ISCA Archive
http://www.isca-speech.org/archive

4th European Conference on
Speech Communication and Technology
EUROSPEECH '95
Madrid, Spain, September 18-21, 1995

CONDITIONAL SPLIT VECTOR QUANTIZATION OF LSP PARAMETERSWITH MULTIPLE SEARCH

Dong-il CHANG, KyungMin NA and Souguil ANN

Department of Electronics Engineering, Seoul National University
Seoul 151-742, Korea
Tel/Fax: +82-2-880-7279/+82-2-882-3906
e-mail: comlab@alliant.snu.ac.kr

10.21437/Eurospeech.1995-271

ABSTRACT

In this paper, we propose an efficient vector quantization scheme of line spectral pair parameters which combines conditional split VQ with a multiple search. At first, we show the efficiency of mutiple search for two VQ schemes of LSP, split VQ and conditional split VQ. In spite of improved performance, the unavoidable increase in computational complexity makes the SVQ with multiple search unattractive. However, the reduced complexity of conditional split VQ make the combination with a multiple search realizable and very attractive. Experimental results are presented to show the efficiency of the proposed scheme, the multiple search conditional split VQ (MS-CONSVQ).

1. INTRODUCTION

To achieve transparent quantization of LSP parameters at a rate below 30 bits/frame, many VQ schemes have been proposed. Among them, multistage VQ (MSVQ) and split VQ (SVQ) have been the most successful schemes so far. Recently, LeBlanc *et al.* reported that they could enhance the performance of tree searched MSVQ (TS-MSVQ) scheme by adopting an *M-L* search and an iterative codebook design algorithm [1]. These structures are adopted for reducing the inefficiencies of the conventional TS-MSVQ in codebook search and design procedure, respectively. Especially, by introducing the *M-L* search, they could enhance the performance of TS-MSVQ considerably.

In the conventional SVQ, the inefficiencies of codebook search and design procedure are also exist. To reduce the inefficiencies of the conventional fixed splitting scheme, we have proposed a conditional splitting scheme and a conditional SVQ (CONSVQ) [2]. The CONSVQ requires considerably reduced computational complexity while showing nearly equal performance to that of SVQ.

In this paper, we show the efficiency of a multiple search for SVQ and CONSVQ. As we can see in the later section, the performances of SVQ and CONSVQ are considerably improved by introducing a multiple search. However, introducing a multiple search also causes the undesirable increase in computational complexity. Accordingly, the SVQ scheme with multiple search is hardly realizable because of its huge computational complexity.

On the contrary, the increase in computational complexity from multiple search can be nearly compensated in CONSVQ. As a result, we construct a very efficient VQ scheme, the multiple search CONSVQ (MS-CONSVQ), which shows considerably improved performance compared with SVQ while maintaining computational complexity.

This paper is organized as follows. The brief description of SVQ and CONSVQ is presented in Section 2. In Section 3, a multiple search algorithm for SVQ and CONSVQ is proposed, and the overall structure of MS-CONSVQ is summarized in Section 4. Experimental results are presented in Section 5 and conclusions are followed in Section 6.

2. SVQ AND CONSVQ

In the conventional SVQ, an LSP vector is split into subvectors in the order of frequency location [3]. This fixed splitting scheme is based on the experimental results that low frequency LSPs are more important than the high frequency LSPs [4]. Though providing considerably low spectral distortion, it is not the best way to get the minimum spectral distortion. To achieve better performance, we should select the most sensitive LSPs and quantize them more precisely. In general, LSPs constructing spectral formants are more sensitive than LSPs constructing spectral valleys. In addition, the LSPs constructing formants are generally more important in perception than LSPs constructing valleys.

However, the relative importance in terms of sensitivity and perception has a time-varying property. Therefore, a time-varying selection scheme is required instead of a fixed scheme. Besides, a proper index representing the sensitivity and perceptual importance of each LSP is also required. In our earlier work, we propose the LSP perceptual importance index (LPII), and realize the time-varying splitting scheme by using a classified VQ structure [2].

According to the earlier experimental results, the spectral sensitivity of an LSP is an exponential-like decaying function of the differences between it and its adjacent LSPs [5]. Besides, two or three closely located LSPs construct a spectral peak, and an isolated LSP constructs a spectral valley generally [2][3]. From these characteristics, therefore, we can find that we should quantize the closely located LSPs more accurately to achieve lower spectral distortion than the isolated ones.

In this paper, LPII $L(\omega_i)$ is defined by

$$L(\omega_i) = W(\omega_i) \cdot \max\{d(\Delta\omega_i), d(\Delta\omega_{i+1})\} \qquad (1)$$

where $\Delta\omega_i = \omega_i - \omega_{i-1}$, and $W(\omega_i)$ is a weighting function representing the importance of lower indexed LSPs. The function of distance $d(\Delta\omega_i)$ is also defined by

$$d(\Delta\omega_i) = \begin{cases} C, & \text{if } \Delta\omega_i \le T_1 \\ 0, & \text{if } T_1 < \Delta\omega_i \le T_2 \\ -C, & \text{if } T_2 < \Delta\omega_i \end{cases} \qquad (2)$$

where $C > 0$ and $T_1 < T_2$. Earlier experimental results have shown that we can represent the relative importance of each LSP efficiently by using this function [2]. An example of LPII with $W(\omega_i) = 1$ is illustrated in Figure 1. In Figure 1, three constants C, T_1 and T_2 are fixed as 10, $\pi/22$ and $\pi/11$, respectively.

By using the LPII in (1), we construct the subvectors according to the order of importance, and quantize them independently with different levels of precision. In this conditional splitting scheme, we should also transmit information about splitting status in addition to the VQ index. Therefore, we should restrict the number of splitting conditions within a proper range. The restriction of the number of conditions can be efficiently realized by using a classified VQ structure. By dividing the input space into several finite subspaces, and dealing with each subspace independently, we can easily restrict the number of conditions without introducing serious inefficiency to the conditional splitting.

3. PROPOSED MULTIPLE SEARCH

A. Multiple search algorithm for SVQ and CONSVQ.
The average spectral distortion that is widely used as a performance measure is evaluated by [1]

$$SD = \frac{1}{N_1 - N_0}\sqrt{\sum_{k=N_0}^{N_1-1}\left(10\log_{10}\left|\frac{A_q(k)}{A(k)}\right|^2\right)^2} \qquad (3)$$

where $A_q(k)$ is FFT of LPC filter constructed by the quantized LSP parameters and $A(k)$ is FFT of original

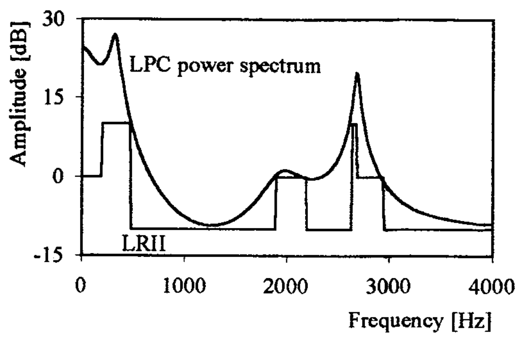

Figure 1. A comparison of LPII with its corresponding LPC power spectrum.

LPC filter. Two indexes N_0 and N_1 correspond to two cut-off frequencies of input band-pass filter, respectively. Because of the huge required computations, we cannot use (3) as a distance measure in codebook search procedure. In spite of the performance degradation, therefore, we generally use simple distance measures such as mean square error (MSE) or weighted MSE (WMSE).

The degradation of performance resulted from the inconsistency between performance measure and distance measure can be reduced by using a multiple search algorithm such as *M-L* search in TS-MSVQ. Instead of selecting a path at each stage, *M-L* search traces M paths which provide the minimum overall distances. Among selected M paths, we choose a path as a codevector with the average spectral distortion as a final selection criterion. By using the *M-L* search in TS-MSVQ, the possibility of selecting optimal codevector with respect to the performance measure in (3) is considerably increased.

In SVQ, the codevector is given as a concatenation of subvectors selected at each stage. The selected subvector at a stage has no effect on the selection at the other stages in SVQ, whereas it affects the selection at the remaining stages in MSVQ. In spite of the independency in selection, a combination of selected subvectors with minimum distances may not provide the best performance in terms of (3) in SVQ. This means that a multiple search can improve the overall performance of SVQ.

In SVQ, we should adopt a multiple search which selects multiple candidates at each stage, and consider all combination with overall performance measure at the final stage. If we choose M_i candidates at i-th stage among all L stages, the number of combinations M_{total} is given by

$$M_{\text{total}} = \prod_{i=0}^{L-1} M_i . \qquad (4)$$

For every combination, we should compute the average spectral distortion in (3), and select a combination achieving the lowest distortion as a codevector. Since the basic structure of CONSVQ is identical with that of SVQ. this multiple search can be also used in CONSVQ.

B. Computational complexity of multiple search.

The computational complexity is defined as the number of arithmetic operations such as multiplies, adds and multiply-adds, etc. In case of using MSE distance measure, the computational complexity of SVQ with L subvectors is given as follows:

$$C_{\mathrm{SVQ}} = \sum_{i=0}^{L-1} n_i N_i \tag{5}$$

where n_i and N_i is the dimension and the size of the i-th codebook, respectively. Similarly, the estimated computational complexity of multiple search C_{mult} is given as follows:

$$C_{\mathrm{mult}} = \left(M_{\mathrm{total}} + 1\right) C_{\mathrm{FFT}} + M_{\mathrm{total}} \cdot C_{\mathrm{SD}}. \tag{6}$$

The C_{FFT} and C_{SD} are the computational complexity required in computing fast Fourier transform (FFT) and the terms inside square root in (3), respectively. Since square root and constant multiply do not change the order of magnitudes of distortion, we just compute and compare the inner term of square root to compare the distortion of each combination. The complexity C_{FFT} which is represented the numbers of multiplies and adds is varied according to the method for computation of FFT. If we use an efficient method for computing FFT with low complexity [6], the complexity required in computing FFT is given by

$$C_{\mathrm{FFT}} = C_{\mathrm{multiply}} + C_{\mathrm{add}} = 2N\log_2 N - 4N + 6. \tag{7}$$

The complexity C_{SD} may be varied according to the implementation methods, since computing the inner term of square root in (3) requires the computation of logarithm and floating-point division repetitively. If we use small modification replacing the logarithm with the polynomial approximation, we can reduce the computation complexity C_{SD}. In addition, we also achieve additional reduction of computational complexity by keeping N as low as possible. According to our experimental results, N should be equal to or greater than 128, and third order approximation of logarithm has provided nearly the same performance as the exact logarithm.

Additionally, if we exploit the dynamic-M search in this multiple search, the complexity C_{mult} can be more reduced [7]. By excluding the candidates with higher MSE distance than a threshold, we can reduce the number of combinations to be examined in the final stage without degradation of performance. A time-varying threshold determined by the minimum distance of each stage shows good efficiency in exclusion of subvectors. However, it is not the scope of this paper.

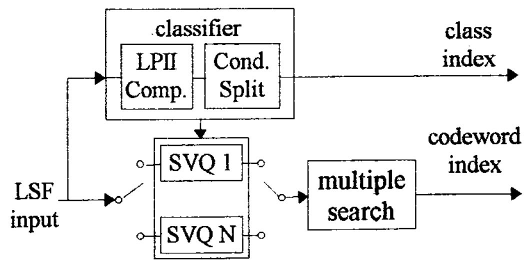

Figure 2. The structure of MS-CONSVQ.

4. THE STRUCTURE OF MS-CONSVQ

In this section, we present the structure of MS-CONSVQ in detail. The classifier of CONSVQ can be a phonetic classifier or a simple VQ of feature vectors, etc. An earlier study of ours shows that a simple VQ of LPII can classify vectors of LSP parameters efficiently [2]. In this paper, we also use a simple VQ of LPII with non-uniform spectral weighting $W(\omega_i)$. This weighting is identically defined with the weighting used in [1]. The number of classes should be determined considering the trade-off between performance and computational complexity. Since the case of 3 classes has provided good performance and considerably reduced computational complexity, we set the number of classes as 3. For each class, an independent SVQ is designed and used for encoding input LSP vectors. The bit allocation of CONSVQ with two subvectors is summarized in Table 1.

Table 1. Bit allocation of CONSVQ operating at N bits/frame with 2 subvectors and 3 classes.

class	class index	subvector 1	subvector 2
1	1	$(N/2)$	$(N/2)-1$
2	2	$(N/2)-1$	$(N/2)-1$
3	2	$(N/2)-1$	$(N/2)-1$

In case of 3 classes, the CONSVQ has reduced computational complexity by about 40 % compared with the conventional SVQ while maintaining the performance [7]. The overall structure of MS-CONSVQ is illustrated in Figure 2.

5. EXPERIMENTAL RESULTS

In our experiments, we use the structure of the conventional SVQ with two subvectors of dimension 4 and 6, respectively. The database is constructed from 8 different FM radio stations with 29 minute's length. The first 25 minutes of database is used for training VQs and the remaining 4 minutes of database is used for testing. In Table 2, we present the quantization results obtained by SVQ and CONSVQ at 24 and 26 bits/frame, respectively.

Table 2. Average spectral distortion and outliers for SVQ and CONSVQ at 24 and 26 bits/frame.

rate	SD	outlier (%)	
(bits/frame)	(dB)	2 - 4 dB	> 4 dB
SVQ-24	1.31	9.21	0.06
CONSVQ-24	1.34	11.1	0.00
SVQ-26	1.18	5.81	0.00
CONSVQ-26	1.21	6.18	0.00

In Table 3 and 4, we present the quantization results of SVQ and CONSVQ at 24 bits/frame as varying the numbers of selected subvector, M_1 and M_2, respectively..

Table 3. Average spectral distortion and outliers for SVQ with various M_1 and M_2.

(M_1, M_2)	SD	outlier (%)	
	(dB)	2 - 4 dB	> 4 dB
(3, 1)	1.28	8.26	0.06
(4, 2)	1.24	6.12	0.03
(4, 3)	1.22	5.44	0.00
(6, 6)	1.21	4.74	0.00

Table 4. Average spectral distortion and outliers for CONSVQ with various M_1 and M_2.

(M_1, M_2)	SD	outlier (%)	
	(dB)	2 - 4 dB	> 4 dB
(3, 1)	1.30	9.24	0.00
(4, 2)	1.25	6.55	0.00
(4, 3)	1.24	5.57	0.00
(6, 6)	1.22	4.89	0.00

As we can see in Table 3 and 4, the improvements in performances by using a multiple search are significant in both SVQ and CONSVQ. In addition to the decease in spectral distortion, multiple search significantly reduces the percentages of outlier, which is very meaningful in real applications. However, the improvements resulting from multiple search are not significant unless M_1 and M_2 are sufficiently large. In case of $M_1 = 4$ and $M_2 = 2$, that is $M_{total} = 8$, the amount of decrease in spectral distortion get significant. In this case, the additional computational complexity in our implementation is nearly the same as the overall computational complexity of 24 bit SVQ. This increased complexity by multiple search can be compensated in CONSVQ at a rate above 26 bits/frame. On the contrary, the increased complexity makes the MS-SVQ unattractive one.

6. CONCLUSIONS

In this paper, we show the efficiencies of multiple search for SVQ and CONSVQ. Though the improvement of performance is considerable, it is difficult to use a multiple search by combining with SVQ because of its huge additional computational complexity. On the contrary, the increased complexity stemmed from the multiple search can be nearly compensated in CONSVQ at a rate above 26 bits/frame. Therefore, we can construct a more efficient quantization scheme of LSP, the MS-CONSVQ, that provides considerably improved performance while maintaining the computational complexity.

REFERENCES

[1] W.P. LeBlanc, B. Bhattacharya, S.A. Mahmoud and V. Cuperman, "Efficient Search and Design Procedures for Robust Multi-Stage VQ of LPC Parameters for 4 kb/s Speech Coding," *IEEE Trans. on Speech and Audio Processing*, vol. 1, pp. 373-385, 1993.

[2] D. Chang, Y. Cho and S. Ann, "Efficient Quantization of LSF Parameters Using Classified SVQ Combined with Conditional Splitting," *Proc. ICASSP-95*, pp. 736-739, 1995.

[3] K. Paliwal and B. Atal, "Efficient Vector Quantization of LPC Parameters at 24 Bits/Frame," *Proc. ICASSP-91*, pp. 661-664, 1991.

[4] G.S. Kang and L.J. Fransen, "Application of Line-Spectrum Pairs to Low-Bit-Rate Speech Encoders," *Proc. ICASSP-85*, pp. 244-247, 1985.

[5] F.K. Soong and B.H. Juang, "Optimal Quantization of LSF Parameters," *IEEE Trans. on Speech and Audio Processing*, vol. 1, pp. 15-24, 1993.

[6] M. Vetterli and H.J. Nussbaumer, "Simple FFT and DCT Algorithms with Reduced Number of Operations," *Signal Processing* (North Holland), vol. 6, pp. 267-278, 1984.

[7] F. Kossentini and M.J.T. Smith, "A Fast Searching Technique for Multistage Residual Vector Quantizers," *IEEE Signal Processing Letters*, vol. 1, pp. 114-116, 1994.