

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 008 258**

51 Int. Cl.:

**G10L 19/008** (2013.01)

**G10L 19/02** (2013.01)

**G10L 21/02** (2013.01)

**G10L 21/0316** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **09.03.2022 PCT/EP2022/055983**

87 Fecha y número de publicación internacional: **15.09.2022 WO22189481**

96 Fecha de presentación y número de la solicitud europea: **09.03.2022 E 22713618 (1)**

97 Fecha y número de publicación de la concesión europea: **04.12.2024 EP 4305617**

54 Título: **Decorrelacionador de audio, sistema de procesamiento y método para decorrelacionar una señal de audio**

30 Prioridad:

**11.03.2021 EP 21162142**

**20.10.2021 EP 21203832**

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

**21.03.2025**

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR  
FÖRDERUNG DER ANGEWANDTEN  
FORSCHUNG E.V. (100.00%)  
Hansastraße 27c  
80686 München, DE**

72 Inventor/es:

**DISCH, SASCHA;  
ANEMÜLLER, CARLOTTA y  
HERRE, JÜRGEN**

74 Agente/Representante:

**RUÍZ GALLEGOS, Natalia**

ES 3 008 258 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**

Decorrelacionador de audio, sistema de procesamiento y método para decorrelacionar una señal de audio

5 La presente invención se refiere a un decorrelacionador para una señal de audio, a un sistema de procesamiento que tiene tal decorrelacionador, a un método de decorrelación y a un producto de programa informático. La presente invención se refiere en particular a un decorrelacionador de señales de audio.

10 En la codificación de audio perceptual, los decorrelacionadores son un componente importante para la codificación paramétrica de audio espacial. Las soluciones conocidas se refieren a decorrelacionadores conocidos de codificación paramétrica de audio espacial como estereofónico paramétrico o envolvente de MPEG. Los decorrelacionadores, como se describe en [1] o [2] utilizan filtros de reverberación (reverb) en el dominio del tiempo computacionalmente costosos con una larga respuesta de impulso. Los decorrelacionadores, tal como se describe en [3] o [4], requieren el uso de un banco de filtros espejo en cuadratura (QMF, Quadrature Mirror Filterbank) considerable y filtros de rejilla computacionalmente costosos.

20 Por lo tanto, existe una necesidad de un decorrelacionador, un sistema de procesamiento que tenga tal decorrelacionador y un método para decorrelacionar porciones de una señal de audio que permitan un bajo retardo de procesamiento y/o decorrelación de baja complejidad computacional.

Un objeto de la presente invención es proporcionar un decorrelacionador, un sistema de procesamiento y un método para decorrelación que permitan un bajo retardo de procesamiento y/o decorrelación con una baja complejidad y alta calidad perceptual, especialmente en el procesamiento de señales que contienen transitorios.

25 Este objetivo se logra mediante el objeto definido en las reivindicaciones independientes.

30 Un hallazgo de la presente invención es que dividir una representación de frecuencia en una pluralidad de partes y para su procesamiento, es decir, retardar cada una de las partes con una unidad de retardo separada, permite un bajo retardo de procesamiento, ya que el cómputo de las diferentes partes se puede llevar a cabo en paralelo. Al mismo tiempo, tales operaciones en el dominio de la frecuencia requieren una baja complejidad computacional.

35 De acuerdo con una realización, un decorrelacionador comprende una pluralidad de unidades de retardo, en donde cada unidad de retardo está configurada para recibir una parte de una representación de frecuencia que se basa en una señal de audio, en donde cada unidad de retardo está configurada para retardar la parte recibida para proporcionar una parte retardada. El decorrelacionador comprende un conformador de envolvente configurado para recibir una combinación de señales que se basa en las partes retardadas de la representación de frecuencia, para recibir la representación de frecuencia de la señal de audio, para ajustar una energía de las partes retardadas con respecto a la representación de frecuencia de la señal de audio y para proporcionar una representación de frecuencia de forma combinada.

40 De acuerdo con una realización, diferentes partes de la representación de frecuencia comprenden un número igual o diferente de bins de frecuencia. En donde un mismo número de bins de frecuencia puede permitir un mismo tiempo de procesamiento, un número diferente de bins de frecuencia puede permitir una adaptación a los requisitos de aplicación.

45 De acuerdo con una realización, el decorrelacionador comprende un cambiador de fase configurado para cambiar la fase de la representación de frecuencia de la señal de audio, o para cambiar la fase de la señal de audio en un dominio del tiempo para obtener una señal de audio con cambio de fase. El cambio de fase puede permitir que se perciba una reverberación y, por lo tanto, una alta calidad de audio.

50 De acuerdo con una realización, el cambiador de fase está configurado para un cambio de fase de la representación de frecuencia de la señal de audio y comprende una pluralidad de filtros de paso total (Allpass), en donde cada filtro de paso total está configurado para cambiar la fase de una parte asociada de la representación de frecuencia de la señal de audio. Es decir, el filtro de paso total puede estar asociado y adaptado a la parte respectiva de la señal de audio, lo que puede permitir una alta calidad de audio en general.

55 De acuerdo con una realización, un filtro de paso total de la pluralidad de filtros de paso total comprende un conjunto de estructuras de filtro de paso total que se conectan en serie entre sí, es decir, utilizando filtros IIR de Schroeder. Las estructuras de filtro de paso total están adaptadas para proporcionar diferentes retardos de tiempo. Alternativamente o adicionalmente, las estructuras de filtro de paso total comprenden una estructura anidada de filtro de paso total.

60 De acuerdo con una realización, un número de estructuras de filtro de paso total y/o una circuitería de la estructura de filtro de paso total es equivalente o diferente entre diferentes filtros de paso total. Esto permite una alta flexibilidad del decorrelacionador.

## ES 3 008 258 T3

De acuerdo con una realización, los diferentes retardos de tiempo se basan en un múltiplo de número primo de una frecuencia de muestreo local utilizada para obtener la representación de frecuencia de la señal de audio. Esto permite que se perciba una alta calidad de audio.

5

De acuerdo con una realización, el conjunto de estructuras de filtro de paso total comprende un número de cuatro estructuras de filtro de paso total y están adaptadas para proporcionar un retardo de 1, 2, 3 y 5 unidades de tiempo. Tal unidad de tiempo se puede basar en un tamaño de bloque de la conversión en el dominio de la frecuencia. Por ejemplo, al utilizar un tamaño de bloque de 256 con una superposición del 50%, una unidad de tiempo puede resultar en 128 muestras a 48 kHz = 2,7 ms. Otras unidades de tiempo razonables pueden ser, por ejemplo, 32 o 64 muestras u otros valores. Preferiblemente, las unidades de tiempo son lo suficientemente cortas para permitir una resolución de tiempo suficiente en la conformación posterior de envolvente de tiempo/frecuencia. En una solución alternativa, se proporciona un retardo de 1, 3, 5 y 7 por las cuatro estructuras de filtro de paso total. Esto permite evitar superposiciones en el dominio del tiempo.

10

15

De acuerdo con una realización, un factor de ganancia del filtro de paso total se adapta a un valor con una magnitud, es decir, valores positivos o negativos, de 0,7 dentro de un margen de tolerancia. El margen de tolerancia es, por ejemplo, del 20%, 10% o 5%.

20

De acuerdo con una realización, el cambiador de fase está configurado para cambiar la fase de la señal de audio en un dominio del tiempo, en donde el cambiador de fase comprende un conjunto de estructuras de filtro de paso total que se conectan en serie entre sí, en donde las estructuras de filtro de paso total están adaptadas para proporcionar diferentes retardos de tiempo. Alternativamente o adicionalmente, las estructuras de filtro de paso total comprenden una estructura anidada de filtro de paso total.

25

De acuerdo con una realización, los diferentes retardos de tiempo de paso total se basan en un múltiplo de número primo de un recíproco de una frecuencia de muestreo utilizada para obtener la representación de frecuencia de la señal de audio. Al igual que en el dominio de la frecuencia, también puede obtenerse la misma ventaja en el dominio del tiempo. En el dominio del tiempo, diferentes retardos de tiempo se pueden basar en un número primo que se obtiene multiplicando cada uno de un conjunto de números primos mínimos, por ejemplo, 1, 2, 3 y 5 como un conjunto de ejemplo o 1, 3, 5 y 7 como otro conjunto de ejemplo con un factor de muestreo descendente utilizado para generar las partes de la representación de frecuencia de la señal de audio para obtener un resultado intermedio y para utilizar un siguiente número primo con respecto al resultado intermedio. Como un siguiente número primo, se puede entender una distancia más cercana, por ejemplo, para obtener el siguiente valor primo más grande o siguiente valor primo más pequeño. En el ejemplo dado, los valores 131, 257, 383 y 641 se pueden obtener para el primer conjunto y 131, 383, 641 y 907 se pueden obtener para el segundo conjunto de ejemplo. Aquí, una unidad de tiempo puede ser 1 muestra. La muestra puede referirse a una frecuencia de muestreo es, por ejemplo, de 48 kHz. En otras realizaciones, la frecuencia de muestreo también puede ser de 44,1 kHz o 32 kHz u otros valores.

30

35

40

De acuerdo con una realización, el decorrelacionador comprende una primera unidad de conversión para obtener la representación de frecuencia de la señal de audio de la señal de audio para el conformador de envolvente y que comprende una segunda unidad de conversión para obtener una representación de frecuencia de la señal de audio reverberada, en donde las partes de la representación de frecuencia forman partes de la representación de frecuencia de la señal de audio reverberada. Esto permite generar la señal utilizada formada directamente en el decorrelacionador.

45

De acuerdo con una realización, el decorrelacionador está adaptado para implementar adicionalmente un retardo igual y predefinido para un subconjunto de todas las partes de la representación de frecuencia. Es decir, un retardo que es igual para las partes respectivas o también se pueden aplicar comúnmente líneas de retardo en un módulo de retardo común que permite unidades de retardo simples en las líneas de retardo respectivas para una parte asociada.

50

De acuerdo con una realización, las unidades de retardo asociadas con una parte espectral de la pluralidad de unidades de retardo están configuradas para retardar la parte asociada de la representación de frecuencia de manera diferente cuando se compara con unidades de retardo asociadas con otras partes espectrales. Esto permite una alta calidad percibida al tratar diferentes porciones de frecuencia de manera diferente.

55

De acuerdo con una realización, la unidad de retardo está configurada para retardar partes de la representación de frecuencia que comprenden frecuencias más bajas con un retardo de tiempo superior cuando se compara con partes de la representación de frecuencia que comprende frecuencias más altas.

60

De acuerdo con una realización, una relación entre diferentes retardos de tiempo es lineal, logarítmica y/o con base en un redondeo en muestras de sub-banda. Esto permite que se perciba una alta calidad.

De acuerdo con una realización, el decorrelacionador comprende una unidad de conversión para recibir y convertir la

## ES 3 008 258 T3

- señal de audio o una versión reverberada de la señal de audio en las partes al llevar a cabo una transformada discreta de Fourier (DFT, Discrete Fourier Transform) por bloques de tiempo, o transformada de Fourier de tiempo corto (STFT, Short-Time Fourier Transform), en donde la unidad de conversión está configurada para convertir bloques que tengan una superposición del 50% dentro de un margen de tolerancia. Tal conversión por bloques permite retardos cortos para una parte respectiva que se obtenga y para un tratamiento paralelo de las diferentes partes.
- 5
- De acuerdo con una realización, el conformador de envolvente está configurado para operar en un dominio de sub-banda y con una resolución temporal de menos de 4 milisegundos.
- 10
- De acuerdo con una realización, el decorrelacionador comprende una etapa de procesamiento de señal configurada para recibir una señal con base en la representación de frecuencia conformada combinada, por ejemplo, como una señal monoaural, y para procesar la señal monoaural al menos como una señal estereofónica. Esto permite una percepción mejorada de un oyente.
- 15
- De acuerdo con una realización, el decorrelacionador comprende una etapa de procesamiento de señal configurada para procesar la representación de frecuencia conformada combinada al menos en una señal estereofónica y para modelado de extensión de la fuente con base en dicha señal estereofónica al menos, por ejemplo, en el dominio de la frecuencia.
- 20
- De acuerdo con una realización, un sistema de procesamiento comprende un decorrelacionador como se describe en el presente documento y una etapa de procesamiento para transformar una señal descompuesta media/lateral a una señal descompuesta izquierda/derecha.
- 25
- De acuerdo con las realizaciones, el sistema de procesamiento puede llevar a cabo la supresión de transitorios para suprimir ecos, por ejemplo, pre-ecos y pos-ecos provocados por un transitorio. Tal manejo de transitorios puede comprender silenciar la salida de un decorrelacionador y, en correspondencia, amplificar una salida de una unidad de compensación de retardo proporcionando lo necesario para una porción de la señal descompuesta izquierda/derecha y que es paralelo con el decorrelacionador y se conecta a la etapa de procesamiento.
- 30
- De acuerdo con una realización, un método comprende recibir una pluralidad de partes de una representación de frecuencia que se basa en una señal de audio, retardar cada una de las partes recibidas para proporcionar una pluralidad de partes retardadas y recibir y combinar señales que se basan en las partes retardadas de la representación de frecuencia. El método comprende recibir la representación de frecuencia de la señal de audio y ajustar una energía de las partes retardadas en relación con la representación de frecuencia de la señal de audio. Se proporciona una
- 35
- representación de frecuencia conformada combinada.
- De acuerdo con una realización, se proporciona un programa informático o producto de programa informático o medio de almacenamiento no transitorio que almacena en el mismo instrucciones para llevar a cabo instrucciones respectivas para ejecutar tal método, cuando se ejecuta en un ordenador.
- 40
- Se definen realizaciones convenientes adicionales en las reivindicaciones dependientes.
- Las realizaciones ventajosas se describen a mayor detalle haciendo referencia a los dibujos de acompañamiento, en los cuales:
- 45
- la figura 1 muestra un diagrama de bloques esquemático de un decorrelacionador de acuerdo con una realización;
- la figura 2 muestra un diagrama de bloques esquemático de un decorrelacionador que comprende una unidad de conversión para generar una representación de frecuencia de una señal en el dominio del tiempo de acuerdo con una
- 50
- realización;
- la figura 3 muestra un diagrama de bloques esquemático de una decorrelación que comprende adicionalmente un pre-retardo de acuerdo con una realización;
- 55
- la figura 4 muestra un diagrama de bloques esquemático de un filtro de paso total de acuerdo con una realización;
- la figura 5 muestra un diagrama de bloques esquemático de una estructura anidada de filtro de paso total de acuerdo con una realización;
- 60
- la figura 6 muestra un diagrama de bloques esquemático de un decorrelacionador que comprende un cambiador de fase configurado para operar en el dominio del tiempo de acuerdo con una realización;
- la figura 7 muestra un diagrama de bloques esquemático de un decorrelacionador que se conecta a un modelado de extensión de la fuente de acuerdo con una realización;

la figura 8 muestra un diagrama de bloques esquemático de un sistema de procesamiento de acuerdo con una realización;

- 5 la figura 9 muestra un diagrama de bloques esquemático de un sistema de procesamiento configurado para manejo de transitorios de acuerdo con una realización; y

la figura 10 muestra un diagrama de bloques esquemático de un método de acuerdo con una realización.

- 10 Los elementos iguales o equivalentes o elementos con funcionalidad igual o equivalente se denotan en la siguiente descripción por medio de números de referencia iguales o equivalentes incluso si se presentan en diferentes figuras.

En la siguiente descripción, se establece una pluralidad de detalles para proporcionar una explicación más completa de realizaciones de la presente invención. Sin embargo, será evidente para aquellos experimentados en la materia que las realizaciones de la presente invención se pueden practicar sin estos detalles específicos. En otros casos, se muestran estructuras y dispositivos bien conocidos en forma de diagrama de bloques en lugar de en detalle con el fin de evitar oscurecer las realizaciones de la presente invención. Adicionalmente, las características de las diferentes realizaciones descritas en adelante en el presente documento se pueden combinar entre sí, a menos que se indique específicamente lo contrario.

- 20 La figura 1 muestra un diagrama de bloques esquemático de un decorrelacionador 10 de acuerdo con una realización. El decorrelacionador 10 comprende un número de al menos dos unidades de retardo  $12_1$  a  $12_n$  con  $n > 1$ . Aunque la figura 1 ilustra un número de dos unidades de retardo 12, el número es preferiblemente mayor, por ejemplo, 4, 8, 16 u otros valores que se obtendrán con una potencia de 2, en donde las realizaciones no se limitan a tales números. Es decir, las realizaciones también pueden comprender un número de 3, 5, 7 o 9 unidades de retardo 12. Cada unidad de retardo está configurada para recibir una parte asociada  $14_1$  a  $14_n$  de una representación de frecuencia 14 que se basa en una señal de audio. Por ejemplo, la representación de frecuencia 14 puede ser o puede comprender un espectro que se obtiene por medio de una transformada de Fourier tal como una transformada discreta de Fourier, DFT, o una transformada de Fourier de tiempo corto, STFT. Las partes  $14_1$  a  $14_n$  se pueden obtener, por ejemplo, como una sub-banda del espectro, es decir, una parte de la representación en el dominio de la frecuencia. Por ejemplo, tal parte  $14_1$  a  $14_n$  se puede obtener utilizando una ventana apropiada.

Cada unidad de retardo  $12_1$  a  $12_n$  está configurada para retardar la parte  $14_1$  a  $14_n$  recibida para proporcionar una parte retardada  $14'_1$  a  $14'_n$ , es decir, para tener un retardo en el dominio del tiempo.

- 35 El decorrelacionador 10 comprende además un conformador de envolvente 16 configurado para recibir señales que se basan en las partes retardadas  $14'_1$  a  $14'_n$ . Tales señales pueden ser las partes retardadas  $14'_1$  a  $14'_n$  en sí o variantes procesadas de las mismas. El conformador de envolvente 16 está configurado para combinar las señales recibidas. Adicionalmente, el conformador de envolvente está configurado para recibir la representación de frecuencia 14 de la señal de audio. El conformador de envolvente 16 está configurado para ajustar una energía de las partes retardadas  $14'_1$  a  $14'_n$  en relación con la representación de frecuencia 14 de la señal de audio. El conformador de envolvente 16 está configurado para proporcionar una representación de frecuencia conformada combinada 18. En la representación de frecuencia conformada combinada 18, las partes  $14_1$  a  $14_n$  respectivas, señales que resultan de la misma respectivamente, se pueden decorrelacionar respectivamente entre sí y/o con respecto a la representación de frecuencia 14.

Aunque el conformador de envolvente 16 se ilustra para recibir la representación de frecuencia 14 combinada, como una alternativa, el conformador de envolvente 16 puede recibir la información respectiva al recibir las partes  $14_1$  a  $14_n$  posiblemente no retardadas o comúnmente tratadas.

- 50 La figura 2 muestra un diagrama de bloques esquemático de un decorrelacionador 20 de acuerdo con una realización. El decorrelacionador 20 está configurado para recibir una señal de audio 22. El decorrelacionador 20 puede comprender una unidad de conversión 24 configurada para generar la representación de frecuencia 14 mostrada en la figura 1. La unidad de conversión 24 puede proporcionar lo necesario para que las partes  $14_1$  a  $14_{16}$  se obtengan por medio de, por ejemplo, una STFT. Por ejemplo, la representación de frecuencia puede comprender un número de 129 bins de frecuencia en total. Alternativamente, se pueden utilizar 128 bins. Por ejemplo, se pueden utilizar dos tipos de transformadas digitales de Fourier (DFT, Digital Fourier Transforms), la llamada "apilada uniformemente" y una "apilada de manera desigual". Por ejemplo, como DFT "estándar", se puede considerar que la versión apilada uniformemente tiene, en el ejemplo proporcionado, 129 bandas (127 complejas, una real y una imaginaria). La apilada de manera desigual puede comprender 128 bandas (complejas). Ambas transformadas se pueden utilizar en las realizaciones descritas en el presente documento. Las partes  $14_1$  a  $14_{16}$  pueden comprender, parcialmente o completamente, un número igual o diferente de bins. Por ejemplo, la parte  $14_1$  puede comprender del primer al noveno bin, por ejemplo, 9 bins. La parte  $14_2$  comprende, por ejemplo, los bins 10 a 19 y, por lo tanto, un número de diez bins. La adaptación o selección con respecto al número de bins se puede basar en la frecuencia de muestreo que, en el

ejemplo que se ilustra, es de 48 kHz, la superposición que es, por ejemplo, del 50% y/o un número de partes 14 a 14<sub>16</sub> que se van a generar. Las partes 14<sub>1</sub> a 14<sub>16</sub> pueden comprender un número igual o diferente de bins de frecuencia, de tal manera que también se pueden generar algunas o todas las partes 14<sub>1</sub> a 14<sub>16</sub> de modo que comprendan un igual número de bins de frecuencia.

5

El decorrelacionador 20 comprende además una sección de retardo 25 que tiene líneas de retardo 12<sub>1</sub> a 12<sub>16</sub>, cada línea de retardo 12<sub>1</sub> a 12<sub>16</sub> se asocia con una parte 14<sub>1</sub> a 14<sub>16</sub> específica y configurada para recibir dicha parte, una versión procesada de la misma, respectivamente. Las unidades de retardo 12<sub>1</sub> a 12<sub>16</sub> pueden estar asociadas con una parte espectral 14<sub>1</sub> a 14<sub>16</sub> respectiva. Tal unidad de retardo 12<sub>1</sub> a 12<sub>16</sub> puede estar configurada para retardar la parte asociada de la representación de frecuencia 14 de manera diferente cuando se compara con las unidades de retardo asociadas con otras partes espectrales. Alternativamente o adicionalmente, una relación entre diferentes retardos de tiempo puede ser una lineal, logarítmica y/o basada en un redondeo en muestras de superbanda.

10

El decorrelacionador 20 comprende además un cambiador de fase 26 que está acoplado a la sección de retardo 25, el cambiador de fase 26 configurado para recibir las partes retardadas 14'<sub>1</sub> a 14'<sub>16</sub>. El cambio de fase que utiliza el cambiador de fase 26 puede permitir una reverberación en las partes de la señal. Sin embargo, de acuerdo con las realizaciones, una secuencia de la sección de retardo 25 y la sección de reverberación 26 también se puede cambiar de tal manera que una parte 14<sub>1</sub> a 14<sub>16</sub> respectiva primero puede ser objeto de un filtro de reverberación y posteriormente retardarse.

15

20

El cambiador de fase 26 puede estar configurado para cambiar la fase de la representación de frecuencia 14 de la señal de audio, una versión procesada, por ejemplo retardada, de la misma. El cambio de fase también se puede llevar a cabo antes de convertir la señal de audio 22 al dominio de la frecuencia, un cambiador de fase correspondiente puede estar configurado para cambiar la fase de la señal de audio 22 en el dominio del tiempo para obtener una señal de audio con cambio de fase. En la configuración corta, donde el cambiador de fase 26 está configurado para cambiar la fase de la representación de frecuencia de la señal de audio 14, la versión retardada de la misma, respectivamente, el cambiador de fase puede comprender una pluralidad de filtros de paso total 28<sub>1</sub> a 28<sub>16</sub>. En el ejemplo mostrado, los filtros de paso total 28<sub>1</sub> a 28<sub>16</sub> están configurados para recibir las partes retardadas 14'<sub>1</sub> a 14'<sub>16</sub>. El término filtro de paso total se debe entender en el sentido de que el intervalo de frecuencias que se van a pasar corresponde al intervalo de frecuencias de la parte 14<sub>1</sub> a 14<sub>16</sub> respectiva. En donde, este puede incluir ejemplos donde cada uno de los filtros de paso total 28<sub>1</sub> a 28<sub>16</sub> pasa el intervalo de frecuencias completo proporcionado en la representación de frecuencia, la banda de paso de diferentes filtros de paso total 28<sub>1</sub> a 28<sub>16</sub> también pueden diferir entre sí con base en los diferentes bins de frecuencia contenidos en las partes 14<sub>1</sub> a 14<sub>16</sub> respectivas.

25

30

35

Cada uno de los filtros de paso total 28<sub>1</sub> a 28<sub>16</sub> está configurado para cambiar la fase de una parte asociada de la representación de frecuencia de la señal de audio.

Es decir, un número de estructuras de filtro de paso total y/o una circuitería de la estructura de filtro de paso total puede ser el mismo, es decir, igual o comparable, o puede, alternativamente, ser diferente entre diferentes filtros de paso total 28<sub>1</sub> a 28<sub>16</sub>.

40

Un retardo de tiempo proporcionado por las líneas de retardo 12<sub>1</sub> a 12<sub>16</sub> puede ser igual o puede ser diferente para diferentes partes 14<sub>1</sub> a 14<sub>16</sub>. Como se indica en la figura 2, partes de la representación de frecuencia que comprenden frecuencias más bajas se pueden retardar con un mayor retardo de tiempo cuando se compara con partes de la representación de frecuencia que comprenden frecuencias más altas. Del bin 1 a bins superiores, puede aumentar una frecuencia representada. Como se representa en el dominio Z, el retardo de tiempo puede disminuir con un aumento de frecuencias.

45

50

Las señales 32<sub>1</sub> a 32<sub>16</sub> pueden comprender un resultado del retardo y el cambio de fase, por ejemplo, como una salida de los filtros de paso total 28<sub>1</sub> a 28<sub>16</sub>.

El conformador de envolvente 16 puede estar configurado para recibir las señales 32<sub>1</sub> a 32<sub>16</sub> y una versión no filtrada o no retardada de las mismas, es decir, las partes 14<sub>1</sub> a 14<sub>16</sub>, es decir, la representación de frecuencia de la señal de audio 22. Las partes 14 a 14<sub>16</sub> se pueden entender como sub-bandas. El conformador de envolvente 16 puede estar configurado para operar en un dominio de sub-banda. Por ejemplo, una resolución temporal del conformador de envolvente 16 puede ser como mucho o menor que 4 milisegundos, por ejemplo, 4 milisegundos, 3,5 milisegundos, 3 milisegundos o menos.

55

60

El decorrelacionador 20 puede comprender otra unidad de conversión 35 que puede proporcionar lo necesario para una operación inversa cuando se compara con la unidad de conversión 24. Por ejemplo, la tasa de conversión 34 puede llevar a cabo una transformada de Fourier de tiempo corto inversa (iSTFT, inverse Short Term Fourier Transform). La representación de frecuencia de forma combinada 18 puede comprender información con respecto al dominio de la frecuencia que está presente en cada uno de los bins de tal manera que la representación de frecuencia conformada combinada 18 se pueda tratar correspondientemente para la salida de la unidad de conversión 24. Es

decir, la unidad de conversión 34 puede recibir las versiones procesadas de las partes 14<sub>1</sub> a 14<sub>16</sub> de la representación de frecuencia 14 y para sintetizar una señal sintetizada 36 de las versiones procesadas 14'<sub>1</sub> a 14'<sub>16</sub> con base en, por ejemplo, un procedimiento de agregar superposición. La señal 36 se puede proporcionar, por ejemplo, en una interfaz 38 del decorrelacionador 20.

5

El conformador de envolvente 16 puede estar configurado para conformar bins espectrales en tiempo y/o frecuencia. La conformación se puede llevar a cabo por el conformador de envolvente 26 para bins individuales y/o para grupos de bins, por ejemplo, implementando un procesamiento de conformación común interdependiente o al menos por grupos.

10

Haciendo referencia nuevamente a la unidad de conversión 24, la misma se puede configurar para recibir y convertir la señal de audio 22 o una versión reverberada de la misma en las partes 14 a 14<sub>16</sub>, en donde el número de 16 es un ejemplo solamente. La versión reverberada de la señal de audio 22 puede ser una entrada en caso de que el cambiador de fase 26 opere en el dominio del tiempo y por lo tanto se puede disponer proceso arriba de la unidad de conversión 24. La unidad de conversión 24 puede llevar a cabo una transformada discreta de Fourier, DFT, por bloques de tiempo, o una transformada de Fourier de tiempo corto, STFT. La unidad de conversión puede estar configurada para convertir bloques que tienen una superposición de, por ejemplo, el 50% dentro de un margen de tolerancia. Por ejemplo, el margen de tolerancia puede ser del 0% tanto como sea posible, como mucho del 5%, como mucho del 10%, como mucho del 15% o más.

20

Los bloques pueden comprender una longitud de bloque de, por ejemplo, 128 muestras, 256 muestras o 512 muestras, en donde puede ser preferible un valor de 256.

25

La figura 3 muestra un diagrama de bloques esquemático de una decorrelación 30. Cuando se compara con el decorrelacionador 20, el decorrelacionador 30 puede comprender adicionalmente un pre-retardo 42, en donde el término pre-retardo no limita el retardo a que se implemente directamente antes o posteriormente a cualquier bloque específico. El pre-retardo 42 puede estar ubicado en cualquier etapa previa al conformador de envolvente 16, preferiblemente y cuando esté operando en el dominio de la frecuencia, después de la unidad de conversión 24. Es decir, por ejemplo, una secuencia entre los filtros de paso total de la reverberación o cambiador de fase 26 y el pre-retardo 42 se puede intercambiar cuando se compara con la ilustración de la figura 3. El pre-retardo 42 o el bloque de retardo 42 se puede configurar para implementar adicionalmente un retardo igual y predefinido para un subconjunto o todas las partes 14<sub>1</sub> a 14<sub>16</sub> de la representación de frecuencia. Esto puede permitir la implementación del mismo retardo para cada parte 14<sub>1</sub> a 14<sub>16</sub> o un grupo de las mismas para combinar el procesamiento en esta etapa y utilizar las líneas de retardo 12<sub>1</sub> a 12<sub>16</sub> para agregar un retardo probablemente individual que difiera del retardo común implementado en el bloque 42. Por ejemplo, el pre-retardo 42 está configurado para permitir un pre-retardo constante para todas las bandas espectrales.

30

35

40

La figura 4 muestra un diagrama de bloques esquemático de un filtro de paso total 40 de acuerdo con una realización que se puede operar al menos como parte de uno de los filtros 28<sub>1</sub> a 28<sub>16</sub> del decorrelacionador 20 y/o 30. El filtro de paso total 40 puede comprender una estructura de un filtro IIR de Schroeder, por ejemplo, y puede comprender una ramificación hacia adelante 46 en combinación con una ramificación hacia atrás 48 en combinación con un bloque de retardo 52 para proporcionar una señal de salida 54 respectiva que se base en una señal de entrada 44 del filtro de paso total 40. Un filtro de paso total 28 del decorrelacionador 20 y/o 30 puede comprender uno o más de tales filtros de paso total 40 que se conectan en serie entre sí. Para proporcionar diferentes retardos de tiempo en diferentes filtros de paso total 28<sub>1</sub> a 28<sub>16</sub>, se pueden conectar en serie un número diferente de estructuras de filtro de paso total 14.

45

En otras palabras, la figura 4 muestra una etapa de filtro de paso total.

50

La figura 5 muestra un diagrama de bloques esquemático de una estructura de filtro de paso total 50 que es una estructura anidada de filtro de paso total. Alternativamente o adicionalmente a una estructura de filtro de paso total 40, una o más estructuras de filtro de paso total 50 pueden formar al menos una parte de un filtro de paso total 28<sub>1</sub> a 28<sub>16</sub> del decorrelacionador 20 y/o 30. Aunque se muestran todos bloques de retardo 52, y 52<sub>2</sub>, pueden estar presentes un número diferente y especialmente mayor de bloques de retardo 52 dando como resultado posiblemente en un mayor número de ramificaciones hacia adelante 46 y/o ramificaciones hacia atrás 48. Además, se pueden adoptar las ganancias  $g_1$ - $g_1$  y/o  $g_2$ - $g_2$ .

55

60

Cuando se considera, por ejemplo, conectar en serie los bloques de retardo 52 en una o más estructuras de filtro de paso total 40 y/o una o más estructuras de filtro de paso total 50, se pueden implementar diferentes filtros de paso total 28<sub>1</sub> a 28<sub>16</sub> para que comprendan un retardo de tiempo diferente cuando se compara con otros filtros de paso total. Por ejemplo, los diferentes retardos de diferentes estructuras de filtro de paso total y/o circuiterías de estructuras de filtro de paso total se pueden basar en un múltiplo de número primo de una frecuencia de muestreo total, por ejemplo, 48 kHz, utilizada para obtener la representación de frecuencia 14 de la señal de audio 22. Por ejemplo, un conjunto de estructuras de filtro de paso total que forman al menos una parte de un filtro de paso total puede comprender un número de cuatro estructuras de filtro de paso total, por ejemplo, estructuras de filtro de paso total 40. Los diferentes

bloques de retardo en las mismas se pueden adaptar para proporcionar un retardo de 1, 2, 3 y 5. De acuerdo con un ejemplo diferente, el número de cuatro estructuras de filtro de paso total pueden proporcionar un retardo de 1, 3, 5 y 7 unidades en el dominio Z. Estos valores pueden formar un conjunto de valores primos, es decir, se puede agrupar un número de 2, 3, 4, 5 o más valores primos.

5

Cuando se transfiere esta realización, los conjuntos de valores primos, respectivamente, a las posibles operaciones de los filtros de paso total en el dominio del tiempo, los retardos de tiempo se basan en un múltiplo de número primo de un recíproco de una frecuencia de muestreo utilizada para obtener la representación de frecuencia de la señal de audio en una realización. Por ejemplo, los diferentes retardos de tiempo se pueden basar en un número primo que se obtiene multiplicando cada uno de un conjunto de números primos mencionados, por ejemplo, 1, 2, 3 y 5 o 1, 3, 5 y 7 con un factor de reducción de resolución utilizado para generar las partes de la representación de frecuencia de la señal de audio para obtener un resultado intermedio. En lugar del resultado intermedio, se puede utilizar un siguiente número primo con respecto al resultado intermedio. Por ejemplo, cuando se hace referencia al factor de muestreo descendente de 128 y considerando los conjuntos de números primos anteriores, tal resultado puede ser el retardo de 131, 257, 383 y 641 por un lado y 131, 383, 641 y 907 por el otro, en donde cada retardo puede referirse a una multiplicación con 1 muestra en la frecuencia de muestreo que es, para una frecuencia de muestreo de 48 kHz aproximadamente de 20,8  $\mu$ s. Otros conjuntos de números primos son posibles sin limitación.

10

15

Cuando se hace referencia, por ejemplo, a la figura 4, el factor de ganancia g del filtro de paso total se puede adaptar a un valor de 0,7 dentro de un margen de tolerancia de, por ejemplo,  $\pm 20\%$ ,  $\pm 10\%$  o  $\pm 5\%$ . Sin embargo, el valor de ganancia también puede tener un valor negativo de, por ejemplo, -0,7 dentro del margen de tolerancia mencionado. Es decir, el factor de ganancia se puede adaptar a un valor dentro de una magnitud de 0,7 dentro del margen de tolerancia.

20

25

En otras palabras, adicionalmente a la configuración de paso en serie de la figura 4, también una configuración anidada en la cual el elemento de retardo de un filtro de paso total de Schroeder se reemplaza por otra configuración de paso total interior o se puede implementar una combinación de ambas configuraciones. La figura 5 muestra una etapa de filtro de paso total anidada simple.

30

La figura 6 muestra un diagrama de bloques esquemático de un decorrelacionador 60 de acuerdo con una realización. El decorrelacionador 60 comprende el cambiador de fase 26 configurado para operar en el dominio del tiempo. Una estructura de filtro de paso total 28' puede estar configurada para utilizar los siguientes números primos respectivos cuando se comparan con los conjuntos de números primos descritos en relación con el decorrelacionador 20 y/o 30. Para asegurar una operación precisa del decorrelacionador 60, el mismo puede comprender las unidades de conversión 24<sub>1</sub> y 24<sub>2</sub>. Mientras que la unidad de conversión 24 puede proporcionar lo necesario para la representación de frecuencia de la señal de audio, la unidad de conversión 24<sub>2</sub> puede recibir la señal de audio reverberada o con cambio de fase 22' proporcionada por el cambiador de fase 28'. Las partes 14''<sub>1</sub> a 14''<sub>16</sub> obtenidas se pueden retardar por medio de las unidades de retardo 12<sub>1</sub> a 12<sub>16</sub> llegando a una entrada comparable para el conformador de envolvente 16 cuando se compara con el decorrelacionador 20 y/o 30 mientras se permite una reverberación basada en el dominio del tiempo. Es decir, las partes de la representación de frecuencia pueden formar partes de la representación de frecuencia de la señal de audio reverberada 22'.

35

40

De acuerdo con las realizaciones, un decorrelacionador como se describe en el presente documento se puede combinar con funcionalidad adicional, es decir, la señal de salida se puede procesar adicionalmente.

45

En otras palabras, la figura 6 muestra una implementación alternativa de un decorrelacionador con respecto a la figura 2.

50

Además, los decorrelacionadores inventivos se pueden combinar con procesamiento de manejo de transitorios. Los transitorios pueden provocar artefactos en la señal estereofónica decorrelacionada tales como pos-ecos o efectos de panoramización no deseados. Para mitigar esto, un manejo de transitorios se puede combinar con el decorrelacionador descrito en el presente documento. El manejo de transitorios puede silenciar la salida del decorrelacionador para conservar la forma de onda inicial directa y suprimir el pos-eco provocado por el pre-retardo.

55

La figura 7 muestra un diagrama de bloques esquemático de un decorrelacionador 70 de acuerdo con una realización. El decorrelacionador 70 comprende al menos una parte del decorrelacionador 10, en donde alternativamente o adicionalmente se pueden disponer al menos partes del decorrelacionador 20, 30 y/o 60. El decorrelacionador 70 puede comprender una etapa de procesamiento de señal 56 configurada para procesar la representación de frecuencia conformada combinada 18 o una señal basada en la misma. La representación de frecuencia conformada combinada 18 se puede considerar como una señal monoaural, es decir, puede representar un solo canal. A partir de la señal monoaural recibida, la etapa de procesamiento puede proporcionar al menos las señales 58 y 58<sub>2</sub> que representan una señal estereofónica.

60

Un extensor de la fuente 58 que modela el efecto perceptual de una fuente de sonido extendida espacialmente de una

señal monoaural de una fuente puntual y una versión decorrelacionada de la misma se puede acoplar al decorrelacionador 70. El extensor de la fuente 58 puede comprender los filtros 64<sub>1</sub> a 64<sub>2</sub> que permiten un modelado de extensión de la fuente con base en la señal estereofónica que tiene las señales 58, y 58<sub>2</sub>. El modelado de extensión de la fuente se puede llevar a cabo, por ejemplo, en el dominio de la frecuencia y puede resultar en las señales de salida estereofónica 64<sub>1</sub>, por ejemplo, un canal izquierdo y 64<sub>2</sub>, por ejemplo, un canal derecho. Se debe tener en cuenta que el extensor de la fuente 58 también puede formar parte del decorrelacionador 70.

En otras palabras, la figura 7 muestra un diagrama de bloques esquemático del procesamiento de extensión de la fuente.

La figura 8 muestra un diagrama de bloques esquemático de un sistema de procesamiento 80 de acuerdo con una realización. El sistema de procesamiento 80 puede comprender el decorrelacionador 10. Alternativamente o adicionalmente, se puede disponer el decorrelacionador 20, 30, 60 y/o 70. El sistema de procesamiento 80 comprende una etapa de procesamiento 66 configurada para transformar una señal descompuesta media/lateral 68 a una señal descompuesta izquierda/derecha 72. Es decir, la señal descompuesta media/lateral 68 puede comprender al menos una primera señal 74<sub>1</sub>, por ejemplo, que representa una de la porción media/intermedia o lateral de una segunda señal 74<sub>2</sub> que representa la otra porción. La etapa de procesamiento 66 puede estar configurada para transformar las señales 74<sub>1</sub> a 74<sub>2</sub> y posiblemente señales adicionales en al menos las señales 76<sub>1</sub> a 76<sub>2</sub> que representan un canal izquierdo y un canal derecho. Un canal, por ejemplo, el canal izquierdo L (Left), se puede obtener, por ejemplo, agregando el componente medio M y el componente lateral M+S (Mid+Side); mientras el otro, por ejemplo, el canal derecho se puede obtener sustrayendo un componente del otro, por ejemplo, M-S. De acuerdo con un planteamiento diferente, ambos canales se pueden obtener utilizando el 50% o un factor de 0,5 de los mismos, es decir, 0,5(M+S) y 0,5(M-S). Otros factores y/o reglas de determinación son posibles.

De acuerdo con una realización, la señal 74 es proporcionada por el decorrelacionador del sistema de procesamiento 80. La otra señal 74<sub>2</sub> puede ser proporcionada por una unidad de compensación de retardo 78 que se conecta en paralelo al decorrelacionador 10 y está configurada para también recibir la señal de audio 22. Por lo tanto, la unidad de compensación de retardo 78 se conecta con la etapa de procesamiento 66. La unidad de compensación de retardo 78 puede estar configurada para proporcionar un retardo de tiempo que sea comparable con el decorrelacionador. Preferiblemente, para realizaciones en el dominio de la frecuencia, el retardo es igual al retardo de procesamiento introducido por el análisis/síntesis de STFT del decorrelacionador. Sin embargo, el decorrelacionador 10 puede proporcionar lo necesario para el procesamiento de señal adicional que conduce a una decorrelación tal que la señal 74<sub>2</sub> pueda comprender un retardo similar cuando se compara con la señal 74<sub>1</sub>. De acuerdo con una realización, la señal 74<sub>2</sub> puede estar sin procesar con excepción del retardo de tiempo.

El decorrelacionador 10 en el sistema de procesamiento 80 puede proporcionar la representación de frecuencia conformada combinada como al menos una parte de la señal descompuesta media/lateral a la etapa de procesamiento 66. La etapa de procesamiento 66 puede transformar la representación de frecuencia conformada combinada junto con la señal de retardo 74<sub>2</sub> en la señal descompuesta izquierda/derecha en el dominio de la frecuencia. La salida de la etapa de procesamiento 66 puede ser una señal L/R 72. El decorrelacionador 10 en sí puede producir una señal monoaural S (lateral, componente 18), en relación con que solo sea parte de la misma. Con el manejo de transitorios, la parte directa M (74<sub>2</sub>; 74'<sub>2</sub>) y la salida del decorrelacionador S (señal 18) se puede acoplar estrechamente, ya que la señal S se silenciará y se "reemplazará" por una señal M amplificada (señal 74'<sub>2</sub>). Como consecuencia, ambas unidades, el decorrelacionador y la "unidad de separación de canales" 66 se acoplan estrechamente y así la etapa de procesamiento 66 finalmente proporciona la señal estereofónica decorrelacionada. Si el decorrelacionador se operará independiente con salida monoaural, por ejemplo, sin la etapa de procesamiento 66, entonces la señal directa de retardo compensado, sin ninguna escala, se agregaría directamente a la salida monoaural para llenar el espacio silenciado y proporcionar una señal "completa".

En otras palabras, la figura 8 muestra un decorrelacionador en configuración M/S a L/R, con compensación de retardo de entrada monoaural (señal media).

La figura 9 muestra un diagrama de bloques esquemático de un sistema de procesamiento 90 de acuerdo con una realización. Cuando se compara con el sistema de procesamiento 80, el sistema de procesamiento 90 comprende un supresor de transitorios 82 configurado para detectar un transitorio en la señal de audio 22 o la representación de frecuencia 14 de la misma en una entrada del decorrelacionador. El supresor de transitorios puede comprender una unidad de detección de transitorios 84 configurada para recibir la señal de audio 22 o la representación de frecuencia de la misma. La unidad de detección de transitorios 84 puede detectar un transitorio en la señal de audio, por ejemplo, procesando la señal de audio 22. El supresor de transitorios 82 puede comprender además una unidad de silenciamiento 86 configurada para recibir la representación de frecuencia conformada combinada 18 y silenciarla con base en una señal de control. Sin embargo, se debe tener en cuenta que también se puede obtener un efecto igual o comparable cuando se controla el decorrelacionador 10 o el decorrelacionador contenido en el sistema de procesamiento 90 para silenciar la salida del decorrelacionador. Es decir, la unidad de silenciamiento 86 también puede formar parte del decorrelacionador. Sin embargo, la señal 74 que forma la entrada de la etapa de procesamiento

66 se puede silenciar con base en un transitorio detectado en la señal de audio 22. El supresor de transitorios 82 puede estar configurado para silenciar temporalmente la porción proporcionada por el decorrelacionador para suprimir ecos en la etapa de procesamiento 66, en donde los ecos pueden estar relacionados con pre-ecos y/o pos-ecos. Cuando se opera en el dominio del tiempo, se puede utilizar una ventana para un silenciamiento suave para evitar que se provoquen transitorios adicionales por el silenciamiento. Si se hace en el dominio de la frecuencia, la ventana de STFT, que se describe en relación con los decorrelacionadores 20, 30 y 60, puede proporcionar lo necesario para tal efecto automáticamente, es decir, de una manera sinérgica.

Con respecto a la etapa de procesamiento 66, silenciar la salida del decorrelacionador 10 podría conducir a un cambio no deseado en la energía de entrada de la etapa de procesamiento de señal 66. Para evitar los efectos negativos, un amplificador 82 se puede conectar entre la unidad de compensación de retardo 78 y la etapa de procesamiento de señal 66 para amplificar temporalmente la señal  $74_2$  para obtener la señal amplificada  $74'_2$ . La amplificación de la señal  $74_2$  puede ser condicional a silenciar la salida del decorrelacionador 10. Es decir, el supresor de transitorios 82 puede estar configurado para amplificar la porción de la unidad de compensación de retardo 78 que corresponde a silenciar la porción del decorrelacionador.

Un nivel de amplificación puede ser fijo o puede ser controlado. De acuerdo con un ejemplo, se aplica, el factor de amplificación del amplificador 82 puede ser un factor de  $\sqrt[2]{2}$  cuando se compara con una porción no silenciada del decorrelacionador. Es decir, cuando se silencia la salida del decorrelacionador, el amplificador 82 puede amplificar la señal  $74_2$  por  $\sqrt[2]{2}$  mientras que no se amplifica la señal  $74_2$  durante tiempos donde el silenciamiento está apagado, es decir,  $g = 1$ .

Opcionalmente, y para evitar efectos no deseados durante la supresión de transitorios, el supresor de transitorios 82 puede estar configurado para suprimir un transitorio detectado en la señal de audio y para suprimir un transitorio siguiente no antes de un tiempo de inhibición predefinido. Por ejemplo, el supresor de transitorios 82 puede comprender una unidad de control 92 configurada para controlar y/o aplicar un tiempo de espera, una histéresis y/o un tiempo de inhibición. Por ejemplo, el tiempo de espera puede ser más corto cuando se compara con el tiempo de inhibición. El tiempo de espera puede referirse a un tiempo durante el cual la salida del decorrelacionador 10 se silencia en respuesta a un transitorio detectado, es decir, una propiedad determinada por la unidad de detección de transitorios 84. El tiempo de inhibición puede ser más largo cuando se compara con el tiempo de espera, para evitar efectos no deseados. Por ejemplo, el contador de espera, es decir, el tiempo para silenciar, puede ser de 1, 2, 4, 6, 7 u 8 bloques, mientras que el tiempo de inhibición puede ser de al menos dos veces el tiempo, por ejemplo, al menos 14, al menos 20, al menos 30 o 56 bloques o cualquier otra duración de tiempo.

De acuerdo con un ejemplo, la unidad de control 92 también puede proporcionar lo necesario para una histéresis para mitigar el encendido/apagado de supresión de transitorios para señales de audio como trenes de pulsos de baja frecuencia. Es decir, el tiempo de inhibición proporcionado por la unidad de control 92 puede ser un primer tiempo de inhibición. El supresor de transitorios 82 puede estar configurado para reiniciar el tiempo de inhibición como un segundo tiempo de inhibición que es más largo que el primer tiempo de inhibición en caso de que ocurra un transitorio durante el primer tiempo de inhibición. Es decir, incluso si el tiempo de espera ha transcurrido, pero el tiempo de inhibición no ha transcurrido todavía y en caso de que se determine un nuevo transitorio (independientemente de si ha transcurrido el tiempo de espera o no), se puede reiniciar el temporizador de inhibición. Opcionalmente, el temporizador de inhibición reiniciado puede ser más largo cuando se compara con el temporizador de inhibición cancelado. En otras palabras, cuando se detecta un transitorio por primera vez, entonces se inician un contador de espera y un contador de inhibición. El transitorio se puede silenciar hasta que el contador de espera haya alcanzado su conteo de detención, por ejemplo, 8 bloques. Después, el contador de espera se puede reiniciar y el silenciamiento se puede detener. El contador de inhibición puede alcanzar su conteo de detención/reinicio mucho más tarde en el tiempo, por ejemplo, 56 bloques. Si durante dicho proceso de conteo de inhibición en curso se detecta un nuevo transitorio, entonces se reinicia el contador de inhibición, pero con un valor de conteo de detención más alto, por ejemplo, 64 bloques. De esta forma, se implementa histéresis por medio de conmutación condicional y modificaciones en el conteo de detención. Es decir, durante la ejecución del contador de inhibición, se puede desactivar una nueva activación de supresión de transitorios o silenciamiento.

El supresor de transitorios 82 puede estar configurado para operar en el dominio de la frecuencia. Alternativamente o adicionalmente, el supresor de transitorios 82 puede estar configurado para silenciar la porción del decorrelacionador durante un tiempo más largo cuando se compara con un pre-retardo del decorrelacionador. Es decir, en caso de que se detecte un transitorio en la señal de audio 22, entonces el silenciamiento debería seguir en efecto cuando el transitorio llegue a la salida del decorrelacionador.

En otras palabras, los decorrelacionadores de acuerdo con las realizaciones operan en el dominio de la transformada de Fourier de tiempo corto (STFT) en bloques de transformada superpuestos con una duración corta. Esto permite un retardo de procesamiento pequeño de unos cuantos milisegundos, por ejemplo, 2,7 milisegundos suponiendo un

- tamaño de transformada de 256 y frecuencia de muestreo de 48 kHz, contrario al alto retardo del decorrelacionador PS/MDS como se describe en [2] o [3] que puede llegar a un tiempo de retardo de 13,3 milisegundos en una frecuencia de muestreo de 48 kHz. Por otra parte, los decorrelacionadores descritos se pueden implementar utilizando filtros de paso total de un cómputo muy bajo y por lo tanto pueden ser computacionalmente mucho más eficientes que la decorrelación en el dominio del tiempo como se describe en [1] o [2]. Si se requiere o desea procesamiento espectral más proceso abajo, por ejemplo, un modelado de extensión de la fuente, los decorrelacionadores descritos se pueden hacer interactuar directamente con esta etapa de procesamiento en el dominio de STFT para lograr una baja complejidad computacional.
- Los decorrelacionadores como se describen en el presente documento pueden por lo tanto proporcionar lo necesario para un retardo de procesamiento corto y una complejidad computacional moderada. Los decorrelacionadores se pueden combinar con procesamiento adicional proceso bajo, para modelar objetos de audio que tengan una dimensión espacial, las llamadas fuentes de sonido extendido espacialmente (SESS, Spatially Extended Sound Sources) con una propiedad perceptual de "extensión de la fuente".
- En otras palabras, la figura 2 y la figura 9 muestran realizaciones preferidas de la presente invención. En primer lugar, la señal de entrada o señal de audio (sonido de una fuente puntual, por ejemplo) se puede alimentar al decorrelacionador 20 que comprende una DFT por bloques de tiempo con, por ejemplo, longitud de bloque de 256 muestras y, por ejemplo, una superposición del 50%. Después, los bins espectrales de la DFT se retardan en el tiempo para una duración dependiente de la frecuencia, donde las frecuencias bajas pueden tener un mayor retardo y las frecuencias altas pueden tener un menor retardo. Por ejemplo, el retardo puede ser de 16 muestras de sub-banda (42,7 milisegundos a 48 kHz) para frecuencias bajas y puede disminuir hasta 1 muestra de sub-banda para los bins más altos, es decir,  $z^{-1}$ . La disminución en el retardo a través del tiempo puede ser lineal, logarítmica o de otra manera con redondeo a números enteros de muestras de sub-banda. Después, cada bin se envía a través de un filtro de paso total, que comprende preferiblemente una cadena simple de filtros de paso total o una estructura anidada de filtros de paso total. Un filtro de paso total de ejemplo se muestra en la figura 4. Una estructura diferente se muestra en la figura 5. Con respecto a la figura 4, una cadena posible puede comprender o constar de cuatro de dichos filtros de paso total. El parámetro  $g$  se puede elegir para que sea, por ejemplo, de 0,7 y los retardos  $M_i$  pueden ser números primos. Obsérvese que la figura 4 muestra la primera parte de la cadena, es decir,  $M_1$ . Como estos filtros pueden operar en bandas espectrales de muestreo descendente, por ejemplo, el factor de muestreo descendente 128, los retardos pueden ser muy bajos, por ejemplo, números primos 1, 2, 3 y 5 o, como otro ejemplo, 1, 3, 5 y 7. A continuación, se puede aplicar una conformación envolvente de tiempo/frecuencia. Las señales de entrada a la conformación de envolvente pueden ser los bins de DFT directamente y sus versiones retardadas y filtradas. Finalmente, una IDFT con superposición añadida puede sintetizar la señal de salida. La señal de salida se puede procesar adicionalmente en el dominio del tiempo para obtener una señal estereofónica izquierda/derecha a partir de una señal de entrada monoaural en una configuración como se muestra en la figura 8. Alternativamente, la señal estereofónica izquierda/derecha se puede ensamblar en el dominio de la frecuencia de DFT y procesar adicionalmente en el dominio de la frecuencia, por ejemplo, para una extensión de la fuente/modelado de SESS por medio de convolución rápida, si es benéfico para la eficiencia computacional general.
- Una configuración para modelado de extensión de la fuente se muestra en la figura 7. En contraste con otras realizaciones, la realización alternativa que tiene retardos  $M_i$  se puede elegir como números primos que son aproximadamente 128 veces (que corresponde con el factor de muestreo descendente mencionado anteriormente) mayores que los escogidos en el dominio de la sub-banda, por ejemplo, 131, 257, 383 y 641 (para el conjunto de valores primos 1, 2, 3 y 5) o 131, 383, 641 y 907 (para el conjunto de valores primos 1, 3, 5 y 7). Para diferentes conjuntos de valores primos con un número diferente de números primos y/o diferentes números primos, se pueden escoger valores correspondientes. Además, la realización alternativa puede requerir una STFT para obtener la entrada de señal directa al conformador de envolvente de tiempo/frecuencia.
- La figura 9 muestra un decorrelacionador de ejemplo en configuración M/S a L/R con procesamiento de manejo de transitorios. Aspectos de estas realizaciones son:
- una detección de transitorios detecta la presencia de un transitorio aislado
  - si se detecta un transitorio, el sonido decorrelacionado se silencia durante un "tiempo de espera" y la señal directa de retardo compensado se amplifica en consecuencia. para compensar el efecto de adición coherente, se aplica un factor de  $2\sqrt{2}$  para amplificar la señal directa donde esta reemplaza la señal decorrelacionada
  - para evitar activar los trenes de pulsos rápidos, que se perciben como tonos, una inhibición previene la activación por el siguiente transitorio durante un cierto "tiempo de inhibición"; el tiempo de inhibición se reinicia por cada nueva detección de transitorio durante el "tiempo de espera"
  - una histéresis evita alternar la detección de transitorios (por ejemplo, aumentando el "tiempo de inhibición" en caso de una inhibición reactivada)

## ES 3 008 258 T3

la detección de transitorios, el silenciamiento, la amplificación directa de sonido, la inhibición de detección y la histéresis se pueden implementar convenientemente en el dominio de la STFT:

- 5 la superposición del bloque de STFT proporciona una transición gradual suave  
el tiempo de silenciamiento es mayor que el pre-retardo del decorrelacionador  
10 el contador de bloques de silenciamiento para silenciar la señal decorrelacionada y amplificar la señal directa  
inhibir el contador de bloques para inhibir la detección de transitorios  
la histéresis para evitar alternar en la detección de transitorios
- 15 Las realizaciones de la presente invención se refieren a  
un aparato/método para decorrelación de una señal de audio  
un decorrelacionador, que incluye  
20 un par de DFT/IDFT (opcional, si interactúan directamente con el procesamiento de SESS en el dominio de la frecuencia)  
retardos en el dominio de la sub-banda; preferiblemente las frecuencias bajas tienen un mayor retardo y las frecuencias  
25 altas tienen un menor retardo; distribución de retardo a lo largo de la frecuencia: lineal, logarítmica, etc.  
filtros de paso total en el dominio de la sub-banda; opcionalmente: las frecuencias bajas pueden tener un retardo/orden superior y las frecuencias altas tienen un retardo/orden menor; los filtros de paso total de orden superior se pueden  
30 realizar por medio de una etapa de filtros de paso total de orden inferior  
filtros IIR de Schroeder cortos en el dominio de la sub-banda de DFT (de reducción de resolución) que utilizan números primos de retardo enteros pequeños en combinación con retardos de frecuencia variante  
un ajustador de envolvente de T/F con alta resolución de tiempo (<4 ms) que trabajar en el dominio de la sub-banda;  
35 medir la energía antes y después del procesamiento de retardo/paso total; ajustar la energía de la señal de sub-banda para (tanto como sea posible) coincidir con la energía de la señal de sub-banda original  
un decorrelacionador de bajo retardo como parte del modelado/procesamiento de "extensión de la fuente" (contrario al decorrelacionador envolvente de MPEG)  
40 una interfaz para el procesamiento de extensión de la fuente de proceso bajo a través del tiempo o dominio de la frecuencia de DFT para eficiencia computacional.  
Implementación alternativa: Filtros de paso total antes de los retardos ("pos-retardos")  
45 La figura 10 muestra un diagrama de bloques esquemático de un método 1000 de acuerdo con una realización que se puede implementar, por ejemplo, por un decorrelacionador descrito en el presente documento. El método 1000 comprende una etapa 1010 en la cual se recibe una pluralidad de partes que se basan en una señal de audio. En 1020, cada una de las partes recibidas se retarda para proporcionar una pluralidad de partes retardadas. 1030  
50 comprende recibir y combinar señales que se basan en las partes retardadas de la representación de frecuencia. 1040 comprende recibir la representación de frecuencia de la señal de audio. 1050 comprende ajustar una energía de las partes retardadas en relación con la representación de frecuencia de la señal de audio. 1060 comprende proporcionar una representación de frecuencia conformada combinada, por ejemplo, utilizando el conformador de envolvente 16.  
55 Aunque algunos aspectos se han descrito en el contexto de un aparato, es claro que estos aspectos también representan una descripción del método correspondiente, en donde un bloque o dispositivo corresponde a una etapa de un método o una característica de una etapa de un método. Análogamente, los aspectos descritos en el contexto de una etapa de un método también representan una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente.  
60 La señal de audio codificada inventiva se puede almacenar en un medio de almacenamiento digital o se puede transmitir en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión alámbrico tal como Internet.

5 Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención se pueden implementar en hardware o en software. La implementación se puede llevar a cabo utilizando un medio de almacenamiento digital, por ejemplo, un disquete, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria flash, que tenga señales de control legibles electrónicamente almacenadas en el mismo, que coopere (o sea capaz de cooperar) con un sistema informático programable de tal manera que se lleve a cabo el método respectivo.

10 Algunas realizaciones de acuerdo con la invención comprenden un portador de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema informático programable de tal manera que se lleve a cabo uno de los métodos descritos en el presente documento.

15 Generalmente, las realizaciones de la presente invención se pueden implementar como un producto de programa informático con un código de programa, el código de programa siendo operativo para llevar a cabo uno de los métodos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa se puede almacenar, por ejemplo, en un portador legible por máquina.

Otras realizaciones comprenden el programa informático para llevar a cabo uno de los métodos descritos en el presente documento, almacenado en un portador legible por máquina.

20 En otras palabras, una realización del método inventivo es, por lo tanto, un programa informático que tiene un código de programa para llevar a cabo uno de los métodos descritos en el presente documento, cuando el programa informático se ejecuta en un ordenador.

25 Una realización adicional de los métodos inventivos es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para llevar a cabo uno de los métodos descritos en el presente documento.

30 Una realización adicional del método inventivo es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa informático para llevar a cabo uno de los métodos descritos en el presente documento. El flujo de datos o la secuencia de señales se puede configurar, por ejemplo, para transferirse por medio de una conexión de comunicación de datos, por ejemplo, a través de Internet.

Una realización adicional comprende un medio de procesamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los métodos descritos en el presente documento.

35 Una realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para llevar a cabo uno de los métodos descritos en el presente documento.

40 En algunas realizaciones, un dispositivo lógico programable (por ejemplo, una matriz de puertas programable en campo) se puede utilizar para llevar a cabo algunas o todas las funcionalidades de los métodos descritos en el presente documento. En algunas realizaciones, una matriz de puertas programable en campo puede cooperar con un microprocesador con el fin de llevar a cabo uno o más de los métodos descritos en el presente documento. Generalmente, los métodos se llevan a cabo preferiblemente por medio de cualquier aparato de hardware.

#### 45 Referencias

[1] W. Oomen, E. Schuijers, B. den Brinker y J. Breebaart, "Advances in Parametric Coding for High-Quality Audio", Paper 5852, (marzo de 2003).

50 [2] J. Breebaart, S. van de Par, A. Kohlrausch y E. Schuijers, "High-quality Parametric Spatial Audio Coding at Low Bitrates", Paper 6072, (mayo de 2004). QMF domain PS:

[3] H. Purnhagen, J. Engdegard, J. Roden y L. Liljeryd, "Synthetic Ambience in Parametric Stereo Coding", Paper 6074, (mayo de 2004).

55 [4] J. Herre, K. Kjörling, J. Breebaart, C. Faller, S. Disch, H. Purnhagen, J. Koppens, J. Hilpert, J. Rödén, W. Oomen, K. Linzmeier y KO. SE. Chong, "MPEG Surround-The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding", J. Audio Eng. Soc., vol. 56, no. 11, págs. 932-955, (noviembre de 2008).

60 [5] Shi Chuang ET AL, "An overview of directivity control methods of the parametric array loudspeaker", APSIPA Transactions on Signal and Information Processing, vol. 3, no. 1, 1 de enero de 2014.

**REIVINDICACIONES**

1. Decorrelacionador, que comprende:
  - 5 una pluralidad de unidades de retardo (12), en donde cada unidad de retardo (12) está configurada para recibir una parte (14<sub>1</sub>-14<sub>n</sub>) de una representación de frecuencia que se basa en una señal de audio (22); en donde cada unidad de retardo (12) está configurada para retardar la parte (14<sub>1</sub>-14<sub>n</sub>) recibida para proporcionar una parte retardada (14'<sub>1</sub>-14'<sub>n</sub>); y
  - 10 un conformador de envolvente (16) configurado para recibir y combinar señales que se basan en las partes retardadas (14'<sub>1</sub>-14'<sub>n</sub>) de la representación de frecuencia; para recibir la representación de frecuencia de la señal de audio (22); para ajustar una energía de las partes retardadas (14'<sub>1</sub>-14'<sub>n</sub>) en relación con la representación de frecuencia de la señal de audio (22); y para proporcionar una representación de frecuencia conformada combinada.
- 15 2. Decorrelacionador de acuerdo con la reivindicación 1, que comprende además un cambiador de fase (26) configurado para cambiar de fase la representación de frecuencia (14) de la señal de audio (22); y comprende una pluralidad de filtros de paso total, en donde cada filtro de paso total (28) está configurado para cambiar de fase de una parte asociada (14<sub>1</sub>-14<sub>n</sub>) de la representación de frecuencia de la señal de audio (22).
- 20 3. Decorrelacionador de acuerdo con la reivindicación 1, que comprende además un cambiador de fase (26) configurado para cambiar de fase la señal de audio (22) en un dominio del tiempo para obtener una señal de audio cambiada de fase (22); en donde el cambiador de fase (26) comprende un conjunto de estructuras de filtro de paso total (40; 50) tales como filtros IIR de Schroeder, que se conectan en serie entre sí; en donde las estructuras de filtro de paso total (40; 50) están adaptadas para proporcionar diferentes retardos de tiempo; o
- 25 en donde las estructuras de filtro de paso total (40; 50) comprenden una estructura anidada de filtros de paso total.
- 30 4. Decorrelacionador de acuerdo con una de las reivindicaciones anteriores, que está adaptado para obtener un número de 16 partes (14<sub>1</sub>-14<sub>n</sub>) de la representación de frecuencia.
- 35 5. Decorrelacionador de acuerdo con una de las reivindicaciones anteriores, que está adaptado para obtener la representación de frecuencia con un número de 128 o 129 bins de frecuencia.
- 40 6. Decorrelacionador de una de las reivindicaciones anteriores, en donde el decorrelacionador está adaptado para implementar adicionalmente un retardo igual y predefinido para un subconjunto o todas las partes (14<sub>1</sub>-14<sub>n</sub>) de la representación de frecuencia.
- 45 7. Decorrelacionador de acuerdo con una de las reivindicaciones anteriores, que comprende una unidad de conversión (24) para recibir y convertir la señal de audio (22) o una versión reverberada de la señal de audio (22) en las partes (14<sub>1</sub>-14<sub>n</sub>) al llevar a cabo una transformada discreta de Fourier, DFT, por bloques de tiempo, o una transformada de Fourier de tiempo corto, STFT; en donde la unidad de conversión (24) está configurada para convertir bloques que tienen una superposición del 50 % dentro de un margen de tolerancia.
- 50 8. Decorrelacionador de acuerdo con una de las reivindicaciones anteriores, que comprende una unidad de conversión (24) para recibir y convertir la señal de audio (22) o una versión reverberada de la señal de audio (22) en las partes (14<sub>1</sub>-14<sub>n</sub>) al llevar a cabo una transformada discreta de Fourier, DFT, por bloques de tiempo, o una transformada de Fourier de tiempo corto, STFT; en donde los bloques comprenden una longitud de bloque de 256 muestras.
- 55 9. Decorrelacionador de acuerdo con una de las reivindicaciones anteriores, que comprende una unidad de conversión inversa (34) para recibir versiones procesadas de las partes de la representación de frecuencia (14) y para sintetizar una señal sintetizada a partir de las versiones procesadas con base en un procedimiento de superposición añadida.
- 60 10. Decorrelacionador de acuerdo con una de las reivindicaciones anteriores, en donde el conformador de envolvente (16) está configurado para operar en un dominio de sub-banda y con una resolución temporal de menos de 4 ms.
11. Decorrelacionador de acuerdo con una de las reivindicaciones anteriores, en donde el conformador de envolvente (16) es para conformar bins espectrales en tiempo y/o en frecuencia individualmente o como un grupo, por ejemplo, implementando un procesamiento de conformación común interdependiente o al menos

por grupos.

12. Decorrelacionador de acuerdo con una de las reivindicaciones anteriores, que comprende una etapa de procesamiento de señal (66) configurada para recibir una señal con base en la representación de frecuencia conformada combinada como una señal monoaural y para procesar la señal monoaural al menos como una señal estereofónica.
13. Sistema de procesamiento, que comprende:
- un decorrelacionador de acuerdo con una de las reivindicaciones anteriores; y
- una etapa de procesamiento (66) para transformar una señal descompuesta media/lateral a una señal descompuesta izquierda/derecha.
14. Sistema de procesamiento de acuerdo con la reivindicación 13, en donde una porción (74<sub>1</sub>) de la señal descompuesta media/lateral es proporcionada por el decorrelacionador y la otra porción (74<sub>2</sub>) es proporcionada por una unidad de compensación de retardo (78) que se conecta en paralelo con el decorrelacionador y se conecta con la etapa de procesamiento (66);
- el sistema de procesamiento que comprende un supresor de transitorios (82) configurado para detectar un transitorio en la señal de audio (22) o la representación de frecuencia (14) de la misma en una entrada del decorrelacionador;
- en donde el supresor de transitorios (82) está configurado para silenciar temporalmente la porción (74<sub>1</sub>) proporcionada por el decorrelacionador para suprimir ecos en la etapa de procesamiento.
15. Sistema de procesamiento de acuerdo con la reivindicación 14, en donde el supresor de transitorios (82) está configurado para amplificar la porción de la unidad de compensación de retardo que corresponde a silenciar la porción del decorrelacionador.
16. Sistema de procesamiento de acuerdo con la reivindicación 15, en donde el supresor de transitorios (82) está configurado para amplificar la porción de la unidad de compensación de retardo por un factor de  $\sqrt{2}$  cuando se compara con una porción no silenciada del decorrelacionador.
17. Sistema de procesamiento de acuerdo con una de las reivindicaciones 14 a 16, en donde el supresor de transitorios (82) está configurado para suprimir un transitorio detectado y para suprimir un transitorio siguiente no antes de un tiempo de inhibición predefinido.
18. Sistema de procesamiento de acuerdo con una de las reivindicaciones 14 a 17, en donde el tiempo de inhibición es un primer tiempo de inhibición; en donde el supresor de transitorios (82) está configurado para reiniciar el tiempo de inhibición como un segundo tiempo de inhibición que es más largo que el primer tiempo de inhibición en caso de que ocurra un transitorio durante el primer tiempo de inhibición.
19. Sistema de procesamiento de acuerdo con una de las reivindicaciones 14 a 18, en donde el supresor de transitorios (82) está configurado para operar en el dominio de la frecuencia.
20. Sistema de procesamiento de acuerdo con una de las reivindicaciones 14 a 19, en donde el supresor de transitorios (82) está configurado para silenciar la porción del decorrelacionador durante un tiempo más largo cuando se compara con un pre-retardo del decorrelacionador.
21. Método que comprende:
- recibir (1010) una pluralidad de partes de una representación de frecuencia que se basa en una señal de audio;
- retardar (1020) cada una de las partes recibidas para proporcionar una pluralidad de partes retardadas; y
- recibir (1030) y combinar señales que se basan en las partes retardadas de la representación de frecuencia;
- recibir (1040) la representación de frecuencia de la señal de audio;
- ajustar (1050) una energía de las partes retardadas en relación con la representación de frecuencia de la señal de audio; y

proporcionar (1060) una representación de frecuencia conformada combinada.

- 5 22. Programa informático para realizar, cuando se ejecuta en un ordenador o un procesador, el método de la reivindicación 21.

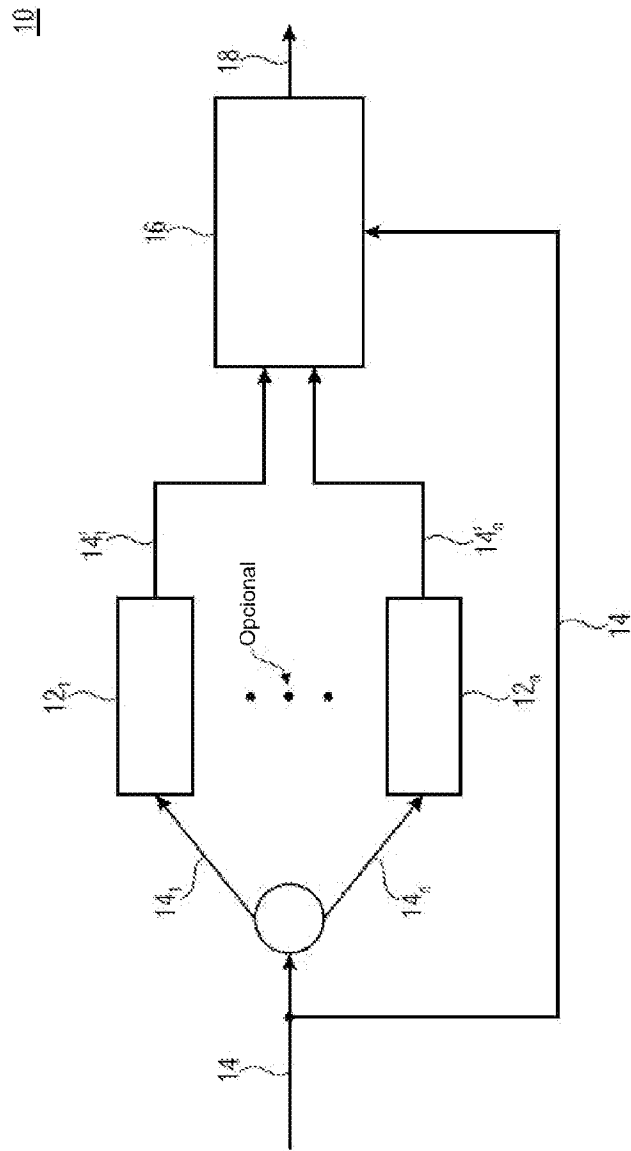


Fig. 1

20

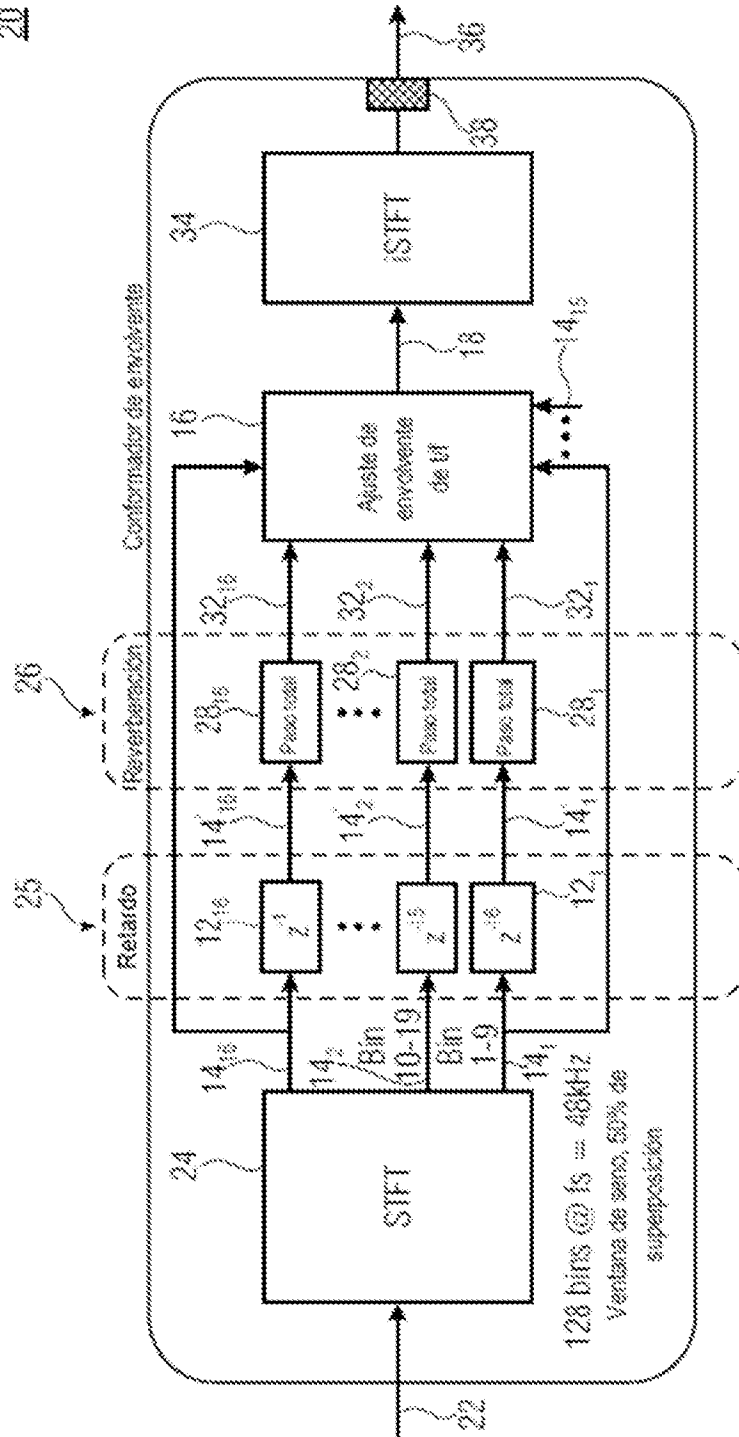


Fig. 2

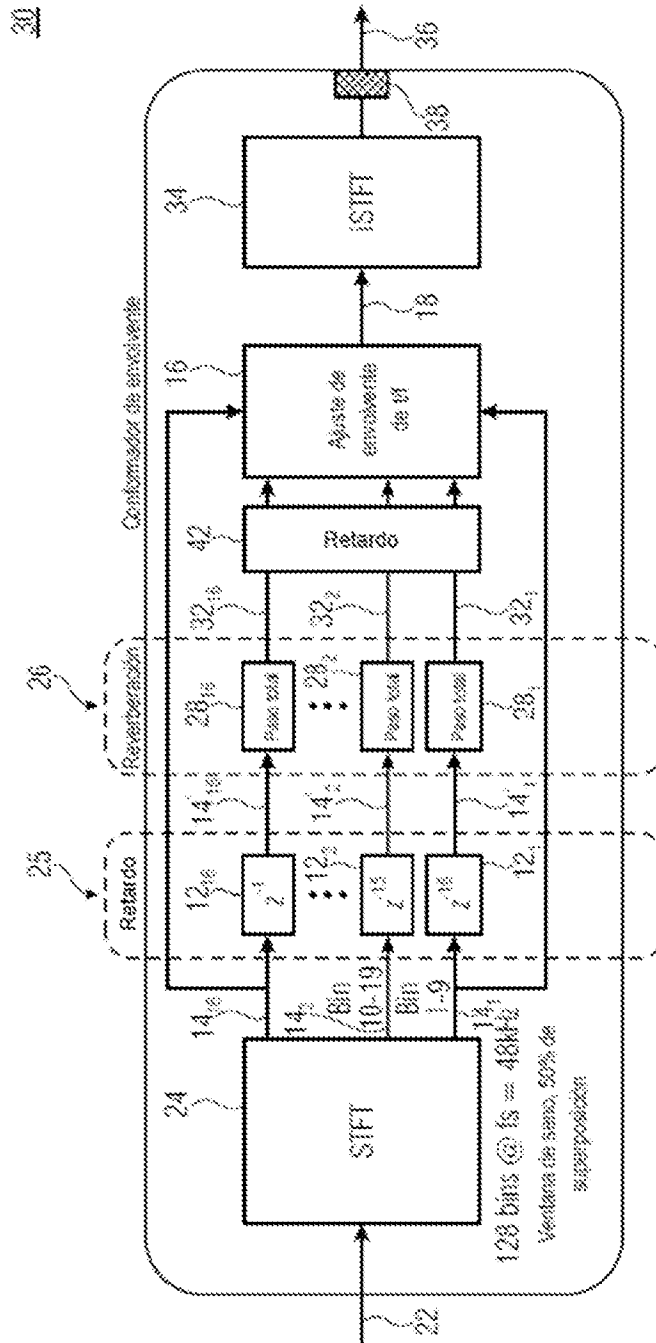


Fig. 3

40

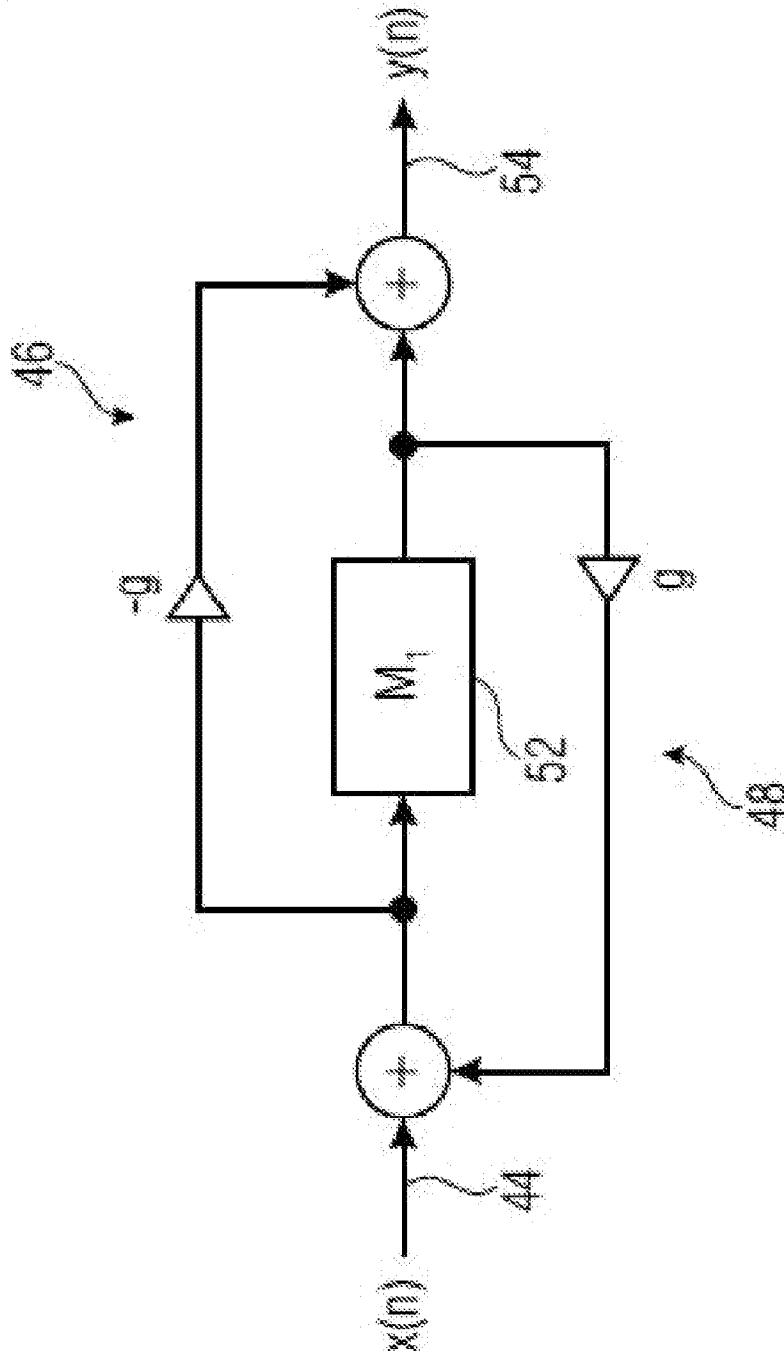


Fig. 4

50

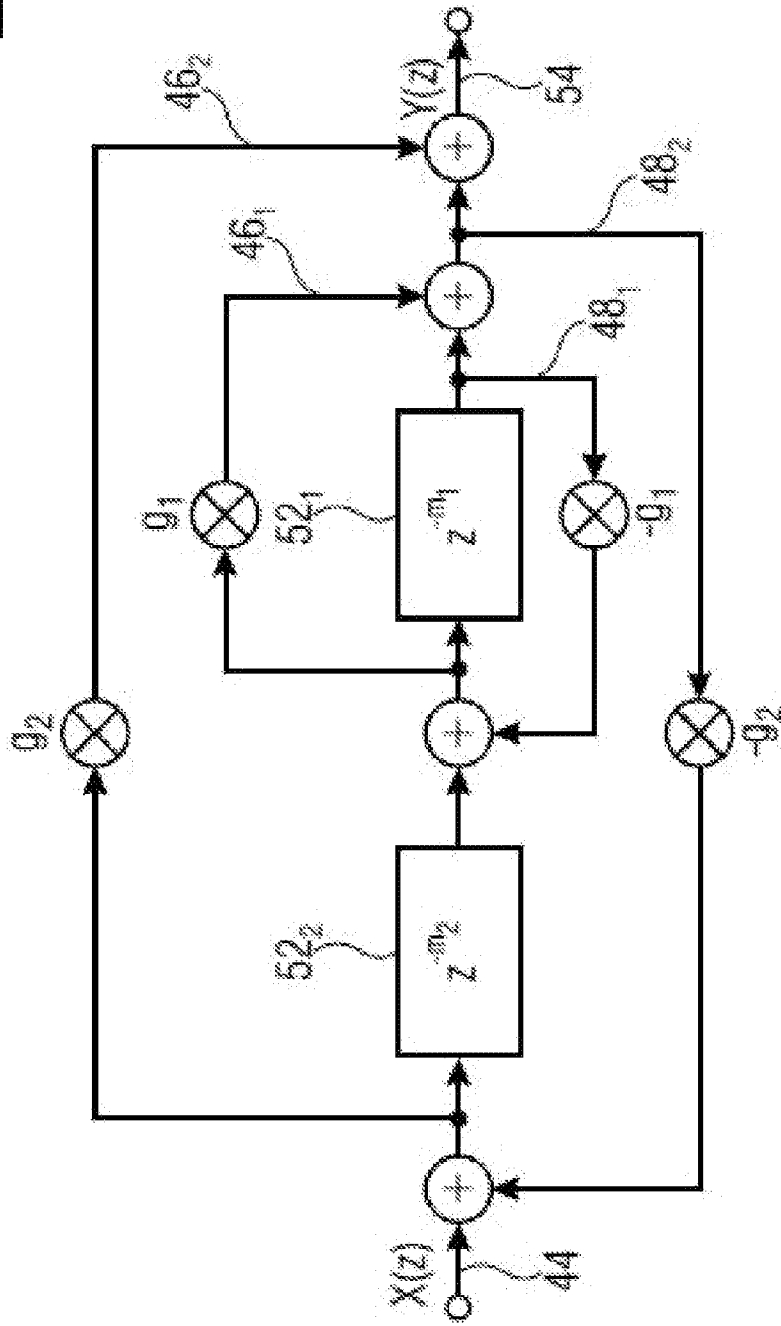


Fig. 5

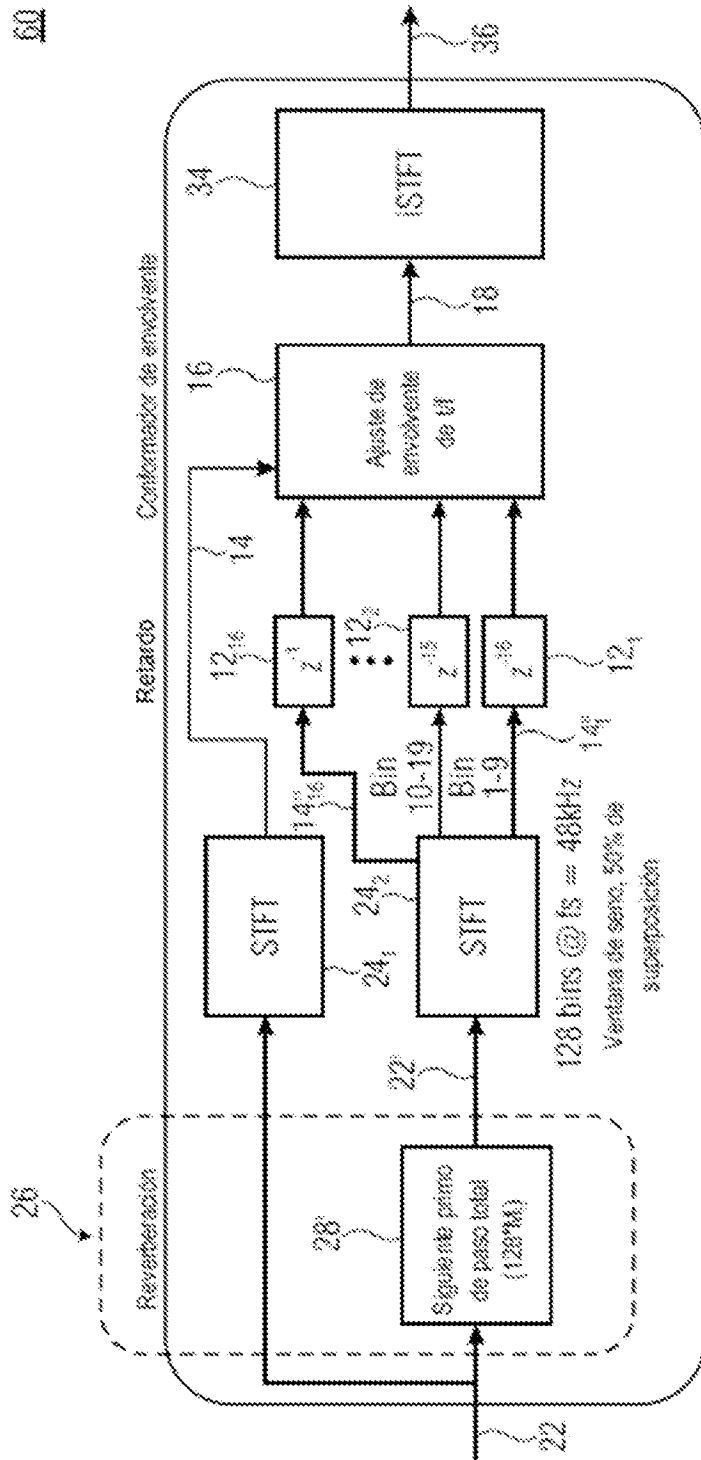


Fig. 6

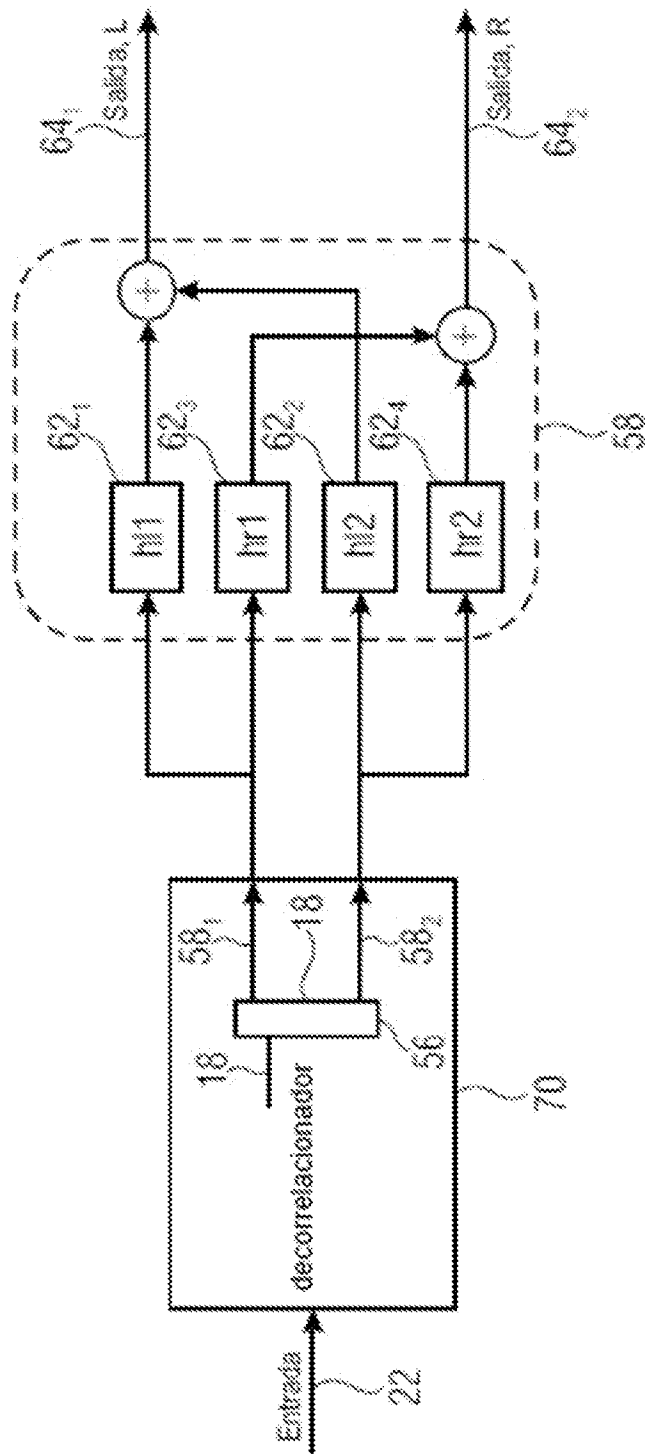


FIG. 7

80

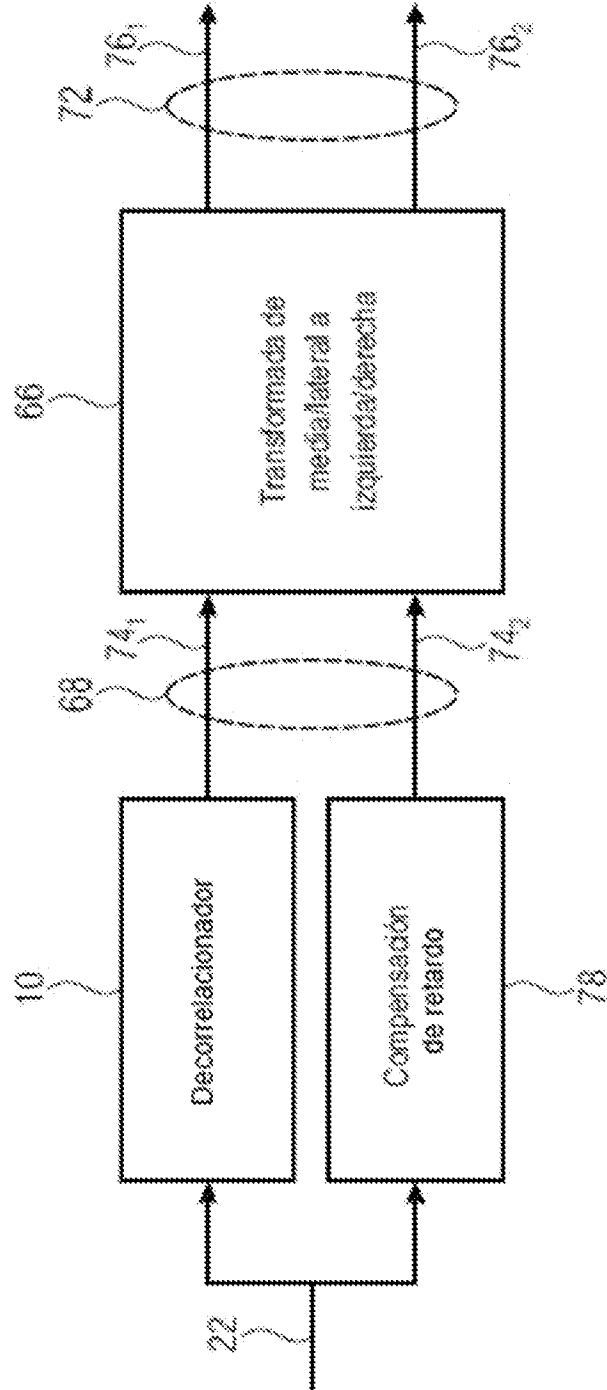


Fig. 8

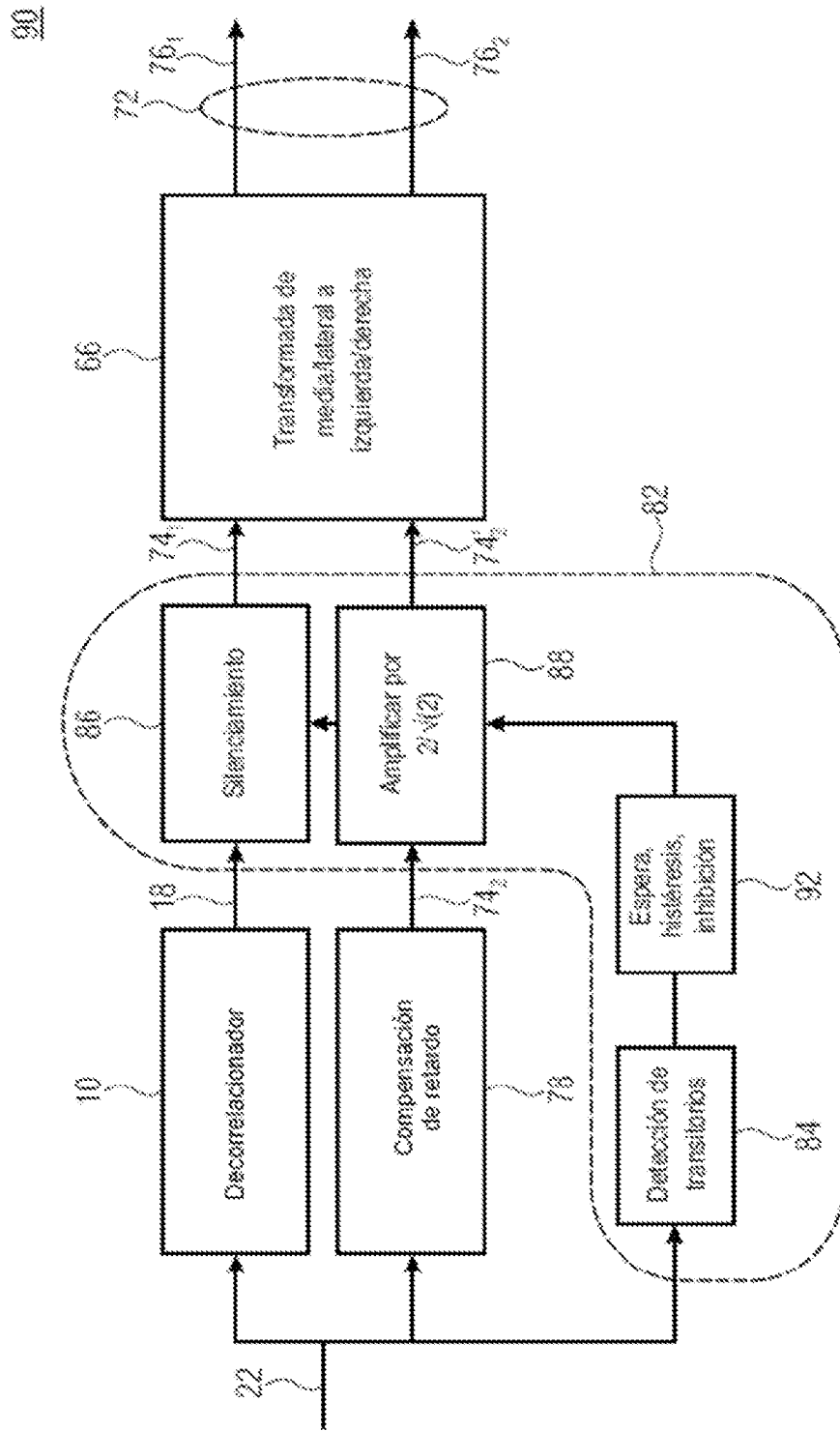


Fig. 9

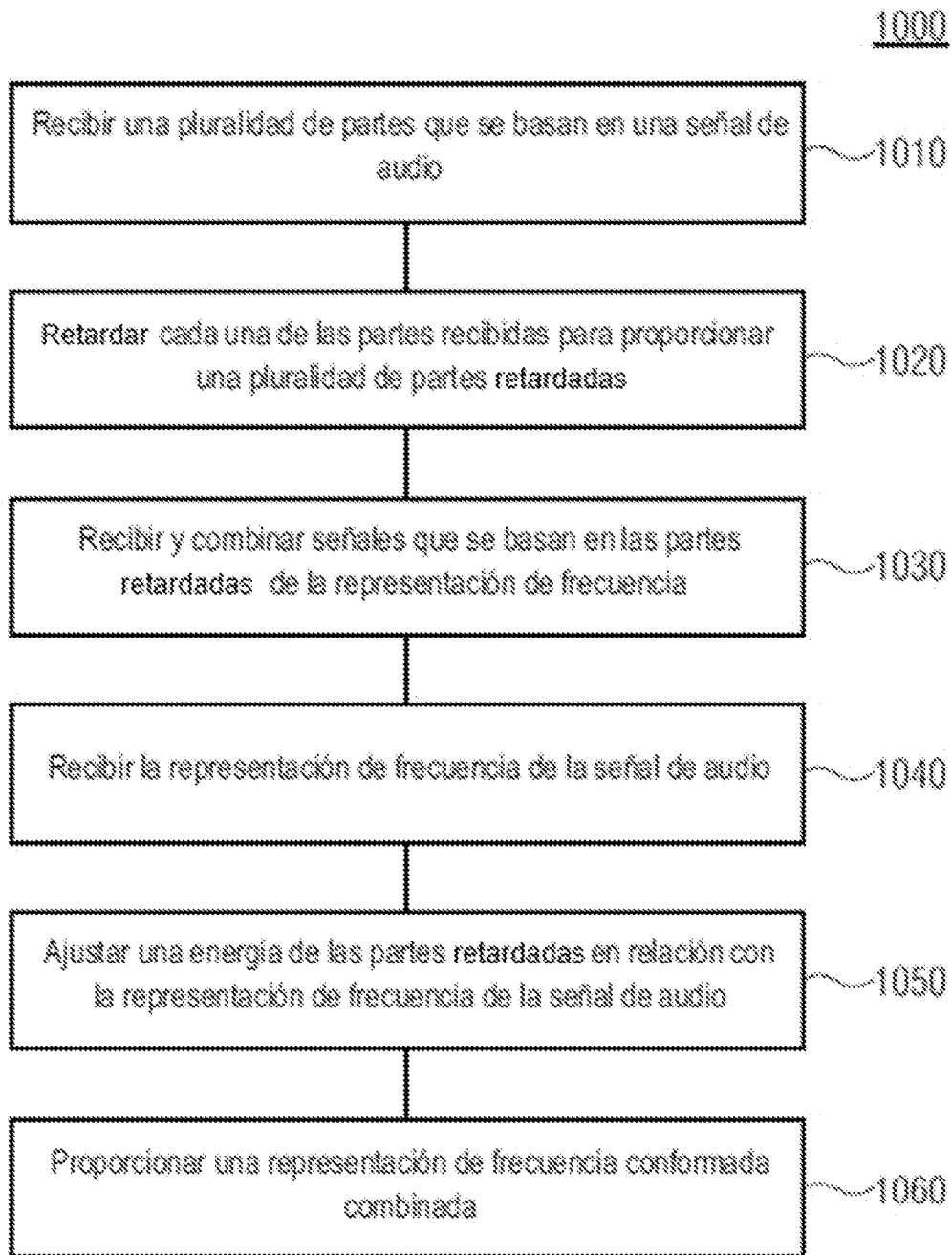


Fig. 10