

**短学期个人总结报告**

**项目名称：音乐播放器**

**组别： 刘畅组**

**学号：3130000157**

**姓名：杨光照**

目录

[1 产品设计 3](#_Toc456962068)

[2 技术实现 3](#_Toc456962069)

[2.1 语音识别 3](#_Toc456962070)

[2.2 手势识别 3](#_Toc456962071)

[2.3 树莓派平台整合 3](#_Toc456962072)

[3 个人心得 4](#_Toc456962073)

# 1 产品设计

如今生活中电子产品层出不穷，各种功能花样翻新推出，但其控制方式一直未能得到有效的改进，譬如电视机、空调等依旧依赖遥控操作，文字依旧通过键盘输入。所以我们试图创造一个将各种控制方式结合在一起的平台解决方案，将人与设备之间的多条路径打通，集合语音、手势、web端等多种操作方式，适应多种条件下对设备的有效调控。

该控制平台解决方案并不局限于某一种特殊的设备，因为平台可以根据设备的变换提供其相应的控制方式，只需调用平台接口再对通信信号进行适配便可实现对设备的控制。

# 2 技术实现

本次项目中我主要负责语音识别、手势识别和树莓派平台集成的三项开发任务。其中语音识别基于百度语音平台，手势识别通过LeapMotion实现，编程语言使用Python。

## 2.1 语音识别

语音识别主要由录音模块recoder.py、语音识别模块voice.py组成，录音部分使用麦克风对外部环境进行收音，若其达到程序设置好的声音阈值、连续性等一系列标准后，便进行录音，并将声音保存为一段采样率为48kHz的initial.wav文件。

语音识别部分通过向百度语音识别API借口发送一段包括录音文件、录音参数、识别语言等信息的JSON数据进行识别操作，并得到其返回的包含识别结果、信息等的JSON数据。由于百度语音仅支持8kHz或16kHz的WAV文件，而受麦克风限制树莓派智能录制48kHz的WAV文件，因此需要使用SOX软件将initial.wav转换为符合标准的voice.wav文件，在进行识别。

## 2.2 手势识别

手势识别基于LeapMotion硬件实现，该硬件与传统身体识别硬件Kinect相比有着成本低、专注于手指跟踪的优点，适用于仅适用手掌操作的场合。

本项目需要识别手掌顺、逆时针旋转、手指KEY TAP并对手掌坐标进行追踪，手掌旋转和KEY TAP通过调用官方SDK的gesture类函数进行识别，并调用position类函数对手掌位置进行跟踪，将其与0.5s前的手掌位置进行比较，根据三个维度坐标的变化发出上、下一首、音量加、减信号。

由于树莓派配置较低，无法完全正常驱动LeapMotion工作，只能退而求其次，将LeapMotion识别工作放在笔记本上进行，再将信号通过串口通讯传给树莓派处理。

## 2.3 树莓派平台整合

树莓派上使用mpg123软件实现音乐播放，PyAudio库对麦克风进行适配。其主要运行raspberry.py文件，该文件将语音、手势、web集合，综合实现音乐播放器。

# 3 个人心得

通过这次短学期，我对Linux、树莓派和python都在不同程度上有了更深的理解，三周时间里遇到的各种问题也都借助搜索引擎一一解决，最终的成品虽然还有这不小的遗憾，但也是一款较为成熟的控制平台解决方案。