

# 基于 GSM 的智能家居语音控制系统<sup>①</sup>

杨叶芬, 叶成景

(广东科学技术职业学院, 珠海 519090)

**摘要:** 在 Android 手机平台上提出了一种基于 GSM 的智能家居语音控制系统设计方案. 该方案通过 Android 手机的语音识别, 使得智能家居具备了远程语音控制的功能, 大大提高了智能家居的可操作性, 满足了智能家居的安防要求. 经测试, 证明了本系统的实时性、有效性和稳定性, 同时, 可以为人机交互提供一种切实可行的参考方案, 具有广泛的应用领域.

**关键词:** 语音识别; GSM 模块; 智能家居; Android 手机

## Intelligent Home Voice Control System Based on the GSM

YANG Ye-Fen, YE Cheng-Jing

(Guangdong Vocational Institute of Science and Technology, Zhuhai 519090, China)

**Abstract:** A design program of intelligent home voice control system based on GSM is presented on Android mobile phone platform. Via the voice recognition of Android mobile phone, the intelligent home can have a function of being controlled by remote voice through this program, which greatly improves the operability and the security requirements of the intelligent home. Test results prove the features of real-time, effectiveness and stability of the system. Meanwhile, this system can also provide a practical reference solution for human-computer interaction, having a wide range of application.

**Key words:** voice recognition; GSM module; intelligent home; Android mobile phone

## 引言

随着计算机网络技术的发展, “智能家居”越来越被人们所重视. 从发展趋势看, 智能家居的日益普及将是一种必然<sup>[1]</sup>. 但是我国智能家居仍处于初步发展阶段, 其平台性、应用的兼容性以及用户体验度都需要不断进行改善. Android 作为优秀的开源软件整体解决方案, 广泛应用于手机、平板电脑等设备. 开发出功能完善、操作简便、系统稳定的智能家居, 正在成为全球智能家居的发展热点<sup>[2]</sup>.

同时, 未来 10 年内, 语音识别技术将进入工业、家电、通信、汽车电子、医疗、家庭服务、消费电子产品等各个领域<sup>[3]</sup>. 语音识别技术的应用已经成为一个具有竞争性的新兴高技术产业<sup>[4]</sup>. 语音识别技术经过全球半个多世纪的研究, 目前已经发展到了接近实用阶段, 语音识别技术开始尝试从实验室演示系统走

向实用化商品<sup>[5]</sup>.

本系统结合语音识别和移动通信技术, 在 Android 手机平台上, 设计开发了一款智能家居语音控制系统, 该系统能够很好地满足现代控制领域智能化的特点, 通过手机语音控制将会极大地替代手机用户的短信方式, 大大提高了智能家居的可操作性; 同时, 实现了实时远程信息监控, 具有广阔的应用前景<sup>[6-8]</sup>.

## 1 架构设计

本系统采用 Android 手机的传声器进行语音录入, 获取录入的语音信号的声压值并将其保存为 txt 文本形式, 在设置口令时, 将保存的声压值文本作为匹配模板(即识别口令的基础语音模板); 在声控口令时, 将声压值文本作为检测模板, 与之前保存的匹配模板进行比较, 比较过程就是语音识别过程. 识别时, 采用

① 基金项目: 广东省高等职业教育一类品牌专业资助项目(2016gzpp007); 广东省高等职业教育教学改革项目(201401099)

收稿时间: 2016-05-18; 收到修改稿时间: 2016-07-14 [doi:10.15888/j.cnki.csa.005615]

DTW 算法作为语音识别引擎对两个模板进行分析,从而识别出语音控制命令.系统反馈给用户识别的结果,并且将识别出的命令以短信方式发送,GSM 模块收到命令后,智能家居主控机根据实际需求进行识别结果的处理,调用通信串口执行相应的功能模块,并向手机端返回处理结果,如图 1 所示.

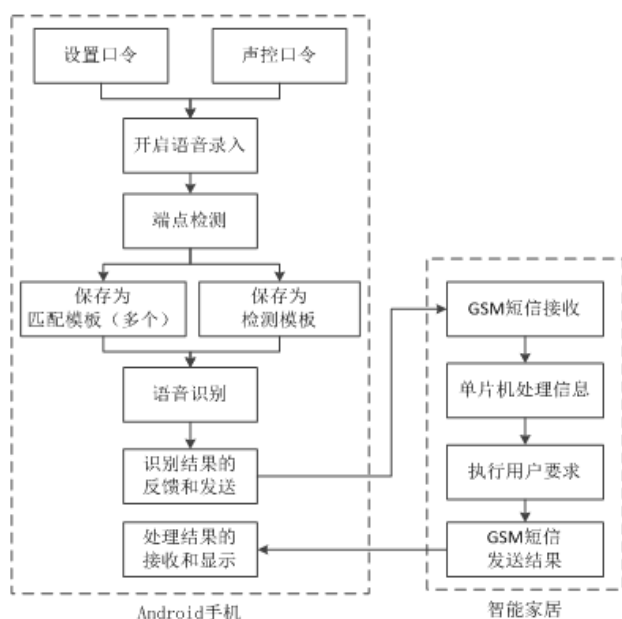


图 1 系统架构

### 1.1 语音录入

语音录入是采用 Android 手机的 AudioRecord 进行的音频采集,采集后的音频格式为 PCM-16bit 的格式,该格式文件是无压缩的声音源,信号是由 0、1 等符号构成的数字信号.

### 1.2 端点检测

得到 PCM 格式文件后,需要对语音片段进行端点检测.端点检测即是检测语音片段有效部分开始与结束的两个端点,检测方法是提取语音片段的振幅,利用振幅提取有效的语音片段(无静音或杂音的语音片段),并保存该语音片段的声压级为 txt 文本作为语音模板.

### 1.3 模板保存

得到有效的语音片段和振幅值后,不能仅仅依靠这两者进行语音识别,因为振幅值的数据量非常大,用于语音识别会严重影响识别效率,所以需要根据由振幅值得到的语音片段,通过提取片段的声压级,保存成 txt 文本形式,从而作为语音识别的模板.模板分

为匹配模板和检测模板,匹配模板为设置口令时保存的模板,相当于正确的“密码”;而检测模板为用户使用时说出的声控口令得到的模板,即输入的“密码”.

### 1.4 语音识别

语音识别,即是利用 DTW 算法对数据进行处理,处理的数据为设置口令得到的匹配模板和声控口令得到的检测模板,经过算法处理得到识别结果,针对智能家居的功能模块,其结果中包括设备的名称,如电灯、电视等,和需要执行的开、关等操作.

### 1.5 识别结果的反馈和发送

识别结果的处理是指将语音识别引擎得到的结果转化为设备的 ID 号与要执行的操作命令,以短信方式发送到 GSM 模块.

### 1.6 GSM 短信接收

GSM 模块中设置手机号码与语音录入的手机进行通讯.当 GSM 模块接收到手机发出的短信时提取短信内容进行分析,并译码后响应执行相应的操作.

### 1.7 单片机处理信息

单片机从 GSM 模块中提取短信息分析的结果做出相应的动作去控制相应的设备.当单片机动作完成后通过 GSM 模块向设定好的手机号码发出短信,表明单片机完成手机之前发出的口令.

## 2 系统的软件实现

### 2.1 系统软件处理过程

本系统软件部分采用 Java 语言和基于 Android 手机平台的应用系统开发,实现语音识别功能.为了让功能模块更好地运行于不同平台(软件、硬件),从平台无关性的角度出发,基于 Android SDK(Software Development Kit)设计,将语音识别模块化,并做成以 APK(Android Package)形式安装.把语音识别制作成系统里面的服务,供软件本身和其它软件的使用.

本系统的语音识别功能是在 Android 下实现的, Voice Control Service 模块在 Android 空间下基于语音识别库来实现语音识别功能, Voice Control Service 在整个系统软件中的层次关系如图 2 所示.

语音识别模块的设计可分为语音识别引擎库调用、Voice Control Service 实现以及 UI 设计 3 个部分.

#### 2.1.1 语音识别引擎库调用

Voice Recognition JNI 对语音识别引擎提供的 API 进行封装,向 Voice Control Service 提供语音控制的

Java API. 封装接口主要包括识别引擎的初始化、识别器的创建、识别结果的回调处理等. 在初始化识别引擎后, 根据应用需求创建本地识别器, 启动开始识别过程, 通过 Java 回调函数通知 Java 层获取语音识别结果, 根据应用场景分别进行处理.

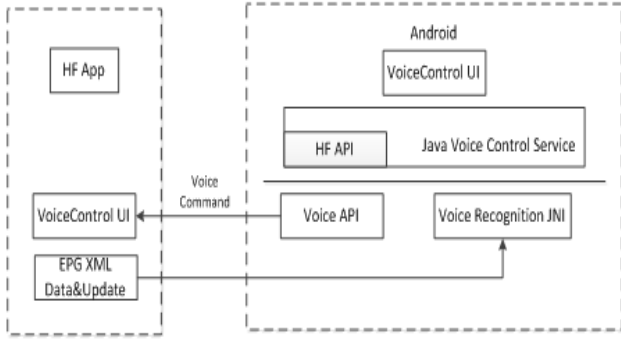


图2 语音识别层次关系图

### 2.1.2 Voice Control Service 实现

语音识别引擎通过匹配得到语音识别结果的字符串, 在 Framework 层还需创建语音识别本地 Service, 对识别结果进行处理.

语音识别服务针对不同声音源识别到系统接受到语音命令. 当系统接收到语音时, 若软件语音识别服务开启, 则对语音数据进行识别.

根据语音识别回调函数解析返回的内容和语音识别状态信息, 软件将识别结果转变成信息, 通过 GSM 模块转发到智能家居的主控机.

### 2.1.3 Voice Control Service UI 设计

组件 Voice UI 是语音识别过程、结果的信息显示. Voice UI 开启/关闭与 Service 同步, 具体是当用户通过菜单开启、关闭 Voice 功能时, Android 端发送信息到智能家居模块端的事件将开启、关闭 Voice Service 和 UI. UI 包括 Android 应用程序的 UI 设计.

Voice Control Service 与 Voice Control Service UI 通过广播的方式进行通信, 基于以下两点:

- 1) Voice Control Service UI 注册 Service 广播, 接收广播的数据进行界面更新.
- 2) Voice Control Service UI 可以用按钮对智能家居的开关进行管理.

## 2.2 语音识别原理

目前常用的语音识别算法有动态时间规整(DTW)技术、矢量量化(VQ)技术、隐马尔可夫模型(HMM)、

基于神经网络(ANN)的识别方法. 在孤立词语的识别方面, DTW 算法与其他算法的识别效果相差很小, 而且由于 DTW 算法计算量较小, 无需前期的长期训练, 便于算法的移植, 并且易于在微处理器上实现并保证其实时性. 为提高系统识别率、灵敏度和安全性, 本系统是基于小识别词汇量的特定人语音识别, 采用 DTW 算法实现, 解决了发音长短不一的模板匹配问题, 其识别原理图如图 3 所示.

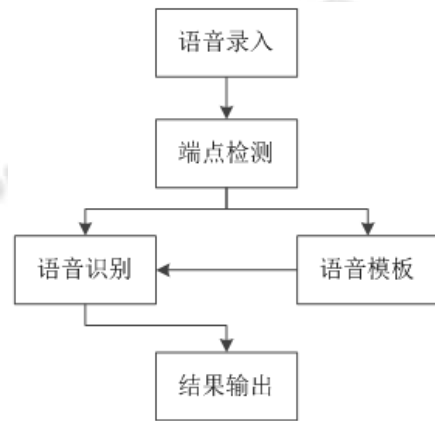


图3 语音识别原理图

### 2.2.1 端点检测

语音信号一般可分为无声段、清音段和浊音段. 无声段是背景噪声段, 平均能量最低; 浊音段为声带振动发出对应的语音信号段, 平均能量最高; 清音段是空气在口腔中的摩擦、冲击或爆破而发出的语音信号段, 平均能量居于前两者之间. 清音段和无声段的波形特点有明显的不同, 无声段信号变化较为缓慢, 而清音段信号在幅度上变化剧烈, 穿越零电平次数也多. 经验表明, 通常清音段过零率最大. 端点检测就是首先判断有声还是无声, 如果有声, 则还要判断是清音还是浊音. 为正确地实现端点检测, 一般综合利用短时能量和过零率两个特征, 采用双门限检测法.

#### (1) 短时能量

设第  $n$  帧语音信号  $X_n(m)$  的短时能量用  $E_n$  表示, 则其计算公式为:

$$E_n = \frac{1}{2} \sum_{m=0}^{N-1} x_n^2(m) \quad (1)$$

其中,  $N$  为信号帧长.

#### (2) 过零率

$$Z_n = \frac{1}{2} \sum_{m=0}^{N-1} |\text{sgn}[x_n(m)] - \text{sgn}[x_n(m-1)]| \quad (2)$$

其中,  $\text{sgn}[]$ 是符号函数, 即:

$$\text{sgn}[x]= \begin{cases} 1, & (x > 0) \\ -1, & (x < 0) \end{cases} \quad (3)$$

根据信号设置三个阈值: 能量阈值 TL、TH; 过零率阈值 ZCR, 当某帧信号大于 TL 或者大于 ZCR 时, 认为信号的开始、起点, 当大于 TH 时, 则认为正式的语音信号, 如果保持一段时间, 则确认这信号即所需信号。

### 2.2.2 语音识别

语音识别是按照一定的准则和测试与匹配模板进行比较, 通过判断得出识别结果. 处理步骤如图 4 所示, 将待测的语音信号, 经与匹配模板同样的处理步骤求得声压级后再与前述的匹配模板对比, 找出最相似的匹配模板(口令)作为辨识的结果。

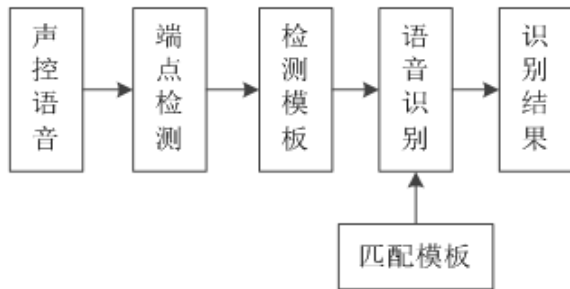


图 4 语音识别的处理步骤

对于语音识别部分, 采用比较模板声压级与输入语音的声压级的相似程度来实现. 根本思想是寻找一个最佳函数  $\phi(i_n)$ , 使得该函数满足模板声压级与输入语音的声压级的累积距离最短, 其公式为:

$$D = \min_{\phi(i_n)} \sum_{i_n} (d(T(i_n)), R(\phi(i_n))) \quad (4)$$

其中,  $(d(T(i_n)), R(\phi(i_n)))$ 表示待测矢量第  $i_n$  帧与训练模板中第  $i_n$  帧之间的距离。

## 3 系统的硬件实现

### 3.1 系统硬件组成

本系统硬件部分主要是智能家居模块的主控机, 由 51 单片机 STC89C52、传感器单元电路、GSM 模块和固态继电器控制具体设备等组成, 如图 5 所示。

### 3.2 GSM 模块

GSM 模块是将 GSM 射频芯片、基带处理芯片、存储器、功放器件等集成在一块线路板上, 具有独立

的操作系统、GSM 射频处理、基带处理并提供标准接口的功能模块. 因此, GSM 模块具有发送 SMS 短信的基于 GSM 网络进行通信的功能。

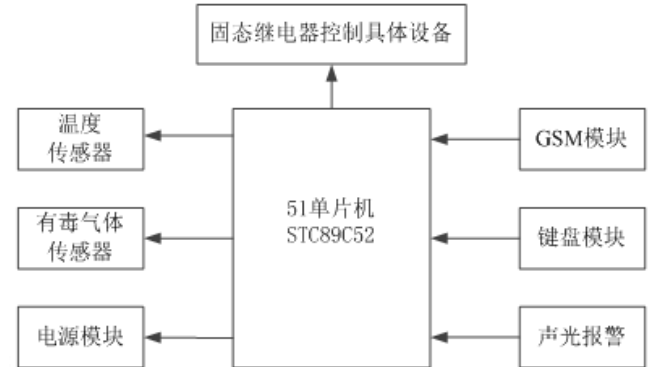


图 5 主机硬件结构框图

开发人员可以方便使用单片机通过 RS232 串口与 GSM 模块通信, 使用标准的 AT 命令来控制 GSM 模块实现发送和接收短信的无线通信功能。

在未触发状态, 单片机会不断检测判断 GSM 是否收到短信, 如果 GSM 收到短信, 则单片机会检测到 GSM 发过来的指令“+CMTI”。具体代码如下:

```
if(((s5=='+')&&(s4=='C')&&(s3=='M')&&(s2=='T')
&&(s1=='I'))||(p37==0))
{
    send(); //提取 GSM 接收的短信
}
```

在手机端进行语音识别时, 会根据不同的识别结果, 给 GSM 发送不同控制命令, 比如手机端识别到语音“电灯”, 发送字节命令“L”; 识别到语音家居编号“1”, 发送字节命令“01”; 识别到语音信号“开”, 发送字节命令“O”。单片机会根据 GSM 提取的短信进行判断, 并执行相应的操作. 具体代码如下:

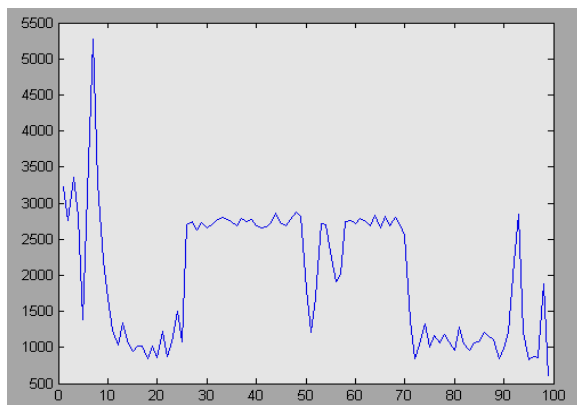
```
if((s1=='O')&&(s4=='L')&&(s2=='01'))
{
    p21=0; //若命令“L01O”, 则开第一盏电灯
}
```

## 4 系统测试结果分析

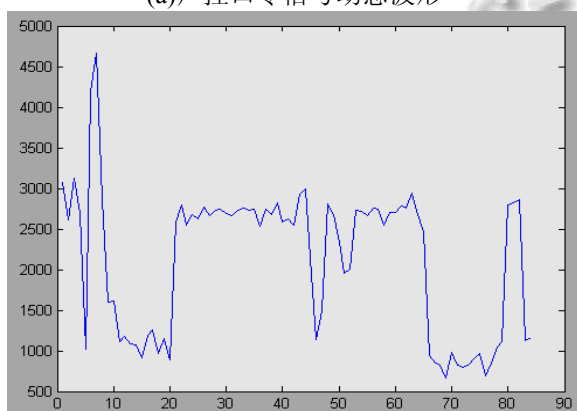
### 4.1 系统软件测试

系统软件测试部分是 Android 手机语音识别功能测试. 如图 6 所示为特定人录入设置口令和声控口令的信号动态波形. 图中 X 轴为采样点数, Y 轴为声压级。

可以看到语音信号的两次比较结果是相似的,即两段语音片段的声压级结果有着很高的相似度,这说明用于语音识别的结果匹配是可信的。



(a)声控口令信号动态波形



(b)设置口令信号动态波形

图6 声控口令和设置口令的信号动态波形图

为了更好地检测语音识别效果,实验中选择两个人在实验室安静和较为噪声的环境下分别进行测试,对设置口令的匹配模板随机测试,每组测试50遍,由系统输出识别结果的方式得到测试数据,测试结果如表1所示。从表1测试数据可以看出,语音识别率平均达到96.5%,证明了语音识别模块的可行性,可以满足语音控制的目标。

表1 口令模板测试结果

	特定人1	特定人2
安静环境	50/50	49/50
噪声环境	48/50	46/50
识别率/(%)	98%	95%

#### 4.2 系统硬件测试

系统硬件测试部分是智能家居远程控制功能测试。通过对语音识别后的结果与硬件接收的结果的一致性

进行比较,即判断手机端发送的命令与硬件端接受的命令是否一致,接着在两者一致的基础上观察判断硬件对家居设备操作是否成功,测试数据如表2所示。

表2 智能家居远程控制测试结果

发送命令	获取命令	命令结果 (一致/不一致)	控制结果 (成功/失败)
L01O	L01O	一致	成功
L03C	L03C	一致	成功
T01O	T01C	一致	成功
T02O	T02C	一致	成功

表2中,L表示电灯,T表示电视,数字为家居编号,O表示开,C表示关。

#### 5 结语

本文通过分析语音识别的基本原理和GSM技术在智能家居中的应用,在Android手机平台上设计了一款智能家居语音控制系统;重点阐述了系统的设计方案以及主要硬件组成和软件实现思路,通过对特定人语音识别的训练和测试表明,本系统能够以较高的识别率和灵敏度完成对智能家居的远程控制,系统实时输出识别结果,识别速度快,识别率可以达到96%,在语音远程控制领域具有广阔的市场前景和应用推广价值。

#### 参考文献

- 高小平.中国智能家居的现状与发展趋势.低压电器,2005(4):18-21.
- 付蔚,刘谋黎,蔡林沁.智能家居系统嵌入式家庭服务器设计与实现.自动化与仪表,2013,28(4):37-40.
- 姜干新,陈伟.嵌入式语音识别系统中的DTW在线并行算法.计算机应用研究,2010,27(3):977-980.
- 崔毓菁.语音识别系统速度优化算法研究[学位论文].北京:北京邮电大学,2008.
- 郑微,李正周,田蕾.基于DSP的智能语音控制系统设计.传感器与微系统,2012,31(2):134-136.
- 邓响,程小辉.面向物联网的智能家居系统设计.桂林理工大学学报,2012,32(2):259-264.
- 曾志永,凌振宝,王君.基于GSM技术的智能家居系统的设计.电子技术应用,2005,31(9):31-33.
- 徐子豪,张腾飞.基于语音识别和无线传感网络的智能家居系统设计.计算机测量与控制,2012,20(1):180-182.