



emASR嵌入式语音识别&语音唤醒软件 用户手册

北京字音天下科技有限公司



010-62986600



010-62969399



www.tts168.com.cn



字音天下官方订阅号

重要声明

版权声明

版权归北京宇音天下科技有限公司所有，保留所有权利。

商标声明

北京宇音天下科技有限公司的产品是北京宇音天下科技有限公司专有。在提及其他公司及其产品时将使用各自公司所拥有的商标，这种使用的目的仅限于引用。本文档可能涉及北京宇音天下科技有限公司的专利（或正在申请的专利）、商标、版权或其他知识产权，除非得到北京宇音天下科技有限公司的明确书面许可协议，本文档不授予使用这些专利（或正在申请的专利）、商标、版权或其他知识产权的任何许可协议。

不作保证声明

北京宇音天下科技有限公司不对此文档中的任何内容作任何明示或暗示的陈述或保证，而且不对特定目的的适销性及适用性或者任何间接、特殊或连带的损失承担任何责任。本手册内容若有变动，恕不另行通知。本手册例子中所用的公司、人名和数据若非特别声明，均属虚构。未得到北京宇音天下科技有限公司明确的书面许可，不得为任何目的、以任何形式或手段（电子的或机械的）复制或传播手册的任何部分。

保密声明

本文档（包括任何附件）包含的信息是保密信息。接收人了解其获得的本文档是保密的，除用于规定的目的外不得用于任何目的，也不得将本文档泄露给任何第三方。

本软件产品受最终用户许可协议（EULA）中所述条款和条件的约束，该协议位于产品文档和/或软件产品的联机文档中，使用本产品，表明您已阅读并接受了EULA的条款。

版权所有：北京宇音天下科技有限公司

目 录

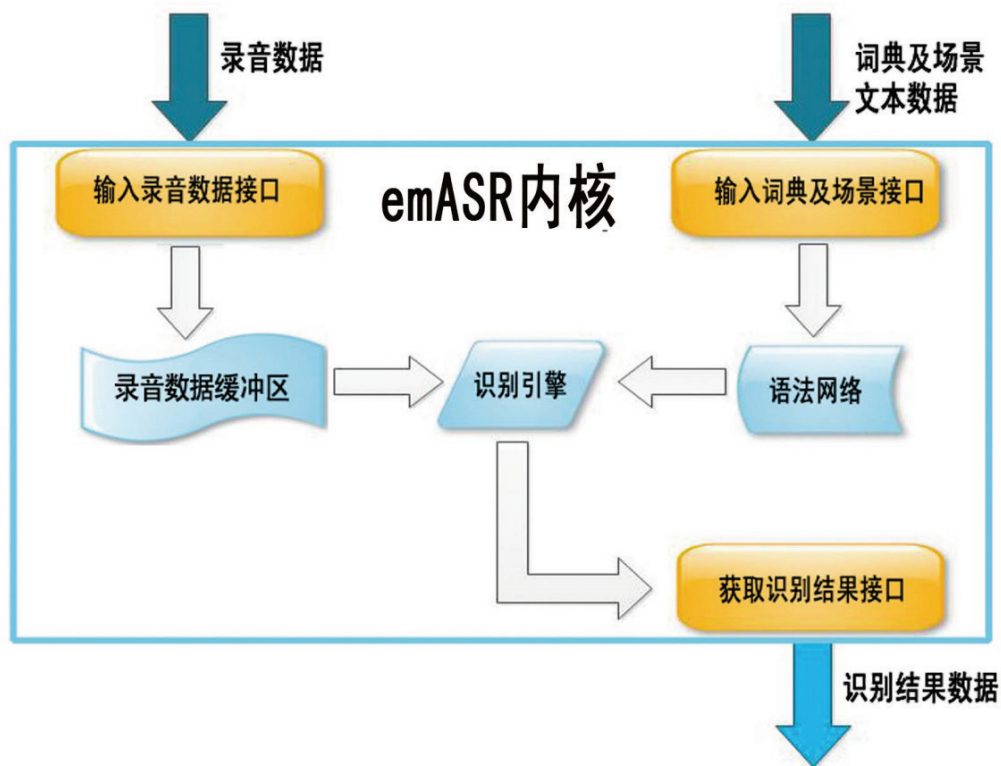
1	概述.....	4
2	系统示意图.....	4
3	主要应用领域.....	5
4	产品功能描述.....	5
4.1	总体功能.....	5
4.2	语音识别.....	5
4.3	语音唤醒.....	7
5	产品技术指标.....	7
5.1	产品整体特性.....	7
5.2	移植所需环境要求.....	8
6	API接口简述.....	9
7	合作&联系方式.....	9

1 概述

emASR 是一款支持语音唤醒及非特定人命令词识别的嵌入式语音识别引擎，可为项目研发定制开发 API 接口——主要利用嵌入式设备平台系统的运算、存储资源，以软件模块方式实现与客户软件系统的整合，从而为用户的智能终端系统实现嵌入式语音识别功能。emASR 软件不受操作系统限制，不增加用户额外的硬件成本，特别是在识别率、系统抗噪性、拒识性、鲁棒性等方面有了很大的提升，其识别效果已达到工控级应用需求。

emASR 软件使用简单、自定义灵活、具有命令词识别率高，语音唤醒率高的超强能力，是一款真正面向室内环境下的高端语音识别引擎，使用户实现无需触碰，直接开启语音交互，打造交互新格局。它将推动 ASR 语音识别技术的行业应用走向更深入、更广泛！我们将以专业化的服务态度和雄厚的语音技术实力，为行业中每个客户量身订做语音识别软核产品。

2 系统示意图



3 主要应用领域

- 智能家居语音控制
- 可穿戴设备
- 车载电子
- 手机
- 语音导游
- 智能仪器
- 故事机
- 智能交互机器人
- GPS 导航仪
- 高端玩具
- 智能玩具
- 插卡音箱
- 学习机
- 安防控制

4 产品功能描述

4.1 总体功能

- **标准API接口**

采用开放式的体系结构，提供标准的开发接口API，二次开发简单

- **完全与平台无关**

emASR有着优良的兼容性，内核完全与平台无关。任何平台只需配置、编译，即可完成移植。emASR能适用绝大部分平台，例如 Windows CE、Linux、MTK系列、Android、IOS、Nucleus、Windows Mobile、Windows Mobile 2003等。

- **资源占用低**

考虑到嵌入式环境下宝贵的资源空间，emASR在系统架构、代码实现、资源结构等方面进行了大量的优化工作，在运算资源和存储资源的需求方面，都较传统的识别系统有了极大的降低，即便是一些低端的平台，emASR也能从容运行。

4.2 语音识别

- **非特定人命令词识别**

非特定人的自然语音识别,对识别语音无特殊要求，不论男女老幼，普通话汉语均可享受识别功能。

- **支持中文、英文、中英文混合识别**

词条无中英文限制，可进行中文、英文、中英混合词条的识别，使用范围广，简单灵活。

- **准确的端点检测**

能在复杂的应用环境中，准确地检测出语音信号的开始及结束，在检测到语音信号结束后自动给出识别结果，可以应用在各种不同的使用场合。

- **识别率高，带拒识功能，系统抗噪性好**

具有五级识别拒识，适用于安静、办公和嘈杂等多种环境下使用；emASR对以下几种声音进行较好的拒识：

- 对非人声（例如：鼓掌声，关门声，敲桌子声）的拒识
- 对不在词表里的人声词条的拒识
- 对在词表里的非常含糊的人声词条的拒识；

- **支持10个词典，上万条词条命令，支持可识别词条的动态刷新**

允许可识别命令词条，在识别过程中进行动态更新，可以对识别词条进行动态的增删，从而实现应用程序具有动态语音识别的能力。

- **词条更新支持携带命令ID的语义模式**

- 基本词条更新模式：

词条编辑格式为“开灯|请开灯|把灯打开|关灯|请关灯|把灯关了|亮一点|再亮一点”。

- 携带命令ID语义模式：（“关灯”“请关灯”“把灯关了”属于同一语义，均会返回命令ID=2）

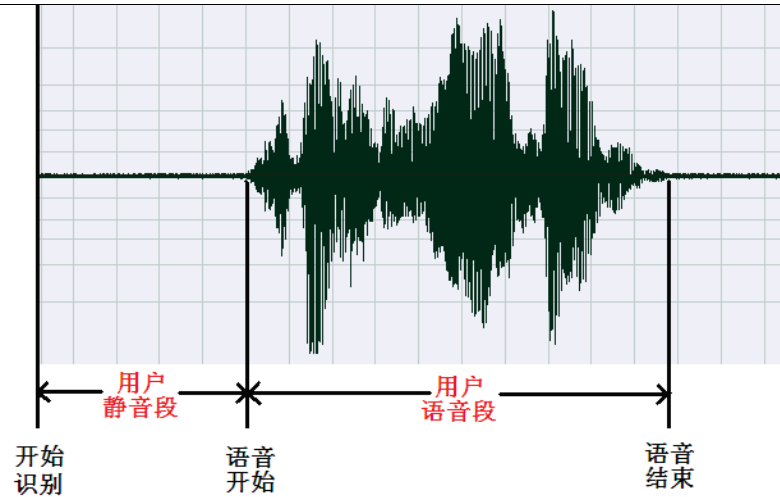
词条编辑格式为“开灯 1|把灯打开 1|关灯 2|请关灯 2|把灯关了 2|亮一点 3|再亮一点 3”。

- **词典词条更新支持多种编码格式**

可以采用GB2312、GBK、BIG5、Unicode和UTF-16等多种编码方式。

- **反应超时检测和语音超时检测**

emASR在规定时间内未检测到有效语音数据会返回反应超时提示，可以有效控制无操作情况下自动超时退出；emASR检测到语音数据超过规定时间会返回语音超时提示，避免长时间无效语音及噪音对用户处理机制的干扰；



4.3 语音唤醒

- **无需触碰，直接唤醒设备**
语音唤醒，无需触碰，直接进入，您只要说出设定的唤醒词，简单语音呼叫就可以唤醒设备
- **支持多种唤醒名（可定制）**
内置6种语音唤醒名，也可以为用户量身打造自定义唤醒名，使交互模式更简单。内置6种唤醒名分别为：云宝、Hi云宝、小播、Hi小播、百灵、Hi百灵。
- **支持动态切换唤醒名**
用户可以动态设置当前唤醒名，随意切换唤醒名。

5 产品技术指标

5.1 产品整体特性

项目	指标
词典文本类别	普通书面文本
输入文本代码页	简体中文 GB2312; 简体中文 GBK; 繁体中文 Big5; Unicode; UTF-16;
语音来源要求	普通话（无特殊要求）
输入语音格式	PCM signed 16bit
输入语音采样率	16KHz 单通道, 8KHz 单通道

输出结果	结果输出状态、词条 ID 号、命令 ID 号、匹配度
平均无故障工作时间	不小于 72 小时

5.2 移植所需环境要求

平台特性	支持范围
最小可访问单元	8bit
编译器	C 编译器，支持递归调用，支持 32 位值运算
字节序	Little endian/Big endian

※ 数据情况仅供参考。具体平台、具体应用以实际配置为准。

			功能	主频	堆内存	资源空间
中英文识别和唤醒	16K 采样率	小资源	不支持动态更新词条 (出厂后词条不再变化)	160MHz	>680K Byte	>400K Byte
			支持动态更新词条	160MHz	>930K Byte	>1400K Byte
		大资源	不支持动态更新词条 (出厂后词条不再变化)	360MHz	>2600K Byte	>2400K Byte
			支持动态更新词条	360MHz	>3000K Byte	>3400K Byte
	8K 采样率	中资源	不支持动态更新词条 (出厂后词条不再变化)	160MHz	>1600K Byte	>1400K Byte
			支持动态更新词条	160MHz	>2000K Byte	>2400K Byte

6 API接口简述

API接口功能	说明	反馈
创建对象	创建识别唤醒对象	成功, 错误
销毁对象	销毁识别唤醒对象	成功, 错误
识别词条更新	更新选定词典的命令词条	成功, 错误
设置参数	设置参数: 识别距离模式, 用户静音上限 用户语音上限, 拒识级别	成功, 错误
开始语音识别	启动语音识别功能	识别成功: 匹配度, 词条ID, 命令ID 识别错误: 静音超时, 语音超时, 识别拒识
停止语音识别	停止当前已经启动的语音识别	成功, 错误
开始语音唤醒	开启语音唤醒功能	成功, 错误
停止语音唤醒	关闭语音唤醒功能	成功, 错误

7 合作&联系方式

emASR 软件针对用户的智能平台进行定制, 销售方式为: 技术授权

联系方式:

北京字音天下科技有限公司 (总部)	张红英	010-62986600	15321801789
珠海字音天下科技有限公司 (子公司)	何海滨	0756-6918288	18575636388
深圳办事处		0755-22197482	