ISSN (Online) 2321–2004
ISSN (Print) 2321–5526

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 3, Issue 6, June 2015

COMPUTER CONTROL WITH VOICE COMMAND USING MATLAB

Yogesh Watile[1], Pratik Ghotkar [2], Bhushan Rohankar[3]

Assistant Professor, Department of Electronics & Telecommunication, DMIETR, Wardha, India[1, 2]

Assistant Professor, Department of Electronics & Telecommunication, GNIET, Nagpur, India[3]

Abstract: This paper describe Computer control with voice command using soft computing tool Matlab. The system enlightens upon the invention as well as technological advancement in the field of voice controlled computer using Matlab. In this paper the concept of voice recognition is use to control the computer. Initially analog voice input will be given to computer using microphone which will be recorded in samples and the MFCC feature is extracted from that voice command and it is stored in the database. The real time input voice command to be executed will be send to the computer that will be compared with stored data base by MFCC algorithm and if the match is found then particular operation is performed. This system will be helpful for physically disabled persons and also for the peoples who needs simple access to their computer.

Keywords: Command Execution, Database Generation, Feature Extraction, Speech Recognition, MFCC.

I. INTRODUCTION

In this designed system, it is aimed to control a computer with speech commands using MATLAB. The idea behind this system for computer control with voice command is taken from the Voice operated Robot using MATLAB [1][2]. If the robot can operated using voice command then a similar system can be developed for controlling a computer using voice command. So we have designed a system for controlling the computer with voice commands. There are some systems that use gesture control system and the tongue motion detection system for operating the system application for the ease of physically handicapped person but this system is much complicated.The main goal of this designed system is to introduce hearing sensor and also the speech synthesis to the computer such that it is capable to interact with human through spoken natural language. The context of speech recognition refers to system where a person can speak via a microphone to a computer. The computer translates the spoken words into either text or commands to execute functions in the computer.

The intelligent speech recognition system enables the computer to understand spoken instructions. The speech recognition system using Matlab is trained in such a way that it recognizes defined commands and the designed computer will operate based on the instruction through the speech commands.

In this system, initially a database is generated by feature extraction of the commands. The database commands are linked to the applications in computer. Whenever the voice command is given to the computer, it is compared with the Database by using MFCC algorithm.

If the command matches with the reference database command then the application linked to that command will be executed. The main focus of designed system is to provide an opportunity to the handicapped to operate computer with voice command with less efforts.

II. SYSTEM OVERVIEW

Like any other pattern recognition systems, this computer control with vice command systems also involves two phases namely, training and testing. Training is the

process of familiarizing the system with the voice characteristics of the speakers registering. Testing is the actual recognition task. The block diagram shows the training as well as the execution phase .Feature is extracted from the voice characteristics of the speaker during the training utterances and is used for building the reference models. During testing, similar feature are extracted from the test utterance, and the degree of their match with the reference is obtained using MFCC matching technique. The level of match is used to arrive at the decision.

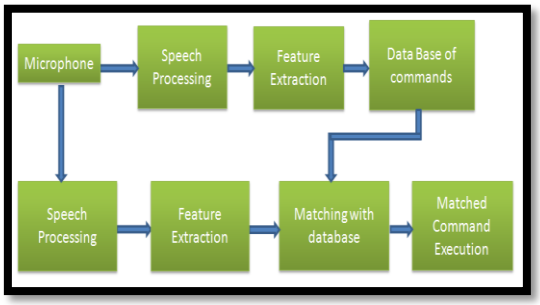

Fig.1: System block diagram

The block diagram for computer control with voice command using Matlab is shown in fig.1. It consist of microphone, speech processing (speech recorder, analog to digital convertor), feature extraction, data base commands, matching with database to compare the command in the data base with input voice and matched command execution

A. Microphone

Microphone is used as a transducer which convert sound signal into electrical signal. Most microphone today use

ISSN (Online) 2321–2004
ISSN (Print) 2321–5526
INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 3, Issue 6, June 2015

electromagnetic induction capacitance change, piezoelectricity to produce an electrical signal from air pressure variations. Microphones typically need to be connected to a preamplifier before the signal can be amplified with an audio power amplifier or recorder. When the diaphragm in the microphone vibrates, it causes other components in the microphone to vibrate. These vibrations are converted into an electrical current which becomes the audio signal.

B. Speech Processing

Audio Signals received from the microphones is sampled at the rate of 16 KHz for discritization.

C. Feature Extraction

For Speech recognitions, the most commonly used acoustic feature is Mel-Scale Frequency Cepstral Coefficient (MFCC). This technique is often used to create the fingerprint of the sound files. The MFCC are based on the known variation of the human ear's critical bandwidth frequencies with filters spaced linearly at low frequencies and logarithmically at high frequencies used to capture the important characteristics of speech. Studies have shown that human perception of the frequency contents of sounds for speech signals does not follow a linear scale. Thus for each tone with an actual frequency (f) measured in Hz, a subjective pitch is measured on a scale called the Mel scale. The Mel-frequency scale is linear frequency spacing below 1000 Hz and a logarithmic spacing above 1000 Hz. As a reference point, the pitch of a 1 kHz tone, 40 dB above the perceptual hearing threshold, is defined as 1000 Mels. MFCC Extraction is shown in figure below.

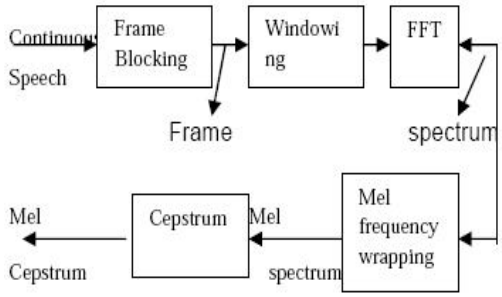

Fig.2: Block Diagram MFCC

The procedure for extracting Mel-frequency Cepstrum Coefficient is:

a) Take the Fourier transform of the signal, (done for each window)
b) Map the log amplitudes of the spectrum onto the Mel scale. Triangular overlapping windows are used.
c) Mel scale: To convert between f hertz into m mel:
d) m= 1127.01048 $\log_e$ (1+f/700)
e) For m mel into f hertz:
f) f=700 (exp (m/1127.01048) – 1)
g) Take the Discrete Cosine Transform of the list of Mel log-amplitudes
h) The amplitudes of the resulting spectrum are the MFCCs.

D. Database of Command

In database we store command in speech commands like open, close, move, play. When user gives command it is automatically compared in database. Initially voice input will give with the help of microphone; this voice will be in the form of analog voice signal. The voice input will be recorded as analog samples in the Matlab. The analog samples are then converted into digital data. This data will be stored in the computer as a digital fingerprint which can be used for the reference of the command. In this system the speech command are stored in .wav format. The location of file or application that is to be open is link with that command.

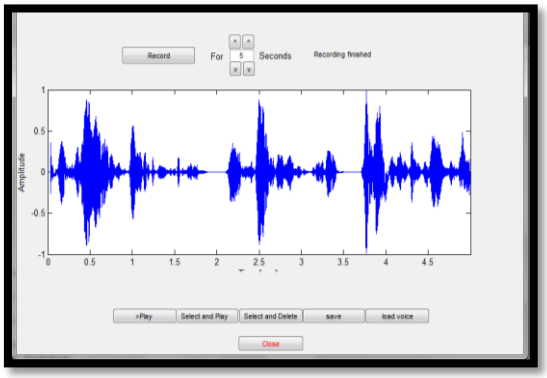

Fig. 3: Guide for generation of database and sound editing

Fig.3 shows the guide for generation of data base and sound editing. Guide "gui.fig" is the function in Matlab which provide the window to record and store the voice commands in the database. It is also use to edit the voice command files stored in the database. Each entry in the database is designed to correspond to a single command utterance. The entries in the database were designed to be a feature matrix containing features extracted from the pre-recorded samples. There was one entry for all the pre-recorded commands. The entries in the database correspond to each command. The entries are in form of feature matrices extracted from the vocal commands. The formation of the feature matrices is summarized in the 'Feature Extraction'.

E. Matching & Command Execution

In database all command with their features are stored for the matching process in testing phase. When the command is given to through the microphone, then given command is compared with the command stored in database by "strcmp" function. Each command in the database is link with the particular application that is to be executed in computer. Do operation class is used to link the command stored in database with the application. The command that matches with the reference database commands perfectly that command will be executed. The fig.4 shows the window to start the speech recognition process.

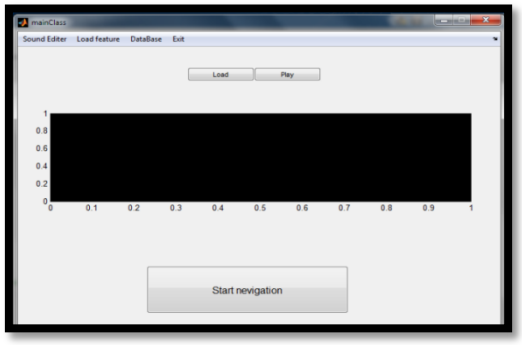

Fig.4: Start Voice Recognition

ISSN (Online) 2321–2004
ISSN (Print) 2321–5526
INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 3, Issue 6, June 2015

III. WORKING

This computer control with vice command systems also involves two phases namely, training and testing. Training is the process of familiarizing the system with the voice characteristics of the speakers registering. Testing is the actual recognition task. The block diagram of training phase is shown in Fig.5. Feature vectors representing the voice characteristics of the speaker are extracted from the training utterances and are used for building the reference models. During testing, similar feature vectors are extracted from the test utterance, and the degree of their match with the reference is obtained using some matching technique. The level of match is used to arrive at the decision.

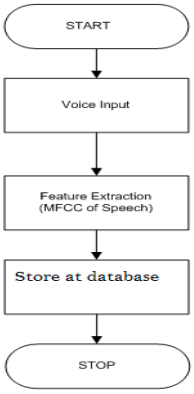

Fig.5: flow diagram of Training phase

A. Training phase

The training phase of the system is also called as database generation phase of the system. The flow diagram for the training phase is shown in fig.5. Training phase involve input speech signal through microphone, speech processing, feature extraction & extracted feature will stored in database.

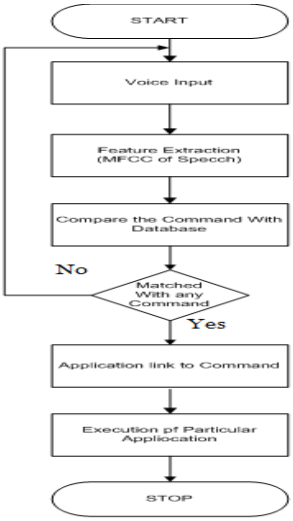

fig.6: flow diagram of testing phase

B. Testing Phase

The testing phase of the system is also called as execution phase of the system. The flow diagram for the testing phase is shown in fig.6. The testing phase includes the real time input speech signal through the microphone, speech processing, feature extraction, matching with database and execution of matched command. When real time speech signal given to the system, MFCC features can be derived from the applied speech signal. If extracted feature matched with feature stored in data base for any particular command, then particular application related to this command will be opened.

IV. RESULT

Fig.7: Snap shot of power point application

The computer file or application can be open using a voice command with this system. When the command to open power point presentation is given to through the microphone, given command is compared with the database command by "strcmp" function. If the match is found then "do operation" function will execute the matched application from the data base, then the power point application will be opened.

In this system we have implemented the application calling for music, word, paint and folder opening. Further it can be implemented for the operating all computer applications.

V. CONCLUSION

Thus we have implemented the voice recognition system for the operating the computer using MATLAB. It can be used by handicapped persons to operate the computer system in easier manner.

REFERENCES

[1]. Aseem Saxsena, Amit K. Sinha, Shashank Cakrawarti, Surabhi Charu, "Speech Recognition Using MFCC and Matlab", Nov-2013, Suresh Gyan Vihar University, Jaipur, Rajasthan, India, International Journal of Advances In Computer Science and Cloud Computing.

[2]. Shraddha D. Gosavi, Uday Pandit Khot, Seema Shah, "Speech Recognition for Robotic Control", Sep-Oct 2013, Department of Electronics & Telecommunication Engineering, Vidyalankar Institute of Technology, Mumbai, Int. Journal of Engineering Research and Application, Vol. 3, Issue 5.

[3]. Suraj Jadhav, Shashank Kava, Sanket Khandare, Sayali Marawar, Savitha Upadhya, "Voice Activated Calculator", Department of Electronics & Telecommunications Engineering, February 2013, Fr.C.Rodrigues Institute of Technology, Vashi, India, International Journal of Emerging Technology and Advanced Engineering.

[4]. [Ms.S.D.Suryawanshi, Mr.J.S.Chithode , Ms.S.S.Pethkar ,Electronics department, Bharti Vidyapith University , "Voice operated intelligent wheel chair" (May-2013),International journal of advance research in computer science & software engineering.

[5]. David Beckstrom, Harun AliKunal, Jadav Nicole, Su Lee, Zorawar Bhatia, "Speech Actvated Appliacnces", Spring 2007, University of Victoria, Faculty of Engineering.