

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 938 858**

51 Int. Cl.:

G10L 21/0388 (2013.01)

G10L 25/90 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.01.2010** **E 21209274 (6)**

97 Fecha y número de publicación de la concesión europea: **23.11.2022** **EP 3992966**

54 Título: **Transposición armónica mejorada de producto cruzado**

30 Prioridad:

16.01.2009 US 14522309 P

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
17.04.2023

73 Titular/es:

DOLBY INTERNATIONAL AB (100.0%)
77 Sir John Rogerson's Quay, Block C, Grand
Canal Docklands
Dublin, D02 VK60, IE

72 Inventor/es:

VILLEMOES, LARS y
HEDELIN, PER

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 938 858 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Transposición armónica mejorada de producto cruzado

5 Referencia cruzada a solicitudes relacionadas

Esta solicitud es una solicitud divisional europea de la solicitud de patente europea 19171998.8 (referencia: D08072EP04), para la que se presentó el formulario de la EPO 1001 el 30 de abril de 2019.

10 Campo técnico

La presente invención se refiere a sistemas de codificación de audio que utilizan un procedimiento de transposición armónica para una reconstrucción de alta frecuencia (HFR).

15 Antecedentes de la invención

Las tecnologías HFR, tal como la tecnología de replicación de banda espectral (SBR), permiten mejorar significativamente la eficacia de codificación de los códecs de audio perceptivos tradicionales. En combinación con la codificación de audio avanzada (AAC) de MPEG-4, forma un códec de audio muy eficaz, que ya se utiliza en el sistema de radio por satélite XM y en la radio digital mundial. La combinación de la AAC y la SBR se llama aacPlus. Es parte de la norma MPEG-4, en la que se denomina perfil AAC de alta eficacia. En general, la tecnología HFR puede combinarse con cualquier códec de audio perceptivo de manera compatible con versiones anteriores y futuras, ofreciendo por tanto la posibilidad de actualizar sistemas de radiodifusión ya establecidos como la Capa-2 de MPEG usada en el sistema DAB Eureka. Los procedimientos de transposición HFR también pueden combinarse con códecs de voz para permitir voz de banda ancha a velocidades binarias ultra bajas.

La idea básica subyacente a la HFR es la observación de que normalmente hay una estrecha correlación entre las características del intervalo de altas frecuencias de una señal y las características del intervalo de bajas frecuencias de la misma señal. Por tanto, una buena aproximación para la representación del intervalo de altas frecuencias de entrada original de una señal puede conseguirse mediante una transposición de señal desde el intervalo de bajas frecuencias hasta el intervalo de altas frecuencias.

Este concepto de transposición se estableció en el documento WO 98/57436 como un procedimiento para recrear una banda de altas frecuencias a partir de una banda de frecuencias más bajas de una señal de audio. Puede obtenerse un ahorro sustancial en la velocidad binaria usando este concepto en la codificación de audio y/o en la codificación de voz. A continuación se hará referencia a la codificación de audio, pero debe observarse que los procedimientos y sistemas descritos pueden aplicarse igualmente a la codificación de voz y a una codificación unificada de voz y audio (USAC).

En un sistema de codificación de audio basado en HFR, una señal de bajo ancho de banda se presenta a un codificador de forma de onda principal y las frecuencias más altas se generan en el lado del descodificador usando la transposición de la señal de bajo ancho de banda e información complementaria adicional, que está codificada normalmente a velocidades binarias muy bajas y que describe la forma espectral objetivo. Para velocidades binarias bajas, en las que el ancho de banda de la señal codificada principal es estrecho, es cada vez más importante recrear una alta banda, es decir, el intervalo de altas frecuencias de la señal de audio, con características agradables desde un punto de vista perceptivo. A continuación se mencionan dos variantes de procedimientos de reconstrucción de frecuencias armónicas, una denominada transposición armónica y la otra denominada modulación de banda lateral única.

El principio de transposición armónica definido en el documento WO 98/57436 es que un senoide de frecuencia ω se correlaciona con un senoide de frecuencia $T\omega$, donde $T > 1$ es un entero que define el orden de la transposición. Una característica atractiva de la transposición armónica es que amplía un intervalo de frecuencias origen, formando un intervalo de frecuencias destino, en un factor igual al orden de transposición, es decir, en un factor igual a T . La transposición armónica funciona bien para material musical complejo. Además, la transposición armónica presenta bajas frecuencias de cruce, es decir, un gran intervalo de altas frecuencias superiores a la frecuencia de cruce puede generarse a partir de un intervalo relativamente pequeño de bajas frecuencias inferiores a la frecuencia de cruce.

A diferencia de la transposición armónica, una HFR basada en modulación de banda lateral única (SSB) correlaciona un senoide de frecuencia ω con un senoide de frecuencia $\omega + \Delta\omega$, donde $\Delta\omega$ es un desplazamiento de frecuencia fijo. Se ha observado que, dada una señal principal con un bajo ancho de banda, puede generarse un artefacto de llamada disonante a partir de la transposición SSB. También debe observarse que para una baja frecuencia de cruce, es decir, un pequeño intervalo de frecuencias origen, la transposición armónica requerirá un menor número de ajustes con el fin de llenar un intervalo de frecuencias objetivo deseado en comparación con la transposición basada en SSB. A modo de ejemplo, si debe llenarse el intervalo de altas frecuencias de $(\omega, 4\omega]$, usar una transposición armónica de orden de transposición $T = 4$ puede llenar este

intervalo de frecuencias a partir de un intervalo de bajas frecuencias de $(\frac{1}{4}\omega, \omega)$. Por otro lado, una transposición basada en SSB que usa el mismo intervalo de bajas frecuencias debe usar un desplazamiento de frecuencia de $\Delta\omega = \frac{3}{4}\omega$ y es necesario repetir el proceso cuatro veces para llenar el intervalo de altas frecuencias $(\omega, 4\omega]$.

5 Por otro lado, como se indica en el documento WO 02/052545 A1, la transposición armónica tiene desventajas para señales con una estructura periódica prominente. Tales señales son superposiciones de sinusoides relacionados de manera armónica con frecuencias $\Omega, 2\Omega, 3\Omega, \dots$, donde Ω es la frecuencia fundamental. Tras la transposición armónica de orden T , los sinusoides de salida tienen frecuencias $T\Omega, 2T\Omega, 3T\Omega, \dots$, que, en el caso de $T > 1$, es solo un subconjunto estricto de la serie armónica total deseada. En lo que respecta a la calidad de audio resultante, normalmente se percibirá un tono "fantasma" correspondiente a la frecuencia fundamental transpuesta $T\Omega$. Con frecuencia, la transposición armónica da como resultado un carácter sonoro "metálico" de la señal de audio codificada y descodificada. La situación puede mitigarse hasta cierto punto añadiendo varios órdenes de transposición $T = 2, 3, \dots, T_{\max}$ a la HFR, pero este procedimiento es complejo desde un punto de vista computacional si van a evitarse la mayor parte de huecos espectrales.

15 Una solución alternativa para evitar la aparición de tonos "fantasma" cuando se usa la transposición armónica se ha presentado en el documento WO 02/052545 A1. La solución consiste en usar dos tipos de transposición, es decir, una transposición armónica típica y una "transposición de impulsos" especial. El procedimiento descrito conmuta a la "transposición de impulsos" dedicada en aquellas partes de la señal de audio detectadas como periódicas a modo de un tren de impulsos. El problema de este enfoque es que la aplicación de la "transposición de impulsos" en material musical complejo degrada normalmente la calidad en comparación con una transposición armónica basada en un banco de filtros de alta resolución. Por tanto, los mecanismos de detección tienen que ajustarse de manera muy conservadora para que la transposición de impulsos no se use con material complejo. Inevitablemente, los instrumentos y veces de un solo tono se clasifican algunas veces como señales complejas, invocándose de ese modo la transposición armónica y, por lo tanto, perdiéndose armónicos. Además, si la conmutación se produce en la parte central de una señal de un solo tono, o de una señal con un tono dominante en un fondo complejo más débil, la propia conmutación entre los dos procedimientos de transposición, que tienen propiedades de llenado de espectro muy diferentes, generará artefactos audibles. Otra variante para realizar una reconstrucción de frecuencia armónica se propone en el documento US 2004/0028244 A1.

30

Resumen de la invención

La invención se define como en las reivindicaciones independientes adjuntas. Otras realizaciones se definen en las reivindicaciones dependientes.

35

La presente invención proporciona un procedimiento y un sistema para completar la serie armónica resultante de la transposición armónica de una señal periódica. La transposición en el dominio de frecuencia comprende la etapa de correlacionar señales de subbanda modificadas de manera no lineal procedentes de un banco de filtros de análisis con subbandas seleccionadas de un banco de filtros de síntesis. La modificación no lineal comprende una modificación de fase o rotación de fase, que en un dominio de banco de filtros complejos puede obtenerse mediante una ley de potencia seguida de un ajuste de magnitud. Mientras que la transposición de la técnica anterior modifica una subbanda de análisis a la vez por separado, la presente invención añade una combinación no lineal de al menos dos subbandas de análisis diferentes para cada subbanda de síntesis. La separación entre las subbandas de análisis que van a combinarse puede estar relacionada con la frecuencia fundamental de una componente dominante de la señal que va a transponerse. En la forma más general, la descripción matemática de la invención es que un conjunto de componentes de frecuencia $\omega_1, \omega_2, \dots, \omega_K$, se usa para crear una nueva componente de frecuencia

45

$$\omega = T_1\omega_1 + T_2\omega_2 + \dots + T_K\omega_K,$$

50

donde los coeficientes $T_1, T_2, \dots, T_K$ son órdenes de transposición enteros cuya suma es el orden de transposición total $T = T_1 + T_2 + \dots + T_K$. Este efecto se obtiene modificando las fases de K señales de subbanda elegidas de manera adecuada mediante los factores $T_1, T_2, \dots, T_K$ y recombinando el resultado en una señal con una fase igual a la suma de las fases modificadas. Es importante observar que todas estas operaciones de fase están bien definidas y no son ambiguas ya que los órdenes de transposición individuales son números enteros, y que algunos de estos números enteros pueden ser incluso negativos siempre que el orden de transposición total satisfaga que $T \geq 1$.

55

Los procedimientos de la técnica anterior corresponden al caso de $K = 1$, y la invención actual insta a usar $K \geq 2$. El texto descriptivo trata principalmente el caso de $K = 2, T \geq 2$ ya que es suficiente para solucionar la mayor parte de los problemas específicos existentes. Sin embargo, debe observarse que los casos en los que $K > 2$ también se contemplan y se tratan en el presente documento.

60

Breve descripción de los dibujos

A continuación se describirá la presente invención mediante ejemplos ilustrativos, los cuales no limitan el alcance de la invención. Se describirá con referencia a los dibujos adjuntos, en los que:

La Fig. 1 ilustra el funcionamiento de un descodificador de audio mejorado HFR.

La Fig. 2 ilustra el funcionamiento de un transposicionador armónico que usa varios órdenes.

La Fig. 3 ilustra el funcionamiento de un transposicionador armónico de dominio de frecuencia (FD).

La Fig. 4 ilustra el funcionamiento del uso inventivo del procesamiento de términos cruzados.

La Fig. 5 ilustra un procesamiento directo de la técnica anterior.

La Fig. 6 ilustra un procesamiento no lineal directo de la técnica anterior de una única subbanda.

La Fig. 7 ilustra los componentes del procesamiento inventivo de términos cruzados.

La Fig. 8 ilustra el funcionamiento de un bloque de procesamiento de términos cruzados.

La Fig. 9 ilustra el procesamiento no lineal inventivo realizado en cada uno de los sistemas MISO de la Fig. 8.

Las Fig. 10 a 18 ilustran el efecto de la invención en la transposición armónica de señales periódicas a modo de ejemplo.

La Fig. 19 ilustra la resolución en tiempo-frecuencia de una transformada de Fourier de tiempo corto (STFT).

La Fig. 20 ilustra la progresión en el tiempo a modo de ejemplo de una función de ventana y su transformada de Fourier usada en el lado de síntesis.

La Fig. 21 ilustra la STFT de una señal de entrada sinusoidal.

La Fig. 22 ilustra la función de ventana y su transformada de Fourier según la Fig. 20 usadas en el lado de análisis.

Las Fig. 23 y 24 ilustran la determinación de subbandas apropiadas de banco de filtros de análisis para la mejora de términos cruzados de una subbanda de banda de filtro de análisis.

Las Fig. 25, 26 y 27 ilustran resultados experimentales del procedimiento descrito de transposición armónica de términos cruzados y de términos directos.

Las Fig. 28 y 29 ilustran realizaciones de un codificador y un descodificador, respectivamente, que usan los esquemas de transposición armónica mejorada descritos en el presente documento.

La Fig. 30 ilustra una realización de una unidad de transposición mostrada en las Fig. 28 y 29.

Descripción de realizaciones preferidas

Las realizaciones descritas a continuación son simplemente ejemplos de los principios de la presente invención para la denominada transposición armónica mejorada de producto cruzado. Debe entenderse que las modificaciones y variaciones de las disposiciones y los detalles descritos en este documento resultarán evidentes a los expertos en la técnica. Por lo tanto, la invención solo está limitada por el alcance de las reivindicaciones de patente adjuntas y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones de este documento.

La Fig. 1 ilustra el funcionamiento de un descodificador de audio mejorado HFR. El descodificador de audio principal 101 proporciona una señal de audio de bajo ancho de banda que se introduce en un muestreador ascendente 104 que puede ser necesario para producir una contribución de salida de audio final a la velocidad de muestreo total deseada. Tal muestreo ascendente es necesario en sistemas de doble velocidad, en los que el códec de audio principal limitado por banda funciona a la mitad de la velocidad de muestreo de audio externa, mientras que la parte HFR se procesa a la frecuencia de muestreo total. Por consiguiente, en un sistema de una única velocidad, este muestreador ascendente 104 se omite. La salida de bajo ancho de banda de 101 también se envía al transposicionador o unidad de transposición 102, que proporciona una señal transpuesta, es decir, una señal que comprende el intervalo de altas frecuencias deseado. Esta señal transpuesta puede conformarse

en tiempo y en frecuencia mediante el ajustador de envolvente 103. La salida de audio final es la suma de la señal principal de bajo ancho de banda y la señal transpuesta de envolvente ajustada.

La Fig. 2 ilustra el funcionamiento de un transposicionador armónico 201, que corresponde al transposicionador 102 de la Fig. 1, que comprende varios transposicionadores de diferente orden de transposición T . La señal que va a transponerse se pasa al banco de transposicionadores individuales 201-2, 201-3,..., 201- $T_{\max}$ que tienen órdenes de transposición de $T = 2, 3, \dots, T_{\max}$, respectivamente. Normalmente, un orden de transposición $T_{\max} = 3$ es suficiente para la mayoría de aplicaciones de codificación de audio. Las contribuciones de los diferentes transposicionadores 201-2, 201-3,..., 201- $T_{\max}$ se suman en 202 para proporcionar la salida de transposicionador combinada. En una primera realización, esta operación de suma puede comprender la suma de las contribuciones individuales. En otra realización, las contribuciones se ponderan con diferentes pesos, de manera que se mitiga el efecto de sumar múltiples contribuciones a determinadas frecuencias. Por ejemplo, las contribuciones de tercer orden pueden sumarse con una ganancia inferior a las contribuciones de segundo orden. Finalmente, la unidad de suma 202 puede sumar las contribuciones de manera selectiva, dependiendo de la frecuencia de salida. Por ejemplo, la transposición de segundo orden puede usarse en un primer intervalo de frecuencias objetivo más bajas, y la transposición de tercer orden puede usarse en un segundo intervalo de frecuencias objetivo más altas.

La Fig. 3 ilustra el funcionamiento de un transposicionador armónico de dominio de frecuencia (FD), tal como uno de los bloques individuales de 201, es decir, uno de los transposicionadores 201- T de orden de transposición T . Un banco de filtros de análisis 301 proporciona subbandas complejas que se someten a un procesamiento no lineal 302 que modifica la fase y/o la amplitud de la señal de subbanda según el orden de transposición elegido T . Las subbandas modificadas se introducen en un banco de filtros de síntesis 303, que proporciona la señal de dominio de tiempo transpuesta. En caso de múltiples transposicionadores paralelos de diferentes órdenes de transposición, tal como se muestra en la Fig. 2, algunas operaciones de banco de filtros pueden compartirse entre diferentes transposicionadores 201-2, 201-3,..., 201- $T_{\max}$. La compartición de las operaciones de banco de filtros puede realizarse para análisis o síntesis. En caso de síntesis compartida 303, la suma 202 puede realizarse en el dominio de subbanda, es decir, antes de la síntesis 303.

La Fig. 4 ilustra el funcionamiento del procesamiento de términos cruzados 402 además del procesamiento directo 401. El procesamiento de términos cruzados 402 y el procesamiento directo 401 se llevan a cabo en paralelo en el bloque de procesamiento no lineal 302 del transposicionador armónico de dominio de frecuencia de la Fig. 3. Las señales de salida transpuestas se combinan, por ejemplo se suman, para proporcionar una señal transpuesta conjunta. Esta combinación de señales de salida transpuestas puede consistir en la superposición de las señales de salida transpuestas. Opcionalmente, la suma selectiva de términos cruzados puede implementarse en el cálculo de ganancia.

La Fig. 5 ilustra en mayor detalle el funcionamiento del bloque de procesamiento directo 401 de la Fig. 4 en el transposicionador armónico de dominio de frecuencia de la Fig. 3. Las unidades de única entrada y única salida (SISO) 401-1,..., 401- n ,..., 401- N correlacionan cada subbanda de análisis de un intervalo origen con una subbanda de síntesis de un intervalo destino. Según la Fig. 5, una subbanda de análisis de índice n se correlaciona mediante la unidad SISO 401- n con una subbanda de síntesis del mismo índice n . Debe observarse que el intervalo de frecuencias de la subbanda con índice n en el banco de filtro de síntesis puede variar dependiendo de la versión o tipo exactos de la transposición armónica. En la versión o tipo ilustrado en la Fig. 5, la separación entre frecuencias del banco de análisis 301 es un factor T más pequeño que el del banco de síntesis 303. Por tanto, el índice n del banco de síntesis 303 corresponde a una frecuencia que es T veces mayor que la frecuencia de la subbanda con el mismo índice n del banco de análisis 301. A modo de ejemplo, una subbanda de análisis $[(n-1)\omega, n\omega]$ se transpone formando una subbanda de síntesis $[(n-1)T\omega, nT\omega]$.

La Fig. 6 ilustra el procesamiento no lineal directo de una única subbanda incluida en cada una de las unidades SISO de 401- n . La no linealidad del bloque 601 lleva a cabo una multiplicación de la fase de la señal de subbanda compleja por un factor igual al orden de transposición T . La unidad de ganancia opcional 602 modifica la magnitud de la señal de subbanda modificada en fase. En términos matemáticos, la salida y de la unidad SISO 401- n puede escribirse en función de la entrada x en el sistema SISO 401- n y el parámetro de ganancia g de la siguiente manera:

$$y = g \cdot v^T, \text{ donde } v = x / |x|^{1-1/T}. \quad (1)$$

Esto también puede escribirse como:

$$y = g \cdot |x| \cdot \left(\frac{x}{|x|} \right)^T.$$

Expresado en palabras, la fase de la señal de subbanda compleja x se multiplica por el orden de transposición T y la amplitud de la señal de subbanda compleja x se modifica mediante el parámetro de ganancia g .

La Fig. 7 ilustra los componentes del procesamiento de términos cruzados 402 para una transposición armónica de orden T . Hay $T-1$ bloques de procesamiento de términos cruzados en paralelo, 701-1,..., 701- r ,..., 701-($T-1$), cuyas salidas se suman en la unidad de suma 702 para producir una salida combinada. Como ya se ha mencionado en la sección de introducción, un objetivo es correlacionar un par de sinusoides de frecuencias (ω , $\omega + \Omega$) con un senoide de frecuencia $(T-r)\omega + r(\omega + \Omega) = T\omega + r\Omega$, donde la variable r varía entre 1 y $T-1$. Dicho de otro modo, dos subbandas del banco de filtros de análisis 301 se correlacionan con una subbanda del intervalo de altas frecuencias. Para un valor particular de r y un orden de transposición dado T , esta etapa de correlación se lleva a cabo en el bloque de procesamiento de términos cruzados 701- r .

La Fig. 8 ilustra el funcionamiento de un bloque de procesamiento de términos cruzados 701- r para un valor fijo $r = 1, 2, \dots, T-1$. Cada subbanda de salida 803 se obtiene en una unidad de múltiples entradas y única salida (MISO) 800- n a partir de dos subbandas de entrada 801 y 802. Para una subbanda de salida 803 de índice n , las dos entradas de la unidad MISO 800- n son las subbandas $n - p_1$, 801, y $n + p_2$, 802, donde p_1 y p_2 son desplazamientos de índice enteros positivos, que dependen del orden de transposición T , la variable r y el parámetro de tono de mejora de producto cruzado Ω . La convención de numeración de subbandas de análisis y síntesis va en consonancia con la de la Fig. 5, es decir, la separación en frecuencia del banco de análisis 301 es un factor T más pequeño que el del banco de síntesis 303 y, por consiguiente, los comentarios anteriores referentes a variaciones del factor T siguen siendo válidos.

Con relación al uso del procesamiento de términos cruzados, deben considerarse las siguientes observaciones. El parámetro de tono Ω no tiene que conocerse con una alta precisión y, ciertamente, sin una mejor resolución de frecuencia que la resolución de frecuencia obtenida por el banco de filtros de análisis 301. De hecho, en algunas realizaciones de la presente invención, el parámetro de tono de mejora de producto cruzado subyacente Ω no se introduce en el descodificador. En cambio, el par elegido de desplazamientos de índice enteros (p_1 , p_2) se selecciona a partir de una lista de posibles candidatos siguiendo un criterio de optimización tal como la maximización de la magnitud de salida de producto cruzado, es decir, la maximización de la energía de la salida de producto cruzado. A modo de ejemplo, para valores dados de T y r , puede usarse una lista de candidatos dada por la fórmula $(p_1, p_2) = (rl, (T-r)l)$, $l \in L$, donde L es una lista de enteros positivos. Esto se muestra en mayor detalle posteriormente en el contexto de la fórmula (11). Todos los enteros positivos son, en principio, válidos como candidatos. En algunos casos, la información de tono puede ayudar a identificar qué l escoger como desplazamientos de índice apropiados.

Además, incluso aunque el procesamiento de producto cruzado de ejemplo ilustrado en la Fig. 8 sugiere que los desplazamientos de índice aplicados (p_1 , p_2) son los mismos para un determinado intervalo de subbandas de salida, por ejemplo las subbandas de síntesis $(n-1)$, n y $(n+1)$ se forman a partir de subbandas de análisis que tienen una distancia fija $p_1 + p_2$, éste no tiene por qué ser el caso. De hecho, los desplazamientos de índice (p_1 , p_2) pueden diferir para cada subbanda de salida. Esto significa que para cada subbanda n puede seleccionarse un valor Ω diferente del parámetro de tono de mejora de producto cruzado.

La Fig. 9 ilustra el procesamiento no lineal realizado en cada una de las unidades MISO 800- n . La operación producto 901 crea una señal de subbanda con una fase igual a la suma ponderada de las fases de las dos señales de subbanda de entrada complejas y una magnitud igual al valor medio generalizado de las magnitudes de las dos muestras de subbanda de entrada. La unidad de ganancia opcional 902 modifica la magnitud de las muestras de subbanda modificadas en fase. En términos matemáticos, la salida y puede escribirse en función de las entradas u_1 801 y u_2 802 de la unidad MISO 800- n y el parámetro de ganancia g de la siguiente manera:

$$y = g \cdot v_1^{T-r} v_2^r, \text{ donde } v_m = u_m / |u_m|^{1-1/T}, \text{ para } m=1, 2. \quad (2)$$

Esto también puede escribirse como:

$$y = \mu(|u_1|, |u_2|) \cdot \left(\frac{u_1}{|u_1|} \right)^{T-r} \left(\frac{u_2}{|u_2|} \right)^r,$$

donde $\mu(|u_1|, |u_2|)$ es una función de generación de magnitud. Expresado en palabras, la fase de la señal de subbanda compleja u_1 se multiplica por el orden de transposición $T-r$, y la fase de la señal de subbanda compleja u_2 se multiplica por el orden de transposición r . La suma de esas dos fases se usa como la fase de la salida y

cuya magnitud se obtiene mediante la función de generación de magnitud. En comparación con la fórmula (2), la función de generación de magnitud se expresa como la media geométrica de magnitudes modificadas por el parámetro de ganancia g , es decir, $\mu(|u_1|, |u_2|) = g \cdot |u_1|^{1-r/T} |u_2|^{r/T}$. Al permitir que el parámetro de ganancia dependa de las entradas, esto cubre por tanto todas las posibilidades.

5

Debe observarse que la fórmula (2) se obtiene a partir del resultado subyacente de que un par de sinusoides de frecuencias $(\omega, \omega+\Omega)$ van a correlacionarse con un senoide de frecuencia $T\omega + r\Omega$, lo que puede escribirse también como $(T-r)\omega + r(\omega+\Omega)$.

- 10 A continuación se ofrece una descripción matemática de la presente invención. Por simplicidad se consideran señales de tiempo continuo. Se supone que el banco de filtros de síntesis 303 consigue una reconstrucción perfecta a partir de un banco de filtros de análisis modulado complejo correspondiente 301 con una función de ventana simétrica de valor real o filtro de prototipo $w(t)$. El banco de filtros de síntesis usará normalmente, pero no siempre, la misma ventana en el proceso de síntesis. Se supone que la modulación va a ser del tipo
- 15 apilamiento par, que el salto está normalizado a uno y que la separación entre frecuencias angulares de las subbandas de síntesis está normalizada a π . Por tanto, se obtendrá una señal objetivo $s(t)$ en la salida del banco de filtros de síntesis si las señales de subbanda de entrada en el banco de filtros de síntesis se proporcionan como señales de subbanda de síntesis $y_n(k)$,

$$y_n(k) = \int_{-\infty}^{\infty} s(t)w(t-k)\exp[-in\pi(t-k)]dt. \quad (3)$$

20

Debe observarse que la fórmula (3) es un modelo matemático de tiempo continuo normalizado de las operaciones habituales en un banco de filtros de análisis de subbanda modulado complejo, tal como una transformada discreta de Fourier (DFT) basada en ventanas, denotada también como transformada de Fourier de tiempo corto (STFT). Con una ligera modificación del argumento del valor exponencial complejo de la fórmula (3), se obtienen modelos de tiempo continuo para un banco de filtros de espejo en cuadratura (QMF) (seudo) modulado complejo y una transformada de coseno discreta modificada compleja (CMDCT), también denominada DFT basada en ventanas con apilamiento impar. El índice de subbanda n abarca todos los enteros no negativos para el caso de tiempo continuo. Para los homólogos de tiempo discreto, la variable de tiempo t se muestrea en la etapa $1/N$, y el índice de subbanda n está limitado por N , donde N es el número de subbandas del banco de filtros, que es igual al lapso de tiempo discreto del banco de filtros. En el caso de tiempo discreto, también se requiere un factor de normalización relacionado con N en la operación de transformación si no está incorporado en el escalado de la ventana.

- 35 Para una señal de valor real, hay tantas muestras de subbanda complejas como muestras de valor real para el modelo de banco de filtros elegido. Por lo tanto, hay un sobremuestreo (o redundancia) total en un factor de dos. También pueden usarse bancos de filtros con un mayor grado de sobremuestreo, pero el sobremuestreo se mantiene en un nivel bajo en la presente descripción de realizaciones para facilitar la explicación.

- 40 Las etapas principales implicadas en el análisis de banco de filtros modulado correspondiente a la fórmula (3) son que la señal se multiplica por una ventana centrada en torno a un tiempo $t = k$, y la señal de ventana resultante se correlaciona con cada uno de los sinusoides complejos $\exp[-in\pi(t-k)]$. En implementaciones de tiempo discreto, esta correlación se implementa de manera eficaz a través de una transformada rápida de Fourier. Las etapas algorítmicas correspondientes para el banco de filtros de síntesis son ampliamente conocidas por los expertos en la técnica y consisten en modulación de síntesis, división en ventanas de síntesis y operaciones de solapamiento y suma.

- La Fig. 19 ilustra la posición en tiempo y en frecuencia correspondiente a la información transportada por la muestra de subbanda $y_n(k)$ para una selección de valores de índice de tiempo k y de índice de subbanda n .
- 50 Como un ejemplo, la muestra de subbanda $y_5(4)$ se representa mediante el rectángulo oscuro 1901.

Para un senoide, $s(t) = A\cos(\omega t + \theta) = \text{Re}\{C\exp(i\omega t)\}$, las señales de subbanda de (3) son para un n suficientemente grande con una buena aproximación, expresadas de la siguiente forma

$$y_n(k) = Ce^{ik\omega} \int_{-\infty}^{\infty} w(t)\exp[-i(n\pi - \omega)t]dt = Ce^{ik\omega} \hat{w}(n\pi - \omega), \quad (4)$$

55

donde el símbolo ' $\hat{\cdot}$ ' denota la transformada de Fourier, es decir, $\hat{w}$ es la transformada de Fourier de la función de ventana w .

En sentido estricto, la fórmula (4) solo se cumple si se añade un término con $-\omega$ en lugar de ω . Este término se desprecia según la suposición de que la respuesta de frecuencia de la ventana disminuye de manera suficientemente rápida y de que la suma de ω y n no se aproxima a cero.

La Fig. 20 ilustra la apariencia típica de una ventana w , 2001, y de su transformada de Fourier $\hat{w}$, 2002.

La Fig. 21 ilustra el análisis de un único senoide correspondiente a la fórmula (4). Las subbandas afectadas principalmente por el senoide a una frecuencia ω son aquéllas con índice n de manera que $n\pi - \omega$ es un valor pequeño. En el ejemplo de la Fig. 21, la frecuencia es $\omega = 6,25\pi$, como se indica mediante la línea discontinua horizontal 2101. En ese caso, las tres subbandas para $n = 5, 6, 7$, representadas mediante los signos de referencia 2102, 2103, 2104, respectivamente, contienen señales de subbanda significativas distintas de cero. El oscurecimiento de esas tres subbandas refleja la amplitud relativa de los sinusoides complejos en cada subbanda obtenida a partir de la fórmula (4). Una sombra más oscura significa mayor amplitud. En el ejemplo concreto, esto significa que la amplitud de la subbanda 5, es decir 2102, es menor comparada con la amplitud de la subbanda 7, es decir 2104, que, de nuevo, es menor que la amplitud de la subbanda 6, es decir 2103. Es importante observar que varias subbandas diferentes de cero pueden ser necesarias, por lo general, para poder sintetizar un senoide de alta calidad en la salida del banco de filtro de síntesis, especialmente en casos en los que la ventana tiene una apariencia como la ventana 2001 de la Fig. 20, con una duración de tiempo relativamente corta y notables curvas laterales de frecuencia.

Las señales de subbanda de síntesis $y_n(k)$ también pueden determinarse como resultado del banco de filtros de análisis 301 y del procesamiento no lineal, es decir, el transposicionador armónico 302 ilustrado en la Fig. 3. En el lado del banco de filtros de análisis, las señales de subbanda de análisis $x_n(k)$ pueden representarse en función de la señal origen $z(t)$. Para una transposición de orden T , un banco de filtros de análisis modulado complejo con ventana $w_T(t) = w(t/T)/T$, un salto de uno y una etapa de frecuencia de modulación, que es T veces menor que la etapa de frecuencia del banco de síntesis, se aplica a la señal origen $z(t)$. La Fig. 22 ilustra la apariencia de la ventana escalada w_T 2201 y su transformada de Fourier $\hat{w}_T$ 2202. En comparación con la Fig. 20, la ventana de tiempo 2201 está ensanchada y la ventana de frecuencia 2202 está comprimida.

El análisis del banco de filtros modificado da lugar a las señales de subbanda de análisis $x_n(k)$:

$$x_n(k) = \int_{-\infty}^{\infty} z(t) w_T(t - k) \exp\left[-i \frac{n\pi}{T} (t - k)\right] dt \quad (5)$$

Para un senoide, $z(t) = B \cos(\xi t + \varphi) = \text{Re}\{D \exp(i\xi t)\}$, se observa que las señales de subbanda de (5) para un n suficientemente grande con una buena aproximación se obtienen de la siguiente manera:

$$x_n(k) = D \exp(ik\xi) \hat{w}(n\pi - T\xi). \quad (6)$$

Por tanto, enviando estas señales de subbanda al transposicionador armónico 302 y aplicando la regla de transposición directa (1) a (6) se obtiene

$$\tilde{y}_n(k) = gD \left(\frac{D}{|D|}\right)^{T-1} \left(\frac{\hat{w}(n\pi - T\xi)}{|\hat{w}(n\pi - T\xi)|}\right)^{T-1} \cdot \exp(ikT\xi) \hat{w}(n\pi - T\xi). \quad (7)$$

De manera ideal, las señales de subbanda de síntesis $y_n(k)$ obtenidas mediante la fórmula (4) y las señales de subbanda no lineales obtenidas a través de la transposición armónica $\tilde{y}_n(k)$ de la fórmula (7) deben coincidir.

Para órdenes de transposición impares T , el factor que contiene la influencia de la ventana en (7) es igual a uno, ya que la transformada de Fourier de la ventana tiene supuestamente un valor real, y $T-1$ es un número par. Por lo tanto, la fórmula (7) puede hacerse corresponder exactamente con la fórmula (4) con $\omega = T\xi$, para todas las subbandas, de manera que la salida del banco de filtros de síntesis con señales de subbanda de entrada según la fórmula (7) es un senoide con una frecuencia $\omega = T\xi$, amplitud $A = gB$ y fase $\theta = T\varphi$, donde B y φ se

determinan a partir de la fórmula: $D = B \exp(i\varphi)$, que tras su inserción se obtiene $gD \left(\frac{D}{|D|}\right)^{T-1} = gB \exp(iT\varphi)$. Por tanto,

se obtiene una transposición armónica de orden T de la señal origen sinusoidal $z(t)$.

Para T par, la correspondencia es más aproximada, pero sigue dependiendo de la parte con valor positivo de la respuesta de frecuencia de ventana $\hat{w}$, que para una ventana simétrica de valores reales incluye la curva principal más importante. Esto significa que incluso para valores pares de T también se obtiene una transposición armónica de la señal origen sinusoidal $z(t)$. En el caso particular de una ventana gaussiana, $\hat{w}$ es siempre positiva y, por consiguiente, no hay diferencia en el funcionamiento para órdenes de transposición pares e impares.

- 10 De manera similar a la fórmula (6), el análisis de un senoide de frecuencia $\xi + \Omega$, es decir, la señal origen sinusoidal $z(t) = B \cos((\xi + \Omega)t + \varphi) = \text{Re}\{E \exp(i(\xi + \Omega)t)\}$, es

$$x'_n(k) = E \exp(ik(\xi + \Omega)) \hat{w}(n\pi - T(\xi + \Omega)). \quad (8)$$

- 15 Por lo tanto, introducir las dos señales de subbanda $u_1 = x_{n-p_1}(k)$, correspondiente a la señal 801 de la Fig. 8, y $u_2 = x'_{n+p_2}(k)$, correspondiente a la señal 802 de la Fig. 8, en el procesamiento de producto cruzado 800-n ilustrado en la Fig. 8, y aplicar la fórmula de producto cruzado (2) proporciona la señal de subbanda de salida 803

$$\hat{y}_n(k) = g \exp[ik(T\xi + r\Omega)] M(n, \xi), \quad (9)$$

20 donde

$$M(n, \xi) = \frac{D^{T-r} E'}{|D^{T-r} E'|^{1-1/T}} \frac{\hat{w}((n-p_1)\pi - T\xi)^{T-r} \hat{w}((n+p_2)\pi - T(\xi + \Omega))^r}{|\hat{w}((n-p_1)\pi - T\xi)^{T-r} \hat{w}((n+p_2)\pi - T(\xi + \Omega))^r|^{1-1/T}}. \quad (10)$$

- 25 A partir de la fórmula (9) puede observarse que la evolución de fase de la señal de subbanda de salida 803 del sistema MISO 800-n sigue la evolución de fase de un análisis de un senoide de frecuencia $T\xi + r\Omega$. Esto se cumple independientemente de la elección de los desplazamientos de índice p_1 y p_2 . De hecho, si la señal de subbanda (9) se introduce en un canal de subbanda n correspondiente a la frecuencia $T\xi + r\Omega$, es decir, si $n\pi \approx T\xi + r\Omega$, entonces la salida será una contribución a la generación de un senoide de frecuencia $T\xi + r\Omega$. Sin embargo, es ventajoso asegurarse de que cada contribución sea significativa y de que las contribuciones se sumen de manera beneficiosa. Estos aspectos se describirán a continuación.

- 30 Dado un parámetro de tono de mejora de producto cruzado Ω , pueden obtenerse elecciones adecuadas para desplazamientos de índice p_1 y p_2 con el fin de que la magnitud compleja $M(n, \xi)$ de (10) se aproxime a $\hat{w}(n\pi - (T\xi + r\Omega))$ para un intervalo de subbandas n , en cuyo caso la salida final se aproximará a un senoide de frecuencia $T\xi + r\Omega$. Una primera consideración acerca de las curvas principales impone que los tres valores de $(n - p_1)\pi - T\xi$, $(n + p_2)\pi - T(\xi + \Omega)$, $n\pi - (T\xi + r\Omega)$ sean pequeños simultáneamente, lo que da lugar a las igualdades de aproximación

$$p_1 \approx r \frac{\Omega}{\pi} \quad \text{y} \quad p_2 \approx (T - r) \frac{\Omega}{\pi}. \quad (11)$$

- 40 Esto significa que cuando se conoce el parámetro de tono de mejora de producto cruzado Ω , los desplazamientos de índice pueden aproximarse mediante la fórmula (11), permitiendo de ese modo una selección simple de subbandas de análisis. Un análisis más minucioso de los efectos de la elección de los desplazamientos de índice p_1 y p_2 según la fórmula (11) en la magnitud del parámetro $M(n, \xi)$ según la fórmula (10) puede realizarse para importantes casos especiales de funciones de ventana $w(t)$, tales como la ventana gaussiana y una ventana de seno. Se observa que la aproximación deseada a $\hat{w}(n\pi - (T\xi + r\Omega))$ es muy buena para varias subbandas con $n\pi \approx T\xi + r\Omega$.

- 50 Debe observarse que la relación (11) está calibrada para una situación a modo de ejemplo en la que el banco de filtros de análisis 301 tiene una separación entre subbandas de frecuencia angulares de π/T . En el caso general, la interpretación resultante de (11) es que el espacio origen de términos cruzados $p_1 + p_2$ es un entero que se

aproxima a la frecuencia fundamental subyacente Ω , medida en unidades de la separación entre subbandas del banco de filtro de análisis, y que el par (p_1, p_2) se elige como un múltiplo de $(r, T-r)$.

Para la determinación del par de desplazamientos de índice (p_1, p_2) en el descodificador, pueden usarse los siguientes modos:

1. Un valor de Ω puede obtenerse en el proceso de codificación y transmitirse explícitamente al descodificador con una precisión suficiente para obtener los valores enteros de p_1 y p_2 mediante un procedimiento de redondeo adecuado, que puede seguir los principios de que:

◦ $p_1 + p_2$ se aproxima a $\Omega/\Delta\omega$, donde $\Delta\omega$ es la separación entre frecuencias angulares del banco de filtros de análisis; y

◦ p_1 / p_2 se elige para que se aproxime a $r/(T-r)$.

2. Para cada muestra de subbanda objetivo, el par de desplazamientos de índice (p_1, p_2) puede obtenerse en el descodificador a partir de una lista predeterminada de valores candidatos, tal como $(p_1, p_2) = (rl, (T-r)l)$, $l \in L$, $r \in \{1, 2, \dots, T-1\}$, donde L es una lista de enteros positivos. La selección puede basarse en una optimización de magnitud de salida de términos cruzados, por ejemplo una maximización de la energía de la salida de términos cruzados.

3. Para cada muestra de subbanda objetivo, el par de desplazamientos de índice (p_1, p_2) puede obtenerse a partir de una lista reducida de valores candidatos mediante una optimización de magnitud de salida de términos cruzados, donde la lista reducida de valores candidatos se obtiene en el proceso de codificación y se transmite al descodificador.

Debe observarse que la modificación de fase de las señales de subbanda u_1 y u_2 se lleva a cabo con una ponderación $(T-r)$ y r , respectivamente, pero la distancia de índice de subbanda p_1 y p_2 se eligen de manera proporcional a r y $(T-r)$, respectivamente. Por tanto, la subbanda más cercana a la subbanda de síntesis n recibe la modificación de fase más significativa.

Un procedimiento ventajoso para el procedimiento de optimización para los modos 2 y 3 descritos anteriormente puede ser considerar la optimización de máximos y mínimos:

$$\max\{\min\{|x_{n-p_1}(k)|, |x_{n+p_2}(k)|\} : (p_1, p_2) = (rl, (T-r)l), l \in L, r \in \{1, 2, \dots, T-1\}\}, \quad (12)$$

y usar el par ganador junto con su valor correspondiente de r para generar la contribución de producto cruzado para un índice de subbanda objetivo n dado. En el modo 2 orientado a la búsqueda en el descodificador, y también parcialmente en el 3, la suma de los términos cruzados para diferentes valores r se realiza preferentemente de manera independiente, ya que puede haber riesgo de añadir contenido a la misma subbanda varias veces. Por otro lado, si se usa la frecuencia fundamental Ω para seleccionar las subbandas, como en el modo 1, o si solo se permite un pequeño intervalo de distancias de índice de subbanda, como puede ser el caso del modo 2, este problema particular de añadir contenido varias veces a la misma subbanda puede evitarse.

Además, debe observarse que en las realizaciones de los esquemas de procesamiento de términos cruzados descritos anteriormente, una modificación adicional de descodificador de la ganancia de producto cruzado g puede ser beneficiosa. Por ejemplo, se hace referencia a las señales de subbanda de entrada u_1, u_2 de la unidad MISO de productos cruzados según la fórmula (2) y a la señal de subbanda de entrada x de la unidad SISO de transposición según la fórmula (1). Si estas tres señales van a introducirse en la misma subbanda de síntesis de salida como se muestra en la Fig. 4, donde el procesamiento directo 401 y el procesamiento de producto cruzado 402 proporcionan componentes para la misma subbanda de síntesis de salida, puede ser deseable fijar la ganancia de producto cruzado g a cero, es decir, la unidad de ganancia 902 de la Fig. 9, si

$$\min(|u_1|, |u_2|) < q|x|, \quad (13)$$

para un umbral predefinido $q > 1$. Dicho de otro modo, la suma de productos cruzados solo se lleva a cabo si la magnitud de subbanda de entrada de términos directos $|x|$ es pequeña en comparación con ambos términos de entrada de producto cruzado. En este contexto, x es la muestra de subbanda de análisis para el procesamiento de términos directos que da lugar a una salida en la misma subbanda de síntesis que el producto cruzado en consideración. Esto puede ser una precaución para no mejorar adicionalmente una componente armónica que ya

se haya optimizado mediante la transposición directa.

El procedimiento de transposición armónica explicado a grandes rasgos en el presente documento se describirá a continuación para configuraciones espectrales a modo de ejemplo con el fin de ilustrar las mejoras con respecto a la técnica anterior. La Fig. 10 ilustra el efecto de una transposición armónica directa de orden $T = 2$. El diagrama superior 1001 ilustra las componentes de frecuencia parciales de la señal original mediante flechas verticales situadas en múltiplos de la frecuencia fundamental Ω . Ilustra la señal origen, por ejemplo en el lado del codificador. El diagrama 1001 está segmentado en un intervalo de frecuencias origen en el lado izquierdo con las frecuencias parciales Ω , 2Ω , 3Ω , 4Ω , 5Ω y en un intervalo de frecuencias objetivo en el lado derecho con frecuencias parciales 6Ω , 7Ω , 8Ω . El intervalo de frecuencias origen se codificará y se transmitirá normalmente al descodificador. Por otro lado, el intervalo de frecuencias objetivo en el lado derecho, que comprende las frecuencias parciales 6Ω , 7Ω , 8Ω superiores a la frecuencia de cruce 1005 del procedimiento HFR, normalmente no se transmitirán al descodificador. Un objeto del procedimiento de transposición armónica es reconstruir el intervalo de frecuencias objetivo por encima de la frecuencia de cruce 1005 de la señal origen a partir del intervalo de frecuencias origen. Por consiguiente, el intervalo de frecuencias objetivo, y especialmente las frecuencias parciales 6Ω , 7Ω , 8Ω del diagrama 1001 no están disponibles como entradas en el transposicionador.

Como se ha indicado anteriormente, el procedimiento de transposición armónica tiene como objetivo regenerar las componentes de señal 6Ω , 7Ω , 8Ω de la señal origen a partir de las componentes de frecuencia disponibles en el intervalo de frecuencias origen. El diagrama inferior 1002 muestra la salida del transposicionador en el intervalo de frecuencias objetivo del lado derecho. Tal transposicionador puede estar situado, por ejemplo, en el lado del descodificador. Las frecuencias parciales 6Ω y 8Ω se regeneran a partir de las frecuencias parciales 3Ω y 4Ω mediante una transposición armónica que usa un orden de transposición $T = 2$. Como resultado de un efecto de ensanchamiento espectral de la transposición armónica, ilustrado aquí mediante las flechas de puntos 1003 y 1004, falta la frecuencia parcial objetivo 7Ω . Esta frecuencia parcial objetivo 7Ω no puede generarse usando el procedimiento de transposición armónica subyacente de la técnica anterior.

La Figura 11 ilustra el efecto de la invención en una transposición armónica de una señal periódica en un caso en el que un transposicionador armónico de segundo orden se ha mejorado mediante un único término cruzado, es decir, $T = 2$ y $r = 1$. Como se ha mencionado en el contexto de la Fig. 10, un transposicionador se usa para generar las frecuencias parciales 6Ω , 7Ω , 8Ω del intervalo de frecuencias objetivo por encima de la frecuencia de cruce 1105 del diagrama inferior 1102 a partir de las frecuencias parciales Ω , 2Ω , 3Ω , 4Ω , 5Ω del intervalo de frecuencias origen inferior a la frecuencia de cruce 1105 del diagrama 1101. Además de la salida del transposicionador de la técnica anterior de la Figura 10, la componente de frecuencia parcial 7Ω se regenera a partir de una combinación de las frecuencias parciales origen 3Ω y 4Ω . El efecto de la suma de productos cruzados se ilustra mediante las flechas discontinuas 1103 y 1104. En lo que respecta a las fórmulas, una tiene $\omega = 3\Omega$ y, por lo tanto, $(T-r)\omega + r(\omega+\Omega) = T\omega + r\Omega = 6\Omega + \Omega = 7\Omega$. Como puede observarse en este ejemplo, todas las frecuencias parciales objetivo pueden regenerarse usando el procedimiento HFR inventivo descrito en el presente documento.

La Fig. 12 ilustra una posible implementación de un transposicionador armónico de segundo orden de la técnica anterior en un banco de filtros modulado para la configuración espectral de la Fig. 10. Las respuestas de frecuencia estilizadas de las subbandas de banco de filtros de análisis se muestran mediante líneas de puntos, por ejemplo el signo de referencia 1206, en la diagrama superior 1201. Las subbandas están enumeradas mediante el índice de subbanda, mostrándose los índices 5, 10 y 15 en la Fig. 12. Para el ejemplo dado, la frecuencia fundamental Ω es igual a 3,5 veces la separación entre frecuencias de subbanda de análisis. Esto se ilustra con el hecho de que la frecuencia parcial Ω del diagrama 1201 está situada entre las dos subbandas con índice de subbanda 3 y 4. La frecuencia parcial 2Ω está situada en el centro de la subbanda con índice de subbanda 7, etc.

El diagrama inferior 1202 muestra las frecuencias parciales regeneradas 6Ω y 8Ω superpuestas con las respuestas de frecuencia estilizadas, por ejemplo el signo de referencia 1207, de subbandas de banco de filtros de síntesis seleccionadas. Como se ha descrito anteriormente, estas subbandas tienen una separación entre frecuencias $T = 2$ veces mayor. Por consiguiente, las respuestas de frecuencia también se escalan en el factor $T = 2$. Tal y como se ha mencionado anteriormente, el procedimiento de procesamiento de términos directos de la técnica anterior modifica la fase de cada subbanda de análisis, es decir, de cada subbanda inferior a la frecuencia de cruce 1205 del diagrama 1201, en un factor $T = 2$, y correlaciona el resultado con la subbanda de síntesis de mismo índice, es decir, una subbanda superior a la frecuencia de cruce 1205 del diagrama 1202. Esto se simboliza en la Fig. 12 mediante flechas de puntos en diagonal, por ejemplo la flecha 1208 para la subbanda de análisis 1206 y la subbanda de síntesis 1207. El resultado de este procesamiento de términos directos para subbandas con índices de subbanda 9 a 16 de la subbanda de análisis 1201 es la regeneración de las dos frecuencias parciales objetivo 6Ω y 8Ω en la subbanda de síntesis 1202 a partir de las frecuencias parciales origen 3Ω y 4Ω . Como puede observarse en la Fig. 12, la contribución principal a la frecuencia parcial objetivo 6Ω

proviene de las subbandas con los índices de subbanda 10 y 11, es decir, las señales de referencia 1209 y 1210, y la contribución principal a la frecuencia parcial objetivo 8Ω proviene de la subbanda con índice de subbanda 14, es decir, el signo de referencia 1211.

- 5 La Fig. 13 ilustra una posible implementación de una etapa de procesamiento adicional de términos cruzados en el banco de filtros modulado de la Fig. 12. La etapa de procesamiento de términos cruzados corresponde a la descrita para señales periódicas con la frecuencia fundamental Ω en relación con la Fig. 11. El diagrama superior 1301 ilustra las subbandas de análisis, cuyo intervalo de frecuencias origen va a transponerse al intervalo de frecuencias objetivo de las subbandas de síntesis del diagrama inferior 1302. Se considera el caso particular de
10 generación de las subbandas de síntesis 1315 y 1316, que rodean a la frecuencia parcial 7Ω , a partir de las subbandas de análisis. Para un orden de transposición $T = 2$, puede seleccionarse un posible valor de $r = 1$. Elegir la lista de valores candidatos (p_1, p_2) como un múltiplo de $(r, T-r) = (1, 1)$ de manera que $p_1 + p_2$ se

$\frac{\Omega}{\Delta\omega} = \frac{\Omega}{(\Omega/3.5)} = 3.5$,
aproxima a $\frac{\Omega}{\Delta\omega} = \frac{\Omega}{(\Omega/3.5)} = 3.5$, es decir, la frecuencia fundamental Ω en unidades de la separación entre frecuencias de subbanda de análisis, da lugar a la elección de $p_1 = p_2 = 2$. Como se indicó en el contexto de la
15 Fig. 8, una subbanda de síntesis con el índice de subbanda n puede generarse a partir del producto de términos cruzados de las subbandas de análisis con el índice de subbanda $(n - p_1)$ y $(n + p_2)$. Por consiguiente, para la subbanda de síntesis con índice de subbanda 12, es decir, el signo de referencia 1315, un producto cruzado se forma a partir de las subbandas de análisis con índice de subbanda $(n - p_1) = 12 - 2 = 10$, es decir, el signo de referencia 1311, y $(n + p_2) = 12 + 2 = 14$, es decir, el signo de referencia 1313. Para la subbanda de síntesis con
20 índice de subbanda 13, un producto cruzado se forma a partir de las subbandas de análisis con índice $(n - p_1) = 13 - 2 = 11$, es decir, el signo de referencia 1312, y $(n + p_2) = 13 + 2 = 15$, es decir, el signo de referencia 1314. Este proceso de generación de productos cruzados se simboliza mediante los pares de flechas discontinuas/de puntos en diagonal, es decir los pares de signos de referencia 1308, 1309 y 1306, 1307, respectivamente.

- 25 Como puede observarse en la Fig. 13, la frecuencia parcial 7Ω está situada principalmente en la subbanda 1315 con índice 12 y solo de manera secundaria en la subbanda 1316 con índice 13. Por consiguiente, para respuestas de filtro más realistas, habrá más términos directos y/o cruzados en torno a la subbanda de síntesis 1315 con índice 12, que se suman de manera beneficiosa a la síntesis de un senoide de alta calidad de frecuencia $(T-r)\omega + r(\omega+\Omega) = T\omega + r\Omega = 6\Omega + \Omega = 7\Omega$, que términos en torno a la subbanda de síntesis 1316 con
30 índice 13. Además, tal y como se ha señalado en el contexto de la fórmula (13), una suma ciega de todos los términos cruzados con $p_1 = p_2 = 2$ puede dar lugar a componentes de señal no deseadas para señales de entrada académicas y menos periódicas. Por consiguiente, este fenómeno de componentes de señal no deseadas puede requerir la aplicación de una regla de cancelación adaptativa de producto cruzado, tal como la regla proporcionada por la fórmula (13).

- 35 La Fig. 14 ilustra el efecto de transposición armónica de orden $T = 3$ de la técnica anterior. El diagrama superior 1401 ilustra las componentes de frecuencia parciales de la señal original mediante flechas verticales situadas en múltiplos de la frecuencia fundamental Ω . Las frecuencias parciales 6Ω , 7Ω , 8Ω , 9Ω están en el intervalo objetivo superior a la frecuencia de cruce 1405 del procedimiento HFR y, por lo tanto, no están disponibles como
40 entradas en el transposicionador. El objetivo de la transposición armónica es regenerar esas componentes de señal a partir de la señal del intervalo origen. El diagrama inferior 1402 muestra la salida del transposicionador en el intervalo de frecuencias objetivo. Las frecuencias parciales 6Ω , es decir, el signo de referencia 1407, y 9Ω , es decir, el signo de referencia 1410, se han regenerado a partir de las frecuencias parciales 2Ω , es decir, el signo de referencia 1406, y 3Ω , es decir, el signo de referencia 1409. Como resultado de un efecto de ensanchamiento
45 espectral de la transposición armónica, ilustrado aquí mediante las flechas de puntos 1408 y 14011, respectivamente, faltan las frecuencias objetivo 7Ω y 8Ω .

- La Fig. 15 ilustra el efecto de la invención para la transposición armónica de una señal periódica en un caso en el que un transposicionador armónico de tercer orden se ha mejorado mediante la suma de dos términos cruzados
50 diferentes, es decir, $T = 3$ y $r = 1,2$. Además de la salida de transposicionador de la técnica anterior de la Fig. 14, la componente de frecuencia parcial 7Ω , 1508, se regenera mediante el término cruzado para $r = 1$ a partir de una combinación de las frecuencias parciales origen 2Ω , 1506, y 3Ω , 1507. El efecto de la suma de productos cruzados se ilustra mediante las flechas discontinuas 1510 y 1511. En lo que respecta a las fórmulas, una tiene $\omega = 2\Omega$, $(T-r)\omega + r(\omega+\Omega) = T\omega + r\Omega = 6\Omega + \Omega = 7\Omega$. Asimismo, la componente de frecuencia parcial 8Ω , 1509, se regenera mediante el término cruzado para $r = 2$. Esta componente de frecuencia parcial 1509 en el intervalo
55 objetivo del diagrama inferior 1502 se genera a partir de las componentes de frecuencia parciales 2Ω , 1506, y 3Ω , 1507, en el intervalo de frecuencias origen del diagrama superior 1501. La generación del producto de términos cruzados se ilustra mediante las flechas 1512 y 1513. En lo que respecta a las fórmulas, una tiene $(T-r)\omega + r(\omega+\Omega) = T\omega + r\Omega = 6\Omega + 2\Omega = 8\Omega$. Como puede observarse, todas las frecuencias parciales objetivo
60 pueden regenerarse usando el procedimiento HFR inventivo descrito en el presente documento.

La Fig. 16 ilustra una posible implementación de un transposicionador armónico de tercer orden en un banco de filtros modulado para la situación espectral de la Fig. 14. Las respuestas de frecuencia estilizadas de las

subbandas de banco de filtros de análisis se muestran mediante líneas de puntos en el diagrama superior 1601. Las subbandas están enumeradas mediante los índices de subbanda 1 a 17, cuyas subbandas 1606, con índice 7, 1607, con índice 10, y 1608, con índice 11, se señalan a modo de ejemplo. Para el ejemplo dado, la frecuencia fundamental Ω es igual a 3,5 veces la separación entre frecuencias de subbanda de análisis $\Delta\omega$. El diagrama inferior 1602 muestra la frecuencia parcial regenerada superpuesta a las respuestas de frecuencia estilizadas de las subbandas seleccionadas de banco de filtros de síntesis. A modo de ejemplo se hace referencia a las subbandas 1609, con índice de subbanda 7, 1610, con índice de subbanda 10, y 1611, con índice de subbanda 11. Tal y como se ha descrito anteriormente, estas subbandas tienen una separación entre frecuencias $\Delta\omega$ que es $T = 3$ veces mayor. Por consiguiente, las respuestas de frecuencia también se escalan de manera correspondiente.

El procesamiento de términos directos de la técnica anterior modifica la fase de las señales de subbanda en un factor $T = 3$ para cada subbanda de análisis y correlaciona el resultado con la subbanda de síntesis de mismo índice, como se indica mediante las flechas de puntos en diagonal. El resultado de este procesamiento de términos directos para las subbandas 6 a 11 es la regeneración de las dos frecuencias parciales objetivo 6Ω y 9Ω a partir de las frecuencias parciales origen 2Ω y 3Ω . Como puede observarse en la Fig. 16, la contribución principal a la frecuencia parcial objetivo 6Ω proviene de la subbanda con índice 7, es decir, el signo de referencia 1606, y las contribuciones principales a la frecuencia parcial objetivo 9Ω proviene de subbandas con índice 10 y 11, es decir, los signos de referencia 1607 y 1608, respectivamente.

La Fig. 17 ilustra una posible implementación de una etapa de procesamiento adicional de términos cruzados para $r = 1$ en el banco de filtros modulado de la Fig. 16 que da lugar a la regeneración de la frecuencia parcial 7Ω . Como se mencionó en el contexto de la Fig. 8, los desplazamientos de índice (p_1, p_2) pueden seleccionarse como un múltiplo de $(r, T-r) = (1, 2)$, de manera que $p_1 + p_2$ se aproxima a 3,5, es decir, la frecuencia fundamental Ω en unidades de la separación entre frecuencias de subbanda de análisis $\Delta\omega$. Dicho de otro modo, la distancia relativa, es decir, la distancia en el eje de frecuencia dividida por la separación entre frecuencias de subbanda de análisis $\Delta\omega$, entre las dos subbandas de análisis que contribuyen a la subbanda de síntesis que va a generarse, debe aproximarse lo más posible a la frecuencia fundamental relativa, es decir, la frecuencia fundamental Ω dividida por la separación entre frecuencias de subbanda de análisis $\Delta\omega$. Esto también se expresa mediante la fórmula (11) y da lugar a la elección de $p_1 = 1, p_2 = 2$.

Tal y como se muestra en la Fig. 17, la subbanda de síntesis con índice 8, es decir, el signo de referencia 1710, se obtiene a partir de un producto cruzado formado a partir de las subbandas de análisis con índice $(n - p_1) = 8 - 1 = 7$, es decir, el signo de referencia 1706, y $(n + p_2) = 8 + 2 = 10$, es decir, el signo de referencia 1708. Para la subbanda de síntesis con índice 9, un producto cruzado se forma a partir de subbandas de análisis con índice $(n - p_1) = 9 - 1 = 8$, es decir, el signo de referencia 1707, y $(n + p_2) = 9 + 2 = 11$, es decir, el signo de referencia 1709. Este proceso de formación de productos cruzados se representa mediante los pares de flechas discontinuas/de puntos en diagonal, es decir, los pares de flechas 1712, 1713 y 1714, 1715, respectivamente. Puede observarse en la Fig. 17 que la frecuencia parcial 7Ω está situada de manera más prominente en la subbanda 1710 que en la subbanda 1711. Por consiguiente, se espera que para respuestas de filtro realistas haya más términos cruzados en torno a la subbanda de síntesis con índice 8, es decir, la subbanda 1710, que se suman de manera beneficiosa a la síntesis de un senoide de alta calidad de frecuencia $(T-r)\omega + r(\omega+\Omega) = T\omega + r\Omega = 6\Omega + \Omega = 7\Omega$.

La Fig. 18 ilustra una posible implementación de una etapa de procesamiento adicional de términos cruzados para $r = 2$ en el banco de filtros modulado de la Fig. 16, que da lugar a la regeneración de la frecuencia parcial 8Ω . Los desplazamientos de índice (p_1, p_2) pueden seleccionarse como un múltiplo de $(r, T-r) = (2, 1)$, de manera que $p_1 + p_2$ se aproxima a 3,5, es decir, la frecuencia fundamental Ω en unidades de la separación entre frecuencias de subbanda de análisis $\Delta\omega$. Esto da lugar a la elección de $p_1 = 2, p_2 = 1$. Como se muestra en la Fig. 18, la subbanda de síntesis con índice 9, es decir, el signo de referencia 1810, se obtiene a partir de un producto cruzado formado a partir de las subbandas de análisis con índice $(n - p_1) = 9 - 2 = 7$, es decir, el signo de referencia 1806, y $(n + p_2) = 9 + 1 = 10$, es decir, el signo de referencia 1808. Para la subbanda de síntesis con índice 10, un producto cruzado se forma a partir de subbandas de análisis con índice $(n - p_1) = 10 - 2 = 8$, es decir, el signo de referencia 1807, y $(n + p_2) = 10 + 1 = 11$, es decir, el signo de referencia 1809. Este proceso de formación de productos cruzados se representa mediante los pares de flechas discontinuas/de puntos en diagonal, es decir los pares de flechas 1812, 1813 y 1814, 1815, respectivamente. Puede observarse en la Fig. 18 que la frecuencia parcial 8Ω está situada ligeramente de manera más prominente en la subbanda 1810 que en la subbanda 1811. Por consiguiente, se espera que para respuestas de filtro realistas haya más términos directos y/o cruzados en torno a la subbanda de síntesis con índice 9, es decir, la subbanda 1810, que se suman de manera beneficiosa a la síntesis de un senoide de alta calidad de frecuencia $(T-r)\omega + r(\omega+\Omega) = T\omega + r\Omega = 2\Omega + 6\Omega = 8\Omega$.

A continuación se hace referencia a las Figuras 23 y 24, que ilustran el procedimiento de selección basado en la optimización de máximos y mínimos (12) para el par de desplazamientos de índice (p_1, p_2) y r según esta regla

para $T = 3$. El índice de subbanda objetivo elegido es $n = 18$ y el diagrama superior ilustra un ejemplo de la magnitud de una señal de subbanda para un índice de tiempo dado. La lista de enteros positivos viene dada en este caso mediante los siete valores de $L = \{2, 3, \dots, 8\}$.

5 La Fig. 23 ilustra la búsqueda de candidatos con $r = 1$. La subbanda objetivo o de síntesis se muestra con el índice $n = 18$. La línea de puntos 2301 resalta la subbanda con índice $n = 18$ en el intervalo de subbandas de análisis superior y el intervalo de subbandas de síntesis inferior. Los posibles pares de desplazamientos de índice son $(p_1, p_2) = \{(2, 4), (3, 6), \dots, (8, 16)\}$, para $l = 2, 3, \dots, 8$, respectivamente, y los pares de índices correspondientes de muestra de magnitud de subbanda de análisis, es decir, la lista de pares de índices de subbanda que se consideran para determinar el término cruzado óptimo son $\{(16, 22), (15, 24), \dots, (10, 34)\}$. El conjunto de flechas ilustra los pares en cuestión. Como un ejemplo, se muestra el par $(15, 24)$ denotado mediante los signos de referencia 2302 y 2303. Calculando el mínimo de estos pares de magnitudes se obtiene la lista $(0, 4, 1, 0, 0, 0, 0)$ de magnitudes mínimas respectivas para la posible lista de términos cruzados. Puesto que la segunda entrada para $l = 3$ es máxima, el par $(15, 24)$ gana entre los candidatos con $r = 1$, y esta selección se ilustra mediante las flechas gruesas.

Por otro lado, la Fig. 24 ilustra la búsqueda de candidatos con $r = 2$. La subbanda objetivo o de síntesis se muestra con el índice $n = 18$. La línea de puntos 2401 resalta la subbanda con índice $n = 18$ en el intervalo de subbandas de análisis superior y el intervalo de subbandas de síntesis inferior. En este caso, los posibles pares de desplazamientos de índice son $(p_1, p_2) = \{(4, 2), (6, 3), \dots, (16, 8)\}$ y los pares de índices correspondientes de muestra de magnitud de subbanda de análisis son $\{(14, 20), (12, 21), \dots, (2, 26)\}$, cuyo par $(6, 24)$ está representado mediante los signos de referencia 2402 y 2403. Calculando el mínimo de estos pares de magnitudes se obtiene la lista $(0, 0, 0, 0, 3, 1, 0)$. Puesto que la quinta entrada es máxima, es decir, $l = 6$, el par $(6, 24)$ gana entre los candidatos con $r = 2$, como se ilustra mediante las flechas gruesas. En términos generales, puesto que el mínimo del par de magnitudes correspondiente es más pequeño que el del par de subbandas seleccionadas para $r = 1$, la selección final del índice de subbanda objetivo $n = 18$ es el par $(15, 24)$ y $r = 1$.

Debe observarse además que cuando la señal de entrada $z(t)$ es una serie armónica con una frecuencia fundamental Ω , es decir, con una frecuencia fundamental que corresponde al parámetro de tono de mejora de producto cruzado, y Ω es suficientemente grande en comparación con la resolución de frecuencia del banco de filtros de análisis, las señales de subbanda de análisis $x_n(k)$ dadas por la fórmula (6) y $x'_n(k)$ dadas por la fórmula (8) son buenas aproximaciones del análisis de la señal de entrada $z(t)$, donde la aproximación es válida en diferentes regiones de subbanda. A partir de una comparación de las fórmulas (6) y (8 a 10) se deduce que una evolución de fase armónica a lo largo del eje de frecuencia de la señal de entrada $z(t)$ se extrapolará correctamente mediante la presente invención. Esto se cumple, en particular, para un tren de impulsos puros. Para la calidad de audio de salida, esto es una característica atractiva para señales a modo de tren de impulsos, tales como las producidas por las voces humanas y algunos instrumentos musicales.

Las figuras 25, 26 y 27 ilustran el rendimiento de una implementación a modo de ejemplo de la transposición inventiva para una señal armónica en el caso de $T = 3$. La señal tiene una frecuencia fundamental de 282,35 Hz y su espectro de magnitud en el intervalo objetivo considerado de 10 a 15 kHz se ilustra en la Fig. 25. Un banco de filtros de $N = 512$ subbandas se usa a una frecuencia de muestreo de 48 kHz para implementar las transposiciones. El espectro de magnitud de la salida de un transposicionador directo de tercer orden ($T = 3$) se ilustra en la Fig. 26. Como puede observarse, cada tercer armónico se reproduce con alta fidelidad como se predijo a través de la teoría descrita anteriormente, y el tono percibido será de 847 Hz, tres veces el original. La Fig. 27 muestra la salida de un transposicionador que aplica productos de términos cruzados. Todos los armónicos se han recreado con imperfecciones debido a los aspectos aproximativos de la teoría. En este caso, las curvas laterales están 40 dB por debajo del nivel de la señal y esto es más que suficiente para la regeneración de contenido de alta frecuencia, el cual no puede distinguirse, desde un punto de vista perceptivo, de la señal armónica original.

A continuación se hace referencia a la Fig. 28 y a la Fig. 29, que ilustran un codificador 2800 a modo de ejemplo y un decodificador 2900 a modo de ejemplo, respectivamente, para una codificación unificada de voz y audio (USAC). A continuación se describe la estructura general del codificador 2800 y del decodificador 2900 USAC: en primer lugar, puede haber un pre/posprocesamiento común que consiste en una unidad funcional *MPEG Surround* (MPEGS) para tratar el procesamiento estéreo o multicanal, y una unidad de SBR mejorada (eSBR) 2801 y 2901, respectivamente, que trata la representación paramétrica de las frecuencias de audio más altas de la señal de entrada y que puede usar los procedimientos de transposición armónica descritos en el presente documento. Por otro lado, hay dos bifurcaciones, donde una consiste en una trayectoria de herramienta de codificación de audio avanzada (AAC) modificada y la otra consiste en una trayectoria basada en la codificación de predicción lineal (dominio LP o LPC), que a su vez incluye una representación en el dominio de frecuencia o una representación en el dominio de tiempo de la señal residual LPC. Todos los espectros transmitidos para AAC y LPC pueden representarse en el dominio MDCT después de la cuantificación y la codificación aritmética. La representación en el dominio de tiempo usa un esquema de codificación de excitación ACELP.

La unidad de replicación de banda espectral mejorada (eSBR) 2801 del codificador 2800 puede comprender los sistemas de reconstrucción de alta frecuencia descritos en el presente documento. En particular, la unidad eSBR 2801 puede comprender un banco de filtros de análisis 301 para generar una pluralidad de señales de subbanda de análisis. Estas señales de subbanda de análisis pueden transponerse después en una unidad de procesamiento no lineal 302 para generar una pluralidad de señales de subbanda de síntesis que pueden introducirse después en un banco de filtros de síntesis 303 para generar una componente de alta frecuencia. En la unidad eSBR 2801, en el lado de codificación, puede determinarse un conjunto de información acerca de cómo generar una componente de alta frecuencia a partir de la componente de baja frecuencia que se ajuste mejor a la componente de alta frecuencia de la señal original. Este conjunto de información puede comprender información acerca de características de la señal, tal como una frecuencia fundamental predominante Ω , acerca de la envolvente espectral de la componente de alta frecuencia, y puede comprender información acerca de cómo combinar de manera óptima señales de subbanda de análisis, es decir, información tal como un conjunto limitado de pares de desplazamientos de índice (p_1 , p_2). Datos codificados relacionados con este conjunto de información se fusionan con la otra información codificada en un multiplexor de flujo de bits y se reenvían como un flujo de audio codificado a un descodificador 2900 correspondiente.

El descodificador 2900 mostrado en la Fig. 29 comprende además una unidad de replicación de ancho de banda espectral mejorada (eSBR) 2901. Esta unidad eSBR 2901 recibe el flujo de bits de audio codificado o la señal codificada desde el codificador 2800 y usa los procedimientos descritos en el presente documento para generar una componente de alta frecuencia de la señal, que se fusiona con la componente de baja frecuencia descodificada para proporcionar una señal descodificada. La unidad eSBR 2901 puede comprender los diferentes componentes descritos en el presente documento. En particular, puede comprender un banco de filtros de análisis 301, una unidad de procesamiento no lineal 302 y un banco de filtros de síntesis 303. La unidad eSBR 2901 puede usar información acerca de la componente de alta frecuencia proporcionada por el codificador 2800 para llevar a cabo la reconstrucción de altas frecuencias. Tal información puede ser una frecuencia fundamental Ω de la señal, la envolvente espectral de la componente de alta frecuencia original y/o información acerca de las subbandas de análisis que van a usarse para generar las señales de subbanda de síntesis y, en última instancia, la componente de alta frecuencia de la señal descodificada.

Además, las Fig. 28 y 29 ilustran posibles componentes adicionales de un codificador/descodificador USAC, tales como:

- una herramienta de desmultiplexor de datos útiles de flujo de bits, que separa los datos útiles del flujo de bits en partes para cada herramienta y proporciona a cada una de las herramientas información de datos útiles de flujo de bits relacionada con esa herramienta;
- una herramienta de descodificación sin ruido de factor de escala, que toma información del desmultiplexor de datos útiles de flujo de bits, analiza sintácticamente la información y descodifica los factores de escala codificados DPCM y de Huffman;
- una herramienta de descodificación espectral sin ruido, que toma información del desmultiplexor de datos útiles de flujo de bits, analiza sintácticamente la información, descodifica los datos codificados de manera aritmética y reconstruye los espectros cuantificados;
- una herramienta de cuantificador inverso, que toma los valores cuantificados para los espectros y convierte los valores enteros en los espectros reconstruidos no escalados; este cuantificador es preferentemente un cuantificador de compresión-expansión cuyo factor de compresión-expansión depende del modo de codificación principal elegido;
- una herramienta de llenado con ruido, que se usa para rellenar huecos espectrales de los espectros descodificados que se producen cuando valores espectrales se cuantifican a cero debido a, por ejemplo, una fuerte restricción en la demanda de bits en el codificador;
- una herramienta de reescalado, que convierte la representación de números enteros de los factores de escala en los valores definitivos y que multiplica los espectros no escalados y cuantificados de manera inversa por los factores de escala pertinentes;
- una herramienta M/S, como la descrita en la especificación ISO/IEC 14496-3;
- una herramienta de conformación de ruido temporal (TNS), como la descrita en la especificación ISO/IEC 14496-3;
- una herramienta de banco de filtros/conmutación de bloques, que aplica la inversa de la correlación de frecuencias que se llevó a cabo en el codificador; una transformada de coseno discreta modificada inversa (IMDCT) se usa preferentemente para la herramienta de banco de filtros;

- una herramienta de banco de filtros degradado en el tiempo/conmutación de bloques, que sustituye a la herramienta habitual de banco de filtros/conmutación de bloques cuando se habilita el modo de degradación en el tiempo; el banco de filtros es preferentemente el mismo (IMDCT) que para el banco de filtros habitual; adicionalmente, las muestras de dominio de tiempo en ventanas se correlacionan desde el dominio de tiempo degradado al dominio de tiempo lineal mediante un remuestreo variable en el tiempo;
- una herramienta *MPEG Surround* (MPEGs), que produce múltiples señales a partir de una o más señales de entrada aplicando un procedimiento de mezclado ascendente sofisticado al (a las) señal(es) de entrada controlada(s) por parámetros espaciales apropiados; en el contexto de la USAC se usa preferentemente MEGPS para codificar una señal multicanal, transmitiendo información complementaria paramétrica junto con una señal transmitida mezclada de manera descendente;
- una herramienta clasificadora de señales, que analiza la señal de entrada original y genera a partir de la misma información de control que activa la selección de los diferentes modos de codificación; el análisis de la señal de entrada depende normalmente de la implementación y trata de elegir el modo de codificación principal óptimo para una trama de señal de entrada dada; la salida del clasificador de señales también puede usarse opcionalmente para influir en el comportamiento de otras herramientas, por ejemplo *MEGP Surround*, SBR mejorada, banco de filtros degradado en el tiempo y otras;
- una herramienta de filtros LPC, que produce una señal en el dominio de tiempo a partir de una señal de dominio de excitación filtrando la señal de excitación reconstruida a través de un filtro de síntesis de predicción lineal; y
- una herramienta ACELP, que proporciona una manera de representar de manera eficaz una señal de excitación en el dominio de tiempo combinando un elemento predictivo a largo plazo (palabra de código adaptativa) con una secuencia a modo de impulsos (palabra de código innovadora).

La Fig. 30 ilustra una realización de las unidades eSBR mostradas en las Fig. 28 y 29. A continuación se describirá la unidad eSBR 3000 en el contexto de un descodificador, donde la entrada en la unidad eSBR 3000 es la componente de baja frecuencia, también conocida como la banda baja, de una señal y posible información adicional relacionada con características de señal específicas, tales como la frecuencia fundamental Ω y/o posibles valores de desplazamiento de índice (p_1 , p_2). En el lado del codificador, la entrada en la unidad eSBR será normalmente la señal completa, mientras que la salida será información adicional relacionada con las características de señal y/o los valores de desplazamiento de índice.

En la Fig. 30, la componente de baja frecuencia 3013 se introduce en un banco de filtros QMF para generar bandas de frecuencia QMF. Estas bandas de frecuencia QMF no deben confundirse con las subbandas de análisis descritas en este documento. Las bandas de frecuencia QMF se usan con el fin de manipular y fusionar las componentes de baja frecuencia y de alta frecuencia de la señal en el dominio de frecuencia en lugar de en el dominio de tiempo. La componente de baja frecuencia 3014 se introduce en la unidad de transposición 3004, que corresponde a los sistemas de reconstrucción de altas frecuencias descritos en el presente documento. La unidad de transposición 3004 también puede recibir información adicional 3011, tal como la frecuencia fundamental Ω de la señal codificada y/o posibles pares de desplazamientos de índice (p_1 , p_2) para la selección de subbanda. La unidad de transposición 3004 genera una componente de alta frecuencia 3012, también conocida como banda alta, de la señal, que se transforma en el dominio de frecuencia mediante un banco de filtros QMF 3003. Tanto la componente de baja frecuencia transformada QMF como la componente de alta frecuencia transformada QMF se introducen en una unidad de manipulación y mezcla 3005. Esta unidad 3005 puede llevar a cabo un ajuste de envolvente de la componente de alta frecuencia y combina la componente de alta frecuencia ajustada y la componente de baja frecuencia. La señal de salida combinada vuelve a transformarse al dominio de tiempo mediante un banco de filtros QMF inversos 3001.

Normalmente, los bancos de filtros QMF comprenden 64 bandas de frecuencia QMF. Sin embargo, debe observarse que puede ser beneficioso muestrear de manera descendente la componente de baja frecuencia 3013, de manera que el banco de filtros QMF 3002 solo necesita 32 bandas de frecuencia QMF. En tales casos, la componente de baja frecuencia 3013 tiene un ancho de banda de $f_s/4$, donde f_s es la frecuencia de muestreo de la señal. Por otro lado, la componente de alta frecuencia 3012 tiene un ancho de banda de $f_s/2$.

El procedimiento y el sistema descritos en el presente documento pueden implementarse como software, firmware y/o hardware. Determinados componentes pueden implementarse, por ejemplo, como software que se ejecuta en un procesador o microprocesador de señales digitales. Otro componente puede implementarse, por ejemplo, como hardware o como circuitos integrados de aplicación específica. Las señales encontradas en los procedimientos y sistemas descritos pueden almacenarse en medios tales como memorias de acceso aleatorio o medios de almacenamiento óptico. Pueden transferirse a través de redes, tales como redes de radio, redes de satélites, redes inalámbricas o redes cableadas, por ejemplo Internet. Dispositivos típicos que utilizan el

procedimiento y el sistema descritos en el presente documento son descodificadores de televisión u otros equipos en las instalaciones del cliente que descodifican señales de audio. En el lado de codificación, el procedimiento y el sistema pueden usarse en estaciones de radiodifusión, por ejemplo en sistemas de cabeceras de vídeo.

5

El presente documento describe un procedimiento y un sistema para llevar a cabo una reconstrucción de altas frecuencias de una señal en función de la componente de baja frecuencia de esa señal. Usando combinaciones de subbandas de la componente de baja frecuencia, el procedimiento y el sistema permiten la reconstrucción de frecuencias y de bandas de frecuencia que no pueden generarse mediante los procedimientos de transposición conocidos en la técnica. Además, el procedimiento y el sistema HTR descritos permiten usar bajas frecuencias de cruce y/o la generación de grandes bandas de alta frecuencia a partir de estrechas bandas de baja frecuencia.

10

REIVINDICACIONES

1. Un sistema para descodificar una señal de audio, comprendiendo el sistema:

5 un decodificador principal (101) para descodificar una componente de baja frecuencia de la señal de audio;

un banco de filtros de análisis (301) para proporcionar una pluralidad de señales de subbanda de análisis de la componente de baja frecuencia de la señal de audio;

10 una unidad de recepción de selección de subbanda para recibir información asociada con una frecuencia fundamental Ω de la señal de audio, y para seleccionar, en respuesta a la información, una primera (801) y una segunda (802) señal de subbanda de análisis de la pluralidad de señales de subbanda de análisis, de las que se genera una señal de subbanda de síntesis (803);

15 una unidad de procesamiento no lineal (302) para generar la señal de subbanda de síntesis con una frecuencia de síntesis, una magnitud y una fase:

20 determinando la magnitud de la señal de subbanda de síntesis a partir de un valor de la media geométrica de las magnitudes de la primera y la segunda señales de subbanda de análisis modificadas por un parámetro de ganancia y

25 determinando la fase de la señal de subbanda de síntesis a partir de una suma ponderada de las fases de la primera y la segunda señales de subbanda de análisis; y

un banco de filtros de síntesis (303) para generar una componente de alta frecuencia de la señal de audio a partir de la señal de subbanda de síntesis.

2. El sistema de acuerdo con la reivindicación 1, que comprende además:

30 un muestreador ascendente (104) para llevar a cabo un muestreo ascendente de la componente de baja frecuencia para proporcionar una componente de baja frecuencia muestreada de manera ascendente;

35 un ajustador de envolvente (103) para conformar la componente de alta frecuencia; y

una unidad de suma de componentes para determinar una señal de audio descodificada como la suma de la componente de baja frecuencia muestreada de manera ascendente y la componente ajustada de alta frecuencia.

40 3. El sistema de acuerdo con la reivindicación 2, que comprende además una unidad de recepción envolvente para recibir información relacionada con la envolvente de la componente de alta frecuencia de la señal de audio.

- 45 4. El sistema de acuerdo con la reivindicación 3, que comprende además:

una unidad de entrada para recibir la señal de audio, que comprende la componente de baja frecuencia; y

50 una unidad de salida para proporcionar la señal de audio descodificada, que comprende la componente de baja frecuencia y la componente de alta frecuencia generada.

55 5. El sistema de acuerdo con la reivindicación 1, en el que el banco de filtros de análisis (301) presenta una separación entre frecuencias que está asociada con la frecuencia fundamental Ω de la señal de audio.

6. Un procedimiento para descodificar una señal de audio, comprendiendo el procedimiento:

descodificar una componente de baja frecuencia de la señal de audio;

60 proporcionar una pluralidad de señales de subbanda de análisis de la componente de baja frecuencia de la señal de audio;

65 recibir información asociada con una frecuencia fundamental Ω de la señal de audio que permite la selección de una primera (801) y una segunda (802) señal de subbanda de análisis a partir de la pluralidad de señales de subbanda de análisis;

generar una señal de subbanda de síntesis con una frecuencia de síntesis, una magnitud y una fase:

- 5 determinando la magnitud de la señal de subbanda de síntesis a partir de un valor de la media geométrica de las magnitudes de la primera y la segunda señales de subbanda de análisis modificadas por un parámetro de ganancia y
- 10 determinando la fase de la señal de subbanda de síntesis a partir de una suma ponderada de las fases de la primera y la segunda señales de subbanda de análisis; y
- 10 generar (303) una componente de alta frecuencia de la señal de audio a partir de la señal de subbanda de síntesis.
- 15 7. Un medio de almacenamiento que comprende un programa de software adaptado para ejecutarse en un procesador y para realizar las etapas de procedimiento de la reivindicación 6 cuando se llevan a cabo en un dispositivo informático.

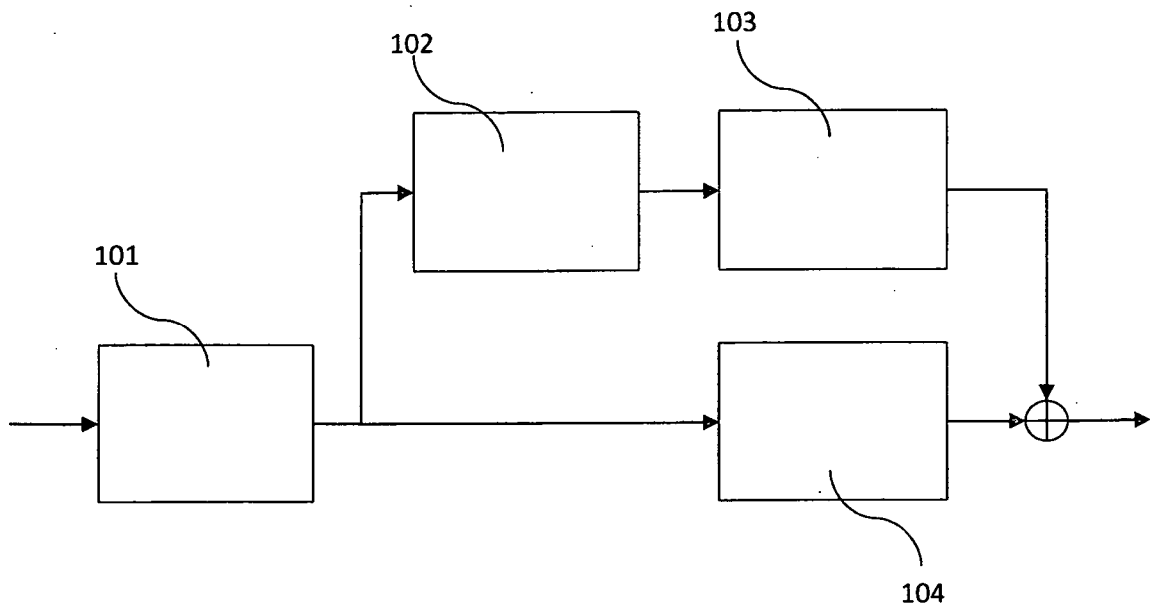


Fig. 1

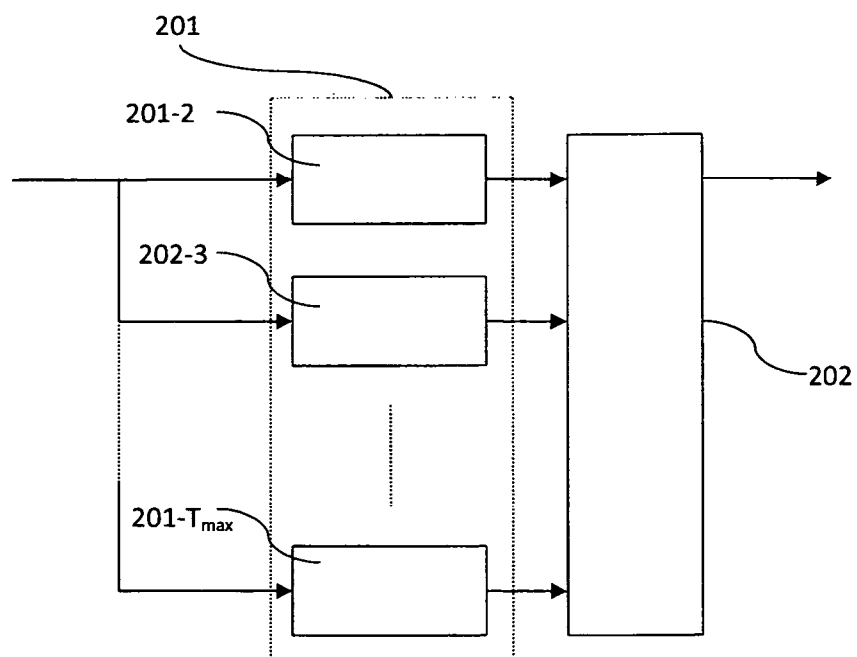


Fig. 2

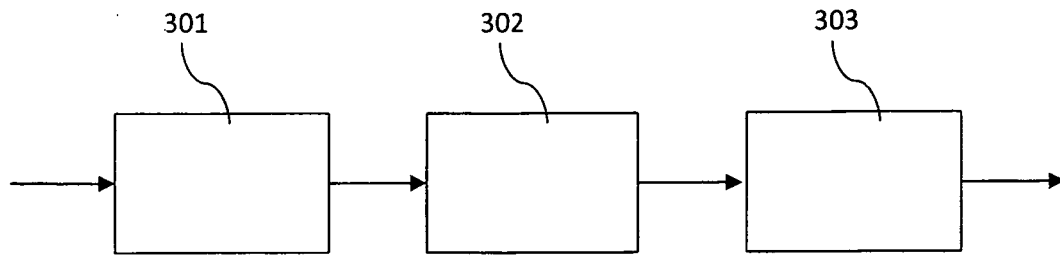


Fig. 3

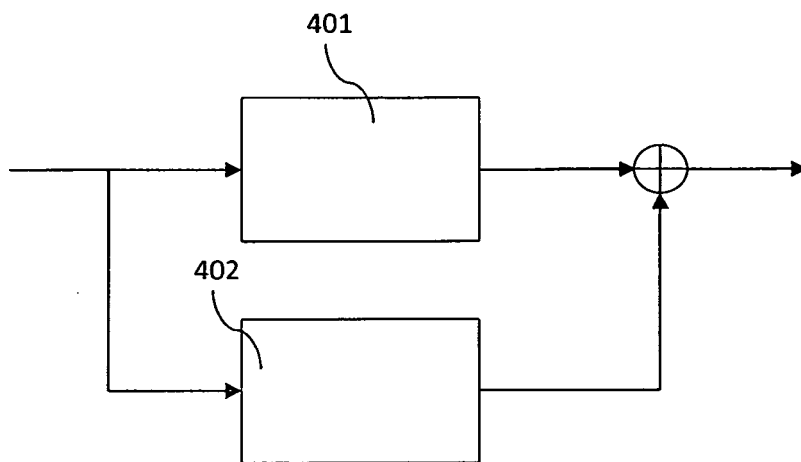


Fig. 4

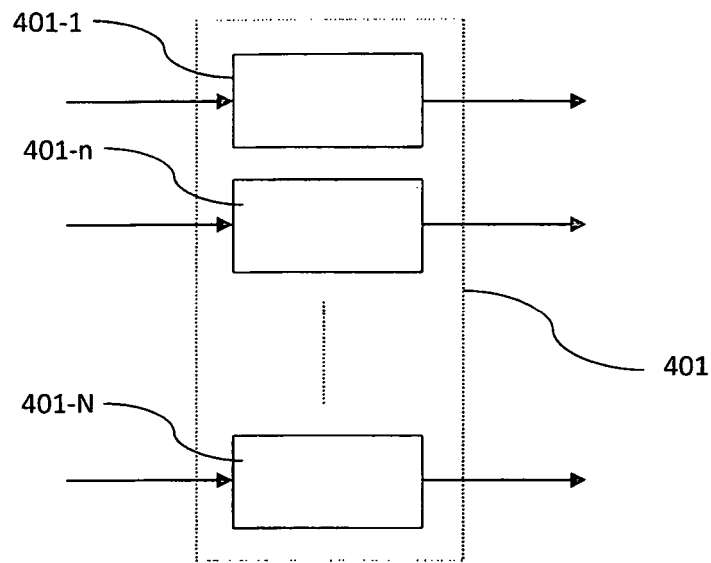


Fig. 5

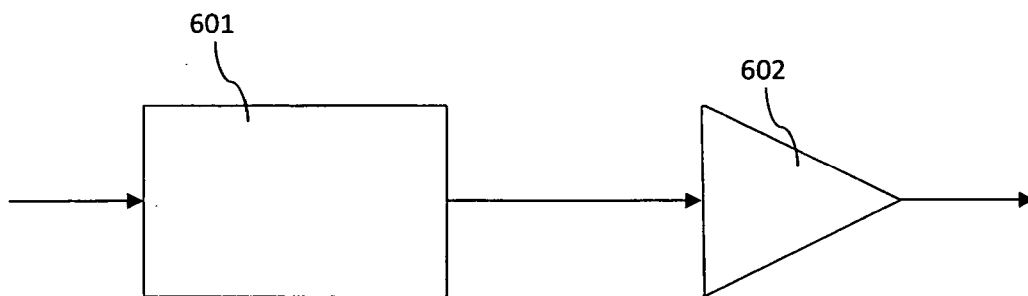


Fig. 6

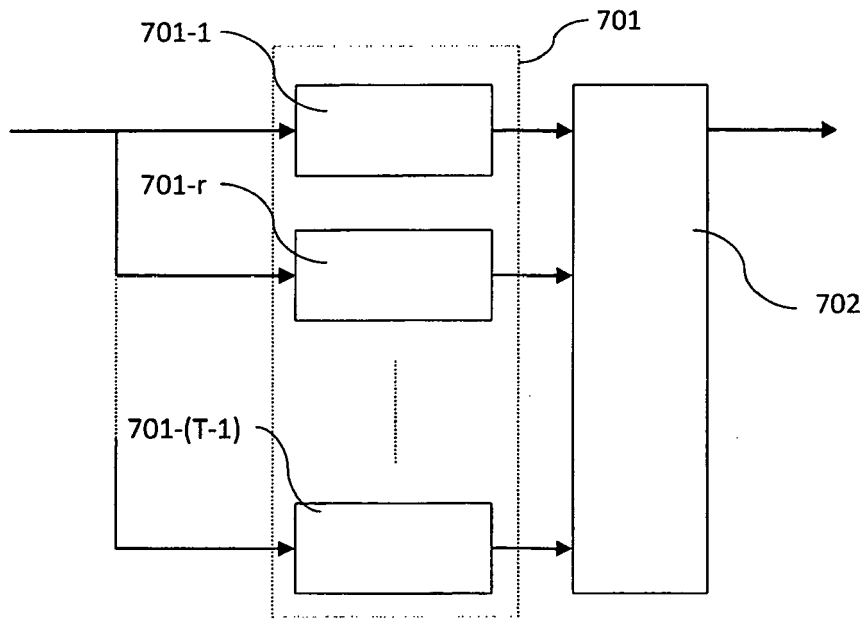


Fig. 7

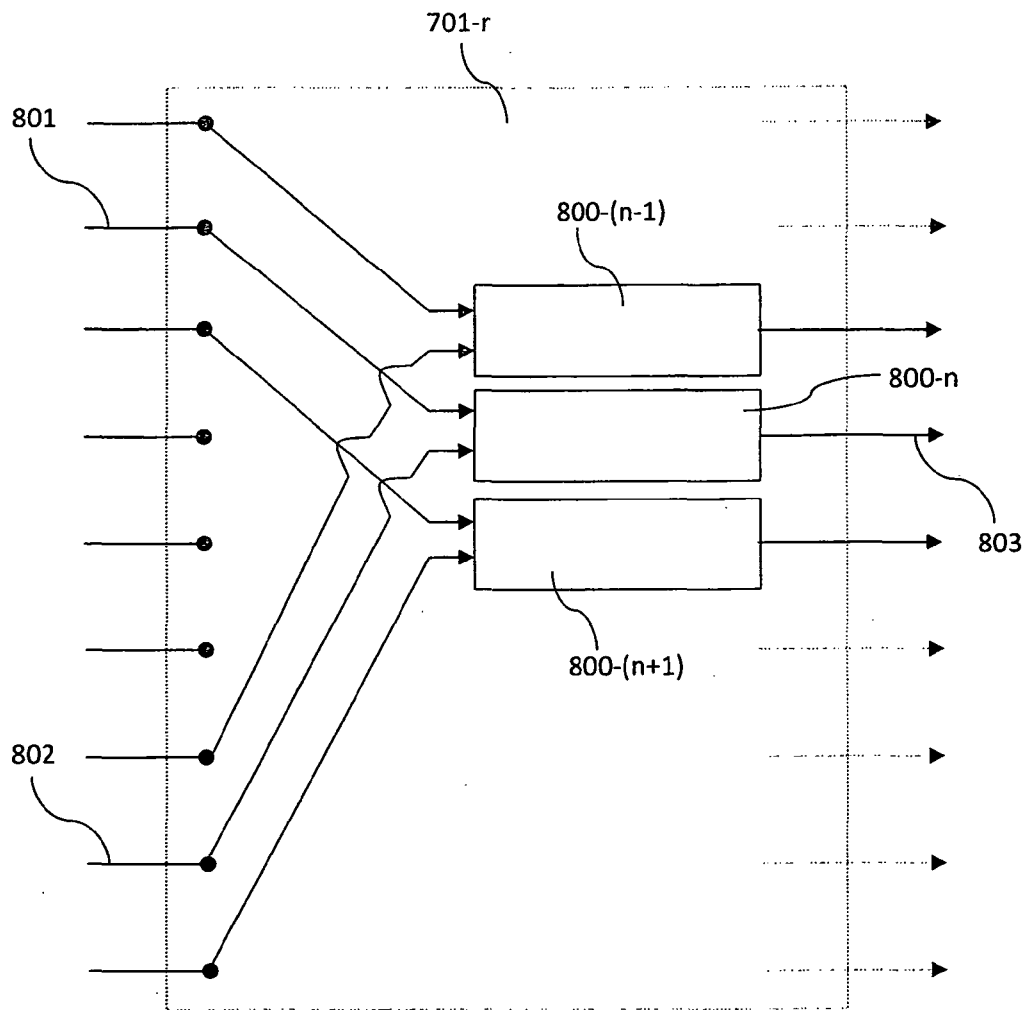


Fig. 8

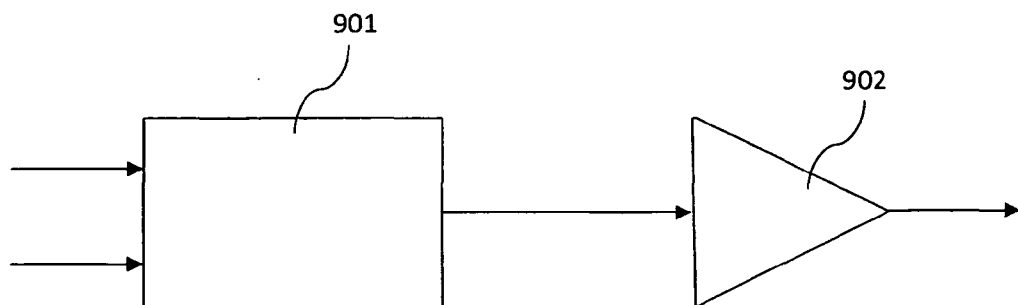


Fig. 9

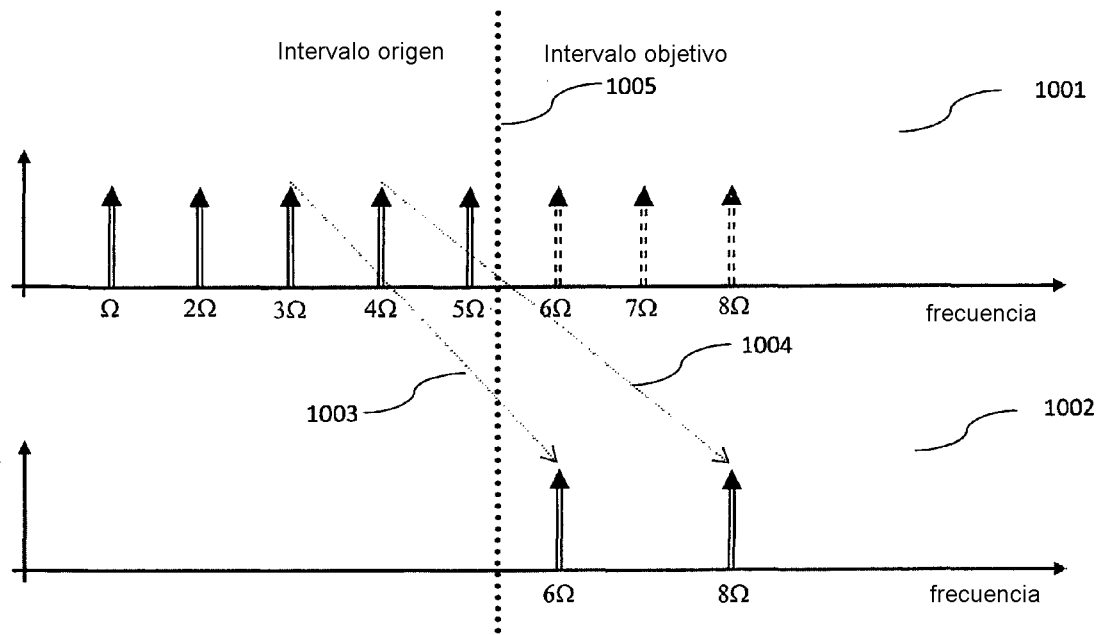


Fig. 10

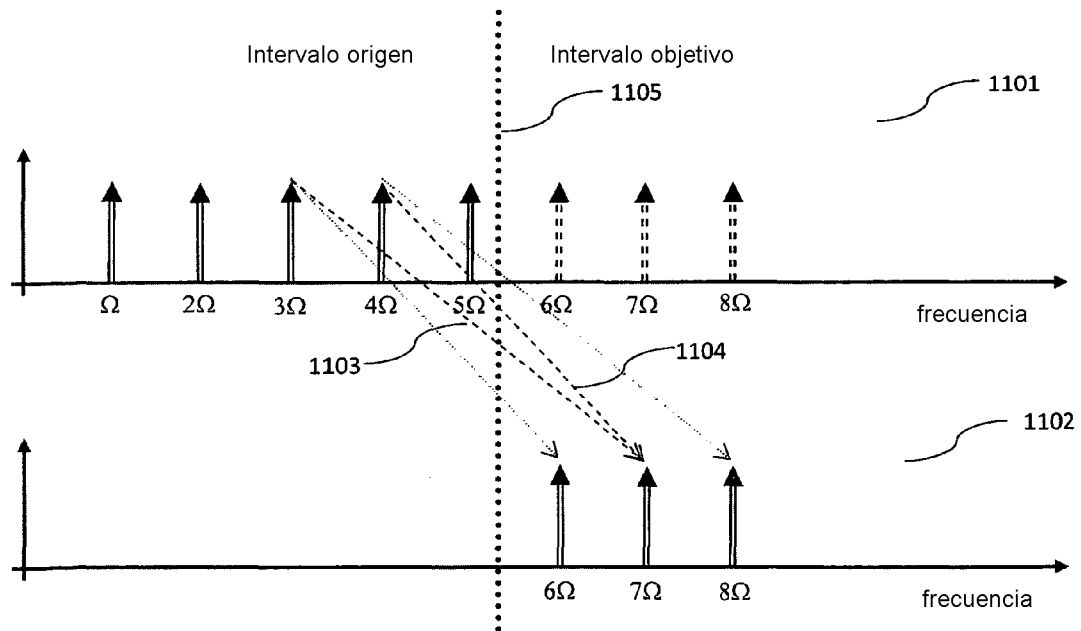


Fig. 11

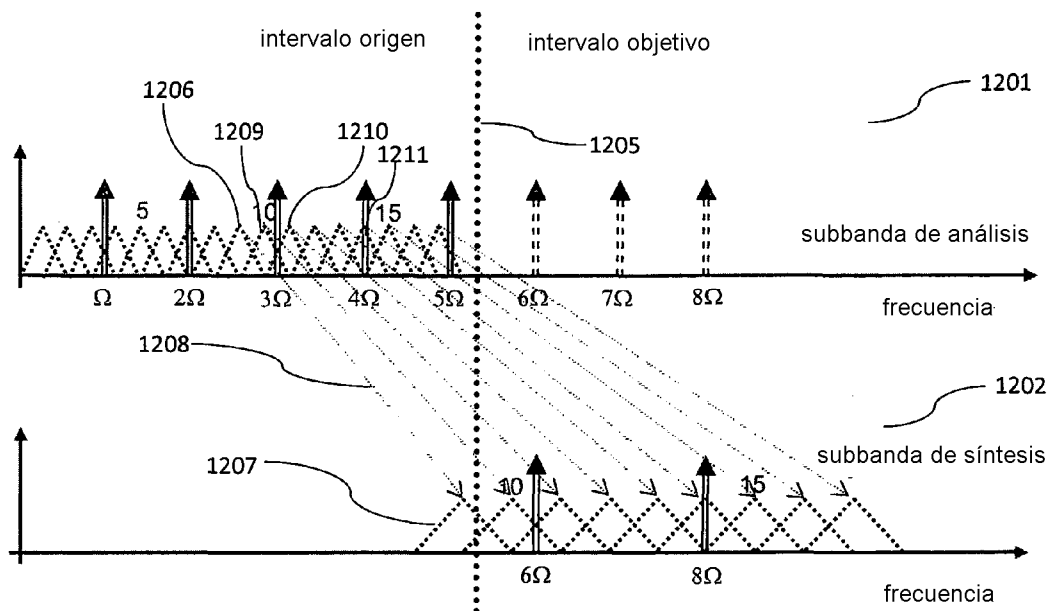


Fig. 12

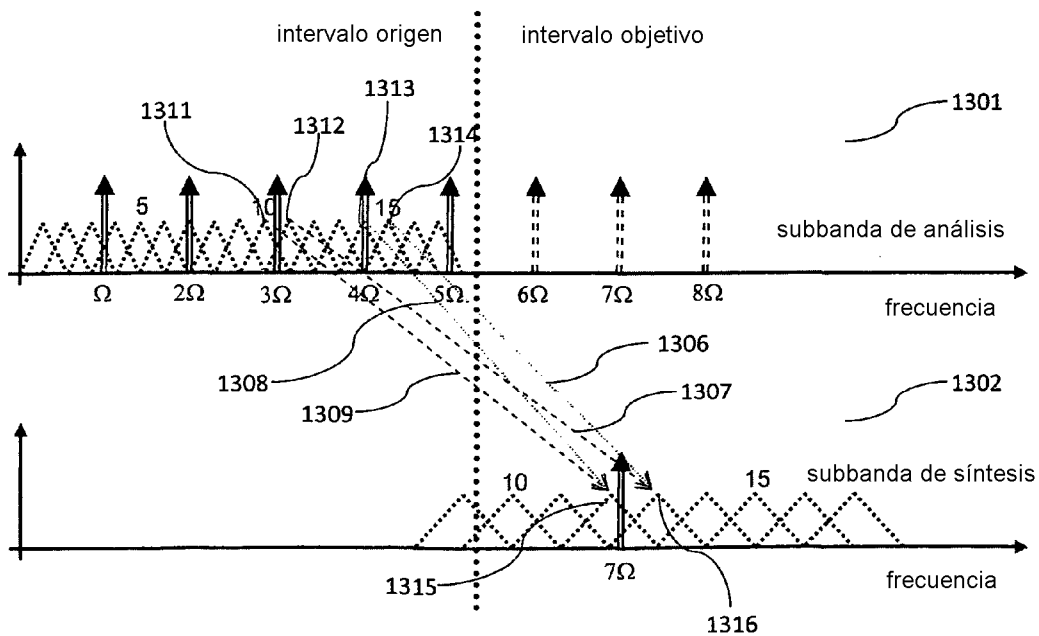


Fig. 13

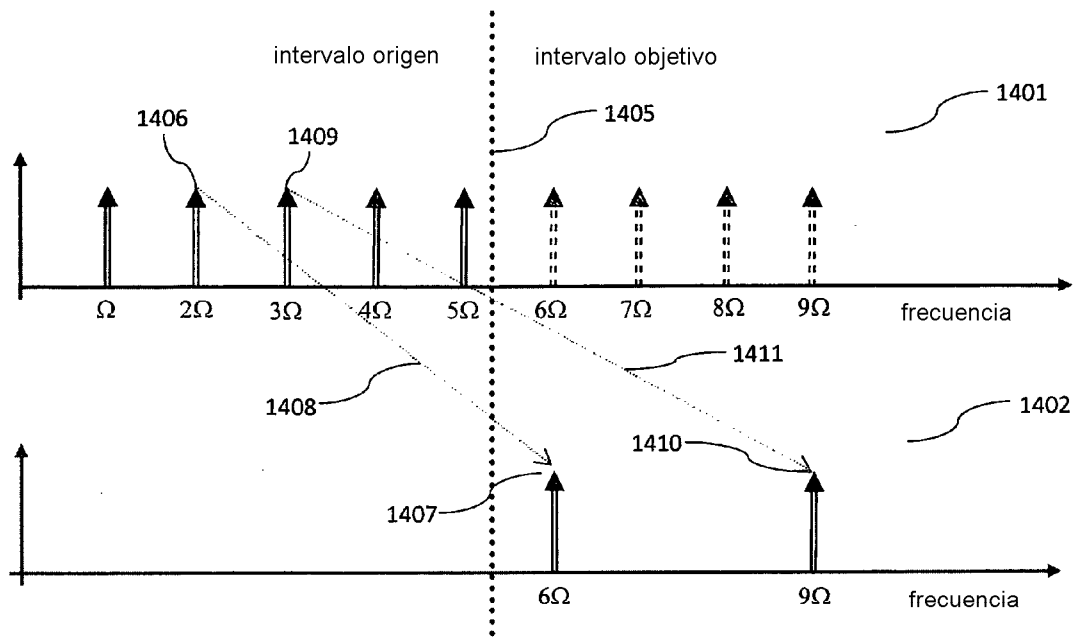


Fig. 14

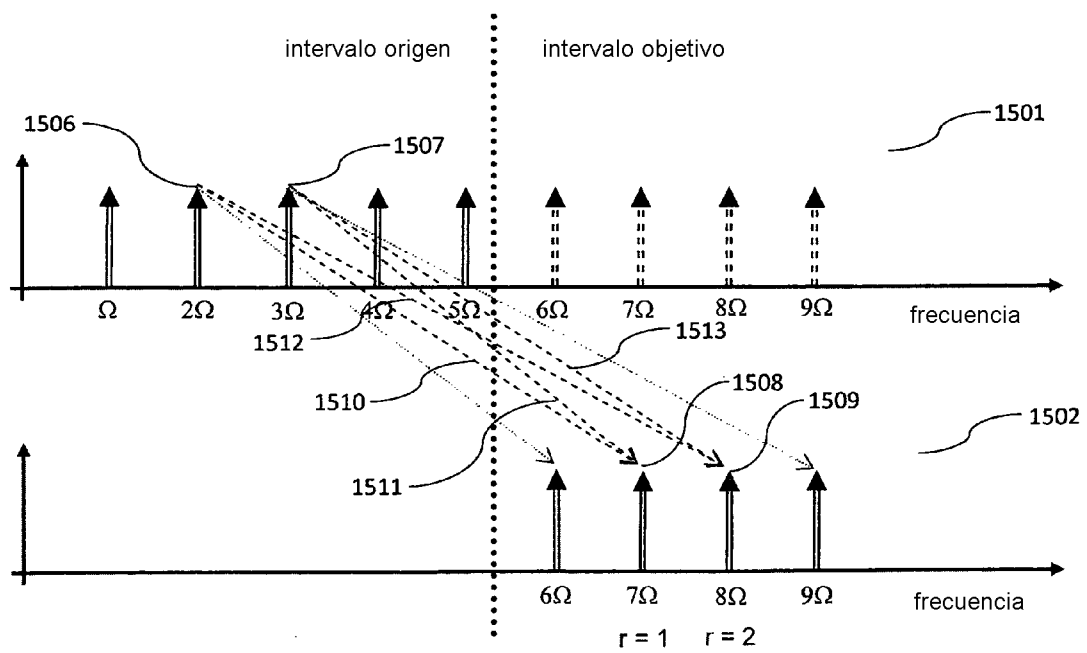


Fig. 15

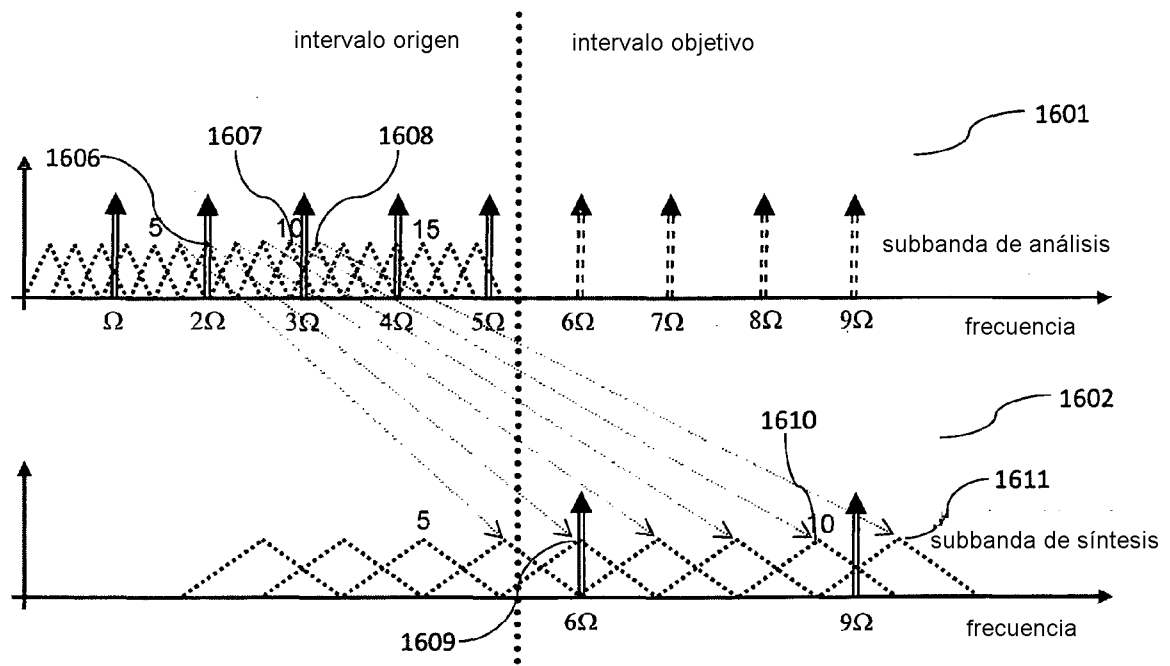


Fig. 16

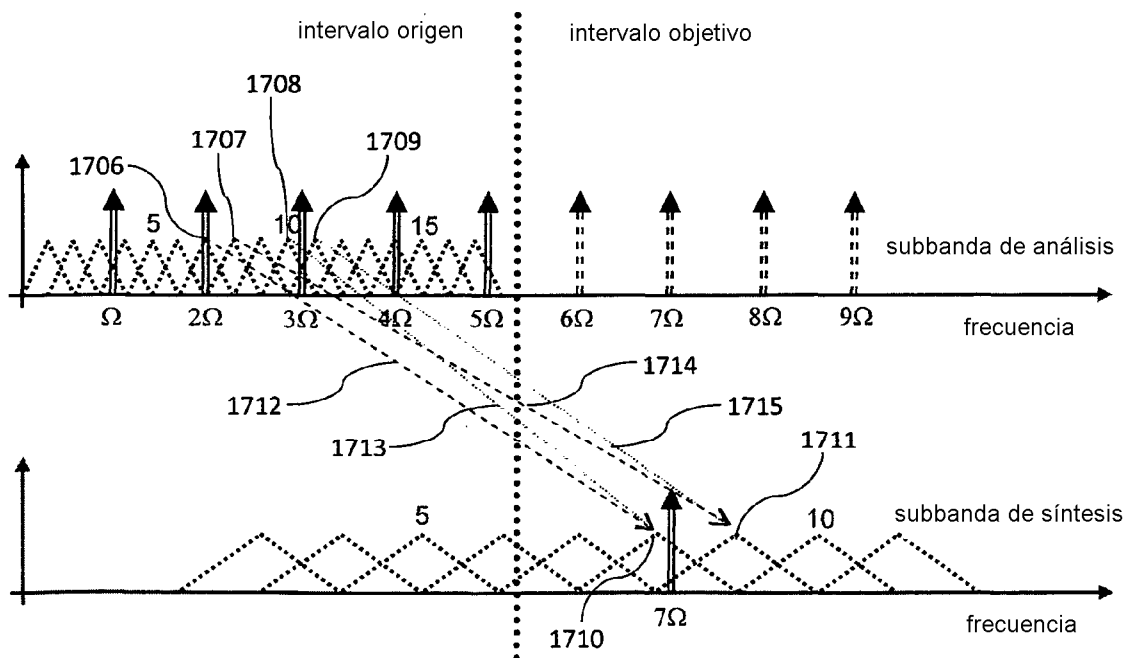


Fig. 17

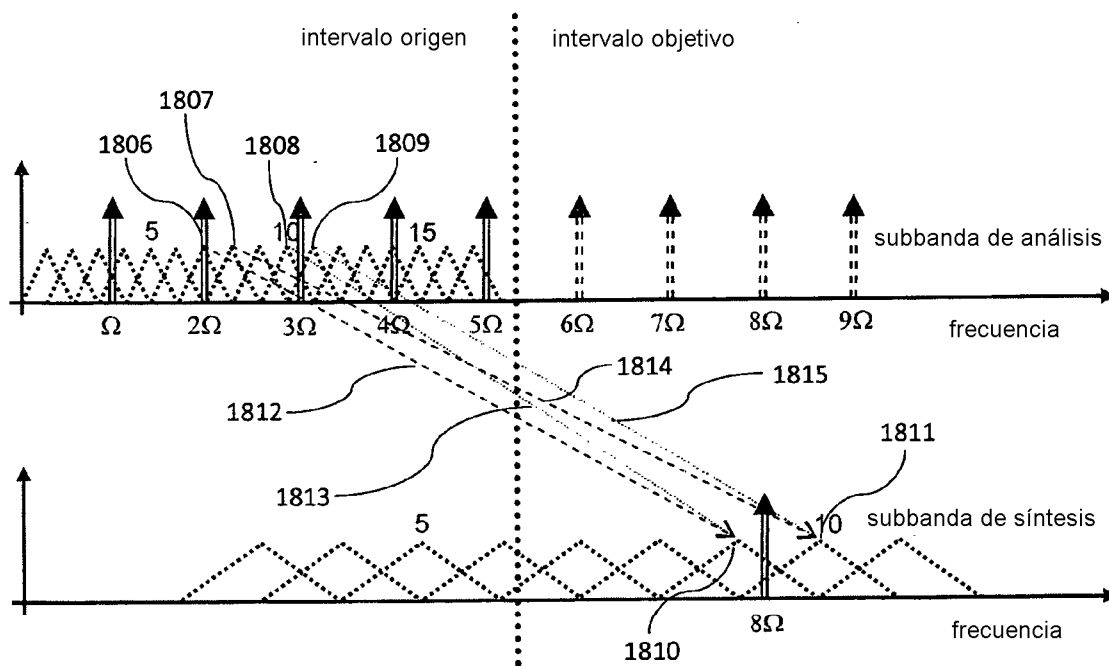


Fig. 18

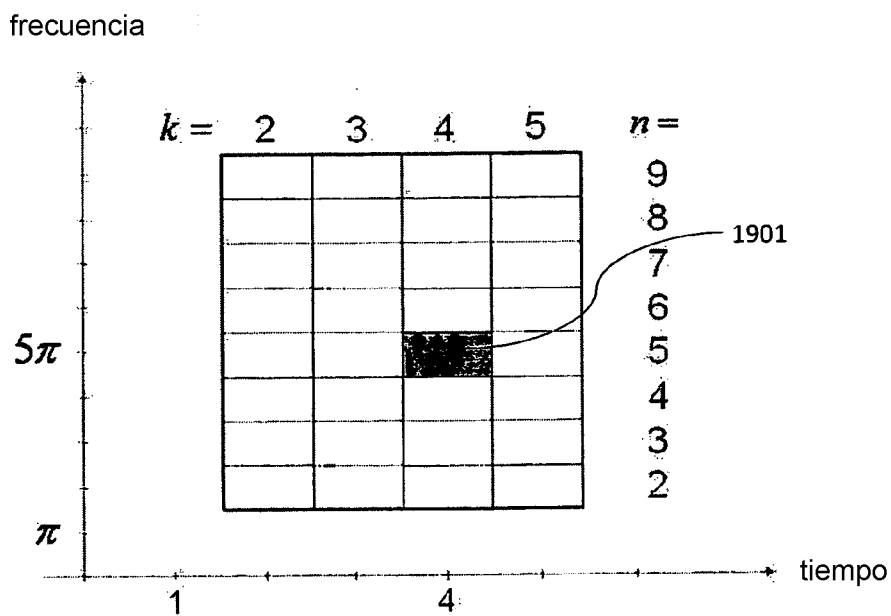


Fig. 19

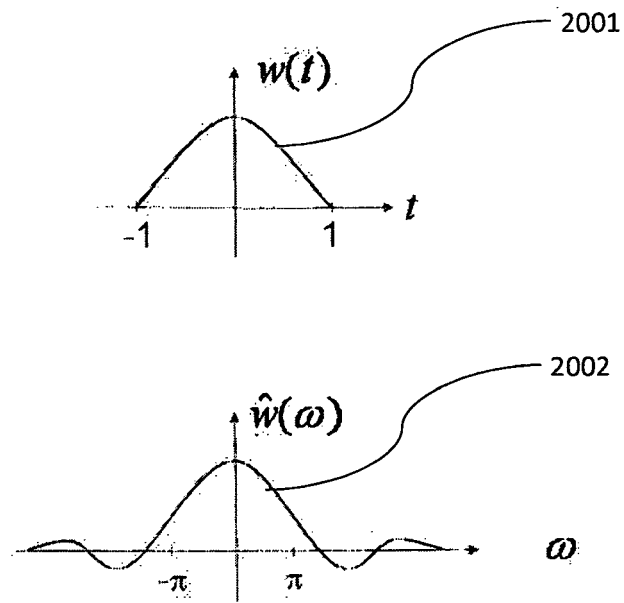


Fig. 20

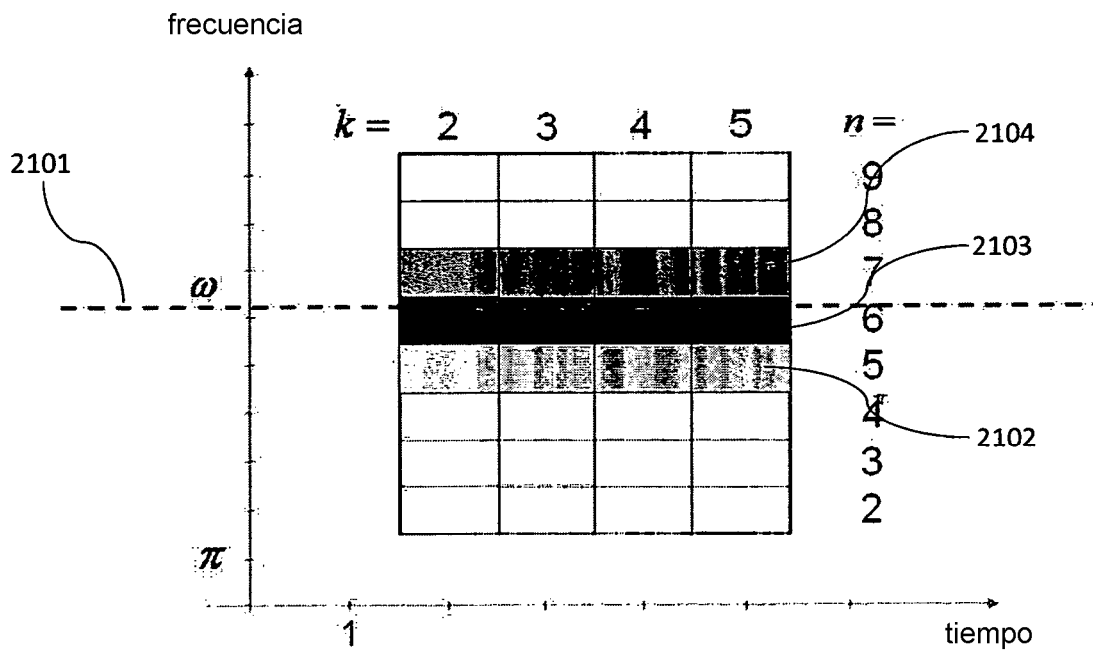


Fig. 21

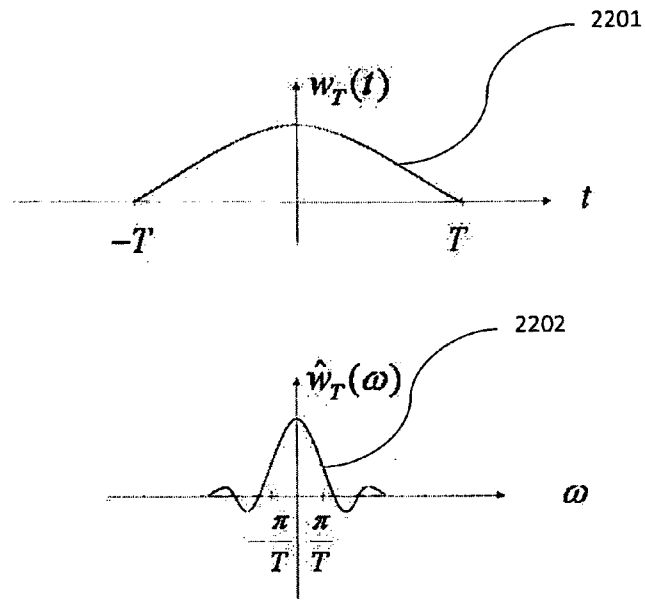


Fig. 22

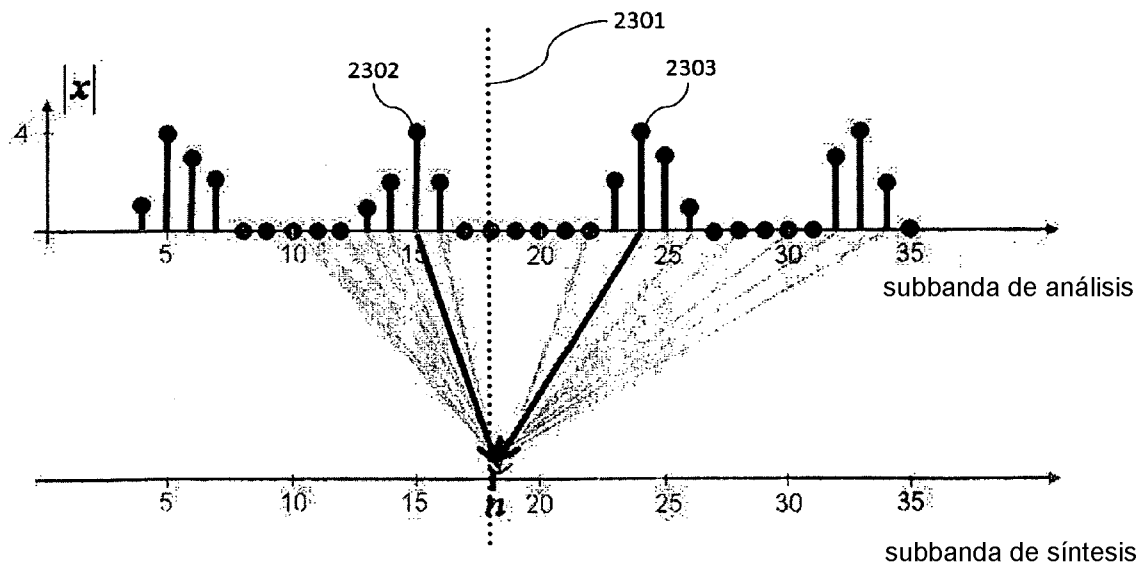


Fig. 23

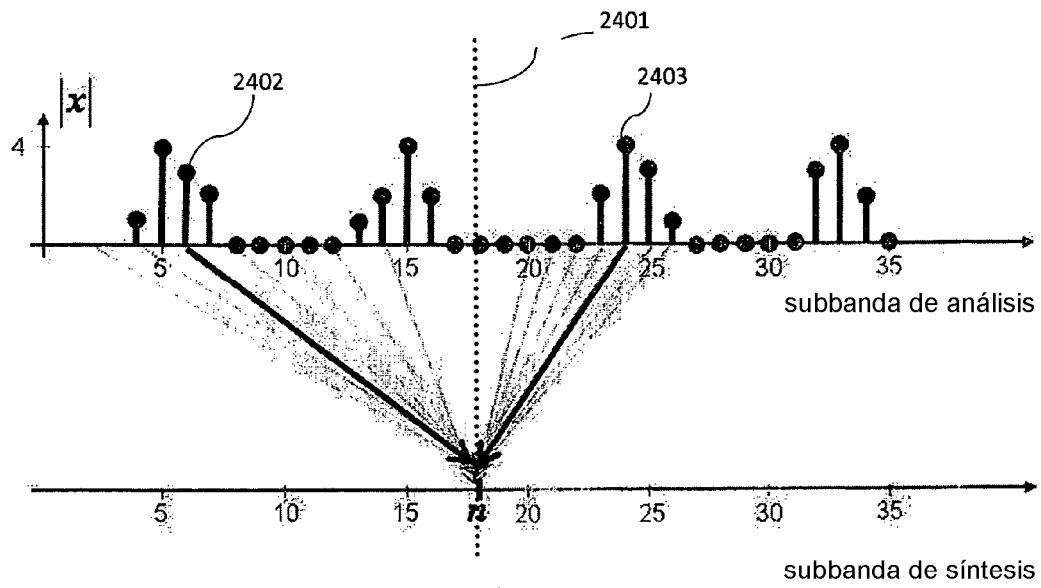


Fig. 24

Espectro de señal

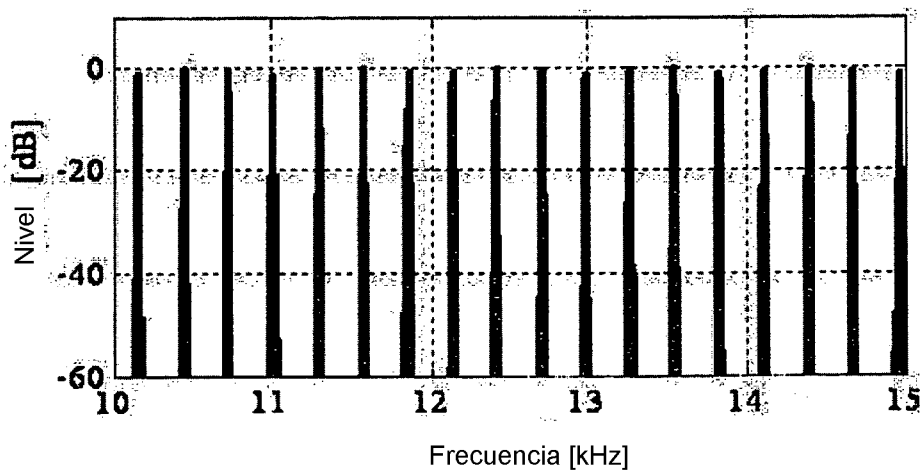


Fig. 25

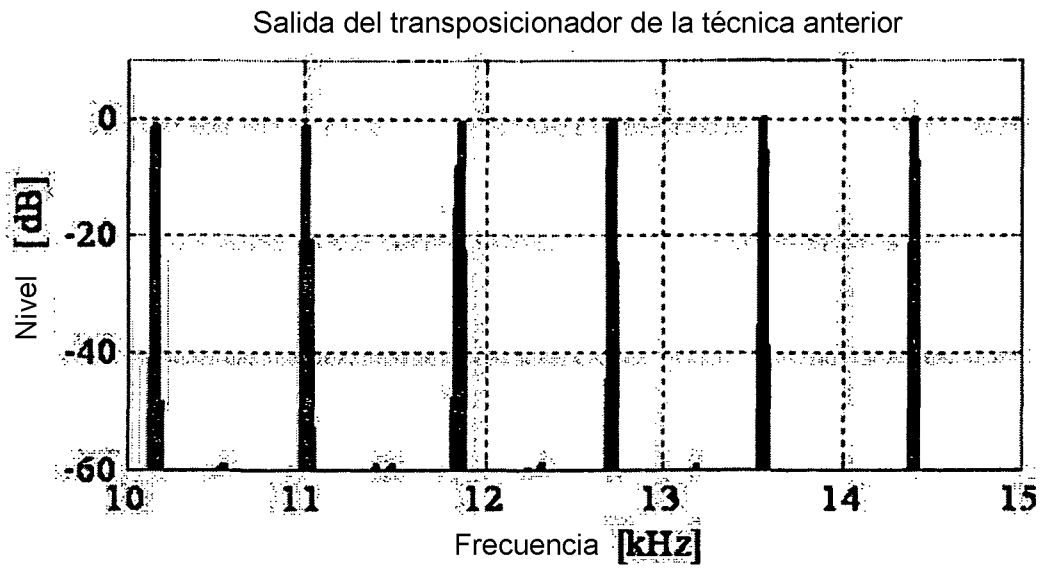


Fig. 26

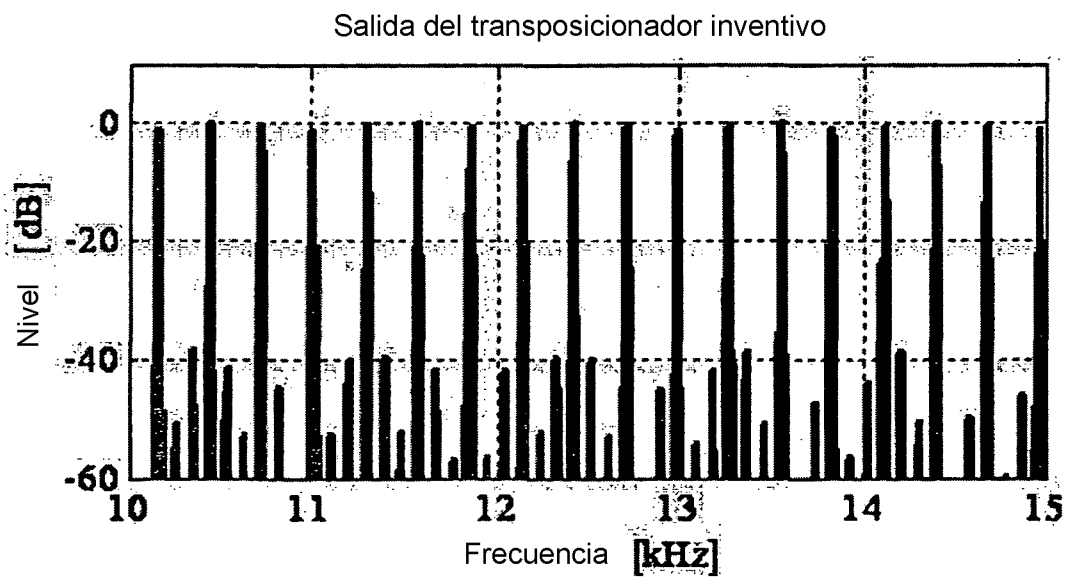


Fig. 27

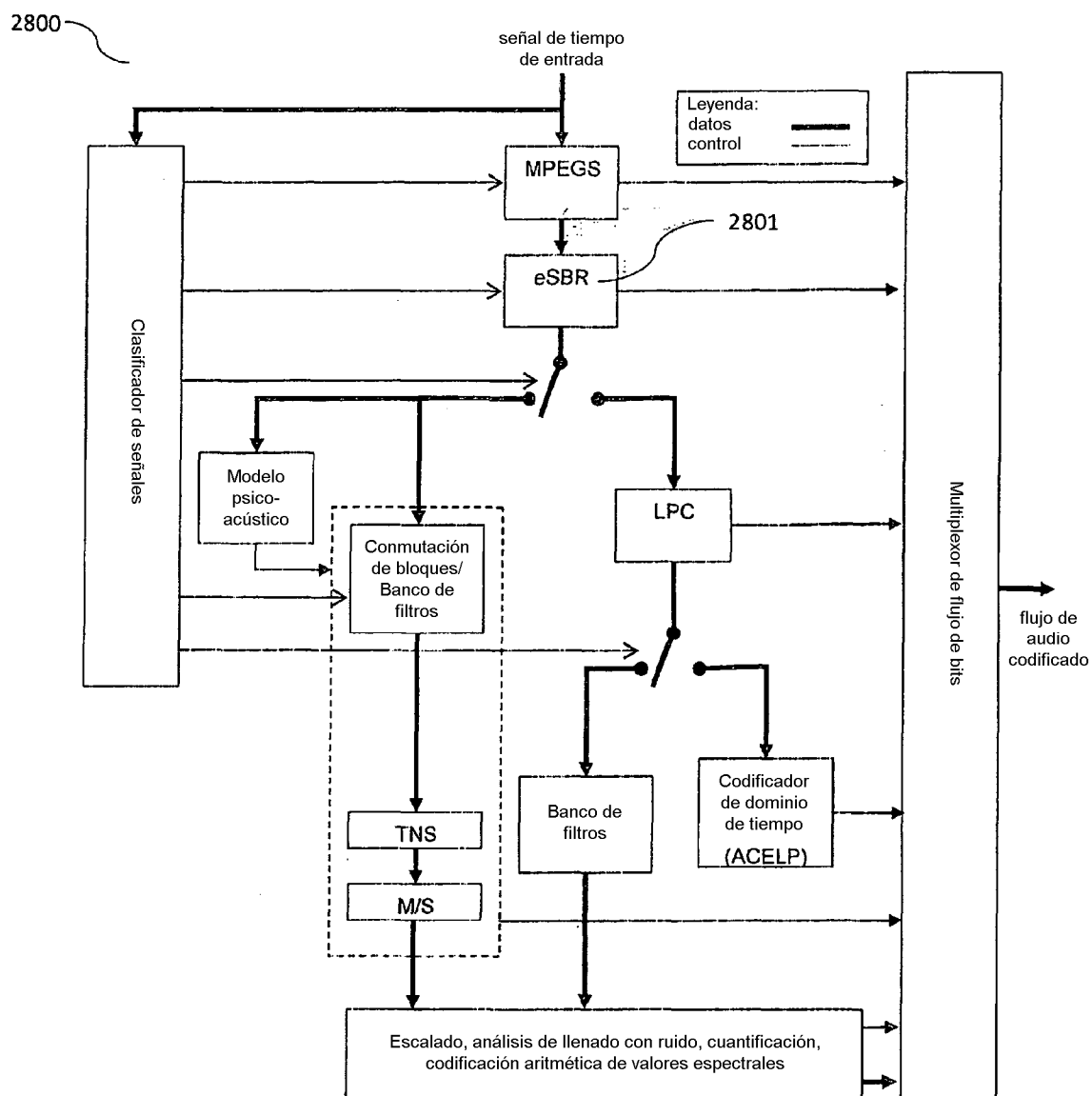


Fig. 28

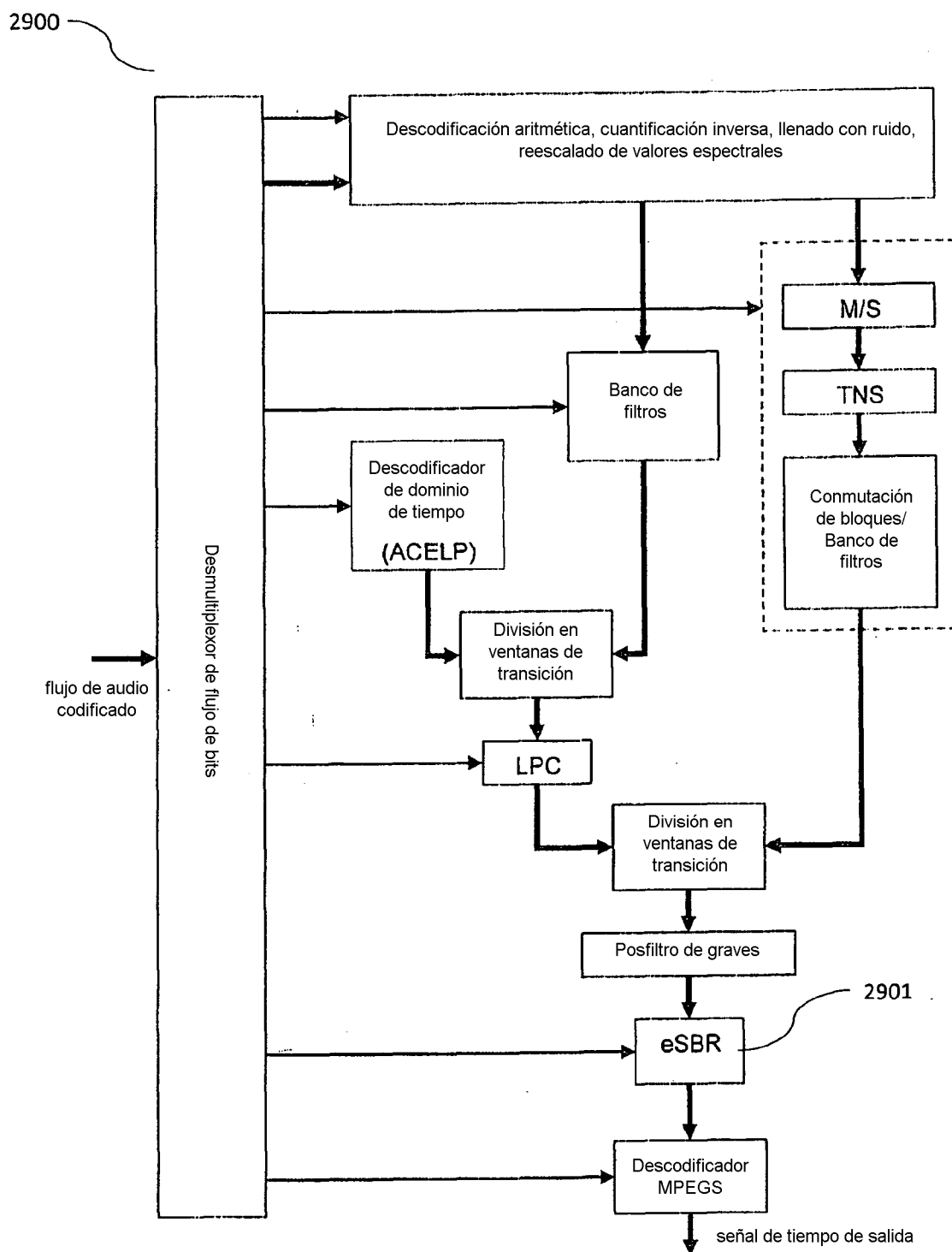


Fig. 29

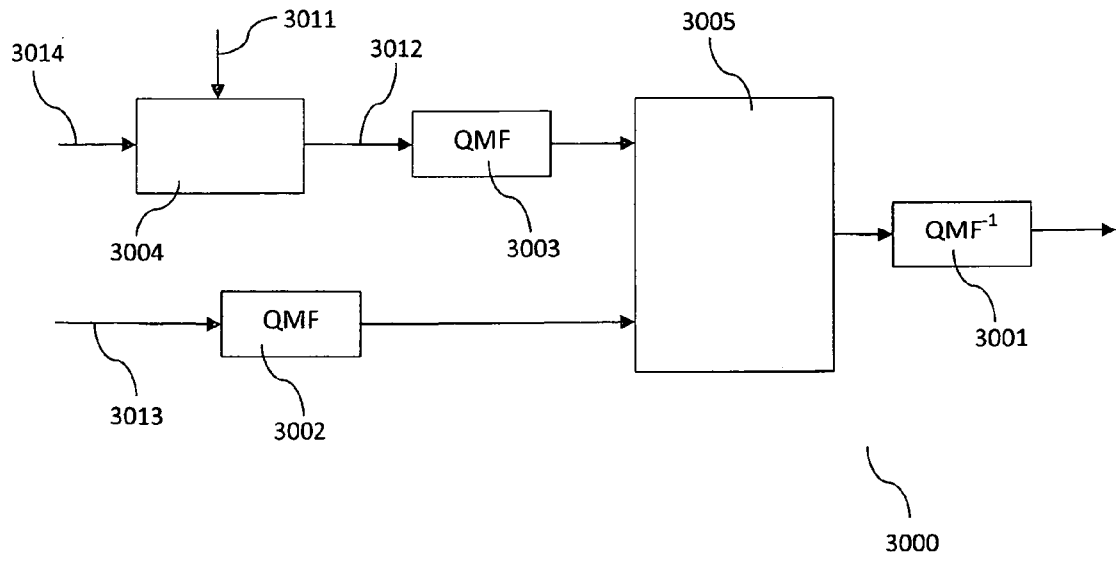


Fig. 30