

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 673 319**

51 Int. Cl.:

G10L 19/02 (2013.01)

G10L 19/26 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.02.2013** **PCT/EP2013/053831**

87 Fecha y número de publicación internacional: **06.09.2013** **WO13127801**

96 Fecha de presentación y número de la solicitud europea: **26.02.2013** **E 13705826 (9)**

97 Fecha y número de publicación de la concesión europea: **21.03.2018** **EP 2820647**

54 Título: **Control de coherencia de fase para señales armónicas en códecs de audio perceptual**

30 Prioridad:

27.02.2012 US 201261603773 P
27.07.2012 EP 12178265

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
21.06.2018

73 Titular/es:

**FRAUNHOFER GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.0%)**
Hansastraße 27 c
80686 München, DE

72 Inventor/es:

DISCH, SASCHA;
HERRE, JÜRGEN;
EDLER, BERND y
NAGEL, FREDERIK

74 Agente/Representante:

SALVA FERRER, Joan

ES 2 673 319 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

Control de coherencia de fase para señales armónicas en códecs de audio perceptual

- 5 **[0001]** La presente invención se refiere a un aparato y procedimiento para generar una señal de salida de audio y en particular a un aparato y procedimiento para implementar control de coherencia de fase para señales armónicas en códecs de audio perceptual.
- 10 **[0002]** El procesamiento de señal de audio se vuelve más y más importante. En particular, la codificación de audio perceptual ha proliferado como corriente principal habilitando la tecnología digital para todos los tipos de aplicaciones que proporcionan audio y multimedia a consumidores utilizando canales de transmisión o almacenamiento con capacidad limitada. Se requiere que los códecs de audio perceptual modernos proporcionen calidad de audio satisfactoria a proporciones de *bits* cada vez más bajas. A su vez, se tiene que tratar con ciertos artefactos de codificación que son más tolerables para la mayoría de los oyentes.
- 15 **[0003]** Uno de estos artefactos es la pérdida de coherencia de fase con respecto a la frecuencia (coherencia de fase "vertical"), véase [8]. Para muchas señales estacionarias, el deterioro resultante en la calidad de señal de audio subjetiva usualmente es más bien pequeño. Sin embargo, en los sonidos tonales armónicos que están constituidos por muchos componentes espectrales que son percibidos por el sistema auditivo humano como una sola combinación, la distorsión perceptual resultante es desagradable.
- 20 **[0004]** Las señales típicas, donde la conservación de coherencia de fase vertical (VPC) es importante, son habla de voz, instrumentos de latón o cuerdas arqueadas, por ejemplo "instrumentos" que, por la naturaleza de su producción de sonido física, producen sonido que es rico en su contenido de sobretono y fase-bloqueado entre los sobretonos armónicos. Especialmente a muy bajas velocidades de *bits*, donde el presupuesto de *bits* es extremadamente limitado, el uso de los códecs del estado de la técnica frecuentemente debilita de manera sustancial la VPC de los componentes espectrales. Sin embargo, en las señales mencionadas anteriormente, la VPC es una guía auditiva perceptual importante y una alta VPC de la señal debe ser conservada.
- 30 **[0005]** En lo sucesivo, se considera la codificación de audio perceptual de acuerdo con el estado de la técnica. En el estado de la técnica, la codificación de audio perceptual sigue varios temas comunes, incluyendo el procesamiento de uso de tiempo/frecuencia-dominio, reducción de redundancia (codificación de entropía) y remoción de irrelevancia a través del aprovechamiento pronunciado de efectos perceptuales (véase [1]). Comúnmente, la señal de entrada es analizada mediante un banco de filtros de análisis que convierten la señal de dominio de tiempo a una representación espectral, por ejemplo, una representación de tiempo/frecuencia. La conversión a coeficientes espectrales permite procesar selectivamente los componentes de señal dependiendo de su contenido de frecuencia, por ejemplo, diferentes instrumentos con sus estructuras de sobretono individuales.
- 35 **[0006]** En paralelo, la señal de entrada es analizada con respecto a sus propiedades perceptuales. Por ejemplo, se puede calcular un umbral de enmascaramiento tiempo- y frecuencia-dependiente. El umbral de enmascaramiento tiempo/frecuencia dependiente puede ser alimentado a una unidad de codificación por medio de un umbral de codificación objetivo en forma de un valor de energía absoluto o una proporción de máscara a señal (MSR) para cada banda de frecuencia y cuadro de tiempo de codificación.
- 45 **[0007]** Los coeficientes espectrales alimentados por el banco de filtro de análisis son cuantificados para reducir la velocidad de datos necesaria para representar la señal. Esta etapa implica una pérdida de información e introduce una distorsión de codificación (error, ruido) a la señal. Con el fin de minimizar el impacto audible de este ruido de codificación, los tamaños de etapa del codificador se controlan de acuerdo con los umbrales de codificación objetivos para cada banda de frecuencia y cuadro. Idealmente, el ruido de codificación inyectado a cada banda de frecuencia es más bajo que el umbral de codificación (enmascaramiento) y así ninguna degradación en audio subjetivo es perceptible (eliminación de irrelevancia). Este control del ruido de cuantificación con respecto a la frecuencia y tiempo de acuerdo con requisitos psicoacústicos conduce a un efecto de formación de ruido sofisticado y es lo que hace al codificador un codificador de audio perceptual.
- 50 **[0008]** Posteriormente, los codificadores de audio modernos efectúan codificación de entropía, por ejemplo, codificación de Huffman o codificación aritmética sobre los datos espectrales cuantificados. La codificación de entropía es una etapa de codificación sin pérdidas que ahorra además velocidad de *bits*.
- 55 **[0009]** Finalmente, todos los datos espectrales codificados y parámetros adicionales relevantes, por ejemplo, información lateral, por ejemplo, los ajustes del cuantificador para cada banda de frecuencia, son empaquetados

conjuntamente a una corriente de *bits*, que es la representación codificada final destinada para almacenamiento o transmisión de archivos.

[0010] Ahora, se considera la extensión de ancho de banda de acuerdo con el estado de la técnica. En la codificación de audio perceptual basada en bancos de filtros, la parte principal de la velocidad de *bits* consumida es usualmente gastada en los coeficientes espectrales cuantificados. Así, a velocidades de *bits* muy bajas, no suficientes *bits* pueden estar disponibles para representar todos los coeficientes con la precisión requerida para obtener reproducción perceptualmente sin deteriorar. Mediante esto, los requisitos de baja velocidad de *bits* establecen efectivamente un límite al ancho de banda de audio que puede ser obtenido mediante codificación de audio perceptual.

[0011] La extensión de ancho de banda (véase [2]) elimina esta limitación fundamental de hace mucho tiempo. La idea central de la extensión de ancho de banda es complementar un códec perceptual banda-limitado mediante un procesador de alta frecuencia adicional que transmite y restaura el contenido de alta frecuencia faltante en una forma paramétrica compacta. El contenido de alta frecuencia puede ser generado con base en una sola modulación de banda lateral de la señal de banda base, véase, por ejemplo [3] o en la aplicación de técnicas de desplazamiento de tono, por ejemplo el "vocoder" en [4].

[0012] Especialmente para bajas velocidades de *bits*, se han diseñado esquemas de codificación paramétricos que codifican los componentes sinusoidales (sinusoides) mediante una representación paramétrica compacta (véase, por ejemplo [9], [10], [11] y [12]). Dependiendo del codificador individual, el residuo restante es sometido además a codificación paramétrica o es codificado en forma de onda.

[0013] En lo sucesivo, se considera la codificación de audio espacial paramétrica de acuerdo con el estado de la técnica. Como la extensión de ancho de banda de las señales de audio, la codificación de audio espacial (SAC) sale del dominio de codificación en forma de onda y en lugar de esto se enfoca en la entrega de una réplica perceptualmente satisfactoria de la imagen de sonido espacial original. Una escena de sonido percibida por un oyente humano es determinada esencialmente por las diferencias entre las señales del oído del oyente (las llamadas diferencias inter-aurales) sin consideración de si la escena consiste de fuentes de audio reales o si es reproducida vía dos o más altavoces que proyectan sonido fantasma. En lugar de codificar discretamente las señales de canal de entrada de audio individuales, un sistema basado en SAC captura la imagen espacial de una señal de audio de multicanal a un conjunto compacto de parámetros que pueden ser usados para sintetizar una representación de multicanal de alta calidad a partir de una señal de mezcla descendente transmitida (véase, por ejemplo, [5], [6] y [7]).

[0014] Debido a su naturaleza paramétrica, la codificación de audio espacial no conserva la forma de onda. Como consecuencia, es difícil obtener una calidad totalmente sin deteriorar para todos los tipos de señales de audio. No obstante, la codificación de audio espacial es una estrategia extremadamente poderosa que proporciona ganancia sustancial a velocidades de *bits* bajas e intermedias.

[0015] Los efectos de audio digital tales como efectos de estiramiento de tiempo o efectos de desplazamiento de tono son usualmente obtenidos mediante la aplicación de técnicas de dominio de tiempo como solapamiento-adición sincronizados (SOLA) o mediante la aplicación de técnicas de dominio de frecuencia, por ejemplo, al emplear un "vocoder". Además, sistemas híbridos han sido propuestos en el estado de la técnica que aplican un procesamiento de SOLA en la subbandas. Los "vocoders" y sistemas híbridos usualmente sufren de un artefacto llamado "phasiness" que puede ser atribuido a la pérdida de coherencia de fase vertical. Algunas publicaciones se refieren a mejoras en la calidad de sonido de algoritmos de estiramiento de tiempo al conservar la coherencia de fase vertical donde es importante (véase, por ejemplo [14] y [15]).

[0016] El uso de los códecs de audio perceptual del estado de la técnica frecuentemente debilita la coherencia de fase vertical (VPC) de los componentes espectrales de una señal de audio, especialmente a bajas velocidades de *bits*, donde se aplican técnicas de codificación paramétricas. Sin embargo, en ciertas señales, VPC es una guía perceptual importante. Como resultado, la calidad perceptual de tales sonidos está deteriorada.

[0017] Los codificadores de audio del estado de la técnica usualmente comprometen la calidad perceptual de las señales de audio al menospreciar propiedades de fase importantes de la señal que van a ser codificadas (véase, por ejemplo [1]). La codificación burda de los coeficientes espectrales transmitidos en un codificador de audio ya puede alterar la VPC de la señal decodificada. Además, especialmente debido a la aplicación de técnicas de codificación paramétricas, tales como extensión de ancho de banda (véase [2], [3] y [4]), la codificación de multicanal paramétrica (véase, por ejemplo [5], [6] y [7]) o codificación paramétrica de componentes sinusoidales (véase [9], [10], [11] y [12]), la coherencia de fase con respecto a la frecuencia está frecuentemente deteriorada.

[0018] El resultado es un sonido sordo que parece proceder de una distancia lejana y así evoca poca involucración del oyente [13]. Existen muchos tipos de componentes de señal donde la coherencia de fase vertical es importante. Las señales típicas donde VPC es importante son, por ejemplo, tonos con un contenido de sobre tono armónico rico, tal como habla de voz, instrumentos de latón o cuerdas arqueadas.

5

[0019] El objeto de la presente invención es proporcionar conceptos mejorados para el procesamiento de señal de audio y en particular proporcionar conceptos mejorados para el control de coherencia de fase para señales armónicas en códecs de audio perceptual. El objeto de la presente invención es resuelto por un decodificador de acuerdo con la reivindicación 1, por un codificador de acuerdo con la reivindicación 8, por un sistema de acuerdo con la reivindicación 14, por un procedimiento de acuerdo con la reivindicación 15, por un procedimiento para decodificación de acuerdo con la reivindicación 16, por un programa informático de acuerdo con la reivindicación 17. Se proporciona un decodificador para decodificar una señal de audio codificada para obtener una señal de audio fase-ajustada. El decodificador comprende una unidad de decodificación y una unidad de ajuste de fase. La unidad de decodificación está adaptada para decodificar la señal de audio codificada para obtener una señal de audio decodificada. La unidad de ajuste de fase está adaptada para ajustar la señal de audio decodificada para obtener la señal de audio fase-ajustada. La unidad de ajuste de fase está configurada para recibir información de control dependiendo de la coherencia de fase vertical de la señal de audio codificada. Además, la unidad de ajuste de fase está adaptada para ajustar la señal de audio decodificada con base en la información de control.

[0020] En una realización, la unidad de ajuste de fase puede estar configurada para ajustar la señal de audio decodificada cuando la información de control indica que el ajuste de fase está activado.

[0021] La unidad de ajuste de fase puede estar configurada para no ajustar la señal de audio decodificada cuando la información de control indica que el ajuste de fase está desactivado.

25

[0022] En otra realización, la unidad de ajuste de fase puede estar configurada para recibir la información de control, la información de control comprende un valor de intensidad que indica una intensidad de un ajuste de fase. Además, la unidad de ajuste de fase puede estar configurada para ajustar la señal de audio decodificada en base al valor de intensidad.

30

[0023] De acuerdo con una realización adicional, el decodificador puede comprender además un banco de filtros de análisis para descomponer la señal de audio decodificada a una pluralidad de señales de subbanda de una pluralidad de subbandas. La unidad de ajuste de fase puede estar configurada para determinar una pluralidad de primeros valores de fase de la pluralidad de señales de subbanda. Además, la unidad de ajuste de fase puede estar adaptada para ajustar la señal de audio codificada al modificar al menos algunos de la pluralidad de los primeros valores de fase para obtener segundos valores de fase de la señal de audio fase-ajustada.

35

[0024] En otra realización, la unidad de ajuste de fase puede ser configurada para ajustar al menos algunos de los valores de fase al aplicar la fórmula:

40

$$\begin{aligned} p_x'(f) &= p_x(f) - dp(f) \text{ y} \\ dp(f) &= \alpha * (p_0(f) + \text{const}), \end{aligned}$$

donde f es una frecuencia que indica una de las subbandas que tiene la frecuencia f como una frecuencia central, donde $p_x(f)$ es uno de los primeros valores de fase de una de las señales de subbanda de una de las subbandas que tiene la frecuencia f como la frecuencia central, donde $p_x'(f)$ es uno de los segundos valores de fase de una de las señales de subbanda de una de las subbandas que tiene la frecuencia f como la frecuencia central, donde const es un primer ángulo en el intervalo $-\pi \leq \text{const} \leq \pi$, donde α es un número real en el intervalo $0 \leq \alpha \leq 1$ y donde $p_0(f)$ es un segundo ángulo en el intervalo $-\pi \leq p_0(f) \leq \pi$, donde el segundo ángulo $p_0(f)$ es asignado a una de las subbandas que tiene la frecuencia f como la frecuencia central. Alternativamente, el ajuste de fase anterior también se puede llevar a cabo mediante multiplicación de una señal de subbanda compleja (por ejemplo, los coeficientes espectrales complejos de una transformada de Fourier discreta) por un término de fase exponencial $e^{jdp(f)}$, donde j es el número imaginario unitario.

[0025] De acuerdo con otra realización, el decodificador puede comprender además un banco de filtros de síntesis. La señal de audio fase-ajustada puede ser una señal de audio de dominio espectral fase-ajustada que es representada en un dominio espectral. El banco de filtro de síntesis puede estar configurado para transformar la señal de audio de dominio espectral fase-ajustada del dominio espectral a un dominio de tiempo para obtener una señal de audio de dominio de tiempo fase-ajustada.

55

[0026] En una realización, el decodificador puede estar configurado para decodificar información de control de VPC.

[0027] Además, de acuerdo con otra realización, el decodificador puede estar configurado para aplicar información de control para obtener una señal decodificada con una VPC mejor conservada que en los sistemas convencionales.

[0028] Además, el decodificador puede estar configurado para manipular la VPC direccionada mediante mediciones en el decodificador y/o información de activación contenida en la corriente de *bits*.

[0029] Además, se proporciona un codificador para codificar información de control basada en una señal de audio de entrada. El codificador comprende una unidad de transformación, un generador de información de control y una unidad de codificación. La unidad de transformación está adaptada para transformar la señal de entrada de audio de un dominio de tiempo a un dominio espectral para obtener una señal de audio transformada que comprende una pluralidad de señales de subbanda que son asignadas a una pluralidad de subbandas. El generador de información de control es apto para generar la información de control de tal manera que la información de control indica una coherencia de fase vertical de la señal de audio transformada. La unidad de codificación está adaptada para codificar la señal de audio transformada y la información de control.

[0030] En una realización, la unidad de transformación del codificador comprende un banco de filtros de coclear para transformar la señal de entrada de audio del dominio de tiempo al dominio espectral para obtener la señal de audio transformada que comprende la pluralidad de señales de subbanda.

[0031] De acuerdo con una realización adicional, el generador de información de control puede estar configurado para determinar una envolvente de subbanda para cada una de la pluralidad de señales de subbanda para obtener una pluralidad de envolventes de señal de subbanda. Además, el generador de información de control puede estar configurado para generar una envolvente combinada en base a la pluralidad de envolventes de señal de subbanda. Además, el generador de información de control puede estar configurado para generar la información de control en base a la envolvente combinada.

[0032] En otra realización, el generador de información de control puede estar configurado para generar un número caracterizante en base a la envolvente combinada. Además, el generador de información de control puede estar configurado para generar la información de control, de tal manera que la información de control indica que el ajuste de fase está activado cuando el número caracterizante es mayor que un valor de umbral. Además, el generador de información de control puede estar configurado para generar la información de control de tal manera que la información de control indica que el ajuste de fase está desactivado cuando el número caracterizante es menor o igual que el valor de umbral.

[0033] De acuerdo con una realización adicional, el generador de información de control puede estar configurado para generar la información de control al calcular una proporción de la media geométrica de la envolvente de combinada con una media aritmética de la envolvente combinada.

[0034] Alternativamente, el valor máximo de la envolvente combinada puede ser comparado con un valor medio de la envolvente combinada. Por ejemplo, se puede formar una proporción de máxima/media, por ejemplo una proporción del valor máximo de la envolvente combinada con el valor medio de la envolvente combinada.

[0035] En una realización, el generador de información de control puede estar configurado para generar la información de control de tal manera que la información de control comprende un valor de intensidad que indica el grado de coherencia de fase vertical de las señales de subbanda.

[0036] Un codificador de acuerdo con una realización puede estar configurado para llevar a cabo una medición de la VPC en el lado del codificador, por medio, por ejemplo, de mediciones fase y/o derivados de fase con respecto a la frecuencia.

[0037] Además, el codificador de acuerdo con una realización puede estar configurado para llevar a cabo una medición de la saliencia perceptual de la coherencia de fase vertical.

[0038] Además, el codificador de acuerdo con una realización puede estar configurado para llevar a cabo una derivación de información de activación de la saliencia de coherencia de fase y/o mediciones de VPC.

[0039] Además, el codificador de acuerdo con una realización puede estar configurado para extraer pistas de VPC adaptables tiempo-frecuencia o información de control.

[0040] Además, el codificador de acuerdo con una realización puede estar configurado para determinar una representación compacta de información de control de VPC.

[0041] En realizaciones, la información de control de VPC puede ser transmitida en una corriente de *bits*.

[0042] Además, se proporciona un sistema. El sistema comprende un codificador de acuerdo con una de las realizaciones anteriormente descritas y al menos un decodificador según una de las realizaciones anteriormente descritas. El codificador está configurado para transformar una señal de entrada de audio para obtener una señal de audio transformada. Además, el codificador está configurado para codificar la señal de audio transformada para obtener una señal de audio codificada. Además, el codificador está configurado para codificar la información de control que indica una coherencia de fase vertical de la señal de audio transformada. Además, el codificador está dispuesto para alimentar la señal de audio codificada y la información de control en al menos un decodificador. Al menos un decodificador está configurado para decodificar la señal de audio codificada para obtener una señal de audio decodificada. Además, al menos un decodificador está configurado para ajustar la señal de audio decodificada basada en la información de control codificada para obtener una señal de audio de fase-ajustada.

[0043] En realizaciones, la VPC puede ser medida en el lado del codificador, transmitida como información lateral compacta apropiada junto con la señal de audio codificada y la VPC de la señal es restaurada en el decodificador. De acuerdo con realizaciones alternativas, la VPC es manipulada en el decodificador direccionado por la información de control generada en el decodificador y/o guiada por información de activación transmitida del codificador en la información lateral. El procesamiento de VPC puede ser tiempo-frecuencia selectivo de tal manera que la VPC es solamente restaurada donde es perceptualmente beneficioso.

[0044] Además, se proporciona un procedimiento para decodificar una señal de audio codificada para obtener una señal de audio fase-ajustada. El procedimiento para decodificación comprende:

30 - Recibir información de control, donde la información de control indica una coherencia de fase vertical de la señal de audio codificada.

- Decodificar la señal de audio codificada para obtener una señal de audio decodificada.

35 - Ajustar la señal de audio decodificada para obtener la señal de audio de fase-ajustada en base a la información de control.

[0045] Además, se proporciona un procedimiento para codificar información de control con base en una señal de entrada de audio. El procedimiento para codificación comprende:

40 - Transformar la señal de entrada de audio de un dominio de tiempo a un dominio espectral para obtener una señal de audio transformada que comprende una pluralidad de señales de subbanda que son asignadas a una pluralidad de subbandas.

45 - Generar la información de control de tal manera que la información de control indica una coherencia de fase vertical de la señal de audio transformada.

- Codificar la señal de audio transformada y la información de control.

50 **[0046]** Además, se proporciona un programa informático para implementar uno de los procedimientos descritos anteriormente cuando el programa informático es ejecutado en un ordenador o un procesador de señal.

[0047] En realizaciones, se proporcionan medios para conservar la coherencia de fase vertical (VPC) de las señales cuando la VPC ha sido comprometida por un procesamiento de señal, codificación o proceso de transmisión.

55 **[0048]** En algunas realizaciones, el sistema de la invención mide la VPC de la señal de entrada antes de su codificación, transmite información lateral compacta apropiada junto con la señal de audio codificada y restaura la VPC de la señal en el decodificador en base a la información lateral compacta transmitida. Alternativamente, el procedimiento de la invención manipula la VPC en el decodificador direccionada por información de control generada en el decodificador y/o guiada por información de activación transmitida del codificador en la información lateral.

[0049] En otras realizaciones, la VPC de una señal deteriorada puede ser procesada para restaurar su VPC original al usar un proceso de ajuste de VPC que es controlado al analizar la señal deteriorada en sí misma.

- 5 **[0050]** En ambos casos, dicho procesamiento puede ser tiempo-frecuencia selectivo de tal manera que la VPC es solamente restaurada donde es perceptualmente beneficioso.

[0051] La calidad de sonido mejorada de los codificadores de audio perceptual se proporciona a costes de información lateral moderados. Además de los codificadores de audio perceptual, la medición y restauración de la
10 VPC es también beneficiosa para efectos de audio digital basados en "vocoders" de fase, como estiramiento de tiempo o desplazamiento de tono.

[0052] Las realizaciones se proporcionan en las reivindicaciones dependientes.

- 15 **[0053]** En lo sucesivo, se describen realizaciones con respecto a las figuras en las cuales:

La figura 1a ilustra un decodificador para decodificar una señal de audio codificada para obtener una señal de audio fase-ajustada de acuerdo con una realización.

- 20 La figura 1b ilustra un decodificador para decodificar una señal de audio decodificada para obtener una señal de audio fase-ajustada de acuerdo con otra realización.

La figura 2 ilustra un codificador para codificar información de control basada en una señal de entrada de audio de acuerdo con una realización.

- 25 La figura 3 ilustra un sistema de acuerdo con una realización que comprende un codificador y al menos un decodificador.

La figura 4 ilustra un sistema de procesamiento de audio con procesamiento de VPC de acuerdo con una realización.

- 30 La figura 5 ilustra un codificador de audio perceptual y un decodificador de acuerdo con una realización.

La figura 6 ilustra un generador de control de VPC de acuerdo con una realización.

- 35 La figura 7 ilustra un aparato para procesar una señal de audio para obtener una segunda señal de audio de acuerdo con un ejemplo.

La figura 8 ilustra un sistema de procesamiento de audio de procesamiento de VPC de acuerdo con otro ejemplo.

- 40 **[0054]** La figura 1a ilustra un decodificador para decodificar una señal de audio codificada para obtener una señal de audio fase-ajustada de acuerdo con una realización. El decodificador comprende una unidad de decodificación 110 y una unidad de ajuste de fase 120. La unidad de decodificación 110 está adaptada para decodificar la señal de audio codificada para obtener una señal de audio decodificada. La unidad de ajuste de fase 120 está adaptada para ajustar la señal de audio decodificada para obtener la señal de audio fase-ajustada. Además, la unidad
45 de ajuste de fase 120 está configurada para recibir información de control dependiendo de la coherencia de fase vertical (VPC) de la señal de audio codificada. Además, la unidad de ajuste de fase 120 está adaptada para ajustar la señal de audio decodificada en base a la información de control.

- [0055]** La realización de la figura 1a toma en cuenta que, para ciertas señales de audio, es importante restaurar
50 la coherencia de fase vertical de la señal codificada. Por ejemplo, cuando la porción de señal de audio comprende habla de voz, instrumentos de latón o cuerdas arqueadas, la conservación de coherencia de fase vertical es importante. Para ese propósito, la unidad de ajuste de fase 120 está adaptada para recibir información de control que depende de la VPC de la señal de audio codificada.

- 55 **[0056]** Por ejemplo, cuando las porciones de señal codificadas comprenden habla de voz, instrumentos de latón o cuerdas arqueadas, entonces la VPC de la señal codificada es alta. En tales casos, la información de control puede indicar que el ajuste de fase está activado.

[0057] Otras porciones de señal pueden no comprender señales tonales impulso-semejantes o transitorias y la VPC de tales porciones de señal puede ser baja. En tales casos, la información de control puede indicar que el ajuste de fase esta desactivado.

5 **[0058]** En otras realizaciones, la información de control puede comprender un valor de intensidad. Tal valor de intensidad puede indicar la intensidad del ajuste de fase que debe ser efectuado. Por ejemplo, el valor de intensidad puede ser un valor α con $0 \leq \alpha \leq 1$. Si $\alpha = 1$ o cercano a 1, esto puede indicar un alto valor de intensidad. Ajustes de fase significativos serán llevados a cabo por la unidad de ajuste de fase 120. Si α es cercano a 0, solamente ajustes de fase menores serán llevados a cabo por la unidad de ajuste de fase 120. Si $\alpha = 0$, no se llevará a cabo ningún
10 ajuste de fase.

[0059] La figura 1b ilustra un decodificador para decodificar una señal de audio codificada para obtener una señal de audio fase-ajustada de acuerdo con otra realización. Además de la unidad de decodificación 110 y la unidad de ajuste de fase 120, el decodificador de la figura 1b comprende un banco de filtros de análisis 115 y un banco de
15 filtro de síntesis 125.

[0060] El banco de filtro de análisis 115 está configurado para descomponer la señal de audio decodificada en una pluralidad de señales de subbanda de una pluralidad de señales de sub- banda. La unidad de ajuste de fase 120 de la figura 1b puede estar configurada para determinar una pluralidad de primeros valores de fase de la pluralidad de
20 señales de subbanda. Además, la unidad de ajuste de fase 120 puede estar adaptada para ajustar la señal de audio codificada al modificar al menos algunos de la pluralidad de los primeros valores de fase para obtener segundos valores de fase de la señal de audio fase-ajustada.

[0061] La señal de audio fase-ajustada puede ser una señal de audio de dominio espectral fase-ajustada que
25 es representada en un dominio espectral. El banco de filtros de síntesis 125 de la figura 1b puede estar configurado para transformar la señal de audio de dominio espectral fase-ajustada del dominio espectral a un dominio de tiempo para obtener una señal de audio de dominio de tiempo fase-ajustada.

[0062] La figura 2 ilustra un codificador correspondiente para codificar información de control en base a una
30 señal de entrada de audio de acuerdo con una realización. El codificador comprende una unidad de transformación 210, un generador de información de control 220 y una unidad de codificación 230. La unidad de transformación 210 está adaptada para transformar la señal de entrada de audio de un dominio de tiempo a un dominio espectral para obtener una señal de audio transformada que comprende una pluralidad de señales de subbanda que son asignadas a una pluralidad de subbandas. El generador de información de control 220 es apto para generar la información de
35 control de tal manera que la información de control indica una coherencia de fase vertical (VPC) de la señal de audio transformada. La unidad de codificación 230 está adaptada para codificar la señal de audio transformada y la información de control.

[0063] El codificador de la figura 2 está adaptado para codificar información de control que depende de la
40 coherencia de fase vertical de la señal de audio que se va a codificar. Para generar la información de control, la unidad de transformación 210 del codificador transforma la señal de entrada de audio en un dominio espectral, de tal manera que la señal de audio transformada resultante comprende una pluralidad de señales de subbanda de una pluralidad de subbandas.

45 **[0064]** Después de esto, el generador de información de control 220 determina a continuación información que depende de la coherencia de fase vertical de la señal de audio transformada.

[0065] Por ejemplo, el generador de información de control 220 puede clasificar una porción de señal de audio particular como una porción de señal donde la VPC es alta y por ejemplo, establecida a un valor de $\alpha=1$. Para otras
50 porciones de señal, el generador de información de control 220 puede clasificar una porción de señal de audio particular como una porción de señal donde la VPC es baja y por ejemplo, ajustada a un valor $\alpha=0$.

[0066] En otras realizaciones, el generador de información de control 220 puede determinar un valor de intensidad que depende de la VPC de la señal de audio transformada. Por ejemplo, el generador de información de control puede asignar un valor de intensidad con respecto a una porción de señal examinada, donde el valor de
55 intensidad depende de la VPC de la porción de señal. En el lado de un decodificador, el valor de intensidad puede ser empleado a continuación para determinar si solamente pequeños ajustes de fase serán llevados a cabo o si fuertes ajustes de fase serán llevados a cabo con respecto a los valores de fase de subbanda de una señal de audio decodificada para restaurar la VPC original de la señal de audio.

[0067] La figura 3 ilustra otra realización. En la figura 3, se proporciona un sistema. El sistema comprende un codificador 310 y al menos un decodificador. Mientras que la figura 3 ilustra solamente un solo decodificador 320, otras realizaciones pueden comprender más de un decodificador. El codificador 310 de la figura 3 puede ser un codificador de la realización de la figura 2. El decodificador 320 de la figura 3 puede ser el decodificador de la realización de la figura 1a o de la realización de la figura 1b. El codificador 310 de la figura 3 está configurado para transformar una señal de entrada de audio para obtener una señal de audio transformada (no mostrada). Además, el codificador 310 está configurado para codificar la señal de audio transformada para obtener una señal de audio codificada. Además, el codificador está configurado para codificar información de control que indica una coherencia de fase vertical de la señal de audio transformada. El codificador está dispuesto para alimentar la señal de audio codificada y la información de control en al menos un decodificador.

[0068] El decodificador 320 de la figura 3 está configurado para decodificar la señal de audio codificada para obtener una señal de audio decodificada (no mostrada). Además, el decodificador 320 está configurado para ajustar la señal de audio decodificada con base en la información de control codificada para obtener una señal de audio fase-ajustada.

[0069] Resumiendo lo anterior, las realizaciones descritas anteriormente tienen como objetivo conservar la coherencia de fase vertical de las señales especialmente en porciones de señal con un alto grado de coherencia de fase vertical.

[0070] Los conceptos propuestos mejoran la calidad perceptual que es alimentada por un sistema de procesamiento de audio, en lo siguiente, también referido como "sistema de audio", al medir las características de VPC de la señal de entrada al sistema de procesamiento de audio y al ajustar la VPC de la señal recibida producida por el sistema de audio en base a las características de VPC medidas para formar una señal de salida final, de tal manera que se obtiene la VPC propuesta de la señal de la salida final.

[0071] La figura 4 muestra un sistema de procesamiento de audio general que es mejorado por la realización descrita anteriormente. En particular, la figura 4 ilustra un sistema para procesamiento de VPC. De la señal de entrada de un sistema de audio 410, un generador de control de VPC 420 mide la VPC y/o su saliencia perceptual y genera una información de control de VPC. La salida del sistema de audio 410 es alimentada a una unidad de ajuste de VPC 430 y la información de control de VPC es mostrada en la unidad de ajuste de VPC 430 para restablecer la VPC.

[0072] Como un caso práctico importante, este concepto puede ser aplicado por ejemplo a códecs de audios convencionales al medir la VPC y/o la saliencia perceptual de corriente de fase y el lado del codificador, transmitir información lateral compacta apropiada junto con la señal de audio codificada y restaurar la señal de la VPC en el decodificador, con base en la información lateral compacta transmitida.

[0073] La figura 5 ilustra un codificador y decodificador de audio perceptual de acuerdo con una realización. En particular, la figura 5 ilustra un códec de audio perceptual que implementa un procesamiento de VPC de dos lados.

[0074] En el lado del codificador se ilustran una unidad de codificación 510, un generador de control de VPC 520 y una unidad de multiplex de corrientes de luz 530. En el lado del decodificador se ilustran una unidad de desmultiplexión de la corriente de *bits* 540, una unidad de decodificación 550 y una unidad de ajuste de VPC 560.

[0075] En el lado del codificador, se genera información de control de control de VPC por el generador de control de VPC 520 y es codificada como información lateral compacta que es multiplexada por la unidad de multiplexión 530 a la corriente de *bits* junto con la señal de audio codificada. La generación de información de control de VPC puede ser tiempo-frecuencia selectiva de tal manera que la VPC es solamente medida y la información de control es solamente codificada donde es perceptualmente beneficioso.

[0076] En el lado del decodificador, la información de control de VPC es extraída por la unidad de desmultiplexión de corriente de *bits* 540 de la corriente de *bits* y es aplicada en la unidad de ajuste de VPC 560 con el fin de restablecer la VPC apropiada.

[0077] La figura 6 ilustra algunos detalles de una implementación posible de un generador de control de VPC 600. En la señal de audio de entrada, la VPC es medida por una unidad de medición de VPC 610 y la saliencia perceptual de la VPC es medida por una unidad de medición de saliencia de VPC 620. De estas, se deriva la información de control de VPC por una unidad de derivación de información de control de VPC 630. La entrada de audio puede comprender más de una señal de audio, por ejemplo, además de la primera entrada de audio, una

segunda entrada de audio que comprende una versión procesada de la primera señal de entrada (véase la figura 5) puede ser aplicada al generador de control de VPC.

[0078] En algunas realizaciones, el lado del codificador puede comprender un generador de control de VPC para medir la VPC de la señal de entrada y/o medición de la saliencia perceptual de la VPC de la señal de entrada. El generador de control de VPC puede proporcionar información de control de VPC para controlar el ajuste de la VPC en el lado del decodificador. Por ejemplo, la información de control puede señalar la habilitación o deshabilitación del ajuste de VPC del lado del decodificador o la información de control puede determinar la intensidad del ajuste de la VPC del lado del decodificador.

[0079] Ya que la coherencia de fase vertical es importante para la calidad subjetiva de la señal de audio, si la señal es tonal y/o armónica y si su tono no cambia demasiado rápidamente, una implementación típica de una unidad de control de VPC puede incluir un detector de tono o un detector de armonía o al menos un detector de variación de tono, proporcionando una medida de la intensidad de tono.

[0080] Además, la información de control generada por el generador de control de VPC puede señalar la intensidad de la VPC de la señal original. De otra manera, la información de control puede señalar un parámetro de modificación que acciona el ajuste de la VPC del decodificador de tal manera que, después del ajuste de la VPC del lado del decodificador, la VPC percibida de la señal original es restaurada aproximadamente. Alternativa o adicionalmente, uno o varios valores de VPC objetivo que se van a restablecer pueden ser señalados.

[0081] La información de control de VPC puede ser transmitida de forma compacta del codificador al lado del decodificador, por ejemplo, al incrustarla a la corriente de *bits* como información lateral adicional.

[0082] En realizaciones, el decodificador puede estar configurado para leer la información de control de VPC proporcionada por el generador de control de VPC del lado del codificador. Para este propósito, el decodificador puede leer la información de control de VPC de la corriente de *bits*. Además, el decodificador puede estar configurado para procesar la salida del decodificador de audio regular dependiendo de la información de control de VPC al emplear una unidad de ajuste de VPC. Además, el decodificador puede estar configurado para integrar la señal de audio procesada como la señal de audio de salida.

[0083] En lo sucesivo, se proporciona un generador de control de VPC del lado del codificador de acuerdo con una realización.

[0084] Se pueden identificar señales periódicas casi estacionarias que exhiben una alta VPC mediante el uso de un detector de tono (como son bien conocidos de por ejemplo codificación de habla o análisis de señal de música) que alimenta una medición de la intensidad de tono y/o el grado de periodicidad. La VPC real puede ser medida mediante la aplicación de un banco de filtros de coclear, una detección de envolvente de subbanda subsecuente seguida por una suma de envolventes de coclear a través de la frecuencia. Si por ejemplo, las envolventes de subbanda son coherentes, la suma alimenta una señal no plana temporalmente, mientras que las envolventes de subbanda no coherentes se suman a una señal temporalmente más planas. A partir de la evaluación combinada (por ejemplo, al comparar con umbrales predefinidos, respectivamente) de la intensidad de tono y/o grado de periodicidad y medida de VPC, la información de control de VPC puede ser derivada, que está constituida por ejemplo por una bandera de señal que denota "ajuste de VPC encendido" o de otra manera "ajuste de VPC apagado".

[0085] Los eventos semejantes a impulsos en un dominio de tiempo exhiben una coherencia de fase fuerte a pesar de sus representaciones espectrales. Por ejemplo, un impulso de Dirac Fourier-transformado tiene un espectro plano con fases que se incrementan linealmente. Lo mismo es cierto para una serie de impulsos periódicos que tienen una frecuencia base de f_0 . Aquí, el espectro es un espectro de línea. Estas líneas individuales que tienen una distancia de frecuencia de f_0 son también coherentes de fase. Cuando su coherencia de fase es alterada (las magnitudes permanecen sin modificar), la señal de dominio de tiempo resultante ya no es una serie de impulsos de Dirac, sino que en lugar de esto, los impulsos han sido ampliados significativamente en el tiempo. Esta modificación es audible y es particularmente relevante para sonidos que son similares a una serie de impulsos, por ejemplo, habla de voz, instrumentos de latón o cuerdas arqueadas.

[0086] Por consiguiente, la VPC puede ser medida indirectamente al determinar la no llanura local de una envolvente de una señal de audio en el tiempo (los valores absolutos de la envolvente pueden ser considerados).

[0087] Al sumar las envolventes de subbanda a través de la frecuencia, se puede determinar si las envolventes se suman hasta una envolvente combinada plana (baja VPC) o a una envolvente combinada no plana (alta VPC). El

concepto propuesto es particularmente ventajoso si las envolventes sumadas se relacionan con bandas de frecuencia auralmente exactas adaptadas perceptualmente.

[0088] La información de control puede por ejemplo ser generada entonces al calcular una proporción de una media geométrica de la envolvente combinada con una media aritmética de la envolvente combinada.

[0089] Alternativamente, el valor máximo de la envolvente combinada puede ser comparado con un valor medio de la envolvente combinada. Por ejemplo, una proporción máxima/media puede ser formada, por ejemplo una proporción del valor máximo de la envolvente combinada al valor medio de la envolvente combinada.

[0090] En lugar de formar una envolvente combinada, por ejemplo, una suma de envolventes, los valores de fase del espectro de la señal de audio que serán codificados por sí mismos serán examinados en cuanto a predictibilidad. Una alta predictibilidad indica una alta VPC. Una baja predictibilidad indica una baja VPC.

[0091] El uso de un banco de filtros de coclear es particularmente ventajoso con respecto a señales de audio, si la VPC o la saliencia de VPC será definida como una medida psicoacústica. Puesto que la elección de un ancho de banda de filtro particular define, cuales tonos parciales del espectro se relacionan con una subbanda común y así contribuyen conjuntamente para formar una cierta envolvente de subbanda, los filtros adaptados perceptualmente pueden modelar el procesamiento interno del sistema auditivo humano más exactamente.

[0092] La diferencia en percepción aurale entre una señal fase-coherente y una señal fase-incoherente que tiene los mismos espectros de magnitud es además dependiente de la dominancia de los componentes espectrales armónicos en la señal (o en la pluralidad de señales). Una baja frecuencia base, por ejemplo de 100 Hz de aquellos componentes armónicos incrementa la diferencia que una alta frecuencia base reduce la diferencia, debido a que una frecuencia base baja da como resultado más sobretonos siendo asignados a la misma subbanda. Aquellos sobretonos en la misma subbanda otra vez se suman y su envolvente de subbanda puede ser examinado.

[0093] Además, la amplitud de los sobretonos es relevante. Si la amplitud de los sobretonos es alta, el incremento de la envolvente de dominio de tiempo se vuelve más agudo. La señal se vuelve más semejante a impulso y así, la VPC se vuelve incrementadamente importante, por ejemplo, la VPC se vuelve más alta.

[0094] En lo sucesivo, se proporciona una unidad de ajuste de VPC del lado del decodificador de acuerdo con una realización. Tal unidad de ajuste de VPC puede comprender información de control que comprende una bandera de información de control de VPC.

[0095] Si la bandera de información de control de VPC denota "ajuste de VPC apagado" ningún procesamiento de VPC dedicado es aplicado ("pasante" o alternativamente un simple retardo) si la bandera dice "ajuste de VPC encendido" el segmento de señal es descompuesto por un banco de filtro de análisis y se inicia una medición de la fase $p_0(f)$ de cada línea espectral a la frecuencia f . A partir de esto, desplazamientos de ajuste de fase $dp(f) = \alpha * (p_0(f) + \text{const})$ son calculados donde "const" denota un ángulo en radianes entre $-\pi$ y π . Para dicho segmento de señal y los siguientes segmentos consecutivos, donde "ajuste de VPC encendido" es señalado, las fases $p_x(f)$ de las líneas espectrales $x(f)$ son ajustadas a continuación para ser $p_x'(f) = p_x(f) - dp(f)$. La señal ajustada en VPC es finalmente convertida al dominio de tiempo por un banco de filtros de síntesis.

[0096] El concepto está basado en la idea de llevar a cabo una medición inicial para determinar la desviación de una respuesta de fase ideal. Esta desviación es compensada más tarde. α puede ser un ángulo en el intervalo de $0 \leq \alpha \leq 1$, $\alpha=0$ significa ninguna compensación, $\alpha=1$ significa plena compensación con respecto a la respuesta de fase ideal. La respuesta de fase ideal puede por ejemplo ser la respuesta de fases resultante en una respuesta de fase con una llanura máxima. "const" es un ángulo aditivo fijo que no cambia la coherencia de fase, pero que permite direccionar fases absolutas alternativas y así generar señales correspondientes, por ejemplo la transformada de Hilbert de la señal cuando const es 90° .

[0097] La figura 7 ilustra un aparato para procesar una primera señal de audio para obtener una segunda señal de audio de acuerdo con otra realización. El aparato comprende un generador de información de control 710 y una unidad de ajuste de fase 720. El generador de información de control 710 está adaptado para generar información de control, de tal manera que la información de control indica una coherencia de fase vertical de la primera señal de audio. La unidad de ajuste de fase es apta para ajustar la primera señal de audio para obtener la segunda señal de audio. Además, la unidad de ajuste de fase 720 es apta para ajustar la primera señal de audio en base a la información de control.

[0098] La figura 7 es una realización de un solo lado. La determinación de la información de control y los ajustes de fase llevados a cabo no son divididos entre el codificador (señalización de información de control) y el decodificador (ajuste de fase). En lugar de esto, la generación de información de control y el ajuste de fase son llevados cabo por un solo aparato o sistema.

[0099] En la figura 8, la VPC es manipulada en el decodificador direccionado por la información de control también generada en el lado del decodificador ("sistema de un solo lado"), donde la información de control es generada al analizar la señal de audio decodificada. En la figura 8, se ilustra un códec de audio perceptual con un procesamiento de VPC de un solo lado de acuerdo con un ejemplo.

[0100] Un sistema de un solo lado de acuerdo con ejemplos se ilustra por ejemplo por la figura 7 y la figura 8 puede tener las siguientes características:

La salida de cualquier proceso de procesamiento de señal existente o de un sistema de audio, por ejemplo la señal de salida de un decodificador de audio, es procesada sin tener acceso a información de control de VPC que es generada con acceso a una señal sin deteriorar/original (por ejemplo, en el lado del codificador). En lugar de esto, la información de control de VPC puede ser generada directamente de la señal dada, por ejemplo de la salida de un sistema de audio, por ejemplo un decodificador (la información de control de VPC puede ser generada "ciegamente").

[0101] La información de control de VPC para controlar el ajuste de VPC puede comprender, por ejemplo, señales para habilitar/deshabilitar la unidad de ajuste de VPC o para determinar la intensidad de ajuste de VPC o la información de control de VPC puede comprender uno o varios valores de VPC objetivos que se van a establecer.

[0102] Además, el procesamiento puede ser efectuado en una etapa de ajuste de VPC (una unidad de ajuste de VPC) que usa la información de control de VPC generada ciegame y la alimenta a la salida como la salida del sistema.

[0103] En lo sucesivo, se proporciona una realización de un generador de control de VPC del lado del decodificador. El generador de control del lado del decodificador puede ser bastante similar al generador de control del lado del codificador. Puede comprender, por ejemplo, un detector de tono que alimenta una medición de la intensidad de tono y/o el grado de periodicidad y una comparación con un umbral predefinido. Sin embargo, el umbral puede ser diferente del usado en el generador de control del lado del codificador puesto que el generador de VPC del lado del decodificador opera en la señal ya distorsionada de VPC. Si la distorsión de VPC es suave, también la VPC restante puede ser medida y comparada con un umbral dado con el fin de generar información de control de VPC.

[0104] De acuerdo con una realización preferida, si la VPC medida es alta, se aplica modificación de VPC con el fin de incrementar adicionalmente la VPC de la señal de salida y si la VPC medida es baja, no se aplica ninguna modificación de VPC. Puesto que la conservación de la VPC es más importante para señales tonales y armónicas, tal procesamiento de VPC de acuerdo con una realización preferida, un detector de tono o al menos un detector de variación de tono puede ser usado, proporcionando una medida de la intensidad del tono dominante.

[0105] Finalmente, la estrategia de dos lados y la estrategia de un solo lado pueden ser combinadas, donde el proceso de ajuste de VPC es controlado tanto por la información de control de VPC transmitida derivada de una señal original/sin deteriorar como información extraída de los procesos de la señal de audio (por ejemplo, decodificada). Por ejemplo, un sistema combinado resulta de tal combinación.

[0106] Aunque algunos aspectos han sido descritos en el contexto de un aparato, está claro que estos aspectos también representan una descripción del procedimiento correspondiente, donde un bloque o dispositivo corresponde a una etapa del procedimiento o un elemento de la etapa del procedimiento. Análogamente, los aspectos descritos del contexto de una etapa del procedimiento también representan una descripción de un bloque o ítem o elemento correspondiente de un aparato correspondiente.

[0107] Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden ser implementadas en elementos físicos o en elementos de programación. La implementación puede ser efectuada utilizando un medio de almacenamiento digital, por ejemplo disco *floppy*, un DVD, un CD, un ROM, un PROM, un EPROM, un EEPROM o una memoria *flash* que tiene señales de control que se pueden leer electrónicamente almacenadas en el mismo que cooperan (o son capaces de cooperar) con un sistema informático programable de tal manera que se efectúa el procedimiento respectivo.

[0108] Algunas realizaciones de acuerdo con la invención comprenden un soporte de datos que tiene señales de control que se pueden leer electrónicamente, que son capaces de cooperar con un sistema informático programable, de tal manera que se efectúa uno de los procedimientos descritos en esta invención.

5 **[0109]** En general, las realizaciones de la presente invención pueden ser implementadas como un producto de programa informático con un código de programa, el código del programa es operativo para efectuar uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede, por ejemplo, ser almacenado en un soporte que se puede leer por máquina.

10 **[0110]** Otras realizaciones comprenden el programa informático para efectuar uno de los procedimientos descritos en esta invención, almacenados en un soporte que se puede leer por máquina o un medio de almacenamiento no transitorio.

[0111] En otras palabras, una realización del procedimiento de la invención es por consiguiente un programa
15 informático que tiene un código de programa para efectuar uno de los procedimientos descritos en la presente, cuando el programa informático se ejecuta en un ordenador.

[0112] Una realización adicional de los procedimientos de la invención es por consiguiente un soporte de datos
20 (o un medio de almacenamiento digital o un medio que se puede leer por ordenador) que comprende, grabado en el mismo, el programa informático para efectuar uno de los procedimientos de esta invención.

[0113] Una realización adicional del procedimiento de la invención es, por consiguiente, una corriente de datos o una secuencia de señales que representan el programa informático para efectuar uno de los procedimientos descritos en esta invención. La corriente de datos o la secuencia de señales puede, por ejemplo, estar configurada
25 para ser transferida vía una conexión de comunicación de datos, por ejemplo, vía internet.

[0114] Una realización adicional comprende un medio de procesamiento, por ejemplo un ordenador o un dispositivo lógico programable, configurado para o adaptado para efectuar uno de los procedimientos descritos en esta invención.

30 **[0115]** Una realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para efectuar uno de los procedimientos descritos en esta invención.

[0116] En algunas realizaciones, un dispositivo lógico programable (por ejemplo, un arreglo de compuerta
35 programable en el campo) puede ser usado para efectuar algunas o todas las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones, un arreglo de compuerta programable en el campo puede cooperar con un microprocesador con el fin de efectuar uno de los procedimientos descritos en esta invención. En general, los procedimientos son efectuados preferiblemente por cualquier aparato de *hardware*.

40 **[0117]** Las realizaciones descritas anteriormente son solamente ilustrativas para los principios de la presente invención. Se comprenderá que modificaciones y variaciones de los arreglos y los detalles son descritos en esta invención serán evidentes a otros experimentados en la técnica. Se pretende por consiguiente estar limitado solamente por el alcance de las reivindicaciones de patente pendientes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en esta invención.

45

REFERENCIAS

[0118]

- 5 [1] Painter, T.; Spanias, A. Perceptual coding of digital audio, Proceedings of the IEEE, 88(4), 2000; pp. 451- 513.
- [2] Larsen, E.; Aarts, R. Audio Bandwidth Extension: Application of psychoacoustics, signal processing and loudspeaker design, John Wiley and Sons Ltd, 2004, Chapters 5, 6.
- [3] Dietz, M.; Liljeryd, L.; Kjørling, K.; Kunz, O. Spectral Band Replication, a Novel Approach in Audio Coding, 112th
10 AES Convention, April 2002, Preprint 5553.
- [4] Nagel, F.; Disch, S. ; Rettelbach, N. A Phase Vocoder Driven Bandwidth Extension Method with Novel Transient Handling for Audio Codecs, 126th AES Convention, 2009.
- 15 [5] Faller, C.; Baumgarte, F. Binaural Cue Coding- Part II: Schemes and applications, IEEE Trans. On Speech and Audio Processing, Vol. 11, No. 6, Nov. 2003.
- [6] Schuijers, E.; Breebaart, J.; Purnhagen, H.; Engdegard, J. Low complexity parametric stereo coding, 116th AES
20 Convention, Berlin, Germany, 2004; Preprint 6073.
- [7] Herre, J.; Kjørling, K.; Breebaart, J. et al. MPEG Surround - The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding, Journal of the AES, Vol. 56, No. 10 11, November 2008; pp. 932-955.
- [8] Laroche, J.; Dolson, M., "Phase-vocoder: about this phasiness business," Applications of Signal Processing to Audio
25 and Acoustics, 1997. 1997 IEEE ASSP Workshop on, vol., no., pp.4 pp., 19-22, Oct 1997.
- [9] Purnhagen, H.; Meine, N., "HILN-the MPEG-4 parametric audio coding tools," Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on, vol.3, no., pp.201-204 vol.3, 2000
- 30 [10] Oomen, Werner; Schuijers, Erik; den Brinker, Bert; Breebaart, Jeroen," Advances in Parametric Coding for High-Quality Audio," Audio Engineering Society Convention 114, preprint, Amsterdam/NL, March 2003.
- [11] van Schijndel, N.H.; van de Par, S.; , "Rate- distortion optimized hybrid sound coding," Applications of Signal Processing to Audio and Acoustics, 2005. IEEE Workshop on, vol., no., pp. 235-238, 16-19 Oct. 2005.
- 35 [12] <http://people.xiph.org/~xiphmont/demo/ghost/demo.html>.
- [13] D. Griesinger 'The Relationship between Audience Engagement and the ability to Perceive Pitch, Timbre, Azimuth and Envelopment of Multiple Sources' Tonmeister Tagung 2010.
- 40 [14] D. Dorran and R. Lawlor, "Time-scale modification of music using a synchronized subband/timedomain approach," IEEE International Conference on Acoustics, Speech and Signal Processing, pp. IV 225- IV 228, Montreal, May 2004.
- [15] J. Laroche, "Frequency-domain techniques for high quality voice modification," Proceedings of the International
45 Conference on Digital Audio Effects, pp. 328- 322, 2003.

REIVINDICACIONES

1. Un decodificador para decodificar una señal de audio codificada para obtener una señal de audio fase-ajustada, que comprende:
 - 5 una unidad de decodificación (110), para decodificar la señal de audio codificada para obtener una señal de audio decodificada y
caracterizado por
 una unidad de ajuste de fase (120; 430; 560) para ajustar la señal de audio decodificada para obtener la señal de audio fase-ajustada,
 - 10 donde la unidad de ajuste de fase (120; 430; 560) está configurada para recibir información de control dependiendo de una coherencia de fase vertical de la señal de audio codificada y
 donde la unidad de ajuste de fase (120; 430; 560) está adaptada para ajustar la señal de audio decodificada con base en la información de control.
- 15 2. Un decodificador de acuerdo con la reivindicación 1,
 donde la unidad de ajuste de fase (120; 430; 560) está configurada para ajustar la señal de audio decodificada cuando la información de control indica que el ajuste de fase está activado y
 donde la unidad de ajuste de fase (120; 430; 560) está configurada para no ajustar la señal de audio decodificada cuando la información de control indica que el ajuste de fase esta desactivado.
- 20 3. Un decodificador de acuerdo con la reivindicación 1,
 donde la unidad de ajuste de fase (120; 430; 560) está configurada para recibir la información de control, donde la información de control comprende un valor de intensidad que indica la intensidad del ajuste de fase y
 donde la unidad de ajuste de fase (120; 430; 560) está configurada para ajustar la señal de audio decodificada con
 25 base en el valor de la intensidad.
4. Un decodificador de acuerdo con una de las reivindicaciones 1 a 3,
 donde el decodificador comprende además un banco de filtro de análisis para descomponer la señal de audio decodificada en una pluralidad de señales de subbanda de una pluralidad de subbandas,
 30 donde la unidad de ajuste de fase (120; 430; 560) está configurada para determinar una pluralidad de primeros valores de fase de la pluralidad de señales de subbanda y
 donde la unidad de ajuste de fase (120; 430; 560) está adaptada para ajustar la señal de audio codificada al modificar al menos algunos de la pluralidad de los primeros valores de fase para obtener segundos valores de fase de la señal de audio fase-ajustada.
- 35 5. Un decodificador de acuerdo con la reivindicación 4,
 donde la unidad de ajuste de fase (120; 430; 560) está configurada para ajustar al menos algunos de los valores de fase al aplicar las fórmulas:

$$px'(f) = px(f) - dp(f) \text{ y}$$

$$dp(f) = \alpha * (p0(f) + \text{const}),$$
 40 donde f es una frecuencia que indica la una de las subbandas que tiene la frecuencia f como una frecuencia central,
 donde px(f) es uno de los primeros valores de fase de una de las señales de subbanda de una de las subbandas que tienen la frecuencia f como la frecuencia central,
 donde px'(f) es uno de los segundos valores de fase de una de las señales de subbanda de una de las subbandas que
 45 tienen la frecuencia f como la frecuencia central,
 donde const es un primer ángulo en el intervalo $-\pi \leq \text{const} \leq \pi$,
 donde α es un número real en el intervalo $0 \leq \alpha \leq 1$; y
 donde p0(f) es un segundo ángulo en el intervalo de $-\pi \leq p0(f) \leq \pi$, donde el segundo ángulo p0(f) es asignado a una de las subbandas que tiene la frecuencia f como la frecuencia central.
- 50 6. Un decodificador de acuerdo con la reivindicación 4,
 donde la unidad de ajuste de fase (120; 430; 560) está configurada para ajustar al menos algunos de los valores de fase al multiplicar al menos algunas de la pluralidad de señales de subbanda por un término de fase exponencial,
 donde el término de fase exponencial es definido por la fórmula $e^{-jdp(f)}$,
 55 donde la pluralidad de señales de subbanda son señales de subbanda complejas y donde j es el número imaginario unitario.
7. Un decodificador de acuerdo con una de las reivindicaciones anteriores,
 donde el decodificador comprende además un banco de filtros de síntesis (125),

donde la señal de audio fase-ajustada es una señal de audio de dominio espectral fase-ajustada que se representa en un dominio espectral y

donde el banco de filtros de síntesis (125) está configurado para transformar la señal de audio de dominio espectral fase-ajustada del dominio espectral a un dominio de tiempo para obtener una señal de audio de dominio de tiempo fase-ajustada.

8. Un codificador para codificar información de control en base a una señal de entrada de audio, que comprende:

una unidad de transformación (210) para transformar la señal de entrada de audio de un dominio de tiempo a un dominio espectral para obtener una señal de audio transformada que comprende una pluralidad de señales de subbanda que son asignadas a una pluralidad de subbandas, estando el codificador **caracterizado por**

un generador de información de control (220; 420; 520; 600) para generar la información de control, de tal manera que la información de control indica una coherencia de fase vertical de la señal de audio transformada y una unidad de codificación (230) para codificar la señal de audio transformada y la información de control.

9. Un codificador de acuerdo con la reivindicación 8, donde la unidad de transformación (210) comprende un banco de filtro de coclear para transformar la señal de entrada de audio del dominio de tiempo al dominio espectral para obtener la señal de audio transformada que comprende la pluralidad de señales de subbanda.

10. Un codificador de acuerdo con la reivindicación 8 o 9, donde el generador de información de control (220; 420; 520; 600) está configurado para determinar una envolvente de subbanda para cada una de la pluralidad de señales de subbanda para obtener una pluralidad de envolventes de señal de subbanda,

donde el generador de información de control (220; 420; 520; 600) está configurado para generar una envolvente combinada basada en la pluralidad de envolventes de señal de subbanda y donde el generador de información de control (220; 420; 520; 600) está configurado para generar la información de control con base en la envolvente combinada.

11. Un codificador de acuerdo con la reivindicación 10, donde el generador de información de control (220; 420; 520; 600) está configurado para generar un número caracterizante basado en la envolvente combinada y

donde el generador de información de control (220; 420; 520; 600) está configurado para generar la información de control, de tal manera que la información de control indica que el ajuste de fase está activado cuando el número caracterizante es mayor que un valor de umbral y

donde el generador de información de control (220; 420; 520; 600) está configurado para generar la información de control, de tal manera que la información de control indica que el ajuste de fase está desactivado cuando el número caracterizante es menor o igual al valor de umbral.

12. Un codificador de acuerdo con la reivindicación 10 u 11, donde el generador de información de control (220; 420; 520; 600) está configurado para generar la información de control al calcular la proporción de una media geométrica de la envolvente combinada con una media aritmética de la envolvente combinada.

13. Un codificador de acuerdo con una de las reivindicaciones 8 a 12, donde el generador de información de control (220; 420; 520; 600) está configurado para generar la información de control, de tal manera que la información de control comprende un valor de intensidad que indica el grado de coherencia de fase vertical de las señales de subbanda.

14. Un sistema que comprende, un codificador (310) de acuerdo con una de las reivindicaciones 8 a 13 y al menos un decodificador (320) de acuerdo con una de las reivindicaciones 1 a 7, donde el codificador (310) está configurado para transformar una señal de entrada de audio para obtener una señal de audio transformada,

donde el codificador (310) está configurado para codificar la señal de audio transformada para obtener una señal de audio codificada,

donde el codificador (310) está configurado para codificar información de control que indica una coherencia de fase vertical de la señal de audio transformada,

donde el codificador (310) está dispuesto para alimentar la señal de audio codificada y la información de control al menos a un decodificador,

donde el al menos un decodificador (320) está configurado para decodificar la señal de audio codificada para obtener una señal de audio decodificada y
donde el al menos un decodificador (320) está configurado para ajustar la señal de audio decodificada en base a la información de control codificada para obtener una señal de audio fase-ajustada.

5

15. Un procedimiento para decodificar una señal de audio codificada para obtener una señal de audio fase-ajustada, que comprende:
la recepción de información de control, donde la información de control indica una coherencia de fase vertical de la señal de audio codificada,

10 la decodificación de la señal de audio codificada para obtener una seña de audio decodificada, y

estando el procedimiento **caracterizado por**

el ajuste de la señal de audio decodificada para obtener la señal de audio fase-ajustada con base en la información de control.

15 16. Un procedimiento para codificar información de control basada en una señal de entrada de audio, que comprende:

la transformación de la señal de entrada de audio de un dominio de tiempo a un dominio espectral para obtener una señal de audio transformada que comprende una pluralidad de señales de subbanda que son asignadas a una pluralidad de subbandas, estando el procedimiento **caracterizado por**

20 la generación de la información de control, de tal manera que la información de control indica una coherencia de fase vertical y la señal de audio transformada y

la codificación de la señal de audio transformada y la información de control.

17. Un programa informático para implementar un procedimiento de acuerdo con la reivindicación 15 o 16,

25 cuando es ejecutado por un ordenador o procesador de señal.

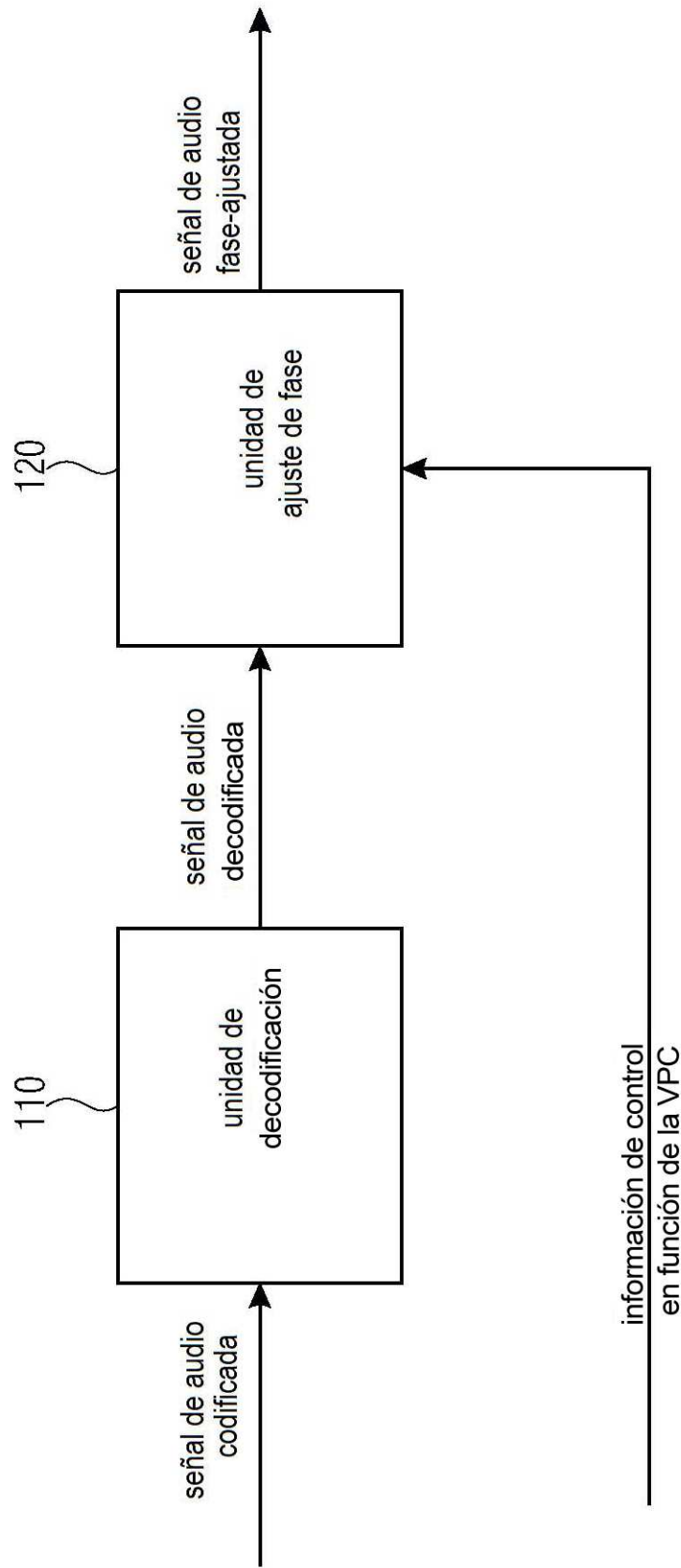


FIG 1A

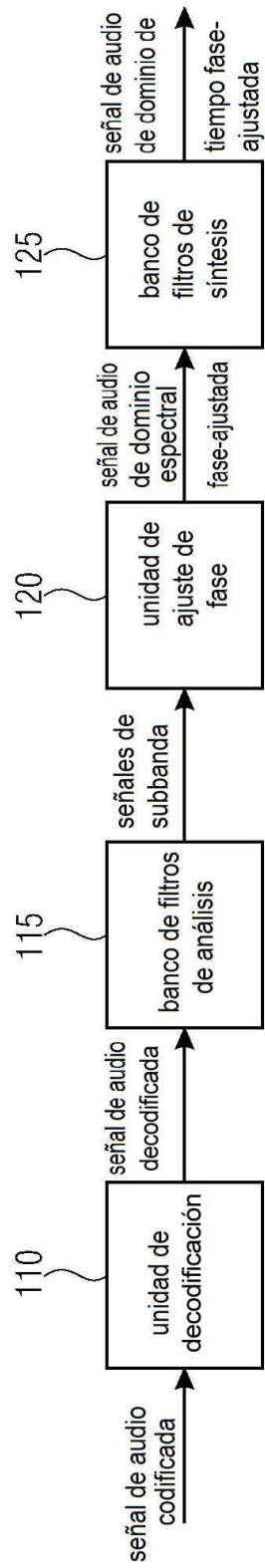


FIG 1B

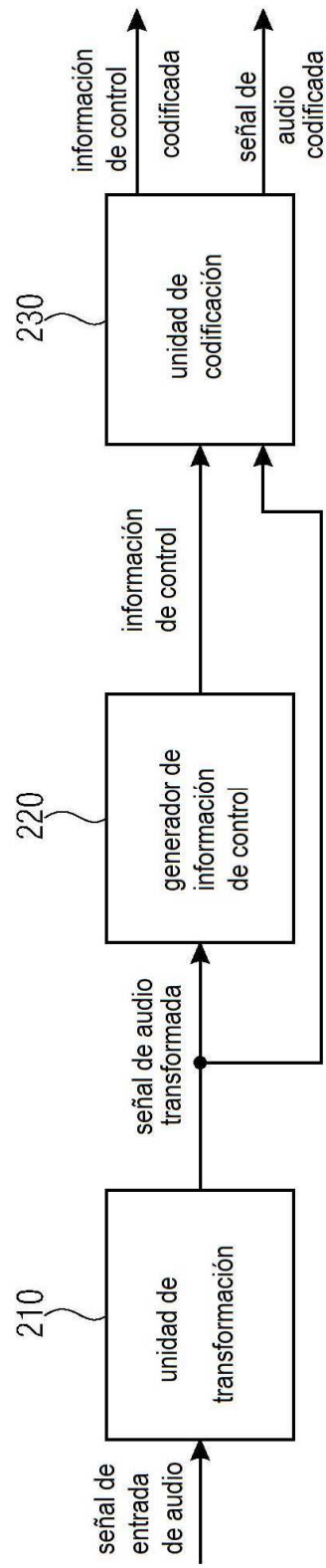


FIG 2

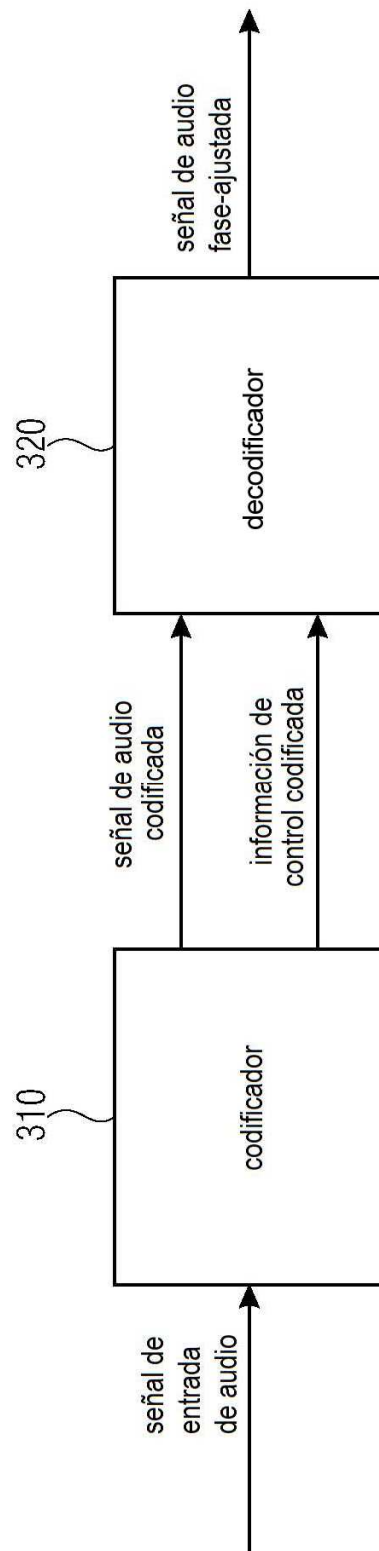


FIG 3

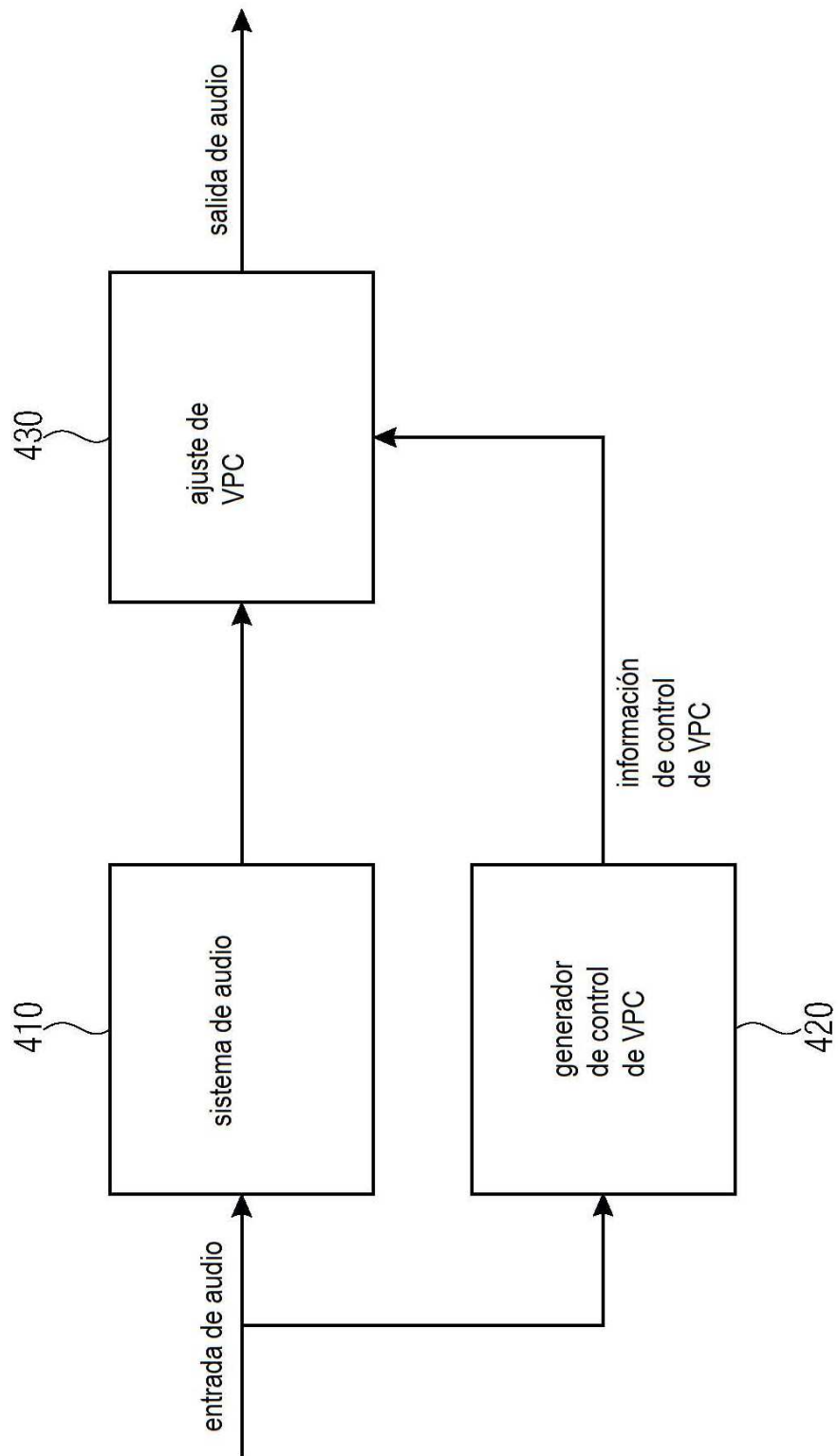


FIG 4

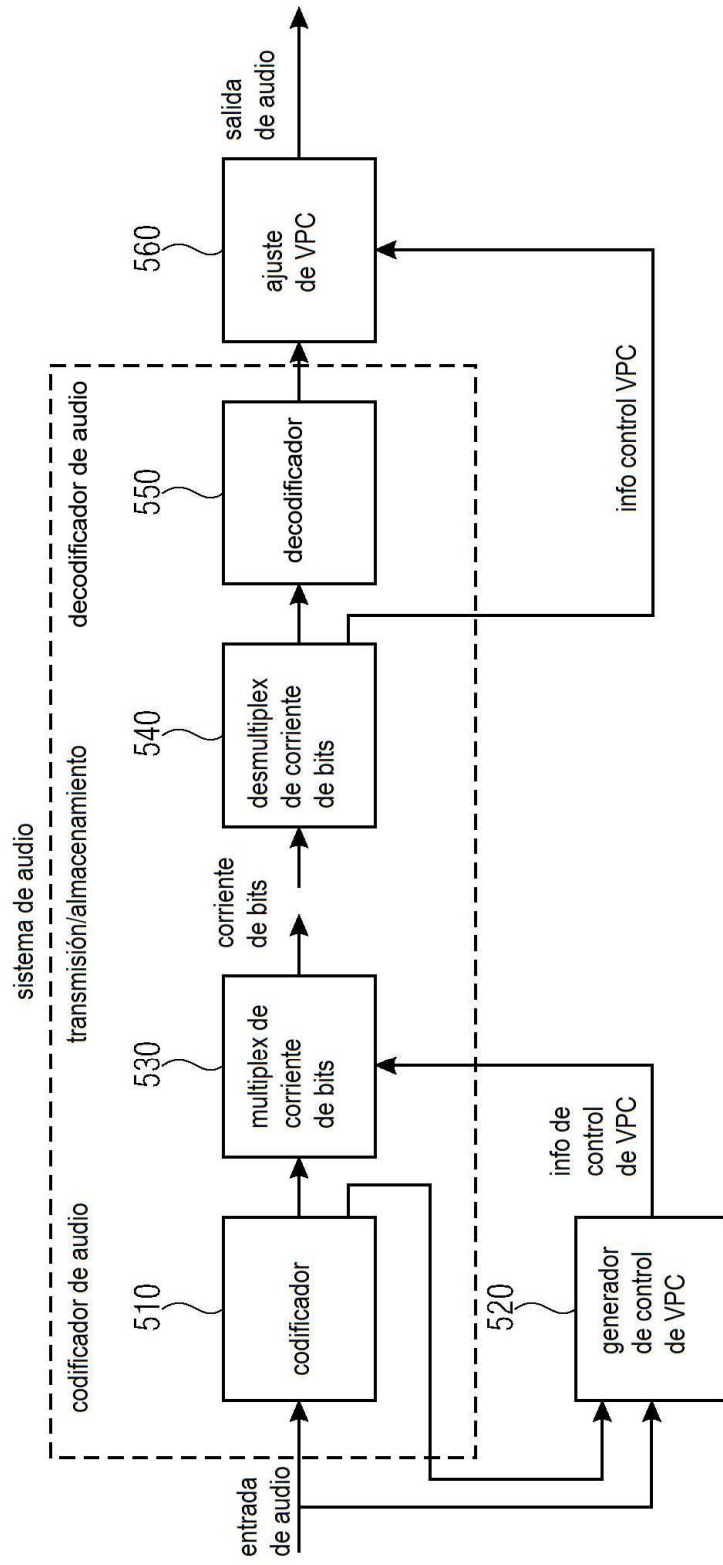


FIG 5

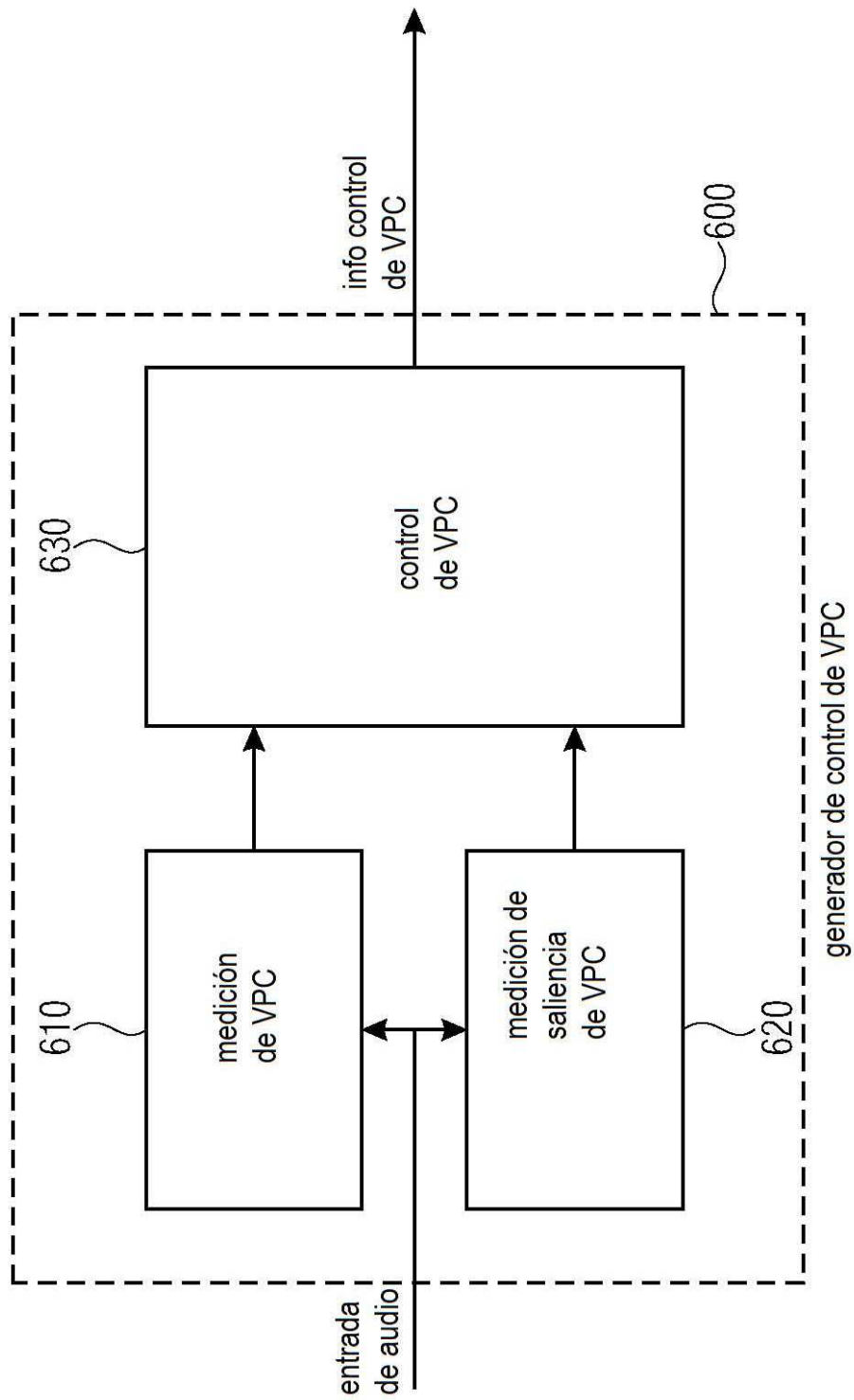


FIG 6

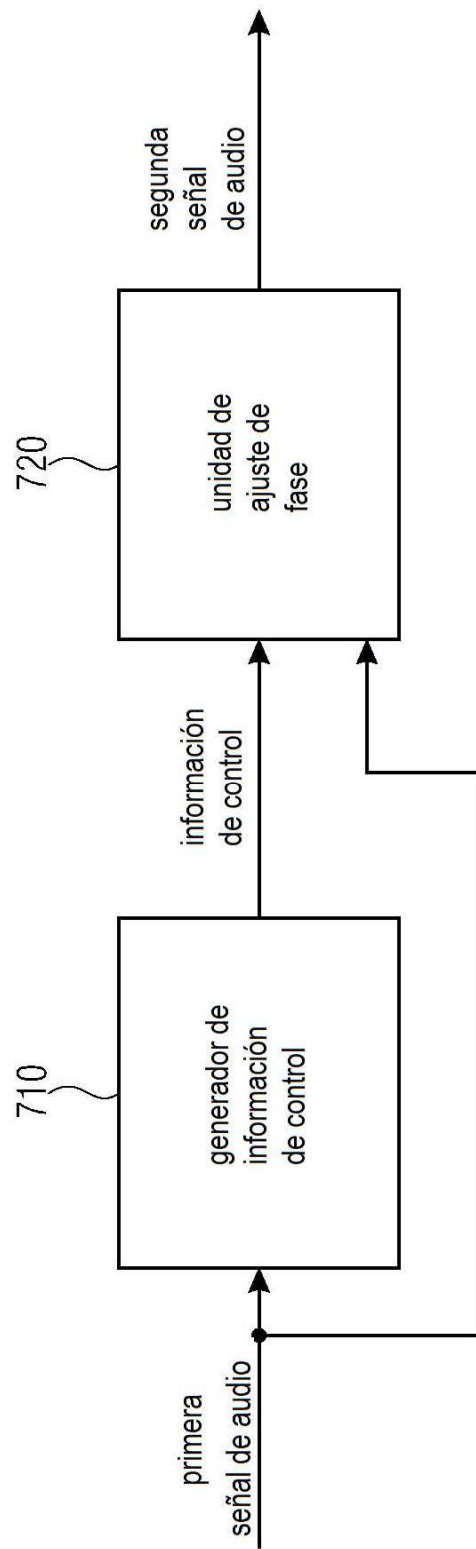


FIG 7

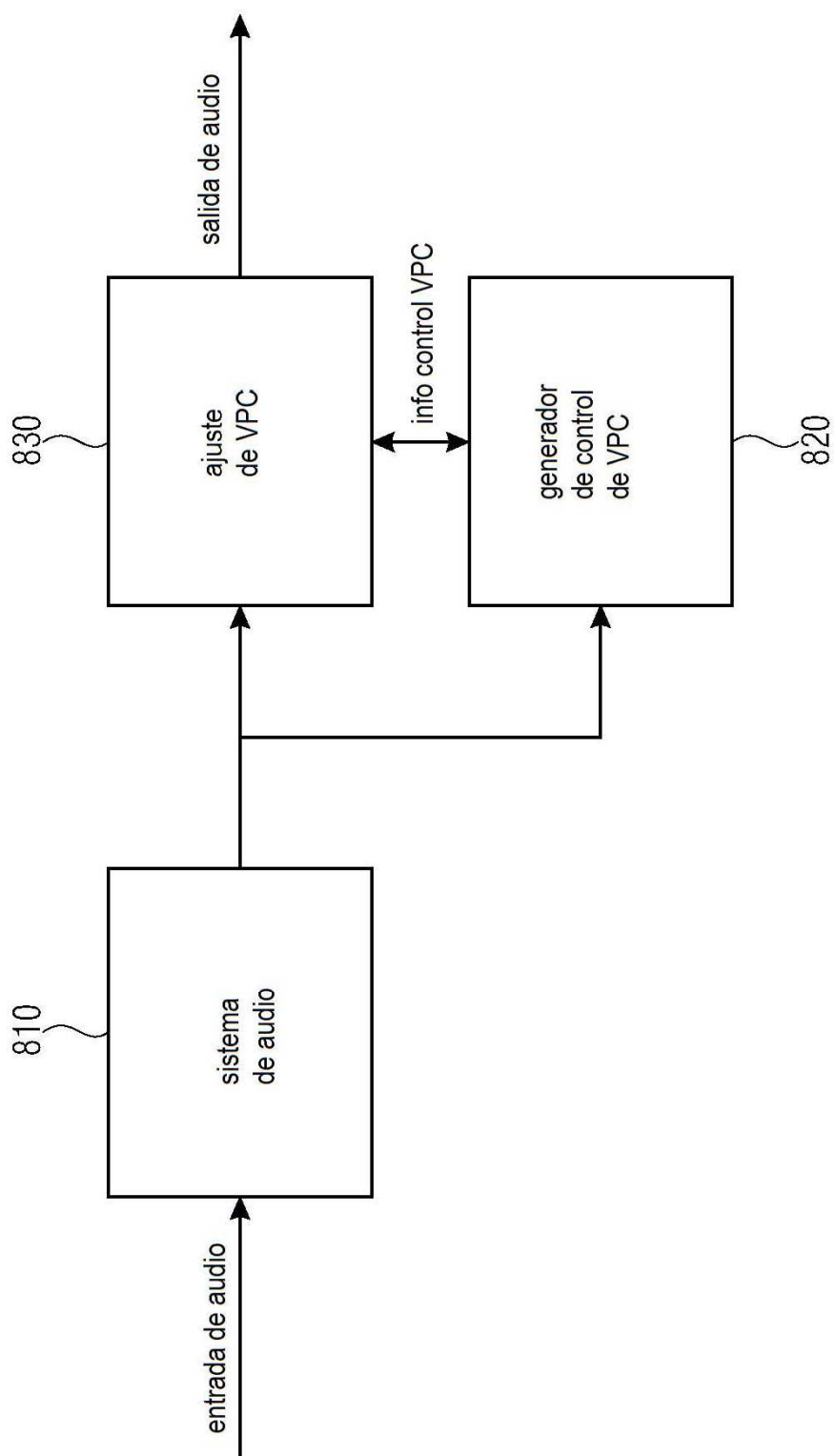


FIG 8