

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 703 199**

51 Int. Cl.:

G10L 19/02 (2013.01)

G10L 19/20 (2013.01)

G10L 19/22 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.06.2015 PCT/CN2015/082076**

87 Fecha y número de publicación internacional: **30.12.2015 WO15196968**

96 Fecha de presentación y número de la solicitud europea: **23.06.2015 E 15811228 (4)**

97 Fecha y número de publicación de la concesión europea: **26.09.2018 EP 3144933**

54 Título: **Método de codificación de audio y aparato**

30 Prioridad:

24.06.2014 CN 201410288983

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

07.03.2019

73 Titular/es:

**HUAWEI TECHNOLOGIES CO. LTD. (100.0%)
Huawei Administration Building Bantian
Longgang District
Shenzhen, Guangdong 518129, CN**

72 Inventor/es:

WANG, ZHE

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 703 199 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Método de codificación de audio y aparato

Campo técnico

5 La presente invención se refiere al campo de las tecnologías de procesamiento de señales y, más específicamente, a un método de codificación de audio y a un aparato.

Antecedentes

10 En la técnica anterior, un codificador híbrido se utiliza habitualmente para codificar una señal de audio en un sistema de comunicaciones de voz. Específicamente, el codificador híbrido habitualmente incluye dos subcodificadores. Un subcodificador es adecuado para codificar una señal de voz y el otro codificador es adecuado para codificar una señal de no voz. Para una señal de audio recibida, cada uno de los subcodificadores del codificador híbrido codifica la señal de audio. El codificador híbrido compara directamente la calidad de las señales de audio codificadas para seleccionar un subcodificador óptimo. Sin embargo, tal método de codificación de bucle cerrado tiene una alta complejidad de operación.

15 El documento WO 2008/045846 A1 da a conocer técnicas para codificar eficientemente una señal de entrada. En un diseño, un codificador generalizado codifica la señal de entrada (p. ej., una señal de audio) en base a al menos un detector y a múltiples codificadores. El al menos un detector puede incluir un detector de actividad de señal, un detector de señal similar al ruido, un detector de dispersión, algún otro detector o una combinación de los mismos. Los múltiples codificadores pueden incluir un codificador de silencio, un codificador de señal similar al ruido, un codificador de dominio del tiempo, un codificador de dominio de transformada, algún otro codificador o una combinación de los mismos. Las características de la señal de entrada se pueden determinar en base al por lo menos un detector. Se puede seleccionar un codificador entre los múltiples codificadores en base a las características de la señal de entrada. La señal de entrada puede codificarse en base al codificador seleccionado. La señal de entrada puede incluir una secuencia de tramas y, la detección y la codificación, pueden realizarse para cada una de las tramas.

25 El documento WO2012-024379 A2 da a conocer que un esquema para inyectar ruido en elementos no codificados de un espectro, se controla de acuerdo con una medida de una distribución de energía del espectro original entre las ubicaciones de los elementos no codificados.

Resumen

30 En vista de la técnica anterior, es un objeto de la invención proporcionar un método de codificación de audio y un aparato, que puedan reducir la complejidad de codificación y asegurar que la codificación sea de una precisión relativamente alta. Este objeto se resuelve mediante el método de codificación de acuerdo con la reivindicación 1 y el aparato de acuerdo con la reivindicación 4.

35 De acuerdo con un primer aspecto, se proporciona un método de codificación de audio, donde el método incluye: determinar la dispersión de distribución, en un espectro, de energía de N tramas de audio de entrada, donde las N tramas de audio incluyen una trama de audio actual y N es un número entero positivo; y determinar, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual, donde el primer método de codificación es un método de codificación que está basado en la transformada de tiempo-frecuencia y la cuantificación del coeficiente de transformada y que no está basado en predicción lineal y, el segundo método de codificación, es un método de codificación basado en predicción lineal.

45 Con referencia al primer aspecto, en una primera manera de implementación posible del primer aspecto, la dispersión de distribución, en un espectro, de energía de N tramas de audio de entrada incluye: dividir un espectro de cada una de las N tramas de audio en P envolventes espectrales, donde P es un número entero positivo; y determinar un parámetro de dispersión general de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde el parámetro de dispersión general indica la dispersión de distribución, en el espectro, de la energía de las N tramas de audio.

50 Con referencia a la primera manera de implementación posible del primer aspecto, en una segunda manera de implementación posible del primer aspecto, el parámetro de dispersión general incluye un primer ancho de banda mínimo; la determinación de un parámetro de dispersión general de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio incluye: determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio de acuerdo con el energía de las P envolventes espectrales de cada una de las N tramas de audio, donde el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio es el primer ancho de banda mínimo; y la determinación, de acuerdo con la

dispersión de distribución, en el espectro, de la energía de las N tramas de audio, de si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual incluye: cuando el primer ancho de banda mínimo es menor que un primer valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; o cuando el primer ancho de banda mínimo es mayor que el primer valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

Con referencia a la segunda manera de implementación posible del primer aspecto, en una tercera manera de implementación posible del primer aspecto, la determinación de un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio incluye: ordenar la energía de las P envolventes espectrales de cada una de las tramas de audio en orden descendente; determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio; y determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio, un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de las N tramas de audio.

Con referencia a la primera manera de implementación posible del primer aspecto, en una cuarta manera de implementación posible del primer aspecto, el parámetro de dispersión general incluye una primera proporción de energía; la determinación de un parámetro de dispersión general de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio incluye: seleccionar P_1 envolventes espectrales de las P envolventes espectrales de cada una de las N tramas de audio; y determinar la primera proporción de energía de acuerdo con la energía de las P_1 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio, donde P_1 es un número entero positivo menor que P; y la determinación, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, de si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual incluye: cuando la primera proporción de energía es mayor que un segundo valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; o cuando la primera proporción de energía es menor que el segundo valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

Con referencia a la cuarta manera de implementación posible del primer aspecto, en una quinta manera de implementación posible del primer aspecto, la energía de una cualquiera de las P_1 envolventes espectrales es mayor que la energía de una cualquiera de las otras envolventes espectrales en las P envolventes espectrales excepto las P_1 envolventes espectrales.

Con referencia a la primera manera de implementación posible del primer aspecto, en una sexta manera de implementación posible del primer aspecto, el parámetro de dispersión general incluye un segundo ancho de banda mínimo y un tercer ancho de banda mínimo; la determinación de un parámetro de dispersión general de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio incluye: determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la segunda proporción de energía preestablecida de las N tramas de audio y determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la tercera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la segunda proporción de energía preestablecida de las N tramas de audio se utilizan como el segundo ancho de banda mínimo, el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la tercera proporción de energía preestablecida de las N tramas de audio se utiliza como el tercer ancho de banda mínimo y, la segunda proporción preestablecida, es menor que la tercera proporción preestablecida; y la determinación, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, de si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual incluye: cuando el segundo ancho de banda mínimo es menor que un tercer valor preestablecido y el tercer ancho de banda mínimo es menor que un cuarto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; cuando el tercer ancho de banda mínimo es menor que un quinto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; o cuando el tercer ancho de banda mínimo es mayor que un sexto valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual, donde el cuarto valor preestablecido es mayor o igual que el tercer valor preestablecido, el quinto valor preestablecido es menor que el cuarto valor preestablecido y el sexto valor preestablecido es mayor que el cuarto valor preestablecido.

Con referencia a la sexta manera de implementación posible del primer aspecto, en una séptima manera de implementación posible del primer aspecto, la determinación de un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la segunda proporción de energía preestablecida de las N tramas de audio y determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la tercera proporción

de energía preestablecida de las N tramas de audio de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio incluye: ordenar la energía de las P envolventes espectrales de cada una de las tramas de audio en orden descendente; determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la segunda proporción preestablecida de cada una de las N tramas de audio; determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la segunda proporción preestablecida de cada una de las N tramas de audio, un valor promedio de anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la segunda proporción preestablecida de las N tramas de audio; determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la tercera proporción preestablecida de cada una de las N tramas de audio; y determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de cada una de las N tramas de audio, un valor promedio de anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la tercera proporción preestablecida de las N tramas de audio.

Con referencia a la primera manera de implementación posible del primer aspecto, en una octava manera de implementación posible del primer aspecto, el parámetro de dispersión general incluye una segunda proporción de energía y una tercera proporción de energía; la determinación de un parámetro de dispersión general de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio incluye: seleccionar P_2 envolventes espectrales de las P envolventes espectrales de cada una de las N tramas de audio; determinar la segunda proporción de energía de acuerdo con la energía de las P_2 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio; seleccionar P_3 envolventes espectrales de las P envolventes espectrales de cada una de las N tramas de audio; y determinar la tercera proporción de energía de acuerdo con la energía de las P_3 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio, donde P_2 y P_3 son números enteros positivos menores que P y P_2 es menor que P_3 ; y la determinación, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, de si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual incluye: cuando la segunda proporción de energía es mayor que un séptimo valor preestablecido y la tercera proporción de energía es mayor que un octavo valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; cuando la segunda proporción de energía es mayor que un noveno valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; o cuando la tercera proporción de energía es menor que un décimo valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

Con referencia a la octava manera de implementación posible del primer aspecto, en una novena manera de implementación posible del primer aspecto, las P_2 envolventes espectrales son P_2 envolventes espectrales que tienen energía máxima en las P envolventes espectrales; y las P_3 envolventes espectrales son P_3 envolventes espectrales que tienen energía máxima en las P envolventes espectrales.

Con referencia al primer aspecto, en una décima manera de implementación posible del primer aspecto, la dispersión de distribución de la energía en el espectro incluye la dispersión global, la dispersión local y la ráfaga a corto plazo de distribución de la energía en el espectro.

Con referencia a la décima manera de implementación posible del primer aspecto, en una decimoprimera manera de implementación posible del primer aspecto, N es 1 y las N tramas de audio son la trama de audio actual; y la determinación de dispersión de distribución, en el espectro, de energía de N tramas de audio de entrada incluye: dividir un espectro de la trama de audio actual en Q subbandas; y determinar un parámetro de dispersión de ráfaga de acuerdo con la energía máxima de cada una de las Q subbandas del espectro de la trama de audio actual, donde el parámetro de dispersión de ráfaga se utiliza para indicar la dispersión global, la dispersión local y la ráfaga a corto plazo de la trama de audio actual.

Con referencia a la decimoprimera manera de implementación posible del primer aspecto, en una decimosegunda manera de implementación posible del primer aspecto, el parámetro de dispersión de ráfaga incluye: una proporción global de máximo a promedio de cada una de las Q subbandas, una proporción local de máximo a promedio de cada una de las Q subbandas y una fluctuación de energía a corto plazo de cada una de las Q subbandas, donde la proporción global de máximo a promedio se determina de acuerdo con la energía máxima en la subbanda y la energía promedio de todas las subbandas de la trama de audio actual, la proporción local de máximo a promedio se determina de acuerdo con la energía máxima en la subbanda y la energía promedio en la subbanda y la fluctuación de energía máxima a corto plazo se determina de acuerdo con la energía máxima en la subbanda y la energía máxima en una banda de frecuencia específica de una trama de audio antes de la trama de audio; y la determinación, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, de si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual incluye: determinar si hay una primera subbanda en las Q subbandas, donde una proporción local de máximo a promedio de la primera subbanda es mayor que un decimoprimer valor preestablecido, una proporción global de máximo a promedio de la primera subbanda es mayor que un decimosegundo valor preestablecido y una fluctuación

de energía máxima a corto plazo de la primera subbanda es mayor que un decimotercer valor preestablecido; y cuando la primera subbanda está en las Q subbandas, determinar utilizar el primer método de codificación para codificar la trama de audio actual.

5 Con referencia al primer aspecto, en una decimotercera manera de implementación posible del primer aspecto, la dispersión de distribución de la energía en el espectro incluye características limitadas por banda de distribución de la energía en el espectro.

10 Con referencia a la decimotercera manera de implementación posible del primer aspecto, en una decimocuarta manera de implementación posible del primer aspecto, la determinación de la dispersión de distribución, en el espectro, de energía de N tramas de audio de entrada incluye: determinar una frecuencia de demarcación de cada una de las N tramas de audio; y determinar un parámetro de dispersión limitada por banda de acuerdo con la frecuencia de demarcación de cada una de las N tramas de audio.

15 Con referencia a la decimocuarta manera de implementación posible del primer aspecto, en una decimoquinta manera de implementación posible del primer aspecto, el parámetro de dispersión limitada por banda es un valor promedio de las frecuencias de demarcación de las N tramas de audio; y la determinación, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, de si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual incluye: cuando se determina que el parámetro de dispersión limitada por banda de las tramas de audio es menor que un decimocuarto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual.

20 De acuerdo con un segundo aspecto, una realización de la presente invención proporciona un aparato, donde el aparato incluye: una unidad de obtención, configurada para obtener N tramas de audio, donde las N tramas de audio incluyen una trama de audio actual y N es un número entero positivo; y una unidad de determinación, configurada para determinar la dispersión de distribución, en el espectro, de energía de las N tramas de audio obtenidas por la unidad de obtención; y la unidad de determinación está además configurada para determinar, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual, donde el primer método de codificación es un método de codificación que está basado en la transformada de frecuencia-tiempo y la cuantificación del coeficiente de transformada y que no está basado en la predicción lineal y, el segundo método de codificación, es un método de codificación basado en predicción lineal.

25 Con referencia al segundo aspecto, en una primera manera de implementación posible del segundo aspecto, la unidad de determinación está configurada específicamente para dividir un espectro de cada una de las N tramas de audio en P envolventes espectrales y determinar un parámetro de dispersión general de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde P es un número entero positivo y el parámetro de dispersión general indica la dispersión de distribución, en el espectro, de la energía de las N tramas de audio.

30 Con referencia a la primera manera de implementación posible del segundo aspecto, en una segunda manera de implementación posible del segundo aspecto, el parámetro de dispersión general incluye un primer ancho de banda mínimo; la unidad de determinación está configurada específicamente para determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio es el primer ancho de banda mínimo; y la unidad de determinación está configurada específicamente para: cuando el primer ancho de banda mínimo es menor que un primer valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; y cuando el primer ancho de banda mínimo es mayor que el primer valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

35 Con referencia a la segunda manera de implementación posible del segundo aspecto, en una tercera manera de implementación posible del segundo aspecto, la unidad de determinación está configurada específicamente para: ordenar la energía de las P envolventes espectrales de cada una de las tramas de audio en orden descendente; determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio; y determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio, un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de las N tramas de audio.

40 Con referencia a la primera manera de implementación posible del segundo aspecto, en una cuarta manera de implementación posible del segundo aspecto, el parámetro de dispersión general incluye una primera proporción de energía; la unidad de determinación está configurada específicamente para seleccionar P_1 envolventes espectrales

de las P envolventes espectrales de cada una de las N tramas de audio y determinar la primera proporción de energía de acuerdo con la energía de las P_1 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio, donde P_1 es un número entero positivo menor que P; y la unidad de determinación está configurada específicamente para: cuando la primera proporción de energía es mayor que un
 5 segundo valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; y cuando la primera proporción de energía es menor que el segundo valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

Con referencia a la cuarta manera de implementación posible del segundo aspecto, en una quinta manera de implementación posible del segundo aspecto, la unidad de determinación está configurada específicamente para
 10 determinar las P_1 envolventes espectrales de acuerdo con la energía de las P envolventes espectrales, donde la energía de una cualquiera de las P_1 envolventes espectrales es mayor que la energía de una cualquiera de las otras envolventes espectrales en las P envolventes espectrales excepto las P_1 envolventes espectrales.

Con referencia a la primera manera de implementación posible del segundo aspecto, en una sexta manera de implementación posible del segundo aspecto, el parámetro de dispersión general incluye un segundo ancho de
 15 banda mínimo y un tercer ancho de banda mínimo; la unidad de determinación está configurada específicamente para determinar un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la segunda proporción de energía preestablecida de las N tramas de audio y determinar un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la tercera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde
 20 el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la segunda proporción de energía preestablecida de las N tramas de audio se utiliza como el segundo ancho de banda mínimo, el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la tercera proporción de energía preestablecida de las N tramas de audio se utiliza como el tercer ancho de banda mínimo y, la segunda proporción preestablecida, es menor que la tercera proporción preestablecida; y la unidad de determinación está configurada específicamente para: cuando el segundo ancho de banda mínimo es menor que un tercer valor preestablecido y el
 25 tercer ancho de banda mínimo es menor que un cuarto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; cuando el tercer ancho de banda mínimo es menor que un quinto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; y cuando el tercer ancho de banda mínimo es mayor que un sexto valor preestablecido, determinar utilizar el
 30 segundo método de codificación para codificar la trama de audio actual, donde el cuarto valor preestablecido es mayor o igual que el tercer valor preestablecido, el quinto valor preestablecido es menor que el cuarto valor preestablecido y el sexto valor preestablecido es mayor que el cuarto valor preestablecido.

Con referencia a la sexta manera de implementación posible del segundo aspecto, en una séptima manera de implementación posible del segundo aspecto, la unidad de determinación está configurada específicamente para:
 35 ordenar la energía de las P envolventes espectrales de cada una de las tramas de audio en orden descendente; determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la segunda proporción preestablecida de cada una de las N tramas de audio; determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la
 40 segunda proporción preestablecida de cada una de las N tramas de audio, un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la segunda proporción preestablecida de las N tramas de audio; determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la tercera proporción preestablecida de cada una de las N
 45 tramas de audio; y determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de cada una de las N tramas de audio, un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la tercera proporción preestablecida de las N tramas de audio.

Con referencia a la primera manera de implementación posible del segundo aspecto, en una octava manera de implementación posible del segundo aspecto, el parámetro de dispersión general incluye una segunda proporción de
 50 energía y una tercera proporción de energía; la unidad de determinación está configurada específicamente para: seleccionar P_2 envolventes espectrales de las P envolventes espectrales de cada una de las N tramas de audio, determinar la segunda proporción de energía de acuerdo con la energía de las P_2 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio, seleccionar P_3 envolventes
 55 espectrales de las P envolventes espectrales de cada una de las N tramas de audio y determinar la tercera proporción de energía de acuerdo con la energía de las P_3 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio, donde P_2 y P_3 son números enteros positivos menores que P y P_2 es menor que P_3 ; y la unidad de determinación está configurada específicamente para: cuando la segunda proporción de energía es mayor que un séptimo valor preestablecido y la tercera proporción de energía es mayor que un octavo valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; cuando la segunda proporción de energía es mayor que un noveno valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; y cuando la tercera proporción de
 60 energía es mayor que un décimo valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

energía es menor que un décimo valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

5 Con referencia a la octava manera de implementación posible del segundo aspecto, en una novena manera de implementación posible del segundo aspecto, la unidad de determinación está configurada específicamente para determinar, a partir de las P envolventes espectrales de cada una de las N tramas de audio, P_2 envolventes espectrales que tienen energía máxima y determinar, a partir de las P envolventes espectrales de cada una de las N tramas de audio, P_3 envolventes espectrales que tienen energía máxima.

10 Con referencia al segundo aspecto, en una décima manera de implementación posible del segundo aspecto, N es 1 y las N tramas de audio son la trama de audio actual; y la unidad de determinación está configurada específicamente para dividir un espectro de la trama de audio actual en Q subbandas y determinar un parámetro de dispersión de ráfaga de acuerdo con la energía máxima de cada una de las Q subbandas del espectro de la trama de audio actual, donde el parámetro de dispersión de ráfaga se utiliza para indicar la dispersión global, la dispersión local y la ráfaga a corto plazo de la trama de audio actual.

15 Con referencia a la décima manera de implementación posible del segundo aspecto, en una decimoprimer manera de implementación posible del segundo aspecto, la unidad de determinación está configurada específicamente para determinar una proporción global de máximo a promedio de cada una de las Q subbandas, una proporción local de máximo a promedio de cada una de las Q subbandas y una fluctuación de energía a corto plazo de cada una de las Q subbandas, donde la proporción global de máximo a promedio se determina por la unidad de determinación de acuerdo con la energía máxima en la subbanda y la energía promedio de todas las subbandas de la trama de audio actual, la proporción local de máximo a promedio se determina por la unidad de determinación de acuerdo con la energía máxima en la subbanda y la energía promedio en la subbanda y, la fluctuación de energía máxima a corto plazo, se determina de acuerdo con la energía máxima en la subbanda y la energía máxima en una banda de frecuencia específica de una trama de audio antes de la trama de audio; y la unidad de determinación está configurada específicamente para: determinar si hay una primera subbanda en las Q subbandas, donde una proporción local de máximo a promedio de la primera subbanda es mayor que un decimoprimer valor preestablecido, una proporción global de máximo a promedio de la primera subbanda es mayor que un decimosegundo valor preestablecido y una fluctuación de energía máxima a corto plazo de la primera subbanda es mayor que un decimotercer valor preestablecido; y cuando la primera subbanda está en las Q subbandas, determinar utilizar el primer método de codificación para codificar la trama de audio actual.

30 Con referencia al segundo aspecto, en una decimosegunda manera de implementación posible del segundo aspecto, la unidad de determinación está configurada específicamente para determinar una frecuencia de demarcación de cada una de las N tramas de audio; y la unidad de determinación está configurada específicamente para determinar un parámetro de dispersión limitada por banda de acuerdo con la frecuencia de demarcación de cada una de las N tramas de audio.

35 Con referencia a la decimosegunda manera de implementación posible del segundo aspecto, en una decimotercera manera de implementación posible del segundo aspecto, el parámetro de dispersión limitada por banda es un valor promedio de las frecuencias de demarcación de las N tramas de audio; y la unidad de determinación está configurada específicamente para: cuando se determina que el parámetro de dispersión limitada por banda de las tramas de audio es menor que un decimocuarto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual.

De acuerdo con las soluciones técnicas anteriores, cuando se codifica una trama de audio, se considera la dispersión de distribución, en un espectro, de energía de la trama de audio, lo cual puede reducir complejidad de codificación y asegurar que la codificación es de precisión relativamente alta.

Breve descripción de los dibujos

45 Para describir las soluciones técnicas en las realizaciones de la presente invención con mayor claridad, en lo siguiente, se introducen brevemente los dibujos adjuntos que describen de realizaciones de la presente invención. Aparentemente, los dibujos que adjuntos en la siguiente descripción muestran simplemente algunas realizaciones de la presente invención.

50 La FIG. 1 es un diagrama de flujo esquemático de un método de codificación de audio de acuerdo con una realización de la presente invención;

la FIG. 2 es un diagrama de bloques estructural de un aparato de acuerdo con una realización de la presente invención; y

la FIG. 3 es un diagrama de bloques estructural de un aparato de acuerdo con una realización de la presente invención.

Descripción de las realizaciones

Lo siguiente describe clara y completamente las soluciones técnicas en las realizaciones de la presente invención con referencia a los dibujos adjuntos que muestran realizaciones de la presente invención. Aparentemente, las realizaciones descritas son simplemente una parte en lugar de todas las realizaciones de la presente invención.

5 La FIG. 1 es un diagrama de flujo esquemático de un método de codificación de audio de acuerdo con una realización de la presente invención.

101: determinar la dispersión de distribución, en un espectro, de energía de las N tramas de audio de entrada, donde las N tramas de audio incluyen una trama de audio actual y N es un número entero positivo.

10 102: determinar, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual, donde el primer método de codificación es un método de codificación que está basado en la transformada de tiempo-frecuencia y la cuantificación del coeficiente de transformada y que no está basado en predicción lineal y, el segundo método de codificación, es un método de codificación basado en predicción lineal.

15 De acuerdo con el método mostrado en la FIG. 1, cuando se codifica una trama de audio, se considera la dispersión de distribución, en un espectro, de energía de la trama de audio, lo cual puede reducir la complejidad de codificación y asegurar que la codificación sea de una precisión relativamente alta.

20 Durante la selección de un método de codificación apropiado para una trama de audio, se puede considerar la dispersión de distribución, en un espectro, de energía de la trama de audio. Puede haber tres tipos de dispersión de distribución, en un espectro, de energía de una trama de audio: dispersión general, dispersión de ráfaga y dispersión limitada por banda.

25 Opcionalmente, en una realización, se puede seleccionar un método de codificación apropiado para la trama de audio actual utilizando la dispersión general. En este caso, la dispersión de distribución, en el espectro, de energía de N tramas de audio de entrada incluye: dividir un espectro de cada una de las N tramas de audio en P envolventes espectrales, donde P es un número entero positivo; y determinar un parámetro de dispersión general de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde el parámetro de dispersión general indica la dispersión de distribución, en el espectro, de la energía de las N tramas de audio.

30 Específicamente, un valor promedio de anchos de banda mínimos de distribución, en un espectro, de energía de proporción específica de N tramas de audio consecutivas de entrada puede definirse como la dispersión general. Un ancho de banda menor indica una dispersión general más fuerte y un ancho de banda mayor indica una dispersión general más débil. En otras palabras, una dispersión general más fuerte indica que la energía de una trama de audio está más centralizada y una dispersión general más débil indica que la energía de una trama de audio está más dispersa. La eficiencia es alta cuando se utiliza el primer método de codificación para codificar una trama de audio cuya dispersión general es relativamente fuerte. Por lo tanto, se puede seleccionar un método de codificación apropiado determinando la dispersión general de una trama de audio, para codificar la trama de audio. Para ayudar a determinar la dispersión general de una trama de audio, la dispersión general se puede cuantificar para obtener un parámetro de dispersión general. Opcionalmente, cuando N es 1, la dispersión general es un ancho de banda mínimo de distribución, en un espectro, de energía de proporción específica de la trama de audio actual.

35 Opcionalmente, en una realización, el parámetro de dispersión general incluye un primer ancho de banda mínimo. En este caso, la determinación de un parámetro de dispersión general de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio incluye: determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio es el primer ancho de banda mínimo. La determinación, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, de si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual incluye: cuando el primer ancho de banda mínimo es menor que un primer valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; o cuando el primer ancho de banda mínimo es mayor que el primer valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual. Opcionalmente, en una realización, cuando N es 1, las N tramas de audio son la trama de audio actual y el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio es un ancho de banda mínimo de distribución, en el espectro, de la primera proporción de energía preestablecida de la trama de audio actual.

55 Una persona experta en la técnica puede entender que, el primer valor preestablecido y la primera proporción preestablecida se pueden determinar de acuerdo con un experimento de simulación. Un primer valor preestablecido apropiado y una primera proporción preestablecida se pueden determinar por medio de un experimento de

simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación o el segundo método de codificación. Generalmente, un valor de la primera proporción preestablecida es generalmente un número entre 0 y 1 y relativamente cercano a 1, por ejemplo, 90 % u 80 %. La selección del primer valor preestablecido se relaciona con el valor de la primera proporción preestablecida y, también, se relaciona con una tendencia de selección entre el primer método de codificación y el segundo método de codificación. Por ejemplo, un primer valor preestablecido correspondiente a una primera proporción preestablecida relativamente grande, es generalmente mayor que un primer valor preestablecido correspondiente a una primera proporción preestablecida relativamente pequeña. Para otro ejemplo, un primer valor preestablecido correspondiente a una tendencia para seleccionar el primer método de codificación es generalmente mayor que un primer valor preestablecido correspondiente a una tendencia para seleccionar el segundo método de codificación.

La determinación de un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio incluye: ordenar la energía de las P envolventes espectrales de cada una de las tramas de audio en orden descendente; determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio; y determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio, un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de las N tramas de audio. Por ejemplo, una señal de audio de entrada es una señal de banda ancha muestreada a 16 kHz y la señal de entrada se ingresa en una trama de 20 ms. Cada una de las tramas de señal es de 320 puntos de muestreo en el dominio del tiempo. La transformada de tiempo-frecuencia se realiza en una señal de dominio del tiempo. Por ejemplo, la transformada de tiempo-frecuencia se realiza por medio de la transformada rápida de Fourier (Fast Fourier Transformation, FFT), para obtener 160 S(k) envolventes espectrales, es decir, 160 coeficientes de espectro de energía de FFT, donde $k = 0, 1, 2, \dots, 159$. Se encuentra un ancho de banda mínimo a partir de las S(k) envolventes espectrales de una manera que una proporción que la energía en el ancho de banda representa en la energía total de la trama es la primera proporción preestablecida. Específicamente, determinar un ancho de banda mínimo de distribución, en un espectro, de la primera proporción de energía preestablecida de una trama de audio de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de la trama de audio incluye: acumular secuencialmente la energía de los intervalos de frecuencia en las S(k) envolventes espectrales en orden descendente; y comparar la energía obtenida después de cada acumulación con la energía total de la trama de audio y, si una proporción es mayor que la primera proporción preestablecida, finalizar el proceso de acumulación, donde una cantidad de veces de acumulación es el ancho de banda mínimo. Por ejemplo, la primera proporción preestablecida es del 90% y, si una proporción que una suma de energía obtenida después de 30 veces de acumulación representa en la energía total supera el 90 %, una proporción que una suma de energía obtenida después de 29 veces de acumulación representa en la energía total es menor que el 90 % y una proporción que una suma de energía obtenida después de 31 veces de acumulación representa en la energía total excede la proporción que la suma de energía obtenida después de 30 veces de acumulación representa en la energía total, se puede considerar que un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de la trama de audio es 30. El proceso de determinación del ancho de banda mínimo anterior se ejecuta para cada una de las N tramas de audio, para determinar por separado los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la primera proporción preestablecida de las N tramas de audio, incluyendo la trama de audio actual, y calcular el valor promedio de los N anchos de banda mínimos. El valor promedio de los N anchos de banda mínimos se puede referir como el primer ancho de banda mínimo y, el primer ancho de banda mínimo, se puede utilizar como el parámetro de dispersión general. Cuando el primer ancho de banda mínimo es menor que el primer valor preestablecido, se determina utilizar el primer método de codificación para codificar la trama de audio actual.

Opcionalmente, en otra realización, el parámetro de dispersión general puede incluir una primera proporción de energía. En este caso, la determinación de un parámetro de dispersión general de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio incluye: seleccionar P_1 envolventes espectrales de las P envolventes espectrales de cada una de las N tramas de audio; y determinar la primera proporción de energía de acuerdo con la energía de las P_1 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio, donde P_1 es un número entero positivo inferior a P. La determinación, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, de si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual, incluye : cuando la primera proporción de energía es mayor que un segundo valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; o cuando la primera proporción de energía es menor que el segundo valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual. Opcionalmente, en una realización, cuando N es 1, las N tramas de audio son la trama de audio actual, y la determinación de la primera proporción de energía de acuerdo con la energía de P_1 Las envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio incluye: determinar la

primera proporción de energía de acuerdo con la energía de las P_1 envolventes espectrales de la trama de audio actual y la energía total de la trama de audio actual.

Específicamente, la primera proporción de energía se puede calcular mediante la siguiente fórmula:

$$\left\{ \begin{array}{l} R_1 = \frac{\sum_{n=1}^N r(n)}{N} \\ r(n) = \frac{E_{p1}(n)}{E_{todo}(n)} \end{array} \right. \quad \text{Fórmula 1.1}$$

5 donde R_1 representa la primera proporción de energía, $E_{p1}(n)$ representa una suma de energía de las P_1 envolventes espectrales seleccionadas en una $n^{\text{ésima}}$ trama de audio, $E_{todo}(n)$ representa la energía total de la $n^{\text{ésima}}$ trama de audio y $r(n)$ representa una proporción que la energía de las P_1 envolventes espectrales de la $n^{\text{ésima}}$ trama de audio en las N tramas de audio representa en la energía total de la trama de audio.

10 Una persona experta en la técnica puede entender que, el segundo valor preestablecido y la selección de las P_1 envolventes espectrales se pueden determinar de acuerdo con un experimento de simulación. Un segundo valor preestablecido apropiado, un valor apropiado de P_1 y un método apropiado para seleccionar las P_1 envolventes espectrales, se pueden determinar por medio de un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando una trama de audio que cumple la condición anterior se codifica utilizando el primer método de codificación o el segundo método de codificación. En general, el valor de P_1 puede ser un número relativamente pequeño. Por ejemplo, P_1 se selecciona de manera que una proporción de P_1 a P es inferior al 20 %. Para el segundo valor preestablecido, generalmente no se selecciona un número correspondiente a una proporción excesivamente pequeña. Por ejemplo, no se selecciona un número inferior al 10 %. La selección del segundo valor preestablecido está relacionada con el valor de P_1 y una tendencia de selección entre el primer método de codificación y el segundo método de codificación. Por ejemplo, un segundo valor preestablecido que corresponde a un P_1 relativamente grande es generalmente mayor que un segundo valor preestablecido que corresponde a un P_1 relativamente pequeño. Para otro ejemplo, un segundo valor preestablecido correspondiente a una tendencia para seleccionar el primer método de codificación es generalmente menor que un segundo valor preestablecido correspondiente a una tendencia para seleccionar el segundo método de codificación. Opcionalmente, en una realización, la energía de una cualquiera de las P_1 envolventes espectrales es mayor que la energía de una cualquiera de las restantes ($P - P_1$) envolventes espectrales en las P envolventes espectrales.

Por ejemplo, una señal de audio de entrada es una señal de banda ancha muestreada a 16 kHz y la señal de entrada se ingresa en una trama de 20 ms. Cada una de las tramas de señal es de 320 puntos de muestreo en el dominio del tiempo. La transformada de tiempo-frecuencia se realiza en una señal de dominio del tiempo. Por ejemplo, la transformada de tiempo-frecuencia se realiza por medio de la transformada rápida de Fourier, para obtener 160 $S(k)$ envolventes espectrales, donde $k = 0, 1, 2, \dots, 159$. Se seleccionan P_1 envolventes espectrales de las 160 envolventes espectrales y se calcula una proporción que una suma de energía de las P_1 envolventes espectrales representa en la energía total de la trama de audio. El proceso anterior se ejecuta para cada una de las N tramas de audio. Es decir, se calcula una proporción que una suma de energía de las P_1 envolventes espectrales de cada una de las N tramas de audio representa en la respectiva energía total. Se calcula un valor promedio de las proporciones. El valor promedio de las proporciones es la primera proporción de energía. Cuando la primera proporción de energía es mayor que el segundo valor preestablecido, se determina utilizar el primer método de codificación para codificar la trama de audio actual. Cuando la primera proporción de energía es menor que el segundo valor preestablecido, se determina utilizar el segundo método de codificación para codificar la trama de audio actual. La energía de una cualquiera de las P_1 envolventes espectrales es mayor que la energía de una cualquiera de las otras envolventes espectrales en las P envolventes espectrales excepto las P_1 envolventes espectrales. Opcionalmente, en una realización, el valor de P_1 puede ser 20.

Opcionalmente, en otra realización, el parámetro de dispersión general puede incluir un segundo ancho de banda mínimo y un tercer ancho de banda mínimo. En este caso, la determinación de un parámetro de dispersión general de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio incluye: determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la segunda proporción de energía preestablecida de N tramas de audio y determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la tercera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la segunda proporción de energía preestablecida de las N tramas de audio se utiliza como el segundo ancho de banda mínimo, el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la tercera proporción de energía preestablecida de las N tramas de audio se utiliza como el tercer ancho de banda mínimo y la segunda proporción preestablecida es menor que la tercera

proporción preestablecida. La determinación, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, de si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual incluye: cuando el segundo ancho de banda mínimo es menor que un tercer valor preestablecido y el tercer ancho de banda mínimo es menor que un cuarto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; cuando el tercer ancho de banda mínimo es menor que un quinto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; o cuando el tercer ancho de banda mínimo es mayor que un sexto valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual. El cuarto valor preestablecido es mayor o igual que el tercer valor preestablecido, el quinto valor preestablecido es menor que el cuarto valor preestablecido y el sexto valor preestablecido es mayor que el cuarto valor preestablecido. Opcionalmente, en una realización, cuando N es 1, las N tramas de audio son la trama de audio actual. La determinación de un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la segunda proporción de energía preestablecida de las N tramas de audio como el segundo ancho de banda mínimo incluye: determinar un ancho de banda mínimo de distribución, en el espectro, de la segunda proporción de energía preestablecida de la trama de audio actual como el segundo ancho de banda mínimo. La determinación de un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la tercera proporción de energía preestablecida de las N tramas de audio como el tercer ancho de banda mínimo incluye: determinar un ancho de banda mínimo de distribución, en el espectro, de la tercera proporción de energía preestablecida de la trama de audio actual como el tercer ancho de banda mínimo.

Una persona experta en la técnica puede entender que, el tercer valor preestablecido, el cuarto valor preestablecido, el quinto valor preestablecido, el sexto valor preestablecido, la segunda proporción preestablecida y la tercera proporción preestablecida se pueden determinar de acuerdo con un experimento de simulación. Los valores preestablecidos apropiados y las proporciones preestablecidas pueden determinarse por medio de un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación o el segundo método de codificación.

La determinación de un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la segunda proporción de energía preestablecida de las N tramas de audio y determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la tercera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio incluye: ordenar la energía de las P envolventes espectrales de cada una de las tramas de audio en orden descendente; determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la segunda proporción preestablecida de cada una de las N tramas de audio; determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la segunda proporción preestablecida de cada una de las N tramas de audio, un valor promedio de anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la segunda proporción preestablecida de las N tramas de audio; determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la tercera proporción preestablecida de cada una de las N tramas de audio; y determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de cada una de las N tramas de audio, un valor promedio de anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la tercera proporción preestablecida de las N tramas de audio. Por ejemplo, una señal de audio de entrada es una señal de banda ancha muestreada a 16 kHz y la señal de entrada se ingresa en una trama de 20 ms. Cada una de las tramas de señal es de 320 puntos de muestreo en el dominio del tiempo. La transformada de tiempo-frecuencia se realiza en una señal de dominio del tiempo. Por ejemplo, la transformada de tiempo-frecuencia se realiza por medio de la transformada rápida de Fourier, para obtener 160 $S(k)$ envolventes espectrales, donde $k = 0, 1, 2, \dots, 159$. Se encuentra un ancho de banda mínimo de las $S(k)$ envolventes espectrales de una manera que una proporción que la energía en el ancho de banda representa en la energía total de la trama es la segunda proporción preestablecida. Se sigue encontrando un ancho de banda de las $S(k)$ envolventes espectrales de manera que una proporción que la energía en el ancho de banda representa en la energía total es la tercera proporción preestablecida. Específicamente, determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de la trama de audio, un ancho de banda mínimo de distribución, en un espectro, de energía que representa no menos que la segunda proporción preestablecida de una trama de audio y un ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de la trama de audio incluye: acumular secuencialmente la energía de los intervalos de frecuencia en las $S(k)$ envolventes espectrales en orden descendente. La energía obtenida después de cada acumulación se compara con la energía total de la trama de audio y, si una proporción es mayor que la segunda proporción preestablecida, una cantidad de veces de acumulación es un ancho de banda mínimo que no es menor que la segunda proporción preestablecida. La acumulación continúa y, si una proporción de energía obtenida después de la acumulación con la energía total de la trama de audio, es mayor que la tercera proporción preestablecida, la acumulación finaliza y una cantidad de veces de acumulación es un ancho de banda mínimo que cumple no ser menor que la tercera proporción preestablecida. Por ejemplo, la segunda proporción preestablecida es del 85 % y la tercera proporción preestablecida es del 95 %. Si una proporción que una suma de energía obtenida después de 30

veces de acumulación representa en la energía total supera el 85 %, se puede considerar que el ancho de banda mínimo de distribución, en el espectro, de la segunda proporción de energía preestablecida de la trama de audio es 30. La acumulación continúa y, si una proporción que una suma de energía obtenida después de 35 veces de acumulación representa en la energía total es del 95 %, se puede considerar que el ancho de banda mínimo de distribución, en el espectro, de la tercera proporción de energía preestablecida de la trama de audio es 35. El proceso anterior se ejecuta para cada una de las N tramas de audio, para determinar por separado los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la segunda proporción preestablecida de las N tramas de audio, incluyendo la trama de audio actual y los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de las N tramas de audio incluyendo la trama de audio actual. El valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la segunda proporción preestablecida de las N tramas de audio, es el segundo ancho de banda mínimo. El valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de las N tramas de audio, es el tercer ancho de banda mínimo. Cuando el segundo ancho de banda mínimo es menor que el tercer valor preestablecido y el tercer ancho de banda mínimo es menor que el cuarto valor preestablecido, se determina utilizar el primer método de codificación para codificar la trama de audio actual. Cuando el tercer ancho de banda mínimo es menor que el quinto valor preestablecido, se determina utilizar el primer método de codificación para codificar la trama de audio actual. Cuando el tercer ancho de banda mínimo es mayor que el sexto valor preestablecido, se determina utilizar el segundo método de codificación para codificar la trama de audio actual.

Opcionalmente, en otra realización, el parámetro de dispersión general incluye una segunda proporción de energía y una tercera proporción de energía. En este caso, la determinación de un parámetro de dispersión general de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio incluye: seleccionar P_2 envolventes espectrales de las P envolventes espectrales de cada una de las N tramas de audio; determinar la segunda proporción de energía de acuerdo con la energía de las P_2 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio; seleccionar P_3 envolventes espectrales de las P envolventes espectrales de cada una de las N tramas de audio; y determinar la tercera proporción de energía de acuerdo con la energía de las P_3 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio. La determinación, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, de si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual incluye: cuando la segunda proporción de energía es mayor que un séptimo valor preestablecido y la tercera proporción de energía es mayor que un octavo valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; cuando la segunda proporción de energía es mayor que un noveno valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; o cuando la tercera proporción de energía es menor que un décimo valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual. P_2 y P_3 son números enteros positivos de menos de P y P_2 es menor que P_3 . Opcionalmente, en una realización, cuando N es 1, las N tramas de audio son la trama de audio actual. La determinación de la segunda proporción de energía de acuerdo con la energía de las P_2 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio incluye: determinar la segunda proporción de energía de acuerdo con la energía de las P_2 envolventes espectrales de la trama de audio actual y la energía total de la trama de audio actual. La determinación de la tercera proporción de energía de acuerdo con la energía de las P_3 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio incluye: determinar la tercera proporción de energía de acuerdo con la energía de las P_3 envolventes espectrales de la trama de audio actual y la energía total de la trama de audio actual.

Una persona experta en la técnica puede entender que, los valores de P_2 y P_3 , el séptimo valor preestablecido, el octavo valor preestablecido, el noveno valor preestablecido y el décimo valor preestablecido se pueden determinar de acuerdo con un experimento de simulación. Los valores preestablecidos apropiados pueden determinarse por medio de un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación o el segundo método de codificación. Opcionalmente, en una realización, las P_2 envolventes espectrales pueden ser P_2 envolventes espectrales que tienen energía máxima en las P envolventes espectrales; y las P_3 envolventes espectrales pueden ser P_3 envolventes espectrales que tienen energía máxima en las P envolventes espectrales.

Por ejemplo, una señal de audio de entrada es una señal de banda ancha muestreada a 16 kHz y la señal de entrada se ingresa en una trama de 20 ms. Cada una de las tramas de señal es de 320 puntos de muestreo en el dominio del tiempo. La transformada de tiempo-frecuencia se realiza en una señal de dominio del tiempo. Por ejemplo, la transformada de tiempo-frecuencia se realiza por medio de la transformada rápida de Fourier, para obtener 160 S(k) envolventes espectrales, donde $k = 0, 1, 2, \dots, 159$. Se seleccionan P_2 envolventes espectrales de las 160 envolventes espectrales y se calcula una proporción que una suma de energía de las P_2 envolventes espectrales representa en la energía total de la trama de audio. El proceso anterior se ejecuta para cada una de las N tramas de audio. Es decir, se calcula una proporción que una suma de energía de las P_2 envolventes espectrales de cada una de las N tramas de audio representa en la respectiva energía total. Se calcula un valor promedio de las proporciones. El valor promedio de las proporciones es la segunda proporción de energía. Las P_3 envolventes espectrales se seleccionan de las 160 envolventes espectrales y se calcula una proporción que una suma de

energía de las P_3 envolventes espectrales representa en la energía total de la trama de audio. El proceso anterior se ejecuta para cada una de las N tramas de audio. Es decir, se calcula una proporción que una suma de energía de las P_3 envolventes espectrales de cada una de las N tramas de audio representa en la respectiva energía total. Se calcula un valor promedio de las proporciones. El valor promedio de las proporciones es la tercera proporción de energía. Cuando la segunda proporción de energía es mayor que el séptimo valor preestablecido y la tercera proporción de energía es mayor que el octavo valor preestablecido, se determina utilizar el primer método de codificación para codificar la trama de audio actual. Cuando la segunda proporción de energía es mayor que el noveno valor preestablecido, se determina utilizar el primer método de codificación para codificar la trama de audio actual. Cuando la tercera proporción de energía es menor que el décimo valor preestablecido, se determina utilizar el segundo método de codificación para codificar la trama de audio actual. Las P_2 envolventes espectrales pueden ser P_2 envolventes espectrales que tienen energía máxima en las P envolventes espectrales; y las P_3 envolventes espectrales pueden ser P_3 envolventes espectrales que tienen energía máxima en las P envolventes espectrales. Opcionalmente, en una realización, el valor de P_2 puede ser 20 y el valor de P_3 puede ser 30.

Opcionalmente, en otra realización, se puede seleccionar un método de codificación apropiado para la trama de audio actual utilizando la dispersión de ráfaga. Para la dispersión de ráfaga, se debe considerar la dispersión global, la dispersión local y la ráfaga a corto plazo de distribución, en un espectro, de energía de una trama de audio. En este caso, la dispersión de distribución de la energía en el espectro puede incluir la dispersión global, la dispersión local y la ráfaga a corto plazo de distribución de la energía en el espectro. En este caso, un valor de N puede ser 1 y las N tramas de audio son la trama de audio actual. La dispersión de distribución, en un espectro, de energía de N tramas de audio de entrada incluye: dividir un espectro de la trama de audio actual en Q subbandas; y determinar un parámetro de dispersión de ráfaga de acuerdo con la energía máxima de cada una de las Q subbandas del espectro de la trama de audio actual, donde el parámetro de dispersión de ráfaga se utiliza para indicar la dispersión global, la dispersión local y la ráfaga a corto plazo de la trama de audio actual. El parámetro de dispersión de ráfaga incluye: una proporción global de máximo a promedio de cada una de las Q subbandas, una proporción local de máximo a promedio de cada una de las Q subbandas y una fluctuación de energía a corto plazo de cada una de las Q subbandas, donde la proporción global de máximo a promedio se determina de acuerdo con la energía máxima en la subbanda y la energía promedio de todas las subbandas de la trama de audio actual, la proporción local máximo a promedio se determina de acuerdo con la energía máxima en la subbanda y la energía promedio en la subbanda y, la fluctuación de energía máxima a corto plazo, se determina de acuerdo con la energía máxima en la subbanda y la energía máxima en una banda de frecuencia específica de una trama de audio antes de la trama de audio. La determinación, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, de si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual incluye: determinar si hay una primera subbanda en las Q subbandas, donde una proporción local de máximo a promedio de la primera subbanda es mayor que un decimoprimer valor preestablecido, una proporción global de máximo a promedio de la primera subbanda es mayor que un decimosegundo valor preestablecido y una fluctuación de energía máxima a corto plazo de la primera subbanda es mayor que un decimotercer valor preestablecido; y cuando primera subbanda está en las Q subbandas, determinar utilizar el primer método de codificación para codificar la trama de audio actual. La proporción global de máximo a promedio de cada una de las Q subbandas, la proporción local de máximo a promedio de cada una de las Q subbandas y la fluctuación de energía a corto plazo de cada una de las Q subbandas, respectivamente, representan la dispersión global, la dispersión local y la ráfaga a corto plazo.

Específicamente, la proporción global de máximo a promedio se puede determinar utilizando la siguiente fórmula:

$$p2s(i) = e(i) / \left(\frac{1}{P} * \sum_{k=0}^{P-1} s(k) \right) \quad \text{Fórmula 1.2}$$

donde $e(i)$ representa la energía máxima de una $i^{\text{ésima}}$ subbanda en las Q subbandas, $s(k)$ representa la energía de una $k^{\text{ésima}}$ envolvente espectral en las P envolventes espectrales y $p2s(i)$ representa una proporción global de máximo a promedio de la $i^{\text{ésima}}$ subbanda.

La proporción local de máximo a promedio se puede determinar utilizando la siguiente fórmula:

$$p2a(i) = e(i) / \left(\frac{1}{h(i) - l(i) + 1} * \sum_{k=l(i)}^{h(i)} s(k) \right) \quad \text{Fórmula 1.3}$$

donde $e(i)$ representa la energía máxima de la $i^{\text{ésima}}$ subbanda en las Q subbandas, $s(k)$ representa la energía de la $k^{\text{ésima}}$ envolvente espectral en las P envolventes espectrales, $h(i)$ representa un índice de una envolvente espectral que está incluida en la $i^{\text{ésima}}$ subbanda y que tiene una frecuencia más alta, $l(i)$ representa un índice de una

envolvente espectral que está incluida en la $i^{\text{ésima}}$ subbanda y que tiene una frecuencia más baja, $p2a(i)$ representa una proporción local de máximo a promedio de la $i^{\text{ésima}}$ subbanda y $h(i)$ es menor o igual que $P - 1$.

La fluctuación de energía máxima a corto plazo se puede determinar utilizando la siguiente fórmula:

$$\text{dev}(i) = (2 * e(i)) / (e_1 + e_2) \quad \text{Fórmula 1.4}$$

5 donde $e(i)$ representa la energía máxima de la $i^{\text{ésima}}$ subbanda en las Q subbandas de la trama de audio actual y e_1 y e_2 representan la energía máxima de bandas de frecuencia específicas de tramas de audio antes de la trama de audio actual. Específicamente, suponiendo que la trama de audio actual es una $M^{\text{ésima}}$ trama de audio, se determina una envolvente espectral en la que se encuentra la energía máxima de la $i^{\text{ésima}}$ subbanda de la trama de audio actual. Se supone que la envolvente espectral en la que se encuentra la energía máxima es i_1 . Se determina la energía máxima dentro de un rango de una $(i_1 - t)^{\text{ésima}}$ envolvente espectral a una $(i_1 + t)^{\text{ésima}}$ envolvente espectral en una $(M - 1)^{\text{ésima}}$ trama de audio y la energía máxima es e_1 . Del mismo modo, se determina la energía máxima dentro de un rango de una $(i_1 - t)^{\text{ésima}}$ envolvente espectral a una $(i_1 + t)^{\text{ésima}}$ envolvente espectral en una $(M - 2)^{\text{ésima}}$ trama de audio y la energía máxima es e_2 .

15 Una persona experta en la técnica puede entender que, el decimoprimer valor preestablecido, el decimosegundo valor preestablecido y el decimotercera valor preestablecido se pueden determinar de acuerdo con un experimento de simulación. Los valores preestablecidos apropiados pueden determinarse por medio de un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación.

20 Opcionalmente, en otra realización, se puede seleccionar un método de codificación apropiado para la trama de audio actual utilizando la dispersión limitada por banda. En este caso, la dispersión de distribución de la energía en el espectro incluye la dispersión limitada por banda de distribución de la energía en el espectro. En este caso, la determinación de la dispersión de distribución, en el espectro, de energía de N tramas de audio de entrada incluye: determinar una frecuencia de demarcación de cada una de las N tramas de audio; y determinar un parámetro de dispersión limitada por banda de acuerdo con la frecuencia de demarcación de cada una de las N tramas de audio. El parámetro de dispersión limitada por banda puede ser un valor promedio de las frecuencias de demarcación de las N tramas de audio. Por ejemplo, una $N_i^{\text{ésima}}$ trama de audio es una cualquiera de las N tramas de audio y un rango de frecuencia de la $N_i^{\text{ésima}}$ trama de audio es de F_b a F_e , donde F_b es menor que F_e . Suponiendo que F_b es una frecuencia de inicio, un método para determinar una frecuencia de demarcación de la $N_i^{\text{ésima}}$ trama de audio puede ser la búsqueda de una frecuencia F_s partiendo de F_b , donde F_s cumple las siguientes condiciones: una proporción de una suma de energía de F_b a F_s con la energía total de la $N_i^{\text{ésima}}$ trama de audio no es menor que una cuarta proporción preestablecida y una proporción de una suma de energía de F_b a cualquier frecuencia menor que F_s con la energía total de la $N_i^{\text{ésima}}$ trama de audio es menor que la cuarta proporción preestablecida, donde F_s es la frecuencia de demarcación de la $N_i^{\text{ésima}}$ trama de audio. El paso de determinación de la frecuencia de demarcación anterior se realiza para cada una de las N tramas de audio. De esta manera, se pueden obtener las N frecuencias de demarcación de las N tramas de audio. La determinación, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, de si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual incluye: cuando se determina que el parámetro de dispersión limitada por banda de las tramas de audio es menor que un decimocuarto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual.

40 Una persona experta en la técnica puede entender que la cuarta proporción preestablecida y el decimocuarto valor preestablecido pueden determinarse de acuerdo con un experimento de simulación. Se puede determinar un valor preestablecido apropiado y una proporción preestablecida de acuerdo con un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación. En general, un número menor que 1 pero cercano a 1, por ejemplo, 95 % o 99 %, se selecciona como un valor de la cuarta proporción preestablecida. Para la selección del decimocuarto valor preestablecido, generalmente, no se selecciona un número correspondiente a una frecuencia relativamente alta. Por ejemplo, en algunas realizaciones, si un rango de frecuencia de una trama de audio es de 0 Hz a 8 kHz, se puede seleccionar un número menor que una frecuencia de 5 kHz como el decimocuarto valor preestablecido.

50 Por ejemplo, se puede determinar la energía de cada una de las P envolventes espectrales de la trama de audio actual y se busca una frecuencia de demarcación de una frecuencia baja a una frecuencia alta de manera que una proporción que la energía, que es menor que la frecuencia de demarcación, representa en la energía total de la trama de audio actual es la cuarta proporción preestablecida. Suponiendo que N es 1, la frecuencia de demarcación de la trama de audio actual es el parámetro de dispersión limitada por banda. Suponiendo que N es un número entero mayor que 1, se determina que el valor promedio de la frecuencia de demarcación de las N tramas de audio es el parámetro de dispersión limitada por banda. Una persona experta en la técnica puede entender que determinar la frecuencia de demarcación mencionada anteriormente es simplemente un ejemplo. Alternativamente, el método

de determinación de frecuencia de demarcación puede ser buscar una frecuencia de demarcación de una frecuencia alta a una frecuencia baja o puede ser otro método.

Además, para evitar una conmutación frecuente entre el primer método de codificación y el segundo método de codificación, se puede ajustar además un periodo de retención. Para una trama de audio en el período de retención, se puede utilizar un método de codificación utilizado para una trama de audio en una posición de inicio del período de retención. De esta manera, se puede evitar una disminución de la calidad de conmutación causada por la conmutación frecuente entre diferentes métodos de codificación.

Si una duración de la duración de la retención es L , L tramas de audio después de la trama de audio actual pertenecen a un período de retención de la trama de audio actual. Si la dispersión de distribución, en un espectro, de la energía de una trama de audio que pertenece al período de retención es diferente de la dispersión de distribución, en un espectro, de la energía de una trama de audio en una posición de inicio del período de retención, la trama de audio sigue codificándose utilizando un método de codificación que es el mismo que el utilizado para la trama de audio en la posición de inicio del período de retención.

La duración del período de retención puede actualizarse de acuerdo con la dispersión de distribución, en un espectro, de energía de una trama de audio en el período de retención, hasta que la duración del período de retención sea 0.

Por ejemplo, si se determina utilizar el primer método de codificación para una i ésima trama de audio y una duración de un período de retención preestablecido es L , el primer método de codificación se utiliza para una $(i + 1)$ ésima trama de audio a una $(i + L)$ ésima trama de audio. Entonces, se determina la dispersión de distribución, en un espectro, de energía de la $(i + 1)$ ésima trama de audio y se recalcula el período de retención de acuerdo con la dispersión de distribución, en el espectro, de la energía de la $(i + 1)$ ésima trama de audio. Si la $(i + 1)$ ésima trama de audio todavía cumple una condición para utilizar el primer método de codificación, un período de retención posterior es todavía el período L de retención preestablecido. Es decir, el período de retención comienza a partir de una $(L + 2)$ ésima trama de audio a una $(i + 1 + L)$ ésima trama de audio. Si la $(i + 1)$ ésima trama de audio no cumple la condición para utilizar el primer método de codificación, se vuelve a determinar el período de retención de acuerdo con la dispersión de distribución, en el espectro, de la energía de la $(i + 1)$ ésima trama de audio. Por ejemplo, se vuelve a determinar que el período de retención es $L - L1$, donde $L1$ es un número entero positivo menor o igual que L . Si $L1$ es igual a L , la duración del período de retención se actualiza a 0. En este caso, el método de codificación se vuelve a determinar de acuerdo con la dispersión de distribución, en el espectro, de la energía de la $(i + 1)$ ésima trama de audio. Si $L1$ es un número entero menor que L , se vuelve a determinar el método de codificación de acuerdo con la dispersión de distribución, en un espectro, de energía de una $(i + 1 + L - L1)$ ésima trama de audio. Sin embargo, debido a que la $(i + 1)$ ésima trama de audio está en un período de retención de la i ésima trama de audio, la $(i + 1)$ ésima trama de audio se codifica todavía utilizando el primer método de codificación. $L1$ se puede referir como un parámetro de actualización de retención y un valor del parámetro de actualización de retención se puede determinar de acuerdo con la dispersión de distribución, en un espectro, de energía de una trama de audio de entrada. De esta manera, la actualización del período de retención está relacionada con la dispersión de distribución, en un espectro, de la energía de una trama de audio.

Por ejemplo, cuando se determina un parámetro de dispersión general y el parámetro de dispersión general es un primer ancho de banda mínimo, el período de retención puede volver a determinarse de acuerdo con un ancho de banda mínimo de distribución, en un espectro, de la primera proporción de energía preestablecida de una trama de audio. Se supone que se determina utilizar el primer método de codificación para codificar la i ésima trama de audio, y un período de retención preestablecido es L . Se determina un ancho de banda mínimo de distribución, en un espectro, de la primera proporción de energía preestablecida de cada una de las H tramas de audio consecutivas que incluyen la $(i + 1)$ ésima trama de audio, donde H es un número entero positivo mayor que 0. Si la $(i + 1)$ ésima trama de audio no cumple la condición para utilizar el primer método de codificación, se determina una cantidad de tramas de audio cuyos anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida son menores que un decimoquinto valor preestablecido (la cantidad se denomina brevemente como el primer parámetro de retención). Cuando un ancho de banda mínimo de distribución, en un espectro, de la primera proporción de energía preestablecida de una $(L + 1)$ ésima trama de audio es mayor que un decimosexto valor preestablecido y menor que un decimoséptimo valor preestablecido y, el primer parámetro de retención, es menor que un decimoctavo valor preestablecido, se resta 1 a la duración del período de retención, es decir, el parámetro de actualización de retención es 1. El decimosexto valor preestablecido es mayor que el primer valor preestablecido. Cuando el ancho de banda mínimo de distribución, en el espectro, de la primera proporción de energía preestablecida de la $(L + 1)$ ésima trama de audio es mayor que el decimoséptimo valor preestablecido y menor que un decimonoveno valor preestablecido y, el primer parámetro de retención, es menor que el decimoctavo valor preestablecido, se resta 2 a la duración del período de retención, es decir, el parámetro de actualización de retención es 2. Cuando el ancho de banda mínimo de distribución, en el espectro, de la primera proporción de energía preestablecida de la $(L + 1)$ ésima trama de audio es mayor que el decimonoveno valor preestablecido, el período de retención se ajusta a 0. Cuando el primer parámetro de retención y el ancho de banda mínimo de distribución, en el espectro, de la primera proporción de energía preestablecida de la $(L + 1)$ ésima trama de audio no

cumplen uno o más del decimosexto valor preestablecido al decimonoveno valor preestablecido, el período de retención permanece sin cambios.

5 Una persona experta en la técnica puede entender que, el período de retención preestablecido se puede fijar de acuerdo con un estado real y el parámetro de actualización de retención, también, se puede ajustar de acuerdo con un estado real. El decimoquinto valor preestablecido al decimonoveno valor preestablecido se pueden ajustar de acuerdo con un estado real, de modo que se pueden ajustar diferentes períodos de retención.

10 Del mismo modo, cuando el parámetro de dispersión general incluye un segundo ancho de banda mínimo y un tercer ancho de banda mínimo, o el parámetro de dispersión general incluye una primera proporción de energía o el parámetro de dispersión general incluye una segunda proporción de energía y una tercera proporción de energía, se pueden configurar un correspondiente período de retención preestablecido, un correspondiente parámetro de actualización de retención y un parámetro relacionado utilizado para determinar el parámetro de actualización de retención, de modo que se pueda determinar un correspondiente período de retención y se evita la conmutación frecuente entre métodos de codificación.

15 Cuando el método de codificación se determina de acuerdo con la dispersión de ráfaga (es decir, el método de codificación se determina de acuerdo con la dispersión global, la dispersión local y la ráfaga a corto plazo de distribución, en un espectro, de energía de una trama de audio), se puede ajustar un correspondiente período de retención, un correspondiente parámetro de actualización de retención y un parámetro relacionado utilizado para determinar el parámetro de actualización de retención, para evitar la conmutación frecuente entre métodos de codificación. En este caso, el período de retención puede ser menor que el período de retención ajustado en el caso del parámetro de dispersión general.

20 Cuando el método de codificación se determina de acuerdo a una característica limitada por banda de distribución de energía en un espectro, se pueden ajustar un correspondiente período de retención, un correspondiente parámetro de actualización de retención y un parámetro relacionado utilizado para determinar el parámetro de actualización de retención, para evitar la conmutación frecuente entre métodos de codificación. Por ejemplo, se puede calcular una proporción de energía de una envolvente espectral baja de una trama de audio de entrada con la energía de todas las envolventes espectrales y, el parámetro de actualización de retención, se determina de acuerdo con la proporción. Específicamente, la proporción de la energía de la envolvente espectral baja con la energía de todas las envolventes espectrales se puede determinar mediante la siguiente fórmula:

$$R_{baja} = \frac{\sum_{k=0}^y s(k)}{\sum_{k=0}^{P-1} s(k)}$$

Fórmula 1.5

30 donde R_{baja} representa la proporción de la energía de la envolvente espectral baja con la energía de todas las envolventes espectrales, $s(k)$ representa la energía de la $k^{ésima}$ envolvente espectral, y representa un índice de la envolvente espectral más alta de una banda de frecuencia baja y P indica que la trama de audio está dividida en P envolventes espectrales en total. En este caso, si R_{baja} es mayor que un vigésimo valor preestablecido, el parámetro de actualización de retención es 0. De lo contrario, si R_{baja} es mayor que un vigesimoprimer valor preestablecido, el parámetro de actualización de retención puede tener un valor relativamente pequeño, donde el vigésimo valor preestablecido es mayor que el vigesimoprimer valor preestablecido. Si R_{baja} no es mayor que el vigesimoprimer valor preestablecido, el parámetro de retención puede tener un valor relativamente grande. Una persona experta en la técnica puede entender que, el vigésimo valor preestablecido y el vigesimoprimer valor preestablecido se pueden determinar de acuerdo con un experimento de simulación y el valor del parámetro de actualización de retención, también, se puede determinar de acuerdo con un experimento. En general, un número que es una proporción excesivamente pequeña, generalmente, no se selecciona como el vigesimoprimer valor preestablecido. Por ejemplo, un número mayor que el 50 % puede seleccionarse en general. El vigésimo valor preestablecido oscila entre el vigesimoprimer valor preestablecido y 1.

45 Además, cuando el método de codificación se determina de acuerdo con una característica limitada por banda de distribución de energía en un espectro, la frecuencia de demarcación de una trama de audio de entrada se puede determinar, además, y el parámetro de actualización de retención se determina de acuerdo con la frecuencia de demarcación, donde la frecuencia de demarcación puede ser diferente de una frecuencia de demarcación utilizada para determinar un parámetro de dispersión limitada por banda. Si la frecuencia de demarcación es menor que un vigesimosegundo valor preestablecido, el parámetro de actualización de retención es 0. De lo contrario, si la frecuencia de demarcación es menor que un vigesimoterter valor preestablecido, el parámetro de actualización de retención tiene un valor relativamente pequeño. El vigesimoterter valor preestablecido es mayor que el vigesimosegundo valor preestablecido. Si la frecuencia de demarcación es mayor que el vigesimoterter valor preestablecido, el parámetro de actualización de retención puede tener un valor relativamente grande. Una persona

- experta en la técnica puede entender que, el vigesimosegundo valor preestablecido y el vigesimotercer valor preestablecido se pueden determinar de acuerdo con un experimento de simulación y el valor del parámetro de actualización de retención, también, se puede determinar de acuerdo con un experimento. En general, un número correspondiente a una frecuencia relativamente alta no se selecciona como el vigesimotercer valor preestablecido.
- 5 Por ejemplo, si un rango de frecuencia de una trama de audio es de 0 Hz a 8 kHz, se puede seleccionar un número menor que una frecuencia de 5 kHz como el vigesimotercer valor preestablecido.
- La FIG. 2 es un diagrama de bloques estructural de un aparato de acuerdo con una realización de la presente invención. El aparato 200 mostrado en la FIG. 2 puede realizar los pasos en la FIG. 1. Como se muestra en la FIG. 2, el aparato 200 incluye una unidad 201 de obtención y una unidad 202 de determinación.
- 10 La unidad 201 de obtención está configurado para obtener N tramas de audio, donde las N tramas de audio incluyen una trama de audio actual y N es un número entero positivo.
- La unidad 202 de determinación está configurado para determinar la dispersión de distribución, en el espectro, de energía de las N tramas de audio obtenidas por la unidad 201 de obtención.
- 15 La unidad 202 de determinación está configurada además para determinar, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual, donde el primer método de codificación es un método de codificación que está basado en la transformada de tiempo-frecuencia y la cuantificación del coeficiente de transformada y que no está basado en la predicción lineal y, el segundo método de codificación, es un método de codificación basado en predicción lineal.
- 20 De acuerdo con el aparato mostrado en la FIG. 2, cuando se codifica una trama de audio, se considera la dispersión de distribución, en un espectro, de energía de la trama de audio, lo que puede reducir la complejidad de codificación y asegurar que la codificación sea de una precisión relativamente alta.
- Durante la selección de un método de codificación apropiado para una trama de audio, se puede considerar la dispersión de distribución, en un espectro, de energía de la trama de audio. Puede haber tres tipos de dispersión de distribución, en un espectro, de energía de una trama de audio: dispersión general, dispersión de ráfaga y dispersión limitada por banda.
- 25 Opcionalmente, en una realización, se puede seleccionar un método de codificación apropiado para la trama de audio actual utilizando la dispersión general. En este caso, la unidad 202 de determinación está configurada específicamente para dividir un espectro de cada una de las N tramas de audio en P envolventes espectrales y determinar un parámetro de dispersión general de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde P es un número entero positivo y el parámetro de dispersión general indica la dispersión de distribución, en el espectro, de la energía de las N tramas de audio.
- 30 Específicamente, un valor promedio de anchos de banda mínimos de distribución, en el espectro, de energía de proporción específica de N tramas de audio consecutivas de entrada puede definirse como la dispersión general. Un ancho de banda menor indica una dispersión general más fuerte y un ancho de banda mayor indica una dispersión general más débil. En otras palabras, una dispersión general más fuerte indica que la energía de una trama de audio está más centralizada y una dispersión general más débil indica que la energía de una trama de audio está más dispersa. La eficiencia es alta cuando se utiliza el primer método de codificación para codificar una trama de audio cuya dispersión general es relativamente fuerte. Por lo tanto, se puede seleccionar un método de codificación apropiado, determinando la dispersión general de una trama de audio, para codificar la trama de audio. Para ayudar a determinar la dispersión general de una trama de audio, la dispersión general se puede cuantificar para obtener un parámetro de dispersión general. Opcionalmente, cuando N es 1, la dispersión general es un ancho de banda mínimo de distribución, en un espectro, de energía de proporción específica de la trama de audio actual.
- 35 Opcionalmente, en una realización, el parámetro de dispersión general incluye un primer ancho de banda mínimo. En este caso, la unidad 202 de determinación está configurada específicamente para determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio es el primer ancho de banda mínimo. La unidad 202 de determinación está configurada específicamente para: cuando el primer ancho de banda mínimo es menor que un primer valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; y cuando el primer ancho de banda mínimo es mayor que el primer valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual.
- 40 45 50 55 Una persona experta en la técnica puede entender que, el primer valor preestablecido y la primera proporción preestablecida se pueden determinar de acuerdo con un experimento de simulación. Un primer valor preestablecido apropiado y una primera proporción preestablecida se pueden determinar por medio de un experimento de

simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación o el segundo método de codificación.

La unidad 202 de determinación está configurada específicamente para: ordenar la energía de las P envolventes espectrales de cada una de las tramas de audio en orden descendente; determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio; y determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio, un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de las N tramas de audio. Por ejemplo, una señal de audio obtenida por la unidad 201 de obtención es una señal de banda ancha muestreada a 16 kHz y la señal de audio obtenida se obtiene en una trama de 20 ms. Cada una de las tramas de señal es de 320 puntos de muestreo en el dominio del tiempo. La unidad 202 de determinación puede realizar una transformada de tiempo-frecuencia en una señal de dominio del tiempo, por ejemplo, realizar la transformada de tiempo-frecuencia por medio de la transformada rápida de Fourier (Fast Fourier Transformation, FFT), para obtener 160 S(k) envolventes espectrales, es decir, 160 coeficientes de espectro de energía de FFT, donde $k = 0, 1, 2, \dots, 159$. La unidad 202 de determinación puede encontrar un ancho de banda mínimo de las S(k) envolventes espectrales de una manera que una proporción que la energía en el ancho de banda representa en la energía total de la trama es la primera proporción preestablecida. Específicamente, la unidad 202 de determinación puede acumular secuencialmente energía de intervalos de frecuencia en las S(k) envolventes espectrales en orden descendente; y comparar la energía obtenida después de cada acumulación con la energía total de la trama de audio y, si una proporción es mayor que la primera proporción preestablecida, finalizar el proceso de acumulación, donde una cantidad de veces de acumulación es el ancho de banda mínimo. Por ejemplo, la primera proporción preestablecida es del 90 % y, si una proporción que una suma de energía obtenida después de 30 veces de acumulación representa en la energía total supera el 90 %, se puede considerar que un ancho de banda mínimo de energía que representa no menos que primera proporción preestablecida de la trama de audio es 30. La unidad 202 de determinación puede ejecutar el proceso de determinación de ancho de banda mínimo anterior para cada una de las N tramas de audio, para determinar por separado los anchos de banda mínimos de la energía que representa no menos que la primera proporción preestablecida de las N tramas de audio, incluyendo la trama de audio actual. La unidad 202 de determinación puede calcular un valor promedio de los anchos de banda mínimos de la energía que representa no menos que la primera proporción preestablecida de las N tramas de audio. El valor promedio de los anchos de banda mínimos de la energía que representa no menos que la primera proporción preestablecida de las N tramas de audio puede denominarse como el primer ancho de banda mínimo y, el primer ancho de banda mínimo, se puede utilizar como el parámetro de dispersión general. Cuando el primer ancho de banda mínimo es menor que el primer valor preestablecido, la unidad 202 de determinación puede determinar utilizar el primer método de codificación para codificar la trama de audio actual. Cuando el primer ancho de banda mínimo es mayor que el primer valor preestablecido, la unidad 202 de determinación puede determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

Opcionalmente, en otra realización, el parámetro de dispersión general puede incluir una primera proporción de energía. En este caso, la unidad 202 de determinación está configurada específicamente para seleccionar P_1 envolventes espectrales de las P envolventes espectrales de cada una de las N tramas de audio y determinar la primera proporción de energía de acuerdo con la energía de las P_1 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio, donde P_1 es un número entero positivo menor que P. La unidad 202 de determinación está configurada específicamente para: cuando la primera proporción de energía es mayor que un segundo valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; y cuando la primera proporción de energía es menor que el segundo valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual. Opcionalmente, en una realización, cuando N es 1, las N tramas de audio son la trama de audio actual y la unidad 202 de determinación está configurada específicamente para determinar la primera proporción de energía de acuerdo con la energía de las P_1 envolventes espectrales de la trama de audio actual y la energía total de la trama de audio actual. La unidad 202 de determinación está configurada específicamente para determinar las P_1 envolventes espectrales de acuerdo con la energía de las P envolventes espectrales, donde la energía de una cualquiera de las P_1 envolventes espectrales es mayor que la energía de una cualquiera de las otras envolventes espectrales en las P envolventes espectrales excepto las P_1 envolventes espectrales.

Específicamente, la unidad 202 de determinación puede calcular la primera proporción de energía utilizando la siguiente fórmula:

$$\left\{ \begin{array}{l} R_1 = \frac{\sum_{n=1}^N r(n)}{N} \\ r(n) = \frac{E_{p1}(n)}{E_{todo}(n)} \end{array} \right.$$

Fórmula 1.6

donde R_1 representa la primera proporción de energía, $E_{p1}(n)$ representa una suma de energía de las P_1 envolventes espectrales seleccionadas en una $n^{\text{ésima}}$ trama de audio, $E_{todo}(n)$ representa la energía total de la $n^{\text{ésima}}$ trama de audio y $r(n)$ representa una proporción que la energía de las P_1 envolventes espectrales de la $n^{\text{ésima}}$ trama de audio en las N tramas de audio representa en la energía total de la trama de audio.

Una persona experta en la técnica puede entender que, el segundo valor preestablecido y la selección de las P_1 envolventes espectrales se pueden determinar de acuerdo con un experimento de simulación. Un segundo valor preestablecido apropiado, un valor apropiado de P_1 y un método apropiado para seleccionar las P_1 envolventes espectrales se pueden determinar por medio de un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando una trama de audio que cumple la condición anterior se codifica utilizando el primer método de codificación o el segundo método de codificación. Opcionalmente, en una realización, las P_1 envolventes espectrales pueden ser P_1 envolventes espectrales que tienen energía máxima en las P envolventes espectrales.

Por ejemplo, una señal de audio obtenida por la unidad 201 de obtención es una señal de banda ancha muestreada a 16 kHz y la señal de audio obtenida se obtiene en una trama de 20 ms. Cada una de las tramas de la señal es de 320 puntos de muestreo en el dominio del tiempo. La unidad 202 de determinación puede realizar una transformada de tiempo-frecuencia en una señal de dominio del tiempo, por ejemplo, realizar una transformada de tiempo-frecuencia por medio de la transformada rápida de Fourier, para obtener 160 $S(k)$ envolventes espectrales, donde $k = 0, 1, 2, \dots, 159$. La unidad 202 de determinación puede seleccionar P_1 envolventes espectrales de las 160 envolventes espectrales y calcular una proporción que una suma de energía de las P_1 envolventes espectrales representa en la energía total de la trama de audio. La unidad 202 de determinación puede ejecutar el proceso anterior para cada una de las N tramas de audio, es decir, calcular una proporción que una suma de energía de las P_1 envolventes espectrales de cada una de las N tramas de audio representa en la respectiva energía total. La unidad 202 de determinación puede calcular un valor promedio de las proporciones. El valor promedio de las proporciones es la primera proporción de energía. Cuando la primera proporción de energía es mayor que el segundo valor preestablecido, la unidad 202 de determinación puede determinar utilizar el primer método de codificación para codificar la trama de audio actual. Cuando la primera proporción de energía es menor que el segundo valor preestablecido, la unidad 202 de determinación puede determinar utilizar el segundo método de codificación para codificar la trama de audio actual. Las P_1 envolventes espectrales pueden ser P_1 envolventes espectrales que tienen energía máxima en las P envolventes espectrales. Es decir, la unidad 202 de determinación está configurada específicamente para determinar, a partir de las P envolventes espectrales de cada una de las N tramas de audio, P_1 envolventes espectrales que tienen energía máxima. Opcionalmente, en una realización, el valor de P_1 puede ser 20.

Opcionalmente, en otra realización, el parámetro de dispersión general puede incluir un segundo ancho de banda mínimo y un tercer ancho de banda mínimo. En este caso, la unidad 202 de determinación está configurada específicamente para determinar un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la segunda proporción de energía preestablecida de las N tramas de audio y determinar un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la tercera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la segunda proporción de energía preestablecida de las N tramas de audio se utiliza como el segundo ancho de banda mínimo, el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la tercera proporción de energía preestablecida de las N tramas de audio se utiliza como el tercer ancho de banda mínimo y la segunda proporción preestablecida es menor que la tercera proporción preestablecida. La unidad 202 de determinación está específicamente configurada para: cuando el segundo ancho de banda mínimo es menor que un tercer valor preestablecido y el tercer ancho de banda mínimo es menor que un cuarto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; cuando el tercer ancho de banda mínimo es menor que un quinto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; y cuando el tercer ancho de banda mínimo es mayor que un sexto valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual. Opcionalmente, en una realización, cuando N es 1, las N tramas de audio son la trama de audio actual. La unidad 202 de determinación puede determinar un ancho de banda mínimo de distribución, en el espectro, de la segunda proporción de energía preestablecida de la trama de audio actual como el segundo ancho de banda mínimo. La unidad 202 de

determinación puede determinar un ancho de banda mínimo de distribución, en el espectro, de la tercera proporción de energía preestablecida de la trama de audio actual como el tercer ancho de banda mínimo.

5 Una persona experta en la técnica puede entender que, el tercer valor preestablecido, el cuarto valor preestablecido, el quinto valor preestablecido, el sexto valor preestablecido, la segunda proporción preestablecida y la tercera proporción preestablecida se pueden determinar de acuerdo con un experimento de simulación. Los valores preestablecidos apropiados y las proporciones preestablecidas pueden determinarse por medio de un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación o el segundo método de codificación.

10 La unidad 202 de determinación está configurada específicamente para: ordenar la energía de las P envolventes espectrales de cada una de las tramas de audio en orden descendente; determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la segunda proporción preestablecida de cada una de las N tramas de audio; determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la segunda proporción preestablecida de cada una de las N tramas de audio, un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la segunda proporción preestablecida de las N tramas de audio; determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la tercera proporción preestablecida de cada una de las N tramas de audio; y determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de cada una de las N tramas de audio, un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la tercera proporción preestablecida de las N tramas de audio. Por ejemplo, una señal de audio obtenida por la unidad 201 de obtención es una señal de banda ancha muestreada a 16 kHz y la señal de audio obtenida se obtiene en una trama de 20 ms. Cada una de las tramas de señal es de 320 puntos de muestreo en el dominio del tiempo. La unidad 202 de determinación puede realizar una transformada de tiempo-frecuencia en una señal de dominio del tiempo, por ejemplo, realizar la transformada de tiempo-frecuencia por medio de la transformada rápida de Fourier, para obtener 160 S(k) envolventes espectrales, donde $k = 0, 1, 2, \dots, 159$. La unidad 202 de determinación puede encontrar un ancho de banda mínimo de las S(k) envolventes espectrales de manera que una proporción que la energía en el ancho de banda representa en la energía total de la trama no sea menor que la segunda proporción preestablecida. La unidad 202 de determinación puede continuar encontrando un ancho de banda de las S(k) envolventes espectrales de manera que una proporción que la energía en el ancho de banda representa en la energía total no sea menor que la tercera proporción preestablecida. Específicamente, la unidad 202 de determinación puede acumular secuencialmente energía de intervalos de frecuencia en las S(k) envolventes espectrales en orden descendente. La energía obtenida después de cada acumulación se compara con la energía total de la trama de audio y, si una proporción es mayor que la segunda proporción preestablecida, una cantidad de veces de acumulación es un ancho de banda mínimo que no es menor que la segunda proporción preestablecida. La unidad 202 de determinación puede continuar la acumulación. Si una proporción de energía obtenida después de la acumulación con la energía total de la trama de audio es mayor que la tercera proporción preestablecida, la acumulación finaliza y, una cantidad de veces de acumulación, es un ancho de banda mínimo que no es menor que la tercera proporción preestablecida. Por ejemplo, la segunda proporción preestablecida es del 85 % y la tercera proporción preestablecida es del 95 %. Si una proporción que una suma de energía obtenida después de 30 veces de acumulación representa en el total de energía excede el 85 %, se puede considerar que el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la segunda proporción preestablecida de la trama de audio es 30. La acumulación continúa y, si una proporción que una suma de energía obtenida después de 35 veces de acumulación representa en la energía total es del 95 %, se puede considerar que el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de la trama de audio es 35. La unidad 202 de determinación puede ejecutar el proceso anterior para cada una de las N tramas de audio. La unidad 202 de determinación puede determinar por separado los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la segunda proporción preestablecida de las N tramas de audio, incluyendo la trama de audio actual y los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de las N tramas de audio, incluyendo la trama de audio actual. El valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la segunda proporción preestablecida de las N tramas de audio es el segundo ancho de banda mínimo. El valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de las N tramas de audio es el tercer ancho de banda mínimo. Cuando el segundo ancho de banda mínimo es menor que el tercer valor preestablecido y el tercer ancho de banda mínimo es menor que el cuarto valor preestablecido, la unidad 202 de determinación puede determinar utilizar el primer método de codificación para codificar la trama de audio actual. Cuando el tercer ancho de banda mínimo es menor que el quinto valor preestablecido, la unidad 202 de determinación puede determinar utilizar el primer método de codificación para codificar la trama de audio actual. Cuando el tercer ancho de banda mínimo es mayor que el primer valor preestablecido, la unidad 202 de determinación puede determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

Opcionalmente, en otra realización, el parámetro de dispersión general incluye una segunda proporción de energía y una tercera proporción de energía. En este caso, la unidad 202 de determinación está específicamente configurada para: seleccionar P_2 envolventes espectrales de las P envolventes espectrales de cada una de las N tramas de audio, determinar la segunda proporción de energía de acuerdo con la energía de las P_2 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio, seleccionar P_3 envolventes espectrales de las P envolventes espectrales de cada una de las N tramas de audio y determinar la tercera proporción de energía de acuerdo con la energía de las P_3 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio, donde P_2 y P_3 son números enteros positivos menores que P y P_2 es menor que P_3 . La unidad 202 de determinación está configurada específicamente para: cuando la segunda proporción de energía es mayor que un séptimo valor preestablecido y la tercera proporción de energía es mayor que un octavo valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; cuando la segunda proporción de energía es mayor que un noveno valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; y cuando la tercera proporción de energía es menor que un décimo valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual. Opcionalmente, en una realización, cuando N es 1, las N tramas de audio son la trama de audio actual. La unidad 202 de determinación puede determinar la segunda proporción de energía de acuerdo con la energía de las P_2 envolventes espectrales de la trama de audio actual y la energía total de la trama de audio actual. La unidad 202 de determinación puede determinar la tercera proporción de energía de acuerdo con la energía de las P_3 envolventes espectrales de la trama de audio actual y la energía total de la trama de audio actual.

Una persona experta en la técnica puede entender que, los valores de P_2 y P_3 , el séptimo valor preestablecido, el octavo valor preestablecido, el noveno valor preestablecido y el décimo valor preestablecido se pueden determinar de acuerdo con un experimento de simulación. Los valores preestablecidos apropiados pueden determinarse por medio de un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación o el segundo método de codificación. Opcionalmente, en una realización, la unidad 202 de determinación está configurada específicamente para determinar, a partir de las P envolventes espectrales de cada una de las N tramas de audio, las P_2 envolventes espectrales que tienen energía máxima y determinar, a partir de las P envolventes espectrales de cada una de las N tramas de audio, las P_3 envolventes espectrales que tienen energía máxima.

Por ejemplo, una señal de audio obtenida por la unidad 201 de obtención es una señal de banda ancha muestreada a 16 kHz y la señal de audio obtenida se obtiene en una trama de 20 ms. Cada una de las tramas de señal es de 320 puntos de muestreo en el dominio del tiempo. La unidad 202 de determinación puede realizar una transformada de tiempo-frecuencia en una señal de dominio del tiempo, por ejemplo, realizar una transformada de tiempo-frecuencia por medio de la transformada rápida de Fourier, para obtener 160 $S(k)$ envolventes espectrales, donde $k = 0, 1, 2, \dots, 159$. La unidad 202 de determinación puede seleccionar P_2 envolventes espectrales de las 160 envolventes espectrales y calcular una proporción que una suma de energía de las P_2 envolventes espectrales representa en la energía total de la trama de audio. La unidad 202 de determinación puede ejecutar el proceso anterior para cada una de las N tramas de audio, es decir, calcular una proporción que una suma de energía de las P_2 envolventes espectrales cada una de las N tramas de audio representa en la respectiva energía total. La unidad 202 de determinación puede calcular un valor promedio de las proporciones. El valor promedio de las proporciones es la segunda proporción de energía. La unidad 202 de determinación puede seleccionar P_3 envolventes espectrales de las 160 envolventes espectrales y calcular una proporción que una suma de energía de las P_3 las envolventes espectrales representa en la energía total de la trama de audio. La unidad 202 de determinación puede ejecutar el proceso anterior para cada una de las N tramas de audio, es decir, calcular una proporción que una suma de energía de las P_3 envolventes espectrales de cada una de las N tramas de audio representa en la respectiva energía total. La unidad 202 de determinación puede calcular un valor promedio de las proporciones. El valor promedio de las proporciones es la tercera proporción de energía. Cuando la segunda proporción de energía es mayor que el séptimo valor preestablecido y la tercera proporción de energía es mayor que el octavo valor preestablecido, la unidad 202 de determinación puede determinar utilizar el primer método de codificación para codificar la trama de audio actual. Cuando la segunda proporción de energía es mayor que el noveno valor preestablecido, la unidad 202 de determinación puede determinar utilizar el primer método de codificación para codificar la trama de audio actual. Cuando la tercera proporción de energía es menor que el décimo valor preestablecido, la unidad 202 de determinación puede determinar utilizar el segundo método de codificación para codificar la trama de audio actual. Las P_2 envolventes espectrales pueden ser P_2 envolventes espectrales que tienen energía máxima en las P envolventes espectrales; y las P_3 envolventes espectrales pueden ser P_3 envolventes espectrales que tienen energía máxima en las P envolventes espectrales. Opcionalmente, en una realización, el valor de P_2 puede ser 20 y el valor de P_3 puede ser 30.

Opcionalmente, en otra realización, se puede seleccionar un método de codificación apropiado para la trama de audio actual utilizando la dispersión de ráfaga. Para la dispersión de ráfaga, se deben considerar la dispersión global, la dispersión local y la ráfaga a corto plazo de distribución, en un espectro, de energía de una trama de audio. En este caso, la dispersión de distribución de la energía en el espectro puede incluir la dispersión global, la dispersión local y la ráfaga a corto plazo de distribución de la energía en el espectro. En este caso, un valor de N puede ser 1 y las N tramas de audio son la trama de audio actual. La unidad 202 de determinación está configurada

específicamente para dividir un espectro de la trama de audio actual en Q subbandas y determinar un parámetro de dispersión de ráfaga de acuerdo con la energía máxima de cada una de las Q subbandas del espectro de la trama de audio actual, donde el parámetro de dispersión de ráfaga se utiliza para indicar la dispersión global, la dispersión local y la ráfaga a corto plazo de la trama de audio actual.

5 Específicamente, la unidad 202 de determinación está configurada específicamente para determinar una proporción global de máximo a promedio de cada una de las Q subbandas, una proporción local de máximo a promedio de cada una de las Q subbandas y una fluctuación de energía a corto plazo de cada una de las Q subbandas, donde la proporción global de máximo a promedio se determina por la unidad 202 de determinación de acuerdo con la energía máxima en la subbanda y la energía promedio de todas las subbandas de la trama de audio actual, la proporción local de máximo a promedio se determina por la unidad 202 de determinación de acuerdo con la energía máxima en la subbanda y la energía promedio en la subbanda y, la fluctuación de energía máxima a corto plazo, se determina de acuerdo con la energía máxima en la subbanda y la energía máxima en una banda de frecuencia específica de una trama de audio antes de la trama de audio. La proporción global de máximo a promedio de cada una de las Q subbandas, la proporción local de máximo a promedio de cada una de las Q subbandas y la fluctuación de energía a corto plazo de cada una de las Q subbandas, respectivamente, representan la dispersión global, la dispersión local y la ráfaga a corto plazo. La unidad 202 de determinación está configurada específicamente para: determinar si hay una primera subbanda en las Q subbandas, donde una proporción local de máximo a promedio de la primera subbanda es mayor que un decimoprimer valor preestablecido, una proporción global de máximo a promedio de la primera subbanda es mayor que un decimosegundo valor preestablecido y una fluctuación de energía máxima a corto plazo de la primera subbanda es mayor que un decimotercer valor preestablecido; y cuando la primera subbanda está en las Q subbandas, determinar utilizar el primer método de codificación para codificar la trama de audio actual.

Específicamente, la unidad 202 de determinación puede calcular la proporción global de máximo a promedio utilizando la siguiente fórmula:

$$p2s(i) = e(i) / \left(\frac{1}{P} * \sum_{k=0}^{P-1} s(k) \right) \quad \text{Fórmula 1.7}$$

donde e(i) representa la energía máxima de una i^{ésima} subbanda en las Q subbandas, s(k) representa la energía de una k^{ésima} envolvente espectral en las P envolventes espectrales y p2s(i) representa una proporción global de máximo a promedio de la i^{ésima} subbanda.

La unidad 202 de determinación puede calcular la proporción local de máximo a promedio utilizando la siguiente fórmula:

$$p2a(i) = e(i) / \left(\frac{1}{h(i) - l(i) + 1} * \sum_{k=l(i)}^{h(i)} s(k) \right) \quad \text{Fórmula 1.8}$$

donde e(i) representa la energía máxima de la i^{ésima} subbanda en las Q subbandas, s(k) representa la energía de la k^{ésima} envolvente espectral en las P envolventes espectrales, h(i) representa un índice de una envolvente espectral que está incluida en la i^{ésima} subbanda y que tiene una frecuencia más alta, l(i) representa un índice de una envolvente espectral que está incluida en la i^{ésima} subbanda y que tiene una frecuencia más baja, p2a(i) representa una proporción local de máximo a promedio de la i^{ésima} subbanda y h(i) es menor o igual que P - 1.

La unidad 202 de determinación puede calcular la fluctuación de energía máxima a corto plazo utilizando la siguiente fórmula:

$$dev(i) = (2 * e(i)) / (e_1 + e_2) \quad \text{Fórmula 1.9}$$

donde e(i) representa la energía máxima de la i^{ésima} subbanda en las Q subbandas de la trama de audio actual y e₁ y e₂ representan la energía máxima de bandas de frecuencia específicas de tramas de audio antes de la trama de audio actual. Específicamente, suponiendo que la trama de audio actual es una M^{ésima} trama de audio, se determina una envolvente espectral en la que se encuentra la energía máxima de la i^{ésima} subbanda de la trama de audio actual. Se supone que la envolvente espectral en la que se encuentra la energía máxima es i₁. Se determina la energía máxima dentro de un rango de una (i₁ - t)^{ésima} envolvente espectral a una (i₁ + t)^{ésima} envolvente espectral en una (M - 1)^{ésima} trama de audio y la energía máxima es e₁. Del mismo modo, se determina la energía máxima dentro de un rango de una (i₁ - t)^{ésima} envolvente espectral a una (i₁ + t)^{ésima} envolvente espectral en una (M - 2)^{ésima} trama de audio y la energía máxima es e₂.

Una persona experta en la técnica puede entender que, el decimoprimer valor preestablecido, el decimosegundo valor preestablecido y el decimotercer valor preestablecido se pueden determinar de acuerdo con un experimento de simulación. Los valores preestablecidos apropiados pueden determinarse por medio de un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación.

Opcionalmente, en otra realización, se puede seleccionar un método de codificación apropiado para la trama de audio actual utilizando la dispersión limitada por banda. En este caso, la dispersión de distribución de la energía en el espectro incluye la dispersión de distribución limitada por banda de la energía en el espectro. En este caso, la unidad 202 de determinación está configurada específicamente para determinar una frecuencia de demarcación de cada una de las N tramas de audio. La unidad 202 de determinación está configurada específicamente para determinar un parámetro de dispersión limitada por banda de acuerdo con la frecuencia de demarcación de cada una de las N tramas de audio.

Una persona experta en la técnica puede entender que, la cuarta proporción preestablecida y el decimocuarto valor preestablecido pueden determinarse de acuerdo con un experimento de simulación. Se puede determinar un valor preestablecido apropiado y una proporción preestablecida de acuerdo con un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación.

Por ejemplo, la unidad 202 de determinación puede determinar la energía de cada una de las P envolventes espectrales de la trama de audio actual y buscar una frecuencia de demarcación de una frecuencia baja a una frecuencia alta de manera tal que una proporción que la energía, que es menor que la frecuencia de demarcación, representa en la energía total de la trama de audio actual es la cuarta proporción preestablecida. El parámetro de dispersión limitada por banda puede ser un valor promedio de las frecuencias de demarcación de las N tramas de audio. En este caso, la unidad 202 de determinación está configurada específicamente para: cuando se determina que el parámetro de dispersión limitada por banda de las tramas de audio es menor que un decimocuarto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual. Suponiendo que N es 1, la frecuencia de demarcación de la trama de audio actual es el parámetro de dispersión limitada por banda. Suponiendo que N es un número entero mayor que 1, la unidad 202 de determinación puede determinar que el valor promedio de las frecuencias de demarcación de las N tramas de audio es el parámetro de dispersión limitada por banda. Una persona experta en la técnica puede entender que la determinación de la frecuencia de demarcación mencionada anteriormente es simplemente un ejemplo. Alternativamente, el método de determinación de la frecuencia de demarcación puede ser buscar una frecuencia de demarcación de una frecuencia alta a una frecuencia baja o puede ser otro método.

Además, para evitar una conmutación frecuente entre el primer método de codificación y el segundo método de codificación, la unidad 202 de determinación puede estar configurada además para ajustar un período de retención. La unidad 202 de determinación puede estar configurada para: para una trama de audio en el período de retención, utilizar un método de codificación utilizado para una trama de audio en una posición de inicio del período de retención. De esta manera, se puede evitar una disminución de la calidad de conmutación causada por la conmutación frecuente entre diferentes métodos de codificación.

Si la duración de retención del período de retención es L, la unidad 202 de determinación puede estar configurada para determinar que L tramas de audio después de la trama de audio actual pertenecen a un período de retención de la trama de audio actual. Si la dispersión de distribución, en un espectro, de energía de una trama de audio que pertenece al período de retención es diferente de la dispersión de distribución, en un espectro, de energía de una trama de audio en una posición de inicio del período de retención, la unidad 202 de determinación puede estar configurada para determinar que la trama de audio todavía está codificada utilizando un método de codificación que es el mismo que el que se utiliza para la trama de audio en la posición de inicio del período de retención.

La duración del período de retención puede actualizarse de acuerdo con la dispersión de distribución, en un espectro, de energía de una trama de audio en el período de retención, hasta que la duración del período de retención sea 0.

Por ejemplo, si la unidad 202 de determinación determina utilizar el primer método de codificación para una $l^{\text{ésima}}$ trama de audio y una duración de un período de retención preestablecido es L, la unidad 202 de determinación puede determinar que el primer método de codificación se utiliza para una $(l + 1)^{\text{ésima}}$ trama de audio a una $(l + L)^{\text{ésima}}$ trama de audio. Entonces, la unidad 202 de determinación puede determinar la dispersión de distribución, en un espectro, de energía de la $(l + 1)^{\text{ésima}}$ trama de audio y volver a calcular el período de retención de acuerdo con la dispersión de distribución, en el espectro, de la energía de la $(l + 1)^{\text{ésima}}$ trama de audio. Si la $(l + 1)^{\text{ésima}}$ trama de audio todavía cumple una condición para utilizar el primer método de codificación, la unidad 202 de determinación puede determinar que un período de retención posterior es todavía el período L de retención preestablecido. Es decir, el período de retención comienza a partir de una $(L + 2)^{\text{ésima}}$ trama de audio a una $(l + 1 + L)^{\text{ésima}}$ trama de audio. Si la $(l + 1)^{\text{ésima}}$ trama de audio no cumple la condición de utilizar el primer método de codificación, la unidad 202 de determinación puede volver a determinar el período de retención de acuerdo con la dispersión de

distribución, en el espectro, de la energía de la $(l + 1)^{\text{ésima}}$ trama de audio. Por ejemplo, la unidad 202 de determinación puede volver a determinar que el período de retención es $L - L1$, donde $L1$ es un número entero positivo menor o igual que L . Si $L1$ es igual a L , la duración del período de retención se actualiza a 0. En este caso, la unidad 202 de determinación puede volver a determinar el método de codificación de acuerdo con la dispersión de distribución, en el espectro, de la energía de la $(l + 1)^{\text{ésima}}$ trama de audio. Si $L1$ es un número entero menor que L , la unidad 202 de determinación puede volver a determinar el método de codificación de acuerdo con la dispersión de distribución, en un espectro, de energía de una $(l + 1 + L - L1)^{\text{ésima}}$ trama de audio. Sin embargo, debido a que la $(l + 1)^{\text{ésima}}$ trama de audio está en un período de retención de la $l^{\text{ésima}}$ trama de audio, la $(l + 1)^{\text{ésima}}$ trama de audio todavía se codifica utilizando el primer método de codificación. Se puede hacer referencia a $L1$ como un parámetro de actualización de retención y se puede determinar un valor del parámetro de actualización de retención de acuerdo con la dispersión de distribución, en un espectro, de energía de una trama de audio de entrada. De esta manera, la actualización del período de retención está relacionada con la dispersión de distribución, en un espectro, de energía de una trama de audio.

Por ejemplo, cuando se determina un parámetro de dispersión general y el parámetro de dispersión general es un primer ancho de banda mínimo, la unidad 202 de determinación puede volver a determinar el período de retención de acuerdo con un ancho de banda mínimo de distribución, en un espectro, de la primera proporción de energía preestablecida de una trama de audio. Se supone que se determina utilizar el primer método de codificación para codificar la $l^{\text{ésima}}$ trama de audio y un período de retención preestablecido es L . La unidad 202 de determinación pueden determinar un ancho de banda mínimo de distribución, en un espectro, de la primera proporción de energía preestablecida de cada una de las H tramas de audio consecutivas incluyendo la $(l + 1)^{\text{ésima}}$ trama de audio, donde H es un número entero positivo mayor que 0. Si la $(l + 1)^{\text{ésima}}$ trama de audio no cumple la condición de utilizar el primer método de codificación, la unidad 202 de determinación puede determinar una cantidad de tramas de audio cuyos anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida sean menores que un decimoquinto valor preestablecido (la cantidad se refiere brevemente como un primer parámetro de retención). Cuando un ancho de banda mínimo de distribución, en un espectro, de la primera proporción de energía preestablecida de una $(L + 1)^{\text{ésima}}$ trama de audio es mayor que un decimosexto valor preestablecido y menor que un decimoséptimo valor preestablecido y, el primer parámetro de retención es menor que un decimoctavo valor preestablecido, la unidad 202 de determinación puede restar 1 a la duración del período de retención, es decir, el parámetro de actualización de retención es 1. El decimosexto valor preestablecido es mayor que el primer valor preestablecido. Cuando el ancho de banda mínimo de distribución, en el espectro, de la primera proporción de energía preestablecida de la $(L + 1)^{\text{ésima}}$ trama de audio es mayor que el decimoséptimo valor preestablecido y menor que un decimonoveno valor preestablecido y, el primer parámetro de retención es menor que el decimoctavo valor preestablecido, la unidad 202 de determinación puede restar 2 a la duración del período de retención, es decir, el parámetro de actualización de retención es 2. Cuando el ancho de banda mínimo de distribución, en el espectro, de la primera proporción de energía preestablecida de la $(L + 1)^{\text{ésima}}$ trama de audio es mayor que el decimonoveno valor preestablecido, la unidad 202 de determinación puede ajustar el período de retención a 0. Cuando el primer parámetro de retención y el ancho de banda mínimo de distribución, en el espectro, de la primera proporción de energía preestablecida de la $(L + 1)^{\text{ésima}}$ trama de audio no cumple uno o más del decimosexto valor preestablecido al decimonoveno valor preestablecido, la unidad 202 de determinación puede determinar que el período de retención permanece sin cambios.

Una persona experta en la técnica puede entender que, el período de retención preestablecido se puede ajustar de acuerdo con un estado real y el parámetro de actualización de retención, también, se puede ajustar de acuerdo con un estado real. El decimoquinto valor preestablecido al decimonoveno valor preestablecido se pueden ajustar de acuerdo con un estado real, de modo que se pueden ajustar diferentes períodos de retención.

Del mismo modo, cuando el parámetro de dispersión general incluye un segundo ancho de banda mínimo y un tercer ancho de banda mínimo o el parámetro de dispersión general, incluye una primera proporción de la energía o el parámetro de dispersión general incluye una segunda proporción de energía y una tercera proporción de energía, la unidad 202 de determinación puede ajustar un correspondiente período de retención preestablecido, un correspondiente parámetro de actualización de retención y un parámetro relacionado utilizado para determinar el parámetro de actualización de retención, de modo que se puede determinar un correspondiente período de retención y se evita la conmutación frecuente entre métodos de codificación.

Cuando el método de codificación se determina de acuerdo con la dispersión de ráfaga (es decir, el método de codificación se determina de acuerdo con la dispersión global, la dispersión local y la ráfaga a corto plazo de distribución, en un espectro, de energía de una trama de audio), la unidad 202 de determinación puede ajustar un correspondiente período de retención, un correspondiente parámetro de actualización de retención y un parámetro relacionado utilizado para determinar el parámetro de actualización de retención, para evitar la conmutación frecuente entre métodos de codificación. En este caso, el período de retención puede ser menor que el período de retención ajustado en el caso del parámetro de dispersión general.

Cuando el método de codificación se determina de acuerdo con una característica limitada por banda de distribución de energía en un espectro, la unidad 202 de determinación puede ajustar un correspondiente período de retención, un correspondiente parámetro de actualización de retención y un parámetro relacionado utilizado para determinar el

parámetro de actualización de retención, para evitar la conmutación frecuente entre métodos de codificación. Por ejemplo, la unidad 202 de determinación puede calcular una proporción de energía de una envolvente espectral baja de una trama de audio de entrada con la energía de todas las envolventes espectrales y determinar el parámetro de actualización de retención de acuerdo con la proporción. Específicamente, la unidad 202 de determinación puede determinar la proporción de la energía de la envolvente espectral baja con la energía de todas las envolventes espectrales utilizando la siguiente fórmula:

$$R_{\text{baja}} = \frac{\sum_{k=0}^y s(k)}{\sum_{k=0}^{P-1} s(k)}$$

Fórmula 1.10

donde R_{baja} representa la proporción de la energía de la envolvente espectral baja con la energía de todas las envolventes espectrales, $s(k)$ representa la energía de la $k^{\text{ésima}}$ envolvente espectral, y representa un índice de una envolvente espectral más alta de una banda de frecuencia baja y P indica que la trama de audio está dividida en P envolventes espectrales en total. En este caso, si R_{baja} es mayor que un vigésimo valor preestablecido, el parámetro de actualización de retención es 0. Si R_{baja} es mayor que un vigesimoprimer valor preestablecido, el parámetro de actualización de retención puede tener un valor relativamente pequeño, donde el vigésimo valor preestablecido es mayor que el vigesimoprimer valor preestablecido. Si R_{baja} no es mayor que el vigesimoprimer valor preestablecido, el parámetro de retención puede tener un valor relativamente grande. Una persona experta en la técnica puede entender que, el vigésimo valor preestablecido y el vigesimoprimer valor preestablecido se pueden determinar de acuerdo con un experimento de simulación y el valor del parámetro de actualización de retención, también, se puede determinar de acuerdo con un experimento.

Además, cuando el método de codificación se determina de acuerdo con una característica limitada por banda de distribución de energía en un espectro, la unidad 202 de determinación puede determinar, además, la frecuencia de demarcación de una trama de audio de entrada y determinar el parámetro de actualización de retención de acuerdo con la frecuencia de demarcación, donde la frecuencia de demarcación puede ser diferente de una frecuencia de demarcación utilizada para determinar un parámetro de dispersión limitada por banda. Si la frecuencia de demarcación es menor que un vigesimosegundo valor preestablecido, la unidad 202 de determinación puede determinar que el parámetro de actualización de retención es 0. Si la frecuencia de demarcación es menor que un vigesimotercer valor preestablecido, la unidad 202 de determinación puede determinar que el parámetro de actualización de retención tiene un valor relativamente pequeño. Si la frecuencia de demarcación es mayor que el vigesimotercer valor preestablecido, la unidad 202 de determinación puede determinar que el parámetro de actualización de retención puede tener un valor relativamente grande. Una persona experta en la técnica puede entender que, el vigesimosegundo valor preestablecido y el vigesimotercer valor preestablecido se pueden determinar de acuerdo con un experimento de simulación y el valor del parámetro de actualización de retención, también, se puede determinar de acuerdo con un experimento.

La FIG. 3 es un diagrama de bloques estructural de un aparato de acuerdo con una realización de la presente invención. El aparato 300 mostrado en la FIG. 3 puede realizar los pasos en la FIG. 1. Como se muestra en la FIG. 3, el aparato 300 incluye un procesador 301 y una memoria 302.

Los componentes en el aparato 300 se acoplan utilizando un sistema 303 de bus. El sistema 303 de bus incluye además un bus de suministro de energía, un bus de control y un bus de señal de estado, además de un bus de datos. Sin embargo, para facilitar la descripción clara, todos los buses están marcados como el sistema 303 de bus en la FIG. 3.

El método descrito en las realizaciones anteriores de la presente invención puede aplicarse al procesador 301 o implementarse por el procesador 301. El procesador 301 puede ser un chip de circuito integrado y tiene una capacidad de procesamiento de señales. En un proceso de implementación, los pasos del método se pueden completar utilizando un circuito lógico integrado de hardware en el procesador 301 o una instrucción en una forma de software. El procesador 301 puede ser un procesador de propósito general, un procesador de señal digital (Digital Signal Processor, DSP), un circuito integrado de aplicación específica (Application Specific Integrated Circuit, ASIC), una matriz de puertas programables en campo (Field Programmable Gate Array, FPGA) u otro dispositivo lógico programable, una puerta discreta, o un dispositivo lógico de transistor, o un componente de hardware discreto. El procesador 301 puede implementar o ejecutar métodos, pasos y diagramas de bloques lógicos dados a conocer en las realizaciones de la presente invención. El procesador de propósito general puede ser un microprocesador o el procesador puede ser cualquier procesador común, y similares. Los pasos de los métodos dados a conocer con referencia a las realizaciones de la presente invención pueden ejecutarse y completarse directamente por medio de un procesador de decodificación de hardware, o pueden ejecutarse y completarse utilizando una combinación de módulos de hardware y de software en el procesador de decodificación. El módulo de software puede ubicarse en un

5 medio de almacenamiento que es maduro en la técnica, tal como una memoria de acceso aleatorio (Random Access Memory, RAM), una memoria flash, una memoria de solo lectura (Read-Only Memory, ROM), una memoria de solo lectura programable, o una memoria programable borrrable eléctricamente, o un registro. El medio de almacenamiento está ubicado en la memoria 302. El procesador 301 lee las instrucciones de la memoria 302 y completa los pasos del método en combinación con el hardware del mismo.

El procesador 301 está configurado para obtener N tramas de audio, donde las N tramas de audio incluyen una trama de audio actual y N es un número entero positivo.

El procesador 301 está configurado para determinar la dispersión de distribución, en el espectro, de energía de las N tramas de audio obtenidas por el procesador 301.

10 El procesador 301 está configurado además para determinar, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual, donde el primer método de codificación es un método de codificación que está basado en la transformada de tiempo-frecuencia y la cuantificación del coeficiente de transformada y que no está basado en la predicción lineal y, el segundo método de codificación, es un método de
15 codificación basado en predicción lineal.

De acuerdo con el aparato mostrado en la FIG. 3, cuando se codifica una trama de audio, se considera la dispersión de distribución, en un espectro, de energía de la trama de audio, lo que puede reducir la complejidad de codificación y asegurar que la codificación es de una precisión relativamente alta.

20 Durante la selección de un método de codificación apropiado para una trama de audio, se puede considerar la dispersión de distribución, en un espectro, de energía de la trama de audio. Puede haber tres tipos de dispersión de distribución, en un espectro, de energía de una trama de audio: dispersión general, dispersión de ráfaga y dispersión limitada por banda.

Opcionalmente, en una realización, se puede seleccionar un método de codificación apropiado para la trama de audio actual utilizando la dispersión general. En este caso, el procesador 301 está configurado específicamente para dividir un espectro de cada una de las N tramas de audio en P envolventes espectrales y determinar un parámetro de dispersión general de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde P es un número entero positivo y el parámetro de dispersión general indica la dispersión de
25 distribución, en el espectro, de la energía de las N tramas de audio.

Específicamente, un valor promedio de anchos de banda mínimos de distribución, en el espectro, de energía de proporción específica de N tramas de audio consecutivas de entrada, se puede definir como la dispersión general. Un ancho de banda menor indica una dispersión general más fuerte y un ancho de banda mayor indica una dispersión general más débil. En otras palabras, una dispersión general más fuerte indica que la energía de una trama de audio está más centralizada y una dispersión general más débil indica que la energía de una trama de audio está más dispersa. La eficiencia es alta cuando se utiliza el primer método de codificación para codificar una trama de audio cuya dispersión general es relativamente fuerte. Por lo tanto, se puede seleccionar un método de codificación apropiado, determinando la dispersión general de una trama de audio, para codificar la trama de audio. Para ayudar a determinar la dispersión general de una trama de audio, la dispersión general puede cuantificarse para obtener un parámetro de dispersión general. Opcionalmente, cuando N es 1, la dispersión general es un ancho de banda mínimo de distribución, en un espectro, de energía de proporción específica de la trama de audio actual.

30 Opcionalmente, en una realización, el parámetro de dispersión general incluye un primer ancho de banda mínimo. En este caso, el procesador 301 está configurado específicamente para determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la primera proporción de
35 energía preestablecida de las N tramas de audio es el primer ancho de banda mínimo. El procesador 301 está configurado específicamente para: cuando el primer ancho de banda mínimo es menor que un primer valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; y cuando el primer ancho de banda mínimo es mayor que el primer valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

40 Una persona experta en la técnica puede entender que, el primer valor preestablecido y la primera proporción preestablecida pueden determinarse de acuerdo con un experimento de simulación. Un primer valor preestablecido apropiado y una primera proporción preestablecida pueden determinarse por medio de un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación o el segundo método de codificación.

55 El procesador 301 está configurado específicamente para: ordenar la energía de las P envolventes espectrales de cada una de las tramas de audio en orden descendente; determinar, de acuerdo con la energía, ordenada en orden

descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio; y determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio, un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de las N tramas de audio. Por ejemplo, una señal de audio obtenida por el procesador 301 es una señal de banda ancha muestreada a 16 kHz y la señal de audio obtenida se obtiene en una trama de 30 ms. Cada una de las tramas de señal es de 330 puntos de muestreo en el dominio del tiempo. El procesador 301 puede realizar la transformada de tiempo-frecuencia en una señal de dominio del tiempo, por ejemplo, realizar la transformada de tiempo-frecuencia por medio de la transformada rápida de Fourier (Fast Fourier Transformation, FFT), para obtener 130 S(k) envolventes espectrales, es decir, 130 coeficientes de espectro de energía de FFT, donde k = 0, 1, 2, ..., 159. El procesador 301 puede encontrar un ancho de banda mínimo de las S(k) envolventes espectrales de una manera que una proporción de que representa la energía en el ancho de banda en la energía total de la trama es la primera proporción preestablecida. Específicamente, el procesador 301 puede acumular secuencialmente la energía de los intervalos de frecuencia en las S(k) envolventes espectrales en orden descendente; y comparar la energía obtenida después de cada acumulación con la energía total de la trama de audio y, si una proporción es mayor que la primera proporción preestablecida, finalizar el proceso de acumulación, donde una cantidad de veces de acumulación es el ancho de banda mínimo. Por ejemplo, la primera proporción preestablecida es del 90 % y, si una proporción que una suma de energía obtenida después de 30 veces de acumulación representa en la energía total supera el 90 %, se puede considerar que un ancho de banda mínimo de energía que representa no menos que la primera proporción preestablecida de la trama de audio es 30. El procesador 301 puede ejecutar el proceso de determinación de ancho de banda mínimo anterior para cada una de las N tramas de audio, para determinar por separado los anchos de banda mínimos de la energía que representa no menos que la primera proporción preestablecida de las N tramas de audio, incluyendo la trama de audio actual. El procesador 301 puede calcular un valor promedio de los anchos de banda mínimos de la energía que representa no menos que la primera proporción preestablecida de las N tramas de audio. El valor promedio de los anchos de banda mínimos de la energía que representa no menos que la primera proporción preestablecida de las N tramas de audio, puede referirse como el primer ancho de banda mínimo y, el primer ancho de banda mínimo, se puede utilizar como el parámetro de dispersión general. Cuando el primer ancho de banda mínimo es menor que el primer valor preestablecido, el procesador 301 puede determinar utilizar el primer método de codificación para codificar la trama de audio actual. Cuando el primer ancho de banda mínimo es mayor que el primer valor preestablecido, el procesador 301 puede determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

Opcionalmente, en otra realización, el parámetro de dispersión general puede incluir una primera proporción de energía. En este caso, el procesador 301 está configurado específicamente para seleccionar P₁ envolventes espectrales de las P envolventes espectrales de cada una de las N tramas de audio y determinar la primera proporción de energía de acuerdo con la energía de las P₁ envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio, donde P₁ es un número entero positivo menor que P. El procesador 301 está configurado específicamente para: cuando la primera proporción de energía es mayor que un segundo valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; y cuando la primera proporción de energía es menor que el segundo valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual. Opcionalmente, en una realización, cuando N es 1, las N tramas de audio son la trama de audio actual y el procesador 301 está configurado específicamente para determinar la primera proporción de energía de acuerdo con la energía de las P₁ envolventes espectrales de la trama de audio actual y la energía total de la trama de audio actual. El procesador 301 está configurado específicamente para determinar las P₁ envolventes espectrales de acuerdo con la energía de las P envolventes espectrales, donde la energía de una cualquiera de las P₁ envolventes espectrales es mayor que la energía de una cualquiera de las otras envolventes espectrales en las P envolventes espectrales excepto las P₁ envolventes espectrales.

Específicamente, el procesador 301 puede calcular la primera proporción de energía utilizando la siguiente fórmula:

$$\left\{ \begin{array}{l} R_1 = \frac{\sum_{n=1}^N r(n)}{N} \\ r(n) = \frac{E_{p1}(n)}{E_{all}(n)} \end{array} \right. \quad \text{Formula 1.6}$$

donde R₁ representa la primera proporción de energía, E_{p1}(n) representa una suma de energía de las P₁ envolventes espectrales seleccionadas en una n^{ésima} trama de audio, E_{todo}(n) representa la energía total de la n^{ésima} trama de audio y r(n) representa una proporción que la energía de las P₁ envolventes espectrales de la n^{ésima} trama de audio en las N tramas de audio representa en la energía total de la trama de audio.

Una persona experta en la técnica puede entender que, el segundo valor preestablecido y la selección de las P_1 envolventes espectrales se pueden determinar de acuerdo con un experimento de simulación. Un segundo valor preestablecido apropiado, un valor apropiado de P_1 y un método apropiado para seleccionar las P_1 envolventes espectrales se pueden determinar por medio de un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando una trama de audio que cumple la condición anterior se codifica utilizando el primer método de codificación o el segundo método de codificación. Opcionalmente, en una realización, las P_1 envolventes espectrales pueden ser P_1 envolventes espectrales que tienen energía máxima en las P envolventes espectrales.

Por ejemplo, una señal de audio obtenida por el procesador 301 es una señal de banda ancha muestreada a 16 kHz y la señal de audio obtenida se obtiene en una trama de 30 ms. Cada una de las tramas de señal es de 330 puntos de muestreo en el dominio del tiempo. El procesador 301 puede realizar la transformada de tiempo-frecuencia en una señal de dominio del tiempo, por ejemplo, realizar una la transformada de tiempo-frecuencia por medio de la transformada rápida de Fourier, para obtener 130 $S(k)$ envolventes espectrales, donde $k = 0, 1, 2, \dots, 159$. El procesador 301 puede seleccionar P_1 envolventes espectrales de las 130 envolventes espectrales y calcular una proporción que una suma de energía de las P_1 envolventes espectrales representa en la energía total de la trama de audio. El procesador 301 puede ejecutar el proceso anterior para cada una de las N tramas de audio, es decir, calcular una proporción que una suma de energía de las P_1 envolventes espectrales de cada una de las N tramas de audio representa en la respectiva energía total. El procesador 301 puede calcular un valor promedio de las proporciones. El valor promedio de las proporciones es la primera proporción de energía. Cuando la primera proporción de energía es mayor que el segundo valor preestablecido, el procesador 301 puede determinar utilizar el primer método de codificación para codificar la trama de audio actual. Cuando la primera proporción de energía es menor que el segundo valor preestablecido, el procesador 301 puede determinar utilizar el segundo método de codificación para codificar la trama de audio actual. Las P_1 envolventes espectrales pueden ser P_1 envolventes espectrales que tienen energía máxima en las P envolventes espectrales. Es decir, el procesador 301 está configurado específicamente para determinar, a partir de las P envolventes espectrales de cada una de las N tramas de audio, P_1 envolventes espectrales que tienen energía máxima. Opcionalmente, en una realización, el valor de P_1 puede ser 30.

Opcionalmente, en otra realización, el parámetro de dispersión general puede incluir un segundo ancho de banda mínimo y un tercer ancho de banda mínimo. En este caso, el procesador 301 está configurado específicamente para determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la segunda proporción de energía preestablecida de las N tramas de audio y determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la tercera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de las P envolventes espectrales de cada una de las N tramas de audio, donde el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la segunda proporción de energía preestablecida de las N tramas de audio se utiliza como el segundo ancho de banda mínimo, el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la tercera proporción de energía preestablecida de las N tramas de audio se utiliza como el tercer ancho de banda mínimo y la segunda proporción preestablecida es menor que la tercera proporción preestablecida. El procesador 301 está configurado específicamente para: cuando el segundo ancho de banda mínimo es menor que un tercer valor preestablecido y el tercer ancho de banda mínimo es menor que un cuarto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; cuando el tercer ancho de banda mínimo es menor que un quinto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; y cuando el tercer ancho de banda mínimo es mayor que un sexto valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual. Opcionalmente, en una realización, cuando N es 1, las N tramas de audio son la trama de audio actual. El procesador 301 puede determinar un ancho de banda mínimo de distribución, en el espectro, de la segunda proporción de energía preestablecida de la trama de audio actual como el segundo ancho de banda mínimo. El procesador 301 puede determinar un ancho de banda mínimo de distribución, en el espectro, de la tercera proporción de energía preestablecida de la trama de audio actual como el tercer ancho de banda mínimo.

Una persona experta en la técnica puede entender que, el tercer valor preestablecido, el cuarto valor preestablecido, el quinto valor preestablecido, el sexto valor preestablecido, la segunda proporción preestablecida y la tercera proporción preestablecida se pueden determinar de acuerdo con un experimento de simulación. Los valores preestablecidos apropiados y las proporciones preestablecidas pueden determinarse por medio de un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación o el segundo método de codificación.

El procesador 301 está configurado específicamente para: ordenar la energía de las P envolventes espectrales de cada una de las tramas de audio en orden descendente; determinar, de acuerdo con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la segunda proporción preestablecida de cada una de las N tramas de audio; determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la segunda proporción preestablecida de cada una de las N tramas de audio, un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la segunda proporción preestablecida de las N tramas de audio; determinar, de acuerdo

con la energía, ordenada en orden descendente, de las P envolventes espectrales de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la tercera proporción preestablecida de cada una de las N tramas de audio; y determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de cada una de las N tramas de audio, un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la tercera proporción preestablecida de las N tramas de audio. Por ejemplo, una señal de audio obtenida por el procesador 301 es una señal de banda ancha muestreada a 16 kHz y la señal de audio obtenida se obtiene en una trama de 30 ms. Cada una de las tramas de señal es de 330 puntos de muestreo en el dominio del tiempo. El procesador 301 puede realizar la transformada de tiempo-frecuencia en una señal de dominio del tiempo, por ejemplo, realizar una transformada de tiempo-frecuencia por medio de la transformada rápida de Fourier, para obtener 130 $S(k)$ envolventes espectrales, donde $k = 0, 1, 2, \dots, 159$. El procesador 301 puede encontrar un ancho de banda mínimo de las $S(k)$ envolventes espectrales de manera que una proporción que la energía en el ancho de banda representa en la energía total de la trama no sea menor que la segunda proporción preestablecida. El procesador 301 puede continuar encontrando un ancho de banda de las $S(k)$ envolventes espectrales de manera que una proporción que la energía en el ancho de banda representa en la energía total no sea menor que la tercera proporción preestablecida. Específicamente, el procesador 301 puede acumular secuencialmente energía de intervalos de frecuencia en las $S(k)$ envolventes espectrales en orden descendente. La energía obtenida después de cada acumulación se compara con la energía total de la trama de audio y, si una proporción es mayor que la segunda proporción preestablecida, una cantidad de veces de acumulación es un ancho de banda mínimo que no es menor que la segunda proporción preestablecida. El procesador 301 puede continuar la acumulación. Si una proporción de energía obtenida después de la acumulación con la energía total de la trama de audio es mayor que la tercera proporción preestablecida, la acumulación se finaliza y, una cantidad de veces de acumulación es un ancho de banda mínimo que no es menor que la tercera proporción preestablecida. Por ejemplo, la segunda proporción preestablecida es del 85 % y la tercera proporción preestablecida es del 95 %. Si una proporción que una suma de energía obtenida después de 30 veces de acumulación representa en el total de energía excede el 85 %, se puede considerar que el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la segunda proporción preestablecida de la trama de audio es 30. La acumulación continúa y, si una proporción que una suma de energía obtenida después de 35 veces de acumulación representa en la energía total es del 95 %, se puede considerar que el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de la trama de audio es 35. El procesador 301 puede ejecutar el proceso anterior para cada una de las N tramas de audio. El procesador 301 puede determinar por separado los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la segunda proporción preestablecida de las N tramas de audio, incluyendo la trama de audio actual y los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de las N tramas de audio, incluyendo la trama de audio actual. El valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la segunda proporción preestablecida de las N tramas de audio es el segundo ancho de banda mínimo. El valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la energía que representa no menos que la tercera proporción preestablecida de las N tramas de audio es el tercer ancho de banda mínimo. Cuando el segundo ancho de banda mínimo es menor que el tercer valor preestablecido y el tercer ancho de banda mínimo es menor que el cuarto valor preestablecido, el procesador 301 puede determinar utilizar el primer método de codificación para codificar la trama de audio actual. Cuando el tercer ancho de banda mínimo es menor que el quinto valor preestablecido, el procesador 301 puede determinar utilizar el primer método de codificación para codificar la trama de audio actual. Cuando el tercer ancho de banda mínimo es mayor que el sexto valor preestablecido, el procesador 301 puede determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

Opcionalmente, en otra realización, el parámetro de dispersión general incluye una segunda proporción de energía y una tercera proporción de energía. En este caso, el procesador 301 está configurado específicamente para: seleccionar P_2 envolventes espectrales de las P envolventes espectrales de cada una de las N tramas de audio, determinar la segunda proporción de energía de acuerdo con la energía de las P_2 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio, seleccionar P_3 envolventes espectrales de las P envolventes espectrales de cada una de las N tramas de audio y determinar la tercera proporción de energía de acuerdo con la energía de las P_3 envolventes espectrales de cada una de las N tramas de audio y la energía total de las respectivas N tramas de audio, donde P_2 y P_3 son números enteros positivos menores que P y P_2 es menor que P_3 . El procesador 301 está configurado específicamente para: cuando la segunda proporción de energía es mayor que un séptimo valor preestablecido y la tercera proporción de energía es mayor que un octavo valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; cuando la segunda proporción de energía es mayor que un noveno valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; y cuando la tercera proporción de energía es menor que un décimo valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual. Opcionalmente, en una realización, cuando N es 1, las N tramas de audio son la trama de audio actual. El procesador 301 puede determinar la segunda proporción de energía de acuerdo con la energía de las P_2 envolventes espectrales de la trama de audio actual y la energía total de la trama de audio actual. El procesador 301 puede determinar la tercera proporción de energía de acuerdo con la energía de las P_3 envolventes espectrales de la trama de audio actual y la energía total de la trama de audio actual.

Una persona experta en la técnica puede entender que, los valores de P_2 y P_3 , el séptimo valor preestablecido, el octavo valor preestablecido, el noveno valor preestablecido y el décimo valor preestablecido se pueden determinar de acuerdo con un experimento de simulación. Los valores preestablecidos apropiados pueden determinarse por medio de un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación o el segundo método de codificación. Opcionalmente, en una realización, el procesador 301 está configurado específicamente para determinar, a partir de las P envolventes espectrales de cada una de las N tramas de audio, P_2 envolventes espectrales que tienen energía máxima y determinar, a partir de las P envolventes espectrales de cada una de las N tramas de audio, las P_3 envolventes espectrales que tienen energía máxima.

Por ejemplo, una señal de audio obtenida por el procesador 301 es una señal de banda ancha muestreada a 16 kHz y la señal de audio obtenida se obtiene en una trama de 30 ms. Cada una de las tramas de señal es de 330 puntos de muestreo en el dominio del tiempo. El procesador 301 puede realizar una transformada de tiempo-frecuencia en una señal de dominio del tiempo, por ejemplo, realizar una transformada de tiempo-frecuencia por medio de la transformada rápida de Fourier, para obtener 130 $S(k)$ envolventes espectrales, donde $k = 0, 1, 2, \dots, 159$. El procesador 301 puede seleccionar P_2 envolventes espectrales de las 130 envolventes espectrales y calcular una proporción que una suma de energía de las P_2 envolventes espectrales representa en la energía total de la trama de audio. El procesador 301 puede ejecutar el proceso anterior para cada una de las N tramas de audio, es decir, calcular una proporción que una suma de energía de las P_2 envolventes espectrales de cada una de las N tramas de audio representa en la respectiva energía total. El procesador 301 puede calcular un valor promedio de las proporciones. El valor promedio de las proporciones es la segunda proporción de energía. El procesador 301 puede seleccionar P_3 envolventes espectrales de las 130 envolventes espectrales y calcular una proporción que una suma de energía de las P_3 envolventes espectrales representa en la energía total de la trama de audio. El procesador 301 puede ejecutar el proceso anterior para cada una de las N tramas de audio, es decir, calcular una proporción que una suma de energía de las P_3 envolventes espectrales de cada una de las N tramas de audio representa en la respectiva energía total. El procesador 301 puede calcular un valor promedio de las proporciones. El valor promedio de las proporciones es la tercera proporción de energía. Cuando la segunda proporción de energía es mayor que el séptimo valor preestablecido y la tercera proporción de energía es mayor que el octavo valor preestablecido, el procesador 301 puede determinar utilizar el primer método de codificación para codificar la trama de audio actual. Cuando la segunda proporción de energía es mayor que el noveno valor preestablecido, el procesador 301 puede determinar utilizar el primer método de codificación para codificar la trama de audio actual. Cuando la tercera proporción de energía es menor que el décimo valor preestablecido, el procesador 301 puede determinar utilizar el segundo método de codificación para codificar la trama de audio actual. Las P_2 envolventes espectrales pueden ser P_2 envolventes espectrales que tienen energía máxima en las P envolventes espectrales; y las P_3 envolventes espectrales pueden ser P_3 envolventes espectrales que tienen energía máxima en las P envolventes espectrales. Opcionalmente, en una realización, el valor de P_2 puede ser 30 y el valor de P_3 puede ser 30.

Opcionalmente, en otra realización, se puede seleccionar un método de codificación apropiado para la trama de audio actual utilizando la dispersión de ráfaga. Para la dispersión de ráfaga, se debe considerar la dispersión global, la dispersión local y la ráfaga a corto plazo de distribución, en un espectro, de energía de una trama de audio. En este caso, la dispersión de distribución de la energía en el espectro puede incluir la dispersión global, la dispersión local y la ráfaga a corto plazo de distribución de la energía en el espectro. En este caso, un valor de N puede ser 1 y las N tramas de audio son la trama de audio actual. El procesador 301 está configurado específicamente para dividir un espectro de la trama de audio actual en Q subbandas y determinar un parámetro de dispersión de ráfaga de acuerdo con la energía máxima de cada una de las Q subbandas del espectro de la trama de audio actual, donde el parámetro de dispersión de ráfaga se utiliza para indicar la dispersión global, la dispersión local y la ráfaga a corto plazo de la trama de audio actual.

Específicamente, el procesador 301 está configurado específicamente para determinar una proporción global de máximo a promedio de cada una de las Q subbandas, una proporción local de máximo a promedio de cada una de las Q subbandas y una fluctuación de energía a corto plazo de cada una de las Q subbandas, donde la proporción global de máximo a promedio se determina por el procesador 301 de acuerdo con la energía máxima en la subbanda y la energía promedio de todas las subbandas de la trama de audio actual, la proporción local de máximo a promedio se determina por el procesador 301 de acuerdo con la energía máxima en la subbanda y la energía promedio en la subbanda y, la fluctuación de energía máxima a corto plazo, se determina de acuerdo con la energía máxima en la subbanda y la energía máxima en una banda de frecuencia específica de una trama de audio antes de la trama de audio. La proporción global de máximo a promedio de cada una de las Q subbandas, la proporción local de máximo a promedio de cada una de las Q subbandas y la fluctuación de energía a corto plazo de cada una de las Q subbandas, respectivamente, representan la dispersión global, la dispersión local y la ráfaga a corto plazo. El procesador 301 está configurado específicamente para: determinar si hay una primera subbanda en las Q subbandas, donde una proporción local de máximo a promedio de la primera subbanda es mayor que un decimoprimer valor preestablecido, una proporción global de máximo a promedio de la primera subbanda es mayor que un decimosegundo valor preestablecido y, una fluctuación de energía máxima a corto plazo de la primera subbanda, es mayor que un decimotercer valor preestablecido; y cuando la primera subbanda está en las Q subbandas, determinar utilizar el primer método de codificación para codificar la trama de audio actual.

Específicamente, el procesador 301 puede calcular la proporción global de máximo a promedio utilizando la siguiente fórmula:

$$p2s(i) = e(i) / \left(\frac{1}{P} * \sum_{k=0}^{P-1} s(k) \right) \quad \text{Formula 1.7}$$

5 donde e(i) representa la energía máxima de una $i^{\text{ésima}}$ subbanda en las Q subbandas, s(k) representa energía de una $k^{\text{ésima}}$ envolvente espectral en las P envolventes espectrales y p2s(i) representa una proporción global de máximo a promedio de la $i^{\text{ésima}}$ subbanda.

El procesador 301 puede calcular la proporción local de máximo a promedio utilizando la siguiente fórmula:

$$p2a(i) = e(i) / \left(\frac{1}{h(i) - l(i) + 1} * \sum_{k=l(i)}^{h(i)} s(k) \right) \quad \text{Formula 1.8}$$

10 donde e(i) representa la energía máxima de la $i^{\text{ésima}}$ subbanda en las Q subbandas, s(k) representa la energía de la $k^{\text{ésima}}$ envolvente espectral en las P envolventes espectrales, h(i) representa un índice de una envolvente espectral que está incluida en la $i^{\text{ésima}}$ subbanda y que tiene una frecuencia más alta, l(i) representa un índice de una envolvente espectral que está incluida en la $i^{\text{ésima}}$ subbanda y que tiene una frecuencia más baja, p2a(i) representa una proporción local de máximo a promedio de la $i^{\text{ésima}}$ subbanda y h(i) es menor o igual que P - 1.

El procesador 301 puede calcular la fluctuación de energía máxima a corto plazo utilizando la siguiente fórmula:

$$15 \quad dev(i) = (2 * e(i)) / (e_1 + e_2) \quad \text{Formula 1.9}$$

20 donde e(i) representa la energía máxima de la $i^{\text{ésima}}$ subbanda en las Q subbandas de la trama de audio actual y e_1 y e_2 representan la energía máxima de bandas de frecuencia específicas de tramas de audio antes de la trama de audio actual. Específicamente, suponiendo que la trama de audio actual es una $M^{\text{ésima}}$ trama de audio, se determina una envolvente espectral en la que se encuentra la energía máxima de la $i^{\text{ésima}}$ subbanda de la trama de audio actual. Se supone que la envolvente espectral en la que se encuentra la energía máxima es i_1 . Se determina la energía máxima dentro de un rango de una $(i_1 - t)^{\text{ésima}}$ envolvente espectral a una $(i_1 + t)^{\text{ésima}}$ envolvente espectral en una $(M - 1)^{\text{ésima}}$ trama de audio y la energía máxima es e_1 . Del mismo modo, se determina la energía máxima dentro de un rango de una $(i_1 - t)^{\text{ésima}}$ envolvente espectral a una $(i_1 + t)^{\text{ésima}}$ envolvente espectral en una $(M - 2)^{\text{ésima}}$ trama de audio y la energía máxima es e_2 .

25 Una persona experta en la técnica puede entender que, el decimoprimer valor preestablecido, el decimosegundo valor preestablecido y el decimotercer valor preestablecido se pueden determinar de acuerdo con un experimento de simulación. Los valores preestablecidos apropiados pueden determinarse por medio de un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación.

30 Opcionalmente, en otra realización, se puede seleccionar un método de codificación apropiado para la trama de audio actual utilizando la dispersión limitada por banda. En este caso, la dispersión de distribución de la energía en el espectro incluye la dispersión de distribución limitada por banda de la energía en el espectro. En este caso, el procesador 301 está configurado específicamente para determinar una frecuencia de demarcación de cada una de las N tramas de audio. El procesador 301 está configurado específicamente para determinar un parámetro de dispersión limitada por banda de acuerdo con la frecuencia de demarcación de cada una de las N tramas de audio.

35 Una persona experta en la técnica puede entender que, la cuarta proporción preestablecida y el decimocuarto valor preestablecido pueden determinarse de acuerdo con un experimento de simulación. Se puede determinar un valor preestablecido apropiado y una proporción preestablecida de acuerdo con un experimento de simulación, de modo que se puede obtener un buen efecto de codificación cuando se codifica una trama de audio que cumple la condición anterior utilizando el primer método de codificación.

40 Por ejemplo, el procesador 301 puede determinar la energía de cada una de las P envolventes espectrales de la trama de audio actual y buscar una frecuencia de demarcación de una frecuencia baja a una frecuencia alta de manera que una proporción de que la energía que es menor que la frecuencia de demarcación representa en la energía total de la trama de audio actual, es la cuarta proporción preestablecida. El parámetro de dispersión limitada por banda puede ser un valor promedio de las frecuencias de demarcación de las N tramas de audio. En este caso, el procesador 301 está configurado específicamente para: cuando se determina que el parámetro de dispersión limitada por banda de las tramas de audio es menor que un decimocuarto valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual. Suponiendo que N es 1, la frecuencia de demarcación de la trama de audio actual es el parámetro de dispersión limitada por banda. Suponiendo que N es un número entero mayor que 1, el procesador 301 puede determinar que el valor promedio de las frecuencias de

demarcación de las N tramas de audio es el parámetro de dispersión limitada por banda. Una persona experta en la técnica puede entender que la determinación de la frecuencia de demarcación mencionada anteriormente es simplemente un ejemplo. Alternativamente, el método de determinación de la frecuencia de demarcación puede ser buscar una frecuencia de demarcación de una frecuencia alta a una frecuencia baja o puede ser otro método.

5 Además, para evitar una conmutación frecuente entre el primer método de codificación y el segundo método de codificación, el procesador 301 puede estar configurado además para ajustar un período de retención. El procesador 301 puede estar configurado para: para una trama de audio en el período de retención, utilizar un método de codificación utilizado para una trama de audio en una posición de inicio del período de retención. De esta manera, se puede evitar una disminución de la calidad de conmutación causada por la conmutación frecuente entre diferentes métodos de codificación.

10 Si una duración de retención del período de retención es L, el procesador 301 puede estar configurado para determinar que las L tramas de audio después de la trama de audio actual pertenecen a un período de retención de la trama de audio actual. Si la dispersión de distribución, en un espectro, de energía de una trama de audio que pertenece al período de retención es diferente de la dispersión de distribución, en un espectro, de energía de una trama de audio en una posición de inicio del período de retención, el procesador 301 puede estar configurado para determinar que la trama de audio todavía está codificada utilizando un método de codificación que es el mismo que el que se utiliza para la trama de audio en la posición de inicio del período de retención.

15 La duración del período de retención puede actualizarse de acuerdo con la dispersión de distribución, en un espectro, de energía de una trama de audio en el período de retención, hasta que la duración del período de retención sea 0.

20 Por ejemplo, si el procesador 301 determina utilizar el primer método de codificación para una $(l+1)$ ésima trama de audio y una duración de un período de retención preestablecido es L, el procesador 301 puede determinar que el primer método de codificación se utiliza para una $(l+1)$ ésima trama de audio a una $(l+L)$ ésima trama de audio. Entonces, el procesador 301 puede determinar la dispersión de distribución, en un espectro, de energía de la $(l+1)$ ésima trama de audio y volver a calcular el período de retención de acuerdo con la dispersión de distribución, en el espectro, de la energía de la $(l+1)$ ésima trama de audio. Si la $(l+1)$ ésima trama de audio todavía cumple una condición para utilizar el primer método de codificación, el procesador 301 puede determinar que un período de retención posterior es todavía el período L de retención preestablecido. Es decir, el período de retención comienza a partir de una $(l+2)$ ésima trama de audio a una $(l+1+L)$ ésima trama de audio. Si la $(l+1)$ ésima trama de audio no cumple la condición de utilizar el primer método de codificación, el procesador 301 puede volver a determinar el período de retención de acuerdo con la dispersión de distribución, en el espectro, de la energía de la $(l+1)$ ésima trama de audio. Por ejemplo, el procesador 301 puede volver a determinar que el período de retención es $L - L_1$, donde L_1 es un número entero positivo menor o igual que L. Si L_1 es igual a L, la duración del período de retención se actualiza a 0. En este caso, el procesador 301 puede volver a determinar el método de codificación de acuerdo con la dispersión de distribución, en el espectro, de la energía de la $(l+1)$ ésima trama de audio. Si L_1 es un número entero menor que L, el procesador 301 puede volver a determinar el método de codificación de acuerdo con la dispersión de distribución, en un espectro, de energía de una $(l+1+L-L_1)$ ésima trama de audio. Sin embargo, debido a que la $(l+1)$ ésima trama de audio está en un período de retención de la $(l+1)$ ésima trama de audio, la $(l+1)$ ésima trama de audio todavía se codifica utilizando el primer método de codificación. L_1 se puede referir como un parámetro de actualización de retención y, un valor del parámetro de actualización de retención, se puede determinar de acuerdo con la dispersión de distribución, en un espectro, de energía de una trama de audio de entrada. De esta manera, la actualización del período de retención está relacionada con la dispersión de distribución, en un espectro, de energía de una trama de audio.

25 Por ejemplo, cuando se determina un parámetro de dispersión general y, el parámetro de dispersión general, es un primer ancho de banda mínimo, el procesador 301 puede volver a determinar el período de retención de acuerdo con un ancho de banda mínimo de distribución, en un espectro, de la primera proporción de energía preestablecida de una trama de audio. Se supone que se determina utilizar el primer método de codificación para codificar la $(l+1)$ ésima trama de audio y un período de retención preestablecido es L. El procesador 301 puede determinar un ancho de banda mínimo de distribución, en un espectro, de la primera proporción de energía preestablecida de cada una de las H tramas de audio consecutivas, incluyendo la $(l+1)$ ésima trama de audio, donde H es un número entero positivo mayor que 0. Si la $(l+1)$ ésima trama de audio no cumple la condición de utilizar el primer método de codificación, el procesador 301 puede determinar una cantidad de tramas de audio cuyos anchos de banda mínimos de distribución, en un espectro, de la primera proporción de energía preestablecida sean menores que un decimoquinto valor preestablecido (la cantidad se refiere brevemente como un primer parámetro de retención). Cuando un ancho de banda mínimo de distribución, en un espectro, de la primera proporción de energía preestablecida de una $(l+1)$ ésima trama de audio es mayor que un decimosexto valor preestablecido y menor que un decimoséptimo valor preestablecido y, el primer parámetro de retención es menor que un decimoctavo valor preestablecido, el procesador 301 puede restar 1 a la duración del período de retención, es decir, el parámetro de actualización de retención es 1. El decimosexto valor preestablecido es mayor que el primer valor preestablecido. Cuando el ancho de banda mínimo de distribución, en el espectro, de la primera proporción de energía preestablecida de la $(l+1)$ ésima trama de audio es mayor que el decimoséptimo valor preestablecido y menor que un decimonoveno valor preestablecido y, el primer parámetro de retención es menor que el decimoctavo valor preestablecido, el procesador 301 puede restar 2 a la

duración del período de retención, es decir, el parámetro de actualización de retención es 2. Cuando el ancho de banda mínimo de distribución, en el espectro, de la primera proporción de energía predefinida de la $(L + 1)^{ésima}$ trama de audio es mayor que el decimonoveno valor preestablecido, el procesador 301 puede ajustar el período de retención a 0. Cuando el primer parámetro de retención y el ancho de banda mínimo de distribución, en el espectro, de la primera proporción de energía preestablecida de la $(L + 1)^{ésima}$ trama de audio no cumple con uno o más del decimosexto valor preestablecido al decimonoveno valor preestablecido, el procesador 301 puede determinar que el período de retención permanece sin cambios.

Una persona experta en la técnica puede entender que, el período de retención preestablecido puede ajustarse de acuerdo con un estado real y el parámetro de actualización de retención, también, puede ajustarse de acuerdo con un estado real. El decimoquinto valor preestablecido al decimonoveno valor preestablecido pueden ajustarse de acuerdo con un estado real, de modo que se pueden ajustar diferentes períodos de retención.

Del mismo modo, cuando el parámetro de dispersión general incluye un segundo ancho de banda mínimo y un tercer ancho de banda mínimo o, el parámetro de dispersión general incluye una primera proporción de la energía o el parámetro de dispersión general incluye una segunda proporción de energía y una tercera proporción de energía, el procesador 301 puede ajustar un correspondiente período de retención preestablecido, un correspondiente parámetro de actualización de retención y un parámetro relacionado utilizado para determinar el parámetro de actualización de retención, de modo que se puede determinar un período de retención correspondiente y se evita la conmutación frecuente entre métodos de codificación.

Cuando el método de codificación se determina de acuerdo con la dispersión de ráfaga (es decir, el método de codificación se determina de acuerdo con la dispersión global, la dispersión local y la ráfaga a corto plazo de distribución, en un espectro, de energía de una trama de audio), el procesador 301 puede ajustar un correspondiente período de retención, un correspondiente parámetro de actualización de retención y un parámetro relacionado utilizado para determinar el parámetro de actualización de retención, para evitar la conmutación frecuente entre métodos de codificación. En este caso, el período de retención puede ser menor que el período de retención ajustado en el caso del parámetro de dispersión general.

Cuando el método de codificación se determina de acuerdo con una característica limitada por banda de distribución de energía en un espectro, el procesador 301 puede ajustar un correspondiente período de retención, un correspondiente parámetro de actualización de retención y un parámetro relacionado utilizado para determinar el parámetro de actualización de retención, para evitar la conmutación frecuente entre métodos de codificación. Por ejemplo, el procesador 301 puede calcular una proporción de energía de una envolvente espectral baja de una trama de audio de entrada con la energía de todas las envolventes espectrales y determinar el parámetro de actualización de retención de acuerdo con la proporción. Específicamente, el procesador 301 puede determinar la proporción de la energía de la envolvente espectral baja con la energía de todas las envolventes espectrales utilizando la siguiente fórmula:

$$R_{low} = \frac{\sum_{k=0}^y s(k)}{\sum_{k=0}^{P-1} s(k)} \quad \text{Formula 1.10}$$

donde R_{baja} representa la proporción de la energía de la envolvente espectral baja con la energía de todas las envolventes espectrales, $s(k)$ representa la energía de la $k^{ésima}$ envolvente espectral, y representa un índice de la envolvente espectral más alta de una banda de frecuencia baja y P indica que la trama de audio está dividida en P envolventes espectrales en total. En este caso, si R_{baja} es mayor que un vigésimo valor preestablecido, el parámetro de actualización de retención puede tener un valor relativamente pequeño, donde el vigésimo valor preestablecido es mayor que el vigesimoprimer valor preestablecido. Si R_{baja} no es mayor que el vigesimoprimer valor preestablecido, el parámetro de retención puede tener un valor relativamente grande. Una persona experta en la técnica puede entender que, el vigésimo valor preestablecido y el vigesimoprimer valor preestablecido se pueden determinar de acuerdo con un experimento de simulación y el valor del parámetro de actualización de retención, también, se puede determinar de acuerdo con un experimento.

Además, cuando el método de codificación se determina de acuerdo con una característica limitada por banda de distribución de energía en un espectro, el procesador 301 puede determinar además una frecuencia de demarcación de una trama de audio de entrada y determinar el parámetro de actualización de retención de acuerdo con la frecuencia de demarcación, donde la frecuencia de demarcación puede ser diferente de la frecuencia de demarcación utilizada para determinar un parámetro de dispersión limitada por banda. Si la frecuencia de demarcación es menor que un vigesimosegundo valor preestablecido, el procesador 301 puede determinar que el parámetro de actualización de retención es 0. Si la frecuencia de demarcación es menor que un vigesimotercer valor preestablecido, el procesador 301 puede determinar que el parámetro de actualización de retención tiene un valor relativamente pequeño. Si la frecuencia de demarcación es mayor que el vigesimotercer valor preestablecido, el

procesador 301 puede determinar que el parámetro de actualización de retención puede tener un valor relativamente grande. Una persona experta en la técnica puede entender que, el vigesimosegundo valor preestablecido y el vigesimotercer valor preestablecido se pueden determinar de acuerdo con un experimento de simulación y el valor del parámetro de actualización de retención, también, se puede determinar de acuerdo con un experimento.

5 Una persona con experiencia ordinaria en la técnica puede ser consciente de que, en combinación con los ejemplos descritos en las realizaciones dadas a conocer en esta memoria descriptiva, las unidades y los pasos de algoritmo pueden implementarse mediante hardware electrónico o una combinación de software informático y hardware electrónico. Si las funciones se realizan por hardware o software depende de aplicaciones particulares y condiciones de restricción de diseño de las soluciones técnicas. Una persona experta en la técnica puede utilizar diferentes
10 métodos para implementar las funciones descritas para cada aplicación particular, pero no debe considerarse que la implementación va más allá del alcance de la presente invención.

Puede ser claramente entendido por una persona experta en la técnica que, para el propósito de la descripción conveniente y breve, para un proceso de trabajo detallado del sistema anterior, el aparato y la unidad, se puede hacer referencia a un correspondiente proceso en las realizaciones del método anterior y los detalles no se describen en el presente documento.
15

En las diversas realizaciones proporcionadas en la presente solicitud, debe entenderse que el sistema, el aparato y el método dados a conocer pueden implementarse de otras maneras. Por ejemplo, la realización del aparato descrito es meramente ejemplar. Por ejemplo, la división de unidades es simplemente una división de funciones lógicas y puede ser otra división en la implementación real. Por ejemplo, una pluralidad de unidades o componentes pueden
20 combinarse o integrarse en otro sistema, o algunas características pueden ignorarse o no realizarse. Además, los acoplamientos mutuos mostrados o discutidos o acoplamientos directos o conexiones de comunicación, pueden implementarse a través de algunas interfaces. Los acoplamientos indirectos o las conexiones de comunicación entre los aparatos o unidades pueden implementarse en forma electrónica, mecánica u otras formas.

Las unidades descritas como partes separadas pueden o pueden no estar físicamente separadas y las partes que se muestran como unidades pueden o pueden no ser unidades físicas, pueden estar ubicadas en una posición, o se pueden distribuir en una pluralidad de unidades de red. Se puede seleccionar una parte o la totalidad de las unidades de acuerdo con las necesidades reales para lograr los objetivos de las soluciones de las realizaciones.
25

Además, las unidades funcionales en las realizaciones de la presente invención pueden integrarse en una unidad de procesamiento, o cada una de las unidades puede existir sola físicamente, o dos o más unidades están integradas en una unidad.
30

Cuando las funciones se implementan en forma de una unidad funcional de software y se venden o utilizan como un producto independiente, las funciones pueden almacenarse en un medio de almacenamiento legible por computadora. En base a tal entendimiento, las soluciones técnicas de la presente invención, esencialmente, o la parte que contribuye a la técnica anterior, o una parte de las soluciones técnicas, pueden implementarse en forma
35 de un producto de software. El producto de software se almacena en un medio de almacenamiento e incluye varias instrucciones para instruir a un dispositivo de computadora (que puede ser una computadora personal, un servidor o un dispositivo de red) o un procesador para realizar la totalidad o parte de los pasos de los métodos descritos en las realizaciones de la presente invención. El medio de almacenamiento anterior incluye: cualquier medio que pueda almacenar código de programa, tal como una unidad flash USB, un disco duro extraíble, una memoria de solo lectura (ROM, Read-Only Memory), una memoria de acceso aleatorio (RAM, Random Access Memory), un disco magnético, o un disco óptico.
40

Las descripciones anteriores son meramente realizaciones específicas de la presente invención, pero no pretenden limitar la presente invención.

REIVINDICACIONES

1. Un método de codificación de audio, en donde el método comprende:

determinar (101) la dispersión de distribución, en un espectro, de energía de N tramas de audio de entrada, en donde la dispersión de distribución se determina para cada una de las N tramas de audio de entrada, en donde las N tramas de audio comprenden una trama de audio actual y N es un número entero positivo; y

determinar (102), de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual, en donde el primer método de codificación es un método de codificación que está basado en la transformada de tiempo-frecuencia y la cuantificación del coeficiente de transformada y que no está basado en la predicción lineal y, el segundo método de codificación, es un método de codificación basado en predicción lineal;

en donde determinar (101) la dispersión de distribución, en el espectro, de energía de N tramas de audio de entrada comprende:

dividir un espectro de cada una de las N tramas de audio en P coeficientes de espectro de energía de FFT, donde P es un número entero positivo; y

determinar un parámetro de dispersión general de acuerdo con la energía de los P coeficientes de espectro de energía de FFT de cada una de las N tramas de audio, en donde el parámetro de dispersión general indica la dispersión de distribución, en el espectro, de la energía de las N tramas de audio;

en donde el parámetro de dispersión general comprende un primer ancho de banda mínimo;

la determinación de un parámetro de dispersión general de acuerdo con la energía de los P coeficientes de espectro de energía de FFT de cada una de las N tramas de audio comprende:

determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de los P coeficientes de espectro de energía de FFT de cada una de las N tramas de audio, en donde un ancho de banda mínimo se encuentra a partir de los P coeficientes de espectro de energía de FFT de manera que una proporción que la energía en el ancho de banda representa en la energía total de una trama es la primera proporción preestablecida, en donde el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio es el primer ancho de banda mínimo; y

determinar (102), de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual comprende:

cuando el primer ancho de banda mínimo es menor que un primer valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; o cuando el primer ancho de banda mínimo es mayor que el primer valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

2. El método de acuerdo con la reivindicación 1, en donde determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de los P coeficientes de espectro de energía de FFT de cada una de las N tramas de audio comprende:

ordenar la energía de los P coeficientes de espectro de energía de FFT de cada una de las tramas de audio en orden descendente;

determinar, de acuerdo con la energía, ordenada en orden descendente, de los P coeficientes de espectro de energía de FFT de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio; y

determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio, un valor promedio de los anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de las N tramas de audio.

3. El método de acuerdo con la reivindicación 2, en donde, la determinación del ancho de banda mínimo comprende:

acumular secuencialmente energía de intervalos de frecuencia en los P coeficientes de espectro de energía de FFT en orden descendente; y comparar la energía obtenida después de cada acumulación con la energía total de la trama de audio y, si una proporción es mayor que la primera proporción preestablecida, finalizar el proceso de acumulación, donde una cantidad de veces de acumulación es el ancho de banda mínimo.

4. Un aparato, en donde el aparato comprende:

una unidad (201) de obtención, configurada para obtener N tramas de audio, en donde la dispersión de distribución se determina para cada una de las N tramas de audio de entrada, en donde las N tramas de audio comprenden una trama de audio actual y N es un número entero positivo;

5 una unidad (202) de determinación, configurada para determinar la dispersión de distribución, en un espectro, de energía de las N tramas de audio obtenidas por la unidad de obtención; y

10 la unidad (202) de determinación está además configurada para determinar, de acuerdo con la dispersión de distribución, en el espectro, de la energía de las N tramas de audio, si utilizar un primer método de codificación o un segundo método de codificación para codificar la trama de audio actual, en donde el primer método de codificación es un método de codificación que está basado en la transformada de tiempo-frecuencia y la cuantificación del coeficiente de transformada y que no está basado en la predicción lineal y, el segundo método de codificación, es un método de codificación basado en predicción lineal;

15 la unidad (202) de determinación está configurada específicamente para dividir un espectro de cada una de las N tramas de audio en P coeficientes de espectro de energía de FFT y determinar un parámetro de dispersión general de acuerdo con la energía de los P coeficientes de espectro de energía de FFT de cada una de las N tramas de audio, donde P es un número entero positivo y el parámetro de dispersión general indica la dispersión de distribución, en el espectro, de la energía de las N tramas de audio;

en donde el parámetro de dispersión general comprende un primer ancho de banda mínimo;

20 la unidad (202) de determinación está configurada específicamente para determinar un valor promedio de anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio de acuerdo con la energía de los P coeficientes de espectro de energía de FFT de cada una de las N tramas de audio, en donde un ancho de banda mínimo se encuentra a partir de los P coeficientes de espectro de energía de FFT, de manera que una proporción que la energía del ancho de banda representa en la energía total de una trama es la primera proporción preestablecida, en donde el valor promedio de los anchos de banda mínimos de distribución, en el espectro, de la primera proporción de energía preestablecida de las N tramas de audio es el primer ancho de banda mínimo; y

25 la unidad (202) de determinación está configurada específicamente para: cuando el primer ancho de banda mínimo es menor que un primer valor preestablecido, determinar utilizar el primer método de codificación para codificar la trama de audio actual; o cuando el primer ancho de banda mínimo es mayor que el primer valor preestablecido, determinar utilizar el segundo método de codificación para codificar la trama de audio actual.

30 5. El aparato de acuerdo con la reivindicación 4, en donde la unidad (202) de determinación está configurada específicamente para:

ordenar la energía de los P coeficientes de espectro de energía de FFT de cada una de las tramas de audio en orden descendente;

35 determinar, de acuerdo con la energía, ordenada en orden descendente, de los P coeficientes de espectro de energía de FFT de cada una de las N tramas de audio, un ancho de banda mínimo de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio; y

40 determinar, de acuerdo con el ancho de banda mínimo de distribución, en el espectro, de la energía que representa no menos que la primera proporción preestablecida de cada una de las N tramas de audio, un valor promedio de anchos de banda mínimos de distribución, en el espectro, de energía que representa no menos que la primera proporción preestablecida de las N tramas de audio.

6. El aparato de acuerdo con la reivindicación 5, en donde, para determinar el ancho de banda mínimo, la unidad (202) de determinación está configurada específicamente para:

45 acumular secuencialmente la energía de los intervalos de frecuencia en los P coeficientes de espectro de energía de FFT en orden descendente;

comparar la energía obtenida después de cada acumulación con la energía total de la trama de audio, y

finalizar el proceso de acumulación si una proporción es mayor que la primera proporción preestablecida, donde una cantidad de veces de acumulación es el ancho de banda mínimo.

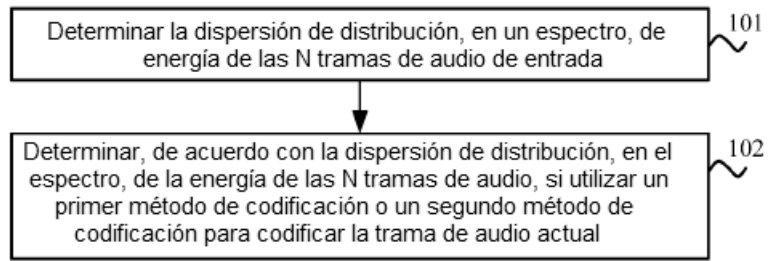


FIG. 1

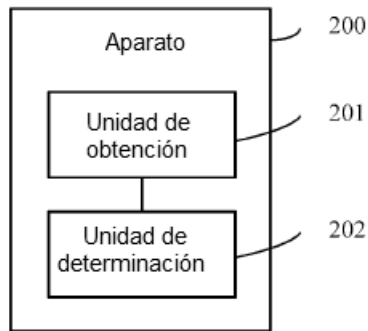


FIG. 2

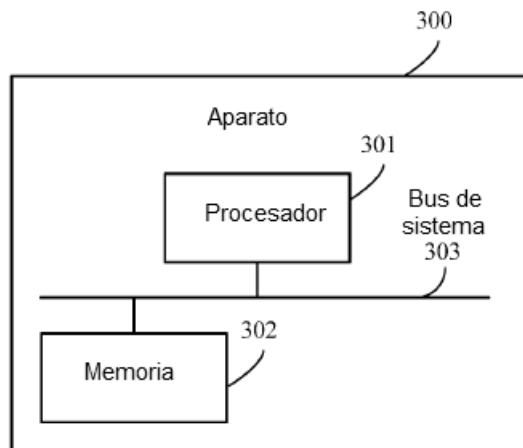


FIG. 3