Improving Classification Between Time-Domain Coding and Frequency Domain Coding

~~[1] This application claims the benefit of U.S. Non Provisional Application Serial No. 14/511,943, filed October 10, 2014, entitled "Improving Classification Between Time-Domain Coding and Frequency Domain Coding", which claims the benefit of U.S. Provisional Application Serial No. 62/029,437, filed on July 26, 2014, entitled "Improving Classification Between Time-Domain Coding and Frequency Domain Coding for High Bit Rates", both of which are hereby incorporated herein by reference.~~

TECHNICAL FIELD

[2] The present invention is generally in the field of signal coding. In particular, the present invention is in the field of improving classification between time-domain coding and frequency domain coding.

BACKGROUND

[3] Speech coding refers to a process that reduces the bit rate of a speech file. Speech coding is an application of data compression of digital audio signals containing speech. Speech coding uses speech-specific parameter estimation using audio signal processing techniques to model the speech signal, combined with generic data compression algorithms to represent the resulting modeled parameters in a compact bitstream. The objective of speech coding is to achieve savings in the required memory storage space, transmission bandwidth and transmission power by reducing the number of bits per sample such that the decoded (decompressed) speech is perceptually indistinguishable from the original speech.

[4] However, speech coders are lossy coders, i.e., the decoded signal is different from the original. Therefore, one of the goals in speech coding is to minimize the distortion (or perceptible loss) at a given bit rate, or minimize the bit rate to reach a given distortion.

[5] Speech coding differs from other forms of audio coding in that speech is a much simpler signal than most other audio signals, and a lot more statistical information is available about the properties of speech. As a result, some auditory information which is relevant in audio

CLAIMS

1. A method for processing speech signals prior to encoding a digital signal comprising audio data, the method comprising:

selecting frequency domain coding or time domain coding based on

a coding bit rate to be used for coding the digital signal and

a short pitch lag detection of the digital signal.

2. The method of claim 1, wherein the short pitch lag detection comprises detecting whether the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, wherein the pitch lag limit is a minimum allowable pitch for a Code Excited Linear Prediction Technique (CELP) algorithm for coding the digital signal.

3. The method of claim 1, wherein the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, and wherein selecting frequency domain coding or time domain coding comprises:

selecting frequency domain coding for coding the digital signal when a coding bit rate is higher than an upper bit rate limit, and

selecting time domain coding for coding the digital signal when the coding bit rate is lower than a lower bit rate limit.

4. The method of claim 3, wherein the coding bit rate is higher than the upper bit rate limit when the coding bit rate is greater than or equal to 46200 bps, and wherein the coding bit rate is lower than a lower bit rate limit when the coding bit rate is less than 24.4 kbps.

5. The method of claim 1, wherein the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, and wherein selecting frequency domain coding or time domain coding comprises:

selecting frequency domain coding for coding the digital signal when coding bit rate is intermediate between a lower bit rate limit and an upper bit rate limit, and wherein a voicing periodicity is low.

6. The method of claim 1, wherein the digital signal does not comprise a short pitch signal for which the pitch lag is shorter than a pitch lag limit, and wherein selecting frequency domain coding or time domain coding comprises:

selecting time domain coding for coding the digital signal when the digital signal is classified as unvoiced speech or normal speech.

7. The method of claim 1, wherein the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, and wherein selecting frequency domain coding or time domain coding comprises:

selecting time domain coding for coding the digital signal when coding bit rate is intermediate between a lower bit rate limit and an upper bit rate limit and a voicing periodicity is very strong.

8. The method of claim 1, further comprising coding the digital signal using the selected frequency domain coding or the selected time domain coding.

9. The method of claim 1, wherein selecting frequency domain coding or time domain coding based on the pitch lag of the digital signal comprises detecting for short pitch signal based on determining a parameter for detecting lack of very low frequency energy or a parameter for spectral sharpness.

10. A method for processing speech signals prior to encoding a digital signal comprising audio data, the method comprising:

selecting frequency domain coding for coding the digital signal when a coding bit rate is higher than an upper bit rate limit; and

selecting time domain coding for coding the digital signal when the coding bit rate is lower than a lower bit rate limit, wherein the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit.

11. The method of claim 10, wherein the coding bit rate is higher than the upper bit rate limit when the coding bit rate is greater than or equal to 46200 bps, and wherein the coding bit rate is lower than a lower bit rate limit when the coding bit rate is less than 24.4 kbps.

12. The method of claim 10, further comprising coding the digital signal using the selected frequency domain coding or the selected time domain coding.

13. An apparatus for processing speech signals prior to encoding a digital signal comprising audio data, the apparatus comprising a coding selector configured to select frequency domain coding or time domain coding based on a coding bit rate to be used for coding the digital signal and a short pitch lag detection of the digital signal.

14. The apparatus of claim 13, wherein when the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, the coding selector is configured to

select frequency domain coding for coding the digital signal when a coding bit rate is higher than an upper bit rate limit, and

select time domain coding for coding the digital signal when the coding bit rate is lower than a lower bit rate limit.

15. The apparatus of claim 13, wherein when the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, the coding selector is configured to

select frequency domain coding for coding the digital signal when coding bit rate is intermediate between a lower bit rate limit and an upper bit rate limit, and wherein a voicing periodicity is low.

16. The apparatus of claim 13, wherein when the digital signal does not comprise a short pitch signal for which the pitch lag is shorter than a pitch lag limit, the coding selector is configured to

select time domain coding for coding the digital signal when the digital signal is classified as unvoiced speech or normal speech.

17. The apparatus of claim 13, wherein when the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, the coding selector is configured to

select time domain coding for coding the digital signal when coding bit rate is intermediate between a lower bit rate limit and an upper bit rate limit and a voicing periodicity is very strong.

18. The apparatus of claim 13, wherein the apparatus further comprising a coding unit which is configured to code the digital signal using the frequency domain coding selected by the selector or the time domain coding selected by the selector.

19. A method for processing speech signals prior to encoding, the method comprising:

selecting time domain coding for coding a digital signal comprising audio data when the digital signal does not comprise short pitch signal and the digital signal is classified as unvoiced speech or normal speech;

selecting frequency domain coding for coding the digital signal when coding bit rate is intermediate between a lower bit rate limit and an upper bit rate limit, and the digital signal comprises short pitch signal and voicing periodicity is low; and

selecting time domain coding for coding the digital signal when coding bit rate is intermediate and the digital signal comprises short pitch signal and a voicing periodicity is very strong.

20. The method of claim 19, further comprising coding the digital signal using the selected frequency domain coding or the selected time domain coding.

CLAIMS

1. A method for processing speech signals prior to encoding a digital signal comprising audio data, the method comprising:

selecting frequency domain coding or time domain coding based on

a coding bit rate to be used for coding the digital signal and

a short pitch lag detection of the digital signal; wherein the short pitch lag detection comprises detecting whether the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, wherein the pitch lag limit is a minimum allowable pitch for a Code Excited Linear Prediction (CELP) algorithm for coding the digital signal.

2. The method of claim 1, wherein the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, and wherein selecting frequency domain coding or time domain coding comprises:

selecting time domain coding for coding the digital signal when the coding bit rate is lower than a lower bit rate limit.

3. The method of claim 2, wherein the coding bit rate is lower than a lower bit rate limit when the coding bit rate is less than 24.4 kbps.

4. The method of claim 1, wherein the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, and wherein selecting frequency domain coding or time domain coding comprises:

selecting frequency domain coding for coding the digital signal when a coding bit rate is higher than an upper bit rate limit, and

5. The method of claim 4, wherein the coding bit rate is higher than the upper bit rate limit when the coding bit rate is greater than or equal to 46200 bps.

6. The method of claim 1, wherein the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, and wherein selecting frequency domain coding or time domain coding comprises:

selecting frequency domain coding for coding the digital signal when coding bit rate is intermediate between a lower bit rate limit and an upper bit rate limit, and wherein a voicing periodicity is low.

7. The method of claim 1, wherein the digital signal does not comprise a short pitch signal for which the pitch lag is shorter than a pitch lag limit, and wherein selecting frequency domain coding or time domain coding comprises:

selecting time domain coding for coding the digital signal when the digital signal is classified as unvoiced speech or normal speech.

8. The method of claim 1, wherein the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, and wherein selecting frequency domain coding or time domain coding comprises:

selecting time domain coding for coding the digital signal when coding bit rate is intermediate between a lower bit rate limit and an upper bit rate limit and a voicing periodicity is very strong.

9. The method of claim 1, further comprising coding the digital signal using the selected frequency domain coding or the selected time domain coding.

10. The method of claim 1, wherein selecting frequency domain coding or time domain coding based on the pitch lag of the digital signal comprises detecting for short pitch signal based on determining a parameter for detecting lack of very low frequency energy or a parameter for spectral sharpness.

11. An apparatus for processing speech signals prior to encoding a digital signal comprising audio data, the apparatus comprising a coding selector configured to select frequency domain coding or time domain coding based on a coding bit rate to be used for coding the digital

signal and a short pitch lag detection of the digital signal, wherein the short pitch lag detection comprises detecting whether the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, wherein the pitch lag limit is a minimum allowable pitch for a Code Excited Linear Prediction (CELP) algorithm for coding the digital signal.

12. The apparatus of claim 11, wherein when the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, the coding selector is configured to

select time domain coding for coding the digital signal when the coding bit rate is lower than a lower bit rate limit.

13. The apparatus of claim 12, wherein the coding bit rate is lower than a lower bit rate limit when the coding bit rate is less than 24.4 kbps.

14. The apparatus of claim 11, wherein when the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, the coding selector is configured to

select frequency domain coding for coding the digital signal when a coding bit rate is higher than an upper bit rate limit.

15. The apparatus of claim 14, wherein the coding bit rate is higher than the upper bit rate limit when the coding bit rate is greater than or equal to 46200 bps.

16. The apparatus of claim 11, wherein when the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, the coding selector is configured to

select frequency domain coding for coding the digital signal when coding bit rate is intermediate between a lower bit rate limit and an upper bit rate limit, and wherein a voicing periodicity is low.

17. The apparatus of claim 11, wherein when the digital signal does not comprise a short pitch signal for which the pitch lag is shorter than a pitch lag limit, the coding selector is configured to

select time domain coding for coding the digital signal when the digital signal is classified as unvoiced speech or normal speech.

18. The apparatus of claim 11, wherein when the digital signal comprises a short pitch signal for which the pitch lag is shorter than a pitch lag limit, the coding selector is configured to

select time domain coding for coding the digital signal when coding bit rate is intermediate between a lower bit rate limit and an upper bit rate limit and a voicing periodicity is very strong.

19. The apparatus of claim 11, wherein the apparatus further comprising a coding unit which is configured to code the digital signal using the frequency domain coding selected by the selector or the time domain coding selected by the selector.