

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 706 490**

51 Int. Cl.:

G10L 19/00 (2013.01)

G10L 19/008 (2013.01)

G10L 19/025 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.07.2011 E 16196394 (7)**

97 Fecha y número de publicación de la concesión europea: **07.11.2018 EP 3144932**

54 Título: **Un aparato para la codificación de una señal de audio que tiene una pluralidad de canales**

30 Prioridad:

25.08.2010 US 376980 P

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

29.03.2019

73 Titular/es:

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

72 Inventor/es:

**KUNTZ, ACHIM;
DISCH, SASCHA;
HERRE, JÜRGEN;
KÜCH, FABIAN y
HILPERT, JOHANNES**

74 Agente/Representante:

SALVÀ FERRER, Joan

ES 2 706 490 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

Un aparato para la codificación de una señal de audio que tiene una pluralidad de canales

5 **[0001]** La presente invención se refiere al campo de procesamiento de audio y decodificación de audio, en particular a la decodificación de una señal que comprende componentes transitorios.

[0002] El procesamiento y/o la decodificación de audio ha avanzado de muchas maneras. En particular, se hacen cada vez más importantes las aplicaciones de audio espacial. Con frecuencia se usa procesamiento de señal
10 de audio para decorrelacionar o renderizar señales. Además, la decorrelación y la renderización de señales se emplean en el proceso de mezclado ascendente mono a estéreo, mezclado ascendente mono/estéreo a multicanal, reverberación artificial, ampliación de estéreo o mezclado/renderización interactiva con el usuario.

[0003] Diversos sistemas de procesamiento de señal de audio emplean decorrelacionadores. Un ejemplo
15 importante es la aplicación de sistemas decorrelacionadores en decodificadores de audio espacial paramétricos para recuperar propiedades de decorrelación específicas entre dos o más señales que son reconstruidas a partir de una o varias señales de mezcla descendente. La aplicación de decorrelacionadores mejora significativamente la calidad perceptual de la señal de salida, por ejemplo, cuando se compara con estéreo de intensidad. Específicamente, el uso de decorrelacionadores permite la correcta síntesis de sonido espacial con una amplia imagen de sonido, varios
20 objetos de sonido concurrentes y/o atmósfera. Sin embargo, también se sabe que los decorrelacionadores introducen artefactos como cambios en la estructura temporal de la señal, timbre, etc.

[0004] Otros ejemplos de aplicación de decorrelacionadores en procesamiento de audio son, por ejemplo, la generación de reverberación artificial para cambiar la impresión espacial o el uso de decorrelacionadores en
25 sistemas de cancelación de eco acústico multicanal para mejorar el comportamiento de convergencia.

[0005] En la Figura 1 se ilustra un estado típico de una aplicación de la técnica de un decorrelacionador en un mezclador ascendente mono a estéreo, por ejemplo, aplicado en Estéreo Paramétrico (PS), donde se proporciona una señal de entrada mono M (una señal "seca") a un decorrelacionador 110. El decorrelacionador 110
30 decorrelaciona la señal de entrada mono M según un procedimiento de decorrelación para proporcionar una señal decorrelacionada D (una señal "mojada") en su salida. La señal decorrelacionada D es alimentada a un mezclador 120 como una primera señal de entrada de mezclador junto con la señal mono seca M como una segunda señal de entrada de mezclador. Además, una unidad de control de mezcla ascendente 130 alimenta parámetros de control de mezcla ascendente al mezclador 120. El mezclador 120 genera entonces dos canales de salida L y R (L = canal de salida
35 estéreo izquierdo; R = canal de salida estéreo derecho) según una matriz de mezclado H. Los coeficientes de la matriz de mezclado pueden ser fijos, dependientes de la señal o controlados por un usuario.

[0006] Alternativamente, la matriz de mezclado es controlada por información lateral que es transmitida junto con el mezclado descendente que contiene una descripción paramétrica sobre cómo mezclar de forma ascendente
40 las señales del mezclado descendente para formar la salida multicanal deseada. Esta información lateral espacial usualmente es generada durante el proceso de mezclado descendente mono en un codificador de señal compatible.

[0007] Este principio se aplica ampliamente en codificación de audio espacial, por ejemplo, Estéreo Paramétrico, véase, por ejemplo, de J. Breebaart, S. van de Par, A. Kohlrausch, E. Schuijers, "Codificación de Audio
45 Espacial Paramétrico de Alta Calidad a baja cantidad de bits transmitidos" ("High-Quality Parametric Spatial Audio Coding at Low Bitrates") en Actas de la 116ª Convención de AES, Berlín, Preimpresión 6072, mayo de 2004.

[0008] Un resumen de avances recientes en la codificación de audio multicanal se presenta en BREEBAART J ET AL: "MPEG spatial audio coding / MPEG surround: Overview and current status" (codificación de audio espacial
50 de MPEG/envolvente de MPEG: resumen y estado actual), AUDIO ENGINEERING SOCIETY CONVENTION PAPER (Convención de la sociedad de ingenieros de sonido), NUEVA YORK, NY, EE.UU., 7 de octubre de 2005, páginas 1-15. Una estructura del estado actual de la técnica típica adicional, de un decodificador estéreo paramétrico está ilustrada en la Figura 2, en la que el proceso de decorrelación se realiza en el dominio de la transformación. Un banco de filtros de análisis 210 transforma una señal de entrada mono a un dominio de la
55 transformación, por ejemplo, al dominio de la frecuencia. La decorrelación de la señal de entrada mono transformada M es realizada a continuación por un decorrelacionador 220 el cual genera una señal decorrelacionada D. Tanto la señal de entrada mono transformada M como la señal decorrelacionada D, son alimentadas a una matriz de mezclado 230. La matriz de mezclado 230 genera entonces dos señales de salida L y R tomando en cuenta parámetros de mezclado ascendente, las cuales son proporcionadas por la unidad de modificación de parámetros
60 240, la cual está provista de parámetros espaciales y está acoplada a una unidad de control de parámetros 250. En

la Figura 2, los parámetros espaciales pueden ser modificados por un usuario o herramientas adicionales, por ejemplo, post-procesamiento para renderización/presentación binaural. En este ejemplo, los parámetros de mezclado ascendente son combinados con los parámetros procedentes de los filtros binaurales para formar los parámetros de entrada para la matriz de mezclado ascendente. Finalmente, las señales de salida generadas por la matriz de mezclado 230 son alimentadas al banco de filtros de síntesis 260, el cual determina la señal de salida estéreo.

[0009] La salida L/R de la matriz de mezclado 230 es computada a partir de la señal de entrada modo M y la señal decorrelacionada D según una regla de mezclado, por ejemplo, aplicando la siguiente fórmula:

$$\begin{bmatrix} L \\ R \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} M \\ D \end{bmatrix}$$

[0010] En la matriz de mezclado, la magnitud del sonido decorrelacionado alimentado a la salida es controlada sobre la base de los parámetros transmitidos, por ejemplo, Correlación/Coherencia Inter-Canal (ICC) y/o configuraciones fijas o definidas por el usuario.

[0011] Conceptualmente, la señal de salida de la salida del decorrelacionador D reemplaza una señal residual que idealmente permitiría una decodificación perfecta de las señales L/R originales. Utilizar la salida de decorrelacionador D en lugar de una señal residual en el mezclador ascendente da por resultado un ahorro de cantidad de bits transmitidos que de otra manera se hubiera requerido para transmitir la señal residual. El objetivo del decorrelacionador es de ese modo generar una señal D a partir de la señal mono M, la cual exhiba propiedades similares a la señal residual que es reemplazada por D.

[0012] En consecuencia, del lado del codificador, se extraen dos tipos de parámetros espaciales: Un primer grupo de parámetros comprende parámetros de correlación/coherencia (por ejemplo, ICCs = parámetros de Correlación/Coherencia Inter-Canal) que representan la coherencia o la correlación cruzada entre dos canales de entrada que serán conectados. Un segundo grupo de parámetros comprende parámetros de diferencia de nivel (por ejemplo, ILDs = parámetros de Diferencia de Nivel Inter-Canal) que representan la diferencia de nivel entre los dos canales de entrada.

[0013] Además, una señal de mezcla descendente es generada mezclando de forma descendente los dos canales de entrada. Asimismo, se genera una señal residual. Las señales residuales son señales que pueden ser usadas para regenerar las señales originales empleando adicionalmente la señal de mezcla descendente y una matriz de mezclado ascendente. Por ejemplo, cuando N señales son mezcladas de forma descendente a 1 señal, el mezclado descendente típicamente es 1 de los N componentes que resultan del mapeo de las N señales de entrada. Los componentes restantes que resultan del mapeo (por ejemplo, N-1 componentes) son señales residuales y permiten reconstruir las N señales originales mediante un mapeo inverso. El mapeo puede ser, por ejemplo, una rotación. El mapeo debe ser conducido de tal forma que la señal de mezcla descendente sea maximizada y las señales residuales sean minimizadas, por ejemplo, similar a una transformación de eje principal. Por ejemplo, la energía de la señal de mezcla descendente debe ser maximizada y las energías de las señales residuales deben ser minimizadas. Cuando se mezclan de forma descendente dos señales a 1 señal, el mezclado descendente es normalmente uno de los dos componentes que resultan del mapeo de las 2 señales de entrada. El componente restante que resulta del mapeo es la señal residual y permite reconstruir las 2 señales originales mediante un mapeo inverso.

[0014] En algunos casos, la señal residual puede representar un error asociado con representar las dos señales mediante sus parámetros de mezclado descendente y asociados. Por ejemplo, la señal residual puede ser una señal de error la cual representa el error entre los canales originales L, R y los canales L', R', resultantes de mezclar de forma ascendente la señal de mezcla descendente que fue generada en base a los canales originales L y R.

[0015] En otras palabras, una señal residual puede ser considerada como una señal en el dominio del tiempo o un dominio de la frecuencia o un dominio sub-banda, que junto con la señal de mezcla descendente sola o con la señal de mezcla descendente y la información paramétrica permite una reconstrucción correcta o casi correcta de un canal original. Casi correcto se debe entender como que la reconstrucción con la señal residual teniendo una energía mayor que cero es más cercana al canal original comparado con una reconstrucción mediante el uso del mezclado de forma descendente sin la señal residual o mediante el uso del mezclado de forma descendente y la

información paramétrica sin la señal residual.

[0016] Considerando MPEG Surround (MPS), se emplean estructuras similares a PS llamadas cajas Uno a Dos (cajas OTT) en árboles de decodificación de audio espacial. Esto se puede ver como una generalización del concepto de mezclado ascendente mono a estéreo a esquemas de codificación/decodificación de audio espacio multicanal. En MPS, también existen sistemas de mezclado ascendente de dos a tres (cajas TTT) que pueden aplicar decorrelacionadores dependiendo del modo de operación TTT. En el documento de J. Herre, K. Kjörling, J. Breebaart, et al., "MPEG surround - la norma ISO/MPEG para codificación de audio multicanal eficaz y compatible" ("MPEG surround—the ISO/MPEG standard for efficient and compatible multi-channel audio coding") de las Actas de la 122ª Convención de AES, Viena, Austria, mayo de 2007, se describen detalles.

[0017] Con respecto a la Codificación de Audio Direccional (DirAC), la DirAC se refiere a un esquema de codificación de campo de sonido paramétrico que no está ligado a un número fijo de canales de salida de audio con posiciones de altavoz fijas. La DirAC aplica decorrelacionadores en el reproductor DirAC, es decir, en el decodificador de audio espacial para sintetizar componentes no coherentes de campos de sonido. Se puede encontrar más información relacionada con codificación de audio direccional en el documento de Pulkki, Ville: "Renderización de Sonido Espacial con Codificación de Audio Direccional" ("Spatial Sound Reproduction with Directional Audio Coding") en J. Audio Eng. Soc., Vol. 55, No. 6, 2007.

[0018] Con respecto a decorrelacionadores del estado actual de la técnica en decodificadores de audio espacial, se hace referencia a la Norma Internacional SO/IEC "Tecnología de Información-tecnologías de audio MPEG" ("Information Technology- MPEG audio technologies) – Parte 1: MPEG Surround", ISO/IEC 23003-1:2007 y también al documento de J. Engdegard, H. Purnhagen, J. Röden, L.Liljeryd, "Atmósfera Sintética en Codificación Paramétrica de Estéreo" ("Synthetic Ambience in Parametric Stereo Coding") en Actas de la 116ª de AES Convention, Berlín, Preimpresión, mayo de 2004. Se usan estructuras reticulares pasa todo (lattice allpass structures) IIR como decorrelacionadores en decodificadores de audio espacial tipo MPS como se describe en el documento de J. Herre, K. Kjörling, J. Breebaart, et al., "MPEG Surround - la norma ISO/MPEG para codificación de audio multicanal eficaz y compatible" ("MPEG surround—the ISO/MPEG standard for efficient and compatible multi-channel audio coding") de las Actas de la 122ª Convención de AES, Viena, Austria, mayo de 2007, y se describen en la Norma Internacional ISO/IEC "Tecnología de Información - tecnologías de audio MPEG - Parte 1: MPEG Surround" ("Information Technology- MPEG audio technologies – Part1: MPEG Surround"), ISO/IEC 23003-1:2007. Otros decorrelacionadores del estado actual de la técnica aplican retrasos (potencialmente dependientes de la frecuencia) para decorrelacionar señales o convolucionar las señales de entrada, por ejemplo, con erupciones de ruido que decaen exponencialmente. Para un panorama de los decorrelacionadores para sistemas de mezclado ascendente de audio espacial del estado actual de la técnica, véase "Atmósfera Sintética en Codificación Estéreo Paramétrica" ("Synthetic Ambience in Parametric Stereo Coding") en Actas de la 116ª Convention de AES, Berlín, Preimpresión, mayo de 2004.

[0019] Otra técnica de procesamiento de señales es "procesamiento de mezclado ascendente semántico" (semantic upmix processing). El procesamiento de mezclado ascendente semántico es una técnica para descomponer señales en componentes con diferentes propiedades semánticas (es decir, clases de señal) y aplicar diferentes estrategias de mezclado ascendente a los diferentes componentes de señal. Los diferentes algoritmos de mezclado ascendente pueden ser optimizados según las diferentes propiedades semánticas para mejorar el esquema de procesamiento de señal global. Este concepto se describe en el documento WO/2010/017967, "Un aparato para determinar una señal de audio multicanal-canal de salida espacial" (An apparatus for determining a spatial output multichannel-channel audio signal), solicitud de patente internacional PCT/EP2009/005828, 11.8.2009, 11.6.2010 (FH090802PCT).

[0020] Un esquema de codificación de audio espacial adicional es el "procedimiento de permutación temporal", según se describe en el documento de Hotho, G., van de Par, S., y Breebaart, J.: "Codificación multicanal de señales de aplauso" ("Multichannel coding of applause signals"), EURASIP Journal on Advances in Signal Processing, enero de 2008, art. 10. DOI=<http://dx.doi.org/10.1155/2008/>. En este documento se propone un esquema de codificación de audio espacial que se adecua a medida a la codificación/decodificación de señales de tipo aplauso. Este esquema se basa en la similaridad perceptual de segmentos de una señal de audio monofónica, especialmente una señal de mezcla descendente de un codificador de audio espacial. La señal de audio monofónica es segmentada en segmentos de tiempo que se superponen. Estos segmentos son permutados temporalmente pseudo aleatoriamente (mutuamente independientes para n canales de salida) dentro de un "superbloque" para formar los canales de salida decorrelacionados.

[0021] Una técnica de codificación de audio espacial adicional es el "procedimiento de intercambio y retraso

temporal" ("temporal delay and swapping method"). En el documento " DE 10 2007 018032 A: 20070417, Erzeugung dekorrelierter Signale", 17.4.2007, 23.10.2008 (FH070414PDE), se propone un esquema que también se adecua a medida para la codificación/decodificación de señales de tipo aplauso para la presentación binaural. Este esquema también se basa en la similaridad perceptual de segmentos de una señal de audio monofónica y retrasa canales de salida uno con respecto a otro. Para evitar una influencia de localización hacia el canal central, son barridos periódicamente los canales delantero y posterior.

[0022] En general, se sabe que las señales de tipo aplauso estéreo o multicanal codificadas/decodificadas en codificadores de audio espacial paramétricos dan por resultado reducida calidad de señal (véase, por ejemplo, Hotho, G., van de Par, S., y Breebaart, J.: "Codificación multicanal de señales de aplauso" ("Multichannel coding of applause signals"), EURASIP Journal on Advances in Signal Processing, enero de 2008, art. 10. DOI=<http://dx.doi.org/10.1155/2008/531693>, ver también DE 10 2007 018032 A). Las señales de tipo aplauso se caracterizan por contener mezclas temporalmente densas de componentes transitorios procedentes de diferentes direcciones. Ejemplos de tales señales son aplausos, el sonido de la lluvia, caballos galopando, etc. Las señales de tipo aplauso también contienen con frecuencia componentes de sonido procedentes de fuentes de sonido distantes, que se funden perceptualmente en un campo de sonido de fondo, suave, tipo ruido.

[0023] Las técnicas de decorrelación del estado actual de la técnica empleadas en decodificadores de audio espacial tipo MPEG Surround contienen estructuras reticulares pasa todo. Estas actúan como generadores de reverberación y consecuentemente son adecuadas para generar sonidos homogéneos, suaves, de tipo ruido, inmersivos (como colas de reverberación del ambiente) Sin embargo, hay ejemplos de campos de sonido con una estructura espacio-temporal no homogénea que aun son inmersores para el oyente: un ejemplo prominente son campos de sonido de tipo aplauso que crean envolvimiento del oyente no solo mediante campos de tipo ruido homogéneos, sino también mediante secuencias bastante densas de palmadas singulares procedentes de diferentes direcciones. Por consiguiente, el componente no homogéneo de los campos de sonido de aplauso puede estar caracterizado por una mezcla espacialmente distribuida de componentes transitorios. Obviamente, estas palmadas separadas no son homogéneas, suaves y de tipo ruido en absoluto.

[0024] Debido a su comportamiento de tipo reverberación, los decorrelacionadores de retícula pasa todo son incapaces de generar un campo de sonido inmersivo con las características, por ejemplo, del aplauso. En cambio, cuando se aplica a señales de tipo aplauso, tienden a embadurnar temporalmente los componentes transitorios de la señales. El resultado no deseado es un campo de sonido inmersivo de tipo ruido sin la estructura espacio-temporal distintiva de los campos de sonido de tipo aplauso. Además, eventos transitorios como un aplauso aislado podría evocar artefactos campanilla (N. del T. fenómeno de Gibbs) de los filtros del decorrelacionador.

[0025] Un sistema según Hotho, G., van de Par, S., y Breebaart, J.: "Codificación multicanal de señales de aplauso" ("Multichannel coding of applause signals"), EURASIP Journal on Advances in Signal Processing, enero de 2008, art. 10. DOI=<http://dx.doi.org/10.1155/2008/531693>, exhibirá degradación perceptible del sonido de salida debido a una cierta cualidad repetitiva en las señales de audio de salida. Eso ocurre por el hecho de que uno y el mismo segmento de la señal de entrada aparece inalterado en cada canal de salida (aunque en un punto de tiempo diferente). Asimismo, para evitar una densidad de aplauso aumentada, hay que dejar caer algunos canales originales en el mezclado ascendente y, por consiguiente, podría faltar algún evento auditivo importante en el mezclado ascendente resultante. El procedimiento solo es aplicable si es posible encontrar segmentos de señal que compartan las mismas propiedades perceptuales, es decir, segmentos de señal que suenen similares. El procedimiento en general difícilmente cambia la estructura temporal de las señales, lo cual podría ser aceptable solo para muy pocas señales. En el caso de aplicar el esquema a señales que no son del tipo aplauso (por ejemplo, debido a la mala clasificación de señal), la permutación temporal conducirá con mayor frecuencia a resultados inaceptables. La permutación temporal limita además la aplicabilidad a casos donde se pueden mezclar varios segmentos de señal juntos sin artefactos como ecos o filtrado peine. Desventajas similares valen para el procedimiento descrito en el documento DE 10 2007 018032 A.

[0026] El procesamiento de mezcla ascendente semántico descrito en el documento WO/2010/017967 separa los componentes transitorios de las señales antes de la aplicación de decorrelacionadores. La señal remanente (sin componentes transitorios) es alimentada a la decorrelación convencional y al procesador de mezclado ascendente, mientras las señales transitorias son manejadas de diferente manera; estas últimas son distribuidas (por ejemplo aleatoriamente) a diferentes canales de la señal de salida estéreo o multicanal mediante aplicación de técnicas de paneo de amplitud. El paneo de amplitud muestras varias desventajas: el paneo de amplitud no necesariamente produce una señal de salida que es cercana a la original. La señal de salida puede ser cercana a la original si la distribución de los componentes transitorios en la señal original pueden ser descritos por las leyes de paneo de amplitud. Es decir: el paneo de amplitud solo puede renderizar eventos puramente paneados

en amplitud correctamente, pero sin diferencias de fase o tiempo entre componentes transitorios en diferentes canales de salida.

5 **[0027]** Asimismo, la aplicación de la estrategia de paneo de amplitud en MPS requeriría puentear no solo el decorrelacionador sino también la matriz de mezclado ascendente. Como la matriz de mezclado ascendente refleja los parámetros espaciales (correlaciones intercanal: ICCs, diferencias de nivel intercanal ILDs) que son necesarios para sintetizar una salida mezclada de forma ascendente que muestra las propiedades espaciales correctas, el sistema de paneo en sí mismo tiene que aplicar alguna regla para sintetizar señales de salida con las propiedades espaciales correctas. No se conoce una regla genérica para hacerlo. Además, esta estructura agrega complejidad ya
10 que hay que cuidar los parámetros espaciales dos veces: uno, para parte no transitoria de la señal y, segundo, para la parte transitoria paneada en amplitud de la señal.

[0028] Por lo tanto es un objetivo de la presente invención proporcionar un concepto mejorado para generar una señal decorrelacionada. El objetivo de la presente invención es resuelto mediante un aparato según la
15 reivindicación 1, mediante un procedimiento según la reivindicación 4 y mediante un programa informático según la reivindicación 7.

[0029] Un aparato según una forma de realización comprende un separador transitorio para separar una señal de entrada en un primer componente de señal y un segundo componente de señal de tal forma que el primer
20 componente de señal comprenda porciones de señal transitorias de la señal de entrada y de tal forma que el segundo componente de señal comprenda porciones de señal no transitorias de la señal de entrada. El separador transitorio puede separar los diferentes componentes de señal uno de otro para permitir que los componentes de señal que contienen transitorios puedan ser procesados de manera diferente que los componentes de señal que no contienen transitorios.

25 **[0030]** El aparato comprende además un decorrelacionador transitorio para decorrelacionar componentes de la señal que contienen transitorios según un procedimiento de decorrelación el cual es particularmente adecuado para decorrelacionar componentes de la señal que contienen transitorios. Asimismo, el aparato comprende un segundo decorrelacionador para decorrelacionar los componentes de la señal que no contienen transitorios.

30 **[0031]** Así, el aparato es capaz de procesar componentes de la señal usando un decorrelacionador estándar o bien, alternativamente procesar componentes de la señal usando el decorrelacionador transitorio particularmente adecuado para procesar componentes de la señal transitorios. En una forma de realización, el separador transitorio decide si un componente de la señal es alimentado al decorrelacionador estándar o al decorrelacionador transitorio.

35 **[0032]** Asimismo, el aparato puede estar adaptado para separar un componente de señal de tal forma que el componente de señal es alimentado parcialmente al decorrelacionador transitorio y parcialmente es alimentado al segundo decorrelacionador.

40 **[0033]** Asimismo, el aparato comprende una unidad de combinación para combinar los componentes de señal entregados por el decorrelacionador estándar y el decorrelacionador transitorio para generar una señal de combinación decorrelacionada.

45 **[0034]** En una forma de realización, el aparato comprende una unidad de recepción para recibir información de fase, en la que el decorrelacionador transitorio está adaptado para aplicar la información de fase al primer componente de la señal. La información de fase podría ser generada por un codificador adecuado.

[0035] En una forma de realización, el separador transitorio está adaptado tanto para alimentar una porción de señal considerada de una señal de entrada de aparato en el decorrelacionador de transitorio como para alimentar
50 la porción de señal considerada en el segundo decorrelacionador dependiendo de la información de separación de componente transitorio la cual, o bien indica que la porción de señal considerada contiene un transitorio, o la cual indica que la porción de señal considerada no contiene un transitorio. Tal forma de realización permite un fácil procesamiento de información de separación de componente transitorio.

55 **[0036]** En otra forma de realización, el separador transitorio está adaptado para alimentar parcialmente una porción de señal considerada de una señal de entrada de aparato, al decorrelacionador de transitorio y para alimentar parcialmente la porción de señal considerada al segundo decorrelacionador. La magnitud de la porción de señal considerada que es alimentada al separador transitorio y la magnitud de porción de señal considerada que es alimentada al segundo decorrelacionador dependen de la información de separación transitoria. Por este medio se
60 puede tomar en cuenta la intensidad de un componente transitorio.

[0037] En una forma de realización adicional, el separador transitorio está adaptado para separar una señal de entrada de aparato la cual está representada en un dominio de la frecuencia. Esto permite el procesamiento (separación y decorrelación) de componente transitorio dependiente de la frecuencia. Así, ciertos componentes de
 5 señal de una primera banda de frecuencia pueden ser procesados según un procedimiento de decorrelación de componente transitorio, mientras componentes de señal de otra banda de frecuencia pueden ser procesados según otro procedimiento, por ejemplo, decorrelación convencional. Por consiguiente, en una forma de realización el separador transitorio está adaptado para separar una señal de entrada de aparato en base a información de
 10 separación transitoria dependientes de la frecuencia. Sin embargo, en una forma de realización el separador transitorio está adaptado para separar una señal de entrada de aparato en base a información de separación independiente de la frecuencia. Esto permite procesamiento más eficaz de señal con componente transitorio.

[0038] En otra forma de realización, el separador transitorio puede estar adaptado para separar una señal de entrada de aparato la cual está representada en un dominio de la frecuencia de tal forma que todas las porciones de
 15 señal de la señal de entrada de aparato dentro de un primer rango de frecuencia son alimentadas al segundo decorrelacionador. Un aparato correspondiente está adaptado, por lo tanto, para restringir el procesamiento de señal transitoria a componentes de señal con frecuencias de señal en un segundo rango de frecuencia, mientras ningún componente de señal con frecuencias de señal en el primer rango de frecuencia es alimentado al decorrelacionador transitorio (sino en el segundo decorrelacionador).

[0039] En una forma de realización adicional, el decorrelacionador transitorio puede estar adaptado para decorrelacionar el primer componente de señal aplicando información de fase que representa una diferencia de fase entre una señal residual y una señal de mezcla descendente. Del lado del codificador se puede emplear una matriz de mezclado "reversa" para crear una señal de mezcla descendente y una señal residual, por ejemplo, desde los dos
 25 canales de una señal estéreo, como se ha explicado más arriba. Mientras la señal de mezcla descendente puede ser transmitida al decodificador, la señal residual puede ser descartada. Según una forma de realización, la diferencia de fase empleada por el decorrelacionador transitorio puede ser la diferencia de fase entre la señal residual y la señal de mezcla descendente. Así puede ser posible reconstruir una señal residual "artificial" aplicando la fase original de la residual a la mezclada de forma descendente. En una forma de realización, la diferencia de fase
 30 puede relacionarse con una cierta banda de frecuencia, es decir, puede ser dependiente de la frecuencia. Alternativamente, una diferencia de fase no se relaciona con ciertas bandas de frecuencias pero puede ser aplicada como un parámetro de banda ancha independiente de la frecuencia.

[0040] En una forma de realización adicional se podría aplicar el término de fase al primer componente de
 35 señal multiplicando el término de fase con el primer componente de señal.

[0041] En una forma de realización adicional el segundo decorrelacionador puede ser un decorrelacionador convencional, por ejemplo, un decorrelacionador IIR de retícula.

[0042] En una forma de realización, el aparato comprende un mezclador que está adaptado para recibir señales de entrada y también está adaptado para generar señales de salida en base a las señales de entrada y a una regla de mezclado. Se alimenta una señal de entrada de aparato a un separador transitorio y, a continuación, es decorrelacionada por un separador transitorio y/o un segundo decorrelacionador como se describe más arriba. La
 40 unidad de combinación y el mezclador pueden ser acomodados de modo que la señal de combinación decorrelacionada sea alimentada al mezclador como una primera señal de entrada de mezclador. Una segunda señal de entrada de mezclador puede ser la señal de entrada de aparato o una señal derivada de la señal de entrada de aparato. Como el proceso de decorrelación ya está completo cuando la señal de combinación decorrelacionada es alimentada al mezclador, el mezclador no tiene que tomar en cuenta la decorrelación del
 45 componente transitorio. Por lo tanto, se puede emplear un mezclador convencional.

[0043] En una forma de realización adicional, el mezclador está adaptado para recibir datos de parámetro de correlación/coherencia indicando una correlación o coherencia entre dos señales y está adaptado para generar las
 50 señales de salida en base a los datos de parámetro de correlación/coherencia. En otra forma de realización, el mezclador está adaptado para recibir datos de parámetro de diferencia de nivel indicando una diferencia de energía entre dos señales y está adaptado para generar las señales de salida en base a los datos de parámetro de
 55 diferencia de nivel. En tal forma de realización, el decorrelacionador de componente transitorio, el segundo decorrelacionador y la unidad de combinación no tiene que estar adaptado para procesar tales datos de parámetros, ya que el mezclador se ocupará de procesar los datos correspondientes. Por el otro lado, en tal forma de realización se puede usar un mezclador convencional con procesamiento de parámetros de correlación/coherencia y diferencia
 60 de nivel.

[0044] Ahora se explicarán las formas de realización en más detalle con respecto a las figuras, en las que:

- 5 la Figura 1 ilustra una aplicación del estado actual de la técnica, de un decorrelacionador en un mezclador ascendente mono a estéreo;
- la Figura 2 representa una aplicación del estado actual de la técnica adicional de un decorrelacionador en un mezclador ascendente mono a estéreo;
- la Figura 3 ilustra un aparato para la generación de una señal decorrelacionada según una forma de realización;
- la Figura 4 ilustra un aparato para la decodificación de una señal según una forma de realización;
- 10 la Figura 5 es una vista global de un sistema uno a dos (OTT) según una forma de realización;
- la Figura 6 ilustra un aparato para generar una señal decorrelacionada que comprende una unidad de recepción según una forma de realización adicional;
- la Figura 7 es una vista global de un sistema uno a dos según otra forma de realización adicional;
- la Figura 8 ilustra mapeos ejemplares desde medidas de consistencia de fase a intensidad de separación de
- 15 componente transitorio;
- la Figura 9 es una vista global de un sistema uno a dos según otra forma de realización adicional;
- la Figura 10 ilustra un aparato para codificar una señal de audio que tiene una pluralidad de canales.

[0045] La Figura 3 ilustra un aparato para la generación de una señal decorrelacionada según una forma de realización. El aparato comprende un separador transitorio 310, un decorrelacionador transitorio 320, un decorrelacionador convencional 330 y una unidad de combinación 340. La estrategia de gestión de componente transitorio de esta forma de realización apunta a generar señales decorrelacionadas a partir de señales de audio de tipo aplauso, por ejemplo, para la aplicación en el proceso de mezclado ascendente de decodificadores de audio espacial.

25

[0046] En la Figura 3, una señal de entrada es alimentada al separador de componente transitorio 310. La señal de entrada puede haber sido transformada a un dominio de la frecuencia, por ejemplo, aplicando un banco de filtro QMF híbrido. El separador transitorio 310 puede decidir para cada componente de señal considerado de la señal de entrada, si ésta contiene un componente transitorio o no. Asimismo, el separador 310 que comprende una

30 porción de señal con componente transitorio puede estar dispuesto para alimentar la porción de señal considerada, ya sea al decorrelacionador transitorio 320, si la porción de señal considerada contiene un componente transitorio (componente de señal s1), o puede alimentar la porción de señal considerada al decorrelacionador convencional 330, si la porción de señal considerada no contiene un componente transitorio (componente de señal s2). El separador transitorio 310 también puede estar dispuesto para dividir la porción de señal considerada dependiendo

35 de la existencia de un componente transitorio en la porción de señal considerada y proporcionarlos parcialmente al decorrelacionador transitorio 320 y parcialmente al decorrelacionador convencional 330.

[0047] En una forma de realización, el decorrelacionador transitorio 320 decorrelaciona el componente de señal s1 según un procedimiento de decorrelación de componente transitorio el cual es particularmente adecuado

40 para decorrelacionar componentes de señal transitorios. Por ejemplo, la decorrelación de los componentes de señal transitorios puede ser llevada a cabo aplicando información de fase, por ejemplo aplicando términos de fase. Se explica un procedimiento de decorrelación en el que se aplican términos de fase sobre componentes de señal transitorios abajo, con respecto a la forma de realización de la Figura 5. También se puede emplear tal

45 procedimiento de decorrelación como procedimiento de decorrelación de componente transitorio del decorrelacionador transitorio 320 de la forma de realización de la Figura 3.

[0048] El componente de señal s2, el cual no comprende porciones de señal transitorias, es alimentado al decorrelacionador convencional 330. El decorrelacionador convencional 330 puede decorrelacionar entonces el componente de señal s2 según un procedimiento de decorrelación convencional, por ejemplo, aplicando estructuras

50 reticulares pasa todo, por ejemplo, un filtro (de respuesta de impulso infinito) IRR reticular.

[0049] Después de ser decorrelacionado mediante el decorrelacionador convencional 330, el componente de señal decorrelacionado procedente del decorrelacionador convencional 330 es alimentado a la unidad de combinación 340. El componente de señal transitorio decorrelacionado procedente del decorrelacionador transitorio

55 320 también es alimentado a la unidad de combinación 340. La unidad de combinación 340 combina entonces ambos componentes de señal decorrelacionados, por ejemplo, sumando ambos componentes de señal, para obtener una señal de combinación decorrelacionada.

[0050] En general, un procedimiento que decorrelaciona una señal que contiene componentes transitorios

60 según una forma de realización, puede realizarse del siguiente modo:

En una etapa de separación, la señal de entrada es separada en dos componentes: un componente s1 contiene los transitorios de la señal de entrada, otro componente s2 contiene la parte restante (no transitoria) de la señal de entrada. El componente no transitorio s2 de la señal puede ser procesado como en sistemas sin aplicar el procedimiento de decorrelación del decorrelacionador transitorio de esta forma de realización. Es decir, la señal sin transitorios s2 puede ser alimentada a una o varias estructuras de procesamiento de señal que decorrelacionan en forma convencional como estructuras pasa todo de IIR reticulares.

[0051] Asimismo, el componente de señal que contiene los transitorios (la transmisión de transitorios s1) es alimentado a una estructura de "decorrelacionador transitorio" que decorrelaciona la transmisión de transitorios mientras que mantiene las propiedades especiales de señal mejor que la estructuras decorrelacionadoras convencionales. La decorrelación de la transmisión de transitorios es llevada a cabo aplicando información de fase a una alta resolución temporal. Preferiblemente, la información de fase comprende términos de fase. Asimismo, se prefiere que la información de fase pueda ser proporcionada por un codificador.

[0052] Además, las señales de salida de ambos, el decorrelacionador convencional y el decorrelacionador transitorio, son combinadas para formar la señal decorrelacionada que podría ser utilizada en el proceso de mezclado ascendente de codificadores de audio espacial. Los elementos (h11, h12, h21, h22) de la matriz de mezclado (Mmix) del decodificador de audio espacial pueden permanecer sin cambio.

[0053] La Figura 4 ilustra un aparato para decodificar una señal de entrada de aparato según una forma de realización, en donde la señal de entrada de aparato es alimentada al separador transitorio 410. El aparato comprende un separador transitorio 410, un decorrelacionador transitorio 420, un decorrelacionador convencional 430, una unidad de combinación 440 y un mezclador 450. El separador transitorio 410, el decorrelacionador transitorio 420, el decorrelacionador convencional 430 y la unidad de combinación 440 de esta forma de realización pueden ser similares al separador 310, el decorrelacionador transitorio 320, el decorrelacionador convencional 330 y la unidad de combinación 340 de la forma de realización de la Figura 3, respectivamente. Una señal de combinación decorrelacionada generada por la unidad de combinación 440 es alimentada a un mezclador 450 como una primera señal de entrada de mezclador. Además, la señal de entrada de aparato que ha sido alimentada al separador transitorio 410, también es alimentada al mezclador 450 como una segunda señal de entrada de mezclador. Alternativamente, la señal de entrada de aparato no es alimentada directamente al mezclador 450, sino que se alimenta una señal derivada de la señal de entrada de aparato al mezclador 450. Se puede derivar una señal a partir de la señal de entrada de aparato, por ejemplo, aplicando un procedimiento de procesamiento de señal convencional a la señal de entrada de aparato, por ejemplo, aplicando un filtro. El mezclador 450 de la forma de realización de la Figura 4 está adaptado para generar señales de salida en base a las señales de entrada y una regla de mezclado. Tal regla de mezclado puede ser, por ejemplo, multiplicar las señales de entrada y una matriz de mezclado, por ejemplo aplicando la fórmula

$$\begin{bmatrix} L \\ R \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} M \\ D \end{bmatrix}$$

[0054] El mezclador 450 puede generar canales de salida L, R sobre la base de datos de parámetro de correlación/coherencia, por ejemplo, Correlación/Coherencia Inter-Canal (ICC) y/o datos de parámetro de diferencia de nivel, por ejemplo, Diferencia de Nivel Inter Canal (ILD). Por ejemplo, los coeficientes de una matriz de mezclado pueden depender de los datos de parámetro de correlación/coherencia y/o de los datos de parámetro de diferencia de nivel. En la forma de realización de la Figura 4, el mezclador 450 genera dos canales de salida L y R. Sin embargo, en formas de realización alternativas, el mezclado puede generar una pluralidad de señales de salida, por ejemplo, 2, 4, 5 o 9 señales de salida, las cuales pueden ser señales de sonido envolvente.

[0055] La Figura 5 representa una vista global del sistema de la estrategia de gestión de componente transitorio en un sistema de mezclado ascendente 1 a 2 (OTT), por ejemplo, un decodificador de audio espacial MPS (MPEG Surround). La trayectoria de señal paralela para los componentes transitorios separados según una forma de realización, está comprendida en una caja de gestión de componente transitorio en forma de U. Se alimenta una señal de entrada de aparato DMX en un separador transitorio 510. La señal de entrada de aparato puede ser representada en un dominio de la frecuencia. Por ejemplo, una señal de entrada del dominio del tiempo puede haber sido transformada a un dominio de la frecuencia aplicando un banco de filtros QMF como se usa en MPEG Surround. El separador transitorio 510 puede alimentar a continuación los componentes de la señal de entrada de aparato DMX en un decorrelacionador transitorio 520 y/o en un decorrelacionador reticular IIR. Los componentes de la señal de entrada de aparato son decorrelacionados entonces por el decorrelacionador transitorio 520 y/o el

decorrelacionador reticular IIR 530. Después, los componentes de señal decorrelacionados D1 y D2 son combinados mediante una unidad de combinación 540, por ejemplo, sumando ambos componentes de señal, para obtener una señal de combinación decorrelacionada D. La señal de combinación decorrelacionada es alimentada en un mezclador 552 como una primera señal de entrada de mezclador D. Asimismo, la señal de entrada de aparato DMX (o alternativamente: una señal derivada de la señal de entrada de aparato DMX) es alimentada también al mezclador 552 como una segunda señal de entrada de mezclador. El mezclador 552 genera entonces una primera y una segunda señal "seca", dependiendo de la señal de entrada de aparato DMX. El mezclador 552 genera también una primera y una segunda señal "mojada" dependiendo de la señal de combinación decorrelacionada D. Las señales, generadas por el mezclador 552 también pueden ser generadas en base a parámetros transmitidos, por ejemplo, Correlación/Coherencia Inter-Canal (ICC) y/o datos de parámetro de diferencia de nivel, por ejemplo, Diferencia de Nivel Inter Canal (ILD). En una forma de realización, las señales generadas por el mezclador 552 pueden ser proporcionadas a una unidad de modelación 554 la cual modela las señales proporcionadas en base a datos de modelado temporal proporcionados. En otras formas de realización no tiene lugar modelado de señal. Las señales generadas a continuación son proporcionadas a una primera 556 y una segunda unidad sumadora las cuales combinan las señales proporcionadas para generar una primera señal de salida L y una segunda señal de salida R, respectivamente.

[0056] Los principios de procesamiento mostrados en la Figura 5 pueden ser aplicados en sistemas de mezclado ascendente mono a estéreo (por ejemplo, codificadores de audio estéreo) así como también en disposiciones multicanal (por ejemplo, MPEG Surround). En formas de realización, el esquema de manipulación de componente transitorio propuesto puede ser aplicado como un mejoramiento de calidad a sistemas de mezclado ascendente existentes, ya que solo se introduce una trayectoria de señal de decorrelacionador paralela sin alterar el proceso de mezclado ascendente en sí mismo.

[0057] La separación de señal en componentes transitorios y componentes no transitorios es controlada mediante parámetros que podrían ser generados en un codificador y/o el decodificador de audio espacial. El decorrelacionador transitorio 520 utiliza información de fase, por ejemplo, términos de fase que podrían ser obtenidos en un codificador o en un decodificador de audio espacial. Más abajo se describen posibles variantes para obtener parámetros de gestión de componentes transitorios (es decir, parámetros de separación transitoria como posiciones de componentes transitorios o intensidad de separación y parámetros de decorrelación de componentes transitorios como información de fase).

[0058] La señal de entrada puede ser representada en un dominio de la frecuencia. Por ejemplo, una señal puede haber sido transformada a un dominio de la frecuencia empleando un banco de filtros de análisis. Se puede aplicar un banco de filtros QMF para obtener una pluralidad de señales sub-banda a partir de una señal del dominio del tiempo.

[0059] Para la mejor calidad perceptual, el procesamiento de señal transitoria puede ser restringido, preferiblemente, a frecuencias de señal en un rango de frecuencia limitado. Un ejemplo sería limitar el rango de procesamiento a índices de banda de frecuencia $k \geq 8$ de un bando de filtros QMF híbrido como se usa en MPS, similar a la limitación de banda de frecuencia de modelado de envolvente guiado (GES) en MPS.

[0060] En lo sucesivo se explicarán con más detalle formas de realización de un separador transitorio 520. El separador transitorio 510 divide la señal de entrada DMX en componentes transitorios y no transitorios s_1 y s_2 , respectivamente. El separador transitorio 510 puede emplear información de separación transitoria para dividir la señal de entrada DMX, por ejemplo, un parámetro de separación de componente transitorio $\beta[n]$. La división de la señal de entrada DMX se puede hacer de tal manera que la suma de los componentes s_1+s_2 , sea igual a la señal de entrada DMX:

$$s_1[n] = DMX[n] \cdot \beta[n]$$

$$s_2[n] = DMX[n] \cdot (1 - \beta[n])$$

donde n es el índice de tiempo de señales sub-banda muestreadas de forma descendente y valores válidos para el parámetro de separación de componente transitorio que varía en el tiempo $\beta[n]$. están en el rango $[0, 1]$. $\beta[n]$ puede ser un parámetro independiente de la frecuencia. Un separador transitorio 510 que está adaptado para separar una señal de entrada de aparato en base a un parámetro de separación independiente de la frecuencia puede alimentar todas las porciones de señal sub-banda con índice de tiempo n ya sea al decorrelacionador transitorio 520 o al

segundo decorrelacionador dependiendo del valor de $\beta[n]$.

[0061] Alternativamente, $\beta[n]$ puede ser un parámetro dependiente de la frecuencia. Un separador transitorio 510 que está adaptado para separar una señal de entrada de aparato en base a una información de separación de 5 componentes transitorios dependiente de la frecuencia puede procesar porciones de señal sub-banda.

[0062] Asimismo, la dependencia de la frecuencia puede ser usada, por ejemplo, para limitar el rango de frecuencia del procesamiento de componente transitorio como se menciona en la sección de más arriba.

10 **[0063]** En una forma de realización, la información de separación transitoria puede ser un parámetro que o bien indica que una porción de señal considerada de una señal de entrada DMX contiene un componente transitorio, o bien que indica que la porción de señal considerada no contiene un componente transitorio. El separador transitorio 510 alimenta la porción de señal considerada al decorrelacionador transitorio 520, si la información de separación de componente transitorio indica que la porción de señal considerada contiene un componente transitorio. Alternativamente, el separador transitorio 510 alimenta la porción de señal considerada al segundo decorrelacionador, por ejemplo, el decorrelacionador de retícula IIR 530, si la información de separación de 15 componente transitorio indica que la porción de señal considerada contiene un componente transitorio.

[0064] Por ejemplo, se puede emplear un parámetro de separación transitoria $\beta[n]$ como una información de separación de componente transitorio que puede ser un parámetro binario. n es el índice de tiempo de una porción de señal considerada de la señal de entrada DMX. $\beta[n]$ puede ser 1 (indicando que la porción de señal considerada será alimentada al decorrelacionador transitorio) o 0 (indicando que la porción de señal considerada será alimentada al segundo decorrelacionador). Restringiendo $\beta[n]$ a $\beta \in \{0, 1\}$ resultan ser difíciles las decisiones de componente transitorio/no transitorio, es decir: componentes que son tratados como transitorios son completamente separados 25 de la entrada ($\beta = 1$).

[0065] En otra forma de realización, el separador transitorio 510 está adaptado para alimentar parcialmente una porción de señal considerada de la señal de entrada de aparato, al decorrelacionador transitorio 520 y para alimentar parcialmente la porción de señal considerada al segundo decorrelacionador 530. La magnitud de la 30 porción de señal considerada que es alimentada al separador transitorio 520 y la magnitud de porción de señal considerada que es alimentada al segundo decorrelacionador 530 depende de la información de separación transitoria. En una forma de realización, $\beta[n]$ tiene que estar en el intervalo $[0, 1]$. En otra forma de realización, $\beta[n]$ puede estar restringido a $\beta[n] \in [0, \beta_{\max}]$, donde $\beta_{\max} < 1$, resulta una separación parcial de los componentes transitorios, conduciendo a un efecto menos pronunciado del esquema de gestión de componentes transitorios. Por 35 lo tanto, cambiar β_{\max} permite que sea indistinto entre la salida del procesamiento de mezclado ascendente convencional sin gestión de componentes transitorios y el procesamiento de mezclado ascendente que incluye la gestión de componentes transitorios.

[0066] En lo sucesivo se explicará con más detalle un decorrelacionador transitorio 520 según una forma de 40 realización.

[0067] Un decorrelacionador transitorio 520 según una forma de realización crea una señal de salida que está suficientemente decorrelacionada de la entrada. No altera la estructura temporal de aplausos singulares/componentes transitorios (sin dispersión horizontal en el tiempo, sin retraso). En cambio, conduce a una 45 distribución espacial de los componentes de señal transitoria (después del proceso de mezclado ascendente), que es similar a la distribución espacial en la señal original (no codificada). El decorrelacionador transitorio 520 puede encargarse de compromisos de cantidad de bits transmitidos en contraste con calidad (por ejemplo, distribución de transitorio espacial a baja cantidad de bits transmitidos ↔ cerca de la original (casi transparente) a alta cantidad de bits transmitidos. Asimismo, esto se logra con baja complejidad computacional.

50 **[0068]** Como se ha explicado más arriba, del lado del codificador se puede emplear una matriz de mezclado "reversa" para crear una señal de mezcla descendente y una señal residual, por ejemplo, desde los dos canales de una señal estéreo. Mientras la señal de mezcla descendente puede ser transmitida al decodificador, la señal residual puede ser descartada. Según una forma de realización, la diferencia de fase entre la señal residual y la señal de 55 mezcla descendente puede ser determinada, por ejemplo, mediante un decodificador, y puede ser empleada por un decodificador cuando decorrelaciona una señal. Por este medio puede ser posible entonces reconstruir una señal residual "artificial" aplicando la fase original de la residual a la mezcla descendente.

[0069] A continuación se explicará un procedimiento de decorrelación correspondiente del decorrelacionador 60 transitorio 520 según una forma de realización: según un procedimiento de decorrelación de componente transitorio,

se puede emplear un término de fase. Se logra la decorrelación simplemente multiplicando la transmisión transitoria por términos de fase a alta resolución temporal, por ejemplo, a resolución de tiempo de señal sub-banda en sistemas del dominio de la transformación como MPS:

$$5 \quad D1[n] = s1[n] \cdot e^{j \cdot \Delta\phi[n]}$$

[0070] En esta ecuación, n es el índice de tiempo de señales sub-banda muestreadas de forma descendente. $\Delta\phi$ refleja idealmente la diferencia de fase entre el mezclado descendente y residual. Por lo tanto, los transitorios residuales son reemplazados por una copia de los transitorios procedentes del mezclado descendente, modificados de tal forma que exhiban la fase original.

[0071] Aplicar la información de fase inherentemente da como resultado un paneo de los componentes transitorios a la posición original en el proceso de mezclado de forma ascendente. Como ejemplo ilustrativo, considere el caso ICC=0, ILD=0: La parte transitoria de las señales de salida dicen entonces:

$$15 \quad L[n] = c \cdot (s[n] + D1[n]) = c \cdot s[n] \cdot (1 + e^{j \cdot \Delta\phi[n]})$$

$$R[n] = c \cdot (s[n] - D1[n]) = c \cdot s[n] \cdot (1 - e^{j \cdot \Delta\phi[n]})$$

20 [0072] Para $\Delta\phi=0$ resulta $L=2c*s$, $R=0$, mientras $\Delta\phi=\pi$ conduce a $L=0$, $R=2c*s$. Otros valores de $\Delta\phi$, ICC, e ILD conducen a diferentes relaciones de nivel y fase entre los componentes transitorios renderizados.

[0073] Los valores de $\Delta\phi[n]$ pueden ser aplicados como parámetros de banda ancha independientes de la frecuencia o como parámetros dependientes de la frecuencia. En caso de señales de tipo aplauso sin componentes tonales, pueden ser ventajosos valores de $\Delta\phi[n]$ de banda ancha debido a menores demandas de velocidad de datos y gestión consistente de componentes transitorios de banda ancha (consistencia sobre la frecuencia).

[0074] La estructura de gestión de componente transitorio de la Figura 5 está acomodada de tal forma que solo el decorrelacionador convencional 530 es puentado con respecto a los componentes de señal transitorios mientras la matriz de mezclado permanece inalterada. Así, los parámetros espaciales (ICC, ILD) también son inherentemente tomados en cuenta para las señales transitorias, por ejemplo, la ICC controla automáticamente el ancho de la distribución de transitorio renderizado.

[0075] Considerando el aspecto de cómo obtener información de fase, en una forma de realización, la información de fase puede ser recibida desde un codificador.

[0076] La Figura 6 ilustra una forma de realización de un aparato para generar una señal decorrelacionada. El aparato comprende un separador transitorio 610, un decorrelacionador transitorio 620, un decorrelacionador convencional 630, una unidad de combinación 640 y una unidad de recepción 650. El separador transitorio 610, el decorrelacionador convencional 630 y la unidad de combinación 640 son similares al separador transitorio 310, el decorrelacionador convencional 330 y la unidad de combinación 340 de la forma de realización de la Figura 3. Sin embargo, la Figura 6 ilustra además una unidad de recepción 650 que está adaptada para recibir información de fase. La información de fase puede haber sido transmitida por un codificador (no mostrado). Por ejemplo, un codificador puede haber computado la diferencia de fase entre señales residual y de mezcla descendente (fase relativa de la señal residual con respecto a una de mezcla descendente). La diferencia de fase puede haber sido calculada para ciertas bandas de frecuencias o banda ancha (por ejemplo, en un dominio del tiempo). El codificador puede codificar apropiadamente los valores de fase mediante cuantificación uniforme o no uniforme y potencialmente codificar sin pérdidas. Después, el codificador puede transmitir los valores de fase codificados al sistema de decodificación de audio espacial. Obtener la información de fase de un codificador es ventajoso ya que la información de fase original está disponible a continuación en un decodificador (excepto por el error de cuantificación).

[0077] La unidad de recepción 650 alimenta la información de fase en el decorrelacionador transitorio 620 el cual usa la información de fase cuando decorrelaciona un componente de señal. Por ejemplo, la información de fase puede ser un término de fase y el decorrelacionador transitorio 620 puede multiplicar un componente de señal transitorio recibido por el término de fase.

- [0078]** En caso de transmitir información de fase $\Delta\phi[n]$ desde un codificador al decodificador, se puede reducir la tasa de datos requerida del siguiente modo: la información de fase $\Delta\phi[n]$ puede ser aplicada solo a los componentes transitorios de una señal en el decodificador. Por lo tanto, la información de fase solo necesita estar disponible en el decodificador siempre que haya componentes transitorios en la señal que va a ser decorrelacionada. La transmisión de la información de fase puede ser limitada entonces por el codificador, de tal forma que solo la información necesaria sea transmitida al decodificador. Esto se puede hacer aplicando una detección de transitorios en el codificador como se describe más abajo. La información de fase $\Delta\phi[n]$ solo es transmitida para puntos en el tiempo n , para los cuales se han detectado componentes transitorios en el codificador.
- 10 **[0079]** Considerando el aspecto de separación transitoria, en una forma de realización, la separación transitoria puede ser conducida por el codificador.
- [0080]** Según una forma de realización, la información de separación transitoria (a lo que también se refiere como "información de transitorios") puede obtenerse de un codificador. El codificador puede aplicar procedimientos de detección de transitorios como se describe en "Uso de Supresión de Transitorios en Algoritmos de Mezcla Ascendente Multicanal Ciega" ("Using Transient Suppression in Blind Multi-channel Up-mix Algorithms") de Andreas Walther, Christian Uhle, Sascha Disch, en Proc. 122^o Convención de AES, Viena, Austria, mayo de 2007, ya sea para las señales de entrada de codificador o para las señales de mezcla descendente. La información de transitorios es transmitida a continuación al decodificador y preferiblemente es obtenida, por ejemplo, a la resolución de tiempo de las señales sub-banda muestreadas de forma descendente.
- 15 **[0081]** La información de transitorios puede comprender preferiblemente una simple decisión binaria (transitorio/no-transitorio) para cada muestra de señal en el tiempo. Esta información puede ser representada también, preferiblemente, mediante las posiciones de transitorio en el tiempo y las duraciones del transitorio.
- 20 **[0082]** La información de transitorios puede ser codificada sin pérdidas (por ejemplo, codificación por longitud de series (RLE), codificación entrópica) para reducir la tasa de datos que es necesaria para transmitir la información de transitorios desde el codificador al decodificador.
- 30 **[0083]** La información de transitorios puede ser transmitida como información de banda ancha o como información dependiente de la frecuencia a una resolución de frecuencia. Transmitir la información de transitorios como parámetros de banda ancha reduce la tasa de datos de información de transitorios y potencialmente mejora la calidad de audio debido a la gestión consistente de transitorios de banda ancha.
- 35 **[0084]** En lugar de la decisión binaria (transitorio/no-transitorio), también puede ser transmitida la intensidad de los transitorios, por ejemplo, cuantificada en dos o cuatro escalones. La intensidad de transitorios puede controlar entonces la separación de los transitorios en el decodificador de audio espacial del siguiente modo: los transitorios fuertes son separados completamente de la entrada del decorrelacionador reticular IIR, mientras que los transitorios más débiles solo son separados parcialmente.
- 40 **[0085]** La información de transitorios solo puede ser transmitida, si el codificador detecta señales tipo aplauso, por ejemplo usando sistemas de detección de aplauso como se describe en "Detección de Sonido de Aplauso con Baja Latencia" ("Applause Sound Detection with Low Latency") de Christian Uhle, en la 127^o Convención de la Sociedad de Ingeniería de Audio, Nueva York, 2009.
- 45 **[0086]** El resultado de la detección para la similaridad de señal de entrada a señales de tipo aplauso también puede ser transmitido a menor resolución de tiempo (por ejemplo, a la tasa de actualización de parámetros espaciales en MPS) al decodificador para controlar la intensidad de la separación de transitorios. El resultado de la detección de aplauso puede ser transmitido como un parámetro binario (es decir, como una decisión drástica) o como un parámetro no binario (es decir, como una decisión blanda). Este parámetro controla la intensidad de separación en el decodificador de audio espacial. Por lo tanto, permite encender/apagar (drásticamente o gradualmente) la gestión de transitorios en el decodificador. Esto permite evitar artefactos que podrían ocurrir, por ejemplo, cuando se aplica un esquema de gestión de transitorios de banda ancha a señales que contienen componentes tonales. La Figura 7 ilustra un aparato para la decodificación de una señal según una forma de realización. El aparato comprende un separador transitorio 710, un decorrelacionador transitorio 720, un decorrelacionador reticular IIR 730, una unidad de combinación 740, un mezclador 752, una unidad de modelación opcional 754, una primera unidad sumadora 756 y una segunda unidad sumadora 758, los cuales corresponden al separador transitorio 510, el decorrelacionador transitorio 520, el decorrelacionador reticular IIR 530, la unidad de combinación 540, el mezclador 552 la unidad de modelación opcional 554, la primera unidad sumadora 556 y la
- 50 55 60 segunda unidad sumadora 558 de la forma de realización de la Figura 5, respectivamente. En la forma de realización

de la Figura 7, un codificador obtiene información de fase e información de posición de transitorio y transmite la información a un aparato para decodificación. No se transmiten señales residuales. La Figura 7 ilustra una configuración de mezclado ascendente 1 a 2 como una caja OTT en MPS. Se puede aplicar en un codificador-decodificador estéreo para mezclar de forma ascendente a partir de mezcla descendente mono a una salida estéreo según una forma de realización. En la forma de realización de la Figura 7 se transmiten tres parámetros de gestión de transitorios como parámetros independientes de la frecuencia desde el codificador al decodificador, como se puede ver en la Figura 7:

Un primer parámetro de gestión de transitorio que va a ser transmitido es la decisión binaria transitorio/no-transitorio de un detector de transitorios que se ejecuta en el codificador. Se usa para controlar la separación de transitorios en el decodificador. En un esquema simple, la decisión binaria transitorio/no-transitorio puede ser transmitida como una bandera binaria por muestra de tiempo sub-banda sin codificación adicional.

[0087] Otro parámetro de gestión de transitorio que va a ser transmitido es el valor de fase (o los valores de fase) $\Delta\phi[n]$ que se necesita para el decorrelacionador de transitorios. $\Delta\phi[n]$ solo es transmitida para instantes n , para los cuales se han detectado componentes transitorios en el codificador. Los valores $\Delta\phi$ son transmitidos como índices de un cuantificador con una resolución de, por ejemplo, 3 bits por muestra.

[0088] Otro parámetro de gestión de transitorio que va a ser transmitido es la intensidad de separación (es decir, la intensidad de efecto del esquema de gestión de transitorios). Esta información es transmitida a alguna resolución temporal como los parámetros espaciales ILD, ICC.

[0089] La cantidad de bits transmitidos BR necesaria para transmitir decisiones de separación de transitorios e información de fase de banda ancha desde el codificador al decodificador, puede ser estimada por sistemas de tipo MPS según:

25

$$BR = BR_{\text{banderas de separación de transitorios}} + BR_{\Delta\phi} \approx (f_s / 64) + \sigma \cdot Q \cdot f_s / 64 = (1 + \sigma \cdot Q) \cdot f_s / 64$$

donde σ es la densidad de transitorios (fracción ranuras de tiempo (=muestras de tiempo de sub-banda) que están marcados como transitorios), Q es el número de bits por valor de fase transmitido, y f_s es la tasa de muestreo. Cabe señalar que $(f_s/64)$ es la tasa de muestreo de las señales sub-banda muestreadas de forma descendente.

[0090] Se ha medido $E\{\sigma\} < 0,25$ para un conjunto de diversos elementos de aplauso representativos, donde $E\{\cdot\}$ indica la media sobre la duración del elemento. Un compromiso razonable entre exactitud de los valores de fase y cantidad de bits transmitidos de parámetro, es $Q=3$. Para reducir la tasa de datos de parámetro, los ICCs y los ILDs pueden ser transmitidos como indicios de banda ancha (broadband cues). La transmisión de los ICCs y los ILDs como indicios de banda ancha es especialmente aplicable para señales no tonales como aplauso.

[0091] Adicionalmente, los parámetros para señalar la intensidad de separación son transmitidos a la tasa de actualización de los ICCs/ILDs. Para cuadros espaciales largos en MPS (32 veces 64 muestras) e intensidades de separación cuantificadas en 4 escalones, este da como resultado una cantidad de bits transmitidos adicional de

40

$$BR_{\text{intensidad de separación de transitorios}} = (f_s / (64 \cdot 32)) \cdot 2$$

[0092] El parámetro de intensidad de separación puede ser establecido en el codificador a partir de los resultados de algoritmos de análisis de señal que evalúan la similitud a señales de tipo aplauso, la tonalidad, u otras características de señal que indican beneficios o problemas potenciales cuando se aplica la decorrelación de componentes transitorios de la forma de realización.

[0093] Los parámetros transmitidos para gestión de transitorios pueden ser sometidos a codificación sin pérdidas para reducir la redundancia, dando como resultado una menor cantidad de bits transmitidos de parámetro (por ejemplo, codificación por longitud de series de información de separación de transitorios, codificación entrópica).

[0094] Volviendo al aspecto de obtención de información de fase, en una forma de realización, la información de fase puede obtenerse en un decodificador.

55

[0095] En tal forma de realización, el aparato para decodificación no obtiene información de fase de un

codificador, sino que puede determinar la información de fase por sí mismo. Por lo tanto, no es necesario transmitir información de fase lo que resulta en una tasa reducida de transmisión global.

[0096] En una forma de realización se obtiene información de fase en un decodificador basado en MPS a partir de datos de "Modulación de Envolvente Guiada (GES)" (Guided Envelope Shaping) Esto solo es aplicable si se transmiten datos GES, es decir, si la opción GES está activada en un codificador. La opción GES está disponible por ejemplo, en sistemas MPS. El cociente de valores de envolvente GES entre los canales de salida refleja posiciones de paneo para los componentes transitorios a alta resolución temporal. El cociente de envolvente GES (GESR) puede ser mapeado a la información de fase necesaria para la gestión de componentes transitorios. En GES, el mapeo puede ser realizado según una regla de mapeo obtenida empíricamente de estadísticas acumulativas de distribución de fase relativa a GESR para un conjunto representativo de señales de prueba apropiadas. Determinar la regla de mapeo es un paso para diseñar el sistema de gestión de componentes transitorios, no cuando se ejecuta el sistema de gestión de componentes transitorios. Por lo tanto, es ventajoso que no hay necesidad de pagar costes adicionales de transmisión para datos de fase si se necesitan datos DES para la aplicación de la opción GES de alguna manera. Compatibilidad hacia atrás de serie de bits en el tiempo se logra con series de bits en el tiempo/decodificadores MPS. Sin embargo, la información de fase extraída de datos GES no es exacta (por ejemplo, el signo de la fase estimada es desconocido) como la información de fase que podría ser obtenida en el codificador.

[0097] En una forma de realización adicional, la información de fase puede ser obtenida también en un decodificador, pero a partir de residuales de banda no completa transmitidas. Esto es aplicable, por ejemplo, si se transmiten señales residuales de banda limitada (típicamente cubriendo un rango de frecuencia hasta una cierta frecuencia de transición) en un esquema de codificación MPS. En tal forma de realización, se calcula la relación de fase entre la señal de mezcla descendente y la residual transmitida en la(s) banda(s) residual(es), es decir, para frecuencias para las cuales son transmitidas las señales residuales. Además, la información de fase procedente de banda(s) residual(es) a la(s) banda(s) no-residual(es) es extrapolada (y/o posiblemente interpolada). Una posibilidad es mapear la relación de fase obtenida en la(s) banda(s) residual(es) a un valor de la relación de fase independiente de la frecuencia que se usa a continuación para el decorrelacionador transitorio. Esto da como resultado el beneficio de que no aparecen costes adicionales de transmisión para datos de fase, si de alguna manera se transmiten residuales de banda no completa. Sin embargo, se debe considerar que la corrección de la estimación de fase depende del ancho de la(s) banda(s) de frecuencia donde se transmiten las señales residuales. La corrección de las estimaciones de fase también depende de la consistencia de la relación de fase entre la señal de mezcla descendente y la residual a lo largo del eje de frecuencia. Para señal con componentes claramente transitorios, usualmente se encuentra alta consistencia.

[0098] En una forma de realización adicional, la información de fase es obtenida en un decodificador empleando información de corrección adicional transmitida desde el codificador. Una forma de realización así es similar a las dos formas de realización anteriores (fase de GES, fase de residuales) pero adicionalmente es necesario generar datos de corrección en el codificador que se transmiten al decodificador. Los datos de corrección permiten reducir el error de estimación de fase que puede ocurrir en las dos variantes descritas antes (fase de GES, fase de residuales). Asimismo, los datos de corrección pueden ser derivados de estimar el error de estimación de fase del lado decodificador en el codificador. Los datos de corrección pueden ser este error de estimación estimado (potencialmente codificado). Asimismo, con respecto a la estrategia de estimación de fase de datos de GES, los datos de corrección simplemente pueden ser el signo correcto de los valores de fase generados por codificador. Esto permite generar términos de fase con el signo correcto en el decodificador. El beneficio de tal estrategia es que debido a los datos de corrección, la exactitud de la información de fase recuperable en el decodificador está mucho más cerca de la información de fase generada en el codificador. Sin embargo, la entropía de la información de corrección es menor que la entropía de la información de fase correcta en sí misma. Así, se disminuye la cantidad de bits transmitidos de parámetros cuando se compara con transmitir directamente la información de fase obtenida en el codificador.

[0099] En otra forma de realización, se obtiene información/término de fase de un proceso (pseudo-) aleatorio en un decodificador. El beneficio de tal estrategia es que no hay necesidad de transmitir información de fase alguna con alta resolución temporal. Esto da como resultado una tasa reducida de datos. En una forma de realización, un procedimiento simple es generar valores de fase con una distribución aleatoria uniforme en un intervalo $[-180^\circ, 180^\circ]$.

[0100] En una forma de realización adicional se miden las propiedades estadísticas de la distribución de fase en el codificador. Estas propiedades son codificadas y transmitidas a continuación (a baja resolución temporal) al decodificador. Se generan valores de fase aleatorios en el decodificador los cuales son sometidos a las propiedades

estadísticas transmitidas. Estas propiedades podrían ser la media, varianzas u otras medidas estadísticas de la distribución de fase estadística.

[0101] Cuando se ejecuta más de una instancia de decorrelacionador en paralelo (por ejemplo, para un
5 mezclado ascendente multicanal), se debe tener cuidado para asegurar salidas de decorrelacionador mutuamente decorrelacionadas. En una forma de realización, en la que se generan múltiples vectores de valores de fase (pseudo-) aleatorios (en lugar de un solo vector) para todas menos para la primera instancia de decorrelacionador, se selecciona un conjunto de vectores que da como resultado la menor correlación del valor de fase a través de todas las instancias de decorrelacionador.

10

[0102] En caso de transmitir información de corrección de fase desde un codificador al decodificador, se puede reducir la tasa de datos requerida del siguiente modo:

La información de corrección de fase solo necesita estar disponible en el decodificador siempre que haya componentes transitorios en la señal que vayan a ser decorrelacionados. La transmisión de la información de corrección de fase puede ser limitada entonces por el codificador, de tal forma que solo la información necesaria sea transmitida al decodificador. Esto se puede hacer aplicando una detección de transitorios en el codificador como se ha descrito más arriba. La información de corrección de fase solo es transmitida para puntos en el tiempo n , para los cuales se han detectado componentes transitorios en el codificador.

15

20 **[0103]** Volviendo al aspecto de separación transitoria, en una forma de realización, la separación transitoria puede ser conducida por el decodificador.

[0104] En tal forma de realización, también se puede obtener información de separación transitoria en el decodificador, por ejemplo, aplicando un procedimiento de detección de componente transitorio como se describe en

25

"Uso de Supresión de Transitorios en Algoritmos de Mezcla ascendente Multicanal Ciega" ("Using Transient Suppression in Blind Multi-channel Up-mix Algorithms") de Andreas Walther, Christian Uhle, Sascha Disch, en Proc. 122^o Convención de AES, Viena, Austria, mayo de 2007, para la señal de mezcla descendente que está disponible en el decodificador de audio espacial antes de mezclar de forma ascendente una señal de salida estéreo o multicanal. En este caso, no se tiene que transmitir información de componente transitorio alguna, lo cual ahorra
30 tasa de datos de transmisión.

[0105] Sin embargo, realizar la detección de componente transitorio al decodificar podría causar problemas cuando, por ejemplo se estandariza el esquema de gestión de componentes transitorios: por ejemplo, podría ser

35

difícil encontrar un algoritmo de detección de componentes transitorios que dé como resultado exactamente la misma detección de componentes transitorios cuando se implementa en diferentes plataformas/arquitecturas que involucran diferentes precisiones numéricas, esquemas de redondeo, etc. Para estandarización, tal comportamiento predecible del decodificador es con frecuencia imprescindible. Asimismo, el algoritmo de detección de componentes transitorios estandarizado podría fallar para algunas señales de entrada, causando distorsiones intolerables en las señales de salida. Entonces podría ser difícil corregir el algoritmo que falla después de la estandarización sin
40 construir un decodificador que no sea conforme al estándar. Esta cuestión podría ser menos severa si al menos se transmitiera un parámetro que controla la separación transitoria a baja resolución temporal (por ejemplo, a la tasa de actualización de parámetro espacial del MPS) desde el codificador al decodificador.

[0106] En una forma de realización adicional, la separación transitoria también es conducida por

45

decodificador y se transmiten residuales de banda no completa. En esta forma de realización, la separación transitoria conducida por decodificador puede ser refinada empleando estimaciones de fase obtenidas de residuales de banda no-completa transmitidas (véase más arriba). Cabe señalar que este refinamiento puede ser aplicado en el decodificador sin transmitir datos adicionales desde el codificador al decodificador.

[0107] En esta forma de realización, los términos de fase que son aplicados a un decorrelacionador

50

transitorio son obtenidos extrapolando los valores de fase correctos desde las bandas residuales a frecuencias donde no hay residuales disponibles. Un procedimiento es calcular un valor medio de fase (potencialmente por ejemplo, ponderado por potencia de señal) de los valores de fase que pueden ser calculados para aquellas frecuencias donde hay señales residuales disponibles. El valor medio de fase puede ser aplicado como un
55 parámetro independiente de la frecuencia en el decorrelacionador transitorio.

[0108] Siempre que la relación de fase correcta entre la mezclada de forma descendente y la residual es independiente de la frecuencia, el valor medio de fase representa una buena estimación del valor correcto de fase.

60

Sin embargo, en el caso de una relación de fase que no es consistente a lo largo del eje de frecuencia, el valor medio de fase puede ser una estimación menos correcta, posiblemente que conduce a artefactos audibles y valores

de fase incorrectos.

[0109] La consistencia de la relación de fase entre la mezcla de forma descendente y la residual transmitida a lo largo del eje de frecuencia, por lo tanto, puede ser usada como una medida de fiabilidad de la estimación de fase extrapolada que se aplica en el decorrelacionador transitorio. Para bajar el riesgo de artefactos audibles, se puede usar la medida de consistencia obtenida en el decodificador para controlar la intensidad de separación transitoria en el decodificador, por ejemplo, del siguiente modo:

Los componentes transitorios para los cuales la correspondiente información de fase (es decir, información de fase para el mismo índice de tiempo n) es consistente a lo largo de la frecuencia, son totalmente separados de la entrada de decorrelacionador convencional y son totalmente alimentados al decorrelacionador transitorio. Dado que grandes errores de estimación de fase son improbables, se usa todo el potencial de la gestión de componentes transitorios.

[0110] Los componentes transitorios para los cuales la correspondiente información de fase es menos consistente a lo largo de la frecuencia, solo son separados parcialmente, conduciendo a un efecto menos prominente del esquema de gestión de componentes transitorios.

[0111] Los componentes transitorios para los cuales la correspondiente información de fase es muy inconsistente a lo largo de la frecuencia, no son separados, conduciendo al comportamiento estándar de un sistema de mezclado ascendente convencional sin la gestión de componentes transitorios propuesta. Así, no pueden ocurrir artefactos debido a grandes errores de estimación de fase.

[0112] Las medidas de consistencia para la información de fase pueden ser deducidas, por ejemplo, a partir de la varianza (posiblemente ponderada por potencia de señal) de la desviación estándar de la información de fase a lo largo de la frecuencia.

[0113] Dado que solo unas pocas frecuencias pueden estar disponibles para las cuales se transmiten las señales residuales, la medida de la consistencia puede tener que ser estimada solo a partir de unas pocas muestras a lo largo de la frecuencia, conduciendo a una medida de la consistencia que solo rara vez alcanza valores extremos ("perfectamente consistente" o "perfectamente inconsistente"). De este modo, la medida de la consistencia puede ser linealmente o no linealmente distorsionada antes de ser usada para controlar la intensidad de separación transitoria. En una forma de realización, se implementa un umbral característico como se ilustra en la Figura 8, ejemplo de la derecha.

[0114] La Figura 8 representa diferentes mapeos ejemplares de medidas de consistencia de fase con respecto a intensidad de separación transitoria, ilustrando el impacto de las variantes para obtener parámetros de gestión de componentes transitorios sobre la robustez a mala clasificación de componentes transitorios. Las variantes para obtener la información de separación transitoria y la información de fase listada más arriba difieren en tasa de datos de parámetro y, por lo tanto, representan diferentes puntos operativos en términos de cantidad de bits transmitidos global de un codificador-decodificador que implementa la técnica de gestión de componentes transitorios propuesta. Aparte de esto, la elección de la fuente para obtener la información de fase también afecta a aspectos tales como robustez a clasificaciones de componentes transitorios falsas: gestionar una señal no transitoria como una transitoria causa muchas menos distorsiones audibles si la información de fase correcta es aplicada para gestionar el componente transitorio. Así, un error de clasificación de señal causa artefactos menos severos en el escenario de valores de fase transmitidos cuando se compara con el escenario de generación de fase aleatoria en el decodificador.

[0115] La Figura 9 es una vista global de un sistema Uno a Dos con gestión de componentes transitorios según otra forma de realización, en el que se transmiten señales residuales de banda estrecha. Se estima el dato de fase $\Delta\phi$ de la relación de fase entre la señal de mezcla descendente (DMX) y la residual en la(s) banda(s) de frecuencia de la señal residual. Opcionalmente, se transmiten datos de corrección de fase para reducir el error de estimación de fase.

[0116] La Figura 9 ilustra un separador transitorio 910, un decorrelacionador transitorio 920, un decorrelacionador reticular IIR 930, una unidad de combinación 940, un mezclador 952, una unidad de modelación opcional 954, una primera unidad sumadora 956 y una segunda unidad sumadora 958, los cuales corresponden al separador transitorio 510, el decorrelacionador transitorio 520, el decorrelacionador reticular IIR 530, la unidad de combinación 540, el mezclador 552 la unidad de modelación opcional 554, la primera unidad sumadora 556 y la segunda unidad sumadora 558 de la forma de realización de la Figura 5, respectivamente. La forma de realización de la Figura 8 comprende además una unidad de estimación de fase 960. La unidad de estimación de fase 960 recibe una señal de entrada DMX, una señal residual "residual" y, opcionalmente, datos de corrección de fase. En

base a la información recibida, la unidad de información de fase calcula los datos de fase $\Delta\phi$. Opcionalmente, la unidad de estimación de fase determina también información de consistencia de fase y pasa la información de consistencia de fase al separador transitorio 910. Por ejemplo, la información de consistencia de fase puede ser usada por el separador transitorio para controlar la intensidad de separación transitoria.

5

[0117] La forma de realización de la Figura 9 aplica el hallazgo de que si se transmiten residuales dentro del esquema de codificación de una manera de banda no-completa, la diferencia de fase media ponderada por potencia de señal entre la residual y la mezclada de forma descendente ($\Delta\phi_{\text{residual_bands}}$) puede ser aplicada como información de fase de banda ancha a los componentes transitorios separados ($\Delta\phi = \Delta\phi_{\text{low residual_bands}}$). En este caso, no hay que transmitir información de fase adicional, reduciendo la demanda de cantidad de bits transmitidos para la gestión de componentes transitorios. En la forma de realización de la Figura 9, la estimación de fase a partir de las bandas residuales puede desviarse considerablemente de la estimación de fase de banda ancha más precisa que está disponible en el codificador. Una opción es, por lo tanto, transmitir datos de corrección de fase (por ejemplo, $\Delta\phi_{\text{correction}} - \Delta\phi_{\text{residual_bands}}$) de modo que la $\Delta\phi$ correcta esté disponible en el decodificador. Sin embargo, como la corrección de $\Delta\phi$ puede mostrar una menor entropía que $\Delta\phi$, la tasa de datos de parámetros necesaria puede ser menor que la tasa que sería necesaria para transmitir $\Delta\phi$. (Este concepto es similar al uso general de predicción en codificación: en lugar de codificar directamente datos, se codifica un error de predicación con menor entropía. En la forma de realización de la Figura 9, el paso de predicción es la extrapolación de la fase a partir de las bandas de frecuencia de residuales con respecto a bandas de no residuales). La consistencia de la diferencia de fase en las bandas de frecuencia de residual ($\Delta\phi_{\text{residual_bands}}$) a lo largo del eje de frecuencia puede ser usada para controlar la intensidad de separación transitoria.

[0118] En formas de realización, un decodificador puede recibir información de fase de un codificador, o el decodificador puede determinar por sí mismo la información de fase. Asimismo, el decodificador puede recibir información de separación transitoria fase de un codificador, o el decodificador puede determinar por sí mismo la información de separación transitoria.

[0119] En formas de realización, un aspecto de la gestión de componentes transitorios es la aplicación del concepto de "decorrelación semántica" descrito en el documento WO/2010/017967 junto con el "decorrelacionador transitorio", el cual está basado en multiplicar la entrada con términos de fase. Se mejora la calidad perceptual de señales de tipo aplauso renderizadas ya que ambos pasos de procesamiento evitan alterar la estructura temporal de las señales con componente transitorio. Asimismo, la distribución espacial de componentes transitorios así como las relaciones de fase entre componentes transitorios, es reconstruida en los canales de salida. Asimismo, las formas de realización son también computacionalmente eficaces y pueden ser integradas fácilmente en sistemas de mezclado de forma ascendente de tipo PS o MPS. En formas de realización, la gestión de componentes transitorios no afecta al proceso de matriz de mezclado, de modo que todas las propiedades de renderización espacial que están definidas por la matriz de mezclado también son aplicadas a la señal con componente transitorio.

[0120] En formas de realización se aplica un esquema de decorrelación innovador el cual es particularmente adecuado para la aplicación en sistemas de mezclado de forma ascendente, el cual es particularmente adecuado para la aplicación de esquemas de codificación de audio espacial como PS o MPS y el cual mejora la calidad perceptual de las señales de salida en el caso de señales de tipo aplauso, es decir, señales que contienen mezclas densas de componentes transitorios distribuidos espacialmente y/o pueden ser vistas como una implementación particularmente mejorada del marco de trabajo genérico de "decorrelación semántica". Asimismo, en formas de realización, un esquema de decorrelación innovador reconstruye la distribución espacial/temporal de los componentes transitorios similar a la distribución en la señal original, preserva la estructura temporal de las señales transitorias, permite variar la cantidad de bits transmitidos versus el compromiso de calidad y/o es idealmente adecuado para una combinación con opciones MPS como residuales de banda no-completa o GES. Las combinaciones son complementarias, es decir, la información de opciones de MPS estándar se vuelve a utilizar para la gestión de componentes transitorios.

[0121] La Figura 10 ilustra un aparato para la codificación de una señal de audio que tiene una pluralidad de canales. Dos canales de entrada L, R son alimentados a un mezclador descendente 1010 y a un calculador de señal residual 1020. En otras formas de realización, una pluralidad de canales es alimentada al mezclador 1010 y al calculador de señal residual 1020, por ejemplo, 3, 5 o 9 canales envolventes. El mezclador descendente 1010 mezcla de forma descendente entonces los dos canales L, R, para obtener una señal de mezcla descendente. Por ejemplo, el mezclador descendente 1010 puede emplear una matriz de mezclado y realizar una multiplicación de matriz de la matriz de mezclado y los dos canales de entrada L, R, para obtener una señal de mezcla descendente. La señal de mezcla descendente puede ser transmitida a un decodificador.

60

- [0122]** Además, el generador de señal residual 1020 está adaptado para calcular una señal adicional a la cual se refiere como señal residual. Señales residuales son señales que pueden ser usadas para regenerar las señales originales empleando adicionalmente la señal de mezcla descendente y una matriz de mezcla ascendente. Por ejemplo, cuando N señales son mezcladas de forma descendente a 1 señal, el mezclado descendente típicamente es 1 de los N componentes que resultan del mapeo de las N señales de entrada. Los componentes restantes que resultan del mapeo (por ejemplo, N-1 componentes) son señales residuales y permiten reconstruir las N señales originales mediante un mapeo inverso. El mapeo puede ser, por ejemplo, una rotación. El mapeo debe ser conducido de tal forma que la señal de mezcla descendente sea maximizada y las señales residuales sean minimizadas, por ejemplo, similar a una transformación de eje principal. Por ejemplo, la energía de la señal de mezcla descendente debe ser maximizada y las energías de las señales residuales deben ser minimizadas. Cuando se mezclan de forma descendente dos señales a 1 señal, el mezclado descendente es normalmente uno de los dos componentes que resultan del mapeo de las 2 señales de entrada. El componente restante que resulta del mapeo es la señal residual y permite reconstruir las 2 señales originales mediante un mapeo inverso.
- [0123]** En algunos casos, la señal residual puede representar un error asociado con representar las dos señales mediante sus parámetros de mezclado descendente y asociados. Por ejemplo, la señal residual puede ser una señal de error la cual representa el error entre los canales originales L, R y los canales L', R', resultantes de mezclar de forma ascendente la señal de mezcla descendente que fue generada en base a los canales originales L y R.
- [0124]** En otras palabras, una señal residual puede ser considerada como una señal en el dominio del tiempo o un dominio de la frecuencia o un dominio sub-banda, que junto con la señal de mezcla descendente sola o con la señal de mezcla descendente y la información paramétrica permite una reconstrucción correcta o casi correcta de un canal original. Casi correcto se debe entender como que la reconstrucción con la señal residual teniendo una energía mayor que cero es más cercana al canal original en comparación con una reconstrucción mediante el uso de la mezcla descendente sin la señal residual o mediante el uso de la mezcla descendente y la información paramétrica sin la señal residual.
- [0125]** Además, el codificador comprende un calculador de información de fase 1030. La señal de mezcla descendente y la señal residual son alimentadas al calculador de información de fase 1030. El calculador de información de fase calcula entonces información sobre la diferencia de fase entre la señal de mezcla descendente y la residual para obtener información de fase. Por ejemplo, el calculador de información de fase puede aplicar funciones que calculan una correlación cruzada de la señal de mezcla descendente y la residual.
- [0126]** Asimismo, el codificador comprende un generador de salida 1040. La información de fase generada por el calculador de información de fase 1030 es alimentada en el generador de salida 1040. El generador de salida 1040 entrega entonces la información de fase.
- [0127]** En una forma de realización, el aparato comprende además un cuantificador de información de fase para cuantificar la información de fase. La información de fase generada por el calculador de información de fase puede ser alimentada en el cuantificador para cuantificar la información de fase. El cuantificador de información de fase cuantifica entonces la información de fase. Por ejemplo, la información de fase puede ser mapeada a 8 valores diferentes, por ejemplo a uno de los valores 0, 1, 2, 3, 4, 5, 6 o 7. Los valores pueden representar las diferencias de fases 0, $\pi/4$, $\pi/2$, $3\pi/4$, π , $5\pi/4$, $3\pi/2$ y $7\pi/4$, respectivamente. La información de fase cuantificada puede ser alimentada entonces al generador de salida 1040.
- [0128]** En una forma de realización adicional, el aparato comprende además un codificador sin pérdidas. La información de fase procedente del calculador de información de fase 1040 o la información de fase cuantificada procedente del cuantificador de información de fase, puede ser alimentada al codificador sin pérdidas. El codificador sin pérdidas está adaptado para codificar información de fase aplicando codificador sin pérdidas. Se puede emplear cualquier esquema de codificación sin pérdidas. Por ejemplo, el codificador puede emplear codificación aritmética. El codificador sin pérdidas alimenta entonces la información de fase codificada sin pérdidas al generador de salida 1040.
- [0129]** Con respecto al decodificador y codificador y los procedimientos de las formas de realización descritas, se menciona lo siguiente:
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 a un rasgo de un paso de procedimiento. Análogamente, los aspectos descritos en el contexto de una etapa de procedimiento representan también una descripción de un bloque o elemento

correspondiente o rasgo de un aparato correspondiente.

- [0130]** Dependiendo de ciertos requisitos de implementación, las formas de realización de la invención pueden ser implementadas 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 EPROM, una EEPROM o una memoria FLASH, los cuales tienen unas señales de control electrónicamente legibles guardadas en ellos, las cuales cooperan (o son capaces de cooperar) con un sistema de computación programable de modo que se ejecuta el respectivo procedimiento.
- 10 **[0131]** Algunas formas de realización según la invención comprenden un soporte de datos que tiene señales de control legibles electrónicamente, las cuales son capaces de cooperar con un sistema informático programable, de tal forma que uno de los procedimientos descritos en esta invención sea ejecutado.
- 15 **[0132]** Generalmente, las formas de realización de la presente invención pueden ser implementadas como un programa informático con un código de programa, siendo el código de programa operativo para ejecutar uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede ser almacenado, por ejemplo, en un soporte legible por una máquina.
- 20 **[0133]** Otras formas de realización comprenden el programa informático para ejecutar uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por una máquina o un medio de almacenamiento no transitorio.
- 25 **[0134]** En otras palabras, una forma de realización del procedimiento inventivo es, por lo tanto, un programa informático que tiene un código de programa para ejecutar uno de los procedimientos descritos en esta invención, cuando el programa informático se ejecuta en un ordenador.
- 30 **[0135]** Una forma de realización adicional de los procedimientos inventivos es, por lo 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 ejecutar uno de los procedimientos descritos en esta invención.
- 35 **[0136]** Una forma de realización adicional del procedimiento inventivo es, por lo tanto, una transmisión de datos o una secuencia de señales que representan el programa informático para ejecutar uno de los procedimientos descritos en esta invención. La transmisión de datos o la secuencia de señales pueden ser configuradas, por ejemplo, para ser transferidos vía una conexión de comunicación de datos, por ejemplo, vía Internet.
- 40 **[0137]** Una forma de realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado para o adaptado para ejecutar uno de los procedimientos descritos en esta invención.
- 45 **[0138]** Una forma de realización adicional comprende un ordenador que tiene instalado en él el programa informático para ejecutar uno de los procedimientos descritos en esta invención.
- [0139]** En algunas formas de realización se puede usar un dispositivo lógico programable (por ejemplo una matriz de puertas programable de campo) para realizar algunas o todas las funcionalidades de los procedimientos descritos en esta invención. En algunas formas de realización, compuertaza matriz de puertas programable de campo puede cooperar con un microprocesador para realizar uno de los procedimientos descritos en esta invención. Generalmente, los procedimientos son realizados preferiblemente mediante algún aparato de hardware.
- 50 **[0140]** Las formas de realización que se han descrito más arriba son meramente ilustrativas para los principios de la presente invención. Se entiende que las modificaciones y variaciones posibles de las disposiciones y de los detalles descritos en esta invención serán evidentes para los expertos en la materia. Por lo tanto, es la intención que la invención esté limitada solo por el alcance de las siguientes reivindicaciones de patente y no por los detalles específicos presentados a título de descripción y la explicación de las formas de realización en esta invención.
- 55

REIVINDICACIONES

1. Un aparato para la codificación de una señal de audio que tiene una pluralidad de canales, que comprende:
- 5 un mezclador descendente (1010) para el mezclado de forma descendente de la pluralidad de canales para obtener una señal de mezcla descendente;
- un calculador de señal residual (1020) adaptado para calcular una señal residual;
- un calculador de información de fase (1030) adaptado para calcular la información sobre una diferencia de fase entre la mezcla descendente y la señal residual para obtener información de fase; y
- 10 un generador de salida (1040) para generar la información de fase.
2. Un aparato para la codificación de una señal de audio según la reivindicación 1, en el que el aparato comprende además un cuantificador de información de fase para cuantificar la información de fase.
- 15 3. Un aparato para la codificación de una señal de audio según la reivindicación 1 o 2, en el que el aparato comprende además un codificador sin pérdidas adaptado para codificar la información de fase sin pérdidas mediante la aplicación de la codificación sin pérdidas.
4. Un procedimiento para la codificación de una señal de audio que tiene una pluralidad de canales, que
- 20 comprende:
- mezclado de forma descendente de la pluralidad de canales para obtener una señal de mezcla descendente;
- cálculo de una señal residual;
- cálculo de información sobre una diferencia de fase entre la mezcla descendente y la señal residual para obtener información de fase; y
- 25 generación de la información de fase.
5. Un procedimiento para la codificación de una señal de audio según la reivindicación 4, en el que el procedimiento comprende además la etapa de cuantificación de la información de fase.
- 30 6. Un procedimiento para la codificación de una señal de audio según la reivindicación 4 o 5, en el que el procedimiento comprende además la etapa de codificación de la información de fase sin pérdidas mediante la aplicación de la codificación sin pérdidas.
7. Un programa informático para la implementación del procedimiento según una de las reivindicaciones
- 35 4 a 6.

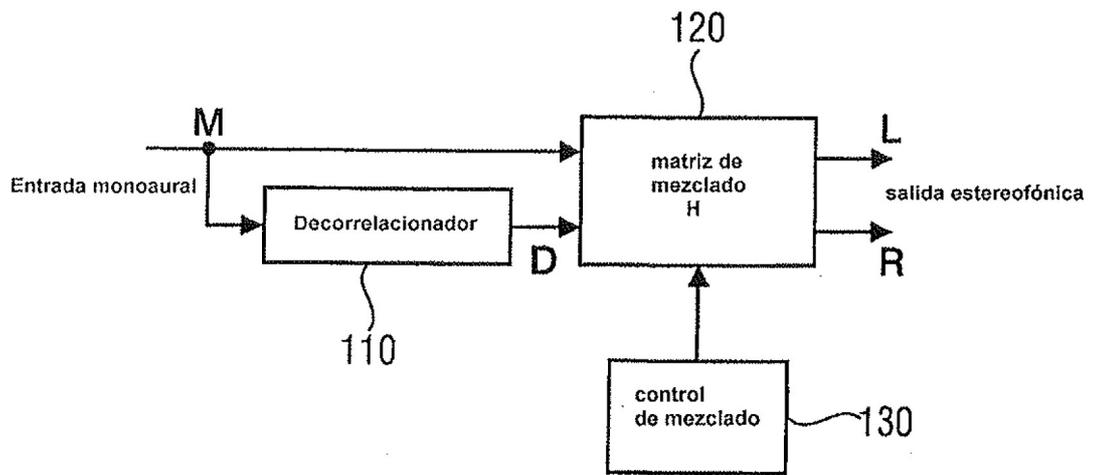


FIG 1
(ESTADO DE LA TÉCNICA)

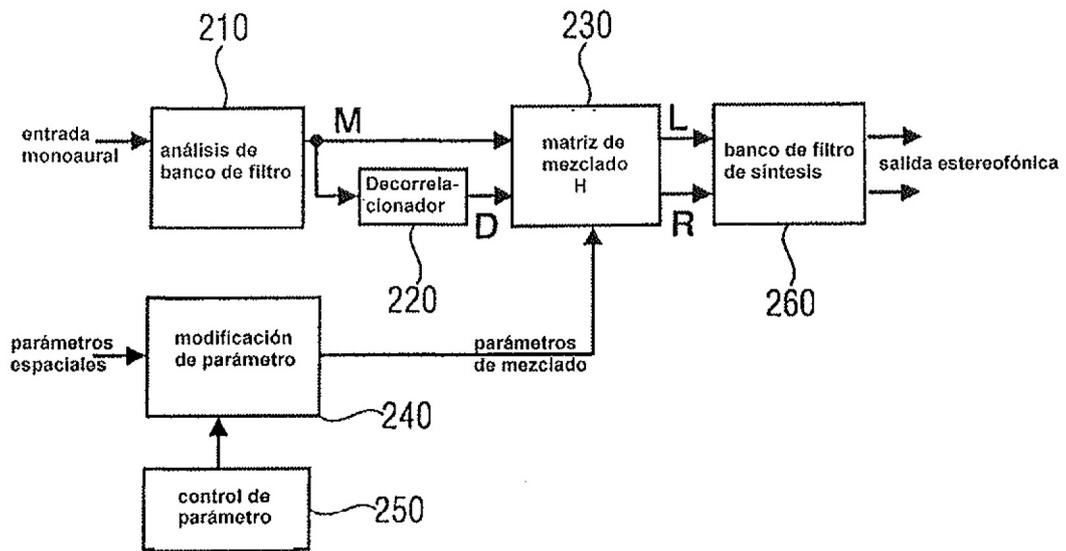
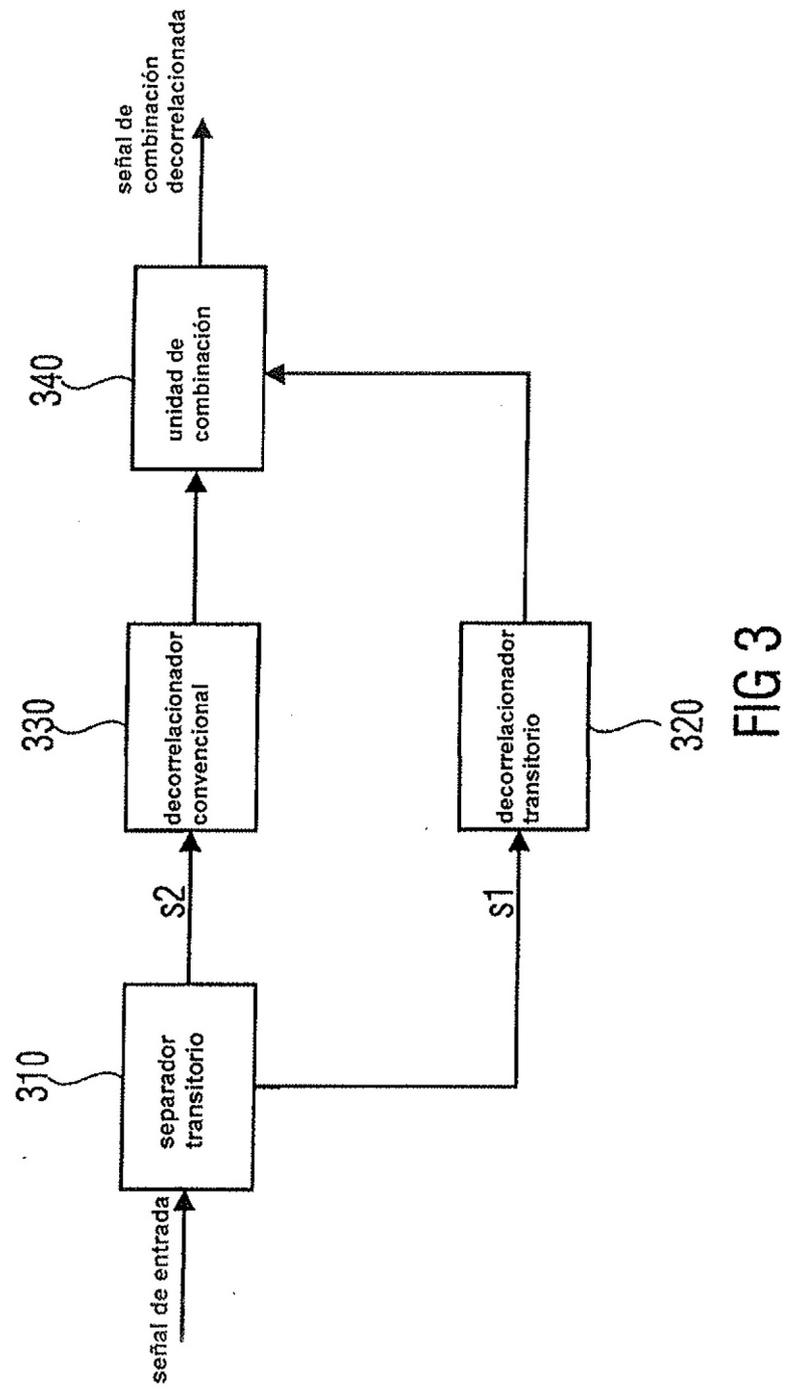


FIG 2
(ESTADO DE LA TÉCNICA)



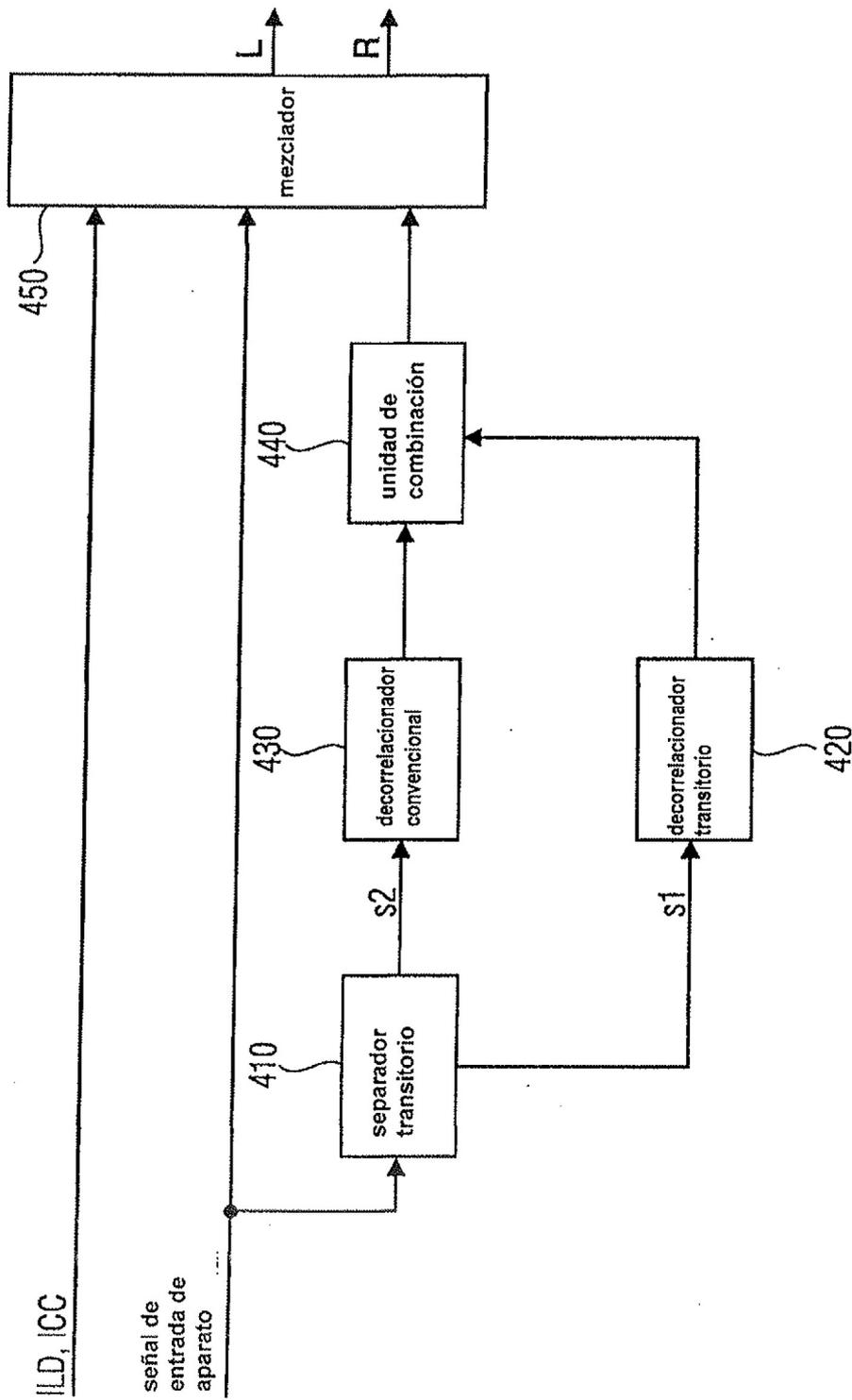


FIG 4

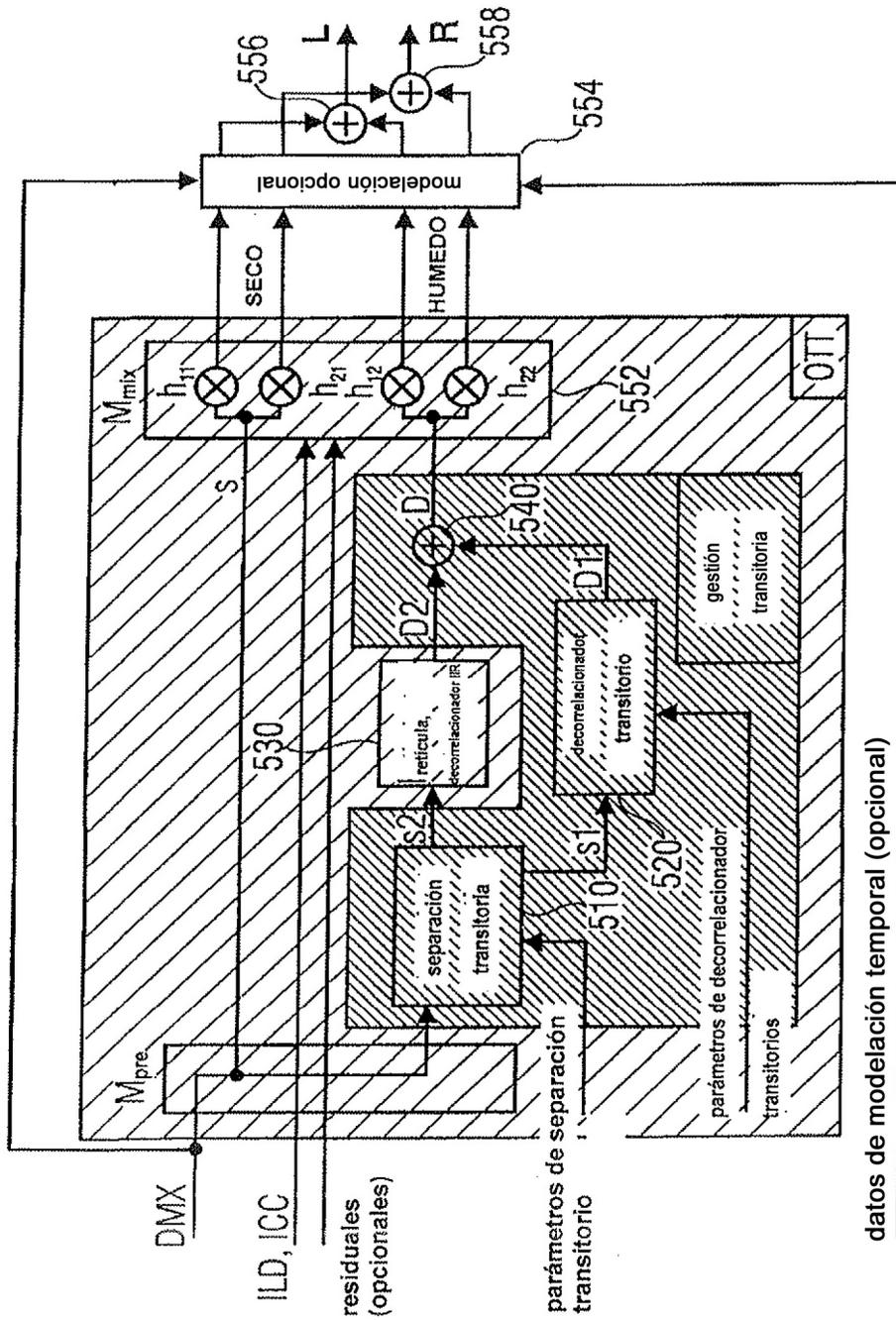


FIG 5

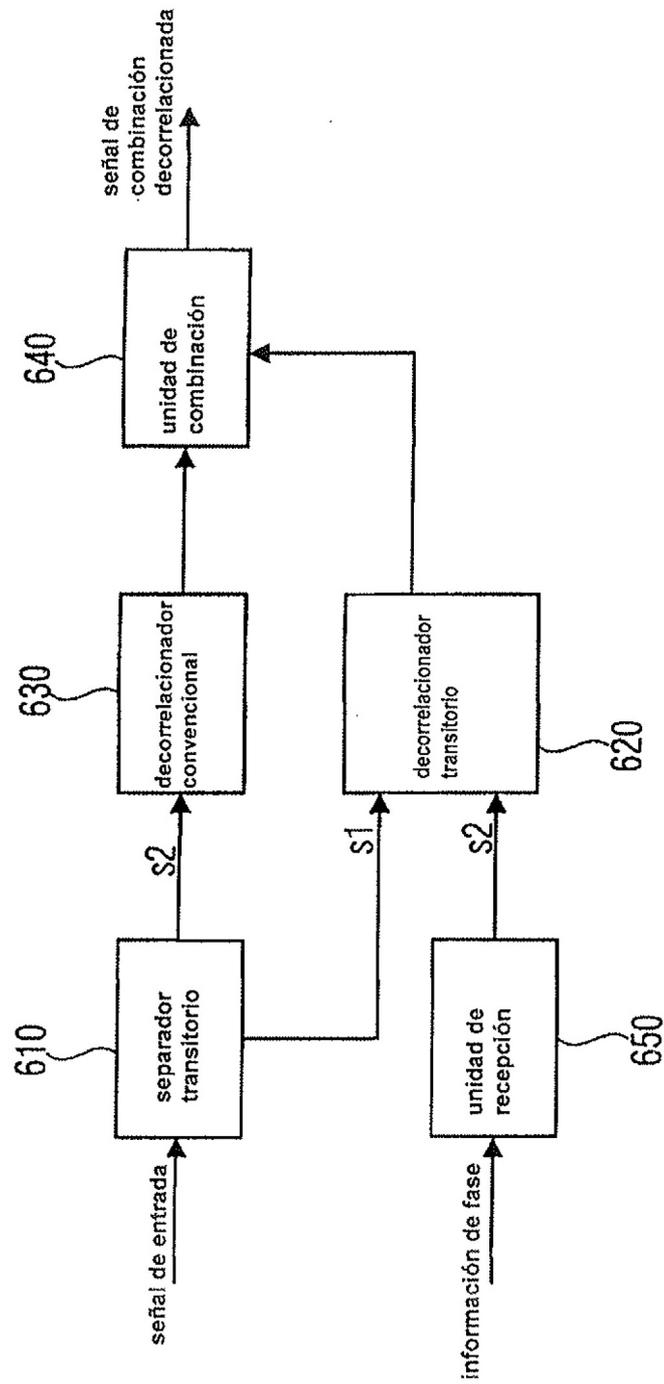


FIG 6

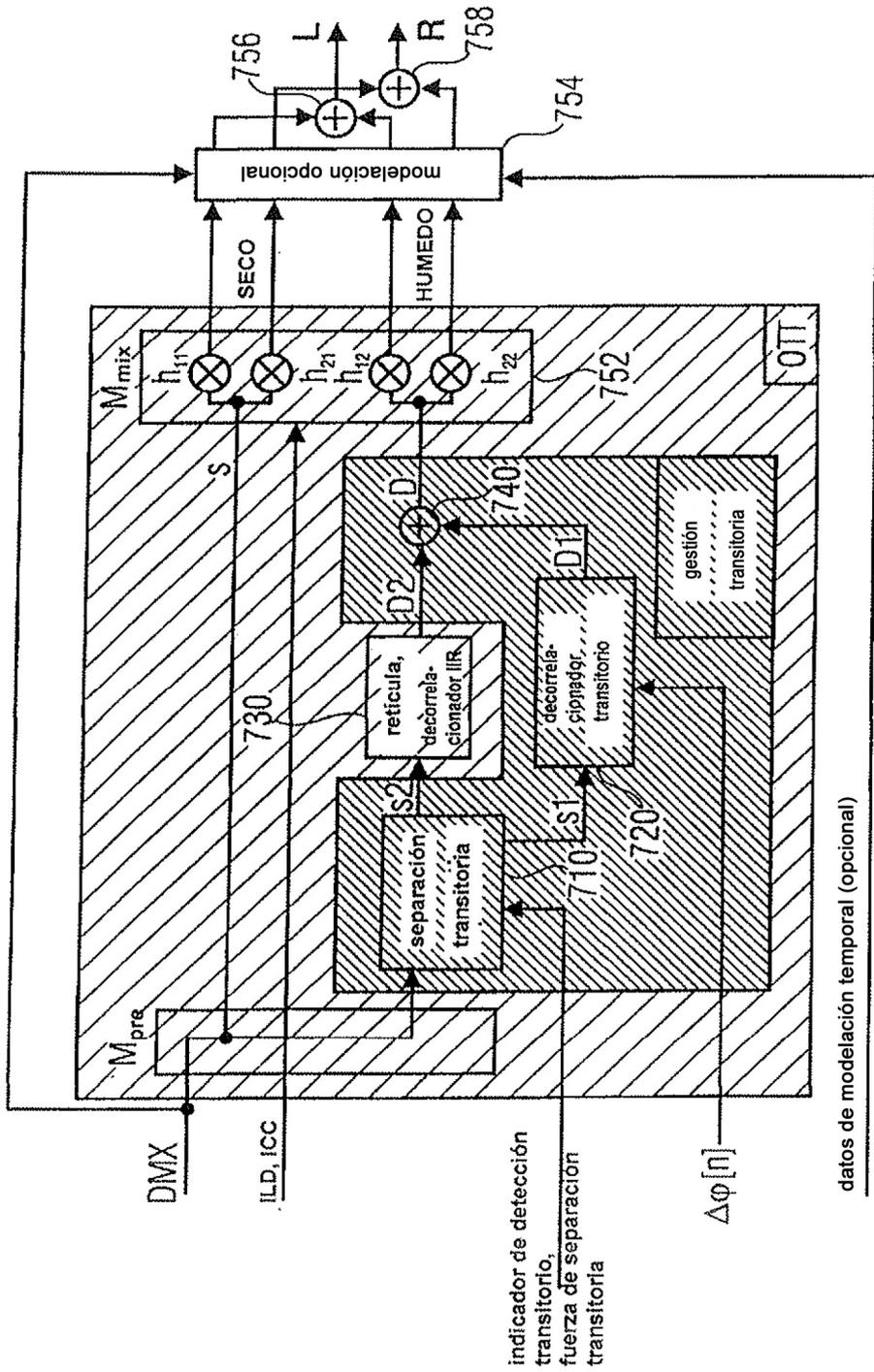


FIG 7

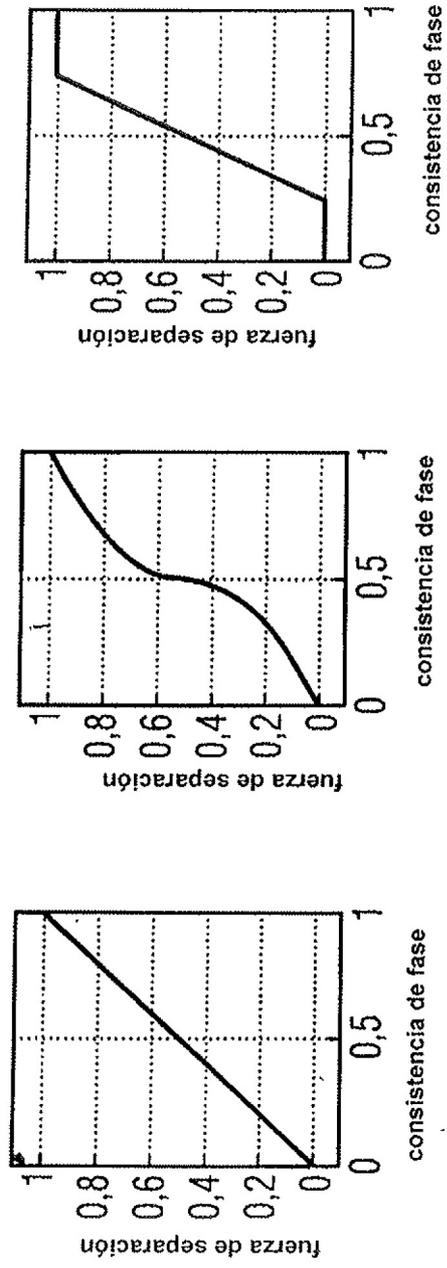


FIG 8

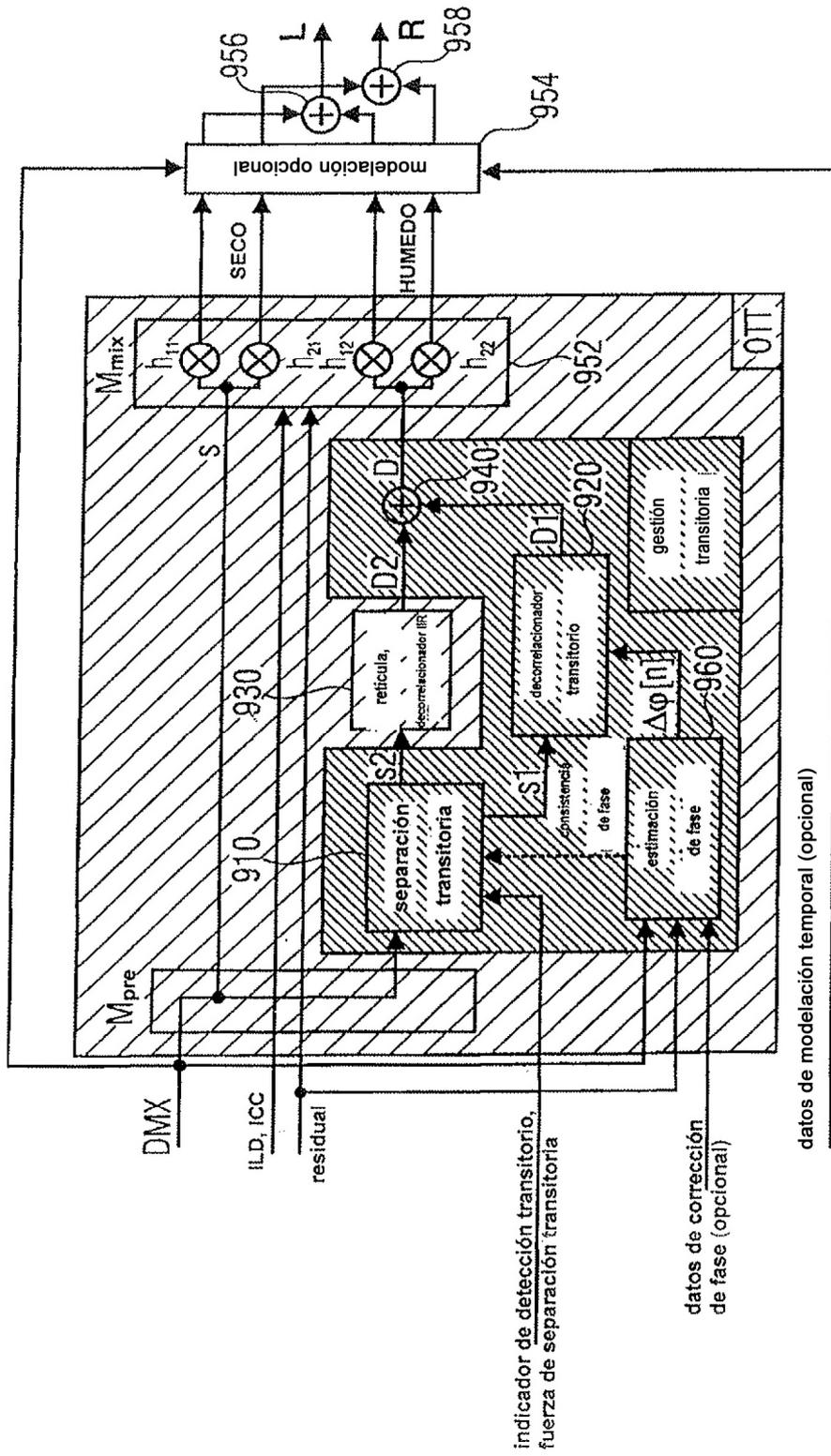


FIG 9

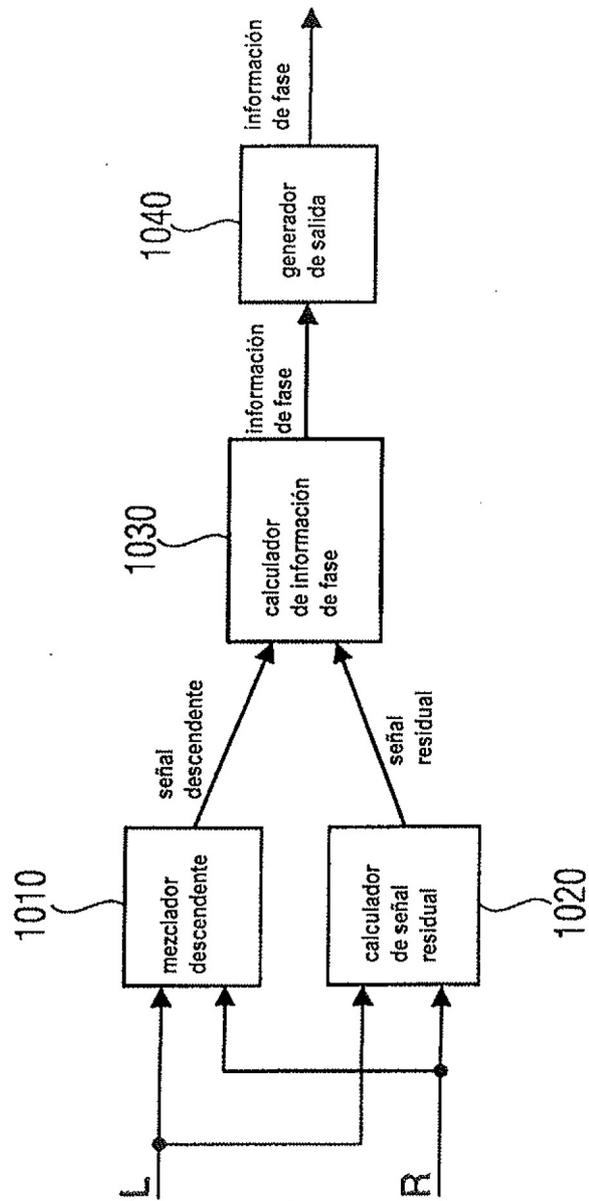


FIG 10