

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 593 822**

51 Int. Cl.:

**G10L 19/00** (2013.01)

**H04S 3/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **09.06.2008** E 10013592 (0)

97 Fecha y número de publicación de la concesión europea: **10.08.2016** EP 2278582

54 Título: **Método y aparato para procesar una señal de audio**

30 Prioridad:

**08.06.2007 US 942967 P**

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

**13.12.2016**

73 Titular/es:

**LG ELECTRONICS INC. (100.0%)  
20, Yeouido-dong, Yeongdeungpo-gu  
Seoul 150-721, KR**

72 Inventor/es:

**JUNG, YANG WON y  
OH, HYEN O**

74 Agente/Representante:

**CURELL AGUILÁ, Mireia**

ES 2 593 822 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 y aparato para procesar una señal de audio.

**5 Campo técnico**

La presente invención se refiere a un método y a un aparato para procesar una señal de audio, y más particularmente, a un aparato para procesar una señal de audio y a un método para ello. Aunque la presente invención es adecuada para una amplia variedad de aplicaciones, resulta particularmente adecuada para procesar la señal de audio recibida a través de un medio digital, una señal de radiodifusión o similares.

**Antecedentes de la técnica**

En general, en el procesado de una señal de audio basada en objetos, un objeto individual que compone una señal de entrada se procesa como un objeto independiente. En este caso, debido a que puede existir correlación entre los objetos, se habilita una codificación más eficiente si la misma se lleva a cabo utilizando la correlación.

El documento WO 2008/069596 A constituye técnica anterior de acuerdo con el Artículo 54(3) EPC y se puede considerar que da a conocer un método para procesar una señal de audio, que comprende: recibir una señal de submezcla (*downmix*), una primera información multicanal, y una información de objeto; procesar la señal de submezcla utilizando la información de objeto y una información de mezcla; y transmitir una de entre la primera información multicanal y una segunda información multicanal de acuerdo con la información de mezcla, en donde se da a conocer que la segunda información de canales se genera usando la información de objeto y la información de mezcla.

**Divulgación del problema técnico de la invención**

El objetivo de la presente invención es aumentar la eficiencia en el procesado de una señal de audio.

**30 Solución técnica**

Por consiguiente, la presente invención se refiere a un aparato para procesar una señal de audio y a un método para lo mismo, que eliminan sustancialmente uno o más de los problemas debidos a limitaciones y desventajas de la técnica relacionada.

Uno de los objetivos de la presente invención es proporcionar un método de procesado de una señal, por el cual la señal se puede procesar más eficientemente utilizando un parámetro auxiliar en el procesado de una señal de audio basada en objetos.

Otro objetivo de la presente invención es proporcionar un método de procesado de una señal, por el cual la señal se puede procesar más eficientemente controlando parcialmente la señal de objeto.

Otro objetivo de la presente invención es proporcionar un método de procesado de una señal, por el cual una señal de audio basada en objetos se procesa utilizando la correlación entre objetos.

Otro objetivo de la presente invención es proporcionar un método de obtención de información que indica correlación entre objetos agrupados.

Otro objetivo de la presente invención es proporcionar un método de transmisión de una señal, por el cual la señal se puede transmitir más eficientemente.

Otro objetivo de la presente invención es proporcionar un método de procesado de una señal, por el cual pueden obtenerse varios efectos de sonido.

Un objetivo adicional de la presente invención es proporcionar un método de procesado de una señal, el cual permite que un usuario modifique una señal de mezcla utilizando una señal fuente.

En la descripción que se ofrece a continuación se expondrán características y ventajas adicionales de la invención, y las mismas resultarán evidentes de forma parcial a partir de la descripción, o se pueden asimilar al poner en práctica la invención. Los objetivos y otras ventajas de la invención se materializarán y alcanzarán por medio de la estructura que se indica particularmente en la descripción publicada y sus reivindicaciones, así como en los dibujos adjuntos.

De acuerdo con la presente invención, se proporcionan un método y un aparato según las reivindicaciones independientes.

Preferentemente, un método de procesado de una señal de audio incluye recibir información de submezcla de por lo

- 5 menos una señal de objeto submezclada (*downmixed*), obtener información lateral que incluye información de objeto, e información de mezcla, generar una pluralidad de informaciones de canales sobre la base de la información lateral y la información de mezcla, y generar una señal de canal de salida a partir de la información de submezcla utilizando la pluralidad de informaciones de canales, en donde la información de objeto incluye por lo menos una de entre información de nivel de la señal de objeto, información de correlación de la señal de objeto, información de ganancia de la señal de objeto e información suplementaria de la misma.
- 10 Preferentemente, la información suplementaria incluye información de diferencia entre un valor real de la información de ganancia de la señal de objeto y un valor de estimación de la misma.
- 15 Preferentemente, la información de mezcla se genera sobre la base de por lo menos una de información de posición de la señal de objeto, la información de ganancia de la señal de objeto e información de configuración de reproducción de la señal de objeto.
- 20 Preferentemente, el método incluye además determinar si llevar a cabo un proceso inverso utilizando la información de objeto y la información de mezcla, y cuando se lleva a cabo el proceso inverso de acuerdo con la determinación, obtener un valor de ganancia del proceso inverso para la compensación de ganancia, en donde si el número de objetos modificados es mayor que el correspondiente de objetos no modificados, el proceso inverso indica que la compensación de ganancia se lleva a cabo en referencia al objeto no modificado, y en donde la señal de canal de salida se genera basándose en el valor de ganancia del proceso inverso.
- 25 Preferentemente, la información de nivel de la señal de objeto incluye la información de nivel modificada sobre la base de la información de mezcla, y la pluralidad de informaciones de canales se genera sobre la base de la información de nivel modificada.
- 30 Más preferentemente, si la magnitud de una señal de objeto específica se amplifica o atenúa en referencia a un umbral preestablecido, la información de nivel modificada se genera multiplicando la información de nivel de la señal de objeto por una constante mayor que 1.
- 35 Preferentemente, un método de procesado de una señal de audio incluye recibir información de submezcla de por lo menos una señal de objeto submezclada, obtener información lateral que incluye información de objeto, e información de mezcla, generar una pluralidad de informaciones de canales sobre la base de la información lateral obtenida y la información de mezcla obtenida, y generar una señal de canal de salida a partir de la información de submezcla utilizando la pluralidad de informaciones de canales, en donde la información de objeto incluye por lo menos una de entre información de nivel de la señal de objeto, información de correlación de la señal de objeto e información de ganancia de la señal de objeto, y en donde por lo menos una de la información de objeto y la información de mezcla se cuantifica.
- 40 Preferentemente, el método incluye además obtener información de acoplamiento que indica si un objeto se agrupa con otro objeto, en donde la información de correlación de la señal de objeto se obtiene basándose en la información de acoplamiento.
- 45 Más preferentemente, el método incluye además obtener una meta-información común para objetos agrupados sobre la base de la información de acoplamiento.
- 50 En este caso, la meta-información incluye el número de caracteres de los meta-datos e información de cada carácter de los meta-datos.
- 55 Preferentemente, un método de procesado de una señal de audio incluye recibir información de submezcla de por lo menos una señal de objeto submezclada, obtener información lateral que incluye información de objeto e información de acoplamiento, e información de mezcla, generar una pluralidad de informaciones de canales sobre la base de la información lateral obtenida y la información de mezcla obtenida, y generar una señal de canal de salida a partir de la información de submezcla utilizando la pluralidad de informaciones de canales, en donde la información de objeto se discrimina en una señal de objeto independiente y una señal de objeto de fondo, en donde la información de objeto incluye por lo menos una de entre información de nivel de la señal de objeto, información de correlación de la señal de objeto e información de ganancia de la señal de objeto, y en donde la información de correlación de la señal de objeto se obtiene basándose en la información de acoplamiento.
- 60 Preferentemente, la señal de objeto independiente incluye una señal de objeto vocal.
- 65 Preferentemente, la señal de objeto de fondo incluye una señal de objeto de acompañamiento.
- Preferentemente, la señal de objeto de fondo incluye por lo menos una señal basada en canales.
- Preferentemente, la señal de objeto se discrimina en la señal de objeto independiente y la señal de objeto de fondo sobre la base de información de indicador.

Preferentemente, la señal de audio se recibe en forma de una señal de radiodifusión.

Preferentemente, la señal de audio se recibe a través de un medio digital.

Preferentemente, un soporte de grabación legible por ordenador incluye un programa grabado en el mismo, en donde el programa se proporciona para ejecutar el método de la presente invención.

Preferentemente, un aparato para procesar una señal de audio incluye una unidad de procesado de submezcla que recibe información de submezcla de por lo menos una señal de objeto submezclada, una unidad generadora de información que obtiene información lateral que incluye información de objeto, e información de mezcla, de manera que la unidad generadora de información genera una pluralidad de informaciones de canales basándose en la información lateral obtenida y la información de mezcla obtenida, y una unidad de descodificación multicanal que genera una señal de canal de salida a partir de la información de submezcla utilizando la pluralidad de informaciones de canales, en donde la información de objeto incluye por lo menos una de entre información de nivel de la señal de objeto, información de correlación de la señal de objeto, información de ganancia de la señal de objeto e información suplementaria de la misma.

Preferentemente, un aparato para procesar una señal de audio incluye una unidad de procesado de submezcla que recibe información de submezcla de por lo menos una señal de objeto submezclada, una unidad generadora de información que obtiene información lateral que incluye información de objeto e información de mezcla, de manera que la unidad generadora de información genera una pluralidad de informaciones de canales basándose en la información lateral obtenida y la información de mezcla obtenida, y una unidad de descodificación multicanal que genera una señal de canal de salida a partir de la información de submezcla utilizando la pluralidad de informaciones de canales, en donde la información de objeto incluye por lo menos una de entre información de nivel de la señal de objeto, información de correlación de la señal de objeto e información de ganancia de la señal de objeto, y en donde por lo menos una de entre la información de objeto y la información de mezcla está cuantificada.

Preferentemente, un aparato para procesar una señal de audio incluye una unidad de procesado de submezcla que recibe información de submezcla de por lo menos una señal de objeto submezclada, una unidad generadora de información que obtiene información lateral que incluye información de objeto e información de acoplamiento, e información de mezcla, de manera que la unidad generadora de información genera una pluralidad de informaciones de canales basándose en la información lateral y la información de mezcla, y una unidad de descodificación multicanal que genera una señal de canal de salida a partir de la información de submezcla utilizando la pluralidad de informaciones de canales, en donde la señal de objeto se discrimina en una señal de objeto independiente y una señal de objeto de fondo, en donde la información de objeto incluye por lo menos una de entre información de nivel de la señal de objeto, información de correlación de la señal de objeto e información de ganancia de la señal de objeto, y en donde la información de correlación de la señal de objeto se obtiene basándose en la información de acoplamiento.

Debe entenderse que tanto la anterior descripción general como la siguiente descripción detallada son ejemplificativas y explicativas, y están destinadas a proporcionar una explicación adicional de la invención según se reivindica.

#### **Efectos ventajosos**

Por consiguiente, la presente invención aporta los siguientes efectos o ventajas. En primer lugar, en caso de señales de objeto que presentan una correlación fuerte entre ellas, la invención puede aumentar la eficiencia en el procesado de una señal de audio utilizando la correlación. En segundo lugar, al transmitir información de atributos detallada sobre cada objeto, un objeto específico del usuario se puede controlar de manera directa y precisa.

#### **Descripción de los dibujos**

Los dibujos adjuntos, que se incluyen para aportar una comprensión adicional de la invención y se incorporan a esta memoria descriptiva constituyendo parte de la misma, ilustran formas de realización de la invención, y, junto con la descripción, sirven para explicar los fundamentos de la invención.

En los dibujos:

la figura 1 es un diagrama de un aparato de procesado de señales de audio de acuerdo con una forma de realización de la presente invención;

la figura 2 es un diagrama para explicar un método de generación de una señal de canal de salida utilizando información de mezcla de acuerdo con una forma de realización de la presente invención;

la figura 3 es un diagrama de flujo para explicar un método más eficiente de procesado de señales de audio de

acuerdo con una forma de realización de la presente invención;

5 la figura 4 es un diagrama de bloques esquemático de un aparato de procesado de señales de audio para transmitir una señal de objeto más eficientemente, de acuerdo con una forma de realización de la presente invención;

la figura 5 es un diagrama de flujo para explicar un método de procesado de una señal de objeto utilizando un control inverso de acuerdo con una forma de realización de la presente invención;

10 la figura 6 y la figura 7 son diagramas de bloques de un aparato de procesado de señales de audio para procesar una señal de objeto utilizando un control inverso de acuerdo con otra forma de realización de la presente invención;

15 la figura 8 es un diagrama estructural de un flujo continuo de bits que contiene meta-información sobre un objeto, de acuerdo con una forma de realización de la presente invención;

la figura 9 es un diagrama de estructura sintáctica para transmitir una señal de audio eficientemente según una forma de realización de la presente invención;

20 las figuras 10 a 12 son diagramas para explicar un proceso de codificación sin pérdidas para transmitir potencia de una fuente de acuerdo con una forma de realización de la presente invención; y

la figura 13 es un diagrama para explicar una interfaz de usuario de acuerdo con una forma de realización de la presente invención.

25

### **Modo óptimo**

### **Modo para llevar la invención a la práctica**

30 A continuación se hará referencia detalladamente a las formas de realización preferidas de la presente invención, cuyos ejemplos se ilustran en los dibujos adjuntos.

35 Como terminología utilizada en la presente invención se selecciona la terminología general que se usa en la actualidad y a nivel global. Además, aparecen terminologías seleccionadas arbitrariamente por el solicitante para casos especiales, para las cuales se explican los significados detallados en la descripción de las formas de realización preferidas de la presente invención. Por tanto, la presente invención debe interpretarse, no con las denominaciones de las terminologías, sino con sus significados.

40 Específicamente, la información que se describe en esta exposición debe interpretarse como terminología que incluye valores, parámetros, coeficientes, elementos y similares, y puede considerarse de manera diferente para no limitar la presente invención.

45 La figura 1 es un diagrama de un aparato de procesado de señales de audio de acuerdo con una forma de realización de la presente invención.

En referencia a la figura 1, un aparato de procesado de señales de audio de acuerdo con una forma de realización de la presente invención puede incluir una unidad generadora de información 110, una unidad de procesado de submezcla 120 y un descodificador multicanal 130.

50 La unidad generadora de información 110 recibe información lateral que incluye información de objeto (OI) y similares, a través de un flujo continuo de bits de la señal de audio, y también puede recibir información de mezcla (MXI) a través de la interfaz de usuario. En este caso, la información de objeto (OI) es la información sobre un objeto incluido en una señal de submezcla y puede incluir información de nivel del objeto, información de correlación del objeto, información de ganancia del objeto, meta-información y similares.

55 La información de nivel del objeto se genera a partir de la normalización de un nivel de objeto usando información de referencia. La información de referencia se corresponde con uno de los niveles del objeto, más particularmente con el más alto de todos los niveles del objeto. La información de correlación del objeto indica correlación entre dos objetos. La información de correlación del objeto puede indicar que dos objetos son señales de canales diferentes de una salida estereofónica que tiene el mismo origen. La información de ganancia del objeto indica un valor sobre la contribución por parte de un objeto para un canal de cada señal de submezcla, y más particularmente, un valor para modificar la contribución por parte de un objeto.

60 Por otra parte, la información preestablecida (PI) puede incluir la información generada sobre la base de información de posición preestablecida, información de ganancia preestablecida, información de configuración de reproducción y similares.

65

La información de posición preestablecida puede indicar información fijada para controlar una posición o la panoramización de cada objeto. La información de ganancia preestablecida es la información fijada para controlar una ganancia de cada objeto, e incluye un factor de ganancia por cada objeto. En este caso, el factor de ganancia por objeto puede variar según el tiempo.

La información preestablecida (PI) puede significar que la información de posición del objeto, la información de ganancia del objeto y la información de configuración de reproducción, que se corresponden con un modo específico, se preestablecen para obtener un efecto de campo acústico o efecto de sonido específico para una señal de audio. Por ejemplo, un modo de karaoke en la información preestablecida puede incluir información de ganancia preestablecida que fija una ganancia del objeto vocal a 0. El modo de estadio en la información preestablecida puede incluir información de posición preestablecida e información de ganancia preestablecida para proporcionar un efecto de que una señal de audio se encuentra en un espacio amplio. De este modo, a un usuario se le facilita el controlar la ganancia o panoramización del objeto seleccionando un modo específico a partir de la información preestablecida (PI) sin ajustar la ganancia o panoramización de cada objeto.

La unidad de procesado de submezcla 120 recibe información de submezcla (en lo sucesivo en la presente, denominada señal de submezcla (DMX)), y a continuación procesa la señal de submezcla (DMX) utilizando información de procesado de submezcla (DPI). Para ajustar la panoramización o ganancia del objeto, se puede procesar la señal de submezcla (DMX).

El decodificador multicanal 130 recibe la submezcla procesada, y a continuación puede generar una señal de diversos canales mezclando en sentido ascendente (*upmixing*) la señal de submezcla procesada con el uso de información multicanal (MI).

La señal de submezcla utilizada en la presente invención puede incluir una señal monofónica, una señal estereofónica o una señal de audio de diversos canales. Por ejemplo, suponiendo que la señal estereofónica se fija a  $\bar{x}_1(n)$  y  $\bar{x}_2(n)$ , la misma se puede representar como una suma de señales fuente, donde "n" indica un índice de tiempo. Por tanto, la señal estereofónica se puede representar en forma de la fórmula 1.

[Fórmula 1]

$$\begin{aligned}\bar{x}_1(n) &= \sum_{i=1}^I a_i \bar{s}_i(n) \\ \bar{x}_2(n) &= \sum_{i=1}^I b_i \bar{s}_i(n),\end{aligned}$$

En este caso, "I" indica el número de señales fuente incluidas en la señal estereofónica y la  $\bar{s}_i(n)$  indica una señal fuente. Además, "a<sub>i</sub>" y "b<sub>i</sub>" son valores para determinar una panoramización en amplitud y una ganancia para cada señal fuente, respectivamente. Cada  $\bar{s}_i(n)$  puede ser independiente. La  $\bar{s}_i(n)$  puede ser una señal fuente pura o puede incluir una señal fuente pura a la cual se le añaden componentes de señal con un poco de reverberación y un efecto de sonido. Por ejemplo, una componente de señal de reverberación específica se puede representar en forma de dos señales fuente, es decir, una señal mezclada en un canal izquierdo y una señal mezclada en un canal derecho.

Una forma de realización de la presente invención puede modificar una señal estereofónica que incluye señales fuente con el fin de remezclar M señales fuente ( $0 \leq M \leq I$ ). Las señales fuente se pueden remezclar en una señal estereofónica con diferentes factores de ganancia. Una señal de remezcla se puede representar en forma de la fórmula 2.

[Fórmula 2]

$$\tilde{y}_1(n) = \sum_{i=1}^M c_i \tilde{s}_i(n) + \sum_{i=M+1}^I a_i \tilde{s}_i(n)$$

$$\tilde{y}_2(n) = \sum_{i=1}^M d_i \tilde{s}_i(n) + \sum_{i=M+1}^I b_i \tilde{s}_i(n),$$

En la fórmula 2, "c<sub>i</sub>" y "d<sub>i</sub>" son factores de ganancia nuevos para M señales fuente que se van a remezclar. El "c<sub>i</sub>" y el "d<sub>i</sub>" pueden ser proporcionados por un lado de descodificador.

5 Según una forma de realización de la presente invención, una señal transportada de canal de entrada se puede modificar en una señal de canal de salida sobre la base de información de mezcla.

10 En este caso, la información de mezcla (MXI) puede indicar la información generada sobre la base de información de posición del objeto, información de ganancia del objeto, información de configuración de reproducción o similares. En este caso, la información de posición del objeto puede indicar la información introducida por un usuario para controlar una posición o panoramización de cada objeto. La información de ganancia del objeto puede indicar la información introducida por un usuario para controlar una ganancia de cada objeto. Además, la información de configuración de reproducción es la información que incluye el número de altavoces, posiciones de altavoces, información del entorno (posición virtual del altavoz) y similares. La información de configuración de reproducción es introducida por un usuario, se almacena de antemano o se recibe desde otro dispositivo.

15 La información de mezcla puede indicar directamente el alcance con el que un objeto específico se incluye en un canal de salida específico o puede indicar un valor de diferencia para un estado de un canal de entrada. La información de mezcla puede usar el mismo valor dentro de un contenido individual o un valor variable en el tiempo. En el caso de que la información de mezcla sea variable en el tiempo, es posible utilizar la información de mezcla introduciendo un estado de inicio, un estado final y un tiempo de variación. Además, también es posible utilizar la información de mezcla introduciendo un índice de tiempo de un instante de temporización variable y un valor para un estado del instante de temporización.

20 Por motivos de claridad y comodidad de la descripción, una forma de realización de la presente invención describe un caso en el que la información de mezcla indica el alcance con el que un objeto específico se incluye en un canal de salida específico en la forma que se muestra en la fórmula 1. En este caso, cada canal de salida se puede construir como la fórmula 2. En este caso, con el fin de discriminar a<sub>i</sub> y b<sub>i</sub> con respecto a c<sub>i</sub> y d<sub>i</sub>, supóngase que a<sub>i</sub> y b<sub>i</sub> son ganancias de mezcla y supóngase que c<sub>i</sub> y d<sub>i</sub> son ganancias de mezcla de reproducción.

Supóngase que la información de mezcla no viene dada como la ganancia de mezcla de reproducción sino como ganancia y panoramización. La ganancia (g<sub>i</sub>) y la panoramización (l<sub>i</sub>) pueden venir dadas como la fórmula 3.

35 [Fórmula 3]

$$g_i = 10 \log_{10}(c_i^2 + d_i^2)$$

$$l_i = 20 \log_{10}(d_i/c_i)$$

40 Por tanto, c<sub>i</sub> y d<sub>i</sub> se pueden obtener utilizando a<sub>i</sub> y b<sub>i</sub>. Además, se pone de manifiesto que la expresión relacional entre la ganancia y la panoramización y la ganancia de mezcla se puede representar de una forma diferente.

La figura 2 es un diagrama para explicar un método de generación de una señal de canal de salida utilizando información de mezcla de acuerdo con una forma de realización de la presente invención.

45 La unidad de procesado de submezcla 120 mostrada en la figura 1 puede obtener una señal de canal de salida multiplicando una señal de canal de entrada por un coeficiente específico. En referencia a la figura 2, supóngase que x<sub>1</sub> y x<sub>2</sub> son señales de canal de entrada y supóngase que y<sub>1</sub> e y<sub>2</sub> son señales de canal de salida, las señales de canal de salida reales pueden representarse en forma de la fórmula 4.

50 [Fórmula 4]

$$y1\_hat = w11 * x1 + w12 * x2$$

$$y2\_hat = w21 * x1 + w22 * x2$$

5 En fórmula 4,  $y_i\_hat$  indica un valor de salida que se va a discriminar con respecto a un valor teórico obtenido a partir de la fórmula 2. "w11~w22" pueden significar factores de ponderación. Además,  $x_i$ ,  $w_{ij}$  e  $y_i$  se pueden corresponder con señales de frecuencia específicas en un momento específico, respectivamente.

Una forma de realización de la presente invención proporciona un método de obtención de un canal de salida eficiente utilizando factores de ponderación.

10 Los factores de ponderación se pueden estimar de diversas maneras. En particular, la presente invención puede usar una estimación de mínimos cuadrados. En este caso, el error de estimación generado se puede definir con la fórmula 5.

[Fórmula 5]

15

$$e1 = y1 - y1\_hat$$

$$e2 = y2 - y2\_hat$$

Los factores de ponderación se pueden generar por subbanda para reducir al mínimo los errores cuadráticos medios  $E\{e1^2\}$  y  $E\{e2^2\}$ . En este caso, si el error de estimación es ortogonal a  $x_1$  y  $x_2$ , se puede usar el hecho de que el error cuadrático medio se reduce al mínimo. Por otra parte,  $w_{11}$  y  $w_{12}$  se pueden representar con la fórmula 6.

20

[Fórmula 6]

$$w_{11} = \frac{E\{x_1^2\}E\{x_1y_1\} - E\{x_1x_2\}E\{x_2y_1\}}{E\{x_1^2\}E\{x_2^2\} - E\{x_1x_2\}^2}$$

$$w_{12} = \frac{E\{x_1x_2\}E\{x_2y_1\} - E\{x_1^2\}E\{x_2y_1\}}{E\{x_1x_2\}^2 - E\{x_1^2\}E\{x_2^2\}}$$

25

Además,  $E\{x_1y_1\}$  y  $E\{x_2y_1\}$  se pueden generar con la fórmula 7.

[Fórmula 7]

$$E\{x_1y_1\} = E\{x_1^2\} + \sum_{n=1}^M a_n(a_n - a_1)E\{e_n^2\}$$

$$E\{x_2y_1\} = E\{x_1x_2\} + \sum_{n=1}^M b_n(a_n - a_1)E\{e_n^2\}$$

30

Así mismo,  $w_{21}$  y  $w_{22}$  se pueden representar con la fórmula 8.

[Fórmula 8]

35

$$w_{21} = \frac{E\{x_1^2\}E\{x_1y_2\} - E\{x_1x_2\}E\{x_2y_2\}}{E\{x_1^2\}E\{x_2^2\} - E\{x_1x_2\}^2}$$

$$w_{22} = \frac{E\{x_1x_2\}E\{x_2y_2\} - E\{x_1^2\}E\{x_2y_2\}}{E\{x_1x_2\}^2 - E\{x_1^2\}E\{x_2^2\}}$$

Además,  $E\{x_2y_1\}$  y  $E\{x_2y_2\}$  se pueden generar con la fórmula 9.

40

[Fórmula 9]

$$E\{x_1y_2\} = E\{x_1x_2\} + \sum_{n=1}^M a_n(d_n - b_1)E\{e_n^2\}$$

$$E\{x_2y_2\} = E\{x_2^2\} + \sum_{n=1}^M b_n(d_n - b_1)E\{e_n^2\}$$

45 De acuerdo con una forma de realización de la presente invención, para configurar información lateral o generar una señal de salida en una codificación basada en objetos, se puede usar información de energía (o información de nivel) de una señal de objeto.



Por ejemplo, en caso de que se configure información lateral, es posible transportar energía de una señal de objeto, un valor de energía relativo entre señales de objeto o un valor de energía relativo entre una señal de objeto y una señal de canal. Por otra parte, en caso de que se genere una señal de salida, se puede usar energía de una señal de objeto.

5 Utilizando señal de canal de entrada, información lateral e información de mezcla, se puede generar una señal de canal de salida que tiene un efecto de sonido específico. En el proceso para generar la señal de canal de salida, se puede usar información de energía de una señal de objeto. La información de energía de la señal de objeto se puede incluir en la información lateral o se puede estimar usando la información lateral y la señal de canal. Por otra parte, es posible usar la información de energía de la señal de objeto modificándola.

Se propone un método de modificación de la información de energía de la señal de objeto de acuerdo con una forma de realización de la presente invención, para mejorar la calidad de la señal de canal de salida. De acuerdo con la presente invención, se puede modificar información de energía bajo el control de un usuario.

15 En referencia a la fórmula 7 y la fórmula 9, puede observarse que la información de energía  $E\{s_i^2\}$  de una señal de objeto se usa para obtener factores de ponderación  $w_{11}$ ~ $w_{22}$  para la generación de una señal de canal de salida. Una forma de realización de la presente invención se refiere a un método de generación de una señal de salida usando coeficientes del propio canal  $w_{11}$  y  $w_{22}$  y coeficientes de canal cruzado  $w_{21}$  y  $w_{12}$ . En caso de usar otro método, según se ha mencionado en la anterior descripción, se pone de manifiesto que hay disponible información de energía de una señal de objeto.

20 En un proceso para obtener factores de ponderación de un canal de salida, la presente invención propone un método de modificación para utilizar información de nivel (o información de energía) de una señal de objeto. Por ejemplo, está disponible la fórmula 10.

[Fórmula 10]

$$\begin{aligned} E\{x_1*y_1\} &= E\{x_1^2\} + \sum \{a_i*(c_i - a_i)E\_mod\{s_i^2\}\} \\ E\{x_2*y_1\} &= E\{x_1*x_2\} + \sum \{b_i*(c_i - a_i)E\_mod\{s_i^2\}\} \\ E\{x_1*y_2\} &= E\{x_1*x_2\} + \sum \{a_i*(d_i - b_i)E\_mod\{s_i^2\}\} \\ E\{x_2*y_2\} &= E\{x_2^2\} + \sum \{b_i*(d_i - b_i)E\_mod\{s_i^2\}\} \end{aligned}$$

30 La información de nivel modificada ( $E\_mod$ ) es aplicable independientemente según una señal de objeto o es aplicable de manera idéntica a cada señal de objeto.

35 La información de nivel modificada de la señal de objeto se puede generar basándose en información de mezcla. Además, se puede generar una pluralidad de informaciones de canales basándose en la información de nivel modificada. Por ejemplo, en caso de que se cambie la magnitud de una señal de objeto específica de manera notable, se puede obtener información de nivel modificada multiplicando información de nivel de la señal de objeto específica por un valor predeterminado. En este caso, se puede determinar si la magnitud de la señal de objeto específica se amplifica o atenúa considerablemente en referencia a un umbral preestablecido. Por ejemplo, el umbral preestablecido puede ser un valor relativo a una magnitud de otra señal de objeto. Para otro caso, el umbral preestablecido puede ser un valor específico de acuerdo con la sicología perceptiva del ser humano o un valor calculado de acuerdo con diversas pruebas. Además, el valor predeterminado, por el cual se multiplica la información de nivel de la señal de objeto específica, puede incluir una constante mayor de 1. En la siguiente descripción, se explicarán detalladamente los casos anteriores.

45 “ $E\_mod\{s_i^2\}$ ” de la fórmula 10 se puede modificar como la fórmula 11 utilizando  $E\{S_i^2\}$ .

[Fórmula 11]

$$50 \quad E\_mod\{s_i^2\} = \alpha * E\{s_i^2\}$$

En la fórmula 11, “alpha” se puede proporcionar de acuerdo con la relación con información de mezcla de reproducción y ganancia de mezcla original de la manera siguiente. En caso de que la información de energía de una señal de objeto se modifique independientemente según cada señal de objeto, se pone de manifiesto que alpha se puede representar como  $\alpha_i$ . Por ejemplo, si  $s_i$  se atenúa considerablemente, puede ser  $\alpha > 1$ . Si  $s_i$  se atenúa o amplifica apropiadamente, puede ser  $\alpha = 1$ . Si  $s_i$  se amplifica considerablemente, puede ser  $\alpha < 1$ .

60 En este caso, se puede conocer la atenuación o amplificación de  $s_i$  a través de la relación entre ganancias de mezcla originales  $a_i$  y  $b_i$  y ganancias de mezcla de reproducción  $c_i$  y  $d_i$ . Por ejemplo, si  $a_i^2 + b_i^2 > c_i^2 + d_i^2$ , se atenúa  $s_i$ . Por el contrario, si  $a_i^2 + b_i^2 < c_i^2 + d_i^2$ ,  $s_i$  se amplifica. Por tanto, es posible ajustar el valor de alpha por medio del

esquema que se representa con las fórmulas 12 a 14.

[Fórmula 12]

$$5 \quad (a_i^2 + b_i^2) / (c_i^2 + d_i^2) > \text{Thr\_atten}$$

$$\alpha = \alpha\_atten, \alpha\_atten > 1$$

[Fórmula 13]

$$10 \quad (a_i^2 + b_i^2) / (c_i^2 + d_i^2) < \text{Thr\_boost}$$

$$\alpha = \alpha\_boost, \alpha\_boost > 1$$

[Fórmula 14]

$$15 \quad \text{Thr\_atten} > (a_i^2 + b_i^2) / (c_i^2 + d_i^2) > \text{Thr\_boost}$$

$$\alpha = 1$$

En este caso, Thr\_atten y Thr\_boost pueden significar umbrales. Cada uno de los umbrales puede ser un valor específico de acuerdo con la psicología perceptiva del ser humano o un valor calculado según diversas pruebas. Además, alpha\_atten puede presentar la característica de  $\alpha\_atten \geq \alpha\_boost$ .

En la presente invención, se puede utilizar alpha\_atten para permitir que  $E\_mod\{s_i^2\}$  obtenga una ganancia de 2 dB en comparación con la de  $E\{s_i^2\}$ .

25 Por otra parte, en la presente invención, se puede usar  $10^{0.2}$  como valor de alpha\_atten.

De acuerdo con otra realización de la presente invención, se puede usar una  $E\_mod\{s_i^2\}$  para obtener factores de ponderación en lugar de usar la misma  $E\_mod\{s_i^2\}$ .

30 Por ejemplo, está disponible la fórmula 15.

[Fórmula 15]

$$E\{x_1*y_1\} = E\{x_1^2\} + \sum [a_1*(c_1 - a_1)E\_mod1\{s_1^2\}]$$

$$E\{x_2*y_1\} = E\{x_1*x_2\} + \sum [b_1*(c_1 - a_1)E\_mod1\{s_1^2\}]$$

$$E\{x_1*y_2\} = E\{x_1*x_2\} + \sum [a_1*(d_1 - b_1)E\_mod2\{s_1^2\}]$$

$$E\{x_2*y_2\} = E\{x_2^2\} + \sum [b_1*(d_1 - b_1)E\_mod2\{s_1^2\}]$$

35 Así mismo,  $E\_mod1\{s_i^2\}$  y  $E\_mod2\{s_i^2\}$  de la fórmula 15 se pueden modificar como la fórmula 16.

[Fórmula 16]

$$40 \quad E\_mod1\{s_i^2\} = \alpha_1 * E\{s_i^2\}$$

$$E\_mod2\{s_i^2\} = \alpha_2 * E\{s_i^2\}$$

En este caso, E\_mod1 y alpha1 son valores que contribuyen a la generación de y1, y E\_mod2 y alpha2 son valores que contribuyen a la generación de y2.

45 La  $E\_mod_i\{s_i^2\}$  usada para la fórmula 11 se puede utilizar discriminándola de la manera siguiente. Por ejemplo, supóngase que  $S_i$  se atenúa/amplifica para un canal de una señal de canal de salida solamente. En este caso, no es necesario modificar y utilizar  $E\{s_i^2\}$  para un canal opuesto. En caso afirmativo, si  $s_i$  se suprime solamente para un canal izquierdo, se puede utilizar el valor de E\_mod para las w11 y w12 utilizadas en la generación de una señal de canal de salida izquierdo solamente. En este caso, se puede utilizar  $\alpha_1 = \alpha\_atten$  y  $\alpha_2 = 1$ . Además, las fórmulas 12 a 14 son utilizables como condición para determinar un valor de alpha\_i. En particular, determinando el alcance con el que se atenúa/amplifica una señal de objeto específica en un canal de salida específico, se puede usar el valor de alpha\_i.

55 La fórmula 17 y la fórmula 18 están disponibles para otra forma de realización de la presente invención.

[Fórmula 17]

$$\begin{aligned} E\{x_1*y_1\} &= E\{x_1^2\} + \sum [a_i*(c_i - a_i)E_{\text{mod11}}\{s_i^2\}] \\ E\{x_2*y_1\} &= E\{x_1*x_2\} + \sum [b_i*(c_i - a_i)E_{\text{mod21}}\{s_i^2\}] \\ E\{x_1*y_2\} &= E\{x_1*x_2\} + \sum [a_i*(d_i - b_i)E_{\text{mod12}}\{s_i^2\}] \\ E\{x_2*y_2\} &= E\{x_2^2\} + \sum [b_i*(d_i - b_i)E_{\text{mod22}}\{s_i^2\}] \end{aligned}$$

[Fórmula 18]

$$\begin{aligned} E_{\text{mod11}}\{s_i^2\} &= \alpha_{11} * E\{s_i^2\} \\ E_{\text{mod21}}\{s_i^2\} &= \alpha_{21} * E\{s_i^2\} \\ E_{\text{mod12}}\{s_i^2\} &= \alpha_{12} * E\{s_i^2\} \\ E_{\text{mod22}}\{s_i^2\} &= \alpha_{22} * E\{s_i^2\} \end{aligned}$$

5 Según otra forma de realización de la presente invención, en caso de que se solicite una atenuación/amplificación excesiva, se puede modificar y usar  $E\{s_i^2\}$  para la potenciación de la calidad de la señal de canal de salida. Sin embargo, en caso de usar un canal cruzado, puede que se solicite el uso de la  $E\{s_i^2\}$  sin modificarla. Así, se puede satisfacer la solicitud fijando el uso de  $\alpha_{21} = \alpha_{12} = 1$ .

10 Por el contrario, se puede solicitar que la información de energía de una señal de objeto se modifique no para un canal propio sino para un canal cruzado. En este caso, se puede satisfacer la solicitud fijando el uso de  $\alpha_{11} = \alpha_{22} = 1$ .

15 Aunque no se ha explicado como ejemplo, con un método similar al correspondiente de la descripción anterior, es posible usar  $\alpha_{11}$  para  $\alpha_{22}$  como valores arbitrarios. Además, para la selección de los valores de alpha se pueden utilizar una señal de canal de entrada, información lateral, información de mezcla de reproducción y similares. Por otra parte, para la selección de los valores de alpha se puede utilizar la relación entre una ganancia de mezcla original y una ganancia de mezcla de reproducción.

20 En los ejemplos, el valor de alpha es igual o superior a 1. Además, se entiende que puede hacerse uso de un caso en el que el valor de alpha sea menor de 1.

25 Al mismo tiempo, en un codificador, en información lateral se incluye posiblemente información de energía de una señal de objeto o se incluye posiblemente un valor de energía relativo entre una señal de objeto y una señal de canal. Si es así, el codificador puede configurar la información lateral modificando información de energía de una señal de objeto. Por ejemplo, se puede configurar la información lateral modificando energía de una señal de objeto específica o energía de señales de objeto completas para aumentar al máximo un efecto de reproducción. En este caso, un decodificador puede llevar a cabo el procesado de la señal reconstruyendo la modificación.

30 Por ejemplo, considérese un caso en el que  $E_{\text{mod}}\{s_i^2\}$  se transmite como información lateral a través de la transformación por medio de la fórmula 11. En este caso, un decodificador puede obtener  $E\{s_i^2\}$  dividiendo  $E_{\text{mod}}\{s_i^2\}$  por alpha. Al hacer esto, el decodificador puede usar la  $E_{\text{mod}}\{s_i^2\}$  y/o  $E\{s_i^2\}$  transmitidas selectivamente. El valor de alpha se puede transmitir incluyéndolo en la información lateral. Alternativamente, el valor de alpha puede ser estimado por el decodificador utilizando una señal de canal de entrada transportada e información lateral.

35 Según una forma de realización de la presente invención, se pueden utilizar factores de ponderación para generar un efecto de sonido específico de cada usuario. En este caso, los factores de ponderación se pueden utilizar únicamente de forma parcial. Para la selección de los factores de ponderación, se puede usar la relación entre canales de entrada, características de canales de entrada, características de información lateral transmitida, información de mezcla, características de un factor de ponderación estimado. Por motivos de claridad y comodidad, se supone que  $w_{11}$  y  $w_{22}$  son coeficientes del canal propio y  $w_{12}$  y  $w_{21}$  son coeficientes de canal cruzado.

40 Según una forma de realización de la presente invención, en caso de que no se usen factores de ponderación de forma parcial o se usen los factores de ponderación de forma parcial, se puede volver a estimar los factores de ponderación usados. Por ejemplo, después de que se haya realizado una estimación de  $w_{11}$ ,  $w_{12}$ ,  $w_{21}$  y  $w_{22}$ , si se determina el uso de solamente un coeficiente de canal propio, puede que resulte posible usar  $w_1$  y  $w_2$  después de la estimación de  $w_1$  y  $w_2$  en lugar de utilizar  $w_{11}$  y  $w_{22}$ . En caso de que no se use el coeficiente de canal cruzado, esto es debido a que  $y_i$  se modifica de acuerdo con la fórmula 18 y a que se cambia la estimación de mínimos cuadrados correspondiente.

[Fórmula 18]

$$\begin{aligned} y_1 &= w_1 * x_1 \\ y_2 &= w_2 * x_2 \end{aligned}$$

60

En este caso,  $w_1$  y  $w_2$ , que reducen al mínimo  $e_i$ , se pueden estimar de acuerdo con la fórmula 19.

[Fórmula 19]

$$w_1 = E\{x_1 \cdot y_1\} / E\{x_1^2\}$$

$$w_2 = E\{x_2 \cdot y_2\} / E\{x_2^2\}$$

Al mismo tiempo, en caso de que se usen factores de ponderación de forma parcial,  $y_i$  se modela de manera que resulte adecuado para el caso y se realiza la estimación de un factor de ponderación óptimo para su uso.

Se explican de la manera siguiente varias formas de realización destinadas a utilizar factores de ponderación.

Como primera forma de realización, se puede disponer de un método basado en la coherencia de un canal de entrada.

Si la correlación entre canales de una señal de entrada es muy alta, las señales, que están incluidas en canales, respectivamente, pueden ser muy similares entre sí. En ese caso, se puede obtener un efecto como si se usase un coeficiente de canal cruzado, a pesar de usar solamente un coeficiente de canal propio.

Por ejemplo, se puede estimar el alcance de la correlación utilizando la fórmula 20.

[Fórmula 20]

$$P_i = E\{x_1 \cdot x_2\} / \sqrt{E\{x_1^2\}E\{x_2^2\}}$$

En este caso, si un valor de  $P_i$  es mayor que un umbral, es decir, si  $P_i > P_{i\_Threshold}$ , cada uno de entre  $w_{12}$  y  $w_{21}$  se puede fijar a 0.  $P_{i\_Threshold}$  puede significar un umbral. Por ejemplo, el umbral puede ser un valor específico de acuerdo con la psicología perceptiva del ser humano o un valor calculado de acuerdo con varias pruebas. Como  $w_{11}$  y  $w_{22}$  se pueden usar el  $w_{11}$  y  $w_{22}$  convencionales. Alternativamente, se pueden usar factores de ponderación tales que sean diferentes de  $w_{11}$  y  $w_{22}$ , como  $w_{11} = w_1$  y  $w_{22} = w_2$ . Además,  $w_1$  y  $w_2$  se pueden hallar a través de un método que se representa con la fórmula 19.

Como segundo método, se puede disponer de un método que utiliza la norma de un factor de ponderación.

En la presente forma de realización, se puede seleccionar un factor de ponderación, el cual será utilizado por la unidad de procesado de submezcla 120, utilizando la norma de factores de ponderación.

En primer lugar, se pueden hallar los factores de ponderación  $w_{11}$ – $w_{22}$  que incluyen factores de ponderación para los cuales se utiliza un canal cruzado. En este caso, la norma de los factores de ponderación se puede hallar con la fórmula 21.

[Fórmula 21]

$$A = w_{11}^2 + w_{12}^2 + w_{21}^2 + w_{22}^2$$

Además, se pueden hallar factores de ponderación  $w_1$  y  $w_2$  para los cuales no se utiliza el canal cruzado. En este caso, la norma de los factores de ponderación se puede hallar con la fórmula 22.

[Fórmula 22]

$$B = w_1^2 + w_2^2$$

En este caso, si  $A < B$ , se pueden usar los factores de ponderación  $w_{11}$ – $w_{22}$ . Si  $B < A$ , se pueden usar los factores de ponderación  $w_1$  y  $w_2$ . Concretamente, comparando un caso en el que se usan cuatro factores de ponderación y un caso en el que se usan factores de ponderación parciales entre sí, se puede seleccionar el método más eficiente. Si se usa el método anterior, se puede evitar un caso en el que un sistema llega a ser inestable debido a magnitudes considerablemente grandes de los factores de ponderación.

Como tercera forma de realización, se puede disponer de un método de utilización de energía de un canal de entrada.

Si  $w_{11}$ – $w_{22}$  se hallan por medio de un método convencional para un caso en el que un canal específico no tiene energía, es decir, un caso en el que una señal existe por ejemplo solamente en un canal, puede generarse un resultado no deseado. En este caso, debido a que un canal de entrada que no tiene energía no puede contribuir a la

salida, el factor de ponderación del canal de entrada que no tiene energía se puede fijar a 0.

El hecho de si un canal específico tiene o no energía se puede estimar con el método representado con la fórmula 23.

5

[Fórmula 23]

$$E\{x_i^2\} < \text{Umbral}$$

10 En este caso, se pueden estimar  $w_{11}$  y  $w_{12}$  con un método nuevo de tal manera que se considere que  $x_2$  es el caso en el que no hay energía, en lugar de usar el valor que se halla con el método convencional. Así mismo, el valor de umbral puede significar un umbral. Por ejemplo, el valor de umbral puede incluir un valor específico de acuerdo con la psicología perceptiva del ser humano o un valor calculado según varias pruebas.

15 Por ejemplo, si  $x_2$  no tiene energía, una señal de salida se puede generar con la fórmula 24.

[Fórmula 24]

$$y_{1\_hat} = w_{11} * x_1$$

$$y_{2\_hat} = w_{21} * x_2$$

20

Además,  $w_{11}$  y  $w_{21}$  se pueden estimar con la fórmula 25.

[Fórmula 25]

$$w_{11} = E\{x_1 * y_1\} / E\{x_1^2\}$$

$$w_{21} = E\{x_1 * y_2\} / E\{x_1^2\}$$

25

En este caso, resulta que  $w_{12} = w_{22} = 0$ .

30 Como cuarta forma de realización, se puede disponer de un método en el que se usa información de ganancia de mezcla.

35 En relación con un caso en el que para la codificación basada en objetos es necesario un factor de ponderación para un canal cruzado, puede existir un caso en el que la señal de salida de un canal propio no se genera a partir de una señal de entrada del canal propio. Esto puede ocurrir si una señal incluida solamente en un canal (o una señal principalmente incluida en un canal) se transmite al otro canal. Concretamente, puede ocurrir en caso de que se intente modificar una característica de panoramización correspondiente para una entrada en la que un objeto específico se panoramiza a un canal específico.

40 En este caso, se puede obtener un efecto de sonido específico solamente si se usa un factor de ponderación para un canal cruzado. Además, son necesarios un método de detección de dicho caso y un método de determinación de cómo utilizar el factor de ponderación. En la presente forma de realización, se proponen un método de detección y un método de uso del factor de ponderación.

45 Por ejemplo, se puede suponer un caso en el que una señal de objeto procesada es monofónica. En primer lugar, se puede determinar si una señal de objeto es monofónica. Si la señal de objeto es monofónica, se puede determinar si la misma se panoramiza hacia el lateral. En este caso, la determinación de la panoramización lateral se puede llevar a cabo utilizando  $a_i/b_i$ . En particular, si  $a_i/b_i = 1$ , se puede observar que la señal de objeto está incluida en cada canal al mismo nivel. Esto puede significar que la señal de objeto está situada en el centro en un espacio acústico. Si  $a_i/b_i < \text{Thr}_B$ , puede observarse que la señal de objeto se panoramiza hacia el lateral (derecho) dirigida por  $b_i$ .  
50 Por el contrario, si  $a_i/b_i > \text{Thr}_A$ , puede observarse que la señal de objeto se panoramiza hacia el lateral (izquierdo) dirigida por  $a_i$ . En este caso, un valor de  $\text{Thr}_A$  o  $\text{Thr}_B$  puede significar un valor de umbral. Por ejemplo, el valor de umbral puede ser un valor específico de acuerdo con la psicología perceptiva del ser humano o un valor calculado de acuerdo con varias pruebas.

55 Como consecuencia de la determinación, si se lleva a cabo la panoramización (*panning*) lateral, se determina si la panoramización se va a cambiar por medio de una ganancia de mezcla de reproducción. El hecho de si la panoramización se va a cambiar se puede determinar comparando un valor de  $a_i/b_i$  con un valor de  $c_i/d_i$ . Por ejemplo, supóngase un estado en el que  $a_i/b_i$  se panoramiza hacia la derecha. Si  $c_i/d_i$  se panoramiza aún más hacia la derecha, puede que no sea necesario un coeficiente de canal cruzado. Sin embargo, si  $c_i/d_i$  se panoramiza hacia la izquierda, se puede incluir la componente de señal de objeto en un canal de salida izquierdo utilizando el coeficiente de canal cruzado.

60

En caso de que se compare el valor de  $a_i/b_i$  con el valor de  $c_i/d_i$ , la sensibilidad de la comparación se puede ajustar aplicando un factor de ponderación adecuado a  $a_i/b_i$  o  $c_i/d_i$ . Por ejemplo, en lugar de comparar  $c_i/d_i$  con  $a_i/b_i$ , se puede usar la fórmula 26.

5 [Fórmula 26]

$$(c_i / d_i) * \alpha > a_i / b_i$$

$$(c_i / d_i) * \beta < a_i / b_i$$

10 En caso de que se use la fórmula 26, se puede ajustar la sensibilidad al uso de un coeficiente de canal cruzado ajustando apropiadamente  $\alpha$  y  $\beta$ .

15 Por otra parte, aunque se cambie la panoramización de la señal de objeto panoramizada lateralmente, si la señal de objeto no tiene suficiente energía, es posible utilizar solamente un coeficiente de canal propio en lugar de utilizar un coeficiente de canal cruzado. Por ejemplo, si una señal de objeto, la cual se panoramiza en el lateral y cuya panoramización se cambia por medio de una ganancia de mezcla de reproducción, existe en una parte frontal de un contenido correspondiente y si después de esto la señal de objeto no existe, puede usarse un coeficiente de canal cruzado solamente para la sección en la que existe la señal de objeto.

20 Tal como propone la forma de realización de la presente invención, usando información de energía de un objeto correspondiente, es posible seleccionar si se utiliza un coeficiente de canal cruzado. La energía del objeto correspondiente se puede transmitir en forma de información lateral o se puede estimar usando información lateral transmitida y una señal de entrada.

25 Como quinta forma de realización, se puede disponer de un método de uso de características del objeto.

En caso de que una señal de objeto sea una señal de objeto de diversos canales, la misma se puede procesar de acuerdo con la característica de la señal de objeto. Por motivos de claridad y comodidad de la siguiente descripción, se supone que la señal de objeto es una señal de objeto estereofónica.

30 Como primer ejemplo, se genera una señal de objeto monofónica submezclando una señal de objeto estereofónica, y la relación entre canales de la señal de objeto estereofónica original se procesa representándola como información sub-lateral. En este caso, información sub-lateral es una terminología que se debe discriminar con respecto a la información lateral convencional, e indica un concepto secundario con respecto a la información lateral en un aspecto jerárquico. En la codificación basada en objetos, si como información lateral se utiliza información de energía del objeto, entonces puede utilizarse la energía de la señal de objeto monofónica como información lateral.

35 Como segundo ejemplo, puede procesarse cada canal de una señal objeto en una única señal de objeto monofónica independiente. Por ejemplo, en caso de que como información lateral se utilice información de energía de una señal de objeto, entonces puede utilizarse la energía de cada canal como información lateral. En este caso, el número de informaciones laterales a transmitir se puede incrementar a un número mayor que el del primer ejemplo.

40 En el caso del primer ejemplo, se puede determinar si se utiliza un coeficiente de canal cruzado de acuerdo con el "método de uso de información de ganancia de mezcla" correspondiente a la cuarta forma de realización antes descrita. En este caso, se puede utilizar información sub-lateral junto con la información de ganancia de mezcla.

45 En el caso del segundo ejemplo, si una señal de objeto de canal izquierdo es  $s_i$ , una señal de objeto de canal derecho puede resultar ser  $s_{i+1}$ . En el caso de la señal de objeto de canal izquierdo, resulta ser  $b_{-1} = 0$ . En el caso de la señal de objeto de canal derecho, resulta ser  $a_{i+1} = 0$ . En particular, en el caso del segundo ejemplo, aunque la señal de objeto se procesa como dos objetos monofónicos, debido a que está incluida solamente en un canal, presenta las características de " $b_{-1} = a_{i+1} = 0$ ".

50 Para llevar a cabo la codificación basada en objetos sobre la señal de objeto estereofónica en el segundo ejemplo, están disponibles los dos siguientes tipos de método.

55 Como primer método, se puede disponer de un caso en el que no se usa un coeficiente de canal cruzado. Por ejemplo, supóngase que la ganancia de mezcla de reproducción viene dada con la fórmula 27.

[FIG. 27]

60

$$c_i = \alpha$$

$$c_{i+1} = \beta$$

65 En el caso de una señal de objeto estereofónica, la misma se puede representar como  $a_{i+1} = 0$ . En este caso, si  $c_{i+1}$  no es cero, una señal de objeto  $s_{i+1}$  incluida en el lado derecho debería incluirse en el lado izquierdo. Por tanto, resultar necesario un coeficiente de canal cruzado.

Sin embargo, en el caso de una señal de objeto estereofónica, se puede suponer que los componentes incluidos en canales respectivos son similares entre sí. Esto se puede representar con la fórmula 28.

5 [Fórmula 28]

$$\begin{aligned} \hat{c}_i &= c_i + c_{i+1}, \\ \hat{c}_{i+1} &= 0 \end{aligned}$$

Por tanto, es posible no utilizar un coeficiente de canal cruzado.

10 Así mismo, un coeficiente de canal cruzado puede no usarse a través del siguiente procesado que se representa con la fórmula 29.

[FIG. 29]

15

$$\begin{aligned} \hat{d}_i &= 0 \\ \hat{d}_{i+1} &= d_i + d_{i+1} \end{aligned}$$

Como segundo método, se puede disponer de un método en el que se utilizan coeficiente de canal cruzado.

20 En caso de intentar incluir en una señal de salida derecha una señal que está incluida en un lado izquierdo de una señal de objeto estereofónica, debe utilizarse un coeficiente de canal cruzado. Por lo tanto, analizando la ganancia de mezcla de reproducción, se puede usar un coeficiente de canal cruzado únicamente si ello es necesario.

25 Como otro ejemplo, en el caso de una señal de objeto estereofónica, se puede usar además la característica de la señal de objeto. En el caso de una señal de objeto estereofónica, una señal sobre una banda de frecuencia específica en una zona de tiempo específica se puede configurar de tal manera que señales muy similares entre sí construyan las señales de canal respectivas. En este caso, si un valor que indica la correlación de una señal de objeto estereofónica en un descodificador es mayor que un umbral, es posible el procesado que se representa con la fórmula 28 ó la fórmula 29 en lugar de usar un coeficiente de canal cruzado.

30

Para analizar la correlación entre canales, se puede usar un método de medición de la coherencia entre canales o similar. Alternativamente, un codificador puede incluir en un flujo continuo de bits información sobre la coherencia entre canales de una señal de objeto estereofónica. Alternativamente, un codificador procesa una señal de objeto estereofónica para obtener una señal monofónica en un dominio del tiempo/frecuencia con una coherencia elevada. Además, el codificador lleva a cabo una codificación sobre la señal de objeto estereofónica procesándola para obtener una señal estereofónica en el dominio del tiempo/frecuencia con una baja coherencia.

35

Como sexta forma de realización, se puede disponer de un método de uso de un coeficiente selectivo.

40 Por ejemplo, se envía una señal izquierda a un canal derecho. Si una señal derecha no está incluida en el canal izquierdo, puede resultar mejor utilizar no  $w_{12}$  sino  $w_{21}$ . Por tanto, en lugar de utilizar cada coeficiente cruzado a pesar de usar coeficientes de canal cruzado, se pueden permitir los cruces necesarios únicamente comprobando la ganancia de una mezcla original y la ganancia de mezcla de reproducción.

45 Tal como se ha mencionado en la anterior descripción, si se cambia la panoramización de un objeto específico, es posible usar un coeficiente de canal cruzado necesario para permitir la panoramización solamente. Si una panoramización de otro objeto se dirige en una dirección opuesta, es posible usar tanto uno como otro de entre los dos coeficientes de canal cruzado.

50 Por ejemplo, en caso de que se usen  $w_{11}$ ,  $w_{12}$  y  $w_{22}$ , es decir, en caso de que no se use  $w_{21}$ , entonces  $w_{11}$ ,  $w_{12}$  y  $w_{22}$  pueden diferir con respecto a los  $w_{11}$ ,  $w_{12}$  y  $w_{22}$  del caso en el que se utilizan cuatro coeficientes  $w_{11}$ - $w_{22}$  en su totalidad. En este caso, tal como se ha mencionado en la anterior descripción,  $w_{11}$ ,  $w_{12}$  y  $w_{22}$  son utilizables modelando  $y_{1\_hat}$  e  $y_{2\_hat}$  y mediante estimación de mínimos cuadrados. En este caso, puesto que se usan  $w_{11}$  y  $w_{12}$ , el  $y_{1\_hat}$  es equivalente al correspondiente de un caso general. Por tanto,  $w_{11}$  y  $w_{12}$  pueden usar los valores previos tal como se encuentren. Sin embargo, puesto que se usa solamente  $w_{22}$ ,  $y_{2\_hat}$  es idéntico al correspondiente del caso en el que se utiliza solamente  $w_2$ . Por tanto,  $w_{22}$  puede usar el correspondiente de la fórmula 11.

55

Por lo tanto, la presente invención propone un método en el que se permite un coeficiente de canal cruzado mono-direccional solamente de acuerdo con las necesidades. Para determinar esto, son utilizables una ganancia de

60

mezcla original y una ganancia de mezcla de reproducción.

Por otra parte, en el caso de que se use un coeficiente de canal cruzado mono-direccional, se puede llevar a cabo de nuevo la estimación de factores de ponderación.

5 Como séptima forma de realización, se puede disponer de un método en el que se usa un coeficiente de canal cruzado solamente.

10 Para una señal de entrada que tiene una característica de panoramización extrema, en caso de que cada señal de objeto se panoramice en una dirección opuesta, el uso de  $w_{21}$  y  $w_{12}$  solamente puede ser más eficiente que el uso de  $w_{11}$ ~ $w_{22}$ . Para usar un coeficiente de canal cruzado solamente, están disponibles las siguientes condiciones. La primera condición se corresponde con si una ganancia de mezcla de una señal de entrada se panoramiza hacia el lateral. La segunda condición se corresponde con si una señal de objeto panoramizada lateralmente se panoramiza en una dirección opuesta. La tercera condición se corresponde con la relación entre el número de objetos que satisfacen tanto la primera como la segunda condiciones y el número total de objetos. Además, una cuarta condición se corresponde con un estado de panoramización original del objeto que no consigue satisfacer tanto la primera como la segunda condiciones y un estado de panoramización solicitado. Sin embargo, en el caso de la cuarta condición, si una panoramización original se panoramiza hacia el lateral y si una panoramización solicitada se panoramiza hacia el mismo lado, puede que no resulte ventajoso usar solamente un coeficiente de canal cruzado.

20 Por otra parte, los diversos métodos antes descritos se pueden utilizar selectivamente juntos o de forma parcial.

La figura 3 es un diagrama de flujo para explicar un método más eficiente de procesamiento de señales de audio, de acuerdo con una forma de realización de la presente invención.

25 En primer lugar, se puede recibir información de mezcla descendente en la cual se submezcla por lo menos una señal de objeto [S310]. Además, se puede obtener información lateral, en la cual se incluye información de objeto, e información de mezcla [S320].

30 En este caso, la información de objeto puede incluir por lo menos una de entre información de nivel de la señal de objeto, información de correlación, información de ganancia y su información suplementaria. La información suplementaria puede incluir información suplementaria de información de nivel, información suplementaria de información de correlación e información suplementaria de información de ganancia. Por ejemplo, la información suplementaria de la información de ganancia puede incluir información de diferencia entre un valor real de la información de ganancia de la señal de objeto y un valor estimado de la misma.

35 La información de mezcla se puede generar basándose en por lo menos una de entre la información de posición, la información de ganancia y la información de configuración de reproducción de la señal de objeto.

40 Se puede generar una pluralidad de informaciones de canales sobre la base de la información lateral y la información de mezcla [S330]. Además, se puede generar una señal de canal de salida a partir de la información de submezcla utilizando la pluralidad de informaciones de canales [S340]. En la siguiente descripción se explican formas de realización detalladas.

45 La figura 4 es un diagrama de bloques esquemático de un aparato de procesamiento de señales de audio para transmitir una señal de objeto más eficientemente, de acuerdo con una forma de realización de la presente invención.

50 En referencia a la figura 4, el aparato de procesamiento de señales de audio puede incluir principalmente un codificador de remezcla mejorada 400, una unidad de codificación de señales de mezcla 430, una unidad de descodificación de señales de mezcla 440, una unidad de generación de parámetros 450 y una unidad de renderización de remezcla 460. Además, el codificador de remezcla mejorada 400 puede incluir una unidad de generación de información lateral 410 y una unidad de codificación de remezcla 420.

55 La información lateral puede ser necesaria para generar factores de ponderación al llevar a cabo la renderización en la unidad de renderización de remezcla 460. Por ejemplo, la información lateral puede incluir valores de estimación de ganancia de mezcla ( $a_{i\_est}$ ,  $b_{i\_est}$ ), ganancias de mezcla de reproducción ( $c_i$ ,  $d_i$ ), energía ( $P_s$ ) de una señal fuente y similares. La unidad de generación de parámetros 450 puede generar los factores de ponderación utilizando la información lateral.

60 De acuerdo con una forma de realización de la presente invención, el codificador de remezcla mejorada 400 puede transmitir el valor de estimación de la ganancia de mezcla ( $a_i$ ,  $b_i$ ), es decir, los valores de estimación de ganancia de mezcla ( $a_{i\_est}$ ,  $b_{i\_est}$ ) como información lateral. Valor de estimación de ganancia de mezcla significa que el valor de ganancia de mezcla ( $a_i$ ,  $b_i$ ) se estima utilizando una señal de mezcla y señales de objeto respectivas. En caso de transmitir el valor de estimación de ganancia de mezcla, se pueden generar factores de ponderación  $w_{11}$ ~ $w_{22}$  utilizando el valor de estimación de ganancia de mezcla y  $c_i/d_i$ . De acuerdo con otra forma de realización, un codificador puede tener un valor real de  $a_i/b_i$  usado para mezcla de hecho señales de objeto respectivas como

65



información independiente. Por ejemplo, en caso de que un codificador genere una señal de mezcla por sí mismo o en caso de que una señal de mezcla se genere externamente, se puede transmitir información independiente de control de mezcla que indica que se usa  $a_i/b_i$  para un valor preestablecido.

- 5 Por ejemplo, si  $c_i/d_i$  significa una escena de remezcla especificada por un usuario y si  $a_i/b_i$  significa una señal mezclada, se puede llevar a cabo una renderización concreta sobre la base de una diferencia entre los dos valores.

10 Por ejemplo, si la información de control indica que  $c_i=1$  y  $d_i=1,5$  para un objeto específico de  $a_i=1$  y  $b_i=1$ , esto puede significar que una señal de canal izquierdo se mantiene intacta ya que ( $a_i \rightarrow c_i$ ), y puede significar que una ganancia de una señal de canal derecho ( $b_i \rightarrow d_i$ ) se amplifica en 0,5.

15 Sin embargo, si se transmiten solamente los valores de estimación de ganancia de mezcla ( $a_{i\_est}$ ,  $b_{i\_est}$ ) en lugar de  $a_i/b_i$  en el ejemplo anterior, puede generarse un problema. Debido a que los valores de estimación de ganancia de mezcla ( $a_{i\_est}$ ,  $b_{i\_est}$ ) se estiman a través del cálculo en el codificador, los mismos pueden tener valores diferentes con respecto a los valores reales  $a_i$  y  $b_i$ , es decir,  $a_{i\_est}=0,9$  y  $b_{i\_est}=1,1$ . En este caso, en el descodificador, a diferencia de la intención real del usuario (amplificación de un canal derecho en 0,5 solamente), el canal izquierdo se amplifica con una ganancia de +0,1 correspondiente a una diferencia entre  $a_{i\_est}$  y  $c_i$ , y el canal derecho se amplifica en +0,4. A saber, el control puede resultar diferente con respecto a la intención del usuario. Por lo tanto, una señal se puede reconstruir más específicamente si se transmiten los valores reales de  $a_i$  y  $b_i$  así como los valores de estimación de ganancia de mezcla ( $a_{i\_est}$ ,  $b_{i\_est}$ ).

20 Al mismo tiempo, si una entrada de usuario se introduce como ganancia y panoramización en lugar de comunicarse por interfaz como  $c_i/d_i$ , un descodificador puede aplicar la ganancia y la panoramización transformando dicha ganancia y dicha panoramización en la forma de  $c_i/d_i$ . En este caso, la transformación se puede llevar a cabo en referencia a  $a_i/b_i$  o  $a_{i\_est}/b_{i\_est}$ .

25 Según otra forma de realización, en caso de que se transmitan  $a_i/b_i$ ,  $a_{i\_est}$  y  $b_{i\_est}$ , los mismos se pueden transmitir en forma de un valor de diferencia entre  $a_i$  y  $a_{i\_est}$  y un valor de diferencia entre  $b_i$  y  $b_{i\_est}$ , en lugar de transmitirse como señales de PCM, respectivamente. Esto es debido a que  $a_i$  y  $a_{i\_est}$  y  $b_i$  y  $b_{i\_est}$  presentan características muy similares. Por ejemplo, se pueden transmitir  $a_i$ ,  $a_{i\_delta}=a_i - a_{i\_est}$ , y  $b_i$ ,  $b_{i\_delta}=b_i - b_{i\_est}$ .

30 Según una forma de realización de la presente invención, en la transmisión de información de mezcla se puede transmitir un valor cuantificado. Por ejemplo, cuando un descodificador lleva a cabo una remezcla utilizando una relación relativa entre  $a_i/b_i$  y  $c_i/d_i$ , el valor realmente transmitido puede ser un valor cuantificado de  $a_{i\_q}/b_{i\_q}$ . En este caso, si el  $a_{i\_q}/b_{i\_q}$  cuantificado se compara con el número real  $c_i/d_i$ , puede generarse un error de nuevo. Por tanto,  $c_i/d_i$  puede utilizar un valor cuantificado de  $c_{i\_q}/d_{i\_q}$  también.

35 Al mismo tiempo, en general un usuario puede introducir  $c_i/d_i$  en un descodificador. Por otra parte, el mismo se puede transmitir como un valor preestablecido incluyéndolo en un flujo continuo de bits. En este caso, el flujo continuo de bits se puede transmitir por separado o junto con información lateral.

40 El flujo continuo de bits transportado desde un codificador puede incluir un flujo continuo de bits individual unificado que contenga una señal de submezcla, información de objeto e información preestablecida. La información de objeto y la información preestablecida se pueden almacenar en un área lateral del flujo continuo de bits de la señal de submezcla. Alternativamente, la información de objeto y la información preestablecida se pueden almacenar o transmitir como una secuencia de bits independiente. Por ejemplo, un primer flujo continuo de bits puede transportar una señal de submezcla. Un segundo flujo continuo de bits puede transportar información de objeto e información preestablecida. De acuerdo con otra forma de realización, un primer flujo continuo de bits puede transportar una señal de submezcla e información de objeto. Además, la información preestablecida puede ser transportada por separado por un segundo flujo continuo de bits. De acuerdo con otra forma de realización, la señal de submezcla, la información de objeto y la información preestablecida pueden ser transportadas por tres flujos continuos de bits independientes, respectivamente.

45 Los flujos continuos de bits primero, segundo e independiente pueden ser idénticos o se pueden transmitir con velocidades de bits diferentes. En particular, después de la reconstrucción de una señal de audio, la información preestablecida se separa con respecto a la señal de submezcla o la información de objeto y a continuación se almacena o transmite.

50 De acuerdo con otra forma de realización de la presente invención,  $c_i/d_i$  puede ser un valor variable en el tiempo si así fuera necesario. En particular, puede ser un valor de ganancia representado en función del tiempo. Por lo tanto, para representar un parámetro de mezcla de usuario que indique una ganancia de mezcla de reproducción en forma de un valor según el tiempo, el mismo se puede introducir como una indicación de tiempo que indique un instante de temporización de aplicación.

55 En este caso, un índice de tiempo puede ser un valor que indique un instante de temporización en un eje del tiempo en el cual se aplica un  $c_i/d_i$  sucesivo. Alternativamente, un índice de tiempo puede ser un valor que indique una

posición de una muestra de una señal de audio mezclada. Alternativamente, en la representación de la señal de audio mediante una unidad de trama, el índice de tiempo puede ser un valor que indique una posición de trama. En caso de un valor de muestra, el mismo se puede representar solamente mediante una unidad de muestra específica.

5 En general, la aplicación de  $c_i/d_i$  correspondiente a un índice de tiempo puede continuar hasta que aparezcan un índice de tiempo y un  $c_i/d_i$  nuevos. Al mismo tiempo, se puede utilizar un valor de intervalo de tiempo en lugar del índice de tiempo. Además, el intervalo de tiempo puede significar una sección en la cual se aplica un  $c_i/d_i$  correspondiente.

10 Por otra parte, se puede definir información de indicador, la cual indica si se va a llevar a cabo una remezcla, dentro de un flujo continuo de bits. Si la información de indicador indica falso,  $c_i/d_i$  no se transmite en una sección correspondiente sino que se puede dar salida a una señal estereofónica según el  $a_i/b_i$  original. En particular, en la sección correspondiente no se puede proceder con un proceso de remezcla. En caso de que se construya un flujo continuo de bits de  $c_i/d_i$  con el método anterior, se puede reducir al mínimo la velocidad de bits. Además, también se  
15 puede evitar la ejecución de una remezcla no deseada.

La figura 5 es un diagrama de flujo para explicar un método de procesado de una señal de objeto utilizando control inverso de acuerdo con una forma de realización de la presente invención.

20 En la ejecución de una codificación basada en objetos, puede existir un caso en el que solamente es necesario controlar señales de objeto parciales. Por ejemplo, como en el caso de a capela, hay disponible un mezclado en el que se deja una señal de objeto específica pero se suprime el resto de señales de objeto. Cuando existe una voz junto con música de fondo, el volumen de la música de fondo se reduce para potenciar la escucha de la voz. Concretamente, el caso anterior puede corresponderse con un caso en el que el número de señales de objeto  
25 cambiadas es mayor que el número de señales de objeto no cambiadas, o con un caso más complicado. Si es así, se lleva a cabo un procesado inverso y a continuación se compensa la ganancia total, con lo cual puede mejorarse adicionalmente la calidad de sonido. Por ejemplo, en el caso de a capela, después de que se haya amplificado solamente una señal de objeto vocal, la ganancia total se puede compensar para que se corresponda con un valor de ganancia de una señal de objeto vocal original.

30 En referencia a la figura 5, en primer lugar, se puede recibir información de submezcla en la cual se submezcla por lo menos una señal de objeto [S510]. Además, se puede obtener información lateral, en la cual se incluye información de objeto, e información de mezcla [S520].

35 En este caso, la información de objeto puede incluir por lo menos una de entre información de nivel de la señal de objeto, información de correlación, información de ganancia y su información suplementaria. La información suplementaria puede incluir información suplementaria de información de nivel, información suplementaria de información de correlación e información suplementaria de información de ganancia. Por ejemplo, la información suplementaria de la información de ganancia puede incluir información de diferencia entre un valor real de la  
40 información de ganancia de la señal de objeto y un valor estimado de la misma. Además, la información de mezcla se puede generar basándose en por lo menos una de entre información de posición, información de ganancia e información de configuración de reproducción de la señal de objeto.

45 La señal de objeto se puede discriminar en una señal de objeto independiente y una señal de objeto de fondo. Por ejemplo, usando información de indicador, se puede determinar si la señal de objeto es una señal de objeto independiente o una señal de objeto de fondo. La señal de objeto independiente puede incluir una señal de objeto vocal. La señal de objeto de fondo puede incluir una señal de objeto de acompañamiento. Además, la señal de objeto de fondo puede incluir por lo menos una señal basada en canales. Por otra parte, usando información de objeto mejorada, se pueden discriminar entre sí la señal de objeto independiente y la señal de objeto de fondo. Por  
50 ejemplo, la información de objeto mejorada puede incluir una señal residual.

Se puede determinar si se lleva a cabo el procesado inverso utilizando la información de objeto y la información de mezcla [S530]. En caso de que el número de objetos cambiados sea mayor que el correspondiente de objetos no cambiados, el procesado inverso significa que la ganancia se compensa en referencia a los objetos no cambiados.  
55 Por ejemplo, en caso de que se intente cambiar la ganancia de un objeto de acompañamiento, si el número de objetos de acompañamiento que se van a cambiar es mayor que el correspondiente de objetos vocales no cambiados, se puede cambiar la ganancia del objeto vocal que presenta el número más pequeño a la inversa. Así, si se lleva a cabo el proceso inverso, se puede obtener un valor de ganancia de procesado inverso para la compensación de ganancia [S540]. Además, se puede generar una señal de canal de salida basándose en el valor de ganancia de procesado inverso [S550].  
60

La figura 6 y la figura 7 son diagramas de bloques de un aparato de procesado de señales de audio para procesar una señal de objeto utilizando control inverso de acuerdo con otra forma de realización de la presente invención.

65 En referencia a la figura 6, el aparato de procesado de señales de audio puede incluir una unidad de control de proceso inverso 610, una unidad de generación de parámetros 620, una unidad de renderización de remezcla 630 y

una unidad de procesado inverso 640.

La determinación sobre si llevar a cabo un procesado inverso la puede ejecutar la unidad de control de proceso inverso 610 utilizando  $a_i/b_i$  y  $c_i/d_i$ . Si, de acuerdo con la determinación, se lleva a cabo el procesado inverso, la unidad de generación de parámetros 620 genera factores de ponderación correspondientes  $w_{11}\sim w_{22}$ , calcula un valor de ganancia de procesado inverso según la compensación de ganancia, y a continuación transmite el valor calculado a la unidad de procesado inverso 640. Además, la unidad de renderización de remezcla 630 lleva a cabo una renderización basándose en los factores de ponderación.

Por ejemplo, supóngase que  $a_i/b_i$  y  $c_i/d_i$  se proporcionan en la forma siguiente:  $a_i/b_i = \{1/1, 1/1, 1/0, 0/1\}$ ; y  $c_i/d_i = \{1/1, 0, 1/0, 1, 0, 1/0, 0/0, 1\}$ . Esto suprimirá el resto de señales de objeto en 1/10 excepto una primera señal de objeto. Si es así, se puede obtener una señal más próxima a una señal más específica utilizando la siguiente relación inversa de factores de ponderación ( $c_{i\_rev}/d_{i\_rev}$ ) y una ganancia de procesado inverso. En este caso,  $c_{i\_rev}/d_{i\_rev} = \{10/10, 1/1, 1/0, 0/1\}$  y  $reverse\_gain = 0,1$ .

Según otra forma de realización de la presente invención, en un flujo continuo de bits se puede incluir información de indicador que indique la complejidad de una señal de objeto específica. Por ejemplo, se puede definir un `complex_object_flag` que indique la presencia o no presencia de complejidad de una señal de objeto. La presencia o no presencia de complejidad se puede determinar en referencia a un valor fijo o a un valor relativo.

Por ejemplo, supóngase que una señal de audio incluye dos señales de objeto, siendo una de las señales de objeto música de fondo, tal como acompañamiento de MR (música grabada), y siendo la otra voces. La música de fondo puede ser una señal de objeto complicado que se ha construido con la combinación de instrumentos musicales, en un nivel mucho mayor que la voz. En este caso, si se transmite la información de `complex_object_flag`, la unidad de control de proceso inverso puede determinar si se lleva a cabo el procesado inverso de una manera sencilla. En particular, si  $c_i/d_i$  solicita una implementación a capela suprimiendo la música de fondo en -24 dB, se puede generar una señal específica amplificando la voz en +24 dB a la inversa y a continuación fijando una ganancia de procesado inverso a -24 dB, de acuerdo con la información del indicador. Este método se puede aplicar en conjunto a un tiempo completo o a bandas completas, o se puede aplicar selectivamente solo a un tiempo específico o a una banda específica.

En la siguiente descripción, se explica un método de ejecución del procesado inverso en caso de que exista una panoramización extrema de acuerdo con otra forma de realización de la presente invención.

Por ejemplo, se puede recibir una solicitud de remezcla para desplazar la mayoría de objetos de un canal izquierdo a la derecha, y desplazar objetos de un canal derecho a la izquierda. En este caso, en lugar del método antes descrito, puede resultar más eficiente llevar a cabo una remezcla en un estado de intercambio después de intercambiar los canales izquierdo y derecho.

En referencia a la figura 7, el aparato de procesado de señales de audio puede incluir una unidad de control de proceso inverso 710, una unidad de intercambio de canales 720, una unidad de renderización de remezcla 730 y una unidad de generación de parámetros 740.

La unidad de control de proceso inverso 710 puede determinar si se intercambian señales de objeto a través del análisis de  $a_i/b_i$  y  $c_i/d_i$ . Si resulta preferible, de acuerdo con la determinación, llevar a cabo el intercambio, la unidad de intercambio de canales 720 realiza el intercambio de canales. La unidad de renderización de remezcla 730 lleva a cabo una renderización utilizando la señal de audio con canales intercambiados. En este caso, pueden generarse factores de ponderación  $w_{11}\sim w_{22}$  en referencia a los canales intercambiados.

Por ejemplo, supóngase que  $a_i/b_i = \{1/0, 1/0, 0,5/0,5, 0/1\}$  y  $c_i/d_i = \{0/1, 0,1/0,9, 0,5/0,5, 1/0\}$ . Si se va a llevar a cabo la panoramización anterior, debería realizarse una panoramización muy extrema sobre las señales de objeto 1<sup>a</sup>, 2<sup>a</sup> y 4<sup>a</sup>. En este caso, si se lleva a cabo un intercambio de canales según la presente invención, no es necesario cambiar las señales de objeto 1<sup>a</sup>, 3<sup>a</sup> y 4<sup>a</sup> pero la señal de objeto 2<sup>a</sup> se debe ajustar de forma precisa.

Este método se puede aplicar en conjunto a un tiempo completo o a bandas completas, o se puede aplicar selectivamente a un tiempo específico o a una banda específica solamente.

Se propone un método de procesado eficiente de señales de objeto que tienen una alta correlación de acuerdo con otra forma de realización de la presente invención.

Con frecuencia puede suceder que las señales de objeto para la remezcla incluyan señales de objeto estereofónicas. En caso de la señal de objeto estereofónica, se transmite un parámetro independiente considerando cada canal (L/R) como un objeto monofónico independiente, y se puede llevar a cabo la remezcla utilizando el parámetro transmitido. Al mismo tiempo, en la remezcla, se puede transmitir información que indique qué tipos de dos objetos se acoplan para una señal de objeto estereofónica con el fin de construir la señal de objeto estereofónica. Por ejemplo, la información se puede definir como `src_type`. Además, se puede transmitir el `src_type`

por cada objeto.

En otra situación, puede darse el caso de que las señales de canal izquierdo y derecho entre las señales de objeto estereofónicas tengan de hecho prácticamente el mismo valor. En este caso, la gestión de la señal de canal izquierdo/derecho como una señal de objeto monofónica facilita el remezclado, en lugar de gestionar la señal de canal izquierdo/derecho como una señal de objeto estereofónica, y esto puede reducir la velocidad de bits requerida para la transmisión.

Por ejemplo, si se introduce una señal de objeto estereofónica, puede determinarse si considerar la misma como una señal de objeto monofónica o una señal de objeto estereofónica dentro de un codificador de remezcla. Además, en una secuencia de bits se puede incluir un parámetro correspondiente. En este caso, en caso de procesarla como señal de objeto estereofónica, es necesario un par de  $a_i/b_i$  para los canales izquierdo y derecho, respectivamente. En este caso, es preferible que el  $b_i$  para el canal izquierdo sea cero. Además, es preferible que el  $a_i$  para el canal derecho sea cero. Por otra parte, también es necesario un par de potencias ( $P_s$ ) de fuente.

Para otra situación, si las señales de objeto izquierda y derecha son sustancialmente las mismas señales o si son señales que tienen una correlación elevada, se puede generar una señal de objeto virtual resultante de una suma de las dos señales. Por otra parte,  $a_i/b_i$  y  $P_s$  se generan y transmiten en referencia a la señal de objeto virtual. Si  $a_i/b_i$  y  $P_s$  se transmiten con un método del tipo mencionado, se puede reducir la velocidad de bits. Cuando se lleva a cabo una renderización en un descodificador, se pueden omitir acciones de panoramización innecesarias. Por lo tanto, el descodificador puede funcionar de forma más estable.

En este caso, una señal de submezcla monofónica puede generarse de diversas maneras. Por ejemplo, puede disponerse de un método en el que se suman entre sí una señal de objeto izquierda y una señal de objeto derecha. Alternativamente, puede disponerse de un método en el que la señal de objeto sumada se divide por un valor de ganancia normalizado. Por tanto, los valores del  $a_i/b_i$  y  $P_s$  transmitidos pueden variar de acuerdo con cómo se generan.

Por otra parte, se puede transmitir información que tiene la capacidad de discriminar si una señal de objeto específica es monofónica o estereofónica, o si una señal de objeto específica, que era estereofónica, se renderiza en una señal monofónica por medio de un codificador. En este caso, se puede mantener la compatibilidad en caso de que  $c_i/d_i$  se comuniquen por interfaz en un descodificador. Por ejemplo, en el caso monofónico, se puede determinar  $src\_type=0$ . En caso de una señal de canal izquierdo en estereofonía, se puede determinar  $src\_type=1$ . En caso de una señal de canal derecho en estereofonía, se puede determinar  $src\_type=2$ . En caso de submezcla una señal estereofónica en una señal monofónica, se puede determinar  $src\_type=3$ .

Al mismo tiempo, un descodificador puede recibir  $c_i/d_i$  para una señal de canal izquierdo y  $c_i/d_i$  para una señal de canal derecho, con vistas al control de una señal de objeto estereofónica. En caso de " $src\_type=3$ " de la señal de objeto, puede resultar preferible que el  $c_i/d_i$  correspondiente a la señal de canal izquierdo y el  $c_i/d_i$  correspondiente a la señal de canal derecho se sumen entre sí. El tipo de adición puede adoptar el método de generar la señal de objeto virtual.

Este método se puede aplicar en conjunto a un tiempo completo o a bandas completas, o se puede aplicar selectivamente solo a un tiempo específico o una banda específica.

De acuerdo con otra forma de realización de la presente invención, en caso de que cada señal de objeto se adapte a cada señal de canal según 1:1, se puede reducir la cantidad de transmisión utilizando información de indicador. En este caso, la renderización se puede llevar a cabo a través de un proceso de mezcla simple en lugar de aplicar cada algoritmo de remezcla para la renderización concreta.

Por ejemplo, si hay dos señales de objeto Obj 1 y Obj 2, y si  $a_i/b_i$  para el Obj 1 y el Obj 2 es  $\{1/0, 0/1\}$ , Obj 1 existe en una señal de canal izquierdo de una señal mezclada solamente y Obj 2 existe en una señal de canal derecho de la señal mezclada solamente. En este caso, puesto que, a partir de la señal mezclada, se puede extraer una potencia de fuente ( $P_s$ ), no es necesario transmitir la misma por separado. Por otra parte, en caso de que se lleve a cabo una renderización, se pueden obtener directamente factores de ponderación ( $w_{11}\sim w_{22}$ ) a partir de las relaciones de  $c_i/d_i$  y  $a_i/b_i$ , y no se solicita por separado una operación usando PS. Por lo tanto, en el caso del ejemplo anterior, el procesamiento se facilita adicionalmente utilizando información de indicador relevante.

La figura 8 es un diagrama estructural de un flujo continuo de bits que contiene meta-información sobre un objeto de acuerdo con una forma de realización de la presente invención.

En la codificación de audio basada en objetos, se puede recibir meta-información sobre un objeto. Por ejemplo, en el proceso para submezclar una pluralidad de objetos en señales monofónicas o estereofónicas, de cada una de las señales de objeto se puede extraer meta-información. Además, la meta-información se puede controlar por medio de una selección realizada por un usuario.

En este caso, meta-información puede significar meta-datos. En particular, los meta-datos son los datos sobre datos y pueden significar datos para describir el atributo de un recurso de información. Concretamente, los meta-datos, que no son los propios datos (por ejemplo, vídeo, audio, etcétera) que se van a almacenar sustancialmente, significan datos para proporcionar información asociada de manera directa o indirecta a los datos correspondientes.

5 Si se utilizan dichos meta-datos, se puede comprobar si datos específicos de cada usuario son correctos, y se pueden encontrar datos específicos de manera sencilla y rápida. Concretamente, se garantiza una simplificación de la gestión en términos de posesión de datos, o se garantiza una simplificación de la búsqueda en términos de usar datos.

10 En la codificación de audio basada en objetos, meta-información puede significar la información que indica atributos de un objeto. Por ejemplo, la meta-información puede indicar si cada una de una pluralidad de señales de objeto que constituyen una fuente de sonido se corresponde con un objeto vocal o un objeto de fondo. Además, la meta-información puede indicar si el objeto vocal es un objeto para un canal izquierdo o un canal derecho. Por otra parte, la meta-información puede indicar si el objeto de fondo se corresponde con un objeto de piano, un objeto de batería, un objeto de guitarra u otro objeto de instrumento musical.

15 Al mismo tiempo, flujo continuo de bits puede significar un grupo de parámetros o datos, o puede significar un flujo continuo de bits general comprimido para su transmisión o almacenamiento. Por otra parte, el flujo continuo de bits se puede interpretar en un sentido amplio para indicar un tipo de parámetro antes de representarlo como flujo continuo de bits. Un dispositivo de decodificación puede obtener información de objeto a partir del flujo continuo de bits basado en objetos. En la siguiente descripción, se explicará información incluida en el flujo continuo de bits basado en objetos.

20 En referencia a la figura 8, un flujo continuo de bits basado en objetos puede incluir un encabezamiento y datos. El encabezamiento 1 puede incluir meta-información, información de parámetros y similares. La meta-información puede incluir la siguiente información. Por ejemplo, la meta-información puede incluir un nombre de objeto, un índice de objeto que indica un objeto, información de atributos detallada sobre el objeto (característica del objeto), información sobre el número de objetos, información de descripción de meta-datos, información sobre el número de caracteres de los meta-datos (número de caracteres), información de caracteres de los meta-datos (un único carácter), información de indicador de meta-datos y similares.

25 En este caso, nombre del objeto puede significar la información que indica el atributo de dicho objeto como objeto vocal, objeto de instrumento musical, objeto de guitarra, objeto de piano y similares. El índice de objeto que indica un objeto puede significar la información para asignar un índice a información de atributos sobre el objeto. Por ejemplo, se asigna un índice a cada nombre de instrumento musical para definir una tabla de antemano. La información de atributos detallada sobre el objeto (característica del objeto) puede significar información de atributo individual sobre un subobjeto. En este caso, sub-objeto puede significar cada uno de entre objetos similares, cuando los objetos similares se agrupan en un único objeto de grupo. Por ejemplo, en el caso de un objeto vocal, existe información que indica un objeto de canal izquierdo e información que indica un objeto de canal derecho.

30 Por otra parte, la información de número de los objetos (número de objeto) puede significar el número de objetos para transmitir parámetros de señales de audio basadas en objetos. La información de descripción de meta-datos puede significar la información de descripción de meta-datos para un objeto codificado. La información de caracteres de meta-datos (un único carácter) puede significar cada carácter de los meta-datos de un único objeto. La información de indicador de meta-datos puede significar un indicador que indica si se transmitirá información de meta-datos de objetos codificados.

35 Al mismo tiempo, la información de parámetros puede incluir una frecuencia de muestreo, el número de subbandas, el número de señales fuente, un tipo de fuente y similares. Además, la información de parámetros puede incluir selectivamente información de configuración de reproducción de una señal fuente.

40 Los datos pueden incluir por lo menos datos de una trama. Si fuera necesario, los datos pueden incluir un encabezamiento (Encabezamiento 2) junto con los datos de trama. En este caso, el Encabezamiento 2 puede incluir informaciones que es necesario actualizar.

45 Los datos de trama pueden incluir información sobre un tipo de datos incluido en cada trama. Por ejemplo, en caso de un primer tipo de datos (Tipo 0), los datos de trama pueden incluir información mínima. En particular, los datos de trama pueden incluir la potencia de fuente asociada a información lateral solamente. En caso de un segundo tipo de datos (Tipo 1), los datos de trama pueden incluir adicionalmente ganancias actualizadas. En caso de un tercer o cuarto tipo de datos, los datos de trama se pueden asignar como área reservada para un uso futuro. Si el flujo continuo de bits se usa para una emisión de radiodifusión, el área reservada puede incluir información (por ejemplo, frecuencia de muestreo, número de subbanda, etcétera) necesaria para encontrar una sintonización de una señal de radiodifusión.

50 La figura 9 es un diagrama de estructura sintáctica para transmitir una señal de audio eficientemente, de acuerdo con una forma de realización de la presente invención.

Se transportan tantas potencias de fuente ( $P_s$ ) como número de particiones (bandas de frecuencias) dentro de una trama. La partición es una banda no uniforme basada en un modelo de sonido psicológico. Además, en general se usan aproximadamente 20 particiones. Por tanto, se transportan 20 potencias de fuente por cada señal fuente. Toda potencia de fuente cuantificada tiene un valor positivo. Además, el transporte de la potencia de fuente por codificación diferencial resulta más ventajoso que el transporte de la potencia de fuente como una señal de PCM lineal. Por otra parte, la potencia de fuente se puede transportar selectivamente seleccionando una codificación óptima de entre codificación diferencial en el tiempo, codificación diferencial en frecuencia y PBC (codificación basada en valores piloto). En el caso de una fuente estereofónica, se puede enviar un valor de diferencia desde una fuente acoplada. En este caso, el valor de diferencia de la potencia de fuente puede tener un signo positivo o negativo.

El valor de potencia de fuente codificado diferencialmente se transporta a través de una codificación Huffman. En este caso, una tabla de codificación Huffman incluye una tabla que trata solamente con valores positivos, o una tabla que trata con valores tanto positivos como negativos. En caso de usar una tabla sin signos que tenga solamente los valores positivos, se transporta por separado un bit correspondiente al signo.

La presente invención propone un método de transporte de un bit de signo en el uso de una tabla de Huffman sin signos.

Sin tener que transportar un bit de signo para cada muestra de valor diferencia, se puede(n) transportar en conjunto bit(s) de signo para 20 valores de diferencia correspondientes a una única partición. En este caso, se puede transportar un indicador `uni_sign` que indica si se usa un mismo signo para el(los) bit(s) de signo transportado(s). Si el `uni_sign` se fija a 1, esto significa que los signos de los 20 valores de diferencia son iguales entre sí. En ese caso, sin tener que transportar 1 bit de signo por cada muestra, se transporta solamente un bit de signo completo de un bit. Si el `uni_sign` se fija a 0, se transporta un bit de signo por valor de diferencia. En este caso, el bit de signo no se transporta para una muestra que tenga el valor de diferencia fijado a 0. Si los 20 valores de diferencia son todos ellos cero, no se transporta el indicador `uni_sign`.

Con el método anterior, se puede reducir el número de bits requeridos para la transmisión de bits de signo en un área en la que los signos tienen los mismos valores de diferencia, respectivamente. En caso de un valor de potencia de fuente real, puesto que una señal fuente tiene una característica transitoria en el dominio del tiempo, el valor de diferencia en el tiempo presenta frecuentemente un único signo. Por lo tanto, el método de transmisión de señales de acuerdo con la presente invención tiene una buena eficiencia.

Las figuras 10 a 12 son diagramas para explicar un proceso de codificación sin pérdidas, con el fin de transmitir potencia de fuente de acuerdo con una forma de realización de la presente invención.

En referencia a la figura 10, se muestra un proceso de codificación sin pérdidas para transmitir una potencia de fuente. Después de que se haya generado una señal diferencial sobre un eje de tiempo o de frecuencia, se lleva a cabo una codificación sobre un valor de PCM diferencial utilizando el libro de códigos de Huffman más ventajoso en el aspecto de la compresión.

En caso de que todos los valores diferenciales sean cero, esto se puede considerar como un caso de Huff\_AZ. En este caso, los valores de diferencia no se transmiten realmente, y un descodificador puede saber que todos ellos son cero por el hecho de que se ha adoptado Huff\_AZ. Es relativamente probable que la magnitud de un valor diferencial sea pequeña. Además, es también relativamente probable que un valor diferencial tenga un valor de cero. Por lo tanto, el método de codificación de Huffman 2D/4D para codificar cada par de dos o cuatro valores diferenciales puede resultar eficiente. Los valores atributos máximos para la codificación por tabla pueden diferir entre sí. En general, es preferible que la tabla de 4D tenga un valor máximo muy bajo fijado a 1.

En caso de la codificación Huffman sin signos, es aplicable el método de codificación con signos que utiliza el `uni_sign` antes mencionado.

Al mismo tiempo, hay disponible selectivamente una tabla de Huffman en cada dimensión, de entre una pluralidad de tablas que tienen entre sí características y estadísticas diferentes. Además, se puede usar una tabla diferente según `FREQ_DIFF` o `TIME_DIFF`. Dentro de un flujo continuo de bits se puede incluir por separado un indicador que indique qué tipo de señal diferencial o codificación Huffman se usa.

Para minimizar el derroche en el uso de bits, se puede definir que no se utiliza una combinación específica de métodos de codificación utilizando un indicador. Por ejemplo, si la combinación de `Freq_diff` y `Huff_4D` apenas se usa, no se adopta la codificación según la combinación correspondiente.

Puesto que la combinación de indicadores se usa frecuentemente, se pueden comprimir adicionalmente datos transmitiendo un índice correspondiente a través de la codificación Huffman.

En referencia a la figura 11, se muestra otro ejemplo de un método de codificación sin pérdidas. En un método de codificación diferencial, pueden existir varios ejemplos. Por ejemplo, CH\_DIFF es un método de transmisión que utiliza un valor diferencial entre fuentes correspondientes a canales de una señal de objeto estereofónica. Además, puede haber una codificación diferencial basada en valores piloto, una codificación diferencial en el tiempo y similares. En el caso de la codificación diferencial en el tiempo, se añade un método de codificación en el cual se selecciona el uso de FWD BWD. En el caso de la codificación Huffman, se añade una codificación Huffman con signos.

En general, en el procesado de una señal de objeto estereofónica, se puede procesar cada canal de una señal de objeto como una señal de objeto independiente. Por ejemplo, el procesado se puede llevar a cabo de tal manera que se considere una primera señal de canal (por ejemplo, un canal izquierdo) como señal de objeto monofónica independiente de  $s_i$ , y se considere una segunda señal de canal (por ejemplo, un canal derecho) como una señal de objeto monofónica independiente de  $s_{i+1}$ . En ese caso, la potencia de una señal de objeto transportada resulta ser  $Ps_i$  o  $Ps_{i+1}$ . Sin embargo, en el caso de una señal de objeto estereofónica, las características entre dos canales son frecuentemente similares entre sí. Por lo tanto, puede resultar ventajoso que tanto  $Ps_i$  como  $Ps_{i+1}$  se consideren juntos en la codificación. La figura 10 muestra un ejemplo para este acoplamiento. La codificación de  $Ps_i$  sigue el método que se muestra en la figura 8 y la figura 9, la codificación de  $Ps_{i+1}$  encuentra una diferencia entre  $Ps_i$  y  $Ps_{i+1}$ , y la diferencia se codifica y se transmite.

A continuación se explica un método de procesado de una señal de audio usando la similitud entre canales, de acuerdo con otra forma de realización de la presente invención.

Como primera forma de realización, se puede disponer de un método en el que se usan potencias de fuente y una diferencia de nivel entre canales. La potencia de fuente de un canal específico se cuantifica, y a continuación es enviada. La potencia de fuente de otro canal se puede obtener a partir de un valor relativo a la potencia de fuente del canal específico. En este caso, el valor relativo puede incluir una relación de potencia (por ejemplo,  $Ps_{i+1}/Ps_i$ ) o un valor diferencial entre valores resultantes de aplicar el logaritmo sobre valores de potencia. Por ejemplo, el valor diferencial incluye  $10\log_{10}(Ps_{i+1}) - 10\log_{10}(Ps_i) = 10\log_{10}(Ps_{i+1}/Ps_i)$ . Alternativamente, se puede transmitir un valor de diferencia de índice después de la cuantificación.

Si se usa la anterior forma, las potencias de fuente de canales de una señal estereofónica tienen valores muy similares entre sí. Además, resulta muy ventajosa para la cuantificación y la transmisión con compresión. Si el valor diferencial se halla antes de la cuantificación, se puede transmitir una potencia de fuente más precisa.

Como segunda forma de realización, se puede disponer de un método en el que se usa la potencia de fuente o una suma y diferencia de una señal original. En este caso, la eficiencia de transmisión es mejor que en la transmisión de una señal de canal original. Además, puede resultar eficiente en el aspecto del equilibrio del error de cuantificación.

En referencia a la figura 12, se puede usar un acoplamiento solamente para un dominio de frecuencia específico. Además, en un flujo continuo de bits se puede incluir información sobre un dominio de frecuencia en el que ha tenido lugar el acoplamiento. En general, por ejemplo, los canales izquierdo y derecho tienen características similares en una señal sobre una banda de frecuencia baja. Además, en una señal sobre una banda de frecuencia alta puede haber una gran diferencia entre los canales izquierdo y derecho. Por lo tanto, si se lleva a cabo el acoplamiento sobre una banda de frecuencia, se puede aumentar la eficiencia de compresión. A continuación se explican varios métodos de llevar a cabo el acoplamiento.

Por ejemplo, el acoplamiento se puede llevar a cabo sobre una señal en una banda de baja frecuencia solamente. En este caso, puesto que el acoplamiento se lleva a cabo solamente sobre una banda preestablecida, resulta innecesario transmitir por separado información sobre la banda en la cual se aplica el acoplamiento. Alternativamente, se puede disponer de un método de transmisión de información en una banda en la que se ha llevado a cabo el acoplamiento. Un codificador determina arbitrariamente una banda sobre la cual llevar a cabo el acoplamiento, y la información sobre la banda en la que se lleva a cabo el acoplamiento se puede incluir en un flujo continuo de bits.

Alternativamente, se puede disponer de un método en el que se usa un índice de acoplamiento. Se asigna el índice a una posible combinación de bandas en las que se produce acoplamiento y a continuación el índice se transmite. Por ejemplo, en caso de que el procesado se lleve a cabo dividiendo una banda en 20 bandas de frecuencia, se puede saber qué bandas se acoplan de acuerdo con el índice que se muestra en la tabla 1.

[Tabla 1]

índice	0	1	2	3
acoplamiento	0~3 banda	0~7 banda	0~12 banda	0~19 banda

Como índice se puede usar un índice predeterminado. Alternativamente, se puede transmitir una tabla de índices

determinando un valor óptimo de un contenido correspondiente. Alternativamente, se puede usar un valor independiente para cada señal de objeto estereofónica.

5 A continuación se explica un método de obtención de información que indica correlación entre objetos agrupados según una forma de realización de la presente invención.

10 En primer lugar, en el procesado de una señal de audio basada en objetos, un objeto individual que construye una señal de entrada se procesa como objeto independiente. Por ejemplo, en el caso de una señal estereofónica que constituya una voz, la señal de canal izquierdo o la señal de canal derecho se procesa reconociéndola, cada una de ellas, como un objeto individual. Si una señal de objeto se configura con este método, puede existir correlación entre objetos que tengan el mismo origen. Si se lleva a cabo una codificación usando la correlación, será posible una codificación más eficiente. Por ejemplo, puede existir correlación entre un objeto construido con una señal de canal izquierdo de una señal estereofónica y un objeto construido con una señal de canal derecho de la misma. Además, se transmite información sobre la correlación para su uso.

15 Agrupando objetos que presentan la correlación entre ellos y transmitiendo información común a los objetos agrupados una vez, es posible una codificación más eficiente.

20 Cuando un objeto individual forma parte de un objeto estereofónico o de diversos canales, *bsRelatedTo*, que es la información transportada por un flujo continuo de bits, puede ser la información que indica otros objetos que se corresponden con una parte del mismo objeto estereofónico o de diversos canales. *bsRelatedTo* puede obtener información de 1 bit a partir de un flujo continuo de bits. Por ejemplo si *bsRelatedTo[i][j]=1*, esto puede significar que el objeto *i* y *j* se corresponden con canales del mismo objeto estereofónico u objeto de diversos canales.

25 Basándose en el valor de *bsRelatedTo*, se puede comprobar si objetos constituyen un grupo. Comprobando el valor de *bsRelatedTo* para cada objeto, se puede comprobar la información sobre correlación entre objetos. Para los objetos agrupados en los que existe correlación, es posible una codificación más eficiente transmitiendo la misma información (por ejemplo, meta-información) una vez.

30 La figura 13 es un diagrama para explicar una interfaz de usuario de acuerdo con una forma de realización de la presente invención.

35 En primer lugar, una ventana de control principal puede incluir un área de lista musical, un área de control de reproducción general y un área de control de remezcla. Por ejemplo, el área de lista musical puede incluir por lo menos una muestra musical. El área de control de reproducción general puede controlar la Reproducción, la Pausa, la Parada, el FF (avance rápido), Reb (rebobinado), el Cambio de Posición, el Volumen y similares. El área de control de remezcla puede incluir un área de sub-ventana. El área de sub-ventana puede incluir un área de control mejorada. Además, en el área de control mejorada se puede controlar un elemento específico de cada usuario.

40 En el caso de un reproductor de CD, un usuario puede escuchar la música cargando un CD en el reproductor de CD. En el caso de un reproductor de PC, si un usuario carga un disco en un PC, se ejecuta automáticamente un reproductor de remezcla. Además, la música a reproducir se puede seleccionar a partir de una lista de archivos del reproductor. El reproductor lee la fuente de sonido de PCM grabada en el CD y un archivo \*.rms para su reproducción automática. La capa puede llevar a cabo un control de remezcla completo, así como un control de reproducción general. Como ejemplos del control de remezcla completo, están el control de pistas o un control de panoramización. Además, puede haber disponible un control de remezcla sencillo. En caso de entrar en un modo de control de remezcla sencillo, son controlables varias funciones. Por ejemplo, el modo de control de remezcla sencillo puede significar una ventana de control sencillo con capacidad de controlar de manera sencilla un objeto específico, tal como un karaoke y música a capela. En el área de sub-ventana, un usuario puede llevar a cabo un control detallado.

50 Tal como se ha mencionado en la anterior descripción, un aparato de procesado de señales según la presente invención se proporciona a un transmisor/receptor de radiodifusión multimedia, tal como DMB (radiodifusión multimedia digital), y se usa en la decodificación de una señal de audio, una señal de datos y similares. Por otra parte, el transmisor/receptor de radiodifusión multimedia puede incluir un terminal móvil de comunicaciones.

55 Por otra parte, un aparato de procesado de señales según la presente invención se puede implementar en un programa grabado en un soporte en forma de códigos legibles por ordenador. Los soportes legibles por ordenador incluyen todos los tipos de dispositivos de grabación en los cuales se almacenan datos legibles por un sistema de ordenador. Los soportes legibles por ordenador incluyen ROM, RAM, CD-ROM, cintas magnéticas, discos flexibles, dispositivos de almacenamiento óptico de datos, y similares, por ejemplo, e incluyen también implementaciones de tipo onda portadora (por ejemplo, transmisión por Internet). Además, un flujo continuo de bits generado por el método de procesado de señales se almacena en un soporte de grabación legible por ordenador o se puede transmitir a través de una red de comunicaciones por cable/inalámbricas.

65



**REIVINDICACIONES**

Método de procesado de una señal de audio, que comprende:

- 5 recibir (S310) una señal de submezcla que comprende por lo menos una señal de objeto;
- obtener (S320) información lateral que incluye información de objeto;
- obtener (S320) información de mezcla,  $ci/di$ ;
- 10 generar (S330) una pluralidad de informaciones de canales basándose en la información lateral y la información de mezcla; y
- 15 generar (S340) una señal multicanal a partir de la señal de submezcla utilizando la pluralidad de informaciones de canales,

en el que:

- 20 la información de objeto incluye por lo menos una de entre información de nivel de la señal de objeto, información de correlación de la señal de objeto, e información de ganancia de la señal de objeto, y
- la información de mezcla ( $ci/di$ ) es por lo menos una de entre información introducida por un usuario en un descodificador y transmitida como un valor preestablecido incluido en un flujo continuo de bits desde un codificador, el flujo continuo de bits es un flujo continuo de bits individual unificado que contiene la señal de submezcla, la información de objeto y el valor preestablecido.
- 25

2. Método según la reivindicación 1, en el que la información lateral incluye un valor de ganancia de mezcla ( $a_i, b_i$ ), siendo el valor de ganancia de mezcla ( $a_i, b_i$ ) un valor para determinar una panoramización en amplitud y una ganancia para cada señal de objeto.

3. Método según la reivindicación 2, en el que, si el valor de ganancia de mezcla,  $a_i, b_i$ , es un valor cuantificado ( $a_{i\_q}, b_{i\_q}$ ), la información de mezcla ( $ci/di$ ) es cambiada a un valor cuantificado ( $ci\_q, di\_q$ ) para comparar el valor cuantificado ( $a_{i\_q}, b_{i\_q}$ ).

4. Soporte legible por ordenador que comprende unas partes de código que, cuando son ejecutadas en un procesador, configuran el procesador para llevar a cabo todas las etapas de un método según cualquiera de las reivindicaciones de método anteriores.

5. Aparato (100) para procesar una señal de audio, que comprende:

- 40 una unidad de procesado de submezcla (120) configurada para recibir una señal de submezcla que comprende por lo menos una señal de objeto;
- 45 una unidad de generación de información (110) configurada para:
  - obtener información lateral que incluye información de objeto,
  - obtener información de mezcla,  $ci/di$ , y
  - 50 - generar una pluralidad de informaciones de canales basándose en la información lateral obtenida y la información de mezcla obtenida; y

una unidad de descodificación multicanal (130) configurada para generar una señal multicanal a partir de la señal de submezcla utilizando la pluralidad de informaciones de canales,

55 en el que:

- 60 la información de objeto incluye por lo menos una de entre información de nivel de la señal de objeto, la información de correlación de la señal de objeto y la información de ganancia de la señal de objeto, y
- la información de mezcla ( $ci/di$ ) es por lo menos una de entre información introducida por un usuario y transmitida como un valor preestablecido incluido en un flujo continuo de bits desde un codificador, el flujo continuo de bits es un flujo continuo de bits individual unificado que contiene la señal de submezcla, la información de objeto y el valor preestablecido.
- 65

6. Aparato según la reivindicación 5, en el que la información lateral incluye un valor de ganancia de mezcla ( $a_i, b_i$ ),

siendo el valor de ganancia de mezcla ( $a_i$ ,  $b_i$ ) un valor para determinar una panoramización en amplitud y una ganancia para cada señal de objeto.

- 5 7. Aparato según la reivindicación 6, en el que, si el valor de ganancia de mezcla,  $a_i$ ,  $b_i$ , es un valor cuantificado ( $a_{i\_q}$ ,  $b_{i\_q}$ ), la información de mezcla,  $c_i/d_i$ , es cambiada a un valor cuantificado ( $c_{i\_q}$ ,  $d_{i\_q}$ ) para comparar el valor cuantificado ( $a_{i\_q}$ ,  $b_{i\_q}$ ).

FIG. 1

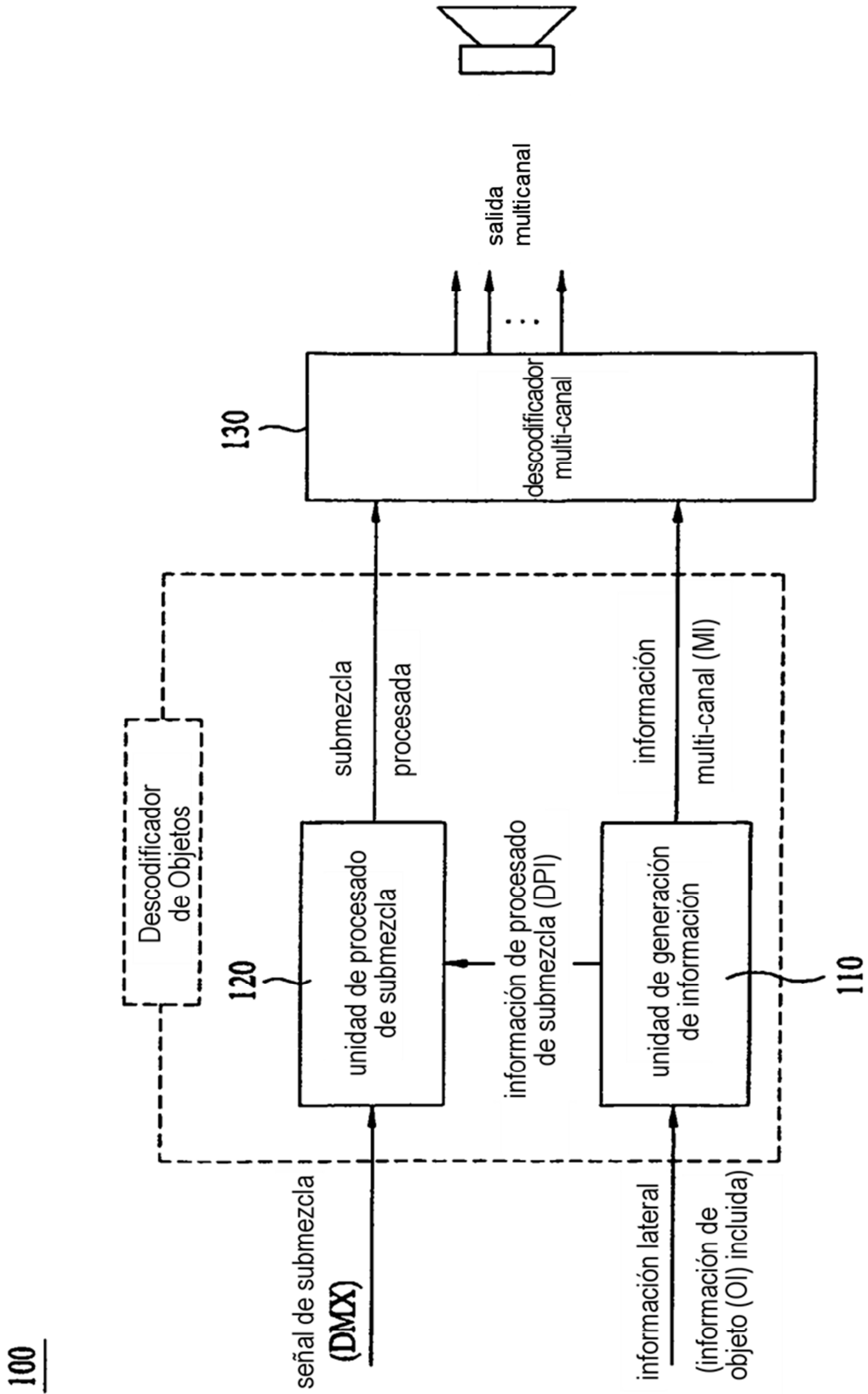
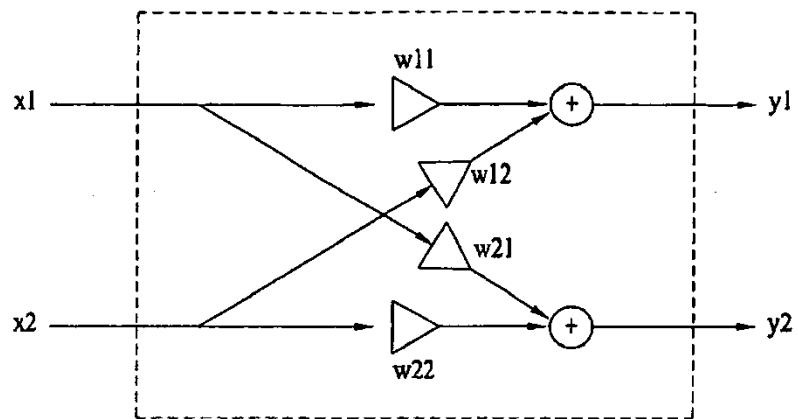


FIG. 2



**FIG. 3**

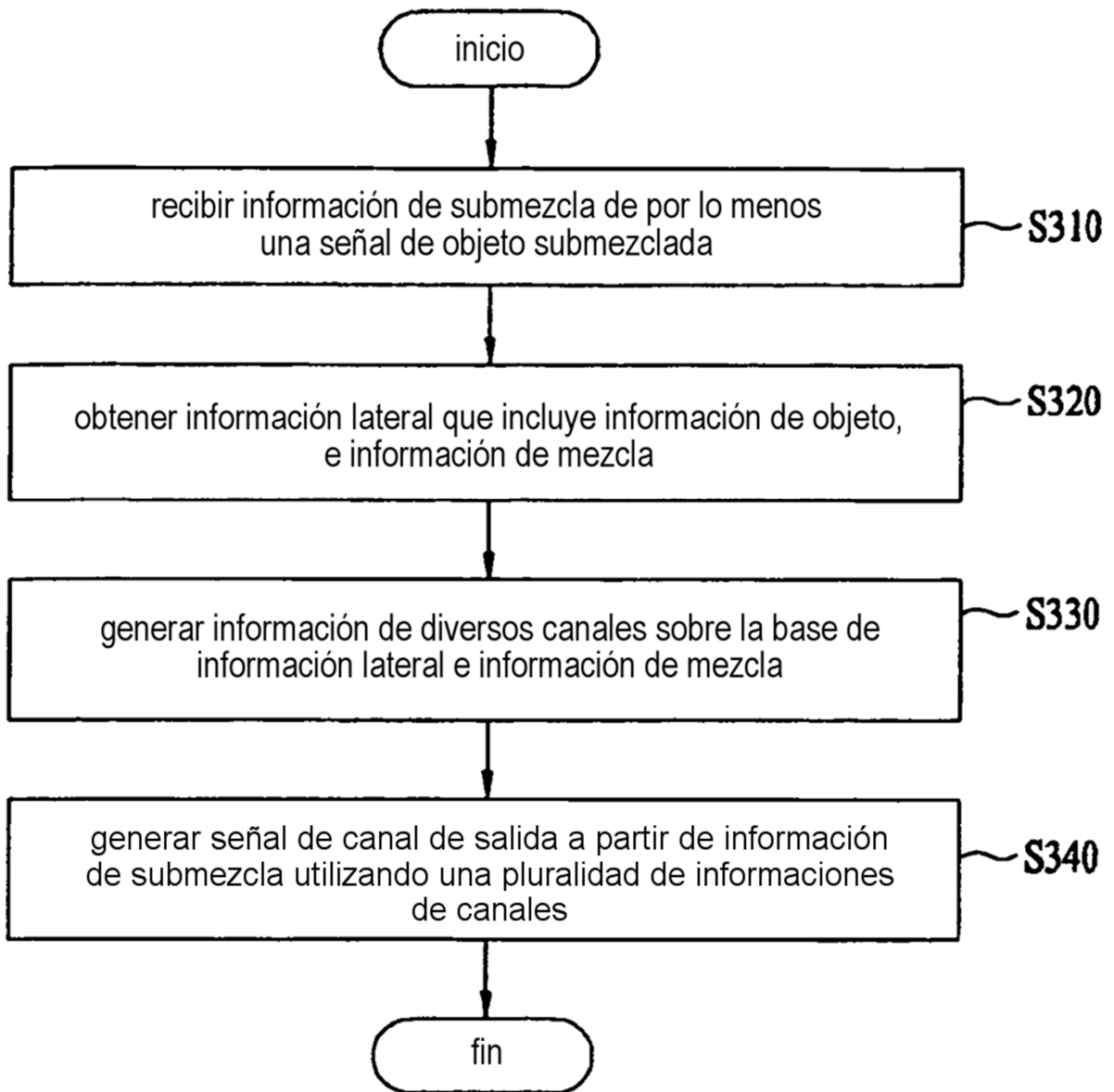


FIG. 4

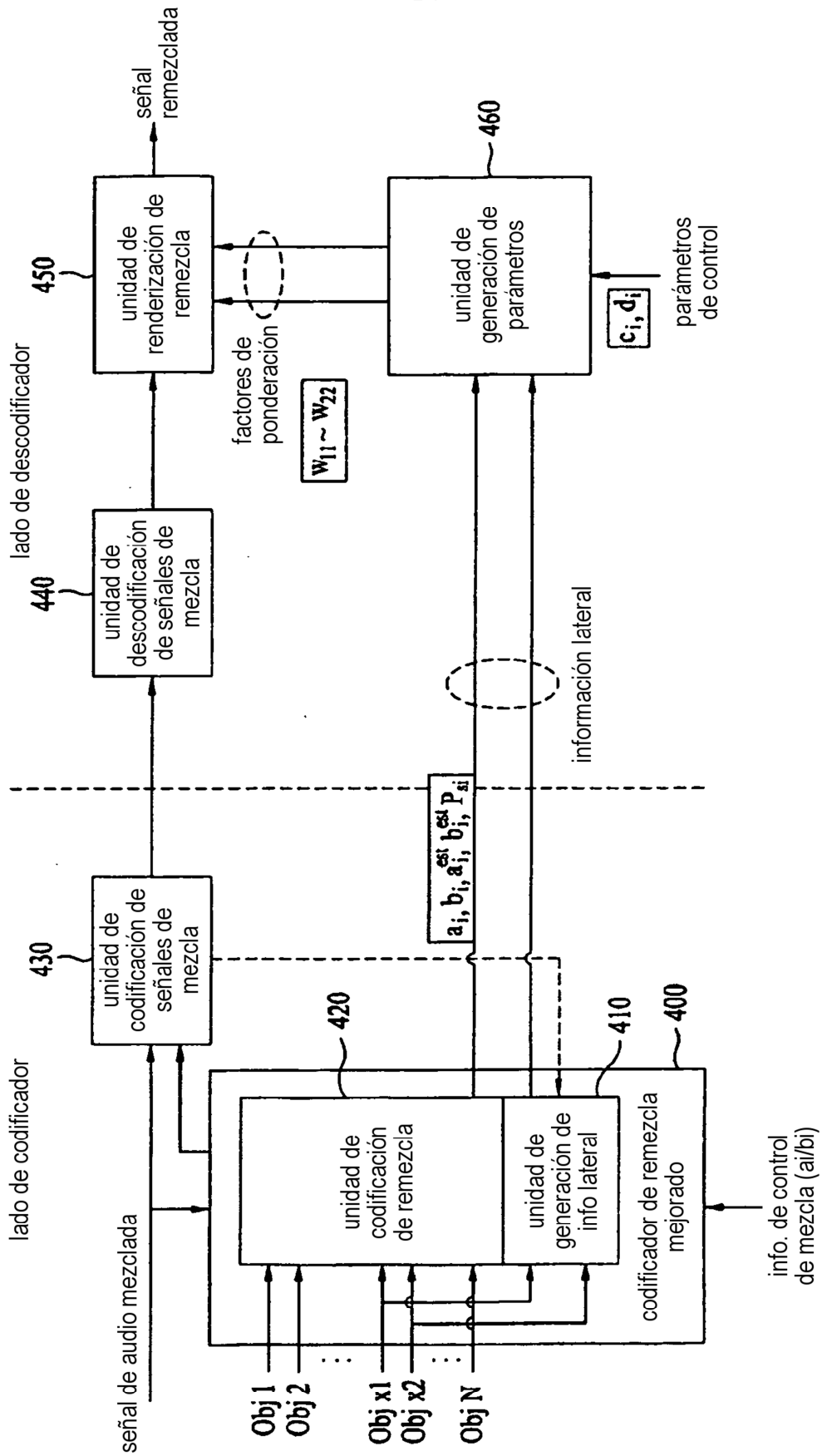


FIG.5

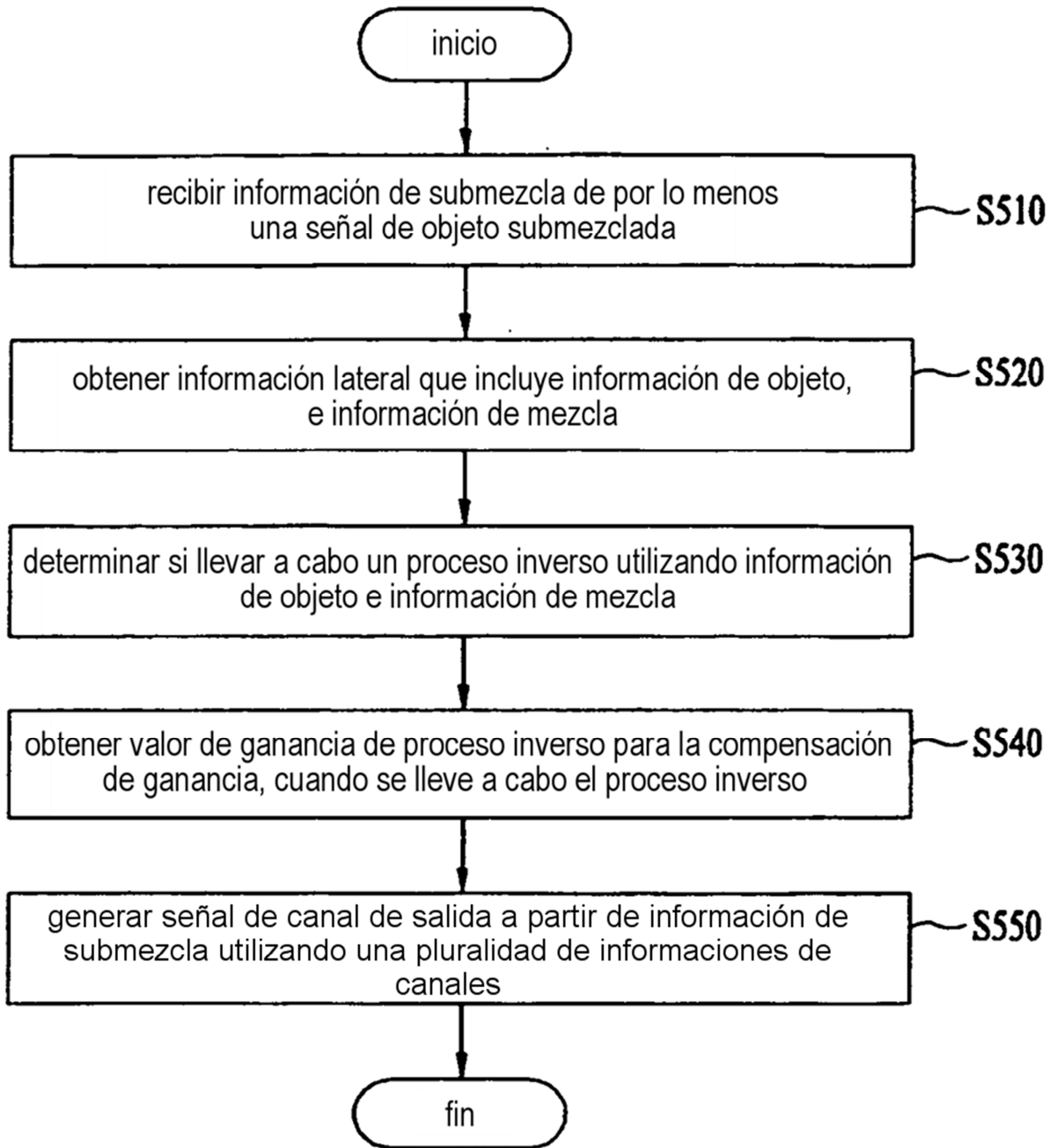


FIG. 6

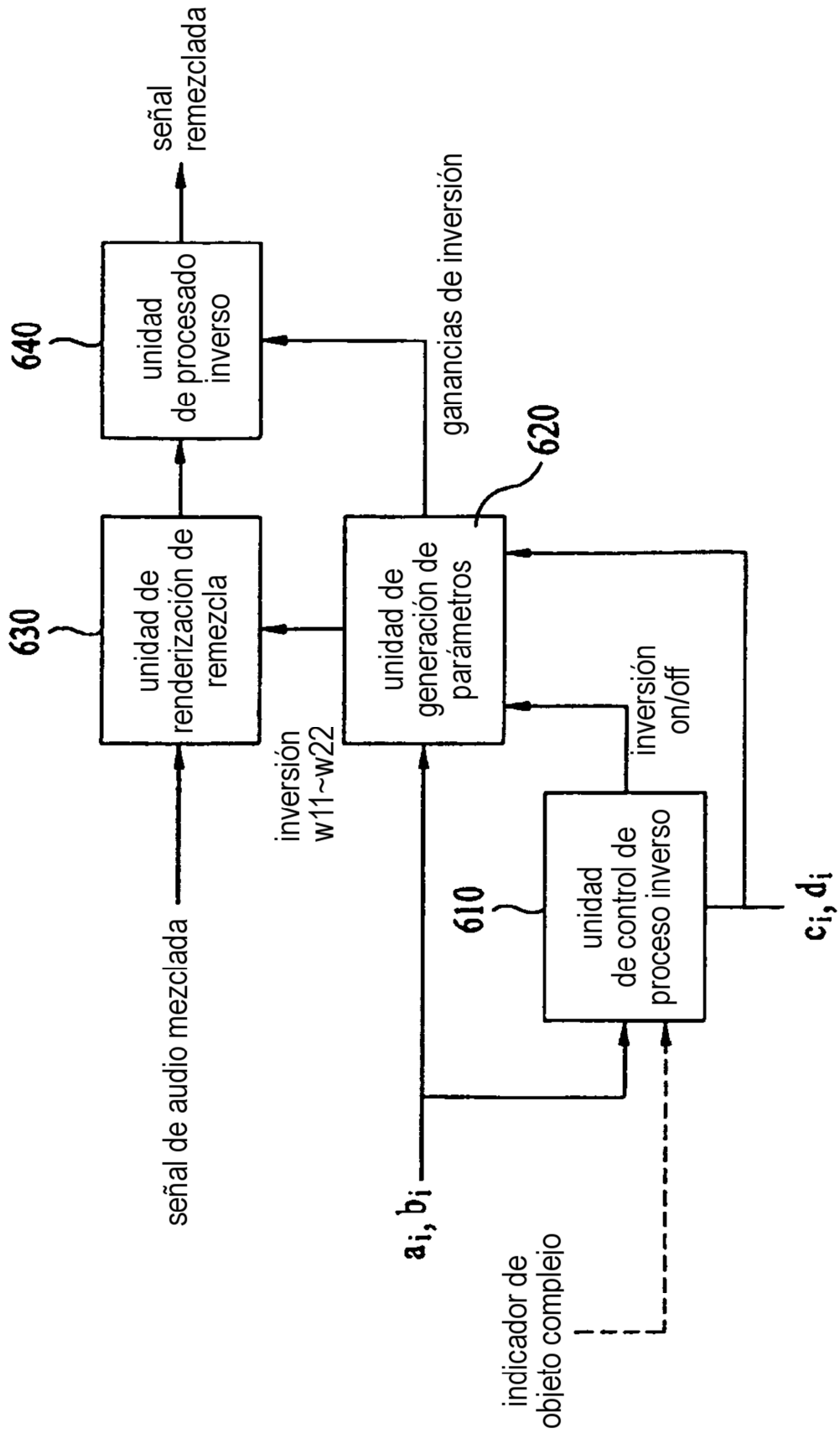




FIG. 7

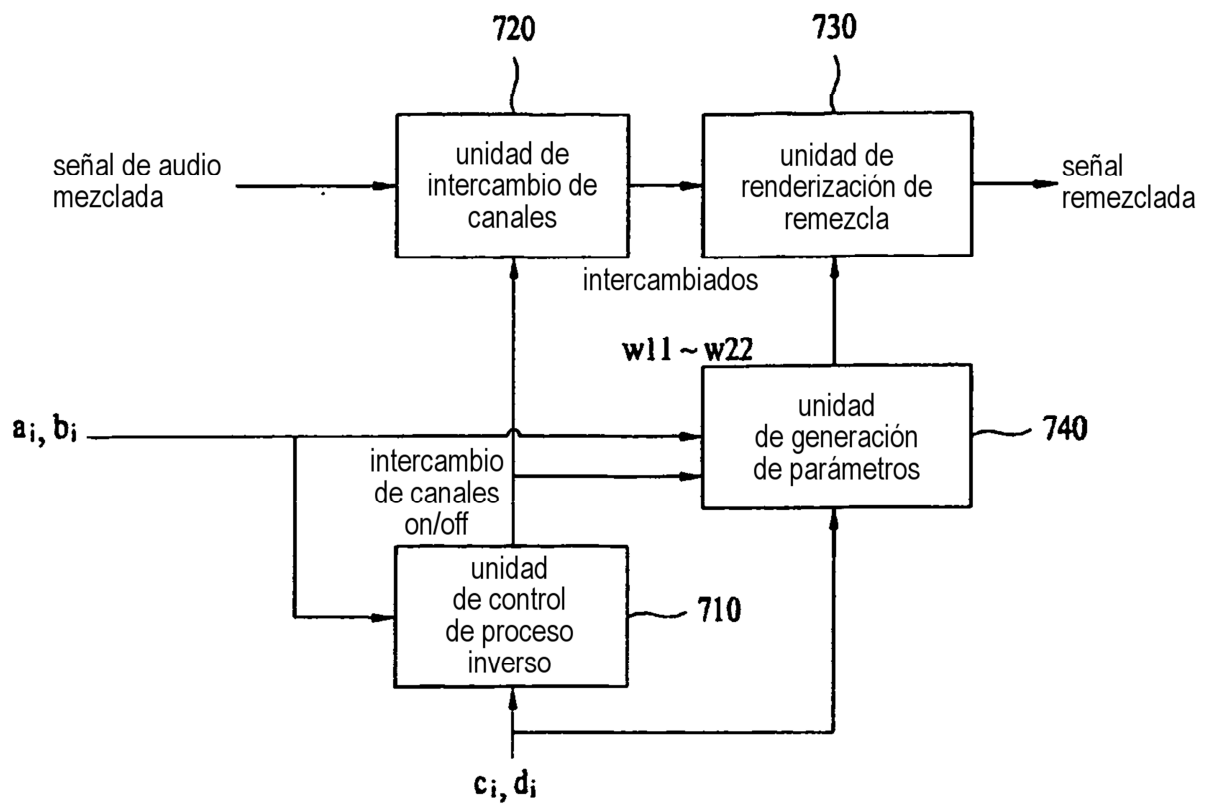


FIG. 8

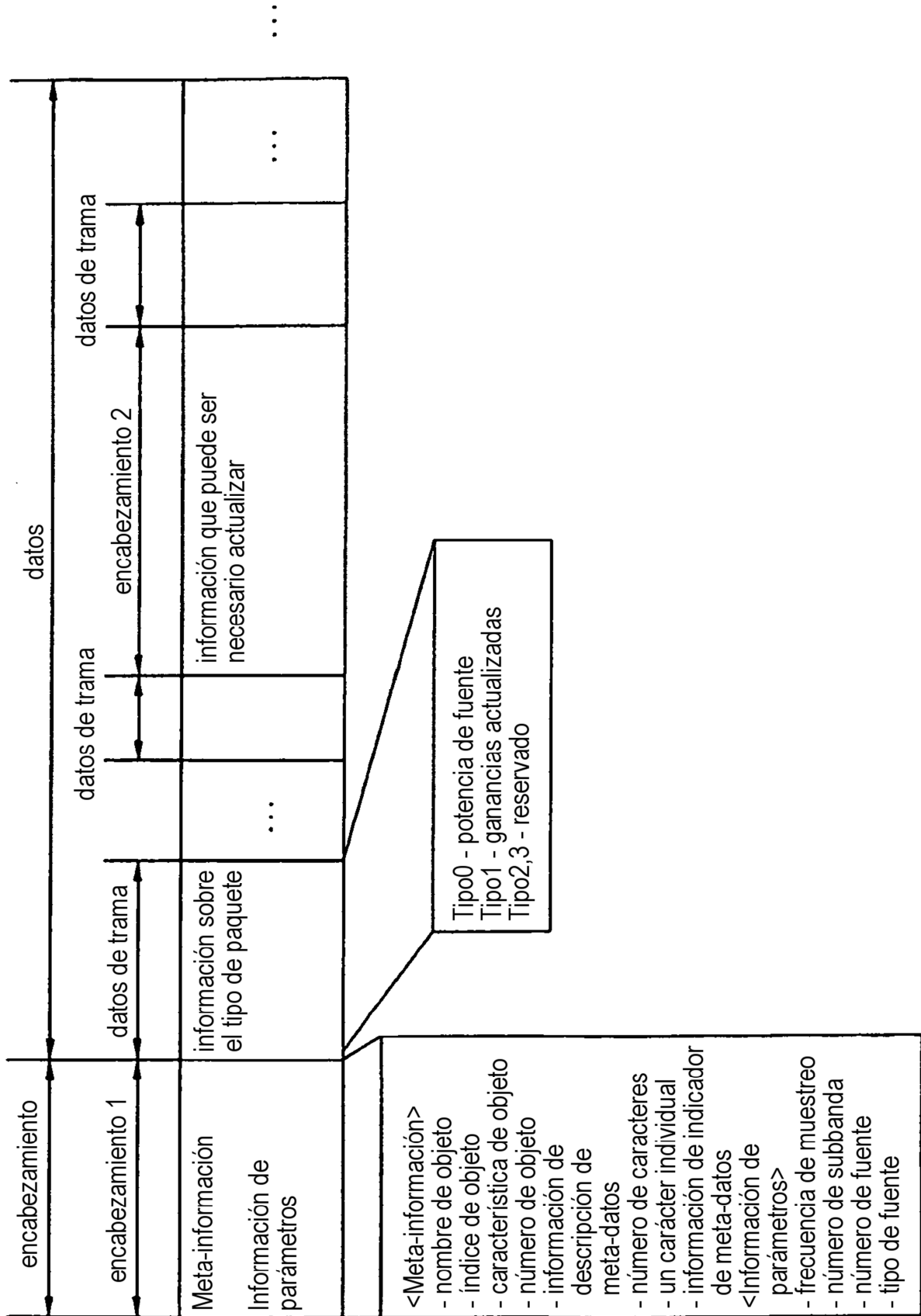


FIG. 9

```
val = Hdecoding(hcb_type_t);
if( sum (val) !=0 ) {
    uni_sign;
    if( uni_sign == TRUE ) {
        sign_for_all;
    }
    else {
        for( k=0; k < npart; k++ ) {
            if( val[k] ) {
                sign_per_val;
            }
        }
    }
}
```

FIG. 10

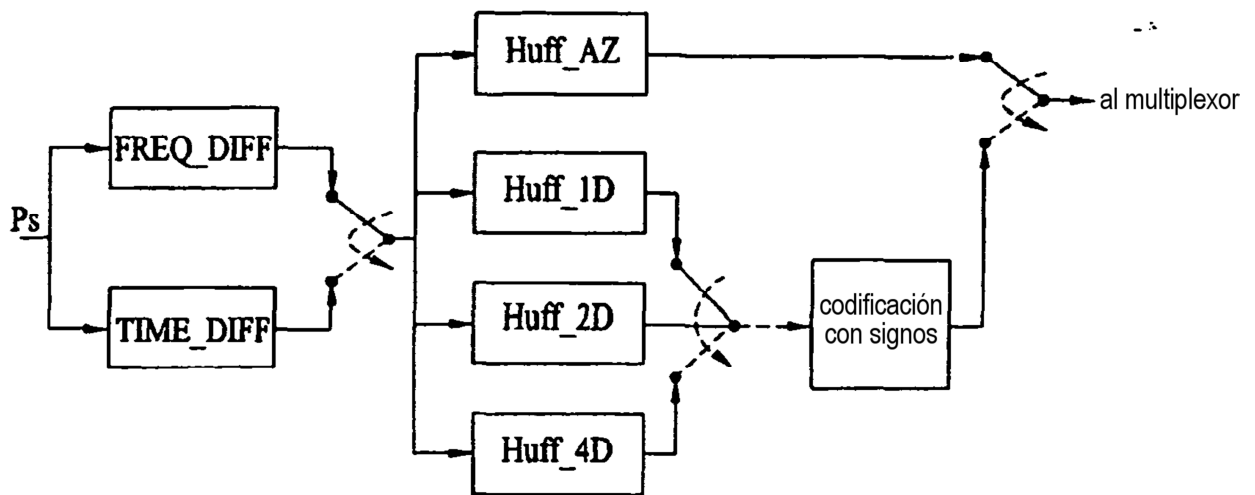


FIG. 11

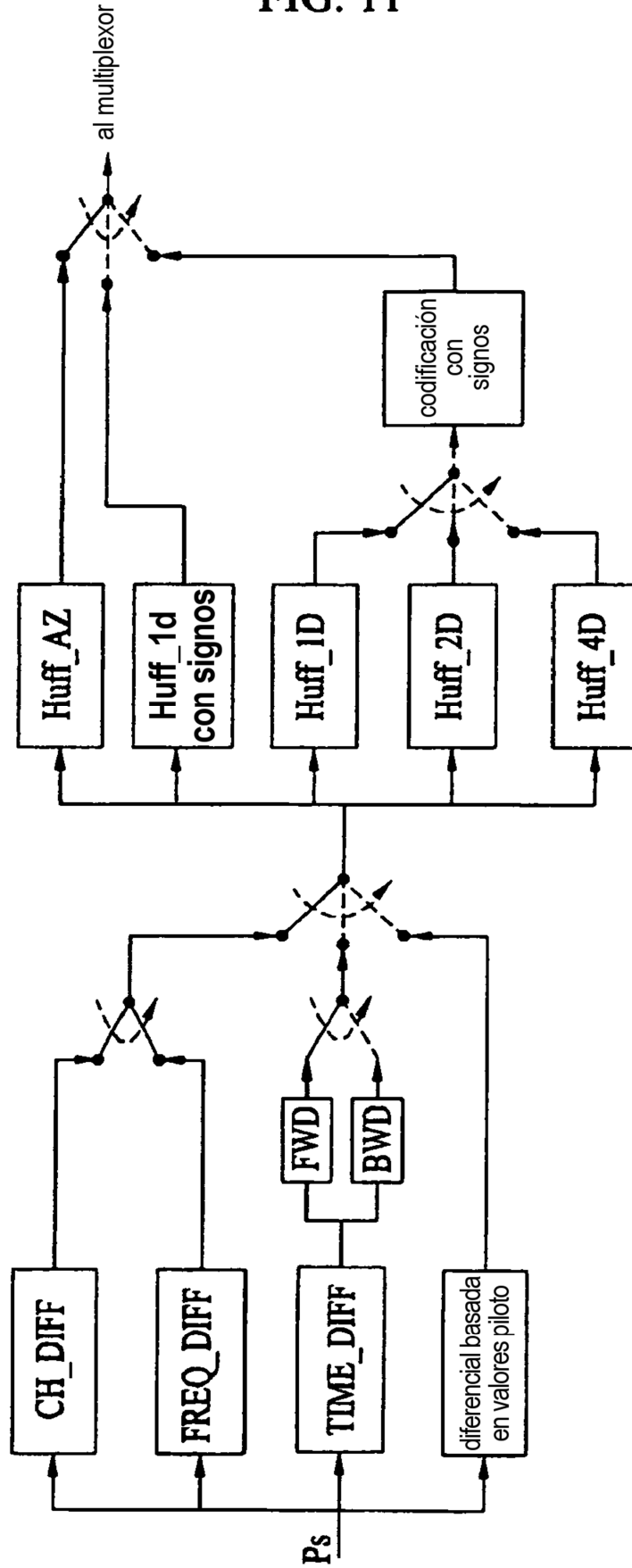
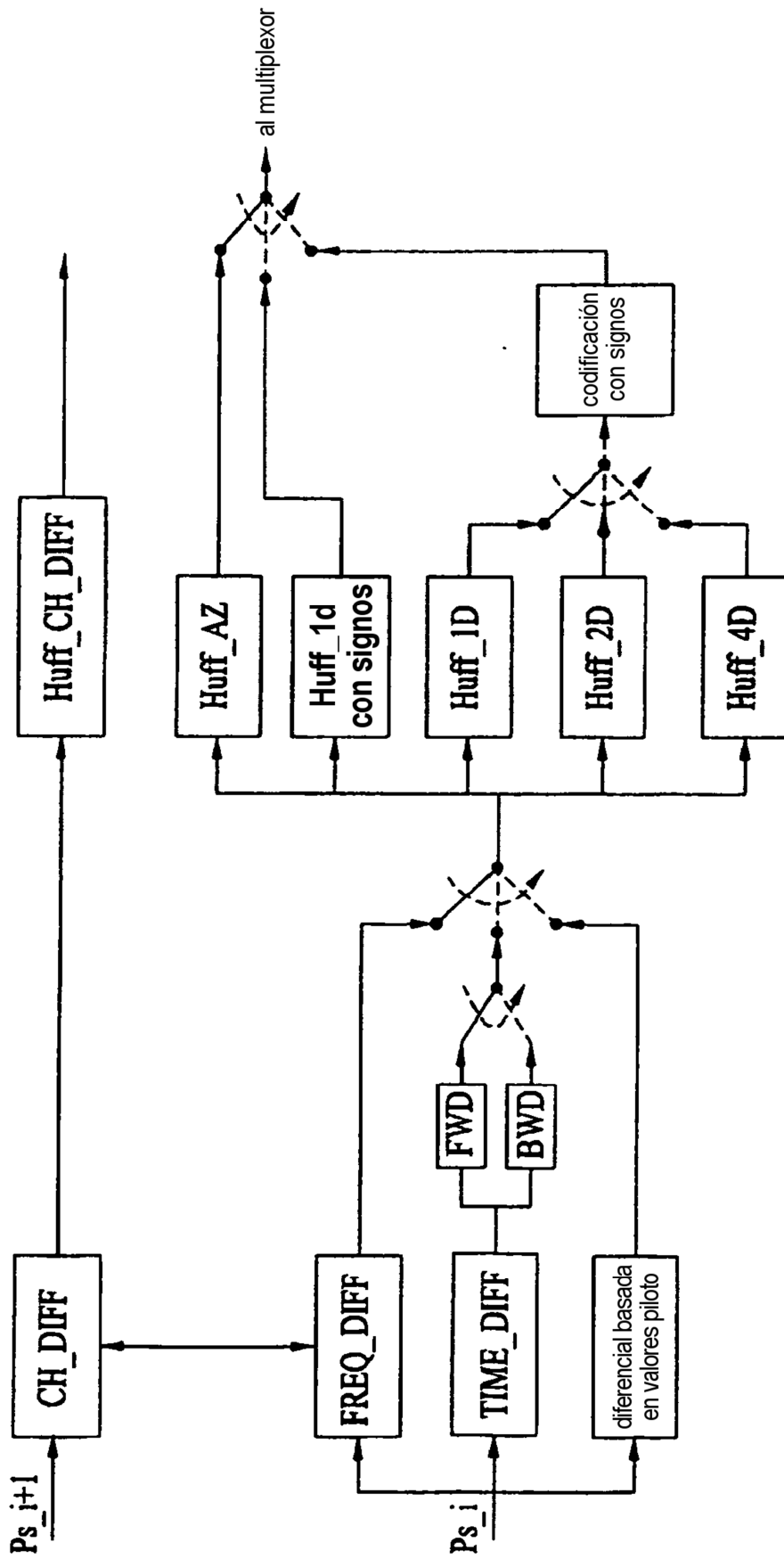


FIG. 12



**FIG. 13**

