Proceedings of the World Congress on Engineering 2008 Vol I
WCE 2008, July 2 - 4, 2008, London, U.K.

Interleaving Concealment Method for CELP-Based Coders in Packet Networks

Fatiha Merazka

Abstract—**This paper presents an interleaving concealment method to improve speech quality deterioration caused by packet losses for CELP based coders. We applied the proposed scheme to the standard ITU-T G729 standard speech coder to evaluate the proposed method. The perceptual evaluation of speech quality (PESQ) and enhanced modified bark spectral distortion (EMBSD) tests under various packet loss conditions confirm that the proposed algorithm is superior to the concealment algorithm embedded in the G729. The spectral distortion measure is also used as an objective distortion measure; the obtained results prove that the interleaving method is better at the expense of extra delay.**

Index Terms— **VoIP, ITU G729, interleaving concealment, Spectral distortion measure, EMBSD, PESQ**

I. INTRODUCTION

In packet-switched telephony applications such as voice over IP (VoIP) some packets may be lost due to excess of the transmission capacity or 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 quality degradation. Recently, many error concealment algorithms for Code Excited Linear Predictive (CELP) type coders were proposed in order to minimize the quality degradation and the error propagation problem. Some of them tried to accurately estimate the excitation signal of the missing packets by a voicing classification [1][2]. Others efficiently estimated the gain parameters of the lost and successive frames [3][4].

In this paper, we present an interleaving concealment scheme for CELP based coders. We apply this method to the ITU-TG729 Conjugate-Structure Algebraic CELP (CS-CELP) speech coder [5] that is widely used in VoIP applications.

We compare the performance of the proposed algorithm with embedded standard method by measuring the average spectral distortion of Line Spectrum Frequencies (LSF) [6], before applying interleaving method and after frame concealment.

Manuscript received March 22, 2008. This work was in part, accomplished when visiting as associate, the international center for theoretical physics (ICTP) Trieste, Italy.

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:213-21-247187; e-mail: fmerazka@hotmail.com).

We also use the perceptual evaluation of the speech quality (PESQ) [7] and measure the enhanced modified bark spectral distortion [8]. This paper is organized as fellows. In section 2, we briefly review frame erasure concealment algorithm embedded in the ITU-T G729 standard speech coder. The proposed method is presented in section 3. Comparison and evaluation results are presented in section 4. Section 5 concludes our work.

II. FRAME ERASURE CONCEALMENT OF G729

In the G729 speech coder, an erased frame is reconstructed using the speech coding parameters of the previous received good frame [5]. Once frame erasure is detected, the new parameters are generated by analyzing the spectral parameters of the last good speech frame. The method replaces the missing excitation signal of the erased frame by taking one of the similar characteristics, while gradually decaying its energy. If n-th frame is detected as an erased frame, the G.729 repeats the spectral parameters of the last received good frame to the erased frame. In addition, an adaptive codebook gain and a fixed codebook gain are obtained by multiplying predefined attenuation factors by the gains of the previous frame. To avoid excessive periodicity a long term prediction lag is increased by one to the value of the previous frame.

III. DESCRIPTION OF THE INTERLEAVING CONCEALMENT METHOD

In interleaving method, the data in N consecutive frames can be mixed together before transmission [9]. This way, loss of a packet destroys only a few bits from each frame. Assuming the coder is more robust to bit errors than frame erasures (which is generally true), this approach may lessen the effect of loss. However, it does so at the expense of the substantial delays. Figure 1 shows 4 interleaved frames before and after transmission. At the coder side, the frames of 80 bits are divided into 4 sub-frames of 20 bits each and interleaved as shown in Fig.1. After transmission, the loss of a single packet from an interleaved stream results in multiple small gaps in the reconstructed stream.

Fig. 2 shows an example of speech quality degradation when frame loss is occurred. A frame of 10 ms is erased. Interleaving method spread the loss in small gaps. The proposed method gives an improved waveform shape.

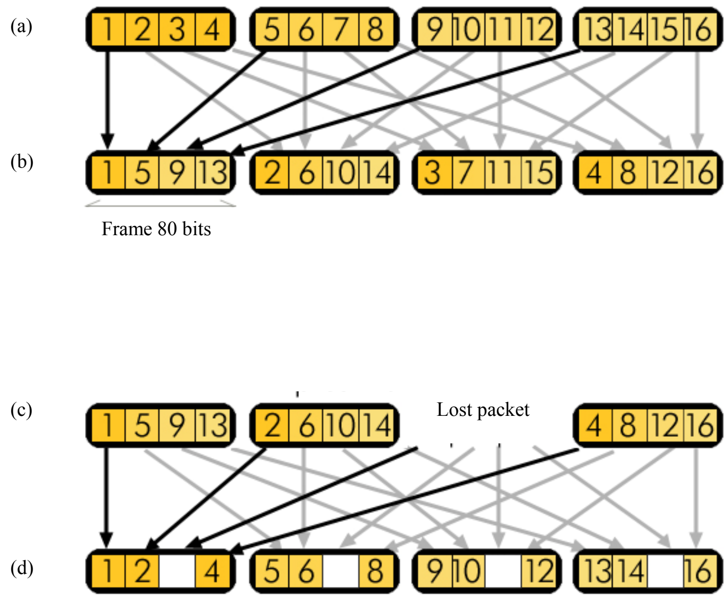

Fig. 1. Interleaving packet concealment scheme, (a) original frames, (b) frames interleaved, (c) frame loss, (d) reconstructed frames.

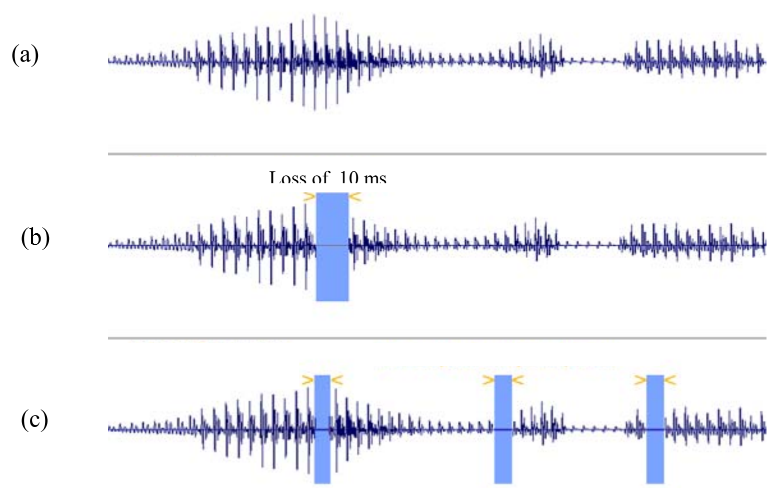

Fig. 2. Example of speech quality degradation due to frame loss, (a) original speech, (b) frame loss with G729, (c) reconstructed speech by interleaving

IV. Experimental Results

In this section we compare the performance of the proposed method with that of the embedded method in the G729. We simulate real-time voice over packet networks where each packet contains one frame. Packet loss is approximated by a Markov random process which emphasizes the bursty nature of Internet packet loss as in Fig. 3. 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" and five loss rates are simulated as given in Table I.

Figs. 4 and 5 show the LSF performance under several loss rates for female and male speakers respectively from TIMIT database [10]. Outliers are tabulated in Tables II and III for female and male speakers respectively.

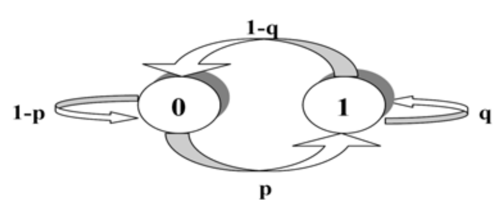

Fig. 3. Two-state Markov model.

Table I. Simulated loss rates

rate(%)	p	q
0	0.00	0.00
10	0.10	0.15
20	0.20	0.30
30	0.30	0.35
40	0.30	0.40

ISBN:978-988-98671-9-5

TABLE II. OUTLIERS OF LSF SPECTRAL DISTORTION WITH PACKET LOSS

Loss Rate (%)	Original G729			4 frames interleaved		
	Av. Spect Dist (dB)	Outliers (%)		Av. Spect Dist (dB)	Outliers (%)	
		2-4 dB	>4dB		2-4 dB	>4dB
0	1,22	6,95	0,05	1,22	6,95	0,05
10	2,94	8,51	2,54	2,54	7,14	1,95
20	3,96	35,05	23,85	3,63	28,80	18,94
30	4,63	45,65	26,25	4,25	40,34	26,25
40	5,05	47,80	26,30	4,67	44,11	26,30

These results show that 0.28 dB and 0.40 dB improvements on average spectral distortion over the original G729 for female and male speakers respectively. The number of outliers is substantially reduced under frame erasures for both female and male speakers.

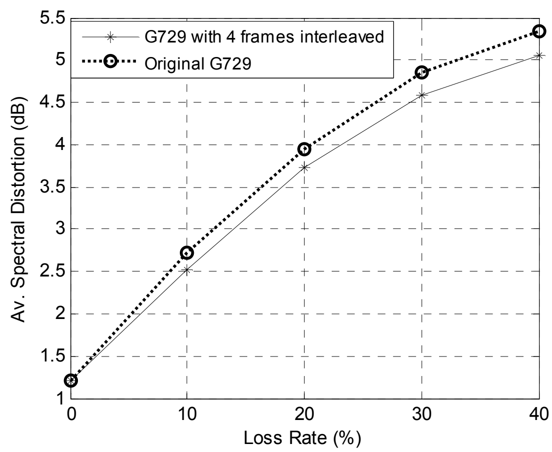

Fig. 4. Comparison of Average Spectral distortion for G729 decoded female speakers with embedded method (dash line) and the proposed method (4 fames interleaved) (solid line) under different loss rates.

TABLE III. OUTLIERS OF LSF SPECTRAL DISTORTION WITH PACKET LOSS

Loss Rate (%)	Original G729			4 frames interleaved		
	Av. Spect Dist (dB)	Outliers (%)		Av. Spect Dist (dB)	Outliers (%)	
		2-4 dB	>4dB		2-4 dB	>4dB
0	1,21	6,70	0,20	1,21	6,70	0,20
10	2,72	23,85	9,95	2,52	19,63	8,83
20	3,94	39,40	14,80	3,73	32,27	13,96
30	4,85	45,85	17,85	4,59	41,38	16,48
40	5,34	47,80	21,30	5,06	44,18	19,71

We use PESQ for an objective quality measure. Figs. 6 and 7 show comparison results for female and male speakers respectively. As the packet loss rate increases, the PESQ scores of the two algorithms decrease. The scores of the proposed algorithm are higher than the embedded method in the G729 standard coder.

We also performed an EMBSD test and the results are depicted in Figs. 8 and 9. As the packet loss rate increases, the EMBSD of the two methods increase. It is shown that the proposed algorithm is always better than the embedded method in the G729 standard coder.

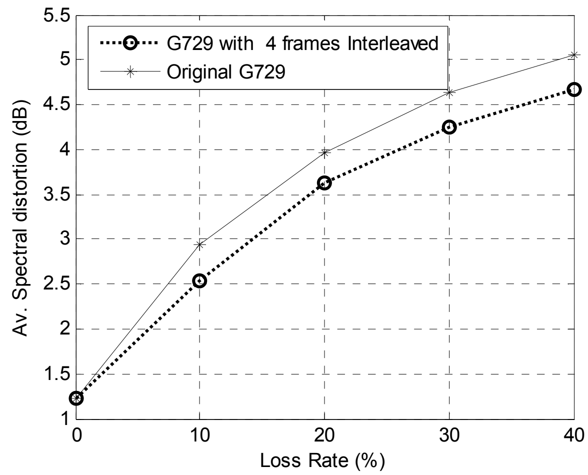

Fig. 5. Comparison of Average Spectral distortion for G729 decoded male speakers with embedded method (dash line) and the proposed method (4 fames interleaved) (solid line) under different loss rates.

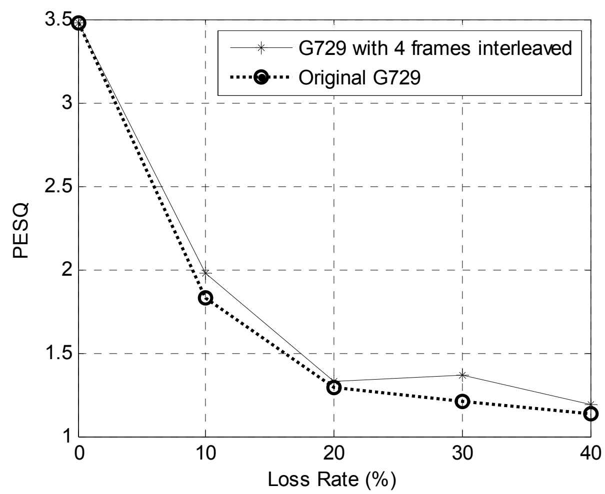

Fig. 6. Comparison of PESQ for female speakers decoded with original G729 (dash line) and the proposed method (4 fames interleaved) (solid line) under different loss rates.

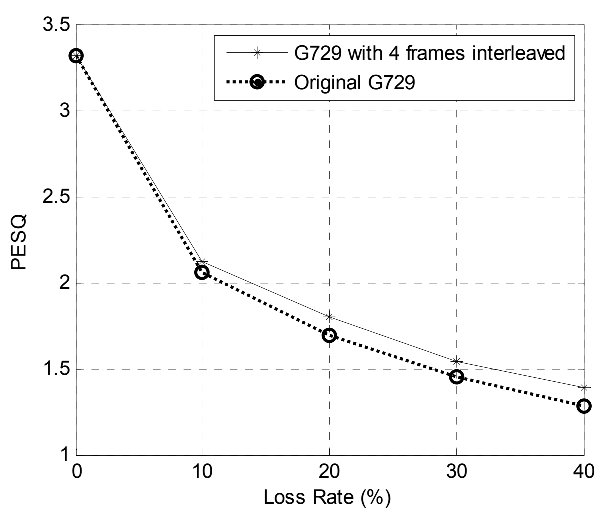

Fig. 7. Comparison of PESQ for male speakers decoded with original G729 (dash line) and the proposed method (4 fames interleaved) (solid line) under different loss rates.

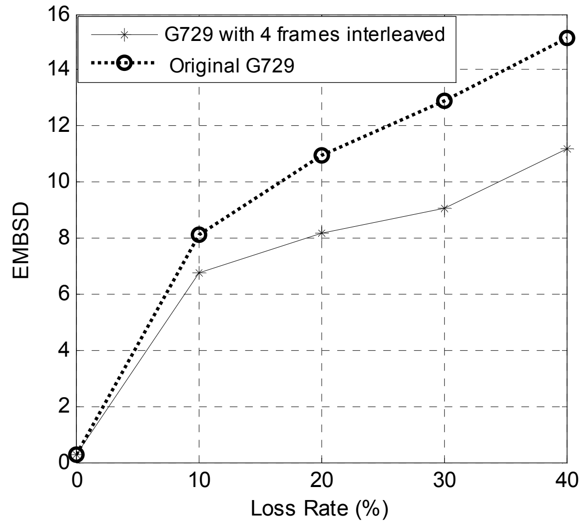

Fig. 8. Comparison of EMBSD for female speakers decoded with original G729 (dash line) and the proposed method (4 fames interleaved) (solid line) under different loss rates.

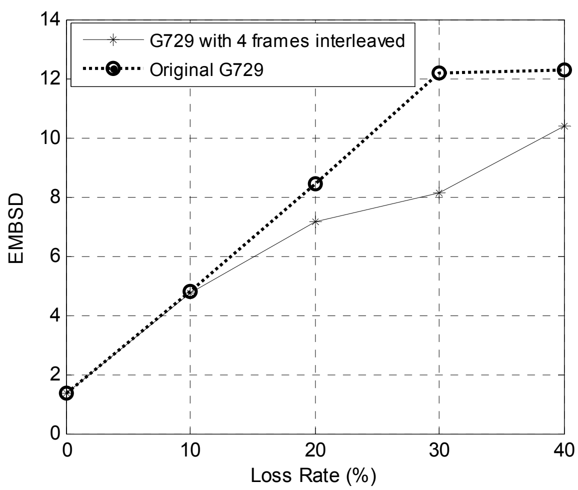

Fig. 9. Comparison of EMBSD for male speakers decoded with original G729 (dash line) and the proposed method (4 fames interleaved) (solid line) under different loss rates.

V. CONCLUSION

In this paper we have presented an efficient method for reconstructing the missing frames for CELP based coders and compared its performance with the embedded algorithm in the standard G729 coder. The proposed algorithm interleaves 4 frames in order to spread out the error. From PESQ measurement and EMBSD tests under a variety of frame erasure conditions, we found that the proposed method, improved significantly the speech quality compared to the embedded algorithm in the standard G729 coder. The objective measures given by the average spectral distortion measure verify that the interleaving method is better at the expense of extra delay.

REFERENCES

[1] A. Huaain and V. Cupeman, "Reconstruction of missing packets for CELP-based speech coders," in Proc. ICASSP-95, vol. 1, 1995, pp.245-248.
[2] Jhing-Fa Wang, Jia-Ching Wang, Jar-Ferr Yang and Jian-Jia Wang, "A voicing-driven packet loss recovery algorithm for analysis-by-synthesis predictive speech coders over Internet."in Multimedia, IEEE Transaction , vol. 3, pp.98-107, March 2001
[3] Hong Kook Kim and Hong-Goo Kang, "A Frame Erasure Concealment Algorithm Based on Gain Parameter Reestimation for CELP coders," in 1EEE Signal Processing Letters, vol. 8, pp.252-256, Sept 2001.
[4] De Martin, J.C, Unno, T. and Viswanathan, V. ''Improved frame erasure concealment for CELP-based coders," in Proc.ICASSP'OO, vol. 3, pp.1483-1486.
[5] ITU, ITU-T G.729: CS-ACELP Speech Coding at 8 kbit/s, ITU 1998.
[6] F.Itakura, "Line spectrum representation of linear predictive coefficients of speech signals", J.Acoust. Soc. Amer., vol. 57, suppl. 1, p. S35(A),1975.
[7] W. Yang, "Enhanced Modified Bark Spectral Distortion (EMBSD): An Objective Speech Quality Measurement Based on Audible Distortion and Cognition Model," Ph.D Dissertation, Temple University, USA, May 1999.
[8] ITU-T Draft Rec P.862 "Perceptual evaluation of speech quality (PESQ), an objective method of end-to-end speech quality assessment of narrowband telephone networks and speech codecs," May. 2000.
[9] J. L. Ramsey, "Realization of optimum interleavers," IEEE Trans. Info. Theory, vol. IT-16, May 1970, pp. 338–45.
[10] [10] NIST,Timit Speech Corpus, NIST 1990.