Voice Based Biometric System Feature Extraction Using MFCC and LPC Technique

Mr. Vijay K. Kale[1] , Dr. Prapti D.Deshmukh[2], Mr.Hanumant R.Gite[3]

[1,3]Department of Computer Science & IT, Dr.Babasaheb Ambedkar Marathwada University, Aurangabad, (MS), India
[2]MGM'S Dr.G.Y.Pathrikar College of Computer Science & IT, Aurangabad, (MS), India

***Abstract*— *Now a day, interest in using biometric technologies for person authentication in security systems has grown rapidly.Voice is one of the most promising and mature biometric modalities for secured access control this paper gives an experimental overview of techniques used for feature extraction in speaker recognition. The research in speaker recognition have been evolved starting from short time features reflecting spectral properties of speech low-level or physical traits to the high level features (behavioral traits) such as prosody, phonetic information, conversational patterns etc. first give a brief overview of Speech processing and voice biometric relation and then describe some feature extraction technique. We have performed experiment for feature extraction of MFCC, LPC techniques.*

Keywords*— *MFCC, LPC, Biometric, Feature, Voice.

I. INTRODUCTION

Now a day in various field computerizations growing number of speaker-recognition tasks with such a technologies as speaker verification and speaker identification. Voice biometrics uses the features of a person's voice to find out the speaker's identity. Systems performing this function have been applied to real-world security applications for more than a decade. Their use is increasing rapidly in a broad spectrum of industries, including financial services, retail, corrections, even entertainment. Voice-biometrics systems can be categorized as belonging in two industries: speech processing and biometric security. Human voice conveys information about the language being spoken and the emotion and gender for the identity of the speaker. Speaker recognition is a process where a person is recognized on the basis of his voice signals [1, 2]. The Objective of speaker recognition is to determine which speaker is present based on the individual's utterance. This is in contrast with speaker verification, where the objective is to verify the person's claimed identity based
On his or her utterance. Speaker identification and speaker verification fall under the general category of Speaker recognition [3, 4].In speaker identification there is two types, one is text dependent and another is text independent. Speaker identification is divided into two components: feature extraction and feature.

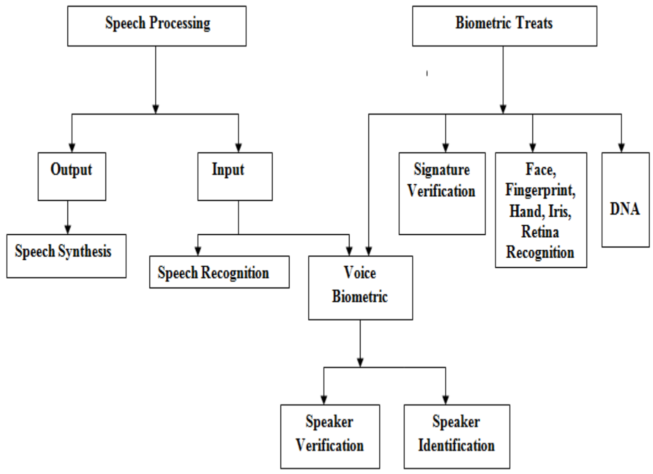

Fig 1.1:Voice Biometric Tree Structure

Historically, all speaker recognition systems have been mainly based on acoustic cues that are nothing but physical traits extracted from spectral characteristics of speech signals. So far the features derived from the speech spectrum have proven to be the most effective in automatic systems, because the spectrum reflects the geometry of system that generates the signal. Therefore the variability in the dimensions of the vocal track is reflected in the variability of the spectra between the speakers. However, studies [5] have proved that there is a large amount of information suitable for speaker recognition being the top part related to learned traits and the bottom part to physical traits.

II. VOICE BIOMETRICS TYPES

1. **Speaker verification:** Speaker-verification systems authenticate that a person is who she or he claims to be.
2. **Speaker Identification:** Speaker identification Assigns an identity to the voice of an unknown speaker

III. SPEECH RECOGNITION PROCESS

The speaker Reorganization system may be viewed as working in a four stages

- Analysis

ISSN: 2349-6495

- Feature extraction
- Modeling
- Testing

Feature Extraction is the most important part of speech recognition since it plays an important role to separate one speech from other. Because every speech has different individual characteristics embedded in utterances. These characteristics can be extracted from a wide range of feature extraction techniques proposed and successfully exploited for speech recognition task. But extracted feature should meet some criteria while dealing with the speech signal such as:

- Easy to measure extracted speech features
- It should not be susceptible to mimicry
- It should show little fluctuation from one speaking environment to another
- It should be stable over time
- It should occur frequently and naturally in speech

From human speech production mechanism, it is possible to identify individual using the speech data. Speech contains speaker specific information due to vocal track and excitation source. Larynx is the major excitation source, whereas vocal track is the major resonant structure. Speaker information is due to particular shape, size and dynamics of vocal track and also the excitation source. These features related to physiological nature of human speech production are called physical traits, which are used in state-of-art systems. However human speaker recognition relies on other sources of information like speaking style, pronunciation etc. Such features are referred to as behavioral traits. Further, the behavioral traits like how the vocal tract and excitation source are controlled during speech production are also unique for each speaker. The information about the behavioral trait is also embedded into the speech signal and can be used for speaker recognition. Thus the information present in speech signal carries the identity of speaker at different levels. To properly represent speech data, it is necessary to analyses it using suitable analysis techniques. The analysis techniques aim at selecting proper frame size and shift for analysis and also at extracting the relevant features in the feature extraction stage [6].

IV. FEATURE EXTRACTION

Types of Features:

A vast number of features have been proposed for speaker recognition. We divide them into the following classes:

- Spectral features
- Dynamic features
- Source features
- Suprasegmental features
- High-level features

Spectral features are descriptors of the short-term speech spectrum, and they reflect more or less the physical characteristics of the vocal tract. Dynamic features relate to time evolution of spectral (and other) features. Source features refer to the features of the glottal voice source. Super a segmental features span over several segments. Finally, high-level features refer to symbolic type of information, such as characteristic word usage. The most widely used feature extraction techniques are explained below.

V. MEL FREQUENCY CEPSTRAL COEFFICIENT (MFCC)

A block diagram of an MFCC feature extraction is shown (Fig. 2).This coefficient has a great success in speaker recognition application. The MFCC [7] [8] is the most evident example of a feature set that is extensively used in speech recognition. As the frequency bands are positioned logarithmically in MFCC [10], it approximates the human system response more closely than any other system. Technique of computing MFCC is based on the short-term analysis, and thus from each frame a MFCC vector is computed.

In order to extract the coefficients the speech sample is taken as the input and hamming window is applied to minimize the discontinuities of a signal. Then DFT will be used to generate the Mel filter bank. MFCC can be computed by using the formula.

$$Mel(f) = 2595 * \log10(1 + f/700) \quad (1)$$

The Mel-frequency cepstrum (MFC) is a representation of the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear Mel scale of frequency. Mel-frequency cepstral coefficients (MFCCs) are coefficients that collectively make up an MFC. The difference between the cepstrum and the Mel-frequency cepstrum is that in MFC the frequency bands are equally spaced on the Mel scale, which approximates the human auditory system's response. MFCCs are commonly used as features in speech recognition system. To enhance the accuracy and efficiency of the extraction processes, speech signals are normally pre-processed before features are extracted. [12].The following figure 2 shows the steps involved in MFCC feature extraction.

ISSN: 2349-6495

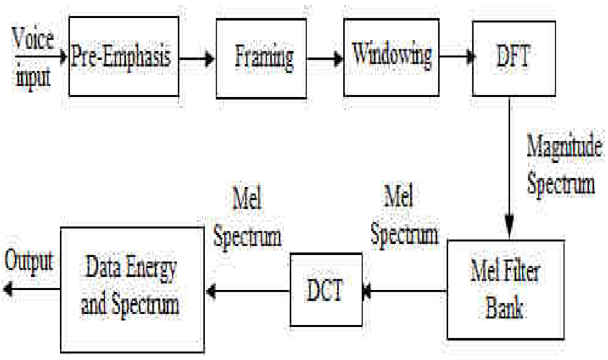

Fig.2: Block diagram of Mel frequency cepstral

VI. LINEAR PREDICTIVE CODING (LPC)

Linear prediction is a mathematical computational operation which is linear combination of several previous samples. LPC [7] [8] of speech has become the predominant technique for estimating the basic parameters of speech. It provides both inaccurate estimate of the speech parameters and it is also an efficient computational model of speech. The basic idea behindLPC is that a speech sample can be approximated as a linear combination of past speech samples. Through minimizing the sum of squared differences (over a finite interval) between the actual speech samples and predicted values, a unique set of parameters or predictor coefficients can be determined. These coefficients form the basis for LPC of speech [10]. The following figure 3 shows the steps involved in LPC feature extraction

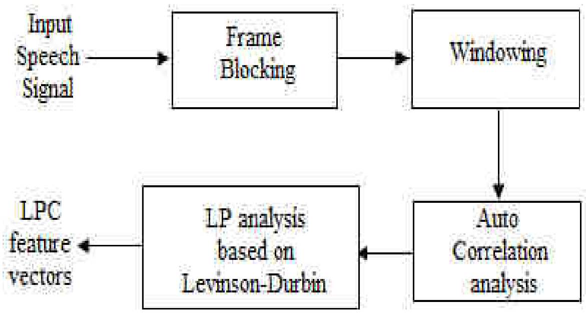

Fig.3: Block diagram of linear predictive coding

VII. EXPERIMENTAL RESULTS

The experiment part perform two experiment first experiment perform on MFCC Feature extraction technique .here use the KVKMMBR speech standard Database. The database contain recording of WAV file. In the recording word like pronounce character A to Z .another subject number from zero to nine digits. The MFCC thirteen features extracted from wav file .Then calculate the Mean, mode and Standard Deviation of extracted feature matrix. The MFCC Thirteen Feature manse one energy cofficent, and Twelve Spectral coefficient The second feature extraction technique use LPC in this experiment also use same database and extract the twenty features. All experiment performs in MATLAB software. The sample of feature of both techniques is shown below.

1. MFCC Feature Extraction:

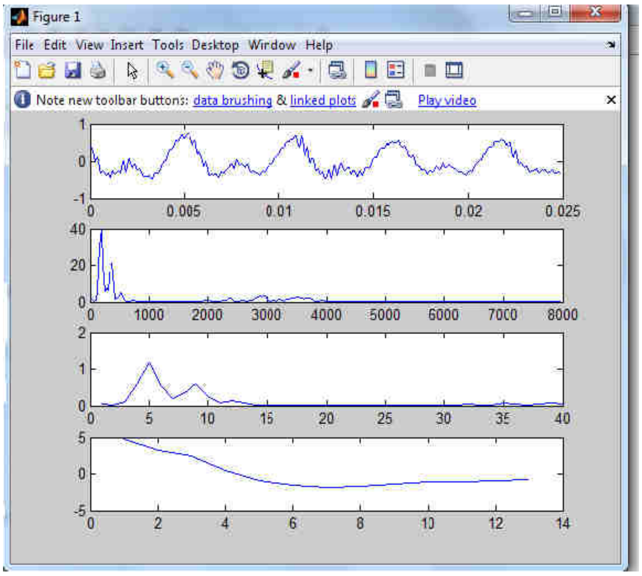

Fig.4:Sample1 MFCC Feature Extraction character 'A' wav

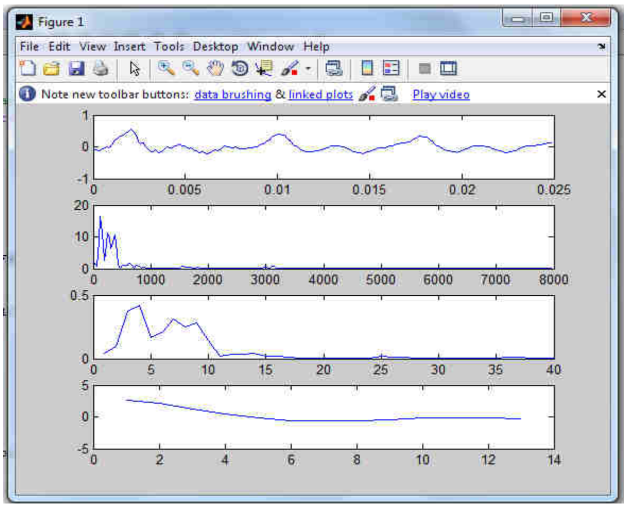

Fig.5:Sample2 MFCC Feature Extraction Digit'one.wav

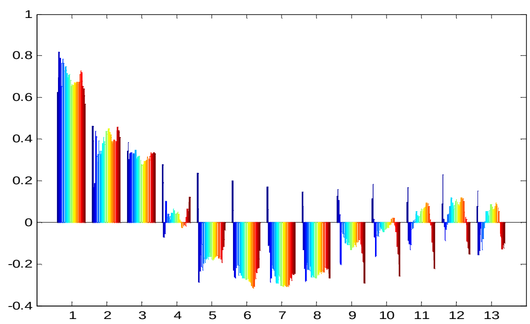

Fig.6:Sample 1 MFCC Feature Graph of character' A' wav

ISSN: 2349-6495

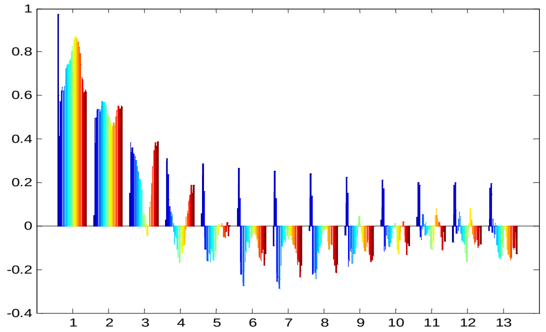

Fig.7:Sample 2 MFCC Feature Graph of Digit 'one.Wav

Table 1. MFCC Extracted Feature

MFCC Feature	Mean	Mode	Median	Standard Deviation
Sample 1 Char. 'A'	-0.3522	-0.3522	-0.0355	0.2765
Sample 2 Degit'One'	-0.2864	-0.2864	0.1252	0.2411

2. LPC Feature Extraction:

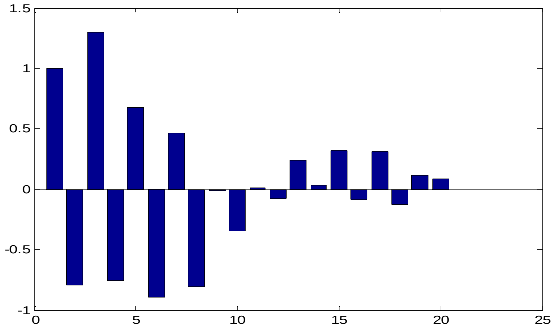

Fig.8:LPC Coefficient Feature Graph of character 'A'wav

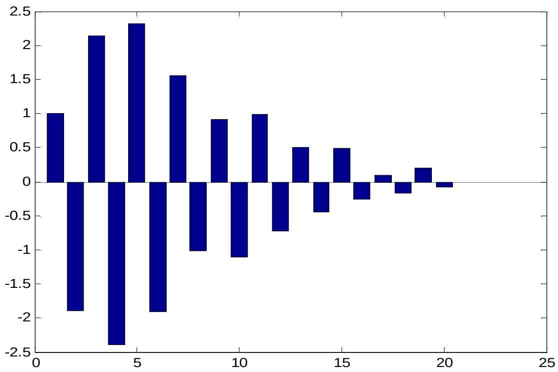

Fig.9:LPC Coefficient Feature Graph of Degit 'one'wav

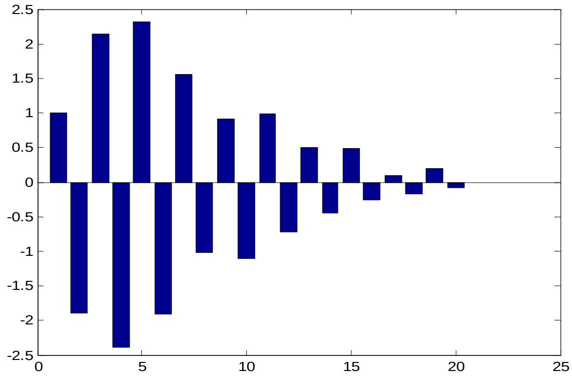

Fig.10:LPC Cepstrum Feature Graph of character 'A'wav

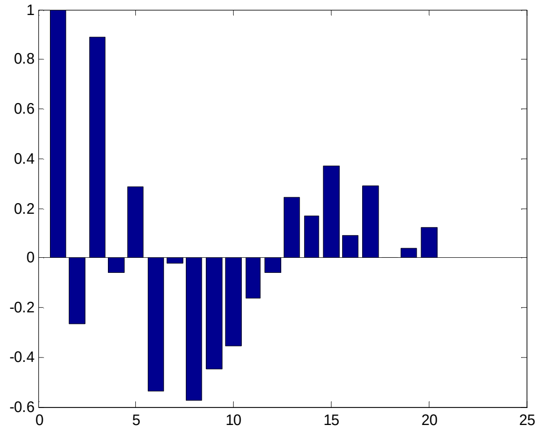

Fig.11.LPC Cepstrum Feature Graph of Digit 'one.wav

Table 2. LPC coefficient Feature

LPC Coefficient Feature	Mean	Mode	Median	Standard Deviation
Sample 1 Char. 'A'	-0.8936	0.0257	-0.8936	0.5808
Sample 2 Degit'One'	-2.3833	0.0075	-2.3833	1.2915

Table.3:LPC Cepstrum Feature

LPC Cepstrum Feature	Mean	Mode	Median	Standard Deviation
Sample 1 Char. 'A'	-0.5756	0.0205	0.5756	0.4109
Sample 2 Degit'One'	-1.3543	0.0051	-1.3543	0.6771

VIII. CONCLUSION

In this paper, we have introduced Voice biometric tree structure and types. Also presented an overview of the feature extraction methods of MFCC and LPC.Then perform Experiment for MFCC and LPC feature extraction in MATLAB software. The sample result of feature is shown in this Paper. The result part show value of Mean, Mode, Median and Standard Deviation, the graphical representation also show the extracted feature. The advantages of this work are introduced how the feature extraction perform in MATLAB. Our proposed work is developing Voice biometric system for identiofation of person.

IX. ACKNOWLEDGEMENTS

We Authors would like to acknowledge and thanks to University Grants Commission (UGC), India for granting UGC SAP (II) DRS Phase-I & Phase-II F. No. 3-42/2009 & 4-15/2015/DRS-II Biometrics: Multimodal System Development Laboratory facility and One Time Research Grant F. No. 4-10/2010 (BSR)& 19-132/2014 (BSR) supporting to this work and Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India.

REFERENCES

[1] Campbell J.P. and Jr. "Speaker recognition: A Tutorial" Proceeding of the IEEE. Vol 85, 1437-1462 1997.
[2] S.Furui. "Fifty years of progress in speech and speaker recognition," Proc. 148th ASA Meeting, 2004.s
[3] A. Rosenberg, "Automatic speaker recognition: A review," Proc. IEEE, vol. 64, pp. 475487, Apr.1976.
[4] G. Doddington, "Speaker recognition-Identifying people by their voices," Proc. IEEE, vol. 73, pp. 1651-1664, 1985
[5] Marcos Faundez-Zanuy and Enric Monte-Moreno, "State-of-the-art in Speaker Recognition ", IEEE A&E Systems Magazine, May 2005.
[6] Jayanna, H.S. and S.R.M. Prasanna, 2009. Analysis, feature extraction, modeling and testing techniques for speaker recognition. IETE Technical Review, 2009, Volume 26, issue 3
[7] Corneliu Octavian DUMITRU, Inge GAVAT,"A Comparative Study of Feature Extraction Methods Applied to Continuous Speech Recognition in Romanian Language", 48th International Symposium ELMAR-2006, 07-09 June 2006, Zadar, Croatia.
[8] DOUGLAS O'SHAUGHNESSY, "Interacting with Computers by Voice: Automatic Speech Recognition and Synthesis", Proceedings of the IEEE, VOL. 91, NO. 9, September 2003, 0018-9219/03 2003 IEEE.
[9] A.P.Henry Charles &G.Devaraj, "Alaigal-A Tamil Speech Recognition", Tamil Internet 2004, Singapore.
[10] N.UmaMaheswari, A.P.Kabilan, R.Venkatesh, "A Hybrid model of Neural Network Approach for Speaker independent Word Recognition", International Journal of Computer Theory and Engineering, Vol.2, No.6, December, 2010 1793-8201.
[11] Tejal Chauhan, Hemant Soni, Sameena Zafar, "A Review of Automatic Speaker Recognition System", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-4 September 2013.