ISSN 1392–124X (print), ISSN 2335–884X (online) INFORMATION TECHNOLOGY AND CONTROL, 2014, T. 43, Nr. 2

An Efficient Search Strategy for ACELP Algebraic Codebook by Means of Reduced Candidate Mechanism and Iteration-Free Pulse Replacement

Chao-Ping Chu[1], Cheng-Yu Yeh[2,*], Shaw-Hwa Hwang[1]

[1] *Department of Electrical Engineering, National Taipei University of Technology*
1, Sec. 3, Chung-hsiao E. Rd., Taipei 10608, Taiwan, R.O.C.
e-mail: t100319011@ntut.edu.tw, hsf@ntut.edu.tw

[2] *Department of Electrical Engineering, National Chin-Yi University of Technology*
57, Sec. 2, Zhongshan Rd., Taiping Dist., Taichung 41170, Taiwan, R.O.C.
e-mail: cy.yeh@ncut.edu.tw

crossref http://dx.doi.org/10.5755/j01.itc.43.2.4946

Abstract. This work aims to present a combined version of reduced candidate mechanism (RCM) and iteration-free pulse replacement (IFPR) as a novel and efficient way to enhance the performance of algebraic codebook search in an algebraic code-excited linear-prediction (ACELP) speech coder. As the first step, individual pulse contribution in each track is given by RCM, and the number N of candidate pulses is then specified. Subsequently, the replacement of a pulse is performed through the search over the sorted top N pulses by IFPR, and those of 2 to 4 pulses are carried out by a standard IFPR. Implemented on a G.729A speech codec, this proposal requires as few as 24 searches, a search load tantamount to 7.5% of G.729A, 37.5% of the global pulse replacement method (iteration=2), 50% of IFPR, but still provides a comparable speech quality in any case. The aim of significant search performance improvement is hence achieved in this work..

Keywords: speech codec; algebraic codebook search; reduced candidate mechanism (RCM); iteration-free pulse replacement.

1. Introduction

An algebraic code-excited linear-prediction (ACELP) based speech coding technique [1-3] is the type of technique most widely applied to digital speech communication systems, and serves as a mainstream technique adopted in a great number of speech coding standards due to the double advantage of low bit rates and high speech quality. Yet, the price paid is a high computational complexity requirement, particularly in an algebraic codebook search. The reason is quite simply that it necessitates a tremendous computational load when conducting a full search over the algebraic codebook to locate the optimal pulses. As suggested in [4], the computational load is dominated by two parts, namely, the load in a search process, and the load during the algorithm initialization phase. The former and the latter, respectively, account for 74.9 and the remaining 25.1% of the entire computational load. Provided that there is a way to reduce the computational load to a great extent, an ACELP based coding technique can be extensively applied to an embedded system on a handheld device. In this way, a high performance embedded system is not seen as required, making electronic devices cost competitive. Moreover, due to a computational load reduction, the aim of energy saving is reached for an extended operation time period.

For this sake, full search scheme is hardly adopted in most prominent speech coding standards. There have been a great number of studies proposed on search load reduction, say, the focus search in G.729 [3], and the depth-first search in G.729A, among other approaches. In recent times, a number of studies on this issue cover the least important pulse replacement [5], where the least significant pulse is replaced in an iterative manner, both the global pulse replacement (GPR) [6] and the iteration-free pulse replacement (IFPR) [7] developed on the basis of [5], and the reduced candidate mechanism (RCM) approach [8], a piece of our prior work. In RCM, individual pulse contribution is evaluated in the associated track and

* Corresponding author

sorted in descending order. Subsequently, a full search is performed on the sorted top N pulses treated as candidates. In this way, the optimal pulse combination is acquired following N searches, that is, a significant reduction in search complexity is achieved.

As a combined version of RCM [8] and IFPR [7], a novel search algorithm is presented in this study as an efficient means to further enhance the search performance. Demonstrated on a G.729A speech codec, the performance is then compared with a number of existing prominent approaches in terms of search complexity and speech quality.

2. Algebraic codebook search

With the determination of an optimal codevector as the goal of the algebraic codebook search, the codebook in G.729 is configured as tabulated in Table 1, on the basis of which each codevector contains 4 nonzero pulses extracted out of associated track. Each pulse's amplitude can be either +1 or -1.

Table 1. A structured algebraic codebook in G.729

Track	Pulse	Sign	Positions
T_0	i_0	s_0	m_0: 0, 5, 10, 15, 20, 25, 30, 35
T_1	i_1	s_1	m_1: 1, 6, 11, 16, 21, 26, 31, 36
T_2	i_2	s_2	m_2: 2, 7, 12, 17, 22, 27, 32, 37
T_3	i_3	s_3	m_3: 3, 8, 13, 18, 23, 28, 33, 38 4, 9, 14, 19, 24, 29, 34, 39

The optimal codevector $\mathbf{c}_k = \{c_k(n)\}$ is thus found by minimizing the mean squared weighted error between the original and the synthesized speeches [2, 3], defined as

$$\varepsilon_k = \|\mathbf{x} - g\mathbf{H}\mathbf{c}_k\|^2 \quad (1)$$

where $\mathbf{x}$ denotes the target vector, g a scaling gain factor, and $\mathbf{H}$ a lower triangular convolution matrix. It can be shown that the optimal codevector is the one maximizing the term Q_k:

$$Q_k = \frac{(\mathbf{x}^{\mathrm{T}}\mathbf{H}\mathbf{c}_k)^2}{\mathbf{c}_k^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{c}_k} = \frac{(\mathbf{d}\mathbf{c}_k)^2}{\mathbf{c}_k^{\mathrm{T}}\mathbf{\Phi}\mathbf{c}_k} \quad (2)$$

where $\mathbf{d} = \mathbf{x}^{\mathrm{T}}\mathbf{H}$, the correlation function, is expressed as

$$d(n) = \sum_{i=n}^{L-1} x(i)h(i-n), 0 \le n \le L-1 \quad (3)$$

where L is the speech subframe size. The correlations of $h(n)$ are contained in the symmetric matrix $\mathbf{\Phi} = \mathbf{H}^T\mathbf{H}$, where the entries are given by

$$\phi(i,j) = \sum_{i=n}^{L-1} h(n-i)h(n-j)$$

$$0 \le i \le L-1; i \le j \le L-1. \quad (4)$$

It takes a total of 8192 (8*8*8*16) searches, a tremendous computational load, to conduct a full search, i.e. repeated computations and comparisons in (2), for the identification of the optimal codevector. Therefore, a focused search method is adopted in G.729 to reduce the search times to below 1440. However, the number of searches is further reduced to 320, adopting a depth-first tree search method in G.729A.

3. Proposed approach

As referred to previously, this proposal features a combination of RCM and IFPR approaches. It is a significant finding in RCM that a pulse with a high contribution is more likely to serve as one of the optimal pulses in the associated track, whereby the hit probability can be elevated when conducting a search task. In the evaluation of individual pulse contribution, (2) is simplified into (5), where the numerator of (5) is derived from (2) and (3), and the denominator of (5) is derived from (2) and (4), respectively. Just as in (2), a higher value of Q_k^i represents a higher contribution of the ith pulse.

$$Q_k^i = \frac{d^2(i)}{\phi(i,i)}, 0 \le i \le L-1 \quad (5)$$

The aim of search performance enhancement is reached by IFPR in a non-iterative manner. In this proposal, individual pulse contribution in each track is evaluated by RCM, the value of N is determined, and then the search task is performed through IFPR. Yet, the replacement of a pulse is fulfilled through the search over the aforementioned top N pulses, while those of 2 to 4 pulses are performed by a standard IFPR subsequently.

It takes a standard IFPR 36 searches to replace a pulse, accounting for 75%, the maximum percentage, of IFPR search load. In view of this, the aim of this work is to reduce the number of searches required in IFPR to 4(N-1) from the original 36 when a pulse is replaced. The implementation of this proposal is stated as follows.

Step 1. Eq. (5) yields the individual pulse contribution, following which all the pulses in associated tracks are sorted by contributions.

Step 2. The No. 1 ranked pulse in each track is labeled as the initial codevector, and the value of N is then determined for RCM.

Step 3. A sequence of searches are conducted over the top N pulses in each track for the replacement of a pulse.

Step 4. Subsequent searches are fulfilled for the replacements of 2 to 4 pulses.

Step 5. The search process terminates at the moment the combination of the optimal pulses is done.

4. Experimental results

There are three experiments conducted in this work. The first is a search accuracy comparison between the full search and other search approaches. Subsequently, the second is a computational complexity comparison between the preceding search approaches. Lastly, the third is that various approaches are compared with ITU-T P.862 perceptual evaluation of speech quality (PESQ) [9] and ITU-T P.862.1 mean opinion score, listening quality objective (MOS-LQO) [10] as an objective measure of speech quality. The test objects are those selected out of a speech database in Chinese language, containing 9,650 syllables out of 100 sentences for a duration over 41 minutes and 495,608 subframes.

For the brevity of the following discussion, the RCM approach with N candidate pulses is abbreviated as RCM-N, $1 \leq N \leq 8$. For instance, RCM-1 symbolizes the one with merely a candidate pulse extracted out of each track. Similarly, the GPR approach with the number R of repetitions is designated as GPR-R.

Tabulated in Table 2 is the search accuracy analysis between various approaches, that is, the hit probability of individual approach against the optimal pulse identified through a full search. During the search process, the best case is the one to successfully locate 4 intended pulses, the all right case, and the worst is to locate none, the all wrong case. As tabulated in Table 2, tacking the all right case as an instance, the accuracies made by G.729A, GPR-2, IFPR and RCM-2 are 68.6438%, 76.1053%, 68.0824% and 50.3579%, respectively, while that by the proposed method falls between 17.3353% (N=1) and 67.9775% (N=8).

Tabulated in Table 3 is the comparison of the search complexity, that is, the number of searches performed and those required in the evaluation of Q_k defined in (2). It is found that G.729A requires 320 searches, GPR-2 64, IFPR 48, RCM-2 16, and the proposed method a number somewhere between 1 (N=1) and 40 (N=8). Accordingly, the search complexity is reduced as intended.

Table 3. Comparison of the search complexity between various methods

Methods		Search complexity
G.729A		320
GPR	R=1	37
	R=2	64
	R=3	91
	R=4	118
IFPR		48
RCM	N=1	1
	N=2	16
	N=3	81
	N=4	256
Proposed method	N=1	1
	N=2	16
	N=3	20
	N=4	24
	N=5	28
	N=6	32
	N=7	36
	N=8	40

Table 2. Comparison of the search accuracy between various methods

Methods		Search accuracy for locating various number of intended pulses (%)				
		1 pulse	2 pulses	3 pulses	4 pulses (all right)	0 pulse (all wrong)
G.729A		98.3475	92.1456	80.9918	**68.6438**	1.6525
GPR	R=1	98.7032	90.7750	76.1053	**55.0718**	1.2968
	R=2	98.4946	91.8335	80.4779	**76.1053**	1.5054
	R=3	98.5246	92.1081	81.2227	**80.4779**	1.4754
	R=4	98.5283	92.1547	81.3393	**81.2227**	1.4717
IFPR		98.6810	92.4963	80.0048	**68.0824**	1.3190
RCM	N=1	99.3295	89.0873	55.8720	**17.3353**	0.6705
	N=2	98.3009	90.8486	73.6532	**50.3579**	1.6991
	N=3	98.8545	94.0394	83.5763	**70.7329**	1.1455
	N=4	99.2014	95.9617	89.2187	**81.8716**	0.7986
Proposed method	N=1	99.3295	89.0873	55.8720	**17.3353**	0.6705
	N=2	98.4431	90.7483	71.7603	**45.5449**	1.5569
	N=3	98.6021	91.8785	76.5752	**57.7495**	1.3979
	N=4	98.6461	92.2211	78.3002	**62.6435**	1.3539
	N=5	98.6616	92.3730	79.1315	**65.1234**	1.3384
	N=6	98.6774	92.4577	79.5883	**66.5500**	1.3226
	N=7	98.6830	92.4993	79.8454	**67.4134**	1.3170
	N=8	98.6855	92.5231	80.0217	**67.9775**	1.3145

Table 4. Comparison of the speech quality between various methods

Methods		PESQ		MOS-LQO	
		Mean	**STD**	**Mean**	**STD**
G.729A		3.8126	0.0838	3.9502	0.1026
GPR	R=1	3.8145	0.0757	3.9521	0.0883
	R=2	3.8324	0.0777	3.9725	0.0906
	R=3	3.8376	0.0739	3.9786	0.0859
	R=4	3.8382	0.0749	3.9793	0.0867
IFPR		3.8276	0.0761	3.9672	0.0884
RCM	N=1	3.7083	0.0807	3.8250	0.1026
	N=2	3.8064	0.0819	3.9423	0.0971
	N=3	3.8256	0.0740	3.9651	0.0868
	N=4	3.8331	0.0735	3.9735	0.0859
Proposed method	N=1	3.7083	0.0807	3.8250	0.1026
	N=2	3.7997	0.0715	3.9353	0.0859
	N=3	3.8162	0.0702	3.9544	0.0826
	N=4	3.8224	0.0694	3.9617	0.0810
	N=5	3.8249	0.0712	3.9644	0.0837
	N=6	3.8264	0.0706	3.9661	0.0820
	N=7	3.8278	0.0757	3.9674	0.0883
	N=8	3.8282	0.0722	3.9681	0.0840

Tabulated in Table 4 are the comparisons of PESQ and MOS-LQO, each including the mean and the standard deviation (STD). In comparison with MOS-LQO, G.729A, all the approaches provide a comparable speech quality within a 1% deviation, except that RCM-1 exhibits a 3% drop.

Furthermore, it is a point worthy of mention that with a marginal variation in MOS-LQO, a low level of search complexity signifies a superior system performance, e.g. the cases N=3 and N=4 in this proposal. For instance, it merely takes 24 searches at N=4, a figure tantamount to 7.5% of that required in G.729A, 37.5% in GPR-2, and 50% in IFPR, but provides comparable speech quality.

5. Conclusion

A combined version of RCM and IFPR is presented in this work as a novel way to perform an efficient algebraic codebook search. Implemented on a G.729A speech codec, the presented novel search strategy requires as few as 24 searches at N = 4, a search load tantamount to 7.5% of G.729A, 37.5% of GPR-2, and 50% of IFPR, but still provides a comparable speech quality. Thus, this proposal is validated as a superior candidate in the aspect of search performance. Furthermore, this improved G.729A speech codec can be utilized to improve the VoIP performance on smart phone. As a consequence, the energy efficiency requirement is met for an extended operation time period due to computational load reduction.

Acknowledgment

This research was financially supported by the National Science Council under Grant No. NSC 102-2221-E-167-027, Taiwan, Republic of China.

References

[1] **J. P. Adoul, P. Mabilleau, M. Delprat, S. Morissette.** Fast CELP coding based on algebraic codes. In: *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1987, pp.1957-1960.

[2] **C. Laflamme, J. P. Adoul, R. Salami, S. Morissette, P. Mabilleau.** 16 kbps wideband speech coding technique based on algebraic CELP. In: *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1991, pp.13-16.

[3] **R. Salami, C. Laflamme, J. P. Adoul, A. Kataoka, S. Hayashi, T. Moriya, C. Lamblin, D. Massaloux, S. Proust., P. Kroon, Y. Shoham.** Design and description of CS-ACELP: A toll quality 8 kb/s speech coder. *IEEE Trans. Speech and Audio Processing*, 1998, Vol. 6, No. 2, 116-130.

[4] **S. M. Tsai, J. F. Yang.** Efficient algebraic code-excited linear predictive codebook search. In: *IEE Proceedings Vision, Image and Signal Processing*, 2006, Vol. 153, No. 6, pp. 761-768.

[5] **H. Park, Y. Choi, D. Lee.** Efficient codebook search method for ACELP speech codecs. In: *Proc. IEEE Workshop on Speech Coding*, 2002, pp. 17-19.

[6] **E. D. Lee, M. S. Lee, D. Y. Kim.** Global pulse replacement method for fixed codebook search of ACELP speech codec. In: *Proc. Second IASTED International Conference on Communications, Internet, and Information Technology*, 2003, pp.372-375.

[7] **E. D. Lee, S. H. Yun, S. I. Lee, J. M. Ahn.** Iteration-free pulse replacement method for algebraic codebook

search. *Electronics Letters*, 2007, Vol. 43, No. 1, 59-60.

[8] **C. Y. Yeh, Y. J. Su.** Reduced candidate mechanism for an algebraic code-excited linear-prediction codebook search. *IET Communications*, 2012, Vol. 6, No. 17, 2864-2869.

[9] **ITU-T Recommendation P.862.** Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs, 2001.

[10] **ITU-T Recommendation P.862.1.** Mapping function for transforming P.862 raw result scores to MOS-LQO, 2003.

Received August 2013.