Lecture Notes in Signal Science, Internet and Education

Proceedings of the 7th WSEAS International Conference on Signal, Speech and Image Processing (SSIP '07)

7th WSEAS International Conference on Multimedia, Internet & Video Technologies (MIV '07)

7th WSEAS International Conference on Distance Learning and Web Engineering (DIWEB '07)

Beijing, China, September 15-17, 2007

Hosted and Sponsored by Beijing Jiaotong University

The Conference Gratefully Acknowledges the Support of K. C. WONG Education Foundation, Hong Kong

Improved Characteristic Waveform Decomposition and Novel Bit Reduction Scheme for WI Coders

KEUNSEOK CHO, HEESIK YANG, SANGBAE JEONG, MINSOO HAHN
School of Engineering
Information and Communications University
119, Munjiro, Yuseong-gu, Daejeon, 305-732
REPUBLIC OF KOREA
E-mail: {cks1223, sheik, sangbae, mshahn}@icu.ac.kr

Abstract: In this paper, we proposed an improved SEW/REW decomposition method with pitch-dependent phase generation and a noble VBR scheme for bit rate reduction in waveform interpolative coders. The proposed decomposition is performed in the magnitude domain to reduce spectral distortions. The phase of the characteristic waveforms is generated after classifying the signal into silence, unvoiced and voiced speech using the pitch value. The proposed VBR scheme is achieved by substituting white Gaussian noises with the excitation signal of silence and unvoiced speech and allocating bit rates variably. Experimental results show that our proposed algorithm achieves the improved speech quality while reducing the required bit rate compared to the conventional methods.

Key-Words: Characteristic waveform decomposition, Excitation signal, Speech coders, Variable bit rates, Waveform interpolation

1 Introduction

These days, the focus on the research of speech coders is to achieve high quality speech at low bit rates. The waveform interpolation (WI) coder has a hybrid structure that uses the filter model plus the excitation signal as a waveform. By the merit about achieving high quality speech at low bit rates, the WI coder can be used as a speech coder in digital wireless communication and can be used to construct low-capacity speech synthesizers. In the WI coder, the characteristic waveforms (CWs), i.e., the pitch length segments of the residual signal are used. By taking advantage of the differences in human perception between a slowly evolving waveform (SEW) and a rapidly evolving waveform (REW) extracted by the decomposition of the CW, the separate quantization of them is offered to get high coding efficiency [1]. The SEW characterizes voiced speech and the REW represents noise-like unvoiced speech [2]. Many researches on the speech quality improvement and the bit rate reduction have been executed. One of the past researches on the generation of the REW phase is based on the SEW/REW energy ratio for the improvement of the speech quality [3]. Source-controlled variable bit rate (SC-VBR) was proposed for the WI coders to reduce the required bit rates. The SC-VBR scheme is performed by the allocation of the variable bit rate according to the types of speech, that is, voiced, unvoiced, silence, and transition [4].

In conventional WI coders, the phase information is not transmitted to the decoder for the bit rate reduction. So, the SEW phase is generated by the fixed phase and the REW phase, by the random phase. The decomposition of the CW in the discrete time Fourier series (DTFS) domain cannot preserve the magnitude response of the excitation signal. The speech quality degradation can be caused by the non-transmission of the phase information and the variation of the magnitude components by the CW decomposition. In this paper, we propose a modified SEW/REW decomposition method with pitch-dependent phase generation and a noble VBR scheme for the speech quality improvement and the required bit reduction. To evaluate the performance, the perceptual evaluation of speech quality (PESQ) is used.

This paper is organized as follows: Section 2 describes the CW decomposition and phase generation method in conventional WI coders. Section 3 discusses the proposed SEW/REW decomposition scheme with pitch-dependent phase information generation and a novel VBR scheme. Section 4 and 5 describes the experimental results and the conclusion of this paper, respectively.