

AIUI 产品白皮书



讯飞开放平台

2016 年 6 月 6 日

目录

AIUI 产品白皮书	1
1. 产品定义	4
2. 产品框架	4
3. 核心基础	4
3.1 交互全链路贯穿	4
3.2 紧密融合业务场景	4
4. 产品特性	5
4.1 远场精准识别	5
4.1.1 麦克风阵列降噪技术	5
4.1.2 远场识别引擎	5
4.1.3 方言口音适配	5
4.1.4 语义上下文自修正	5
4.2 全双工持续交互	5
4.2.1 持续录音,连续识别	5
4.2.2 拒识 - 过滤无效语音	5
4.2.3 基于语义的智能断句	6
4.2.4 随时可打断,支持主动式交互	6
4.3 上下文对话	6
4.3.1 上下文理解	6
4.3.2 基于内容提问	6
4.3.3 多对话场景管理	6
4.3.4 跨场景信息共享	6
4.3.5 长时记忆	7
4.4 个性化可扩展的交互	7
4.4.1 产品特性定制	7
4.4.2 用户个性化支持	7
4.4.3 交互模式可扩展	7
4.5 简单易用的产品集成	7
4.5.1 软硬一体化	7
4.5.2 云端配置定制一体化	7
5. 组件与服务	7
5.1 AIUI 核心模块（硬件）	7
5.1.1 麦克风阵列音频处理模块	8
5.1.2 核心处理模块	8
5.1.3 加密芯片	8
5.2 AIUI 核心算法	8
5.2.1 麦克风阵列算法	8
5.2.2 语音唤醒	8
5.2.3 交互核心算法	8
5.2.4 离线语音合成	8
5.3 云端服务	8
5.3.1 连续语音远场识别引擎	8



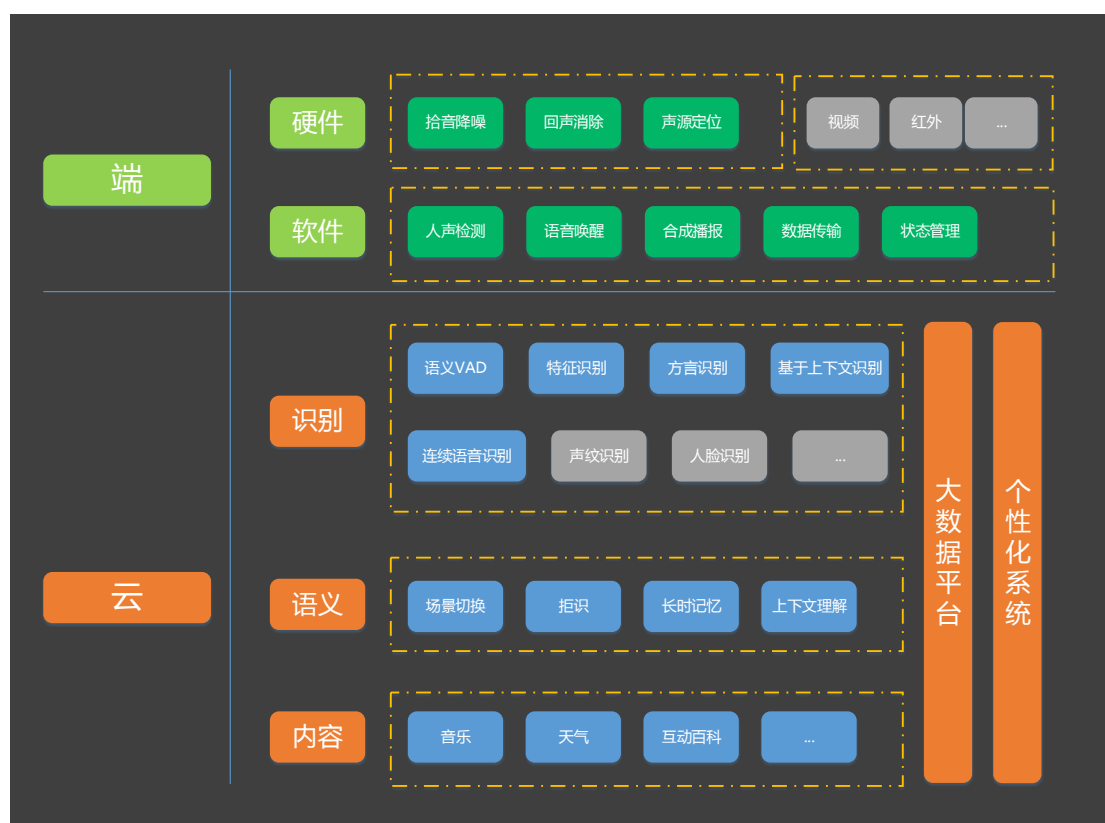
	5.3.2 雅典娜语义理解引擎.....	8
6.	适合领域.....	9
7.	联系方式.....	9

1. 产品定义

AIUI 是科大讯飞提供的一套人机智能交互解决方案，旨在实现人机交互无障碍，使人与机器之间可以通过语音、图像、手势等自然交互方式，进行持续，双向，自然地沟通。可以广泛应用多种智能硬件领域，如智能家居、机器人、智能穿戴等。

AIUI 是一套软硬一体的模块方案，贯穿了科大讯飞对整个人机交互流程的理解。方案包括前端声学处理、语音识别、语音合成、云端连续识别远场引擎、语义理解平台、内容平台、用户个性化系统等。在具备语音交互能力的同时，AIUI 也具备了第三方在平台上进行灵活能力配置、业务扩展的能力。

2. 产品框架



3. 核心基础

3.1 交互全链路贯穿

交互能力单点突破转变成链路整体演进思路，AIUI 方案贯穿整个交互链路，从宏观的角度进行综合优化，进而大大提升交互效果。

3.2 紧密融合业务场景

结合实际业务需求，不断改进、优化交互流程，不仅仅让人机交互更加流畅，更重要的是解决在生活中遇到的实际问题。

4. 产品特性

4.1 远场精准识别

4.1.1 麦克风阵列降噪技术

- ✧ 波束形成技术，有效去除各类噪声。
- ✧ 回声消除技术，消除自身播放的各类声音。
- ✧ 声源定位技术，准确定位声源角度，形成有效波束。
- ✧ 去除混响技术，有效去除混响，去除干扰。
- ✧ 适应多种构型，可以根据产品需求自由选择。

- 4 麦线性构型
- 5 麦（4+1）环形构型
- 8 麦（7+1）环形构型

4.1.2 远场识别引擎

适配远场拾音特征，优化识别算法，有效提升识别率。

- ✧ 1 米识别率达到 95%以上。
- ✧ 3 米识别率达到 92%以上。
- ✧ 5 米识别率在 90%左右。

4.1.3 方言口音适配

- ✧ 方言口音普通话自适应。对于非标准的，有一定的方言特点的普通话，具备很高的识别率。
- ✧ 支持客家语、四川话、河南话、东北话、天津话、湖南(长沙)话、山东（济南）话、湖北（武汉）话、安徽（合肥）话、江西（南昌）话、闽南语、陕西（西安）话、江苏（南京）话、山西（太原）话、上海话等接近 20 种方言的精准识别，识别率在 90%以上。开发者可以让用户根据实际情况，手动切换方言引擎。

4.1.4 语义上下文自修正

- ✧ 结合上下文动态修正。根据语音识别结果以及上下文对识别结果进行动态修正，使结果更符合当前语境。
- ✧ 应用场景热词优化。支持开发者针对应用所在的场景的热词进行整理，并通过开放平台免费的上传到云端，以提高场景内的热词语音识别率。

4.2 全双工持续交互

4.2.1 持续录音,连续识别

持续监听语音源，并将语音进行录音上传到云端。云端对接收的语音进行不间断的识别处理。

4.2.2 拒识 - 过滤无效语音

通过端点检测及语义处理等多种途径对上传的语音进行有效判断，将无效的语音进行过滤。

- ✧ 嗯、啊、哦等语气词。
- ✧ 与业务场景无关的语音。

4.2.3 基于语义的智能断句

通过语义对整个句子进行智能化断句。

- ✧ 对于短句，在结束后，很快判断出语句已经结束，并及时各处反馈。例子见图 1。

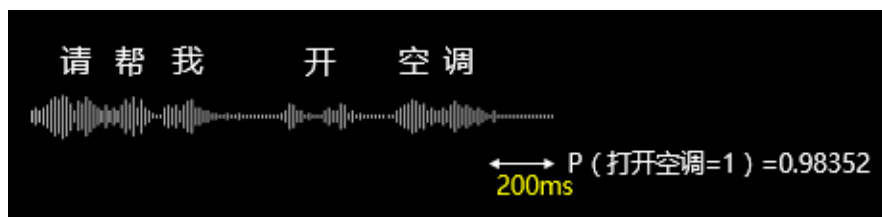


图 1 短句断句示例

- ✧ 对于长句中出现的语气词，及不自然的停顿，可以进行智能化的语句切分，从而获取完整的语音信息数据。例子见图 2。

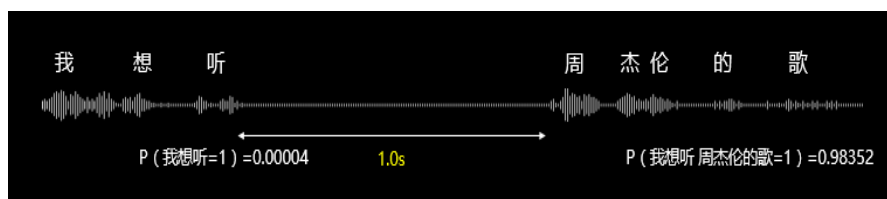


图 2 长句断句示例

4.2.4 随时可打断,支持主动式交互

- ✧ 用户可随时打断当前进行中的对话，进行下一次的对话。
- ✧ 机器可以根据场景，主动与用户进行交互。

4.3 上下文对话

4.3.1 上下文理解

结合交互的上下文，进行语义理解，以准确理解说话人的意图。

4.3.2 基于内容提问

对判断说话人意图所缺失的信息，会主动的针对所需要并且缺失的内容进行主动提问。。

4.3.3 多对话场景管理

同时管理多个对话场景，各个场景之间根据用户意图进行场景自适应级场景切换。

4.3.4 跨场景信息共享

场景切换之后，上一场景的信息可以保留到当前场景。

4.3.5 长时记忆

长时间的记忆用户交互历史，从而分析出用户的使用习惯及其他信息，并将这些记忆利用在交互中，使交互更智能。

4.4 个性化可扩展的交互

4.4.1 产品特性定制

- ✧ 唤醒词资源可自由定制，开发者可以根据自己的产品特性，为产品配置不同的唤醒。支持 3 个唤醒词同时使用，增强型的方案的唤醒率在 90% 左右。
- ✧ 合成发音人定制，提供萌萌、小燕、小宇、小梦、楠楠、台普 晓琳、晓倩（东北话）、小蓉（四川话）、小坤（河南话）、小强（湖南话）、小英（陕西话）、小梅（粤语）、凯瑟琳（美式英语）共 12 种发音人，开发者可自由选择。在此基础上，开发者还可以自选声音源，与 AIUI 团队有偿合作，训练制作发音人音库。
- ✧ 业务对话场景定制。AIUI 提供音乐、天气、诗词、火车票+飞机票、智能家居 5 个场景，根据开发者需求，逐渐开放更多的场景。在此基础上，开发者还可以联系 AIUI 团队有偿合作，进行产品所需场景的定制。

4.4.2 用户个性化支持

- ✧ 基于用户身份的个性化支持

提供用户身份识别服务，根据用户的语音输入，判断客户的性别、年龄等基本用户属性，并可以根据这些属性提供个性化的支持。基于用户环境的个性化支持与传感器配合使用，对用户所处的生活环境进行数据收集，并根据这些环境数据，给出个性化的交互。

4.4.3 交互模式可扩展

除语音之外的，类似于红外、人脸、手势等各类交互模式可自由扩展。

4.5 简单易用的产品集成

4.5.1 软硬一体化

讯飞开放平台提供软硬一体的 AIUI 模块，开发者可以忽略交互能力及流程调度，直接处理业务交互状态及结果。

4.5.2 云端配置定制一体化

交互的配置、语义及场景定制、开发者产品数据的分析查看均可以在平台上面进行。

5. 组件与服务

5.1 AIUI 核心模块（硬件）

5.1.1 麦克风阵列音频处理模块

处理多麦克风的音频汇聚。

5.1.2 核心处理模块

安装 AIUI 软件客户端。

5.1.3 加密芯片

保护 AIUI 算法。

5.2 AIUI 核心算法

5.2.1 麦克风阵列算法

包含远场拾音降噪、回声消除、声源定位等功能，支持多种麦克风构型。开发者可自由选择构型。

- ✧ 4 麦线性
- ✧ 4+1 环形
- ✧ 7+1 环形

5.2.2 语音唤醒

提供加强型唤醒引擎机资源。唤醒率可以达到 90%左右。
支持 3 个及 3 个以下不同语音唤醒资源。开发者可自由定制个性化唤醒词。

5.2.3 交互核心算法

包含持续录音、连续识别等语音交互前端处理功能

5.2.4 离线语音合成

提供离线语音合成服务。提供萌萌、小燕、小宇、小梦、楠楠、台普 晓琳、晓倩（东北话）、小蓉（四川话）、小坤（河南话）、小强（湖南话）、小英（陕西话）、小梅（粤语）、凯瑟琳（美式英语）共 12 种发音人，开发者可自由选择。在此基础上，开发者还可以自选声音源，与 AIUI 团队有偿合作，训练制作发音人音库。。

5.3 云端服务

5.3.1 连续语音远场识别引擎

支持语音连续识别，并对经过降噪算法处理后的远场音频有着非常好的适配性，大大提升语音识别率。

- 识别率 1 米左右可以达到 95%;
- 识别率 3 米左右可以达到 93%;
- 识别率在 5 米左右可以达到 90%。

5.3.2 雅典娜语义理解引擎

支持上下文理解，以及拒识、智能断句，同时支持多种语音场景。

- ✧ 音乐场景
- ✧ 天气场景
- ✧ 对诗场景
- ✧ 智能家居场景
- ✧ 订票系统场景

6. 适合领域

适合需要语音智能交互的各个领域，包括智能机器人、智能家居、智能穿戴、智能音箱、智能手表等。如图 3 所示：

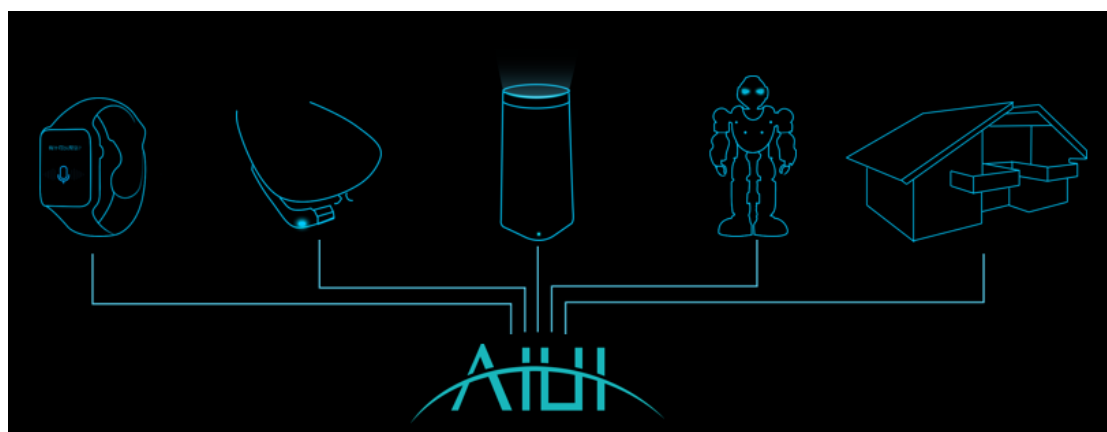


图 3 使用场景示例

7. 联系方式

官方网站: www.xfyun.cn

方案地址: www.xfyun.cn/aiui/index

商务合作: mssp_support@iflytek.com

技术咨询: bbs.xfyun.cn

开发者交流群: 153789256 / 385413567 / 91104836