

# 翻译耳机

## 功能

### 录音功能

- 外部录音：录制外部声音，可在app播放转写
- 系统录音：录制系统声音（通话录音、视频录音等），可在app播放转写

### 录音实时转写

- 可选语言/领域：可以选择不同的输入语言，不同专业领域
- 实时转文字：录音实时转为文字。并保存至云端。
- 声纹识别：识别人物声纹。录入声纹后可标明不同人物所说的话。

### 同声传译

- 录音实时转文字，将文字翻译成目标语言并播放出来。

### 会议纪要

- 会议记录摘要

## 相关技术实现

### 录音相关

- 设备/app实现，上传云端

### 录音实时转写

- 实时转文字：流式音频+asr实时输出结果
- 可选语言/领域：asr能力？
- 声纹识别：？

### 同声传译

- 第一种方式：asr实时结果（根据实时结果，划分转写范围进入翻译api）+百度翻译api+tts
  - 需确定asr结果分段导入翻译api逻辑（根据标点符号？根据字数？）
  - 会有一定延时，具体效果体验待测试。
  - 技术组成：
    - asr+翻译api（按字计算, 49元/100w字）+tts
- 第二种方式：asr实时结果（根据实时结果，划分转写范围进入翻译api）+文心千帆大模型实现翻译+tts
  - 需确定asr结果分段导入翻译api逻辑（根据标点符号？根据字数？）
  - 输入需增加prompt

- 现在你是一个专业的英汉互译器，输入中文将其翻译成英文，输入英文将其翻译成中文。我输入的内容是[今天的天气怎么样]
- 会有一定延时，具体效果体验待测试。
- 技术组成：
  - asr+文心大模型api（按token 0.012元/1000token）+tts
- 第三种方式：百度实时语音翻译（通过输入语音实时文本转写和翻译）+tts（翻译tts播放）
  - 百度完整的能力，延时相对较低，体验会好
  - 技术组成
    - 百度实时翻译api+tts
      - 百度实时翻译api，按时间计费（<https://ai.baidu.com/ai-doc/MT/TI9pjqsym>）
 

规格（小时）	价格（元）	并发限制
100	950	10
1000	9000	10
5000	42500	10

## 会议纪要

- 文心大模型会议摘要
  - 设计prompt