

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 644 131**

51 Int. Cl.:

G10L 19/00 (2013.01)

G10L 19/08 (2013.01)

G10L 19/032 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.06.2013 PCT/EP2013/062809**

87 Fecha y número de publicación internacional: **03.01.2014 WO14001182**

96 Fecha de presentación y número de la solicitud europea: **19.06.2013 E 13730249 (3)**

97 Fecha y número de publicación de la concesión europea: **02.08.2017 EP 2867892**

54 Título: **Predicción lineal basada en una codificación de audio utilizando un estimador mejorado de distribución de probabilidad**

30 Prioridad:

28.06.2012 US 201261665485 P

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

27.11.2017

73 Titular/es:

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

72 Inventor/es:

**BÄCKSTRÖM, TOM;
HELMRICH, CHRISTIAN;
FUCHS, GUILLAUME;
MULTRUS, MARKUS y
DIETZ, MARTIN**

74 Agente/Representante:

SALVA FERRER, Joan

ES 2 644 131 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

Predicción lineal basada en una codificación de audio utilizando un estimador mejorado de distribución de probabilidad.

5

[0001] La presente invención se relaciona con la codificación de audio basada en la predicción lineal y, en particular, la codificación de audio basada en la predicción lineal usando codificación de espectro.

[0002] El enfoque clásico para la cuantificación y codificación en el dominio de la frecuencia es tomar 10 ventanas (superpuestas) de la señal, realizar una transformación de tiempo-frecuencia, aplicar un modelo perceptivo y cuantificar las frecuencias individuales con un codificador de entropía, como un codificador aritmético [1]. El modelo perceptivo es básicamente una función de ponderación que se multiplica sobre las líneas espectrales de forma que los errores en cada línea espectral ponderada tengan un impacto perceptivo igual. Todas las líneas 15 ponderadas pueden así ser cuantificadas con la misma precisión, y la precisión general determina el compromiso entre la calidad perceptual y el consumo de bits.

[0003] En AAC y en el modo de dominio de frecuencia de USAC (no-TCX), el modelo perceptivo se definió a nivel de banda de forma que un grupo de líneas espectrales (la banda espectral) tendría la misma ponderación. Estas ponderaciones se conocen como factores de escala, ya que definen por qué factor se escala la banda. 20 Además, los factores de escala se codificaron de forma diferencial.

[0004] En el dominio TCX, las ponderaciones no se codifican usando factores de escala, sino mediante un modelo LPC [2] que define el envolvente espectral, que es la forma general del espectro. El LPC se utiliza porque permite un intercambio suave entre TCX y ACELP. Sin embargo, el LPC no se corresponde bien con el modelo 25 perceptivo, que debería ser mucho más suave, donde un proceso conocido como ponderación se aplica al LPC de forma que el LPC se corresponde aproximadamente con el modelo perceptivo deseado.

[0005] En el dominio TCX de USAC, las líneas espectrales se codifican mediante un codificador aritmético. Un codificador aritmético se basa en asignar probabilidades a todas las configuraciones posibles de la señal, de 30 forma que los valores de alta probabilidad pueden codificarse con un número pequeño de bits, de forma que el consumo de bits se minimiza. Para estimar la distribución de probabilidad de las líneas espectrales, el códec emplea un modelo de probabilidad que predice la distribución de la señal en base a líneas anteriores ya codificadas en el espacio tiempo-frecuencia. Las líneas anteriores se conocen como el contexto de la línea actual a codificar [3].

35 **[0006]** Recientemente, NTT propuso un procedimiento para mejorar el contexto del codificador aritmético (comparar [4]). Se basa en el uso de la LTP para determinar las posiciones aproximadas de las líneas armónicas (filtro-comp) y re-disponiendo las líneas espectrales de forma que la predicción de la magnitud desde el contexto es más eficiente. La Solicitud de Patente Europea EP 2077550 A1 describe un sistema de codificación de audio en el cual las distribuciones de probabilidad para los valores de las líneas espectrales a ser codificadas o decodificadas se 40 determinan en base al parámetro de ganancia de la predicción a largo plazo (LTP), donde una variación espectral más grande se asume para ganancias LTP más altas. En general, cuanto mejor sea la estimación de la distribución de probabilidad, más eficiente será la compresión conseguida por la codificación de entropía. Debería ser favorable tener un concepto a mano que posibilitaría la consecución de una estimación de distribución de probabilidad de calidad similar a la obtenible usando cualquiera de las técnicas indicadas anteriormente, pero con una menor 45 complejidad.

[0007] Por tanto, es un objetivo de la presente invención proporcionar un esquema de codificación de audio basado en una predicción lineal que tenga características mejoradas. Este objetivo se consigue mediante la materia 50 de las reivindicaciones independientes.

[0008] Es una conclusión básica de la presente invención que la codificación de audio basada en la predicción lineal puede mejorarse codificando un espectro compuesto por una pluralidad de componentes espectrales usando una estimación de distribución de probabilidad determinada para cada uno de los componentes 55 espectrales a partir de la información del coeficiente de predicción lineal. En particular, la información del coeficiente de predicción lineal está disponible de todas formas. Por tanto, puede utilizarse para determinar la estimación de distribución de probabilidad tanto en la parte de codificación como en la de decodificación. La determinación de esta última puede implementarse de una forma computacionalmente sencilla usando, por ejemplo, una parametrización apropiada para la estimación de la distribución de probabilidad en la pluralidad de los componentes espectrales. En general, la eficiencia de codificación como proporciona la codificación de entropía es compatible con las

estimaciones de distribución de probabilidad como se consigue usando selección de contexto, pero su derivación es menos compleja. Por ejemplo, la derivación puede ser puramente analíticamente y/o no requerir ninguna información sobre los atributos de las líneas espectrales adyacentes como los valores espectrales codificados/decodificados previamente de las líneas espectrales adyacentes como es el caso de la selección de contexto espacial. Esto, a su vez, 5
renderiza la paralelización de los procesos de computación de forma más fácil, por ejemplo. Además, pueden ser necesarios menos requisitos de memoria y menos accesos a la memoria.

[0009] De acuerdo con una realización de la presente aplicación, el espectro, los valores espectrales del cual están codificados por entropía usando la estimación de probabilidad determinada como simplemente indicada, 10
puede ser una excitación codificada de transformación obtenida usando la información del coeficiente de predicción lineal.

[0010] De acuerdo con una realización de la presente aplicación, por ejemplo, el espectro es una excitación codificada de transformación definida, sin embargo, en un dominio ponderado perceptualmente. Es decir, el espectro 15
codificado con entropía usando la estimación de distribución de probabilidad determinada se corresponde con un espectro de señales de audio pre-filtrado usando una función de transformación correspondiente a un filtro de síntesis de predicción lineal ponderada perceptualmente definido por la información del coeficiente de predicción lineal y para cada uno de la pluralidad de componentes espectrales una pluralidad de parámetros de distribución se determina de forma que los parámetros de distribución de probabilidad siguen, por ejemplo, son una versión a 20
escala, a una función que depende de un producto de una función de transferencia del filtro de síntesis de predicción lineal y una inversa de una función de transferencia de la modificación ponderada perceptualmente del filtro de síntesis de predicción lineal. Para cada uno de la pluralidad de componentes espectrales, la estimación de distribución de la pluralidad es entonces una función parametrizable parametrizada con el parámetro de distribución de probabilidad del componente espectral respectivo. De nuevo, la información del coeficiente de predicción lineal 25
está disponible de todas formas, y la derivación del parámetro de distribución de probabilidad puede implementarse como un proceso puramente analítico y/o un proceso que no requiere ninguna interdependencia entre los valores espectrales en diferentes componentes espectrales del espectro.

El parámetro de distribución de probabilidad se determina adicionalmente de forma que los parámetros de distribución de probabilidad sigan espectralmente una función que depende multiplicativamente de la estructura fina 30
espectral que a su vez se determina utilizando la predicción a largo plazo (LTP). De nuevo, en algunos códecs basados en la predicción lineal, la información LTP está disponible de todas formas y más allá de esto, la determinación de los parámetros de distribución de probabilidad es factible que pueda realizarse puramente de forma analítica y/o sin interdependencias entre la codificación de los valores espectrales de los diferentes componentes espectrales del espectro. Cuando se combina en uso LTP con la codificación de excitación codificada 35
de transformación perceptual, la eficiencia de la codificación se mejora aún más mientras que la complejidad aumenta moderadamente.

[0011] Las implementaciones y realizaciones ventajosas están sujetas a las reivindicaciones dependientes. Las realizaciones preferidas de la presente solicitud se describen más adelante con respecto a las figuras, entre las 40
cuales,

la Fig. 1 muestra un diagrama de bloque de un codificador de audio basado en predicción lineal de acuerdo con una realización

La Fig. 2 muestra un diagrama de bloque de un determinador de espectro de la Fig. 1 de acuerdo con una 45
realización;

La Fig. 3a muestra diferentes funciones de transferencia que se producen en la descripción del modo de funcionamiento de los elementos mostrados en las Figs. 1 y 2 cuando se implementan las mismas usando la codificación perceptual;

La Fig. 3b muestra las funciones de la Fig. 3a ponderadas, sin embargo, usando la inversa del modelo perceptivo;

La Fig. 4 muestra un diagrama de bloque que ilustra el funcionamiento interno del estimador de distribución de probabilidad de la Fig. 1 de acuerdo con una realización utilizando codificación perceptual;

La Fig. 5a muestra un gráfico que ilustra una señal de audio original tras el filtrado pre-énfasis y su envolvente estimado;

La Fig. 5b muestra un ejemplo de una función LTP utilizada para estimar más de cerca el envolvente de acuerdo con 55
una realización;

La Fig. 5c muestra un gráfico que ilustra el resultado de la estimación del envolvente, aplicando la función LTP de la Fig. 5b al ejemplo de la Fig. 5a;

La Fig. 6 muestra un diagrama de bloque del funcionamiento interno del estimador de distribución de probabilidad 16 en una realización adicional usando codificación perceptual además de procesamiento LTP;

La Fig. 7 muestra un diagrama de bloque de un decodificador de audio basado en predicción lineal de acuerdo con una realización;

La Fig. 8 muestra un diagrama de bloque de un decodificador de audio basado en predicción lineal de acuerdo con una realización adicional;

5 La Fig. 9 muestra un diagrama de bloque del filtro de la Fig. 8 de acuerdo con una realización;

La Fig. 10 muestra un diagrama de bloque de una estructura más detallada de una parte del codificador de la Fig. 1 posicionada en una etapa de cuantificación y codificación de entropía, y el estimador de distribución de probabilidad 14 de acuerdo con una realización; y la Fig. 11 muestra un diagrama de bloque de una parte dentro del

10 mismo que se corresponde con la parte en la cual la Fig. 10 está ubicada en el lado de codificación, es decir, ubicada en el estimador de distribución de probabilidad 102 y la etapa de decodificación por entropía y descuantificación 104 de acuerdo con una realización.

[0012] Antes de describir las diferentes realizaciones de la presente solicitud, las ideas subyacentes en las mismas se explican de forma ejemplar contra el antecedente indicado en la parte introductoria de la especificación de la presente solicitud. Las características específicas que, derivadas de la comparación con las técnicas de comparación concreta como USAC, no serán tratadas como restrictivas del alcance de la presente solicitud y sus realizaciones.

20 **[0013]** En el enfoque USAC para codificación aritmética, el contexto prácticamente predice la distribución de magnitud de las siguientes líneas. Es decir, las líneas espectrales o componentes espectrales se analizan en dimensiones espectrales durante la codificación/decodificación, y la distribución de magnitud se predice continuamente dependiendo de los vales espectrales codificados/decodificados previamente. Sin embargo, el LPC ya codifica la misma información de forma explícita, sin necesidad de predicción. Por tanto, emplear el LPC en vez de este contexto arrojaría un resultado similar, sin embargo, a una complejidad computacional inferior o al menos con la posibilidad de conseguir una menor complejidad. De hecho, ya que a una tasa de bits baja el espectro consiste esencialmente de unos a ceros, el contexto casi siempre será muy escaso y carecerá de información útil. Por tanto, en teoría, el LPC debería de hecho ser una fuente mucho mejor para las estimaciones de magnitud ya que la plantilla de los valores espectrales adyacentes ya codificados/decodificados, usados para la estimación de la

25 distribución de probabilidad, meramente se completa escasamente con información útil. Además, la información LPC ya está disponible en el codificador y el decodificador, donde tiene un coste cero en términos de consumo de bits.

[0014] El modelo LPC solo define la forma envolvente espectral, que es la magnitud relativa de cada línea, pero no la magnitud absoluta. Para definir una distribución de probabilidad para una única línea, siempre necesitamos la magnitud absoluta, es decir un valor para la variación de señal (o una medida similar). Una parte esencial de la mayoría de los modelos de cuantificador espectral basado en LPC debería por tanto ser una escalación del envolvente de LPC, de forma que la variación deseada (y, por tanto, el consumo de bits deseado), se alcance. Esta escalación debería normalmente realizarse en el codificador y el decodificador ya que las distribuciones de probabilidad para cada línea entonces dependen del LPC escalado.

40 **[0015]** Como se ha descrito anteriormente, el modelo perceptivo (LPC ponderado) puede utilizarse para definir el modelo perceptivo, es decir, la cuantificación puede realizarse en el dominio perceptivo de forma que el error de cuantificación esperado en cada línea espectral causa aproximadamente una cantidad igual de distorsión perceptual. Por tanto, si es así, el modelo LPC se transforma al dominio perceptivo también multiplicándolo con el LPC ponderado como se define a continuación. En las realizaciones descritas a continuación, se asume a menudo que el envolvente del LPC se transforma para el dominio perceptivo.

[0016] Así, es posible aplicar un modelo de probabilidad independiente para cada línea espectral. Es razonable asumir que las líneas espectrales no tienen correlación de fase predecible, donde sea suficiente para modelar solo la magnitud. Ya que se puede presumir que el LPC codifica la magnitud de forma eficiente, tener un codificador aritmético basado en contexto probablemente no mejorará la eficiencia de la estimación de magnitud.

[0017] Por tanto, es posible aplicar un codificador de entropía basado en contexto, como del que depende el contexto, o incluso consiste en, el envolvente LPC.

55 **[0018]** Además del envolvente LPC, el LTP puede utilizarse para inferir la información del envolvente. Después de todo, el LTP puede corresponderse a un filtro de peine en el dominio de frecuencia. A continuación, se explican algunos detalles prácticos.

[0019] Tras haber explicado algunos pensamientos que llevaron a la idea en la que se basan las realizaciones descritas a continuación, la descripción de estas realizaciones ahora comienza con respecto a la Fig. 1, que muestra una realización de un codificador de audio basado en predicción lineal de acuerdo con una realización de la presente solicitud. El codificador de audio basado en predicción lineal de la Fig. 1 se indica
5 generalmente usando el signo de referencia 10 y comprende un analizador de predicción lineal 12, una estimación de distribución de probabilidad 14, un determinador de espectro 16 y una etapa de codificación de cuantificación y entropía 18. El codificador de audio basado en predicción lineal 10 de la Fig. 1, recibe una señal de audio que se codificará en, por ejemplo, una entrada 20 y produce un flujo de datos 22 que de forma correspondiente tiene una señal de audio codificada en el mismo. El analizador LP 12 y el determinador de espectro 16 están, como se
10 muestra en la Fig. 1, directamente o indirectamente conectados con la entrada 20. El estimador de distribución de probabilidad 14 se conecta entre el analizador LP 12 y la etapa de cuantificación y codificación de entropía 18, y la etapa de cuantificación y codificación de entropía 18, a su vez, se conecta a una salida del determinador del espectro 16. Como puede verse en la Fig. 1, el analizador LP 12 y la etapa de cuantificación y codificación de entropía 18 contribuyen a la formación/generación del flujo de datos 22. Como se describirá con más detalle a
15 continuación, el codificador 10 puede comprender opcionalmente un filtro pre-énfasis 24, que puede conectarse entre la entrada 20 y el analizador LP 12 y/o el determinador de espectro 16. Además, el determinador de espectro 16 puede conectarse opcionalmente a la salida del analizador LP 12.

[0020] En particular, el analizador LP 12 está configurado para determinar la información del coeficiente de predicción lineal basada en la señal de audio entrante en la entrada 20. Como se describe en la Fig. 1, el analizador LP 12 puede realizar análisis de predicción lineal en la señal de audio en la entrada 20 directamente o sobre alguna
20 versión modificada de la misma, como, por ejemplo, una versión pre-enfatizada de la misma obtenida mediante el filtro de pre-énfasis 24. El modo de funcionamiento del analizador LP 12, por ejemplo, implica una división en ventanas de la señal entrante para obtener una secuencia de partes con sistema de ventanas de la señal que será
25 analizada por LP, una determinación de autocorrelación para determinar la autocorrelación de cada parte en ventana y sistema de ventana con desfase, que es opcional, para aplicar una función de ventana con desfase sobre las autocorrelaciones. La estimación del parámetro de predicción lineal puede entonces realizarse sobre las autocorrelaciones o la salida de la ventana con desfase, es decir, las funciones de autocorrelación con sistema de
30 ventanas. La estimación del parámetro de predicción lineal puede, por ejemplo, implicar el rendimiento de un algoritmo Wiener-Levinson-Durbin u otro algoritmo adecuado sobre las autocorrelaciones (con sistema de ventanas desfasadas) para así obtener coeficientes de predicción lineal por autocorrelación, es decir, por parte con sistema de
35 ventanas de la señal que será analizada por LP. Es decir, en la salida del analizador LP 12, el resultado de los coeficientes LPC que son, como se describirá más adelante, usados por el estimador de distribución de probabilidad 14 y, opcionalmente el determinador de espectro 16. El analizador LP 12 puede configurarse para cuantificar el
40 coeficiente de predicción lineal para la inserción en el flujo de datos 22. La cuantificación de los coeficientes de predicción lineal puede realizarse en otro dominio distinto al dominio de coeficiente de predicción lineal como, por ejemplo, en un par espectral de línea o dominio de frecuencia espectral de línea. Los coeficientes de predicción lineal cuantificados pueden codificarse en el flujo de datos 22. La información del coeficiente de predicción lineal usada realmente por el estimador de distribución de probabilidad 14 y, opcionalmente, por el determinador de
45 espectro 16, puede tener en cuenta la pérdida de cuantificación, es decir, puede ser la versión cuantificada que se transmite sin pérdidas a través del flujo de datos. Es decir, este último puede en realidad usar como información del coeficiente de predicción lineal los coeficientes de predicción lineal cuantificada tal como los obtiene el analizador de predicción lineal 12. Meramente con fines de integridad, se nota que existe una enorme cantidad de posibilidades de realizar la determinación de la información del coeficiente de predicción por medio del analizador de predicción lineal
50 12. Por ejemplo, pueden usarse otros algoritmos distintos al algoritmo Wiener-Levinson-Durbin. Además, una estimación de la autocorrelación local de la señal que se analizará por LP puede obtenerse en base a una descomposición espectral de la señal que será analizada por LP. En WO 2012/110476 A1, por ejemplo, se describe que la autocorrelación puede obtenerse dividiendo la señal que será analizada por LP en ventanas, sometiendo
55 cada parte dividida en ventanas a una MDCT, determinando el espectro de potencia por espectro MDCT y realizando una ODFT inversa para realizar la transición desde el dominio MDCT a una estimación de la autocorrelación. Para resumir, el analizador LP 12 proporciona información del coeficiente de predicción lineal y el flujo de datos 22 expresa o comprende esta información del coeficiente de predicción lineal. Por ejemplo, el flujo de datos 22 expresa la información del coeficiente de predicción lineal en la resolución temporal que se determina mediante la frecuencia de la parte con sistema de ventanas que acabamos de mencionar, donde las partes con
sistema de ventanas pueden, como se conoce en la técnica, superponerse unas a otras, como por ejemplo con una superposición del 50%.

[0021] En cuanto al filtro de pre-énfasis 24, se nota que el mismo puede, por ejemplo, implementarse usando filtrado FIR. El filtro de pre-énfasis 24 puede, por ejemplo, tener una función de transferencia de pase alto. De

acuerdo con una realización, el filtro de pre-énfasis 24 se realiza como un filtro de paso alto de orden $n-t$, como, por ejemplo, $H(z) = 1 - \alpha z^{-1}$ con α establecido, por ejemplo, a 0,68.

[0022] A continuación, se describirá el determinador de espectro. El determinador de espectro 16 se configura para determinar un espectro compuesto por una pluralidad de componentes espectrales basados en la señal de audio en la entrada 20. El espectro describirá la señal de audio. De forma similar al analizador de predicción lineal 12, el determinador de espectro 16 puede operar sobre la señal de audio 20 directamente, o sobre alguna versión modificada de la misma, como por ejemplo la versión filtrada por pre-énfasis de la misma. El determinador de espectro 16 puede usar cualquier transformación para determinar el espectro como, por ejemplo, una transformación solapada o incluso una transformación solapada críticamente muestreada, como, por ejemplo, una MDCT, aunque existen otras posibilidades. Es decir, el determinador de espectro 16 puede someter la señal para que sea descompuesta espectralmente a sistema de ventanas para así obtener una secuencia de partes en ventanas y someter cada parte en ventanas a una transformación respectiva como una MDCT. La frecuencia con sistema de ventanas del determinador de espectro 16, es decir, la resolución temporal de la descomposición espectral puede diferir de la resolución temporal en la cual el analizador LP 12 determina la información del coeficiente de predicción lineal.

[0023] El determinador de espectro 16 así produce un espectro compuesto de una pluralidad de componentes espectrales. En particular, el determinador de espectro 16 puede producir, por parte con ventanas que está sometida a una transformación, una secuencia de valores espectrales, es decir, un valor espectral por componente espectral, por ejemplo, por línea espectral de frecuencia. Los valores espectrales pueden ser valores complejos o valores reales. Los valores espectrales son valores reales en el caso de que se utilice una MDCT, por ejemplo, En particular, los valores espectrales pueden estar señalizados, es decir, pueden ser una combinación de señal y magnitud.

[0024] Como se ha indicado anteriormente, la información del coeficiente de predicción lineal forma una predicción a corto plazo del envolvente espectral de la señal analizada por LP y puede, por tanto, servir como base para determinar, para cada una de las pluralidades de componentes espectrales, una estimación de la distribución de probabilidad, es decir, una estimación de cómo, estadísticamente, la probabilidad de que el espectro en el componente espectral respectivo, asuma un cierto valor espectral posible, y varíe sobre el dominio de posibles valores espectrales. La determinación se realiza mediante el estimador de distribución de probabilidad 14. Existen diferentes posibilidades con respecto a los detalles de determinación de la estimación de distribución de probabilidad. Por ejemplo, aunque el determinador de espectro 16 podría implementarse para determinar el espectrograma de la señal de audio o la versión pre-enfatizada de la señal de audio, de acuerdo con las realizaciones descritas adicionalmente a continuación, el determinador de espectro 16 se configura para determinar, como el espectro, una señal de excitación, es decir, una señal residual obtenida mediante filtrado basado en LP de la señal de audio o alguna versión modificada de la misma, como la versión filtrada por pre-énfasis de la misma. En particular, el determinador de espectro 16 puede configurarse para determinar el espectro de la señal entrante al determinador de espectro 16, tras filtrar la señal entrante usando una función de transferencia que depende de, o es igual a, una inversa de un filtro de síntesis de predicción lineal definido por la información del coeficiente de predicción lineal, es decir, el filtro de análisis de predicción lineal. Alternativamente, el codificador de audio basado en LP puede ser un codificador de audio basado en LP perceptivo y el determinador de espectro 16 puede configurarse para determinar el espectro de la señal entrante al determinador de espectro 16, tras filtrar la señal entrante usando una función de transferencia que depende de, o es igual a, una inversa de un filtro de síntesis de predicción lineal definido por la información del coeficiente de predicción lineal, pero ha sido modificado para, por ejemplo, corresponderse a la inversa de una estimación de un umbral de enmascaramiento. Es decir, el determinador de espectro 16 podría configurarse para determinar el espectro de la señal entrante, filtrado con una función de transferencia que se corresponde a la inversa de un filtro de síntesis de predicción lineal modificado de forma preceptiva. En tal caso, el determinador de espectro 16 reduce comparativamente el espectro en las regiones espectrales donde el enmascaramiento preceptivo es más alto en relación con las regiones espectrales donde el enmascaramiento perceptivo es más bajo. Al usar la información del coeficiente de predicción lineal, el estimador de distribución de probabilidad 14 aún es capaz, sin embargo, de estimar el envolvente del espectro determinado por el determinador del espectro 16, es decir, teniendo en cuenta la modificación preceptiva del filtro de síntesis de predicción a la hora de determinar la estimación de la distribución de probabilidad. Los detalles en este aspecto se describen adicionalmente a continuación.

[0025] Además, como se describirá con más detalle a continuación, el estimador de distribución de probabilidad 14 es capaz de usar la predicción a largo plazo para obtener una información de estructura fina sobre el espectro, para así obtener una mejor estimación de la distribución de probabilidad por componente espectral. El

parámetro o parámetros LTP se envían, por ejemplo, a decodificación para así permitir una reconstrucción de la información de estructura fina. Los detalles en este aspecto se describen más adelante.

[0026] En cualquier caso, la etapa de cuantificación y codificación de entropía 18 se configura para codificar por cuantificación y entropía el espectro usando la estimación de distribución de probabilidad como se determina para cada una de las pluralidades de los componentes espectrales mediante el estimador de distribución de probabilidad 14. Para ser más precisos, la etapa de cuantificación y codificación de entropía 18 recibe desde el determinador del espectro 16 un espectro 26 compuesto por componentes espectrales k , o para ser más precisos, una secuencia de espectros 26 a alguna frecuencia temporal que se corresponde con la frecuencia de la parte con sistema de ventanas mencionada anteriormente de las partes con sistema de ventanas sujetas a la transformación. En particular, la etapa 18 puede recibir un valor de señal por valor espectral en el componente espectral k y una magnitud correspondiente $|x_k|$ por componente espectral k .

[0027] Por otro lado, la etapa de cuantificación y codificación de entropía 18 recibe, por componente espectral k , una estimación de distribución de probabilidad 28 que define, para cada valor posible que el valor espectral puede asumir, una estimación del valor de probabilidad que determina la probabilidad del valor espectral en el componente espectral k teniendo este muy posible valor. Por ejemplo, la estimación de la distribución de probabilidad determinada mediante el estimador de distribución de probabilidad 14, se concentra en las magnitudes de los valores espectrales solamente, y determina, por tanto, los valores de probabilidad para los valores positivos, incluyendo cero, únicamente. En particular, la etapa de cuantificación y codificación de entropía 18 cuantifica los valores espectrales, por ejemplo, utilizando una regla de cuantificación que es igual para todos los componentes espectrales. Los niveles de magnitud para los componentes espectrales k , obtenidos de este modo, se definen por tanto sobre un dominio de enteros incluyendo cero hasta, opcionalmente, algún valor máximo. La estimación de la distribución de probabilidad podría, para cada componente espectral k , definirse sobre este dominio de posibles enteros i , es decir, $p(k, i)$ sería la estimación de probabilidad para el componente espectral k y estará definido sobre el entero $i \in [0; \max]$ con el entero $k \in [0; k_{\max}]$ con k_{\max} siendo el componente espectral máximo y $p(k; i) \in [0; 1]$ para todo k, i y la suma sobre $p(k, i)$ sobre todo $i \in [0; \max]$ siendo uno para todo k .

[0028] La etapa de cuantificación y codificación de entropía 18 puede, por ejemplo, utilizar un tamaño de paso de cuantificación constante para la cuantificación con el tamaño del paso siendo igual para todos los componentes espectrales k . Cuanto mejor sea la estimación de la distribución de probabilidad 28, mejor será la eficiencia de compresión conseguida por la etapa de cuantificación y codificación de entropía 18.

[0029] Francamente, el estimador de la distribución de probabilidad 14 puede usar la información del coeficiente de predicción lineal proporcionada por el analizador LP 12 para conseguir una información sobre un envolvente 30, o forma aproximada, del espectro 26. Usando esta estimación 30 del envolvente o forma, el estimador 14 puede obtener una medida de dispersión 32 para cada componente espectral k mediante, por ejemplo, una escalación apropiada, usando un factor de escala común igual para todos los componentes espectrales, el envolvente. Estas medidas de dispersión en los componentes espectrales k pueden servir como parámetros para las parametrizaciones de las estimaciones de distribución de probabilidad para cada componente espectral k . Por ejemplo, $p(k, i)$ puede ser $f(i, l(k))$ para todo k con $l(i)$ siendo la medida de dispersión determinada en el componente espectral k , con $f(i, l)$ siendo, para cada l fija, una función apropiada de la variable i como una función monótonica como, como se define más adelante, una función Gaussiana o Laplace definida para los valores positivos i incluyendo cero, mientras l es el parámetro de la función que mide la "pendiente" o "amplitud" de la función como se explicará más adelante de forma más precisa. Usando las parametrizaciones parametrizadas, la etapa de cuantificación y codificación de entropía 18 puede así codificar por entropía de forma eficiente los valores espectrales del espectro a un flujo de datos 22. Como se verá claramente a partir de la descripción explicada más adelante con más detalle, la determinación de la estimación de distribución de probabilidad 28 puede implementarse puramente analíticamente y/o sin requerir interdependencias entre los valores espectrales de los diferentes componentes espectrales del mismo espectro 26, es decir, de forma independiente de los valores espectrales de los diferentes componentes espectrales relacionados con el mismo instante de tiempo. La etapa de cuantificación y codificación de entropía 18 podría por tanto realizar la codificación por entropía de los valores espectrales cuantificados o niveles de magnitud, respectivamente, en paralelo. La codificación por entropía real puede, a su vez, ser una codificación aritmética o una codificación de longitud variable o alguna otra forma de codificación por entropía como codificación por entropía de partición de intervalo de probabilidad o similar. En efecto, la etapa de cuantificación y codificación de entropía 18 codifica por entropía cada valor espectral en un cierto componente espectral k usando la estimación de distribución de probabilidad 28 para ese componente espectral k , de forma que un consumo de bits para un valor espectral k respectivo para su codificación a un flujo de datos 22 es inferior dentro de las partes del dominio de posibles valores del valor espectral en el componente espectral k donde la probabilidad

indicada por la estimación de distribución de probabilidad 28 es más alta, y el consumo de bits es mayor en las partes del dominio de valores posibles donde la probabilidad indicada por la estimación de distribución de probabilidad 28 es más baja. En el caso de codificación aritmética, por ejemplo, puede utilizarse la codificación aritmética basada en tabla. En el caso de codificación de longitud variable, la etapa de cuantificación y codificación de entropía puede seleccionar y aplicar diferentes tablas de palabras claves que asignen los posibles valores a las palabras claves dependiendo de la estimación de distribución de probabilidad 28 determinada por el estimador de distribución de probabilidad 14 para el componente espectral k respectivo.

10 **[0030]** La Fig. 2 muestra una implementación posible del determinador de espectro 16 de la Fig. 1. De acuerdo con la Fig. 2, el determinador de espectro 16 comprende un determinador de factores de escala 34, un transformador 36 y un conformador espectral 38. El transformador y el conformador espectral 38 se conectan en serie uno al otro entre la entrada y la salida del determinador del espectro 16 mediante el cual el determinador del espectro 16 se conecta entre la entrada 20 y la etapa de cuantificación y codificación de entropía 18 en la Fig. 1.

15 **[0031]** El determinador del factor de escala 34 está, a su vez, conectado entre el analizador LP 12 y una entrada adicional del conformador espectral 38 (consultar la Fig. 1). El determinador de factores de escala 34 se configura para usar la información del coeficiente de predicción lineal para determinar así los factores de escala. El transformador 36 descompone espectralmente la señal que recibe, para obtener un espectro original. Como se ha indicado anteriormente, la señal entrante puede ser la señal de audio original en la entrada 20 o, por ejemplo, una versión pre-enfaticada de la misma. Como también se ha indicado anteriormente, el transformador 36 puede someter internamente a la señal para que sea transformada a un sistema con ventanas, con partes, usando partes superpuestas, mientras que transforma individualmente cada parte con sistema de ventanas. Como se ha esbozado anteriormente, una MDCT puede usarse para la transformación. Es decir, el transformador 36 produce un valor espectral X'_k por componente espectral k y el conformador espectral 38 se configura para dar forma espectralmente a este espectro original escalando el espectro usando factores de escala, es decir, escalando cada valor espectral original X'_k usando los factores de escala S_k producidos por el determinador de factores de escala 34 para así obtener un valor espectral respectivo X_k , que luego se somete a codificación y cuantificación de entropía en el estado 18 de la Fig. 1.

30 **[0032]** La resolución espectral en la cual el determinador de factores de escala 34 determina los factores de escala, no coincide necesariamente con la resolución definida por el componente espectral k . Por ejemplo, un agrupamiento motivado perceptualmente de los componentes espectrales en los grupos espectrales como bandas Bark, puede formar la resolución espectral en la cual los factores de escala, es decir las ponderaciones espectrales por las cuales los valores espectrales del espectro producido por el transformador 36 se ponderan, se determinan.

35 **[0033]** El determinador de factores de escala 34 está configurado para determinar los factores de escala de forma que los mismos representen, o se aproximen, a una función de transferencia que depende de una inversa de un filtro de síntesis de predicción lineal definido por la información de coeficiente de predicción lineal. Por ejemplo, el determinador de factores de escala 34 puede estar configurado para usar los coeficientes de predicción lineal obtenidos desde el analizador LP 12 en, por ejemplo, su forma cuantificada en la cual están también disponibles en el lado de decodificación a través de un flujo de datos 22, como una base para la conversión LPC a MDCT que, a su vez, puede involucrar una ODFT. Naturalmente, también existen alternativas. En el caso de las alternativas indicadas anteriormente, donde el codificador de audio de la Fig. 1 es un codificador de audio basado en predicción lineal perceptual, el determinador de factores de escala 34 puede configurarse para realizar una ponderación motivada perceptualmente de los LPC primero antes de realizar la conversión a los factores de escala usando, por ejemplo, una ODFT. Sin embargo, pueden existir otras posibilidades. Como se explicará con más detalle más adelante, la función de transferencia del filtrado resultante de la escalación espectral por el conformador espectral 38 puede depender, a través de la determinación del factor de escala realizada por el determinador de factores de escala 34, en la inversa del filtro de síntesis de predicción lineal $1/A(z)$ definida por la información del coeficiente de predicción lineal de forma que la función de transferencia sea una inversa de una función de transferencia de $1/A(k \cdot z)$, donde k denota una constante que puede ser, por ejemplo, 0,92.

[0034] Para comprender mejor la relación mutua entre la funcionalidad del determinador del espectro por un lado y el estimador de distribución de probabilidad 14 por el otro, y la forma en que esta relación lleva al funcionamiento efectivo de la etapa de codificación y cuantificación de entropía 18 en el caso del codificador de audio basado en predicción lineal que actúa como codificador de audio basado en predicción lineal perceptual, pueden consultarse las Figs. 3a y 3b. La Fig. 3a muestra un espectro original 40. Aquí, es ejemplarmente el espectro de la señal de audio ponderado por la función de transferencia del filtro de pre-énfasis. Para ser más precisos, la Fig. 3a muestra la magnitud del espectro 40 trazada sobre los componentes espectrales o líneas espectrales k . En el

mismo gráfico, la Fig. 3 muestra la función de transferencia del filtro de síntesis de predicción lineal $A(z)$ multiplicada por la función de transferencia del filtro de pre-énfasis 24, el producto resultante siendo denotado por 42. Como puede verse, la función 42 se aproxima al envolvente o forma gruesa del espectro 40. En la Fig. 3a, la modificación motivada perceptualmente del filtro de síntesis de predicción lineal se muestra, como $A(0,92z)$ en el caso ejemplar 5 mencionado antes. Este "modelo perceptual" se denota mediante la referencia 44. La función 44 así representa una estimación simplificada de un umbral de enmascaramiento de la señal de audio teniendo en cuenta al menos oclusiones espectrales. El determinador de factores espectrales 34 determina los factores de escala para aproximar la inversa del modelo perceptual 44. El resultado de multiplicar las funciones 40 a 44 de la Fig. 3a con la inversa del modelo perceptual 44 se muestra en la Fig. 3b. Por ejemplo, 46 muestra el resultado de multiplicar el espectro 40 con la inversa de 44 y así se corresponde con el espectro ponderado perceptualmente como resultado del conformador espectral 38 en el caso de que el codificador 10 actúe como codificador basado en predicción lineal perceptual como se ha descrito anteriormente. Como función multiplicadora 44 con la inversa de la misma resulta en una función constante, el producto resultante se representa como siendo plano en la Fig. 3b; véase 50.

15 **[0035]** Ahora volviendo al estimador de distribución de probabilidad 14, el mismo tiene acceso a la información del coeficiente de predicción lineal como se ha descrito anteriormente. El estimador 14 es por tanto capaz de calcular la función 48 resultante de la función multiplicadora 42 con la inversa de la función 44. Esta función 48 puede servir, como se puede ver en la Fig. 3b, como una estimación del envolvente o la forma gruesa del pre-filtrado 46 como resultado del conformador espectral 38.

20 **[0036]** Por tanto, el estimador de distribución de probabilidad 14 podría operar como se ilustra en la Fig. 4. En particular, el estimador de distribución de probabilidad 14 podría someter los coeficientes de la predicción lineal que definen el filtro de síntesis de predicción lineal $1/A(z)$ a la ponderación perceptual 64 de forma que ésta se corresponda con un filtro de síntesis de predicción lineal modificada perceptualmente $1/A(k-z)$. Ambos, los 25 coeficientes de predicción lineal no ponderados y los ponderados se someten a la conversión LPC a ponderación espectral 60 y 62 respectivamente, y el resultado se somete, por el componente espectral k , a división. El cociente resultante se somete opcionalmente a alguna derivación de parámetro 68 donde los cocientes para los componentes espectrales k se someten individualmente, es decir para cada k , a alguna función de asignación para obtener como resultado un parámetro de distribución de probabilidad que representa una medida, por ejemplo, para la dispersión 30 de la estimación de distribución de probabilidad. Para ser más precisos, las conversiones ponderadas LPC a espectral 60, 62 aplicadas a los coeficientes de predicción lineal no ponderados y ponderados resulta en ponderaciones s_k y para los componentes espectrales k . Las conversiones 60, 62 pueden, como se ha indicado anteriormente, realizarse a una resolución espectral más baja que la resolución espectral definida por los componentes espectrales k , pero puede utilizarse por ejemplo la interpolación para suavizar el cociente resultante q_k sobre el componente espectral k . La derivación del parámetro entonces resulta en un parámetro de distribución de probabilidad π_k por componente espectral mediante, por ejemplo, escalar todo q_k usando un factor de escala común para todo k . La etapa de codificación y cuantificación de entropía 18 puede entonces usar estos parámetros de distribución de probabilidad π_k para la codificación de entropía del espectro con forma espectral de la cuantificación. En particular, como π_k es una medida para una dispersión de la estimación de la distribución de la probabilidad del 40 valor de espectro envolvente x_k o al menos su magnitud, una función parametrizable, como la mencionada anteriormente $f(i,l(k))$, puede usarse mediante la etapa de codificación y cuantificación de entropía 18 para determinar, para cada componente espectral k , la estimación de distribución de probabilidad 28 usando π_k como función parametrizable, es decir, como $l(k)$. Preferiblemente, la parametrización de la función parametrizable es tal que el parámetro de distribución de probabilidad, por ejemplo, $l(k)$, es realmente una medida para una dispersión de 45 la estimación de distribución de la probabilidad, es decir, el parámetro de distribución de probabilidad mide una anchura de la función parametrizable de distribución de probabilidad. En una realización específica descrita más adelante, se utiliza una distribución Laplace como la función parametrizable, por ejemplo, $f(i,l(k))$.

[0037] Con respecto a la Fig. 1, se nota que el estimador de distribución de probabilidad 14 puede insertar 50 adicionalmente información en el flujo de datos 22, lo que permite que el lado de decodificación aumente la calidad de la estimación de distribución de probabilidad 28 para los componentes espectrales k individuales en comparación con la calidad proporcionada únicamente basada en la información del LPC. En particular, de acuerdo con estos detalles de implementación específicos descritos de forma ejemplar explicados más adelante, el estimador de distribución de probabilidad 14 puede usar predicción a largo plazo para obtener una estimación espectralmente más 55 fina 30 del envolvente o forma del espectro 26 en el caso de que el espectro 26 represente una excitación codificada por transformación como el espectro resultante del filtrado con una función de transformación correspondiente con una inversa del modelo perceptual o la inversa del filtro de síntesis de predicción lineal.

- [0038]** Por ejemplo, véanse las Figs. 5a y 5c para ilustrar esto último, la funcionalidad opcional del estimador de distribución de probabilidad 14. La Fig. 5a muestra, como la Fig. 3a, el espectro de señales de audio originales 40 y el modelo LPC $A(z)$ incluyendo el pre-énfasis. Es decir, tenemos la señal original 40 y su envolvente 42 incluyendo pre-énfasis. La Fig. 5b muestra, como un ejemplo del resultado del análisis LTP realizado por el estimador de distribución de probabilidad 14, un filtro de peine LTP 70, es decir una función de peine sobre los componentes espectrales k parametrizados, por ejemplo, mediante un valor de ganancia LTP que describe la frecuencia valle a pico a/b y un parámetro de retardo LTP que define el paso o distancia entre los picos de la función de peine 70, es decir, c . El estimador de distribución de probabilidad 14 puede determinar los parámetros LTP que acabamos de mencionar de forma que multiplicar la función de peine LTP 70 con la estimación basada en el coeficiente de predicción lineal 30 del espectro 26 estime de forma más aproximada el espectro real 26. La multiplicación de la función de peine LTP 70 con el modelo LPC 42 se muestra de forma ejemplar en la Fig. 5c y puede verse que el producto 72 de la función de peine LTP 70 y el modelo LPC 42 se aproxima más de cerca a la forma real del espectro 40.
- [0039]** En el caso de combinar la funcionalidad LTP del estimador de distribución de probabilidad 14 con el uso del dominio perceptual, el estimador de distribución de probabilidad 14 puede operar como se muestra en la Fig. 6. El modo de operación coincide ampliamente con el mostrado en la Fig. 4. Es decir, los coeficientes LPC que definen el filtro de síntesis de predicción lineal $1/A(z)$ se someten a la conversión LPC a ponderación espectral 60 y 62, es decir una vez directamente y la otra vez tras ser perceptualmente ponderados 64. Los factores de escala resultantes se someten a división 66 y los cocientes resultantes q_k se multiplican usando el multiplicador 47 con la función de peine LTP 70, la ganancia LTP de los parámetros y el desfase LTP del cual se determina mediante el estimador de distribución de probabilidad 14 de forma apropiada y se inserta en el flujo de datos 22 para su acceso por parte del lado de decodificación. El producto resultante $l_k \cdot q_k$ con l_k denotando la función de peine LTP en el componente espectral k , se somete a continuación a la derivación del parámetro de distribución de probabilidad 68 para obtener los parámetros de distribución de probabilidad π_k . Por favor, tenga en cuenta que, en la siguiente descripción del lado de decodificación, se hace referencia, inter alias, a la Fig. 6 con respecto a la funcionalidad del lado del decodificador de la estimación de la distribución de probabilidad. En este aspecto, por favor tenga en cuenta que, en el lado del decodificador, los parámetros LTP se determinan mediante la optimización y similar y se insertan en el flujo de datos 22, mientras que el lado de decodificación meramente tiene que leer los parámetros LTP desde el flujo de datos.
- [0040]** Tras haber descrito varias realizaciones para un codificador de audio basado en la predicción lineal con respecto a las Figs. 1 a 6, la siguiente descripción se concentra en el lado de decodificación. La Fig. 7 muestra una realización para un decodificador de audio basado en la predicción lineal 100. Comprende un estimador de distribución de probabilidad 102 y una etapa de decodificación y cuantificación de entropía 104. El decodificador de audio basado en predicción lineal tiene acceso al flujo de datos 22 y aunque el estimador de distribución de probabilidad 102 está configurado para determinar, para cada una de las pluralidades de componentes espectrales k , una estimación de distribución de probabilidad 28 a partir de la información del coeficiente de predicción lineal contenida en el flujo de datos 22, la etapa decodificación entropía y descuantificación 104 está configurada para decodificar la entropía y descuantificar el espectro 26 del flujo de datos 22 usando la estimación de distribución de probabilidad como se determina para cada una de las pluralidades de componentes espectrales k mediante el estimador de distribución de probabilidad 102. Es decir, tanto el estimador de distribución de probabilidad 102 como la etapa decodificación entropía y descuantificación 104 tienen acceso al flujo de datos 22 y el estimador de distribución de probabilidad 102 tiene su salida conectada a una entrada de la etapa decodificación entropía y descuantificación 104. En la salida de la última, se obtiene el espectro 26.
- [0041]** Debería tenerse en cuenta que, naturalmente, el resultado del espectro por la etapa decodificación entropía y descuantificación 104 puede someterse a un posterior procesamiento dependiendo de la aplicación. La salida del decodificador 100 no necesita ser necesariamente, sin embargo, la señal de audio que se codifica en el flujo de datos 22, en el dominio temporal para, por ejemplo, reproducirse usando altavoces. En vez de eso, el decodificador de audio basado en predicción lineal 100 puede interconectarse con la salida de, por ejemplo, el mezclador de un sistema de conferencias, un decodificador multi-canal o multi-objeto o similar, y esta interconexión puede estar en el dominio espectral. Alternativamente, el espectro o alguna versión post-procesada del mismo puede estar sujeto a una conversión espectral a tiempo mediante una conversión de descomposición espectral como una transformación inversa usando un proceso de superposición/adición como se describe más adelante.
- [0042]** Como el estimador de distribución de probabilidad 102 tiene acceso a la misma información LPC que el estimador de distribución de probabilidad 14 en el lado de codificación, el estimador de distribución de probabilidad 102 opera igual que el estimador correspondiente en el lado de codificación, excepto por, por ejemplo, la determinación del parámetro LTP adicional en el lado de codificación, el resultado de dicha determinación se

señaliza al lado de decodificación mediante el flujo de datos 22. La etapa decodificación entropía y descuantificación 104 se configura para usar la estimación de distribución de probabilidad en la decodificación de entropía de los valores espectrales del espectro 62, como los niveles de magnitud desde el flujo de datos 22, y la descuantifica igualmente para todos los componentes espectrales para así obtener el espectro 26. En cuanto a las diferentes
5 posibilidades para implementar la codificación de entropía, se hace referencia a las declaraciones anteriores que sobre la codificación de entropía. Además, la misma regla de cuantificación se aplica en una dirección inversa relativa a la usada en el lado de codificación de forma que todas las alternativas y detalles descritos anteriormente con respecto a la codificación y cuantificación de entropía se aplicarán también para las realizaciones del decodificador de forma correspondiente. Es decir, por ejemplo, la etapa decodificación entropía y descuantificación
10 puede configurarse para usar un tamaño de paso de cuantificación constante para descuantificar los niveles de magnitud y puede usar, por ejemplo, decodificación aritmética.

[0043] Como se ha denotado anteriormente, el espectro 26 puede representar una excitación de codificación de transformación y por tanto la Fig.8 muestra que el decodificador de audio basado en predicción lineal puede
15 comprender adicionalmente un filtro 106 que también tiene acceso a la información LPC y al flujo de datos 22, y está conectado a la salida de la etapa decodificación entropía y descuantificación 104 para recibir el espectro 26 y la salida del espectro de una señal de audio post-filtrada/reconstruida en su salida. En particular, el filtro 106 se configura para dar forma al espectro 26 de acuerdo con una función de transferencia dependiendo de un filtro de síntesis de predicción lineal definido por la información del coeficiente de predicción lineal. Para ser aún más
20 precisos, el filtro 106 puede implementarse mediante la concatenación del determinador de factores de escala 34 y el conformador espectral 38, con el conformador espectral 38 recibiendo el espectro 26 desde la etapa 104 y produciendo la señal post-filtrada, es decir la señal de audio reconstruida. La única diferencia sería que la escalación realizada dentro del filtro 106 sería exactamente la inversa de la escalación realizada por el conformador espectral 38 en el lado de codificación, es decir, donde el conformador espectral 38 en el lado de codificación realiza, por
25 ejemplo, una multiplicación usando los factores de escala, y en el filtro 106 se realizará una división por los factores de escala o viceversa.

[0044] La última circunstancia se muestra en la Fig. 9, que muestra una realización para el filtro 106 de la Fig. 8. Como puede verse, el filtro 108 puede comprender un determinador de factores de escala 110 que opera, por
30 ejemplo, como el determinador de factores de escala 34 de la Fig.2, y un conformador espectral 112 que, como se ha indicado anteriormente, aplica los factores de escala para el determinador del factor de escala 110 a la entrada del espectro, inversamente relativo al conformador espectral 38.

[0045] La Fig. 9 ilustra que el filtro 106 puede además comprender de forma ejemplar un transformador
35 inverso 114, un añadidor de superposición 116 y un filtro de de-énfasis 118. Estos últimos componentes 114 y 118 podrían conectarse secuencialmente a la salida del conformador espectral 112 en el orden en que se mencionan, donde el filtro de de-énfasis 118 o la superposición/añadidor 116 y el filtro de de-énfasis 118 podrían, de acuerdo con una alternativa adicional, dejarse fuera.

[0046] El filtro de de-énfasis 118 realiza la inversa del filtrado pre-énfasis del filtro 24 en la Fig. 1 y la
40 superposición/añadidor 116 puede, como se conoce en la técnica, resultar en una cancelación del solapamiento en el caso de que la transformación inversa usada dentro del transformador inverso 114 sea una transformación traslapada muestreada críticamente. Por ejemplo, el transformador inverso 114 podría someter a cada espectro 26 recibido desde el conformador espectral 112 a una frecuencia temporal en la cual dichos espectros se codifican
45 dentro del flujo de datos 22 a una transformación inversa para obtener partes con sistema de ventanas que, a su vez, se superponen-añaden por la superposición/añadidor 115 para obtener una versión de señal tiempo-dominio. El filtro de de-énfasis 118, al igual que el filtro pre-énfasis 24, puede implementarse como un filtro FIR.

[0047] Tras haber descrito las realizaciones de la presente solicitud con respecto a las figuras, a continuación
50 se proporciona una descripción más matemática de las realizaciones de la presente solicitud con esta descripción finalizando en la descripción correspondiente de las Figs. 10 y 11. En particular, en las realizaciones descritas a continuación se asume que la binarización unaria de los valores espectrales del espectro con codificación aritmética binaria de las celdas de las secuencias de celdas resultantes se utiliza para codificar el espectro.

[0048] En particular, en los detalles ejemplares descritos a continuación, que deben comprenderse como
55 transferibles a las realizaciones descritas anteriormente, se ha decidido ejemplarmente calcular la estructura del envolvente 30 en 64 bandas cuando la longitud de la trama, es decir, la frecuencia de espectro en la cual el espectro 26 se actualiza dentro del flujo de datos 22, es de 256 muestras y 80 bandas cuando la longitud de la trama es de 320 muestras. Si el modelo LPC es $A(z)$, entonces el LPC ponderado es, por ejemplo, $A(\gamma z)$ con $\gamma = 0.92$ y el término

de filtro de pre-énfasis 24 asociado es $(1 - 0.68z^{-1})$, por ejemplo, donde las constantes pueden variar en base a la aplicación. El envolvente 30 y el dominio perceptual es así

$$\frac{A(0.92z)(1 - 0.68z^{-1})}{A(z)}, \quad (1)$$

5

[0049] De esta forma, la función de transferencia del filtro definido por la fórmula (1) se corresponde con la función 48 en la Fig. 3b y es el resultado de la computación en las Figs. 4 y 6 en el resultado del divisor 66.

[0050] Debería notarse que las Figs. 4 y 6 representan el modo de funcionamiento del estimador de distribución de probabilidad 14 y el estimador de distribución de probabilidad 102 de la Fig. 7. Además, en el caso de que se utilicen el filtro pre-énfasis 24 y el filtro de-énfasis 118, la conversión LPC a ponderación espectral 60 tiene la función de filtro pre-énfasis en cuenta de forma que, al final, represente el producto de las funciones de transferencia del filtro de síntesis y el filtro pre-énfasis.

[0051] En cualquier caso, la transformación tiempo-frecuencia del filtro definido por la fórmula (1) debería calcularse de forma que el envolvente final sea la frecuencia alineada con la representación espectral de la señal de entrada. Además, debería observarse de nuevo que el estimador de distribución de probabilidad puede computar meramente la magnitud absoluta del envolvente o la función de transferencia del filtro de la fórmula (1). En ese caso, el componente de fase puede descartarse.

20

[0052] En el caso de calcular el envolvente para las bandas espectrales y no para las líneas individuales, el envolvente aplicado a las líneas espectrales tendrá un planteamiento por etapas continuo. Para obtener un envolvente aún más continuo, es posible interpolar o suavizar el envolvente. Sin embargo, debería observarse que las bandas espectrales continuas en sus etapas proporcionan una reducción en la complejidad computacional. Por tanto, esto es un equilibrio entre la precisión y la complejidad.

25

[0053] Como se ha indicado anteriormente, el LTP también puede usarse para inferir un envolvente más detallado. Algunos de los principales desafíos de aplicar la información armónica a la forma del envolvente son:

30 1) Seleccionar la codificación y precisión de la información LTP como el desfase LTP y la ganancia LTP. Por ejemplo, podría usarse la misma codificación que en ACELP.

2) El LTP puede corresponderse a un filtro de peine en el dominio de frecuencia. Sin embargo, las realizaciones anteriores o cualquier otra realización de acuerdo con la presente invención no estarán restringida al uso de un filtro de peine de la misma forma que el LTP. También podrían usarse otras funciones.

35 3) Además de la forma de filtro de peine del LTP, también es posible decidir aplicar el LTP de forma diferente en diferentes regiones de frecuencia. Por ejemplo, los picos armónicos son normalmente más prominentes en frecuencias más bajas. Tendría entonces sentido aplicar el modelo armónico en la frecuencia baja con una amplitud más alta que en las frecuencias altas.

4) Como se ha indicado anteriormente, la forma del envolvente se calcula a nivel de bandas. Sin embargo, un filtro de peine en LTP ciertamente tendrá una estructura y frecuencia mucho más detalladas que las que tienen los valores de envolvente estimados a nivel de banda. En la implementación del modelo armónico, es entonces beneficioso reducir la complejidad computacional.

[0054] En las realizaciones anteriores, puede utilizarse el supuesto según el cual las líneas individuales o más específicamente las magnitudes del espectro 26 en los componentes espectrales k , se distribuyen de acuerdo con la distribución Laplace, es decir, la distribución exponencial señalizada. En otras palabras, la anteriormente mencionada $f(i,l(k))$ puede ser una función Laplace. Ya que la señal del espectro 26 en el componente espectral k puede siempre codificarse por un bit, y la probabilidad de ambas señales puede asumirse de forma segura que será 0,5, la señal siempre puede codificarse por separado, y debemos considerar solo la distribución exponencial.

50

[0055] En general, sin ninguna información anterior, la primera opción para cualquier distribución sería la distribución normal. La distribución exponencial, sin embargo, tiene mucha más probabilidad de tener una masa cercana a cero que la distribución normal y por tanto describe una señal más dispersa que la distribución normal. Ya que uno de los principales objetivos de las transformaciones tiempo-frecuencia es conseguir una señal dispersa, entonces una distribución de probabilidad que describa señales dispersas está bien garantizada. Además, la distribución exponencial también proporciona ecuaciones que son fácilmente tratables en formato analítico. Estos

55

dos segmentos proporcionan la base para usar la distribución exponencial. Las siguientes derivaciones pueden naturalmente modificarse fácilmente para otras distribuciones.

[0056] Una variable distribuida exponencialmente x tiene la función de densidad de probabilidad ($x \geq 0$):

5

$$f(x; \lambda) = \lambda e^{-\lambda x} \quad (2)$$

y la función de distribución acumulativa

10

$$F(x; \lambda) = 1 - e^{-\lambda x}. \quad (3)$$

[0057] La entropía de una variable exponencial es $1 - \ln(\lambda)$, donde el consumo de bits esperado de una línea única, incluyendo signo, sería $\log_2(2e\lambda)$. Sin embargo, éste es un valor teórico que se conserva para variables discretas solo λ cuando es grande.

15

[0058] El consumo real de bits puede estimarse mediante simulaciones, pero no hay disponible una fórmula analítica precisa. Un consumo de bits aproximado es, por tanto, $\log_2(2e\lambda + 0,15 + 0,035/\lambda)$ para $\lambda > 0,08$.

[0059] Es decir, las realizaciones descritas anteriormente con el estimador de distribución de probabilidad en los lados de codificación y decodificación pueden usar una distribución Laplace como una función parametrizable para determinar la estimación de distribución de probabilidad. El parámetro de escala λ de la distribución Laplace puede servir como el parámetro de distribución de probabilidad mencionado anteriormente, es decir, como π_k .

[0060] A continuación, se describe una posibilidad para realizar la escalación de envolvente. Un enfoque se basa en realizar una primera suposición para la escalación, calculando su consumo de bits y mejorando la escalación iterativamente hasta que esté lo suficientemente cerca del nivel deseado. En otras palabras, los estimadores de distribución de probabilidad mencionados anteriormente en el lado de codificación y decodificación podrían realizar los siguientes pasos.

[0061] Digamos que f_k es el valor de envolvente para la posición k . El valor promedio del envolvente es

entonces $\hat{f} = \frac{1}{N} \sum_k f_k$ donde N es el número de líneas espectrales. Si el consumo de bits deseado es b , entonces el primer intento de suposición de escalación puede resolverse fácilmente a partir de

$$\frac{b}{N} = \log_2 \left(2e\hat{f}g_0 + 0,15 - 0,035(\hat{f}g_0)^{-1} \right),$$

[0062] El consumo estimado de bits b_k para la iteración k y con la escalación g_k es entonces

35

$$b_k = \sum_h \log_2(2ef_h g_k + 0,15 + 0,035(f_h g_k)^{-1}) \quad (4)$$

[0063] La operación de logaritmo es computacionalmente compleja, así que, en vez de eso, podemos calcular

40

$$b_k = \log_2 \prod_h (2ef_h g_k + 0,15 + 0,035(f_h g_k)^{-1}) \quad (5)$$

[0064] Incluso aunque el término del producto es un número muy grande y su cálculo en punto fijo requiere mucha administración, sigue siendo menos complejo que un número grande de operaciones $\log_2()$.

[0065] Para reducir aún más la complejidad, podemos estimar el consumo de bits mediante $\log_2(2e \cdot \lambda)$, donde el consumo total de bits es $b = \log_2 \prod 2e \cdot f_k \cdot g$. A partir de esta ecuación, el coeficiente de escalación g puede resolverse fácilmente de forma analítica, donde la iteración de escalación del envolvente no es necesaria.

[0066] En general, no existe una forma analítica de resolver g_k a partir de la Ec. 5, donde debe usarse un método iterativo. Si se utiliza la búsqueda de bisección, entonces si $b_0 < b$, entonces el tamaño del paso inicial es $2^{(b-b_0)/N} - 1$ y de lo contrario el tamaño del paso es $1 - 2^{(b-b_0)/N}$. Mediante este enfoque, la búsqueda de bisección converge típicamente en 5-6 iteraciones.

[0067] El envolvente debe escalarse igualmente en el codificador y en el decodificador. Ya que las distribuciones de probabilidad se derivan del envolvente, incluso una diferencia de 1 bit en la escalación en el codificador y el decodificador causaría que el decodificador aritmético produzca un resultado aleatorio. Es por tanto muy importante que la implementación opere exactamente igual en todas las plataformas. En la práctica, esto requiere que el algoritmo se implemente con operaciones de entero y punto fijo.

[0068] Aunque el envolvente ya ha sido escalado de forma que la expectativa del consumo de bits sea igual al nivel deseado, las líneas espectrales reales en general no coincidirán con la asignación de bits sin la escalación. Incluso si se escalara la señal de forma que su variación coincidiera con la variación del envolvente, la distribución de muestra invariablemente diferiría de la distribución del modelo, donde el consumo deseado de bits no se alcanza. Es por tanto necesario escalar la señal de forma que cuando se cuantifique y codifique, el consumo final de bits alcance el nivel deseado. Ya que esto normalmente debe realizarse de forma iterativa (no existe una solución analítica), el proceso se conoce como bucle de régimen.

[0069] Hemos elegido comenzar con una escalación de primera suposición de forma que la variación del envolvente y la señal escalada coincidan. Simultáneamente, podemos encontrar esa línea espectral, que tiene la menor probabilidad de acuerdo con nuestro modelo de probabilidad. Debe tenerse cuidado de que el valor de probabilidad más pequeño no esté por debajo de la precisión de máquina. Esto así establece un límite en el factor de escala que se estimará en el bucle de régimen.

[0070] Para el bucle de régimen, de nuevo empleamos la búsqueda bisección, de forma que el tamaño del paso comience a la mitad del factor de escala inicial. Luego el consumo de bits se calcula en cada iteración como una suma de todas las líneas espectrales y la precisión de cuantificación se actualiza dependiendo de cómo de cerca estemos a la asignación de bits.

[0071] En cada iteración, la señal se cuantifica primero con la escalación actual. Segundo, cada línea se codifica con el codificador aritmético. De acuerdo con el modelo de probabilidad, la probabilidad de que una línea x_k se cuantifique a cero es de $p(x_k = 0) = 1 - \exp(-.5/f_k)$, donde f_k es el valor del envolvente (= desviación estándar de la línea espectral). El consumo de bits de dicha línea es naturalmente $-\log_2 p(x_k = 0)$. Un valor x_k no cero tiene la probabilidad $p(|x_k| = q) = \exp((q + .5)/f_k) - \exp((q - .5)/f_k)$. La magnitud puede entonces codificarse con $\log_2(p(|x_k| = q))$ bits, más un bit por el signo.

[0072] De esta forma, puede calcularse el consumo de bits de todo el espectro. Además, debe tenerse en cuenta que podemos establecer un límite K de forma que todas las líneas $k > K$ sean cero. Es entonces suficiente codificar las primeras líneas K . El decodificador puede entonces deducir que, si las primeras líneas K han sido decodificadas, pero no hay bits adicionales disponibles, entonces las líneas restantes deben ser todas cero. Por tanto, no es necesario transmitir el límite K , sino que puede deducirse de la secuencia de bits. De esta forma, podemos evitar las líneas de codificación que son cero, donde ahorramos bits. Ya que para las señales de voz y audio sucede frecuentemente que la parte superior del espectro se cuantifica a cero, es beneficioso comenzar desde las frecuencias bajas, y hasta donde sea posible, usar todos los bits para las líneas K .

[0073] Debe tenerse en cuenta de que ya que los valores del envolvente f_k son iguales dentro de una banda, podemos reducir fácilmente la complejidad precalculando los valores que son necesarios para cada línea en una banda. Específicamente, en las líneas de codificación, el término $\exp(.5/f_k)$ siempre es necesario y es igual dentro de cada banda. Además, este valor no cambia dentro del bucle de régimen, donde puede calcularse fuera del bucle de régimen y el mismo valor puede usarse también para la cuantificación final.

[0074] Además, ya que el consumo de bits de una línea es $\log_2()$ de la probabilidad, podemos, en vez de calcular la suma de los logaritmos, calcular el logaritmo de un producto. De esta forma se ahorra en complejidad de nuevo. Además, ya que el bucle de régimen es una función solo del codificador, las operaciones de punto flotante nativas pueden usarse en vez de punto fijo.

5

[0075] Haciendo referencia a lo anterior, se consulta la Fig. 10, que muestra una subparte del codificador explicada anteriormente con respecto a las figuras, cuya parte es responsable de realizar la escalación del envolvente mencionada anteriormente y el bucle de régimen de acuerdo con una realización. En particular, la Fig. 10 muestra elementos de la etapa de codificación y cuantificación de entropía 18 por un lado y el estimador de distribución de probabilidad 14 por el otro. Un binarizador de binarización unaria 130 somete a las magnitudes de los valores espectrales x_k del espectro 26 a los componentes espectrales k para una binarización unaria, por tanto, generando, para cada magnitud en el componente espectral k , una secuencia de celdas. El codificador aritmético binario 132 recibe estas secuencias de celdas, es decir, una por componente espectral k , y las somete a codificación aritmética binaria. Tanto el binarizador de binarización unaria 130 y el codificador aritmético binario 132 son parte de la etapa de codificación y cuantificación de entropía 18. La Fig. 10 muestra el derivador del parámetro 68, que es responsable de realizar la escalación mencionada anteriormente para escalar los valores de estimación del envolvente q_k , o como también se denotaron anteriormente por f_k , de forma que resulte en parámetros de distribución de probabilidad escalada correctamente π_k o usando notación recién usada $g_k f_k$. Como se ha descrito anteriormente usando la fórmula (5), el derivador binario 68 determina el valor de escalación g_k de forma iterativa, de forma que la estimación analítica del consumo de bits sea un ejemplo del cual se representa mediante la ecuación (5), alcance alguna tasa de bits objetivo para todo el espectro 26. Como nota adicional, deberá notarse que k como se utiliza en conexión con la ecuación (5) denotaba el número de paso de la iteración, mientras que en otros sitios la variable k debía denotar la línea espectral o el componente k . Más allá de eso, debería notarse que el derivador de parámetros 68 no escala necesariamente los valores de envolvente originales derivados ejemplarmente como se muestra en las Figs. 4 y 6, sino que podría modificar directa e iterativamente los valores del envolvente usando, por ejemplo, modificadores aditivos.

[0076] En cualquier caso, el codificador aritmético binario 132 aplica, para cada componente espectral, la estimación de distribución de probabilidad como se define mediante el parámetro de distribución de probabilidad π_k , o como se ha usado de forma alternativa anteriormente, $g_k f_k$, para todas las celdas de la binarización unaria de la magnitud respectiva de los valores espectrales x_k .

[0077] Además, como se ha descrito anteriormente, un verificador de bucle de régimen 134 puede proporcionarse para verificar el consumo real de bits producido usando los parámetros de distribución de probabilidad como determina el derivador de parámetros 68 como primera suposición. El verificador de bucle de régimen 134 verifica la suposición al estar conectado entre el codificador aritmético binario 132 y el derivador de parámetros 68.

[0078] Si el consumo real de bits supera el consumo de bits permitido a pesar de la estimación realizada por el derivador de parámetros 68, el verificador de bucle de régimen 134 corrige los primeros valores supuestos de los parámetros de distribución de parámetro π_k (o $g_k f_k$), y la codificación aritmética binaria 132 real de las binarizaciones unarias se realiza de nuevo.

[0079] La Fig. 11 muestra con fines de exhaustividad una parte similar del decodificador de la Fig. 8. En particular, el derivador de parámetros 68 opera en el lado de codificación y decodificación de la misma forma y por tanto se muestra igual en la Fig. 11. En vez de usar una concatenación del binarizador de binarización unaria seguido del codificador aritmético, el lado de decodificación se utiliza la disposición secuencial inversa, es decir, la etapa de decodificación y descuantificación de entropía 104 de acuerdo con la Fig. 11 ejemplarmente comprende un decodificador aritmético binario 136 seguido de un desbinarizador de dispositivo de binarización unaria 138. El decodificador aritmético binario 136 recibe la parte del flujo de datos 22 que codifica aritméticamente el espectro 26. El resultado del decodificador aritmético binario 136 es una secuencia de secuencias de celdas, es decir, una secuencia de celdas de una cierta magnitud de valor espectral en el componente espectral k seguida por la secuencia de bits de la magnitud del valor espectral del siguiente componente espectral $k + 1$ y así sucesivamente. El desbinarizador de binarización unaria 138 realiza la desbinarización, es decir, produce las magnitudes desbinarizadas de los valores espectrales en el componente espectral k e informa al decodificador aritmético binario 136 al principio y al final de las secuencias de celda de las magnitudes individuales de los valores espectrales. Al igual que el codificador aritmético binario 132, el decodificador aritmético binario 136 utiliza, por decodificación aritmética binaria, las estimaciones de distribución de parámetros definidas por los parámetros de distribución de

parámetro, es decir, el parámetro de distribución de probabilidad π_k (gkfk), para todas las celdas que pertenezcan a una magnitud respectiva de un valor espectral del componente espectral k.

5 **[0080]** Como también se ha descrito anteriormente, el codificador y decodificador pueden explorar el hecho de que ambos lados pueden estar informados sobre la tasa de bits máxima disponible para que ambos lados puedan explorar la circunstancia de que la codificación real de las magnitudes de los valores espectrales del espectro 26 puede frustrarse cuando cruzan desde la frecuencia más baja a la frecuencia más alta, tan pronto como se alcance la tasa de bits máxima disponible en la secuencia de bits 22. Por convención, la magnitud no transmitida puede establecerse a cero.

10 **[0081]** Con respecto a las realizaciones descritas más recientemente, se tiene en cuenta que, por ejemplo, la primera escalación supuesta del envolvente para obtener los parámetros de distribución de probabilidad pueden usarse sin el bucle de régimen para obedecer la tasa bits algo constante como, por ejemplo, si el escenario de la aplicación no requiere conformidad.

15 **[0082]** Aunque algunos aspectos se han descrito en el contexto de un aparato, está claro que estos aspectos también representan una descripción del procedimiento correspondiente, donde un bloque o dispositivo se corresponde con un paso del procedimiento o una característica de un paso del procedimiento. Semejantemente, los aspectos descritos en el contexto de un paso del procedimiento representan una descripción de un bloque correspondiente o elemento o característica de un aparato correspondiente. Algunos o todos los pasos del procedimiento pueden ejecutarse por (o usando) un aparato de hardware, como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, alguno o más de los pasos del procedimiento más importantes pueden ejecutarse por dicho aparato.

25 **[0083]** La señal de audio codificada de la invención puede almacenarse en un medio de almacenamiento digital o pueden transmitirse sobre un medio de transmisión, como un medio de transmisión inalámbrica o un medio de transmisión por cable como Internet.

30 **[0084]** Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo, un disco floppy, un DVD, un Blue-Ray, un CD, un ROM, un PROM, un EPROM, un EEPROM o una memoria flash, con señales de control legibles electrónicamente almacenadas en el mismo, que cooperan (o son capaces de cooperar) con un sistema informático programable de forma que se realice el procedimiento respectivo. Por tanto, el medio de almacenamiento digital puede ser legible por ordenador.

35 **[0085]** Algunas realizaciones de acuerdo con 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 forma que uno de los procedimientos descritos en el presente se lleva a cabo.

40 **[0086]** Generalmente, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, el código de programa siendo operativo para realizar uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código del programa puede, por ejemplo, almacenarse en un soporte legible por máquina.

45 **[0087]** Otras realizaciones comprenden el programa informático para realizar uno de los procedimientos descritos en el presente, almacenado en un soporte legible por máquina.

50 **[0088]** En otras palabras, una realización del procedimiento de la invención es, por tanto, un programa informático que tiene un código de programa para realizar uno de los procedimientos descritos en el presente, cuando el programa informático se ejecuta en un ordenador.

[0089] Una realización adicional del procedimiento de la invención es, por tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para realizar uno de los procedimientos descritos en el mismo. El soporte de datos, el medio de almacenamiento digital o el medio grabado son normalmente tangibles y/o no transitorios.

[0090] Una realización adicional del procedimiento de la invención es, por tanto, un flujo de datos o una secuencia de señales que representa el programa informático para realizar uno de los procedimientos descritos en el

presente. El flujo de datos de la secuencia de las señales puede, por ejemplo, estar configurado para transferirse mediante una conexión de comunicación de datos, por ejemplo, por Internet.

[0091] Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado o adaptado para realizar uno de los procedimientos descritos en el presente.

[0092] Una realización adicional comprende un ordenador que tiene instalado el programa informático para realizar uno de los procedimientos descritos en el presente.

[0093] Una realización adicional de acuerdo con la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, de forma electrónica u óptica) un programa informático para realizar uno de los procedimientos descritos en el presente a un receptor. El receptor puede, por ejemplo, ser un ordenador, un dispositivo móvil, un dispositivo de memoria o similar. El aparato o sistema puede, por ejemplo, comprender un servidor de archivos para transferir el programa informático al receptor.

[0094] En algunas realizaciones, un dispositivo lógico programable (por ejemplo, una Matriz de puertas programables por campo o FPGA) puede utilizarse para realizar algunas o todas las funcionalidades de los procedimientos descritos en el presente. En algunas realizaciones, la FPGA puede cooperar con un microprocesador para realizar uno de los procedimientos descritos en el presente. Generalmente, los procedimientos se realizan preferiblemente mediante un aparato de hardware.

[0095] Las realizaciones descritas anteriormente son meramente ilustrativas para los principios de la presente invención. Se comprende que las modificaciones y variaciones de las disposiciones y los detalles descritos en el presente serán aparentes para aquellos expertos en la técnica. Es la intención, por tanto, estar limitados solo por el alcance de las reivindicaciones de patente inminentes y no por los detalles específicos presentados mediante la descripción y explicación de las realizaciones del presente.

Referencias

30

[0096]

[1] ISO/IEC 23003-3:2012, "MPEG-D (MPEG audio technologies), Part 3: Unified speech and audio coding," 2012.

[2] J. Makhoul, "Linear prediction: A tutorial review," Proc. IEEE, vol. 63, no. 4, pp. 561-580, April 1975.

35 [3] G. Fuchs, V. Subbaraman, and M. Multrus, "Efficient context adaptive entropy coding for real-time applications," in Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on, May 2011, pp. 493-496.

[4] US8296134 and WO2012046685.

REIVINDICACIONES

1. Decodificador de audio basado en predicción lineal, que comprende:
- 5 un estimador de distribución de probabilidad (102) configurado para determinar, para cada pluralidad de componentes espectrales, una estimación de distribución de probabilidad (28) a partir de la información del coeficiente de predicción lineal contenido en un flujo de datos (22) en el cual la señal de audio se codifica;
- una etapa de decodificación y descuantificación de entropía (104) configurada para decodificar y descuantificar la
- 10 entropía de un espectro (26) compuesta por una pluralidad de componentes espectrales desde el flujo de datos (22) usando la estimación de distribución de probabilidad (28) como determina cada una de las pluralidades de componentes espectrales; y un filtro configurado para dar forma al espectro (26) de acuerdo con una función de transferencia dependiendo de un filtro de síntesis de predicción lineal definido por la información del coeficiente de predicción lineal, **caracterizado porque** el estimador de distribución de probabilidad (102) está configurado para
- 15 determinar una estructura fina espectral (70) a partir de los parámetros de predicción a largo plazo contenidos en el flujo de datos y para determinar, para cada una de las pluralidades de los componentes espectrales, un parámetro de distribución de probabilidad de forma que los parámetros de distribución de probabilidad sigan espectralmente una función (72) que depende multiplicativamente de la estructura fina espectral (70), donde, para cada una de las pluralidades de componentes espectrales, la estimación de distribución de probabilidad (28) es una función
- 20 parametrizable parametrizada con el parámetro de distribución de probabilidad del componente espectral respectivo.
2. El decodificador de audio basado en predicción lineal de acuerdo con la reivindicación 1, que comprende, además:
- 25 un determinador de factores de escala (110) configurado para determinar los factores de escala basado en la información del coeficiente de predicción lineal; y un conformador espectral (112) configurado para formar espectralmente el espectro escalando el espectro usando los factores de escala, donde el determinador del factor de escala se configura para determinar los factores de escala de forma que representen una función de transferencia dependiendo del filtro de síntesis de predicción lineal definido por la información del coeficiente de predicción lineal.
- 30 3. El decodificador de audio basado en predicción lineal de acuerdo con las reivindicaciones 1 o 2, donde la dependencia de la función de transferencia en el filtro de síntesis de predicción lineal definido por la información del coeficiente de predicción lineal es tal que la función de transferencia se pondera perceptualmente.
- 35 4. El decodificador de audio basado en predicción lineal de acuerdo con las reivindicaciones 1 o 3, donde la dependencia de la función de transferencia en el filtro de síntesis de predicción lineal $1/A(z)$ definido por la predicción lineal es tal que la función de transferencia es una función de transferencia de $1/A(k \cdot z)$, donde k es una constante.
- 40 5. El decodificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones anteriores, donde el estimador de distribución de probabilidad está configurado de forma que la estructura fina espectral es una estructura tipo peine definida por los parámetros de predicción a largo plazo.
6. El decodificador de audio basado en predicción lineal de acuerdo con cualquiera de las
- 45 reivindicaciones anteriores, donde los parámetros de predicción a largo plazo comprenden la ganancia de predicción a largo plazo y un paso de predicción a largo plazo.
7. El decodificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones 1 a 6, donde, por cada una de las pluralidades de componentes espectrales, la función
- 50 parametrizable se define de forma que el parámetro de distribución de probabilidad es una medida para una dispersión de la estimación de distribución de probabilidad.
8. El decodificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones 1 a 7, donde, por cada una de las pluralidades de componentes espectrales, la función
- 55 parametrizable es una distribución Laplace, y el parámetro de distribución de probabilidad del componente espectral respectivo forma un parámetro de escala de la distribución Laplace respectiva.
9. El decodificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones 1 a 8 además comprende un filtro de de-énfasis.

10. El decodificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones anteriores, donde la etapa de decodificación y descuantificación de entropía (104) está configurada para, decodificar y descuantificar la entropía del espectro de la pluralidad de componentes espectrales, tratar el signo y la pluralidad en los componentes espectrales por separado usando la estimación de distribución de probabilidad como se determina para cada una de las pluralidades de componentes espectrales para la magnitud.
11. El decodificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones, donde la etapa de decodificación y descuantificación de entropía (104) se configura para usar la estimación de distribución de probabilidad en la decodificación de entropía a un nivel de magnitud del espectro por componente espectral y descuantificar los niveles de magnitud para todos los componentes espectrales de forma que se obtenga el espectro.
12. El decodificador de audio basado en predicción lineal de acuerdo con la reivindicación 11, donde la etapa de decodificación y cuantificación de entropía (104) se configura para usar un tamaño de paso de cuantificación constante para descuantificar los niveles de magnitud.
13. El decodificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones anteriores, que además comprende un transformador inverso configurado para someter al espectro a una transformación inversa muestreada críticamente con valor real para obtener una parte de señal con solapamiento tiempo-dominio; y un añadidor de superposición configurado para someter a la parte de señal con solapamiento tiempo-dominio a un proceso superponer y añadir con una parte tiempo-dominio precedente y/o subsiguiente para reconstruir la señal de audio.
14. Codificador de audio basado en predicción lineal que comprende:
- un analizador de predicción lineal (12) configurado para determinar la información del coeficiente de predicción lineal;
 - un estimador de distribución de probabilidad (14) configurado para determinar, para cada una de las pluralidades de componentes espectrales, una estimación de distribución de probabilidad a partir de la información del coeficiente de predicción lineal; y un determinador de espectro (16) configurado para determinar un espectro compuesto por la pluralidad de componentes espectrales desde una señal de audio;
 - una etapa de cuantificación y codificación de entropía (18) configura para cuantificar y codificar la entropía del espectro usando la estimación de distribución de probabilidad como se determina para cada una de las pluralidades de componentes espectrales, donde el determinador de espectro (16) está configurado para dar forma a un espectro original de la señal de audio de acuerdo con una función de transferencia que depende de una inversa de un filtro de síntesis de predicción lineal definido por la información del coeficiente de predicción lineal, y **caracterizado porque** el codificador de audio basado en predicción lineal además comprende un predictor a largo plazo configurado para determinar los parámetros de predicción a largo plazo, y el estimador de distribución de probabilidad (14) está configurado para determinar una estructura fina espectral (70) a partir de los parámetros de predicción a largo plazo y determinar, para cada una de las pluralidades de componentes espectrales, un parámetro de distribución de probabilidad de forma que los parámetros de distribución de probabilidad sigan espectralmente una función (72) que depende de un producto de una función de transferencia del filtro de síntesis de predicción lineal, una inversa de una función de transferencia de una modificación ponderada perceptualmente del filtro de síntesis de predicción lineal, y la estructura espectral fina (70), donde, para cada una de las pluralidades de componentes espectrales, la estimación de la distribución de probabilidad (28) es una función parametrizable con el parámetro de distribución de probabilidad del componente espectral respectivo.
15. El codificador de audio basado en predicción lineal de acuerdo con la reivindicación 14, donde el determinador de espectro (16) comprende:
- un determinador de factores de escala (34) configurado para determinar los factores de escala basado en la información del coeficiente de predicción lineal;
 - un transformador (36) configurado para descomponer espectralmente la señal de audio para obtener el espectro original; y un conformador espectral (38) configurado para dar forma espectralmente al espectro original escalando el espectro usando los factores de escala, donde el determinador de factores de escala (34) se configura para

determinar los factores de escala de forma que la conformación espectral por el conformador espectral usando los factores de escala se corresponde con una función de transferencia que depende de una inversa del filtro de síntesis de predicción lineal definido por la información del coeficiente de predicción lineal.

- 5 16. El codificador de audio basado en predicción lineal de acuerdo con las reivindicaciones 14 o 15, donde la dependencia de la función de transferencia en la inversa del filtro de síntesis de predicción lineal definido por la predicción lineal es tal que la función de transferencia se pondera perceptualmente.
17. El codificador de audio basado en predicción lineal de acuerdo con las reivindicaciones 14 a 16, donde
10 la dependencia de la función de transferencia en la inversa del filtro de síntesis de predicción lineal $1/A(z)$ definido por la información del coeficiente de predicción lineal es tal que la función de transferencia es una inversa de una función de transferencia de $1/A(k-z)$, donde k es una constante.
18. El codificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones
15 14 a 17, donde el estimador de distribución de probabilidad está configurado de forma que la estructura fina espectral es una estructura tipo peine definida por los parámetros de predicción a largo plazo.
19. El codificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones
20 14 a 18, donde los parámetros de predicción a largo plazo comprenden una ganancia de predicción a largo plazo y un paso de predicción a largo plazo.
20. El codificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones
14 a 19, donde, por cada una de las pluralidades de componentes espectrales, la función parametrizable se define
25 de forma que el parámetro de distribución de probabilidad es una medida para una dispersión de la estimación de distribución de probabilidad.
21. El codificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones
14 a 20, donde, por cada una de las pluralidades de componentes espectrales, la función parametrizable es una
30 distribución Laplace, y el parámetro de distribución de probabilidad del componente espectral respectivo forma un parámetro de escala de la distribución Laplace respectiva.
22. El codificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones
14 a 21, que además comprende un filtro pre-énfasis (24) configurado para someter la señal de audio a un pre-
35 énfasis.
23. El codificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones
14 a 22, donde la etapa de codificación y cuantificación de entropía está configurada para, al cuantificar y codificar la
40 entropía del espectro de la pluralidad de componentes espectrales, tratar el signo y la magnitud en la pluralidad de componentes espectrales de forma independiente usando la estimación de distribución de probabilidad como se determina para cada una de las pluralidades de componentes espectrales para la magnitud.
24. El codificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones
14 a 23, donde la etapa de codificación y cuantificación de entropía (18) se configura para cuantificar el espectro
45 igualmente para todos los componentes espectrales para obtener los niveles de magnitud para los componentes espectrales y usar la estimación de distribución de probabilidad en la codificación de entropía de los niveles de magnitud del espectro por componente espectral.
25. El codificador de audio basado en predicción lineal de acuerdo con la reivindicación 24, donde la etapa
50 de codificación y cuantificación de entropía está configurada para usar un tamaño de paso de cuantificación constante para la cuantificación.
26. El codificador de audio basado en predicción lineal de acuerdo con cualquiera de las reivindicaciones
14 a 25, donde el transformador está configurado para realizar una transformación muestreada críticamente con
55 valor real.
27. Procedimiento de decodificación de audio basada en predicción lineal, que comprende:

determinar, para cada pluralidad de componentes espectrales, una estimación de distribución de probabilidad (28) a partir de la información del coeficiente de predicción lineal contenido en un flujo de datos (22) en el cual la señal de audio se codifica; y

- 5 decodificar y descuantificar la entropía de un espectro (26) compuesta por una pluralidad de componentes espectrales desde el flujo de datos (22) usando la estimación de distribución de probabilidad como determina cada una de las pluralidades de componentes espectrales, el procedimiento comprendiendo además dar forma al espectro (26) de acuerdo con una función de transferencia dependiendo de un filtro de síntesis de predicción lineal definido por la información del coeficiente de predicción lineal, **caracterizado porque** la determinación de la
- 10 estimación de distribución de probabilidad (28) comprende determinar una estructura fina espectral (70) a partir de los parámetros de predicción a largo plazo contenidos en el flujo de datos y determinar, para cada una de las pluralidades de los componentes espectrales, un parámetro de distribución de probabilidad de forma que los parámetros de distribución de probabilidad sigan espectralmente una función (72) que depende multiplicativamente de la estructura fina espectral (70), donde, para cada una de las pluralidades de componentes espectrales, la
- 15 estimación de distribución de probabilidad es una función parametrizable parametrizada con el parámetro de distribución de probabilidad del componente espectral respectivo.

28. Procedimiento de codificación de audio basada en predicción lineal, que comprende:

- 20 determinar la información del coeficiente de predicción lineal;

determinar, para cada una de las pluralidades de componentes espectrales, una estimación de distribución de probabilidad (28) a partir de la información del coeficiente de predicción lineal; determinar un espectro compuesto por la pluralidad de componentes espectrales desde una señal de audio;

- 25 cuantificar y codificar la entropía del espectro usando la estimación de distribución de probabilidad como se determina para cada una de las pluralidades de componentes espectrales, donde el determinador de espectro comprende dar forma a un espectro original de la señal de audio de acuerdo con una función de transferencia que depende de una inversa de un filtro de síntesis de predicción lineal definido por la información del coeficiente de
- 30 predicción lineal, y **caracterizado porque** el procedimiento además comprende determinar los parámetros de predicción a largo plazo, y la determinación de la estimación de distribución de probabilidad (28) comprende determinar una estructura espectral fina (70) a partir de los parámetros de predicción a largo plazo y determinar, para cada una de las pluralidades de componentes espectrales, un parámetro de distribución de probabilidad de forma que los parámetros de distribución de probabilidad sigan espectralmente una función (72) que depende de un
- 35 producto de una función de transferencia del filtro de síntesis de predicción lineal, una inversa de una función de transferencia de una modificación ponderada perceptualmente del filtro de síntesis de predicción lineal, y la estructura espectral fina (70), donde, para cada una de las pluralidades de los componentes espectrales, la estimación de la distribución de probabilidad (28) es una función parametrizable parametrizada con el parámetro de distribución de probabilidad del componente espectral respectivo.

- 40
29. Programa informático que tiene un código de programa para realizar, cuando se ejecuta en un ordenador, un procedimiento de acuerdo con la reivindicación 27 o 28.

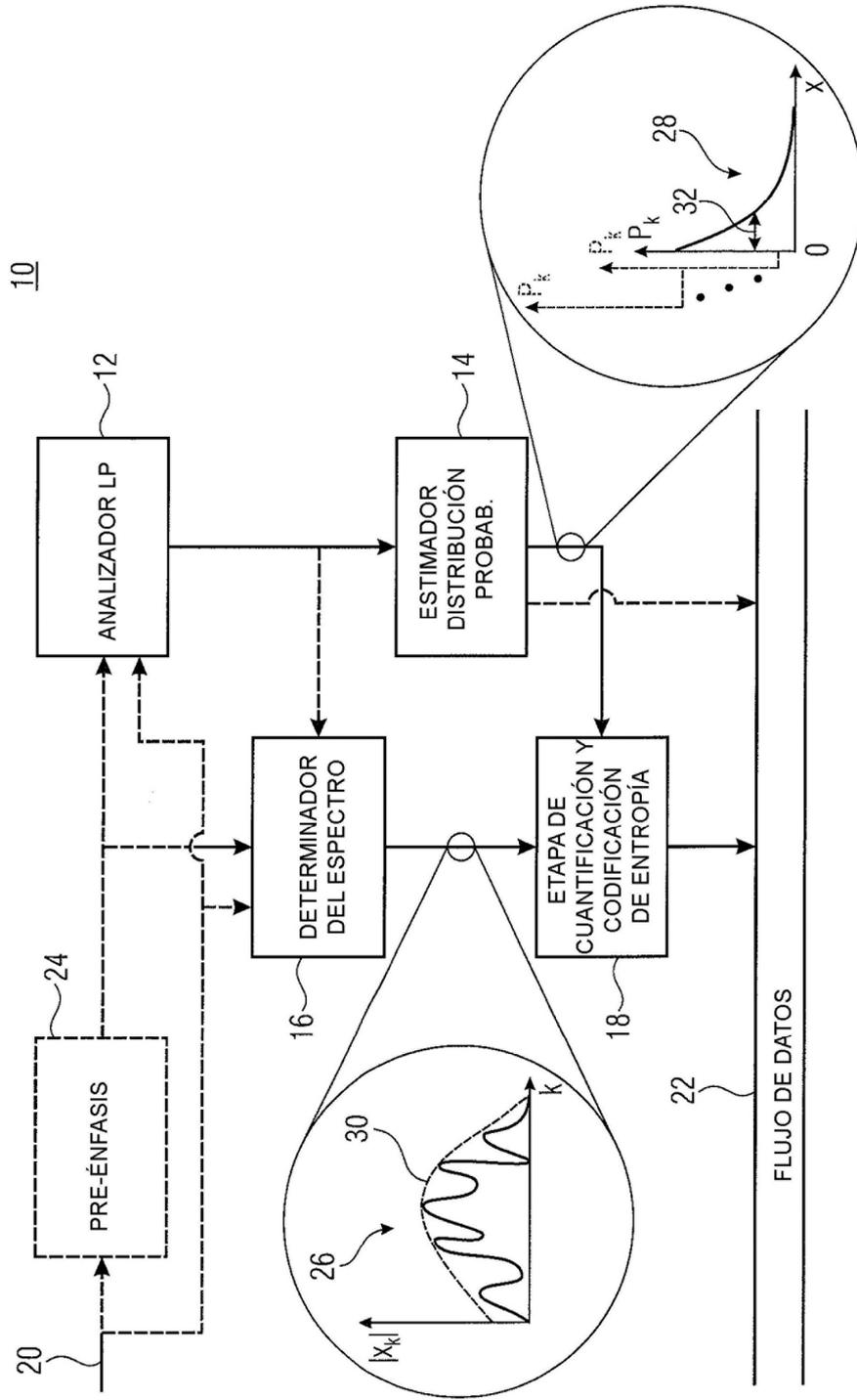


FIG 1

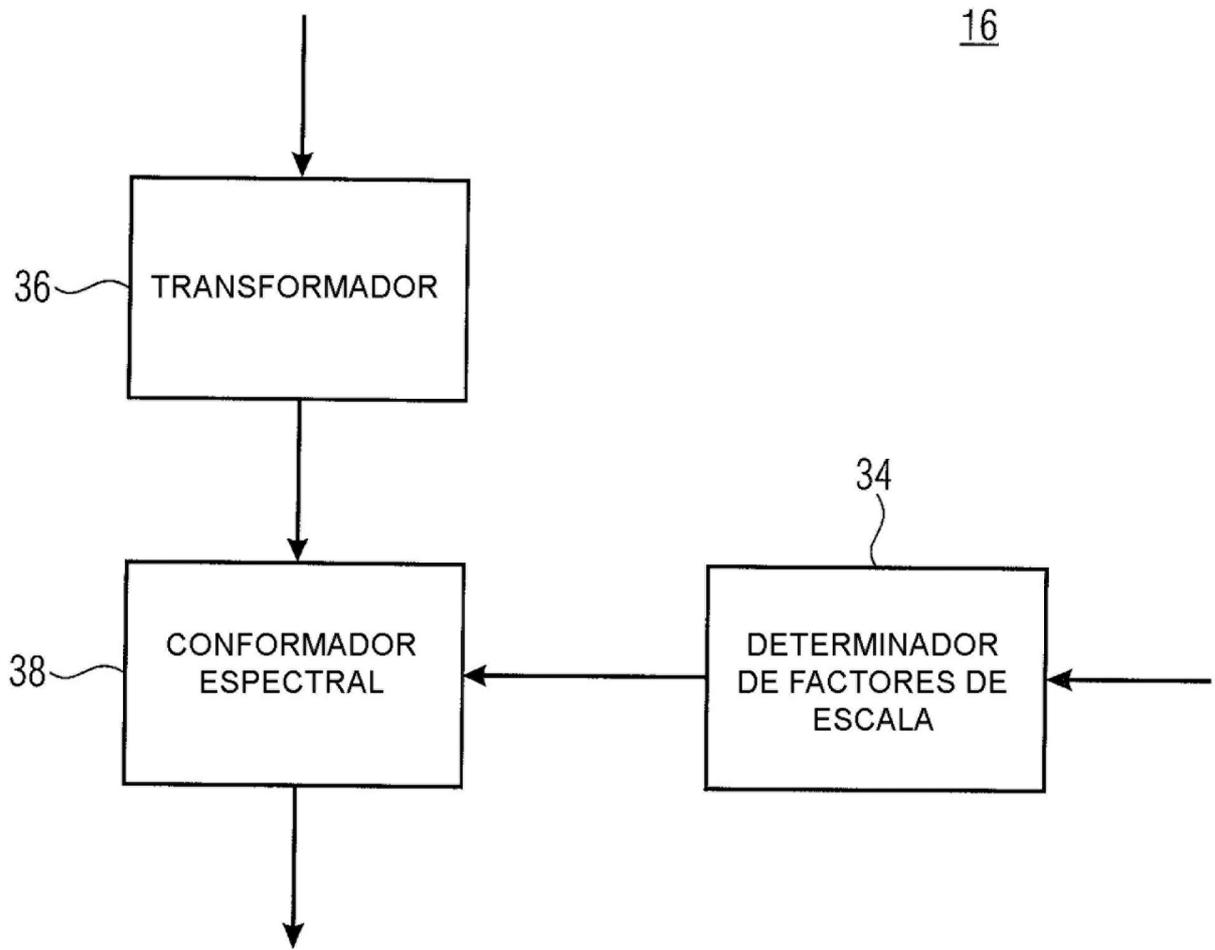


FIG 2

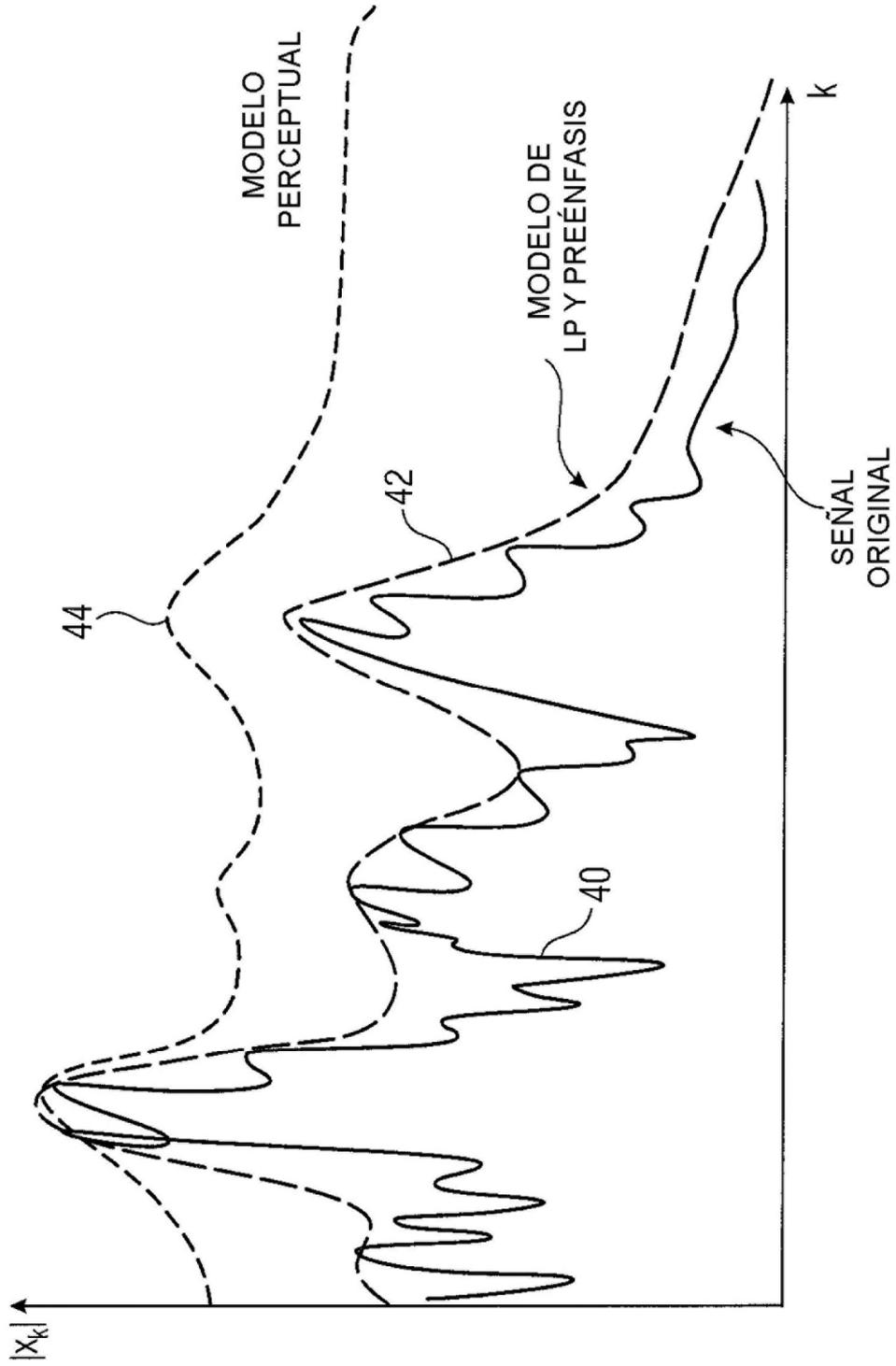


FIG 3A

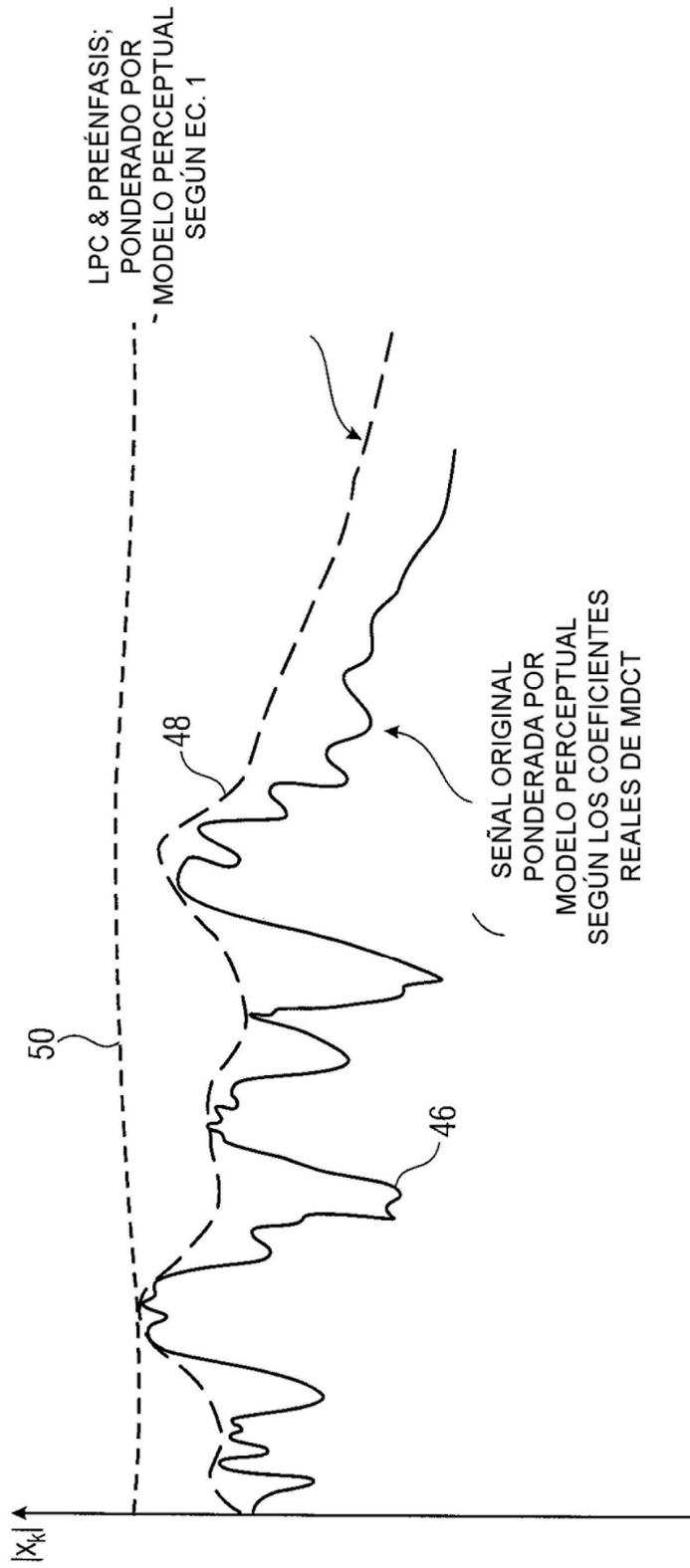


FIG 3B

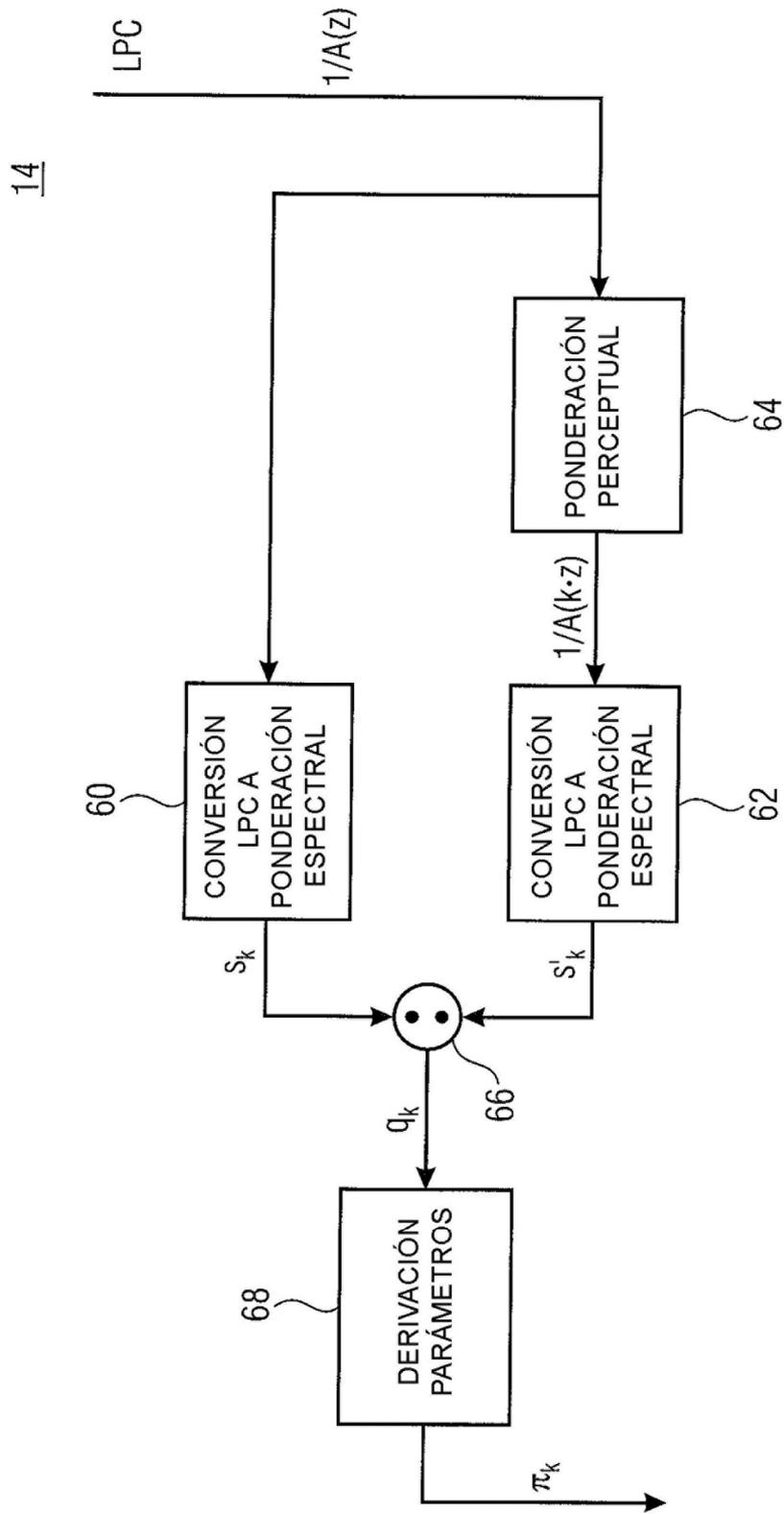


FIG 4

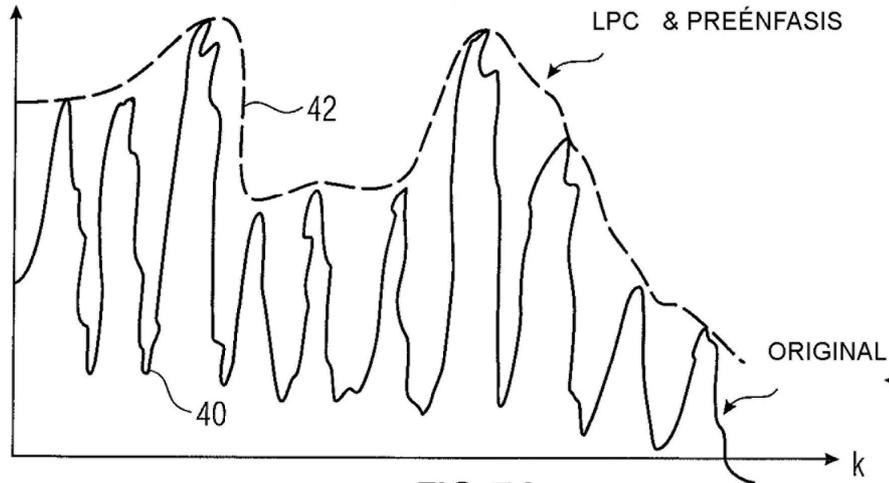


FIG 5A

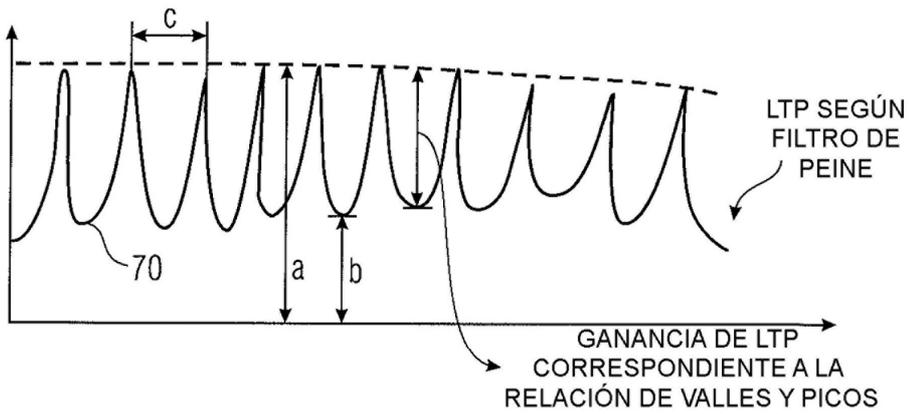


FIG 5B

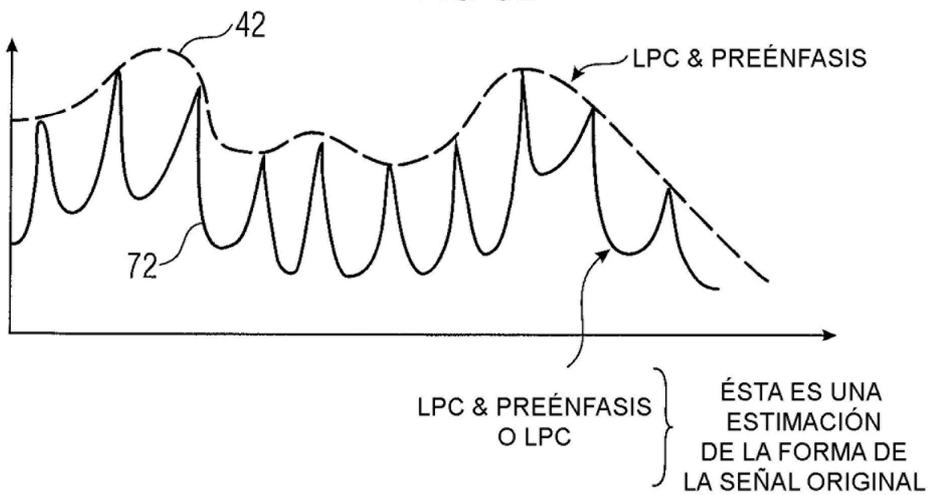


FIG 5C

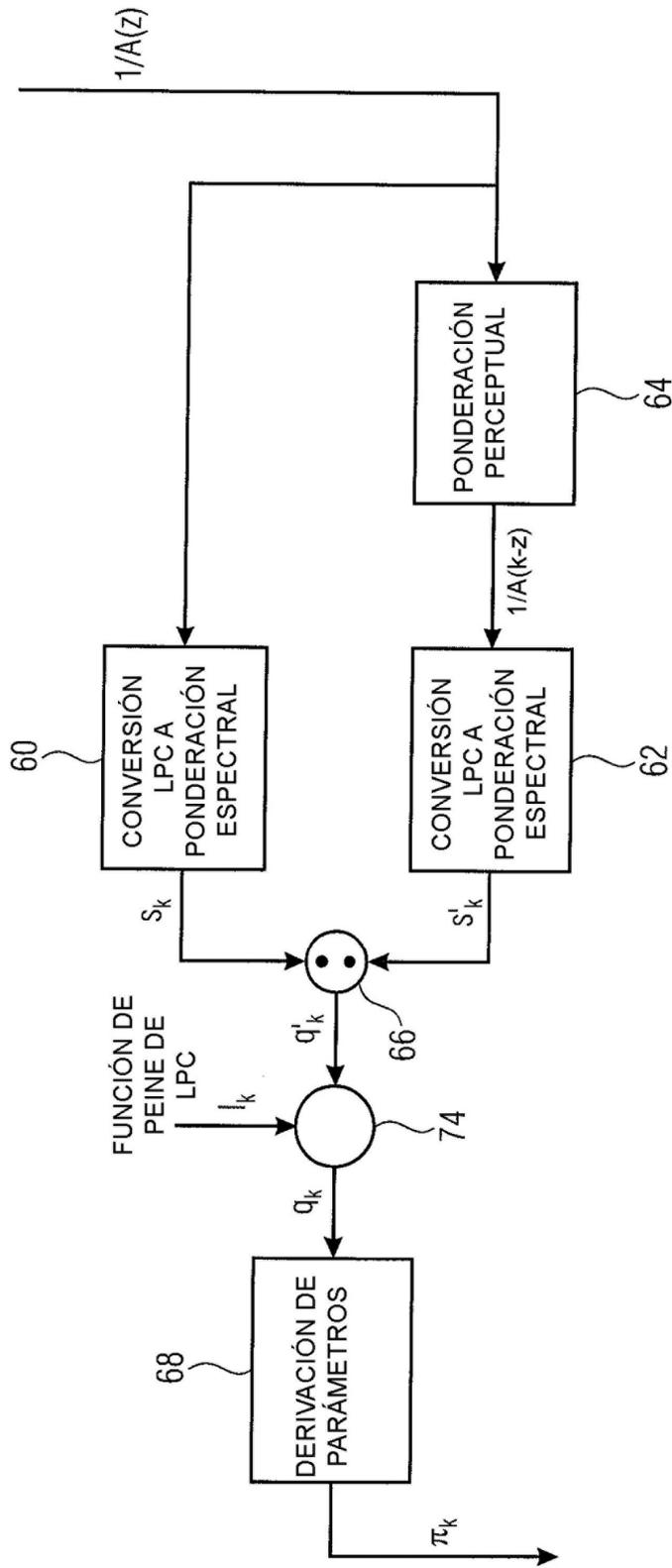


FIG 6

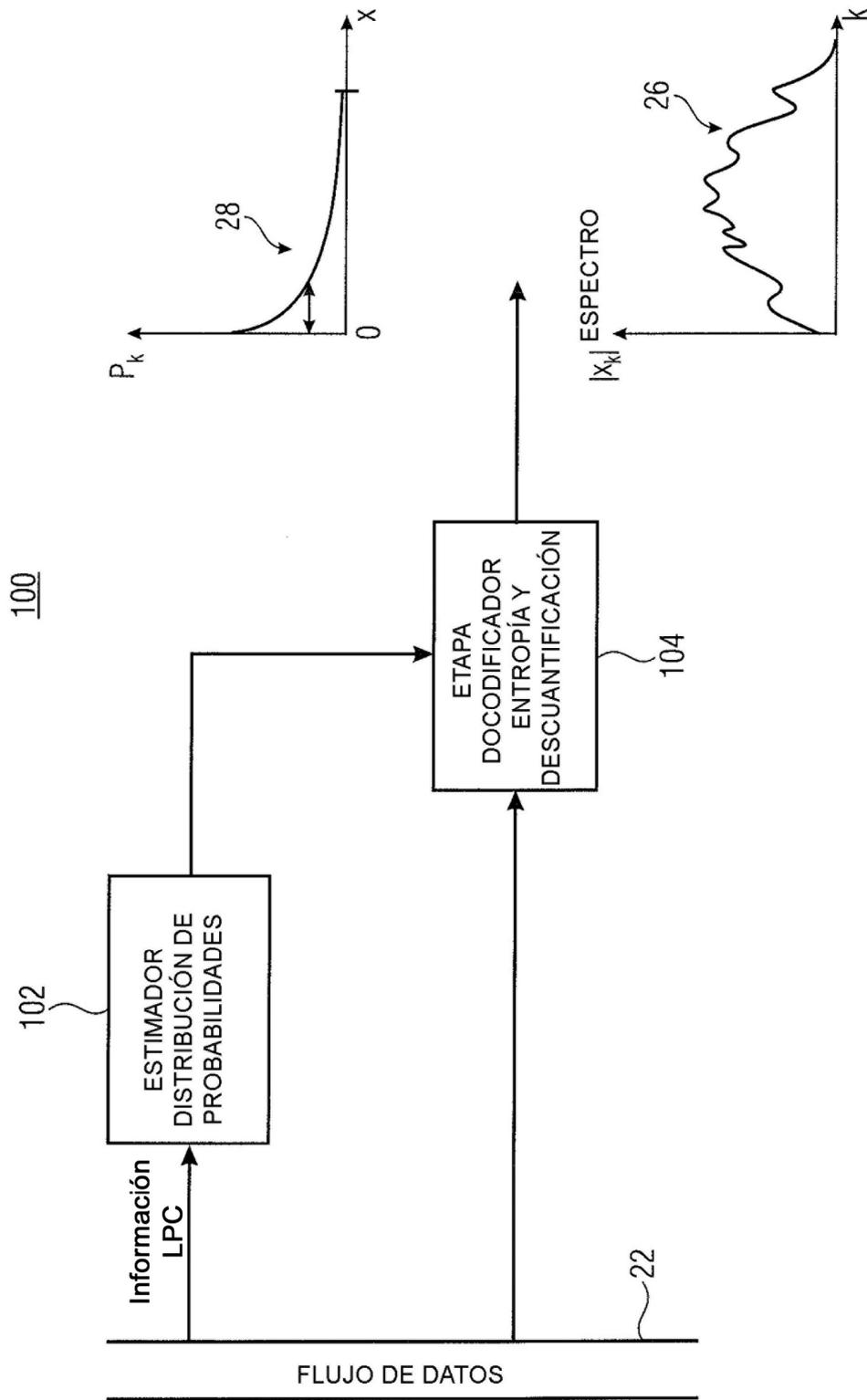


FIG 7

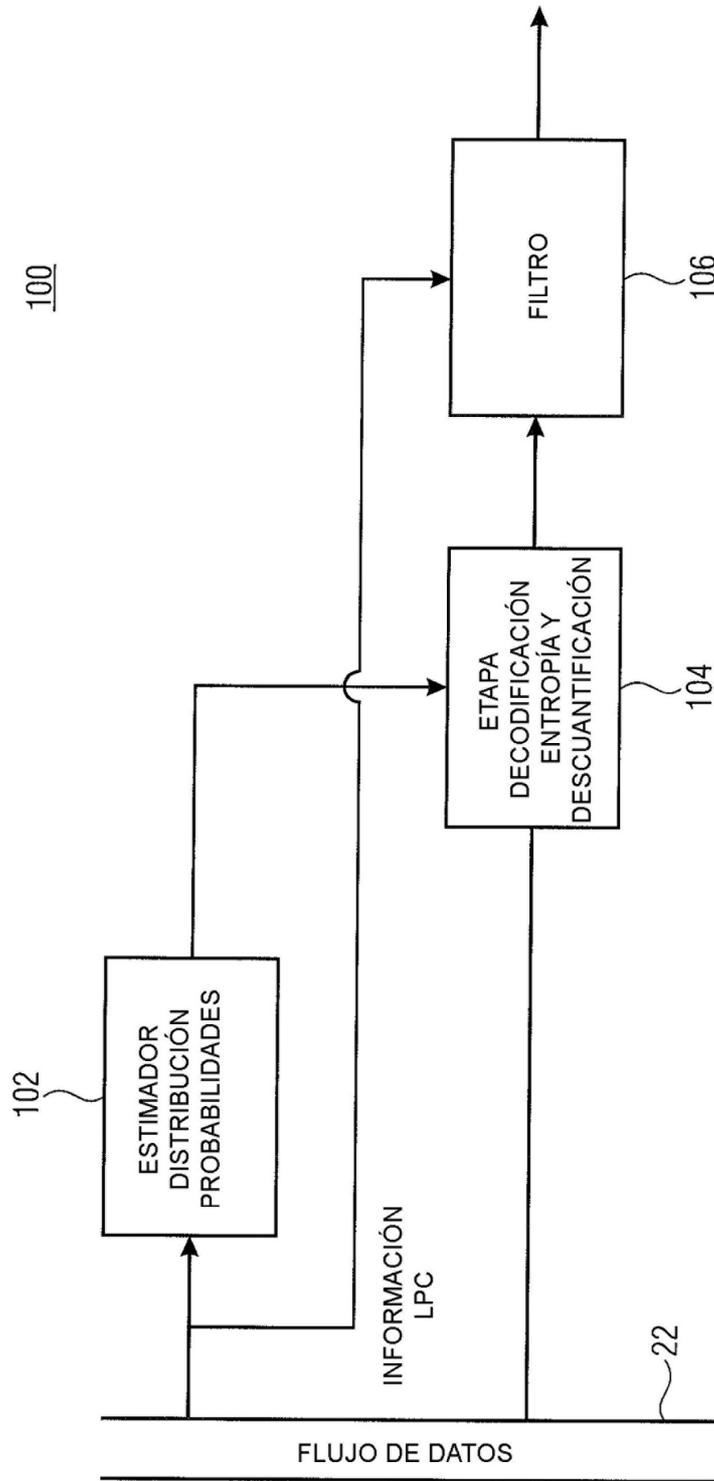


FIG 8

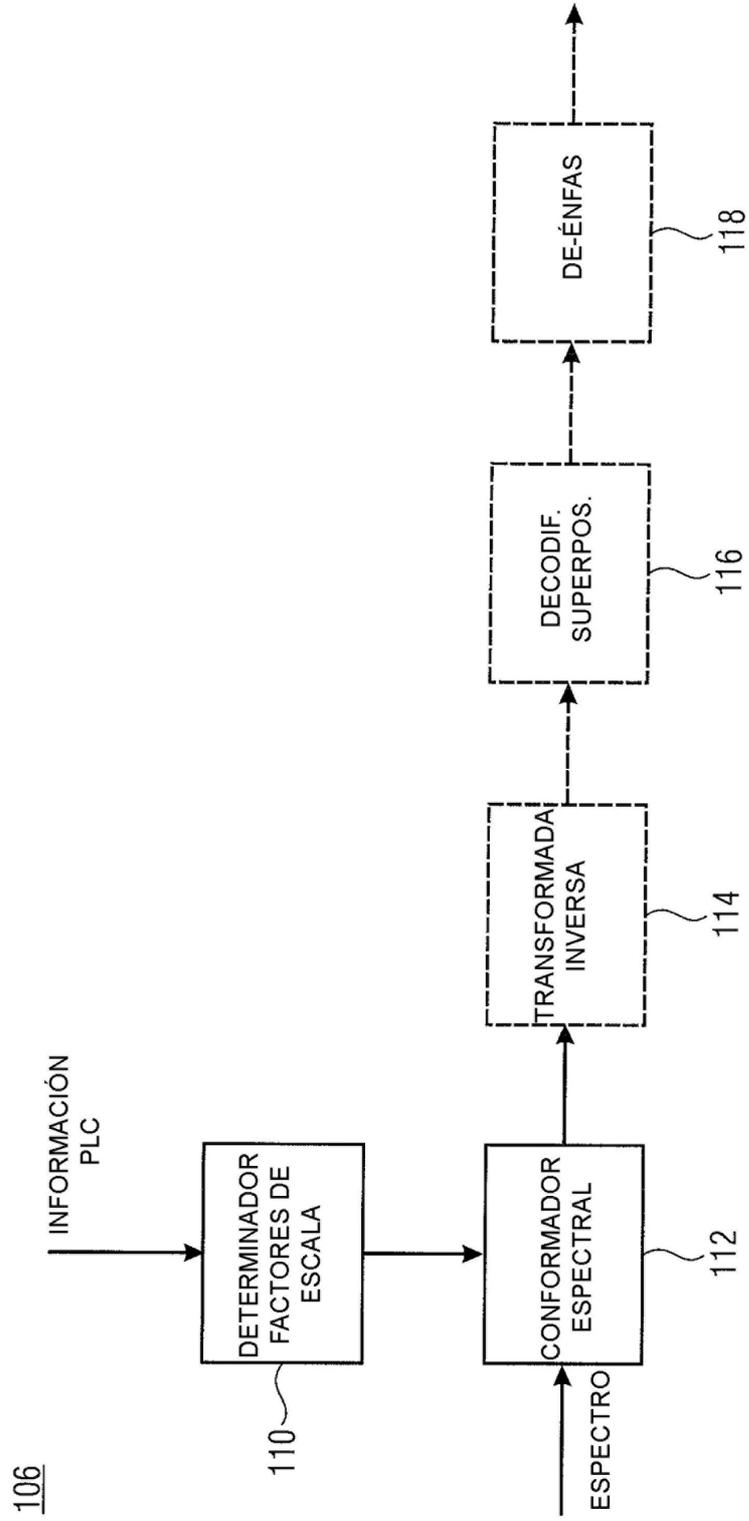


FIG 9

106

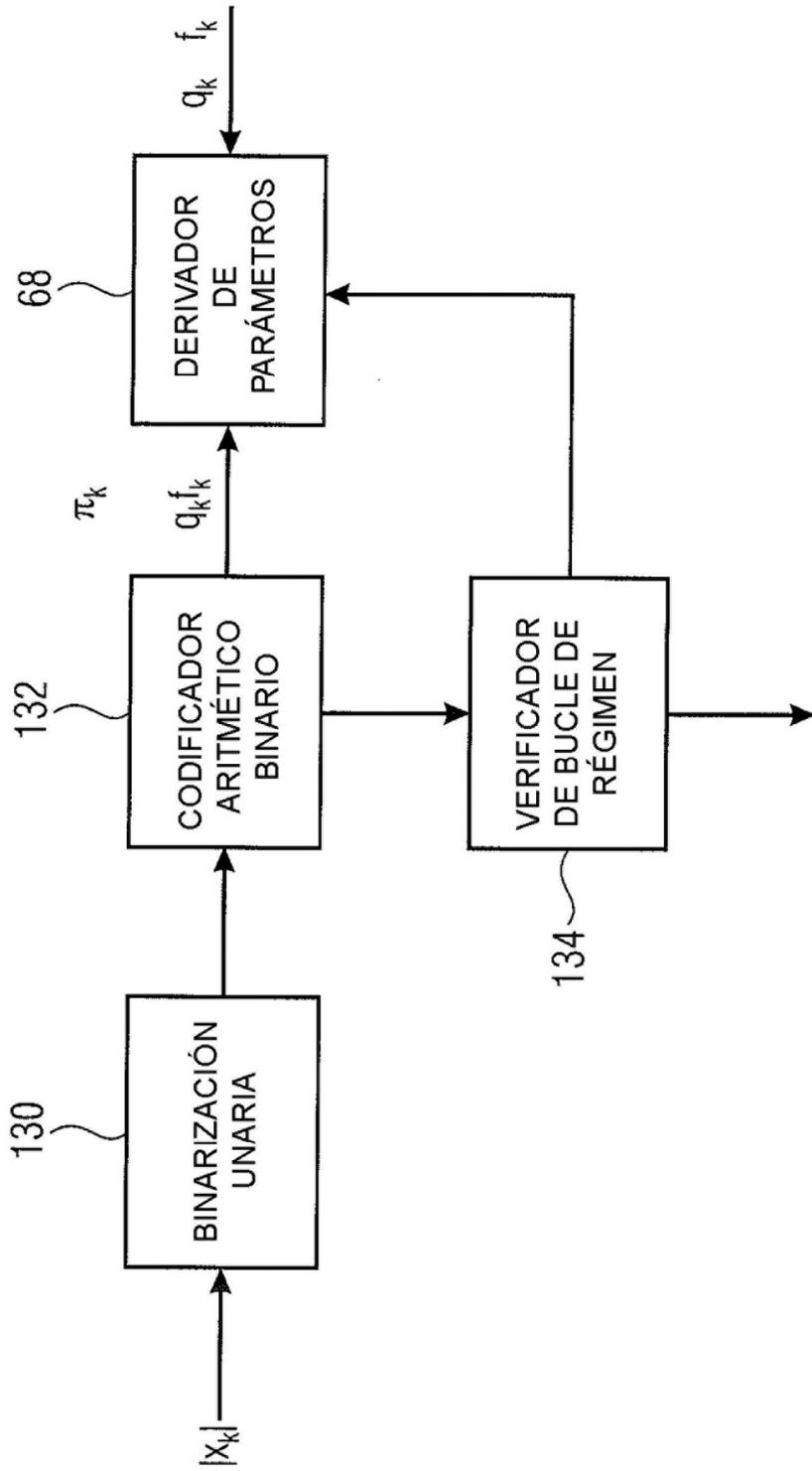


FIG 10

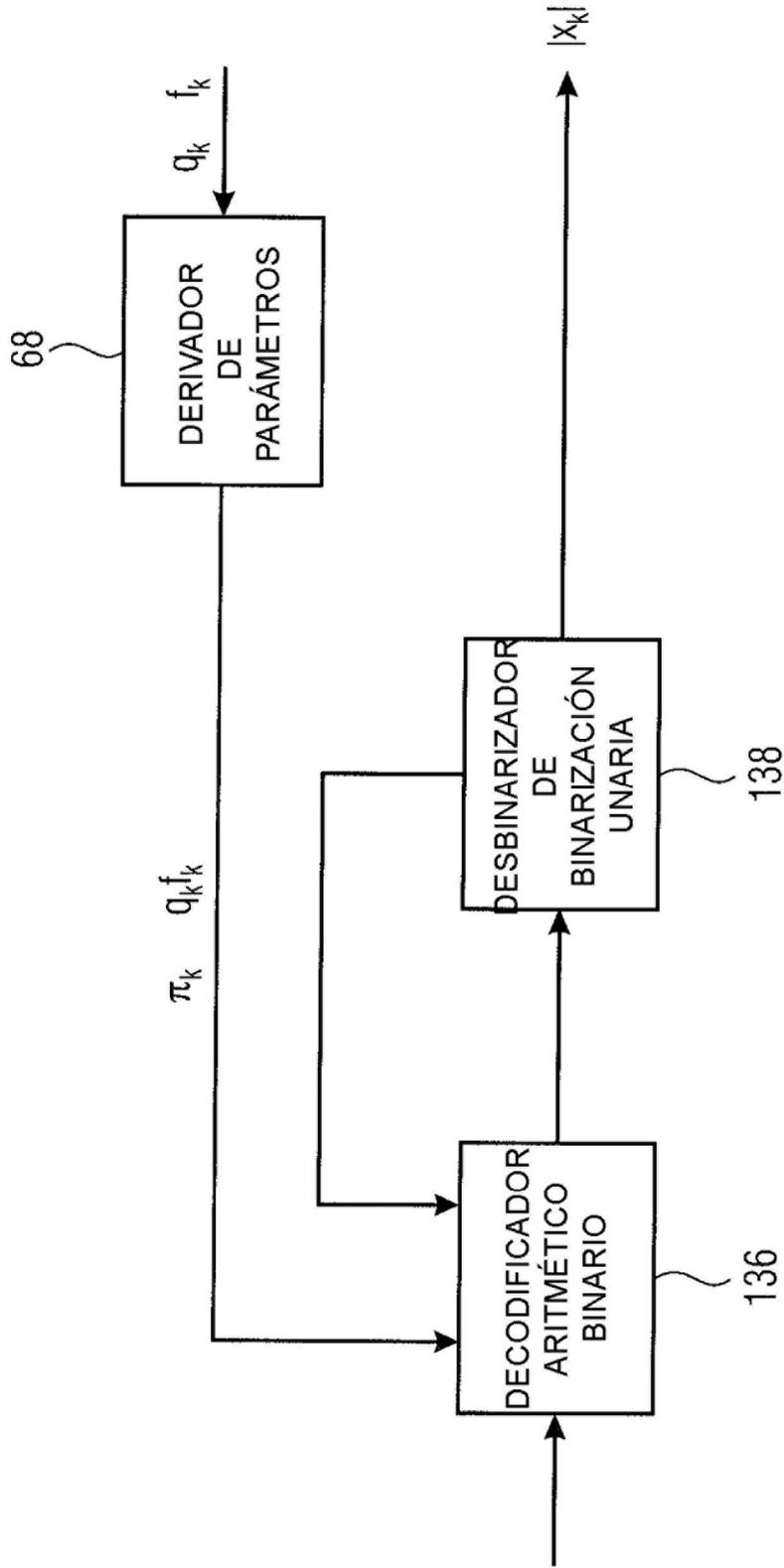


FIG 11