

智能家居语音控制系统的设计

Design of the Voice Control System for Smart Home

付蔚 唐鹏光 李倩

(重庆邮电大学工业物联网与网络化控制教育部重点实验室,重庆 400065)

摘要: 对非特定人语音识别技术在智能家居系统中的应用进行了研究,并结合物联网的相关特点和最新发展动向,设计了智能家居语音控制系统。系统由移动终端控制软件和嵌入式便携语音控制器构成,其中移动终端控制软件采用了语音云技术,嵌入式便携语音控制终端采用 LD3320+STM32+CC2430 的系统构架并结合 802.15.4E 无线通信协议进行设计。实践证明,该语音控制系统实现了对智能家居可控设备的多样化、全方位语音控制。

关键词: 智能家居 语音控制 语音云 无线通信 MCU

中图分类号: TP29 **文献标志码:** A

Abstract: The application of speaker-independent voice recognition technology in smart home system is researched. Combining with the associated features and the latest development trends of Internet of things, the voice control system for smart home has been designed. The system consists of mobile terminal control software and embedded portable voice controller; in mobile terminal control software, the voice cloud technology is adopted. The embedded voice control terminal is designed by using the systematic framework of LD3320+STM32+CC2430, and combining with 802.15.4E wireless communication protocol. The practice verifies that this voice control system implements diversified and comprehensive voice control for controllable devices in smart home.

Keywords: Smart home Voice control Voice cloud Wireless communication Micro controller unit(MCU)

0 引言

智能家居是将家庭中各种通信设备、家用电器和安防设备连接成为一个家庭智能化系统,以便对上述设备进行监视、控制管理和信息交换,提供安全、高效、舒适、便利的家居环境^[1]。

传统的智能家居控制仍然依赖于人工触发式控制,语音控制技术的引入将使得智能家居的控制不再单调。无论用户身处室内或室外都可以通过语音控制技术,实现对家庭内部设备的控制。语音是语言的声学表现,是人类交流信息最自然、最有效、最方便的手段。语音控制技术的使用将“人与家居”整合在一起。语音控制技术将改变现有的用户与家居设备的信息交流方式,解放用户的双手,使用户的家居生活具备更人性化的魅力。语音控制技术让人与机器之间的沟通交流更加智能化,用户可以像与人说话一样,用自然化的

人类语言操作语音控制系统,从而实现语音控制系统的人机交互。

1 语音控制系统设计方案

本文设计的智能家居语音控制系统主要是采用先进的非特定人语音识别^[3]技术,并结合对识别到的内容解析和控制指令下发来实现对智能家居内部可控设备的控制,为用户打造更加贴心、舒适的个性化智能家居体验。

整个智能家居语音控制系统主要由移动终端语音控制软件和嵌入式便携语音控制器两部分构成。

安装有移动终端语音控制软件的终端在接收到本地语音信息后,通过网络(3G、GPRS、WiFi)将本地语音信息发送到讯飞语音云进行识别,再将云端反馈的数据解析并与本地设置指令匹配,获得正确的控制指令。通过服务器和网关将这些指令转发给相应的设备进行动作指令的执行,我们可以随时随地对家庭内部设备进行语音控制。

嵌入式便携语音控制器在家庭内部通过家庭网关融入到智能家居无线传感网中,终端在本地获取语音信息后进行识别并将相应的控制指令转发给网关,经由网关和服务器转发给相应的设备进行动作指令的执行^[4]。

国家“863”计划基金资助项目(编号:2012AA040602);

科技部国际科技合作专项基金资助项目(编号:2012DFG12240);

重庆市工程技术研究中心建设基金资助项目(编号:cstc2011pt-gc40006)。

修改稿收到日期:2013-05-06。

第一作者付蔚(1981-),女,2008年毕业于重庆邮电大学控制理论与控制工程专业,获硕士学位,讲师;主要从事计算机网络技术、物联网技术的研究。

系统网络拓扑结构如图1所示。

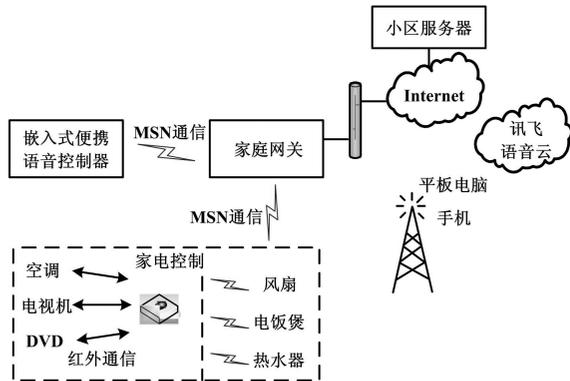


图1 系统网络拓扑结构图

Fig.1 Structure of the system network topology

2 语音控制器设计方案

2.1 硬件设计

整个嵌入式便携语音控制器的硬件电路主要包括MCU电路、语音识别电路、WSN电路和电源管理电路四个部分。由于嵌入式便携语音控制终端必须具备便携、微型化、低功耗的特点^[5],这就要求MCU芯片具有更高的集成性和低功耗性,便于进行低功耗功能设置。因此,MCU采用的芯片是基于Cortex-M3架构的STM32F103RX。

STM32F103RX芯片作为32位标准精简指令集(RISC)处理器,在通常8位和16位系统的存储空间上发挥了ARM内核的高性能;内部集成了Flash存储、SRAM、丰富的I/O接口和先进的中断响应系统等,更加适合应用在本方案中^[6]。

本方案中的语音控制芯片采用的是ICRoute的LD3320,WSN通信采用的芯片是CC2430。终端经由按钮触发启动,MIC在接收到外部语音信息后,首先经过语音识别模块对信息进行解析,以获取需要的数据;然后将数据通过串口发送给MCU,MCU经过处理后解析出相应的设备动作指令;之后经串口将信息发送给CC2430,最后发给家庭网关,从而实现了对智能家居内部可控家电的无线语音控制。语音控制器硬件框图如图2所示。

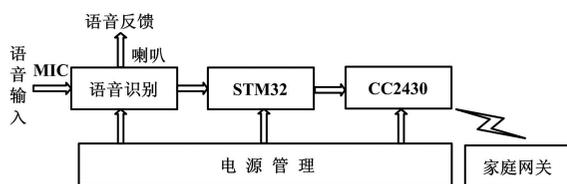


图2 语音控制器硬件框图

Fig.2 Block diagram of the hardware of voice controller

2.1.1 语音识别模块设计

语音识别模块电路参照ICRoute发布的LD3320数据手册进行设计。芯片内部集成了快速稳定的优化算法,不再需要外接Flash、RAM,大大简化了电路的设计。同时LD3320也不需要用户在使用过程中事先训练和录音来完成非特定人语音识别,识别准确率也比较高。LD3320采用并行方式直接与STM32芯片相连,并均采用1kΩ电阻上拉。A₀用于数据段与地址段的判断,控制信号RDB、WRB、CSB,复位信号RSTB和中断返回信号INTB直接与STM32芯片相连。LD3320与STM32芯片采用同一个外部时钟(8MHz),发光二极管D₁、D₂都用于复位后的上电指示,MBS引脚接一个RC电路作为MIC偏置,用于保证能输出一个浮动电压给MIC^[7]。

语音信息通过MIC输入后先进行频谱分析,提取出声音的语音特征信息;然后语音识别器将关键词列表中的数据与采集到的语音特征信息进行对比匹配,找出其中得分最高的关键词语作为语音识别结果输出。

LD3320芯片语音识别原理如图3所示。

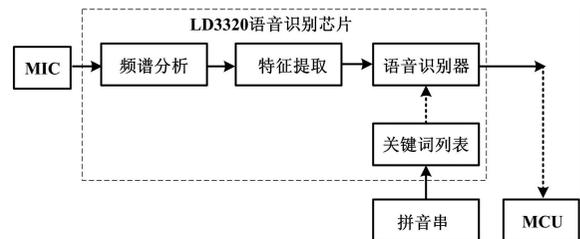


图3 LD3320芯片语音识别原理图

Fig.3 Speech recognition principle of LD3320

2.1.2 WSN通信电路设计

无线传感网通信协议采用820.15.4E^[8],选用型号为CC2430的无线SOC芯片。该芯片有丰富的片上资源,其内部已经集成了必要的电路,使用时只需要采用较为简单的外围电路,即可实现数据的无线收发。CC2430外围电路原理图如图4所示。图4中,Y₁为32.768kHz晶振,其与两个电容(C₁、C₂)组合构成32.768kHz晶振电路;Y₂为32MHz的晶振,其与两个电容(C₃、C₄)组合构成32MHz晶振电路;电容C₅和电感L₁、L₂、L₃以及一个PCB微波传输线组合构成WSN电路的非平衡变压电路,外围电压都配置滤波电容,以保证芯片工作的稳定性。整个电路设计满足RF输入/输出匹配电阻(50Ω)的指标。另外,在电压脚和地脚都增加了滤波电容,用于为芯片稳定工作提供保障。

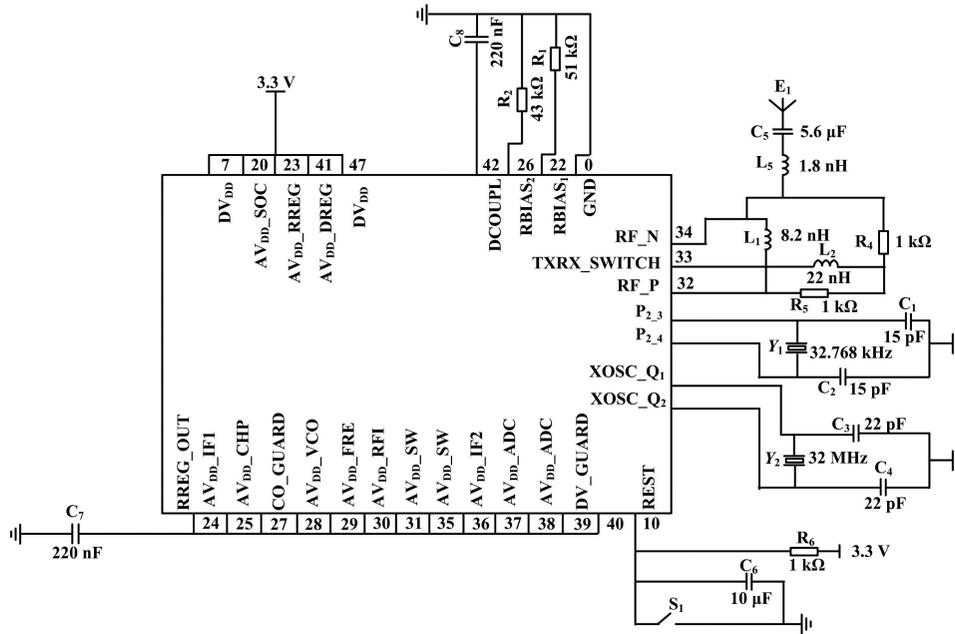


图 4 CC2430 外围电路原理图

Fig. 4 The principle of CC2430 peripheral circuit

2.1.3 电源管理电路设计

由于嵌入式便携移动终端需要具备便携、无线、低功耗等特性,所以在电源管理部分就采取了大容量、高性能的锂电池供电。电源管理电路主要包括锂电池充电电路、锂电池保护电路、电压转换电路三个部分。其中,锂电池充电电路主要采用了 MAX1555 芯片和相应的电阻、电容来实现。锂电池保护电路由 FS8820P 电池保护芯片、电阻、电容组成。由于终端硬件电路所用芯片工作电压均为 3.3 V,而锂电池输入电压为 0 ~ 4.2 V,所以需要采用 MAX8881 进行电压转换。

2.2 软件设计

嵌入式便携语音控制终端的软件设计主要分为 MCU 程序设计和语音识别程序设计,其中语音识别程序设计是整个程序设计的核心。参考 ICRoute 的 LD332X 开发手册,在本方案中采用中断方式工作,整个程序主要分为通用初始化、语音识别初始化、写入识别列表、开始识别、响应中断几个流程。

① 通用初始化和语音识别初始化:在初始化程序中,主要完成芯片的软复位、模式设定、时钟频率设定、FIFO 设定等相关操作。

② 写入识别列表:LD3320 语音识别列表的规则是每个识别条目对应一个特定的编号(编号可以相同,可以不连续,但是数值要小于 256),最多支持 50 个识别条目。每个识别条目是标准普通话的汉语拼音(小写),每 2 个中文字之间用一个空格间隔。

③ 开始识别:开始识别程序的流程图如图 5 所

示,此部分程序主要是设置几个相关的寄存器。ADC 通道即为麦克风输入通道,ADC 增益也就是麦克风音量,可设定值 00H ~ 7FH,ICRoute 官方建议设置值为 40H ~ 6FH。ADC 增益值越大,代表 MIC 音量越大,识别启动越敏感,但可能带来更多误识别;反之,ADC 增益值越小,代表 MIC 音量越小,需要近距离说话才能启动识别功能,优点是对远处的干扰语音没有反应。本方案采取的 ADC 增益值为 44H。

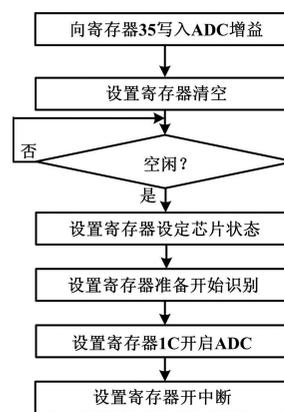


图 5 开始识别程序流程图

Fig. 5 The flowchart of recognition beginning process

④ 响应中断:当麦克风采集到声音时,就会产生一个中断信号。而中断程序要根据寄存器的值分析结果(BA 寄存器的值为候选识别结果个数,C5 寄存器里的识别结果是得分最高、最可能正确的答案)。

为了给用户提供更加贴心的语音控制服务,本

方案设计了对话管理模块(通过定义二维数组实现),以存储待识别语句和待执行命令。根据 LD3320 芯片自身特点,我们定义的对话管理数组如表 1 所示。

表 1 对话管理模块数组

Tab. 1 The design of dialogue management module array

名称	存储内容	定义数组
识别数组	待识别语句	INT Rec[50][13]
行为数组	对应行为编号	INT Act[50][8]

3 移动终端语音控制软件设计

本方案的移动终端控制软件主要基于 IOS 和 Android 两大移动操作系统的手机、平板电脑,将讯飞语音云引入移动终端语音控制软件开发中,不需要在本地建立大型语音库。这种设计在简化软件开发过程的同时,又扩大了软件使用人群^[8-10]。软件工作流程图如图 6 所示。

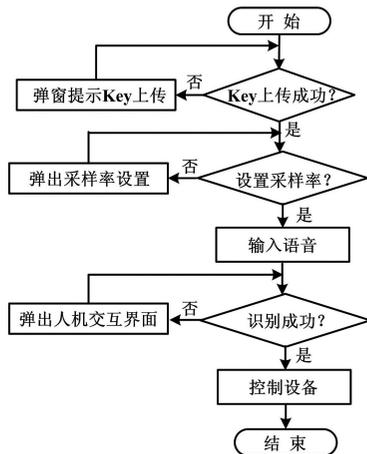


图 6 移动终端语音控制软件流程图

Fig. 6 Flowchart of mobile terminal voice control software

从图 6 可以看出,当开启软件后,首先要设置语音采样率,一般手机都只支持 8 kHz、16 kHz 两种采样率。为了获得更好的识别效果,默认设置都为 16 kHz,一些移动终端平台可能存在 MIC 性能参数不同,有的需要设置语音信息采样率(8 kHz、11 kHz、16 kHz、22 kHz)。点击上传后,手机和平板先将本地家庭内部的库文件(名称为 Keys,默认从 SD 卡读取,如果 SD 卡中不存在,则读取程序包 aassets 下的 Keys)上传至云端;然后,再将采集到的用户本地语音信息上传到云端进行识别,识别后在云端与用户家庭内部库文件进行匹配,并反馈回终端(确保识别信息是家庭内部已经设置的指令内容);最后,终端根据相应的控制信息下发相应的指令给服务

器,服务器进行动作指令的接收与转发,进而实现对智能家居内部可控设备的控制。

4 系统测试

4.1 嵌入式便携移动终端测试

针对嵌入式便携移动终端测试,目前主要对通信距离测试、射频穿透性、识别准确性三个方面进行了测试^[10],通过测试为终端后续的优化设计奠定基础。

4.1.1 通信距离测试

通信距离的测试是在设备带天线增益情况下进行的,嵌入式便携移动终端每一次语音识别都要将识别后的控制指令发送给网关,所以我们在不同距离下使用设备 50 次,记录控制指令成功转发给网关的次数,得到设备在不同距离下的丢包率。测试结果如表 2 所示。

表 2 通信距离测试结果

Tab. 2 The test results of communication distance

序号	通信距离/m	丢包率/%
1	25	8.3
2	50	16.7
3	75	31.7

从表 2 可以看出,设备丢包率随着通信距离的增加而变大。造成这一结果的原因是多方面的,不仅与测试环境中存在大气、电磁干扰、空气湿度、阻挡物、多径损耗有关,而且与天线增益、设备高度、性能匹配和协议自身原因都有一定的关系。

4.1.2 射频穿透性测试

射频穿透性测试主要是模拟设备与网关之间的墙体障碍物数目在 0 道、1 道、2 道不同情形下使用设备 50 次,记录在不同情形下控制指令成功转发给网关的次数,得到设备在不同障碍物数目下的丢包率。测试结果如表 3 所示。

表 3 射频穿透性测试结果

Tab. 3 The results of RF penetration tests

序号	障碍	丢包率/%
1	0 道墙	3.3
2	1 道墙	10.0
3	2 道墙	18.3

从表 3 可以看出,在实际的家居环境中,随着设备与网关之间墙体障碍物数目的减少,设备丢包率也随之降低,其原因是在家居环境中物理空间的电磁干扰性小,并且采用高可靠性的无线组网协议,降低了设备组网在多障碍情况下的丢包率。

4.1.3 识别准确性测试

识别准确性的测试主要是选取男生和女生两个不同音色的人分别在安静环境和嘈杂环境两种家庭场景中使用设备,分别记录在不同情况下设备指令识别成功的准确率。测试结果如表4所示。

表4 设备识别准确性测试结果

Tab.4 The results of device recognition accurateness

场景	性别	准确率/%
安静	男	91
	女	90
嘈杂	男	85
	女	82

从表4可以看出,在安静环境中,设备对非特定人的语音识别准确率高达90%以上;在嘈杂环境中有所降低,这主要是由于杂音对语音信息的干扰造成。

4.2 移动终端控制软件功能测试

移动终端语音控制软件主要借助讯飞语音云技术来实现,终端设备需要联网,不同的网络状况会对识别结果造成一定的延时和误差。我们在三种不同网络类型下,针对男生和女生两个不同音色的非特定人使用软件进行控制,分别记录在不同网络类型下设备识别的成功和平均延时,测试结果如表5所示。

表5 移动终端控制软件功能测试结果

Tab.5 The results of device recognition accurateness

网络类型	性别	成功率/%	平均延时/s
GPRS	男	52	4.5
	女	49	4.8
3G	男	88	21.0
	女	85	2.4
WiFi	男	95	1.5
	女	92	1.7

从表5可以看出,在三种不同网络类型下,软件识别成功率和平均延时都有所不同,其中在WiFi网络类型下,软件识别成功率最高,平均延时最低;而在GPRS网络类型下,识别成功率最低,平均延时最高,充分说明了移动终端语音控制软件受网络质量影响大。

5 结束语

本文结合实验室已有智能家居平台,通过对非特定人语音识别技术和相关软硬件知识的研究,分别通过讯飞语音云技术和嵌入式非特定人语音识别技术设计了智能家居语音控制系统,用以实现本地与远程多样化的智能家居语音控制。目前,系统已经应用在实验室智能家居平台中,用户可以在室内和室外两种不同情境中通过本系统来实现对智能家居内部可控设备的语音控制。从目前的使用状况来看,系统的性能和功能都相对稳定,但从物联网技术^[11]的长远发展来说仍然需要进一步的测试,再不断地优化和完善智能家居语音控制系统。

参考文献

- [1] 高小平. 中国智能家居的现状与发展趋势[J]. 低压电器,2005(4): 18-21.
- [2] 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[J]. 计算机测量与控制,2012,20(1):180-182.
- [3] 黄涛,胡宾. 基于SPCE061A单片机的非特定人语音识别设计[J]. 微计算机信息,2006(6):19-21.
- [4] William S. 无线通信与网络[M]. 2版. 何军,译. 北京:清华大学出版社,2010:397-403.
- [5] Tompros S, Mouratidis N, Draaijer M, et al. Enabling application of energy saving appliances of the home environment [J]. IEEE Network, 2009, 23(6): 8-16.
- [6] 陈福彬,伍毅,王铁流. 基于STM32的人行通道智能控制与管理系统设计[J]. 自动化与仪表,2012(1):45-48.
- [7] 陈喜春. 基于LD3320语音识别专用芯片实现的语音控制[J]. 电子技术,2011(11):20-21.
- [8] IEEE Std. IEEE 802. 15. 4-2003 Part 15. 4: wireless medium access control (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (LR-WPANs) [EB/OL]. [2012-11-18]. <http://www.ZIGBEE.org>.
- [9] 高戈坤. 科大讯飞. 移动互联网进入语音时代[J]. 通信世界, 2010(42):34.
- [10] 赵威威. 嵌入式语音识别及控制技术在智能家居系统中的应用[D]. 重庆:重庆大学,2007.
- [11] 严萍,张兴敢. 基于物联网技术的智能家居系统[J]. 南京大学学报:自然科学版,2012(1):27-31.

《自动化仪表》中文核心期刊 中国科技核心期刊

邮发代号: 4-304, 2014年定价: 15.00元, 全年价: 180.00元; 国外代号: M 721

欢迎赐稿, 欢迎订阅, 欢迎宝贵建议, 欢迎惠刊各类广告