

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 649 481**

51 Int. Cl.:

**G10L 19/008** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.09.2014 PCT/EP2014/068611**

87 Fecha y número de publicación internacional: **02.04.2015 WO15043891**

96 Fecha de presentación y número de la solicitud europea: **02.09.2014 E 14758881 (8)**

97 Fecha y número de publicación de la concesión europea: **18.10.2017 EP 3050054**

54 Título: **Procesamiento de señal de audio para generar una señal de mezcla descendente**

30 Prioridad:

**27.09.2013 EP 13186480**  
**21.03.2014 EP 14161059**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**12.01.2018**

73 Titular/es:

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

72 Inventor/es:

**ADAMI, ALEXANDER;**  
**HABETS, EMANUEL y**  
**HERRE, JÜRGEN**

74 Agente/Representante:

**SALVA FERRER, Joan**

**ES 2 649 481 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

Procesamiento de señal de audio para generar una señal de mezcla descendente

- 5 **[0001]** La presente invención se refiere al procesamiento de señales de audio y, en particular, a la mezcla descendente de una pluralidad de señales de entrada a una señal de mezcla descendente.
- [0002]** En el procesamiento de señales, a menudo se vuelve necesario mezclar dos o más señales a una señal de suma. El procedimiento de mezcla por lo general viene junto con algunas deficiencias de señal, especialmente si dos señales, que se van a mezclar, contienen partes de señal similares pero desfasadas. Si se suman aquellas señales, la señal resultante contiene varias distorsiones de filtro de peine. Para impedir aquellas distorsiones, se han sugerido diferentes procedimientos que son ya sea muy costosos en términos de complejidad computacional o están basados en la aplicación de un término o ganancia de corrección a la señal ya deteriorada.
- 10 **[0003]** La conversión de señales de audio multi-canal en un número menor de canales normalmente implica mezclar varios canales de audio. La ITU, por ejemplo, recomienda utilizar una matriz de mezcla pasiva, en el dominio del tiempo con ganancias estáticas para una conversión descendente de una cierta configuración multi-canal a otra [1]. En [2] se propone una estrategia algo similar.
- 20 **[0004]** Para incrementar la inteligibilidad de diálogo, se propone en [3] una estrategia combinada de uso de la mezcla descendente basada en la ITU y una mezcla descendente basada en matriz. También, codificadores de audio utilizan una mezcla descendente pasiva de canales, por ejemplo en algunos módulos paramétricos [4, 5, 6].
- [0005]** La estrategia descrita en [7] realiza una medición de volumen de cada canal de entrada y salida, es decir de todos los canales antes y después del proceso de mezcla. Al tomar la relación de la suma de las energías de entrada (es decir energía de los canales que se supone que se van a mezclar) y la energía de salida (es decir energía de los canales mezclados), se pueden derivar ganancias de tal forma que se reduce la pérdida de energía de señal y los efectos de coloración.
- 25 **[0006]** La estrategia descrita en [8] realiza una mezcla descendente pasiva que se transforma después al dominio de la frecuencia. La mezcla descendente se analiza entonces por una etapa de corrección espacial que trata de detectar y corregir cualquier inconsistencia espacial a través de modificaciones a las diferencias de nivel inter-canal y diferencias de fase inter-canal. Entonces, se aplica un ecualizador a la señal para asegurar que la señal de mezcla descendente tiene la misma potencia que la señal de entrada. En el último paso, la señal de mezcla
- 30 **[0007]** Se describe una estrategia diferente en [9, 10], donde dos señales, que se van a someter a mezcla descendente, se transforman al dominio de la frecuencia y se construye un par de valores deseados/reales. El valor deseado se calcula como la raíz de la suma de las energías individuales, mientras el valor real se calcula como la raíz de la energía de la señal de suma. Los dos valores se comparan entonces y dependiendo del valor real que es mayor o menor que el valor deseado, se aplica una corrección diferente al valor real.
- 35 **[0008]** De forma alternativa, existen procedimientos que tienen como objetivo alinear las fases de las señales, de tal forma que no se presentan efectos de cancelación de señales debido a las diferencias de fase. Estos procedimientos se propusieron por ejemplo para codificadores estéreo paramétricos [11, 12, 13].
- [0009]** Una mezcla descendente pasiva como se hace en [1, 2, 3, 4, 5, 6] es la estrategia más directa para mezclar señales. Pero si no se realiza ninguna acción adicional, las señales de mezcla descendente resultantes pueden sufrir de distorsiones de filtro de peine y pérdida de señal grave.
- 50 **[0010]** Las estrategias descritas en [7, 8, 9, 10] realizan una mezcla descendente pasiva, en el sentido de mezclar equitativamente ambas señales, en el primer paso. Después, se aplican algunas correcciones a la señal sometida a mezcla descendente. Esto puede ayudar a reducir los efectos de filtro de peine, pero por otro lado introducirán distorsiones de modulación. Esto se provoca por términos/ganancias de corrección que varían rápidamente con el paso del tiempo. Además, un cambio de fase de 180 grados entre las señales que se van a someter a mezcla descendente resulta aún en una mezcla descendente de valor cero y no se puede compensar al aplicar, por ejemplo, una ganancia de corrección.
- 55 **[0011]** Una estrategia de alineación de fases, tal como se menciona en [11, 12, 13], puede ayudar a evitar

cancelación indeseada de señales; pero debido a que aún se realiza un procedimiento de acumulación simple de las señales alineadas en fase puede presentarse cancelación y efectos de filtro de peine si no se estiman apropiadamente las fases. De forma adicional, la estimación robusta de las relaciones de fase entre dos señales no es una tarea fácil y es computacionalmente intensiva, especialmente si se hace para más de dos señales.

5

**[0012]** Las estrategias de mezcla descendente mediante el uso de un proceso ortogonal Gram-Schmidt [14] son conocidas también en la bibliografía.

**[0013]** Es un objeto de la presente invención proporcionar un concepto mejorado para someter a mezcla descendente una pluralidad de señales de entrada a una señal de mezcla descendente.

10

**[0014]** Este objeto se logra mediante un dispositivo según la reivindicación 1, un sistema según la reivindicación 16, un procedimiento según la reivindicación 17 o un programa informático de la reivindicación 18. Realizaciones adicionales de acuerdo con la invención son definidas por las reivindicaciones dependientes. El dispositivo se describirá en esta invención en el dominio del tiempo-frecuencia, pero todas las consideraciones también son verdaderas para señales en el dominio del tiempo. Una primera señal de entrada y segunda señal de entrada son las señales que se van a mezclar, donde la primera señal de entrada sirve como señal de referencia. Ambas señales se alimentan en un extractor de disparidad, donde se rechazan partes de señal correlacionadas de la segunda señal de entrada con respecto a la segunda señal de entrada y solamente se pasan las partes de señal no correlacionadas de la segunda señal de entrada a la salida del extractor.

15

**[0015]** La mejora del concepto propuesto radica en la forma que se mezclan las señales. En el primer paso, se selecciona una señal para servir como una referencia. Entonces se determina, qué parte de la señal de referencia ya está presente dentro de la otra, y solamente aquellas partes, que no están presentes en la señal de referencia (es decir la señal no correlacionada), se añaden a la referencia para construir la señal de mezcla descendente. Ya que solamente se combinan las partes de señal poco correlacionadas o no correlacionadas con respecto a la referencia con la referencia, se reduce al mínimo el riesgo de introducir efectos de filtro de peine.

25

**[0016]** Como un resumen, se propone un concepto novedoso para mezclar dos señales a una señal de mezcla descendente. El procedimiento novedoso tiene como objetivo evitar la creación de distorsiones de mezcla descendente, como efectos de filtro de peine. Además, el procedimiento propuesto es computacionalmente eficiente.

30

**[0017]** En algunas realizaciones de la invención el combinador comprende un sistema modificador de escala de energía configurado de tal forma que en la relación de la energía de la mezcla descendente y las energías sumadas de la primera señal de entrada y la segunda señal de entrada es independiente de la correlación de la primera señal de entrada y la segunda señal de entrada. Tal dispositivo modificador de escala de energía puede asegurar que el proceso de mezcla descendente es conservador de energía (es decir, la señal de mezcla descendente contiene la misma cantidad de energía que la señal estéreo original) o al menos que el sonido percibido permanece igual independientemente de la correlación de la primera señal de entrada y la segunda señal de entrada.

40

**[0018]** En realizaciones de la invención el sistema modificador de escala de energía comprende un primer dispositivo modificador de escala de energía configurado para modificar la escala de la primera señal de entrada con base en un primer factor de escala a fin de obtener una señal de entrada a escala.

45

**[0019]** En algunas realizaciones de la invención el sistema modificador de escala de energía comprende un primer proveedor de factor de escala configurado para proporcionar el primer factor de escala, en el que el primer proveedor de factor de escala se diseña preferentemente como un procesador configurado para calcular el primer factor de escala dependiendo de la primera señal de entrada, la segunda señal de entrada, la señal extraída y/o un factor de escala para la señal extraída. Durante la mezcla descendente, se puede modificar la escala de la señal de referencia (primera señal de entrada) para conservar el nivel de energía total o para mantener el nivel de energía independiente de la correlación de las señales de entrada de forma automática.

50

**[0020]** En realizaciones de la invención el sistema modificador de escala de energía comprende un segundo dispositivo modificador de escala de energía configurado para modificar la escala de la señal extraída con base en un segundo factor de escala a fin de obtener una señal extraída a escala.

55

**[0021]** En algunas realizaciones de la invención el sistema modificador de escala de energía comprende un segundo proveedor de factor de escala configurado para proporcionar el segundo factor de escala, en el que el

segundo proveedor de factor de escala se diseña preferentemente como una interfaz de hombre-máquina configurada para introducir de forma manual el segundo factor de escala.

**[0022]** El segundo factor de escala se puede ver como un ecualizador. En general, éste se puede hacer dependiente de la frecuencia y en realizaciones preferidas de forma manual mediante un ingeniero de sonido. Desde luego, son posibles varias relaciones de mezcla descendente y éstas dependen en gran medida de la experiencia y/o gusto del ingeniero de sonido.

**[0023]** De forma alternativa, el segundo proveedor de factor de escala se diseña preferentemente como un procesador configurado para calcular el primer factor de escala dependiendo de la primera señal de entrada, la segunda señal de entrada y/o la señal extraída.

**[0024]** En algunas realizaciones de la invención el combinador comprende un dispositivo sumador para sacar la señal de mezcla descendente con base en la primera señal de entrada y con base en la señal extraída. Ya que solamente se añaden a la referencia partes de señal poco correlacionadas o incluso no relacionadas con respecto a la referencia, se reduce al mínimo el riesgo de introducir efectos de filtro de peine. Además, el uso de un dispositivo sumador es computacionalmente eficiente.

**[0025]** De acuerdo con la invención el extractor de disparidad comprende un estimador de similitud configurado para proporcionar coeficientes de filtro para obtener las partes de señal de la primera señal de entrada que están presentes en la segunda señal de entrada de la primera señal de entrada y un reductor de similitud configurado para reducir las partes de señal de la primera señal de entrada que están presentes en la segunda señal de entrada con base en los coeficientes de filtro. En tales implementaciones, el extractor de disparidad consiste en dos sub-etapas: un estimador de similitud y un reductor de similitud. La primera señal de entrada y la segunda señal de entrada se alimentan en una etapa de estimación de similitud, donde se estiman las partes de señal de la primera señal de entrada que están presentes dentro de la segunda señal de entrada y se representan por los coeficientes de filtro resultantes. Los coeficientes de filtro, la primera señal de entrada y la segunda señal de entrada se alimentan en el reductor de similitud donde se suprimen y/o cancelan, respectivamente las partes de señal de la segunda señal de entrada que son similares a la primera señal de entrada. Esto da como resultado que la señal extraída sea una estimación para la parte de señal no correlacionada de la segunda señal de entrada con respecto a la primera señal de entrada.

**[0026]** En algunas realizaciones de la invención el reductor de similitud comprende una etapa de cancelación que tiene un dispositivo de cancelación de señal configurado para sustraer las partes de señal obtenidas de la primera señal de entrada que están presentes en la segunda señal de entrada o una señal derivada de las partes de señal obtenidas de la segunda señal de entrada o de una señal derivada de la segunda señal de entrada. Este concepto está relacionado con un procedimiento que se utiliza en el tema de cancelación adaptativa de ruido pero con la diferencia de que no se utiliza, como se propuso de forma original, para cancelar el ruido o componente no correlacionado sino en su lugar para cancelar la parte de señal correlacionada, que da como resultado la señal extraída.

**[0027]** En algunas realizaciones de la invención la etapa de cancelación comprende un dispositivo de filtro complejo configurado para filtrar la primera señal de entrada al utilizar coeficientes de filtro de valores complejos. La ventaja de esta estrategia es que se pueden modelar los cambios de fase.

**[0028]** En algunas realizaciones de la invención la etapa de cancelación comprende un dispositivo de cambio de fase configurado para alinear la fase de la segunda señal de entrada con la fase de la primera señal de entrada. Para fases opuestas entre la primera señal de entrada y la segunda señal de entrada además con caídas repentinas de señal de la primera señal de entrada, pueden presentarse saltos de fase y efectos de cancelación de señal dentro de la señal de mezcla descendente. Este efecto se puede reducir de forma drástica al alinear la fase de la segunda señal de entrada hacia la primera señal de entrada. Tal etapa de cancelación puede ser llamada etapa de cancelación inversa de fases alineadas.

**[0029]** De acuerdo con la invención el reductor de similitud comprende una etapa de supresión de señal que tiene un dispositivo de supresión de señales configurado para multiplicar la segunda señal de entrada con un factor de ganancia de supresión a fin de obtener la señal extraída. Se ha observado que se pueden reducir distorsiones audibles debido a errores de estimación en los coeficientes de filtro mediante estas características.

**[0030]** En algunas realizaciones de la invención la etapa de supresión de señal comprende un dispositivo de

cambio de fase configurado para alinear la fase de la segunda señal de entrada con la fase de la primera señal de entrada. Los factores de ganancia de supresión son de valores reales y por lo tanto no tienen influencia en las relaciones de fase de las dos señales de entrada, pero ya que se tienen que estimar de todas formas los coeficientes de filtro de valores complejos, se puede obtener información adicional acerca de la fase relativa entre las señales de entrada. Esta información se puede utilizar para ajustar la fase de la segunda señal de entrada hacia la primera señal de entrada. Esto se puede hacer dentro de la etapa de supresión de señal antes de que se apliquen las ganancias de supresión, en la que se cambia la fase de la segunda señal de entrada mediante la fase estimada de los factores de filtro de valores complejos mencionados anteriormente. Tal etapa de supresión puede ser llamada etapa de supresión inversa de fases alineadas.

10

**[0031]** En algunas realizaciones de la invención se alimenta una señal de salida de la etapa de cancelación a una entrada de la etapa de supresión de señal a fin de obtener la señal extraída o se alimenta una señal de salida de la etapa de supresión de señal a una entrada de la etapa de cancelación a fin de obtener la señal extraída. Se puede utilizar una estrategia combinada de uso de cancelación así como supresión de componentes de señal coherentes para incrementar de forma adicional la calidad de la señal de mezcla descendente. La señal de mezcla descendente resultante se puede obtener al realizar primero un procedimiento de cancelación y después aplicando un procedimiento de supresión. En otras realizaciones, la señal de mezcla descendente resultante se puede obtener al realizar primero un procedimiento de supresión y después aplicando un procedimiento de cancelación. De esta forma, se pueden reducir de forma adicional las partes de señal en la señal extraída, que están correlacionadas con la primera señal. Se puede modificar la escala de energía de la señal extraída así como de la primera señal de entrada como antes.

15

20

**[0032]** En algunas realizaciones de la invención las partes de señal de la primera señal de entrada que están presentes en la segunda señal de entrada se ponderan antes de que se sustraigan de la segunda señal de entrada dependiendo de un factor de ponderación. Un factor de ponderación puede ser en general dependiente del tiempo y la frecuencia pero también se puede elegir como constante. En algunas realizaciones, aquí también se puede utilizar el módulo de cancelación inversa de fases alineadas con una pequeña modificación: la ponderación con el factor de ponderación se tiene que realizar de forma análoga después del filtrado con el valor absoluto de los coeficientes de filtro.

30

**[0033]** En algunas realizaciones de la invención el dispositivo de cambio de fases se configura para alinear la fase de la segunda señal de entrada con la fase de la primera señal de entrada dependiendo del factor de ponderación.

35

**[0034]** En algunas realizaciones de la invención el dispositivo de cambio de fase se configura para alinear la fase de la segunda señal de entrada con la fase de la primera señal de entrada solamente, si el factor de ponderación es menor o igual a un umbral predefinido. La invención se refiere además a un sistema de procesamiento de señales de audio para la mezcla descendente de una pluralidad de señales de entrada a una señal de mezcla descendente que comprende al menos un primer dispositivo según la invención y un segundo dispositivo según la invención, en el que la señal de mezcla descendente del primer dispositivo se alimenta al segundo dispositivo como una primera señal de entrada o como una segunda señal de entrada. Para someter a mezcla descendente una pluralidad de canales de entrada, se puede utilizar una cascada de una pluralidad de dispositivos de mezcla descendente de dos canales.

40

45

**[0035]** Además, la invención se refiere a un procedimiento de procesamiento de señal de audio para la mezcla descendente de una primera señal de entrada y una segunda señal de entrada a una señal de mezcla descendente y un programa informático, tal como se define en las reivindicaciones 17 y 18, respectivamente. Las realizaciones preferidas se tratan posteriormente con respecto a los dibujos adjuntos, en los que:

50

La Figura 1 ilustra una primera realización de un dispositivo de procesamiento de señales de audio;

La Figura 2 ilustra la primera realización en más detalle;

La Figura 3 ilustra un reductor de similitud y un combinador de la primera realización;

La Figura 4 ilustra un reductor de similitud de una segunda realización;

La Figura 5 ilustra un reductor de similitud y un combinador de una tercera realización;

55

La Figura 6 ilustra un reductor de similitud de una cuarta realización;

La Figura 7 ilustra un reductor de similitud y un combinador de una quinta realización;

La Figura 8 ilustra un reductor de similitud y un combinador de una sexta realización; y

La Figura 9 ilustra una cascada de una pluralidad de dispositivos de procesamiento de señales de audio.

**[0036]** La Figura 1 muestra una descripción de sistema de alto nivel del dispositivo de mezcla descendente novedoso propuesto 1. El dispositivo se describe en el dominio del tiempo-frecuencia, donde  $k$  y  $m$  corresponden a índices de frecuencia y tiempo respectivamente, pero todas las consideraciones también son verdaderas para señales en el dominio del tiempo. Una primera señal de entrada  $X_1(k, m)$  y segunda señal de entrada  $X_2(k, m)$  son las señales de entrada que se van a mezclar, donde la primera señal de entrada  $X_1(k, m)$  puede servir como señal de referencia. Ambas señales  $X_1(k, m)$  y  $X_2(k, m)$  se alimentan en un extractor de disparidad 2, donde se rechazan partes de señal correlacionadas con respecto a  $X_1(k, m)$  y  $X_2(k, m)$  o al menos se reducen y solamente se extrae y

se pasa la señal no correlacionada o las partes poco correlacionadas  $\hat{U}_2(k, m)$  a la salida del extractor. Entonces, la primera señal de entrada  $X_1(k, m)$  se modifica en escala utilizando un primer dispositivo modificador de escala de energía 4 para cumplir algunas limitaciones de energía predefinidas, que da como resultado una señal de referencia a escala  $X_{1s}(k, m)$ . Los factores de escala necesarios  $G_{Ex}(k, m)$  se proporcionan por el proveedor de factor de escala

5. La parte de señal extraída  $\hat{U}_2(k, m)$  también se puede modificar en escala utilizando un segundo dispositivo modificador de escala de energía 6, que da como resultado una parte de señal no correlacionada a escala

$\hat{U}_{2s}(k, m)$ . Los factores de escala correspondientes  $G_{Eu}(k, m)$  se proporcionan por el segundo proveedor de

15 factor de escala 7. Los factores de escala  $G_{Eu}(k, m)$  se pueden determinar preferentemente de forma manual

mediante un ingeniero de sonido. Ambas señales a escala  $X_{1s}(k, m)$  y  $\hat{U}_{2s}(k, m)$  se suman utilizando un

dispositivo sumador 8 para formar la señal de mezcla descendente deseada  $\tilde{X}_D(k, m)$ .

**[0037]** La Figura 2 muestra una descripción de sistema de nivel medio del dispositivo propuesto 1. En algunas implementaciones, el extractor de disparidad 2 consiste en dos sub-etapas: un estimador de similitud 9 y un reductor de similitud 10 como se representa en la Figura 2. La primera señal de entrada  $X_1(k, m)$  y la segunda señal de entrada  $X_2(k, m)$  se alimentan en una etapa de estimación de similitud 9, donde las partes de señal de  $X_1(k, m)$

que están presentes dentro de  $X_2(k, m)$  se estiman y representan por los coeficientes de filtro resultantes  $W_k(l)$

con  $l = 0 \dots L - 1$  y  $L$  que es la longitud de filtro. Los coeficientes de filtro  $W_k(l)$ , la primera señal de entrada  $X_1(k, m)$  y la segunda señal de entrada  $X_2(k, m)$  se alimentan en el reductor de similitud 10, donde las partes de señal de  $X_2(k, m)$  que son similares a  $X_1(k, m)$  se suprimen y/o cancelan al menos parcialmente, respectivamente. Esto da

como resultado la señal residual  $\hat{U}_2(k, m)$ , que es una estimación para la parte de señal no correlacionada de  $X_2(k, m)$  con respecto a  $X_1(k, m)$ .

30 **[0038]** El modelo de señal asume que la segunda señal de entrada  $X_2(k, m)$  es una mezcla de una versión

ponderada o filtrada  $W^l(k, m)X_1(k, m)$  de la primera señal de entrada  $X_1(k, m)$  y una señal independiente

inicialmente desconocida  $U_2(k, m)$  con  $E\{X_1 U_2^* \} = 0$ . Por lo tanto,  $X_2(k, m)$  se considera que consiste en la suma de una parte de señal correlacionada y una parte de señal no correlacionada con respecto a  $X_1(k, m)$ :

$$X_2(k, m) = W^l(k, m) \cdot X_1(k, m) + U_2(k, m) \quad (1)$$

35

**[0039]** Las letras mayúsculas indican señales transformadas en frecuencia y  $k$  y  $m$  son los índices de

frecuencia y tiempo respectivamente. Ahora la señal de mezcla descendente deseada  $\tilde{X}_D(k, m)$  se puede definir como:

$$\tilde{X}_D(k, m) = G_{Ex}(k, m)X_1(k, m) + G_{Eu}(k, m)\hat{U}_2(k, m), \quad (2)$$

(2)

donde  $\hat{U}_2(k, m)$  es una estimación de  $U_2(k, m)$  y donde  $G_{E_r}(k, m)$  y  $G_{E_n}(k, m)$  son factores de escala

para ajustar las energías de la señal de referencia  $X_1(k, m)$  y la parte de señal extraída  $\hat{U}_2(k, m)$  de la otra señal de entrada  $X_2(k, m)$  según las limitaciones predefinidas. De forma adicional, se pueden utilizar para

5 ecualizar las señales. En algunos escenarios esto puede volverse necesario, especialmente para  $\hat{U}_2(k, m)$ . En el resto de este documento los índices de tiempo-frecuencia  $(k, m)$  se omitirán para claridad.

**[0040]** El objetivo primordial es obtener el componente de señal  $U_2$ , que no está correlacionado con  $X_1$ . Esto se puede hacer al utilizar un procedimiento que se usa en el tema de cancelación adaptativa de ruido pero con la diferencia de que no se utiliza, como se propuso de forma original, para cancelar el ruido o componente no correlacionado, sino en su lugar la parte de señal correlacionada, que da como resultado la estimación  $\hat{U}_2$  de  $U_2$ .

**[0041]** La Figura 3 representa un reductor de similitud 10 que tiene una etapa de cancelación 10a y un combinador 3 de la primera realización de tal sistema. La ventaja de esta estrategia es que se permite que  $W$  sea complejo y por lo tanto se pueden modelar los cambios de fase.

$$\hat{U}_2 = X_2 - WX_1 \quad (3)$$

**[0042]** Para determinar  $\hat{U}_2$ , se necesita una ganancia compleja estimada  $W$  para la ganancia compleja inicialmente desconocida  $W'$ . Esto se realiza al reducir al mínimo la energía de la señal extraída  $\hat{U}_2$  en el sentido de la mínima media cuadrática (MMS):

$$\begin{aligned} J(W) &= E\{|X_2 - WX_1|^2\} \\ &= E\{(X_2 - WX_1)(X_2 - WX_1)^*\} \\ &= E\{X_2X_2^* - X_2W^*X_1^* - WX_1X_2^* + WX_1W^*X_1^*\} \end{aligned} \quad (4)$$

**[0043]** El ajuste de la derivada parcial de  $J(W)$  con respecto a  $W'$  a cero conduce a los coeficientes de filtro deseados, es decir:

$$\frac{\partial}{\partial W^*} J(W) = E\{X_2X_1^*\} - W E\{|X_1|^2\} = 0 \quad (5)$$

$$\Rightarrow W = \frac{E\{X_2X_1^*\}}{E\{|X_1|^2\}} \quad (6)$$

**[0044]** En una realización, el módulo de cancelación 10a, destacado por el rectángulo discontinuo gris en la Figura 3, se puede reemplazar por un bloque de cancelación inversa de fases alineada 10a' como se representa en la Figura 4, en el que la etapa de cancelación 10a' comprende un dispositivo de cambio de fase 13 configurado para alinear la fase de la segunda señal de entrada  $X_2$  con la fase de la primera señal de entrada  $X_1$  y un dispositivo de filtro absoluto 11' configurado para filtrar una primera señal de entrada alineada ( $X_2$  al utilizar coeficientes de filtro de valor absoluto  $|W|$ ).

**[0045]** Para fase opuesta de la primera señal de entrada  $X_1$  y la segunda señal de entrada  $X_2$  además con caídas repentinas de señal de la primera señal de entrada  $X_1$ , pueden presentarse saltos de fase y efectos de

cancelación de señal dentro de la señal de mezcla descendente  $\hat{X}_D$ . Este efecto se puede reducir drásticamente al alinear la fase de la segunda señal de entrada  $X_2$  hacia la fase de la primera señal de entrada  $X_1$ . Además, sólo se utiliza el valor absoluto de  $W$  para realizar el filtrado de  $X_1$  y por lo tanto también la cancelación.

**[0046]** La Figura 5 ilustra un reductor de similitud 10 y un combinador 3 de una tercera realización, según la invención, en donde el reductor de similitud 10 comprende una etapa de supresión de señal 10b que tiene un dispositivo de supresión de señal 14 configurado para multiplicar la segunda señal de entrada  $X_2$  con un factor de ganancia de supresión ( $G$ ) a fin de obtener la señal extraída  $\hat{U}_2$ .

**[0047]** En la práctica, la señal extraída  $\hat{U}_2$  obtenida utilizando (3) podría contener distorsiones audibles debido a errores de estimación en la ganancia compleja  $W$ . Como una alternativa, se puede derivar un estimador 9 (véase la Figura 2) para obtener una estimación  $\hat{U}_2$  de  $U_2$  en el sentido del error cuadrático medio mínimo (MMSE). La Figura 5 muestra un diagrama de bloques de la estrategia propuesta.

**[0048]** La señal extraída  $\hat{U}_2$  está dada entonces por

$$G = \arg \min_G E \left\{ |U_2 - \hat{U}_2|^2 \right\} \quad G \in \mathbf{R} \quad (8)$$

$$\begin{aligned} J(G) &= E \left\{ |U_2 - \hat{U}_2|^2 \right\} = E \left\{ |U_2 - GX_2|^2 \right\} = E \left\{ |U_2 - GWX_1 - GU_2|^2 \right\} \\ &= E \left\{ (U_2 - GWX_1 - GU_2)(U_2 - GWX_1 - GU_2)^* \right\} \\ &= E \left\{ |U_2|^2 \right\} - G E \left\{ |U_2|^2 \right\} + G^2 E \left\{ |WX_1|^2 \right\} - G E \left\{ |U_2|^2 \right\} + G^2 E \left\{ |U_2|^2 \right\} \\ &= \Phi_{U_2}(1 - 2G + G^2) + G^2 \Phi_{WX_1} \end{aligned} \quad (9)$$

15 **[0049]** El ajuste de la derivada parcial de  $J(G)$  con respecto a  $G$  a cero conduce a las ganancias deseadas:

$$\frac{\partial}{\partial G} J(G) = \Phi_{U_2}(-2 + 2G) + 2G\Phi_{WX_1} \stackrel{!}{=} 0 \quad (10)$$

$$\begin{aligned} 2\Phi_{U_2}(-1 + G) + 2G\Phi_{WX_1} &= 0 \\ -\Phi_{U_2} + \Phi_{U_2}G + G\Phi_{WX_1} &= 0 \\ G \cdot (\Phi_{U_2} + \Phi_{WX_1}) &= \Phi_{U_2} \\ G &= \frac{\Phi_{U_2}}{\Phi_{U_2} + \Phi_{WX_1}} = \frac{\Phi_{U_2}}{\Phi_{X_2}} \end{aligned} \quad (11)$$

**[0050]** Según (12), se puede sustituir la energía de  $X_2$  por la suma de las energías de la versión filtrada de  $X_1$  y la señal no correlacionada  $U_2$ :

$$\begin{aligned} \Phi_{X_2} &= E \left\{ |X_2|^2 \right\} = E \left\{ (WX_1 + U_2)(WX_1 + U_2)^* \right\} \\ &= E \left\{ |WX_1|^2 \right\} + E \left\{ |U_2|^2 \right\} = \Phi_{WX_1} + \Phi_{U_2}. \end{aligned} \quad (12)$$

20 **[0051]** Para las ganancia  $G$ , esto conduce a

$$G = \frac{\Phi_{U_2}}{\Phi_{U_2} + \Phi_{WX_1}} = \frac{1}{1 + \frac{\Phi_{WX_1}}{\Phi_{U_2}}} = \frac{1}{1 + \underbrace{\frac{1}{\text{SNR}_{U_2(WX_1)}}}_{\text{a priori SNR}}}, \quad 0 \leq G \leq 1 \quad (13)$$

con  $\text{SNR}_{U_2(WX_1)}$  que es la SNR a priori de  $X_2$ . Las ganancias de filtro complejas  $W$  se determinan utilizando (6).

**[0052]** En una realización, el módulo de supresión 10b, destacado por el rectángulo gris discontinuo en la



Figura 5, se puede reemplazar por un módulo de supresión inversa de fases alineada 10b' que comprende un dispositivo de cambio de fase 15 configurado para alinear la fase de la segunda señal de entrada  $X_2$  con la fase de la primera señal de entrada  $X_1$ .

- 5 **[0053]** La Figura 6 ilustra un reductor de similitud 10b' que tiene este dispositivo de cambio de fase 15 como la cuarta realización. Las ganancias de supresión  $G$  son de valor real y, por tanto, no tienen ninguna influencia en las relaciones de fase de las dos señales  $X_1$  y  $X_2$ . Pero ya que los coeficientes de filtro  $W$  se tienen que estimar de todas formas, se puede adquirir información adicional acerca de la fase relativa entre las señales de entrada. Esta información se puede utilizar para ajustar la fase de  $X_2$  hacia la fase de  $X_1$ . Esto se hace dentro del bloque de  
10 supresión inversa de fase alineada 10b'; antes de que se apliquen las ganancias de supresión  $G$ , se cambia la fase de  $X_2$  por la fase estimada de  $W$ . Con una alineación de fases, la señal  $\hat{U}_2$  se puede expresar como

$$\begin{aligned}\hat{U}_2 &= X_2 \cdot e^{-j\angle\hat{W}} \cdot G \\ &= \left( |W| \cdot e^{j(\angle W - \angle\hat{W})} X_1 + U_2 \cdot e^{-j\angle\hat{W}} \right) \cdot G,\end{aligned}\quad (14)$$

que muestra que el componente residual de  $X_1$  dentro de  $\hat{U}_2$  está en fase con respecto a  $X_1$  siempre que  $\angle W$  se estime correctamente.

- 15 **[0054]** Una estrategia combinada de uso de cancelación así como supresión de componentes de señal coherentes se representa en la Figura 7, en la que una señal de salida  $\hat{U}_2$  de la etapa de cancelación 10a se alimenta a una entrada de la etapa de supresión de señal 10b a fin de obtener la señal extraída  $\hat{U}_2$ . La etapa de cancelación 10a comprende un dispositivo de ponderación configurado para ponderar las partes de señal obtenidas  $WX_1$  de la primera señal de entrada  $X_1$  que están presentes en la segunda señal de entrada ( $X_2$ ).  
20

- [0055]** Aquí, la señal de mezcla descendente resultante  $\hat{X}_D$  se obtiene al realizar un procedimiento de cancelación ponderada, primero y, después aplicando una ganancia de supresión. La señal resultante  $\hat{U}_2$  así como  $X_1$  se modifican en escala de energía como antes. Debido al factor de ponderación  $\gamma$ , la señal  $\hat{U}_2$  después de la etapa de cancelación contiene aún algunas partes de señal correlacionadas con  $X_1$ . Para reducir de forma adicional  
25 aquellas partes de señal, se deriva la ganancia de supresión  $G_c$  para la estrategia combinada:

$$G_c = \arg \min_{G_c} E \left\{ \left| U_2 - \hat{U}_2 \right|^2 \right\}, \quad G_c \in \mathbb{R} \quad (15)$$

$$J'(G_c) = E \left\{ \left| U_2 - \hat{U}_2 \right|^2 \right\} = \Phi_{U_2} - G_c \Phi_{U_2} + (1 - \gamma)^2 G_c^2 \Phi_{WX_1} - G_c \Phi_{U_2} + G_c^2 \Phi_{U_2} \quad (16)$$

$$\frac{\partial}{\partial G} J'(G_c) = -\Phi_{U_2} + 2(1 - \gamma)^2 G_c \Phi_{WX_1} - \Phi_{U_2} + 2G_c \Phi_{U_2} \stackrel{!}{=} 0 \quad (17)$$

$$G_c = \frac{1}{1 + (1 - \gamma)^2 \frac{\Phi_{WX_1}}{\Phi_{U_2}}} = \frac{1}{1 + (1 - \gamma)^2 \frac{1}{\text{SNR}_{U_2 WX_1}}} \quad (18)$$

- [0056]** El parámetro  $\gamma$  es en general dependiente del tiempo y la frecuencia pero también se puede elegir como constante. Una posibilidad para determinar un  $\gamma$  dependiente del tiempo y frecuencia es:

$$\gamma = 1 - \frac{|E \{ X_2 X_1^* \}|}{\sqrt{\Phi_{X_1} \Phi_{X_2}}} \quad (19)$$

- 30 **[0057]** La Figura 8 ilustra un reductor de similitud 10 y un combinador 3 de una sexta realización. Según esta realización la correlación cruzada normalizada en (19) se alimenta como entrada a una función de asignación cuya salida se puede utilizar para determinar los  $\gamma$ -valores reales. Para la asignación, se puede utilizar una función logística que se puede definir como:

$$f(i) = A_l + \frac{A_u - A_l}{\left(1 + (-1 + (\frac{A_u}{Y_0})^v) \cdot e^{-R(i+M)}\right)^{\frac{1}{v}}}, \quad (20)$$

donde  $i$  define los datos de entrada,  $A_u$  y  $A_l$  la asíntota superior e inferior,  $R$  es la velocidad de crecimiento,  $v > 0$  influye en la velocidad de crecimiento máxima cerca de la asíntota,  $f_0$  especifica el valor de salida para  $f(0)$  y  $M$  es el punto de datos  $i$  de crecimiento máximo. En tal realización,  $\gamma$  se determina por

$$\gamma = 1 - f\left(\frac{|\mathbf{E}\{X_2 X_1^*\}|}{\sqrt{\Phi_{X_1} \Phi_{X_2}}} - 0.5\right) \quad (21)$$

5 **[0058]** En una realización, aquí también se puede utilizar el módulo de cancelación inversa de fases alineadas 10a' con una pequeña modificación. La ponderación con  $\gamma$  se tiene que realizar de forma análoga después del filtrado con el valor absoluto de  $W$ .

10 **[0059]** Una sexta realización mostrada en la Figura 8 comprende una aplicación más sofisticada del procesamiento de fase inversa. Afecta solamente a los intervalos de tiempo-frecuencia que se asignaron para suprimirse principalmente, es decir  $\gamma$  está por debajo de un cierto umbral  $\Gamma_{um}$ . Por esa razón, se introduce una bandera  $F$  definida por

$$F = \begin{cases} 1 & \gamma \leq \Gamma_{um} \\ 0 & \text{de otra forma} \end{cases} \quad (22)$$

**[0060]** En una realización, el módulo de cancelación de fase alineada inversa 10a' se puede utilizar aquí también con una pequeña modificación. La ponderación con  $\gamma$  se debe realizar de forma análoga después del filtrado con el valor absoluto de  $W$ .

**[0061]** En algunas realizaciones el proveedor de factor de escala 7 proporciona  $G_{E_{x_1}}$ , por lo cual se puede controlar la cantidad de energía de la señal no correlacionada  $\hat{U}_2$  con respecto a  $X_1$  que contribuye a la señal de mezcla descendente  $\hat{X}_D$ . Estos factores de escala  $G_{E_{x_1}}$  se pueden ver como un ecualizador. En general, éste se puede hacer dependiente de la frecuencia y en la realización preferida de forma manual por un ingeniero de sonido. Desde luego, son posibles muchas relaciones de mezcla diferentes y éstas dependen en gran medida de la experiencia y/o gusto del ingeniero de sonido. De forma alternativa, los factores de escala  $G_{E_{x_1}}$  pueden ser una función de las señales  $X_1$ ,  $X_2$ , y  $\hat{U}_2$ .

25 **[0062]** En algunas realizaciones el proveedor de factor de escala 4 proporciona  $G_{E_x}$ , mediante lo cual se puede controlar la cantidad de energía de la primera señal de entrada  $X_1$  que contribuye a la señal de mezcla

descendente  $\hat{X}_D$ . Si el proceso de mezcla descendente debe ser conservador de energía (es decir, la señal de mezcla descendente contiene la misma cantidad de energía que la señal estéreo original) o al menos si el nivel de sonido percibido debe permanecer igual, se requiere procesamiento adicional. Se hace la siguiente consideración con la objeción de mantener constante el nivel de sonido percibido de las partes de señal individuales en la señal de mezcla descendente. En la realización preferida, se modifica la escala de la energía según una consideración de

mezcla descendente-energía óptima derivada. Se pueden considerar dos señales  $X_1^e$  y  $X_2^e$  y asumir que van a estar altamente correlacionadas como sería el caso, por ejemplo, para una fuente de toma panorámica de amplitud con

35  $\mathbf{E}\{X_1^e X_2^{e*}\} \neq 0$ . La señal  $X_2^e$  se puede expresar como  $X_2^e = a \cdot X_1^e$  de tal forma que la señal de mezcla

descendente  $X_D^c$  resulta en

$$\begin{aligned} X_D^c &= X_1^c + X_2^c \\ &= X_1^c + a \cdot X_1^c \\ &= (1 + a) \cdot X_1^c. \end{aligned} \tag{23}$$

[0063] La energía de  $X_D^c$  está dada por

$$E \left\{ |X_D^c|^2 \right\} = (1 + a)^2 \cdot E \left\{ |X_1^c|^2 \right\}. \tag{24}$$

5

[0064] Ahora se asume que las dos señales van a estar completamente no correlacionadas con

$E\{X_1^u X_2^{u*}\} = 0$ . La señal de mezcla descendente  $X_D^u$  resulta en

$$X_D^u = X_1^u + X_2^u. \tag{25}$$

[0065] La energía de  $X_D^u$  es dada por

$$\begin{aligned} E \left\{ |X_D^u|^2 \right\} &= E \left\{ |X_1^u|^2 \right\} + E \left\{ |X_2^u|^2 \right\} \\ &= E \left\{ |X_1^u|^2 \right\} + b \cdot E \left\{ |X_1^u|^2 \right\} \\ &= (1 + b) \cdot E \left\{ |X_1^u|^2 \right\}. \end{aligned} \tag{26}$$

10

[0066] A partir de estas consideraciones, se puede ver que la energía de una mezcla descendente óptima de las partes de señal correlacionadas resultaría en

$$E \left\{ |X_{D_o}^c|^2 \right\} = E \left\{ |X_1|^2 \right\} + E \left\{ |W X_1|^2 \right\}, \tag{27}$$

con  $W$  que corresponde a  $a$  en (23) y para las partes de señal no correlacionadas, se tiene que realizar una simple adición de la energía. La energía de mezcla descendente óptima final con respecto al modelo de

15 señal asumido y la señal de mezcla descendente deseada en (1) y (2) resultaría entonces en

$$\begin{aligned} E \left\{ |X_D^o|^2 \right\} &= E \left\{ |X_{D_o}^c|^2 \right\} + E \left\{ |U_2|^2 \right\} \\ &= E \left\{ |X_1|^2 \right\} + E \left\{ |W X_1|^2 \right\} + E \left\{ |U_2|^2 \right\}. \end{aligned} \tag{28}$$

[0067] A fin de asegurarse de que  $X_D^o$  y  $\tilde{X}_D$  contengan la misma cantidad de energía, se introdujeron los factores de modificación de escala de energía  $G_{E_x}$  y  $G_{E_u}$ , donde el último se proporciona por el proveedor de factor de escala U2. La señal de mezcla descendente real  $\tilde{X}_D$  se calcula como

$$\hat{X}_D = G_{E_x} \cdot X_1 + G_{E_u} \cdot \hat{U}_2. \quad (29)$$

**[0068]** Dada la energía de mezcla descendente óptima y  $G_{E_x}$ , ahora se puede derivar  $G_{E_x}$  del siguiente modo:

$$E \left\{ |X_D^o|^2 \right\} \stackrel{!}{=} E \left\{ |\hat{X}_D|^2 \right\} \quad (30)$$

$$\Phi_{X_1} + \Phi_{WX_1} + \Phi_{U_2} = G_{E_x}^2 \cdot \Phi_{X_1} + G_{E_u}^2 \cdot \Phi_{\hat{U}_2} \quad (31)$$

$$\begin{aligned} G_{E_x} &= \sqrt{\frac{\Phi_{X_1} + \Phi_{WX_1} + \Phi_{U_2} - G_{E_u}^2 \cdot \Phi_{\hat{U}_2}}{\Phi_{X_1}}} \\ &= \sqrt{1 + \frac{\Phi_{WX_1}}{\Phi_{X_1}} + \frac{\Phi_{U_2}}{\Phi_{X_1}} - G_{E_u}^2 \frac{\Phi_{\hat{U}_2}}{\Phi_{X_1}}} \end{aligned} \quad (32)$$

5

**[0069]** Con (12) la parte media de la ecuación (32) se identifica como

$$\frac{\Phi_{WX_1}}{\Phi_{X_1}} + \frac{\Phi_{U_2}}{\Phi_{X_1}} = \frac{\Phi_{X_2}}{\Phi_{X_1}}$$

por lo que se convierte en

$$G_{E_x} = \sqrt{1 + \frac{\Phi_{X_2}}{\Phi_{X_1}} - G_{E_u}^2 \frac{\Phi_{\hat{U}_2}}{\Phi_{X_1}}}. \quad (33)$$

10 **[0070]** Para someter a mezcla descendente múltiples canales de entrada  $X_1, X_2, X_3$ , se puede utilizar una cascada de múltiples etapas de mezcla descendente de dos canales 1. En la Figura 9, se muestra un ejemplo para tres señales de entrada  $X_1, X_2, X_3$ .

**[0071]** La señal de mezcla descendente final  $\hat{X}_{D_2}$  para un sistema de dos etapas resulta en

$$\begin{aligned} \hat{X}_{D_2} &= G_{E_{\hat{X}_{D_1}}} \hat{X}_{D_1} + G_{E_{U_3}} U_3 \\ &= G_{E_{\hat{X}_{D_1}}} (G_{E_{z_1}} X_1 + G_{E_{U_2}} U_2) + G_{E_{U_3}} U_3 \\ &= G_{E_{\hat{X}_{D_1}}} G_{E_{z_1}} X_1 + G_{E_{\hat{X}_{D_1}}} G_{E_{U_2}} U_2 + G_{E_{U_3}} U_3 \end{aligned} \quad (34)$$

15 **[0072]** Las características clave de una realización de la invención son:

- Considerar  $X_1$  como una señal de referencia y considerar  $X_2$  como una mezcla de una versión filtrada de  $X_1$  y, por tanto, una parte de señal correlacionada  $WX_1$  y una parte de señal no correlacionada  $U_2$  con respecto a  $X_1$ .

20 • Separación/Descomposición de  $X_2$  en sus dos componentes de señal mencionados anteriormente. Extracción de disparidad de  $X_1$  y  $X_2$  mediante

- estimación de la similitud de  $X_1$  y  $X_2$ , que resulta en un coeficiente de filtro  $W$  y
- reducción de similitud por supresión de partes de señal correlacionadas o una combinación de ambas, que resulta en una parte de señal no correlacionada estimada  $\hat{U}_2$ .

25

- Modificación de escala de energía de  $X_1$  para satisfacer un nivel de energía predefinido.

- Modificación de escala de energía de  $\hat{U}_2$ .
- Suma de las señales a escala de energía para formar la señal de mezcla descendente deseada  $X_D$ .
- Procesamiento en bandas de frecuencia.

5 **[0073]** Las características de implementación opcionales son:

- Supresión inversa de fase alineada o cancelación inversa de fase alineada.
- Cascada de dos o más bloques de mezcla descendente para realizar una mezcla descendente multi-canal.
- Solamente supresión inversa de fase alineada parcialmente aplicada.

10

**[0074]** Aunque se han descrito algunos aspectos en el contexto de un aparato, es claro que estos aspectos también representan una descripción del procedimiento correspondiente, donde un bloque o dispositivo corresponde a un paso de procedimiento o una característica de un paso de procedimiento. De forma análoga, aspectos descritos en el contexto de un paso de procedimiento también representan una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente.

15

**[0075]** Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención se pueden implementar en hardware o software. La implementación se puede realizar utilizando un medio de almacenamiento no transitorio tal como un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM, o una memoria FLASH, que tiene señales de control legibles de forma electrónica almacenadas en la misma, que cooperan (o son capaces de cooperar) con un sistema informático programable de tal forma que se realiza el procedimiento respectivo. Por tanto, el medio de almacenamiento digital puede ser legible por ordenador.

20

**[0076]** Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control legibles de forma electrónica, que son capaces de cooperar con un sistema informático programable, de tal forma que se realiza uno de los procedimientos descritos en esta invención.

25

**[0077]** En general, las realizaciones de la presente invención se pueden implementar como un producto de programa informático con un código de programa, el código de programa que es operativo para realizar uno de los procedimientos cuando el producto de programa informático se ejecuta en un equipo. El código de programa se puede almacenar, por ejemplo, en un soporte legible por máquina.

30

**[0078]** Otras realizaciones comprenden el programa informático para realizar uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por máquina.

35

**[0079]** En otras palabras, una realización del procedimiento inventivo es, por tanto, un programa informático que tiene un código de programa para realizar uno de los procedimientos descritos en esta invención, cuando el programa informático se ejecuta en un equipo.

40

**[0080]** Una realización adicional del procedimiento inventivo es, por tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para realizar uno de los procedimientos descritos en esta invención. El soporte de datos, el medio de almacenamiento digital o el medio grabado son convencionalmente tangibles y/o no transitorios.

45

**[0081]** Una realización adicional del procedimiento de la invención es, por tanto, un flujo de datos o una secuencia de señales que representan el programa informático para realizar uno de los procedimientos descritos en esta invención. El flujo de datos o la secuencia de señales se pueden configurar, por ejemplo, para que se transfieran mediante una conexión de comunicación de datos, por ejemplo, a través de Internet.

50

**[0082]** Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador o un dispositivo lógico programable, configurado para, o adaptado para, realizar uno de los procedimientos descritos en esta invención.

55

**[0083]** Una realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para realizar uno de los procedimientos descritos en esta invención.

**[0084]** Una realización adicional según la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, de forma electrónica u óptica) un programa informático para realizar uno de los procedimientos descritos en esta invención a un receptor. El receptor puede ser, por ejemplo, un ordenador, un dispositivo móvil, un dispositivo de memoria o similares. El aparato o sistema, por ejemplo, puede comprender un servidor de archivos para transferir el programa informático al receptor.

**[0085]** En algunas realizaciones, se puede utilizar un dispositivo lógico programable (por ejemplo, una matriz de puertas programables en el campo) para realizar algunas o todas las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones, una matriz de puertas programables en el campo puede cooperar con un microprocesador a fin de realizar uno de los procedimientos descritos en esta invención. En general, los procedimientos se realizan preferentemente mediante cualquier aparato de hardware.

**[0086]** Las realizaciones descritas anteriormente son simplemente ilustrativas de los principios de la presente invención. Se entiende que serán evidentes modificaciones y variaciones de las disposiciones y detalles descritos en esta invención para otros expertos en la técnica. Por tanto, se propone que sólo se limite por el alcance de las reivindicaciones de patente inminentes y no por los detalles específicos presentados a título de descripción y explicación de las realizaciones en esta invención.

Signos de Referencia:

20

**[0087]**

1	dispositivo de procesamiento de señal de audio
2	extractor de disparidad
3	combinador
4	primer dispositivo modificador de escala de energía
5	primer proveedor de factor de escala
6	segundo dispositivo modificador de escala de energía
7	segundo proveedor de factor de escala
8	dispositivo sumador
9	estimador de similitud
10	reductor de similitud
10a	etapa de cancelación
10a'	etapa de cancelación
10b	etapa de supresión
10b'	etapa de supresión
11	dispositivo de filtro complejo
11'	dispositivo de filtro absoluto
12	dispositivo de cancelación de señal
13	dispositivo de cambio de fase
14	dispositivo de supresión
15	dispositivo de cambio de fase
16	dispositivo de ponderación
$X_1$	primera señal de entrada
$X_2$	segunda señal de entrada
	señal de mezcla descendente
$\hat{X}_D$	
$\hat{U}_2$	señal extraída
$G_{E_x}$	primer factor de escala
$X_{1s}$	una primera señal de entrada a escala
$W$	coeficientes de filtro
$WX_1$	partes de señal de la primera señal de entrada que están presentes en la segunda señal de entrada ( $X_2$ )
$X'_2$	señal derivada de la segunda señal de entrada
$\gamma$	factor de ponderación
$\gamma WX_1$	partes de señal ponderadas de la primera señal de entrada que están presentes en la segunda señal de entrada ( $X_2$ )

Referencias:**[0088]**

- 5 [1] ITU-R BS.775-2, "Multichannel Stereophonic Sound System With And Without Accompanying Picture," 07/2006.
- [2] R. Dressler, (05.08.2004) Dolby Surround Pro Logic II Decoder Principles of Operation. Disponible [En línea]: [http://www.dolby.com/uploadedFiles/Assets/US/Doc/Professional/209\\_Dolby\\_Surround\\_Pro\\_Logic\\_II\\_Decoder\\_Principles\\_of\\_Operation.pdf](http://www.dolby.com/uploadedFiles/Assets/US/Doc/Professional/209_Dolby_Surround_Pro_Logic_II_Decoder_Principles_of_Operation.pdf).
- 10 [3] K. Lopatka, B. Kunka, y A. Czyzewski, "Novel 5.1 Downmix Algorithm with Improved Dialogue Intelligibility," en la 134-ta Convención de la AES, 2013.
- [4] J. Breebaart, K. S. Chong, S. Disch, C. Faller, J. Herre, J. Hilpert, K. Kjörling, J. Koppens, K. Linzmeier, W. Oomen, H. Purnhagen, y J. Rödén, "MPEG Surround – the ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding," *J. Audio Eng. Soc.*, vol. 56, no. 11, pp. 932–955, 2007.
- 15 [5] M. Neuendorf, M. Multrus, N. Rellerbach, R. J. Fuchs Guillaume, J. Lecomte, Wilde Stefan, S. Bayer, S. Disch, C. Helmrich, R. Lefebvre, P. Gournay, B. Bessette, J. Lapierre, K. Kjörling, H. Purnhagen, L. Villemoes, W. Oomen, E. Schuijers, K. Kikuri, T. Chinen, T. Norimatsu, C. K. Seng, E. Oh, M. Kim, S. Quackenbush, y B. Grill, "MPEG Unified Speech and Audio Coding - The ISO/MPEG Standard for High-Efficiency Audio Coding of all Content Types," *J. Audio Eng. Soc.*, vol. 132-da Convención, 2012.
- 20 [6] C. Faller y F. Baumgarte, "Binaural Cue Coding-Part II: Schemes and Applications," *Speech and Audio Processing*, *IEEE Transactions en*, vol. 11, no. 6, pp. 520–531, 2003.
- [7] F. Baumgarte, "Equalization for Audio Mixing," *Patente US 7,039,204 B2*, 2003.
- [8] J. Thompson, A. Warner, y B. Smith, "An Active Multichannel Downmix Enhancement for Minimizing Spatial and Spectral Distortions", en la 127-ma Convención de la AES, Octubre 2009.
- 25 [9] G. Stoll, J. Groh, M. Link, J. Deigmöller, B. Runow, M. Keil, R. Stoll, M. Stoll, y C. Stoll, "Method for Generating a Downward-Compatible Sound Format," *Patente de Estados Unidos US2012/0 014 526*, 2012.
- [10] B. Runow y J. Deigmöller, "Optimierter Stereo-Dowmix von 5.1-Mehrkanalproduktionen: An optimized Stereo-Downmix of a 5.1 multichannel audio production," en 25. Convención Internacional Tonmeistertagung-VDT, 2008.
- [11] Samsudin, E. Kurniawati, Ng Boon Poh, F. Sattar, and S. George, "A Stereo to Mono Downmixing Scheme for MPEG-4 Parametric Stereo Encoder," in *Acoustics, Speech and Signal Processing*, 2006. *ICASSP 2006 Proceedings. 2006 IEEE International Conference en*, vol. 5, 2006, p. V. 2.
- 30 [12] M. Kim, E. Oh, y H. Shim, "Stereo audio coding improved by phase parameters," en la 129-na Convención de la AES, 2010.
- [13] W. Wu, L. Miao, Y. Lang, y D. Virette, "Parametric Stereo Coding Scheme with a New Downmix Method and Whole Band Inter Channel Time/Phase Differences," *Acoustics, Speech and Signal Processing*, *IEEE Transactions en*, pp. 556–560, 2013.
- 35 [14] Der-Pei Chen ET AL: "Gram-Schmidt-based Downmixer and Decorrelator in the MPEG Surround Coding", 128-va AES CONVENTION, Convention Paper 8067, 22 de mayo de 2010 (2010-05-22).

**REIVINDICACIONES**

1. Un dispositivo de procesamiento de señales de audio (1) para la mezcla descendente de una primera señal de entrada ( $X_1$ ) y una segunda señal de entrada ( $X_2$ ) a una señal de mezcla descendente ( $\tilde{X}_D$ ), en el que la primera señal de entrada ( $X_1$ ) y la segunda señal de entrada ( $X_2$ ) están al menos parcialmente correlacionadas, que comprende:
  - un extractor de disparidad (2) configurado para recibir la primera señal de entrada ( $X_1$ ) y la segunda señal de entrada ( $X_2$ ) así como para sacar una señal extraída ( $\hat{U}_2$ ), que está menos correlacionada con respecto a la primera señal de entrada ( $X_1$ ) que la segunda señal de entrada ( $X_2$ ) y
  - un combinador (3) configurado para combinar la primera señal de entrada ( $X_1$ ) y la señal extraída ( $\hat{U}_2$ ) a fin de obtener la señal de mezcla descendente ( $\tilde{X}_D$ ), en el que el extractor de disparidad (2) comprende un estimador de similitud (9) configurado para proporcionar coeficientes de filtro ( $W$ ,  $|W|$ ) para obtener partes de señal ( $WX1$ ,  $|WX1|$ ) de la primera señal de entrada ( $X_1$ ) que están presentes en la segunda señal de entrada ( $X_2$ ) de la primera señal de entrada ( $X_1$ ), en el que el extractor de disparidad (2) comprende un reductor de similitud (10) configurado para reducir las partes de señal obtenidas ( $WX1$ ,  $|WX1|$ ) de la primera señal de entrada que están presentes en la segunda señal de entrada ( $X_2$ ) con base en los coeficientes de filtro ( $W$ ,  $|W|$ ), en el que el reductor de similitud (10) comprende una etapa de supresión de señal (10b, 10b') que tiene un dispositivo de supresión de señales (14) configurado para multiplicar la segunda señal de entrada ( $X_2$ ) o una señal ( $X'_2$ ) derivada de la segunda señal de entrada ( $X_2$ ) con un factor de ganancia de supresión ( $G$ ) a fin de obtener la señal extraída ( $\hat{U}_2$ ), en el que el factor de ganancia de supresión ( $G$ ) se elige de tal forma que se reduce al mínimo un error cuadrático medio entre la señal extraída ( $\hat{U}_2$ ) y una parte de señal ( $U_2$ ) de la segunda señal de entrada ( $X_2$ ), que no está correlacionada con la primera señal de entrada ( $X_1$ ).
2. Un dispositivo según la reivindicación anterior, en el que el combinador (3) comprende un sistema modificador de escala de energía (4, 5, 6, 7) configurado de tal forma que la relación de la energía de la mezcla descendente ( $\tilde{X}_D$ ) y las energías sumadas de la primera señal de entrada ( $X_1$ ) y la segunda señal de entrada ( $X_2$ ) es independiente de la correlación de la primera señal de entrada ( $X_1$ ) y la segunda señal de entrada ( $X_2$ ).
3. Un dispositivo según la reivindicación anterior, en el que el sistema modificador de escala de energía (4, 5, 6, 7) comprende un primer dispositivo modificador de escala de energía (4) configurado para modificar la escala de la primera señal de entrada ( $X_1$ ) con base en un primer factor de escala ( $G_{E_x}$ ) a fin de obtener una señal de entrada a escala ( $X_{1s}$ ).
4. Un dispositivo según la reivindicación anterior, en el que el sistema modificador de escala de energía (4, 5, 6, 7) comprende un primer proveedor de factor de escala (5) configurado para proporcionar el primer factor de escala ( $G_{E_x}$ ), en el que el primer proveedor de factor de escala (5) se diseña preferentemente como un procesador (5) configurado para calcular el primer factor de escala ( $G_{E_x}$ ) dependiendo de la primera señal de entrada ( $X_1$ ), la segunda señal de entrada ( $X_2$ ) y/o la señal extraída ( $\hat{U}_2$ ).
5. Un dispositivo según una de las reivindicaciones 2 a 4, en el que el sistema modificador de escala de energía (4, 5, 6, 7) comprende un segundo dispositivo modificador de escala de energía (6) configurado para modificar la escala de la señal extraída ( $\hat{U}_2$ ) con base en un segundo factor de escala ( $G_{E_u}$ ) a fin de obtener una señal extraída a escala ( $\hat{U}_{2s}$ ).
6. Un dispositivo según la reivindicación anterior, en el que el sistema modificador de escala de energía



(4, 5, 6, 7) comprende un segundo proveedor de factor de escala (7) configurado para proporcionar el segundo factor de escala ( $G_{E_u}$ ), en el que el segundo proveedor de factor de escala (7) se diseña preferentemente como una interfaz de hombre-máquina configurada para introducir de forma manual el segundo factor de escala ( $G_{E_u}$ ).

5 7. Un dispositivo según una de las reivindicaciones anteriores, en el que el combinador (3) comprende un dispositivo sumador (8) para sacar la señal de mezcla descendente ( $\tilde{X}_D$ ) con base en la primera señal de entrada ( $X_1$ ) y con base en la señal extraída ( $\hat{U}_2$ ).

8. Un dispositivo según una de las reivindicaciones anteriores, en el que el reductor de similitud (10) comprende una etapa de cancelación (10a, 10a') que tiene un dispositivo de cancelación de señal (12) configurado para sustraer las partes de señal obtenidas ( $WX1, |WX1|$ ) de la primera señal de entrada ( $X_1$ ) que están presentes en la segunda señal de entrada ( $X_2$ ) o una señal ( $\gamma WX1$ ) derivada de las partes de señal obtenidas ( $WX1, |WX1|$ ) de la segunda señal de entrada ( $X_2$ ) o de una señal ( $X'_2$ ) derivada de la segunda señal de entrada ( $X_2$ ).

15 9. Un dispositivo según la reivindicación 8, en el que la etapa de cancelación (10a) comprende un dispositivo de filtro complejo (11) configurado para filtrar la primera señal de entrada ( $X_1$ ) al utilizar coeficientes de filtro de valores complejos  $W$ .

10. Un dispositivo según la reivindicación 8 ó 9, en el que la etapa de cancelación (10a') comprende un dispositivo de cambio de fase (13) configurado para alinear la fase de la segunda señal de entrada ( $X_2$ ) con la fase de la primera señal de entrada ( $X_1$ ).

11. Un dispositivo según una de las reivindicaciones 8 a 10, en el que se alimenta una señal de salida ( $\hat{U}'_2$ ) de la etapa de cancelación (10a) a una entrada de la etapa de supresión de señal (10b) a fin de obtener la  
25 señal extraída ( $\hat{U}_2$ ), o en el que se alimenta una señal de salida de la etapa de supresión de señal (10b) a una entrada de la etapa de cancelación (10a) a fin de obtener la señal extraída ( $\hat{U}_2$ ).

12. Un dispositivo según la reivindicación anterior, en el que la etapa de cancelación (10a) comprende un dispositivo de ponderación (16) configurado para ponderar las partes de señal obtenidas ( $WX1, |WX1|$ ) de la primera  
30 señal de entrada ( $X_1$ ) que están presentes en la segunda señal de entrada ( $X_2$ ) dependiendo de un factor de ponderación ( $\gamma$ ).

13. Un dispositivo según una de las reivindicaciones anteriores, en el que la etapa de supresión de señal (10b') comprende un dispositivo de cambio de fase (15) configurado para alinear la fase de la segunda señal de  
35 entrada ( $X_2$ ) con la fase de la primera señal de entrada ( $X_1$ ).

14. Un dispositivo según la reivindicación 10 y 12, en el que el dispositivo de cambio de fase (13) se configura para alinear la fase de la segunda señal de entrada ( $X_2$ ) con la fase de la primera señal de entrada ( $X_1$ ) dependiendo del factor de ponderación ( $\gamma$ ).

40 15. Un dispositivo según la reivindicación anterior, en el que el dispositivo de cambio de fase (13) se configura para alinear la fase de la segunda señal de entrada ( $X_2$ ) con la fase de la primera señal de entrada ( $X_1$ ) solamente, si el factor de ponderación ( $\gamma$ ) es menor o igual a un umbral predefinido ( $\Gamma$ ).

45 16. Un dispositivo de procesamiento de señales de audio para la mezcla descendente de una pluralidad de señales de entrada ( $X_1, X_2, X_3$ ) a una señal de mezcla descendente ( $\tilde{X}_{D2}$ ) que comprende al menos un primer dispositivo (1) según una de las reivindicaciones anteriores y un segundo dispositivo (1') según una de las reivindicaciones anteriores, en el que la señal de mezcla descendente ( $\tilde{X}_{D1}$ ) del primer dispositivo se alimenta al  
segundo dispositivo como una primera señal de entrada ( $\tilde{X}_{D1}$ ) o como una segunda señal de entrada.

50

17. Un procedimiento de procesamiento de señal de audio para la mezcla descendente de una primera señal de entrada ( $X_1$ ) y una segunda señal de entrada ( $X_2$ ) a una señal de mezcla descendente ( $\tilde{X}_D$ ) que comprende los pasos de:

extraer una señal extraída ( $\hat{U}_2$ ) de la segunda señal de entrada ( $X_2$ ), en la que la señal extraída ( $\hat{U}_2$ ) está menos correlacionada con respecto a la primera señal de entrada ( $X_1$ ) que la segunda señal de entrada ( $X_2$ )

sumar la primera señal de entrada ( $X_1$ ) y la señal extraída ( $\hat{U}_2$ ) a fin de obtener la señal de mezcla descendente ( $\tilde{X}_D$ )

proporcionar coeficientes de filtro ( $W, |W|$ ) para obtener partes de señal ( $WX_1, |WX_1|$ ) de la primera señal de entrada ( $X_1$ ) que están presentes en la segunda señal de entrada ( $X_2$ ) de la primera señal de entrada ( $X_1$ ),

10 reducir las partes de señal obtenidas ( $WX_1, |WX_1|$ ) de la primera señal de entrada que están presentes en la segunda señal de entrada ( $X_2$ ) con base en los coeficientes de filtro ( $W, |W|$ ),

multiplicar la segunda señal de entrada ( $X_2$ ) o una señal ( $X'_2$ ) derivada de la segunda señal de entrada ( $X_2$ ) con un factor de ganancia de supresión ( $G$ ) a fin de obtener la señal extraída ( $\hat{U}_2$ ),

en el que el factor de ganancia de supresión ( $G$ ) se elige de tal forma que se reduce al mínimo un error cuadrático

15 medio entre la señal extraída ( $\hat{U}_2$ ) y una parte de señal ( $U_2$ ) de la segunda señal de entrada ( $X_2$ ), cuya parte de señal ( $U_2$ ) no está correlacionada con la primera señal de entrada ( $X_1$ ).

18. Un programa informático para implementar el procedimiento de procesamiento de señal de audio de la reivindicación 17 cuando se ejecuta en un ordenador o procesador de señales.

20

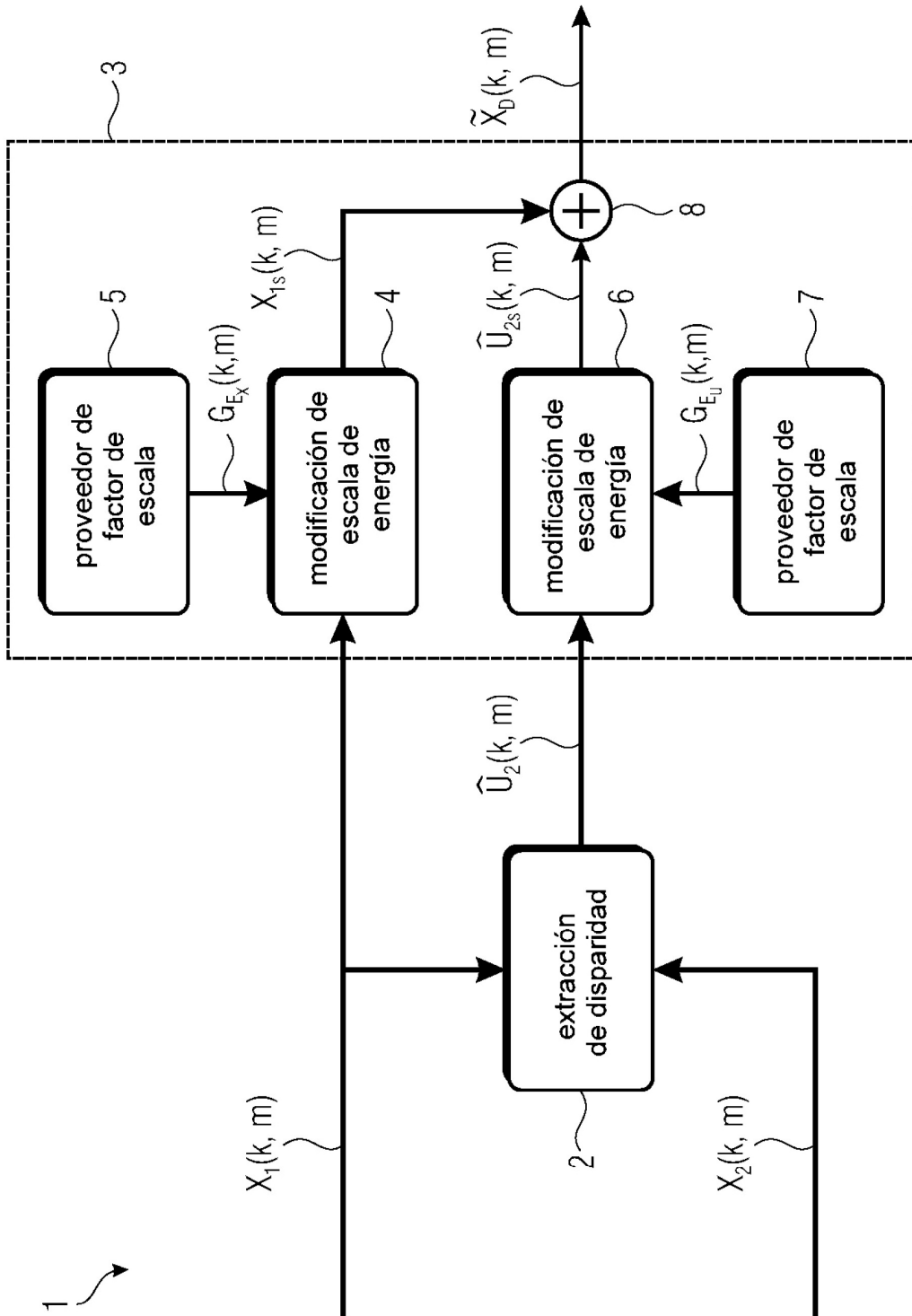


FIG 1

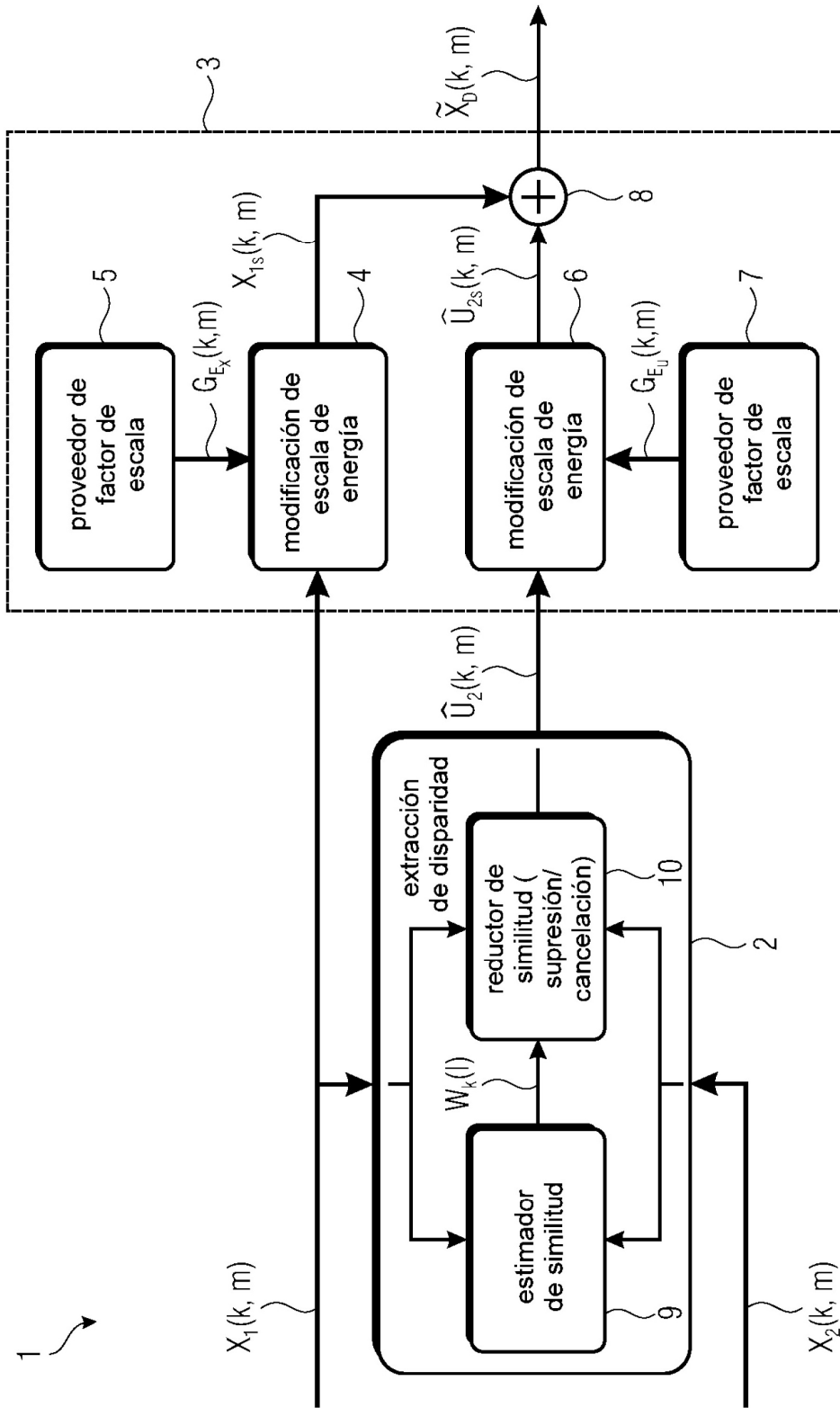


FIG 2

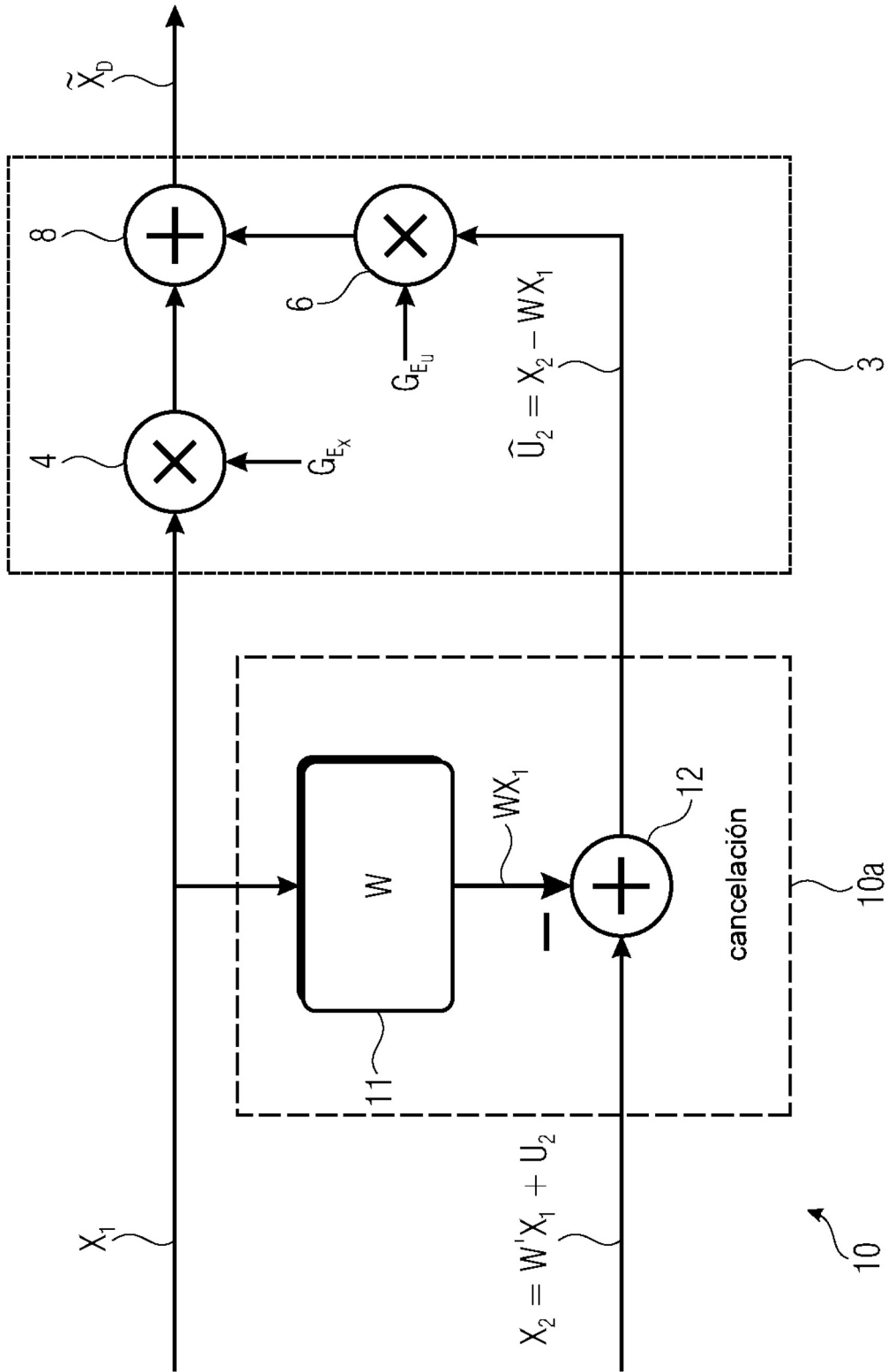


FIG 3

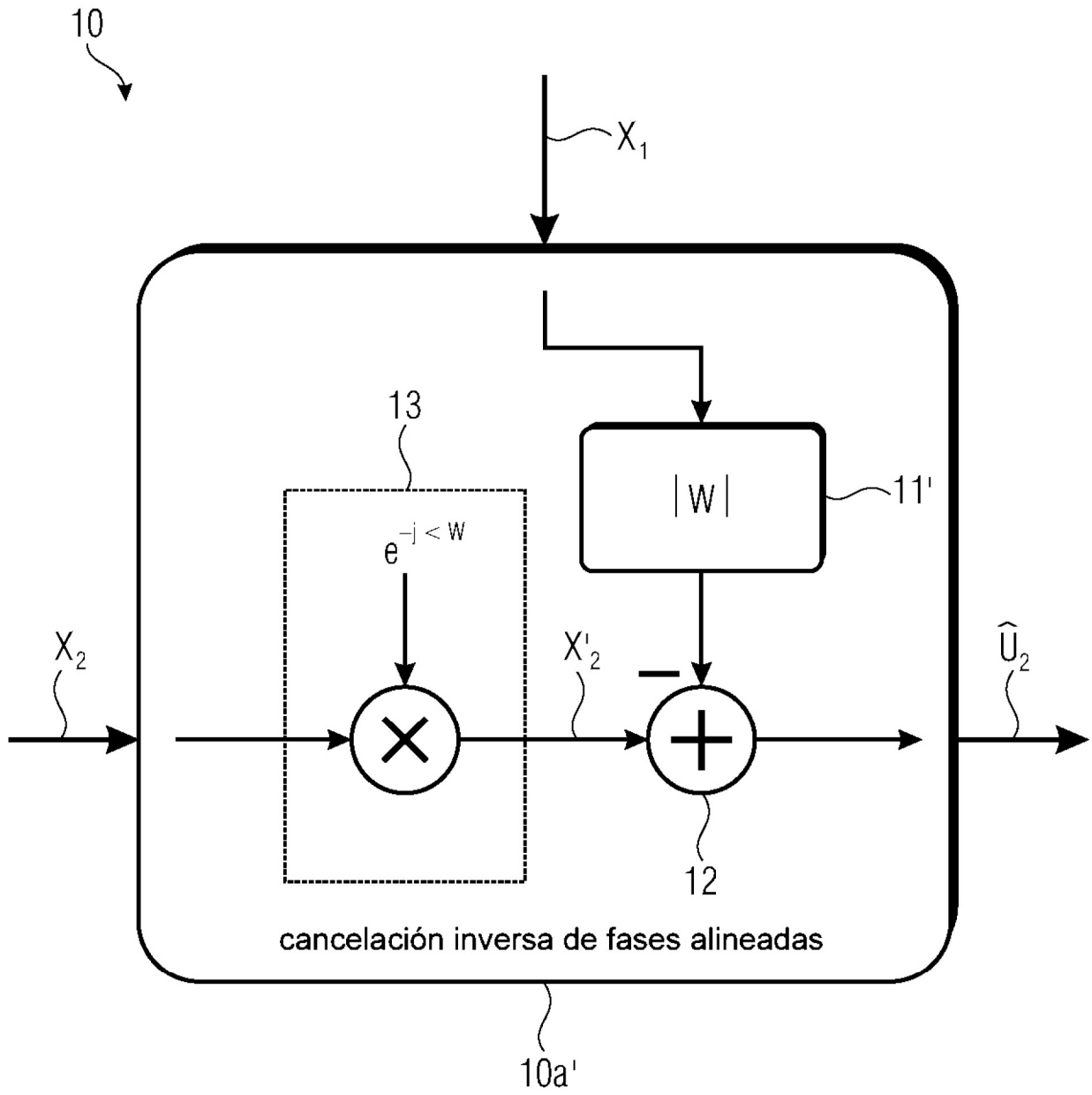


FIG 4

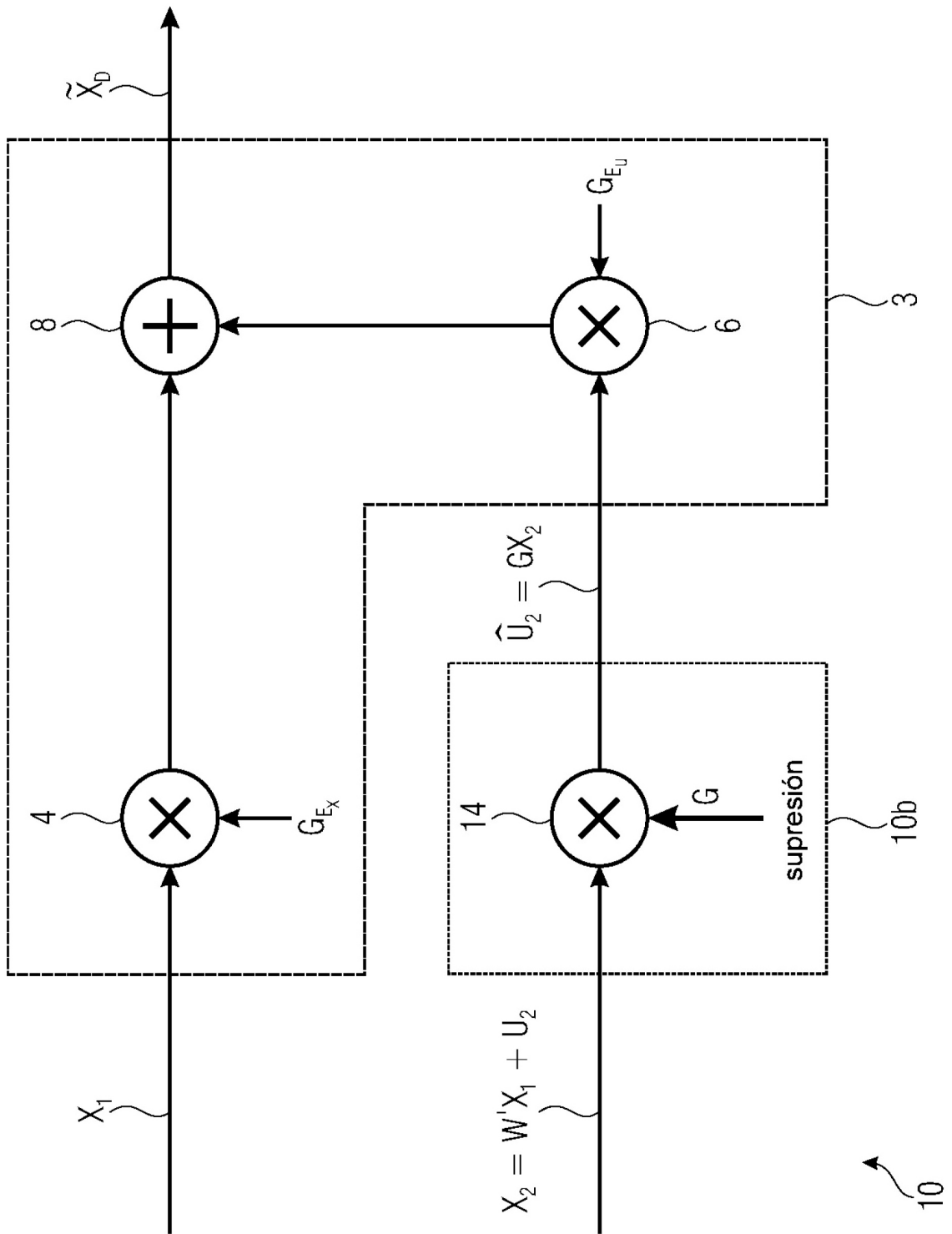


FIG 5

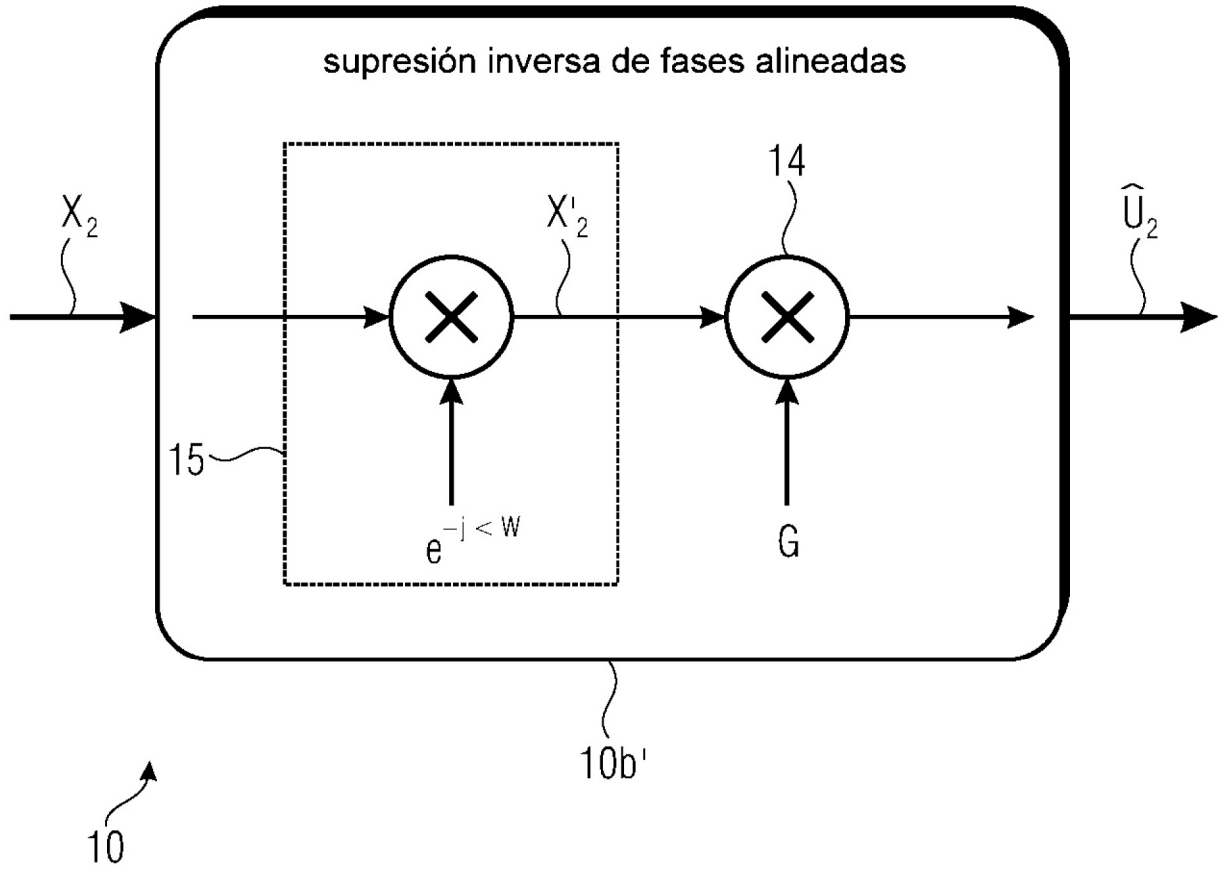


FIG 6



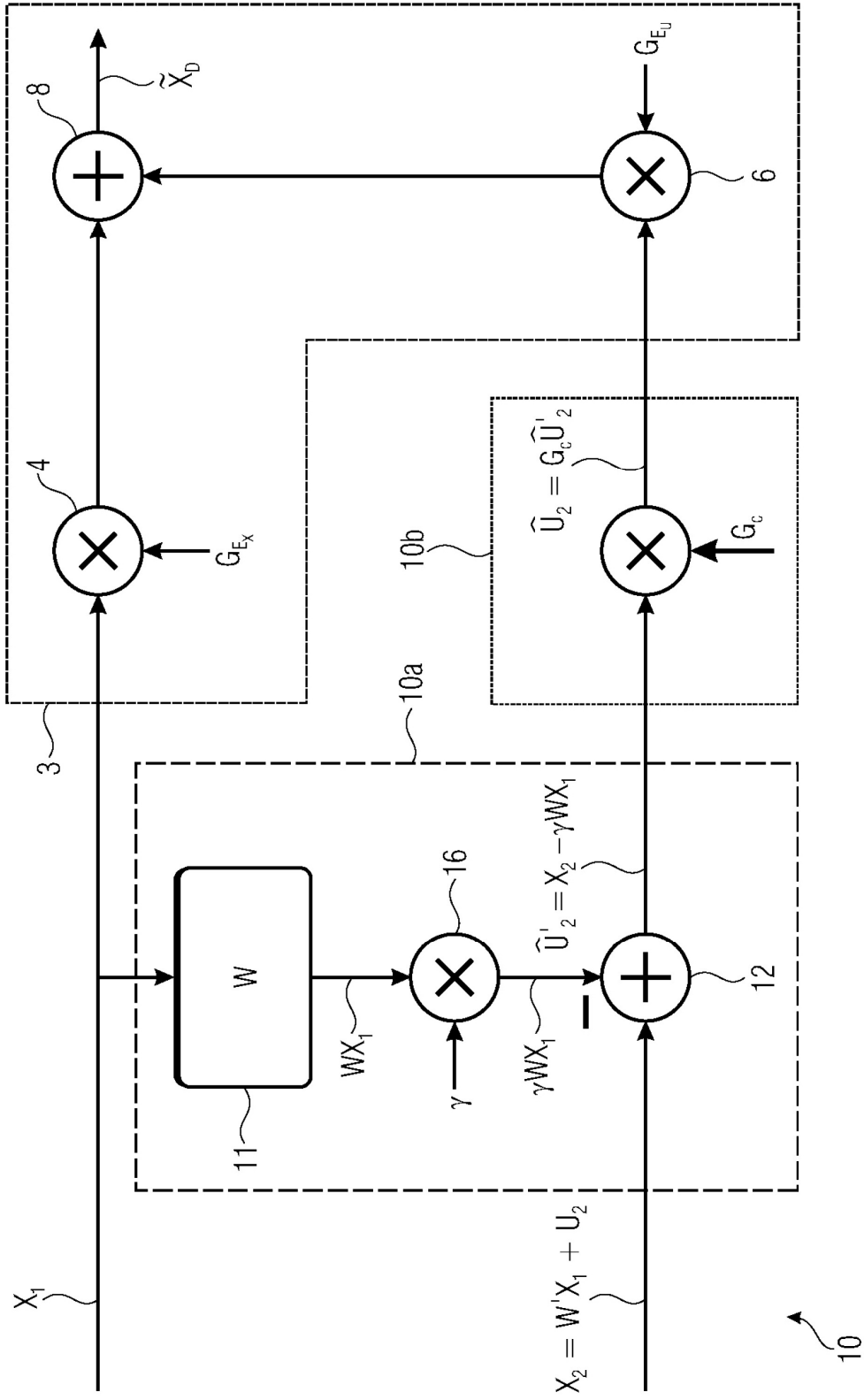


FIG 7

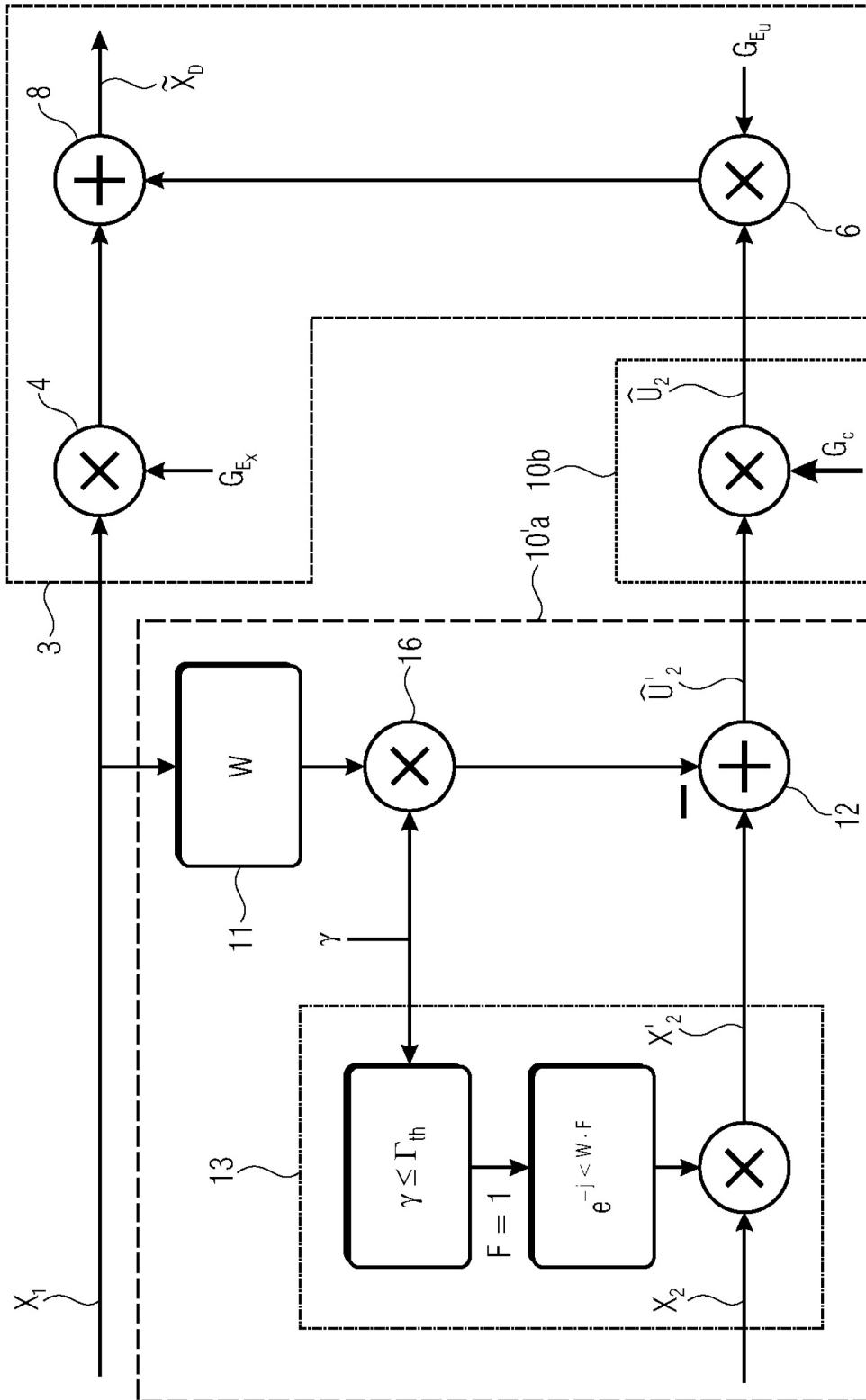


FIG 8

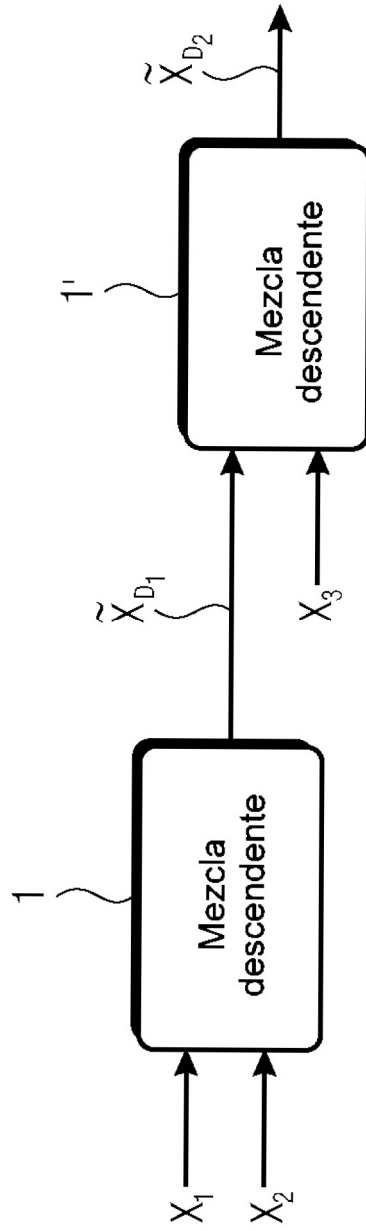


FIG 9