

文章编号:1671 - 251X(2010)04 - 0123 - 05

基于语音处理技术的人机交互系统的设计

杨 帅^{1,2}, 薛 岚^{1,2}, 吴洪兵^{1,2}, 陈俊生¹

(1. 淮安信息职业技术学院电气工程系,
2. 江苏电子产品装备制造工程技术研究开发中心,江苏 淮安 223003)

摘要:提出了一种采用 SPCE061A 单片机构建基于语音处理技术的人机交互系统的设计方案,给出了系统结构,介绍了系统音频输入电路、音频输出电路的设计,详细介绍了语音播放及语音识别功能的软件实现方案。实验测试结果表明,该系统性能良好,语音识别正确率较高,但易受训练和识别环境的影响。

关键词:人机交互; 语音处理; 语音播放; 语音识别; SPCE061A

中图分类号:TP273 **文献标识码:**B

Design of Man-machine Interaction System Based on Voice Processing Technology

YANG Shuai^{1,2}, XUE Lan^{1,2}, WU Hong-bing^{1,2}, CHEN Jun-sheng¹

(1. Dept. of Electrical Engineering of Huaian College of Information and Technology, Huaian 223003, China. 2. Jiangsu Engineering Technical Research and Development Center for Equipment Manufacturing of Electronic Products, Huaian 223003, China)

Abstract: A design scheme of man-machine interaction system based on voice processing technology was proposed which was designed with SPCE061A single-chip microcomputer. The system's structure was given and circuits of audio input and output were introduced, especially the scheme of software realization of functions of voice playing and voice recognition. The experiment result showed that the system has good performance and high correct rate of voice recognition, while it is easy to be influenced by environments of training and recognition.

Key words: man-machine interaction, voice processing, voice playing, voice recognition, SPCE061A

0 引言

近些年来人机交互系统层出不穷,但大多数人机交互系统采用鼠标、键盘方式,而采用语音处理技术克服了人机交互必须直接接触的局限,应用在人机之间具有障碍的场合具有重大的现实意义。语音处理技术是一门新兴的技术,它不仅包括语音的录制和播放,还涉及语音的压缩编码和解码、语音的识别等各种处理技术。传统实现语音控制的途径有2种:一种是单片机扩展设计,另一种是借助于专门的语音处理芯片。普通的单片机往往不能实现这么复杂的过程和算法,即使勉强实现也要加很多的外

围器件。专门的语音处理芯片功能比较单一且价格较高,除语音之外的其它功能几乎无法实现。本文采用独具语音特色的 SPCE061A 单片机设计了一种基于语音处理技术的人机交互系统,该系统外围电路少,成本低,语音识别正确率较高。

1 系统结构

基于语音处理技术的人机交互系统结构如图1所示。

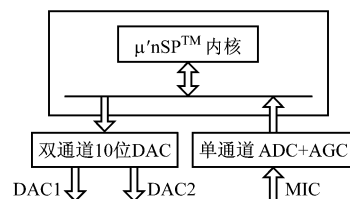


图1 基于语音处理技术的人机交互系统结构
系统中央处理器选用凌阳公司生产的

收稿日期:2009 - 12 - 03

基金项目:江苏省2008年高等学校实践创新计划项目(325109)

作者简介:杨 帅(1981 -),男,辽宁锦州人,硕士,主要研究方向为现场总线与嵌入式系统。E-mail:lnyangshuai@163.com

SPCE061A 单片机, 它采用 μ nSPTM (Microcontroller and Signal Processor) 核心处理器, 具有较高的处理速度, 能够完成 16 位算术逻辑运算、16 \times 16 位硬件乘法运算和 DSP 内积滤波运算, 能够快速处理复杂的数字信号^[1~2]; 集成了 8 路 10 位精度 ADC, 其中 1 路为音频转换通道, 能够轻松地将语音信号采集到芯片内部, 并且内置自动增益环节 (AGC), 随时跟踪、监视前置放大器输出的音频信号电平, 以便使采集的音频信号保持在最佳电平, 又可使谐波减至最小, 从而保证对语音信号的较好控制。

系统的工作过程主要包括语音采集和语音识别 2 个部分, 其原理分别如图 2、图 3 所示。

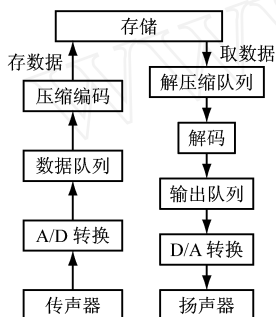


图 2 语音采集工作原理

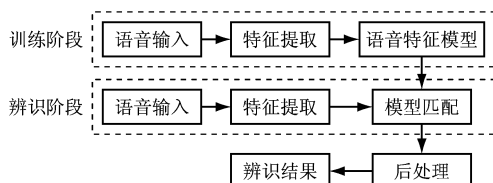


图 3 语音识别原理

2 系统硬件设计

2.1 音频输入电路

由于 SPCE061A 集成了音频输入专用 ADC 以及 AGC 放大电路, 因此, 其外部电路比较简单, 如图 4 所示。其中 VMIC 用于提供传声器的电源, V_{SS} 为系统的模拟地, VCM 为参考电压, MICP 脚和 MICN 脚分别为传声器 X1 的正、负极的输入引脚。当对着传声器讲话时, MICP 脚和 MICN 脚将随着传声器输入的声音产生变化的波形, 并在 SPCE061A 的 2 个端口处形成 2 路反相波形, 送到 SPCE061A 内部的运算放大器进行音频放大, 经过放大的音频信号通过 ADC 转化为数字量, 保存到相应的寄存器中^[3~4], 然后对数字音频信号进行压缩、识别、播放等处理。

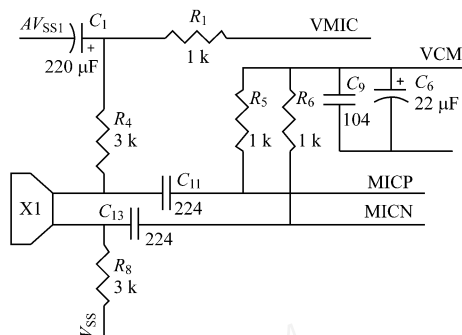


图 4 音频输入电路

2.2 音频输出电路

SPCE061A 内置 2 路 10 位的 DAC, 只需外接功放电路即可实现语音播放功能, 如图 5 所示。其中 VDDH 为参考电压, V_{SS} 为系统模拟地。音频信号由 SPCE061A 的 DAC 引脚输出送入到音频输出电路的 DAC 端, 通过音量电位器 R_9 的调节端输入到集成音频功率放大器 SPY0030, 经音频放大后, 音频信号从 SPY0030 输出端外接扬声器播放声音。SPY0030 是凌阳公司生产的一款音频放大芯片, 工作范围为 2.4 ~ 6 V, 最大输出功率可达 700 mW。

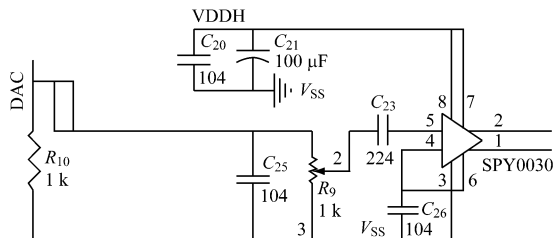


图 5 音频输出电路

3 系统软件设计

基于语音控制的人机交互系统主要包括语音播放和语音识别功能。语音播放功能采用凌阳语音播放算法 SACM_480 实现, 通过调用 Sacmv26e.lib 语音播放库中的函数实现语音信号的模数转换、压缩编码、解码、填充队列、数模转化等功能。语音识别分为语音训练和语音识别 2 个过程, 通过调用 bsrv222SDL.lib 语音识别库中的函数实现特征模型导入和导出等。

3.1 软件总体设计

系统软件的总体设计流程如图 6 所示。语音命令的特征模型保存在 SPCE061A 的内部 RAM 中。由于 SPCE061A 的存储资源有限, 系统所需 RAM 空间可能被旧的特征模型数据占满, 新的特征模型则无法保存到 RAM 中, 所以首先利用 BSR_DeleteSDGroup(0) 函数将 RAM 空间中所有的特征模型删除, 释放出所需空间。然后检测训练标志位

BS_Flag, 当 BS_Flag = 0xFFFF 时, 表明系统未经训练, 需要调用 BSR_Train (int CommandID, int TraindMode) 训练函数来完成语音模型的建立, 其中 CommandID 为命令序号, 范围为 0x100 ~ 0x104, 并且对于每组训练语句都是唯一的; TraindMode 为训练次数, 为 1 表示使用者训练 1 次, 为 2 表示训练者训练 2 次, 为了增强可靠性, 最好训练 2 次, 否则识别的命令就会倾向于噪音。BS_Flag 为其它值表明系统已经训练, 即可装载语音模型, 进入识别状态,

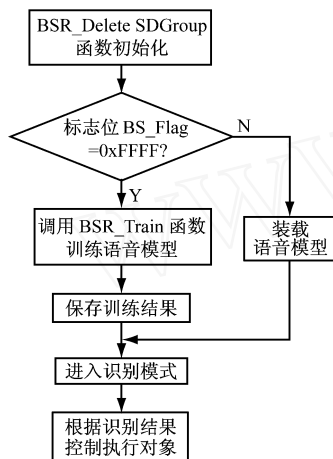


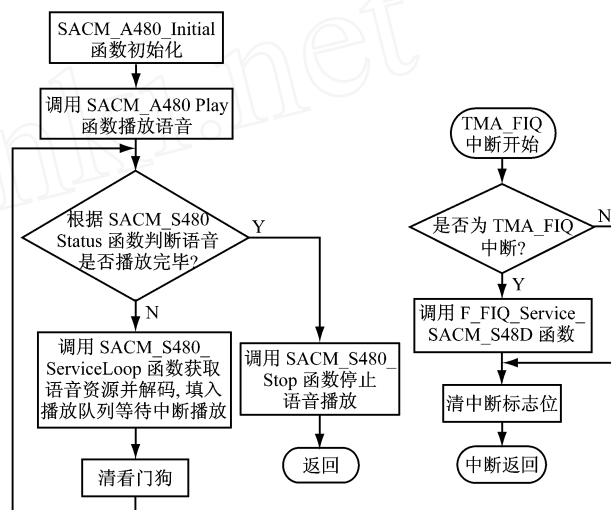
图6 系统软件的总体设计流程

3.2 语音播放程序设计

系统实现人机交互的条件之一就是语音播放。系统在训练和识别过程以及控制执行对象过程中要播放提示信息, 使操作者更好地与系统交互。实现语音播放首先要制定语音播放资源, 可以采用 Windows 操作系统自带的录音机录制语音. wav 文件资源, 将. wav 文件的属性设置成“PCM/ 8 kHz/ 16 位/ 单声道/ 15 KB/s”格式, 否则 SPCE061A 无法播放语音资源。由于 SPCE061A 内部存储资源有限, 所以还要将. wav 文件通过凌阳压缩工具 Compress Tool 将其压缩成. 48k 文件。最后将. 48k 文件添加到所建工程的资源列表中。

语音播放程序流程如图 7 所示。通过 SACM_A480_Initial(int Init_index) 函数选定播放模式, Init_index = 0 表示手动方式, Init_index = 1 表示自动方式, 这里选择自动方式。然后调用 SACM_A480_Play(int Speech_Index, int Channel, int Ramp_Set) 函数设定语音资源序号、声道、音量是否可以调节等, Speech_Index 的值按照. 48k 文件在资源列表中的位置 N 设定, 即 Speech_Index = N - 1; 选择 DAC1 和 DAC2 双通道播放, 设定 Channel = 3; 允许音量增减, 设定 Ramp_Set = 3; 通过函数

SACM_S480_Status(void) 的返回值是否为 1 判断语音是否播放完毕, 如果没有播放完毕则调用语音播放系统服务函数 SACM_S480_ServiceLoop(void) 将语音资源送入解码队列、解压缩后送入播放队列等待中断播放, 如果播放完毕则调用 SACM_S480_Stop(void) 函数停止语音播放^[5]。语音播放中断必须放在 TMA_FIQ 中断源上且通过调用函数 F_FIQ_Service_SACM_S480 执行。



(a) 主程序流程

(b) TMA_FIQ 中断流程

图7 语音播放程序流程

3.3 语音识别程序设计

系统实现人机交互的条件之二就是语音识别, 即系统能够识别人的口令并按照口令执行动作。SPCE061A 配有专用的麦克接口用于语音训练和语音识别, 16 位的定时/计数器用于语音信号的控制采样, 内置的硬件乘法器和内积运算保证了识别算法的运行。

训练模式启动后, 系统播放语音提示, 提示用户语音训练已启动, 接下来用户可按照系统提示依次对各条命令进行训练。在训练过程中, 如训练成功则由语音提示进行下一条命令进行训练, 若失败, 也会提示用户继续训练该条命令, 全部命令成功训练完毕后系统将准备进行语音识别。当向控制器发出语音命令时, 声波通过麦克接口输入, 将相应的信号传递到 SPCE061A, 经编解码电路和数字信号处理后, 通过相关程序与预先植入的语音库中的命令进行比较识别, 根据识别的结果进行判断, 转换为能被系统识别的信号, 从而对被监控系统进行控制^[6~7]。

(1) 语音训练

语音训练程序流程如图 8 所示。语音训练通过调用函数 int BSR_Train 来完成, 训练采用 2 次训练

获取结果的方式,训练成功返回 0,没有声音返回 - 1,训练需要再说一遍返回 - 2,训练环境太吵返回 - 3,数据库满返回 - 4,2 次训练的命令不同返回 - 5。系统首先会提示“第一条命令”,这时训练者可以告诉系统第一条命令,然后系统会提示“请再说一遍”,这时重说一次命令。如果 2 次的声音音色和响度基本相同,系统就能够成功地建立语音模型,此时第一条口令训练成功;如果没能成功建立模型,则提示“声音不同”。另外若训练口令时周围环境太吵,系统提示“环境太吵”;若训练口令声音太小,则提示“没有声音”。口令训练结束后,系统提示“准备就绪”。

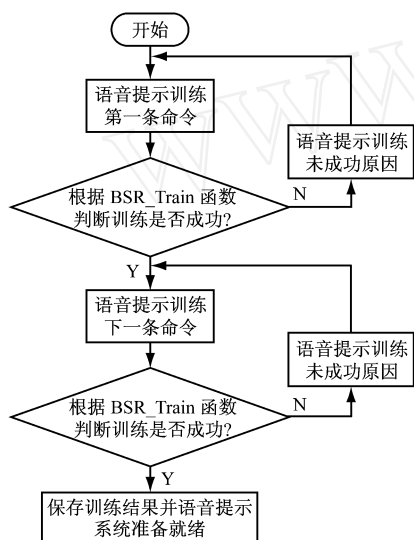


图 8 语音训练程序流程

(2) 语音识别

语音识别程序流程如图 9 所示。通过调用 void BSR_InitRecognizer(int AudioSource) 函数完成初始化,其中 AudioSource 为 0 表示 MIC 语音输入,为 1 表示 LINE_IN 模拟电压输入。当主程序调用该函数时,语音识别器便打开 8 kHz 采样频率的 FIQ_TMA 中断,并将采样得到的语音数据填入语音识别数据队列中。函数 int BSR_GetResult() 完成语音识别处理,当无命令识别出来时,函数返回 0;识别器未初始化或识别未激活时返回 - 1;识别不合格时返回 - 2;识别成功时返回命令的序号 ID,根据 ID 号执行对应的动作。

4 测试与结论

测试方法:系统分 10 次测试,每次测试 5 种口令识别,每次测试正确识别口令的数据如表 1 所示。

从表 1 可看出,系统的口令识别正确率在 90% 以上。由于语音信号具有不确定性,因此,语音识别

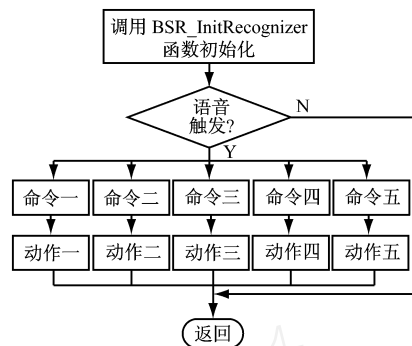


图 9 语音识别程序流程

表 1 系统每次测试正确识别口令的数据

	测试 1	测试 2	测试 3	测试 4	测试 5
正确识别 口令数	4	5	5	5	4
	测试 6	测试 7	测试 8	测试 9	测试 10
正确识别 口令数	5	4	5	5	4

过程中会出现一定的误差和不确定性。经分析得出以下结论:

(1) 系统口令识别的正确率与训练环境紧密相关,在训练过程中应尽量保持周围环境安静,否则可能导致识别正确率下降;

(2) 系统口令识别的正确率与人和麦克的距离有关,在条件允许下的情况下应尽量保证人与麦克的距离最小;

(3) 使用 8 kHz 的采样频率即可获得比较好的效果,且占用存储空间适中,因此,语音播放速率要与语音采样速率相同,即 8 kHz,这样才能保证系统正常播放语音。

另外在系统调试过程中发现,当程序首次编译时,会提示“ The external symbol ‘ T_SACM_A2000_SpeechTable ’ has not a public definition ”信息,解决方法就是在工程中打开 Resource.asm 文件,在文件结尾处加入语音资源索引表。如果编译出现“ cannot locate DATA section automatically ”信息,表明被播放的语音资源较多,需要扩展存储器。

5 结语

本文介绍的基于语音处理技术的人机交互系统利用声音实现了人机交互功能,克服了鼠标、键盘受地域限制的缺点。如果采用无线麦克,则使用更加简便。该系统成本低,外围电路少,易操作,语音识别率较高,实际应用前景广阔。

文章编号:1671-251X(2010)04-0127-03

西门子 PLC 在索道动态监控系统中的应用

马秋环, 徐文尚, 滕景忠, 段 锋

(山东科技大学信电学院, 山东 青岛 266510)

摘要:介绍了一种基于 PLC 的索道动态监控系统的硬件组成及结构设计,详细阐述了系统功能实现的原理以及系统程序的具体实现。该系统结构简单,实时性能好,且具有 Internet 远程监控功能。

关键词:索道;吊厢;视频监控;PLC;条码扫描器

中图分类号:TP273

文献标识码:B

0 引言

索道动态监控系统主要是对各条索道的吊厢进行实时监控,利用硬件设备采集扫描枪、抓拍、速度、出站等信号,动态显示当前索道吊厢的吊厢号、吊厢的抓拍图片、吊厢距离地面的高度以及吊厢与上下站的距离、吊厢抱索器的开合次数、运载索道总的运行距离等。

本文介绍的索道动态监控系统采用西门子 S7-200 系列 PLC(CPU226)作为控制器,通过与采用 VC 开发的上位机的有机结合,实现了对索道运行情况的实时监控。

1 系统硬件组成及结构

1.1 系统硬件组成

根据监控的具体要求,本系统主要由西门子 S7-200 PLC CPU226(以下简称 CPU226)、条码扫描

器、速度传感器、2 个接近开关和 1 个高精度摄像头组成。

选用西门子 S7-226 作为控制系统核心,是因为它除了能够满足监控系统要求外,还具有成本低、稳定性能好、抗干扰能力强、程序编写调试方便等特点。由于工作环境恶劣,笔者采用 3800g 型工业条码扫描器,它扫描范围大,分辨率高,扫描速度快,稳定性好。速度传感器型号为 LM18-3005NAYT,用来检测索道同步轮的转速。接近开关选用 PRL18-5DN 型,将其作为 PLC 的数字量输入。高精度摄像头用来抓拍吊厢的图片。

1.2 系统结构

基于西门子 PLC 的索道动态监控系统结构如图 1 所示。

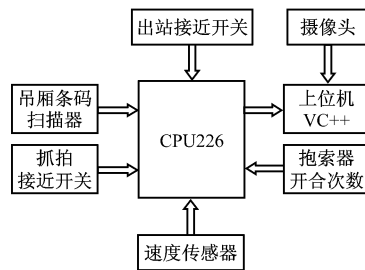


图 1 基于西门子 PLC 的索道动态监控系统结构

收稿日期:2009-12-24

作者简介:马秋环(1983-),女,山东菏泽人,山东科技大学信电学院检测技术与自动化装置专业在读硕士研究生,研究方向为自动控制系统。E-mail:maqiuahuan@126.com

参考文献:

- [1] 罗亚非. 凌阳 16 位单片机应用基础[M]. 北京:北京航空航天大学出版社,2006.
- [2] 杨 帅,薛 岚,史宜巧,等. 基于音频信息采集的 LonWorks 节点的研究[J]. 计算机技术与发展, 2009(8):205-207.
- [3] 汪 敏. 凌阳 16 位单片机实验与实践[M]. 北京:北京航空航天大学出版社,2007.
- [4] 杨 帅,薛 岚,高安邦,等. 基于 SPCE061A 的智能小车机器人的设计[J]. 机械制造与自动化,2009, 38(6):161-163.
- [5] 李小白. 凌阳 16 位单片机 C 语言开发[M]. 北京:北京航空航天大学出版社,2006.
- [6] 杨 帅,薛 岚,高安邦,等. 基于 SPCE061A 智能小车机器人语音辨识系统的设计[J]. 山西电子技术, 2009(4):5-6,19.
- [7] 杨 帅,薛 岚,高安邦,等. 基于 SPCE061A 智能小车机器人语音播放系统的研究[J]. 装备制造技术, 2009(7):38-39.