SIGNAL ENCODING AND DECODING METHOD AND APPARATUS, AND USER EQUIPMENT, NETWORK SIDE DEVICE AND STORAGE MEDIUM

TECHNICAL FIELD

[0001] The disclosure relates to the field of communication technologies, in particular to a signal encoding and decoding method and apparatus, an encoding device, a decoding device and a storage medium.

BACKGROUND

[0002] Since a 3D audio can enable users to have better stereoscopic perception and spatial immersion experience, the 3D audio has been widely used. When creating the end-to-end 3D audio experience, mixed-format audio signals are usually collected at an acquisition end, and the mixed-format audio signals may include, for example, at least two formats of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal, and then the collected signals may be encoded and decoded, and finally rendered into binaural signals or multi-speaker signals according to a capability of a playback device (such as the terminal capability) for playback.

[0003] In the related art, a method for encoding a mixed-format audio signal may refer to processing each format by a corresponding encoding kernel, that is, the channel-based audio signal is processed using a channel signal encoding kernel, and the object-based audio signal is processed using an object signal encoding kernel, the scene-based audio signal is processed using a scene signal encoding kernel.

[0004] However, in the related art, when performing the encoding, parameter information such as the control information of the encoding end, the characteristic of the input mixed-format audio signal, the advantages and disadvantages of audio signals in different formats, and the actual playback requirement of the playback end are not considered, resulting in a low encoding efficiency for the mixed-format audio signal.

SUMMARY

[0005] A signal encoding and decoding method and apparatus, a user equipment (UE), a network side device, and a storage medium proposed in the disclosure aim to solve a technical problem of low data compression rate and inability to save bandwidth caused by

the encoding method in the related art.

[0006] An aspect of embodiments of the disclosure provides a signal encoding and decoding method, which is applied to an encoding end. The method includes:

[0007] obtaining an audio signal in a mixed format, in which the audio signal in the mixed format comprises at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal;

[0008] determining, based on signal characteristics of audio signals in different formats, an encoding mode of the audio signal in each format; and

[0009] encoding the audio signal in each format using the encoding mode of the audio signal in each format to obtain encoded signal parameter information of the audio signal in each format, writing the encoded signal parameter information of the audio signal in each format into an encoded stream and sending the encoded stream to a decoding end.

[0010] Another aspect of the embodiments of the disclosure provides a signal encoding and decoding method, which is applied to a decoding end. The method includes:

[0011] receiving an encoded stream sent by an encoding end; and

[0012] decoding the encoded stream to obtain an audio signal in a mixed format, in which the audio signal in the mixed format comprises at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[0013] Another aspect of embodiments of the disclosure provides a signal encoding and decoding apparatus. The apparatus includes:

[0014] an obtaining module, configured to obtain an audio signal in a mixed format, in which the audio signal in the mixed format comprises at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal;

[0015] a determining module, configured to determine, based on signal characteristics of audio signals in different formats, an encoding mode of the audio signal in each format; and

[0016] an encoding module, configured to encode the audio signal in each format using the encoding mode of the audio signal in each format to obtain encoded signal parameter information of the audio signal in each format, write the encoded signal parameter information of the audio signal in each format into an encoded stream and send the encoded stream to a decoding end.

[0017] Another aspect of embodiments of the disclosure provides a signal encoding and

decoding apparatus. The apparatus includes:

[0018] a receiving module, configured to receive an encoded stream sent by an encoding end; and

[0019] a decoding module, configured to decode the encoded stream to obtain an audio signal in a mixed format, in which the audio signal in the mixed format comprises at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[0020] Another aspect of embodiments of the disclosure provides a communication apparatus. The communication apparatus includes: a processor and a memory. A computer program is stored in the memory, and the processor is configured to execute the computer program stored in the memory, to cause the apparatus to perform the method described in the aspect of the embodiments of the disclosure.

[0021] Another aspect of embodiments of the disclosure provides a communication apparatus. The communication apparatus includes: a processor and a memory. A computer program is stored in the memory, and the processor is configured to execute the computer program stored in the memory, to cause the apparatus to perform the method described in another aspect of the embodiments of the disclosure.

[0022] Another aspect of embodiments of the disclosure provides a communication apparatus including a processor and an interface circuit.

[0023] The interface circuit is configured to receive code instructions and transmit the code instructions to the processor.

[0024] The processor is configured to run the code instructions to execute the method described in the aspect of the embodiments of the disclosure.

[0025] Another aspect of embodiments of the disclosure provides a communication apparatus including a processor and an interface circuit.

[0026] The interface circuit is configured to receive code instructions and transmit the code instructions to the processor.

[0027] The processor is configured to run the code instructions to execute the method described in another aspect of the embodiments of the disclosure.

[0028] Another aspect of embodiments of the disclosure provides a computer-readable storage medium for storing instructions. When the instructions are executed, the method described in the aspect of the embodiments of the disclosure is implemented.

[0029] Another aspect of embodiments of the disclosure provides a computer-readable storage medium for storing instructions. When the instructions are executed, the method described in another aspect of the embodiments of the disclosure is implemented.

[0030] To sum up, in the signal encoding and decoding method and apparatus, the encoding device, the decoding device, and the storage medium provided by the embodiments of the present disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiments of the present disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

BRIEF DESCRIPTION OF THE DRAWINGS

[0031] The above and/or additional aspects and advantages of the present disclosure will become apparent and understandable from the following description of the embodiments in combination with the accompanying drawings, in which:

FIG. 1a is a flowchart of an encoding and decoding method provided by an embodiment of the disclosure;

FIG. 1b is a schematic diagram of a collection layout of microphones of a collection end provided by an embodiment of the disclosure;

FIG. 1c is a schematic diagram of a playback layout of speakers of a playback end corresponding to FIG. 1b provided by an embodiment of the disclosure;

FIG. 2a is a flowchart of another encoding and decoding method provided by an embodiment of the disclosure;

FIG. 2b is a flowchart of a signal encoding method provided by an embodiment of the disclosure;

FIG. 3 is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 4a is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 4b is a flowchart of a signal encoding method for an object-based audio signal provided by an embodiment of the disclosure;

FIG. 5a is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 5b is a flowchart of another signal encoding method for an object-based audio signal provided by an embodiment of the disclosure;

FIG. 6a is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 6b is a flowchart of another signal encoding method for an object-based audio signal provided by an embodiment of the disclosure;

FIG. 7a is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 7b is a diagram of an algebraic codebook excited linear prediction (ACELP) encoding principle provided by another embodiment of the disclosure;

FIG. 7c is a diagram of a frequency domain encoding principle provided by another embodiment of the disclosure;

FIG. 7d is a flowchart of an encoding method for a second type of object signal set according to an embodiment of the disclosure;

FIG. 8a is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 8b is a flowchart of another encoding method for a second type of object signal set provided by an embodiment of the disclosure;

FIG. 9a is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 9b is a flowchart of another encoding method for a second type of object signal set provided by an embodiment of the disclosure;

FIG. 10 is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 11a is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 11b is a flowchart of a signal decoding method provided by an embodiment of the disclosure;

FIG. 12a is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 12b, 12c to 12d are flowcharts of a decoding method for an object-based audio signal provided by an embodiment of the disclosure;

FIG. 12e to 12f are flowcharts of a decoding method for a second type of object signal set provided by an embodiment of the disclosure;

FIG. 13 is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 14 is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 15 is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 16 is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 17 is a flowchart of an encoding and decoding method provided by another embodiment of the disclosure;

FIG. 18 is a block diagram of an encoding and decoding apparatus provided by an embodiment of the disclosure;

FIG. 19 is a block diagram of an encoding and decoding apparatus provided by another embodiment of the disclosure;

FIG. 20 is a block diagram of a user equipment provided by an embodiment of the disclosure; and

FIG. 21 is a block diagram of a network side device provided by an embodiment of the disclosure.

DETAILED DESCRIPTION

[0032] The technical solutions in the embodiments of the disclosure will be clearly and completely described below with reference to the accompanying drawings in the embodiments of the disclosure. Obviously, the described embodiments are only part of the embodiments of the disclosure, and not all of the embodiments. Based on the embodiments in the disclosure, other embodiments obtained by those skilled in the art without inventive work fall within the scope of protection of this disclosure.

[0033] The terms used in the disclosure are only for the purpose of describing specific embodiments, and are not intended to limit the embodiments of the disclosure. The singular forms of “a” and “the” used in the disclosure and appended claims are also intended to include plural forms, unless the context clearly indicates other meanings. It should also be understood that the term “and/or” as used herein refers to and includes any or all possible combinations of one or more associated listed items.

[0034] It is understandable that although the terms “first”, “second”, and “third” may be used in the embodiments of the disclosure to describe various information, the information should not be limited to these terms. These terms are only used to distinguish the same type of information from each other. For example, without departing from the scope of the disclosure, the first information may also be referred to as the second information, and similarly, the second information may also be referred to as the first information. Depending on the context, the term “if” as used herein can be interpreted as “when”, “while” or “in response to determining”.

[0035] The encoding and decoding method and apparatus, the user equipment, the network side device, and the storage medium provided by the embodiments of the present disclosure will be described in detail below with reference to the accompanying drawings.

[0036] FIG. 1 is a flowchart of a signal encoding and decoding method according to an embodiment of the disclosure. The method is performed by an encoding end. As illustrated in FIG. 1, the signal encoding and decoding method may include the following steps.

[0037] At step 101, an audio signal in a mixed format is obtained. The audio signal in the mixed format includes at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[0038] In an embodiment of the disclosure, the encoding end may be a user equipment (UE, terminal device) or a base station, and the UE may be a device that provides voice and/or data connectivity to a user. The terminal device can communicate with one or more

core networks via a radio access network (RAN). The UE can be an Internet of Things (IoT) terminal, such as a sensor device, a mobile phone (or called "cellular" phone) and a computer with the IoT terminal, for example, may be a fixed, portable, pocket, hand-held, built-in computer or vehicle-mounted device, for example, a station (STA), a subscriber unit, a subscriber station, a mobile station, a mobile, a remote station, an access point, a remote terminal, an access terminal, a user terminal, or a user agent. Alternatively, the UE may also be a device of an unmanned aerial vehicle. Alternatively, the UE may also be a vehicle-mounted device, for example, may be a trip computer with a wireless communication function, or a wireless terminal connected externally to the trip computer. Alternatively, the UE may also be a roadside device, for example, may be a street lamp, a signal lamp, or other roadside devices with a wireless communication function.

[0039] In an embodiment of the present disclosure, the above-mentioned three formats of audio signals are specifically distinguished based on collection formats of signals, and the application scenarios of the audio signals in different formats will also be different.

[0040] Specifically, in an embodiment of the disclosure, a main application scenario of the above-mentioned channel-based audio signal may be a scenario in which collection layout of microphones and a playback layout of speakers that are the same are respectively pre-set at the collection end and the playback end. For example, FIG. 1b is a schematic diagram of the collection layout of microphones at the collection end provided by an embodiment of the disclosure, which can be used to collect channel-based audio signals in a 5.0 format. FIG. 1c is a schematic diagram of the playback layout of speakers at the playback end corresponding to FIG. 1b provided by an embodiment of the disclosure, which can play back the channel-based audio signals in the 5.0 format collected by the collection end in FIG. 1b.

[0041] In another embodiment of the disclosure, the above-mentioned object-based audio signal is typically obtained by performing sound recording on a sounding object using an independent microphone, and a main application scenario of the above-mentioned object-based audio signal may be a scenario in which independent control operations need to be performed on the audio signal at the playback end, such as sound switch, volume adjustment, sound image orientation adjustment, frequency band equalization processing and other control operations.

[0042] In another embodiment of the disclosure, a main application scenario of the

above-mentioned scene-based audio signal may be a scenario in which a complete sound field where the collection end is located needs to be recorded, such as live recording of a concert, live recording of a football game, and the like.

[0043] At step 102, based on signal characteristics of audio signals in different formats, an encoding mode of the audio signal in each format is determined.

[0044] In an embodiment of the disclosure, the above-mentioned step "determining, based on signal characteristics of audio signals in different formats, an encoding mode of the audio signal in each format" may include: determining an encoding mode of the channel-based audio signal based on the signal characteristic of the channel-based audio signal; determining an encoding mode of the object-based audio signal based on the signal characteristic of the object-based audio signal; and determining an encoding mode of the scene-based audio signal based on the signal characteristic of the scene-based audio signal.

[0045] It should be noted that, in an embodiment of the disclosure, for the audio signals in different formats, methods for determining corresponding encoding modes based on the signal characteristics are different. The method for determining the encoding mode of the audio signal in each format based on the signal characteristic of the audio signal in each format will be described in detail in the following embodiments.

[0046] At step 103, the audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into an encoded stream and the encoded stream is sent to a decoding end.

[0047] In an embodiment of the disclosure, encoding the audio signal in each format using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format may include:

[0048] encoding the channel-based audio signal using the encoding mode of the channel-based audio signal;

[0049] encoding the object-based audio signal using the encoding mode of the object-based audio signal;

[0050] encoding the scene-based audio signal using the encoding mode of the scene-based audio signal.

[0051] Further, in an embodiment of the disclosure, when the encoded signal parameter

information of the audio signals in various formats is written into the encoded stream, determined side information parameters corresponding to the audio signals in various formats may be written into the encoded stream. The side information parameter is configured to indicate an encoding mode corresponding to the audio signal in a corresponding format.

[0052] In an embodiment of the disclosure, by writing the side information parameters corresponding to the audio signals in various formats into the encoded stream and sending the encoded stream to the decoding end, the decoding end may determine the encoding code corresponding to the audio signal in each format based on the side information parameters corresponding to the audio signals in various formats, and may decode, based on the encoding mode, the audio signal in each format using a corresponding decoding mode subsequently.

[0053] It should be noted that, in an embodiment of the disclosure, for the object-based audio signal, the corresponding encoded signal parameter information may retain partial object signals. For the scene-based audio signal and the channel-based audio signal, the corresponding encoded signal parameter information does not need to retain the signal in the original format, but is converted to the signal in another format.

[0054] In conclusion, in the signal encoding and decoding method provided in the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[0055] FIG. 2a is a flowchart of another signal encoding and decoding method according to an embodiment of the disclosure. The method is performed by an encoding end. As illustrated in FIG. 2a, the signal encoding and decoding method may include the following steps.

[0056] At step 201, an audio signal in a mixed format is obtained. The audio signal in the mixed format includes at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[0057] At step 202, in response to the audio signal in the mixed format including a channel-based audio signal, an encoding code of the channel-based audio signal is determined based on a signal characteristic of the channel-based audio signal.

[0058] In an embodiment of the disclosure, the method for determining the encoding mode of the channel-based audio signal based on the signal characteristic of the channel-based audio signal may include:

[0059] obtaining a number of object signals included in the channel-based audio signal and determining whether the number of the object signals included in the channel-based audio signal is less than a first threshold (for example, which may be 5).

[0060] In an embodiment of the disclosure, when the number of the object signals included in the channel-based audio signal is less than the first threshold, the method for determining the encoding mode of the channel-based audio signal may be at least one of the following solutions.

[0061] Solution 1, each object signal in the channel-based audio signal is encoded using the object signal encoding kernel.

[0062] Solution 2, input first command line control information is obtained, and the object signal encoding kernel is used to encode at least part of object signals in the channel-based audio signal based on the first command line control information. The first command line control information is configured to indicate object signals that need to be encoded among the object signals included in the channel-based audio signal. The number of the object signals that need to be encoded is greater than or equal to 1, and less than or equal to the total number of the object signals included in the channel-based audio signal.

[0063] It can be seen that, in an embodiment of the disclosure, when it is determined that the number of the object signals included in the channel-based audio signal is less than the first threshold, all or a part of the object signals in the channel-based audio signal may

be encoded, so that the encoding difficulty can be greatly reduced and the encoding efficiency can be improved.

[0064] In another embodiment of the disclosure, when the number of the object signals included in the channel-based audio signal is not less than the first threshold, the method for determining the encoding mode of the channel-based audio signal may be at least one of the following solutions.

[0065] Solution 3, the channel-based audio signal is converted into a first audio signal in another format (for example, it may be a scene-based audio signal or an object-based audio signal). A number of channels of the first audio signal in another format is less than or equal to a number of channels of the channel-based audio signal. The encoding kernel corresponding to the first audio signal in another format is used to encode the first audio signal in another format. For example, in an embodiment of the disclosure, when the channel-based audio signal is a channel-based audio signal in the 7.1.4 format (the total number of channels is 13), the first audio signal in another format may be, for example, an FOA (First Order Ambisonics, also called first-order high-fidelity stereo) signal (the total number of channels is 4), then the total number of channels of the signal to be encoded can be changed from 13 to 4 by converting the channel-based audio signal in the 7.1.4 format into the FOA signal, thereby greatly reducing the encoding difficulty and improving the encoding efficiency.

[0066] Solution 4, input first command line control information is obtained, and the object signal encoding kernel is used to encode at least part of the object signals in the channel-based audio signal based on the first command line control information. The first command line control information is configured to indicate object signals that need to be encoded among the object signals included in the channel-based audio signal, the number of the object signals that need to be encoded is greater than or equal to 1, and less than or equal to the total number of the object signals included in the channel-based audio signal.

[0067] Solution 5, input second command line control information is obtained, and the object signal encoding kernel is used to encode at least part of channel signals in the channel-based audio signal based on the second command line control information. The second command line control information is configured to indicate channel signals that need to be encoded among the channel signals included in the channel-based audio signal, and the number of the channel signals that need to be encoded is greater than or equal to 1, and less

than or equal to the total number of the channel signals included in the channel-based audio signal.

[0068] It can be seen that, in an embodiment of the disclosure, when it is determined that the number of the object signals included in the channel-based audio signal is large, if the channel-based audio signal is directly encoded, then the encoding complexity is high. In this case, only part of the object signals in the channel-based audio signal may be encoded, and/or only part of the channel signals in the channel-based audio signal may be encoded, and/or the channel-based audio signal may be converted into a signal with fewer channels for encoding, which can greatly reduce the encoding complexity and optimize the encoding efficiency.

[0069] At step 203, in response to the audio signal in the mixed format including an object-based audio signal, an encoding code of the object-based audio signal is determined based on a signal characteristic of the object-based audio signal.

[0070] Detailed description of step 203 may be introduced in the following embodiments.

[0071] At step 204, in response to the audio signal in the mixed format including a scene-based audio signal, an encoding code of the scene-based audio signal is determined based on a signal characteristic of the scene-based audio signal.

[0072] In an embodiment of the disclosure, determining the encoding mode of the scene-based audio signal based on the signal characteristic of the scene-based audio signal includes:

[0073] obtaining a number of object signals included in the scene-based audio signal and determining whether the number of the object signals included in the scene-based audio signal is less than a second threshold (for example, which may be 5).

[0074] In an embodiment of the disclosure, when the number of the object signals included in the scene-based audio signal is less than the second threshold, the method for determining the encoding mode of the scene-based audio signal may be at least one of the following solutions.

[0075] Solution a, each object signal in the scene-based audio signal is encoded using the object signal encoding kernel.

[0076] Solution b, input fourth command line control information is obtained, and the object signal encoding kernel is used to encode at least part of object signals in the scene-

based audio signal based on the fourth command line control information. The fourth command line control information is configured to indicate object signals that need to be encoded among the object signals included in the scene-based audio signal. The number of the object signals that need to be encoded is greater than or equal to 1, and less than or equal to the total number of the object signals included in the scene-based audio signal.

[0077] It can be seen that, in an embodiment of the disclosure, when it is determined that the number of the object signals included in the scene-based audio signal is less than the second threshold, all or a part of the object signals in the scene-based audio signal may be encoded, so that the encoding difficulty can be greatly reduced and the encoding efficiency can be improved.

[0078] In another embodiment of the disclosure, when the number of the object signals included in the scene-based audio signal is not less than the second threshold, the method for determining the encoding mode of the scene-based audio signal may be at least one of the following solutions.

[0079] Solution c, the scene-based audio signal is converted into a second audio signal in another format. A number of channels of the second audio signal in another format is less than or equal to a number of channels of the scene-based audio signal. The scene signal encoding kernel is used to encode the second audio signal in another format.

[0080] Solution d, a low-order conversion is performed on the scene-based audio signal, so as to convert the scene-based audio signal into a scene-based audio signal with a lower order than a current order of the scene-based audio signal, and the scene signal encoding kernel is used to encode the scene-based audio signal with the lower order. It should be noted that, in an embodiment of the disclosure, when the low-order conversion is performed on the scene-based audio signal, the scene-based audio signal may also converted into a signal in another format through the low-order conversion. As an example, the 3rd-order scene-based audio signal can be converted into a low-order channel-based audio signal in a 5.0 format. In this case, the total number of channels of the signal to be encoded is changed from 16((3+1)*(3+ 1)) to 5, which greatly reduces the encoding complexity and improves the encoding efficiency.

[0081] It can be seen that, in an embodiment of the disclosure, when it is determined that the number of the object signals included in the scene-based audio signal is large, if the scene-based audio signal is directly encoded, the encoding complexity is high. In this case,

the scene-based audio signal can be converted into a signal with a small number of channels before performing the encoding, and/or the scene-based audio signal can be converted into a low-order signal before performing the encoding, thereby greatly reducing the encoding complexity and optimizing the encoding efficiency.

[0082] At step 205, the audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into an encoded stream and the encoded stream is sent to a decoding end.

[0083] For the related description of step 205, reference may be made to the foregoing embodiments, which is not elaborated in the embodiment of the disclosure.

[0084] Finally, based on the above contents, FIG. 2b is a flowchart of a signal encoding method provided by an embodiment of the present disclosure. In combination with the above contents and FIG. 2b, it can be seen that when the encoding end receives an audio signal in a mixed format, the audio signals in various formats can be obtained by the signal characteristic analysis, and then, based on the command line control information (that is, the above-mentioned first command line control information, and/or the second command line control information (which will be introduced later), and/or the fourth command line control information), the corresponding encoding kernels are adopted to encode the audio signals in various formats using the corresponding encoding modes, and the encoded signal parameter information of the audio signals in various formats is written into the encoded stream and the encoded stream is sent to the decoding end.

[0085] In conclusion, in the signal encoding and decoding method provided by the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the

audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[0086] FIG. 3 is a flowchart of a signal encoding and decoding method provided by an embodiment of the present disclosure. The method is performed by an encoding end. As illustrated in FIG. 3, the signal encoding and decoding method may include the following steps.

[0087] At step 301, an audio signal in a mixed format is obtained. The audio signal in the mixed format includes at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[0088] At step 302, in response to the audio signal in the mixed format including an object-based audio signal, a signal characteristic analysis is performed on the object-based audio signal to obtain an analysis result.

[0089] In an embodiment of the disclosure, the signal characteristic analysis may be an analysis of cross-correlation parameter values of signals. In another embodiment of the disclosure, the characteristic analysis may be an analysis of a frequency-band bandwidth range of the signals. The analysis of the cross-correlation parameter values and the analysis of the frequency-band bandwidth range will be described in detail in following embodiments.

[0090] At step 303, a classification is performed on the object-based audio signal to obtain a first type of object signal set and a second type of object signal set. Each of the first type of object signal set and the second type of object signal set includes at least one object-based audio signal.

[0091] Since the object-based audio signal may include different types of object signals, and the subsequent encoding modes for different types of object signals will be different, in an embodiment of the disclosure, the different types of object signals in the object-based audio signal can be classified to obtain the first type of object signal set and the second type of object signal set, and then the corresponding encoding modes can be determined respectively for the first type of object signal set and the second type of object signal set. The classification manner for the first type of object signal set and the second type of object signal set will be described in detail in following embodiments.

[0092] At step 304, an encoding mode corresponding to the first type of object signal set is determined.

[0093] In an embodiment of the disclosure, when a different classification manner for the first type of object signal set is used in the above step 303, a different encoding mode of the first type of object signal set may be determined in this step. The specific method of "determining the encoding mode corresponding to the first type of object signal set" will be described in following embodiments.

[0094] At step 305, a classification is performed on the second type of object signal set based on the analysis result to obtain at least one object signal subset, and the encoding mode corresponding to each object signal subset is determined based on the classification result. The object signal subset includes at least one object-based audio signal.

[0095] If a different signal characteristic analysis is used in step 302, a different classification manner for the object based audio signal and a different method for determining the encoding mode corresponding to each object signal subset can be used in this step.

[0096] Specifically, in an embodiment of the disclosure, if the signal characteristic analysis used in step 302 is the analysis of the cross-correlation parameter values of the signals, then the classification manner for the second type of object signal set in this step can be a classification manner based on the cross-correlation parameter values of the signals, and the method for determining the encoding mode corresponding to each object signal subset may be determining the encoding mode corresponding to each object signal subset based on the cross-correlation parameter values of the signals.

[0097] In another embodiment of the disclosure, if the signal characteristic analysis used in step 302 is the analysis of the frequency-band bandwidth range of the signals, the classification manner for the second type of object signal set in this step may be a classification manner based on the frequency-band bandwidth range of the signals, and the method for determining the encoding mode corresponding to each object signal subset may be determining the encoding mode corresponding to each object signal subset based on the frequency-band bandwidth range of the signals.

[0098] The above-mentioned “the classification manner based on the cross-correlation parameter values of the signals or the frequency-band bandwidth range of the signals” and “determining the encoding mode corresponding to each object signal subset based on the

cross-correlation parameter values of the signals or the frequency-band bandwidth range of the signals" will be described in detail in following embodiments.

[0099] At step 306, the audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into an encoded stream and the encoded stream is sent to a decoding end.

[00100] It should be noted that, in an embodiment of the disclosure, when a different classification manner for the second type of object signal set is used in step 307, the encoding situation of the above-mentioned second type of object signal subset may be different.

[00101] Accordingly, in an embodiment of the disclosure, the above-mentioned method for writing the encoded signal parameter information of the audio signal in each format into the encoded stream and sending the encoded stream to the decoding end may include the following steps.

[00102] Step 1, a classification side information parameter is determined. The classification side information parameter is configured to indicate the classification manner for the second type of object signal set.

[00103] Step 2, a side information parameter corresponding to the audio signal in each format is determined. The side information parameter is configured to indicate the encoding mode corresponding to the audio signal of the corresponding format.

[00104] Step 3, code stream multiplexing is performed on the classification side information parameter, the side information parameter corresponding to the audio signal in each format, and the encoded signal parameter information of the audio signal in each format, to obtain the encoded stream, and the encoded stream is sent to the decoding end.

[00105] In an embodiment of the disclosure, by sending the classification side information parameter and the side information parameters corresponding to the audio signals in various formats to the decoding end, the decoding end can determine the encoding situation corresponding to the object signal subset in the second type of object signal set based on the classification side information parameter, and the encoding mode corresponding to each object signal subset is determined based on the side information parameter corresponding to each object signal subset, so that based on the encoding situation and the encoding mode, the object-based audio signal can be subsequently decoded using the

corresponding decoding mode and decoding mode, and the decoding end can also determine the encoding modes corresponding to the channel-based audio signal and the scene-based audio signal based on the side information parameter corresponding to the audio signal in each format, and then realize decoding of the channel-based audio signal and the scene-based audio signal.

[00106] In conclusion, in the signal encoding and decoding method provided by the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00107] FIG. 4a is a flowchart of a signal encoding and decoding method provided by another embodiment of the present disclosure. The method is performed by an encoding end. As illustrated in FIG. 4a, the signal encoding and decoding method may include the following steps.

[00108] At step 401, an audio signal in a mixed format is obtained. The audio signal in the mixed format includes at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[00109] At step 402, in response to the audio signal in the mixed format including an object-based audio signal, a signal characteristic analysis is performed on the object-based audio signal to obtain an analysis result.

[00110] For the description of steps 401-402, reference may be made to the foregoing embodiments, which is not elaborated in the embodiment of the present disclosure.

[00111] At step 403, one or more signals that need not to be individually operated and processed in the object-based audio signal are classified into a first type of object signal set, and remaining signals are classified into a second type of object signal set. Each of the first type of object signal set and the second type of object signal set includes at least one object-based audio signal.

[00112] At step 404, it is determined that an encoding mode corresponding to the first type of object signal set includes: performing first pre-rendering processing on an object-based audio signal in the first type of object signal set, and encoding the signal after the first pre-rendering processing using a multi-channel encoding kernel.

[00113] In an embodiment of the disclosure, the first pre-rendering processing may include: performing signal format conversion processing on an object-based audio signal to convert the object-based audio signal into a channel-based audio signal.

[00114] At step 405, a classification is performed on the second type of object signal set based on an analysis result to obtain at least one object signal subset, and an encoding mode corresponding to each object signal subset is determined based on the classification result. The object signal subset includes at least one object-based audio signal.

[00115] At step 406, the audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into an encoded stream and the encoded stream is sent to a decoding end.

[00116] For the description of steps 405-406, reference may be made to the foregoing embodiments, which is not elaborated in the embodiment of the present disclosure.

[00117] Finally, based on the above contents, FIG. 4b is a flowchart of a signal encoding method for an object-based audio signal provided by an embodiment of the present disclosure. In combination with the above contents and FIG. 4b, a characteristic analysis can be performed on the object-based audio signal and then the classification is performed on the object-based audio signal to obtain the first type of object signal set and the second type of object signal set. The first pre-rendering processing is performed on the first type of object signal set, and the multi-channel encoding kernel is used for encoding. The classification is performed on the second type of object signal set based on the analysis result to obtain at least one object signal subset (such as object signal subset 1, object signal subset 2 ... object

signal subset n), and then the at least one object signal subset is encoded respectively.

[00118] In conclusion, in the signal encoding and decoding method provided by the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00119] FIG. 5a is a flowchart of a signal encoding and decoding method provided by an embodiment of the present disclosure. The method is performed by an encoding end. As illustrated in FIG. 5a, the signal encoding and decoding method may include the following steps.

[00120] At step 501, an audio signal in a mixed format is obtained. The audio signal in the mixed format includes at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[00121] At step 502, in response to the audio signal in the mixed format including an object-based audio signal, a signal characteristic analysis is performed on the object-based audio signal to obtain an analysis result.

[00122] For the description of steps 501-502, reference may be made to the foregoing embodiments, which is not elaborated in the embodiment of the present disclosure.

[00123] At step 503, one or more signals belonging to a background sound in the object-based audio signal are classified into a first type of object signal set, and remaining signals are classified into a second type of object signal set. Each of the first type of object signal set and the second type of object signal set includes at least one object-based audio

signal.

[00124] At step 504, it is determined that an encoding mode corresponding to the first type of object signal set includes: performing second pre-rendering processing on an object-based audio signal in the first type of object signal set, and encoding the signal after the second pre-rendering processing using a high order ambisonics (HOA) encoding kernel.

[00125] In an embodiment of the disclosure, the second pre-rendering processing may include: performing signal format conversion processing on an object-based audio signal to convert the object-based audio signal into a scene-based audio signal.

[00126] At step 505, a classification is performed on the second type of object signal set based on an analysis result to obtain at least one object signal subset, and an encoding mode corresponding to each object signal subset is determined based on the classification result. The object signal subset includes at least one object-based audio signal.

[00127] At step 506, the audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into an encoded stream and the encoded stream is sent to a decoding end.

[00128] For the description of steps 505-506, reference may be made to the foregoing embodiments, which is not elaborated in the embodiment of the present disclosure.

[00129] Finally, based on the above contents, FIG. 5b is a flowchart of another signal encoding method for an object-based audio signal provided by an embodiment of the present disclosure. In combination with the above contents and FIG. 5b, a characteristic analysis can be performed on the object-based audio signal and then the classification is performed on the object-based audio signal to obtain the first type of object signal set and the second type of object signal set. The second pre-rendering processing is performed on the first type of object signal set, and the HOA encoding kernel is used for encoding. The classification is performed on the second type of object signal set based on the analysis result to obtain at least one object signal subset (such as object signal subset 1, object signal subset 2 ... object signal subset n), and then the at least one object signal subset is encoded respectively.

[00130] In conclusion, in the signal encoding and decoding method provided by the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio

signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00131] FIG. 6a is a flowchart of a signal encoding and decoding method provided by an embodiment of the present disclosure. The method is performed by an encoding end. The embodiment of FIG. 6a is different from the embodiments of FIG. 4a and FIG. 5a in that, in the embodiment, the first type of object signal set is further divided into a first object signal subset and a second object signal subset. As illustrated in FIG. 6a, the signal encoding and decoding method may include the following steps.

[00132] At step 601, an audio signal in a mixed format is obtained. The audio signal in the mixed format includes at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[00133] At step 602, a signal characteristic analysis is performed on an object-based audio signal to obtain an analysis result.

[00134] At step 603, one or more signals that need not to be individually operated and processed in the object-based audio signal are classified into the first object signal subset, one or more signals belonging to a background sound in the object-based audio signal are classified into the second object signal subset, and remaining signals are classified into a second type of object signal set. Each of the first object signal subset, the second object signal subset and the second type of object signal set includes at least one object-based audio signal.

[00135] At step 604, encoding codes of the first object signal subset and the second object signal subset in the first type of object signal set are determined.

[00136] In an embodiment of the disclosure, it is determined that the encoding mode corresponding to the first object signal subset in the first type of object signal set includes: performing first pre-rendering processing on an object-based audio signal in the first object signal subset, and encoding the signal after the first pre-rendering processing using a multi-channel encoding kernel. The first pre-rendering processing includes: performing signal format conversion processing on the object-based audio signal to convert it into a channel-based audio signal.

[00137] In an embodiment of the disclosure, it is determined that the encoding mode corresponding to the second object signal subset in the first type of object signal set includes: performing second pre-rendering processing on an object-based audio signal in the second object signal subset, and encoding the signal after the second pre-rendering processing using the HOA encoding kernel. The second pre-rendering processing includes: performing signal format conversion processing on the object-based audio signal to convert it into a scene-based audio signal.

[00138] At step 605, a classification is performed on the second type of object signal set based on an analysis result to obtain at least one object signal subset, and an encoding mode corresponding to each object signal subset is determined based on the classification result. The object signal subset includes at least one object-based audio signal.

[00139] At step 606, the audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into an encoded stream and the encoded stream is sent to a decoding end.

[00140] For the description of steps 601-606, reference may be made to the foregoing embodiments, which is not elaborated in the embodiment of the present disclosure.

[00141] Finally, based on the above contents, FIG. 6b is a flowchart of another signal encoding method for an object-based audio signal provided by an embodiment of the present disclosure. In combination with the above contents and FIG. 6b, a characteristic analysis can be performed on the object-based audio signal and then the classification is performed on the object-based audio signal to obtain a first type of object signal set and a second type of object signal set. The first type of object signal set includes a first object signal subset and a second object signal subset. The first pre-rendering processing is performed on the first type

of object signal set, and the multi-channel encoding kernel is used for encoding. The second pre-rendering processing is performed on the second type of object signal set, and the HOA encoding kernel is used for encoding. The classification is performed on the second type of object signal set based on the analysis result to obtain at least one object signal subset (such as object signal subset 1, object signal subset 2 ... object signal subset n), and then the at least one object signal subset is encoded respectively.

[00142] In conclusion, in the signal encoding and decoding method provided by the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00143] FIG. 7a is a flowchart of a signal encoding and decoding method provided by an embodiment of the present disclosure. The method is performed by an encoding end. As illustrated in FIG. 7a, the signal encoding and decoding method may include the following steps.

[00144] At step 701, an audio signal in a mixed format is obtained. The audio signal in the mixed format includes at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[00145] At step 702, in response to the audio signal in the mixed format including an object-based audio signal, high-pass filtering processing is performed on the object-based audio signal.

[00146] In an embodiment of the disclosure, a filter may be used to perform the high-

pass filtering processing on the object signal.

[00147] A cut-off frequency of the filter is set to 20Hz (Hertz). A filtering formula adopted by the filter can be expressed as the following formula (1):

$$H_{20}(Z)=\frac{b_0+b_1z^{-1}+b_2z^{-2}}{1+a_1z^{-1}+a_2z^{-2}} \quad (1)$$

[00148] where, a_1, a_2, b_0, b_1, and b_2 are all constants, for example, b_0=0.9981492, b_1=-1.9963008, b_2=0.9981498, a_1=1.9962990, a_2=-0.9963056.

[00149] At step 703, a correlation analysis is performed on the signal after the high-pass filtering processing to determine cross-correlation parameter values between object-based audio signals.

[00150] In an embodiment of the disclosure, the above-mentioned correlation analysis may be calculated using the following formula (2):

$$\eta_{xy}=\frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}} \quad (2)$$

[00151] η_{xy} is configured to indicate the cross-correlation parameter value of object-based audio signal X and object-based audio signal Y. Both X_i and Y_i are configured to indicate the i-th object-based audio signal. $\bar{X}$ is configured to indicate an average value of a signal sequence of the object-based audio signal X, $\bar{Y}$ is configured to indicate an average value of a signal sequence of the object-based audio signal Y.

[00152] It should be noted that the above-mentioned method of "calculating the cross-correlation parameter value using the formula (2)" is an optional implementation provided by an embodiment of the present disclosure, and it should be recognized that other methods for calculating cross-correlation between parameter values between object signals in the related art can also be applied in the disclosure.

[00153] At step 704, a classification is performed on the object-based audio signal to obtain a first type of object signal set and a second type of object signal set. Each of the first type of object signal set and the second type of object signal set includes at least one object-based audio signal.

[00154] At step 705, an encoding mode corresponding to the first type of object signal set is determined.

[00155] For the description of steps 704-705, reference may be made to the foregoing embodiments, which is not elaborated in the embodiment of the present disclosure.

[00156] At step 706, a classification is performed on the second type of object signal set based on an analysis result to obtain at least one object signal subset, and an encoding

mode corresponding to each object signal subset is determined based on the classification result. The object signal subset includes at least one object-based audio signal.

[00157] In an embodiment of the disclosure, performing the classification on the second type of object signal set to obtain at least one object signal subset and determining the encoding mode corresponding to each object signal subset based on the classification result includes:

[00158] setting a normalized correlation degree interval based on correlation degrees; and performing the classification on the second type of object signal set based on the cross-correlation parameter values of the signals and the normalized correlation degree interval to obtain the at least one object signal subset.

[00159] Then the corresponding encoding mode can be determined based on a correlation degree corresponding to the object signal subset.

[00160] It can be understood that the number of the normalized correlation degree intervals is determined according to the division of the correlation degrees, which is not limited in the disclosure. Further, the lengths of different normalized correlation degree intervals are not limited in the disclosure. The corresponding number of normalized correlation degree intervals and different interval lengths can be set according to different divisions of the correlation degrees.

[00161] In an embodiment of the disclosure, the correlation degrees can be classified into four correlation degrees, including weak correlation, real correlation, significant correlation, and high correlation. Table 1 is a classification table for normalized correlation degree intervals provided by an embodiment of the disclosure.

normalized correlation degree interval	correlation degree
0.00 ~±0.30	weak correlation
±0.30-±0.50	real correlation
±0.50-±0.80	significant correlation
±0.80-±1.00	high correlation

Table 1

[00162] Based on the above contents, as an example, the object signals having the cross-correlation parameter values within the first interval are classified into an object signal set 1, and an independent encoding mode corresponding to the object signal set 1 is determined.

[00163] The object signals having the cross-correlation parameter values within the second interval are classified into an object signal set 2, and a joint encoding mode 1 corresponding to the object signal set 2 is determined.

[00164] The object signals having the cross-correlation parameter values within the third interval are classified into an object signal set 3, and a joint encoding mode 2 corresponding to the object signal set 3 is determined.

[00165] The object signals having the cross-correlation parameter values within the fourth interval are classified into an object signal set 4, and a joint encoding mode 3 corresponding to the object signal set 4 is determined.

[00166] In an embodiment of the disclosure, the first interval may be [0.00 ~±0.30), the second interval may be [±0.30-±0.50), the third interval may be [±0.50-±0.80), and the fourth interval may be [±0.80-±1.00]. When the cross-correlation parameter value between the object signals is within the first interval, it means that the object signals are weakly correlated. In this case, in order to ensure the encoding accuracy, an independent encoding mode is used for encoding. When the cross-correlation parameter value between the object signals is within the second interval, the third interval, or the fourth interval, it means that the cross-correlation between the object signals is high, and in this case, a joint encoding mode can be used for encoding to ensure the compression rate to save bandwidth.

[00167] In an embodiment of the disclosure, the encoding mode corresponding to the object signal subset includes the independent encoding mode or the joint encoding mode.

[00168] In an embodiment of the disclosure, the independent encoding mode corresponds to a time-domain processing manner or a frequency-domain processing manner.

[00169] When an object signal in the object signal subset is a speech signal or a speech-like signal, the independent encoding mode adopts the time-domain processing manner.

[00170] When an object signal in the object signal subset is an audio signal in another format other than the speech signal or the speech-like signal, the independent encoding mode adopts the frequency-domain processing manner.

[00171] In an embodiment of the disclosure, the above-mentioned time-domain processing manner may be implemented by using an ACELP encoding model. FIG. 7b is a schematic diagram of an ACELP encoding principle provided by an embodiment of the present disclosure. For details about the ACELP encoder principle, reference can be made to

the introduction in the prior art, which is not elaborated in the embodiment of the disclosure.

[00172] In an embodiment of the disclosure, the above-mentioned frequency-domain processing manner may include a transform domain processing manner. FIG. 7c is a schematic diagram of a frequency-domain encoding principle provided by an embodiment of the present disclosure. With reference to FIG. 7c, the input object signal can be converted to the frequency domain by performing MDCT transformation through a transformation module. A transformation formula and an inverse transformation formula of the MDCT transformation are expressed by the following formula (3) and formula (4) respectively.

$$\sum_{k=0} \qquad 2n \qquad 2 \qquad 2 \tag{3}$$

$$n \sum_{m=0} \qquad 2n \qquad 2 \tag{4}$$

[00173] A psychoacoustic model is used to adjust each frequency band for the object signal which is transformed into the frequency domain, and a quantization module is used to quantize an envelope coefficient of each frequency band through bit allocation to obtain quantized parameters. Finally, an entropy encoding module is used to perform entropy encoding on the quantized parameters to output the encoded object signal.

[00174] At step 707, the audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into an encoded stream and the encoded stream is sent to a decoding end.

[00175] In an embodiment of the disclosure, encoding the audio signal in each format using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format may include:

[00176] encoding the channel-based audio signal using the encoding mode of the channel-based audio signal;

[00177] encoding the object-based audio signal using the encoding mode of the object-based audio signal;

[00178] encoding the scene-based audio signal using the encoding mode of the scene-based audio signal.

[00179] In an embodiment of the disclosure, the above method of encoding the object-based audio signal using the encoding mode of the object-based audio signal may include:

[00180] encoding one or more signals in the first type of object signal set using the encoding mode corresponding to the first type of object signal set;

[00181] performing preprocessing on one or more object signal subsets in the second type of object signal set, and encoding all the object signal subsets after the preprocessing in the second type of object signal set using respective encoding modes and using a same object signal encoding kernel.

[00182] Based on the above contents, FIG. 7d is a flowchart of an encoding method for the second type of object signal set provided in an embodiment of the present disclosure.

[00183] In conclusion, in the signal encoding and decoding method provided in the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00184] FIG. 8a is a flowchart of a signal encoding and decoding method according to an embodiment of the disclosure. The method is performed by an encoding end. As illustrated in FIG. 8a, the signal encoding and decoding method may include the following steps.

[00185] At step 801, an audio signal in a mixed format is obtained. The audio signal in the mixed format includes at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[00186] At step 802, in response to the audio signal in the mixed format including an object-based audio signal, a frequency-band bandwidth range of an object signal is

analyzed.

[00187] At step 803, a classification is performed on the object-based audio signal to obtain a first type of object signal set and a second type of object signal set. Each of the first type of object signal set and the second type of object signal set includes at least one object-based audio signal.

[00188] At step 804, an encoding mode corresponding to the first type of object signal set is determined.

[00189] At step 805, a classification is performed on the second type of object signal set based on an analysis result to obtain at least one object signal subset, and an encoding mode corresponding to each object signal subset is determined based on the classification result. The object signal subset includes at least one object-based audio signal.

[00190] In an embodiment of the disclosure, performing the classification on the second type of object signal set based on the analysis result to obtain at least one object signal subset, and determining the encoding mode corresponding to each object signal subset based on the classification result may include:

[00191] determining bandwidth intervals corresponding to different frequency-band bandwidths;

[00192] performing the classification on the second type of object signal set to obtain the at least one object signal subset based on the frequency-band bandwidth range of the object signal and the bandwidth intervals corresponding to different frequency-band bandwidths, and determining a corresponding encoding mode based on a frequency-band bandwidth corresponding to the at least one object signal subset.

[00193] The frequency-band bandwidth of the signal usually includes narrowband, wideband, ultra-wideband and full-band. The bandwidth interval corresponding to the narrowband may be the first interval, the bandwidth interval corresponding to the wideband may be the second interval, the bandwidth interval corresponding to the ultra-wideband may be the third interval, and the bandwidth interval corresponding to the full band may be the fourth interval. Then, the classification can be performed on the second type of object signal set to obtain at least one object signal subset by determining the bandwidth interval to which the frequency-band bandwidth range of the object signal belongs. Afterwards, the corresponding encoding mode is determined according to the frequency-band bandwidth corresponding to at least one object signal subset. The narrowband, wideband, ultra-

wideband and full-band correspond to the narrowband encoding mode, wideband encoding mode, ultra-wideband encoding mode and full-band encoding mode, respectively.

[00194] It should be noted that, the lengths of different bandwidth intervals are not limited in the embodiment of the disclosure, and bandwidth intervals between different frequency-band bandwidths may overlap.

[00195] As an example, the object signals having the frequency-band bandwidths within the first interval are classified into an object signal set 1, and the narrowband encoding mode corresponding to the object signal set 1 is determined.

[00196] The object signals having the frequency-band bandwidths within the second interval are classified into an object signal set 2, and the wideband encoding mode corresponding to the object signal set 2 is determined.

[00197] The object signals having the frequency-band bandwidths within the third interval are classified into an object signal set 3, and the ultra-wideband encoding mode corresponding to the object signal set 3 is determined.

[00198] The object signals having the frequency-band bandwidths within the fourth interval are classified into an object signal set 4, and the full band encoding mode corresponding to the object signal set 4 is determined.

[00199] In an embodiment of the disclosure, the first interval may be 0~4kHz, the second interval may be 0~8kHz, the third interval may be 0~16kHz, and the fourth interval may be 0~20kHz. When the frequency-band bandwidth of the object signal is within the first interval, it means that the object signal is a narrowband signal, and then it may be determined that the encoding mode corresponding to the object signal may include performing encoding using fewer bits (i.e., using the narrowband encoding mode). When frequency-band bandwidth of the object signal is within the second interval, it means that the object signal is a wideband signal, and then it may be determined that the encoding mode corresponding to the object signal may include performing encoding using more bits (i.e., using the wideband encoding mode). When the frequency-band bandwidth of the object signal is within the third interval, it means that the object signal is an ultra-wideband signal, and then it may be determined that the encoding mode corresponding to the object signal may include performing encoding using relative more bits (i.e., using the ultra-wideband encoding mode). When frequency-band bandwidth of the object signal is within the fourth interval, it means that the object signal is a full band signal, and then it may be determined that the encoding

mode corresponding to the object signal may include performing encoding using much more bits (i.e., using the full band encoding mode).

[00200] By using different bits to encode the signal for different frequency-band bandwidth signals, the compression rate of the signal can be ensured and the bandwidth can be saved.

[00201] At step 806, the audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into an encoded stream and the encoded stream is sent to a decoding end.

[00202] In an embodiment of the disclosure, encoding the audio signal in each format using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format may include:

[00203] encoding the channel-based audio signal using the encoding mode of the channel-based audio signal;

[00204] encoding the object-based audio signal using the encoding mode of the object-based audio signal;

[00205] encoding the scene-based audio signal using the encoding mode of the scene-based audio signal.

[00206] In an embodiment of the disclosure, the above method of encoding the object-based audio signal using the encoding mode of the object-based audio signal may include:

[00207] encoding one or more signals in the first type of object signal set using the encoding mode corresponding to the first type of object signal set;

[00208] performing preprocessing on one or more object signal subsets in the second type of object signal set, and encoding different object signal subsets after the preprocessing using respective encoding modes and using different object signal encoding kernels.

[00209] Based on the above contents, FIG. 8b is a flowchart of an encoding method for the second type of object signal set provided in an embodiment of the present disclosure.

[00210] In conclusion, in the signal encoding and decoding method provided in the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio

signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00211] FIG. 9a is a flowchart of a signal encoding and decoding method according to an embodiment of the disclosure. The method is performed by an encoding end. As illustrated in FIG. 9a, the signal encoding and decoding method may include the following steps.

[00212] At step 901, an audio signal in a mixed format is obtained. The audio signal in the mixed format includes at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[00213] At step 902, in response to the audio signal in the mixed format including an object-based audio signal, a frequency-band bandwidth range of an object signal is analyzed.

[00214] At step 903, a classification is performed on the object-based audio signal to obtain a first type of object signal set and a second type of object signal set. Each of the first type of object signal set and the second type of object signal set includes at least one object-based audio signal.

[00215] At step 904, an encoding mode corresponding to the first type of object signal set is determined.

[00216] At step 905, input third command line control information is obtained. The third command line control information is configured to indicate a frequency-band bandwidth range to be encoded corresponding to the object-based audio signal.

[00217] At step 906, a classification is performed on the second type of object signal

set by combining the third command line control information and an analysis result to obtain at least one object signal subset, and the encoding mode corresponding to each object signal subset is determined based on the classification result.

[00218] In an embodiment of the disclosure, performing the classification on the second type of object signal set by combining the third command line control information and the analysis result to obtain the at least one object signal subset, and determining the encoding mode corresponding to each object signal subset based on the classification result may include:

[00219] when a frequency-band bandwidth range indicated by the third command line control information is different from a frequency-band bandwidth range obtained from the analysis result, performing the classification on the second type of object signal set preferentially based on the frequency-band bandwidth range indicated by the third command line control information, and determining the encoding mode corresponding to each object signal set based on the classification result;

[00220] when a frequency-band bandwidth range indicated by the third command line control information is the same as a frequency-band bandwidth range obtained from the analysis result, performing the classification on the second type of object signal set based on the frequency-band bandwidth range indicated by the third command line control information and the frequency-band bandwidth range obtained from the analysis result, and determining the encoding mode corresponding to each object signal set based on the classification result.

[00221] In an embodiment of the disclosure, it is assumed that the analysis result of the object signal is an ultra-wideband signal, and the frequency-band bandwidth range indicated by the third command line control information of the object signal is a full band signal. In this case, the object signal can be classified into the object signal subset 4 based on the third command line control information, and it is determined that the encoding mode corresponding to the object signal subset 4 is the full band encoding mode.

[00222] At step 907, the audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into an encoded stream and the encoded stream is sent to a decoding end.

[00223] In an embodiment of the disclosure, encoding the audio signal in each

format using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format may include:

[00224] encoding the channel-based audio signal using the encoding mode of the channel-based audio signal;

[00225] encoding the object-based audio signal using the encoding mode of the object-based audio signal;

[00226] encoding the scene-based audio signal using the encoding mode of the scene-based audio signal.

[00227] In an embodiment of the disclosure, the above method of encoding the object-based audio signal using the encoding mode of the object-based audio signal may include:

[00228] encoding one or more signals in the first type of object signal set using the encoding mode corresponding to the first type of object signal set;

[00229] performing preprocessing on one or more object signal subsets in the second type of object signal set, and encoding different object signal subsets after the preprocessing using respective encoding modes and using different object signal encoding kernels.

[00230] Based on the above contents, FIG. 9b is a flowchart of another encoding method for the second type of object signal set provided in an embodiment of the present disclosure.

[00231] In conclusion, in the signal encoding and decoding method provided in the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive

encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00232] FIG. 10 is a flowchart of a signal encoding and decoding method according to an embodiment of the disclosure. The method is performed by a decoding end. As illustrated in FIG. 10, the signal encoding and decoding method may include the following steps.

[00233] At step 1001, an encoded stream sent by an encoding end is received.

[00234] In an embodiment of the disclosure, the decoding end may be a UE or a base station.

[00235] At step 1002, the encoded stream is decoded to obtain an audio signal in a mixed format, in which the audio signal in the mixed format includes at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[00236] In conclusion, in the signal encoding and decoding method provided in the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00237] FIG. 11a is a flowchart of a signal encoding and decoding method according to an embodiment of the disclosure. The method is performed by a decoding end. As illustrated in FIG. 11a, the signal encoding and decoding method may include the following steps.

[00238] At step 1101, an encoded stream sent by an encoding end is received.

[00239] At step 1102, a code stream analysis is performed on the encoded stream to obtain a classification side information parameter, a side information parameter corresponding to an audio signal in each format, and encoded signal parameter information of the audio signal in each format.

[00240] The classification side information parameter is configured to indicate a classification manner for a second type of object signal set of the object-based audio signal. The side information parameter is configured to indicate an encoding mode corresponding to the audio signal in each format.

[00241] At step 1103, encoded signal parameter information of the channel-based audio signal is decoded based on the side information parameter corresponding to the channel-based audio signal.

[00242] In an embodiment of the disclosure, decoding the encoded signal parameter information of the channel-based audio signal based on the side information parameter corresponding to the channel-based audio signal may include: determining an encoding mode corresponding to the channel-based audio signal based on the side information parameter corresponding to the channel-based audio signal; and decoding the encoded signal parameter information of the channel-based audio signal using a corresponding decoding mode based on the encoding mode corresponding to the channel-based audio signal.

[00243] At step 1104, encoded signal parameter information of the scene-based audio signal is decoded based on the side information parameter corresponding to the scene-based audio signal.

[00244] In an embodiment of the disclosure, decoding the encoded signal parameter information of the scene-based audio signal based on the side information parameter corresponding to the scene-based audio signal may include: determining an encoding mode corresponding to the scene-based audio signal based on the side information parameter corresponding to the scene-based audio signal; and decoding the encoded signal parameter information of the scene-based audio signal using a corresponding decoding mode based on the encoding mode corresponding to the scene-based audio signal.

[00245] At step 1105, encoded signal parameter information of the object-based audio signal is decoded based on the classification side information parameter and the side information parameter corresponding to the object-based audio signal.

[00246] The detailed implementations of step 1105 will be described in the following

embodiments.

[00247] Finally, based on the above contents, FIG. 11b is a flowchart of a signal decoding method provided in an embodiment of the present disclosure.

[00248] In conclusion, in the signal encoding and decoding method provided in the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00249] FIG. 12a is a flowchart of a signal encoding and decoding method according to an embodiment of the disclosure. The method is performed by a decoding end. As illustrated in FIG. 12a, the signal encoding and decoding method may include the following steps.

[00250] At step 1201, an encoded stream sent by an encoding end is received.

[00251] At step 1202, a code stream analysis is performed on the encoded stream to obtain a classification side information parameter, a side information parameter corresponding to an audio signal in each format, and encoded signal parameter information of the audio signal in each format.

[00252] At step 1203, encoded signal parameter information corresponding to a first type of object signal set and encoded signal parameter information corresponding to the second type of object signal set are determined from the encoded signal parameter information of the object-based audio signal.

[00253] In an embodiment of the disclosure, the encoded signal parameter

information corresponding to the first type of object signal set and the encoded signal parameter information corresponding to the second type of object signal set are determined from the encoded signal parameter information of the object-based audio signal based on the side information parameter corresponding to the object-based audio signal.

[00254] At step 1204, the encoded signal parameter information corresponding to the first type of object signal set is decoded based on a side information parameter corresponding to the first type of object signal set.

[00255] In detail, in an embodiment of the disclosure, decoding the encoded signal parameter information corresponding to the first type of object signal set based on the side information parameter corresponding to the first type of object signal set may include: determining an encoding mode corresponding to the first type of object signal set based on the side information parameter corresponding to the first type of object signal set; and decoding the encoded signal parameter information of the first type of object signal set using a corresponding decoding mode based on the encoding mode corresponding to the first type of object signal set.

[00256] At step 1205, the encoded signal parameter information corresponding to the second type of object signal set is decoded based on the classification side information parameter and the side information parameter corresponding to the second type of object signal set.

[00257] In an embodiment of the disclosure, decoding the encoded signal parameter information corresponding to the second type of object signal set based on the classification side information parameter and the side information parameter corresponding to the second type of object signal set may include the following steps.

[00258] Step a, determining the classification manner for the second type of object signal set based on the classification side information parameter;

[00259] With reference to the above embodiments, for different classification manners for the second type of object signal set, the encoding conditions are different. In detail, in an embodiment of the disclosure, when the classification manner for the second type of object signal set is a classification manner based on cross-correlation parameter values of signals, the encoding situation corresponding to the encoding end includes encoding all object signal sets using respective encoding modes and using the same encoding kernel.

[00260] In another embodiment of the disclosure, when the classification manner for the second type of object signal set is a classification manner based on a frequency-band bandwidth range, the encoding situation corresponding to the encoding end includes encoding different object signal sets using respective encoding modes and using different encoding kernels.

[00261] Therefore, in this step, it is necessary to determine, based on the classification side information parameter, the classification manner for the second type of object signal set in the encoding process, so as to determine the encoding situation in the encoding process. Subsequently, the decoding can be performed based on the encoding situation.

[00262] Step b, the encoded signal parameter information corresponding to each object signal subset in the second type of object signal set is decoded based on the classification manner for the second type of object signal set and the side information parameter corresponding to the second type of object signal set.

[00263] In an embodiment of the disclosure, decoding the encoded signal parameter information corresponding to each object signal subset in the second type of object signal set based on the classification manner for the second type of object signal set and the side information parameter corresponding to the second type of object signal set may include:

[00264] determining the encoding situation in the encoding process based on the classification manner, and determining the corresponding decoding situation based on the encoding situation, and then according to the corresponding decoding situation and based on the encoding mode corresponding to the encoded signal parameter information corresponding to each object signal subset, using a corresponding decoding mode to decode the encoded signal parameter information corresponding to each object signal subset.

[00265] In detail, in an embodiment of the disclosure, if it is determined based on the classification side information parameter that the encoding situation in the encoding process includes: encoding all object signal subsets using the corresponding encoding modes and using the same encoding kernel, then it is determined that the decoding condition of the decoding process includes: decoding the encoded signal parameter information corresponding to all object signal subsets using the same decoding kernel. In the decoding process, the encoded signal parameter information corresponding to the object signal subset is decoded using a corresponding decoding mode based on the encoding mode corresponding

to the encoded signal parameter information corresponding to each object signal subset.

[00266] In another embodiment of the disclosure, if it is determined based on the classification side information parameter that the encoding situation in the encoding process includes: encoding different object signal subsets using the corresponding encoding modes and using different encoding kernels, then it is determined that the decoding mode of the decoding process includes: decoding the encoded signal parameter information corresponding to each object signal subset using different decoding kernels. In the decoding process, specifically, the encoded signal parameter information corresponding to each object signal subset is decoded using a corresponding decoding mode based on the encoding mode corresponding to the encoded signal parameter information corresponding to each object signal subset.

[00267] Finally, based on the above contents, and FIGs. 12b, 12c and 12d are flowcharts of a decoding method for an object-based audio signal according to embodiments of the present disclosure. FIGs. 12e and 12f are flowcharts of a decoding method for a second type of object signal set provided by embodiments of the present disclosure.

[00268] In conclusion, in the signal encoding and decoding method provided in the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00269] FIG. 13 is a flowchart of a signal encoding and decoding method according to an embodiment of the disclosure. The method is performed by a decoding end. As

illustrated in FIG. 13, the signal encoding and decoding method may include the following steps.

[00270] At step 1301, an encoded stream sent by an encoding end is received.

[00271] At step 1302, the encoded stream is decoded to obtain an audio signal in a mixed format, in which the audio signal in the mixed format includes at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[00272] At step 1303, post-processing is performed on the decoded object-based audio signal.

[00273] In conclusion, in the signal encoding and decoding method provided in the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00274] FIG. 14 is a flowchart of a signal encoding and decoding method according to an embodiment of the disclosure. The method is performed by an encoding end. As illustrated in FIG. 14, the signal encoding and decoding method may include the following steps.

[00275] At step 1401, an audio signal in a mixed format is obtained. The audio signal in the mixed format includes at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[00276] At step 1402, in response to the audio signal in the mixed format including a channel-based audio signal, an encoding code of the channel-based audio signal is

determined based on a signal characteristic of the channel-based audio signal.

[00277] In an embodiment of the disclosure, the method for determining the encoding mode of the channel-based audio signal based on the signal characteristic of the channel-based audio signal may include:

[00278] obtaining a number of object signals included in the channel-based audio signal and determining whether the number of the object signals included in the channel-based audio signal is less than a first threshold (for example, which may be 5).

[00279] In an embodiment of the disclosure, when the number of the object signals included in the channel-based audio signal is less than the first threshold, the method for determining the encoding mode of the channel-based audio signal may be at least one of the following solutions.

[00280] Solution 1, each object signal in the channel-based audio signal is encoded using the object signal encoding kernel.

[00281] Solution 2, input first command line control information is obtained, and the object signal encoding kernel is used to encode at least part of object signals in the channel-based audio signal based on the first command line control information. The first command line control information is configured to indicate object signals that need to be encoded among the object signals included in the channel-based audio signal. The number of the object signals that need to be encoded is greater than or equal to 1, and less than or equal to the total number of the object signals included in the channel-based audio signal.

[00282] It can be seen that, in an embodiment of the disclosure, when it is determined that the number of the object signals included in the channel-based audio signal is less than the first threshold, all or a part of the object signals in the channel-based audio signal may be encoded, so that the encoding difficulty can be greatly reduced and the encoding efficiency can be improved.

[00283] In another embodiment of the disclosure, when the number of the object signals included in the channel-based audio signal is not less than the first threshold, the method for determining the encoding mode of the channel-based audio signal may be at least one of the following solutions.

[00284] Solution 3, the channel-based audio signal is converted into a first audio signal in another format (for example, it may be a scene-based audio signal or an object-based audio signal). A number of channels of the first audio signal in another format is less

than or equal to a number of channels of the channel-based audio signal. The encoding kernel corresponding to the first audio signal in another format is used to encode the first audio signal in another format. For example, in an embodiment of the disclosure, when the channel-based audio signal is a channel-based audio signal in the 7.1.4 format (the total number of channels is 13), the first audio signal in another format may be, for example, an FOA (First Order Ambisonics, also called first-order high-fidelity stereo) signal (the total number of channels is 4), then the total number of channels of the signal to be encoded can be changed from 13 to 4 by converting the channel-based audio signal in the 7.1.4 format into the FOA signal, thereby greatly reducing the encoding difficulty and improving the encoding efficiency.

[00285] Solution 4, input first command line control information is obtained, and the object signal encoding kernel is used to encode at least part of the object signals in the channel-based audio signal based on the first command line control information. The first command line control information is configured to indicate object signals that need to be encoded among the object signals included in the channel-based audio signal, the number of the object signals that need to be encoded is greater than or equal to 1, and less than or equal to the total number of the object signals included in the channel-based audio signal.

[00286] Solution 5, input second command line control information is obtained, and the object signal encoding kernel is used to encode at least part of channel signals in the channel-based audio signal based on the second command line control information. The second command line control information is configured to indicate channel signals that need to be encoded among the channel signals included in the channel-based audio signal, and the number of the channel signals that need to be encoded is greater than or equal to 1, and less than or equal to the total number of the channel signals included in the channel-based audio signal.

[00287] It can be seen that, in an embodiment of the disclosure, when it is determined that the number of the object signals included in the channel-based audio signal is large, if the channel-based audio signal is directly encoded, then the encoding complexity is high. In this case, only part of the object signals in the channel-based audio signal may be encoded, and/or only part of the channel signals in the channel-based audio signal may be encoded, and/or the channel-based audio signal may be converted into a signal with fewer channels for encoding, which can greatly reduce the encoding complexity and optimize the encoding

efficiency.

[00288] At step 1403, the channel-based audio signal is encoded using the encoding mode of the channel-based audio signal to obtain encoded signal parameter information of the channel-based audio signal, the encoded signal parameter information of the channel-based audio signal is written into an encoded stream and the encoded stream is sent to a decoding end.

[00289] For the related description of step 1403, reference may be made to the foregoing embodiments, which is not elaborated in the embodiment of the disclosure.

[00290] In conclusion, in the signal encoding and decoding method provided by the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00291] FIG. 15 is a flowchart of another signal encoding and decoding method according to an embodiment of the disclosure. The method is performed by an encoding end. As illustrated in FIG. 15, the signal encoding and decoding method may include the following steps.

[00292] At step 1501, an audio signal in a mixed format is obtained. The audio signal in the mixed format includes at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[00293] At step 1502, in response to the audio signal in the mixed format including a scene-based audio signal, an encoding code of the scene-based audio signal is determined

based on a signal characteristic of the scene-based audio signal.

[00294] In an embodiment of the disclosure, the method for determining the encoding mode of the scene-based audio signal based on the signal characteristic of the scene-based audio signal may include:

[00295] obtaining a number of object signals included in the scene-based audio signal and determining whether the number of the object signals included in the scene-based audio signal is less than a second threshold (for example, which may be 5).

[00296] In an embodiment of the disclosure, when the number of the object signals included in the scene-based audio signal is less than the second threshold, the method for determining the encoding mode of the scene-based audio signal may be at least one of the following solutions.

[00297] Solution a, each object signal in the scene-based audio signal is encoded using the object signal encoding kernel.

[00298] Solution b, input fourth command line control information is obtained, and the object signal encoding kernel is used to encode at least part of object signals in the scene-based audio signal based on the fourth command line control information. The fourth command line control information is configured to indicate object signals that need to be encoded among the object signals included in the scene-based audio signal. The number of the object signals that need to be encoded is greater than or equal to 1, and less than or equal to the total number of the object signals included in the scene-based audio signal.

[00299] It can be seen that, in an embodiment of the disclosure, when it is determined that the number of the object signals included in the scene-based audio signal is less than the second threshold, all or a part of the object signals in the scene-based audio signal may be encoded, so that the encoding difficulty can be greatly reduced and the encoding efficiency can be improved.

[00300] In another embodiment of the disclosure, when the number of the object signals included in the scene-based audio signal is not less than the second threshold, the method for determining the encoding mode of the scene-based audio signal may be at least one of the following solutions.

[00301] Solution c, the scene-based audio signal is converted into a second audio signal in another format. A number of channels of the second audio signal in another format is less than or equal to a number of channels of the scene-based audio signal. The scene

signal encoding kernel is used to encode the second audio signal in another format.

[00302] Solution d, a low-order conversion is performed on the scene-based audio signal, so as to convert the scene-based audio signal into a scene-based audio signal with a lower order than a current order of the scene-based audio signal, and the scene signal encoding kernel is used to encode the scene-based audio signal with the lower order. It should be noted that, in an embodiment of the disclosure, when the low-order conversion is performed on the scene-based audio signal, the scene-based audio signal may also converted into a signal in another format through the low-order conversion. As an example, the 3rd-order scene-based audio signal can be converted into a low-order channel-based audio signal in a 5.0 format. In this case, the total number of channels of the signal to be encoded is changed from 16((3+1)*(3+ 1)) to 5, which greatly reduces the encoding complexity and improves the encoding efficiency.

[00303] It can be seen that, in an embodiment of the disclosure, when it is determined that the number of the object signals included in the scene-based audio signal is large, if the scene-based audio signal is directly encoded, the encoding complexity is high. In this case, the scene-based audio signal can be converted into a signal with a small number of channels before performing the encoding, and/or the scene-based audio signal can be converted into a low-order signal before performing the encoding, thereby greatly reducing the encoding complexity and optimizing the encoding efficiency.

[00304] At step 1503, the scene-based audio signal is encoded using the encoding mode of the scene-based audio signal to obtain encoded signal parameter information of the scene-based audio signal, the encoded signal parameter information of the scene-based audio signal is written into an encoded stream and the encoded stream is sent to a decoding end.

[00305] For the related description of step 1503, reference may be made to the foregoing embodiments, which is not elaborated in the embodiment of the disclosure.

[00306] In conclusion, in the signal encoding and decoding method provided by the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information

of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00307] FIG. 16 is a flowchart of a signal encoding and decoding method according to an embodiment of the disclosure. The method is performed by a decoding end. As illustrated in FIG. 16, the signal encoding and decoding method may include the following steps.

[00308] At step 1601, an encoded stream sent by an encoding end is received.

[00309] At step 1602, a code stream analysis is performed on the encoded stream to obtain a classification side information parameter, a side information parameter corresponding to an audio signal in each format, and encoded signal parameter information of the audio signal in each format.

[00310] At step 1603, encoded signal parameter information of the channel-based audio signal is decoded based on the side information parameter corresponding to the channel-based audio signal.

[00311] In conclusion, in the signal encoding and decoding method provided by the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the

audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00312] FIG. 17 is a flowchart of a signal encoding and decoding method according to an embodiment of the disclosure. The method is performed by a decoding end. As illustrated in FIG. 17, the signal encoding and decoding method may include the following steps.

[00313] At step 1701, an encoded stream sent by an encoding end is received.

[00314] At step 1702, a code stream analysis is performed on the encoded stream to obtain a classification side information parameter, a side information parameter corresponding to an audio signal in each format, and encoded signal parameter information of the audio signal in each format.

[00315] At step 1703, encoded signal parameter information of the scene-based audio signal is decoded based on the side information parameter corresponding to the scene-based audio signal.

[00316] In conclusion, in the signal encoding and decoding method provided by the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00317] FIG. 18 is a block diagram of a signal encoding and decoding apparatus according to an embodiment of the disclosure. The apparatus is applied to an encoding end.

As illustrated in FIG. 18, the apparatus may include:

[00318] an obtaining module 1801, configured to obtain an audio signal in a mixed format, in which the audio signal in the mixed format comprises at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal;

[00319] a determining module 1802, configured to determine, based on signal characteristics of audio signals in different formats, an encoding mode of the audio signal in each format; and

[00320] an encoding module 1803, configured to encode the audio signal in each format using the encoding mode of the audio signal in each format to obtain encoded signal parameter information of the audio signal in each format, write the encoded signal parameter information of the audio signal in each format into an encoded stream and send the encoded stream to a decoding end.

[00321] In conclusion, in the signal encoding and decoding apparatus provided by the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal), the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00322] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00323] determine an encoding mode of the channel-based audio signal based on a signal characteristic of the channel-based audio signal;

[00324] determine an encoding mode of the object-based audio signal based on a

signal characteristic of the object-based audio signal; and

[00325] determine an encoding mode of the scene-based audio signal based on a signal characteristic of the scene-based audio signal.

[00326] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00327] obtain a number of object signals included in the channel-based audio signal;

[00328] determine whether the number of the object signals included in the channel-based audio signal is less than a first threshold;

[00329] in response to the number of the object signals included in the channel-based audio signal being less than the first threshold, determine that the encoding mode of the channel-based audio signal is at least one of:

[00330] encoding each object signal in the channel-based audio signal using an object signal encoding kernel;

[00331] obtaining input first command line control information, and encoding at least part of the object signals in the channel-based audio signal using the object signal encoding kernel based on the first command line control information, in which the first command line control information is configured to indicate object signals that need to be encoded among the object signals included in the channel-based audio signal, and a number of the object signals that need to be encoded is greater than or equal to 1 and less than the number of the object signals included in the channel-based audio signal.

[00332] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00333] obtain a number of object signals included in the channel-based audio signal;

[00334] determine whether the number of the object signals included in the channel-based audio signal is less than a first threshold;

[00335] in response to the number of the object signals included in the channel-based audio signal being not less than the first threshold, determine that the encoding mode of the channel-based audio signal is at least one of:

[00336] converting the channel-based audio signal into a first audio signal in another format, and encoding the first audio signal in another format using an encoding kernel corresponding to the first audio signal in another format, in which a number of channels of the first audio signal in another format is less than a number of channels of the channel-based

audio signal;

[00337] obtaining input first command line control information, and encoding at least part of the object signals in the channel-based audio signal using an object signal encoding kernel based on the first command line control information, in which the first command line control information is configured to indicate object signals that need to be encoded among the object signals included in the channel-based audio signal, and a number of the object signals that need to be encoded is greater than or equal to 1 and less than the number of the object signals included in the channel-based audio signal;

[00338] obtaining input second command line control information, and encoding at least part of channel signals in the channel-based audio signal using the object signal encoding kernel based on the second command line control information, in which the second command line control information is configured to indicate channel signals that need to be encoded among the channel signals included in the channel-based audio signal, and a number of the channel signals that need to be encoded is greater than or equal to 1 and less than a number of the channel signals included in the channel-based audio signal.

[00339] Alternatively, in an embodiment of the disclosure, the encoding module is further configured to:

[00340] encode the channel-based audio signal using the encoding mode of the channel-based audio signal.

[00341] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00342] perform a signal characteristic analysis on the object-based audio signal to obtain an analysis result;

[00343] perform a classification on the object-based audio signal to obtain a first type of object signal set and a second type of object signal set, in which each of the first type of object signal set and the second type of object signal set includes at least one object-based audio signal;

[00344] determine an encoding mode corresponding to the first type of object signal set; and

[00345] perform a classification on the second type of object signal set based on the analysis result to obtain at least one object signal subset, and determining an encoding mode corresponding to each object signal subset based on a classification result, in which the object

signal subset includes at least one object-based audio signal.

[00346] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00347] classify one or more signals that need not to be individually operated and processed in the object-based audio signal into the first type of object signal set, and classifying remaining signals into the second type of object signal set.

[00348] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00349] determine that the encoding mode corresponding to the first type of object signal set includes: performing first pre-rendering processing on an object-based audio signal in the first type of object signal set, and encoding the signal after the first pre-rendering processing using a multi-channel encoding kernel.

[00350] The first pre-rendering processing includes: performing signal format conversion processing on an object-based audio signal to convert the object-based audio signal into a channel-based audio signal.

[00351] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00352] classify one or more signals belonging to a background sound in the object-based audio signal into the first type of object signal set, and classify remaining signals into the second type of object signal set.

[00353] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00354] determine that the encoding mode corresponding to the first type of object signal set includes: performing second pre-rendering processing on an object-based audio signal in the first type of object signal set, and encoding the signal after the second pre-rendering processing using a high order ambisonics (HOA) encoding kernel.

[00355] The second pre-rendering processing comprises: performing signal format conversion processing on an object-based audio signal to convert the object-based audio signal into a scene-based audio signal.

[00356] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00357] classify one or more signals that need not to be individually operated and

processed in the object-based audio signal into the first object signal subset, classify one or more signals belonging to a background sound in the object-based audio signal into the second object signal subset, and classify remaining signals into the second type of object signal set.

[00358] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00359] determine that an encoding mode corresponding to the first object signal subset in the first type of object signal set includes: performing first pre-rendering processing on an object-based audio signal in the first object signal subset, and encoding the signal after the first pre-rendering processing using a multi-channel encoding kernel; in which the first pre-rendering processing includes: performing signal format conversion processing on an object-based audio signal to convert the object-based audio signal into a channel-based audio signal; and

[00360] determine that an encoding mode corresponding to the second object signal subset in the first type of object signal set includes: performing second pre-rendering processing on an object-based audio signal in the second object signal subset, and encoding the signal after the second pre-rendering processing using an HOA encoding kernel; in which the second pre-rendering processing includes: performing signal format conversion processing on an object-based audio signal to convert the object-based audio signal into a scene-based audio signal.

[00361] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00362] perform high-pass filtering processing on object-based audio signals; and

[00363] perform a correlation analysis on the signals after the high-pass filtering processing to determine cross-correlation parameter values between the object-based audio signals.

[00364] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00365] set a normalized correlation degree interval based on correlation degrees; and

[00366] perform the classification on the second type of object signal set based on the cross-correlation parameter values of the object-based audio signals and the normalized

correlation degree interval to obtain the at least one object signal subset, and determine the corresponding encoding mode based on a correlation degree corresponding to the at least one object signal subset.

[00367] Alternatively, in an embodiment of the disclosure,

[00368] the encoding mode corresponding to the object signal subset includes an independent encoding mode or a joint encoding mode.

[00369] Alternatively, in an embodiment of the disclosure, the independent encoding mode corresponds to a time-domain processing manner or a frequency-domain processing manner.

[00370] In response to an object signal in the object signal subset being a speech signal or a speech-like signal, the independent encoding mode adopts the time-domain processing manner.

[00371] In response to an object signal in the object signal subset being an audio signal in another format other than the speech signal or the speech-like signal, the independent encoding mode adopts the frequency-domain processing manner.

[00372] Alternatively, in an embodiment of the disclosure, the encoding module is further configured to:

[00373] encode the object-based audio signal using the encoding mode of the object-based audio signal;

[00374] in which encoding the object-based audio signal using the encoding mode of the object-based audio signal includes:

[00375] encoding one or more signals in the first type of object signal set using an encoding mode corresponding to the first type of object signal set;

[00376] performing preprocessing on one or more object signal subsets in the second type of object signal set, and encoding all the object signal subsets after the preprocessing in the second type of object signal set using respective encoding modes and using a same object signal encoding kernel.

[00377] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00378] analyze a frequency-band bandwidth range of the object signal.

[00379] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00380] determine bandwidth intervals corresponding to different frequency-band bandwidths;

[00381] perform the classification on the second type of object signal set to obtain the at least one object signal subset based on the frequency-band bandwidth range of the object-based audio signal and the bandwidth intervals corresponding to different frequency-band bandwidths, and determine a corresponding encoding mode based on a frequency-band bandwidth corresponding to the at least one object signal subset.

[00382] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00383] obtain input third command line control information, in which the third command line control information is configured to indicate a frequency-band bandwidth range to be encoded corresponding to the object-based audio signal;

[00384] perform the classification on the second type of object signal set by combining the third command line control information and the analysis result to obtain the at least one object signal subset, and determine the encoding mode corresponding to each object signal subset based on the classification result.

[00385] Alternatively, in an embodiment of the disclosure, the encoding module is further configured to:

[00386] encode the object-based audio signal using the encoding mode of the object-based audio signal;

[00387] in which encoding the object-based audio signal using the encoding mode of the object-based audio signal includes:

[00388] encoding one or more signals in the first type of object signal set using the encoding mode corresponding to the first type of object signal set;

[00389] performing preprocessing on object signal subsets in the second type of object signal set, and encoding different object signal subsets after the preprocessing using respective encoding modes and using different object signal encoding kernels.

[00390] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00391] obtain a number of object signals included in the scene-based audio signal;

[00392] determine whether the number of the object signals included in the scene-based audio signal is less than a second threshold;

[00393] in response to the number of the object signals included in the scene-based audio signal being less than the second threshold, determine that the encoding mode of the scene-based audio signal is at least one of:

[00394] encoding each object signal in the scene-based audio signal using an object signal encoding kernel;

[00395] obtaining input fourth command line control information, and encoding at least part of the object signals in the scene-based audio signal using the object signal encoding kernel based on the fourth command line control information, in which the fourth command line control information is configured to indicate object signals that need to be encoded among the object signals included in the scene-based audio signal, and a number of the object signals that need to be encoded is greater than or equal to 1 and less than the number of the object signals included in the scene-based audio signal.

[00396] Alternatively, in an embodiment of the disclosure, the determining module is further configured to:

[00397] obtain a number of object signals included in the scene-based audio signal;

[00398] determine whether the number of the object signals included in the scene-based audio signal is less than a second threshold;

[00399] in response to the number of the object signals included in the scene-based audio signal being not less than the second threshold, determine that the encoding mode of the scene-based audio signal is at least one of:

[00400] convert the scene-based audio signal into a second audio signal in another format, and encode the second audio signal in another format using a scene signal encoding kernel, in which a number of channels of the second audio signal in another format is smaller than a number of channels of the scene-based audio signal;

[00401] perform a low-order conversion on the scene-based audio signal to covert the scene-based audio signal to a scene-based audio signal with a lower order than a current order of the scene-based audio signal, and encode the scene-based audio signal with the lower order using the scene signal encoding kernel.

[00402] Alternatively, in an embodiment of the disclosure, the encoding module is further configured to:

[00403] encode the scene-based audio signal using the encoding mode of the scene-based audio signal.

[00404] Alternatively, in an embodiment of the disclosure, the encoding module is further configured to:

[00405] determine a classification side information parameter, in which the classification side information parameter is configured to indicate a classification manner for the second type of object signal set;

[00406] determine a side information parameter corresponding to the audio signal in each format, in which the side information parameter is configured to indicate the encoding mode corresponding to the audio signal in each format;

[00407] perform code stream multiplexing on the classification side information parameter, the side information parameter corresponding to the audio signal in each format, and the encoded signal parameter information of the audio signal in each format to obtain the encoded stream, and send the encoded stream to the decoding end.

[00408] FIG. 19 is a block diagram of a signal encoding and decoding apparatus according to an embodiment of the disclosure. The apparatus is applied to a decoding end. As illustrated in FIG. 19, the apparatus may include:

[00409] a receiving module 1901, configured to receive an encoded stream sent by an encoding end;

[00410] a decoding module 1902, configured to decode the encoded stream to obtain an audio signal in a mixed format, in which the audio signal in the mixed format comprises at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

[00411] In conclusion, in the signal encoding and decoding apparatus provided by the embodiment of the disclosure, firstly, the audio signal in the mixed format is obtained, and the audio signal in the mixed format includes at least one format of the channel-based audio signal, the object-based audio signal, and the scene-based audio signal. The encoding mode of the audio signal in each format is determined based on signal characteristics of the audio signals in different formats. The audio signal in each format is encoded using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format, the encoded signal parameter information of the audio signal in each format is written into the encoded stream and the encoded stream is sent to the decoding end. It can be seen that, in the embodiment of the disclosure, when encoding the audio signal in the mixed format (also called the mixed-format audio signal),

the audio signals in different formats are reorganized and analyzed based on the characteristics of the audio signals in different formats, and for the audio signals in different formats, adaptive encoding modes are determined and then the corresponding encoding kernels are used for encoding, thereby achieving a better encoding efficiency.

[00412] Alternatively, in an embodiment of the disclosure, the apparatus is further configured to:

[00413] perform a code stream analysis on the encoded stream to obtain a classification side information parameter, a side information parameter corresponding to an audio signal in each format, and encoded signal parameter information of the audio signal in each format.

[00414] The classification side information parameter is configured to indicate a classification manner for a second type of object signal set of the object-based audio signal, and the side information parameter is configured to indicate an encoding mode corresponding to the audio signal in each format.

[00415] Alternatively, in an embodiment of the disclosure, the decoding module is further configured to:

[00416] decode encoded signal parameter information of the channel-based audio signal based on a side information parameter corresponding to the channel-based audio signal;

[00417] decode encoded signal parameter information of the object-based audio signal based on the classification side information parameter and a side information parameter corresponding to the object-based audio signal; and

[00418] decode encoded signal parameter information of the scene-based audio signal based on a side information parameter corresponding to the scene-based audio signal.

[00419] Alternatively, in an embodiment of the disclosure, the decoding module is further configured to:

[00420] determine, from the encoded signal parameter information of the object-based audio signal, encoded signal parameter information corresponding to a first type of object signal set and encoded signal parameter information corresponding to the second type of object signal set;

[00421] decode the encoded signal parameter information corresponding to the first type of object signal set based on a side information parameter corresponding to the first type of object signal set; and

[00422] decode the encoded signal parameter information corresponding to the second type of object signal set based on the classification side information parameter and the side information parameter corresponding to the second type of object signal set.

[00423] Alternatively, in an embodiment of the disclosure, the decoding module is further configured to:

[00424] determine the classification manner for the second type of object signal set based on the classification side information parameter;

[00425] decode the encoded signal parameter information corresponding to the second type of object signal set based on the classification manner for the second type of object signal set and the side information parameter corresponding to the second type of object signal set.

[00426] Alternatively, in an embodiment of the disclosure, the classification side information parameter indicates that the classification manner for the second type of object signal set is based on cross-correlation parameter values; the decoding module is further configured to:

[00427] decode the encoded signal parameter information of all signals in the second type of object signal set using a same object signal decoding kernel based on the classification manner for the second type of object signal set and the side information parameter corresponding to the second type of object signal set.

[00428] Alternatively, in an embodiment of the disclosure, the classification side information parameter indicates that the classification manner for the second type of object signal set is based on a frequency-band bandwidth range; the decoding module is further configured to:

[00429] decode the encoded signal parameter information of different signals in the second type of object signal set using different object signal decoding kernels based on the classification manner for the second type of object signal set and the side information parameter corresponding to the second type of object signal set.

[00430] Alternatively, in an embodiment of the disclosure, the apparatus is further configured to:

[00431] perform post-processing on the decoded object-based audio signal.

[00432] Alternatively, in an embodiment of the disclosure, the decoding module is further configured to:

[00433] determine an encoding mode corresponding to the channel-based audio signal based on the side information parameter corresponding to the channel-based audio signal; and

[00434] decode the encoded signal parameter information of the channel-based audio signal using a corresponding decoding mode based on the encoding mode corresponding to the channel-based audio signal.

[00435] Alternatively, in an embodiment of the disclosure, the decoding module is further configured to:

[00436] determine an encoding mode corresponding to the scene-based audio signal based on the side information parameter corresponding to the scene-based audio signal; and

[00437] decode the encoded signal parameter information of the scene-based audio signal using a corresponding decoding mode based on the encoding mode corresponding to the scene-based audio signal.

[00438] FIG. 20 is a block diagram of a user equipment UE2000 according to an embodiment of the disclosure. The UE2000 may be a mobile phone, a computer, a digital broadcasting terminal, a message transceiver device, a game console, a tablet device, a medical device, a fitness device and a personal digital assistant, etc.

[00439] As illustrated in FIG. 20, the UE2000 may include one or more of the following components: a processing component 2002, a memory 2004, a power component 2006, a multimedia component 2008, an audio component 2010, an input/output (I/O) interface 2012, a sensor component 2013, and a communication component 2016.

[00440] The processing component 2002 typically controls overall operations of the UE2000, such as the operations associated with display, telephone calls, data communications, camera operations, and recording operations. The processing component 2002 may include one or more processors 2020 to execute instructions to implement all or part of the steps in the above described method for reporting location-related information. Moreover, the processing component 2002 may include one or more modules which facilitate the interaction between the processing component 2002 and other components. For example, the processing component 2002 may include a multimedia module to facilitate the interaction between the multimedia component 2008 and the processing component 2002.

[00441] The memory 2004 is configured to store various types of data to support the operation of the UE2000. Examples of such data include instructions for any applications or

methods operated on the UE2000, contact data, phonebook data, messages, pictures, videos, etc. The memory 2004 may be implemented using any type of volatile or non-volatile memory devices, or a combination thereof, such as a Static Random-Access Memory (SRAM), an Electrically-Erasable Programmable Read Only Memory (EEPROM), an Erasable Programmable Read Only Memory (EPROM), a Programmable Read Only Memory (PROM), a Read Only Memory (ROM), a magnetic memory, a flash memory, a magnetic or optical disk.

[00442] The power component 2006 provides power to various components of the UE2000. The power component 2006 may include a power management system, one or more power sources, and any other components associated with the generation, management, and distribution of power in the UE2000.

[00443] The multimedia component 2008 includes a screen providing an output interface between the UE2000 and the user. In some embodiments, the screen may include a Liquid Crystal Display (LCD) and a Touch Panel (TP). If the screen includes the touch panel, the screen may be implemented as a touch screen to receive input signals from the user. The touch panel includes one or more touch sensors to sense touches, swipes, and gestures on the touch panel. The touch sensor may not only sense a boundary of a touch or swipe action, but also sense a period of time and a pressure associated with the touch or swipe action. In some embodiments, the multimedia component 2008 includes a front-facing camera and/or a rear-facing camera. When the UE2000 is in an operating mode, such as a shooting mode or a video mode, the front-facing camera and/or the rear-facing camera can receive external multimedia data. Each front-facing camera and rear-facing camera may be a fixed optical lens system or has a focal length and optical zoom capability.

[00444] The audio component 2010 is configured to output and/or input audio signals. For example, the audio component 2010 includes a microphone (MIC) configured to receive an external audio signal when the UE2000 is in an operation mode, such as a call mode, a recording mode, and a voice recognition mode. The received audio signal may be further stored in the memory 2004 or transmitted via the communication component 2016. In some embodiments, the audio component 2010 further includes a speaker to output audio signals.

[00445] The I/O interface 2012 provides an interface between the processing component 2002 and peripheral interface modules, such as a keyboard, a click wheel, buttons,

and the like. The buttons may include, but are not limited to, a home button, a volume button, a starting button, and a locking button.

[00446] The sensor component 2013 includes one or more sensors to provide status assessments of various aspects of the UE2000. For instance, the sensor component 2013 may detect an open/closed status of the UE2000, relative positioning of components, e.g., the display and the keypad, of the UE2000, a change in position of the UE2000 or a component of the UE2000, a presence or absence of user contact with the UE2000, an orientation or an acceleration/deceleration of the UE2000, and a change in temperature of the UE2000. The sensor component 2013 may include a proximity sensor configured to detect the presence of nearby objects without any physical contact. The sensor component 2013 may also include a light sensor, such as a Complementary Metal Oxide Semiconductor (CMOS) or Charge-Coupled Device (CCD) image sensor, for use in imaging applications. In some embodiments, the sensor component 2013 may also include an accelerometer sensor, a gyroscope sensor, a magnetic sensor, a pressure sensor, or a temperature sensor.

[00447] The communication component 2016 is configured to facilitate communication, wired or wirelessly, between the UE2000 and other devices. The UE2000 can access a wireless network based on a communication standard, such as Wi-Fi, 2G or 3G, or a combination thereof. In an exemplary embodiment, the communication component 2016 receives a broadcast signal from an external broadcast management system or broadcast associated information via a broadcast channel. In an exemplary embodiment, the communication component 2016 further includes a Near Field Communication (NFC) module to facilitate short-range communication. For example, the NFC module may be implemented based on a RF Identification (RFID) technology, an Infrared Data Association (IrDA) technology, an Ultra-Wide Band (UWB) technology, a Blue Tooth (BT) technology, and other technologies.

[00448] In the exemplary embodiment, the UE2000 may be implemented with one or more Application Specific Integrated Circuits (ASICs), Digital Signal Processors (DSPs), Digital Signal Processing Devices (DSPDs), Programmable Logic Devices (PLDs), Field Programmable Gate Arrays (FPGAs), controllers, micro-controllers, microprocessors or other electronic components, for performing the above method.

[00449] FIG. 21 is a block diagram of a network side device 2100 according to an embodiment of the disclosure. The network side device 2100 may be provided as a network

side device. As illustrated in FIG. 21, the network side device 2100 includes a processing component 2111 consisting of one or more processors, and memory resources represented by a memory 2132 for storing instructions that may be executed by the processing component 2122, such as applications. The applications stored in the memory 2132 may include one or more modules each corresponding to a set of instructions. In addition, the processing component 2122 is configured to execute the instructions to implement the method applied to the network side device as described above, for example, the method illustrated in FIG. 1.

[00450] The network side device 2100 may also include a power component 2126 configured to perform power management of the network side device 2100, a wired or wireless network interface 2150 configured to connect the network side device 2100 to a network, and an input/output (I/O) interface 2158. The network side device 2100 may operate based on an operating system stored in the memory 2132, such as Windows Server TM, Mac OS XTM, UnixTM, LinuxTM, FreeBSDTM, or the like.

[00451] In the above embodiments of the disclosure, the methods according to the embodiments of the disclosure are described from the perspectives of the network side device and the UE, respectively. In order to realize each of the functions in the methods according to the above embodiments of the disclosure, the network side device and the UE may include a hardware structure, a software module, and realize each of the above functions in the form of hardware structure, software module, or a combination of hardware structure and software module. A certain function of the above functions may be performed in the form of hardware structure, software module, or a combination of hardware structure and software module.

[00452] An embodiment of the disclosure further provides a communication apparatus. The communication apparatus may include a transceiver module and a processing module. The transceiver module may include a sending module and/or a receiving module. The sending module is configured to implement a sending function. The receiving module is configured to implement a receiving function. The transceiver module may implement the sending function and/or the receiving function.

[00453] The communication apparatus may be a terminal device (such as the terminal device in the foregoing method embodiments), or may be an apparatus in the terminal device, or may be an apparatus capable of being used in combination with the terminal device. Alternatively, the communication apparatus may be a network device, or may be an apparatus in the network device, or may be an apparatus capable of being used in

combination with the network device.

[00454] An embodiment of the disclosure further provides another communication apparatus. The communication apparatus may be a network device or a terminal device (such as the terminal device in the foregoing method embodiments), or may be a chip, a chip system or a processor that supports the network device to realize the above-described methods, or may be a chip, a chip system or a processor that supports the terminal device to realize the above-described methods. The apparatus may be used to realize the methods described in the above method embodiments with reference to the description of the above-described method embodiments.

[00455] The communication device may include one or more processors. The processor may be a general purpose processor or a dedicated processor, such as, a baseband processor and a central processor. The baseband processor is used for processing communication protocols and communication data. The central processor is used for controlling the communication apparatus (e.g., a base station, a baseband chip, a terminal device, a terminal device chip, a DU, or a CU), executing computer programs, and processing data of the computer programs.

[00456] Optionally, the communication apparatus may include one or more memories on which computer programs may be stored. The processor executes the computer programs to cause the communication apparatus to perform the methods described in the above method embodiments. Alternatively, the memory may also store data. The communication apparatus and the memory may be provided separately or may be integrated together.

[00457] Optionally, the communication apparatus may also include a transceiver and an antenna. The transceiver may be referred to as a transceiver unit, a transceiver machine, or a transceiver circuit, for realizing a transceiver function. The transceiver may include a receiver and a transmitter. The receiver may be referred to as a receiving machine or a receiving circuit, for realizing the receiving function. The transmitter may be referred to as a transmitter machine or a transmitting circuit, for realizing the transmitting function.

[00458] Optionally, the communication apparatus may also include one or more interface circuits. The interface circuits are used to receive code instructions and transmit them to the processor. The processor runs the code instructions to cause the communication apparatus 70 to perform the method described in the method embodiments.

[00459] The communication apparatus is the terminal device (such as the terminal device in the foregoing method embodiments): the processor is configured to perform the method illustrated in any of FIG. 1 to FIG. 4.

[00460] The communication apparatus is the network device, and the transceiver is configured to perform the method illustrated in any of FIG. 5 to FIG. 7.

[00461] In an implementation, the processor may include a transceiver for implementing the receiving and sending functions. The transceiver may be, for example, a transceiver circuit, an interface, or an interface circuit. The transceiver circuit, the interface, or the interface circuit for implementing the receiving and sending functions may be separated or may be integrated together. The transceiver circuit, the interface, or the interface circuit described above may be used for reading and writing code/data, or may be used for signal transmission or delivery.

[00462] In an implementation, the processor may store a computer program. When the computer program runs on the processor, the communication apparatus is caused to perform the methods described in the method embodiments above. The computer program may be solidified in the processor, and in such case the processor may be implemented by hardware.

[00463] In an implementation, the communication apparatus may include circuits. The circuits may implement the sending, receiving or communicating function in the preceding method embodiments. The processor and the transceiver described in this disclosure may be implemented on integrated circuits (ICs), analog ICs, radio frequency integrated circuits (RFICs), mixed signal ICs, application specific integrated circuits (ASICs), printed circuit boards (PCBs), and electronic devices. The processor and the transceiver can also be produced using various IC process technologies, such as complementary metal oxide semiconductor (CMOS), nMetal-oxide-semiconductor (NMOS), positive channel metal oxide semiconductor (PMOS), bipolar junction transistor (BJT), bipolar CMOS (BiCMOS), silicon-germanium (SiGe), gallium arsenide (GaAs) and so on.

[00464] The communication apparatus in the description of the above embodiments may be a network device or a terminal device (such as the terminal device in the foregoing method embodiments), but the scope of the communication apparatus described in the disclosure is not limited thereto, and the structure of the communication apparatus may not be limited. The communication apparatus may be a stand-alone device or may be part of a

larger device. For example, the described communication apparatus may be:

[00465] (1) a stand-alone IC, a chip, a chip system or a subsystem;

[00466] (2) a collection of ICs including one or more ICs, optionally, the collection of ICs may also include storage components for storing data and computer programs;

[00467] (3) an ASIC, such as a modem;

[00468] (4) a module that can be embedded within other devices;

[00469] (5) a receiver, a terminal device, a smart terminal device, a cellular phone, a wireless device, a handheld machine, a mobile unit, an in-vehicle device, a network device, a cloud device, an artificial intelligence device, and the like; and

[00470] (6) others.

[00471] The case where the communication apparatus may be a chip or a chip system is described with reference to the schematic structure of the chip. The chip includes a processor and an interface. There may be one or more processors, and there may be multiple interfaces.

[00472] Alternatively, the chip further includes a memory, and the memory is configured to store necessary computer programs and data.

[00473] It is understood by those skilled in the art that various illustrative logical blocks and steps listed in the embodiments of the disclosure may be implemented by electronic hardware, computer software, or a combination of both. Whether such function is implemented by hardware or software depends on the particular application and the design requirements of the entire system. Those skilled in the art may, for each particular application, use various methods to implement the described function, but such implementation should not be understood as beyond the scope of protection of the embodiments of the disclosure.

[00474] The embodiments of the present disclosure further provide a system, in which the system includes a communication apparatus acting as a terminal device (such as the first terminal device in the foregoing method embodiments) and a communication apparatus acting as a network device in the foregoing embodiment, or the system include a communication apparatus acting as a terminal device (such as the first terminal device in the foregoing method embodiments) and a communication apparatus acting as a network device in the foregoing embodiment.

[00475] The present disclosure further provides a readable storage medium having stored thereon instructions that, when executed by a computer, the functions of any of the

foregoing method embodiments are implemented.

[00476] The present disclosure further provides a computer program product. When the computer program product is executed by a computer, the function of any of the method embodiments described above is implemented.

[00477] The above embodiments may be implemented in whole or in part by software, hardware, firmware, or any combination thereof. When implemented using software, the above embodiments may be implemented, in whole or in part, in the form of a computer program product. The computer program product includes one or more computer programs. When loading and executing the computer program on the computer, all or part of processes or functions described in the embodiments of the disclosure is implemented. The computer may be a general-purpose computer, a dedicated computer, a computer network, or other programmable devices. The computer program may be stored in a computer-readable storage medium or transmitted from one computer-readable storage medium to another computer-readable storage medium. For example, the computer program may be transmitted from one web site, computer, server, or data center to another web site, computer, server, or data center, in a wired manner (e.g., using coaxial cables, fiber optics, or digital subscriber lines (DSLs) or wireless manner (e.g., using infrared wave, wireless wave, or microwave). The computer-readable storage medium may be any usable medium to which the computer is capable to access or a data storage device such as a server integrated by one or more usable mediums and a data center. The usable medium may be a magnetic medium (e.g., a floppy disk, a hard disk, and a tape), an optical medium (e.g., a high-density digital video disc (DVD)), or a semiconductor medium (e.g., a solid state disk (SSD)).

[00478] Those skilled in the art can understand that the first, second, and other various numerical numbers involved in the disclosure are only described for the convenience of differentiation, and are not used to limit the scope of the embodiments of the disclosure, or used to indicate the order of precedence.

[00479] The term “at least one” in the disclosure may also be described as one or more, and the term “multiple” may be two, three, four, or more, which is not limited in the disclosure. In the embodiments of the disclosure, for a type of technical features, “first”, “second”, and “third”, and “A”, “B”, “C” and “D” are used to distinguish different technical features of the type, the technical features described using the “first”, “second”, and “third”, and “A”, “B”, “C” and “D” do not indicate any order of precedence or magnitude.

[00480] Other embodiments of the disclosure will be apparent to those skilled in the art from consideration of the specification and practice of the disclosure disclosed here. The disclosure is intended to cover any variations, usages, or adaptations of the embodiments of the disclosure following the general principles thereof and including such departures from the disclosure as come within known or customary practice in the art. It is intended that the specification and embodiments are considered as exemplary only, with a true scope and spirit of the disclosure being indicated by the following claims.

[00481] It will be appreciated that the disclosure is not limited to the exact construction that has been described above and illustrated in the accompanying drawings, and that various modifications and changes can be made without departing from the scope thereof. It is intended that the scope of the disclosure only be limited by the appended claims.

WHAT IS CLAIMED IS:

1. A signal encoding and decoding method, performed by an encoding end, comprising:

obtaining an audio signal in a mixed format, wherein the audio signal in the mixed format comprises at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal;

determining, based on signal characteristics of audio signals in different formats, an encoding mode of the audio signal in each format; and

encoding the audio signal in each format using the encoding mode of the audio signal in each format to obtain encoded signal parameter information of the audio signal in each format, writing the encoded signal parameter information of the audio signal in each format into an encoded stream and sending the encoded stream to a decoding end.

2. The method of claim 1, wherein determining, based on the signal characteristics of the audio signals in different formats, the encoding mode of the audio signal in each format comprises:

determining an encoding mode of the channel-based audio signal based on a signal characteristic of the channel-based audio signal;

determining an encoding mode of the object-based audio signal based on a signal characteristic of the object-based audio signal; and

determining an encoding mode of the scene-based audio signal based on a signal characteristic of the scene-based audio signal.

3. The method of claim 2, wherein determining the encoding mode of the channel-based audio signal based on the signal characteristic of the channel-based audio signal comprises:

obtaining a number of object signals included in the channel-based audio signal;

determining whether the number of the object signals included in the channel-based audio signal is less than a first threshold;

in response to the number of the object signals included in the channel-based audio signal being less than the first threshold, determining that the encoding mode of the channel-based audio signal is at least one of:

encoding each object signal in the channel-based audio signal using an object signal encoding kernel;

obtaining input first command line control information, and encoding at least part of the object signals in the channel-based audio signal using the object signal encoding kernel based on the first command line control information, wherein the first command line control information is configured to indicate object signals that need to be encoded among the object signals included in the channel-based audio signal, and a number of the object signals that need to be encoded is greater than or equal to 1 and less than the number of the object signals included in the channel-based audio signal.

4. The method of claim 2, wherein determining the encoding mode of the channel-based audio signal based on the signal characteristic of the channel-based audio signal comprises:

obtaining a number of object signals included in the channel-based audio signal;

determining whether the number of the object signals included in the channel-based audio signal is less than a first threshold;

in response to the number of the object signals included in the channel-based audio signal being not less than the first threshold, determining that the encoding mode of the channel-based audio signal is at least one of:

converting the channel-based audio signal into a first audio signal in another format, and encoding the first audio signal in another format using an encoding kernel corresponding to the first audio signal in another format, wherein a number of channels of the first audio signal in another format is smaller than a number of channels of the channel-based audio signal;

obtaining input first command line control information, and encoding at least part of the object signals in the channel-based audio signal using an object signal encoding kernel based on the first command line control information, wherein the first command line control information is configured to indicate object signals that need to be encoded among the object signals included in the channel-based audio signal, and a number of the object signals that need to be encoded is greater than or equal to 1 and less than the number of the object signals included in the channel-based audio signal;

obtaining input second command line control information, and encoding at least part of channel signals in the channel-based audio signal using the object signal encoding kernel based on the second command line control information, wherein the second command line control information is configured to indicate channel signals that need to be encoded among

the channel signals included in the channel-based audio signal, and a number of the channel signals that need to be encoded is greater than or equal to 1 and less than a number of the channel signals included in the channel-based audio signal.

5. The method of claim 2, wherein determining the encoding mode of the object-based audio signal based on the signal characteristic of the object-based audio signal comprises:

performing a signal characteristic analysis on the object-based audio signal to obtain an analysis result;

performing a classification on the object-based audio signal to obtain a first type of object signal set and a second type of object signal set, wherein each of the first type of object signal set and the second type of object signal set comprises at least one object-based audio signal;

determining an encoding mode corresponding to the first type of object signal set; and

performing a classification on the second type of object signal set based on the analysis result to obtain at least one object signal subset, and determining an encoding mode corresponding to each object signal subset based on a classification result, wherein the object signal subset comprises at least one object-based audio signal.

6. The method of claim 5, wherein performing the classification on the object-based audio signal to obtain the first type of object signal set and the second type of object signal set comprises one of the following:

classifying one or more signals that need not to be individually operated and processed in the object-based audio signal into the first type of object signal set, and classifying remaining signals into the second type of object signal set;

classifying one or more signals belonging to a background sound in the object-based audio signal into the first type of object signal set, and classifying remaining signals into the second type of object signal set; or

classifying one or more signals that need not to be individually operated and processed in the object-based audio signal into a first object signal subset, classifying one or more signals belonging to a background sound in the object-based audio signal into a second object signal subset, and classifying remaining signals into the second type of object signal set, wherein the first type of object signal set comprises the first object signal subset and the

second object signal subset.

7. The method of claim 6, wherein determining the encoding mode corresponding to the first type of object signal set comprises one of the following:

performing first pre-rendering processing on an object-based audio signal in the first type of object signal set, and encoding the signal after the first pre-rendering processing using a multi-channel encoding kernel; wherein, the first type of object signal set comprises the one or more signals that need not to be individually operated and processed; wherein, the first pre-rendering processing comprises: performing signal format conversion processing on an object-based audio signal to convert the object-based audio signal into a channel-based audio signal;

performing second pre-rendering processing on an object-based audio signal in the first type of object signal set, and encoding the signal after the second pre-rendering processing using a high order ambisonics (HOA) encoding kernel; wherein, the first type of object signal set comprises the one or more signals belonging to the background sound; and the second pre-rendering processing comprises: performing signal format conversion processing on an object-based audio signal to convert the object-based audio signal into a scene-based audio signal; or

performing first pre-rendering processing on an object-based audio signal in the first object signal subset, and encoding the signal after the first pre-rendering processing using a multi-channel encoding kernel; wherein the first pre-rendering processing comprises: performing signal format conversion processing on an object-based audio signal to convert the object-based audio signal into a channel-based audio signal; and performing second pre-rendering processing on an object-based audio signal in the second object signal subset, and encoding the signal after the second pre-rendering processing using an HOA encoding kernel; wherein the second pre-rendering processing comprises: performing signal format conversion processing on an object-based audio signal to convert the object-based audio signal into a scene-based audio signal.

8. The method of claim 7, wherein performing the signal characteristic analysis on the object-based audio signal to obtain the analysis result comprises:

performing high-pass filtering processing on object-based audio signals; and

performing a correlation analysis on the signals after the high-pass filtering processing to determine cross-correlation parameter values between the object-based audio signals.

9. The method of claim 8, wherein performing the classification on the second type of object signal set based on the analysis result to obtain at least one object signal subset, and determining the encoding mode corresponding to each object signal subset based on the classification result comprises:

setting a normalized correlation degree interval based on correlation degrees; and

performing the classification on the second type of object signal set based on the cross-correlation parameter values of the object-based audio signals and the normalized correlation degree interval to obtain the at least one object signal subset, and determining the corresponding encoding mode based on a correlation degree corresponding to the at least one object signal subset.

10. The method of claim 9, wherein the encoding mode corresponding to the object signal subset comprises an independent encoding mode or a joint encoding mode.

11. The method of claim 9, wherein encoding the audio signal in each format using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format comprises:

encoding the object-based audio signal using the encoding mode of the object-based audio signal;

wherein encoding the object-based audio signal using the encoding mode of the object-based audio signal comprises:

encoding one or more signals in the first type of object signal set using an encoding mode corresponding to the first type of object signal set;

performing preprocessing on one or more object signal subsets in the second type of object signal set, and encoding all the object signal subsets after the preprocessing in the second type of object signal set using respective encoding modes and using a same object signal encoding kernel.

12. The method of claim 7, wherein performing the signal characteristic analysis on the

object-based audio signal to obtain the analysis result comprises:

analyzing a frequency-band bandwidth range of the object-based audio signal.

13. The method of claim 12, wherein performing the classification on the second type of object signal set based on the analysis result to obtain at least one object signal subset, and determining the encoding mode corresponding to each object signal subset based on the classification result comprises one of the following:

determining bandwidth intervals corresponding to different frequency-band bandwidths; performing the classification on the second type of object signal set to obtain the at least one object signal subset based on the frequency-band bandwidth range of the object-based audio signal and the bandwidth intervals corresponding to different frequency-band bandwidths, and determining a corresponding encoding mode based on a frequency-band bandwidth corresponding to the at least one object signal subset; or

obtaining input third command line control information, wherein the third command line control information is configured to indicate a frequency-band bandwidth range to be encoded corresponding to the object-based audio signal; and performing the classification on the second type of object signal set by combining the third command line control information and the analysis result to obtain the at least one object signal subset, and determining the encoding mode corresponding to each object signal subset based on the classification result.

14. The method of claim 12, wherein encoding the audio signal in each format using the encoding mode of the audio signal in each format to obtain the encoded signal parameter information of the audio signal in each format comprises:

encoding the object-based audio signal using the encoding mode of the object-based audio signal;

wherein encoding the object-based audio signal using the encoding mode of the object-based audio signal comprises:

encoding one or more signals in the first type of object signal set using the encoding mode corresponding to the first type of object signal set;

performing preprocessing on object signal subsets in the second type of object signal set, and encoding different object signal subsets after the preprocessing using respective

encoding modes and using different object signal encoding kernels.

15. The method of claim 2, wherein determining the encoding mode of the scene-based audio signal based on the signal characteristic of the scene-based audio signal comprises:

obtaining a number of object signals included in the scene-based audio signal;

determining whether the number of the object signals included in the scene-based audio signal is less than a second threshold;

in response to the number of the object signals included in the scene-based audio signal being less than the second threshold, determining that the encoding mode of the scene-based audio signal is at least one of:

encoding each object signal in the scene-based audio signal using an object signal encoding kernel;

obtaining input fourth command line control information, and encoding at least part of the object signals in the scene-based audio signal using the object signal encoding kernel based on the fourth command line control information, wherein the fourth command line control information is configured to indicate object signals that need to be encoded among the object signals included in the scene-based audio signal, and a number of the object signals that need to be encoded is greater than or equal to 1 and less than the number of the object signals included in the scene-based audio signal.

16. The method according to claim 2, wherein determining the encoding mode of the scene-based audio signal based on the signal characteristic of the scene-based audio signal comprises:

obtaining a number of object signals included in the scene-based audio signal;

determining whether the number of the object signals included in the scene-based audio signal is less than a second threshold;

in response to the number of the object signals included in the scene-based audio signal being not less than the second threshold, determining that the encoding mode of the scene-based audio signal is at least one of:

converting the scene-based audio signal into a second audio signal in another format, and encoding the second audio signal in another format using a scene signal encoding kernel, wherein a number of channels of the second audio signal in another format is smaller than a

number of channels of the scene-based audio signal;

performing a low-order conversion on the scene-based audio signal to covert the scene-based audio signal to a scene-based audio signal with a lower order than a current order of the scene-based audio signal, and encoding the scene-based audio signal with the lower order using the scene signal encoding kernel.

17. The method of claim 4 or 5 or 15, wherein writing the encoded signal parameter information of the audio signal in each format into the encoded stream and sending the encoded stream to the decoding end comprises:

determining a classification side information parameter, wherein the classification side information parameter is configured to indicate a classification manner for the second type of object signal set;

determining a side information parameter corresponding to the audio signal in each format, wherein the side information parameter is configured to indicate the encoding mode corresponding to the audio signal in each format;

performing code stream multiplexing on the classification side information parameter, the side information parameter corresponding to the audio signal in each format, and the encoded signal parameter information of the audio signal in each format to obtain the encoded stream, and sending the encoded stream to the decoding end.

18. A signal encoding and decoding method, performed by a decoding end, comprising:

receiving an encoded stream sent by an encoding end; and

decoding the encoded stream to obtain an audio signal in a mixed format, wherein the audio signal in the mixed format comprises at least one format of a channel-based audio signal, an object-based audio signal, and a scene-based audio signal.

19. The method of claim 18, further comprising:

performing a code stream analysis on the encoded stream to obtain a classification side information parameter, a side information parameter corresponding to an audio signal in each format, and encoded signal parameter information of the audio signal in each format;

wherein, the classification side information parameter is configured to indicate a classification manner for a second type of object signal set of the object-based audio signal,

and the side information parameter is configured to indicate an encoding mode corresponding to the audio signal in each format.

20. The method of claim 19, wherein decoding the encoded stream to obtain the audio signal in the mixed format comprises:

decoding encoded signal parameter information of the channel-based audio signal based on a side information parameter corresponding to the channel-based audio signal;

decoding encoded signal parameter information of the object-based audio signal based on the classification side information parameter and a side information parameter corresponding to the object-based audio signal; and

decoding encoded signal parameter information of the scene-based audio signal based on a side information parameter corresponding to the scene-based audio signal.

21. The method of claim 20, wherein decoding the encoded signal parameter information of the object-based audio signal based on the classification side information parameter and the side information parameter corresponding to the object-based audio signal comprises:

determining, from the encoded signal parameter information of the object-based audio signal, encoded signal parameter information corresponding to a first type of object signal set and encoded signal parameter information corresponding to the second type of object signal set;

decoding the encoded signal parameter information corresponding to the first type of object signal set based on a side information parameter corresponding to the first type of object signal set; and

decoding the encoded signal parameter information corresponding to the second type of object signal set based on the classification side information parameter and the side information parameter corresponding to the second type of object signal set.

22. The method of claim 21, wherein decoding the encoded signal parameter information corresponding to the second type of object signal set based on the classification side information parameter and the side information parameter corresponding to the second type of object signal set comprises:

determining the classification manner for the second type of object signal set based on the classification side information parameter;

decoding the encoded signal parameter information corresponding to the second type of object signal set based on the classification manner for the second type of object signal set and the side information parameter corresponding to the second type of object signal set.

23. The method of claim 22, wherein the classification side information parameter indicates that the classification manner for the second type of object signal set is based on cross-correlation parameter values or a frequency-band bandwidth range;

wherein decoding the encoded signal parameter information corresponding to the second type of object signal set based on the classification manner for the second type of object signal set and the side information parameter corresponding to the second type of object signal set comprises one of the following:

decoding the encoded signal parameter information of all signals in the second type of object signal set using a same object signal decoding kernel based on the classification manner for the second type of object signal set and the side information parameter corresponding to the second type of object signal set; or

decoding the encoded signal parameter information of different signals in the second type of object signal set using different object signal decoding kernels based on the classification manner for the second type of object signal set and the side information parameter corresponding to the second type of object signal set.

24. The method of any one of claims 21-23, further comprising:

performing post-processing on the decoded object-based audio signal.

25. The method of claim 20, wherein decoding the encoded signal parameter information of the channel-based audio signal based on the side information parameter corresponding to the channel-based audio signal comprises:

determining an encoding mode corresponding to the channel-based audio signal based on the side information parameter corresponding to the channel-based audio signal; and

decoding the encoded signal parameter information of the channel-based audio signal using a corresponding decoding mode based on the encoding mode corresponding to the

channel-based audio signal.

26. The method of claim 20, wherein decoding the encoded signal parameter information of the scene-based audio signal based on the side information parameter corresponding to the scene-based audio signal comprises:

determining an encoding mode corresponding to the scene-based audio signal based on the side information parameter corresponding to the scene-based audio signal; and

decoding the encoded signal parameter information of the scene-based audio signal using a corresponding decoding mode based on the encoding mode corresponding to the scene-based audio signal.

27. A communication apparatus, comprising a processor and a memory, wherein a computer program is stored in the memory, and the processor executes the computer program stored in the memory, to cause the apparatus to perform the method of any one of claims 1 to 17.

28. A communication apparatus, comprising a processor and a memory, wherein a computer program is stored in the memory, and the processor executes the computer program stored in the memory, to cause the apparatus to perform the method of any one of claims 18 to 26.