First Libyan international Conference on Engineering Sciences & Applications (FLICESA_LA)
13 – 15 March 2023, Tripoli – Libya

Performance Improvement of Vector Quantization Methods Using Residual LSF

Hiba Faraj
department of electric and electronic
computer technology collag
Tripoli, Libya
hebafaraj1975@gmail.com

Selma Ozaydin
department of electric and electronic
Cankaya university
Ankara, Turkey
selmaozaydin@yahoo.com

Alia Andi
department of electric and electronic
high institute of industrial technology
Tripoli, Libya
Alia_eletri@yahoo.com

Basma Larbah
department of electric and electronic
computer technology collag
Tripoli, Libya
basmalarbah@gmail.com

***Abstract*: For low bit-rate linear predictive speech coding, it is best to quantize the LPC parameters using as few bits as possible with transparent quantization quality. To reduce the costs associated with applying vector quantization (VQ), many structured schemes have been proposed - among them the split VQ (SVQ) and the multistage VQ (MSVQ) schemes that offer substantial saving in both memory and computational cost. The present paper is an attempt to analyze these two methods, both of which employ an M-L search algorithm to significantly reduce the computational complexity and produce identical even improved quality with sequential search in terms of spectral distortion and outliers. A comparison is made of four different VQ methods, namely (SVQ, MSVQ, RSVQ, and R_MSVQ) all evaluated in terms of SD complexity and memory requirement. Simulation results show that a 3-stage split residual vector quantization can achieve a transparent quantization quality at 24 bit/frame. And MSVQ gives better performance in SD than SVQ in terms of SD, and R_MSVQ gave the best SD of all methods specially when compared to RSVQ.**

Keywords: Multistage vector quantization, Split vector quantization, Residual vector quantization, Codebook design

I. INTRODUCTION

VQ is a data reduction method, which means that in speech coding, quantization is required to reduce the number of bits used to represent a sample of speech signal. When less number of bits is used to represent a sample the bit-rate, complexity and memory requirement gets reduced. can be as simple as possible. The challenge in VQ is to map an N-dimensional real-valued vector x = $[x_1,x_2.....x_N]$ to another real-valued vector y_i= $[y_{i1},y_{i2}\ ,.....y_{iN}]$. Typically, all possible values of y_i are composed into a reconstruction codebook Y = $\{y_1,y_2.....y_L\}$. Hence, if L is the size of the codebook, it is called an L-level codebook or L-level quantizer. Vector quantization (VQ) is an important application of distance measures as it can reduce the dimension of an input data upon encoding each input vector as a single number. In a VQ algorithm, a set of training data is taken, and centroids are found by using a clustering algorithm. Among these clustering methods is the k-means clustering algorithm, which begins by choosing k vectors as the starting central points; then, it assigns each training input vector to the closest of these centers to gradually work out new central points by taking the means of each cluster. These three steps are repeated until the central point's cease to change any further or significantly. In speech coding, quantization is required to reduce the number of bits used to represent a sample of speech signal. When fewer bits are used to represent a sample, the bit-rate, complexity and memory requirement are reduced as well. Thus, a compromise must be made between the reduction in bit-rate and the quality of the speech signal. The parameters used in the analysis and synthesis of these signals are the LPC coefficients. In speech coding, quantization is not performed directly on the LPC coefficients, but carried out by transforming the LPC coefficients to other forms to ensure filter stability after quantization. The alternative to LPC coefficients is the use of line spectral frequency (LSF) parameters for filter stability after quantization. [3-7].

Speech coding refers to the process of reducing the bit rate of digital speech representations for transmission or storage while maintaining a speech quality that is acceptable for the application. Most of the speech coders reported in the literature are based on linear prediction (LP) analysis [1]. For LP-based vocoders, the bit rate reduction is strongly tied to efficient quantization of the LPC filter coefficients $\{a_j\}$. The Line Spectral Frequencies (LSF), an equivalent representation of $\{a_j\}$ and more suitable for quantization and interpolation, can alternatively be used for the same purpose. In this sense, the Multi-Stage Vector Quantization (MSVQ) of LSF parameters presented in [4] has an efficient quantization performance at 22-24 bits per 20ms frames. Furthermore, such a multi-stage structure has more flexibility than a single stage VQ in terms of search complexity, codebook storage and channel error protection. Very low-rate speech communication systems require efficient fixed-rate and low delay coding methods which operate at lower bit rates. The residual multistage vector quantization (R_MSVQ) is a combination of two product code vector quantization techniques, namely the Residual vector quantization technique and the Multistage vector quantization technique namely the Residual vector quantization technique and the Multistage vector quantization technique. In this paper, the R_MSVQ is implemented using an LSF coefficient vector split into stages, each of the employs an M-search for quantization according to a multistage structure. The residual LSF parameters of a current frame are predicted from the quantized LSF parameters of the previous frames and, then, the residual LSF vectors are coded with an MSVQ codebook. R_MSVQ offers satisfactory performance and reaches the transparency required for speech hence its employment in the present study to improve the spectral distortion

II. CODEBOOK DESIGN SPEECH

signal VQs require codebook generation which, in this study, are designed using an iterative algorithm called the Linde, Buzo, and Gray (LBG) algorithm. The input to the LBG algorithm is a training vectors clustered into a set of codebook vectors. The speech signals used to obtain training vectors must be free of background noise, and they are ideally recorded in soundproof booths, computer rooms, and open environments. In this work, the speech signals are taken from the TIMIT database. The codebook generation using the LBG algorithm requires the formation of an initial codebook, which is the centroid or means obtained from the training sequence. This centroid is, then, split into two centroids or codewords using the splitting method. In turn, the iterative LBG algorithm splits these two codewords into four, four into eight, and the process continues till the required number of codewords in the codebook are obtained [5-8].

III. RESIDUAL MULTISTAGE VECTOR QUANTIZATION (R_MSVQ)

Multistage Vector Quantization (MSVQ) is an evolution of the basic VQ technique, otherwise regarded as multistep, residual or cascaded vector quantization. It is a cascaded VQ encoder where the output of the VQ of a stage is the input of the next. It preserves the features of the VQ technique while reducing computational complexity and memory requirements while improving the quality. In MSVQ, each stage has its own codebook. The codebook of the first stage is created using the training sequence as the input, and the codebook of each remaining stage is created using the quantization error of its previous stage. As for VQ, we used the LBG algorithm to create the codebooks. Let y be the signal to be quantized and assume the number of stages as 3 (st = 3) which illustrates the MSVQ encoding and decoding level. Note that, here, the input of the V Q of the first stage is y, the input of the V Q of the second stage is the quantization error of the previous stage ($e_1 = y - \hat{y}_1$), and the input of the V Q of the third stage is the quantization eerror of stage two ($e_2 = e_1 - \hat{e}_1$). The MSVQ encoder provides three indexes as the observation vector (y), the quantization error of the first stage (e1), and the quantization error of the second stage (e2) the MSVQ decoder uses the codebook of each stage to find the corresponding codewords of the indexes and, then, provide the quantize vectors of the stages as $\hat{y}_1$, $\hat{e}_1$, and $\hat{e}_2$. Accordingly, the quantized observation vector is $\hat{y} = \hat{y}_1 + \hat{e}_1 + \hat{e}_2$. In this study, we designed the residual MSVQ (R_MSVQ) algorithm, with the related flowchart appearing in Fig1, As for the R_MSVQ, the following details describe the approach:

- R_MSVQ is a hybrid of two product code vector quantization techniques; namely, the residual vector quantization technique (RVQ) and the multistage vector quantization (MVQ) technique.

- The R_MSVQ implemented employs LSF coefficients. Here, the LSF coefficient vector is split into stages, where M-L search is used according to a multistage structure.

- The residual LSF parameters of the current frame are predicted from the quantized LSF parameters of the previous ones and, later, residual LSF vectors are coded with an MSVQ codebook.

- In the R_MSVQ method, the residual LSF parameters of the current frame are predicted from the quantized LSF parameters of the previous frames using interframe correlation feature of spectrum parameters [5-8] and then residual LSF vectors are coded with a MSVQ codebook. Firstly, the LSF parameter vector is obtained by transforming the 10th order LPC parameter vector. Next, the average LSF vector of the training set x_{DC} is subtracted from the LSF vector $x^{(i)}$ belonging to the i^{th} frame. By defining the mean removed LSF vectors ($z^{(i)}= x^{(i)}- x_{DC}$) and its quantized version $z^{\wedge(i)}= x^{\wedge(i)}- x_{DC}$, the residual LSF vector e(i) is calculated using

$$e^{(i)}= z^{(i)}- r^{(i)} \qquad (1)$$

Where i=1, 2,…and r(0) = 0. The quantized residual vector $e^{\wedge(i)}$ is found by quantizing $e^{(i)}$ with a VQ codebook. Depending on how (i) is computed, various prediction schemes can be proposed.

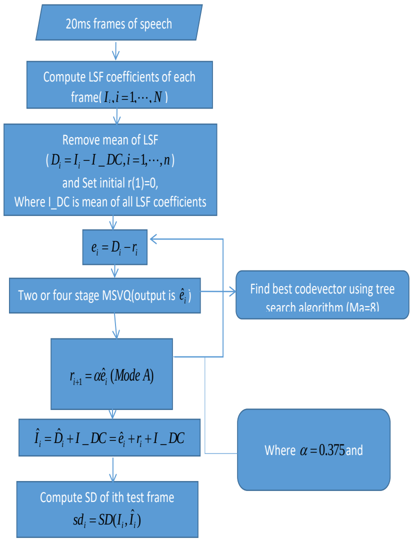

Figure1. flowchart for RMSVQ

The LSF parameter vector is obtained by transforming a 10th order LPC parameter vector. Next, the long-term average LSF vector (obtained by averaging the LSF vectors in the training set) I_{DC} is subtracted from the LSF vector belonging to the ith frame I_i to obtain a differential LSF vector d_i given by (1):

$$d_i = I_i - I_{DC} \qquad (2)$$

IV. RESIDUAL SPLIT VECTOR QUANTIZATION (RSVQ)

This technique is developed to improve the performance of split vector quantization. As stated earlier, R_SVQ is also a hybrid of two product code vector quantization techniques, namely the residual vector quantization technique and the split vector quantization technique. In RSVQ, the vector dimensions to be quantized are reduced by means of splitting, and the bits allocated to the quantizer are divided among the stages and splits of each stage. All sub-vectors in the quantizer are connected in cascade to form a multi-stage structure. Obviously, splitting the input vector into sub-vectors will lead to performance degradation because of the decline in the statistical dependence between the sub-vectors. However, if an analysis procedure is

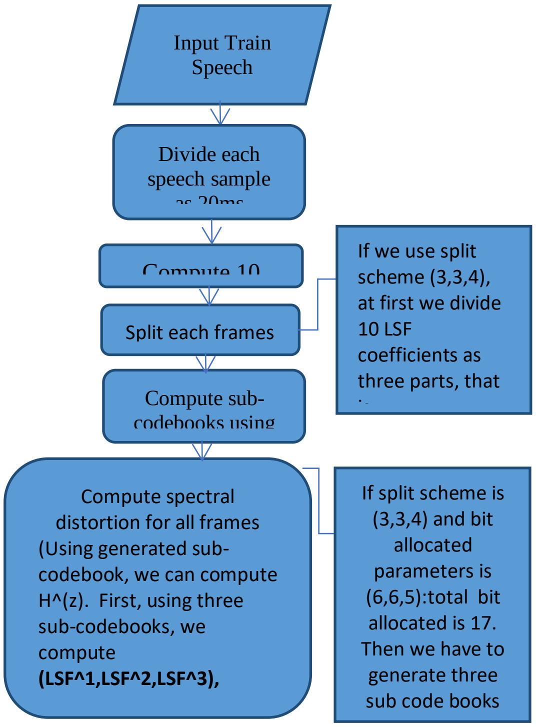

Figure 2. SD for 3-stage RMSVQ and RSVQ

devised in such a way to allow for coupling the forward and backward prediction errors across successive stages, then the statistical dependence can be preserved during the quantization performance of the SVQ to approach to that of a single stage VQ. The details of the RSVQ are explained in the following low chart fig 2.

❑ In Split Vector Quantization, the training sequence used for codebook generation is split into vectors of smaller dimensions.
❑ Each split of the training sequence is used to generate separate sub codebooks.
❑ Computing LSF coefficient and removing the mean from the LSF coefficient.
❑ Each VQ stage operates on the residual vector of the previous stage similar to the residual VQ scheme.
❑ The use of a split vector quantizer makes the less availability of bits at each split of the vector quantizer as a result, the complexity and memory requirements are greatly reduced, but the dependencies that exist across the dimensions (splits) of a vector is lost. Consequently, the spectral distortion is slightly increased.

V. PERFORMANCE EVALUATION

The quality of the speech signal is an important parameter in speech coders, measured in terms of spectral distortion for an objective performance evaluation. The spectral distortion is measured between the LPC power spectrum of the quantized and unquantized speech signals. In this work, the experiment is carried out using the standard TIMIT speech database containing American and British English speeches sampled at 8 kHz. In order to design SVQ or MSVQ codebooks, the LSF vectors are generated from the TIMIT database. The sentences are uttered by both female and male speakers. The frame length is about 30 ms with hamming window and overlap 10ms. A 10^{th}-order LPC analysis based on an autocorrelation method is carried out for every 20 ms frame. The resulting coefficients of the 10^{th}-order LP polynomial A(z) are, then, converted into the LSF domain (Figure 4). The LSF input data is used to design the required codebooks. The input data used in the experiments consists of two separate databases, namely for training and testing. The former consisted of 90,000 LSF vectors for clean speeches, and the latter consisted of 43,300 LSF vectors for clean speech and 32,400 LSF vectors for noisy speech. After designing the codebooks, the spectral distortion (SD) values are calculated over the frequency band of 100-3800Hz for 8 kHz sampled speech. [8]

Table1. SD for RSVQ for 3-splits

Number of bits	**SD for (3,3,4)**	**SD for (4,3,3)**	**SD for (4,4,2)**
14	2.325	2.277	2.286
16	2.078	2.077	2.147
18	1.881	1.910	1.986
20	1.712	1.669	1.727
22	1.468	1.485	1.595
24	1.297	1.380	1.499

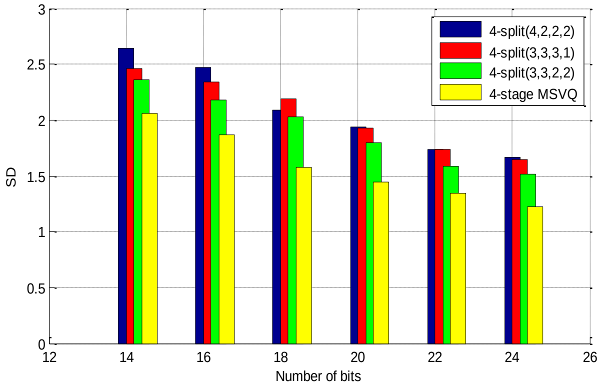

Figure 3. SD for 3-stage RMSVQ and RSVQ

Table2. SD for RSVQ for 4-splits

Number of bits	**SD for (4,2,2,2)**	**SD for (3,3,3,1)**	**SD for (3,3,2,2)**
14	2.643	2.457	2.363
16	2.475	2.339	2.182
18	2.092	2.190	2.026
20	1.937	1.930	1.801
22	1.734	1.735	1.587
24	1.663	1.648	1.515

Table3. SD for RMSVQ for 3 & 4splits

Number of bits	SD for 3-stages	SD for 4-stages
14	1.94	2.06
16	1.80	1.87
18	1.61	1.58
20	1.43	1.44
22	1.31	1.34
24	0.98	1.22

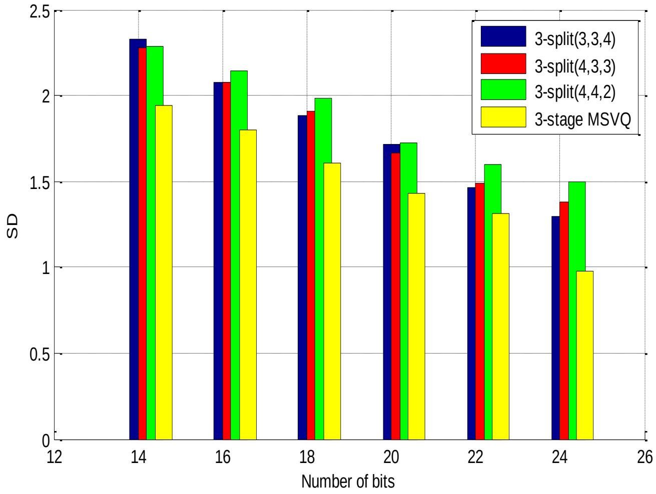

Figure 4. SD for 4-stage RMSVQ and RSVQ

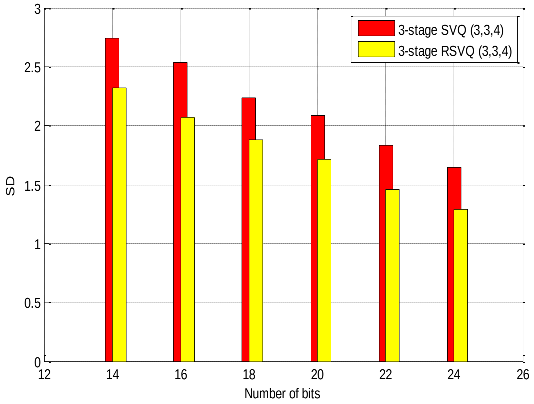

Figure 5. SD for 3- stages RSVQ and SVQ

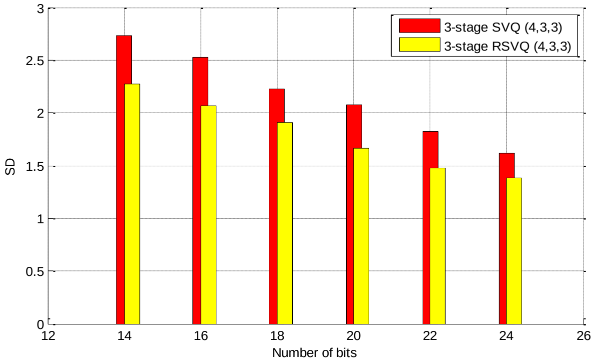

Figure 4. SD for 3- stages RSVQ and SVQ

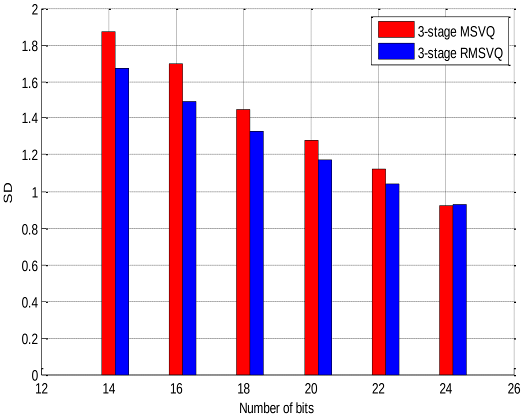

Figure 6. SD for 3- stage R_MSVQ and MSVQ

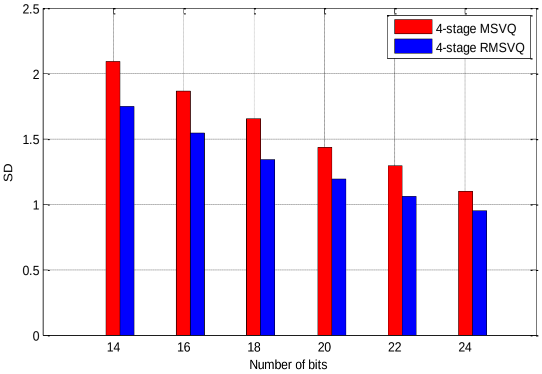

Figure 7. SD for 4- stage R_MSVQ and MSVQ

6. Analysis of the Tables and Figures:

A residual split vector quantization (RSVQ) and residual multistage vector quantization (R_MSVQ) scheme for sequential quantization of LPC coefficients was presented. When we compare both (RSVQ) and (R_MSVQ) as we can see from tables R_MSVQ give best performance than RSVQ in terms of SD. As we can see from table 1,2 and 3 calculating the SD in all number of bits from 14 to 24 R_MSVQ give very good performance in terms of SD when compared with RSVQ about 27% improvement. And as we can see from literature studies that MSVQ is better than SVQ in terms of SD and comparing this literature results for (SVQ, MSVQ) with new methods (RSVQ, R_MSVQ) from the table results R_MSVQ gives the best result when compared with all algorithms as we can see from figures [1-6].

VI. Conclusion

The RSVQ and R_MSVQ schemes for sequential quantization of LPC coefficients is presented. From the results of the designed codebooks with these methods, it can be said that RSVQ outperforms SVQ by about 27%. In 3-stage the best result of SD is the split (4,3,3) for both the SVQ and RSVQ algorithms. Comparing these two results (SVQ and RSVQ) reveals that RSVQ provides lower SD, and that R_MSVQ also yields lower SD when compared to MSVQ. Based on the figures, R_MSVQ performs better than RSVQ, which in turn performs better in terms of computation complexity and memory requirement. RSVQ with SVQ according to the bit allocation from14 to 24 gives better performance than SVQ.

REFERENCE

[1] So.Stephen and K. K.Paliwal ,"Efficient product code vectoquantization usng switched split vectorquantizer ," Digital Signal Processing journal, Elsevier, Vol. 17, Issue 1, Jan 2007, pp.138-171.

[2] F .K.Soong and B.H.J uang, "Line spectrum pair (LSP) and speech dat a compression," in: Proc. IEEE Int.Conf.Acoust ,Speech, Signal Processing, Vol 9, Issue1, March 1984, pp. 3740.

[3] P.Kabal and R.P.Ram a chandran,"The computation of line spectral frequencies using Chebyshev polynomials," IEEE Trans.Acoust . Speech Signal Process, Vol. 34, Issue 6, Dec 1986, pp.1419-1426.

[4] Hong Kook Kim and Hwang Soo Lee, "Interlacing properties of line spectrum pair frequencies," IEEE Trans. Speech Audi o Process, Vol. 7, Issue 1, Jan 1999, pp.87– 91.

[5] W.B. Kleijn,T.Backstrom,P.Alku,"Online spectral frequencies," IEEE Signal Processing Letters, Vol 10, Issue 3, March 2003, pp.7577.

[6] Y. Linde, A. Buzo and R.M.Gray, "An algorithm for vector quantizer design," IEEE Trans. Commun,Vol 28, Issue1, Jan.1980, pp.84 – 95.

[7] A. Buzo, A.H.Gray,R.M. Gray and J . Markel , "Speech coding based upon vector quantization," IEEE Trans. Acoust. Speech Signal Process,Vol. 28, Issue 5, Oct 1980, pp.562-574.

[8] N.Sugamura and N.Farvardin,"Quantizer designin LSP speech analysis and synthesis," in:Proc. IEEE Int.Conf.Acoust,Speech,Signal Processing, Vol. 1, Apr 1988, pp. 398– 401.

[9] A. Gersh, R.M Gray, Vector Quantization and Signal Comperssion, Kluwer Acadmic Publisher, dordrecht, 1992

[10] N. Kitawaki, H. Nagabuchi and K. Itoh, "Objective quality evaluation for low-bit rate speech coding systems," 21 IEEE Journal on Selected Areas in Communications, Vol 6, Issue 2, Feb 1988, pp. 242–248.

[11] P. R. Kanawade ; S. S. Gundal"Tree structured vector quantization based technique for speech compression,"2017,23 International Conference on Data Management, Analytics and Innovation (ICDMAI)

[12] M. Shafi V. Makwana ; A. B. Nandurbarkar ; K. R. Parmar"Speech compression using tree structured vector 25 quantization,"2014 2nd International Conference on Device,Circuit and Systems(ICDCS)

[13] J. Tribolet, P. Noll, B. McDermott and R. Crochiere, "A study of complexity and quality of speech waveform 27 coders," in: Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, Vol. 3, Apr 1978, pp. 586–590.

[14] J. Tribolet, P. Noll, B. McDermott and R. Crochiere, "A study of complexity and quality of speech waveform 29 coders," in: Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, Vol. 3, Apr 1978, pp. 586–590.