CONTROLING OF DEVICES THROUGH VOICE RECOGNITION USING MATLAB

Ajay K. Yadav[1], Alokit Upadhyay[2], Shekhar Rastogi[3], Sonali Shukla[4]

[1,2,3,4] Department of EI Engineering, FET, MJP Rohilkhand University, Bareilly, (India)

ABSTRACT

This paper proposes a creative work focused on designing an intelligent living-space with automatic speech recognition system to control all home appliances basically electrical and electronic devices. System is designed to help people with disabilities to do their works at home. We are introducing a new technique in which we use speech recognition principles to generate control commands at the parallel port. The programming languages we used in this research are MATLAB. Two main parts in the system introduced is the voice train process and voice recognition process. The main objectives are to design a GUI (Graphical User Interface) and a source code that can recognize human voice. This system used speaker independent method; any users who had trained their voice in the system can use the system. This is because, system built to recognize user's speech and not user's identity. System user record their voice by using microphone, after recording, system processes the voice, store it, and compare it. Then system sends the data to hardware that has been connected to home electrical appliances through serial connector.

Keywords: Automatic Speech Recognition, MATLAB, GUI (Graphical User Interface), Source Code, Serial Connector.

I. INTRODUCTION

Speech is a natural medium of communication for humans, and in the last decade various speech technologies like automatic speech recognition (ASR), voice response systems etc. have considerably matured. While this has been true since the dawn of civilization, the invention and widespread use of the telephone, audio-phonic storage media, and radio has given even further importance to speech communication and speech processing. Reliable speech recognition is a hard problem, requiring a combination of many techniques; however modern methods have been able to achieve an impressive degree of accuracy. [1]. The advances in digital signal processing technology has led the use of speech processing in many different application areas like speech compression, enhancement, synthesis, and recognition. [2].

This paper details the construction and building of a stand-alone and very less expensive speech recognition technique[3] that may be used to control just about anything such as electrical appliances, robots, test instruments, Fans, Light, TV's, etc.

Speech recognition is a multileveled pattern recognition task, in which signals are examined and structured into a sequence of sub word units (e.g., phonemes), words, phrases, and sentences. [4]. To control and command an appliance by speaking to mice, will make it easier, while increasing the efficiency and effectiveness of working with that device. [5] It is useful for physically handicap persons in their day to day activity. So our main goal here is to develop the MATLAB Based Automatic Speech Recognition System which is able to decrease the noise level up to .8 dB in some commands and more Users Friendly System based GUI so that every user doesn't need any major training sessions before using the system. [6]

II. PROPOSED SYSTEM

The proposed system aims at designing a device controlled voice recognition system for the following reasons. One of the major problems in our present day society is wastage of energy, whereas energy consumption is continuously increasing year by year. Nowadays, some people may be too lazy or too busy to get up and turn off a particular appliance. Hence, the smart device control system will be useful as one will only have to speak to turn off a device, thereby saving energy as well as one's time. Moreover, old or disabled persons may experience difficulties in going around the house to turn on/off their appliances, especially if they live alone. It will be much easier for them to use the voice control system. It will also help blind people as they will be able to turn on a fan or a radio without relying on others. The system is designed in such a way that it is easy to install and use. The proposed method is to use MATLAB for speech processing and recognition. The output will be sent, through the XBee transceivers, to the control part, where a microcontroller will select the required device according to the input voice command.

The system can be divided into 3 main parts:

1. Audio processing part -Windows OS and MATLAB
2. Transmission part - XBee transceivers
3. Control part - microcontroller and relay

2.1 Block Diagram of The Proposed System

The voice commands will first be recorded and processed using MATLAB according to the voice recognition method used. After successful identification of the commands, control characters will be wirelessly sent through the XBee transceivers to the microcontroller, which will in turn activate the corresponding relay. As a result, home appliances could be turned on or off depending on the voice command given.

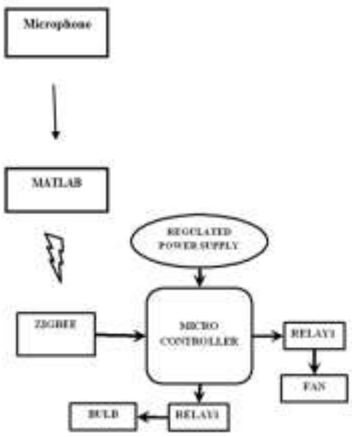

Fig. 1. Block Diagram of Proposed System

2.2 Description of Proposed System

2.2.1 Hardware Design

In this section, we will present the hardware description of the 3 main parts that form our Voice Recognition Device Control System.

(i) Audio Processing

This basically includes microphone used as an input device to record training words while record mode and input source for speech processor operation. Speech processor HM 2007 is mother board for this module, which is responsible for speech recognition. HY-6264 8-bit static RAM device is used for data back-up stored in speech processor when power is cut. Latch 74573 is used to separate data signals of processor to be fed to the microcontroller 89C51/52. This unit consists of two vital subsystems, Speech Processor and Microcontroller. The main element of the design consist of speech recognition ASIC and ZIGBEE transceiver module. The speech recognition IC needs to be programmed for various speech commands. This is done by using a MIC, analog speech signal from MIC are stored in the internal memory of the IC after being digitized using ADC blocks. This learning process need to be done for a number of speech commands. Once the learning is completed the IC is ready to accept the commands. A command issued by the user through MIC will be digitized and compared with the digitized commands already stored in the internal memory of IC. When a match is received, microcontroller status will be updated accordingly. The microcontroller in turn will generate a specific data pertaining to a given equipment/appliance and speech command which will transmitted through RF channel using ZIGBEE transmitter.

(ii) Microcontroller module

This module receives signal from speech processor and decision making regarding switching ON/OFF appliance is made here. Assembly program is used to do the same and burnt in microcontroller. Wireless module is connected to one of the port of microcontroller. Generally port 3 is used for transmission of signal.

(iii) Wireless (zigbee) module

Zigbee protocol is the communication protocol that is used in this system for wireless communication. Zigbee offers 250kbps as maximum baud rate. However, 115200 bps was used for sending and receiving as this was the highest speed that the UART of the microcontroller could be programmed to operate at.

(iv) Appliance control module

Once the speech commands are recognized, control charterers are sent to the specified appliance address through Zigbee communication protocol. Each appliance that has to be controlled has a relay controlling circuit.

III. SOFTWARE DESIGN

Software design for this system involves training for speech processor as HM 2007 is ASIC processor and assembly programming for microcontroller to work on incoming signals from speech processor. It receives control commands from processor and switches ON/OFF relay connected to the particular appliance. Our system simultaneously works with MATLAB code, where speech recognition will be carried out using MFCC algorithm. If voice command is matched with the sample stored in system then and then only code will generate matched waveform and appliance will be activated.

(v) MFCC algorithm

Mel frequency Cepstral Coefficient (MFCC) is used to extract the features from voice and Vector quantization technique to identify the speaker. Voice has an infinitive amount of information, we have to determine who is the person speaking based on the features of the person´s voice. So an analysis in the frequency domain can be a more viable option. Extract the parametric representation of voice signals is a vital process for the recognition performance. MFCC is a technique based on human hearing behaviour that cannot recognize frequencies over

1Khz. MFCC are based on the difference of frequencies that the human ear can distinguish. The signal is expressed in the MEL scale, this scale is based on the perception of the pitches in an equally spaced intervals judged by observers. This scale uses a filter that is spaced linearly at frequencies below 1000 Hz.

(vi) FRAMING

It is the segmentation of the speech samples in boxes within the range of 20 ms to 40 ms. The voice signal is divide in frames of N samples. Adjacent frames are separated by M(M<N). Distinctive values used are M= 100 and N= 256.

MFCC algorithm follows following steps:

1) Automatically detects an isolated word from your speech (input utterance). It does this by calculating energy and the number of zero crossings on a frame-by frame basis, and compares these values to a threshold.

2) For each frame in a word, it applies a window function followed by a pre-emphasis filter. It then calculates Mel Frequency Cepstral Coefficients (MFCC) and their delta and delta-delta coefficients for each frame, and uses these as feature vectors.

3) For each word's training data, we estimate the parameters of a Gaussian Mixture Model to fit the distribution of training vectors. We therefore train a model to represent each word.

IV. FLOW DIAGRAM OF PROPOSED MODEL

Working of this system is based upon flow diagram given below.

- The MATLAB is first initialized in order to take the command to be followed with the help of the microphone and the template of the input command signal is created.
- Next, we have to determine whether a voice signal is actually present. For this purpose, a “noise-gate” is used, which filters out low frequency noise signals. Thus, noise is replaced by silence, which helps us in effectively determining when an actual voice signal is present.
- If it is valid, all samples that are beyond a certain threshold level are stored in a buffer until the voice signal is no longer present.
- Now the template of command received is compared with the templates of dictionary using Euclidian distance i.e. the accumulation of the square of each difference between the value of dictionary template and that of command template at each sample points.
- The formula can be given as

Euclidian Distance=Σi=1(dic[i]-com[i])^2

After calculating Euclidian distance for each dictionary template, these distances are sorted in the ascending order to find out the smallest distance among them which corresponds to a particular dictionary template which is the template belonging to a particular dictionary command.

- Then the device detects that particular command given by the operator. If the command given by the operator does not match with any of the dictionary command then the device should not follow that command.
- A specific port of the microcontroller is set as the output port. This port will be used for controlling the turning on and off of the relays
- Global interrupts are then enabled and the UART is initialized, which involves baud rate being set to 9600 bps and data frame format being given, that is, 1 stop bit, 8-bit data, no parity, etc.

- The microcontroller then waits for an interrupt to occur. Whenever a control character is received, the Interrupt Service Routine (ISR) runs and those characters are compared with predetermined ones.
- For example, if character “a” is received, the first pin of port C is set high or low; similarly, if character “b" is received, the second pin of port C is activated or deactivated accordingly. If any other characters are received, no action is taken. After having activated/deactivated the corresponding pins, the microcontroller again waits for the next character to be received.

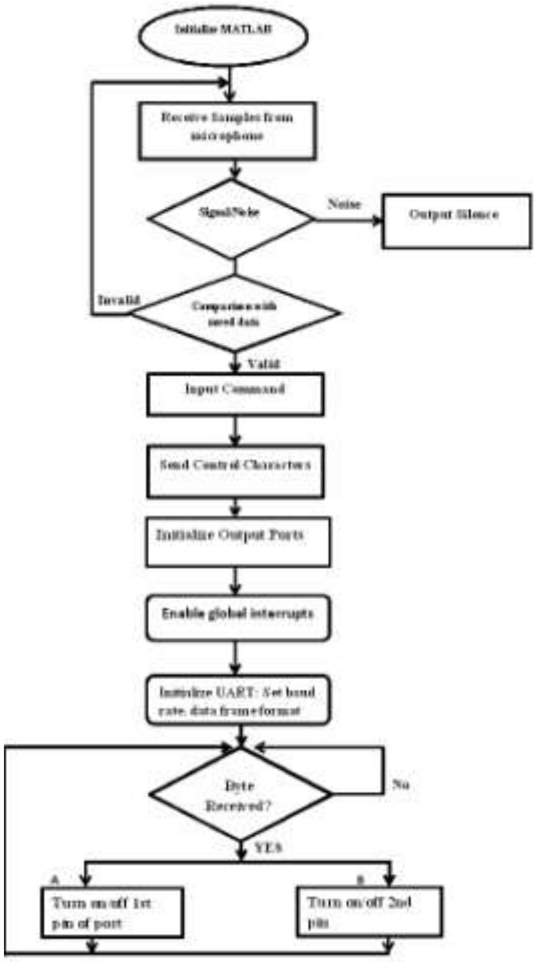

Fig. 2 Flow Diagram of Proposed System

V. CONCLUSION

After studying the proposed work we present a simple technique for design and development of a system for reliable recognition of voice with the following results:

- The system implements Automatic Speech Recognition using speech processor and MATLAB coding. This system possesses high efficiency.
- The recognized voice commands are displayed on GUI in text format.
- As voice command is recognized by MATLAB, the specific data is send to microcontroller. The controlling of devices using microcontroller is implemented according to received data
- The system developed can be used to control AC and DC appliances through voice.
- The system is targeted at elderly and disabled people to help ease their life.

- The proposed system therefore provides solutions for the problems faced by old or disabled persons in daily life and makes their life easier and more comfortable by proposing a cost effective and reliable solution.
- This system can be made highly efficient and effective (noise free) if stringent environmental conditions are maintained.
- The setup for maintaining these environmental conditions will be a onetime investment for any real life application.

REFERENCES

[1] S. K. Hasnain, A. Beg and S. Awan, "Frequency Analysis of Urdu Spoken Numbers Using MATLAB and Simulink" Journal of Science & Technology PAF KIET ISSN 1994- 862x,pp: 46-48, Dec. 2007

[2] Martinez, J., Perez, H, Escamilla, E..Suzuki, M.M – "Speaker recognition using Mel frequency Cepstral Coefficients (MFCC) and Vector quantization (VQ) techniques" - 22nd International Conference on Electrical Communications and Computers (CONIELECOMP), pp: 248-251, 27-29 Feb 2012

[3] Z. Goh et al., "Kalman-Filtering Speech Enhancement Method Based on a Voiced-Unvoiced Speech Model", IEEE Trans. Speech Audio Processing, vol. 7, pp. 510-524, Sep 1999

[4] Y. Romanyshyn, V. Hudym, "Wavelet transforms applications for speech signals processing" , 6th International Conference on The Experience of Designing and Application of CAD Systems in Microelectronics, 2001. CADSM 2001, pp: 297-298, 12-17, Feb 2001

[5] Arvinder Singh, Gagandeep Singh, "Speech Recognition Based System to Control Electrical Appliances" International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 2, pp 81-83, August 2012

[6] Punit Kumar Sharma, Dr. B.R. Lakshmikantha and K. Shanmukha Sundar, "Real Time Control of DC Motor Drive using Speech Recognition" India International Conference on power electronics, (IICPE), pp. 1-5, 28-30, Jan 2010

[7] Punchihewa, A., Arshad, Z.M. "Voice command interpretation for robot control" 5th International Conference on Automation, Robotics and Applications (ICARA), pp: 90-95, 6-8, Dec 2011

[8] Haleem, M.S. "Voice Controlled Automation System" Multitopic Conference, 2008. INMIC 2008. IEEE International, pp: 508-512, 23-24, Dec 2008

[9] Katti and Anusuya," Speech Recognition by Machine: A Review", International Journal of Computer Science and Information Security, Volume: 64, pp: 501 - 531 IEEE, 2009

[10] A.K. Gnanasekar, Jayavelu, P, Nagarajan V., "Speech recognition based wireless automation of home loads with fault identification for physically challenged" International Conference on Communications and Signal Processing (ICCSP), pp: 128-132, 4-5, April 2012