AUDIO ENCODING METHOD AND APPARATUS

~~[0001] This application claims priority to Chinese Patent Application No. 201410288983.3, filed with the Chinese Patent Office on June 24, 2014 and entitled "AUDIO ENCODING METHOD AND APPARATUS", which is incorporated herein by reference in its entirety.~~

TECHNICAL FIELD

[0002] Embodiments of the present invention relate to the field of signal processing technologies, and more specifically, to an audio encoding method and an apparatus.

BACKGROUND

[0003] In the prior art, a hybrid encoder is usually used to encode an audio signal in a voice communications system. Specifically, the hybrid encoder usually includes two sub encoders. One sub encoder is suitable to encoding a speech signal, and the other encoder is suitable to encoding a non-speech signal. For a received audio signal, each sub encoder of the hybrid encoder encodes the audio signal. The hybrid encoder directly compares quality of encoded audio signals to select an optimum sub encoder. However, such a closed-loop encoding method has high operation complexity.

SUMMARY

[0004] Embodiments of the present invention provide an audio encoding method and an apparatus, which can reduce encoding complexity and ensure that encoding is of relatively high accuracy.

[0005] According to a first aspect, an audio encoding method is provided, where the method includes: determining sparseness of distribution, on spectrums, of energy of N input audio frames, where the N audio frames include a current audio frame, and N is a positive integer; and determining, according to the sparseness of distribution, on the spectrums, of the energy of the N audio frames, whether to use a first encoding method or a second encoding method to encode the current audio frame, where the first encoding method is an encoding method that is based on

CLAIMS

~~What is claimed is:~~

1. An audio encoding method, wherein the method comprises:

determining sparseness of distribution, on spectrums, of energy of N input audio frames, wherein the N audio frames comprise a current audio frame, and N is a positive integer; and

determining, according to the sparseness of distribution, on the spectrums, of the energy of the N audio frames, whether to use a first encoding method or a second encoding method to encode the current audio frame, wherein the first encoding method is an encoding method that is based on time-frequency transform and transform coefficient quantization and that is not based on linear prediction, and the second encoding method is a linear-predication-based encoding method.

2. The method according to claim 1, wherein the determining sparseness of distribution, on spectrums, of energy of N input audio frames comprises:

dividing a spectrum of each of the N audio frames into P spectral envelopes, wherein P is a positive integer; and

determining a general sparseness parameter according to energy of the P spectral envelopes of each of the N audio frames, wherein the general sparseness parameter indicates the sparseness of distribution, on the spectrums, of the energy of the N audio frames.

3. The method according to claim 2, wherein the general sparseness parameter comprises a first minimum bandwidth;

the determining a general sparseness parameter according to energy of the P spectral envelopes of each of the N audio frames comprises:

determining an average value of minimum bandwidths, distributed on the spectrums, of first-preset-proportion energy of the N audio frames according to the energy of the P spectral envelopes of each of the N audio frames, wherein the average value of the minimum bandwidths, distributed on the spectrums, of the first-preset-proportion energy of the N audio frames is the first minimum bandwidth; and

the determining, according to the sparseness of distribution, on the spectrums, of the energy of the N audio frames, whether to use a first encoding method or a second encoding method to encode the current audio frame comprises:

when the first minimum bandwidth is less than a first preset value, determining to use the first encoding method to encode the current audio frame; or when the first minimum bandwidth is greater than the first preset value, determining to use the second encoding method to encode the current audio frame.

11. The method according to claim 1, wherein the sparseness of distribution of the energy on the spectrums comprises global sparseness, local sparseness, and short-time burstiness of distribution of the energy on the spectrums.

12. The method according to claim 11, wherein N is 1, and the N audio frames are the current audio frame; and

the determining sparseness of distribution, on spectrums, of energy of N input audio frames comprises:

dividing a spectrum of the current audio frame into Q sub bands; and

determining a burst sparseness parameter according to peak energy of each of the Q sub bands of the spectrum of the current audio frame, wherein the burst sparseness parameter is used to indicate global sparseness, local sparseness, and short-time burstiness of the current audio frame.

13. The method according to claim 12, wherein the burst sparseness parameter comprises: a global peak-to-average proportion of each of the Q sub bands, a local peak-to-average proportion of each of the Q sub bands, and a short-time peak energy fluctuation of each of the Q sub bands, wherein the global peak-to-average proportion is determined according to the peak energy in the sub band and average energy of all the sub bands of the current audio frame, the local peak-to-average proportion is determined according to the peak energy in the sub band and average energy in the sub band, and the short-time peak energy fluctuation is determined according to the peak energy in the sub band and peak energy in a specific frequency band of an audio frame before the audio frame; and

the determining, according to the sparseness of distribution, on the spectrums, of the energy of the N audio frames, whether to use a first encoding method or a second encoding method to encode the current audio frame comprises:

determining whether there is a first sub band in the Q sub bands, wherein a local peak-to-average proportion of the first sub band is greater than an eleventh preset value, a global peak-to-average proportion of the first sub band is greater than a twelfth preset value, and a short-time peak energy fluctuation of the first sub band is greater than a thirteenth preset value; and

when there is the first sub band in the Q sub bands, determining to use the first encoding method to encode the current audio frame.

14. The method according to claim 1, wherein the sparseness of distribution of the energy on the spectrums comprises band-limited characteristics of distribution of the energy on the spectrums.

15. The method according to claim 14, wherein the determining sparseness of distribution, on spectrums, of energy of N input audio frames comprises:

determining a demarcation frequency of each of the N audio frames; and

N audio frames, and determine the third energy proportion according to energy of the P_3 spectral envelopes of each of the N audio frames and the total energy of the respective N audio frames, wherein P_2 and P_3 are positive integers less than P, and P_2 is less than P_3; and

the determining unit is specifically configured to: when the second energy proportion is greater than a seventh preset value and the third energy proportion is greater than an eighth preset value, determine to use the first encoding method to encode the current audio frame; when the second energy proportion is greater than a ninth preset value, determine to use the first encoding method to encode the current audio frame; and when the third energy proportion is less than a tenth preset value, determine to use the second encoding method to encode the current audio frame.

26. The apparatus according to claim 25, wherein the determining unit is specifically configured to determine, from the P spectral envelopes of each of the N audio frames, P_2 spectral envelopes having maximum energy, and determine, from the P spectral envelopes of each of the N audio frames, P_3 spectral envelopes having maximum energy.

27. The apparatus according to claim 17, wherein N is 1, and the N audio frames are the current audio frame; and

the determining unit is specifically configured to divide a spectrum of the current audio frame into Q sub bands, and determine a burst sparseness parameter according to peak energy of each of the Q sub bands of the spectrum of the current audio frame, wherein the burst sparseness parameter is used to indicate global sparseness, local sparseness, and short-time burstiness of the current audio frame.

28. The apparatus according to claim 27, wherein the determining unit is specifically configured to determine a global peak-to-average proportion of each of the Q sub bands, a local peak-to-average proportion of each of the Q sub bands, and a short-time peak energy fluctuation of each of the Q sub bands, wherein the global peak-to-average proportion is determined by the determining unit according to the peak energy in the sub band and average energy of all the sub bands of the current audio frame, the local peak-to-average proportion is determined by the determining unit according to the peak energy in the sub band and average energy in the sub band, and the short-time peak energy fluctuation is determined according to the peak energy in the sub band and peak energy in a specific frequency band of an audio frame before the audio frame; and

the determining unit is specifically configured to: determine whether there is a first sub band in the Q sub bands, wherein a local peak-to-average proportion of the first sub band is greater than an eleventh preset value, a global peak-to-average proportion of the first sub band is greater than a twelfth preset value, and a short-time peak energy fluctuation of the first sub band is greater than a