this embodiment of the present invention imposes no limitation on another algorithm for obtaining the downmixed signal.

[0066] In Embodiment 1 of the present invention, the N^{th}-frame stereo parameter set is encoded, so that a decoder can restore the N^{th}-frame downmixed signal. Optionally, to improve compression efficiency during encoding, the encoder encodes a stereo parameter used for obtaining the N^{th}-frame downmixed signal in the N^{th}-frame stereo parameter set. For example, the generated N^{th}-frame stereo parameter set includes the ITD, the ILD, the IPD, and the IC. If the encoder mixes the N^{th}-frame audio signals on the two channels into the N^{th}-frame downmixed signal according to only the ILD and the IPD in the N^{th}-frame stereo parameter set based on the predetermined first algorithm, to improve the compression efficiency, the encoder may encode only the ILD and the IPD in the N^{th}-frame stereo parameter set.

[0067] Step 102: The encoder detects whether the N^{th}-frame downmixed signal includes a speech signal, and if the N^{th}-frame downmixed signal includes the speech signal, performs step 103, or if the N^{th}-frame downmixed signal does not include the speech signal, performs step 104.

[0068] For ease of detecting, by the encoder, whether the N^{th}-frame downmixed signal includes the speech signal, optionally, the encoder directly detects, by means of voice activity detection (Voice Activity Detection, VAD), whether the N^{th}-frame downmixed signal includes the speech signal.

[0069] Optionally, a method for indirectly detecting, by the encoder, whether the N^{th}-frame downmixed signal includes the speech signal is: The encoder directly detects, by means of VAD, whether the N^{th}-frame audio signals include the speech signal. Specifically, if detecting that an audio signal on one of the two channels includes the speech signal, the encoder determines that a downmixed signal obtained by mixing audio signals on the two channels includes the speech signal. Only when determining that neither of the audio signals on the two channels includes the speech signal, the encoder determines that the downmixed signal obtained by mixing the audio signals on the two channels does not include the speech signal. It should be noted that in such an indirect detection manner, a sequence between step 102 and step 100 or step 101 is not limited, provided that step 100 precedes step 101.

[0070] Step 103: The encoder encodes the N^{th}-frame downmixed signal, and performs step 107.

[0071] The encoder encodes the N^{th}-frame downmixed signal to obtain an N^{th}-frame bitstream.

[0072] Because discontinuous encoding is performed on the downmixed signal in Embodiment 1 of the present invention, a bitstream includes two frame types: a first-type frame and a second-type frame. The first-type frame includes a downmixed signal, and the second-type frame does not include a downmixed signal. The N^{th}-frame bitstream obtained in step 103 is the first-type

used for the encoding and a parameter used for the encoding.

[0084] The preset speech frame encoding condition may be: duration between the N^{th}-frame downmixed signal and an M^{th}-frame downmixed signal is not greater than preset duration. The M^{th}-frame downmixed signal includes the speech signal, and the M^{th}-frame downmixed signal is a frame of downmixed signal that includes the speech signal and that is closest to the N^{th}-frame downmixed signal. The preset SID encoding condition may be encoding an odd-number frame. When N of the N^{th}-frame downmixed signal is an odd number, the encoder determines that the N^{th}-frame downmixed signal satisfies the preset SID encoding condition.

[0085] Step 106: The encoder skips encoding the N^{th}-frame downmixed signal, and performs step 109.

[0086] Specifically, the N^{th}-frame bitstream obtained in step 106 is the second-type frame.

[0087] The encoder determines that the N^{th}-frame downmixed signal does not satisfy the preset audio frame encoding condition. Specifically, the encoder determines that the N^{th}-frame downmixed signal does not satisfy the preset speech frame encoding condition, and does not satisfy the preset SID encoding condition.

[0088] In this embodiment of the present invention, the encoder does not encode the N^{th}-frame downmixed signal. Specifically, the N^{th}-frame bitstream does not include the N^{th}-frame downmixed signal.

[0089] When the encoder does not encode the N^{th}-frame downmixed signal, the encoder may encode the N^{th}-frame stereo parameter set, or may not encode the N^{th}-frame stereo parameter set.

[0090] In Embodiment 1 of the present invention, a description is made by using an example in which the encoder does not encode the N^{th}-frame downmixed signal, but encodes the N^{th}-frame stereo parameter set. However, optionally, when the encoder does not encode the N^{th}-frame downmixed signal, the encoder may not encode the N^{th}-frame stereo parameter set either. Specifically, when the encoder encodes neither the N^{th}-frame stereo parameter set nor the N^{th}-frame downmixed signal, for a manner of obtaining the N^{th}-frame downmixed signal and the N^{th}-frame stereo parameter set by the decoder, refer to Embodiment 2 of the present invention.

[0091] Step 107: The encoder sends an N^{th}-frame bitstream to a decoder.

[0092] In order that the decoder can restore the N^{th}-frame downmixed signal to the N^{th}-frame audio signals on the two channels after obtaining, by means of decoding, the N^{th}-frame downmixed signal, the N^{th}-frame bitstream includes both the N^{th}-frame stereo parameter set and the N^{th}-frame downmixed signal.

[0093] Step 108: If determining that the N^{th}-frame bitstream is a first-type frame, the decoder decodes the N^{th}-frame bitstream, to obtain the N^{th}-frame downmixed signal and the N^{th}-frame stereo

ILD(9), five IPDs are selected from the IPD(0), …, and the IPD(9), one ITD is selected from the ITD(0) and the ITD(1), and the selected parameters are encoded. Alternatively, the preset stereo parameter dimension reduction rule is that five ILDs and five IPDs are selected. Alternatively, if the preset stereo parameter dimension reduction rule is that frequency-domain resolution of the ILDs, frequency-domain resolution of the IPDs, and time-domain resolution of the ITDs are reduced, ILDs in neighboring sub frequency bands in the ILD(0), …, and the ILD(9) are combined. For example, an average value of the ILD(0) and the ILD(1) is calculated to obtain a new ILD(0), an average value of the ILD(2) and the ILD(3) is calculated to obtain a new ILD(1), …, and an average value of the ILD(8) and the ILD(9) is calculated to obtain a new ILD(4). A sub frequency band corresponding to the new ILD(0) is obtained by combining sub frequency bands corresponding to the original ILD(0) and the original ILD(1), …, and a sub frequency band corresponding to the new ILD(4) is obtained by combining corresponding to the original ILD(8) and the original ILD(9). According to the same method, IPDs in neighboring sub frequency bands in the IPD(0), …, and the IPD(9) are combined, to obtain a new IPD(0), …, and a new IPD(4); and an average value of the ITD(0) and the ITD(1) is also calculated to obtain a new ITD(0). A time-domain signal corresponding to the new ITD(0) is obtained by combining corresponding to the original ITD(0) and the original ITD(1). The new ILD(0), …, and the new ILD(4), the new IPD(0), …, and the new IPD(4), and the new ITD(0) are encoded. Alternatively, if the preset stereo parameter dimension reduction rule is that frequency-domain resolution of the ILDs is reduced, ILDs in neighboring sub frequency bands in the ILD(0), …, and the ILD(9) are combined. For example, an average value of the ILD(0) and the ILD(1) is calculated to obtain a new ILD(0), an average value of the ILD(2) and the ILD(3) is calculated to obtain a new ILD(1), …, and an average value of the ILD(8) and the ILD(9) is calculated to obtain a new ILD(4). A sub frequency band corresponding to the new ILD(0) is obtained by combining corresponding to the original ILD(0) and the original ILD(1), …, and a sub frequency band corresponding to the new ILD(4) is obtained by combining corresponding to the original ILD(8) and the original ILD(9). Then, the new ILD(0), …, and the new ILD(4) are encoded.

[0134] Step 208: The encoder encodes the N^{th}-frame downmixed signal according to a preset SID encoding rate, but skips encoding the at least one stereo parameter in the N^{th}-frame stereo parameter set, and performs step 211.

[0135] Step 209: The encoder encodes the at least one stereo parameter in the N^{th}-frame stereo parameter set, but skips encoding the N^{th}-frame downmixed signal, and performs step 215.

[0136] Step 210: The encoder encodes neither the N^{th}-frame downmixed signal nor the

N^{th}-frame stereo parameter set, and performs step 217.

[0137] In Embodiment 2 of the present invention, the encoder performs encoding to obtain a bitstream. The bitstream includes four different types of frames, that is, a third-type frame, a fourth-type frame, a fifth-type frame, and a sixth-type frame. The third-type frame includes a stereo parameter set, but does not include a downmixed signal, the fourth-type frame includes neither a downmixed signal nor a stereo parameter set, the fifth-type frame includes both a downmixed signal and a stereo parameter set, and the sixth-type frame includes a downmixed signal, but does not include a stereo parameter set. Each of the fifth-type frame and the sixth-type frame is one case of a type frame including a downmixed signal, and each of the third-type frame and the fourth-type frame is one case of a type frame including no downmixed signal.

[0138] Specifically, an N^{th}-frame bitstream obtained in step 203, step 205, or step 207 is the fifth-type frame, an N^{th}-frame bitstream obtained in step 208 is the sixth-type frame, an N^{th}-frame bitstream obtained in step 209 is the third-type frame, and an N^{th}-frame bitstream obtained in step 211 is the fourth-type frame.

[0139] Step 211: The encoder sends an N^{th}-frame bitstream to a decoder, where the N^{th}-frame bitstream includes the N^{th}-frame downmixed signal and the N^{th}-frame stereo parameter set.

[0140] Step 212: The decoder receives the N^{th}-frame bitstream, decodes the N^{th}-frame bitstream if determining that the N^{th}-frame bitstream is a fifth-type frame, to obtain the N^{th}-frame downmixed signal and the N^{th}-frame stereo parameter set, and performs step 218.

[0141] For a specific implementation of determining, by the decoder, which type frame the N^{th}-frame bitstream is, refer to Embodiment 1 of the present invention.

[0142] Specifically, the decoder decodes the N^{th}-frame bitstream according to a rate corresponding to the N^{th}-frame bitstream. Specifically, if the encoder encodes the N^{th}-frame downmixed signal according to 13.2 kbps, the decoder decodes a bitstream of the N^{th}-frame downmixed signal in the N^{th}-frame bitstream according to 13.2 kbps. If the encoder encodes the N^{th}-frame stereo parameter set according to 4.2 kbps, the decoder decodes a bitstream of the N^{th}-frame stereo parameter set in the N^{th}-frame bitstream according to 4.2 kbps.

[0143] Step 213: The encoder sends an N^{th}-frame bitstream to a decoder, where the N^{th}-frame bitstream includes the N^{th}-frame downmixed signal.

[0144] Step 214: The decoder decodes the N^{th}-frame bitstream if determining that the N^{th}-frame bitstream is a sixth-type frame, to obtain the N^{th}-frame downmixed signal, determines, according to a preset second rule, k-frame stereo parameter sets in at least one-frame stereo parameter set preceding an N^{th}-frame stereo parameter set, obtains the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on a predetermined sixth algorithm, and performs step 218.

[0145] Specifically, using a stereo parameter in the N^{th}-frame stereo parameter set as an example, a stereo parameter set stipulated in the preset second rule is a frame of stereo parameter set that is closest to P and that is obtained by means of decoding, and an N^{th}-frame stereo parameter P is obtained according to the following algorithm:

$$P = \tilde{P}^{[-1]} + \delta ,$$

where P represents the N^{th}-frame stereo parameter, $\tilde{P}^{[-1]}$ represents a frame of stereo parameter that is closest to P and that is obtained by means of decoding, and δ represents a random number whose absolute value is relatively small. For example, δ may be a random number between $-\tilde{P}^{[-1]} \times 5\%$ and $+\tilde{P}^{[-1]} \times 5\%$.

[0146] It should be noted that this embodiment of the present invention imposes no limitation on the method for estimating stereo parameters in the N^{th}-frame stereo parameter set.

[0147] Step 215: The encoder sends an N^{th}-frame bitstream to a decoder, where the N^{th}-frame bitstream includes the at least one stereo parameter in the N^{th}-frame stereo parameter set.

[0148] Step 216: The decoder decodes the N^{th}-frame bitstream if determining that the N^{th}-frame bitstream is a third-type frame, to obtain the at least one stereo parameter in the N^{th}-frame stereo parameter set, determines, according to a preset first rule, m-frame downmixed signals in at least one-frame downmixed signal preceding the N^{th}-frame downmixed signal, obtains the N^{th}-frame downmixed signal according to the m-frame downmixed signals based on a predetermined second algorithm, where m is a positive integer greater than 0, and performs step 218.

[0149] Specifically, an average value of an $(N-3)^{th}$-frame downmixed signal, an $(N-2)^{th}$-frame downmixed signal, and an $(N-1)^{th}$-frame downmixed signal is used as the N^{th}-frame downmixed signal, or an $(N-1)^{th}$-frame downmixed signal is directly used as the N^{th}-frame downmixed signal, or the N^{th}-frame downmixed signal is estimated according to another algorithm.

[0150] In addition, the $(N-1)^{th}$-frame downmixed signal may be directly used as the N^{th}-frame downmixed signal, or the N^{th}-frame downmixed signal is calculated according to the $(N-1)^{th}$-frame downmixed signal and a preset offset value based on a preset algorithm.

[0151] Step 217: After receiving an N^{th}-frame bitstream, a decoder determines that the N^{th}-frame bitstream is a fourth-type frame, determines, according to a preset second rule, k-frame stereo parameter sets in at least one-frame stereo parameter set preceding an N^{th}-frame stereo parameter set, and obtains the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on a predetermined sixth algorithm; and

[0152] determines, according to a preset first rule, m-frame downmixed signals in at least one-frame downmixed signal preceding the N^{th}-frame downmixed signal, and obtains the N^{th}-frame downmixed signal according to the m-frame downmixed signals based on a predetermined second

that is of a stereo parameter and that is stipulated in the first stereo parameter set generation manner is not lower than frequency-domain resolution that is of a corresponding stereo parameter and that is stipulated in the second stereo parameter set generation manner.

[0158] Specifically, frequency-domain precision or time-domain precision of a stereo parameter set obtained in the first stereo parameter set generation manner is higher than that of a stereo parameter set obtained in the second stereo parameter set generation manner.

[0159] In addition, in a multichannel audio signal processing method in Embodiment 3 of the present invention, when detecting that an N^{th}-frame downmixed signal includes a speech signal, an encoder encodes the N^{th}-frame downmixed signal according to a speech encoding rate, and encodes an N^{th}-frame stereo parameter set; or when an encoder detects that an N^{th}-frame downmixed signal does not include a speech signal: if the N^{th}-frame downmixed signal satisfies a preset speech frame encoding condition, the encoder encodes the N^{th}-frame downmixed signal according to a speech encoding rate, and encodes an N^{th}-frame stereo parameter set, or if the N^{th}-frame downmixed signal does not satisfy a preset speech frame encoding condition, but satisfies a preset SID encoding condition, the encoder encodes the N^{th}-frame downmixed signal according to an SID encoding rate, and encodes at least one stereo parameter in an N^{th}-frame stereo parameter set, or if the N^{th}-frame downmixed signal satisfies neither a preset speech frame encoding condition nor a preset SID encoding condition, the encoder encodes neither the N^{th}-frame downmixed signal nor an N^{th}-frame stereo parameter set.

[0160] It should be understood that a difference between Embodiment 3 of the present invention and Embodiment 1 of the present invention or between Embodiment 3 of the present invention and Embodiment 2 of the present invention lies in: The encoder does not perform determining on a stereo parameter set, and encodes the stereo parameter set regardless of which manner is used to encode a downmixed signal.

[0161] In Embodiment 3 of the present invention, a bitstream obtained after the encoder encodes the downmixed signal includes two types of frames: a first-type frame and a second-type frame. The first-type frame includes both a downmixed signal and a stereo parameter set, and the second-type frame includes neither a downmixed signal nor a stereo parameter set. Specifically, for a method for restoring the bitstream to audio signals on two channels by a decoder after receiving the bitstream, refer to Embodiment 2 of the present invention and Embodiment 1 of the present invention.

[0162] Based on Embodiment 3 of the present invention, optionally, when the N^{th}-frame downmixed signal satisfies neither the preset speech frame encoding condition nor the preset SID encoding condition, the encoder determines whether the N^{th}-frame stereo parameter set satisfies a

preset stereo parameter encoding condition, and if the N^{th}-frame stereo parameter set satisfies the preset stereo parameter encoding condition, the encoder does not encode the N^{th}-frame downmixed signal, but encodes at least one stereo parameter in the N^{th}-frame stereo parameter set, or if the N^{th}-frame stereo parameter set does not satisfy the preset stereo parameter encoding condition, the encoder encodes neither the N^{th}-frame downmixed signal nor the N^{th}-frame stereo parameter set.

[0163] A bitstream obtained based on the foregoing encoding method includes three types of frames: a first-type frame, a third-type frame, and a fourth-type frame. The first-type frame includes both a downmixed signal and a stereo parameter set, the third-type frame does not include a downmixed signal, but includes a stereo parameter set, and the fourth-type frame includes neither a downmixed signal nor a stereo parameter set. Specifically, for a method for restoring the bitstream to audio signals on two channels by a decoder after receiving the bitstream, refer to Embodiment 2 of the present invention and Embodiment 1 of the present invention.

[0164] A difference between the foregoing technical solution and Embodiment 2 of the present invention lies in: When the N^{th}-frame downmixed signal satisfies neither the preset speech frame encoding condition nor the preset SID encoding condition, the encoder determines whether the N^{th}-frame stereo parameter set satisfies the preset stereo parameter encoding condition.

[0165] Optionally, in a multichannel audio signal processing method in Embodiment 4 of the present invention, when detecting that an N^{th}-frame downmixed signal includes a speech signal, an encoder encodes the N^{th}-frame downmixed signal according to a speech encoding rate, and encodes an N^{th}-frame stereo parameter set; or when an encoder detects that an N^{th}-frame downmixed signal does not include a speech signal: if the N^{th}-frame downmixed signal satisfies a preset speech frame encoding condition, the encoder encodes the N^{th}-frame downmixed signal according to a speech encoding rate, and encodes an N^{th}-frame stereo parameter set, or if the N^{th}-frame downmixed signal does not satisfy a preset speech frame encoding condition, but satisfies a preset SID encoding condition, the encoder determines whether an N^{th}-frame stereo parameter set satisfies a preset stereo parameter encoding condition, and when the N^{th}-frame stereo parameter set satisfies the preset stereo parameter encoding condition, the encoder encodes the N^{th}-frame downmixed signal according to an SID encoding rate, and encodes at least one stereo parameter in the N^{th}-frame stereo parameter set, or when the N^{th}-frame stereo parameter set does not satisfy a preset stereo parameter encoding condition, the encoder encodes the N^{th}-frame downmixed signal according to an SID encoding rate, but does not encode the N^{th}-frame stereo parameter set; or if the N^{th}-frame downmixed signal satisfies neither a preset speech frame encoding condition nor a preset SID encoding condition, the encoder encodes neither the N^{th}-frame downmixed signal nor an N^{th}-frame stereo parameter set.

parameter detection unit 340 determines that the N^{th}-frame stereo parameter set satisfies a preset stereo parameter encoding condition; or skip encoding the stereo parameter set if the parameter detection unit 340 determines that the N^{th}-frame stereo parameter set does not satisfy a preset stereo parameter encoding condition.

[0177] Optionally, the parameter encoding unit 330 is configured to: obtain X target stereo parameters according to the Z stereo parameters in the N^{th}-frame stereo parameter set based on a preset stereo parameter dimension reduction rule, and encode the X target stereo parameters. X is a positive integer greater than 0 and less than or equal to Z.

[0178] Specifically, when the parameter encoding unit 330 includes a first parameter encoding unit 331 and a second parameter encoding unit 332, the second parameter encoding unit 332 is configured to: obtain the X target stereo parameters according to the Z stereo parameters in the N^{th}-frame stereo parameter set based on the preset stereo parameter dimension reduction rule, and encode the X target stereo parameters.

[0179] Optionally, based on FIG. 3a and FIG. 3b, as shown in FIG. 3c, the parameter generation unit 320 of the encoder includes a first parameter generation unit 321 and a second parameter generation unit 322. When the signal detection unit 300 detects that the N^{th}-frame audio signals include the speech signal, or the signal detection unit 300 detects that the N^{th}-frame audio signals do not include the speech signal and the N^{th}-frame audio signals satisfy the preset speech frame encoding condition, the signal detection unit 300 instructs the first parameter generation unit 321 to generate the N^{th}-frame stereo parameter set. When the signal detection unit 300 detects that the N^{th}-frame audio signals do not include the speech signal, and the N^{th}-frame audio signals do not satisfy the preset speech frame encoding condition, the signal detection unit 300 instructs the second parameter generation unit 322 to generate the N^{th}-frame stereo parameter set. Specifically, it is pre-stipulated that the first parameter generation unit 321 obtains the N^{th}-frame stereo parameter set according to the N^{th}-frame audio signals based on a first stereo parameter set generation manner, and the second parameter generation unit 322 obtains the N^{th}-frame stereo parameter set according to the N^{th}-frame audio signals based on a second stereo parameter set generation manner.

[0180] The first stereo parameter set generation manner and the second stereo parameter set generation manner satisfy at least one of the following conditions:

[0181] A quantity that is of types of stereo parameters included in a stereo parameter set and that is stipulated in the first stereo parameter set generation manner is not less than a quantity that is of types of stereo parameters included in a stereo parameter set and that is stipulated in the second stereo parameter set generation manner, a quantity that is of stereo parameters included in a stereo parameter set and that is stipulated in the first stereo parameter set generation

shown in FIG. 4.

[0216] Persons skilled in the art should understand that the embodiments of the present invention may be provided as a method, a system, or a computer program product. Therefore, the present invention may use a form of hardware only embodiments, software only embodiments, or embodiments with a combination of software and hardware. Moreover, the present invention may use a form of a computer program product that is implemented on one or more computer-usable storage media (including but not limited to a disk memory, a CD-ROM, an optical memory, and the like) that include computer-usable program code.

[0217] The present invention is described with reference to the flowcharts and/or block diagrams of the method, the device (system), and the computer program product according to the embodiments of the present invention. It should be understood that computer program instructions may be used to implement each process and/or each block in the flowcharts and/or the block diagrams and implement a combination of a process and/or a block in the flowcharts and/or the block diagrams. These computer program instructions may be provided for a general-purpose computer, a dedicated computer, an embedded processor, or a processor of another programmable data processing device to generate a machine, so that the instructions executed by the computer or the processor of the another programmable data processing device generate an apparatus for implementing a specific function in one or more processes in the flowcharts and/or in one or more blocks in the block diagrams.

[0218] These computer program instructions may be stored in a computer readable memory that can instruct the computer or the another programmable data processing device to work in a specific manner, so that the instructions stored in the computer readable memory generate an artifact that includes an instruction apparatus. The instruction apparatus implements a specific function in one or more processes in the flowcharts and/or in one or more blocks in the block diagrams.

[0219] These computer program instructions may be loaded onto the computer or the another programmable data processing device, so that a series of operations and steps are performed on the computer or the another programmable device, to generate computer-implemented processing. Therefore, the instructions executed on the computer or the another programmable device provide steps for implementing a specific function in one or more processes in the flowcharts and/or in one or more blocks in the block diagrams.

[0220] Although some embodiments of the present invention have been described, persons skilled in the art can make changes and modifications to these embodiments once they learn the basic inventive concept. Therefore, the following claims are intended to be construed as to cover the embodiments and all changes and modifications falling within the scope of the present invention.

[0221] Obviously, persons skilled in the art can make various modifications and variations to the present invention without departing from the scope of the present invention. The present invention is intended to cover these modifications and variations provided that they fall within the scope of protection defined by the following claims and their equivalent technologies.

CLAIMS

1. A multichannel audio signal processing method implemented by an encoder, comprising:

mixing N^{th}-frame audio signals on two of a plurality of channels based on a first algorithm to obtain an N^{th}-frame downmixed signal;

detecting whether the N^{th}-frame downmixed signal comprises a speech signal, wherein N is a positive integer greater than zero;

encoding the N^{th}-frame downmixed signal when detecting that the N^{th}-frame downmixed signal comprises the speech signal;

encoding the N^{th}-frame downmixed signal when the encoder detects that the N^{th}-frame downmixed signal does not comprise the speech signal and when determining that the N^{th}-frame downmixed signal satisfies a preset audio frame encoding condition; and

skipping the N^{th}-frame downmixed signal when determining that the N^{th}-frame downmixed signal does not satisfy the preset audio frame encoding condition.

2. The multichannel audio signal processing method of claim 1, wherein encoding the N^{th}-frame downmixed signal comprises:

encoding the N^{th}-frame downmixed signal according to a preset speech frame encoding rate when detecting that the N^{th}-frame downmixed signal comprises the speech signal;

encoding the N^{th}-frame downmixed signal according to the preset speech frame encoding rate when determining that the N^{th}-frame downmixed signal satisfies a preset speech frame encoding condition; and

encoding the N^{th}-frame downmixed signal according to a preset silence insertion descriptor (SID) frame encoding rate when determining that the N^{th}-frame downmixed signal does not satisfy the preset speech frame encoding condition and satisfies a preset SID encoding condition, wherein the preset SID frame encoding rate is less than or equal to the preset speech frame encoding rate.

3. The multichannel audio signal processing method of claim 2, further comprising:

detecting that the N^{th}-frame audio signals comprise the speech signal;

obtaining an N^{th}-frame stereo parameter set according to the N^{th}-frame audio signals based on a first stereo parameter set generation manner, and encoding the N^{th}-frame stereo parameter set when detecting that the N^{th}-frame audio signals comprise the speech signal;

determining that the N^{th}-frame audio signals satisfy the preset speech frame encoding condition;

obtaining the N^{th}-frame stereo parameter set according to the N^{th}-frame audio signals based on the first stereo parameter set generation manner, and encoding the N^{th}-frame stereo parameter set when detecting that the N^{th}-frame audio signals do not comprise the speech signal and when determining that the N^{th}-frame audio signals satisfy the preset speech frame encoding condition;

obtaining the N^{th}-frame stereo parameter set according to the N^{th}-frame audio signals based on a second stereo parameter set generation manner when detecting that the N^{th}-frame audio signals do not comprise the speech signal and when determining that the N^{th}-frame audio signals do not satisfy the preset speech frame encoding condition;

encoding at least one stereo parameter in the N^{th}-frame stereo parameter set when determining that the N^{th}-frame stereo parameter set satisfies a preset stereo parameter encoding condition; and

skipping encoding the stereo parameter set when determining that the N^{th}-frame stereo parameter set does not satisfy the preset stereo parameter encoding condition,

wherein the first stereo parameter set generation manner and the second stereo parameter set generation manner satisfy at least one of the following conditions:

a quantity of types of stereo parameters comprised in a stereo parameter set stipulated in the first stereo parameter set generation manner is not less than a quantity of types of stereo parameters comprised in a stereo parameter set stipulated in the second stereo parameter set generation manner;

a quantity of stereo parameters comprised in the stereo parameter set stipulated in the first stereo parameter set generation manner is not less than a quantity of stereo parameters comprised in the stereo parameter set stipulated in the second stereo parameter set generation manner;

a time-domain resolution of a stereo parameter stipulated in the first stereo parameter set generation manner is higher than or equal to a time-domain resolution of a corresponding stereo parameter stipulated in the second stereo parameter set generation manner; or

a frequency-domain resolution of the stereo parameter stipulated in the first stereo parameter set generation manner is higher than or equal to a frequency-domain resolution of the corresponding stereo parameter is stipulated in the second stereo parameter set generation manner.

4. The multichannel audio signal processing method of claim 1, further comprising:

obtaining an N^{th}-frame stereo parameter set according to the N^{th}-frame audio signals, wherein the N^{th}-frame stereo parameter set comprises Z stereo parameters, wherein the Z stereo parameters comprise a parameter used to mix the N^{th}-frame audio signals, and wherein Z is a positive integer greater than zero;

encoding the N^{th}-frame stereo parameter set when detecting that the N^{th}-frame downmixed signal comprises the speech signal;

determining that the N^{th}-frame stereo parameter set satisfies a preset stereo parameter encoding condition;

encoding at least one stereo parameter in the N^{th}-frame stereo parameter set when detecting that the N^{th}-frame downmixed signal does not comprise the speech signal and when determining that the N^{th}-frame stereo parameter set satisfies the preset stereo parameter encoding condition; and

skipping encoding the stereo parameter set when detecting that the N^{th}-frame downmixed signal does not comprise the speech signal and when determining that the N^{th}-frame stereo parameter set does not satisfy the preset stereo parameter encoding condition.

5. The multichannel audio signal processing method of claim 4, wherein encoding the at least one stereo parameter in the N^{th}-frame stereo parameter set comprises:

obtaining X target stereo parameters according to the Z stereo parameters in the N^{th}-frame stereo parameter set based on a preset stereo parameter dimension reduction rule, wherein X is a positive integer greater than zero and less than or equal to Z; and

encoding the X target stereo parameters.

6. The multichannel audio signal processing method of claim 4, wherein encoding the N^{th}-frame stereo parameter set comprises encoding the N^{th}-frame stereo parameter set according to a first encoding manner, and wherein encoding the at least one stereo parameter in the N^{th}-frame stereo parameter set comprises:

encoding the at least one stereo parameter in the N^{th}-frame stereo parameter set according to the first encoding manner when the N^{th}-frame downmixed signal satisfies the preset audio frame encoding condition; and

encoding the at least one stereo parameter in the N^{th}-frame stereo parameter set according to a second encoding manner when the N^{th}-frame downmixed signal does not satisfy the preset audio frame encoding condition,

wherein an encoding rate stipulated in the first encoding manner is greater than or equal to an encoding rate stipulated in the second encoding manner, or wherein a quantization precision stipulated in the first encoding manner is higher than or equal to a quantization precision stipulated in the second encoding manner for any stereo parameter in the N^{th}-frame stereo parameter set.

7. The multichannel audio signal processing method of claim 4, further comprising:

determining that the at least one stereo parameter in the N^{th}-frame stereo parameter set comprises an inter-channel level difference (ILD), wherein the preset stereo parameter encoding condition comprises $D_L \geq D_0$ when determining that the at least one stereo parameter in the N^{th}-frame stereo parameter set comprises the ILD, wherein D_L represents a degree by which the ILD deviates from a first standard, wherein the first standard is determined based on a second algorithm according to T-frame stereo parameter sets preceding the N^{th}-frame stereo parameter set, and wherein T is a positive integer greater than zero;

determining that the at least one stereo parameter in the N^{th}-frame stereo parameter set comprises an inter-channel time difference (ITD), wherein the preset stereo parameter encoding condition comprises $D_T \geq D_1$ when determining that the at least one stereo parameter in the N^{th}-frame stereo parameter set comprises the ITD, wherein D_T represents a degree by which the ITD deviates from a second standard, and wherein the second standard is determined based on a third algorithm according to the T-frame stereo parameter sets preceding the N^{th}-frame stereo parameter set; and

determining that the at least one stereo parameter in the N^{th}-frame stereo parameter set comprises an inter-channel phase difference (IPD), wherein the preset stereo parameter encoding condition comprises $D_P \geq D_2$ when determining that the at least one stereo parameter in the N^{th}-frame stereo parameter set comprises the IPD, wherein D_P represents a degree by which the IPD deviates from a third standard, and wherein the third standard is determined based on a fourth algorithm according to the T-frame stereo parameter sets preceding the N^{th}-frame stereo parameter set.

8. The multichannel audio signal processing method of claim 7, wherein D_L, D_T, and D_P respectively satisfy the following expressions:

$$D_L = \sum_{m=0}^{M-1}\left(ILD(m) - \frac{1}{T}\sum_{t=1}^{T} ILD^{[-t]}(m)\right);$$

$$D_T = ITD - \frac{1}{T}\sum_{t=1}^{T} ITD^{[-t]}(m); \text{ and}$$

$$D_P \sum_{m=0}^{M-1}\left(IPD(m) - \frac{1}{T}\sum_{t=1}^{T} IPD^{[-t]}(m)\right),$$

wherein ILD(m) is a first level difference generated when the N^{th}-frame audio signals are respectively transmitted on two channels in an m^{th} sub frequency band, wherein M is a total quantity of sub frequency bands occupied for transmitting the N^{th}-frame audio signals, wherein

$$\frac{1}{T}\sum_{t=1}^{T} ILD^{[-t]}(m)$$

is an average value of ILDs in the T-frame stereo parameter sets preceding the N^{th}-frame stereo parameter set in the m^{th} sub frequency band, wherein $ILD^{[-t]}(m)$ is a second level difference generated when t^{th}-frame audio signals preceding the N^{th}-frame audio signals are respectively transmitted on the two channels in the m^{th} sub frequency band, wherein the ITD is a first time difference generated when the N^{th}-frame audio signals are respectively transmitted on the two channels, wherein

$$\frac{1}{T}\sum_{t=1}^{T} ITD^{[-t]}$$

is an average value of ITDs in the T-frame stereo parameter sets preceding the N^{th}-frame stereo parameter set, wherein $ITD^{[-t]}$ is a second time difference generated when the t^{th}-frame audio signals preceding the N^{th}-frame audio signals are respectively transmitted on the two channels, wherein IPD(m) is a first phase difference generated when some of the N^{th}-frame audio signals are respectively transmitted on the two channels in the m^{th} sub frequency band, wherein

$$\frac{1}{T}\sum_{t=1}^{T} IPD^{[-t]}(m)$$

is an average value of IPDs in the T-frame stereo parameter sets preceding the N^{th}-frame stereo parameter set in the m^{th} sub frequency band, and wherein $IPDI^{[-t]}(m)$ is a second phase difference generated when the t^{th}-frame audio signals preceding the N^{th}-frame audio signals are respectively transmitted on the two channels in the m^{th} sub frequency band.

9. A multichannel audio signal processing method implemented by a decoder, comprising:

receiving a bitstream, wherein the bitstream comprises at least two frames, wherein the at least two frames comprise at least one first-type frame or at least one second-type frame, wherein the first-type frame comprises a downmixed signal, and wherein the second-type frame does not comprise the downmixed signal;

decoding the N^{th}-frame bitstream when determining that the N^{th}-frame bitstream is the first-type frame to obtain an N^{th}-frame downmixed signal;

according to a preset first rule, m-frame downmixed signals in at least one-frame downmixed signal preceding the N^{th}-frame downmixed signal, and obtaining the N^{th}-frame downmixed signal according to the m-frame downmixed signals based on a first algorithm when determining that the N^{th}-frame bitstream is the second-type frame, wherein m is a positive integer greater than zero, wherein N is a positive integer greater than one, and wherein the N^{th}-frame downmixed signal is received from an encoder after mixing N^{th}-frame audio signals on two of a plurality of channels based on a second algorithm.

10. The multichannel audio signal processing method of claim 9, wherein the first-type frame comprises the downmixed signal and a stereo parameter set, wherein the second-type frame comprises the stereo parameter set and does not comprise the downmixed signal, and wherein the multichannel audio signal processing method further comprises:

obtaining an N^{th}-frame stereo parameter set after decoding the N^{th}-frame bitstream when determining that the N^{th}-frame bitstream is the first-type frame;

decoding the N^{th}-frame bitstream to obtain the N^{th}-frame stereo parameter set when determining that the N^{th}-frame bitstream is the second-type frame; and

restoring the N^{th}-frame downmixed signal to the N^{th}-frame audio signals according t at least one stereo parameter in the N^{th}-frame stereo parameter set based on a third algorithm.

11. The multichannel audio signal processing method of claim 9, wherein the first-type frame comprises the downmixed signal and a stereo parameter set, wherein the second-type frame comprises neither the downmixed signal nor the stereo parameter set, and wherein the multichannel audio signal processing method further comprises:

obtaining an N^{th}-frame stereo parameter set after decoding the N^{th}-frame bitstream when determining that the N^{th}-frame bitstream is the first-type frame;

determining, according to a preset second rule, k-frame stereo parameter sets in at least one-frame stereo parameter set preceding the N^{th}-frame stereo parameter set, and obtaining the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on a fourth algorithm after determining that the N^{th}-frame bitstream is the second-type frame, wherein k is a positive integer greater than zero; and

restoring the N^{th}-frame downmixed signal to the N^{th}-frame audio signals according to at least one stereo parameter in the N^{th}-frame stereo parameter set based on third algorithm.

12. The multichannel audio signal processing method of claim 9, wherein the first-type frame comprises the downmixed signal and a stereo parameter set, wherein a third-type frame comprises the stereo parameter set and does not comprise the downmixed signal, wherein a fourth-type frame comprises neither the downmixed signal nor the stereo parameter set, wherein each of the third-type frame and the fourth-type frame is one case of the second-type frame, and wherein the multichannel audio signal processing method further comprises:

obtaining an N^{th}-frame stereo parameter set after decoding the N^{th}-frame bitstream when determining that the N^{th}-frame bitstream is the first-type frame;

decoding the N^{th}-frame bitstream to obtain the N^{th}-frame stereo parameter set when determining that the N^{th}-frame bitstream is the third-type frame;

determining, according to a preset second rule, k-frame stereo parameter sets in at least one-frame stereo parameter set preceding the N^{th}-frame stereo parameter set, and obtaining the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on a fourth algorithm when determining that the N^{th}-frame bitstream is the fourth-type frame, wherein k is a positive integer greater than zero; and

restoring the N^{th}-frame downmixed signal to the N^{th}-frame audio signals according to at least one stereo parameter in the N^{th}-frame stereo parameter set based on a third algorithm.

13. The multichannel audio signal processing method of claim 9, wherein a fifth-type frame comprises the downmixed signal and a stereo parameter set, wherein a sixth-type frame comprises the downmixed signal and does not comprise the stereo parameter set, wherein each of the fifth-type frame and the sixth-type frame is one case of the first-type frame, wherein the second-type frame comprises neither the downmixed signal nor the stereo parameter set, and wherein the multichannel audio signal processing method further comprises:

decoding the N^{th}-frame bitstream to obtain an N^{th}-frame stereo parameter set when determining that the N^{th}-frame bitstream is the fifth-type frame;

determining, according to a preset second rule, k-frame stereo parameter sets in at least one-frame stereo parameter set preceding the N^{th}-frame stereo parameter set, and obtaining the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on a fourth algorithm when determining that the N^{th}-frame bitstream is the sixth-type frame;

determining, according to the preset second rule, the k-frame stereo parameter sets in the at least one-frame stereo parameter set preceding the N^{th}-frame stereo parameter set, and obtaining the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on the fourth algorithm after determining that the N^{th}-frame bitstream is the second-type frame, wherein is a positive integer greater than zero; and

restoring the N^{th}-frame downmixed signal to the N^{th}-frame audio signals according to at least one stereo parameter in the N^{th}-frame stereo parameter set based on third algorithm.

14. The multichannel audio signal processing method of claim 9, wherein a fifth-type frame comprises the downmixed signal and a stereo parameter set, wherein a sixth-type frame comprises the downmixed signal and does not comprise the stereo parameter set, wherein each of the fifth-type frame and the sixth-type frame is one case of the first-type frame, wherein a third-type frame comprises the stereo parameter set and does not comprise the downmixed signal, wherein a fourth-type frame comprises neither the downmixed signal nor the stereo parameter set, wherein each of the third-type frame and the fourth-type frame is one case of the second-type frame, and wherein the multichannel audio signal processing method further comprises:

decoding the N^{th}-frame bitstream to obtain an Nm-frame stereo parameter set when determining that the N^{th}-frame bitstream is the fifth-type frame;

determining, according to a preset second rule, k-frame stereo parameter sets in at least one-frame stereo parameter set preceding the N^{th}-frame stereo parameter set, and obtaining the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on a fourth algorithm when determining that the N^{th}-frame bitstream is the sixth-type frame;

decoding the N^{th}-frame bitstream to obtain the N^{th}-frame stereo parameter set when determining that the N^{th}-frame bitstream is the third-type frame;

determining, according to the preset second rule, the k-frame stereo parameter sets in the at least one-frame stereo parameter set preceding the N^{th}-frame stereo parameter set, and obtaining the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on the fourth algorithm when determining that the N^{th}-frame bitstream is the fourth-type frame, wherein k is a positive integer greater than zero; and

restoring the N^{th}-frame downmixed signal to the N^{th}-frame audio signals according to at least one stereo parameter in the N^{th}-frame stereo parameter set based on third algorithm.

15. An encoder, comprising:

a memory configured to store instructions; and

a processor coupled to the memory, wherein the instructions cause the processor to be configured to:

mix N^{th}-frame audio signals on two of a plurality of channels based on a first algorithm to obtain an N^{th}-frame downmixed signal;

detect whether the N^{th}-frame downmixed signal comprises a speech signal, wherein N is a positive integer greater than zero;

encode the N^{th}-frame downmixed signal when the N^{th}-frame downmixed signal comprises the speech signal;

encode the N^{th}-frame downmixed signal when the N^{th}-frame downmixed signal satisfies a preset audio frame encoding condition and when detecting that the N^{th}-frame downmixed signal does not comprise the speech signal; and

skip encoding the N^{th}-frame downmixed signal when the N^{th}-frame downmixed signal does not satisfy the preset audio frame encoding condition and when detecting that the N^{th}-frame downmixed signal does not comprise the speech signal.

16. The encoder of claim 15, wherein the instructions further cause the processor to be configured to:

encode the N^{th}-frame downmixed signal according to a preset speech frame encoding rate when detecting that the N^{th}-frame downmixed signal comprises the speech signal;

encode the N^{th}-frame downmixed signal according to the preset speech frame encoding rate when the N^{th}-frame downmixed signal satisfies a preset speech frame encoding condition; and

encode the N^{th}-frame downmixed signal according to a preset silence insertion descriptor (SID) frame encoding rate when the N^{th}-frame downmixed signal does not satisfy the preset speech frame encoding condition and satisfies a preset SID encoding condition, wherein the preset SID frame encoding rate is less than or equal to the preset speech frame encoding rate.

17. The encoder of claim 16, wherein the instructions further cause the processor to be configured to:

obtain an N^{th}-frame stereo parameter set according to the N^{th}-frame audio signals based on a first stereo parameter set generation manner, and encode the N^{th}-frame stereo parameter set when detecting that the N^{th}-frame audio signals comprise the speech signal, or when detecting that the N^{th}-frame audio signals do not comprise the speech signal and when the N^{th}-frame audio signals satisfy the preset speech frame encoding condition;

obtain the N^{th}-frame stereo parameter set according to the N^{th}-frame audio signals based on a second stereo parameter set generation manner when detecting that the N^{th}-frame audio signals do not comprise the speech signal and when the N^{th}-frame audio signals do not satisfy the preset speech frame encoding condition;

encode at least one stereo parameter in the N^{th}-frame stereo parameter set when the N^{th}-frame stereo parameter set satisfies a preset stereo parameter encoding condition; and

skip encoding the stereo parameter set when the N^{th}-frame stereo parameter set does not satisfy the preset stereo parameter encoding condition, wherein the first stereo parameter set generation manner and the second stereo parameter set generation manner satisfy at least one of the following conditions:

a quantity of types of stereo parameters comprised in a stereo parameter set stipulated in the first stereo parameter set generation manner is greater than or equal to a quantity of types of stereo parameters comprised in a stereo parameter set stipulated in the second stereo parameter set generation manner;

a quantity of stereo parameters comprised in the stereo parameter set stipulated in the first stereo parameter set generation manner is greater than or equal to a quantity stereo parameters comprised in the stereo parameter set stipulated in the second stereo parameter set generation manner;

a time-domain resolution of the stereo parameter stipulated in the first stereo parameter set generation manner is higher than or equal to a time-domain resolution of a corresponding stereo parameter stipulated in the second stereo parameter set generation manner; or

a frequency-domain resolution of the stereo parameter stipulated in the first stereo parameter set generation manner is higher than or equal to frequency-domain resolution of the corresponding stereo parameter stipulated in the second stereo parameter set generation manner.

18. The encoder of claim 15, wherein the instructions further cause the processor to be configured to:

obtain an N^{th}-frame stereo parameter set according to the N^{th}-frame audio signals, wherein the N^{th}-frame stereo parameter set comprises Z stereo parameters, wherein the Z stereo parameters comprise a parameter used when mixing the N^{th}-frame audio signal, and wherein Z is a positive integer greater than zero;

encode the N^{th}-frame stereo parameter set when detecting that the N^{th}-frame downmixed signal comprises the speech signal;

encode at least one stereo parameter in the N^{th}-frame stereo parameter set when the N^{th}-frame stereo parameter set satisfies a preset stereo parameter encoding condition and when detecting that the N^{th}-frame downmixed signal does not comprise the speech signal; and

skip encoding the stereo parameter set when the N^{th}-frame stereo parameter set does not satisfy the preset stereo parameter encoding condition and when detecting that the N^{th}-frame downmixed signal does not comprise the speech signal.

19. The encoder of claim 18, wherein the instructions further cause the processor to be configured to:

obtain X target stereo parameters according to the Z stereo parameters in the N^{th}-frame stereo parameter set based on a preset stereo parameter dimension reduction rule; and

encode the X target stereo parameters, wherein X is a positive integer greater than zero and less than or equal to Z.

20. The encoder of claim 18, wherein the instructions further cause the processor to be configured to:

encode the N^{th}-frame stereo parameter set according to a first encoding manner when detecting that the N^{th}-frame downmixed signal comprises the speech signal and the N^{th}-frame downmixed signal satisfies the preset audio frame encoding condition; and

encode the at least one stereo parameter in the N^{th}-frame stereo parameter set according to a second encoding manner when the N^{th}-frame downmixed signal does not satisfy the preset audio frame encoding condition,

wherein an encoding rate stipulated in the first encoding manner is greater than or equal to an encoding rate stipulated in the second encoding manner, or wherein a quantization precision stipulated in the first encoding manner is higher than or equal to a quantization precision stipulated in the second encoding manner for any stereo parameter in the N^{th}-frame stereo parameter set.

21. The encoder of claim 18, wherein the instructions further cause the processor to be configured to:

determine that the preset stereo parameter encoding condition comprises $D_L \geq D_0$ when the at least one stereo parameter in the N^{th}-frame stereo parameter set comprises an inter-channel level difference (ILD), wherein D_L represents a degree by which the ILD deviates from a first standard, wherein the first standard is determined based on a second algorithm according to T-frame stereo parameter sets preceding the N^{th}-frame stereo parameter set, and wherein T is a positive integer greater than zero;

determine that the preset stereo parameter encoding condition comprises $D_T \geq D_1$ when the at least one stereo parameter in the N^{th}-frame stereo parameter set comprises an inter-channel time difference (ITD) wherein D_T represents a degree by which the ITD deviates from a second standard, and wherein the second standard is determined based on a third algorithm according to the T-frame stereo parameter sets preceding the N^{th}-frame stereo parameter set; and

determine that the preset stereo parameter encoding condition comprises $D_P \geq D_2$ when the at least one stereo parameter in the N^{th}-frame stereo parameter set comprises an inter-channel phase difference (IPD), wherein D_P represents a degree by which the IPD deviates from a third standard, and wherein the third standard is determined based on a fourth algorithm according to the T-frame stereo parameter sets preceding the N^{th}-frame stereo parameter set.

22. The encoder of claim 21, wherein D_L, D_T, and D_P respectively satisfy the following expressions:

$$D_L = \sum_{m=0}^{M-1}\left(ILD(m) - \frac{1}{T}\sum_{t=1}^{T} ILD^{[-t]}(m)\right);$$

$$D_T = ITD - \frac{1}{T}\sum_{t=1}^{T} ITD^{[-t]}(m); \text{ and}$$

$$D_P \sum_{m=0}^{M-1}\left(IPD(m) - \frac{1}{T}\sum_{t=1}^{T} IPD^{[-t]}(m)\right),$$

wherein ILD(m) is a first level difference generated when the N^{th}-frame audio signals are respectively transmitted on two channels in an m^{th} sub frequency band, wherein M is a total quantity of sub frequency bands occupied for transmitting the N^{th}-frame audio signals, wherein

$$\frac{1}{T}\sum_{t=1}^{T} ILD^{[-t]}(m)$$

is an average value of ILDs in the T-frame stereo parameter sets preceding the N^{th}-frame stereo parameter set in the mh sub frequency band, wherein $ILD^{[-t]}(m)$ is a second level difference generated when t^{th}-frame audio signals preceding the N^{th}-frame audio signals are respectively transmitted on the two channels in the m^{th} sub frequency band, wherein the ITD is a first time difference generated when the N^{th}-frame audio signals are respectively transmitted on the two channels, wherein

$$\frac{1}{T}\sum_{t=1}^{T} ITD^{[-t]}$$

is an average value of ITDs in the T-frame stereo parameter sets preceding the N^{th}-frame stereo parameter set, wherein $ITD^{[-t]}$ is a second time difference generated when the t^{th}-frame audio signals preceding the N^{th}-frame audio signals are respectively transmitted on the two channels, wherein IPD(m) is a first phase difference generated when some of the N^{th}-frame audio signals are respectively transmitted on the two channels in the m^{th} sub frequency band, wherein

$$\frac{1}{T}\sum_{t=1}^{T} IPD^{[-t]}(m)$$

is an average value of IPDs in the T-frame stereo parameter sets preceding the N^{th}-frame stereo parameter set in the m^{th} sub frequency band, and wherein $IPD^{[-t]}(m)$ is a second phase difference generated when the t^{th}-frame audio signals preceding the N^{th}-frame audio signals are respectively transmitted on the two channels in the m^{th} sub frequency band.

23. A decoder, comprising:

a memory configured to store instructions; and

a processor coupled to the memory, wherein the instructions cause the processor to be configured to:

receive a bitstream, wherein the bitstream comprises at least two frames, wherein the at least two frames comprise at least one first-type frame or at least one second-type frame, wherein the first-type frame comprises a downmixed signal, and wherein the second-type frame does not comprise the downmixed signal; and

N^{th}-frame bitstream to obtain an N^{th}-frame downmixed signal when the N^{th}-frame bitstream is the first-type frame; and

determine, according to a preset first rule, m-frame downmixed signals in at least one-frame downmixed signal preceding the N^{th}-frame downmixed signal, and obtain the N^{th}-frame downmixed signal according to the m-frame downmixed signals based on a first algorithm when the N^{th}-frame bitstream is the second-tyve frame, wherein m is a positive integer greater than zero, wherein N is a positive integer greater than one, and wherein the N^{th}-frame downmixed signal is received from an encoder after mixing N^{th}-frame audio signals on two of a plurality of channels based on a second algorithm.

24. The decoder of claim 23, wherein the first-type frame comprises the downmixed signal and a stereo parameter set, wherein the second-type frame comprises the stereo parameter set and does not comprise the downmixed signal, and wherein the instructions further cause the processor to be configured to:

decode the N^{th}-frame bitstream to obtain an N^{th}-frame stereo parameter set when the N^{th}-frame bitstream is the first-type frame;

decode the N^{th}-frame bitstream to obtain the N^{th}-frame stereo parameter set when the N^{th}-frame bitstream is the second-type frame; and

restore the N^{th}-frame downmixed signal to the N^{th}-frame audio signals according to at least one stereo parameter in the N h-frame stereo parameter set based on a third algorithm.

25. The decoder of claim 23, wherein the first-type frame comprises the downmixed signal and a stereo parameter set, wherein the second-type frame comprises neither the downmixed signal nor the stereo parameter set, and wherein the instructions further cause the processor to be configured to:

decode the N^{th}-frame bitstream to obtain an N^{th}-frame stereo parameter set when the N^{th}-frame bitstream is the first-type frame;

determine, according to a preset second rule, k-frame stereo parameter sets in at least one-frame stereo parameter set preceding the N^{th}-frame stereo parameter set, and obtain the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on a fourth algorithm when the N^{th}-frame bitstream is the second-type frame, wherein k is a positive integer greater than zero; and

the N^{th}-frame downmixed signal to the N^{th}-frame audio signals according to at least one stereo parameter in the N^{th}-frame stereo parameter set based on a third algorithm.

26. The decoder of claim 23, wherein the first-type frame comprises the downmixed signal and a stereo parameter set, wherein a third-type frame comprises the stereo parameter set and does not comprise the downmixed signal, wherein a fourth-type frame comprises neither the downmixed signal nor the stereo parameter set, wherein each of the third-type frame and the fourth-type frame is one case of the second-type frame, and wherein the instructions further cause the processor to be configured to:

decode the N^{th}-frame bitstream to obtain an N^{th}-frame stereo parameter set when the N^{th}-frame bitstream is the first-type frame;

decode the N^{th}-frame bitstream to obtain the N^{th}-frame stereo parameter set when the N^{th}-frame bitstream is the third-type frame;

determine, according to a preset second rule, k-frame stereo parameter sets in at least one-frame stereo parameter set preceding the N^{th}-frame stereo parameter set, and obtain the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on a fourth algorithm when the N^{th}-frame bitstream is the fourth-type frame, wherein k is a positive integer greater than zero; and

restore the N^{th}-frame downmixed signal to the N^{th}-frame audio signals according to at least one stereo parameter in the N^{th}-frame stereo parameter set based on a third algorithm.

27. The decoder of claim 23, wherein a fifth-type frame comprises both the downmixed signal and a stereo parameter set, wherein a sixth-type frame comprises the downmixed signal and does not comprise the stereo parameter set, wherein each of the fifth-type frame and the sixth-type frame is one case of the first-type frame, wherein the second-type frame comprises neither the downmixed signal nor the stereo parameter set, and wherein the instructions further cause the processor to be configured to:

decode the N^{th}-frame bitstream to obtain an N^{th}-frame stereo parameter set when the N^{th}-frame bitstream is the fifth-type frame;

determine, according to a preset second rule, k-frame stereo parameter sets in at least one-frame stereo parameter set preceding the N^{th}-frame stereo parameter set, and obtain the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on a fourth algorithm when the N^{th}-frame bitstream is the sixth-type frame;

determine, according to the preset second rule, the k-frame stereo parameter sets in the at least one-frame stereo parameter set preceding the N^{th}-frame stereo parameter set, and obtain the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on the fourth algorithm when the N^{th}-frame bitstream is the second-type frame, wherein k is a positive integer greater than zero; and

restore the N^{th}-frame downmixed signal to the N^{th}-frame audio signals according to at least one stereo parameter in the N^{th}-frame stereo parameter set based on a third algorithm.

28. The decoder of claim 23, wherein a fifth-type frame comprises both the downmixed signal and a stereo parameter set, wherein a sixth-type frame comprises the downmixed signal and does not comprise the stereo parameter set, wherein each of the fifth-type frame and the sixth-type frame is one case of the first-type frame, wherein a third-type frame comprises the stereo parameter set and does not comprise the downmixed signal, wherein a fourth-type frame comprises neither the downmixed signal nor the stereo parameter set, wherein each of the third-type frame and the fourth-type frame is one case of the second-type frame, and wherein the instructions further cause the processor to be configured to:

decode the N^{th}-frame bitstream to obtain an N^{th}-frame stereo parameter set when the N^{th}-frame bitstream is the fifth-type frame;

determine, according to a preset second rule, k-frame stereo parameter sets in at least one-frame stereo parameter set preceding the N^{th}-frame stereo parameter set, and obtain the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on a fourth algorithm when the N^{th}-frame bitstream is the sixth-type frame;

decode the N^{th}-frame bitstream to obtain the N^{th}-frame stereo parameter set when the N^{th}-frame bitstream is the third-type frame;

determine, according to the preset second rule, the k-frame stereo parameter sets in the at least one-frame stereo parameter set preceding the N^{th}-frame stereo parameter set, and obtain the N^{th}-frame stereo parameter set according to the k-frame stereo parameter sets based on the fourth algorithm when the N^{th}-frame bitstream is the fourth-type frame, wherein k is a positive integer greater than zero; and

restore the N^{th}-frame downmixed signal to the N^{th}-frame audio signals according to at least one stereo parameter in the N^{th}-frame stereo parameter set based on a third algorithm.