Computers and Electrical Engineering 60 (2017) 116–125

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

A low computational complexity bandwidth extension method for mobile audio coding☆

Bo Hang*, Changqing Kang, Yi Wang

School of Mathematics and Computer Science, Hubei University of Arts and Science,Xiangyang, Hubei, 441053, China

ARTICLE INFO

Article history:
Received 7 August 2016
Revised 27 April 2017
Accepted 5 May 2017
Available online 11 May 2017

Keywords:
Mobile audio
Bandwidth extension
Computational complexity
Time-frequency transformation

ABSTRACT

Mobile devices always demand low computational complexity because computing resources, such as processor and battery, in mobile are always limited. The bandwidth extension (BWE) algorithm in mobile audio codec standard of China was proposed to improve audio quality in mobile communication. But the computational complexity of the algorithm is too high to implement in mobile devices. By analyzing the BWE algorithm, we discover that the main reason of high computational complexity is the frequently usage of time-frequency transformation. Then a low computational complexity scheme is proposed, which include algorithm optimization and code optimization. The experiment results show that computation time consumption ratio of BWE module in encoder and decoder are decreased by 4.5 and 14.3 percentage points respectively, without reducing the overall audio codec subjective quality, which is be conductive to the algorithm implement in mobile audio field.

1. Introduction

Traditional mobile audio coding methods in a mobile environment can not compress full band audio signal, due to the limited bandwidth resource and limited computing resources in mobile devices [1]. Since the semantic information of the audio signal is mainly concentrated in the low frequency portion, and therefore the traditional mobile audio encoder only encode the low-frequency audio portion of the signal, and high-frequency part is discarded. But the high-frequency signal helps improve intelligibility, naturalness of speech, enhance audio timbre, and also contains a wealth of spatial location information [2]. Therefore, in recent years, researchers have introduced a high-frequency bandwidth extension (BWE in mobile audio coding algorithm) technology to solve the problem of high-frequency signal lost [3].

Bandwidth extension technique can be divided into two categories: "blinded" bandwidth extension and "non-blind" bandwidth extension. Based on human auditory characteristics, the human ear is more sensitive to low-frequency signal than the high-frequency signal in the audible range, and the main information in a speech signal is contained in low frequency signals below 4000 Hz. Therefore, the sampling rates of early speech encoders are not more than 8000 Hz, in which narrowband speech signal are obtained. However, due to the high-frequency portion of the speech more than 4000 Hz play an important role for recognize unvoiced speech such as /f/ and /s/, and speech without high-frequency signal will sound bright enough. Therefore, a higher sampling rate to get broadband speech signal can significantly improve the intelligibility

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. J. Wen.

* Corresponding author.
E-mail address: bohang@163.com (B. Hang).

http://dx.doi.org/10.1016/j.compeleceng.2017.05.003

and naturalness of the speech. For compatible with existing narrowband speech signal, high-frequency bandwidth extension is applied to a narrowband speech signal to expand the narrowband signal to wideband speech signal. The methods predict high-frequency signal characteristics by low-frequency signal characteristics to rebuild the lost high-frequency signals, based on the statistical correlation between high-frequency and low-frequency signals. These methods, which do not require any side information from encoder, are called “blinded” the bandwidth extension methods [4].

Early codecs for audio signal, such as music, encoded the whole band signal, including low and high frequency signal. The encoding rate of whole band codec is too high for real time network communication. In order to reduce the coding rate of the audio signal, the introduction of bandwidth extension technique to reconstruct the high frequency portion of the audio signal can achieve low bit rate audio coding algorithm. However, since the statistical characteristics between the high frequency and low frequency of audio signal are not same as the speech signal, and high-frequency audio signals more complex, so blinded speech bandwidth extension algorithms can not be applied to all types of audio signals. So the audio bandwidth extension algorithm needs increase side information on a small number of high-frequency parameters in addition to the core coder, applying parameters coding method to reconstruct the high frequency portion of the audio signal. These methods are called “non-blinded” bandwidth extension methods [5].

In the research of mobile audio bandwidth extension algorithms, researchers have proposed a variety of algorithms. In 2002, in order to enhance the high frequency band audio signal parameters encoding quality, Liljeryd etc. proposed a Spectral Bandwidth Replication method (SBR) [6]. The low-frequency signal and high frequency signal are divided equal number of sub band signals and duplicate the low frequency sub-band, which has a higher signal correlation between high band and low band. And then the high-frequency characteristic parameter can be used to adjust the high frequency sub-band signals to synthesized high frequency signal. The method can reconstruct the high frequency portion of the audio signal, which has been accepted by MPEG (Moving Picture Experts Group) international audio standard. The shortage of the SBR method is that its decoding computational complexity is comparable with the AAC (Advanced Audio Coding) narrowband core decoder [7,8]. In 2004, 3GPP standard organization introduced audio codec standard AMR-WB + . The bandwidth extension method in AMR-WB + use linear prediction factors and time domain energy parameters to reconstruct high frequency signal [9], the rate as low as 0.8 kbps. In 2007 Bernd proposed a two stages, time domain and frequency domain, smoothing method as Time Domain Band-Width Extension method (TDBWE). Bernd believes that time domain and frequency domain energy play important roles for quality of bandwidth extension. When encoding, parameters of sub-frame energy in time domain and sub-band energy in frequency domain are computed. When decoding, the sub-frame energy and sub-band energy parameters and low frequency residual signal are used to reconstruct high-band signal. This simple, flexible bandwidth extension method has been adopted by G.729.1 audio standards [10]. However, due to not fully considered the correlation between low and high frequency, the bit rate of TDBWE is as high as 2 kbps, much higher than the AMR-WB + .

AVS P10 is a mobile audio codec standard developed independently by China. The algorithm used in a low-bit-rate bandwidth extension algorithm, by extracting high frequency linear prediction coefficient and sub-band energy factors for parameter coding and reconstruction of the high frequency signal [11]. The encoding quality of AVS P10 is comparable with AMR-WB + with same bit rate 0.8 kbps, which can meet the requirements of the mobile communication audio codec for mobile network bandwidth [12]. But the computational complexity of the algorithm is higher, which demand higher mobile device computing resources. In the limited computing power of mobile device cases, it is difficult to meet the needs of real-time codec [13].

To resolve the high computing complexity problem of AVS P10 bandwidth extension algorithm, we proposed a new low complexity method of high-frequency bandwidth extension, to reduce the complexity of encoding and decoding algorithms without compromising the quality of the high-frequency signal coding. The codec with the proposed new BWE method will meet the real-time requirements of mobile communication better.

This article describes the following four parts. The first part introduces the AVS P10 standard audio codec bandwidth extension algorithms, and analyzes the reasons of high computational complexity; the second part proposes the bandwidth extension algorithm complexity optimization method based on the analysis presented in the first part; the third part gives the performance testing result of optimized algorithm; and the final part is the conclusions.

2. Analysis of AVS P10 bandwidth extension algorithm computational complexity

We tested the computational complexity of AVS P10 encoder and decoder algorithms. According to test results, the computation time consumption of high frequency bandwidth extension algorithm accounted for 13.3% of the entire coding algorithm in encoder, and 36.8% in decoder. Thus, AVS P10 high frequency bandwidth extension algorithm although only increases a few bit rates, but increases much higher computational complexity. We will further analyze the reasons in following sections.

AVS P10 high-frequency bandwidth extension principle is to use the high frequency signal envelope information and gain information to adjust the low band excitation signal so as to generate a reconstructed high frequency signal. In encoder, side information, including spectral envelope and gain information of high frequency signals, were extracted in frequency domain. In decoder, high frequency spectral envelope information were used to adjust the low frequency excitation signals to rebuild the high frequency base signals, which were adjusted by gain factors to get high frequency signals in the frequency domain. At last, frequency domain reconstructed high frequency signals were converted to time domain to obtain time domain high-frequency signals. Encoding and decoding diagram are shown in Figs. 1 and 2 respectively.