ICS 33.160.99
CCS L 84

T/SILA

上 海 浦 东 智 能 照 明 联 合 会 团 体 标 准

T/SILA 006—2022

面板本地语音控制规范

Offline voice control specification of smart panel

2022－08－03 发布　　2022－08－03 实施

上海浦东智能照明联合会　　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　本地语音控制系统架构和要求 1
4.1　本地语音控制概述 1
4.2　系统组成 1
4.3　本地语音控制面板功能 2
5　唤醒 2
5.1　唤醒词 2
5.2　个性化唤醒词 2
5.3　退出唤醒时长 2
6　语音口令 2
6.1　一般要求 2
6.2　意图语料列表 2
7　串口协议 7
7.1　通信要求 7
7.2　工作模式 7
7.3　格式定义 8
8　测试和验收 11
8.1　测试配置 11
8.2　测试方法 12
8.3　验收标准 13
8.4　噪声类型 14
参考文献 16

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。

本文件由上海浦东智能照明联合会提出并归口。

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。

本文件起草单位：上海浦东智能照明联合会、杭州博联智能科技股份有限公司、成都启英泰伦科技有限公司、上海华镇电子科技有限公司、北京探境科技有限公司、欧普照明股份有限公司、欧智通科技股份有限公司、利尔达科技集团股份有限公司、北京如影智能科技有限公司、广东三雄极光照明股份有限公司、上海时代之光照明电器检测有限公司、广州力控智电信息科技有限公司、贝云(厦门)智能科技有限公司、惠州雷士光电科技有限公司、横店集团得邦照明股份有限公司、上海三思电子工程有限公司、江苏英索纳通信科技有限公司、广州河东科技有限公司、佛山照明智达电工科技有限公司、宁波雷利照明科技有限公司、深圳市奇脉电子技术有限公司、上海屹店智能科技有限公司、广东顺德胜上智能家居有限公司、珠海雷特科技股份有限公司、宁波赛耐比光电科技有限公司。

本文件主要起草人：刘宗孺、庄晓波、安波、曾义、陈磊、陈龙、陈镇、代照亮、程朝炜、刘金冬、郭磊、何振超、黄嵩、解创、康海洋、李金晶、李坤坤、廖冬年、林小科、凌志云、刘红彬、马小平、南宫荣、钱庄、苏伟健、孙岛、汤超、翁步升、叶飞、邹联报、洪艳君、方俊华。

面板本地语音控制规范

1 范围

本文件规定了面板本地语音控制的唤醒方式，语料口令，响应等技术要求。
本文件适用于智能家居场景的面板本地语音控制。

2 规范性引用文件

本文件没有规范性引用文件。

3 术语和定义

下列术语和定义适用于本文件。

3.1
唤醒 wake up
用户通过一个固定的单词，让面板进入语音对话的状态。

3.2
意图 intent
用户发起口令的目的。

3.3
本地语音 offline voice control
不依赖云平台进行语音识别和意图理解。

3.4
语料 text corpus
意图的具体口令说法。

4 本地语音控制系统架构和要求

4.1 本地语音控制概述

本地语音控制是指语音口令的识别和意图识别不依赖云平台，所有算法都在产品以离线形式在内部完成。

4.2 系统组成

面板本地语音系统主要构成见图1：

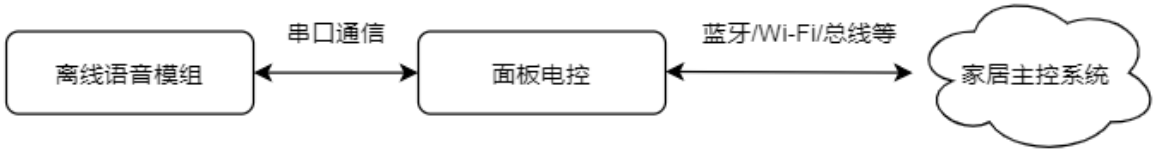

图1 系统构成

说明：
系统包括三部分：离线语音模组，面板电控，家居主控系统。
离线语音模组负责与用户进行对话，识别用户的意图，转换为串口指令，发送给面板电控。
面板电控收到指令后，按自己产品协议将相关意图进行适当的变换后，通过本地通信方式发送给家居主控系统进行控制家居设备或者直接完成家居设备的相关控制。

4.3 本地语音控制面板功能

本地语音控制面板功能如下：

a) 本地语音控制面板可以通过唤醒词进行唤醒
b) 本地语音控制面板要求面板被唤醒后才可以接收用户的语音口令。
c) 本地语音控制面板要求唤醒后在规定的时间段内支持连续口令控制。
d) 本地语音面板在一定时间内无语音控制命令进入待机模式。

5 唤醒

5.1 唤醒词

本文件要求各语音模组内置一个默认唤醒词：智能面板。

5.2 个性化唤醒词

各产品可以根据实际需要，定制产品个性化唤醒词。

5.3 退出唤醒时长

各产品可以根据实际需要设计退出唤醒时间，默认要求为15秒。

6 语音口令

6.1 一般要求

要求针对相同的意图，语音模组输出的口令编号和意图编号相同。

要求资源够用的情况下，语音模组按优先级内置推荐的意图列表。

要求语音模组的口令编号和意图编号统一分配和使用相同编码规则。

针对相同的意图，语音模组推荐使用默认的语料；可以根据资源和算法情况调整语料的顺序和个别单词。比如“打开主卧的灯”这个意图对应的语料也可以是“主卧灯打开”。

口令回复由产品自行定义。

6.2 意图语料列表

6.2.1 动作列表

动作列表见表1：

表1 动作列表

动作编号 （10 进制）	动作
1	打开
2	关闭
3	调亮
4	调暗

6.2.2 空间列表

空间列表见表2：

表2　空间列表

空间编号 （10 进制）	空间
0	无
1	主卧
2	次卧
3	客厅
4	玄关
5	卫生间
6	阳台

6.2.3　品类和场景列表

品类和场景列表见表3：

表3　品类和场景列表

品类/场景编号(10 进制)	品类/场景
1	灯
2	筒灯
3	台灯
4	射灯
5	灯带
6	顶灯
7	吊灯
8	风扇灯
9	所有的/全部的灯
10	回家
11	离家
12	睡眠
13	起床

6.2.4　推荐意图列表

推荐意图列表见表4：

表4　推荐意图列表

口令编号 （10 进制）	意图以及默认语料	意图编号 （动作+空间+品类场景，16 进制）
1	打开所有的灯	010009（打开+无空间+所有的灯），下同不再赘述。
2	关闭所有的灯	
3	开灯	
4	关灯	
5	打开回家场景	
6	打开离家场景	
7	打开睡眠场景	
8	打开起床场景	
9	我出门了	
10	我回家了	
11	我睡觉了	
12	我起床了	
13	打开筒灯	
14	打开台灯	
15	打开射灯	
16	打开灯带	
17	打开顶灯	
18	打开吊灯	
19	打开风扇灯	
20	关闭筒灯	
21	关闭台灯	
22	关闭射灯	
23	关闭灯带	
24	关闭顶灯	
25	关闭风扇灯	
26	关闭吊灯	
27	调亮筒灯	
28	调亮台灯	
29	调亮射灯	
30	调亮灯带	

表4　推荐意图列表（续）

口令编号 （10 进制）	意图以及默认语料	意图编号 （动作+空间+品类场景，16 进制）
31	调亮顶灯	
32	调亮吊灯	
33	调亮风扇灯	
34	调暗筒灯	
35	调暗台灯	
36	调暗射灯	
37	调暗灯带	
38	调暗顶灯	
39	调暗吊灯	
40	调暗风扇灯	
41	打开主卧的灯	
42	关闭主卧的灯	
43	打开次卧的灯	
44	关闭次卧的灯	
45	打开客厅的灯	
46	关闭客厅的灯	
47	打开玄关的灯	
48	关闭玄关的灯	
49	打开卫生间的灯	
50	关闭卫生间的灯	
51	打开书房的灯	
52	关闭书房的灯	
53	打开阳台的灯	
54	关闭阳台的灯	

6.2.5　参考意图列表

参考意图列表见表5：

表5　参考意图列表

口令编号（10 进制）	意图以及默认语料	意图编号（动作+空间+品类场景，16 进制）
100	打开主卧的筒灯	010102（打开+无空间+所有的灯），下同不再赘述。
101	打开主卧的台灯	

表5　参考意图列表（续）

口令编号（10 进制）	意图以及默认语料	意图编号（动作+空间+品类场景，16 进制）
102	打开主卧的射灯	
103	打开主卧的灯带	
104	打开主卧的顶灯	
105	关闭主卧的筒灯	
106	关闭主卧的台灯	
107	关闭主卧的射灯	
108	关闭主卧的灯带	
109	关闭主卧的顶灯	
110	调亮主卧的筒灯	
111	调亮主卧的台灯	
112	调亮主卧的射灯	
113	调亮主卧的灯带	
114	调亮主卧的顶灯	
115	调暗主卧的筒灯	
116	调暗主卧的台灯	
117	调暗主卧的射灯	
118	调暗主卧的灯带	
119	调暗主卧的顶灯	
120	打开次卧的筒灯	
121	打开次卧的台灯	
122	打开次卧的射灯	
123	打开次卧的灯带	
124	打开次卧的顶灯	
125	关闭次卧的筒灯	
126	关闭次卧的台灯	
127	关闭次卧的射灯	
128	关闭次卧的灯带	
129	关闭次卧的顶灯	
130	调亮次卧的筒灯	
131	调亮次卧的台灯	
132	调亮次卧的射灯	
133	调亮次卧的灯带	
134	调亮次卧的顶灯	
135	调暗次卧的灯	
136	调暗次卧的筒灯	
137	调暗次卧的台灯	
138	调暗次卧的射灯	
139	调暗次卧的灯带	
140	调暗次卧的顶灯	
141	打开客厅的吊灯	
142	打开客厅的顶灯	
143	打开客厅的风扇灯	

表5　参考意图列表（续）

口令编号（10 进制）	意图以及默认语料	意图编号（动作+空间+品类场景，16 进制）
144	打开客厅的筒灯	
145	打开客厅的台灯	
146	打开客厅的射灯	
147	打开客厅的灯带	
148	关闭客厅的吊灯	
149	关闭客厅的顶灯	
150	关闭客厅的风扇灯	
151	关闭客厅的筒灯	
152	关闭客厅的台灯	
153	关闭客厅的射灯	
154	关闭客厅的灯带	
155	调亮客厅的灯	
156	调亮客厅的吊灯	
157	调亮客厅的顶灯	
158	调亮客厅的风扇灯	
159	调亮客厅的筒灯	
160	调亮客厅的台灯	
161	调亮客厅的射灯	
162	调亮客厅的灯带	
163	调暗客厅的灯	
164	调暗客厅的吊灯	
165	调暗客厅的顶灯	
166	调暗客厅的风扇灯	
167	调暗客厅的筒灯	
168	调暗客厅的台灯	
169	调暗客厅的射灯	
170	调暗客厅的灯带	
171	打开空调	
172	关闭空调	
173	打开窗帘	
174	关闭窗帘	

7　串口协议

7.1　通信要求

串口通信参数，要求如下:

a)　波特率：9600 bit/s；
b)　数据位：8 bit；
c)　奇偶校验位：无；
d)　停止位：1；
e)　流控：无。

7.2　工作模式

设备支持“工作模式”配置和查询，可支持包括但不限于以下：

a) 0：独立设备（默认）；
b) 0x1：带空间属性设备；
c) 0x2：带类型属性设备；
d) 0x4：带独立主控模块属性设备；
e) ... 预留；
f) 0x80：独立扩展属性，不再按以上位配置属性。

说明：

示例：

0x81 使用《扩展通讯协议 1》
0x82 使用《扩展通讯协议 2》
0： 使用当前通讯协议的独立设备
0x3： 带空间和类型属性，但不带主控模块的应用场景

根据“工作模式”定义：

a) 不带0x80属性中，0（独立设备）是没有回复的；
b) 不带0x80属性中，除0外其余配置需要有指令回复；
c) 带0x80属性中，是否支持回复可根据类型单独进行约定，如无明确约定默认为支持回复指定。

7.3 格式定义

7.3.1 指令格式说明

语音口令在被语音模块处理完成后，通过二进制指令方式以产品自有的接口协议发送到接收端或者硬件总线模组。二进制格式由各口令内的编码按口令格式排列构成。

指令格式，包括：帧头+帧长+序号+设备+类型+内容+校验和。采用小端方式收发数据，若有多字节数据，低字节在前，见表6：

表6 指令格式

索引	名称	内容	值	说明
0	帧头	起始标志	0xFCA6	包的起始标志，以“0xFCA6”作为起始标志
2	帧长	数据长度	N	序号+设备+类型+内容+校验和的字节长度（N）
4	序号	包的序号	0～65535	每发送完一个消息，内容加1，用以检测丢包情况
6	设备	设备地址	0～65535	发送本条数据包的设备地址
8	类型	命令类型	0～255	命令类型的编号
9～N+2	内容	命令内容	N-6字节	命令附加的数据命令或参数
N+3	校验和	校验和	1字节	累加和校验方法，前面所有字节相加取低8位

说明：

以下本协议中部分项目明细与本规范整体规定保持一致：

a) 本协议中“动作”定义请参考“6.2.1 动作列表”， 其中 0xff 为无效命令；
b) 本协议中“空间”定义列表请参考“6.2.2 空间列表”，其中 0xff 指代所有设备；
c) 本协议中“品类场景”定义列表请参考“6.2.3 品类和场景列表”，其中 0xff 指代所有品类；
d) 本协议中“唤醒词 ID”待具体项目中统一详细定义。

7.3.2 设备地址定义

设备地址 16 比特，高 8 比特标识“空间”信息，低 8 比特标识设备“品类”。同空间+同品类设备认为是同一组设备，不作地址区分，示例如下：

a) 0x203 表示次卧的台灯；

b) 0xff03 表示所有位置的台灯。

7.3.3 命令类型定义

7.3.3.1 命令类型

命令类型的定义，见表7：

表7 命令类型

编码	说明
0x01	唤醒
0x02	设备控制
0x03	信息查询
0x04	设备配置
……	
0x10	命令反馈
0x11	查询反馈
……	

7.3.3.2 不同命令类型，对应的命令格式的定义

7.3.3.2.1 唤醒命令

唤醒命令，见表8：

表8 唤醒命令

帧头	帧长	序号	设备	类型	内容	校验和
0xA6 0xFC	7	顺序号	0	0x01	唤醒词ID	校验和

说明：

“命令内容”仅包含 1 字节长度的唤醒词 ID。

7.3.3.2.2 设备控制

设备控制命令，见表9：

表9 设备控制命令

帧头	帧长	序号	设备	类型	内容				校验和
0xA6 0xFC	10	顺序号	地址	0x02	动作	空间	品类场景	参数	校验和

说明：

“命令内容”包含：长度均为一个字节的动作，空间，品类场景，动作参数，当前规划目标地址中已包含设备类型属性。参数可选，如详细配置 60%亮度。

7.3.3.2.3 信息查询

信息类型包括：芯片，产品 vendor 厂家，版本等，见表 10-11。

表10　信息查询

帧头	帧长	序号	设备	类型	内容		校验和
0xA6 0xFC	9	顺序号	地址	0x03	信息类型	设备	校验和

说明：

“命令内容”包含：

a)　1 字节长度信息类型；

b)　2 字节长度目标设备地址。

表11　信息类型

信息类型	信息长度	说明
0x01	4	版本
0x02	4	芯片型号
0x03	4	Vendor 厂家
0x04	4	工作模式
0x05	4	设备地址
0x06	4	当前状态
…		

7.3.3.2.4　设备配置

设备配置，见表 12：

表12　设置配置

帧头	帧长	序号	设备	类型	内容			校验和
0xA6 0xFC	12	顺序号	地址	0x04	配置	参数	设备	校验和

说明：

“命令内容”包含：

a)　1 字节长度配置信息；

b)　4 字节长度配置参数；

c)　2 字节长度目标设备地址。

此配置除了配置设备参数外，还可以作控制模块对设备的控制。包含但不限于调指定目标设备灯的亮度、色温等，以方便带中控模块的控制模块通过手机 app 等对指定设备进行配置和调节。具体参数根据全局配置类型待定，参数也跟随配置类型而有所差异。

“配置”信息 ID 同“表 11 信息类型”。

7.3.3.2.5　命令反馈

命令反馈，见表13：

表13　命令反馈

帧头	帧长	序号	设备	类型	内容		校验和
0xA6 0xFC	8	顺序号	地址	0x10	命令类型	结果	校验和

说明：

此顺序号同对所回复命令的顺序号（即同所接收到的“控制模块”发来的“设备配置”命令的顺序号）。

7.3.3.2.6　查询反馈

查询反馈，见表14：

表14　查询反馈

帧头	帧长	序号	设备	类型	内容		校验和
0xA6 0xFC	12	顺序号	地址	0x11	信息类型	信息内容	校验和

说明：

此顺序号同对所回复命令的顺序号（即同所接收到的“控制模块”发来的“信息查询”命令的顺序号）。

8　测试和验收

8.1　测试配置

如图2所示，人工嘴（声音源）位于语音模块麦克风正前方[1]，水平直线距离L米[2]。人工嘴（声音源）距离地面120-150cm；噪声源（监听音箱/电视）[3]、语音模块和精密噪音计位于同一平面处（距离地面80-100cm）；噪声源（监听音箱/电视）与语音模块麦克风距离≥150cm，精密噪音计与语音模块麦克风尽量靠近（两者之间距离≤5cm），但不能与语音模块麦克风接触。

图2　测试环境[123]

[1] 人工嘴的位置与角度可以根据实际的场景而定；

[2] L 根据实际场景而定，一般为 1/3/5 米；

[3] 噪声源可以通过监听音箱或电视播放，噪声源的位置与角度可以根据实际的情况而定。

8.2 测试方法

根据声音源不同，主要分为人工测试和自动化测试，人工测试需要真人现场说测试集，自动化测试需要提前录制测试集。

8.2.1 测试集

测试集的性别、年龄、口音分布标准。见表15。

表15 测试集

语言	占比	年龄分布（男女比例 1:1）		
		青年（15-30 岁）	中年（30-50 岁）	老年（50 岁以上）
标准普通话	30%	40%	40%	20%
带方言口音普通话	70%	20%	40%	40%
注：方言口音有广东口音、四川口音、上海口音、湖南口音、湖北口音、江西口音、山东口音、河南口音、安徽口音、江苏口音、浙江口音、福建口音、北京口音、陕西口音				

8.2.2 识别率/唤醒率测试方法与步骤

根据测试要求，改变人工嘴距离语音模组的位置与角度，构建不同声学场景，噪声源（监听音箱/电视）播放噪声集，人工嘴播放对应的测试集，记录测试的数据。

计算方法：

识别率 =（正确识别指令数/输入指令总数）×100%

唤醒率 =（正确唤醒率次数/输入指令总数）×100%

步骤：

a) 使用噪声源（监听音箱/电视）连续播放噪声集，同时通过人工嘴按一定的时间间隔对测试集中的指令进行逐条播放；
b) 记录测试数据；
c) 统计、计算测试结果。

8.2.3 误唤醒测试方法与步骤

根据测试要求，改变人工嘴距离语音模块的位置与角度，构建不同声学场景，人工嘴播放对应的测试集，统计误唤醒次数。

步骤：

a) 使用人工嘴播放测试集（播放声音中不含唤醒词）；
b) 统计误唤醒次数。

8.2.4 响应时间测试方法与步骤

如图 3 所示，搭建好测试环境，打开语音录制工具，播放测试集，播报完成后，利用语音录制工具计算出语音指令与播报之间的时间间隔即为响应时间。

步骤：

a) 使用人工嘴播放测试集；
b) 记录测试数据；
c) 计算相应时间。

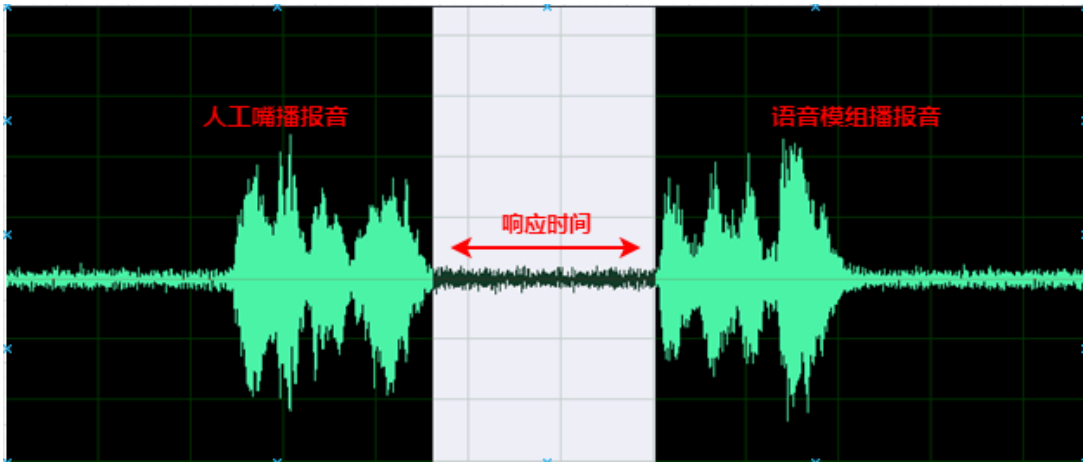

图3　响应时间测量

8.3　验收标准

8.3.1　唤醒率/识别率测试标准

唤醒率和识别率测试标准，见表16：

表16　唤醒率和识别率测试标准

测试项	环境噪声	噪声指标	信噪比 SNR	角度	距离	唤醒率	识别率	备注
本地识别率/唤醒率测试	安静环境	人声：62±3 dB(A) 底噪：40±5 dB	NA	0～90°	1 m	98%	95%	选测
					3 m	95%	95%	必测
					5 m	92%	90%	选测
	人声类噪声	人声：70±5 dB(A) 噪声：55±5 dB	≥15 dB	0～90°	1 m	90%	90%	选测
					3 m	90%	90%	必测
					5 m	88%	88%	选测
	稳态噪声	人声：70±5 dB(A) 噪声：55±5 dB		0～90°	1 m	90%	90%	选测
					3 m	90%	90%	必测
					5 m	88%	88%	选测
	人声 + 稳态噪声	人声：70±5 dB(A) 噪声：55±5 dB		0～90°	1 m	90%	90%	选测
					3 m	88%	85%	必测
					5 m	85%	80%	选测

8.3.2 误唤醒测试标准

误唤醒测试标准，见表17：

表17 误唤醒测试标准

测试项	测试环境	测试时长	指标	误唤醒噪声集说明
误唤醒测试	噪声 60±5 dB	24 小时	≤3 次	1）包括 24 小时时长噪声语料集 2）误唤醒噪声集不出现唤醒词语音
	安静 ≤45 dB	72 小时	0 次	/

说明：
24小时时长噪声语料集，包括：
a) 4 小时的新闻；
b) 4 小时的电视剧；
c) 4 小时的综艺节目；
d) 4 小时的音乐（纯音乐或歌曲）；
e) 8 小时的环境噪声集（设备所处环境）。

8.3.3 响应时间测试标准

响应时间测试标准，见表18：

表18 响应时间测试标准

测试项	测试环境	响应时间
响应时间	安静	≤700 ms
	噪声	500～1000 ms

8.4 噪声类型

测试中常用噪声如表19，主要分为人声干扰和非人声干扰，非人声噪声中有稳态类噪声，如风扇、烟机噪声。测试唤醒率/识别率时，新闻和风扇噪声是必测项，其他根据实际情况选测。

表19　噪声类型

类别	常见可选噪声源	说明
非人声噪声	水流	
	脚步	
	开关门	
	风扇	必测
	烟机	
	窗外马路	
	炒菜	
人声噪声	新闻	必测
	音乐	
	影视剧	
	广告	

参 考 文 献

[1] GB/T 16915.1—2014 家用和类似用途固定式电气装置的开关 第1部分：通用要求（IEC 60669-1:2007, MOD）
[2] GB/T 36464.2—2018 信息技术 智能语音交互系统 第2部分：智能家居
[3] T/SETEA 0005—2021 智能家居产品安全智能家居设备通用要求