Int. J. Elec&Elecn.Eng&Telcomm. 2017

International Journal of Electrical and Electronic Engineering & Telecommunications

ISSN 2319 – 2518 www.ijeetc.com
Vol. 6, No. 2, April 2017

Research Paper

A REVIEW OF DIFFERENT SPEECH CODING METHODS

Kaladharan N[1]*

*Corresponding Author: **Kaladharan N***

Speech coding is the act of converting the speech signal in more compressed form, which can be transmittedthrough less number of bits with highest quality. Speech coding is an algorithm used to analyze the special, non- stationary and intelligent speech signal in order to extract its important parameters and to compress it for the maximum utilization of available bandwidth. Speech coding is a lossy type of coding and hence the output signal does not exactly sound like the input. Speech coding is useful for long distance communication, message encryption and quality of speech. Speech coding has been major issue in the area of digital speech processing and telecommunication fields.

Keywords: Speech coding, LPC, Sub band, MELP

INTRODUCTION

Speech coding is a process of obtaining compressed form of speech signal. It allows efficient transmission of speech signal over a band limited wired or wireless channel. Most important applications of several speech coding systems is security. Maximum of the speech coding techniques are constructed on the lossy coding techniques, in which irrelevant information is removed. Speech coding either one enhances the quality of a speech signal at a particular bit-rate or minimizes the bit-rate at a certainquality (Abhinav Kumar, 2014). While reducing the at once in a size has to preserve lucidity and quality of original speech signal. Speech coding is very important in Cellular, teleconferencing, wireless communication, and audio for videophones or video teleconferencing systems and Mobile Communication (Srinivasa Rao P *et al*. (2014).The main objective of speech coding is to represent analog form of speech waveform in to digital one with minimum bits as possible.

This paper is organized as follows, Section II describes survey of speech coding, Section III represent Speech Coding, Section IV

[1] Assistant Professor, Department of Electrical and Electronics Engineering, Annamalai University, India.

Classification of coding techniques, Section V refers performance of evaluation and Section VI gives the Conclusion.

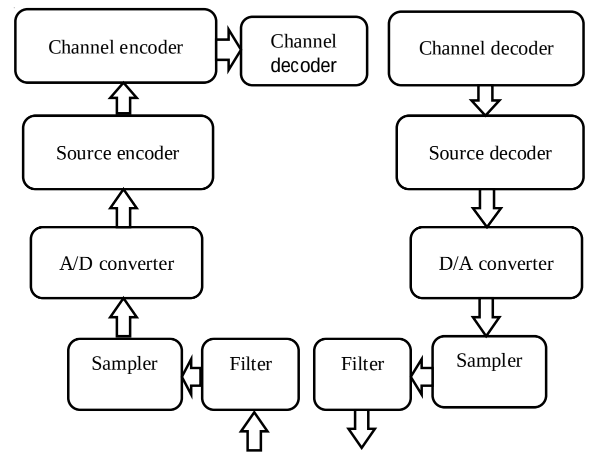
SURVEY OF SPEECH CODING

A speech coder is one which converts a digitized speech signal into anoblique representation and transmits it in the form of frames and receiving end the speech decoder receives the coded frames and performs synthesis to restructure the speech signal. The speech coders divergemostly in bit-rate, intricacy, delay and perceptual quality of the synthesized speech with realistic quality. The factorsshould consider while designing an efficient speech coding mechanism are transmission quality, multiuser capability, low storage requirements, bandwidth maintenance and speech quality enhancement with higher signal to noise ratio (Jagtap S K *et al*. (2015). Several properties for speech codes normally attains in design are Low bit-rate,high speech quality, strength to different speakers languages, channel faults, low memory requests, fewer calculation difficulty, low coding delay etc.Waveformcoders are used in landline telephone services with low complexity and sub band coders are medium in nature and where use in tele conferencing audio. LPC having high complexity having the range of 2.0-4.8 kbps, where used in satellite transmission purpose (Mohit Narayanbhai Raja *et al*., 2015)

SPEECH CODING

Speech coding is a talent of compressing and then converting speech signals for effective transmission or stowage. Speech coding system initially having the input speech signal which is in the form of analog in nature is need to digitized using a sampler, analog to digital converter and a filter. Speech coding techniques are mainly two types which are lossless and lossy coding methods. The lossy coding technique have the reconstructed speech signal perceptually different from the original speech signal however the lossless coding technique, the reconstructed signal at the decoder end has exactly the same character as the input speech signal.Wave form, parameter, hybrid and vocodors coding methods are mostly practicedin speech coding techniques. Wave form coders uses samplebysample coding scheme and preserve its output related to the input waveform (Pretty Varghese *et al.,* 2015). Regularly the speech coding techniques are based on the lossy coding technique for it removes the information which is irrelevant from the perceptual quality[9] [10].The goal track of all speech coders is to minimize the partiality at a given bit rate with high quality.

Figure 1: Block Diagram of Speech Coding



These coders have squat compared to parameter and hybrid coders. Parameter coding model depends on certain speech parameters analysis of these speech parameters before transmission and synthesizing at the receiving end for reconstruction of the original signal is crucialpart (Ravindrbabu A *et al*., 2015).

The block diagram of a speech coding system is shown in Figure 1. Sampler helps to converts the analog speech signal in to discrete signal from the filter region. Discrete item will be given to the Analog to digital converter for the purpose of output is digitized. The digital items are through the source and channel encoder, encodes the value in secure form to the authenticate channel. Then the decoder, decodes the digital input and converts in to analog form as output speech.

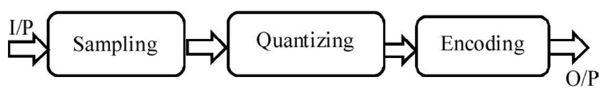
TYPES OF SPEECH CODING WAYS

A. Waveform Coding

Pulse Code Modulation

PCM is a method used to digitally epitomize sampled analogsignals.PCM is the best waveform coding technique which quantizes and encode every sample of speech valueto a finite number of bits. This is the method of analog to digital conversionand also non differential. The operations that it involves are sampling, quantizing and coding based on quantized levels shown in Figure.

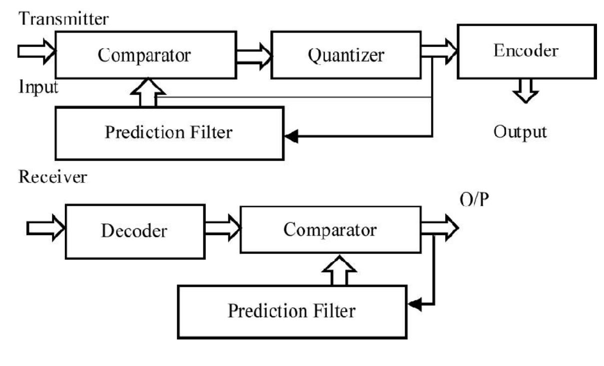
Figure 2: Block Diagram of PCM Coding



Differential PCM

Differential PCM system having the procedure, the sampled input signal is stored in a predictor, and sends it through a differentiator, the differentiator compares the previous sample signal with the current sample signal and sends this difference to the quantizing and coding phase of PCM each samples are compared to a prediction result, and the change is called the prediction residual.

Figure 3: Block Diagram of DPCM Coding



Since the difference between input samples is less than a whole input sample, the number of bits required for transmission is reduced (Yadav Monika S *et al*., 2016). DPCM quantizes the difference signal using uniform quantization. Uniform quantization produces an SNR that is slight for small input sample signals and large for huge input sample signals. Hence, voice quality is well at higher signals.

Adaptive DPCM

ADPCM uses linear prediction that means it uses the preceding samples to predict the current sample. It then calculates the difference between the current sample and its predication and quantizes the difference. Decoder multiplies this number by the quantization step to obtain the reconstructed audio sample. The method used is efficient because the quantization step is updated on the time, by both encoder and decoder, according to the varying amplitude of the input speech samples. Adaptive differential pulse code modulation is improved version of PCM where 4 bits per sample are used encoding the pulse creek. ADPCM codecs require appreciable less storage space. In this both quantizer and predictor are adjust in nature.

Delta Modulation

Delta modulation is a disparity quantization scheme that uses two levels of quantization. By using a single bit to represent each sample, the sample rate and the bit rate are equivalent. Accordingly, sample rate is directly related to signal worth. Applying adaptive techniques to delta modulation, quantizer allows for unceasing step size adjustment. By adjusting the quantization step size, the coder is able to signify low amplitude signals with greater accuracy.

Transform Coding

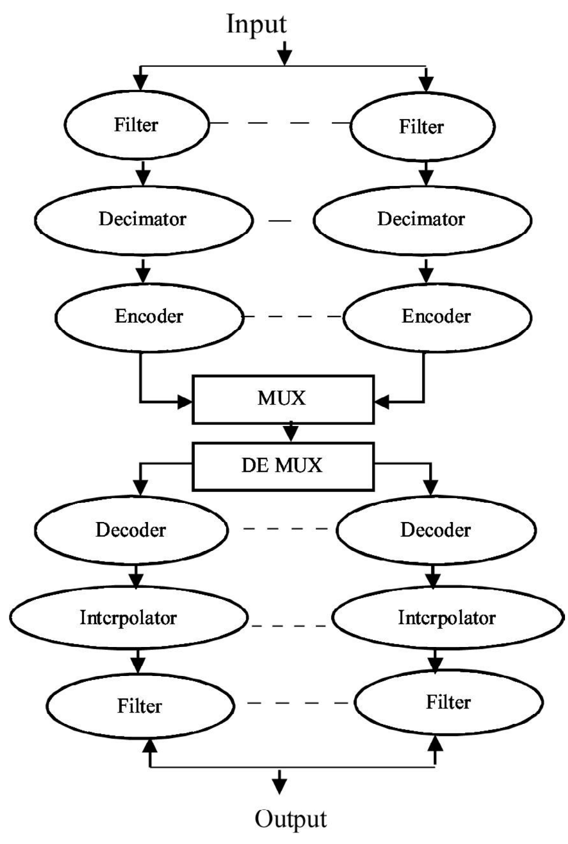
In this method, each section is signified by a set of transform coefficients, which are separately quantized and transmitted. Bit rates in the range 9.6 kbps to 20 kbps .The basic principle of in this of a block of speech samples is functioned on by a discrete entity transform and the resulting transform coefficients are quantized and coded for transmission to the receiver speech waveform.

Sub Band Coding

Sub band coding is the waveform coding spectral domain type. It is the process of decomposing of speech signal, where speech is usually divided into four or eight sub bands by a bank of filters, and each sub band is tasted at a band pass Nyquist rate and encoded with dissimilar accuracy in accordance to a perceptual norms. To reduce number of samples, sampling rate of the signal in each sub band is reduced by obliteration. Sub band coding can be used for coding speech at bit rates in the range of 9.6 Kbps to 32 Kbps. In this range speech quality is roughly equivalent to that of ADPCM at an equivalent bit. The advantage of sub band coding is that each of the sub band can be encoded separately using different coders based on perceptual appearances.

A sub band spectral analysis technique was established that significantly shrinks the complexity of computing the perceptual model.

Figure 4: Block Diagram of Sub band Coding

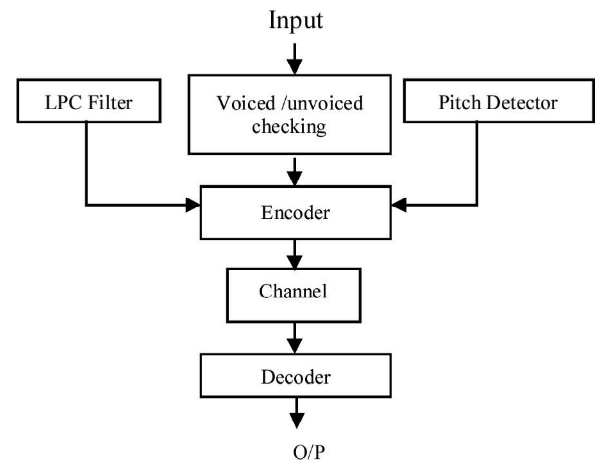


B. Parameter Coding

Linear Predictive Coding

Linear predictive coding, a prevailing, worthy quality, low bit rate speech analysis compression technique for encoding a speech signal. The basic approach is to find a set of predictor coefficients that curtail the mean squared error over a small segment of speech waveform. It has two main components LPC analysis namely encoding and LPC synthesis decoding. The goal line of the LPC analysis is to estimate whether the speech signal is enunciated or tacit, to find the pitch of each frame and to the parameters needed to build the source filter model. These parameters are transmitted to the receiver will carry out LPC synthesis using the received parameters. It is an auto regressive method of speech coding, in which the speech signal at a precise instant is represented by a linear sum of the previous samples linear prediction estimates the current sample by conjoining past few samples linearly. Although auto correlation and covariance methods have been mostly used to determine LP coefficients. Speech coding or compression is generally conducted with the use of voice coders or vocodors.

Figure 5: Block Diagram of LPC



Mixed Excitation Linear Predictive

The Mixed Excitation Linear Prediction algorithm is the linear prediction based parametric speech coder normally works at the range of 2.4 kbps. The MELP vocodors is based on the customary LPC parametric model, but also includes four additional topographies. These are mixed excitation, aperiodic pulses, pulse dispersion, and adaptive spectral enhancement. Aperiodic pulses are most repeatedly used during alteration regions between voiced and unvoiced segments of the speech signal. . In MELP speech coding the input speech signal limits are predictable first which are then used to combine speech signal at the output.

Residual Excited Linear Prediction

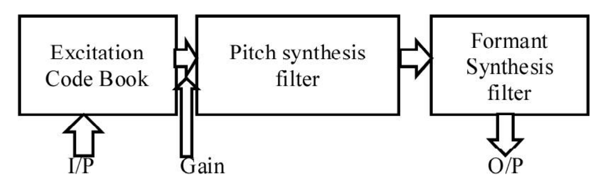
The Residual Excited Linear Prediction vocoder is one of the Linear Predictive Coding based vocodors. Its firmness rate is moderate because the RELP vocoder needs to encode a sequence of residual signals for exhilarating the vocal tract model synthesized from speech signals. However, the quality of synthesized speech is greater to other kinds of LPC vocodors. The system is healthy since there is no need to analyze whether the sound is voiced or unvoiced nor to analyze the pitch epoch.

C. Hybrid

Coded Excited Linear Prediction

CELP is a fit structured closed loop analysis by synthesis hybrid coding technique which combines the advantages of both techniques waveform and parametric to afford a robust

Figure 6: Block Diagram of CELP



low bit speech coder for narrow band and medium band speech coding.

Impression of CELP was intuitive as a stab to improve on LPC coder. Most popular coding systems in the range of 4-8 Kbps bit rate use CELP (Rhutuja Jage *et al*., 2016). Examining of an excitation codebook to offer a consistent excitation structure during encoding is the vital behindhand CELP functioning. This technique is widely used for fee quality speech at 16Kbps. Presently the CELP is used very effectively in MPEG-4 audio speech coding conversion.

Conjugate Structure Algebraic Code Excited Linear Prediction

It is the utmost present hybrid coding technique which is currently positioned in almost all the latest VoIP applications. It operates at 8Kbps and provides near Toll quality performance of the speech signal. The coder is based on code excited linear prediction model. The CS-ACELP coder procedures input signals on a frame by frame and sub frame by sub frame root. The algorithm exploits vector quantization method, both the adaptive and fixed codebook are vector quantized to form conjugate structure.

D. VOCODORS

Vocodors system is based on the analysis synthesis technique, used to imitate human speech. The vocodors was originally developed as a speech coder for telecommunications applications, for the purpose of being to code speech for transmission. The vocodors are further classified as channel vocodors, formant vocodors.

PERFORMANCE OF EVALUATION

Signal to Noise Ratio

Signal to noise ratio is a compute used in signal processing that evaluates the level of a desired wave to the level of background noise .Signal-to-noise ratio is sometimes used to refer to the part of useful information to extraneous data in a discussion or replace. Signal to noise ratio is defined as the power percentage between a noteworthy background noises.

Mean Opinion Score Measure

The mean opinion score presents an arithmetical measure of the quality of human speech. The system uses skewed tests that are mathematically averaged to obtain a quantitative indicator of the system performance. MOS is determined, a number of listeners rate the quality of test words by male and female speakers. A listener gives each sentence a rating as follows: 1.Bad 2. Poor 3. Fair 4. Good 5. Excellent. The MOS is the arithmetic mean of all the individual scores, desired by the value from bad to best.

CONCLUSION

In this paper, four broad categories of speech codes were studied namely waveform coding, parameter type, hybrid method and vocodors .The ultimate goal to design a speech coder is to achieve the best possible speech quality low bit rate, with limitations on intricacy and delay. Each coder has its own advantages and weaknesses. As per this review most of the speech coding system has suggest the CELP and LPC method speech coding is efficient. Hence, in future use and apply the CELP as well as the LPC method by hybrid technique approach from that which result is better for a best speech coding.

REFERENCES

1. Abhinav Kumar (2014), "Study of Different Types Coders for GSM", *International Journal of Engineering and Technical Research*, Vol. 2, No. 2.
2. Bankar Santosh F *et al*. (2015), "Residual Excited Linear Predictive Coding", *International Journal of Engineering Research & Technology*, Vol. 4, No. 5, May.
3. Gibson Jerry D (2016), "Speech Compression", *Information*.
4. Jagtap S K *et al*. (2015), *Speech Coding techniques*, Science Direct.
5. Kiran Chavan *et al*. (2016), "Speech Coding Techniques", *International Journal of Technical Research and Applications*.
6. Manas Ray *et al*. (2015), "Speech Coding Techniques for VoIP Applications: A Technical Review", *World Applied Sciences Journal,* Vol. 33, No. 5, pp. 736-743.
7. Milos Cernak *et al*. (2016), "Cognitive Speech Coding", IDIAP Research, November 07.
8. Mohit Narayanbhai Raja *et al*. (2015), "Linear Predictive Coding", *International Journal of Engineering Sciences & Research Technology*, April.
9. Nimisha Susan Jacob *et al*. (2015), "Performance Analysis of CS-ACELP Speech Coder", *International Journal of Engineering and Advanced Technology*, Vol. 4, No. 5.
10. Oliver Jokisch, *et al*. (2016), "Audio and Speech Coding & Transcoding in Web Real-Time Communication".
11. Pretty Varghese *et al* (2015), "Advanced Voice Excited Linear Predictive Coding With Noise Reduction", *International Journal of Current Engineering and Technology*.
12. Priyanka Kadam *et al* (2016), "Voice Coding Technique: A Review", *International Journal of Multidisciplinary Research and Information*.
13. Rashmirekha Ram *et al*. (2015), "Speech Coding Techniques: A Comparative Study", *IJECT*, Vol. 6, No. 3, September.
14. Ravindrbabu A *et al*. (2015), "Speech Signal Analysis and Synthesis by Using LPC Technique", *International Journal of Engineering Trends and Technology*, Vol. 22, No. 1, April.
15. Rhutuja Jage, *et al*. (2016), " Implementation of CELP and MELP Speech Coding Techniques", *International Journal of Advanced Research in Computer and Communication Engineering,* Vol. 5, No. 6, June.
16. Srinivasa Rao P *et al*. (2014), "Speech Coding", *International Journal of Future*

Generation Communication and Networking, Vol. 7, No. 5.

17. Sukriti Sharma *et al*. (2015), "Lossless Linear Predictive Coding for Speech Signals", *International Journal of science, Technology & Management*, Vol. 04, Special Issue No. 01, March.

18. Yadav Monika S *et al*. (2016), "Speech Coding Technique and Analysis of Speech Codec Using CS-ACELP", *International Journal of Engineering Science & Advanced Technology*, Vol. 6, No. 3.