2008資訊科技國際研討會

The Novel Search Algorithm of G.729's Fixed Codebook

Chi-Jung Huang
Department of Electrical Engineering
National Taipei University of Technology
Taipei, Taiwan, ROC
t6319010@ntut.edu.tw

Shaw-Hwa Hwang
Department of Electrical Engineering
National Taipei University of Technology
Taipei, Taiwan, ROC
hsf@ntut.edu.tw

***Abstract*—In this paper, a new search algorithm for the G.729's fixed codebook is proposed. In this approach, the contribution of each impulse located in the 4 groups is pre-estimation. Then the Top-N search algorithm is employed to search the sub-optimal fixed codeword. The search space of Top-8 is 8192(=8*8*8*16) and is same as the full search. The search space of Top-4 is equal to 512(=4*4*4*8). More than 93% of search space is reduced. The correct rate between Top-4 and full search is 82.7%. The MOS of listening test is 3.95 and 4.24 for Top-4 and full search, respectively. These experimental results confirm the good performance of our new search algorithm.**

Keywords-component; G.729; Fixed Codebook Search

I. INTRODUCTION

Speech compression is widely used in digital communication system, such as mobile phone, VoIP, video conference technology, etc. Speech compression reduces the redundant data and saves the bandwidth of communication channel. Recently, due to the rapidly growing of Internet, many applications move to the Internet. The VoIP is an important application and trend for future's personal communication system.

Due to the low bit rate, good speech quality, and the reasonable computation requirement, the standard G.729 is the default and important speech codec for the VoIP. The G.729 is designed by the ITU. It adopts the CS-ACELP (Code structure Algebraic Code Excited Linear Prediction) algorithm and makes the good speech quality. It converts 80 sample points into 80 bits and achieves the bit-rate with 8kbps.

Since the G.729 is based on the CELP (code excited linear prediction) model, each frame contains linear prediction coefficients, excitation codebook indices, and gain parameters that are used by the decoder structure in order to generate the synthesized speech. The codebook indices include the index of adaptive codeword and the index of stochastic codeword. The adaptive codeword is obtained from the long-term prediction and the stochastic codeword is estimated from the short-term prediction. The short-term prediction needs the close-loop search on the fixed codebook. It will need more computation than the long-term prediction.

However, the computation requirement of G.729 is about 10~20 MIPS. Although the G.729A with computation improvement version is proposed, it still needs about 10 MIPS. It can't be easily implemented in the low-end DSP chip. In the past, many researches pay their effort on the computation reduction. Most of paper paid their effort on the improvement of pitch detection. The VQ-based LSP quantization process got more attention also. The computation of fixed codebook is also dramatically large. The search space of G.729's fixed codebook is 8192. Moreover, the search space of these 4 pulse's sign is 16. Total search space is 131072. In the G.729's document, no any effective algorithm is proposed to improve the loading of computation. The large computation requirement of G.729's fixed codebook makes the G.729 not realistic for product and market.

In this paper, a novel search algorithm is proposed to improve the computation of G.729's fixed codebook. The theoretic analysis and experimental result are made to inspect and verify the performance of our search algorithm.

In the following section, the novel search algorithm of G.729's fixed codebook is given first. The experimental results and analysis is given in the section IV. The conclusion and discussion will be given in the last section.

II. THE NEW SEARCH ALGORITHM OF FIXED CODEBOOK

In the traditional G.729, the fixed codeword is estimated for each sub-frame. The length of sub-frame is 40 sample points (5ms). Each sample point of fixed codeword can be assigned with a value of -1, 0, or +1. Thus, the combination of fixed codeword is equal to $3^{40}(;\ 10^{19})$. It is dramatically large and not realistic. Thus, in the G.729's specification, these 40 positions of codeword are classified into 4 groups. The detail classification of position for each group is listed in Table.1. Only four positions located into each group are associated with non-zero value. The other position is associated with zero value. The combination of position in Table.1 is equal to 8192(8*8*8*16). The combination of sign of these four positions is equal to $16(=2^4)$. The total combination of Table.1 is equal to $131072(=8192*16;\ 10^{5.12})$. Thus, in the approach of 4 groups, the search space is reduced from 10^{19} to $10^{5.12}$ for each sub-frame. The search space is $10^{7.42}$ per 1 second. However, the search space of 4 groups approach is still dramatic large. The implementation of G.729 on DSP can't be easily achieved.

Table.1 The Structure of Fixed codebook.

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

In the G.729's C code designed by ITU, the search space of the first subframe is 105 times and the second subrame is 75 times. Most of the 8192 search space is not done. The search result is not optimal. In the G.729 document, The total search times can't exceed 180 times for 1 frame. The average search time of subframe is 90 times. Not only in the G.729's C code, but also in the G.729 document, only a small search space is done. In the past, none of algorithm is proposed to effectively or correctly search the partial search space.

In this paper, the novel search algorithm is proposed to reduce the search space of fixed codeword. In this approach, the contribution and the sign of each position is pre-estimated by close-loop analysis. The criterion of contribution is the root mean square error (RMSE) between the long-term synthesized error and the filter output of impulse for each position. Then, these 40 positions are divided into 4 groups according to the Table.1. Each group is sorted by the minimum RMSE. For each group, the best position will have the maximum contribution to reduce the RMSE and improve the synthesized speech. Thus, the best position will be aligned in the first candidate. In our approach, the fixed codeword is also same as G.729 standard. Each sub-frame will select 4 un-zero positions and locate in the 4 groups, respectively. The Top-N algorithm will be employed to select the best N-candidate position. Thus, the search space of our approach for each sub-frame is listed as Table 2.

Table-2. The search space of Top-N.

Top-N	Search Space	Total
N=1	$C_1^1 * C_1^1 * C_1^1 * C_1^2$	2(0.02%)
N=2	$C_1^2 * C_1^2 * C_1^2 * C_1^4$	32(0.39%)
N=3	$C_1^3 * C_1^3 * C_1^3 * C_1^6$	162(1.98%)
N=4	$C_1^4 * C_1^4 * C_1^4 * C_1^8$	512(6.25%)
N=5	$C_1^5 * C_1^5 * C_1^5 * C_1^{10}$	1250(15.3%)
N=6	$C_1^6 * C_1^6 * C_1^6 * C_2^{12}$	2592(31.6%)
N=7	$C_1^7 * C_1^7 * C_1^7 * C_2^{14}$	4802(58.6%)
N=8	$C_1^8 * C_1^8 * C_1^8 * C_2^{16}$	8192(100%)

The rough step of G.729 is listed as below. The detail description on the fast search algorithm of fixed codebook for each sub-frame is listed in the step 3~9.

Step #1: Estimated LSP, Gain, and Pitch.

Step #2: Calculate the Long-Term Contribution

Step #3: Estimated the predictive Error $e_{LP}[n]$ after Long-term Process

Step #4: Estimated the Impulse response of LP filter $h[n]$.

Step #5: Estimate optimal gain $\hat{g}_i$ at i-th position from the following equation

$$\frac{\partial \sum_{n=0}^{L} (e_{LP}[n] - h[n-i])^2}{\partial g_i} = 0$$

Step #6: Estimate the minimum RMSE D_i at the i-th position from the following equation.

$$\hat{D}_i = \sum_{n=0}^{L} (e_{LP}[n] - \hat{g}_i h[n-i])^2$$

Step #7: Repeat Step 4~6 for i=0,1,…,239.

Step #8: Reorder the position of Table 1 according to their minimum RMSE D_i for each group. Moreover, the sign of each position is determined by the optimal gain $\hat{g}_i$. If the $\hat{g}_i$ is positive, the sign of i-th position is +1. If the $\hat{g}_i$ is negative, the sign of i-th position is -1.

Step #9: In the group 1~3, the Top-N candidates are selected. In the group 4, the Top-2N candidates are selected. All of these 4N candidates are checked and estimated for the optimal codeword. The search space is equal to the value of $C_1^N * C_1^N * C_1^N * C_1^{2N} (= 2N^4)$.

Step #10: Packet Parameter into Bit stream.

In our novel search algorithm, the search space is reduced dramatically. For example, in the case of Top-3, more than 98% of search space is reduced. In the case of Top-7, 41.4% of search space is also reduced. The performance of computation saving is good. The other performance will be analyzed in the next section.

III. EXPERIMENTIAL RESULTS

In this paper, the novel search algorithm on G.729's fixed codebook is firstly proposed. The search space is reduced dramatically. However, the speech quality and correction rate of fixed codebook must be inspected to confirm the performance of our approach.

In our experiment model, two speech databases are selected. One is uttered by male and the other is female. The speech database is generated from the reading process on newspaper. The length of database is 1325.4 and 1839.7 seconds for male and female, respectively.

In the analysis process of speech quality, the MOS of synthesized speech for each Top-N is examined by listening test. Totally 20 persons are invited to join the process of listening test. The Table 3 lists the MOS vs. Top-N. The MOS of Top-8 is 4.24. However, the computation requirement is dramatically large. It can't be implemented in general DSP chip or PC. The computation saving and MOS of Top-1 and Top-2 can be accepted in the real application.

Table-3. The MOS of synthesized speech.

Algorithm	Computation Saving	MOS
Top-1	99.98%	3.82
Top-2	99.61%	3.91
Top-3	92.02%	3.90
Top-4	93.75%	3.95
Top-5	84.7%	4.11
Top-6	68.4%	4.23
Top-7	42.4%	4.23
Top-8(Full Search)	0%	4.24
G.729(90 times)	98.9%	3.65
Original		4.25

In the analysis of correction rate, the full search result is employed to examine the Top-N solution. The correction rate of codeword between Top-8 and Top-N is checked. The Table 5 lists the correction rate vs. Top-N.

Table.5 The correction rate between Top-8 and Top-N

Algorithm	Search Space	Correction rate
Top-1	2(0.02%)	47.5%
Top-2	32(0.39%)	62.6%
Top-3	162(1.98%)	73.9%
Top-4	512(6.25%)	82.7%
Top-5	1250(15.3%)	90.1%
Top-6	2592(31.6%)	95.4%
Top-7	4802(58.6%)	98.5%
Top-8(Full Search)	8192(100%)	100%
G.729(90 times)	90(1.1%)	1.92%

In the Table-5, the correct rate between Top-8(Full search) and the traditional G.729 is 1.92%. It is reasonable. Because the search space of traditional G.729 is about 1.1% of full space and no effective search algorithm is employed. Moreover, the correction rate of Top-1~7 is larger than the search space. The Top-1's search space is 0.02%. It is very small. However, the correction rate is about 47.5%. These results confirm that the Top-N algorithm will correctly group the optimal position and reduce the search space of fixed codebook. The novel search algorithm on G.729's fixed codebook proposed in this paper is confirmed with good performance.

In another inspection of the speech quality, the fixed codeword will not have great contribution on the speech quality. For example, in the Top-1, the correction rate is only 47.5%. However, the speech quality is still not bad. The MOS of Top-1 is 3.8. In the other case, the MOS of the traditional G.729 is 3.65. It is also reasonable good. From the above experimental results, it points out that the short-term prediction has only slight influence on the speech quality. But the search space of G.729's fixed codeword is dramatically large. Thus, the effective search algorithm is urgently needed to reduce the search space. In the other experimental result, the MOS is 1.5 for the synthesized speech without short-term prediction. Thus, the above experimental results confirm the unimportant property of fixed codeword. Moreover, the G.729 is not lack of fixed codebook. The short-term prediction can be simplified. But it can't be removed.

IV. Conclusion

In this paper, the novel search algorithm on G.729's fixed codebook is proposed. Those experimential results cofirm the good performance. It not only reduces the computation dramatically, but also maintains the good speech quality. This algorithm make the implementation of G.729 become easily for general DSP chip.

Moreover, some important conclusions are obtained. First, due to the large search space, the effective search must be employed in the G.729's fixed codebook. Second, the contribution of fixed codebook is very slight. Third, the G.729 codec is not lack of fixed codebook.

References

[1] Arian Maleki and Kambiz Nayebi, "A New Efficient PNLMS Based Algorithm for Adaptive Line Echo Cancellation", Seventh International Symposium on Signal Processing and Its Application, pp.555~558, Vol.2, July, 1~4, 2003

[2] S.H. Hwang and S.H. Chen, "Fast Encoding Algorithm for VQ-Based Image Coding, " IEE Electronics Letter,pp.1618-1619,1990.

[3] G.. Ruggeri , F. Beritelli and S. Casale, "Hybrid Multi-Mode/Multi-Rate CS-ACELP Speech Coding For Adaptive Voice Over IP", IEEE,2001.

[4] Cheng Deyuan, "An 8kb/s Low Complexity ACELP Speech Codec" ICSP'96.

[5] R.Salami,C.Laflamme,and J-P.Adoul "ACELP Speech Coding At 8kbit/s With A 10ms Frame A Candidate For CCITT Standardization" Quebec,Canada J1k 2R1.

[6] Rand El-Koualy and Salwa H. El-Ramly "A New Low Bit Rate Low Delay Algebraic CELP (ACELP) Coder" , Seventeenth National Radio Science Conference ,Feb.22-24,2000,Minufiya University ,Egypt.

[7] Nam Kyu Ha, "A Fast Search Method of Algebraic Codebook by Reordering Search Sequence ,"pp.21~24, ICASSP'99,1999.

[8] Redwan Salami, Claude Laflamme, etc.,"ITU-T G.729 AnnexA:Reduced Complexity 8kb/s CS-ACELP Codec for Digital Simultaneous Voice and Data",pp.56~63,IEEE Communication Magazine,Sep.1997.

[9] Cheng Deyuan, "An 8kb/s Low Complexity ACELP Speech Codec", pp.671~674,Proceedings of ICSP,96,1996.

[10] ITU-T G.723,"Dual Rate Speech Coder for Multimedia Communication Transmitting at 5.3 & 6.3kbit/s," ITU-T,Oct,1995.

[11] ITU-T G.729, "Coding of Speech at 8kbit/s Using CS-ACELP," Mar.,1996.

[12] N.Naja, J. M. Boucher,and S. Saoudi "Fast LSP Vector Quantization Algorithms Comparison",pp.1127~1130, ICASSP'94, 1994.

[13] A. Kataoka, J. Ikedo, and S. Hayashi, "LSP and Gain Quantization for the proposed ITU-T 8kb/s Speech Coding Standard",pp.7~8.

[14] M. A. Ferrer-Ballester and A. R. Figueiras-Vidal, "Efficient Adaptive Vector Quantization of LPC Parameters",pp.314~317, IEEE Trans. On Speech and Audio,No.4, July, 1995.

[15] Schroeder M. R.f and B. S. Atal, "Code-excited Linear prediction(CELP): High quality at very low bit rates," pp.937~940, ICASSP85, 1985.

[16] J. D.Markel, and A. H. Gray, "Linear Prediction of speech ," New York,1980.

[17] B. S. Atal, and J. R. Remde, "A New Model of LPC Excitation for Producing Natural Sounding Speech at Low Bit Rates,"pp614~617,ICASSP82,1982.

[18] P.Kroon, E. D. Deprettere and R. J. Sluyter, "Regular-Pulse Excitation : A Novel Approach to Efficient Multipulse Coding of Speech , " IEEE Trans.On ASSP,Vol.34, No.5, PP.1054-1063,1986.

[19] A. V.McCree,etc…, "A2.4kb/s MELP Coder Candidate for the New U.S.Federal Standard," pp.200~203, ICASSP96, 1996.

[20] R.Salami, C.Laflamme,etc.., "Design and Description of CS-ACELP: A Toll Quality 8kb/s Speech Coder ," IEEE Trans . On Speech and Audio Processing,Vo1.6, No.2,Mar.,1998.

[21] Nakhai, M**.**, Marvasti, F**.,**1999**. "**A Hybrid Speech Coder Based on CELP and Sinusoidal Coding **" .** IEICE Trans **.** Inf**.** & Syst**.,** Vol**.** E82-D**,** No.8, August, 1190-1199 .

[22] Kim, H. k., Lee , H. S.,1999. " A Low-Bit Rate Extension Algorithm to the 8kbit/s CS-ACELP Based on Adaptive Fixed Codebook Modeling " **.** IEICE Trans **.** Inf**.** & Syst**.,** Vol**.** E82-D**,** No.7, July, 1087-1092 .

[23] N. Kitawaki, M. Honda, and K. Itoh, " Speech-quality assessment methods methods for speech-coding system", IEEE Commun . Mag.,vol. 22, pp. 26-33,1984.

[24] Kurihara, A. S., Hayashi, S. and Moriya, T., 1996. "Improved CELP-based coding in a noisy environment using a trained sparse conjugate codebook ". IEICE Trans. Inf & Syst., Vol. E79-D, No. 2, February, 123-129.

[25] Suddle, M.R., Kondoz, A. M. and Evans, B. G., 1991. "DSP Implementation of low bit-rate CELP based speech coders ".Digital Processing of Signals in Communication Sixth International Conference on IEE, 309-314.

[26] Gerson, I. A. and Jasiuk, M. A., 1991. " Vector sum excited linear prediction (VSELP)". IEEE Trans. On Signal Processing, Vol. 39, No. 7,July, 1503-1515.

[27] Adoul, J. P., Mabilleau, P., Delprat, M. and Morissette, S., 1987. "Fast CELP coding based on algebraic codes".IEEE Int . Conf. On Acoust., Speech and Signal Processing,1957-1960.