ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Development of a Speaker Independent Speech Recognition System for Remotely Voice Operated Control Systems

Teetla Asha
M.Tech Student,
Department of ECE,
St. Martin's Engineering College,
Hyderabad, Telangana, India.

S.Ravi Kumar
Associate Professor,
Department of ECE,
St. Martin's Engineering College,
Hyderabad, Telangana, India.

Abstract:

In this present era everything is reduce to minute particles right from home to industrial area. A home automation system combines electrical devices in a house with each other. In this paper, the implementation of an isolated speaker independent speech recognition system for remotely control the electrical devices (loads) are presented. The proposed ISRS is to control different devices remotely by giving voice commands.

The system is tested and found better suited for the physically challenged persons, in order to do the regular activities without having others help. The most common thing that they need to do control the loads without much human involvement wherever they are required by using user defined voice commands to switch either ON or OFF different electrical devices. The proposed system has an application in voice control robots, appliances and access control systems. The project is built around LPC2148 microcontroller; PSoC based Voice Recognition module and RF Tx/Rx (Radio Frequency Transmitter and Receiver).

Keywords: ***RF Tx/Rx, PSoC and LPC 2148 Microcontroller.***

I. INTRODUCTION

The speech recognition system was developed for the first time in early 1950s [1] and [2] and Isolated speech recognition system targeted for voiced operated appliances, robots and access control have been implemented using many hardware platforms such as microprocessors [3,4] and digital signal processors[5] and FPGAs [6,7] and PSoCs [8,9]. Programmable System on Chip (PSoC) is employed in a number of applications. As they have to be cost effective, they have limited storage and computational power. The existing literature covered on speech recognition techniques with different algorithms and complexity.

The motivation of this paper is to propose an isolated speech recognition system for physically challenged (except dumb) persons in order to control remotely the home appliances by using user defined voice commands. The proposed system is remotely control the loads based on voice commands. The distance of the coverage depends on the power output of the driver circuit and the present system is targeting for four loads. The distance can be increased at the cost of the transmitted power. The number of loads to be controlled can be increased based on the application at the cost of complexity.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

The remainder of this paper is organized as follows .The overview of isolated speech recognition system and proposed system model is presented in Section II and Section III respectively. Section IV shows the implementation details and results and Section V summarizes the conclusions.

II. OVERVIEW OF THE ISOLATED SPEECH RECOGNITION SYSTEM

In this section, the basic concept of the isolated speech recognition system is discussed. Most speech-recognition systems are distinguished in to isolated or continuous. An isolated-word system operates on single words at a time - requiring a halt between saying each word.

This is the simplest form of recognition to perform because the end points are simpler to find the pronunciation of a word tends not affect others. Thus, because the occurrences of words are more regular they are easier to recognize. A continuous speech system operates on speech in which words are connected together, i.e. not separated by halts. Continuous speech is more difficult to handle because of a variety of effects. First, it is very difficult to find the start and end points of words. Another problem is coarticulation. The making of each phoneme is affected by the production of surrounding phonemes, and similarly the start and end of commands are affected by the preceding and following commands. The recognition of continuous speech is also affected by the rate of speech (fast speech tends to be harder).

Speech-recognition systems can be further divided as speaker-dependent or speaker-independent. A speaker-dependent system only recognizes speech from single particular speaker’s voice, whereas a speaker-independent system can recognize speech from anybody. The selection of the system is either speaker dependent or speaker independent depending on the application. The proposed system is speaker independent ISR system is targeted for remotely voice operated control systems.

III. SYSTEM MODEL

In this section, the system transmitter and receiver block diagrams are presented at first, and then the hardware details are outlined.

A. Transmitter

As shown in Fig.1, the transmitter block diagram of proposed isolated speech recognition system.

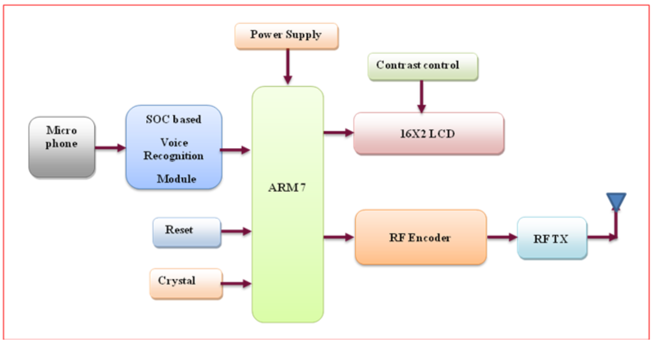


Fig.1. Transmitter Block diagram

Speech recognition or identification system required a "voice reference template" to be created so that it can be compared against subsequent voice identification. To do this, the speaker must speak a set of phrase a number of times while the system builds the template. It consist several parameters in recognizing one's voice pattern, such as pitch, dynamics and waveform. Here, the reference template is created for user defined voice commands by using PSoC based voice recognition module (VRM), which receives the speech signal from microphone is attached to a VRM that recognizes spoken commands. These commands are then converted into analog electrical signals. The analog signals are converted to digital patterns, which are decoded by template matching or feature analysis using Hidden Markov Models. The output data is stored and entered into a program which will be loaded in to the microcontroller. The module can take 5 different voice commands as a one group like that totally 15 different voice commands can be stored. The selection of group is controlled by software or hardware configuration pins.

During Recognition process, the micro- controller receives the data from VRM and compares it with the stored user defined commands. If the received data is

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

matched with stored patterns then the corresponding code will be forwarded to RF module for further processing. Suppose, if the received data is not matching with the stored pattern then it will be written zeros on the output.

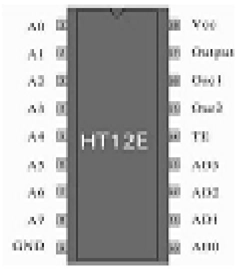


Fig .2. Encoder HT-12E

The data sent from the microcontroller can be sent directly to RF transmitter module but it cannot work exactly at all times in other words there will be no proper synchronization between the transmitter and the micro controller unit. Hence, the transmitter is unable to accept this data as this will be not in the radio frequency range. For this, we need an intermediate device which can accept the input from the microcontroller, process it in the range of radio frequency range and then send to the transmitter. Thus, an encoder is used. The encoder used here is HT12E shown in Fig.2.

Fig. 3. RF Transmitter STT-433

The encoder receives the data from microcontroller and converted it in to the particular frequency format which will be forwarded to STT-433 RF transmitter is as shown in Fig.3. Finally, the RF signal is transmitted through air on 433MHz by using antenna.

B. Receiver

The receiver block diagram of isolated speech recognition system is as shown in Fig.4.

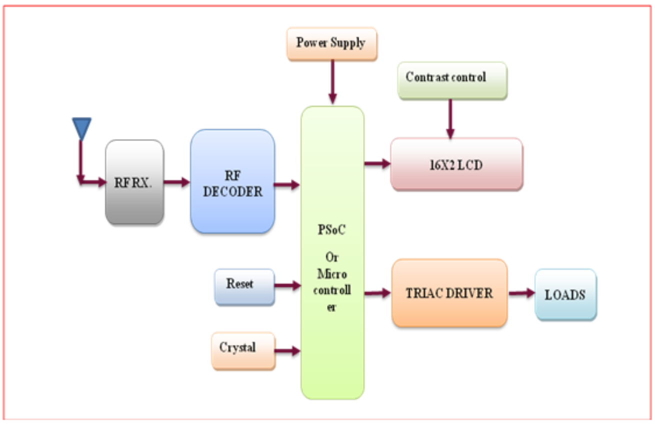


Fig.4. Receiver Block Diagram

The RF receiver STR-433 of the system is as shown in Fig.5. It receives the RF signal from the antenna port on frequency 433MHz. The received signal is converted in to digital format by the RF receiver and forwarded it to RF decoder for further processing. Similarly, like transmitter the receiver module requires a decoder for decodes the received information. The decoder used here is HT12D.

The received data from the data line of the receiver is fed to the decoder. The decoder decodes the received information and is given to PSoC or Microcontroller. The received data is processed according to the applications.

Fig. 5. RF Receiver STR-433

The PSoC or microcontroller reads the digital data and compares it with the preloaded training patterns for execute the pre assigned application. Based on the application the PSoC generates the commands. The configuration of the PSoC is depended on the application and type of the signals required to control the systems. Here, the system is targeted for control the home appliances. So, the appliances or loads are controlled by commands received from the PSoC through TRIAC (BT136). Traic can be triggered by voltage being applied to its gate. Once triggered, the

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

device continues to conduct until current through it drops below a certain threshold limit value, the holding current, such as at the end of a half-cycle of alternating current (AC). This makes the TRIAC a very convenient switch for AC circuits.Finally, the loads will be operated based on the controlled voice inputs.

C. Overview of ARM7 Microcontroller

LPC2141/42/44/46/48microcontrollers consists of 16-bit/32-bit ARM7TDMI-CPU with real-time emulation and embedded trace support, that combine controller with embedded high speed flash memory ranging from 32kB to 512kB. A 128-bit wide memory interface and distinctive accelerator architecture enable 32-bit code execution at the maximum clock rate. For high critical code size applications, the alternative 16bit thumb mode reduces code by more than 30 % with minimal performance penalty. Due to their small size and low power consumption, LPC2141/42/44/46/48 are ideal for applications where miniaturization is key requirement, such as access control and point-of-sale.

Serial communications interfaces ranging from USB2.0 full-speed device, multiple UARTs, Serial peripheral Interface, SSP to I2C-bus and on-chip SRAM of 8kB up to 40kB, make these devices well suited for communication gateways and protocol converters, soft modems, Voice recognition and low end imaging, giving both large buffer size and high processing power. Various 32-bit timers, either single or dual 10-bit ADC(s), 10-bit DAC, PWM channels and 45 fast GPIO lines with up to nine edge or level sensitive external interrupt pins make these controllers suitable for home automation, industrial control and medical systems.

D. Overview of PSoC

Programmable System-on-Chips (PSoCs) designed and implemented by Cypress Semiconductor [8] contain a microcontroller, programmable analog blocks such as Op-amps, A/D, D/A converters, I/O drivers and digital blocks such as Universal Digital Block(UDB), CAN,PWM in a single chip as shown in Fig.7. PSoC has larger internal memory (256 KB flash, 64 KB SRAM) and has the high performance 32-bit ARM cortex- M3 CPU(CY8C55) and high performance configurable digital system with many components including interfaces such as USB and multi master I2C. Hence, it is chosen in this paper.

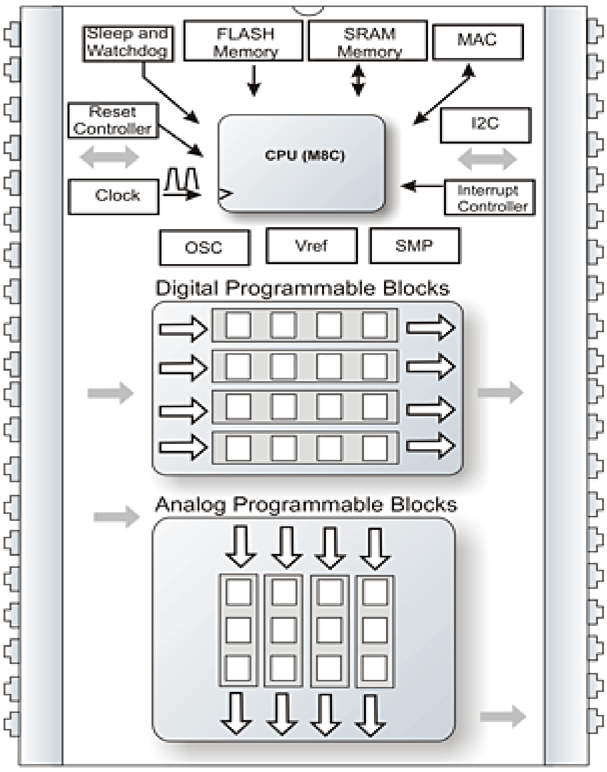


Fig.7. PSoC Architecture

For the development and implementation of systems using PSoC5, PSoC designer software is used. It is a visual embedded design tool and integrated development Environment (IDE). In order to support a hardware and software environment, PSoC designer combines a C based development flow with automatically generated software Application programmable interfaces (APIs) for the components in the design. APIs reduce coding errors and ensure proper interaction with the peripheral so that the whole firmware development is faster, easier and less error prone. PSoC designer has a powerful, modern debugger, built into the IDE which is used to display the values of the variables at each point after execution.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

E. Overview of Voice Recognition Module

The Voice recognition module is used for the recognize voice as shown in Fig.8. This module can be configured by sending commands through serial port. Configuration will be not erased after powered off. The serial data format contains 8 data bits, no parity and 1 stop bit. The default baud rate is 9600 and baud rate can be changed. The module can store up to 15 pieces of voice commands. Those 15 pieces are divided into 3 groups, with 5 in each group. First, need to train the module with voice recognition module group by group. After that, we should import one group before it would recognize the 5 voice instructions within that group. If we need to implement instructions in other groups, we should import the group first.

Fig.8. Voice Recognition Module

Before using it, we should train it by recording voice instructions. Each voice commands has the maximum length of 1300ms, which ensures that more commands can be recorded. Once you start recording, you couldn't stop the recording process until all 5 voice instructions recording of one group should be done. Also, once you start recording, the previous voice commands in that group will be erased. In training state, this module shouldn't reply to any other serial commands. It has 15 voice instructions in 3 groups. Each time we need to import the group before it could recognize commands in that group. That indicates, this module could recognize 5 voice instructions at the same time. For better results, place the MIC close to the module and make voice instructions distinctive.

IV. IMPLEMENTATION DETAILS AND RESULTS

The transmitter and receiver hardware is presented in figure 9 and 10.

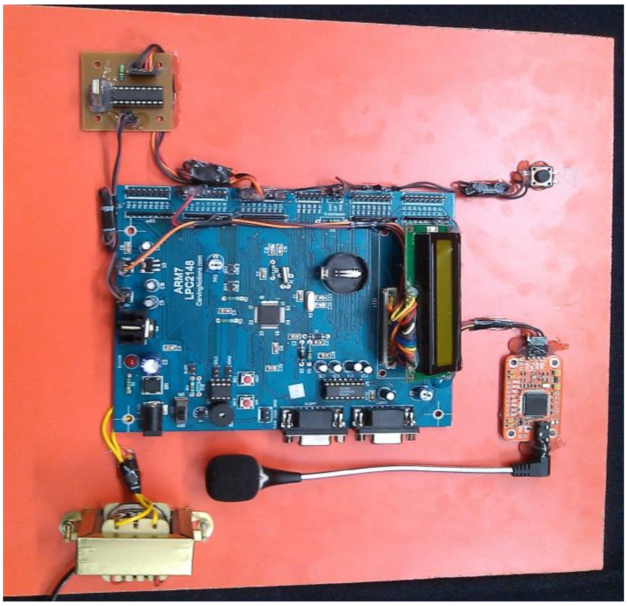
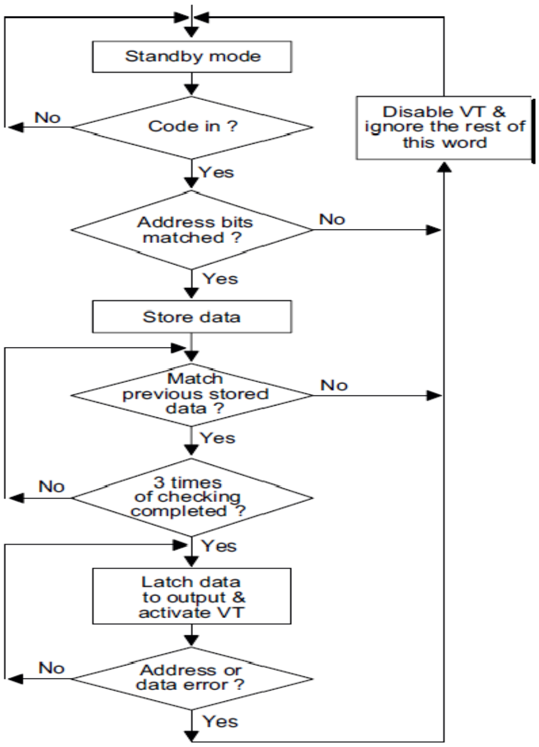


Fig. 9 Transmitter

First, the voice recognition module is put it in to the training mode and train with voice commands. Once the training is completed, store the data and deactivate the training command. The flow chart of training process is presented in flow chart 1.



Flow chart .1.

During recognition process, the received voice command from the speaker will go to the microcontroller through the voice recognition module. The controller receives the commands and compares with the existing trained commands. If the transmitted commands are matching with the trained commands, the microcontroller pass the received data information to the encoder for further processing otherwise it will write simply zeros on the data bus of the encoder. The encoder will convert the data according to the RF module standard and forward it to the RF transmitter. The RF transmitter receives the data and transmitted on 433MHz carrier through channel.

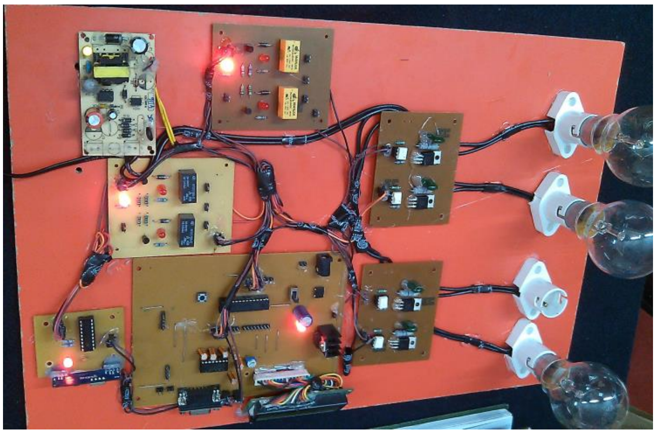


Fig.10. Receiver

At receiver side, the RF module receives the signal on 433MHz carrier. The received signal is converted in to digital format by the RF receiver and forwarded it to RF decoder for further processing. The decoder decodes the received information and is given to PSoC for further processing. The Psoc verifies the received data and compares it with the existing trained data; if the received data and trained data are matching, the PSoC generates the command to the particular load for performing the action.

The Hardware is catered for the future extension of the applications. The software was implemented using Keil IDE, Psoc design tool and Voice recognition IDE. The Hardware components ARM7, PSoC, Voice recognition module, RF Transceiver, Triac's, Mike, Loads, Regulated power supply and discrete passive components are used for developing this system.

V. CONCLUSION

The isolated speaker independent speech recognition system was successfully developed using LPC2148, Voice recognition module, RF Tx/Rx and PSoC controller. The developed system was tested by connecting four loads at receiver and operated remotely by using voice commands. The results were found satisfactory. This project can be widely used for an embedded application to operate the loads remotely by giving the voice commands. The proposed system can be integrated with any home appliances, industry machineries, automobiles and robots.

REFERENCES

[1] D.B. Fry, “Theoretical aspects of mechanical speech recognition, "Journal of the British Institution of Radio Engineers, vol.19, no.4.211-218, 1959.

[2]. H.F. Olson, H. Belar, "Time compensation for speed of talking in speech recognition machines, " IRE Transactions on Audio, vol.AU-8, no.3. 87-90, May-June 1960.

[3] Ackenhusen, J.Rabiner, L, "Microprocessor implementation of an LPC-based isolated word recognizer,” IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP '81.vol.6. 746-749, Apr 1981

[4] N. Hataoka, H. Kokubo, Y. Obuchi, A. Amano, "Compact and robust speech recognition for embedded use on microprocessors, " IEEE Workshop on Multimedia Signal Processing. 288-291, 9-11 Dec. 2002.

[5] J. Manikandan, B. Venkataramani, K. Girish, H. Karthic, V. Siddharth, "Hardware Implementation of Real-Time Speech Recognition System Using TMS320C6713 DSP", 24th International Conference on VLSI Design (VLSI Design).250-255, 2011

[6] J.Manikandan, B.Venkataramani, P.Preeti, G.Sananda, and K.V.Sadhana, "Implementation of a phoneme recognition system using zero-crossing and

magnitude sum function," IEEE Region 10 Conference TENCON 2009, pp.1-5,

[7] V. Amudha, B. Venkataramani, J. Manikandan, "FPGA implementation of isolated digit recognition system using modified back propagation algorithm,", International Conference on Electronic Design ICED 2008, pp.1-6,

[8] Cypress http://www.cypress.com

[9] V. Naresh, B. Venkataramani, Abhishek Karan and J. Manikandan, "PSoC based isolated speech recognition system", International conference on Communication and Signal Processing, April 3-5, 2013, India

Author Details

Ms. Teetla Asha received B.Tech degree in Electronics and Communication Engineering (ECE) from JNTUH, Hyderabad, Telangana, India and is pursuing M.Tech degree in Embedded systems at the St. Martin's Engineering College, Dhullapally, Hyderabad, India.

Mr. S.Ravi Kumar working as a Associate Professor in St. Martin's Engineering College, Dhullapally, Hyderabad, India. He received B.Tech degree in Electronics and Communication Engineering (ECE) from JNTU, Andhrapradesh, India in 2002 and the M.Tech degree in Radar and Microwave Engineering from Andhra University, Visakhapatnam, Andhrapradesh, India in 2005. He is currently pursuing Ph.D degree at the JNTUH, Hyderabad, India.