(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 374 486**

(51) Int. Cl.:
G10L 19/02 (2006.01)
G10L 21/02 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Número de solicitud europea: **09013051 .9**
(96) Fecha de presentación: **15.10.2009**
(97) Número de publicación de la solicitud: **2234103**
(97) Fecha de publicación de la solicitud: **29.09.2010**

(54) Título: **DISPOSITIVO Y MÉTODO PARA MANIPULAR UNA SEÑAL DE AUDIO.**

(30) Prioridad:
26.03.2009 US 163609

(45) Fecha de publicación de la mención BOPI:
17.02.2012

(45) Fecha de la publicación del folleto de la patente:
17.02.2012

(73) Titular/es:
Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.
Hansastrasse 27c
80686 München, DE

(72) Inventor/es:
Disch, Sascha;
Nagel, Frederik;
Neuendorf, Max;
Helmrich, Christian y
Zorn, Dominik

(74) Agente: **Arizti Acha, Monica**

ES 2 374 486 T3

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

DESCRIPCION

Dispositivo y método para manipular una señal de audio.

La presente invención se refiere a un esquema para manipular una señal de audio modificando fases de valores espectrales de la señal de audio tal como dentro de un esquema de extensión de ancho de banda (BWE).

En Faller, C. *et al*.: "Efficient Representation of Spatial Audio Using Perceptual Parametrization", Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop del 21-24 de octubre del 2001, Piscataway, N.J., USA, págs. 199-202, XP010566909, se describe un esquema para la colocación simultánea de varias fuentes en el espacio auditivo. El esquema se basa en una suposición acerca de la relevancia de indicaciones de ubicación en diferentes bandas críticas. Dada la señal de suma de varias fuentes, es decir una señal monofónica, y un conjunto de parámetros (información secundaria) el esquema puede generar una señal binaural colocando en el espacio las fuentes contenidas en la señal monofónica. Potenciales aplicaciones para el esquema son la realización de conferencias de sobremesa con múltiples hablantes y la codificación de audio.

El documento WO 2007/016107 A2 da a conocer un método de codificación de audio en el que un codificador recibe una pluralidad de canales de entrada y genera uno o más canales de salida de audio y uno o más parámetros que describen las relaciones espaciales deseadas entre una pluralidad de canales de audio que pueden derivarse del uno o más canales de salida de audio. El método comprende detectar cambios en características de señal con respecto al tiempo en uno o más de la pluralidad de canales de entrada de audio, identificar como límites de evento auditivo cambios en características de señal con respecto al tiempo en el uno o más de la pluralidad de canales de entrada de audio, un segmento de audio entre límites consecutivos que constituyen un evento auditivo en el canal o canales, y generar todos o algunos de los uno o más parámetros al menos parcialmente en respuesta a los eventos auditivos y/o al grado de cambio en las características de señal asociadas con los límites de evento auditivo. También se da a conocer un método de mezclado ascendente o mezclador ascendente de audio sensible al evento auditivo.

El documento US 6.549.884 B1 da a conocer un sistema para la modulación de tono de una señal de audio en el que se realiza un nuevo muestreo en el dominio de frecuencia. El sistema incluye un método para la modulación de tono de una señal convirtiendo la señal a una representación del dominio de frecuencia e identificando entonces una región específica en la representación del dominio de frecuencia. La región está ubicada en una primera ubicación de frecuencia. A continuación, la región se desplaza a una segunda ubicación de frecuencia para formar una representación del dominio de frecuencia ajustada. Finalmente, la representación del dominio de frecuencia ajustada se transforma a una señal en el dominio del tiempo que representa la señal de entrada con tono modulado.

En Disch, S. *et al.*: "An Amplitude- and Frequency- Modulation Vocoder for Audio Signal Processing", Proc. 11th International Conference on Digital Audio Effects (DAFX-08), 1 de septiembre del 2008, págs. 1-7, XP002574848, se describe un método de análisis/síntesis que proporciona medios apropiados para reducir las diferencias entre procesamiento de audio de forma de onda y paramétrico posibilitando una transición sin interrupción entre ambos enfoques.

Con frecuencia el almacenamiento o transmisión de señales de audio se somete a restricciones estrictas de tasa de bits. En el pasado, se forzó a los codificadores a reducir drásticamente el ancho de banda de audio transmitido cuando estaba disponible sólo una tasa de bits muy baja. En la actualidad, códecs de audio de módem pueden codificar señales de banda ancha usando métodos de extensión de ancho de banda, tal como se describe en M. Dietz, L. Liljeryd, K. Kjörling y O. Kunz, "Spectral Band Replication, a novel approach in audio coding", en 112th AES Convention, Munich, mayo del 2002; S. Meltzer, R. Böhm y F. Henn, "SBR enhanced audio codecs for digital broadcasting such as "Digital Radio Mondiale" (DRM)", en 112th AES Convention, Munich, mayo del 2002; T. Ziegler, A. Ehret, P. Ekstrand y M. Lutzky, "Enhancing mp3 with SBR: Features and Capabilities of the new mp3PRO Algorithm", en 112th AES Convention, Munich, mayo del 2002; International Standard ISO/IEC 14496-3:2001/FPDAM 1, "Bandwidth extension", ISO/IEC, 2002. Speech bandwidth extension method and apparatus Vasu Iyengar *et al.*; E. Larsen, R. M. Aarts, y M. Danessis. Efficient high-frequency bandwidth extension of music and speech. En AES 112th Convention, Munich, Alemania, mayo del 2002; R. M. Aarts, E. Larsen y O. Ouweltjes. A unified approach to low-and high frequency bandwidth extension. En AES 115th Convention, Nueva York, Estados Unidos, octubre del 2003; K. Käyhkö. A Robust Wideband Enhancement for Narrowband Speech Signal. Research Report, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, 2001; 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. En AES 112th Convention, Munich, Alemania, mayo del 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 estadounidense 08/951.029, Ohmori, *et al*. Audio band width extending system and method y patente estadounidense 6895375, Malah, D & Cox, R. V.: System for bandwidth extension of Narrow-band speech. Estos algoritmos se basan en una representación paramétrica del contenido de alta frecuencia (HF), que se genera a partir de la parte de baja frecuencia (LF) codificada de forma de onda de la señal decodificada por medio de transposición en la región espectral de HF ("parcheado") y aplicación de un posprocesamiento controlado por parámetros.

Recientemente, un nuevo algoritmo que emplea vocodificador de fase como se describe, por ejemplo, en M. Puckette.

Phase-locked Vocoder. IEEE ASSP Conference on Applications of Signal Processing to Audio and Acoustics, Mohonk 1995.", Röbel, A.: Transient detection and preservation in the phase vocoder; citeseer.ist.psu.edu/679246.html; Laroche L., Dolson M.: "Improved phase vocoder timescale modification of audio", IEEE Trans. Speech and Audio Processing, vol. 7, n.º 3, págs. 323-332 y la patente estadounidense 6549884 Laroche, J. & Dolson, M.: Phase-vocoder pitch-shifting for the patch generation, se ha presentado en Frederik Nagel, Sascha Disch, "A harmonic bandwidth extension method for audio codecs", ICASSP International Conference on Acoustics, Speech and Signal Processing, IEEE CNF, Taipei, Taiwán, abril del 2009. Sin embargo, este método denominado "harmonic bandwidth extension" (HBE) es propenso a degradaciones de calidad de transitorios contenidos en la señal de audio, tal como se describe en Frederik Nagel, Sascha Disch, Nikolaus Rettelbach, "A phase vocoder driven bandwidth extension method with novel transient handling for audio codecs", 126th AES Convention, Munich, Alemania, mayo del 2009, puesto que no se garantiza que la coherencia vertical sobre subbandas se conserve en el algoritmo de vocodificador de fase convencional y, además, el nuevo cálculo de las fases de transformada discreta de Fourier (DFT) debe realizarse en bloques de tiempo aislados de una transformación que supone implícitamente periodicidad circular.

Se conoce que pueden observarse específicamente dos clases de artefactos debido al procesamiento de vocodificador de fase basado en bloque. Éstas, en particular, son la dispersión de la forma de onda y el solapamiento (*aliasing*) temporal debido a efectos de convolución cíclica temporales de la señal debido a la aplicación de fases nuevamente calculadas.

En otras palabras, debido a la aplicación de una modificación de fase en los valores espectrales de la señal de audio en el algoritmo de BWE, un transitorio contenido en un bloque de la señal de audio puede envolverse alrededor del bloque, es decir convolucionarse cíclicamente de vuelta al bloque. Esto da como resultado solapamiento temporal y, por consiguiente, lleva a una degradación de la señal de audio.

Por tanto, deben emplearse métodos para un tratamiento especial para partes de señal que contienen transitorios. Sin embargo, puesto que especialmente se ejecuta el algoritmo de BWE en el lado de decodificador de una cadena de códec, la complejidad computacional es un tema importante. Por consiguiente, las medidas contra la degradación de señal de audio recién mencionadas no deben realizarse a costa de una complejidad computacional muy aumentada.

El objeto de la presente invención es proporcionar un esquema para manipular una señal de audio modificando fases de valores espectrales de la señal de audio, por ejemplo, en el contexto de un esquema BWE que posibilita lograr un mejor equilibrio entre reducción de la degradación recién mencionada y la complejidad computacional.

Este objeto se logra mediante un dispositivo según la reivindicación 1 o un método según la reivindicación 17, o un programa informático según la reivindicación 18.

La idea básica subyacente a la presente invención es que el mejor equilibrio mencionado anteriormente puede lograrse cuando al menos un bloque rellenado de muestras de audio que tiene valores rellenados y valores de señal de audio se genera antes de modificar las fases de los valores espectrales del bloque rellenado. Mediante esta medida, puede evitarse que se produzca una desviación del contenido de señal hacia los bordes de bloque debido a la modificación de fase y un solapamiento de tiempo correspondiente o al menos conseguir que sea menos probable, y por tanto se mantiene la calidad de audio con poco esfuerzo.

El concepto de la invención para manipular una señal de audio se basa en generar una pluralidad de bloques consecutivos de muestras de audio, comprendiendo la pluralidad de bloques consecutivos al menos un bloque rellenado de muestras de audio, teniendo el bloque rellenado valores rellenados y valores de señal de audio. El bloque rellenado se convierte entonces en una representación espectral que tiene valores espectrales. Los valores espectrales se modifican entonces para obtener una representación espectral modificada. Finalmente, la representación espectral modificada se convierte en una señal de audio en el dominio del tiempo modificada. Entonces, puede eliminarse el intervalo de valores que se usó para el rellenado.

Según una realización de la presente invención, el bloque rellenado se genera insertando valores rellenados que preferiblemente consisten en valores cero antes o después de un bloque de tiempo.

Según una realización, los bloques rellenados se limitan a aquéllos que contienen un evento transitorio, limitando así la sobrecarga de complejidad computacional adicional para estos eventos. De manera más precisa, un bloque se procesa, por ejemplo, de una forma avanzada mediante un algoritmo de BWE, cuando se detecta un evento transitorio en este bloque de la señal de audio, en forma de un bloque rellenado, mientras se procesa otro bloque de la señal de audio como un bloque no rellenado que tiene sólo valores de señal de audio de una manera convencional de un algoritmo de BWE cuando no se detecta el evento transitorio en el bloque. Conmutando de manera adaptativa entre procesamiento convencional y procesamiento avanzado, puede reducirse significativamente el esfuerzo computacional promedio, lo que permite, por ejemplo, una velocidad y memoria de procesador reducidas.

Según realizaciones de la presente invención, los valores rellenados están dispuestos antes y/o después de un bloque de tiempo en que se detecta un evento transitorio, de modo que el bloque rellenado se adapta a una conversión entre el dominio del tiempo y de la frecuencia mediante un primer y un segundo convertidor, realizado, por ejemplo, a través de un procesador de DFT y de IDFT, respectivamente. Una solución preferible sería disponer el rellenado simétricamente alrededor del bloque de tiempo.

Según una realización, el al menos un bloque rellenado se genera añadiendo valores rellenados tales como valores cero a un bloque de muestras de audio de la señal de audio. Alternativamente, se usa una función de ventana de análisis que tiene al menos una zona de guarda añadida a una posición inicial de la función de ventana o una posición final de la función de ventana para formar un bloque rellenado aplicando esta función de ventana de análisis a un bloque de muestras de audio de la señal de audio. La función de ventana puede comprender, por ejemplo, una ventana de Hann con zonas de guarda.

En lo sucesivo, se explican realizaciones de la presente invención con referencia a los dibujos adjuntos, en los que:

la figura 1 muestra un diagrama de bloques de una realización para manipular una señal de audio;

la figura 2 muestra un diagrama de bloques de una realización para realizar una extensión de ancho de banda usando la señal de audio;

la figura 3 muestra un diagrama de bloques de una realización para ejecutar un algoritmo de extensión de ancho de banda usando diferentes factores de BWE;

la figura 4 muestra un diagrama de bloques de una realización adicional para convertir un bloque rellenado o un bloque no rellenado usando un detector de transitorios;

la figura 5 muestra un diagrama de bloques de una implementación de una realización de la figura 4;

la figura 6 muestra un diagrama de bloques de una implementación adicional de una realización de la figura 4;

la figura 7a muestra un gráfico de un bloque de señal a modo de ejemplo antes y después de una modificación de fase para ilustrar un efecto de una modificación de fase en una forma de onda de señal con un transitorio centrado en un bloque de tiempo;

la figura 7b muestra un gráfico de un bloque de señal a modo de ejemplo antes y después de una modificación de fase para ilustrar un efecto de una modificación de fase en una forma de onda de señal con el transitorio en las proximidades de una primera muestra de un bloque de tiempo;

la figura 8 muestra un diagrama de bloques de una visión general de una realización adicional de la presente invención;

la figura 9a muestra un gráfico de una función de ventana de análisis a modo de ejemplo en forma de una ventana de Hann con zonas de guarda estando caracterizadas las zonas de guarda por ceros constantes, usándose la ventana en una realización alternativa de la presente invención;

la figura 9b muestra un gráfico de una función de ventana de análisis a modo de ejemplo en forma de una ventana de Hann con zonas de guarda estando caracterizadas las zonas de guarda por oscilaciones, usándose la ventana en una realización alternativa adicional de la presente invención;

la figura 10 muestra una ilustración esquemática para una manipulación de una banda espectral de una señal de audio en un esquema de extensión de ancho de banda;

la figura 11 muestra una ilustración esquemática para una operación de suma por solapamiento en el contexto de un esquema de extensión de ancho de banda;

la figura 12 muestra un diagrama de bloques y una ilustración esquemática para una implementación de una realización alternativa basada en la figura 4; y

la figura 13 muestra un diagrama de bloques de una implementación de extensión de ancho de banda armónica típica (HBE).

La figura 1 ilustra un aparato para manipular una señal de audio según una realización de la presente invención. El aparato comprende un divisor 102 en ventanas, que tiene una entrada 100 para una señal de audio. El divisor 102 en ventanas se implementa para generar una pluralidad de bloques consecutivos de muestras de audio, que comprende al menos un bloque rellenado. El bloque rellenado, en particular, tiene valores rellenados y valores de señal de audio. El bloque rellenado presente en una salida 103 del divisor 102 en ventanas se suministra a un primer convertidor 104, que se implementa para convertir el bloque 103 rellenado en una representación espectral que tiene valores espectrales. Los valores espectrales en la salida 105 del primer convertidor 104 se suministran entonces a un modificador 106 de fase. El modificador 106 de fase se implementa para modificar fases de los valores 105 espectrales para obtener una representación espectral modificada en 107. La salida 107 se suministra finalmente a un segundo convertidor 108, que se implementa para convertir la representación 107 espectral modificada en una señal 109 de audio en el dominio del tiempo modificada. La salida 109 del segundo convertidor 108 puede conectarse a un diezmador adicional, que se requiere para un esquema de extensión de ancho de banda, tal como se trata en relación con las figuras 2, 3 y 8.

La figura 2 muestra una ilustración esquemática de una realización para ejecutar un algoritmo de extensión de ancho de banda usando un factor de extensión de ancho de banda (σ). En este caso, la señal 100 de audio se alimenta al divisor

102 en ventanas, que comprende un procesador 110 de ventana de análisis y un rellenador 112 posterior. En una realización, el procesador 110 de ventana de análisis se implementa para generar una pluralidad de bloques consecutivos que tienen el mismo tamaño. La salida 111 del procesador 110 de ventana de análisis se conecta adicionalmente al rellenador 112. En particular, el rellenador 112 se implementa para rellenar un bloque de la pluralidad de bloques consecutivos en la salida 111 del procesador 110 de ventana de análisis para obtener el bloque rellenado en la salida 103 del rellenador 112. En este caso, el bloque rellenado se obtiene insertando valores rellenados en posiciones temporales específicas antes de una primera muestra de bloques consecutivos de muestras de audio o después de una última muestra del bloque consecutivo de muestras de audio. El bloque 103 rellenado se convierte adicionalmente por el primer convertidor 104 para obtener una representación espectral en la salida 105. Adicionalmente, se usa un filtro 114 paso banda, que se implementa para extraer la señal 113 paso banda de la representación 105 espectral o la señal 100 de audio. Una característica paso banda del filtro 114 paso banda se selecciona de manera que la señal 113 paso banda se limita a un intervalo de frecuencia objetivo apropiado. En este caso, el filtro 114 paso banda recibe un factor de extensión de ancho de banda (σ) que también está presente en la salida 115 de un modificador 106 de fase aguas abajo. En una realización de la presente invención, se usa un factor de extensión de ancho de banda (σ) de 2,0 para ejecutar el algoritmo de extensión de ancho de banda. En caso de que la señal 100 de audio tenga, por ejemplo, un intervalo de frecuencia de 0 a 4 kHz, el filtro 114 paso banda extraerá el intervalo de frecuencia de 2 a 4 kHz, de modo que la señal 113 paso banda se transformará mediante el algoritmo de BWE posterior a un intervalo de frecuencia objetivo de 4 a 8 kHz siempre que, por ejemplo, el factor de extensión de ancho de banda (σ) de 2,0 se aplique para seleccionar un filtro 114 paso banda apropiado (véase la figura 10). La representación espectral de la señal paso banda en la salida 113 del filtro 114 paso banda comprende información de amplitud e información de fase, que se procesa adicionalmente en un ajustador 116 a escala y el modificador 106 de fase, respectivamente. El ajustador 116 a escala se implementa para ajustar a escala los valores 113 espectrales de la información de amplitud por un factor, dependiendo el factor de una característica de suma por solapamiento en la medida en que se tiene en cuenta una relación de una primera distancia de tiempo (a) para una suma por solapamiento aplicada por el divisor 102 en ventanas y una distancia de tiempo (b) diferente aplicada por un sumador 124 por solapamiento aguas abajo.

Por ejemplo, si existe una característica de suma por solapamiento con una suma por solapamiento de seis veces de bloques consecutivos de muestras de audio que tienen la primera distancia de tiempo (a), y una relación de la segunda distancia de tiempo (b) con respecto a la primera distancia de tiempo (a) de b/a=2, entonces el factor de b/a x 1/6 se aplicará mediante el ajustador 116 a escala para ajustar a escala los valores espectrales en la salida 113 (véase la figura 11) suponiendo una ventana de análisis rectangular.

Sin embargo, este ajuste a escala de amplitud específico sólo puede aplicarse cuando se realiza un diezmado aguas abajo posteriormente a la suma por solapamiento. En caso de que se realice el diezmado antes de la suma por solapamiento, el diezmado puede tener un efecto sobre las amplitudes de los valores espectrales que generalmente debe tener en cuenta el ajustador 116 a escala.

El modificador 106 de fase está configurado para ajustar a escala o multiplicar, respectivamente, las fases de los valores 113 espectrales de la banda de la señal de audio por el factor de extensión de ancho de banda (σ), de modo que al menos una muestra de un bloque consecutivo de muestras de audio se convoluciona cíclicamente en el bloque.

El efecto de convolución cíclica basada en una periodicidad circular, que es un efecto secundario no deseado de la conversión por el primer convertidor 104 y el segundo convertidor 108 se muestra en la figura 7 mediante el ejemplo de un transitorio 700 centrado en la ventana 704 de análisis (figura 7a) y un transitorio 702 en las proximidades de un borde de la ventana 704 de análisis (figura 7b).

La figura 7a muestra el transitorio 700 centrado en la ventana 704 de análisis, es decir dentro del bloque consecutivo de muestras de audio que tiene una longitud 706 de muestra que incluye, por ejemplo, 1001 muestras con una primera muestra 708 y una última muestra 710 del bloque consecutivo. La señal 700 original se indica mediante una delgada línea discontinua. Tras la conversión por el primer convertidor 104 y la aplicación posterior de una modificación de fase, por ejemplo, mediante el uso de un vocodificador de fase al espectro de la señal original, el transitorio 700 se desplazará y convolucionará cíclicamente de vuelta a la ventana 704 de análisis tras la conversión por el segundo convertidor 108, es decir de manera que el transitorio 701 cíclicamente convolucionado aún se ubicará dentro de la ventana 704 de análisis. El transitorio 701 cíclicamente convolucionado se indica mediante la línea gruesa denominada “no de guarda”.

La figura 7b muestra la señal original que contiene un transitorio 702 cerca de la primera muestra 708 de la ventana 704 de análisis. La señal original que tiene un transitorio 702, de nuevo, se indica mediante la delgada línea discontinua. En este caso, tras la conversión por el primer convertidor 104 y la aplicación posterior de la modificación de fase, el transitorio 702 se desplazará y convolucionará cíclicamente de vuelta a la ventana 704 de análisis tras la conversión por el segundo convertidor 108, de modo que se obtendrá un transitorio 703 cíclicamente convolucionado, que se indica mediante la línea gruesa denominada “no de guarda”. En este caso, se genera el transitorio 703 cíclicamente convolucionado porque se desplaza al menos una parte del transitorio 702 antes de la primera muestra 708 de la ventana 704 de análisis debido a la modificación de fase, que da como resultado una envoltura circular del transitorio 703 cíclicamente convolucionado. En particular, tal como puede observarse en la figura 7b, la parte del transitorio 702 que se desplaza fuera de la ventana 704 de análisis se produce de nuevo (parte 705) a la izquierda de la última muestra

710 de la ventana 704 de análisis debido al efecto de periodicidad circular.

La representación espectral modificada que comprende la información de amplitud modificada de la salida 117 del ajustador 116 a escala y la información de fase modificada de la salida 107 del modificador 106 de fase se suministran al segundo convertidor 108, que está configurado para convertir la representación espectral modificada en la señal de audio en el dominio del tiempo modificada presente en la salida 109 del segundo convertidor 108. La señal de audio en el dominio del tiempo modificada en la salida 109 del segundo convertidor 108 puede suministrarse entonces a un eliminador 118 de relleno. El eliminador 118 de relleno se implementa para eliminar aquellas muestras de la señal de audio en el dominio del tiempo modificada, que corresponden a las muestras de los valores rellenados insertados para generar el bloque rellenado en la salida 103 del divisor 102 en ventanas antes de que la modificación de fase se aplique mediante el procesamiento aguas abajo del modificador 106 de fase. De manera más precisa, se eliminan muestras en aquellas posiciones temporales de la señal de audio en el dominio del tiempo modificada, que corresponden a las posiciones temporales especificadas para las que se insertan valores rellenados antes de la modificación de fase.

En una realización de la presente invención, los valores rellenados se insertan simétricamente antes de la primera muestra 708 del bloque consecutivo y después de la última muestra 710 del bloque consecutivo de muestras de audio, tal como se muestra, por ejemplo, en la figura 7, de modo que se forman dos zonas 712, 714 de guarda simétricas, que encierran el bloque consecutivo centrado que tiene la longitud 706 de muestra. En este caso simétrico, las zonas de guarda o “intervalos 712, 714 de guarda”, respectivamente, pueden eliminarse preferiblemente del bloque rellenado mediante el eliminador 118 de relleno tras la modificación de fase de los valores espectrales y su conversión posterior en la señal de audio en el dominio del tiempo modificada, para obtener el bloque consecutivo sólo sin los valores rellenados en la salida 119 del eliminador 118 de relleno.

En una implementación alternativa, puede ser que los intervalos de guarda no se eliminen mediante el eliminador 118 de relleno de la salida 109 del segundo convertidor 108, de modo que la señal de audio en el dominio del tiempo modificada del bloque rellenado tendrá la longitud 716 de muestra incluyendo la longitud 706 de muestra del bloque consecutivo centrado y las longitudes 712, 714 de muestra de los intervalos de guarda. Esta señal puede procesarse adicionalmente en etapas de procesamiento posteriores hasta un sumador 124 por solapamiento, tal como se muestra en el diagrama de bloques de la figura 2. En el caso de que el eliminador 118 de relleno no esté presente, este procesamiento, incluyendo la operación en los intervalos de guarda, también puede interpretarse como un sobremuestreo de la señal. Aunque no se requiera el eliminador 118 de relleno en realizaciones de la presente invención, es ventajoso usarlo tal como se muestra en la figura 2, porque la señal presente en la salida 119 ya tendrá la misma longitud de muestra que el bloque consecutivo original o bloque no rellenado, respectivamente, presente en la salida 111 del procesador 110 de ventana de análisis antes del rellenado mediante el rellenador 112. Por tanto, las etapas de procesamiento posteriores se adaptarán fácilmente a la señal en la salida 119.

Preferiblemente, la señal de audio en el dominio del tiempo modificada en la salida 119 del eliminador 118 de relleno se suministra a un diezmador 120. El diezmador 120 se implementa preferiblemente mediante un convertidor de tasa de bits de muestra simple que opera usando el factor de extensión de ancho de banda (σ) para obtener una señal en el dominio del tiempo diezmada en la salida 121 del diezmador 120. En este caso, la característica de diezmado depende de la característica de modificación de fase provista por el modificador 106 de fase en la salida 115. En una realización de la presente invención, el factor de extensión de ancho de banda $\sigma=2$ se suministra por el modificador 106 de fase a través de la salida 115 al diezmador 120, de modo que se eliminará cada segunda muestra de la señal de audio en el dominio del tiempo modificada en la salida 119, dando como resultado la señal en el dominio del tiempo diezmada presente en la salida 121.

La señal en el dominio del tiempo diezmada presente en la salida 121 del diezmador 120 se alimenta posteriormente a un divisor 122 en ventanas de síntesis, que se implementa para aplicar una función de ventana de síntesis, por ejemplo, a la señal en el dominio del tiempo diezmada, donde la función de ventana de síntesis se hace coincidir con una función de análisis aplicada por el procesador 110 de ventana de análisis del divisor 102 en ventanas. En este caso, la función de ventana de síntesis puede hacerse coincidir con la función de análisis de tal manera que la aplicación de la función de síntesis compensa el efecto de la función de análisis. Alternativamente, el divisor 122 en ventanas de síntesis también puede implementarse para operar en la señal de audio en el dominio del tiempo modificada en la salida 109 del segundo convertidor 108.

La señal en el dominio del tiempo diezmada y dividida en ventanas a partir de la salida 123 del divisor 122 en ventanas de síntesis se suministra entonces a un sumador 124 por solapamiento. En este caso, el sumador 124 por solapamiento recibe información acerca de la primera distancia de tiempo para la operación de suma por solapamiento (a) aplicada por el divisor 102 en ventanas y el factor de extensión de ancho de banda (σ) aplicado por el modificador 106 de fase en la salida 115. El sumador 124 por solapamiento aplica una distancia de tiempo (b) diferente que es mayor que la primera distancia de tiempo (a) a la señal en el dominio del tiempo diezmada y dividida en ventanas. En caso de que el diezmado se realice tras la suma por solapamiento, puede cumplirse la condición $\sigma=b/a$ según un esquema de extensión de ancho de banda. Sin embargo, en la realización tal como se muestra en la figura 2, el diezmado se realiza antes de la suma por solapamiento, de modo que el diezmado puede tener un efecto en la condición anterior que generalmente debe tenerse en cuenta por el sumador 124 por solapamiento.

Preferiblemente, el aparato mostrado en la figura 2 está configurado para ejecutar un algoritmo de BWE, que

comprende un factor de extensión de ancho de banda (σ), controlando el factor de extensión de ancho de banda (σ) una expansión de frecuencia de una banda de la señal de audio a una banda de frecuencia objetivo. De esta manera, la señal en el intervalo de frecuencia objetivo que depende del factor de extensión de ancho de banda (σ) puede obtenerse en la salida 125 del sumador 124 por solapamiento.

En el contexto de un algoritmo de BWE, se implementa un sumador 124 por solapamiento para inducir un ensanchamiento temporal de la señal de audio separando los bloques consecutivos de una señal de entrada en el dominio del tiempo en mayor medida que los bloques en solapamiento consecutivos originales de la señal de audio para obtener una señal ensanchada.

En caso de que se realice el diezmado después de la suma por solapamiento, un ensanchamiento temporal por un factor de 2,0, por ejemplo, llevará a una señal ensanchada con dos veces la duración de la señal 100 original de audio. El diezmado posterior con un correspondiente factor de diezmado de 2,0, por ejemplo, llevará a una señal con extensión de ancho de banda y diezmada de nuevo con la duración original de la señal 100 de audio. Sin embargo, en caso de que el diezmador 120 se sitúe antes del sumador 124 por solapamiento tal como se muestra en la figura 2, el diezmador 120 puede configurarse para operar en un factor de extensión de ancho de banda (σ) de 2,0, de modo que, por ejemplo, cada segunda muestra se elimine de su señal de entrada en el dominio del tiempo, lo que da como resultado una señal en el dominio del tiempo diezmada con la mitad de la duración de la señal 100 original de audio. Simultáneamente, una señal filtrada paso banda en el intervalo de frecuencia de, por ejemplo, 2 a 4 kHz se extenderá en su ancho de banda por un factor 2,0, llevando a una señal 121 en el intervalo de frecuencia objetivo correspondiente, por ejemplo, de 4 a 8 kHz después del diezmado. Posteriormente, la señal con extensión de ancho de banda y diezmada puede ensancharse temporalmente a la duración original de la señal 100 de audio mediante el sumador 124 por solapamiento aguas abajo. El procesamiento anterior, esencialmente, se relaciona con el principio de un vocodificador de fase.

La señal en el intervalo de frecuencia objetivo obtenida a partir de la salida 125 del sumador 124 por solapamiento se suministra posteriormente a un ajustador 130 de envolvente. Basándose en los parámetros transmitidos recibidos en la entrada 101 del ajustador 130 de envolvente derivados de la señal 100 de audio, se implementa el ajustador 130 de envolvente para ajustar la envolvente de la señal en la salida 125 del sumador 124 por solapamiento de una manera determinada, de modo que se obtiene una señal corregida en la salida 129 del ajustador 130 de envolvente, que comprende una envolvente ajustada y/o una tonalidad corregida.

La figura 3 muestra un diagrama de bloques de una realización de la presente invención, en la que el aparato está configurado para ejecutar un algoritmo de extensión de ancho de banda usando diferentes factores de BWE (σ) como, por ejemplo, σ=2, 3, 4,.... Inicialmente, los parámetros de algoritmo de extensión de ancho de banda se reenvían a través de la entrada 128 a todos los dispositivos que operan conjuntamente en los factores de BWE (σ). Éstos son, en particular, el primer convertidor 104, el modificador 106 de fase, el segundo convertidor 108, el diezmador 120 y el sumador 124 por solapamiento, tal como se muestra en la figura 3. Como se describió anteriormente, los dispositivos de procesamiento consecutivo para ejecutar el algoritmo de extensión de ancho de banda se implementan para operar de tal manera, que para diferentes factores de BWE (σ) en la entrada 128 correspondiente se obtienen señales de audio en el dominio del tiempo modificadas en las salidas 121-1, 121-2, 121-3,..., del diezmador 120, caracterizadas por diferentes bandas o intervalos de frecuencia objetivo, respectivamente. Entonces, las diferentes señales de audio en el dominio del tiempo modificadas se procesan por el sumador 124 por solapamiento basándose en los diferentes factores de BWE (σ), llevando a diferentes resultados de suma por solapamiento en las salidas 125-1, 125-2, 125-3,..., del sumador 124 por solapamiento. Estos resultados de suma por solapamiento se combinan finalmente mediante un combinador 126 en su salida 127 para obtener una señal combinada que comprende las diferentes bandas de frecuencia objetivo.

Para una vista ilustrativa, el principio básico del algoritmo de extensión de ancho de banda se representa en la figura 10. En particular, la figura 10 muestra esquemáticamente cómo el factor de BWE (σ) controla, por ejemplo, el desplazamiento de frecuencia entre una parte 113-1, 113-2, 113-3 de la banda de la señal 100 de audio y una banda 125-1, 125-2 ó 125-3 de frecuencia objetivo, respectivamente.

En primer lugar, en caso de que σ=2, se extrae una señal 113-1 filtrada paso banda con un intervalo de frecuencia de, por ejemplo, 2 a 4 kHz de la banda inicial de la señal 100 de audio. La banda de la señal 113-1 filtrada paso banda se transforma entonces para dar la primera salida 125-1 del sumador 124 por solapamiento. La primera salida 125-1 tiene un intervalo de frecuencia de 4 a 8 kHz correspondiente a una extensión de ancho de banda de la banda inicial de la señal 100 de audio por un factor 2,0 (σ=2). Esta banda superior para σ=2 también puede denominarse “primera banda parcheada”. A continuación, en caso de que σ=3, se extrae una señal 113-2 filtrada paso banda con el intervalo de frecuencia de 8/3 a 4 kHz, que se transforma entonces para dar la segunda salida 125-2 después del sumador 124 por solapamiento que se caracteriza por un intervalo de frecuencia de 8 a 12 kHz. La banda superior de la salida 125-2 correspondiente a una extensión de ancho de banda por un factor 3,0 (σ=3) también puede denominarse “segunda banda parcheada”. A continuación, en caso de que σ=4, se extrae la señal 113-3 filtrada paso banda con un intervalo de frecuencia de 3 a 4 kHz, que se transforma entonces para dar la tercera salida 125-3 con un intervalo de frecuencia de 12 a 16 kHz después del sumador 124 por solapamiento. La banda superior de la salida 125-3 correspondiente a una extensión de ancho de banda por un factor 4,0 (σ=4) también puede denominarse “tercera banda parcheada”. Así se obtienen las bandas parcheadas primera, segunda y tercera cubriendo bandas de frecuencia consecutivas hasta una

frecuencia máxima de 16 kHz, que se requiere preferiblemente para manipular la señal 100 de audio en el contexto de un algoritmo de extensión de ancho de banda de alta calidad. En principio, el algoritmo de extensión de ancho de banda también puede ejecutarse para valores superiores del factor de BWE $\sigma>4$, produciendo incluso más bandas de alta frecuencia. Sin embargo, teniendo en cuenta tales bandas de alta frecuencia no darán como resultado generalmente una mejora adicional de la calidad de percepción de la señal manipulada de audio.

Tal como se muestra en la figura 3, los resultados 125-1, 125-2, 125-3, ... de suma por solapamiento, basados en los diferentes factores de BWE (σ), se combinan adicionalmente mediante un combinador 126, de modo que se obtiene una señal combinada en la salida 127 que comprende las diferentes bandas de frecuencia (véase la figura 10). En este caso, la señal combinada en la salida 127 consiste en la banda parcheada de alta frecuencia transformada, que oscila entre la frecuencia máxima (f_{max}) de la señal 100 de audio y σ veces la frecuencia máxima ($o x f_{max}$), como, por ejemplo, desde 4 hasta 16 kHz (figura 10).

El ajustador 130 de envolvente aguas abajo está configurado como anteriormente para modificar la envolvente de la señal combinada basándose en parámetros transmitidos desde la señal de audio presente en la entrada 101, llevando a una señal corregida en la salida 129 del ajustador 130 de envolvente. La señal corregida suministrada por el ajustador 130 de envolvente en la salida 129 se combina adicionalmente con la señal 100 original de audio mediante un combinador 132 adicional con el fin de obtener finalmente una señal manipulada extendida en su ancho de banda en la salida 131 del combinador 132 adicional. Tal como se muestra en la figura 10, el intervalo de frecuencia de la señal con extensión de ancho de banda en la salida 131 comprende la banda de la señal 100 de audio y las diferentes bandas de frecuencia obtenidas a partir de la transformación según el algoritmo de extensión de ancho de banda que oscila, en total, por ejemplo, entre 0 y 16 kHz (figura 10).

En una realización de la presente invención según la figura 2, el divisor 102 en ventanas está configurado para insertar valores rellenados en posiciones temporales específicas antes de una primera muestra de un bloque consecutivo de muestras de audio o después de una última muestra del bloque consecutivo de muestras de audio, donde una suma de varios valores rellenados y varios valores en el bloque consecutivo es al menos 1,4 veces el número de valores en el bloque consecutivo de muestras de audio.

En particular, con respecto a la figura 7, una primera parte del bloque rellenado que tiene la longitud 712 de muestra se inserta antes de la primera muestra 708 del bloque 704 consecutivo centrado que tiene la longitud 706 de muestra, mientras que una segunda parte del bloque rellenado que tiene la longitud 714 de muestra se inserta después del bloque 704 consecutivo centrado. Obsérvese que en la figura 7 el bloque 704 consecutivo o la ventana de análisis, respectivamente, se indica mediante “región de interés” (ROI), en la que las líneas continuas, verticales que cruzan las muestras 0 y 1000 indican los bordes de la ventana 704 de análisis, donde se cumple la condición de periodicidad circular.

Preferiblemente, la primera parte del bloque rellenado a la izquierda del bloque 704 consecutivo tiene el mismo tamaño que la segunda parte del bloque rellenado a la derecha del bloque 704 consecutivo, donde el tamaño total del bloque rellenado tiene una longitud 716 de muestra (por ejemplo, desde la muestra -500 hasta la muestra 1500), que es el doble de grande que la longitud 706 de muestra del bloque 704 consecutivo centrado. Se muestra en la figura 7b, por ejemplo, que un transitorio 702 ubicado originalmente cerca del borde izquierdo de la ventana 704 de análisis se desplazará en el tiempo debido a una modificación de fase aplicada por el modificador 106 de fase, de modo que se obtendrá un transitorio 707 desplazado centrado alrededor de la primera muestra 708 del bloque 704 consecutivo centrado. En este caso, el transitorio 707 desplazado se ubicará en su totalidad dentro del bloque rellenado que tiene la longitud 716 de muestra, evitando así la convolución circular o envoltura circular provocada por la modificación de fase aplicada.

Si, por ejemplo, la primera parte del bloque rellenado a la izquierda de la primera muestra 708 del bloque 704 consecutivo centrado no es lo suficientemente grande como para albergar completamente un posible desplazamiento en el tiempo del transitorio, este último convolucionará cíclicamente, lo que significa que al menos parte del transitorio reaparecerá en la segunda parte del bloque rellenado a la derecha de la última muestra 710 del bloque 704 consecutivo. Esta parte del transitorio, sin embargo, puede eliminarse preferiblemente por el eliminador 118 de relleno tras aplicar el modificador 106 de fase en las etapas posteriores del procesamiento. Sin embargo, la longitud 716 de muestra del bloque rellenado debe ser al menos 1,4 veces más grande que la longitud 706 de muestra del bloque 704 consecutivo. Se considera que la modificación de fase aplicada por el modificador 106 de fase como, por ejemplo, la realizada por un vocodificador de fase, siempre lleva a un desplazamiento en el tiempo hacia tiempos negativos, es decir, a un desplazamiento hacia la izquierda en el eje de tiempo/muestra.

En realizaciones de la presente invención, los convertidores 104, 108 primero y segundo se implementan para operar en una longitud de conversión, que corresponde a la longitud de muestra del bloque rellenado. Por ejemplo, si el bloque consecutivo tiene una longitud de muestra N, mientras que el bloque rellenado tiene una longitud de muestra de al menos 1,4xN, tal como, por ejemplo, 2N, la longitud de conversión aplicada por el primer y el segundo convertidor 104, 108 también será 1,4xN, por ejemplo, 2N.

En principio, sin embargo, la longitud de conversión del primer convertidor y el segundo convertidor 104, 108 debe elegirse dependiendo del factor de BWE (σ) en la medida en que cuanto mayor sea el factor de BWE (σ) mayor deberá

ser la longitud de conversión. Sin embargo, es preferiblemente suficiente usar una longitud de conversión tan grande como la longitud de muestra del bloque rellenado, incluso si la longitud de conversión no es lo suficientemente grande para evitar cualquier clase de efectos de convolución cíclica para valores mayores del factor de BWE tal como, por ejemplo, para $\sigma>4$. Esto es porque en caso de que ($\sigma>4$), el solapamiento temporal de eventos transitorios por convolución cíclica, por ejemplo, es insignificante en las bandas parcheadas de alta frecuencia transformadas y no influirá significativamente en la calidad de percepción.

En la figura 4, se muestra una realización que comprende un detector 134 de transitorios, que se implementa para detectar un evento transitorio en un bloque de la señal 100 de audio, tal como, por ejemplo, en el bloque 704 consecutivo de muestras de audio que tiene la longitud 706 de muestra, tal como se muestra en la figura 7.

Específicamente, el detector 134 de transitorios está configurado para determinar si un bloque consecutivo de un bloque de audio contiene un evento transitorio, que se caracteriza por un cambio repentino de la energía de la señal 100 de audio en el tiempo, tal como, por ejemplo, un aumento o una disminución de energía en más de por ejemplo el 50% de una parte temporal a la siguiente parte temporal.

La detección de transitorios, por ejemplo, puede basarse en un procesamiento selectivo de frecuencia tal como una operación cuadrada de partes de alta frecuencia de una representación espectral que representa una medida de la potencia contenida en la banda de alta frecuencia de la señal 100 de audio y una comparación posterior del cambio temporal en la potencia con un umbral predeterminado.

Además, por un lado, el primer convertidor 104 está configurado para convertir el bloque rellenado en la salida 103 del rellenador 112, cuando el evento transitorio, tal como, por ejemplo, el evento 702 transitorio de la figura 7b se detecta por el detector 134 de transitorios en un determinado bloque 133-1 de la señal 100 de audio, que corresponde al bloque rellenado. Por otro lado, el primer convertidor 104 está configurado para convertir un bloque no rellenado que tiene sólo valores de señal de audio en la salida 133-2 del detector 134 de transitorios, donde el bloque no rellenado corresponde al bloque de la señal 100 de audio, cuando el evento transitorio no se detecta en el bloque.

En este caso, el bloque rellenado comprende valores rellenados, tales como, por ejemplo, valores cero insertados a la izquierda y derecha del bloque 704 consecutivo centrado de la figura 7b, y valores de señal de audio que residen dentro del bloque 704 consecutivo centrado de la figura 7b. El bloque no rellenado, sin embargo, comprende sólo valores de señal de audio, tal como, por ejemplo, aquellos valores de muestras de audio que residen dentro del bloque 704 consecutivo de la figura 7b.

En la realización anterior, en la que la conversión por el primer convertidor 104 y por tanto, también etapas de procesamiento posteriores basadas en la salida 105 del primer convertidor 104 son dependientes de la detección del evento transitorio, el bloque rellenado en la salida 103 del rellenador 112 se genera sólo para determinados bloques de tiempo seleccionados de la señal 100 de audio (es decir, bloques de tiempo que contienen un evento transitorio), para los que se anticipa el rellenado antes de la manipulación adicional de la señal 100 de audio de manera ventajosa en términos de calidad de percepción.

En realizaciones adicionales de la presente invención, la elección de la trayectoria de señal apropiada para el procesamiento posterior tal como se indica mediante “evento no transitorio” o “evento transitorio,” respectivamente, en la figura 4 se realiza con el uso del conmutador 136 tal como se muestra en la figura 5, que se controla por la salida 135 del detector 134 de transitorios que contiene información acerca de la detección del evento transitorio, incluyendo la información de si se detecta o no el evento transitorio en el bloque de la señal 100 de audio. Esta información del detector 134 de transitorios se reenvía mediante el conmutador 136 o bien a la salida 135-1 del conmutador 136 indicada por “evento transitorio” o bien a la salida 135-2 del conmutador 136 indicada por “evento no transitorio”. En este caso, las salidas 135-1, 135-2 del conmutador 136 en la figura 5 corresponden de manera idéntica a las salidas 133-1, 133-2 del detector 134 de transitorios en la figura 4. Como anteriormente, el bloque rellenado en la salida 103 del rellenador 112 se genera a partir del bloque 135-1 de la señal 100 de audio en el que se detecta el evento transitorio por el detector 134 de transitorios. Además, el conmutador 136 está configurado para alimentar el bloque rellenado generado por el rellenador 112 en la salida 103 al primer subconvertidor 138-1 cuando se detecta el evento transitorio por el detector 134 de transitorios y para alimentar el bloque no rellenado en la salida 135-2 a un segundo subconvertidor 138-2 cuando el evento transitorio no se detecta por el detector 134 de transitorios. En este caso, el primer subconvertidor 138-1 está adaptado para realizar una conversión del bloque rellenado usando una primera longitud de conversión, tal como, por ejemplo, 2N, mientras que el segundo subconvertidor 138-2 está adaptado para realizar una conversión del bloque no rellenado usando una segunda longitud de conversión, tal como, por ejemplo, N. Como el bloque rellenado tiene una mayor longitud de muestra que el bloque no rellenado, la segunda longitud de conversión es más corta que la primera longitud de conversión. Finalmente, se obtienen una primera representación espectral en la salida 137-1 del primer subconvertidor 138-1 o una segunda representación espectral en la salida 137-2 del segundo subconvertidor 138-2, respectivamente, que puede procesarse adicionalmente en el contexto del algoritmo de extensión de ancho de banda, tal como se ilustró anteriormente.

En una realización alternativa de la presente invención, el divisor 102 en ventanas comprende un procesador 140 de ventana de análisis, que está configurado para aplicar una función de ventana de análisis a un bloque consecutivo de muestras de audio, tal como, por ejemplo, el bloque 704 consecutivo de la figura 7. La función de ventana de análisis

aplicada por el procesador 140 de ventana de análisis, en particular, comprende al menos una zona de guarda en una posición inicial de la función de ventana, tal como, por ejemplo, la parte de tiempo que comienza en la primera muestra 718 (es decir, la muestra -500) de la función 709 de ventana a la izquierda del bloque 704 consecutivo de la figura 7b, o en una posición final de la función de ventana, tal como, por ejemplo, la parte de tiempo que termina en la última muestra 720 (es decir, la muestra 1500) de la función 709 de ventana en el lado derecho del bloque 704 consecutivo de la figura 7b.

La figura 6 muestra una realización alternativa de la presente invención que comprende además un conmutador 142 de ventana de guarda, que está configurado para controlar el procesador 140 de ventana de análisis dependiendo de la información acerca de la detección de transitorios tal como se proporciona mediante la salida 135 del detector 134 de transitorios. El procesador 140 de ventana de análisis se controla en la medida en que se genera un primer bloque consecutivo en la salida 139-1 del conmutador 142 de ventana de guarda que tiene un primer tamaño de ventana cuando se detecta el evento transitorio por el detector 134 de transitorios y se genera un bloque consecutivo adicional en la salida 139-2 del conmutador 142 de ventana de guarda que tiene un segundo tamaño de ventana cuando no se detecta el evento transitorio por el detector 134 de transitorios. En este caso, el procesador 140 de ventana de análisis está configurado para aplicar la función de ventana de análisis, tal como, por ejemplo, una ventana de Hann con una zona de guarda tal como se representa mediante la figura 9a, al bloque consecutivo en la salida 139-1 o el bloque consecutivo adicional en la salida 139-2, de modo que se obtiene un bloque rellenado en la salida 141-1 o un bloque no rellenado en la salida 141-2, respectivamente.

En la figura 9a, el bloque rellenado en la salida 141-1, por ejemplo, comprende una primera zona 910 de guarda y una segunda zona 920 de guarda, donde los valores de las muestras de audio de las zonas 910, 920 de guarda se establecen a cero. En este caso, las zonas 910, 920 de guarda rodean una zona 930 correspondiente a las características de la función de ventana, en este caso, por ejemplo, dadas por la forma característica de la ventana de Hann. Alternativamente, con respecto a la figura 9b, los valores de las muestras de audio de las zonas 940, 950 de guarda también pueden oscilar alrededor de cero. Las líneas verticales en la figura 9 indican una primera muestra 905 y una última muestra 915 de la zona 930. Además, las zonas 910, 940 de guarda comienzan con la primera muestra 901 de la función de ventana, mientras que las zonas 920, 950 de guarda terminan con la última muestra 903 de la función de ventana. La longitud de muestra 900 de la ventana completa que tiene una parte de ventana de Hann centrada, que incluyen las zonas 910, 920 de guarda, de la figura 9a, por ejemplo, es dos veces más grande que la longitud de muestra de la zona 930.

En el caso de que se detecte el evento transitorio por el detector 134 de transitorios, se procesa el bloque consecutivo en la salida 139-1 en la medida en que se pondera por la forma característica de la función de ventana de análisis tal como, por ejemplo, la ventana 901 de Hann normalizada con las zonas 910, 920 de guarda tal como se muestra en la figura 9a, mientras que en el caso de que no se detecte el evento transitorio por el detector 134 de transitorios, se procesa el bloque consecutivo en la salida 139-2 en la medida en que se pondera por la forma característica de la zona 930 de sólo la función de ventana de análisis tal como, por ejemplo, la zona 930 de la ventana 901 de Hann normalizada de la figura 9a.

En caso de que el bloque rellenado o bloque no rellenado en las salidas 141-1, 141-2 se genere mediante el uso de la función de ventana de análisis que comprende la zona de guarda que acaba de mencionarse, los valores rellenados o valores de señal de audio se originan de la ponderación de las muestras de audio por la zona de guarda o la zona no de guarda (característica) de la función de ventana, respectivamente. En este caso, tanto los valores rellenados como los valores de señal de audio representan valores ponderados, donde específicamente los valores rellenados son aproximadamente cero. Específicamente, el bloque rellenado o el bloque no rellenado en las salidas 141-1, 141-2 pueden corresponder a los de las salidas 103, 135-2 en la realización mostrada en la figura 5.

Por la ponderación debida a la aplicación de la función de ventana de análisis, el detector 134 de transitorios y el procesador 140 de ventana de análisis deben disponerse preferiblemente de tal manera que la detección del evento transitorio por el detector 134 de transitorios tenga lugar antes de que se aplique la función de ventana de análisis por el procesador 140 de ventana de análisis. Por el contrario, la detección del evento transitorio se verá influida significativamente debido al proceso de ponderación, que es especialmente el caso para un evento transitorio ubicado dentro de las zonas de guarda o cerca de los bordes de la zona no de guarda (característica), ya que en esta región, los factores de ponderación correspondientes a los valores de la función de ventana de análisis siempre se acercan a cero.

El bloque rellenado en la salida 141-1 y el bloque no rellenado en la salida 141-2 se convierten posteriormente en sus representaciones espectrales en las salidas 143-1, 143-2, usando el primer subconvertidor 138-1 con la primera longitud de conversión y el segundo subconvertidor 138-2 con la segunda longitud de conversión, donde la primera y la segunda longitud de conversión corresponden a las longitudes de muestra de los bloques convertidos, respectivamente. Las representaciones espectrales en las salidas 143-1, 143-2 pueden procesarse adicionalmente como en las realizaciones tratadas anteriormente.

La figura 8 muestra una visión general de una realización de la implementación de extensión de ancho de banda. En particular, la figura 8 incluye el bloque 800 indicado por “señal de audio/parámetros adicionales” que proporciona la señal 100 de audio indicada por el bloque de salida “datos de audio de baja frecuencia (LF)”. Además, el bloque 800 proporciona parámetros decodificados que pueden corresponder a la entrada 101 del ajustador 130 de envolvente en

las figuras 2 y 3. Los parámetros en la salida 101 del bloque 800 pueden usarse posteriormente para el ajustador 130 de envolvente y/o un corrector 150 de tonalidad. El ajustador 130 de envolvente y el corrector 150 de tonalidad están configurados para aplicar, por ejemplo, una distorsión predeterminada a la señal 127 combinada para obtener la señal 151 distorsionada, que puede corresponder a la señal 129 corregida de las figuras 2 y 3.

El bloque 800 puede comprender información secundaria acerca de la detección de transitorios provista en el lado de codificador de la implementación de extensión de ancho de banda. En este caso, esta información secundaria se transmite adicionalmente mediante un flujo 810 de bits que se indica mediante la línea discontinua al detector 134 de transitorios en el lado de decodificador.

Preferiblemente, sin embargo, la detección de transitorios se realiza en la pluralidad de bloques consecutivos de muestras de audio en la salida 111 del procesador 110 de ventana de análisis denominado en este caso dispositivo 102-1 de “formación de tramas”. En otras palabras, la información secundaria transitoria o bien se detecta en el detector 134 de transitorios que representa el decodificador o bien se transfiere en el flujo 810 de bits a partir del codificador (línea discontinua). La primera solución no aumenta la tasa de bits que va a transmitirse, mientras que la última facilita la detección, ya que la señal original está todavía disponible.

Específicamente, la figura 8 muestra un diagrama de bloques de un aparato que se configura para realizar una implementación de extensión de ancho de banda armónica (HBE), tal como se muestra en la figura 13, que se combina con el conmutador 136, controlado por el detector 134 de transitorios, para ejecutar un procesamiento adaptativo de señal, dependiendo de la información acerca de la aparición de un evento transitorio en la salida 135.

En la figura 8, la pluralidad de bloques consecutivos en la salida 111 del dispositivo 102-1 de formación de tramas se suministra a un dispositivo 102-2 divisor en ventanas de análisis, que está configurado para aplicar una función de ventana de análisis que tiene una forma de ventana predeterminada, tal como, por ejemplo, una ventana de coseno elevado, que se caracteriza por flancos menos profundos en comparación con una forma de ventana rectangular aplicada normalmente en una operación de formación de tramas. Dependiendo de la decisión de conmutación indicada por “transitorio” o “no transitorio” obtenida con el conmutador 136, el bloque 135-1 que incluye el evento transitorio o el bloque 135-2 que no incluye el evento transitorio, respectivamente, de la pluralidad de bloques divididos en ventanas consecutivos (es decir, en tramas y ponderados) en la salida 811 del dispositivo 102-2 divisor en ventanas de análisis, tal como se detecta por el detector 134 de transitorios, se procesan adicionalmente tal como se trató en detalle anteriormente. Especialmente, se usa preferiblemente un dispositivo 102-3 de rellenado con ceros, que puede corresponder al rellenador 112 de la ventana 102 en las figuras 2, 4 y 5 para insertar valores cero fuera del bloque 135-1 de tiempo, de modo que se obtiene un bloque 803 rellenado con ceros, que puede corresponder al bloque 103 rellenado, con la longitud de muestra 2N dos veces más grande que la longitud de muestra N del bloque 135-2 de tiempo. En este caso, el detector 134 de transitorios se indica mediante “detector de posición transitoria”, ya que puede usarse para determinar la “posición” (es decir, ubicación de tiempo) del bloque 135-1 consecutivo con respecto a la pluralidad de bloques consecutivos en la salida 811, es decir, el bloque de tiempo respectivo que contiene el evento transitorio puede identificarse a partir de la secuencia de bloques consecutivos en la salida 811.

En una realización, el bloque rellenado siempre se genera a partir de un bloque consecutivo específico para el que se detecta el evento transitorio, independientemente de su ubicación dentro del bloque. En este caso, el detector 134 de transitorios está configurado simplemente para determinar (identificar) el bloque que contiene el evento transitorio. En una realización alternativa, el detector 134 de transitorios puede configurarse además para determinar la ubicación particular del evento transitorio con respecto al bloque. En la primera realización, puede usarse una implementación más simple del detector 134 de transitorios, mientras que en la última realización, puede reducirse la complejidad computacional del procesamiento, ya que el bloque rellenado se generará y procesará adicionalmente sólo si un evento transitorio se ubica en una ubicación particular, preferiblemente cerca de un borde de bloque. En otras palabras, en la última realización, sólo se necesitarán zonas de guarda o rellenado con ceros si un evento transitorio se ubica cerca de los bordes de bloque (es decir, si se producen transitorios fuera del centro).

El aparato de la figura 8, esencialmente, proporciona un método para contrarrestar el efecto de convolución cíclica introduciendo denominados “intervalos de guarda” rellenando con ceros ambos extremos de cada bloque de tiempo antes de entrar en el procesamiento de vocodificador de fase. En este caso, el procesamiento de vocodificador de fase comienza con la operación del primer o el segundo subconvertidor 138-1, 138-2, que comprende, por ejemplo, un procesador de FFT que tiene una longitud de conversión de 2N o N, respectivamente.

Específicamente, el primer convertidor 104 puede implementarse para realizar una transformada de Fourier a corto plazo (STFT) del bloque 103 rellenado, mientras que el segundo convertidor 108 puede implementarse para realizar una STFT inversa basándose en la magnitud y fase de la representación espectral modificada en la salida 105.

Con respecto a la figura 8, después de que se hayan calculado las fases nuevas y, por ejemplo, se realice la síntesis de la transformada discreta de Fourier inversa (IDFT) o STFT inversa, los intervalos de guarda se desprenden simplemente de la parte central del bloque de tiempo, que se procesa adicionalmente en la etapa de suma por solapamiento (OLA) del vocodificador. Alternativamente, los intervalos de guarda no deben eliminarse, sino procesarse adicionalmente en la etapa de OLA. Esta operación también puede considerarse de manera eficaz como un sobremuestreo de la señal.

Como resultado de la implementación según la figura 8, se obtiene una señal manipulada extendida en ancho de banda en la salida 131 del combinador 132 adicional. Posteriormente, puede usarse un dispositivo 160 de formación de tramas adicional para modificar la formación de tramas (es decir, el tamaño de ventana de la pluralidad de bloques de tiempo consecutivos) del audio manipulado en la señal de salida 131 indicada por "señal de audio con alta frecuencia (HF)" de una manera predeterminada, por ejemplo, de manera que el bloque consecutivo de muestras de audio en la salida 161 del dispositivo 160 de formación de tramas adicional tendrá el mismo tamaño de ventana que la señal 800 de audio inicial.

La posible ventaja de usar intervalos de guarda en este contexto mientras se procesan transitorios por un vocodificador de fase tal como se expone, por ejemplo, en la realización de la figura 8, se visualiza a modo de ejemplo en la figura 7. La lámina a) muestra el transitorio centrado en la ventana de análisis ("con líneas discontinuas delgadas" se indica la señal original). En este caso, el intervalo de guarda no tiene un efecto significativo en el procesamiento puesto que la ventana también puede albergar el transitorio modificado ("continua delgada" usando intervalos de guarda, "continua gruesa" sin intervalos de guarda). Sin embargo, tal como se muestra en la lámina b), si el transitorio está fuera del centro ("con líneas discontinuas delgadas" se indica la señal original), se desplazará en el tiempo mediante la manipulación de fase durante el procesamiento de vocodificador. Si este desplazamiento no puede albergarse directamente por el intervalo de tiempo cubierto por la ventana, se produce una envoltura circular ("continua gruesa" sin intervalos de guarda) que eventualmente lleva a una colocación errónea de (partes de) el transitorio, degradando así la calidad de audio de percepción. Sin embargo, el uso de intervalos de guarda evita efectos de convolución circular albergando las partes desplazadas en la zona de guarda ("continua delgada" usando intervalos de guarda).

Como una alternativa a la implementación de rellenado con ceros anterior, pueden usarse las ventanas con zonas de guarda (véase la figura 9) tal como se mencionó anteriormente. En el caso de las ventanas con zonas de guarda, en uno o ambos lados de las ventanas los valores son aproximadamente cero. Pueden ser exactamente cero u oscilar alrededor de cero con la posible ventaja de no desplazar ceros desde la zona de guarda a la ventana a través de la adaptación de fase sino valores pequeños. La figura 9 muestra ambos tipos de ventanas. Particularmente, en la figura 9, la diferencia entre las funciones 901, 902 de ventana es que en la figura 9a la función 901 de ventana comprende las zonas 910, 920 de guarda cuyos valores de muestra son exactamente cero, mientras que en la figura 9b la función 902 de ventana comprende las zonas 940, 950 de guarda cuyos valores de muestra oscilan alrededor de cero. Por tanto, en este último caso, se desplazarán valores pequeños en lugar de valores cero a través de la adaptación de fase desde la zona 940 ó 950 de guarda a la zona 930 de la ventana.

Tal como se mencionó anteriormente, la aplicación de intervalos de guarda puede aumentar la complejidad computacional debido a sus equivalentes a un sobremuestreo puesto que las transformaciones de análisis y síntesis deben calcularse en bloques de señal de longitud sustancialmente extendida (habitualmente un factor de 2). Por un lado, esto garantiza una calidad de percepción mejorada al menos para bloques de señal transitorios, pero esto sólo se produce en bloques seleccionados de una señal de audio de música promedio. Por otro lado, se aumenta regularmente la capacidad de procesamiento a lo largo de todo el procesamiento de la totalidad de la señal.

Las realizaciones de la invención se basan en el hecho de que el sobremuestreo es sólo ventajoso para determinados bloques de señal seleccionados. Específicamente, las realizaciones proporcionan un método de procesamiento adaptativo de señal novedoso que comprende un mecanismo de detección y aplica sobremuestreo sólo a aquellos bloques de señal en los que en efecto mejora la calidad de percepción. Además, como el procesamiento de señal conmuta de manera adaptativa entre procesamiento convencional y procesamiento avanzado, la eficacia del procesamiento de señal en el contexto de la presente invención puede aumentarse significativamente, reduciendo así el esfuerzo computacional.

Para ilustrar la diferencia entre el procesamiento convencional y el procesamiento avanzado, a continuación se realizará la comparación de una implementación de extensión de ancho de banda armónica típica (HBE) (figura 13) con la implementación de la figura 8.

La figura 13 representa una visión general de HBE. En este caso, las etapas de vocodificador de múltiples fases operan en la misma frecuencia de muestreo que la totalidad del sistema. La figura 8, sin embargo, muestra la manera de procesar aplicando rellenado con ceros/sobremuestreo sólo a aquellas partes de la señal, en las que es verdaderamente beneficioso y da como resultado una calidad de percepción mejorada. Esto se logra mediante una decisión de conmutación, que es preferiblemente dependiente de una detección de ubicación de transitorio que elige la trayectoria de señal apropiada para el procesamiento posterior. Comparado con la HBE mostrada en la figura 13, la detección 134 de ubicación de transitorio (desde la señal o flujo de bits), el conmutador 136 y la trayectoria de señal en el lado derecho, comenzando con la operación de rellenado con ceros aplicada por el rellenador 102-3 con ceros y terminando con la eliminación de rellenado (opcional) realizada por el eliminador 118 de relleno, se ha añadido en las realizaciones tal como se ilustra en la figura 8.

En una realización de la presente invención, el divisor 102 en ventanas está configurado para generar una pluralidad 111 de bloques consecutivos de muestras de audio que forman una secuencia temporal, que comprende al menos un primer par 145-1 de un bloque 133-2, 141-2 no rellenado y un bloque 103, 141-1 rellenado consecutivo y un segundo par 145-2 de un bloque 103, 141-1 rellenado y un bloque 133-2, 141-2 no rellenado consecutivo (véase la figura 12). El primer y el segundo par de bloques 145-1, 145-2 consecutivos se procesan adicionalmente en el contexto de la

implementación extensión de ancho de banda, hasta que se obtienen sus muestras de audio diezmadas correspondientes en las salidas 147-1, 147-2 del diezmador 120, respectivamente. Las muestras 147-1, 147-2 de audio diezmadas se alimentan posteriormente al sumador 124 por solapamiento, que está configurado para sumar bloques en solapamiento de las muestras 147-1, 147-2 de audio diezmadas del primer par 145-1 o del segundo par 145-2.

Alternativamente, el diezmador 120 también puede colocarse después del sumador 124 por solapamiento tal como se describe anteriormente de manera correspondiente.

Entonces, para el primer par 145-1, una distancia de tiempo b', que puede corresponder a la distancia de tiempo b de la figura 2, entre una primera muestra 151, 155 del bloque 133-2, 141-2 no rellenado y una primera muestra 153, 157 de los valores de señal de audio del bloque 103, 141-1 rellenado, respectivamente, se suministra por el sumador 124 por solapamiento, de modo que se obtiene una señal en el intervalo de frecuencia objetivo del algoritmo de extensión de ancho de banda en la salida 149-1 del sumador 124 por solapamiento.

Para el segundo par 145-2, la distancia de tiempo b' entre una primera muestra 153, 157 de los valores de señal de audio del bloque 103, 141-1 rellenado y una primera muestra 151, 155 del bloque 133-2, 141-2 no rellenado, respectivamente, se suministra por el sumador 124 por solapamiento, de modo que se obtiene una señal en el intervalo de frecuencia objetivo del algoritmo de extensión de ancho de banda en la salida 149-2 del sumador 124 por solapamiento.

Nuevamente, en el caso de que el diezmador 120 se sitúe antes del sumador 124 por solapamiento en la cadena de procesamiento tal como se muestra en la figura 2, debe tenerse en cuenta un posible efecto del diezmado en la correspondencia a la distancia de tiempo b'.

Debe observarse que aunque la presente invención se ha descrito en el contexto de diagramas de bloques en los que los bloques representan componentes de hardware real o lógico, la presente invención también puede implementarse mediante un método implementado por ordenador. En este último caso, los bloques representan etapas del método correspondientes donde estas etapas representan las funcionalidades realizadas por los bloques de hardware lógico o físico correspondientes.

Las realizaciones descritas son meramente ilustrativas para los principios de la presente invención. Se entiende que serán evidentes para otros expertos en la técnica modificaciones y variaciones de las disposiciones y los detalles descritos en el presente documento. Por tanto, sólo pretende quedar limitada por el alcance de las reivindicaciones de patente adjuntas y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en el presente documento.

Dependiendo de determinados requisitos de implementación de los métodos de la invención, los métodos de la invención pueden implementarse en hardware o en software. La implementación puede realizarse usando un medio de almacenamiento digital, en particular un disco, un DVD o un CD con señales de control legibles electrónicamente almacenadas en los mismos, que operan conjuntamente con sistemas informáticos programables, de manera que se realicen los métodos de la invención. Generalmente, la presente, por tanto, puede implementarse como un producto de programa informático con el código de programa almacenado en un soporte legible por máquina, operándose el código de programa para realizar los métodos de la invención cuando el producto programa informático se ejecuta en un ordenador. En otras palabras, los métodos de la invención son, por tanto, un programa informático que tiene un código de programa para realizar al menos uno de los métodos de la invención cuando el programa informático se ejecuta en un ordenador. La señal de audio procesada de la invención puede almacenarse en cualquier medio de almacenamiento legible por máquina, tal como un medio de almacenamiento digital.

Las ventajas del procesamiento novedoso son que las realizaciones mencionadas anteriormente, es decir el aparato, los métodos o programas informáticos, descritos en esta solicitud evitan un procesamiento computacional demasiado complejo y costoso cuando no es necesario. Utiliza una detección de ubicación de transitorio que identifica los bloques de tiempo que contienen, por ejemplo, eventos transitorios fuera del centro y conmuta al procesamiento avanzado, por ejemplo al procesamiento sobremuestreado usando intervalos de guarda, sin embargo, sólo en aquellos casos, en los que da como resultado una mejora en cuanto a calidad de percepción.

El procesamiento presentado es útil en cualquier aplicación de procesamiento de audio basada en bloques, por ejemplo vocodificadores de fase, o aplicaciones de sonido envolvente paramétricas (Herre, J.; Faller, C.; Ertel, C.; Hilpert, J.; Hölzer, A.; Spenger, C, "MP3 Surround: Efficient and Compatible Coding of Multi-Canal Audio", 116th Conv. Aud. Eng. Soc., mayo del 2004), donde efectos de convolución circular temporales llevan a solapamiento y, al mismo tiempo, la capacidad de procesamiento es un recurso limitado.

Las aplicaciones más destacadas son decodificadores de audio, que con frecuencia se implementan en dispositivos portátiles y de ese modo operan con un suministro de energía por batería.

REIVINDICACIONES

1. Aparato para manipular una señal (100) de audio, que comprende:

un divisor (102) en ventanas para generar una pluralidad (111; 811) de bloques consecutivos de muestras de audio, comprendiendo la pluralidad (111; 811) de bloques consecutivos al menos un bloque (103; 803; 141-1; 902) rellenado de muestras de audio, teniendo el bloque (103; 803; 141-1; 902) rellenado valores rellenados y valores de señal de audio;

un primer convertidor (104) para convertir el bloque (103; 803; 141-1; 902) rellenado en una representación (105) espectral que tiene valores espectrales;

un modificador (106) de fase para modificar fases de los valores espectrales para obtener una representación (107) espectral modificada; y

un segundo convertidor (108) para convertir la representación (107) espectral modificada en una señal (109) de audio en el dominio del tiempo modificada,

estando configurado el aparato para ejecutar un algoritmo de extensión de ancho de banda, comprendiendo el algoritmo de extensión de ancho de banda un factor de extensión de ancho de banda (σ), controlando el factor de extensión de ancho de banda (σ) un desplazamiento de frecuencia entre una banda (113-1, 113-2, 113-3, ...) de la señal (100) de audio y una banda (125-1, 125-2, 125-3, ...) de frecuencia objetivo, estando configurado el modificador (106) de fase para ajustar a escala las fases de valores espectrales de la banda (113-1, 113-2, 113-3, ...) de la señal (100) de audio por el factor de extensión de ancho de banda (σ), de modo que al menos una muestra de un bloque consecutivo de muestras de audio se convoluciona cíclicamente en el bloque, y

comprendiendo el aparato adicionalmente un detector (134) de transitorios para determinar un evento (700, 701, 702, 703, 705, 707) transitorio en la señal (100) de audio,

estando configurado el primer convertidor (104) para convertir el bloque (103; 803; 141-1; 902) rellenado, cuando el transitorio (134) detecta el evento (700, 701, 702, 703, 705, 707) transitorio en un bloque (133-1; 135-1) de la señal (100) de audio correspondiente al bloque (103; 803; 141-1; 902) rellenado, y

estando configurado el primer convertidor (104) para convertir un bloque (133-2; 135-2; 141-2; 930) no rellenado que tiene sólo valores de señal de audio, correspondiendo el bloque (133-2; 135-2; 141-2; 930) no rellenado al bloque de la señal (100) de audio, cuando el transitorio (700, 701, 702, 703, 705, 707) no se detecta en el bloque.

2. Aparato según la reivindicación 1, que comprende además:

un diezmador (120) para diezmar la señal (109) de audio en el dominio del tiempo modificada o bloques sumados por solapamiento de muestras de audio en el dominio del tiempo modificadas para obtener una señal (121) en el dominio del tiempo diezmada, en el que una característica de diezmado depende de una característica de modificación de fase aplicada por el modificador (106) de fase.

3. Aparato según la reivindicación 2, que está adaptado para realizar una extensión de ancho de banda usando la señal (100) de audio, que comprende además:

un filtro (114) paso banda para extraer una señal (113) paso banda de la representación (105) espectral o de la señal (100) de audio, en el que se selecciona una característica paso banda del filtro (114) paso banda dependiendo de una característica de modificación de fase aplicada por el modificador (106) de fase, de modo que la señal (113) paso banda se transforma mediante un procesamiento posterior a un intervalo (125-1, 125-2, 125-3) de frecuencia objetivo no incluido en la señal (100) de audio.

4. Aparato según la reivindicación 2, que comprende además:

un sumador (124) por solapamiento para sumar bloques (121-1, 121-2, 121-3) en solapamiento de muestras de audio diezmadas o muestras de audio en el dominio del tiempo modificadas para obtener una señal (125) en un intervalo (125-1, 125-2, 125-3) de frecuencia objetivo de un algoritmo de extensión de ancho de banda.

5. Aparato según la reivindicación 4, que comprende además:

un ajustador (116) a escala para ajustar a escala los valores espectrales por un factor, dependiendo el factor de una característica de suma por solapamiento en la medida en que se tiene en cuenta una relación entre la primera distancia de tiempo (a) para una suma por solapamiento aplicada por el divisor (102) en ventanas y una distancia de tiempo (b) diferente aplicada por el sumador (124) por solapamiento y las características de ventana.

6. Aparato según la reivindicación 1, en el que el divisor (102) en ventanas comprende:

un procesador (110; 102-1, 102-2; 140) de ventana de análisis para generar una pluralidad (111; 811) de bloques consecutivos que tienen el mismo tamaño; y

un rellenador (112; 102-3) para rellenar un bloque (133-1; 135-1) de la pluralidad (111; 811) de bloques consecutivos de muestras de audio para obtener el bloque (103; 803; 141-1; 902) rellenado insertando valores rellenados en posiciones temporales específicas antes de una primera muestra (708) de un bloque (133-1; 135-1; 704) consecutivo de muestras de audio o después de una última muestra (710) del bloque (133-1; 135-1; 704) consecutivo de muestras de audio.

7. Aparato según la reivindicación 1, en el que el divisor (102) en ventanas está configurado para insertar valores rellenados en posiciones temporales específicas antes de una primera muestra (708) de un bloque (133-1; 135-1; 704) consecutivo de muestras de audio o después de una última muestra (710) del bloque (133-1; 135-1; 704) consecutivo de muestras de audio, comprendiendo el aparato adicionalmente:

un eliminador (118) de relleno para eliminar muestras en posiciones temporales de la señal (109) de audio en el dominio del tiempo modificada, correspondiendo las posiciones temporales a las posiciones temporales especificadas aplicadas por el divisor (102) en ventanas.

8. Aparato según la reivindicación 1 ó 2, que comprende además:

un divisor (122) en ventanas de síntesis para dividir en ventanas la señal (121) en el dominio del tiempo diezmada o la señal (109) de audio en el dominio del tiempo modificada que tiene una función de ventana de síntesis coincidente con una función de análisis aplicada por el divisor (102) en ventanas.

9. Aparato según la reivindicación 1, en el que el divisor (102) en ventanas está configurado para insertar valores rellenados en posiciones temporales específicas antes de una primera muestra (708) de un bloque (133-1; 135-1; 704) consecutivo de muestras de audio o después de una última muestra (710) del bloque (133-1; 135-1; 704) consecutivo de muestras de audio, en el que una suma de varios valores rellenados y varios valores en el bloque (133-1; 135-1; 704) consecutivo de muestras de audio es al menos 1,4 veces el número de valores en el bloque (133-1; 135-1; 704) consecutivo de muestras de audio.

10. Aparato según la reivindicación 7, en el que el divisor (102) en ventanas está configurado para insertar simétricamente los valores rellenados antes de la primera muestra (708) del bloque (133-1; 135-1; 704) consecutivo de muestras de audio y después de la última muestra (710) del bloque (133-1; 135-1; 704) consecutivo centrado de muestras de audio, de modo que el bloque (103; 803; 141-1; 902) rellenado está adaptado para una conversión por el primer convertidor (104) y el segundo convertidor (108).

11. Aparato según la reivindicación 1, en el que el divisor (102) en ventanas está configurado para aplicar una función (709; 902) de ventana que tiene al menos una zona (712, 714; 910, 920; 940, 950) de guarda en la posición (718; 901) inicial de la función (709; 902) de ventana o en la posición (720; 903) final de la función (709; 902) de ventana.

12. Aparato según la reivindicación 2, estando configurado el aparato para ejecutar un algoritmo de extensión de ancho de banda, comprendiendo el algoritmo de extensión de ancho de banda un factor de extensión de ancho de banda (σ), controlando el factor de extensión de ancho de banda (σ) un desplazamiento de frecuencia entre una banda (113-1, 113,-2, 113-3, ...) de la señal (100) de audio y una banda (125-1, 125-2, 125-3, ...) de frecuencia objetivo,

estando configurados el primer convertidor (104), el modificador (106) de fase, el segundo convertidor (108) y el diezmador (120) para operar usando diferentes factores de extensión de ancho de banda (σ), de modo que se obtienen diferentes señales (121-1, 121-2, 121-3, ...) de audio temporales modificadas que tienen diferentes bandas (125-1, 125-2, 125-3, ...) de frecuencia objetivo,

comprendiendo además un sumador (124) por solapamiento para realizar una suma por solapamiento basándose en los diferentes factores de extensión de ancho de banda (σ), y

un combinador (126) para combinar resultados (125-1, 125-2, 125-3, ...) de suma por solapamiento para obtener una señal (127) combinada que comprende las diferentes bandas (125-1, 125-2, 125-3) de frecuencia objetivo.

13. Aparato según la reivindicación 1, en el que el divisor (102) en ventanas comprende:

un rellenador (112; 102-3) para insertar valores rellenados en posiciones temporales específicas antes de una primera muestra (708) de un bloque (133-1; 135-1; 704) consecutivo de muestras de audio o después de una última muestra (710) del bloque (133-1; 135-1; 704) consecutivo de muestras de audio, comprendiendo el aparato adicionalmente:

un conmutador (136) que está controlado por el detector (134) de transitorios, estando configurado el conmutador (136) para controlar el rellenador (112; 102-3) de modo que se genera un bloque (103; 803) rellenado cuando se detecta un evento (700, 701, 702, 703, 705, 707) transitorio por el detector (134) de transitorios, teniendo el bloque (103; 803) rellenado valores rellenados y valores de señal de audio, y para controlar el rellenador (112; 102-3), de modo que se genera un bloque (133-2; 135-2) no rellenado cuando no se detecta el evento (700, 701, 702, 703, 705, 707) transitorio por el detector (134) de transitorios, teniendo el bloque (133-2; 135-2) no rellenado sólo valores de señal de audio,

comprendiendo el primer convertidor (104) un primer subconvertidor (138-1) y un segundo subconvertidor (138-2),

estando configurado el conmutador (136) adicionalmente para alimentar el bloque (103; 803) rellenado al primer subconvertidor (138-1) para realizar una conversión que tiene una primera longitud de conversión cuando se detecta el evento (700, 701, 702, 703, 705, 707) transitorio por el detector (134) de transitorios y para alimentar el bloque (133-2; 135-2) no rellenado al segundo subconvertidor (138-2) para realizar una conversión que tiene una segunda longitud más corta que la primera longitud cuando no se detecta el evento (700, 701, 702, 703, 705, 707) transitorio por el detector (134) de transitorios.

14. Aparato según la reivindicación 1, en el que el divisor (102) en ventanas comprende un procesador (110; 102-1, 102-2; 140) de ventana de análisis para aplicar una función de ventana de análisis a un bloque (139-1, 139-2) consecutivo de muestras de audio, pudiendo controlarse el procesador de ventana de análisis de modo que la función de ventana de análisis comprende una zona (712, 714; 910, 920; 940, 950) de guarda en una posición (718; 901) inicial de la función (709; 902) de ventana o una posición (720; 903) final de la función (709; 902) de ventana, comprendiendo el aparato adicionalmente:

un conmutador (142) de ventana de guarda que está controlado por el detector (134) de transitorios, estando configurado el conmutador (142) de ventana de guarda para controlar el procesador (110; 102-1, 102-2; 140) de ventana de análisis, de modo que se genera un bloque (141-1; 902) rellenado de un bloque consecutivo de muestras de audio mediante el uso de la función de ventana de análisis que comprende la zona de guarda, teniendo el bloque (141-1; 902) rellenado valores rellenados y valores de señal de audio cuando se detecta un evento (700, 701, 702, 703, 705, 707) transitorio por el detector (134) de transitorios, y para controlar el procesador (102-1, 102-2; 140) de ventana de análisis, de modo que se genera un bloque (141-2; 930) no rellenado, teniendo el bloque (141-2; 930) no rellenado sólo valores de señal de audio, cuando no se detecta el evento (700, 701, 702, 703, 705, 707) transitorio por el detector (134) de transitorios,

comprendiendo el primer convertidor (104) un primer subconvertidor (138-1) y un segundo subconvertidor (138-2),

estando configurado el conmutador (142) de ventana de guarda además para alimentar el bloque (141-1; 902) rellenado al primer subconvertidor (138-1) para realizar una conversión que tiene una primera longitud de conversión cuando se detecta un evento (700, 701, 702, 703, 705, 707) transitorio por el detector (134) de transitorios y para alimentar el bloque (141-2; 930) no rellenado al segundo subconvertidor (138-2) para realizar una conversión que tiene una segunda longitud más corta que la primera longitud cuando no se detecta el evento (700, 701, 702, 703, 705, 707) transitorio por el detector (134) de transitorios.

15. Aparato según la reivindicación 4 ó 12, que comprende además:

un ajustador (130) de envolvente para ajustar la envolvente de la señal (125) en un intervalo (125-1, 125-2, 125-3) de frecuencia objetivo o la señal (129) combinada basándose en parámetros (101) transmitidos para obtener una señal (129) corregida; y

un combinador (132) adicional para combinar la señal (100; 102-1) de audio y la señal (129) corregida para obtener una señal (131) manipulada extendida en ancho de banda.

16. Aparato según la reivindicación 1, en el que el divisor (102) en ventanas está configurado para generar una pluralidad (111; 811) de bloques consecutivos de muestras de audio, comprendiendo la pluralidad (111; 811) de bloques consecutivos al menos un primer par (145-1) de un bloque (133-2; 135-2; 141-2; 930) no rellenado y un bloque (103; 803; 141-1; 902) rellenado consecutivo y un segundo par (145-2) de un bloque (103; 803; 141-1; 902) rellenado y un bloque (133-2; 135-2; 141-2; 930) no rellenado consecutivo, comprendiendo el aparato adicionalmente:

un diezmador (120) para diezmar las muestras de audio en el dominio del tiempo modificadas o bloques sumados por solapamiento de muestras de audio en el dominio del tiempo modificadas del primer par (145-1) para obtener las muestras (147-1) de audio diezmadas del primer par (145-1) o para diezmar las muestras de audio en el dominio del tiempo modificadas o bloques sumados por solapamiento de muestras de audio en el dominio del tiempo modificadas del segundo par (145-2) para obtener las muestras (147-2) de audio diezmadas del segundo par (145-2), y

un sumador (124) por solapamiento, estando configurado el sumador (124) por solapamiento para sumar bloques en solapamiento de las muestras (147-1, 147-2) de audio diezmadas o muestras de audio en el dominio del tiempo modificadas del primer par (145-1) o del segundo par (145-2), en el que para el primer par (145-1) la distancia de tiempo (b') entre una primera muestra (151) del bloque (133-2; 135-2; 141-2; 930) no rellenado y una primera muestra (153) de los valores de señal de audio del bloque (103; 803141-1; 902) rellenado se suministra mediante el sumador (124) por solapamiento, o en el que para el segundo par (145-2) una distancia de tiempo (b') entre una primera muestra (153) de los valores de señal de audio del bloque (103; 803; 141-1; 902) rellenado y una primera muestra (157) del bloque (133-2; 135-2; 141-2; 930) no rellenado se suministra mediante el sumador (124) por solapamiento, para obtener una señal en un intervalo de frecuencia objetivo del algoritmo de extensión de ancho de banda.

17. Método para manipular una señal de audio, que comprende:

generar (102) una pluralidad (111; 811) de bloques consecutivos de muestras de audio, comprendiendo la pluralidad (111; 811) de bloques consecutivos al menos un bloque (103; 803) rellenado de muestras de audio, teniendo el bloque (103; 803) rellenado valores rellenados y valores de señal de audio;

convertir (104) el bloque (103; 803) rellenado en una representación espectral que tiene valores espectrales;

modificar (106) fases de los valores espectrales para obtener una representación (107) espectral modificada; y

convertir (108) la representación (107) espectral modificada en una señal (109) de audio en el dominio del tiempo (105) modificada,

ejecutar un algoritmo de extensión de ancho de banda, comprendiendo el algoritmo de extensión de ancho de banda un factor de extensión de ancho de banda (σ), controlando el factor de extensión de ancho de banda (σ) un desplazamiento de frecuencia entre una banda (113-1, 113-2, 113-3, ...) de la señal (100) de audio y una banda (125-1, 125-2, 125-3, ...) de frecuencia objetivo, en el que la etapa de modificación (106) comprende ajustar a escala fases de valores espectrales de la banda (113-1, 113-2, 113-3, ...) de la señal (100) de audio por el factor de extensión de ancho de banda (σ), de modo que al menos una muestra de un bloque consecutivo de muestras de audio se convoluciona cíclicamente en el bloque, y

determinar un evento (700, 701, 702, 703, 705, 707) transitorio en la señal (100) de audio,

comprendiendo la etapa de conversión (104) convertir el bloque (103; 803; 141-1; 902) rellenado, cuando el transitorio (134) detecta el evento (700, 701, 702, 703, 705, 707) transitorio en un bloque (133-1; 135-1) de la señal (100) de audio correspondiente al bloque (103; 803; 141-1; 902) rellenado, y

comprendiendo la etapa de conversión (104) convertir un bloque (133-2; 135-2; 141-2; 930) no rellenado que tiene sólo valores de señal de audio, correspondiendo el bloque (133-2; 135-2; 141-2; 930) no rellenado al bloque de la señal (100) de audio, cuando no se detecta el transitorio (700, 701, 702, 703, 705, 707) en el bloque.

18. Programa informático que tiene un código de programa adaptado para realizar el método según la reivindicación 17, cuando el programa informático se ejecuta en un ordenador.

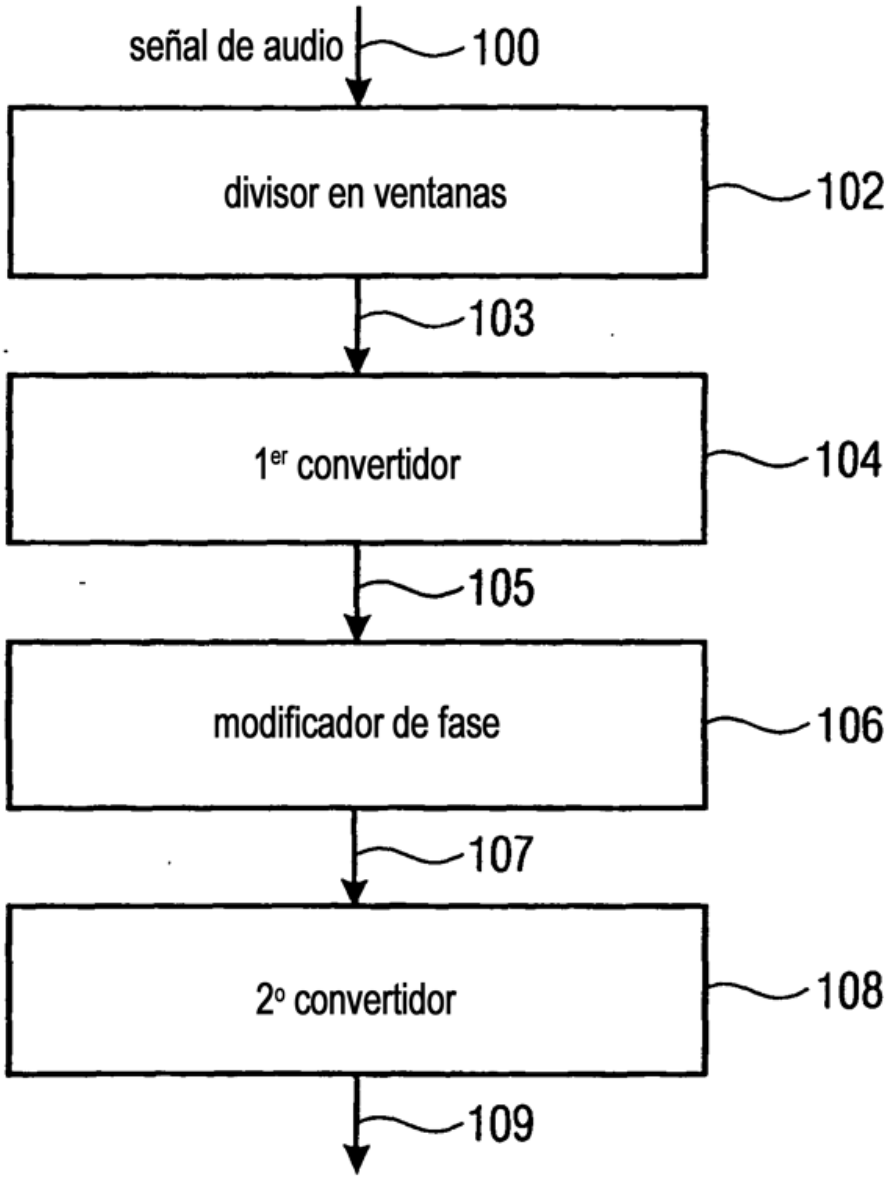
señal de audio
100
divisor en ventanas
102
103
1er convertidor
104
105
modificador de fase
106
107
2º convertidor
108
109

FIGURA 1

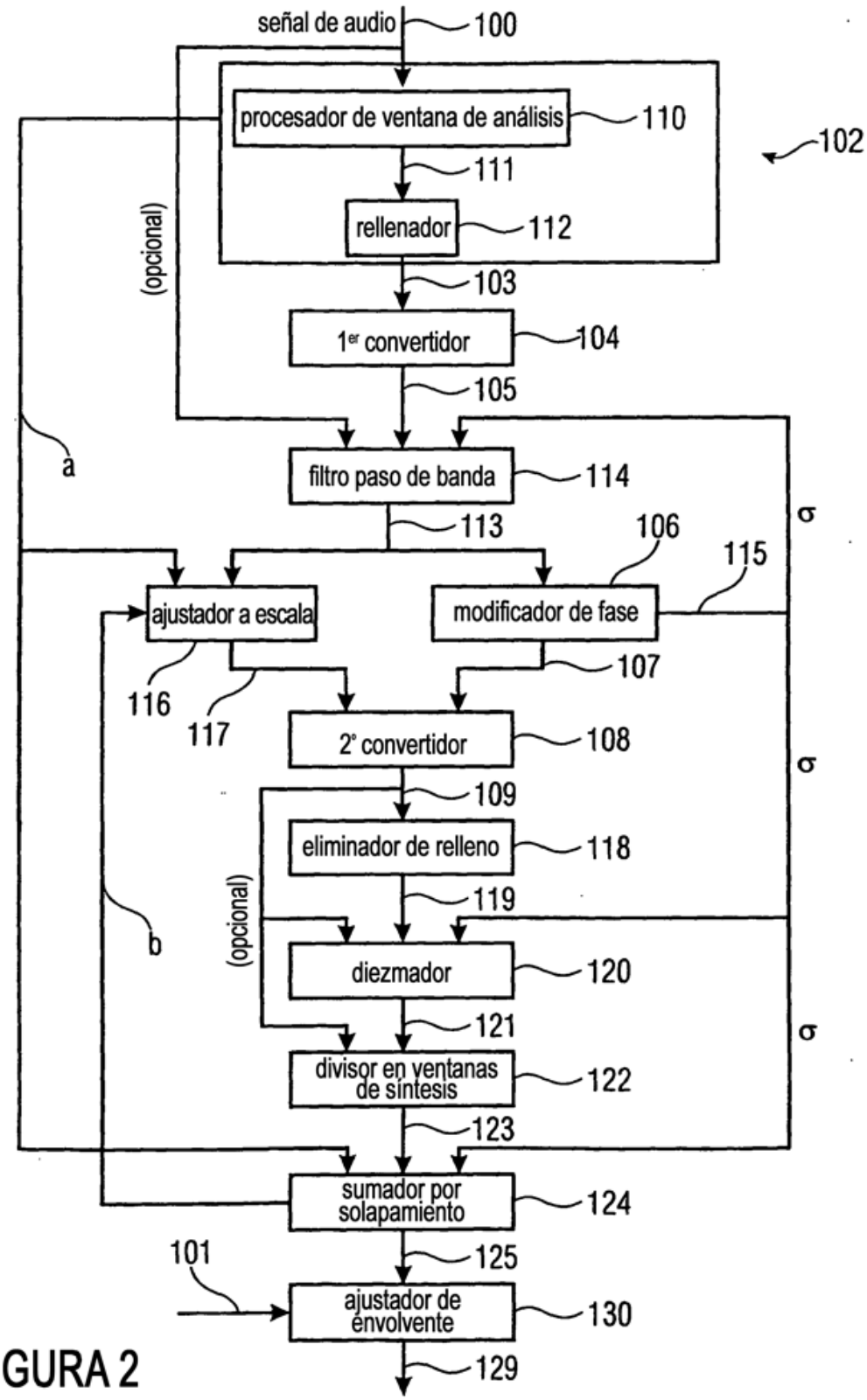
señal de audio
100
procesador de ventana de análisis
110
102
111
rellenador
112
(opcional)
103
1er convertidor
104
105
a
filtro paso de banda
114
113
106
115
σ
ajustador a escala
modificador de fase
107
116
117
2° convertidor
108
σ
109
eliminador de relleno
118
119
(opcional)
b
diezmador
120
121
σ
divisor en ventanas de síntesis
122
123
sumador por solapamiento
124
101
125
ajustador de envolvente
130
129

FIGURA 2

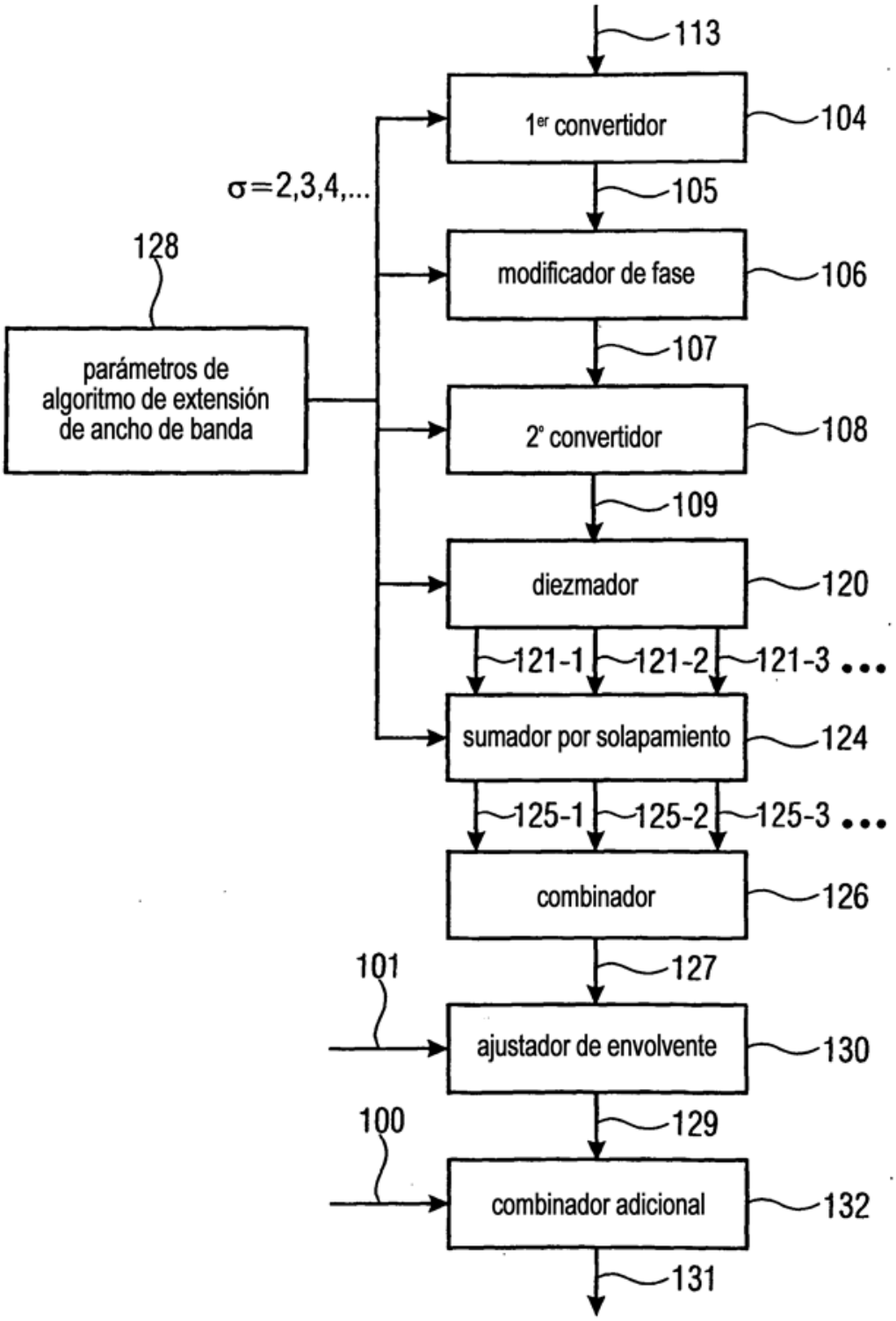
113
1er convertidor
104
σ=2,3,4,...
105
128
modificador de fase
106
parámetros de algoritmo de extensión de ancho de banda
107
2° convertidor
108
109
diezmador
120
121-1
121-2
121-3 ...
sumador por solapamiento
124
125-1
125-2
125-3 ...
combinador
126
101
127
ajustador de envolvente
130
100
129
combinador adicional
132
131

FIGURA 3

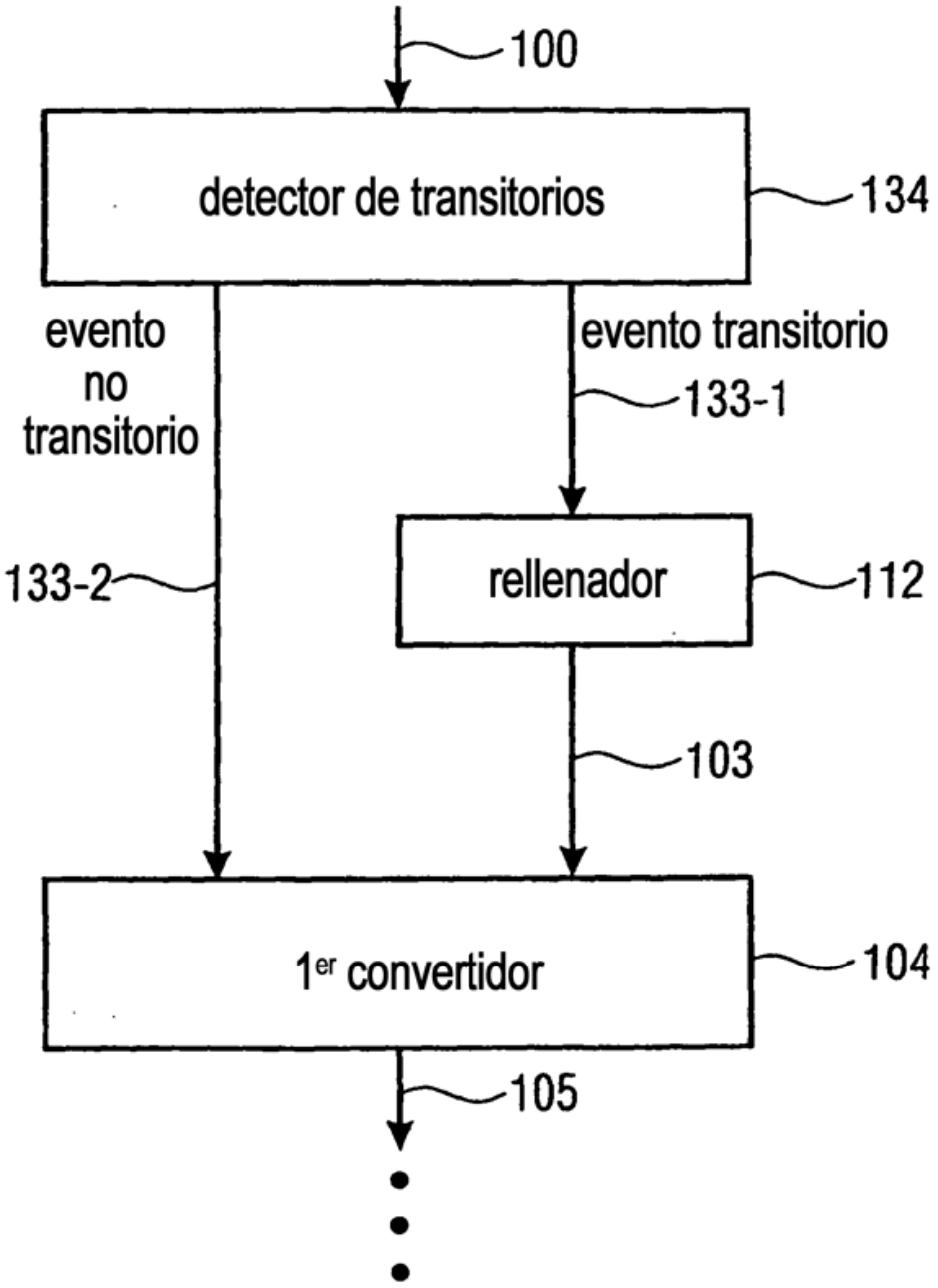
100
detector de transitorios
134
evento
no
transitorio
evento transitorio
133-1
133-2
rellenador
112
103
1er convertidor
104
105

FIGURA 4

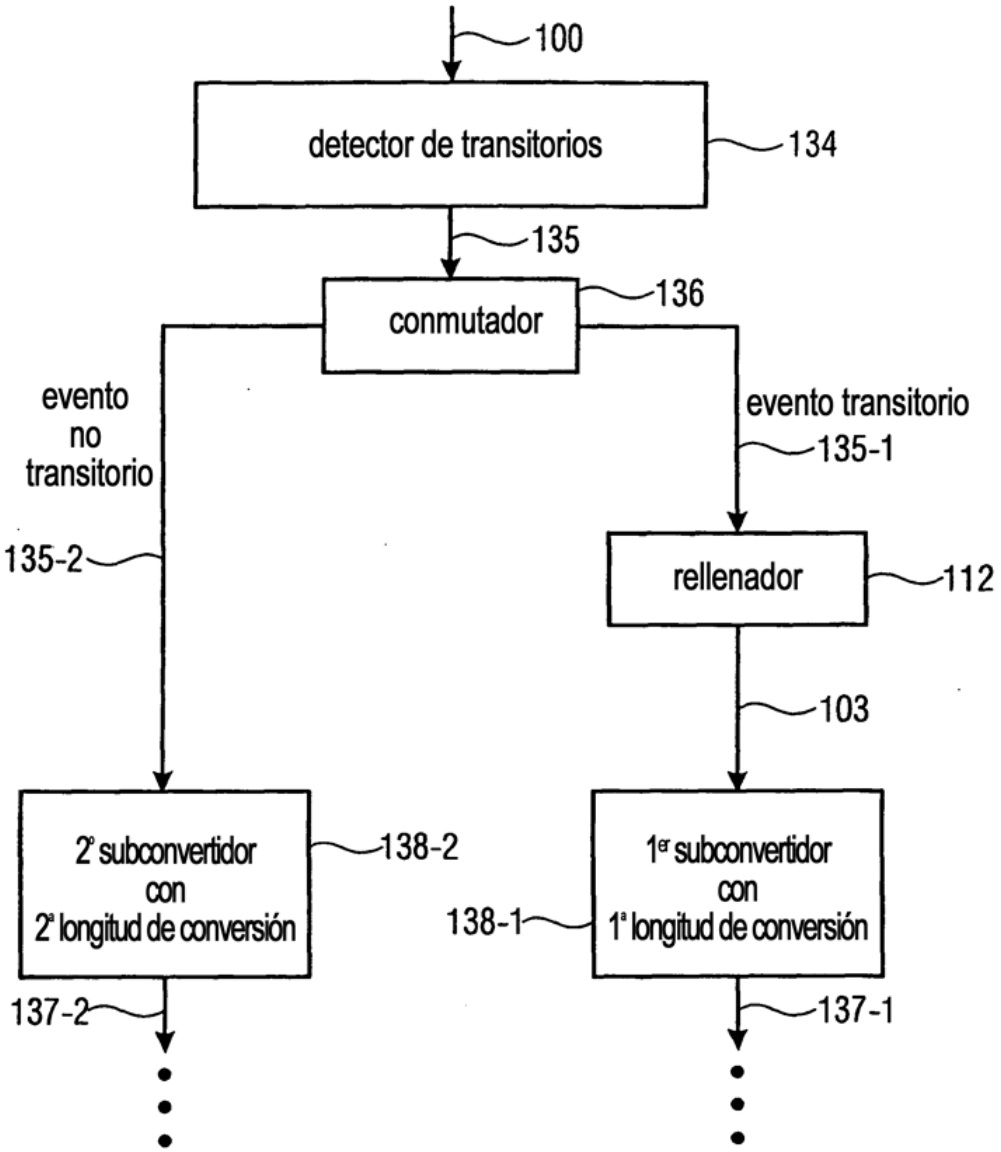
100
detector de transitorios
134
135
conmutador
136
evento
no
transitorio
evento transitorio
135-1
135-2
rellenador
112
103
2º subconvertidor
con
2ª longitud de conversión
138-2
1er subconvertidor
con
1ª longitud de conversión
138-1
137-2
137-1

FIGURA 5

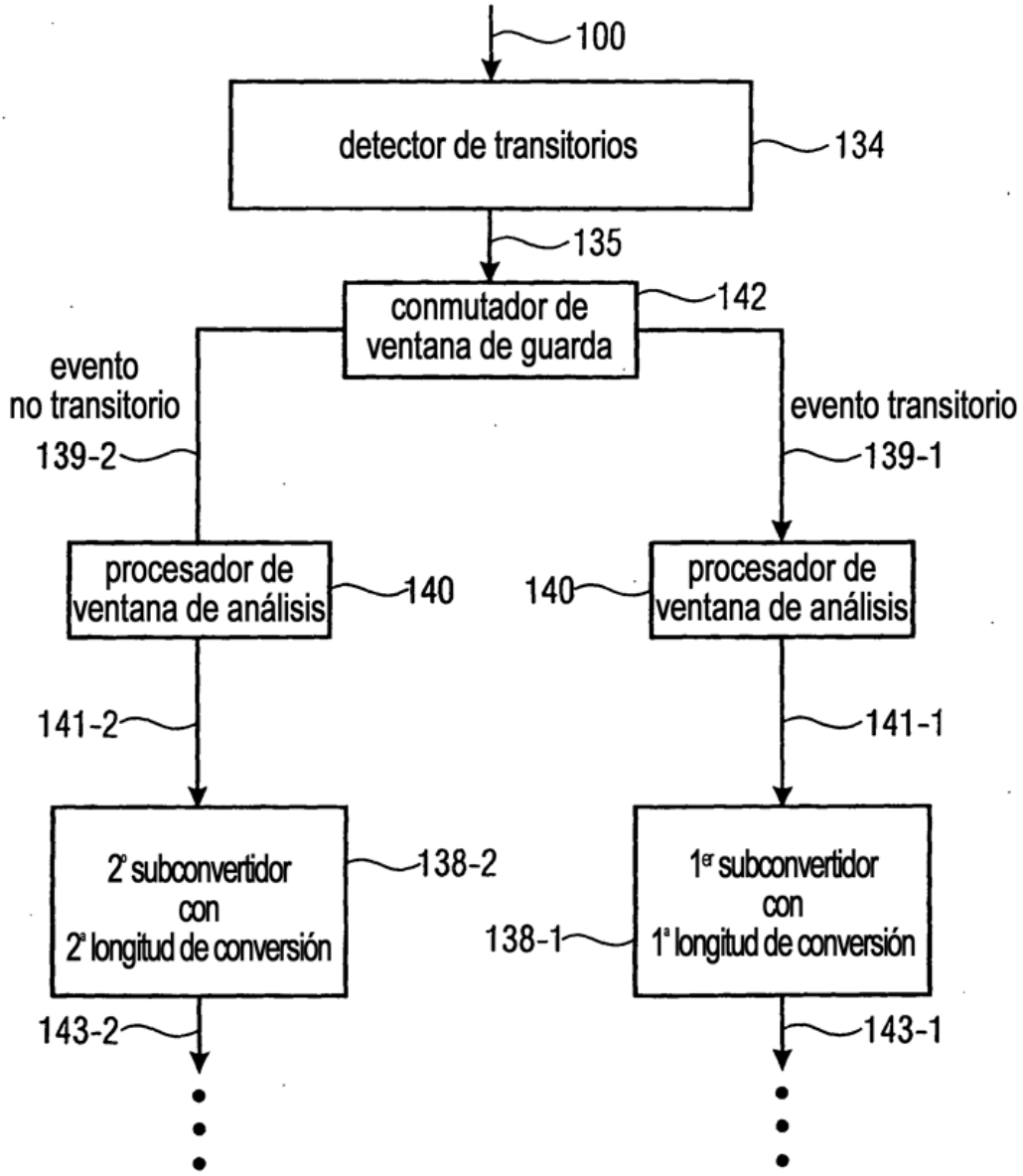
100
detector de transitorios
134
135
conmutador de ventana de guarda
142
evento no transitorio
139-2
evento transitorio
139-1
procesador de ventana de análisis
140
140
procesador de ventana de análisis
141-2
141-1
2º subconvertidor con 2ª longitud de conversión
138-2
1er subconvertidor con 1ª longitud de conversión
138-1
143-2
143-1

FIGURA 6

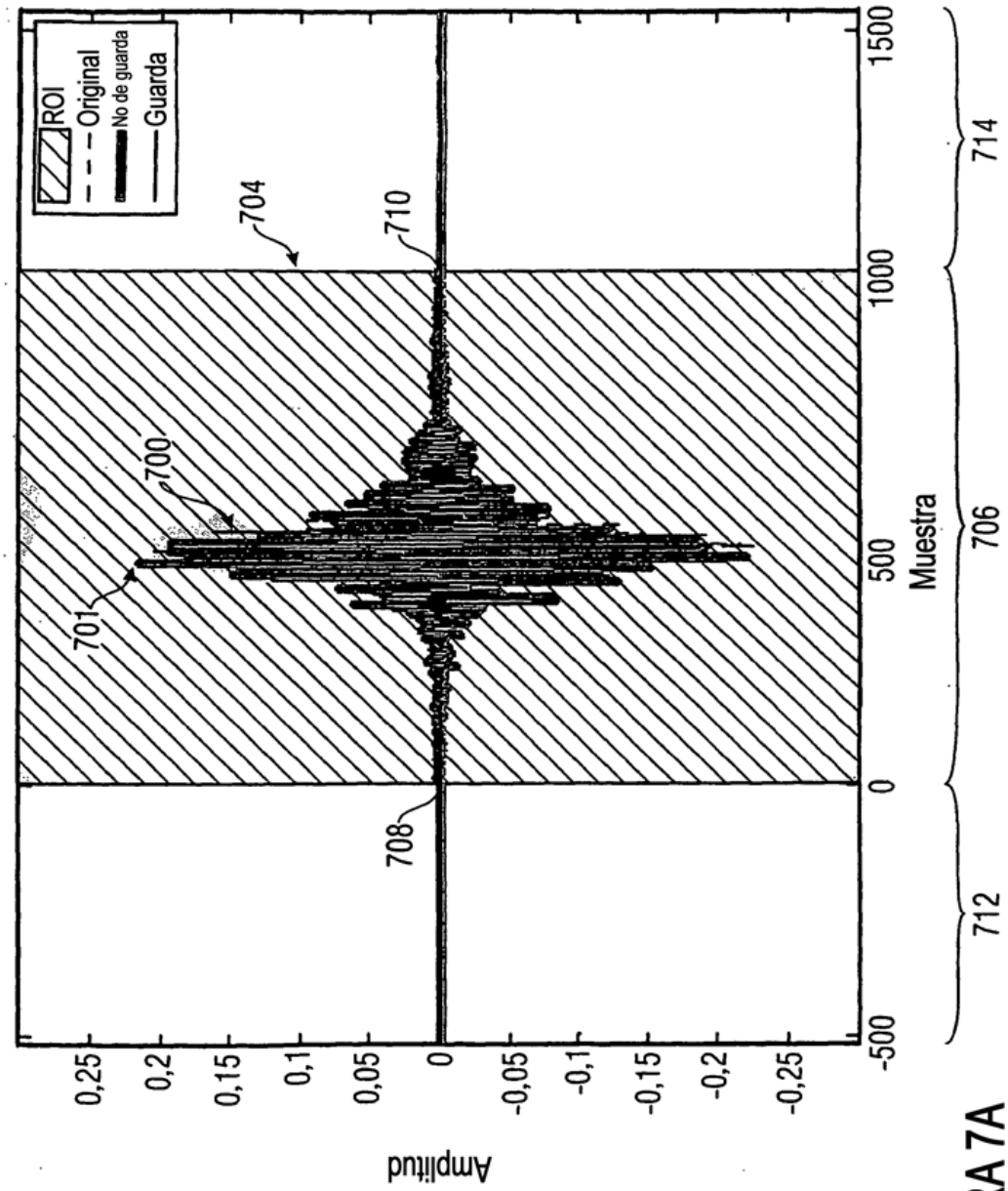
ROI
Original
No de guarda
Guarda
704
710
700
701
708
0,25
0,2
0,15
0,1
0,05
0
-0,05
-0,1
-0,15
-0,2
-0,25
Amplitud
-500
0
500
1000
1500
Muestra
712
706
714

FIGURA 7A

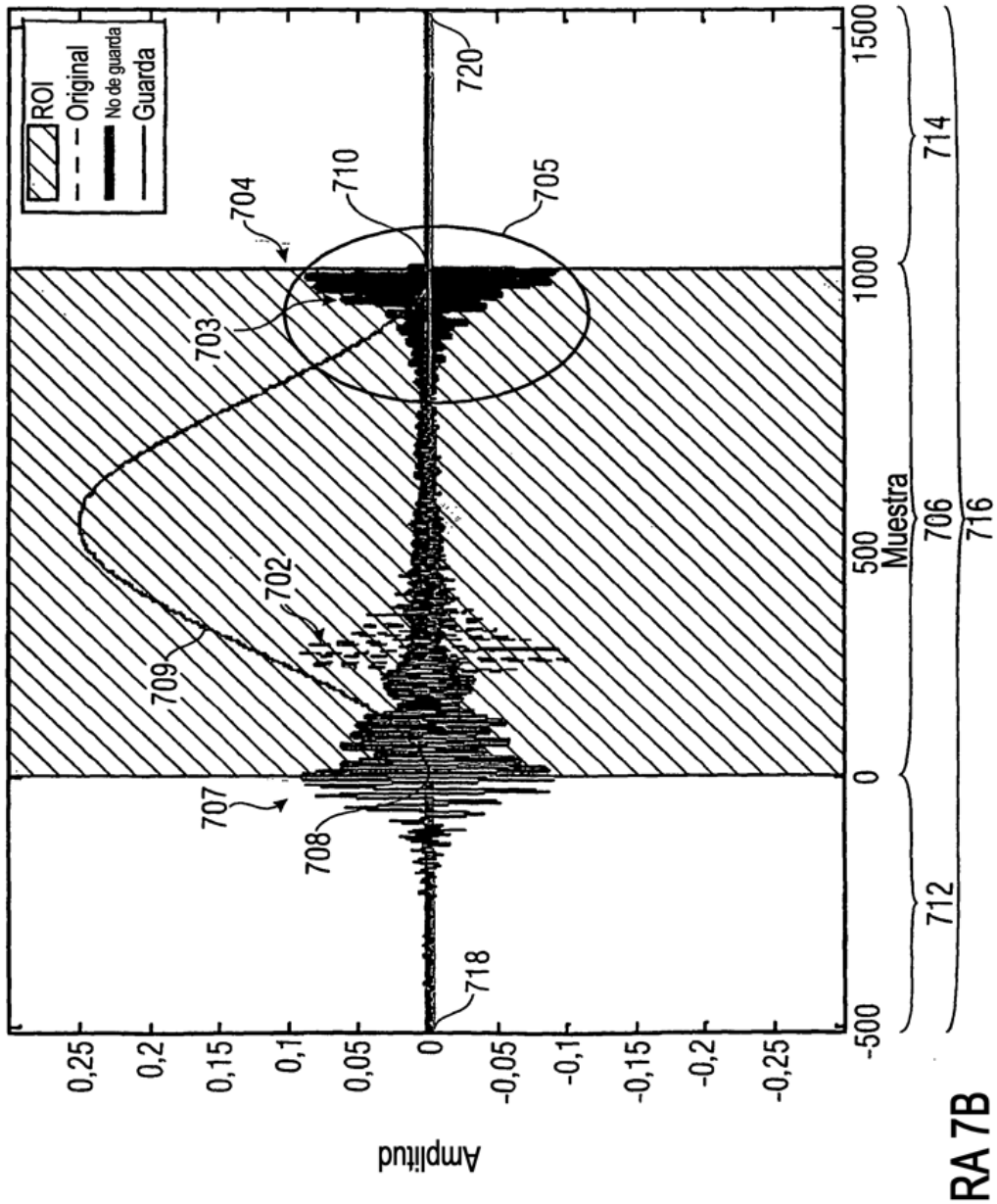
ROI
Original
No de guarda
Guarda
Amplitud
0,25
0,2
0,15
0,1
0,05
0
-0,05
-0,1
-0,15
-0,2
-0,25
-500
0
500
1000
1500
Muestra
702
703
704
705
706
707
708
709
710
712
714
716
718
720

FIGURA 7B

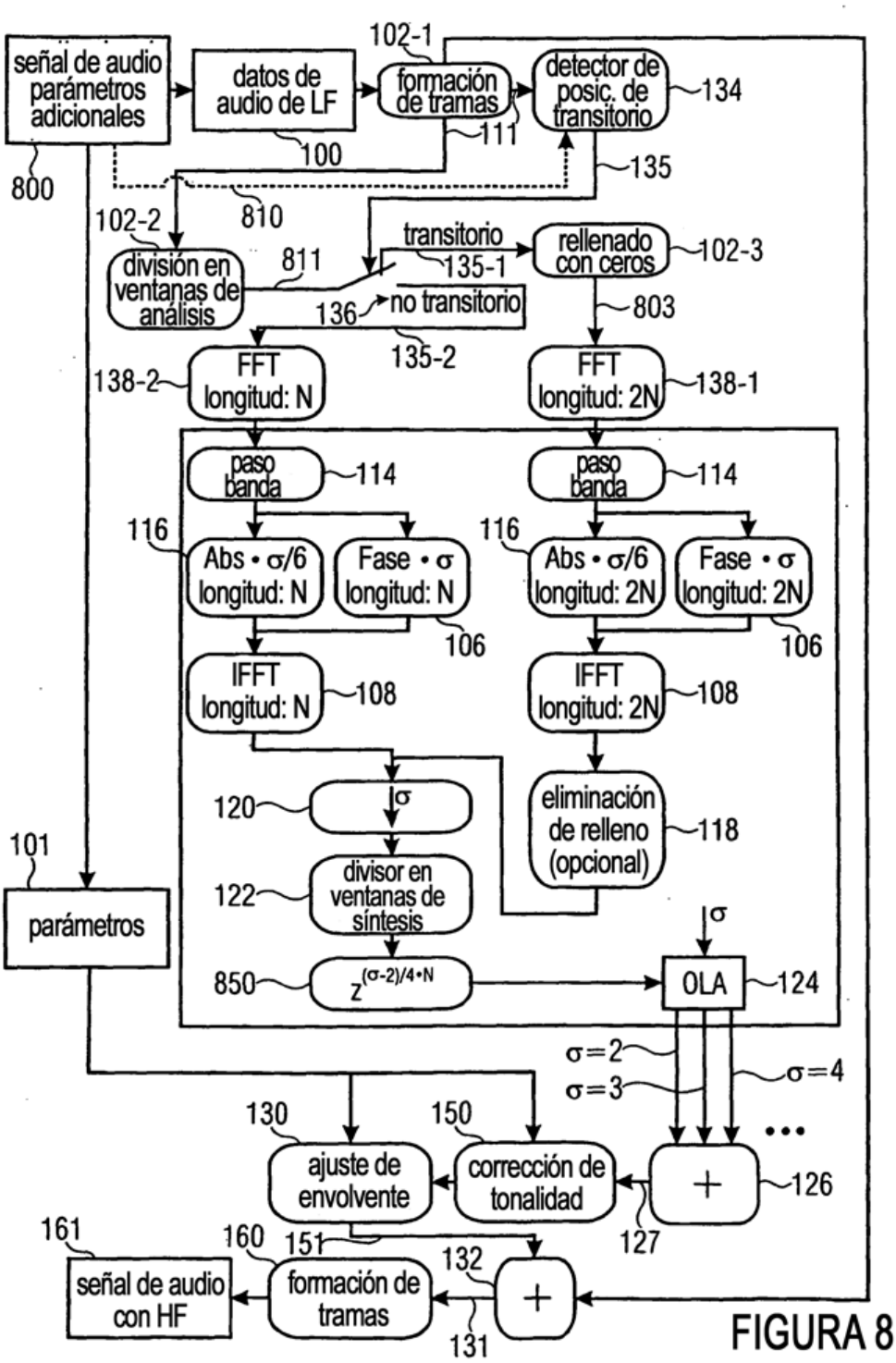
señal de audio parámetros adicionales
800
datos de audio de LF
100
102-1
formación de tramas
111
detector de posic. de transitorio
134
135
810
102-2
división en ventanas de análisis
811
transitorio
135-1
rellenado con ceros
102-3
803
136
no transitorio
135-2
138-2
FFT longitud: N
FFT longitud: 2N
138-1
paso banda
114
116
Abs • σ/6 longitud: N
Fase • σ longitud: N
Abs • σ/6 longitud: 2N
Fase • σ longitud: 2N
106
IFFT longitud: N
108
IFFT longitud: 2N
120
σ
eliminación de relleno (opcional)
118
101
parámetros
122
divisor en ventanas de síntesis
850
$z^{(\sigma-2)/4 \cdot N}$
OLA
124
σ=2
σ=3
σ=4
130
150
ajuste de envolvente
corrección de tonalidad
+
126
127
161
160
151
132
señal de audio con HF
formación de tramas
131

FIGURA 8

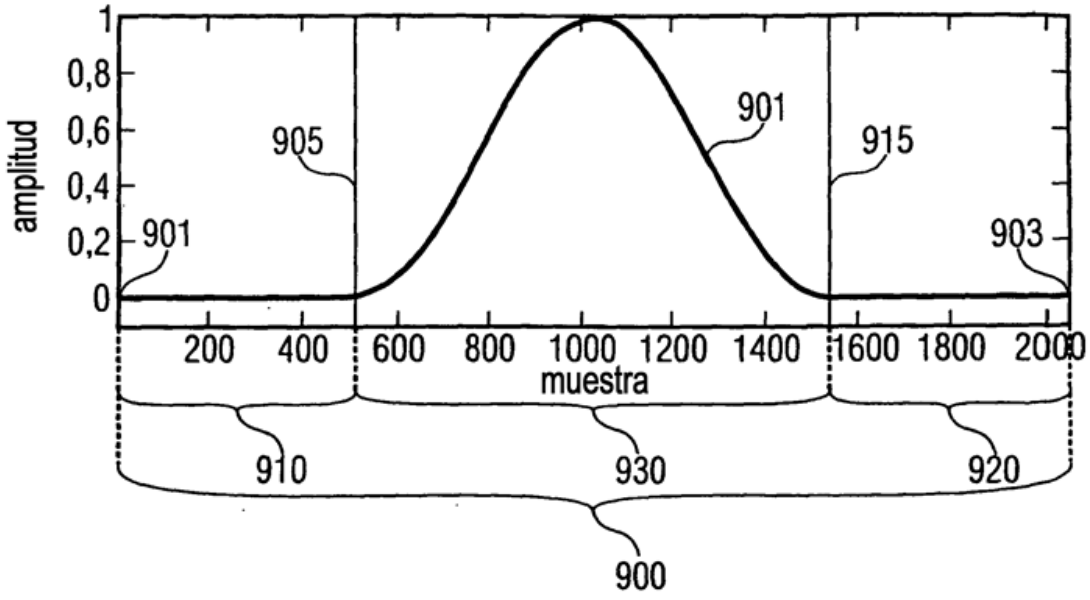

FIGURA 9A

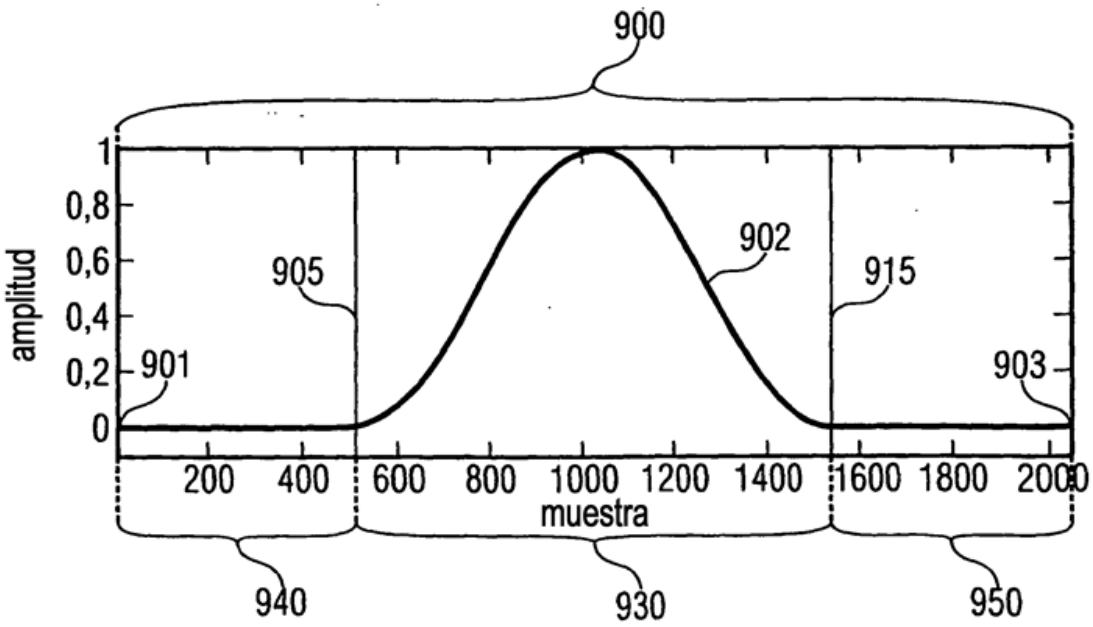

FIGURA 9B

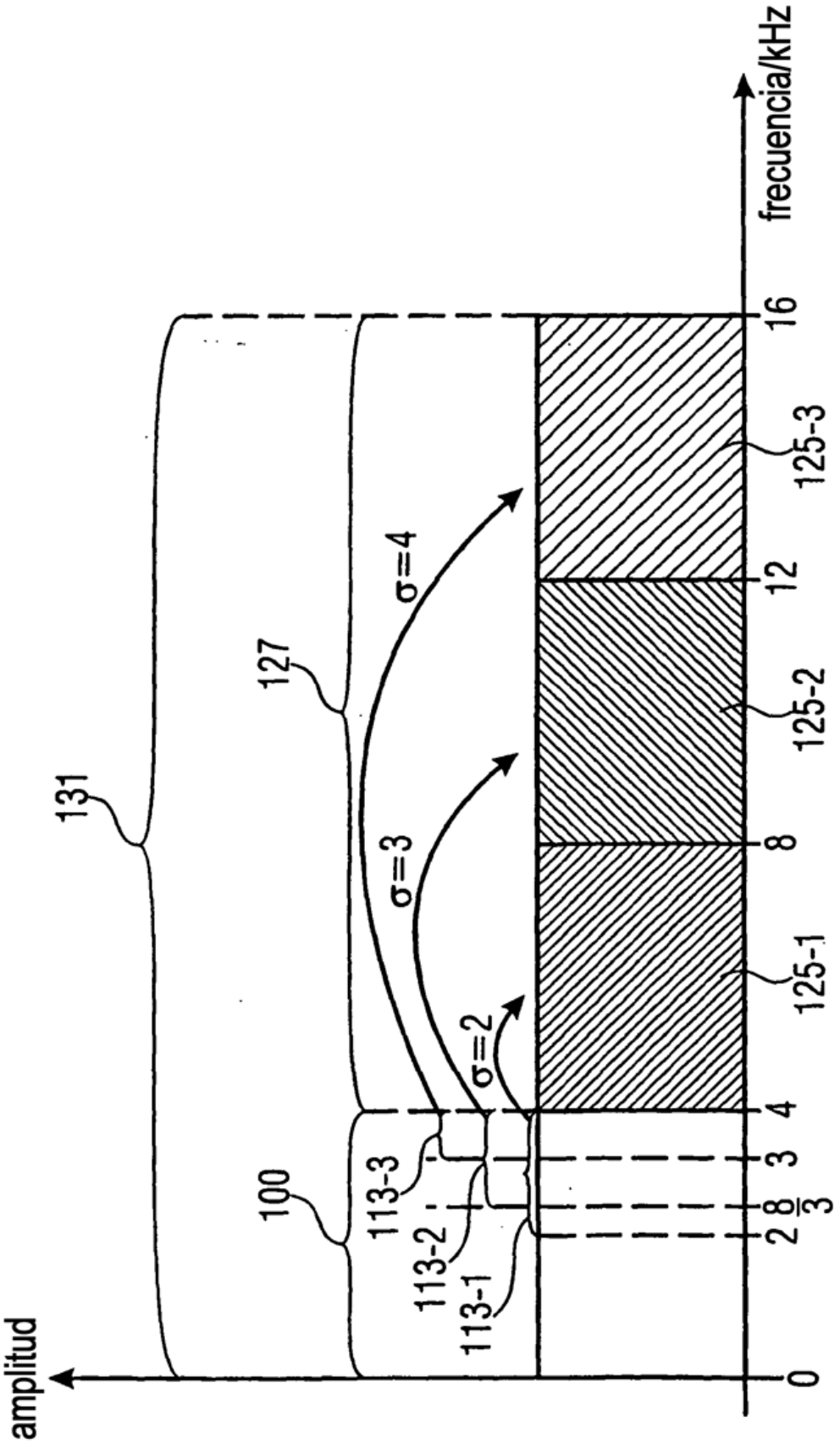
amplitud
frecuencia/kHz
131
127
100
113-3
113-2
113-1
σ=4
σ=3
σ=2
125-1
125-2
125-3
0
2 8/3
3
4
8
12
16

FIGURA 10

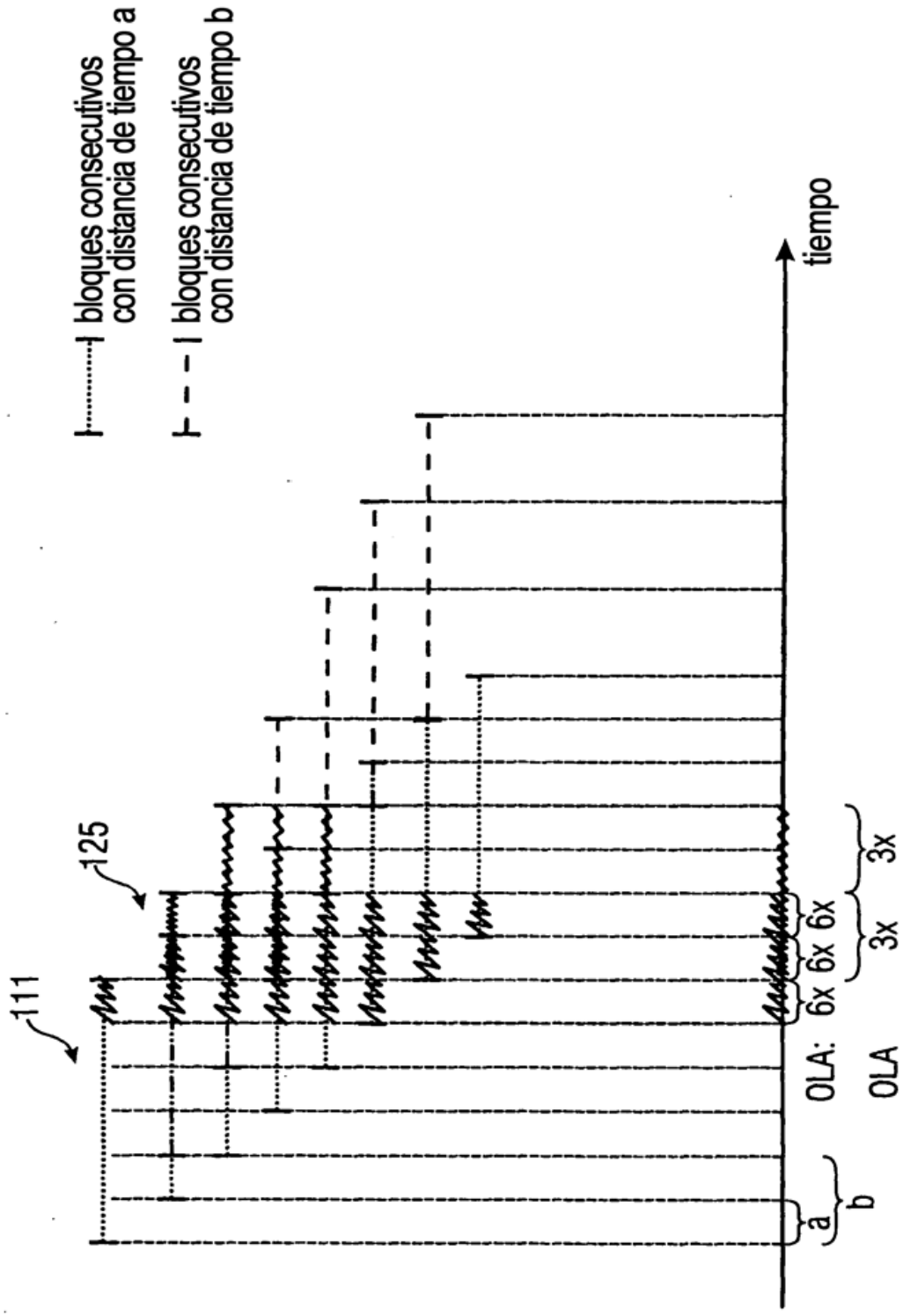
bloques consecutivos con distancia de tiempo a
bloques consecutivos con distancia de tiempo b
tiempo
111
125
a
b
OLA:
OLA
6x
6x
6x
3x
3x

FIGURA 11

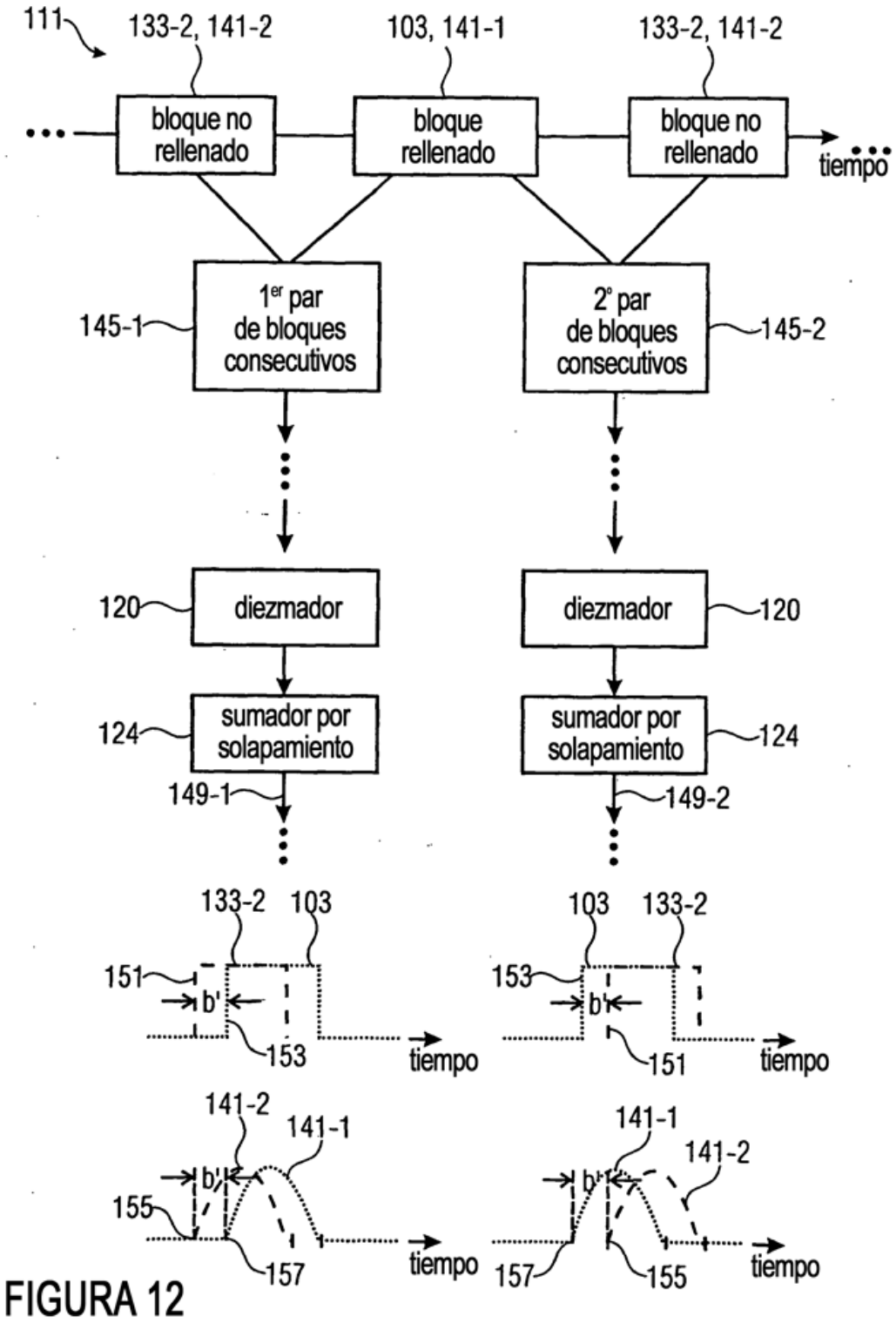
111
133-2, 141-2
103, 141-1
133-2, 141-2
bloque no rellenado
bloque rellenado
bloque no rellenado
tiempo
145-1
1er par de bloques consecutivos
2º par de bloques consecutivos
145-2
120
diezmador
diezmador
120
124
sumador por solapamiento
sumador por solapamiento
124
149-1
149-2
133-2
103
103
133-2
151
b'
153
tiempo
153
b'
151
tiempo
141-2
141-1
b'
155
157
tiempo
141-1
141-2
b'
157
155
tiempo

FIGURA 12

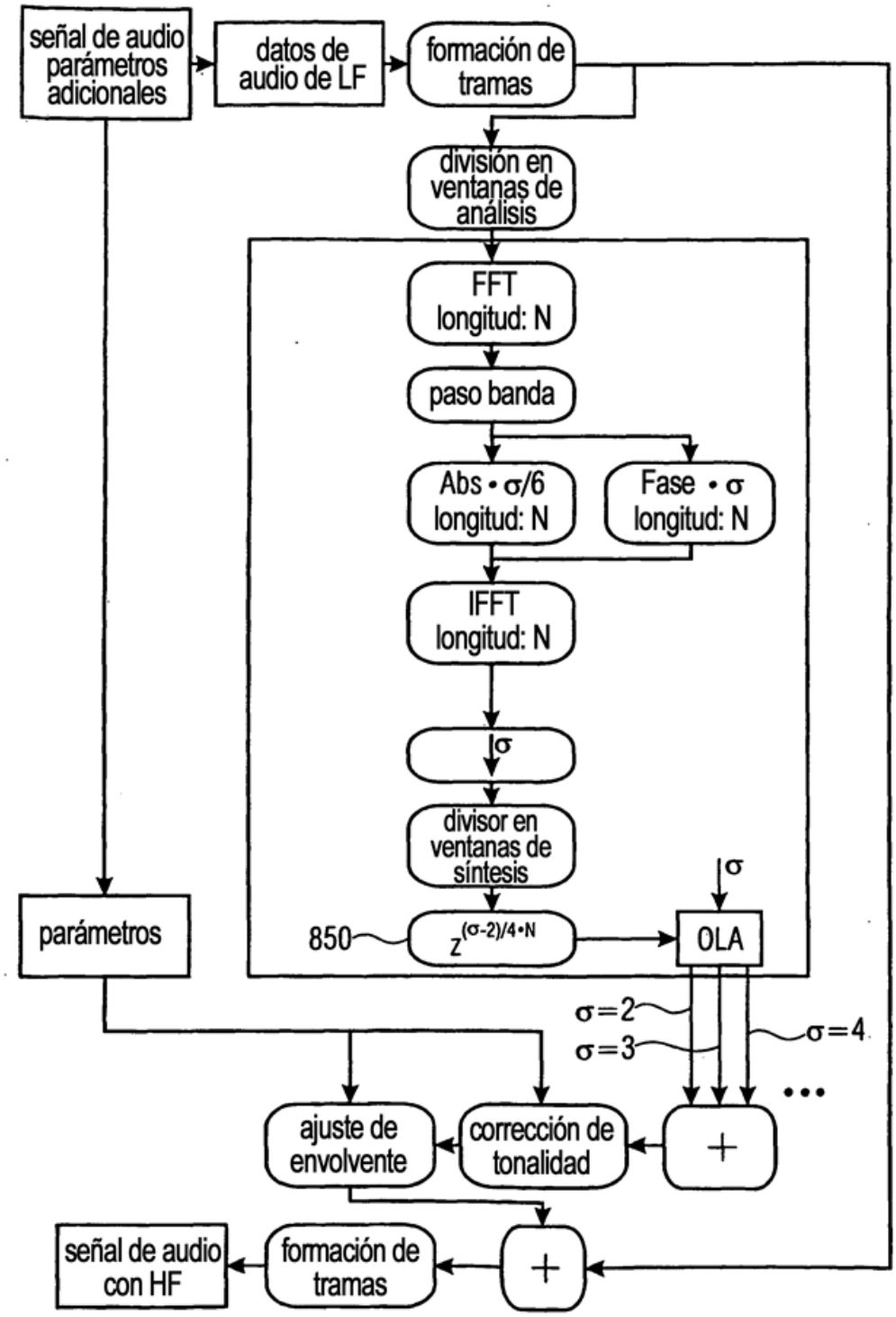
señal de audio parámetros adicionales
datos de audio de LF
formación de tramas
división en ventanas de análisis
FFT longitud: N
paso banda
Abs • σ/6 longitud: N
Fase • σ longitud: N
IFFT longitud: N
σ
divisor en ventanas de síntesis
σ
parámetros
850
z^(σ-2)/4•N
OLA
σ=2
σ=3
σ=4
...
ajuste de envolvente
corrección de tonalidad
+
señal de audio con HF
formación de tramas
+

FIGURA 13