SPEAKER RECOGNIZED SECURITYBASED VOTING MACHINE

Prof. Neha Mathais[1], Nilesh Avhad[2], Suraj Ware[3], Gaurav Thakre[4]

[1,2,3,4] Electronics and Telecommunication Engineering,

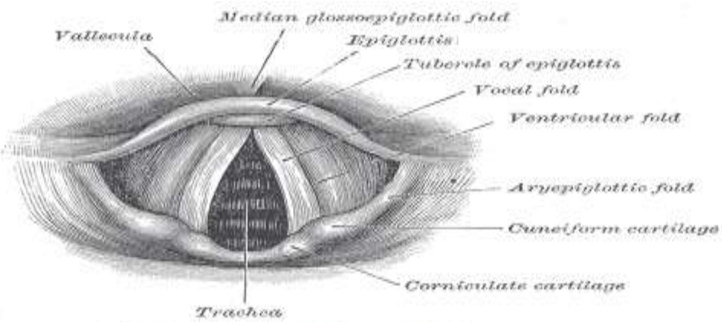
JSPM Rajarshishahu College of Engineering, Tathawade, Pune, (India)

ABSTRACT

Speech processing is one of the important area of digital signal processing. The objective of the automatic speaker reorganization is to extract, characterize, recognize information about speaker identity. The Mel Frequency Cepstrum Coefficient (MFCC) used for design a text dependent speaker identification system. This voice reorganization system is used for improvisation of security, scalability, &flexibility of electronic voting machine model.

I. INTRODUCTION

The human speech or voice is produce due to vocal cord in the larynx. The human speech contain numerous discriminative feature that can be used to identify speaker. Speech contains energy from zero to 5khz. The Mel frequency cepstrumcoefficient(MFCC) frature has been used for designing a text dependent speaker identification system. The extracted speech features of the speaker are quantized to a number of centroids using vector quantization algorithm. The Euclidean distance between the MFCC's of each speaker in training phase to the centroids of individual speaker in testing phase is measured and the speaker is identified according to the minimum Euclidean distance. The code is developed in the MATLAB environmental & perform the identification satisfactorily. This features are used for security purpose to cast their important votes. 2. HUMAN VOICE:- Human voice can be subdivide into 2 parts:-1)The lungs 2)The vocal fold within the larynx. The lungs produce adequate air flow & air pressure to vibrate the vocal fold. Larynx is major source of sound through the rhythmic opening and closing of vocal fold. The vocal folds are vibrating values in which air flow lungs into available pulses that from the laryngeal sound source. The adult male and female have different size of vocal fold. The male vocal fold are between 17mm to 25mm in length having frequency 125hurtz and female vocal cord are between 12.5mm to 17.5mm in length having frequency 210hurtz. The sound of each individual's voice is entirely unique not only because of actual shape and size of an individual vocal cord but also due to size & shape of rest of that person body.



View of vocal cord (vocak fold)

II. SPEAKER RECOGNIZATION

Anatomical structure of the vocal cord is uniquefor everyperson and hence the voice information available in the speech signal can be used to identify the speaker. Voice comes under biometric identity category.Using voice for identity has major advantage that the remote person authentication. Speakerreorganization methods can be two types text-independent and text-dependent method bothmethodshave its advantage as well as disadvantage.In text independent method speaker model capture characteristics of somebody's speech irrespective of what one is saying. In text dependent method, identity is based on speaking one or more specific phrases like codes,password,number's etc.

It involves two phases namely training and testing. Training phase consist process of familiaring the system with the voice characteristics of speaker registration. Testing is actual recognizing task.

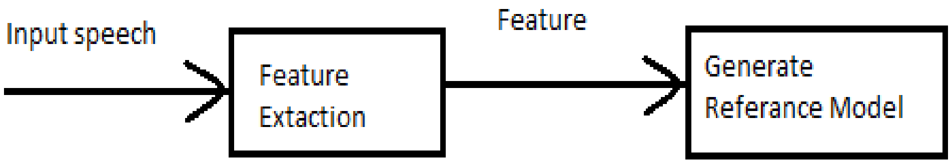


Fig.1. The block diagram of training phase.

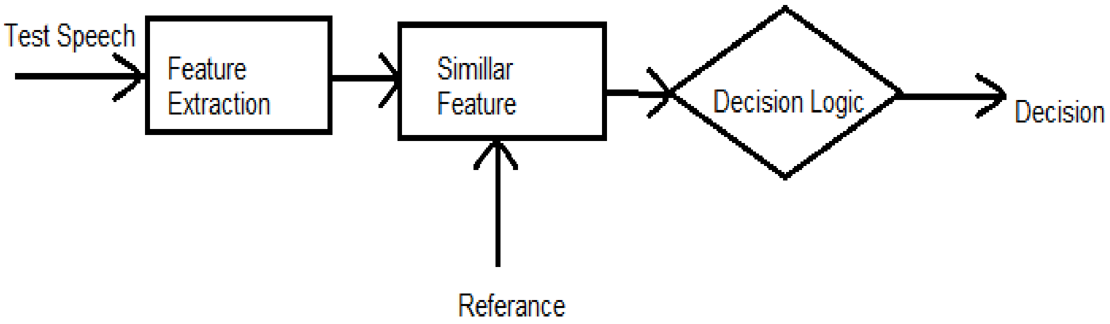


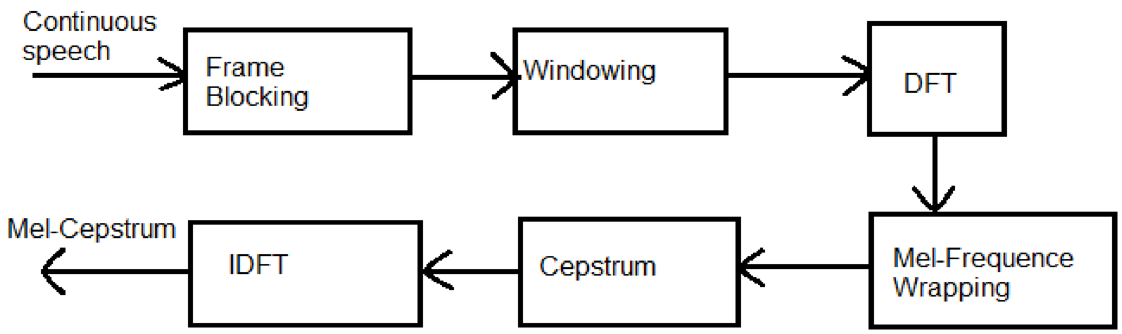
Fig.2. The block diagram of the testing phase.

III. TECHNIQUES OF FEATURE EXTRACTION

The features can be extracted either directly from the time domain signal or from a transformation domain depending upon the choice of thesignal analysis approach. Some of the audio features that have been successfully used for audio classification include Mel-frequency cepstral coefficients (MFCC), Linear predictive coding (LPC), Local discriminant bases (LDB).

IV. MEL-FREQUENCE CEPSTRUM COEFFICIENT

The extraction and selection of the best parametric representation of acoustic signals is an important task in the design of any speech recognition system; it significantly affects the recognition performance. A compact representation would be provided by a set of mel-frequency cepstrum coefficients (MFCC), which are the results of a cosine transform of the real logarithm of the short-term energy spectrum expressed on a mel-frequency scale. The MFCCs are proved more efficient.



Block Diagram of MFCC Processer

V. MEL-FREQUENCE WRAPPING

The speech signal consists of tones with differentfrequencies. For each tone with an actual Frequency, f, measured in Hz, a subjective pitch is measured on the 'Mel' scale. The mel-frequencyscale is a linear frequency spacing below 1000Hz and a logarithmic spacing above 1000Hz. As a reference point, the pitch of a 1kHz tone, 40Db above the perceptual hearing threshold, is defined as 1000 mels. Therefore we can use the following formula to compute the mels for a given frequency f inHz.

mel(f)= 2595*log10(1+f/700) ……….. (1)

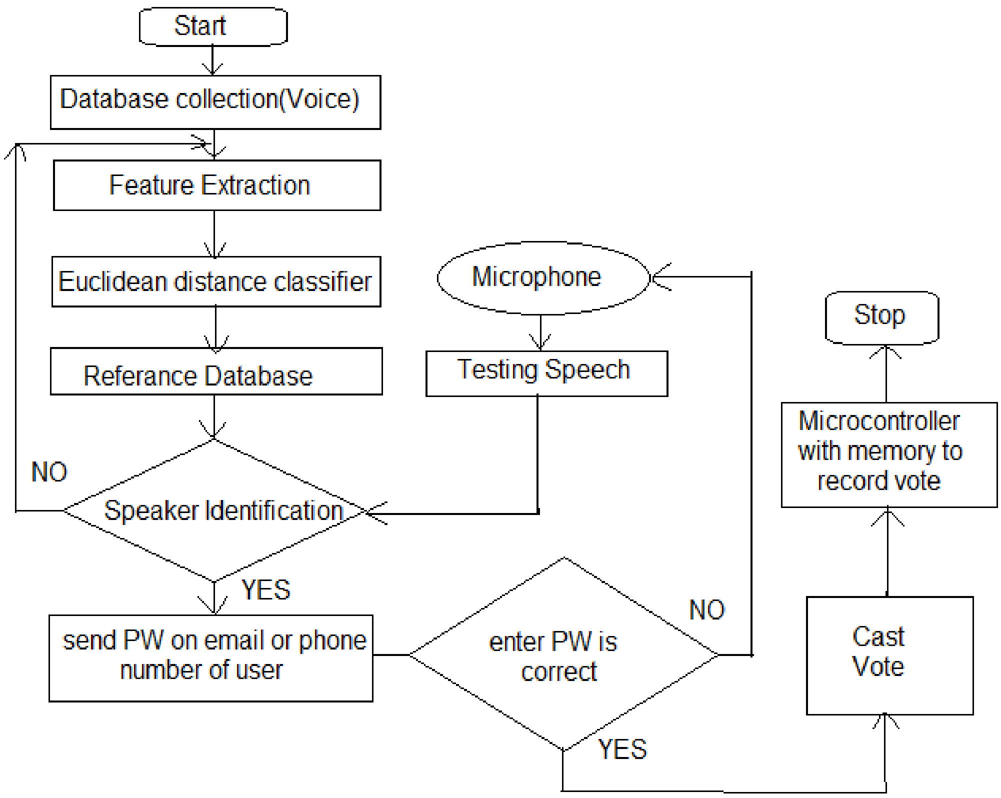
One approach to simulating the subjective spectrum is to use a filter bank, one filter for each desired mel-frequency component. The filter bank has a triangular bandpass frequency response, and the spacing as well as the bandwidth is determined by a constant mel-frequency interval.

VI. CEPSTRUM

In the final step, the log mel spectrum has to be converted back to time. The result is called the mel frequency cepstrum coefficients (MFCCs). The cepstral representation of the speech spectrum provides a good representation of the local spectral properties of the signal for the given frame analysis. Because the mel spectrum coefficients are real numbers(and so are their logarithms), they may be converted to the time domain using the Discrete Cosine Transform (DCT). The MFCCs may be calculated using this equation

$$C_n = \sum_{k=1}^{k} (\log S_k) \cos[n(k-1/2)\pi/k]$$

where n=1,2,….K The number of melcepstrum coefficients, K, is typically chosen as 20. The first component, *c* is excluded from the DCT since it represents the mean value of the input signal which carries little speaker specific information. By applying the procedure described above, for each speech frame of about 30 ms with overlap, a set of mel-frequency cepstrum coefficients is computed. This set of coefficients is called an *acoustic vecto*r. These acoustic vectors can be used to represent and recognize the voice characteristic of the speaker [4]. Therefore each input utterance is transformed into a sequence of acoustic vectors. The next section describes how these acoustic vectors can be used to represent and recognize the voice characteristic of a speaker.



Structural Flow Of Speaker Recognized Security Based Voting Machine

VII. OBJECTIVE OF THE PROJECT WORK

Making a simple electronic device used to record/cast votes with help of mobile phone based on voice recognition system for security in place of ballot papers and boxes as well as electronic voting machine which were used earlier in conventional voting system. It eliminates the possibility of invalid and doubtful votes which, in many cases, are the root causes of controversies and election petitions. It makes the process of counting of votes much faster than the conventional system and voter can cast their vote from anywhere. It reduces to a great extent the quantity of paper used thus saving a large number of trees making the process eco-friendly.

Security:-The system is free from intentional tamper. It is not possible to hack the machine. Though this factor depends on the personnel integrity, attempts should be made to make the model as secure as possible. In this model every user is provided with a password. The votes will be successful only after successful verification of voice recognition and then password.

Reliability:-The machine registers the votes faithfully. A vote is never altered. A valid vote is never eliminated, from the final tally and an invalid vote is not counted. Vote counting is flawless. The final vote tally must be perfect. Most importantly the votes are stored in EEPROM memory, where the numbers of votes are stored permanently.

Scalability:-It is easy to use the basic design for any number of voters. The model is able to handle increasing voter participation without any stress on performance.

Flexibility:-The design is such that it can be put to use in various polling systems, with different requirements and mechanisms.

VIII. IMPLEMENTATION

We decided to implement these modules in Matlab. After examining literature, the number of mfcc coefficient was decided to be 28. We collect 10 speech samples from our faculty for testing. We compute 28 mfcc coefficients of all speech samples and stored them.For classification purpose we use Euclidean distance classifier. The speaker was identified as the closest matching mfcc coefficients stored.

IX. CONCLUSION

The conclusion of the given paper is as follows:-

1) To create a speaker recognition system, and apply it to a speech of an unknown speaker.
2) Investigate the extracted features of the unknown speech and then compare them to the stored extracted features for each different speaker in order to identify the unknown speaker.
3) Apply the speaker recognition system for security purpose to cast vote on basis of phone based voting machine.

X. RESULTS

Project result obtained is as follows:-

1) We tested our code on 10 speakers and achieved 100% accuracy.
2) We implemented a Euclidean distance classifier for speaker identification.

REFERANCES

[1] M. Pandit and J. Kittler, "Feature selection for a dtw-based speaker verification system, in *Proceedings of IEEE Int.Conf. Acoust. Speech and Signal Processing*,**2:** 769-772(1998).

[2] S. Furui, "An overview of speaker recognition technology, in Automatic Speech and Speaker Recognition (C.H. Lee, F.K. Soong, and K.K. Paliwal,eds), ch.2 pp.31-56Boston :Kluwer Academic, (1996).

[3] Automatic speaker recognition by S.Khan, MohdRafibullslam, M. Faizul, D. Doll. *3rd international conference on electrical and computer engineering* (ICECE), 28-30th Dec. (2004), Dhaka, Bangladesh.

[4] Speaker recognition using MFCC by S. Khan, MohdRafibullslam, M. Faizul, D. Doll, presented in *IJCSES (International Journal of Computer Science and Engineering System)* **2**(1)**:** 2008.

[5] Speaker identification using MFCC coefficients –MohdRasheedur Hassan, Mustafa Zamil, MohdBolamKhabsani, MohdSaifurRehman. *3rd international conference on electrical and computer engineering* (ICECE), (2004).

[6] Premakanthan and W.B. Mikhael, Speaker verification/ recognition and the importance of selective feature extraction: Review, *Proceedings of the 44th IEEE 2001,Midwest Symposium*, **1:** 14-17(2001).

[7] F .Bimbot, J. Bonastre, C. Fredouille, G. Gravier, I. Magrin-Chagnolleau,S. Meignier, T. Merlin, J. Ortega-Garcia, D. Petrovska-Delacretaz,and D. Reynolds, "A tutorial on text-independent speaker verification,"*EURASIP J. Appl. Signal Process*., vol. 2004, no. 4, pp. 430–451, 2004.

[8] Zhong-Xuan, Yuan & Bo-Ling, Xu& Chong-Zhi, Yu. (1999). "Binary Quantization of Feature Vectors for Robust Text-Independent Speaker Identification" in *IEEE Transactions on Speech and AudioProcessing, Vol. 7, No. 1*, January 1999. IEEE, New York, NY, U.S.A.

[9] F. Soong, E. Rosenberg, B. Juang, and L. Rabiner,"A Vector Quantization Approach to Speaker Recognition", AT&T Technical Journal, vol. 66,March/April 1987, pp. 14-26