

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 749 575**

51 Int. Cl.:

H04S 3/00 (2006.01)

H04S 5/00 (2006.01)

G10L 19/00 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **30.04.2004 E 17173334 (8)**

97 Fecha y número de publicación de la concesión europea: **28.08.2019 EP 3244638**

54 Título: **Procesamiento avanzado basado en un banco de filtros complejo, exponencial y modulado**

30 Prioridad:

30.04.2003 SE 0301273

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

23.03.2020

73 Titular/es:

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

72 Inventor/es:

**ENGDEGARD, JONAS y
VILLEMOES, LARS**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 749 575 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

Procesamiento avanzado basado en un banco de filtros complejo, exponencial y modulado

Campo técnico

5 La presente invención hace referencia a sistemas de codificación de fuente de audio, pero los mismos métodos podrían ser aplicados, asimismo, en muchos otros sectores técnicos. Se presentan diferentes técnicas que son útiles para los sistemas de codificación de audio que utilizan representaciones paramétricas de propiedades estéreo.

Antecedentes de la invención y técnica anterior

10 La presente invención hace referencia a la codificación paramétrica de la imagen estéreo de una señal de audio. Parámetros típicos utilizados para describir las propiedades de la imagen estéreo son la diferencia de intensidad entre canales (IID – Inter-channel Intensity Difference, en inglés), diferencia de tiempo entre canales (ITD - Inter-channel Time Difference, en inglés) y coherencia entre canales (IC - Inter-channel Coherence, en inglés). Con el fin de reconstruir la imagen estéreo en base a estos parámetros, se requiere un método que pueda reconstruir el nivel correcto de correlación entre los dos canales, de acuerdo con el parámetro IC. Esto se consigue mediante un método de descorrelación.

15 Existen un par de métodos disponibles para la creación de señales descorrelacionadas. De manera ideal, se desea una función lineal e invariante en el tiempo (LTI – Linear Time Invariant, en inglés) con respuesta de frecuencia pasa todo. Un método obvio para conseguir esto es mediante la utilización de un retardo constante. No obstante, mediante la utilización de un retardo, o de cualquier otra función pasa todo de LTI, resultará en una respuesta no pasa todo después de agregar la señal no procesada. En el caso de un retardo, el resultado será un filtro de peine típico. El filtro de peine a menudo produce un sonido “metálico” no deseable que, incluso si el efecto de ensanchamiento del estéreo puede ser eficiente, reduce mucho la naturalidad del original.

20 Los métodos en el dominio de la frecuencia para generar una señal descorrelacionada mediante la agregación de una secuencia aleatoria a los valores de IID a lo largo del eje de la frecuencia, donde se utilizan diferentes secuencias para los diferentes canales de audio, también son conocidos a partir de la técnica anterior. Un problema con la descorrelación en el dominio de frecuencia mediante las modificaciones de la secuencia aleatoria es la introducción de pre-ecos. Las pruebas subjetivas han mostrado que, para señales no estacionarias, los pre-ecos son, con mucho, más molestos que los post-ecos, lo que también está bien soportado por los principios psicoacústicos establecidos. Este problema podría ser reducido adaptando dinámicamente los tamaños de transformación a las características de la señal en términos de contenido transitorio. No obstante, cambiar los tamaños de transformación siempre es una decisión dura (es decir, binaria) que afecta a la totalidad del ancho de banda de señal, y que puede ser difícil de conseguir de una manera robusta.

25 La publicación de solicitud de patente de los Estados Unidos US 2003/0219130 A1 da a conocer una codificación y síntesis de audio basada en la coherencia. En particular, una escena auditiva es sintetizada a partir de una señal de audio mono modificando, para cada banda crítica, un parámetro de la escena auditiva tal como una diferencia de nivel inter-aural (ILD – Inter-aural Level Difference, en inglés) y/o una diferencia de tiempo inter-aural (ITD) para cada subbanda dentro de la banda crítica, en la que la modificación se basa en una coherencia estimada media para la banda crítica. La modificación basada en la coherencia produce escenas auditivas que tienen anchos de objeto, que coinciden con mayor precisión con los anchos de los objetos en la escena auditiva de entrada original. Los parámetros estéreo son los parámetros BCC bien conocidos, donde BCC significa codificación de indicación binaural (Binaural Cue Coding, en inglés). Cuando se generan dos canales de salida descorelacionados diferentes, los coeficientes de frecuencia obtenidos mediante una transformación discreta de Fourier se agrupan en una sola banda crítica. en base a la medida de coherencia entre canales, los factores de ponderación son multiplicados por una secuencia pseudoaleatoria que se elige preferiblemente de tal manera que la varianza sea aproximadamente constante para todas las bandas críticas, y la media es “0” dentro de cada banda crítica. La misma secuencia se aplica a los coeficientes espectrales de cada trama diferente.

30 La publicación de solicitud de patente de los Estados Unidos US 6005946 da a conocer una pluralidad de (al menos dos) señales de diferentes tipos que se forman a partir de la señal de entrada mono por filtrado pasa todo por subbanda con retardos enteros.

Compendio de la invención

50 El objetivo de la presente invención es proporcionar un concepto de descodificación para señales multicanal codificadas de manera paramétrica o un concepto de codificación para generar las señales que resultan en una buena calidad de audio y una buena eficiencia de codificación.

55 Este objetivo se consigue mediante un aparato para generar una señal de descorrelación de acuerdo con la reivindicación 1, un método para generar una señal de descorrelación de acuerdo con la reivindicación 7, o un programa informático de acuerdo con la reivindicación 9.

La presente invención se basa en encontrar que, en el lado de descodificación, se obtiene una buena señal de descorrelación para generar un primer y un segundo canal de una señal multicanal basada en la señal mono, cuando se utiliza un filtro de reverberación, que introduce un retardo entero o, preferiblemente, fraccional, en la señal de entrada. De manera importante, este filtro de reverberación no se aplica a la totalidad de la señal de entrada. Por el contrario, se aplican varios filtros de reverberación a varias subbandas de la señal de entrada original, es decir, la señal mono, de tal manera que el filtrado de reverberación utilizando los filtros de reverberación no se aplica en un dominio de tiempo o en el dominio de la frecuencia, es decir, en el dominio que se alcanza, cuando se aplica una transformada de Fourier. De acuerdo con la invención, el filtrado de reverberación utilizando filtros de reverberación para las subbandas se realiza de manera individual en el dominio de la subbanda.

Una señal de subbanda incluye una secuencia de al menos dos muestras de subbanda, representando la secuencia de las muestras de subbanda un ancho de banda de la señal de subbanda, que es menor que el ancho de banda de la señal de entrada. Naturalmente, el ancho de banda de la frecuencia de una señal de subbanda es mayor que el ancho de banda de la frecuencia atribuido a un coeficiente de frecuencia obtenido mediante una transformada de Fourier. Las señales de subbanda están generadas, preferiblemente, por medio de un banco de filtros que tiene, por ejemplo, 32 o 64 canales de banco de filtros, mientras que una FFT tendría, para el mismo ejemplo, 1.024 o 2.048 coeficientes de frecuencia, es decir, canales de frecuencia.

Las señales de subbanda son señales de subbanda obtenidas mediante el filtrado de subbanda de un bloque de muestras de la señal de entrada.

Puesto que el filtrado de reverberación no se aplica a toda la señal, sino que se aplica a nivel de subbanda, se evita un sonido "metálico" causado por el filtrado de peine.

En casos, en los que un período de muestra entre dos muestras de subbanda posteriores de la subbanda es demasiado grande para una buena impresión de sonido en el extremo del descodificador, se utilizan retardos fraccionales en un filtro de reverberación, tales como un retardo comprendido entre 0,1 y 0,9 y, preferiblemente, entre 0,2 y 0,8 del período de muestreo de la señal de subbanda. Se observa que, en caso de muestreo crítico, y cuando se generan 64 señales de subbanda utilizando un banco de filtros que tiene 64 canales de banco de filtros, el período de muestreo en una señal de subbanda es 64 veces mayor que el período de muestreo de la señal de entrada original.

Se debe observar, en el presente documento, que los retardos son una parte integral del proceso de filtrado utilizado en el dispositivo de reverberación. La señal de salida está formada por una multitud de versiones retardadas de la señal de entrada. Se prefiere retardar las señales mediante fracciones del período de muestreo de la subbanda, a fin de conseguir un buen dispositivo de verberación en el dominio de la subbanda.

En realizaciones preferidas de la presente invención, el retardo, y, preferiblemente, el retardo fraccional introducido por cada filtro de reverberación en cada subbanda es igual para todas las subbandas. Sin embargo, los coeficientes del filtro son diferentes para cada subbanda. Se prefiere utilizar filtros de IIR. Dependiendo de la situación real, el retardo fraccional y los coeficientes de filtro para los diferentes filtros se pueden determinar de manera empírica utilizando pruebas de audición.

Las subbandas filtradas mediante el conjunto de filtros de reverberación constituyen una señal de descorrelación que se mezclará con la señal de entrada original, es decir, la señal mono, para obtener un canal izquierdo descodificado y un canal derecho descodificado. Esta mezcla de una señal de descorrelación con la señal original se lleva a cabo en base a un parámetro de coherencia entre canales transmitido junto con la señal codificada de manera paramétrica. Para obtener diferentes canales izquierdo y derecho, es decir, diferentes primer y segundo canales, mezclar la señal de descorrelación con una señal mono para obtener el primer canal de salida es diferente de mezclar la señal de descorrelación con la señal mono para obtener el segundo canal de salida.

Para obtener una mayor eficiencia en el lado de la codificación, la codificación multicanal se realiza utilizando una determinación adaptativa del conjunto de parámetros estéreo. Para este fin, un codificador incluye, además, un medio para calcular la señal mono y, además, un medio para generar un conjunto de parámetros estéreo, un medio para determinar una validez de los conjuntos de parámetros estéreo para posteriores porciones de los canales izquierdo y derecho. Preferiblemente, el medio para la determinación es operativo para activar el medio para la generación, cuando se determina que el conjunto de parámetros estéreo ya no es válido, de tal manera que se calcula un segundo conjunto de parámetros estéreo para porciones de los canales izquierdo y derecho que comienzan en un segundo límite de tiempo. Este segundo límite de tiempo está determinado, asimismo, por el medio para determinar una validez.

La señal de salida codificada incluye la señal mono, un primer conjunto de parámetros estéreo y un primer límite de tiempo asociado con el primer conjunto de parámetros y el segundo conjunto de parámetros estéreo, y el segundo límite de tiempo asociado con el segundo conjunto de parámetros estéreo. En el lado de la descodificación, el descodificador utilizará un conjunto de parámetros estéreo válido hasta que se alcanza un nuevo límite de tiempo. Cuando se alcanza este nuevo límite de tiempo, las operaciones de descodificación se llevan a cabo utilizando el nuevo conjunto de parámetros estéreo.

5 En comparación con los métodos de la técnica anterior, que realizaron un procesamiento por bloques y, por lo tanto, una determinación inteligente de los conjuntos de parámetros estéreo, la determinación adaptativa según la invención de conjuntos de parámetros estéreo para diferentes límites de tiempo determinados en el lado del codificador proporciona una alta eficiencia de codificación, por una parte, y una alta calidad de codificación, por otra parte. Esto se debe al hecho de que, para señales relativamente estacionarias, el mismo conjunto de parámetros estéreo se puede utilizar para muchos bloques de las muestras de la señal mono sin introducir errores audibles. Por otra parte, por lo que se refiere a señales no estacionarias, la determinación adaptativa de parámetros estéreo según la invención proporciona una mejor resolución en tiempo, para que cada porción de señal tenga su conjunto de parámetros estéreo óptimo.

10 La presente invención proporciona una solución a los problemas de la técnica anterior mediante la utilización de una unidad de reverberación tal como un descorrelacionador implementado con líneas de retardo fraccionales en un banco de filtros, y utilizando un ajuste de nivel adaptativo de la señal reverberada descorrelacionada.

Posteriormente, se esbozan varios aspectos que no forman parte de la presente invención.

15 Un aspecto de la invención es un método para retardar una señal mediante: filtrado de una señal de valor real en el dominio del tiempo mediante la parte de análisis de un banco de filtros complejo; modificación de las señales de subbanda de valor complejo obtenidas a partir del filtrado; y filtrado de las señales de subbanda de valor complejo modificadas mediante la parte de síntesis del banco de filtros; y adopción de la parte real de la señal de salida de valor complejo en el dominio de tiempo, donde la señal de salida es la suma de las señales obtenidas a partir del filtrado de la síntesis.

20 Otro aspecto de la invención es un método para modificar las señales de subbanda de valor complejo filtrando cada señal de subbanda de valor complejo con un filtro de respuesta de impulsos finitos de valor complejo, en el que el filtro de respuesta de impulsos finitos para la subbanda número n viene dado por una transformada discreta de

$$H_n(\omega) = \begin{cases} \exp(-i\pi(n+1/2)\tau)G_\tau(\omega), & \text{para } n \text{ par;} \\ \exp(-i\pi(n+1/2)\tau)G_\tau(\omega+\pi), & \text{para } n \text{ impar.} \end{cases}$$

25 Fourier de la forma $\exp(-i\pi(n+1/2)\tau)G_\tau(\omega+\pi)$, en la que el parámetro $\tau = T/L$, y en la que el banco de filtros de síntesis tiene L subbandas y el retardo deseado está medido en T en unidades de la muestra de la señal de salida.

Otro aspecto de la invención es un método para modificar las señales de subbanda de valor complejo mediante filtrado, donde el filtro $G_\tau(\omega)$ cumple, aproximadamente, $V_\tau(\omega)G_\tau(\omega) + V_\tau(\omega+\pi)G_\tau(\omega+\pi) = 1$, donde $V_\tau(\omega)$ es la

$$v_\tau(k) = A i^k \sum_l p(l)p(l-T-Lk)$$

transformada discreta de Fourier en el tiempo de la secuencia $v_\tau(k)$, y $p(l)$ es el filtro prototipo de dicho banco de filtros complejo, y A es un factor de normalización real apropiado.

30 Otro aspecto de la invención es un método para modificar las señales de subbanda de valor complejo mediante filtrado, donde el filtro $G_\tau(\omega)$ cumple $G_\tau(-\omega) = G_\tau(\omega+\pi)^*$, de tal manera que las muestras de respuesta de impulsos de índice par tienen un valor real, y las muestras de respuesta de impulsos de índice impar tienen un valor imaginario puro.

35 Otro aspecto de la invención es un método para codificar las propiedades estéreo de una señal de entrada mediante, en un codificador, calcular los parámetros de la retícula de tiempos que describen la ubicación en el tiempo para cada conjunto de parámetros estéreo, donde el número de conjuntos de parámetros estéreo es arbitrario, y mediante, en un descodificador, aplicar síntesis estéreo paramétrica de acuerdo con esa retícula de tiempo.

40 Otro aspecto de la invención es un método para la codificación de las propiedades estéreo de una señal de entrada, en la que la localización en el tiempo para el conjunto de parámetros estéreo está señalizada, en el caso de que una indicación de tiempo para el conjunto de parámetros estéreo coincida con el inicio de una trama, de manera explícita, en lugar de señalar el puntero de tiempo.

45 Otro aspecto de la invención es un método para la generación de descorrelación estéreo para la reconstrucción estéreo paramétrica mediante, en un descodificador, aplicar un proceso de reverberación artificial para sintetizar la señal lateral.

Otro aspecto de la invención es un método para la generación de una descorrelación estéreo para la reconstrucción estéreo paramétrica mediante, en el descodificador, llevar a cabo el proceso de reverberación en un banco de filtros modulado complejo utilizando ajuste de retardo de fase en cada canal del banco de filtros.

50 Otro aspecto de la invención es un método para la generación de descorrelación estéreo para la reconstrucción estéreo paramétrica mediante, en el descodificador, la utilización por parte del proceso de reverberación de un

detector diseñado para encontrar señales en las que la cola de reverberación podría no ser deseada y dejar que la cola de reverberación sea atenuada o eliminada.

Breve descripción de los dibujos

5 La presente invención se describirá, a continuación, a modo de ejemplos ilustrativos, que no limitan el alcance de la invención, haciendo referencia a los dibujos adjuntos, en los cuales:

la figura 1 ilustra un diagrama de bloques del aparato según la invención;

la figura 2 ilustra un diagrama de bloques del medio para generar una señal descorrelacionada;

la figura 3 ilustra el análisis de un solo canal y la síntesis del par de canales estéreo en base a las señales de subbanda estéreo reconstruidas de acuerdo con la presente invención;

10 la figura 4 ilustra un diagrama de bloques de la división en segmentos de tiempo de los conjuntos de parámetros estéreo paramétricos, en base a la característica de la señal; y

la figura 5 ilustra un ejemplo de la división en segmentos de tiempo de los conjuntos de parámetros estéreo paramétricos, en base a la característica de la señal.

Descripción de realizaciones preferidas

15 Las realizaciones que se describen a continuación son meramente ilustrativas para los principios de la presente invención para codificación estéreo paramétrica. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en este documento resultarán evidentes para otros expertos en la técnica. La intención es, por lo tanto, estar limitados solo por el alcance de las reivindicaciones inminentes de la patente, y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en este documento.

20 Retardar una señal en una fracción de una muestra se puede conseguir mediante varios métodos de interpolación de la técnica anterior. No obstante, surgen casos especiales cuando la señal original está disponible como muestras de valor complejo sobremuestreadas. Llevar a cabo un retardo fraccional en el banco de qmf aplicando solo el retardo de fase en un factor para cada canal de qmf correspondiente a un retardo de tiempo constante, resulta en artefactos importantes.

25 Esto se puede evitar de manera eficiente mediante la utilización de un filtro de compensación de acuerdo con un enfoque novedoso que permite aproximaciones de alta calidad a retardos arbitrarios en cualquier banco de filtros modulado, exponencial y complejo. A continuación, sigue una descripción detallada.

Un modelo de tiempo continuo

30 Para facilitar los cálculos, un banco de filtros de L bandas, modulado, exponencial, complejo, se modelará en este caso mediante una transformada continua por ventanas de tiempo utilizando las formas de onda de síntesis

$$u_{n,k}(t) = v(t-k) \exp[i\pi(n+1/2)(t-k+\theta)], \tag{1}$$

35 donde n, k son enteros con $n \geq 0$ y θ es un término de fase fija. Se obtienen resultados para señales de tiempo discreto mediante un muestreo adecuado de la variable t con separación de $1/L$. Se supone que la ventana de valor real $v(t)$ se elige de tal manera que para señales de valor reales $x(t)$ tiene una precisión muy alta tal que

$$x(t) = 2 \operatorname{Re} \left\{ \sum_{n=0}^{\infty} \sum_{k=-\infty}^{\infty} c_n(k) u_{n,k}(t) \right\} \tag{2}$$

$$c_n(k) = \int_{-\infty}^{\infty} x(t) u_{n,k}^*(t) dt, \tag{3}$$

si

40 donde * denota conjugación compleja. Asimismo, se supone que $v(t)$ está esencialmente limitado por la banda al intervalo de frecuencia $[-\pi, \pi]$. Considérese la modificación de cada banda de frecuencia n filtrando las muestras discretas de análisis de tiempo $c_n(k)$ con un filtro con respuesta de impulso $h_n(k)$,

$$d_n(k) = \sum_l h_n(l) c_n(k-l). \tag{4}$$

A continuación, la síntesis modificada

$$y(t) = 2 \operatorname{Re} \left\{ \sum_{n=0}^{\infty} \sum_{k=-\infty}^{\infty} d_n(k) u_{n,k}(t) \right\} \quad (5)$$

se puede calcular en el dominio de frecuencia para ser

$$\hat{y}(\omega) = H(\omega) \hat{x}(\omega), \quad (6)$$

donde $\hat{f}(\omega)$ denota transformadas de Fourier de $f(t)$ y

$$H(\omega) = \sum_{n=-\infty}^{\infty} H_n(\omega) |\hat{v}(\omega - \pi(n+1/2))|^2. \quad (7)$$

5 En este caso, $H_n(\omega) = \sum_k h_n(k) \exp(-ik\omega)$ es la transformada de Fourier en tiempo discreto del filtro aplicado en la banda de frecuencia n para $n \geq 0$ y

$$H_n(\omega) = H_{-1-n}(-\omega)^* \quad \text{para } n < 0. \quad (8)$$

10 Se debe observar, en este caso, que el caso especial $H_n(\omega) = 1$ conduce a $H(\omega) = 1$ en (7) debido al diseño especial de la ventana $v(t)$. Otro caso de interés es $H_n(\omega) = \exp(-i\omega)$ que conduce a $H(\omega) = \exp(-i\omega)$, de tal manera que $y(t) = x(t-1)$.

La solución propuesta

Con el fin de conseguir un retardo de tamaño τ de tal manera que $y(t) = x(t-\tau)$, el problema es diseñar filtros $H_n(\omega)$ para $n \geq 0$ de tal manera que

$$15 \quad H(\omega) = \exp(-i\tau\omega), \quad (9)$$

donde $H(\omega)$ viene dado por (7) y (8) La solución particular propuesta, en este caso, es aplicar los filtros

$$H_n(\omega) = \begin{cases} \exp(-i\pi(n+1/2)\tau) G_\tau(\omega), & \text{para } n \text{ par;} \\ \exp(-i\pi(n+1/2)\tau) G_\tau(\omega + \pi), & \text{para } n \text{ impar.} \end{cases} \quad (10)$$

En este caso, $G_\tau(-\omega) = G_\tau(\omega + \pi)^*$ implica consistencia con (8) para todo n . La inserción de (10) en el lado derecho de (7) conduce a

$$20 \quad H(\omega) = \exp(-i\omega\tau) [V_\tau(\omega) G_\tau(\omega) + V_\tau(\omega + \pi) G_\tau(\omega + \pi)] \quad (11)$$

donde $V_\tau(\omega) = \sum_n b(\omega - \pi(2n+1/2))$ siendo $b(\omega) = \exp(i\tau\omega) |\hat{v}(\omega)|^2$. Cálculos elementales muestran que $V_\tau(\omega)$ es la transformada de Fourier en tiempo discreto de

$$v_\tau(k) = i^k \int_{-\infty}^{\infty} v(t) v(t-\tau-k) dt. \quad (12)$$

Se pueden obtener muy buenas aproximaciones al retardo perfecto resolviendo el sistema lineal

$$25 \quad V_\tau(\omega) G_\tau(\omega) + V_\tau(\omega + \pi) G_\tau(\omega + \pi) = 1 \quad (13)$$

en el sentido de mínimos cuadrados con un filtro FIR

$G_\tau(\omega) = \sum_{k=-N}^M g_\tau(k) \exp(-ik\omega)$. En términos de coeficientes de filtro, la ecuación (13) se puede escribir

$$2 \sum_l v_\tau(2k-l) g_\tau(l) = \delta[k], \quad (14)$$

donde $\delta[k] = 1$ para $k = 0$ y $\delta[k] = 0$ para $k \neq 0$.

30 En el caso de un banco de filtros de L bandas en tiempo discreto con filtro prototipo $p(k)$, el retardo obtenido en unidades de muestra es $L\tau$ y el cálculo (12) se sustituye por

$$v_r(k) = i^k \sum_l p(l) p(l - T - Lk), \quad (15)$$

donde T es el número entero más cercano a $L\tau$. En este caso $p(k)$ se extiende por ceros fuera de su soporte. Para un filtro prototipo de longitud finita, solo de manera finita, muchos $v_r(k)$ son diferentes de cero, y (14) es un sistema de ecuaciones lineales. El número de incógnitas $g_r(k)$ se elige, típicamente, para ser un número pequeño. Para buenos diseños de banco de filtros QMF, entre 3 y 4 coeficientes ya dan muy buen rendimiento de retardo. Además, la dependencia de los coeficientes de filtro $g_r(k)$ del parámetro de retardo τ a menudo puede ser modelada con éxito mediante polinomios de bajo orden.

Retícula de tiempo adaptativo de señalización para parámetros estéreo

Los sistemas estéreo paramétricos siempre conllevan compromisos en términos de tiempo limitado o de resolución de frecuencia, con el fin de minimizar los datos transportados. Sin embargo, es bien conocido a partir de la psicoacústica que algunas indicaciones espaciales pueden ser más importantes que otras, lo que conduce a la posibilidad de descartar las indicaciones menos importantes. Por lo tanto, la resolución en tiempo no tiene que ser constante. Se puede conseguir una ganancia grande en la velocidad de bits permitiendo que la retícula de tiempo se sincronice con las indicaciones espaciales. Se puede hacer fácilmente enviando un número variable de conjuntos de parámetros para cada trama de datos que corresponde a un segmento de tiempo de tamaño fijo. Para sincronizar los conjuntos de parámetros con las indicaciones espaciales correspondientes, se deben enviar datos adicionales de la retícula de tiempo describiendo la ubicación en el tiempo para cada conjunto de parámetros. La resolución de dichos punteros de tiempo se podría elegir para ser bastante bajos, para mantener la cantidad total de datos en un mínimo. Un caso especial en el que una indicación de tiempo para un conjunto de parámetros coincide con el comienzo de una trama podría ser señalizado de manera explícita para evitar enviar ese puntero de tiempo.

La figura 4 ilustra un aparato según la invención para llevar a cabo un análisis paramétrico para segmentos de tiempo que tienen límites de tiempo que dependen de variables y de señal. El aparato según la invención incluye el medio 401 para dividir la señal de entrada en uno o varios segmentos de tiempo. Los límites de tiempo que separan los segmentos de tiempo están proporcionados por el medio 402. El medio 402 utiliza un detector especialmente diseñado firmado para extraer indicaciones espaciales que sean relevante para decidir dónde establecer los límites de tiempo. El medio 401 envía todas las señales de entrada divididas en uno o varios segmentos de tiempo. Esta salida es introducida en el medio 403 para el análisis paramétrico separado para cada segmento de tiempo. el medio 403 envía de un conjunto de parámetros por cada segmento de tiempo que se está analizando.

La figura 5 ilustra un ejemplo de cómo el generador de retícula de tiempo puede funcionar para una hipotética señal de entrada. En este ejemplo se utiliza un conjunto de parámetros por cada trama de datos, si no existe otra información sobre el límite de tiempo. Por lo tanto, cuando no existe otra información de límite de tiempo, se utilizan los límites de tiempo inherentes de la trama de datos. Los límites de tiempo representados en la figura 5 son la salida del medio 402 en la figura 4. Los límites de tiempo representados en la figura 5 son proporcionan por el medio 401 en la figura 4.

El aparato para codificar una señal estéreo para obtener una señal de salida mono y el conjunto de parámetros estéreo incluye el medio para calcular la señal mono combinando un canal izquierdo y uno derecho de las señales estéreo mediante suma ponderada. Adicionalmente, un medio 403 está generando un primer conjunto de parámetros estéreo utilizando una porción del canal izquierdo y una porción del canal derecho, estando conectadas las porciones que empiezan en un primer límite de tiempo al medio para la determinación de la validez del primer conjunto de parámetros estéreo para porciones posteriores del canal izquierdo y del canal derecho.

El medio para la determinación está formado conjuntamente por los medios 402 y 401 en la figura 1.

En particular, el medio para la determinación funciona para generar un segundo límite de tiempo y para activar el medio de generación, cuando se determina que este primer conjunto de parámetros estéreo ya no es válido, de tal manera que se genera un segundo conjunto de parámetros estéreo para porciones de los canales izquierdo y derecho que comienzan en el segundo el límite de tiempo.

No mostrados en la figura 4 están dispuestos medios para emitir la señal mono, el primer conjunto de parámetros estéreo y el primer límite de tiempo asociado con el primer conjunto de parámetros estéreo y el segundo conjunto de parámetros estéreo y el segundo límite de tiempo asociado con el segundo conjunto de parámetros estéreo como la señal estéreo codificada de manera paramétrica. El medio para la determinación de la validez de un conjunto de parámetros estéreo puede incluir un detector transitorio, puesto que existe una alta probabilidad de que, después de un transitorio, un nuevo parámetro estéreo tenga que ser generado, puesto que una señal ha cambiado su forma de manera significativa. De manera alternativa, el medio para la determinación de una validez puede incluir un dispositivo de análisis por síntesis, que está adaptado para descodificar la señal mono y el conjunto de parámetros estéreo para obtener un canal izquierdo descodificado y un canal derecho descodificado, para comparar el canal izquierdo descodificado y el canal derecho descodificado con el canal izquierdo y el canal derecho, y para activar el medio para la generación, cuando el canal izquierdo descodificado y el canal derecho descodificado es diferente del canal izquierdo y del canal derecho en más del umbral predeterminado.

Trama de datos 1: el segmento de tiempo correspondiente al conjunto de parámetros 1 comienza al inicio de la trama de datos 1, puesto que no existe ninguna otra información sobre el límite de tiempo en esta trama de datos.

5 Trama de datos 2: existen dos límites de tiempo en esta trama de datos. El segmento de tiempo correspondiente al conjunto de parámetros 2 comienza en el primer límite de tiempo en esta trama de datos. El segmento de tiempo correspondiente al conjunto de parámetros 3 comienza en el segundo límite de tiempo en esta trama de datos.

Trama de datos 3: existe un límite de tiempo en esta trama de datos. El segmento de tiempo correspondiente al conjunto de parámetros 4 comienza en el límite de tiempo en esta trama de datos.

10 Trama de datos 4: existe un límite de tiempo en esta trama de datos. Este límite de tiempo coincide con el límite de inicio de la trama de datos 4 y no tiene que ser señalado, puesto que esto se maneja con el caso predeterminado. Por lo tanto, esta señal de límite de tiempo puede ser eliminada. El segmento de tiempo correspondiente al conjunto de parámetros 5 comienza al inicio de la trama de datos 4, incluso sin señalar este límite de tiempo.

Uso de la reverberación artificial como método de descorrelación para la reconstrucción estéreo paramétrica

15 Una parte vital para realizar la síntesis estéreo en un sistema estéreo paramétrico es disminuir la coherencia entre los canales izquierdo y derecho con el fin de crear la amplitud de la imagen estéreo. Esto se puede realizar agregando una versión filtrada de la señal mono original a la señal lateral, donde las señales lateral y mono se definen por:

mono = (izquierda + derecha) / 2, y

lateral = (izquierda - derecha) / 2, respectivamente.

20 Con el fin de no cambiar demasiado el timbre, el filtro en cuestión debe ser preferiblemente, de carácter pasa todo. Un enfoque que tiene éxito es utilizar filtros pasa todo similares utilizados para procesos de reverberación artificial. Los algoritmos de reverberación artificial requieren, generalmente, una alta resolución de tiempo para proporcionar una respuesta de impulso que es satisfactoriamente difusa en el tiempo. Existen grandes ventajas en basar un algoritmo de reverberación artificial en un banco de filtros complejo, tal como el banco qmf complejo. El banco de
25 filtros ofrece excelentes posibilidades para dejar que las propiedades de reverberación sean selectivas en frecuencia en términos, por ejemplo, de equalización de la reverberación, tiempo de decaimiento, densidad y timbre. No obstante, las implementaciones del banco de filtros, generalmente, intercambian resolución de tiempo por una resolución de mayor frecuencia, que, normalmente, hace que sea difícil implementar un proceso de reverberación que sea lo suficientemente suave en el tiempo. Para hacer frente a este problema, un método novedoso sería utilizar una aproximación de retardo fraccional aplicando solo retardo de fase por un factor para cada canal qmf
30 correspondiente a un retardo de tiempo constante. Este primitivo método de retardo fraccional introduce una importante difuminación en el tiempo que, afortunadamente, es muy deseada en este caso. La difuminación en el tiempo contribuye a la difusión en el tiempo, que es altamente deseable para los algoritmos de reverberación, y aumenta a medida que el retardo de fase se acerca a $\pi / 2$ o $-\pi / 2$.

35 Los procesos de reverberación artificial son, por razones naturales, procesos con una respuesta de impulso infinita, y ofrecen decaimientos exponenciales. En [PCT/SE02/01372] se señala que, si se utiliza una unidad de reverberación para generar una señal estéreo, el decaimiento de la reverberación a veces puede no ser deseado después, justo, del final de un sonido. Sin embargo, estas colas de reverberación no deseadas pueden ser fácilmente atenuadas o eliminadas por completo simplemente alterando la ganancia de la señal de reverberación. Se puede utilizar un detector diseñado para encontrar terminaciones de sonido para ese propósito. Si la unidad de
40 reverberación genera artefactos en algunas señales específicas, por ejemplo, transitorios, un detector para esas señales también se puede utilizar para atenuar los mismos.

45 La figura 1 ilustra un aparato según la invención para el método de descorrelación de señales tal como se utiliza en un sistema estéreo paramétrico. El aparato de la invención incluye un medio 101 para proporcionar una pluralidad de señales de subbanda. El medio para proporcionar puede ser un banco de filtros QMF complejo, en el que cada señal está asociada con un índice de subbanda.

Las señales de subbanda emitidas por el medio 101 de la figura 1 son introducidas en un medio 102 para proporcionar una señal descorrelacionada 102, y en un medio 103 y 106 para modificar la señal de subbanda. La salida de 102 es introducida en un medio 104 y 105 para modificar la señal, y las salidas de 103, 104, 105 y 106 son introducidas en un medio para agregar, 107 y 108, las señales de subbanda.

50 En la realización de la invención actualmente descrita, el medio para modificar 103, 104, 105 y 106, las señales de subbanda, ajusta el nivel de la señal descorrelacionada y la señal no procesada que es la salida de 101, multiplicando la señal de subbanda con un factor de ganancia, de tal manera que cada suma de cada par da como resultado una señal con la cantidad de señal descorrelacionada proporcionada por los parámetros de control. Se debe observar que los factores de ganancia utilizados en los medios para modificar, 103 a 106, no están limitados a
55 un valor positivo. También puede ser un valor negativo.

La salida del medio para agregar las señales de subbanda 107 y 108, es introducida en los medios 109 y 110 para proporcionar una señal en el dominio del tiempo. La salida de 109 corresponde al canal izquierdo de la señal estéreo reconstruida y la salida de 110 corresponde al canal derecho de la señal estéreo reconstruida. En la realización descrita en el presente documento, se utiliza el mismo correlacionador se utiliza para ambos canales de salida, mientras que los medios para agregar la señal descorrelacionada con la señal no procesada están separados para los dos canales de salida. La realización descrita actualmente asegura de este modo que las dos señales de salida pueden ser idénticas, así como estar completamente descorrelacionadas, dependiendo de los datos de control proporcionados al medio para ajustar los niveles de las señales y los datos de control proporcionados al medio para sumar las señales.

En la figura 2, se muestra un diagrama de bloques del medio para proporcionar una señal descorrelacionada. La señal de subbanda de entrada es introducida en el medio para filtrar una señal de subbanda 201. En la realización de la presente invención descrita actualmente, la etapa de filtrado es una unidad de reverberación que incorpora filtrado mediante un filtro pasa todo. Los coeficientes de filtro utilizados vienen dados por el medio para proporcionar coeficientes de filtro 202. El índice de subbanda de la señal de subbanda procesada actualmente es introducido en 202. En una realización de la presente invención se calculan diferentes coeficientes de filtro en base al índice de subbanda proporcionado a 202. La etapa de filtrado en 201, se basa en muestras retardadas de la señal de subbanda de entrada, así como en muestras retardadas de las señales intermedias en el procedimiento de filtrado.

Una característica esencial de la presente invención es que, en 203, se proporciona un medio para proporcionar retardo de muestra de subbanda entero y retardo de muestra de subbanda fraccional. La salida de 201 es introducida en un medio para ajustar el nivel de la señal de subbanda 204, y también en un medio para estimar las características de señal de la señal de subbanda 205. En una realización preferida de la presente invención, las características estimadas son el comportamiento transitorio de la señal de subbanda. En esta realización, un transitorio detectado es señalado al medio para ajustar el nivel de una señal de subbanda 204, de tal manera que el nivel de la señal se reduce durante pasajes transitorios. La salida de 204 es la señal descorrelacionada introducida en 104 y 105 de la figura 1.

En la figura 3, se muestran el banco de filtros de análisis único y los dos bancos de filtros de síntesis. El banco de filtros de análisis 301, actúa sobre la señal de entrada mono, mientras que los bancos de filtros de síntesis 302 y 303 actúan sobre las señales estéreo reconstruidas.

La figura 1, por lo tanto, muestra el aparato según la invención para generar una señal de descorrelación que se indica mediante la referencia 102. Tal como se muestra en la figura 1 o 3, este aparato incluye un medio para proporcionar una pluralidad de señales de subbanda, en el que una señal de subbanda incluye la secuencia de al menos dos muestras de subbanda, representando la secuencia de las muestras de subbanda un ancho de banda de la señal de subbanda que es menor que el ancho de banda de la señal de entrada. Cada señal de subbanda es introducida en el medio 201 para filtrado. Cada medio 201 para filtrado incluye un filtro de reverberación, de tal manera que se obtiene una pluralidad de señales de subbanda reverberadas, en el que la pluralidad de señales de subbanda reverberadas juntas representan la señal de descorrelación. Preferiblemente, tal como se muestra en la figura 2, puede tener lugar un post-procesamiento por cada subbanda de las señales de subbanda reverberadas que es llevado a cabo por el bloque 204, que está controlado por el bloque 205.

Cada filtro de reverberación está configurado para un cierto retardo y, preferiblemente, un retardo fraccional, y cada filtro de reverberación tiene varios coeficientes de filtro, que, tal como se muestra en la figura 2, dependen del índice de subbanda. Esto significa que se prefiere utilizar el mismo retardo para cada subbanda, pero utilizar conjuntos diferentes de coeficientes de filtro para las diferentes subbandas. Esto es simbolizado por los medios 203 y 202 en la figura 2, aunque se debe indicar, en este caso, que los retardos y los coeficientes de filtro son determinados preferiblemente, de manera fija enviando un dispositivo de descorrelación, en el que los retardos y los coeficientes de filtro se pueden determinar utilizando empíricamente pruebas de audición, etc.

La figura 1 muestra un descodificador multicanal e incluye el aparato de la invención para generar la señal de correlación, que está designado con 102 en la figura 1. El descodificador multicanal mostrado en la figura 1 está destinado a descodificar una señal mono y una medida de coherencia entre canales asociada, representando la medida de coherencia entre canales una coherencia entre una pluralidad de canales originales, en los que la señal mono se deriva de la pluralidad de canales originales. El bloque 102 en la figura 1 constituye un generador, para generar una señal de descorrelación para la señal mono. Los bloques 103, 104, 105, 106 y 107 y 108 constituyen un mezclador, para mezclar la señal mono y la señal de descorrelación de acuerdo con el primer modo de mezcla, para obtener una primera señal de salida descodificada y, de acuerdo con el segundo modo de mezcla, para obtener una segunda señal de salida descodificada, en el que el mezclador funciona para determinar el primer modo de mezcla y el segundo modo de mezcla en base a la medida de coherencia entre canales transmitida como una información lateral a la señal mono.

El mezclador funciona, preferiblemente, para mezclar en un dominio de subbanda en base a medidas de coherencia entre canales separadas para diferentes subbandas. En este caso, el descodificador multicanal adicional comprende, además, medios 109 y 110 para convertir las primera y segunda señales de salida descodificadas del dominio de subbanda en un dominio de tiempo para obtener una primera señal de salida descodificada y una

segunda señal de salida descodificada en el dominio del tiempo. Por lo tanto, el medio 102 de la invención para generar una señal de descorrelación y el descodificador multicanal de la invención tal como se muestra en la figura 1 funcionan en el dominio de la subbanda y llevan a cabo, justo como última etapa, una conversión del dominio de la subbanda al dominio del tiempo.

5

Dependiendo de la situación real, el dispositivo de la invención se puede implementar en hardware o en software o en un firmware, incluyendo componentes de hardware y componentes de software. Cuando se implementa parcial o totalmente en software, la invención también es un programa informático que tiene un código legible por ordenador para llevar a cabo los métodos de la invención cuando es ejecutado en un ordenador.

10

REIVINDICACIONES

1. Aparato para generar una señal de descorrelación utilizando una señal de entrada, que comprende:

5 un banco de filtros de subbanda, para proporcionar una pluralidad de señales de subbanda de valor complejo, en el que una señal de subbanda incluye una secuencia de al menos dos muestras de subbanda, representando la secuencia de las muestras de subbanda un ancho de banda de la señal de subbanda, que es menor que el ancho de banda de la señal de entrada, en el que la señal de entrada incluye un bloque de un número predeterminado de muestras de entrada, siendo el número de muestras de subbanda en una señal de subbanda menor que el número de muestras de entrada; y

10 un filtro de reverberación, para obtener una pluralidad de señales de subbanda reverberadas de la pluralidad de señales de subbanda, teniendo el filtrado del filtro de reverberación una característica de pasa todo, funcionando el filtro de reverberación para aplicar un retardo fraccional a cada señal de subbanda y estando adaptado para tener diferentes conjuntos de coeficientes de filtro para cada señal de subbanda, en el que la pluralidad de señales de subbanda reverberadas juntas representan la señal de descorrelación.

15 2. El aparato de la reivindicación 1, en el que el retardo fraccional es mayor de cero y menor que un período de muestreo.

3. El aparato de la reivindicación 2, en el que el retardo fraccional es menor que 0,9 veces el período de muestreo de la señal de subbanda y mayor que 0,1 veces el período de muestreo de la señal de subbanda.

4. El aparato de acuerdo con cualquiera de las reivindicaciones anteriores, en el que el filtro de reverberación funciona para introducir retardos de fase predeterminados en las señales de subbanda.

20 5. Método para generar una señal de descorrelación utilizando una señal de entrada, que comprende:

proporcionar (101) una pluralidad de señales de subbanda de valor complejo, en el que una señal de subbanda incluye una secuencia de al menos dos muestras de subbanda, representando la secuencia de muestras de subbanda un ancho de banda de la señal de subbanda, que es menor que el ancho de banda de la señal de entrada, en el que la señal de entrada incluye un bloque de un número predeterminado de muestras de entrada, siendo el número de muestras de subbanda en una señal de subbanda menor que el número de muestras de entrada; y

30 filtrar (201) cada señal de subbanda utilizando un filtro de reverberación para obtener una pluralidad de señales de subbanda reverberadas, teniendo el filtrado del filtro de reverberación una característica de pasa todo, en el que el filtro de reverberación funciona para aplicar un retardo fraccional a cada señal de subbanda, y en el que el filtro de reverberación está adaptado para tener diferentes conjuntos de coeficientes de filtro para cada señal de subbanda, en el que una pluralidad de señales de subbanda reverberadas juntas representan la señal de descorrelación.

6. Un programa informático que tiene un código legible por ordenador para llevar a cabo, cuando es ejecutado en un ordenador, el método de la reivindicación 5.

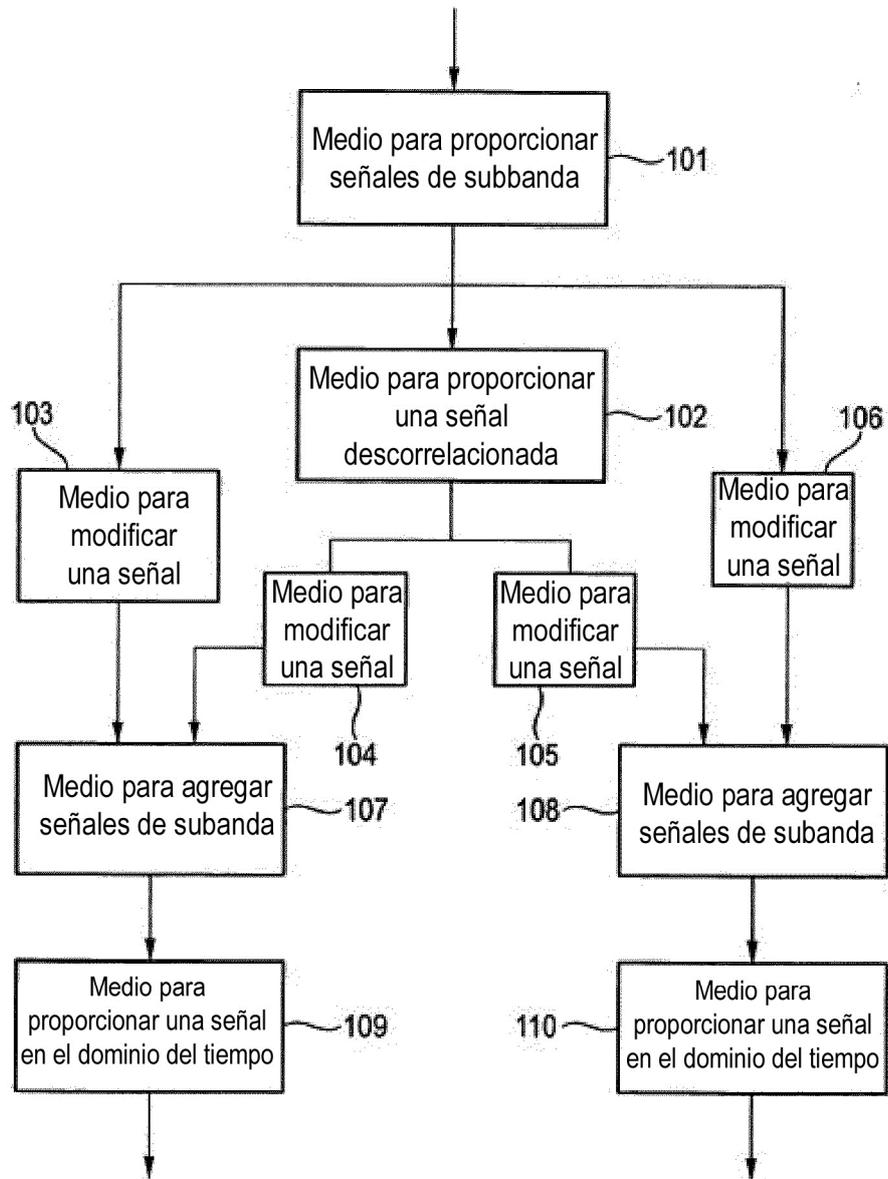


FIG. 1

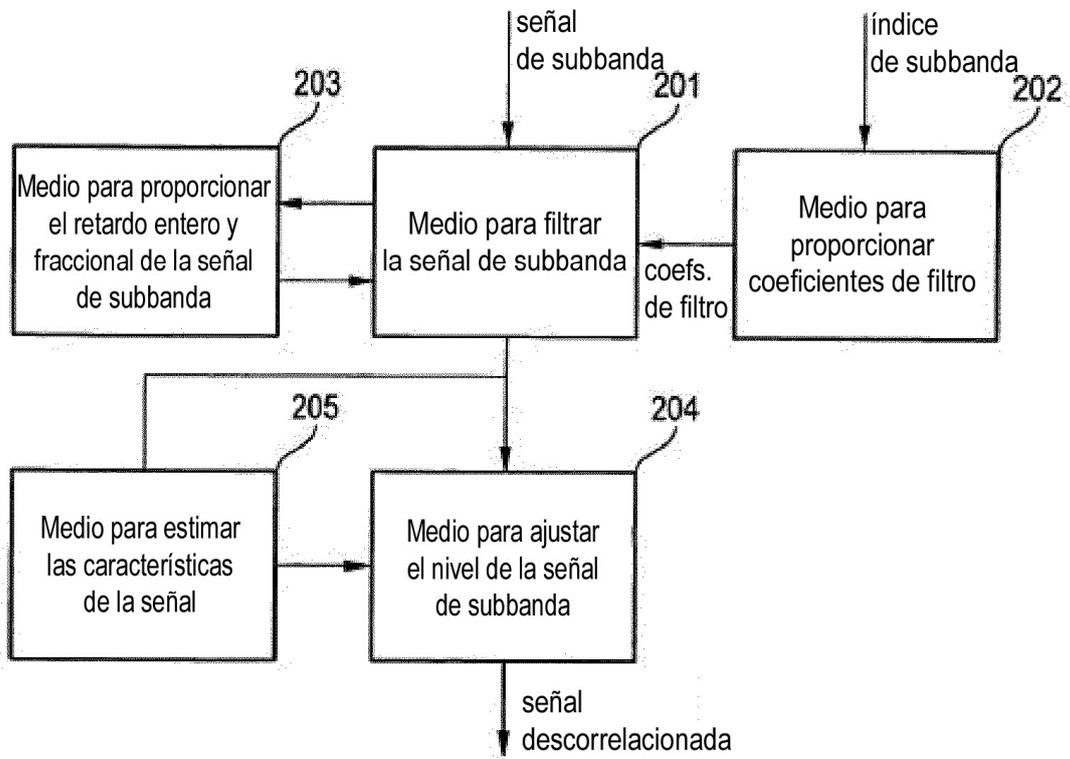


FIG. 2

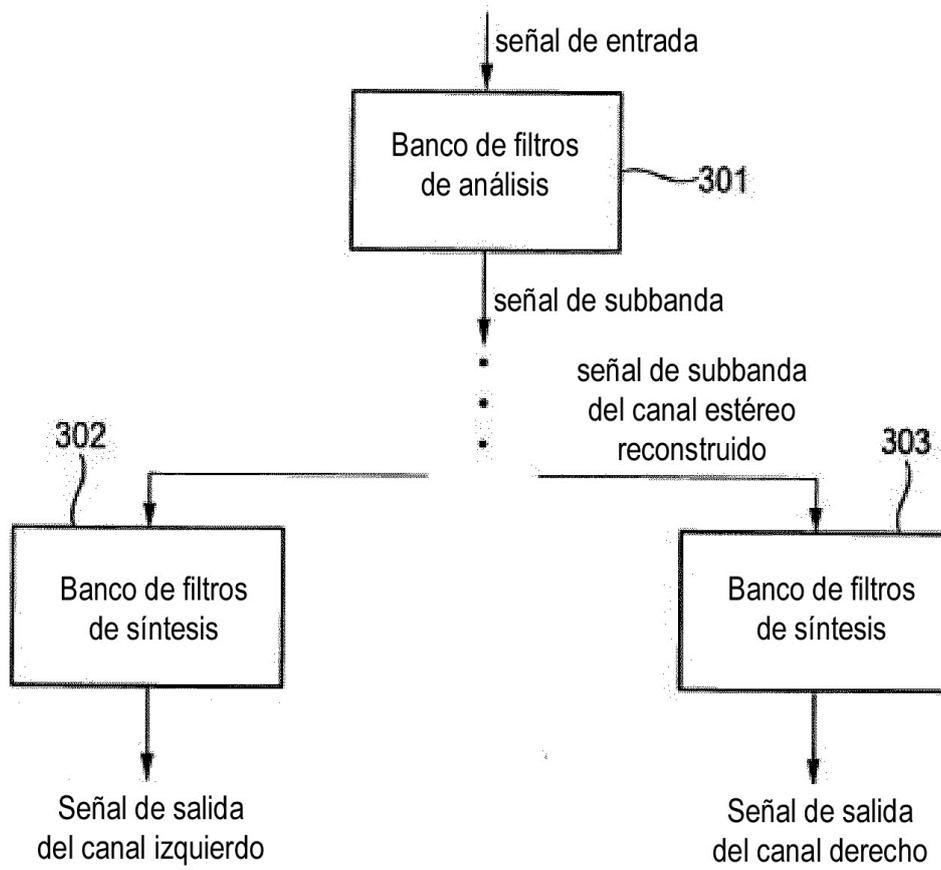


FIG. 3

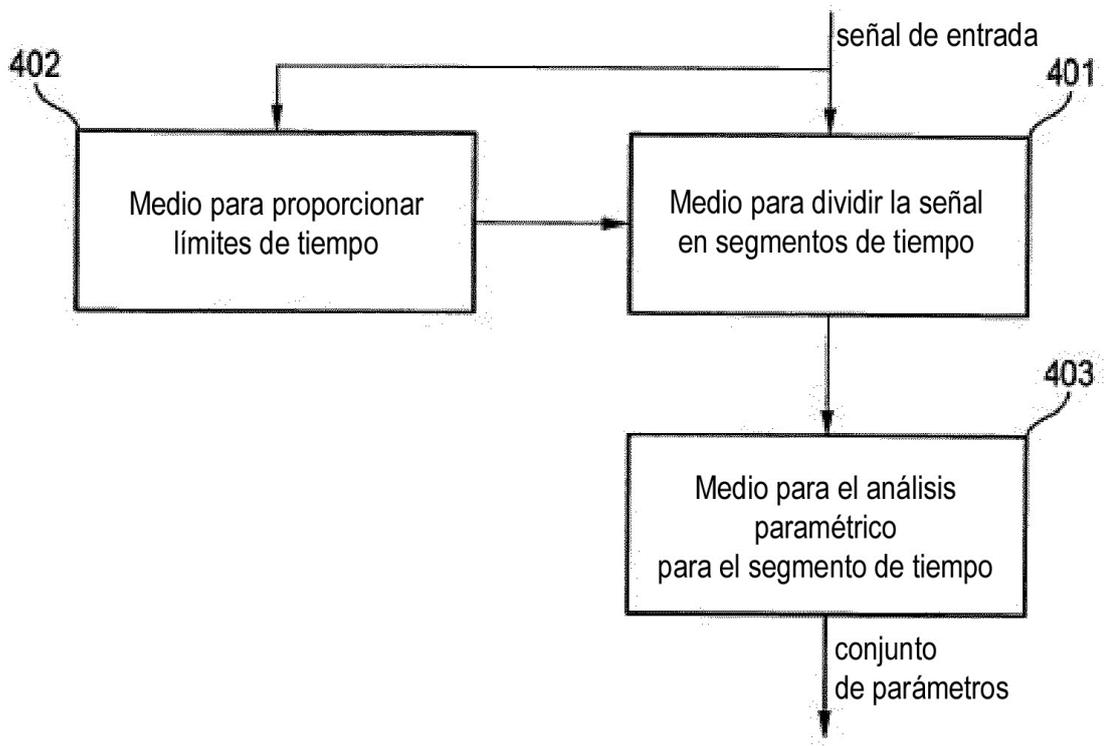


FIG. 4

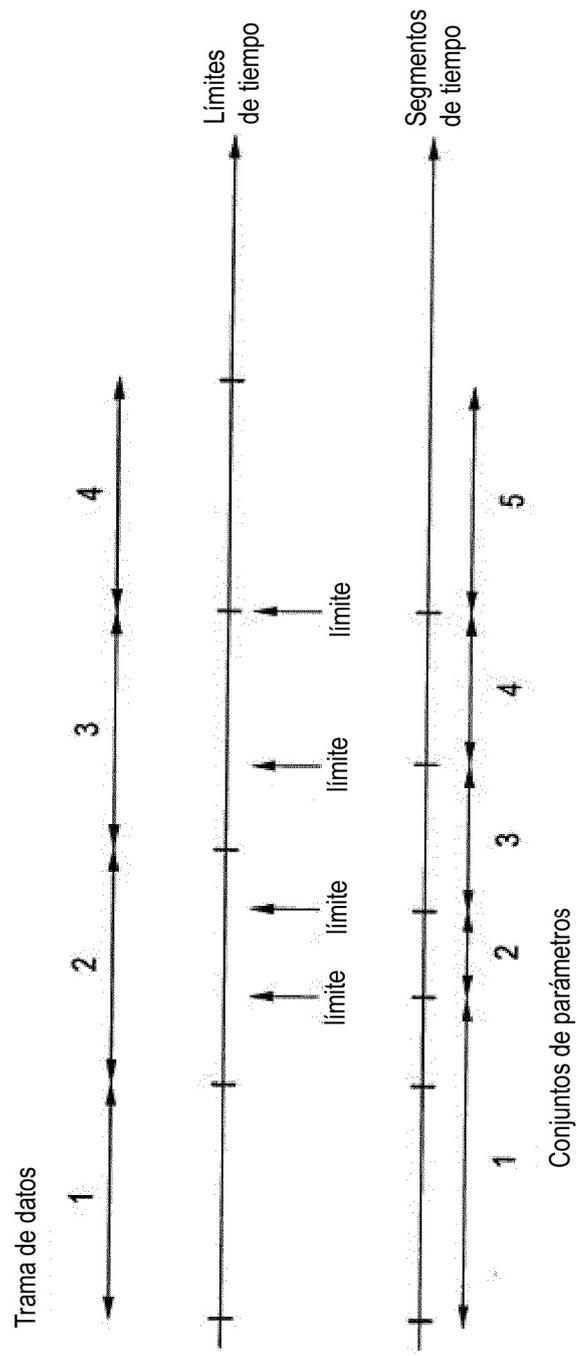


FIG. 5