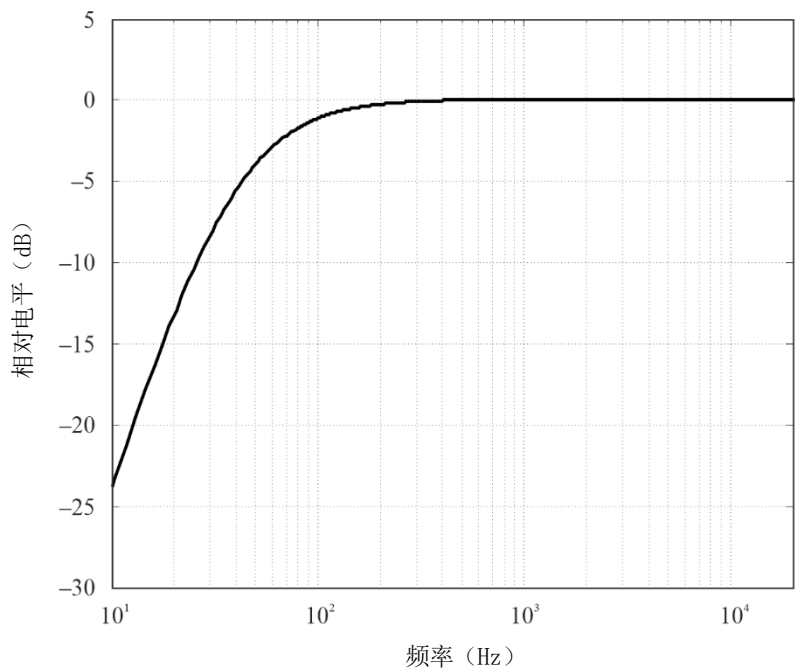


图 4 第二级预滤波加权滤波器的频率响应曲线

表 2 第二级预滤波加权滤波器系数

系数	系数值	系数	系数值
—	—	b_0	1.0
a_1	-1.99004745483398	b_1	-2.0
a_2	0.99007225036621	b_2	1.0

在一个测量周期 T 内，对预滤波后的信号计算平均功率 z_i ，计算公式见公式（1）。

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

式中：

- z_i —— 经两级预滤波后，在测量周期 T 内，声道 i 输入信号的平均功率；
- T —— 测量周期；
- y_i —— 经过两级预滤波后，第 i 个输入声道的瞬时信号幅度， $i \in I$ ， $I = \{L, R, C, Ls, Rs\}$ （输入声道的集合）。

测量周期 T 内的响度 L_K 的定义见公式（2）。

$$L_K = -0.691 + 10 \log_{10} (\sum_i G_i \cdot z_i) \dots\dots\dots (2)$$

式中：

- L_K —— K 加权下的响度测量值，单位为LKFS；
- G_i —— 声道 i 的加权系数；
- z_i —— 经两级预滤波后，在测量周期 T 内，声道 i 输入信号的平均功率。

为进行门限处理，将测量周期 T 分为一系列交互重叠的施加门限的处理块，每个块为持续时间 $T_g=400\text{ms}$ 的连续音频采样（精确到最接近的采样点）集合，各个相邻块应有75%的时间重叠。

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

在测量周期 T_g 内，第 i 个输入声道的第 j 个处理块的平均功率 z_{ij} 的计算见公式（3）。

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

式中:

z_{ij} ——在测量周期 T_g 内, 第 i 个输入声道的第 j 个处理块的平均功率;

T_g ——处理块的时长400ms;

y_i ——经两级预滤波后, 声道 i 在第 j 个处理块内的瞬时信号幅度;

$\text{step} = 1 - 0.75$ (重叠率);

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

第 j 个门限处理块的平均响度 l_j 定义见公式(4)。

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

式中:

l_j ——第 j 个门限处理块的平均响度;

G_i ——声道 i 的加权系数;

z_{ij} ——第 i 个输入声道的第 j 个处理块的平均功率。

响度测量周期 T 内, 进行门限处理后的平均响度 L_{KG} 的定义见公式(5)。

$$L_{KG} = -0.691 + 10 \log_{10} \left(\sum_i G_i \cdot \left(\frac{1}{|J_g|} \sum_{J_g} z_{ij} \right) \right) \quad \dots\dots\dots (3)$$

式中:

L_{KG} ——进行门限处理后的平均响度;

G_i ——声道 i 的加权系数;

$|J_g|$ ——集合 J_g 中元素的个数;

J_g ——门限处理块索引集合, $J_g = \{j: l_j > \Gamma\}$, Γ 为门限阈值;

z_{ij} ——在测量周期 T 内, 第 i 个输入声道的第 j 个处理块的平均功率。

门限处理分为两步, 第一步采用绝对门限, 第二步采用相对门限:

a) 绝对门限处理: 绝对门限阈值 $\Gamma_a = -70\text{LKFS}$, 低于该阈值的处理块不进入相对门限处理过程;

b) 相对门限处理: 相对门限阈值 Γ_r 为经过绝对门限处理的整体响度减去10dB, 见公式(6)。

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

式中:

Γ_r ——相对门限阈值;

G_i ——声道 i 的加权系数;

$|J_g|$ ——集合 J_g 中元素的个数;

J_g ——门限处理块索引集合, $\Gamma_a = -70\text{LKFS}$, $J_g = \{j: l_j > \Gamma_a\}$;

z_{ij} ——在测量周期 T 内, 第 i 个输入声道的第 j 个处理块的平均功率。

经过两步门限处理后的响度值即为最终测量结果, 见公式(7)。

$$L'_{KG} = -0.691 + 10 \log_{10} \left(\sum_i G_i \cdot \left(\frac{1}{|J_g|} \sum_{J_g} z_{ij} \right) \right) \quad \dots\dots\dots (7)$$

式中:

L'_{KG} ——经过两步门限处理后的平均响度;

G_i ——声道 i 的加权系数;

$|J_g|$ ——集合 J_g 中元素的个数;

J_g ——门限处理块索引集合, $J_g = \{j: l_j > \Gamma_r \text{ 且 } l_j > \Gamma_a\}$;

z_{ij} ——在测量周期 T 内, 第 i 个输入声道的第 j 个处理块的平均功率。

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。由公式（2）计算出的响度数值结果应以单位“LKFS”标识。信号电平每增加1dB，响度值读数相应增加1LKFS，在此意义上，LKFS单位与dB相当。

如果将0dB FS的1kHz（准确值为997Hz，见注1和注2）正弦波信号分别输入到左声道、中置声道或右声道，显示的响度值为-3.01LKFS。

注1：公式（2）中的常数-0.691用于抵消997Hz频率下的K加权增益。

注2：GY/T 285规定，测量的参考频率为997Hz，在非严格场景下，表述为标称频率1kHz。

音频各声道的加权系数应符合表3的规定。

表 3 音频各声道加权系数

声道	加权系数, 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)

本文件规定的算法对于典型的广播内容的响度测量有效，但是通常不适用于测量纯音信号的主观响度，响度测量（估算）效果见附录B。

6 先进声音系统扬声器配置的扩展响度测量算法

本章规定了用于先进声音系统中任意扬声器配置和布局下的客观响度测量算法，是对第5章规定的3/2多声道基本算法的扩展。算法增加了输入声道数量，并修改了基本算法的步骤c)：声道加权求和（除LFE声道外，每个声道都有一个取决于其所处位置的水平角和俯仰角的加权系数 G_i ）。

GB/T 42758中规定的先进声音系统扬声器配置下的客观响度测量算法框图，见图5。N是除LFE声道外的输入声道数。该算法的第一、第二和第四步（滤波、计算均方功率和门限处理）独立于声道位置，与3/2多声道所用基本算法的对应阶段完全相同。

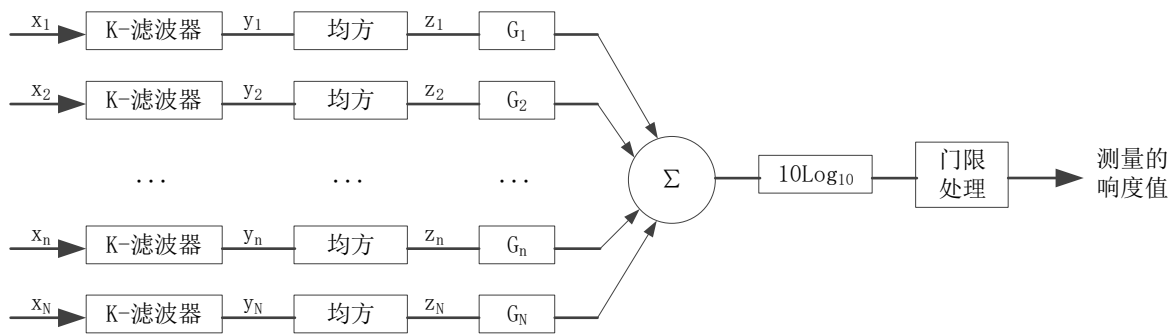


图 5 先进声音系统扬声器配置下的客观响度测量算法框图

声道的位置相关加权系数 G_i 应符合表4的规定。 G_i 取决于声道位置的方向，该方向由水平角(θ)和俯仰角(ϕ)确定。

表4 声道的位置相关加权系数

俯仰角 (φ)	水平角 (θ)		
	$ \theta < 60^\circ$	$60^\circ \leq \theta \leq 120^\circ$	$120^\circ < \theta \leq 180^\circ$
$ \varphi \leq 30^\circ$	1.00 ($\pm 0\text{dB}$)	1.41 (+1.5dB)	1.00 ($\pm 0\text{dB}$)
其他	1.00 ($\pm 0\text{dB}$)		

根据表4，GB/T 42758中规定的扬声器配置下的位置相关加权系数见表5。

表5 GB/T 42758中规定的扬声器配置下的位置相关加权系数

扬声器标签	加权系数		扬声器配置									
			A	B	C	D	E	F	G	H	I	J
			0+2+0	0+5+0	2+5+0	4+5+0	4+5+1	3+7+0	4+9+0	9+10+3	0+7+0	4+7+0
M+000	1.00	(±0.0dB)		X	X	X	X	X	X	X	X	X
M+SC	1.00	(±0.0dB)							X			
M-SC	1.00	(±0.0dB)							X			
M+030	1.00	(±0.0dB)	X	X	X	X	X	X	X	X	X	X
M-030	1.00	(±0.0dB)	X	X	X	X	X	X	X	X	X	X
M+060	1.41	(+1.5dB)								X		
M-060	1.41	(+1.5dB)								X		
M+090	1.41	(+1.5dB)						X	X	X	X	X
M-090	1.41	(+1.5dB)						X	X	X	X	X
M+110	1.41	(+1.5dB)		X	X	X	X					
M-110	1.41	(+1.5dB)		X	X	X	X					
M+135	1.00	(±0.0dB)						X	X	X	X	X
M-135	1.00	(±0.0dB)						X	X	X	X	X
M+180	1.00	(±0.0dB)								X		
U+000	1.00	(±0.0dB)								X		
U+030	1.00	(±0.0dB)			X	X	X					
U-030	1.00	(±0.0dB)			X	X	X					
U+045	1.00	(±0.0dB)						X	X	X		X
U-045	1.00	(±0.0dB)						X	X	X		X
U+090	1.00	(±0.0dB)								X		
U-090	1.00	(±0.0dB)								X		
U+110	1.00	(±0.0dB)				X	X					
U-110	1.00	(±0.0dB)				X	X					
U+135	1.00	(±0.0dB)							X	X		X
U-135	1.00	(±0.0dB)							X	X		X
U+180	1.00	(±0.0dB)						X		X		
T+000	1.00	(±0.0dB)								X		
B+000	1.00	(±0.0dB)					X			X		
B+045	1.00	(±0.0dB)								X		

表5（续）

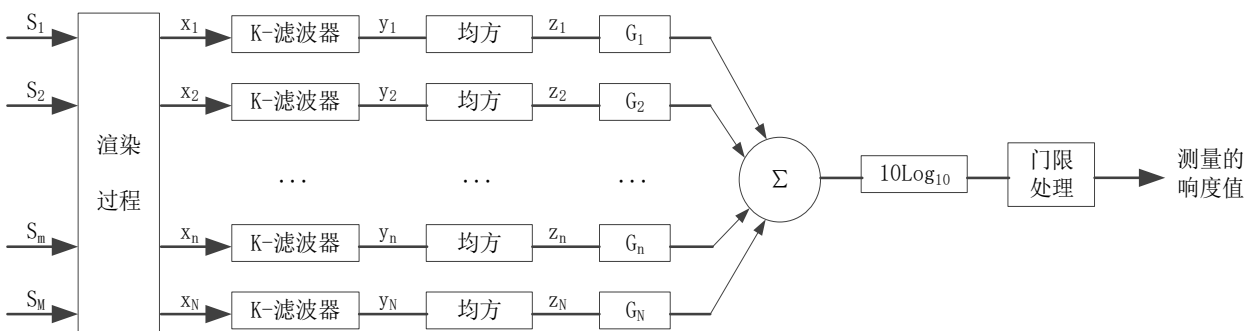
扬声器标签	加权系数		扬声器配置									
			A	B	C	D	E	F	G	H	I	J
			0+2+0	0+5+0	2+5+0	4+5+0	4+5+1	3+7+0	4+9+0	9+10+3	0+7+0	4+7+0
B-045	1.00	(±0.0dB)								X		
注：X表示扬声器配置A~J中包含该扬声器标签。												

7 基于对象的音频信号或基于声道与对象混合的音频信号的响度测量算法

本章规定了基于对象的音频信号或基于声道与对象混合的音频信号（包括先进声音系统的此类信号）的客观响度测量算法。

测量响度前，宜先将基于对象的音频信号或基于声道与对象混合的音频信号渲染到特定的扬声器配置。对于先进声音系统，宜使用GB/T 42758规定的扬声器配置。对渲染输出的音频信号进行客观响度测量时，采用第5章和第6章规定的响度测量算法。整个响度测量过程的简化示意图见图6。

客观响度可能因渲染条件的不同而有所差异。因此，测量中使用的扬声器配置和渲染算法宜予以记录。因渲染条件的不同而带来的响度测量差异情况见附录C。



注：S₁, ..., S_M是渲染前的输入信号；x₁, ..., x_N是渲染后的输出信号。

图6 基于对象的音频信号或基于声道与对象混合的音频信号的客观响度测量算法简化框图

8 真峰值电平测量算法

8.1 通则

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测量算法，该算法假定采样率为48kHz。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），该数值可能高于48kHz时域采样中的最大样本值。

真峰值电平测量算法的处理步骤如下：

- a) 衰减：12.04dB 的衰减；
- b) 4 倍过采样；
- c) 低通滤波；
- d) 取绝对值；
- e) 转换为 dB TP。

8.2 框图

信号真峰值电平测量算法框图见图7。

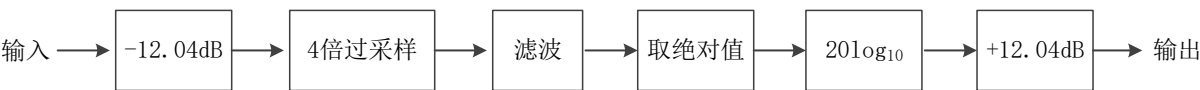


图7 信号真峰值电平测量算法框图

8.3 详细说明

对输入信号施加12.04dB的衰减（对应2bit右移）的目的是为可能采用整型运算的后续信号处理提供动态余量。若采用浮点运算，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，提高采样率后的信号能更准确地呈现音频样本所表征的实际波形。优先采用更高采样率和过采样倍数。如果输入信号本身具有较高采样率，则可相应降低过采样倍数（例如，对于采样率为96kHz的输入信号，2倍过采样即可）。

一组可满足要求的滤波器系数（适用于48阶、4相位、FIR插值滤波器）见表6。

表6 过采样滤波器系数

相位 0	相位 1	相位 2	相位 3
0.0017089843750	-0.0291748046875	-0.0189208984375	-0.0083007812500
0.0109863281250	0.0292968750000	0.0330810546875	0.0148925781250
-0.0196533203125	-0.0517578125000	-0.0582275390625	-0.0266113281250
0.0332031250000	0.0891113281250	0.1015625000000	0.0476074218750
-0.0594482421875	-0.1665039062500	-0.2003173828125	-0.1022949218750
0.1373291015625	0.4650878906250	0.7797851562500	0.9721679687500
0.9721679687500	0.7797851562500	0.4650878906250	0.1373291015625
-0.1022949218750	-0.2003173828125	-0.1665039062500	-0.0594482421875
0.0476074218750	0.1015625000000	0.0891113281250	0.0332031250000
-0.0266113281250	-0.0582275390625	-0.0517578125000	-0.0196533203125
0.0148925781250	0.0330810546875	0.0292968750000	0.0109863281250
-0.0083007812500	-0.0189208984375	-0.0291748046875	0.0017089843750

对信号样本取绝对值是将负值样本取反；至此，信号为单极性，负值被相同幅度的正值所取代。

经过四个步骤（衰减、过采样、滤波和取绝对值）处理后，得到的数值与原始样本值处于相同数域（例如24位整数）。此后需对初始的12.04dB衰减进行补偿，从而使处理过程的整体增益归一化。

若将衰减后的数值放大12.04dB（对应2bit左移），通常需要将该数值转换为一种能够表示超出原始格式满量程范围的数值格式。采用浮点格式执行计算步骤可满足这一要求。除了对结果进行放大外，另一种替代方案是对仪表刻度进行相应校准。

遵循本真峰值电平算法且过采样率至少为192kHz的真峰值表，将测量结果转换为对数标度后，应以“dB TP”为单位显示读数，“dB TP”表示相对于100%满刻度的真峰值测量分贝值。具体实现方式为：计算经过衰减、过采样、滤波和绝对值处理后的信号的“ $20\log_{10}$ ”对数值，随后加上12.04dB补偿值。数字音频信号准确峰值电平测量的技术要点见附录D。

附 录 A
(资料性)

本文件与 ITU-R BS. 1770-5:2023 结构编号对照一览表

表 A.1 给出了本文件与 ITU-R BS. 1770-5:2023 结构编号对照一览表。

表 A.1 本文件与 ITU-R BS. 1770-5:2023 结构编号对照情况

本文件结构编号	ITU-R BS. 1770-5:2023 结构编号
前言	—
—	范围
—	关键词
—	考虑到, 进一步考虑到, 建议, 进一步建议
1	—
2	—
3	—
4	—
5	附录 1
6	附录 3
7	附录 4
8	附录 2
附录 A	—
附录 B	附录 1 的附件 1
附录 C	附录 4 的附件 1
附录 D	附录 2 的附件 1
参考文献	参考文献

附录 B

（资料性）

多声道响度测量算法描述

B.1 概述

本附录描述了一种用于客观测量音频信号主观响度的算法。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，具有实现简单、成本低的特点。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。

在电视和无线电广播领域，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切换，这种节目内容的变化可能导致主观响度产生显著波动；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。因此，有必要对音频信号的响度进行测量和控制。

该多声道算法基于 $Leq(RLB)$ 算法的直接拓展，并保留了单声道 $Leq(RLB)$ 算法的极低计算复杂度。

B.2 $Leq(RLB)$ 算法设计

$Leq(RLB)$ 响度算法设计的核心理念是保持极简结构。 $Leq(RLB)$ 算法框图见图B.1，它由一个高通滤波器和一个计算功率平均值的处理单元组成。滤波器的输出信号被送入处理单元，处理单元对信号进行能量累加并计算其时间平均值。

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：所有处理过程均可通过计算需求极低的简易时域模块完成。

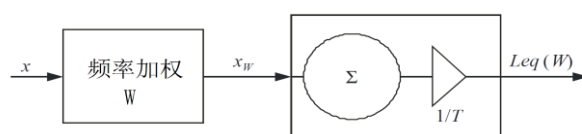


图 B.1 基于能量的简易响度测量框图

图B.1的 $Leq(RLB)$ 算法，实质上是等效声级（ Leq ）测量的一个频率加权形式， Leq 定义见公式（B.1）。

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

式中：

$Leq(W)$ ——频率加权等效电平，单位为分贝（dB）， W 代表频率加权，此处采用的是RLB加权；

T ——测量间隔；

x_W ——加权滤波器的输出信号；

x_{Ref} ——相应的基准量。

B.3 主观测试

B.3.1 主观测试安排

为了评估多声道响度测量算法的优劣，有必要开展正式主观测试以构建主观评价数据库。通过该数据库，可以验证各类响度测量算法预测主观测试结果的能力。该数据库收录了涵盖广泛节目类型的

单声道、立体声和多声道节目序列的主观响度评价结果。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、影视剧、体育赛事、新闻播报、音效和广告，其中还涉及多种语言的语音片段。

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B. 2。

参考信号始终以60dBA声压级播放，该声压级数值通常被认为是在家庭实际收看电视的典型值。

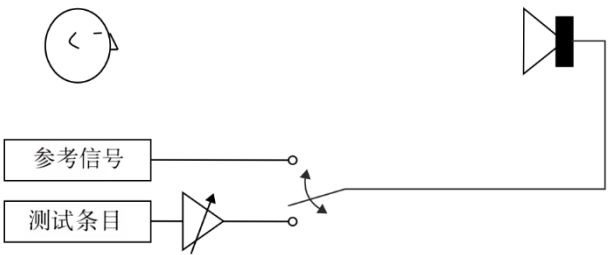


图 B. 2 主观测试方法示意图

测试采用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B. 3，测试条目的电平增益能够以0.25dB的步进调整。选择标记为“1”的按钮播放参考信号，该参考信号的电平保持固定不变。

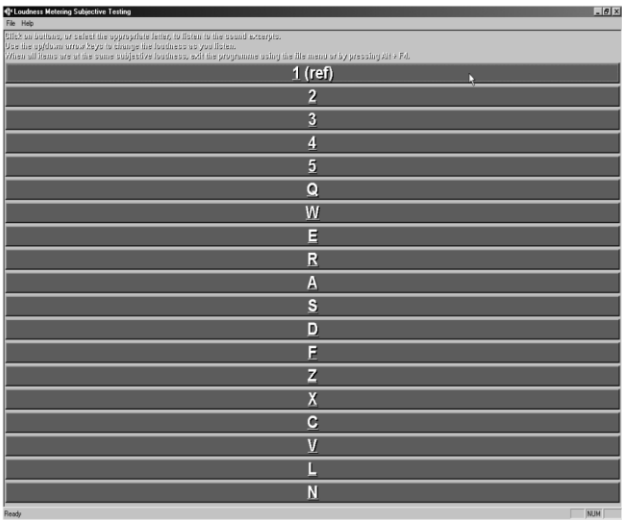


图 B. 3 主观测试系统的用户界面截屏图

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。鉴于多数测试条目混合了语音和其他声音（如音乐，背景，噪声等），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，提供给每位评价员的测试序列顺序随机，这意味着没有两位评价员面对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。

B.3.2 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。其中的测试序列和参考序列由无线电通信WP 6P SRG3成员组成的一个三人专门小组所挑选，参考信号由英语女性语音组成。序列由置于评价员正前方的单个扬声器播放。

沿用建立首个数据集时采用的主观测试方法和参考信号，20名评价员对96个新增单声道序列进行了响度评价，测试结果形成了第二个数据集。

第三个数据集包括144个音频序列的相对响度值（48个单声道条目、48个立体声和48个多声道条目）。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（双单声道模式）播放—这种设计旨在涵盖单声道信号的两种收听方式。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的规定，用于第三个数据集的扬声器配置见图B.4。

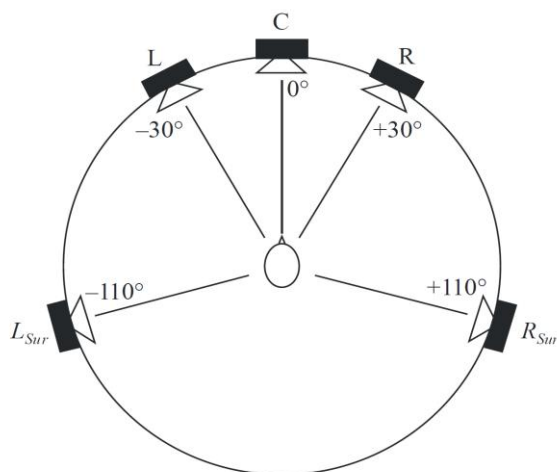


图 B.4 用于第三个数据集的扬声器配置

前两个数据集限于单声道序列，因此无需考虑声像因素。而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）和各种环境类型（例如干声与混响声）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要更为困难，而且，序列中包含的各种声像类型和环境类型加剧了测试难度。因此，研究者曾担心这些因素可能导致评价员不堪重负。但初步的测试结果表明该项测试是可控的，事实上，20名评价员确实提供了一致的结果。

B.4 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，ITU的研究结果表明该算法成功地应用于单声道。而多声道响度算法的设计需应对若干新的挑战，其中之一便是要求它必须同样适用于单声道、双单声道和立体声信号。

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

在多声道响度算法中，各独立音频声道的响度均由单声道 $Leq(RLB)$ 算法分别测量，多声道响度算法框图见图B. 5。在采用 $Leq(RLB)$ 之前对各个声道进行预滤波。

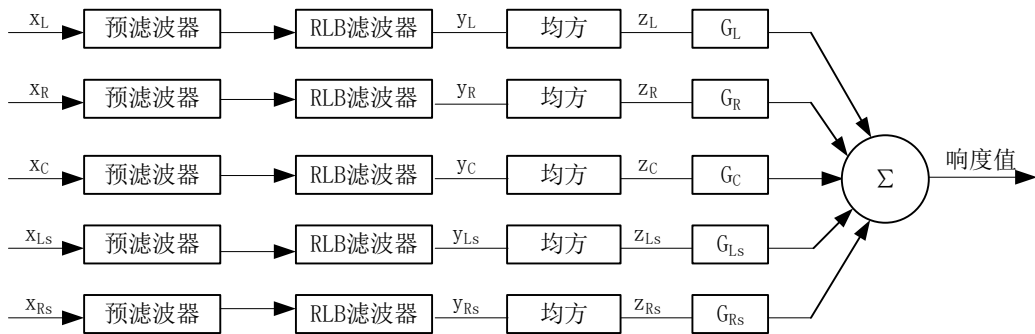


图 B. 5 多声道响度算法框图

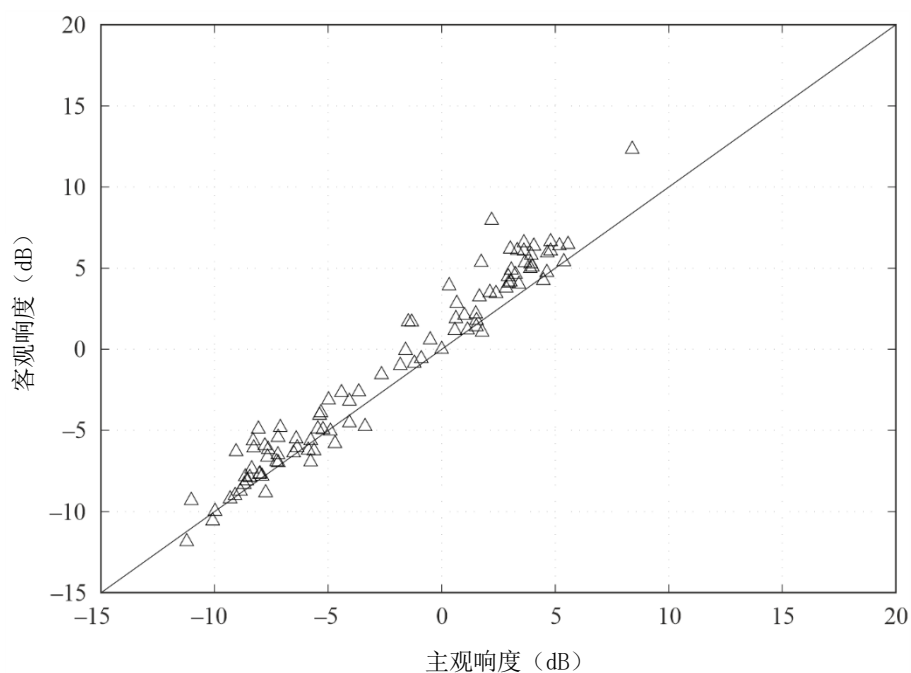
引入预滤波器的目的是考虑人体头部对入射声波的声学效果。这里，头部被模型化为一个刚性的球体。所有声道采用相同的预滤波器。根据信号到达人耳的角度对预滤波产生的各路响度值进行加权 (G_i)，然后求和（在线性域），最终获得综合响度测量值。这种加权机制是为了反映这样一个事实：从收听者后方到达的声音比从前方到达的声音在感觉上响度更高。图B. 5中“预滤波器”与“RLB滤波器”的组合即构成K加权处理。

该多声道响度算法的核心优势在于其简洁性，整个算法完全由基础信号处理单元组成，可在低成本硬件上通过时域处理轻松实现。另一个关键优势是其可扩展性，由于各个声道采用完全相同的处理流程，所以很容易实现一种能够容纳1~ M 任意声道数的响度计。此外，由于各个声道的贡献值是在响度层面而非信号层面进行叠加，算法不依赖于声道之间的相位或相关性，这使得本文件算法更具通用性和鲁棒性。

B. 5 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。

图B. 6~图B. 8分别给出了符合本文件算法的响度计对三个数据集客观测量结果与主观测试结果的对比性能。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文件响度计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。

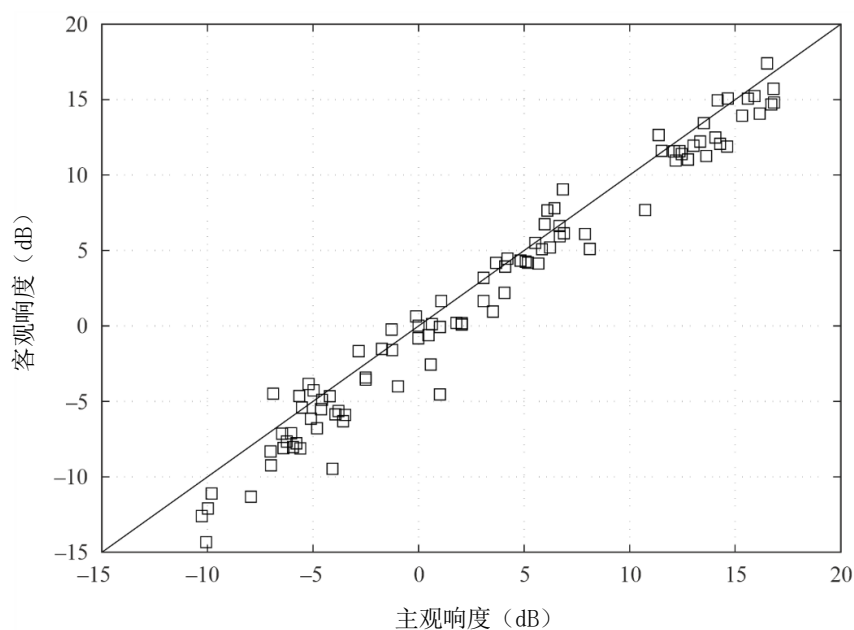


△ 第一个数据集

图 B. 6 第一个（单声道）数据集客观测量与主观测试结果的对比 ($r = 0.979$)

本文件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，主观响度评分与客观响度测量结果之间的相关系数达到 $r=0.979$ ，见图B. 6。

对于第二个数据集，主观响度评分与客观响度测量结果之间的相关性也非常好（ $r=0.985$ ），见图B. 7。该数据集中约半数序列为音乐内容。



□ 第二个数据集

图 B. 7 第二个（单声道）数据集客观测量与主观测试结果的对比 ($r=0.985$)

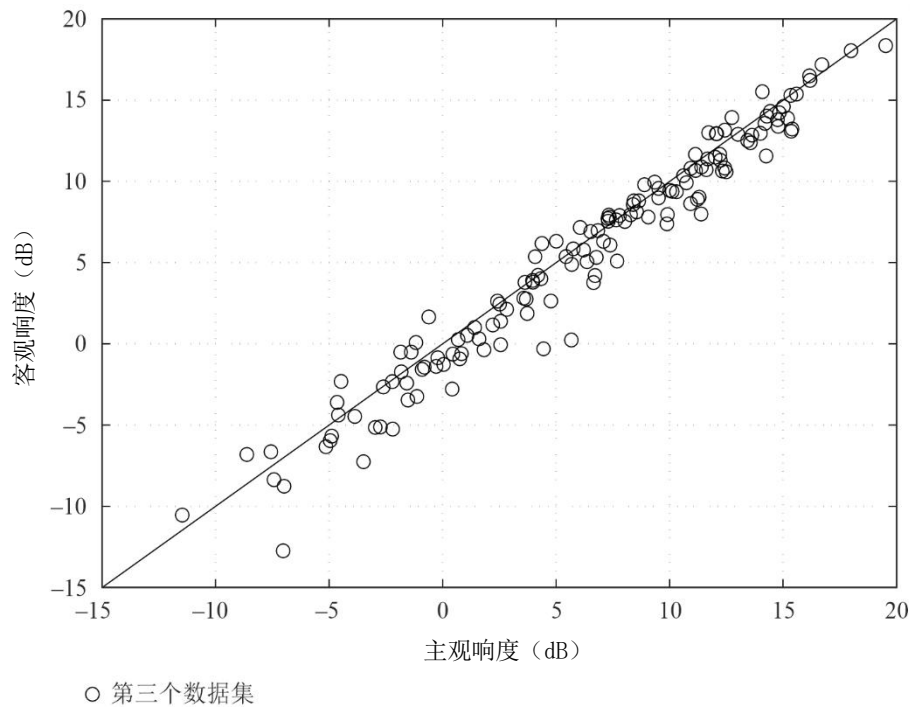


图 B.8 第三个（单声道、双单声道、立体声和多声道）数据集客观测量
与主观测试结果的对比（ $r = 0.980$ ）

第三个数据集的结果见图B.8，该数据集包括单声道、双单声道、立体声和多声道信号。图B.8和图B.9中呈现的多声道结果对应的是本文件算法在环绕声道的加权值设定为4dB（原始方案）时的表现，而非最终标准确定的1.5dB。经验证，加权值从4.0dB调整为1.5dB对结果未产生显著影响。该算法同样表现出色，主客观相关系数达 $r=0.980$ 。

为全面评估算法性能，有必要考察其在整个主观数据库（包括全部336段音频序列）中的表现，为此，图B.9整合了三个数据集的测试结果。该算法在整个数据集上主客观吻合性非常好，总相关系数达到 $r=0.977$ 。

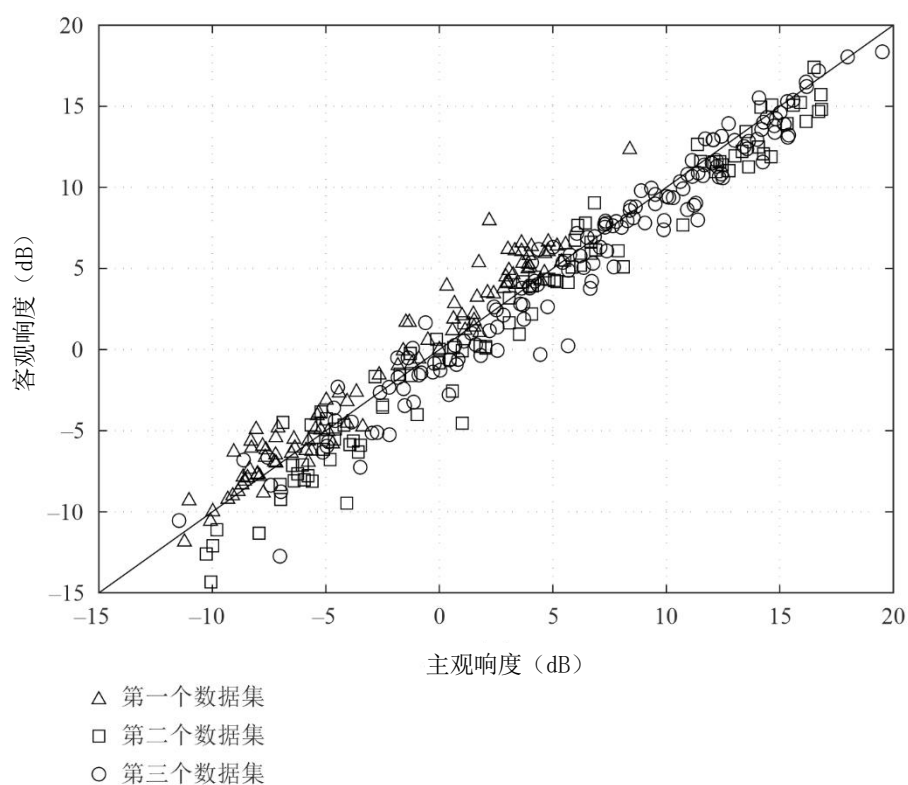


图 B.9 全部三个数据集客观测量与主观测试结果的对比 ($r=0.977$)

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。该主观数据库涵盖了广泛的节目类型, 包括音乐、影视剧、体育赛事、新闻播报、音效和广告, 其中还包含多种语言的语音片段。测试结果证实了, 符合本文件的响度计适用于单声道、双单声道、立体声以及多声道信号。

附录 C
(资料性)

基于对象的音频信号或基于声道与对象混合的音频信号的响度测量算法与渲染条件的关系

C.1 概述

本附录描述了基于声道与对象混合的音频信号的客观响度测量结果因渲染条件不同而产生的差异，以及主观测试的结果。

C.2 客观响度测量结果因渲染条件不同而产生的差异

共准备了8个完整的节目作为原始声音素材，这些节目包含多语言对白、增强对白、多视角对白和背景声等多种音频对象。原始素材基于极坐标系的位置元数据，并采用符合GB/T 42758规定的系统H（9+10+3）监听系统制作完成。将原始完整声音素材的客观响度统一调整为-24.0LKFS，调整条件为：使用ITU-R ADM渲染器（IAR），并依据ITU-R BS.2127规定的极坐标参数，在系统H扬声器配置下渲染后调整。

从原始声音素材中截取12s~16s片段，共生成31个测试素材。根据ITU-R BS.2127规定的转换算法，制作了这些片段的笛卡尔坐标系版本。每个测试素材使用5种渲染算法（极坐标系IAR、笛卡尔坐标系IAR、渲染器A、默认增益渲染器B和无增益调整渲染器B，渲染到6种扬声器配置（系统A（0+2+0）、系统B（0+5+0）、系统I（0+7+0）、系统D（4+5+0）、系统J（4+7+0）和系统H（9+10+3））。

注：渲染器B不支持系统H（9+10+3）配置。对于系统J（4+7+0）配置，渲染器B在默认增益和无增益调整模式下表现相同。

表C.1显示了客观响度测量结果因渲染条件不同而产生的差异。当采用单一渲染算法时，因扬声器配置不同导致的客观响度差异最大值可达2.1LU~4.1LU。另一方面，当采用单一扬声器配置时，因渲染算法不同导致的客观响度差异最大值可达0.3LU（9+10+3系统）~5.9LU（0+2+0立体声系统）。当扬声器数量少于音频信号数量时，客观响度会受到渲染算法的显著影响。因此，通常记录客观响度测量所用渲染条件。

表C.1 客观响度测量结果因渲染条件不同而产生的差异

渲染算法	最大值（绝对值）/平均值（绝对值）	扬声器配置	最大值（绝对值）/平均值（绝对值）	
			不包括默认增益渲染器B	不包括无增益调整渲染器B
极坐标系IAR	3.1LU/1.1LU	0+2+0	5.9LU/1.6LU	3.6LU/1.9LU
笛卡尔坐标系IAR	3.1LU/1.1LU	0+5+0	3.9LU/1.0LU	2.9LU/1.7LU
渲染器A	2.1LU/0.6LU	0+7+0	2.3LU/0.7LU	1.7LU/0.9LU
默认增益渲染器B	4.1LU/1.9LU	4+5+0	1.9LU/0.6LU	1.7LU/0.9LU
无增益调整渲染器B	3.8LU/1.0LU	4+7+0	1.7LU/0.5LU	1.7LU/0.5LU
—	—	9+10+3	0.3LU/0.0LU	0.3LU/0.0LU
注：扬声器配置9+10+3的结果不包括默认增益渲染器B和无增益调整渲染器B。				

C.3 主观测试

C.3.1 测试方法

主观测试的设备和过程基于前期的主观测试方法（见附录B）。

C.3.2 测试素材的影响

为研究测试素材的影响，测试1采用了31个测试素材，通过极坐标系IAR渲染到3种扬声器配置（0+2+0、0+5+0和9+10+3）。图C.1显示了针对两组初始响度值不同的31个测试素材，本文件规定的响度算法的性能。与基于声道音频信号的结果相比，即使在响度测量中使用经过渲染处理的音频信号，结果也没有显著差异。

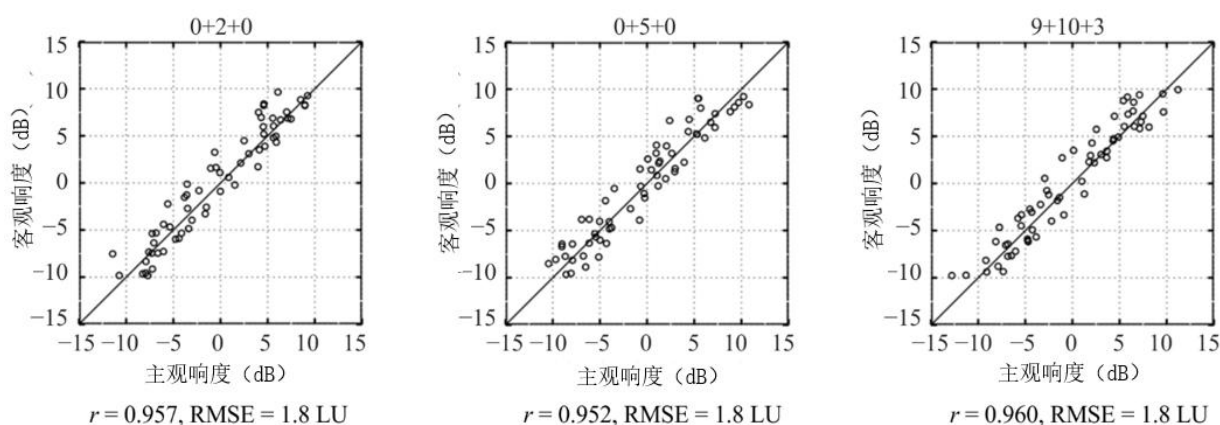
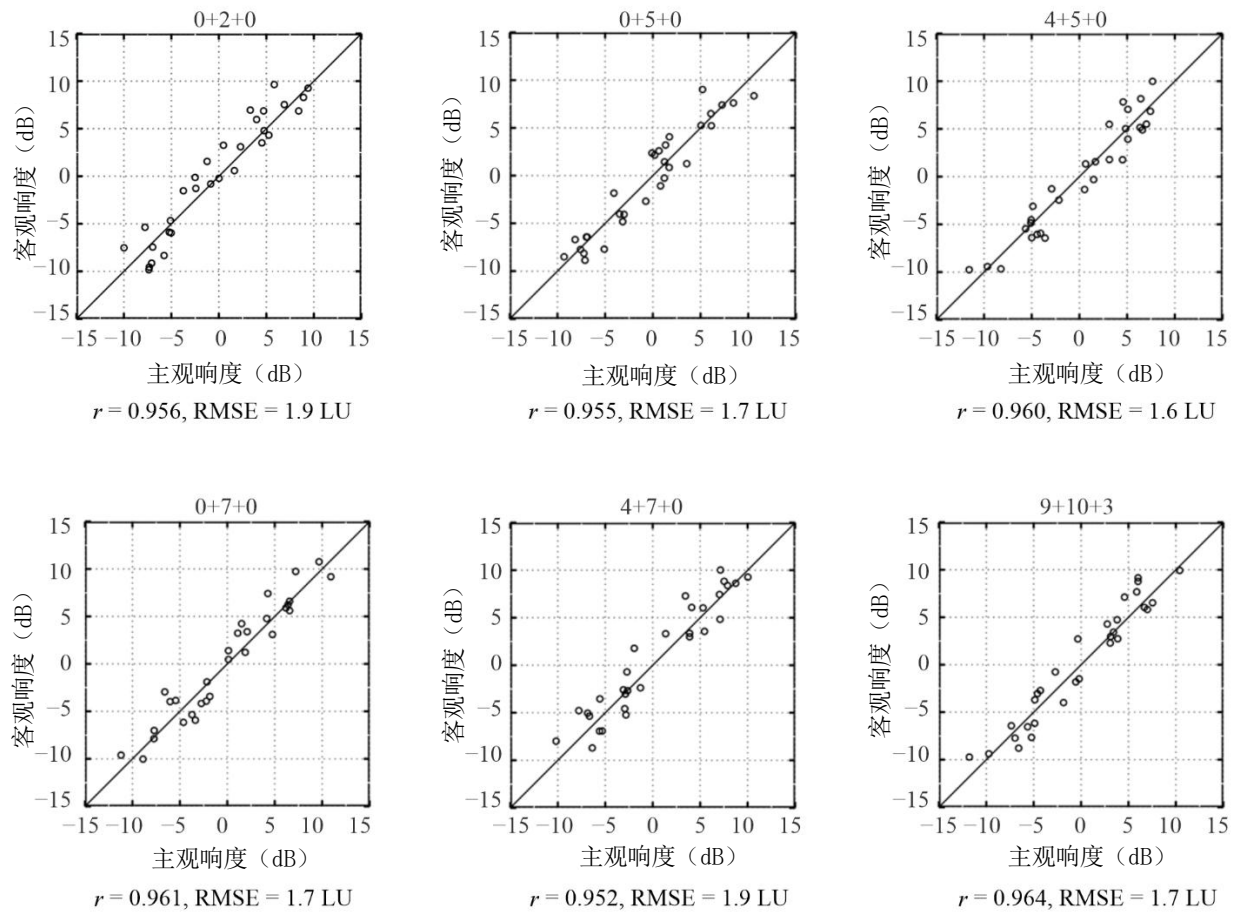


图 C.1 31 个测试素材在三种扬声器配置下的结果

C.3.3 扬声器配置的影响

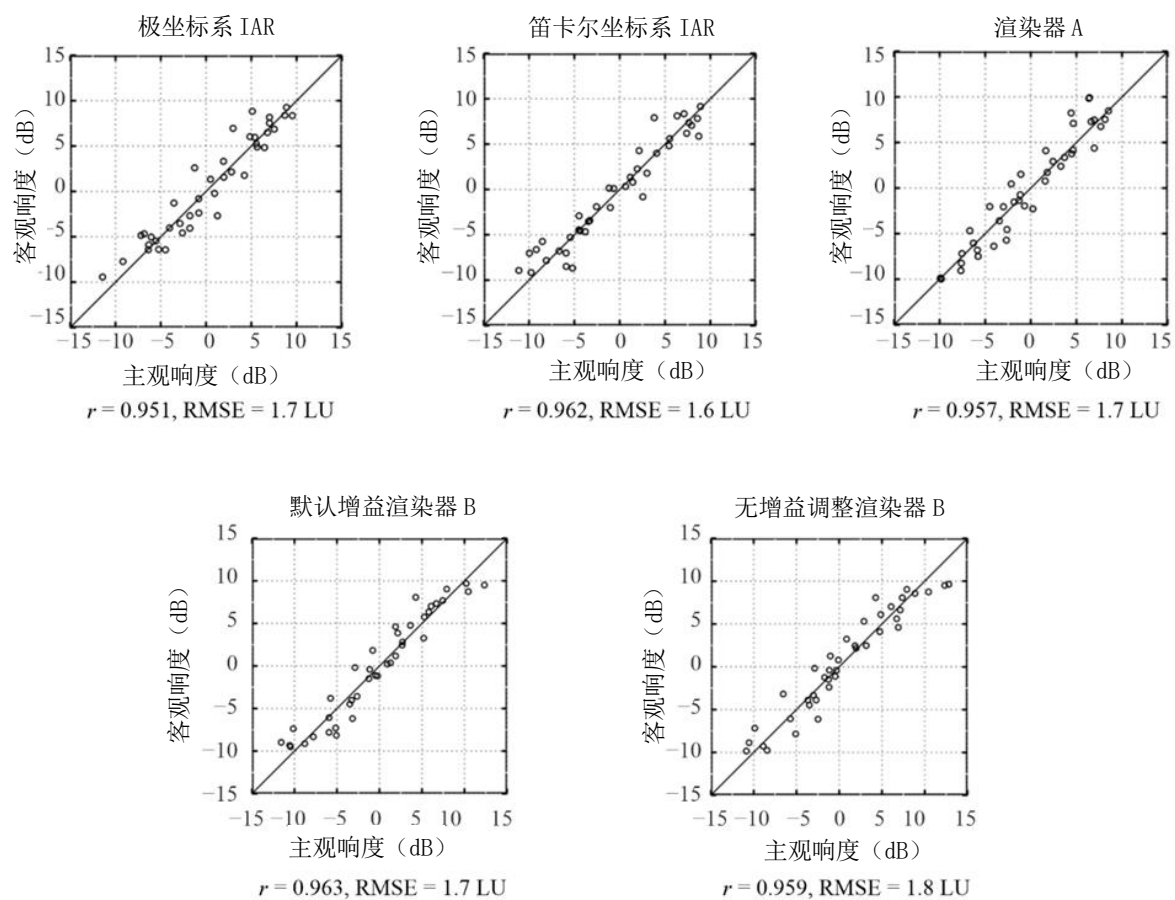
为研究扬声器配置的影响，测试2采用了15个测试素材，通过极坐标系IAR渲染到6种扬声器配置（0+2+0、0+5+0、4+5+0、0+7+0、4+7+0和9+10+3）。图C.2显示了本文件规定的响度算法在6种扬声器配置下的性能。通过比较不同扬声器配置间的结果可知，这些配置对结果未产生显著影响。另一项使用4种扬声器配置（0+5+0、4+5+0、0+7+0和4+7+0）的测试也显示出相同的趋势。



图C.2 6种扬声器配置的结果

C.3.4 渲染算法的影响

为研究渲染算法的影响，测试3采用了5个测试素材，通过5种渲染算法（极坐标系IAR、笛卡尔坐标系IAR、渲染器A，默认增益渲染器B和无增益调整渲染器B）渲染到4种扬声器配置（0+2+0、0+5+0、4+5+0和4+7+0）。图C.3显示了本文件规定的响度算法在5种渲染算法下的性能。通过比较不同渲染算法之间的结果可知，尽管客观响度值因渲染条件不同而存在差异，但不同渲染算法对响度测量结果并未产生显著影响。



图C.3 5种渲染算法的结果

附录 D

(资料性)

数字音频信号准确峰值电平测量的技术要点

D.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“真峰值”。

采样峰值表的工作原理通常是：将每个输入采样点的绝对值（整流后的值）与仪表的当前读数进行比较，如果新采样值更大，则替换当前读数；否则将当前读数乘以一个略小于 1 的常数，以实现指数衰减。这种峰值表因其实现简单而应用普遍，但它们并不能始终准确反映音频信号的真峰值。

因此，在节目峰值准确测量至关重要的应用场合，采用采样峰值表易产生测量偏差与应用隐患。当前大多数数字峰值仪表均属于采样峰值类型，此特性通常不为操作人员所知晓。

问题的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非恰巧位于采样时刻，造成采样峰值表错失捕捉信号的真正峰值。

一些常见的采样峰值表应用异常如下。

- 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往往差异显著。同样地，重复播放一段数字录音，在测量前进行采样率转换，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。
- 意外过载。采样信号中即使没有达到或接近数字满刻度电平的采样值，仍可能包含过载成分，因此，采样峰值表的过载指示并不可靠。这些过载可能会在后续处理中（例如在特定数模转换器或采样率转换过程中）导致削波，即便此前采样峰值表并未记录到这些过载（甚至在该节点监听时也无法察觉）。
- 纯音信号的读数偏低和跳动。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。

D.2 对测量准确性的影响

一般而言，输入信号频率越高，采样峰值表的潜在误差越大。

对于连续纯音信号，采样峰值表的读数偏低现象可通过理论分析予以证实。例如，频率为四分之一采样率且处于不利相位条件的纯音，采样峰值表读数比信号的真实峰值低 3dB；当频率升至二分之一采样率时，读数偏差在理论上可趋于无穷大；然而，绝大多数的数字音频信号在该频点所含能量极低，这既源于抗混叠滤波器在该频点处的大幅抑制，也符合真实声音信号高频能量自然衰减的物理特性。

连续单音信号频率不接近低整数分之一倍采样率时，由于差拍频率（ $n \cdot f_{\text{tone}}$ 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。

对于单个瞬态信号，上述机制无法掩盖其读数偏低的问题，因此，瞬态信号的频率越高，潜在的读数偏低幅度就越大。在“真实”声音中，富含高频成分的瞬态信号十分常见，其读数偏差通常可达数个分贝。

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，因此在原始采样率较高时，采样峰值表的低读现象会有所缓解。

D.3 解决方案

为了测量采样信号的真正峰值，有必要对信号进行“过采样”（或“上采样”），从本质上重建现有采样点之间的原始信号，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) 可接受的最大读数偏低误差；
- b) 待测最高频率与采样率的比（最大“归一化频率”）。

如果已知这些参数，就可以通过直观的“方格纸”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）。问题简化为仅考虑在最大归一化频率正弦曲线波峰两侧对称分布的一对过采样值产生的读数偏低，这是“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 $20\log_{10}\left(\cos\left(\frac{2\pi \cdot f_{norm} \cdot f_s}{n \cdot f_s \cdot 2}\right)\right)$ ，（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）为 $20\log_{10}\left(\cos\left(\frac{\pi \cdot f_{norm}}{n}\right)\right)$ 。

由以上公式可得到表 D.1，表中涵盖了主要的应用场景。

表 D.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 $f_{norm}=0.45$	读数偏低（dB）最大值 $f_{norm}=0.5$
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

D.4 真峰值表的实现

为实现过采样，在原始采样点之间插入零值样本，以生成符合目标过采样率的数据流，随后施加低通“内插”滤波器抑制高于目标最大归一化频率的成分。此时再对过采样信号运用采样峰值算法，将得到满足最大低读误差要求的真峰值仪表。

从过采样器的实现角度而言，其通常采用对称FIR滤波器作为低通滤波器。当这类滤波器用于高质量音频时，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

极大的阻带衰减和陡峭的过渡带，需要计算大量的“抽头”系数。同时，为了保持动态范围且使失真最小化，还必须保留较长的字长。

然而，真峰值仪表中过采样器的输出仅用于驱动读数显示或柱状图指示而非听觉重现，其精度要求自然不同。只要通带纹波与阻带杂散分量叠加造成的读数误差不超出目标容限即可。这使得所需抽头数量得以大幅缩减，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。同样，除非需要低振幅下的精确数值输出，只需足够的字长保证柱状图最低刻度处的目标精度即可。

因此，适用于多声道的过采样器完全能在普通低成本DSP或FPGA中轻松实现，甚至可采用更简化的处理器架构。另一方面，过采样仪表也可采用专为数模转换器设计的高精度过采样芯片实现，虽然这种方案对硅材料与功耗的利用效率较低，但此类器件成本低廉且易于获取。

确定一种特定仪表规格所需的抽头数量及抽头系数的最简方法，是使用Remez或Meteor等递归FIR滤波器设计程序。

峰值仪表通常还需具备消除输入直流分量的功能，因为传统音频仪表均配备隔直功能。反之，若以消除过载为目的监测真实峰值信号，则必须保留并测量直流分量。必要时，通过在仪表输入端植入低阶IIR高通滤波器，即可以较低计算量实现直流分量滤除。

某些应用场景要求在施加某种加权滤波器后测量峰值信号幅度，以强化特定频段的影响效果。具体实现方案取决于加权滤波器的特性。

参 考 文 献

- [1] GY/T 285 数字音频设备音频特性测量方法
 - [2] ITU-R BS.775-1 Multichannel stereophonic sound system with and without accompanying picture
 - [3] ITU-R BS.2127 Audio Definition Model renderer for advanced sound systems
 - [4] ITU-R BS.2217 Compliance material for Recommendation ITU-R BS.1770
 - [5] Sizle A., Sporer T., Liebetrau J., and Oode S. (2015), "Progress in Standardization of 3D-Audio Loudness in ITU-R BS.1770," Proceedings of the 3rd International Conference on Spatial Audio, paper 013.
 - [6] Norcross S., Nanda S., Cohen Z. (2016), "ITU-R BS.1770 Based Loudness for Immersive Audio", the 140th Convention of the Audio Engineering Society, Convention Paper 9500.
 - [7] Iwasaki T., Kubo H., and Oode S. (2022), "Loudness of Next-Generation Audio contents depending on rendering conditions," the 154th Convention of the Audio Engineering Society, Convention Express Paper 77.
 - [8] Benjamin E. (2004) Preferred Listening Levels and Acceptance Windows for Dialog Reproduction in the Domestic Environment, 117th Convention of the Audio Engineering Society, San Francisco, Preprint 6233.
 - [9] Soulodre G.A. (2004) Evaluation of Objective Loudness Meters, 116th Convention of the Audio Engineering Society, Berlin, Preprint 6161.
-