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.

[0006] With reference to the first aspect, in a first possible implementation manner of the first aspect, the determining sparseness of distribution, on spectrums, of energy of N input audio frames includes: dividing a spectrum of each of the N audio frames into P spectral envelopes, where 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, where the general sparseness parameter indicates the sparseness of distribution, on the spectrums, of the energy of the N audio frames.

[0007] With reference to the first possible implementation manner of the first aspect, in a second possible implementation manner of the first aspect, the general sparseness parameter includes 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 includes: determining an average value of minimum bandwidths of distribution, 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, where the average value of the minimum bandwidths of distribution, 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 includes: 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.

[0008] With reference to the second possible implementation manner of the first aspect, in a third possible implementation manner of the first aspect, the determining an average value of minimum bandwidths of distribution, 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 includes: sorting the energy of the P spectral envelopes of each audio frame in descending order; determining, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the first preset proportion of each of the N audio frames; and determining, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the first preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the first preset proportion of the N audio frames.

[0009] With reference to the first possible implementation manner of the first aspect, in a fourth possible implementation manner of the first aspect, the general sparseness parameter includes a first energy proportion; the determining a general sparseness parameter according to energy of the P spectral envelopes of each of the N audio frames includes: selecting P_1 spectral envelopes from the P spectral envelopes of each of the N audio frames; and determining the first energy proportion according to energy of the P_1 spectral envelopes of each of the N audio frames and total energy of the respective N audio frames, where P_1 is a positive integer less than P; 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 includes: when the first energy proportion is greater than a second preset value, determining to use the first encoding method to encode the current audio frame; or when the first energy proportion is less than the second preset value, determining to use the second encoding method to encode the current audio frame.

[0010] With reference to the fourth possible implementation manner of the first aspect, in a fifth possible implementation manner of the first aspect, energy of any one of the P_1 spectral envelopes is greater than energy of any one of the other spectral envelopes in the P spectral envelopes except the P_1 spectral envelopes.

[0011] With reference to the first possible implementation manner of the first aspect, in a sixth possible implementation manner of the first aspect, the general sparseness parameter includes a second minimum bandwidth and a third minimum bandwidth; the determining a general sparseness parameter according to energy of the P spectral envelopes of each of the N audio frames includes: determining an average value of minimum bandwidths of distribution, on the spectrums, of second-preset-proportion energy of the N audio frames and determining an average value of minimum bandwidths of distribution, on the spectrums, of third-preset-proportion energy of the N audio frames according to the energy of the P spectral envelopes of each of the N audio frames, where the average value of the minimum bandwidths of distribution, on the spectrums, of the second-preset-proportion energy of the N audio frames is used as the second minimum bandwidth, the average value of the minimum bandwidths of distribution, on the spectrums, of the third-preset-proportion energy of the N audio frames is used as the third minimum bandwidth, and the second preset proportion is less than the third preset proportion; 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 includes: when the second minimum bandwidth is less than a third preset value and the third minimum bandwidth is less than a fourth preset value, determining to use the first encoding method to encode the current audio frame; when the third minimum bandwidth is less than a fifth

preset value, determining to use the first encoding method to encode the current audio frame; or when the third minimum bandwidth is greater than a sixth preset value, determining to use the second encoding method to encode the current audio frame, where the fourth preset value is greater than or equal to the third preset value, the fifth preset value is less than the fourth preset value, and the sixth preset value is greater than the fourth preset value.

[0012] With reference to the sixth possible implementation manner of the first aspect, in a seventh possible implementation manner of the first aspect, the determining an average value of minimum bandwidths of distribution, on the spectrums, of second-preset-proportion energy of the N audio frames and determining an average value of minimum bandwidths of distribution, on the spectrums, of third-preset-proportion energy of the N audio frames according to the energy of the P spectral envelopes of each of the N audio frames includes: sorting the energy of the P spectral envelopes of each audio frame in descending order; determining, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the second preset proportion of each of the N audio frames; determining, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the second preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the second preset proportion of the N audio frames; determining, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the third preset proportion of each of the N audio frames; and determining, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the third preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the third preset proportion of the N audio frames.

[0013] With reference to the first possible implementation manner of the first aspect, in an eighth possible implementation manner of the first aspect, the general sparseness parameter includes a second energy proportion and a third energy proportion; the determining a general sparseness parameter according to energy of the P spectral envelopes of each of the N audio frames includes: selecting P_2 spectral envelopes from the P spectral envelopes of each of the N audio frames; determining the second energy proportion according to energy of the P_2 spectral envelopes of each of the N audio frames and total energy of the respective N audio frames; selecting P_3 spectral envelopes from the P spectral envelopes of each of the N audio frames; and determining the

[0022] With reference to the second aspect, in a first possible implementation manner of the second aspect, the determining unit is specifically configured to divide a spectrum of each of the N audio frames into P spectral envelopes, and determine a general sparseness parameter according to energy of the P spectral envelopes of each of the N audio frames, where P is a positive integer, and the general sparseness parameter indicates the sparseness of distribution, on the spectrums, of the energy of the N audio frames.

[0023] With reference to the first possible implementation manner of the second aspect, in a second possible implementation manner of the second aspect, the general sparseness parameter includes a first minimum bandwidth; the determining unit is specifically configured to determine an average value of minimum bandwidths of distribution, 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, where the average value of the minimum bandwidths of distribution, on the spectrums, of the first preset proportion energy of the N audio frames is the first minimum bandwidth; and the determining unit is specifically configured to: when the first minimum bandwidth is less than a first preset value, determine to use the first encoding method to encode the current audio frame; and when the first minimum bandwidth is greater than the first preset value, determine to use the second encoding method to encode the current audio frame.

[0024] With reference to the second possible implementation manner of the second aspect, in a third possible implementation manner of the second aspect, the determining unit is specifically configured to: sort the energy of the P spectral envelopes of each audio frame in descending order; determine, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the first preset proportion of each of the N audio frames; and determine, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the first preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the first preset proportion of the N audio frames.

[0025] With reference to the first possible implementation manner of the second aspect, in a fourth possible implementation manner of the second aspect, the general sparseness parameter includes a first energy proportion; the determining unit is specifically configured to select P_1 spectral envelopes from the P spectral envelopes of each of the N audio frames, and determine the first energy proportion according to energy of the P_1 spectral envelopes of each of the N audio frames and total energy of the respective N audio frames, where P_1 is a positive integer less than P; and the determining unit is specifically configured to: when the first energy proportion is greater than a second preset value, determine to use the first encoding method to encode the current audio

frame; and when the first energy proportion is less than the second preset value, determine to use the second encoding method to encode the current audio frame.

[0026] With reference to the fourth possible implementation manner of the second aspect, in a fifth possible implementation manner of the second aspect, the determining unit is specifically configured to determine the P_1 spectral envelopes according to the energy of the P spectral envelopes, where energy of any one of the P_1 spectral envelopes is greater than energy of any one of the other spectral envelopes in the P spectral envelopes except the P_1 spectral envelopes.

[0027] With reference to the first possible implementation manner of the second aspect, in a sixth possible implementation manner of the second aspect, the general sparseness parameter includes a second minimum bandwidth and a third minimum bandwidth; the determining unit is specifically configured to determine an average value of minimum bandwidths of distribution, on the spectrums, of second-preset-proportion energy of the N audio frames and determine an average value of minimum bandwidths of distribution, on the spectrums, of third-preset-proportion energy of the N audio frames according to the energy of the P spectral envelopes of each of the N audio frames, where the average value of the minimum bandwidths of distribution, on the spectrums, of the second-preset-proportion energy of the N audio frames is used as the second minimum bandwidth, the average value of the minimum bandwidths of distribution, on the spectrums, of the third-preset-proportion energy of the N audio frames is used as the third minimum bandwidth, and the second preset proportion is less than the third preset proportion; and the determining unit is specifically configured to: when the second minimum bandwidth is less than a third preset value and the third minimum bandwidth is less than a fourth preset value, determine to use the first encoding method to encode the current audio frame; when the third minimum bandwidth is less than a fifth preset value, determine to use the first encoding method to encode the current audio frame; and when the third minimum bandwidth is greater than a sixth preset value, determine to use the second encoding method to encode the current audio frame, where the fourth preset value is greater than or equal to the third preset value, the fifth preset value is less than the fourth preset value, and the sixth preset value is greater than the fourth preset value.

[0028] With reference to the sixth possible implementation manner of the second aspect, in a seventh possible implementation manner of the second aspect, the determining unit is specifically configured to: sort the energy of the P spectral envelopes of each audio frame in descending order; determine, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the second preset proportion of each of the N audio frames; determine, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts

for not less than the second preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the second preset proportion of the N audio frames; determine, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the third preset proportion of each of the N audio frames; and determine, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the third preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the third preset proportion of the N audio frames.

[0029] With reference to the first possible implementation manner of the second aspect, in an eighth possible implementation manner of the second aspect, the general sparseness parameter includes a second energy proportion and a third energy proportion; the determining unit is specifically configured to: select P_2 spectral envelopes from the P spectral envelopes of each of the N audio frames, determine the second energy proportion according to energy of the P_2 spectral envelopes of each of the N audio frames and total energy of the respective N audio frames, select P_3 spectral envelopes from the P spectral envelopes of each of the 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, where 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.

[0030] With reference to the eighth possible implementation manner of the second aspect, in a ninth possible implementation manner of the second aspect, 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.

[0031] With reference to the second aspect, in a tenth possible implementation manner of the second aspect, 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

[0045] During selection of an appropriate encoding method for an audio frame, sparseness of distribution, on a spectrum, of energy of the audio frame may be considered. There may be three types of sparseness of distribution, on a spectrum, of energy of an audio frame: general sparseness, burst sparseness, and band-limited sparseness.

[0046] Optionally, in an embodiment, an appropriate encoding method may be selected for the current audio frame by using the general sparseness. In this case, the determining sparseness of distribution, on spectrums, of energy of N input audio frames includes: dividing a spectrum of each of the N audio frames into P spectral envelopes, where 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, where the general sparseness parameter indicates the sparseness of distribution, on the spectrums, of the energy of the N audio frames.

[0047] Specifically, an average value of minimum bandwidths of distribution, on spectrums, of specific-proportion energy of N input consecutive audio frames may be defined as the general sparseness. A smaller bandwidth indicates stronger general sparseness, and a larger bandwidth indicates weaker general sparseness. In other words, stronger general sparseness indicates that energy of an audio frame is more centralized, and weaker general sparseness indicates that energy of an audio frame is more disperse. Efficiency is high when the first encoding method is used to encode an audio frame whose general sparseness is relatively strong. Therefore, an appropriate encoding method may be selected by determining general sparseness of an audio frame, to encode the audio frame. To help determine general sparseness of an audio frame, the general sparseness may be quantized to obtain a general sparseness parameter. Optionally, when N is 1, the general sparseness is a minimum bandwidth of distribution, on a spectrum, of specific-proportion energy of the current audio frame.

[0048] Optionally, in an embodiment, the general sparseness parameter includes a first minimum bandwidth. In this case, the determining a general sparseness parameter according to energy of the P spectral envelopes of each of the N audio frames includes: determining an average value of minimum bandwidths of distribution, 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, where the average value of the minimum bandwidths of distribution, on the spectrums, of the first preset proportion energy of the N audio frames is the first minimum bandwidth. 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 includes: 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. Optionally, in an embodiment, when N is 1, the N audio frames are the current audio frame, and the average value of the minimum bandwidths of distribution, on the spectrums, of the first preset proportion energy of the N audio frames is a minimum bandwidth of distribution, on the spectrum, of first preset proportion energy of the current audio frame.

[0049] A person skilled in the art may understand that, the first preset value and the first preset proportion may be determined according to a simulation experiment. An appropriate first preset value and first preset proportion may be determined by means of a simulation experiment, so that a good encoding effect can be obtained when an audio frame meeting the foregoing condition is encoded by using the first encoding method or the second encoding method. Generally, a value of the first preset proportion is generally a number between 0 and 1 and relatively close to 1, for example, 90% or 80%. The selection of the first preset value is related to the value of the first preset proportion, and also related to a selection tendency between the first encoding method and the second encoding method. For example, a first preset value corresponding to a relatively large first preset proportion is generally greater than a first preset value corresponding to a relatively small first preset proportion. For another example, a first preset value corresponding to a tendency to select the first encoding method is generally greater than a first preset value corresponding to a tendency to select the second encoding method.

[0050] The determining an average value of minimum bandwidths of distribution, 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 includes: sorting the energy of the P spectral envelopes of each audio frame in descending order; determining, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the first preset proportion of each of the N audio frames; and determining, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the first preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the first preset proportion of the N audio frames. For example, an input audio signal is a wideband signal sampled at 16 kHz, and the input signal is input in a frame of 20 ms. Each frame of signal is 320 time domain sampling points. Time-frequency transform is performed on a time domain signal. For example, time-frequency transform is performed by means of fast Fourier transform (Fast Fourier Transformation, FFT), to obtain 160 spectral envelopes S(k), that is, 160 FFT energy spectrum coefficients, where k=0, 1, 2, ..., 159. A minimum bandwidth is found from the spectral envelopes S(k) in a manner that a proportion that energy on the bandwidth accounts for in total energy of the

frame is the first preset proportion. Specifically, determining a minimum bandwidth of distribution, on a spectrum, of first preset proportion energy of an audio frame according to energy, sorted in descending order, of P spectral envelopes of the audio frame includes: sequentially accumulating energy of frequency bins in the spectral envelopes S(k) in descending order; and comparing energy obtained after each time of accumulation with the total energy of the audio frame, and if a proportion is greater than the first preset proportion, ending the accumulation process, where a quantity of times of accumulation is the minimum bandwidth. For example, the first preset proportion is 90%, and if a proportion that an energy sum obtained after 30 times of accumulation accounts for in the total energy exceeds 90%, a proportion that an energy sum obtained after 29 times of accumulation accounts for in the total energy is less than 90%, and a proportion that an energy sum obtained after 31 times of accumulation accounts for in the total energy exceeds the proportion that the energy sum obtained after 30 times of accumulation accounts for in the total energy, it may be considered that a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the first preset proportion of the audio frame is 30. The foregoing minimum bandwidth determining process is executed for each of the N audio frames, to separately determine the minimum bandwidths of distribution, on the spectrums, of the energy that accounts for not less than the first preset proportion of the N audio frames including the current audio frame, and calculate the average value of the N minimum bandwidths. The average value of the N minimum bandwidths may be referred to as the first minimum bandwidth, and the first minimum bandwidth may be used as the general sparseness parameter. When the first minimum bandwidth is less than the first preset value, it is determined to use the first encoding method to encode the current audio frame. When the first minimum bandwidth is greater than the first preset value, it is determined to use the second encoding method to encode the current audio frame.

[0051] Optionally, in another embodiment, the general sparseness parameter may include a first energy proportion. In this case, the determining a general sparseness parameter according to energy of the P spectral envelopes of each of the N audio frames includes: selecting P_1 spectral envelopes from the P spectral envelopes of each of the N audio frames; and determining the first energy proportion according to energy of the P_1 spectral envelopes of each of the N audio frames and total energy of the respective N audio frames, where P_1 is a positive integer less than P. 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 includes: when the first energy proportion is greater than a second preset value, determining to use the first encoding method to encode the current audio frame; or when the first energy proportion is less than the second preset value, determining to use the second encoding method to

Time-frequency transform is performed on a time domain signal. For example, time-frequency transform is performed by means of fast Fourier transform, to obtain 160 spectral envelopes S(k), where k=0, 1, 2, ..., 159. P_1 spectral envelopes are sclected from the 160 spectral envelopes, and a proportion that an energy sum of the P_1 spectral envelopes accounts for in total energy of the audio frame is calculated. The foregoing process is executed for each of the N audio frames. That is, a proportion that an energy sum of the P_1 spectral envelopes of each of the N audio frames accounts for in respective total energy is calculated. An average value of the proportions is calculated. The average value of the proportions is the first energy proportion. When the first energy proportion is greater than the second preset value, it is determined to use the first encoding method to encode the current audio frame. When the first energy proportion is less than the second preset value, it is determined to use the second encoding method to encode the current audio frame. Energy of any one of the P_1 spectral envelopes is greater than energy of any one of the other spectral envelopes in the P spectral envelopes except the P_1 spectral envelopes. Optionally, in an embodiment, the value of P_1 may be 20.

[0055] Optionally, in another embodiment, the general sparseness parameter may include a second minimum bandwidth and a third minimum bandwidth. In this case, the determining a general sparseness parameter according to energy of the P spectral envelopes of each of the N audio frames includes: determining an average value of minimum bandwidths of distribution, on the spectrums, of second-preset-proportion energy of the N audio frames and determining an average value of minimum bandwidths of distribution, on the spectrums, of third-preset-proportion energy of the N audio frames according to the energy of the P spectral envelopes of each of the N audio frames, where the average value of the minimum bandwidths of distribution, on the spectrums, of the second-preset-proportion energy of the N audio frames is used as the second minimum bandwidth, the average value of the minimum bandwidths of distribution, on the spectrums, of the third-preset-proportion energy of the N audio frames is used as the third minimum bandwidth, and the second preset proportion is less than the third preset proportion. 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 includes: when the second minimum bandwidth is less than a third preset value and the third minimum bandwidth is less than a fourth preset value, determining to use the first encoding method to encode the current audio frame; when the third minimum bandwidth is less than a fifth preset value, determining to use the first encoding method to encode the current audio frame; or when the third minimum bandwidth is greater than a sixth preset value, determining to use the second encoding method to encode the current audio frame. The fourth preset value is greater than or equal to the third preset value, the fifth preset value is

less than the fourth preset value, and the sixth preset value is greater than the fourth preset value. Optionally, in an embodiment, when N is 1, the N audio frames are the current audio frame. The determining an average value of minimum bandwidths of distribution, on the spectrums, of second-preset-proportion energy of the N audio frames as the second minimum bandwidth includes: determining a minimum bandwidth of distribution, on the spectrum, of second-preset-proportion energy of the current audio frame as the second minimum bandwidth. The determining an average value of minimum bandwidths of distribution, on the spectrums, of third-preset-proportion energy of the N audio frames as the third minimum bandwidth includes: determining a minimum bandwidth of distribution, on the spectrum, of third-preset-proportion energy of the current audio frame as the third minimum bandwidth.

[0056] A person skilled in the art may understand that, the third preset value, the fourth preset value, the fifth preset value, the sixth preset value, the second preset proportion, and the third preset proportion may be determined according to a simulation experiment. Appropriate preset values and preset proportions may be determined by means of a simulation experiment, so that a good encoding effect can be obtained when an audio frame meeting the foregoing condition is encoded by using the first encoding method or the second encoding method.

[0057] The determining an average value of minimum bandwidths of distribution, on the spectrums, of second-preset-proportion energy of the N audio frames and determining an average value of minimum bandwidths of distribution, on the spectrums, of third-preset-proportion energy of the N audio frames according to the energy of the P spectral envelopes of each of the N audio frames includes: sorting the energy of the P spectral envelopes of each audio frame in descending order; determining, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the second preset proportion of each of the N audio frames; determining, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the second preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the second preset proportion of the N audio frames; determining, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the third preset proportion of each of the N audio frames; and determining, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the third preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the third preset proportion of the N audio frames. For example, an input audio signal is a wideband signal sampled at 16 kHz,

and the input signal is input in a frame of 20 ms. Each frame of signal is 320 time domain sampling points. Time-frequency transform is performed on a time domain signal. For example, time-frequency transform is performed by means of fast Fourier transform, to obtain 160 spectral envelopes S(k), where k=0, 1, 2, ..., 159. A minimum bandwidth is found from the spectral envelopes S(k) in a manner that a proportion that energy on the bandwidth accounts for in total energy of the frame is the second preset proportion. A bandwidth continues to be found from the spectral envelopes S(k) in a manner that a proportion that energy on the bandwidth accounts for in the total energy is the third preset proportion. Specifically, determining, according to energy, sorted in descending order, of P spectral envelopes of the audio frame, a minimum bandwidth of distribution, on a spectrum, of energy that accounts for not less than the second preset proportion of an audio frame and a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the third preset proportion of the audio frame includes: sequentially accumulating energy of frequency bins in the spectral envelopes S(k) in descending order. Energy obtained after each time of accumulation is compared with total energy of the audio frame, and if a proportion is greater than the second preset proportion, a quantity of times of accumulation is a minimum bandwidth that meets being not less than the second preset proportion. The accumulation is continued, and if a proportion of energy obtained after accumulation to the total energy of the audio frame is greater than the third preset proportion, the accumulation is ended, and a quantity of times of accumulation is a minimum bandwidth that meets being not less than the third preset proportion. For example, the second preset proportion is 85%, and the third preset proportion is 95%. If a proportion that an energy sum obtained after 30 times of accumulation accounts for in the total energy exceeds 85%, it may be considered that the minimum bandwidth of distribution, on the spectrum, of the second-preset-proportion energy of the audio frame is 30. The accumulation is continued, and if a proportion that an energy sum obtained after 35 times of accumulation accounts for in the total energy is 95%, it may be considered that the minimum bandwidth of distribution, on the spectrum, of the third-preset-proportion energy of the audio frame is 35. The foregoing process is executed for each of the N audio frames, to separately determine the minimum bandwidths of distribution, on the spectrums, of the energy that accounts for not less than the second preset proportion of the N audio frames including the current audio frame and the minimum bandwidths of distribution, on the spectrums, of the energy that accounts for not less than the third preset proportion of the N audio frames including the current audio frame. The average value of the minimum bandwidths of distribution, on the spectrums, of the energy that accounts for not less than the second preset proportion of the N audio frames is the second minimum bandwidth. The average value of the minimum bandwidths of distribution, on the spectrums, of the energy that accounts for not less than the third preset proportion of the N

audio frames is the third minimum bandwidth. When the second minimum bandwidth is less than the third preset value and the third minimum bandwidth is less than the fourth preset value, it is determined to use the first encoding method to encode the current audio frame. When the third minimum bandwidth is less than the fifth preset value, it is determined to use the first encoding method to encode the current audio frame. When the third minimum bandwidth is greater than the sixth preset value, it is determined to use the second encoding method to encode the current audio frame.

[0058] Optionally, in another embodiment, the general sparseness parameter includes a second energy proportion and a third energy proportion. In this case, the determining a general sparseness parameter according to energy of the P spectral envelopes of each of the N audio frames includes: selecting P_2 spectral envelopes from the P spectral envelopes of each of the N audio frames; determining the second energy proportion according to energy of the P_2 spectral envelopes of each of the N audio frames and total energy of the respective N audio frames; selecting P_3 spectral envelopes from the P spectral envelopes of each of the N audio frames; and determining 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. 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 includes: when the second energy proportion is greater than a seventh preset value and the third energy proportion is greater than an eighth preset value, determining to use the first encoding method to encode the current audio frame; when the second energy proportion is greater than a ninth preset value, determining to use the first encoding method to encode the current audio frame; or when the third energy proportion is less than a tenth preset value, determining to use the second encoding method to encode the current audio frame. P_2 and P_3 are positive integers less than P, and P_2 is less than P_3. Optionally, in an embodiment, when N is 1, the N audio frames are the current audio frame. The determining the second energy proportion according to energy of the P_2 spectral envelopes of each of the N audio frames and total energy of the respective N audio frames includes: determining the second energy proportion according to energy of P_2 spectral envelopes of the current audio frame and total energy of the current audio frame. The determining 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 includes: determining the third energy proportion according to energy of P_3 spectral envelopes of the current audio frame and the total energy of the current audio frame.

[0059] A person skilled in the art may understand that, values of P_2 and P_3, the seventh preset value, the eighth preset value, the ninth preset value, and the tenth preset value may be determined

[0073] For example, when a general sparseness parameter is determined and the general sparseness parameter is a first minimum bandwidth, the hangover period may be re-determined according to a minimum bandwidth of distribution, on a spectrum, of first preset proportion energy of an audio frame. It is assumed that it is determined to use the first encoding method to encode the I^{th} audio frame, and a preset hangover period is L. A minimum bandwidth of distribution, on a spectrum, of first preset proportion energy of each of H consecutive audio frames including the $(I+1)^{th}$ audio frame is determined, where H is a positive integer greater than 0. If the $(I+1)^{th}$ audio frame does not meet the condition of using the first encoding method, a quantity of audio frames whose minimum bandwidths of distribution, on spectrums, of first preset proportion energy are less than a fifteenth preset value (the quantity is briefly referred to as a first hangover parameter) is determined. When a minimum bandwidth of distribution, on a spectrum, of first preset proportion energy of an $(L+1)^{th}$ audio frame is greater than a sixteenth preset value and is less than a seventeenth preset value, and the first hangover parameter is less than an eighteenth preset value, the hangover period length is subtracted by 1, that is, the hangover update parameter is 1. The sixteenth preset value is greater than the first preset value. When the minimum bandwidth of distribution, on the spectrum, of the first preset proportion energy of the $(L+1)^{th}$ audio frame is greater than the seventeenth preset value and is less than a nineteenth preset value, and the first hangover parameter is less than the eighteenth preset value, the hangover period length is subtracted by 2, that is, the hangover update parameter is 2. When the minimum bandwidth of distribution, on the spectrum, of the first preset proportion energy of the $(L+1)^{th}$ audio frame is greater than the nineteenth preset value, the hangover period is set to 0. When the first hangover parameter and the minimum bandwidth of distribution, on the spectrum, of the first preset proportion energy of the $(L+1)^{th}$ audio frame do not meet one or more of the sixteenth preset value to the nineteenth preset value, the hangover period remains unchanged.

[0074] A person skilled in the art may understand that, the preset hangover period may be set according to an actual status, and the hangover update parameter also may be adjusted according to an actual status. The fifteenth preset value to the nineteenth preset value may be adjusted according to an actual status, so that different hangover periods may be set.

[0075] Similarly, when the general sparseness parameter includes a second minimum bandwidth and a third minimum bandwidth, or the general sparseness parameter includes a first energy proportion, or the general sparseness parameter includes a second energy proportion and a third energy proportion, a corresponding preset hangover period, a corresponding hangover update parameter, and a related parameter used to determine the hangover update parameter may be set, so

[0084] During selection of an appropriate encoding method for an audio frame, sparseness of distribution, on a spectrum, of energy of the audio frame may be considered. There may be three types of sparseness of distribution, on a spectrum, of energy of an audio frame: general sparseness, burst sparseness, and band-limited sparseness.

[0085] Optionally, in an embodiment, an appropriate encoding method may be selected for the current audio frame by using the general sparseness. In this case, the determining unit 202 is specifically configured to divide a spectrum of each of the N audio frames into P spectral envelopes, and determine a general sparseness parameter according to energy of the P spectral envelopes of each of the N audio frames, where P is a positive integer, and the general sparseness parameter indicates the sparseness of distribution, on the spectrums, of the energy of the N audio frames.

[0086] Specifically, an average value of minimum bandwidths of distribution, on spectrums, of specific-proportion energy of N input consecutive audio frames may be defined as the general sparseness. A smaller bandwidth indicates stronger general sparseness, and a larger bandwidth indicates weaker general sparseness. In other words, stronger general sparseness indicates that energy of an audio frame is more centralized, and weaker general sparseness indicates that energy of an audio frame is more disperse. Efficiency is high when the first encoding method is used to encode an audio frame whose general sparseness is relatively strong. Therefore, an appropriate encoding method may be selected by determining general sparseness of an audio frame, to encode the audio frame. To help determine general sparseness of an audio frame, the general sparseness may be quantized to obtain a general sparseness parameter. Optionally, when N is 1, the general sparseness is a minimum bandwidth of distribution, on a spectrum, of specific-proportion energy of the current audio frame.

[0087] Optionally, in an embodiment, the general sparseness parameter includes a first minimum bandwidth. In this case, the determining unit 202 is specifically configured to determine an average value of minimum bandwidths of distribution, 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, where the average value of the minimum bandwidths of distribution, on the spectrums, of the first preset proportion energy of the N audio frames is the first minimum bandwidth. The determining unit 202 is specifically configured to: when the first minimum bandwidth is less than a first preset value, determine to use the first encoding method to encode the current audio frame; and when the first minimum bandwidth is greater than the first preset value, determine to use the second encoding method to encode the current audio frame.

[0088] A person skilled in the art may understand that, the first preset value and the first preset proportion may be determined according to a simulation experiment. An appropriate first preset

value and first preset proportion may be determined by means of a simulation experiment, so that a good encoding effect can be obtained when an audio frame meeting the foregoing condition is encoded by using the first encoding method or the second encoding method.

[0089] The determining unit 202 is specifically configured to: sort the energy of the P spectral envelopes of each audio frame in descending order; determine, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the first preset proportion of each of the N audio frames; and determine, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the first preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the first preset proportion of the N audio frames. For example, an audio signal obtained by the obtaining unit 201 is a wideband signal sampled at 16 kHz, and the obtained audio signal is obtained in a frame of 20 ms. Each frame of signal is 320 time domain sampling points. The determining unit 202 may perform time-frequency transform on a time domain signal, for example, perform time-frequency transform by means of fast Fourier transform (Fast Fourier Transformation, FFT), to obtain 160 spectral envelopes S(k), that is, 160 FFT energy spectrum coefficients, where k=0, 1, 2, ..., 159. The determining unit 202 may find a minimum bandwidth from the spectral envelopes S(k) in a manner that a proportion that energy on the bandwidth accounts for in total energy of the frame is the first preset proportion. Specifically, the determining unit 202 may sequentially accumulate energy of frequency bins in the spectral envelopes S(k) in descending order; and compare energy obtained after each time of accumulation with the total energy of the audio frame, and if a proportion is greater than the first preset proportion, end the accumulation process, where a quantity of times of accumulation is the minimum bandwidth. For example, the first preset proportion is 90%, and if a proportion that an energy sum obtained after 30 times of accumulation accounts for in the total energy exceeds 90%, it may be considered that a minimum bandwidth of energy that accounts for not less than the first preset proportion of the audio frame is 30. The determining unit 202 may execute the foregoing minimum bandwidth determining process for each of the N audio frames, to separately determine the minimum bandwidths of the energy that accounts for not less than the first preset proportion of the N audio frames including the current audio frame. The determining unit 202 may calculate an average value of the minimum bandwidths of the energy that accounts for not less than the first preset proportion of the N audio frames. The average value of the minimum bandwidths of the energy that accounts for not less than the first preset proportion of the N audio frames may be referred to as the first minimum bandwidth, and the first minimum bandwidth may be used as the

spectral envelopes may be determined by means of a simulation experiment, so that a good encoding effect can be obtained when an audio frame meeting the foregoing condition is encoded by using the first encoding method or the second encoding method. Optionally, in an embodiment, the P_1 spectral envelopes may be P_1 spectral envelopes having maximum energy in the P spectral envelopes.

[0093] For example, an audio signal obtained by the obtaining unit 201 is a wideband signal sampled at 16 kHz, and the obtained audio signal is obtained in a frame of 20 ms. Each frame of signal is 320 time domain sampling points. The determining unit 202 may perform time-frequency transform on a time domain signal, for example, perform time-frequency transform by means of fast Fourier transform, to obtain 160 spectral envelopes S(k), where k=0, 1, 2, ..., 159. The determining unit 202 may select P_1 spectral envelopes from the 160 spectral envelopes, and calculate a proportion that an energy sum of the P_1 spectral envelopes accounts for in total energy of the audio frame. The determining unit 202 may execute the foregoing process for each of the N audio frames, that is, calculate a proportion that an energy sum of the P_1 spectral envelopes of each of the N audio frames accounts for in respective total energy. The determining unit 202 may calculate an average value of the proportions. The average value of the proportions is the first energy proportion. When the first energy proportion is greater than the second preset value, the determining unit 202 may determine to use the first encoding method to encode the current audio frame. When the first energy proportion is less than the second preset value, the determining unit 202 may determine to use the second encoding method to encode the current audio frame. The P_1 spectral envelopes may be P_1 spectral envelopes having maximum energy in the P spectral envelopes. That is, the determining unit 202 is specifically configured to determine, from the P spectral envelopes of each of the N audio frames, P_1 spectral envelopes having maximum energy. Optionally, in an embodiment, the value of P_1 may be 20.

[0094] Optionally, in another embodiment, the general sparseness parameter may include a second minimum bandwidth and a third minimum bandwidth. In this case, the determining unit 202 is specifically configured to determine an average value of minimum bandwidths of distribution, on the spectrums, of second-preset-proportion energy of the N audio frames and determine an average value of minimum bandwidths of distribution, on the spectrums, of third-preset-proportion energy of the N audio frames according to the energy of the P spectral envelopes of each of the N audio frames, where the average value of the minimum bandwidths of distribution, on the spectrums, of the second-preset-proportion energy of the N audio frames is used as the second minimum bandwidth, the average value of the minimum bandwidths of distribution, on the spectrums, of the third-preset-proportion energy of the N audio frames is used as the third minimum bandwidth, and the second preset proportion is less than the third preset

proportion. The determining unit 202 is specifically configured to: when the second minimum bandwidth is less than a third preset value and the third minimum bandwidth is less than a fourth preset value, determine to use the first encoding method to encode the current audio frame; when the third minimum bandwidth is less than a fifth preset value, determine to use the first encoding method to encode the current audio frame; and when the third minimum bandwidth is greater than a sixth preset value, determine to use the second encoding method to encode the current audio frame. Optionally, in an embodiment, when N is 1, the N audio frames are the current audio frame. The determining unit 202 may determine a minimum bandwidth of distribution, on the spectrum, of second-preset-proportion energy of the current audio frame as the second minimum bandwidth. The determining unit 202 may determine a minimum bandwidth of distribution, on the spectrum, of third-preset-proportion energy of the current audio frame as the third minimum bandwidth.

[0095] A person skilled in the art may understand that, the third preset value, the fourth preset value, the fifth preset value, the sixth preset value, the second preset proportion, and the third preset proportion may be determined according to a simulation experiment. Appropriate preset values and preset proportions may be determined by means of a simulation experiment, so that a good encoding effect can be obtained when an audio frame meeting the foregoing condition is encoded by using the first encoding method or the second encoding method.

[0096] The determining unit 202 is specifically configured to: sort the energy of the P spectral envelopes of each audio frame in descending order; determine, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the second preset proportion of each of the N audio frames; determine, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the second preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the second preset proportion of the N audio frames; determine, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the third preset proportion of each of the N audio frames; and determine, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the third preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the third preset proportion of the N audio frames. For example, an audio signal obtained by the obtaining unit 201 is a wideband signal sampled at 16 kHz, and the obtained audio signal is obtained in a frame of 20 ms. Each frame of signal is 320 time domain sampling points. The determining unit 202 may perform time-frequency transform on

a time domain signal, for example, perform time-frequency transform by means of fast Fourier transform, to obtain 160 spectral envelopes S(k), where k=0, 1, 2, ..., 159. The determining unit 202 may find a minimum bandwidth from the spectral envelopes S(k) in a manner that a proportion that energy on the bandwidth accounts for in total energy of the frame is not less than the second preset proportion. The determining unit 202 may continue to find a bandwidth from the spectral envelopes S(k) in a manner that a proportion that energy on the bandwidth accounts for in the total energy is not less than the third preset proportion. Specifically, the determining unit 202 may sequentially accumulate energy of frequency bins in the spectral envelopes S(k) in descending order. Energy obtained after each time of accumulation is compared with the total energy of the audio frame, and if a proportion is greater than the second preset proportion, a quantity of times of accumulation is a minimum bandwidth that is not less than the second preset proportion. The determining unit 202 may continue the accumulation. If a proportion of energy obtained after accumulation to the total energy of the audio frame is greater than the third preset proportion, the accumulation is ended, and a quantity of times of accumulation is a minimum bandwidth that is not less than the third preset proportion. For example, the second preset proportion is 85%, and the third preset proportion is 95%. If a proportion that an energy sum obtained after 30 times of accumulation accounts for in the total energy exceeds 85%, it may be considered that the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the second preset proportion of the audio frame is 30. The accumulation is continued, and if a proportion that an energy sum obtained after 35 times of accumulation accounts for in the total energy is 95%, it may be considered that the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the third preset proportion of the audio frame is 35. The determining unit 202 may execute the foregoing process for each of the N audio frames. The determining unit 202 may separately determine the minimum bandwidths of distribution, on the spectrums, of the energy that accounts for not less than the second preset proportion of the N audio frames including the current audio frame and the minimum bandwidths of distribution, on the spectrums, of the energy that accounts for not less than the third preset proportion of the N audio frames including the current audio frame. The average value of the minimum bandwidths of distribution, on the spectrums, of the energy that accounts for not less than the second preset proportion of the N audio frames is the second minimum bandwidth. The average value of the minimum bandwidths of distribution, on the spectrums, of the energy that accounts for not less than the third preset proportion of the N audio frames is the third minimum bandwidth. When the second minimum bandwidth is less than the third preset value and the third minimum bandwidth is less

[0111] The hangover period length may be updated according to sparseness of distribution, on a spectrum, of energy of an audio frame in the hangover period, until the hangover period length is 0.

[0112] For example, if the determining unit 202 determines to use the first encoding method for an I^{th} audio frame and a length of a preset hangover period is L, the determining unit 202 may determine that the first encoding method is used for an $(I+1)^{th}$ audio frame to an $(I+L)^{th}$ audio frame. Then, the determining unit 202 may determine sparseness of distribution, on a spectrum, of energy of the $(I+1)^{th}$ audio frame, and re-calculate the hangover period according to the sparseness of distribution, on the spectrum, of the energy of the $(I+1)^{th}$ audio frame. If the $(I+1)^{th}$ audio frame still meets a condition of using the first encoding method, the determining unit 202 may determine that a subsequent hangover period is still the preset hangover period L. That is, the hangover period starts from an $(L+2)^{th}$ audio frame to an $(I+1+L)^{th}$ audio frame. If the $(I+1)^{th}$ audio frame does not meet the condition of using the first encoding method, the determining unit 202 may re-determine the hangover period according to the sparseness of distribution, on the spectrum, of the energy of the $(I+1)^{th}$ audio frame. For example, the determining unit 202 may re-determine that the hangover period is L–L1, where L1 is a positive integer less than or equal to L. If L1 is equal to L, the hangover period length is updated to 0. In this case, the determining unit 202 may re-determine the encoding method according to the sparseness of distribution, on the spectrum, of the energy of the $(I+1)^{th}$ audio frame. If L1 is an integer less than L, the determining unit 202 may re-determine the encoding method according to sparseness of distribution, on a spectrum, of energy of an $(I+1+L–L1)^{th}$ audio frame. However, because the $(I+1)^{th}$ audio frame is in a hangover period of the I^{th} audio frame, the $(I+1)^{th}$ audio frame is still encoded by using the first encoding method. L1 may be referred to as a hangover update parameter, and a value of the hangover update parameter may be determined according to sparseness of distribution, on a spectrum, of energy of an input audio frame. In this way, hangover period update is related to sparseness of distribution, on a spectrum, of energy of an audio frame.

[0113] For example, when a general sparseness parameter is determined and the general sparseness parameter is a first minimum bandwidth, the determining unit 202 may re-determine the hangover period according to a minimum bandwidth of distribution, on a spectrum, of first preset proportion energy of an audio frame. It is assumed that it is determined to use the first encoding method to encode the I^{th} audio frame, and a preset hangover period is L. The determining unit 202 may determine a minimum bandwidth of distribution, on a spectrum, of first preset proportion energy of each of H consecutive audio frames including the $(I+1)^{th}$ audio frame, where H is a positive integer greater than 0. If the $(I+1)^{th}$ audio frame does not meet the condition of using the first encoding method, the determining unit 202 may determine a quantity of audio

frames whose minimum bandwidths of distribution, on spectrums, of first preset proportion energy are less than a fifteenth preset value (the quantity is briefly referred to as a first hangover parameter). When a minimum bandwidth of distribution, on a spectrum, of first preset proportion energy of an $(L+1)^{th}$ audio frame is greater than a sixteenth preset value and is less than a seventeenth preset value, and the first hangover parameter is less than an eighteenth preset value, the determining unit 202 may subtract the hangover period length by 1, that is, the hangover update parameter is 1. The sixteenth preset value is greater than the first preset value. When the minimum bandwidth of distribution, on the spectrum, of the first preset proportion energy of the $(L+1)^{th}$ audio frame is greater than the seventeenth preset value and is less than a nineteenth preset value, and the first hangover parameter is less than the eighteenth preset value, the determining unit 202 may subtract the hangover period length by 2, that is, the hangover update parameter is 2. When the minimum bandwidth of distribution, on the spectrum, of the first preset proportion energy of the $(L+1)^{th}$ audio frame is greater than the nineteenth preset value, the determining unit 202 may set the hangover period to 0. When the first hangover parameter and the minimum bandwidth of distribution, on the spectrum, of the first preset proportion energy of the $(L+1)^{th}$ audio frame do not meet one or more of the sixteenth preset value to the nineteenth preset value, the determining unit 202 may determine that the hangover period remains unchanged.

[0114] A person skilled in the art may understand that, the preset hangover period may be set according to an actual status, and the hangover update parameter also may be adjusted according to an actual status. The fifteenth preset value to the nineteenth preset value may be adjusted according to an actual status, so that different hangover periods may be set.

[0115] Similarly, when the general sparseness parameter includes a second minimum bandwidth and a third minimum bandwidth, or the general sparseness parameter includes a first energy proportion, or the general sparseness parameter includes a second energy proportion and a third energy proportion, the determining unit 202 may set a corresponding preset hangover period, a corresponding hangover update parameter, and a related parameter used to determine the hangover update parameter, so that a corresponding hangover period can be determined, and frequent switching between encoding methods is avoided.

[0116] When the encoding method is determined according to the burst sparseness (that is, the encoding method is determined according to global sparseness, local sparseness, and short-time burstiness of distribution, on a spectrum, of energy of an audio frame), the determining unit 202 may set a corresponding hangover period, a corresponding hangover update parameter, and a related parameter used to determine the hangover update parameter, to avoid frequent switching between

[0122] The processor 301 is configured to obtain N audio frames, where the N audio frames include a current audio frame, and N is a positive integer.

[0123] The processor 301 is configured to determine sparseness of distribution, on the spectrums, of energy of the N audio frames obtained by the processor 301.

[0124] The processor 301 is further configured to determine, 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 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.

[0125] According to the apparatus shown in FIG. 3, when an audio frame is encoded, sparseness of distribution, on a spectrum, of energy of the audio frame is considered, which can reduce encoding complexity and ensure that encoding is of relatively high accuracy.

[0126] During selection of an appropriate encoding method for an audio frame, sparseness of distribution, on a spectrum, of energy of the audio frame may be considered. There may be three types of sparseness of distribution, on a spectrum, of energy of an audio frame: general sparseness, burst sparseness, and band-limited sparseness.

[0127] Optionally, in an embodiment, an appropriate encoding method may be selected for the current audio frame by using the general sparseness. In this case, the processor 301 is specifically configured to divide a spectrum of each of the N audio frames into P spectral envelopes, and determine a general sparseness parameter according to energy of the P spectral envelopes of each of the N audio frames, where P is a positive integer, and the general sparseness parameter indicates the sparseness of distribution, on the spectrums, of the energy of the N audio frames.

[0128] Specifically, an average value of minimum bandwidths of distribution, on spectrums, of specific-proportion energy of N input consecutive audio frames may be defined as the general sparseness. A smaller bandwidth indicates stronger general sparseness, and a larger bandwidth indicates weaker general sparseness. In other words, stronger general sparseness indicates that energy of an audio frame is more centralized, and weaker general sparseness indicates that energy of an audio frame is more disperse. Efficiency is high when the first encoding method is used to encode an audio frame whose general sparseness is relatively strong. Therefore, an appropriate encoding method may be selected by determining general sparseness of an audio frame, to encode the audio frame. To help determine general sparseness of an audio frame, the general sparseness may be quantized to obtain a general sparseness parameter. Optionally, when N is 1, the general sparseness is a minimum bandwidth of distribution, on a spectrum, of specific-proportion energy of the current audio frame.

[0129] Optionally, in an embodiment, the general sparseness parameter includes a first minimum bandwidth. In this case, the processor 301 is specifically configured to determine an average value of minimum bandwidths of distribution, 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, where the average value of the minimum bandwidths of distribution, on the spectrums, of the first preset proportion energy of the N audio frames is the first minimum bandwidth. The processor 301 is specifically configured to: when the first minimum bandwidth is less than a first preset value, determine to use the first encoding method to encode the current audio frame; and when the first minimum bandwidth is greater than the first preset value, determine to use the second encoding method to encode the current audio frame.

[0130] A person skilled in the art may understand that, the first preset value and the first preset proportion may be determined according to a simulation experiment. An appropriate first preset value and first preset proportion may be determined by means of a simulation experiment, so that a good encoding effect can be obtained when an audio frame meeting the foregoing condition is encoded by using the first encoding method or the second encoding method.

[0131] The processor 301 is specifically configured to: sort the energy of the P spectral envelopes of each audio frame in descending order; determine, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the first preset proportion of each of the N audio frames; and determine, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the first preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the first preset proportion of the N audio frames. For example, an audio signal obtained by the processor 301 is a wideband signal sampled at 16 kHz, and the obtained audio signal is obtained in a frame of 30 ms. Each frame of signal is 330 time domain sampling points. The processor 301 may perform time-frequency transform on a time domain signal, for example, perform time-frequency transform by means of fast Fourier transform (Fast Fourier Transformation, FFT), to obtain 130 spectral envelopes S(k), that is, 130 FFT energy spectrum coefficients, where k=0, 1, 2, ..., 159. The processor 301 may find a minimum bandwidth from the spectral envelopes S(k) in a manner that a proportion that energy on the bandwidth accounts for in total energy of the frame is the first preset proportion. Specifically, the processor 301 may sequentially accumulate energy of frequency bins in the spectral envelopes S(k) in descending order;

from the P spectral envelopes of each of the N audio frames, P_1 spectral envelopes having maximum energy. Optionally, in an embodiment, the value of P_1 may be 30.

[0136] Optionally, in another embodiment, the general sparseness parameter may include a second minimum bandwidth and a third minimum bandwidth. In this case, the processor 301 is specifically configured to determine an average value of minimum bandwidths of distribution, on the spectrums, of second-preset-proportion energy of the N audio frames and determine an average value of minimum bandwidths of distribution, on the spectrums, of third-preset-proportion energy of the N audio frames according to the energy of the P spectral envelopes of each of the N audio frames, where the average value of the minimum bandwidths of distribution, on the spectrums, of the second-preset-proportion energy of the N audio frames is used as the second minimum bandwidth, the average value of the minimum bandwidths of distribution, on the spectrums, of the third-preset-proportion energy of the N audio frames is used as the third minimum bandwidth, and the second preset proportion is less than the third preset proportion. The processor 301 is specifically configured to: when the second minimum bandwidth is less than a third preset value and the third minimum bandwidth is less than a fourth preset value, determine to use the first encoding method to encode the current audio frame; when the third minimum bandwidth is less than a fifth preset value, determine to use the first encoding method to encode the current audio frame; and when the third minimum bandwidth is greater than a sixth preset value, determine to use the second encoding method to encode the current audio frame. Optionally, in an embodiment, when N is 1, the N audio frames are the current audio frame. The processor 301 may determine a minimum bandwidth of distribution, on the spectrum, of second-preset-proportion energy of the current audio frame as the second minimum bandwidth. The processor 301 may determine a minimum bandwidth of distribution, on the spectrum, of third-preset-proportion energy of the current audio frame as the third minimum bandwidth.

[0137] A person skilled in the art may understand that, the third preset value, the fourth preset value, the fifth preset value, the sixth preset value, the second preset proportion, and the third preset proportion may be determined according to a simulation experiment. Appropriate preset values and preset proportions may be determined by means of a simulation experiment, so that a good encoding effect can be obtained when an audio frame meeting the foregoing condition is encoded by using the first encoding method or the second encoding method.

[0138] The processor 301 is specifically configured to: sort the energy of the P spectral envelopes of each audio frame in descending order; determine, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the second preset proportion of each of the N audio frames; determine, according to the minimum bandwidth

of distribution, on the spectrum, of the energy that accounts for not less than the second preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the second preset proportion of the N audio frames; determine, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the third preset proportion of each of the N audio frames; and determine, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the third preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the third preset proportion of the N audio frames. For example, an audio signal obtained by the processor 301 is a wideband signal sampled at 16 kHz, and the obtained audio signal is obtained in a frame of 30 ms. Each frame of signal is 330 time domain sampling points. The processor 301 may perform time-frequency transform on a time domain signal, for example, perform time-frequency transform by means of fast Fourier transform, to obtain 130 spectral envelopes S(k), where k=0, 1, 2, ..., 159. The processor 301 may find a minimum bandwidth from the spectral envelopes S(k) in a manner that a proportion that energy on the bandwidth accounts for in total energy of the frame is not less than the second preset proportion. The processor 301 may continue to find a bandwidth from the spectral envelopes S(k) in a manner that a proportion that energy on the bandwidth accounts for in the total energy is not less than the third preset proportion. Specifically, the processor 301 may sequentially accumulate energy of frequency bins in the spectral envelopes S(k) in descending order. Energy obtained after each time of accumulation is compared with the total energy of the audio frame, and if a proportion is greater than the second preset proportion, a quantity of times of accumulation is a minimum bandwidth that is not less than the second preset proportion. The processor 301 may continue the accumulation. If a proportion of energy obtained after accumulation to the total energy of the audio frame is greater than the third preset proportion, the accumulation is ended, and a quantity of times of accumulation is a minimum bandwidth that is not less than the third preset proportion. For example, the second preset proportion is 85%, and the third preset proportion is 95%. If a proportion that an energy sum obtained after 30 times of accumulation accounts for in the total energy exceeds 85%, it may be considered that the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the second preset proportion of the audio frame is 30. The accumulation is continued, and if a proportion that an energy sum obtained after 35 times of accumulation accounts for in the total energy is 95%, it may be considered that the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the third preset proportion of the audio frame is 35. The processor 301

may execute the foregoing process for each of the N audio frames. The processor 301 may separately determine the minimum bandwidths of distribution, on the spectrums, of the energy that accounts for not less than the second preset proportion of the N audio frames including the current audio frame and the minimum bandwidths of distribution, on the spectrums, of the energy that accounts for not less than the third preset proportion of the N audio frames including the current audio frame. The average value of the minimum bandwidths of distribution, on the spectrums, of the energy that accounts for not less than the second preset proportion of the N audio frames is the second minimum bandwidth. The average value of the minimum bandwidths of distribution, on the spectrums, of the energy that accounts for not less than the third preset proportion of the N audio frames is the third minimum bandwidth. When the second minimum bandwidth is less than the third preset value and the third minimum bandwidth is less than the fourth preset value, the processor 301 may determine to use the first encoding method to encode the current audio frame. When the third minimum bandwidth is less than the fifth preset value, the processor 301 may determine to use the first encoding method to encode the current audio frame. When the third minimum bandwidth is greater than the sixth preset valuc, the processor 301 may determine to use the second encoding method to encode the current audio frame.

[0139] Optionally, in another embodiment, the general sparseness parameter includes a second energy proportion and a third energy proportion. In this case, the processor 301 is specifically configured to: select P_2 spectral envelopes from the P spectral envelopes of each of the N audio frames, determine the second energy proportion according to energy of the P_2 spectral envelopes of each of the N audio frames and total energy of the respective N audio frames, select P_3 spectral envelopes from the P spectral envelopes of each of the 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, where P_2 and P_3 are positive integers less than P, and P_2 is less than P_3. The processor 301 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. Optionally, in an embodiment, when N is 1, the N audio frames are the current audio frame. The processor 301 may determine the second energy proportion according to energy of P_2 spectral envelopes of the current audio frame and total energy of the current audio frame. The processor 301

[0152] If a hangover length of the hangover period is L, the processor 301 may be configured to determine that L audio frames after the current audio frame all belong to a hangover period of the current audio frame. If sparseness of distribution, on a spectrum, of energy of an audio frame belonging the hangover period is different from sparseness of distribution, on a spectrum, of energy of an audio frame at a start position of the hangover period, the processor 301 may be configured to determine that the audio frame is still encoded by using an encoding method that is the same as that used for the audio frame at the start position of the hangover period.

[0153] The hangover period length may be updated according to sparseness of distribution, on a spectrum, of energy of an audio frame in the hangover period, until the hangover period length is 0.

[0154] For example, if the processor 301 determines to use the first encoding method for an I^{th} audio frame and a length of a preset hangover period is L, the processor 301 may determine that the first encoding method is used for an $(I+1)^{th}$ audio frame to an $(I+L)^{th}$ audio frame. Then, the processor 301 may determine sparseness of distribution, on a spectrum, of energy of the $(I+1)^{th}$ audio frame, and re-calculate the hangover period according to the sparseness of distribution, on the spectrum, of the energy of the $(I+1)^{th}$ audio frame. If the $(I+1)^{th}$ audio frame still meets a condition of using the first encoding method, the processor 301 may determine that a subsequent hangover period is still the preset hangover period L. That is, the hangover period starts from an $(L+2)^{th}$ audio frame to an $(I+1+L)^{th}$ audio frame. If the $(I+1)^{th}$ audio frame does not meet the condition of using the first encoding method, the processor 301 may re-determine the hangover period according to the sparseness of distribution, on the spectrum, of the energy of the $(I+1)^{th}$ audio frame. For example, the processor 301 may re-determine that the hangover period is L–L1, where L1 is a positive integer less than or equal to L. If L1 is equal to L, the hangover period length is updated to 0. In this case, the processor 301 may re-determine the encoding method according to the sparseness of distribution, on the spectrum, of the energy of the $(I+1)^{th}$ audio frame. If L1 is an integer less than L, the processor 301 may re-determine the encoding method according to sparseness of distribution, on a spectrum, of energy of an $(I+1+L-L1)^{th}$ audio frame. However, because the $(I+1)^{th}$ audio frame is in a hangover period of the I^{th} audio frame, the $(I+1)^{th}$ audio frame is still encoded by using the first encoding method. L1 may be referred to as a hangover update parameter, and a value of the hangover update parameter may be determined according to sparseness of distribution, on a spectrum, of energy of an input audio frame. In this way, hangover period update is related to sparseness of distribution, on a spectrum, of energy of an audio frame.

[0155] For example, when a general sparseness parameter is determined and the general sparseness parameter is a first minimum bandwidth, the processor 301 may re-determine the

hangover period according to a minimum bandwidth of distribution, on a spectrum, of first preset proportion energy of an audio frame. It is assumed that it is determined to use the first encoding method to encode the I^{th} audio frame, and a preset hangover period is L. The processor 301 may determine a minimum bandwidth of distribution, on a spectrum, of first preset proportion energy of each of H consecutive audio frames including the $(I+1)^{th}$ audio frame, where H is a positive integer greater than 0. If the $(I+1)^{th}$ audio frame does not meet the condition of using the first encoding method, the processor 301 may determine a quantity of audio frames whose minimum bandwidths of distribution, on spectrums, of first preset proportion energy are less than a fifteenth preset value (the quantity is briefly referred to as a first hangover parameter). When a minimum bandwidth of distribution, on a spectrum, of first preset proportion energy of an $(L+1)^{th}$ audio frame is greater than a sixteenth preset value and is less than a seventeenth preset value, and the first hangover parameter is less than an eighteenth preset value, the processor 301 may subtract the hangover period length by 1, that is, the hangover update parameter is 1. The sixteenth preset value is greater than the first preset value. When the minimum bandwidth of distribution, on the spectrum, of the first preset proportion energy of the $(L+1)^{th}$ audio frame is greater than the seventeenth preset value and is less than a nineteenth preset value, and the first hangover parameter is less than the eighteenth preset value, the processor 301 may subtract the hangover period length by 2, that is, the hangover update parameter is 2. When the minimum bandwidth of distribution, on the spectrum, of the first preset proportion energy of the $(L+1)^{th}$ audio frame is greater than the nineteenth preset value, the processor 301 may set the hangover period to 0. When the first hangover parameter and the minimum bandwidth of distribution, on the spectrum, of the first preset proportion energy of the $(L+1)^{th}$ audio frame do not meet one or more of the sixteenth preset value to the nineteenth preset value, the processor 301 may determine that the hangover period remains unchanged.

[0156] A person skilled in the art may understand that, the preset hangover period may be set according to an actual status, and the hangover update parameter also may be adjusted according to an actual status. The fifteenth preset value to the nineteenth preset value may be adjusted according to an actual status, so that different hangover periods may be set.

[0157] Similarly, when the general sparseness parameter includes a second minimum bandwidth and a third minimum bandwidth, or the general sparseness parameter includes a first energy proportion, or the general sparseness parameter includes a second energy proportion and a third energy proportion, the processor 301 may set a corresponding preset hangover period, a corresponding hangover update parameter, and a related parameter used to determine the hangover update parameter, so that a corresponding hangover period can be determined, and frequent switching between encoding methods is avoided.

CLAIMS

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

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 FFT energy spectrum coefficients, wherein P is a positive integer; and

determining a general sparseness parameter according to energy of the P FFT energy spectrum coefficients 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;

wherein the general sparseness parameter comprises a first minimum bandwidth;

the determining a general sparseness parameter according to energy of the P FFT energy spectrum coefficients of each of the N audio frames comprises:

determining an average value of minimum bandwidths of distribution, on the spectrums, of first preset proportion energy of the N audio frames according to the energy of the P FFT energy spectrum coefficients of each of the N audio frames, wherein, a minimum bandwidth is found from the P FFT energy spectrum coefficients in a manner that a proportion that energy on the bandwidth accounts for in total energy of a frame is the first preset proportion, wherein the average value of the minimum bandwidths of distribution, 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.

2. The method according to claim 1, wherein the determining an average value of minimum bandwidths of distribution, on the spectrums, of first preset proportion energy of the N audio frames according to the energy of the P FFT energy spectrum coefficients of each of the N audio frames comprises:

sorting the energy of the P FFT energy spectrum coefficients of each audio frame in descending order;

determining, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the first preset proportion of each of the N audio frames; and

determining, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the first preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the first preset proportion of the N audio frames.

3. The method according to claim 2, wherein, the determining the minimum bandwidth comprises:

sequentially accumulating energy of frequency bins in the P FFT energy spectrum coefficients in descending order; and comparing energy obtained after each time of accumulation with the total energy of the audio frame, and if a proportion is greater than the first preset proportion, ending the accumulation process, where a quantity of times of accumulation is the minimum bandwidth.

4. An apparatus, wherein the apparatus comprises:

an obtaining unit, configured to obtain N audio frames, wherein the N audio frames comprise a current audio frame, and N is a positive integer; and

a determining unit, configured to determine sparseness of distribution, on the spectrums, of energy of the N audio frames obtained by the obtaining unit; and

the determining unit is further configured to determine, 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;

the determining unit is specifically configured to divide a spectrum of each of the N audio frames into P FFT energy spectrum coefficients, and determine a general sparseness parameter

according to energy of the P FFT energy spectrum coefficients of each of the N audio frames, wherein P is a positive integer, and the general sparseness parameter indicates the sparseness of distribution, on the spectrums, of the energy of the N audio frames;

wherein the general sparseness parameter comprises a first minimum bandwidth;

the determining unit is specifically configured to determine an average value of minimum bandwidths of distribution, on the spectrums, of first preset proportion energy of the N audio frames according to the energy of the P FFT energy spectrum coefficients of each of the N audio frames, wherein, a minimum bandwidth is found from the P FFT energy spectrum coefficients in a manner that a proportion that energy on the bandwidth accounts for in total energy of a frame is the first preset proportion, wherein the average value of the minimum bandwidths of distribution, on the spectrums, of the first preset proportion energy of the N audio frames is the first minimum bandwidth; and

the determining unit is specifically configured to: when the first minimum bandwidth is less than a first preset value, determine 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, determine to use the second encoding method to encode the current audio frame.

5. The apparatus according to claim 4, wherein the determining unit is specifically configured to:

sort the energy of the P FFT energy spectrum coefficients of each audio frame in descending order;

determine, according to the energy, sorted in descending order, of the P spectral envelopes of each of the N audio frames, a minimum bandwidth of distribution, on the spectrum, of energy that accounts for not less than the first preset proportion of each of the N audio frames; and

determine, according to the minimum bandwidth of distribution, on the spectrum, of the energy that accounts for not less than the first preset proportion of each of the N audio frames, an average value of minimum bandwidths of distribution, on the spectrums, of energy that accounts for not less than the first preset proportion of the N audio frames.

6. The apparatus according to claim 5, wherein, to determine the minimum bandwidth, the determine unit is specifically configured to:

sequentially accumulate energy of frequency bins in the P FFT energy spectrum coefficients in descending order;

compare energy obtained after each time of accumulation with the total energy of the audio frame, and

end the accumulation process if a proportion is greater than the first preset proportion, where a quantity of times of accumulation is the minimum bandwidth.