

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 826 324**

51 Int. Cl.:

G10L 21/04 (2013.01)

G10L 19/022 (2013.01)

G10L 21/038 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **12.03.2010** **E 17175871 (7)**

97 Fecha y número de publicación de la concesión europea: **26.08.2020** **EP 3246919**

54 Título: **Transposición armónica mejorada**

30 Prioridad:

28.01.2009 SE 0900087

18.09.2009 US 243624 P

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
18.05.2021

73 Titular/es:

DOLBY INTERNATIONAL AB (100.0%)
Apollo Building, 3E, Herikerbergweg 1-35
1101 CN Amsterdam Zuidoost, NL

72 Inventor/es:

EKSTRAND, PER y
VILLEMOES, LARS FALCK

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 826 324 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

Campo técnico

- 5 La presente invención se refiere a la transposición de señales en frecuencia y/o al estiramiento/compresión de una señal en el tiempo y, en particular, a la codificación de señales de audio. Dicho de otro modo, la presente invención se refiere a una modificación en la escala de tiempo y/o en la escala de frecuencia. Más en particular, la presente invención se refiere a procedimientos de reconstrucción de altas frecuencias (HFR), incluido un transponedor armónico de dominio de frecuencia.

Antecedentes de la invención

- 10 Las tecnologías HFR, tales como la tecnología de replicación de banda espectral (SBR), permiten mejorar significativamente la eficiencia de codificación de los códecs de audio perceptuales tradicionales. Enfoques ejemplares que usan SBR se pueden encontrar en NPL1 y NPL2. En combinación con la codificación de audio avanzada (AAC) de MPEG-4, la tecnología HFR crea un códec de audio muy eficiente, que ya se usa en el sistema de Radio por Satélite XM y en la Radio Digital Mundial, y también normalizado en el 3GPP, el Fórum DVD, y otros. La combinación de AAC y SBR
15 se denomina aacPlus. Es parte de la norma MPEG-4, en la que se denomina perfil AAC de alta eficiencia (HE-AAC). En general, la tecnología HFR puede combinarse con cualquier códec de audio perceptual de manera compatible con versiones anteriores y nuevas, ofreciendo así la posibilidad de actualizar sistemas de difusión ya establecidos, como la Capa 2 de MPEG usada en el sistema Eureka DAB. Los procedimientos de transposición HFR también pueden combinarse con códecs de voz para permitir voz de banda ancha a tasas de transferencia de bits ultrabajas.

- 20 La idea básica subyacente en la HRF 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 al intervalo de altas frecuencias.

- 25 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 lograrse un ahorro sustancial en la tasa de transferencia de bits usando este concepto de codificación de audio y/o codificación de voz. A continuación, se hará referencia a la codificación de audio, pero debe apreciarse que los procedimientos y sistemas descritos pueden aplicarse igualmente a la codificación de voz y a la codificación de voz y de audio unificada (USAC).
30

- 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 central para su codificación, y frecuencias más altas se regeneran en el lado del descodificador usando la transposición de la señal de bajo ancho de banda e información lateral adicional, que se codifica normalmente a tasas de transferencia de bits muy bajas y que describe la forma espectral objetivo. Para tasas de transferencia de bits bajas,
35 donde el ancho de banda de la señal codificada central es estrecho, cada vez es más importante reproducir o sintetizar una banda alta, es decir, el intervalo de altas frecuencias de la señal de audio, con características perceptivamente agradables.

- En la técnica anterior hay varios procedimientos para la reconstrucción de altas frecuencias que usan, por ejemplo, transposición armónica o estiramiento en el tiempo. Un procedimiento está basado en codificadores de voz de fase que
40 funcionan según el principio de llevar a cabo un análisis de frecuencia con una resolución de frecuencias suficientemente altas. Una modificación de señal se lleva a cabo en el dominio de frecuencia antes de volver a sintetizar la señal. La modificación de señales puede ser una operación de estiramiento en el tiempo o de transposición.

- Uno de los problemas subyacentes que existen con estos procedimientos son las contradictorias restricciones de una resolución de alta frecuencia prevista con el fin de conseguir una transposición de alta calidad para sonidos estacionarios y la respuesta en el tiempo del sistema para sonidos transitorios o percusivos. Dicho de otro modo, aunque el uso de una resolución de alta frecuencia es beneficioso para la transposición de señales estacionarias, tal resolución de alta frecuencia requiere normalmente grandes tamaños de ventana, los cuales son perjudiciales cuando se tratan partes transitorias de una señal. Un enfoque para abordar este problema puede ser cambiar de manera adaptativa las ventanas del transponedor, por ejemplo, usando conmutación de ventanas, en función de las características de las señales de
45 entrada. Normalmente, ventanas largas se usarán para partes estacionarias de una señal, con el fin de conseguir una resolución de alta frecuencia, mientras que ventanas cortas se usarán para partes transitorias de la señal, con el fin de implementar una buena respuesta transitoria, es decir, una buena resolución temporal, del transponedor. Sin embargo, este enfoque tiene la desventaja de que es necesario incorporar medidas de análisis de señales, tales como la detección de transitorios o similares, en el sistema de transposición. Tales medidas de análisis de señales implican con frecuencia una etapa de decisión, por ejemplo, una decisión acerca de la presencia de un transitorio, que activa la conmutación del procesamiento de señales. Además, tales medidas afectan normalmente a la fiabilidad del sistema y pueden introducir artefactos de señal cuando se conmuta el procesamiento de señales, por ejemplo, cuando se conmuta entre tamaños de
50 ventana.

La presente invención soluciona los problemas antes mencionados relativos al rendimiento transitorio de la transposición armónica sin necesidad de conmutar ventanas. Además, la transposición armónica mejorada se consigue con una baja complejidad adicional.

5 NPL1: Max Neuendorf et al.: "Detailed Technical Description of Reference Model 0 of the CfP on Unified Speech and Audio Coding (USAC)"; 86. MPEG Meeting; 13-10-2008 - 17-10-2008; BUSAN; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/ WG11), nº M15867, 9 Octubre 2008 (09-10-2008).

NPL2: Lars Villemoes et al.: "Core experiment proposal on the USAC eSBR module", 87. MPEG Meeting; 2-2-2009 - 6-2-2009; LAUSANNE; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11), nº M16142, 29 Enero 2009 (29-01-2009).

10 Compendio de la invención

La presente invención se refiere al problema del rendimiento transitorio mejorado para una transposición armónica, así como a diversas mejoras en procedimientos conocidos de transposición armónica. Además, la presente invención indica cómo la complejidad adicional puede mantenerse al mínimo a la vez que se consiguen las mejoras propuestas.

Entre otras cosas, la presente invención puede comprender al menos uno de los siguientes aspectos:

- 15 - sobremuestrear en frecuencia mediante un factor en función del factor de transposición del punto de funcionamiento del transponedor;
- elegir de manera apropiada la combinación de ventanas de análisis y síntesis; y
- garantizar una alineación en el tiempo de diferentes señales transpuestas para los casos en que tales señales están combinadas.

20 La invención está definida por las reivindicaciones independientes 1, 7 y 9. Según un aspecto de la invención, se describe un sistema para generar una señal de salida transpuesta a partir de una señal de entrada usando un factor de transposición T. La señal de salida transpuesta puede ser una versión estirada en el tiempo y/o desplazada en frecuencia de la señal de entrada. Con respecto a la señal de entrada, la señal de salida transpuesta puede estirarse en el tiempo mediante el factor de transposición T. Como alternativa, las componentes de frecuencia de la señal de salida transpuesta pueden desplazarse de manera ascendente mediante el factor de transposición T.

25 El sistema puede comprender una ventana de análisis de longitud L que extrae L muestras de la señal de entrada. Normalmente, las L muestras de las señales de entrada son muestras de la señal de entrada, por ejemplo, una señal de audio, en el dominio de tiempo. Las L muestras extraídas se denominan trama de la señal de entrada. El sistema comprende además una unidad de transformación de análisis de orden $M = F \cdot L$ que transforma las L muestras de dominio de tiempo en M coeficientes complejos, donde F es un factor de sobremuestreo de frecuencia. Los M coeficientes complejos son normalmente coeficientes en el dominio de frecuencia. La transformación de análisis puede ser una transformada de Fourier, una transformada rápida de Fourier, una transformada discreta de Fourier, una transformada de ondículas o una etapa de análisis de un banco de filtros (posiblemente modulado). El factor de sobremuestreo F está basado en o es una función del factor de transposición T.

35 La operación de sobremuestreo también puede denominarse relleno con ceros de la ventana de análisis mediante $(F-1) \cdot L$ ceros adicionales. También puede considerarse como la elección de un tamaño de una transformación de análisis M que es mayor que el tamaño de la ventana de análisis en un factor F.

40 El sistema puede comprender además una unidad de procesamiento no lineal que modifica la fase de los coeficientes complejos usando el factor de transposición T. La modificación de la fase puede comprender multiplicar la fase de los coeficientes complejos por el factor de transposición T. Además, el sistema puede comprender una unidad de transformación de síntesis de orden M que transforma los coeficientes modificados en M muestras modificadas y una ventana de síntesis de longitud L para generar la señal de salida. La transformada de síntesis puede ser una transformada de Fourier inversa, una transformada rápida de Fourier inversa, una transformada discreta de Fourier inversa, una transformada de ondículas inversa o una etapa de síntesis de un banco de filtros (posiblemente) modulado.

45 Normalmente, la transformada de análisis y la transformada de síntesis están relacionadas entre sí, por ejemplo, para conseguir una reconstrucción perfecta de una señal de entrada cuando el factor de transposición T es igual a 1.

Según otro aspecto de la invención, el factor de sobremuestreo F es proporcional al factor de transposición T. En particular, el factor de sobremuestreo F puede ser mayor que o igual a $(T+1)/2$. Esta selección del factor de sobremuestreo F garantiza que artefactos de señal no deseados, por ejemplo, ecos previos y posteriores, que pueden haberse generado por la transposición, sean rechazados por la ventana de síntesis.

50 Debe observarse que, en términos más generales, la longitud de la ventana de análisis puede ser L_a y la longitud de la ventana de síntesis puede ser L_s . Además, en tales casos, puede ser beneficioso seleccionar el orden de la unidad de transformación M según el orden de transposición T, es decir, como una función del orden de transposición T. Además, puede ser beneficioso seleccionar M de modo que sea mayor que la longitud media de la ventana de análisis y de la

ventana de síntesis, es decir, mayor que $(L_a + L_s)/2$. En una forma de realización, la diferencia entre el orden de la unidad de transformación M y la longitud de ventana media es proporcional a $(T-1)$. En una forma de realización adicional, M se selecciona de modo que sea mayor que o igual a $(T L_a + L_s)/2$. Debe observarse que el caso en que la longitud de la ventana de análisis y de la ventana de síntesis es igual, es decir, $L_a = L_s = L$, es un caso especial del caso genérico anterior. En lo que respecta al caso genérico, el factor de sobremuestreo F puede ser

$$F \geq 1 + (T-1) \frac{L_a}{L_s + L_a}$$

El sistema puede comprender además una unidad de paso de análisis que desplaza la ventana de análisis en un paso de análisis de S_a muestras a lo largo de la señal de entrada. Como resultado de la unidad de paso de análisis se genera una sucesión de tramas de la señal de entrada. Además, el sistema puede comprender una unidad de paso de síntesis que desplaza la ventana de síntesis y/o tramas sucesivas de la señal de salida en un paso de síntesis de S_s muestras. Como resultado, se genera una sucesión de tramas desplazadas de la señal de salida que pueden solaparse y sumarse en una unidad de solapamiento y suma.

Dicho de otro modo, la ventana de análisis puede extraer o aislar L, o de manera más genérica L_a , muestras de la señal de entrada, por ejemplo, multiplicando un conjunto de L muestras de la señal de entrada por coeficientes de ventana distintos de cero. Tal conjunto de L muestras puede denominarse trama de señal de entrada o trama de la señal de entrada. La unidad de paso de análisis desplaza la ventana de análisis a lo largo de la señal de entrada y, por lo tanto, selecciona una trama diferente de la señal de entrada, es decir, genera una secuencia de tramas de la señal de entrada. La distancia de las muestras entre tramas sucesivas viene dada por el paso de análisis. De manera similar, la unidad de paso de síntesis desplaza la ventana de síntesis y/o las tramas de la señal de salida, es decir, genera una secuencia de tramas desplazadas de la señal de salida. La distancia de las muestras entre tramas sucesivas de la señal de salida viene dada por el paso de síntesis. La señal de salida puede determinarse solapando la secuencia de tramas de la señal de salida y sumando valores de muestra que coinciden en el tiempo.

Según un aspecto adicional de la invención, el paso de síntesis es T veces el paso de análisis. En tales casos, la señal de salida corresponde a la señal de entrada estirada en el tiempo mediante el factor de transposición T. Dicho de otro modo, al seleccionar que el paso de síntesis sea T veces mayor que el paso de análisis, puede obtenerse un desplazamiento en el tiempo o estiramiento en el tiempo de la señal de salida con respecto a la señal de entrada. Este desplazamiento en el tiempo es de orden T.

Dicho de otro modo, el sistema antes mencionado puede describirse de la siguiente manera: Usando una unidad de ventana de análisis, una unidad de transformación de análisis y una unidad de paso de análisis con un paso de análisis S_a , una serie o secuencia de conjuntos de M coeficientes complejos puede determinarse a partir de una señal de entrada. El paso de análisis define el número de muestras por el que la ventana de análisis se desplaza a lo largo de la señal de entrada. Puesto que el tiempo transcurrido entre dos muestras sucesivas viene dado por la tasa de muestreo, el paso de análisis también define el tiempo transcurrido entre dos tramas de la señal de entrada. En consecuencia, también el tiempo transcurrido entre dos conjuntos sucesivos de M coeficientes complejos viene dado por el paso de análisis S_a .

Tras pasar por la unidad de procesamiento no lineal en la que la fase de los coeficientes complejos puede modificarse, por ejemplo, multiplicándola por el factor de transposición T, la serie o secuencia de conjuntos de M coeficientes complejos puede volverse a convertir al dominio de tiempo. Cada conjunto de M coeficientes complejos modificados puede transformarse en M muestras modificadas usando la unidad de transformación de síntesis. En una operación de solapamiento y suma subsiguiente que implica a la unidad de ventana de síntesis y a la unidad de paso de síntesis con un paso de síntesis S_s , la serie de conjuntos de M muestras modificadas puede solaparse y sumarse para formar la señal de salida. En esta operación de solapamiento y suma, conjuntos sucesivos de M muestras modificadas pueden ser desplazados por S_s muestras entre sí antes de que puedan multiplicarse por la ventana de síntesis y posteriormente sumarse para proporcionar la señal de salida. Por consiguiente, si el paso de síntesis S_s es T veces el paso de análisis S_a , la señal puede estirarse en el tiempo conforme a un factor T.

Según un aspecto adicional de la invención, la ventana de síntesis se obtiene a partir de la ventana de análisis y el paso de síntesis. En particular, la ventana de síntesis puede venir dada por la fórmula:

$$v_s(n) = v_a(n) \left(\sum_{k=-\infty}^{\infty} (v_a(n - k \cdot \Delta t))^2 \right)^{-1},$$

donde $v_s(n)$ es la ventana de síntesis, $V_a(n)$ es la ventana de análisis y Δt es el paso de síntesis S_s . La ventana de análisis y/o de síntesis puede ser una de entre una ventana gaussiana, una ventana de coseno, una ventana de Hamming, una ventana de Hann, una ventana rectangular, una ventana de Bartlett, una ventana de Blackman, una ventana que tiene la función

$$v(n) = \text{sen} \left(\frac{\pi}{L} (n + 0.5) \right), 0 \leq n < L$$

donde en el caso de diferentes longitudes de la ventana de análisis y de la ventana de síntesis, L puede ser L_a o L_s , respectivamente.

Según otro aspecto de la invención, el sistema comprende además una unidad de contracción que lleva a cabo, por ejemplo, una conversión de tasa de la señal de salida según el orden de transposición T, obteniéndose así una señal de salida transpuesta. Seleccionando que el paso de síntesis sea T veces el paso de análisis, puede obtenerse una señal de salida estirada en el tiempo como se ha descrito anteriormente. Si la tasa de muestreo de la señal estirada en el tiempo aumenta en un factor T, o si la señal estirada en el tiempo se submuestra en un factor T, puede generarse una señal de salida transpuesta que corresponde a la señal de entrada, desplazada en frecuencia por el factor de transposición T. La operación de submuestreo puede comprender la etapa de seleccionar solamente un subconjunto de muestras de la señal de salida. Normalmente, solo se conserva cada $T^{\text{ésima}}$ muestra de la señal de salida. Como alternativa, la tasa de muestreo puede aumentarse por un factor T, es decir, se interpreta que la tasa de muestreo es T veces mayor. Dicho de otro modo, un nuevo muestreo o la conversión de la tasa de muestreo significa que la tasa de muestreo ha cambiado, o bien a un valor superior o un valor inferior. Submuestreo significa una conversión de tasa a un valor inferior.

Según un aspecto adicional de la invención, el sistema puede generar una segunda señal de salida a partir de la señal de entrada. El sistema puede comprender una segunda unidad de procesamiento no lineal que modifica la fase de los coeficientes complejos usando un segundo factor de transposición T_2 y una segunda unidad de paso de síntesis que desplaza la ventana de síntesis y/o las tramas de la segunda señal de salida por un segundo paso de síntesis. Modificar la fase puede comprender multiplicar la fase por un factor T_2 . Las tramas de la segunda señal de salida pueden generarse a partir de una trama de la señal de entrada modificando la fase de los coeficientes complejos usando el segundo factor de transposición y transformando los segundos coeficientes modificados en M segundas muestras modificadas y aplicando la ventana de síntesis. La segunda señal de salida puede generarse en la unidad de solapamiento y suma aplicando el segundo paso de síntesis a la secuencia de tramas de la segunda señal de salida.

La segunda señal de salida puede contraerse en una segunda unidad de contracción que lleva a cabo, por ejemplo, una conversión de tasa de la segunda señal de salida mediante el segundo orden de transposición T_2 . Esto proporciona una segunda señal de salida transpuesta. En resumen, una primera señal de salida transpuesta puede generarse usando el primer factor de transposición T y una segunda señal de salida transpuesta puede generarse usando el segundo factor de transposición T_2 . Estas dos señales de salida transpuestas pueden fusionarse después en una unidad de combinación para proporcionar la señal de salida transpuesta global. La operación de fusión puede comprender sumar las dos señales de salida transpuestas. Tal generación y combinación de una pluralidad de señales de salida transpuestas puede ser beneficiosa para obtener buenas aproximaciones de la componente de señal de alta frecuencia que va a sintetizarse. Debe observarse que cualquier número de señales de salida transpuestas puede generarse usando una pluralidad de órdenes de transposición. Esta pluralidad de señales de salida transpuestas puede fusionarse después, por ejemplo, sumarse, en una unidad de combinación para proporcionar una señal de salida transpuesta global.

Puede ser beneficioso que la unidad de combinación pondere la primera y la segunda señal de salida transpuestas antes de su fusión. La ponderación puede llevarse a cabo de manera que la energía o la energía por ancho de banda de la primera y la segunda señal de salida transpuestas corresponda a la energía o energía por ancho de banda de la señal de entrada, respectivamente.

Según un aspecto adicional de la invención, el sistema puede comprender una unidad de alineamiento que aplica un desfase de tiempo a la primera y a la segunda señales de salida transpuestas antes de que se introduzcan en la unidad de combinación. Tal desfase de tiempo puede comprender el desplazamiento de las dos señales de salida transpuestas entre sí en el dominio de tiempo. El desfase de tiempo puede depender del orden de transposición y/o de la longitud de las ventanas. En particular, el desfase de tiempo puede determinarse como

$$\frac{(T-2)L}{4}.$$

Según otro aspecto de la invención, el sistema de transposición antes descrito puede incluirse en un sistema que descodifica una señal multimedia recibida que comprende una señal de audio. El sistema de descodificación puede comprender una unidad de transposición que corresponde al sistema antes descrito, donde la señal de entrada es normalmente una componente de baja frecuencia de la señal de audio y la señal de salida es una componente de alta frecuencia de la señal de audio. Dicho de otro modo, la señal de entrada es normalmente una señal de paso bajo con un determinado ancho de banda, y la señal de salida es una señal de paso banda normalmente de un ancho de banda mayor. Además, puede comprender un descodificador central para descodificar la componente de baja frecuencia de la señal de audio a partir del flujo de bits recibido. Tal descodificador central puede basarse en un esquema de codificación tal como Dolby E, Dolby Digital o AAC. En particular, tal sistema de descodificación puede ser un descodificador que descodifica una señal multimedia recibida que comprende una señal de audio y otras señales, tales como por ejemplo de

vídeo.

Debe observarse que la presente invención también describe un procedimiento para transponer una señal de entrada por un factor de transposición T. El procedimiento corresponde al sistema antes descrito y puede comprender cualquier combinación de los aspectos antes mencionados. Puede comprender las etapas de extraer muestras de la señal de entrada usando una ventana de análisis de longitud L, y de seleccionar un factor de sobremuestreo F en función del factor de transposición T. Puede comprender además las etapas de transformar las L muestras desde el dominio de tiempo al dominio de frecuencia, proporcionando F * L coeficientes complejos, y de modificar la fase de los coeficientes complejos con el factor de transposición T. En etapas adicionales, el procedimiento puede transformar los F * L coeficientes complejos modificados al dominio de tiempo, proporcionando F * L muestras modificadas, y puede generar la señal de salida usando una ventana de síntesis de longitud L. Debe observarse que el procedimiento también puede estar adaptado a longitudes generales de la ventana de análisis y de síntesis, es decir, a L_a y L_s genéricas, como se ha indicado anteriormente.

Según un aspecto adicional de la invención, el procedimiento puede comprender las etapas de desplazar la ventana de análisis en un paso de análisis de S_a muestras a lo largo de la señal de entrada, y/o de desplazar la ventana de síntesis y/o las tramas de la señal de salida en un paso de síntesis de S_s muestras. La señal de salida puede estirarse en el tiempo con respecto a la señal de entrada en un factor T seleccionando que el paso de síntesis sea T veces el paso de análisis. Cuando se ejecuta una etapa adicional que lleva a cabo una conversión de tasa de la señal de salida por el orden de transposición T, puede obtenerse una señal de salida transpuesta. Tal señal de salida transpuesta puede comprender componentes de frecuencia que se han desplazado de manera ascendente en un factor T con respecto a las componentes de frecuencia correspondientes de la señal de entrada.

El procedimiento puede comprender además etapas para generar una segunda señal de salida. Esto puede implementarse modificando la fase de los coeficientes complejos usando un segundo factor de transposición T_2 ; desplazando la ventana de síntesis y/o las tramas de la segunda señal de salida mediante un segundo paso de síntesis puede generarse una segunda señal de salida usando el segundo factor de transposición T_2 y el segundo paso de síntesis. Una segunda señal de salida transpuesta puede generarse realizando una conversión de tasa de la segunda señal de salida por el segundo orden de transposición T_2 . Finalmente, fusionando la primera y la segunda señales de salida transpuestas puede obtenerse una señal de salida fusionada o transpuesta global que incluye componentes de señal de alta frecuencia generadas mediante dos o más transposiciones con diferentes factores de transposición.

Según otros aspectos de la invención, la invención describe un programa de software adaptado para ejecutarse en un procesador y para realizar las etapas de procedimiento de la presente invención cuando se lleva a cabo en un dispositivo informático. La invención describe además 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 invención cuando se lleva a cabo en un dispositivo informático. Además, la invención describe un producto de programa informático que comprende instrucciones ejecutables para llevar a cabo el procedimiento de la invención cuando se ejecutan en un ordenador.

Según un aspecto adicional, se describe otro procedimiento y sistema para transponer una señal de entrada en un factor de transposición T. Este procedimiento y sistema puede usarse de manera autónoma o en combinación con los procedimientos y sistemas antes descritos. Cualquiera de las características descritas en el presente documento puede aplicarse a este procedimiento/sistema y viceversa.

El procedimiento puede comprender la etapa de extraer una trama de muestras de la señal de entrada usando una ventana de análisis de longitud L. Después, la trama de la señal de entrada puede transformarse desde el dominio de tiempo al dominio de frecuencia proporcionando M coeficientes complejos. La fase de los coeficientes complejos puede modificarse con el factor de transposición T, y los M coeficientes complejos modificados pueden transformarse al dominio de tiempo proporcionando M muestras modificadas. Finalmente, una trama de una señal de salida puede generarse usando una ventana de síntesis de longitud L. El procedimiento y sistema puede usar una ventana de análisis y una ventana de síntesis que son diferentes entre sí. La ventana de análisis y la ventana de síntesis pueden ser diferentes en cuanto a su forma, su longitud, el número de coeficientes que definen las ventanas y/o los valores de los coeficientes que definen las ventanas. De este modo pueden obtenerse grados de libertad adicionales en la selección de las ventanas de análisis y de síntesis, de tal modo puede reducirse o eliminarse el solapamiento de la señal de salida transpuesta.

Según otro aspecto, la ventana de análisis y la ventana de síntesis son bi-ortogonales entre sí. La ventana de síntesis $v_s(n)$ puede venir dada por:

$$v_s(n) = c \frac{v_a(n)}{s(n \bmod \Delta t_s)}, \quad 0 \leq n < L,$$

donde c es una constante, $V_a(n)$ es la ventana de análisis (311), Δt_s es un paso de tiempo de la ventana de síntesis y $s(n)$ viene dado por:

$$s(m) = \sum_{i=0}^{L/(\Delta t_s)-1} v_a^2(m + \Delta t_s i), \quad 0 \leq m < \Delta t_s.$$

El paso de tiempo de la ventana de síntesis Δt_s corresponde normalmente al paso de síntesis S_s .

Según un aspecto adicional, la ventana de análisis puede seleccionarse de manera que su transformada z tenga dos ceros en el círculo unitario. Preferentemente, la transformada z de la ventana de análisis solo tiene dos ceros en el círculo unitario. A modo de ejemplo, la ventana de análisis puede ser una ventana de seno al cuadrado. En otro ejemplo, la ventana de análisis de longitud L puede determinarse mediante convolución de dos ventanas de seno de longitud L , lo que proporciona una ventana de seno al cuadrado de longitud $2L-1$. En una etapa adicional se añade un cero a la ventana de seno al cuadrado, lo que proporciona una ventana base de longitud $2L$. Finalmente, la ventana base puede volver a muestrearse usando interpolación lineal, lo que proporciona una ventana de simetría par de longitud L como ventana de análisis.

Los procedimientos y sistemas 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 y/o como circuitos integrados de aplicación específica. Las señales encontradas en los procedimientos y sistemas descritos pueden almacenarse en medios tales como memoria de acceso aleatorio o medios de almacenamiento óptico. Pueden transferirse mediante redes, tales como redes radioeléctricas, redes de satélites, redes inalámbricas o redes cableadas, por ejemplo, Internet. Dispositivos típicos que usan el procedimiento y sistema descritos en el presente documento son descodificadores u otros equipos en las instalaciones de los clientes que descodifiquen señales de audio. En el lado de la codificación, el procedimiento y sistema puede usarse en estaciones de radiodifusión, por ejemplo, en sistemas de distribución de televisión o vídeo.

Debe observarse que las formas de realización y los aspectos de la invención descritos en este documento pueden combinarse de manera arbitraria. En particular, debe observarse que los aspectos descritos para un sistema también pueden aplicarse al procedimiento correspondiente perteneciente a la presente invención.

Breve descripción de los dibujos

A continuación, se describirá la presente invención a través de ejemplos ilustrativos, que no limitan el alcance de la invención definida en las reivindicaciones adjuntas, haciendo referencia a los dibujos adjuntos, en los que:

la Fig. 1 ilustra un Dirac en una posición particular tal y como aparece en las ventanas de análisis y de síntesis de un transponedor armónico;

la Fig. 2 ilustra un Dirac en una posición diferente tal y como aparece en las ventanas de análisis y de síntesis de un transponedor armónico;

la Fig. 3 ilustra un Dirac para la posición de la Fig. 2 tal y como aparece según la presente invención;

la Fig. 4 ilustra el funcionamiento de un descodificador de audio mejorado HFR;

la Fig. 5 ilustra el funcionamiento de un transponedor armónico usando varios órdenes;

la Fig. 6 ilustra el funcionamiento de un transponedor armónico de dominio de frecuencia (FD);

la Fig. 7 muestra una sucesión de ventanas de análisis y de síntesis;

la Fig. 8 ilustra ventanas de análisis y síntesis con diferentes pasos;

la Fig. 9 ilustra el efecto del nuevo muestreo en el paso de síntesis de ventanas;

las Fig. 10 y 11 ilustran formas de realización de un codificador y un descodificador, respectivamente, usando los esquemas de transposición armónica mejorada descritos en el presente documento; y

la Fig. 12 ilustra una forma de realización de una unidad de transposición mostrada en las Fig. 10 y 11.

Descripción detallada

Las formas de realización descritas a continuación simplemente ilustran los principios de la presente invención para lograr una transposición armónica mejorada. Debe entenderse que modificaciones y variaciones de las disposiciones y de los detalles descritos en el presente documento resultarán evidentes a otros expertos en la técnica. Por lo tanto, solo estarán limitados por el alcance de las reivindicaciones de patente adjuntas y no por los detalles específicos presentados a modo de descripción y explicación de las formas de realización en el presente documento.

A continuación, se describirá el principio de transposición armónica en el dominio de frecuencia y las mejoras propuestas como se enseñan por la presente invención. Un elemento importante de la transposición armónica es el estiramiento en el tiempo mediante un factor de transposición entero T que conserva la frecuencia de las sinusoides. Dicho de otro modo, la transposición armónica está basada en el estiramiento en el tiempo de la señal subyacente en un factor T . El estiramiento en el tiempo se lleva a cabo de manera que se mantengan las frecuencias de las sinusoides que componen

la señal de entrada. Tal estiramiento en el tiempo puede llevarse a cabo usando un codificador de voz de fase. El codificador de voz de fase está basado en una representación en el dominio de frecuencia mediante un banco de filtros DFT divididos en ventanas con una ventana de análisis $v_a(n)$ y una ventana de síntesis $v_s(n)$. Tal transformación de análisis/síntesis también se denomina transformada de Fourier de tiempo corto (STFT).

5 Una transformada de Fourier de tiempo corto se lleva a cabo sobre una señal de entrada de dominio de tiempo para obtener una serie de tramas espectrales solapadas. Para minimizar posibles efectos de banda lateral, deben seleccionarse ventanas de análisis/síntesis apropiadas, por ejemplo, ventanas gaussianas, ventanas de coseno, ventanas de Hamming, ventanas de Hann, ventanas rectangulares, ventanas de Bartlett, ventanas de Blackman y otras. El retardo de tiempo en que se capta cada trama espectral a partir de la señal de entrada se determina tamaño de salto o
10 paso. La STFT de la señal de entrada se denomina etapa de análisis y da lugar a una representación en el dominio de frecuencia de la señal de entrada. La representación en el dominio de frecuencia comprende una pluralidad de señales de subbanda, donde cada señal de subbanda representa una determinada componente de frecuencia de la señal de entrada.

La representación en el dominio de frecuencia de la señal de entrada puede procesarse después de la manera deseada. Con el fin de estirar en el tiempo la señal de entrada, cada señal de subbanda puede estirarse en el tiempo, por ejemplo, retardando las muestras de señales de subbanda. Esto puede conseguirse usando un tamaño de salto de síntesis mayor que el tamaño de salto de análisis. La señal de dominio de tiempo puede reconstruirse llevando a cabo una transformada (rápida) de Fourier inversa en todas las tramas, seguido de una acumulación sucesiva de las tramas. Esta operación de la etapa de síntesis se denomina operación de solapamiento y suma. La señal de salida resultante es una versión
20 estirada en el tiempo de la señal de entrada, que comprende las mismas componentes de frecuencia que la señal de entrada. Dicho de otro modo, la señal de salida resultante tiene la misma composición espectral que la señal de entrada, pero es más lenta que la señal de entrada, es decir, su progresión está estirada en el tiempo.

La transposición a frecuencias más altas puede obtenerse posteriormente, o de manera integrada, mediante un submuestreo de las señales estiradas. Como resultado, la señal transpuesta tiene la longitud en el tiempo de la señal inicial, pero comprende componentes de frecuencia que están desplazadas hacia arriba mediante un factor de transposición predefinido.

En términos matemáticos, el codificador de voz de fase puede describirse de la siguiente manera. Una señal de entrada $x(t)$ se muestrea a una tasa de muestreo R para proporcionar la señal de entrada discreta $x(n)$. Durante la etapa de análisis, una STFT se determina para la señal de entrada $x(n)$ en instantes de tiempo t_a^k de análisis particulares para valores sucesivos k . Los instantes de tiempo de análisis se seleccionan preferentemente de manera uniforme mediante
30 $t_a^k = k \cdot \Delta t_a$, donde Δt_a es el factor de salto de análisis o el paso de análisis. En cada uno de estos instantes de tiempo t_a^k de análisis se calcula una transformada de Fourier en una parte dividida en ventanas de la señal original $x(n)$, donde la ventana de análisis $v_a(t)$ está centrada en torno a t_a^k , es decir,

$v_a(t - t_a^k)$. Esta parte dividida en ventanas de la señal de entrada $x(n)$ se denomina trama. El resultado es la
35 representación STFT de la señal de entrada $x(n)$, que puede denotarse como:

$$X(t_a^k, \Omega_m) = \sum_{n=-\infty}^{\infty} v_a(n - t_a^k) x(n) \exp(-j\Omega_m n),$$

donde

$$\Omega_m = 2\pi \frac{m}{M}$$

es la frecuencia central de la m ésima señal de subbanda del análisis STFT y M es el tamaño de la transformada discreta de Fourier (DFT). En la práctica, la función de ventana $v_a(n)$ tiene un lapso de tiempo limitado, es decir, solo cubre un número limitado de muestras L , que es normalmente igual al tamaño M de la DFT. En consecuencia, la suma anterior
40 tiene un número finito de términos. Las señales de subbanda $X(t_a^k, \Omega_m)$ son tanto una función de tiempo, mediante el índice k , como de frecuencia, mediante la frecuencia central Ω_m de subbanda.

La etapa de síntesis puede llevarse a cabo en los instantes de tiempo de síntesis t_s^k que están normalmente distribuidos
45 de manera uniforme según $t_s^k = k \cdot \Delta t_s$, donde Δt_s es el factor de salto de síntesis o paso de síntesis. En cada uno de estos instantes de tiempo de síntesis, una señal de corta duración $y_k(n)$ se obtiene sometiendo a una transformada de Fourier inversa la señal

$Y(t_s^k, \Omega_m)$, de subbanda STFT que puede ser idéntica a

$X(t_a^k, \Omega_m)$, en los instantes de tiempo t_s^k de síntesis

Sin embargo, normalmente las señales de subbanda STFT se modifican, por ejemplo, se estiran en el tiempo y/o se modulan en fase y/o en amplitud, de modo que la señal $X(t_a^k, \Omega_m)$ de subbanda de análisis es diferente de la señal $Y(t_s^k, \Omega_m)$ de subbanda de síntesis

En una forma de realización preferida, las señales de subbanda STFT se modulan en fase, es decir, la fase de las señales de subbanda STFT es modificada. La señal de síntesis $y_k(n)$ de corta duración puede denotarse como

$$y_k(n) = \frac{1}{M} \sum_{m=0}^{M-1} Y(t_s^k, \Omega_m) \exp(j\Omega_m n).$$

La señal $y_k(n)$ de corta duración puede considerarse como una componente de la señal de salida global $y(n)$ que

comprende las señales $Y(t_s^k, \Omega_m)$ de subbanda de síntesis

para $m = 0, \dots, M-1$ en el instante de tiempo t_s^k de síntesis

Es decir, la señal $y_k(n)$ de corta duración es la DFT inversa para una trama de señal específica. La señal de salida global $y(n)$ puede obtenerse solapando y sumando señales $y_k(n)$ de corta duración divididas en ventanas en todos los instantes de tiempo t_s^k de síntesis

Es decir, la señal de salida $y(n)$ puede denotarse como

$$y(n) = \sum_{k=-\infty}^{\infty} v_s(n - t_s^k) y_k(n - t_s^k),$$

Donde $v_s(n - t_s^k)$ es la ventana de síntesis centrada alrededor del instante de tiempo t_s^k de síntesis

Debe observarse que la ventana de síntesis tiene normalmente un número limitado de muestras L , de modo que la suma antes mencionada solo comprende un número limitado de términos.

A continuación, se describe la implementación del estiramiento en el tiempo en el dominio de frecuencia. Un punto de partida adecuado para describir los aspectos del estirador en el tiempo es considerar el caso en que $T=1$, es decir, el caso en que el factor de transposición T es igual a 1 y en el que no se produce ningún estiramiento. Suponiendo que el paso de tiempo de análisis Δt_a y el paso de tiempo de síntesis Δt_s del banco de filtros DFT son iguales, es decir, $\Delta t_a = \Delta t_s = \Delta t$, el efecto combinado de análisis seguido de síntesis es el de una modulación de amplitud con la función periódica Δt

$$K(n) = \sum_{k=-\infty}^{\infty} q(n - k\Delta t), \quad (1)$$

donde $q(n) = v_a(n)v_s(n)$ es el producto por puntos de las dos ventanas, es decir, el producto por puntos de la ventana de análisis y de la ventana de síntesis. Resulta ventajoso elegir las ventanas de modo que $K(n) = 1$ u otro valor constante, ya que así el banco de filtros DFT dividido en ventanas consigue una reconstrucción perfecta. Si se proporciona la ventana de análisis $v_a(n)$, y si la ventana de análisis tiene una duración suficientemente larga en comparación con el paso Δt , puede obtenerse una reconstrucción perfecta eligiendo la ventana de síntesis según

$$v_s(n) = v_a(n) \left(\sum_{k=-\infty}^{\infty} (v_a(n - k \cdot \Delta t))^2 \right)^{-1}. \quad (2)$$

Para $T > 1$, es decir, para un factor de transposición mayor que 1, un estiramiento en el tiempo puede obtenerse realizando el análisis en el paso

$\Delta t_a = \frac{\Delta t}{T}$ mientras que el paso de síntesis se mantiene en $\Delta t_s = \Delta t$. Dicho de otro modo, un estiramiento en el tiempo en

un factor T puede obtenerse aplicando un factor de salto o paso en la etapa de análisis que es T veces más pequeño que el factor de salto o paso en la etapa de síntesis. Como puede observarse en las fórmulas proporcionadas anteriormente, el uso de un paso de síntesis que es T veces mayor que el paso de análisis desplazará las señales de síntesis de corta duración $y_k(n)$ en intervalos T veces mayores en la operación de solapamiento y suma. Esto dará finalmente como resultado un estiramiento en el tiempo de la señal de salida $y(n)$.

Debe observarse que el estiramiento en el tiempo mediante el factor T puede implicar además una multiplicación de fase por un factor T entre el análisis y la síntesis. Dicho de otro modo, el estiramiento en el tiempo mediante un factor T implica una multiplicación de fase por un factor T de las señales de subbanda.

A continuación, se describe cómo la operación de estiramiento en el tiempo descrita anteriormente puede convertirse en una operación de transposición armónica. La modificación de la escala de tonos o transposición armónica puede obtenerse llevando a cabo una conversión de tasa de muestras de la señal de salida estirada en el tiempo $y(n)$. Para llevar a cabo una transposición armónica mediante un factor T, una señal de salida $y(n)$, que es una versión estirada en el tiempo según el factor T de la señal de entrada $x(n)$, puede obtenerse usando el procedimiento de codificación de voz de fase descrito anteriormente. La transposición armónica puede obtenerse entonces submuestreando la señal de salida $y(n)$ por un factor T o convirtiendo la tasa de muestreo de R a TR. Dicho de otro modo, en lugar de interpretar que la señal de salida $y(n)$ tiene la misma tasa de muestreo que la señal de entrada $x(n)$ pero con una duración de T veces, puede interpretarse que la señal de salida $y(n)$ tiene la misma duración, pero de T veces la tasa de muestreo. Puede interpretarse entonces que el submuestreo subsiguiente de T consiste en hacer que la tasa de muestreo de salida sea igual a la tasa de muestreo de entrada de modo que las señales puedan sumarse finalmente. Durante estas operaciones, debe tenerse cuidado cuando se submuestree la señal transpuesta para que no se produzcan solapamientos.

Cuando se supone que la señal de entrada $x(n)$ es una senoide y cuando se utiliza una ventana de análisis simétrica $v_a(n)$, el procedimiento de estiramiento en el tiempo basado en el codificador de voz de fase antes descrito funcionará correctamente para valores impares de T y dará como resultado una versión estirada en el tiempo de la señal de entrada $x(n)$, que presenta la misma frecuencia. En combinación con un submuestreo subsiguiente, se obtendrá una senoide $y(n)$ con una frecuencia que es T veces la frecuencia de la señal de entrada $x(n)$.

Para valores pares de T, el procedimiento de estiramiento en el tiempo/transposición armónica descrito anteriormente será más aproximado, ya que lóbulos laterales de valor negativo de la respuesta de frecuencia de la ventana de análisis $v_a(n)$ se reproducirán con diferente fidelidad mediante la multiplicación de fase. Los lóbulos laterales negativos se deben normalmente a que la mayoría de ventanas prácticas (o filtros de prototipo) tienen numerosos ceros discretos ubicados en el círculo unitario, dando como resultado desfases de 180 grados. Cuando se multiplican los ángulos de fase usando factores de transposición pares, los desfases pasan a ser normalmente 0 (o, en vez de ello, múltiplos de 360) grados, dependiendo del factor de transposición usado. Dicho de otro modo, cuando se usan factores de transposición pares, los desfases desaparecen. Esto dará lugar normalmente a un solapamiento en la señal de salida transpuesta $y(n)$. Un escenario particularmente desventajoso puede surgir cuando una senoide está ubicada en una frecuencia correspondiente a la parte superior del primer lóbulo lateral del filtro de análisis. Dependiendo del rechazo de este lóbulo en la respuesta de magnitud, el solapamiento será más o menos audible en la señal de salida. Debe observarse que, para factores T pares, reducir el paso global Δt mejora normalmente el rendimiento del estirador en el tiempo a expensas de una mayor complejidad desde el punto de vista computacional.

En el documento EP0940015B1 / WO98/57436 titulado "Source coding enhancement using spectral band replication" se describe un procedimiento sobre cómo evitar el solapamiento que se produce con un transponedor armónico cuando se usan factores de transposición pares. Este procedimiento, denominado bloqueo de fase relativa, evalúa la diferencia de fase relativa entre canales adyacentes y determina si una senoide está invertida en fase en cualquier canal. La detección se realiza usando la ecuación (32) del documento EP0940015B1. Los canales detectados como invertidos en fase se corrigen después de multiplicar los ángulos de fase por el factor de transposición real.

A continuación, se describe un procedimiento novedoso para evitar el solapamiento cuando se usan factores de transposición T pares y/o impares. A diferencia del procedimiento de bloqueo de fase relativa del documento EP0940015B1, este procedimiento no requiere la detección y corrección de ángulos de fase. La solución novedosa al problema anterior utiliza ventanas de transformada de análisis y de síntesis que no son idénticas. En el caso de reconstrucción perfecta (PR), esto corresponde a una transformada bi-ortogonal/banco de filtros en lugar de a una transformada ortogonal/banco de filtros.

Para obtener una transformada bi-ortogonal dada una determinada ventana de análisis $v_a(n)$, la ventana de síntesis $v_s(n)$ se elige de la siguiente manera:

$$\sum_{i=0}^{L/(\Delta t_s)-1} v_a(m + \Delta t_s i) v_s(m + \Delta t_s i) = c, \quad 0 \leq m < \Delta t_s$$

donde c es una constante, Δt_s es el paso de tiempo de síntesis y L es la longitud de ventana. Si la secuencia $s(n)$ se define como

$$s(m) = \sum_{i=0}^{L/(\Delta t_s)-1} v_a^2(m + \Delta t_s i), \quad 0 \leq m < \Delta t_s,$$

es decir, $v_a(n) = v_s(n)$ se usa para una división en ventanas tanto de análisis como de síntesis, entonces la condición para una transformada ortogonal es

$$s(m) = c, \quad 0 \leq m < \Delta t_s.$$

- 5 Sin embargo, a continuación, se introduce otra secuencia $w(n)$, donde $w(n)$ es una medida de cuánto se desvía la ventana de síntesis $v_s(n)$ con respecto a la ventana de análisis $v_a(n)$, es decir, en cuánto se diferencia la transformada bi-ortogonal con respecto al caso ortogonal. La secuencia $w(n)$ viene dada por

$$w(n) = \frac{v_s(n)}{v_a(n)}, \quad 0 \leq n < L.$$

La condición de la reconstrucción perfecta viene dada entonces por

$$10 \quad \sum_{i=0}^{L/(\Delta t_s)-1} v_a^2(m + \Delta t_s i) w(m + \Delta t_s i) = c, \quad 0 \leq m < \Delta t_s.$$

Para una posible solución, puede restringirse que $w(n)$ sea periódica con el paso de tiempo de síntesis Δt_s , es decir, $w(n) = w(n + \Delta t_s i)$, $\forall i, n$. Entonces, se obtiene

$$\sum_{i=0}^{L/(\Delta t_s)-1} v_a^2(m + \Delta t_s i) w(m + \Delta t_s i) = w(m) \sum_{i=0}^{L/(\Delta t_s)-1} v_a^2(m + \Delta t_s i) = w(m) s(m) = c, \\ 0 \leq m < \Delta t_s.$$

La condición en la ventana de síntesis $v_s(n)$ es, por tanto

$$15 \quad v_s(n) = w(n \bmod \Delta t_s) v_a(n) = c \frac{v_a(n)}{s(n \bmod \Delta t_s)}, \quad 0 \leq n < L.$$

Obteniendo las ventanas de síntesis $v_s(n)$ de la manera antes descrita se consigue una libertad mucho mayor a la hora de diseñar la ventana de análisis $v_a(n)$. Esta libertad adicional puede usarse para diseñar un par de ventanas de análisis/síntesis que no presentan solapamiento de la señal transpuesta.

- 20 Para obtener un par de ventanas de análisis/síntesis que supriman el solapamiento con factores de transposición pares, a continuación, se describen varias formas de realización. Según una primera forma de realización, las ventanas o filtros de prototipo se crean lo bastante largos como para atenuar el nivel del primer lóbulo lateral en la respuesta de frecuencia por debajo de un determinado nivel de "solapamiento". El paso de tiempo Δt_a de análisis será en este caso solamente una (pequeña) fracción de la longitud de ventana L . Esto da normalmente como resultado una distribución de transitorios, por ejemplo, en señales percutivas.

- 25 Según una segunda forma de realización, la ventana de análisis $v_a(n)$ se elige para que tenga dos ceros en el círculo unitario. La respuesta de fase resultante de dos ceros es un desfase de 360 grados. Estos desfases se mantienen cuando los ángulos de fase se multiplican por los factores de transposición, independientemente de si los factores de transposición son pares o impares. Cuando se obtiene un filtro de análisis $v_a(n)$ apropiado y homogéneo, que presenta dos ceros en el círculo unitario, la ventana de síntesis se obtiene a partir de las ecuaciones descritas anteriormente.

- 30 En un ejemplo de la segunda forma de realización, el filtro de análisis / la ventana $v_a(n)$ es la "ventana de seno al cuadrado", es decir, la ventana de seno

$$v(n) = \text{sen} \left(\frac{\pi}{L} (n + 0.5) \right), 0 \leq n < L$$

- 35 realizada la convolución con ella misma como $v_a(n) = v(n) \otimes v(n)$. Sin embargo, debe observarse que el filtro/ la ventana $v_a(n)$ resultante tendrá una simetría impar de longitud $L_a = 2L - 1$, es decir, un número impar de coeficientes de filtro/ventana. Cuando un filtro/ventana con una longitud par es más apropiado/a, en particular un filtro de simetría par, el filtro puede obtenerse realizando la convolución en primer lugar de dos ventanas de seno de longitud L . Después se añade un cero al final del filtro resultante. Después, el filtro de longitud $2L$ se muestrea de nuevo usando interpolación lineal con respecto a un filtro de simetría par y longitud L que sigue teniendo dos ceros solamente en el círculo unitario.

Se ha descrito en términos generales cómo puede seleccionarse un par de ventanas de análisis y de síntesis de manera

que el solapamiento en la señal de salida transpuesta pueda evitarse o reducirse significativamente. El procedimiento es particularmente relevante cuando se usan factores de transposición pares.

Otro aspecto a tener en cuenta en el contexto de transponedores armónicos basados en codificadores de voz es el desenrollado de fase. Debe observarse que aunque debe tenerse mucho cuidado en relación con los problemas de desenrollado de fase en los codificadores de voz de fase de propósito general, el transponedor armónico tiene operaciones de fase definidas de manera no ambigua cuando se usan factores de transposición enteros T . Por tanto, en formas de realización preferidas, el orden de transposición T es un valor entero. En caso contrario pueden aplicarse técnicas de desenrollado de fase, donde el desenrollado de fase es un proceso mediante el cual se usa el incremento de fase entre dos tramas consecutivas para estimar la frecuencia instantánea de una senoide cercana en cada canal.

Aún otro aspecto a tener en cuenta cuando se aborda la transposición de señales de audio y/o voz, es el procesamiento de secciones de señal estacionarias y/o transitorias. Normalmente, para poder transponer señales de audio estacionarias sin artefactos de intermodulación, la resolución de frecuencia del banco de filtros DFT tiene que ser bastante alta y, por lo tanto, las ventanas son largas en comparación con los transitorios de las señales de entrada $x(n)$, concretamente señales de audio y/o voz. Como resultado, el transponedor tiene una mala respuesta transitoria. Sin embargo, como se describirá a continuación, este problema puede resolverse modificando los parámetros de diseño de ventana, de tamaño de transformada y de paso de tiempo. Por tanto, a diferencia de muchos procedimientos de última generación para la mejora de la respuesta transitoria de codificadores de voz de fase, la solución propuesta no se basa en ninguna operación de adaptación de señal, tal como la detección de transitorios.

A continuación, se describe la transposición armónica de señales transitorias usando codificadores de voz. Como punto de partida, se tiene en cuenta una señal transitoria de prototipo, un pulso Dirac de tiempo discreto en el instante de tiempo $t = t_0$,

$$\delta(t - t_0) = \begin{cases} 1, & t = t_0 \\ 0, & t \neq t_0 \end{cases},$$

La transformada de Fourier de tal pulso de Dirac tiene una magnitud unitaria y una fase lineal con una pendiente proporcional a t_0 .

$$X(\Omega_m) = \sum_{n=-\infty}^{\infty} \delta(n - t_0) \exp(-j\Omega_m n) = \exp(-j\Omega_m t_0).$$

Tal transformada de Fourier puede considerarse como la etapa de análisis del codificador de voz de fase descrito anteriormente, donde se usa una ventana de análisis $v_a(n)$ plana de duración infinita. Para generar una señal de salida $y(n)$ que está estirada en el tiempo mediante un factor T , es decir, un pulso de Dirac $\delta(t - Tt_0)$ en el instante de tiempo $t = Tt_0$, la fase de las señales de subbanda de análisis debe multiplicarse por el factor T con el fin de obtener la señal de subbanda de síntesis $Y(\Omega_m) = \exp(-j\Omega_m Tt_0)$ que proporciona el pulso de Dirac deseado $\delta(t - Tt_0)$ como una salida de una transformada de Fourier inversa.

Esto muestra que el funcionamiento de la multiplicación de fase de las señales de subbanda de análisis por un factor T da lugar al desplazamiento de tiempo deseado de un pulso de Dirac, es decir, de una señal de entrada transitoria. Debe observarse que para señales transitorias más realistas que comprenden más de una muestra distinta de cero, deben realizarse las operaciones adicionales de estiramiento en el tiempo de las señales de subbanda de análisis en un factor T . Dicho de otro modo, deben usarse diferentes tamaños de salto en el lado de análisis y de síntesis.

Sin embargo, debe observarse que las consideraciones anteriores se refieren a una etapa de análisis/síntesis que usa ventanas de análisis y de síntesis de longitudes infinitas. De hecho, un transponedor teórico con una ventana de duración infinita proporcionaría el estiramiento correcto de un pulso de Dirac $\delta(t - t_0)$. Para un análisis dividido en ventanas de duración finita, la situación se aleatoriza por el hecho de que cada bloque de análisis debe interpretarse como un intervalo de periodo de una señal periódica con un periodo igual al tamaño de la DFT.

Esto se ilustra en la Fig. 1, que muestra el análisis y la síntesis 100 de un pulso de Dirac $\delta(t - t_0)$. La parte superior de la Fig. 1 muestra la entrada en la etapa de análisis 110, y la parte inferior de la Fig. 1 muestra la salida de la etapa de síntesis 120. El gráfico superior y el inferior representan el dominio de tiempo. La ventana de análisis 111 y la ventana de síntesis 121 estilizadas se ilustran como ventanas triangulares (de Bartlett). El pulso de entrada $\delta(t - t_0)$ 112 en el instante de tiempo $t = t_0$ se muestra en el gráfico superior 110 como una flecha vertical. Se supone que el bloque de transformada DFT tiene un tamaño $M = L$, es decir, el tamaño de la transformada DFT se elige para que sea igual al tamaño de las ventanas. La multiplicación de fase de las señales de subbanda por el factor T producirá el análisis DFT de un pulso de Dirac $\delta(t - Tt_0)$ en $t = Tt_0$, aunque periodizado con respecto a un tren de pulsos de Dirac de periodo L . Esto se debe a la longitud finita de la ventana aplicada y la transformada de Fourier. El tren de pulsos periodizado de periodo L se ilustra mediante las flechas discontinuas 123, 124 en el gráfico inferior.

En un sistema del mundo real, donde las ventanas de análisis y de síntesis tienen una longitud finita, el tren de pulsos solo contiene realmente algunos pulsos (dependiendo del factor de transposición), un pulso principal, es decir, el término

deseado, algunos pulsos previos y algunos pulsos posteriores, es decir, los términos no deseados. Los pulsos previos y los pulsos posteriores aparecen porque la DFT es periódica (con L). Cuando un pulso está ubicado dentro de una ventana de análisis, de modo que la fase compleja queda enrollada cuando se multiplica por T (es decir, el pulso se desplaza fuera del final de la ventana y vuelve a enrollarse al principio), se produce un pulso no deseado. Los pulsos no deseados pueden tener, o no, la misma polaridad que el pulso de entrada, dependiendo de la ubicación en la ventana de análisis y del factor de transposición.

Esto puede observarse desde un punto de vista matemático cuando se transforma el pulso de Dirac $\delta(t - t_0)$ situado en el intervalo $-L/2 < t_0 < L/2$ usando una DFT de longitud L centrada alrededor de $t = 0$,

$$X(\Omega_m) = \sum_{n=-L/2}^{L/2-1} \delta(n - t_0) \exp(-j\Omega_m n) = \exp(-j\Omega_m t_0).$$

Las señales de subbanda de análisis se multiplican en fase por un factor T para obtener las señales de subbanda de síntesis $Y(\Omega_m) = \exp(-j\Omega_m T t_0)$. Después, la DFT inversa se aplica para obtener la señal de síntesis periódica:

$$y(n) = \frac{1}{L} \sum_{m=-L/2}^{L/2-1} \exp(-j\Omega_m T t_0) \exp(j\Omega_m n) = \sum_{k=-\infty}^{\infty} \delta(n - T t_0 + kL).$$

es decir, un tren de pulsos de Dirac de periodo L.

En el ejemplo de la Fig. 1, la división en ventanas de síntesis usa una ventana finita $v_s(n)$ 121. La ventana de síntesis finita 121 capta el pulso deseado $\delta(t - T t_0)$ en $t = T t_0$ ilustrado como una flecha continua 122, y cancela las otras contribuciones que se muestran como flechas discontinuas 123, 124.

A medida que la etapa de análisis y de síntesis se desplaza a lo largo del eje de tiempo según el factor de salto o el paso de tiempo Δt , el pulso $\delta(t - t_0)$ 112 tendrá otra posición con respecto al centro de la ventana de análisis respectiva 111. Como se ha descrito anteriormente, la operación para conseguir estiramiento en el tiempo consiste en desplazar el pulso 112 a T veces su posición con respecto al centro de la ventana. Mientras esta posición esté dentro de la ventana 121, esta operación de estiramiento en el tiempo garantiza que todas las contribuciones se sumen en un único pulso sintetizado y estirado en el tiempo $\delta(t - T t_0)$ en $t = T t_0$.

Sin embargo, en la situación de la Fig. 2 se produce el problema de que el pulso $\delta(t - t_0)$ 212 se aleja hacia el borde del bloque DFT. La Fig. 2 ilustra una configuración 200 de análisis/síntesis similar a la de la Fig. 1. El gráfico superior 210 muestra la entrada en la etapa de análisis y la ventana de análisis 211, y el gráfico inferior 220 ilustra la salida de la etapa de síntesis y de la ventana de síntesis 221. Cuando se estira en el tiempo el pulso de Dirac 212 de entrada en un factor T, el pulso de Dirac 222 estirado en el tiempo, es decir, $\delta(t - T t_0)$, está fuera de la ventana de síntesis 221. Al mismo tiempo, otro pulso de Dirac 224 del tren de pulsos, es decir, $\delta(t - T t_0 + L)$ en el instante de tiempo $t = T t_0 - L$, es captado por la ventana de síntesis. Dicho de otro modo, el pulso de Dirac 212 de entrada no está retardado a un instante de tiempo T veces posterior, sino que se dirige hacia delante a un instante de tiempo anterior al pulso de Dirac 212 de entrada. El efecto final en la señal de audio es la aparición de un eco previo en una distancia de tiempo de la escala de ventanas de transponedor más bien largas, es decir, en un instante de tiempo $t = T t_0 - L$ que es $L - (T-1)t_0$ anterior al pulso de Dirac 212 de entrada.

El principio de la solución propuesta por la presente invención se describe con referencia a la Fig. 3. La Fig. 3 ilustra un escenario 300 de análisis/síntesis similar al de la Fig. 2. El gráfico superior 310 muestra la entrada en la etapa de análisis con la ventana de análisis 311, y el gráfico inferior 320 muestra la salida de la etapa de síntesis con la ventana de síntesis 321. La idea básica de la invención es adaptar el tamaño de la DFT para evitar ecos previos. Esto puede conseguirse fijando el tamaño M de la DFT de manera que la ventana de síntesis no capte imágenes de pulsos de Dirac no deseados del tren de pulsos resultante. El tamaño de la transformada DFT 301 aumenta a $M = FL$, donde L es la longitud de la función 302 de ventana y el factor F es un factor de sobremuestreo en el dominio de frecuencia. Dicho de otro modo, el tamaño de la transformada DFT 301 se selecciona para que sea mayor que el tamaño 302 de ventana. En particular, el tamaño de la transformada DFT 301 puede seleccionarse para que sea mayor que el tamaño 302 de ventana de la ventana de síntesis. Debido a la mayor longitud de la transformada DFT 301, el periodo del tren de pulsos que comprende los pulsos de Dirac 322, 324 es FL. Al seleccionar un valor suficientemente grande de F, es decir, al seleccionar un factor de sobremuestreo en el dominio de frecuencia suficientemente grande, pueden cancelarse contribuciones no deseadas al estiramiento de pulsos. Esto se muestra en la FIG. 3, donde el pulso de Dirac 324 en el instante de tiempo $t = T t_0 - FL$ está fuera de la ventana de síntesis 321. Por lo tanto, la ventana de síntesis 321 no capta el pulso de Dirac 324 y, como consecuencia, pueden evitarse ecos previos.

Debe observarse que, en una forma de realización preferida, la ventana de síntesis y la ventana de análisis tienen las mismas longitudes "nominales". Sin embargo, cuando se usa un muestreo de nuevo implícito de la señal de salida descartando o insertando muestras en las bandas de frecuencia de la transformada o banco de filtros, el tamaño de la ventana de síntesis será normalmente diferente del tamaño de análisis, dependiendo del muestreo de nuevo o del factor de transposición.

El valor mínimo de F, es decir, el factor mínimo de sobremuestreo en el dominio de frecuencia, puede deducirse a partir de la Fig. 3. La condición para no captar imágenes de pulsos de Dirac no deseados puede formularse de la siguiente

manera: Para cualquier pulso de entrada $\delta(t - t_0)$ en la posición $t = t_0 < \frac{L}{2}$, es decir, para cualquier pulso de entrada comprendido dentro de la ventana de análisis 311, la imagen no deseada $\delta(t - Tt_0 + FL)$ en el instante de tiempo $t = Tt_0 - \frac{L}{2}$.

5 FL debe estar ubicada a la izquierda del borde izquierdo de la ventana de síntesis en

Asimismo, debe satisfacerse la condición $T \frac{L}{2} - FL \leq -\frac{L}{2}$ lo que da lugar a la regla

$$F \geq \frac{T+1}{2}. \quad (3)$$

10 Como puede observarse en la fórmula (3), el factor mínimo F de sobremuestreo en el dominio de frecuencia es una función del factor T de transposición/estiramiento en el tiempo. Más específicamente, el factor mínimo F de sobremuestreo en el dominio de frecuencia es proporcional al factor T de transposición/estiramiento en el tiempo.

Repitiendo la línea de pensamiento anterior para el caso en que las ventanas de análisis y de síntesis tienen longitudes diferentes, se obtiene una fórmula más general. Sean L_A y L_S las longitudes de las ventanas de análisis y de síntesis, respectivamente, y sea M el tamaño de DFT utilizado. La regla que amplía la fórmula (3) es entonces

$$M \geq \frac{TL_A + L_S}{2}. \quad (4)$$

15 El que esta regla sea una extensión de (3) puede verificarse insertando $M = FL$, y $L_A = L_S = L$ en (4) y dividiendo por L en ambos lados de la ecuación resultante.

El anterior análisis se lleva a cabo para un modelo bastante especial de un transitorio, es decir, un pulso de Dirac. Sin embargo, el razonamiento puede extenderse para mostrar que cuando se usa el esquema de estiramiento en el tiempo descrito anteriormente, señales de entrada que tienen una envolvente espectral casi plana y que se desvanecen fuera de un intervalo de tiempo [a,b] se estirarán para formar señales de salida que son pequeñas fuera del intervalo [Ta,Tb]. También puede comprobarse estudiando espectrogramas de señales de audio y/o voz, reales en los que los ecos previos desaparecen en las señales estiradas cuando se respeta la regla antes descrita para seleccionar un factor de sobremuestreo de dominio de frecuencia apropiado. Un análisis más cuantitativo también revela que los ecos previos se reducen cuando se usan factores de sobremuestreo de dominio de frecuencia que son ligeramente inferiores al valor impuesto por la condición de la fórmula (3). Esto se debe a que funciones $v_s(n)$ de ventana típicas son pequeñas cerca de sus bordes, lo que atenúa ecos previos no deseados situados cerca de los bordes de las funciones de ventana.

En resumen, la presente invención enseña una nueva manera de mejorar la respuesta transitoria de transponedores armónicos de dominio de frecuencia, o estiradores de tiempo, introduciendo una transformada sobremuestreada, donde la cantidad de sobremuestreo depende del factor de transposición elegido.

30 A continuación, se describe en mayor detalle la aplicación de una transposición armónica según la invención en descodificadores de audio. Un caso de uso común de un transponedor armónico se produce en un sistema de códecs de audio/voz que utiliza la denomina extensión de ancho de banda o regeneración de alta frecuencia (HFR). Debe observarse que, aunque puede hacerse referencia a la codificación de audio, los procedimientos y sistemas descritos pueden aplicarse igualmente en la codificación de voz y en la codificación de voz y audio unificada (USAC).

35 En tales sistemas HFR, el transponedor puede usarse para generar una componente de señal de alta frecuencia a partir de una componente de señal de baja frecuencia proporcionada por el denominado descodificador central. La envolvente de la componente de alta frecuencia puede conformarse en tiempo y frecuencia en función de la información lateral transportada en el flujo de bits.

40 La Fig. 4 ilustra el funcionamiento de un descodificador de audio mejorado HFR. El descodificador 401 de audio central proporciona una señal de audio de bajo ancho de banda que se introduce en un muestreador ascendente 404 que puede ser necesario para producir una contribución de salida de audio final con la tasa de muestreo total deseada. Tal muestreo ascendente se requiere en sistemas de doble tasa, donde el códec de audio central de banda limitada funciona a la mitad de la tasa 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 sola tasa se omite este muestreador ascendente 404. La salida de bajo ancho de banda de 401 también se envía al transponedor o unidad de transposición 402, que proporciona una señal transpuesta, es decir, una señal que comprende el intervalo de altas frecuencias deseado. La señal transpuesta puede conformarse en tiempo y frecuencia mediante el ajustador 403 de envolvente. La salida de audio final es la suma de la señal central de bajo ancho de banda y de la señal transpuesta de envolvente ajustada.

Como se ha descrito en el contexto de la Fig. 4, la señal de salida del descodificador central puede sobremuestrearse como una etapa de preprocesamiento mediante un factor de 2 en la unidad 402 de transposición. Una transposición por un factor T da como resultado una señal que tiene T veces la longitud de la señal no transpuesta, en caso de estiramiento en el tiempo. Para conseguir el desplazamiento de tono deseado o la transposición de frecuencia a frecuencias T veces superiores, se realiza posteriormente un submuestreo o una conversión de tasa de la señal estirada en el tiempo. Como se ha mencionado anteriormente, esta operación puede conseguirse usando diferentes pasos de análisis y síntesis en el codificador de voz de fase.

El orden de transposición global puede obtenerse de diferentes maneras. Una primera posibilidad es sobremuestrear la señal de salida del descodificador mediante un factor de 2 en la entrada del transponedor, como se ha indicado anteriormente. En tales casos, la señal estirada en el tiempo necesitaría submuestrearse por un factor T con el fin de obtener la señal de salida deseada, cuya frecuencia está transpuesta por un factor T . Una segunda posibilidad sería omitir la etapa de preprocesamiento y llevar a cabo directamente las operaciones de estiramiento en el tiempo en la señal de salida del descodificador central. En tales casos, las señales transpuestas deben submuestrearse por un factor $T/2$ para mantener el factor de muestreo ascendente global de 2 y para conseguir una transposición de frecuencia por un factor T . Dicho de otro modo, el muestreo ascendente de la señal del descodificador central puede omitirse cuando se lleva a cabo un submuestreo de la señal de salida del transponedor 402 de $T/2$ en lugar de T . Sin embargo, debe observarse que la señal central sigue necesitando ser sobremuestreada en el muestreador ascendente 404 antes de combinar la señal con la señal transpuesta.

También debe observarse que el transponedor 402 puede usar diferentes factores de transposición enteros con el fin de generar la componente de alta frecuencia. Esto se muestra en la Fig. 5, que ilustra el funcionamiento de un transponedor armónico 501, que corresponde al transponedor 402 de la Fig. 4, que comprende varios transponedores de diferente orden de transposición o de diferente factor de transposición T . La señal que va a transponerse se transfiere al banco de transponedores individuales 501-2, 501-3, ..., 501- $T_{\max}$ que tienen órdenes de transposición $T=2,3,\dots,T_{\max}$, respectivamente. Normalmente, un orden de transposición $T_{\max}=4$ satisface la mayoría de aplicaciones de codificación de audio. Las contribuciones de los diferentes transponedores 501-2, 501-3, ..., 501- $T_{\max}$ se suman en 502 para proporcionar la salida de transponedor combinada. En una primera forma de realización, esta operación de suma puede comprender la adición de las contribuciones individuales. En otra forma de realización, las contribuciones se ponderan con diferentes pesos, de manera que se mitiga el efecto de añadir múltiples contribuciones a determinadas frecuencias. Por ejemplo, la contribución de tercer orden puede añadirse con una ganancia más baja que la contribución de segundo orden. Finalmente, la unidad de suma 502 puede añadir 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. 6 ilustra el funcionamiento de un transponedor armónico, tal como uno de los bloques individuales de 501, es decir, uno de los transponedores 501- T de orden de transposición T . Una unidad 601 de paso de análisis selecciona tramas sucesivas de la señal de entrada que va a transponerse. Estas tramas se superponen, por ejemplo, se multiplican, en una unidad 602 de ventana de análisis con una ventana de análisis. Debe observarse que las operaciones de seleccionar tramas de una señal de entrada y de multiplicar las muestras de la señal de entrada con una función de ventana de análisis puede llevarse a cabo en una única etapa, por ejemplo, usando una función de ventana que se desplaza a lo largo de la señal de entrada según el paso de análisis. En la unidad 603 de transformación de análisis, las tramas divididas en ventanas de la señal de entrada se transforman al dominio de frecuencia. La unidad 603 de transformación de análisis puede, por ejemplo, realizar una DFT. El tamaño de la DFT se selecciona para que sea F veces mayor que el tamaño L de la ventana de análisis, generándose así $M=F*L$ coeficientes complejos en el dominio de frecuencia. Estos coeficientes complejos se modifican en la unidad 604 de procesamiento no lineal, por ejemplo, multiplicando su fase por el factor de transposición T . La secuencia de coeficientes complejos de dominio de frecuencia, es decir, los coeficientes complejos de la secuencia de tramas de la señal de entrada pueden considerarse como señales de subbanda. La combinación de la unidad 601 de paso de análisis, la unidad 602 de ventana de análisis y la unidad 603 de transformación de análisis puede considerarse como una etapa de análisis combinada o banco de filtros de análisis.

Los coeficientes modificados o las señales de subbanda modificadas se transforman de nuevo en el dominio de tiempo usando la unidad 605 de transformación de síntesis. Para cada conjunto de coeficientes complejos modificados, esto proporciona una trama de muestras modificadas, es decir, un conjunto de M muestras modificadas. Usando la unidad 606 de ventana de síntesis pueden extraerse L muestras de cada conjunto de muestras modificadas, lo que proporciona una trama de la señal de salida. En general, una secuencia de tramas de la señal de salida puede generarse para la secuencia de tramas de la señal de entrada. Esta secuencia de tramas se desplazan entre sí por el paso de síntesis en la unidad 607 de paso de síntesis. El paso de síntesis puede ser T veces mayor que el paso de análisis. La señal de salida se genera en la unidad 608 de solapamiento y suma, donde las tramas desplazadas de la señal de salida están solapadas y las muestras del mismo instante de tiempo se suman. Al recorrer el sistema anterior, la señal de entrada puede estirarse en el tiempo por un factor T , es decir, la señal de salida puede ser una versión estirada en el tiempo de la señal de entrada.

Finalmente, la señal de salida puede contraerse en el tiempo usando la unidad 609 de contracción. La unidad 609 de contracción puede realizar una conversión de tasa de muestreo de orden T , es decir, puede aumentar la tasa de muestreo de la señal de salida en un factor T , manteniendo al mismo tiempo el número de muestras sin cambios. Esto

proporciona una señal de salida transpuesta que tiene la misma longitud en el tiempo que la señal de entrada, pero que comprende componentes de frecuencia que se han desplazado de manera ascendente en un factor T con respecto a la señal de entrada. La unidad 609 de combinación también puede realizar una operación de submuestreo por un factor T, es decir, solo puede conservar cada $T^{\text{ésima}}$ muestra, mientras las otras muestras se descartan. Esta operación de submuestreo también puede ir acompañada de una operación de filtro de paso bajo. Si la tasa de muestreo global no varía, entonces la señal de salida transpuesta comprende componentes de frecuencia que se han desplazado de manera ascendente en un factor T con respecto a las componentes de frecuencia de la señal de entrada.

Debe observarse que la unidad 609 de contracción puede llevar a cabo una combinación de conversión de tasa y de submuestreo. A modo de ejemplo, la tasa de muestreo puede aumentarse en un factor de 2. Al mismo tiempo, la señal puede submuestrearse en un factor T/2. En general, tal combinación de conversión de tasa y de submuestreo también genera una señal de salida que es una transposición armónica de la señal de entrada por un factor T. En general, puede decirse que la unidad 609 de contracción realiza una combinación de conversión de tasa y/o de submuestreo con el fin de conseguir una transposición armónica por el orden de transposición T. Esto es particularmente útil cuando se lleva a cabo una transposición armónica de la salida de bajo ancho de banda del descodificador 401 de audio central. Como se ha descrito anteriormente, tal salida de bajo ancho de banda puede haberse submuestreado por un factor de 2 en el descodificador y, por lo tanto, puede necesitar un muestro ascendente en la unidad 404 de muestreo ascendente antes de fusionarse con la componente de alta frecuencia reconstruida. Sin embargo, puede ser beneficioso para reducir la complejidad de cálculo a la hora de realizar una transposición armónica en la unidad 402 de transposición usar la salida de bajo ancho de banda "no sobremuestreada". En tales casos, la unidad 609 de contracción de la unidad 402 de transposición puede realizar una conversión de tasa de orden 2 y, de ese modo, realizar de manera implícita la operación de muestreo ascendente requerida de la componente de alta frecuencia. En consecuencia, las señales de salida transpuestas de orden T se submuestrean en la unidad 609 de contracción por el factor T/2.

En el caso de múltiples transponedores paralelos de diferentes órdenes de transposición, tales como los mostrados en la Fig. 5, algunas operaciones de transformación o de banco de filtros pueden compartirse entre diferentes transponedores 501-2, 501-3,..., 501- T_{max} . La compartición de operaciones de banco de filtros puede realizarse preferentemente para el análisis con el fin de obtener implementaciones más eficaces de las unidades 402 de transposición. Debe observarse que una manera preferida de muestrear de nuevo las salidas de diferentes transponedores consiste en descartar bins DFT o canales de subbanda antes de la etapa de síntesis. De esta manera pueden omitirse los filtros de muestreo de nuevo y reducirse la complejidad cuando se lleva a cabo una DFT inversa/ banco de filtros de síntesis de tamaño más pequeño.

Como se acaba de mencionar, la ventana de análisis puede ser común para las señales de diferentes factores de transposición. Cuando se usa una ventana de análisis común, un ejemplo del paso de ventanas 700 aplicado a la señal de banda baja se muestra en la Fig. 7. La Fig. 7 muestra un paso de ventanas de análisis 701, 702, 703 y 704, que están desplazadas entre sí por el factor de salto de análisis o el paso de tiempo de análisis Δt_a .

Un ejemplo del paso de ventanas aplicado a la señal de banda baja, por ejemplo, la señal de salida del descodificador central, se muestra en la Fig. 8(a). El paso con el que la ventana de análisis de longitud L se mueve para cada transformada de análisis se denota como Δt_a . Cada transformada de análisis de este tipo y la parte dividida en ventanas de la señal de entrada también se denominan trama. La transformada de análisis convierte/transforma la trama de señales de entrada en un conjunto de coeficientes FFT complejos. Después de la transformada de análisis, los coeficientes FFT complejos pueden transformarse de coordenadas cartesianas a coordenada polares. El conjunto de coeficientes FFT para tramas subsiguientes conforma las señales de subbanda de análisis. Para cada uno de los factores de transposición $T = 2, 3, \dots, T_{\text{max}}$ usados, los ángulos de fase de los coeficientes FFT se multiplican por el factor de transposición respectivo T y vuelven a transformarse a coordenadas cartesianas. Por tanto, habrá un conjunto diferente de coeficientes FFT complejos que representan una trama particular para cada factor de transposición T. Dicho de otro modo, para cada uno de los factores de transposición $T = 2, 3, \dots, T_{\text{max}}$ y para cada trama, se determina un conjunto independiente de coeficientes FFT. En consecuencia, para cada orden de transposición T se genera un conjunto diferente de señales $Y(t_i)$ de subbanda de síntesis.

En las etapas de síntesis, los pasos de síntesis Δt_s de las ventanas de síntesis se determinan en función del orden de transposición T usado en el transponedor respectivo. Como se ha descrito anteriormente, la operación de estiramiento en el tiempo también implica estiramiento en el tiempo de las señales de subbanda, es decir, el estiramiento en el tiempo de la sucesión de tramas. Esta operación puede realizarse eligiendo un factor de salto de síntesis o paso de síntesis Δt_s que es mayor que el paso de análisis Δt_a por un factor T. Por consiguiente, el paso de síntesis Δt_{sT} para el transponedor de orden T viene dado por $\Delta t_{sT} = T\Delta t_a$. Las Fig. 8(b) y 8(c) muestran el paso de síntesis Δt_{sT} de ventanas de síntesis para los factores de transposición $T=2$ y $T=3$, respectivamente, donde $\Delta t_{s2} = 2\Delta t_a$ y $\Delta t_{s3} = 3\Delta t_a$.

La Fig. 8 indica además el tiempo de referencia t_r que se ha "estirado" por un factor $T=2$ y $T=3$ en las Fig. 8(b) y 8(c) en comparación con la Fig. 8(a), respectivamente. Sin embargo, en las salidas este tiempo de referencia t_r tiene que alinearse para los dos factores de transposición. Para alinear la salida, la señal transpuesta de tercer orden, es decir, la Fig. 8(c), tiene que submuestrearse o su tasa tiene que convertirse con el factor 3/2. Este submuestreo da lugar a una transposición armónica con respecto a la señal transpuesta de segundo orden. La Fig. 9 ilustra el efecto del muestreo de nuevo en el paso de síntesis de ventanas para $T=3$. Si se supone que la señal analizada es la señal de salida de un

descodificador central que no se ha sobremuestreado, entonces la señal de la Fig. 8(b) se ha transpuesto en frecuencia de manera eficaz por un factor de 2 y la señal de la Fig. 8(c) se ha transpuesto en frecuencia de manera eficaz por un factor de 3.

A continuación, se aborda el aspecto del alineamiento de tiempo de secuencias transpuestas de diferentes factores de transposición cuando se usan ventanas de análisis comunes. Dicho de otro modo, se aborda el aspecto de alinear las señales de salida de los transponedores de frecuencia que utilizan un orden de transposición diferente. Cuando se usan los procedimientos descritos anteriormente, las funciones de Dirac $\delta(t - t_0)$ se estiran en el tiempo, es decir, se desplazan a lo largo del eje de tiempo, en la cantidad de tiempo dada por el factor de transposición T aplicado. Con el fin de convertir la operación de estiramiento en el tiempo en una operación de desplazamiento de frecuencia, se lleva a cabo un diezmado o submuestreo usando el mismo factor de transposición T . Si tal diezmado según el factor de transposición u orden de transposición T se lleva a cabo en la función de Dirac estirada en el tiempo $\delta(t - Tt_0)$, el pulso de Dirac submuestreado se alinearán en el tiempo con respecto al tiempo 710 de referencia cero en la parte central de la primera ventana de análisis 701. Esto se ilustra en la Fig. 7.

Sin embargo, cuando se usan diferentes órdenes de transposición T , los diezmados darán como resultado diferentes desfases para la referencia cero, a no ser que la referencia cero esté alineada con el tiempo "cero" de la señal de entrada. En consecuencia, es necesario realizar un ajuste de desfase de tiempo de las señales transpuestas diezmadas, antes de que puedan sumarse en la unidad 502 de suma. Como ejemplo, se supone un primer transponedor de orden $T = 3$ y un segundo transponedor de orden $T = 4$. Además, se supone que la señal de salida del descodificador central no está sobremuestreada. Después, el transponedor diezma la señal de tercer orden estirada en el tiempo en un factor de $3/2$ y la señal de cuarto orden estirada en el tiempo en un factor de 2. Se interpretará simplemente que la señal de segundo orden estirada en el tiempo, es decir $T = 2$, presenta una frecuencia de muestreo más alta en comparación con la señal de entrada, es decir, una frecuencia de muestreo más alta de factor 2, lo que hace que el tono de la señal de salida esté desplazado en un factor de 2.

Puede observarse que con el fin de alinear las señales transpuestas y submuestreadas, es necesario aplicar desfases de

tiempo por $\frac{(T-2)L}{4}$ a las señales transpuestas antes del diezmado, es decir, para transposiciones de tercer y cuarto orden deben aplicarse, respectivamente, desfases de $\frac{L}{4}$ y

$\frac{L}{2}$. Para verificar esto en un ejemplo concreto, se supondrá que la referencia cero para una señal de segundo orden estirada en el tiempo corresponde al instante de tiempo o muestra $\frac{L}{2}$, es decir, a la referencia cero 710 de la Fig. 7. Esto se debe a que no se ha utilizado ningún diezmado. Para una señal de tercer orden estirada en el tiempo, la referencia se convertirá en

$\frac{L}{2} \left(\frac{2}{3} \right) = \frac{L}{3}$, debido a un submuestreo por un factor de $\frac{3}{2}$. Si el desfase de tiempo según la regla antes mencionada se añade antes del diezmado, la referencia se convertirá en $\left(\frac{L}{2} + \frac{L}{4} \right) \left(\frac{2}{3} \right) = \frac{L}{2}$. Esto significa que la referencia de la señal transpuesta submuestreada está alineada con la referencia cero 710. De manera similar, para la transposición de

cuarto orden sin desfase, la referencia cero corresponde a

$$\frac{L}{2} \left(\frac{1}{2} \right) = \frac{L}{4},$$

pero cuando se usa el desfase propuesto, la referencia se convierte en $\left(\frac{L}{2} + \frac{L}{2} \right) \left(\frac{1}{2} \right) = \frac{L}{2}$, que está alineada de nuevo con la referencia cero de 2º orden 710, es decir, la referencia cero para la señal transpuesta usando $T = 2$.

Otro aspecto a tener en cuenta cuando se usan simultáneamente múltiples órdenes de transposición se refiere a las ganancias aplicadas a las secuencias transpuestas de diferentes factores de transposición. Dicho de otro modo, se aborda el aspecto de combinar las señales de salida de los transponedores de diferente orden de transposición. Hay dos principios a la hora de seleccionar la ganancia de las señales transpuestas, los cuales pueden tenerse en cuenta bajo

diferentes enfoques teóricos. En primer lugar, se supone que las señales transpuestas conservan la energía, lo que significa que se conserva la energía total en la señal de banda baja que posteriormente se transpone para formar una señal de banda alta transpuesta por un factor T. En este caso, la energía por ancho de banda debe reducirse en el factor de transposición T ya que la señal está estirada en frecuencia en la misma cantidad T. Sin embargo, las sinusoides, cuya energía está dentro de un ancho de banda infinitamente pequeño, conservarán su energía tras la transposición. Esto se debe a que de la misma manera que un pulso de Dirac se mueve en el tiempo mediante el transponedor cuando se estira en el tiempo, es decir, de la misma manera que la duración en el tiempo del pulso no varía mediante la operación de estiramiento en el tiempo, una sinusoide se desplaza en frecuencia cuando se transpone, es decir, la duración en frecuencia (dicho de otro modo, el ancho de banda) no varía por la operación de transposición de frecuencia. Es decir, aunque la energía por ancho de banda se reduzca en T, la sinusoide tiene toda su energía en un punto de frecuencia, de modo que se conservará la energía por puntos.

La otra opción a la hora de seleccionar la ganancia de las señales transpuestas es mantener la energía por ancho de banda después de la transposición. En este caso, el ruido blanco de banda ancha y los transitorios mostrarán una respuesta de frecuencia plana después de la transposición, mientras que la energía de las sinusoides aumentará en un factor T.

Un aspecto adicional de la invención es la elección de ventanas de codificador de voz de fase de análisis y de síntesis cuando se usan ventanas de análisis comunes. Resulta beneficioso elegir con cuidado las ventanas de codificador de voz de fase de análisis y de síntesis, es decir $v_a(n)$ y $v_s(n)$. No solo la ventana de síntesis $v_s(n)$ cumple la fórmula 2, antes mencionada, con el fin de permitir una reconstrucción perfecta. Además, la ventana de análisis $v_a(n)$ también debe rechazar de manera adecuada los niveles de lóbulo lateral. En caso contrario, los términos de "solapamiento" no deseados se oirán normalmente como interferencias con los términos principales para sinusoides de frecuencia variable. Tales términos de "solapamiento" no deseados también pueden aparecer para sinusoides estacionarias en el caso de factores de transposición pares, como se ha mencionado anteriormente.

La presente invención propone el uso de ventanas de seno debido a su buen porcentaje de rechazo de lóbulos laterales. Por tanto, se propone que la ventana de análisis sea

$$v_a(n) = \sin\left(\frac{\pi}{L}(n + 0.5)\right), 0 \leq n < L \quad (4)$$

Las ventanas de síntesis $v_s(n)$ serán o bien idénticas a la ventana de análisis $v_a(n)$ o vendrán dadas por la anterior fórmula (2) si al tamaño de salto de síntesis Δt_s no es un factor de la longitud L de ventana de análisis, es decir, si la longitud L de ventana de análisis no es divisible en un número entero por el tamaño de salto de síntesis. A modo de ejemplo, si $L=1024$ y $\Delta t_s = 384$, entonces $1024/384 = 2,667$ ya no es un número entero. Debe observarse que también es posible seleccionar un par de ventanas de análisis y de síntesis bi-ortogonales como las descritas anteriormente. Esto puede ser beneficioso para la reducción del solapamiento de la señal de salida, en particular cuando se usan órdenes pares T de transposición.

A continuación, se hace referencia a la Fig. 10 y a la Fig. 11, que ilustran un codificador 1000 a modo de ejemplo y un decodificador 1100 a modo de ejemplo, respectivamente, para la codificación de voz y audio unificada (USAC). La estructura general del codificador USAC 1000 y del decodificador 1100 se describe de la siguiente manera: En primer lugar, puede haber un procesamiento previo/posterior común que consiste en que una unidad funcional MPEG Surround (MPEGs) gestione un procesamiento estéreo o multicanal y que unidades 1001 y 1101 de replicación de banda espectral mejorada (eSBR), respectivamente, gestionen la representación paramétrica de las frecuencias de audio superiores en la señal de entrada y usen los procedimientos de transposición armónica descritos en el presente documento. Por tanto hay dos enfoques, uno que consiste en una trayectoria de herramienta de codificación de audio avanzada (AAC) modificada y otro que consiste en una trayectoria basada en 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 LPC residual. Todos los espectros transmitidos para AAC y LPC pueden representarse en el dominio MDCT seguidos de cuantificación y codificación aritmética. La representación en el dominio de tiempo puede usar un esquema de codificación de excitación ACELP.

La unidad 1001 de replicación de banda espectral mejorada (eSBR) del codificador 1000 puede comprender componentes de reconstrucción de alta frecuencia descritos en el presente documento. En algunas formas de realización, la unidad eSBR 1001 puede comprender una unidad de transposición descrita en el contexto de las Figs. 4, 5 y 6. Los datos codificados relacionados con la transposición armónica, por ejemplo el orden de transposición usado, la cantidad de sobremuestreo de dominio de frecuencia necesario o las ganancias utilizadas, pueden obtenerse en el codificador 1000 y fusionarse con la otra información codificada en un multiplexor de flujo de bits y reenviarse como un flujo de audio codificado a un decodificador 1100 correspondiente.

El decodificador 1100 mostrado en la Fig. 11 comprende además una unidad 1101 de replicación de ancho de banda espectral mejorada (eSBR). Esta unidad eSBR 1101 recibe el flujo de bits de audio codificado o la señal codificada desde el codificador 1000 y usa los procedimientos descritos en el presente documento para generar una componente de alta frecuencia o banda alta de la señal, que se fusiona con la componente de baja frecuencia decodificada o banda baja

para proporcionar una señal descodificada. La unidad eSBR 1101 puede comprender las diferentes componentes descritas en el presente documento. En particular, puede comprender la unidad de transposición descrita en el contexto de las Figs. 4, 5 y 6. La unidad eSBR 1101 puede usar información de la componente de alta frecuencia proporcionada por el codificador 1000 a través del flujo de bits con el fin de llevar a cabo la reconstrucción de alta frecuencia. Tal información puede ser la envolvente espectral de la componente de alta frecuencia original para generar las señales de subbanda de síntesis y, en última instancia, la componente de alta frecuencia de la señal descodificada, así como el orden de transposición usado, la cantidad de sobremuestreo de dominio de frecuencia necesario o las ganancias utilizadas.

Además, las Figs. 10 y 11 ilustran posibles componentes adicionales de un codificador/descodificador USAC, tal como:

- 10 - una herramienta de desmultiplexación de datos útiles de flujos de bits, que separa los datos útiles de flujos de bits en partes para cada herramienta, y proporciona a cada una de las herramientas información de datos útiles de flujos de bits relacionada con esa herramienta;
- 15 - una herramienta de descodificación sin ruido de factor de escala, que toma información del desmultiplexor de datos útiles de flujos de bits, analiza esa información, y descodifica los factores de escala codificados de Huffman y DPCM;
- 20 - una herramienta de descodificación sin ruido espectral, que toma información del desmultiplexor de datos útiles de flujos de bits, analiza esa información, descodifica los datos codificados de manera aritmética y reconstruye los espectros cuantificados;
- 25 - una herramienta de cuantificación inversa, 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;
- 30 - una herramienta de relleno de ruido, que se usa para llenar huecos espectrales de los espectros descodificados, que se producen cuando los valores espectrales se cuantifican a cero, por ejemplo, debido a una fuerte restricción en la demanda de bits en el codificador;
- 35 - una herramienta de doble escalado, que convierte la representación de números enteros de los factores de escala a los valores reales, y multiplica los espectros cuantificados de manera inversa no escalados por los factores de escala pertinentes;
- 40 - una herramienta M/S, como la descrita en la norma ISO/IEC 14496-3;
- 45 - una herramienta de conformación de ruido temporal (TNS), como la descrita en la norma ISO/IEC 14496-3;
- 50 - una herramienta de conmutación de banco/bloque de filtros, que aplica la inversa de la correlación de frecuencia que se llevó a cabo en el codificador; una transformada discreta del coseno modificada e inversa (IMDCT) se usa preferentemente para la herramienta de banco de filtros;
- una herramienta de conmutación de banco/bloque de filtros de distorsión de tiempo, que sustituye a la herramienta de conmutación de banco/bloque de filtros normal cuando se habilita el modo de distorsión de tiempo; el banco de filtros es preferentemente el mismo (IMDCT) que el banco de filtros normal; además, las muestras de dominio de tiempo divididas en ventanas se correlacionan desde el dominio de tiempo distorsionado al dominio de tiempo lineal mediante un doble muestreo 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 sofisticado de mezcla ascendente en la(s) señal(es) de entrada controlada(s) por parámetros espaciales apropiados; en el contexto USAC se usa preferentemente MPEGS para codificar una señal multicanal mediante la transmisión de información lateral paramétrica junto con una señal transmitida mezclada de manera descendente;
- una herramienta de clasificación 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 tratará 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 MPEG Surround, SBR mejorada, banco de filtros de distorsión de tiempo, et.;
- una herramienta de filtros LPC, que produce una señal de dominio de tiempo a partir de una señal de dominio de excitación mediante el filtrado de 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 eficiente una señal de excitación de dominio de tiempo combinando un predictor a largo plazo (palabra de código adaptativa) con una

secuencia a modo de pulsos (palabra de código de innovación).

La Fig. 12 ilustra una forma de realización de las unidades eSBR mostradas en las Figs. 10 y 11. A continuación se describirá la unidad eSBR 1200 en el contexto de un descodificador, donde la entrada en la unidad eSBR 1200 es la componente de baja frecuencia, también conocida como banda baja, de una señal.

5 En la Fig. 12, la componente 1213 de baja frecuencia se introduce en un banco de filtros QMF con el fin de 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 propósito de manipular y fusionar la componente de baja y alta frecuencia de la señal en el dominio de frecuencia, en lugar de en el dominio de tiempo. La componente 1214 de baja frecuencia se introduce en la unidad 1204 de transposición, que corresponde a los sistemas de reconstrucción de alta frecuencia descritos en el presente documento. La unidad 1204 de transposición genera una
10 componente 1212 de alta frecuencia, también denominada banda alta, de la señal, que se transforma al dominio de frecuencia mediante un banco de filtros QMF 1203. Tanto la componente de baja frecuencia transformada QMF como la componente de alta frecuencia transformada QMF se introducen en una unidad 1205 de manipulación y fusión. Esta unidad 1205 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
15 dominio de tiempo mediante un banco 1201 de filtros QMF inversos.

Normalmente, el banco 1202 de filtros QMF comprende 32 bandas de frecuencia QMF. En tales casos, la componente 1213 de baja frecuencia tiene un ancho de banda de $f_s / 4$, donde $f_s / 2$ es la frecuencia de muestreo de la señal 1213. La componente 1212 de alta frecuencia tiene normalmente un ancho de banda de $f_s / 2$ y se filtra a través del banco QMF
20 1203, que comprende 64 bandas de frecuencia QMF.

En el presente documento se ha descrito un procedimiento de transposición armónica. Este procedimiento de transposición armónica está muy adaptado particularmente a la transposición de señales transitorias. Comprende la combinación de sobremuestreo de dominio de frecuencia con la transposición armónica usando codificadores de voz. La operación de transposición depende de la combinación de la ventana de análisis, del paso de ventana de análisis, del
25 tamaño de transformada, de la ventana de síntesis, del paso de ventana de síntesis, así como de ajustes de fase de la señal analizada. Con este procedimiento pueden evitarse efectos no deseados, tales como ecos previos y posteriores. Además, el procedimiento no usa medidas de análisis de señales, tales como detección de transitorios, que normalmente introducen distorsiones de señal debido a discontinuidades en el procesamiento de las señales. Además, el procedimiento propuesto solo tiene una complejidad computacional reducida. El procedimiento de transposición
30 armónica según la invención puede mejorarse además mediante una selección apropiada de ventanas de análisis/síntesis, valores de ganancia y/o alineación de tiempo.

REIVINDICACIONES

1. Un sistema para descodificar un flujo de bits de audio codificado que comprende una señal de audio; en donde el sistema comprende una unidad (402) de transposición para transponer una señal de audio (312) de entrada por un factor T de transposición, en donde la señal de audio (312) de entrada es una componente de baja frecuencia de la señal de audio; comprendiendo la unidad (402) de transposición:
 - medios (601, 602) para extraer una trama de L muestras de dominio de tiempo de la señal de audio (312) de entrada utilizando una ventana de análisis (311) de longitud L,
 - medios (603) para transformar las L muestras de dominio de tiempo en M coeficientes complejos de dominio de frecuencia;
 - una unidad (604) de procesamiento no lineal para modificar una fase de los coeficientes complejos de dominio de frecuencia utilizando el factor T de transposición;
 - medios (605) para transformar los M coeficientes complejos de dominio de frecuencia modificados en M muestras de dominio de tiempo modificadas; y
 - un medio (606) para generar una trama de L muestras de salida de dominio de tiempo a partir de M muestras de dominio de tiempo modificadas utilizando una ventana de síntesis (321); en donde las L muestras de salida de dominio de tiempo forman una trama de una señal de salida; en donde la señal de salida es una componente de alta frecuencia de la señal de audio;
 - en donde $M=F*L$, siendo F un factor de sobremuestreo de dominio de frecuencia determinado en respuesta a la información de sobremuestreo del dominio de frecuencia recibida en el flujo de bits de audio codificado; en donde F se basa en, o es una función del factor de transposición T; y en donde el sistema está configurado para fusionar la componente de alta frecuencia con la componente de baja frecuencia descodificada.
2. El sistema de la reivindicación 1, en donde la transformación de las L muestras de dominio de tiempo en M coeficientes complejos de dominio de frecuencia comprende un relleno de ceros de la ventana de análisis (311) por $(F-1)*L$ ceros adicionales.
3. El sistema de la reivindicación 1 o 2, en el que un tamaño M de transformación de análisis se elige más grande que la longitud de la ventana L de análisis en un factor F.
4. El sistema de cualquier reclamación anterior, comprendiendo además la unidad (402) de transposición:
 - una unidad (601) de paso de análisis para desplazar la ventana de análisis mediante un paso de análisis a lo largo de la señal de audio de entrada, generando así tramas sucesivas de la señal de audio de entrada;
 - una unidad (607) de paso de síntesis para desplazar las tramas sucesivas de las L muestras de salida de dominio de tiempo en un paso de síntesis; y
 - una unidad (608) de superposición-suma para superponer y sumar las tramas sucesivas desplazadas de las L muestras de salida de dominio de tiempo, generando así la señal de salida.
5. El sistema de la reivindicación 4, comprendiendo además la unidad (402) de transposición una unidad (609) de contracción para aumentar la tasa de muestreo de la señal de salida en el orden T de transposición; produciendo por ello una señal de salida transpuesta.
6. El sistema de la reivindicación 5, en donde el paso de síntesis es T veces el paso de análisis.
7. Un procedimiento para descodificar un flujo de bits de audio codificado recibido que comprende una señal de audio; en donde el procedimiento comprende transponer una señal de audio (312) de entrada en un factor T de transposición, en donde la señal de audio (312) de entrada es una componente de baja frecuencia descodificada de la señal de audio; en donde transponer la señal de audio (312) de entrada comprende:
 - extraer una trama de L muestras de dominio de tiempo de la señal de audio (312) de entrada utilizando una ventana de análisis (311) de longitud L,
 - transformar las L muestras de dominio de tiempo en M coeficientes complejos de dominio de frecuencia,
 - modificar una fase de los coeficientes complejos de dominio de frecuencia utilizando el factor T de transposición;
 - transformar los M coeficientes complejos de dominio de frecuencia modificados en M muestras de dominio de

tiempo modificadas; y

- 5 – generar una trama de L muestras de salida de dominio de tiempo a partir de las M muestras de dominio de tiempo modificadas utilizando una ventana de síntesis (321); en donde las L muestras de salida de dominio de tiempo forman una trama de una señal de salida; en donde la señal de salida es una componente de alta frecuencia de la señal de audio; en donde $M=F*L$, siendo F un factor de sobremuestreo de dominio de frecuencia determinado en respuesta a la información de sobremuestreo del dominio de frecuencia recibida en el flujo de bits de audio codificado; en donde F se basa en, o es una función del factor T de transposición; y

en donde el procedimiento comprende además la fusión de la componente de alta frecuencia con la componente de baja frecuencia descodificada.

- 10 8. El procedimiento de la reivindicación 7, en donde la transformación de las L muestras de dominio de tiempo en M coeficientes complejos de dominio de frecuencia se lleva a cabo realizando una de una transformada de Fourier, una transformada rápida de Fourier, una transformada discreta de Fourier, una transformada de ondículas.

- 15 9. Un medio de almacenamiento que comprende un programa de software adaptado para su ejecución en un procesador y para llevar a cabo las etapas del procedimiento de la reivindicación 7 o la reivindicación 8, cuando es ejecutado en un dispositivo informático.

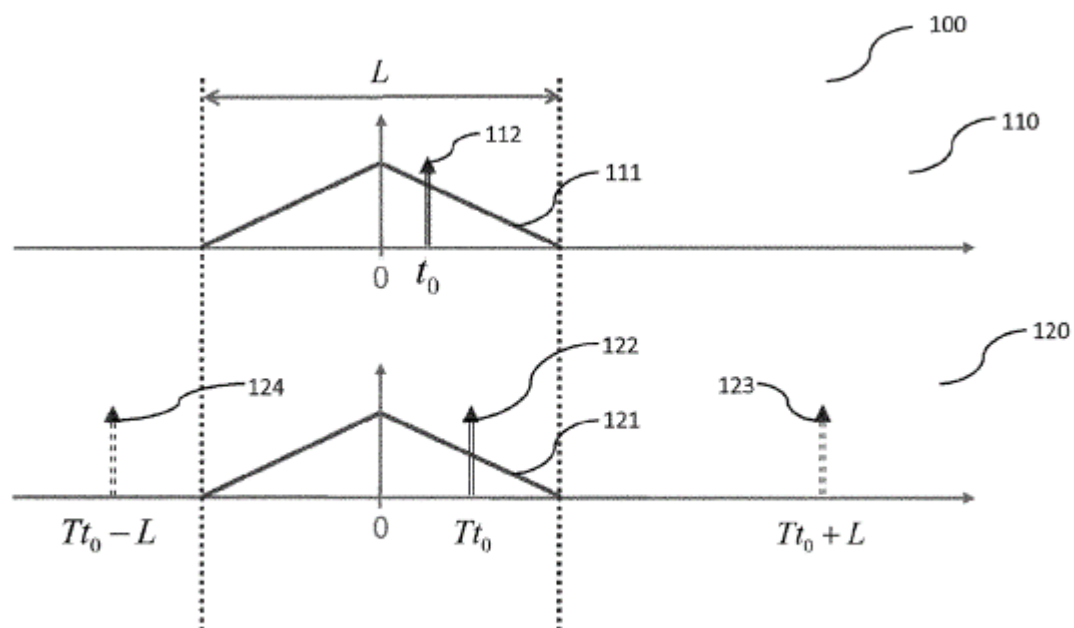


Fig. 1

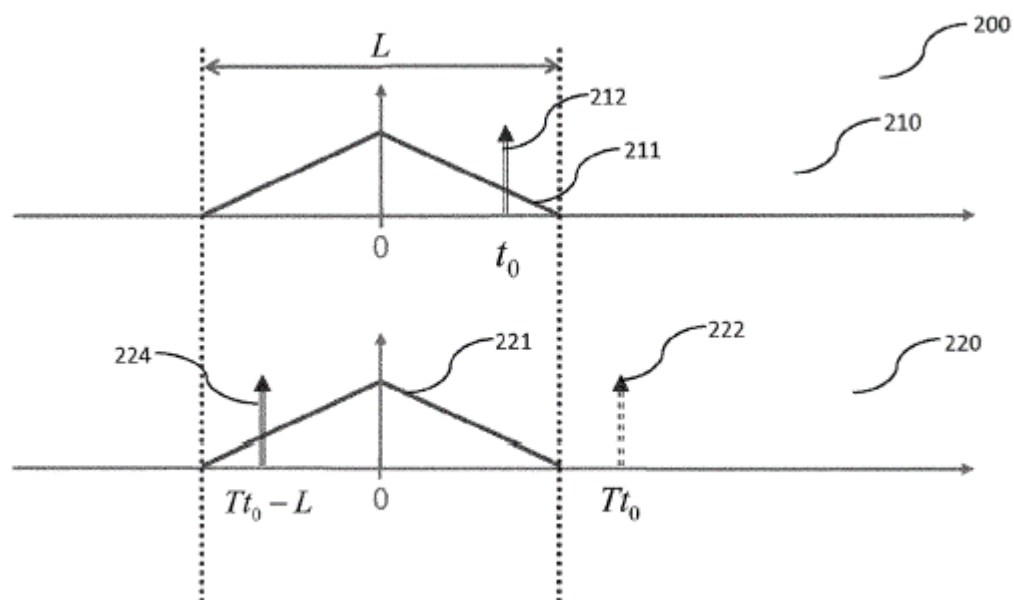


Fig. 2

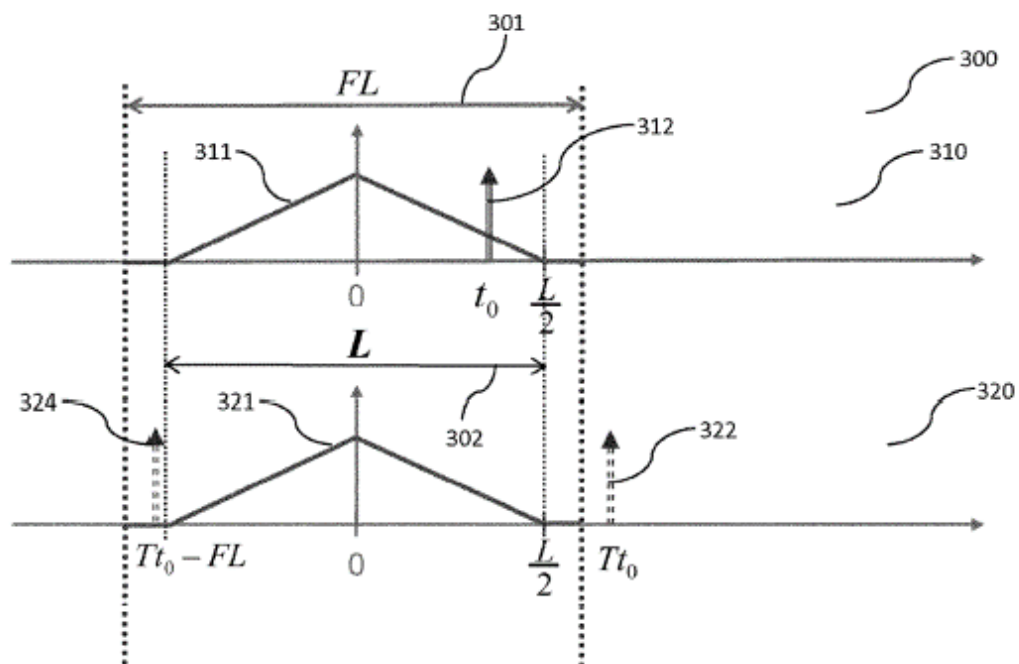


Fig. 3

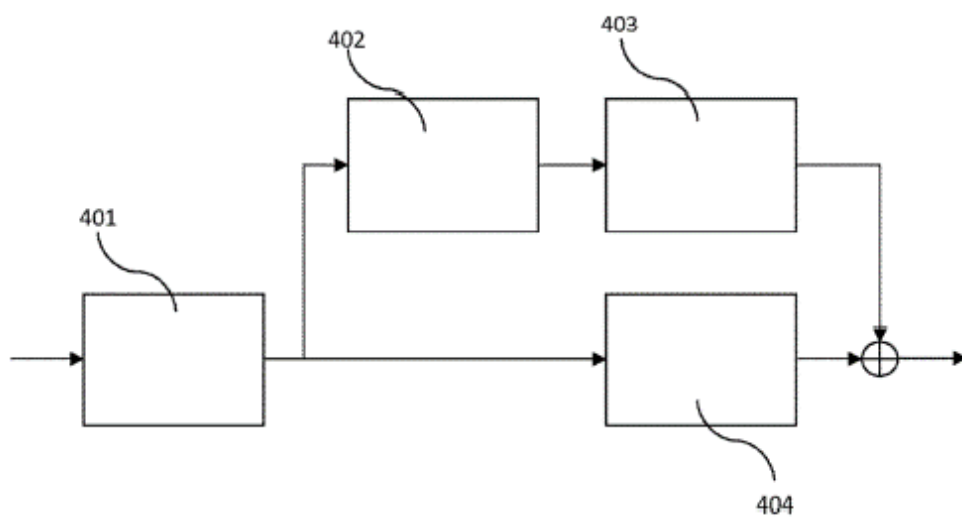


Fig. 4

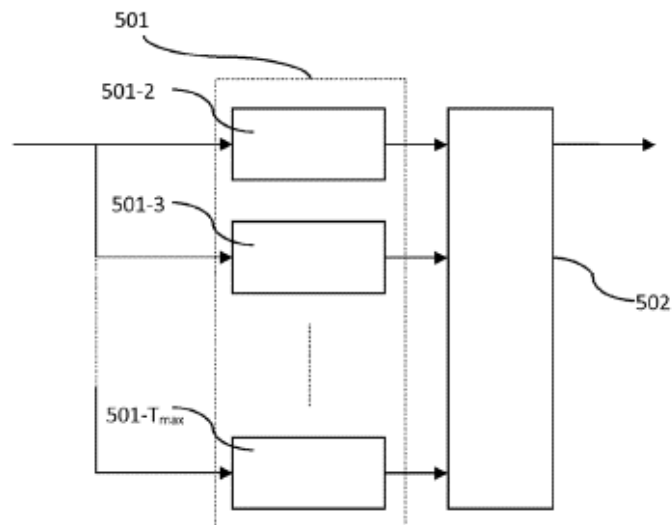


Fig. 5

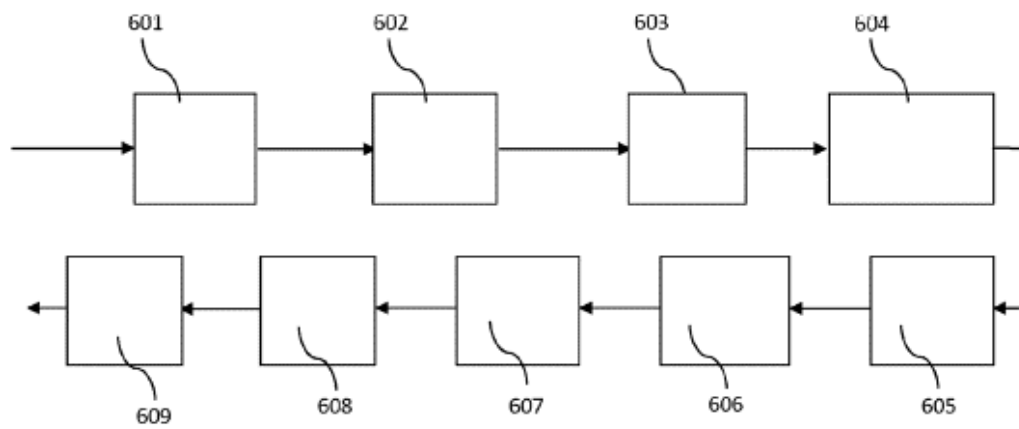


Fig. 6

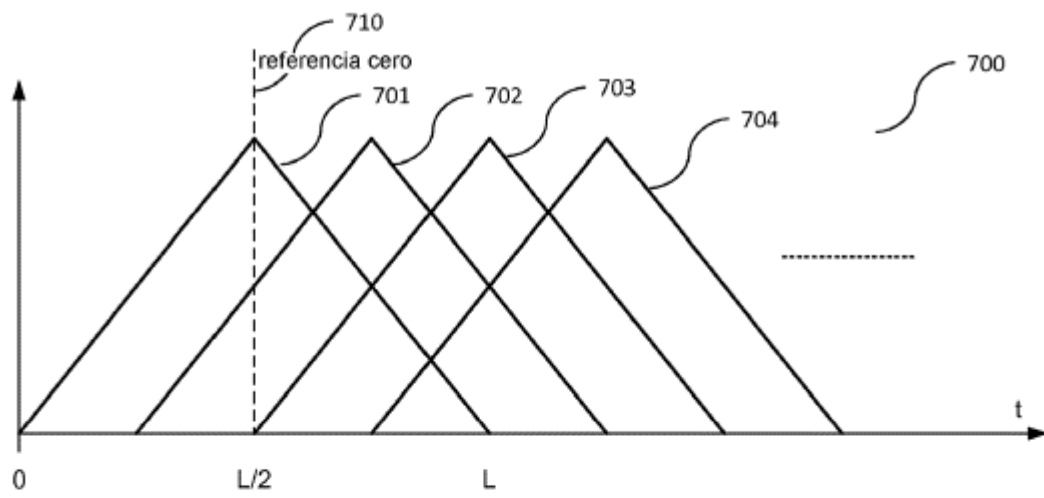
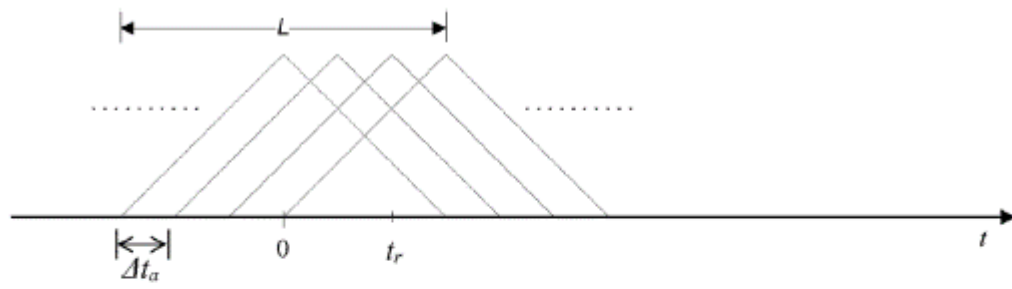
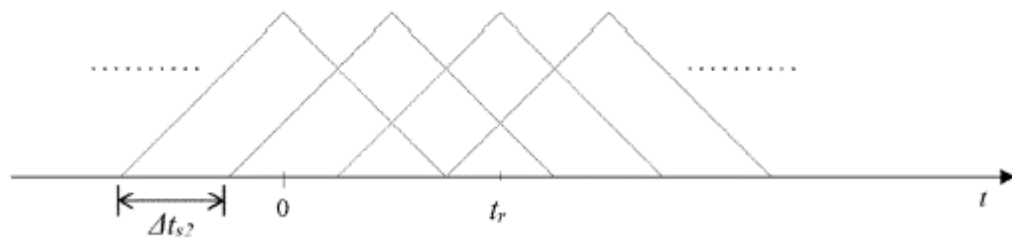


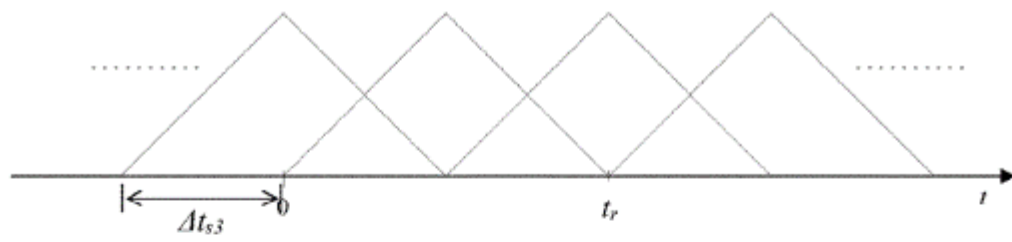
Fig. 7



(a)



(b)



(c)

Fig. 8

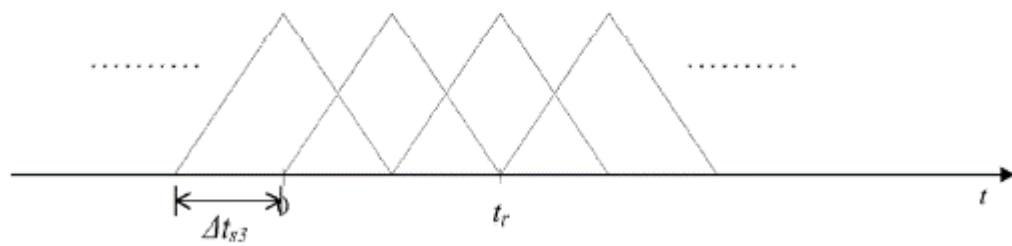


Fig. 9

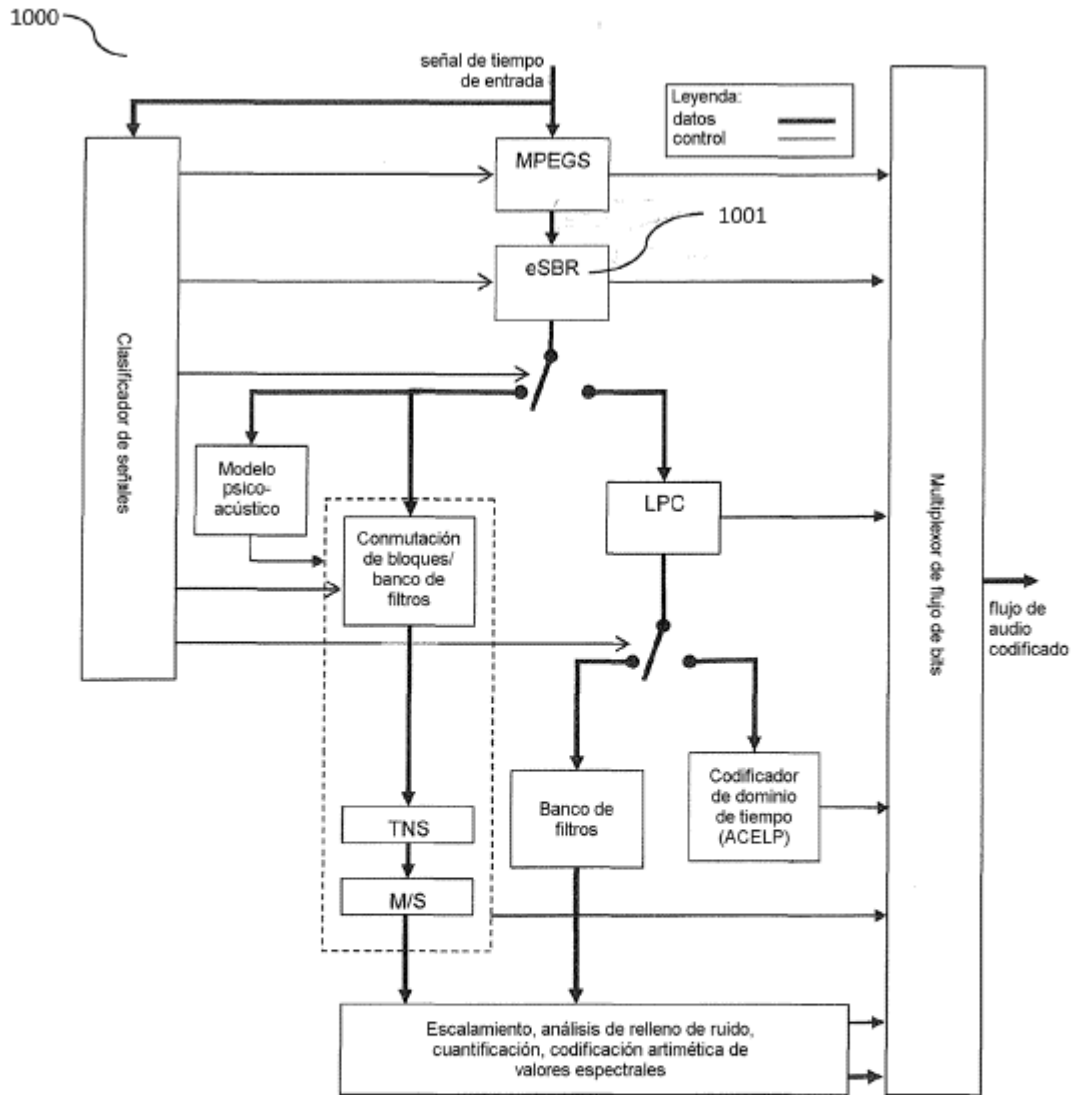


Fig. 10

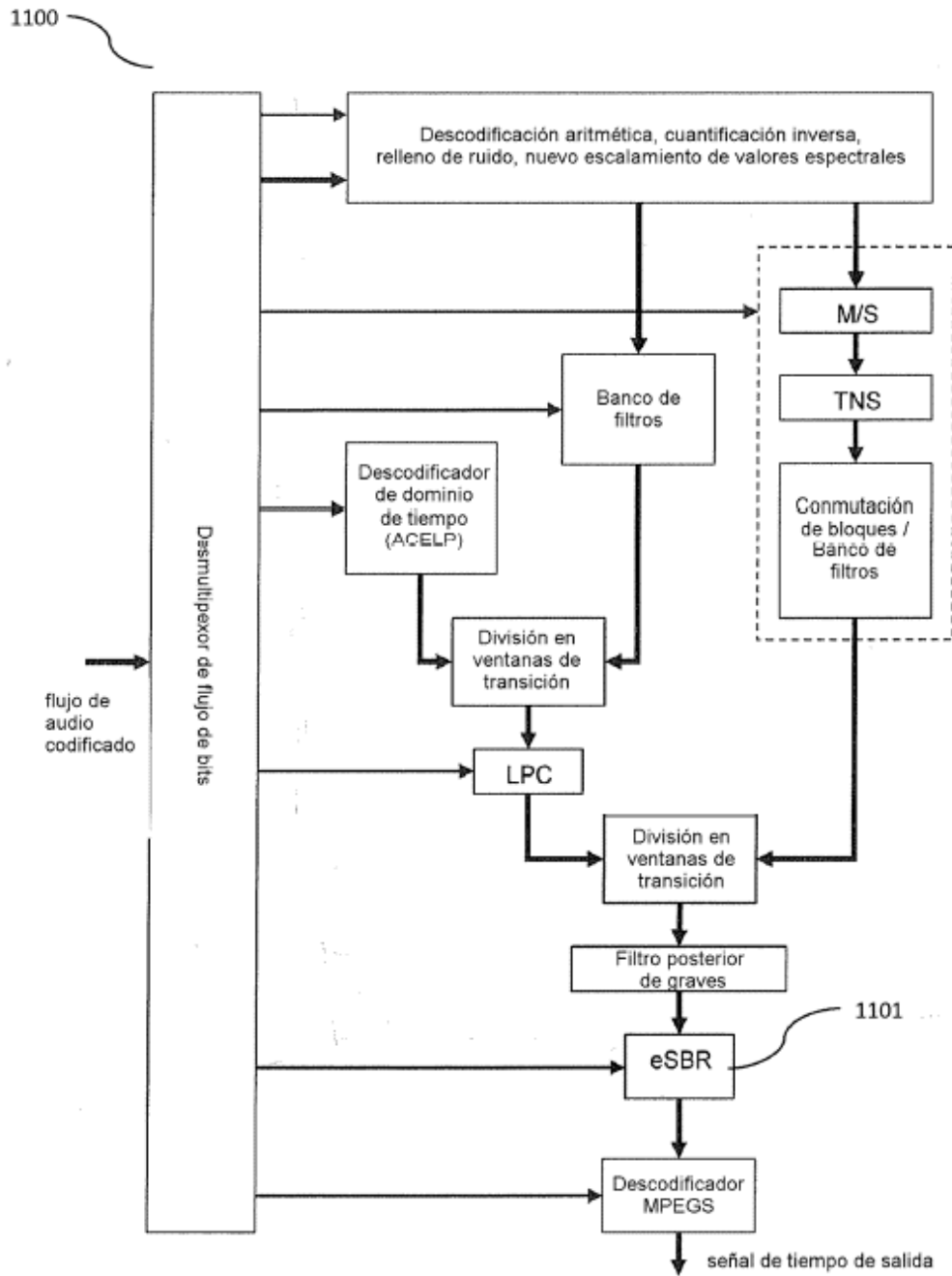


Fig. 11

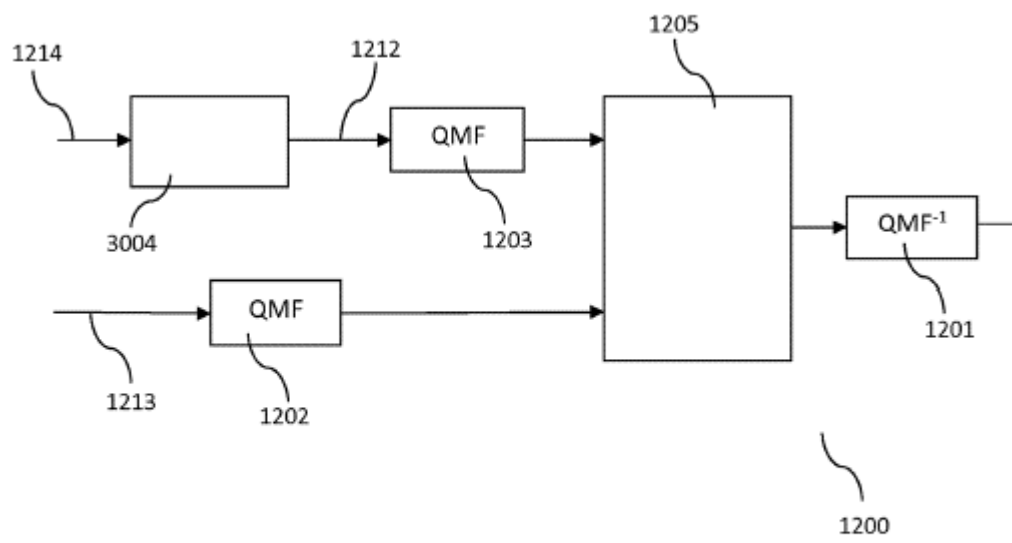


Fig. 12