

基于 AVR 单片机的语音识别系统设计

苏宝林

(绥化学院, 黑龙江 绥化 152061)

摘要:系统以 AVR 单片机为控制核心, 实现对人的语音的识别控制。系统采用的主控芯片为 Atmel 公司的 ATMEGA128, 语音识别功能采用 ICRout 公司的单芯片 LD3320。LD3320 内部集成语音识别算法, 无需外部 FLASH, RAM 资源, 可以很好地完成非特定人的语音识别任务。同时该芯片内部集成了 MP3 播放功能, 支持 MPEG 等格式, 可实现语音提示或 MP3 歌曲的播放功能。由于内部含有 16 位 A/D、D/A 转换器和功放电路, 所以不需要外接功放电路就可以产生清晰的声音。该系统已经预留好各种接口, 具有良好的扩展性。

关键词:AVR 单片机; LD3320; 语音识别; 非特定人语言识别

中图分类号:TN912.2-34

文献标识码:A

文章编号:1004-373X(2012)11-0136-03

Design of phonetic recognition system based on AVR

SU Bao-lin

(Suihua University, Suihua 152061, China)

Abstract: The system with AVR MCU as a control core, realizes the speech recognition control. The system uses the main control chip Atmel ATMEGA128, ICRout single chip LD3320 is adopted for speech recognition, in which the speech recognition algorithm is integrated without external FLASH and RAM resources. It's a good way to complete the speech recognition task of non-specific person. MP3 player is integrated in the chip to support MPEG format. It can be used for voice prompt playback or MP3 songs. Due to internal containing 16 A/D, D/A converter and a power amplifier circuit, it can produce clear sound without any external power amplifier circuit. The system has reserved a variety of interfaces and has a good scalability.

Keywords: AVR MCU; LD3320; phonetic recognition; non-specific person phonetic recognition

0 引言

传统的人机交互依靠复杂的键盘或按钮来实现, 随着科技的发展, 一些新型的人机交互方式也随之诞生, 带给人们全新的体验。基于语音识别的人机交互方式是目前热门的技术之一。但是语音识别功能算法复杂、计算量大, 一般在计算机上实现, 即使是嵌入式方面, 多数方案也需要运算能力强的 ARM 或 DSP, 并且外扩 RAM、FLASH 等资源, 增加了硬件成本, 这些特点无疑限制了语音识别技术的应用, 尤其是嵌入式领域。

本系统采用的主控 MCU 为 Atmel 公司的 ATMEGA128, 语音识别功能则采用 ICRout 公司的单芯片 LD3320。LD3320 内部集成优化过的语音识别算法, 无需外部 FLASH, RAM 资源, 可以很好地完成非特定人的语音识别任务。

1 整体方案设计

1.1 语音识别原理

在计算机系统中, 语音信号本身的不确定性、动态性和连续性是语音识别的难点。主流的语音识别技术是基于统计模式识别的基本理论, 原理如图 1 所示。

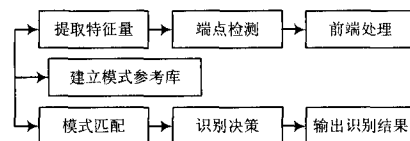


图 1 语音识别过程图

语音识别通常需要两个阶段完成。第一阶段是训练, 主要是提取语音特征, 用户往往需要进行几次语音训练, 经过预处理和特征提取后获得相应特征参数。第二阶段是识别, 识别过程就是将输入的语音特征参数和模型库中的参数进行相似性比较, 最后输出匹配度最高的特征参数完成识别过程。

收稿日期: 2011-11-29

基金项目: 绥化学院 2011 年科学技术研究资助项目

(K1102004)

2 硬件电路设计

硬件框架如图 2 所示,电路主要由主控制器电路和语音识别电路组成。ATMEGA128 控制 LD3320 语音识别电路,输出结果由 ATMEGA128 处理,然后通过总线来控制不同的设备^[1]。

2.1 控制器电路

控制器选用 Atmel 公司生产的 ATMEGA128 芯片,采用先进的 RISC 结构,内置 128 KB FLASH, 4 KB SRAM, 4 KB E²PROM 等丰富资源^[2]。该芯片是业界高性能、低功耗的 8 位微处理器,并在 8 位单片机市场有着广泛应用^[3]。

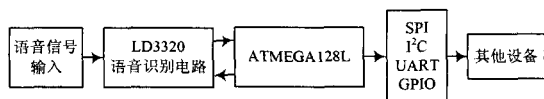


图 2 硬件电路框架图

2.2 LD3320 语音识别电路^[4]

LD3320 芯片是一款“语音识别”专用芯片。该芯片集成了语音识别处理器和一些外部电路,包括 A/D、D/A 转换器、麦克风接口、声音输出接口等,而且可以播放 MP3。不需要外接任何的辅助芯片如 FLASH, RAM 等,直接集成到产品中即可以实现语音识别、声控、人机对话功能^[5]。

图 3 为 LD3320 电路原理图,与 MCU 通信采用 SPI 总线方式,时钟不能超过 1.5 MHz。

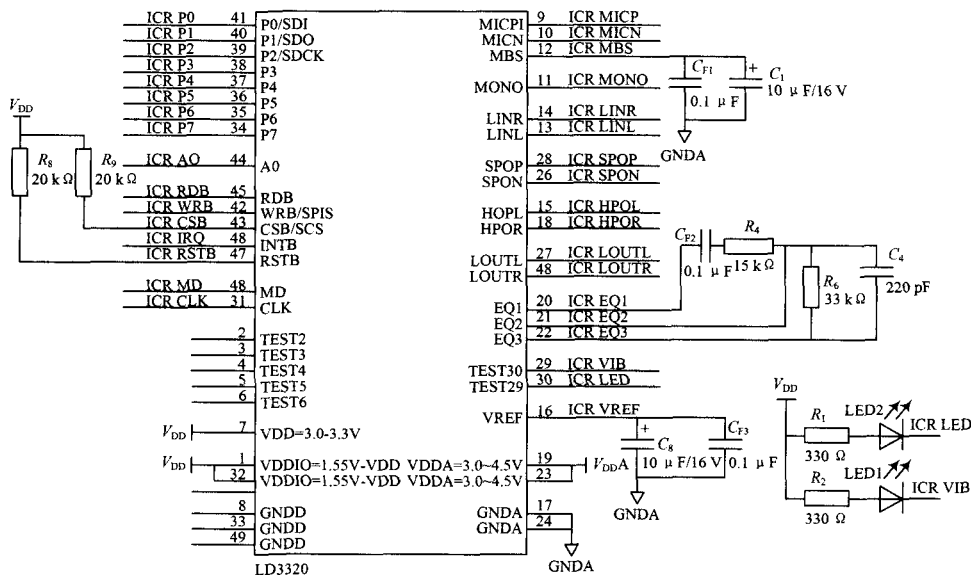


图 3 LD3320 电路原理图

麦克风工作电路如图 4 所示,音频输出只需将扬声器连接到 SPOP 和 SPON 即可。使用 SPI 总线方式时,LD3320 的 MD 要设为高电平,SPIS 设为低电平。SPI 总线的引脚有 SDI, SDO, SDCK 以及 SCS。INTB

为中断端口,当有识别结果或 MP3 数据不足时,会触发中断,通知 MCU 处理。RSTB 引脚是 LD3320 复位端,低电平有效。LED1, LED2 作为上电指示灯。

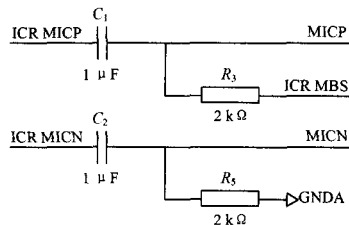


图 4 麦克风电路

3 软件系统设计

软件设计主要有两部分,分别为移植 LD3320 官方代码和编写语音识别应用程序。

3.1 移植 LD3320 源代码^[5]

LD3320 源代码是基于 51 单片机实现的, SPI 部分采用的是软件模拟方式,但在播放 MP3 数据时会有停顿现象,原因是 51 单片机主频较低,导致 SPI 速率很慢,不能及时更新 MP3 数据^[6]。移植到 ATMEGA128 需要修改底层寄存器读写函数、中断函数等。底层驱动在 Reg_RW.c 文件中,首先在 Reg_RW.h 使用 HARD_PARA_PORT 宏定义,以支持硬件 SPI。然后在 Reg_RW.c 文件中找到 HARD_PARA_PORT 对应条件宏的代码段,保留 AVR 的 SPI 接口代码。

3.2 应用程序实现^[5]

在代码中预先设定几个单词:“你好”,“播放音乐”,“打开”。当用户说“播放音乐”时,MCU 控制 LD3320 播放一段音乐,如果是其他词语,则在串口中打印识别结果,然后再次转换到语音识别状态。

3.2.1 MP3 播放代码

LD3320 支持 MP3 数据播放,播放声音的操作顺序为:

通用初始化→MP3 播放用初始化→调节播放音量→开始播放。

将 MP3 数据顺序放入数据寄存器,芯片播放完一定数量的数据时会发出中断请求,在中断函数中连续送入声音数据,直到声音数据结束。MP3 播放函数实现代码如下:

```

void PlaySound()
{
    nMp3StartPos = 0;
    nMp3Size = 12200;           //MP3 文件大小
    LD_Init_MP3();              //MP3 播放初始化
    LD_AdjustMIX2SPVolume(15);  //设置音量
    LD_play();                  //开始播放
}

```

由于 MCU 容量限制,选取测试的 MP3 文件不能太大。首先在计算机上将 MP3 文件的二进制数据转为标准 C 数组格式文件,然后将该文件加入工程中。源代码中 MP3 文件存储在外扩的 SPI FLASH 中,工程中需要注释和移除全部相关代码。MP3 数据读取函数是 LD_ReloadMp3Data(),只需将读取的 SPI FLASH 数据部分改成以数组数据读取的方式即可^[7]。

3.2.2 语音识别程序

LD3320 语音识别芯片完成的操作顺序为:通用初始化→ASR 初始化→添加关键词→开启语音识别。在源代码中的 RunASR()函数已经实现了上面的过程,直接调用即可开启语音识别功能。

RunASR()函数代码如下:

```

uint8 RunASR()
{
    uint8 i=0;
    uint8 asrflag=0;
    for (i=0; i<5; i++)
    {
        LD_AsrStart();
        delay(100);
        if (LD_AsrAddFixed()==0)
        {
            LD_reset();
            delay(100);
            continue;
        }
        asrflag=1;
        break;
    }
    return asrflag;
}

```

用户说完话后,LD3320 通过打分的方式,将关键词列表中特征最相似的一个作为输出。然后 LD3320 会产生一个中断信号,此时 MCU 跳入中断函数读取 C5 寄存器的值,该值即为识别结果,得到结果后,用户可以根据数值来实现一些功能,比如读取到 1,说明是

“播放音乐”,那么可以调用前面的 PlaySound()函数来播放音乐。

语音识别控制的关键点在于语音识别的准确率。表 1 给出了测试结果,当然也可以在识别列表中加入更多的关键词来做测试。通过测试结果可以看出,LD3320 的识别率在 95%上,能够满足用户需求。

表 1 语音识别测试结果

识别词语	你好	打开	播放声音
测试次数	100	100	100
识别率/%	96	98	95

4 结 语

本文讨论了基于 AVR 单片机的语音识别系统设计的可行性,并给出了设计方案。通过多次测试结果表明,本系统具有电路运行稳定,语音识别率高,成本低等优点。同时借助于 LD3320 的 MP3 播放功能,该系统具有一定的交互性和娱乐性。移植性方面,系统通过简单的修改,可以很方便地将 LD3320 驱动程序移植到各种嵌入式系统中。随着人们对人工智能功能的需求,语音识别技术将越来越受到人们的关注,相信不久的将来,语音识别将会拥有更广阔的应用。

参 考 文 献

- [1] 郭天祥. 新概念 51 单片机 C 语言教程[M]. 北京:电子工业出版社,2010.
- [2] 马潮. AVR 单片机嵌入式系统原理与应用实践[M]. 北京:北京航空航天大学出版社,2007.
- [3] Atmel. ATMEGA128 数据手册[EB/OL]. [2008-05-12]. <http://www.atmel.com>.
- [4] 苏鹏,周凤余,陈磊. 基于 STM32 的嵌入式语音识别模块设计[J]. 单片机与嵌入式系统应用,2011(2):42-45.
- [5] ICRout. LD332X 数据手册[EB/OL]. [2010-03-10]. http://www.icroute.com/web_cn/Download.html.
- [6] 张毅刚. 新编 MCS-51 单片机应用设计[M]. 哈尔滨:哈尔滨工业大学出版社,2003.
- [7] 沈文,詹卫前. AVR 单片机 C 语言开发入门指导[M]. 北京:清华大学出版社,2003.

作者简介:苏宝林 男,1977 年出生,天津人,硕士,讲师。主要研究方向为单片机技术、EDA 技术。