Digital Signal Representation of Speech Signal

Mrs. Smita Chopde [1], Mrs. Pushpa U S[2]
[1,2.] EXTC Department, Mumbai University

***Abstract*— Delta modulation is a waveform coding techniques which the data rate to alarge extent in data communication ; the problem encountered in delta modulation is the slope overload error , which is inherent in the system. In order for the signal to have good fidelity, the slope –overload error need to be as small as possible . Hence there is need for adaptive techniques to be applied to delta modulation to reduce noise .Adaptive delta modulation reduce the slope overload error at the same time increase the dynamic range and the tracking capabilities of fixed step size delta modulation . The adaptive algorithm adjust the step size (from the range of step size) to the power level of the signal and thus enhance the dynamic range of the coding system. This paper discusses the experiment worked using quantization delta modulation and adaptive modulation and their improvements with each other .**

***Keywords*— Quantization, Delta modulation, Adaptive delta modulation ,Digital communication and Signal Processing .**

I. INTRODUCTION

The endeavour to communicate among themselves is a prime action of the human beings. Desire to convey one'spoint of view to relate one to another is a never ending process. Of all forms of communication ,speech has evolved ,over along period of time as the primary means of communication between human being.The convenience and popularity of communication through speech is due to few features of speech.

- Speech is natural
- Speech is efficient
- Speech is flexible
- Speech is robust
- Speech signals are easily manageable

A. Mechanism of Speech Production

To communicate , a speaker must produce a speech signal in the form of a sound pressure wave that travels from speaker mouth to the listener's ears.The communication process has several stages . The process begins with the thoughtthat the speaker wants to speak to a listeners , the thought is converted into a linguistic structure by choosing appropriate words and phrase to express its meaning. The word is expressed as a sequence of sound units local and global sound effects and finally human brain issues several commands to form suitable articulatory gesture producing acoustic vibration in the physical world . The acoustic vibration are captured at the listeners ears and results in the perception of speaker intended meaning into listeners 's mind.

Thus in the whole the process of speech originates at the linguistic level in the speakers mind , descends to physiological level during pronunciation and then to acoustic level during transmission. The listener brings it back to physiological level during hearing and finally the sensation produce at the inner ear ends at the same linguistic level in the listener's brain .

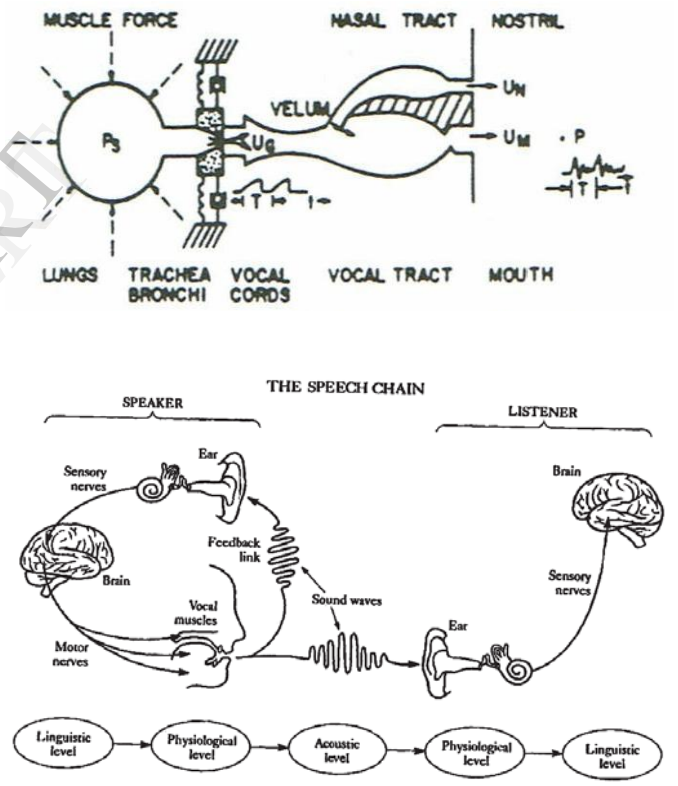


Figure1speech chain

B. Anatomy and Physiology of Human Speech production system.

Speech is an acoustic sound pressure wave that originates from coordinated movements of anatomical structures which make human speech production system. The components of this system are lungs trachea(windpipe),larynx(organ of voice production),pharyngeal cavity(throat),oral cavity (mouth) and nasal cavity(noise). The pharyngeal and oral cavities are

grouped together and referred as vocal tract .Vocal cords velum,tongue ,teeth,jaws ,and lips make together speech production system and are called as articulators.

.A principal features of speech sounds is the manner of excitation to the vocal tract. Voiced and unvoiced excitation are the elemental components for excitation..During normal speech production ,air pressure builds up below the larynx , by the efforts of muscles and lungs ,which stats pushing the vocal chord apart. The glottal slit begins to open and accumulated air is forced up from the lungs through the glottis (the space between vocal folds) and through the possible narrowing in the vocal tract . Increased airflow reduces local air pressure and vocal cords try to close due to their elastic tension . This cycle repeats and the vocal cords get positioned in the way the air flows through the glottis setting the vocal folds into vibration. The sound produced in the process is called as voiced sounds and the time period between successive vocal fold opening (rate of vibration)is referred to as fundamental frequency or pitch .

Unvoiced sounds are produced when vocal tract is sufficiently closed by formation of constriction at the same point along it and air is forced through this producing turbulence as for the word fish during the last sound.

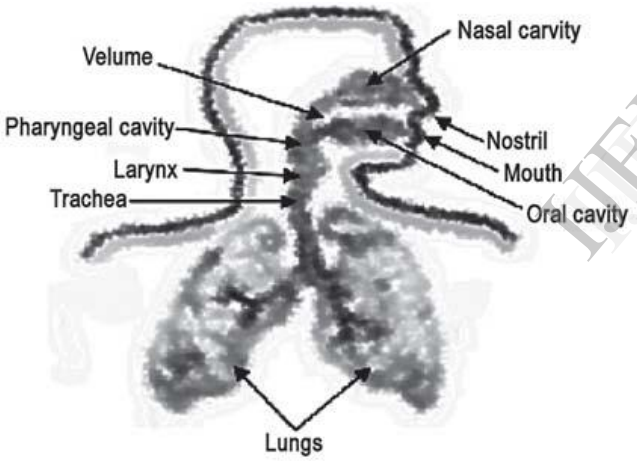


Figure2:Anatomy of human speech production

II. METHODS

A. Uniform Quantization

The quantization ranges and levels may be choosen in a variety of ways depending on the intended applications of the digital representation. With uniform quantization , the dynamic range of the signal R is divided into L equal sized intervals each with length Δ the quantization step- size. The input(unquantized value) and the output (quantized value)relationship in uniform quatizer is shown in figure3In figure3 x_i represents the right boundary of interval i and $\widehat{x_i}$ the quantization level of this interval.

$$x_i - x_{i-1} = \Delta \quad (3.1)$$

$$\widehat{x_i} - \widehat{x_{i-1}} = \Delta \quad (3.2)$$

Any value in the *i* –th interval is mapped into the middle in this interval , i.e.

$$Q(x) = \widehat{x_i} = X_{min}+(i-1)\Delta+\Delta, \text{ if } x_{i-1}<=x<x_i \quad (3.3)$$

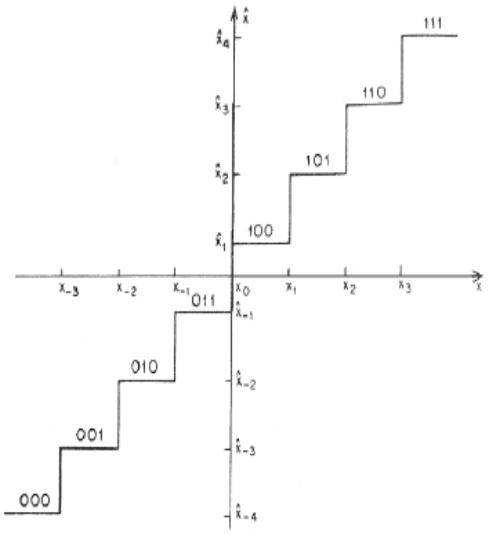


Figure3 3-bit quantizer

Given the signal range R a uniform quantizer has only one parameter ; the number of levels N or the quantization step size Δ and related by

$$\Delta = R/N \quad (3.4)$$

The number of levels N is generally chosen to be of the form 2^B so as to make the most of the efficient use of B –bit binary code words. If the signal has symmetrical probability density function so that $|x(n)| \leq X_{max}$, or R=2 X_{max} then

$$\Delta = 2\,X_{max}\,/2^B \quad (3.5)$$

The quantized samples are represented as

$$\widehat{x(n)} = x(n) + e(n) \quad (3.6)$$

x(n) is the unquantized sample e(n) quantization error or noise.

Signal to Noise Ratio (SNR)=6B _ 7.2(db) (3.7)

This shows that every additional bit contributes 6db improvement in SNR . the actual SNR for the given value B depends on the relation between Xmax and σ_x which depends on the probability distribution of the signal . In order to improve the fidelity of representation with uniform quantization so that it is acceptable perceptually , it is necessary to use more bits .

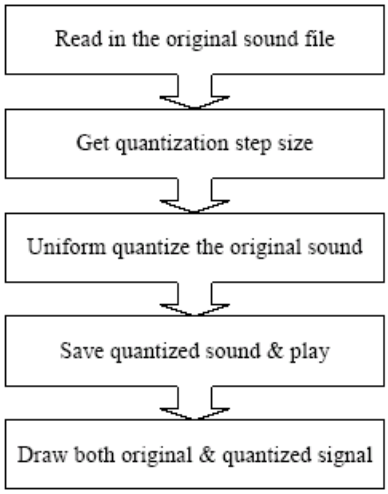


Figure4. The scheme representing the process of quantization

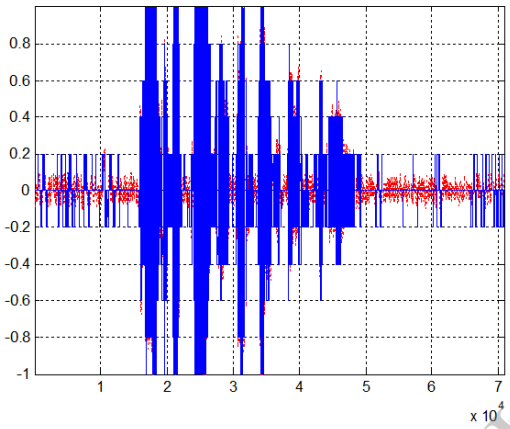


Figure5 Uniform quantization over the entire length of signal.

Firure5 represents the complete length of the signal quantized using uniform quantization

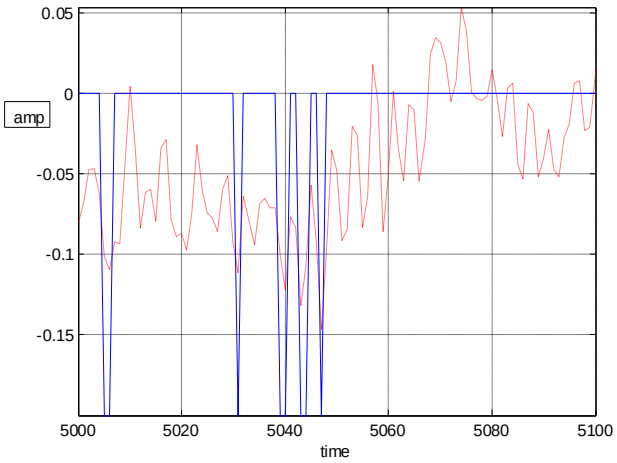


Figure6 shows the uniform quantization on fixed length of signal stretched over time interval t=5000 to t=5100.to digitize the signal the number of bits required are more.

B. Predictive Coding

In a typical speech waveform ,adjacent samples take similar value ,except at transition between difference phonemes..One way to exploit this correlation is by linear predictive coding. It first predicts a present sample x(n)using linear combination of previously reconstructed samples $\hat{x}(n-k)$ so that

$$x_p(n) = \sum a_k \hat{x}(n-k) \qquad (3.8)$$

The error between the actual samples value and the predicted ones,

$$d(n) = x(n) - x_p(n) \qquad (3.9)$$

quantized to $\hat{d}(n)$ and coded into codeword *c(n)*.

In the decoder the same predicted value is first produced from previously decoded samples . This value is then added to the decoded quantized value for the current samples i.e.

$$\hat{x}(n) = x_p(n) + \hat{e}(n) \qquad (3.10)$$

- Delta Modulation

A simple predictive coding system is the delta modulation (DM) system . In case the quantizer for the prediction error has only two levels and the step size is fixed . The positive quantization level is represented by *c(n) =0* and the negative value by *c(n)=1*.The $\hat{d}(n)$ is given by

$$\hat{d}(n) = \Delta \text{ or } c(n)=0 \quad if\ d(n)\geq 0$$
$$\hat{d}(n) = -\Delta \text{ or } c(n)=1 \quad if\ d(n)< 0 \qquad (3.11)$$

A simple first order prediction is used i.e. $x_p(n) = \hat{x}(n-1)$.

$\hat{x}(n)$ satisfies the difference equation

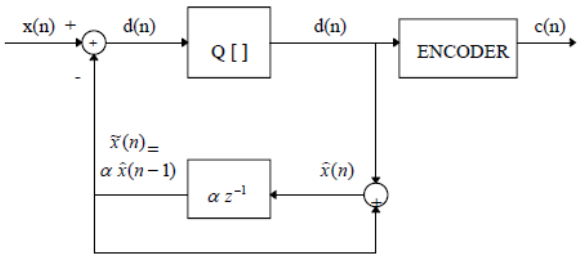
$$\hat{x}(n) = \propto \hat{x}(n-1) + \hat{d}(n) \qquad (3.12)$$

With α =1, the above equation is the digital equivalent of integration i.e. it represents the accumulation of positive and negative increments of magnitude Δ .The input to the quantizer is

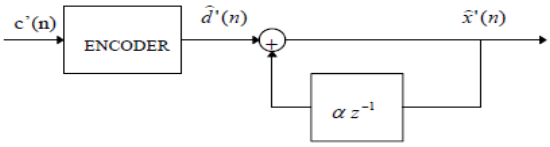
$$d(n) = x(n) - \hat{x}(n-1) = x(n) - x(n-1) - e(n-1) \qquad (3.13)$$

Thus except for the quantization error in $\hat{x}(n-1)$,d(n) is a first order backward difference of x(n) , which can be viewed as a digital approximation to the derivative of the input , the inverse of the digital integration process.

Because the error quantization is only two-level , the delta modulation has a bit rate of 1 bit/sample. If it is applied to a 16 bit/sample sequence , then it leads to a compression ratio (CR)



Encoder



Decoder

Figure7.showing block diagram of a delta modulation system

for the delta modulation to work well , the step size must be chosen properly to match the signal variation . This is a difficult task as the signal characteristics often change from tone to tone The quantization process of delta modulation with a fine step size is shown in figure 7. When the step size is too small the quantize signal lags below the actual signal magnitude . If the step size is too large in the latter portion , will cause the quantized signal to oscillate about the actual signal. . For a better performance , the step size should be adaptively adjusted .

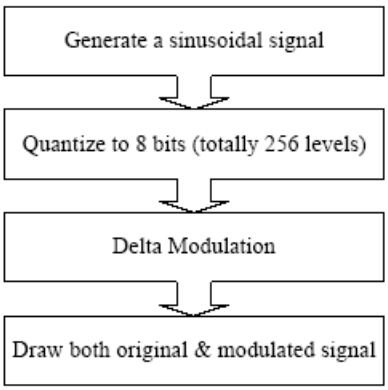
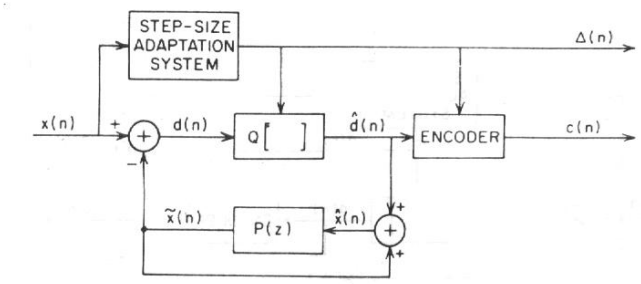


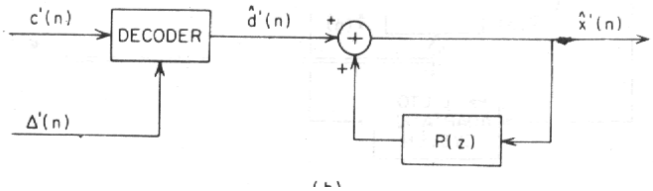
Figure8 showing the Flowchart for performing the delta modulation

- Adaptive Delta Modulation

The scheme shown in this paper have a feedback type scheme in which the step size for two level quantizer is adapted from the output code words. The general form of such system is shown in figure 8 . Such schemes maintain the advantage that no synchronization of bit pattern is recquired as there is no presence of errors, the step size information can be derived from the codeword sequence at both the transmitter and receiver .The use of adaptive quantization in delta modulation is discussed through specific adaptation algoritm.



Encoder



Decoder

Figure9.General form of adaptive delta modulation.

The scheme which is extensively studied by N.S.Jayant . Jayant's algorithm for adaptive modulation of quantization scheme is the step size is given as

$$\Delta(n) = M\Delta(n-1) \qquad (3.14)$$

$$\Delta_{min} \leq \Delta(n) \leq \Delta_{max} \qquad (3.15)$$

The algorithm for choosing step size is

M=P>1 if $c(n)=c(n-1)$

M=P<1 if $c(n) \neq$ c(n-1)

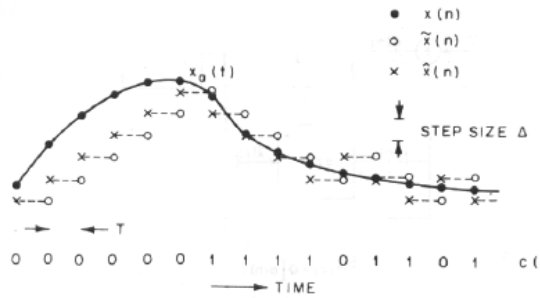


Figure10 Quantization done by Adaptive Delta Modulation

Figure 10 shows how waveform would be quantized by an adaptive delta modulation . for convenience the parameter of the system are set to P=2,Q=1/2,α=1, and the minimum step size is shown in figure. It can be that the region of large positive slope still runs of0's but in this case the step size increases exponentially so as to follow the increase in slop of the waveform. The region of granularity to the right of figure is again signalled by an alternating sequence of 0's and 1's but in this case step size fall rapidly to minimum (Δ_{min}) and remains there as long as the slope is small.

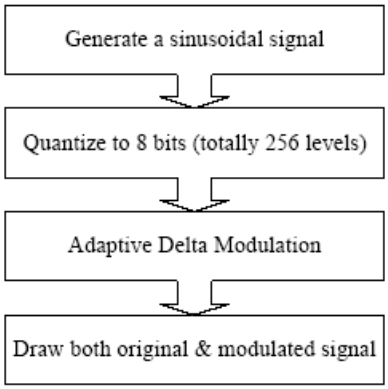


Figure 11 Flowchart for performing Adaptive delta Modulation

III. RESULTS

A. *Delta Modulation*

Figure 10 represents the delta modulation being performed on the complete sentence recorded at frequency of 8kHz for duration of 2 seconds.

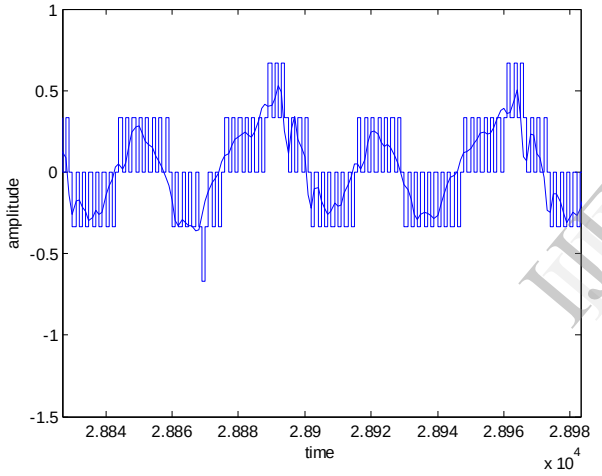


Figure 12 Delta modulation performed on the complete signal.

Figure 12 shows the delta modulation performed on the female voice speaker recorded for duration of 2 seconds over the complete length of signal .The experiment shows the of granular noise and hunting error is to very large extent ,which cause the information to be lost.

Therefore if the step size is made adaptive this problem gets ruled out .

B. *Adaptive Delta Modulation*

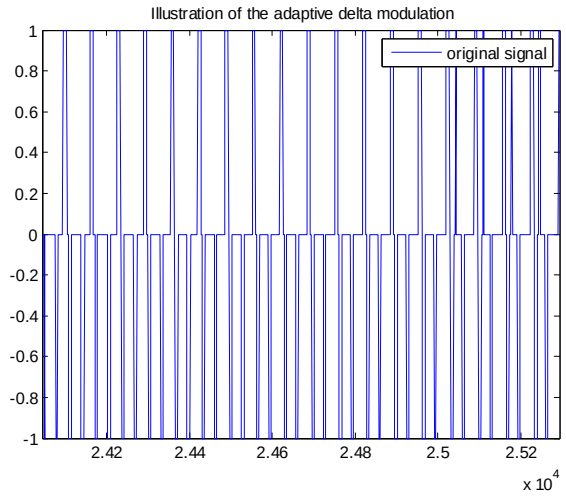


Figure 13. Adaptive delta modulation performed on the complete sentence .

Figure13 represents the adaptive delta modulation being performed on the complete sentence recorded at frequency of 8KHz for duration of 2 seconds.

IV. CONCLUSION

The above experiments conducted showed that the adaptive delta modulation gave better result than the delta modulation as the step size for quantization is adaptive to the input signal given to the system. The signal to noise ratio is also gets improved to larger extent in adaptive delta modulation . These can made useful for vocoder and encoder ,encoding the speech signal for transmitting the signal to larger distance .

REFERENCES

1] Makhoul.J."Linear Prediction :A Tutorial Review ", Proceedings of the IEEE,Vol.63.No4,pp561-580,Apr1975.
2] Rabiner and Schifer,"Digital Signal Processing" Printice and Hall Publication.
3] B.P.Lathi and Zhi Ding Modern Digital and Analog communication systems Oxford University Press , International Fourth edition 2010.
4] N.S.Jayant ,Adaptive quantization with one word memory..The Bell System Technical Journal,Vol.52.N0.7.pp1119-1144 .
5] Douglas O' Shaughnessy, 'Speech Communication' University Press, Second edition.
6] H.R.Schinder, 'Delta Modulation," IEEE spectrum,vol.7,pp69-78 October 1970.
7] L.R.Rabiner,R.W.Schafer,"Digital Processing of Speech Signal"Pearson edition
8] J.G.Prokis,'Digital Signal Processing" Pearson edition