



LD3320 评估板说明书

语音识别芯片/声控芯片
单芯片/非特定人/动态编辑识别列表
语音识别解决方案

用声音去沟通 VUI (Voice User Interface)

ICRoute 用声音去沟通
VUI (Voice User Interface)

Web : www.icroute.com
Tel : 021-68546025
Mail: info@icroute.com

一、语音识别芯片介绍

LD3320 是一颗基于非特定人语音识别（SI-ASR: Speaker-Independent Automatic Speech Recognition）技术的语音识别/声控芯片。提供了真正的单芯片语音识别解决方案。

LD3320 芯片上集成了高精度的 A/D 和 D/A 接口，不再需要外接辅助的 Flash 和 RAM，即可以实现语音识别/声控/人机对话功能。并且，识别的关键词语列表是可以动态编辑的。

基于 LD3320，可以在任何的电子产品中，甚至包括最简单的 51 作为主控芯片的系统中，轻松实现语音识别/声控/人机对话功能。为所有的电子产品增加 VUI（Voice User Interface）语音用户操作界面。



二、语音识别评估板介绍

语音识别评估板，目的是为了用最简单的方式，向用户展示出 LD3320 芯片所具有的语音识别功能。该评估板并不具备让用户再二次开发的功能。用户如果需要根据自己的产品定义来进行具体测试，需要自行用 LD3320 芯片进行开发。

1. 评估板上集成了 LD3320 用来作语音识别，主控 MCU 选用的是单片机 STC10L08XE。演示的软件程序全部烧录在 MCU 的内置 Flash 中。主控 MCU 直接控制 LD3320 完成所有和语音识别相关的工作。（主控 MCU 通过读/写 LD3320 的寄存器完成操作。例如在编辑关键词语时，主控 MCU 通过把关键词语的拼音串设置进 LD3320 相关寄存器来完成编辑）

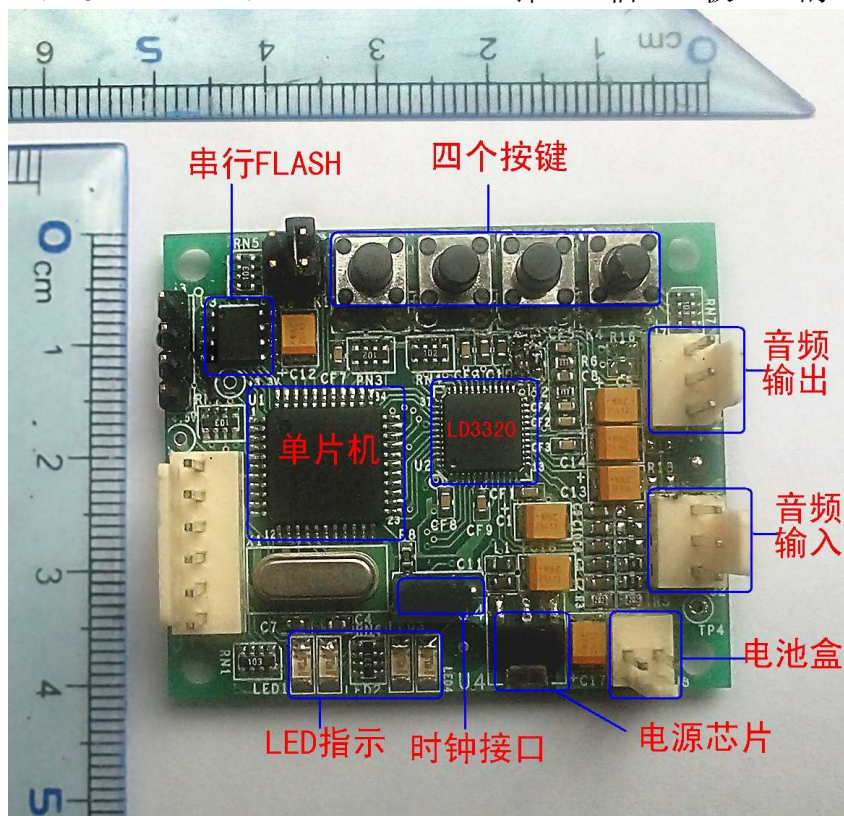
2. 评估板上另外集成了一片串行存储芯片，是华邦的 W25Q40AVSNIG，大小为 512Kbyte。里面存储的是演示程序在运行过程中播放的声音素材。格式为 MP3 格式。为了节省存储空间，所有的录音声音都为 8kbps 的 MP3，音乐声音也只是 16Kbps 的 MP3（相当于 2Kbyte/秒），所以音效不是最佳。如果客户追求最佳音效，可以用最高码率的 MP3 以演绎 LD3320 的播放效果。

3. 在电源方面，评估板使用了 NS 公司的 LM1117-3.3 芯片，为 LD3320 提供 3.3V 的电压。用户应该为评估板提供稳定的电源输入，建议输入电压为 5V 直流电源。由于评估板使用了 LD3320 的内置功放（Speaker 模式）来推动外置喇叭播放声音，而且目前演示程序设置的播音音量为 12（最大为 15，最小为 0），峰值时会需要有约 300mA 的电流来驱动播放。所以请用户注意提供足够功率的电源来保证评估板的正常运行。（请注意电源的正负极）

4. 在音频输入和输出方面，评估板配套了一个柱极体 MIC 和一个 0.5W 的小喇叭。需要说明的是，语音识别的效果和多种因素相关，这些因素也包括外接麦克风的灵敏性能和信噪比性能。目前随评估板配套的柱极体 MIC 是市面上最普通的售价大约为 0.3RMB 的 MIC。用户如果需要测试，可以自行配置高灵敏度或者单一指向性的麦克风，接入评估板进行测试。请注意麦克风的极性正反，如果接反，会导致没有声音输入 LD3320。

5. 在语音识别芯片的音频输入灵敏度的参数设置上，目前评估板设置的参数相对比较低，也就是对于比较近距离的声音有比较好的识别效果，对于比较距离远的声音会不灵敏。用户将来在自行开发时，可以自行设置相关参数来调整性能以满足具体产品的应用场景需求。在评估板上，如果换用高灵敏度的麦克风，也可以有效地扩大语音识别的作用距离，用户如果有兴趣可以自行测试。（举例说明，奥林巴斯的 ME52 麦克风就属于高灵敏度的麦克风）

6. 在评估板上 MCU 通过并口的方式和 LD3320 相连接，晶振是 22.1184MHz。LD3320 评估板的照片：

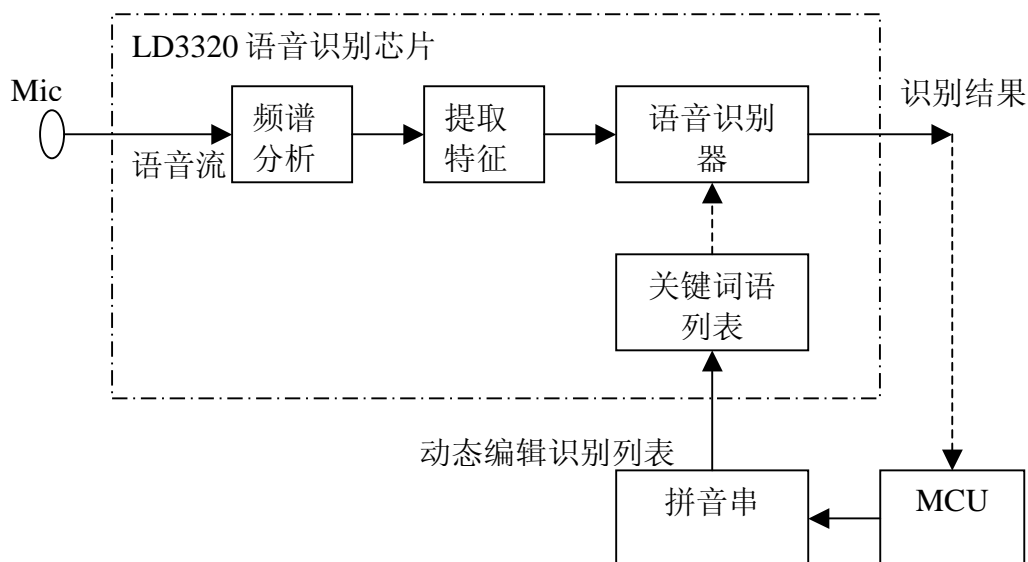


三、语音识别芯片的工作原理

（本节是讲述语音识别芯片相关的技术原理，用户也可以浏览第五节直接运行评估板的 Demo）

LD3320 提供的语音识别技术，是基于“关键词语列表”的识别技术，关于 ASR 技术的详细说明，请参考网站：
http://www.icroute.com/web_cn/Tech ASR.html。

语音识别芯片完成的工作就是：把通过 MIC 输入的声音进行频谱分析->提取语音特征->和关键词语列表中的关键词语进行对比匹配->找出得分最高的关键词语作为识别结果输出。



语音识别芯片能在两种情况下给出识别结果：

- 1) 外部送入预定时间的语音数据后（比如 5 秒钟的语音数据），芯片对这些语音数据运算分析后，给出识别结果
- 2) 外部送入语音数据流，语音识别芯片通过端点检测 VAD（voice activity detection）检测出用户停止说话，把用户开始说话到停止说话之间的语音数据进行运算分析后，给出识别结果

对于第一种情况，可以理解为设定了一个定时录音（比如为 5 秒钟），芯片在 5 秒钟后，会停止把声音送入识别引擎，并且根据已送入引擎的语音数据计算出一个识别结果。

对于第二种情况，需要了解 VAD 的工作原理：

VAD（Voice Activity Detection）技术是在一段语音数据流中，判断出哪个时间点是人声的开始，哪个时间点是人声的结束。判断的依据是，在

背景声音的基础上有了语音发音，则视为声音的开始。而后，检测到一段持续时间的背景音（比如 600 毫秒），则视为人声说话结束。通过 VAD 判断出人声说话的区域后，语音识别芯片会把这期间的声音数据进行识别处理后，计算出识别结果。

需要说明的是，除了以上两种情况外，语音识别算法无法“主动”地判断出是否识别出了一个结果。这是因为，在计算的过程中的任何时刻，语音识别器都会对已送入识别芯片的声音数据进行分析，并根据匹配程度为识别列表中的关键词语进行打分，最匹配的打分最高。

但是，由于识别算法不知道用户后面是否还继续说语音，所以无法“主动”地判断已经识别出一个结果。

比如，识别列表中有两个关键词语分别是“李晓”和“李晓明”，当用户说到“晓”这个音节时，在识别芯片内部是“李晓”的得分最高，但此时识别芯片是无法给出识别结果。随着用户说出“明”音节，则成为得分最高。此时，用户停止说话，或者是定时录音已到，使得识别芯片可以判断出用户已经停止说话了，才能给出识别结果是“李晓明”。

四、语音识别芯片用户使用模式

用户可以通过编程，设置两种不同的用户使用模式：“触发识别模式”和“循环识别模式”：

触发识别模式：

系统的主控 MCU 在接受到外界一个触发后（比如用户按动某个按键），启动 LD3320 芯片的一个定时识别过程（比如 5 秒钟），要求用户在这个定时过程中说出要识别的语音关键词语。过了这个过程后，需要用户再次触发才能再次启动一个识别过程。

循环识别模式：

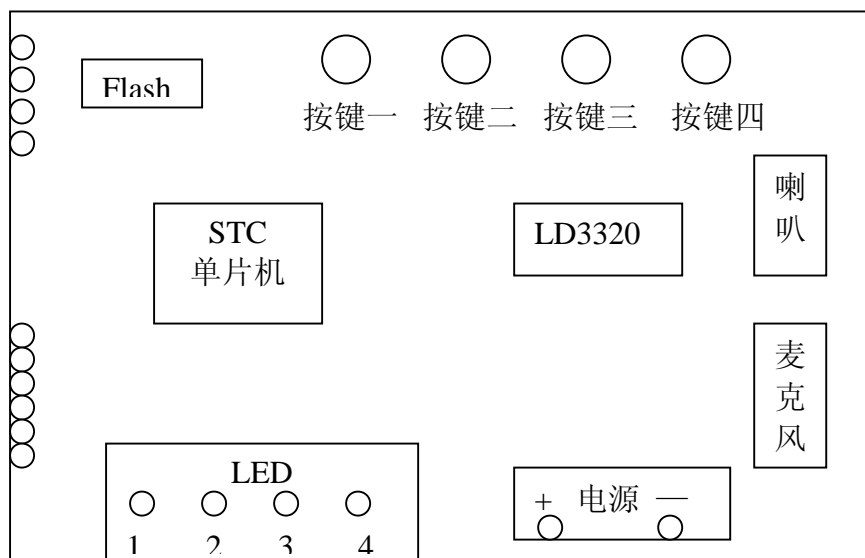
系统的主控 MCU 反复启动识别过程。如果没有人说话没有识别结果，则每次识别过程的定时到时后再启动一个识别过程；如果有识别结果，则根据识别作相应处理后（比如播放某个声音作为回答）再启动一个识别过程。

一般来说，触发识别适合于识别精度要求比较高的场合。外界触发后，产品可以播放提示音或者其他方式来提示用户在接下来的几秒钟内说出要识别的内容，这样来引导用户在规定的时间内只说出要识别的内容，从而保证比较高的识别率。

而循环识别比较适合于需要始终进行语音监控的场合，或者没有按键等其他设备控制识别开始的场合。而这种状态，识别准确度会有一定下降，在循环识别的过程中，用户的其他说话声音，或者外界的其他声音，都有可能被识别引擎误识别出错误的结果，需要产品的控制逻辑都作相应的处理。

五、LD3320 评估板 Demo 使用说明

评估板的布局如下图所示：



评估板的 STC 单片机内共设置了三个内置 Demo 程序：

按键一：语音点歌（循环识别模式）

按键二：智能家居命令操作（循环识别模式）

按键三：城市区号查询（触发识别模式）

评估板上电后，LED3, 4 会点亮。同时会播放“叮”的一声，此时用户可以按键来选择合适的 Demo。

1) 语音点歌（循环识别模式）

按键一进入 Demo 后，播放声音“请说出您喜欢的歌手”，随后进入循环识别模式，用户可以说出以下 12 位歌手名字中的任意一个：

刘德华
张学友
周杰伦
蔡依林
林忆莲
蔡琴
王力宏
林俊杰
梁静茹
邓丽君
王菲
陈奕迅

当用户说完歌手名字后，评估板会播放识别出的结果，并播放该歌手的一段 MP3 歌曲。

2) 智能家居命令操作（循环识别模式）

按键二进入 Demo 后，播放声音“请说出要操作的家用电器”，随后进入循环识别模式，用户可以说出以下 4 个命令中的任意一个（此时状态为下页命令列表中的 S20 状态）：

空调
照明
微波炉
电磁炉

当用户说完命令后，评估板会根据识别结果播放相应的提示音进入二级菜单，用户可以再说出二级菜单中的命令。

命令列表见随后的列表。

3) 城市区号查询（触发识别模式）

按键三进入 Demo 后，播放声音“请在按键后说出要查询的城市名字”，随后进入触发识别模式。

用户每次按动按键四，会播放“叮”的提示音，随后用户可以说出 50 个城市名字中的一个。评估板会根据识别结果播放该城市的名称和电话区号。城市名字列表见随后的识别句列表。

识别句列表:

刘德华
张学友
周杰伦
蔡依林
林忆莲
蔡琴
王力宏
林俊杰
梁静茹
邓丽君
王菲
陈奕迅

S20: 102 空调 转 S21
103 照明 转 S22
104 微波炉 转 S23
105 电磁炉 转 S24

S21: 112 开机
113 关闭
114 调高温度
115 降低温度
116 不对 转 S20
117 退出 转 S20
118 返回 转 S20

S22: 122 卧室
123 客厅
124 厕所
125 睡觉
126 关闭
127 不对 转 S20
128 退出 转 S20
129 返回 转 S20

S23: 132 热奶
133 解冻
134 大火一分钟
135 中火两分钟
136 不对 转 S20
137 退出 转 S20
138 返回 转 S20

S24: 142 火锅,
143 慢炖,
144 炒菜,
145 关机
146 调高温度,
147 降低温度
148 不对 转 S20
149 退出 转 S20
150 返回 转 S20

201 北京
202 上海
203 天津
204 重庆
205 广州
206 杭州
207 成都
208 南京
209 武汉
210 沈阳
211 济南
212 哈尔滨
213 郑州
214 石家庄
215 长沙
216 长春
217 福州
218 西安
219 昆明
220 南昌
221 合肥
222 太原
223 呼和浩特
224 南宁
225 乌鲁木齐

226 兰州
227 贵阳
228 银川
229 海口
230 西宁
231 拉萨
232 台湾
233 苏州
234 厦门
235 青岛
236 深圳
237 香港
238 澳门
239 三亚
240 无锡
241 桂林
242 黄山
243 丽江
244 凤凰
245 宁波
246 九寨沟
247 张家界
248 大连
249 西双版纳
250 珠海

使用完评估板并有初步体验和疑问的用户，可以参考阅读“语音识别芯片 LD3320 高阶秘籍.pdf” [http://www.icroute.com/doc/语音识别芯片 LD3320 高阶秘籍.pdf](http://www.icroute.com/doc/语音识别芯片_LD3320高阶秘籍.pdf) 增强对语音识别芯片的深入理解。

本评估板仅作评估功能使用，不能作为二次开发使用。

ICRoute 对每个评估板收取押金，在收取押金后，把评估板提供给用户评估。在用户完成评估板的评估工作后，把评估板返回给 ICRoute，把押金再退还给用户。

建议大家可以在使用完评估板后把评估板返回给 ICRoute 公司，循环使用评估板，以利于环保和节约。

如果有问题，请联系 ICRoute：

Tel : 021-68546025

Mail: info@icroute.com