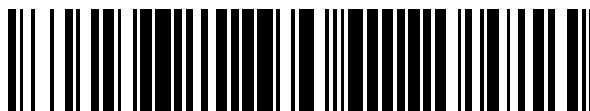


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 398 627**

51 Int. Cl.:

G10L 21/02 (2006.01)

G10L 19/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.06.2009 E 09776811 (3)**

97 Fecha y número de publicación de la concesión europea: **05.12.2012 EP 2301028**

54 Título: **Aparato y procedimiento para calcular una cantidad de envolventes espectrales**

30 Prioridad:

11.07.2008 US 79841

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
20.03.2013

73 Titular/es:

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

72 Inventor/es:

**NEUENDORF, MAX;
GRILL, BERNHARD;
KRAEMER, ULRICH;
MULTRUS, MARKUS;
POPP, HARALD;
RETTTELACH, NIKOLAUS;
NAGEL, FREDERIK;
LOHWASSER, MARKUS;
GAYER, MARC;
JANDER, MANUEL y
BACIGALUPO, VIRGILIO**

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 398 627 T3

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

DESCRIPCIÓN

Aparato y procedimiento para calcular una cantidad de envolventes espectrales

[0001] La presente invención se relaciona con un aparato y un procedimiento para calcular una cantidad de envolventes espectrales, un codificador de audio y un procedimiento para codificar señales de audio.

5 **[0002]** La codificación de audio y la codificación de voz naturales constituyen dos mayores tareas de los codificadores para las señales de audio. La codificación de audio natural es comúnmente utilizada para señales de música o señales arbitrarias a tasa de bits media y generalmente ofrece amplios anchos de banda de audio. Por otro lado, los codificadores de voz están básicamente limitados a la reproducción de voz, pero también pueden utilizarse a una tasa de bits muy baja. El discurso en banda ancha ofrece una mayor mejora de calidad subjetiva que el discurso de banda estrecha. El aumento de ancho de banda no sólo mejora la inteligibilidad y naturaleza del discurso, sino también el reconocimiento del interlocutor. La codificación de voz de banda ancha es, por lo tanto, un tema importante para la próxima generación de sistemas telefónicos. Además, debido al tremendo crecimiento del campo de la multimedia, la transmisión de música y demás señales diferentes a las de voz de alta calidad en los sistemas telefónicos es una característica deseable.

10 **[0003]** Para reducir de manera drástica la tasa de bits, se puede realizar una codificación de fuente utilizando codificadores de audio perceptuales de banda dividida. Estos codificadores de audio naturales explotan la irrelevancia perceptual y redundancia estadística en la señal. Asimismo, es común reducir la tasa de muestreo y, por lo tanto el ancho de banda de audio. Es también común disminuir la cantidad de niveles de composición, permitiendo en forma ocasional una distorsión de cuantización auditiva y emplear una degradación del campo estéreo a través de la codificación de intensidad. El uso excesivo de dichos procedimientos da por resultado una degradación perceptual engorrosa. Con el fin de mejorar el desempeño de codificación, se utiliza la replicación de banda espectral como procedimiento eficiente para generar señales de frecuencia alta en un codificador basado en la reconstrucción de frecuencia alta (HFR, por su sigla en inglés).

15 **[0004]** La replicación de banda espectral (SBR, por su sigla en inglés) comprende una técnica que obtuvo popularidad como complemento de los codificadores de audio perceptuales populares como el MP3 y la codificación de audio avanzada (AAC, por su sigla en inglés). Un ejemplo de enfoque SBR se describe en US 2008/0120116A1. SBR comprende un procedimiento de ampliación de ancho de banda donde la banda baja (banda base o banda núcleo) del espectro es codificada utilizando un codificador de vanguardia, mientras que la banda superior (o banda alta) es parametrizada en forma gruesa utilizando pocos parámetros. SBR hace uso de una correlación entre la banda baja y la banda alta prediciendo la señal con señal más ancha desde la banda inferior utilizando los rasgos de la banda alta extraída. Este procedimiento es usualmente suficiente, ya que el oído humano es menos sensible a las distorsiones en la banda más alta en comparación con la banda inferior. Los nuevos codificadores de audio, por lo tanto, codifican el espectro menor, utilizando, por ejemplo, MP3 o AAC (sigla en inglés correspondiente a la traducción codificación de audio avanzada), mientras que la banda mayor es codificada utilizando SBR. La clave para el algoritmo SBR es la información utilizada para describir la porción de frecuencia mayor de la señal. El objetivo primario del diseño de este algoritmo es reconstruir el mayor espectro de banda sin introducir artefactos (que en este contexto significa una pérdida de datos por compresión defectuosa) y proveer una buena resolución espectral y temporal. Por ejemplo, un banco de filtro polifase con valor complejo de 64 bandas es utilizado en la porción de análisis y en el codificador; el banco de filtro es utilizado para obtener, por ejemplo, muestras de energía de la banda alta de la señal de entrada original. Estos muestreos de energía pueden ser utilizados como valores de referencia para un esquema de ajuste de envolventes en el decodificador.

25 **[0005]** Las envolventes espectrales se refieren a una distribución espectral grosera de la señal en sentido general y comprenden por ejemplo, coeficientes de filtro en un codificador basado en la predicción lineal o un conjunto de promedios de tiempo-frecuencia de muestras de sub-bandas en un codificador de sub-banda. Los datos de la envolvente se refieren, a su vez, a la envolvente espectral cuantizada y codificada. Especialmente si la banda de frecuencia menor es codificada con baja, los datos de la envolvente constituyen una mayor parte de la secuencia de bits. Por ende, es importante representar la envolvente espectral en forma compacta al utilizar especialmente velocidades binarias menores.

30 **[0006]** La replicación de banda espectral utiliza herramientas, basadas en una replicación de, por ejemplo, secuencias de armonía, truncadas durante la codificación. Asimismo, ajusta la envolvente espectral de la banda alta generada y aplica un filtrado inverso y agrega componentes de ruido y armonía con el fin de recrear las características espectrales de la señal original. Por lo tanto, la entrada de la herramienta SBR comprende, por ejemplo los datos de la envolvente cuantizados, datos de control varios, una señal en dominio de tiempo desde el codificador de núcleo (por ejemplo AAC o MP3). La salida de la herramienta SBR es una señal en dominio de tiempo o una representación de señal en dominio QMF (por su sigla en inglés) (QMF = filtro espejo en cuadratura) como, por ejemplo, en caso de utilizar la herramienta de sonido envolvente MPEG. La descripción de los elementos de secuencia de bits para la carga útil SBR puede hallarse en las Normas ISO/IEC 14496-3:2005, sub-cláusula 4.5.2.8 y comprenden entre otros datos, datos de ampliación SBR, un encabezado SBR e indica el número de envolventes SBR dentro de un cuadro SBR.

[0007] Para la implementación de un SBR del lado del codificador, se realiza un análisis en la señal de entrada. La información obtenida de este análisis es utilizada para elegir la resolución de tiempo/frecuencia apropiada y la resolución de cuadro SBR actual. El algoritmo calcula los bordes de tiempo de inicio y fin de la envolvente SBR en el cuadro SBR actual, el número de envolventes SBR como su resolución de frecuencia. Las diferentes resoluciones de frecuencia se calculan de la manera descrita, por ejemplo, en la Norma ISO/IEC 14496 3 sub-cláusula 4.6.18.3. El algoritmo también calcula el número de pisos de ruido para el cuadro SBR dado y los bordes de tiempo de inicio y fin de los mismos. Los bordes de tiempo de inicio y fin de los pisos de ruido deberían ser un sub-grupo de los bordes de tiempo de inicio y fin de las envolventes espectrales. El algoritmo divide el cuadro SBR actual en cuatro clases:

FIXFIX – El borde de tiempo frontal y posterior son iguales a los límites del cuadro SBR nominal. Todos los bordes de tiempo de la envoltura SBR en el cuadro están uniformemente distribuidos en el tiempo. El número de envolventes es una potencia entera de dos (1,2,4,8,...).

FIXVAR – El borde de tiempo frontal es igual al límite de cuadro nominal frontal. El borde de tiempo posterior es variable y puede definirse por elementos de secuencia de bits. Todos los bordes de tiempo de la envolvente SBR entre el borde de tiempo frontal y el borde de tiempo posterior pueden especificarse como la distancia relativa en franjas de tiempo al borde previo, comenzando desde el borde de tiempo posterior.

VARFIX – El borde de tiempo frontal es variable y se define por elementos de secuencia de bits. El borde de tiempo posterior es igual al límite de cuadro nominal. Todos los bordes de tiempo de la envolvente SBR entre el borde de tiempo frontal y el borde de tiempo posterior pueden especificarse en la secuencia de bits la distancia relativa en franjas de tiempo al borde previo, comenzando desde el borde de tiempo frontal.

VARVAR – El borde de tiempo frontal y posterior son variables y pueden definirse en la secuencia de bits. Todos los bordes de tiempo de la envolvente SBR entre el borde de tiempo frontal y el borde de tiempo posterior pueden también ser especificados. Los bordes de tiempo relativos comenzando desde el borde de tiempo frontal se especifican como la distancia relativa al borde previo. Los bordes de tiempo relativos comenzando desde el borde de tiempo posterior se especifican como la distancia relativa al borde previo.

[0008] No existen restricciones en las transiciones de clase de cuadro SBR, es decir, toda secuencia de clases es permitida en la Norma. Sin embargo, de acuerdo con esta Norma, la cantidad máxima de envolventes SBR por cuadro SBR es restringida a 4 para la clase FIXFIX y 5 para la clase VARVAR. Las clases FIXVAR y VARFIX están sintácticamente limitadas a cuadro envolventes SBR. Las envolventes espectrales del cuadro SBR se estiman por el segmento de tiempo con la resolución de frecuencia dada por la rejilla de tiempo/frecuencia. La envolvente SBR se estima promediando las muestras de sub-bandas cuadráticas con las regiones de tiempo/frecuencia dadas.

[0009] Los transitorios reciben en SBR, en general, un tratamiento específico empleando envolventes específicas de longitudes variables. Los transitorios pueden ser definidos por porciones dentro de las señales convencionales, donde un fuerte aumento en la energía aparece dentro de un corto período de tiempo, que puede o no estar limitado en una región específica de frecuencia. Los ejemplos para transitorios son golpes de castañuelas y de instrumentos de percusión, pero también ciertos sonidos de la voz humana como, por ejemplo, las letras: P, T, K, La detección de este tipo de transitorios es implementada siempre de la misma forma o por el mismo algoritmo (utilizando un valor umbral de transitorios), independiente de la señal, ya sea clasificado como voz o música. Además, una posible distinción entre el discurso vocalizado y no vocalizado no tiene influencia en el mecanismo de detección de transitorios convencional o clásico.

[0010] En consecuencia, en caso de detectar un transitorio, los datos SBR deberían ajustarse para que un decodificador pueda replicar el transitorio detectado en forma apropiada. En WO 01/26095, un aparato y procedimiento se divulga para la codificación de envolvente espectral, que tiene en cuenta un transitorio detectado en la señal de audio. En este procedimiento convencional, un muestreo de tiempo y frecuencia no uniforme de la envolvente espectral se logra agrupando en forma adaptativa muestras de sub-bandas desde un banco de filtro con tamaño fijo en bandas de frecuencia y segmentos de tiempo, cada uno de ellos genera una muestra de envolvente. El sistema correspondiente falla en segmentos de largo tiempo y resolución de alta frecuencia, pero en las cercanías de un transitorio, se utilizan segmentos de tiempo menores, donde se pueden usar pasos de mayor frecuencia con el fin de mantener el tamaño de los datos dentro de los límites. En caso de detectar un transitorio, el sistema conmuta desde un cuadro FIXFIX a un cuadro FIXVAR seguido de un cuadro VARFIX de modo que un borde de envolvente se fije justo antes del transitorio detectado. Este procedimiento se repite toda vez que se detecta un transitorio.

[0011] En caso que la fluctuación de energía cambie sólo lentamente, el detector de transitorio no detectará el cambio. Estos cambios pueden, sin embargo, ser lo suficientemente fuertes para generar artefactos perceptibles si no se tratan en forma apropiada. Una solución simple sería disminuir el valor umbral en el detector de transitorio. Sin embargo, esto resultaría en una conmutación frecuente entre diferentes cuadros (FIXFIX a FIXVAR + VARFIX). En consecuencia, una cantidad significativa de datos adicionales deben ser transmitidos implicando una pobre eficiencia de codificación – especialmente si el lento incremento perdura en el tiempo (por ejemplo por múltiples cuadros). Esto no es aceptable, ya que la señal no comprende la complejidad, que justificaría una velocidad mayor de datos y por lo tanto no constituye una opción para solucionar el problema.

[0012] Un objetivo de la presente invención es por lo tanto proveer un aparato, que permita una codificación eficiente sin artefactos perceptibles, especialmente para señales que comprenden una energía que varía lentamente, que es demasiado baja para ser detectada por el detector de transitorios.

[0013] Este objetivo se alcanza mediante el uso del aparato de acuerdo con la reivindicación 1, el codificador de acuerdo con la reivindicación 11, un procedimiento para calcular la cantidad de envolventes espectrales de acuerdo con la reivindicación 13 o un procedimiento para generar una secuencia de datos de acuerdo con la reivindicación 14.

[0014] La presente invención se basa en el principio que la calidad perceptiva de la señal de audio transmitida puede incrementarse ajustando de modo flexible la cantidad de envolventes espectrales dentro de un cuadro SBR de acuerdo con una señal dada. Este objetivo se alcanza comparando la señal de audio de las porciones de tiempo vecinas dentro del cuadro SBR. La comparación se realiza determinando distribuciones de energía para la señal de audio dentro de las porciones de tiempo, y un valor de decisión mide una desviación de las distribuciones de energía de las dos porciones de tiempo vecinas. Dependiendo de si los valores de decisión violan un valor umbral, un borde de envolvente se encuentra entre las porciones de tiempo vecinas. El otro borde de la envolvente puede estar al inicio o al final del cuadro SBR o, alternativamente, también entre otras dos porciones de tiempo vecinas dentro del cuadro SBR.

[0015] En consecuencia, el cuadro SBR no está adaptado o cambia como, por ejemplo, en un aparato convencional donde un cambio del cuadro FIXFIX al cuadro FIXVAR o al cuadro VARFIX se lleva a cabo con el fin de tratar transitorios. En cambio, las formas de realización usan una variada cantidad de envolventes, por ejemplo dentro de los cuadros FIXFIX, con el fin de tomar en cuenta las fluctuaciones variantes de la señal de audio para que aún señales que varían lentamente puedan resultar en una cantidad cambiante de envolventes y así permitir una mejor calidad de audio a producir por la herramienta SBR en un decodificador. Las envolventes determinadas pueden, por ejemplo, cubrir porciones de igual longitud de tiempo dentro del cuadro SBR. Por ejemplo, el cuadro SBR puede dividirse en una cantidad predeterminada de porciones de tiempo (las cuales pueden, por ejemplo, comprender 4, 8 u otras potencias enteras de 2).

[0016] La distribución de energía espectral de cada porción de tiempo puede cubrir sólo la banda de frecuencia superior, replicada por SBR. Por otro lado, la e distribución de energía espectral puede también estar relacionada con toda la banda de frecuencia (superior e inferior), donde la banda de frecuencia superior puede o no ser ponderada más que la banda de frecuencia inferior. Mediante este procedimiento, sólo una violación del valor umbral puede ser suficiente para aumentar la cantidad de envolventes o para utilizar la máxima cantidad de envolventes dentro del cuadro SBR.

[0017] Otras formas de realización pueden también comprender una herramienta clasificadora de señal, que analiza la señal de entrada original y genera información de control desde la misma, que provoca la selección de diferentes modos de codificación. Los diferentes modos de codificación pueden, por ejemplo, comprender un codificador de voz y un codificador de audio general. El análisis de la señal de entrada depende de la implementación con el objetivo de elegir el óptimo modo de codificación de núcleo para un cuadro de señal de entrada dado. El óptimo modo se relaciona con equilibrar una alta calidad perceptiva utilizando a la vez una baja tasa de bits para la codificación. La entrada de la herramienta clasificadora de señal puede ser la señal de entrada no modificada original y/o parámetros adicionales que dependen de la implementación. La salida de la herramienta clasificadora de señal puede, por ejemplo, ser una señal de control para controlar la selección del codificador de núcleo.

[0018] Si, por ejemplo, la señal se identifica o clasifica como de voz, la resolución tipo tiempo de la ampliación de ancho de banda (BWE) puede aumentar (por ejemplo por más envolventes) para que una fluctuación de energía del tipo tiempo (fluctuación lenta o fuerte) pueda ser mejor tomada en cuenta.

[0019] Este procedimiento toma en cuenta que diferentes señales con diferentes características de tiempo/frecuencia poseen diferentes demandas en las características de ampliación de ancho de banda. Por ejemplo, las señales de transitorios (que aparecen, por ejemplo, en señales de voz) necesitan una resolución temporal fina de BWE, la frecuencia de cruce (que significa el borde de la frecuencia superior del codificador de núcleo) debería ser lo más alta posible. Especialmente en el discurso vocalizado, una estructura temporal distorsionada puede disminuir la calidad percibida. Por otro lado, las señales tonales necesitan a menudo una reproducción estable de componentes espectrales y un patrón armónico concordante de las porciones de frecuencia alta reproducida. La reproducción estable de las partes tonales limita el ancho de banda del codificador de núcleo - no necesita un BWE con el temporal fino, sino una resolución espectral más fina. En un diseño de codificador de núcleo de voz/ audio conmutado, es además posible utilizar la decisión del codificador de núcleo para adaptar, las características temporales y espectrales de BWE y también adaptar el ancho del codificador de núcleo a las características de la señal.

[0020] Si todas las envolventes comprenden la misma longitud en el tiempo, dependiendo de la violación detectada (en qué momento), la cantidad de envolventes podrá diferir de cuadro en cuadro. Las formas de realización determinan la cantidad de envolventes para un cuadro SBR, por ejemplo, de la siguiente manera. Es posible comenzar con una partición con una cantidad máxima posible de envolventes (por ejemplo, 8) y reducir la cantidad de envolventes paso a paso para que dependiendo de la señal de entrada, no se utilicen más envolventes de las necesarias para permitir una reproducción de la señal en una calidad alta perceptivo.

5 **[0021]** Por ejemplo, una violación ya detectada en el primer borde de las porciones de tiempo dentro del cuadro pueden dar por resultado un máximo número de envolventes, mientras que una violación sólo detectada en el segundo borde puede dar por resultado la mitad del número máximo de envolventes. Con el fin de reducir los datos a transmitir, en otras formas de realización el valor umbral puede depender del instante de tiempo (es decir, dependiendo de qué borde es actualmente analizado). Por ejemplo, entre la primera y segunda porción de tiempo (primer borde) y entre la tercera y cuarta porción de tiempo (tercer borde) el valor umbral puede en ambos casos ser mayor al umbral entre la segunda y tercera porción de tiempo (segundo borde). En consecuencia, estadísticamente habrá más violaciones en el segundo borde que en el primer o tercer borde y, por lo tanto, pocas envolventes son posibles, lo que sería de preferencia (para más detalles ver a continuación).

10 **[0022]** En otras formas de realización la longitud en el tiempo de una porción de tiempo de la cantidad predeterminada de subsecuentes porciones de tiempo es igual a una longitud mínima en el tiempo, para la cual una sola envolvente es determinada, y en al cual el calculador de valor de decisión está adaptado para calcular un valor de decisión para dos porciones de tiempo vecinas con la mínima longitud de tiempo.

15 **[0023]** Otras formas de realización comprenden un procesador de información para proporcionar información lateral adicional, la información lateral adicional comprende el primer borde de envolvente y el segundo borde de envolvente dentro de al secuencia de tiempo de la señal de audio. En otras formas de realización el detector está adaptado para investigar en un orden temporal cada borde entre las porciones de tiempo vecinas.

20 **[0024]** Las formas de realización también utilizan el aparato para calcular la cantidad de envolventes dentro de un codificador. El codificador comprende el aparato para calcular la cantidad de la envolvente espectral y un calculador de envolvente utiliza esta cantidad para calcular los datos de la envolvente espectral para un cuadro SBR. Las formas de realización también comprenden un procedimiento para calcular la cantidad de envolventes y un procedimiento para codificar una señal de audio.

25 **[0025]** Por lo tanto, el uso de envolventes dentro de los cuadros FIXFIX tiene por objetivo un mejor modelamiento de la fluctuación de energía, que no están cubiertos por dichos tratamientos de transitorios, ya que son demasiado lentos con el fin de detectarlos como transitorios o clasificarlos como transitorios. Por otro lado, son lo suficientemente rápidos para originar artefactos si no se tratan en forma apropiada, debido a la resolución del tipo tiempo insuficiente. En consecuencia, el tratamiento de envolvente de acuerdo con la presente invención tendrá en cuenta fluctuaciones de energía que varían lentamente y no sólo las fluctuaciones de energía fuertes o rápidas, que son características de las transitorios. Por ende, las formas de realización de la presente invención permiten una codificación más eficiente en una mejor calidad, especialmente para señales con energía de variación lenta, cuya intensidad de fluctuación es demasiado baja para ser detectada por el detector de transitorios convencional.

Breve Descripción de los Dibujos

35 **[0026]** La presente invención será descrita por los ejemplos ilustrados. Las características de la invención serán apreciadas con facilidad y tendrán una mejor comprensión al hacer referencia a la siguiente descripción detallada, que debería ser considerada con referencia a los dibujos adjuntos, donde:

La Fig. 1 muestra un diagrama de bloque de un aparato para calcular una cantidad de envolventes espectrales de acuerdo con formas de realización de la presente invención;

La Fig. 2 muestra un diagrama de bloque de un módulo SBR que comprende un calculador de cantidad de envolvente;

40 Las Figs. 3a y 3b muestran diagramas de bloques de un codificador que comprende un calculador de cantidad de envolvente;

La Fig. 4 ilustra la partición de un cuadro SBR en una cantidad predeterminada de cantidad de porciones de tiempo;

45 Las Figs. 5a a 5c muestran otras particiones para un cuadro SBR que comprende tres envolventes que cubren diferentes cantidades de porciones de tiempo;

Las Figs. 6a y 6b ilustran la distribución de energía espectral para señales dentro de las porciones de tiempo vecinas; y

Las Figs. 7a a 7c muestran un codificador que comprende un conmutador opcional de audio/voz que da por resultado diferentes resoluciones temporales para una señal de audio.

50 Descripción Detallada de la invención

[0027] Las formas de realización descritas a continuación son meramente ilustrativas para el principio de la presente invención para mejorar la replicación de banda espectral, por ejemplo, utilizada dentro de un codificador de audio. Se entiende que las modificaciones y variaciones de las disposiciones y detalles descritos en la presente resultarán

evidentes para los expertos en la técnica. Por lo tanto se intenta no quedar limitado a los detalles específicos presentados como descripción y explicación de las formas de realización establecidas en la presente.

[0028] La Fig. 1 muestra un aparato 100 para calcular una cantidad 102 de envolventes espectrales 104. Las envolventes espectrales 104 derivan del codificador de replicación de banda espectral, donde el codificador está adaptado para codificar una señal de audio 105 utilizando una pluralidad de valores de muestra dentro de una cantidad predeterminada de porciones de tiempo subsecuentes 110 en el cuadro de replicación de banda espectral (cuadro SBR) que se extiende desde un tiempo inicial t_0 a un tiempo final t_n . La cantidad predeterminada de porciones de tiempo subsecuentes 110 se dispone en una secuencia de tiempo dada por la señal de audio 105.

[0029] El aparato 100 comprende un calculador de valor de decisión 120 para determinar un valor de decisión 125, donde el valor de decisión 125 mide una desviación en la distribución de energía espectral de un par de porciones de tiempo vecinas. El aparato 100 además comprende un detector de violación 130 para detectar una violación 135 de valor umbral por un valor de decisión 125. Además, el aparato 100 comprende un procesador 140 (primer procesador de determinación de borde) para determinar un primer borde de envolvente 145 entre el par de porciones de tiempo vecinas cuando se detecta una violación 135 del valor umbral. El aparato 100 también comprende un procesador 150 (segundo procesador de determinación de borde) para determinar un segundo borde de envolvente 155 entre un par diferente de porciones de tiempo vecinas o en el tiempo inicial t_0 o tiempo final t_n para una envolvente 104 con el primer borde de t envolvente 145 basado en una violación 135 del valor umbral para el otro par o basado en una posición temporal del par o el otro par en el cuadro SBR. Finalmente, el aparato 100 comprende un procesador 160 (procesador de cantidad de envolvente) para establecer la cantidad 102 de envolventes espectrales 104 con el primer borde de envolvente 145 y el segundo borde de envolvente 155.

[0030] Otras formas de realización comprenden un aparato 100, donde una longitud de tiempo de una porción de tiempo de la cantidad predeterminada de la porción de tiempo subsecuente 110 es igual a la longitud mínima en el tiempo para la cual se determina una sola envolvente 104. Asimismo, el calculador de valor de decisión 120 está adaptado para calcular un valor de decisión 25 para dos porciones de tiempo vecinas con longitud mínima en el tiempo.

[0031] La Fig. 2 muestra una forma de realización para una herramienta SBR que comprende el calculador de cantidad de envolvente 100 (Fig. 1), que determina la cantidad de 102 envolventes espectrales 104 al procesar la señal de audio 105. La cantidad 102 es ingresada en un calculador de envolvente 210, que calcula los datos de la envolvente 205 desde la señal de audio 105. Utilizando la cantidad 102, el calculador de envolvente 210 dividirá el cuadro SBR en porciones cubiertas por una envolvente espectral 104 y para cada envolvente espectral 104 el calculador de envolvente 210 calcula los datos de envolvente data 205. Los datos de la envolvente data comprenden, por ejemplo, la envolvente espectral cuantizada y codificada, y estos datos son necesarios del lado del decodificador para generar la señal de banda alta y para aplicar filtrado inverso, agregando componentes de ruido y armónicos para replicar las características espectrales de la señal original.

[0032] La Fig. 3a muestra una forma de realización para un codificador 300, el codificador 300 comprende SBR módulos relacionados 310, un banco de análisis QMF 320, un aparato para sub-muestreo 330, un codificador de núcleo AAC 340 y una secuencia formateador de bits de carga útil 350. Además, el codificador 300 comprende el calculador de datos de la envolvente 210. El codificador 300 comprende una entrada para muestras PCM (señal de audio 105; PCM = modulación de código de pulso), conectada al banco de análisis QMF 320, y a los módulos relacionados con SBR 310 y con el aparato para sub-muestreo 330. El banco de análisis QMF 320, a su vez, se conecta con el calculador de datos de la envolvente 210, al cual, a su vez, se conecta con el formateador de secuencia de bits de carga útil 350. El aparato para sub-muestreo 330 se conecta con el codificador de núcleo AAC 340, el cual, a la vez, se conecta con el formateador de secuencia de bits de carga útil 350. Finalmente, el módulo relacionado con SBR 310 se conecta con el calculador de datos de la envolvente 210 y con el codificador de núcleo AAC 340.

[0033] En consecuencia, el codificador 300 realiza un sub-muestreo de la señal de audio 105 para generar componentes en la banda de frecuencia de núcleo (en el aparato para sub-muestreo 330), los que son ingresados en el codificador de núcleo AAC 340, el cual codifica la señal de audio en la banda de frecuencia de núcleo y envía la señal codificada al formateador de secuencia de bits de carga útil 350 en el cual la señal de audio codificada de la banda de frecuencia de núcleo es agregada a la secuencia de audio codificada 355. Por otro lado, la señal de audio 105 se analiza mediante el banco de análisis QMF 320, que extrae los componentes de frecuencia de la banda de frecuencia alta e ingresa las señales en el calculador de datos de la envolvente 210. Por ejemplo, un banco 64 de sub-banda QMF 320 realiza el filtrado de sub-banda de la señal de entrada. La salida desde el banco de filtro (es decir las muestras de sub-bandas) es valuada en forma compleja y por lo tanto, sometida a un sobre-muestreo por un factor de dos en comparación con un banco QMF regular.

[0034] Los módulos relacionados con SBR 310 controlan el calculador de datos de la envolvente 210 proporcionando, por ejemplo, la cantidad 102 de envolventes 104 al el calculador de datos de la envolvente 210. Utilizando la cantidad 102 y los componentes de audio generados por el banco de análisis QMF 320, el calculador de datos de la envolvente 210 calcula los datos de la envolvente 205 y envía los datos de la envolvente data 205 al formateador de secuencia de bits de carga útil 350, que combina los datos de la envolvente 205 con los componentes codificados por el codificador de núcleo 340 en la secuencia de audio codificada 355.

[0035] La Fig. 3a muestra por lo tanto parte de la herramienta SBR codificada estimando varios parámetros usados por el procedimiento de reconstrucción de y the frecuencia alta en el decodificador.

[0036] La Fig. 3b muestra un ejemplo para el módulo relacionado con SBR 310, que comprende el calculador de cantidad de envolvente 100 (ver Fig. 1) y opcionalmente otros módulos SBR 360. Los módulos relacionados con SBR-310 reciben la señal de audio 105 y emiten la cantidad 102 de envolventes 104, y además otros datos generados por otros módulos SBR 360.

[0037] Los otros módulos SBR 360 pueden, por ejemplo, comprender un detector de transitorios convencional adaptado para detectar transitorios en la señal de audio 105 y puede además obtener la cantidad y/o posiciones de las envolventes para que los módulos SBR puedan o no calcular parte de los parámetros utilizados por el procedimiento de reconstrucción de frecuencia alta en el decodificador (parámetro SBR).

Como se estableció previamente dentro de SBR una unidad de tiempo SBR (cuadro SBR) puede dividirse en varios bloques de datos, llamadas envolventes. Si esta división o partición es uniforme, es decir que todas las envolventes 104 tienen el mismo tamaño y la primera envolvente comienza y la última envolvente termina con un límite de cuadro, el cuadro SBR se define como cuadro FIXFIX.

[0038] La Fig. 4 ilustra dicha partición para un cuadro SBR en una cantidad 102 de envolventes espectrales 104. El cuadro SBR cubre un período de tiempo entre el tiempo inicial t_0 y el tiempo final t_n y, en las formas de realización se muestra en la Fig. 4, dividido en 8 porciones de tiempo, una primera porción de tiempo 111, una segunda porción de tiempo 112, . . ., una séptima porción de tiempo 117 y una octava porción de tiempo 118. Las 8 porciones de tiempo 110 están separadas por 7 bordes, que significa que un borde 1 se encuentra entre la primera y segunda porción de tiempo 111, 112, un borde 2 se encuentra entre la segunda porción 112 y la tercera porción 113, y así hasta un borde 7 se encuentra entre la séptima porción 117 y la octava porción 118.

[0039] En la Norma ISO/IEC 14496-3, la cantidad máxima de envolventes 104 en un cuadro FIXFIX se restringe a cuatro (ver sub-parte 4, párrafo 4.6.18.3.6). En general, la cantidad de envolventes 104 en el cuadro FIXFIX podría ser una potencia de dos (por ejemplo, 1, 2, 4), donde los cuadros FIXFIX sólo se utilizan si en el mismo cuadro no se han detectado transitorios. En implementaciones del codificador convencional AACX de alta eficiencia, por otro lado, la cantidad máxima de envolventes 104 se limita a dos, aún si la especificación de la norma permite en forma teórica hasta cuatro envolventes. Esta cantidad de envolventes 104 por cuadro puede aumentar, por ejemplo, a ocho (ver Fig. 4), para que un cuadro FIXFIX pueda comprender 1, 2, 4 u 8 envolventes (u otra potencia de 2). Por supuesto, otra cantidad 102 de envolventes 104 es también posible para que la cantidad máxima de envolventes 104 (cantidad predeterminada) pueda sólo restringirse por la resolución de tiempo del banco de filtro QMF que posee 32 franjas de tiempo QMF por cuadro SBR.

[0040] La cantidad 102 de envolventes 104 puede, por ejemplo, ser calculada de la siguiente manera. El calculador de valor de decisión 120 mide las desviaciones en la distribución de energía espectral de pares de porciones de tiempo vecinas 110. Por ejemplo, esto significa que un calculador de valor de decisión 120 calcula una primera distribución de energía espectral para la primera porción de tiempo 111, calcula una segunda distribución de energía espectral desde los datos espectrales dentro de la segunda porción de tiempo 112, y así sucesivamente. Luego, la primera distribución de energía espectral y la segunda distribución de energía espectral se comparan y de este valor de comparación 125 se deriva el valor de decisión, donde el valor de decisión 125 se relaciona, en este ejemplo, con el borde 1 entre la primera porción de tiempo 111 y la segunda porción de tiempo 112. El mismo procedimiento puede aplicarse a la segunda porción de tiempo 112 y la tercera porción de tiempo 113 para que estas dos porciones de tiempo vecinas además de las dos distribuciones de energía espectrales sean derivadas y estas dos distribuciones de energía espectrales son, a la vez, comparadas con el calculador de valor de decisión 120 para derivar otro valor de decisión 125.

[0041] Como próximo paso, el detector 130 comparará los valores de decisión derivados 125 con un valor umbral y si el valor umbral es violado, el detector 130 detectará una violación 135. Si el detector 130 detecta una violación 135, el procesador 140 determina un primer borde de envolvente 145. Por ejemplo, si el detector 130 detecta una violación en el borde 1 entre la primera porción de tiempo 111 y la segunda porción de tiempo 112, el primer borde envolvente 145a se encuentra en el tiempo del borde 1.

[0042] En la forma de realización de la Fig. 4, donde sólo varias posibilidades para gránulos/bordes son permitidas, esto significaría que todo el proceso ha terminado, y todos los bordes están determinados como lo indican las envolventes menores en 104a, 104b. En este caso los bordes estarían en todo los tiempos 0, 1, 2, ..., n.

[0043] Cuando, sin embargo, el primer borde debe determinarse por ejemplo en el instante de tiempo 4, se debe realizar la búsqueda del segundo borde. Como se indica en la Fig. 4, el segundo borde podría estar en 3, 2, 0. En caso que el borde se encuentre en 3, todo el procedimiento ha terminado, ya que las envolventes menores 104a, 104b están determinadas. En caso que el borde se encuentre en 2, la búsqueda debe continuar ya que no es seguro aún que las envolventes medias (indicadas por 145a) puedan ser utilizadas. Aún en caso que el borde se encuentre en 0, no se determina que en la segunda mitad, es decir, entre 4 y n, no haya un borde. Si no existe borde en la segunda mitad, las envolventes más amplias pueden ser determinadas. Si existe un borde por ejemplo en 5, las envolventes menores deben ser utilizadas. Si existe un borde sólo en 6, las envolventes medias deben ser utilizadas.

[0044] Cuando, sin embargo, se permite un patrón completamente flexible o más flexible para las envolventes, el procedimiento continua, cuando un primer borde en 1 ha sido determinado. Luego, el procesador 150 determina un segundo borde de envolvente 155, que se encuentra entre otro par de porciones de tiempo vecinas o coincide con el tiempo inicial t_0 o tiempo final t_n . En las formas de realización como se muestra en la Fig. 4, el segundo borde envolvente 155a coincide con el tiempo inicial t_0 (que produce una primera envolvente 104a) y otro segundo borde envolvente 155b coincide con el borde 2 entre la segunda porción de tiempo 112 y la tercera porción de tiempo 113 (que produce una segunda envolvente 104b). Si no se detecta violación en el borde 1 entre la primera porción de tiempo 111 y la segunda porción de tiempo 112, el detector 130 continuará investigando el borde 2 entre la segunda porción de tiempo 112 y la tercera porción de tiempo 113. Si no se detecta violación, otra envolvente 104c se extiende desde el tiempo de inicio t_0 al borde 2.

[0045] De acuerdo con las formas de realización de la invención, para un par de envolventes vecinas, dicho valor de decisión 125 mide la desviación de la distribución de energía espectral, donde cada distribución de energía espectral se refiere a una porción de la señal de audio dentro de la porción de tiempo. En el ejemplo de 8 envolventes, existe un total de 7 medidas (= 7 bordes entre porciones de tiempo vecinas) o, en general, si hay n envolventes, existen $n-1$ medidas (valores de decisión 125). Cada valor de decisión 125 puede ser comparada con un valor umbral y si el valor de decisión 125 (medida) viola el valor umbral, un borde de envolvente borde estará situado entre las dos envolventes vecinas. Dependiendo de la definición del valor de decisión 125 y del valor umbral, la violación puede representar que un valor de decisión 125 se encuentra sobre o debajo del valor umbral. En caso que el valor de decisión 125 se encuentra debajo el valor umbral, la distribución espectral no puede variar fuertemente de envolvente a envolvente. Por ende, no es necesario un borde de envolvente en esta posición (= momento en el tiempo).

[0046] En una forma de realización preferida, la cantidad 102 de envolventes 104 comprende una potencia de dos, y además cada envolvente comprende un período de tiempo igual. Esto implica que existen cuatro posibilidades: una primera posibilidad consiste en que todo el cuadro SBR está cubierto por una sola envolvente (no se muestra en la Fig. 4), la segunda posibilidad consiste en que el cuadro SBR está cubierto por 2 envolventes, la tercera posibilidad consiste en que el cuadro SBR está cubierto por 4 envolventes y la última posibilidad consiste en que el cuadro SBR está cubierto por 8 envolventes (mostrados en la Fig. 4 desde la parte superior a la inferior).

[0047] Puede resultar ventajoso investigar los bordes dentro de un orden específico, ya que si existe violación en un borde impar (borde 1, borde 3, borde 5, borde 7), la cantidad de envolventes será siempre ocho (bajo la suposición de envolventes de igual tamaño). Por otro lado, si existe violación en el borde 2 y borde 6, hay cuatro envolventes y, finalmente, si existe violación sólo en el borde 4, dos envolventes serán codificadas y no hay violación en ninguno de los 7 bordes, todo el cuadro SBR está cubierto por un solo envolvente. Por ende, el aparato 100 puede investigar primero el borde 1, 3, 5, 7 y si se detecta violación en estos bordes, el aparato 100 puede investigar el próximo cuadro SBR, ya que en este caso todo el cuadro SBR será codificado por la cantidad máxima de envolventes. Después de investigar estos bordes impares y no se detecta violación en los bordes impares, el detector 130 puede investigar, como próximo paso, el borde 2 y borde 6, para que si se detecta violación en uno de estos dos bordes, la cantidad de envolventes será cuatro y el aparato 100 podrá nuevamente cambiar al próximo cuadro SBR. Como último paso, si no se detectan violaciones hasta los bordes 1, 2, 3, 5, 6, 7, el detector 130 puede investigar el borde 4 y si se detecta una violación en el borde 4, la cantidad de envolventes se fija en dos.

[0048] Para el caso general (de porciones de tiempo n , donde n es un número par) este procedimiento puede también ser re-expresado de la siguiente forma. Si, por ejemplo, en los bordes impares no se detecta violación y por lo tanto el valor de decisión 125 puede ser menor al valor umbral implicando que las envolventes vecinas (separadas por esos bordes) comprenden diferencias que no son fuertes con respecto a la distribución de energía espectral, no hay necesidad de dividir el cuadro SBR en envolventes n y, en cambio, las envolventes $n/2$ podrán ser suficientes. Si además, el detector 130 no detecta violación en los bordes, los cuales son dos veces un número impar (por ejemplo en los bordes 2, 6, 10, ...), no hay necesidad de colocar un borde de envolvente en estas posiciones y, por ende, la cantidad de envolventes puede además reducirse por un factor de 2, es decir a $n/4$. Estos procedimientos continúan paso a paso (el próximo paso sería el borde, que es 4 veces un número impar, es decir 4, 12, ...). Si en todos estos bordes no se detecta violación una sola envolvente para todo el cuadro SBR será suficiente.

[0049] Sin embargo, si uno de los valores de decisión 125 en los bordes impares se encuentra por sobre el valor umbral, las envolventes n deberían ser consideradas, ya que sólo después un borde de envolvente estará en la posición correspondiente (ya que se supone que todas las envolventes tienen la misma longitud). En este caso, las envolventes n serán calculadas aún si los demás valores de decisión 125 se encuentran por debajo del valor umbral.

[0050] El detector 130 puede, sin embargo, considerar también todos los bordes y considerar todos los valores de decisión 125 para todas las porciones de tiempo 110 para calcular la cantidad de envolventes 104.

[0051] Como el aumento en la cantidad de envolventes 102 también implica un incremento en la cantidad de datos a transmitir, el valor umbral de decisión para el borde de envolvente correspondiente, el cual conlleva una alta cantidad de envolventes 104 puede también aumentar. Esto significa que el valor umbral en el borde 1, 3, 5 y 7 puede opcionalmente ser mayor al valor umbral en los bordes 2 y 6, que a la vez pueden ser mayores al valor umbral en el borde 4. Los valores umbrales mayores o menores se refieren al caso en que la violación del valor umbral sea más o menos posible. Por ejemplo un mayor valor umbral implica que la desviación en la distribución de energía espectral

entre dos porciones de tiempo vecinas puede ser más tolerable que con un valor umbral menor y por lo tanto para un mayor valor umbral se necesitan desviaciones más severas en la distribución de energía espectral para demandar otras envolventes.

5 **[0052]** El valor umbral escogido puede además depender de la señal en cuanto a si la señal se clasifica en señal de voz o señal de audio general. Sin embargo, no es el caso que el valor umbral de decisión siempre sea reducido (o aumentado) si la señal se clasifica como de voz. Dependiendo de la aplicación, sin embargo, podrá ser ventajoso si para una señal de audio general, el valor umbral es alto para que en este caso la cantidad de envolventes sea generalmente menor a la señal de voz.

10 **[0053]** La Fig. 5 ilustra otras formas de realización donde la longitud de las envolventes varía sobre los cuadros SBR. En la Fig. 5a, se muestra un ejemplo con tres envolventes 104, una primera envolvente 104a, una segunda envolvente 104b y una tercera envolvente 104c. La primera envolvente 104a se extiende desde el tiempo inicial t_0 al borde 2 en el tiempo t_2 , la segunda envolvente 104b se extiende desde el borde 2 en el tiempo t_2 al borde 5 en el tiempo t_5 y la tercera envolvente 104c se extiende del borde 5 en el tiempo t_5 al tiempo final t_n . Si todas las porciones de tiempo se encuentran nuevamente con la misma longitud y si el cuadro SBR es, nuevamente dividido en ocho porciones de tiempo, la primera envolvente 104a cubre la primera y segunda porción de tiempo 111, 112, la segunda envolvente 104b cubre la tercera, la cuarta y la quinta porción de tiempo 113 a 115 y la tercera envolvente 104c cubre la sexta, la séptima y la octava porción de tiempo. Por lo tanto, la primera envolvente 104a es menor a la segunda y la tercera envolvente 104b y 104c.

20 **[0054]** La Fig. 5b muestra otra forma de realización con sólo dos envolventes, una primera envolvente 104a se extiende desde el tiempo inicial t_0 al primer tiempo t_1 y una segunda envolvente 104b se extiende desde el primer tiempo t_1 al tiempo final. Por ende, la segunda envolvente 104b se extiende por 7 porciones de tiempo, mientras que la primera envolvente 104a se extiende sólo por una sola porción de tiempo (la primera porción de tiempo 111).

25 **[0055]** La Fig. 5c muestra, nuevamente, una forma de realización con tres envolventes 104, donde la primera envolvente 104a se extiende desde el tiempo inicial t_0 al segundo tiempo t_2 , la segunda envolvente 104b se extiende del segundo tiempo t_2 al cuarto tiempo t_4 y la tercera envolvente 104c se extiende del cuarto tiempo t_4 al tiempo final t_n .

30 **[0056]** Estas formas de realización pueden, por ejemplo, ser utilizadas en caso que los bordes de envolventes 104 se coloquen sólo entre porciones de tiempo vecinas donde se detecte una violación del valor umbral en el tiempo inicial y final t_0 , t_n . Esto significa que en la Fig. 5a, se detecta una violación en el tiempo t_2 y se detecta una violación en el tiempo t_5 , mientras que no se detectan violaciones en los momentos de tiempo restantes t_1 , t_3 , t_4 , t_6 y t_7 . De manera similar, en la Fig. 5b, sólo se detecta una violación en el tiempo t_1 , resultando en un borde para la primera envolvente 104a y para la segunda envolvente 104b y en la Fig. 5c, se detecta una violación sólo en el segundo tiempo t_2 y el cuarto tiempo t_4 .

35 **[0057]** Para que un decodificador pueda utilizar los datos de la envolvente y replicar e consecuencia la banda espectral mayor, el decodificador necesita la posición de las envolventes 104 y de los bordes de envolvente correspondientes. En las formas de realización previas, basadas en dicha norma, donde todas las envolventes 104 comprenden la misma longitud y por lo tanto, fue suficiente transmitir la cantidad de envolventes para que el decodificador pueda decidir donde debe colocarse un borde de envolvente. En estas formas de realización como muestra la Fig. 5 sin embargo, el decodificador necesita información en cuyo tiempo se coloca un borde de envolvente y por lo tanto la información lateral adicional puede colocarse en la secuencia de datos para que al utilizar la información lateral, el decodificador puede retener momentos de tiempo donde se coloca un borde y una envolvente comienza y termina. Esta información adicional comprende el tiempo t_2 y t_5 (en el caso de la Fig. 5a), el tiempo t_1 (en el caso de la Fig. 5b) y el tiempo t_2 y t_4 (en caso de la Fig. 5c).

45 **[0058]** Las Figs. 6a y 6b muestran una forma de realización para el calculador de valor de decisión 120 utilizando la distribución de energía espectral en la señal de audio 105.

50 **[0059]** La Fig. 6a muestra un primer grupo de valor de muestra 610 para la señal de audio en una porción de tiempo dada, por ejemplo, la primera porción de tiempo 111 y compara esta señal de audio tomada por muestreo con un segundo grupo de muestras de señal de audio 620 en la segunda porción de tiempo 112. La señal de audio se transformó en dominio de frecuencia para que los grupos de valores de muestras 610, 620 o sus niveles P se muestren como función de la frecuencia f . Las bandas de frecuencia mayores y menores son separadas por la frecuencia de cruce f_0 implicando que para frecuencias mayores a f_0 los valores de muestra no serán transmitidos. El decodificador debería en cambio replicar estos valores de muestra utilizando los datos SBR. Por otro lado, las muestras debajo de la frecuencia de cruce f_0 son codificados, por ejemplo, por el codificador AAC y transmitidos al decodificador.

55 **[0060]** El decodificador puede utilizar valores de muestra desde la banda de frecuencia baja con el fin de replicar los componentes de frecuencia alta. Por lo tanto, con el fin de medir la desviación del primer grupo de muestras 610 en la primera porción de tiempo 111 y el segundo grupo de muestras 620 en la segunda porción de tiempo 112, no podrá ser suficiente considerar sólo los valores de muestra en la banda de frecuencia alta (para $f > f_0$), sino también tener en cuenta los componentes de frecuencia en la banda de frecuencia baja. En general, se espera una buena replicación de

calidad si hay una correlación entre los componentes de frecuencia en la banda de frecuencia alta con respecto a los componentes de frecuencia en la banda de frecuencia baja. En un primer paso, puede ser suficiente considerar sólo valores de muestra en la banda de frecuencia alta (sobre la frecuencia de cruce f_0) y calcular una correlación ente el primer grupo de valores de muestra 610 con un segundo grupo de valores de muestra 620.

5 **[0061]** La correlación puede calcularse utilizando procedimientos estadísticos estándar y puede comprender, por ejemplo, el cálculo de la llamada función de correlación cruzada u otras mediciones estadísticas para la similitud de las dos señales. Existe también el coeficiente de correlación de momento del producto de Pearson, que puede utilizarse para estimar una correlación de dos señales. Los coeficientes de Pearson se conocen también como coeficientes de correlación de muestra. En general, una correlación indica la fuerza y dirección de una relación lineal entre dos variables aleatorias –en este caso, las dos distribuciones de muestras 610 y 620. Por lo tanto, la correlación se refiere a la salida de dos variables aleatorias desde la independencia. En este sentido amplio, existen varios coeficientes que miden el grado de correlación adaptado la naturaleza de los datos para que diferentes coeficientes sean usados para diferentes situaciones.

10 **[0062]** La Fig. 6b muestra un tercer grupo de valores de muestra 630 y un cuarto grupo de valores de muestra 640, que, por ejemplo, pueden esta relacionados con los valores de muestra en la tercera porción de tiempo 113 y la cuarta porción de tiempo 114. Nuevamente, con el fin de comparar los dos grupos de muestras (o señales), dos porciones de tiempo vecinas son consideradas. En contraposición con el caso de la Fig. 6a, en la Fig. 6b un valor umbral T es introducido para que sólo se consideren los valores de muestra cuyos niveles P se encuentren por encima (o más en general violan) el valor umbral T (para los cuales $P > T$ permanece).

15 **[0063]** En esta forma de realización la desviación en la distribución de energía espectral puede medirse simplemente contando la cantidad de valores de muestra que violan este valor umbral T y el resultado puede fijar el valor de decisión 125. Este simple procedimiento producirá una correlación entre ambas señales sin realizar un detallado análisis estadístico de varios grupos de valores de muestra en las distintas porciones de tiempo 110. En forma alternativa, un análisis estadístico, por ejemplo como se mencionó anteriormente, puede aplicarse a las muestras que violan el valor umbral T solamente.

20 **[0064]** Las Figs. 7a a 7c muestran otra forma de realización donde el codificador 300 comprende una unidad de decisión de conmutación 370 y una unidad codificadora estéreo 380. Además, el codificador 300 también comprende la herramienta de ampliación de ancho de banda como, por ejemplo, el calculador de datos de la envolvente 210 y los módulos relacionados con SBR 310. La unidad de decisión de conmutación 370 provee una señal de decisión de conmutación 371 que conmuta entre un codificador de audio 372 y un codificador de voz 373. Cada uno de estos códigos puede codificar la señal de audio en la banda de frecuencia del núcleo utilizando diferentes cantidades de valores de muestra (por ejemplo 1024 para una resolución mayor o 256 para una resolución menor). La señal de decisión de conmutación 371 es también suministrada a la herramienta de ampliación de ancho de banda (BWE, por su sigla en inglés) herramienta 210, 310. La herramienta BWE 210, 310 usará la decisión de conmutación 371 con el fin de, por ejemplo, ajustar los valores umbrales para determinar la cantidad 102 de envolventes espectrales 104 y para encender/apagar un detector de transitorios opcional. La señal de audio 105 es ingresada en la unidad de decisión de conmutación 370 e ingresada en una codificación estéreo 380 para que la codificación estéreo 380 pueda producir los valores de muestra, ingresados en la unidad de ampliación de ancho de banda 210, 310. Dependiendo de la decisión 371 generada por la unidad de decisión de unidad de conmutación 370, la herramienta de ampliación de ancho de banda 210, 310 generará los daos de replicación de banda espectral, que a la vez son enviados a un codificador de audio 372 o codificador de voz 373.

25 **[0065]** La señal de decisión de conmutación 371 depende de la señal y puede obtenerse mediante la unidad de decisión de conmutación 370 analizando la señal de audio, por ejemplo utilizando un detector de transitorios u otros detectores, que pueden o no comprender un valor umbral variable. De manera alternativa, la señal de decisión de conmutación 371 puede también ajustarse en forma manual o ser obtenida de una secuencia de datos (incluidos en la señal de audio).

30 **[0066]** La salida del codificador de audio 372 y del codificador de voz 373 pueden ser ingresadas nuevamente en el formateador de secuencia de bits 350 (ver Fig. 3a).

35 **[0067]** La Fig. 7b muestra un ejemplo para la señal de decisión de conmutación 371, que detectas una señal de audio para un período de tiempo debajo de un primer tiempo t_a y por encima de un segundo tiempo t_b . Entre el primer tiempo t_a y el segundo tiempo t_b , la unidad de decisión de conmutación 370 detecta una señal de voz que implica diferentes valores discretos para la señal de decisión de conmutación 371.

40 **[0068]** En consecuencia, como lo muestra la Fig. 7c, durante el tiempo, la señal de audio es detectada, lo que significa que para tiempos antes de t_a , la resolución temporal de la codificación es baja, mientras que durante el período donde una señal de voz es detectada (entre el primer tiempo t_a y el segundo tiempo t_b), la resolución temporal es incrementada. Un aumento en la resolución temporal implica una ventana de análisis menor en el dominio de tiempo. La resolución temporal aumentada implica además la cantidad aumentada mencionada anteriormente de envolventes espectrales (ver descripción en Fig. 4).

- 5 [0069] Para señales de voz que necesitan una representación temporal exacta de las frecuencias altas, el valor umbral de decisión (por ejemplo usado en la Fig. 4) para transmitir una cantidad mayor de grupos de parámetros es controlado por la unidad de decisión de conmutación 370. Para señales de voz y del tipo voz, codificadas con la parte de codificación de voz o dominio de tiempo 373 del codificador de núcleo conmutado, el valor umbral de decisión que utiliza más grupos de parámetros puede, por ejemplo, ser reducido y, por lo tanto la resolución temporal es incrementada. Sin embargo, este no siempre es el caso mencionado previamente. La adaptación de la resolución de tipo temporal es independiente de la estructura subyacente del codificador (la cual no fue utilizada en la Fig. 4). Esto significa que el procedimiento descrito se utiliza también dentro de un sistema en el cual el módulo SBR comprende sólo un codificador de núcleo.
- 10 [0070] Aunque algunos aspectos han sido descritos en el contexto de un aparato, queda claro que estos aspectos también representan una descripción del procedimiento correspondiente, donde un bloque o dispositivo corresponde a un paso del procedimiento o rasgo de un paso del procedimiento. En forma análoga, los aspectos descritos en el contexto de un paso del procedimiento también representan una descripción de un bloque o ítem correspondiente o rasgo de un aparato correspondiente.
- 15 [0071] La señal de audio codificada de la invención puede almacenarse en un medio de almacenamiento digital o transmitida en un medio de transmisión como un medio de transmisión inalámbrico o por cable como la Internet
- 20 [0072] Dependiendo de ciertos requisitos de implementación, las formas de realización de la invención pueden implementarse en hardware o software. La implementación puede realizarse utilizando un medio de almacenamiento digital por ejemplo un disquete, DVD, CD, memoria ROM, PROM, EPROM, EEPROM o FLASH, con señales de control capaces de ser leídas en forma electrónica almacenados en ellos, que cooperan (o son capaces de cooperar) con un sistema de computación programable tal que se lleve a cabo el procedimiento respectivo
- [0073] Algunas formas de realización de acuerdo con la invención comprenden un portador de datos con señales de control capaces de ser leídas en forma electrónica, capaces de cooperar con un sistema de computación programable tal que se lleve a cabo uno de los procedimientos descritos.
- 25 [0074] Generalmente, las formas de realización de la presente invención pueden implementarse como un producto de programa de computación con un código de programa, el código de programa es operativo para desarrollar uno de los procedimientos cuando el producto de programa de computación es utilizado en una computadora. El código de programa puede por ejemplo ser almacenado en un portador susceptible de ser leído por una máquina.
- 30 [0075] Otras formas de realización comprenden el programa de computación para desarrollar uno de los procedimientos descritos en la presente, almacenado en un portador susceptible de ser leído por una máquina.
- [0076] En otras palabras, una forma de realización del procedimiento de invención consiste, por lo tanto, en un programa de computación con un código de programa para desarrollar uno de los procedimientos descritos en la presente, cuando el programa de computación es utilizado en una computadora.
- 35 [0077] Otra forma de realización del procedimiento de invención consiste, por lo tanto, un portador de datos (o medio almacenador digital, o medio capaz de ser leído por una computadora) que comprende, el programa de computación grabado en los mismos, para desarrollar uno de los procedimientos descritos en la presente.
- [0078] Otra forma de realización del procedimiento de invención consiste, por lo tanto, en una secuencia de datos o secuencia de señales que representan el programa de computación para desarrollar uno de los procedimientos descritos en la presente. La secuencia de datos o secuencia de señales puede por ejemplo ser configurada para ser transferida a través de una conexión de comunicación de datos, por ejemplo por Internet.
- 40 [0079] Otra forma de realización comprende un medio de procesamiento, por ejemplo una computadora, o un dispositivo lógico programable, configurado o adaptado para desarrollar uno de los procedimientos descritos en la presente.
- 45 [0080] Otra forma de realización comprende una computadora con programa de computación instalado en la misma, para desarrollar uno de los procedimientos descritos en la presente.
- [0081] En algunas formas de realización, un dispositivo lógico programable (por ejemplo una matriz de compuertas programables en obra) puede utilizarse para desarrollar una de las funcionalidades de los procedimientos descritos en la presente. En algunas formas de realización, una matriz de compuertas programables en obra puede cooperar con un microprocesador con el fin de desarrollar uno de los procedimientos descritos en la presente. Generalmente, los procedimientos son preferentemente desarrollados por cualquier aparato de hardware.
- 50 [0082] Las formas de realización descritas anteriormente son meramente ilustrativas para los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en la presente resultarán evidentes para los expertos en la técnica. Por lo tanto, se intenta limitar sólo por el alcance de las reivindicaciones de la patente inminentes y no por los detalle específicos presentados a modo de descripción y explicación de las formas de realización de la presente.
- 55

REIVINDICACIONES

1. Un aparato (100) para calcular una cantidad (102) de envolventes espectrales (104) derivadas por el codificador de replicación de banda espectral (SBR), donde el codificador SBR está adaptado para codificar una señal de audio (105) utilizando una pluralidad de valores de muestra dentro de una cantidad predeterminada de porciones de tiempo subsecuentes (110) en un cuadro SBR que se extiende desde un tiempo inicial (t_0) a un tiempo final (t_n), la cantidad predeterminada de porciones de tiempo subsecuentes (110) se disponen en una secuencia de tiempo dada por la señal de audio (105), el aparato (100) comprende:
- Un calculador de valor de decisión (120) para determinar un valor de decisión (125), el valor de decisión (125) mide una desviación en la distribución de energía espectral de un par de porciones de tiempo vecinas;
- Un detector (130) para detectar una violación (135) de un valor umbral por medio del valor de decisión (125);
- Un procesador (140) para determinar un primer borde de envolvente (145) entre el par de porciones de tiempo vecinas cuando se detecta violación (135) del valor umbral;
- Un procesador (150) para determinar un segundo borde de envolvente (155) entre un par de porciones de tiempo vecinas diferente o en el tiempo inicial (t_0) o en el tiempo final (t_n) para una envolvente con el primer borde de envolvente (145) basado en la violación (135) del valor umbral para el otro par o basado en una posición temporal del par o del par diferente en el cuadro SBR; y
- Un procesador de cantidad (160) para establecer la cantidad (102) de envolventes espectrales (104) con el primer borde de envolvente (145) y el segundo borde de envolvente (155),
- donde la cantidad predeterminada de porciones de tiempo (110) es igual a n con bordes $n-1$ entre porciones de tiempo vecinas (110), numeradas y ordenadas con respecto al tiempo para que los bordes comprendan bordes pares e impares, y donde el procesador de cantidad (160) está adaptado para establecer n como la cantidad (102) de envolvente espectral (104) si el detector (130) detecta la violación (135) en un borde impar, o
- donde el detector (150) está adaptado para determinar el segundo borde (155) de modo tal que las envolventes espectrales (104) comprendan una misma longitud temporal y la cantidad (102) de envolventes espectrales (104) sea una potencia de dos, o
- en el que el aparato (100) comprende además una unidad de decisión de conmutación (370) configurada para proveer una señal de decisión de conmutación (371), la señal de decisión de conmutación (371) señala una señal de audio del tipo voz y una señal de audio general del tipo audio, donde el detector (130) está adaptado para disminuir el valor umbral para señales de audio del tipo voz.
2. El aparato (100) de la reivindicación 1, donde una longitud en el tiempo de una porción de tiempo de la cantidad predeterminada de porciones de tiempo subsecuentes (110) es igual a la longitud mínima en el tiempo, para la cual se determina una sola envolvente, y en la cual el calculador de valor de decisión (120) está adaptado para calcular un valor de decisión (125) para dos porciones de tiempo vecinas con la longitud mínima en el tiempo.
3. El aparato (100) de la reivindicación 1 o reivindicación 2, donde el procesador (140) está adaptado para fijar el primer borde (145) en la primera violación detectada (135), y donde el procesador (150) está adaptado para fijar el segundo borde envolvente (155) después de comparar al menos otro valor de decisión (125) con el valor umbral.
4. El aparato (100) de la reivindicación 3, que además comprende un procesador de información para proveer información lateral adicional, la información lateral adicional comprende el primer borde envolvente (145) y el segundo borde envolvente (155) dentro de la secuencia de tiempo de la señal de audio (105).
5. El aparato (100) de una de las reivindicaciones precedentes, donde el detector (130) está adaptado para investigar en un orden temporal cada borde entre las porciones de tiempo vecinas (110).
6. El aparato (100) de la reivindicación 1, donde el detector (130) está adaptado para detectar primero la violación (135) en los bordes impares.
7. El aparato (100) de la reivindicación 1, donde la cantidad predeterminada es igual a 8, y donde el procesador de cantidad (160) está adaptado para establecer la cantidad (102) de las envolventes espectrales (104) a 1, 2, 4 u 8 de modo tal que cada envolvente espectral (104) comprende una misma longitud temporal.
8. El aparato (100) de la reivindicación 1 o la 7, donde el detector (130) está adaptado para utilizar un valor umbral, que depende de una posición temporal de la violación (135) de modo tal que en una posición temporal que produce una gran cantidad de envolventes espectrales (104) se utiliza un valor umbral mayor que el utilizado para una posición temporal que produce una cantidad menor de envolventes espectrales (104).
9. El aparato (100) de una de las reivindicaciones precedentes, que además comprende un detector de transitorios con un valor umbral de transitorio, el valor umbral de transitorio es mayor al valor umbral y/o además

comprende un calculador de datos de envolvente (210), el calculador de datos de envolvente (210) está adaptado para calcular datos de la envolvente espectral para una envolvente espectral (104) que se extiende desde el primer borde de envolvente (145) al segundo borde de envolvente (155).

10. Un codificador (300) para codificar una señal de audio (105) que comprende:

5 Un codificador de núcleo (340) para codificar la señal de audio (105) dentro de una banda de frecuencia de núcleo;

Un aparato (100) para calcular una cantidad (102) de envolventes espectrales (104) de acuerdo con una de las reivindicaciones 1 a 9; y

Un calculador de datos de envolvente (210) para calcular datos de la envolvente basados en la señal de audio (105) y la cantidad (102).

10 11. Un procedimiento para calcular una cantidad (102) de envolventes espectrales (104) a derivar por el codificador (SBR) de replicación de banda espectral donde el codificador SBR está adaptado para codificar una señal de audio (105) utilizando una pluralidad de valores de muestra dentro de una cantidad predeterminada de porciones de tiempo subsecuentes (110) en un cuadro SBR que se extiende desde un tiempo inicial (t_0) a un tiempo final (t_n), la cantidad predeterminada de porciones de tiempo subsecuentes (110) se dispone en una secuencia de tiempo dada por la señal de audio (105), comprendiendo el procedimiento:

1.5 La determinación de un valor de decisión (125), el valor de decisión (125) mide una desviación en la distribución de energía espectral de un par de porciones de tiempo vecinas;

La detección de una violación (135) de un valor umbral por medio del valor de decisión (125);

2.0 La determinación de un primer borde de envolvente (145) entre el par de porciones de tiempo vecinas cuando se detecta la violación (135) del valor umbral;

La determinación de un segundo borde de envolvente (155) entre el par diferente de porciones de tiempo vecinas o en el tiempo inicial (t_0) o en el tiempo final (t_n) para una envolvente con el primer borde de envolvente (145) basado en la violación (135) del valor umbral para el otro par o basado en una posición temporal del par o del par diferente en el cuadro SBR; y

2.5 El establecimiento de cantidad (102) de envolventes espectrales (104) que posee el primer borde de envolvente (145) y el segundo borde de envolvente (155),

3.0 donde la cantidad predeterminada de porciones de tiempo (110) es igual a n con bordes $n-1$ entre porciones de tiempo vecinas (110), numeradas y ordenadas con respecto al tiempo para que los bordes comprendan bordes pares e impares, y donde n se establece como la cantidad (102) de envolventes espectrales (104) si el detector (130) detecta la violación (135) en un borde impar, o

donde el segundo borde (155) se determina de modo tal que las envolventes espectrales (104) comprendan una misma longitud temporal y la cantidad (102) de envolventes espectrales (104) sea una potencia de dos, o

3.5 que comprende además una etapa de proporcionar una señal de decisión de conmutación (371), la señal de decisión de conmutación (371) señala una señal de audio del tipo voz y una señal de audio general del tipo audio, donde se disminuye el valor umbral para señales de audio del tipo voz.

12. Un programa de computación para desarrollar, cuando es utilizado en un procesador, un procedimiento de acuerdo con la reivindicación 11.

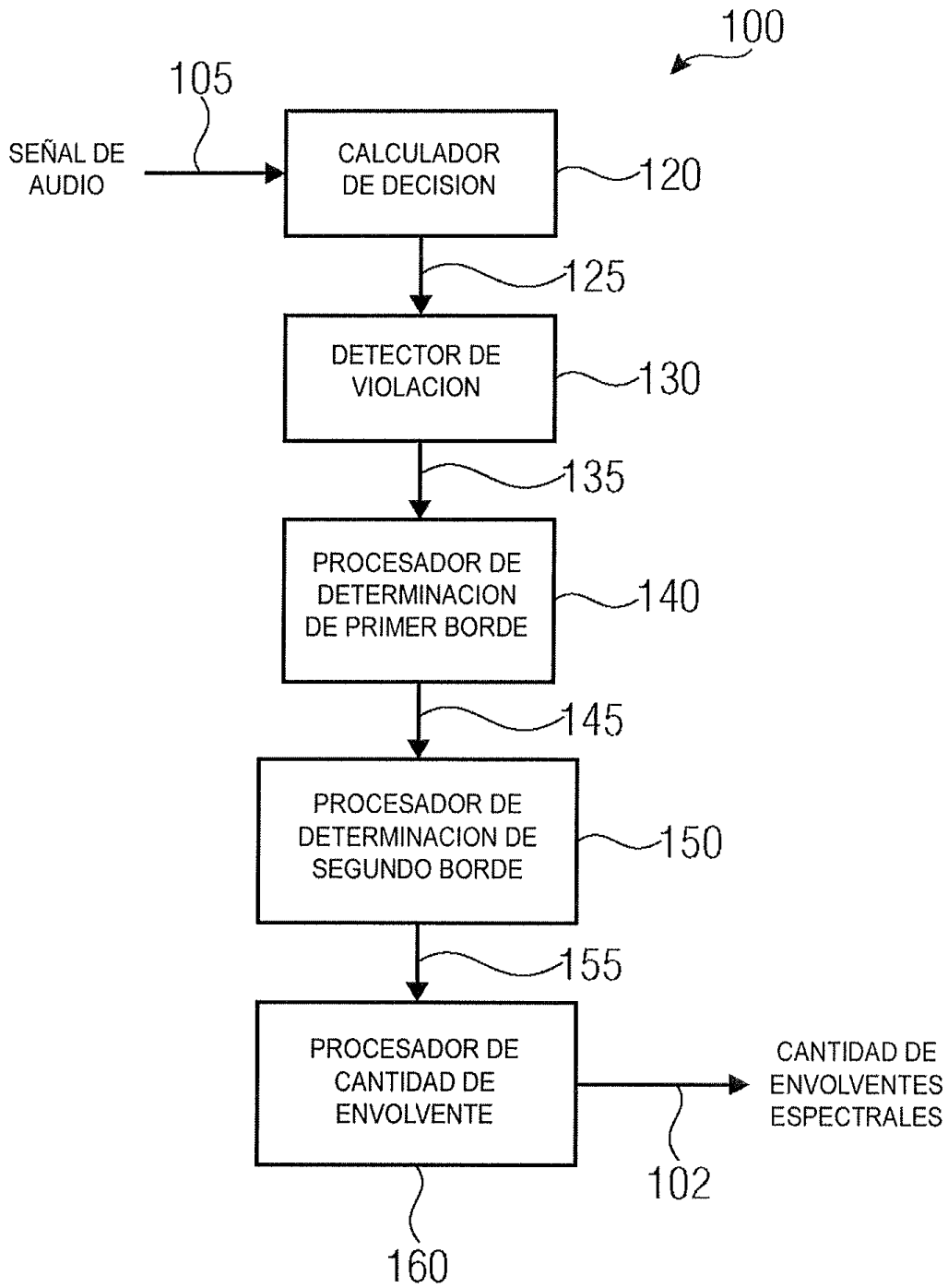


FIG 1

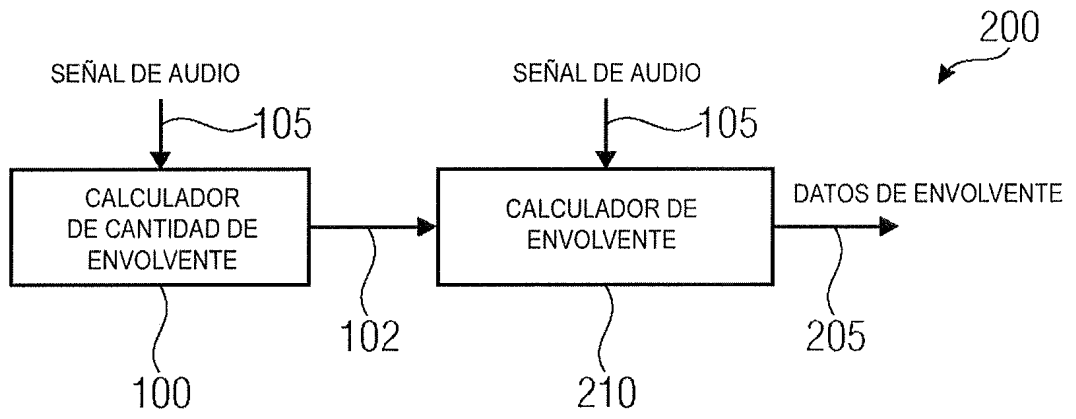


FIG 2

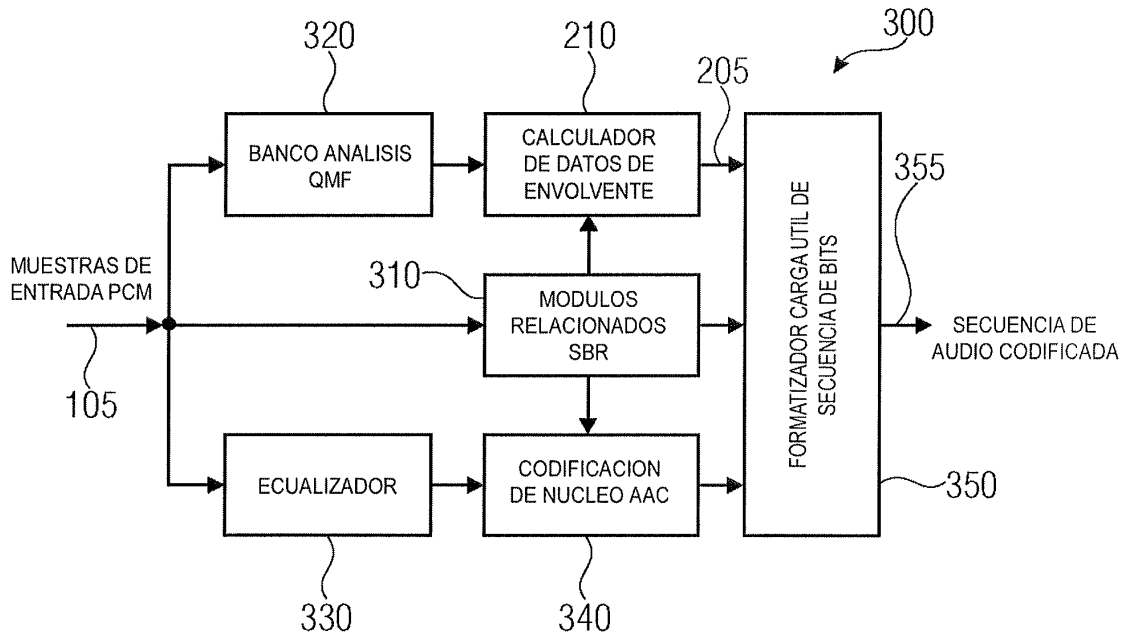


FIG 3A

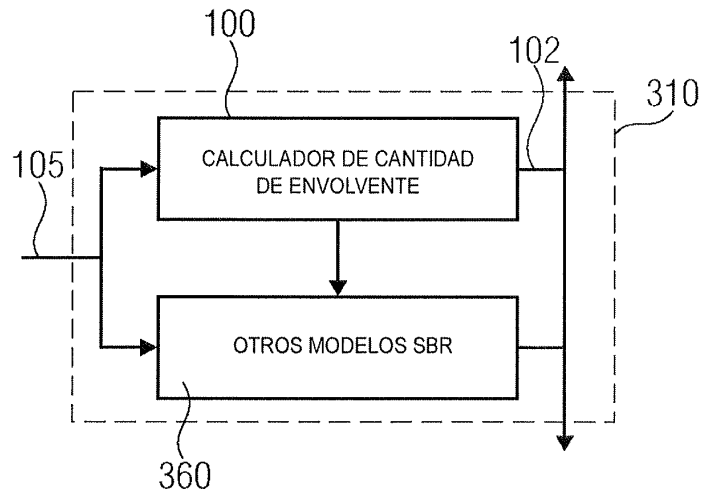


FIG 3B

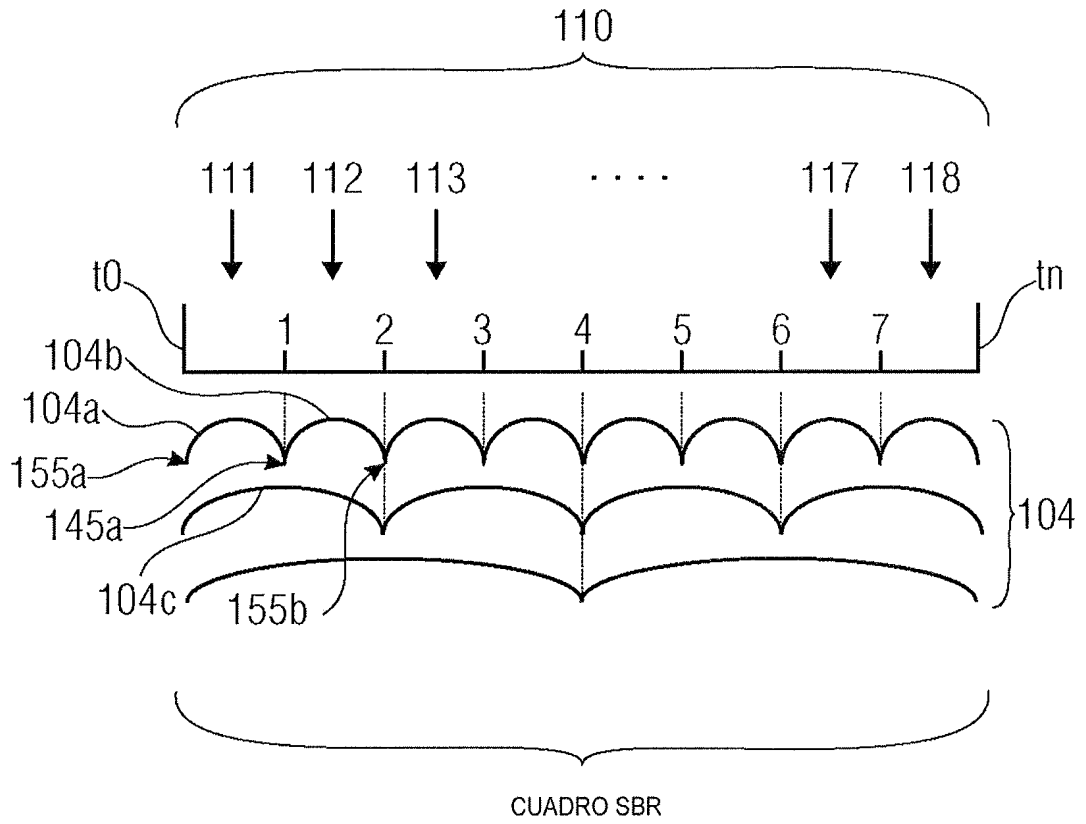


FIG 4

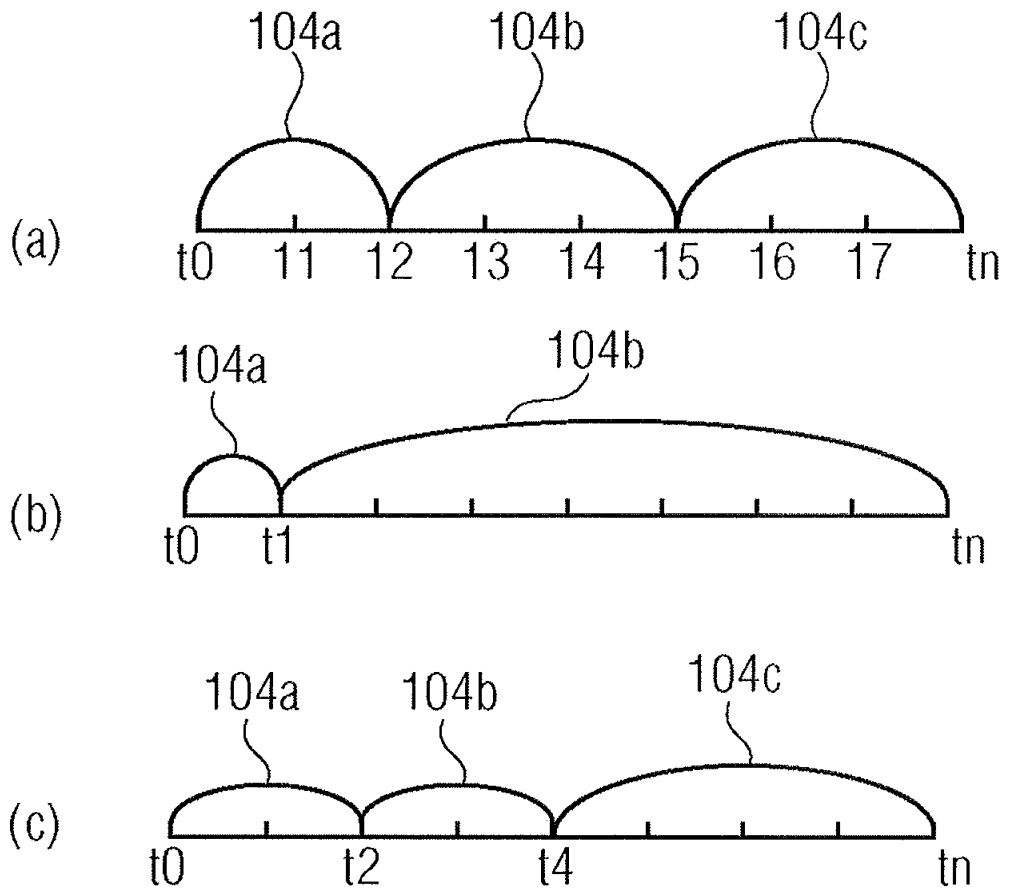


FIG 5

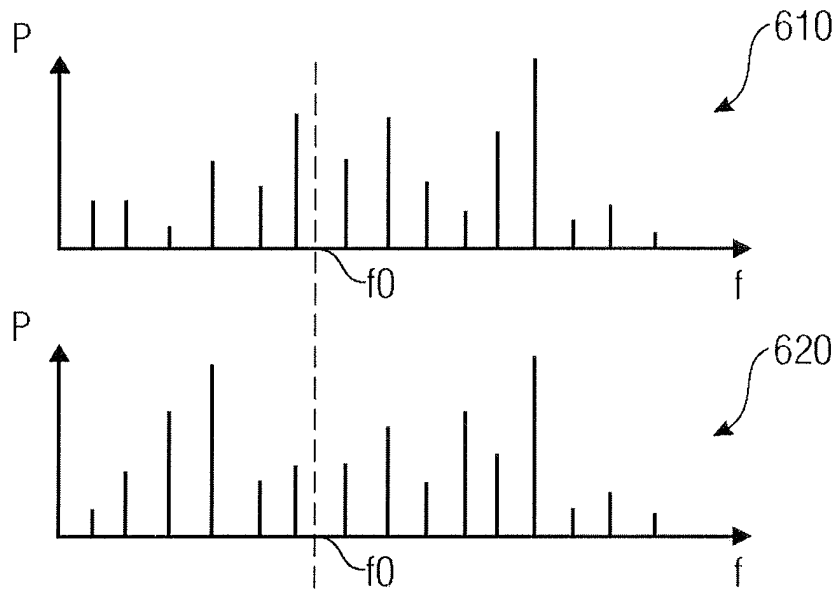


FIG 6A

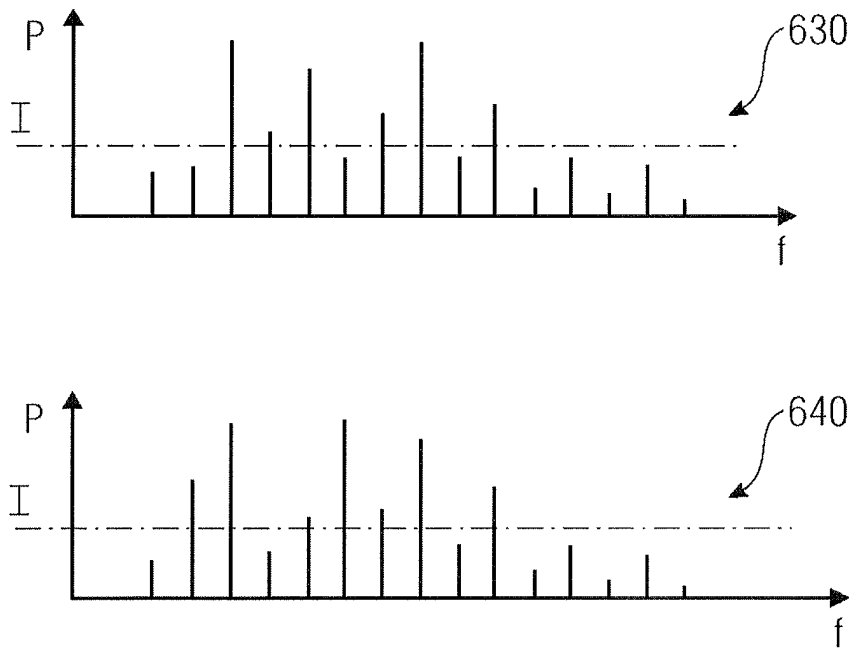


FIG 6B

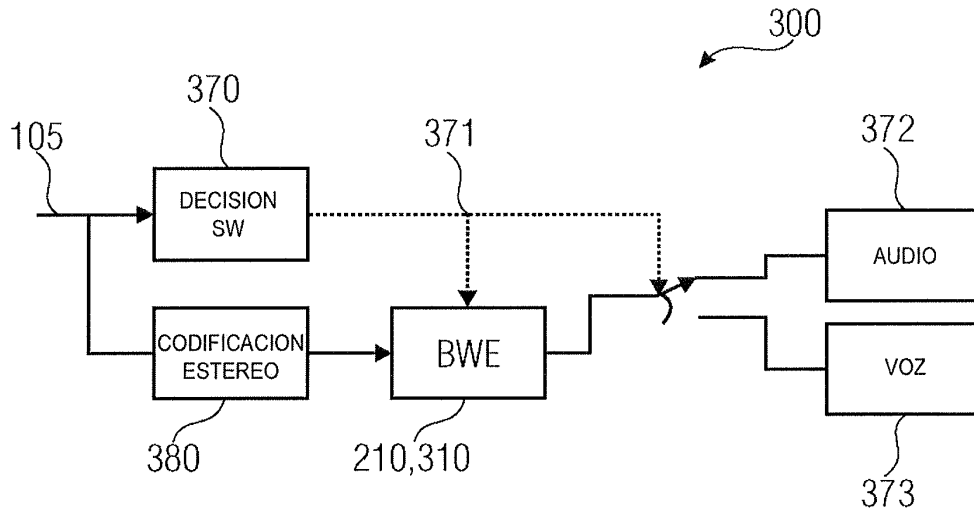


FIG 7A

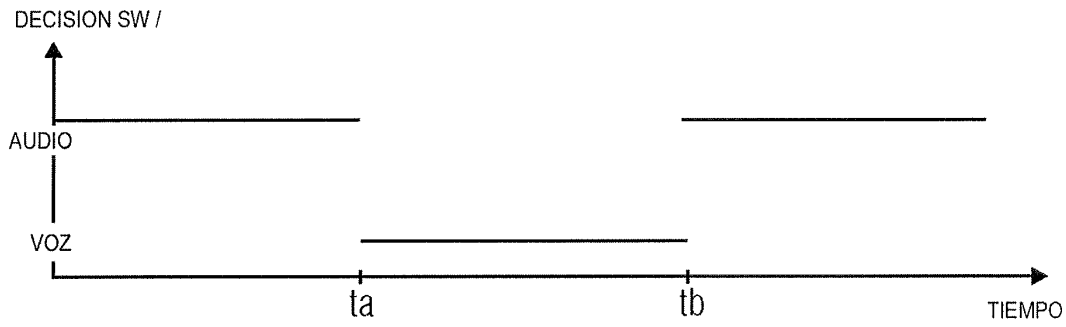


FIG 7B

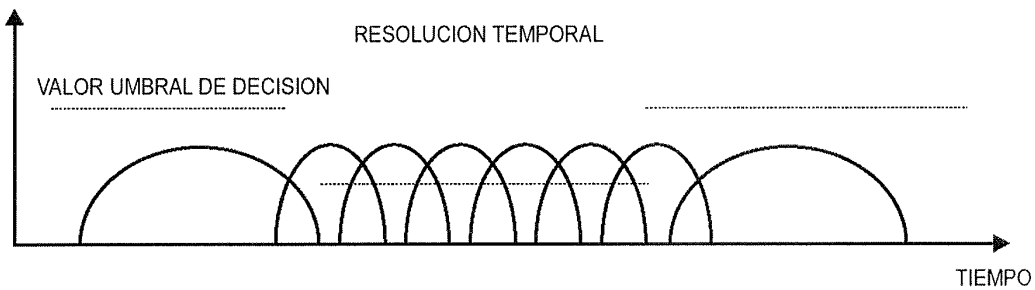


FIG 7C