

Bitrate Scalability for Multi-Pulse Based Code Excited Linear Prediction Speech Coder

¹Suphattharachai Chomphan and ²Chutarat Chompunth

¹Department of Electrical Engineering, Faculty of Engineering at Si Racha, Kasetsart University, 199 M.6, Tungsukhla, Si Racha, Chonburi, 20230, Thailand

²School of Social and Environmental Development, National Institute of Development Administration, 118 M.3, Serithai Road, Klong-Chan, Bangkok, 10240, Thailand

Abstract: Problem statement: The bitrate scalability is an important functionality that is defined in the Moving Picture Expert Group (MPEG-4) requirements. Since the traffic in a communication network varies with time, the capability of varying the bitrate of a speech coder is needed. **Approach:** This study proposes an approach to vary the bitrate by using a bitrate scalable tool which is attached with the core section of Multi-Pulse based Code Excited Linear Predictive (MP-CELP) coder. The bitrates scalable tool employs multi-stage excitation coding based on an embedded-coding approach. The multi-pulse excitation codebook at each stage is adaptively produced depending on the selected excitation signal at the previous stage. **Results:** The experimental results show that the speech quality of the proposed coder with the bitrate scalable tools is improved above the speech quality of the conventional coder. **Conclusion:** From the study, the proposed approach gives a tool to improve the speech compression and can also support a variety of coder bitrates.

Key words: Thai speech, Multi-Pulse based Code Excited Linear Predictive (MP-CELP), speech compression, bitrates scalability, MOS scores

INTRODUCTION

Nowadays, the digital communications are widely developed (Jabrane *et al.*, 2007). The audio, still image, video or text information can be transmitted through wire and wireless networks, meanwhile, the number of users to access these networks increases considerably. As a result, channel capacity should be increased, signal compression aims at treating this situation (Chompun *et al.*, 2000). The multimedia applications such as videophone and videoconference on ATM and Internet are widely used, therefore, the high quality speech codes are highly demanded. These kinds of applications require specific considerations for packet loss. To treat this problem, a bitrate-scalable coder is realized where the synthesized speech signal can be decoded from the received packets, which contain only a part of the whole encoded bitstream. One of standardization activities for such areas is undergoing at the MPEG-4 (Nomura *et al.*, 1998; Chomphan, 2010b).

The MP-CELP coder has been proposed to be a scalable coder. This speech coder employs the multi-pulse excitation which the number of pulses in fixed-

entry codebook is selective for bitrate scalability and multiple bitrates functionality according to the MPEG-4 CELP speech coder requirements, (Nomura *et al.*, 1998; Chomphan, 2010b).

In the MP-CELP core coder, amplitudes or signs for multi-pulse excitation are simultaneously vector quantized. Moreover, to improve speech quality for background noise conditions, the adaptive pulse location restriction method are utilized (Ozawa and Serizawa, 1998). This coder operates at various bitrates ranging from 4-12 kbps by applying the flexibility in multi-pulse excitation coding (Chomphan, 2010b).

This study explains mainly on a bitrate scalable tonal language speech coder based on a multi-pulse based code excited linear predictive coding (Taumi *et al.*, 1996; Ozawa *et al.*, 1997). The bitrate scalability is focused and illustrated in details.

MATERIALS AND METHODS

MP-CELP core coder: The MP-CELP core coder achieves a high coding performance by introducing a

Corresponding Author: Suphattharachai Chomphan, Department of Electrical Engineering, Faculty of Engineering at Si Racha, Kasetsart University, 199 M.6, Tungsukhla, Si Racha, Chonburi, 20230, Thailand

multi-pulse vector quantization as depicted in Fig. 1 (Taumi *et al.*, 1996; Ozawa *et al.*, 1997). The input speech of 10 ms frame is processed through Linear Prediction (LP) and pitch analysis. The LP coefficients are quantized in the Line Spectrum Pairs (LSP) domain. The pitch delay is encoded by using an adaptive codebook. The residual signal for LP and the pitch analysis is encoded by the multi-pulse excitation scheme. The multi-pulse excitation signal is composed of several non-zero pulses. The pulse positions are restricted in the algebraic-structure codebook and determined by an analysis-by-synthesis approach, (Laflamme *et al.*, 1991; Chomphan, 2010a). The pulse signs and positions are encoded, while the gains for pitch predictor and the multi-pulse excitation are normalized by the frame energy and encoded.

Bitrates scalable tool: The study applies 3 stages of the bitrate scalable tools according to the MPEG-4CELP requirement as illustrated in Fig. 2-4 (Chomphan, 2010b). One-stage bitrate scalable MP-CELP coder is illustrated in Fig. 2. Subsequently, two-stage bitrate scalable MP-CELP coder is illustrated in Fig. 3. Finally, three-stage bitrate scalable MP-CELP coder is illustrated in Fig. 4. As for one-stage bitrate scalable MP-CELP coder, the bitrate scalable tool encodes the residual signal produced at the MP-CELP core coder utilizing the multi-pulse vector quantization. As for two-stage and three-stage bitrate scalable MP-CELP coders, the bitrate scalable tool encodes the residual signal produced at the previous stage as shown in Fig. 3 and 4. Adaptive pulse position control is employed to change the algebraic-structure codebook at each excitation-coding stage depending on the encoded multi-pulse excitation at the previous stage. The algebraic-structure codebook is adaptively controlled to inhibit the same pulse positions as those of the multi-pulse excitation in the MP-CELP core coder or the previous stage. The pulse positions are determined so that the perceptually weighted distortion between the residual signal and output signal from the scalable tool is minimized. The LP synthesis and perceptually weighted filters are commonly used for both the MP-CELP core coder and the scalable tool.

The supporting core bitrates of this coder are 5600, 8200, 12200 bps, for the core coder with one pulse in fixed codebook, five pulses in fixed codebook and ten pulses in fixed codebook, respectively. As for bitrates scalable tool, each stage increases the bitrates of 800 bps. In the core coder with one pulse in fixed codebook, the 1, 2, 3 stages of scalability operate at the total bitrates of 6400, 7200 and 8000 bps respectively.

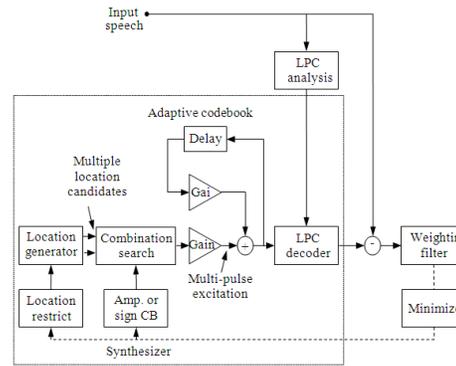


Fig. 1: MP-CELP core coder

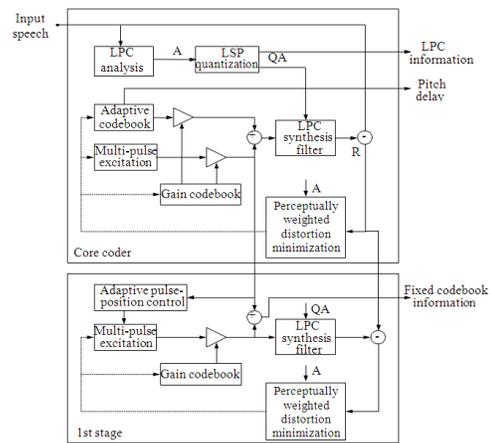


Fig. 2: One-stage bitrate scalable MP-CELP coder

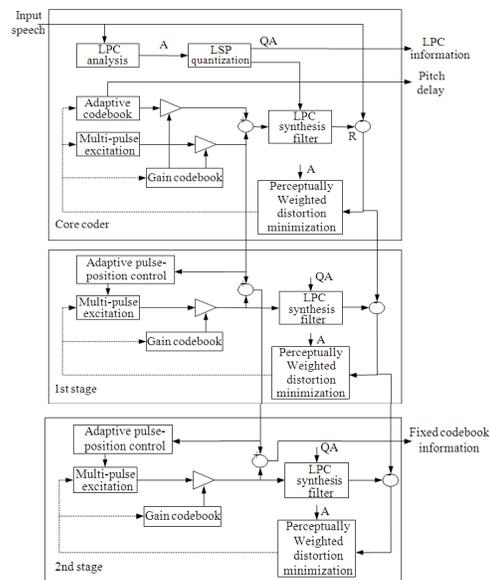


Fig. 3: Two-stage bitrate scalable MP-CELP coder

Table 1: Subjective speech quality (MOS Score)

Core bitrate (bps)	Number of stages	Toll bitrate (bps)	Male MOS score		Female MOS score	
			English	Thai	English	Thai
5600	-	5600	3.16	3.02	3.15	3.01
	1	6400	3.25	3.08	3.22	3.03
	2	7200	3.35	3.21	3.28	3.16
	3	8000	3.50	3.35	3.33	3.28
8200	-	8200	3.61	3.43	3.44	3.39
	1	9000	3.62	3.50	3.54	3.50
	2	9800	3.72	3.56	3.61	3.48
	3	10600	3.78	3.61	3.71	3.56
12200	-	12200	3.89	3.78	3.84	3.72
	1	13000	3.98	3.81	3.88	3.80
	2	13800	4.01	3.88	3.91	3.83
	3	14600	4.02	3.93	3.96	3.92

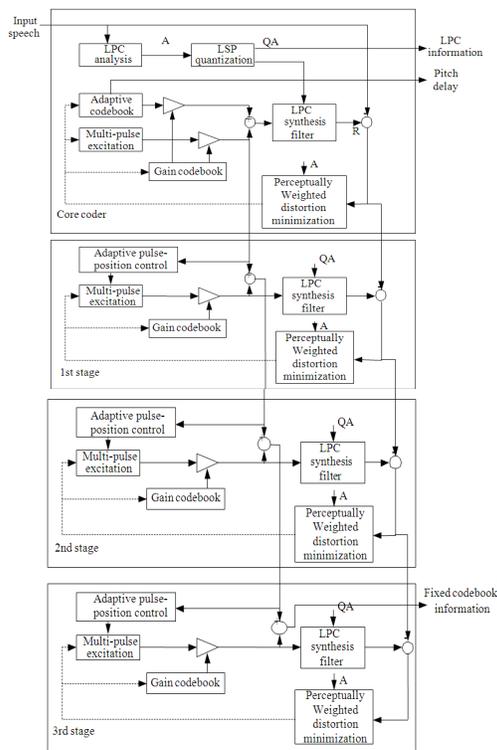


Fig. 4: Three-stage bitrate scalable MP-CELP coder

Moreover, in the core coder with five pulses in fixed codebook, the 1, 2, 3 stages of scalability operate at the total bitrates of 9000, 9800 and 10600 bps respectively. Finally, in the core coder with ten pulses in fixed codebook, the 1, 2, 3 stages of scalability operate at the total bitrates of 13000, 13800 and 14600 bps respectively.

RESULTS

The coding quality of the bitrate-scalable coder was evaluated subjectively by using 36 tested

sentences from 10 men and 10 women in English and Thai. Table 1 summarizes the Mean Opinion Score (MOS) of the bitrate-scalable coder.

DISCUSSION

From the experimental results, Table 1 shows that the coding quality is improved when the bitrate scalability is applied. In other words, when the bitrate is increased by the bitrate scalable tool the MOS score is also increased. The female coding quality is a little bit lower than that of male coding quality. Comparing between English and Thai, the coding quality of Thai is lower than that of English. The reason is that Thai is a tonal language where the pitch change must be more considered. The pitch resolution of Thai should be treated with higher level.

CONCLUSION

In this study, a bitrate scalable speech coder is described. The compression algorithm consists of a MP-CELP core coder and the bitrates scalable tools. The proposed approach can support a number of coding rates. Moreover, the coding quality is improved when the bitrate scalability is applied.

REFERENCES

Chomphan, S., 2010a. Performance evaluation of multi-pulse based code excited linear predictive speech coder with bitrate scalable tool over additive white gaussian noise and rayleigh fading channels. *J. Comput. Sci.*, 6: 1438-1442. DOI: 10.3844/jcssp.2010.1438.1442

Chomphan, S., 2010b. Multi-Pulse based code excited linear predictive speech coder with fine granularity scalability for tonal language. *J. Comput. Sci.*, 6: 1288-1292. DOI: 10.3844/jcssp.2010.1288.1292

- Chompun, S., S. Jitapunkul, D. Tancharoen and T. Srithanasan, 2000. Thai speech compression using CS-ACELP coder based on ITU G.729 standard. Proceedings of the 4th Symposium on Natural Language Processing, (SNLP'00), Thailand, pp: 1-5.
- Jabrane, Y., R. Iqdour, B.A.E. Said and N. Naja, 2007. MAI cancellation in DS/CDMA using a new approach on WDS. *Am. J. Applied Sci.*, 4: 736-740. DOI: 10.3844/ajassp.2007.736.740
- Laflamme, C., J.P. Adoul, R. Salami, S. Morissette and P. Mabilieu, 1991. 16 kbps wideband speech coding technique based on algebraic CELP. Proceedings of the International Conference on Acoustics, Speech and Signal Processing, Apr. 14-17, IEEE Xplore Press, Toronto, Ont., Canada, pp: 13-16. DOI: 10.1109/ICASSP.1991.150267
- Nomura, T., M. Iwadare, M. Serizawa and K. Ozawa, 1998. A bitrate and bandwidth scalable CELP coder. Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, May 12-15, IEEE Xplore Press, Seattle, USA., pp: 341-344. DOI: 10.1109/ICASSP.1998.674437
- Ozawa, K., T. Nomura and M. Serizawa, 1997. MP-CELP speech coding based on multipulse vector quantization and fast search. *Elect. Commun. Japan Part III: Fundamental Elect. Sci.*, 80: 55-63. DOI: 10.1002/(SICI)1520-6440(199711)80:11<55::AID-ECJC6>3.0.CO;2-R
- Ozawa, K. and M. Serizawa, 1998. High quality multipulse based CELP speech coding at 6.4 kb/s and its subjective evaluation. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, May 12-15, IEEE Xplore Press, Seattle, USA., pp: 153-156. DOI: 10.1109/ICASSP.1998.674390
- Taumi, S., K. Ozawa, T. Nomura and M. Serizawa, 1996. Low-delay CELP with multi-pulse VQ and fast search for GSM EFR. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, May 7-10, IEEE Xplore Press, Atlanta, USA., pp: 562-565. DOI: 10.1109/ICASSP.1996.541158