

Aitalk 3.0

嵌入式语音识别系统 产品白皮书

重要声明

版权声明

版权所有 © 2008, 安徽科大讯飞信息科技股份有限公司, 保留所有权利。

商标声明

安徽科大讯飞信息科技股份有限公司的产品是安徽科大讯飞信息科技股份有限公司专有。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 档可能涉及安徽科大讯飞信息科技股份有限公司的专利(或正在申请的专利)、商标、版权 或其他知识产权,除非得到安徽科大讯飞信息科技股份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

安徽科大讯飞信息科技股份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。未得到安徽科大讯飞信息科技股份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© 安徽科大讯飞信息科技股份有限公司

Copyrights © Anhui USTC iFLYTEK CO., LTD.

目录

1	产品简介	4
2	产品功能	4
2.1	电话号码语音输入	4
2.2	电台号码语音输入	5
2.3	非特定人语音标签	5
2.4	为中国人设计的英文识别	5
2.5	支持中文、英文、中英文混合识别	5
2.6	支持句式功能	6
2.7	对符号、中文姓氏智能判断	6
2.8	同音字识别	6
2.9	支持动态识别命令词	
2.10	非特定人识别	6
2.11	I 77 II CER	
2.12	拒识	7
2.13	内置超时控制功能	7
3	联系我们	8



1 产品简介

Aitalk 3.0 产品是科大讯飞推出的命令词语音识别产品。为最终用户、合作伙伴使用语音识别时进行了仔细的考虑:

> 为最终用户

Aitalk 3.0 提供了更多的功能支持,最终用户可以更多地享受到语音技术带来的便利。 Aitalk 3.0 提供的新功能包括: 电话号码输入、FM 调频输入、非特定人语音标签。

Aitalk 3.0 对**车载环境优化**,相对识别率提升 30%以上。

Aitalk 3.0 对中国人说英文的发音习惯,收集了大量数据并开展了专题研究,是**为中国人设计的英文识别引擎**。实验表明,相对识别率提升 50%以上。

> 为合作伙伴

Aitalk 3.0 设计了全新的集成架构,尽量减少应用层冗余代码,**编码更精简,集成更方**便。

Aitalk 3.0 支持结构化的语法描述文件输入,可以使交互设计工程师独立于研发工程师工作,**优化语音交互**;独立的语法描述还可以分离程序逻辑与描述数据,工程的**可维护性**得到提高。

2 产品功能

2.1 电话号码语音输入

用户一次说完整个电话号码,Aitalk 3.0 就能自动完成由语音向文本的转换,输出用户所说的电话号码的内容。

Aitalk 3.0 对于中国大陆地区电话号码、区号、IP 电话、手机号码的编号规则进行了适应,国内用户使用时更加方便快捷。

电话号码输入对于提高车载安全性有重要的意义。在开车过程中,用户可以随时启动语音识别,输入电话号码后拨打电话。从而免去了中途停车拨号的麻烦。



2.2 电台号码语音输入

FM 电台号码输入功能可以支持用户说出电台频率时,直接将收音机调到特定的频率。如用户说"调频九十八点八",语音识别就可以帮助将收音机调节到调频 98.8MHz。

Aitalk 3.0 覆盖了绝大部分中国人说电台号码的习惯。使用 Aitalk 3.0,你也可以定制用户说电台号码的方法。

FM 电台号码输入,能够扩展语音在汽车等设备中语音识别的应用范围,进一步提升娱乐系统的用户体验。

2.3 非特定人语音标签

语音标签(Voice Tag),就是为某个命令词指定发音,如对于命令词"张三",如果绑定了语音标签"小张",用户在使用语音识别时,说"小张"即可快速定位到"张三"。同样的原理,用户在使用导航的过程中,可以为某一个地名指定语音标签"公司",或"家",便可以在导航过程中说"导航到公司",使用非常便捷。

语音标签分为两种,特定人语音标签和非特定人语音标签。特定人语音标签是传统的、基于模板匹配的技术实现的,当受到噪声干扰,或者用户音色变化(如感冒)时,往往不能正确识别,实用性受到一定的限制。非特定人语音标签是基于最新的语音识别技术开发的,不受噪声、用户音色变化的影响的语音标签,甚至在更换用户时,仍能正确识别。

Aitalk 3.0 基于科大讯飞自主研发的语音识别技术实现了非特定人语音标签。

2.4 为中国人设计的英文识别

针对中国人说英语的特点,收集了大量录音,并进行了专题研究分析。中国人在读英文单词的,有的带有一些地方口音,有的部分音标有误读的情况。

产品对中国人的发音习惯定制后,英文识别上的相对识别率提升50%以上。

2.5 支持中文、英文、中英文混合识别

能同时支持中文、英文、中英文混合的命令词输入,并正确识别。

特别是中英文混合命令词输入,是语音识别的技术难点。在 Aitalk 3.0 中,诸如"Tony 黄"



之类的命令词, 也能被识别引擎所准确识别。

2.6 支持句式功能

相对于用户直接说人名 "Tony 黄",用户说出"打电话给 Tony 黄的手机"句子将会显得更加自然。Aitalk 3.0 支持句式功能,对于以上例子,可以将句式"打电话给{联系人}的手机"输入引擎,对于手机中的联系人,用户可以用说"句子"的方式来拨打电话了。

2.7 对符号、中文姓氏智能判断

对于常见符号的读法,引擎能自动决策并正确处理,如"3/4"。 自动决策中文的姓氏的特殊读法,如"仇晓庆"、"区建国"等。

2.8 同音字识别

如果识别集中含有拼音相同的词条(指汉语拼音相同,音调可以不同的词条,例如"张雷"和"张磊"), Aitalk 3.0 CE 能够正确进行处理,给出所有正确的识别结果。

例如: 手机电话簿中存有"张蕾"和"张磊"两条联系人名,用户说"张蕾"时,"张蕾"和"张磊"都会作为结果输出,输出顺序为这两词条添加到识别词集的顺序。

2.9 支持动态识别命令词

识别词条可以以文本形式动态添加。如手机中的联系人名、MP3 播放器中的歌曲名称等,都可以添加到引擎的识别集中,识别引擎可对加入到识别集中的所有词条进行识别,无需用户进行录音训练。

针对用户实际的使用场景,如联系人名的增删,歌曲的更新等操作,Aitalk 3.0 CE 支持识别词条的动态更新,只需要应用程序将更新后的识别词条加入或替换识别集合即可进行识别。

2.10 非特定人识别

不需要使用者录音训练,直接说出命令词的文本内容,即进行识别。



2.11 自动信道匹配

将外界声音转换为数字信号的系统,称为信道。一般情况下,信道包括 MIC,增益控制电路、A/D 等,在少数情况下,还包括降噪处理(Noise Reduction)、自动语音检测(Voice Activity Detection)等。

Aitalk 3.0 能自动匹配多种信道,并能对信道进行不断学习,使引擎在信道上的表现越来越好。Aitalk 3.0 会自动保存匹配参数(记录和读取数据需要客户配合)。

Aitalk 3.0 还提供信道定制(CMN)服务,使得在特定信道上,开始工作时就达到较好的状态。

2.12 拒识

对识别过程中外界的一些噪音干扰,如开、关门声,咳嗽声等,具有一定的拒识效果,提高了识别应用系统的抗干扰能力。

2.13 内置超时控制功能

用户在使用语音识别功能时,往往需要在规定的时间(如 **4S**)内进行语音输入。当超过规定的时间后,语音识别功能将自动取消。



3 联系我们

更加详细的产品信息请与科大讯飞取得联系:

● 联系地址

安徽省合肥市高新技术产业开发区黄山路 616 号讯飞语音大厦 安徽科大讯飞信息科技股份有限公司嵌入式事业部 邮编: 230088

● 联系电话

+86-551-5331878

● 传真

- +86-551-5331816
- +86-551-5331801
- +86-551-5331802

● 电子邮件

fae @iflytek.com

讯飞语音 沟通无限

安徽科大讯飞信息科技股份有限公司凭借深厚的语音识别技术积累以及雄厚的研发能力,将持续不断的研究嵌入式语音技术,提供最新最好的代表用户需求的语音识别系统,不断追求为用户创造价值的目标,推动语音信息化时代的发展。