

A Simple System Design of Gesture Recognition to Speech Expression

Xiaohui He, Xinliang Cao*, Na Zhao, Shouyu Luo

School of Physics and Electricity Information, Yan'an University, Yan'an Shaanxi
Email: *caoxinliang874@163.com

Received: Nov. 1st, 2019; accepted: Nov. 19th, 2019; published: Nov. 26th, 2019

Abstract

This study aims to convert the gestures of deaf and dumb into “dubbing” and to help the deaf-mute people communicate with ordinary people. It would complete the simple sign language recognition system based on MSP430F5529 microcontroller control sensor chip—FDC2214. The system uses a capacitive sensing chip to collect and process signals with gesture gestures changes, and then transmits these signals to the microcontroller through a SPI. It can be output into two modes. On the one hand, the change of the gesture is output in the form of text through the display screen, so that the person who does not understand the sign language understands the language of the deaf person; on the other hand, it can process the change of the frequency by the single chip microcomputer. It is found the corresponding relationship between the gesture and the pre-stored language segment in the voice chip, and the corresponding voice announcement of the changed gesture is controlled by the instruction. The system is designed to recognize the correct output of ten gestures and daily languages.

Keywords

Sensor, Gesture, Recognition, Speech

一种简单手语识别及语音表达系统设计

何晓慧, 曹新亮*, 赵娜, 骆守宇

延安大学物电学院, 陕西 延安
Email: *caoxinliang874@163.com

收稿日期: 2019年11月1日; 录用日期: 2019年11月19日; 发布日期: 2019年11月26日

*通讯作者。

摘要

本研究旨在为聋哑人的手语释义“配音”，以帮助聋哑人和普通人进行交流，设计基于MSP430F5529单片机控制传感芯片FDC2214的简单手语识别系统。该系统采用电容式传感芯片来采集并处理手势变化的信号，再将这些信号通过串口线传送给单片机处理。它可以分两种模式输出，一方面将手势的变化情况通过显示屏以文字的形式输出，以便不懂手语的人理解聋哑人的语言；另一方面它可以通过单片机对频率的变化进行处理，找出手势与语音芯片中事先存储语言片段的对应关系，通过指令控制对变化的手势做出相应的语音播报。所设计的系统能够识别十种手势和日常用语的正确输出。

关键词

传感器，手势，识别，语音

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

我国聋哑症的发病率约为 2%！按年均人口出生率计算，连同出生后 2 至 3 岁婴幼儿，每年总的群体达 5700 万，听损伤的发病人数约为 17 万。这一群体因听力障碍、缺乏对外界事物的全面感知，导致大部分人语言表达的缺失。为了解决聋哑人与普通人思想交流的壁垒问题，手语的语音翻译技术应用而生。

目前，现有技术主要有以下几类[1]，第一，由美国大学生研发的“Sign Aloud”智能手语翻译手套，当中含有内置传感器[2]，可以记录佩戴者的手部位置和动作。第二，是由电子科技大学的学生以数字图像处理技术为出发点，通过对手势图像的识别，来实现对手语的翻译，但这种方法受环境影响因素较大，且后期对图像的处理，相对复杂；并且此技术在图像生成与传递的过程中会使视频图像变得模糊，对于后期处理的难度增大；第三，是 MotionSavvy 的公司推出的一款 UNI 平板，该平板内置一个 Leap Motion 动作传感器[3] [4]，可通过它将聋哑人手势识别处理并转化为文字，然后通过语音合成[5]朗读出来。但在日常生活的使用过程中，很不方便，且每月需支付 20 美元的订阅费用于软件更新服务，这些价钱对于本就生活不便的聋哑人人群来说，是一笔不小的开支。研发一种价格便宜、能被大部分聋哑人人群经济条件所能接受的便携式聋哑人简单手语识别系统是现实需要。

该项研究对打破聋哑人与正常人之间的沟通壁垒、提高弱势群体的生活和工作质量具有重要的意义和实用价值。

2. 系统组成

2.1. 系统硬件组成

本设计是基于 MSP430F5529 单片机控制[6]传感芯片 FDC2214 研发简单手语识别装置，利用电容式传感芯片控制的感应板来收集手势的变化情况，再通过串口线将采集到的信息传给单片机进行处理[7]，分两种模式输出：一部分通过显示屏以文字的形式输出；另一部分，通过 ISD2560 语音模块，将手语要表达的简要意思“说出来”。如图 1 所示。

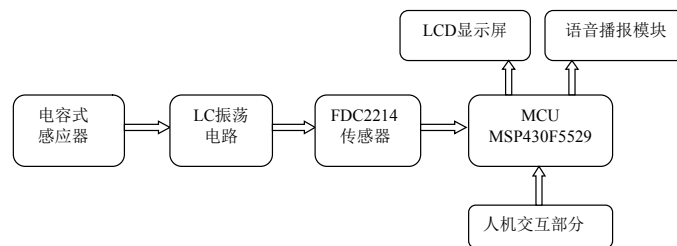


Figure 1. The system hardware block diagram
图 1. 系统硬件组成框图

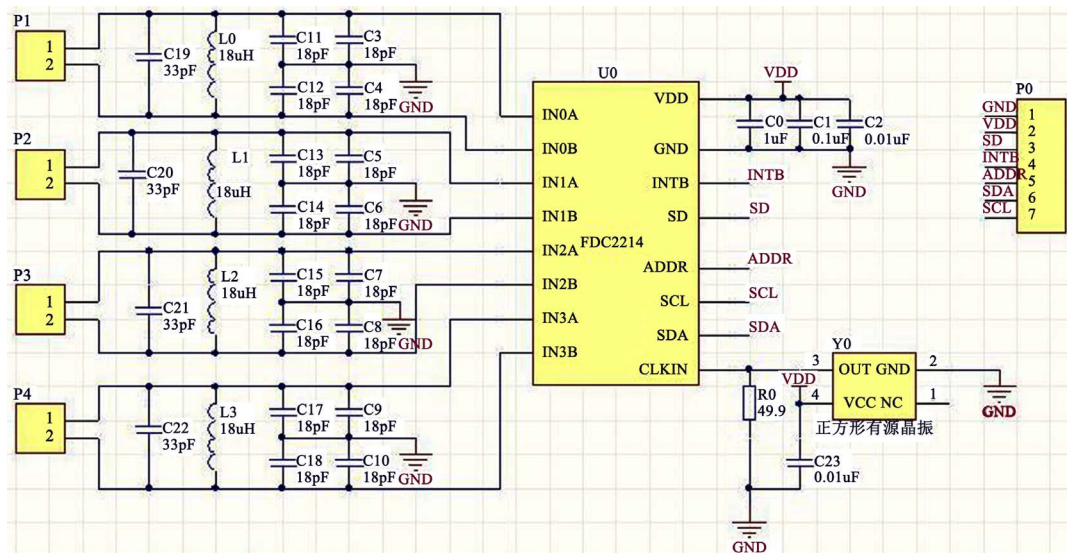


Figure 2. Circuit diagram of the capacitive sensing subsystem
图 2. 电容传感子系统电路构成图

2.1.1. 系统感测部分设计

图 1 中, 被测电容传感端即“FDC2214 的传感平面”为导体材质, 当人手接近该导体传感平面时, 传感端的电容发生了变化, 这就会导致 LC 电路振荡频率的变化, 从而反映手势状态的判定。

图 2 为传感芯片 FDC2214 的外围电路图, P1、P2、P3、P4 分别感应手势的变化情况, 当人手接近该导体传感平面时, 传感端的电容发生了变化, 这就会导致 LC 电路振荡频率的变化, 以频率变化量的大小来反映手势的变化情况, 从而做出对手势变化的判定。

FDC2214 是基于 LC 谐振电路原理的一个电容检测传感器。其基本原理如图 3 所示, FDC2214 由前端谐振电路驱动器组成, 后面是一个多路复用器, 通过序列进行排序有源通道, 将他们连接到测量和数字化传感器频率的核心。在多通道模式下操作时, FDC2214 顺序采用活动通道; 在单通道模式下, FDC2214 则采用单个通道。在芯片每个检测通道的输入端连接一个电感和电容, 组成 LC 电路, 被测电容传感端与 LC 电路相连接, 将产生一个振荡频率, 根据该频率值可计算出被测电容值。

2.1.2. 系统语音播报部分设计

ISD2560 是自带电可擦除只读存储器 (EEPROM), 可以将不同电平的模拟量通过采样直接把采样结果存储在这个只读存储器 (ROM) 的单元中。ISD2560 舍掉了一般数字存储芯片的模数转换 (ADC) 和数模转换 (DAC) 部分, 典型采样频率为 8 kHz。因此, ISD2560 避免了一般固体录音电路因量化和压缩造成的附加噪声, 而使得这种固态录音技术还原声音更加清晰、逼真。

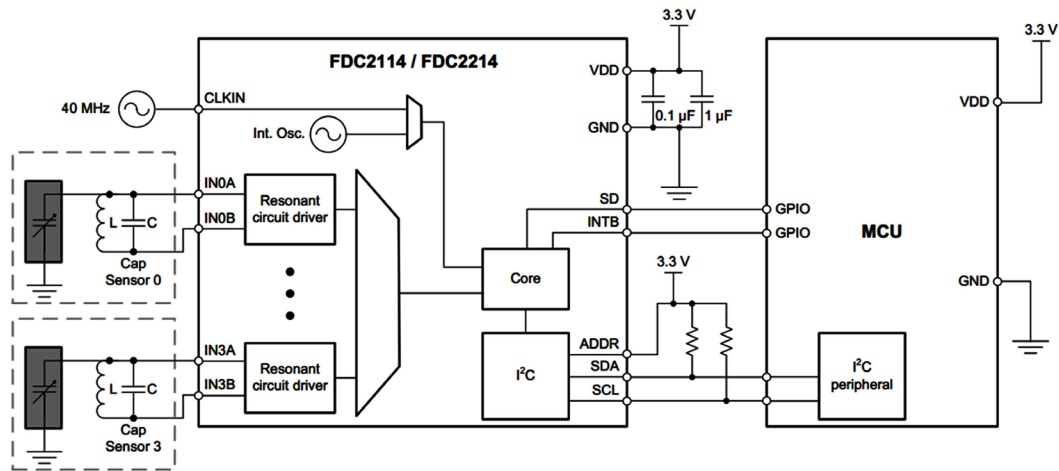


Figure 3. FDC2214 structure and its typical circuit diagram
图 3. 传感芯片 FDC2214 结构与典型电路图

一般地，固体录放电路的存储容量是一定的。当采样频率较低，存储容量满载时，对应录音/放音时间就长；当采样频率较高、存储容量满载时，对应录音/放音时间就变短。但由于同频带内采样频率与采样点的密集程度相关，采样频率越高音质质量越好、采样频率降低音质就会变差。ISD2560 允许录放次数很多，达 10 万次以上，它因具有音质逼真，便携低功耗，可多次擦除、重复录存等特点而得到广泛应用。

ISD2560 内部结构可分为：放大、采样存储、滤波、逻辑控制、存储器、语音输出等 6 个部分。

其中，放大器部分由前置放大器、AGC 和模拟放大器构成；采样存储部分由内部振荡器和分频定时器构成；滤波部分是由防混叠滤波器和平滑滤波器构成；逻辑控制部分由逻辑控制线、地址缓冲器组成；语音输出部分是由混合器和驱动器组成；存储器部分由模拟收发器、地址解码器和存储阵列构成，存储阵列是 480 k 字节的 EEPROM。ISD2560 有 600 个地址对应着内部存储单元被均匀划分的 600 行。此外，ISD2560 配设有微控制器输入接口。通过操纵地址和控制线来实现复杂的信息处理功能，能以最小段长为单位灵活组合分段。ISD2560 内部原理图如图 4 所示。

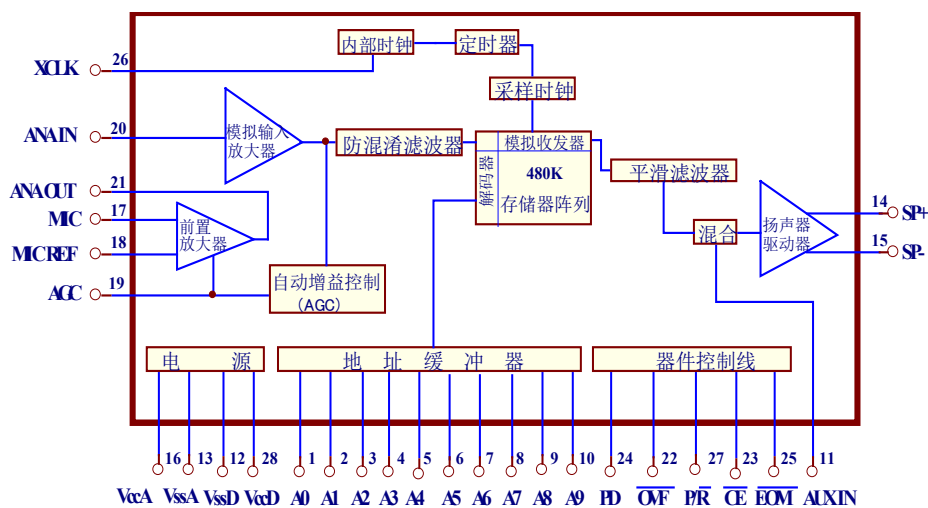


Figure 4. ISD2560 internal schematic
图 4. ISD2560 内部原理图

由图 4 可知 ISD2560 内部集成了高精度的振荡电路, 无需外部配置晶振, 具有录、放双向功能。当录音时, 语音经 MIC 拾音变化成电信号, 此音频微弱电信号经电容隔除直流分量后输入给前置放大器放大后由 ANAOUT 脚输出, 再进入自动增益 AGC 放大器, 然后通过信号整体电平的搬移, 使之适应于存储电路的动态范围。为使电路实现对信号的不失真采样, 放大后的信号还需一次抗混叠滤波, 才能送入存储单元, 该抗混叠滤波器为一五极点抗干扰滤波器, 高频频限为 3.4 K, 于是该器件典型带宽为 3.4 K。调理完毕的信号在内部时钟的作用下以闭环控制形式送入模拟存储阵列。如图 5 所示:

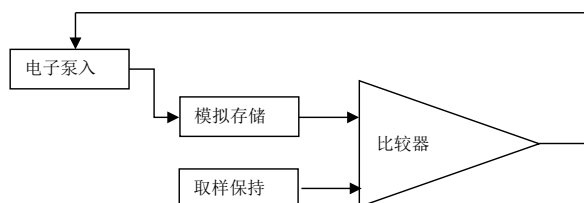


Figure 5. Closed-loop memory circuit
图 5. 闭环存储电路

录音模式下, 音频调理信号经采样保持电路一次取样保持瞬间电平, 同时电子被泵入模拟存储单元。以取样保持的瞬时电平将作为基准与泵入模拟存储单元的电平比较, 当两者电平相等时则停止向 EPROM 中写入数据(即泵入电子), 这样模拟信息在比较中得到了存储, 继续完成下一次采样所得基准的比较存储, 储存结果类似音量指示标志。

在放音模式下, 通过取样脉冲的作用, 顺序地从模拟矩阵中读出录入的模拟电压并恢复为原始波形, 经五极点平滑滤波器后入混合器, 以便与外不输入信号混合, 而后通过功放驱动送扬声器发声, 恢复出原录入声音。

2.2. 系统工作流程

图 6 是系统工作流程图。此装置通过 FDC2214 的传感平面来感应聋哑人手势变化, 手势变化会引起电容量的变化, 利用电容式传感芯片感测并调理这种变化的电容值; 再将电容式传感芯片处理后的信号传送给单片机设置频率变化。最终根据变化信息将手势变化的动作在屏幕上以文字的形式显示出来, 通过对变化的频率界定手势并进行定义, 然后给语音芯片指令使之对手势做出相对应表达意思的语音播报, 以实现手势变化 - 图文信息 - 语音播报三者的转变。

聋哑人简单手语识别系统的具体实施步骤如下:

- 1) 使用者在感应板上进行简单的手语演示, 连接在感应板上的电容式传感芯片会对手势变化引起的电容变化信号进行收集并处理;
- 2) 将电容式传感芯片所收集到的信息传送给单片机处理, 设置频率变化的阈值, 并进行数据定义;
- 3) 根据变化的阈值信息将手势变化的动作在屏幕上以文字的形式显示出来, 通过对变化的频率进行数据定义, 然后给语音芯片指令让其对变化的手势做出相对应的语音播报。

3. 系统测试结果

将(如图 7 所示)10 种手势分别置于感应极板上, 通过对传感数据进行采集, 发现可以甄别这些不同手势。

再将事先录入(如表 1 中列写)的语言表达内容, 通过软件把特定手势与语言存放地址对应起来, 由 MCU 检测手势的频率, 并寻找语音芯片存储对应表达内容, 控制输出相应的语音, 同时, 根据检测手势在显示屏上输出语言表达的文字形式。测试结果表明: 系统能过按照表 1 设定正确输出结果。

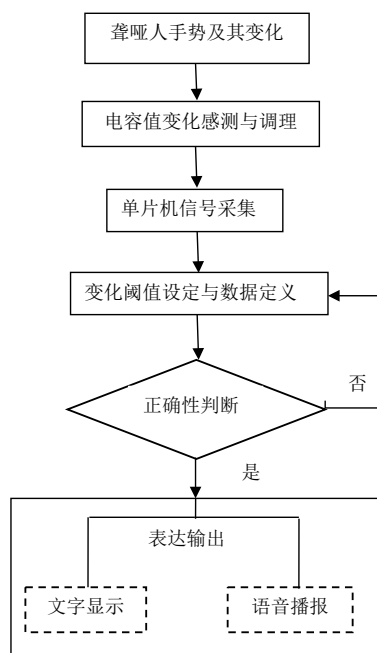


Figure 6. System work flow chart
图 6. 系统工作流程图

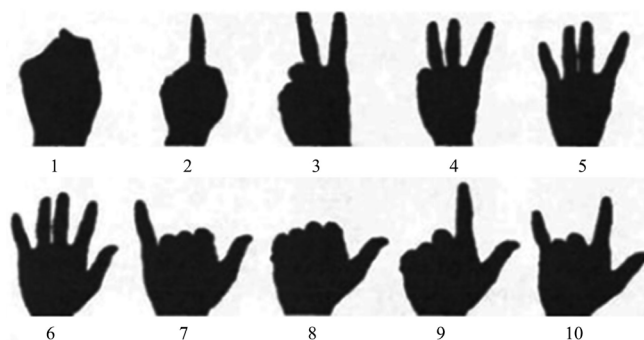
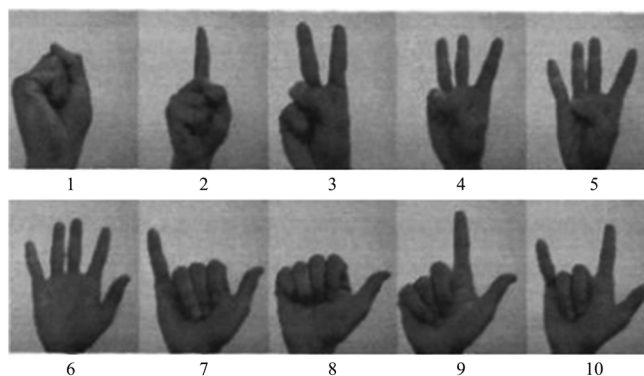


Figure 7. The schematic for gesture
图 7. 手语示意图

Table 1. Gestures and expressions
表 1. 手势与表达内容

手势编号	示意	对应语言存储内容	表达输出结果	
			文字输出	语音输出
0	什么也没有	没什么, 就这样了	正确	正确
1	向我走来	请过来一下	正确	正确
2	分叉	我有不同意见	正确	正确
3	三生有幸	幸会, 见到你很高兴	正确	正确
4	四季发财	祝您好运	正确	正确
5	不遗余力	我会全力以赴	正确	正确
6	大小相差甚远	差的太远了	正确	正确
7	一马当先	太好了	正确	正确
8	瞄准	这就是目的	正确	正确
9	隔三差五	经常联系	正确	正确

4. 总结

系统通过电容式传感芯片检测并处理所接收到的手势变化引起的电容变化值; 利用单片机设置频率变化的阈值与数据定义。通过变化的阈值信息可以将手势变化的动作在屏幕上以文字的形式显示出来, 通过对变化的频率进行数据定义, 然后语音芯片指令可以让其对变化的手势做出相对应的语音播报。实验验证结果说明, 系统实现了简单手势变化转为文字或语音输出。

该系统是以手势投影在电容极板上的面积大小来判断手势类型的, 而且还要根据具体人手的大小进行先验学习, 手势类型数目和适应性受限。改变这两个不足, 寄希望于图像处理的手势识别技术。

基金项目

国家级大学生创新训练计划项目(201910719022); 延安大学研究生教改项目。

参考文献

- [1] 武霞, 张崎, 许艳旭. 手势识别研究发展现状综述[J]. 电子科技, 2013, 26(6): 171-174.
- [2] 韩久强, 周兴鹏. 传感器与检测技术[M]. 北京: 清华大学出版社, 2010: 10-110.
- [3] 关然, 徐向民, 罗雅愉, 等. 基于计算机视觉的手势检测识别技术[J]. 计算机应用与软件, 2013, 30(1): 155-159+164.
- [4] 黄健. 基于电容感测高温隔离触控模块的研制[J]. 宇航计测技术, 2017, 37(5): 40-43.
- [5] 贺国秀, 薛家秀, 赵炜. 基于单片机的语音提醒系统设计开发[J]. 智能计算机与应用, 2015, 5(1): 46-47.
- [6] 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, 2009: 202-264.
- [7] 廖义奎. ARM Cortex-M4 嵌入式实战开发精解——基于 STM32F4 [M]. 北京: 北京航空航天大学出版社, 2013: 27-53.