



LD3320 语音识别芯片

(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 : 021-58796509

Mail: sales@51sensor.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。

主要特色功能：

非特定人语音识别技术：不需要用户进行录音训练

可动态编辑的识别关键词语列表：只需要把识别的关键词语以字符串的形式传进芯片，即可以在下次识别中立即生效。比如，用户在 51 等 MCU 的编程中，简单地通过设置芯片的寄存器，把诸如“你好”这样的识别关键词的内容动态地传入芯片中，芯片就可以识别这样设定的关键词语了。

真正单芯片解决方案：不需要任何外接的辅助 Flash 和 RAM，真正降低系统成本。

内置高精度 A/D 和 D/A 通道：不需要外接 AD 芯片，只需要把麦克风接在芯片的 AD 引脚上；可以播放声音文件，并提供 550mW 的内置放大器。

高准确度和实用的语音识别效果

支持用户自由编辑 50 条关键词语 在同一时刻，最多在 50 条关键词语中进行识别，终端用户可以根据场景需要，随时编辑和更新这 50 条关键词语的内容。



LD3320 芯片外观

二、语音识别介绍

语音识别 ASR 技术，是基于关键词语列表识别的技术。只需要设定好要识别的关键词语列表，并把这些关键词语以字符的形式传送到 LD3320 内部，就可以对用户说出的关键词语进行识别。不需要用户作任何地录音训练。

语音识别 ASR 现实意义：

ASR 技术最重要的现实意义就在于提供了一种脱离按键，键盘，鼠标的基于语音的用户界面 VUI: Voice User Interface。

使得用户对于产品的操作更快速，更自然。

非特定人语音识别技术 ASR，是对几十 G 的语音数据经语言学家语音模型分析，科学家建立数学模型，并经过反复训练提取基元语音的细节特征，以及提取各基元间的特征差异。可以得到在统计概率最优化意义上的各个基元语音特征。最后才由资深工程师将算法以及语音模型转化为硬件芯片以应用在嵌入式系统中。

基于领先的语音识别核心引擎，经过在各大主流手机产品中三年来超过一千万份的量产检验，LD3320 提供了可以真正实用的语音识别/声控功能。

ASR 技术是基于关键词语列表识别的技术。

每次识别的过程，就是把用户说出的语音内容，通过频谱转换为语音特征，和这个关键词语列表中的条目进行一一匹配，最优匹配的一条作为识别结果。比如在手机的应用中，这个关键词语列表的内容就是电话本中的人名/手机的菜单命令/T 卡中的歌曲名字。

不论这个列表的条目内容是什么，只需要用户设置相关的寄存器，就可以把相应的待识别条目内容以字符形式传递给识别引擎。

LD3320 可以识别列表中的关键词，用户说的语音可以是这个列表中任意的关键词语，而且不需要用户在识别前进行任何训练。

识别引擎不关心关键词语列表中的关键词语的内容，可以是命令，人名，歌曲名字，操作指令等等任何的汉字字符串。

每条关键词语最大可以支持的字数，从算法角度是限制在 30 字以内。但是从实际情况来看，用户一口气说超过 8 个字以上的条目时，几乎肯定会出现说错字/说漏字/说多字/打嗝/停顿等情况，这些情况都会严重影响识别并造成识别错误。因而一般来说，如果要获得理想的识别效果，建议每条关键词语的字数不要过长，避免影响效果。

ASR 技术上的局限性在于：

ASR 不是听写系统

ASR 不能把人的声音转换为确定的拼音串

ASR 不能任意地识别人的说话内容

ASR 不能识别关键词列表中的某个关键词语的一部分内容，除非这一部分内容自身也是一个关键词语

ASR 不可以识别与关键词语列表中列出的关键词语不相符的情况

比如，前后加了“嗯”，“阿”之类的语气词

比如，只说出了关键词语中的一部分而不是整个关键词语

比如，没有列在关键词语列表中的词语

三、应用场景

由于 LD3320 可以动态编辑的识别关键词语列表，因此其可以应用的范围大大超过了那些不可以改变识别列表的芯片。同时，LD3320 是一颗真正的单芯片解决方案，不需要辅助的外围 Flash/RAM/AD，只需要有一颗 51 级别的单片机芯片就可以控制 LD3320，从而降低了系统的成本。

其应用场景主要包括：

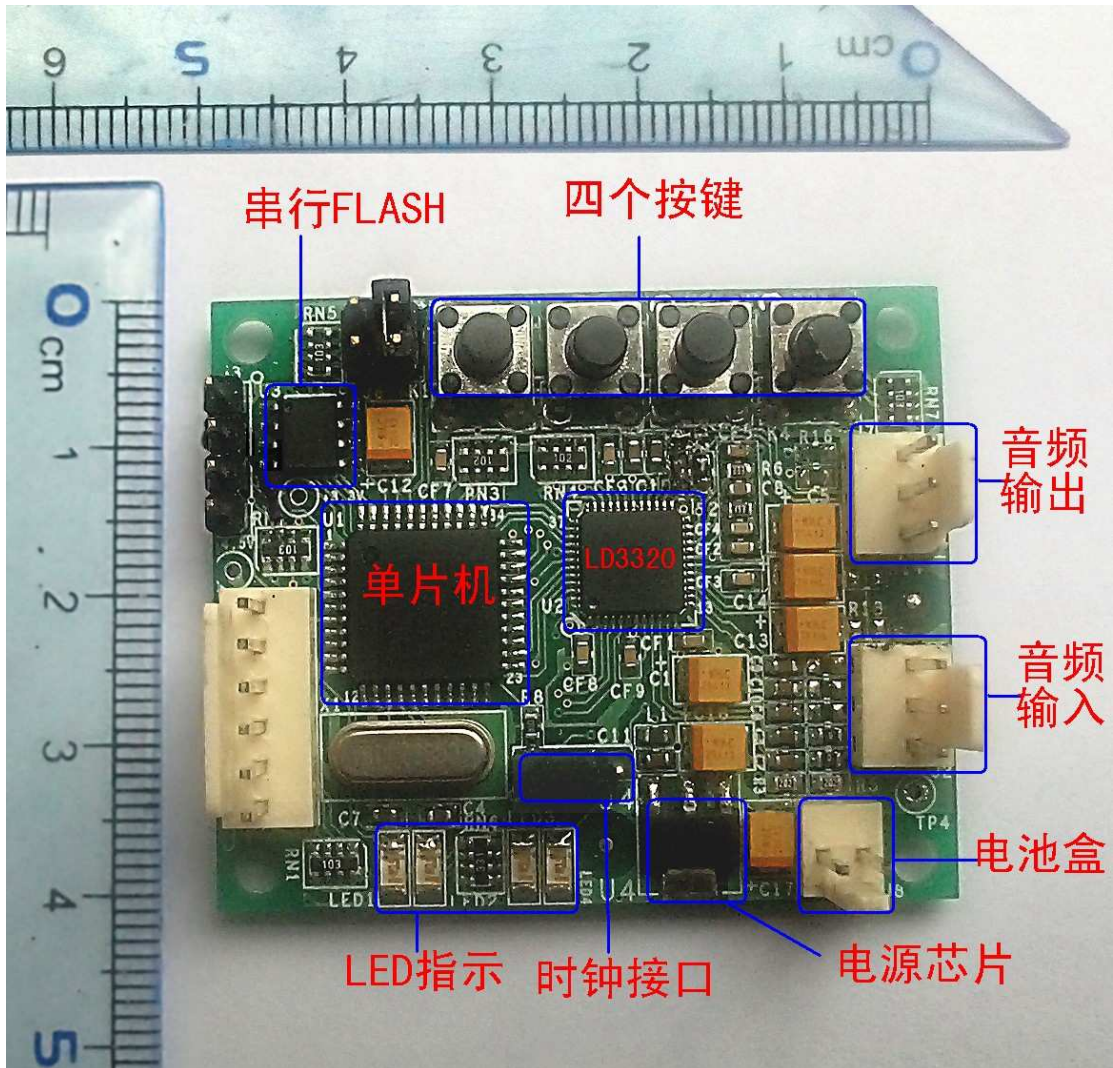
1. 智能玩具/对话玩具
2. 导航仪
3. MP3/MP4
4. 数码像框
5. 机顶盒/彩电遥控器
6. 电磁炉/微波炉/洗衣机/智能家电操作
7. PMP/游戏机
8. 自动售货机
9. 地铁自动售票机
10. 导游机
11. 楼宇电视的广告点播

(请参考“LD3320 语音识别芯片应用场景”)

四、主要技术参数

1. 内置单声道 mono 16-bit A/D 模数转换
2. 内置双声道 stereo 16-bit D/A 数模转换
3. 内置 20mW 双声道耳机放大器输出
4. 内置 550mW 单声道扬声器放大器输出
5. 支持并行接口或者 SPI 接口
6. 内置锁相电路 PLL，输入主控时钟频率为 2MHz - 34MHz
7. 工作电压：(VDD: for internal core) 3.3V
8. 48pin 的 QFN 7*7 标准封装
9. 省电模式耗电：1uA

LD3320 评估板的照片：（其中，控制 MCU 目前选用的是 STC 的 51 单片机）



五、联系我们

T 021-58796509