Voice Recognition and Its Application in Online Exam for P.H People

Mr. Vishal D. Jamunkar¹, Mr. Abhijeet A. Sawankar², Mr. Pavan A. Buradkar³ Department of Electronics & Telecommunication Engineering, J.D.I.E.T Yavatmal (Mah) 445001 jamunkarvisha491@gmail.com1, abhijeetsawankar@gmail.com2, pavanburadkar1141990@gmail.com3.

Abstract-Voice Recognition Is A fascinating Field spanning several areas of computer science and mathematics. Reliable speech recognition is a hard problem, requiring a combination of many complex techniques however moderrn method have been able to achive an impressive degree of accuracy. On the other hand ,today most of the companies or institute are conducting their examination online to be a part of this best ever growing world. In his system user can give any available examination at any accessible center as per his/her chice and authority also can condense manpower and process delay overhead. This paper offer one way to conduct online examination . In addition the course of this paper it has been tried to authenticate one registered user and to make the authentication process persistent throughout the examination interlude.

1. INTRODUCTION

Humans have distinctive features that distinguish one person from another. Biometric recognition or simply biometrics, refers to automated methods of recognizing a person based on psychological or behavioral characteristics. The use of biometrics for human identification has a lot of advantages because the features explored are part of the personal information that, in most cases, is impossible to fake, share or forget like, for example passwords or PINs. The way an individual talk is one of those distinctive features that may be used for recognition. Voice, usually considered as a form of behavioral biometric is in fact a combination of both physiological and behavioral biometrics.

A voice biometric is a numerical model of the sound, pattern and rhythm of an individual's voice. A voice biometric or "voice print," is as unique to an individual as a finger or palm print. Any Authentication application that employs a voice channel during the Authentication session is able to add voice biometric authentication to the process for even higher levels of authentication and security. Voice verification technology uses the different characteristics of a person's voice to discriminate between speakers. Speech recognition allows providing input to an application with voice

Speaker Recognition is among the widely used biometrics when it comes to our behavioral characteristics. The speaker recognition area has a long and rich scientific basis with over 30 years of research, development, and evaluations. Over the last decade, speaker recognition technology has made debut in several commercial products. There are two phases in speaker recognition, identification and verification. In speaker identification, we want to determine a speaker (from a set of registered speakers) given an utterance. Verification, a simpler problem, involves determining whether a given speaker matches a given utterance. Quite general, Speaker Verification (SV), is the process of verifying the claimed identity of a registered speaker by using their voice characteristics. Timbre or Harmonic Profile (tone quality or tone colour) Duration of sound means length of sound in time. Duration of sound also deals with sound periodicity.

2. EXISTING ONLINE EXAMINATION SYSTEM

The manual procedure used for conducting examination is time consuming process and error prone due to human limitations. Online Examination System is mainly used to test the domain knowledge of the students and employees with respect to a particular technology. For complete automation, the old manual exam procedure is changed to Online Web Based Examination System . All those Online Examination system finds use of mouse and keyboard only, but none of them uses voice as an essential tool. Through this seminar, I have proposed a system that uses voice of user for attending the online test. Specially, the physically challenged people, those can't use keyboard and mouse for attending the online examination, will now be able to use their voice for the said test. The proposed system will also authenticate the user through verifying the voice and will find an alternative to interact with the system though out the examination.

3. OVERVIEW OF THE PROPOSED METHOD

As voice of the user is the only verifiable entity to authenticate, the proposed method creates voice templates through registration of the user at the first time and it also authenticates the user by recognition of the created voice template. After recognition and authentication, the system proceeds for next phase i.e. the question answer module. Hence, the proposed system can be separated in the following parts:

3.1 Voice Recognition

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 recognized what you are saying. Voice or speech recognition is the ability of a machine or program to receive and interpret dictation, or to understand and carry out spoken commands. In the proposed system only voice will be interaction tool to a user with the system for registration and verification.

i) At first, we have to provide the user details as input in the form of voice asked by system.

ii) The system will then generate a ".wav" file and the generated file will be saved in the database for future references.

iii) At the time of log in by the user, user needs to provide the same information given at the time of registration and the system compares the recorded voice with the one saved in database. If both match, user logs in successfully, otherwise not. After successful log in, the set of questions from database appears on the screen which is read by system itself. And the user will speak out the answer as option A, B, C or D which goes to the database where it matches with the previously stored answer for comparison. Based on the comparison result, the system keeps record of the user's score.

For correct matching at the time of registration, system stores user voice for A, B, C and D. These files are compared for matching the questions answer later. 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 it's power spectra and then truncate it to form a 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. The voice comparison and analysis in our proposed system has been done in the following way illustrated with an example: We have recorded the word ``hello" twice by the same speaker. The sounds are stored in ``test1.wav" and `test2.wav". We start by reading them into MATLAB vectors Y1 and Y2.

Y1=wavread('test1');

Y2=wavread('test2');

Next, we plot both signals in one figure. To do this we type

subplot(2,1,1);	plot(Y1)
subplot(2,1,2);	plot(Y2)

The result is shown in below figure :



two signals Y1 and Y2



3.2 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 more tricky 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

plot(X2)

Comparing these graphs we have seen that, the bulk of the signals are at different positions. Matching these positions is difficult. The signal Y2 starts at approximately 16,000 and ends at 40,000, Y1 starts at 20,000 and ends at 44,000. So we create two new vectors.

X1(1:24000)=Y1(20001:44000);

X2(1:24000)=Y2(16001:40000); and plot them again using subplot(2,1,1); plot(X1)

subplot(2,1,2); The result is shown in below figure :



Fig 2: The two ``cleaned-up" signals X1 and X2

Now the two signals match approximately and we can try to compare them directly.

Figure plot(abs(X1-X2));

Produces a graph of the difference shown in figure below. As we see in the figure, this can take value up to 1.8 whereas the original signal has only values between -0.9 and 0.9. The problem seems to be that we need to exactly match the two figures. Since this seems to be an impossible task, we continue this using the Fourier transform. WAV files are produced by sampling 22050 times per second. i.e the vector Y1 with 55,130 entries represents approximately 2.5 seconds. In order to compare the fast Fourier Transforms of two vectors, they need to be of the same length. Otherwise, we would get a slight shift in frequencies. We will base our length on 60,000, which represents approximately 3 seconds worth of signal.

To compute the Fourier Transform based on this length we proceed as

Z1=fft(Y1,60000);

Z2=fft(Y2,60000);



To visualize this we compute the power spectra and plot them as:



Fig 4: The power spectra p1 and p2.

We see that in both spectra the peaks are at approximately the same locations. The index along the X-axis represents the frequency. Since the length of the signal was approximately 3 seconds, \hat{k} =500 corresponds to 499 cycles per 3 seconds or approximately 160Hz. We also see that everything happens for $\hat{k} < 2500$, which corresponds to frequencies below about 900 Hz. We recall, that the C-major scale ranges from 261 Hz to 523 Hz, and conclude that most action of our vocal cords also happens under 900 Hz. It therefore makes sense to

truncate the Fourier Transforms Z1 and Z2 at approximately 2500.



Fig 5: The truncated power spectra p1 and p2

It seen that three major peaks in the two graphs of Figure, which agree in location. However, there are two distinct differences in the two graphs. The second graph shows much more in the region for &mathackappa < 500. But, these are Frequencies below 170 Hz. This represents low frequency noise and is not necessarily part of the voice. We will filter tt.

 $\begin{array}{l} t1(1:500) = zeros(500,1);\\ t2(1:500) = zeros(500,1);\\ Q1(1:500) = zeros(500,1);\\ Q2(1:500) = zeros(500,1); \end{array}$

The other difference is height of the peaks. This represents in some way the volume.



. Fig 6: The cleaned up power spectra Q1 and Q2 $\,$

For authentication of voice an efficient spectrum analysis to continuous-time signal based on DFT, the Fast Fourier Transform (FFT) method is adopted. It breaks down DFT of N-points sequence to the one of shorter sequence gradually, which will greatly improve the calculating speed of DFT. This huge improvement made many DFT-based algorithms practical; FFTs are of great importance to a wide variety of applications, from digital signal processing and solving partial differential equations to algorithms for quick multiplication of large integers.

4. PROPOSED METHODS ARCHITECTURE

The architecture for the proposed examination system is

given herewith



Fig. 7.Architecture of the proposed system

[3] Liang Zhang, Yue-ting Zhuang, Zhen-Ming Yuan, Guohua Zhan, "A Web-Based

5. SIMULATION RESULT

In the paper, the full programming simulation is being done using simple Mat lab Programming. Below some of the snapshots are given from the Matlab Program which controls the voice driven examination system scheme as a simulator: After the start of the proposed system it will first registered the user through voice. Then for login it will ask the user to give requisite information. If same voice and information recognized user will logged in as shown in figure



Fig 8. Spectra comparison of the user voice

6. CONCLUSION

The proposed system is developed to conduct an Online Examination using voice which will be helpful for people who are blinds and the people who don't use keyboard for interaction with the system.

The proposed method will help for voice recognition where we take voice as input through microphone and then register for Online examination. Then at the time of loging user will logged in to the online examination system through voice authentication, which uses FFT for comparing input voice of the user with the template voice. Next the user will proceed with the questions displayed and read by the system.

REFERENCES

[1] Anil K. Jain, Arun Ross, Salil Prabhakar "An Introduction to Biometric Recognition" IEEE Transactions on circuit and systems for video technology, Vol 14, No. 1, January 2004.

[2] B. Schneier, "The Uses and Abuses of Biometrics," Comm. ACM, vol. 42, no. 8, p. 136, Aug. 1999.

[4] N.K. Ratha, J.H. Connell, and R.M. Bolle, "A Biometrics-Based Secure Authentication System," Proc. IEEE Workshop Automatic Identification Advanced Technologies, pp. 70-73, Oct. 1999. Examination and Evaluation System for Computer Education"Proceedings.