**FEATURE EXTRACTION FOR SPEECH RECOGNITION**

**ABSTRACT:**

With the advancement of digital signal processing hardware and software, automatic speech recognition (ASR) has advanced significantly. But even with all these advancements, machines still fall short of human performance in terms of accuracy and speed, particularly when it comes to speaker-independent speech recognition. Therefore, speaker independent speech is the subject of a sizable percentage of speech recognition research nowadays recognizing difficulty. Due to the complexity of its applications and the limits of current methods of speech synthesis. In this paper, we briefly go over the speech recognition method known as signal modelling. It is then a summary of the fundamental techniques used in signal modelling. Additional frequently used temporal and we go into great length about spectral analysis feature extraction strategies

**EXISTING SYSTEM:**

One of the most basic ideas in voice processing is this. It can be viewed as an imperfect representation of the early phases of transduction in the human auditory system. Motives for representing filter banks According to "place theory," the position of the basilar membrane's maximal displacement for stimuli like pure tones is inversely correlated with the tone's frequency logarithm. Human perception research have revealed that, unless one of the components of a complex sound goes beyond the bandwidth, its frequencies cannot be independently detected within the range of some nominal frequency. Critical bandwidth is the name given to this bandwidth.

**DISADVANTAGES:**

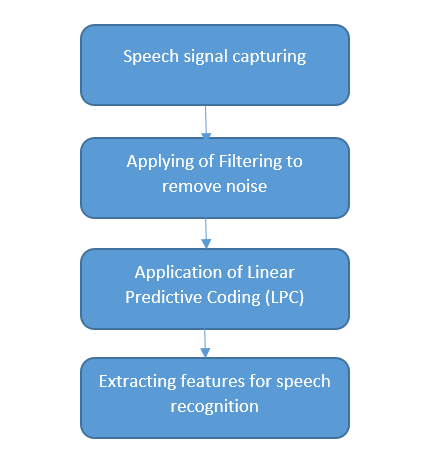
1. Feature extraction using Band filter bank analysis is a complex process.

2. Features extracted using filter bank analysis will generates very complex vector.

3. Feature vector size determines the efficiency of training and testing, filter bank analysis feature vector results in time consuming for training and not that accurate at test results.

**PROPOSED METHOD:**

A voice sample can be roughly represented as a linear combination of prior speech samples, according to the fundamental principle behind the linear predictive coding (LPC). A distinct set of predictor coefficients is established by minimizing the sum of the squared discrepancies (during a finite period) between the actual speech samples and the ones that were linearly predicted. The output of a linear, time-varying system triggered by either random noise or quasi-periodic pulses is used to model speech (during unvoiced speech). The robust, trustworthy, and accurate linear prediction approach offers a way for predicting the variables that define the linear time-varying system that represents the vocal track. The majority of recognition systems use the auto-regressive all-pole model (AR).



**Block diagram of proposed method**

**ADVANTAGES:**

1. The main advantage is the dimensions are reduced compared to other existing feature generating processes.
2. Minimizes time for any algorithm or recognition technique’s training.
3. Extracting the better features that maximizes the accuracy and sensitivity.

**APPLICATIONS:**

1. Machine Learning based techniques
2. Deep Learning based techniques
3. Detection of objects
4. Classification of objects, signals and images

.

**Software & Hardware Requirements:**

**Software:** Matlab R2018a.

**Hardware:**

**Operating Systems:**

• Windows 10

• Windows 7 Service Pack 1

• Windows Server 2019

• Windows Server 2016

**Processors:**

Minimum: Any Intel or AMD x86-64 processor

Recommended: Any Intel or AMD x86-64 processor with four logical cores and AVX2 instruction set support

**Disk:**

Minimum: 2.9 GB of HDD space for MATLAB only, 5-8 GB for a typical installation

Recommended: An SSD is recommended a full installation of all Math Works products may take up to 29 GB of disk space

**RAM:**

Minimum: 4 GB

Recommended: 8