

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 023 486**

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 23210729 (2)

97 Fecha y número de publicación de la concesión europea: **26.02.2025** EP 4300495

54 Título: **Transposición armónica mejorada de productos de cruce**

30 Prioridad:

16.01.2009 US 14522309 P

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

02.06.2025

73 Titular/es:

**DOLBY INTERNATIONAL AB (100.00%)
77 Sir John Rogerson's Quay, Block C, Grand
Canal Docklands
Dublin, D02 VK60, IE**

72 Inventor/es:

**VILLEMOS, LARS y
HEDELIN, PER**

74 Agente/Representante:

ELZABURU, S.L.P

ES 3 023 486 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 productos de cruce

Campo técnico

5 La presente invención se relaciona con los sistemas de codificación de audio que hacen uso de un método de transposición armónica para la reconstrucción de altas frecuencias (HFR).

Antecedentes de la invención

10 Las tecnologías HFR, tales como la tecnología de Replicación de Banda Espectral (SBR), permiten mejorar de manera significativa la eficiencia de codificación de los códec de audio tradicionales. En combinación con la Codificación de Audio Avanzada (AAC) MPEG-4 forma un muy eficiente códec de audio, que ya está en uso dentro del Sistema de Radio por Satélite XM y la Radio Digital Mundial. La combinación de AAC y SBR es denominada aacPlus. Es parte del estándar MPEG-4 donde es referido como el Perfil AAC de Alta Eficiencia. En general, la tecnología HFR se puede combinar con un códec de audio perceptivo de una manera compatible hacia atrás y hacia delante, ofreciendo así la posibilidad de mejorar los sistemas de difusión ya establecidos como el MPEG capa 2 usado en el sistema DAB Eureka. Los métodos de transposición HFR se pueden combinar también con códec de voz para permitir voz de banda ancha a tasas de bits ultra bajas.

15 La idea básica detrás de la HFR es la observación de que normalmente hay presente una fuerte correlación entre las características del rango de alta frecuencia de una señal y las características del rango de baja frecuencia de la misma señal. Por tanto, una buena aproximación para la representación del rango de entrada original de alta frecuencia de una señal puede ser alcanzado mediante una transposición de señal desde el rango de baja frecuencia al rango de alta frecuencia.

20 Este concepto de transposición fue establecido en el documento WO 98/57436, como un método para recrear una banda de alta frecuencia desde una banda de frecuencia inferior de una señal de audio. Se puede obtener un ahorro sustancial en la tasa de bits usando este concepto en la codificación de audio y/o la codificación de voz. A continuación, se hará referencia a la codificación de audio, pero se debería observar que los métodos y sistemas descritos son igualmente aplicables a la codificación de voz y la codificación de audio y voz unificada (USAC).

25 En un sistema de codificación de audio basado en HFR, se presenta una señal de bajo ancho de banda a un codificador de forma de onda de núcleo y se regeneran las frecuencias superiores en el lado del decodificador usando una transposición de la señal de bajo ancho de banda e información complementaria adicional, que se codifica normalmente a tasas de bits muy bajas y que describe la forma espectral objetivo. Para tasas de bits bajas, donde el ancho de banda de la señal codificada del núcleo es estrecho, resulta cada vez más importante recrear una banda superior, esto es el rango superior de frecuencia de la señal de audio, con características perceptualmente agradables. Se mencionan dos variantes de métodos de reconstrucción de frecuencia armónica a continuación, una es referida como transposición armónica y la otra es referida como modulación de banda lateral única.

30 El principio de la transposición armónica definido en el documento WO 98/57436 es que una senoide con una frecuencia ω se hace corresponder con una senoide con frecuencia $T\omega$ donde $T > 1$ es un número entero que define el orden de transposición. Una característica atractiva de la transposición armónica es que extiende un rango de frecuencia de origen en un rango de frecuencia objetivo por un factor igual al orden de transposición, esto es por un factor igual a T . La transposición armónica se desempeña bien para material musical complejo. Además, la transposición armónica muestra unas bajas frecuencias de corte, esto es se puede generar un gran rango de alta frecuencia por encima de la frecuencia de corte a partir de un relativamente pequeño rango de baja frecuencia por debajo de la frecuencia de corte.

35 En contraste a la transposición armónica, una modulación de banda lateral única (SSB) basada en HFR hace corresponder una senoide con la frecuencia ω a una senoide con frecuencia $\omega + \Delta\omega$ donde $\Delta\omega$ es un desplazamiento de frecuencia fijo. Se ha observado que, dada una señal de núcleo con un bajo ancho de banda, puede producirse un artefacto de sonido disonante a partir de la transposición SSB. Se debería observar también que para una frecuencia de corte baja, esto es un rango de frecuencia de origen pequeño, la transposición armónica requerirá un menor número de correcciones para completar un rango de frecuencia objetivo deseado que la transposición basada en SSB. A modo de ejemplo, si se debe completar el rango de alta frecuencia de $(\omega, 4\omega]$, entonces usar un orden de transposición $T = 4$ de la transposición armónica puede completar este rango de frecuencia a partir de un rango de baja frecuencia de $(\frac{1}{4}\omega, \omega]$. Por otro lado, una transposición basada en SSB que usa el mismo rango de baja frecuencia debe usar un desplazamiento de frecuencia $\Delta\omega = \frac{3}{4}\omega$ y es necesario repetir el proceso cuatro veces para completar el rango de alta frecuencia $(\omega, 4\omega]$.

40 Por otro lado, como ya se apuntó en el documento WO 02/052545 A1, la transposición armónica tiene inconvenientes para las señales con una estructura periódica prominente. Dichas señales son superposiciones de las sinusoides relacionadas de manera armónica con frecuencias $\Omega, 2\Omega, 3\Omega, \dots$, donde Ω es la frecuencia fundamental.

55 Tras la transposición armónica de orden T , las sinusoides de salida tienen frecuencias $T\Omega, 2T\Omega, 3T\Omega, \dots$, que, en

5 caso de $T > 1$, es sólo un subconjunto estricto de las series armónicas completas deseadas. En términos de calidad de audio resultante normalmente se percibirá un tono "fantasma" correspondiente a la frecuencia fundamental transpuesta $T\Omega$. A menudo la transposición armónica resulta en un carácter de sonido "metálico" de la señal de audio codificada y decodificada. La situación se puede aliviar en un cierto grado añadiendo varios órdenes de transposición $T = 2, 3, \dots, T_{\max}$ a la HFR, pero éste método es computacionalmente complejo si se han de evitar la mayoría de los saltos espectrales.

10 Una solución alternativa para evitar la aparición de los tonos "fantasma" al usar la transposición armónica se ha presentado en el documento WO 02/052545 A1. La solución consiste en usar dos tipos de transposición, esto es una transposición armónica típica y una "transposición de pulso" especial. El método descrito enseña a cambiar a la "transposición de pulso" dedicada para las partes de la señal de audio que se detectan que son periódicas con una naturaleza de tren de pulsos. El problema con este enfoque es que la solicitud de "transposición de pulso" en el material de música complejo a menudo degrada la calidad comparado con la transposición armónica basada en un banco de filtros de gran resolución. Por lo tanto, los mecanismos de detección se han de sintonizar de forma bastante conservadora de manera tal que la transposición de pulso no se use para el material complejo. Inevitablemente, los instrumentos y voces de tono único se clasificarán a veces como señales complejas, invocando la transposición armónica y por lo tanto faltando armónicos. Además, si se produce el cambio en el medio de una señal de un solo tono, o de una señal con un tono dominante en un fondo complejo más débil, el cambio en sí entre los dos métodos de transposición que tienen propiedades de relleno de espectro muy diferentes generará artefactos audíbles. Otra variante para realizar la reconstrucción de frecuencia de armónico se propone en el documento US 2004/0028244 A1.

Compendio de la invención

25 Las realizaciones de la presente invención se definen por las reivindicaciones independientes. Características adicionales de las realizaciones de la invención se presentan en las reivindicaciones dependientes. En lo siguiente, las partes de la descripción y de los dibujos que se refieren a realizaciones anteriores que no necesariamente comprenden todas las características para implementar las realizaciones de la invención reivindicada no se representan como realizaciones de la invención sino como ejemplos útiles para entender las realizaciones de la invención.

Breve descripción de los dibujos

30 La presente invención será descrita ahora a modo de ejemplos ilustrativos, no limitando el alcance de la invención. Será descrita con referencia a los dibujos adjuntos, en los cuales:

La Fig. 1 ilustra la operación de un decodificador de audio mejorado HFR;

La Fig. 2 ilustra la operación de un transponedor armónico que usa varios órdenes;

La Fig. 3 ilustra la operación de un transponedor armónico en el dominio de la frecuencia (FD);

La Fig. 4 ilustra la operación del uso inventivo de procesamiento de términos de cruce;

35 La Fig. 5 ilustra el procesamiento directo de la técnica anterior;

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

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

La Fig. 8 ilustra la operación de un bloque de procesamiento de términos de cruce;

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

40 Las Fig. 10 – 18 ilustra el efecto de la invención para la transposición armónica de las señales periódicas ejemplares;

La Fig. 19 ilustra la resolución tiempo-frecuencia de una Transformada de Fourier de Tiempo Reducido (STFT).

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

45 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 usada en el lado de análisis;

Las Fig. 23 y 24 ilustran la determinación de las subbandas apropiadas del banco de filtros de análisis para la mejora de los términos de cruce de la subbanda de la banda del filtro de síntesis;

Las Fig. 25, 26, y 27 ilustran resultados experimentales del método descrito de transposición armónica de los

términos directos y de cruce;

Las Fig. 28 y 29 ilustran las realizaciones de un codificador y un decodificador, respectivamente, que usan los esquemas de transposición armónica mejorados explicados en términos generales en el presente documento; y

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

5 Descripción de las realizaciones preferidas

Las realizaciones anteriormente descritas son simplemente ilustrativas para los principios de la presente invención para la así denominada TRANSPOSICIÓN ARMÓNICA MEJORADA DE PRODUCTOS DE CRUCE. Se entiende que serán evidentes las modificaciones y variaciones de las disposiciones y los detalles descritos en la presente memoria para aquellos expertos en la técnica. Es la intención, por lo tanto, estar limitado sólo por el alcance de las próximas reivindicaciones de patente y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en la presente memoria.

La Fig. 1 ilustra la operación de un decodificador de audio mejorado HFR. El decodificador 101 de audio de núcleo emite una señal de audio de bajo ancho de banda que se alimenta a un muestreador ascendente 104 que se puede requerir para producir una contribución de salida de audio final a la tasa de muestreo completa deseada. Dicho muestreo ascendente se requiere para los sistemas de doble tasa, donde el códec de audio de núcleo limitado en banda está operando a la mitad de la tasa de muestreo de audio externa, mientras que la parte HFR se procesa a la frecuencia de muestreo completa. Por consiguiente, para un sistema de tasa única, se omite este muestreador ascendente 104. La emisión de bajo ancho de banda 101 se envía también al transponedor o unidad 102 de transposición que emite la señal transpuesta, esto es la señal que comprende el rango de alta frecuencia deseado. A esta señal transpuesta se le puede dar forma en el tiempo y la frecuencia mediante el ajustador 103 de envolvente. La emisión de audio final es la suma de la señal de núcleo de bajo ancho de banda y la señal transpuesta ajustada de la envolvente.

La Fig. 2 ilustra la operación de un transponedor 201 armónico, que corresponde al transponedor 102 de la Fig. 1, que comprende diversos transponedores de diferentes órdenes T de transposición. La señal a ser transpuesta se pasa al banco de transponedores 201-2, 201-3, ..., 201- T_{max} individuales que tienen órdenes de transposición $T = 2, 3, \dots, T_{max}$, respectivamente. Normalmente un orden $T_{max} = 3$ de transposición es suficiente para la mayoría de las aplicaciones de codificación de audio. Las contribuciones de los diferentes transponedores 201-2, 201-3, ..., 201- T_{max} se suman en 202 para producir la salida combinada del transponedor. En una primera realización, la operación de suma puede comprender la adición de las contribuciones individuales. En otra realización, las contribuciones se ponderan con diferentes pesos, de manera que el efecto de la adición de las múltiples contribuciones a ciertas frecuencias se mitiga. Por ejemplo, las contribuciones de tercer orden se pueden añadir con una ganancia inferior que las contribuciones de segundo orden. Finalmente, la unidad 202 de suma puede añadir las contribuciones de manera selectiva dependiendo de la frecuencia de salida. Por ejemplo, la transposición de segundo orden se puede usar para un primer rango de frecuencia objetivo inferior, y la transposición de tercer orden se puede usar para un segundo rango de frecuencia objetivo superior.

La Fig. 3 ilustra la operación de un transponedor armónico del dominio de la frecuencia (FD), tal como uno de los bloques individuales de 201, esto es uno de los transponedores 201- T de orden de transposición T . Un banco 301 de filtros de análisis emite las subbandas complejas que se envían al procesamiento 302 no lineal, que modifica la fase y/o la amplitud de la señal de subbanda según el orden T de transposición elegido. Las subbandas modificadas se alimentan al banco 303 de filtros de síntesis que emite la señal transpuesta en el dominio del tiempo. En el caso de múltiples transponedores paralelos de diferentes órdenes de transposición tal como se muestra en la Fig. 2, algunas operaciones del banco de filtros se pueden compartir entre los diferentes transponedores 201-2, 201-3, ..., 201- T_{max} . La compartición de las operaciones del banco de filtros se puede hacer para análisis o síntesis. En el caso de la síntesis 303 compartida, la suma 202 se puede realizar en el dominio de la subbanda, esto es antes de la síntesis 303.

La Fig. 4 ilustra la operación del procesamiento 402 de los términos de cruce además del procesamiento 401 directo. El procesamiento 402 de los términos de cruce y el procesamiento 401 directo se realiza en paralelo dentro del bloque 302 de procesamiento no lineal del transponedor armónico del dominio de la frecuencia de la Fig. 3. Las señales de salida transpuestas se combinan, por ejemplo se suman, para proporcionar una señal conjunta transpuesta. Esta combinación de señales de salida transpuesta puede consistir en la superposición de las señales de salida transpuestas. De manera opcional, la adición selectiva de los términos de cruce se puede implementar en el cálculo de la ganancia.

La Fig. 5 ilustra en más detalle la operación del bloque 401 de procesamiento directo de la Fig. 4 dentro del transponedor armónico en el dominio de la frecuencia de la Fig. 3. Las unidades 401-1, ..., 401-n, ..., 401-N de Entrada Única Salida Única (SISO) hacen corresponder cada subbanda de análisis de un rango de origen en una subbanda de síntesis de un rango objetivo. Según la Fig. 5, una subbanda de análisis de índice n es hecha corresponder por la unidad 401-n SISO a una subbanda de síntesis del mismo índice n . Se debería observar que el rango de frecuencia de la subbanda con índice n en el banco de filtros de síntesis puede variar dependiendo de la

versión o el tipo exacto de la transposición armónica. En la versión o tipo ilustrado en la Fig. 5, el espaciado de frecuencia del banco 301 de análisis es un factor T menor que el del banco 303 de síntesis. Por tanto, el índice n en el banco 303 de síntesis corresponde a una frecuencia, que es T veces mayor que la frecuencia de la subbanda con el mismo índice n en el banco 301 de análisis. A modo de ejemplo, se transpone una subbanda de análisis [(n-1)ω, nω] en una subbanda de síntesis [(n-1)Tω, nTω].

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

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

Esto se puede escribir también como:

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

En otras palabras, la fase de la señal x de subbanda compleja es multiplicada por el orden T de transposición y la amplitud de la señal x de subbanda compleja es modificada por el parámetro g de ganancia.

La Fig. 7 ilustra las componentes del procesamiento 402 de términos de cruce para una transposición armónica de orden T. Existen T - 1 bloques de procesamiento de términos de cruce en paralelo, 701-1, ..., 701-r, ... 701-(T-1), cuyas salidas se suman en la unidad 702 de suma para producir una salida combinada. Como ya se apuntó en la sección introductoria, es un objetivo hacer corresponder una pareja de sinusoides con frecuencias (ω, ω+Ω) a una sinusoide con frecuencia (T-r)ω + r(ω+Ω) = Tω + rΩ, en donde la variable r varía desde 1 a T-1. En otras palabras, las dos subbandas del banco 301 de filtros de análisis se han de hacer corresponder a una subbanda del rango de alta frecuencia. Para un valor concreto de r y un orden T de transposición dado, este paso de correspondencia se realiza en el bloque 701-r de procesamiento de términos de cruce.

La Fig. 8 ilustra la operación de un bloque 701-r de procesamiento de términos de cruce para un valor fijo r = 1, 2, ..., T-1. Cada subbanda 803 de salida se obtiene en una unidad 800-n de múltiples entradas salida única (MISO) a partir de las dos subbandas 801 y 802 de entrada. Para una subbanda 803 de salida de índice n, las dos entradas de la unidad 800-n MISO son subbandas n - p₁, 801, y n + p₂, 802, donde p₁ y p₂ son desplazamientos de índice enteros positivos, que dependen del orden T de transposición, la variable r, y el parámetro Ω de tono mejorado de productos de cruce. La convención de numeración de la subbanda de análisis y síntesis se mantiene en línea con el de la Fig. 5, esto es, el espaciado en la frecuencia del banco 301 de análisis es un factor T menor que el del banco 303 de síntesis y por consiguiente los comentarios anteriores dados sobre las variaciones del factor T siguen siendo relevantes.

En relación con el uso del procesamiento de términos de cruce, se deberían considerar las siguientes observaciones. El parámetro Ω de tono no tiene que ser conocido con una gran precisión, y ciertamente no con una mejor resolución de frecuencia que la resolución de frecuencia obtenida mediante el banco 301 de filtros de análisis. De hecho, en algunas realizaciones de la presente invención, el parámetro Ω de tono mejorado de productos de cruce subyacente no se introduce en el decodificador para nada. En su lugar, la pareja elegida de desplazamientos (p₁, p₂) de índice enteros 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 productos de cruce, esto es la maximización de la energía de la salida de productos de cruce. A modo de ejemplo, para los valores dados de T y r, se podría usar una lista de candidatos dada por la fórmula (p₁, p₂) = (r*l*, (T-r)*l*), *l* ∈ L, donde L es una lista de enteros positivos. Esto se muestra con mayor detalle más adelante en el contexto de la fórmula (11). Todos los enteros positivos son en principio candidatos válidos. En algunos casos la información de tono puede ayudar a identificar que *l* elegir como desplazamientos de índice apropiados.

Además, incluso aunque el procesamiento de los productos de cruce de ejemplo ilustrado en la Fig. 8 sugiere que los desplazamientos (p₁, p₂) de índice aplicados son los mismos para un cierto rango de subbandas de salida, por ejemplo, las subbandas (n-1), n y (n+1) de síntesis están compuestas a partir de subbandas de análisis que tienen una distancia p₁ + p₂ fija, este no debe ser necesariamente el caso. De hecho, los desplazamientos (p₁, p₂) de índice pueden diferir para cada una y todas las subbandas de salida. Esto significa que para cada subbanda n se puede seleccionar un valor Ω diferente del parámetro de tono de mejora de productos de cruce.

La Fig. 9 ilustra el procesamiento no lineal contenido en cada una de las unidades 800-n MISO. La operación 901 de producto 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 902 de ganancia opcional modifica la magnitud de las muestras de

subbanda de fase modificada. En términos matemático, la salida y se puede escribir como una función de las entradas u_1 801 y u_2 802 a la unidad 800-n MISO y el parámetro g de ganancia como sigue,

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

Esto se puede escribir también como:

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

5 donde $\mu(|u_1|, |u_2|)$ es una función de generación de magnitud. En otras palabras, la fase de la señal u_1 de subbanda compleja es multiplicada por el orden $T - r$ de transposición y la fase de la señal u_2 de subbanda compleja es multiplicada por el orden r de transposición. La suma de las dos fases se usa como la fase de la salida y cuya magnitud es obtenida mediante la función de generación de la magnitud. Comparando con la fórmula (2) la función de generación de la magnitud se expresa como la media geométrica de las magnitudes modificada por el parámetro g de ganancia, esto es $\mu(|u_1|, |u_2|) = g \cdot |u_1|^{1-r/T} |u_2|^{r/T}$. Permitir al parámetro de ganancia depender de las entradas cubre por supuesto todas las posibilidades.

10 Se debería observar que la fórmula (2) resulta a partir del objetivo subyacente en que se han de hacer corresponder un par de sinusoides con frecuencias $(\omega, \omega + \Omega)$ a una sinusoide con frecuencia $T\omega + r\Omega$, que se puede escribir también como $(T - r)\omega + r(\omega + \Omega)$.

15 En el texto siguiente, se explicará en términos generales la descripción matemática de la presente invención. Por simplicidad, se consideran las señales de tiempo continuo. Se supone que el banco 303 de filtros de síntesis consigue la reconstrucción perfecta a partir de un banco 301 de filtros de análisis modulados complejos correspondiente con una función de ventana simétrica o filtro de valor real o un filtro $w(t)$ prototipo. El banco de filtros de síntesis usará a menudo, pero no siempre, la misma ventana en el proceso de síntesis. La modulación se supone que es de un tipo apilado uniformemente, el paso se normaliza a uno y el espaciado de la frecuencia angular de las subbandas de síntesis se normaliza a π . Por tanto, se conseguirá una señal $x(t)$ objetivo en la salida del banco del filtro de síntesis si las señales de subbanda de entrada al banco del filtro de síntesis están dadas mediante las señales $y_n(k)$ de subbanda de síntesis.

$$25 \quad y_n(k) = \int_{-\infty}^{\infty} s(t)w(t-k)\exp[-in\pi(t-k)]dt. \quad (3)$$

30 Observe que la fórmula (3) es un modelo matemático en tiempo continuo normalizado de las operaciones normales en un banco de filtros de análisis de subbanda modulados complejos, tal como una Transformada de Fourier Discreta (DFT) con ventana, también denotada como Transformada de Fourier de tiempo reducido (STFT). Con una pequeña modificación en el argumento de la exponencial compleja de la fórmula (3), se obtienen modelos en tiempo continuo para el Banco de Filtros de Espejo en Cuadratura (QMF) (pseudo) modulados complejos y la Transformada de Coseno Discreta Modificada Complejizada (CMDCT), también denotada como DFT en ventana apilada particularmente en ventana. El índice n de subbanda funciona a través de todos los números enteros negativos para el caso de tiempo continuo. Para las contrapartes en tiempo discreto, la variable t de tiempo se muestra en el paso $1 / N$, y el índice n de subbanda es limitado por N, donde N es el número de subbandas en el banco de filtros, que es igual al paso de tiempo discreto del banco de filtros. En el caso de tiempo discreto, se requiere también un factor de normalización relacionado con N en la operación de transformada si no se incorpora en el escalado de la ventana.

35 Para una señal de valor real, existen tantas muestras de subbanda complejas fuera como muestras de valor real dentro para el modelo del banco de filtro elegido. Por lo tanto, existe un sobre muestreo total (o una redundancia) por un factor de dos. Los bancos de filtros con un mayor grado de sobre muestreo se pueden emplear también, pero el sobre muestreo se mantiene pequeño en la presente descripción de las realizaciones por claridad de exposición.

40 Los pasos principales involucrados en el análisis del banco de filtros modulados correspondiente a la fórmula (3) son que la señal es multiplicada por una ventana centrada alrededor del tiempo $t = k$, y la señal en ventana resultante se correlaciona con cada una de las sinusoides complejas $\exp[-in\pi(t-k)]$. En implementaciones de tiempo discreto esta correlación se implementa de manera eficiente a través de la Transformada Rápida de Fourier. Los siguientes pasos algorítmicos para el banco de filtros de síntesis son bien conocidos para aquellos expertos en la técnica, y consisten de operaciones de modulación de síntesis, de ventana de síntesis, y de adición de superposición.

45 La Fig. 19 ilustra la posición en el tiempo y la frecuencia correspondiente a la información transportada por la muestra $y_n(k)$ de subbanda para la selección de los valores del índice k de tiempo y el índice n de subbanda. Como ejemplo, la muestra $y_5(4)$ de subbanda es representada mediante el rectángulo 1901 oscuro.

50 Para una sinusoide, $s(t) = \text{Acos}(\omega t + \theta) = \text{Re}\{C\exp(i\omega t)\}$, las señales de subbanda de (3) son para una n

suficientemente grande con una buena aproximación dada por

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

donde el sombrero denota la transformada de Fourier, esto es \hat{w} es la transformada de Fourier de la función w de ventana.

- 5 Estrictamente hablando, la fórmula (4) es cierta sólo si añade un término con $-\omega$ en lugar de ω . Este término se abandona en base a la suposición de que la respuesta en frecuencia de la ventana decae suficientemente rápido, y que la suma de ω y n no está cerca de cero.

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

- 10 La Fig. 21 ilustra el análisis de una única senoide que corresponde a la fórmula (4). Las subbandas que se ven afectadas principalmente por la senoide a la frecuencia ω son aquellas con índice n de manera tal que $n\pi - \omega$ sea pequeño. Para el ejemplo de la Fig. 21, la frecuencia es $\omega=6,25\pi$ tal como se indica mediante la línea 2101 horizontal discontinua. En ese caso, las tres subbandas para $n = 5, 6, 7$, representadas por los signos 2102, 2103, 2104 de referencia, contienen, respectivamente señales de subbanda distintas de cero significativas. El conformado de estas tres subbandas refleja la amplitud relativa de las senoideas complejas dentro de cada subbanda obtenida a partir de la fórmula (4). Un tono más oscuro implica una mayor amplitud. En el ejemplo concreto, esto implica que la amplitud de la subbanda 5, esto es 2102, es inferior comparada con la amplitud de la subbanda 7, esto es 2104, que de nuevo es inferior que la amplitud de la subbanda 6, esto es 2103. Es importante observar que pueden ser necesarias en general diversas subbandas distintas de cero para ser capaz de sintetizar una senoide de alta calidad en la salida del banco de filtros de síntesis, especialmente en los 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 lóbulos laterales significativos en frecuencia.

- 15 Las señales $y_n(k)$ de la subbanda de síntesis se pueden determinar también como resultado del banco 301 de filtros de análisis y el procesamiento no lineal esto es el transponedor 302 armónico ilustrado en la Fig. 3. En el lado del banco de filtros de análisis, las señales $x_n(k)$ de subbanda de análisis se pueden representar como una función de la señal $z(t)$ de origen. Para una transposición de orden T , se aplica un banco de filtros de análisis modulados complejos con ventana $w_T(t) = w(t/T)/T$, un paso uno, y un paso de frecuencia de modulación, que es T veces más preciso que el paso de frecuencia del banco de síntesis, en la señal $z(t)$ de origen. La Fig. 22 ilustra la apariencia de la ventana w_T 2201 escalada y su transformada \hat{w}_T 2202 de Fourier. Comparada con la Fig. 20, la ventana 2201 de tiempo se despliega y la ventana 2202 de frecuencia se comprime.

- 30 El análisis mediante el banco de filtros modificado da lugar a las señales $x_n(k)$ de subbanda:

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

Para una senoide, $z(t) = B \cos(\xi t + \varphi) = \text{Re}\{D \exp(i\xi t)\}$, se encuentra que las señales de subbanda de (5) para una n suficientemente grande con una buena aproximación vienen dadas por

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

- 35 Por tanto, enviar estas señales de subbanda al transponedor 302 armónico y aplicando la regla (1) a (6) de transposición produce

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

Las señales $y_n(k)$ de subbanda de síntesis dadas por la fórmula (4) y las señales de subbanda no lineales obtenidas a través de la transposición $\tilde{y}_n(k)$ armónica dadas por la fórmula (7) idealmente deberían coincidir.

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

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

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

Para T pares, la coincidencia es más aproximada, pero aún se mantiene en la parte de valor positivo de la respuesta \hat{w} en frecuencia de la ventana, que para una ventana de valor real simétrica incluye el lóbulo principal más importante. Esto implica que también para valores pares de T se obtiene una transposición armónica de la señal z(t) de origen sinusoidal. En el caso particular de una ventana Gaussiana, \hat{w} es siempre positivo y por consiguiente, no existe diferencia en el rendimiento para órdenes pares o impares de la transposición.

De manera similar a la fórmula (6), el análisis de una senoide con frecuencia $\xi + \Omega$, esto es la señal de origen sinusoidal $z(t) = B' \cos((\xi + \Omega)t + \varphi')$ es $\text{Re}\{E \exp(i(\xi + \Omega)t)\}$

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

Por lo tanto, alimentando las dos señales $u_1 = x_{n-p_1}(k)$, que corresponde a la señal 801 en la Fig. 8, y $u_2 = x'_{n+p_2}(k)$, que corresponde a la señal 802 en la Fig. 8, en el procesamiento 800-n de productos de cruce ilustrado en la Fig. 8 y aplicando la fórmula (2) de productos de cruce produce la señal 803 de subbanda de salida

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

donde

$$M(n, \xi) = \frac{D^{T-r} E^r \hat{w}((n-p_1)\pi - T\xi)^{T-r} \hat{w}((n+p_2)\pi - T(\xi + \Omega))^r}{|D^{T-r} E^r|^{1-1/T} |\hat{w}((n-p_1)\pi - T\xi)^{T-r} \hat{w}((n+p_2)\pi - T(\xi + \Omega))^r|^{1-1/T}}. \quad (10)$$

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

Dado un parámetro Ω de tono de mejora de producto de cruce las elecciones adecuadas para los desplazamientos p_1 y p_2 de índice se pueden derivar en orden para la magnitud $M(n, \xi)$ compleja de (10) para aproximar $\hat{w}(n\pi - (T\xi + r\Omega))$ para un rango de subbandas n, en cuyo caso la salida final se aproximará a una senoide en la frecuencia $T\xi + r\Omega$. Una primera consideración sobre los lóbulos principales impone todos los tres valores de $(n - p_1)\pi - T\xi$, $(n + p_2)\pi - T(\xi + \Omega)$, $n\pi - (T\xi + r\Omega)$ sean menores de manera simultánea, lo que lleva a las igualdades aproximadas

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

Esto significa que cuando se conoce el parámetro Ω de tono mejorado de productos de cruce, los desplazamientos de índice se pueden aproximar mediante la fórmula (11), permitiendo de este modo una selección simple de las subbandas de análisis. Se puede realizar un análisis más completo de los efectos de la elección de los desplazamientos p_1 y p_2 según la fórmula (11) sobre la magnitud del parámetro $M(n, \xi)$ según la fórmula (10) para casos especiales importantes de funciones w(t) de ventana tales como la ventana Gaussiana y una ventana de seno. Se encuentra que la aproximación deseada de $\hat{w}(n\pi - (T\xi + r\Omega))$ es muy buena para diversas subbandas con $n\pi \approx T\xi + r\Omega$.

Se debería observar que la relación (11) se calibra a la situación ejemplar donde el banco 301 de filtros de análisis tiene un espaciado de subbanda de frecuencia angular de π/T . En el caso general, la interpretación resultante de (11) es que la extensión $p_1 + p_2$ de origen de términos de cruce es un número entero que aproxima la frecuencia Ω fundamental subyacente, medida en unidades de espaciado de subbanda del banco de filtros de análisis, y que la pareja (p_1, p_2) se elige como un múltiple de $(r, T - r)$.

Para la determinación de la pareja (p_1, p_2) de desplazamientos de índice en el decodificador se pueden usar los siguientes modos:

1. Un valor de Ω se puede derivar en el proceso de codificación y transmitirse explícitamente al decodificador con una precisión suficiente para derivar los valores enteros de p_1 y p_2 por medio de un procedimiento de redondeado adecuado, que puede seguir los principios de que

- o $p_1 + p_2$ se aproxima a $\Omega/\Delta\omega$, donde $\Delta\omega$ es el espaciado de la frecuencia angular del banco de filtros de análisis; y
- o p_1/p_2 se elige para aproximarse a $r/(T - r)$.

5 2. Para cada muestra de subbanda de objetivo, la pareja (p_1, p_2) de desplazamientos de índice se puede derivar en el decodificador a partir de una lista predeterminada de valores candidatos tales 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 números enteros positivos. La selección se puede basar en una optimización de la magnitud de salida de términos de cruce, por ejemplo, una maximización de la energía de la salida de los términos de cruce.

10 3. Para cada muestra de subbanda objetivo, la pareja (p_1, p_2) de desplazamientos de índice se puede derivar a partir de una lista reducida de valores de candidatos mediante una optimización de la magnitud de salida de los términos de cruce, donde la lista reducida de valores de candidatos se deriva en el proceso de codificación y se transmite al decodificador.

15 Se debería observar que la modificación de fase de las señales de subbanda u_1 y u_2 se realiza con un peso $(T-r)$ y r , respectivamente, pero las distancias p_1 y p_2 de índice de subbanda se eligen proporcionales a r y $(T - r)$, respectivamente. Por tanto, la subbanda más cercana a la subbanda n de síntesis recibe la modificación de fase más fuerte.

Un método ventajoso para el procedimiento de optimización para los modos 2 y 3 explicados en términos generales anteriormente puede ser considerar la optimización Max-Min:

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

20 y para usar la pareja ganadora entre sí con su valor correspondiente de r para construir la contribución del producto de cruce para un índice n de subbanda objetivo dado. En los modos orientados de búsqueda del decodificador 2 y parcialmente también en el 3, la adición de términos de cruce para diferentes valores de r se hace preferiblemente de manera independiente, ya que puede existir riesgo de añadir contenido a la misma subbanda varias veces. Si, por otro lado, se usa la frecuencia Ω fundamental para seleccionar las subbandas como en el modo 1 o si sólo se permiten un rango reducido de distancias de índice de subbanda como puede ser el caso en el modo 2, se puede evitar este problema concreto de añadir contenido a la misma subbanda varias veces.

25 Además, se debería observar también que para las realizaciones de los esquemas de procesamiento de términos de cruce explicados en términos generales anteriormente puede ser beneficiosa una modificación del decodificador adicional de la ganancia g de los productos de cruce. Por ejemplo, las señales u_1, u_2 de subbanda de entrada se refieren a la unidad MISO de productos de cruce dada por la fórmula (2) y la señal x de subbanda de entrada a la unidad SISO de transposición dada por la fórmula (1). Si las tres señales se han de alimentar a la misma subbanda de síntesis de salida como se muestra en la Fig. 4, donde el procesamiento 401 directo y el procesamiento 402 de productos de cruce proporciona los componentes para la misma subbanda de síntesis de salida, puede ser deseable establecer la ganancia g de los productos de cruce a cero, esto es la unidad 902 de ganancia de la Fig. 9, si

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

35 para un umbral predefinido $q > 1$. En otras palabras, la adición del producto de cruce se realiza sólo si la magnitud $|x|$ de subbanda de entrada del término directo es pequeña comparada con ambos términos de entrada del producto de cruce. En este contexto, x es la muestra de subbanda de análisis para el procesamiento de término directo que lleva a una salida en la misma subbanda de síntesis que el producto de cruce bajo consideración. Esto puede ser una precaución para no mejorar más un componente armónico que ya ha sido equipado por la transposición directa.

40 A continuación, el método de transposición armónica explicado en términos generales en el presente documento se describirá para las configuraciones espectrales ejemplares para ilustrar las mejoras sobre la técnica anterior. La Fig. 10 ilustra el efecto de la transposición armónica directa de orden $T = 2$. El diagrama 1001 superior representa las componentes de frecuencia parciales de la señal original mediante flechas verticales posicionadas en múltiplos de la frecuencia Ω fundamental. Éste ilustra la señal de origen, por ejemplo, en el lado del codificador. El diagrama 1001 se segmenta en un rango de frecuencia de origen del lado izquierdo con frecuencias $\Omega, 2\Omega, 3\Omega, 4\Omega, 5\Omega$ parciales y un rango de frecuencia objetivo del lado derecho con frecuencias $6\Omega, 7\Omega, 8\Omega$ parciales. El rango de frecuencia de origen se codificará y transmitirá normalmente al decodificador. Por otro lado, el rango de frecuencia objetivo del lado derecho, que comprende las frecuencias $6\Omega, 7\Omega, 8\Omega$ parciales por encima de la frecuencia 1005 de corte del método HFR, normalmente no se transmitirá al decodificador. Es un objetivo del método de transposición armónica reconstruir el rango de frecuencia objetivo por encima de la frecuencia 1005 de corte de la señal de origen a partir del rango de frecuencia de origen. Por consiguiente, el rango de frecuencia objetivo, y especialmente las frecuencias $6\Omega, 7\Omega, 8\Omega$ parciales en el diagrama 1001 no están disponibles como entrada al transponedor.

Como se explicó en términos generales anteriormente, es un objetivo del método de transposición armónica regenerar las componentes 6Ω , 7Ω , 8Ω de señal de la señal de origen a partir de las componentes de frecuencia disponibles en el rango de frecuencia de origen. El diagrama 1002 inferior muestra la salida del transponedor en el rango de frecuencia objetivo del lado derecho. Dicho transponedor se puede ubicar por ejemplo en el lado del decodificador. Las parciales a frecuencias 6Ω y 8Ω se regeneran a partir de las parciales a frecuencias 3Ω y 4Ω mediante transposición armónica usando un orden de transposición $T = 2$. Como resultado del efecto de extensión espectral de la transposición armónica, representada aquí por las flechas 1003 y 1004 punteadas, se pierde la parcial objetivo a 7Ω . Esta parcial objetivo a 7Ω no se puede generar usando el método de transposición armónica de la técnica anterior subyacente.

La Figura 11 ilustra el efecto de la invención para la transposición armónica de una señal periódica en el caso en que un transponedor armónico de segundo orden sea mejorado mediante un término de cruce único, esto es $T = 2$ y $r = 1$. Tal como se explicó en términos generales en el contexto de la Fig. 10, se usa un transponedor para generar las parciales 6Ω , 7Ω , 8Ω en el rango de frecuencia objetivo por encima de la frecuencia 1105 de corte en el diagrama 1102 inferior a partir de las parciales Ω , 2Ω , 3Ω , 4Ω , 5Ω en el rango de frecuencia de origen por debajo de la frecuencia 1105 de corte del diagrama 1101. Además de la salida del transponedor de la técnica anterior de la Figura 10, la componente de frecuencia parcial a 7Ω se regenera a partir de una combinación de las parciales de origen a 3Ω y 4Ω . El efecto de la adición del producto de cruce es representado por las flechas 1103 y 1104 punteadas. En términos de 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 se puede ver a partir de este ejemplo, todas las parciales objetivo se pueden regenerar usando el método HFR inventivo explicado en términos generales en el presente documento.

La Fig. 12 ilustra una posible implementación de un transponedor armónico de segundo orden de la técnica anterior en un banco de filtros modulados para la configuración espectral de la Fig. 10. Las respuestas de frecuencia estilizadas de las subbandas del banco de filtros de análisis son mostradas por las líneas punteadas, por ejemplo el signo 1206 de referencia, en el diagrama 1201 superior. Las subbandas son enumeradas mediante el índice de subbanda, de los cuales los índices 5, 10 y 15 se muestran en la Fig. 12. Para el ejemplo dado, la frecuencia Ω fundamental es igual a 3,5 veces el espaciado de la frecuencia de la subbanda de análisis. Esto es ilustrado mediante el hecho de que la Ω parcial en el diagrama 1201 se posiciona entre las dos subbandas con el índice 3 y 4 de subbanda. La 2Ω parcial se posiciona en el centro de la subbanda con índice 7 de subbanda y así sucesivamente.

El diagrama 1202 inferior muestra las parciales 6Ω y 8Ω regeneradas superpuestas con las respuestas de frecuencia estilizadas, por ejemplo el signo 1207 de referencia, de las subbandas del banco de filtros de síntesis. Como se describió anteriormente, estas subbandas tienen un espaciado de frecuencia $T = 2$ veces más grande. Por consiguiente, también las respuestas de frecuencia son escaladas por el factor $T = 2$. Como se explicó en términos generales anteriormente, el método de procesamiento de términos directos de la técnica anterior modifica la fase de cada subbanda de análisis, esto es de cada subbanda por debajo de la frecuencia 1205 de corte en el diagrama 1201, por un factor $T = 2$ y hace corresponder el resultado en la subbanda de síntesis con el mismo índice, esto es una subbanda por encima de la frecuencia 1205 de corte en el diagrama 1202. Esto es simbolizado en la Fig. 12 mediante flechas punteadas diagonales, por ejemplo, la flecha 1208 para la subbanda 1206 de análisis y la subbanda 1207 de síntesis. El resultado de este procesamiento del término directo para las subbandas con índices 9 a 16 de subbanda a partir de la subbanda 1201 de análisis es la regeneración de las dos parciales objetivo a las frecuencias 6Ω y 8Ω en la subbanda 1202 de síntesis a partir de las parciales de origen a las frecuencias 3Ω y 4Ω . Como se puede ver a partir de la Fig. 12, la contribución principal de la parcial 6Ω objetivo viene de las subbandas con los índices 10 y 11 de subbanda, esto es los signos 1209 y 1210 de referencia, y la contribución principal a la parcial 8Ω objetivo viene de la subbanda con índice 14 de subbanda, esto es el signo 1211 de referencia.

La Fig. 13 ilustra una posible implementación de un paso adicional de procesamiento de términos de cruce en el banco de filtros modulados de la Fig. 12. El paso de procesamiento de términos de cruce corresponde al descrito para señales periódicas con la frecuencia Ω fundamental en relación a la Fig. 11. El diagrama 1301 superior ilustra las subbandas de análisis, de las cuales el rango de frecuencia de origen se ha de transponer en el rango de frecuencia objetivo de las subbandas de síntesis en el diagrama 1302 inferior. Se considera el caso concreto de la generación de las subbandas 1315 y 1316 de síntesis, que rodean a la parcial 7Ω , a partir de las subbandas de análisis. Para un orden de transposición $T = 2$, se puede seleccionar un posible valor $r = 1$. Elegir la lista de valores

$$\frac{\Omega}{\Delta\omega} = \frac{\Omega}{(\Omega/3.5)} = 3.5$$

(p_1, p_2) candidatos como un múltiplo $(r, T - r) = (1, 1)$ de manera tal que $p_1 + p_2$ aproxima $\frac{\Omega}{\Delta\omega} = 3.5$, esto es la frecuencia Ω fundamental en unidades de espaciado de frecuencia de subbanda de análisis, lleva a la elección $p_1 = p_2 = 2$. Como se explicó en términos generales en el contexto de la Fig. 8, una subbanda de síntesis con el índice n de subbanda se puede generar a partir del producto de términos de cruce de las subbandas de análisis con el índice de subbandas $(n-p_1)$ y $(n+p_2)$. Por lo tanto. Para la subbanda de síntesis con el índice 12 de subbanda, esto es el signo 1315 de referencia, un producto de cruce se forma a partir de las subbandas de análisis con el índice de subbanda $(n - p_1) = 12 - 2 = 10$, esto es el signo 1311 de referencia, y $(n + p_2) = 12 + 2 = 14$, esto es el signo 1313 de referencia. Para la subbanda de síntesis con índice 13 de subbanda, se forma un producto de cruce a partir de las subbandas de análisis con un índice $(n - p_1) = 13 - 2 = 11$, esto es el signo 1312 de referencia, y $(n + p_2) = 13 + 2 = 15$, esto es el signo 1314 de referencia. Este proceso de generación del producto de cruce es simbolizado mediante las parejas de flechas discontinuas / punteadas diagonales, esto es las parejas 1308, 1309 y

1306, 1307, respectivamente de signos de referencia.

Como se puede ver a partir de la Fig. 13, la parcial 7Ω se coloca principalmente dentro de la subbanda 1315 con índice 12 y sólo de manera secundaria en la subbanda 1316 con índice 13. Por consiguiente, para unas respuestas del filtro más realistas, existirán más términos directos y/o de cruce alrededor de la subbanda 1315 de síntesis con índice 12 que se añadirán de manera beneficiosa a la síntesis de un senoide de alta calidad a la frecuencia $(T - r)\omega + r(\omega + \Omega) = T\omega + r\Omega = 6\Omega + \Omega = 7\Omega$ que los términos alrededor de la subbanda 1316 de síntesis con índice 13. Además, como se destacó en el contexto de la fórmula (13), una adición ciega de todos los términos de cruce con $p_1 = p_2 = 2$ podría llevar a componentes de señal no deseados para señales de entrada menos periódicas y académicas. Por consiguiente, este fenómeno de componentes de señal no deseados puede requerir la aplicación de una regla de cancelación de productos de cruce adaptativa tal como la regla dada por la fórmula (13).

La Fig. 14 ilustra el efecto de la transposición armónica de la técnica anterior de orden $T = 3$. El diagrama 1401 superior representa los componentes de frecuencia parciales de la señal original mediante flechas verticales posicionadas a múltiplos de la frecuencia Ω fundamental. Las parciales $6\Omega, 7\Omega, 8\Omega, 9\Omega$ están en el rango objetivo por encima de la frecuencia 1405 de corte del método HFR y por lo tanto no están disponibles como entradas al transponedor. El objetivo de la transposición armónica es regenerar estas componentes de señal a partir de la señal en el rango de origen. El diagrama 1402 inferior muestra la salida del transponedor en el rango de frecuencia objetivo. Las parciales a frecuencias 6Ω , esto es el signo 1407 de referencia, y 9Ω , esto es el signo 1410 de referencia, se han regenerado a partir de las parciales a las frecuencias 2Ω , esto es el signo 1406 de referencia, y 3Ω , esto es el signo 1409 de referencia. Como resultado del efecto de expansión espectral de la transposición armónica, representada aquí mediante las flechas 1408 y 1411 punteadas, respectivamente, las parciales objetivo a 7Ω y 8Ω se pierden.

La Fig. 15 ilustra el efecto de la invención para la transposición armónica de una señal periódica en el caso en que un transponedor armónico de tercer orden sea mejorado mediante la adición de dos términos de cruce diferentes, esto es $T = 3$ y $r = 1, 2$. Además de la salida del transponedor de la técnica anterior de la Fig. 14, la componente 1508 de frecuencia parcial en 7Ω es regenerada mediante el término de cruce para $r = 1$ a partir de una combinación de las parciales 1506 y 1507 de origen a 2Ω y 3Ω . El efecto de la adición de productos de cruce es representado mediante las flechas 1510 y 1511 punteadas. En términos de fórmulas, se tiene que con $\omega = 2\Omega$, $(T - r)\omega + r(\omega + \Omega) = T\omega + r\Omega = 6\Omega + \Omega = 7\Omega$. Igualmente, la componente 1509 de frecuencia parcial a 8Ω es regenerada por el término de cruce para $r = 2$. Esta componente 1509 de frecuencia parcial en el rango objetivo del diagrama 1502 inferior es generado a partir de las componentes parciales 1506 a 2Ω y 1507 a 3Ω en el rango de frecuencia de origen del diagrama 1501 superior. La generación del producto de términos de cruce es representada mediante las flechas 1512 y 1513. En términos de fórmulas, se tiene que $(T - r)\omega + r(\omega + \Omega) = T\omega + r\Omega = 6\Omega + 2\Omega = 8\Omega$. Como se puede ver, todas las parciales objetivo se pueden regenerar usando el método HFR inventivo descrito en el presente documento.

La Fig. 16 ilustra una posible implementación de un transponedor armónico de tercer orden de la técnica anterior en un banco de filtros modulados para la situación espectral de la Fig. 14. Las respuestas de frecuencia estilizada de las subbandas del banco de filtros de análisis son mostradas por líneas punteadas en el diagrama 1601 superior. Las subbandas son enumeradas por los índices 1 a 17 de subbanda de los cuales la subbandas 1606, con índice 7, la 1607, con índice 10 y la 1608, con índice 11 son referenciadas de una manera ejemplar. Para el ejemplo dado, la frecuencia Ω fundamental es igual a 3,5 veces el espaciado $\Delta\omega$ de frecuencia de la subbanda de análisis. El diagrama 1602 inferior muestra la frecuencia parcial regenerada superpuesta con las respuestas de frecuencia estilizada de las subbandas del banco de filtros de síntesis seleccionadas. A modo de ejemplo, se referencian la subbanda 1609, con índice 7, la 1610, con índice 10 de subbanda y la 1611, con índice 11 de subbanda. Como se describe anteriormente, estas subbandas tienen un espaciado $\Delta\omega$ de frecuencia $T = 3$ veces más grande. Correspondientemente, también se escalan las respuestas en frecuencia, por consiguiente

El procesamiento de términos directos de la técnica anterior modifica la fase de las señales de subbanda por un factor $T = 3$ para cada subbanda de análisis y hace corresponder el resultado en la subbanda de síntesis con el mismo índice, tal como se simboliza mediante las flechas punteadas diagonales. 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 parciales de origen a frecuencia 2Ω y 3Ω . Tal como se puede ver a partir de la Fig. 16, la contribución principal a la parcial 6Ω objetivo viene de la subbanda con índice 7, esto es el signo 1606 de referencia, y las contribuciones principales a la parcial 9Ω objetivo viene de las subbandas con índice 10 y 11, esto es los signos 1607 y 1608 de referencia, respectivamente.

La Fig. 17 ilustra una posible implementación de un paso de procesamiento de términos de cruce adicional para $r = 1$ en el banco de filtros modulados de la Fig. 16 que lleva a la regeneración de la parcial a 7Ω . Tal como se explicó en términos generales en el contexto de la Fig. 8 los desplazamientos (p_1, p_2) de índice pueden ser seleccionados como un múltiplo de $(r, T - r) = (1, 2)$, de manera tal que $p_1 + p_2$ se aproxima a 3,5, esto es la frecuencia Ω fundamental en unidades del espaciado $\Delta\omega$ de frecuencia de subbanda de análisis. En otras palabras, la distancia relativa, esto es la distancia en el eje de frecuencia dividida por el espaciado $\Delta\omega$ de frecuencia de subbanda de análisis, entre las dos subbandas de análisis que contribuyen a la subbanda de síntesis que se ha de generar, debería aproximar mejor la frecuencia fundamental relativa, esto es la frecuencia Ω fundamental dividida entre el espaciado $\Delta\omega$ de frecuencia de subbanda de análisis. Esto es expresado también mediante las fórmulas (11) y lleva a la elección $p_1 = 1, p_2 = 2$.

Tal como se muestra en la Fig. 17, la subbanda de síntesis con índice 8, esto es el signo 1710 de referencia, se obtiene a partir de un producto de cruce formado a partir de las subbandas de análisis con índice $(n - p_1) = 8 - 1 = 7$, esto es el signo 1706 de referencia, y $(n + p_2) = 8 + 2 = 10$, esto es el signo 1708 de referencia. Para la subbanda de síntesis con índice 9, se forma un producto de cruce a partir de las subbandas de análisis con índice $(n - p_1) = 9 - 1 = 8$, esto es el signo 1707 de referencia, y $(n + p_2) = 9 + 2 = 11$, esto es el signo 1709 de referencia. Este proceso de formación de productos de cruce es simbolizado mediante las parejas de flechas discontinuas/punteadas diagonales, esto es las parejas 1712, 1713 y 1714, 1715 de flechas, respectivamente. Se puede ver a partir de la Fig. 17 que la frecuencia 7Ω parcial se posiciona de manera más prominente en la subbanda 1710 que en la subbanda 1711. Por consiguiente, se ha de esperar que para unas respuestas realistas del filtro, existirán más términos de cruce alrededor de la subbanda de síntesis con índice 8, esto es la subbanda 1710, que se añaden de manera beneficiosa a la síntesis de una sinusoide de alta calidad a la frecuencia $(T - r)\omega + r(\omega + \Omega) = T\omega + r\Omega = 6\Omega + \Omega = 7\Omega$.

La Fig. 18 ilustra una posible implementación de un paso de procesamiento de términos de cruce adicional para $r = 2$ en el banco de filtros modulados de la Fig. 16 que lleva a la regeneración de la frecuencia parcial en 8Ω . Los desplazamientos (p_1, p_2) de índice se pueden seleccionar como un múltiplo de $(r, T - r) = (2, 1)$, de manera tal que $p_1 + p_2$ se aproxime a 3,5, esto es la frecuencia Ω fundamental en unidades de espaciado $\Delta\omega$ de frecuencia de subbanda de análisis. Esto lleva a la elección $p_1 = 2, p_2 = 1$. Como se muestra en la Fig. 18, la subbanda de síntesis con índice 9, esto es el signo 1810 de referencia, se obtiene a partir del producto de cruce formado a partir de las subbandas de análisis con índice $(n - p_1) = 9 - 2 = 7$, esto es el signo 1806 de referencia, y $(n + p_2) = 9 + 1 = 10$, esto es el signo 1808 de referencia. Para la subbanda de síntesis con índice 10, se forma un producto de cruce a partir de las subbandas de análisis con índice $(n - p_1) = 10 - 2 = 8$, esto es el signo 1807 de referencia, y $(n + p_2) = 10 + 1 = 11$, esto es el signo 1809 de referencia. Este proceso de formación de productos de cruce es simbolizado mediante las parejas de flechas discontinuas/punteadas diagonales, esto es las parejas 1812, 1813 y 1814, 1815 de flechas, respectivamente. Se puede ver a partir de la Fig. 18 que la frecuencia 8Ω parcial se posiciona ligeramente de manera más prominente en la subbanda 1810 que en la subbanda 1811. Por consiguiente, se ha de esperar que para tener respuestas del filtro realistas, existirán más términos directos y/o de cruce alrededor de la subbanda de síntesis con índice 9, esto es la subbanda 1810, que se añaden de manera beneficiosa para la síntesis de una sinusoide de alta calidad a la 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 la optimización Max-Min en base al procedimiento (12) de selección para la pareja (p_1, p_2) de desplazamiento de índice y r según esta regla para $T = 3$. El índice de subbanda objetivo elegido es $n = 18$ y el diagrama superior suministra un ejemplo de la magnitud de una señal de subbanda para un índice de tiempo dado. La lista de enteros posibles está dada aquí por los siete valores $L = \{2, 3, \dots, 8\}$.

La Fig. 23 ilustra la búsqueda para candidatos con $r = 1$. La subbanda objetivo o de análisis se muestra con el índice $n = 18$. La línea 2301 punteada destaca la subbanda con el índice $n = 18$ en el rango de subbanda de análisis superior y el rango de subbanda de síntesis inferior. Las posibles parejas de desplazamiento de índice son $(p_1, p_2) = \{(2, 4), (3, 6), \dots, (8, 16)\}$, para $l = 2, 3, \dots, 8$, respectivamente, y las correspondientes parejas de índice de muestra de magnitud de subbanda de análisis, esto es la lista de parejas de índice de subbanda que se consideran para determinar el término de cruce óptimo, son $\{(16, 22), (15, 24), \dots, (10, 34)\}$. El conjunto de flechas ilustra las parejas bajo consideración. Como un ejemplo, se muestra la pareja $(15, 24)$ denotada por los signos 2302 y 2303 de referencia. La evaluación del mínimo de estas parejas de magnitud da la lista $(0, 4, 1, 0, 0, 0, 0)$ para las respectivas magnitudes mínimas para la posible lista de términos de cruce. Ya que la segunda entrada para $l = 3$ es máxima, la pareja $(15, 24)$ gana de entre los candidatos con $r = 1$, y esta selección es representada mediante las flechas gruesas.

La Fig. 24 ilustra de manera similar la búsqueda para candidatos con $r = 2$. La subbanda objetivo o de síntesis se muestra con el índice $n = 18$. La línea 2401 punteada destaca la subbanda con el índice $n = 18$ en el rango de subbanda de análisis superior y el rango de subbanda de síntesis inferior. En este caso, las posibles parejas de desplazamiento de índice son $(p_1, p_2) = \{(4, 2), (6, 3), \dots, (16, 8)\}$ y las correspondientes parejas de índice de muestra de magnitud de subbanda de análisis son $\{(14, 20), (12, 21), \dots, (2, 26)\}$, de las cuales la pareja $(6, 24)$ es representada mediante los signos 2402 y 2403 de referencia. La evaluación del mínimo de estas parejas de magnitud da la lista $(0, 0, 0, 0, 3, 1, 0)$. Ya que la quinta entrada es máxima, esto es $l = 6$, la pareja $(6, 24)$ gana de entre los candidatos con $r = 2$, tal como se representa mediante las flechas gruesas. En general, ya que el mínimo de la pareja de magnitud correspondiente es menor que el de la pareja de subbanda seleccionada para $r = 1$, la selección final para el índice $n = 18$ de subbanda objetivo cae en la pareja $(15, 24)$ y $r = 1$.

Se debería observar además que cuando la señal $z(t)$ de entrada es una serie armónica con una frecuencia Ω fundamental, esto es con una frecuencia fundamental que corresponde con el parámetro de tono mejorado de producto de cruce, y Ω es suficientemente grande comparado con la resolución de frecuencia del banco de filtros de análisis, las señales $x_n(k)$ de subbanda de análisis dadas por la fórmula (6) y $x'_n(k)$ dada por la fórmula (8) son buenas aproximaciones del análisis de la señal $z(t)$ de entrada donde la aproximación es válida en diferentes regiones de subbanda. A continuación, a partir de la comparación de las fórmulas (6) y (8-10) de que una evolución de fase armónica a lo largo del eje de frecuencia de la señal $z(t)$ de entrada será extrapolada de manera correcta mediante la presente invención. Esto se sostiene en particular para un tren de pulsos puro. Para la calidad de audio de salida, ésta es una característica atractiva para señales de tren de pulsos, como aquellos producidos por las voces humanas y algunos instrumentos musicales.

Las Figuras 25, 26 y 27 ilustran el rendimiento de una implementación ejemplar de la transposición inventiva para una señal armónica en el caso $T=3$. La señal tiene una frecuencia fundamental de 282,35 Hz y su espectro de magnitud en el rango objetivo considerado de 10 a 15 kHz se representa 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 transponedor directo de tercer orden ($T=3$) se representa en la Fig. 26. Como se puede ver, cada tercer armónico se reproduce con gran fidelidad tal como se predijo por la teoría explicada en términos generales anteriormente, y el tono percibido será de 847 Hz, tres veces el original. La Fig. 27 muestra la salida de un transponedor que aplica productos de términos de cruce. Todos los armónicos se han recreado hasta las imperfecciones debido a los aspectos aproximados de la teoría. Para este caso, los lóbulos laterales son de aproximadamente 40 dB por debajo del nivel de señal y esto es más que suficiente para la regeneración del contenido de alta frecuencia que es perceptualmente indistinguible de la señal armónica original.

A continuación, se hace referencia a la Fig. 28 y la Fig. 29 que ilustran un codificador 2800 ejemplar y un decodificador 2900 ejemplar, respectivamente, para codificación de voz y audio unificada (USAC). La estructura general del codificador 2800 y del decodificador 2900 USAC se describen a continuación: Primero existe un pre/post procesamiento que consiste de una unidad funcional MPEG envolvente (MPEGs) para manipular procesamiento estéreo o multi canal y una unidad 2801 y 2901 SBR mejorada (eSBR), respectivamente, que manipula la representación paramétrica de las frecuencias de audio superiores en la señal de entrada y que puede hacer uso de los métodos de transposición armónica explicados en términos generales en el presente documento. Entonces existen dos ramificaciones, una consistente de una ruta de herramienta de Codificación de Audio Avanzada (AAC) modificada y la otra consistente de una codificación de predicción lineal (dominio LP o LPC) basada en la ruta, que a su vez presente bien una representación en el dominio de la frecuencia o una representación en el dominio del tiempo del LPC residual. Todos los espectros transmitidos para ambas, la AAC y la LPC, se pueden representar en el dominio MDCT que sigue a la cuantificación y codificación aritmética. La representación en el dominio del tiempo usa un esquema de codificación por excitación ACELP.

La unidad 2801 de Replicación de Banda Espectral mejorada (eSBR) del codificador 2800 puede comprender los sistemas de reconstrucción de alta frecuencia explicados en términos generales en el presente documento. En concreto, la unidad 2801 eSBR puede comprender un banco 301 de filtros de análisis para generar una pluralidad de señales de subbanda de análisis. Estas señales de subbanda de análisis se pueden transponer entonces en una unidad 302 no lineal para generar una pluralidad de señales de subbanda de síntesis, que se pueden introducir después a un banco 303 de filtros de síntesis para generar una componente de alta frecuencia. En la unidad 2801 eSBR, en el lado de codificación, se puede determinar un conjunto de información sobre cómo generar una componente de alta frecuencia a partir de la componente de baja frecuencia que mejor coincide con la componente de alta frecuencia de la señal original, Este conjunto de información puede comprender información sobre las características de la señal, tal como la frecuencia Ω fundamental predominante, sobre la envolvente espectral de la componente de alta frecuencia, y puede comprender información sobre cómo combinar mejor las señales de subbanda de análisis, esto es información tal como el conjunto limitado de parejas (p_1, p_2) de desplazamiento de índice. Los datos codificados relacionados con este primer conjunto de información se combinan con la otra información codificada en un multiplexor de flujo de bits y se envían como un flujo de audio codificado al decodificador 2900 correspondiente.

El decodificador 2900 mostrado en la Fig. 29 comprende también una unidad de Replicación de Ancho de Banda Espectral mejorada (eSBR) 2901. Esta unidad 2901 eSBR recibe el flujo de bits de audio codificado o la señal codificada del codificador 2800 y usa los métodos explicados en términos generales en el presente documento para generar una componente de alta frecuencia de la señal, que se combina con la componente de baja frecuencia decodificada para producir una señal decodificada. La unidad 2901 eSBR puede comprender las diferentes componentes explicadas en términos generales en el presente documento. En concreto, puede comprender un banco 301 de filtros de análisis, una unidad 302 de procesamiento no lineal y un banco 303 de filtros de síntesis. La unidad 2901 eSBR puede usar información en la componente de alta frecuencia proporcionada por el codificador 2800 para realizar la reconstrucción de alta frecuencia. Dicha información puede ser una frecuencia Ω fundamental de la señal, la envolvente espectral de la componente de alta frecuencia original y/o la información sobre las subbandas de análisis que se han de usar para generar las señales de subbanda de síntesis y por último la componente de alta frecuencia de la señal decodificada.

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

- una herramienta de demultiplexor de carga útil de flujo de bits, que separa la carga útil del flujo de bits en partes para cada herramienta, y proporciona cada una de las herramientas con la información de la carga útil del flujo de bits relacionada con esa herramienta;
- una herramienta de decodificación sin ruido de factor de escala, que toma la información del demultiplexor de la carga útil del flujo de bits, analiza esa información, y decodifica los factores de escala codificados en Huffman y DPCM;
- una herramienta de decodificación sin ruido espectral, que toma la información del demultiplexor de la carga útil del flujo de bits, analiza esa información, decodifica los datos codificados de manera aritmética, y

reconstruye los espectros cuantificados;

- una herramienta cuantificadora inversa, que toma los valores cuantificados para los espectros, y convierte los valores enteros a los espectros reconstruidos, no escalados; este cuantificador es preferiblemente un cuantificador de compresión, cuyo factor de compresión depende del modo de codificación de núcleo elegido;
- 5
- una herramienta de llenado de ruido, que se usa para rellenar huecos espectrales en los espectros decodificados, que se producen cuando los valores espectrales se cuantifican a cero por ejemplo debido a una restricción fuerte sobre la demanda de bits en el codificador;
 - una herramienta de reescalado, que convierte la representación entera de los factores de escala a los valores reales, y multiplica los espectros cuantificados de manera inversa no escalados por los factores de escala;
- 10
- una herramienta M/S, tal como se describe en la ISO/IEC 14496-3;
 - una herramienta de conformado de ruido temporal (TNS), tal como se describe en la ISO/IEC 14496-3;
 - una herramienta de conmutación de banco / bloque de filtros, que aplica la inversa de la correspondencia de frecuencia que fue llevada a cabo en el codificador; se usa preferiblemente una transformada de coseno discreta modificada (IMDCT) para la herramienta de banco de filtros;
- 15
- una herramienta de conmutación de banco / bloque de filtros distorsionados en tiempo, que reemplaza la herramienta de conmutación de banco / bloque de filtros normal cuando se habilita el modo de distorsión; el banco de filtros preferiblemente es el mismo (IMDCT) que para el banco de filtros normal, de manera adicional las muestras en el dominio de tiempo con ventana se hacen corresponder a partir del dominio de tiempo distorsionado al dominio de tiempo lineal mediante el muestreo variante en el tiempo;
- 20
- una herramienta de Envoltente MPEG (MPEGS), que produce múltiples señales a partir de una o más señales de entrada aplicando un procedimiento de mezcla ascendente sofisticado a la señal o señales de entrada controladas por los parámetros espaciales apropiados; en el contexto USAC, MPEGS se usa preferiblemente para codificar una señal multicanal, transmitiendo información lateral paramétrica junto a la señal de mezcla descendente transmitida;
- 25
- una herramienta Clasificadora de Señal, que analiza la señal de entrada original y genera a partir de ésta la información de control que desencadena la selección de los diferentes modos de codificación; el análisis de la señal de entrada es normalmente una implementación dependiente e intentará elegir el modo de codificación de núcleo óptimo para una trama de señal de entrada dada; la salida del clasificador de señal se puede usar de manera opcional para influenciar el comportamiento de las otras herramientas, por
- 30
- ejemplo los bancos de filtros distorsionados en tiempo de envoltente MPEG, de SBR mejorada.
 - una herramienta de filtro LPC, que produce una señal en el dominio del tiempo a partir de una señal en el dominio de excitación mediante el filtrado de la señal de excitación reconstruida a través del filtro de síntesis de predicción lineal; y
- 35
- una herramienta ACELP, que proporciona una manera de representar de manera eficiente una señal de excitación en el dominio del tiempo combinando un predictor a largo plazo (palabra clave adaptativa) con una secuencia como un pulso (palabra clave de innovación).

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

En la Fig. 30 la componente 3013 de baja frecuencia es alimentada a un banco de filtros QMF, para generar las bandas de frecuencia QMF. Estas bandas de frecuencia QMF no se deben confundir con las subbandas de análisis explicadas en términos generales en este documento. Las bandas de frecuencia QMF se usan con el propósito de manipular y combinar la componente de baja y alta frecuencia de la señal en el dominio de la frecuencia, en lugar de en el dominio de tiempo. La componente 3014 de baja frecuencia es alimentada dentro de la unidad 3004 de transposición que corresponde a los sistemas para la reconstrucción de la alta frecuencia explicados en términos generales en el presente documento. La unidad 3004 de transposición puede recibir también información 3011 adicional, tal como la frecuencia Ω fundamental de la señal codificada y/o las posibles parejas (p_1, p_2) de desplazamiento de índice para la selección de subbanda. La unidad 3004 de transposición genera una componente 3012 de alta frecuencia, también conocida como banda superior, de la señal, que se transforma al dominio de la frecuencia mediante un banco 3003 de filtros QMF. Tanto, la componente de baja frecuencia transformada QMF como la componente de alta frecuencia transformada QMF se alimentan en la unidad 3005 de manipulación y

combinación. Esta unidad 3005 puede realizar 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 se vuelve a transformar al dominio del tiempo mediante un banco 3001 de filtros QMF.

5 Normalmente los bancos de filtros QMF comprende 64 bandas de frecuencias QMF. Se debería observar, sin embargo, que puede ser beneficioso muestrear de manera descendente la componente 3013 de baja frecuencia, de manera tal que el banco 3002 de filtros QMF sólo requiera 32 bandas de frecuencia QMF. En dichos casos, la componente 3013 de baja frecuencia 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 3012 de alta frecuencia tiene un ancho de banda de $f_s/2$.

10 El método y sistema descrito en el presente documento se puede implementar como software, firmware y/o hardware. Ciertos componentes se pueden implementar por ejemplo como software que se ejecuta en un procesador de señal digital o un microprocesador. Se pueden implementar otros componentes por ejemplo como hardware y o como circuitos integrados específicos de aplicación. Las señales encontradas en los métodos y sistemas descritos se pueden almacenar en medios tales como redes de radio, redes de satélite, redes inalámbricas o redes por cable, por ejemplo, internet. Los dispositivos típicos que hacen uso del método y sistema descrito en el
15 presente documento son cajas descodificadoras u otro equipo de las instalaciones del cliente que decodifique señales de audio. En el lado de la codificación, se puede usar el método y sistema en estaciones de difusión, por ejemplo, en sistemas de cabecera de video.

20 El presente documento explica en términos generales el método y sistema para realizar la reconstrucción en alta frecuencia de una señal basada en la componente en baja frecuencia de esa señal. Usando combinaciones de subbandas desde la componente de baja frecuencia, el método y sistema permiten la reconstrucción de las frecuencias y las bandas de frecuencia que no se pueden generar mediante los métodos de transposición conocidos en la técnica. Además, el método y sistema HTR descrito permite el uso de frecuencia de corte bajas y/o la generación de grandes bandas de frecuencia a partir de bandas de baja frecuencia estrechas.

REIVINDICACIONES

1. Un sistema para decodificar una señal de audio, comprendiendo el sistema:
 - un decodificador (101) de núcleo para decodificar una componente de baja frecuencia de la señal de audio;
 - 5 un banco (301) de filtros de análisis para proporcionar una pluralidad de señales de subbanda de análisis de la componente de baja frecuencia de la señal de audio;
 - una unidad de recepción de selección de subbanda para recibir la 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 a partir de la pluralidad de señales de subbanda de análisis, a partir de la cual se genera la señal (803) de subbanda de síntesis;
 - 10 una unidad (302) de procesamiento no lineal para generar la señal de subbanda de síntesis con una frecuencia de síntesis, una magnitud y una fase mediante:
 - determinación de la magnitud de la señal de subbanda de síntesis desde un valor medio generalizado de las magnitudes de la primera y la segunda señal de subbanda de análisis, y
 - 15 determinación de la fase de la señal de subbanda de síntesis desde una suma ponderada de las fases de la primera y segunda señales de subbanda de análisis, en donde un primer peso aplicado a la fase de la primera señal de subbanda de análisis corresponde a un primer factor de transposición T-r, y en donde un segundo peso aplicado a la fase de la segunda señal de subbanda de análisis corresponde a un segundo factor de transposición r, en donde T y r son números enteros positivos, $T > 1$ y $1 \leq r < T$, y
 - 20 un banco (303) de filtros de síntesis para generar una componente de alta frecuencia de la señal de audio a partir de la señal de subbanda de síntesis.
2. El sistema según la reivindicación 1, en donde
 - el banco (301) de filtros de análisis tiene N subbandas de análisis con un espaciado de subbanda esencialmente constante de $\Delta\omega$;
 - una subbanda de análisis se asocia con un índice n de subbanda de análisis n, con $n \in \{1, \dots, N\}$;
 - 25 el banco (303) de filtros de síntesis tiene una subbanda de síntesis;
 - la subbanda de síntesis se asocia con un índice n de subbanda de síntesis; y
 - la subbanda de síntesis y la subbanda de análisis con índice n comprenden cada una los rangos de frecuencia que se relacionan los unos con los otros a través de T.
3. El sistema según la reivindicación 2, que comprende además:
 - 30 una ventana (2001) de análisis, que aísla un intervalo de tiempo predefinido de la componente de baja frecuencia alrededor de una instancia k de tiempo predefinida; y
 - una ventana (2201) de síntesis que aísla un intervalo de tiempo predefinido de la componente de alta frecuencia alrededor de una instancia k de tiempo predefinida.
4. El sistema según la reivindicación 3, en donde la ventana (2201) de síntesis es una versión escalada en el tiempo de la ventana (2001) de análisis.
- 35 5. El sistema según la reivindicación 1, que comprende además:
 - un muestreador ascendente (104) para realizar un muestreo ascendente de la componente de baja frecuencia para producir una componente de baja frecuencia sin muestrear;
 - un ajustador (103) de envolvente para dar forma a la componente de alta frecuencia; y
 - 40 una unidad de suma de componentes para determinar una señal de audio decodificada como la suma de la componente de baja frecuencia muestreada de manera ascendente y la componente de alta frecuencia ajustada.
6. El sistema según la reivindicación 5, que comprende además una unidad de recepción de envolvente para recibir la información relacionada con la envolvente de la componente de alta frecuencia de la señal de audio.
7. El sistema según la reivindicación 6, que comprende además:
 - 45 una unidad de entrada para recibir la señal de audio, que comprende la componente de baja frecuencia; y

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

8. El sistema según la reivindicación 1, en donde el banco (301) de filtros de análisis muestra un espaciado de frecuencia que se asocia con la frecuencia fundamental Ω de la señal de audio.

5 9. Un método para decodificar una señal de audio, comprendiendo el método:

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

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

10 recibir la 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 mediante:

determinación de la magnitud de la señal de subbanda de síntesis desde un valor medio generalizado de las magnitudes de la primera y una segunda señal de subbanda de análisis, y

15 determinación de la fase de la señal de subbanda de síntesis desde una suma ponderada de las fases de la primera y segunda señal de subbanda de análisis, en donde un primer peso aplicado a la fase de la primera señal de subbanda de análisis corresponde a un primer factor de transposición $T-r$, y en donde un segundo peso aplicado a la fase de la segunda señal de subbanda de análisis corresponde a un segundo factor de transposición r , en donde T y r son números enteros positivos, $T > 1$ y $1 \leq r < T$, y

20 generar (303) una componente de alta frecuencia de la señal de audio a partir de la señal de subbanda de síntesis.

10. Un medio de almacenamiento que comprende un programa de software adaptado para la ejecución en un procesador y para llevar a cabo las etapas de método de la reivindicación 9 cuando se llevan a cabo en un dispositivo de computación.

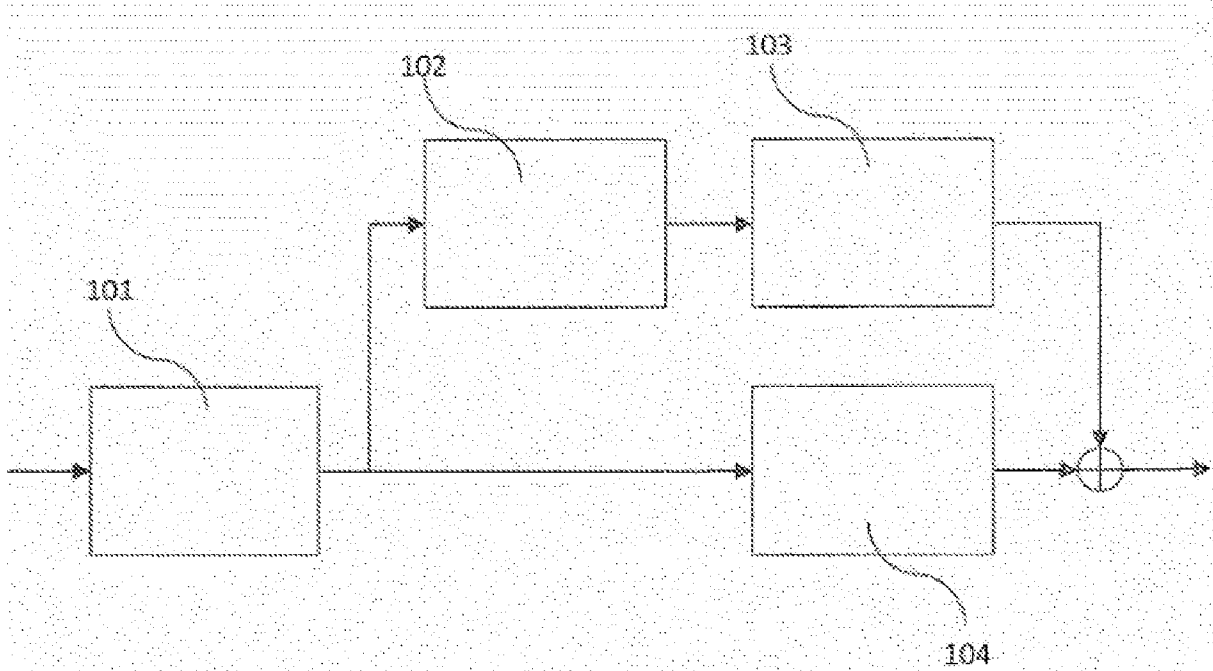


Fig. 1

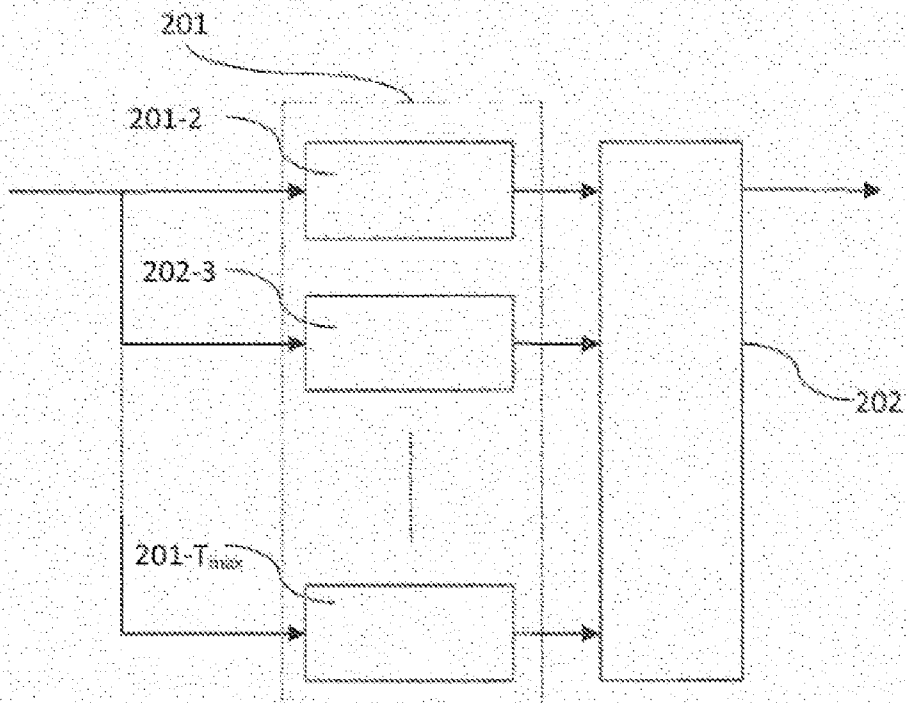


Fig. 2

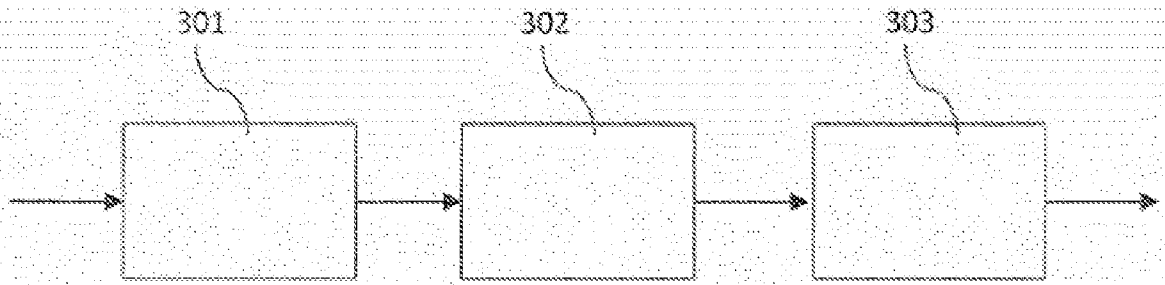


Fig. 3

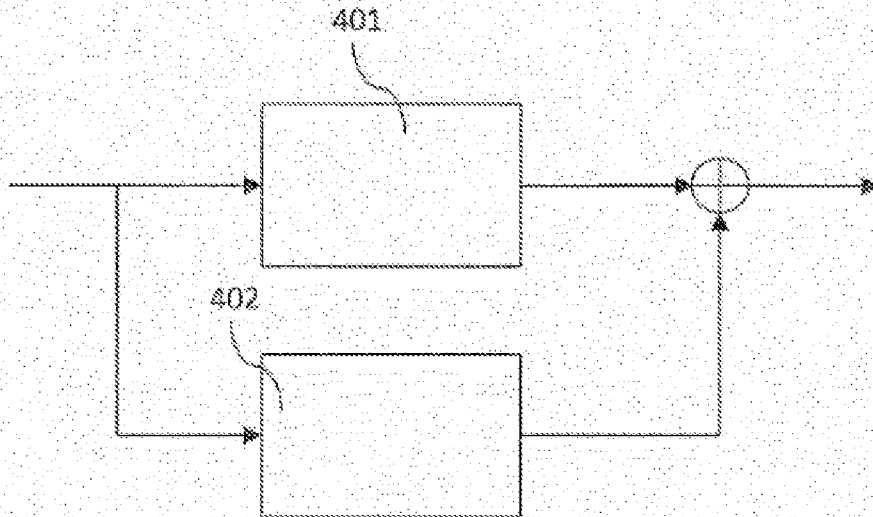


Fig. 4

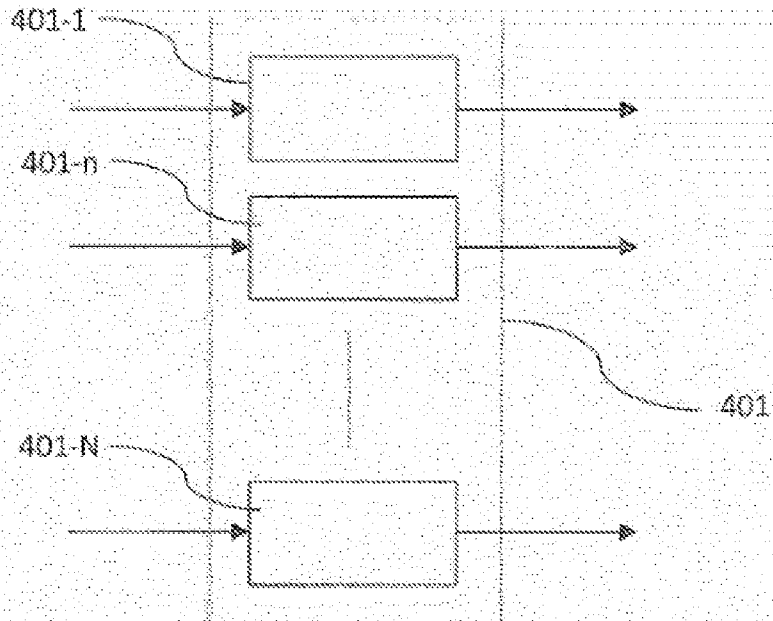


Fig. 5

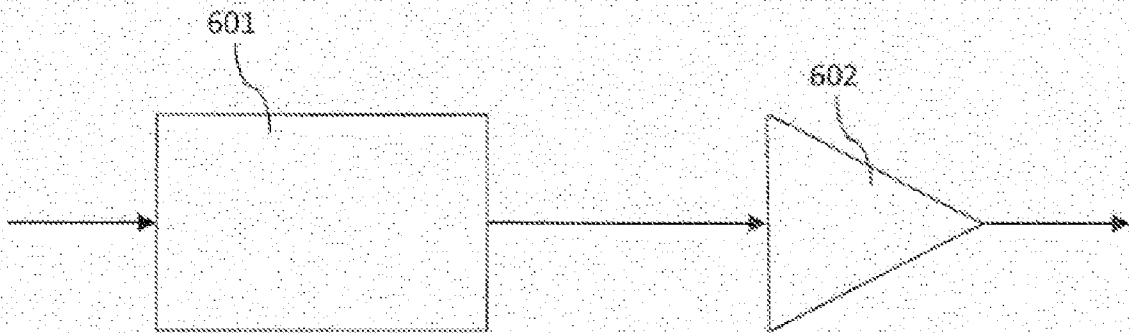


Fig. 6

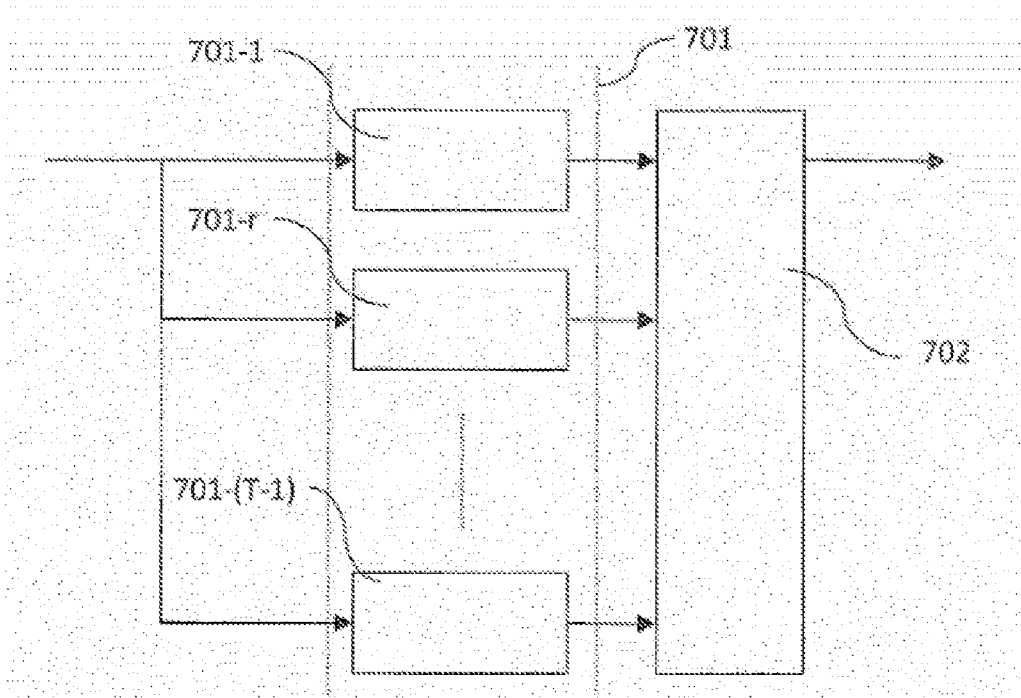


Fig. 7

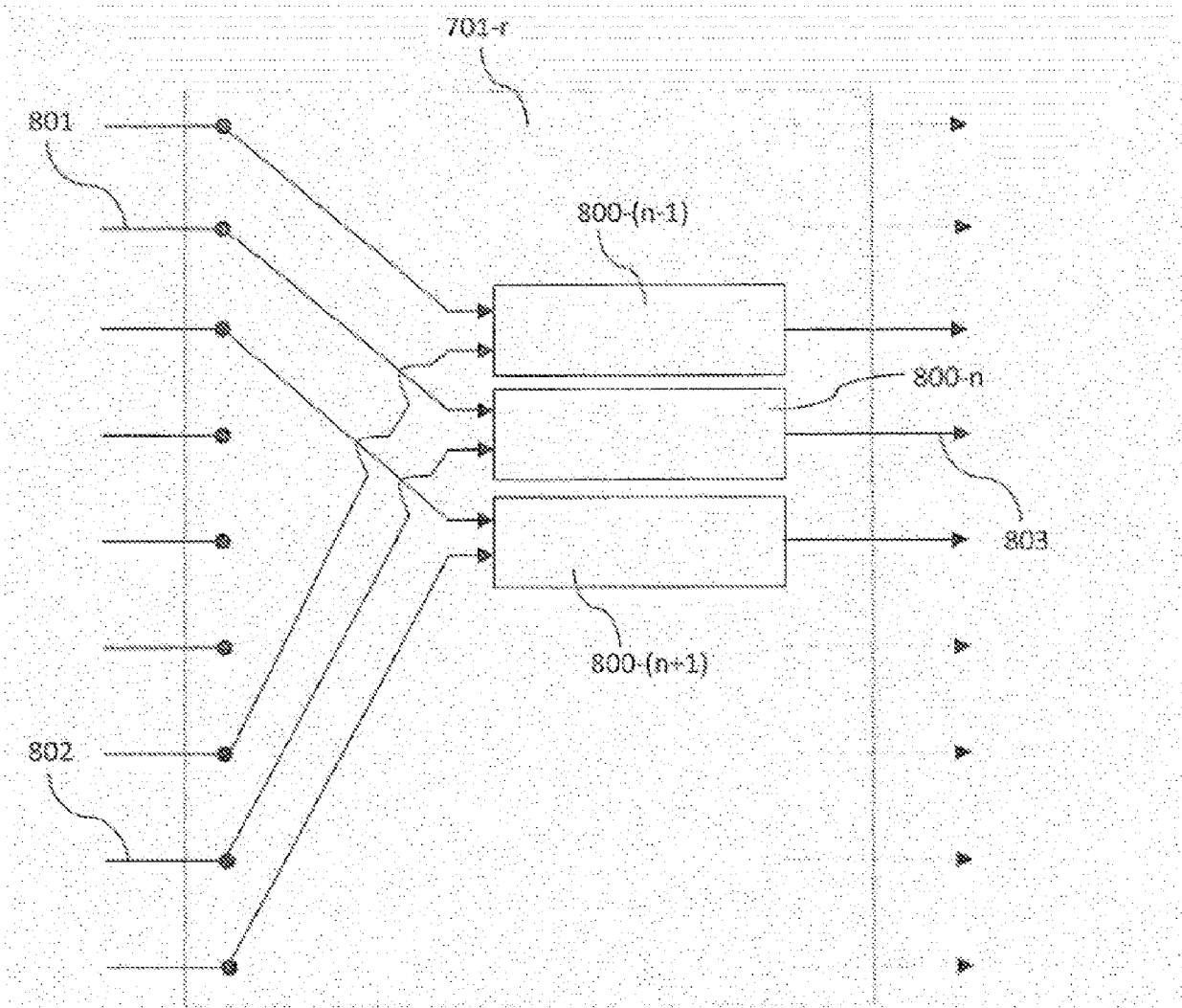


Fig. 8

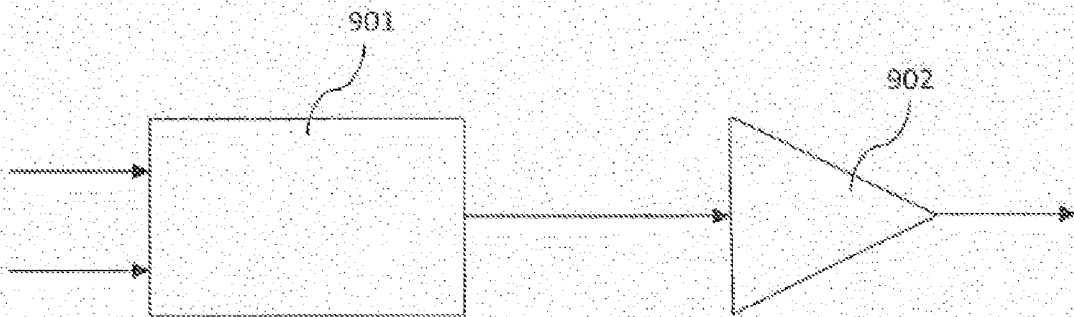


Fig. 9

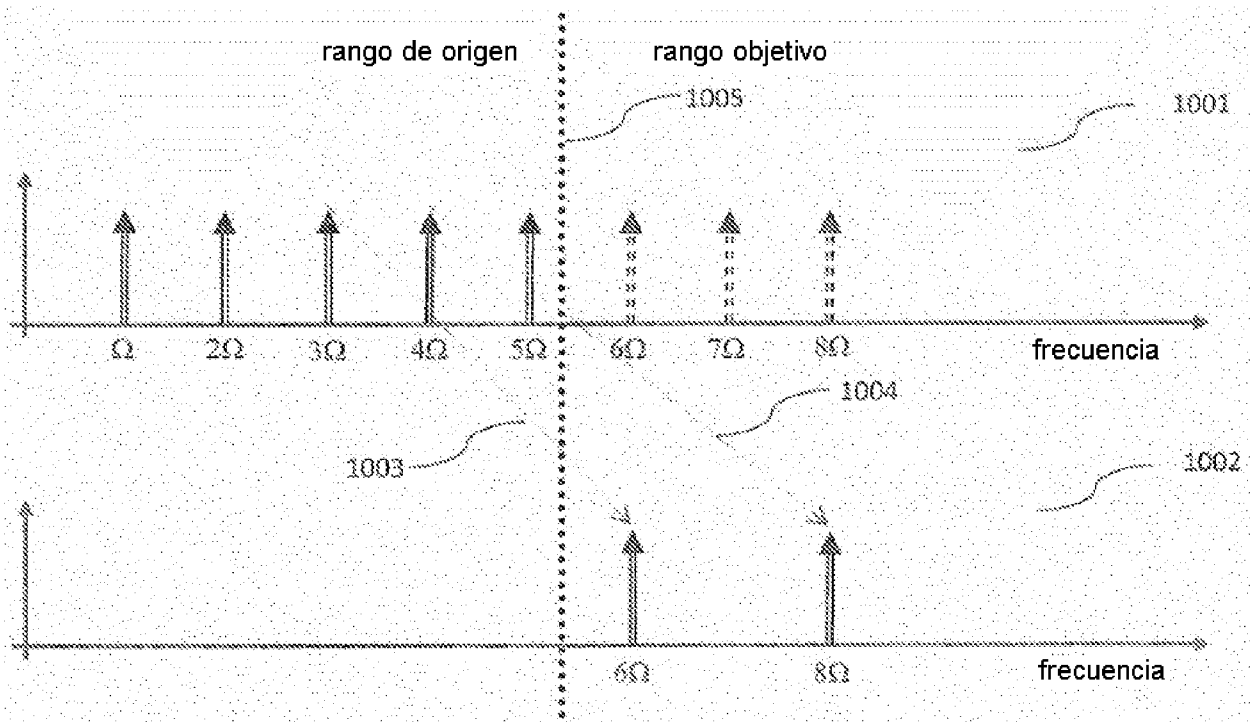


Fig. 10

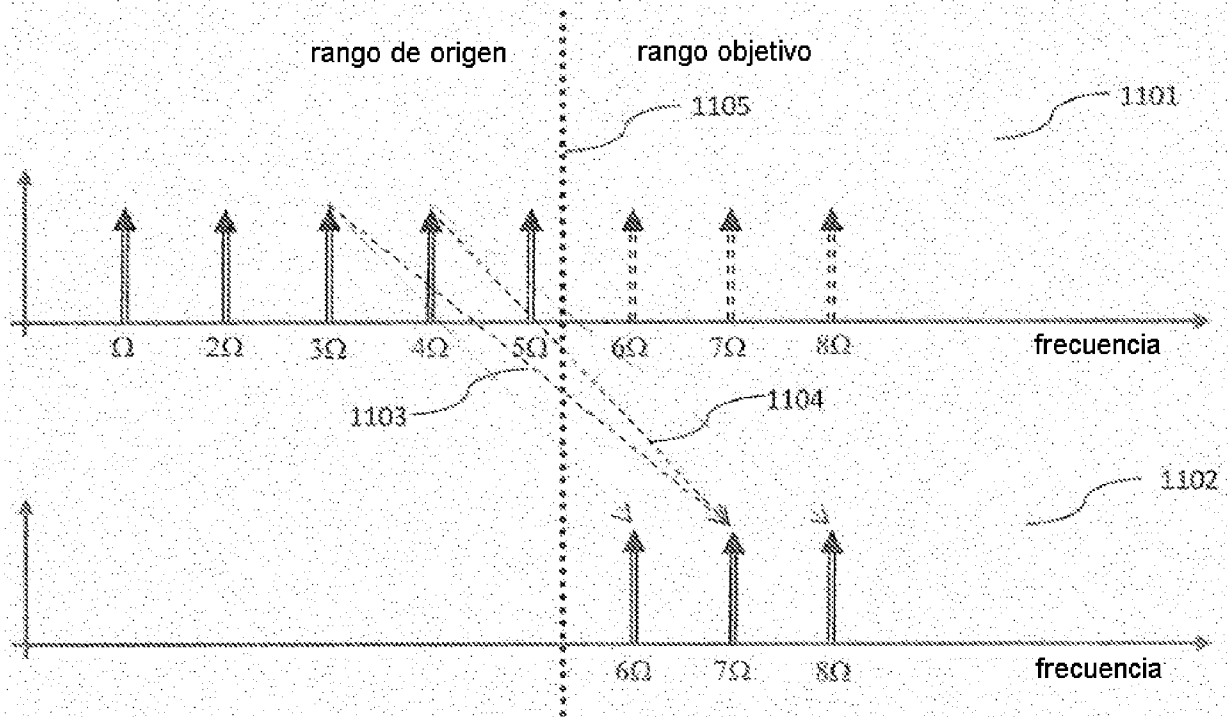


Fig. 11

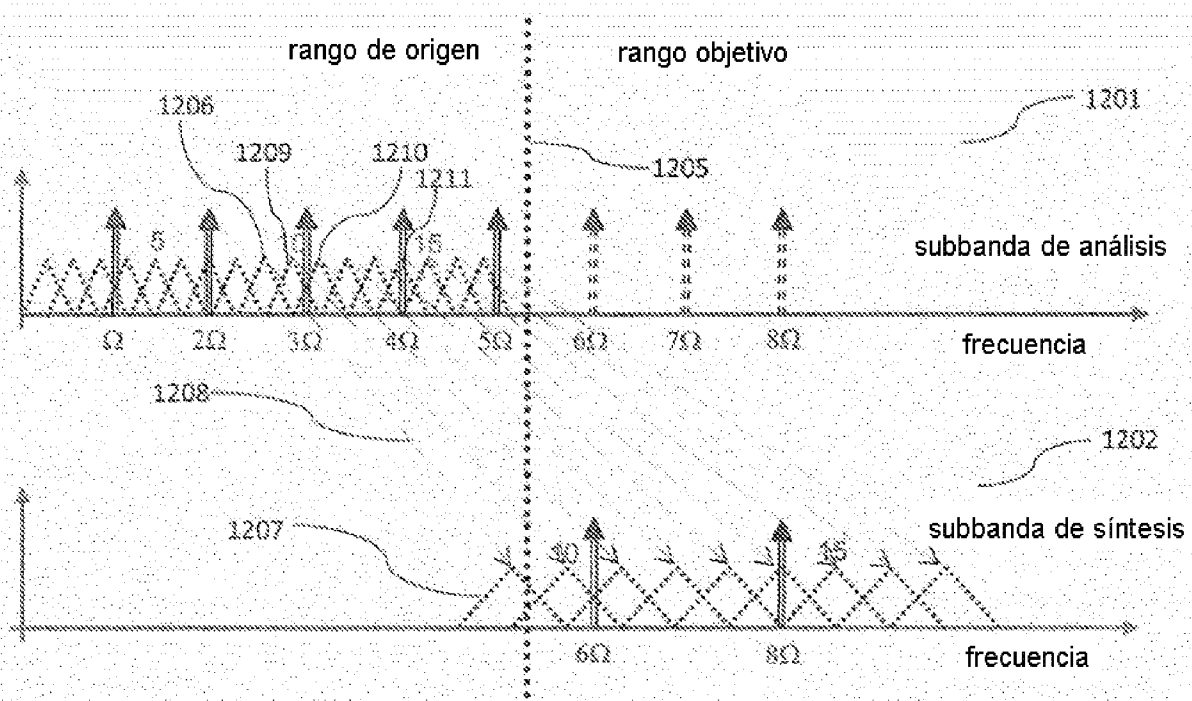


Fig. 12

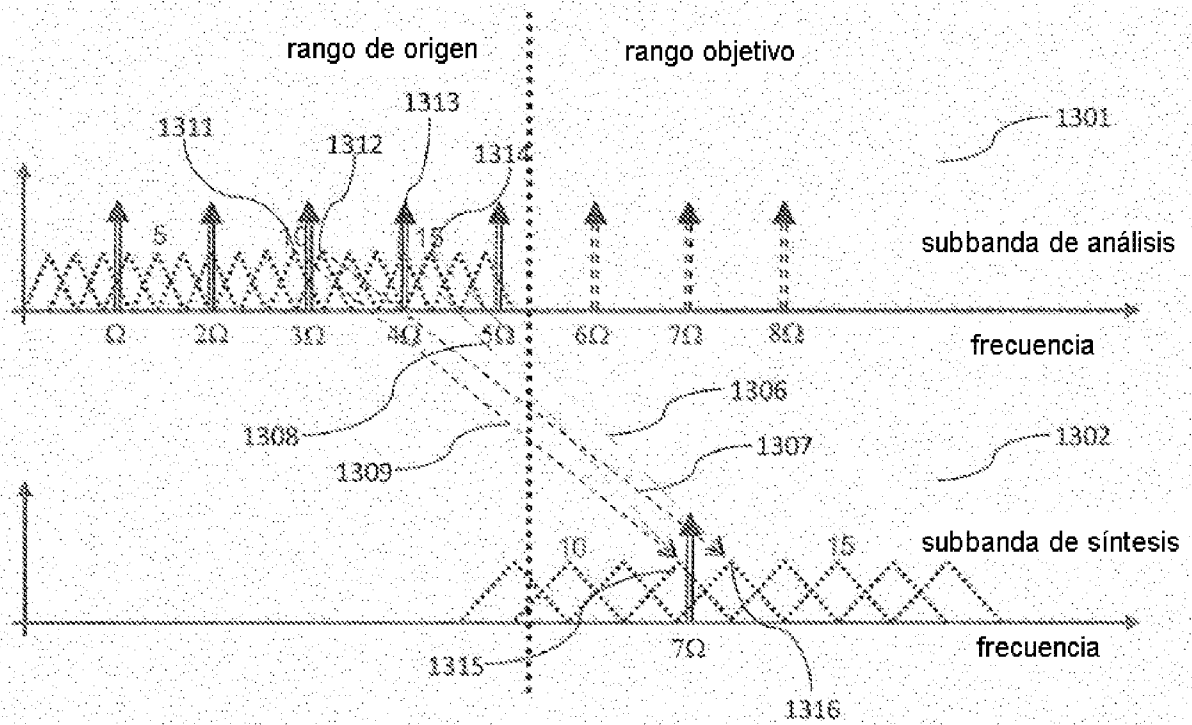


Fig. 13

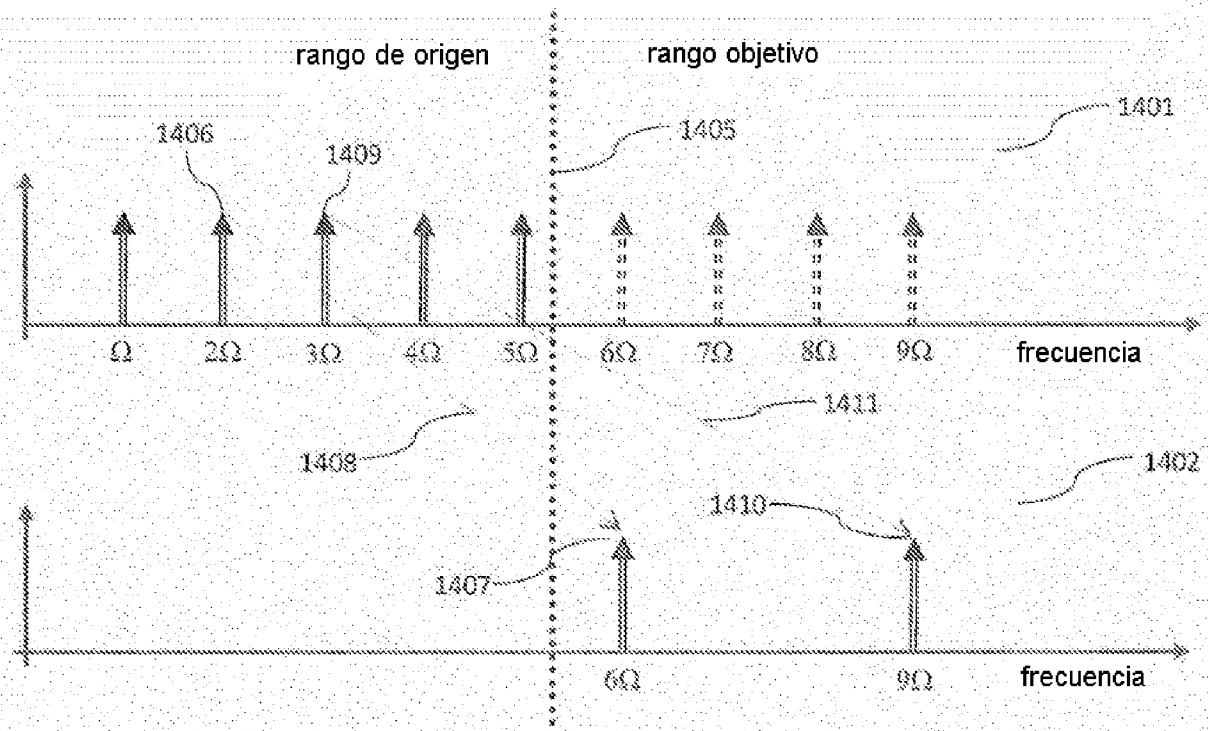


Fig. 14

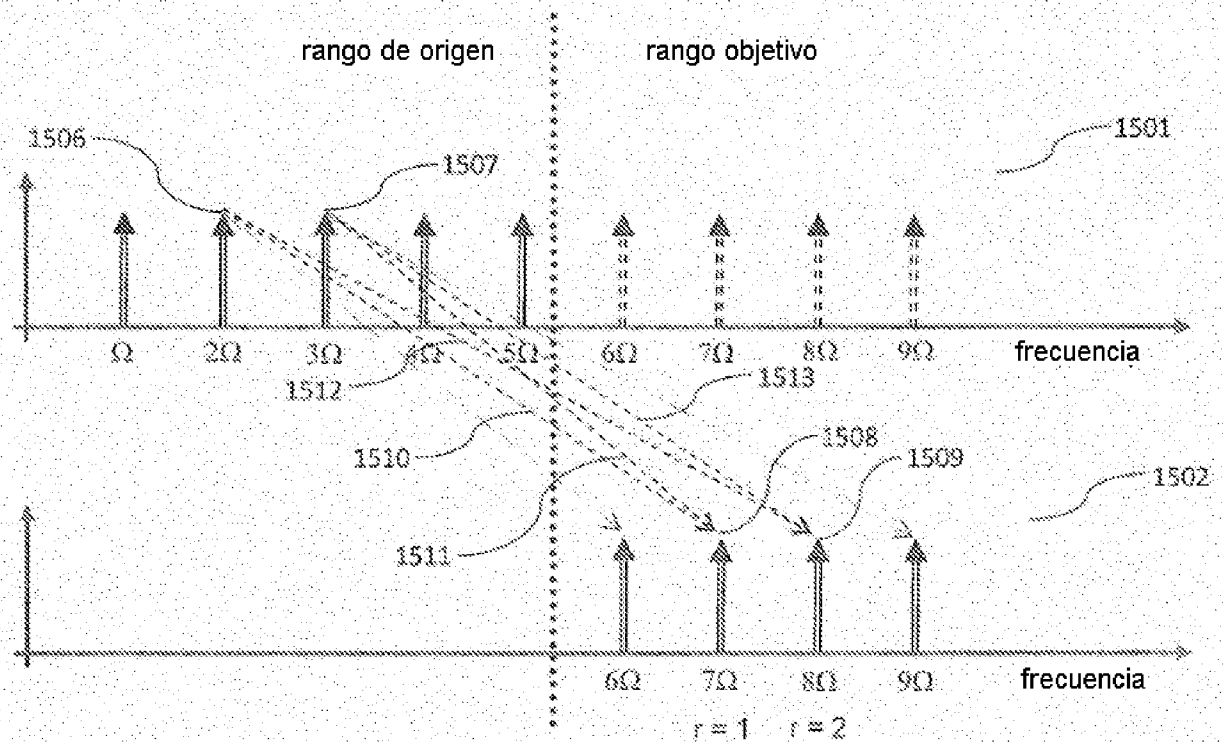


Fig. 15

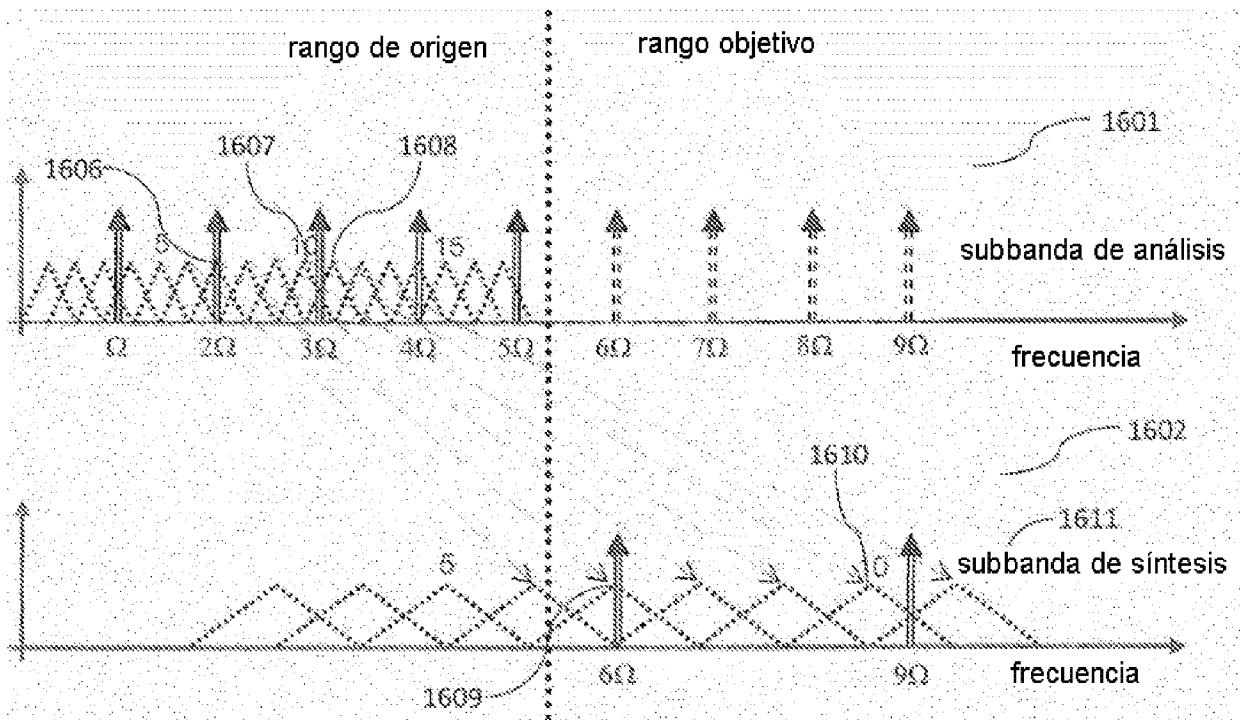


Fig. 16

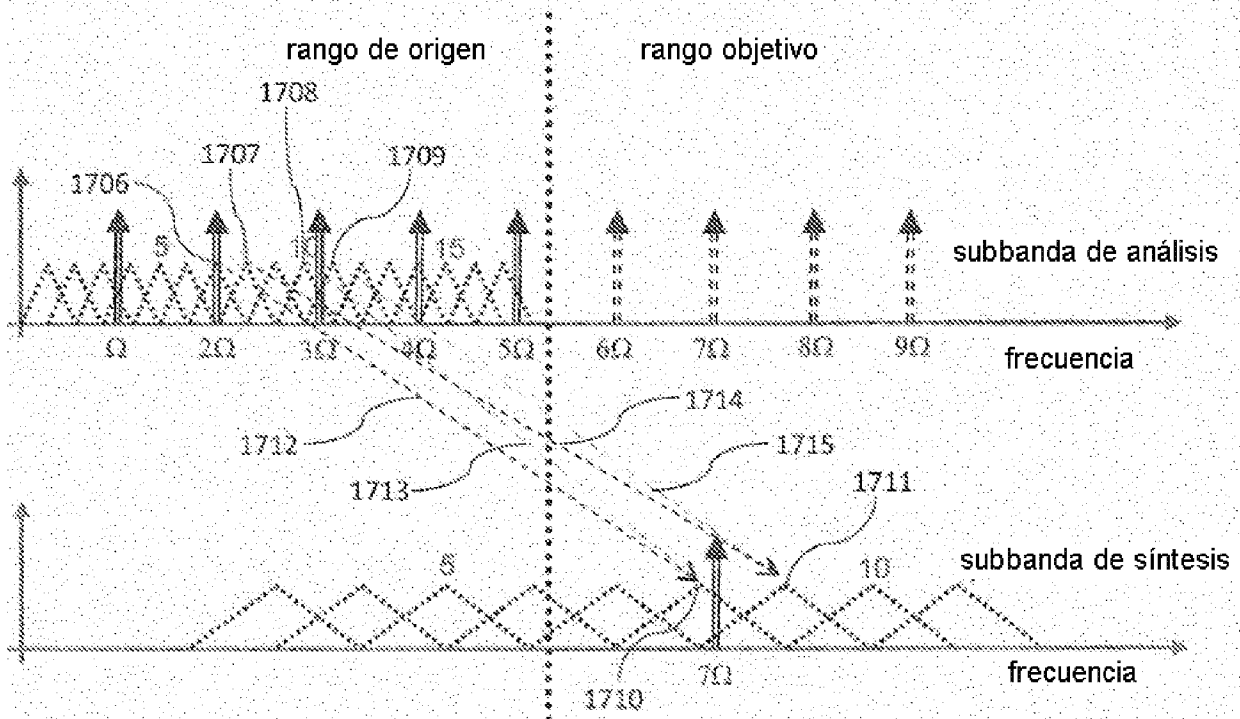


Fig. 17

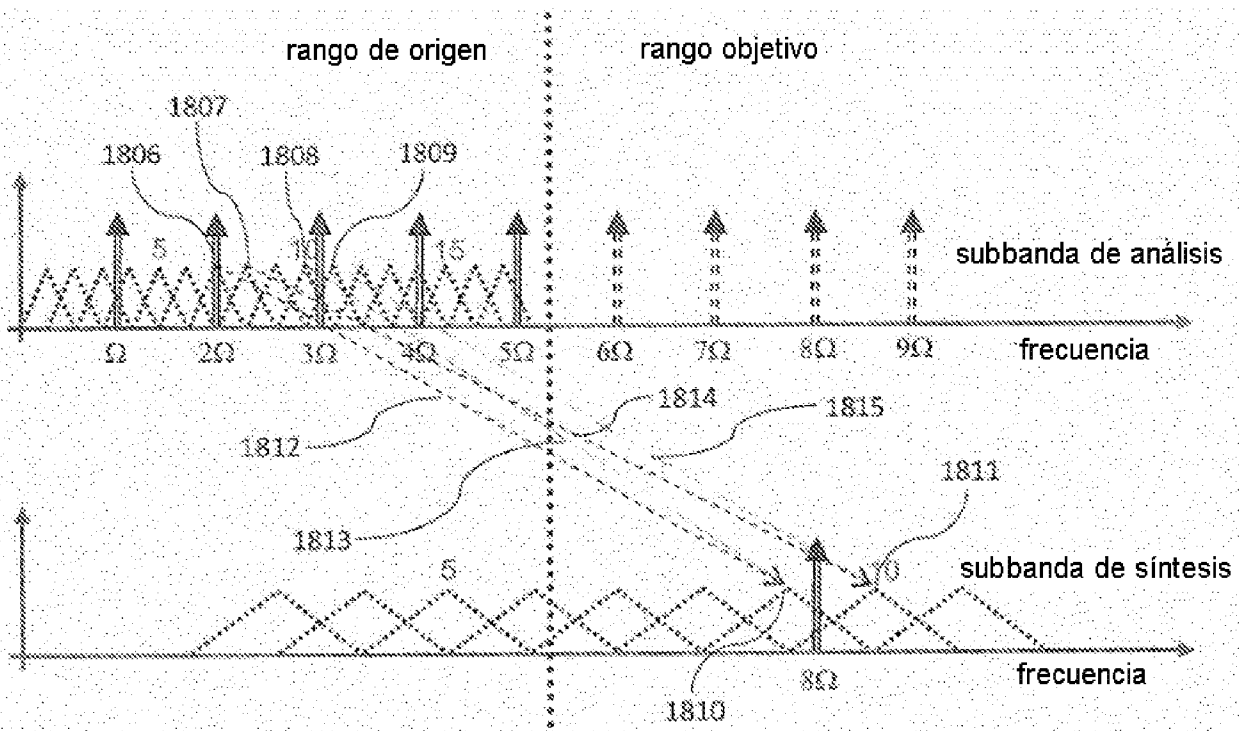


Fig. 18

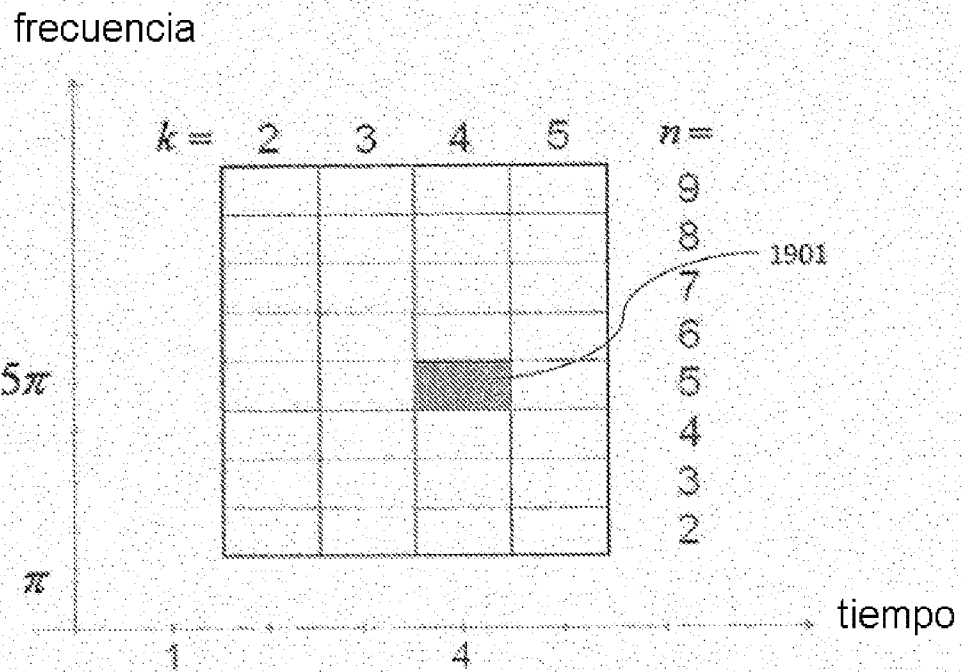


Fig. 19

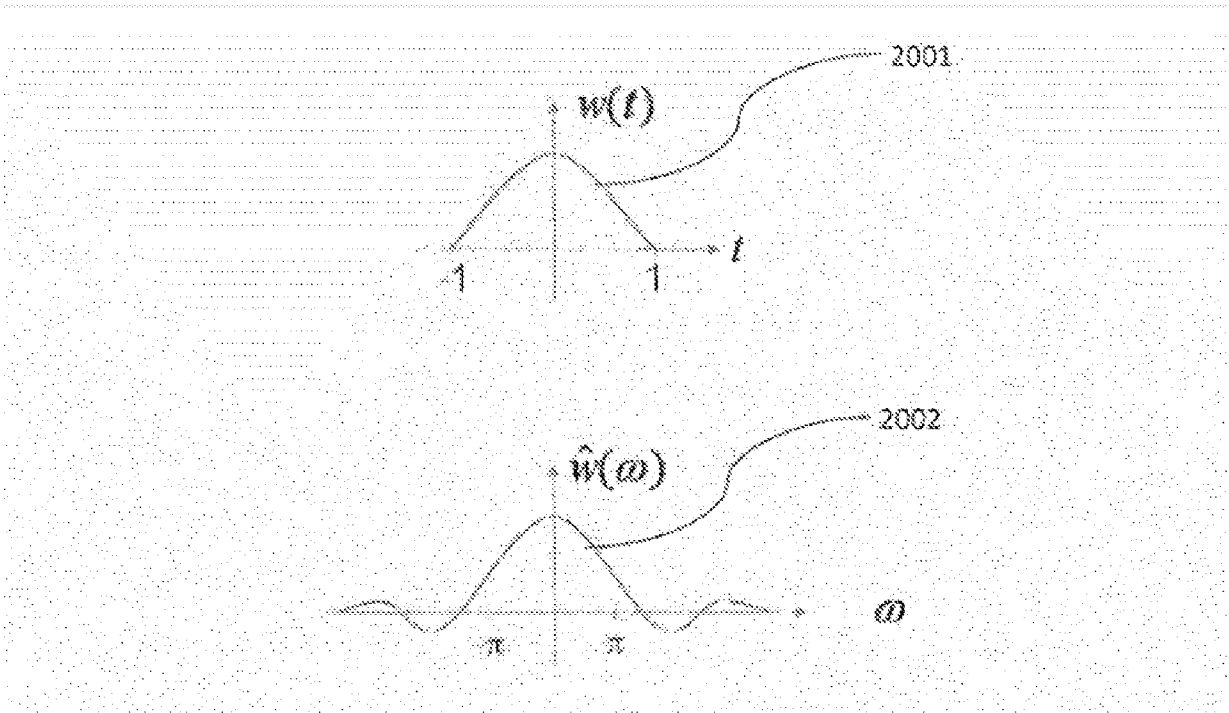


Fig. 20

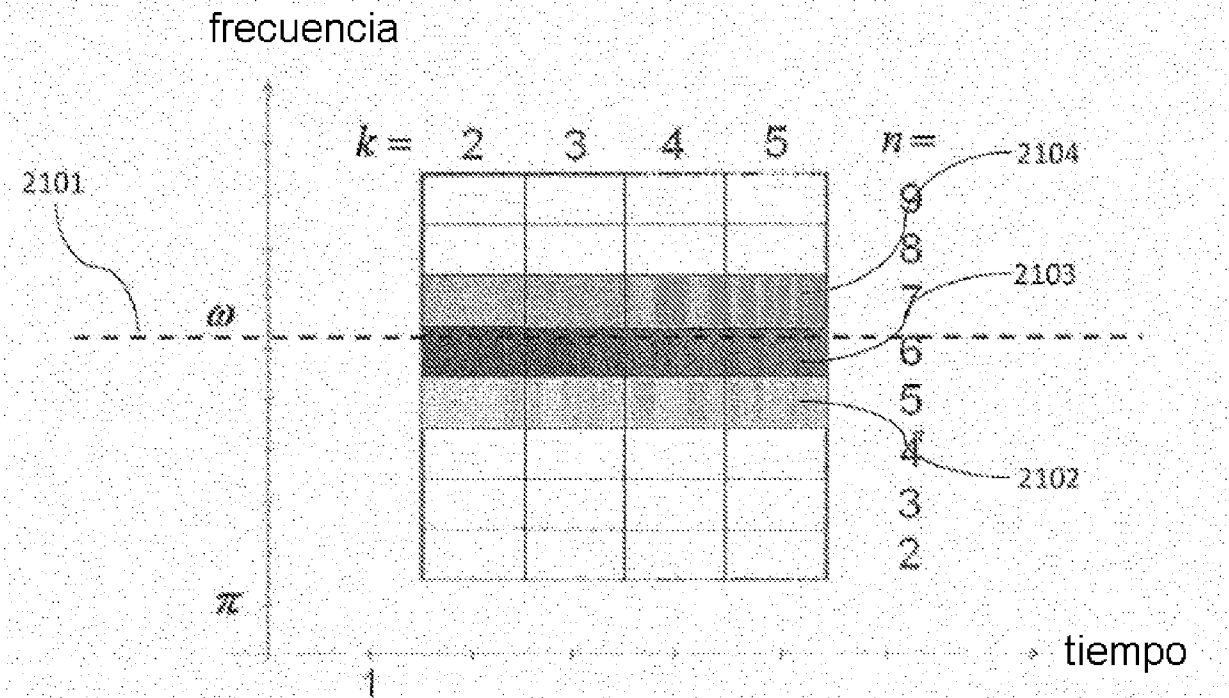


Fig. 21

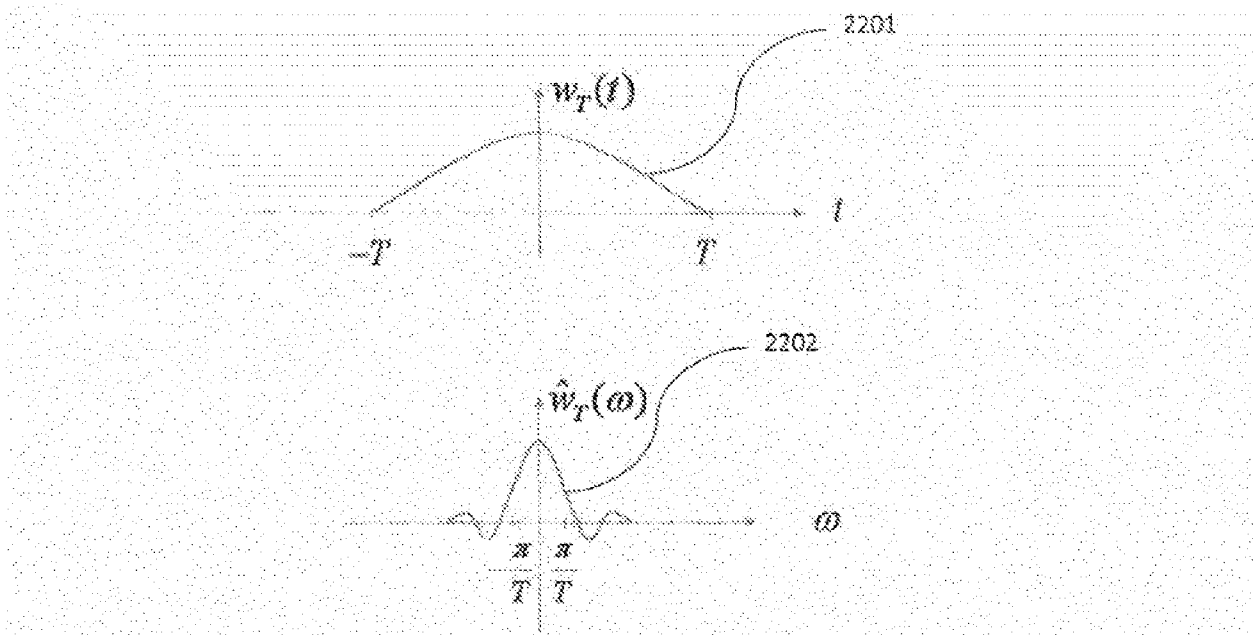


Fig. 22

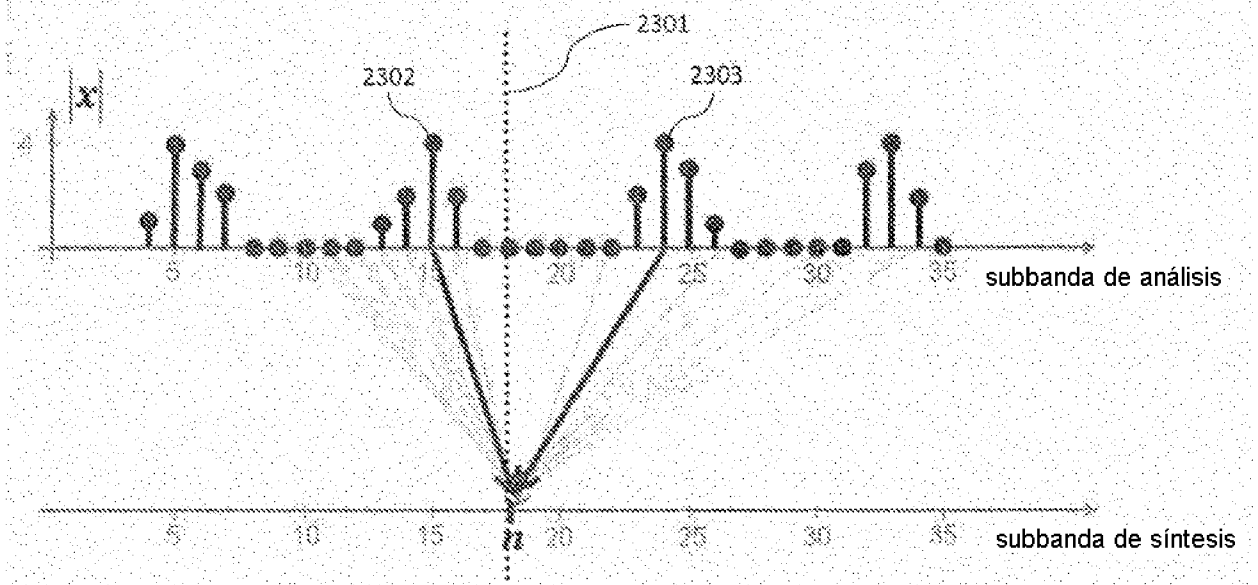


Fig. 23

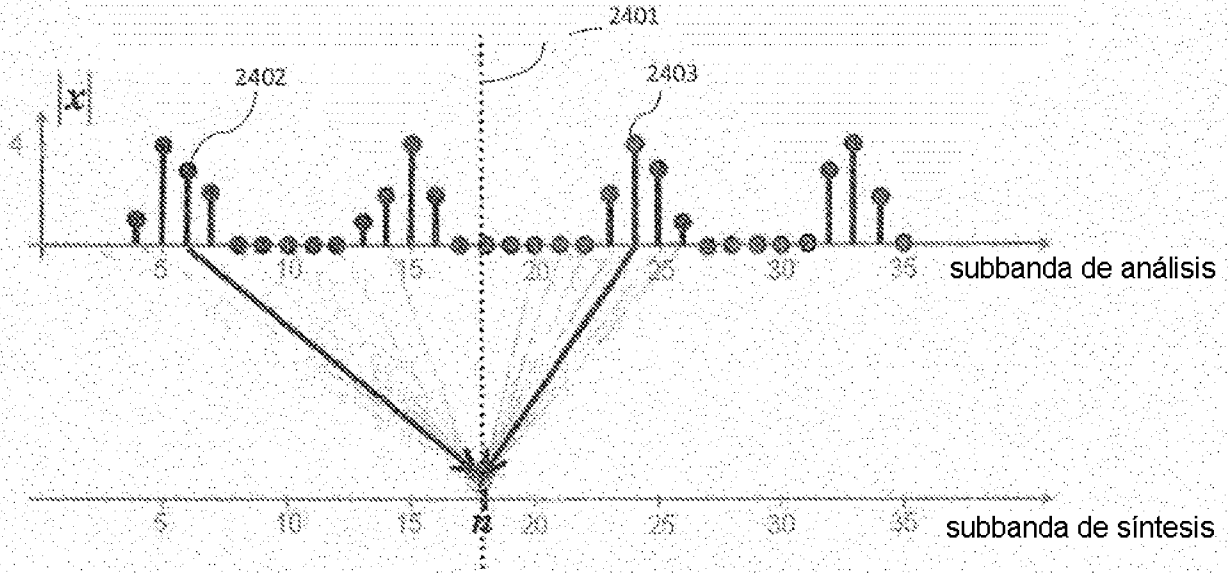


Fig. 24

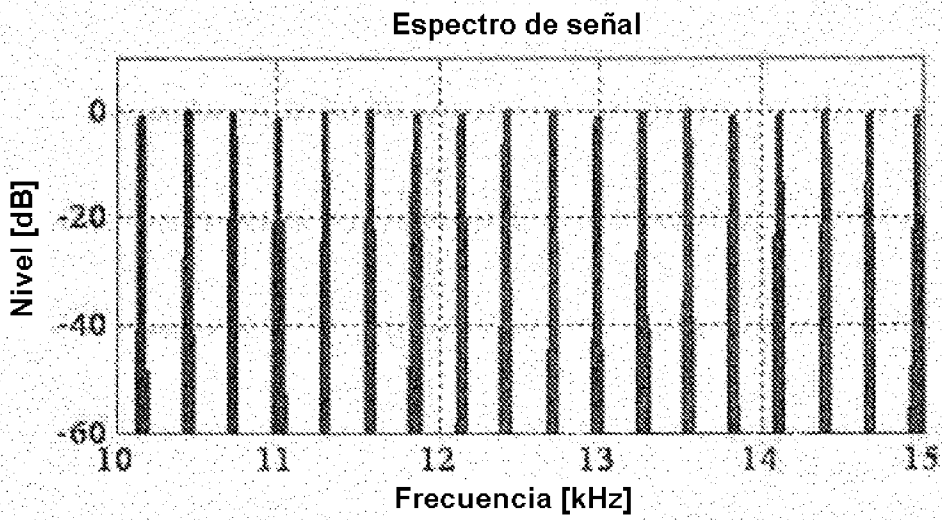


Fig. 25

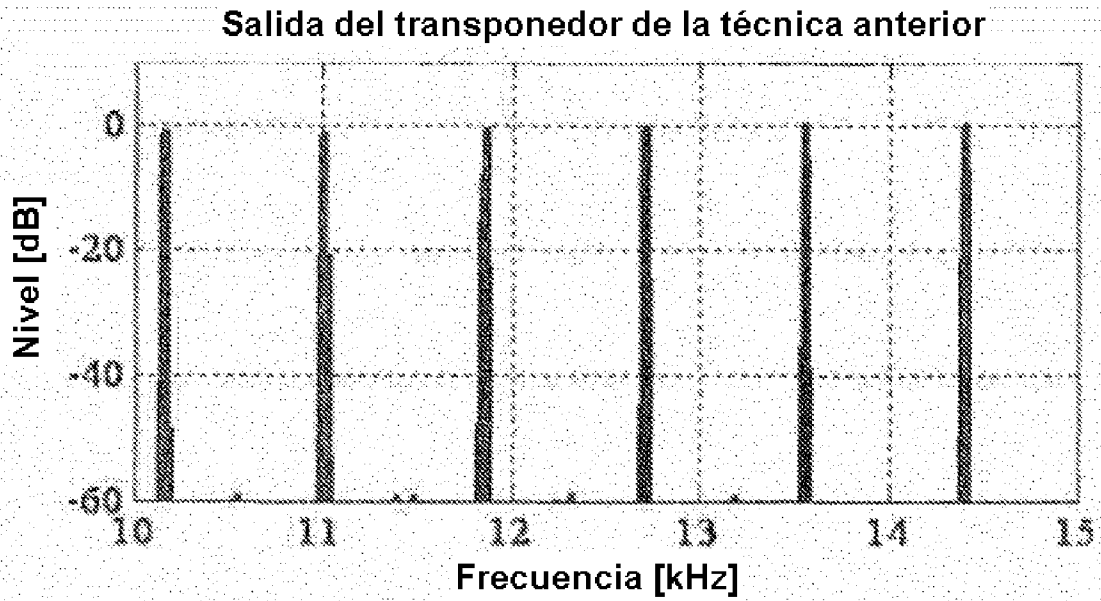


Fig. 26

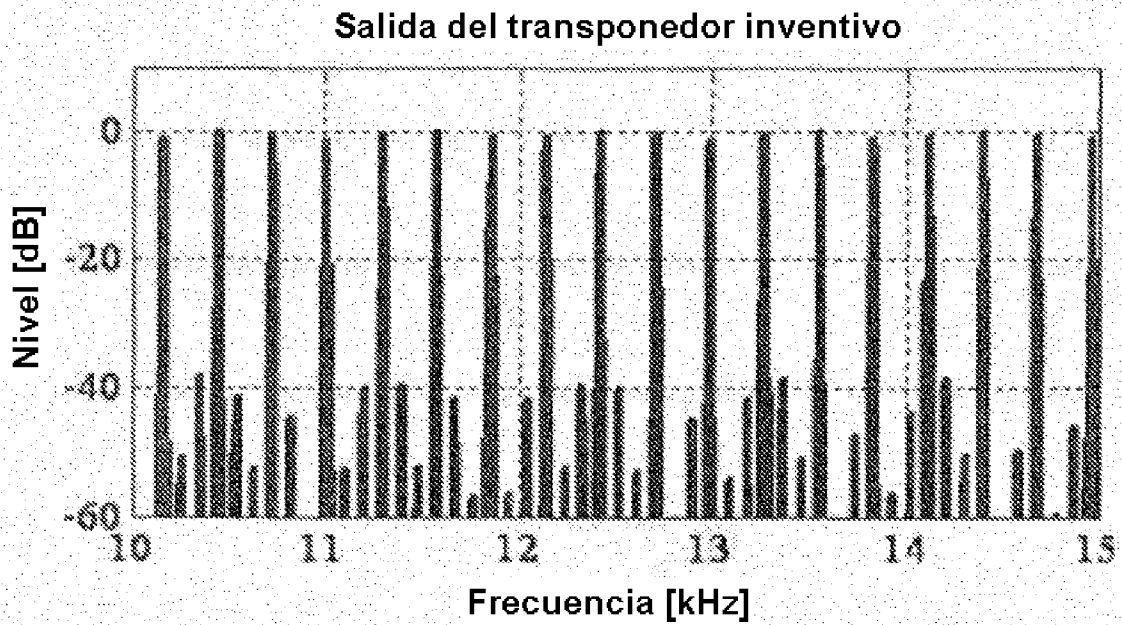


Fig. 27

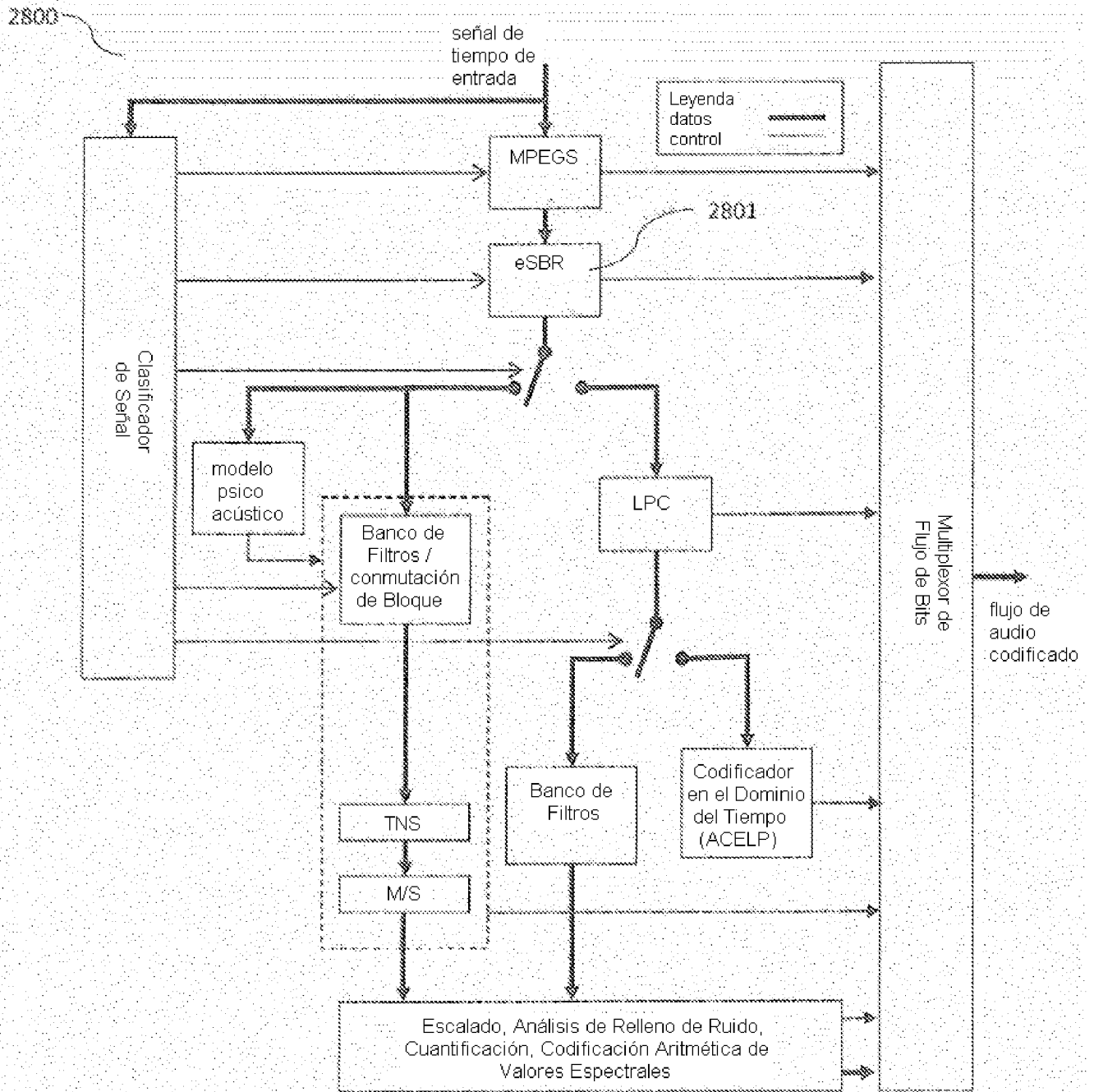


Fig. 28

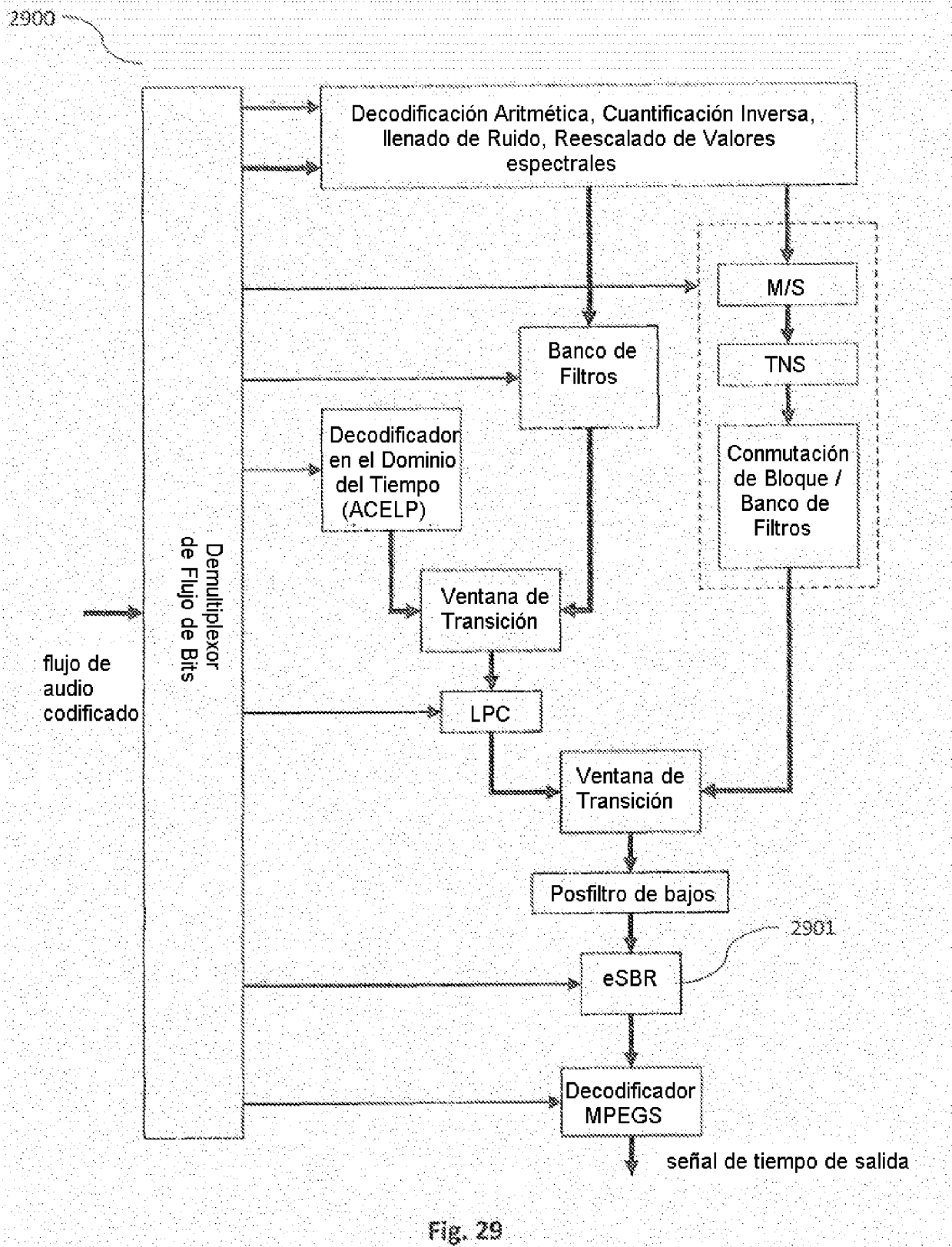


Fig. 29

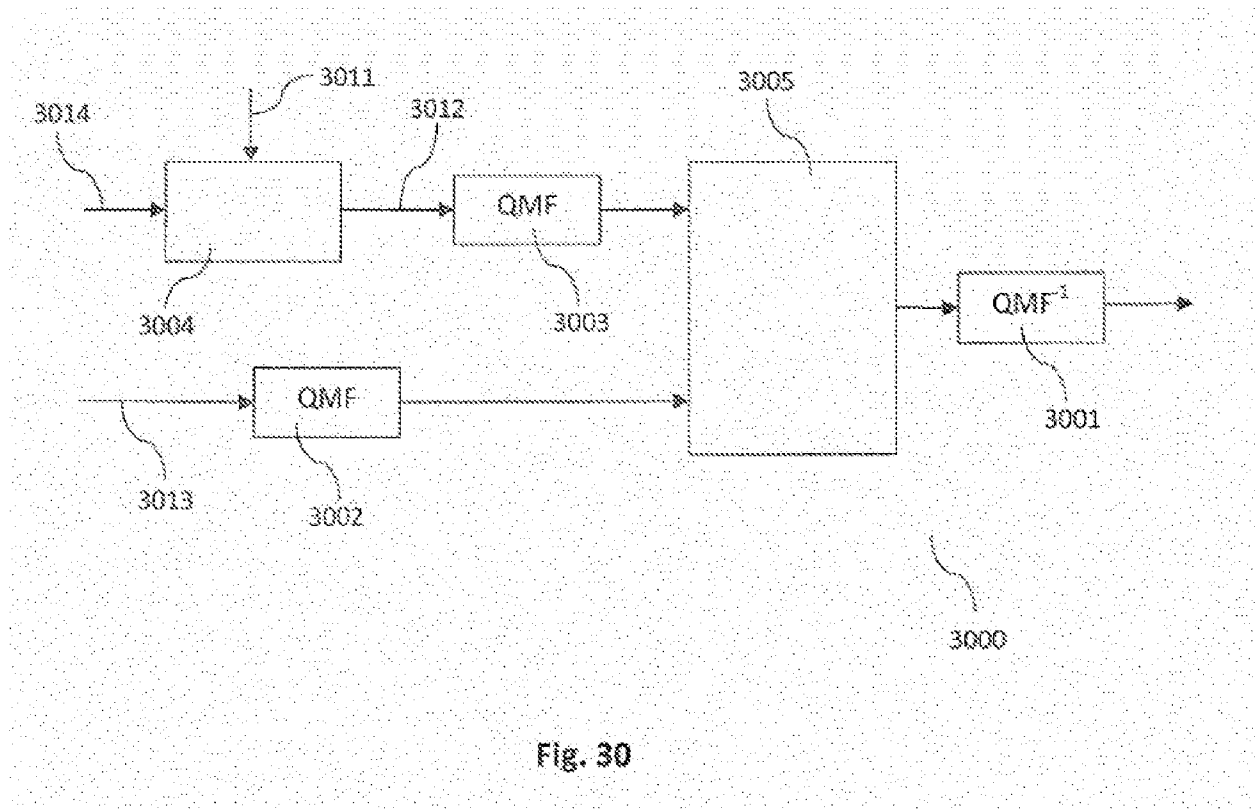


Fig. 30