INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 6.078
(Volume 6, Issue 6)
Available online at: https://www.ijariit.com

Speech Recognition for Crime Analysis

Mehak Sharma
meh7891ak@gmail.com
Vellore Institute of Technology, Vellore, Tamil Nadu

ABSTRACT

Human-Computer Interaction (HCI) is a multidisciplinary field of study focusinology design. We aim to take this further in the field of crime analysis by developing a system for speech recognition. This will be a huge contribution to the crime department where speech recognition will not only help them to handle large amount of data but also in checking the criminal records faster. This is a modern and advanced technique where with the help of voice recognition test file will be compared with database files.

Keywords: *Speech Recognition, Correlation, Noise*

1. INTRODUCTION

Speech is the primary abilities of human being to exchange information with others. Communications in a community are dominated by speech. Because of this, Interactive of speech with the modern technology are require in this current civilization. It is a high demand to implement speech or voice system in the society. Since 60, many researchers have been done to make the computer able to interpret and understand human language. However speech is a difficult task for digital devices to recognize human voice. Study on technical specifications of sound is requiring in understanding the characteristic of human voice.

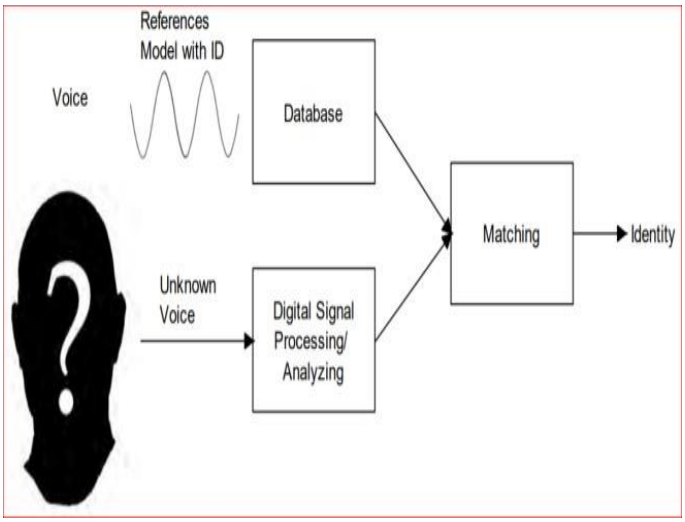

Fig. 1: Typical Voice Recognition Process

Through this, digital signal processing method is require filtering and analyzing the voice signal. Speech recognition is applicable in the real world such as multimedia web portal, identity verification and forensic analysis. In forensic analysis, there are several recognition technique used to identify the suspect contribute in criminal cases such 2 as biometric analysis which include fingerprint and voiceprint. By the analysis of a recorded message obtain in a criminal cases, the possible criminal could be identify. As the result it helps legal authority to enforce the law, protect property, and limit civil disorder.

Figure 1 below show the basic idea for speech recognition, where the unknown voice was compare to the references model. The highest matching score obtained from database compare to unknown voice recognize as the identity of speaker.

2. PROBLEM STATEMENT

In forensic application, voice of the criminal would obtain in a recorded message. These criminal voice needs to be identified and verified However, identification of speaker using speech may not be totally reliable and accurate. Content of recorded message may be blur and not clear. Besides, the Information of speech extracted may affected by the environment (noise) and the speaker status such as ill, influence of drug and so on. In this project, speech recognition with digital signal processing method is use to filter and analyze the content of the recorded voice.

3. SIGNIFICANCE OF THE RESEARCH

To create awareness on the idiosyncrasies of human voice recognition and the popularization of this fresh idea will have tomorrow's merit. The originality of this research in its virgin perspective serves the significance in the crime track and scale dawn.

Additionally, it is a mental stimulus and hence food for thought to the Nigerian populace as crime prevention is everybody's business. Time will tell whether the research rationales have made the right impact, on this train of thought, this work stresses on the collective effort to rid the nation of all undesirable elements, yes peoples' oriented security control is the catch phrase.

4. AUTOMATIC SPEECH RECOGNITION SYSTEM CLASSIFICATION

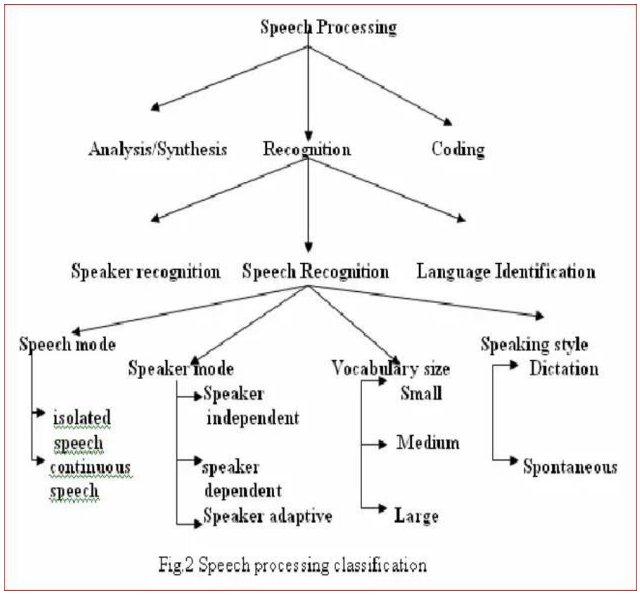

Fig.2 Speech processing classification

5. OBJECTIVE

- In this project, we aim to create a human computer interaction interface where we will be comparing speech samples.
- We will be using MATLAB and correlation for the project.
- Correlation is normally used in signal processing, where we need to compare two signals and need to find the similarity between them. It is also known as the dot product of those two signals.
- In our project we are using correlation to find similarity between our stored signals and the testing signal.
- The audio files with same commands as well as different commands are tested. If the file matches, allowed access plus the matched file present in the database will be spoken.
- If the file doesn't match, the access is denied voice command will be spoken.
- This is a modern and advanced approach in order to check the criminal records by comparing their voices with the earlier stored voices. This helps us to check whether the criminal have any previous records.

6. LITERATURE REVIEW

6.1 Speech production

Speech is nature of sound produce by the human voice mechanisms. Voice speeches introduce to the medium through vibration and transmit to listener. The frequent of vibration of these voices in air medium is known as fundamental frequency. The more frequent the vibration of voice represent in higher fundamental frequency or known as pitch, while lower vibration show the lower pitch. Different person would have different speech pitch. Typically, adult female voice average fundamental frequency is higher compare to male

6.2 Speech analysis

The information grabs form the speech signal contributes to identification of speaker. It is the requirement to identify the characteristic of each individual voice which include vocal tract, excitation source and behavior feature for further processing. Figure 6.1 shows the waveform and spectrum of speaker 1. Figure 6.2 shows the spectrum of same sentences spoke by two different speakers. From the spectrum analysis, the peak for the frequency is different from each other, even the same word is mention. In other word, it means that various speaker have peaks at certain frequency. It is the requirement for the system to analyze the characteristic of spectrum distribution.

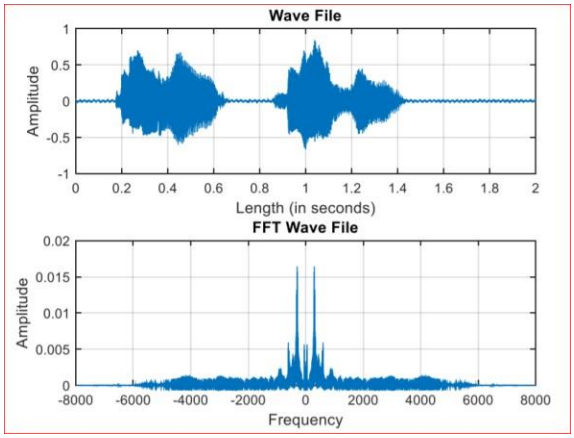

Fig. 6.1 & 6.2: [Typical Voice Recognition Process]

7. METHODOLOGY

- We used nine audio wav files.
- Five of them are the recorded sounds which are already feed in the MATLAB.
- Two are test files which will be recognized by the code.
- The first file is used to check that out of the five uploaded files in the database, which one it is. The test file is a voice command with same wording.
- The second file is used to check whether the different voice command with different wording matches with any file in the database or not.
- The remaining two are success and failure files which will run if we got the recognition or not.
- First of all, what we did was to upload the first five training audio files in the MATLAB software.
- Then we test these five files with the test files with the help of correlation.
- In a wav format, we get two terminologies, amplitude and frequency.
- The amplitude is ignored as it only depicts the loudness of the voice. It doesn't play any kind of role in comparing different sounds.
- Now from the frequency scale, the first column is picked for each file present in the database and compared with the frequency of test file.
- After that we check which one is a maximum match.
- We will also get spectrum graphs of our sound files.
- The matched voice in .wav format as an output is achieved.

8. COMPARISON OF VOICE RECORDINGS

A human can easily recognize a familiar voice. However, getting a computer to distinguish a particular voice among others is a trickier task. Immediately, several problems arise when trying to write a voice recognition algorithm. The majority of these difficulties are due to the fact that it is almost impossible to say a word exactly the same way on two different occasions. Some factors that continuously change in human speech are how fast the word is spoken, emphasizing different parts of the word, etc. Furthermore, suppose that a word could in fact be said the same way on different occasions, and then we would still be left with another major dilemma. Namely, in order to analyze two sound files in time domain, the recordings would have to be aligned just right so that both recordings would begin at precisely the same moment. We start comparison by storing two voices in .wav files. Then we plot both signal and try to match them. Directly comparison done

here by producing wav files through sampling which is done by calculating it's Fourier transform. Next we plot its power spectra and then truncate it to form new power spectra with differences like noise and height of peaks which is to be normalized resulting in a new power spectra. Using mathematical functions we compute and plot an average power spectrum which is also normalized to compare it with two individual voices giving us desired results.

9. ANALYSIS OF MATLAB CODE

The main concept used in the following code is correlation Correlation is normally used in signal processing. It is also known as the dot product of two signals. It can also be used in pattern recognition. The code revolves around the recognition of audio files present in the sample files folder. The wavread function is used instead of the auread function to load a wave file specified by the string filename, returning the sample data in 'y'. The wav extension is appended if no extension is given. Amplitude values are in the range [-1,+1].

The figure below shows:

1) Matched Audio File (.wav format)
2) Graph plot for each file present in database

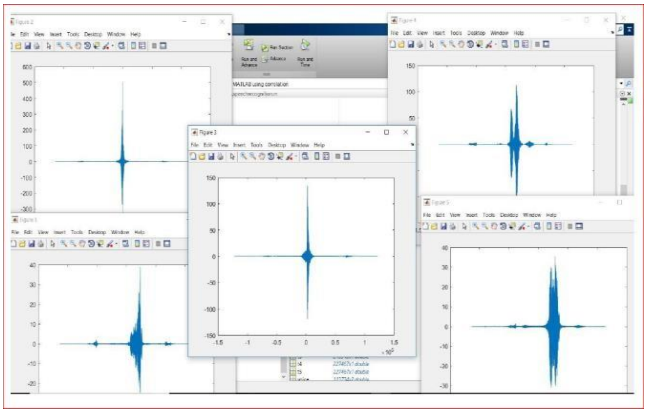

Fig. 7.1: Typical Voice Recognition Process

9.1 Concepts Used in Matlab Code Correlation

Correlation is normally used in signal processing, where you need to compare two signals and need to find the similarity between them. It is also known as the dot product of those two signals. Correlation is a mathematical operation that is very similar to convolution. Just as with convolution, correlation uses two signals to produce a third signal. This third signal is called the cross-correlation of the two input signals. If a signal is correlated with itself, the resulting signal is instead called the autocorrelation.

10. CONCLUSION

To conclude we can say that speech recognition for crime analysis is an upcoming approach which will ease out the work of crime departments as you can compare large amount of data with a speech command more efficiently. This technique aims to detect whether the criminal have any earlier crime records or not. If yes, the access will be allowed, else the access will be denied. Thus, frequency of the test file will be compared with the database files and spectrum graphs will be achieved respectively. This MATLAB code can handle small amount of data as the threshold value is set manually by us. But, this technique can be expanded with the help of a python code with a library called libsndfile.

11. FUTURE SCOPE

Automatic speaker recognition system using MATLAB is an efficient program giving almost 90%of accuracy still there are chances to improve it. i) The main problem to the system is from the external noise. By using some another noise elimination methods, the performance of the system can be improved.

a. Different method of the Silence Remove can be used to improve it.
b. High quality microphone can be used to improve the system accuracy.
c. The system can be tested on larger database.

12. ACKNOWLEDGMENT

I would like to acknowledge thanks to my faculty, Prof. Swathi JN, for providing me with the opportunity of preparing this research paper for the subject of Human and Computer Interaction (CSE4015). This project work binds up a combined experience of working with the audio signal analysis along with learning the working and uses of various components. It taught us how to compare large amount of data with a speech command more efficiently. It would be a bright opportunity to work on any such project work in future too.

13. REFERENCES

[1] J. Kanisha, G. Balakrishanan, "Speech Transaction for Blinds Using Speech-Text-Speech Conversions", Communications in Computer and Information Science book series (CCIS), vol. 131, pp. 43-48, 2011.
[2] C. S. T. Thu, T. Zin, "Implementation of Text to Speech Conversion", International Journal of Engineering Research & Technology, vol. 3, no. 3, pp. 911-915, 2014.
[3] P. S. Shetake, S. A. Patil, P. M. Jadhav, "Review of Text To Speech Conversion Methods", International Journal of Industrial Electronics and Electrical Engineering, vol. 2, no. 8, pp. 29-35, 2014.
[4] P. Khilari, V. P. Bhope, "A Review on Speech To Text Conversion Methods", International Journal of Advanced*Research in Computer Engineering & Technology*, vol. 4, no. 7, pp. 3067-3072, 2015.
[5] A. Joshi, D. Chabbi, M. Suman, S. Kulkarni, "Text To Speech System for Kannada Language", *International Conference on communications and Signal Processing*, 2015.
[6] H. M. Htun, T. Zin, H. M. Tun, "Text To Speech Conversion using Different Speech Synthesis", *International Journal of Scientific & Technology Research*, vol. 4, no. 7, pp. 104-108, 2015.
[7] S. Patil, M. Phonde, S. Prajapati, S. Rane, A. Lahane, "Multilingual Speech and Text Recognition and Translation using Image", *International Journal of Engineering Research & Technology*, vol. 5, no. 4, pp. 85-87, 2016.
[8] S. Patil, M. Phonde, S. Prajapati, S. Rane, A. Lahane, "Multilingual Speech and Text Recognition and Translation using Image", *International Journal of Engineering Research & Technology*, vol. 5, no. 4, pp. 85-87, 2016.
[9] D. B. K. Kamesh, S. Nazma, J. K. R. Sastry, S. Venkateswarlu, "Camera based Text to Speech Conversion Obstacle and Currency Detection for Blind Persons", *Indian Journal of Science and Technology*, vol. 9, no. 30, pp. 1-5, 2016.
[10] B. Sanjana, J. R. Parvin, "Voice Assisted Text Reading System for Visually Impaired Persons Using TTS Method", *IOSR Journal of VLSI and Signal Processing*, vol. 6, no. 3, pp. 15-23, 2016..

