

# 扩声系统语言传输指数 STIPA 论述

· 技术交流 ·

咸爱清<sup>1</sup>, 唐笋翀<sup>2</sup>

(1. 恩缇艾音频设备技术(苏州)有限公司, 江苏 苏州 215168;  
2. 苏州市建设工程质量检测中心有限公司, 江苏 苏州 215129)

**【摘要】**语音的清晰度程度是音质评价的重要指标, 其可以反映厅堂或扩声系统的语音传输质量, 是评价语言传输系统工作质量的重要参考量。可以通过实验的方法来测量, 也可以借助传输系统的物理特性来预测。介绍了语音清晰度指标的由来、相关标准与重要性、测试方法、测试原理以及测量结果的评价。

**【关键词】**语言清晰度; 语音传输指数; STIPA

**【中图分类号】** TB54      **【文献标志码】** A

Discourse Upon Speech Transmission Index for Public Address Systems(STIPA)

XIAN Aiqing<sup>1</sup>, TANG Sunchong<sup>2</sup>

(1. NTiChina, Suzhou Jiangsu 215168, China; 2. Suzhou Building Construction Quality Inspection And Test Center, Suzhou Jiangsu 215129, China)

**【Abstract】** The speech articulation is an important index of sound quality evaluation, which represents the quality of voice transmission of the hall or sound reinforcement system. It can be measured by experimental methods and it can be predicted by the physical properties of the transmission system. The origin of language definition index, related standards and the importance of evaluation, testing method, testing principle and the measurement results are introduced.

**【Key words】** speech articulation; speech transmission index; STIPA

有二者紧密结合才能创造出扩声系统的精品。

## 2 相关标准

国际上, NFPA 72, BS 5839 - 8, VDE 0828 - 1, ISO 7240 - 16/ - 19 等标准要求验证电声音响系统紧急用途: 根据实际情况确定最低水平的语言清晰度, 以防遇到紧急情况。因而, 严格监管下的语言清晰度测量不能仅依靠主观测量, 而必须是采用经过验证的、多少有些复杂的方法, 这些测量方法在 IEC 60268 - 16 中已经被标准化。IEC 60268 - 16 标准中详细介绍了测量语言清晰度的许多方法。相较于 2003 年的版本, 2011 年更新了以下内容:

### 2.1 增加从“A +”到“U”各级别评估

用字母的方式对 STI 值分级, 并介绍了典型的专业应用, 见表 1。

### 2.2 心理声学效果(掩蔽效应)

人耳听觉系统在声压级较低或者较高的时候, 敏感度降低; 声压级较低时(例如 20 ~ 50 dB SPL), 由于听觉阈值作用而导致人耳敏感度降低。当声压级较高时(例如大于 80 dB SPL), 掩蔽效应导致人耳敏感度降低。

听觉掩蔽功能曲线随 IEC 60268 - 16 标准各版本的发展如下:

表 1 不同范围内 STI 的典型应用

级别	范围	典型应用
A +	> 0.76	录音室
A	0.72 ~ 0.76	剧院, 演讲厅, 议会, 法院
B	0.68 ~ 0.72	剧院, 演讲厅, 议会, 法院
C	0.64 ~ 0.68	电话会议, 剧院
D	0.60 ~ 0.64	教室, 音乐厅
E	0.56 ~ 0.60	音乐厅, 现代教堂
F	0.52 ~ 0.56	购物商场广播系统, 公共办公室, 教堂
G	0.48 ~ 0.52	购物商场广播系统, 公共办公室
H	0.44 ~ 0.48	所处声学环境较差的广播系统
I	0.40 ~ 0.44	所处声学环境很差的广播系统
J	0.36 ~ 0.40	不适用广播系统
U	< 0.36	不适用广播系统

- (1) ed4.0, 2011 年发布的版本, 含连续的听觉掩蔽功能 ( $> 80 \text{ dB}$  范围平滑衰减的曲线);  
(2) ed3.0, 2003 年发布的老版本, 含分段的听觉掩蔽功能 ( $> 80 \text{ dB}$  范围阶梯状衰减的曲线);  
(3) ed2.0, 1998 年发布的老版本, 含固定听觉屏蔽功能 ( $> 80 \text{ dB}$  范围平直不衰减曲线)。

具体情况如图 1 所示。

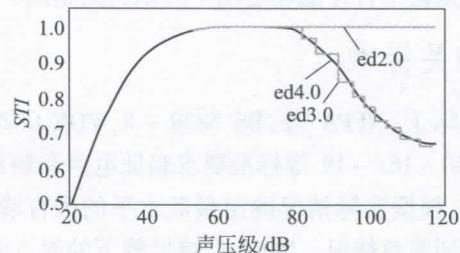


图 1 STI - 声压级趋势图

信号电平在 70 ~ 80 dB 范围内时可获得最佳的语言清晰度。更高的声压级也会使人耳进行自我保护, 这可以从清晰度指数降低上看出, 例如在 70 dB 时 STI 为 1, 在更高声压级时降到 0.7。

在国内, 从公共广播标准 GB 50526 到智能建筑领域标准 GB 50339, 从会议系统标准 GB 50635 到厅堂体育场标准 GB/T 28048, 很多涉及到音响扩声系统的标准规范都要求测量语言清晰度, 所有这些测量方法在 GB/T 4959 中已经被标准化(参照 IEC60268-16 ed3.0 2003)。

### 3 测试方法

尽管频率响应、混响时间、失真、信噪比和响度等

参量都已经被证实和语言清晰度结果相关, 但是即使将这些常规参量测量出来也不能完全确定语言清晰度。将其他因素考虑进去时, 例如声波指向性和环境条件, 又会碰到新的问题: 在不同位置上, 信息可被理解的程度如何? 因此还需要得到不同位置处的语言清晰度结果, 通常需要在各个位置进行测量<sup>[5]</sup>。

测量语言清晰度主要分为主观测量方法与客观评价方法。

#### 3.1 主观测量法

让一个经过训练的人读一定数目字词, 而那些具有代表性的位置上的人则分别写下他们认为已经明白的字词。然后统计分析他们记下的结果, 以百分比的形式表示被理解的字词正确率。具体的操作步骤需按照标准 PBwords, CVC, SRT 或者 GB-T 15508。尽管这种方法可以测试出最直接真实的语言清晰度结果, 但是要进行这样的测量需花费很长时间并且花费巨大, 而且在一些对人体有害的位置这种方法是不可行的。因此, 这些方法主要是用来验证的候补测量方法。

#### 3.2 客观评价方法

语言清晰度是指语言信号经系统传递后, 被听音人所正确了解的语言单位数与发音人发出的语言单位数的比值百分数。在国内外标准中, 与“语言清晰度”相关的术语或者定义有多种不同的表述: 清晰度指数 AI、语言可懂度 SI、语言传输指数 STI、房间声学语言传输指数 RASTI、扩声系统语言传输指数 STIPA 等。

##### 3.2.1 清晰度指数 AI(Articulation Index)

AI 是在给定的语言通道和噪声条件下, 表示对一个听音人获得语言可懂度可利用的正常语言信号的有效比率(部分), 它是一个加权的分数, 由测量或估计的语言频谱和在听音人耳边存在的噪声的有效掩蔽谱可计算 AI, 可参照 GB/T 15485—1995 标准或者 ANSI Std. S3.5—1969\1986\1997 等标准。

特点: 适用于评价诸如噪声干扰、有限带通等频域失真引起的语音清晰度恶化程度, 缺陷是不能反映混响、回声或削峰等时域或非线性失真的影响。

##### 3.2.2 辅音损失率 % AL<sub>cons</sub> (Articulation Loss of Consonants)

根据大量的实地测量和清晰度主观评价测试的结果, 荷兰声学家 Peutz 发现 % AL<sub>cons</sub> 与房间的声学特性相关。其概念是人们讲话发声, 其元音如 a, i, u,

e,o 比辅音 b,p,m,f 等能量大 6 dB 左右(即 4 倍), 而辅音是比较短促的声音, 一般是 20 ms 左右, 母音则为 90 ms 左右, 所以, 人们在讲话时往往因辅音听不清楚, 而影响语言清晰度。辅音损失率越少, 清晰度越高。

**缺陷:**首先由于其是根据实验结果的经验推导, 物理基础不够; 并且对于信噪比低于 25 dB 的情况估计不足, 不能充分反映噪声和混响联合作用时的影响; 其次, 没有考虑早期反射声、单次回响的作用, 且结果仅给出某个频带的清晰度水平, 未考虑参数的频率特性的影响。

### 3.3.3 语言传输指数 STI

从十九世纪四十年代开始, 贝尔实验室开始研究确定语言清晰度的测量技术: 比如 SII(语言清晰度指数)和各种形式的 STI(语言传递指数)等都是非常成熟的语言清晰度测量方法。这些测量方法考虑了很多对语言清晰度有影响的参量, 比如: 信噪比、心理声学效果(掩蔽效应)、声压级、环境噪声、混响时间  $RT_{60}$ 、反射、频率响应、失真等。

## 4 语言传输指数测量原理

STI 测量的基本思想是播放一个合成的测试信号来代替演讲者的声音。测量语言清晰度时需要获取并分析这个信号, 这相当于前边测量时听众的耳朵。大量的调查已经表明了语音特征的改变与语言清晰度之间的关系。这些研究结果被应用到了语言清晰度测量仪器内, 使之直接将语言清晰度以单一数值在 0(不知所云) 和 1(良好的可懂度) 之间显示。语言传输指数值在 0~1 之间, 按主观感觉分为 5 档, 如表 2 和图 2 所示。

表 2 语言传输指数值按照主观感觉分类

STIPA 值	0~0.30	0.30~0.45	0.45~0.60	0.60~0.75	0.75~1.00
等级	Bad(劣)	Poor(差)	Fair(中)	Good(良)	Excellent(优)

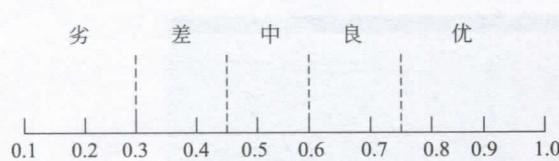


图 2 与可懂度有关的语言传输质量  
和 STI 指数的关系

### 4.1 语言的物理描述

连续的语言信号可以认为是一系列被称为“音位 phoneme”的语音碎片组成的。而“音位”又可以被认为是一个窄带噪声被发音器官的极低频运动所调制而形成, 在时域上表现为一个以某种窄带噪声为载波的调幅波。因而, 每一个音位都有自己的包络函数, 语言所传递的信息就包含在该调幅波包络之中。

#### 4.1.1 语音模型

首先, 测量语言清晰度需要一个语音信号的模型。例如, 语音可能被描述成各种不同的声音频率调制与叠加而成(例如振幅的变化), 如图 3 所示。

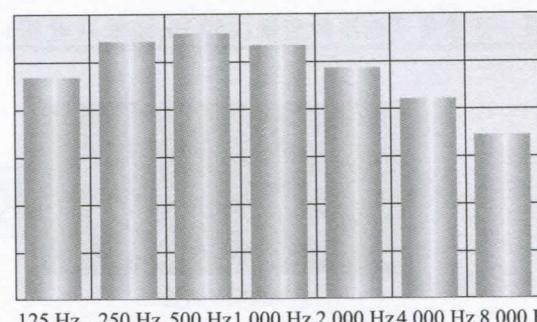


图 3 男性演讲者的平均倍频程频带频谱(截图)

#### 4.1.2 时间调制

由于说话者的“调制”导致每一个倍频程频带的信号电平的变动。图 4 显示语音信号在 250 Hz 频带的包络, 包络的外形由语音内容所决定。包络频率  $F$  约在 0.2~12.5 Hz 之间。

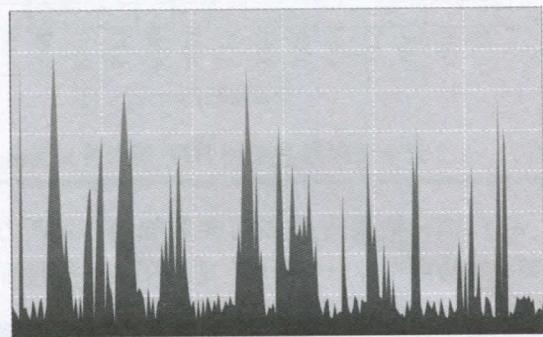


图 4 语音信号的包络(250 Hz 频带)(截图)

#### 4.1.3 频谱

典型的对男性语音在一定周期中的频率分析结果特性如图 3 所示。经过对频谱时间调制强度的研究表明, 说话者对语音频谱的调制频率范围在

0.1~24 Hz, 选择 0.63~12.5 Hz 范围内的系列典型调制频率即可充分代表所有调制特性, 如图 5 所示。

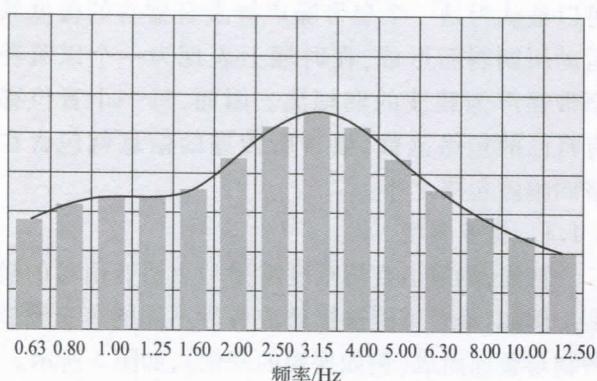


图 5 语音信号包络的频谱(250 Hz 频带)(截图)

#### 4.1.4 调制转移函数 MTF

高语言可懂度要求谱强度调制和总频谱可清晰地保存在听者的耳内。因此,三个核心的理解度测量方法 STI, RASTI 与 STIPA 基于 7 个频带内的 MTF (Modulation Transfer Function) 测量。调制传递函数 MTF 测量量化在个别频带被调制的信号维持的程度。

语音在传输过程中会受到诸如混响、反射、背景噪声等因素的影响,其可以直接反映到测量到的调制指数上,如图 6 所示。

图 7 显示了一个倍频程带的 MTF。这由测得的 1/3 倍频程调制频率导出,结果为 0.63~12.5 Hz 之间的 14 个频点。每个调制传输函数决定了相关倍频程带调制的好坏。



图 6 调制指数受影响示意图

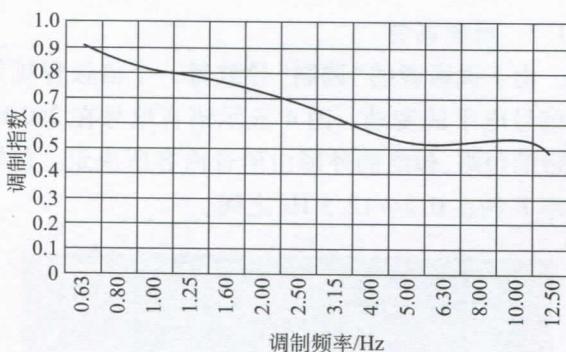


图 7 一个语音倍频程频带的 MTF 调制传递函数

基于 MTF 结果与更多的参数如声压级听力临界、频率响应或精神声学效应 (遮蔽效应) 使得可靠

地测量语言清晰度成为可能,相对的计算基于广泛且深入的主观方法评估。

完整的 MTF 测量需要  $14 \times 7 = 98$  个测量,完整的测量要花费约 15 min。因此,研发出各种不同的方法以减少测试时间。

#### 4.2 STI 的计算

STI 结果为基于完整的 98 套测量而得,如图 8 所示,由于经由这种途径需要相当长的测量周期,实际上不常采用。不过,STI 所代表的是测量语言清晰度最详细的方法,同时大多是用在其他方法在不利的测量环境条件下不能提供可靠的测量结果时采用。

(1) 14 种调制频率,7 种倍频带噪声作为测试信号。

Modulation Frequencies														
	0.63 Hz	0.8 Hz	1 Hz	1.25 Hz	1.6 Hz	2 Hz	2.5 Hz	3.15 Hz	4 Hz	5 Hz	6.3 Hz	8 Hz	10 Hz	12.5 Hz
Octave Bands	125 Hz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
250 Hz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
500 Hz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
1 kHz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2 kHz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
4 kHz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
8 kHz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

图 8 STI 调制指数(截图)

- (2) 计算 98 个 MTF 结果。
- (3) 类似 AI 计算中的加权平均最终得 STI  $[0,1]$ 。

(4) 根据房间脉冲响应可以直接计算 MTF。

实际上,STI 结果能经由计算测量到的脉冲响应 (MLSA) 得到,例如用一台基于计算机的系统,这种途径更快速,但它需要很丰富的经验,特别是系统设置上的表现是线性的,例如,必须不存在任何非线性处理或条件,包含压缩器与限制器等,那是非常罕见的处境。传声器和扬声器在量测过程中不能移动。因为手持式分析仪不需要安装,因此手持式分析仪量测不需支持 MLS 测试。

客观可测的物理参量 STI 能很好地反映和预测真实的语音清晰度水平,不论是信噪比、心理声学效果(掩蔽效应)、声压级、环境噪声、混响时间  $RT_{60}$ 、反射、频率响应、失真、带通限制、削峰等干扰因素,STI 与语音清晰度得分之间始终有着良好的一致性。

#### 4.3 语言传输指数 STI 的简化算法

测量广播系统的 STI 一直非常耗时。一个完整的测量必须执行  $14 \times 7 = 98$  个 MTF 参量的测量,要得到完整的 STI 数值可能需要 15 min,这在实际工程测量中并不实用。由于操作过于复杂并耗时,因此多年以来没有真正可用的测量系统问世。因此,发展出不同的方案以减少测试时间以及使得语言清

晰度测量得以在便携式仪表上实现成为业内共同的需求。

由 STI 法派生出 3 种简缩版:STITEL——适用于长途通信;RASTI——适用于不用电声设备的、仅由房间声学特性决定的清晰度测量;STIPA——适用于扩声系统,称为“扩声系统语言传输指数”。

随着 IEC 60268 - 16,2011,4.0 版标准的生效, RASTI 不再是语言清晰度测量的许可方法,这里笔者着重介绍下 STIPA 测量方法。

#### 4.4 STIPA

由于对安全上更多的注意,新技术方法与 RASTI 的缺陷触发了扬声器制造商 Bose 与 TNO 研究院研发了新的测量公共广播语言清晰度的方法,经过这些努力发展出 STIPA,这个指标可以用便携式仪表快速、精确地测试。如同 RASTI 一样,STIPA 实行很简单的步骤来计算出 MTF,但是 STIPA 以分析所有 7 个频带来决定 MTF,每一个频带通过两个调制频率进行调制,如图 9 所示。

IEC60268 - 16 标准中描述了 STIPA 测试方法,其 125 Hz 和 250 Hz 频带合并起来,深黑色区域标识是未进行调制的区域(即 1.6 Hz 与 8 Hz 调制频率未启用)。

NTi Audio - STIPA 方法(经过 TNO 认证)包含 7 个倍频程带和 14 个调制频率,其所得到的结果精度高于 IEC STIPA 方法,如图 10 所示。

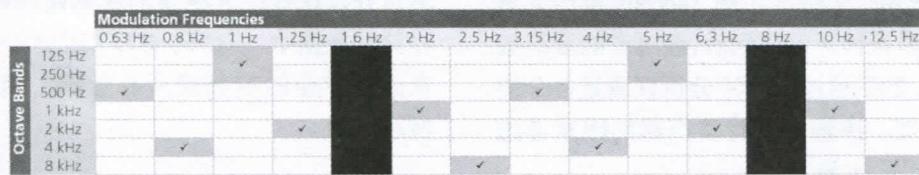


图 9 STIPA 调制指数(截图)

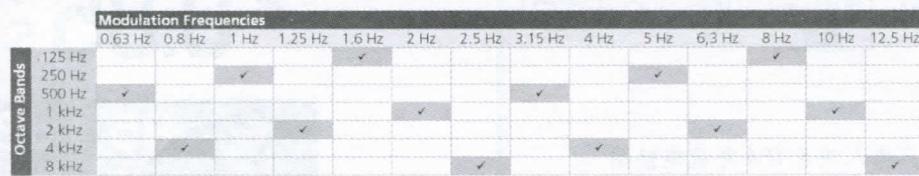


图 10 NTi Audio - STIPA 方法(截图)

假设没有严重的环境噪声脉冲出现与没有大量的非线性失真发生,STIPA 提供如 STI 一样精确的测量结果,但是如果在正常系统工作时间内出现环境噪

声脉冲,就得不到精确的 STIPA。

STIPA(扩声系统语言传输指数)借助便携仪器,支持更快更精确的测试。便携式 STIPA 测试仪,例

如 NTi Audio 的 XL2 音频和声学分析仪,可在 15 s 内测得语言清晰度,因此非常适合大面积区域的测量。

可以以 STIPA 的测量结果来表示清晰度  $Al_{cons}\% = 10^{(1-STI)/0.45}$ ,但是反过来因为不同的测量规范之间存在差异,因此基于  $Al_{cons}\%$  的测量来计算 STIPA 是不合理的。

## 5 语言传输指数 STIPA 测量以及结果评价

### 5.1 测量声源

要测量得到 STIPA,首先要获取 MTF 值,而 MTF 可以通过不同的方法得,最根本的方法就是通过比较传输信道的输入和输出的幅度调制而得出。为此可以采用语声信号、脉冲响应或者专门的测试信号(合成信号)。

(1) 语声信号能在实际环境(如现场演出)中确定 MTF,但不便于测量运算。

(2) 脉冲响应能用于确定 MTF 以及混响或回声对它的影响。但脉冲测试信号不能体现平均语声谱和声级分布。

(3) 可以通过选取适当的人工合成信号来测试从而获得 MTF,这就是 STI 法。

确定了测试信号的形式,还要决定通过哪种形式来产生信号,不管是国际还是国内的标准都介绍了两种形式的测量方法:声输入法与电输入法。

声输入法:使用一个标准声源 TalkBox 模仿人讲话。声源发的信号被传声器收到后,经过音响系统后经由扬声器放出,然后通过 STIPA 分析仪测量各个测量点的语言清晰度。声输入法将传声器的因素考虑在内,系统结构如图 11 所示。

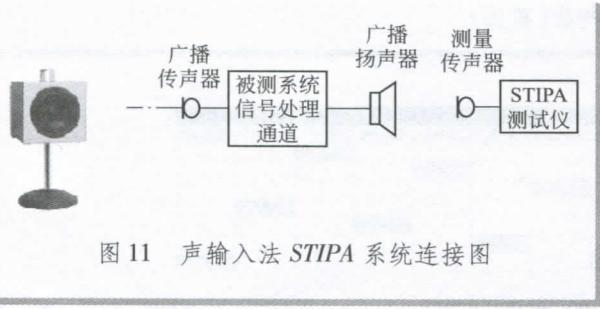


图 11 声输入法 STIPA 系统连接图

什么时候需要使用声输入法测试 STIPA 呢?

(1) 规定需要进行完整的终端至终端系统测量。这是任何情况下都最现实的系统验证。

(2) 电测试信号无法输入。

(3) 测试信号电平没有明确定义。

(4) 扬声器声学环境特性不可忽略且平坦。

(5) 传声器特性、灵敏度和频率响应都未知但需要考虑。

(6) 其他任何在现实环境下需要测试整个信号通道的情形。

电输入法:测试信号直接进入调音台,然后通过扬声器放出,此方法测试的是不含传声器的系统,而测试信号可以通过专业信号发生器(如 MR - PRO)产生,也可以通过专业 CD 播放器产生,当使用 CD 播放器时,请注意:只有高品质的 CD 播放器才能用来播放 STIPA 测试信号,因为只有有限的时间位移( $\pm 20$  ppm)确保稳定的 STIPA 测试结果。应禁用音调控制和振动保护。笔者推荐使用专业 CD 播放器。使用一个 1 kHz 测试信号验证 CD 播放器的时间位移:

(1) 将 NTi Audio 测试 CD 放入 CD 播放器,播放轨道 1,即 1 kHz 测试信号。

(2) 直接将 XL2 分析仪连接至音频输出端,使用 RMS/THD + N 功能测量信号频率。显示出的频率应该在 0.999 98 ~ 1.000 02 kHz 之间。

### 5.2 STIPA 测试

尽管 STIPA 测试的背景原理很复杂,XL2 音频和声学分析仪 STPA 测量功能的操作却非常简单。只要有基本声学知识的操作人员就能轻松进行。分析仪的存储功能可以支持在较大范围的厅堂内进行测量并将在各位置测得的数据储存在仪器内。进入存储的数据中,测得的 MTF(调制传递指数)方便专家进行专业的后处理所有数据。测量界面如图 12 所示<sup>[6]</sup>。



图 12 STIPA 测量界面显示(截图)

### 5.3 STIPA 报告生成工具

STIPA 报告生成工具依据 IEC 60268 - 16 和

VDE0833 标准生成测量报告。直接从仪器导入测量数据,包括测得的环境噪声,得到对应的语言清晰度 STI 值。报告范例如图 13 所示。

### STIPA 测量报告总结



图 13 STIPA 报告生成工具(截图)

## 5.4 STIPA 测量提示<sup>[7]</sup>

### 5.4.1 环境噪声

在测试进行中任何背景噪声都必须足够的静态。测量过程中出现的瞬间脉冲背景噪声,像语音,会造成严重的测量错误。STIPA 测量结果通常过高。在不使用测试信号时直接测量 STIPA 值可测出波动噪声。在有代表性的测试点进行这些测量,如果 STI 值太大(如  $STI > 0.2$ ),最终的测量结果就可能错误。此时,STIPA 测量就必须在没有这些噪声存在时进行。这种情况下就需要利用环境噪声修正功能。

对于情况复杂的场所(如某些区域人较少,其余地方人很多)应测量最嘈杂情况下的 STIPA 值。请参考当地法规(德国 VDE 0833-4),获得有关 STIPA 测量位置和次数的指导:

- (1)  $STI > 0.63$  单次测量即可。
- (2)  $STI < 0.63$  在该点连续进行三次测量。
  - ① 如果三次测量值最大偏差大于 0.03,那还需要再进行三次测量;
  - ② 如果三次测量值最大偏差大于 0.05,则需评估造成不稳定的原因并将其消除;
  - ③ 需要报告三次或六次测量的算数平均值。
  - (3)  $STI > 0.63$  确保了语言清晰度高于 0.5,置

信度达到 95%。

(4) STIPA 值较低可能由以下原因导致:

- ① 过多的混响、回声或反射;
- ② 较差的扬声器指向性或覆盖率;
- ③ 扬声器调节错误,如较低的信噪比。

### 5.4.2 一直有脉冲噪声怎么办

在一个 7 天  $\times$  24 小时期间都在运行的工厂或高速公路旁,脉冲噪声是一直存在的。因此不能正常进行 STIPA 的量测。在这种情况下,必须要在实验室中模拟现场的环境:

① 现实环境中的背景噪声必须被测量到,例如,使用 XL2 的 SLM 模式,在一个足够长的测量时间内求平均值。

② 在实验室中,必须模拟出在某个测量位置上具有一个和现场环境有相同频谱和声压级且没有脉冲噪声的弥散性声场。

③ 在实验室中,尽可能地模拟出现实环境中演讲者和听众所处的环境。

④ 然后就可以进行 STIPA 的测量了,不需要再进行后处理。

⑤ 如果含有自动增益控制(AGC)的系统在原始环境中难以测试,这个步骤需强制使用。

## 6 总结

本文介绍了语言清晰度技术的发展历程,以及相关测量方法以及注意事项,仅以此抛砖引玉。

### 参考文献

- [1] 国家技术监督局. GB/T 15485—1995 语言清晰度指数的计算方法[S]. 北京:中国标准出版社,1995.
- [2] 中华人民共和国国家质量监督检验检疫总局,中国国家标准化管理委员会. GB/T 4959—2011 厅堂扩声特性测量方法[S]. 北京:中国标准出版社,2012.
- [3] 刘春旺,程伟. 体育场语言清晰度的检测[J]. 智能建筑,2011,13(3):47-49.
- [4] 顾克明. 语言清晰度是多功能会场系统设计的重要声学指标[J]. 智能建筑,2011,10(2):44-46.
- [5] 王杰,钟恭良,张承云. STIPA 测试原理简介[J]. 电声技术,2008,32(5):11-13.
- [6] 顾克明,彭妙颜,周锡韬. 会场系统工程[M]. 北京:中国电力出版社,2013.
- [7] BERNO N. XL2 音频与声学分析仪使用手册[G]. (2009-03-06)[2014-05-20]. www.nti-audio.com/XL2.

[责任编辑] 闫雯雯

[收稿日期] 2014-05-21