

# Feature Extraction of Speech Signal using LPC

Amrutha R<sup>1</sup>, Lalitha K<sup>1</sup>, Dr M Shivakumar<sup>2</sup>, Stafford Michahial<sup>1</sup>

Asst Prof, Dept. of EIE, GSSSIETW<sup>1</sup>

Prof & Head Dept of EIE, GSSSIETW<sup>2</sup>

**Abstract:** Speech is an ancient field of study and research is being done on it till date. Automatic Speech recognition system deals with analysis and recognition of the input speech signal by the machine or computer in various environments. To enhance the accuracy and capability of the system various feature extraction techniques are implemented. This research paper provides a brief overview of Speech recognition system and its various phases like analysis, feature extraction, modeling and testing or matching. In addition it also includes detailed and comparative study on Linear Predictive Coding (LPC) feature extraction techniques used in Automatic Speech Recognition systems. The main objective of this paper is to briefly summarize speech recognition system and three feature extraction methods that are an integral part of ASR.

**Keywords:** Automatic Speech Recognition system, Feature Extraction.

## I. INTRODUCTION

Speech has been one of the most potent tools at man's disposal since ancient times. Humans have constantly evolved and expressed themselves through speech. There is plethora of languages used and spoken by man throughout the entire world. Computer speech recognition or Automatic Speech recognition system is a process in which the words spoken by humans or speech signals are translated into words. The words that are recognized by the machine can be the final output or these words can act as input to natural NLP i.e. natural language processing, it also uses algorithm, which is implemented as computer program.[3]

Automatic speech recognition (ASR) has made great strides with the development of digital signal processing hardware and software. But despite of all these advances, machines cannot match the performance of their human counterparts in terms of accuracy and speed, especially in case of speaker independent speech recognition. So today significant portion of speech recognition research is focussed on speaker independent speech recognition problem. The reasons are its wide range of applications, and limitations of available techniques of speech recognition. In this report we briefly discuss the signal modeling approach for speech recognition. It is followed by overview of basic operations involved in signal modeling. Further commonly used temporal and spectral analysis techniques of feature extraction are discussed in detail[4]. Speech is acoustic signal which contains information of idea that is formed in speaker's mind. Speech is bimodal in nature, Automatic Speech Recognition (ASR) only considers acoustic information contained in speech signal. In noisy environment, it is less accurate. Audio Visual Speech Recognition (AVSR) out weights ASR as it uses acoustic and visual information contained in speech. Speech processing can be performed at different three levels. Signal level processing considers the anatomy of human auditory system and process signal

in form of small chunks called frames. In phoneme level processing, speech phonemes are acquired and processed.[5]

## II. IMPLIMENTAION

The speech recognition system can be classified mainly into four phases.[3]

- Analysis
- Feature Extraction
- Modeling
- Testing or Matching

**ANALYSIS:** The first phase of speech recognition system is the analysis phase. The speech signal is provided as input to the machine and it contains various signal information. Analysis is done through various techniques:

- ✓ Segmentation Analysis.
- ✓ Sub-segmental Analysis.
- ✓ Supra-segmental Analysis

**FEATURE EXTRACTION TECHNIQUES:** Feature extraction is an integral part of Automatic speech recognition system. The performance, quality and accuracy of ASR suffer greatly due to increase in background noises and linear distortions.

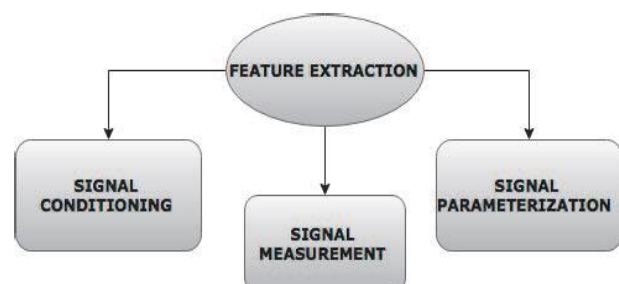


Fig.1: Feature extraction components.

The above figure represents the typical techniques that feature extraction includes-signal conditioning, signal measurement and signal parameterization.[3]

**TECHNIQUE USED TO EXTRACT SPEECH SIGNAL USED IS:**

Here to achieve our requirements there is several methodology i.e.,

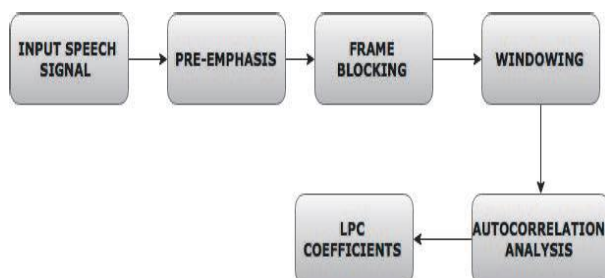
- Linear Predictive Coding
- Relative Spectral filtering.
- Mel-frequency cepstral coefficients .[3]

We are using linear predictive coding for extracting speech signal . As LP is a model based on human speech production. It utilizes a conventional source-filter model, in which the glottal, vocal tract, and lip radiation transfer functions are integrated into one all-pole filter that simulates acoustics of the vocal tract[2].

The principle behind the use of LPC is to minimize the sum of the squared differences between the original speech signal and the estimated speech signal over a finite duration. This could be used to give a unique set of predictor coefficients. These predictor coefficients are estimated every frame, which is normally 20 ms long. The predictor coefficients are represented by  $a_k$ . Another important parameter is the gain (G)[2].

**III. METHODOLOGY**

It is a powerful, robust, accurate, reliable and popular tool for speech recognition, compression and synthesis. The main objective of LPC is frame-based analysis of the input speech signal to generate observational vectors. It is a very simplified method and belongs to spectral analysis part. LPC technique can provide estimation of poles of vocal tract transfer function. Each sample in LPC can be approximated as past samples in linear combination. In order to implement LPC and generate the features the input speech signals needs to pass through pre-emphasizer, the output of pre-emphasizer acts as the input to frame blocking where the signal is blocked into frames of N samples. The next step after frame blocking is windowing where each frame is windowed in order to reduce signal discontinuity at the beginning and end of every frame. Hamming frame is an example of typical frame. After windowing each windowed frame is auto correlated and the highest autocorrelation value gives the order of LPC analysis and finally LPC coefficients are derived.[3]



**Fig 2.:** Block diagram of LPC implementation.

**IV. RESULT AND DISCUSSION**

Automatic Speech Recognition has been under scrutiny for many years but still completely accurate and efficient systems have not been created. In this paper we have studied speech recognition system in depth and also few feature extraction techniques like LPC. It was observed it has technique has its own merits and demerits. There are many limitations with the systems, such as it gets affected by background noise and give less efficient results, it also cannot identify speeches from various users due to speech overlap and the system also undergoes problems while detecting accent and pronunciation of the speakers. It can also be concluded from the study that majority of work in speech recognition has been done for English language and comparatively less work has been carried out for other languages like Arabic and Indian. It can also be observed from the study that for English language the recognition is most accurate and hence the rate is higher as compared to others. Instead of implementing single feature extraction techniques for ASR and there is a need to develop combination of one or more techniques i.e. hybrid techniques that will make the system more reliable, robust and provide more accurate results. Here feature extraction techniques along with detailed study of these feature extraction techniques and the main aim of this paper is to provide researchers working in this area with an understanding of differences between these commonly used feature extraction techniques [3]

**REFERENCES**

- [1] S. Dhingra, G. Nijhawan and P. Pandit, Isolated Speech Recognition using MFCC and DTW, International journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering,8(2), 2013.
- [2] C. Poonkuzhali, R. Karthiprakash, S. Valarmathy and M. Kalamani, An Approach to feature selection algorithm based on Ant Colony Optimization for Automatic Speech Recognition, International journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 11(2), and 2013.
- [3] V. Sharma and P. Sharma, Discrete and continuous Mouse Motion using Vocal and Non-Vocal Characteristics of Human Voice, International journal of Computer Science and Engineering Technology,4,2013.
- [4] Ittichaichareon, S. Suksri and T. Yingthawornsuk, speech Recognition using MFCC, International Conference on Computer Graphics Simulation and Modeling, 2012.
- [5] N.N. Lokhande, N.S. Nehe and P.S. Vikhe , MFCC based Robust features for English word Recognition, IEEE, 2012.
- [6] L. Muda, M. Begam and I. Elamvazuthi, Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping(DTW) Techniques, Journal of Computing, 3(2),2010.
- [7] Anjali, A. Kumar and N. Birla, Voice Command Recognition System based on MFCC and DTW, International Journal of Engineering Science and Technology, 2(12),2010