



中华人民共和国国家标准

GB/T 31070.3—2021

楼宇对讲系统 第3部分：特定应用技术要求

Building intercom systems—Part 3: Technical requirements
for specific applications

2021-05-21 发布

2021-12-01 实施



国家市场监督管理总局 发布
国家标准化管理委员会

目 次

前言	III
引言	IV
1 范围	1
2 规范性引用文件	1
3 术语、定义和缩略语	1
3.1 术语和定义	1
3.2 缩略语	2
4 系统分类	3
4.1 按音频带宽分类	3
4.2 按应用场景分类	3
4.3 按建筑声环境功能区分类	3
5 功能要求	4
5.1 一般要求	4
5.2 基本功能	4
5.3 扩展功能	4
6 性能要求	6
6.1 音频特性	6
6.2 视频特性	8
6.3 电源电压适应性	8
6.4 环境适应性	8
6.5 安全性	9
6.6 电磁兼容性	9
6.7 标志、机械结构要求	9
7 试验方法	10
7.1 试验条件	10
7.2 功能检查	10
7.3 音频特性试验	10
7.4 安全性试验	16
7.5 电磁兼容性试验	16
7.6 标志和机械结构试验	16
8 说明文件	16
附录 A (规范性附录) 应用要求	17
附录 B (规范性附录) 声环境功能区场景噪声仿真试验环境要求	20
附录 C (规范性附录) 基于神经网络的语音质量评价(NNSQE)方法	23
附录 D (资料性附录) 一种典型的 NNSQE 方法——NN MOS 评分系统	26
参考文献	32

前 言

GB/T 31070《楼寓对讲系统》分为4个部分：

- 第1部分：通用技术要求；
- 第2部分：全数字系统技术要求；
- 第3部分：特定应用技术要求；
- 第4部分：应用指南。

本部分为GB/T 31070的第3部分。

本部分按照GB/T 1.1—2009给出的规则起草。

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。

本部分由中华人民共和国公安部提出。

本部分由全国安全防范报警系统标准化技术委员会(SAC/TC 100)归口。

本部分起草单位：公安部第三研究所、公安部第一研究所、厦门狄耐克智能科技股份有限公司、科大讯飞股份有限公司、长沙世邦通信技术有限公司、来邦科技股份公司、珠海三润电子有限公司、厦门立林科技有限公司、深圳市视得安罗格朗电子有限公司、福建省冠林科技有限公司、厦门ABB振威电器设备有限公司、中山市奥敏电子有限公司、北京中盾安全技术开发公司、国家安全防范报警系统产品质量监督检验中心(上海)、国家安全防范报警系统产品质量监督检验中心(北京)。

本部分主要起草人：戎玲、陈平、施巨岭、游寒旭、解桂秋、殷兵、余求理、王庆、邓大航、汤光耀、张达勇、陈谧、张仰鹏、仇锦涛、张济国、庄伟、杨鑫、陈俊健、陈伟。

引 言

GB/T 31070.1—2014 规定了楼宇对讲系统的组成、功能要求、性能要求、试验方法和检验规则等通用技术要求。GB/T 31070.2—2018 规定了全数字楼宇对讲系统的系统结构、功能要求、性能要求、试验方法、说明文件和检验规则。GB/T 31070 的本部分针对在特定场合应用的对讲系统提出了技术要求，这些系统包括但不限于：公共区域报警对讲系统、监室对讲系统、监所探访对讲系统、审讯室对讲系统、岗亭对讲系统、景区报警对讲系统、银行自助设备求助对讲系统、银行窗口对讲系统、防尾随门对讲系统、地铁/动车车厢对讲系统、站台求助对讲系统、高速公路对讲系统、病房医护对讲系统、手术室对讲系统、ICU(重症加强护理病房)对讲系统、养老监护对讲系统、工业企业内部对讲系统、电梯五方对讲系统等。



楼宇对讲系统

第3部分：特定应用技术要求

1 范围

GB/T 31070 的本部分规定了特定应用对讲系统的分类、功能要求、性能要求、试验方法和说明文件。

本部分适用于特定应用对讲系统,其他系统亦可参照执行。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是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 4208—2017 外壳防护等级(IP 代码)

GB/T 14522—2008 机械工业产品用塑料、涂料、橡胶材料人工气候老化试验方法 荧光紫外灯

GB/T 15211—2013 安全防范报警设备 环境适应性要求和试验方法

GB/T 20138 电器设备外壳对外界机械碰撞的防护等级(IK 代码)

GB/T 31070.1—2014 楼宇对讲系统 第1部分:通用技术要求

GB/T 31070.2—2018 楼宇对讲系统 第2部分:全数字系统技术要求

ITU-T P.340 免提终端的传输特性和语音质量参数(Transmission characteristics and speech quality parameters of hands-free terminals)

ITU-T P.501 电话测量用测试信号(Test signals for use in telephony)

ITU-T P.830 电话带宽和宽带数字编解码器的主观性能评估(Subjective performance assessment of telephone-band and wideband digital codecs)

ITU-T P.863—2018 感知客观收听质量评估(Perceptual objective listening quality prediction)

ETSI EG 202 396-1 语音和多媒体传输质量(STQ)背景噪声条件下的语音质量性能 第1部分:背景噪声模拟技术和背景噪声数据库(Speech Processing, Transmission and Quality Aspects (STQ); Speech quality performance in the presence of background noise; Part 1: Background noise simulation technique and background noise database)

ETSI EG 202 396-3 语音和多媒体传输质量(STQ)背景噪声条件下的语音质量性能 第3部分:背景噪声传输-客观测试方法(Speech Processing, Transmission and Quality Aspects (STQ); Speech Quality performance in the presence of background noise Part 3: Background noise transmission-Objective test methods)

3 术语、定义和缩略语

3.1 术语和定义

GB/T 31070.1—2014 和 GB/T 31070.2—2018 界定的以及下列术语和定义适用于本文件。

3.1.1

特定应用对讲系统 **specific applications building intercom system;SABIS**

用于各类声环境功能区 and/或特定环境和/或重要场所,具有选呼、对讲、可视(如有)等功能的电子系统。

3.1.2

对讲终端 **intercom unit**

特定应用对讲系统中能够建立通话的设备,包括管理机和下属对讲终端机。

注1: 不包括电话、智能手机、平板电脑等。

注2: 根据安装的环境分为建筑外部声环境功能区(嘈杂)类和建筑内部声环境功能区(非嘈杂)类。

3.1.3

昼间等效声级 **day-time equivalent sound level**

在昼间时段内测得的等效连续 A 计权声级。

注: 用 L_d 表示,单位为分贝[dB(A)]。

3.1.4

夜间等效声级 **night-time equivalent sound level**

在夜间时段内测得的等效连续 A 计权声级。

注: 用 L_n 表示,单位为分贝[dB(A)]。

3.1.5

全双工对讲 **full duplex**

对讲终端之间音频的发送和接收通信各自独立,语音可以在两个方向同时传输的对讲模式。

3.1.6

音频激活声压 **audio threshold sound pressure**

激励音频信道从空闲状态进入单方向语音传输状态,所需最小的语音声压值。

3.1.7

音频建立时间 **audio build-up time**

在标准激励语音声压下,音频信道从空闲状态进入单方向语音传输状态,幅度达到比正常值低 3 dB 时所需的时间。

3.1.8

通话客观语音评价 **objective speech quality evaluation for communication**

双向对讲状态下语音质量的客观评价得分。

3.1.9

窄带语音 **narrowband speech**

在 200 Hz~4 000 Hz 频率范围内满足全程灵敏度/频率特性的语音。

3.1.10

宽带语音 **broadband speech**

在 100 Hz~8 000 Hz 频率范围内满足全程灵敏度/频率特性的语音。

3.2 缩略语

下列缩略语适用于本文件。

CSS:合成源信号(Composite Source Signal)

EUT:受试设备(Equipment Under Test)

MRP:嘴参考点(Mouth Reference Point)

MU:管理机(Management Unit)

NNSQE:基于神经网络的语音质量评价(Neural Network Speech Quality Evaluation)

NVR:网络硬盘录像机(Network Video Recorder)

POLQA:感知客观收听质量评估(Perceptual Objective Listening Quality Prediction)

SIU:下属对讲终端机(Subordinate Intercom Unit)

4 系统分类

4.1 按音频带宽分类

按音频带宽宽度不同,特定应用对讲系统(以下简称为系统)分为窄带语音系统和宽带语音系统。

4.2 按应用场景分类

按应用场景不同,系统的分类见附录 A 表 A.1。

4.3 按建筑声环境功能区分类

4.3.1 建筑外部声环境功能区分类

按区域的功能特点和环境要求不同,分类如下:

- a) 建筑外部 0 类声环境功能区:特别需要安静,如医疗养老等区域,昼间等效声级 L_d 不大于 50 dB(A),夜间等效声级 L_n 不大于 40 dB(A)。取其昼间等效声级最大值 50 dB(A)为通用条件。
- b) 建筑外部 1 类声环境功能区:需要保持安静,如银行自助设备求助对讲 SIU 等的安装区域,昼间等效声级 L_d 不大于 55 dB(A),夜间等效声级 L_n 不大于 45 dB(A)。取其昼间等效声级最大值 55 dB(A)为通用条件。
- c) 建筑外部 2 类声环境功能区:需要维护安静,如公共区域、岗亭、景区对讲 SIU 等安装区域,昼间等效声级 L_d 不大于 60 dB(A),夜间等效声级 L_n 不大于 50 dB(A)。取其昼间等效声级最大值 60 dB(A)为通用条件。
- d) 建筑外部 3 类声环境功能区:需要防止噪声对周围环境产生严重影响,如站台求助对讲 SIU 等的安装区域,昼间等效声级 L_d 不大于 65 dB(A),夜间等效声级 L_n 不大于 55 dB(A)。取其昼间等效声级最大值 65 dB(A)为通用条件。
- e) 建筑外部 4 类声环境功能区:需要防止交通噪声对周围环境产生严重影响,如公路对讲 SIU 的安装区域,高速公路、一级公路、二级公路、城市快速路、城市主干路、城市次干路、城市轨道交通(地面段)、内河航道两侧区域、铁路干线两侧区域,昼间等效声级 L_d 不大于 70 dB(A),夜间等效声级 L_n 不大于 55 dB(A)。取其昼间等效声级最大值 70 dB(A)为通用条件。

4.3.2 建筑内部声环境功能区分类

按区域的功能特点和环境要求不同,分类如下:

- a) 建筑内部 0 类声环境功能区:审讯室、手术室、ICU(重症加强护理病房)对讲等的安装区域,噪声通过建筑物结构传播至噪声敏感建筑内部时,等效声级不大于 40 dB(A)的区域。取其等效声级最大值 40 dB(A)为通用条件;
- b) 建筑内部 1 类声环境功能区:公共区域、监室监狱、银行、医护养老、电梯对讲,以及岗亭、景区、站台、公路对讲 MU 等的安装区域,噪声通过建筑物结构传播至噪声敏感建筑内部时,等效声级不大于 45 dB(A)的区域。取其等效声级最大值 45 dB(A)为通用条件;
- c) 建筑内部 2 类声环境功能区:地铁/动车车厢、工业对讲等的安装区域,噪声通过建筑物结构传播

至噪声敏感建筑内部时,等效声级不大于 50 dB(A)的区域。取其等效声级最大值 50 dB(A)为通用条件。

5 功能要求

5.1 一般要求

按应用场景不同,系统的基本功能应满足 5.2 的要求。扩展功能配置应满足表 A.1 对应的要求,对应扩展功能应满足 5.3 的要求。

5.2 基本功能

5.2.1 呼叫

对讲终端应能呼叫系统中的其他对讲终端。呼叫过程中,呼叫方应有听觉和/或视觉的提示;被呼叫终端在收到呼叫信号后,应能发出听觉和/或视觉的提示。

5.2.2 对讲

系统应具有双向通话功能,语音清晰、连续且无明显漏字。

5.2.3 夜间操作

对讲终端应能提供夜间操作辅助照明。

5.2.4 防窃听

对讲终端间通话过程中,语音不应被系统中无权限的其他对讲终端窃听。

5.2.5 MU 广播

MU 应具备向 SIU 发起广播的功能,广播类型可包含定时广播、分区广播、喊话等。

5.2.6 外部扩展接口

系统应具备软/硬件扩展接口,宜具备 NVR 接口、状态提示输出接口。

5.2.7 远程维护与升级

系统宜具有远程维护和升级功能,在鉴别合法管理员权限后应能通过 MU 和/或联接的网络对系统中的 SIU 设备进行检测、调试或升级。

5.2.8 开锁及门开超时告警

具有电控开锁功能的系统,当系统控制的门体开启时间超过预设时间时,门体端和 MU 均应立即发出听觉和/或视觉的告警提示信息。

5.3 扩展功能

5.3.1 一键呼叫 MU

SIU 具有一键式呼叫 MU 功能,呼叫方应有等待提示。

5.3.2 紧急呼叫

SIU 具有紧急呼叫功能,呼叫应持续有效或循环呼叫转移直至 MU 接听为止。

5.3.3 全双工对讲

对讲终端具有全双工对讲模式。

5.3.4 录音录像与回放

系统具有对讲通话过程中的录音/录像功能,数据存储于 MU 中,并提供查询与回放功能。

5.3.5 防拆报警

安装于室外和/或室内公共场所和/或监管等特殊场所的 SIU 应具有防拆报警功能,当 SIU 被非法打开或移离安装位置时,应立即发出听觉和/或视觉报警, MU 应能指示相应 SIU 的名称/位置信息。

5.3.6 多级管理

MU 宜分为多个级别,不同的级别拥有不同的权限和管理范围。

5.3.7 监听监视

MU 具有对所管理的 SIU 进行监听/监视的功能。

5.3.8 多方通话

MU 具有主动发起并实现三方(或以上)通话的功能。

5.3.9 一键回拨

MU 具有一键回拨 SIU 的功能。

5.3.10 呼叫等待

系统具有呼叫等待功能。当被呼对讲终端处于占线状态时,呼叫方具有被叫状态提示语音或等待提示语音。带有显示屏的 MU 应具有呼入状态显示和/或呼入队列显示功能,宜具有通话保持功能。

5.3.11 呼叫转移

系统具有呼叫转移功能,可根据预设呼叫转移策略,将 SIU 的呼叫转移到系统内另一台 MU 上。转移事件类型包括无响应转移、关机转移、超时转移、强制转移(即托管)等。

5.3.12 一呼多响

系统配置多台 MU 时, SIU 呼叫 MU, 多台 MU 应能同时振铃。当其中任意一台 MU 接听通话后,其他 MU 应停止响铃。

5.3.13 视频联动

系统具有与视频监控系统联动功能,当 SIU 呼叫 MU 时,系统可自动显示视频监控系统中相关的图像。

5.3.14 在线检测

系统具有自动检测对讲终端在线状态的功能,当 SIU 掉线时应自动在 MU 上显示或提示,并宜具有故障检测功能。

6 性能要求

6.1 音频特性

6.1.1 全程响度评定值

6.1.1.1 窄带语音系统

在 200 Hz~4 000 Hz 范围内应满足下列要求：

- a) 应用于建筑外部声环境功能区免提通话方式的对讲终端： ≤ 20 dB；
- b) 应用于建筑内部声环境功能区免提通话方式的对讲终端： ≤ 23 dB；
- c) 手柄通话方式的对讲终端： ≤ 15 dB。

6.1.1.2 宽带语音系统

在 100 Hz~8 000 Hz 范围内应满足下列要求：

- a) 应用于建筑外部声环境功能区免提通话方式的对讲终端： ≤ 20 dB；
- b) 应用于建筑内部声环境功能区免提通话方式的对讲终端： ≤ 23 dB；
- c) 手柄通话方式的对讲终端： ≤ 15 dB。

6.1.2 全程灵敏度/频率特性

6.1.2.1 窄带语音系统

在 500 Hz~3 400 Hz 范围内，免提通话方式的对讲终端的典型曲线和允差范围应满足 GB/T 31070.1—2014 中图 4 的要求；手柄通话方式的对讲终端，应满足 GB/T 31070.1—2014 中图 5 的要求。

6.1.2.2 宽带语音系统

在 315 Hz~8 000 Hz 范围内的典型曲线为图 1 中虚线，允差范围为图 1 中实线，频率特性允差框线限值如表 1 所示。

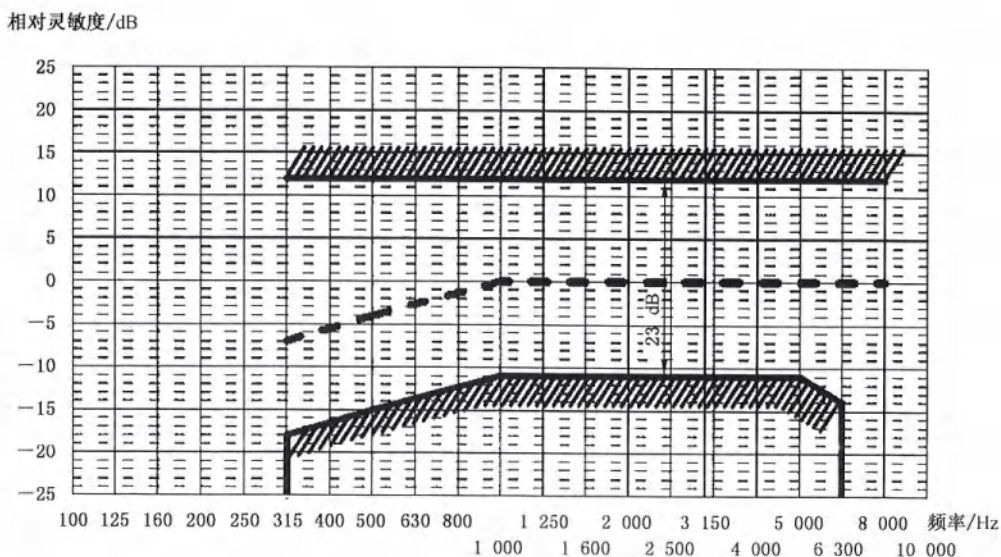


图 1 宽带语音系统全程灵敏度/频率特性模板

表 1 宽带语音系统频率特性允差框线限值表

频率/Hz	框线上限/dB	框线下限/dB
315	12	-18
500	12	-15
1 000	12	-11
5 000	12	-11
6 300	12	-14
8 000	12	—

6.1.3 侧音掩蔽评定值

手柄通话方式的对讲终端,手柄端的侧音掩蔽评定值应不小于 5 dB。

6.1.4 空闲信道噪声

免提通话方式的对讲终端空闲信道噪声应不大于 45 dB(A)。

手柄通话方式的对讲终端空闲信道噪声应不大于 48 dB(A)。

6.1.5 音频延时

系统端对端的音频延时不应大于 300 ms。

6.1.6 回声抑制

系统端对端的回声抑制应不小于 45 dB。

6.1.7 音频转换时间

系统的端对端的音频转换时间不应大于 200 ms。

注:不适用于双工通话及单向通话的系统。

6.1.8 音频激活声压

系统音频激活声压应不大于 -37 dBPa。

6.1.9 音频建立时间

系统音频建立时间不应大于 50 ms。

6.1.10 通话客观语音评价

系统通话客观语音评价得分应不小于 3 分。

6.1.11 声环境功能区通用条件下通话客观语音评价

系统在符合表 A.2 要求的相应声环境功能区条件下,通话客观语音评价得分应不小于 2.5 分。

6.1.12 振铃声压

对讲终端的振铃声压应不小于 62 dB(A)且不大于 106 dB(A)。

6.2 视频特性

具有可视功能的系统,视频特性应满足 GB/T 31070.1—2014 中的 6.2 的要求,全数字系统视频特性应满足 GB/T 31070.2—2018 中的 6.2 的要求。

6.3 电源电压适应性

电源电压适应性应满足 GB/T 31070.1—2014 中的 6.3 的要求。

6.4 环境适应性

6.4.1 常规环境适应性

6.4.1.1 根据使用环境的严酷程度不同,系统各组成部分的常规环境适应性分为三个等级,见表 2。系统常规环境适应性应符合表 A.3 的规定。

表 2 常规环境适应性要求

项目	I 级		II 级		III 级	
	额定值	试验时间	额定值	试验时间	额定值	试验时间
高温试验 (工作状态)	55 ℃	16 h	70 ℃	16 h	70 ℃	16 h
	室内安装的对讲终端和辅助装置: +55 ℃ 16 h					
低温试验 (工作状态)	-10 ℃	16 h	-25 ℃	16 h	-40 ℃	16 h
	室内安装的对讲终端和辅助装置: -10 ℃ 16 h					
恒定湿热试验 (工作状态)	40 ℃ 相对湿度 93%	96 h	40 ℃ 相对湿度 93%	96 h	40 ℃ 相对湿度 93%	96 h
正弦振动试验 (工作状态)	频率 10 Hz~150 Hz 加速度 2 m/s ² 扫频速率 1 oct/min 三轴向各 1 个循环		频率 10 Hz~150 Hz 加速度 5 m/s ² 扫频速率 1 oct/min 三轴向各 1 个循环		频率 10 Hz~150 Hz 加速度 5 m/s ² 扫频速率 1 oct/min 三轴向各 1 个循环	

6.4.1.2 在承受各项气候和机械环境试验后,系统各组成部分应无任何电气故障、结构变形或接触不良现象。每项试验中及试验后系统功能均应正常。

6.4.2 盐雾循环耐久性

6.4.2.1 对于系统中安装于室外,盐雾进入会导致严重后果的各组成部分,其设备外壳应具有抗盐雾腐蚀能力,根据使用环境的严酷程度不同分为二个等级,见表 3。

表 3 盐雾循环耐久性要求

项目	I 级		II 级	
	额定值	试验时间	额定值	试验时间
盐雾循环耐久性试验	盐雾环境:氯化钠 盐溶液浓度:(5±1)% 温度:15℃~35℃ pH 值:6.5~7.2	盐雾环境 2 h、湿热环境 22 h 为 1 次循环,循环 3 次,共 3 d	盐雾环境:氯化钠 盐溶液浓度:(5±1)% 温度:15℃~35℃ pH 值:6.5~7.2	盐雾环境 2 h、湿热环境 166 h 为 1 次循环,循环 4 次,共 28 d
	湿热环境:40℃ 相对湿度:93%		湿热环境:40℃ 相对湿度:93%	

6.4.2.2 盐雾试验后检查,设备外壳表面不应有涂层脱落、露出底材以及生锈、发黄等现象。

6.4.3 UV 防护适应性

系统中安装于室外环境的设备应具有 UV 防护性能,按 GB/T 14522—2008 附录 C 中表 C.1 试验条件暴露周期类型 5 进行 UV 防护适应性试验,试验总时长 168 h,非金属外壳表面不应出现变形、脆化等现象,设备外壳经表面处理(如喷漆)的不应出现龟裂、掉漆、氧化、起泡等现象。

6.5 安全性

系统的安全性应满足 GB/T 31070.1—2014 中的 6.5 的要求。

6.6 电磁兼容性

系统的电磁兼容性应满足 GB/T 31070.1—2014 中的 6.6 的要求。

6.7 标志、机械结构要求

6.7.1 标志

系统设备的标志应满足 GB/T 31070.1—2014 中的 6.7.1 的要求。

6.7.2 机械结构要求

系统设备的机械结构应满足 GB/T 31070.1—2014 中的 6.7.2 的要求。

6.7.3 外壳防护能力

系统中安装于室内环境的设备,外壳防护能力应满足 GB/T 4208—2017 中 IP30 要求。安装于室外环境和/或室内公共场所和/或监管等特殊场所的设备,外壳防护能力按使用环境的严酷等级分为三个等级,应符合表 A.3 对应的等级要求:

- a) I 级:符合 GB/T 4208—2017 中 IP33 的要求;
- b) II 级:符合 GB/T 4208—2017 中 IP44 的要求;
- c) III 级:符合 GB/T 4208—2017 中 IP54 的要求。

6.7.4 外壳抗破坏能力

具有抗破坏能力的系统设备,依据使用环境的严酷等级不同分为四个等级,按产品说明正常安装后,应符合表 A.3 对应等级的要求:

- a) I级:符合 GB/T 20138 中 IK04 的要求;
- b) II级:符合 GB/T 20138 中 IK07 的要求;
- c) III级:符合 GB/T 20138 中 IK07 的要求和 GB/T 31070.1—2014 中 6.7.3.2c)的要求;
- d) IV级:符合 GB/T 20138 中 IK10 的要求和 GB/T 31070.1—2014 中 6.7.3.2c)的要求。

7 试验方法

7.1 试验条件

7.1.1 试验环境条件

试验环境条件如下:

- 温度:15 °C~35 °C;
- 相对湿度:25%~75%;
- 大气压力:86 kPa~106 kPa。

7.1.2 电气连接

EUT 应按制造商推荐的方法进行连接构成试验基本配置,EUT 的配置应满足实现系统功能的要求。

7.2 功能检查

按使用说明书操作,逐项检查系统功能,判定试验结果是否符合第 5 章的要求。

7.3 音频特性试验

7.3.1 试验条件

试验条件满足以下要求:

- a) 试验场地的环境噪声应不大于 40 dB(A),特殊说明除外;
- b) 全程响度评定值、全程灵敏度/频率响应、侧音掩蔽评定值、通话客观语音评价方法试验的测试信号采用 ITU-T P.501 规定的中文语音信号。音频延时、回声抑制、音频转换时间、音频建立时间试验的测试信号采用 ITU-T P.501 规定的 CSS 复合信号,每个 CSS 信号持续 248.62 ms、停顿 101.38 ms,测试信号在 MRP 处产生的平均声压应为-4.7 dBPa;
- c) 采用手柄通话方式的对讲终端,试验安装位置应符合 GB/T 31070.1—2014 附录 A 中 A.1c)的要求;
- d) 采用壁挂免提通话方式的对讲终端,试验安装位置应符合 GB/T 31070.2—2018 中 7.3.2e)的要求;
- e) 采用桌面免提通话方式的对讲终端,试验安装位置应按 ITU-T P.340 建议的配置,见图 2。测试桌应采用坚硬木质制成,桌面应平滑、牢固、水平放置。桌面面积为 1 m²(不小于 0.96 m²)、宽度不小于 800 mm、厚度不小于 20 mm。对讲终端、仿真嘴和测量传声器应按图 2 规定放置,仿真嘴或测量传声器轴线应与图 2 的 BC 线重合;

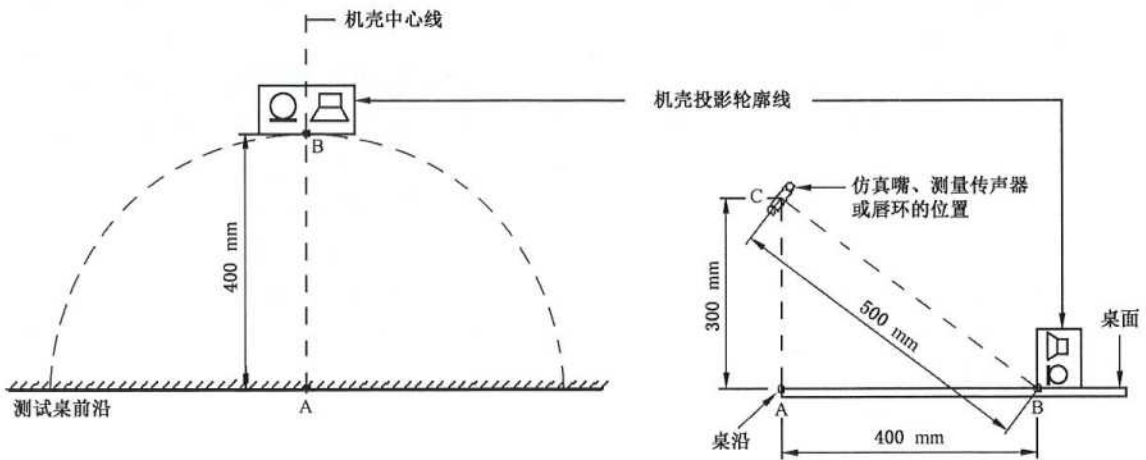


图2 桌面免提通话方式的对讲终端测试位置示意图

- f) 采用桌面鹅颈话筒方式通话的对讲终端，试验安装位置与 7.3.1e) 要求基本一致，因鹅颈话筒安装位置的不同，被测终端、仿真嘴或自由场传声器的位置允许选择图 3a) 或图 3b) 中的一种。图 3a) 中被测终端的机壳中心线应与 AB 线重合；图 3b) 中鹅颈话筒中心线应与自由场传声器中心相交。试验时鹅颈话筒中心线应与 BC 线平行，见图 3c)；

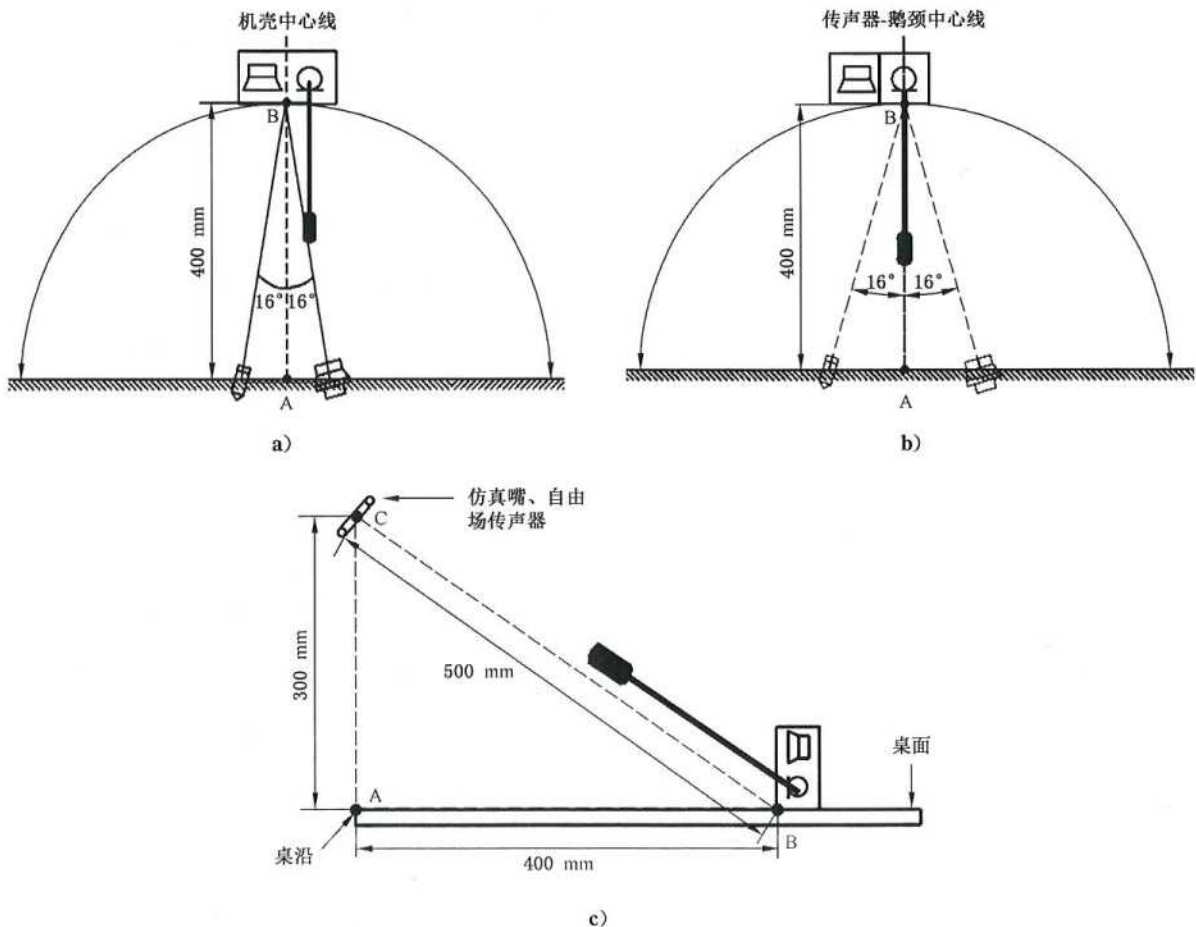


图3 桌面鹅颈话筒通话方式的对讲终端测试位置示意图

- g) 声环境功能区通用条件下通话客观语音评价的试验环境应符合附录 B 要求。动车、地铁内部对讲系统的车厢内场景噪声信号源为 ETSI EG 202 396-1\Binaural_Signals\Inside_Train_Noise3_binaural (0.00—30.00 s); 高速公路对讲系统的高速公路场景噪声信号源为 Roadnoise; 站台求助对讲系统的站台场景噪声信号源为 P.501\Annex B Monaural noises sequences\met_mono1_30 s; 其他场景通用噪声信号源为 ITU-T P.501\Annex B Monaural noises sequences\res_mono1_30 s。采用噪声播放扬声器阵列的方式模拟产生对应的建筑外部声环境功能区/建筑内部声环境功能区的噪声,噪声信号在 MRP 处产生的平均声压应分别对应的建筑外部声环境功能区/建筑内部声环境功能区的通用条件噪声等效声级;
- h) 如无特殊说明,对于通话音量可调 EUT,试验时应将通话音量设置在满足全程响度评定值要求的出厂默认值;
- i) 对讲终端建立通话后,EUT 之间的试验环境应能确保声音不致互相影响。

7.3.2 全程响度评定值试验

7.3.2.1 窄带语音系统

按 GB/T 31070.1—2014 中 A.2 规定的方法进行,判定试验结果是否符合 6.1.1.1 的要求。

7.3.2.2 宽带语音系统

全程响度评定值的试验方法如下:

- a) 按 GB/T 31070.1—2014 中附录 A 的 A.2.1~A.2.3 规定方法,在 100 Hz~8 000 Hz 频率范围内进行嘴参考点 P_m 值、各频率点的声压 P_o 值、耳承输出声压 P_e 测量;
- b) 采用免提通话方式的对讲终端的全程响度评定值的计算采用以下方法:
按公式(1)计算免提通话方式的宽带语音对讲终端的全程灵敏度/频率特性 S_{ro} :

$$S_{ro} = 20 \lg \left(\frac{P_o}{P_m} \right) \text{dB (相对于 1 Pa/Pa)} \quad \dots\dots\dots (1)$$

按公式(2)计算免提通话方式的宽带语音对讲终端的全程响度评定值 OLR:

$$\text{OLR} = -\frac{10}{m} \times \lg \sum_{i=1}^{20} 10^{\frac{m}{10}(S_{ro}-W_{oi})} \quad \dots\dots\dots (2)$$

式中:

W_{oi} ——计算 OLR 的加权系数,不同频率的 W_o 见表 4;

m ——斜率参数, $m=0.175$ 。

判定试验结果是否符合 6.1.1.2a)~b) 的要求;

- c) 采用手柄通话方式的宽带语音对讲终端的全程响度评定值的计算采用以下方法:
按公式(3)计算采用手柄通话方式的宽带语音对讲终端的全程灵敏度/频率特性 S_{re} :

$$S_{re} = 20 \lg \left(\frac{P_e}{P_m} \right) \text{dB (相对于 1 Pa/Pa)} \quad \dots\dots\dots (3)$$

按公式(4)计算手柄通话方式的宽带语音对讲终端全程响度评定值 OLR:

$$\text{OLR} = -\frac{10}{m} \times \lg \sum_{i=1}^{20} 10^{\frac{m}{10}(S_{re}-W_{oi}-L_{Ei})} \quad \dots\dots\dots (4)$$

式中:

W_{oi} ——计算 OLR 的加权系数,不同频率的 W_o 见表 4;

L_{Ei} ——为耳承声泄漏系数,不同频率的 L_E 见表 4;

m ——斜率参数, $m=0.175$ 。

判定试验结果是否符合 6.1.1.2c) 的要求。

表 4 不同频率点的 W_0 和 L_E 系数表

频带序号 i	中心频率/Hz	全程加权系数 W_0	耳承声泄漏系数 L_E
1	100	107.0	20
2	125	80.1	16.5
3	160	65.7	12.5
4	200	66.1	8.4
5	250	60.7	4.9
6	315	68.5	1.0
7	400	55.6	-0.7
8	500	66.9	-2.2
9	630	63.3	-2.6
10	800	63.4	-3.2
11	1 000	65.3	-2.3
12	1 250	73.1	-1.2
13	1 600	70.1	-0.1
14	2 000	82.0	3.6
15	2 500	78.6	7.4
16	3 150	95.4	6.7
17	4 000	76.9	8.8
18	5 000	92.4	10
19	6 300	92.2	12.5
20	8 000	76.7	15

7.3.3 全程灵敏度/频率特性试验

7.3.3.1 窄带语音系统

按 GB/T 31070.1—2014 中 A.3 规定的方法进行,判定试验结果是否符合 6.1.2.1 的要求。

7.3.3.2 宽带语音系统

全程灵敏度/频率特性试验方法如下:

- 采用免提方式通话的对讲终端,按 7.3.2.2a) 方法进行测量,并按公式(1)计算全程灵敏度/频率特性,记录 P_0 。相对于 P_m 的声压波动;
- 采用手柄方式通话的对讲终端,按 7.3.2.2c) 方法进行测量,并按公式(3)计算全程灵敏度/频率特性,记录 P_0 。相对于 P_m 的声压波动;
- 在 315 Hz~8 000 Hz 范围内,判定试验结果是否符合 6.1.2.2 的要求。

注:频率间隔可用 1/3 倍频程,仲裁时可用 1/12 倍频程。

7.3.4 侧音掩蔽评定值试验

按 GB/T 31070.1—2014 中 A.6 规定的方法进行,判定试验结果是否符合 6.1.3 的要求。

7.3.5 空闲信道噪声试验

按 GB/T 31070.1—2014 中 A.7 规定的方法进行,判定试验结果是否符合 6.1.4 的要求。

7.3.6 音频延时试验

按 GB/T 31070.2—2018 中 7.3.3 规定的方法进行,判定试验结果是否符合 6.1.5 的要求。

7.3.7 回声抑制试验

按 GB/T 31070.2—2018 中 7.3.4 规定的方法进行,判定试验结果是否符合 6.1.6 的要求。

7.3.8 音频转换时间试验

按 GB/T 31070.2—2018 中 7.3.5 规定的方法进行,判定试验结果是否符合 6.1.7 的要求。

7.3.9 音频激活声压试验

音频激活声压试验方法如下:

- a) 测试信号组由周期重复的 10 个 CSS 组成,每个 CSS 信号持续 248.62 ms、停顿 451.38 ms,起始第一个 CSS 信号持续高电平部分在 MRP 处产生的平均声压级为 -44 dBPa,声压级由小逐渐增大,相邻两段信号的幅度差为 1 dB,最终一个 CSS 信号持续高电平部分在 MRP 处产生的平均声压级为 -35 dBPa,见图 4;

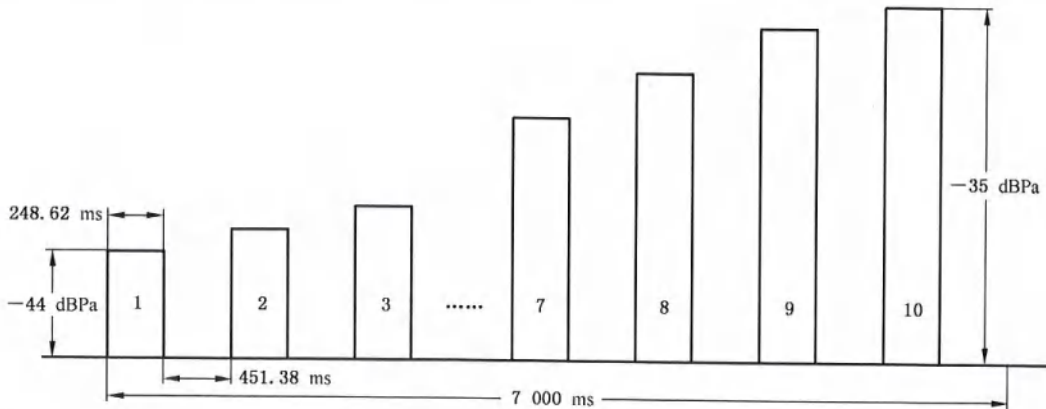


图 4 音频激活声压试验测试信号组

- b) 音频信号发生器产生以上 CSS 测试信号组信号,通过仿真嘴激励本地对讲终端,在远端以传声器采集输出的音频信号,通过音频分析仪记录并分析 CSS 信号的数量,从记录的最终一个 CSS 信号持续高电平部分逆序统计前序的完整的 CSS 信号的数量,前序不完整的 CSS 信号小于音频最小激活声压不予统计,通过计算即可对应 -35 dBPa、-36 dBPa、-37 dBPa、-38 dBPa 直至最小的 -44 dBPa,判定试验结果是否符合 6.1.8 的要求。

7.3.10 音频建立时间试验

音频建立时间试验方法如下:

- a) 测试信号组由周期重复的 4 个 CSS 组成,每个 CSS 信号持续 248.62 ms、停顿 101.38 ms,每一个 CSS 信号持续高电平部分在 MRP 处产生的平均声压级为 -4.7 dBPa,第 1 个 CSS 信号的持续高电平部分用于测量,见图 5;

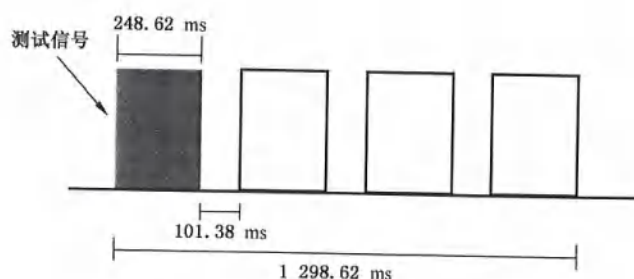


图5 音频建立时间试验测试信号组

- b) 音频信号发生器产生以上 CSS 测试信号组信号,通过仿真嘴激励本地对讲终端,在远端以传声器采集输出的音频信号,通过音频分析仪记录第 1 个 CSS 信号的高电平部分宽度,从下降沿逆序计算至幅度比正常传送信号幅度低 3 dB 处的时间 t ,将测试信号的时长 248.62 ms 减去逆序计算的时间 t ,差值即为音频建立时间。判定试验结果是否符合 6.1.9 的要求。

7.3.11 通话客观语音评价试验

7.3.11.1 客观语音评价方法应选择满足附录 C 要求的 NNSQE 方法或 ITU-T P.863—2018 规定的 POLQA 方法,同一次试验中采用的语音评价方法应一致。

7.3.11.2 通话客观语音评价试验方法如下:

- a) 测试信号采用 ITU-T P.501 建议的中文真人语音信号,测试信号共包含 10 句中文真人语音,奇数语音信号句段分别记为 S_{s1} 、 S_{s3} 、 S_{s5} 、 S_{s7} 、 S_{s9} ,用于本地端仿真嘴的激励,偶数语音信号句段分别记为 S_{s2} 、 S_{s4} 、 S_{s6} 、 S_{s8} 、 S_{s10} ,用于远端仿真嘴的激励,每段语音信号句段测试信号在 MRP 处产生的平均声压应为 -4.7 dBPa。测试信号的包络见图 6;

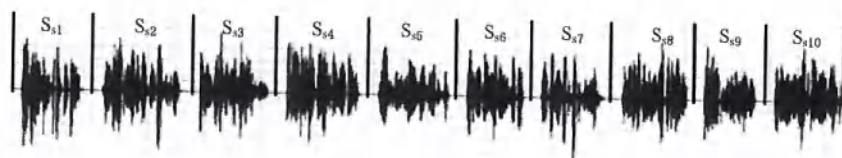


图6 语音测试信号包络及句段分配示意图

- b) 以 S_{s1} 语音信号通过本地端仿真嘴激励本地对讲终端,在远端以传声器采集并通过音频分析仪记录远端对讲终端输出语音信号 S_{r1} 。当 S_{r1} 接收并播放完成时,以 S_{s2} 语音信号通过远端仿真嘴激励远端对讲终端,在本地端以传声器采集并通过音频分析仪记录本地对讲终端输出信号 S_{r2} ;
- c) 当 S_{r2} 接收并播放完成时,继续以 S_{s3} 语音信号通过本地端仿真嘴激励本地对讲终端,按 7.3.11b) 的方法,依次将奇数和偶数语音信号句段在本地和远端对讲终端进行发送,得到全部 10 句输出语音;
- d) 采用选定的客观语音评价方法,分别对 (S_{s1}, S_{r1}) 、 (S_{s2}, S_{r2}) ... (S_{s10}, S_{r10}) 10 对传输前语音和传输后语音进行语音质量评价,分别记录得分;
- e) 5 个奇数句段得分的平均值即为远端通话客观语音评价得分,5 个偶数句段得分的平均值即为本地端通话客观语音评价得分,判定试验结果是否符合 6.1.10 的要求。

7.3.12 声环境功能区通用条件下的通话客观语音评价试验

按产品典型应用场景,对应选择对讲终端的声环境功能区等级,测试环境应符合附录 B 要求,按

GB/T 31070.3—2021

7.3.11 进行试验,判定试验结果是否符合 6.1.11 的要求。

7.3.13 振铃声压试验

按 GB/T 31070.1—2014 中 A.8 规定的方法进行,判定试验结果是否符合 6.1.12 的要求。

7.3.14 视频特性试验

按 GB/T 31070.1—2014 和 GB/T 31070.2—2018 中 7.4 规定的方法进行,判定试验结果是否符合 6.2 的要求。

7.3.15 电源电压适应性试验

按 GB/T 31070.1—2014 中 7.5 规定的方法进行,判定试验结果是否符合 6.3 的要求。

7.3.16 环境适应性试验

7.3.16.1 常规试验

按 GB/T 31070.1—2014 中的 7.6 规定的方法进行,判定试验结果是否符合 6.4.1 的要求。

7.3.16.2 盐雾循环耐久性试验

按 GB/T 15211—2013 第 18 章规定的方法进行,试验结束后允许在常温下放置 2 h 后检查外观。判定试验结果是否符合 6.4.2 的要求。

7.3.16.3 UV 防护适应性试验

按 GB/T 14522—2008 规定的试验方法进行,试验后样品与未试验参照品对比,判定试验结果是否符合 6.4.3 的要求。

7.4 安全性试验

按 GB/T 31070.1—2014 中 7.7 规定的方法进行,判定试验结果是否符合 6.5 的要求。

7.5 电磁兼容性试验

按 GB/T 31070.1—2014 中 7.8 规定的方法进行,判定试验结果是否符合 6.6 的要求。

7.6 标志和机械结构试验

按 GB/T 31070.1—2014 中 7.9、7.10 和 GB/T 20138 规定的方法进行,判定试验结果是否符合 6.7 的要求。

8 说明文件

应符合 GB/T 31070.1—2014 中第 8 章的要求。在产品说明书中还应明确产品具有与附录 A 表格规定的相应功能和性能。

附 录 A
(规范性附录)
应用要求

A.1 扩展功能配置要求

按应用场景不同,系统各类产品的扩展功能配置应符合表 A.1 的要求。

表 A.1 扩展功能配置要求

系统分类		扩展功能要求													
应用场景 大类	应用场景分类	一键 呼叫 MU	紧急 呼叫	全双 工对 讲	录音 录像 与回 放	呼叫 终端 防拆 报警	多级 管理	监听 监视	多方 通话	一键 回拨	呼叫 等待	呼叫 转移	一呼 多响	视频 联动	在线 检测
司法 公安	公共区域报警对讲	●	●	○	●	●	●	●	●	●	●	●	●	●	●
	监室对讲	●	●	○	●	●	●	—	●	●	●	●	○	○	
	监狱探访对讲	—	—	○	●	—	○	●	●	—	—	—	○	○	
	审讯室对讲	●	○	○	○	—	●	●	●	○	●	○	○	○	
	岗亭对讲	●	○	○	○	—	●	○	●	●	●	○	○	○	
	景区报警对讲	●	●	○	●	●	●	●	●	●	●	●	●	●	
金融	银行自助设备求助对讲	●	○	○	○	●	●	●	—	●	●	●	○	●	
	银行窗口对讲	○	○	●	○	○	○	○	○	○	○	○	○	○	
	防尾随门对讲	○	○	○	○	○	—	○	—	○	○	○	○	○	
交通	地铁/动车车厢对讲	●	○	○	○	●	—	○	—	○	●	○	○	○	
	站台求助对讲	●	○	○	○	●	●	●	—	○	●	○	○	○	
	高速公路对讲	●	○	○	○	●	●	●	—	○	●	○	○	○	
医疗 养老	病房医护对讲	●	○	○	○	—	●	—	—	●	●	○	○	—	
	手术室对讲	●	○	○	○	—	●	○	—	○	●	○	○	○	
	ICU(重症加强护理病房)对讲	●	○	○	○	—	○	●	●	○	●	○	○	○	
	养老监护对讲	●	○	○	○	○	●	○	—	○	●	○	○	○	
工业	工业企业内部对讲(车间及嘈杂区域)	●	○	○	○	●	○	○	○	○	●	○	○	○	
	工业企业内部对讲(办公区域)	●	○	○	○	○	○	○	○	○	●	○	○	○	
居住建筑 办公建筑	电梯五方对讲	●	○	○	○	○	●	●	●	●	●	○	●	○	

注：●必要项；○可选择项；—不做要求。

A.2 声环境功能区等级要求

按应用场景不同,系统应确定各组成部分的声环境功能区,对应建筑物外部及内部不同应用环境的声环境功能区等级应符合表 A.2 的规定。

表 A.2 对应建筑外部及建筑内部声环境功能区等级要求

系统分类		声环境功能区	
应用场景大类	应用场景分类	建筑外部声环境功能区	建筑内部声环境功能区
司法 公安	公共区域对讲	外部 2 类或以上	内部 1 类或以上
	监室对讲	—	内部 1 类或以上
	监狱探访对讲	—	内部 1 类或以上
	审讯室对讲	—	内部 0 类或以上
	岗亭对讲	外部 2 类或以上	内部 1 类或以上
	景区报警对讲	外部 2 类或以上	内部 1 类或以上
金融	银行自助设备求助对讲	外部 1 类或以上	内部 1 类或以上
	银行窗口对讲	—	内部 1 类或以上
	防尾随门对讲	—	内部 1 类或以上
交通	地铁/动车车厢对讲	外部 2 类或以上(SIU)	内部 2 类或以上(MU)
	站台求助对讲	外部 3 类或以上	内部 1 类或以上
	高速公路对讲	外部 4 类	内部 1 类或以上
医疗 养老	病房医护对讲	—	内部 1 类或以上
	手术室对讲	—	内部 0 类或以上
	ICU(重症加强护理病房)对讲	—	内部 0 类或以上
	养老监护对讲	外部 0 类或以上	内部 1 类或以上
工业 企业	车间及嘈杂区域对讲	外部 3 类或以上(SIU)	内部 2 类或以上(MU)
	办公区域对讲	外部 1 类或以上(SIU)	内部 1 类或以上(MU)
居住建筑 办公建筑	电梯五方对讲	—	内部 1 类或以上
注:地铁/动车车厢对讲 SIU、工业企业车间及嘈杂区域对讲 SIU 及办公区域对讲 SIU 的安装应用环境超出建筑内部声环境功能区等级要求,因此按建筑外部声环境功能区等级进行要求。其他未列出的系统可根据实际应用场景选择合适的声功能环境区等级。			

A.3 其他特性要求

系统各类产品的其他特性要求应符合表 A.3 的规定。

表 A.3 其他特性要求

系统分类		其他特性要求						
应用场景 大类	应用场景分类	电源电压 适应性	常规环境 适应性	盐雾循环 耐久性	UV 防护 适应性	电磁兼容 要求	外壳防护 等级	外壳抗破 坏能力
司法 公安	公共区域对讲	●	Ⅲ	○	○	●	Ⅲ	Ⅲ
	监室对讲	●	I	○	○	●	Ⅲ	Ⅳ
	监狱探访对讲	●	I	○	○	●	I	Ⅱ
	审讯室对讲	●	I	○	○	●	I	I
	岗亭对讲	●	Ⅱ	○	○	●	I	I
	景区报警对讲	●	Ⅱ	○	○	●	Ⅲ	Ⅲ
金融	银行自助设备求助对讲	●	I	○	○	●	Ⅱ	Ⅲ
	银行窗口对讲	●	I	○	○	●	I	I
	防尾随门对讲	●	I	○	○	●	I	I
交通	地铁/动车车厢对讲	●	I	○	○	●	Ⅱ	Ⅲ
	站台求助对讲	●	Ⅱ	○	○	●	Ⅲ	Ⅲ
	高速公路对讲	●	Ⅲ	○	○	●	Ⅲ	Ⅲ
医疗 养老	病房医护对讲	●	I	○	○	●	I	I
	手术室对讲	●	I	○	○	●	I	I
	ICU(重症加强护理病房)对讲	●	I	○	○	●	I	I
	养老监护对讲	●	I	○	○	●	I	I
工业	工业企业内部对讲(车间及嘈杂区域)	●	I	○	○	●	Ⅱ	Ⅲ
	工业企业内部对讲(办公区域)	●	I	○	○	●	I	I
居住建筑 办公建筑	电梯五方对讲	●	I	○	○	●	Ⅱ	Ⅱ

注 1: ●必要项;○可选项;—不做要求。I/Ⅱ/Ⅲ/Ⅳ为必要项的对应级别。

注 2: 盐雾循环耐久性为推荐性,根据产品特点宜选择符合 I 级或 II 级要求;UV 防护适应性为推荐性。

附录 B
(规范性附录)

声环境功能区场景噪声仿真试验环境要求

B.1 环境配置

除特别声明外,声环境功能区场景噪声仿真试验环境应符合:

- 扬声器数量:5个(其中1个为重低音扬声器);
- 房间面积: $2.5 \times 3 \text{ m}^2 \sim 3.5 \times 4 \text{ m}^2$;
- 房间高度:2.20 m~2.50 m;
- 房间混响时间:在 200 Hz~8 000 Hz 的混响时间应 $\leq 0.7 \text{ s}$,且应 $\geq 0.2 \text{ s}$;
- 本底噪声: $\leq 30 \text{ dB(A)}$ 。

B.2 扬声器布置

B.2.1 扬声器阵列的摆放位置如图 B.1 所示。

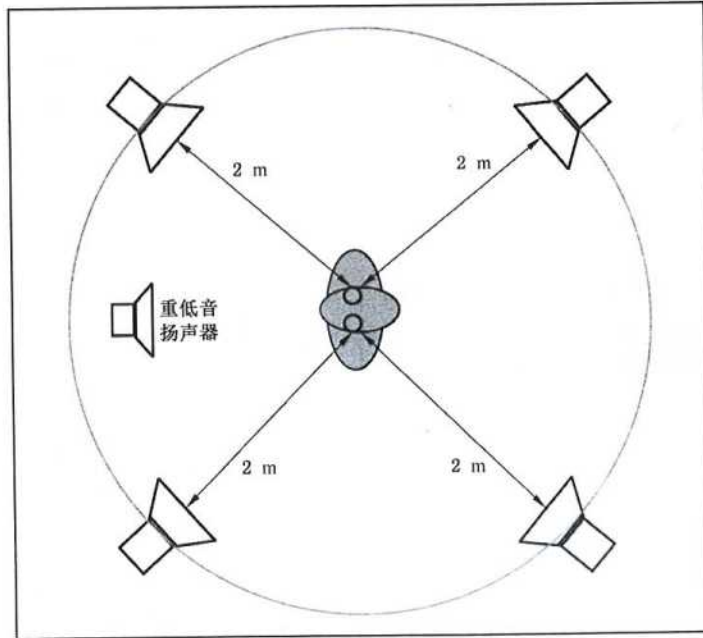


图 B.1 仿真试验环境中的扬声器阵列摆放位置

B.2.2 音频测试装置应摆放在仿真环境中心位置,4个扬声器距离中心位置的距离应相同,在 0.5 m~2.5 m 之间,如图 B.1 所示为 2 m。

B.2.3 相邻位置的扬声器距离应一致,按正方形 4 个顶点位置摆放。

B.2.4 扬声器摆放高度应与音频测试装置声音采集点的高度一致,误差应在 $\pm 15 \text{ cm}$ 之间。

B.2.5 重低音扬声器参考图 B.1 的位置摆放。

B.3 扬声器阵列校准

B.3.1 阵列校准示意图

扬声器阵列校准示意图如图 B.2 所示。

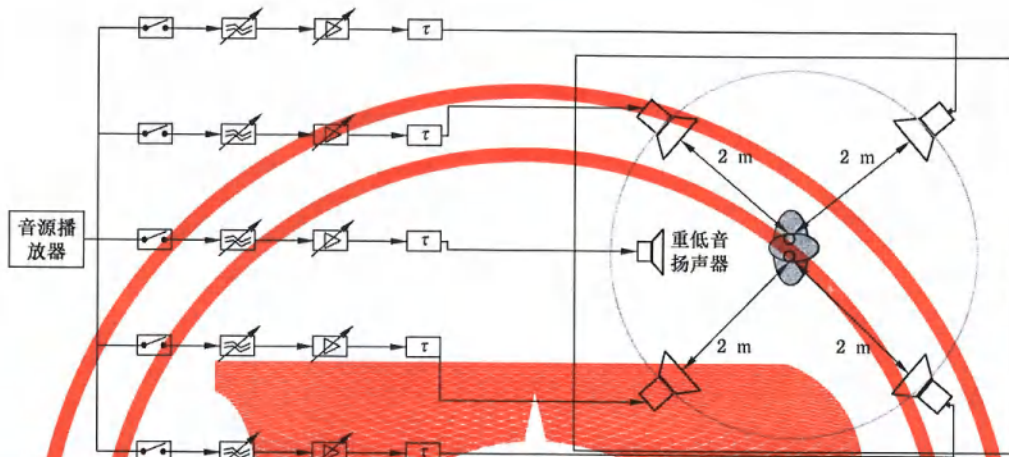


图 B.2 扬声器阵列校准示意图

B.3.2 扬声器单独校准

B.3.2.1 应按左前、左后、右前、右后顺序分别对阵列扬声器进行校准。阵列扬声器校准频率范围为 120 Hz~20 kHz。重低音扬声器的校准频率范围应为:30 Hz~120 Hz。

B.3.2.2 扬声器声压级相对原始噪声源的声压级低 6 dB。

B.3.2.3 校准时可通过补偿滤波器补偿扬声器频响,以满足校准要求。补偿滤波器信号幅度应≤9 dB。补偿滤波器的频响应满足公式(B.1):

$$H_{EQ}(f) = \frac{1}{H_{LS}(f)} \dots\dots\dots (B.1)$$

B.3.2.4 采用粉伪随机噪声作为校准信号。将校准信号通过高通滤波器(120 Hz)输入至阵列扬声器播放,播放时长应不小于 2 s,同时用音频测试装置采集播放的噪声。

B.3.2.5 扬声器频响在 120 Hz~20 kHz 频率范围内应是平坦的,允许容差在±3 dB 之间。如图 B.3 所示。通过公式(B.2)录制噪声和原始噪声的功率密度谱比值:

$$H_{LS}(f) = \frac{S_{LS}(f)}{S_{pn}(f)} \dots\dots\dots (B.2)$$

注: $H_{LS}(f)$ 采用 1/3 倍频程分析, $S_{pn}(f)$ 原始噪声功率谱, $S_{LS}(f)$ 录制噪声功率谱。

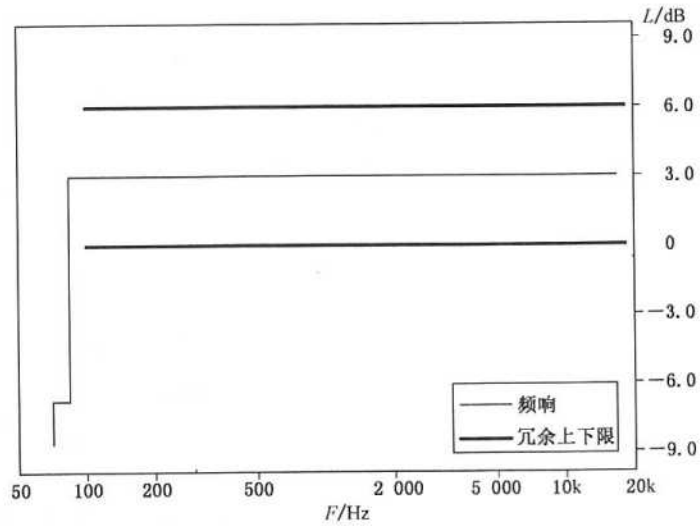


图 B.3 阵列扬声器频响曲线允许容差示意图

B.3.3 阵列校准

B.3.3.1 完成 B.3.2 扬声器单独校准后,应再进行扬声器阵列的阵列校准,即 4 个阵列扬声器同时播放,方法与 B.3.2 相同。

B.3.3.2 如某个阵列扬声器的单独校准不满足要求,应重新调整补偿滤波器。

B.3.3.3 阵列校准的频响应是平坦的,允许容差应在 ± 3 dB 之间。

B.3.4 重低音扬声器校准

B.3.4.1 重低音扬声器的频率范围宜为 30 Hz~120 Hz。

B.3.4.2 重低音扬声器的声压级应与噪声声压级一致。

B.3.4.3 采用粉伪随机噪声作为校准信号,结合低通滤波器(120 Hz, 18 dB/oct)和高通滤波器(50 Hz, 12 dB/oct)进行校准,方法同 B.3.2.4。

B.3.4.4 重低音扬声器校准的频响应是平坦的,允许容差应在 ± 3 dB 之间。

B.3.5 扬声器播放延时补偿

B.3.5.1 应对阵列扬声器的播放延时进行补偿,保证阵列扬声器对中心区域音频测试装置声场的相关性。阵列扬声器的最大播放延时不应大于 50 ms,延时可以为 0 ms。

B.3.5.2 重低音扬声器一般不需要延时补偿。

B.3.6 整体校准

B.3.6.1 采用粉伪随机噪声作为校准信号,将所有滤波器和延时补偿等设置开启,所有扬声器播放校准信号,音频测试装置进行噪声采集,然后用 1/3 倍频程进行分析。

B.3.6.2 如果在 50 Hz~20 kHz 频率范围之内,频率响应曲线的容差在 ± 3 dB 之间,则完成校准。否则,需按 B.3.2~B.3.5 步骤重新进行校准。

附录 C

(规范性附录)

基于神经网络的语音质量评价(NNSQE)方法

C.1 测试流程

C.1.1 基于神经网络的语音质量评价(NNSQE)方法是一种有参的客观语音评价方法。输入信号分为传输前语音和传输后语音,通过神经网络模型分析处理,输出语音得分。NNSQE 方法流程图如图 C.1 所示。

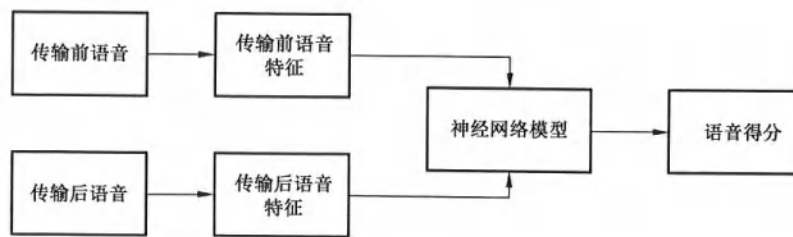


图 C.1 NNSQE 方法流程图

C.1.2 预处理

在预处理阶段,分别对传输前语音和传输后语音提取对应的语音特征,然后将提取的特征输入到神经网络模型。语音特征应包含信噪比(SNR)特征(附录 D 中 D.3.1)。

C.1.3 神经网络模型

神经网络模型是 NNSQE 方法的核心部分,它将传输前后的语音特征作为输入,输出的结果可以衡量传输前后的质量差异。NNSQE 方法所采用的神经网络模型应至少包含但不限于以下要求:输入网络层数应至少包含 3 层(如输入层、中间层、输出层);网络中间层节点数应不少于 50 个。

C.1.4 语音得分

语音得分是指神经网络模型的输出结果,是可以衡量语音质量好坏的得分。该得分用于评价源语音(传输前语音)经过采集、处理、传输、播放等操作后得到的语音(传输后语音)的质量。得分取值范围为 1 分~5 分。质量越好,分数越高。通过结合该得分与人工标注得分进行综合分析,也可实现 NNSQE 方法的准确性评价。

C.2 NNSQE 方法的准确性评价

C.2.1 准确性衡量参数

NNSQE 方法的准确性衡量参数采用平均绝对误差(Mean Absolute Error, MAE)、均方根误差(Root Mean Squared Error, RMSE)和相关系数。

按公式(C.1)计算平均绝对误差:

$$\Delta = (|x_1 - y_1| + |x_2 - y_2| + \dots + |x_n - y_n|) / n \quad \dots\dots\dots(C.1)$$

按公式(C.2)计算均方根误差:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots (C.2)$$

按公式(C.3)计算相关系数:

$$R(X, Y) = Cov(X, Y) / \sqrt{(Var(X) \times Var(Y))} \dots\dots\dots (C.3)$$

C.2.2 准确性衡量参数提取

C.2.2.1 NNSQE 方法的准确性衡量参数提取及评价应按图 C.2 所示方案进行。

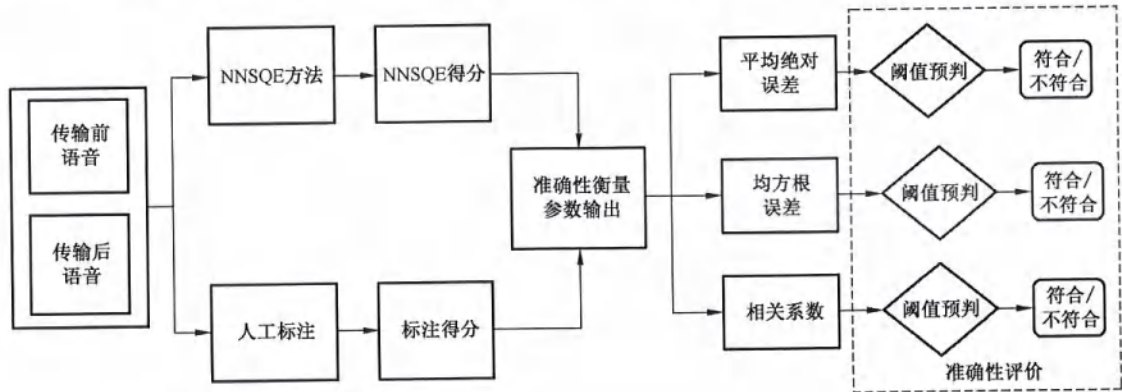


图 C.2 NNSQE 方法的准确性衡量参数提取及评价流程图

C.2.2.2 根据 ITU-T P.830 评分标准要求,组织语音专家和普通标注人员对传输后语音进行评分标注。标注人员由语音专家和普通标注人员组成,语音专家和普通标注人员具有不同的评分权重,对语音专家和普通标注人员的评分进行加权平均后作为测试语音的标注得分。测试语音集中,传输前语音和传输后语音成对出现。测试语音应满足表 C.1 的要求。

表 C.1 NNSQE 方法准确性评价的测试语音要求

序号	要求描述	说明
1	语音条目数≥500 条	传输前语音
2	每条语音长度≥5s	传输前语音
3	每条语音的起始静音≥0.25 s	传输前语音
4	每条语音的尾部静音≥0.25 s	传输前语音
5	每条语音的采样率≥8 kHz	传输前语音
6	测试语音应覆盖不同的性别	传输前语音
7	每条语音的标注人数≥3 人	传输后语音
8	人工标注得分精度 0.1 分以上	传输后语音
9	语音的质量分数在 1 分~5 分之间分布均匀	传输后语音

C.2.2.3 按图 C.2 所示流程,采用 NNSQE 方法对传输前语音和传输后语音进行评分,得到 NNSQE 得分。

C.2.2.4 然后对 NNSQE 得分和人工标注得分进行分析处理,按公式(C.1)、公式(C.2)、公式(C.3)分别计算出平均绝对误差、均方根误差以及相关系数等准确性衡量参数。

C.2.2.5 最后根据准确性衡量参数的指标要求,对衡量参数进行符合性评价。NNSQE 方法的准确性评价应在与训练数据(附录 D 中 D.2)同源的测试数据集合上进行。根据论证实验结果,结合 ITU-T P.863—2018、ETSI EG 202 396-3 等标准的相关要求,准确性衡量参数应满足以下指标要求:平均绝对误差(Δ) ≤ 0.4 ,均方根误差(RMSE) ≤ 0.5 ,相关系数 $[R(X,Y)] \geq 0.93$ 。

C.2.2.6 附录 D 给出了一种典型的满足准确性要求 NNSQE 方法——NN MOS 评分系统,其在测试数据集合上的准确性测试结果为:平均绝对误差 0.18,均方根误差 0.28,相关系数 0.97。

附录 D
(资料性附录)

一种典型的 NNSQE 方法 —— NN MOS 评分系统

D.1 NN MOS 评分系统结构

NN MOS 评分系统是一种典型的 NNSQE 方法,其结构如图 D.1 所示:

- 网络模型:BiLSTM 网络;
- 网络层数:5 层(1 个输入层、3 个隐藏层、1 个输出层);
- 输入层节点数:语音特征维数,包括信噪比、线性预测系数及其参数、基于谱距离评价的特征参数、基于声学感知的特征参数、基于中文音素的瓶颈特征;
- 隐藏层节点数:1 024 个;
- 输出层节点数:1 个。

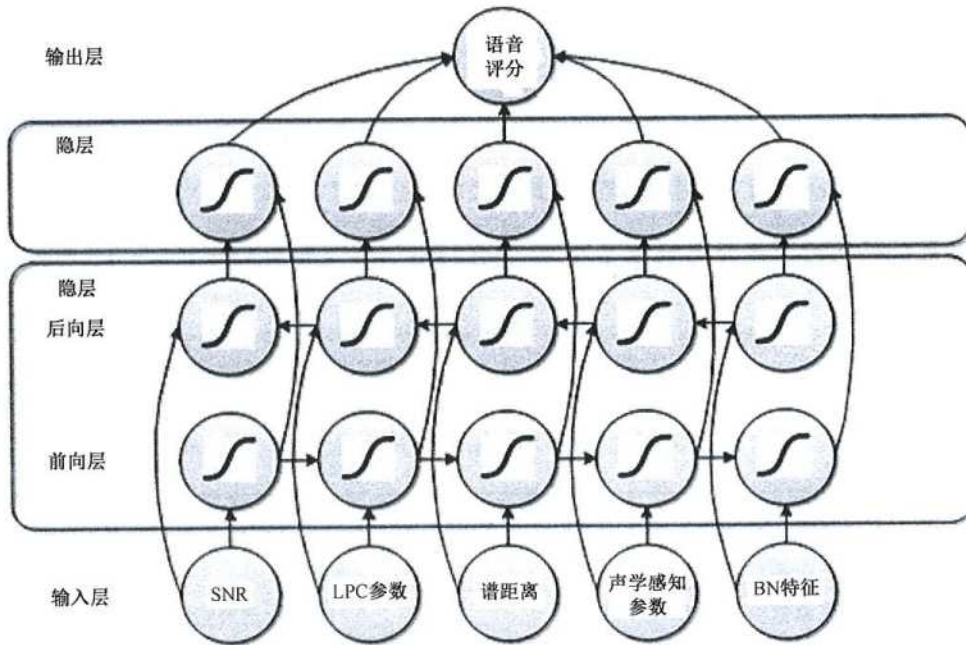


图 D.1 NN MOS 评分系统结构示意图

D.2 训练数据

D.2.1 训练数据采集

NN MOS 评分系统的训练数据来源为真实应用场景的录音,具体采集方法如下:选取应用场景,通过特定应用对讲系统,进行模拟通话,录制通话语音。训练数据集包含了本标准涉及特定应用场景的通话录音,训练数据录音时长达到上千小时。

D.2.2 训练数据时长

具体各场景下的训练数据时长分布如表 D.1。

表 D.1 训练数据各场景数据时长

应用场景大类	应用场景分类	数据时长/h
司法 公安	公共区域报警对讲	60
	监室对讲	40
	监狱探访对讲	46
	审讯室对讲	38
	岗亭对讲	60
	景区报警对讲	53
金融	银行自助设备求助对讲	42
	银行窗口对讲	56
	防尾随门对讲	56
交通	地铁/动车车厢对讲	42
	站台求助对讲	39
	高速公路对讲	53
医疗 养老	病房医护对讲	50
	手术室对讲	37
	ICU(重症加强护理病房)对讲	39
	养老监护对讲	45
工业	工业企业内部对讲(车间及嘈杂区域)	58
	工业企业内部对讲(办公区域)	60
居住建筑 办公建筑	电梯五方对讲	51

D.2.3 训练数据标注

训练数据标注方法同附录 C 中 C.2.2.2。根据 ITU-T P.830 评分标准对训练数据进行人工标注,标注人员由语音专家和普通标注人员组成,语音专家和普通标注人员具有不同的评分权重,对语音专家和普通标注人员的评分进行加权平均后作为训练语音的评分。

D.3 语音特征提取

D.3.1 特征参数

NN MOS 评分系统提取的听觉感知特征参数包括信噪比、线性预测系数及其参数、基于谱距离评价的特征参数、基于声学感知的特征参数、基于中文音素的瓶颈特征等 5 类。

D.3.2 信噪比(SNR)

信噪比的计量单位是 dB,其计算方法如公式(D.1):

$$SNR = 10 \times \text{Log}_{10}(\text{signal}/\text{noise}) \dots\dots\dots (D.1)$$

其中 signal 和 noise 分别表示语音信号和噪声的功率。为获取测试语音的信噪比信息,首先应对原语音和测试语音进行对齐,从而获取语音段以及非语音段信息,然后再根据非语音段的信息估算噪声信号。

D.3.3 线性预测系数(LPC)及其参数

线性预测分析原理认为系统的传递函数符合全极点数字滤波器的形式,因此 t 时刻的信号可以用前若干时刻的信号线性组合来估计。当实际语音的采样值和线性预测采样值之间达到均值方差最小时,即可得到线性预测系数 LPC。通过线性预测分析得到的合成滤波器的系统函数公式(D.2):

$$H(z) = G / (1 - \sum_{i=1}^n (a_i \times z^{-i})) \quad \dots\dots\dots (D.2)$$

D.3.4 基于谱距离评价的特征参数

包括语音频谱的均值、方差、最大值、最小值、一阶差分、二阶差分等参数。其中差分特征的计算公式见公式(D.3):

$$d_t = \frac{\sum_{\theta=1}^{\theta} \theta (c_{t+\theta} - c_{t-\theta})}{2 \sum_{\theta=1}^{\theta} \theta^2} \quad \dots\dots\dots (D.3)$$

D.3.5 基于声学感知的特征参数

D.3.5.1 感知线性预测(PLP)特征

PLP 特征参数是全极点模型预测多项式的一组系数,等效于一种 LPC 系数特征。PLP 特征分析将人耳听觉模型进行工程化处理,应用到频谱分析中。将输入的语音信号经听觉模型处理后所得到的信号替代传统的 LPC 分析所用的时域信号,更有利于抗噪语音特征提取。PLP 特征的提取流程如图 D.2 所示。

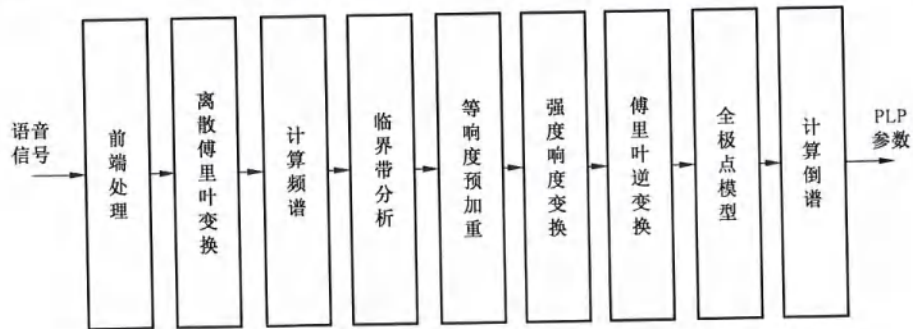


图 D.2 PLP 特征提取流程图

D.3.5.2 梅尔频率倒谱系数(MFCC)特征

MFCC 特征提取包含两个关键步骤:转化到梅尔频率,然后进行倒谱分析。梅尔刻度是一种基于人耳对等距的音高(pitch)变化的感官判断而定的非线性频率刻度。梅尔刻度和频率刻度关系如公式(D.4):

$$m = 2595 \log_{10} \left(1 + \frac{f}{700} \right) \quad \dots\dots\dots (D.4)$$

当在梅尔刻度上是均匀的话,对应频率的距离将会越来越大,梅尔刻度三角滤波器组如图 D.3 所示。

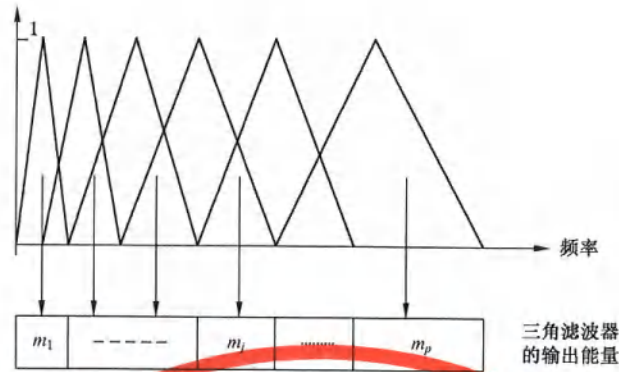


图 D.3 梅尔刻度三角滤波器组

在倒谱分析部分时,首先对时域信号进行傅里叶变换转换到频域,然后再利用梅尔刻度滤波器组将对应频域信号进行切分,最后每个频段对应一个数值,即为倒谱系数。具体的 MFCC 特征提取过程见图 D.4。

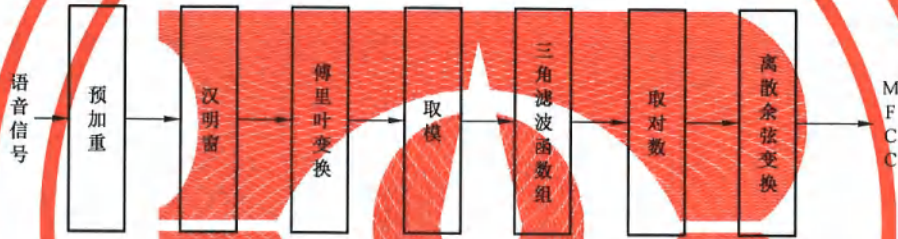


图 D.4 MFCC 特征提取流程图

D.3.6 基于中文音素的瓶颈(BOTTLE NECK)特征

通过中文语音识别深度神经网络提取的包含中文音素特性的 BN 特征。利用基于音素状态信息的提取深瓶颈特征的方法,训练以音素状态为网络输出的,在设计网络结构时为了得到合适维度的底层声学特征,设置一个节点数目较少的隐层,由于其在结构上像是网络的一个瓶颈,因此将该层形象的称为瓶颈层。在网络训练完成以后,就在底层输入特征与输出音素状态之间建立了深层的信息提取关系,底层声学特征中与音素相关的信息被逐层加以抽象,而与音素无关的信息则被逐渐加以抑制。具体的网络结构如图 D.5 所示。

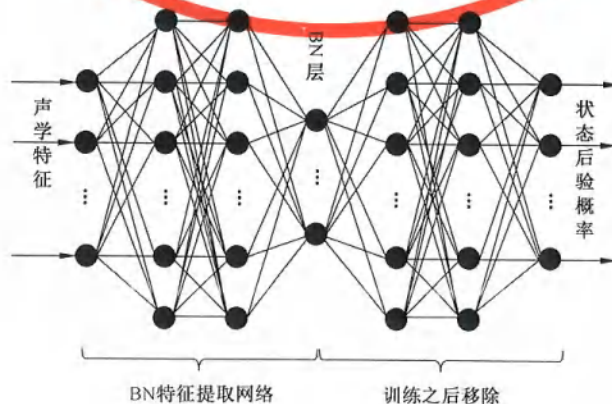


图 D.5 中文音素的瓶颈特征提取网络

D.4 NN MOS 评分系统算法原理

D.4.1 双向长短时记忆(BILSTM)网络建模

考虑到语音的时序特性,采用时序模型双向长短时记忆(BILSTM)网络对语音进行建模,网络整体结构如图 D.1 所示。对语音逐帧提取的不同特征拼接成特征向量 X_t ,逐帧送入 BILSTM 网络单中,其中包括一个前向 LSTM 网络以及一个后向 LSTM 网络(前两个隐层),然后将前两层隐层输出拼接输出到第三层隐层,预测每一帧语音的 MOS 分;最后,输出层对每帧得分进行平均得到最终的语音 MOS 分。

D.4.2 BILSTM 网络结点

BILSTM 网络主要由 LSTM 的门结构组成,门结构如图 D.6 所示,包括细胞(Cell)、输入门(Input-Gate)、输出门(OutputGate)和遗忘门(ForgetGate)。对于前两层隐层,LSTM 网络节点的输入 X_t 即为提取的 t 时刻的音频特征,输出为隐层特征 h_t ;对于第三层隐层,输入前两层隐层的隐层特征,输出 h_t 即为 t 时刻所对应的 MOS 分。

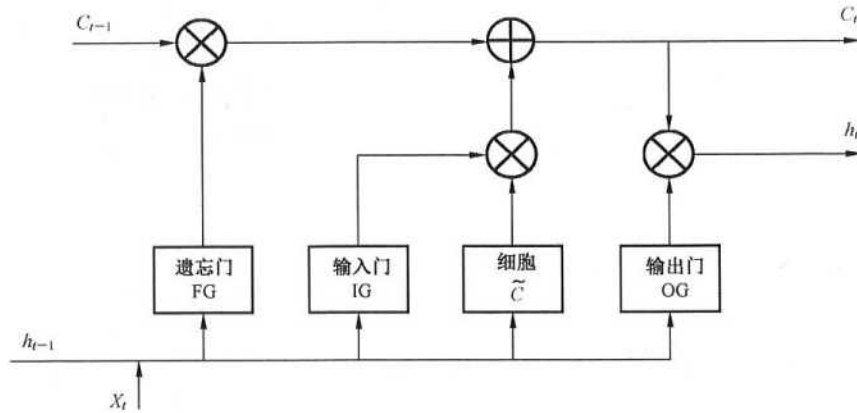


图 D.6 BILSTM 门结构示意图

D.4.3 NN MOS 评分系统原理

D.4.3.1 特征归一化

将语音特征每个维度减去该维度的均值并除以标准差,如公式(D.5)。

$$fea'_i = (fea_i - \mu_i) / \sigma_i \quad \dots\dots\dots (D.5)$$

D.4.3.2 前向计算

将归一化特征参数输入 BILSTM 网络,进行前向计算,如公式(D.6)。

$$\begin{aligned} IG_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ FG_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ \tilde{C}_t &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \\ C_t &= FG_t \times C_{t-1} + IG_t \times \tilde{C}_t \quad \dots\dots\dots (D.6) \\ OG_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t &= OG_t \times \tanh(C_t) \\ MOS &= AveragePooling(N_1, \dots, N_i) \end{aligned}$$

注：其中， IG_t 是 t 时刻输入门的输出， x_t 是 t 时刻上一层的输入， h_{t-1} 是 $t-1$ 时刻该节点的输出， C_{t-1} 是 $t-1$ 时刻 Cell 的输出； FG_t 是 t 时刻输入门的输出； C_t 是 t 时刻 Cell 的输出； OG_t 是 t 时刻输出门的输出； h_t 是 t 时刻该节点的输出， σ 为激活函数 sigmoid。其中， W_i, W_f, W_c, W_o 为 BILSTM 网络模型中需要训练的模型参数。

D.4.3.3 MOS 评分

将训练集通过前向计算得到的最终 MOS 评分，并将输出评分与标注分数进行分析处理，得到残差值及网络的损失函数。

D.4.3.4 更新权重

通过残差值进行反向传播计算，计算网络各层权值 W_i, W_f, W_c, W_o 的梯度，对网络权重进行更新，从而实现网络优化。更新网络权重如公式(D.7)，其中 α 为学习率。

$$W_{new} = W_{old} - \alpha \Delta W \quad \dots\dots\dots (D.7)$$

D.5 NN MOS 评分系统准确性评价

按附录 C 中图 C.2 流程图对 NN MOS 评分系统进行准确性评价，准确性评价结果如图 D.7 所示。结果表明，在和训练数据同源的测试数据集合上测试结果平均绝对误差 0.18，均方根误差 0.28，相关系数 0.97。平均绝对值误差 < 0.4 占 97.18%。NN MOS 评分系统满足附录 C 中 C.2 的准确性评价要求，适用于本部分。

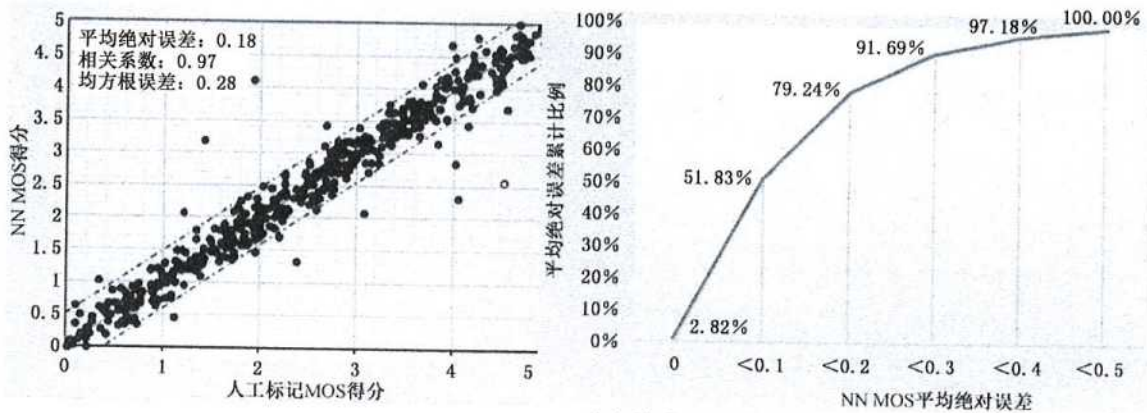


图 D.7 NN MOS 评分系统准确性评价结果

参 考 文 献

- [1] ETSI ES 202 396-1 V1.6.1 (2015-06) Speech and multimedia Transmission Quality (STQ); Speech quality performance in the presence of background noise; Part 1: Background noise simulation technique and background noise database
- [2] ITU-T P.863.1—2019 Application guide for Recommendation
-

中华人民共和国
国家标准
楼宇对讲系统
第3部分：特定应用技术要求
GB/T 31070.3—2021

*

中国标准出版社出版发行
北京市朝阳区和平里西街甲2号(100029)
北京市西城区三里河北街16号(100045)

网址 www.spc.net.cn

总编室:(010)68533533 发行中心:(010)51780238

读者服务部:(010)68523946

中国标准出版社秦皇岛印刷厂印刷
各地新华书店经销

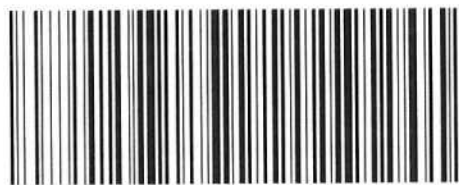
*

开本 880×1230 1/16 印张 2.5 字数 71 千字
2021年5月第一版 2021年5月第一次印刷

*

书号: 155066·1-67718 定价 36.00 元

如有印装差错 由本社发行中心调换
版权专有 侵权必究
举报电话:(010)68510107



GB/T 31070.3-2021



码上扫一扫 正版服务到