LETTER

An Enhanced Distortion Measure Based VBR for Waveform Interpolative Speech Coders

Heesik YANG†a, Student Member, Sangbae JEONG†, and Minsoo HAHN†, Nonmembers

SUMMARY In our previous study, a distortion measure based variable bit rate (DM-VBR) scheme in waveform interpolation (WI) coders was proposed. In this paper, the repetition method is proposed to estimate non-transmitted parameters instead of the extrapolation method. For the further reduction of slowly evolving waveform (SEW) bit rates, the dimensions of the past parameters, which are different from those of the current parameters, are converted to match the dimension of the current ones. Distortions between interpolated sub-frames and original sub-frames are measured for the reduction of the SEW parameters. And the usefulness of several other distortion measures is also investigated instead of the simple log spectral distortion. Experimental results show that the coder adopting the new schemes offers above 41% bit rate reduction with almost unnoticeable output speech degradation.

key words: waveform interpolation, VBR or multimode

1. Introduction

It has been reported that the waveform interpolation (WI) coder showed a relatively good speech quality at very low bit rates compared with CELP-based coders [1], [2]. Nevertheless, further bit-rate reduction without noticeable quality degradation is still necessary due to the bandwidth or memory limitation and customer’s demands for good speech quality at low cost. The speech quality and the bit rate in a WI coder directly rely on the quantization of the parameters. Thus, a rather direct approach to guarantee a good speech quality at low bit rates is to design an appropriate quantization scheme for each parameter. Another approach might be to use the VBR or multi-mode schemes as in other coders [3].

Several VBR schemes in the WI coder have been proposed [4], [5]. The VBR scheme in [4] is called the source controlled-VBR (SC-VBR). The SC-VBR classifies input speech segments into silence, voiced, unvoiced, and transition using a voice activity detector (VAD), the energy ratio between the rapidly evolving waveform (REW) and the SEW at each frame, and the cross-correlation between consecutive frames. The VBR scheme in [5] is called a distortion measure based variable bit rate (DM-VBR) scheme. The target of the DM-VBR is to remove the inter-frame redundancy at the parametric level. The DM-VBR is more robust to the input signal or the channel status than the SC-VBR [5]. Nevertheless, it has shown a relatively small bit rate reduction compared to the SC-VBR.

In this paper, we propose an enhanced DM-VBR in the WI scheme, which aims to acquire more bit rate reduction while maintaining the robustness of the DM-VBR.

In this paper, Sect. 2 describes the DM-VBR scheme and Sect. 3 concerns our proposed WI coder with an improved DM-VBR scheme. Experiments and the results are presented in Sect. 4 while conclusions are given in the final section.

2. Previous Works

The DM-VBR scheme for WI coders was proposed in [5]. The scheme measures the distortion \( D_P \) between the current parameters and the predicted parameters as

\[
D_P(n) = \frac{1}{L} \sum_{k=0}^{L-1} (p_k(n) - \hat{p}_k(n))^2,
\]

where \( p_k(n) \) is the \( k \)-th element of the \( n \)-th parameter vector \( P(n) \) and \( \hat{p}_k(n) \) is the predicted one. The dimension of the parameter vector \( P(n) \) is denoted as \( L \). The predicted parameter vector \( \hat{P}(n) \) is evaluated by the extrapolation using the past two encoded parameter vectors as in (2).

\[
\hat{P}(n) = 2P(n-1) - P(n-2)
\]

The measured distortion \( D_P \) is used as a criterion for the transmission of the current parameter vector. For the reconstruction of the parameter at the decoder, the flag information is transmitted to inform whether the current parameter vector is transmitted or not. At the decoder side, the non-transmitted parameter vectors are reconstructed by the extrapolation utilizing the past two parameter vectors as in (2). Despite of the quite successful performance of the DM-VBR scheme, it showed three undesirable problems. Firstly, the extrapolated parameters in the scheme may have unexpected values. For example, the extrapolation of LSF parameters using past two LSFs can violate the ordering property, which results in the divergence of the reconstructed signal [6]. Extrapolated parameters may contain negative values for powers both in the SEW and the REW. But this is already covered in [5] by performing the extrapolation in the log scale. The second problem is the error propagation. Prediction of a parameter can not avoid an error. If non-transmission of parameters continues, this prediction error...