



(11)

EP 4 336 494 A1

(12)

EUROPEAN PATENT APPLICATION
published in accordance with Art. 153(4) EPC

(43) Date of publication:
13.03.2024 Bulletin 2024/11

(21) Application number: **22810378.4**

(22) Date of filing: **12.05.2022**

(51) International Patent Classification (IPC):
G10L 19/008 (2013.01) **G10L 19/00** (2013.01)
H04S 3/00 (2006.01)

(52) Cooperative Patent Classification (CPC):
G10L 19/00; G10L 19/008; H04S 3/00

(86) International application number:
PCT/CN2022/092518

(87) International publication number:
WO 2022/247651 (01.12.2022 Gazette 2022/48)

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME
Designated Validation States:
KH MA MD TN

(30) Priority: **28.05.2021 CN 202110595367**

(71) Applicant: **Huawei Technologies Co., Ltd.**
Shenzhen, Guangdong 518129 (CN)

(72) Inventors:
• **WANG, Zhi**
Shenzhen, Guangdong 518129 (CN)
• **WANG, Zhe**
Shenzhen, Guangdong 518129 (CN)

(74) Representative: **Thun, Clemens**
Mitscherlich PartmbB
Patent- und Rechtsanwälte
Karlstraße 7
80333 München (DE)

(54) **ENCODING METHOD AND APPARATUS FOR MULTI-CHANNEL AUDIO SIGNALS**

(57) This application provides a multi-channel audio signal encoding method and an apparatus. The multi-channel audio signal encoding method includes: obtaining a to-be-encoded first audio frame, where the first audio frame includes at least five channel signals; obtaining a sum of correlation values of all channel pairs in a target channel pair set, where the target channel pair set includes at least one channel pair, one channel pair includes two of the at least five channel signals, the one channel pair has one correlation value, and the correlation value indicates correlation between the two channel signals in the one channel pair; when the sum of the correlation values is greater than a preset threshold, performing energy equalization processing on at least two of the at least five channel signals to obtain at least two equalized channel signals; and encoding the at least two equalized channel signals to obtain an encoded bitstream. In this application, encoding efficiency of an audio frame can be improved.

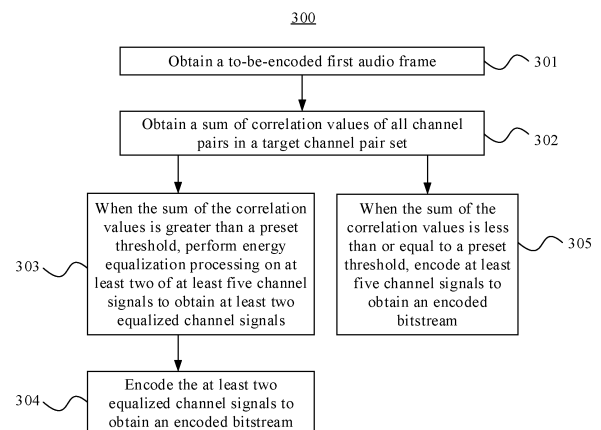


FIG. 3

EP 4 336 494 A1

Description

[0001] This application claims priority to Chinese Patent Application No. 202110595367.2, filed with the China National Intellectual Property Administration on May 28, 2021 and entitled "MULTI-CHANNEL AUDIO SIGNAL ENCODING METHOD AND APPARATUS", which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

[0002] This application relates to audio processing technologies, and in particular, to a multi-channel audio signal encoding method and an apparatus.

BACKGROUND

[0003] Multi-channel audio encoding and decoding is a technology of encoding or decoding audio that includes at least two channels. Common multi-channel audio includes 5.1 channel audio, 7.1 channel audio, 7.1.4 channel audio, 22.2 channel audio, and the like.

[0004] An MPEG surround (MPEG Surround, MPS) standard specifies joint encoding for four channels. However, it still requires encoding and decoding methods for the foregoing various multi-channel audio signals.

SUMMARY

[0005] This application provides a multi-channel audio signal encoding method and an apparatus, to improve encoding efficiency of an audio frame.

[0006] According to a first aspect, this application provides a multi-channel audio signal encoding method. The method includes: obtaining a to-be-encoded first audio frame, where the first audio frame includes at least five channel signals; obtaining a sum of correlation values of all channel pairs in a target channel pair set, where the target channel pair set includes at least one channel pair, one channel pair includes two of the at least five channel signals, the one channel pair has one correlation value, and the correlation value indicates correlation between the two channel signals in the one channel pair; when the sum of the correlation values is greater than a preset threshold, performing energy equalization processing on at least two of the at least five channel signals to obtain at least two equalized channel signals; and encoding the at least two equalized channel signals to obtain an encoded bitstream.

[0007] In this embodiment, to obtain a maximum sum of correlation values, the at least five channel signals included in the audio frame are paired to obtain a target channel pair set. When the sum of the correlation values of the target channel pair set is greater than the preset threshold, the energy equalization processing is performed on the at least two of the at least five channel signals, to perform encoding, to improve encoding efficiency of the audio frame.

[0008] In a possible implementation, the method further includes: when the sum of the correlation values is less than or equal to the preset threshold, encoding the at least five channel signals to obtain an encoded bitstream.

[0009] In this embodiment, if the sum of the correlation values is less than or equal to the preset threshold, it indicates that correlation between two channel signals in the channel pair in the target channel pair set is low, and there is no need to perform encoding in pair, and energy equalization processing does not need to be performed on the at least five channel signals. In this case, an encoded object is the at least five channel signals rather than an equalized channel signal.

[0010] In a possible implementation, the performing energy equalization processing on at least two of the at least five channel signals to obtain at least two equalized channel signals includes: obtaining a fluctuation interval value of the at least five channel signals; determining an energy equalization mode based on the fluctuation interval value of the at least five channel signals; and separately performing energy equalization processing on the at least two channel signals based on the energy equalization mode to obtain the at least two equalized channel signals.

[0011] The fluctuation interval value indicates a difference between energy or amplitude of the at least five channel signals. The energy equalization mode includes a first energy equalization mode and a second energy equalization mode. In the first energy equalization mode, two channel signals of a channel pair are used to obtain two equalized channel signals corresponding to the channel pair. In the second energy equalization mode, two channel signals in one channel pair and at least one channel signal that is not in the channel pair are used to obtain two equalized channel signals corresponding to the channel pair.

[0012] In a possible implementation, the determining an energy equalization mode based on the fluctuation interval value of the at least five channel signals includes: determining the energy equalization mode as the first energy equalization mode when the fluctuation interval value meets a preset condition; or determining the energy equalization mode as the second energy equalization mode when the fluctuation interval value does not meet the preset condition.

[0013] In a possible implementation, the fluctuation interval value includes energy flatness of the first audio frame,

and that the fluctuation interval value meets a preset condition indicates that the energy flatness is less than a first threshold; or the fluctuation interval value includes amplitude flatness of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the amplitude flatness is less than a second threshold; or the fluctuation interval value includes energy deviation of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the energy deviation is not within a first preset range; or the fluctuation interval value includes amplitude deviation of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the amplitude deviation is not within a second preset range.

[0014] In a possible implementation, when the energy equalization mode is the first energy equalization mode, the performing energy equalization processing on at least two of the at least five channel signals to obtain at least two equalized channel signals includes: performing energy equalization processing on channel signals corresponding to the target channel pair set to obtain the at least two equalized channel signals.

[0015] In a possible implementation, the performing energy equalization processing on channel signals corresponding to the target channel pair set to obtain the at least two equalized channel signals includes: calculating, for a current channel pair in the target channel pair set, an average value of energy values or amplitude values of two channel signals included in the current channel pair, and separately performing, based on the average value, energy equalization processing on the two channel signals included in the current channel pair to obtain two corresponding equalized channel signals.

[0016] In this way, when the fluctuation interval value of the at least five channel signals is large, energy equalization may be performed only between two correlated channel signals, so that bit allocation during stereo processing more adapts to a fluctuation interval value of channel signals. This avoids a problem that in a low bit rate encoding environment, encoding noise of a channel pair with high energy may be much greater than encoding noise of a channel pair with low energy due to bit insufficiency, and the channel pair with low energy has bit redundancy.

[0017] In a possible implementation, when the energy equalization mode is the second energy equalization mode, the performing energy equalization processing on at least two of the at least five channel signals to obtain at least two equalized channel signals includes: calculating an average value of energy values or amplitude values of the at least five channel signals, and separately performing energy equalization processing on the at least five channel signals based on the average value to obtain at least five equalized channel signals.

[0018] In a possible implementation, before the determining an energy equalization mode based on the fluctuation interval value of the at least five channel signals, the method further includes: determining whether an encoding bit rate corresponding to the first audio frame is greater than a bit rate threshold; and determining the energy equalization mode as the second energy equalization mode when the encoding bit rate is greater than the bit rate threshold; or determining the energy equalization mode based on the fluctuation interval value when the encoding bit rate is less than or equal to the bit rate threshold.

[0019] In a possible implementation, the method further includes: encoding a channel signal on which energy equalization processing is not performed in the at least five channel signals.

[0020] According to a second aspect, this application provides an encoding apparatus. The apparatus includes: an obtaining module, configured to: obtain a to-be-encoded first audio frame, where the first audio frame includes at least five channel signals; and obtain a sum of correlation values of all channel pairs in a target channel pair set, where the target channel pair set includes at least one channel pair, one channel pair includes two of the at least five channel signals, the one channel pair has one correlation value, and the correlation value indicates correlation between the two channel signals in the one channel pair; a processing module, configured to perform energy equalization processing on at least two of the at least five channel signals to obtain at least two equalized channel signals when the sum of the correlation values is greater than a preset threshold; and an encoding module, configured to encode the at least two equalized channel signals to obtain an encoded bitstream.

[0021] In a possible implementation, the encoding module is further configured to: when the sum of the correlation values is less than or equal to the preset threshold, encode the at least five channel signals to obtain an encoded bitstream.

[0022] In a possible implementation, the processing module is specifically configured to: obtain a fluctuation interval value of the at least five channel signals; determine an energy equalization mode based on the fluctuation interval value of the at least five channel signals; and separately perform energy equalization processing on the at least two channel signals based on the energy equalization mode to obtain the at least two equalized channel signals.

[0023] In a possible implementation, the processing module is specifically configured to: determine the energy equalization mode as a first energy equalization mode when the fluctuation interval value meets a preset condition; or determine the energy equalization mode as a second energy equalization mode when the fluctuation interval value does not meet a preset condition.

[0024] In a possible implementation, the fluctuation interval value includes energy flatness of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the energy flatness is less than a first threshold; or the fluctuation interval value includes amplitude flatness of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the amplitude flatness is less than a second threshold; or the fluctuation interval value includes energy deviation of the first audio frame, and that the fluctuation interval value meets

a preset condition indicates that the energy deviation is not within a first preset range; or the fluctuation interval value includes amplitude deviation of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the amplitude deviation is not within a second preset range.

[0025] In a possible implementation, when the energy equalization mode is the first energy equalization mode, the processing module is specifically configured to perform energy equalization processing on channel signals corresponding to the target channel pair set to obtain the at least two equalized channel signals.

[0026] In a possible implementation, the processing module is specifically configured to calculate, for a current channel pair in the target channel pair set, an average value of energy values or amplitude values of two channel signals included in the current channel pair, and separately perform, based on the average value, energy equalization processing on the two channel signals included in the current channel pair to obtain two corresponding equalized channel signals.

[0027] In a possible implementation, when the energy equalization mode is the second energy equalization mode, the processing module is specifically configured to: calculate an average value of energy values or amplitude values of the at least five channel signals; and separately perform energy equalization processing on the at least five channel signals based on the average value to obtain at least five equalized channel signals.

[0028] In a possible implementation, the processing module is further configured to: determine whether an encoding bit rate corresponding to the first audio frame is greater than a bit rate threshold; and determine the energy equalization mode as the second energy equalization mode when the encoding bit rate is greater than the bit rate threshold; or determine the energy equalization mode based on the fluctuation interval value when the encoding bit rate is less than or equal to the bit rate threshold.

[0029] In a possible implementation, the encoding module is further configured to encode a channel signal on which energy equalization processing is not performed in the at least five channel signals.

[0030] According to a third aspect, this application provides a device. The device includes: one or more processors; and a memory, configured to store one or more programs. When the one or more programs are executed by the one or more processors, the one or more processors are enabled to implement the method according to any possible implementation of the first aspect.

[0031] According to a fourth aspect, this application provides a computer-readable storage medium. The computer-readable storage medium includes a computer program. When the computer program is executed by a computer, the computer is enabled to perform the method according to any one of the possible implementations of the first aspect.

[0032] According to a fifth aspect, an embodiment of this application provides a computer-readable storage medium. The computer-readable storage medium includes an encoded bitstream obtained by using the multi-channel audio signal encoding method according to any possible implementation of the first aspect.

BRIEF DESCRIPTION OF DRAWINGS

[0033]

FIG. 1 is an example of a schematic block diagram of an audio transcoding system 10 to which this application is applied;

FIG. 2 is an example of a schematic block diagram of an audio transcoding device 200 to which this application is applied;

FIG. 3 is a flowchart of an example embodiment of a multi-channel audio signal encoding method according to this application;

FIG. 4a is an example diagram depicting a structure of an encoding apparatus to which a multi-channel audio signal encoding method is applied according to this application;

FIG. 4b is an example diagram depicting a structure of a multi-channel adaptive pairing module;

FIG. 4c is an example diagram depicting a structure of a pairing processing module;

FIG. 5 is an example diagram depicting a structure of a decoding apparatus to which a multi-channel audio decoding method is applied according to this application;

FIG. 6 is a schematic diagram depicting a structure of an encoding apparatus according to an embodiment of this application; and

FIG. 7 is a schematic diagram depicting a structure of a device according to an embodiment of this application.

DESCRIPTION OF EMBODIMENTS

[0034] To make the objectives, technical solutions, and advantages of this application clearer, the following clearly and completely describes the technical solutions of this application with reference to the accompanying drawings in embodiments of this application. It is clear that the described embodiments are merely a part rather than all of embodiments of this application. All other embodiments obtained by persons of ordinary skill in the art based on embodiments

of this application without creative efforts shall fall within the protection scope of this application.

[0035] In this specification, embodiments, claims, and accompanying drawings of this application, terms "first", "second", and the like are merely intended for distinguishing and description, and shall not be understood as an indication or implication of relative importance or an indication or implication of an order. In addition, the terms "include", "have", and any variant thereof are intended to cover non-exclusive inclusion, for example, include a series of steps or units. Methods, systems, products, or devices are not necessarily limited to those steps or units that are literally listed, but may include other steps or units that are not literally listed or that are inherent to such processes, methods, products, or devices.

[0036] It should be understood that in this application, "at least one" means one or more and "a plurality of" means two or more. The term "and/or" is used for describing an association relationship between associated objects, and represents that three relationships may exist. For example, "A and/or B" may represent the following three cases: Only A exists, only B exists, and both A and B exist, where A and B may be singular or plural. The character "/" generally indicates an "or" relationship between the associated objects. "At least one of the following" or a similar expression thereof indicates any combination of the following, including any combination of one or more of the following. For example, at least one of a, b, or c may indicate a, b, c, a and b, a and c, b and c, or a and b and c, where a, b, and c may be singular or plural.

[0037] Explanations of related terms in this application are as follows:

Audio frame: Audio data is in a stream form. In an actual application, to facilitate audio processing and transmission, an audio data amount within one duration is usually selected as a frame of audio. The duration is referred to as a "sampling time period", and a value of the duration may be determined based on a requirement of a codec and a specific application, for example, the duration ranges from 2.5 ms to 60 ms, where ms is millisecond.

[0038] **Audio signal:** The audio signal is a frequency and amplitude change information carrier of a regular sound wave with voice, music, and sound effect. Audio is a continuously changing analog signal, and can be represented by a continuous curve and referred to as a sound wave. A digital signal generated from the audio through analog-to-digital conversion or by using a computer is an audio signal. The sound wave has three important parameters: frequency, amplitude, and phase, which determine characteristics of the audio signal.

[0039] **Channel signals** are independent audio signals that are collected or played in different spatial positions during sound recording or playing. Therefore, a quantity of channels is a quantity of audio sources used during audio recording, or a quantity of loudspeakers used for audio playing.

[0040] A system architecture to which this application is applied is described below.

[0041] FIG. 1 is an example of a schematic block diagram of an audio transcoding system 10 to which this application is applied. As shown in FIG. 1, the audio transcoding system 10 may include a source device 12 and a destination device 14. The source device 12 generates an encoded bitstream. Therefore, the source device 12 may be referred to as an audio encoding apparatus. The destination device 14 may decode the encoded bitstream generated by the source device 12. Therefore, the destination device 14 may be referred to as an audio decoding apparatus.

[0042] The source device 12 includes a coder 20, and optionally, may include an audio source 16, an audio preprocessor 18, and a communication interface 22.

[0043] The audio source 16 may include or may be any type of audio capture device configured to capture real-world speech, music, sound effect, and the like; and/or any type of audio generation device, for example, an audio processor or device configured to generate speech, music, and sound effect. The audio source may be any type of memory or storage that stores the foregoing audio.

[0044] The audio preprocessor 18 is configured to receive (original) audio data 17, and preprocess the audio data 17 to obtain preprocessed audio data 19. For example, preprocessing performed by the audio preprocessor 18 may include pruning or noise reduction. It may be understood that the audio preprocessor 18 may be an optional component.

[0045] The coder 20 is configured to receive the preprocessed audio data 19 and provide encoded audio data 21.

[0046] The communication interface 22 in the source device 12 may be configured to receive the encoded audio data 21 and send the encoded audio data 21 to the destination device 14 through a communication channel 13, to store or directly reconstruct the encoded audio data 21.

[0047] The destination device 14 includes a decoder 30, and optionally, may include a communication interface 28, an audio postprocessor 32, and a playing device 34.

[0048] The communication interface 28 in the destination device 14 is configured to directly receive the encoded audio data 21 from the source device 12, and provide the encoded audio data 21 to the decoder 30.

[0049] The communication interface 22 and the communication interface 28 may be configured to use a direct communication link between the source device 12 and the destination device 14, for example, a direct wired or wireless connection; or use any type of network, for example, a wired network, a wireless network, or any combination thereof, any type of private network and public network, or any type of combination thereof, to send or receive the encoded audio data 21.

[0050] For example, the communication interface 22 may be configured to encapsulate the encoded audio data 21

into a suitable format such as a packet, and/or process the encoded audio data 21 through any type of transmission encoding or processing, to be transmitted over a communication link or a communication network.

[0051] The communication interface 28 corresponds to the communication interface 22. For example, the communication interface 28 may be configured to receive transmitted data, and process the transmitted data through any type of corresponding transmission decoding or processing and/or decapsulation, to obtain the encoded audio data 21.

[0052] The communication interface 22 and the communication interface 28 each may be configured as a unidirectional communication interface indicated by an arrow that is of the corresponding communication channel 13 and that points from the source device 12 to the destination device 14 in FIG. 1 or a bidirectional communication interface; and may be configured to send and receive a message, or the like, to establish a connection, confirm and exchange any other information related to data transmission, such as a communication link and/or encoded audio data.

[0053] The decoder 30 is configured to receive the encoded audio data 21 and provide decoded audio data 31.

[0054] The audio postprocessor 32 is configured to perform postprocessing on the decoded audio data 31 to obtain postprocessed audio data 33. Post-processing performed by the audio postprocessor 32 may include, for example, pruning or resampling.

[0055] The playing device 34 is configured to receive the postprocessed audio data 33, to play audio to a user or a listener. The playing device 34 may be or include any type of player configured to play reconstructed audio, for example, an integrated or external loudspeaker. For example, the loudspeaker may include a horn, a speaker, and the like.

[0056] FIG. 2 is an example of a schematic block diagram of an audio transcoding device 200 to which this application is applied. In an embodiment, the audio transcoding device 200 may be an audio decoder (for example, the decoder 30 in FIG. 1) or an audio coder (for example, the coder 20 in FIG. 1).

[0057] The audio transcoding device 200 includes an ingress port 210 and a receiving unit (Rx) 220 for receiving data; a processor, a logic unit, or a central processing unit 230 for processing data; a transmitting unit (Tx) 240 and an egress port 250 for transmitting data; and a memory 260 for storing data. The audio transcoding device 200 may further include an optical-to-electrical conversion component and an electrical-to-optical (EO) component coupled to the ingress port 210, the receiving unit 220, the transmitting unit 240, and the egress port 250. The components are configured as ingress ports or egress ports of an optical signal or an electrical signal.

[0058] The processor 230 is implemented through hardware and software. The processor 230 may be implemented as one or more CPU chips, cores (for example, a multi-core processor), FPGAs, ASICs, and DSPs. The processor 230 communicates with the ingress port 210, the receiving unit 220, the transmitting unit 240, the egress port 250, and the memory 260. The processor 230 includes a transcoding module 270 (for example, an encoding module or a decoding module). The transcoding module 270 implements the embodiments disclosed in this application, to implement the multi-channel audio signal encoding method provided in this application. For example, the transcoding module 270 implements, processes, or provides various encoding operations. Therefore, the transcoding module 270 substantially improves functions of the audio transcoding device 200, and affects conversion of the audio transcoding device 200 to different states. Alternatively, the transcoding module 270 is implemented by using instructions stored in the memory 260 and executed by the processor 230.

[0059] The memory 260 includes one or more disks, tape drives, and solid state drives, and may be used as an overflow data storage device, to store programs when such programs are selected for execution, and to store instructions and data that are read during program execution. The memory 260 may be volatile and/or non-volatile, and may be a read-only memory (ROM), a random access memory (RAM), a random access memory (ternary content-addressable memory, TCAM), and/or a static random access memory (SRAM).

[0060] Based on the description of the foregoing embodiments, this application provides a multi-channel audio signal encoding method.

[0061] FIG. 3 is a flowchart of an example embodiment of a multi-channel audio signal encoding method according to this application. A process 300 may be executed by the source device 12 in the audio transcoding system 10 or the audio transcoding device 200. The process 300 includes a series of steps or operations. It should be understood that steps in the process 300 may be performed in various sequences and/or simultaneously, and is not limited to an execution sequence shown in FIG. 3. As shown in FIG. 3, the method includes the following steps.

[0062] Step 301: Obtain a to-be-encoded first audio frame.

[0063] The first audio frame in this embodiment may be any frame of to-be-encoded multi-channel audio, and the first audio frame includes five or more channel signals. For example, 5.1 channels include six channel signals: a center (C) channel signal, a front left (left, L) channel signal, a front right (right, R) channel signal, a back left surround (left surround, LS) channel signal, a back right surround (right surround, RS) channel signal, and a 0.1 channel low frequency effects (low frequency effects, LFE) channel signal. 7.1 channels include eight channel signals: a C channel signal, an L channel signal, an R channel signal, an LS channel signal, an RS channel signal, an LB channel signal, an RB channel signal, and an LFE channel signal. An LFE channel is an audio channel ranging from 3 Hz to 120 Hz, which is usually sent to a loudspeaker specially designed for low tones.

[0064] Step 302: Obtain a sum of correlation values of all channel pairs in a target channel pair set.

[0065] The target channel pair set is obtained to obtain a maximum sum of correlation values. The target channel pair set includes at least one channel pair, and the channel pair includes two channel signals in at least five channel signals. One channel pair has one correlation value, and the correlation value indicates correlation between two channel signals of one channel pair.

[0066] Two highly correlated channel signals are encoded together can reduce redundancy and improve encoding efficiency. Therefore, in this embodiment, pairing is determined based on a correlation value of the two channel signals. To find a pairing manner with highest correlation as much as possible, a correlation value between every two of the at least five channel signals in the first audio frame may be first calculated to obtain a correlation value set of the first audio frame. For example, 10 channel pairs in total may be formed for the five channel signals; and correspondingly, the correlation value set may include 10 correlation values.

[0067] Optionally, the correlation values may be normalized. In this way, the correlation values of all channel pairs are limited within a specific range, to set a unified determining standard, for example, a pairing threshold, for the correlation value. The pairing threshold may be set to a value greater than or equal to 0.2 and less than or equal to 1, for example, 0.3. In this way, as long as a normalized correlation value of two channel signals is smaller than the pairing threshold, it is considered that the two channel signals have low correlation and pairing the two channel signals for encoding is not needed.

[0068] In a possible implementation, the correlation value between the two channel signals (for example, ch1 and ch2) may be calculated based on the following formula:

$$\text{corr}(\text{ch1}, \text{ch2}) = \frac{\sum_{i=1}^N (\text{spec}_{\text{ch1}(i)} \times \text{spec}_{\text{ch2}(i)})}{\sqrt{\sum_{i=1}^N (\text{spec}_{\text{ch1}(i)} \times \text{spec}_{\text{ch1}(i)}) \times \sum_{i=1}^N (\text{spec}_{\text{ch2}(i)} \times \text{spec}_{\text{ch2}(i)})}}$$

$\text{corr}(\text{ch1}, \text{ch2})$ is a normalized correlation value between the channel signal ch1 and the channel signal ch2, $\text{spec_ch1}(i)$ is a frequency domain coefficient of an i^{th} frequency of the channel signal ch1, $\text{spec_ch2}(i)$ is a frequency domain coefficient of an i^{th} frequency of the channel signal ch2, and N indicates an integer value that does not exceed a total frequency quantity of one audio frame.

[0069] It should be noted that another algorithm or formula may alternatively be used to calculate the correlation value between the two channel signals. This is not specifically limited in this application.

[0070] The obtaining a pairing manner of the target channel pair set includes: to obtain a maximum sum of correlation values, selecting a channel pair from channel pairs corresponding to at least five channel signals, and adding the channel pair to the target channel pair set. The sum of correlation values of the target channel pair set is a sum of correlation values of all channel pairs of the target channel pair set that are obtained by performing pairing on the at least five channel signals based on the foregoing pairing manner. The pairing manner in this embodiment may include the following two implementations.

(1) Select M largest correlation values from the correlation value set. The M correlation values need to be greater than or equal to the pairing threshold, because a correlation value less than the pairing threshold indicates that correlation between two channel signals in a channel pair corresponding to the correlation value is low, and pairing of the channel signals for encoding is not needed. To improve encoding efficiency, there is no need to select all correlation values greater than or equal to the pairing threshold. Therefore, an upper limit N of M is set, in other words, a maximum of N correlation values are selected.

[0071] N may be an integer greater than or equal to 2, and a maximum value of N cannot exceed a quantity of all channel pairs corresponding to all channel signals of the first audio frame. A larger value of N indicates an increase in a calculation amount. A smaller value of N indicates that a channel pair set may be lost, and encoding efficiency is reduced.

[0072] Optionally, N may be set to a maximum quantity of channel pairs plus 1, that is, $N = \left\lceil \frac{\text{CH}}{2} \right\rceil + 1$, where CH represents a quantity of channel signals included in the first audio frame packet. For example, if a 5.1 channel packet includes five channel signals, $N = 3$. If a 7.1 channel packet includes seven channel signals, $N = 4$.

[0073] Then, M channel pair sets are obtained based on the M correlation values. Each channel pair set includes at least one of M channel pairs corresponding to the M correlation values, and when the channel pair set includes at least two channel pairs, the at least two channel pairs do not include a same channel signal. For example, for the 5.1 channels, three channel pairs (L, R), (R, C), and (LS, RS) corresponding to the maximum correlation value are selected based on the correlation value set. A correlation value of (LS, RS) is less than the pairing threshold, and therefore the channel pair is excluded. In this case, two channel pair sets may be obtained by using the two channel pairs (L, R) and (R, C).

One of the two channel pair sets includes (L, R), and the other includes (R, C).

[0074] Any one of the M channel pairs (for example, a first channel pair) corresponding to M correlation values greater than or equal to the pairing threshold is used as an example. A method for obtaining the M channel pair sets in this embodiment may include: adding the first channel pair to the target channel pair set, where the M channel pair sets include the target channel pair set; and when other channel pairs other than an associated channel pair in the plurality of channel pairs include a channel pair with a correlation value greater than the pairing threshold, selecting a channel pair with a maximum correlation value from the other channel pairs and adding the channel pair to the target channel pair set, where the associated channel pair includes any channel signal included in a channel pair included in the target channel pair set.

[0075] Except the step of adding the first channel pair to the target channel pair set, all the foregoing processes are iterative processing steps:

a: Determine whether the channel pairs other than the associated channel pair in the plurality of channel pairs include a channel pair whose correlation value is greater than the pairing threshold; and

b: If the channel pair whose correlation value is greater than the pairing threshold is included, select a channel pair whose correlation value is maximum from the other channel pairs, and adding the channel pair to the target channel pair set.

[0076] In this case, as long as the other channel pairs include a channel pair whose correlation value is greater than the pairing threshold, step b may be performed iteratively.

[0077] Optionally, to reduce a calculation amount, a correlation value less than the pairing threshold may be deleted from the correlation value set. In this way, a quantity of channel pairs may be reduced, and a quantity of iterations may be further reduced.

[0078] (2) Obtain, based on a plurality of channel pairs, all channel pair sets corresponding to the at least five channel signals, obtain, based on the correlation value set, a sum of correlation values of all channel pairs included in any channel pair set in all the channel pair sets, and determine a channel pair set in all the channel pair sets that is corresponding to a maximum sum of correlation values as a target channel pair set.

[0079] The correlation value set includes correlation values of a plurality of channel pairs of at least five channel signals in the first audio frame, and the plurality of channel pairs are regularly combined (in other words, a plurality of channel pairs in a same channel pair set cannot include a same channel signal) to obtain the plurality of channel pair sets corresponding to the at least five channel signals.

[0080] In a possible implementation, when a quantity of channel signals is an odd number, a quantity of all channel pair sets may be calculated based on the following formula:

$$Pair_{num} = \frac{C_{CH}^2 \times C_{CH-2}^2 \times \dots \times C_3^2}{\frac{A_{CH}^2}{2}}$$

[0081] In a possible implementation, when a quantity of channel signals is an even number, a quantity of all channel pair sets may be calculated based on the following formula:

$$Pair_{num} = \frac{C_{CH}^2 \times C_{CH-2}^2 \times \dots \times C_2^2}{\frac{A_{CH}^2}{2}}$$

[0082] Pair num indicates the quantity of all channel pair sets; and CH indicates a quantity of channel signals for multi-channel processing in the first audio frame, and is a result obtained through multi-channel mask screening.

[0083] Optionally, to reduce a calculation amount, after the correlation value set is obtained, the plurality of channel pair sets may be obtained based on channel pairs other than an uncorrelated channel pair in the plurality of channel pairs. A correlation value of the uncorrelated channel pair is less than the pairing threshold. In this way, when the channel pair set is obtained, a quantity of channel pairs in calculation may be reduced, a quantity of channel pair sets is reduced, and a calculation amount of a sum of correlation values may also be reduced in a subsequent step.

[0084] Step 303: When the sum of the correlation values is greater than the preset threshold, perform energy equalization processing on at least two of the at least five channel signals to obtain at least two equalized channel signals.

[0085] In a possible implementation, a fluctuation interval value of the at least five channel signals may be first obtained, an energy equalization mode is determined based on the fluctuation interval value of the at least five channel signals,

and then energy equalization processing is separately performed on the at least five channel signals based on the energy equalization mode to obtain at least five equalized channel signals.

[0086] The fluctuation interval value indicates a difference between energy or amplitude of the at least five channel signals.

[0087] The energy equalization mode includes a first energy equalization mode and a second energy equalization mode. In the first energy equalization mode, two channel signals of a channel pair are used to obtain two equalized channel signals corresponding to the channel pair. In the second energy equalization mode, two channel signals in one channel pair and at least one channel signal that is not in the channel pair are used to obtain two equalized channel signals corresponding to the channel pair.

[0088] The determining an energy equalization mode based on the fluctuation interval value of the at least five channel signals may include: determining the energy equalization mode as the first energy equalization mode when the fluctuation interval value meets a preset condition; or determining the energy equalization mode as the second energy equalization mode when the fluctuation interval value does not meet the preset condition.

[0089] The fluctuation interval value includes energy flatness of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the energy flatness is less than a first threshold; or the fluctuation interval value includes amplitude flatness of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the amplitude flatness is less than a second threshold; or the fluctuation interval value includes energy deviation of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the energy deviation is not within a first preset range; or the fluctuation interval value includes amplitude deviation of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the amplitude deviation is not within a second preset range.

[0090] In this embodiment of the present invention, the energy flatness represents fluctuation of frame energy that is obtained after energy normalization with a current frame frequency domain coefficient is performed on a plurality of channels screened by a multi-channel screening unit, and may be measured based on a flatness calculation formula.

When energy of all channels of the current frame is the same, the energy flatness of the current frame is 1. When energy of a specific channel of the current frame is 0, the energy flatness of the current frame is 0. Therefore, a value range of inter-channel energy flatness is [0, 1]. Larger fluctuation of inter-channel energy indicates a smaller value of energy flatness. In an implementation, a unified first threshold, for example, 0.483, 0.492, or 0.504, may be set for all channel formats (for example, 5.1, 7.1, 9.1, and 11.1). In another implementation, different first thresholds are set for different channel formats. For example, a first threshold for the 5.1 channel format is 0.511, a first threshold for the 7.1 channel format is 0.563, a first threshold for the 9.1 channel format is 0.608, and a first threshold for the 11.1 channel format is 0.654.

[0091] The amplitude flatness represents fluctuation of frame amplitude obtained after amplitude normalization with a current frame frequency domain coefficient is performed on a plurality of channels screened by a multi-channel screening unit, and may be measured based on a flatness calculation formula. When frame amplitude of all channels is the same, the flatness is 1. When frame amplitude of a specific channel is 0, the flatness is 0. Therefore, a range of the amplitude flatness is [0, 1]. Larger fluctuation of inter-channel amplitude indicates a smaller value of the flatness. In an implementation, a unified second threshold, for example, 0.695, 0.701, or 0.710, may be set for all channel formats (for example, 5.1, 7.1, 9.1, and 11.1). In another implementation, different second thresholds may be provided for different channel formats. For example, a second threshold for the 5.1 channel format may be 0.715, a second threshold for the 7.1 channel format may be 0.753, a second threshold for the 9.1 channel format may be 0.784, and a second threshold for the 11.1 channel format may be 0.809.

[0092] Because there is a square relationship between the amplitude and the energy, there is also a square relationship between the amplitude flatness and the energy flatness, that is, fluctuation of inter-channel frame amplitude corresponding to a square of the amplitude flatness is approximately equivalent to fluctuation of inter-channel frame energy corresponding to the energy flatness.

[0093] In this embodiment, the energy equalization mode may be determined based on the foregoing plurality of types of information indicating a fluctuation interval value of the at least five channel signals, and the information includes energy flatness, amplitude flatness, energy deviation, or amplitude deviation.

(1) Calculate energy values of the at least five channel signals, obtain the energy flatness of the first audio frame based on the energy values of the at least five channel signals, and determine the energy equalization mode as the first energy equalization mode when the energy flatness of the first audio frame is less than the first threshold; or determine the energy equalization mode as the second energy equalization mode when the energy flatness of the first audio frame is greater than or equal to the first threshold.

(2) Calculate amplitude values of the at least five channel signals, obtain the amplitude flatness of the first audio frame based on the amplitude values of the at least five channel signals, and determine the energy equalization mode as the first energy equalization mode when the amplitude flatness of the first audio frame is less than the second threshold; or determine the energy equalization mode as the second energy equalization mode when the

amplitude flatness of the first audio frame is greater than or equal to the second threshold.

(3) Calculate energy values of the at least five channel signals, obtain the energy deviation of the first audio frame based on the energy values of the at least five channel signals, and determine the energy equalization mode as the first energy equalization mode when the energy deviation of the first audio frame is not within the first preset range; or determine the energy equalization mode as the second energy equalization mode when the energy deviation of the first audio frame is within the first preset range.

(4) Calculate amplitude values of the at least five channel signals, obtain the amplitude deviation of the first audio frame based on the amplitude values of the at least five channel signals, and determine the energy equalization mode as the first energy equalization mode when the amplitude deviation of the first audio frame is not within the second preset range; or determine the energy equalization mode as the second energy equalization mode when the amplitude deviation of the first audio frame is within the second preset range.

[0094] It should be noted that another energy equalization mode may alternatively be used in this application. This is not specifically limited herein.

[0095] In a possible implementation, before an energy equalization mode is determined based on the fluctuation interval value of the at least five channel signals, the energy equalization mode may be first determined based on an encoding bit rate corresponding to the first audio frame, that is, whether the encoding bit rate is greater than a bit rate threshold is determined. The energy equalization mode is determined as the second energy equalization mode when the encoding bit rate is greater than the bit rate threshold; or the energy equalization mode is determined based on the fluctuation interval value of the at least five channel signals when the encoding bit rate is less than or equal to the bit rate threshold.

[0096] When the energy equalization mode is the first energy equalization mode, for a current channel pair in the target channel pair set corresponding to the pairing manner, an average value of energy values or amplitude values of two channel signals included in the current channel pair may be calculated; and energy equalization processing is separately performed on the two channel signals based on the average value to obtain two corresponding equalized channel signals.

[0097] In this way, when the fluctuation interval value of the at least five channel signals is large, energy equalization may be performed only between two correlated channel signals, so that bit allocation during stereo processing more adapts to a fluctuation interval value of channel signals. This avoids a problem that in a low bit rate encoding environment, encoding noise of a channel pair with high energy may be much greater than encoding noise of a channel pair with low energy due to bit insufficiency, and the channel pair with low energy has bit redundancy.

[0098] When the energy equalization mode is the second energy equalization mode, an average value of energy values or amplitude values of the at least five channel signals may be calculated, and energy equalization processing is separately performed on the at least five channel signals based on the average value to obtain at least five equalized channel signals.

[0099] It should be noted that, in step 303, energy equalization processing is mainly performed on at least two of the at least five channel signals to obtain at least two equalized channel signals, and the at least two channel signals are channel signals that are paired in the target channel pair set, and remaining channel signals that are not paired in the target channel pair set are directly encoded.

[0100] Step 304: Encode the at least two equalized channel signals to obtain an encoded bitstream.

[0101] Step 305: When the sum of the correlation values is less than or equal to the preset threshold, encode the at least five channel signals to obtain an encoded bitstream.

[0102] In this embodiment, if the sum of the correlation values is less than or equal to the preset threshold, it indicates that correlation between two channel signals in the channel pair in the target channel pair set is low, and there is no need to perform encoding in pair, and energy equalization processing does not need to be performed on the at least five channel signals. In this case, an encoded object is the at least five channel signals rather than an equalized channel signal.

[0103] In this embodiment, to obtain a maximum sum of correlation values, the at least five channel signals included in the audio frame are paired to obtain a target channel pair set. When the sum of the correlation values of the target channel pair set is greater than the preset threshold, the energy equalization processing is performed on the at least five channel signals, to perform encoding, to improve encoding efficiency of the audio frame.

[0104] The following describes, by using two specific embodiments, the process of determining the pairing manner and the energy equalization mode in the method embodiment shown in FIG. 3. A 5.1 channel is used as an example. The 5.1 channel includes a central channel (C), a front left channel (left, L), a front right channel (right, R), a back left surround channel (left surround, LS), a back right surround channel (right surround, RS), and 0.1 channel low frequency effects (low frequency effects, LFE). As shown in Table 1. A channel index is set for the six channel signals.

Table 1

Channel index	Channel signal
0	L
1	R
2	LS
3	RS
4	C
5	LFE

[0105] FIG. 4a is an example diagram depicting a structure of an encoding apparatus to which a multi-channel audio signal encoding method is applied according to this application. The encoding apparatus may be the coder 20 of the source device 12 in the audio transcoding system 10, or may be the transcoding module 270 in the audio transcoding device 200. The encoding apparatus may include a multi-channel adaptive pairing module, a channel encoding module, and a bitstream multiplexing interface.

[0106] Input of the multi-channel adaptive pairing module includes six channel signals (L, R, C, LS, RS, LFE) of 5.1 channels and a multi-channel processing indicator (MultiProcFlag), and outputs six channel signals (M1, S1, M2, S2, C, LFE) after pairing, where M1 and S1 are a channel pair obtained by pairing, M2 and S2 are a channel pair obtained by pairing, and outputs multi-channel side information (sideInfoMc), the multi-channel side information includes a channel pair set.

[0107] The channel encoding module uses mono channel encoding units (or mono-channel channel boxes or mono-channel tools) to encode the channel signals (M1, S1, M2, S2, C and LFE) output by the multi-channel adaptive pairing module, and outputs corresponding encoded channel signals (E1 to E6). In a process of encoding the channel signals by the mono-channel encoding unit, more bits are allocated to a channel signal with higher energy (or a higher amplitude), and fewer bits are allocated to a channel signal with lower energy (or a lower amplitude). Optionally, the channel encoding module may alternatively use a stereo encoding unit, for example, a parametric stereo coder or a lossy stereo encoder, to encode the processed channel signals output by the multi-channel processing module.

[0108] It should be noted that unpaired channel signals (for example, C and LFE) may be directly input into the channel encoding module to obtain the encoded channel signals E5 and E6.

[0109] The bitstream multiplexing interface generates encoded multi-channel signals. The encoded multi-channel signals include the encoded channel signals (E1 to E6) output by the channel encoding module and side information (including the multi-channel side information). Optionally, the bitstream multiplexing interface may process the encoded multi-channel signals into serial signals or serial bitstreams.

[0110] FIG. 4b is an example diagram depicting a structure of a multi-channel adaptive pairing module. As shown in FIG. 4b, the multi-channel adaptive pairing module includes: a multi-channel screening unit, a global correlation value collecting unit, a multi-channel energy equalization selection module, and a pairing processing module.

[0111] The multi-channel screening unit screens five channel signals participating in multi-channel processing, namely, L, R, C, LS, and RS, from the six channel signals (L, R, C, LS, RS and LFE) based on the multi-channel processing indicator (MultiProcFlag).

[0112] The global correlation value statistics unit first calculates a normalized correlation value between any two of the channel signals L, R, C, LS, and RS that participate in multi-channel processing. In this application, a correlation value between two channel signals (for example, a channel signal ch1 and a channel signal ch2) may be calculated based on the following formula:

$$\text{corr}(\text{ch1}, \text{ch2}) = \frac{\sum_{i=1}^N (\text{spec}_{\text{ch1}(i)} \times \text{spec}_{\text{ch2}(i)})}{\sqrt{\sum_{i=1}^N (\text{spec}_{\text{ch1}(i)} \times \text{spec}_{\text{ch1}(i)}) \times \sum_{i=1}^N (\text{spec}_{\text{ch2}(i)} \times \text{spec}_{\text{ch2}(i)})}}$$

[0113] $\text{corr}(\text{ch1}, \text{ch2})$ is a normalized correlation value between the channel signal ch1 and the channel signal ch2, $\text{spec_ch1}(i)$ is a frequency domain coefficient of an i^{th} frequency of the channel signal ch1, $\text{spec_ch2}(i)$ is a frequency domain coefficient of an i^{th} frequency of the channel signal ch2, and N indicates an integer value that does not exceed a total frequency quantity of one audio frame. Then, a maximum sum of correlation values of channel pair sets corresponding to all channel signals participated in multi-channel processing (that is, a sum of correlation values of all channel

pairs included in a channel pair set) and a channel pair set (which is considered as a target channel pair set) corresponding to the maximum sum of correlation values are determined, based on the normalized correlation value between any two channel signals. Finally, global correlation value side information is output, and the global correlation value side information includes the maximum sum of correlation values corr_sum_max and the target channel pair set. It is assumed that the target channel pair set includes (R, C) and (LS, RS), and the maximum sum of correlation values is $\text{corr_sum_max} = \text{corr}(L, R) + \text{corr}(LS, RS)$.

[0114] It should be noted that, after obtaining the normalized correlation value between any two channel signals, the global correlation value statistics unit may screen the correlation values based on a pairing threshold. That is, a correlation value greater than or equal to the pairing threshold is retained, and a correlation value less than the pairing threshold is deleted or set to 0. In this way, a calculation amount can be reduced.

[0115] The multi-channel energy equalization selection module determines, based on an encoding bit rate and the five channel signals, whether energy equalization processing needs to be performed for the five channel signals. A pairing manner of the five channel signals is a global pairing manner. This manner aims to obtain a maximum sum of correlation values. For details, refer to Step 302. When the sum of the correlation values of the target channel pair set is greater than a preset threshold, it is determined that energy equalization needs to be performed on the five channel signals, or when the sum of the correlation values of the target channel pair set is less than or equal to the preset threshold, it is determined that energy equalization does not need to be performed on the five channel signals. When it is determined that energy equalization processing needs to be performed on the five channel signals, an energy equalization mode is determined.

[0116] FIG. 4c is an example diagram depicting a structure of a pairing processing module. As shown in FIG. 5b, the pairing processing module includes a pairing determining device, an energy equalization unit, and a stereo processing box.

[0117] The pairing determining device first calculates an energy value or an amplitude value of each channel signal. In this application, the following formula may be used to calculate the energy value or the amplitude value of the channel signal(ch):

$$\text{energy}(ch) = \sqrt{\sum_{i=1}^N \text{sepc_coeff}(ch, i) \times \text{sepc_coeff}(ch, i)}$$

$\text{energy}(ch)$ is the energy value or the amplitude value of the channel signal ch , $\text{sepc_coeff}(ch, i)$ is a frequency domain coefficient of an i^{th} frequency of the channel signal ch , and N indicates an integer value that does not exceed a total frequency quantity of one audio frame.

[0118] Then, a normalized energy value or amplitude value of each channel signal is calculated. In this application, a normalized energy value or amplitude value of a channel signal(ch) may be calculated based on the following formula:

$$\text{energy_uniform}(ch) = \frac{\text{energy}(ch)}{\text{energy_max}}$$

$\text{energy_uniform}(ch)$ is the normalized energy value or amplitude value of the channel signal ch , and energy_max is a maximum value of energy values or amplitude values of the five channel signals (that is, $\text{energy}(L)$, $\text{energy}(R)$, $\text{energy}(C)$, $\text{energy}(LS)$, and $\text{energy}(RS)$). If $\text{energy_max} = 0$, all values of $\text{energy_uniform}(ch)$ are 0.

[0119] Next, the fluctuation interval value of the five channel signals is calculated. Optionally, the fluctuation interval value may be the energy flatness. In this application, energy flatness of five channel signals may be calculated based on the following formula:

$$\text{efm} = \frac{\sqrt[5]{\prod_{ch=0}^4 \text{energy_uniform}(ch)}}{\frac{1}{5} \sum_{ch=0}^4 \text{energy_uniform}(ch)}$$

efm is the energy flatness of the five channel signals. For a channel indexes of L , R , C , LS , and RS , refer to Table 1.

[0120] Optionally, the fluctuation interval value may alternatively be energy deviation. Based on the normalized energy value or amplitude value $\text{energy_uniform}(ch)$ obtained through the foregoing calculation, in this application, an average energy value or amplitude value of the five channel signals may be calculated based on the following formula:

$$avg_energy_uniform = \frac{1}{5} \times \sum_{ch=0}^4 energy_uniform(ch)$$

avg_energy_uniform is the average energy value or amplitude value of the five channel signals. For a channel index of L, R, C, LS, and RS, refer to Table 1.

[0121] Energy deviation of a channel signal(ch) is calculated based on the following formula:

$$deviation(ch) = \frac{energy_uniform(ch)}{avg_energy_uniform}$$

deviation(ch) is the energy deviation of the channel signal ch. A maximum value of the energy deviation of L, R, C, LS, and RS is determined as the energy deviation (deviation) of the five channel signals.

[0122] Optionally, the fluctuation interval value may alternatively be an amplitude value or amplitude deviation. A principle of the fluctuation interval value is similar to the foregoing energy-related value, and details are not described herein again.

[0123] As described above, the energy equalization mode in this application includes two implementations. A Pair energy equalization mode, is to use a channel pair in the target channel pair set corresponding to the pairing manner determined by the module selection unit, so that two channel signals in one channel pair are used to obtain two equalized channel signals corresponding to the channel pair. An overall energy equalization mode is to use two channel signals in one channel pair and one channel signal that is not in the channel pair to obtain two equalized channel signals corresponding to the channel pair. For a channel signal that is not paired, a corresponding equalized channel signal is the channel signal itself.

[0124] The pairing determining device determines the energy equalization mode based on the fluctuation interval value in the following two determining manners:

(1) When efm is less than the first threshold, the energy equalization mode is the Pair energy equalization mode, or when efm is greater than or equal to the first threshold, the energy equalization mode is the overall energy equalization mode.

(2) When deviation is within a value range [threshold, 1/threshold], the energy equalization mode is the overall energy equalization mode, or when deviation is not within the value range [threshold, 1/threshold], the energy equalization mode is the Pair energy equalization mode. A value range of threshold may be (0, 1).

[0125] Herein, deviation may represent a ratio of frequency domain amplitude of each channel in a current frame to an average value of frequency domain amplitude of all channels in the current frame, that is, the amplitude deviation. When a proportion between frequency domain amplitude of a current channel in a current frame and the average value of the frequency domain amplitude of all channels in the current frame is less than 5 (corresponding to threshold = 0.2), there may be two cases: 1. The frequency domain amplitude of the current channel is less than or equal to the average value of the frequency domain amplitude of all channels in the current frame, and "the frequency domain amplitude of the current channel/the average value of the frequency domain amplitude of all channels in the current frame" that meets the condition is between (0.2, 1], that is, between (threshold, 1]. 2. The frequency domain amplitude of the current channel is greater than the average value of the frequency domain amplitude of all channels in the current frame, and "the frequency domain amplitude of the current channel/the average value of the frequency domain amplitude of all channels in the current frame" that meets the condition is between (1, 5). In combination with the foregoing two cases, when the proportion between the frequency domain amplitude of the current channel and the average value of the frequency domain amplitude of all channels in the current frame is less than 5, a range of "the frequency domain amplitude of the current channel/the average value of the frequency domain amplitude of all channels in the current frame" that meets the condition is between (0.2, 5), that is, between (threshold, 1/threshold), and (threshold, 1/threshold) is the second preset range. The value of threshold may be between (0, 1). A smaller value of threshold indicates larger fluctuation of the frequency domain amplitude of the current channel relative to the average value of the frequency domain amplitude of all channels in the current frame, and a larger value of threshold indicates smaller fluctuation of the frequency domain amplitude of the current channel relative to the average value of the frequency domain amplitude of all channels in the current frame. The value of threshold may be 0.2, 0.15, 0.125, 0.11, 0.1, or the like.

[0126] Alternatively, deviation may represent a ratio of frequency domain energy of each channel to an average value of frequency domain energy of all channels, that is, energy deviation. When a proportion between frequency domain energy of a current channel in a current frame and an average value of frequency domain energy of all channels in the

current frame is less than 25 (threshold = 0.04), there may be two cases: 1. The frequency domain energy of the current channel is less than or equal to the average value of the frequency domain energy of all channels in the current frame, and "the frequency domain energy of the current channel/the average value of the frequency domain energy of all channels in the current frame" that meets the condition is between (0.04, 1], that is, between (threshold, 1]. 2. The frequency domain energy of the current channel is greater than the average value of the frequency domain energy of all channels in the current frame, and "the frequency domain energy of the current channel/the average value of the frequency domain energy of all channels in the current frame" that meets the condition is between (1, 25). In combination with the foregoing two cases, when the proportion between the frequency domain energy of the current channel and the average value of the frequency domain energy of all channels in the current frame is less than 25, the range of "the frequency domain energy of the current channel/the average value of the frequency domain energy of all channels in the current frame" that meets the condition is between (0.04, 25), that is, between (threshold, 1/threshold), and (threshold, 1/threshold) is the first preset range. Herein, threshold may be between (0, 1). A smaller value of threshold indicates larger fluctuation of the frequency domain energy of the current channel relative to the average value of the frequency domain energy of all channels in the current frame, and a larger value of threshold indicates smaller fluctuation of the frequency domain energy of the current channel relative to the average value of the frequency domain energy of all channels in the current frame. The value of threshold may be 0.04, 0.0225, 0.015625, 0.0121, 0.01 or the like.

[0127] Because there is a square relationship between the amplitude and the energy, there is also a square relationship between the amplitude deviation and the energy deviation, that is, fluctuation of inter-channel frame amplitude corresponding to a square of the amplitude deviation is approximately equivalent to fluctuation of inter-channel frame energy corresponding to the energy deviation.

[0128] In another implementation, the first preset range may alternatively be expanded to (0, 1/threshold). In this case, a range of Pair energy equalization is [1/threshold, +∞), indicating that Pair energy equalization is performed when the frequency domain energy of the current channel is greater than the average value of the frequency domain energy of all channels in the current frame, and "the frequency domain energy of the current channel/the average value of the frequency domain energy of all channels in the current frame" is greater than 1/threshold.

[0129] In another implementation, the second preset range may alternatively be expanded to (0, 1/threshold). In this case, a range of Pair amplitude equalization is [1/threshold, +∞), indicating that Pair amplitude equalization is performed when the frequency domain amplitude of the current channel is greater than the average value of the frequency domain amplitude of all channels in the current frame, and "the frequency domain amplitude of the current channel/the average value of the frequency domain amplitude of all channels in the current frame" is greater than 1/threshold.

[0130] It should be noted that the pairing determining device may calculate normalized energy values or amplitude values based on the five channel signals to obtain the energy flatness or energy deviation, or calculate normalized energy values or amplitude values only based on successfully paired channel signals to obtain the energy flatness or energy deviation, or calculate normalized energy values or amplitude values based on some of the five channel signals to obtain the energy flatness or energy deviation. This is not specifically limited in this application.

[0131] A stereo processing unit may use prediction-based or Karhunen-Loeve transform (Karhunen-Loeve Transform, KLT)-based processing, that is, two input channel signals are rotated (for example, by using a 2×2 rotation matrix) to maximize energy compression, to concentrate signal energy in one channel.

[0132] After processing the two input channel signals, the stereo processing unit outputs processed channel signals (P1 to P4) corresponding to the two channel signals and multi-channel side information, and the multi-channel side information includes a sum of correlation values and a target channel pair set.

[0133] FIG. 5 is an example diagram depicting a structure of a decoding apparatus to which a multi-channel audio decoding method is applied according to this application. The decoding apparatus may be the decoder 30 of the destination device 14 in the audio transcoding system 10, or may be the transcoding module 270 in the audio transcoding device 200. The decoding apparatus may include a bitstream de-multiplexing interface, a channel decoding module, and a multi-channel processing module.

[0134] The bitstream de-multiplexing interface receives an encoded multi-channel signal (for example, a serial bitstream bitstream) from an encoding apparatus, and obtains encoded channel signals (E) and multi-channel parameters (SIDE_PAIR) after de-multiplexing, for example, E1, E2, E3, E4, ..., Ei1, and Ei; and SIDE_PAIR1, SIDE_PAIR2, ..., and SIDE_PAIRm.

[0135] The channel decoding module uses mono-channel decoding units (or mono-channel channel boxes or mono-channel tools) to decode the encoded channel signals output by the bitstream de-multiplexing interface, and output decoded channel signals (D). For example, E1, E2, E3, E4, ..., Ei1, and Ei are decoded by the mono-channel decoding units to obtain D1, D2, D3, D4, ..., Di1, and Di.

[0136] The multi-channel processing module includes a plurality of stereo processing units. The stereo processing unit may use prediction-based or KLT-based processing, that is, two input channel signals are reversely rotated (for example, by using a 2 × 2 rotation matrix), to convert the signals to an original signal direction.

[0137] That two specific decoded channel signals in the decoded channel signals output by the channel decoding

module that are to be paired can be identified based on the multi-channel parameter, and the paired decoded channel signals are input into the stereo processing unit. After processing the two input decoded channel signals, the stereo processing unit outputs channel signals(ch) corresponding to the two decoded channel signals. For example, a stereo processing unit 1 processes D1 and D2 based on SIDE_PAIR1 to obtain CH1 and CH2, a stereo processing unit 2 processes D3 and D4 based on SIDE_PAIR2 to obtain CH3 and CH4, ..., and a stereo processing unit m processes Di-1 and Di based on SIDE_PAIRm to obtain CHi-1 and CHi.

[0138] It should be noted that a channel signal (for example, CHj) that is not paired does not need to be processed by a stereo processing unit in the multi-channel processing module, and may be directly output after being decoded.

[0139] FIG. 6 is a schematic diagram depicting a structure of an encoding apparatus according to an embodiment of this application. As shown in FIG. 6, the apparatus may be applied to the source device 12 or the audio transcoding device 200 in the foregoing embodiments. The encoding apparatus in this embodiment may include: an obtaining module 601, a processing module 602, and an encoding module 603.

[0140] The obtaining module 601 is configured to: obtain a to-be-encoded first audio frame, where the first audio frame includes at least five channel signals; and obtain a sum of correlation values of a target channel pair set, where the target channel pair set is obtained for a purpose of obtaining a maximum sum of correlation values, the target channel pair set includes at least one channel pair, one channel pair includes two channel signals in the at least five channel signals, the one channel pair has one correlation value, and the correlation value indicates correlation between the two channel signals in the one channel pair. The processing module 602 is configured to: when the sum of the correlation values is greater than a preset threshold, perform energy equalization processing on the at least five channel signals to obtain at least five equalized channel signals. The encoding module 603 is configured to encode the at least five equalized channel signals.

[0141] In a possible implementation, the encoding module 603 is further configured to encode the at least five channel signals when the sum of the correlation values is less than or equal to the preset threshold.

[0142] In a possible implementation, the processing module 602 is specifically configured to obtain a fluctuation interval value of the at least five channel signals; determine an energy equalization mode based on the fluctuation interval value of the at least five channel signals; and separately perform energy equalization processing on the at least five channel signals based on the energy equalization mode to obtain at least five equalized channel signals.

[0143] In a possible implementation, the processing module 602 is specifically configured to: determine the energy equalization mode as a first energy equalization mode when the fluctuation interval value meets a preset condition; or determine the energy equalization mode as a second energy equalization mode when the fluctuation interval value does not meet a preset condition.

[0144] In a possible implementation, the fluctuation interval value includes energy flatness of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the energy flatness is less than a first threshold; or the fluctuation interval value includes amplitude flatness of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the amplitude flatness is less than a second threshold; or the fluctuation interval value includes energy deviation of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the energy deviation is not within a first preset range; or the fluctuation interval value includes amplitude deviation of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the amplitude deviation is not within a second preset range.

[0145] In a possible implementation, when the energy equalization mode is the first energy equalization mode, the processing module 602 is specifically configured to: calculate, for a current channel pair in the target channel pair set, an average value of energy values or amplitude values of two channel signals included in the current channel pair; and separately perform energy equalization processing on the two channel signals based on the average value to obtain two corresponding equalized channel signals.

[0146] In a possible implementation, when the energy equalization mode is the second energy equalization mode, the processing module 602 is specifically configured to: calculate an average value of energy values or amplitude values of the at least five channel signals; and separately perform energy equalization processing on the at least five channel signals based on the average value to obtain at least five equalized channel signals.

[0147] In a possible implementation, the processing module 602 is further configured to: determine whether an encoding bit rate corresponding to the first audio frame is greater than a bit rate threshold; and determine the energy equalization mode as the second energy equalization mode when the encoding bit rate is greater than the bit rate threshold; or determine the energy equalization mode based on the fluctuation interval value when the encoding bit rate is less than or equal to the bit rate threshold.

[0148] The apparatus in this embodiment may be configured to execute the technical solution of the method embodiment shown in FIG. 3, implementation principles and technical effects of the apparatus and the method embodiment are similar, and details are not described herein.

[0149] FIG. 7 is a schematic diagram depicting a structure of a device according to an embodiment of this application. As shown in FIG. 7, the device may be the encoding device in the foregoing embodiments. The device in this embodiment

may include a processor 701 and a memory 702, and the memory 702 is configured to store one or more programs. When the one or more programs are executed by the processor 701, the processor 701 is enabled to implement the technical solution of the method embodiment shown in FIG. 3.

[0150] In an implementation process, steps in the foregoing method embodiments may be implemented by using a hardware integrated logical circuit in the processor, or by using instructions in a form of software. The processor may be a general-purpose processor, a digital signal processor (digital signal processor, DSP), an application-specific integrated circuit (application-specific integrated circuit, ASIC), a field-programmable gate array (field-programmable gate array, FPGA) or another programmable logic device, a discrete gate or transistor logic device, or a discrete hardware component. A general purpose processor may be a microprocessor, or the processor may also be any conventional processor, or the like. The steps of the methods disclosed in this application may be directly performed by a hardware encoding processor, or may be performed by a combination of hardware and a software module in an encoding processor. The software module may be located in a mature storage medium in the art, such as a random access memory, a flash memory, a read-only memory, a programmable read-only memory, an electrically erasable programmable memory, or a register. The storage medium is located in the memory, and the processor reads information in the memory and completes the steps in the foregoing methods in combination with hardware of the processor.

[0151] The memory in the foregoing embodiments may be a volatile memory or a non-volatile memory, or may include both a volatile memory and a non-volatile memory. The non-volatile memory may be a read-only memory (read-only memory, ROM), a programmable read-only memory (programmable ROM, PROM), an erasable programmable read-only memory (erasable PROM, EPROM), an electrically erasable programmable read-only memory (electrically EPROM, EEPROM), or a flash memory. The volatile memory may be a random access memory (random access memory, RAM), used as an external cache. By way of example but not limitative description, many forms of RAMs may be used, for example, a static random access memory (static RAM, SRAM), a dynamic random access memory (dynamic RAM, DRAM), a synchronous dynamic random access memory (synchronous DRAM, SDRAM), a double data rate synchronous dynamic random access memory (double data rate SDRAM, DDR SDRAM), an enhanced synchronous dynamic random access memory (enhanced SDRAM, ESDRAM), a synchronous link dynamic random access memory (synchlink DRAM, SLDRAM), and a direct rambus dynamic random access memory (direct rambus RAM, DR RAM). It should be noted that the memory of the system and the method described in this specification includes but is not limited to these memories and any memory of another appropriate type.

[0152] Persons of ordinary skill in the art may be aware that, in combination with examples described in embodiments disclosed in this specification, units and algorithm steps may be implemented by electronic hardware or a combination of computer software and electronic hardware. Whether the functions are performed by hardware or software depends on particular applications and design constraints of the technical solutions. Persons skilled in the art may use different methods to implement the described functions for each particular application, but it should not be considered that the implementation goes beyond the scope of this application.

[0153] It may be clearly understood by persons skilled in the art that, for the purpose of convenient and brief description, for a detailed working process of the foregoing system, apparatus, and unit, refer to a corresponding process in the foregoing method embodiments. Details are not described herein again.

[0154] In the several embodiments provided in this application, it should be understood that the disclosed systems, apparatuses, and methods may be implemented in other manners. For example, the foregoing apparatus embodiments are merely examples. For example, division of the units is merely logical function division and may be other division during actual implementation. For example, a plurality of units or components may be combined or integrated into another system, or some features may be ignored or not performed. In addition, the displayed or discussed mutual couplings or direct couplings or communication connections may be implemented by using some interfaces. The indirect couplings or communication connections between the apparatuses or units may be implemented in electronic, mechanical, or another form.

[0155] The units described as separate parts may or may not be physically separate, and parts displayed as units may or may not be physical units, may be located at one position, or may be distributed on a plurality of network units. Some or all of the units may be selected depending on actual requirements to achieve the objectives of the solutions in the embodiments.

[0156] In addition, functional units in embodiments of this application may be integrated into one processing unit, or each of the units may exist alone physically, or two or more units are integrated into one unit.

[0157] When the functions are implemented in the form of a software functional unit and sold or used as an independent product, the functions may be stored in a computer-readable storage medium. Based on such an understanding, the technical solutions in this application essentially, or the part contributing to the conventional technology, or a part of the technical solutions may be implemented in a form of a software product. The computer software product is stored in a storage medium and includes several instructions for instructing a computer device (a personal computer, a server, a network device, or the like) to perform all or a part of the steps of the methods in embodiments of this application. The foregoing storage medium includes any medium that can store program encode, such as a USB flash drive, a removable

hard disk, a read-only memory (read-only memory, ROM), a random access memory (random access memory, RAM), a magnetic disk, or an optical disc.

[0158] The foregoing descriptions are merely specific implementations of this application, but the protection scope of this application is not limited thereto. Any variation or replacement readily figured out by persons skilled in the art within the technical scope disclosed in this application shall fall within the protection scope of this application. Therefore, the protection scope of this application shall be subject to the protection scope of the claims.

Claims

1. A multi-channel audio signal encoding method, comprising:

obtaining a to-be-encoded first audio frame, wherein the first audio frame comprises at least five channel signals; obtaining a sum of correlation values of all channel pairs in a target channel pair set, wherein the target channel pair set comprises at least one channel pair, one channel pair comprises two of the at least five channel signals, the one channel pair has one correlation value, and the correlation value indicates correlation between the two channel signals in the one channel pair; when the sum of the correlation values is greater than a preset threshold, performing energy equalization processing on at least two of the at least five channel signals to obtain at least two equalized channel signals; and encoding the at least two equalized channel signals to obtain an encoded bitstream.

2. The method according to claim 1, wherein the method further comprises:

when the sum of the correlation values is less than or equal to the preset threshold, encoding the at least five channel signals to obtain an encoded bitstream.

3. The method according to claim 1 or 2, wherein the performing energy equalization processing on at least two of the at least five channel signals to obtain at least two equalized channel signals comprises:

obtaining a fluctuation interval value of the at least five channel signals; determining an energy equalization mode based on the fluctuation interval value of the at least five channel signals; and separately performing energy equalization processing on the at least two channel signals based on the energy equalization mode to obtain the at least two equalized channel signals.

4. The method according to claim 3, wherein the determining an energy equalization mode based on the fluctuation interval value of the at least five channel signals comprises:

determining the energy equalization mode as a first energy equalization mode when the fluctuation interval value meets a preset condition; or determining the energy equalization mode as a second energy equalization mode when the fluctuation interval value does not meet a preset condition.

5. The method according to claim 4, wherein the fluctuation interval value comprises energy flatness of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the energy flatness is less than a first threshold; or

the fluctuation interval value comprises amplitude flatness of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the amplitude flatness is less than a second threshold; or the fluctuation interval value comprises energy deviation of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the energy deviation is not within a first preset range; or the fluctuation interval value comprises amplitude deviation of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the amplitude deviation is not within a second preset range.

6. The method according to claim 4 or 5, wherein when the energy equalization mode is the first energy equalization mode, the performing energy equalization processing on at least two of the at least five channel signals to obtain at least two equalized channel signals comprises:

performing energy equalization processing on channel signals corresponding to the target channel pair set to obtain the at least two equalized channel signals.

7. The method according to claim 6, wherein the performing energy equalization processing on channel signals corresponding to the target channel pair set to obtain the at least two equalized channel signals comprises:
calculating, for a current channel pair in the target channel pair set, an average value of energy values or amplitude values of two channel signals comprised in the current channel pair, and separately performing, based on the average value, energy equalization processing on the two channel signals comprised in the current channel pair to obtain two equalized channel signals.
8. The method according to claim 4 or 5, wherein when the energy equalization mode is the second energy equalization mode, the performing energy equalization processing on at least two of the at least five channel signals to obtain at least two equalized channel signals comprises:
calculating an average value of energy values or amplitude values of the at least five channel signals, and separately performing energy equalization processing on the at least five channel signals based on the average value to obtain at least five equalized channel signals.
9. The method according to any one of claims 3 to 8, wherein before the determining an energy equalization mode based on the fluctuation interval value of the at least five channel signals, the method further comprises:
determining whether an encoding bit rate corresponding to the first audio frame is greater than a bit rate threshold;
and
determining the energy equalization mode as the second energy equalization mode when the encoding bit rate is greater than the bit rate threshold; or
determining the energy equalization mode based on the fluctuation interval value when the encoding bit rate is less than or equal to the bit rate threshold.
10. The method according to any one of claims 1 to 9, wherein the method further comprises:
encoding a channel signal on which energy equalization processing is not performed in the at least five channel signals.
11. An encoding apparatus, comprising:
an obtaining module, configured to: obtain a to-be-encoded first audio frame, wherein the first audio frame comprises at least five channel signals; and obtain a sum of correlation values of all channel pairs in a target channel pair set, wherein the target channel pair set comprises at least one channel pair, one channel pair comprises two of the at least five channel signals, the one channel pair has one correlation value, and the correlation value indicates correlation between the two channel signals in the one channel pair;
a processing module, configured to: when the sum of the correlation values is greater than a preset threshold, perform energy equalization processing on at least two of the at least five channel signals to obtain at least two equalized channel signals; and
an encoding module, configured to encode the at least two equalized channel signals to obtain an encoded bitstream.
12. The apparatus according to claim 11, wherein the encoding module is further configured to: when the sum of the correlation values is less than or equal to the preset threshold, encode the at least five channel signals to obtain an encoded bitstream.
13. The apparatus according to claim 11 or 12, wherein the processing module is specifically configured to: obtain a fluctuation interval value of the at least five channel signals; determine an energy equalization mode based on the fluctuation interval value of the at least five channel signals; and separately perform energy equalization processing on the at least two channel signals based on the energy equalization mode to obtain the at least two equalized channel signals.
14. The apparatus according to claim 13, wherein the processing module is specifically configured to: determine the energy equalization mode as a first energy equalization mode when the fluctuation interval value meets a preset condition; or determine the energy equalization mode as a second energy equalization mode when the fluctuation interval value does not meet a preset condition.
15. The apparatus according to claim 14, wherein the fluctuation interval value comprises energy flatness of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the energy flatness is

less than a first threshold; or

the fluctuation interval value comprises amplitude flatness of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the amplitude flatness is less than a second threshold; or
 the fluctuation interval value comprises energy deviation of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the energy deviation is not within a first preset range; or
 the fluctuation interval value comprises amplitude deviation of the first audio frame, and that the fluctuation interval value meets a preset condition indicates that the amplitude deviation is not within a second preset range.

16. The apparatus according to claim 14 or 15, wherein when the energy equalization mode is the first energy equalization mode, the processing module is specifically configured to perform energy equalization processing on channel signals corresponding to the target channel pair set to obtain the at least two equalized channel signals.

17. The apparatus according to claim 16, wherein the processing module is specifically configured to calculate, for a current channel pair in the target channel pair set, an average value of energy values or amplitude values of two channel signals comprised in the current channel pair, and separately perform, based on the average value, energy equalization processing on the two channel signals comprised in the current channel pair to obtain two corresponding equalized channel signals.

18. The apparatus according to claim 14 or 15, wherein when the energy equalization mode is the second energy equalization mode, the processing module is specifically configured to: calculate an average value of energy values or amplitude values of the at least five channel signals; and separately perform energy equalization processing on the at least five channel signals based on the average value to obtain at least five equalized channel signals.

19. The apparatus according to any one of claims 13 to 18, wherein the processing module is further configured to: determine whether an encoding bit rate corresponding to the first audio frame is greater than a bit rate threshold; and determine the energy equalization mode as the second energy equalization mode when the encoding bit rate is greater than the bit rate threshold; or determine the energy equalization mode based on the fluctuation interval value when the encoding bit rate is less than or equal to the bit rate threshold.

20. The apparatus according to any one of claims 11 to 19, wherein the encoding module is further configured to encode a channel signal on which energy equalization processing is not performed in the at least five channel signals.

21. A device, comprising:

one or more processors; and
 a memory, configured to store one or more programs, wherein
 when the one or more programs are executed by the one or more processors, the one or more processors are enabled to implement the method according to any one of claims 1 to 10.

22. A computer-readable storage medium, comprising a computer program, wherein when the computer program is executed on a computer, the computer is enabled to perform the method according to any one of claims 1 to 10.

23. A computer-readable storage medium, comprising a bitstream obtained by using the multi-channel audio signal encoding method according to any one of claims 1 to 10.

10

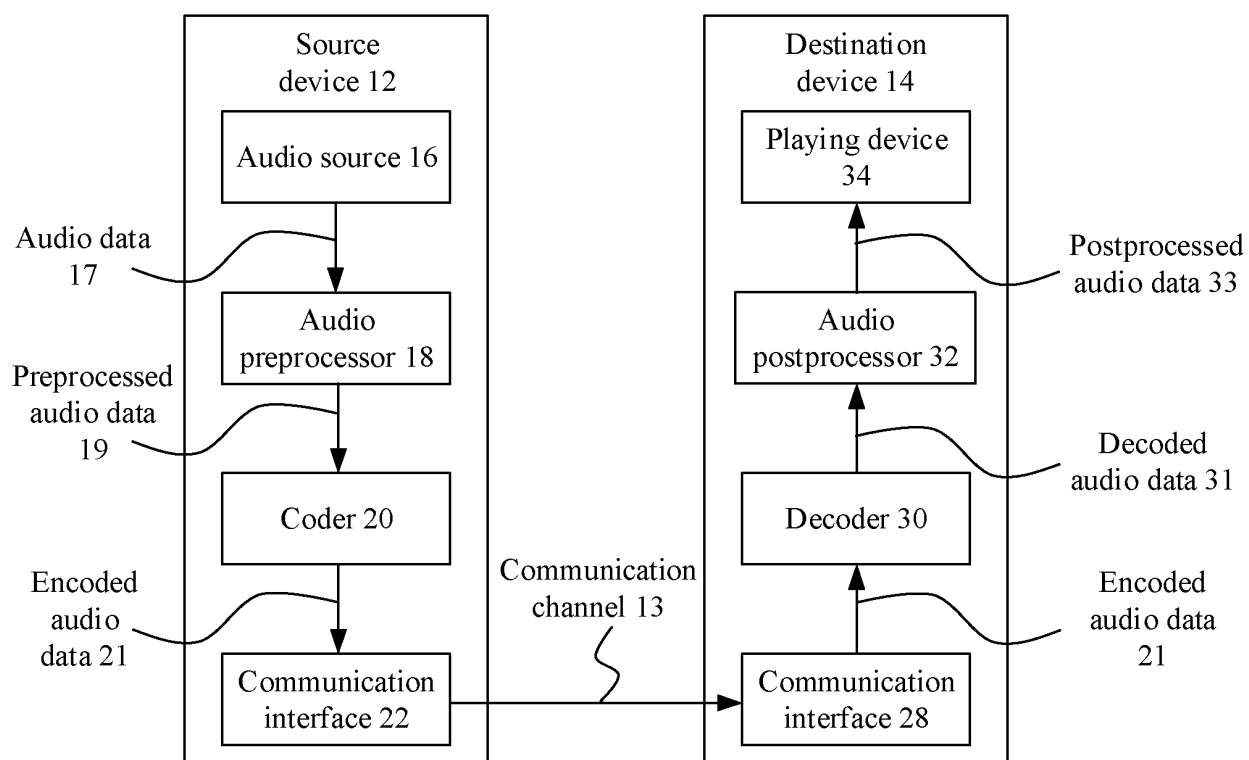


FIG. 1

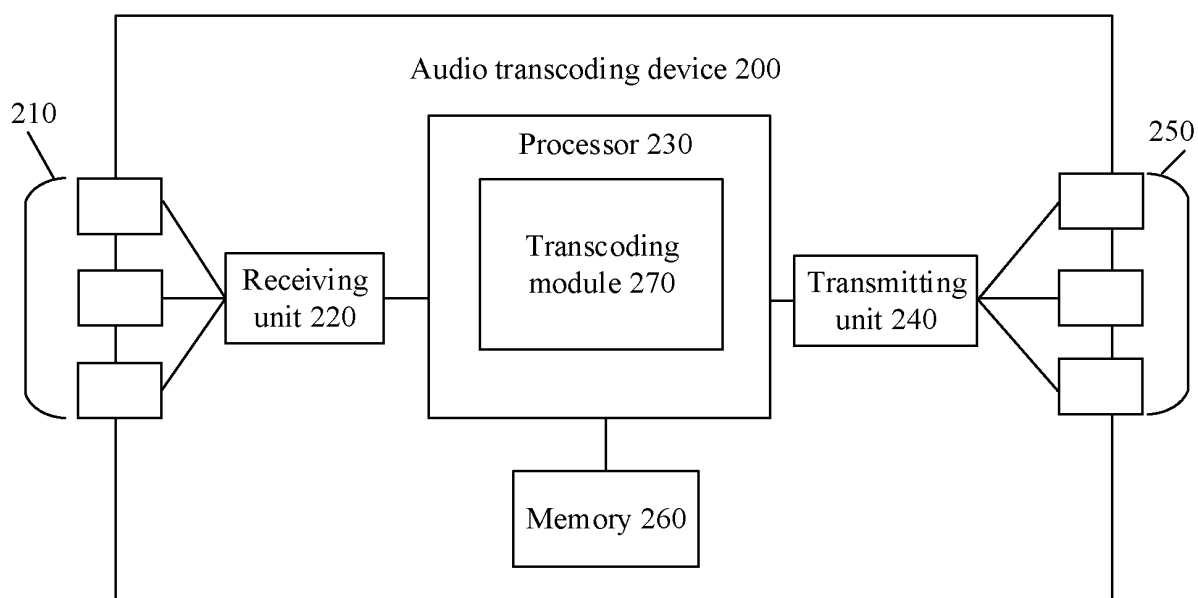


FIG. 2

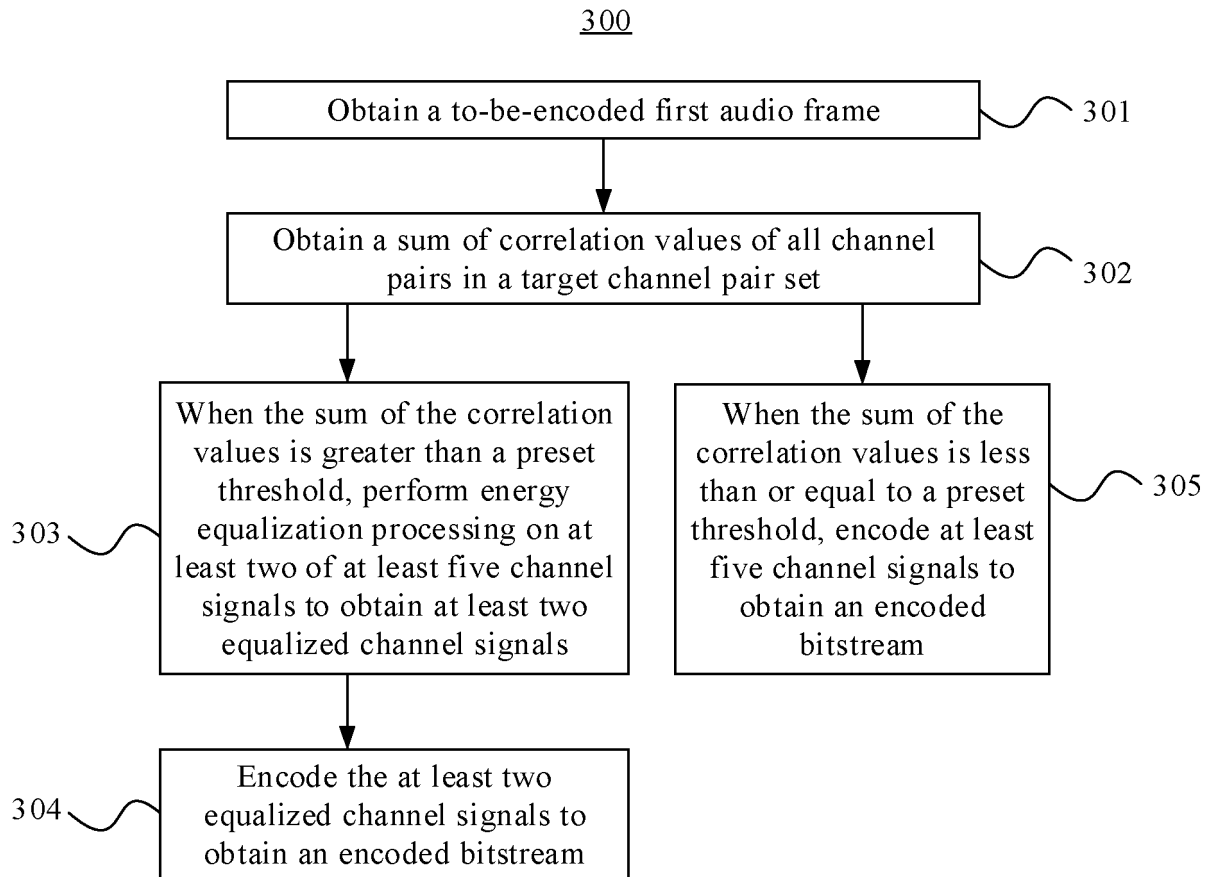


FIG. 3

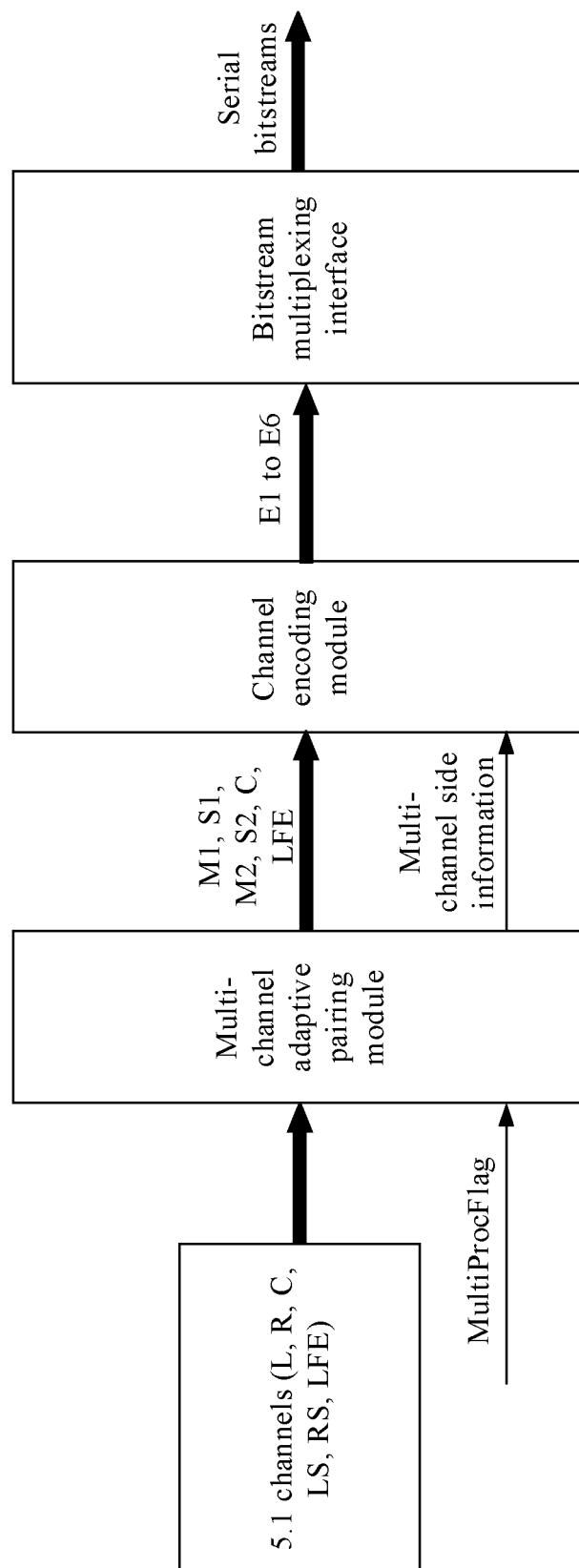


FIG. 4a

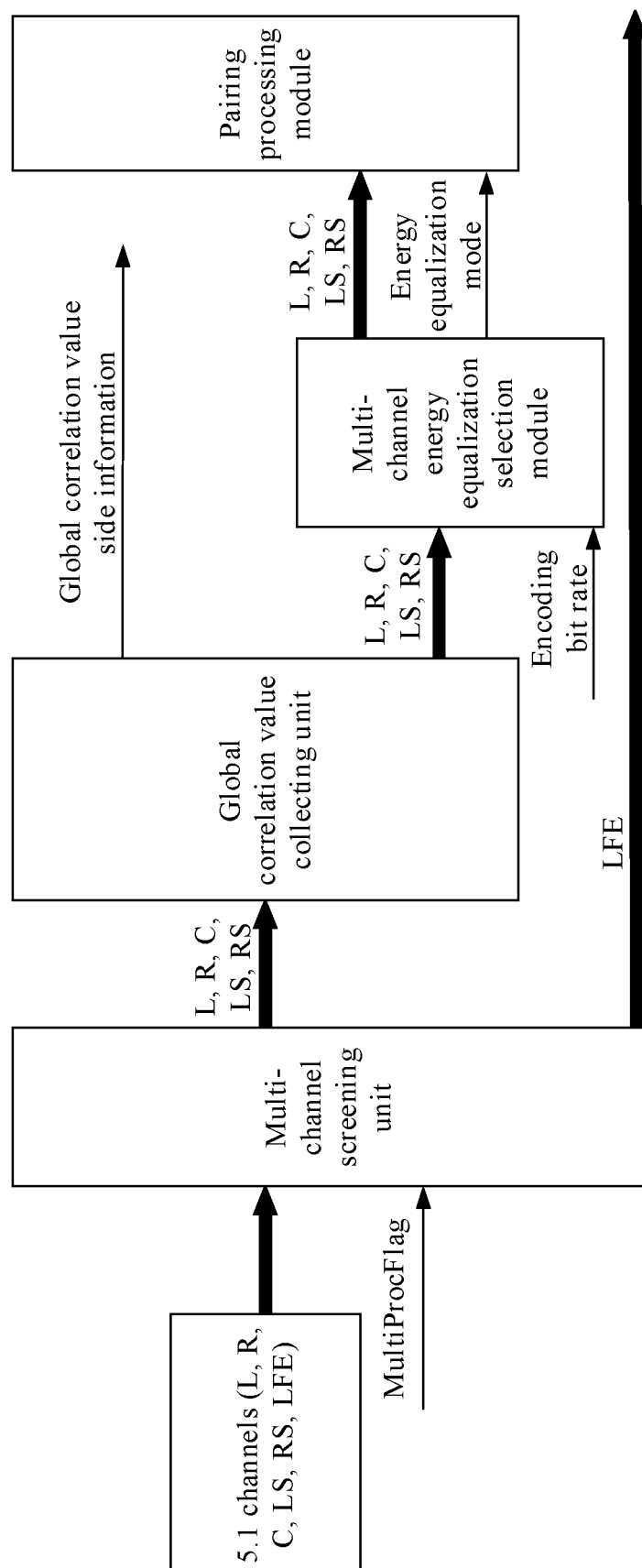


FIG. 4b

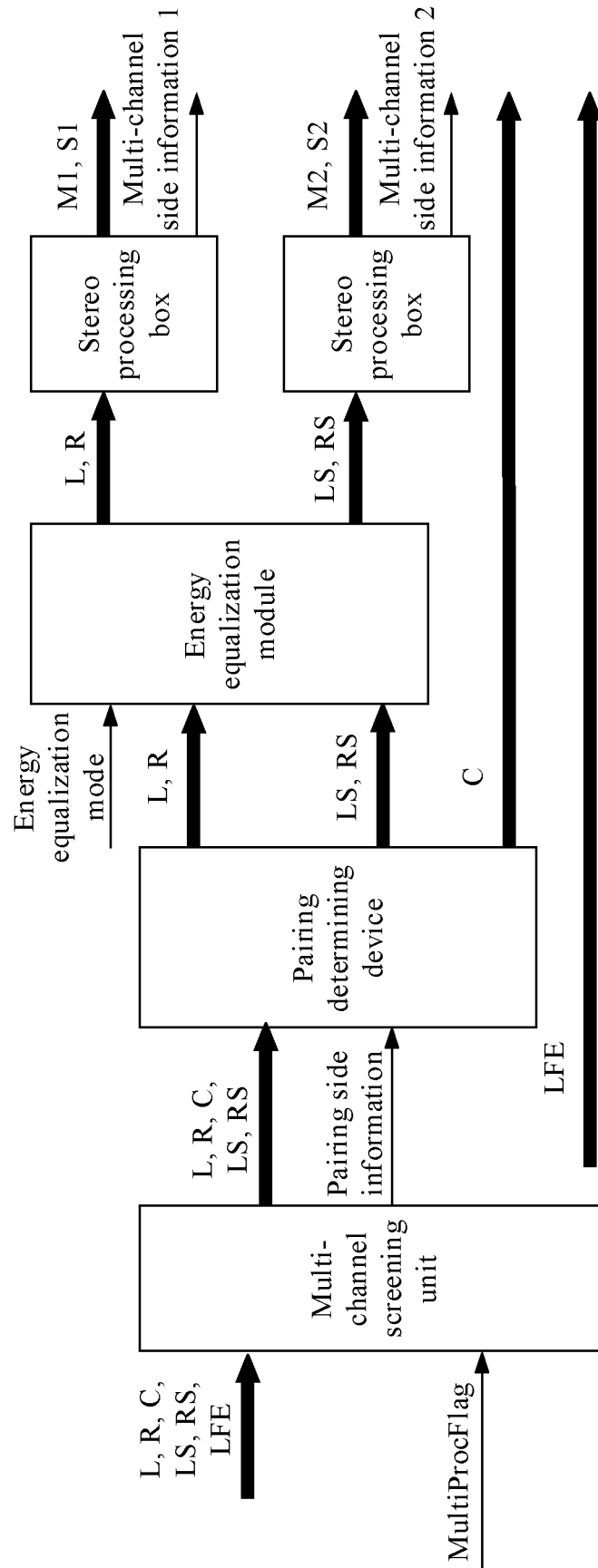


FIG. 4c

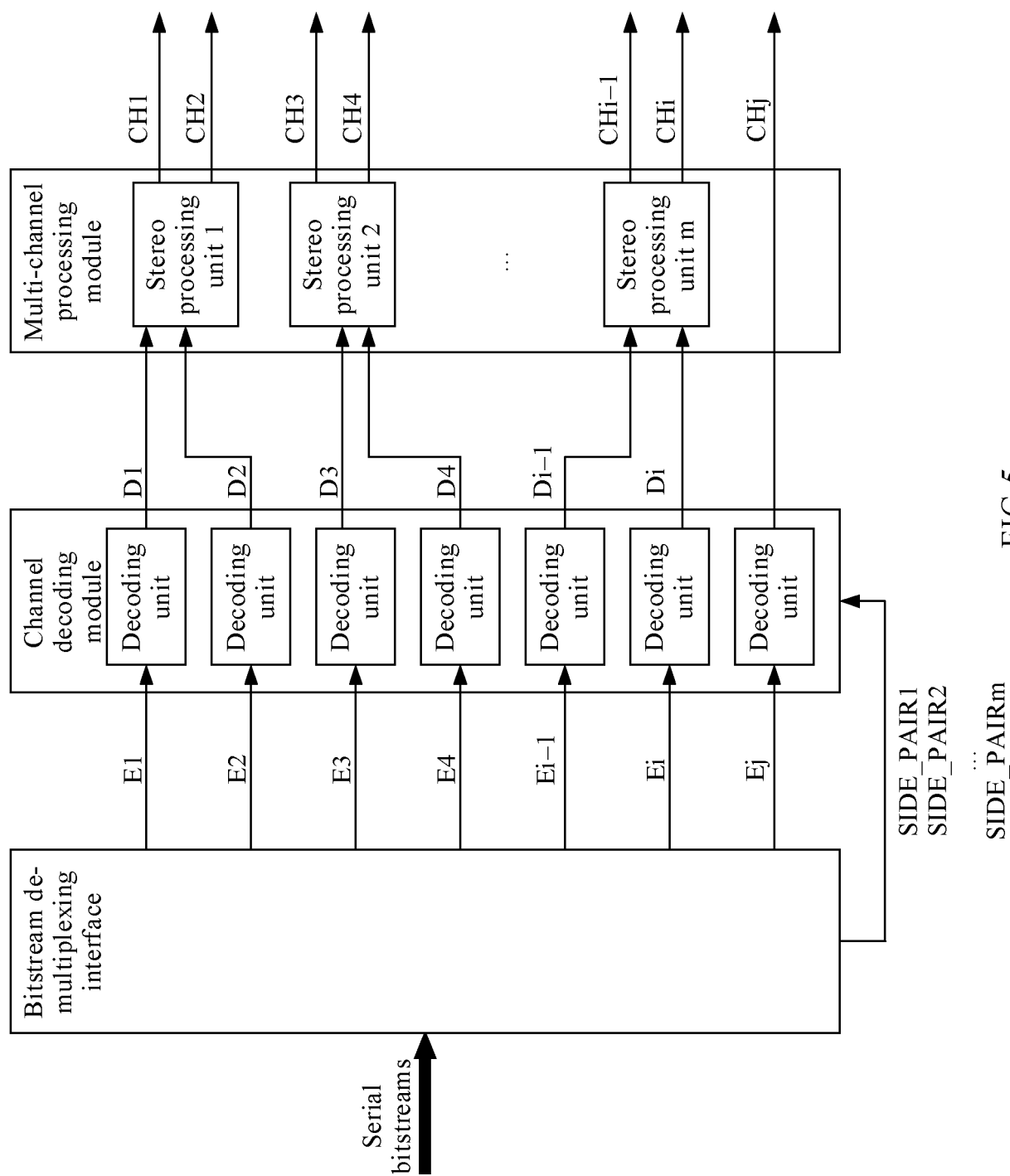


FIG. 5

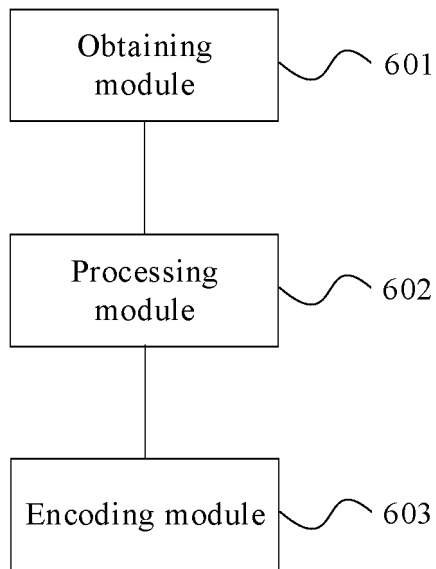


FIG. 6

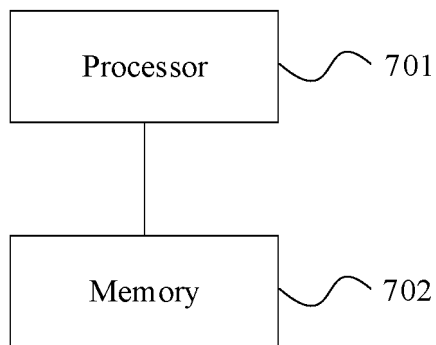


FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/092518

A. CLASSIFICATION OF SUBJECT MATTER

G10L 19/008(2013.01)i; G10L 19/00(2013.01)i; H04S 3/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L; H04S

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS; CNTXT; CNKI; VEN; WOTXT; EPTXT; USTXT: 音频, 帧, 编码, 相关性, 相关值, 相关系数, 相关度, 相干性, 相加, 和, 多, 五, 七, 5.1, 7.1, 22.2, 信道, 声道, 通道, 音道, 音轨, 声路, 能量, 功率, 幅度, 均衡, 平均, speech, sound, voice, surround, multi-channel, multichannel, five, seven, multiple, channel?, encod+, cod+, correlation, ICC, coherent, cross-correlation, coherence, balanc+, energy, threshold value

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 107592937 A (FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E. V.) 16 January 2018 (2018-01-16) description, paragraphs [0005]-[0194], and figures 1-6	1-23
A	EP 1175030 A2 (NOKIA MOBILE PHONES LTD.) 23 January 2002 (2002-01-23) entire document	1-23
A	WO 2009146734 A1 (NOKIA CORP. et al.) 10 December 2009 (2009-12-10) entire document	1-23
A	US 2012029916 A1 (TSUJIKAWA, M. et al.) 02 February 2012 (2012-02-02) entire document	1-23
A	WO 2015009040 A1 (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 22 January 2015 (2015-01-22) entire document	1-23
A	US 2004109471 A1 (TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)) 10 June 2004 (2004-06-10) entire document	1-23

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

12 July 2022

Date of mailing of the international search report

27 July 2022

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088, China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/092518

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2020168232 A1 (PANASONIC INTELLECTUAL PROPERTY CORPORATION OF AMERICA) 28 May 2020 (2020-05-28) entire document	1-23
A	JP H0964751 A (SAMSUNG ELECTRONIC CO., LTD.) 07 March 1997 (1997-03-07) entire document	1-23

Form PCT/ISA/210 (second sheet) (January 2015)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/092518

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 107592937 A	16 January 2018	EP 3067885 A1	14 September 2016
		CA 2978818 A1	15 September 2016
		WO 2016142375 A1	15 September 2016
		TW 201642248 A	01 December 2016
		TW 584271 B1	21 May 2017
		AR 103873 A1	07 June 2017
		AU 2016231238 A1	21 September 2017
		IN 201737030658 A	13 October 2017
		SG 11201707180 A1	30 October 2017
		KR 20170130458 A	28 November 2017
		VN 55341 A	25 December 2017
		EP 3268959 A1	17 January 2018
		MX 2017011495 A1	25 January 2018
		ID 201801507 A	16 February 2018
		US 2018090151 A1	29 March 2018
		BR 112017019187 A2	24 April 2018
		JP 2018513402 W	24 May 2018
		HK 1241552 A0	08 June 2018
		AU 2016231238 B2	02 August 2018
		ZA 201706602 A	31 October 2018
		SG 11201707180 B	02 March 2019
		RU 2017134964 A	05 April 2019
		MX 364419 B	25 April 2019
		EP 3268959 B1	14 August 2019
		US 10388289 B2	20 August 2019
		JP 6600004 B2	30 October 2019
		US 2019333524 A1	31 October 2019
		RU 2711055 C2	14 January 2020
		JP 2020034920 A	05 March 2020
		KR 102109159 B1	12 May 2020
		ES 2769032 T3	24 June 2020
		HK 40010211 A0	03 July 2020
		US 10762909 B2	01 September 2020
		CA 2978818 C	22 September 2020
		HK 1241552 A1	25 September 2020
		US 2021012783 A1	14 January 2021
		CN 112233684 A	15 January 2021
		CN 107592937 B	23 February 2021
EP 1175030 A2	23 January 2002	DE 60132853 D1	03 April 2008
		EP 1175030 A3	23 October 2002
		EP 1175030 B1	20 February 2008
WO 2009146734 A1	10 December 2009	None	
US 2012029916 A1	02 February 2012	JP WO2010092915 A1	16 August 2012
		WO 2010092915 A1	19 August 2010
		US 9064499 B2	23 June 2015
		JP 5605575 B2	15 October 2014
WO 2015009040 A1	22 January 2015	KR 20150009474 A	26 January 2015
		EP 3023984 A1	25 May 2016
		EP 3023984 A4	08 March 2017

Form PCT/ISA/210 (patent family annex) (January 2015)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/092518

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
US 2004109471 A1	10 June 2004	SE 200003285 A	16 March 2002
		WO 0223528 A1	21 March 2002
		AU 8458801 A	26 March 2002
		SE 519981 C2	06 May 2003
		EP 1320849 A1	25 June 2003
		JP 2004509366 A	25 March 2004
		EP 1320849 B1	30 May 2007
		DE 60128711 D1	12 July 2007
		US 7283957 B2	16 October 2007
		DE 60128711 T2	07 February 2008
		JP 4485123 B2	16 June 2010
US 2020168232 A1	28 May 2020	JP WO2018221138 A1	02 April 2020
		WO 2018221138 A1	06 December 2018
		US 11145316 B2	12 October 2021
JP H0964751 A	07 March 1997	JP 3342996 B2	11 November 2002

Form PCT/ISA/210 (patent family annex) (January 2015)

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- CN 202110595367 [0001]