

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 650 549**

51 Int. Cl.:

**G10L 19/035** (2013.01)

**G10L 19/008** (2013.01)

**G10L 19/028** (2013.01)

**H04S 3/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **18.07.2014 PCT/EP2014/065550**

87 Fecha y número de publicación internacional: **29.01.2015 WO15011061**

96 Fecha de presentación y número de la solicitud europea: **18.07.2014 E 14744026 (7)**

97 Fecha y número de publicación de la concesión europea: **06.09.2017 EP 3025341**

54 Título: **Relleno de ruido en codificación de audio multicanal**

30 Prioridad:

**22.07.2013 EP 13177356**  
**18.10.2013 EP 13189450**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**19.01.2018**

73 Titular/es:

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

72 Inventor/es:

**LUIS VALERO, MARIA;  
HELMRICH, CHRISTIAN y  
HILPERT, JOHANNES**

74 Agente/Representante:

**SALVA FERRER, Joan**

ES 2 650 549 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**

Relleno de ruido en codificación de audio multicanal

- 5 **[0001]** La presente solicitud se refiere al relleno de ruido en la codificación de audio multicanal.
- [0002]** Los sistemas modernos de codificación de voz/audio en el dominio de la frecuencia como el códec Opus/Celt de IETF [1], MPEG-4 (HE-)AAC [2] o, en particular, MPEG-D xHE-AAC (USAC) [3], ofrecen medios para codificar tramas de audio que utilizan una transformada larga – un bloque largo – u ocho transformadas cortas  
 10 secuenciales – bloques cortos – en función de la estacionalidad temporal de la señal. Además, para la codificación a bajas tasas de bits estos esquemas ofrecen herramientas para reconstruir los coeficientes de frecuencia de un canal que utiliza un ruido pseudoaleatorio o coeficientes de menor frecuencia del mismo canal. En xHE-AAC, estas herramientas se conocen como relleno de ruido y replicación de banda espectral, respectivamente.
- 15 **[0003]** Sin embargo, para una entrada estereofónica muy tonal o transitoria, el relleno de ruido y/o la replicación de banda espectral por si solas limitan la calidad de la codificación que se consigue a tasas de bits muy bajas, principalmente porque se necesitan transmitir demasiados coeficientes espectrales de ambos canales explícitamente.
- [0004]** El documento US 2013/013321 A1 describe un procedimiento de procesamiento de una señal de audio.  
 20 El procedimiento de procesamiento de una señal de audio comprende: recibir, mediante un aparato de procesamiento de audio, los datos espectrales que incluyen un bloque actual, y una información de tipo de sustitución que indica si aplicar un esquema de predicción de forma a un bloque actual, si la información de tipo de sustitución indica que el esquema de predicción de forma se aplica al bloque actual, recibir información de retardo que indica un intervalo entre los coeficientes espectrales del bloque actual y el vector de forma predictiva de una trama actual o una trama anterior,  
 25 obtener los coeficientes espectrales sustituyendo el agujero espectral incluido en el bloque actual que utiliza el vector de forma predictiva.
- [0005]** El documento US 2009/006103 A1 describe un decodificador de audio que proporciona una combinación de componentes de decodificación que incluye componentes que implementan una decodificación de banda base,  
 30 decodificación de pico espectral, decodificación de extensión de frecuencia y técnicas de decodificación de extensión de canal. El decodificador de audio decodifica un flujo de bits comprimido estructurado en un esquema de sintaxis de flujo de bits que permite que los diversos componentes de decodificación extraigan los parámetros adecuados para su técnica de decodificación respectiva.
- 35 **[0006]** Por lo tanto, el objetivo consiste en proporcionar un concepto para realizar el relleno de ruido en la codificación de audio multicanal que proporciona una codificación más eficaz, especialmente a tasas de bits muy bajas.
- [0007]** El objetivo se consigue con la materia objeto de las reivindicaciones independientes adjuntas.
- 40 **[0008]** La presente solicitud está basada en el hallazgo de que en una codificación de audio multicanal se puede conseguir una mayor eficacia de codificación si el relleno de ruido de las bandas de factor de escala cuantizadas a cero de un canal se realiza utilizando fuentes de relleno de ruido y no ruidos generados artificialmente o réplicas espectrales del mismo canal. En particular, la eficacia en la codificación de audio multicanal puede ser más eficaz si se realiza el relleno de ruido basado en el ruido generado a partir de las líneas espectrales de una trama anterior de,  
 45 o un canal diferente de la trama actual de, la señal de audio multicanal.
- [0009]** Si se utilizan líneas espectrales de una trama anterior cubricadas espectralmente, o líneas espectrales cubricadas espectralmente de otros canales de la señal de audio multicanal, es posible conseguir una calidad más agradable de la señal de audio multicanal reconstruida, especialmente a tasas de bits muy bajas en las que el  
 50 requisito del codificador respecto a una cuantización cero de las líneas espectrales es próximo a una situación en la que se cuantizan a cero las bandas de factor de escala en su totalidad. Gracias a un mejor relleno de ruido, un codificador puede, con menor penalización sobre la calidad, elegir cuantizar a cero las bandas de factor de escala, mejorando así la eficacia de la codificación.
- 55 **[0010]** De acuerdo con una realización de la presente solicitud, la fuente para realizar el relleno de ruido se superpone parcialmente con una fuente utilizada para realizar una predicción estéreo de valor complejo. En particular, la mezcla descendente de una trama anterior puede utilizarse como fuente para el relleno de ruido y coutilizarse como fuente para realizar, o al menos aumentar, la estimación de la parte imaginaria que realiza la predicción compleja entre canales.

**[0011]** De acuerdo con realizaciones, un códec de audio multicanal existente se extiende de modo compatible hacia atrás para señalar, en una base de trama por trama, el uso de relleno de ruido entre canales. Las realizaciones descritas a continuación, por ejemplo, extienden xHE-AAC mediante una señalización de manera compatible hacia atrás, con la señalización de una conmutación activada/desactiva del relleno de ruido entre canales y que aprovecha los estados no utilizados del parámetro de relleno de ruido codificado condicionalmente.

**[0012]** Las implementaciones ventajosas de la presente solicitud son la materia de las reivindicaciones dependientes. Las realizaciones preferentes de la presente solicitud se describen a continuación con respecto a las figuras, entre las cuales:

la Fig. 1 muestra un diagrama de bloques de un decodificador paramétrico en el dominio de la frecuencia según una realización de la presente solicitud;

la Fig. 2 muestra un diagrama esquemático que ilustra la secuencia de espectros que forman los espectrogramas de canales de una señal de audio multicanal con el fin facilitar la comprensión de la descripción del decodificador de la Fig. 1;

la Fig. 3 muestra un diagrama esquemático que ilustra los espectros actuales fuera de los espectrogramas que se muestran en la Fig. 2 en aras de satisfacer la comprensión de la descripción de la Fig. 1;

La Fig. 4 muestra un diagrama de bloques de un decodificador de audio paramétrico en el dominio de la frecuencia según una realización alternativa según la cual la mezcla descendente de la trama anterior se utiliza como base para el relleno de ruido entre canales; y

La Fig. 5 muestra un diagrama de bloques de un codificador de audio paramétrico en el dominio de la frecuencia según una realización.

**[0013]** La Fig. 1 muestra un decodificador de audio en el dominio de la frecuencia según una realización de la presente solicitud. El decodificador generalmente se indica con el signo de referencia 10 y comprende un identificador de banda de factor de escala 12, un decuantizador 14, un rellenador de ruido 16 y un transformador inverso 18 así como un extractor de líneas espectrales 20 y un extractor de factor de escala 22. Otros elementos que podría comprender el decodificador 10 incluyen un predictor estéreo complejo 24, un decodificador MS (tamaño medio) 26 y una herramienta de filtro TNS inverso (modelado del ruido temporal) de la cual se muestran dos ejemplos 28a y 28b en la Fig. 1. Además, se muestra y se describe un proveedor de mezcla descendente en detalle a continuación con el signo de referencia 30.

**[0014]** El decodificador de audio en el dominio de la frecuencia 10 de la Fig. 1 es un decodificador paramétrico que permite el relleno de ruido según el cual una determinada banda de factor de escala cuantizada a cero se rellena con ruido que utiliza un factor de escala de la banda de factor de escala como un medio para controlar el nivel de ruido de relleno en la banda de factor de escala. Aparte de esto, el decodificador 10 de la Fig. 1 representa un decodificador de audio multicanal configurado para reconstruir una señal de audio multicanal a partir un flujo de datos de entrada 30. La Fig. 1, sin embargo, se concentra en los elementos del decodificador 10 que intervienen en la reconstrucción de una de las señales de audio multicanal codificadas en el flujo de datos 30 y emite este canal (salida) en una salida 32. Un signo de referencia 34 indica que un decodificador 10 puede comprender otros elementos o puede comprender determinada canalización del control de operaciones responsable de reconstruir los demás canales de la señal de audio multicanal en la que la descripción presentada a continuación indica cómo la reconstrucción del decodificador 10 del canal de interés en la salida 32 interactúa con la decodificación de los demás canales.

**[0015]** La señal de audio multicanal representada por el flujo de datos 30 puede comprender dos o más canales. A continuación, la descripción de las realizaciones de la presente solicitud se concentra en el caso estéreo en el que la señal de audio multicanal sólo comprende dos canales, pero en principio las realizaciones presentadas a continuación pueden transferirse a realizaciones alternativas con respecto a las señales de audio multicanal y su codificación que comprende más de dos canales.

**[0016]** Como resulta claro de la descripción de la Fig. 1 a continuación, el decodificador 10 de la Fig. 1 es un decodificador por transformada. Es decir, según el decodificador subyacente de la técnica de codificación 10, los canales se codifican en un dominio de transformada utilizando, por ejemplo, una transformada con superposición de canales. Asimismo, en función del creador de la señal de audio, existen fases de tiempo durante las cuales los canales

de la señal de audio representan en gran medida el mismo contenido de audio, desviándose entre sí sólo por cambios menores o determinísticos entre ellos tal como diferentes amplitudes y/o fase con el fin de representar una escena de audio en la cual las diferencias entre los canales permiten el posicionamiento virtual de una fuente de audio de la escena de audio con respecto a las posiciones virtuales del hablante asociadas con los canales de salida de la señal de audio multicanal. En otras fases temporales, sin embargo, los diferentes canales de la señal de audio pueden estar más o menos no correlacionados entre sí y pueden incluso representar, por ejemplo, fuentes de audio completamente diferentes.

**[0017]** A fin de poder justificar la posible relación de variable en el tiempo entre los canales de la señal de audio, el decodificador subyacente del códec de audio 10 de la Fig. 1 permite un uso variable en el tiempo de diferentes mediciones que aprovechan las redundancias entre canales. Por ejemplo, la codificación MS permite activar entre las representaciones de los canales izquierdo y derecho de una señal de audio estéreo tal como están, o como un par de canales M (medio) y S (lateral) que representan la mezcla descendente de los canales izquierdo y derecho y la diferencia de la mitad de los mismos, respectivamente. Es decir, existen continuamente – en un sentido espectrot temporal – espectrogramas de los dos canales transmitidos por el flujo de datos 30, pero el significado de estos canales (transmitidos) pueden cambiar en el tiempo y en la relación con los canales de salida, respectivamente.

**[0018]** La predicción estéreo compleja – otra herramienta de aprovechamiento de la redundancia entre canales – permite, en el dominio espectral, predecir los coeficientes en el dominio de la frecuencia de un canal o las líneas espectrales que utilizan líneas subcubicadas espectralmente de otro canal. Más detalles con respecto a este caso se describen a continuación.

**[0019]** Con el fin de facilitar la comprensión de la descripción siguiente de la Fig. 1 y los componentes que se muestran en la misma, la Fig. 2 muestra, para el caso ejemplar de una señal de audio estéreo representada por el flujo de datos 30, una manera posible de cómo los valores de muestra de las líneas espectrales de los dos canales podrían codificarse en el flujo de datos 30 para que sean procesadas por el decodificador 10 de la Fig. 1. En particular, mientras en la mitad superior de la Fig. 2 se representa el espectrograma 40 de un primer canal de la señal de audio estéreo, la mitad inferior de la Fig. 2 ilustra el espectrograma 42 del otro canal de la señal de audio estéreo. Nuevamente, cabe observar que el “significado” de los espectrogramas 40 y 42 puede cambiar con el tiempo, por ejemplo, debido a una conmutación variable en el tiempo entre un dominio codificado MS y un dominio no codificado MS. En el primer caso, los espectrogramas 40 y 42 se relacionan con un canal M y S, respectivamente, mientras que en el último caso los espectrogramas 40 y 42 se relacionan con canales izquierdo y derecho. La conmutación entre un dominio codificado MS y un dominio no codificado MS puede señalizarse en el flujo de datos 30.

**[0020]** La Fig. 2 muestra que los espectrogramas 40 y 42 pueden codificarse en flujos de datos 30 en una resolución espectrot temporal variable en el tiempo. Por ejemplo, ambos canales (transmitidos) pueden estar, de manera alineada en el tiempo, subdivididos en una secuencia de tramas indicadas que utilizan llaves 44 que pueden tener la misma longitud y colindar entre sí sin superponerse. Como se ha mencionado anteriormente, la resolución espectral en la que los espectrogramas 40 y 42 se representan en el flujo de datos 30 puede cambiar con el tiempo. Preliminarmente, se supone que la que la resolución espectrot temporal cambia con el tiempo de igual manera para los espectrogramas 40 y 42, pero una extensión de esta simplificación es además viable y resulta evidente a partir de la descripción siguiente. El cambio de la resolución espectrot temporal es, por ejemplo, señalizada en el flujo de datos 30 en unidades de las tramas 44. Es decir, la resolución espectrot temporal cambia en unidades de tramas 44. El cambio en la resolución espectrot temporal de los espectrogramas 40 y 42 se consigue conmutando la longitud de transformada y el número de transformadas utilizadas para describir los espectrogramas 40 y 42 dentro de cada trama 44. En el ejemplo de la Fig. 2, las tramas 44a y 44b ejemplifican las tramas en las que una transformada larga ha sido utilizada para muestrear los canales de señal de audio en las mismas, resultando de este modo una resolución espectral alta con un valor de muestra de línea espectral por línea espectral para cada una de dichas tramas por canal. En la Fig. 2, los valores de muestra de las líneas espectrales se indican utilizando pequeñas cruces dentro de las casillas, en la cual las casillas, a la vez, están dispuestas en filas y columnas y representan un matriz espectral temporal con cada fila que se corresponde con una línea espectral y cada columna correspondiente a subintervalos de tramas 44 que se corresponden con las transformadas más cortas que intervienen en la formación de espectrogramas 40 y 42. En particular, la Fig. 2 ilustra, por ejemplo, para la trama 44d, que una trama puede alternativamente estar sujeta a transformadas consecutivas de longitud menor, produciendo de este modo, para dichas tramas como la trama 44d, varios espectros subsiguientes temporalmente de resolución espectral reducida. Se utilizan ejemplos de ocho transformadas cortas para la trama 44d, lo que da como resultado un muestreo espectrot temporal de los espectrogramas 40 y 42 dentro de esa trama 44d, en líneas espectrales separadas entre sí de manera que sólo se rellena cada ocho líneas espectrales, pero con un valor de muestra para cada una de las ocho ventanas de transformada o transformadas de menor longitud utilizadas para transformar la trama 44d. Para fines ilustrativos, la

Fig. 2 muestra que también son viables otros números de transformadas para una trama, como el uso de dos transformadas de una longitud de transformada que, por ejemplo, es la mitad de la longitud de transformada de las transformadas largas para las tramas 44a y 44b, resultando de este modo un muestreo de la matriz espectrot temporal o espectrogramas 40 y 42 en el que se obtienen dos valores de muestreo de línea espectral para cada segunda línea espectral, una de las cuales se relaciona con la transformada delantera y la otra con la transformada posterior.

**[0021]** Las ventanas de transformada de las transformadas en las cuales se subdividen las tramas se ilustran en la Fig. 2 a continuación, con cada espectrograma que utiliza líneas tipo ventanas de superposición. La superposición temporal sirve, por ejemplo, para fines de TDAC (cancelación de aliasing en el dominio del tiempo).

10

**[0022]** Aunque las realizaciones que se describen a continuación pueden implementarse de otra manera, la Fig. 2 ilustra el caso en el que la conmutación entre diferentes resoluciones espectrotemporales para las tramas individuales 44 se realiza de manera que para cada trama 44 el mismo número de valores de línea espectral indicados con las pequeñas cruces en la Fig. 2 resultan para el espectrograma 40 y el espectrograma 42, la diferencia sólo reside en la forma cómo las líneas muestrean espectrot temporalmente el respectivo recuadro espectrot temporal correspondiente a la trama respectiva 44, que se extiende temporalmente en el tiempo de la trama respectiva 44 y que se extiende espectralmente desde la frecuencia cero a la frecuencia máxima  $f_{max}$ .

15

**[0023]** Con las flechas de la Fig. 2, la Fig. 2 ilustra con respecto a la trama 44d que pueden obtenerse similares espectros para todas las tramas 44 distribuyendo adecuadamente los valores de muestra de línea espectral que pertenecen a la misma línea espectral, pero con ventanas de transformada cortas dentro de una trama de un canal sobre las líneas espectrales no ocupadas (vacías) dentro de esa trama hasta la próxima línea espectral ocupada de esa misma trama. Dichos espectros se denominan "espectros entrelazados" en las descripciones siguientes. En las  $n$  transformadas entrelazadas de una trama de un canal, por ejemplo, los valores de línea espectral cubricados espectralmente de las  $n$  transformadas cortas siguen uno tras otro antes de que siga el conjunto de  $n$  valores de las líneas espectrales cubricadas espectralmente de las  $n$  transformadas cortas de la línea espectral subsiguiente espectralmente. Una forma intermedia de entrelazar también sería factible: en vez de entrelazar todos los coeficientes de las líneas espectrales de una trama, sería factible entrelazar sólo los coeficientes de las líneas espectrales de un subconjunto adecuado de las transformadas cortas de una trama 44d. En cualquier caso, siempre que se analizan los espectros de las tramas de los dos canales correspondientes a los espectrogramas 40 y 42, los espectros pueden referirse a los entrelazados o no entrelazados.

20

25

30

35

40

45

**[0024]** Con el fin de codificar de modo eficaz los coeficientes de la línea espectral que representan los espectrogramas 40 y 42 mediante el flujo de datos 30 pasados al decodificador 10, los mismos son cuantizados. Con el fin de controlar el ruido de la cuantización espectrot temporalmente, el tamaño de la etapa de cuantización se controla mediante factores de escala que se determinan en una matriz espectrot temporal. En particular, dentro de la secuencia de espectros de cada espectrograma, las líneas espectrales son agrupadas en grupos de factor de escala consecutivos que no se superponen espectralmente. La Fig. 3 muestra un espectro 46 del espectrograma 40 en la mitad superior del mismo y un espectro cotemporal 48 fuera del espectrograma 42. Como se muestra en la misma, los espectros 46 y 48 se subdividen en bandas de factor de escala a lo largo del eje espectral  $f$  para agrupar las líneas espectrales en grupos no superpuestos. Las bandas de factor de escala se ilustran en la Fig. 3 con llaves 50. Para mayor simplicidad, se supone que los límites entre las bandas de factor de escala coinciden entre los espectros 46 y 48 pero no es necesario que este sea el caso.

**[0025]** Es decir, por medio de la codificación en el flujo de datos 30, los espectrogramas 40 y 42 se subdividen cada uno en una secuencia temporal de espectros y cada uno de estos espectros se subdivide espectralmente en bandas de factor de escala y, en cada banda de factor de escala, el flujo de datos 30 codifica o transmite información sobre un factor de escala correspondiente a la respectiva banda de factor de escala. Los coeficientes de la línea espectral que caen en una respectiva banda de factor de escala 50 se cuantizan con el respectivo factor de escala o, en lo que respecta al decodificador 10, se puede decuantizar con el factor de escala de la banda de factor de escala correspondiente.

50

55

**[0026]** Antes de volver a la Fig. 1 y la descripción del mismo, se supondrá en la descripción siguiente que el canal concretamente tratado, es decir, el mismo en cuya decodificación intervienen los elementos específicos del decodificador de la Fig. 1 excepto 34, es el canal transmitido del espectrograma 40, que como se ha mencionado anteriormente, puede representar uno de los canales izquierdo y derecho, un canal M o un canal S, con el supuesto de que la señal de audio multicanal codificada en el flujo de datos 30 es una señal de audio estéreo.

**[0027]** Si bien el extractor de líneas espectrales 20 se configura para extraer los datos de líneas espectrales, es decir, los coeficientes de las líneas espectrales de las tramas 44 del flujo de datos 30, el extractor de factor de escala 22 está configurado para extraer de cada trama 44 los factores de escala correspondientes. Con este propósito, los extractores 20 y 22 pueden utilizar decodificación por entropía. De acuerdo con una realización, el extractor de factor de escala 22 se configura para extraer secuencialmente los factores de escala de, por ejemplo, el espectro 46 de la Fig. 3, es decir los factores de escala de las bandas de factor de escala 50, a partir del flujo de datos 30 mediante una decodificación por entropía adaptativa al contexto. El orden de la decodificación secuencial puede seguir el orden espectral definido entre las bandas de factor de escala que van, por ejemplo, de baja frecuencia a alta frecuencia. El extractor de factor de escala 22 puede utilizar decodificación por entropía adaptativa al contexto y determinar el contexto para cada factor de escala en función de factores de escala ya extraídos en un vecindario espectral de un factor de escala extraído actual, como por ejemplo, en función del factor de escala de la banda de factor de escala inmediatamente precedente. Alternativamente, el extractor de factor de escala 22 puede de manera predictiva decodificar los factores de escala del flujo de datos 30 como, por ejemplo, mediante decodificación diferencial al mismo tiempo que predice un factor de escala decodificado actual basado en cualquiera de los factores de escala previamente decodificados como el inmediatamente anterior. Notablemente, este procedimiento de extracción de factor de escala es agnóstico con respecto al factor de escala que pertenece a una banda de factor de escala rellena exclusivamente con líneas espectrales cuantizadas a cero, o rellena con líneas espectrales entre las cuales al menos una se cuantiza con un valor diferente de cero. Un factor de escala perteneciente a una banda de factor de escala rellena con líneas espectrales cuantizadas a cero sólo puede servir como base de predicción para un factor de escala decodificado posterior que posiblemente pertenece a una banda de factor de escala rellena con líneas espectrales entre las cuales una es diferente de cero, y se puede predecir sobre la base de un factor de escala previamente decodificado que posiblemente pertenece a una banda de factor de escala rellena con líneas espectrales entre las cuales una es diferente de cero.

**[0028]** Sólo en aras de integridad, se observa que el extractor de líneas espectrales 20 extrae los coeficientes de las líneas espectrales con los cuales las bandas de factor de escala 50 se rellenan igualmente mediante, por ejemplo, codificación por entropía y/o codificación predictiva. La codificación por entropía puede utilizar la adaptabilidad al contexto sobre la base de los coeficientes de las líneas espectrales en un vecindario espectral de coeficientes de líneas espectrales decodificados actuales y, de igual modo, la predicción puede ser una predicción espectral, una predicción temporal o una predicción espectral que predice un coeficiente de línea espectral decodificado actual sobre la base de los coeficientes de líneas espectrales decodificados previamente en un vecindario espectral de los mismos. En aras de una mayor eficacia en la codificación, el extractor de líneas espectrales 20 se configura para realizar la decodificación de las líneas espectrales o coeficientes de líneas en tuplas, que reúnen o agrupan las líneas espectrales a lo largo del eje de frecuencia.

**[0029]** De esta manera, en la salida del extractor de líneas espectrales 20 los coeficientes de líneas espectrales se proporcionan como, por ejemplo, en unidades de espectros de manera que el espectro 46 reúne, por ejemplo, todos los coeficientes de líneas espectrales de una trama correspondiente o, alternativamente, reúne todos los coeficientes de líneas espectrales de determinadas transformadas cortas de una trama correspondiente. En la salida del extractor de factor de escala 22, a su vez, se emiten los factores de escala correspondientes de los espectros respectivos.

**[0030]** EL identificador de la banda de factor de escala 12 así como el decuantizador 14 tienen entradas de línea espectral acopladas a la salida del extractor de líneas espectrales 20, y el decuantizador 14 y el relleno de ruido 16 tienen entradas del factor de escala acopladas a la salida del extractor de factor de escala 22. El identificador de la banda de factor de escala 12 está configurado para identificar las denominadas bandas de factor de escala cuantizadas a cero dentro de un espectro actual 46, es decir las bandas de factor de escala dentro de las cuales todas las líneas espectrales se cuantizan a cero como la banda de factor de escala 50c en la Fig. 3, y las bandas de factor de escala restantes del espectro dentro de las cuales al menos una línea espectral está cuantizada diferente de cero. En particular, en la Fig. 3 los coeficientes de líneas espectrales se indican con áreas sombreadas en la Fig. 3. A partir de la misma se observa que en el espectro 46 todas las bandas de factor de escala, menos la banda de factor de escala 50b, tienen al menos una línea espectral, el coeficiente de línea espectral de la cual se cuantizan a un valor diferente de cero. Más adelante resultará evidente que la banda de factor de escala cuantizada a cero como 50d forma la materia de relleno de ruido entre canales que se describe adicionalmente a continuación. Antes de proceder con la descripción, se observa que el identificador de la banda de factor de escala 12 puede restringir su identificación a solamente un subconjunto adecuado de bandas de factor de escala 50 como las bandas de factor de escala por encima de una determinada frecuencia inicial 52. En la Fig. 3, este aspecto restringe el procedimiento de identificación en las bandas de factor de escala 50d, 50e y 50f.

**[0031]** El identificador de la banda de factor de escala 12 informa al relleno de ruido 16 sobre las bandas de

factor de escala que son bandas de factor de escala cuantizadas a cero. El decuantizador 14 utiliza los factores de escala asociados con un espectro de entrada 46 que decuantiza, o escala, los coeficientes de líneas espectrales de las líneas espectrales del espectro 46 según los factores de escala asociados, es decir, los factores de escala asociados con las bandas de factor de escala 50. En particular, el decuantizador 14 decuantiza y escala los  
 5 coeficientes de líneas espectrales que caen en la banda de factor de escala respectiva con el factor de escala asociado con la banda de factor de escala respectiva. La Fig. 3 debe interpretarse como una muestra del resultado de la decuantización de las líneas espectrales.

10 **[0032]** El rellenador de ruido 16 obtiene la información en las bandas de factor de escala cuantizadas a cero que forman la materia del relleno de ruido siguiente, el espectro decuantizado así como los factores de escala de, al menos, las bandas de factor de escala identificadas como bandas de factor de escala cuantizadas a cero y una señalización que se obtiene a partir del flujo de datos 30 para la trama actual que revela si el relleno de ruido entre canales debe realizarse para la trama actual.

15 **[0033]** El procedimiento de relleno de ruido entre canales que se describe en el ejemplo siguiente en realidad implica dos tipos de relleno de ruido, concretamente, la inserción de un ruido de fondo 54 perteneciente a todas las líneas espectrales que se han cuantizado a cero independientemente de su potencial pertenencia a una banda de factor de escala cuantizada a cero, y el procedimiento real de relleno de ruido entre canales. Aunque esta combinación se describe a continuación, debe enfatizarse que la inserción del ruido de fondo puede omitirse según una realización  
 20 alternativa. Asimismo, la señalización con respecto a la conmutación activada/desactivada del relleno de ruido relacionada con la trama actual y obtenida a partir del flujo de datos 30 podría relacionarse solamente con el relleno de ruido entre canales, o podría controlar la combinación de ambos tipos de relleno de ruido juntos.

25 **[0034]** En lo que respecta a la inserción del ruido de fondo, el rellenador de ruido 16 puede funcionar de la siguiente manera. En particular, el rellenador de ruido 16 puede emplear una generación de ruido artificial como un generador de números pseudoaleatorios u otra fuente de aleatoriedad a fin de rellenar las líneas espectrales, los coeficientes de las líneas espectrales de las cuales eran cero. El nivel de ruido de fondo 54 insertado de esta manera en las líneas espectrales cuantizadas a cero podría establecerse según una señalización explícita dentro del flujo de datos 30 para la trama actual del espectro actual 46. El "nivel" del ruido de fondo 54 puede determinarse mediante el  
 30 valor cuadrático medio (RMS) o la medición de la energía, por ejemplo.

**[0035]** La inserción del ruido de fondo representa de este modo una clase de prerelleno para las bandas de factor de escala que se han identificado como cuantizadas a cero como, por ejemplo, la banda de factor de escala 50d en la Fig. 3. Además afecta a otras bandas de factor de escala más allá de las que se han cuantizado a cero, pero  
 35 estas últimas están sujetas al siguiente relleno de ruido entre canales. Como se describe a continuación, el procedimiento de relleno de ruido entre canales es para rellenar las bandas de factor de escala cuantizadas a cero hasta un nivel que está controlado por el factor de escala de la banda de factor de escala cuantizada a cero respectiva. Esta última puede utilizarse directamente para este fin debido a que todas las líneas espectrales de la banda de factor de escala cuantizada a cero respectiva se han cuantizado a cero. Sin embargo, el flujo de datos 30 puede contener  
 40 una señalización adicional de un parámetro para cada trama o cada espectro 46, que comúnmente se aplica a los factores de escala de todas las bandas de factor de escala cuantizadas a cero de la trama o espectro 46 correspondiente y da como resultado, cuando se aplica a los factores de escala de las bandas de factor de escala cuantizadas a cero por el rellenador de ruido 16, el nivel de relleno respectivo que es individual para las bandas de factor de escala cuantizadas a cero. Es decir, el rellenador de ruido 16 puede modificar, utilizando la misma función  
 45 de modificación, para cada banda de factor de escala cuantizada a cero del espectro 46, el factor de escala de la banda de factor de escala respectiva que utiliza el parámetro recién mencionado contenido en el flujo de datos 30 para ese espectro 46 de la trama actual, a fin de obtener un nivel objetivo de relleno para la banda de factor de escala cuantizada a cero respectiva midiendo, en términos de energía o RMS, por ejemplo, el nivel hasta el cual el procedimiento de relleno de ruido entre canales deberá rellenar la banda de factor de escala cuantizada a cero  
 50 respectiva con (opcionalmente) ruido adicional (además del ruido de fondo 54).

**[0036]** En particular, para realizar el relleno de ruido entre canales 56, el rellenador de ruido 16 obtiene una parte coubicada espectralmente del espectro del otro canal 48, en un estado ya decodificado en gran parte o en su totalidad, y copia la parte del espectro obtenida 48 en la banda de factor de escala cuantizada a cero, en la cual esa  
 55 parte ha sido coubicada espectralmente y escalada de tal manera que el nivel de ruido total resultante dentro de esa banda de factor de escala cuantizada a cero –derivada por una integración sobre las líneas espectrales de la banda de factor de escala respectiva– es igual al nivel objetivo de relleno mencionado anteriormente obtenido a partir del factor de escala de la banda de factor de escala cuantizada a cero. Por medio de esta medición, la tonalidad del relleno de ruido en la banda de factor de escala cuantizada a cero respectiva mejora en comparación con el ruido generado

artificialmente, como el que forma la base del ruido de fondo 54, y es además mejor que una copia/replicación espectral no controlada de líneas de muy baja frecuencia dentro del mismo espectro 46.

**[0037]** Para ser aún más exactos, el rellenador de ruido 16 ubica, para una banda actual como 50d, una parte 5 coubicada espectralmente dentro del espectro 48 del otro canal, escala las líneas espectrales del mismo en función del factor de escala de la banda de factor de escala cuantizada a cero 50d de la manera que se ha descrito antes y que implica, opcionalmente, una compensación adicional o parámetro de factor de ruido contenido en el flujo de datos 30 para la trama o espectro actual 46, de manera que el resultado del mismo rellene la banda de factor de escala cuantizada a cero respectiva 50d hasta el nivel deseado tal como se define con el factor de escala de la banda de 10 factor de escala cuantizada a cero 50d. En la presente realización, esto significa que el relleno se realiza de una manera aditiva en relación con el ruido de fondo 54.

**[0038]** De acuerdo con una realización simplificada, el espectro con relleno de ruido resultante 46 entraría directamente en la entrada del transformador inverso 18 para obtener, para cada ventana de transformada, a las 15 cuales pertenecen los coeficientes de las líneas espectrales del espectro 46, una parte en el dominio del tiempo de la señal de audio en el tiempo del canal respectivo, en la que (no se muestra en la Fig. 1) un procedimiento de suma por superposición puede combinar estas partes en el dominio del tiempo. Es decir, si el espectro 46 es un espectro no entrelazado, los coeficientes de líneas espectrales del cual solo pertenecen a una transformada, el transformador inverso 18 somete esa transformada para que dé como resultado una parte en el dominio del tiempo, y los extremos 20 precedente y posterior de la cual estarían sujetos a un procedimiento de suma por superposición con las partes precedente y posterior en el dominio del tiempo obtenidas mediante la transformación inversa de las transformadas inversas precedente y posterior a fin de realizar, por ejemplo, la cancelación de aliasing en el dominio del tiempo. Sin embargo, si el espectro 46 ha entrelazado en el mismo los coeficientes de líneas espectrales de más de una transformada consecutiva, el transformador inverso 18 sometería a los mismos a transformaciones inversas separadas 25 a fin de obtener una parte en el dominio del tiempo por transformación inversa y, según el orden temporal definido entre ellas, las partes en el dominio del tiempo estarían sometidas a un procedimiento de suma por superposición entre ellas, así como con respecto a las partes en el dominio del tiempo precedente y posterior de otros espectros o tramas.

**[0039]** Sin embargo, en aras de la integridad debe observarse que pueden realizarse otros procesamientos en 30 el espectro con relleno de ruido. En la Fig. 1, el filtro TNS inverso puede realizar un filtrado TNS inverso en el espectro con relleno de ruido. Es decir, controlado mediante los coeficientes de filtro TNS para la trama o espectro actual 46, el espectro obtenido hasta ahora está sujeto a un filtrado lineal en la dirección espectral.

**[0040]** Con o sin el filtrado TNS inverso, el predictor estéreo complejo 24 puede tratar el espectro como 35 predicción residual de la predicción entre canales. Más concretamente, el predictor entre canales 24 puede utilizar una parte coubicada espectralmente del otro canal para predecir el espectro 46 o, al menos, un subconjunto de las bandas de factor de escalas 50 del mismo. El procedimiento de predicción complejo se ilustra en la Fig. 3 con el recuadro de línea discontinua 58 en relación con la banda de factor de escala 50b. Es decir, el flujo de datos 30 puede contener 40 parámetros de predicción entre canales, por ejemplo, cuáles de las bandas de factor de escala 50 se van a predecir entre canales y cuáles no se van a predecir de esta manera. Además, los parámetros de predicción entre canales en el flujo de datos 30 pueden comprender factores de predicción compleja entre canales aplicados mediante el predictor entre canales 24 a fin de obtener el resultado de la predicción entre canales. Estos factores pueden estar en el flujo de datos 30 individualmente para cada banda de factor de escala o, alternativamente para cada grupo de una o más 45 bandas de factor de escala, para las cuales la predicción entre canales se activa o se señala su activación en el flujo de datos 30.

**[0041]** La fuente de predicción entre canales puede, como se indica en la Fig. 3, ser el espectro 48 del otro canal. Para ser más exactos, la fuente de predicción entre canales puede ser la parte del espectro 48 coubicada 50 espectralmente, coubicada en la banda de factor de escala 50b que será predicha entre canales, extendida por una estimación de su parte imaginaria. La estimación de la parte imaginaria puede realizarse sobre la base de la parte coubicada espectralmente 60 del espectro 48 en sí, y/o puede utilizar una mezcla descendente de los canales ya decodificados de la trama anterior, es decir, la trama inmediatamente precedente a la trama decodificada actual a la cual pertenece el espectro 46. En efecto, el predictor entre canales 24 añade a las bandas de factor de escala que serán predichas entre canales, como por ejemplo la banda de factor de escala 50b en la Fig. 3, la señal de predicción 55 obtenida tal como se acaba de describir.

**[0042]** Como se observa en la descripción anterior, el canal al que pertenece el espectro 46 puede ser un canal codificado MS, o puede ser un canal relacionado con un altavoz, como un canal izquierdo o derecho de una señal de audio estéreo. En consecuencia, opcionalmente, un decodificador MS 26 somete el espectro predicho opcionalmente



entre canales 46 a una decodificación MS, en el mismo realiza, por línea espectral o espectro 46, una adición o sustracción con las líneas espectrales espectralmente correspondientes del otro canal correspondiente al espectro 48. Por ejemplo, aunque no se muestra en la Fig. 1, el espectro 48 tal como se muestra en la Fig. 3 se ha obtenido por medio de una parte 34 del decodificador 10 de manera análoga a la descripción presentada anteriormente con respecto al canal al que pertenece el espectro 46, y el módulo de decodificación MS 26 al realizar la decodificación MS, somete los espectros 46 y 48 a una adición de líneas espectrales o una sustracción de líneas espectrales, con ambos espectros 46 y 48 en la misma etapa dentro de la línea de procesamiento, es decir, ambos se han obtenido por predicción entre canales, por ejemplo, o ambos se han obtenido por relleno de ruido o filtrado TNS inverso.

10 **[0043]** Se observa que, opcionalmente, la decodificación MS puede realizarse de manera global con respecto a todo el espectro 46, o individualmente activable con el flujo de datos 30 en unidades de, por ejemplo, bandas de factor de escala 50. En otra palabras, la decodificación MS puede activarse o desactivarse utilizando la señalización respectiva en el flujo de datos 30 en unidades de, por ejemplo, tramas o una resolución espectral más fina como, por ejemplo, individualmente para las bandas de factor de escala de los espectros 46 y/o 48 de los espectrogramas 40 y/o 42, en los que se supone que se definen límites idénticos de las bandas de factor de escala de ambos canales.

**[0044]** Como se ilustra en la Fig. 1, el filtrado TNS inverso con el filtro TNS inverso 28 también podría realizarse después de cualquier procesamiento entre canales como, por ejemplo, una predicción entre canales 58 o la decodificación MS con un decodificador MS 26. El resultado aguas arriba, o aguas abajo de, el procesamiento entre canales podría fijarse, o controlarse, mediante una señalización respectiva en cada trama del flujo de datos 30 o en otro nivel de granularidad. Siempre que se realice un filtrado TNS inverso, los coeficientes de filtro TNS respectivos presentes en el flujo de datos para el espectro actual 46 controlan un filtro TNS, es decir un filtro de predicción lineal que se ejecuta en la dirección espectral para filtrar linealmente el espectro de entrada en el módulo de filtro TNS inverso respectivo 28a y/o 28b.

**[0045]** Por lo tanto, el espectro 46 que llega a la entrada del transformador inverso 18 puede haber estado sujeto a otro procesamiento como el que se ha descrito. Nuevamente, la descripción anterior no debe interpretarse de manera que todas estas herramientas opcionales deban estar presentes, ya sea simultáneamente o no. Estas herramientas pueden estar presentes en el decodificador 10 parcial o colectivamente.

**[0046]** En cualquier caso, el espectro resultante en la entrada del transformador inverso representa la reconstrucción final de la señal de salida del canal y forma la base de la mezcla descendente antes mencionada para la trama actual y que sirve, como se ha descrito con respecto a la predicción compleja 58, como base para la estimación de la posible parte imaginaria de la próxima trama a decodificar. Podrá servir además como reconstrucción final para predecir entre canales otro canal diferente con el que se relacionan los elementos excepto el 34 en la Fig. 1.

**[0047]** La mezcla descendente respectiva se forma mediante un proveedor de mezcla descendente 31 que combina este espectro final 46 con la versión final respectiva del espectro 48. La última entidad, es decir, la versión final respectiva del espectro 48 forma la base para la predicción compleja entre canales en el predictor 24.

**[0048]** La Fig. 4 muestra una alternativa relativa a la Fig. 1 siempre que la base para el relleno de ruido entre canales esté representada por la mezcla descendente de líneas espectrales cúbicadas espectralmente de una trama anterior de manera que, en el caso opcional de utilizar una predicción compleja entre canales, la fuente de esta predicción compleja entre canales se utiliza dos veces, como fuente para el relleno de ruido entre canales así como fuente para la estimación de la parte imaginaria en la predicción compleja entre canales. La Fig. 4 muestra un decodificador 10 con una parte 70 perteneciente a la decodificación del primer canal al que pertenece el espectro 46, y la estructura interna de la otra parte antes mencionada 34, que interviene en la decodificación del otro canal que comprende espectro 48. El mismo signo de referencia ha sido utilizado para los elementos internos de la parte 70 por un lado y 34 por el otro lado. Como se puede observar, la construcción es la misma. En la salida 32 se emite un canal de la señal de audio estéreo, y en la salida del transformador inverso 18 de la segunda parte del decodificador 34 da como resultado el otro canal (salida) de la señal de audio estéreo con esta salida que se indica con el signo de referencia 74. Nuevamente las realizaciones que se han descrito anteriormente pueden ser fácilmente transferidas al caso de utilizar más de dos canales.

**[0049]** El proveedor de mezcla descendente 31 es utilizado por ambas partes 70 y 34 y recibe los espectros cúbicados temporalmente 48 y 46 de los espectrogramas 40 y 42 a fin de formar una mezcla descendente basada en los mismos mediante la suma de estos espectros de una línea espectral con una base de línea espectral, potencialmente formando el promedio de los mismos al dividir la suma en cada línea espectral por el número de

canales sometidos a mezcla descendente, es decir, dos en el caso de la Fig. 4. En la salida del proveedor de mezcla descendente 31, la mezcla descendente de la trama anterior resulta de esta medición. En este sentido se observa que, en el caso de la trama anterior que contiene más de un espectro en cada espectrograma 40 y 42, existen diferentes posibilidades en lo que respecta a la manera cómo funciona el proveedor de la mezcla descendente 31 en ese caso. Por ejemplo, en ese caso el proveedor de la mezcla descendente 31 puede utilizar el espectro de las transformadas posteriores de la trama actual, o puede utilizar un resultado entrelazado de entrelazar todos los coeficientes de las líneas espectrales de la trama actual del espectrograma 40 y 42. El elemento de retardo 74 de la Fig. 4, tal como está conectado a la salida del proveedor de la mezcla descendente 31, muestra que la mezcla descendente así proporcionada en la salida del proveedor de la mezcla descendente 31 forma la mezcla descendente de la trama anterior 76 (véase la Fig. 3 con respecto al relleno de ruido entre canales 56 y la predicción compleja 58, respectivamente). Por lo tanto, la salida del elemento de retardo 74 está conectada a las entradas de los predictores entre canales 24 de las partes del decodificador 34 y 70 por un lado, y a las entradas de los rellenos de ruido 16 de las partes del decodificador 70 y 34, por el otro lado.

15 **[0050]** Es decir, si bien en la Fig. 1 el relleno de ruido 16 recibe el espectro coubicado espectrot temporalmente finalmente reconstruido del otro canal 48 de la misma trama actual como base del relleno de ruido entre canales, en la Fig. 4, en cambio, el relleno de ruido entre canales se realiza a partir de la mezcla descendente de la trama anterior tal como está proporcionada por el proveedor de mezcla descendente 31. La manera cómo se realiza el relleno de ruido entre canales sigue siendo la misma. Es decir, el relleno de ruido entre canales 20 16 toma una parte coubicada espectralmente fuera del espectro respectivo del espectro del otro canal de la trama actual, en el caso de la Fig. 1, y el espectro final decodificado en gran parte, o en su totalidad, tal como se obtiene a partir de la trama anterior que representa la mezcla descendente de la trama anterior, en el caso de la Fig. 4, y añade la misma parte de “fuente” a las líneas espectrales dentro de la banda de factor de escala para ser rellena con ruido como 50d en la Fig. 3, escalada según un nivel de ruido objetivo determinado por el factor de escala de la banda de 25 factor de escala respectiva.

**[0051]** Concluyendo el análisis anterior de las realizaciones que describen el relleno de ruido entre canales en un decodificador de audio, resultará evidente para los lectores expertos en la técnica que, antes de sumar la parte tomada coubicada espectralmente o temporalmente del espectro “fuente” a las líneas espectrales de la banda de factor de escala “objetivo”, puede aplicarse un determinado preprocesamiento a las líneas espectrales “fuente” sin desviarse del concepto general del relleno entre canales. En particular, puede ser de favorable aplicar una operación de filtrado como, por ejemplo, un aplanamiento espectral o una eliminación de la inclinación a las líneas espectrales de la región “fuente” que se añadirán a la banda de factor de escala “objetivo”, como 50d en la Fig. 3, con el fin de mejorar la calidad de audio del procedimiento de relleno de ruido entre canales. De igual modo, y como ejemplo de un espectro 35 decodificado en gran parte (y no totalmente), la parte “fuente” antes mencionada puede obtenerse a partir de un espectro que todavía no se ha filtrado mediante un filtro TNS inverso disponible (es decir, síntesis).

**[0052]** Por lo tanto, las realizaciones anteriores se relacionan con un concepto de relleno de ruido entre canales. A continuación, se describe una posibilidad de cómo construir el concepto anterior de relleno de ruido entre canales 40 en un códec existente, concretamente xHE-AAC, de manera semicompatible hacia atrás. En particular, se describe a continuación una implementación preferente de las realizaciones anteriores, según la cual se construye una herramienta de relleno de estéreo en un códec de audio basado en xHE-AAC con una forma de señalización semicompatible hacia atrás. Mediante el uso de la implementación que se describe a continuación, para determinadas señales estéreo, es viable el relleno de estéreo de los coeficientes de transformada en cualquiera de los dos canales 45 en un códec de audio basado en MPEG-D xHE-AAC (USAC), lo que mejora de este modo la calidad de la codificación de determinadas señales de audio especialmente a tasas de bits bajas. La herramienta de relleno de estéreo se señala de manera semicompatible hacia atrás de manera que los decodificadores xHE-AAC heredados pueden analizar y decodificar el flujo de bits sin errores o pérdidas de audio obvios. Como se ha descrito anteriormente, puede conseguirse una mayor calidad global si un decodificador de audio utiliza una combinación de los coeficientes 50 decodificados/cuantizados previamente de los dos canales estéreo para reconstruir los coeficientes cuantizados a cero (no transmitidos) de cualquiera de los canales decodificados actuales. Por lo tanto será deseable permitir el relleno de estéreo (de los coeficientes de canales anteriores a los actuales) además de la replicación de banda espectral (de los coeficientes de canales de baja a alta frecuencia) y el relleno de ruido (desde una fuente pseudoaleatoria no correlacionada) en codificadores de audio, especialmente xHE-AAC, o codificadores basados en él.

55 **[0053]** Para permitir que los flujos de bits codificados con relleno de estéreo se puedan leer y analizar por decodificadores xHE-AAC heredados, la herramienta de relleno de estéreo deseada debe utilizarse de manera semicompatible hacia atrás: su presencia no debería hacer que se detenga – ni tampoco que no se inicie – la decodificación en los decodificadores heredados La lectura del flujo de bits con la infraestructura xHE-AAC puede

facilitar la adopción del mercado.

**[0054]** Para conseguir el deseo mencionado anteriormente de la semicompatibilidad hacia atrás de la herramienta de relleno de estéreo en el contexto de xHE-AAC o sus posibles derivados, la siguiente implementación 5 implica la función de relleno de estéreo así como la capacidad de señalar la misma mediante sintaxis en el flujo de datos relacionados con el relleno de ruido. La herramienta de relleno de estéreo funcionará en línea con la descripción anterior. En un par de canales con configuración de ventana común, un coeficiente de banda de factor de escala cuantizado a cero es, cuando se ha activado la herramienta de relleno de estéreo, como alternativa (o, tal como se ha descrito, además) al relleno de ruido, reconstruido por una suma o diferencia de los coeficientes de la trama anterior 10 en cualquiera de los dos canales, preferentemente el canal derecho. El relleno de estéreo se realiza de forma similar al relleno de ruido. La señalización se realiza mediante la señalización del relleno de ruido de xHE-AAC. El relleno de estéreo se transfiere mediante información lateral de relleno de ruido de 8-bits. Esto es viable porque la norma MPEG-D USAC [4] establece que se transmitan todos los 8 bits aunque el nivel de ruido a aplicar sea cero. En esa situación, algunos bits de relleno de ruido pueden reutilizarse para la herramienta de relleno de estéreo.

15 **[0055]** La semicompatibilidad hacia atrás relacionada con el análisis del flujo de bits y la reproducción mediante decodificadores xHE-AAC heredados se asegura de la siguiente manera. El relleno de estéreo se señala mediante un nivel de ruido de cero (es decir, los primeros tres bits de relleno de ruido tienen todos un valor de cero) seguido de cinco bits diferentes de cero (que tradicionalmente representan una compensación de ruido) que contienen la 20 información lateral para la herramienta de relleno de estéreo así como el nivel de ruido perdido. Debido a que un decodificador xHE-AAC heredado ignora el valor de la compensación de ruido de 5 bits si el nivel de ruido de 3 bits es cero, la presencia de la señalización de la herramienta de relleno de estéreo sólo tiene efecto en el relleno de ruido en el decodificador heredado: el relleno de ruido se desactiva ya que los primeros tres bits son cero y el resto de la operación de decodificación funciona de la manera prevista. En particular, el relleno de estéreo no se realiza por el 25 hecho de que funciona como un procedimiento de relleno de ruido, que se ha desactivado. Por lo tanto, un decodificador heredado todavía ofrece una decodificación “correcta” del flujo de bits mejorada 30 ya que no necesita silenciar la señal de salida o abortar la decodificación al llegar a una trama con relleno de estéreo activado. Naturalmente, es, sin embargo, imposible proporcionar una reconstrucción correcta y prevista de los coeficientes de las líneas con relleno de estéreo, lo que origina una calidad deteriorada en las tramas afectadas en comparación con la decodificación mediante un decodificador adecuado capaz de manejar adecuadamente la nueva herramienta de 30 relleno de estéreo. Sin embargo, suponiendo que la herramienta de relleno de estéreo se utiliza tal como está prevista, es decir, solo en la entrada de estéreo a tasas de bits bajas, la calidad a través de los decodificadores xHE-AAC deberá ser mejor si las tramas afectadas se dejan debido al silenciamiento u originará otros errores de reproducción obvios.

35 **[0056]** A continuación se presenta una descripción detallada de cómo integrar una herramienta de relleno de estéreo, como extensión, en el códec xHE-AAC.

**[0057]** Cuando se integra dentro de la norma, la herramienta de relleno de estéreo se puede describir de la 40 siguiente manera. En particular, dicha herramienta de relleno de estéreo (SF) representará una nueva herramienta en la parte de dominio de la frecuencia (FD) del audio MPEG-H 3D. En línea con el análisis anterior, el objetivo de dicha herramienta de relleno de estéreo será la reconstrucción paramétrica de los coeficientes espectrales MDCT a tasas de bits bajas, similar a lo que ya se puede conseguir con el relleno de ruido según la sección 7.2 de la norma que se describe en [4]. Sin embargo, a diferencia del relleno de ruido, que emplea una fuente de ruido pseudoaleatorio para 45 generar valores espectrales MDCT de los canales FD, el SF también estará disponible para reconstruir los valores MDCT del canal derecho de un par de canales estéreo codificados conjuntamente que utilizan una mezcla descendente de los espectros MDCT izquierdo y derecho de la trama anterior. El SF, según la implementación que se establece a continuación, se señala de manera semicompatible hacia atrás mediante la información lateral del relleno de ruido que puede analizarse correctamente mediante un decodificador MPEG-D USAC heredado.

50 **[0058]** La descripción de la herramienta puede ser la siguiente manera. Cuando el SF se encuentra activo e una trama FD estéreo conjunta, los coeficientes MDCT de la banda de factor de escala vacíos (es decir, cuantizados totalmente a cero) del canal derecho (segundo), como 50d, se sustituyen por una suma o una diferencia de los coeficientes MDCT correspondientes de los canales decodificados izquierdo y derecho de la trama anterior (si FD). Si 55 el relleno de ruido heredado se encuentra activo para el segundo canal, los valores pseudoaleatorios se añaden también a cada coeficiente. Los coeficientes resultantes de cada banda de factor de escala son escalados de manera que la RMS (la raíz del valor cuadrático medio) de cada banda coincide con el valor transmitido por medio de ese factor de escala de la banda. Véase la sección 7.3 de la norma en [4]. Algunas restricciones operativas pueden proporcionarse para el uso de la nueva herramienta SF en la norma MPEG-D USAC. Por ejemplo, la herramienta SF

puede estar disponible para su uso solo en el canal FD derecho de un par de canales FD común, es decir, un elemento del par de canales que transmite un StereoCoreToolInfo() con common\_window == 1. Además, debido a la señalización semicompatible hacia atrás, la herramienta SF puede estar disponible para su uso sólo cuando el noiseFilling==1 en el contenedor de sintaxis UsacCoreConfig(). Si uno de los canales en el par está en el LPD core\_mode, la herramienta SF puede no utilizarse incluso si el canal derecho se encuentre en el modo FD.

**[0059]** Los siguientes términos y definiciones se utilizan a continuación con el fin de describir en forma más clara la extensión de la norma tal como se describe en [4].

10 **[0060]** En particular, en lo que se refiere a los elementos de los datos, se introduce el elemento de datos nuevo siguiente:

stereo\_filling indicador binario que indica si SF se utiliza en la trama y canal actuales

15 **[0061]** Adicionalmente se introducen nuevos elementos de ayuda:

noise\_offset compensación de relleno de ruido que modifica los factores de escala de bandas cuantizadas a cero (sección 7.2)

20 noise\_level nivel de relleno de ruido que representa la amplitud del ruido de espectro añadido (sección 7.2)  
downmix\_prev[] mezcla descendente (es decir, suma o diferencia) de los canales izquierdo y derecho de la trama anterior

sf\_index[g][sfb] índice de factor de escala (es decir, número entero transmitido) para el grupo de ventana g y banda sfb

25 **[0062]** El procedimiento de decodificación de la norma se extenderá de la siguiente manera. En particular, la decodificación de un canal FD codificado conjuntamente en estéreo con la herramienta SF activada se ejecuta en tres etapas secuenciales de la siguiente manera:

**[0063]** En primer lugar, tendrá lugar la decodificación del indicador stereo\_filling.

30 El indicador stereo\_filling no representa un elemento de flujo de datos independiente sino que deriva de elementos rellenos de ruido, noise\_offset y noise\_level, en un UsacChannelPairElement() y el indicador de common\_window en StereoCoreToolInfo().

Si noiseFilling == 0 o common\_window == 0 o el canal actual es el canal izquierdo (primero) en el elemento, el stereo\_filling es 0 y el procedimiento de relleno de estéreo finaliza. Sino,

```
35 if ((noiseFilling != 0) && (common_window != 0) && (noise_level == 0)) {
    stereo_filling = (noise_offset & 16) / 16;
    noise_level = (noise_offset & 14) / 2;
    noise_offset = (noise_offset & 1) * 16;
40 }
    else {
        stereo_filling = 0;
    }
}
```

45 **[0064]** En otras palabras, si noise\_level == 0, noise\_offset contiene el indicador de stereo\_filling seguido por 4 bits de datos de relleno de ruido que posteriormente son redispuestos. Como esta operación altera los valores de noise\_level y noise\_offset, necesita ejecutarse antes del procedimiento de relleno de ruido de la sección 7.2. Asimismo, el pseudocódigo anterior no se ejecuta en el canal izquierdo (primero) de un UsacChannelPairElement() o ningún otro elemento.

50 **[0065]** A continuación, tendrá lugar el cálculo de downmix\_prev. downmix\_prev[], la mezcla descendente espectral que va a utilizarse para el relleno de estéreo es idéntica a la dm\_x\_re\_prev[] utilizada para la estimación del espectro MDST en la predicción estéreo compleja (sección 7.7.2.3). Esto significa que

55 • Todos los coeficientes de downmix\_prev[ ] deben ser cero si alguno de los canales de la trama y elemento con los cuales se realiza la mezcla descendente –es decir, la trama antes de la decodificada actualmente– utiliza core\_mode == 1 (LPD) o los canales utilizan longitudes de transformada diferentes (split\_transform == 1 o conmutación de bloques en window\_sequence == EIGHT\_SHORT\_SEQUENCE en sólo un canal) o usaIndependencyFlag == 1.

• Todos los coeficientes de `downmix_prev[]` deben ser cero durante el procedimiento de relleno de estéreo si la longitud de transformada del canal ha cambiado de la trama anterior a la actual (es decir, `split_transform == 1` precedida por `split_transform == 0`, o `window_sequence == EIGHT_SHORT_SEQUENCE` precedida por la `window_sequence !=`  
 5 `EIGHT_SHORT_SEQUENCE`, o viceversa) en el elemento actual.

• Si se aplica la división de transformada en los canales de la trama anterior o actual, `downmix_prev[]` representa una mezcla descendente espectral entrelazada línea a línea. Véase la herramienta de división de transformada.

10 • Si la predicción estéreo compleja no se utiliza en la trama y elemento actuales, `pred_dir` es igual a 0.

**[0066]** En consecuencia, la mezcla descendente anterior solo debe calcularse una vez para ambas herramientas, lo cual ahorra complejidad. La sola diferencia entre `downmix_prev[]` y `dmx_re_prev[]` en la sección 7.7.2 es el comportamiento cuando la predicción estéreo compleja no se utiliza en este momento, o cuando está activa pero  
 15 `use_prev_frame == 0`. En este caso, `downmix_prev[]` se calcula para la decodificación de relleno de estéreo según la sección 7.7.2.3 aunque `dmx_re_prev[]` no es necesaria para la decodificación de predicción estéreo compleja y es, por lo tanto, es indefinida/cero.

**[0067]** A continuación, se realizará el relleno de estéreo de las bandas de factor de escala vacíos.  
 20

**[0068]** Si el `stereo_filling == 1`, se lleva a cabo el procedimiento siguiente después del procedimiento de relleno de ruido en todas las bandas de factor de escala inicialmente vacías `sfb[]` por debajo de `max_sfb_ste`, es decir, todas las bandas en las que las líneas MDCT se han cuantizado a cero. En primer lugar, las energías de la `sfb[]` dada y las líneas correspondientes en `downmix_prev[]` se calculan mediante las sumas de los cuadrados de línea.  
 25 Posteriormente, el `sfbWidth` dado que contiene el número de líneas por `sfb[]`,

```

if (energy [sfb] < sfbWidth [sfb]) { /* el nivel de ruido no es máximo, o la banda empieza por debajo de la zona de
relleno de ruido */
facDmx = sqrt((sfbWidth[sfb] - energy[sfb]) / energy_dmx[sfb]);
30 factor = 0.0;
/* si la mezcla descendente anterior no está vacía, añade las líneas de mezcla descendente escaladas hasta que la
banda alcanza la unidad de energía */
for (index = swb_offset[sfb]; index < swb_offset[sfb+1]; index++) {
spectrum>window][index] += downmix_prev>window][index] * facDmx;
35 factor += spectrum>window][index] * spectrum>window][index];
}
if ((factor != sfbWidth [sfb]) && (factor > 0)) { /* la unidad de energía no se alcanzado, por lo que se modifica la banda
*/
factor = sqrt(sfbWidth[sfb] / (factor + 1e-8));
40 for (index = swb_offset[sfb]; index < swb_offset[sfb+1]; index++) {
spectrum>window][index] *= factor;
}}}
    
```

para el espectro de cada ventana de grupo. A continuación, se aplican los factores de escala en el espectro resultante  
 45 en la sección 7.3, con factores de escala de las bandas vacías procesadas como factores de escala regulares.

**[0069]** Una alternativa a la extensión anterior de la norma xHE-AAC utilizará un procedimiento de señalización semicompatible hacia atrás implícito.

50 **[0070]** La implementación anterior en la estructura del código xHE-AAC describe una estrategia que emplea un bit en un flujo de bits para señalar el uso de la nueva herramienta de relleno de estéreo, contenida en `stereo_filling` hacia un decodificador según la Fig. 1. Más concretamente, dicha señalización (llamémosla señalización semicompatible hacia atrás explícita) permite que los datos del flujo de bits heredados siguientes –en este caso información lateral de relleno de ruido– se usen independientemente de la señalización SF: en la presente realización,  
 55 los datos de relleno de ruido no dependen de la información de relleno de estéreo y viceversa. Por ejemplo, los datos de relleno de ruido que consisten en todos ceros (`noise_level = noise_offset = 0`) pueden transmitirse mientras que `stereo_filling` puede señalar cualquier posible valor (siendo un indicador binario, 0 o 1).

**[0071]** En los casos en que una estricta independencia entre los datos heredados y los datos de flujo de bits

de la invención no son necesarios y la señal de la invención es una decisión binaria, la transmisión explícita de un bit de señalización puede evitarse, y dicha decisión binaria puede señalizarse por presencia o ausencia de lo que puede denominarse señalización semicompatible hacia atrás explícita. Tomando la realización anterior como ejemplo, el uso de relleno de estéreo puede transmitirse sencillamente con el empleo de la nueva señalización: Si `noise_level` es cero y, al mismo tiempo, `noise_offset` no es cero, `stereo_filling` se establece igual a 1. Si ambos `noise_level` y `noise_offset` no son cero, `stereo_filling` es igual a 0. Un dependiente de esta señal implícita en la señal heredada con relleno de ruido ocurre cuando tanto `noise_level` como `noise_offset` son cero. En este caso, no queda claro si se utiliza la señal heredada o la nueva señalización implícita SF. Para evitar dicha ambigüedad, el valor de `stereo_filling` debe definirse por adelantado. En el presente ejemplo, es adecuado definir el `stereo_filling = 0` si los datos de relleno de ruido consisten en todos ceros, ya que esto es lo que señalizan los codificadores heredados sin capacidad de relleno de estéreo cuando el relleno de ruido no debe aplicarse en una trama.

**[0072]** Esta cuestión que todavía debe resolverse en el caso de señalización semicompatible hacia atrás implícita consiste en cómo señalizar `stereo_filling == 1` y el no relleno de ruido al mismo tiempo. Como se ha explicado, los datos de relleno ruido no deben ser todos cero, y si se requiere una magnitud de ruido de cero, `noise_level ((noise_offset & 14)/2` (tal como se ha mencionado anteriormente) debe ser igual a 0. Esto solo deja un `noise_offset ((noise_offset & 1)*16` tal como se ha mencionado anteriormente) mayor que 0 como solución. El `noise_offset`, sin embargo, se tiene en cuenta en el caso de relleno de estéreo cuando se aplica a factores de escala, incluso si `noise_level` es cero. Afortunadamente, un codificador puede compensar el hecho de que un `noise_offset` de cero no pueda ser transmisible mediante la alteración de los factores de escala afectados de manera que, al escribir el flujo de bits, contienen una compensación que no se deshace en el decodificador mediante `noise_offset`. Esto permite dicha señalización implícita en la realización anterior al coste de un aumento potencial en la velocidad de datos del factor de escala. Por lo tanto, la señalización del relleno de estéreo en el pseudocódigo de la descripción anterior podría cambiar de la siguiente manera, utilizando el bit de señalización SF guardado para transmitir el `noise_offset` con 2 bits (4 valores) en lugar de 1 bit:

```

if ((noiseFilling) && (common_window) && (noise_level == 0) &&
(noise_offset > 0)) {
stereo_filling = 1;
30 noise_level = (noise_offset & 28) / 4;
noise_offset = (noise_offset & 3) * 8;
}
else {
stereo_filling = 0;
35 }

```

**[0073]** En aras de la integridad, la Fig. 5 muestra un codificador de audio paramétrico según una realización de la presente solicitud. En primer lugar, el codificador de la Fig. 5 que generalmente se indica con el signo de referencia 100 comprende un transformador 102 que realiza la transformación de la versión original y no distorsionada de la señal de audio reconstruida en la salida 32 de la Fig. 1. Como se ha descrito con respecto a la Fig. 2, una transformada superpuesta puede utilizarse con una conmutación entre diferentes longitudes de transformadas con las ventanas de transformadas correspondientes en unidades de tramas 44. La diferente longitud de transformada y las ventanas de transformadas correspondientes se ilustran en la Fig. 2 con el signo de referencia 104. De manera similar a la Fig. 1, la Fig. 5 se concentra en una parte del decodificador 100 responsable para la codificación de un canal de la señal de audio multicanal, mientras que otra parte del decodificador del dominio del canal 100 generalmente se indica con el signo de referencia 106 en la Fig. 5.

**[0074]** En la salida del transformador 102 las líneas espectrales y los factores de escala no se cuantizan y significativamente no se ha producido todavía una pérdida de codificación. La salida del espectrograma por el transformador 102 entra en un cuantizador 108 que está configurado para cuantizar las líneas espectrales de la salida del espectrograma por el transformador 102, espectro por espectro, estableciendo y utilizando los factores de escala preliminares de las bandas de factor de escala. Es decir, en la salida del cuantizador 108, resultan los factores de escala preliminares y los coeficientes de líneas espectrales correspondientes, y una secuencia de un relleno de ruido 16', un filtro TNS inverso opcional 28a', un predictor entre canales 24', un decodificador MS 26' y un filtro TNS inverso 28b' están conectados secuencialmente para proporcionar al codificador 100 de la Fig. 5 la capacidad de obtener una versión final reconstruida del espectro actual tal como se obtiene en el lado del decodificador en la entrada del proveedor de mezcla descendente (véase la Fig. 1). En el caso de utilizar la predicción entre canales 24' y/o utilizar el relleno de ruido entre canales en la versión que forma el ruido entre canales que utiliza la mezcla descendente de la trama anterior, el codificador 100 también comprende un proveedor de mezcla descendente 31' para formar una

mezcla descendente de las versiones finales reconstruidas de los espectros de los canales de la señal de audio multicanal. Por supuesto, para ahorrar cálculos, en lugar de versiones finales, se pueden utilizar las versiones originales no cuantizadas de dichos espectros de los canales mediante el proveedor de mezcla descendente 31' en la formación de la mezcla descendente.

5

**[0075]** El codificador 100 puede utilizar información sobre la versión final reconstruida disponible de los espectros con el fin de realizar una predicción espectral entre tramas como la versión posible antes mencionada de realizar la predicción entre canales utilizando una estimación de parte imaginaria, y/o con el fin de realizar el control de velocidad, es decir, con el fin de determinar, dentro de un bucle de control de velocidad de transmisión, que los posibles parámetros finalmente codificados en flujo de datos 30 mediante el codificador 100 se establecen en un sentido óptimo de velocidad/distorsión.

10

**[0076]** Por ejemplo, uno de dichos parámetros establecidos en el bucle de predicción y/o bucle de control de la velocidad de transmisión del codificador 100 es, para cada banda de factor de escala cuantizada a cero identificada por el identificador 12', el factor de escala de la banda de factor de escala respectiva que ha sido sólo preliminarmente establecida mediante el cuantizador 108. En una predicción y/o bucle de control del codificador 100, el factor de escala de las bandas de factor de escala cuantizadas a cero se establece en un sentido óptimo de velocidad de transmisión/distorsión psicoacústica a fin de determinar el nivel de ruido objetivo mencionado anteriormente junto con, como se ha descrito antes, un parámetro opcional de modificación que también se transfiere en el flujo de datos de la trama correspondiente al lado del decodificador. Debe observarse que este factor de escala puede calcularse utilizando sólo las líneas espectrales del espectro y el canal al cual pertenece (es decir, el espectro "objetivo", tal como se ha descrito anteriormente) o, alternativamente, puede determinarse utilizando ambas líneas espectrales del espectro de canal "objetivo" y, además, las líneas espectrales del otro espectro de canal o espectro de mezcla descendente desde la trama anterior (es decir, el espectro "fuente" tal como se ha introducido antes) obtenido a partir del proveedor de mezcla descendente 31'. En particular para estabilizar el nivel de ruido objetivo y para reducir las fluctuaciones de nivel temporal en los canales de audio decodificados sobre los cuales se aplica el relleno de ruido entre canales, el factor de escala objetivo puede calcularse utilizando una relación entre una medición de energía de las líneas espectrales en la banda de factor de escala "objetivo", y una medición de energía de las líneas espectrales coubicadas en la región "fuente" correspondiente. Finalmente, como se ha observado anteriormente, esta región "fuente" puede proceder de una versión final reconstruida de otro canal o de la mezcla descendente de la trama anterior, o si la complejidad del codificador debe ser reducida, la versión original no cuantizada del mismo otro canal o la mezcla descendente de versiones originales no cuantizadas de los espectros de la trama anterior.

15

20

25

30

**[0077]** En función de ciertos requisitos de implementación, las realizaciones de la invención se pueden implementar en hardware o en software. La implementación se puede realizar utilizando un medio de almacenamiento digital, por ejemplo un disquete, un DVD, un disco Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control legibles electrónicamente almacenadas, que cooperan (o son capaces de cooperar) con un sistema informático programable de manera que se lleva a cabo el procedimiento respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

40

**[0078]** Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema informático programable, de manera que se lleva a cabo uno de los procedimientos descritos en esta invención.

**[0079]** En general, las realizaciones de la presente invención se pueden implementar como un producto de programa informático con un código del programa, el código del programa es operativo para realizar uno de los procedimientos, cuando el producto de programa informático se ejecuta en un ordenador. El código del programa, por ejemplo, se puede almacenar en un soporte legible por máquina.

45

**[0080]** Otras realizaciones comprenden el programa informático para realizar uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por máquina.

50

**[0081]** En otras palabras, una realización del procedimiento de la invención es, por lo tanto, un programa informático que tiene un código del programa para realizar uno de los procedimientos descritos en esta invención, cuando el programa informático se ejecuta en un ordenador.

55

**[0082]** Una realización adicional de los procedimientos de la invención 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 realizar uno de los procedimientos descritos en esta invención. El soporte de datos, el

medio de almacenamiento digital o el medio grabado normalmente son tangibles y/o no transitorios.

5 **[0083]** Una realización adicional del procedimiento de la invención, en consecuencia, es una corriente de datos o una secuencia de señales que representan el programa informático para realizar uno de los procedimientos descritos en esta invención. El flujo de datos o la secuencia de señales pueden, por ejemplo, estar configurados para ser transferidos a través de una conexión de comunicación de datos, por ejemplo, a través de Internet.

10 **[0084]** Otra realización adicional comprende un medio de procesamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado para o adaptado para realizar uno de los procedimientos descritos en esta invención.

**[0085]** Otra realización adicional comprende además un ordenador que tiene instalado el programa informático para realizar uno de los procedimientos descritos en esta invención.

15 **[0086]** Otra realización adicional según la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, en forma electrónica u óptica) un programa informático para realizar uno de los procedimientos descritos en esta invención a un receptor. El receptor, por ejemplo, puede ser un ordenador, un dispositivo móvil, un dispositivo de memoria o similares. El aparato o sistema, por ejemplo, pueden comprender un servidor de archivos para transferir el programa informático al receptor.

20 **[0087]** En algunas realizaciones, un dispositivo lógico programable (por ejemplo, una matriz de puertas programables por campo) se puede utilizar para realizar todas o algunas de las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones una matriz de puertas programables por campo puede cooperar con un microprocesador a fin de llevar a cabo uno de los procedimientos descritos en esta invención. En general los  
25 procedimientos se llevan a cabo, preferentemente, mediante cualquier aparato de hardware.

**[0088]** Las realizaciones descritas anteriormente son meramente ilustrativas de los principios de la presente invención. Debe entenderse que las modificaciones y variaciones de las disposiciones y detalles descritos en esta invención serán evidentes para otros expertos en la técnica. Es la intención, por lo tanto, de estar limitado solo por el  
30 alcance de las reivindicaciones de patente inminentes y no por los detalles específicos presentados a modo de descripción y las explicaciones de las realizaciones de este invención.

#### Referencias

35 **[0089]**

[1] Internet Engineering Task Force (IETF), RFC 6716, "Definition of the Opus Audio Codec," Int. Standard, Sep. 2012. Disponible en línea en <http://tools.ietf.org/html/rfc6716>.

40 [2] International Organization for Standardization, ISO/IEC 14496-3:2009, "Information Technology - Coding of audio-visual objects - Part 3: Audio," Geneva, Switzerland, Aug. 2009.

[3] M. Neuendorf et al., "MPEG Unified Speech and Audio Coding - The ISO/MPEG Standard for High-Efficiency Audio Coding of All Content Types," in Proc. 132nd AES Convention, Budapest, Hungary, Apr. 2012. También aparece en el  
45 Journal de AES, 2013.

[4] International Organization for Standardization, ISO/IEC 23003-3:2012, "Information Technology - MPEG audio - Part 3: Unified speech and audio coding," Geneva, Jan. 2012.



**REIVINDICACIONES**

1. Decodificador de audio paramétrico en el dominio de la frecuencia configurado para identificar (12) unas primeras bandas de factor de escala de un espectro de un primer canal de una trama actual de una señal de audio multicanal, dentro en las cuales todas las líneas espectrales se cuantizan a cero, y unas segundas bandas de factor de escala del espectro, en las cuales, al menos una línea espectral, se cuantiza a un valor diferente de cero;
- rellenar (16) las líneas espectrales dentro de una banda de factor de escala predeterminada de la primera banda de factor de escalas con ruido generado mediante las líneas espectrales de la mezcla descendente de una trama anterior de la señal de audio multicanal, con un ajuste del nivel del ruido que utiliza un factor de escala de la banda de factor de escala predeterminada;
- 10 decuantizar (14) las líneas espectrales dentro de las segundas bandas de factor de escala que utilizan factores de escala de las segundas bandas de factor de escala; y
- transformar inversamente (18) el espectro obtenido de las primeras bandas de factor de escala relleno con ruido cuyo nivel se ha ajustado con los factores de escala de las primeras bandas de factor de escala, y las segundas bandas de factor de escala decuantizadas con los factores de escala de las segundas bandas de factor de escala, a fin de obtener una parte en el dominio del tiempo del primer canal de la señal de audio multicanal.
2. Decodificador de audio paramétrico en el dominio de la frecuencia según la reivindicación 1 configurado además para, en el relleno, ajustar un nivel de una parte coubicada de un espectro de la mezcla descendente de la trama anterior, coubicada espectralmente con respecto a la banda de factor de escala predeterminada, que utiliza el factor de escala de la banda de factor de escala predeterminada, y añadir la parte coubicada, que tiene su nivel ajustado, a la banda de factor de escala predeterminada.
- 25 3. Decodificador de audio paramétrico en el dominio de la frecuencia según la reivindicación 2 configurado además para predecir un subconjunto de las bandas de factor de escala de un canal diferente o mezcla descendente de la trama actual para obtener una predicción entre canales, y utilizar la banda de factor de escala predeterminada rellena con el ruido, y las segundas bandas de factor de escala decuantizadas que utilizan los factores de escala de las segundas bandas de factor de escala como predicción residual de la predicción entre canales para obtener el
- 30 espectro.
4. Decodificador de audio paramétrico en el dominio de la frecuencia según la reivindicación 3 configurado además para, al predecir el subconjunto de las bandas de factor de escala, realizar una estimación de la parte imaginaria del canal diferente o mezcla descendente de la trama actual mediante el espectro de la mezcla descendente de la trama anterior.
- 35 5. Decodificador de audio paramétrico en el dominio de la frecuencia según cualquiera de las reivindicaciones 1 a 4, en el que el canal actual y el otro canal están sujetos a una codificación MS en el flujo de datos, y el decodificador de audio paramétrico en el dominio de la frecuencia está configurado para someter el espectro a la
- 40 decodificación MS.
6. Decodificador de audio paramétrico en el dominio de la frecuencia según cualquiera de las reivindicaciones anteriores, configurado además para extraer secuencialmente los factores de escala de las primeras y segundas bandas de factor de escala a partir de un flujo de datos que utiliza una decodificación por entropía adaptada al contexto con determinación de contexto en función de, y/o mediante una decodificación predictiva con predicción espectral en función de, factores de escala ya extraídos en el vecindario espectral de un factor de escala extraído actual, con los factores de escala dispuestos espectralmente según un orden espectral entre las primeras y segundas bandas de factor de escala.
- 45 7. Decodificador de audio paramétrico en el dominio de la frecuencia según cualquiera de las reivindicaciones anteriores, configurado, además, de modo que el ruido está generado adicionalmente con ruido aleatorio o pseudoaleatorio.
- 50 8. Decodificador de audio paramétrico en el dominio de la frecuencia según la reivindicación 7 configurado, además, para ajustar un nivel de ruido aleatorio o pseudoaleatorio igualmente para la primera banda de factor de escala, según un parámetro de ruido señalado en un flujo de datos de la trama actual.
9. Decodificador de audio paramétrico en el dominio de la frecuencia según cualquiera de las reivindicaciones anteriores configurado, además, para modificar igualmente los factores de escala de las primeras

bandas de factor de escala en relación con los factores de escala de las segundas bandas de factor de escala que utilizan un parámetro modificador señalado en un flujo de datos de la trama actual.

10. Codificador de audio paramétrico en el dominio de la frecuencia configurado para cuantizar las líneas  
5 espectrales de un espectro de un primer canal de una trama actual de una señal de audio multicanal que utiliza factores de escala preliminares de las bandas de factor de escala dentro del espectro;

identificar las primeras bandas de factor de escala en el espectro dentro de las cuales todas las líneas espectrales se  
cuantizan a cero, y las segundas bandas de factor de escala del espectro dentro de las cuales, al menos una línea  
10 spectral, se cuantiza a un valor diferente de cero,  
dentro de una predicción y/o bucle de control de velocidad de transmisión,

rellenar las líneas espectrales dentro de una banda de factor de escala predeterminada de las primeras bandas de  
factor de escala con ruido generado utilizando

15 las líneas espectrales de una mezcla descendente de una trama anterior de la señal de audio multicanal,

con un ajuste del nivel de ruido que utiliza un factor de escala real de la banda de factor de escala predeterminada; y  
señalizar el factor de escala real de la banda de factor de escala predeterminada en lugar del factor de escala  
20 preliminar.

11. Codificador de audio paramétrico en el dominio de la frecuencia según la reivindicación 10 configurado,  
además, para calcular el factor de escala real de la banda de factor de escala predeterminada sobre la base de un  
nivel de una versión no cuantizada de las líneas espectrales del espectro del primer canal dentro de la banda de factor  
25 de escala predeterminada y, además, sobre la base de las líneas espectrales de la mezcla descendente de una trama  
anterior de la señal de audio multicanal, o líneas espectrales de un canal diferente de la trama actual de la señal de  
audio multicanal.

12. Procedimiento de decodificación de audio paramétrico en el dominio de la frecuencia que comprende  
30 identificar las primeras bandas de factor de escala de un espectro de un primer canal de una trama actual de una señal  
de audio multicanal, dentro de las cuales todas las líneas espectrales se cuantizan a cero, y las segundas bandas de  
factor de escala del espectro, dentro de las cuales, al menos una línea spectral, se cuantiza a un valor diferente de  
cero;

35 rellenar las líneas espectrales dentro de una banda de factor de escala predeterminada de las primeras bandas de  
factor de escala con ruido generado utilizando

las líneas espectrales de una mezcla descendente de una trama anterior de la señal de audio multicanal,

40 con un ajuste del nivel de ruido que utiliza un factor de escala de la banda de factor de escala predeterminada;  
decuantizar las líneas espectrales dentro de las segundas bandas de factor de escala que utilizan los factores de  
escala de las segundas bandas de factor de escala; y

transformar inversamente el espectro obtenido de las primeras bandas de factor de escala relleno con ruido cuyo nivel  
se ha ajustado con los factores de escala de las primeras bandas de factor de escala, y las segundas bandas de factor  
45 de escala decuantizadas con los factores de escala de las segundas bandas de factor de escala, a fin de obtener una  
parte en el dominio del tiempo del primer canal de la señal de audio multicanal.

13. Procedimiento de decodificación de audio paramétrico en el dominio de la frecuencia que comprende  
50 cuantizar las líneas espectrales de un espectro de un primer canal de una trama actual de una señal de audio  
multicanal que utiliza los factores de escala preliminares de las bandas de factor de escala dentro del espectro;

identificar las primeras bandas de factor de escala en el espectro dentro de las cuales todas las líneas espectrales se  
cuantizan a cero, y las segundas bandas de factor de escala del espectro dentro de las cuales, al menos una línea  
espectral, se cuantiza a un valor diferente de cero,

55 rellenar las líneas espectrales dentro de una banda de factor de escala predeterminada de las primeras bandas de  
factor de escala con ruido generado utilizando

las líneas espectrales de una mezcla descendente de una trama anterior de la señal de audio multicanal,

con un ajuste del nivel de ruido que utiliza un factor de escala real de la banda de factor de escala predeterminada; y señalar el factor de escala real de la banda de factor de escala predeterminada en lugar del factor de escala preliminar.

5

14. Programa informático que tiene un código de programa que realiza, cuando se ejecuta en un ordenador, un procedimiento según la reivindicación 12 o 13.

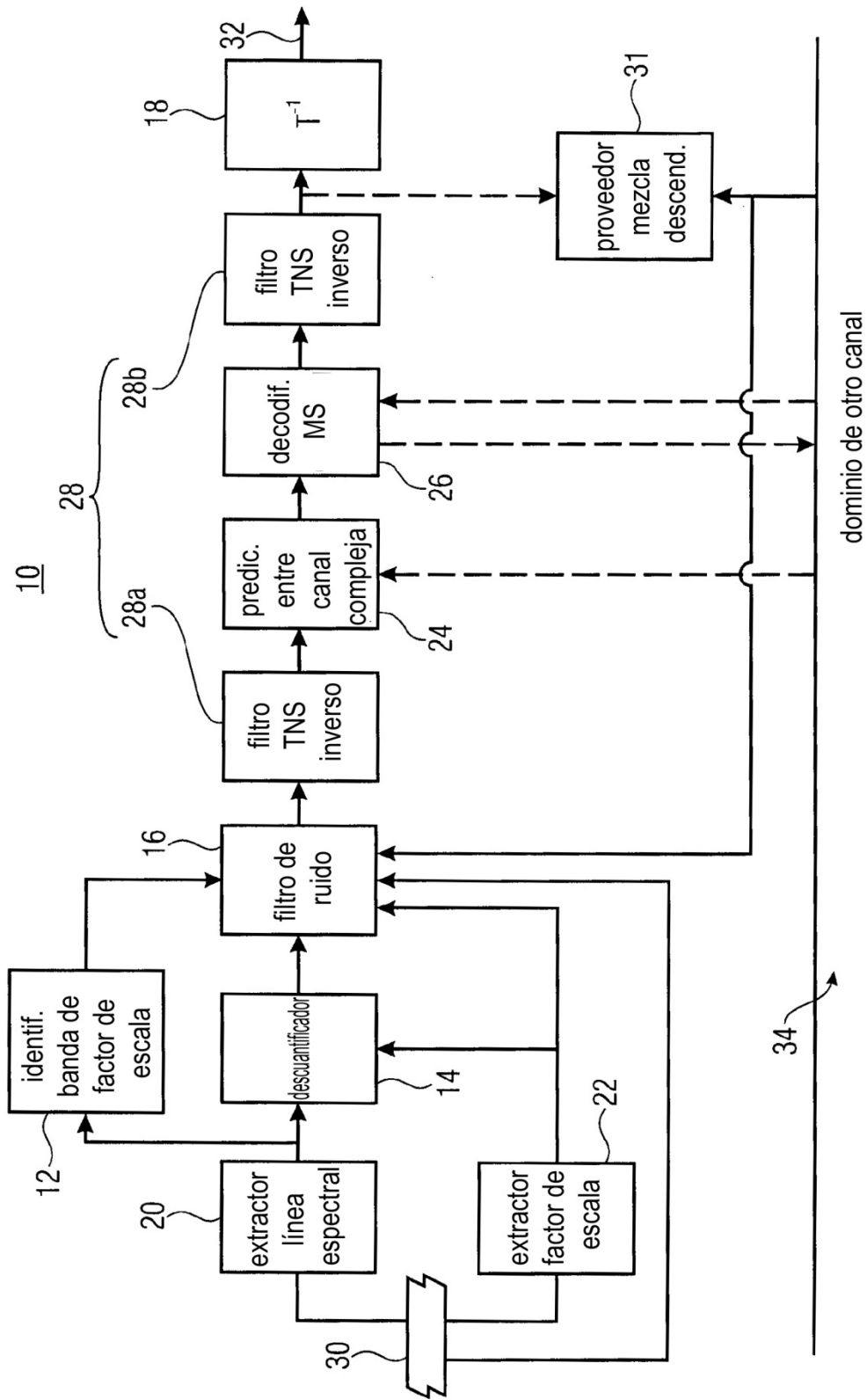


FIG 1

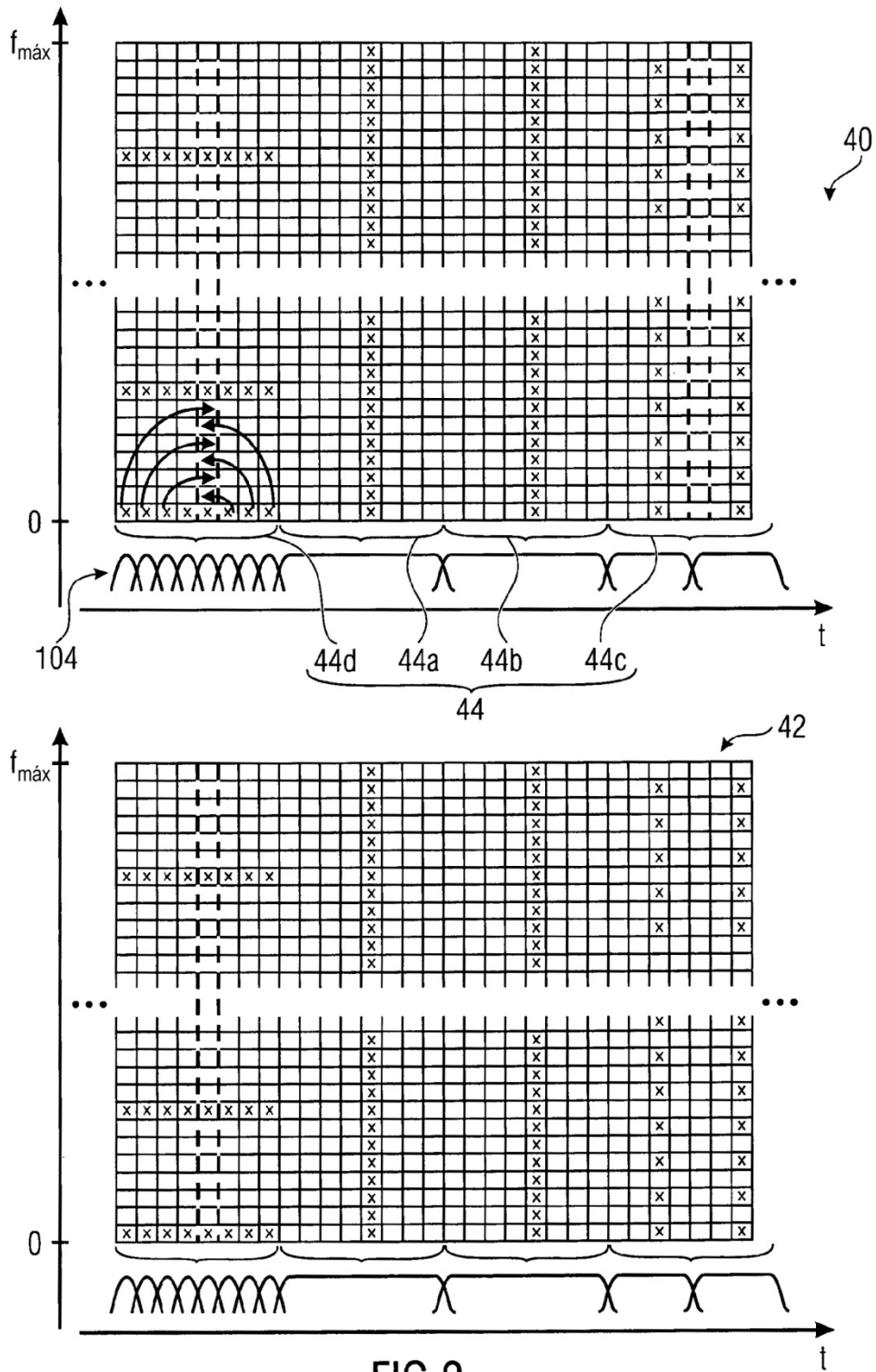


FIG 2

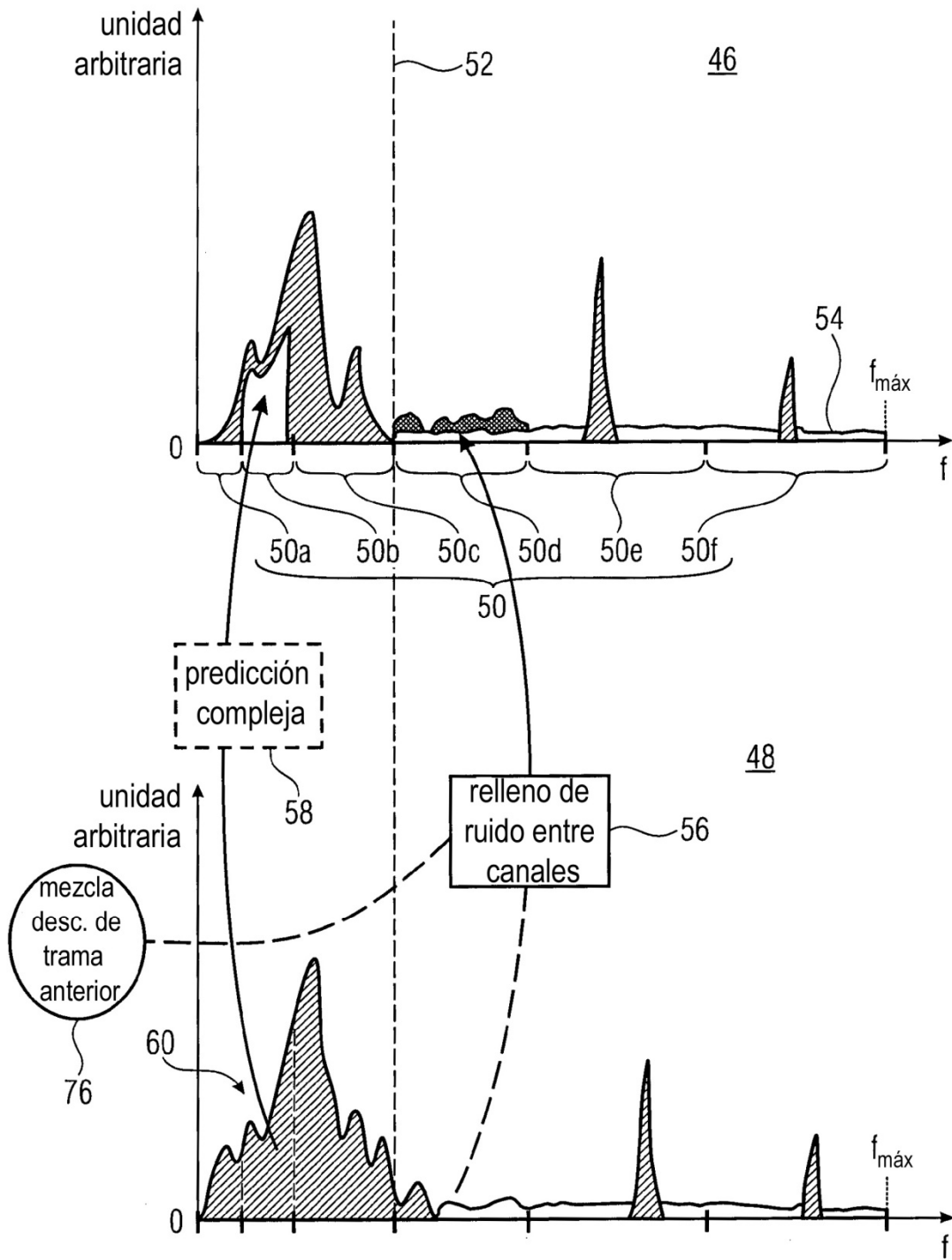
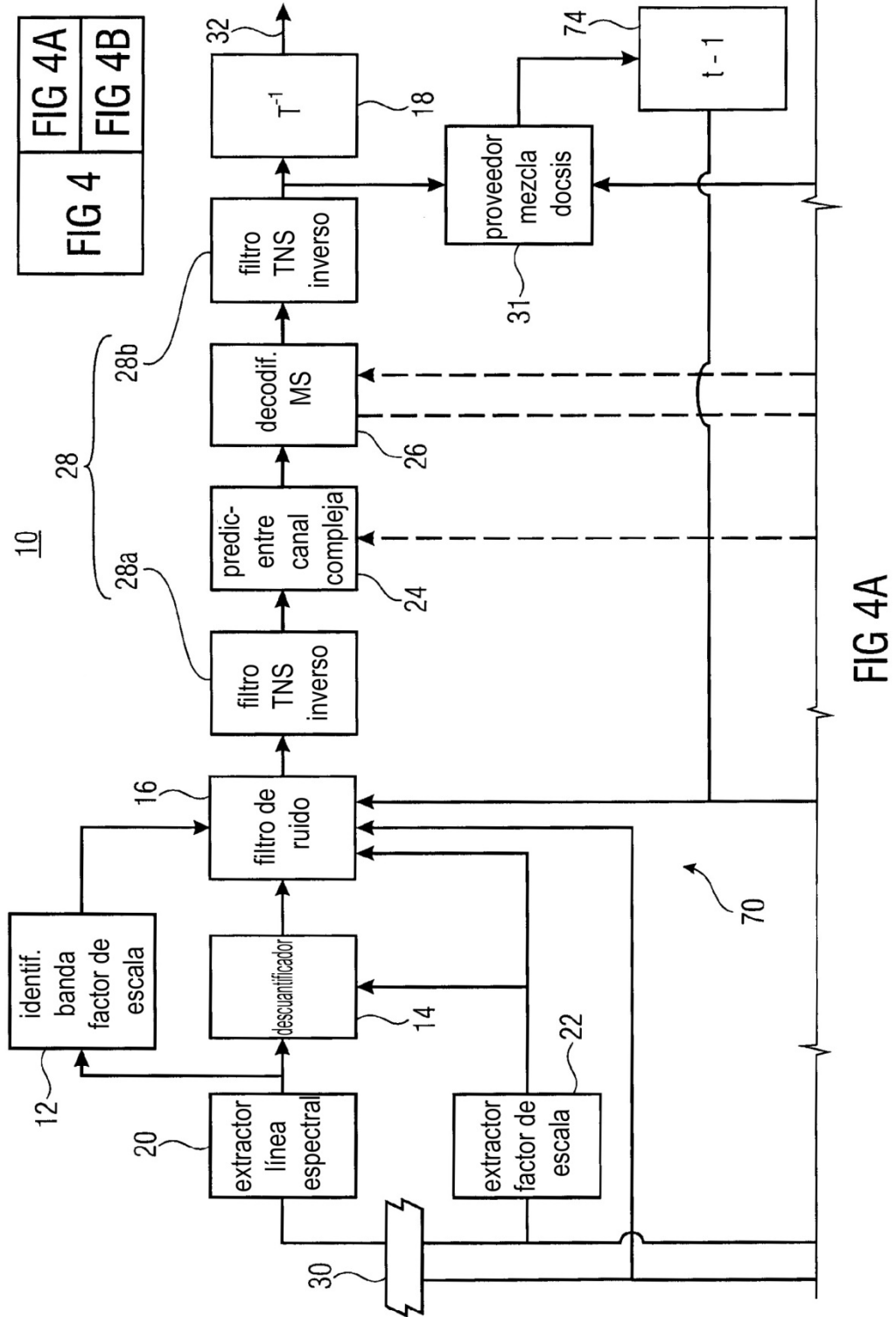


FIG 3



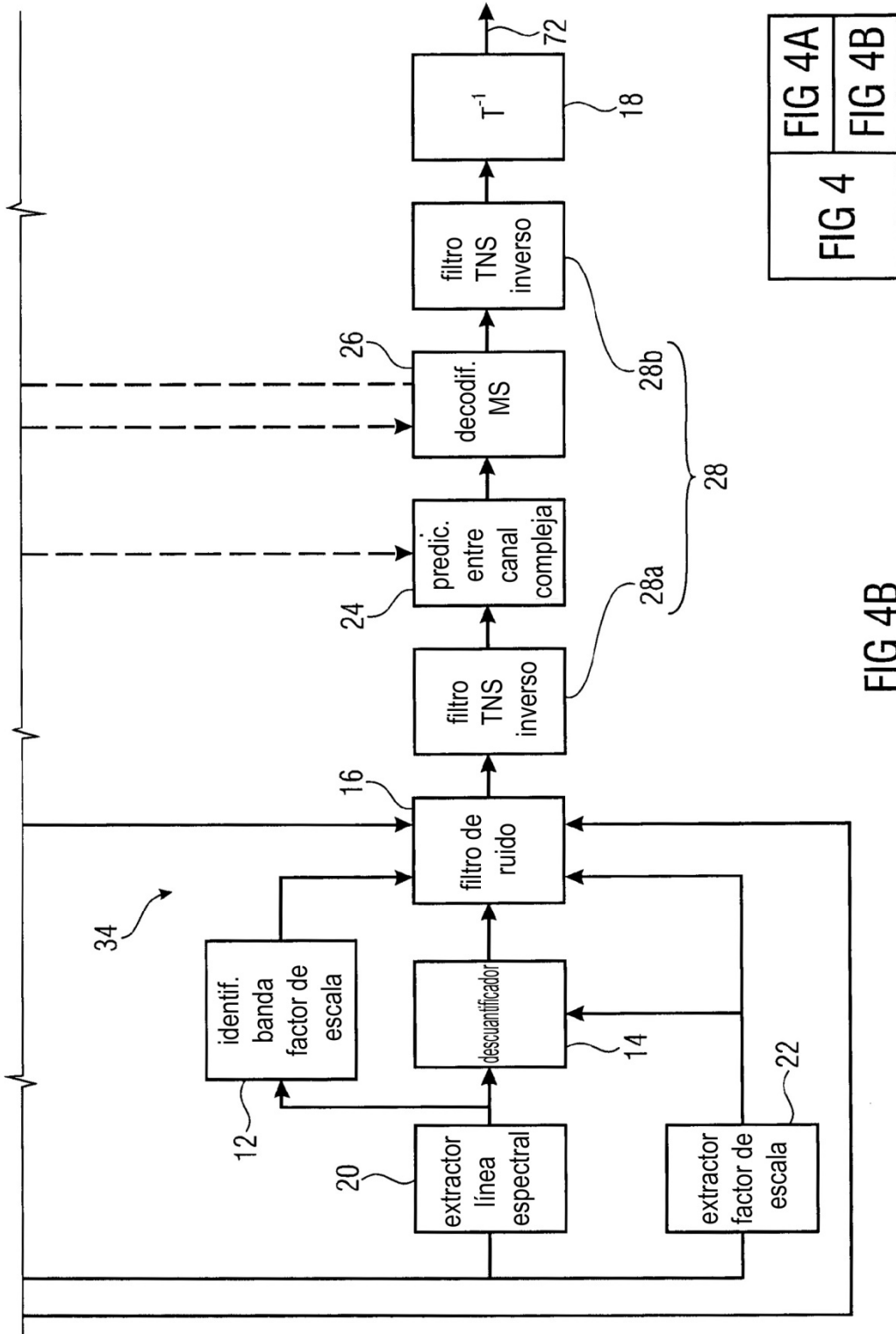


FIG 4	FIG 4A
	FIG 4B

FIG 4B



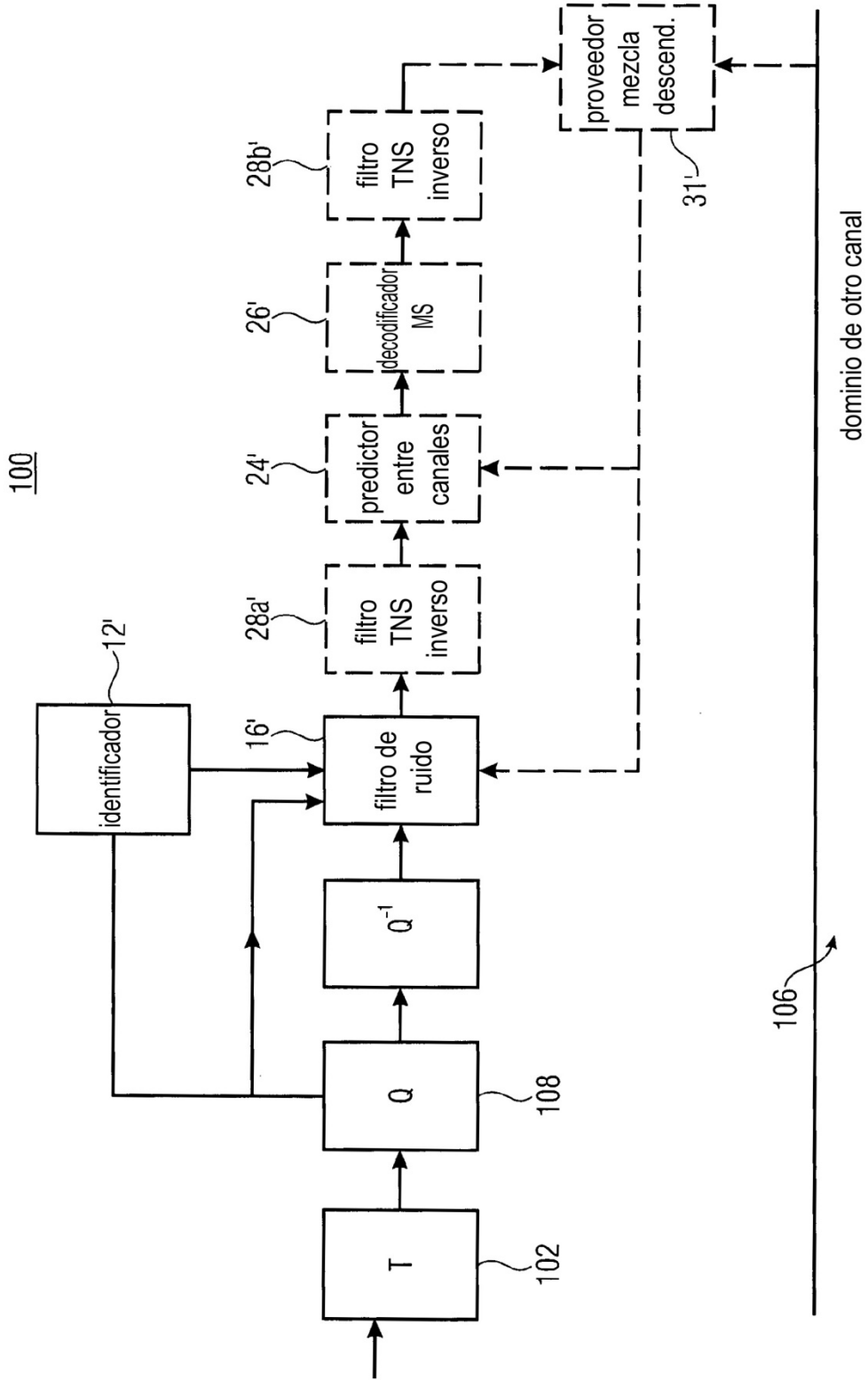


FIG 5