

Voice Recognition System Through MATLAB

¹ Juhi Khichi, ²Monica lamba

¹B.Tech Scholar, ²Assistant Professor

¹juhikhichi@gmail.com, ²lambamonica346@gmail.com

Abstract

As the technology is expanding its roots in multimedia, the peoples are more concerning about the security and speed of their work. By viewing this, the need of voice recognition system is increasing. So, this has become a dominant research topic. It includes training and testing of voices. Sampling, feature extraction, database creation, classifying all these steps are used to recognize a voice. Voice recognition are of two types. First is speaker dependent which require training of software, by reading some pages of text. Second one is speaker independent in which there is no requirement of software training. This can be done in time domain by correlation but, it will lack accuracy and speed, so fast Fourier transform is used to find its features. Although, there are many limitations of it for e.g. variation in person's voice with time, crowded rooms etc.

Keywords – VRS, FFT, ASR.

I. INTRODUCTION

Voice recognition technology is a process of extracting the voice characteristic information from people's voice. Operated through the computer and to recognize the content of voice of a person, to recognize who is speaking. It is also known as automatic speech recognition. The voice characteristic parameters of different people are almost different, such as the loudness, voice amplitude, all of them are different. It's

interdisciplinary involving many fields, where modern speech recognition technology consist of many domains of technology, such as signal processing, theory of information, phonetics, linguistics, artificial intelligence, etc. Using the voice recognition system not only improves the efficiency of the daily life, but also makes people's life more diversified.[4]

2. Types of Voice Recognition

Speaker dependent software is commonly used for dictation software, while speaker independent software more commonly found in telephone application. Speaker dependent software work by learning the unique characteristics of a single person's voice. New users must first train the software by speaking to it. Speaker independent software is designed to recognize anyone's voice, so no training is required. It means that it is only real option for application involving such as interactive voice response system. Where business do not ask person to read pages of text before using the system. The downside is that Speaker independent software is less accurate than Speaker dependent software. [2]

3. Voice Signal Acquiring

Signal acquisition is the first step for the analysis of speech.

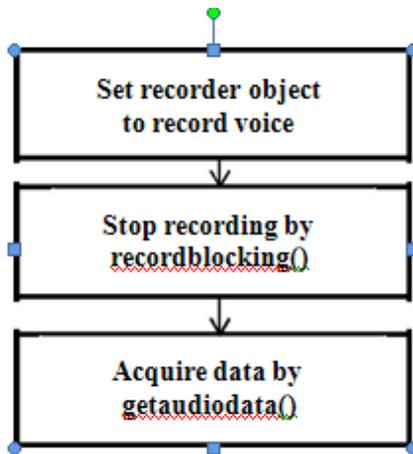


Fig.1 Flow chart for voice recording

4. Feature Extraction

Before the recognition of speech, characteristic parameters of the input speech signal is need to be extracted. The purpose of characteristic parameters extraction is to analyze speech signal processing and removes the redundant information which has nothing to do with speech recognition and obtain the important information. One of the most important parameter of sound is its frequency. The sounds are discriminated from each other by the help of their frequencies.[1]

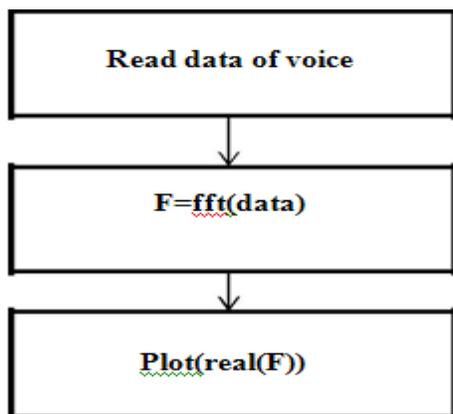


Fig.2 Flow chart for FFT

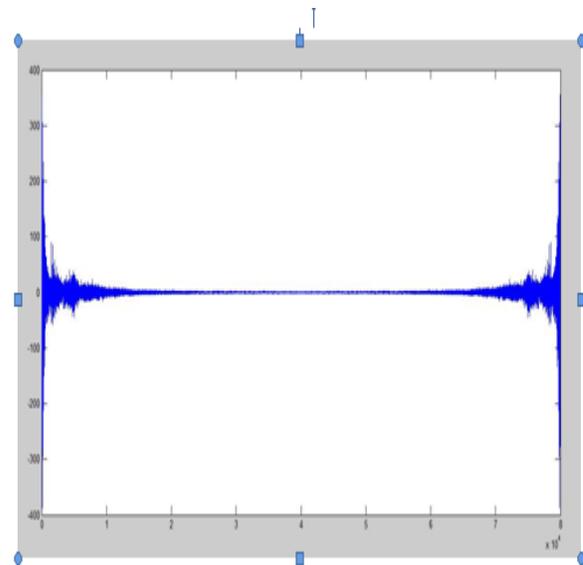


Fig.3 Plot of FFT

By using above FFT we can find maximum pitch of a person by using, $m=\max(\text{real}(F))$; code in MATLAB.[5]

5. Testing

During the testing phase, the input speech is matched with stored reference model and recognition decision is made. Speech recognition is a difficult task and it is still an active research area. Automatic speech recognition works based on the premise that a person’s speech exhibits characteristics that are unique to the speaker. In it the extracted features acquired at the time of training is matched with new recorded voice’s feature.

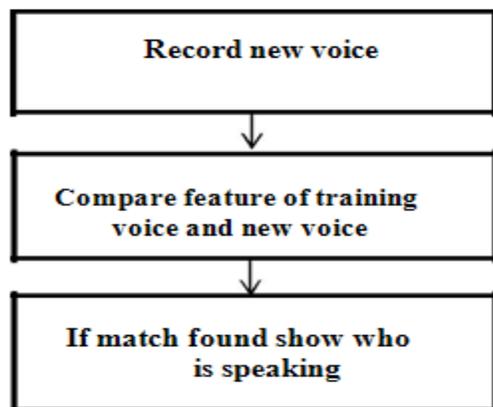


Fig. 4 Flow chart of testing

6. Future Scope

It can be used for dictation, voice dialing, for security of any gadget. It can also be used for navigation and system control. industries use these Voice Recognition Tools for Corporate Customer Relations, Training, Design, Project Management, Public Relations, Advertising, Word Processing, Data Mining, Writing, Translation, Recording, Machine Interface, and that is to just name a few. These applications in Voice Recognition have improved efficiency, saved time and that translates, at least to the corporations that employ the technology, into quarterly profits and improved shareholder's equity, its all been very well received. Transportation, Communication, Energy, Education, Military, Mining, Manufacturing, Policing,

7. Conclusion

Voice recognition is a very vast and difficult field. There are many methods to achieve recognition. One of them is used in this research.

The main role in this is of FFT, this is used to extract feature i.e. pitch. Pitch is one of the characteristic of human voice, which differentiate human's voice from each other. However, there are some limitations of it but, it is the most popular hotspot for present research. Scientists are researching the way to record voice without noise. So, in near future there can be a system with best quality.

8. References

- [1] "M. Bahoura och J. Rouat", A new approach for wavelet speech enhancement. INTERSPEECH, pp. 1937-1940, 2001.
- [2] "M.A. Anusuya and S.K. Katti", Speech Recognition by Machine: A Review", (IJCSIS) International Journal of Computer Science and Information Security, vol. 6, no. 3, pp.181-205, 2009.
- [3]"Preeti Saini, Parneet Kaur", Automatic Speech Recognition: A Review, International Journal of Engineering Trends and Technology- Volume4Issue2-2013
- [4] "Linlin Pan", Dec 2013, Research and simulation on speech recognition by MATLAB
- [5] Topic on "Extraction of Pitch and Formants and its Analysis to identify 3 different emotional states of a person"ijcsi.org/papers/IJCSI-9-4-1-296-299.pdf