

# 基于树莓派的智能语音交互系统设计

吴友兰 胡冬琴

北京电子科技职业学院 亦庄开发区 100029

**摘要:** 本文给出了一套设计基于百度语音的智能语音交互系统的设计方案, 用户不仅可以通过语音控制外部设备的工作, 而且可以实现智能语音对话。该设计控制器采用树莓派为核心控制器, 用户可以通过语音信息的输入来对设备进行控制, 极大的降低了用户操作的复杂性, 提高了设备的使用效率。

**关键词:** 语音识别; 语音合成; 树莓派; AI; 百度语音; 图灵机器人

## 引言:

语音控制是一种有效便捷的控制方式, 让机器设备能够“听懂”人类的语言, 并做出正确的动作和反馈, 给机器设备加上耳朵和嘴巴从而实现人机交互的目标。语音识别技术(ASR技术), 就是让机器通过识别和理解过程把语音信号转变为相应的文本或命令的技术。语音合成技术(TTS技术)是通过机械的、电子的方法产生人造语音的技术。本文是针对语音识别和语音合成技术提供了三种解决方案来实现语音控制机器设备。

## 1 系统方案设计

我们在基于树莓派控制的设备中加入语音控制功能, 摆脱以往必须手动操作的麻烦, 通过语音就能达到控制相关设备的目的。下面我们提出三种方案来加入语音控制, 用户可以根据自己设计产品的情况选择一种进行应用。本次设计的系统框图如下图1。

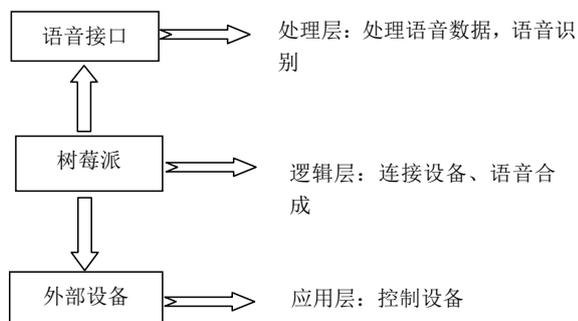


图1 系统框图

本次设计的语音控制系统采用设备+树莓派+语音接口来实现, 分为应用层、逻辑层和处理层。其中语音接口作为处理层, 对树莓派所录的语音信号进行识别,

**作者简介:** 吴友兰, 1978年2月, 女, 汉, 浙江桐乡, 本科, 北京电子科技职业学院, 100029, 讲师, 电子信息; 胡冬琴, 1972年12月, 女, 江西乐平, 汉, 本科, 北京电子科技职业学院, 100029, 讲师, 电子信息。

将识别后的语音信号返回给树莓派。树莓派作为逻辑层, 处理业务逻辑上的问题, 其中包括对处理层返回的识别后的语音信号进行存储, 以及相应操作。外部设备作为应用层, 负责系统的处理和控制在。

我们分别设计三种方案来解决上述问题:

第一个方案中我们需要购买相应的语音合成模块和语音识别模块, 并利用购买方提供的语音合成和语音识别库来编写程序实现处理层语音识别和逻辑层语音合成的功能。这种方案能够实现简单的语音控制, 无法实现智能语音对话功能。

第二个方案中我们只需要在设备中加入普通的USB耳麦和音箱, 利用第三方百度语音平台和图灵机器人来实现语音的识别和合成。这种方案中我们加入了图灵机器人所以能够实现智能语音对话功能。

第三种方案我们需要购买 ReSpeaker 2-Mics Pi HAT 双麦克风扩展板, 使用第三方 Google Assisant 来实现智能语音对话和设备的控制。这种方案配合树莓派使用, 可以搭建一个类似 Google 智能音响、亚马逊 Alexa 音响的智能音响。

## 2 系统方案实施

下面我们针对三种方案的实施做详细的讲解

方案一: 我们可以购买语音合成模块和语音识别模块, 这里我购买的模块如下图2, 大家可以根据需要在网上购买相类似的语音合成和语音识别模块。



图2 语音合成和识别模块

硬件模块包中厂家提供语音识别和语音合成库, 我们可以通过下面两条语句加入库

```
import HiwonderSDK.TTS as TTS  
import HiwonderSDK.ASR as ASR
```

在语音识别 ASR 库中提供了 addWords 函数用于添加唤醒词和关键词, 我们可以通过修改程序自行的更改唤醒词和语音识别的词条。在语音合成库 TTS 中提供了 TTSModuleSpeak 函数用于实现语音合成, 把文本用语音的形式进行播报。在处理层中我们通过 ASR 库中的相关函数对输入语音信号转变为相应的文本传输给树莓派, 树莓派对文本进行存储和语音合成播报, 在应用层中树莓派根据收到的文本信息对外部连接的设备进行控制。比如我们设计人形机器人时就采用这种方法, 通过语音输入前进后退左移右移来控制机器人的运动状态。

方案二: 在硬件的选取上我们采用通用的带 USB 驱动的耳麦和 3.5mm 接口的音箱, 在软件设计上程序主要包括以下四部分如下图 3:



图 3 程序结构图

1) 语音信号的采集和识别文件 yuyinshibie.py: 通过树莓派自带的录音功能来采集语音信号, 并保存。语音信号的识别是通过 python 调用百度语音接口, 识别采集的语音信号, 将识别后的语音信号以 wav 文件的格式存储于树莓派中。

2) 语音机器人的回复文件 yuyinhecheng.py: 读取存放识别语音信号的 wav 文件, 调用图灵机器人接口 Turling.py, 实现机器人智能回复。

3) 电器设备的控制 yuyin.py: 电器设备的控制其实就是对树莓派 GPIO 的控制, 我们可以控制连接在树莓派 GPIO 口上的设备。

该方案中我们是通过百度语音和图灵机器人来实现智能语音控制, 操作简单, 直接可以把写好的程序加入到树莓派控制设备中, 不需要做系统的配置和安装, 用户在现有已经做好的树莓派控制设置中加入语音可以选择这种方案简单易行。

方案三: 我们需要购买 ReSpeaker 2-Mics Pi HAT 双麦克风扩展板如下图 4, 该板是基于 WM8960 开发的低功耗立体声编解码器, 配合树莓派使用可以搭建一个类似 Google 智能音响、亚马逊 Alexa 音响的智能音响。



图 4 ReSpeaker 2-Mics Pi HAT 扩展板

我们需要进行硬件配置和软件的编写, 步骤如下:

1. 烧录系统, 登陆
2. 驱动下载并安装
3. 检查声卡名称是否与源代码 seeed-voicecard 相匹配
4. 录音播放测试
5. 安装 python 和虚拟环境
6. Alexa SDK 和 DuerOs SDK 语音引擎
7. 修改和运行语音程序

我们可以到 snowboy 官网上录制自己的声音来修改唤醒词, 并且通过修改程序添加语音词条来控制设备的工作状态, 这个方案接收的语音信号远, 效果好, 如果在远距离的语音控制中可以采用这种方法。

### 3 系统方案总结

对比三个方案, 方案一中我们调用购买厂家提供的语音识别和合成 SDK 来实现简单的语音控制, 比较适合一些简单的工业控制设备, 方案二我们调用百度语音 SDK 和图灵机器人作为智能控制, 方案三我们采用 Alexa 和百度 DuerOs 作为语音引擎, 开发出语音互动系统和设备的控制系统。方案二和方案三我们都能实现设备的控制和简单的智能对话, 只是在设备的硬件选取和软件的配置上不同, 制作出的效果也略有不同。大家可以根据自己需要选择一种方案来实现自己的语音控制设备的设计和制作。

#### 参考文献:

- [1] 张伟. 面向智能家居的嵌入式语音控制系统的研究[D]. 郑州: 郑州大学, 2015.
- [2] 周旭. 车载蓝牙语音控制系统的设计与实现[D]. 长春: 吉林大学, 2011
- [3] 禹琳琳. 语音识别技术及应用综述[J]. 现代电子技术, 2013, 36(13): 43-45.
- [4] 于俊婷, 刘伍颖, 易绵竹, 李雪, 李娜. 国内语音识别研究综述[J]. 计算机光盘软件与应用, 2014, 17(10): 76-78.
- [5] 汪开元, 胡大春. 基于树莓派的智能家居语音控制系统[J]. 科技与创新, 2017(22): 56-57.