

语音识别模块高级使 用指南

——语音识别模块之
语音识别模块高级使用指南

版本号: V1.1



一、在识别精度要求高的场景中，使用“触发识别”模式

在识别精度要求高的场景中，应该采用“触发识别”模式。原因是：

📖 (1) 用户在每次按热键后，精神处于最集中的状态，此时用户说的语音

命令会比较认真，清晰。避免了用户过于随意的发音导致的识别误差。★

📖 (2) 每次按热键后，产品应该给以一个明显的开始信号，比如发出“当”的一声或者其他提示信号，可以给用户一个明确开始的提示，方便用户掌握说语音命令的时间。

📖 (3) 由于按键触发后，用户就会贴近麦克风并说出语音命令，避免了其他环境声音被录入 LD3320 芯片导致的误识别。

二、 增添“垃圾关键词语”——吸收错误识别

在设定好要识别的关键词语后，为了进一步降低误识别率，可以再添加一些其他的任意词汇进识别列表，用来吸收错误识别，从而达到降低误识别率的目的。可以把这些关键词语称之为“垃圾关键词语”。比如，某个应用场景中，需要识别的关键词语是 4 条，“前进”，“后退”，“开门”，“关门”。在把这 4 个关键词语设置进语音识别模块后，可以再

另外设置 10~30 个词语进语音识别模块，比如“前门”，“后门”，“阿阿阿”，“呜呜”等等。

只有识别结果是 4 个关键词语之内的，才认为识别有效。如果识别结果是“垃圾关键词语”，则说明是其他的声音导致的误识别，产品应该重新开始一次识别过程。

这样，可以非常非常有效地降低误识别率。极大地提高终端用户的主观使用体验。“垃圾关键词语”的选取，最好可以选择一些字数和关键词语一样的词语，用来吸收可能发生的错误识别。需要说明的是：这一方法，即可以应用在“触发识别”模式中，也可以应用在“循环识别”模式中。

三、口令触发模式

在一些应用场合，希望识别精度高，但是又无法要求用户每次都用手按键来“触发识别”。此时，可以采用“口令触发模式”。产品定义一句短语，作为触发口令。比如，可以定义“芝麻开门”作为触发口令。产品在等待用户触发时，启动一个“循环识别”模式，把触发口令“芝麻开门”和其他几十个用来吸收错误的词汇设置进音识别模块。只有当检测到识别出的结果是触发口令时，才认为是终端用户叫了这个口令。此时，给出提示音，并跳转到其他的识别模态，并且把相应的识别列表设置进语音识别模块，提示用户在提示音后几秒钟内说出要执行的操作。

在等待用户的过程时，如果识别的结果是那些用来吸收错误的词汇，则认为是误识别，或者其他的声音干扰，而不进行任何的处理，直接再次进入口令识别模态。

这种口令触发模式，融合了其他两种模式的优点，并且结合第二节提到的“垃圾关键词语”的方法，可以为产品提供更加方便实用的语音操作特性。

四、用拼音标注外文或者方言

语音识别，识别的是“语音”。对于非特定人语音识别来说，在描述关键词语时，是用音标标注出要识别的关键词语。对于目前 Newway 语音识别模块支持的中文识别来说，就是用拼音来描述出关键词语。也就是说，只要是拼音可以拼出的发音，都是可以输入芯片并进行识别的。因此，在某些场合需要识别一些简单的外文或者纯方言发音的时候，可以用拼音标注的方法来实现。

例如，有些场合需要识别一些简单的英文单词，可以用拼音标注：

one -> wan

two ->□ tu

three ->□ si rui

例如，有些场合需要识别一些纯方言发音的词汇，也可以用拼音标注：

上海话的“晚”发音是“ya”，那么“晚报”这个词

汇，用普通话标注是

“wan bao”，如果要标注成上海话发音，就是“ya bao”，这样上海话说的“晚报”也就可以被识别了。值得注意的是：Newway 语音识别模块支持的是中文普通话，有些外文或者方言发音无法用拼音描述，所以 Newway 语音识别模块不一定能够完成所有需要的外文或者方言任务。

