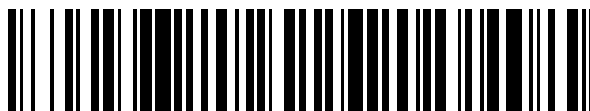


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 649 012**

51 Int. Cl.:

**G10L 21/038** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.01.2009 PCT/EP2009/000329**

87 Fecha y número de publicación internacional: **06.08.2009 WO09095169**

96 Fecha de presentación y número de la solicitud europea: **20.01.2009 E 09705824 (2)**

97 Fecha y número de publicación de la concesión europea: **06.09.2017 EP 2238591**

54 Título: **Procedimiento y dispositivo para extensión de ancho de banda de 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:  
**09.01.2018**

73 Titular/es:

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

72 Inventor/es:

**NAGEL, FREDERIK;**  
**DISCH, SASCHA y**  
**NEUENDORF, MAX**

74 Agente/Representante:

**SALVA FERRER, Joan**

**ES 2 649 012 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

Procedimiento y dispositivo para extensión de ancho de banda de 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 en las cuales la proporción de datos disponible es más bien pequeña.
- [0002]** La codificación adaptada a la audición de las señales de audio para una reducción de datos para un almacenamiento y transmisión eficiente de estas señales han ganado aceptación en muchos campos. Algoritmos de  
10 codificación son conocidos, en particular, tales como "MP3" o "MP4". La codificación usada para esto, en particular cuando se tienen velocidades de bits más bajas, conduce a la reducción de la calidad de audio que es frecuentemente provocada principalmente por una limitación lateral del codificador del ancho de banda de la señal de audio a ser transmitida.
- 15 **[0003]** Es conocido de WO 98 57436 el 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. La banda superior, sin embargo, está solamente caracterizada muy toscamente, es decir, por un conjunto de parámetros que reproduce la envolvente espectral de la banda superior. En el lado del descodificador, la banda superior es luego sintetizada. Para este propósito, se propone una transposición armónica,  
20 en la que la banda inferior de la señal de audio descodificada es suministrada a un banco de filtros. Los canales del banco de filtros de la banda inferior son conectados a canales del banco de filtros de la banda superior o son "parchados", y cada señal de paso de banda parchada es remitida a un ajuste de envolvente. El banco de filtro de síntesis perteneciente a un banco de filtros de análisis especial recibe aquí las señales de paso de banda de la señal de audio en la banda inferior y las señales de paso de banda envolventes-ajustadas de la banda inferior que fueron  
25 parchadas armónicamente en la banda superior. La señal de salida del banco de filtro de síntesis es una señal de audio extendida con respecto a su ancho de banda, que fue transmitida del lado del codificador al lado del descodificador con una proporción de datos muy baja. En particular, los cálculos de bancos de filtros y parches en el dominio del banco de filtros se pueden convertir en un alto esfuerzo computacional.
- 30 **[0004]** Los procedimientos de complejidad reducida para la extensión de ancho de banda de las señales de audio banda-limitadas utilizan en lugar de esto una función de copia de las porciones de señal de baja frecuencia (LF) en el intervalo de alta frecuencia (HF) con el fin de aproximar la información perdida debido a la limitación de banda. Tales métodos son descritos en M. Dietz, L. Liljeryd, K. Kjöring and O. Kunz, "Spectral Band Replication, a novel approach in audio coding," in 112th AES Convention, Munich, May 2002; S. Meltzer, R. Böhm and F. Henn, "SBR  
35 enhanced audio codecs for digital broadcasting tales como "Digital Radio Mondiale" (DRM)," 112th AES Convention, Munich, May 2002; T. Ziegler, A. Ehret, P. Ekstrand and M. Lutzky, "Enhancing mp3 with SBR: Features and Capabilities of the new mp3PRO Algorithm," in 112th AES Convention, Munich, May 2002; International Standard ISO/IEC 14496-3:2001/FPDAM I, "Bandwidth Extension," ISO/IEC, 2002, or "Speech bandwidth extension method and apparatus", Vasu lyengar et al. patente de EE.UU. n.º 5.455.888.
- 40 **[0005]** En estos procedimientos, no se efectúa ninguna transposición armónica, sino que señales de paso de banda sucesivas de la banda inferior son introducidas en canales de banco de filtro sucesivos de la banda superior. Mediante esto, se obtiene una aproximación tosca de la banda superior de la señal de audio. Esta aproximación tosca de la señal es luego en una etapa adicional aproximada a la original, mediante un post-procesamiento  
45 utilizando información de control, ganada de la señal original. Aquí, por ejemplo, los factores de escala sirven para adaptar la envolvente espectral, una filtración inversa y la adición de una alfombra de ruido para adaptar tonalidad y una complementación mediante porciones de señal sinusoidales, como también se describe en el estándar de MPEG-4.
- 50 **[0006]** Aparte de esto, existen métodos adicionales tales como la llamada "extensión de ancho de banda ciega" descrita en E. Larsen, R.M. Aarts, and M. Danessis, "Efficient high-frequency bandwidth extension of music and speech", In AES 112th Convention, Munich, Germany, May 2002, en el que no se usa información en cuanto al intervalo de HF original. Además, también existe el procedimiento de la llamada "Extensión de ancho de banda artificial" que es descrita en: K. Käyhkö, A Robust Wideband Enhancement for Narrowband Speech Signal; Research  
55 Report, Helsinki University of Technology, Laboratory of Acoustics and Audio signal Processing, 2001.
- [0007]** En J. Makinen et al.: AMR-WB+: se describe un nuevo estándar de codificación de audio para servicios de audio móviles de tercera generación Broadcasts, IEEE, ICASSP '05, se describe un procedimiento para la extensión de ancho de banda, en el que la operación de copia de la extensión de ancho de banda con una copia hacia arriba  
60 de señales de paso de banda sucesivas, de acuerdo con la tecnología de SBR, es reemplazada por reflejo, por ejemplo mediante toma de muestras ascendente.
- [0008]** Tecnologías adicionales para extensión de ancho de banda son descritas en los siguientes documentos: R.M. Aarts, E. Larsen, and O. Ouweltjes, "A unified approach to low- and high frequency bandwidth extension", AES 115th  
65 Convention, New York, USA, October 2003; E. Larsen and R.M. Aarts, "Audio Bandwidth Extension – Application to psychoacoustics, Signal Processing and Loudspeaker Design", John Wiley & Sons, Ltd., 2004; E. Larsen, R.M. Aarts,

and M. Danessis, "Efficient high-frequency bandwidth extension of music and speech", AES 112th Convention, Munich, May 2002; J. Makhoul, "Spectral Analysis of Speech by Linear Prediction", IEEE Transactions on Audio and Electroacoustics, AU-21(3), June 1973; solicitud de patente de EE.UU. 08/951,029; patente de EE.UU. n.º 6.895.375.

5 **[0009]** Los procedimientos conocidos de extensión de ancho de banda armónica muestran una alta complejidad. Por otra parte, procedimientos de extensión de ancho de banda de complejidad reducida muestran pérdidas de calidad. En particular, con una baja proporción de bits y en combinación con un ancho de banda bajo del intervalo de LF, se pueden presentar artefactos tales como rugosidad y un timbre percibido como desagradable. Una razón para esto es el hecho de que la porción de HF aproximada está basada en una operación de copia que deja relaciones armónicas  
10 de las porciones de señal tonal sin notar entre sí. Esto se aplica a ambos, a la relación armónica entre LF y HF y también a la relación armónica dentro de la porción de HF misma. Con SBR, por ejemplo, en la frontera entre el intervalo de LF y el intervalo de HF generado, ocurren ocasionalmente impresiones de sonido tosco como porciones tonales copiadas del intervalo de LF en el intervalo de HF, como se ilustra, por ejemplo, en la Fig. 4a, puede ahora en la señal global encontrar porciones tonales del intervalo de LF para estar espectralmente adyacentes  
15 densamente. Así, en la Fig. 4a, se ilustra una señal original con picos en 401, 402, 403, y 404, mientras que una señal de prueba es ilustrada con los picos en 405, 406, 407, y 408. Al copiar porciones tonales del intervalo de LF en el intervalo de HF, en las que en la Fig. 4a la frontera estaba a 4250 Hz, la distancia de los dos picos dejados en la señal de prueba es menor que la frecuencia base subyacente al tramador armónico, que conduce a una percepción de rugosidad.

20 **[0010]** Ya que el ancho de los grupos de frecuencia tono-compensados se incrementa con un incremento de la frecuencia central, como se describe en Zwicker, E. and H. Fastl (1999), Psychoacoustics: Facts and models. Berlin – Springer-Verlag, porciones sinusoidales que caen en el intervalo de LF en diferentes grupos de frecuencia, mediante copia en el intervalo de HF, pueden llegar a caer en el mismo grupo de frecuencia en la presente, lo cual  
25 también conduce a una impresión de audición tosca como puede verse en la Fig. 4b. Aquí es en particular mostrado que la copia del intervalo de LF en el intervalo de HF conduce a una estructura tonal más densa en la señal de prueba en comparación con la original. La señal original es distribuida de manera relativa uniformemente a través del espectro en el intervalo de frecuencia más alto, como se muestra en particular en 410. En contraste, en particular en este intervalo más alto, la señal de prueba 411 está distribuida relativamente no uniforme a través del espectro y así  
30 claramente más tonal que la señal original 410.

**[0011]** El Manual de Erik Larsen y Roland M. Arts: "Audio Bandwidth Extension", del 6 de diciembre, 2005, describe una extensión de ancho de banda para el habla que posee un etapa de duplicación de paso que comprende una toma de muestras descendente y una etapa subsecuente de estiramiento de tiempo, un filtro de desviación  
35 subsecuentemente conectado 5 y un adicionador que es alimentado por una señal original subsecuente para aplicar una compensación de retardo a esta señal original.

**[0012]** Es objeto de la presente invención obtener una extensión de ancho de banda con una alta calidad y todavía simultáneamente obtener un procesamiento de señal con una complejidad más baja, sin embargo, se puede hacer  
40 implementando poco retardo y poco esfuerzo y así también con procesadores que tienen requerimientos de hardware reducidos con respecto a la velocidad del procesador y memoria requerida.

**[0013]** Este objeto es obtenido por un dispositivo para la extensión de ancho de banda según la reivindicación 1 o un procedimiento para la extensión de ancho de banda según la reivindicación 12 o un programa de ordenador de  
45 acuerdo con la reivindicación 13. El concepto de la invención para la extensión de ancho de banda está basado en una dispersión de señal temporal para generar una versión de la señal de audio como una señal de tiempo que es dispersada por un factor de dispersión  $> 1$  y un diezmado subsecuente de la señal de tiempo para obtener una señal transpuesta, que puede luego, por ejemplo, ser filtrada por un filtro de paso de banda simple para extraer una porción de señal de alta frecuencia que puede solamente todavía ser distorsionada o cambiada con respecto a su  
50 amplitud, respectivamente, para obtener una buena aproximación para la porción de alta frecuencia original. La filtración de paso de banda puede alternativamente tener lugar antes de que la dispersión de señal sea efectuada, de tal manera que sólo el intervalo de frecuencia deseado está presente después de la dispersión en la señal dispersada, de tal manera que una filtración de paso de banda después de la dispersión puede ser omitida.

55 **[0014]** Con la extensión de ancho de banda armónica por una parte, los problemas resultantes de una operación de copia o reflejo, o ambos, pueden ser impedidos en base a una continuación y dispersión armónica del espectro utilizando el dispersor de señal para dispersar la señal de tiempo. Por otra parte, una dispersión temporal y diezmado subsecuente puede ser ejecutada más fácil por procesadores simples que por un banco de filtros de análisis/síntesis completo, como es usado por ejemplo con la transposición armónica, en la que se tienen que tomar  
60 adicionalmente decisiones en cuanto a cómo realizar los parches dentro del dominio de banco de filtros.

**[0015]** Preferentemente, para la dispersión de señal, se usa un vocoder de fase para el cual hay implementaciones de menor esfuerzo. Con el fin de obtener extensiones de ancho de banda con factores  $> 2$ , también varios fase-  
65 vocoders pueden ser usados en paralelo, lo que es una ventaja, en particular, con respecto al retardo de la extensión de ancho de banda que tiene que ser bajo en aplicaciones de tiempo real. Alternativamente, otros procedimientos para la dispersión de señal están disponibles, como por ejemplo el método de PSOLA (Adición y Superposición

Síncrona de Paso)

**[0016]** En una realización preferida de la presente invención, la señal de audio de LF es primero extendida en la dirección del tiempo con la frecuencia máxima  $LF_{m\acute{a}x}$  con la ayuda del vocoder de fase, es decir, a un múltiplo entero de la duración convencional de la señal. A partir de esto, en un diezmador descendente, un diezmado de la señal por el factor de la extensión temporal tiene lugar que en total conduce a una dispersión del espectro. Esto corresponde a una transposición de la señal de audio. Finalmente, la señal resultante es filtrada en paso de banda al intervalo (factor de extensión - 1)· $LF_{m\acute{a}x}$  al factor de extensión · $LF_{m\acute{a}x}$ . Alternativamente, las señales de alta frecuencia individuales generadas por la dispersión y diezmado pueden ser sometidas a filtración de paso de banda, de tal manera que al final se suman por completo a través de intervalo de alta frecuencia (es decir de  $LF_{m\acute{a}x}$  a  $k \cdot LF_{m\acute{a}x}$ ). Esto es razonable para el caso en que todavía una densidad espectral más alta de armónicas es deseable.

**[0017]** El procedimiento de extensión de ancho de banda armónica es ejecutado en una realización preferida de la presente invención en paralelo para varios factores de extensión diferentes. Como una alternativa al procesamiento en paralelo, también un solo vocoder de fase puede ser usado para que se ponga en operación serialmente y en el que los resultados intermedios sean guardados en memoria temporal. Así, se puede obtener cualesquiera frecuencias de corte de extensión de ancho de banda. La extensión de la señal puede alternativamente también ser ejecutada directamente en la dirección de frecuencia, es decir, en particular, por una operación doble correspondiente al principio funcional del vocoder de fase.

**[0018]** Ventajosamente, en realizaciones de la invención, no se requiere ningún análisis de la señal con respecto a armonicidad o frecuencia fundamental.

**[0019]** En lo siguiente, realizaciones preferidas de la presente invención son explicadas en más detalle con referencia a los dibujos adjuntos, en los cuales:

La Fig. 1 muestra un diagrama de bloque del concepto inventivo para una extensión de ancho de banda de la señal de audio;

La Fig. 2a muestra un diagrama de bloques de un dispositivo para la extensión de ancho de banda de una señal de audio de acuerdo con un aspecto de la presente invención;

La Fig. 2b muestra una mejora del concepto de la Fig. 2a con detectores transitorios;

La Fig. 3 muestra una ilustración esquemática del procesamiento de señal utilizando espectros en ciertos puntos en el tiempo de una extensión de ancho de banda de la invención;

La Fig. 4a muestra una comparación entre una señal original y una señal de prueba que provee una impresión de sonido tosca;

La Fig. 4b muestra una comparación de la señal original con una señal de prueba que también conduce a una impresión auditiva tosca;

La Fig. 5a muestra una ilustración esquemática de la implementación del banco de filtros de un vocoder de fase;

La Fig. 5b muestra una ilustración detallada del filtro de la Fig. 5a;

La Fig. 5c muestra una ilustración esquemática para la manipulación de la señal de magnitud y la señal de frecuencia en un canal de filtro de la Fig. 5a;

La Fig. 6 muestra una ilustración esquemática de la implementación de transformación de un vocoder 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 de banda de una señal de audio.

**[0020]** La Fig. 1 muestra una ilustración esquemática de un dispositivo o un procedimiento, respectivamente, para la extensión de ancho de banda de una señal de audio. Solo de manera ejemplar, la Fig. 1 es descrita como un dispositivo, aunque la Fig. 1 puede simultáneamente también ser considerada como el diagrama de flujo de un procedimiento para la extensión de ancho de banda. Aquí, la señal de audio es alimentada en el dispositivo en la entrada 100. La señal de audio es suministrada a un dispersor de señal 102 que es implementado para generar una versión de la señal de audio como una señal de tiempo dispersada en el tiempo por un factor de dispersión mayor de 1. El factor de dispersión en la realización ilustrada en la Fig. 1 es suministrado vía una entrada del factor de dispersión 104. La señal de tiempo de audio dispersada presente en una salida 103 del dispersor de señal 102 es suministrada a un diezmador 105 que es implementado para diezmar la señal de tiempo de audio dispersada temporalmente 103 por un factor de decimación que coincide con el factor de dispersión 104. Esto es ilustrado esquemáticamente por la entrada del factor de dispersión 104 en la Fig. 1, que es trazada en líneas discontinuas y lleva al diezmador 105. En una realización, el factor de dispersión en el dispersor de señal es igual al inverso del factor de diezmado. Por ejemplo, si se aplica un factor de dispersión de 2.0 en el dispersor de señal 102, se ejecuta un diezmado con un factor de diezmado de 0.5. Sin embargo, si el diezmado es descrito a efectos de que se efectúa un diezmado por un factor de 2, es decir, que cada segundo valor de muestra es eliminado, entonces en esta ilustración, el factor de diezmado es idéntico al factor de dispersión. Proporciones alternativas entre el factor de dispersión y el factor de diezmado, por ejemplo proporciones enteras o proporciones racionales, pueden también ser usadas dependiendo de la implementación. La extensión de ancho de banda armónica máxima es obtenida, sin embargo cuando el factor de dispersión es igual al factor de diezmado o al inverso del factor de diezmado,

respectivamente.

**[0021]** En una realización preferida de la presente invención, el diezmador 105 es implementado para, por ejemplo, eliminar cada segunda muestra (con un factor de dispersión igual a 2) de tal manera que se tiene como resultado una señal de audio diezmada que tiene la misma duración temporal que la señal de audio original 100. También se pueden usar otros algoritmos de diezmado, por ejemplo, formar valores promedios ponderados o considerar las tendencias del pasado o el 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 es suministrada a un filtro 107, en el que el filtro 107 es implementado 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 ser implementado como un filtro de paso de banda digital, por ejemplo, como un filtro de FIR o IIR o también como un filtro de paso de banda análogo, aunque una implementación digital es preferida. Además, el filtro 107 es implementado de tal manera que extrae el intervalo espectral superior generado por las operaciones 102 y 105 en las que, sin embargo, el intervalo espectral del fondo, que es de cualquier manera cubierto por la señal de audio 100, es suprimido tanto como sea posible. En la implementación, el filtro 107 puede también ser implementado de tal manera que, sin embargo, también extrae porciones de señal con frecuencias como una señal de paso de banda contenida en la señal original 100, en la que la señal de paso de banda extraída contiene por lo menos una banda de frecuencia que no estaba contenida en la señal de audio original 100.

**[0022]** La señal de paso de banda 108, emitida por el filtro 107, es suministrada a un distorsionador 109, que es implementado para distorsionar las señales de paso de banda, de tal manera que la señal de paso de banda comprende una envolvente predeterminada. Esta información de envolvente que puede ser usada para distorsión puede ser introducida externamente y aun proceder de un codificador o puede también ser generada internamente, por ejemplo mediante una extrapolación ciega de la señal de audio 100 o en base a tablas guardadas en el lado del descodificador indexadas con una envolvente de una señal de audio 100. La señal de paso de banda distorsionada 110 emitida por el distorsionador 109 es finalmente suministrada a un combinador 111 que es implementado para combinar la señal de paso de banda distorsionada 110 a la señal de audio original 100 que fue también distorsionada dependiendo de la implementación (la etapa de retardo no es indicada en la Fig. 1), para generar una señal de audio extendida con respecto a su ancho de banda en una salida 112.

**[0023]** En una implementación alternativa, la secuencia de distorsionador 109 y combinador 111 es inversa a la ilustración indicada en la Fig. 1. Aquí, la señal de salida del filtro, es decir, la señal de paso de banda 108, es combinada directamente con la señal de audio 100 y la distorsión de la banda superior de la señal combinada que es emitida a partir del combinador 111 es solamente ejecutada después de combinación por el distorsionador 109. En esta implementación, el distorsionador opera como distorsionador para distorsionar la señal de combinación de tal manera que la señal de combinación comprende una envolvente predeterminada. El combinador es en esta realización implementado así 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 está extendida con respecto a su ancho de banda. En esta realización, en la cual la distorsión solamente tiene lugar después de la combinación, es preferible implementar el distorsionador 109 de tal manera que no influya la señal de audio 100 o el ancho de banda de la señal de combinación, respectivamente, provista por la señal de audio 100, ya que la banda inferior de la señal de audio fue codificada por un codificador de alta calidad y está, en el lado del descodificador, en la síntesis de la banda superior, para hablar de la medida de todas las cosas y no se debe interferir con la extensión de ancho de banda.

**[0024]** Antes de que se ilustren realizaciones detalladas de la presente invención, se ilustra un escenario de extensión de ancho de banda con referencia a las Figs. 7a y 7b, en el cual la presente invención puede ser implementada ventajosamente. Una señal de audio es alimentada en una combinación de paso bajo/paso alto en una entrada 700. La combinación de paso bajo/paso alto por una parte 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 por paso bajo es codificada con un codificador de audio 704. El codificador de audio es, por ejemplo, un codificador de MP3 (capa 3 de MPEG1) o un codificador de AAC, también conocido como codificador de MP4 y descrito en el estándar de MPEG4. Codificadores de audio alternativos que proveen una representación transparente o representación ventajosamente transparente psicoacústicamente de la señal de audio de banda limitada 703 pueden ser usados en el codificador 704 para generar una señal de audio completamente codificada o psicoacústicamente codificada y preferentemente señal de audio codificada transparentemente psicoacústicamente 705, respectivamente. La banda superior de la señal de audio es emitida en una salida 706 por la porción de paso alto del filtro 702, designada por "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, es suministrada a un calculador de parámetros 707 que es implementado 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 representación de un factor de escala para cada grupo de frecuencia psicoacústico 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 preferentemente ser relacionada 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 la

energía espectral está distribuida en una banda, es decir, si la energía espectral en la banda está distribuida relativamente de manera uniforme, en la que entonces existe una señal no tonal en esta banda o si la energía en esta banda está concentrada relativamente fuerte en un cierto sitio en la banda, en el que luego existe una señal más bien tonal para esta banda. Parámetros adicionales consisten en codificar explícitamente picos que sobresalen  
 5 relativamente fuertemente en la banda superior con respecto a su altura y su frecuencia, como el concepto de extensión de ancho de banda, en la reconstrucción sin tal codificación explícita de porciones sinusoidales prominentes en la banda superior, que solamente recupera el mismo muy rudimentariamente o no.

**[0025]** En cualquier caso, el calculador de parámetros 707 es implementado para generar solamente parámetros  
 10 708 para la banda superior que puede ser sometida a etapas de reducción de entropías similares, ya que pueden también ser efectuadas en el codificador de audio 704 para valores espectrales cuantizados, como por ejemplo, codificación diferencial, predicción o codificación de Huffman, etc. La representación de parámetros 708 y la señal de audio 705 son luego suministradas a un formateador de flujo de datos 709 que es implementado para proporcionar un flujo de datos del lado de salida 710 que comúnmente será un tren de bits de acuerdo con un cierto formato como  
 15 es, por ejemplo, normalizado en el estándar de MPEG4.

**[0026]** El lado del descodificador, ya que es especialmente apropiado para la presente invención, es en lo siguiente ilustrado con respecto a la Fig. 7b. El flujo de datos 710 entra a un interpretador de flujo de datos 711 que es implementado para separar la porción de parámetros 708 de la porción de la señal de audio 705. La porción de  
 20 parámetros 708 es descodificada por un descodificador de parámetros 712 para obtener parámetros descodificados 713. En paralelo a esto, la porción de señal de audio 705 es descodificada por un descodificador de audio 714 para obtener la señal de audio que fue ilustrada en 100 en la Fig. 1.

**[0027]** Dependiendo de la implementación, la señal de audio 100 puede ser emitida vía una primera salida 715. En  
 25 la salida 715, una señal de audio con un ancho de banda pequeño y así también una calidad baja puede luego ser obtenida. Para mejorar la calidad, sin embargo, se efectúa la extensión de ancho de banda de la invención 720, que es por ejemplo implementada 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.

**[0028]** En lo siguiente, 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 preferentemente ser usada en el bloque 712 de la Fig. 7b. La Fig. 2a incluye en primer lugar un bloque designado por la "señal de audio y parámetro" que puede corresponder al bloque 711, 712, y 714 de la Fig. 7b y es designado por el bloque 200. El bloque 200 provee la señal de salida 100 también como parámetros descodificados 713 en el lado de salida que puede ser usado para diferentes distorsiones,  
 35 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, es suministrada al combinador 111 para obtener la señal de audio en el lado de salida con un ancho de banda extendido 112.

**[0029]** Preferentemente, el dispersor de señal 102 de la Fig. 1 es implementado por un vocoder de fase 202a. El diezmador 105 de la Fig. 1 es implementado preferentemente por un convertidor de velocidad de muestras simple 205a. El filtro 107 para la extracción de una señal de paso de banda es implementado preferentemente por un filtro de paso de banda simple 107a. En particular, el vocoder de fase 202a y el diezmador de velocidad de muestras 205a se ponen en operación con un factor de dispersión = 2.

**[0030]** Preferentemente, se provee un "tren" adicional que consiste en el vocoder de fase 202b, diezmador 205b y 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 frecuencia 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.

**[0031]** Además de esto, se provee un vocoder de fase k 202c que obtiene una dispersión de la señal de audio por el factor k, en la que k es preferentemente un número entero mayor de 1. Un diezmador 205 es conectado de forma descendente al vocoder de fase 202c, que diezma por el factor k. Finalmente, la señal diezmada es suministrada a un filtro de paso de banda 207c que es implementado 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 que corresponde a k veces la frecuencia máxima de la señal de audio 100. Todas las señales de paso de banda son combinadas por un combinador 209, en las que el combinador 209 puede ser implementado por ejemplo como un adicionador. Alternativamente, el combinador 209 puede también ser implementado como un adicionador ponderado que, dependiendo de la implementación, atenúa bandas superiores más fuertemente que las bandas inferiores, independientemente de la distorsión descendente por los elementos 109a, 109b. Además de esto, el sistema  
 50 ilustrado en la Fig. 2a incluye una etapa de retardo 211 que garantiza que una combinación sincronizada tenga lugar en el combinador 111 que puede ser por ejemplo una adición por muestras.

**[0032]** La Fig. 3 muestra una ilustración esquemática de diferentes espectros que se pueden presentar en el procesamiento ilustrado en la Fig. 1 o en la Fig. 2a. La imagen parcial (1) de la Fig. 3 muestra una señal de audio banda-limitada como está por ejemplo presente en 100 en la Fig. 1, o 703 en la Fig. 7a. Esta señal es preferentemente dispersada por el dispersor de señal 102 a un múltiplo entero de la duración original de la señal y  
 65

subsecuentemente diezmada por el factor entero, que conduce a una dispersión global del espectro como se ilustra en la imagen parcial (2) de la Fig. 3. La porción de HF es ilustrada en la Fig. 3 a medida que es extraída por un filtro de paso de banda que comprende una banda de paso 300. En la tercera imagen parcial (3), la Fig. 3 muestra las variantes en las cuales la señal de paso de banda ya está combinada con la señal de audio original 100 antes de la

5 distorsión de la señal de paso de banda. Así, se tiene como resultado un espectro combinado con una señal de paso de banda sin distorsionar, en el que luego, como se indica en la imagen parcial (4) una distorsión de la banda superior, pero si es posible, ninguna modificación de la banda inferior tiene lugar para obtener la señal de audio 112 con un ancho de banda extendido.

- 10 **[0033]** La señal de LF en la imagen parcial (1) tiene la frecuencia máxima  $LF_{m\acute{a}x}$ . El vocoder de fase 202a efectúa una transposición de la señal de audio, de tal manera que la frecuencia máxima de la señal de audio transpuesta es  $2LF_{m\acute{a}x}$ . Ahora, la señal resultante en la imagen parcial (2) es filtrada por paso de banda al intervalo de  $LF_{m\acute{a}x}$  a  $2LF_{m\acute{a}x}$ . Visto en general, cuando el factor de dispersión es designado por  $k$  ( $k > 1$ ), el filtro de paso de banda comprende una banda de paso de  $(k-1) \cdot LF_{m\acute{a}x}$  a  $k \cdot LF_{m\acute{a}x}$ . El procedimiento ilustrado en la Fig. 3 es repetido para
- 15 diferentes factores de dispersión, hasta que se obtiene la frecuencia más alta deseada  $k \cdot LF_{m\acute{a}x}$  en la que  $k =$  al factor de extensión máximo  $k_{m\acute{a}x}$ .

**[0034]** En lo siguiente, con referencia a las Figs 5 y 6, se ilustran implementaciones preferidas para un vocoder de fase 202a, 202b, 202c según la presente invención. La Fig. 5a muestra una implementación de banco de filtro de un

20 vocoder de fase, en la que una señal de audio es alimentada en una entrada 500 y obtenida 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 descendente 502. Las señales de salida de todos los osciladores de cada canal son combinadas por un combinador, que es implementado por ejemplo como un adicionador e indicado en 503, con el fin de obtener la señal de salida. Cada filtro 501 es implementado de tal manera que provee una señal de amplitud por una parte y una

25 señal de frecuencia por otra parte. La señal de amplitud y la señal de frecuencia son señales en el tiempo que ilustran un desarrollo de la amplitud en un filtro 501 con el paso del tiempo, en tanto que la señal de frecuencia representa un desarrollo de la frecuencia de la señal filtrada por un filtro 501.

**[0035]** Una configuración esquemática del filtro 501 es ilustrada en la Fig. 5b. Cada filtro 501 de la Fig. 5a puede ser

30 establecido como en la Fig. 5b, en la que, sin embargo, solamente las frecuencias  $f_i$  suministradas a los dos mezcladores de entrada 551 y el adicionador 552 son diferentes de canal a canal. Las señales de salida del mezclador son ambas filtradas en paso bajo por los pasos bajos 553, en el que las señales de paso bajo son diferentes ya que fueron generadas por frecuencias de oscilador local (frecuencias LO), que están fuera de fase por  $90^\circ$ . El filtro de paso de bajo superior 553 provee una señal de cuadratura 554, mientras que el filtro inferior 553

35 provee una señal de fase de entrada 555. Estas dos señales, es decir I y Q, son suministradas 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 con el paso del tiempo es emitida en una salida 557. La señal de fase es suministrada a un desenvolvedor de fase 558. En la salida del elemento 558, ya no hay ningún valor de fase presente que esté siempre entre 0 y  $360^\circ$ , sino un valor de fase que se incrementa

40 linealmente. Este valor de fase "sin envolver" es suministrado a un convertidor de fase/frecuencia 559 que puede ser implementado por ejemplo como un simple formador de diferencia de fase que resta una fase de un punto en el tiempo previo de una fase en un punto en el tiempo actual para obtener un valor de frecuencia para el punto en el tiempo actual. Este valor de frecuencia es agregado al valor de frecuencia constante  $f_i$  del canal de filtro  $i$  para obtener un valor de frecuencia que varía temporalmente en la salida 560. El valor de frecuencia en la salida 560

45 tiene un componente directo =  $f_i$  y un componente alternante = la desviación de frecuencia por la cual una frecuencia actual de la señal en el canal de filtro se desvía de la frecuencia promedio  $f_i$ .

**[0036]** Así, como se ilustra en las Figs. 5a y 5b, el vocoder de fase obtiene una separación de la información

50 espectral e información de tiempo. La información espectral está en el canal especial o en la frecuencia  $f_i$  que provee 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 con respecto al tiempo, respectivamente.

**[0037]** La Fig. 5c muestra una manipulación que es ejecutada por el incremento de ancho de banda de acuerdo con la invención, en particular, en el vocoder de fase 202a y en particular en el sitio del circuito ilustrado y trazado en

55 líneas discontinuas en la Fig. 5a.

**[0038]** Para la escala de tiempo, por ejemplo las señales de amplitud  $A(t)$  en cada canal o la frecuencia de las

60 señales  $f(t)$  en cada señal, pueden ser diezgadas o interpoladas, respectivamente. Por propósitos de transposición, ya que es útil para la presente invención, una interpolación, es decir, una extensión o dispersión temporal de las señales  $A(t)$  y  $f(t)$  es efectuada para obtener señales dispersadas  $A'(t)$  y  $f'(t)$ , en las que la interpolación es controlada por el factor de dispersión 104, como se ilustró en la Fig. 1. Por la interpolación de la variación de fase, es decir, el valor antes de la adición de la frecuencia constante por el adicionador 552, la frecuencia de cada oscilador individual 502 en la Fig. 5a no es cambiada. El cambio temporal de la señal de audio global es frenada, sin embargo esto es por el factor 2. El resultado es un tono dispersado temporalmente que tiene el paso original, es decir, la onda

65 fundamental original con su armónica.

**[0039]** Al efectuar el procesamiento de señal ilustrado en la Fig. 5c, en la que tal procesamiento es ejecutado en cada canal de banda de filtro en la Fig. 5 y por la señal temporal resultante después de ser diezmada en el diezmador 105 de la Fig. 1, o en el diezmador 205a en la Fig. 5a, respectivamente, la señal de audio es contraída a su duración original en tanto que todas las frecuencias son duplicadas simultáneamente. Esto conduce a una transposición de paso por el factor 2 en la que, sin embargo, una señal de audio es obtenida y que tiene la misma longitud que la señal de audio original, es decir, el mismo número de muestras.

**[0040]** Como una alternativa a la implementación de banda de filtro ilustrada en la Fig. 5a, una implementación de transformación de un vocoder de fase puede también ser usada. Aquí, la señal de audio 100 es alimentada en un procesador de FFT o más en general, un procesador de Transformación de Fourier de Tiempo Corto 600 como una secuencia de muestras de tiempo. El procesador de FFT 600 es implementado esquemáticamente en la Fig. 6 para efectuar una ventana de tiempo de una señal de audio con el fin de después, por medio de un FFT, calcular tanto el espectro de magnitud, como también un espectro de fase, en el que este cálculo es efectuado para espectros sucesivos que están relacionados con bloques de la señal de audio, y que están fuertemente superpuestos.

**[0041]** En un caso extremo, para cada nueva muestra de señal de audio se puede calcular un nuevo espectro, en la que un nuevo espectro puede ser calculado también, por ejemplo solamente para cada vigésima nueva muestra. Esta distancia a en muestras entre dos espectros es preferentemente dada por un controlador 602. El controlador 602 es implementado adicionalmente para alimentar un procesador de IFFT 604 que es implementado para operar en una operación superpuesta. En particular, el procesador de IFFT 604 es implementado de tal manera que efectúa una Transformación de Fourier de Tiempo Corto inversa al efectuar una IFFT por espectro en base a un espectro de magnitud y un espectro de fase, con el fin de efectuar luego una operación de adición y superposición de la cual se tiene como resultado el intervalo de tiempo. La operación de adición y superposición elimina los efectos de la ventana de análisis.

**[0042]** Una dispersión de la señal de tiempo es obtenida por la distancia  $b$  entre dos espectros, a medida que son procesados por el procesador de IFFT 604, que es mayor que la distancia  $a$  entre los espectros en la generación de los espectros de FFT. La idea básica es dispersar la señal de audio por las FFT inversas simplemente al estar espaciadas adicionalmente más que las FFT de análisis. Como resultado, ocurren cambios espectrales en la señal de audio sintetizada más lentamente que en la señal de audio original.

**[0043]** Sin un reescalamiento de fase en el bloque 606, esto conduciría sin embargo a artefactos de frecuencia. Cuando, por ejemplo, una sola bandeja de frecuencia es considerada para la cual valores de fase sucesivos por  $45^\circ$  son implementados, esto implica que la señal dentro de esta banda de filtro se incrementa en la fase con una velocidad de  $1/8$  de un ciclo, es decir, por  $45^\circ$  por intervalo de tiempo, en la que el intervalo de tiempo aquí es el intervalo de tiempo entre FFT sucesivas. Si ahora las FFT inversas están separadas adicionalmente entre sí, esto significa que el incremento de fase de  $45^\circ$  ocurre a través de un intervalo de tiempo más largo. Esto significa que la frecuencia de esta porción de señal fue reducida no intencionalmente. Para eliminar esta reducción de frecuencia de artefacto, la fase es reescalada por exactamente el mismo factor por el cual la señal de audio fue dispersada en el tiempo. La fase de cada valor espectral de FFT es así incrementada por el factor  $b/a$ , de tal manera que esta reducción de frecuencia no intencional es eliminada.

**[0044]** En tanto que en la realización ilustrada en la Fig. 5c, la dispersión por interpolación de las señales de control de amplitud/frecuencia fue obtenida por un oscilador de señal en la implementación de banco de filtros de la Fig. 5a, la dispersión en la Fig. 6 se obtiene por la distancia entre dos espectros de IFFT que es mayor que la distancia entre dos espectros de FFT, es decir,  $b$  es mayor que  $a$ , en la que, sin embargo, para una prevención de artefacto se ejecuta un reescalamiento de fase de acuerdo con  $b/a$ .

**[0045]** Con respecto a una descripción detallada de vocoders de fase se hace referencia a los siguientes documentos:

"The phase Vocoder: A tutorial", Mark Dolson, Computer Music Journal, vol. 10, no. 4, pp. 14 -- 27, 1986, or "New phase Vocoder techniques for pitch-shifting, harmonizing and other exotic effects", L. Laroche und M. Dolson, Proceedings 1999 IEEE Workshop on applications of signal processing to audio and acoustics, New Paltz, New York, October 17 - 20, 1999, pages 91 to 94; "New approached to transient processing interphase vocoder", A. Röbel, Proceeding of the 6th international conference on digital audio effects (DAFx-03), London, UK, September 8-11, 2003, pages DAFx-1 to DAFx-6; "Phase-locked Vocoder", Meller Puckette, Proceedings 1995, IEEE ASSP, Conference on applications of signal processing to audio and acoustics, o la solicitud de patente de EE.UU n.º 6.549.884.

**[0046]** La Fig. 2b muestra una mejora del sistema ilustrado en la Fig. 2a, en donde se usa un detector transitorio 250 que es implementado 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 la energía de la señal de audio cambia en más del 50% de una porción temporal a la siguiente porción temporal, es decir, se incrementa o disminuye. El umbral de 50% es solamente un ejemplo, sin embargo, y puede también ser de valores más pequeños o mayores. Alternativamente, para una detección transitoria, el cambio de distribución de



energía puede también ser considerado, por ejemplo en la conversión de una vocal a consonante sibilante.

**[0047]** Si una porción transitoria de la señal de audio es determinada, la transposición armónica es dejada y para el intervalo de tiempo transitorio, se ejecuta una conmutación a una operación de copia no armónica o un reflejo no armónico o algún otro algoritmo de extensión de ancho de banda, como se ilustra en 260. Si luego se detecta otra vez que la señal de audio ya no es transitoria, se efectúa otra vez una transposición armónica, como se ilustra por los elementos 102, 105 en la Fig. 1. Esto se ilustra en 270 en la Fig. 2b.

**[0048]** Las señales de salida de los bloques 270 y 260 que llegan desfasados en el tiempo debido al hecho de que una porción temporal de la señal de audio puede ser o bien transitoria o no transitoria, son suministradas a un combinador 260 que es implementado para proporcionar una señal de paso de banda con el tiempo que puede por ejemplo ser suministrada a la corrección de tonalidad en el bloque 109a de la Fig. 2a. Alternativamente, la combinación por bloque 280 puede por ejemplo también ser efectuada después del adicionador 111. Esto significaría, sin embargo, que para un bloque de transformación entero de la señal de audio, se supone una característica transitoria o si la implementación de banco de filtros también opera en base a bloques, para todo un bloque como tal se toma una decisión en favor de si es transitorio o no es transitorio, respectivamente.

**[0049]** Ya que un vocoder de fase 202a, 202b, 202c, como se ilustra en la Fig. 2a y explicado en más detalle en las Figs. 5 y 6, genera más artefactos en el procesamiento de porciones de señal transitorias que en el procesamiento de porciones de señal no transitorias, se efectúa una conmutación a una operación de copia no armónica o reflejo, como se ilustró en la Fig. 2b en 260. Alternativamente, también se puede efectuar un restablecimiento de fase al transitorio, como se describe, por ejemplo, en la publicación de expertos por Laroche citados anteriormente o en la patente estadounidense con n.º 6.549.884.

**[0050]** Como ya se indicó, en los bloques 109a, 109b, después de la generación de la porción de HF del espectro, se efectúa una formación espectral y un ajuste a la medida original de ruido. La formación espectral puede tener lugar, por ejemplo, con la ayuda de los factores de escala, factores de escala dB(A)-ponderados o una predicción lineal, en la que existe la ventaja en la predicción lineal de que ninguna conversión de tiempo/frecuencia y ninguna conversión de frecuencia/tiempo subsecuente es requerida.

**[0051]** La presente invención es ventajosa ya que mediante el uso del vocoder de fase, un espectro con una frecuencia incrementada es dispersada adicionalmente y es siempre continuada armónicamente de manera correcta por la dispersión del entero. Así, el resultado de lo tosco en la frecuencia de corte del intervalo de LF es excluido y las interferencias demasiado densamente ocupadas por porciones de HF del espectro son evitadas. Además, se pueden usar implementaciones de vocoder de fase eficientes que pueden hacerse sin operaciones de parche de banco de filtros.

**[0052]** Alternativamente, otros métodos para la dispersión de señal están disponibles, tales como, por ejemplo, el método de PSOLA (Pitch Synchronous Overlap Add). Adición y Superposición Síncrona de Paso, abreviado PSOLA, es un método de síntesis en el cual las grabaciones de las señales de habla están ubicadas en la base de datos. Como éstas son señales periódicas, las mismas son provistas con información en cuanto a la frecuencia fundamental (paso) y el comienzo de cada periodo es marcado. En la síntesis, estos periodos son cortados con un cierto medio ambiente por medio de una función de ventana y agregados a la señal a ser sintetizada en un sitio apropiado: dependiendo de si la frecuencia fundamental deseada es más alta o más baja que aquélla de la entrada de base de datos, son combinadas, de este modo, más densas o menos densas que en el original. Para ajustar la duración de los periodos audibles, pueden ser omitidos o emitidos en doble. Este método es también llamado TD-PSOLA, en el que TD significa dominio de tiempo (Time Domain) y enfatiza que los métodos operan en el dominio de tiempo. Un desarrollo adicional es el Método de Adición y Superposición de Resíntesis de Multibanda, abreviado MBROLA. Aquí, los segmentos en la base de datos son traídos a una frecuencia fundamental uniforme mediante un pre-procesamiento y la posición de fase de la armónica es normalizada. Mediante esto, en la síntesis de una transición de un segmento al siguiente, se tienen como resultado menos interferencias perceptivas y la calidad de habla obtenida es más alta.

**[0053]** En una alternativa adicional, la señal de audio ya está filtrada en paso de banda antes de la dispersión, de tal manera que la señal después de la dispersión y diezmado ya contiene las porciones deseadas y la filtración de paso de banda subsecuente puede ser omitida. En este caso, el filtro de paso de banda es ajustado de tal manera que la porción de la señal de audio que habría sido filtrada después de la extensión de ancho de banda está todavía contenida en la señal de salida del filtro de paso de banda. El filtro de banda contiene así un intervalo de frecuencia que no está contenido en la señal de audio 106 después de la dispersión y diezmado. La señal con este intervalo de frecuencia 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 dispersada y diezmada derivada de una señal de audio filtrada de paso de banda.

**[0054]** Cabe señalar además que la señal dispersada puede también ser útil en el intervalo de frecuencia de la señal original, por ejemplo, al mezclar la señal original y señal dispersada, así, no se requiere ninguna banda de paso "estricta". La señal dispersada puede luego ser bien mezclada con la señal original en la banda de frecuencia en la

cual se superpone con la señal original con respecto a la frecuencia, para modificar la característica de la señal original en el intervalo de superposición.

- 5 **[0055]** Cabe señalar además que las funcionalidades de distorsión 109 y filtración 107 pueden ser implementadas en un solo bloque de filtro o en dos filtros separados en cascada. Ya que la distorsión tiene lugar dependiendo de la señal, las características de amplitud de este bloque de filtro serán variables. Su característica de frecuencia es, sin embargo independiente de la señal.
- 10 **[0056]** Dependiendo de la implementación, como se ilustra en la Fig. 1, primero, la señal de audio global puede ser dispersada, diezmada y luego filtrada, en la que la filtración corresponde a la operación de los elementos 107, 109. La distorsión es así ejecutada después o simultáneamente a la filtración, en la que para este propósito un bloque de filtro/distorsionador combinado en forma de un filtro digital es apropiado. Alternativamente, antes de la filtración (paso de banda) 107, una distorsión puede tener lugar aquí cuando se usan dos elementos de filtro diferentes.
- 15 **[0057]** Otra vez, alternativamente una filtración de paso de banda puede tener lugar antes de la dispersión, de tal manera que solamente la distorsión 109 sigue después del diezmado. Para estas funciones, dos elementos diferentes son preferidos.
- 20 **[0058]** Otra vez, alternativamente, también en todas las variantes anteriores, la distorsión puede tener lugar después de la combinación de la señal de síntesis con la señal de audio original, como por ejemplo, con un filtro que no tiene, o tiene solo muy poco efecto, sobre la señal a ser filtrada en el intervalo de frecuencia del filtro original que, sin embargo, genera la envolvente deseada en el intervalo de frecuencia extendido. En este caso, otra vez dos elementos diferentes son usados preferentemente para la extracción y distorsión.
- 25 **[0059]** El concepto de la invención es apropiado para todas las aplicaciones de audio en las cuales el ancho de banda total no está disponible. En la propagación de contenidos de audio, como por ejemplo, mediante radio digital, transmisión continua en internet y en aplicaciones de comunicación de audio, el concepto de la invención puede ser usado.
- 30 **[0060]** Dependiendo de las circunstancias, el procedimiento de la invención puede ser implementado para analizar una señal de información en hardware o en software. La implementación puede ser ejecutada en un soporte de almacenamiento digital, en particular un disquete o un CD que tiene señales de control que se pueden leer electrónicamente en el mismo, que pueden cooperar con el sistema de ordenador programable, de tal manera que el procedimiento es efectuado. En general, la invención consiste así en un producto de programa de ordenador con
- 35 códigos de programa para ejecutar el procedimiento almacenado en un portador que puede ser leído por la máquina, cuando el producto de programa de ordenador es ejecutado en un ordenador. En otras palabras, la invención puede así ser realizada como un programa de ordenador que tiene códigos de programa para efectuar el procedimiento, cuando el programa de ordenador es ejecutado en un ordenador.

**REIVINDICACIONES**

1. Un dispositivo para la extensión de ancho de banda de una señal de audio, que comprende:

5 un dispersor de señal (102) para generar una versión de la señal de audio como una señal de tiempo dispersada en el tiempo por un factor de dispersión de 2 para obtener una primera señal dispersada;

un dispersor de señal adicional (202b) implementado para dispersar la señal de audio (100) por un segundo factor de dispersión de 3 para obtener una segunda señal dispersada;

10

un diezmador (105) para diezmar la primera señal dispersada por un factor de diezmado de 2 para obtener una primera señal de audio diezmada (106);

15 un diezmador adicional (205b) implementado para diezmar la segunda señal dispersada por un segundo factor de diezmado de 3 para obtener una segunda señal de audio diezmada;

20 un filtro (107, 109) para extraer una primera señal de paso de banda de la primera señal de audio diezmada (106) que contiene un intervalo de frecuencia que está entre una frecuencia máxima de la señal de audio (100) y dos veces la máxima frecuencia de la señal de audio (100), o para extraer una primera señal de paso de banda de la señal de audio antes de ser generada por el dispersor de señal (102), en el que la primera señal de paso de banda, después de ser generada por el dispersor de señal (102) y diezmada por el diezmador (105), tiene un intervalo de frecuencia que está entre la máxima frecuencia de la señal de audio (100) y dos veces la frecuencia máxima de la señal de audio (100),

25 un filtro (207b) para extraer una segunda señal de paso de banda de la segunda señal diezmada que contiene un intervalo de frecuencia que está entre dos veces la frecuencia máxima de la señal de audio (100) y tres veces la frecuencia máxima de la señal de audio (100), o para extraer una segunda señal de paso de banda de la señal de audio antes de la dispersión del dispersor de señal adicional (202b), en la que la segunda señal de paso de banda, después de ser dispersada por el dispersor de la señal adicional (202b) y diezmada por el diezmador adicional (205b), tiene un intervalo de frecuencia que está entre dos veces la frecuencia máxima de la señal de audio (100) y tres veces la frecuencia máxima de la señal de audio (100); y

30 un combinador (111) para combinar la primera y segunda señales del paso de banda o la primera y segunda señales diezmadas con la señal de audio (100) para obtener la señal de combinación (112) extendida en su ancho de banda por un factor de 3;

en el que la primera y segunda señales del paso de banda son distorsionadas para que la primera y segunda señales del paso de banda comprendan una envolvente predeterminada;

40 o la primera y segunda señales de audio diezmadas son distorsionadas para que las primeras y segundas señales de audio diezmadas comprendan una envolvente predeterminada;

o la señal de combinación es distorsionada para que la señal de combinación comprenda una envolvente predeterminada.

45

2. El dispositivo según la reivindicación 1, en el que el dispersor de señal (102) es implementado para dispersar la señal de audio (100), de tal manera que la altura de la señal de audio no se cambia.

3. El dispositivo según una de las reivindicaciones anteriores, en el que el dispersor de señal (102) o el dispersor de señal (202b) son implementados para dispersar la señal de audio de tal manera que la duración temporal de la señal de audio es aumentada y un ancho de banda de la señal de audio dispersada es igual al ancho de banda de la señal de audio.

4. El dispositivo según una de las reivindicaciones anteriores, en el que el dispersor de señal (102) comprende un vocoder de fase (202a, 202b, 202c).

5. El dispositivo según la reivindicación 4, en el que el vocoder de fase es implementado en un banco de filtros o en una implementación de transformada de Fourier.

60 6. El dispositivo según la reivindicación 1, en el que un grupo adicional de un vocoder de fase adicional (202c), un diezmador descendente (205c) y un filtro de paso de banda descendente (207c) está presente y que son establecidos a un factor de dispersión (k), diferentes de 2 y 3, para generar una señal de paso de banda adicional que puede ser suministrada al adicionador (209).

65 7. El dispositivo según una de las reivindicaciones anteriores, en el que el filtro (107, 109) comprende un distorsionador (109) que es implementado para ejecutar la distorsión en base a parámetros espectrales transmitidos

(713) que describe una envolvente espectral de una banda superior.

8. El dispositivo según una de las reivindicaciones anteriores, que comprende además:

5 un detector transitorio (250) implementado para controlar el dispersor de señal (102) o el diezmador (105), cuando una porción transitoria es detectada en la señal de audio para ejecutar (260) una operación de copia no armónica o una operación reflejo no armónica para generar porciones espectrales más altas.

9. El dispositivo según una de las reivindicaciones anteriores, que comprende además:

10

un módulo de corrección de tonalidad/ruido (109a) que es implementado para manipular la tonalidad o ruido de la señal de paso de banda o la señal de paso de banda distorsionada.

10. El dispositivo según una de las reivindicaciones anteriores, en el que el dispersor de señal (102) comprende una pluralidad de canales de filtro, en el que cada canal de filtro comprende un filtro para generar una señal de magnitud que varía temporalmente (557) y una señal de frecuencia que varía temporalmente (560) y un oscilador (502) controlable por las señales que varían temporalmente, en el que cada canal de filtro comprende un interpolador para interpolar la señal de magnitud que varía temporalmente ( $A(t)$ ), para obtener una señal de magnitud que varía temporalmente interpolada ( $A'(t)$ ) o un interpolador para interpolar la señal de frecuencia por el factor de dispersión (104) para obtener una señal de frecuencia interpolada, y

en el que el oscilador (502) de cada canal de filtro es implementado para ser controlado por la señal de magnitud interpolada o por la señal de frecuencia interpolada.

25 11. El dispositivo según una de las reivindicaciones 1 a 11, en el que el dispersor de señal (102) comprende:

un procesador de FFT (600) para generar espectros sucesivos para superponer bloques de muestras temporales de la señal de audio, en las que los bloques superpuestos están separados entre sí por una primera distancia de tiempo (a);

30 un procesador de IFFT para transformar espectros sucesivos de un intervalo de frecuencia en el intervalo de tiempo para generar bloques superpuestos de muestras de tiempo separadas entre sí por una segunda distancia de tiempo (b) que es mayor que la primera distancia (a) y

un re-escalador de fase (606) para reescalar las fases de los valores espectrales de las secuencias de espectros de FFT generados de acuerdo con una proporción de la primera distancia (a) y la segunda distancia (b).

35

12. Un procedimiento para la extensión de ancho de banda de una señal de audio, que comprende:

generar (102) una versión de la señal de audio como una señal de tiempo dispersada temporalmente por un primer factor de dispersión de 2 para obtener una primera señal dispersada;

40

dispersar la señal de audio (100) por un segundo factor de dispersión de 3 para obtener una segunda señal dispersada;

45 diezmar (105) la primera señal dispersada por un primer factor de diezmado de 2 para obtener una primera señal de audio diezmado;

diezmar adicionalmente una segunda señal dispersada por un segundo factor de diezmado de 3 para obtener una segunda señal de audio diezmada;

50 extraer (107, 109) una primera señal de paso de banda de la primera señal de audio diezmada (106), la primera señal de paso de banda que contiene un intervalo de frecuencia que está entre una frecuencia máxima de la señal de audio (100) y dos veces la frecuencia máxima de la señal de audio (100), o extraer una primera señal de paso de banda de la señal de audio antes de generar (102), en la que la primera señal de paso de banda, después de generar (102) y diezmar (105), contiene un intervalo de frecuencia que está entre la frecuencia máxima de la señal de audio (100) y dos veces la frecuencia máxima de la señal de audio (100), extraer una segunda señal de paso de banda de la segunda señal diezmada que contiene un intervalo de frecuencia que está entre dos veces la frecuencia máxima de la señal de audio (100) y tres veces la frecuencia máxima de la señal de audio (100), o extraer una segunda señal de paso de banda de la señal de audio antes de la dispersión, en la que la segunda señal de paso de banda, después de dispersar y diezmar adicionalmente, tiene un intervalo de frecuencia que está entre dos veces la frecuencia máxima de la señal de audio (100) y tres veces la frecuencia máxima de la señal de audio (100); y

combinar (111) la primera y segunda señales de paso de banda o la primera y segunda señales diezmadas con la señal de audio (100) para obtener la señal de combinación (112) extendida en su ancho de banda por un factor de 3;

65 en la que la primera y segunda señales de paso de banda son distorsionadas para que la primera y segunda señales de paso de banda comprendan una envolvente predeterminada;

o la primera y la segunda señales de audio diezmadas son distorsionadas para que la primera y segunda señales de audio diezmadas comprendan una envolvente predeterminada;

5 o la señal de combinación se distorsiona para que la señal de combinación comprenda una envolvente predeterminada.

13. Un programa de ordenador que tiene un código de programa para efectuar el procedimiento según la reivindicación 12, cuando el programa de ordenador es ejecutado en un ordenador.

10

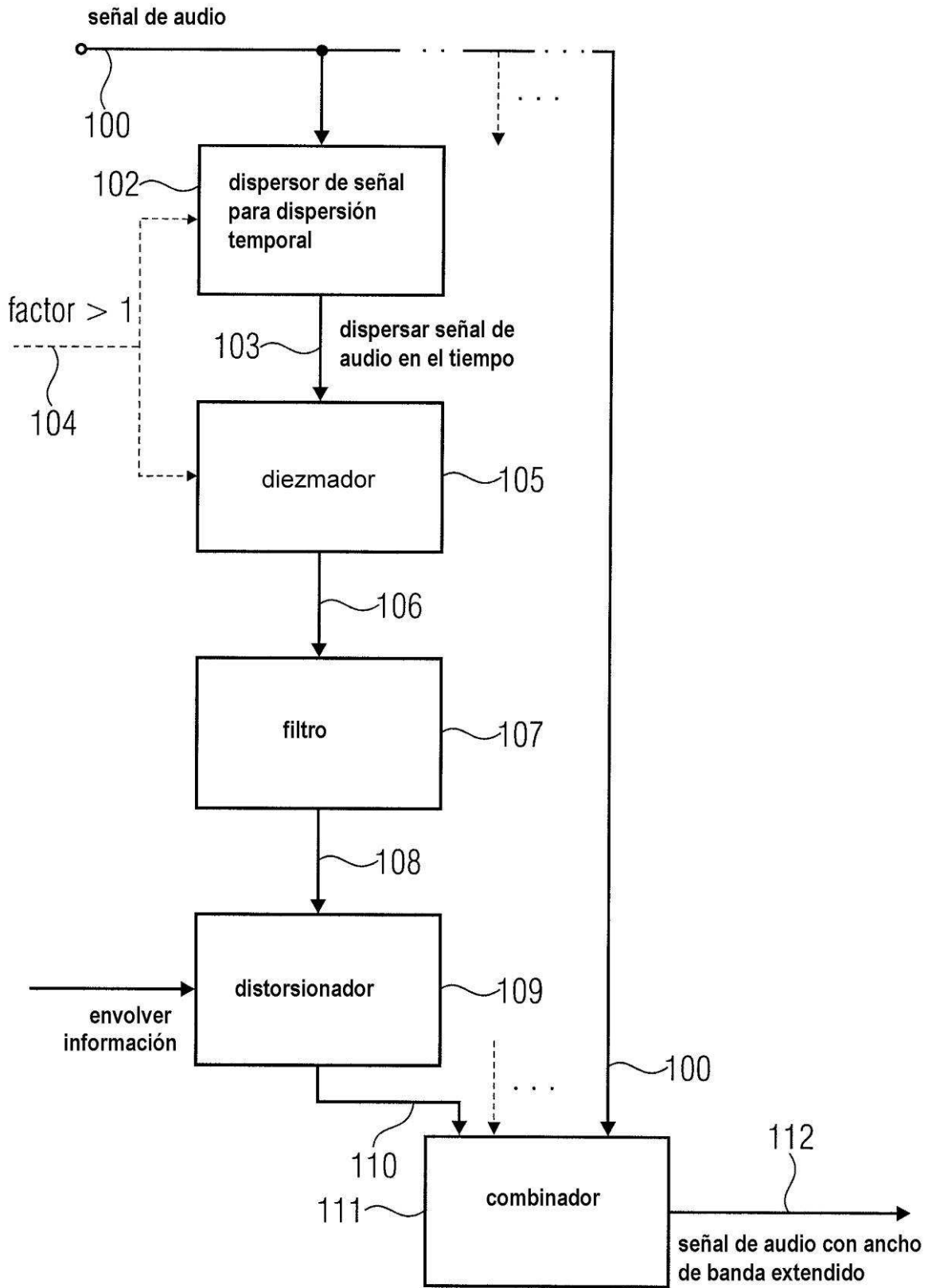


FIG 1

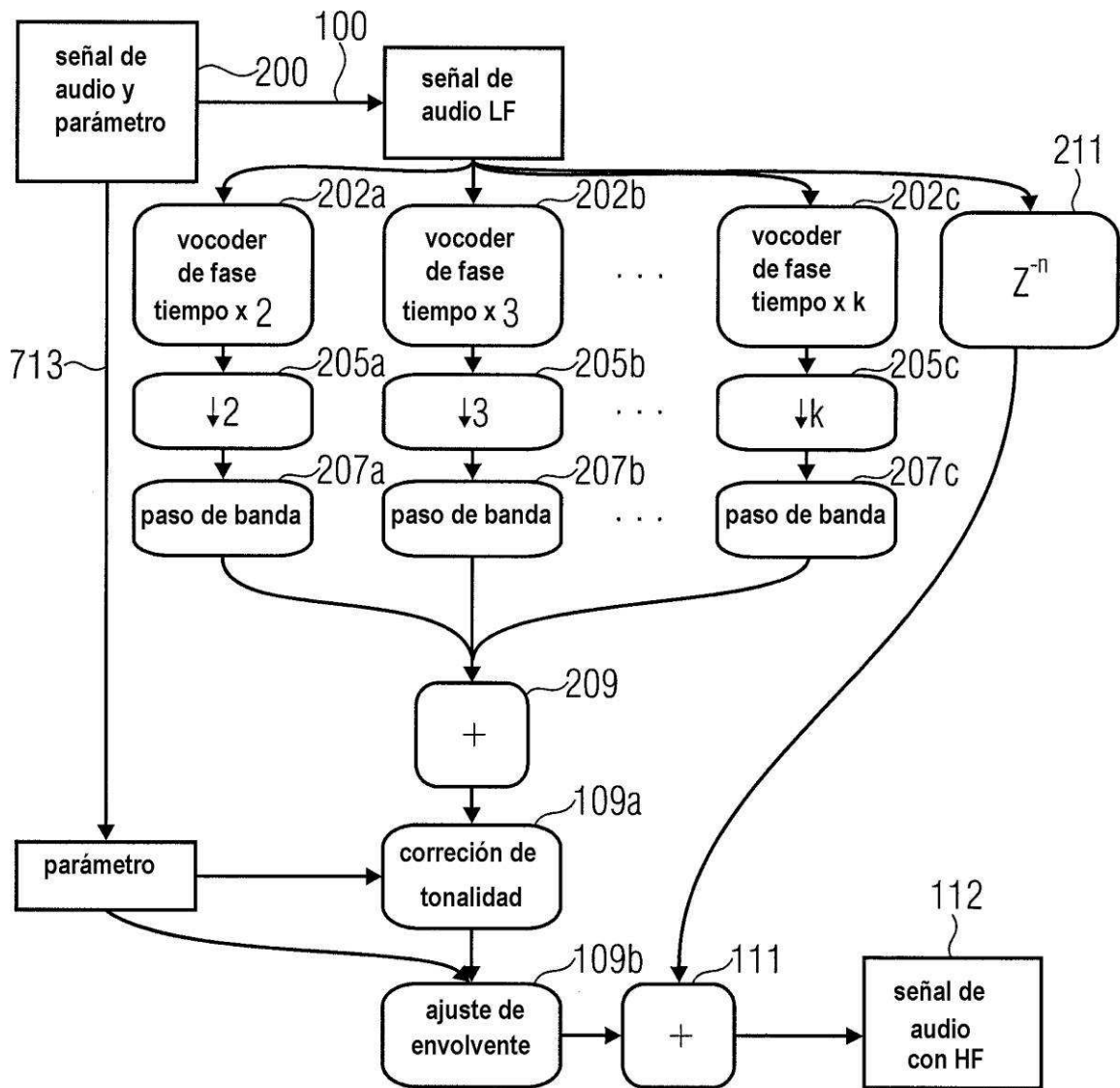


FIG 2A

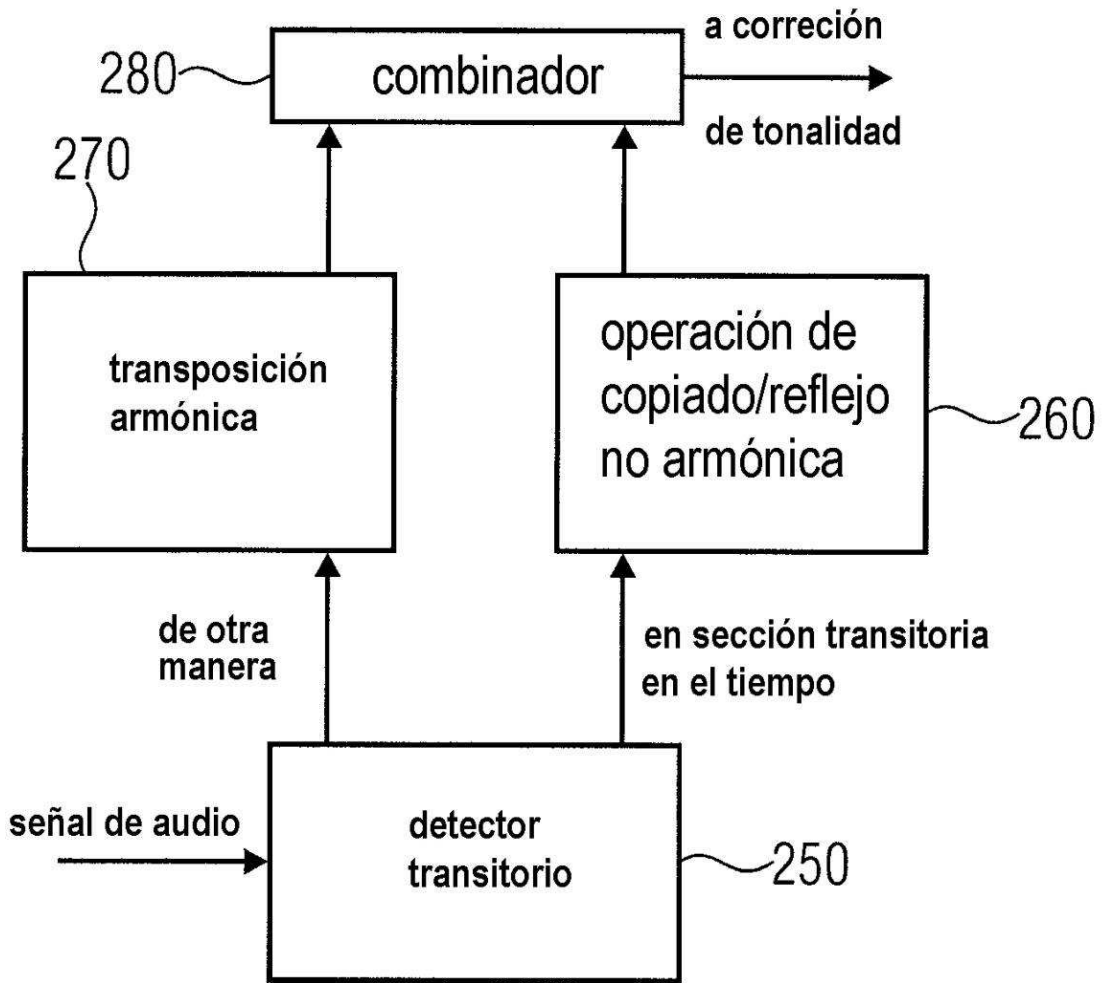


FIG 2B



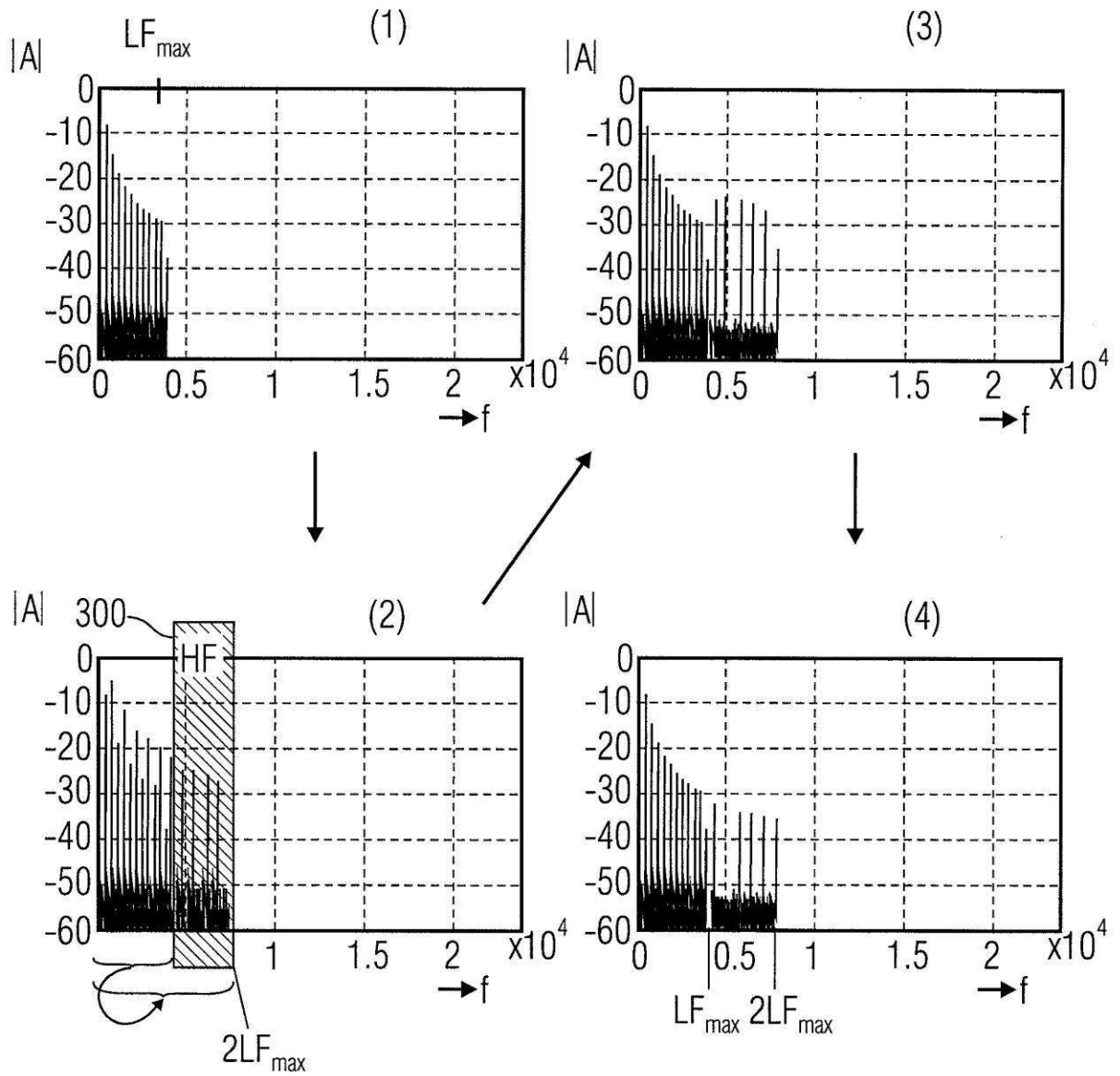


FIG 3

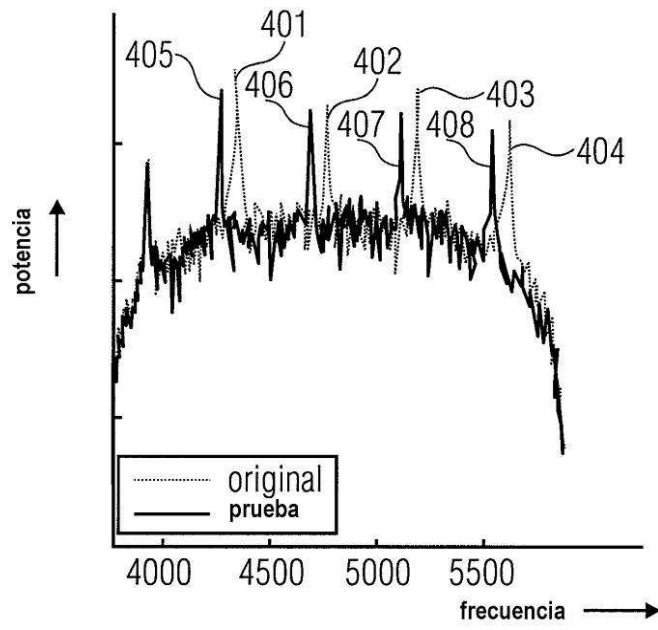


FIG 4A

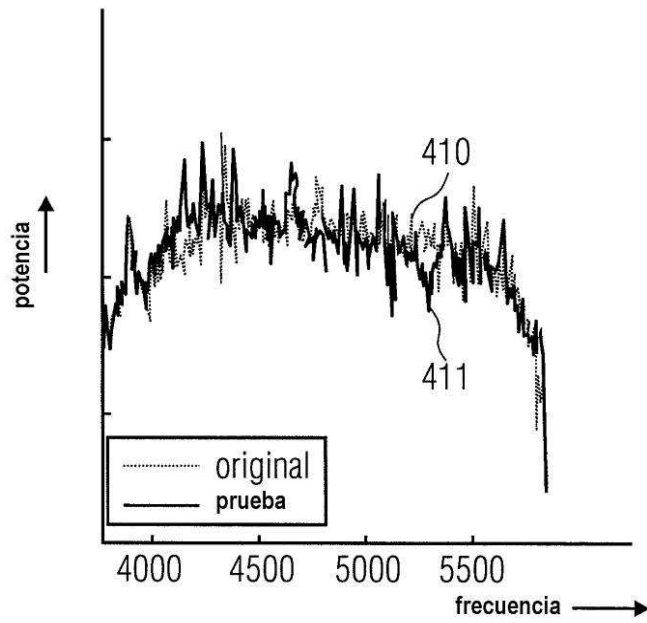
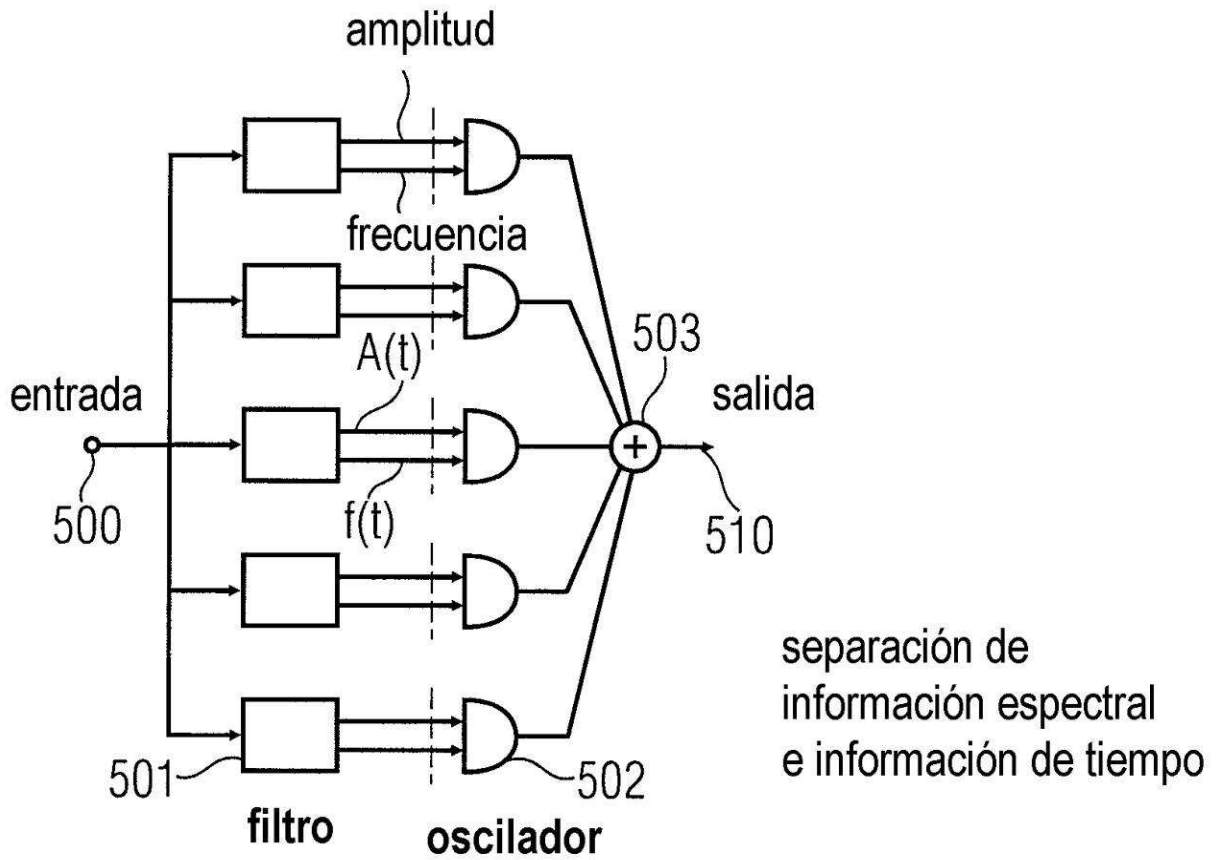


FIG 4B



**FIG 5A**

(implementación de banco de filtros)

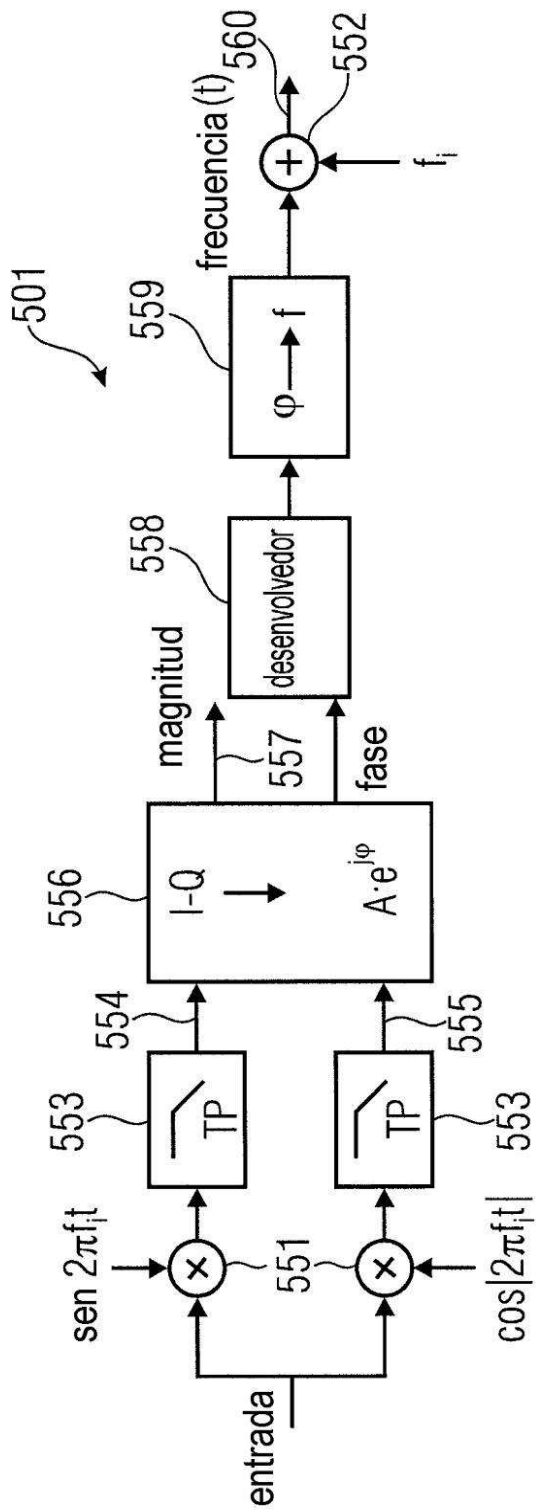


FIG 5B

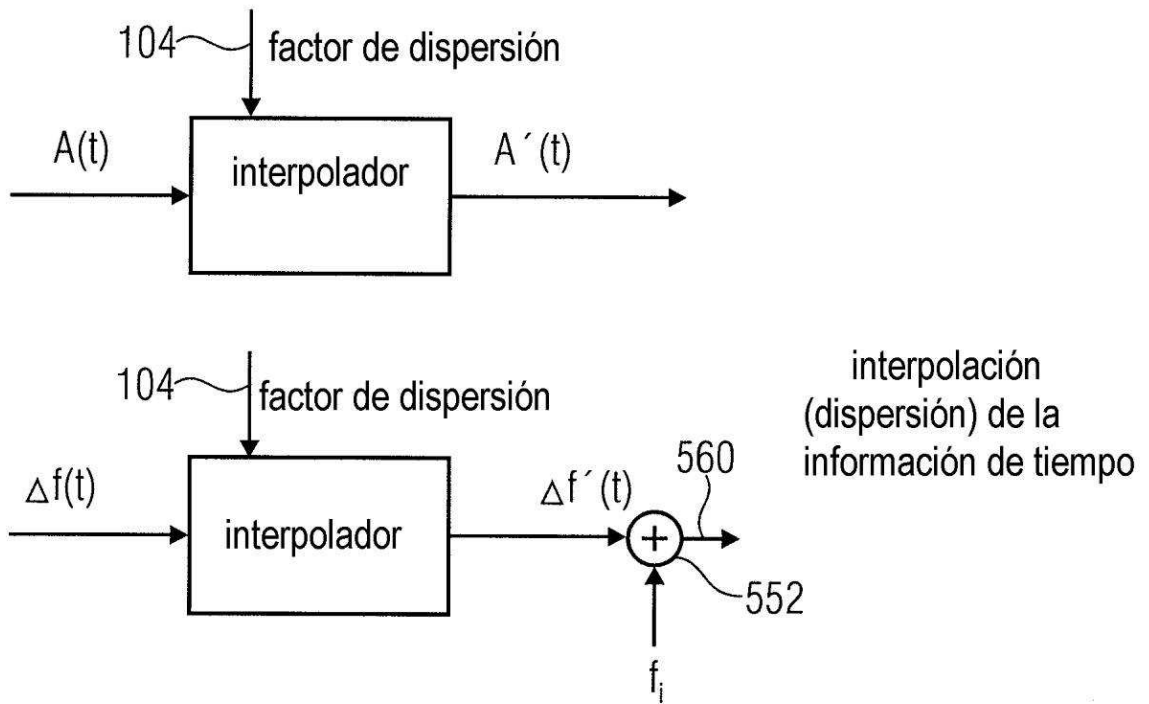
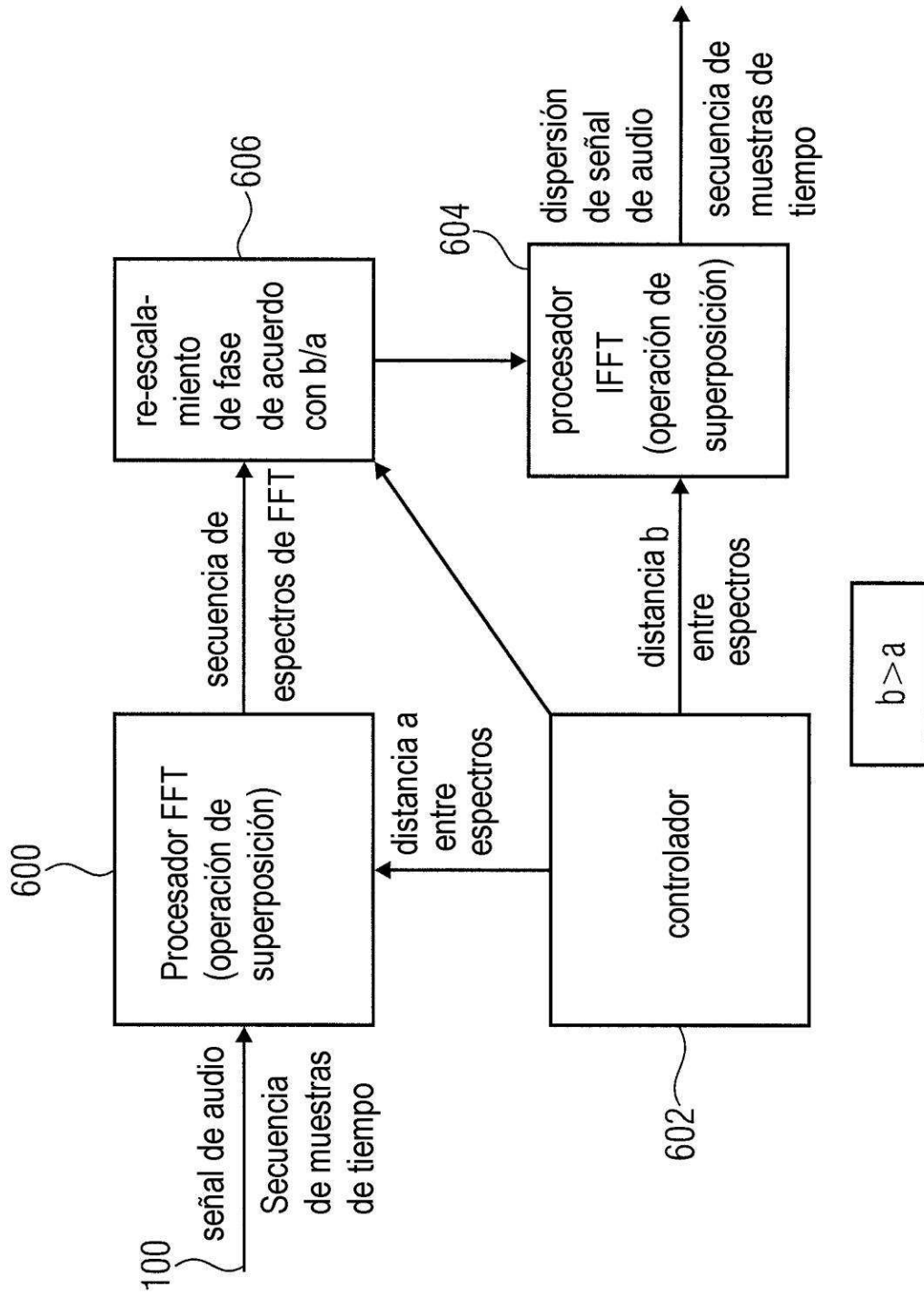
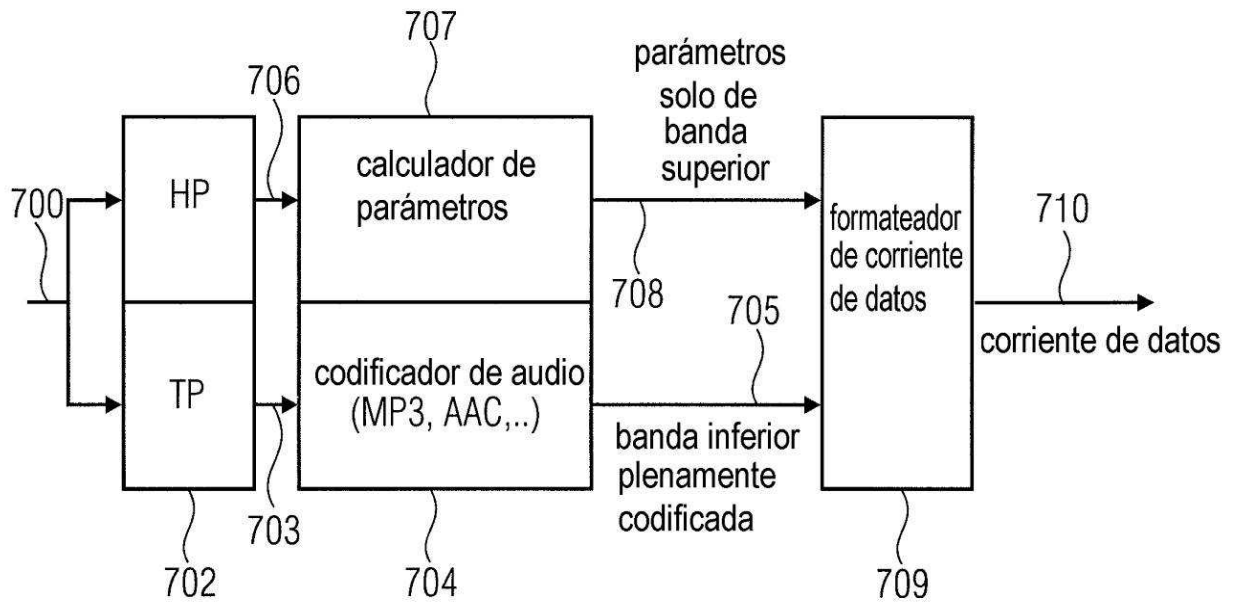


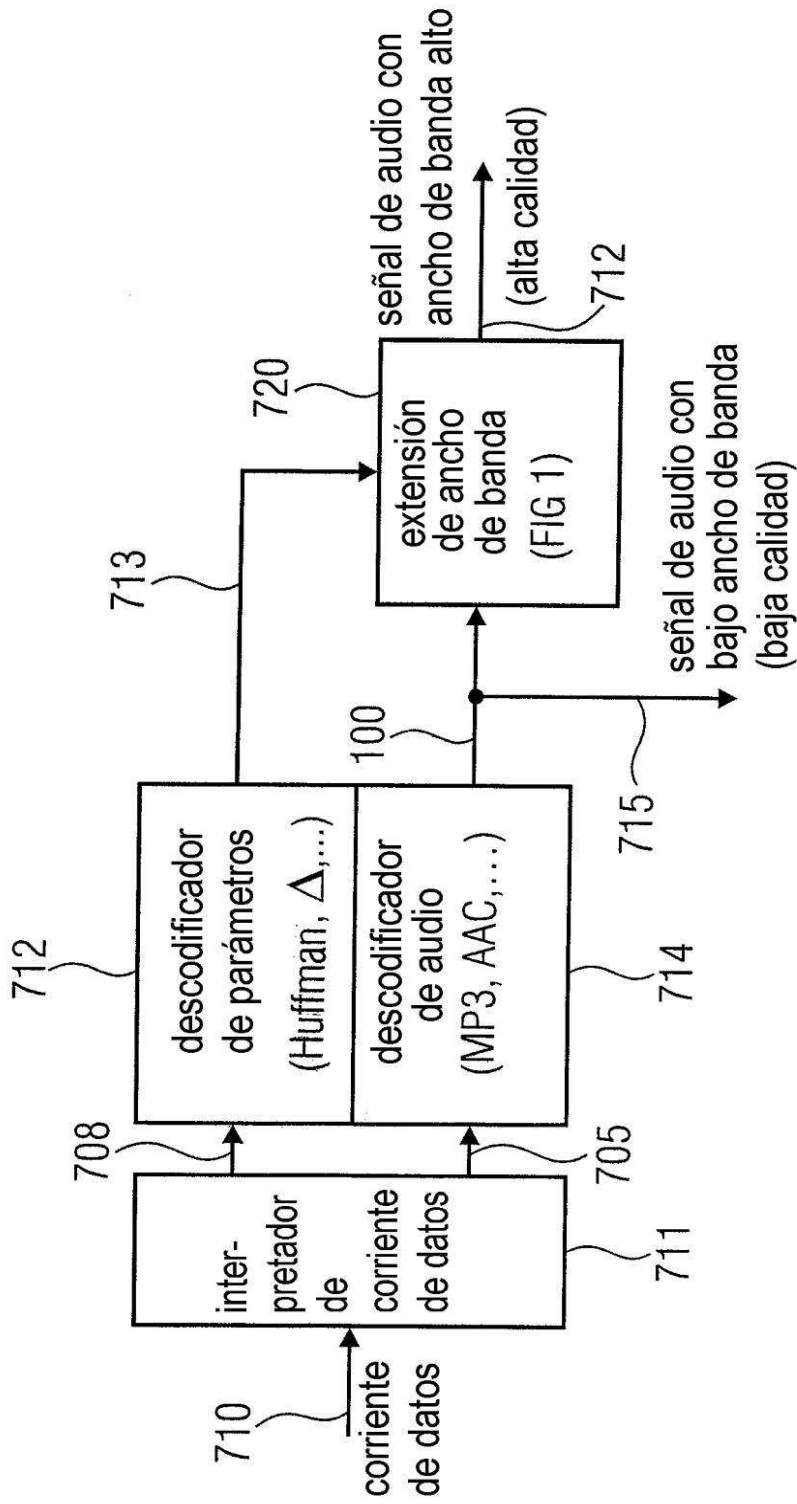
FIG 5C



**FIG 6**  
(implementación de transformación)



**FIG 7A**  
(Lado del codificador)



**FIG 7B**

(Lado del descodificador)