

Embedded Speech Recognition System Design and Optimization

Chandarmouli.K¹, Dineshkumar.S², Dinesh.V³, Dineshkumar.G⁴, Manjula.C⁵

UG scholar, Department of ECE, Adhiyamaan College of Engineering, Hosur, Tamilnadu, India^{1,2,3,4}

Professor, Department of ECE, Adhiyamaan College of Engineering, Hosur, Tamilnadu, India⁵

ABSTRACT: This paper, we present the design and implementation of a distributed sensor network application for embedded, isolated-word, real-time speech recognition. In our system design, we adopt a parameterized-data-flow-based modeling approach to model the functionalities associated with sensing and processing of acoustic data, and we implement the associated embedded software on an off-the-shelf sensor node platform that is equipped with an acoustic sensor. The topology of the sensor network deployed in this work involves a clustered network hierarchy. A customized time division multiple access protocol is developed to manage the wireless channel. We analyze the distribution of the overall computation workload across the network to improve energy efficiency. In our experiments, we demonstrate the recognition accuracy for our speech recognition system to verify its functionality and utility. We also evaluate improvements in networks lifetime to demonstrate the effectiveness of our energy-aware optimization techniques.

I. INTRODUCTION

To control and command an appliance (computer, VCR, TV security system, etc.) by speaking to it, will make it easier, while increasing the efficiency and effectiveness of working with that device. At its most basic level speech recognition allows the user to perform parallel tasks, (i.e. hands and eyes are busy elsewhere) while continuing to work with the computer or appliance. This circuit allows one to experiment with many facets of speech recognition technology. The heart of the circuit is the HM2007 speech recognition integrated circuit. The chip provides the options of recognizing either forty .96 second words or twenty 1.92 second words. This circuit allows the user to choose either the .96 second word length (40 word vocabulary) or the 1.92 second word length (20 word vocabulary). For memory the circuit uses an 8K X 8 static RAM.

The chip has two operational modes; manual mode and CPU mode. The CPU mode is designed to allow the chip to work under a host computer. This is an attractive approach to speech recognition for computers because the speech recognition chip operates as a co-processor to the main CPU. The jobs of listening and recognition doesn't occupy any of the computer's CPU time. When the HM2007 recognizes a command it can signal an interrupt to the host CPU and then relay the command code. The HM2007 chip can be cascaded to provide a larger word recognition library.

The SR-07 circuit we are building operates in the manual mode. The manual mode allows one to build a stand alone speech recognition board that doesn't require a host computer and may be integrated into other devices to utilize speech control.

II. BLOCK DIAGRAM OF EMBEDDED SPEECH RECOGNITION SYSTEM

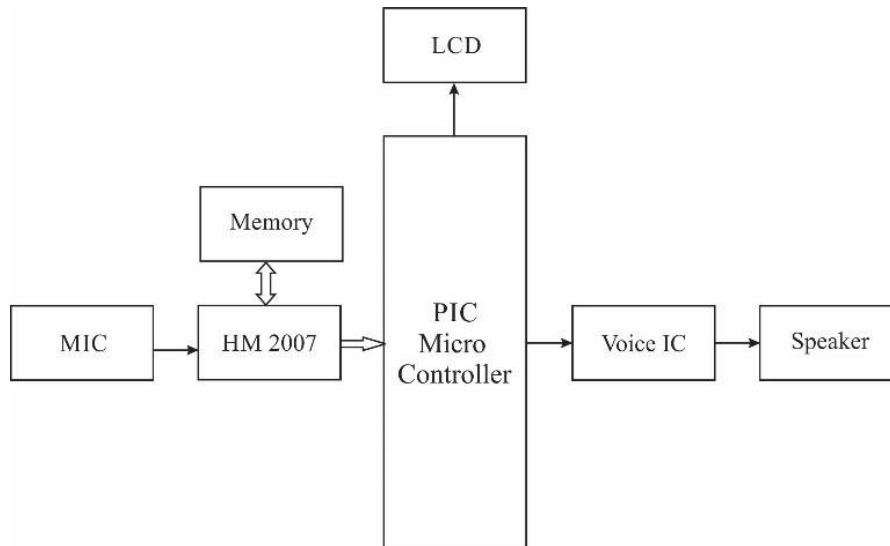


Fig.1 Block diagram of embedded speech recognition system

A. POWER SUPPLY

There are many types of power supply. Most are designed to convert Voltage AC Mains electricity to a suitable low voltage supply for electronic Circuits and other Devices. A power supply can be broken down into a series of blocks, each of which performs a particular function.



Fig.2 Block diagram of power supply

- Transformer converts from 230V to 5V and 12 V for the IC.
- Rectifier converts the power supply from AC to DC.
- Filter is used to compress the noise supply.
- Regulator is used to give the constant 5V to the circuit.

B. HM 2007

The heart of the circuit is the **HM2007 speech recognition integrated circuit**. The chip provides the options of recognizing either forty .96 second words or twenty 1.92 second words. This circuit allows the user to choose either the .96 second word length (40 word vocabulary) or the 1.92 second word length (20 word vocabulary). For memory the circuit uses an 32K X 8 static RAM.

The chip has two operational modes; manual mode and CPU mode. The CPU mode is designed to allow the chip to work under a host computer. This is an attractive approach to **speech recognition** for computers because the **speech recognition** chip operates as a co-processor to the main CPU. The jobs of listening and recognition don't occupy any of the computer's CPU time. When the **HM2007** recognizes a command it can signal an interrupt to the host CPU and then relay the command code.

The **HM2007** chip can be cascaded to provide a larger word recognition library.

C. PIC MICROCONTROLLER

The microcontroller that has been used for this project is from PIC series. PIC microcontroller is the first RISC based microcontroller fabricated in CMOS (complimentary metal oxide semiconductor) that uses separate bus for instruction and data allowing simultaneous access of program and data memory. The main advantage of CMOS and RISC combination is low power consumption resulting in a very small chip size with a small pin count. The main advantage of CMOS is that it has immunity to noise than other fabrication techniques.

Various microcontrollers offer different kinds of memories. EEPROM, EPROM, FLASH etc. are some of the memories of which FLASH is the most recently developed. Technology that is used in pic16F877 is flash technology, so that data is retained even when the power is switched off. Easy Programming and Erasing are other features of PIC 16F877.

D. MEMORY ORGANISATION

There are three memory blocks in each of the PIC16F877 MUC's. The program memory and Data Memory have separate buses so that concurrent access can occur.

2.4.1 PROGRAM MEMORY ORGANISATION

The PIC16f877 devices have a 13-bit program counter capable of addressing 8K *14 words of FLASH program memory. Accessing a location above the physically implemented address will cause a wraparound.

The RESET vector is at 0000h and the interrupt vector is at 0004h.

2.4.2 DATA MEMORY ORGANISATION

The data memory is partitioned into multiple banks which contain the General Purpose Registers and the special functions Registers. Bits RP1 (STATUS<6>) and RP0 (STATUS<5>) are the bank selected bits.

RP1:RP0	Banks
00	0
01	1
10	2
11	3

Table 1. Data memory organisation

Each bank extends up to 7Fh (1238 bytes). The lower locations of each bank are reserved for the Special Function Registers. Above the Special Function Registers are General Purpose Registers, implemented as static RAM. All implemented banks contain special function registers. Some frequently used special function registers from one bank may be mirrored in another bank for code reduction and quicker access

E. LIQUID CRYSTAL DISPLAY (LCD)

Liquid crystal displays (LCDs) have materials, which combine the properties of both liquids and crystals. Rather than having a melting point, they have a temperature range within which the molecules are almost as mobile as they would be in a liquid, but are grouped together in an ordered form similar to a crystal.

An LCD consists of two glass panels, with the liquid crystal material sandwiched in between them. The inner surface of the glass plates are coated with transparent electrodes which define the character, symbols or patterns to be displayed polymeric layers are present in between the electrodes and the liquid crystal, which makes the liquid crystal molecules to maintain a defined orientation angle.

One each polarizes are pasted outside the two glass panels. These polarizes would rotate the light rays passing through them to a definite angle, in a particular direction. When the LCD is in the off state, light rays are rotated by the two polarizes and the liquid crystal, such that the light rays come out of the LCD without any orientation, and hence the LCD appears transparent.

When sufficient voltage is applied to the electrodes, the liquid crystal molecules would be aligned in a specific direction. The light rays passing through the LCD would be rotated by the polarizes, which would result in activating / highlighting the desired characters. The LCD's are lightweight with only a few millimeters thickness. Since the LCD's consume less power, they are compatible with low power electronic circuits, and can be powered for long durations.

The LCD does not generate light and so light is needed to read the display. By using backlighting, reading is possible in the dark. The LCD's have long life and a wide operating temperature range. Changing the display size or the layout size is relatively simple which makes the LCD's more customers friendly.

F. Software Approach

Currently most speech recognition systems available today are programs that use personal computers. The add on programs operate continuously in the background of the computers operating system (windows, OS/2, etc.). These programs require the computer to be equipped with a compatible sound card. The disadvantage in this approach is the necessity of a computer. While these speech programs are impressive, it is not economically viable for manufacturers to add full blown computer systems to control a washing machine or VCR. At best the programs add to the processing required of the computer's CPU. There is a noticeable slow down in the operation and function of the computer when voice recognition is enabled.

III. RESULT AND DISCUSSION

The speech recognition is a technology, which helps to transfer the speech signals into homologous text or command during the process of recognition and comprehending. Recently speech recognition technology becomes a research hot spot for its importance and difficulty.

IV. CONCLUSION

With the development of embedded systems, speech recognition technology based on embedded systems has become a new and important direction in this field. In hardware platform, it analyses and designs the general architecture for the system, mainly including the following three parts, system's control part, input and output part, and storage part; it also introduces the functions of each part as well as their relationship in details. In addition, according to actual needs, it makes the selection for the audio chip, and other external circuit chips, and expands the external circuit. In the aspect of system software, not only it requires its perfect function to meet the needs of this design, but also every parts meet certain compatibility. Namely the system has good stability and can work effectively. The speech recognition technology has developed to a mature stage and has practical products. But the speech recognition under the embedded environment does not develop as fast as people expected. The speech recognition process of embedded system is realized. This system still has a lot of aspects which needs to be improved.

REFERENCES

- [1] Reichl W, chou W, Robust decision tree state tying for continuous speech recognition. IEEE Trans Speech and Audio Processing 2000, 8(5): 555 – 566.
- [2] John M. Paulett, Curtis P. Langlotz. Improving Language models for Radiology Speech Recognition. Journal of Biomedical Informatics, 2009, 1(42):53 – 58.
- [3] K.S.R. Murthy, B. Yegnanarayana. Combining Evidence from Residual Phase and MFCC Features for Speaker Verification. IEEE Signal Processing Letters, 2006, 13(1):52 – 55.
- [4] Mingyu Chu Stephen. System and Method for Likelihood Computation in Multi-stream HMM Speech Recognition. Acoustical Society of America, 2013, 133(1):163-164.
- [5] A. Amrouche, A. Taleb-Amed, J.M.Rouvaen. Improvement of the Speech Recognition In noisy Environment Using a Nonparametric Regression. International Journal of Parallel, 2009, 24(1):49-67.

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization

Volume 6, Special Issue 3, March 2017

5th National Conference on Frontiers in Communication and Signal Processing Systems (NCFCSPS '17)

13th-14th March 2017

Organized by

Department of ECE, Adhiyamaan College of Engineering, Hosur, Tamilnadu, India.

- [6] Li C J. Andersen. Efficient Blind System Identification of Non-Gaussian Autoregressive Models with HMM Modeling of the Excitation. IEEE Transactions on Signal Processing, 2007, 55(6):2432-2445.
- [7] Lichi Yuan. An Improved HMM Speech Recognition Model. IEEE Transaction on Consumer Electronics, 2008, 7(8):1311-1315.
- [8] C.Kim, R.M. Stern. Feature Extraction for Robust Speech Recognition Using a Power-law Nonlinearity and Power-bias Subtraction. INTERSPEECH, 2009. 28-31.
- [9] Hanpu, Jiang jie. HMM application research in the field of natural language processing. Computer technology and development, 2010, 20(2):245-253.
- [10] Anuj Mohamed,K.N.Ramachandran Nair. HMM/ANN Hybrid Model for Continuous Malayalam Speech Recognition. Procedia Engineering, 2012, (30):616-622.