

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 988 633**

51 Int. Cl.:

G10L 21/038 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **20.01.2009** **E 22183878 (2)**

97 Fecha y número de publicación de la concesión europea: **24.07.2024** **EP 4102503**

54 Título: **Dispositivo y procedimiento para una extensión de ancho de banda de una señal de audio**

30 Prioridad:

31.01.2008 US 25129 P
26.03.2008 DE 102008015702

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
21.11.2024

73 Titular/es:

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

72 Inventor/es:

NAGEL, FREDERIK;
DISCH, SASCHA y
NEUENDORF, MAX

74 Agente/Representante:

PONTI & PARTNERS, S.L.P.

ES 2 988 633 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

Dispositivo y procedimiento para una extensión de ancho de banda de una señal de audio

5 **[0001]** La presente invención se refiere al procesamiento de señales de audio y, en particular, al procesamiento de señales de audio en situaciones donde la velocidad de datos disponible es bastante pequeña, o a una extensión de ancho de banda de una señal de audio.

10 **[0002]** La codificación adaptada a la audición de señales de audio para una reducción de datos para un almacenamiento y una transmisión eficientes de estas señales ha ganado aceptación en muchos campos. Se conocen algoritmos de codificación, en particular, como "MP3" o "MP4". La codificación usada para esto, en particular cuando se obtienen las velocidades de bits más bajas, da lugar a una reducción de la calidad de audio, que es frecuentemente provocada principalmente por una limitación en el lado del codificador del ancho de banda de señales de audio que se van a transmitir.

15 **[0003]** A partir del documento WO 98 57436 se conoce el hecho de someter la señal de audio a una limitación de banda en tal situación en el lado del codificador y codificar solamente una banda inferior de la señal de audio por medio de un codificador de audio de alta calidad. Sin embargo, la banda superior se caracteriza únicamente de forma muy tosca, es decir, por un conjunto de parámetros que reproducen la envolvente espectral de la banda superior. A 20 continuación, la banda superior se sintetiza en el lado del decodificador. Para este propósito, se propone una transposición armónica, donde la banda inferior de la señal de audio decodificada se suministra a un banco de filtros. Canales de banco de filtros de la banda inferior están conectados a canales de banco de filtros de la banda superior, o son "parcheados" y cada señal de paso de banda parcheada se somete a un ajuste de envolvente. El banco de filtros de síntesis que pertenece a un banco de filtros de análisis especial recibe aquí señales de paso de banda de la señal 25 de audio en la banda inferior y señales de paso de banda ajustadas por envolvente de la banda inferior que se parchearon armónicamente a la banda superior. La señal de salida del banco de filtros de síntesis es una señal de audio extendida con respecto a su ancho de banda, que se transmitió desde el lado del codificador al lado del decodificador con una velocidad de datos muy baja. En particular, los cálculos de banco de filtros y el parcheo en el dominio de banco de filtros pueden suponer un alto esfuerzo computacional.

30 **[0004]** Procedimientos de complejidad reducida para una extensión de ancho de banda de señales de audio de banda limitada usan, en cambio, una función de copia de porciones de señal de baja frecuencia (LF) en el intervalo de altas frecuencias (HF) con el fin de aproximar información que falta debido a la limitación de banda. Tales procedimientos se describen en el documento de M. Dietz, L. Liljeryd, K. Kjörling y O. Kunz, "*Spectral Band Replication, a novel approach in audio coding*", en la 112ª Convención AES, Munich, Mayo de 2002; en el documento de S. Meltzer, R. Böhm y F. Henn, "*SBR enhanced audio codecs for digital broadcasting such as 'Digital Radio Mondiale' (DRM)*", en la 112ª Convención AES, Munich, Mayo de 2002; en el documento de T. Ziegler, A. Ehret, P. Ekstrand y M. Lutzky, "*Enhancing mp3 with SBR: Features and Capabilities of the new mp3PRO Algorithm*", en la 112ª Convención AES, Munich, Mayo de 2002; en la norma internacional ISO/IEC 14496-3:2001/FPDAM 1, "*Bandwidth Extension*", ISO/IEC, 40 2002, o en la patente de EE.UU. n.º 5.455.888, "*Speech bandwidth extension method and apparatus*", de Vasu lyengar y col.

[0005] En estos procedimientos no se realiza ninguna transposición armónica, pero señales sucesivas de paso de banda de la banda inferior se introducen en canales sucesivos de banco de filtros de la banda superior. Esto logra 45 una aproximación tosca de la banda superior de la señal de audio. A continuación, esta aproximación tosca de la señal se aproxima, en una etapa adicional, a la original mediante un posprocesamiento que utiliza la información de control obtenida de la señal original. Aquí, por ejemplo, los factores de escala sirven para adaptar la envolvente espectral, un filtrado inverso y la adición de una alfombra de ruido para adaptar la tonalidad y una suplementación por porciones de señal sinusoidal, como se describe también en la norma MPEG-4.

50 **[0006]** Aparte de esto, existen procedimientos adicionales tales como la denominada "extensión ciega de ancho de banda", descrita en el documento de E. Larsen, R.M. Aarts y M. Danessis, "*Efficient high-frequency bandwidth extension of music and speech*", en la 112ª Convención AES, Munich, Alemania, mayo de 2002, donde no se usa ninguna información del intervalo de HF original. Además, también existe el procedimiento de la denominada 55 "extensión artificial de ancho de banda", descrito en el documento de K. Käyhkö, *A Robust Wideband Enhancement for Narrowband Speech Signal; Research Report*, Universidad Tecnológica de Helsinki, Laboratorio de acústica y procesamiento de señales de audio, 2001.

[0007] En el documento de J. Makinen y col.: *AMR-WB+: a new audio coding standard for 3rd generation mobile audio services Broadcasts*, IEEE, ICASSP '05, se describe un procedimiento para la extensión de ancho de banda, 60 donde la operación de copia de la extensión de ancho de banda con una copia ascendente de señales sucesivas de paso de banda según la tecnología SBR se reemplaza por duplicación, por ejemplo, mediante muestreo ascendente.

[0008] Otras tecnologías para la extensión de ancho de banda se describen en los siguientes documentos. 65 R.M. Aarts, E. Larsen y O. Ouweltjes, "*A unified approach to low- and high frequency bandwidth extension*", 115ª

Convención AES, Nueva York, EE.UU., octubre de 2003; E. Larsen y R. M. Aarts, "Audio Bandwidth Extension - Application to psychoacoustics, Signal Processing and Loudspeaker Design", John Wiley & Sons, Ltd., 2004; E. Larsen, R.M. Aarts y M. Danessis, "Efficient high-frequency bandwidth extension of music and speech", 112ª Convención AES, Munich, mayo de 2002; J. Makhoul, "Spectral Analysis of Speech by Linear Prediction", IEEE Transactions on Audio and Electroacoustics, AU-21(3), junio de 1973; solicitud de patente de EE.UU. 08/951.029; patente de EE.UU. N.º 6.895.375.

[0009] Los procedimientos conocidos de extensión armónica de ancho de banda muestran una alta complejidad. Por otro lado, procedimientos de extensión de ancho de banda de menor complejidad muestran pérdidas de calidad. En particular, con una velocidad de bits baja y en combinación con un ancho de banda bajo del intervalo de LF, pueden producirse artefactos tales como una irregularidad y un timbre percibidos como desagradables. Una razón de esto es el hecho de que la porción de HF aproximada se basa en una operación de copia que hace que las relaciones armónicas de las porciones de señal tonal pasen desapercibidas entre sí. Esto se aplica tanto a la relación armónica entre LF y HF como también a la relación armónica dentro de la propia porción de HF. Con SBR, por ejemplo, en el límite entre el intervalo de LF y el intervalo de HF generado, ocasionalmente se producen impresiones de sonido irregulares, ya que las porciones tonales copiadas del intervalo de LF en el intervalo de HF, como se ilustra por ejemplo en la Fig. 4a, ahora pueden, en la señal global, encontrar porciones tonales del intervalo de LF como densamente adyacentes de forma espectral. Por lo tanto, en la Fig. 4a, se ilustra una señal original con picos en 401, 402, 403 y 404, mientras que una señal de prueba se ilustra con picos en 405, 406, 407 y 408. Al copiarse porciones tonales del intervalo de LF en el intervalo de HF, donde en la Fig. 4a el límite estaba en 4.250 Hz, la distancia de los dos picos izquierdos en la señal de prueba es menor que la frecuencia base subyacente al barrido de armónicos, lo que da lugar a una percepción de irregularidad.

[0010] A medida que el ancho de los grupos de frecuencia compensados por tono aumenta con un aumento de la frecuencia central, como se describe en el documento de Zwicker, E. y H. Fastl (1999), *Psychoacoustics: Facts and models*. Berlin-Springerverlag, porciones sinusoidales que se encuentran en el intervalo de LF en diferentes grupos de frecuencia, mediante copiado en el intervalo de HF, pueden llegar a estar en el mismo grupo de frecuencias aquí, lo que también da lugar a una impresión auditiva irregular como se puede observar en la Fig. 4b. Aquí se muestra, en particular, que copiar el intervalo de LF en el intervalo de HF da lugar a una estructura tonal más densa en la señal de prueba en comparación con la original. La señal original se distribuye de manera relativamente uniforme a través del espectro en el intervalo de frecuencias más altas, tal como se muestra en particular en 410. Por el contrario, en particular en este intervalo más alto, la señal de prueba 411 se distribuye de manera relativamente no uniforme a través del espectro y, por lo tanto, claramente más tonal que la señal original 410.

[0011] El documento "Audio bandwidth extension", de Erik Larsen and Ronald M. Aarts, John Wiley & Sons, 6 de diciembre de 2005, sección 6.3.4, describe un procedimiento de escalamiento de altura tonal (*pitch*), donde, duplicando la frecuencia de altura tonal, se produce una versión de la señal de excitación, que tiene un doble límite de banda superior en comparación con la señal de excitación de límite de banda. La duplicación de altura tonal comprende un muestreo descendente y un estiramiento de tiempo realizado posteriormente, y la salida del estiramiento de tiempo se introduce en un filtro de paso alto y la señal de salida del filtro de paso alto se suma a una señal de excitación compensada por retardo. La señal de excitación extendida de ancho de banda generada por la adición se introduce en un filtro de síntesis de voz correspondiente al filtro de análisis de voz para obtener una señal de voz extendida de ancho de banda.

[0012] El documento US 6.549.884 describe un procedimiento de desplazamiento de altura tonal de un vocodificador de fase. Una señal se convierte en una representación de dominio de frecuencia y, a continuación, se identifica una región específica en la representación de dominio de frecuencia. A continuación, la región se desplaza a una segunda ubicación de frecuencia para formar una representación de dominio de frecuencia ajustada, y la representación de dominio de frecuencia ajustada se transforma en una señal de dominio de tiempo que representa la señal de entrada con una altura tonal desplazada. Esto elimina la costosa etapa de remuestreo en el dominio del tiempo.

[0013] El objeto de la presente invención es lograr una extensión de ancho de banda con una alta calidad al mismo tiempo que se logra un procesamiento de señales con una menor complejidad que pueda implementarse, sin embargo, con poco retardo y poco esfuerzo y, por lo tanto, también con procesadores que tengan requisitos de hardware menos estrictos con respecto a la velocidad del procesador y la memoria requerida.

[0014] Este objeto se logra mediante un dispositivo para una extensión de ancho de banda de una señal de audio según la reivindicación 1 o un procedimiento para una extensión de ancho de banda de una señal de audio según la reivindicación 15, o un programa informático según la reivindicación 16.

[0015] El concepto inventivo para una extensión de ancho de banda se basa en un ensanchamiento de señal temporal para generar una versión de la señal de audio como una señal de tiempo que se ensancha mediante un factor de ensanchamiento > 1 y un posterior diezmado de la señal de tiempo para obtener una señal transpuesta que a continuación, en una realización, se filtra por un filtro de paso de banda simple para extraer una porción de señal de

alta frecuencia que, en una realización, solo se sigue distorsionando o cambiando con respecto a su amplitud, respectivamente, para obtener una buena aproximación para la porción de alta frecuencia original. En una realización, el filtrado de paso de banda tendrá lugar, de manera alternativa, antes de que se realice el ensanchamiento de señal, de modo que solo el intervalo de frecuencias deseada esté presente después del ensanchamiento en la señal
5 ensanchada, de modo que se pueda omitir un filtrado de paso de banda después del ensanchamiento.

[0016] Con la extensión armónica de ancho de banda, por un lado, los problemas resultantes de una operación de copia o duplicación, o ambas, se pueden prevenir según una continuación armónica y ensanchamiento del espectro utilizando el ensanchador de señales para ensanchar la señal de tiempo. Por otro lado, un ensanchamiento temporal
10 y el posterior diezmado pueden ejecutarse más fácilmente por procesadores simples que un banco de filtros de análisis/síntesis completo, como se usa, por ejemplo, con la transposición armónica, donde, adicionalmente, se deben tomar decisiones sobre cómo debe tener lugar el parcheo dentro del dominio de banco de filtros.

[0017] Preferentemente, para el ensanchamiento de señales, se utiliza un vocodificador de fase para el cual
15 hay implementaciones de menor esfuerzo. Con el fin de obtener extensiones de ancho de banda con factores > 2 , también pueden usarse varios codificadores de fase en paralelo, lo cual es ventajoso, en particular, con respecto al retardo de la extensión de ancho de banda, que tiene que ser bajo en aplicaciones en tiempo real. De manera alternativa, hay disponibles otros procedimientos para el ensanchamiento de señales, tales como, por ejemplo, el procedimiento PSOLA (*Pitch Synchronous Overlap Add*).

[0018] En una realización preferida de la presente invención, la señal de audio de LF se extiende primero en la dirección del tiempo con la frecuencia máxima $LF_{\text{máx}}$ con la ayuda del vocodificador de fase, es decir, a un múltiplo entero de la duración convencional de la señal. A continuación, en un diezmador aguas abajo, tiene lugar un diezmado de la señal mediante el factor de la extensión temporal que, en total, da lugar a un ensanchamiento del espectro. Esto
25 corresponde a una transposición de la señal de audio. Finalmente, la señal resultante se filtra mediante paso de banda con respecto al intervalo de $(\text{factor de extensión} - 1) \cdot LF_{\text{máx}}$ a $\text{factor de extensión} \cdot LF_{\text{máx}}$. De manera alternativa, las señales de alta frecuencia individuales generadas mediante ensanchamiento y diezmado, en una realización, se someterán a un filtrado de paso de banda de modo que al final se superpongan de forma aditiva a través del intervalo completa de alta frecuencia (es decir, de $LF_{\text{máx}}$ a $k \cdot LF_{\text{máx}}$). Esto es razonable en caso de que aún se desee una mayor
30 densidad espectral de armónicos.

[0019] El procedimiento de extensión armónica de ancho de banda se ejecuta en una realización preferida de la presente invención en paralelo para varios factores de extensión diferentes. Como alternativa al procesamiento en paralelo, también se puede usar un vocodificador monofásico que se hace funcionar en serie y donde los resultados
35 intermedios se almacenan en memoria intermedia. Por lo tanto, se pueden lograr frecuencias de corte de extensión de ancho de banda. De manera alternativa, la extensión de la señal también puede ejecutarse directamente en la dirección de frecuencia, es decir, en particular, mediante una operación dual correspondiente al principio funcional del vocodificador de fase.

[0020] De manera ventajosa, en realizaciones de la invención, no se requiere ningún análisis de la señal con respecto a la armonicidad o frecuencia fundamental.

[0021] A continuación, realizaciones preferidas de la presente invención se explican más en detalle con referencia a los dibujos adjuntos, donde:

45 La Fig. 1 muestra un diagrama de bloques del concepto inventivo para una extensión de ancho de banda de una señal de audio;
La Fig. 2a muestra un diagrama de bloques de un dispositivo para una extensión de ancho de banda de una señal de audio según un aspecto de la presente invención;
50 La Fig. 2b muestra una mejora del concepto de la Fig. 2a con detectores de transitorios;
La Fig. 3 muestra una ilustración esquemática del procesamiento de señales usando espectros en determinados puntos en el tiempo de una extensión de ancho de banda inventiva;
La Fig. 4a muestra una comparación entre una señal original y una señal de prueba que proporciona una impresión de sonido irregular;
55 La Fig. 4b muestra una comparación de una señal original con una señal de prueba que también da lugar a una impresión auditiva irregular;
La Fig. 5a muestra una ilustración esquemática de la implementación del banco de filtros de un vocodificador de fase;
La Fig. 5b muestra una ilustración detallada de un filtro de la Fig. 5a;
La Fig. 5c muestra una ilustración esquemática para la manipulación de la señal de magnitud y de la señal de
60 frecuencia en un canal de filtro de la Fig. 5a;
La Fig. 6 muestra una ilustración esquemática de la implementación de transformación de un vocodificador de fase;
La Fig. 7a muestra una ilustración esquemática del lado del codificador en el contexto de la extensión de ancho de banda; y
La Fig. 7b muestra una ilustración esquemática del lado del decodificador en el contexto de una extensión de ancho
65 de banda de una señal de audio.

[0022] La Fig. 1 muestra una ilustración esquemática de un dispositivo o un procedimiento, respectivamente, para una extensión de ancho de banda de una señal de audio. Solo a modo de ejemplo, la Fig. 1 se describe como un dispositivo, aunque la Fig. 1 también puede considerarse simultáneamente como el diagrama de flujo de un procedimiento para una extensión de ancho de banda. Aquí, la señal de audio se introduce en el dispositivo en una entrada 100. La señal de audio se suministra a un ensanchador de señal 102 que se implementa para generar una versión de la señal de audio como una señal de tiempo ensanchada en el tiempo mediante un factor de ensanchamiento mayor que 1. El factor de ensanchamiento en la realización ilustrada en la Fig. 1 se suministra mediante una entrada de factor de ensanchamiento 104. La señal de tiempo de audio ensanchada presente en una salida 103 del ensanchador de señales 102 se suministra a un diezmador 105 que se implementa para diezmara la señal de tiempo de audio ensanchada temporalmente 103 mediante un factor de diezmado que coincide con el factor de ensanchamiento 104. Esto se ilustra esquemáticamente mediante la entrada de factor de ensanchamiento 104 en la Fig. 1, que se traza en líneas discontinuas y conduce al diezmador 105. En una realización, el factor de ensanchamiento en el ensanchador de señales es igual a la inversa del factor de diezmado. Si, por ejemplo, se aplica un factor de ensanchamiento de 2,0 en el ensanchador de señales 102, se ejecuta un diezmado con un factor de diezmado de 0,5. Si, sin embargo, el diezmado se describe en el sentido de que se realiza un diezmado con un factor de 2, es decir, se elimina cada segundo valor de muestra, entonces, en esta ilustración, el factor de diezmado es idéntico al factor de ensanchamiento. También se pueden usar relaciones alternativas entre el factor de ensanchamiento y el factor de diezmado, por ejemplo, relaciones enteras o relaciones racionales, dependiendo de la implementación. La extensión armónica máxima del ancho de banda se logra, sin embargo, cuando el factor de ensanchamiento es igual al factor de diezmado o al inverso del factor de diezmado, respectivamente.

[0023] En una realización preferida de la presente invención, el diezmador 105 se implementa para, por ejemplo, eliminar cada segunda muestra (con un factor de ensanchamiento igual a 2) de modo que se produzca una señal de audio diezmada que tenga la misma longitud temporal que la señal de audio original 100. También se pueden usar otros algoritmos de diezmado, por ejemplo, formando valores promedio ponderados o considerando las tendencias del pasado o del futuro, respectivamente, aunque, sin embargo, se puede implementar un diezmado simple con muy poco esfuerzo mediante la eliminación de muestras. La señal de tiempo diezmada 106 generada por el diezmador 105 se suministra a un filtro 107, donde el filtro 107 se implementa para extraer una señal de paso de banda de la señal de audio diezmada 106, que contiene intervalos de frecuencia que no están contenidos en la señal de audio 100 en la entrada del dispositivo. En la implementación, el filtro 107 puede implementarse como un filtro de paso de banda digital, por ejemplo, como un filtro FIR o IIR, o también como un filtro de paso de banda analógico, aunque se prefiere una implementación digital. Además, el filtro 107 se implementa de manera que extraiga el intervalo espectral superior generado por las operaciones 102 y 105 donde, sin embargo, el intervalo espectral inferior, que de todos modos está cubierto por la señal de audio 100, se suprime tanto como sea posible. En la implementación, el filtro 107 también se puede implementar de manera que, sin embargo, también extraiga porciones de señal con frecuencias como una señal de paso de banda contenida en la señal original 100, donde la señal de paso de banda extraída contiene al menos una banda de frecuencia que no estaba contenida en la señal de audio original 100.

[0024] La señal de paso de banda 108, proporcionada por el filtro 107, se suministra a un distorsionador 109, que se implementa para distorsionar las señales de paso de banda de modo que la señal de paso de banda comprenda una envolvente predeterminada. Esta información de envolvente que se puede usar para las distorsiones se puede introducir externamente, e incluso puede proceder de un codificador o también se puede generar internamente, por ejemplo, mediante una extrapolación ciega de la señal de audio 100, o según tablas almacenadas en el lado del decodificador indexadas con una envolvente de una señal de audio 100. La señal de paso de banda distorsionada 110 proporcionada por el distorsionador 109 finalmente se suministra a un combinador 111, que está implementado para combinar la señal de paso de banda distorsionada 110 con la señal de audio original 100 que también se distorsionó dependiendo de la implementación (la fase de retardo no se indica en la Fig. 1), para generar una señal de audio extendida con respecto a su ancho de banda en una salida 112.

[0025] En una implementación alternativa, la secuencia del distorsionador 109 y el combinador 111 es inversa a la ilustración indicada en la Fig. 1. Aquí, la señal de salida de filtro, es decir, la señal de paso de banda 108, se combina directamente con la señal de audio 100, y la distorsión de la banda superior de la señal combinada que se proporciona desde el combinador 111 solo se ejecuta después de combinarse mediante el distorsionador 109. En esta implementación, el distorsionador funciona como un distorsionador para distorsionar la señal de combinación de modo que la señal de combinación comprenda una envolvente predeterminada. En esta realización, el combinador se implementa de tal manera que combina la señal de paso de banda 108 con la señal de audio 100 para obtener una señal de audio que se extiende con respecto a su ancho de banda. En esta realización, donde la distorsión solo tiene lugar después de la combinación, es preferible implementar el distorsionador 109 de manera que no influya en la señal de audio 100 o el ancho de banda de la señal de combinación, respectivamente, proporcionada por la señal de audio 100, ya que la banda inferior de la señal de audio se codificó mediante un codificador de alta calidad y se encuentra, en el lado del decodificador, en la síntesis de la banda superior, por así decirlo, la medida de todas las cosas, y no debe ser interferida por la extensión de ancho de banda.

[0026] Antes de ilustrar realizaciones detalladas de la presente invención, se ilustra un escenario de extensión

del ancho de banda con referencia a las Figs. 7a y 7b, donde la presente invención puede implementarse de manera ventajosa. Una señal de audio se introduce en una combinación de paso bajo/paso alto en una entrada 700. La combinación de paso bajo/paso alto, por un lado, incluye un paso bajo (LP), para generar una versión filtrada de paso bajo de la señal de audio 700, ilustrada en 703 en la Fig. 7a. Esta señal de audio filtrada de paso bajo está codificada con un codificador de audio 704. El codificador de audio es, por ejemplo, un codificador MP3 (capa 3 de MPEG1) o un codificador AAC, también conocido como codificador MP4 y descrito en la norma MPEG4. Codificadores de audio alternativos que proporcionan una representación transparente o, de manera ventajosa, psicoacústicamente transparente de la señal de audio limitada por banda 703 pueden usarse en el codificador 704 para generar una señal de audio 705 completamente codificada o codificada psicoacústicamente y, preferentemente, codificada psicoacústicamente de forma transparente, respectivamente. La banda superior de la señal de audio se proporciona en una salida 706 por la porción de paso alto del filtro 702, designada como "HP". La porción de paso alto de la señal de audio, es decir, la banda superior o banda de HF, también designada como la porción de HF, se suministra a un calculador de parámetros 707 que se implementa para calcular los diferentes parámetros. Estos parámetros son, por ejemplo, la envolvente espectral de la banda superior 706 en una resolución relativamente tosca, por ejemplo, mediante la representación de un factor de escala para cada grupo de frecuencias psicoacústicas o para cada banda de Bark en la escala de Bark, respectivamente. Un parámetro adicional que puede ser calculado por el calculador de parámetros 707 es la alfombra de ruido en la banda superior, cuya energía por banda puede estar relacionada, preferentemente, con la energía de la envolvente en esta banda. Parámetros adicionales que pueden ser calculados por el calculador de parámetros 707 incluyen una medida de tonalidad para cada banda parcial de la banda superior que indica cómo se distribuye la energía espectral en una banda, es decir, si la energía espectral en la banda se distribuye de manera relativamente uniforme, donde entonces existe una señal no tonal en esta banda, o si la energía en esta banda se concentra de manera relativamente intensa en una ubicación determinada en la banda, donde entonces existe, en cambio, una señal tonal para esta banda. Los parámetros adicionales consisten en la codificación explícita de picos que sobresalen de manera relativamente intensa en la banda superior con respecto a su altura y su frecuencia, ya que el concepto de extensión del ancho de banda, en la reconstrucción sin dicha codificación explícita de porciones sinusoidales prominentes en la banda superior, solo recuperará el mismo de manera muy rudimentaria, o no lo recuperará en absoluto.

[0027] En cualquier caso, el calculador de parámetros 707 se implementa para generar solo parámetros 708 para la banda superior que se pueden someter a etapas de reducción de entropía similares, ya que también se pueden realizar en el codificador de audio 704 para valores espectrales cuantificados, tales como, por ejemplo, codificación diferencial, predicción o codificación de Huffman, etc. A continuación, la representación de parámetros 708 y la señal de audio 705 se suministran a un formateador de flujo de datos 709 que se implementa para proporcionar un flujo de datos lateral de salida 710 que, típicamente, será un flujo de bits según un determinado formato tal como se normaliza, por ejemplo, en la norma MPEG4.

[0028] El lado del decodificador, que es especialmente adecuado para la presente invención, se ilustra a continuación con respecto a la Fig. 7b. El flujo de datos 710 entra en un intérprete de flujos de datos 711, que se implementa para separar la porción de parámetros 708 de la porción de señales de audio 705. La porción de parámetros 708 se decodifica mediante un decodificador de parámetros 712 para obtener parámetros decodificados 713. En paralelo a esto, la porción de señales de audio 705 se decodifica mediante un decodificador de audio 714 para obtener la señal de audio que se ilustró en 100 en la Fig. 1.

[0029] Dependiendo de la implementación, la señal de audio 100 puede proporcionarse a través de una primera salida 715. En la salida 715 se puede obtener una señal de audio con un ancho de banda pequeño y, por lo tanto, también una calidad baja. Para una mejora de la calidad, sin embargo, se realiza la extensión de ancho de banda inventiva 720, que se implementa, por ejemplo, como se ilustra en la Fig. 1, para obtener la señal de audio 112 en el lado de salida con un ancho de banda extendido o alto, respectivamente, y una alta calidad.

[0030] A continuación, con referencia a la Fig. 2a, se ilustra una implementación preferida de la implementación de extensión de ancho de banda de la Fig. 1, que puede usarse preferentemente en el bloque 712 de la Fig. 7b. La Fig. 2a incluye, en primer lugar, un bloque designado como "señal y parámetro de audio", que puede corresponder al bloque 711, 712 y 714 de la Fig. 7b, y está designado como 200. El bloque 200 proporciona la señal de salida 100 así como parámetros decodificados 713 en el lado de salida que pueden usarse para diferentes distorsiones, como por ejemplo para una corrección de tonalidad 109a y un ajuste de envolvente 109B. La señal generada o corregida, respectivamente, por la corrección de tonalidad 109a y el ajuste de envolvente 109b, se suministra al combinador 111 para obtener la señal de audio en el lado de salida con un ancho de banda extendido 112.

[0031] Preferentemente, el ensanchador de señales 102 de la Fig. 1 se implementa mediante un vocodificador de fase 202a. El diezmador 105 de la Fig. 1 se implementa preferentemente mediante un convertidor de velocidad de muestreo simple 205a. El filtro 107 para la extracción de una señal de paso de banda se implementa preferentemente mediante un filtro de paso de banda simple 107a. En particular, el vocodificador de fase 202a y el diezmador de velocidad de muestreo 205a se hacen funcionar con un factor de ensanchamiento = 2.

[0032] Preferentemente, se proporciona un "tren" adicional que consiste en el vocodificador de fase 202b, el

diezmador 205b y el filtro de paso de banda 207b para extraer una señal de paso de banda adicional en la salida del filtro 207b, que comprende un intervalo de frecuencias entre la frecuencia de corte superior del filtro de paso de banda 207a y tres veces la frecuencia máxima de la señal de audio 100.

- 5 **[0033]** Además de esto, se proporciona un vocodificador de fase k 202c que logra un ensanchamiento de la señal de audio con el factor k, donde k es preferentemente un número entero mayor que 1. Un diezmador 205 está conectado aguas abajo al vocodificador de fase 202c, que se diezma mediante el factor k. Finalmente, la señal diezmada se suministra a un filtro de paso de banda 207c que se implementa para tener una frecuencia de corte inferior que es igual a la frecuencia de corte superior de la rama adyacente y que tiene una frecuencia de corte superior
10 que corresponde a la multiplicación por k de la frecuencia máxima de la señal de audio 100. Todas las señales de paso de banda se combinan mediante un combinador 209, donde el combinador 209 puede, por ejemplo, implementarse como un sumador. De manera alternativa, el combinador 209 también puede implementarse como un sumador ponderado que, dependiendo de la implementación, atenúa las bandas superiores más intensamente que las bandas inferiores, independientemente de la distorsión aguas abajo mediante los elementos 109a, 109b. Además
15 de esto, el sistema ilustrado en la Fig. 2a incluye una fase de retardo 211 que garantiza que se produzca una combinación sincronizada en el combinador 111 que puede ser, por ejemplo, una adición de muestras.

- [0034]** La Fig. 3 muestra una ilustración esquemática de diferentes espectros que pueden producirse en el procesamiento ilustrado en la Fig. 1 o Fig. 2a. La imagen parcial (1) de la Fig. 3 muestra una señal de audio limitada por banda tal como la presente, por ejemplo, en 100 en la Fig. 1 o en 703 en la Fig. 7a. Esta señal se ensancha preferentemente por el ensanchador de señales 102 a un múltiplo entero de la duración original de la señal y, posteriormente, se diezma mediante el factor entero, lo que da lugar a un ensanchamiento general del espectro como se ilustra en la imagen parcial (2) de la Fig. 3. La porción de HF se ilustra en la Fig. 3, tal como se extrae mediante un filtro de paso de banda que comprende una banda de paso 300. En la tercera imagen parcial (3), la Fig. 3 muestra las
25 variantes donde la señal de paso de banda ya está combinada con la señal de audio original 100 antes de la distorsión de la señal de paso de banda. Por lo tanto, se obtiene un espectro de combinación con una señal de paso de banda no distorsionada, donde entonces, como se indica en la imagen parcial (4), tiene lugar una distorsión de la banda superior, pero, si es posible, no se produce ninguna modificación de la banda inferior para obtener la señal de audio 112 con un ancho de banda extendido.

- 30 **[0035]** La señal de LF en la imagen parcial (1) tiene la frecuencia máxima $LF_{\text{máx}}$. El vocodificador de fase 202a realiza una transposición de la señal de audio de modo que la frecuencia máxima de la señal de audio transpuesta sea $2LF_{\text{máx}}$. Ahora, la señal resultante en la imagen parcial (2) se filtra por paso de banda a el intervalo de $LF_{\text{máx}}$ a $2LF_{\text{máx}}$. Por lo general, cuando el factor de ensanchamiento se designa como k ($k > 1$), el filtro de paso de banda
35 comprende una banda de paso de $(k-1) \cdot LF_{\text{máx}}$ a $k \cdot LF_{\text{máx}}$. El procedimiento ilustrado en la Fig. 3 se repite para diferentes factores de ensanchamiento, hasta que se logra la frecuencia más alta deseada $k \cdot LF_{\text{máx}}$, donde k = el factor de extensión máximo $k_{\text{máx}}$.

- [0036]** A continuación, con referencia a las Figs. 5 y 6, se ilustran implementaciones preferidas para un
40 vocodificador de fase 202a, 202b, 202c según la presente invención. La Fig. 5a muestra una implementación de banco de filtros de un vocodificador de fase, donde una señal de audio se introduce en una entrada 500 y se obtiene en una salida 510. En particular, cada canal del banco de filtros esquemático ilustrado en la Fig. 5a incluye un filtro de paso de banda 501 y un oscilador aguas abajo 502. Las señales de salida de todos los osciladores de cada canal se combinan mediante un combinador que, por ejemplo, se implementa como un sumador y se indica en 503, con el fin
45 de obtener la señal de salida. Cada filtro 501 se implementa de manera que proporcione una señal de amplitud, por un lado, y una señal de frecuencia, por otro. La señal de amplitud y la señal de frecuencia son señales de tiempo que ilustran un desarrollo de la amplitud en un filtro 501 a lo largo del tiempo, mientras que la señal de frecuencia representa un desarrollo de la frecuencia de la señal filtrada por un filtro 501.

- 50 **[0037]** En la Fig. 5b se ilustra una configuración esquemática del filtro 501. Cada filtro 501 de la Fig. 5a se puede configurar como en la Fig. 5b, donde, sin embargo, solo las frecuencias f_i suministradas a los dos mezcladores de entrada 551 y al sumador 552 son diferentes de canal a canal. Las señales de salida de los mezcladores se filtran por paso bajo mediante filtros de paso bajo 553, donde las señales de paso bajo son diferentes en la medida en que fueron generadas por frecuencias de oscilador local (frecuencias LO), que están desfasadas 90° . El filtro de paso bajo
55 superior 553 proporciona una señal de cuadratura 554, mientras que el filtro inferior 553 proporciona una señal en fase 555. Estas dos señales, es decir, I y Q, se suministran a un transformador de coordenadas 556 que genera una representación de fase de magnitud a partir de la representación rectangular. La señal de magnitud o señal de amplitud, respectivamente, de la Fig. 5a a lo largo del tiempo se proporciona en una salida 557. La señal de fase se suministra a un desenvolvedor de fase 558. En la salida del elemento 558, no hay ningún valor de fase presente que
60 esté siempre entre 0 y 360° , sino un valor de fase que aumenta linealmente. Este valor de fase "desenvuelto" se suministra a un convertidor de fase/frecuencia 559 que puede, por ejemplo, implementarse como un formador de diferencia de fase simple que a una fase en un instante de tiempo actual le resta una fase de un instante de tiempo anterior para obtener un valor de frecuencia para el instante de tiempo actual. Este valor de frecuencia se suma al valor de frecuencia constante f_i del canal de filtro i para obtener un valor de frecuencia que varía temporalmente en la
65 salida 560. El valor de frecuencia en la salida 560 tiene un componente directo = f_i y un componente alterno = la

desviación de frecuencia por la cual una frecuencia de corriente de la señal en el canal del filtro se desvía de la frecuencia promedio fi.

[0038] Por lo tanto, como se ilustra en las Figs. 5a y 5b, el vocodificador de fase logra una separación de la información espectral y la información de tiempo. La información espectral está en el canal especial o en la frecuencia fi que proporciona la porción directa de la frecuencia para cada canal, mientras que la información de tiempo está contenida en la desviación de frecuencia o la magnitud en el tiempo, respectivamente.

[0039] La Fig. 5c muestra una manipulación a medida que se ejecuta para el aumento de ancho de banda según la invención, en particular, en el vocodificador de fase 202a, y en particular, en la ubicación del circuito ilustrado trazado con líneas discontinuas en la Fig. 5a.

[0040] En lo que respecta al escalamiento de tiempo, por ejemplo, las señales de amplitud $A(t)$ en cada canal o la frecuencia de las señales $f(t)$ en cada señal pueden diezmarse o interpolarse, respectivamente. A efectos de transposición, como es útil para la presente invención, se realiza una interpolación, es decir, una extensión o ensanchamiento temporal de las señales $A(t)$ y $f(t)$ para obtener las señales ensanchadas $A'(t)$ y $f'(t)$, donde la interpolación se controla mediante el factor de ensanchamiento 104, como se ilustró en la Fig. 1. Mediante la interpolación de la variación de fase, es decir, el valor antes de la adición de la frecuencia constante por el sumador 552, la frecuencia de cada oscilador individual 502 en la Fig. 5a no cambia. Sin embargo, el cambio temporal de la señal de audio global se ralentiza, es decir, por el factor 2. El resultado es un tono ensanchado temporalmente que tiene la altura tonal original, es decir, la onda fundamental original con sus armónicos.

[0041] Al realizar el procesamiento de señales ilustrado en la Fig. 5c, donde dicho procesamiento se ejecuta en cada canal de banda de filtro en la Fig. 5, y por la señal temporal resultante que, a continuación, se diezma en el diezmadore 105 de la Fig. 1 o en el diezmadore 205a en la Fig. 5a, respectivamente, la señal de audio se reduce de nuevo a su duración original mientras que todas las frecuencias se duplican simultáneamente. Esto da lugar a una transposición de altura tonal por el factor 2 donde, sin embargo, se obtiene una señal de audio que tiene la misma longitud que la señal de audio original, es decir, el mismo número de muestras.

[0042] Como alternativa a la implementación de banda de filtro ilustrada en la Fig. 5a, también se puede usar una implementación de transformación de un vocodificador de fase. En este caso, la señal de audio 100 se introduce en un procesador FFT, o más generalmente, en un procesador de transformada de tiempo corto de Fourier 600 como una secuencia de muestras de tiempo. El procesador FFT 600 se implementa esquemáticamente en la Fig. 6 para una formación de ventana de tiempo de una señal de audio para, a continuación, mediante una FFT, calcular tanto un espectro de magnitud como también un espectro de fase, donde este cálculo se realiza para espectros sucesivos que están relacionados con bloques de la señal de audio, que están sumamente superpuestos.

[0043] En un caso extremo, para cada nueva muestra de señal de audio se puede calcular un nuevo espectro, donde también se puede calcular un nuevo espectro, por ejemplo, solo para cada vigésima muestra nueva. Esta distancia a en muestras entre dos espectros viene dada preferentemente por un controlador 602. El controlador 602 se implementa además para alimentar un procesador IFFT 604 que está implementado para funcionar en una operación de superposición. En particular, el procesador IFFT 604 está implementado de manera que realice una transformada de Fourier inversa de tiempo corto mediante la realización de una IFFT por espectro según un espectro de magnitud y un espectro de fase, con el fin de realizar entonces una operación de suma superpuesta, de la cual se obtiene un intervalo de tiempo. La operación de suma superpuesta elimina los efectos de la ventana de análisis.

[0044] Un ensanchamiento de la señal de tiempo se logra mediante la distancia b entre dos espectros, cuando son procesados por el procesador IFFT 604, siendo mayor que la distancia a entre los espectros en la generación de los espectros FFT. La idea básica es ensanchar la señal de audio mediante FFT inversas simplemente más espaciadas que las FFT de análisis. Como resultado, los cambios espectrales en la señal de audio sintetizada se producen más lentamente que en la señal de audio original.

[0045] Sin embargo, sin un reescalado de fase en el bloque 606, esto produciría artefactos de frecuencia. Cuando, por ejemplo, se considera un único contenedor de frecuencia para el cual se implementan valores de fase sucesivos en 45° , esto implica que la señal dentro de esta banda de filtro aumenta en la fase con una velocidad de $1/8$ de un ciclo, es decir, en 45° por intervalo de tiempo, donde el intervalo de tiempo es aquí el intervalo de tiempo entre FFT sucesivas. Si ahora las FFT inversas se están espaciando más lejos entre sí, esto significa que el aumento de fase de 45° se produce a lo largo de un intervalo de tiempo más largo. Esto significa que la frecuencia de esta porción de señal se redujo involuntariamente. Para eliminar esta reducción de frecuencia de artefacto, la fase se reescala exactamente por el mismo factor por el cual la señal de audio se ensanchó en el tiempo. La fase de cada valor espectral de FFT se incrementa así mediante el factor b/a , de modo que se elimina esta reducción de frecuencia no intencionada.

[0046] Mientras que en la realización ilustrada en la Fig. 5c el ensanchamiento por interpolación de las señales de control de amplitud/frecuencia se logró para un oscilador de señales en la implementación del banco de filtros de la Fig. 5a, el ensanchamiento en la Fig. 6 se logra por la distancia entre dos espectros IFFT que es mayor que la

distancia entre dos espectros FFT, es decir, b es mayor que a , donde, sin embargo, para una prevención de artefactos se ejecuta un reescalado de fase según b/a .

[0047] Con respecto a una descripción detallada de los vocodificadores de fase, se hace referencia a los siguientes documentos:

- 5 "The Phase Vocoder: A tutorial", de Mark Dolson, *Computer Music Journal*, vol. 10, n.º 4, págs. 14 a 27, 1986, o "New phase Vocoder techniques for pitch-shifting, harmonizing and other exotic effects", de L. Laroche y M. Dolson, *Proceedings 1999 IEEE Workshop on applications of signal processing to audio and acoustics*, New Paltz, Nueva York, del 17 al 20 de octubre de 1999, págs. 91 a 94; "New approach to transient processing interphase vocoder", de A. Robel, *Proceeding of the 6th international conference on digital audio effects (DAFx-03)*, Londres (Reino Unido), del 8 al 11 de septiembre de 2003, págs. DAFx-1 a DAFx-6; "Phase-locked Vocoder", Meller Puckette, *Proceedings 1995, IEEE AS- SP, Conference on applications of signal processing to audio and acoustics* o la solicitud de patente de EE.UU. n.º 6.549.884.

- 15 [0048] La Fig. 2b muestra una mejora del sistema ilustrado en la Fig. 2a, donde se utiliza un detector de transitorios 250 que se implementa para determinar si una operación temporal actual de la señal de audio contiene una porción transitoria. Una porción transitoria consiste en el hecho de que la señal de audio cambia mucho en total, es decir, que, por ejemplo, la energía de la señal de audio cambia en más de un 50 % de una porción temporal a la siguiente porción temporal, es decir, aumenta o disminuye. Sin embargo, el umbral del 50 % es solo un ejemplo y
20 también puede ser valores más pequeños o más grandes. De manera alternativa, para una detección de transitorios, también se puede considerar el cambio de la distribución de energía, por ejemplo, en la conversión de vocal a sibilante.

- [0049] Si se determina una porción transitoria de la señal de audio, se deja la transposición armónica, y para el intervalo de tiempo de transitorios, se pasa a ejecutar una operación de copia no armónica o una duplicación no armónica o algún otro algoritmo de extensión de ancho de banda, tal como se ilustra en 260. A continuación, si se detecta de nuevo que la señal de audio ya no es transitoria, se realiza de nuevo una transposición armónica, como se ilustra mediante los elementos 102, 105 en la Fig. 1. Esto se ilustra en 270 en la Fig. 2b.

- [0050] Las señales de salida de los bloques 270 y 260 que llegan desfasadas en el tiempo debido al hecho de que una porción temporal de la señal de audio puede ser transitoria o no transitoria, se suministran a un combinador 280 que está implementado para proporcionar una señal de paso de banda en el tiempo que puede, por ejemplo, suministrarse a la corrección de tonalidad en el bloque 109a en la Fig. 2a. De manera alternativa, la combinación en el bloque 280 también puede realizarse, por ejemplo, después del sumador 111. Sin embargo, esto significaría que para un bloque de transformación completo de la señal de audio se supone una característica transitoria, o si la implementación del banco de filtros también funciona en base a bloques, para un bloque completo se toma una decisión a favor de transitorios o no transitorios, respectivamente.

- [0051] Puesto que un vocodificador de fase 202a, 202b, 202c, como se ilustra en la Fig. 2a y se explica con más detalle en las Figs. 5 y 6, genera más artefactos en el procesamiento de porciones de señales transitorias que en el procesamiento de porciones de señales no transitorias, se realiza un cambio a una operación de copia no armónica o duplicación, como se ilustró en la Fig. 2b en 260. De manera alternativa, también se puede realizar un reajuste de fase con respecto al transitorio, tal como se describe, por ejemplo, en la publicación de expertos de Laroche citada anteriormente, o en la patente de EE.UU. n.º 6.549.884.

- 45 [0052] Como ya se ha indicado, en los bloques 109a, 109b, después de la generación de la porción de HF del espectro, se realiza una formación espectral y un ajuste a la medida original de ruido. La formación espectral puede tener lugar, por ejemplo, con la ayuda de factores de escala, factores de escala ponderados en dB(A) o una predicción lineal, donde la predicción lineal tiene la ventaja de que no se requiere conversión de tiempo/frecuencia y no se requiere una conversión de frecuencia/tiempo posterior.

- 50 [0053] La presente invención es ventajosa en la medida en que mediante el uso del vocodificador de fase, un espectro con una frecuencia creciente se ensancha aún más y siempre continúa correctamente de manera armónica mediante el ensanchamiento de números enteros. Por lo tanto, se excluye el resultado tosco en la frecuencia de corte de el intervalo de LF y se evitan interferencias por porciones de HF muy densamente ocupadas del espectro. Además, se pueden usar implementaciones eficaces del vocodificador de fase, que se pueden realizar sin operaciones de parcheo de banco de filtros.

- [0054] De manera alternativa, hay disponibles otros procedimientos para el ensanchamiento de señales, tales como, por ejemplo, el procedimiento PSOLA (*Pitch Synchronous Overlap Add*). *Pitch Synchronous Overlap Add*, abreviado como PSOLA, es un procedimiento de síntesis donde las grabaciones de señales de voz están ubicadas en la base de datos. En la medida en que se trate de señales periódicas, estas se proporcionan con información acerca de la frecuencia fundamental (altura tonal) y se marca el inicio de cada periodo. En la síntesis, estos periodos se cortan con un entorno determinado por medio de una función de ventana y se añaden a la señal que se va a sintetizar en una ubicación adecuada. Dependiendo de si la frecuencia fundamental deseada es mayor o menor que la de la entrada
65 de la base de datos, se combinan en consecuencia de manera más densa o menos densa que en la original. Para

ajustar la duración de lo audible, los períodos pueden omitirse o proporcionarse doblemente. Este procedimiento también se denomina TD-PSOLA, donde TD significa dominio del tiempo y enfatiza que los procedimientos funcionan en el dominio del tiempo. Otro desarrollo es el procedimiento *MultiBand Resynthesis OverLap Add*, abreviado como MBROLA. Aquí los segmentos en la base de datos son llevados a una frecuencia fundamental uniforme mediante un preprocesamiento y la posición de fase del armónico se normaliza. Por ello, en la síntesis de una transición de un segmento al siguiente, se producen interferencias menos perceptivas y la calidad de la voz conseguida es mayor.

[0055] En una alternativa adicional, la señal de audio ya se ha filtrado por paso de banda antes de ensancharse, de modo que la señal después del ensanchamiento y el diezmado ya contiene las porciones deseadas y el filtrado por paso de banda posterior se omite en dicha realización. En este caso, el filtro de paso de banda se establece de modo que la porción de la señal de audio que se habría filtrado después de la extensión del ancho de banda todavía esté contenida en la señal de salida del filtro de paso de banda. Por lo tanto, el filtro de paso de banda contiene un intervalo de frecuencia que no está contenido en la señal de audio después del ensanchamiento y el diezmado. La señal con este intervalo de frecuencias es la señal deseada que forma la señal de alta frecuencia sintetizada. En esta realización, el distorsionador 109 no distorsionará una señal de paso de banda, sino una señal ensanchada y diezmada derivada de una señal de audio filtrada por paso de banda.

[0056] Cabe señalar, además, que la señal ensanchada también puede ser útil en el intervalo de frecuencias de la señal original, por ejemplo, mezclando la señal original y la señal ensanchada, por lo que no se requiere una banda de paso "estricta". La señal ensanchada puede entonces mezclarse con la señal original en la banda de frecuencias donde se superpone con la señal original con respecto a la frecuencia, para modificar la característica de la señal original en el intervalo de superposición.

[0057] Cabe señalar además que las funcionalidades de distorsión 109 y filtrado 107 se pueden implementar en un solo bloque de filtros o en dos filtros separados en cascada. Puesto que la distorsión tiene lugar dependiendo de la señal, la característica de amplitud de este bloque de filtros será variable. Su característica de frecuencia es, sin embargo, independiente de la señal.

[0058] Dependiendo de la implementación, como se ilustra en la Fig. 1, primero la señal de audio global se ensancha, se diezma y, a continuación, se filtra en una realización, donde el filtrado corresponde a las operaciones de los elementos 107, 109. Por lo tanto, la distorsión se ejecuta después de o simultáneamente al filtrado, donde para este propósito es adecuado un bloque combinado de filtros/distorsionadores en forma de un filtro digital. De manera alternativa, antes del filtrado (de paso de banda) (107) se produce una distorsión en una realización cuando se utilizan dos elementos de filtro diferentes.

[0059] De nuevo, de manera alternativa, un filtrado de paso de banda tiene lugar en una realización antes del ensanchamiento de modo que solo la distorsión (109) sigue después del diezmado. Para estas funciones se prefieren aquí dos elementos diferentes.

[0060] De nuevo en una realización alternativa, además de en todas las variantes anteriores, la distorsión tiene lugar después de la combinación de la señal de síntesis con la señal de audio original, tal como, por ejemplo, con un filtro que no tiene, o solo tiene muy poco efecto, en la señal que se va a filtrar en el intervalo de frecuencias del filtro original, que, sin embargo, genera la envolvente deseada en el intervalo de frecuencias extendido. En este caso, de nuevo se utilizan preferentemente dos elementos diferentes para la extracción y la distorsión.

[0061] El concepto inventivo es adecuado para todas las aplicaciones de audio donde el ancho de banda completo no está disponible. El concepto inventivo se puede utilizar en la propagación de contenidos de audio mediante, por ejemplo, radio digital, transmisión por Internet y aplicaciones de comunicación de audio.

[0062] Dependiendo de las circunstancias, el procedimiento inventivo para una extensión de ancho de banda de una señal de audio puede implementarse en hardware o en software. La implementación puede ejecutarse en un medio de almacenamiento digital, en particular un disco flexible o un CD, que tenga señales de control legibles electrónicamente almacenadas en el mismo, que pueden actuar conjuntamente con el sistema informático programable, de tal manera que se realice el procedimiento. Por lo tanto, la invención consiste, en general, en un producto de programa informático con un código de programa para ejecutar el procedimiento almacenado en un soporte legible por máquina, cuando el producto de programa informático se ejecuta en un ordenador. En otras palabras, la invención puede realizarse por tanto como un programa informático con un código de programa para llevar a cabo el procedimiento cuando el programa informático se ejecuta en un ordenador.

REIVINDICACIONES

1. Un dispositivo para una extensión de ancho de banda de una señal de audio (100, 703), que comprende:
 - 5 un ensanchador de señales (102) para generar una versión de la señal de audio (100, 703) como una señal de tiempo ensanchada en el tiempo mediante un factor de ensanchamiento (104) mayor que 1 para obtener una versión ensanchada temporalmente (103) de la señal de audio (100, 703);
un diezmador (105) para diezmarse la versión ensanchada temporalmente (103) de la señal de audio (100, 703) mediante un factor de diezmado que coincide con el factor de ensanchamiento (104) para obtener una señal de
10 audio diezmada (106)
un filtro (107) para extraer

una señal distorsionada de la señal de audio diezmada (106) para obtener una señal extraída distorsionada, conteniendo la señal extraída distorsionada un intervalo de frecuencias que no está contenida en la señal de
15 audio (100, 703), o
una señal no distorsionada de la señal de audio diezmada (106) para obtener una señal extraída no distorsionada (108), conteniendo la señal extraída no distorsionada un intervalo de frecuencias que no está contenido en la señal de audio (100, 703), o
una señal de la señal de audio (100, 703) antes de un ensanchamiento por el ensanchador de señales (102)
20 para obtener una señal extraída no ensanchada, donde la señal extraída no ensanchada, posterior a un procesamiento por el ensanchador de señales (102) y el diezmador (105), contiene un intervalo de frecuencias que no está contenido en la señal de audio (100, 703), y

un combinador (111) para combinar
25 la señal extraída distorsionada y la señal de audio (100, 703) para obtener una señal de audio (112) extendida en su ancho de banda, o
una señal distorsionada (110) generada a partir de la señal extraída no distorsionada (108) o la señal extraída no ensanchada posterior al procesamiento por el ensanchador de señales (102) y el diezmador (105) por medio
30 de distorsión (109) y la señal de audio (100, 703) para obtener una señal de audio (112) extendida en su ancho de banda, o
la señal extraída no distorsionada (108) y la señal de audio (100, 703) para obtener una señal de combinación, donde una señal de audio generada distorsionando la señal de combinación representa una señal de audio
35 (112) extendida en su ancho de banda,

donde la señal extraída distorsionada, o la señal distorsionada (110), o la señal de audio generada distorsionando la señal de combinación se distorsiona (109) de modo que la señal extraída distorsionada, o la señal distorsionada (110), o la señal de audio generada distorsionando la señal de combinación comprende una envolvente
40 predeterminada.
 2. El dispositivo según la reivindicación 1, donde el ensanchador de señales (102) se implementa para usar como factor de ensanchamiento (104) un factor de ensanchamiento entero mayor que 1, y donde el diezmador (105) se implementa para tomar un factor de diezmado igual a o inverso al factor de ensanchamiento (104).
 - 45 3. El dispositivo según la reivindicación 1 o 2, donde el ensanchador de señales (102) se implementa para ensanchar la señal de audio (100, 703) de modo que un tono de la señal de audio (100, 703) no cambie.
 4. El dispositivo según una de las reivindicaciones anteriores, donde el ensanchador de señales (102) se
50 implementa para ensanchar la señal de audio (100, 703) de forma que una duración temporal de la señal de audio (100, 703) se incremente y que un ancho de banda de la versión ensanchada temporalmente (103) de la señal de audio (100, 703) sea igual a un ancho de banda de la señal de audio (100, 703).
 5. El dispositivo según una de las reivindicaciones anteriores, donde el ensanchador de señal (102)
55 comprende un vocodificador de fase (202a, 202b, 202c).
 6. El dispositivo según la reivindicación 5, donde el vocodificador de fase (202a, 202b, 202c) se implementa en una implementación de banco de filtros o en una implementación de transformada de Fourier.
 - 60 7. El dispositivo según una de las reivindicaciones anteriores, donde el ensanchador de señales (102) está implementado para ensanchar la señal de audio (100, 703) mediante el factor de ensanchamiento igual a 2 para obtener una primera señal ensanchada correspondiente a la versión ensanchada temporalmente (103) de la señal de audio (100, 703),

65 donde está presente un ensanchador de señales adicional (202b), que está implementado para ensanchar la señal

- de audio (100, 703) mediante un factor de ensanchamiento adicional de 3 para obtener una segunda señal ensanchada,
 donde el diezmador (105) está implementado para diezmar la primera señal ensanchada mediante el factor de ensanchamiento de 2 para obtener la señal de audio diezmada (106),
 5 donde está presente un diezmador adicional (205b) que se implementa para diezmar la segunda señal ensanchada mediante el factor de ensanchamiento adicional de 3 para obtener una segunda señal diezmada,
 donde el filtro (107) está implementado para filtrar una primera banda recién generada en la señal de audio diezmada (106) proporcionada por el diezmador (105), correspondiendo la primera banda a la señal no distorsionada (108), o para ejecutar un filtrado antes del ensanchamiento por el ensanchador de señales (102)
 10 para obtener la señal extraída no ensanchada,
 donde existe un segundo filtro de paso de banda (207b) para extraer una segunda banda de la segunda señal diezmada que es nueva con respecto a la primera señal diezmada o para ejecutar un filtrado antes del ensanchamiento por el ensanchador de señales adicional (202b), y
 donde el combinador (209) está configurado para sumar señales extraídas correspondientes a la primera banda y
 15 la segunda banda o para distorsionar las señales extraídas correspondientes a la primera banda y la segunda banda y para sumar señales extraídas distorsionadas.
8. El dispositivo según la reivindicación 7, donde está presente un grupo adicional de un vocodificador de fase adicional (202c), un diezmador aguas abajo (205c) y un filtro de paso de banda aguas abajo (207c) que se
 20 establecen en un factor de ensanchamiento adicional (k) para generar una señal de paso de banda adicional que se suministra al sumador (209).
9. El dispositivo según una de las reivindicaciones anteriores,
- 25 donde el ensanchador de señales (102) está implementado para proporcionar la versión ensanchada temporalmente (103) de la señal de audio (100, 703) como una señal de tiempo como una secuencia de muestras que tiene el ancho de banda completo de la señal de audio (100, 703), y
 donde el diezmador (105) está implementado para obtener la secuencia de muestras como una señal de entrada y para diezmar la misma para obtener la señal de audio diezmada (106).
 30
10. El dispositivo según una de las reivindicaciones anteriores, que comprende un distorsionador (109), donde el distorsionador (109) se implementa para ejecutar la distorsión según los parámetros transmitidos (713).
11. El dispositivo según una de las reivindicaciones anteriores, que comprende, además:
 35 un detector de transitorios (250) implementado para controlar el ensanchador de señales (102) o el diezmador (105), cuando se detecta una porción transitoria en la señal de audio (100, 703) para ejecutar (260) una manera alternativa de generar porciones espectrales más altas.
- 40 12. El dispositivo según una de las reivindicaciones anteriores, que comprende, además:
 un módulo de corrección de tonalidad/ruido (109a) que está implementado para manipular una tonalidad o un ruido de la señal extraída no distorsionada o la señal extraída distorsionada.
- 45 13. El dispositivo según una de las reivindicaciones anteriores,
 donde el ensanchador de señales (102) comprende una pluralidad de canales de filtro, donde cada canal de filtro comprende un filtro para generar una señal de magnitud variable temporalmente (557) y una señal de frecuencia variable temporalmente (560) y un oscilador (502) controlable por las señales variables temporalmente,
 50 donde cada canal de filtro comprende un interpolador para interpolar la señal de magnitud variable temporalmente ($A(t)$), para obtener una señal de magnitud variable temporalmente interpolada ($A'(t)$), o un interpolador para interpolar la señal de frecuencia mediante el factor de ensanchamiento (104) para obtener una señal de frecuencia interpolada, y
 donde el oscilador (502) de cada canal de filtro se implementa para ser controlado por la señal de magnitud interpolada o por la señal de frecuencia interpolada.
 55
14. El dispositivo según una de las reivindicaciones 1 a 12, donde el ensanchador de señales (102) comprende:
 60 un procesador FFT (600) para generar espectros sucesivos para superponer bloques de muestras temporales de la señal de audio (100, 703), donde los bloques superpuestos están separados entre sí por una primera distancia de tiempo (a);
 un procesador IFFT para transformar espectros sucesivos desde un dominio de frecuencia al dominio de tiempo para generar bloques superpuestos de muestras de tiempo separadas entre sí por una segunda distancia de tiempo
 65 (b) que es mayor que la primera distancia (a); y

un reescalador de fase (606) para reescalar las fases de los valores espectrales de las secuencias de espectros FFT generados según una relación de la primera distancia (a) y la segunda distancia (b).

15. Un procedimiento para una extensión de ancho de banda de una señal de audio (100, 703), que
5 comprende:

generar (102) una versión de la señal de audio (100, 703) como una señal de tiempo ensanchada temporalmente mediante un factor de ensanchamiento (104) mayor que 1 para obtener una versión ensanchada temporalmente (103) de la señal de audio (100, 703);

10 diezmar (105) la versión ensanchada temporalmente (103) de la señal de audio (100, 703) mediante el factor de diezmado que coincide con el factor de ensanchamiento (104) para obtener una señal de audio diezmada (106);
extraer (107, 109)

15 una señal distorsionada de la señal de audio diezmada (106) para obtener una señal extraída distorsionada, conteniendo la señal extraída distorsionada un intervalo de frecuencias que no está contenido en la señal de audio (100, 703), o

una señal no distorsionada de la señal de audio diezmada (106) para obtener una señal extraída no distorsionada (108), conteniendo la señal extraída no distorsionada un intervalo de frecuencias que no está contenido en la señal de audio (100, 703), o

20 una señal de la señal de audio (100, 703) antes del ensanchamiento (102) para obtener una señal extraída no ensanchada, donde la señal extraída no ensanchada posterior al ensanchamiento (102) y al diezmado (105) contiene un intervalo de frecuencias que no está contenido en la señal de audio (100, 703); y

combinar (111)

25 la señal extraída distorsionada y la señal de audio (100, 703) para obtener una señal de audio (112) extendida en su ancho de banda, o

una señal distorsionada (110) generada a partir de la señal extraída no distorsionada (108) o la señal extraída no ensanchada posterior al ensanchamiento (102) y al diezmado (105) por medio de distorsión (109) y la señal de audio (100, 703) para obtener una señal de audio (112) extendida en su ancho de banda, o

30 la señal extraída no distorsionada (108) y la señal de audio (100, 703) para obtener una señal de combinación, donde una señal de audio generada distorsionando la señal de combinación representa una señal de audio (112) extendida en su ancho de banda,

35 donde la señal extraída distorsionada, o la señal distorsionada (110), o la señal de audio generada distorsionando la señal de combinación se distorsiona (109) de modo que la señal extraída distorsionada, o la señal distorsionada (110), o la señal de audio generada distorsionando la señal de combinación comprende una envolvente predeterminada.

40 16. Un programa informático que comprende instrucciones que, cuando son ejecutadas por un ordenador, hacen que el ordenador lleve a cabo el procedimiento según la reivindicación 15.

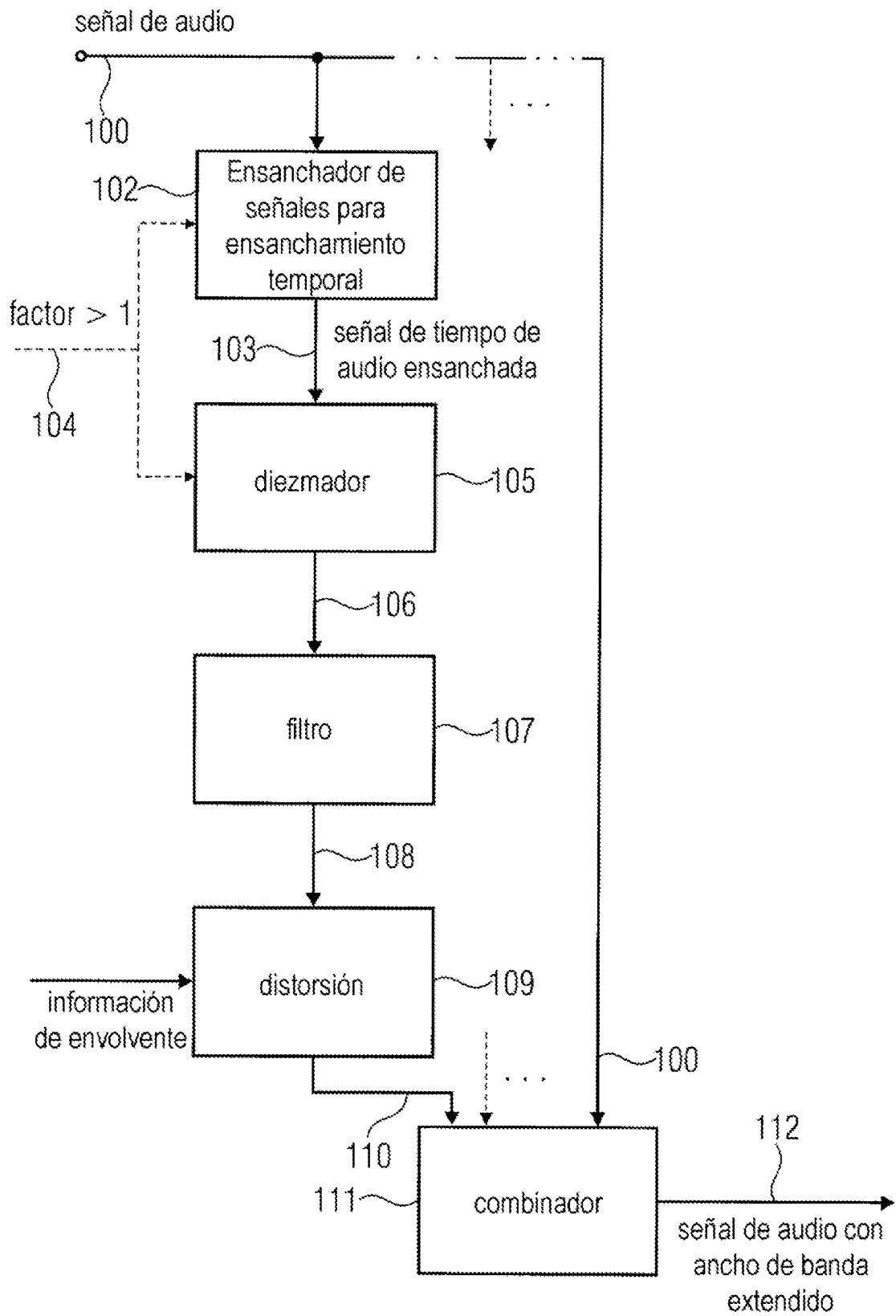


FIGURA 1

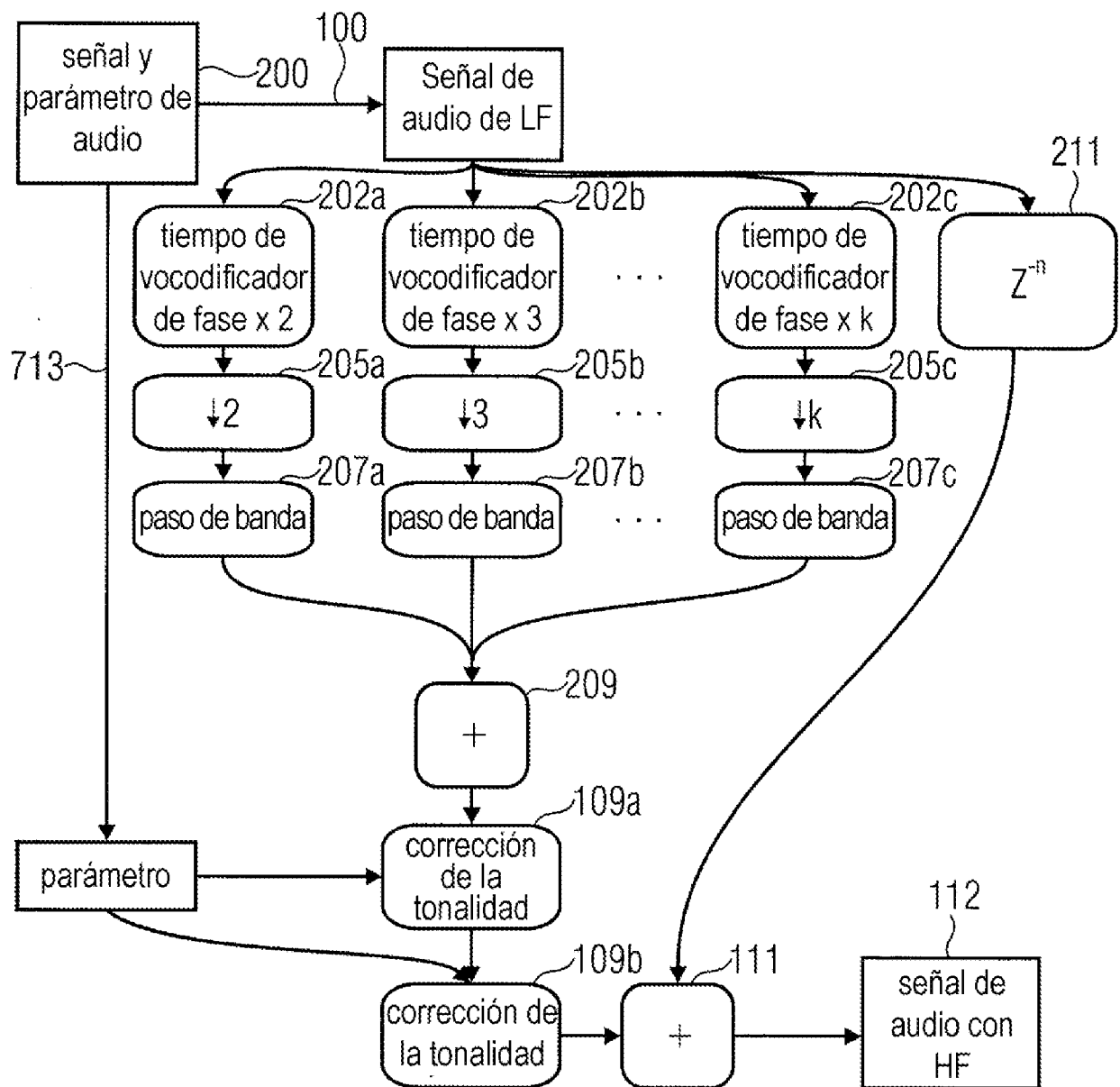


FIGURA 2A

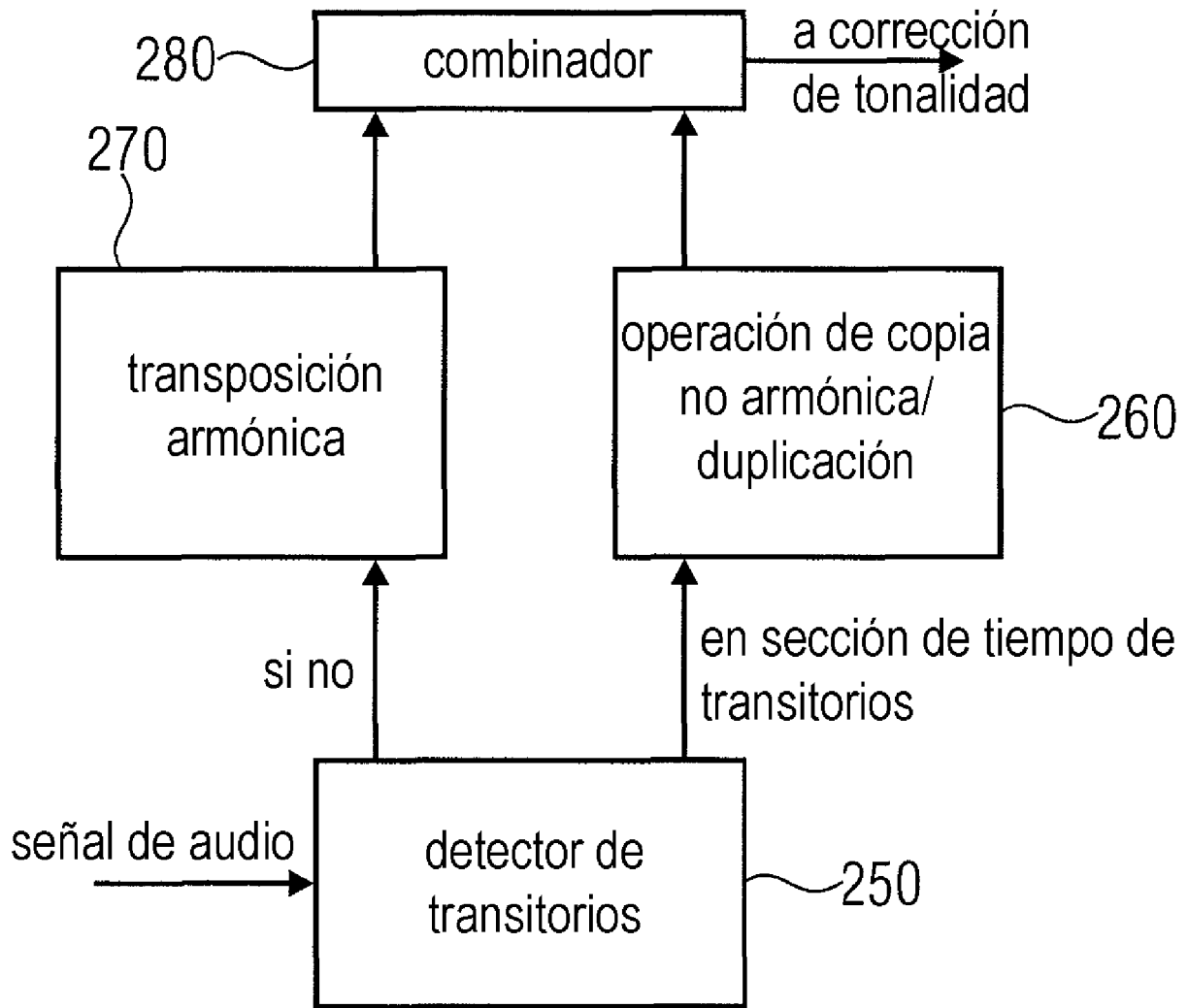


FIGURA 2B

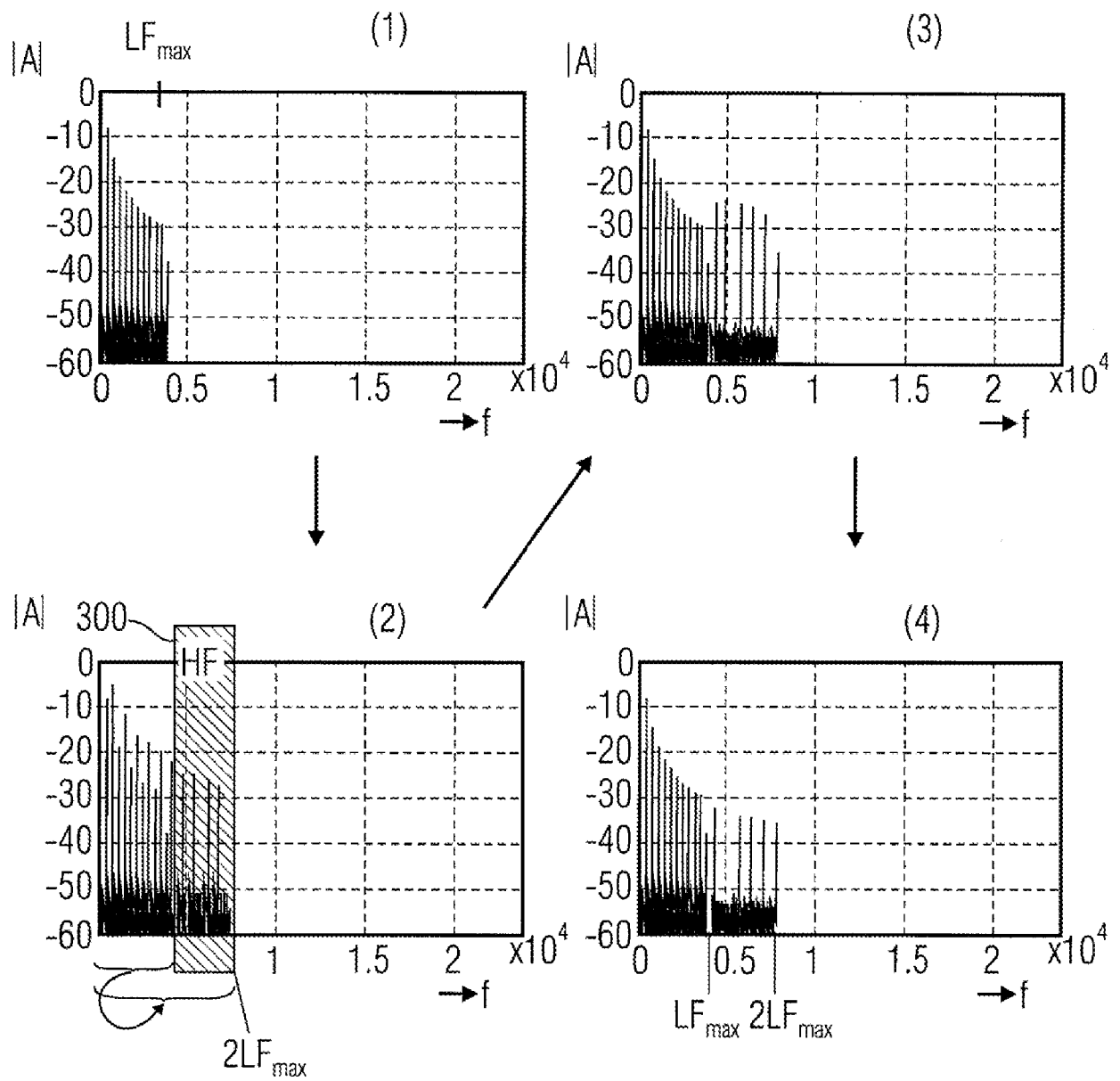


FIGURA 3

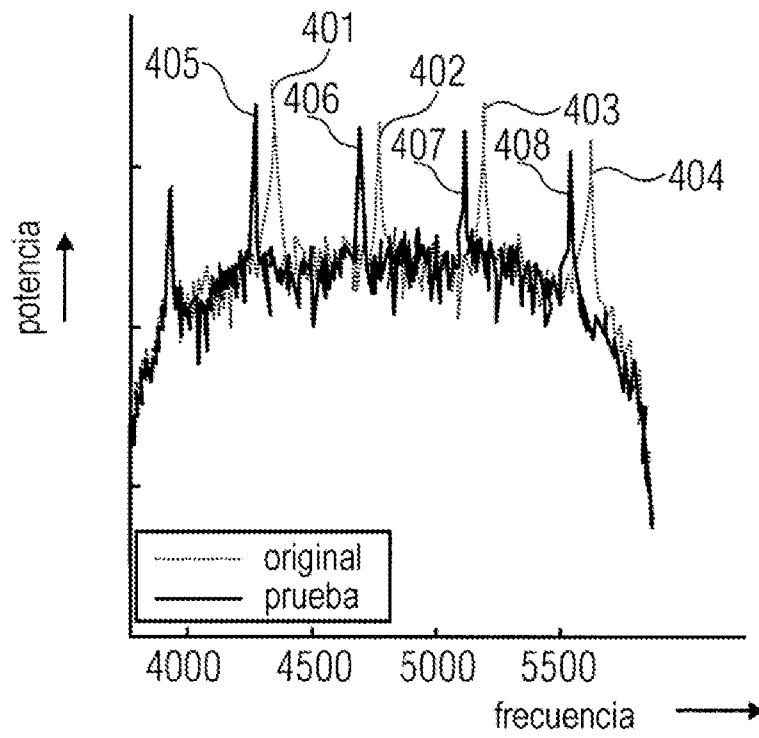


FIGURA 4A

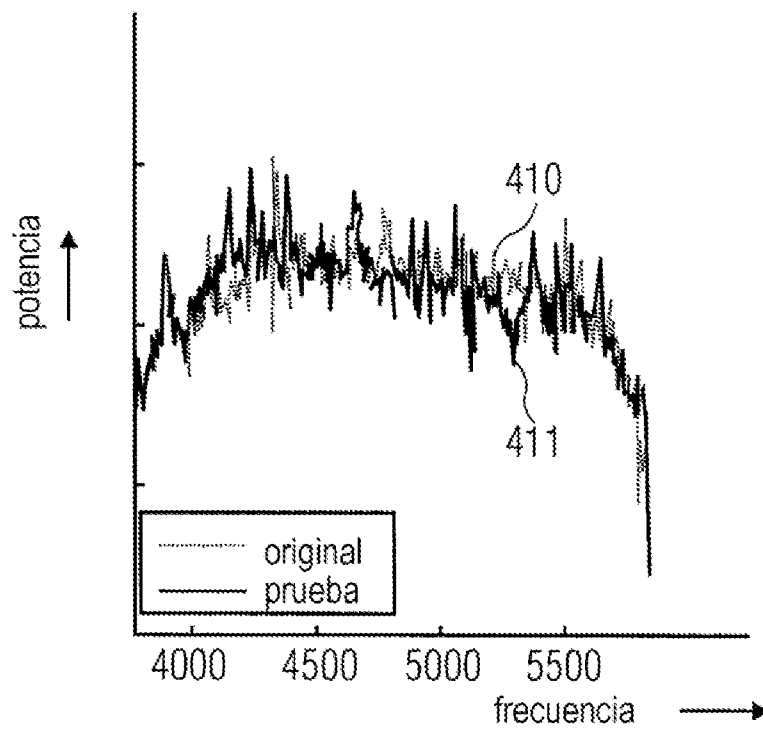


FIGURA 4B

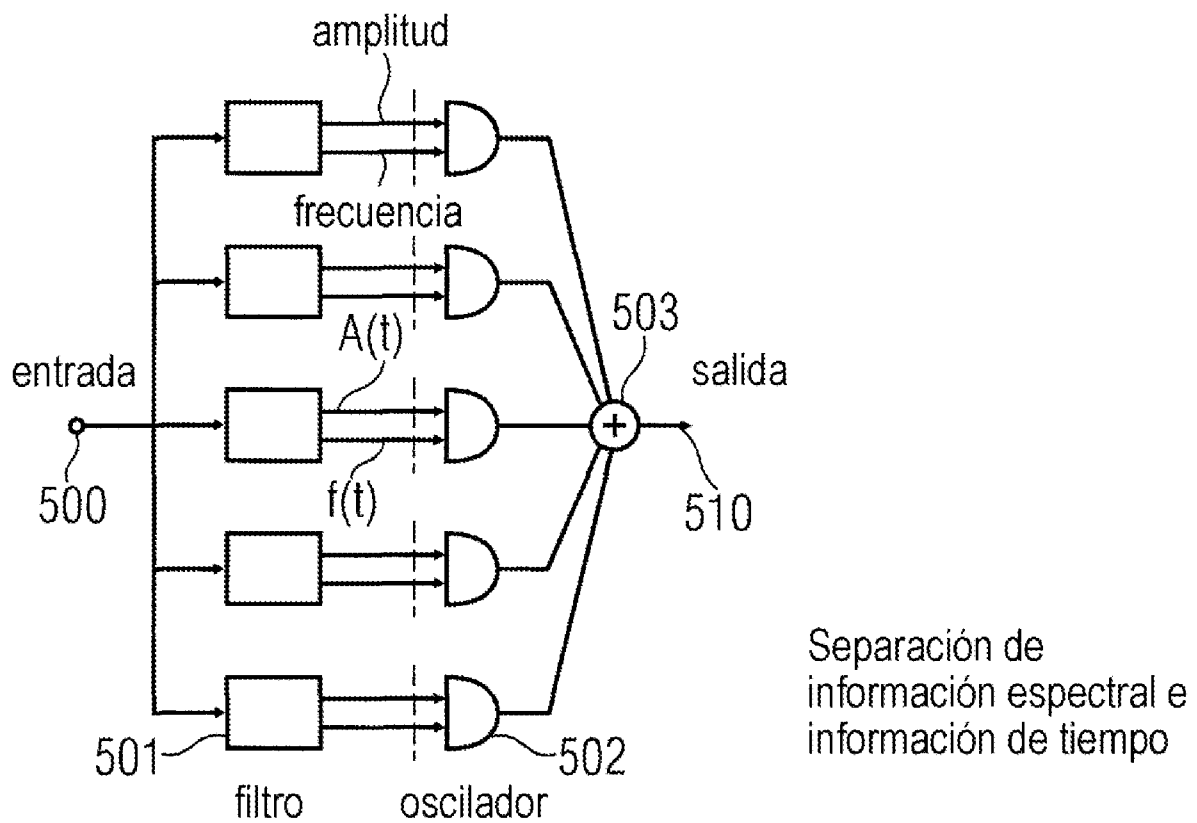


FIGURA 5A
(implementación de banco de filtros)

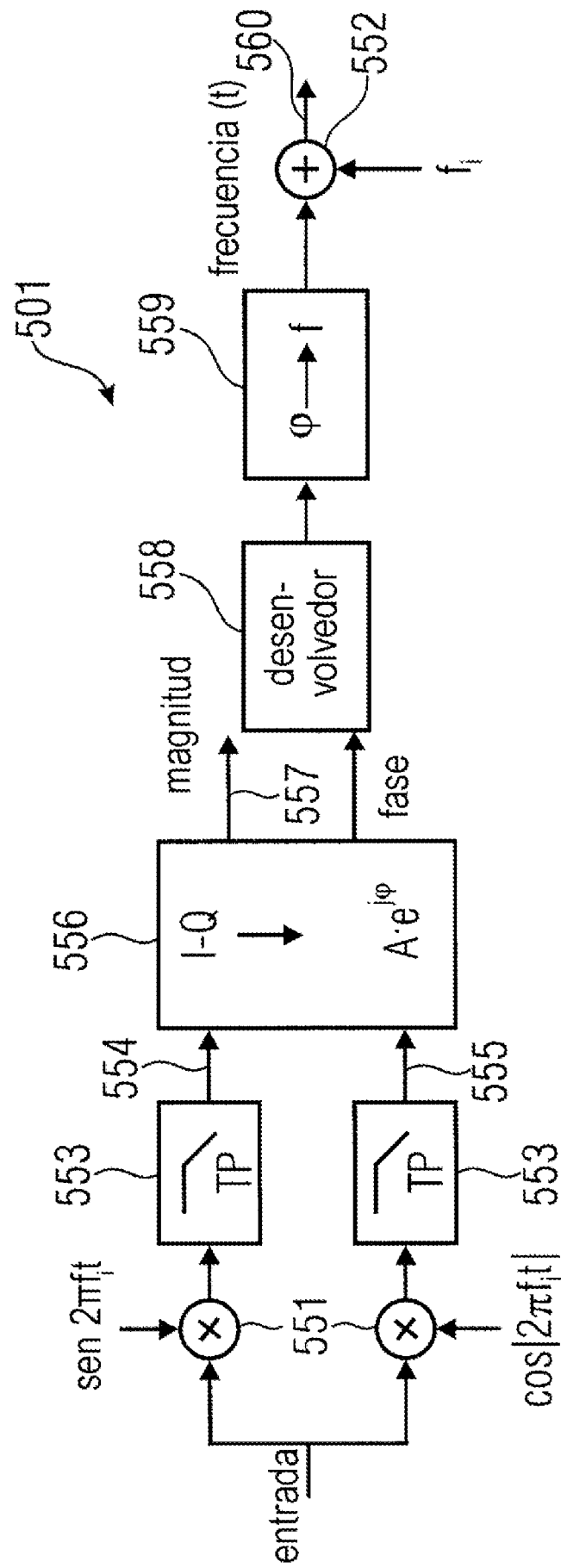


FIGURA 5B

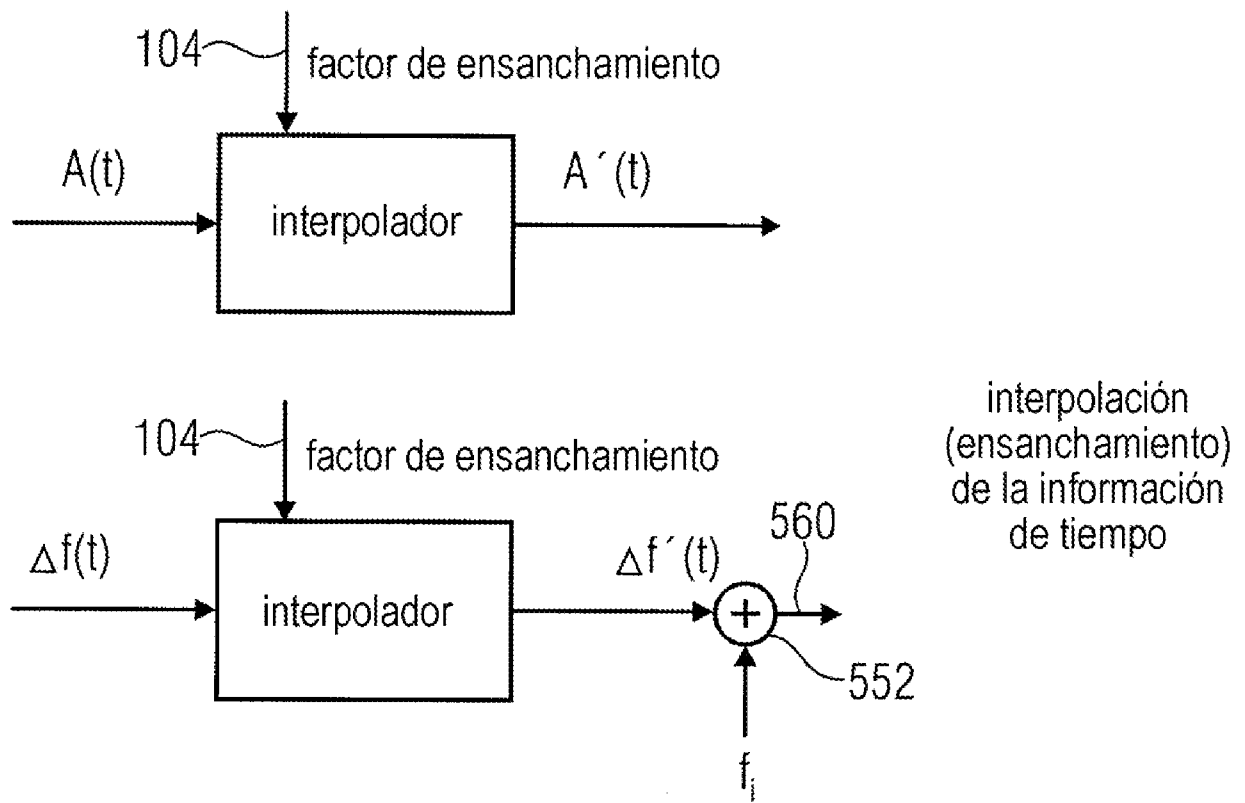


FIGURA 5C

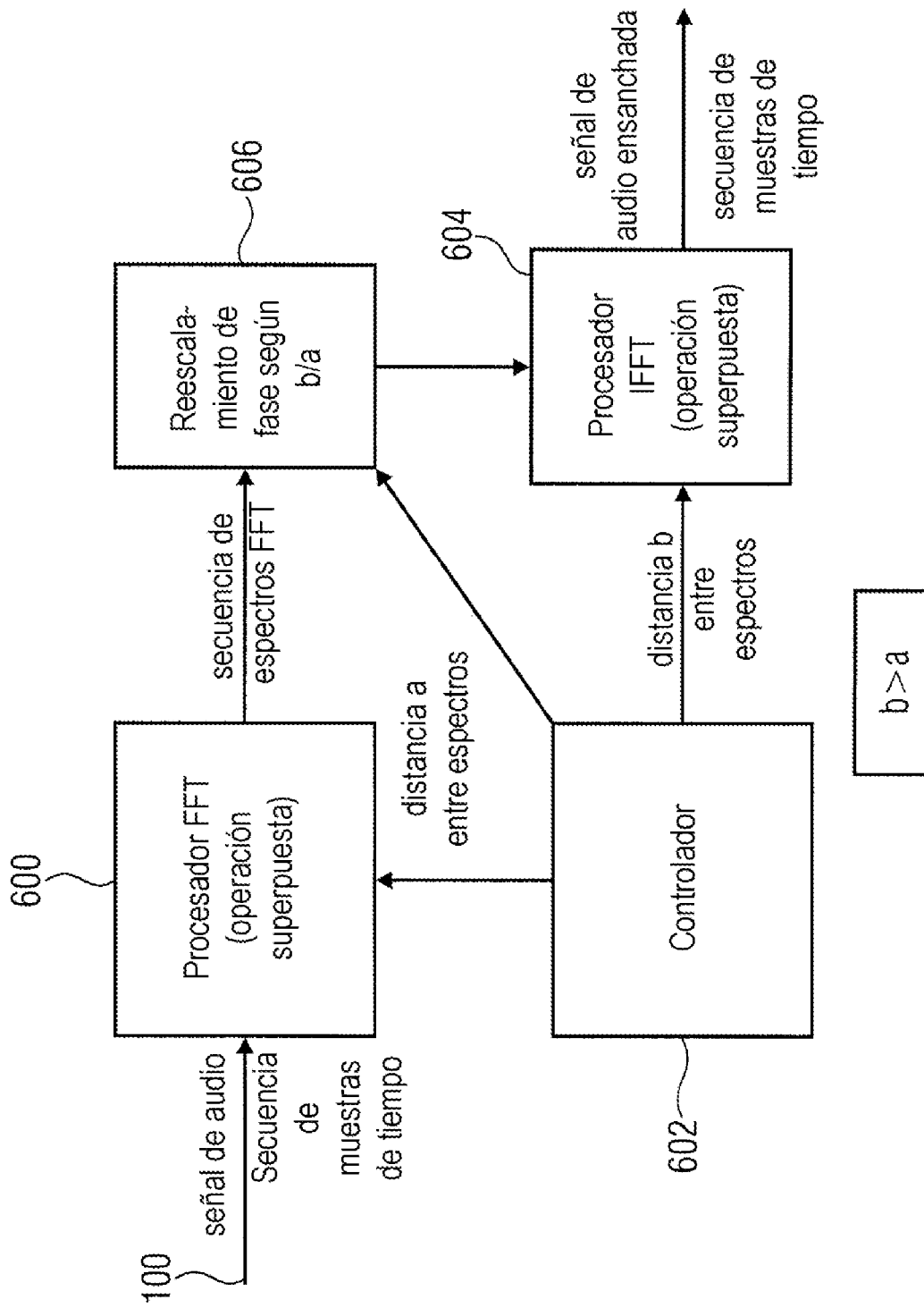


FIGURA 6
(implementación de transformación)

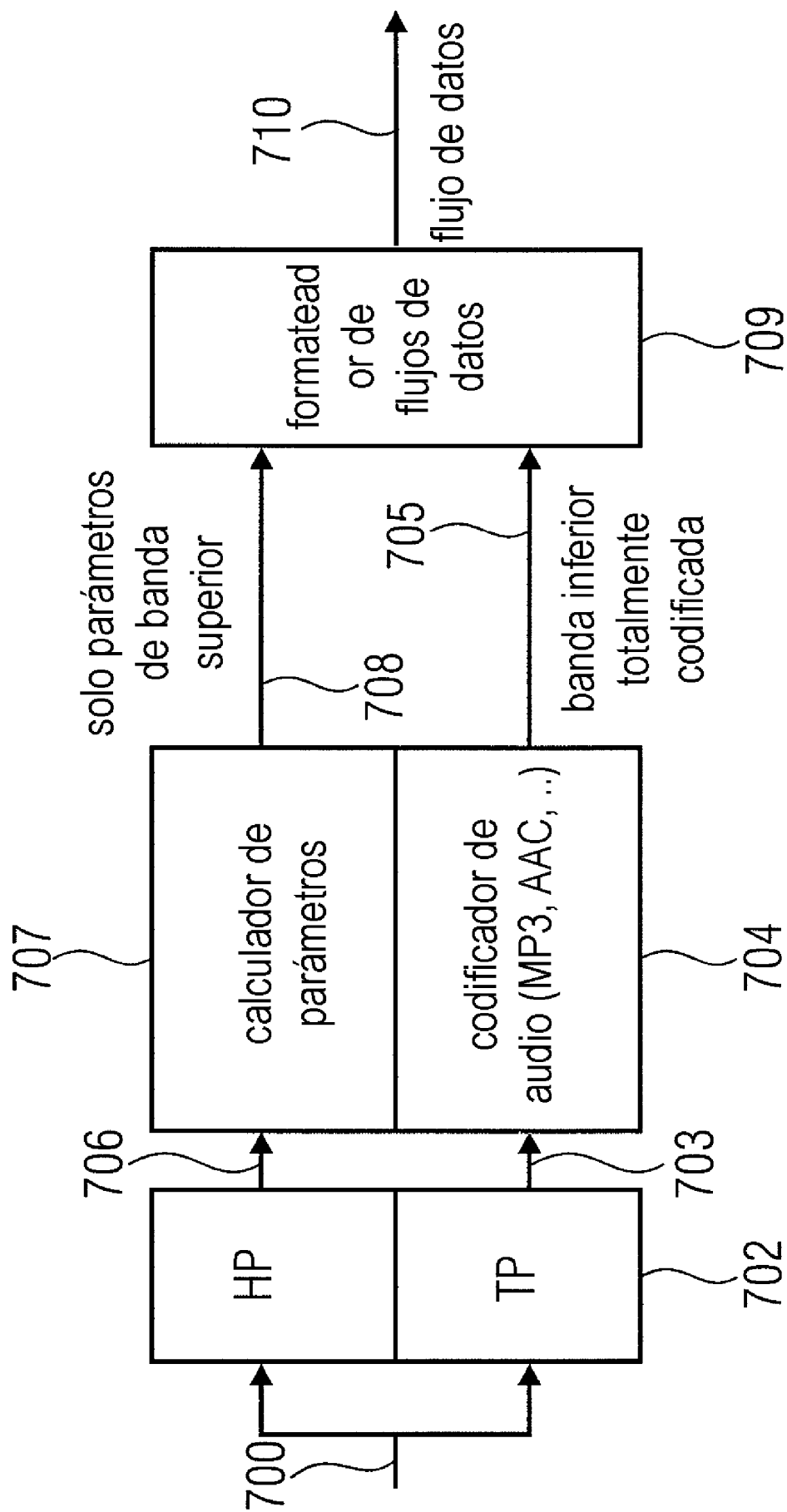


FIGURA 7A
(Lado del codificador)

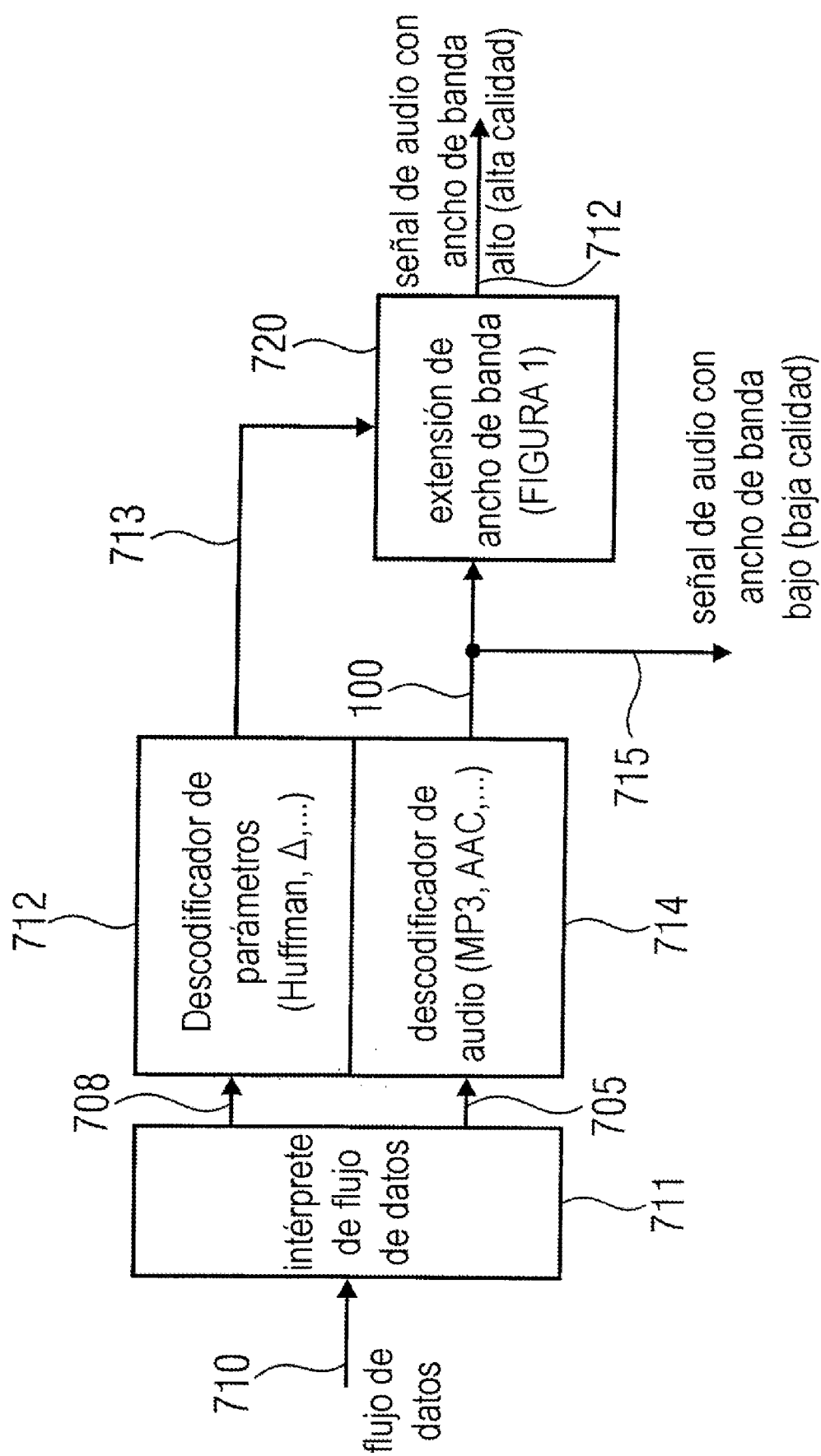


FIGURA 7B
(Lado del descodificador)