

# Home Automation System Based on Speech Recognition

Neha A. Wahile<sup>1</sup>, Priyanka D. Hatwar<sup>2</sup>, Isha M. Padiya<sup>3</sup>

<sup>1,2</sup>UG, <sup>3</sup>Assistant Professor, Electronics and Telecommunication Engineering Department, DES'COET, Maharashtra, India.

**Abstract** – Old aged or disabled persons who can't walk are most sensitive persons and they must be served in a systematic, quick, sophisticated and efficient manner by very little effort. The problem is that there is no anybody who is always with them for 24 hours. Speech recognition can be used to serve the old aged or disable persons and to give a full control to them so that they may control all the appliances of home. Traditional home automation systems are not cost effective and they are not suitable for aging populations or disable persons. This paper presents an effective method to overcome these problems. We have designed and implemented a low-cost, reliable, efficient and secure speech operated system for home appliances especially for persons with disabilities to do their work at home. This system is both software and hardware designed using MATLAB R2009a. This system is divided into three main parts namely voice train process, voice recognition process and integration of hardware with MATLAB. This system used speaker dependent method. This proposed design is novel in the way that it is controlling loads by speech recognition using MATLAB to turn on/off loads via parallel port of a computer.

**Index Terms** – Speaker Identification, Speech operated system, Home Automation system, Home Appliances, Aging populations, Speech Recognition, MATLAB Coding.

## 1. INTRODUCTION

Speech Recognition Systems have become so advanced and main stream that business and health care professionals are turning to speech recognition solutions for everything from providing telephone support to writing medical reports.

In many homes there are many people who are old aged or disabled and they can't walk. And there is no anybody who is always with them for 24 hours. There are people who look after them in periodic intervals. The problem is that when a person visits them then it is might not necessary that they needs them but the old aged or disabled person may need a person when he/she is not present with them. Hence home automation systems play a crucial role for elderly or disable persons, so that they can feel comfortable, independent and secure. Development of automation systems using speaker identification began in the 1960s with exploration into speech analysis using text matching, where characteristics of an individual's voice were thought to be able to characterize the uniqueness of an individual much like a fingerprint. The early systems had many flaws and research ensued to derive a more

reliable method of predicting the correlation between two sets of speech utterances. The home evolutionary developments time from the era in which man became sedentary to stop living inside caves and start building their homes. These evolutionary trends of homes automation are focused on several main issues such as security, culture, leisure, comfort, energy savings, management and economic activities. Over the years much work has been done in the domain of automatic speech recognition for automation systems.

The objective of voice recognition is to determine which speaker is present based on the individual's utterance. Several techniques have been proposed for reducing the mismatch between the testing and training environments. The performance of the speech recognition system is given in terms of error rate as measured for a specified technology. For example, one may be carrying groceries into a house and is unable to manually activate the light switch, consequently, if the lighting system in the house has voice activated technology there in, the person may simply say, for example, "lights on" to activate the lights.

Speech recognition is the process by which a computer (or other type of machine) identifies spoken words. Basically, it means talking to your computer, and having it correctly recognize what you are saying. This is the key to any speech related application. There are a number ways to do this but the basic principle is to somehow extract certain key features from the uttered speech and then treat those features as the key to recognizing the word when it is uttered again.

In this paper, a low cost, reliable, efficient and secure Speaker identification based home automation system is presented which utilizes the use of biometric method such as human voice as a directive to activate any electrical appliances. This objective makes the human's voice as an input to the system and this system is speaker dependent that mean only the real or trained user and right command can activate the appliances. This produces and improves the security level of the system.

## 2. OBEJECTIVES AND GOALS

The main objective of this research is to develop a Speaker Identification based automation system capable of controlling many devices inside a home and office using speech commands

with security of voice command of respective user only with speech commands transmitted and received wirelessly.

### 3. LITERATURE REVIEW

Several techniques and methods are available for Home automation system. The common methods are given as:

#### i) Home Automation System using GSM Technology

Home Automation Systems are mostly developed by using microcontroller as a central controlling unit. The Central Control Unit is the hub and brain of a home automation system. We consider three options for communication with GSM, namely SMS based, GPRS based and DTMF based Home Automation systems. Home appliance control system provides security on detection of intrusion via SMS using GSM technology. In this system, user sends SMS from mobile phone to the GSM module connected with Microcontroller and on the basis of SMS various appliances in the home are turned on/off. This system provides mobility to user so that user can turn on/off appliances from anywhere in the world. However it is not possible to implement this system where the user is old aged or disabled with illness due to the main two reasons.

The first main reason is that to use this system a user must know the use of mobile for sending SMS generally old aged person don't know much about creating and sending SMS and second is providing mobile phone to each old aged or disabled person is not cost effective. GPRS based technology uses a webcam to stream video and pictures of the home to its owner's mobile through GPRS. In GPRS based Home Automation system user has to monitor his/her phone constantly to successfully defend against intrusion detection. In DTMF based Home Automation system user calls a SIM number assigned to the home and presses the digits on their phone's keypad to control the home's devices by generating a DTMF tone.

The tone is received and decoded by the GSM module at home using a DTMF decoder. The decoded instructions are passed to the microcontroller so that user commands can be implemented at home. DTMF-based home security systems also have their security flaws. They are vulnerable to —fuzzing attacks, as described by R.Sasi. This may cause whole home network to crash.

#### ii) Home Automation System Via Gesture Recognition System

Traditional input systems for interaction with machines include keyboards, joystick or the mouse. Those suffering from physical handicaps such as Carpel Tunnel Syndrome, Rheumatoid Arthritis or Quadriplegia may be unable to use such forms of input. In that case Gesture recognition is used for Home Automation. Gesture recognition is not based on voice commands but, rather, allows a device to recognize certain

gestures. This approach does not require any technical knowledge (like in SMS based automation system). Old aged or disabled will use his/her hand to control appliances. By using a simple webcam the images will be taken and will be processed at Laptop in MATLAB software and once a particular gesture is recognized then the corresponding action will be performed.

Although it is a sophisticated solution but when Old aged or disabled person is not able to move hand and when they can only shake hand then hand detection may not accurately detected and the chance of false alarm is more in this approach and mostly the hand gesture recognition is done by detecting the human skin color and so because of this the background of the hand must be a non-skin color background with fixed distance between hand and the camera. Moreover for the smooth working of system there must be a proper arrangement of lighting always. Gesture recognition system can be used in various applications like Virtual reality, games and sign language. Sign language is an important case of communicative gestures. Sign language for the deaf (e.g. American Sign Language) is an example that has received significant attention in the gesture literature.

#### iii) Home Automation System using Bluetooth, WIFI, WSN and Zigbee Technologies

Many Wireless Technologies like RF, Wi-Fi, Bluetooth and Zigbee have been developed and remote monitoring systems using these technologies are popular due to flexibility, low operating charges, etc. Bluetooth looks like an attractive communication technology for creating smart homes. It is cheap, easy, and quick to set up. People are already familiar with the technology; however Bluetooth communication should only be used on occasions where there is a need for quick short-lived network communication with little concern for security.

Limitations include, they have maximum communication range of 100m in ideal conditions, and it has high power consumption. It has serious security concerns such as eavesdropping and weak encryption as discussed by M.Ryan. Other wireless technologies like WIFI, WSN and Zigbee have very high developing and deployment cost due to needs of motes, sensors, and radio transceivers etc., spread over a large area. Further it is difficult to upgrade existing conventional control system with remote control capabilities. Moreover they are commonly used by mobile users, who want to monitor and control their home appliances remotely; hence these technologies are not suitable for aging populations.

### 4. SYSTEM OVERVIEW

The block diagram of Home Automation system based on speech recognition is shown in figure

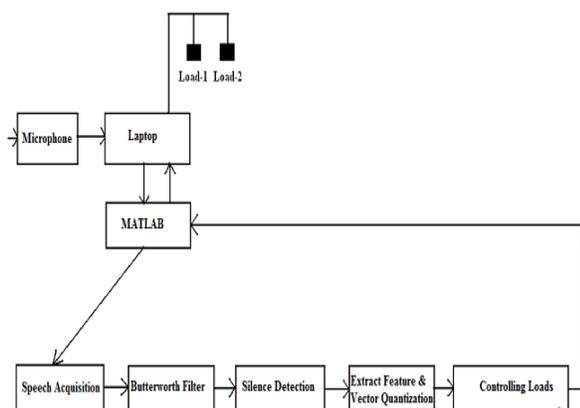


Fig.1: Block Diagram of Speech Operated System for Home Appliances

Speaker Identification based Home Automation system using speech recognition is a low-cost, reliable, efficient and secure method for Home Automation System. This report is divided into two main parts which are Voice Training Process and Voice Testing Process. In voice training process the first step is acquisition of speech. Robust training in which several versions of the sound pattern are used to create a single merged template or statistical model. In voice testing process the user has uttered two different words each process. The system used is speaker dependent method that means user has to record his/her voice before using the system. Various steps involved in Speaker Identification based Home Automation system using speech recognition is shown in figure 1.

#### Voice Training Process:

In voice training process the first step is acquisition of speech. Built in Microphone in laptop is utilized for Speech Acquisition, and then speech acquisition device is installed by simply Connecting the Microphone with laptop via sound card input port. In second step a function is created, which will record speech in MATLAB. In third step recorded speech is played on laptop based audio output device. Fourth step is to write acquired speech in MATLAB and .Wav file is created. In fifth step .wav file is loaded in MATLAB, in order to read the saved speech and in sixth step saved speech is acquired. In seventh step it is filtered out through the Butterworth band pass filter. Butterworth filter is used because it is the best compromise between attenuation and phase response. It has got no ripple in the pass band or the stop band. After that it is saved in the computer memory so that it can be matched with incoming utterance of speech. In this research work user has uttered two training voices to control the load. These uttered words are “CLOSE” and “YES”. Now all above steps are applied to these uttered words. Silence detection or Voice Activity Detection (VAD) is used in speech processing, which is used to detect presence or absence of human speech. VAD is

used here to deactivate some processes when there is a silence or non-speech section in audio session. Short time Fourier transforms is performed successfully so that for each incoming speech, the part of containing high frequency component is extracted. Actually here in MATLAB coding 2500 samples per word are created for feature extraction.

#### Voice Testing Process:

In voice testing process the user has uttered two different words each process. One word is same as which was trained in training phase was “CLOSE” and other one is “OPEN”. Then both uttered signals are further processed and analysed by applying same steps which are already used in Voice Training Process. Like voice training process, 2500 samples per word are also created here for feature extraction. These testing signals are used to match with trained signals to authenticate the desired speech. There are various feature matching techniques used in MATLAB, from which Vector Quantization method is used in this research paper. Vector Quantization is a process of mapping vectors from a big vector space to a finite number of regions in that space. In the testing phase, a speaker specific Vector Quantization codebook is generated for each known speaker by clustering his/her testing acoustic vectors.

#### Speech Recognition System:

A speech recognition roughly consists of two portions. They are speech analysis and pattern recognition.

##### ➤ Speech Analysis:

The purpose of the speech analysis block is to transform the speech waveform into a parsimonious representation which characterizes the time varying properties of the speech. The speech analysis typically includes two modules, namely data acquisition and feature extraction. The data acquisition module usually contains a microphone and a code from which digitized speech data are generated. The feature extraction is the computation of a sequence of feature vectors which provides a compact representation of the given speech signal. The feature extraction is done on short-time basis. The speech signal is separated into overlapped fixed-length frames. From each frame, a set of frequency-domain or cepstral-domain parameters are derived from each frame, to form the so-called feature vector. There are some basic principles and analysis techniques used in the feature extraction module. They are pre-emphasis, frame blocking and windowing, Discrete Fourier Transform (DFT) computation, spectral magnitudes, Mel-frequency filter bank, logarithm of filter energies, Discrete Cosine Transformation (DCT), Cepstral Weighting, and dynamic featurings.

##### ➤ Pattern Recognition:

The speech signal is first analyses and a feature representation is obtained for comparison with either stored reference

templates or statistical models in the pattern matching block. A decision scheme determines the word or phonetic class of the unknown speech based on the matching scores with respect to the stored reference patterns. There are two types of reference patterns. The first type, called a nonparametric reference pattern (or often a template), is a pattern created from one or more spoken tokens of the sound associated with the pattern. The second type, called a statistical reference model, is created as a statistical characterization of the behavior of a collection of tokens of the sound associated with the pattern. The vector quantization model is used as the statistical model. There are three portions in pattern recognition. They are pattern training, pattern matching and maximum selection.

### i) Pattern Training:

Pattern training is the method by which representative sound patterns are converted into reference patterns for use by the pattern matching algorithm. There are several ways in which pattern training can be performed, including: Casual training in which a single sound pattern is used directly to create either a template or a crude statistical model. Robust training in which several versions of the sound pattern are used to create a single merged template or statistical model. Clustering training in which a large number of versions of the sound pattern is used to create one or more templates or a reliable statistical model of the sound pattern.

### ii) Pattern Matching:

Pattern matching refers to the process of assessing the similarity between two speech patterns, one of which represents the unknown speech and one of which represents the reference pattern (derived from the training process) of each element that can be recognized. When the reference pattern is a “typical” utterance template, pattern matching produces a gross similarity (or dissimilarity) score. When the reference pattern consists of a probabilistic model, the process of pattern matching is equivalent to using the statistical knowledge contained in the probabilistic model to assess the likelihood of the speech (which led to the model) being realized as the unknown pattern. Pattern matching refers to the process of assessing the similarity between two speech patterns, one of which represents the unknown speech and one of which represents the reference pattern (derived from the training process) of each element that can be recognized. Firstly, the training set of vectors is used to create the optimal set of codebook vectors for representing the spectral variability observed in the training set. And Then distance is measured between a pair of spectral analysis vectors to able to cluster the training set vectors as well as to classify spectral vectors into unique codebook entries. The next step is a centroid computation procedure. Finally, a classification procedure selects the codebook vectors that closet to the input vector and uses the codebook index as the resulting spectral representation. The classification procedure is essentially a

quantizer. It accepts speech spectral vectors as input and provides the code index of the code vectors that best matches the input.

### iii) Electronic Control System

In this diagram, Speech instruction is firstly taken as input to control home appliances and then a microphone is used to record the person speech. Secondly, the speech instruction is caught and transferred the analog signal to digital signal and the recorded speech is sent to the speech based verification/identification system. Thirdly, the digital information of speech instruction is processed and compared by using the MATLAB programming.

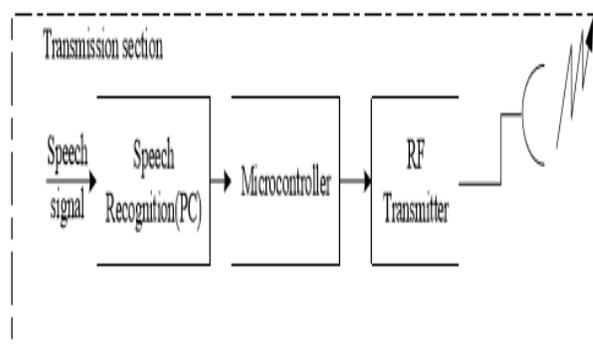


Fig. 3: Block Diagram of Transmission Section of Home Appliances Control System for Speech Recognition

The speech signal is first analyzed and a feature representation is obtained for comparison with either stored reference templates or statistical models in the pattern matching block. Speech Recognition is a technology allowing the computer to identify and understand words spoken by a person using a microphone. Speech Recognition is a technology allowing the computer to identify and understand words spoken by a person using a microphone. Then signal goes to microcontroller unit then the signal is transmitted.

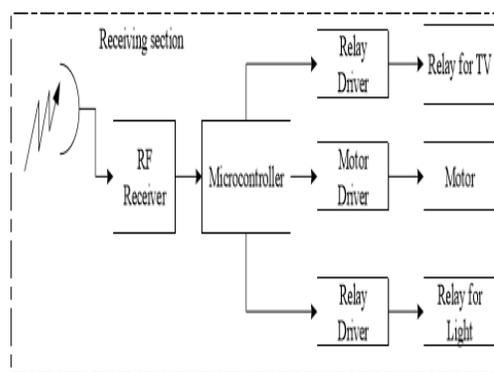


Fig. 4: Block Diagram of Receiving Section Home Appliances Control System for Speech Recognition

In the receiver section receiver accept radio signal and then microcontroller read the signal and then send to drives relay and motor driver. Pulse width modulation (PWM) is a method for binary signals generation, which has two signal periods (high and low). The width (W) of each pulse varies between 0 and the period (T). The main principle is control of power by varying the duty cycle. Here the conduction time to the load is controlled. The duty cycle can be varied from 0 to 1 by varying ton or T. Therefore, the average output voltage  $V_{avr}$  can be changed between 0 and  $V_{in}$  by controlling the duty cycle, thus, the power flow can be controlled. The on-off switching is performed by power MOSFETs.

➤ **Control**

The transmitted control characters are received by the microcontroller and compared with some predefined characters. If there is a match, the microcontroller will switch the corresponding relay and turn on/off the appliance connected to it. On the control side, the microcontroller has to be programmed to be able to receive control characters from the receiver and activate/control the required relays accordingly.

**Advantages**

- Speech is a very natural way to interact & it is not necessary to sit at a keyboard or work with a remote control.
- No training required for users.
- Beneficial for aging population.

**Disadvantages**

- Even the best speech recognition system must make errors. If there is noise of some other sound in the room (e.g. Television), the no. of errors will increase.
- Speech recognition work best if the microphone is close to the user (e.g.in a phone or if the user is wearing a microphone).More distance microphones (e.g. on a table or wall) will tend to increase no. of errors.

**Applications**

- Speech to text processing (word processors or emails)
- Optimizing use of low cost electricity.
- Can be used in all electrical appliances.

5. CIRCUIT DETAILS

Transmission Section: In the transmission section, there are KS232 module, PIC 16F887 and KST-TX01 (Radio Frequency transmitter module). The KS232 module is used to carry the signal from PC to Microcontroller unit. The signal is retransmitted with baud rate 1200 for RF transmission by KST-TX01 module. This module has four pins: supply pin,

GRN pin, and ANT pin. KST-TX01 technical specific data for wireless transmitter module are:

- (1) Transmit power: 1W,
- (2) Operating frequency: 315MHZ~433.92MHZ,
- (3) Operating temperature:-40°C~80°C,
- (4) Operating voltage: 3V~5V and

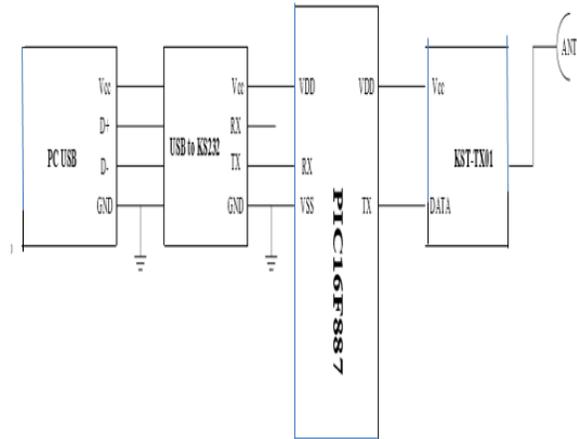


Fig. 5: Circuit Diagram of Transmission Section

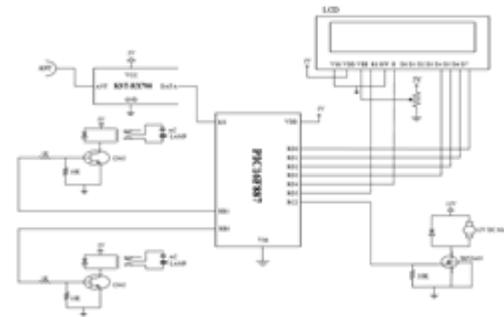


Fig. 6: Circuit Diagram of Receiving Section

Receiving sections: The receiver section consists of KST-RX706 (RF receiver module), PIC microcontroller, relays, relay drivers and motor driver. In this section, KST-RX706 firstly accept radio signal and then microcontroller read radio signal with baud rate 1200. Microcontroller drives relay and motor driver. The speed of motor is controlled by using Pulse Wide Modulation (PWM) module. KST-RX706 firstly accept radio signal and then microcontroller read radio signal with baud rate 1200. Microcontroller drives relay and motor driver. The speed of motor is controlled by using Pulse Wide Modulation.

6. SOFTWARE DETAILS

MATLAB Millions of engineers and scientist worldwide use MATLAB® to analyse and design the systems and products

transforming our world. MATLAB is in automobile active safety systems, interplanetary spacecraft, health monitoring devices, smart power grids, and LTE cellular networks. It is used for machine learning, signal processing, image processing, computer vision, communications, computational finance, control design, robotics, and much more. MATLAB is the easiest and most productive software for engineers and scientists. MATLAB window, whether you're analysing data, developing algorithms, or creating models, MATLAB provides an environment that invites exploration and discovery. It combines a high-level language with a desktop environment tuned for iterative engineering and scientific workflows. The desktop environment invites experimentation, exploration, and discovery. These MATLAB tools and capabilities are all rigorously tested and designed to work together.

**Key Features of MATLAB:**

High-level language for scientific and engineering computing.

- Desktop environment tuned for iterative exploration, design, and problem solving.
- Graphics for visualizing data and tools for creating custom plots.
- Apps for curve fitting, data classification, signal analysis, and many other domain-specific tasks.
- Add-on toolboxes for a wide range of engineering and scientific applications.
- Tools for building applications with custom user interfaces.
- Interfaces to C/C++, Java®, .NET, Python®, SQL, Hadoop®, and Microsoft® Excel®.
- Royalty-free deployment options for sharing MATLAB programs with end users.

7. SOFTWARE IMPLEMENTATION

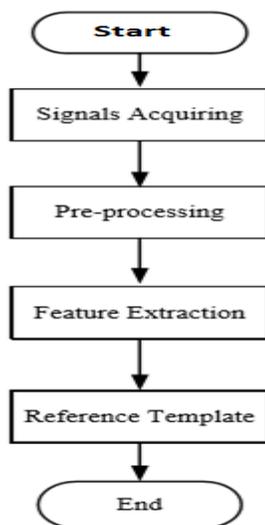


Fig.7: Flow Chart of Training phase

The software implementation part of voice recognition based home automation system implemented using the Arduino controller. It consists of training of voice recognition module. The voice recognition module needs to be trained first with the voice commands before it can be put to recognizing function. This section explains the methods used for speech recognition. These methods are training phase and testing phase.

Initially, the user must prepare the training files. Figure.4.4 shows the flow chart of the step of training phase. In that signal acquiring, pre-processing, features extraction, reference template, these steps are include. The speech files are recorded from the microphone and MFCC features are extracted from the input file. Then these features are stored. In this case, the collection of training files is called database. Then, the user must train the system using the files in the database. This is called training phase or pre-processing.

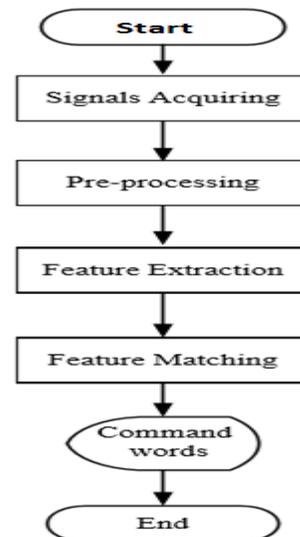
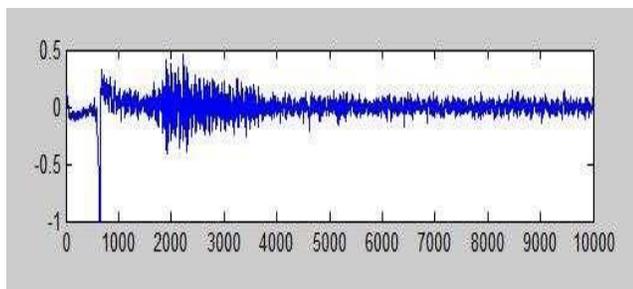


Fig. 8: Flow Chart of Testing Phase

In the testing phase, users have to provide the command words as input. In this case, user may use two ways of testing. If user chooses to use the pre-recorded sound file, one of the samples are loaded from test files and read. Then, the modified MFCC features are extracted from the input file. In the next step, the distances between the modified MFCC features and the stored reference models are calculated using Euclidean Distance. Finally, the minimum distance is selected among the distances between the input vectors and codebook vectors. If this minimum distance falls below the local threshold, the system outputs the command word as result. Otherwise, the system determines it is wrong command word. If the user wants to test the system with spoken commands in real time, the sound file to be recognized is recorded from the microphone. To do so, the user must choose time length typical time length is 2 seconds. In this system, sound files are recorded within this

time length. Then the subsequence processes, as above, are carried out and recognition decision.

## 8. RESULT



Time vs. Amplitude

Fig. 9: Train Signal Uttered as “CLOSE”

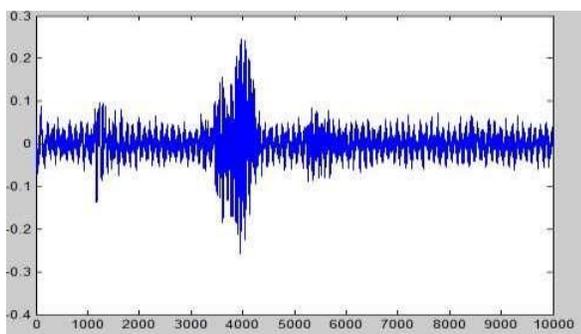


Fig. 10: Train Signal Uttered as “OPEN”

## 9. CONCLUSION

The designed Speech operated system is a low-cost, reliable, efficient and secure. The designed Speech Operated system can be used in various areas of application. Speech operated system can also be used to answer computers in a hands-free environment, like when driving. Speech operated system can be used in tasks that require human-machine interface, for example automatic call processing in the telephone network and data query information systems. The system has two main parts: speech recognition and smart home appliances electronic

control system. Speech recognition is implemented in MATLAB environment. An application for speech command processing is developed.

## 10. FUTURE WORK

In future, a significant research can be carried out in the area of Speech Operated system for Home Appliances. Speaker Identification based Home Automation system with better efficiency can be developed which will also be operate able in noisy environment. Reliable and efficient Home Automation system can be designed in future which will be both speaker dependant as well as speaker independent with maximum efficiency, security and performance.

## REFERENCES

- [1] B.H. Juang & Lawrence R. Rabiner, Automatic Speech Recognition – “A Brief History of the Technology Development”.
- [2] Coucopoulos, Andre; 2007, “Voice processing for home automation systems”, Network system design line, January 2007.
- [3] Gnanasekar. A.K, Jayavelu.P, Nagarajan.V, “Speech Recognition Based Wireless Automation of Home Loads with Fault Identification”, IEEE International conference on communications and signal processing (ICCSP), Vol. 3, pp.128-132, 2012.
- [4] Jie Liu, Jigui Sun, Shengsheng Wang, “Pattern Recognition: An overview”, IJCSNS International Journal of Computer Science and Network Security, vol. 6, no. 6, June 2006.
- [5] LA Roberts, HT Nguyen, EM Silver, “Gesture Activated Home Appliance”- US Patent 6,937,742, 2005 - Google Patents
- [6] M.A. Anusuya, “Speech Recognition by Machine,” International Journal of Computer Science and Information security, Vol.6, No.3, 2009.
- [7] Mishr and Suyash Agrawal, “Recognition Of Voice Using Mel Cepstral Coefficient & Vector Quantization” International Journal of Engineering Research and Applications (IJERA) Vol. 2, Issue 2, Mar-Apr 2012, pp.933-938
- [8] Prabhakar V. Mhadse<sup>1</sup>, Amol C. Wani<sup>2</sup> “Speaker Identification Based Automation System through Speech Recognition”
- [9] Seema Asht and Rajeshwar Dass, “Pattern Recognition Techniques: A Review”, International Journal of Computer Science and Telecommunications, vol. 3, issue 8, August 2012.
- [10] S.J. Arora and R. Singh, “Automatic Speech Recognition” A Review, “International Journal of Computer Applications”, vol.60-No.9, December 2012
- [11] Tycht and Josef Psutka, “Speech Production Based on the Mel-Frequency Cepstral Coefficients” Mulgrew,