ABSTRACT

Vector quantization techniques play a dominant role in compression of speech signals. There exists a variety of vector quantization techniques, each technique has its own advantages and disadvantages and no vector quantization technique is perfect in all aspects till now. This thesis deals with enhancing the performance of the existing vector quantization techniques using hybrid methods. In this thesis two hybrid vector quantization techniques are proposed which are developed from the existing vector quantization techniques. To find the performance of these vector quantization techniques, they are used in linear predictive coder to reduce the bit-rate of the speech signal, without any considerable loss in the quality of reconstructed speech signal. The performance of the vector quantizer is evaluated in terms of the spectral distortion measured in decibels (dB), computational complexity in flops per frame and memory requirements in floats.

The two vector quantization techniques proposed in this thesis are Switched Multistage Vector Quantization technique and Multi Switched Split Vector Quantization technique. Each of these vector quantization techniques can be implemented in two ways. They are hard decision scheme and soft decision scheme. In this work only hard decision scheme is implemented. The work is carried out using the standard TIMIT database.
The performance of the proposed vector quantization techniques is compared with the performance of the existing vector quantization techniques which are: Split Vector Quantization technique, Multistage Vector quantization technique, Split-Multistage Vector Quantization technique, Switched Split Vector Quantization technique using hard decision scheme.

The linear predictive coder (LPC) involves two key mechanisms which are LPC analysis and LPC synthesis. LPC analysis involves in segmenting the speech signal into frames of short duration, Smoothening the beginning and end of each frame using a window, computation of the linear predictive coefficients for each frame and estimation of information regarding pitch of a frame. The window used in LPC analysis is the hamming window, the method used for the computation of linear predictive coefficients is the autocorrelation method using Durbin’s recursive algorithm and the pitch of a frame is estimated using autocorrelation function with center clipping. LPC synthesis involves the reconstruction of a speech signal using the speech parameters like linear predictive coefficients, pitch and gain.

Vector quantization is the process done in between LPC analysis and synthesis. The speech parameters required for vector quantization are the line spectral frequencies (LSF) and are obtained from the linear predictive coefficients using the real root method. The process of vector quantization for all the techniques involves the generation of codebooks. The codebooks are generated using the Linde, Buzo, Gray (LBG) algorithm. LBG algorithm is a non recursive algorithm that
makes vector quantization a powerful tool for compression of speech signals. The speech parameters used for codebook generation are the line spectral frequencies.

Voice banking is a telephone banking service where the speech recognizer at the bank has to recognize the compressed words from a telephone line. Hence to test the probability of recognition for different coded words, in this work the coded outputs obtained using various vector quantization techniques at a particular bit-rate are applied to a speech recognizer for the purpose of recognition. The Speech recognition technique used for recognizing the compressed words is the Linear Hidden Markov Model (HMM), also called as the left-right model or Bakis model. The speech parameters used for recognition are the Linear Prediction Cepstral Coefficients (LPCC). Speech recognition using HMM requires the estimation of the transition and emission matrices. The transition matrix is estimated using the Baum Welch algorithm, the emission matrix is generated using the K-means clustering algorithm and is estimated using the Baum Welch algorithm. The probability of a particular state sequence is calculated using the forward algorithm. The recognition accuracy is measured in terms of the score of recognition and the recognition rate. Score of recognition is the probability of recognition and the recognition rate is the mean percentage score of recognition.