https://doi.org/10.36375/prepare_u.a88

Analysis of Different Aspects of Speech Signal Using Delta Modulation Technique

Subhadeep Basu[1], Bubai Maji

Department of Electronics and Telecommunication Engineering

IIEST Shibpur

1: subhadeepbasu.pg2018@telecom.iiests.ac.in

Abstract: Speech signal is analog in nature. The maximum frequency component present in human voice signal can be taken approximately as 3-4KHz.In many modern day applications, it is often required to process the speech signal in digital domain. For this reason, digital processing of analog speech signal is highly important. Delta Modulation is a scheme that can be used to implement digital representation of analog speech signal. In this work, the speech signal is modulated and demodulated using a very simple technique, known as Delta Modulation(DM). Moreover, several other aspects like silence removal, Signal to Noise ratio (SNR) calculation and Power Spectral Density(PSD) analysis has also been carried out.

Keywords: Speech Signal; Delta Modulation; Silence removal; PSD Analysis.

INTRODUCTION

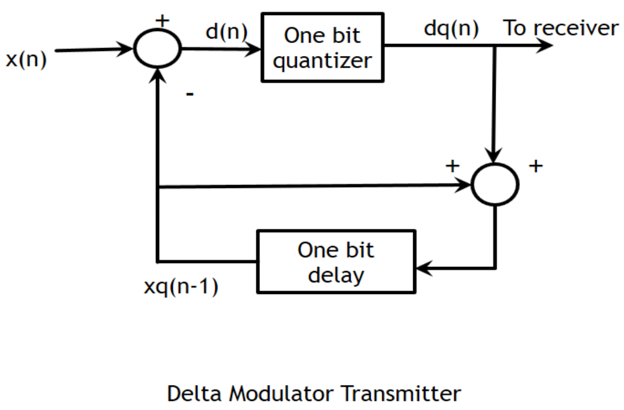
Speech signal consists of sequence of sounds. These sounds carry the information that need to be conveyed. The speech originates from human vocal system that consists of two main parts: the vocal cords(glottis) and vocal tract [1]. Speech signals are very much non-stationary with wide range of multiple frequency components. Any speech signal can be classified in two parts. These are known as voiced sound and unvoiced sound. Voiced sounds are produced by forcing air through the glottis with the vibration of vocal folds [2]. This results in quasi-periodic air puffs that excite the vocal tract. On the other hand, unvoiced sounds are produced by forming a constriction in the vocal tract and forcing air through this constriction ta a high velocity to produce turbulence [2]. This unvoiced parts do not contain any significant information. Hence these parts can be removed from the original speech signal. This is also known popularly as Silence Removal [3-5]. Due to the presence of multiple frequency components in speech signal, it’s difficult to modulate it in digital form. One of the most fundamental way to modulate an analog signal is to consider the Pulse Coded Modulation(PCM) scheme. The limitation to modulate and demodulate voice signal using PCM scheme is bandwidth requirements. PCM requires much larger bandwidth which makes it inefficient for this purpose. There is another popular way to modulate speech signal using Differential Pulse Coded Modulation(DPCM) technique. This is an efficient scheme but the drawback is that this scheme requires complex circuitry. In this regard, Delta Modulation(DM) can be considered as one of the efficient scheme which overcomes the limitations of large bandwidth requirements and also simultaneously involves simple circuitry [3]. If the sampling interval in DPCM is reduced considerably i.e. if we sample a signal in much faster than Nyquist sampling rate, then two adjacent samples of the signal will have high correlation. Thus the difference in the amplitude levels of two adjacent samples will be very less. In case of Delta Modulation, instead of sampling and quantizing the original analog signal, the difference between two adjacent samples are quantized. Hence it’s also known as 1-bit DPCM scheme [3]. Like any other schemes, DM also employs several steps in modulation viz. sampling, quantization, encoding etc. Hence this scheme is also not free from noise which results from quantization. This is also known as quantization noise. Signal to Noise ratio(SNR) is a measure that

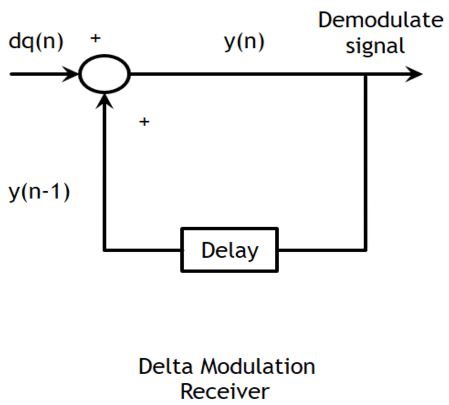
https://doi.org/10.36375/prepare_u.a88

compares the level of desired signal to the level of noise. Power Spectral Density(PSD) is also a popular parameter in speech signal processing. It gives an estimate that how the power of the signal is distributed over the frequency range. In this work, speech signal analysis using Delta modulation and demodulation has performed. Moreover, SNR calculation, silence removal has also been performed.

MODELLING AND SIMULATION

In this study, voice signal has been recorded in microphone and that voice signal has been used as the input analog signal. The work is performed using MATLAB (R 2016). The maximum frequency component present in the voice signal is 3.4KHz. This is a short duration (8 sec) speech signal. This speech signal is sampled at 48KHz, which is much higher than Nyquist sampling rate. After sampling, different samples with varying amplitude levels has been obtained. These samples are normalized. It can be noticed that there are several samples whose amplitude levels are very low. The modulator and demodulator of Delta Modulation is shown in the Fig 1.





(a) (b)

Fig 1: (a) Block diagram of Delta Modulator Transmitter; (b) Block diagram of Delta Modulator Receiver.

These are identified as the silence in the speech signal. Thus the samples whose amplitude levels are less than 0.03V are rejected from this input speech signal. As this signal is sampled at a rate much higher than Nyquist sampling rate, the normalized amplitude levels of two adjacent samples are nearly the same. In this technique the difference between two samples are quantized.

RESULTS AND DISCUSSIONS

The original voice signal contains many frequency components with different amplitude levels. Hence the amplitude value is normalized. The fig shown below explains this. This speech signal contains various unvoiced portions. These do not contain any information. Hence the silence portions are removed from this input signal. Fig 3 explains and compares two speech signal with and without silence removal. The step size is taken as 0.1V in Delta Modulation.

https://doi.org/10.36375/prepare_u.a88

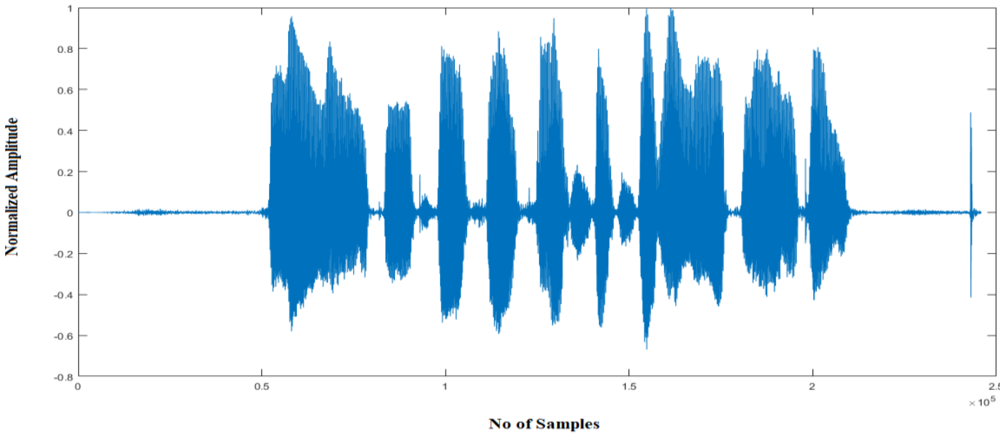


Fig 2.: The normalized input speech signal.

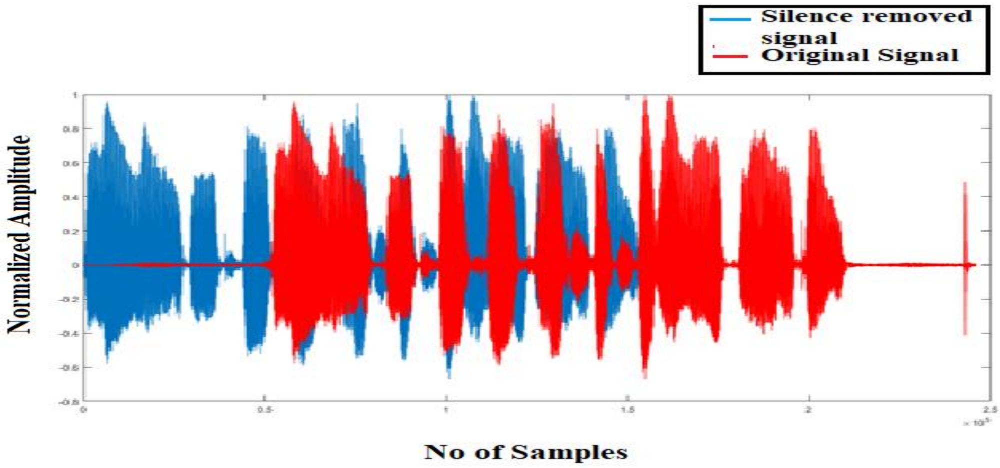


Fig 3: The figure shows original voice signal and signal after silence removal.

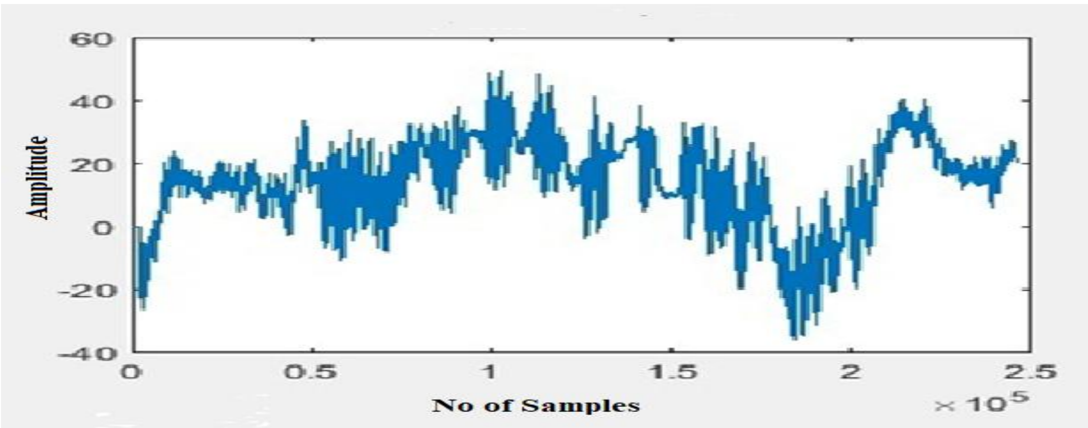


Fig 4.: Demodulated speech signal

As Delta Modulation technique is employed, the difference in the amplitude levels of two samples of the speech signal will be very low. The demodulated signal is mixed with noise. This noise comes from the quantization noise and also from some external noise sources. The above fig(Fig 4) shows the demodulated speech signal. The SNR calculation is also carried out [3]. f_s and f_m are taken as sampling frequency(48KHz) and maximum frequency(3.4KHz) of speech signal. In case of Delta Modulation, the expression of SNR is as follows:

$$\text{SNR}= ((\frac{3}{80}) * (\frac{fs}{fm})^{3} \qquad (1)$$

In dB scale this can also be expressed as 10*log (SNR). In this case, the value of SNR is calculated as 20. 23dB.The PSD estimation of the speech signal can be obtained if Welch PSD technique is employed. The following figure (Fig 5) explains this.

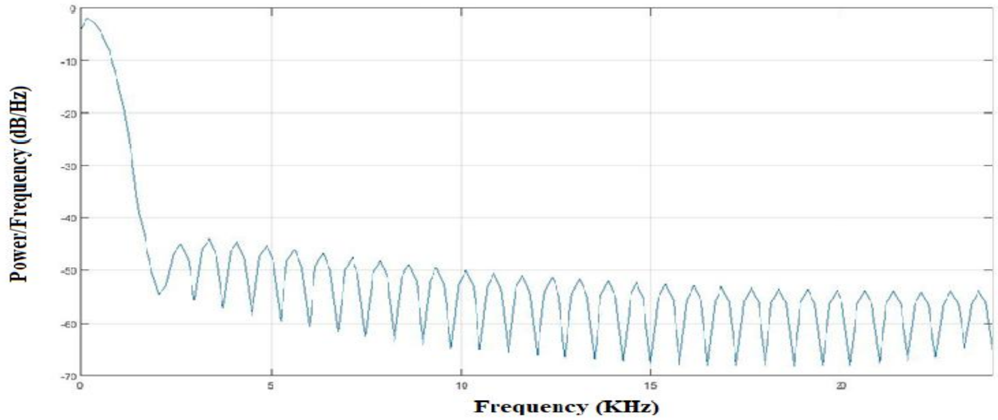


Fig 5.: PSD Analysis of modulated speech signal.

From the above figure, it is clear that the maximum energy content of speech signal is lying within 2KHz frequency range.

CONCLUSION

In this communication the analysis of speech signal using Delta Modulation scheme is presented. This study also deals with SNR analysis, PSD calculation and a comparative analysis of different auto correlated samples. The demodulated signal resembles with the original voice signal. But demodulated signal is mixed with quantization noise. The quantization noise can be reduced using appropriate techniques. Linear predictive coding(LPC) of speech signal can be employed for further improvement. There are also various nonlinear predictive coding techniques that can be employed [5].

REFERENCES

[1] Ming Yang, "Low bit rate speech coding", Potentials, IEEE, vol 23, No. 4, Oct-Nov. 2004 pp. 32-36.

[2] K. Satyapriya, and D. Yugandhar, "Performance Analysis of Speech Coding Techniques", *International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering*, vol. 2, no. 11, pp. 5725-5732, 2013.

[3] S. Haykin, “Communication Systems”, Fourth International Edition, Wiley Student Edition, 2012.pp 218-224.

[4] J.D. Gibson, "Speech coding methods, standards and applications", Circuits and Systems Magazine, IEEE, Vol. 5, No. 4, fourth Quarter 2005. Pp. 30-49.

[5] M. Hasegawa-Johnson, A. Alwan, "Speech Coding: Fundamentals and Applications", *Encyclopedia of Telecommunications*, 2003.