International Journal of Innovative Research in Computer Science & Technology (IJIRCST)
ISSN: 2347-5552, Volume-3, Issue-2, March 2015

Implementation of Speech Recognition System on Raspberry Pi

Sonal Agrawal, Mrs. S.Krithiga

***Abstract*— The speech recognition system was proposed for the first time in early 1950's in which fixed vocabulary were deployed in many applications. Speech recognition techniques have been proposed in the literature for convenient and efficient human-machine interaction Speech recognition system targeted for voiced operated appliances, robots, access control system etc. In this proposed system, the implementation of an Isolated Speech Recognition System (ISRS) is done by using Raspberry pi which has ARM 1176JZFS processor. Recognition performance has been studied using feature extraction techniques such as Zero crossing point (ZCP) and End point detection (EPD). Result of the proposed work is verified by using softwares such as QT creator and MATLAB. This proposed work reduces computation and cost in recognizing the speech**

***Index Terms*— ZCP, EPD, HMM, LPC, Speech Recognition, Raspberry Pi**

I. INTRODUCTION

Various methods have been proposed for recognizing speech which includes Hidden Markov Model (HMM) [1], Linear Predictive Coding (LPC) [2] etc. Some techniques have been implemented using many hardware platforms such as microprocessor, FPGAs, PSoC 5 (ARM cortex-M3) etc. HMM and LPC causes very high computation and FPGA is only digital in nature. Existing technique has been implemented on PSoC 5 [3] which is very costly in comparison to conventional techniques. This paper describes the speech recognition technique developed on Raspberry Pi [4] K4P4G324EB model B which has 512MB SDRAM, 700 MHZ low power ARM11, CPU, GPU, GPIO, operating system by booting up SD card. The Broadcom SOC used in the Raspberry Pi is equivalent to a chip used in smartphone. It is specially used for media. It is called as $25, as it is low in cost and provide high level applications.

II. SYSTEM ARCHITECTURE

The proposed system consists of other hardwares too which include transformer, Voltage regulator, MIC, Amplifier, ADC, ST232, and RS232. Raspberry Pi is used in media so it is more appropriate device for speech recognition. The GPU provides 1 Pixels/s or 1.5 Gtexel/s of graphics processing or 24 GFLOPS of general purpose computing performance.

Speech is captured by microphone which is later amplified by amplifier LM741 and converted in to digital form by using ADC MPC3208. The digitized speech are then extracted by using ZCP [5] and EPD techniques.

Raspberry Pi K4P4G324EB has one GPIO which connect with ADC. It contains one SD card which should be of at least 8GB. Raspberry pi get power supply through a cable which can connect directly to phone charger, laptop and computer. It is the quality of Raspberry Pi that it does not contain motherboard, any problem in SD card will not create problem for the whole board and is even useful for backup. Another card can be used after that.

Interfacing between Raspberry Pi and hardware is done with the help of QT creator software. It is installed in SD card of Raspberry Pi. It is a lightweight tool with advanced C++ code editor and a graphical user interface for debugging C++ function. It creates three files while creating a program named main.cpp, mainwindow.cpp and mainwindow.ui so that data can visible on screen on varying the speech.

Simulation results are obtained using MATLAB. The main advantage of MATLAB is that its graphical output is optimized for interaction. Other calculations like mean, variance, and kurtosis value can easily done on this software. These calculations help in recognizing the speech.

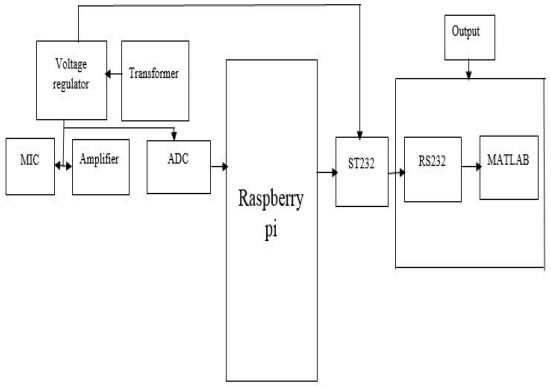

Fig-1: Proposed System

Manuscript received March 25, 2015.
Sonal Agrawal, Electronics and Communication Department, SRM University, Kattankulathur, Chennai, India
Mrs. S.Krithiga, Electronics and Communication Department, SRM University, Kattankulathur, Chennai, India

III. ALGORITHMS

Feature extraction is a process required for recognizing the speech. It extracts the necessary information from speech signal which helps in getting the goal. In this paper two algorithms have been choose for extracting the features of speech.

1) Zero Crossing Point (ZCP)

2) End Point Detection with Zero Crossing point (EPD with ZCP)

These two algorithms have been used through MATLAB. The following parameters have been calculated with both algorithms. If a word does not crosses zero for a single time, these parameters will help in the recognition process.

a) Mean (M) = $(x_1+x_{2.....}x_n)/n$

b) Variance (V) = Σ [(xi-M)2]/n-1

(Where x_i = values at different point of speech and i = 1, 2….. n)

A) Zero Crossing Point

In this paper zero crossing is calculated by calculating the number of time the speech signal is crossing zero and the other calculations also help in extracting the feature. Zero crossing requires less computation in comparison to other features used like Linear Predictive Coefficient (LPC) and vector method. LPC features are extracted from the input speech using the pre-emphasis, frame blocking, windowing, and autocorrelation block which requires more operation for addition and multiplication. In LPC the computation is 13 times higher than zero crossing. So zero crossing is better in recognizing the speech in comparison to linear predictive coding.

Vector quantization is a technique which calculate value at every point of signal so it become more complicated in comparison to zero crossing. Mean is calculated with ZCP to give more exact computation in recognizing the speech.

B) End Point Detection with Zero Crossing Point

End point detection is also another algorithms which extract the features from speech signal. It is used with zero crossing in this paper. It tells the end point of signal which helps in recognizing the speech. It gives more appropriate result in comparison to ZCP. Mean and variance are also calculated to give more exact computation in recognizing the speech. It also gives point to point value of the speech obtained from mic. End point value is that value which it obtain at last for that speech signal. It can also assume by the number of times a speech signal crosses zero because it is after the last zero crossing point.

IV. SIMULATION RESULT

The simulation results obtained after execution are shown in this section. These computations have been done in very less time. The following Fig.3 and Fig.4 show the simulation result. In Fig.3 values of speech signal at different points with other parameters like mean, variance etc. is shown. In Fig.4 graph of different speeches is shown in a single graph. Comparison is done in a single graph so that it can use less memory.

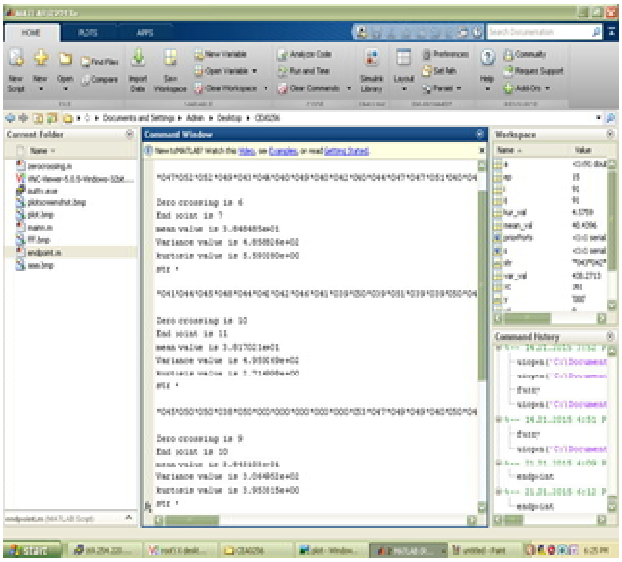

Fig-2: Simulation Result

.

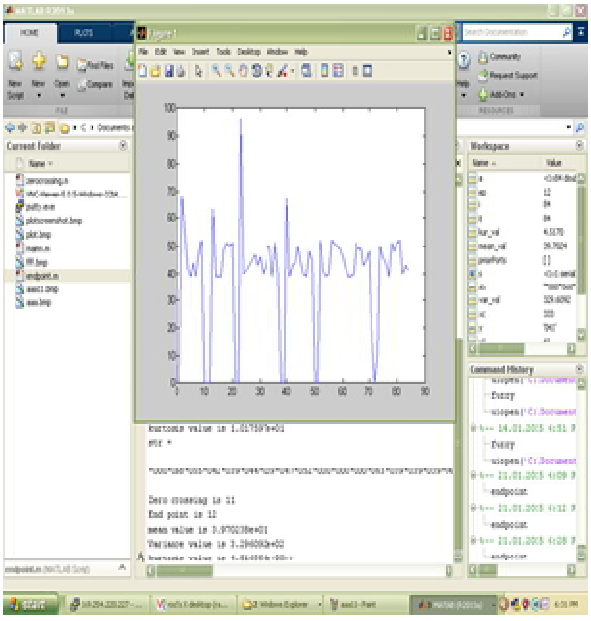

Fig-3: Simulation Result

V. CONCLUSION

The speech recognition is successfully implemented on Raspberry Pi. The algorithms using zero crossing and end point detection results in good accuracy with the reduction in cost and computation. The results achieved using the proposed recognition system is satisfactory and gives a scope for development in future. The number of words for recognition in a single time can be increased according to memory available. The proposed system implemented on Raspberry Pi can be used with home appliances, robots etc. .

REFERENCES

[1] V.Amudha, B.Venkataramani, R.Vinoth Kumar and S. Ravishankar,"SOC Implementation of HMM Based Speaker Independent Isolated Digit Recognition System", in Proc. of IEEE Int. Conf. on VLSI Design VLSI'07, Tiruchirappalli, 2007,

[2] 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, New Jersey, Apr 1981,

[3] 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, India, 2013,

[5] Charles Severance, "Eben Upton: Raspberry Pi," IEEE Journal, Volume 46, Issue 10, 2013,

[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, Trichy, 2009,