Proceedings of the World Congress on Engineering and Computer Science 2009 Vol II
WCECS 2009, October 20-22, 2009, San Francisco, USA

Enhanced Differential Split Vector Quantization of Line Spectrum Pairs for CELP-Type Coders in Packet Networks

Fatiha Merazka

Abstract— In this paper we propose intra-frame quantization methods of Line Spectrum Pairs (LSP) parameters to improve frame erasures for ITU-T G723.1 coder. The standard ITU-T G723.1 coder uses an inter-frame quantization of the LSP parameters which causes error propagation to the next frames. Simulations results show that our intra-frame quantization achieve smaller average spectral distortions than that of the embedded in the G723.1. Enhanced modified bark spectral distortion (EMBSD) tests under various packet loss conditions confirm that the proposed method is superior to the interframe algorithm embedded in the G723.1.

***Index Terms*— VoIP, Error propagation, Intraframe quantization, Spectral distortion measure, EMBSD.**

I. Introduction

The quality of real time voice communications in packet switched transmission or mobile links is degraded by frame erasures. In voice communication over IP networks, the packet loss is caused by the transmission impairments such as the process of the transmission capacity and congestion. Since even a single missing packet may generate an audible artifact in the decoded speech signal, the receiver needs a packet loss concealment method to minimize the quality degradation at the packet loss regions.

Most packet loss concealment (PLC) algorithms embedded in the standards speech coders are based on an extrapolation method or a repetition method in which the speech coding parameters are extrapolated or repeated from the parameters of the last good frame received. Since the lost packet causes the corruption of the long term prediction memory, extra performance degradation may occur from the use of the incorrect memory even at the received frames in the future. Many error concealment schemes for CELP based coders were proposed in order to reduce the quality degradation and the error propagation problem [1]-[6] .

Forward Error Concealment schemes are effective when the network loss is predictable and extra bandwidth is provided. Decoder based concealment is of relevance for bandwidth limited applications. Coded linear prediction (CELP) coded speech frames are adequate for this technique since many coding parameters show good smoothness between frames. Some ITU speech coders [7]-[8] have built-in mechanisms that process the erased frames based on predictive recovery. The parameters of the lost frames are recovered from previous good frames only.

However, both of the coders G723.1 [9] and G729 [10] quantize Line Spectrum Pairs (LSP) parameters via predictive methods. Predictive concealment can cause error propagation to subsequent frames.

In this paper, we propose to quantize the LSP parameters using an intraframe vector quantization method. We apply the proposed scheme to the ITU-T G723.1 Conjugate-structure Algebraic CELP (CS-ACELP) speech coder which is very used in voice over IP (VoIP) applications. We compare the performance of the proposed method with the embedded standard method by measuring the spectral distortion and the enhanced modified bark spectral distortion [11].

This paper is organized as fellows. In section 2, the proposed methods are presented. Comparison and evaluation results are presented in section 3. section 4 concludes our work.

II. The Proposed Methods

Line Spectrum Frequencies (LSF) parameters are well known for their ordering property [12]-[13], which states that within each frame, LSF's are strictly in ascending order with their indexes as shown in figure 1. We can see from this figure that LSP in medium frequencies are more variable than the LSP at high and low frequencies. They are also known for their intraframe and interframe correlation.

The localized sensitivity property of the LSP's makes them ideal for split vector quantization as the individual parts of an LSP vector can be independently quantized without a leakage of quantization distortion from one spectral region to another [12].

We propose intraframe quantization methods to quantize the LSP parameters and compare their performance to the PSVQ method embedded in the standard G723.1.

The first method is a differential split VQ namely DSVQ, the second method is an enhanced DSVQ namely EDSVQ.

Manuscript received July 26, 2009. Dr Fatiha Merazka Author is with the Electronic & Computer Engineering Faculty, University of Science & Technology Houari Boumediene, P.O.Box 32, El Alia, 16111 Algiers, Algeria. Phone: 213-21-247187;fax: 213-21- 24 71 87; e-mail: fmerazka@hotmail.com).

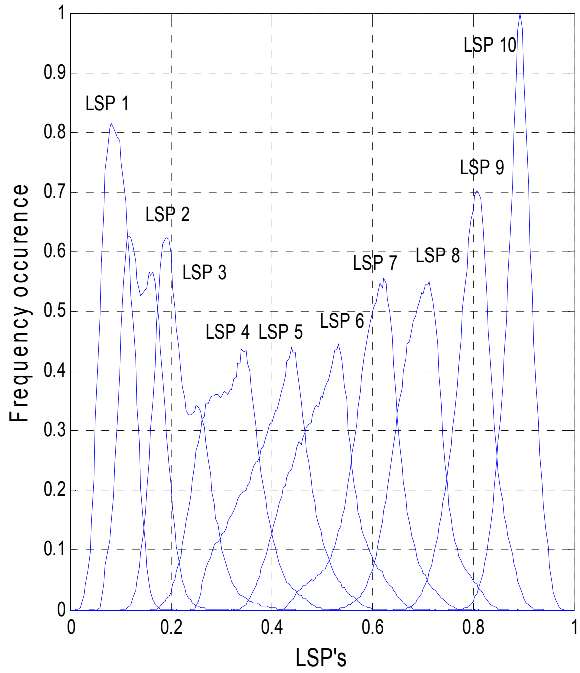

Figure 1.Histograms of LSP parameters.

The DSVQ method is shown in Figure 2 and the basic procedure of it is:

At the coder stage:

1. Compute the differences as:

$$\begin{cases} \Delta p_1 = p'_1 \\ \Delta p_i = p'_i - p'_{i-1} \ldots \ for \quad i = 2,10 \end{cases}$$

2. Quantize Δp^j to $\tilde{\Delta} p^j$:

$$\Delta p^j \xrightarrow{Q^j} \tilde{\Delta} p^j \ldots for \ \ j = 1, j_{\max}$$

At the decoder stage:

3. $\tilde{p}'_i = \tilde{\Delta} p_i + \tilde{p}'_{i-1} \ldots for\, i = 2,10$

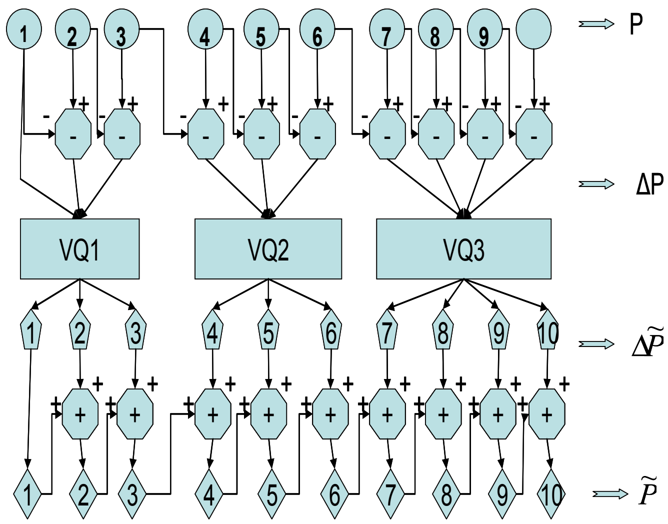

Figure 2 .DSVQ scheme.

The EDSVQ method is shown in Figure 3 and the basic procedure of it is:

At the coder stage:

1. construct the first sub-vector of LSP differences

$$\begin{cases} \Delta p_1^1 = p'^1_1 \\ \Delta p_i^1 = p'^1_i - p'^1_{i-1} \ldots for\, i = 2, i_{\max} \end{cases}$$

2. quantize the first sub-vecor

$$\Delta p_i^1 \xrightarrow{Q1} \tilde{\Delta} p_i^1 \ldots for\, i = 1, i_{\max}$$

3. Construct the last element *LSP* of the first sub-vector.

$$\tilde{p}'^1_{i_{\max}} = \tilde{\Delta} p^1_{i_{\max}} + \tilde{p}'^1_{i_{\max}-1}$$

4. compute ΔLSP of sub-vector j

$$\begin{cases} \Delta p_1^j = p'^j_1 - \tilde{p}'^{j-1}_{i\max} \\ \Delta p_i^j = p'^j_i - p'^j_{i-1} \ldots for\, i = 2, i_{\max} \end{cases}$$

5. quantize the sub-vector j

$$\Delta p_i^j \xrightarrow{Q^j} \tilde{\Delta} p_i^j \ldots pour\, i = 1, i_{\max}$$

ISBN:978-988-18210-2-7

6. compute the last element *LSP* of sub-vector j.

$$\tilde{p}'^{j}_{i\max} = \tilde{\Delta}p^{j}_{i\max} + \tilde{p}'^{j}_{i\max-1}$$

if j >number of sub-vectors, stop

otherwise, j = j+1and goto 4

At the decoder stage:

7. $\tilde{p}'_i = \tilde{\Delta}p_i + \tilde{p}'_{i-1} \ldots for \quad i = 2{,}10$

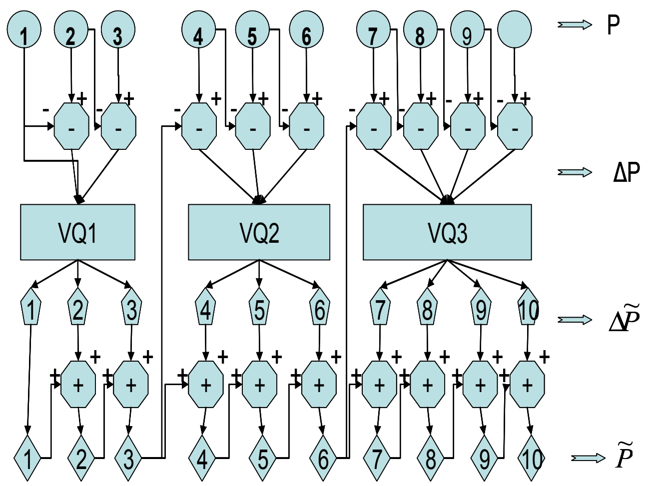

Figure 3 .EDSVQ scheme.

III. Simulation Results

In this section we compare the performance of the proposed methods with that of the embedded method in the G723.1 for speech from TIMIT database [14]. The quantizers of the sub-vectors, as shown in Figures 3 and 4, are trained using the Linde-Buzo-Gray (LBG) algorithm [15]. The performance comparison between DSVQ and PSVQ are evaluated using the average spectral distortion (Av. SD) measure [13] and the enhanced modified bark spectral distortion EMBSD, the results are depicted in table I.

TABLE I. Performance comparison between DSVQ and PSVQ

	Av.SD (dB)	2 – 4 dB	> 4dB	EMBSD
PSVQ	1.84	32.99	4.67	1.551
DSVQ	2.23	40.82	16.76	1.558

From Table I, the performance of PSVQ is slightly better than the proposed method, that's why we use an enhanced version of it namely EDSVQ, which stops the error propagation at the limit of each sub-vector. The bit allocations for the different splitting simulated are presented in Table II. The performance of EDSVQ is shown in Table III.

TABLE II. Bit allocation for EDSVQ

splitting	bit allocation		
	Sub-vector 1	Sub-vector 2	Sub-vector 3
3 3 4	7	8	9
3 4 3	8	9	7
4 3 3	9	8	7

TABLE III. Performance comparison between EDSVQ, DSVQ and PSVQ

Quantization Method		Av.SD (dB)	2 – 4 dB	> 4dB	EMBSD
PSVQ		1.84	32.99	4.67	1.551
DSVQ 334		1.86	34.82	7.76	1.558
EDSVQ	334	1.85	34.35	5.81	1.552
	343	1.86	41.90	7.18	1.555
	433	1.82	33.00	4.19	1.530

From Table III, the results show that EDSVQ method performs better than DSVQ one for the three splitting presented. The EDSVQ method outperforms the PSVQ by 0.02 dB.

We simulate real-time voice over packet networks where each packet contains one frame. Packet loss is approximated by a Gilbert random process which emphasizes the bursty nature of Internet packet loss as in Figure 4. Let state "0" stand for a packet being correctly received and "1" be a packet being erased. Let the P be the transition probability from "0" to "1" and Q be the probability from "1" to "0"

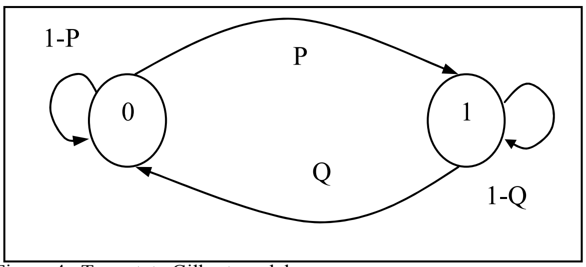

Figure 4. Two-state Gilbert model.

Table IV shows the performance of PSVQ for different loss rates.

The decoder was modified so that if a frame erasure occurs, and if the next frame is not lost as well, interpolative concealment is applied instead of the embedded method in G723.1 [9].

The LSP parameters are linearly interpolated from previous and next good frames. For the frame recovery method from [6] is used. The obtained results are tabulated in Table V.

TABLE IV. PERFORMANCE OF PSVQ METHOD FOR DIFFERENT LOSS RATES

Loss Rates (%)	Av. DS (dB)	2-4 (dB)	> 4 (dB)	EMBSD
0	1.84	32.99	4.67	1.551
10	2.07	36.95	7.71	2.567
20	2.31	41.28	10.98	3.688
30	2.54	45.76	13.93	4.610
40	2.77	48.97	17.44	4.944

TABLE V. PERFORMANCE OF EDSVQ METHOD FOR DIFFERENT LOSS RATES

Loss Rates (%)	Av. SD (dB)	2-4 (dB)	> 4 (dB)	EMBSD
0	1.82	33.55	4.19	1.530
10	1.96	35.43	5.83	2.086
20	2.07	37.54	7.47	2.771
30	2.19	38.99	9.29	3.169
40	2.31	41.05	11.24	3.390

From Table IV and V, results show that our proposed method (EDSVQ) outperforms significantly the embedded method in the standard G723.1.

IV. CONCLUSION

In this paper we have presented an efficient method for stopping error propagation by an intraframe method namely EDSVQ with interpolation for the standard G723.1. Our proposed method outperforms significantly the method embedded in the G723.1 for different loss rates.

REFERENCES

[1] D. J. Goodman et al., “Waveform substitution techniques for recovering missing speech segments in packet voice communications,” IEEE Trans. ASSP, vol. ASSP-34, no. 6, pp. 1440–1448, Dec. 1986.
[2] de Martin, J.C, Unno, T. and Viswanathan, V. “Improved frame erasure concealment for CELP-based coders,” in Proc.ICASSP’00, vol. 3, pp.1483-1486.
[3] Hong Kook Kim and Hong-Goo Kang, “A Frame Erasure Concealment Algorithm Based on Gain Parameter Reestimation for CELP coders,” in 1EEE Signal ProcessingLerrers, vol. 8, pp.252-256, Sept 2001.
[4] P. Gournay and K. D. Anderson, “Performance analysis of a decoder based time scaling algorithm for variable jitter buffering of speech over packet networks,” in Proc. Int. Conf. Acoust., Speech, Signal Process.,Toulouse, France, Mar. 14–19, 2006, pp. I-17–I-20.
[5] N. Jayant and S.W. Christensen, “Effect of packet losses in waveform coded speech and improvements due to an odd-even sample-interpolation procedure,” IEEE Trans. Commun. Vol. 29 NO. 2, Feb. 1981.
[6] J-C. Bolot, “Analysis and control of audio packet loss in the Internet,” NOSSDAV 95.
[7] M. Podolsky, C. Romer and S. McCanne, “Simulation of FEC-based error control for packet audio on the Internet,” Proceedings IEEE Infocom, vol.2, pp. 505-515. April 1998.
[8] J. C. Bolot, S. Fosse-Parisis, and D. Towsley, ”Adaptive FEC-Based Error Control for Interactive Audio in the Internet,” Proceedings IEEE Infocom 1999, New York, NY, March 1999.
[9] ITU, ITU-T G.723.1: Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 and 6.3 kbit/s, ITU 1996.
[10] ITU, ITU-T G.729: CS-ACELP Speech Coding at 8 kbit/s, ITU 1998.
[11] S. Voran, “Objective estimated of perceived speech quality-part II: evaluation and measuring normalized block technique,” IEEE Trans. Speech and audio processing, vol. 7, NO 4, pp. 371-382, July 1999.
[12] F.Itakura, “Line spectrum representation of linear predictive coefficients of speech signals", J.Acoust. Soc. Amer., vol. 57, suppl. 1, p. S35(A),1975.
[13] K. K. Paliwal and B.Atal, "Efficient Vector quantization of LPC Parameters at 24 bits/frame," ICASSP, pp. 661-664, Mar. 1991.
[14] NIST,Timit Speech Corpus, NIST 1990.
[15] Y. Linde, A. Buzo, and R. M. Gray, “An algorithm for vector quantizer design,” IEEE Trans. Commun., Technol., vol. COM-28, No. 1, pp. 84-95, Jan. 1980.