International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

Speech Compression Using LPC-10 and VELPC Based on DCT Technique Compression

Aditi Kar[1], Rasmi Ranjan Samantray[2]

[1]M Tech Scholar, CCEM Kabir Nagar Raipur, Affiliated to CSVTU, Bhilai, Chhattisgarh, India

[2]Assistant Professor, CCEM, Raipur, CCEM Kabir Nagar Raipur, Affiliated to CSVTU, Bhilai, Chhattisgarh, India

Abstract: ***In this paper, we have presented the speech compression techniques i.e. Linear Predictive Coding and its modified version Voice Excited Linear Predictive Coding. Output of LPC-10 is compared with output of VELPC. Here instead of using plain VELPC we are using modified VELPC or we can say that VELPC based on Discrete Cosine Transform technique. The result shown here was obvious since VELPC is modified version of LPC and so its output quality is better as compared to output quality of LPC. LPC is one of the most powerful speech coding techniques and it is widely used for encoding speech signal at a very low bit rate. Pitch period estimation is the weakest link in the LPC method so this drawback is eliminated in VELPC technique. Thus, the bit rate however is somewhat increased in VELPC technique. Therefore Discrete Cosine Transform is used to compress the innovation signal in case of VELPC in order to reduce bit rate to some extent.***

Keywords: Speech signal, Discrete Cosine Transform (DCT), Linear Prediction Coding (LPC), Voice Excited Linear Prediction Coding (VELPC), Bit rate

1. Introduction

Linear Predictive Coding is based on Linear Prediction approach in which the present sample is estimated based on the linear prediction of past samples. The analog signal at the input is first analyzed then the parameters of the analyzer are encoded and transmitted. Then the synthesizer reconstructs the approximated speech signal. The analyzer calculates each frame for respected parameters in the following manner- pitch period estimation, voiced/unvoiced flag, and gain parameter and at last filter coefficients. VELPC is a modification of LPC. Same procedure is followed but in VELPC an innovation signal is transmitted to the receiver along with the filter coefficients.

2. Problem Statement

Whenever there is a sudden rise in the peak of the signal then the prediction error would be large, Since LPC prediction is based on the past samples and the difference between peak of past sample and peak of present sample would be more so this pitch period estimation is the weakest link of LPC. So we need accurate values of pitch period for better prediction. Thus, VELPC gives better results in terms of quality. Although the bit rate would little bit increase in comparisons with LPC output but then also the quality is maintained in VELPC and use of DCT is to compress the innovation signal so as to lower bit rate to some extent.

3. Methodology

These speech coding techniques represent human speech model. In human body, the sound is generated from lungs and the intensity of air coming out of lungs decides the intensity of the voice. If the voice is loud then the pitch is high and so it will be considered as voiced excitation while the intensity of sound is least then the pitch will be zero or low then it will be considered as unvoiced or noise. This all is done inside the vocal cord. Coming out of the vocal cord these voices are shaped or articulated in the vocal tract. These articulations along with the gain are considered as the innovation signal.

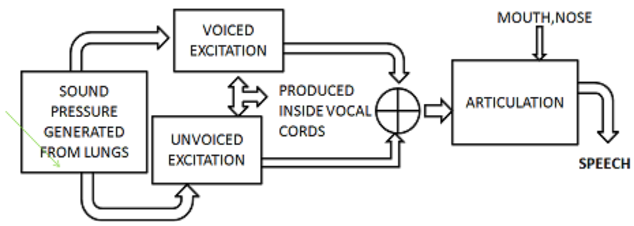

Figure 1: Human speech production model

LPC is a type of parametric coding. In LPC-10, 10 filter coefficients are carried out in LPC analyzer along with the other parameters such as gain, voiced/unvoiced, pitch period. At LPC synthesizer, the reconstructed voiced signal values are excited with impulse train and unvoiced values with white noise which are then combined with the energy of the sound and proceeded to LPC filter.

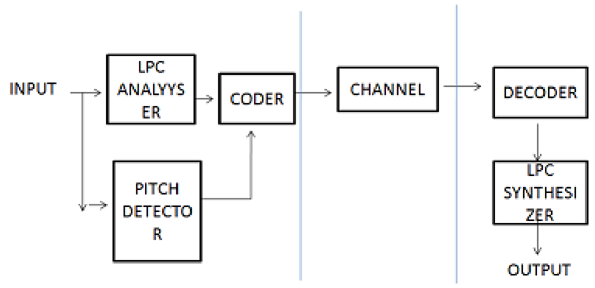

Figure 2: Block diagram of LPC technique

VELPC is a combination of parametric coding as well as waveform coding. Same procedure is followed in VELPC as in LPC but in VELPC the pitch parameter and voicing decision is not evaluated rather an innovation signal or we can say that low pass filtered version of the input signal is carried along with the filter coefficients. Essentially a good pitch period estimation is required in order to have a proper synthesis. DCT is used to compress the innovation signal due to the nature of the source coding i.e. compression. It concentrates most of the energy into its first few coefficients Therefore, extraction of first few coefficients are done.

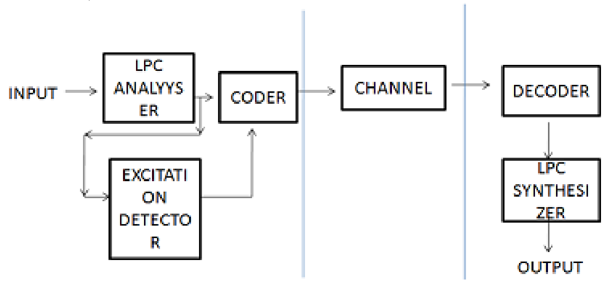

Figure 3: Block diagram of VELPC technique

4. Literature Survey

Khare, Gaurav, Prashant Shekhar, M. Kulkarni et.al [1] titled "Generation of excitation signal in Voice Excited Linear Predictive Coding using Discrete Cosine Transform." Provided method to generate the excitation parameters and gain that provides improvement over plain Linear Predictive Coding(LPC). They employed DCT computation to transmit the excitation signal energy.

Raza, M. Ahsan, Dr. Parvez Akhtar. et.al [2] titled "Implementation of Voice Excited Linear Predictive Coding(VELPC) on TMS320C6711 DSP kit using Simulink RTW(Real Time Workshop) which explained Simulink model of VELPC analysis, VELPC synthesis. This paper briefly explained as the signal is passed through analyzer the filter coefficient and residual signal are generated.

Gupta, Harshita, Divya Gupta. et.al [3] titled LPC and LPCC Method of Feature Extraction in Speech Recognition System." Presented about the automatic speech recognition system According to this system Speech Recognition is the ability to listen what we speak, interpret and perform action according to spoken information. This paper shows the comparative analysis of LPC and LPCC.

Viswanathan, R, W. Russell, J Makhaul. et.al [4] titled "Voice-Excited LPC coders for 9.6 kbps Speech Transmission." considered the use of VELPC for transmission of speech at a bit rate of 9.6 kbps. This paper presented various aspects of speech coders with the goal of maximizing the speech quality. They discussed aspects of speech coders with the goal of maximizing the speech quality. They discussed aspects such as baseband residual verses baseband speech transmission, coding of the baseband signal and high frequency regeneration from the baseband.

Sen, Tosha, Kruti Jay Pncholi, et.al [5] titled "Speech Compression using Voice Excited Linear Predictive Coding." aimed at designing good quality encoder and decoder for long distance communication. They used VELPC and suggested LPC as one of the most powerful speech analysis technique that has become predominant technique for representing speech for low bit rate transmission and storage.

Bharadwaj B, Sai, U. Vagdevi, Yadav Bineetha. et.al [6] titled "Linear Prediction Analysis" discussed about the all pole i.e. the Linear Prediction model. In this paper they used linear prediction to analyze the sound files (A bird chirping, A dog barking, A girl speaking in English and a male speaking in Chinese). and compared predictability of sounds from animals and different languages. This paper aimed at implementation of the Levinson-Durbin algorithm to analyze the four sound files. They used frame size of 20ms to obtain the required results.

5. Equations

In LPC, instead of transmitting the input signal we are sending the error signal and slowly varying parameters a_1 to a_{10} . Now the question arises that why we are sending the error signal e(n) instead of input signal s(n). The answer is that the variance of e(n) is much lower as compared to variance of s(n). We know that a linear predictor estimated the present sample based on the past sample values. So, we can write error signal as

$$e(n)= s(n)- \sum_{k=1}^{p} a_k\, s(n-k) \qquad (1)$$

Here, e(n) is the error signal. s(n) is the present speech sample, s(n-k) is previous sample value, p is the order of the filter and k is index value which varies from 1 to 10 in case of LPC-10 and a_k are LPC filter coefficients varying from k=1 to 10.

To calculate the values of coefficients i.e. a_1 to a_{10}. We have to find MSE i.e. Mean Square Error value.

$$MSE=\sum_n [s(n) - \sum_{k=1}^{p} a_k\, s(n-k)]^2 \qquad (2)$$

Filter coefficients are calculated by differentiating MSE with respect to a_j and equated to zero to get the perfect match for filter coefficients.

$$\frac{\partial MSE}{\partial a_j}=0 \qquad (3)$$

Rearranging and computing the above equation and performing Levinson Durbin recursion for matrix solving and finally we get filter coefficients for each frame.

In LPC, we transmit this filter coefficients with pitch parameter, voicing decision and gain of the signal whereas in VELPC we transmit these filter coefficients with innovation signal. In VELPC, for generation of innovation signal i.e. the low pass filtered version of the input signal we are using a moving average filter and the compression of innovation signal is done by DCT technique. By using low pass filter, higher formants are filtered out. Moving average filter is a simple low pass FIR filter which is generally used for regulating an array of sampled data. It takes M samples of input at a time and takes the average of those to produce a single output point. As the length of the filter increases the smoothness of the output increases.

We know that an speech sample is a sequence of real numbers i.e. X={x_1, x_2 , x_3, ------------------------x_N} and the dct of these samples is Y={y_1, y_2 , y_3, ------------------------ y_N} such that

$$x_n = \sum_{k=1}^{n} y_k\, w(k) \cos\left(\frac{\pi(2n-1)(k-1)}{2N}\right) \qquad (4)$$

Where $w(k)=\begin{cases} \frac{1}{\sqrt{N}}, & k=1 \\ \sqrt{\frac{2}{N}}, & otherwise \end{cases}$

k= bit range from 1 to n
n= index value
N is typically

6. Conclusion

The pros and cons of LPC technique have been explained and based on the output quality we can hereby conclude that VELPC technique is better than LPC technique however with an increased bit rate. For that purpose we used DCT technique to compress the innovation signal in VELPC technique. The result is shown in the below figure.

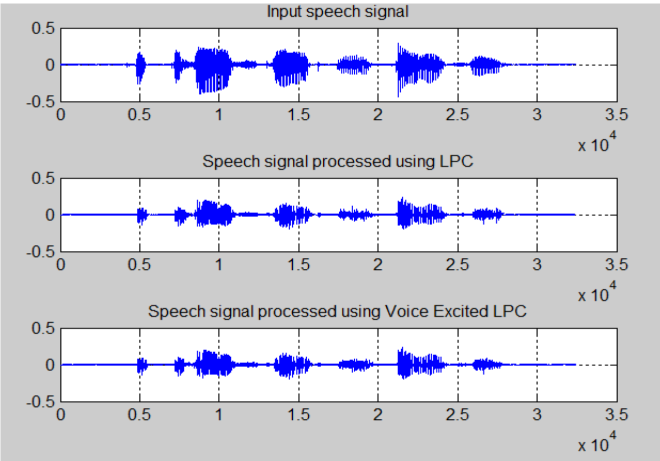

Figure 4: Plot for LPC output and VELPC output

References

[1] Khare, Gaurav, Prashant Shekhar, M. Kulkarni. "Generation of excitation signal in Voice Excited Linear Predictive Coding using Discrete Cosine Transform." TENCON 2005 IEEE Region 10, IEEE, 2005.
[2] Raza, M. Ahsan, Dr. Parvez Akhtar. "Implementation of Voice Excited Linear Predictive Coding (VELP) on TMS320C6711 DSP Kit." 9th International Multitopic Conference, IEEE INMIC 2005.
[3] Gupta, Harshita, Divya Gupta. "LPC and LPCC Method of Feature Extraction in Speech Recognition System." Cloud System and Big data Engineering (Confluence), 2016 6th International Conference. IEEE 2016.
[4] Viswanathan, R., W. Russell, J Makhaul. "Voice-Excited LPC Coders for 9.6 kbps Speech Transmission." Acoustics, Speech and Signal Processing, IEEE International Conference on ICASSP'79, IEEE 2003.
[5] Sen, Tosha, Kruti Jay Pancholi. "Speech Compression using Voice Excited Linear Predictive Coding." JETIR(ISSN-2349-5162). Volume-2, Issue-6, June 2015.
[6] Bharadwaj B, Sai, U.Vagdevi, Yadav Bineetha. "Linear Prediction Analysis." The International Journal of Engineering and Science(IJES). Volume-2, Issue-4, April 2013.
[7] L.R. Rabiner and R.W. Schafer, "Digital Processing OF Speech Signals", Prentice-Hall, Englewood Cliffs, NJ,1978.
[8] Proakis, John G.: 'Digital Communication' (Tata McGraw Hill, Third Edition).
[9] http://nptel.ac.in/syllabus/117105081/(NPTEL Lectures on Digital Voice and Picture Communication).