

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 665 646**

51 Int. Cl.:

G10L 19/02 (2013.01)

G10L 19/032 (2013.01)

G10L 21/038 (2013.01)

G10L 19/038 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **15.07.2014 PCT/EP2014/065173**

87 Fecha y número de publicación internacional: **29.01.2015 WO15010966**

96 Fecha de presentación y número de la solicitud europea: **15.07.2014 E 14738866 (4)**

97 Fecha y número de publicación de la concesión europea: **07.03.2018 EP 3025338**

54 Título: **Codificación de entropía basada en el contexto de valores de muestra de una envolvente espectral**

30 Prioridad:

22.07.2013 EP 13177351
18.10.2013 EP 13189336

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
26.04.2018

73 Titular/es:

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

72 Inventor/es:

GHIDO, FLORIN y
NIEDERMEIER, ANDREAS

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 665 646 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).

Codificación de entropía basada en el contexto de valores de muestra de una envolvente espectral

DESCRIPCIÓN

5 La presente solicitud se refiere a la codificación de entropía basada en el contexto, de valores de muestra de una envolvente espectral, y a su uso en la codificación y compresión de audio.

10 Muchos codificadores de audio con pérdida del estado de la técnica modernos tales como los descritos en [1] y [2] se basan en una transformada de coseno discreta modificada MDCT, y utilizan tanto la reducción de irrelevancia como la reducción de redundancia para minimizar la tasa de bits requerida para una determinada calidad perceptual. La reducción de irrelevancia típicamente aprovecha las limitaciones perceptuales del sistema auditivo humano, para reducir la precisión de representación o eliminar la información de frecuencia que no es perceptualmente relevante. La reducción de redundancia se aplica para aprovechar la correlación o estructura estadística, para lograr la representación más compacta de los datos restantes, típicamente, usando el modelado estadístico en conjunto con la codificación de entropía.

15 Entre otros, se usan los conceptos de la codificación paramétrica para codificar de manera eficiente contenido de audio. Con el uso de la codificación paramétrica, porciones de la señal de audio, por ejemplo, porciones de su espectrograma, se describen usando parámetros en lugar de usando muestras de audio de dominio de tiempo reales o similares. Por ejemplo, pueden sintetizarse porciones del espectrograma de una señal de audio en el lado del decodificador, comprendiendo meramente la corriente de datos parámetros tales como la envolvente espectral y otros parámetros adicionales que controlan la sintetización, para adaptar la porción del espectrograma sintetizado a la envolvente espectral transmitida. Una nueva técnica de este tipo es la Replicación de Banda Espectral (SBR) de acuerdo con la cual se usa un códec núcleo para codificar y transmitir el componente de baja frecuencia de una señal de audio, mientras que se usa una envolvente espectral transmitida en el lado de la decodificación, para conformar/formar replicas espectrales de una reconstrucción del componente de banda de baja frecuencia de la señal de audio para sintetizar el componente de banda de alta frecuencia de la señal de audio en el lado de la decodificación.

20 Una envolvente espectral dentro de la estructura de las técnicas de codificación que se han señalado anteriormente se transmite dentro de una corriente de datos a alguna resolución espectro-temporal adecuada. De una manera similar a la transmisión de valores de muestra de envolvente espectral, los factores de escala para la escala de los coeficientes lineales espectrales o los coeficientes de dominio de frecuencia tales como los coeficientes de MDCT se transmiten de manera análoga en alguna resolución espectro-temporal adecuada que es más basta que la resolución lineal espectral original, más basta, por ejemplo, en un sentido espectral.

25 Una tabla de codificación de Huffman fijada podría usarse para transportar información en las muestras que describen una envolvente espectral o factores de escala o coeficientes de dominio de frecuencia. Un enfoque mejorado es el uso de codificación de contexto, por ejemplo, como se describe en [2] y [3], donde el contexto utilizado para seleccionar la distribución de probabilidad para la codificación de un valor se extiende tanto a través del tiempo como de la frecuencia. Una línea espectral individual tal como un valor de coeficiente de MDCT es la proyección real de una línea espectral compleja, y puede parecer algo aleatoria en su naturaleza, incluso cuando la magnitud de la línea espectral compleja sea constante a través del tiempo, pero la fase varía de una trama a la siguiente. Esto requiere un diagrama esquemático muy complejo de selección de contexto, cuantificación y mapeo para buenos resultados, como se describe en [3].

El documento de patente US6978236 desvela un codificador para codificación de envolvente espectral, en el que se codifican factores de escala de envolvente espectral en cualquiera de la dirección de tiempo o de frecuencia.

30 En la codificación de imagen, los contextos utilizados son típicamente bidimensionales a través de los ejes x e y de una imagen, por ejemplo, como en [4]. En la codificación de imagen, los valores están en el dominio lineal o el dominio de ley de la potencia, por ejemplo, mediante el uso de un ajuste gamma. Además, puede usarse una sola predicción lineal fijada, en cada contexto, como un ajuste de plano y un mecanismo de detección de bordes rudimentario, y puede codificarse el error de predicción. Puede usarse la codificación paramétrica Golomb o Golomb-Rice para la codificación de los errores de predicción. Adicionalmente, se usa la codificación de longitud de serie para compensar las dificultades de la codificación directa de señales de muy baja entropía, por debajo de 1 bit por muestra, por ejemplo, usando un codificador basado en bits.

35 Sin embargo, a pesar de las mejoras en relación con la codificación para factores de escala y/o envolventes espectrales, existe aún la necesidad de buscar un concepto mejorado para la codificación de valores de muestra de una envolvente espectral. En consecuencia, es un objeto de la presente invención proporcionar un concepto para la codificación de valores espectrales de una envolvente espectral.

Este objeto se consigue por la materia objeto de las reivindicaciones independientes 1, 13, 18, 19, 20 y 21. Las

realizaciones que se describen en el presente documento se basan en el hallazgo de que un concepto mejorado para la codificación de valores de muestra de una envolvente espectral puede obtenerse mediante la combinación de la predicción espectro-temporal por una parte, y la codificación de entropía basada en el contexto de los residuales, por otra parte, mientras se determina en particular el contexto para un valor de muestra actual dependiente de una medida para una desviación entre un par de valores de muestra ya codificados/decodificados de la envolvente espectral en una proximidad espectro-temporal del valor de muestra actual. La combinación de la predicción espectro-temporal por una parte, y la codificación de entropía basada en el contexto de los residuales de predicción, dependiendo de la selección de contexto de la medida de desviación, por otra parte, armoniza con la naturaleza de las envolventes espectrales: la suavidad de la envolvente espectral da como resultado compactas distribuciones residuales de predicción, de modo que la intercorrelación espectro-temporal se elimina casi por completo después de la predicción, y puede no considerarse en la selección de contexto con respecto a la codificación de entropía del resultado de predicción. Esto, a su vez, disminuye la sobrecarga para el gestionar los contextos. El uso de la medida de desviación entre valores de muestra ya codificados/decodificados en la proximidad espectro-temporal del valor de muestra actual, sin embargo, aún permite la provisión de una adaptabilidad de contexto que mejora la eficiencia de codificación de entropía de una manera que justifica la sobrecarga adicional causada de ese modo.

De acuerdo con realizaciones descritas a continuación, se combina la predicción lineal con el uso del valor de diferencia como la medida de desviación, manteniendo de esta manera la sobrecarga para la codificación baja

De acuerdo con una realización, la posición de los valores de muestra ya codificados/decodificados utilizados para determinar el valor de diferencia finalmente utilizado para seleccionar/determinar el contexto se selecciona de modo tal que los valores son vecinos entre sí, espectral o temporalmente, de manera co-alineada con el valor de muestra actual, es decir, yacen sobre una línea en paralelo al eje temporal o espectral, y el signo del valor de diferencia se tiene en cuenta adicionalmente cuando se determina/selecciona el contexto. Por esta medida, una clase de "tendencia" en el residual de predicción puede tenerse en cuenta cuando se determina/selecciona el contexto para el valor de muestra actual, mientras que se incrementa de manera meramente razonable la sobrecarga de gestión de contexto.

Las realizaciones preferidas de la presente solicitud se describen a continuación con respecto a las figuras, en las cuales:

Fig. 1 muestra un diagrama esquemático de una envolvente espectral e ilustra su composición, a partir de valores de muestra y un posible orden de decodificación definido entre ellos, al igual que una posible proximidad espectro-temporal para un valor de muestra actualmente codificado/decodificado de la envolvente espectral;

Fig. 2 muestra un diagrama de bloques de un codificador de entropía basado en el contexto para la codificación de valores de muestra de una envolvente espectral de acuerdo con una realización;

Fig. 3 muestra un diagrama esquemático que ilustra una función de cuantificación que puede usarse en la cuantificación de la medida de derivación;

Fig. 4 muestra un diagrama de bloques de un decodificador de entropía basado en el contexto que se ajusta al codificador de la Fig. 2;

Fig. 5 muestra un diagrama de bloques de un codificador de entropía basado en el contexto para la codificación de valores de muestra de una envolvente espectral de acuerdo con una realización adicional;

Fig. 6 muestra un diagrama esquemático que ilustra la colocación del intervalo de valores posibles codificados de entropía del residual de predicción en relación con el intervalo general de posibles valores de los residuales de predicción de acuerdo con una realización que usa la codificación de escape;

Fig. 7 muestra un diagrama de bloques de un decodificador de entropía basado en el contexto que se ajusta al codificador de la Fig. 5;

Fig. 8 muestra una posible definición de una proximidad espectro-temporal usando una cierta notación;

Fig. 9 muestra un diagrama de bloques de un decodificador de audio paramétrico de acuerdo con una realización;

Fig. 10 muestra un diagrama esquemático que ilustra una posible variante de implementación del decodificador paramétrico de la Fig. 9, mostrando la relación entre el intervalo de frecuencia cubierto por la envolvente espectral, por una parte, y la estructura fina que cubre otro intervalo del rango de frecuencia de la señal

de audio general, por otra parte;

- Fig. 11 muestra un diagrama de bloques de un codificador de audio que se ajusta al decodificador de audio paramétrico de la Fig. 9 de acuerdo con la variante de la Fig. 10;
- 5 Fig. 12 muestra un diagrama esquemático que ilustra una variante del decodificador de audio paramétrico de la Fig. 9 cuando soporta IGF (Relleno Inteligente de Hueco);
- Fig. 13 muestra un diagrama esquemático que ilustra un espectro a partir de un espectrograma de estructura fina, es decir, un corte espectral, el relleno de IGF del espectro y su conformación de acuerdo con la envolvente espectral de acuerdo con una realización; y
- 10 Fig. 14 muestra un diagrama de bloques de un codificador de audio que soporta IGF, que se ajusta a la variante del decodificador paramétrico de la Fig. 9 de acuerdo con la Fig. 12.

15 Como una clase de motivación de las realizaciones representadas en el presente documento a continuación, que se aplican generalmente a la codificación de una envolvente espectral, se presentan ahora algunos conceptos que conducen a las realizaciones ventajosas que se indican a continuación, usando el Relleno Inteligente de Hueco (IGF) como un ejemplo. IGF es un nuevo método para mejorar de manera significativa la calidad de una señal codificada incluso a muy bajas tasas de bits. Se hace referencia a la descripción que sigue para detalles. En cualquier caso, IGF trata el hecho de que una parte significativa de un espectro en la región de alta frecuencia se cuantifica a cero, debido típicamente al insuficiente presupuesto de bits. Para preservar lo más posible la estructura fina de la región de frecuencia superior, se usa información de IGF en la región de baja frecuencia como una fuente para reemplazar de manera adaptativa las regiones de destino en la región de alta frecuencia que se cuantifican en su mayoría a cero. Un importante requisito para lograr una buena calidad perceptual es la coincidencia de la envolvente de energía decodificada de los coeficientes espectrales con la de la señal original. Para lograr esto, se calculan las energías espectrales promedio sobre coeficientes espectrales a partir de una o más bandas de factor de escala AAC consecutivas. Calcular energías promedio usando límites definidos por bandas de factor de escala se ve motivado por la ya existente cuidadosa adaptación de dichos límites a fracciones de las bandas críticas, que son características de la audición humana. Las energías promedio se convierten en una representación de escala de dB usando una fórmula similar a la de los factores de escala AAC, y a continuación, se cuantifican de manera uniforme. En IGF, puede usarse de manera opcional una diferente exactitud de cuantificación, de acuerdo con la tasa de bits total solicitada. Las energías promedio constituyen una significativa parte de la información generada por el IGF, de modo que su eficiente representación es de alta importancia para el rendimiento global de IGF.

20 Por lo tanto, en IGF, las energías de factor de escala describen la envolvente espectral. Las Energías de Factor de Escala (SFE) representan valores espectrales que describen la envolvente espectral. Es posible aprovechar propiedades especiales de las SFE cuando se decodifican las mismas. En particular, se ha observado que, en contraste a [2] y [3], las SFE representan valores promedio de líneas espectrales de MDCT y, en consecuencia, sus valores son mucho más "suaves" y linealmente correlacionados con la magnitud promedio de las correspondientes líneas espectrales complejas. Aprovechándose de esta circunstancia, las siguientes realizaciones usan una combinación de la predicción de valores de muestra de envolvente espectral, por una parte, y la codificación de entropía basada en el contexto del residual de predicción usando contextos que dependen de una medida de una desviación de un par de valores de muestra vecinos ya codificados/decodificados de la envolvente espectral, por otra parte. El uso de esta combinación está particularmente adaptado a esta clase de datos que se van a codificar, es decir, la envolvente espectral.

25 Para facilitar la comprensión de las realizaciones representadas adicionalmente a continuación, la Fig. 1 muestra una envolvente espectral 10 y su composición a partir de valores de muestra 12 que muestrean la envolvente espectral de la señal de audio 10 a una cierta resolución espectro-temporal. En la Fig. 1, los valores de muestra 12 se disponen ejemplarmente a lo largo del eje de tiempo 14 y el eje espectral 16. Cada valor de muestra 12 describe o define la altura de la envolvente espectral 10 dentro de una correspondiente pieza espacio-temporal que cubre, por ejemplo, un cierto rectángulo del dominio espacio-temporal de un espectrograma de una señal de audio. Los valores de muestra son, en consecuencia, valores integrativos que se han obtenido mediante la integración de un espectrograma sobre su pieza espectro-temporal asociada. Los valores de muestra 12 pueden medir la altura o intensidad de la envolvente espectral 10 en términos de energía o alguna otra medida física, y pueden definirse en el dominio no logarítmico o el dominio lineal, o en el dominio logarítmico, en el que el dominio logarítmico puede proporcionar ventajas adicionales debido a su característica de suavizar adicionalmente los valores de muestra a lo largo de los ejes 14 y 16, respectivamente.

30 Debe observarse que, en lo que respecta a la siguiente descripción, se asume, con propósitos de ilustración, que solo los valores de muestra 12 se disponen regularmente de manera espectral y temporal, es decir, que las correspondientes piezas espacio-temporales correspondientes a los valores de muestra 12 cubren regularmente una banda de frecuencia 18 de un espectrograma de una señal de audio, si bien dicha regularidad no es obligatoria. En cambio, puede usarse también un muestreo irregular de la envolvente espectral 10 por los valores de muestra 12,

representando cada valor de muestra 12 la media de promedio de la altura de la envolvente espectral 10 dentro de su correspondiente pieza espacio-temporal. Las definiciones de proximidad que se indican adicionalmente a continuación pueden, sin embargo, transferirse a dichas realizaciones alternativas de un muestreo irregular de la envolvente espectral 10. Una breve declaración sobre dicha posibilidad se presenta a continuación.

5 Antes, sin embargo, se observa que la envolvente espectral anteriormente mencionada puede someterse a la codificación y decodificación para la transmisión de codificador a decodificador, por varias razones. Por ejemplo, la envolvente espectral puede usarse con propósitos de escalabilidad para extender una codificación núcleo de una banda de frecuencia baja de una señal de audio, en concreto, extender la banda de frecuencia baja hacia
10 frecuencias más altas, en concreto, a una banda de alta frecuencia con al cual se relaciona la envolvente espectral. En dicho caso, los decodificadores/codificadores de entropía basados en el contexto que se describen a continuación podrían ser parte de un decodificador/codificador de SBR, por ejemplo. Como alternativa, podrían ser parte de codificadores/decodificadores de audio que usan IGF como ya se mencionó. En IGF, una porción de alta frecuencia de un espectrograma de señal de audio se describe adicionalmente usando los valores espectrales que
15 describen la envolvente espectral de porciones de alta frecuencia del espectrograma para poder llenar áreas cuantificadas en cero del espectrograma dentro de la porción de alta frecuencia usando la envolvente espectral. Los detalles en este sentido se describen adicionalmente a continuación.

20 La Fig. 2 muestra el codificador de entropía basado en el contexto para la codificación de valores de muestra 12 de una envolvente espectral 10 de una señal de audio de acuerdo con una realización de la presente solicitud.

El codificador de entropía basado en el contexto de la Fig. 2 se indica generalmente usando el signo de referencia
25 los valores de muestra 12 de la envolvente espectral (Fig. 1). El codificador de entropía 26 tiene una entrada de control conectada a una salida del determinador de contexto 24, y una entrada de datos conectada a una salida del determinador de residual 28. El determinador de residual 28 tiene dos entradas, una de las cuales está conectada a una salida del predictor 22, y la otra proporciona al determinador de residual 28 el acceso a los valores de muestra 12 de la envolvente espectral 10. En particular, el determinador de residual 28 recibe el valor de muestra x
30 actualmente a codificarse en su entrada, mientras que el determinador de contexto 24 y el predictor 22 reciben en sus entradas valores de muestra 12 que han sido codificados y que residen dentro de una proximidad espectro-temporal del valor de muestra actual x .

El predictor 22 está configurado para predecir espectro-temporalmente el valor de muestra actual x de la envolvente
35 espectral 10 para obtener un valor estimado \hat{x} . Como se ilustrará en relación con una realización más detallada señalada a continuación, el predictor 22 puede utilizar predicción lineal. En particular, en la realización de la predicción espectro-temporal, el predictor 22 inspecciona valores de muestra ya codificados en una proximidad espectro-temporal de valor de muestra actual x . Véase, por ejemplo, la Fig. 1. El valor de muestra actual x se ilustra usando un delineado trazado de manera continua en negrita. Usando líneas de puntos, se muestran los valores de
40 muestra en la proximidad espectro-temporal de la muestra actual x que, de acuerdo con una realización, forman una base para la predicción espectro-temporal del predictor 22. "a", por ejemplo, indica el valor de muestra 12 inmediatamente vecino a la muestra actual x , que se co-ubica con respecto a la muestra actual x espectralmente, si bien precede a la muestra actual x temporalmente. Asimismo, el valor de muestra "b" vecino indica el valor de muestra inmediatamente vecino a la muestra actual x , que se co-ubica con respecto al valor de muestra actual x
45 temporalmente, si bien se refiere a frecuencias inferiores en comparación con el valor de muestra actual x , y el valor de muestra "c" en la proximidad espectro-temporal del valor de muestra actual x es valor de muestra vecino más cercano de valor de muestra actual x , que precede a este último temporalmente, y se refiere a frecuencias inferiores. La proximidad espectro-temporal puede incluso abarcar valores de muestra que representan el vecino siguiente de la muestra actual x . Por ejemplo, el valor de muestra "d" se separa del valor de muestra actual x por el valor de
50 muestra "a", es decir, está co-ubicado con respecto al valor de muestra actual x temporalmente, y precede al valor actual x meramente con el valor de muestra "a" posicionado en el medio. De manera similar, el valor de muestra "e" es vecino del valor de muestra x mientras se co-ubica con respecto al valor de muestra actual x temporalmente, y es vecino del valor de muestra x a lo largo del eje espectral 16 meramente con la muestra vecina "b" posicionada en el medio.

55 Como ya se indicó con anterioridad, si bien se asume que los valores de muestra 12 se disponen regularmente a lo largo de los ejes de tiempo y espectrales 14 y 16, esta regularidad no es obligatoria, y la definición de proximidad y la identificación de valores de muestra vecinos puede ampliarse a dicho caso irregular. Por ejemplo, el valor de muestra "a" vecino puede definirse como aquel que es vecino de la esquina izquierda superior de la pieza espectro-temporal de la muestra actual a lo largo del eje temporal precediendo la esquina izquierda superior temporalmente.
60 Pueden usarse definiciones similares para definir también otros vecinos, tales como los vecinos b a e.

Como se indicará en más detalle a continuación, el predictor 22 puede, dependiendo de la posición espectro-temporal del valor de muestra actual x , usar un subconjunto diferente de todos los valores de muestra dentro de la

proximidad espectro-temporal, es decir, un subconjunto de {a, b, c, d, e}. El subconjunto que se usa en realidad, por ejemplo, depende de la disponibilidad de los valores de muestra vecinos dentro de la proximidad espectro-temporal definido por el conjunto {a, b, c, d, e}. Los valores de muestra vecinos a, d y c, por ejemplo, pueden no estar disponibles debido a que el valor de muestra actual x que sucede inmediatamente un punto de acceso aleatorio, es decir, un punto en el tiempo que permite a los decodificadores el inicio de la decodificación, de modo tal que las dependencias de porciones previas de la envolvente espectral 10 se prohíban o impidan. Como alternativa, los valores de muestra vecinos b, c y e pueden no estar disponibles debido a que el valor de muestra actual x representa el borde de baja frecuencia del intervalo 18, de modo tal que la posición del respectivo valor de muestra vecino cae fuera del intervalo 18. En cualquier caso, el predictor 22 puede predecir espectro-temporalmente el valor de muestra actual x mediante la combinación lineal de valores de muestra ya codificados dentro de la proximidad espectro-temporal.

La tarea del determinador de contexto 24 es seleccionar uno de los varios contextos soportados para la codificación de entropía de los residuales de predicción, es decir, $r = x - \hat{x}$. Para este fin, el determinador de contexto 24 determina el contexto para valor de muestra actual x dependiendo de una medida para una desviación entre un par de valores de muestra ya codificados entre a y e en la proximidad espectro-temporal. En las realizaciones específicas que se indican adicionalmente a continuación, la diferencia de un par de valores de muestra dentro de la proximidad espectro-temporal se usa como una medida para una desviación entre, por ejemplo, a - c, b - c, b - e, a - d o similares, si bien, como alternativa, pueden usarse otras medidas de desviación, por ejemplo, un cociente (es decir, a/c, b/c, a/d), la diferencia a la potencia de un valor no igual a uno, tal como un número impar n no igual a uno (es decir, $(a-c)^n$, $(b-c)^n$, $(a-d)^n$), o algún otro tipo de medida de desviación, por ejemplo, a^n-c^n , b^n-c^n , a^n-d^n o $(a/c)^n$, $(b/c)^n$, $(a/d)^n$ con $n \neq 1$. En este punto, n además podría ser cualquier valor superior a 1, por ejemplo.

Como se mostrará en más detalle a continuación, el determinador de contexto 24 puede configurarse para determinar el contexto para el valor de muestra actual x que depende de una primera medida para una desviación entre un primer par de valores de muestra ya codificadas en la proximidad espectro-temporal y una segunda medida para una desviación entre un segundo par de valores de muestra ya codificadas dentro la proximidad espectro-temporal, donde el primer par es vecino entre sí espectralmente, y el segundo par es vecino entre sí temporalmente. Por ejemplo, pueden usarse los valores de diferencia b - c y a - c donde a y c son vecinos entre sí espectralmente, y b y c son vecinos entre sí temporalmente. El mismo conjunto de valores de muestra vecinos, en concreto, {a, c, b}, puede usarse por el predictor 22 para obtener el valor estimado \hat{x} , en concreto, por ejemplo, por una combinación lineal de los mismos. Un conjunto diferente de valores de muestra vecinos puede usarse para la determinación de contexto y/o la predicción, en casos de cierta no disponibilidad de cualquiera de los valores de muestra a, c y/o b. Los factores de la combinación lineal pueden, como se expone adicionalmente a continuación, establecerse de modo tal que los factores son los mismos para diferentes contextos, en el caso de que la tasa de bits a la cual se codifica la señal de audio es mayor que un umbral predeterminado, y los factores se establecen individualmente para los diferentes contextos, en el caso donde la tasa de bits es inferior a un umbral predeterminado.

Como una observación intermedia, debe mencionarse que la definición de la proximidad espectro-temporal puede adaptarse al orden de la codificación/decodificación a lo largo del cual el codificador de entropía basado en el contexto 20 codifica secuencialmente los valores de muestra 12. Como se muestra en la Fig. 1, por ejemplo, el codificador de entropía basado en el contexto puede configurarse para codificar secuencialmente los valores de muestra 12 usando un orden de decodificación 30 que atraviesa los valores de muestra 12 instante de tiempo por instante de tiempo, donde, en cada instante de tiempo, conduce desde la menor hasta la mayor frecuencia. A continuación, los "instantes de tiempo" se indican como "tramas", si bien los instantes de tiempo, como alternativa, podrían denominarse intervalos de tiempo, unidades de tiempo o similares. En cualquier caso, en el uso de dicho atravesamiento espectral antes de la alimentación temporal hacia adelante, la definición de la proximidad espectro-temporal para ampliarse en el tiempo precedente y hacia frecuencias inferiores proporciona la más alta probabilidad factible de que los correspondientes valores de muestra ya hayan sido codificados/decodificados, y estén disponibles. En el presente caso, los valores dentro de la proximidad siempre están ya codificados/decodificados, siempre que estén presentes, si bien esto puede ser diferente para otros pares de proximidad y orden de decodificación. Naturalmente, el decodificador utiliza el mismo orden de decodificación 30.

Los valores de muestra 12 pueden, como ya se observó anteriormente, representar la envolvente espectral 10 en un dominio logarítmico. En particular, los valores espectrales 12 pueden haber sido ya cuantificados hasta valores enteros usando una función de cuantificación logarítmica. En consecuencia, debido a la cuantificación, las medidas de desviación determinadas por el determinador de contexto 24 ya pueden ser números enteros intrínsecamente. Este es el caso, por ejemplo, cuando se usa la diferencia como la medida de desviación. Independientemente de la naturaleza de número entero intrínseca de la medida de desviación determinada por el determinador de contexto 24, el determinador de contexto 24 puede someter la medida de desviación a la cuantificación y determinar el contexto usando la medida cuantificada. En particular, como se indicará a continuación, la función de cuantificación utilizada por el determinador de contexto 24 puede ser constante para valores de la medida de desviación fuera de un intervalo predeterminado, donde el intervalo predeterminado incluye, por ejemplo, cero.

La Fig. 3 muestra de manera ejemplar dicha función de cuantificación 32 que mapea medidas de desviación no cuantificadas a medidas de desviación cuantificadas, donde, en este ejemplo, el mencionado intervalo predeterminado 34 se extiende desde -2,5 hasta 2,5, donde los valores no cuantificados de medida de desviación sobre dicho intervalo se mapean constantemente al valor de medida de desviación cuantificado 3, y los valores de medida de desviación no cuantificados por debajo de dicho intervalo 34 se mapean de manera constante al valor de medida de desviación cuantificado -3. En consecuencia, se distinguen meramente siete contextos que deben soportarse por el codificador de entropía basado en el contexto. En los ejemplos de implementación indicados a continuación, la longitud del intervalo 34 es 5 como se acaba de ejemplificar, donde la cardinalidad del conjunto de posibles valores de los valores de muestra de la envolvente espectral es 2^n (por ejemplo = 128), es decir, superior a 16 veces la longitud del intervalo. En el caso de usar la codificación de escape ilustrada más adelante, el rango de posibles valores de los valores de muestra de la envolvente espectral puede definirse para que sea $[0; 2^n[$ donde n es un número entero seleccionado de modo tal que 2^{n+1} es inferior a la cardinalidad de posibles valores codificables de los valores de residuales de predicción, que se describe a continuación de acuerdo con un ejemplo de implementación específica 311.

El codificador de entropía 26 utiliza el contexto determinado por el determinador de contexto 24 para codificar por entropía de manera eficiente el residual de predicción r que, a su vez, se determina por el determinador de residual 28 basándose en el valor de muestra actual x y el valor estimado \hat{x} , por ejemplo, por medio de la resta. Preferentemente, se usa la codificación aritmética. Los contextos pueden asociarse con distribuciones de probabilidades constantes. Para cada contexto, la distribución de probabilidad asociada con los mismos asigna un cierto valor de probabilidad a cada posible símbolo de un alfabeto de símbolos del codificador de entropía 26. Por ejemplo, el alfabeto de símbolos del codificador de entropía 26 coincide o cubre el rango de posibles valores de residual de predicción r . En realizaciones alternativas, que se indican en más detalle a continuación, puede usarse un cierto mecanismo de codificación de escape para garantizar que el valor r a codificarse por entropía por el codificador de entropía 26 se encuentre dentro del alfabeto de símbolos del codificador de entropía 26. Cuando se usa la codificación aritmética, el codificador de entropía 26 utiliza la distribución de probabilidad del contexto determinado, determinado por el determinador de contexto 24, para subdividir un intervalo de probabilidad actual que representa el estado interno del codificador de entropía 26, en un subintervalo por valor de alfabeto, donde la selección de uno de los subintervalos depende del valor real de r , y emite una corriente de bits codificada aritméticamente que informa al lado de la decodificación sobre actualizaciones de desviación y ancho de intervalo de probabilidad mediante el uso de, por ejemplo, un proceso de renormalización. Como alternativa, sin embargo, el codificador de entropía 26 puede usar, para cada contexto, una tabla de codificación de longitud variable individual que traduce la distribución de probabilidad del respectivo contexto en un correspondiente mapeo de posibles valores de r sobre códigos de una longitud correspondiente a la respectiva frecuencia del respectivo posible valor r . Pueden usarse también otros códecs de entropía.

Por razones de completitud, la Fig. 2 muestra que un cuantificador 36 puede estar conectado delante de la entrada del determinador de residual 28, donde el valor de muestra actual x es entrante para obtener el valor de muestra actual \hat{x} , por ejemplo, como ya se mencionó anteriormente, mediante el uso de una función de cuantificación logarítmica, por ejemplo, aplicada a un valor de muestra x no cuantificado.

La Fig. 4 muestra un decodificador de entropía basado en el contexto de acuerdo con una realización, que se ajusta al codificador de entropía basado en el contexto de la Fig. 2.

El decodificador de entropía basado en el contexto de la Fig. 4 se indica usando el signo de referencia 40, y se construye de manera similar al codificador de la Fig. 2. Por lo tanto, el decodificador de entropía basado en el contexto 40 comprende un predictor 42, un determinador de contexto 44, un decodificador de entropía 46 y un combinador 48. El determinador de contexto 44 y el predictor 42 operan como el predictor 22 y el determinador de contexto 24 del codificador 20 de la Fig. 2. Esto es, el predictor 42 predice espectro-temporalmente el valor de muestra actual x , es decir, aquel actualmente a decodificarse, para obtener el valor estimado \hat{x} , y emite este hacia el combinador 48, y el determinador de contexto 44 determina el contexto para la decodificación de entropía del residual de predicción r de valor de muestra actual x dependiendo de la medida de desviación entre un par de valores de muestra ya decodificados dentro de la proximidad espectro-temporal de valor de muestra x , que informa al decodificador de entropía 46 del contexto determinado por medio de una entrada de control de este último. Por consiguiente, tanto el determinador de contexto 44 como el predictor 42 tienen acceso a los valores de muestra en la proximidad espectro-temporal. El combinador 48 tiene dos entradas conectadas a las salidas del predictor 42 y el decodificador de entropía 46, respectivamente, y una salida para emitir el valor de muestra actual. En particular, el codificador de entropía 46 decodifica por entropía el valor de residual r para los valores de muestra actuales \hat{x} usando el contexto determinado por el determinador de contexto 44, y el combinador 48 combina el valor estimado \hat{x} y el correspondiente valor de residual r para obtener el valor de muestra actual x , por ejemplo, por adición. Solo por razones de completitud, la Fig. 4 muestra que un descuantificador 50 puede suceder la salida del combinador 48 para descuantificar el valor de muestra emitido por el combinador 48, por ejemplo, sometiéndola a una conversión de dominio logarítmico a dominio lineal usando, por ejemplo, una función exponencial.

El decodificador de entropía 46 revierte la codificación de entropía realizada por el codificador de entropía 26. Esto es, el decodificador de entropía además gestiona una cantidad de contextos y usos, para un valor de muestra actual x , un contexto seleccionado por el determinador de contexto 44, donde cada contexto tiene una correspondiente distribución de probabilidad asociada con este, que asigna a cada posible valor de r una cierta probabilidad que es la misma que aquella seleccionada por el determinador de contexto 24 para el codificador de entropía 26.

Cuando se usa la codificación aritmética, el decodificador de entropía 46 revierte, por ejemplo, la secuencia de subdivisión de intervalo del codificador de entropía 26. El estado interno del decodificador de entropía 46 se define, por ejemplo, por el ancho de intervalo de probabilidad del intervalo actual, y un valor desviado que apunta, dentro del intervalo de probabilidad actual, al subintervalo del mismo al cual corresponde el valor real de r del valor de muestra actual x . El decodificador de entropía 46 actualiza el intervalo de probabilidad y el valor de desviado, usando la corriente de bits codificada aritméticamente entrante emitida por el codificador de entropía 26 tal como por medio de un proceso de renormalización y obtiene el valor real de r mediante la inspección del valor de desvío y la identificación del subintervalo dentro del cual cae.

Como ya se mencionó con anterioridad, puede ser ventajoso restringir la codificación de entropía de los valores residuales sobre cierto subintervalo pequeño de posibles valores de residuales de predicción r . La Fig. 5 muestra una modificación del codificador de entropía basado en el contexto de la Fig. 2 para realizar esto. Además de los elementos que se exponen en la Fig. 2, el codificador de entropía de contexto de la Fig. 5 comprende un control conectado entre el determinador de residual 28 y el codificador de entropía 26, en concreto, el control 60, al igual que un manipulador de codificación de escape 62 controlado por medio de un control 60.

La funcionalidad de control 60 se ilustra en la Fig. 5 en una manera de resumen. Como se ilustra en la Fig. 5, el control 60 inspecciona el valor residual r inicialmente determinado, determinado por el determinador de residual 28 basándose en una comparación del valor de muestra x real y su valor estimado \hat{x} . En particular, el control 60 inspecciona si r se encuentra dentro o fuera de un intervalo de valor predeterminado como se ilustra en la Fig. 5 en 64. Véase, por ejemplo, la Fig. 6. La Fig. 6 muestra a lo largo del eje x , posibles valores del residual de predicción inicial r , mientras que el eje y muestra la entropía realmente codificada r . Además, la Fig. 6 muestra el rango de posibles valores del residual de predicción r inicial, en concreto, 66, y el recién mencionado intervalo predeterminado 68 involucrado en la verificación 64. Imagine, por ejemplo, que los valores de muestra 12 son valores enteros entre 0 y 2^{n-1} , ambos inclusive. Entonces, el rango 66 de posibles valores para el residual de predicción r puede ampliarse desde $-(2^n-1)$ hasta 2^n-1 , ambos, inclusive, y los valores absolutos de los términos del intervalo 70 y 72 del intervalo 68 pueden ser menores o iguales que 2^{n-2} , es decir, los valores absolutos de los términos del intervalo pueden ser menores que $1/8$ de la cardinalidad del conjunto de posibles valores dentro del rango 66. En uno de los ejemplos de implementación que se exponen a continuación en relación con xHE-AAC, el intervalo 68 es de -12 a +12 inclusive, los términos del intervalo 70 y 72 son -13 y +13, y la codificación de escape extiende el intervalo 68 por la codificación de un valor absoluto codificado VLC, en concreto, extiende el intervalo 68 a $-/(13 + 15)$ usando 4 bits y a $-/(13 + 15 + 127)$ usando otros 7 bits, si los 4 bits anteriores eran 15. De este modo, el residual de predicción puede codificarse en un rango de $-/+155$, inclusive, para cubrir suficientemente el rango 66 de posibles valores para el residual de predicción que, a su vez, se extiende de -127 a 127. Como puede observarse, la cardinalidad de $[127; 127]$ es 255, y 13, es decir, los valores absolutos de los términos internos 70 y 72, es menor que $32 \approx 255/8$. Cuando se compara la longitud del intervalo 68 con la cardinalidad de posibles valores codificables usando la codificación de escape, es decir, $[-155; 155]$, entonces se descubre que los valores absolutos de los términos internos 70 y 72 pueden seleccionarse ventajosamente para que sean menores que $1/8$ o incluso $1/16$ de dicha cardinalidad (en este punto, 311).

En el caso de que el residual de predicción r inicial resida dentro del intervalo 68, el control 60 hace que el codificador de entropía 26 codifique por entropía este residual de predicción r inicial directamente. No debe tomarse ninguna medida especial. Sin embargo, si r como se proporciona por el determinador de residual 28 está fuera del intervalo 68, se inicia un procedimiento de codificación de escape por medio del control 60. En particular, los valores vecinos inmediatos que son vecinos inmediatos de los términos del intervalo 70 y 72 del intervalo 68 pueden, de acuerdo con una realización, pertenecer al alfabeto de símbolos del codificador de entropía 26 y servir como códigos escape por sí mismos. Esto es, el alfabeto de símbolos del codificador de entropía 26 abarcaría todos los valores del intervalo 68 más los valores inmediatamente vecinos debajo y sobre dicho intervalo 68, como se indica con el paréntesis curvo 74, y el control 60 simplemente reduciría el valor a codificarse por entropía hasta el valor de alfabeto más alto 76 que es vecino inmediato del término superior 72 del intervalo 68, en el caso de que el valor residual r sea superior al término superior 72 del intervalo 68, y avanzaría el valor de alfabeto menor 78 hasta el codificador de entropía 26, para lindar inmediatamente con el término inferior 70 del intervalo 68, en el caso de que el residual de predicción inicial r sea menor que el término menor 70 del intervalo 68.

Mediante el uso de la realización recién señalada, el valor codificado por entropía r corresponde, es decir equivale, al residual de predicción real en el caso de que este se encuentre dentro del intervalo 68. Sin embargo, si el valor codificado por entropía r equivale al valor 76, entonces es evidente que el residual de predicción r real del valor de muestra actual x equivale a 76 o algún valor superior a este último, y si el valor residual codificado por entropía r

- equivale al valor 78, entonces el residual de predicción r real equivale a este valor 78 o algún valor debajo de este. Es decir, hay en realidad dos códigos escape 76 y 78 en dicho caso. En el caso de que el valor inicial r caiga fuera del intervalo 68, el control 60 activa el manipulador de codificación de escape 62 para insertar dentro de la corriente de datos en la cual el codificador de entropía 26 emite su corriente de datos codificada por entropía, una codificación
- 5 que permite al decodificador recuperar el residual de predicción real, ya sea de una manera autocontenida, independiente de que el valor codificado por entropía r sea igual al código escape 76 o 78, o dependiente de lo mismo. Por ejemplo, el manipulador de codificación de escape 62 puede escribir en la corriente de datos el residual de predicción r real directamente usando una representación binaria de suficiente longitud de bits, por ejemplo, de longitud 2^{n+1} , que incluye el signo del residual de predicción r real, o meramente el valor absoluto del residual de predicción real r usando una representación binaria de longitud de bits 2^n usando código escape 76 para la señalización del signo más, y el código escape 78, para la señalización del signo menos. Como alternativa, solo se codifica el valor absoluto de la diferencia entre el valor de residual de predicción inicial r y el valor de código escape 76, en el caso de que el residual de predicción inicial exceda el término superior 72, y el valor absoluto de la diferencia entre el residual de predicción r inicial y el valor del código escape 78, en el caso de que el residual de predicción inicial resida por debajo del término inferior 70. Esto se hace, de acuerdo con un ejemplo de
- 10 implementación, usando codificación condicional: en primer lugar, se codifica $\min(|x-\hat{x}|-13; 15)$ en el caso de codificación de escape, usando cuatro bits, y si $\min(|x-\hat{x}|-13; 15)$ equivale a 15, entonces se codifica $|x-\hat{x}|-13-15$, usando otros siete bits.
- 20 De manera evidente, la codificación de escape es menos compleja que la codificación de los residuales de predicción habituales que yacen dentro del intervalo 68. Por ejemplo, no se usa adaptación de contexto. En cambio, la codificación del valor codificado en el caso de escape puede realizarse mediante simplemente la escritura de una representación binaria para un valor tal como $|r|$ o incluso x , directamente. Sin embargo, el intervalo 68 se selecciona preferentemente de modo tal que el procedimiento escape se produce estadísticamente de manera ocasional, y
- 25 meramente representa "valores atípicos" en la estadística de valores de muestra x .

La Fig. 7 muestra una modificación del decodificador de entropía basado en el contexto de la Fig. 4, que corresponde, o está ajustado, al codificador de entropía de la Fig. 5. De manera similar al codificador de entropía de la Fig. 5, el decodificador de entropía basado en el contexto de la Fig. 7 difiere de aquel expuesto en la Fig. 4 en que

30 un control 71 se conecta entre el decodificador de entropía 46, por una parte, y el combinador 48, por otra parte, en el que el decodificador de entropía de la Fig. 7 comprende adicionalmente un manipulador de código escape 73. De manera similar a la Fig. 5, el control 71 realiza una verificación 74 de si el valor decodificado por entropía r que se emite desde el decodificador de entropía 46 se encuentra dentro del intervalo 68 o corresponde a cierto código escape. Si se aplica esta última circunstancia, el manipulador de código escape 73 se activa por el control 71, para extraer de la corriente de datos que también porta la corriente de datos codificados por entropía decodificados por entropía por el decodificador de entropía 46, el código anteriormente mencionado insertado por el manipulador de código escape 62, por ejemplo, por una representación binaria de suficiente longitud de bits que podría indicar el residual de predicción real r de una manera autocontenida independiente del código escape indicado por el valor decodificado por entropía r , o de manera dependiente del código escape real que el valor decodificado por entropía r asume como ya se explicó en relación con la Fig. 6. Por ejemplo, el manipulador de código escape 73 lee una representación binaria de un valor de la corriente de datos, suma esta al valor absoluto del código escape, es decir, el valor absoluto del término superior o inferior, respectivamente, y utiliza como un signo del valor leído el signo del respectivo término, es decir, el signo más, para el término superior, el signo menos, para el término inferior. La codificación condicional también podría usarse. Esto es, si el valor decodificado por entropía r emitido por el

35 decodificador de entropía 46 se encuentra fuera del intervalo 68, el manipulador de código escape 73 podría leer, en primer lugar, por ejemplo, un valor absoluto de p bits de la corriente de datos, y verificar si este es 2^{p-1} . Si no lo es, el valor decodificado por entropía r se actualiza agregando el valor absoluto de p bits al valor decodificado por entropía r , si el código escape fue el término superior 72, y restando el valor absoluto de p bits del valor decodificado por entropía r , si el código escape fue el término inferior 70. Sin embargo, si el valor absoluto de p bits es 2^{p-1} , entonces se lee otro valor absoluto de q bits, desde la corriente de bits, y el valor decodificado por entropía r se actualiza agregando el valor absoluto de q bits más 2^{p-1} al valor decodificado por entropía r si el código escape fue el término superior 72, y restando el valor absoluto de p bits más 2^{p-1} del valor decodificado por entropía r , si el código escape fue el término inferior 70.

55 Sin embargo, la Fig. 7 muestra además otra alternativa. De acuerdo con esta alternativa, el procedimiento de código escape realizado por los manipuladores de código escape 62 y 72 codifica el valor de muestra x completo directamente, de modo tal que en los casos de código escape, el valor estimado \hat{x} es superfluo. Por ejemplo, una representación de 2^n bits puede ser suficiente en dicho caso, e indicar el valor de x .

60 Solo como una medida de precaución, se observa que otra forma de realizar la codificación de escape sería también factible con estas realizaciones alternativas, no decodificando por entropía ninguno de los valores espectrales cuyo residual de predicción excede o se encuentra fuera del intervalo 68. Por ejemplo, para cada elemento de sintaxis, podría transmitirse una bandera que indica si el mismo se codifica usando la codificación de entropía, o si se usa la codificación de escape. En tal caso, para cada valor de muestra, una bandera indicará la forma de codificación

seleccionada.

A continuación, se describe un ejemplo concreto para la implementación de las realizaciones anteriores. En particular, el ejemplo explícito que se expone a continuación ejemplifica la manera de tratar con la no disponibilidad anteriormente mencionada de ciertos valores de muestra previamente codificados/decodificados en la proximidad espectro-temporal. Adicionalmente, se presentan ejemplos específicos para establecer el posible rango de valores 66, el intervalo 68, la función de cuantificación 32, el rango 34 y así sucesivamente. Más adelante, se describirá que el ejemplo concreto puede usarse en relación con IGF. Sin embargo, se observa que la descripción expuesta a continuación puede transferirse con facilidad a otros casos donde la cuadrícula temporal en la cual se disponen los valores de muestra de la envolvente espectral se define, por ejemplo, por otras unidades de tiempo diferentes de las tramas tales como grupos de intervalos QMF, y la resolución espectral se define asimismo por una subagrupación de subbandas en piezas espectro-temporales.

Indicamos con t (tiempo) el número de tramas a través del tiempo, y f (frecuencia), la posición del respectivo valor de muestra de la envolvente espectral a través de factores de escala (o grupos de factores de escala). Los valores de muestra se denominan valor de SFE, a continuación. Queremos codificar el valor de x , usando información ya disponible de tramas previamente decodificadas en las posiciones $(t - 1)$, $(t - 2)$, ..., y de la trama actual en la posición (t) en las frecuencias $(f - 1)$, $(f - 2)$, La situación se representa nuevamente en la Fig. 8.

Para una trama independiente, establecemos $t = 0$. Una trama independiente es una trama que se califica por sí misma como un punto de acceso aleatorio para una entidad de decodificación. Por lo tanto, representa un instante de tiempo donde es factible el acceso aleatorio en la decodificación en el lado de la decodificación. En términos del eje espectral 16, las primeras SFE 12 asociadas con la menor frecuencia tendrán $f = 0$. En la Fig. 8, los vecinos en tiempo y frecuencia (disponibles tanto en el codificador como en el decodificador) que se usan para calcular el contexto son como fue el caso en la Fig. 1, a, b, c, d y e.

Tenemos varios casos de acuerdo con si $t = 0$ o $f = 0$. En cada caso y en cada contexto, podemos calcular una estimación adaptativa \hat{x} del valor x , basándose en los vecinos, de la siguiente manera:

$t = 0$ $f = 0$	predicción espectro-temporal $\hat{x} = 0$, codificación adaptativa de contexto $r = x - \hat{x}$ usando binario en bruto 7 bits;
$t = 0$ $f = 1$	predicción espectro-temporal $\hat{x} = b$, codificación adaptativa de contexto $r = x - \hat{x}$ usando contexto se01;
$t = 0$ $f \geq 2$	predicción espectro-temporal $\hat{x} = b$, codificación adaptativa de contexto $r = x - \hat{x}$ usando contexto se02[Q(b - e)];
$t = 1$ $f = 0$	predicción espectro-temporal $\hat{x} = a$, codificación adaptativa de contexto $r = x - \hat{x}$ usando contexto se10;
$t \geq 2$ $f = 0$	predicción espectro-temporal $\hat{x} = a$, codificación adaptativa de contexto $r = x - \hat{x}$ usando contexto se02[Q(a - d)];
$t \geq 0$ $f \geq 0$	predicción espectro-temporal $\hat{x} = r \text{INT}(a[q(b-e)][q(a-c)]a + \beta[q(b-e)][q(a-c)]b + \gamma[q(b-e)][q(a-c)]c + \delta[q(b-e)][q(a-c)])$, codificación adaptativa de contexto $r = x - \hat{x}$ usando contexto se11[Q(b - e)][Q(a - c)].

Los valores $b - e$ y $a - c$ representan, como ya se indicó con anterioridad, medidas de desviación. Representan la cantidad esperada de ruido de variabilidad a través de la frecuencia cercana al valor a decodificarse/codificarse, en concreto, x . Los valores $b - c$ y $a - d$ representan la cantidad esperada de ruido de variabilidad a través de tiempo cercano a x . Para reducir significativamente el número total de contextos, estos valores pueden cuantificarse no linealmente antes de usarse para la selección del contexto, por ejemplo, como se expone con respecto a la Fig. 3. El contexto indica la confianza del valor estimado \hat{x} , o de manera equivalente, el pico de la distribución de la codificación. Por ejemplo, la función de cuantificación puede ser según se ilustra en la Fig. 3. Puede definirse como $Q(x)=x$, para $|x| \leq 3$, y $Q(x)=3 \text{ sign}(x)$, para $|x| > 3$. Esta función de cuantificación mapea todos los valores enteros a los siete valores $\{-3, -2, -1, 0, 1, 2, 3\}$. Por favor, obsérvese lo siguiente. Al escribir $Q(x)=x$ ya se ha aprovechado el hecho de que la diferencia de dos enteros es un entero en sí mismo. La fórmula podría escribirse como $Q(x)=\text{rInt}(x)$ para coincidir con la descripción más general presentada anteriormente, y la función en la Fig. 3, respectivamente. Sin embargo, si solo se usa para entradas de enteros para la medida de desviación, $Q(x)=x$ es funcionalmente equivalente a $Q(x)=\text{rInt}(x)$, para el entero x , con $|x| \leq 3$.

Los términos se02[.], se20[.], y se11[.][.] en la tabla anterior son vectores/matrices de contexto. Es decir, cada una de las entradas de estos vectores/matrices son/representan un índice de contexto que indexa uno de los contextos disponibles. Cada uno de estos tres vectores/matrices puede indexar un contexto de un conjunto disjunto de

contextos. Es decir, diferentes conjuntos de contextos pueden elegirse por el determinador de contexto señalado anteriormente, de acuerdo con la condición de disponibilidad. La tabla anterior distingue ejemplarmente entre seis diferentes condiciones de disponibilidad. El contexto correspondiente a se01 y se10 puede corresponder a contextos diferentes de cualquier contexto de los grupos de contextos indexados por se02, se20 y se11, también. El valor estimado de x se calcula como $\hat{x} = \text{rINT}(\alpha a + \beta b + \gamma c + \delta)$. Para tasas de bits superiores, pueden usarse $\alpha=1$, $\beta=-1$, $\gamma=1$, y $\delta=0$, y para tasas de bits menores, puede usarse un conjunto separado de coeficientes para cada contexto, basándose en la información de un conjunto de datos de entrenamiento.

El error de predicción o residual de predicción $r = x - \hat{x}$ puede codificarse usando una distribución separada para cada contexto, derivada usando información extraída de un conjunto de datos de entrenamiento representativo. Pueden usarse dos símbolos especiales en ambos lados de la distribución de codificación 74, en concreto, 76 y 78, para indicar valores positivos o negativos grandes fuera de rango, que se codifican entonces usando una técnica de codificación de escape como ya se indicó con anterioridad. Por ejemplo, de acuerdo con un ejemplo de implementación, se codifica $\min(|x-\hat{x}|-13; 15)$ en el caso de codificación de escape, usando cuatro bits, y si $\min(|x-\hat{x}|-13; 15)$ equivale a 15, entonces se codifica $|x-\hat{x}|-13-15$ usando otros siete bits.

Con respecto a las siguientes figuras, se describen diversas posibilidades en términos de la manera en que pueden crearse los codificadores/decodificadores de entropía anteriormente mencionados basados en el contexto, en respectivos decodificadores/codificadores de audio. La Fig. 9 muestra, por ejemplo, un decodificador paramétrico 80 en el cual podría crearse ventajosamente un decodificador de entropía basado en el contexto 40 de acuerdo con cualquiera de las realizaciones señaladas anteriormente. El decodificador paramétrico 80 comprende, además del decodificador de entropía basado en el contexto 40, un determinador de estructura fina 82 y un conformador espectral 84. Opcionalmente, el decodificador paramétrico 80 comprende un transformador inverso 86. El decodificador de entropía basado en el contexto 40 recibe, como ha señalado con anterioridad, una corriente de datos codificada por entropía 88, codificada de acuerdo con cualquiera de las realizaciones señaladas anteriormente de un codificador de entropía basado en el contexto. La corriente de datos 88, en consecuencia, tiene una envolvente espectral codificada en la misma. El decodificador de entropía basado en el contexto 40 decodifica, de la manera señalada con anterioridad, los valores de muestra de la envolvente espectral de la señal de audio que el decodificador paramétrico 80 busca reconstruir. El determinador de estructura fina 82 está configurado para determinar una estructura fina de un espectrograma de esta señal de audio. Para este fin, el determinador de estructura fina 82 puede recibir información desde fuera, tal como otra porción de una corriente de datos que además comprende la corriente de datos 88. Se describen a continuación alternativas adicionales. En otra alternativa, sin embargo, el determinador de estructura fina 82 puede determinar la estructura fina por sí mismo, usando un proceso aleatorio o pseudoaleatorio. El conformador espectral 84, a su vez, está configurado para conformar la estructura fina de acuerdo con la envolvente espectral definida por los valores espectrales decodificados por el decodificador de entropía basado en el contexto 40. En otras palabras, las entradas del conformador espectral 84 están conectadas a salidas del decodificador de entropía basado en el contexto 40 y el determinador de estructura fina 82, respectivamente, para recibir, de la misma envolvente espectral, por una parte, y la estructura fina del espectrograma de la señal de audio, por otra parte, y el conformador espectral 84 emite, en su salida, la estructura fina del espectrograma conformado de acuerdo con la envolvente espectral. El transformador inverso 86 puede realizar una transformada inversa sobre la estructura fina conformada, de modo tal que emite a una reconstrucción de la señal de audio como su salida.

En particular, el determinador de estructura fina 82 podría estar configurado para determinar la estructura fina del espectrograma usando al menos una de generación de ruido aleatorio artificial, regeneración espectral y decodificación a modo de línea espectral usando predicción espectral y/o entropía espectral-derivación de contexto. Las primeras dos posibilidades se describen con respecto a la Fig. 10. La Fig. 10 ilustra la posibilidad de que la envolvente espectral 10 decodificada por el decodificador de entropía basado en el contexto 40 pertenezca a un intervalo de frecuencia 18 que forma una extensión de frecuencia más alta de un intervalo de frecuencia inferior 90, es decir, el intervalo 18 extiende el intervalo de frecuencia inferior 90 hacia frecuencias superiores, es decir, el intervalo 18 bordea el intervalo 19 en el lado de la frecuencia más alta de este último. En consecuencia, la Fig. 10 muestra la posibilidad de que la señal de audio a reproducirse por el decodificador paramétrico 80 realmente cubre un intervalo de frecuencia 92 del cual el intervalo 18 meramente representa una porción de alta frecuencia del intervalo de frecuencia general 92. Como se muestra en la Fig. 9, el decodificador paramétrico 80, por ejemplo, comprende adicionalmente un decodificador de baja frecuencia 94 configurado para decodificar una corriente de datos de baja frecuencia 96 que acompaña la corriente de datos 88, para obtener la versión de banda de baja frecuencia de la señal de audio en su salida. El espectrograma de esta versión de baja frecuencia se representa en la Fig. 10 usando el signo de referencia 98. En conjunto, esta versión de frecuencia 98 de la señal de audio y la estructura fina formada dentro del intervalo 18 dan como resultado la reconstrucción de las señales de audio del intervalo de frecuencia completo 92, es decir, de su espectrograma a través del intervalo de frecuencia completo 92. Como se indica mediante las líneas de puntos en la Fig. 9, el transformador inverso 86 podría realizar la transformada inversa sobre el intervalo completo 92. En esta estructura, el determinador de estructura fina 82 podría recibir la versión de baja frecuencia 98 del decodificador 94 en el dominio de tiempo o dominio de frecuencia. En el primer caso, el determinador de estructura fina 82 podría someter la versión de baja frecuencia recibida a una

transformación a dominio espectral para obtener el espectrograma 98, y obtener la estructura fina a conformarse por el conformador espectral 84 de acuerdo con la envolvente espectral provista por el decodificador de entropía basado en el contexto 40 usando la regeneración espectral como se ilustra usando la flecha 100. Sin embargo, como ya se indicó con anterioridad, el determinador de estructura fina 82 puede incluso no recibir la versión de baja frecuencia de la señal de audio del decodificador LF 94, y generar la estructura fina solamente usando un proceso aleatorio o pseudoaleatorio.

Un correspondiente codificador paramétrico que se ajusta al decodificador paramétrico de acuerdo con las Figs. 9 y 10 se representa en la Fig. 11. El codificador paramétrico de la Fig. 11 comprende un filtro de cruce de frecuencia 110 que recibe una señal de audio 112 a codificarse, un codificador de banda de alta frecuencia 114 y un codificador de banda de baja frecuencia 116. El filtro de cruce de frecuencia 110 descompone la señal de audio entrante 112 en dos componentes, en concreto, una primera señal 118 que corresponde a una versión filtrada paso alto de una señal de audio entrante 112, y una señal de baja frecuencia 120 que corresponde a una versión filtrada paso bajo de la señal de audio entrante 112, donde las bandas de frecuencias cubiertas por las señales de alta frecuencia y de baja frecuencia 118 y 120 se bordean entre sí a cierta frecuencia de cruce de filtro (compárese con 122 en la Fig. 10). El codificador de banda de baja frecuencia 116 recibe la señal de baja frecuencia 120 y la codifica en una corriente de datos de baja frecuencia, en concreto, 96, y el codificador de banda de alta frecuencia 114 calcula los valores de muestra que describen la envolvente espectral de la señal de alta frecuencia 118 dentro del intervalo de alta frecuencia 18. El codificador de banda de alta frecuencia 114 además comprende el codificador de entropía basado en el contexto anteriormente descrito, para la codificación de estos valores de muestra de la envolvente espectral. El codificador de banda de baja frecuencia 116, por ejemplo, puede ser un codificador de transformada, y la resolución espectro-temporal a la cual el codificador de banda de baja frecuencia 116 codifica la transformada o el espectrograma de la señal de baja frecuencia 120 puede ser mayor que la resolución espectro-temporal a la cual los valores de muestra 12 resuelven la envolvente espectral de la señal de alta frecuencia 118. En consecuencia, el codificador de banda de alta frecuencia 114 emite, entre otros, la corriente de datos 88. Como se muestra mediante la línea de puntos 124 en la Fig. 11, el codificador de banda de baja frecuencia 116 puede emitir información hacia el codificador de banda de alta frecuencia 114, por ejemplo, para controlar el codificador de banda de alta frecuencia 114 con respecto a esta generación de valores de muestra que describen la envolvente espectral, o al menos con respecto a la selección de la resolución espectro-temporal a la cual los valores de muestra muestrean la envolvente espectral.

La Fig. 12 muestra otra posibilidad de realización del decodificador paramétrico 80 de la Fig. 9, y en particular, el determinador de estructura fina 82. En particular, de acuerdo con el ejemplo de la Fig. 12, el determinador de estructura fina 82 en sí mismo recibe una corriente de datos y determina, basándose en esta, la estructura fina del espectrograma de las señales de audio usando la decodificación a nivel de línea espectral utilizando la predicción espectral y/o entropía espectral-derivación de contexto. Esto es, el determinador de estructura fina 82 en sí mismo recupera desde una corriente de datos, la estructura fina en forma de un espectrograma compuesto por una secuencia temporal de espectros de una transformada superpuesta, por ejemplo. Sin embargo, en el caso de la Fig. 12, la estructura fina determinada de esta manera por la estructura fina 82 se refiere a un primer intervalo de frecuencia 130 y coincide con el intervalo de frecuencia completo de la señal de audio, es decir, 92.

En el ejemplo de la Fig. 12, el intervalo de frecuencia 18 al cual se refiere la envolvente espectral 10, se superpone completamente con el intervalo 130. En particular, el intervalo 18 forma una porción de alta frecuencia de intervalo 130. Por ejemplo, muchas de las líneas espectrales dentro del espectrograma 132 recuperadas por el determinador de estructura fina 82 y que cubren el intervalo de frecuencia 130 se cuantificarán a cero, en especial, dentro de la porción de alta frecuencia 18. Sin embargo, para reconstruir la señal de audio con alta calidad, incluso, dentro de la porción de alta frecuencia 18, a una tasa de bits razonable, el decodificador paramétrico 80 aprovecha la envolvente espectral 10. Los valores espectrales 12 de la envolvente espectral 10 describen la envolvente espectral de la señal de audio dentro de la porción de alta frecuencia 18 a una resolución temporal espectral que es más basta que la resolución espectro-temporal del espectrograma 132 decodificado por el determinador de estructura fina 82. Por ejemplo, la resolución espectro-temporal de la envolvente espectral 10 es más basta en términos espectrales, es decir, su resolución espectral es más basta que la granularidad de línea espectral de la estructura fina 132. Como se ha descrito con anterioridad, espectralmente, los valores de muestra 12 de la envolvente espectral 10 pueden describir la envolvente espectral 10 en las bandas de frecuencias 134 en las cuales las líneas espectrales del espectrograma 132 se agrupan, para una escala a nivel de banda de factor de escala de los coeficientes de líneas espectrales, por ejemplo.

El conformador espectral 84 entonces podría, usando los valores de muestra 12, rellenar líneas espectrales dentro de grupos de líneas espectrales o piezas espectro-temporales que corresponden a los respectivos valores de muestra 12 usando mecanismos como la regeneración espectral o la generación de ruido artificial, ajustando el nivel o la energía de la estructura fina resultante dentro del respectivo grupo de piezas/factores de escala espectro-temporales de acuerdo con el correspondiente valor de muestra que describe la envolvente espectral. Véase, por ejemplo, Fig. 13. La Fig. 13 muestra a modo de ejemplo un espectro del espectrograma 132 correspondiente a una trama o su instante de tiempo, tal como el instante de tiempo 136 en la Fig. 12. El espectro se indica ejemplarmente

usando el signo de referencia 140. Como se ilustra en la Fig. 13, algunas de sus porciones 142 se cuantifican a cero. La Fig. 13 muestra la porción de alta frecuencia 18 y la subdivisión de las líneas espectrales del espectro 140 en bandas de factores de escala indicadas por los paréntesis. Usando "x" y "b" y "e", la Fig. 13 ilustra de manera ejemplar que tres valores de muestra 12 describen la envolvente espectral dentro de la porción de alta frecuencia 18 en el instante de tiempo 136 - una para cada banda de factor de escala. Dentro de cada banda de factor de escala correspondiente a estos valores de muestra e, b y x, el determinador de estructura fina 82 genera la estructura fina dentro de al menos las porciones cuantificadas a cero 142 del espectro 140, como se ilustra mediante las áreas sombreadas 144, por ejemplo, mediante la regeneración espectral desde la porción de frecuencia inferior 146, del intervalo de frecuencia completo 130, y a continuación, ajustando la energía del espectro resultante mediante la escala de la estructura fina artificial 144 de acuerdo con, o usando, los valores de muestra e, b y x. Es interesante el hecho de que hay porciones no cuantificadas a cero 148 del espectro 140 en el medio o dentro de las bandas de factores de escala de la porción de alta frecuencia 18, y en consecuencia, usando el relleno inteligente de hueco de acuerdo con la Fig. 12, es factible posicionar picos dentro del espectro 140 incluso en la porción de alta frecuencia 18 del intervalo de frecuencia completo 130 a la resolución de línea espectral y a cualquier posición de línea espectral, sin embargo, con la oportunidad de llenar las porciones cuantificadas a cero 142 usando los valores de muestra x, b y e para la conformación de la estructura fina insertada dentro de estas porciones cuantificadas a cero 142.

Finalmente, la Fig. 14 muestra un posible codificador paramétrico para la alimentación del decodificador paramétrico de la Fig. 9 cuando se representa de acuerdo con la descripción de las Figs. 12 y 13. En particular, en dicho caso, el codificador paramétrico puede comprender un transformador 150 configurado para descomponer espectralmente una señal de audio entrante 152 en el espectrograma completo que cubre el intervalo de frecuencia completo 130. Puede usarse una transformada superpuesta con longitud de transformada posiblemente variable. Un codificador de línea espectral 154 codifica, a resolución de línea espectral, este espectrograma. Para este fin, el codificador de línea espectral 154 recibe tanto la porción de alta frecuencia 18 al igual que el resto de la porción de baja frecuencia desde el transformador 150, donde ambas porciones sin espacios y sin superposición cubren el intervalo de frecuencia completo 130. Un codificador de alta frecuencia paramétrico 156 meramente recibe la porción de alta frecuencia 18 del espectrograma 132 del transformador 150, y genera al menos la corriente de datos 88, es decir, los valores de muestra que describen la envolvente espectral dentro de la porción de alta frecuencia 18.

Esto es, de acuerdo con las realizaciones de las Figs. 12 a 14, el espectrograma de la señal de audio 132 se codifica en una corriente de datos 158 por el codificador de línea espectral 154. En consecuencia, el codificador de línea espectral 154 puede codificar un valor de línea espectral por línea espectral del intervalo completo 130, por instante de tiempo o trama 136. Las cajas pequeñas 160 en la Fig. 12 muestran estos valores de líneas espectrales. A lo largo del eje espectral 16, las líneas espectrales pueden agruparse en bandas de factores de escala. En otras palabras, el intervalo de frecuencia 16 puede subdividirse en bandas de factores de escala compuestas por grupos de líneas espectrales. El codificador de línea espectral 154 puede seleccionar un factor de escala para cada banda de factor de escala dentro de cada instante de tiempo, para escalar los valores de líneas espectrales cuantificados 160 codificados por medio de la corriente de datos 158. A una resolución espectro-temporal que es al menos más basta que la cuadrícula espectro-temporal definida por las instancias de tiempo y las líneas espectrales en las cuales los valores de líneas espectrales 160 se disponen regularmente, y que puede coincidir con el mapa de bits definido por la resolución de factor de escala, el codificador de alta frecuencia paramétrico 156 describe la envolvente espectral dentro de la porción de alta frecuencia 18. Es interesante el hecho de que los valores de líneas espectrales no cuantificados a cero 160, escalados de acuerdo con el factor de escala de la banda de factor de escala donde caen, pueden estar intercalados, a resolución de línea espectral, en cualquier posición dentro de la porción de alta frecuencia 18, y en consecuencia, sobreviven a la síntesis de alta frecuencia en el lado de la decodificación dentro del conformador espectral 84 usando los valores de muestra que describen la envolvente espectral dentro de la porción de alta frecuencia, a medida que el determinador de estructura fina 82 y el conformador espectral 84 restringen, por ejemplo, su síntesis de estructura fina y conformación a las porciones cuantificadas a cero 142 dentro de la porción de alta frecuencia 18 del espectrograma 132. En conjunto, se logra un compromiso muy eficiente entre la tasa de bits gastada, por una parte, y la calidad obtenible, por otra parte.

Como se indica por medio de la flecha de línea de puntos en la Fig. 14, indicado en 164, el codificador de línea espectral 154 puede informar al codificador de alta frecuencia paramétrico 156 sobre, por ejemplo, la versión reconstruible de espectrograma 132 como reconstruible a partir de la corriente de datos 158, con un codificador de alta frecuencia paramétrico 156 usando esta información, por ejemplo, para controlar la generación de los valores de muestra 12 y/o la resolución espectro-temporal de la representación de la envolvente espectral 10 por los valores de muestra 12.

Resumiendo lo anterior, las realizaciones anteriores se aprovechan de las propiedades especiales de valores de muestra de envolventes espectrales, donde, en contraste a [2] y [3], dichos valores de muestra representan valores promedio de líneas espectrales. En todas las realizaciones señaladas anteriormente, las transformadas pueden utilizar MDCT, y en consecuencia, puede usarse una MDCT inversa para todas las transformadas inversas. En cualquier caso, dichos valores de muestra de envolventes espectrales son mucho más "suaves" y están linealmente

correlacionados con la magnitud promedio de las correspondientes líneas espectrales complejas. Además, de acuerdo con al menos algunas de las realizaciones anteriores, los valores de muestra de la envolvente espectral, denominados valores de SFE a continuación, son, de hecho, del dominio dB o, más generalmente, dominio logarítmico, que es una representación logarítmica. Esto mejora adicionalmente la "suavidad" en comparación con los valores en el dominio lineal o el dominio de ley de la potencia para las líneas espectrales. Por ejemplo, en AAC, el exponente de la ley de potencias es 0,75. En contraste a [4], en al menos algunas realizaciones, los valores de muestra de envolvente espectral están en dominio logarítmico, y las propiedades y la estructura de las distribuciones de codificación son significativamente diferentes (de acuerdo con su magnitud, un valor de dominio logarítmico típicamente mapea a un número exponencialmente creciente de valores de dominio lineal). Por lo tanto, al menos algunas de las realizaciones anteriormente descritas se aprovechan de la representación logarítmica en la cuantificación del contexto (se presenta típicamente un número menor de contextos) y en la codificación de las colas de la distribución de cada contexto (las colas de cada distribución son más amplias). En contraste a [2], algunas de las realizaciones anteriores adicionalmente utilizan una predicción lineal fijada o adaptativa en cada contexto, basándose en los mismos datos que los utilizados en el cálculo del contexto cuantificado. Este enfoque es útil al reducir drásticamente el número de contextos, mientras que aún se obtiene óptimo desempeño. En contraste, por ejemplo, a [4], en al menos algunas de las realizaciones, la predicción lineal en el dominio logarítmico tiene un uso y una significación significativamente diferentes. Por ejemplo, permite la predicción perfecta de áreas de espectro de energía constantes, y además, tanto áreas de espectro de desvanecimiento descendente como de desvanecimiento ascendente de la señal. En contraste a [4], algunas de las realizaciones anteriormente descritas utilizan la codificación aritmética que permite la óptima codificación de distribuciones arbitrarias usando información extraída de un conjunto de datos de entrenamiento representativo. En contraste a [2], que también utiliza la codificación aritmética, de acuerdo con las realizaciones anteriores, se codifican los valores de error de predicción en lugar de los valores originales. Aún más, en las realizaciones anteriores, no es necesario utilizar la codificación de mapa de bits. La codificación de mapa de bits, sin embargo, requiere varias etapas de codificación aritmética para cada valor entero. En comparación, de acuerdo con realizaciones anteriores, cada valor de muestra de la envolvente espectral podría codificarse/decodificarse dentro de una etapa que incluye, como se señala anteriormente, el uso opcional de la codificación de escape para valores fuera del centro de la distribución entera de valores de muestra, lo que es mucho más rápido.

Resumiendo brevemente la realización de un decodificador de parámetros que soporta IGF nuevamente, como se ha descrito con anterioridad con respecto a las Figs. 9, 12 y 13, de acuerdo con esta realización, el determinador de estructura fina 82 está configurado para usar la decodificación a nivel de línea espectral usando la predicción espectral y/o entropía espectral-derivación de contexto para derivar la estructura fina 132 del espectrograma de la señal de audio dentro de un primer intervalo de frecuencia 130, en concreto, el intervalo de frecuencia completo. La decodificación a nivel de línea de frecuencia indica el hecho de que el determinador de estructura fina 82 recibe valores de líneas espectrales 160 desde una corriente de datos dispuesta, espectralmente, en tono de línea espectral, para formar un espectro 136 por instante de tiempo correspondiente a una respectiva porción de tiempo. El uso de la predicción espectral, por ejemplo, podría involucrar la codificación diferencial de estos valores de líneas espectrales a lo largo del eje espectral 16, es decir, meramente la diferencia al valor de línea espectral que precede espectralmente de inmediato se decodifica a partir de la corriente de datos y a continuación se agrega a este predecesor. La entropía espectral-derivación de contexto podría indicar el hecho de que el contexto para la decodificación de entropía de un respectivo valor de línea espectral 160 podría depender, es decir, podría seleccionarse de manera aditiva, basándose en los valores ya decodificados de líneas espectrales en la proximidad espectro-temporal, o al menos la proximidad espectral, del valor de línea espectral actualmente decodificado 160. Para rellenar porciones cuantificadas a cero 142 de la estructura fina, el determinador de estructura fina 82 puede utilizar la generación de ruido aleatorio artificial y/o la regeneración espectral. El determinador de estructura fina 82 realiza esta acción meramente dentro de un segundo intervalo de frecuencia 18 que, por ejemplo, puede restringirse a una porción de alta frecuencia del intervalo de frecuencia general 130. Las porciones espectralmente regeneradas pueden tomarse, por ejemplo, de la porción de frecuencia restante 146. El conformador espectral entonces realiza la conformación de la estructura fina así obtenida de acuerdo con la envolvente espectral descrita por los valores de muestra 12 en las porciones cuantificadas en cero. Notablemente, la contribución de las porciones no cuantificadas a cero de la estructura fina dentro del intervalo 18 al resultado de la estructura fina después de la conformación es independiente de la envolvente espectral real 10. Esto significa lo siguiente: o la generación de ruido aleatorio artificial y/o la regeneración espectral, es decir, el relleno, se restringen a las porciones cuantificadas a cero 142 completamente, de manera que en el espectro de estructura fina final meramente las porciones 142 se han rellenado por la generación de ruido aleatorio artificial y/o la regeneración espectral usando la conformación de envolvente espectral, con las contribuciones distintas de cero 148 quedando como están, intercaladas entre las porciones 142, o como alternativa, el resultado de la generación de ruido aleatorio artificial y/o la regeneración espectral, en concreto, la respectiva estructura fina sintetizada se dispone además, de manera aditiva, sobre las porciones 148, y a continuación, la formación de la resultante estructura fina sintetizada de acuerdo con la envolvente espectral 10. Sin embargo, aun en dicho caso, se mantiene la contribución a modo de las porciones no cuantificadas a cero 148 de la estructura fina originalmente decodificada.

Con respecto a la realización de las Figs. 12 a 14, se observa finalmente que el procedimiento o concepto de IGF

que se describe con respecto a estas figuras mejora significativamente la calidad de una señal codificada, incluso, a muy bajas tasas de bits, donde una parte significativa del espectro en la región de alta frecuencia 18 se cuantifica a cero debido, típicamente, al insuficiente presupuesto de bits. Para preservar lo más posible la estructura fina de la región de frecuencia superior 18, la información de IGF, se usa la región de baja frecuencia como una fuente para el reemplazo adaptativo de las regiones de destino de la región de alta frecuencia que fueron más cuantificadas a cero, es decir, las regiones 142. Un requisito importante para lograr una buena calidad perceptual es la coincidencia de la envolvente de energía decodificada de los coeficientes espectrales con aquella de la señal original. Para lograr este objetivo, se calculan las energías espectrales promedio sobre coeficientes espectrales de una o más bandas de factor de escala AAC consecutivas. Los valores resultantes son los valores de muestra 12 que describen la envolvente espectral. El cálculo de los promedios usando límites definidos por bandas de factores de escala se ve motivado por el ya existente cuidadoso ajuste de dichos límites a fracciones de las bandas críticas, que son características de la audición humana. Las energías promedio pueden convertirse, como se ha descrito anteriormente, en una representación logarítmica, tal como una representación de escala de dB usando una fórmula que, por ejemplo, puede ser similar a la ya conocida de los factores de escala AAC, y a continuación, pueden cuantificarse uniformemente. En IGF, puede usarse de manera opcional una diferente exactitud de cuantificación de acuerdo con la tasa de bits total requerida. Las energías promedio constituyen una parte significativa de la información generada por el IGF, de modo que su eficiente representación dentro de la corriente de datos 88 es muy importante para el desempeño general del concepto de IGF.

Si bien se han descrito algunos aspectos en el contexto de un aparato, es evidente que estos aspectos además representan una descripción del correspondiente método, donde un bloque o dispositivo corresponde a una etapa de método o un rasgo de una etapa de método. De manera análoga, los aspectos descritos en el contexto de una etapa de método además representan una descripción de un correspondiente bloque o elemento o rasgo de un correspondiente aparato. Algunas o la totalidad de las etapas de método pueden ejecutarse por (o usando) un aparato de hardware, por ejemplo, un microprocesador, una computadora programable o un circuito electrónico. En algunas realizaciones, una o más de las etapas de método más importantes pueden ejecutarse por dicho aparato.

De acuerdo con 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 disquete, un disco duro, un DVD (disco versátil digital), un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control electrónicamente legibles almacenadas en la misma, que cooperan (o son capaces de cooperar) con un sistema de computadora programable de manera tal para llevar a cabo el respectivo método. Por lo tanto, el medio de almacenamiento digital puede ser legible por computadora.

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

En general, las realizaciones de la presente invención pueden implementarse como un producto de programa de computadora con un código de programa, donde el código de programa es operativo para llevar a cabo uno de los métodos cuando se ejecuta el producto de programa de computadora en una computadora. El código de programa puede almacenarse, por ejemplo, en un portador legible en una máquina.

Otras realizaciones comprenden el programa de computadora para llevar a cabo uno de los métodos que se describen en el presente documento, almacenado en un portador legible en una máquina.

En otras palabras, una realización del método de la invención, por lo tanto, es un programa de computadora que tiene un código de programa para la realización de uno de los métodos que se describen en el presente documento, cuando se ejecuta el programa de computadora en una computadora.

Una realización adicional del método de la invención es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio legible por computadora) que comprende, grabado en el mismo, el programa de computadora para llevar a cabo uno de los métodos que se describen en el presente documento. El portador de datos, el medio de almacenamiento digital o el medio grabado son típicamente tangibles y/o no transitorios.

Una realización adicional del método de la invención es, por lo tanto, una corriente de datos o una secuencia de señales que representan el programa de computadora para llevar a cabo uno de los métodos que se describen en el presente documento. La corriente de datos o la secuencia de señales, por ejemplo, pueden estar configuradas para transferirse por medio de una conexión de comunicación de datos, por ejemplo, por medio de Internet.

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

Una realización adicional comprende una computadora que tiene instalado en la misma el programa de computadora para llevar a cabo uno de los métodos que se describen en el presente documento.

5 Una realización adicional de acuerdo con la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa de computadora para la realización de uno de los métodos descritos en el presente documento, a un receptor. El receptor puede ser, por ejemplo, una computadora, un dispositivo móvil, un dispositivo de memoria o similar. El aparato o el sistema pueden comprender, por ejemplo, un servidor de archivos para transferir el programa de computadora al receptor.

10 En algunas realizaciones, puede usarse un dispositivo lógico programable (por ejemplo, un campo de matrices de puertas programables) para realizar algunas o la totalidad de las funcionalidades de los métodos que se describen en el presente documento. En algunas realizaciones, un campo de matrices de puertas programables puede cooperar con un microprocesador para llevar a cabo uno de los métodos que se describen en el presente documento. En general, los métodos se llevan a cabo, preferentemente, por cualquier aparato de hardware.

15 Las realizaciones anteriormente descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles que se describen en el presente documento serán evidentes para los expertos en la materia. Por lo tanto, se tiene la intención de limitación solo por el alcance de las reivindicaciones de patente inminentes, y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones del presente documento.

Referencias.

25 [1] International Standard ISO/IEC 14496-3: 2005, Information technology - Coding of audio-visual objects - Part 3: Audio, 2005.

[2] International Standard ISO/IEC 23003-3:2012, Information technology - MPEG audio technologies - Part 3: Unified Speech and Audio Coding, 2012.

30 [3] B. Edler y N. Meine: Improved Quantization and Lossless Coding for Subband Audio Coding, AES 118th Convention, mayo de 2005.

35 [4] M.J. Weinberger y G. Seroussi: The LOCO-I Lossless Image Compression Algorithm: Principles and Standardization into JPEG-LS, 1999. Disponible en línea en http://www.hpl.hp.com/research/info_theory/loco/HPL-98-193R1.pdf

REIVINDICACIONES

1. Decodificador de entropía basado en el contexto para decodificar valores de muestra (12) de una envolvente espectral (10) de una señal de audio, configurado para:
 - 5 predecir espectro-temporalmente (42) un valor de muestra actual de la envolvente espectral para obtener un valor estimado del valor de muestra actual;
 - determinar (44) un contexto para el valor de muestra actual dependiente de una medida para una desviación entre un primer par de valores de muestra ya decodificados en la proximidad espectro-temporal y una segunda medida para una desviación entre un segundo par de valores de muestra ya decodificados en la proximidad espectro-temporal, el primer par es vecino entre sí espectralmente, y el segundo par es vecino entre sí temporalmente;
 - 10 decodificar por entropía (46) un valor de residual de predicción del valor de muestra actual usando el contexto determinado; y
 - combinar (48) el valor estimado y el valor de residual de predicción para obtener el valor de muestra actual, en el que el decodificador por entropía basado en contexto está configurado adicionalmente para usar una diferencia de signo entre el par de valores de muestra ya decodificados de la envolvente espectral en la proximidad espectro-temporal del valor de muestra actual para medir la desviación.
2. Decodificador de entropía basado en el contexto de acuerdo con la reivindicación 1, configurado adicionalmente para realizar la predicción espectro-temporal por predicción lineal.
- 20 3. Decodificador de entropía basado en el contexto de acuerdo con la reivindicación 1 o 2, configurado adicionalmente para predecir espectro-temporalmente el valor de muestra actual de la envolvente espectral combinando linealmente los valores de muestra ya decodificados del primer y segundo pares.
- 25 4. Decodificador de entropía basado en el contexto de acuerdo con la reivindicación 3, configurado adicionalmente para establecer factores de la combinación lineal de modo tal que los factores son los mismos para diferentes contextos, en el caso de que la tasa de bits a la que se codifica la señal de audio sea mayor que un umbral predeterminado, y los factores se establecen individualmente para los diferentes contextos, en el caso de que la tasa de bits sea menor que el umbral predeterminado.
- 30 5. Decodificador de entropía basado en el contexto de acuerdo con cualquiera de las reivindicaciones previas, configurado adicionalmente para, al decodificar los valores de muestra de la envolvente espectral, decodificar secuencialmente los valores de muestra usando un orden de decodificación (30) que atraviesa los valores de muestra instantáneo de tiempo por instante, conduciendo, en cada instante de tiempo, de la menor a la más alta frecuencia.
- 35 6. Decodificador de entropía basado en el contexto de acuerdo con cualquiera de las reivindicaciones previas, configurado adicionalmente para, al determinar el contexto, cuantificar la medida para la desviación y determinar el contexto usando la medida cuantificada.
- 40 7. Decodificador de entropía basado en el contexto de acuerdo con la reivindicación 6, configurado adicionalmente para usar una función de cuantificación (32) en la cuantificación de la medida para la desviación, que es constante para valores de la medida para la desviación fuera de un intervalo predeterminado (34), incluyendo el intervalo predeterminado cero.
- 45 8. Decodificador de entropía basado en el contexto de acuerdo con la reivindicación 7, en el que los valores de la envolvente espectral se representan como números enteros, y la longitud del intervalo predeterminado (34) es menor o igual que $1/16$ del número de estados representables de una representación de enteros de los valores de la envolvente espectral.
- 50 9. Decodificador de entropía basado en el contexto de acuerdo con cualquiera de las reivindicaciones previas, configurado adicionalmente para transferir (50) el valor de muestra actual, derivado por la combinación, de un dominio logarítmico a un dominio lineal.
- 55 10. Decodificador de entropía basado en el contexto de acuerdo con cualquiera de las reivindicaciones previas, gestionando el decodificador de entropía basado en el contexto un número de contextos, teniendo cada contexto una distribución de probabilidad asociada con el mismo que asigna a cada posible valor del valor de residual de predicción una respectiva probabilidad, en el que el decodificador de entropía basado en el contexto está configurado adicionalmente para, al decodificar por entropía los valores de residuales, decodificar secuencialmente los valores de muestra a lo largo de un orden de decodificación y usar un conjunto de distribuciones de probabilidad individuales de contexto, que son constantes durante la decodificación secuencial de los valores de muestra de una envolvente espectral.
- 60 11. Decodificador de entropía basado en el contexto de acuerdo con cualquiera de las reivindicaciones previas,

configurado adicionalmente para, al decodificar por entropía del valor de residual, usar un mecanismo de codificación de escape en el caso de que el valor de residual se encuentre fuera de un rango de valores predeterminado (68).

5 12. Decodificador de entropía basado en el contexto de acuerdo con la reivindicación 11, en el que los valores de muestra de la envolvente espectral se representan como números enteros, y el residual de predicción se representa como un número entero, y los valores absolutos de los términos de intervalo (70, 72) del rango de valores predeterminado son menores o iguales que 1/8 del número de estados representables del valor de residual de predicción.

10 13. Decodificador paramétrico que comprende:
 un decodificador de entropía basado en el contexto (40) para decodificar valores de muestra de una envolvente espectral de una señal de audio de acuerdo con cualquiera de las reivindicaciones previas;
 15 un determinador de estructura fina (82) configurado para recibir valores de línea espectral (160) de una corriente de datos dispuesta espectralmente, en tono de línea espectral para determinar una estructura fina de un espectrograma de la señal de audio; y
 un conformador espectral (84) configurado para conformar la estructura fina de acuerdo con la envolvente espectral.

20 14. Decodificador paramétrico de acuerdo con la reivindicación 13, en el que el determinador de estructura fina está configurado para determinar la estructura fina del espectrograma usando al menos una de
 generación de ruido aleatorio artificial,
 25 regeneración espectral, y
 decodificación a nivel de línea espectral usando la predicción espectral y/o la entropía espectral-derivación de contexto.

30 15. Decodificador paramétrico de acuerdo con la reivindicación 13 o 14, que comprende adicionalmente un decodificador de intervalo de frecuencia inferior (94) configurado para decodificar un intervalo de frecuencia inferior (98) del espectrograma de la señal de audio, en el que el codificador de entropía basado en el contexto, el determinador de estructura fina y el conformador espectral están configurados de manera tal que la conformación de la estructura fina de acuerdo con la envolvente espectral se realiza dentro de una extensión de frecuencia superior espectral (18) del intervalo de frecuencia inferior.

35 16. Decodificador paramétrico de acuerdo con la reivindicación 15, en el que el decodificador de intervalo de frecuencia inferior (94) está configurado para determinar la estructura fina del espectrograma usando decodificación a nivel de línea espectral usando predicción espectral y/o entropía espectral-derivación de contexto o descomposición espectral de una señal de audio de banda de frecuencia baja de dominio de tiempo decodificada.

40 17. Decodificador paramétrico de acuerdo con la reivindicación 13 o 14, en el que el determinador de estructura fina está configurado para usar la decodificación a nivel de línea espectral usando la predicción espectral y/o la entropía espectral-derivación de contexto para derivar la estructura fina del espectrograma de la señal de audio dentro de un primer intervalo de frecuencia (130), localizar porciones cuantificadas a cero (142) de la estructura fina dentro de un segundo intervalo de frecuencia (18) que se superpone con el primer intervalo de frecuencia y aplicar la generación de ruido aleatorio artificial y/o la regeneración espectral sobre las porciones cuantificadas a cero (142), en el que el conformador espectral (84) está configurado para realizar la conformación de la estructura fina de acuerdo con la envolvente espectral en las porciones cuantificadas a cero (142).

50 18. Codificador de entropía basado en el contexto para codificar valores de muestra de una envolvente espectral de una señal de audio, configurado para predecir espectro-temporalmente un valor de muestra actual de la envolvente espectral para obtener un valor estimado del valor de muestra actual;
 determinar un contexto para el valor de muestra actual dependiente de una primera medida para una desviación entre un primer par de valores de muestra ya decodificados en la proximidad espectro-temporal y una segunda
 55 medida para una desviación entre un segundo par de valores de muestra ya decodificados en la proximidad espectro-temporal, el primer par es vecino entre sí espectralmente, y el segundo par es vecino entre sí temporalmente;
 determinar un valor de residual de predicción basándose en una desviación entre el valor estimado y el valor de muestra actual; y
 60 codificar por entropía el valor de residual de predicción del valor de muestra actual usando el contexto determinado, en el que el decodificador por entropía basado en contexto está configurado adicionalmente para usar una diferencia de signo entre el par de valores de muestra ya decodificados de la envolvente espectral en la proximidad espectro-temporal del valor de muestra actual para medir la desviación.

19. Método para, usando la decodificación de entropía basada en el contexto, decodificar valores de muestra de una envolvente espectral de una señal de audio, que comprende:
 predecir espectro-temporalmente un valor de muestra actual de la envolvente espectral para obtener un valor estimado del valor de muestra actual;
- 5 determinar un contexto para el valor de muestra actual dependiente de una primera medida para una desviación entre un primer par de valores de muestra ya decodificados en la proximidad espectro-temporal y una segunda medida para una desviación entre un segundo par de valores de muestra ya decodificados en la proximidad espectro-temporal, el primer par es vecino entre sí espectralmente, y el segundo par es vecino entre sí temporalmente;
- 10 decodificar por entropía un valor de residual de predicción del valor de muestra actual usando el contexto determinado; y
 combinar el valor estimado y el valor de residual de predicción para obtener el valor de muestra actual, en el que una diferencia de signo entre el par de valores de muestra ya decodificados de la envolvente espectral en una proximidad espectro-temporal del valor de muestra actual se usa para medir la desviación.
- 15
20. Método para, usando la codificación de entropía basada en el contexto, codificar valores de muestra de una envolvente espectral de una señal de audio, que comprende:
 predecir espectro-temporalmente un valor de muestra actual de la envolvente espectral para obtener un valor estimado del valor de muestra actual;
- 20 determinar un contexto para el valor de muestra actual dependiente de una medida para una desviación entre un primer par de valores de muestra ya decodificados en la proximidad espectro-temporal y una segunda medida para una desviación entre un segundo par de valores de muestra ya decodificados en la proximidad espectro-temporal, el primer par es vecino entre sí espectralmente, y el segundo par es vecino entre sí temporalmente;
- 25 determinar un valor de residual de predicción basándose en una desviación entre el valor estimado y el valor de muestra actual; y
 codificar por entropía el valor de residual de predicción del valor de muestra actual usando el contexto determinado, en el que una diferencia de signo entre el par de valores de muestra ya decodificados de la envolvente espectral en una proximidad espectro-temporal del valor de muestra actual se usa para medir la desviación.
- 30
21. Programa de computadora que tiene un código de programa para la realización, cuando se ejecuta en una computadora, de un método de acuerdo con la reivindicación 19 o 20.

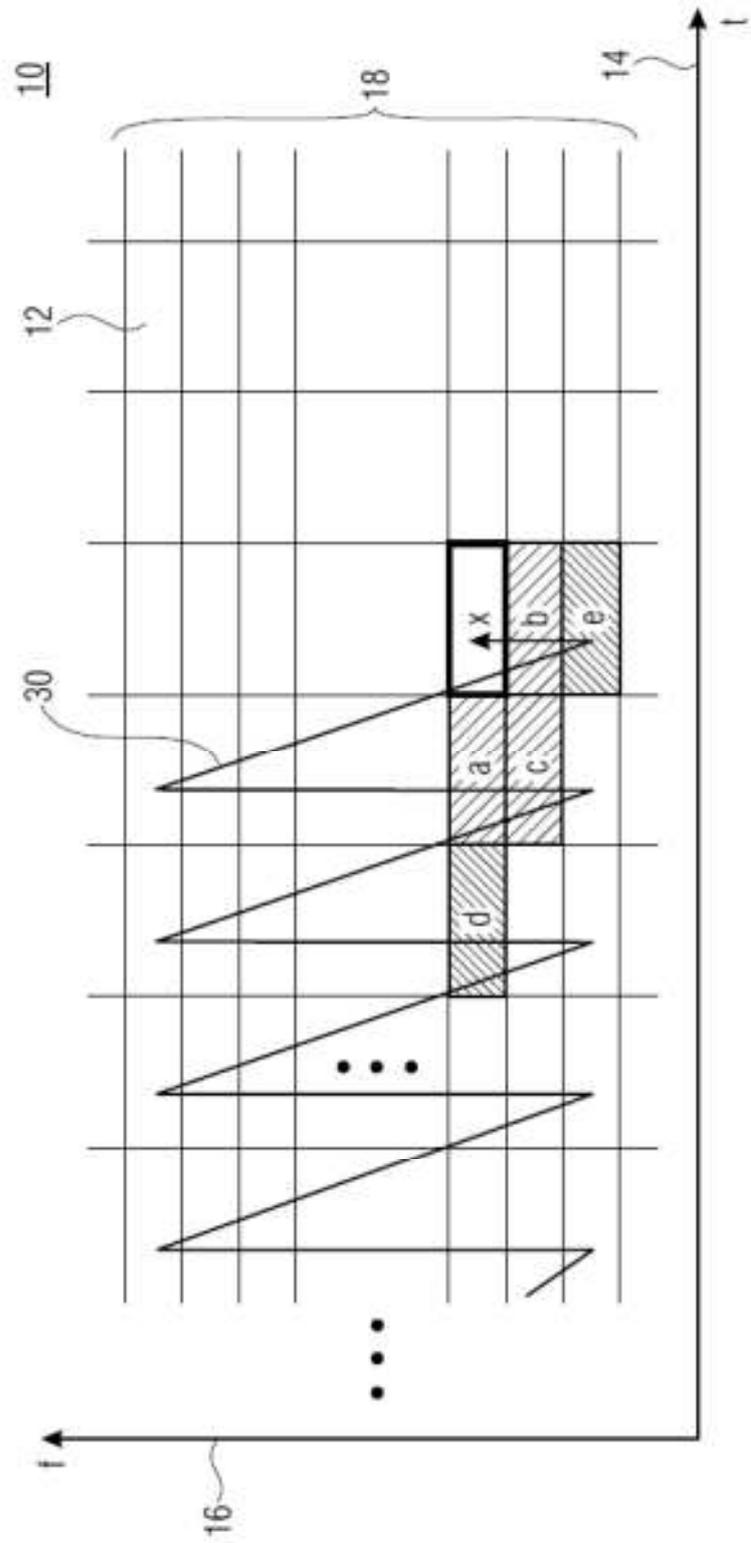


FIG. 1

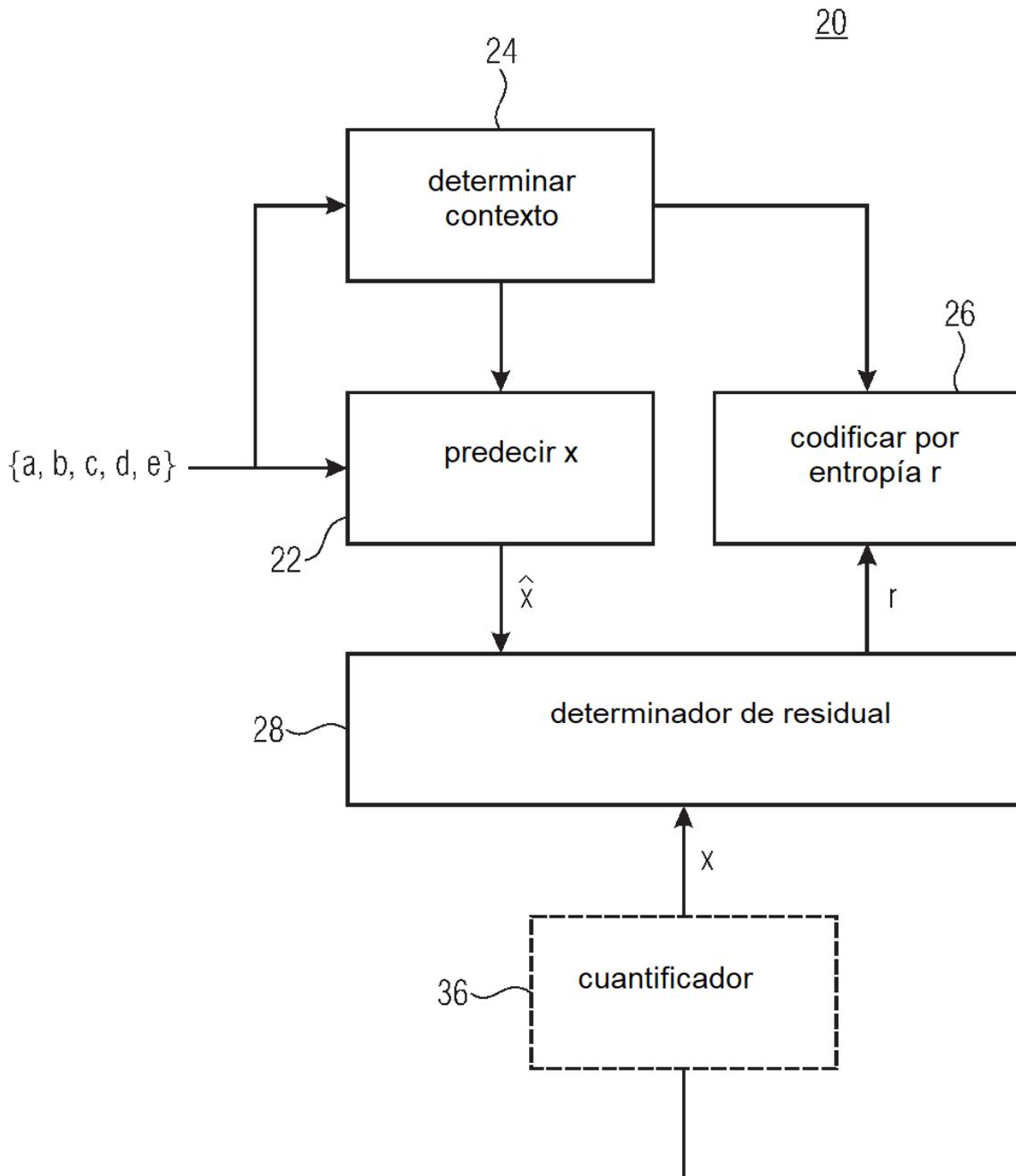


FIG 2

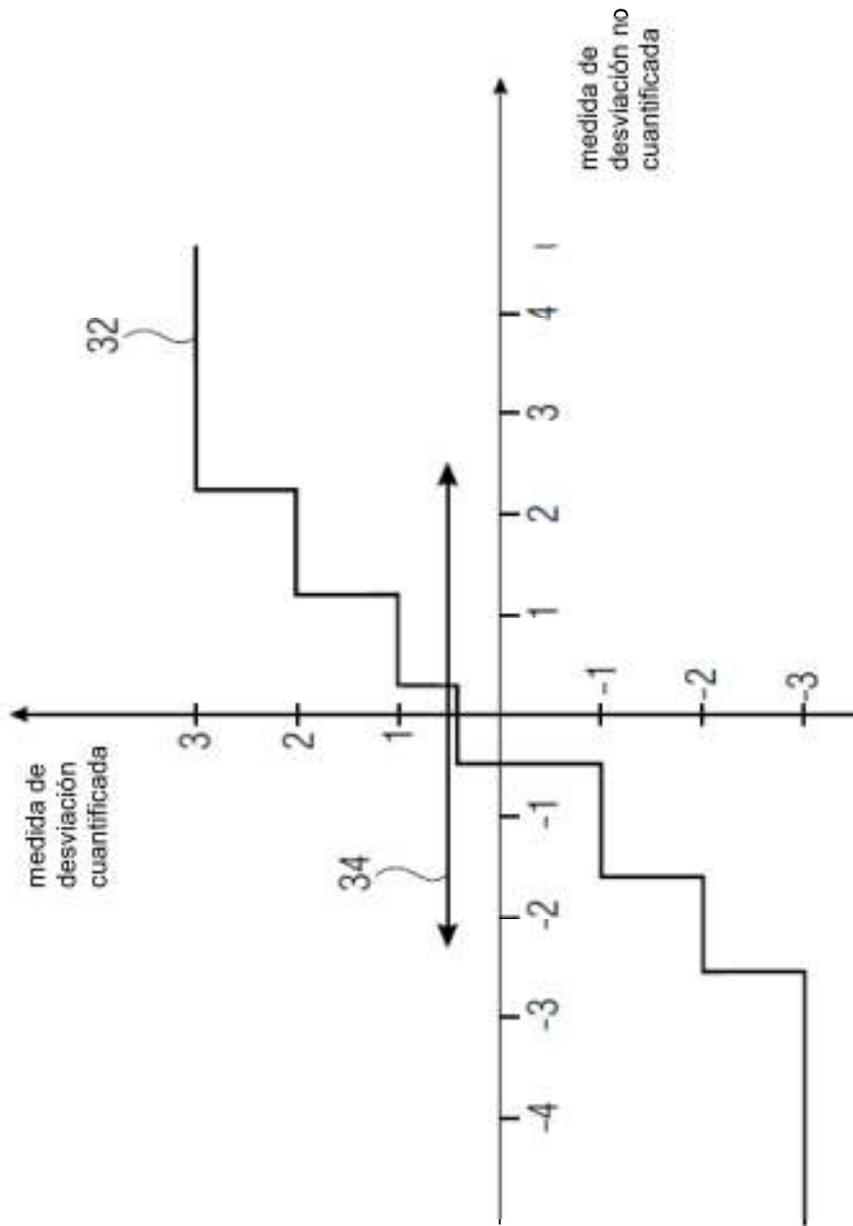


FIG. 3

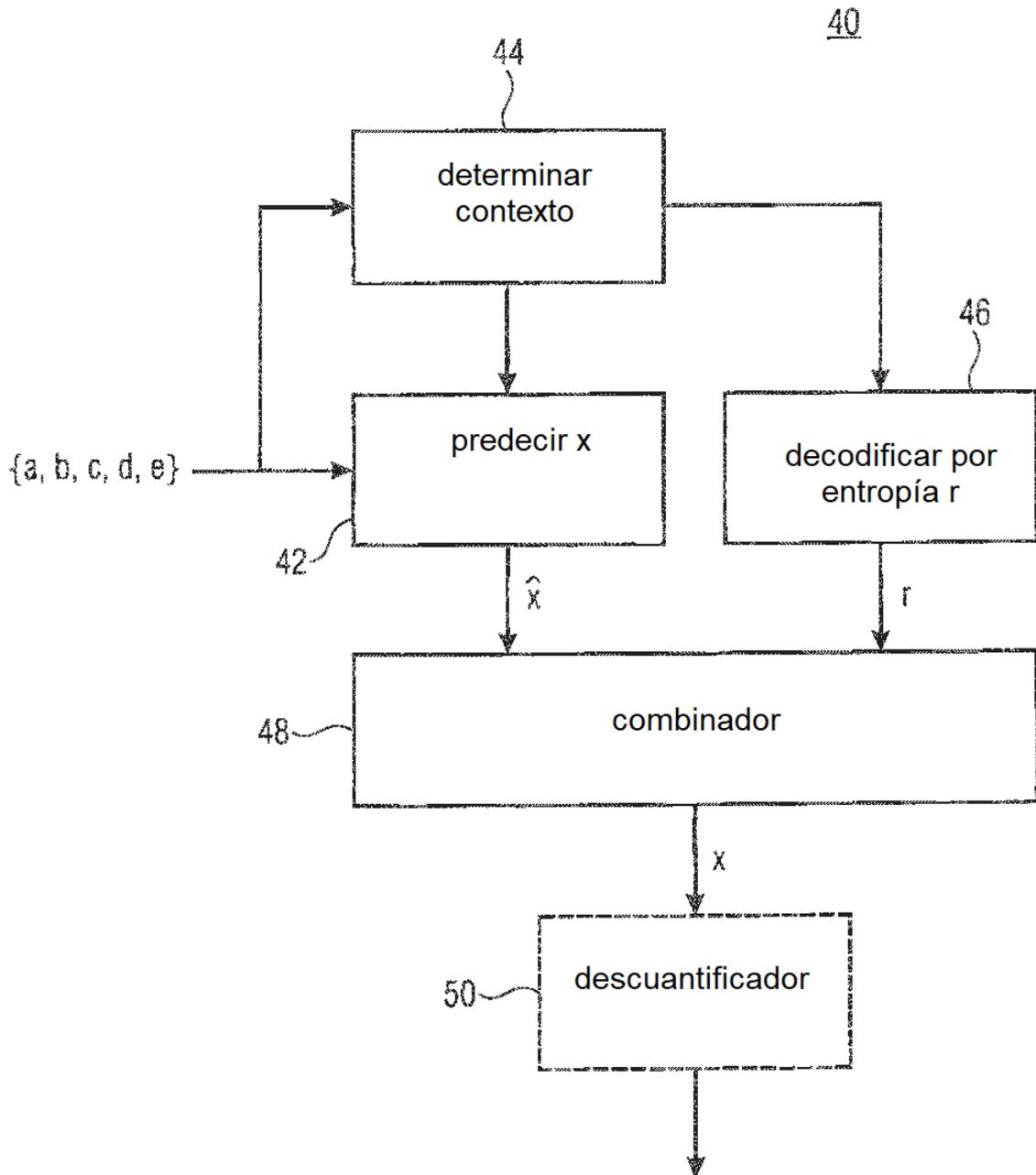


FIG 4

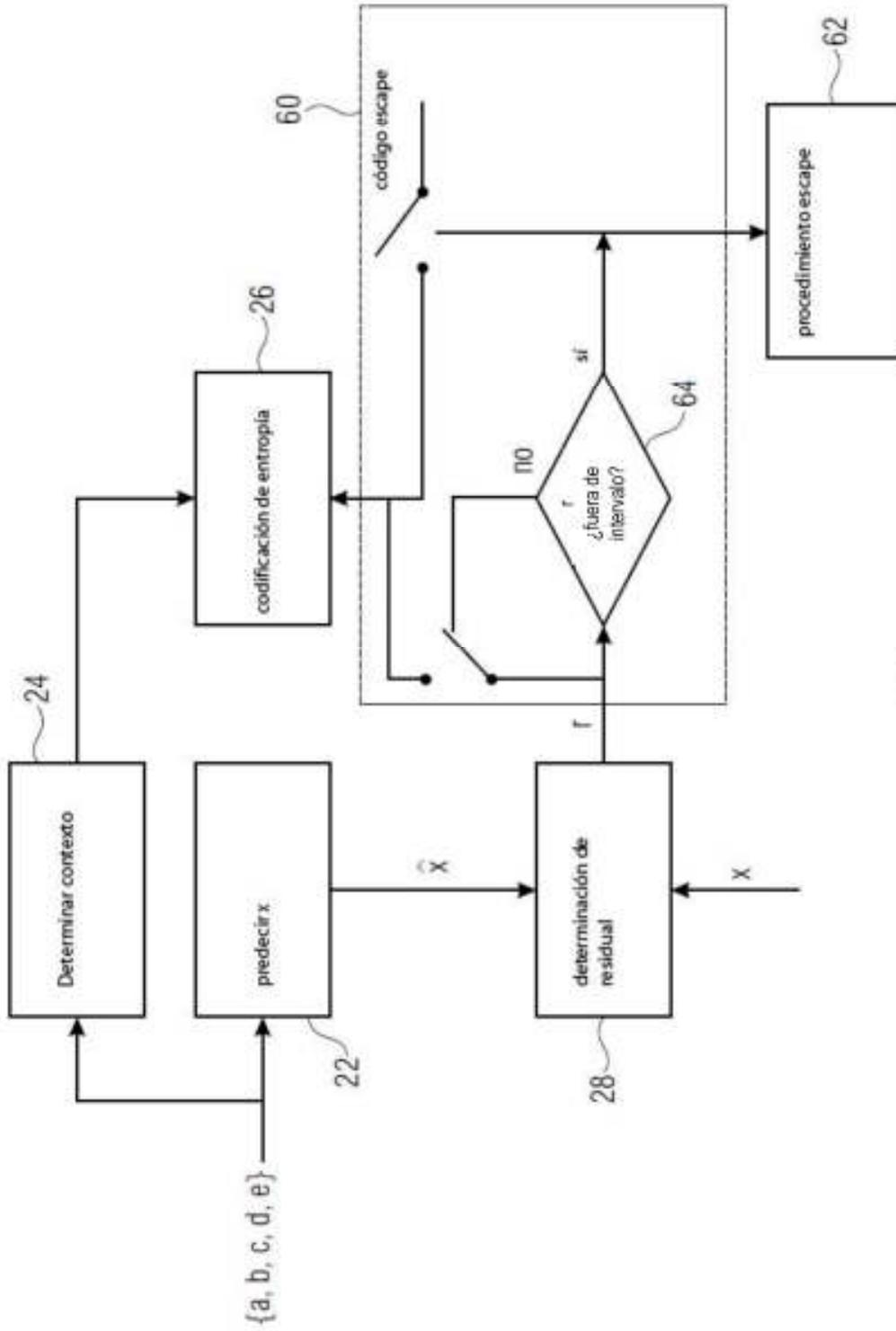


FIG. 5

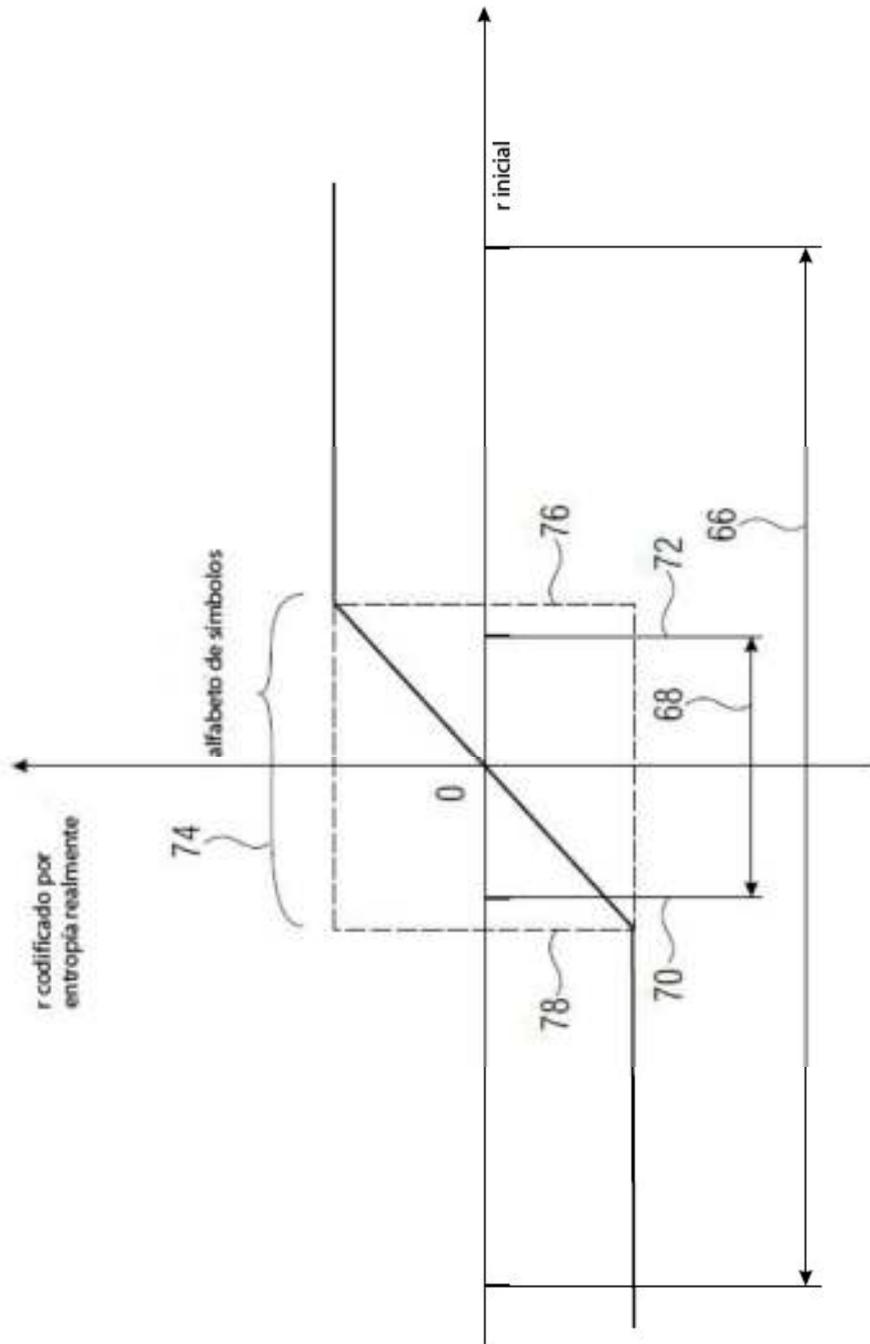


FIG. 6

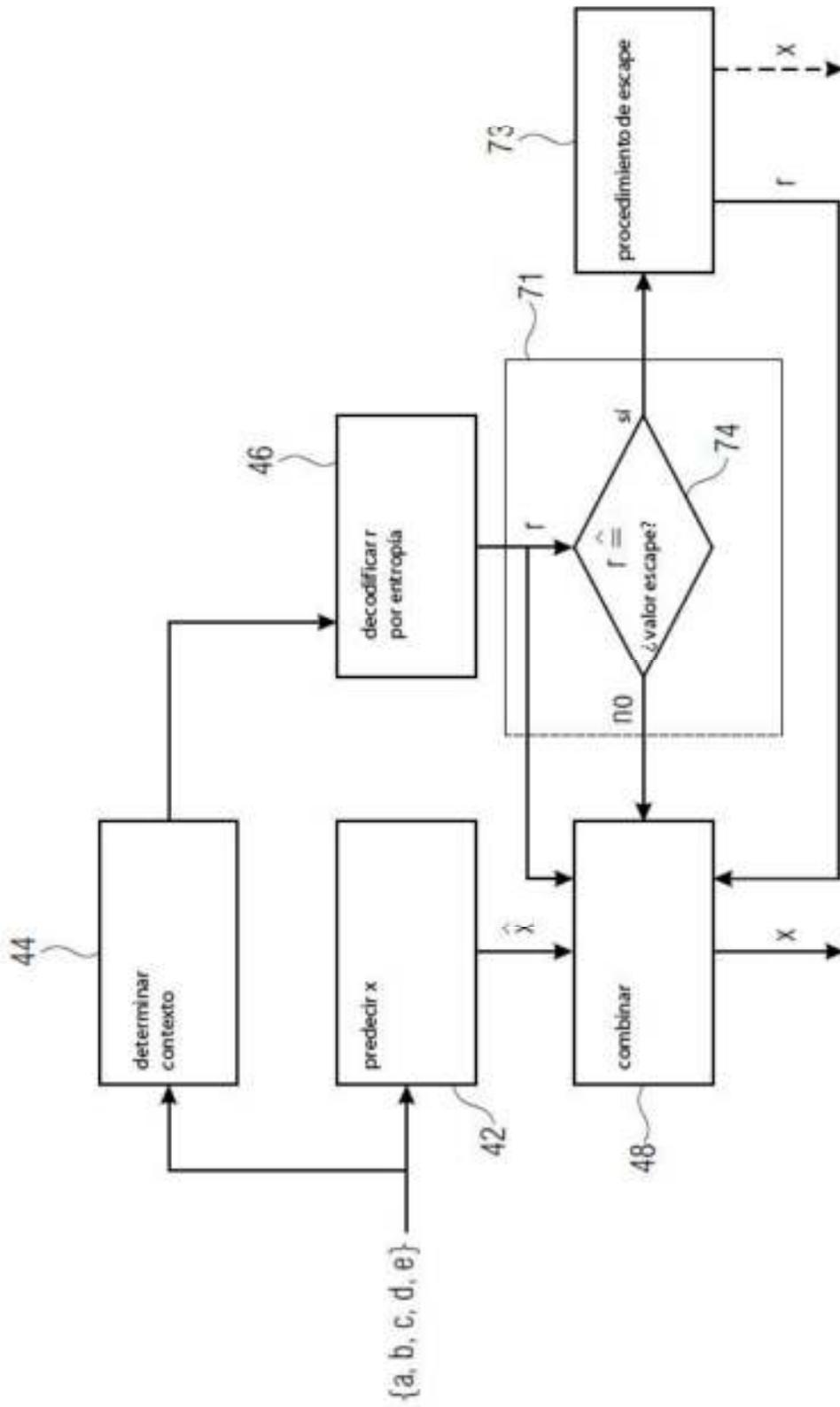


FIG. 7

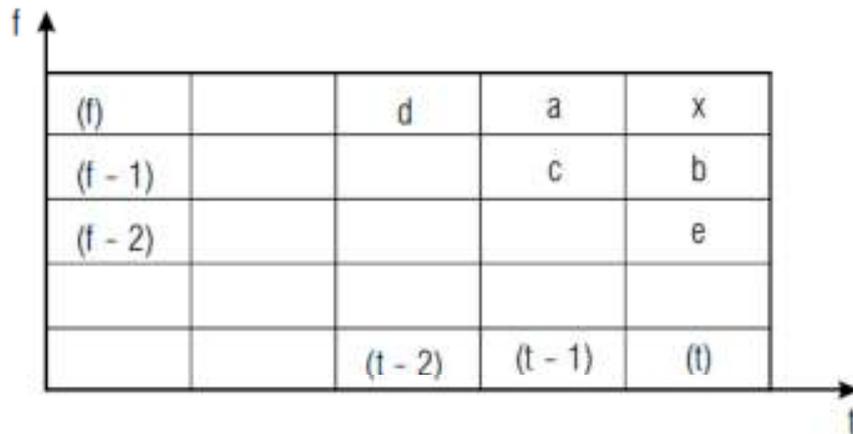


FIG. 8

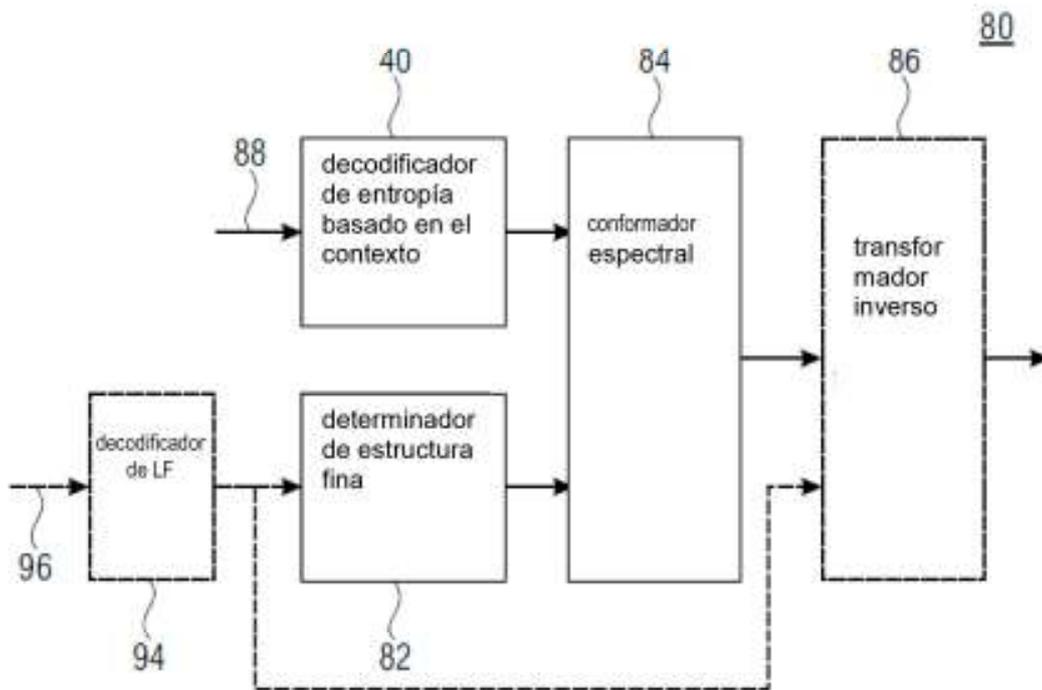


FIG. 9

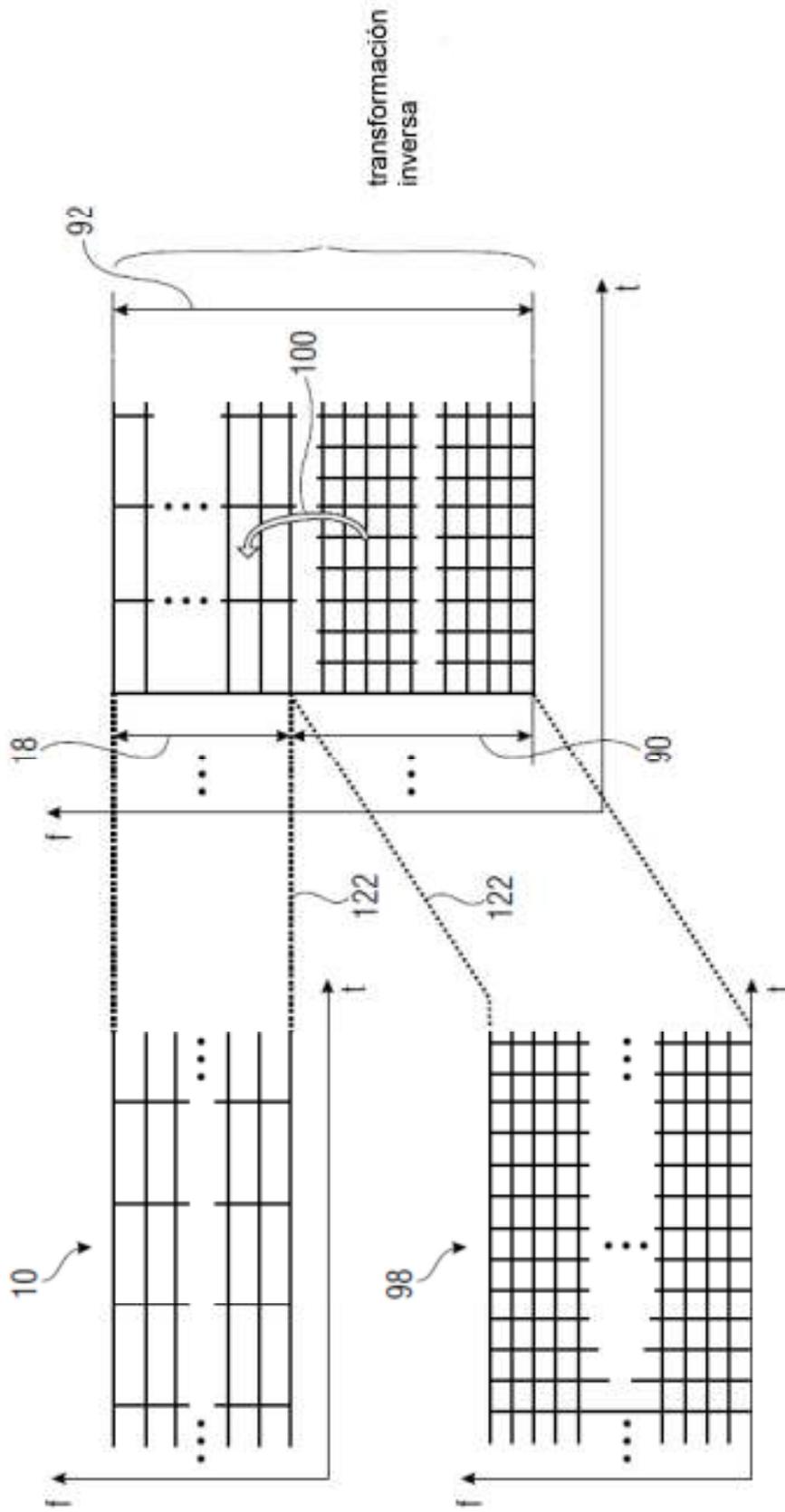


FIG. 10

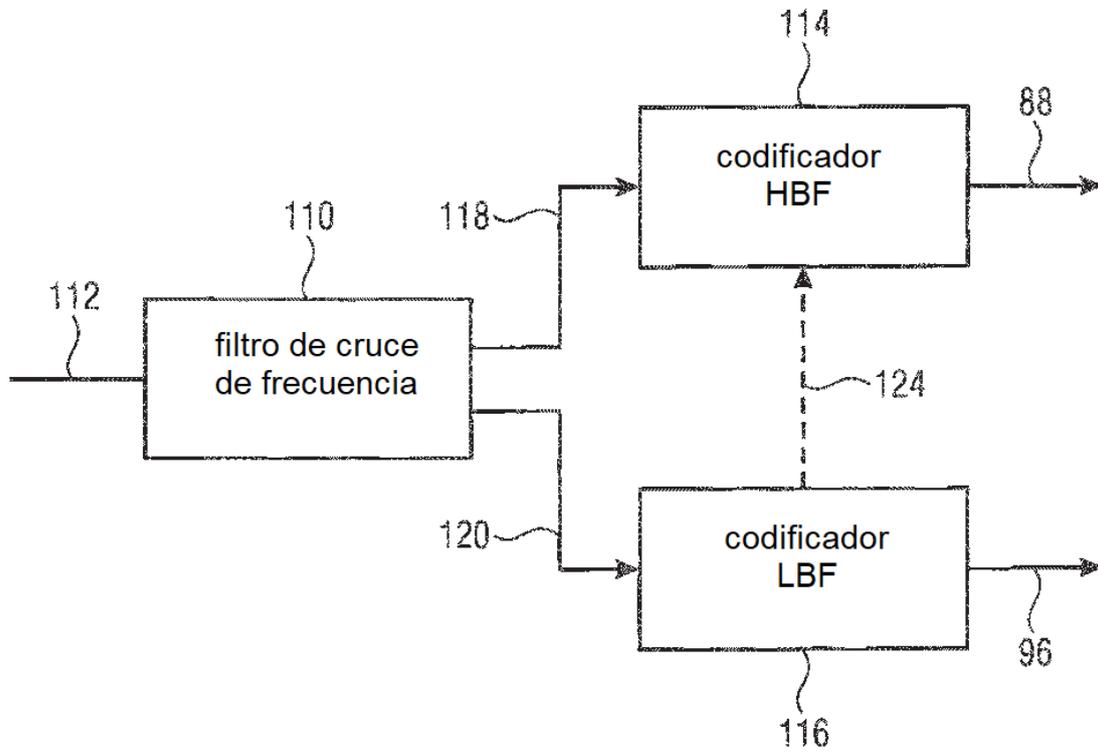


FIG 11

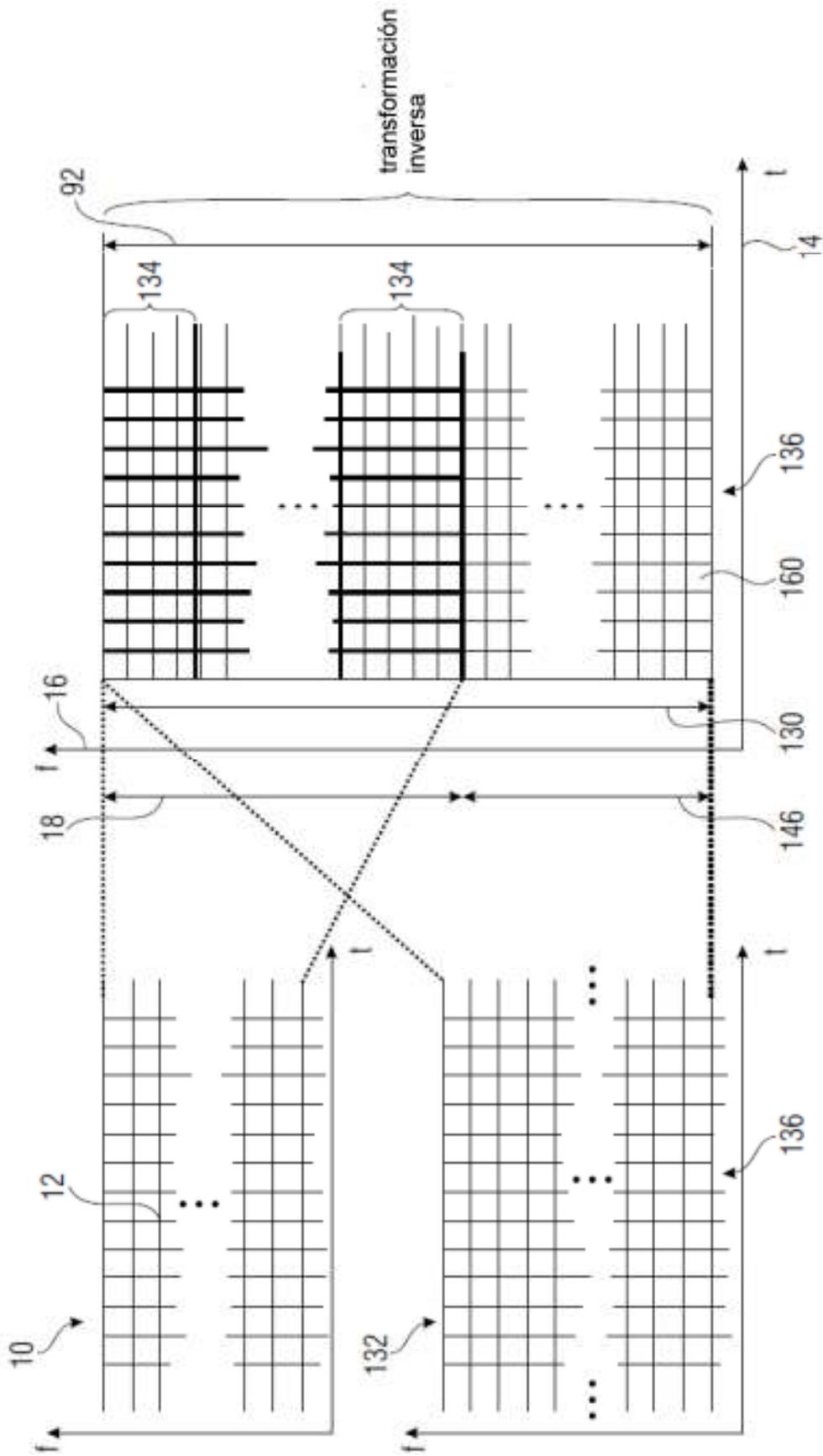


FIG. 12

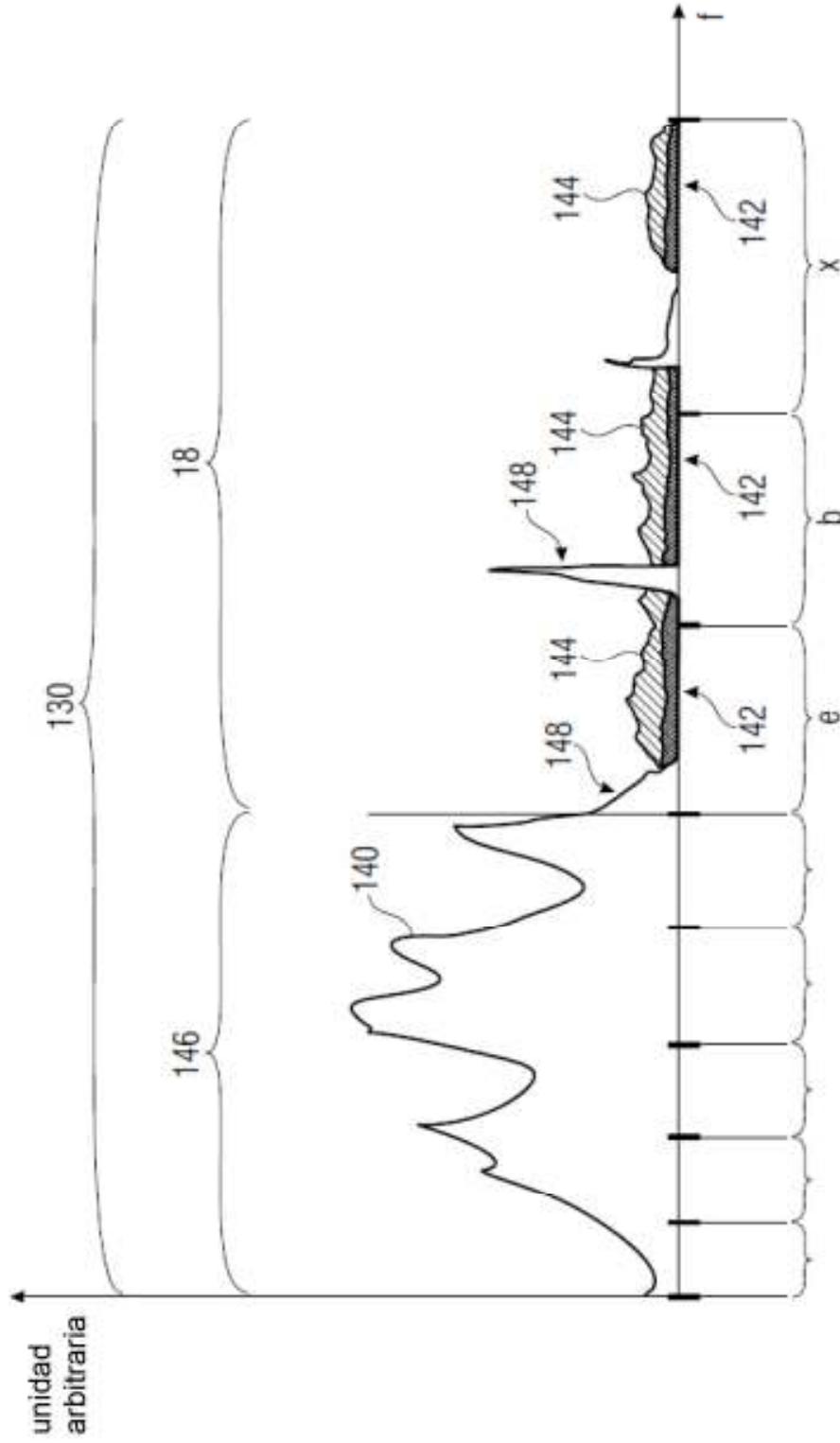


FIG. 13

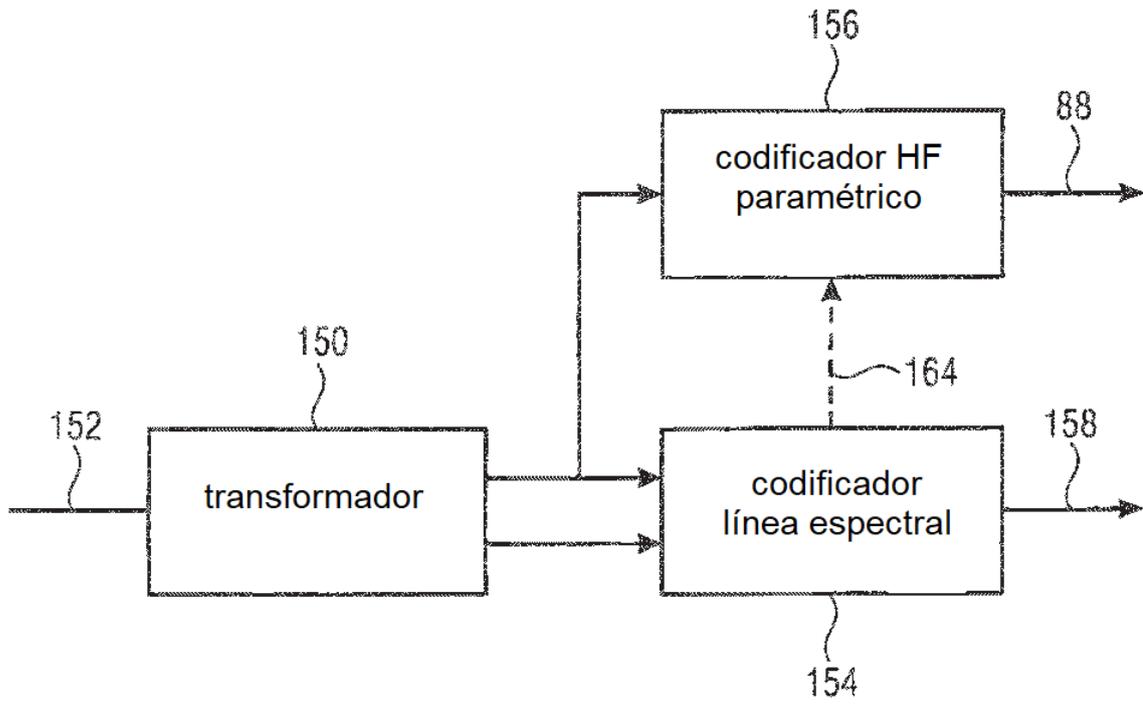


FIG 14