A Bluetooth Based Spoken Query System for Controlling Electronic Devices

Hirakjyoti Sarma*
Dept. of Computer Science and Technology
Bodoland University
Kokrajhar (BTC), India
hiraks.88@gmail.com

Manoj Deka
Dept. of Computer Science and Technology
Bodoland University
Kokrajhar (BTC), India
Manojdeka1@gmail.com

Abstract— This paper discusses a spoken query system developed for accessing some household electronic devices. The developed system enables the user to access these devices remotely by calling the system using a GSM modem/ Mobile phone. The spoken query system has mainly two modules i.e. interactive voice response (IVR) and automatic speech recognition (ASR) modules, which can be developed using open source resources. In the development of ASR models, the system specific data are collected from people of different ages and genders. The issues of data preparation, Training and Testing are performed by using HTK as a speech recognition tool. The system can recognize sample data i.e. the voice commands independent of vocabulary size, noise, and speaker characteristics. The recognized voice command as a text file will be used as an input into the microcontroller that is responsible to control the electronic home appliances. This speaker independent interactive voice response system is based on HMM (Hidden Markov Model) which is found to give a relative improvement of accuracy by more than 50%.

Keywords - IVR, Speech, ISR, HTK, Training, Testing, HMM

I. INTRODUCTION

In today's modern era, home automation is becoming a crucial area in research as information technology is growing so rapidly from the computing to communication. As we know, the speech is a special kind of communicator among all human beings for their communication, that is why in this work, research is emphasized for developing a Speaker independent speech module which controls electronic devices by the voice of human beings i.e. communication between man and machine. At the present time, people want to get comfort as well as easy and less time consuming day to day life. So lots of home automation system controlled by different communication mechanism is introduced to ease the daily life of human being. But most these systems are not convenient, energy efficient, and safety in real time. Therefore although several development leading to Automate the electronic or electrical devices over wireless already developed, but in this work a special focus has been given to introduce voice in terms of speech not only for a natural communication, also for the communication between man and machine. In this entire work, an user friendly spoken query system consists of interactive voice response (IVR) and automatic speech recognition (ASR) modules is designed for controlling some household electronic devices. It is easy and cost effective to integrate the telephone network with automatic speech recognition (ASR) system. As a result developing a spoken query (SQ) system for accessing the appliances this system appears as more appropriate home automation system controlling by speech independent to

speaker. The designed system is a wrapping of two modules containing a SQ (spoken query) system and a hardware interface connected to the electronic devices. The developed spoken query system again contains two definite modules such as an interactive voice response (IVR) and an ASR system based on Hidden Markov Model (HMM) which are developed using open source resources. In the development of ASR system, controlling oriented specific speech data was collected from more 200 peoples of different ages including male and female using the same IVR module. The hardware interface is responsible to take the voice command in terms of text content and switching the electromagnetic relay which controls the electronic devices. The core of this hardware module is low-power consuming, but high-performing AT89S52 Microcontroller with programmable flash memory.

II. SPOKEN QUERY SYSTEM

The Spoken query (SQ) system is developed to provide speaker independent home automation system. The SQ system consists of a server running Asterisk, an ASR system and a controlling information database for all electronic devices connected to the system. Asterisk is open source software on Linux/Unix platform that enables in connecting the server to the telephone network[4]. The Asterisk server consists of an Interactive Voice Response and a computer telephone interface (CTI) card. The CTI card is connected to the integrated services digital network (ISDN) primary rate interface (PRI) digital line Devices such as IP phone, mobile phones or landline can access the Asterisk server through the ISDNPRI line. In our designed

system, a mobile phone is used to access the Asterisk server through the Bluetooth line. The developed system enables the user to make a query to control the devices and check the current status of the devices through a pre-recorded voice response. The system query system interacts with the user with a user friendly call-flow consisting of two major parts. In the first part, the user is prompted to utter electronic device name and then the system will inform the present status of devices and wait for the users' utterance. The block diagram of our SQ system is given in Fig. 1.
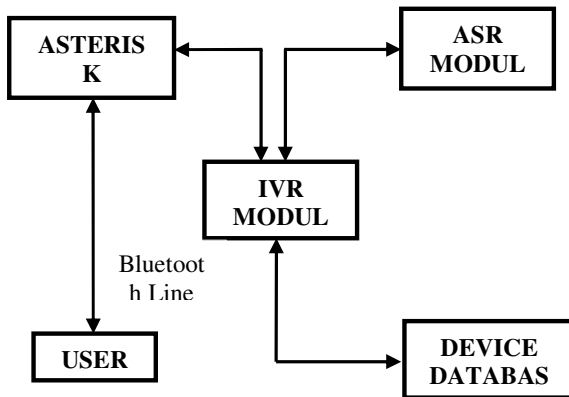
Fig. 1 – Block Diagram of SQ system

A. Device Controlling Database

In this system, all the information regarding the each and every operations performed on electronic devices should be stored for the further uses by another user. To serve this purpose, the MySQL database management system is used. The device controlling information is updated using this database. The query from the user is used to retrieve the device present information that is stored in the database. The status and names of the different devices are disseminated through the pre-recorded voice commands.

III. METHODOLOGY

The proposed system is an automation system for controlling devices consists of two main modules, the speech recognizer server and the hardware interface. User commands are transferred to the home automation server via speech recognizing system. In the home automation server the incoming commands are processed, then digitized and sent to the relevant unit to be processed. These hardware units have also the capability of sending their status back to microcontroller which is connected to the home automation server thus they can be monitored in real time. After receiving the feedbacks from the appliance nodes, the home automation server interprets them and performs the necessary tasks.
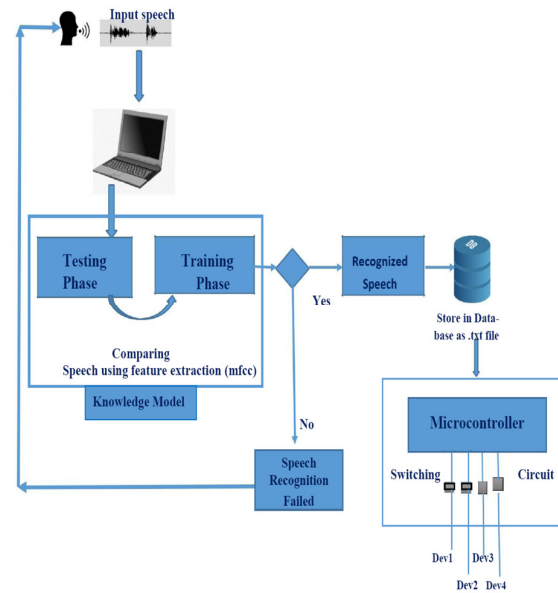

Fig. 2 – Proposed Framework of the system

A. Input Speech

In our proposed system we use speech as command for controlling electronic device. Therefore we have collected our required voice data from different people in different environment. For collection our required data, we have used microphone and wave surfer as a voice recording toolkit. We have tried to avoid the noise in our recording environment. Still in our collected data, we found some noisy samples. Depending on noisy utterance, we have ignored such samples for further processing. We have saved all voice data in .wav file extension with 16000 bit rate. Collection of voice data for speech recognition is very challenging because the accuracy of recognition depends on this collected input speech as well as collecting training data.

B. Testing phase

In testing phase the system collect voice command from microphone or stored database. The voice commands collected from microphone are saved into database in .wav file extension. Then speech recognizer verifies the voice data by feature extracting using MFCC. Then the extracted MFCC of voice data forward training phase.

C. Training Phase

An HMM-based system, like other speech recognition systems, functions by first learning the characteristics of a set of sound units, and then using what it has learned about the units to find the most probable sequence of sound units

for a given speech signal. The process of learning about the sound units is called training.

D. Recognition Phase

After training and testing phases completed, our process goes through a recognition process. As mention earlier we have used HTK as recognition tool. We have trained our collected data in HTK training phase. Again we have stored some samples as testing data in our database in HTK decoding phase. When we give test file in decoding phase then HTK will compare the testing sample with the samples that are trained in HTK training phase. If the test file matches then the recognition will success and then the recognition sample will go to the microcontroller.

IV. IMPLEMENTATION

As mentioned the proposed framework, the implementation of designed system is depend on mainly two modules. The first one is speech recognition and another one is hardware interface for controlling Devices. The speech Recognition module consists of different sub-modules such as data collection, data preparation, execution etc. The hardware module provides an interface to get the recognized command that are to be used for device controlling by using a microcontroller, heart of the switching circuit.

A. Data Collection

We have collected three commands as word "LIGHT","FAN","REFRIGERATOR","TV","MOTOR"," ON","OFF", "STATUS") from 200 people out of which approximately 140 recordings are of male and approximately 60 recordings are of female speakers. The data is recorded with the help of unidirectional microphone using a recording tool wave surfer in .wav extension. The .wav files recorded are saved transcription. The sampling rate used for recording is 16 KHz. A labeling tool wave surfer is used to label the speech waveforms. The label files are used in acoustic model generation phase of the system. Following are the command words that we have collected.

Serial No.	Command
1	LIGHT
2	FAN
3	REFRIGERATOR
4	TV
5	MOTOR
6	ON
7	OFF
8	STATUS

Table-1- Required voice command word

B. Phone Set

Phoneme is the basic or the smallest unit of sound in any language. In the phone set that we have used to develop the speech recognition system, the phone set consists of 8 phonemes.

A list of all the phones that are being used and saved as filename.phone is prepared. Here the entries of the phone list will be light fan Refrigerator Tv motor on off status SIL.

C. Hardware Interface

The hardware interface along with a microcontroller, a LCD and a relay driver is responsible to get the recognized command and control the devices according to the voice given by the user. For managing the microcontroller we have designed and compiled a program using embedded c in keil 4.0 compiler. After that we have burned the program into the microcontroller by using flash magic. After training and testing, the recognized voice should be put into the TTY USB port so that our microcontroller can read and access the voice and can handle the devices.
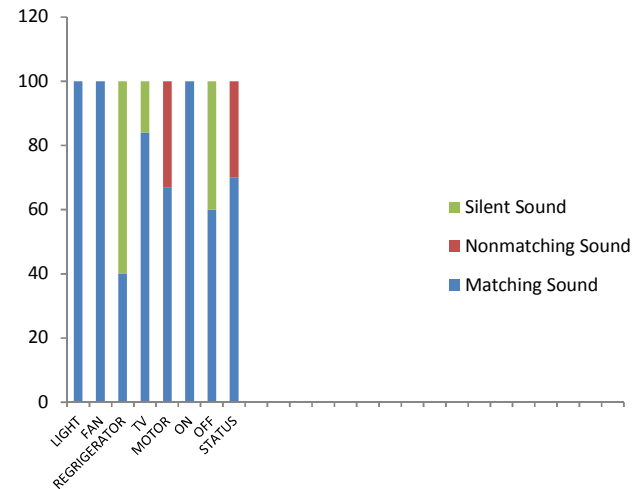
D. Speech Reorganization

Figure 3: Bar diagram of analysis v/s recognition words

In this system, database consists of 8 different words LIGHT, FAN, REFRIGERATOR, TV, MOTOR, ON, OFF and STATUS. Our speech recognition systems consist of total 1000 utterance words taken from different speaker. Including these words we have taken 880 utterance words for training, which are spoken by 200 different users and took them as a trainee in training phase by recognition toolkit. After completion of their training we have tested by new utterance words by the new input different speakers. We took 120 new utterance words from new speakers. After testing phase is completed, we have

compared the training and testing phase. Then we have recognized different kinds of sounds as mention below:

1.   Matching sound: These are the sounds used in the training model which match with the testing sounds.

2.   Non matching sound: These are the sounds used in the training model which do not match with the testing sounds.

3.   Silence sound: These are the sounds used in the training which do not show any outcome.

V. CONCLUSION

The development of a speaker independent spoken query system for accessing the household electronic devices is described in detail in this work. In this system, we use speech as the main communicating media between the machine and the human beings. It has been discovered that there are many people who have a computer phobia. The reasons why many people fear to use speech recognition tools have been due to the inadequate user interfaces. The HTK was used for the implementation of the recognizer. HTK was used because it is open source, more accruable and has been used by many researchers all over the world. A limited grammar and dictionary were constructed to be used by the recognizer. The Speech data was recorded and labeled from 200 different speakers making the training and the testing corpus. We have also explored a set of data to make the system more speaker independent with a gradual improvement of accuracy from more than 50% of present time.

REFERENCES

[1]   Armando Roy Delgado, Rich Picking and Vic Grout, "*Remote-Controlled Home Automation Systems with Different Network Technologies",* Centre for Applied Internet Research (CAIR), University of Wales, NEWI, Wrexham, UK

[2]   L. R. Rabiner, "Applications of Speech Recognition in the Area of Telecommunications,"*IEEE Workshop on Automatic Speech Recognition and Understanding Proceedings*, pp.501–510, 1997.

[3]   S. Goel and M. Bhattacharya, "Speech based dialog query system over asterisk pbx server," in *2nd International Conference on Signal Processing Signal Processing Systems (ICSPS)*, Dalian, Jul 2010.

[4]   Asterisk open source communications, http://www.asterisk.org/home

[5]   Muhammad Ali Mazidi, Janice Gillispie Mazidi, "*The 8051 Microcontroller and Embedded Systems Using Assembly and C* ", Second Edition.

[6]   S. R. Suralkar, Amol C.Wani, Prabhakar V. Mhadse "Speech Recognized Automation System Using Speaker Identification through Wireless Communication", IOSR Journal of Electronics and Communication Engineering (IOSR-JECE), e-ISSN: 2278-2834,p- ISSN:2278-8735.Volume 6, Issue 1
(May. - Jun. 2013), PP 11-18

[7]   Nurul Fadzilah Hasan, Mohd Ruzaimi Mat Rejab and Nurul Hidayah, "Implementation Of Speech Recognition Home Control System Using Arduino", ARPN Journal of Engineering and Applied Sciences, VOL. 10, NO. 23, DECEMBER 2015 ISSN 1819-6608
.

Author's Profile

Dr. Manoj Kumar Deka, received the M.Sc. degree in Instrumentation in the year 2007 and Ph.D. degree in Signal Processing in the year 2011 from Gauhati University, Assam. He achieved INSPIRE Fellowship sponsored by Department of Science & Technology (DST), Govt. of India, New Delhi during his Ph.D. work. Presently he is working as Assistant Professor in the Department of Computer Science & Technology, Bodoland University, Assam. His area of interest includes Signal Processing, Embedded system, Artificial Intelligence.

Mr. Hirakjyoti Sarma, received the B. Tech. degree in Computer Science and Engineering in the year 2011 and M.Tech degree in Information Technology in the year 2013 from Gauhati University, Assam. He has been prusuing Ph.D. in the field of embedded system since 2013 in the department of Computer science and Technology of Bodoland University, Kokrajhar, Assam. His area of interest includes Embedded system, Artificial Intelligence and Operting System.