Adaptive Bandwidth Extension and Apparatus for the Same

~~[1] This application claims priority to U.S. Patent Application No. 14/478,839 filed on September 5, 2014, entitled "Adaptive Bandwidth Extension and Apparatus for the Same", which is a continuation of U.S. Provisional Application No. 61/875,690 filed on September 10, 2013, entitled "Adaptive Selection of Shifting Band Based on Spectral Energy Level for Bandwidth Extension", both of which are incorporated herein by reference as if reproduced in its entirety.~~

TECHNICAL FIELD

[2] The present invention is generally in the field of speech processing, and in particular to adaptive band width extension and apparatus for the same.

BACKGROUND

[3] In modern audio/speech digital signal communication system, a digital signal is compressed at encoder; the compressed information (bitstream) can be packetized and sent to decoder through a communication channel frame by frame. The system of encoder and decoder together is called codec. Speech/audio compression may be used to reduce the number of bits that represent the speech/audio signal thereby reducing the bit rate needed for transmission. Speech/audio compression technology can be generally classified into time domain coding and frequency domain coding. Time domain coding is usually used for coding speech signal or for coding audio signal at low bit rates. Frequency domain coding is usually used for coding audio signal or for coding speech signal at high bit rates. Bandwidth Extension (BWE) can be a part of time domain coding or frequency domain coding in order to generate a high band signal at very low bit rate or at zero bit rate.

[67] Figure 4 illustrates a basic CELP decoder corresponding to the encoder in Figure 53. Figure 4 includes a post-processing block 408 receiving the synthesized speech 407 from the main decoder. This decoder is similar to Figure 3 except the adaptive codebook 307.

[68] For each subframe received, the received coded excitation index, quantized coded excitation gain index, quantized pitch index, quantized adaptive codebook gain index, and quantized short-term prediction parameter index, are used to find the corresponding parameters using corresponding decoders, for example, gain decoder 81, pitch decoder 84, adaptive codebook gain decoder 85, and short-term prediction decoder 83.

[69] In various embodiments, the CELP decoder is a combination of several blocks and comprises coded excitation 402, adaptive codebook 401, short-term prediction 406, and post-processing 408. Every block except post-processing has the same definition as described in the encoder of Figure 3. The post-processing may further include short-term post-processing and long-term post-processing.

[70] As already mentioned, CELP is mainly used to encode speech signal by benefiting from specific human voice characteristics or human vocal voice production model. In order to encode speech signal more efficiently, speech signal may be classified into different classes and each class is encoded in a different way. Voiced/Unvoiced classification or Unvoiced Decision may be an important and basic classification among all the classifications of different classes. For each class, LPC or STP filter is always used to represent the spectral envelope. But the excitation to the LPC filter may be different. Unvoiced signals may be coded with a noise-like excitation. On the other hand, voiced signals may be coded with a pulse-like excitation.

CLAIMS~~WHAT IS CLAIMED IS:~~

1. A method of decoding an encoded audio bitstream and generating frequency bandwidth extension at a decoder, the method comprising:

decoding the audio bitstream to produce a decoded low band audio signal and generate a low band excitation spectrum corresponding to a low frequency band;

selecting a sub-band area from within the low frequency band using a parameter which indicates energy information of a spectral envelope of the decoded low band audio signal;

generating a high band excitation spectrum for a high frequency band by copying a sub-band excitation spectrum from the selected sub-band area to a high sub-band area corresponding to the high frequency band;

using the generated high band excitation spectrum to generate an extended high band audio signal by applying a high band spectral envelope; and

adding the extended high band audio signal to the decoded low band audio signal to generate an audio output signal having an extended frequency bandwidth.

2. The method of claim 1, wherein selecting a sub-band area from within the low frequency band using the parameter which indicates energy information of the spectral envelope comprises identifying the highest quality sub band within the low frequency band by searching an highest energy point of the spectral envelope and selecting the identified highest quality sub band.

3. The method of claim 1, wherein selecting a sub-band area from within the low frequency band using the parameter which indicates energy information of the spectral envelope comprises selecting the sub-band area corresponding to highest spectral envelope energy.

4. The method of claim 1, wherein selecting a sub-band area from within the low frequency band using the parameter which indicates energy information of the spectral envelope comprises identifying a sub band from within the low band by using parameters reflecting an highest energy of the spectral ~~energy~~ envelope or spectral formant peak and selecting the identified sub band.

5. The method of any one of claims 1 to 4, wherein the method of decoding applies a bandwidth extension technology to generate the high frequency band.

6. The method of any one of claims 1 to 5, wherein applying the high band spectral envelope comprises applying a predicted high band filter representing the high band spectral envelope.

7. The method of any one of claims 1 to 6, further comprising:

generating the audio output signal by inverse transforming the frequency domain audio spectrum into time domain.

8. The method of any one of claims 1 to 7, wherein copying the sub-band excitation spectrum from the selected sub-band area to the high sub-band area corresponding to the high frequency band comprises copying low frequency band coefficients of an output from a filter bank analysis to the high sub-band area.

~~9. The method of any one of claims 1 to 8, wherein the audio bitstream comprises voiced speech or harmonic music.~~

~~10~~9. A decoder for decoding an encoded audio bitstream and generating frequency bandwidth, the decoder comprising:

a low band decoding unit configured to decode the audio bitstream to produce a decoded low band audio signal and to generate a low band excitation spectrum corresponding to a low frequency band; and

a band width extension unit coupled to the low band decoding unit and comprising a sub band selection unit and a copying unit, wherein the sub band selection unit is configured to select a sub-band area from within the low frequency band using a parameter which indicates energy information of a spectral envelope of the decoded low band audio signal, wherein the copying unit is configured to generate a high band excitation spectrum for a high frequency band by copying a sub-band excitation spectrum from the selected sub-band area to a high sub-band area corresponding to the high frequency band.

~~11~~10. The decoder of claim ~~10~~9, wherein selecting a sub-band area from within the low frequency band using energy information of the spectral envelope comprises identifying the highest quality sub band within the low frequency band.

~~12~~11. The decoder of claim ~~10~~9, wherein the sub band selection unit is configured to select the sub-band area corresponding to the highest spectral envelope energy.

~~13~~12. The decoder of claim ~~10~~9, wherein the sub band selection unit is configured to identify a sub band from within the low band by using parameters reflecting spectral ~~energy~~ envelope or spectral formant peak.

~~14~~13. The decoder of any one of claims ~~10~~9 to ~~13~~12, further comprising:

a high band signal generator coupled to the copying unit, the high band signal generator configured to apply a predicted high band spectral envelope to generate a high band time domain signal; and

an output generator coupled to the high band signal generator and the low band decoding unit, wherein the output generator is configured to generate an audio output signal by combining a low band time domain signal obtained by decoding the audio bitstream with the high band time domain signal.

~~15~~14. The decoder of claim ~~14~~13, wherein the high band signal generator is configured to apply a predicted high band filter representing the predicted high band spectral envelope.

~~16~~15. The decoder of any one of claims ~~10~~9 to ~~15~~14, further comprising:

a high band spectrum generator coupled to the copying unit, the high band spectrum~~signal~~ generator configured to apply an estimated high band spectral envelope to generate a high band spectrum for the high frequency band using the high band excitation spectrum; and

an output spectrum generator coupled to the high band spectrum generator and the low band decoding unit, wherein the output spectrum generator is configured to generate a frequency domain audio spectrum by combining a low band spectrum obtained by decoding the audio bitstream with the high band spectrum.

~~17~~16. The decoder of claim ~~16~~15, further comprising:

an inverse transform signal generator configured to generate a time domain audio signal by inverse transforming the frequency domain audio spectrum into time domain.

~~18~~17. A decoder for speech processing comprising:

a processor; and

a computer readable storage medium storing programming for execution by the processor, the programming including instructions to:

decode the audio bitstream to produce a decoded low band audio signal and generate a low band excitation spectrum corresponding to a low frequency band,

select a sub-band area from within the low frequency band using a parameter which indicates energy information of a spectral envelope of the decoded low band audio signal,

generate a high band excitation spectrum for a high frequency band by copying a sub-band excitation spectrum from the selected sub-band area to a high sub-band area corresponding to the high frequency band,

use the generated high band excitation spectrum to generate an extended high band audio signal by applying a high band spectral envelope, and

add the extended high band audio signal to the decoded low band audio signal to generate an audio output signal having an extended frequency bandwidth.

~~19. A method of decoding an encoded audio bitstream and generating frequency bandwidth extension at a decoder, the method comprising:~~

~~decoding the audio bitstream to produce a decoded low band audio signal and generate a low band spectrum corresponding to a low frequency band;~~

~~selecting a sub-band area from within the low frequency band using a parameter which indicates energy information of a spectral envelope of the decoded low band audio signal;~~

~~generating a high band spectrum by copying a sub-band spectrum from the selected sub-band area to a high sub-band area;~~

~~using the generated high band spectrum to generate an extended high band audio signal by applying a high band spectral envelope energy; and~~

~~adding the extended high band audio signal to the decoded low band audio signal to generate an audio output signal having an extended frequency bandwidth.~~