

团 体 标 准

T/GDEIIA XX—2022

面向物联网家居产品具有联网功能的语音 模块的技术规范

Technical specification of voice module with networking function for
IoT home products

(征求意见稿)

2022 – XX – XX 发布

2022 – XX – XX 实施

广东省电子信息行业协会 发布

目 次

前言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 目的及系统边界	2
5 总体要求	3
6 实施与控制	3
6.1 语音采集	3
6.2 语音识别	3
6.3 语音合成	4
6.4 交互成功率	4
6.5 误唤醒率	4
6.6 语音打断	4
6.7 响应时间（语音）	4
6.8 响应时间（广域网和局域网）	4
6.9 语音输入与 APP 操作互通控制	5
6.10 指令转换	5
6.11 蓝牙模块性能	5
6.12 蓝牙配网	5
6.13 SoftAP 配网	5
6.14 Wi-Fi 模块性能	5
6.15 Wi-Fi 联网	6
6.16 Wi-Fi 兼容性	6
6.17 信息安全	6
6.18 电磁兼容性	6
6.19 环境可靠性	6

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。

本文件由深圳市中易腾达科技股份有限公司提出。

本文件由广东省电子信息行业协会归口。

本文件起草单位：深圳市中易腾达科技股份有限公司、上海亮牛半导体科技有限公司、佛山电器照明股份有限公司、华勤技术股份有限公司。

本文件主要起草人：赖厚先、罗宏洲。

本文件为首次发布。

面向物联网家居产品具有联网功能的语音模块的技术规范

1 范围

本文件规定了物联网家居产品中具有联网功能的语音模块的语音识别的术语和定义、功能要求、性能要求，射频性能要求，联网功能要求，配网功能要求，语音声控功能与APP操作互通逻辑控制要求，信息安全，电磁兼容性，环境可靠性。

本文件适用于既有语音声控功能又具备联网功能的家居产品。

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 2423.1-2008 电工电子产品环境试验 第2部分：试验方法 试验A：低温
- GB/T 2423.2-2008 电工电子产品环境试验 第2部分：试验方法 试验B：高温
- GB/T 17626.2-2018 电磁兼容 试验和测量技术 静电放电抗扰度试验
- GB/T 17626.3-2016 电磁兼容 试验和测量技术 射频电磁场辐射抗扰度试验
- GB/T 17626.4-2018 电磁兼容 试验和测量技术 电快速瞬变脉冲群抗扰度试验
- GB/T 17626.5-2019 电磁兼容 试验和测量技术 浪涌（冲击）抗扰度试验
- GB/T 17626.6-2017 电磁兼容 试验和测量技术 射频场感应的传导骚扰抗扰度
- GB 15629.1101-2006 信息技术 系统间远程通信和信息交换 局域网和城域网特定要求 第11部分：无线局域网媒体访问控制和物理层规范：5.8 GHz频段高速物理层扩展规范
- GB 15629.1102-2003 信息技术 系统间远程通信和信息交换 局域网和城域网特定要求 第11部分：无线局域网媒体访问控制和物理层规范：5.8 GHz频段高速物理层扩展规范
- GB 15629.1104-2006 信息技术 系统间远程通信和信息交换 局域网和城域网特定要求 第11部分：无线局域网媒体访问控制和物理层规范：2.4 GHz频段更高数据速率扩展规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

语音识别

将人类的声音信号转化为文字或者指令的过程。

3.2

近场语音识别

以麦克风阵列近距离拾取的语音数据作为输入数据，通过语音识别的算法将语音信号转写成文字的技术。

3.3

远场语音识别

以麦克风阵列远距离拾取的语音数据作为输入数据，通过语音识别的算法将语音信号转写成文字的技术。

3.4

语音识别率

将人类的声音信号转化为文字或者指令过程的成功率，一般用百分比表示。

3.5

唤醒词

使家居产品进入语音识别状态的指令。

3.6

唤醒率

在远场语音场景语音识别有效距离内，用户说出特定的语音指令（唤醒词），使家居产品进入语音识别状态的成功率，一般用百分比表示。

3.7

误唤醒率

语音中不存在唤醒词，但是将一些其他语音误判为语音指令（唤醒词）使家居产品进入语音识别状态的次数。

3.8

词条库

用于语音测试的词汇表。

3.9

测试距离

语音输入与被测设备中心点之间的距离。

3.10

SoftAP配网

即利用语音模块的无线芯片，将语音模块进入到softAP模式，开启一个无线局域网，手机（或其它移动设备或APP）通过连入语音模块开启的无线局域网后，向语音模块发送路由器的SSID及Passwords等信息，让语音模块在无屏幕的情况下，获取到路由器的ssid信息，达到联网的目的。

3.11

蓝牙配网

即利用语音模块的无线芯片，通过手机（或其它移动设备或APP）连接此语音模块的蓝牙，连接之后发送Wi-Fi连接所需要的SSID及Passwords，语音模块的无线芯片收到蓝牙信息后，通过SSID及Passwords去进行联网。

4 目的及系统边界

语音模块分为前端语音处理模块、语音处理模块、蓝牙处理模块、Wi-Fi处理模块和应用模块，其中：

- a) 前端语音处理模块包括音频降噪模块、语音唤醒模块、声纹识别模块，负责将语音输入转化为语音流，作为语音处理模块的输入，其中语音处理模块的输入。其中，音频降噪模块负责对拾音器取到的音频进行降噪、去混响、回声消除等处理。

- b) 语音处理模块包括语音识别模块和语音合成模块，其中语音识别模块负责将语音流转换为计算机可识别的文本信息并输出给相关的应用模块；语音合成模块负责将应用模块提供的计算机可识别的文本信息转换成语音流输出。
- c) 应用模块包括指令处理模块和家居应用，其中指令处理模块负责将语音识别模块获得的用户语音指令转换成家居硬件设备支持的命令。家居应用负责实际控制家电设备并给出响应信息；同时指令模块把此命令输出给Wi-Fi模块，由Wi-Fi模块传输到服务器端，并由服务器下发数据信息给APP或手机端，实现UI界面的信息同步。

5 总体要求

语音模块的功能集语音输入、蓝牙配网、Wi-Fi联网、APP控制、语音指令与APP操作互通逻辑与时间响应、各项性能指标以及安全性要求。语音模块系统图见图1。

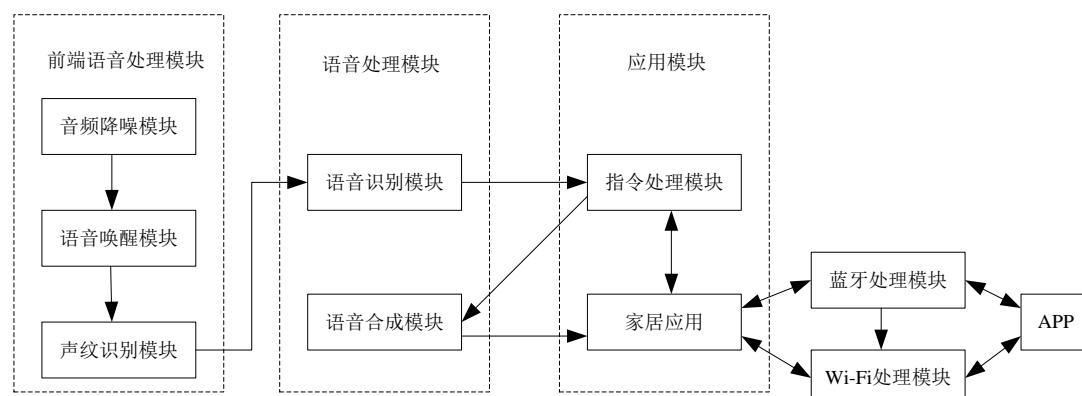


图1

6 实施与控制

6.1 语音采集

通过传声器或传声器阵列等具备语音采集能力的硬件设备对语音进行采集。语音模块的语音采集应支持远场语音采集，并可支持近场语音采集。

6.2 语音识别

语音识别的基本要求包括：

- 识别引擎应支持远场语音采集，并可支持近场语音采集。应支持命令字识别、连续语音识别中的一种。
- 在低噪家居环境（声音强度在52dB以下）中，语音识别句识别正确率（语音识别率）在不同的距离位置1米距离应大于等于95%，3米距离应大于等于93%，5米距离应大于等于90%。
- 在高噪家居环境（声音强度在62dB±3dB）中，语音识别句识别正确率（语音识别率）在不同的距离位置1米距离应大于等于93%，3米距离应大于等于91%，5米距离应大于等于89%。
- 距离如图2所示，其中一个高保真音箱用于播放背景噪声，语音模块放置在MIC处不变，移动另外一个高保真音箱的位置，分别在1米、3米和5米的位置进行测试的值。

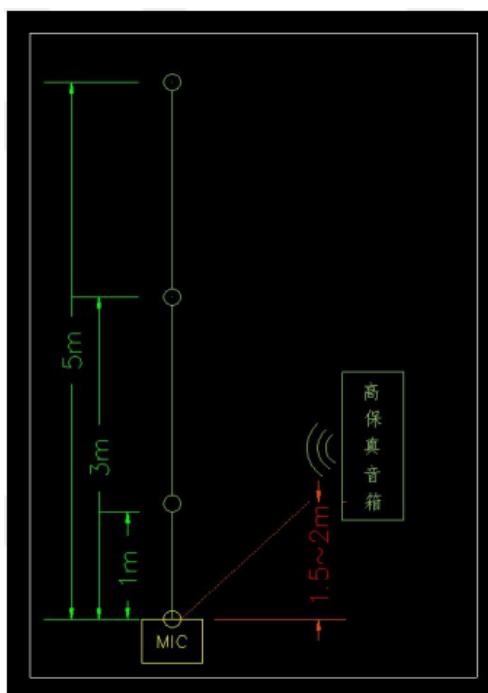


图2

6.3 语音合成

应支持汉语普通话，宜支持英语以及粤语或其他语言，宜支持多音色合成和个性化合成，其中多音色应支持16~65周岁。

6.4 交互成功率

语音模块在语音交互系统的交互过程中，交互目的能够在交互轮次内完成，即可判定本轮交互有成功，反之则判定为本轮交互失败。

交互成功率=[交互成功的次数/(交互成功的次数+交互失败的次数)]×100%

6.5 误唤醒率

语音模块处在待机状态下，计算在没有唤醒词相关唤醒词情况下的误唤醒，在安静家居环境（声音强度在45dB以下）中，误唤醒次数≤1次/24Hr，在噪音家居环境（声音强度>60dB）中，误唤醒次数≤3次/24Hr。

6.6 语音打断

应支持交互过程中的语音打断，实现交互速度与自然度的提高，在语音交互过程中，语音打断成功率的计算方法：被语音交互系统正确响应的次数/交互内容中需要执行打断操作的次数×100%

6.7 响应时间（语音）

平均响应时间应该小于或等于500ms，其中，响应时间的计算方法为语音播报给出结果的时刻-语音输入结束的时刻。

6.8 响应时间（广域网和局域网）

平均响应时间应该小于或等于2s，其中，响应时间的计算方法为：
APP界面上显示状态变化的时刻-语音输入结束的时刻。

网络条件应满足上行带宽不低于100kbit/s，下行带宽不低于50kbit/s，且应保持稳定的连通状态。

6.9 语音输入与 APP 操作互通控制

语音模块在成功处理一次语音的指令输入后，手机APP（或移动设备）UI 操作界面在响应时间（广域网和局域网状态下）内会有对应的状态变化，且要能成功返回状态值上报到服务器端。

6.10 指令转换

应支持将用户的唤醒词转换成家居设备支持的控制命令，同时手机APP（或移动设备）也能实现对家居设备的控制。

6.11 蓝牙模块性能

应支持Bluetooth LE5.1版本，物理层要符合蓝牙SIG的规范。

在屏蔽室的环境下，蓝牙反复连接次数100次，成功率达到100%，在室内空旷距离的环境下，10米距离控制丢包率 $\leq 1\%$ 。

6.12 蓝牙配网

在屏蔽室，语音模块先进入配网，连接次数100次的状态下，蓝牙配网成功率（拿到SSID和Password） $\geq 98\%$ ，蓝牙配网耗时 $\leq 25\text{ms}$ 。

在室外空旷距离的环境下，蓝牙配网拉距最差面 ≥ 50 米，最优面 ≥ 100 米。

在室内空旷距离的环境下，蓝牙配网拉距最差面 ≥ 30 米，最优面 ≥ 50 米。

6.13 SoftAP 配网

在屏蔽室，SoftAP配网成功率 $\geq 95\%$ ，配网耗时 $\leq 25\text{s}$ 。

在弱网配网（信号强度 $\leq -70\text{dBm}$ ），配网100次，成功率应 $\geq 60\%$ 。

6.14 Wi-Fi 模块性能

应支持 802.11 b/g/n 三个频段，每个频段的性能要符合GB 15629.1104-2016 中规定。

在室内空旷距离的环境下：

- a) 语音模块从开始连接路由器到连接成功时间 $\leq 6\text{s}$ 。
- b) Ping丢包率，语音模块信号强度 $\leq -60\text{dbm}$ ，ping大小32Bytes，ping包2000次，丢包率 $< 5\%$ ，延时时间 $< 50\text{ms}$ 。
- c) Ping丢包率，语音模块信号强度 $\leq -60\text{dbm}$ ，ping大小512Bytes，ping包2000次，丢包率 $< 5\%$ ，延时时间 $< 70\text{ms}$ 。
- d) 语音模块出现异常到恢复路由器连接（包括重启时间）时间，测试10次，时间 $< 20\text{s}$ 。
- e) 语音模块掉电重启测试10次，自动恢复连接能力，路由器连接成功率达到100%。
- f) 语音模块失败自动重连测试10次，语音模块连接路由器失败时，自动重连连接成功率达到100%。
- g) 语音模块网络恢复能力测试10次，语音模块正常连接路由器，重启路由器，语音模块回连成功率达到100%。
- h) 语音模块所连接的路由器长时间断电后，模块自动重连（期间不断电）测试3次，路由器断电12小时后，查看APP端显示模块离线，此时路由器上电，模块应主动重连网络，重连后模块在线可控，回连成功率100%。

- i) 语音模块所连接的路由器长时间断电后，模块自动重连（期间断电）测试3次，路由器断电5分钟后，模块断电重启，继续等待1小时，查看APP端显示模块离线，此时路由器上电，模块应主动重连网络，重连后模块在线可控，回连成功率100%。
- j) 路由器配置变更测试10次，语音模块正常连接路由器，改变路由器的无线相关的配置（除SSID和密码，包括模式、信道、恢复出厂设置再重新配置）后，测试模块自动回连功能，回连成功率100%。
- k) 语音模块大数据包测试（Payload \geq 1400bytes），APP发送大数据包给模块，监测是否正常返回，测试10次，成功处理并响应，成功率 \geq 95%。

6.15 Wi-Fi 联网

6.15.1 局域网设备自动上线平均时间，语音模块局域网正常在线，重启模块，测试次数 50 次，模块自动上线成功率 $>$ 95%，模块自动上线平均时间 $<$ 10s，控制一次为成功。

6.15.2 广域网设备自动上线平均时间，语音模块局域网正常在线，重启模块，测试次数 50 次，模块自动上线成功率 $>$ 95%，模块自动上线平均时间 $<$ 10s，控制一次为成功。

6.15.3 语音模块 OTA 升级：

- a) 升级时间语音模块信号强度 $>$ -50dBm，升级10次，时间 \leq 60s。
- b) 升级成功率语音模块信号强度 $>$ -50dBm，升级100次，成功率100%，

6.15.4 正常长运，语音模块信号强度 $>$ -60dBm，长运 7 天，每天掉线次数 $<$ 1 次，重连时间 \leq 300s。

6.16 Wi-Fi 兼容性

手机兼容性，语音模块应支持兼容主流手机和操作系统 \geq 95%。

路由器兼容性，语音模块应支持兼容主流路由器品牌和型号 \geq 95%。

APP版本兼容，语音模块应支持所有APP版本 \geq 100%。

6.17 信息安全

应支持

OPEN/WEP-OPEN/WEP-SHARED/WPA-PSK(TKIP)/WPA-PSK(AES)/WPA2-PSK(TKIP)/WPA2-PSK(AES)/WPA2-PSK(TKIP-AES混合加密)/WPA3-SAE 加密方式。

应支持篡改加密方式不允许响应，使用安全测试功能，篡改加密方式的指令不允许响应。

支持非认证数据包不允许携带随机数。

支持模拟恶意抓包控制，不允许响应。

6.18 电磁兼容性

6.18.1 蓝牙 SRRC 应符合工信部【2002】353 号文件中的规定。

6.18.2 Wi-Fi SRRC 应符合工信部【2002】353 号文件、工信部【2012】620 号文件、工信部【2002】277 号文件和 GB 15629.1102-2003、GB 15629.1104-2006、GB 15629.1101-2006 中的规定。

6.18.3 静电放电抗扰度试验应符合 GB/T 17626.2-2018 的规定。

6.18.4 射频电磁场辐射抗扰度试验应符合 GB/T 17626.3-2016 的规定。

6.18.5 电快速瞬变脉冲群抗扰度试验应符合 GB/T 17626.4-2018 的规定。

6.18.6 浪涌（冲击）抗扰度试验应符合 GB/T 17626.5-2019 的规定。

6.18.7 射频场感应的传导骚扰抗扰度应符合 GB/T 17626.6-2017 的规定。

6.19 环境可靠性

6.19.1 高温贮存

试验设备应符合GB/T 2423.2-2008中的有关规定。

语音模块在不包装，不上电的状态下，放入恒温恒湿箱内，恒温恒湿箱内的温度逐步升温至 $65^{\circ}\text{C} \pm 2^{\circ}\text{C}$ ，湿度 $60\% \pm 5\% \text{RH}$ ，搁置时间24h，然后取出在常温下恢复2h，检查外观，各项功能应正常。

6.19.2 低温贮存

试验设备应符合GB/T 2423.1-2008中的有关规定。

语音模块在不包装，不上电的状态下，放入烤箱内，烤箱内的温度逐步上升，逐渐下降至 $-20^{\circ}\text{C} \pm 2^{\circ}\text{C}$ ，搁置时间24h，然后取出在常温下恢复2h，检查外观，结构，各项功能应正常。

6.19.3 高温运行

语音模块在不包装，上电的状态下，放入恒温恒湿箱内，恒温恒湿箱内的温度逐步升温至 45°C ，湿度 $60\% \pm 5\% \text{RH}$ ，每隔1h取出测试语音输入及联网功能是否正常（重启模块测试10次，允许掉线1次，25s内需要连接成功），循环测试8h，最后取出检查外观，各项功能应正常。

6.19.4 低温运行

语音模块在不包装，上电的状态下，放入恒温恒湿箱内，恒温恒湿箱内的温度逐步下降至 -20°C ，每隔1h取出测试语音输入及联网功能是否正常（重启模块测试10次，允许掉线1次，25s内需要连接成功），循环测试8h，最后取出检查外观，各项功能应正常。
