COM-5 573

MILITARY TECHNICAL COLLEGE

CAIRO - EGYPT

PERFORMANCE STUDY OF DIFFERENT QUANTIZATION SCHEMES IN LPC VOCODER

A. Taha[*] , A. Elmoghazy[**] , A. Elosmany[***] and F. Elmansy[***]

ABSTRACT

We present in this paper a performance study of five different quantization schemes for coding (LPC) of speech. The quantization schemes used in this study are: linear quantization scheme, piece wise linear quantization scheme ,equal area quantization scheme ,log area ratio (LAR) quantization scheme, and inverse sine quantization scheme. The LPC distance measure proposed by Itakura is used for evaluating these quantization schemes.

* Air Forces

** Information System Department

*** Military Technical Col.ege

I. INTRODUCTION

In LPC vocoder, the transmission parameters are the reflection coefficients, the pitch period the residual gain, and the voiced/unvoiced decision. The bulk of the total number of bits per frame is used for coding the receiving end. Therefore, efficient coding of the reflection coefficients is important to minimize the transmission rate for given synthetic speech quality or to optimize the speech quality at a specified transmission rate.

Various quantization schemes of the reflection coefficients have been proposed since research on LPC for speech bandwidth compression began. These are the linear (direct) quantization scheme [1], the piece wise linear quantization scheme [2], the equal area quantization scheme [3], the LAR quantization scheme [4] and the inverse sine quantization scheme [1].

In this paper, these quantization schemes are used individually in LPC vocoder and the performance of the synthetic speech is studied for the same transmission rate. The LPC distance measure (between the original speech signal and the distorted synthetic speech signal due to quantization) proposed by Itakura [3] is used for this performance study.

Section II concerns the description of these five quantization schemes. Section III concerns the description of the experiment used for this study and the obtained results. section IV concerns the main conclusions of this study.

II. QUANTIZATION SCHEMES

A. Linear Quantization Scheme

In this method of quantization ,the range of each reflection coefficient k_i is divided into $Q^{(i)}$ intervals of equal length, say $\Delta^{(i)}$. If the value of k_i falls in the j-th quantizing interval, then the quantized value of k_i is taken to be the midpoint of this interval (Figure 1). The range values of k_i is (-1,+1), then the step size or interval length $\Delta^{(i)}$ is given by :

$$\Delta^{(i)} = 2 / Q^{(i)} \quad , Q^{(i)} = 2^{\beta_i} \tag{1}$$

where β_i is the number of bits assigned for k_i. The quantized output k_{iq} is generated according to :

$$k_{iq} = m_j^{(i)} \quad \text{if} \quad x_{j-1}^{(i)} \le k_i < x_j^{(i)} \tag{2}$$

where

$$x_j^{(i)} = -1+j\Delta^{(i)}, \; x_o^{(i)} = -1, \; x_Q^{(i)} = +1 \tag{3}$$

$$i = 1,2,\ldots,p, \; j = 1,2,\ldots,Q^{(i)}$$

where p is the number of reflection coefficients and

$$m_j^{(i)} = \left[x_{j-1}^{(i)} + x_j^{(i)} \right] / 2 \quad (4)$$

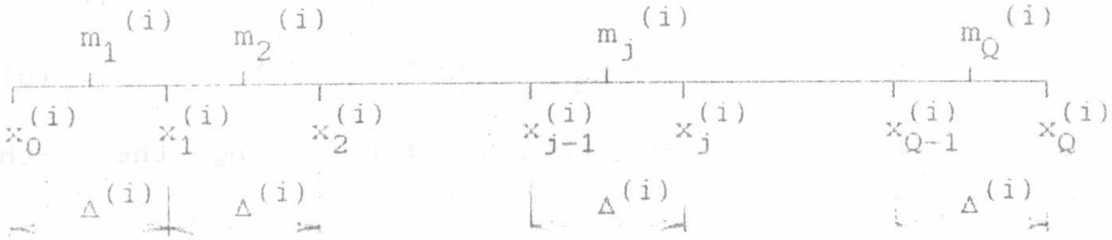

Figure 1. Linear Quantization Scheme.

B. Piece wise Linear Quantization Scheme

The quantization scheme for coding the reflection coefficients in this case uses piece wise linear quantization and requires statistical properties of the reflection coefficients. Rather than making quantizing levels uniform throughout the range of coefficient value, we divide the range of each coefficient into three regions and assign a different number of quantization levels to each region. This is shown in Figure 2. Note that, before the range of each coefficient is divided into three regions, the minimum value, $R_0^{(i)}$, and the maximum value $R_3^{(i)}$, of each coefficient k_i are set on the basis of experimental observation. The range of each coefficient is then divided into three regions by specifying $R_1^{(i)}$ and $R_2^{(i)}$ according to the distribution of the coefficient values and the desired quantization accuracy in each region. Also, the number of quantization levels in each region depends on the frequency of occurrence of the coefficient values and on the quantization accuracy.

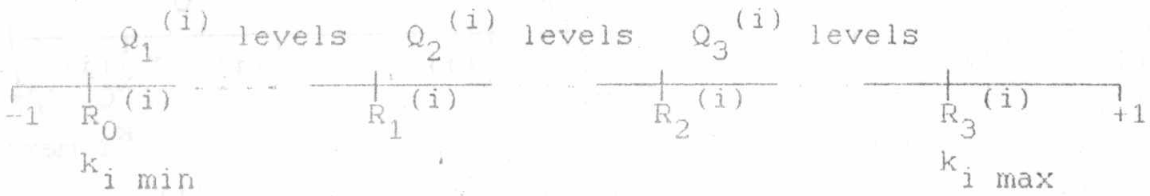

Figure 2. Piece wise Linear Quantization for Coding of a Reflection Coefficient.

Algebraically, the coding of a coefficient with piece wise linear quantization is done, depending on the region in which the coefficient value falls, one of the following equations is used :

$$\bar{k}_i = [[(R_3^{(i)} - k_i)\, Q_3^{(i)} / (R_3^{(i)} - R_2^{(i)}) + 1]], \quad R_2 < k_i \le R_3^{(i)} \quad (5)$$

$$\bar{k}_i = [[(R_2^{(i)} - k_i) Q_2^{(i)} / (R_2^{(i)} - R_1^{(i)}) + Q_3^{(i)} + 1]], \quad R_1^{(i)} < k_i \le R_2^{(i)} \quad (6)$$

$$\bar{k}_i = [[(R_1^{(i)}-k_i)Q_1^{(i)}/(R_1^{(i)}-R_0^{(i)})+Q_2^{(i)}+Q_3^{(i)}+1]], \quad R_0^{(i)} \le k_i \le R_1^{(i)} \qquad (7)$$

where $\bar{k}_i$ is the coded i-th coefficient, and a double bracket [[ ]] indicates integer truncation. It is to be noted that for binary encoding we always have :

$$Q_1^{(i)} + Q_2^{(i)} + Q_3^{(i)} = 2^{\beta_i} \qquad (8)$$

where β_i is the number of bits required for coding the i-th coefficient.

The decoding of $\bar{k}_i$ is done by inverse process. The decoded i-th coefficient $\hat{k}_i$, corresponding to equations (5), (6), and (7) is :

$$\hat{k}_i = R_3^{(i)} - (\bar{k}_i-0.5)(R_3^{(i)} - R_2^{(i)})/Q_3^{(i)}, \quad R_2^{(i)} < k_i \le R_3^{(i)} \qquad (9)$$

$$\hat{k}_i = R_2^{(i)} - (\bar{k}_i-Q_3^{(i)}-0.5)(R_2^{(i)}-R_1^{(i)})/Q_2^{(i)}, \quad R_1^{(i)} < k_i < R_2^{(i)} \qquad (10)$$

$$\hat{k}_i = R_1^{(i)} - (\bar{k}_i-Q_3^{(i)}-Q_2^{(i)}-.5)(R_1^{(i)}-R_0^{(i)})/Q_1^{(i)}, \quad R_0^{(i)} < k_i < R_1^{(i)} \qquad (11)$$

C. Equal-Area Quantization Scheme

The concept of equal area quantization scheme is based on the use of the distribution of coefficient values. The scheme uses an equal probability quantizing algorithm wherein the quantizer levels are made to occur with equal probability of occurrence. If the value of k_i falls within this interval, it will be quantized into the midpoint of this interval (Figure 3).

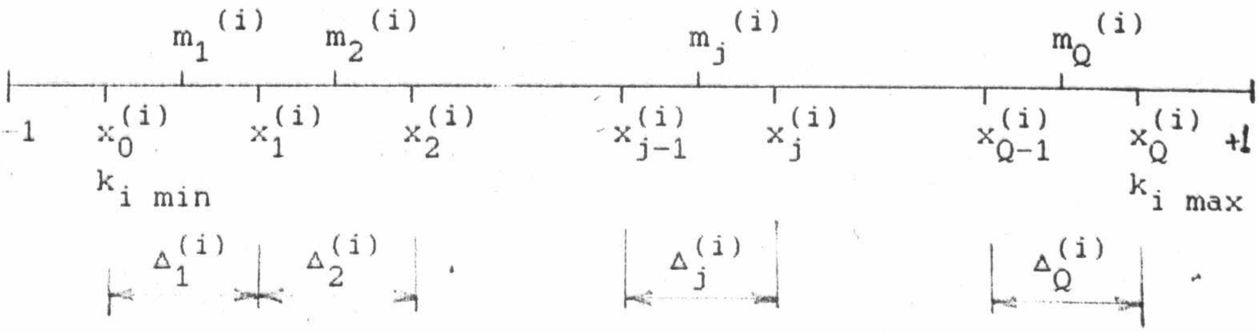

Figure 3. Equal Area Quantization Scheme.

The quantization intervals $\Delta_j^{(i)} = x_j^{(i)} - x_{j-1}^{(i)}$ are determined from :

$$P(x_{j-1}^{(i)} \le k_i < x_j^{(i)}) = \int_{x_{j-1}^{(i)}}^{x_j^{(i)}} p_X(x)\,dx = 1/Q^{(i)}, \qquad (12)$$

$i = 1,2,\ldots,p$, $j = 1,2,\ldots,Q$ and $x_0^{(i)} = k_{imin}$, $x_Q^{(i)} = k_{imax}$, $p_x(x)$ is the probability density function (pdf) of the coefficient k_i.

The quantized output k_{iq} is generated according to :

$$k_{iq} = m_j^{(i)} \quad \text{if} \quad x_{j-1} \leq k_i < x_j \tag{13}$$

where

$$m_j^{(i)} = (x_{j-1}^{(i)} + x_j^{(i)})/2,\ i = 1,2,\ldots,p,\ j = 1,2,\ldots,Q \tag{14}$$

The values of $x_j^{(i)}$, $j = 1,2,\ldots,Q$, $i = 1,2,\ldots,p$ are first determined from equation (12) and are used as an elaborate look-tables for quantization process.

D. Log Area Ratio (LAR) Quantization Scheme

The LAR coding scheme transforms k_i into a new function g_i:

$$g_i = \log\,[(1+k_i)/(1-k_i)],\ i = 1,2,\ldots,Q \tag{15}$$

these nonlinear transformed filter parameters $\{g_i\}$ are then linearly quantized in the range $[g_{imin}, g_{imax}]$. The quantization interval $\Delta^{(i)}$ of the transformed parameter g_i is given by :

$$\Delta^{(i)} = (g_{i\,max} - g_{i\,min}) / Q^{(i)},\ i = 1,2,\ldots,p \tag{16}$$

where $g_{i\,max} = \log\,[(1+k_{i\,max})/(1-k_{i\,max})]$ and

$$g_{i\,min} = \log\,[(1+k_{i\,min})/(1-k_{i\,min})]$$

For decoding process inverse transformation is used.

E. Inverse Sine Quantization Scheme

The inverse sine coding scheme transforms k_i into a new function θ_i, such that :

$$\theta_i = \sin^{-1} k_i,\quad i = 1,2,\ldots,p \tag{17}$$

These nonlinear transformed parameters $\{\theta_i\}$ are then linearly quantized in the range $[\theta_{imin}, \theta_{imax}]$. The quantization interval $\Delta^{(i)}$ of the transformed parameter θ_i is given by :

$$\Delta^{(i)} = (\theta_{imax} - \theta_{imin}) / Q^{(i)},\ i = 1,2,\ldots,p \tag{18}$$

where $\theta_{imax} = \sin^{-1} k_{imax}$ and $\theta_{imin} = \sin^{-1} k_{imin}$, $i = 1,2,\ldots,p$.

For decoding process, inverse transformation is used.

III. ITAKURA DISTANCE MEASURE

The measure developed by Itakura for auto correlation method of linear prediction is given by :

$$d = \ln (AVA^T / BVB^T)$$

where the vectors A and B are augmented LPC coefficient vectors. $A = (1, a_1, a_2, \ldots, a_p)$, $B = (1, b_1, b_2, \ldots, b_p)$. T is the transposition , and

$$V = [v|i-j|], \quad i, j = 0, 1, 2, \ldots, p$$

is the correlation matrix of the distorted speech segment Y_n with elements :

$$v(i) = (1/N) \sum_{n=1}^{N-|i|} Y_n Y_{n+i}$$

and N is the length of the speech segment.
The values a_i and b_i $(i=1,2,\ldots,p)$ are the reflection coefficient values before and after quantization, respectively.

IV. EXPERIMENT AND RESULTS

A. Experiment

This performance study is done using a computer simulation and changing the quantization process in the simulation in each case to be compatible with the used quantization scheme. The used file for testing these quantization schemes is a real speech signal sampled at 10 kHz frequency for the words "SPEECH CODING". The frame length is 200 samples (20 msec frame size). The used computer for performing this study is ICL 2960. The number of bits assigned for the reflection coefficients are fixed for all used quantization schemes and are as follows :
(5,5,4,4,4,4,3,3,3,3).

B. Results

The average LPC distance measure over all frames, over voiced frames , and over unvoiced frames is shown in Table 1.

Quantization Scheme	Average LPC Distance Measure		
	Over all Frames	Over voiced Frames	Over unvoiced Frames
1-Inverse Sine	0.0673	0.0270	0.0949
2-Piecewise Linear	0.0694	0.0284	0.0979
3-LAR	0.0731	0.0300	0.1025
4-Equal Area	0.0741	0.0440	0.0948
5-Linear	0.0942	0.0757	0.1069

C. Comments on the Results

1- The effect of these quantization schemes on the synthetic speech quality in the unvoiced frames is approximately the same.

2- The effect of these quantization schemes on the synthetic speech quality in the voiced frames are quite different and the most efficient scheme is the inverse sine scheme.

3- The most efficient scheme for all frames in the used file is the inverse sine quantization scheme while, the linear quantization scheme is the least efficient one , however it is the simplest one to implement.

V. CONCLUSIONS

Five different quantization schemes for reflection coefficients are studied and evaluated. The evaluation is done in three cases : over all frames, over voiced frames, and over unvoiced frames of the used file. The evaluation criterion is the synthetic speech quality for the same transmission rate. The results indicate that for this file , the most efficient scheme is the inverse sine scheme, while the linear quantization scheme is the least efficient one, however, it is the simplest one to implement.

REFERENCES

[1] J. D. Markel and A. H. Gray, Jr., *Linear Prediction of Speech*, Springer-Verlag, New York, 1976.

[2] C. K. Un and S. C. Yang, "Piecewise Linear Quantization of LPC Reflection Coefficients", *IEEE Int. Conf. Acoust., Speech, Signal Processing*, May 9-11, 1977.

[3] S. S. McCandless, "A New Encoding Technique for the k-Parameters ; A Statistical Approach", *Network Compression Group Note No. 53*, MIT Lincoln Lab., Lexington, Massachusetts, Dec. 1974.

[4] R. Viswanathan and J. Makhoul, "Quantization Properties of Transmission Parameters in Linear Predictive Systems", *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-23, No.3, June 1976.

[5] F. Itakura, "Minimum Prediction Residual Principle Applied To Speech Recognition", *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-23, No. 1, pp. 67-72, Feb. 1975.