

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 376 989**

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

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **09728768 .4**
- 96 Fecha de presentación: **23.03.2009**
- 97 Número de publicación de la solicitud: **2147430**
- 97 Fecha de publicación de la solicitud: **27.01.2010**

54 Título: **Codificación por transformada de audio utilizando corrección tonal**

30 Prioridad:
04.04.2008 US 42314
08.12.2008 EP 08021298

45 Fecha de publicación de la mención BOPI:
21.03.2012

45 Fecha de la publicación del folleto de la patente:
21.03.2012

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

72 Inventor/es:
EDLER, Bernd;
DISCH, Sascha;
GEIGER, Ralf;
BAYER, Stefan;
KRAEMER, Ulrich;
FUCHS, Guillaume;
NEUENDORF, Max;
MULTRUS, Markus;
SCHULLER, Gerald y
POPP, Harald

74 Agente/Representante:
Arizti Acha, Monica

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).

ES 2 376 989 T3

DESCRIPCIÓN

Codificación por transformada de audio utilizando corrección tonal.

Campo de la invención

5 Varias realizaciones de la presente invención se refieren a procesadores de audio para generar una representación procesada de una señal de audio en tramas utilizando el muestreo y remuestreo en función de la altura tonal de las señales.

Antecedentes de la invención y técnica anterior

10 A menudo se utilizan transformadas solapadas moduladas basadas en coseno o seno correspondientes a bancos de filtros modulados, en aplicaciones en codificación de fuente debido a sus propiedades de compactación de energía. Es decir, para tonos armónicos con frecuencias fundamentales (altura tonal) constantes, concentran la energía de señal en un número bajo de componentes espectrales (subbandas), lo que lleva a representaciones eficaces de la señal. En general, la altura tonal de una señal debe entenderse como la frecuencia dominante más baja que puede distinguirse del espectro de la señal. En el modelo de habla común, la altura tonal es la frecuencia de la señal de excitación modulada por la garganta humana. Si sólo estuviera presente una única frecuencia fundamental, el espectro sería extremadamente sencillo, comprendiendo únicamente la frecuencia fundamental y los armónicos superiores. Un espectro de este tipo podría codificarse de manera muy eficaz. Sin embargo, para señales con una altura tonal variable, la energía correspondiente a cada componente armónica se propaga por varios coeficientes de transformada llevando, por tanto, a una reducción de la eficacia de codificación.

20 Se podría intentar mejorar la eficacia de codificación para señales con una altura tonal variable creando en primer lugar una señal discreta en el tiempo con una altura tonal prácticamente constante. Para lograr esto, la tasa de muestreo podría variarse proporcionalmente a la altura tonal. Es decir, podría remuestrearse toda la señal antes de la aplicación de la transformada de modo que la altura tonal fuera lo más constante posible en toda la duración de la señal. Esto podría lograrse con un muestreo no equidistante, en el que los intervalos de muestreo son localmente adaptativos y se eligen de modo que la señal remuestreada, cuando se interpreta en términos de muestras equidistantes, tiene un contorno tonal más próximo a una altura tonal media común que la señal original. En este sentido, el contorno tonal debe entenderse que es la variación local de la altura tonal. La variación local podría parametrizarse, por ejemplo, en función de un número de muestra o tiempo.

25 De manera equivalente, esta operación podría considerarse como un reajuste a escala del eje de tiempo de una señal muestreada o de una continua antes de un muestreo equidistante. Una transformada de tiempo de este tipo también se conoce como alineamiento (*warping*). La aplicación de una transformada de frecuencia a una señal procesada previamente para llegar a una altura tonal casi constante, podría aproximar la eficacia de codificación a la eficacia alcanzable para una señal con una altura tonal generalmente constante.

30 El enfoque anterior, sin embargo, tiene varias desventajas. En primer lugar, una variación de la tasa de muestreo por un intervalo amplio, tal como se requiere por el procesamiento de toda la señal, podría llevar a una variación importante del ancho de banda de la señal debido al teorema de muestreo. En segundo lugar, cada bloque de coeficientes de transformada que representa un número fijo de muestras de entrada representaría entonces un segmento de tiempo de duración variable en la señal original. Esto haría que las aplicaciones con retardo de codificación limitado fueran casi imposibles y, además, daría como resultado dificultades en la sincronización.

35 Un método adicional se propone por los solicitantes de la solicitud de patente internacional 2007/051548. Los autores proponen un método para realizar el alineamiento por tramas. Sin embargo, esto se logra introduciendo limitaciones no deseadas a los contornos de alineamiento aplicables. Por tanto, existe la necesidad de enfoques alternativos para aumentar la eficacia de codificación, manteniendo al mismo tiempo una alta calidad de las señales de audio codificadas y decodificadas.

Sumario de la invención

45 Varias realizaciones de la presente invención permiten un aumento en la eficacia de codificación realizando una transformación local de la señal dentro de cada bloque de señal (trama de audio) con el fin de proporcionar una altura tonal (prácticamente) constante dentro de la duración de cada bloque de entrada que contribuye a un conjunto de coeficientes de transformada en una transformada basada en bloques. Un bloque de entrada de este tipo puede crearse, por ejemplo, por dos tramas consecutivas de una señal de audio cuando se utiliza una transformada de coseno discreta modificada como transformación de dominio de frecuencia.

50 Según varias realizaciones de la presente invención, un procesador de audio según la reivindicación 1 para generar una representación procesada de una señal de audio que tiene una secuencia de tramas comprende: un muestreador adaptado para muestrear la señal de audio dentro de una primera y una segunda trama de la secuencia de tramas, siguiendo la segunda trama a la primera trama, utilizando el muestreador información sobre un contorno tonal de la primera y la segunda trama para derivar una primera representación muestreada y para muestrear la señal de audio dentro de la segunda y una tercera trama, siguiendo la tercera trama a la segunda trama en la secuencia de tramas

- 5 utilizando la información sobre el contorno tonal de la segunda trama e información sobre un contorno tonal de la tercera trama para derivar una segunda representación muestreada; un calculador de ventanas de transformada adaptado para derivar una primera ventana de ajuste a escala para la primera representación muestreada y una segunda ventana de ajuste a escala para la segunda representación muestreada, dependiendo las ventanas de ajuste a escala del muestreo aplicado para derivar la primera representación muestreada o la segunda representación muestreada; y un divisor en ventanas adaptado para aplicar la primera ventana de ajuste a escala a la primera representación muestreada y la segunda ventana de ajuste a escala a la segunda representación muestreada para derivar una representación procesada de las tramas de audio primera, segunda y tercera de la señal de audio.
- 10 Según realizaciones adicionales el divisor en ventanas está adaptado para derivar una primera representación muestreada ajustada a escala aplicando la primera ventana de ajuste a escala a la primera representación muestreada y para derivar una segunda representación muestreada ajustada a escala aplicando la segunda ventana de ajuste a escala a la segunda representación ajustada a escala.
- 15 Según realizaciones adicionales el divisor en ventanas comprende además un transformador de dominio de frecuencia para derivar una primera representación de dominio de frecuencia de una primera representación remuestreada ajustada a escala y para derivar una segunda representación de dominio de frecuencia de una segunda representación remuestreada ajustada a escala.
- 20 Según realizaciones adicionales un procesador de audio comprende además un estimador de altura tonal adaptado para derivar el contorno tonal de las tramas primera, segunda y tercera.
- Según realizaciones adicionales un procesador de audio comprende además una interfaz de salida para proporcionar las representaciones de dominio de frecuencia primera y segunda y el contorno tonal de las tramas primera, segunda y tercera como una representación codificada de la segunda trama.
- Según una realización adicional, se define un procesador de audio según la reivindicación 11.
- 25 Según realizaciones adicionales de la presente invención un método según la reivindicación 13 para procesar una primera representación muestreada de una primera y una segunda trama de una señal de audio que tiene una secuencia de tramas en la que la segunda trama sigue a la primera trama y para procesar una segunda representación muestreada de la segunda trama y de una tercera trama de la señal de audio que sigue a la segunda trama en la secuencia de tramas, comprende: derivar una primera ventana de ajuste a escala para la primera representación muestreada utilizando información sobre un contorno tonal de la primera y la segunda trama y derivar una segunda ventana de ajuste a escala para la segunda representación muestreada utilizando información sobre un contorno tonal de la segunda y la tercera trama, en el que las ventanas de ajuste a escala se derivan de manera que tienen un número idéntico de muestras, en el que un primer número de muestras utilizadas para atenuar la primera ventana de ajuste a escala difiere de un segundo número de muestras utilizadas para intensificar la segunda ventana de ajuste a escala; aplicar la primera ventana de ajuste a escala a la primera representación muestreada y la segunda ventana de ajuste a escala a la segunda representación muestreada; y remuestrear la primera representación muestreada ajustada a escala para derivar una primera representación remuestreada utilizando la información sobre el contorno tonal de la primera y la segunda trama y remuestrear la segunda representación muestreada ajustada a escala para derivar una segunda representación remuestreada utilizando la información sobre el contorno tonal de la segunda y la tercera trama dependiendo el remuestreo de las ventanas de ajuste a escala derivadas.
- 30
- 35
- 40 Según una realización adicional se define un método según la reivindicación 13. Según una realización adicional se define un programa informático según la reivindicación 15.
- Según realizaciones adicionales el método comprende además: sumar la parte de la primera representación remuestreada correspondiente a la segunda trama y la parte de la segunda representación remuestreada correspondiente a la segunda trama para derivar una representación reconstruida de la segunda trama de la señal de audio.
- 45 Cuando se utiliza una transformada solapada modulada, como la transformada de coseno discreta modificada (MDCT), dos bloques sucesivos introducidos en la transformada de dominio de frecuencia se superponen con el fin de permitir una atenuación cruzada de la señal en los bordes del bloque, de modo que se suprimen los artefactos audibles del procesamiento por bloques. Un aumento en el número de coeficientes de transformada en comparación con una transformada sin superposición se evita mediante muestreo crítico. En MDCT, la aplicación de la transformada directa e inversa a un bloque de entrada, sin embargo, no lleva a su reconstrucción completa porque, por el muestreo crítico, se introducen artefactos en la señal reconstruida. La diferencia entre el bloque de entrada y la señal transformada de manera directa e inversa se denomina habitualmente "solapamiento de dominio de tiempo" (*time domain aliasing*). Mediante la superposición de los bloques reconstruidos en medio ancho de bloque después de la reconstrucción y sumando las muestras superpuestas, la señal de entrada puede, no obstante, reconstruirse perfectamente en el esquema MDCT. Según algunas realizaciones, esta propiedad de la transformada de coseno directa modificada puede mantenerse incluso cuando la señal subyacente se alinea en el tiempo por bloques (lo que es equivalente a la aplicación de tasas de muestreo localmente adaptivas).
- 50
- 55
- Como se describió anteriormente, el muestreo con tasas de muestreo localmente adaptivas (una tasa de muestreo

variable) puede considerarse como un muestreo uniforme en una escala de tiempo alineada. En este sentido, una compactación de la escala de tiempo antes del muestreo lleva a una tasa de muestreo menos eficaz, mientras que un estiramiento aumenta la tasa de muestreo eficaz de la señal subyacente.

5 Considerando una transformada de frecuencia u otra transformada, que utiliza superposición y suma en la reconstrucción con el fin de compensar posibles artefactos, aún funciona la cancelación de solapamiento de dominio de tiempo si se aplica el mismo alineamiento (corrección de altura tonal) en la región de superposición de dos bloques sucesivos. Así, la señal original puede reconstruirse después de invertir el alineamiento. Esto es por tanto cierto cuando se eligen diferentes tasas de muestreo local en los dos bloques de transformada superpuestos, porque el solapamiento de dominio de tiempo de la señal de tiempo continua correspondiente sigue anulándose, siempre que se cumpla el teorema de muestreo.

10 En algunas realizaciones, la tasa de muestreo después del alineamiento de tiempo de la señal dentro de cada bloque de transformada se selecciona individualmente para cada bloque. Esto tiene el efecto de que un número fijo de muestras aún representa un segmento de duración fija en la señal de entrada. Además, puede usarse un muestreador, que muestrea la señal de audio dentro de bloques de transformada superpuestos utilizando información sobre el contorno tonal de la señal de modo que la parte de señal superpuesta de una primera representación muestreada y de una segunda representación muestreada tiene un contorno tonal similar o idéntico en cada una de las representaciones muestreadas. El contorno tonal o la información sobre el contorno tonal utilizada para muestrear puede derivarse de manera arbitraria, siempre que existe una interrelación inequívoca entre la información sobre el contorno tonal (el contorno tonal) y la altura tonal de la señal. La información sobre el contorno tonal utilizada puede ser, por ejemplo, la altura tonal absoluta, la altura tonal relativa (el cambio de altura tonal), una fracción de la altura tonal absoluta o una función que depende de manera inequívoca de la altura tonal. Mediante la elección de la información sobre el contorno tonal como se indicó anteriormente, la parte de la primera representación muestreada correspondiente a la segunda trama tiene un contorno tonal similar al contorno tonal de la parte de la segunda representación muestreada correspondiente a la segunda trama. La similitud puede ser, por ejemplo, que los valores de altura tonal de partes de señal correspondientes tengan una relación más o menos constante, es decir, una relación dentro de un intervalo de tolerancia predeterminado. El muestreo puede realizarse por tanto de modo que la parte de la primera representación muestreada correspondiente a la segunda trama tenga un contorno tonal dentro de un intervalo de tolerancia predeterminado de un contorno tonal de la parte de la segunda representación muestreada correspondiente a la segunda trama.

15 20 25 Como la señal dentro de los bloques de transformada puede remuestrearse con diferentes frecuencias de muestreo o intervalos de muestreo, se crean bloques de entrada que pueden codificarse de manera eficaz mediante un algoritmo de codificación de transformada posterior. Esto puede lograrse mientras que, al mismo tiempo, se aplica la información derivada sobre el contorno tonal sin ninguna limitación adicional siempre que el contorno tonal sea continuo.

30 Aunque no se derive ningún cambio de altura tonal relativa dentro de un único bloque de entrada, el contorno tonal puede mantenerse constante dentro de y en los límites de los intervalos de señal o bloques de señal que no tienen un cambio de altura tonal que puede derivarse. Esto puede ser ventajoso cuando falla el seguimiento de altura tonal o es erróneo, lo que podría ser el caso de las señales complejas. Tampoco en este caso, un ajuste de altura tonal o remuestreo previo a la codificación de transformada proporciona artefactos adicionales.

35 40 El muestreo independiente dentro de los bloques de entrada puede lograrse utilizando ventanas de transformada especiales (ventanas de ajuste a escala) aplicadas antes de o durante la transformada de dominio de frecuencia. Según algunas realizaciones, estas ventanas de ajuste a escala dependen del contorno tonal de las tramas asociadas a los bloques de transformada. En términos generales, las ventanas de ajuste a escala dependen del muestreo aplicado para derivar la primera representación muestreada o la segunda representación muestreada. Es decir, la ventana de ajuste a escala de la primera representación muestreada puede depender del muestreo aplicado para derivar la primera ventana de ajuste a escala solamente, del muestreo aplicado para derivar la segunda ventana de ajuste a escala solamente o de ambos, el muestreo aplicado para derivar la primera ventana de ajuste a escala y el muestreo aplicado para derivar la segunda ventana de ajuste a escala. Lo mismo se aplica, *mutatis mutandis*, a la ventana de ajuste a escala para la segunda representación muestreada.

45 50 Esto proporciona la posibilidad de garantizar que no se superpongan más de dos bloques posteriores en ningún momento durante la reconstrucción de superposición y suma, de modo que es posible una cancelación de solapamiento de dominio de tiempo.

55 En particular, las ventanas de ajuste a escala de la transformada se crean, en algunas realizaciones, de modo que pueden tener diferentes formas dentro de cada una de las dos mitades de cada bloque de transformada. Esto es posible siempre que cada mitad de ventana cumpla con la condición de cancelación de solapamiento junto con la mitad de ventana del bloque contiguo dentro del intervalo de superposición común.

Como las tasas de muestreo de los dos bloques de superposición pueden ser diferentes (valores diferentes de las señales de audio subyacentes corresponden a muestras idénticas), ahora el mismo número de muestras puede corresponder a diferentes partes de la señal (formas de señal). Sin embargo, el requisito previo puede cumplirse reduciendo la longitud de transición (muestras) para un bloque con una tasa de muestreo menos eficaz que su bloque

- de superposición asociado. Dicho de otro modo, puede utilizarse un calculador de ventanas de transformada o un método para calcular ventanas de ajuste a escala, que proporcione ventanas de ajuste a escala con un número idéntico de muestras para cada bloque de entrada. Sin embargo, el número de muestras utilizadas para atenuar el primer bloque de entrada puede ser diferente del número de muestras utilizadas para intensificar el segundo bloque de entrada. Por tanto, la utilización de ventanas de ajuste a escala para las representaciones muestreadas de bloques de entrada de superposición (una primera representación muestreada y una segunda representación muestreada), que dependen del muestreo aplicado a los bloques de entrada, permite un muestreo diferente dentro de los bloques de entrada de superposición, conservando al mismo tiempo la capacidad de una reconstrucción de superposición y suma con cancelación de solapamiento de dominio de tiempo.
- 10 Resumiendo, el contorno tonal determinado de manera ideal puede utilizarse sin requerir modificaciones adicionales del contorno tonal mientras que, al mismo tiempo, se permite una representación de los bloques de entrada muestreados, que pueden codificarse de manera eficaz utilizando una transformada de dominio de frecuencia posterior.

Breve descripción de los dibujos

- 15 A continuación se describen varias realizaciones de la presente invención haciendo referencia a las figuras adjuntas, en las que:
- la figura 1 muestra una realización de un procesador de audio para generar una representación procesada de una señal de audio con una secuencia de tramas;
- las figuras 2a a 2d muestran un ejemplo del muestreo de una señal de entrada de audio dependiendo del contorno tonal de la señal de entrada de audio utilizando ventanas de ajuste a escala que dependen del muestreo aplicado;
- 20 la figura 3 muestra un ejemplo de cómo asociar las posiciones de muestreo utilizadas para el muestreo y las posiciones de muestreo de una señal de entrada con muestras equidistantes;
- la figura 4 muestra un ejemplo de un contorno de tiempo utilizado para determinar las posiciones de muestreo para el muestreo;
- la figura 5 muestra una realización de una ventana de ajuste a escala;
- 25 la figura 6 muestra un ejemplo de un contorno tonal asociado a una secuencia de tramas de audio que va a procesarse;
- la figura 7 muestra una ventana de ajuste a escala aplicada a un bloque de transformada muestreado;
- la figura 8 muestra las ventanas de ajuste a escala correspondientes al contorno tonal de la figura 6;
- la figura 9 muestra un ejemplo adicional de un contorno tonal de una secuencia de tramas de una señal de audio que va a procesarse;
- 30 la figura 10 muestra las ventanas de ajuste a escala utilizadas para el contorno tonal de la figura 9;
- la figura 11 muestra las ventanas de ajuste a escala de la figura 10 transformadas a la escala de tiempo lineal;
- la figura 11a muestra un ejemplo adicional de un contorno tonal de una secuencia de tramas;
- la figura 11b muestra las ventanas de ajuste a escala correspondientes a la figura 11a en una escala de tiempo lineal;
- la figura 12 muestra una realización de un método para generar una representación procesada de una señal de audio;
- 35 la figura 13 muestra una realización de un procesador para procesar representaciones muestreadas de una señal de audio compuesta por una secuencia de tramas de audio; y
- la figura 14 muestra una realización de un método para procesar representaciones muestreadas de una señal de audio.

Descripción detallada de realizaciones preferidas

- 40 La figura 1 muestra una realización de un procesador 10 de audio (señal de entrada) para generar una representación procesada de una señal de audio que tiene una secuencia de tramas. El procesador 2 de audio comprende un muestreador 4, que está adaptado para muestrear una señal 10 de audio (señal de entrada) introducida en el procesador 2 de audio para derivar los bloques de señal (representaciones muestreadas) utilizados como base para una transformada de dominio de frecuencia. El procesador 2 de audio comprende además un calculador 6 de ventanas de transformada adaptado para derivar ventanas de ajuste a escala para las representaciones muestreadas proporcionadas por el muestreador 4. Éstas se introducen en un divisor 8 en ventanas, que está adaptado para aplicar las ventanas de ajuste a escala a las representaciones muestreadas derivadas por el muestreador 4. En algunas realizaciones, el divisor en ventanas puede comprender adicionalmente un transformador 8a de dominio de frecuencia para derivar representaciones de dominio de frecuencia de las representaciones muestreadas ajustadas a escala. Éstas pueden procesarse a continuación o transmitirse adicionalmente como una representación codificada de la señal 10 de

audio. El procesador de audio utiliza además un contorno 12 tonal de la señal de audio, que puede proporcionarse al procesador de audio o que, según una realización adicional, puede derivarse por el procesador 2 de audio. El procesador 2 de audio puede comprender, por tanto, opcionalmente un estimador de altura tonal para derivar el contorno tonal.

5 El muestreador 4 puede funcionar en una señal de audio continua o, alternativamente, en una representación previamente muestreada de la señal de audio. En el último caso, el muestreador puede remuestrear la señal de audio proporcionada en su entrada tal como se indica en las figuras 2a a 2d. El muestreador está adaptado para muestrear bloques de audio de superposición contiguos de modo que la parte de superposición tenga el mismo contorno tonal o uno similar dentro de cada uno de los bloques de entrada después del muestreo.

10 El caso de una señal de audio previamente muestreada se elabora con más detalle en la descripción de las figuras 3 y 4.

15 El calculador 6 de ventanas de transformada deriva las ventanas de ajuste a escala para los bloques de audio dependiendo del remuestreo realizado por el muestreador 4. Para ello, un bloque 14 de ajuste de tasa de muestreo opcional puede estar presente con el fin de definir una regla de remuestreo utilizada por el muestreador, que entonces también se proporciona al calculador de ventanas de transformada. En una realización alternativa, el bloque 14 de ajuste de tasa de muestreo puede omitirse y el contorno 12 tonal puede proporcionarse directamente al calculador 6 de ventanas de transformada, que a su vez puede realizar los cálculos apropiados. Además, el muestreador 4 puede comunicar el muestreo aplicado al calculador 6 de ventanas de transformada con el fin de permitir el cálculo de ventanas de ajuste a escala apropiadas.

20 El remuestreo se realiza de modo que un contorno tonal de bloques de audio muestreados que muestrea el muestreador 4 es más constante que el contorno tonal de la señal original de audio dentro del bloque de entrada. Para ello, se evalúa el contorno tonal, tal como se indica para un ejemplo específico en las figuras 2a y 2d.

25 La figura 2a muestra un contorno tonal de disminución lineal como una función de los números de muestras de la señal de audio de entrada previamente muestreada. Es decir, las figuras 2a a 2d ilustran un escenario en el que las señales de audio de entrada ya se proporcionan como valores de muestra. No obstante, las señales de audio antes del remuestreo y después del remuestreo (alineamiento de la escala de tiempo) también se ilustran como señales continuas con el fin de ilustrar el concepto con más claridad. La figura 2b muestra un ejemplo de una señal 16 sinusoidal que tiene una frecuencia de barrido que disminuye de frecuencias más altas a frecuencias más bajas. Este comportamiento corresponde al contorno tonal de la figura 2a, que se muestra en unidades arbitrarias. De nuevo se destaca que el alineamiento de tiempo del eje de tiempo es equivalente a un remuestreo de la señal con intervalos de muestreo localmente adaptivos.

30 Con el fin de ilustrar el procesamiento de superposición y suma, la figura 2b muestra tres tramas 20a, 20b y 20c consecutivas de la señal de audio, que se procesan por bloques con una superposición de una trama (trama 20b). Es decir, se procesa y remuestrea un primer bloque 22 de señal (bloque 1 de señal) que comprende las muestras de la primera trama 20a y la segunda trama 20b y de manera independiente se remuestrea un segundo bloque 24 de señal que comprende las muestras de la segunda trama 20b y la tercera trama 20c. El primer bloque 22 de señal se remuestrea para derivar la primera representación 26 remuestreada mostrada en la figura 2c y el segundo bloque 24 de señal se remuestrea para dar la segunda representación 28 remuestreada mostrada en la figura 2d. Sin embargo, el muestreo se realiza de modo que las partes correspondientes a la trama 20b de superposición tienen el mismo contorno tonal o sólo uno que se desvía ligeramente (idéntico dentro de un intervalo de tolerancia predeterminado) en la primera representación 26 muestreada y la segunda representación 28 muestreada. Esto, evidentemente, sólo es cierto cuando se estima la altura tonal en términos de números de muestra. El primer bloque 22 de señal se remuestrea para dar la primera representación 26 remuestreada, que tiene una altura tonal constante (idealizada). Por tanto, utilizando valores de muestra de la representación 26 remuestreada como entrada para una transformada de dominio de frecuencia, de manera ideal sólo se derivaría un único coeficiente de frecuencia. Esto es evidentemente una representación extremadamente eficaz de la señal de audio. A continuación se comentarán detalles en cuanto a cómo se realiza el remuestreo, haciendo referencia a las figuras 3 y 4. Como es evidente por la figura 2c, el remuestreo se realiza de modo que el eje de las posiciones de muestra (el eje x), que corresponde al eje de tiempo en una representación muestreada de manera equidistante se modifica de modo que la forma de la señal resultante tiene sólo una única frecuencia tonal. Esto corresponde a un alineamiento de tiempo del eje de tiempo y a un muestreo equidistante posterior de la representación alineada en el tiempo de la señal del primer bloque 22 de señal.

35 El segundo bloque 24 de señal se remuestrea de modo que la parte de señal correspondiente a la trama 20b de superposición en la segunda representación 28 remuestreada tiene un contorno tonal idéntico o sólo uno que se desvía ligeramente con respecto a la parte de señal correspondiente de la representación 26 remuestreada. Sin embargo, las tasas de muestreo difieren. Es decir, las formas de señal idénticas dentro de las representaciones remuestreadas se representan mediante diferentes números de muestras. No obstante, cada representación remuestreada, cuando se codifica mediante un codificador de transformada, da como resultado una representación codificada de manera altamente eficaz que tiene sólo un número limitado de coeficientes de frecuencia diferentes de cero.

Debido al remuestreo, las partes de señal de la primera mitad del bloque 22 de señal se desplazan a muestras

pertencientes a la segunda mitad del bloque de señal de la representación remuestreada, tal como se indica en la figura 2c. En particular, el área 30 sombreada y la señal correspondiente a la derecha del segundo pico (indicado por II) se desplaza a la mitad derecha de la representación 26 remuestreada y, por tanto, se representa por la segunda mitad de las muestras de la representación 26 remuestreada. Sin embargo, estas muestras no tienen una parte de señal correspondiente en la mitad izquierda de la representación 28 remuestreada de la figura 2d.

Dicho de otro modo, durante el remuestreo, se determina la tasa de muestreo para cada bloque MDCT de modo que la tasa de muestreo lleva a una duración constante en un tiempo lineal del centro del bloque, que contiene N muestras en el caso de una resolución de frecuencia de N y una longitud de ventana máxima de 2N. En el ejemplo descrito anteriormente de las figuras 2a a 2d, $N = 1024$ y, por consiguiente, $2N = 2048$ muestras. El remuestreo realiza la interpolación de señal real en las posiciones requeridas. Debido a la superposición de dos bloques, que pueden tener diferentes tasas de muestreo, el remuestreo tiene que realizarse dos veces para cada segmento de tiempo (que es igual a una de las tramas 20a a 20c) de la señal de entrada. Puede usarse el mismo contorno tonal, que controla el codificador o el procesador de audio que realiza la codificación, para controlar el procesamiento necesario para invertir la transformada y el alineamiento, como puede implementarse dentro de un decodificador de audio. En algunas realizaciones, el contorno tonal se transmite, por tanto, como información secundaria. Con el fin de evitar un desajuste entre un codificador y un decodificador correspondiente, algunas realizaciones de codificadores utilizan el contorno tonal codificado y, posteriormente, decodificado en lugar del contorno tonal tal como se introdujo o derivó originalmente. Sin embargo, el contorno tonal derivado o introducido puede usarse directamente de manera alternativa.

Con el fin de garantizar que sólo partes de señal correspondientes se superponen en la reconstrucción de superposición y suma, se derivan ventanas de ajuste a escala apropiadas. Estas ventanas de ajuste a escala tienen que explicar el efecto de que diferentes partes de señal de las señales originales se representen dentro de las mitades de ventana correspondientes de las representaciones remuestreadas, como se produce por el remuestreo descrito anteriormente.

Pueden derivarse ventanas de ajuste a escala apropiadas para las señales que van a codificarse, que dependen del muestreo o remuestreo aplicado para derivar las representaciones 26 y 28 remuestreadas primera y segunda. Para el ejemplo de la señal original ilustrada en la figura 2b y el contorno tonal ilustrado en la figura 2a, se proporcionan ventanas de ajuste a escala apropiadas para la segunda mitad de ventana de la primera representación 26 remuestreada y para la primera mitad de ventana de la segunda representación 28 remuestreada mediante la primera ventana 32 de ajuste a escala (su segunda mitad) y mediante la segunda ventana 34 de ajuste a escala, respectivamente (la mitad izquierda de la ventana correspondiente a las primeras 1024 muestras de la segunda representación 28 remuestreada).

Como la parte de señal dentro del área 30 sombreada de la primera representación 26 remuestreada no tiene parte de señal correspondiente en la primera mitad de ventana de la segunda representación 28 remuestreada, la parte de señal dentro del área sombreada tiene que reconstruirse por completo mediante la primera representación 26 remuestreada. En una reconstrucción MDCT, esto puede conseguirse cuando no se utilizan las muestras correspondientes para intensificación o atenuación, es decir, cuando las muestras reciben un factor de ajuste a escala de 1. Por tanto, las muestras de la ventana 32 de ajuste a escala correspondientes al área 30 sombreada, se establecen en una unidad. Al mismo tiempo, el mismo número de muestras deben establecerse en 0 al final de la ventana de ajuste a escala con el fin de evitar que se mezclen esas muestras con las muestras de la primera área 30 sombreada debido a las propiedades inherentes de transformada MDCT y transformada inversa.

Debido al remuestreo (aplicado), que logra un alineamiento de tiempo idéntico del segmento de ventana de superposición, las muestras de la segunda área 36 sombreada tampoco tienen una parte complementaria de señal dentro de la primera mitad de ventana de la segunda representación 28 remuestreada. Por tanto, esta parte de señal puede reconstruirse por completo por la segunda mitad de ventana de la segunda representación 28 remuestreada. Por tanto, es viable establecer las muestras de la primera ventana de ajuste a escala correspondientes a la segunda área 36 sombreada en 0 sin perder información sobre la señal que va a reconstruirse. Cada parte de señal presente dentro de la primera mitad de ventana de la segunda representación 28 remuestreada tiene una parte complementaria correspondiente dentro de la segunda mitad de ventana de la primera representación 26 remuestreada. Por tanto, todas las muestras dentro de la primera mitad de ventana de la segunda representación 28 remuestreada se utilizan para la atenuación cruzada entre las representaciones 26 y 28 remuestreadas primera y segunda, tal como se indica por la forma de la segunda ventana 34 de ajuste a escala.

Resumiendo, el remuestreo dependiente de la altura tonal y la utilización de ventanas de ajuste a escala diseñadas de manera apropiada permite aplicar un contorno tonal óptimo, que no necesita cumplir con limitaciones excepto de ser continuo. Como, para aumentar la eficacia de codificación, sólo son relevantes los cambios de altura tonal relativa, el contorno tonal puede mantenerse constante dentro de y en los límites de intervalos de señal en los que no puede estimarse una altura tonal distinta o en los que no está presente una variación de altura tonal. Algunos conceptos alternativos proponen implementar alineamiento de tiempo con contornos tonales especializados o funciones de alineamiento de tiempo, que tienen restricciones especiales con respecto a sus contornos. Mediante la utilización de las realizaciones de la invención, la eficacia de codificación será mayor, dado que en cualquier momento puede utilizarse el contorno tonal óptimo.

Con respecto a las figuras 3 a 5, a continuación se describirá con más detalle una posibilidad particular para realizar el remuestreo y para derivar las ventanas de ajuste a escala asociadas.

El muestreo se basa, de nuevo, en un contorno 50 tonal de disminución lineal, correspondiente a un número predeterminado de muestras N. La señal 52 correspondiente se ilustra en un tiempo normalizado. En el ejemplo elegido, la señal tiene una duración de 10 milisegundos. Si se procesa una señal previamente muestreada, la señal 52 se muestrea normalmente en intervalos de muestreo equidistantes, tal como se indica por las marcas del eje 54 de tiempo. Si se aplicara alineamiento de tiempo mediante la transformación apropiada del eje 54 de tiempo, la señal 52, en una escala 56 de tiempo alineada, se convertiría en una señal 58, que tiene una altura tonal constante. Es decir, la diferencia de tiempo (la diferencia de números de muestras) entre máximos contiguos de la señal 58 es igual en la nueva escala 56 de tiempo. La longitud de la trama de señal también cambiaría a una nueva longitud de x milisegundos, dependiendo del alineamiento aplicado. Debe indicarse que el dibujo del alineamiento de tiempo sólo se utiliza para visualizar la idea del remuestreo no equidistante utilizado en varias realizaciones de la presente invención que, de hecho, puede implementarse sólo utilizando los valores del contorno 50 tonal.

La siguiente realización, que describe cómo puede realizarse el muestreo se basa, para facilitar la comprensión, en el supuesto de que la altura tonal objetivo a la que se alineará la señal (una altura tonal derivada de la representación remuestreada o muestreada de la señal original) es una unidad. Sin embargo, evidentemente, las siguientes consideraciones pueden aplicarse fácilmente a alturas tonales objetivo arbitrarias de los segmentos de señal procesados.

Suponiendo que el alineamiento de tiempo se aplica en una trama j que comienza en la muestra jN de manera que hace que la altura tonal sea una unidad (1), la duración de trama tras el alineamiento de tiempo correspondería a la suma de las N muestras correspondientes del contorno tonal:

$$D_j = \sum_{i=0}^{N-1} \text{contorno_tonal}_{jN+xi}$$

Es decir, la duración de la señal 58 alineada en el tiempo (el tiempo t' = x en la figura 3) se determina por la fórmula anterior.

Con el fin de obtener N muestras alineadas, el intervalo de muestreo en la trama alineada en el tiempo j es igual a:

$$I_j = N / D_j$$

De manera iterativa, puede construirse un contorno de tiempo, que asocie las posiciones de las muestras originales en relación con la ventana MDCT alineada según:

$$\text{contorno_tiempo}_{i+1} = \text{contorno_tiempo}_i + \text{contorno_tonal}_{jN+i} * I_j$$

Un ejemplo de un contorno de tiempo se proporciona en la figura 4. El eje x muestra el número de muestra de la representación remuestreada y el eje y da la posición de este número de muestreo en unidades de muestras de la representación original. En el ejemplo de la figura 3, el contorno de tiempo se construye, por tanto, con un tamaño de etapa siempre en disminución. La posición de muestra asociada al número de muestra 1 en la representación alineada en el tiempo (eje n') en unidades de las muestras originales es, por ejemplo, de aproximadamente 2. Para el remuestreo no equidistante, dependiente del contorno tonal, se requiere que las posiciones de las muestras de entrada MDCT alineadas sean unidades de la escala de tiempo no alineada original. La posición de la muestra de entrada MDCT alineada i (eje y) puede obtenerse buscando un par de posiciones de muestra originales k y k+1, que definen un intervalo que incluye i:

$$\text{contorno_tiempo}_k \leq i < \text{contorno_tiempo}_{k+1} .$$

Por ejemplo, la muestra i=1 está ubicada en el intervalo definido por la muestra k=0, k+1=1. Una parte fraccionaria u de la posición de muestra se obtiene suponiendo un contorno de tiempo lineal entre k=1 y k+1=1 (eje x). En términos generales, la parte fraccionaria (u) de la muestra i se determina por:

$$u = \frac{\text{contorno_tiempo}_i - \text{contorno_tiempo}_k}{\text{contorno_tiempo}_{k+1} - \text{contorno_tiempo}_k} .$$

Por tanto, la posición de muestreo para el remuestreo no equidistante de la señal 52 original puede derivarse en unidades de posiciones de muestreo originales. Por tanto, la señal puede remuestrearse de modo que los valores remuestreados correspondan a una señal alineada en el tiempo. Este remuestreo puede implementarse, por ejemplo, utilizando un filtro de interpolación polifásico h dividido en P subfiltros h_p con una precisión de 1/P intervalos de muestras originales. Para este fin, puede obtenerse el índice de subfiltro a partir de la posición de muestra fraccionaria:

$$p = \lfloor uP \rfloor,$$

y la muestra de entrada MDCT alineada xw puede calcularse entonces mediante convolución:

$$xw_i = x_k * h_{p,k}.$$

5 Evidentemente, pueden utilizarse otros métodos de remuestreo, tales como, por ejemplo, remuestreo basado en ranuras, interpolación lineal, interpolación cuadrática u otros métodos de remuestreo.

Después de haber derivado las representaciones remuestreadas, se derivan ventanas de ajuste a escala apropiadas de modo que ninguna de las dos ventanas de superposición oscila más de $N/2$ muestras en el área central de la trama MDCT contigua. Como se describió anteriormente, esto puede lograrse utilizando el contorno tonal o los intervalos de muestra correspondientes l_j o, de manera equivalente, las duraciones de trama D_j . La longitud de una superposición "izquierda" de la trama j (es decir la intensificación con respecto a la trama precedente $j-1$) se determina por:

$$\alpha_j = \begin{cases} N/2 & \text{si } D_j \leq D_{j-1} \\ N/2 * D_{j-1} / D_j & \text{si no} \end{cases},$$

y la longitud de la superposición "derecha" de la trama j (es decir la atenuación a la trama posterior $j+1$) se determina por:

$$\sigma_j = \begin{cases} N/2 & \text{si } D_j \leq D_{j+1} \\ N/2 * D_{j+1} / D_j & \text{si no} \end{cases}.$$

15 Por tanto, una ventana resultante para la trama j de longitud $2N$, es decir la longitud de ventana MDCT típica utilizada para el remuestreo de tramas con N muestras (es decir, una resolución de frecuencia de N), consiste en los siguientes segmentos, tal como se ilustra en la figura 5.

$0 \leq i < N/2 - \sigma l_j$	0
$N/2 - \sigma l_j \leq i < N/2 + \sigma l_j$	$W_l(i)$
$N/2 + \sigma l_j \leq i < 3N/2 + \sigma r_j$	1
$3N/2 - \sigma r_j \leq i < 3N/2 + \sigma r_j$	$w_r(i)$
$3N/2 + \sigma r_j \leq i < 2N$	0

20 Es decir, las muestras 0 a $N/2 - \sigma l$ del bloque de entrada j son 0 cuando D_{j+1} es mayor que o igual a D_j . Las muestras en el intervalo $[N/2 - \sigma l; N/2 + \sigma l]$ se utilizan para intensificar la ventana de ajuste a escala. Las muestras en el intervalo $[N/2 + \sigma l; N]$ se establecen en una unidad. La mitad de ventana derecha, es decir la mitad de ventana utilizada para atenuar las $2N$ muestras comprende un intervalo $[N; 3/2N - \sigma r]$, que se establece en una unidad. Las muestras utilizadas para atenuar la ventana están contenidas dentro del intervalo $[3/2N - \sigma r; 3/2N + \sigma r]$. Las muestras en el intervalo $[3/2N + \sigma r; 2N]$ se establecen en 0. En términos generales, se derivan las ventanas de ajuste a escala, que tienen números idénticos de muestras, en las que un primer número de muestras utilizadas para atenuar la ventana de ajuste a escala difiere de un segundo número de muestras utilizadas para intensificar la ventana de ajuste a escala.

25 La forma precisa o los valores de muestra correspondientes a las ventanas de ajuste a escala derivadas pueden obtenerse, por ejemplo, (también para una longitud de superposición no de número entero) a partir de una interpolación lineal de mitades de ventana prototipo, que especifican la función ventana en posiciones de muestra de número entero (o en una rejilla fija con una resolución temporal incluso mayor). Es decir, las ventanas prototipo se ajustan a escala en el tiempo para dar las longitudes de intensificación y atenuación requeridas de $2\sigma l_j$ o $2\sigma r_j$, respectivamente.

30 Según una realización adicional de la presente invención, la parte de ventana de atenuación puede determinarse sin utilizar información sobre el contorno tonal de la tercera trama. Para ello, el valor de D_{j+1} puede limitarse a un límite predeterminado. En algunas realizaciones, el valor puede establecerse en un número predeterminado fijo y la parte de ventana de intensificación del segundo bloque de entrada puede calcularse basándose en el muestreo aplicado para derivar la primera representación muestreada, la segunda representación muestreada y el número predeterminado o el límite predeterminado para D_{j+1} . Esto puede utilizarse en aplicaciones en las que los tiempos de retardo bajos son de gran importancia, porque cada bloque de entrada puede procesarse sin conocimiento del bloque posterior.

En una realización adicional de la presente invención, la longitud variable de las ventanas de ajuste a escala puede utilizarse para cambiar entre bloques de entrada de diferente longitud.

5 Las figuras 6 a 8 ilustran un ejemplo que tiene una resolución de frecuencia de $N=1024$ y una altura tonal de disminución lineal. La figura 6 muestra la altura tonal como una función del número de muestra. Como es evidente, la disminución de la altura tonal es lineal y oscila entre 3500 Hz y 2500 Hz en el centro del bloque 1 MDCT (bloque 100 de transformada), entre 2500 Hz y 1500 Hz en el centro del bloque 2 MDCT (bloque 102 de transformada) y entre 1500 Hz y 500 Hz en el centro del bloque 3 MDCT (bloque 104 de transformada). Esto corresponde a las siguientes duraciones de trama en la escala de tiempo alineada (dado en unidades de duración (D_2) de bloque 102 de transformada):

$$D_1 = 1,5D_2; D_3 = 0,5D_2.$$

10 Dado lo anterior, el segundo bloque 102 de transformada tiene una longitud de superposición izquierda $\sigma_2 = N/2 = 512$, porque $D_2 < D_1$ y una longitud de superposición derecha $\sigma_2 = N/2 \times 0,5 = 256$. La figura 7 muestra la ventana de ajuste a escala calculada que tiene las propiedades descritas anteriormente.

15 Además, la longitud de superposición derecha del bloque 1 es igual a $\sigma_1 = N/2 \times 2/3 = 341,33$ y la longitud de superposición izquierda del bloque 3 (bloque 104 de transformada) es $\sigma_3 = N/2 = 512$. Como es evidente, la forma de las ventanas de transformada sólo depende del contorno tonal de la señal subyacente. La figura 8 muestra las ventanas eficaces en el dominio de tiempo no alineado (es decir lineal) para los bloques 100, 102 y 104 de transformada.

20 Las figuras 9 a 11 muestran un ejemplo adicional para una secuencia de cuatro bloques 110 a 113 de transformada consecutivos. Sin embargo, el contorno tonal tal como se indica en la figura 9 es ligeramente más complejo, teniendo la forma de una función sinusoidal. Para la resolución de frecuencia a modo de ejemplo $N(1024)$ y una longitud de ventana máxima de 2048, las funciones ventana adaptadas de manera correspondiente (calculadas) en el dominio de tiempo alineado se dan en la figura 10. Sus formas eficaces correspondientes en una escala de tiempo lineal se ilustran en la figura 11. Puede observarse que todas las figuras muestran funciones ventana cuadradas con el fin de ilustrar las capacidades de reconstrucción del procedimiento de superposición y suma mejor cuando las ventanas se aplican dos veces (antes de MDCT y después de IMDCT). La propiedad de cancelación de solapamiento de dominio de tiempo de las ventanas generadas puede reconocerse a partir de las simetrías de transiciones correspondientes en el dominio alineado. Como se determinó anteriormente, las figuras también ilustran que pueden seleccionarse intervalos de transición más cortos en bloques en los que la altura tonal disminuye hacia los límites, correspondiendo a intervalos de muestreo en aumento y, por tanto, a formas eficaces estiradas en el dominio de tiempo lineal. Un ejemplo de este comportamiento puede observarse en la trama 4 (bloque 113 de transformada), en el que la función ventana abarca menos de las 2048 muestras máximas. Sin embargo, debido a los intervalos de muestreo, que son inversamente proporcionales a la altura tonal de la señal, la duración posible máxima se cubre con la limitación de que sólo dos ventanas sucesivas pueden superponerse en cualquier punto del tiempo.

Las figuras 11a y 11b proporcionan un ejemplo adicional de un contorno tonal (información de contorno tonal) y sus ventanas de ajuste a escala correspondientes en una escala de tiempo lineal.

35 La figura 11a da el contorno 120 tonal, como una función de números de muestra, que se indican en el eje x. Es decir, la figura 11a da información de contorno de alineamiento para tres bloques 122, 124 y 126 de transformación consecutivos.

40 La figura 11b ilustra las ventanas de ajuste a escala correspondientes para cada uno de los bloques 122, 124 y 126 de transformada en una escala de tiempo lineal. Las ventanas de transformada se calculan dependiendo del muestreo aplicado a la señal correspondiente a la información de contorno tonal ilustrada en la figura 11a. Estas ventanas de transformada vuelven a transformarse en la escala de tiempo lineal, con el fin de proporcionar la ilustración de la figura 11b.

45 Dicho de otro modo, la figura 11b ilustra que las ventanas de ajuste a escala que han vuelto a transformarse pueden superar el borde de trama (líneas continuas de la figura 11b) cuando se vuelven a alinear o a transformar a la escala de tiempo lineal. Esto puede considerarse en el codificador proporcionando algunas muestras de entrada más, más allá de los bordes de trama. En el decodificador, la memoria intermedia de salida puede ser lo suficientemente grande como para almacenar las muestras correspondientes. Una forma alternativa de considerar esto puede ser acortar el intervalo de superposición de la ventana y utilizar regiones de ceros y unos en su lugar, de modo que la parte diferente de cero de la ventana no supere el borde de trama.

50 Como además se hace evidente a partir de la figura 11b, las intersecciones de las ventanas que se han vuelto a alinear (los puntos de simetría para el solapamiento de dominio de tiempo) no se ven alteradas por el alineamiento de tiempo, puesto que permanecen en las posiciones 512, 3×512 , 5×512 , 7×512 "no alineadas". Éste también es el caso para las ventanas de ajuste a escala correspondientes en el dominio alineado, porque también son simétricas para posiciones dadas por un cuarto y tres cuartos de la longitud de bloque de transformada.

55 Una realización de un método para generar una representación procesada de una señal de audio que tiene una secuencia de tramas puede estar caracterizada por las etapas ilustradas en la figura 12.

5 En una etapa 200 de muestreo, la señal de audio se muestrea dentro de una primera y una segunda trama de la secuencia de tramas, siguiendo la segunda trama a la primera trama, utilizando información sobre un contorno tonal de la primera y la segunda trama para derivar una primera representación muestreada y la señal de audio se muestrea dentro de la segunda y una tercera trama, siguiendo la tercera trama a la segunda trama en la secuencia de tramas, utilizando información sobre el contorno tonal de la segunda trama e información sobre un contorno tonal de la tercera trama para derivar una segunda representación muestreada.

10 En una etapa 202 de cálculo de ventana de transformada, se deriva la primera ventana de ajuste a escala para la primera representación muestreada y se deriva la segunda ventana de ajuste a escala para la segunda representación muestreada, en la que las ventanas de ajuste a escala dependen del muestreo aplicado para derivar las representaciones muestreadas primera y segunda.

En una etapa 204 de división en ventanas, la primera ventana de ajuste a escala se aplica a la primera representación muestreada y la segunda ventana de ajuste a escala se aplica a la segunda representación muestreada.

15 La figura 13 muestra una realización de un procesador 290 de audio para procesar una primera representación muestreada de una primera y una segunda trama de una señal de audio que tiene una secuencia de tramas en la que la segunda trama sigue a la primera trama y para procesar adicionalmente una segunda representación muestreada de la segunda trama y de una tercera trama que sigue a la segunda trama en la secuencia de tramas, que comprende:

20 un calculador 300 de ventanas de transformada adaptado para derivar una primera ventana de ajuste a escala para la primera representación 301a muestreada utilizando información sobre un contorno 302 tonal de la primera y la segunda trama y para derivar una segunda ventana de ajuste a escala para la segunda representación 301b muestreada utilizando información sobre un contorno tonal de la segunda y la tercera trama, en el que las ventanas de ajuste a escala tienen números idénticos de muestras y en el que un primer número de muestras utilizadas para atenuar la primera ventana de ajuste a escala difiere de un segundo número de muestras utilizadas para intensificar la segunda ventana de ajuste a escala;

25 el procesador 290 de audio comprende además un divisor 306 en ventanas adaptado para aplicar la primera ventana de ajuste a escala a la primera representación muestreada y para aplicar la segunda ventana de ajuste a escala a la segunda representación muestreada. El procesador 290 de audio comprende además un remuestreador 308 adaptado para remuestrear la primera representación muestreada ajustada a escala para derivar una primera representación remuestreada utilizando la información sobre el contorno tonal de la primera y la segunda trama y para remuestrear la segunda representación muestreada ajustada a escala para derivar una segunda representación remuestreada, utilizando la información sobre el contorno tonal de la segunda y la tercera trama de modo que una parte de la primera representación remuestreada correspondiente a la segunda trama tiene un contorno tonal dentro de un intervalo de tolerancia predeterminado de un contorno tonal de la parte de la segunda representación remuestreada correspondiente a la segunda trama. Para derivar la ventana de ajuste a escala, el calculador 300 de ventanas de transformada puede o bien recibir el contorno 302 tonal directamente o bien recibir información del remuestreo desde un ajustador 310 de tasa de muestra opcional, que recibe el contorno 302 tonal y que deriva una estrategia de remuestreo.

30 En una realización adicional de la presente invención, un procesador de audio comprende además un sumador 320 opcional, que está adaptado para sumar la parte de la primera representación remuestreada correspondiente a la segunda trama y la parte de la segunda representación remuestreada correspondiente a la segunda trama para derivar una representación reconstruida de la segunda trama de la señal de audio como señal 322 de salida. La primera representación muestreada y la segunda representación muestreada podrían proporcionarse, en una realización, como salida al procesador 290 de audio. En una realización adicional, el procesador de audio puede comprender, opcionalmente, un transformador 330 de dominio de frecuencia inversa, que puede derivar las representaciones muestreadas primera y segunda de representaciones de dominio de frecuencia de las representaciones muestreadas primera y segunda proporcionadas a la entrada del transformador 330 de dominio de frecuencia inversa.

45 La figura 14 muestra una realización de un método para procesar una primera representación muestreada de una primera y una segunda trama de una señal de audio que tiene una secuencia de tramas en la que la segunda trama sigue a la primera trama y para procesar una segunda representación muestreada de la segunda trama y de una tercera trama que sigue a la segunda trama en la secuencia de tramas. En una etapa 400 de creación de ventanas, se deriva una primera ventana de ajuste a escala para la primera representación muestreada utilizando información sobre un contorno tonal de la primera y la segunda trama y se deriva una segunda ventana de ajuste a escala para la segunda representación muestreada utilizando información sobre un contorno tonal de la segunda y la tercera trama, en el que las ventanas de ajuste a escala tienen números idénticos de muestras y en el que un primer número de muestras utilizadas para atenuar la primera ventana de ajuste a escala difiere de un segundo número de muestras utilizadas para intensificar la segunda ventana de ajuste a escala.

55 En una etapa 402 de ajuste a escala, la primera ventana de ajuste a escala se aplica a la primera representación muestreada y la segunda ventana de ajuste a escala se aplica a la segunda representación muestreada.

En una operación 402 de remuestreo, la primera representación muestreada ajustada a escala se remuestrea para derivar una primera representación remuestreada utilizando la información sobre el contorno tonal de las tramas primera

5 y segunda y la segunda representación muestreada ajustada a escala se remuestrea para derivar una segunda representación remuestreada utilizando la información sobre el contorno tonal de las tramas segunda y tercera de modo que una parte de la primera representación remuestreada correspondiente a la primera trama tiene un contorno tonal dentro de un intervalo de tolerancia predeterminado de un contorno tonal de la parte de la segunda representación remuestreada correspondiente a la segunda trama.

Según una realización adicional de la invención, el método comprende una etapa 406 de síntesis opcional en la que la parte de la primera representación remuestreada correspondiente a la segunda trama y la parte de la segunda representación remuestreada correspondiente a la segunda trama se combinan para derivar una representación reconstruida de la segunda trama de la señal de audio.

10 Resumiendo, las realizaciones comentadas anteriormente de la presente invención permiten aplicar un contorno tonal óptimo a una señal de audio continua o previamente muestreada para remuestrear o transformar la señal de audio en una representación, que puede codificarse dando como resultado una representación codificada con alta calidad y una tasa de transmisión de bits baja. Para ello, puede codificarse la señal remuestreada utilizando una transformada de dominio de frecuencia. Ésta podría ser, por ejemplo, la transformada de coseno discreta modificada en las realizaciones anteriores. Sin embargo, alternativamente podrían usarse otras transformadas de dominio de frecuencia u otras transformadas para derivar una representación codificada de una señal de audio con una tasa de transmisión de bits baja.

15 No obstante, también es posible utilizar diferentes transformadas de frecuencia para lograr el mismo resultado tal como, por ejemplo, una transformada rápida de Fourier o una transformada de coseno discreta para derivar la representación codificada de la señal de audio.

Evidentemente el número de muestras, es decir los bloques de transformada utilizados como entrada para la transformada de dominio de frecuencia no está limitado al ejemplo particular utilizado en las realizaciones descritas anteriormente. En su lugar, puede utilizarse una longitud de trama de bloque arbitraria tal como, por ejemplo, bloques constituidos por 256, 512, 1024 bloques.

25 Pueden utilizarse técnicas arbitrarias para muestrear o para remuestrear las señales de audio para la implementación en realizaciones adicionales de la presente invención.

30 Un procesador de audio utilizado para generar la representación procesada puede recibir, tal como se ilustra en la figura 1, la señal de audio y la información sobre el contorno tonal como entradas separadas, por ejemplo, como flujos de bits de entrada separados. En realizaciones adicionales, sin embargo, la señal de audio y la información sobre el contorno tonal pueden proporcionarse dentro de un flujo de bits entrelazado, de modo que la información de la señal de audio y el contorno tonal se multiplexen mediante el procesador de audio. Las mismas configuraciones pueden implementarse para el procesador de audio que deriva una reconstrucción de la señal de audio basándose en las representaciones muestreadas. Es decir, las representaciones muestreadas pueden introducirse como un flujo de bits unido con la información de contorno tonal o como dos flujos de bits separados. El procesador de audio podría comprender además un transformador de dominio de frecuencia con el fin de transformar las representaciones remuestreadas en coeficientes de transformada, que entonces se transmiten junto con un contorno tonal como una representación codificada de la señal de audio, como para transmitir de manera eficaz una señal de audio codificada a un decodificador correspondiente.

35 Las realizaciones descritas anteriormente, por motivos de simplicidad, suponen que la altura tonal objetivo a la que se remuestrea la señal es una unidad. Evidentemente la altura tonal puede ser cualquier otra altura tonal arbitraria. Como la altura tonal puede aplicarse sin limitaciones del contorno tonal, es además posible aplicar un contorno tonal constante en caso de que no pueda derivarse un contorno tonal o en caso de que no se proporcione un contorno tonal.

40 Dependiendo de determinados requisitos de implementación de los métodos inventivos, los métodos inventivos pueden implementarse en hardware o en software. La implementación puede realizarse utilizando un medio de almacenamiento digital, en particular un disco, DVD o un CD que tenga señales de control electrónicamente legibles almacenadas en los mismos, que actúen conjuntamente con un sistema informático programable de modo que se realicen los métodos inventivos. En general, la presente invención es, por tanto, un producto de programa informático con un código de programa almacenado en un soporte legible por máquina, siendo el código de programa operativo para realizar los métodos inventivos cuando el producto de programa informático se ejecuta en un ordenador. Dicho de otro modo, los métodos inventivos son, por tanto, un programa informático que tiene un código de programa para realizar al menos uno de los métodos inventivos cuando el programa informático se ejecuta en un ordenador.

45 Aunque lo anterior se ha mostrado y descrito particularmente con referencia a realizaciones particulares del mismo, los expertos en la técnica entenderán que pueden realizarse diversos cambios en la forma y los detalles sin apartarse del alcance definido por las reivindicaciones adjuntas.

55

REIVINDICACIONES

1. Procesador de audio para generar una representación procesada de una señal de audio que tiene una secuencia de tramas, comprendiendo el procesador de audio:

5

un muestreador adaptado para muestrear la señal de audio dentro de una primera y una segunda trama de la secuencia de tramas, siguiendo la segunda trama a la primera trama, utilizando el muestreador información sobre un contorno tonal de la primera y la segunda trama para derivar una primera representación muestreada y para muestrear la señal de audio dentro de la segunda y una tercera trama, siguiendo la tercera trama a la segunda trama en la secuencia de tramas utilizando la información sobre el contorno tonal de la segunda trama e información sobre un contorno tonal de la tercera trama para derivar una segunda representación muestreada;

10

un calculador de ventanas de transformada adaptado para derivar una primera ventana de ajuste a escala para la primera representación muestreada y una segunda ventana de ajuste a escala para la segunda representación muestreada, dependiendo las ventanas de ajuste a escala del muestreo aplicado para derivar la primera representación muestreada o la segunda representación muestreada; y

15

un divisor en ventanas adaptado para aplicar la primera ventana de ajuste a escala a la primera representación muestreada y la segunda ventana de ajuste a escala a la segunda representación muestreada para derivar una representación procesada de las tramas de audio primera, segunda y tercera de la señal de audio.

20

2. Procesador de audio según la reivindicación 1, en el que el muestreador es operativo para muestrear la señal de audio de modo que un contorno tonal dentro de las representaciones muestreadas primera y segunda es más constante que un contorno tonal de la señal de audio dentro de las tramas primera, segunda y tercera correspondientes.

25

3. Procesador de audio según la reivindicación 1, en el que el muestreador es operativo para remuestrear una señal de audio muestreada que tiene N muestras en cada una de las tramas primera, segunda y tercera, de modo que cada una de las representaciones muestreadas primera y segunda comprende 2N muestras.

30

4. Procesador de audio según la reivindicación 3, en el que el muestreador es operativo para derivar una muestra i de la primera representación muestreada en una posición dada por la fracción u entre las posiciones de muestreo originales k y (k+1) de las 2N muestras de las tramas primera y segunda, dependiendo la fracción u de un contorno de tiempo que asocia las posiciones de muestreo utilizadas por el muestreador y las posiciones de muestreo originales de la señal de audio muestreada de las tramas primera y segunda.

5. Procesador de audio según la reivindicación 4, en el que el muestreador es operativo para utilizar un contorno de tiempo derivado del contorno tonal p_i de las tramas según la siguiente ecuación:

$$\text{contorno_tiempo}_{i+1} = \text{contorno_tiempo}_i + (p_i I),$$

35

donde un intervalo de tiempo de referencia I para la primera representación muestreada se deriva de un indicador de altura tonal D derivado del contorno tonal p_i según:

$$D = \sum_{i=0}^{2N-1} p_i, \quad I = 2N / D.$$

40

6. Procesador de audio según la reivindicación 1, en el que el calculador de ventanas de transformada está adaptado para derivar ventanas de ajuste a escala con números idénticos de muestras, en el que un primer número de muestras utilizadas para atenuar la primera ventana de ajuste a escala difiere de un segundo número de muestras utilizadas para intensificar la segunda ventana de ajuste a escala.

45

7. Procesador de audio según la reivindicación 1, en el que el calculador de ventanas de transformada está adaptado para derivar una primera ventana de ajuste a escala en la que un primer número de muestras es menor que un segundo número de muestras de la segunda ventana de ajuste a escala cuando las tramas primera y segunda combinadas tienen una altura tonal media mayor que las tramas combinadas segunda y tercera o para derivar una primera ventana de ajuste a escala en la que el primer número de muestras es mayor que el segundo número de muestras de la segunda ventana de ajuste a escala cuando las tramas combinadas primera y segunda tienen una altura tonal media menor que las tramas combinadas segunda y tercera.

50

8. Procesador de audio según la reivindicación 6, en el que el calculador de ventanas de transformada está adaptado para derivar ventanas de ajuste a escala en las que un número de muestras antes de las muestras utilizadas para atenuar y en las que un número de muestras después de las muestras utilizadas para intensificar se establecen en una unidad y en las que el número de muestras después de las muestras utilizadas para atenuar y antes de las muestras utilizadas para intensificar se establecen en 0.

9. Procesador de audio según la reivindicación 8, en el que el calculador de ventanas de transformada está adaptado para derivar el número de muestras utilizadas para intensificar y utilizadas para atenuar dependiendo de un primer indicador de altura tonal D_j de las tramas primera y segunda que tienen muestras 0, ..., $2N-1$ y de un segundo indicador de altura tonal D_{j+1} de la segunda y la tercera que tienen muestras N , ..., $3N-1$, de modo que el número de muestras utilizadas para intensificar es:

$$N \quad \text{si} \quad D_{j+1} \leq D_j$$

o

$$N \times \frac{D_j}{D_{j+1}} \quad \text{si} \quad D_{j+1} > D_j;$$

y

el primer número de muestras utilizadas para atenuar es:

$$N \quad \text{si} \quad D_j \leq D_{j+1}$$

o

$$N \times \frac{D_{j+1}}{D_j} \quad \text{si} \quad D_j > D_{j+1}$$

donde los indicadores de altura tonal D_j y D_{j+1} se derivan del contorno tonal p_i según las siguientes ecuaciones:

$$D_{j+1} = \sum_{i=N}^{3N-1} p_i \quad \text{y} \quad D_j = \sum_{i=0}^{2N-1} p_i .$$

10. Procesador de audio según la reivindicación 8, en el que el calculador de ventanas es operativo para derivar el primer y segundo número de muestras volviendo a muestrear una ventana de intensificación y atenuación predeterminada con números iguales de muestras con respecto al primer y segundo número de muestras.

11. Procesador de audio para procesar una primera representación muestreada de una primera y una segunda trama de una señal de audio que tiene una secuencia de tramas en la que la segunda trama sigue a la primera trama y para procesar una segunda representación muestreada de la segunda trama y de una tercera trama de la señal de audio que sigue a la segunda trama en la secuencia de tramas, que comprende:

un calculador de ventanas de transformada adaptado para derivar una primera ventana de ajuste a escala para la primera representación muestreada utilizando información sobre un contorno tonal de la primera y la segunda trama y para derivar una segunda ventana de ajuste a escala para la segunda representación muestreada utilizando información sobre un contorno tonal de las tramas segunda y tercera, en el que las ventanas de ajuste a escala tienen un número idéntico de muestras y en el que un primer número de muestras utilizadas para atenuar la primera ventana de ajuste a escala difiere de un segundo número de muestras utilizadas para intensificar la segunda ventana de ajuste a escala;

un divisor en ventanas adaptado para aplicar la primera ventana de ajuste a escala a la primera representación muestreada y para aplicar la segunda ventana de ajuste a escala a la segunda representación muestreada; y

un remuestreador adaptado para remuestrear la primera representación muestreada ajustada a escala para derivar una primera representación remuestreada utilizando la información sobre el contorno tonal de la primera y la segunda trama y para remuestrear la segunda representación muestreada ajustada a escala para derivar una segunda representación remuestreada utilizando la información sobre el contorno tonal de las tramas segunda y tercera, dependiendo el remuestreo de las ventanas

de ajuste a escala derivadas.

- 5 12. Procesador de audio según la reivindicación 11, que comprende además un sumador adaptado para sumar la parte de la primera representación remuestreada correspondiente a la segunda trama y la parte de la segunda representación remuestreada correspondiente a la segunda trama para derivar una representación reconstruida de la segunda trama de la señal de audio.
- 10 13. Método para generar una representación procesada de una señal de audio que tiene una secuencia de tramas que comprende:
- muestrear la señal de audio dentro de una primera y una segunda trama de la secuencia de tramas, siguiendo la segunda trama a la primera trama, utilizando el muestreo información sobre un contorno tonal de la primera y la segunda trama para derivar una primera representación muestreada;
- muestrear la señal de audio dentro de la segunda y una tercera trama, siguiendo la tercera trama a la segunda trama en la secuencia de tramas, utilizando el muestreo la información sobre el contorno tonal de la segunda trama e información sobre un contorno tonal de la tercera trama para derivar una segunda representación muestreada;
- 15 derivar una primera ventana de ajuste a escala para la primera representación muestreada y una segunda ventana de ajuste a escala para la segunda representación muestreada, dependiendo las ventanas de ajuste a escala de los muestreos aplicados para derivar la primera representación muestreada o la segunda representación muestreada; y
- 20 aplicar la primera ventana de ajuste a escala a la primera representación muestreada y aplicar la segunda ventana de ajuste a escala a la segunda representación muestreada.
14. Método para procesar una primera representación muestreada de una primera y una segunda trama de una señal de audio que tiene una secuencia de tramas en la que la segunda trama sigue a la primera trama y para procesar una segunda representación muestreada de la segunda trama y de una tercera trama de la señal de audio que sigue a la segunda trama en la secuencia de tramas, que comprende:
- 25 derivar una primera ventana de ajuste a escala para la primera representación muestreada utilizando información sobre un contorno tonal de la primera y la segunda trama y derivar una segunda ventana de ajuste a escala para la segunda representación muestreada utilizando información sobre un contorno tonal de la segunda y la tercera trama, en el que las ventanas de ajuste a escala se derivan de manera que tienen un número idéntico de muestras, en el que un primer número de muestras utilizadas para atenuar la primera ventana de ajuste a escala difiere de un segundo número de muestras utilizadas para intensificar la segunda ventana de ajuste a escala;
- 30 aplicar la primera ventana de ajuste a escala a la primera representación muestreada y la segunda ventana de ajuste a escala a la segunda representación muestreada; y
- 35 remuestrear la primera representación muestreada ajustada a escala para derivar una primera representación remuestreada utilizando la información sobre el contorno tonal de la primera y la segunda trama y remuestrear la segunda representación muestreada ajustada a escala para derivar una segunda representación remuestreada utilizando la información sobre el contorno tonal de la segunda y la tercera trama, dependiendo el remuestreo de las ventanas de ajuste a escala derivadas.
- 40 15. Programa informático que comprende medios de código de programa que cuando se ejecuta en un ordenador hace que dicho ordenador ejecute las etapas de un método según las reivindicaciones 13 ó 14.

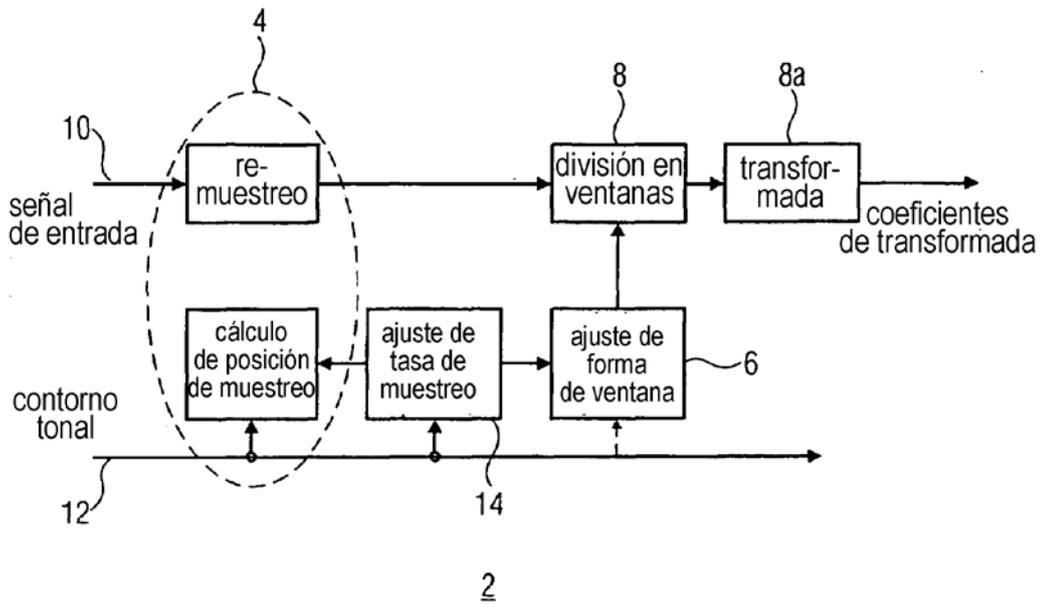


FIG 1

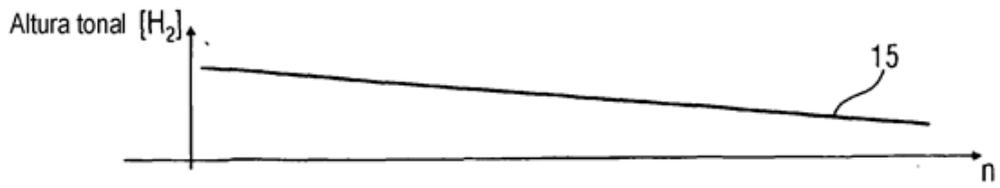


FIG 2A

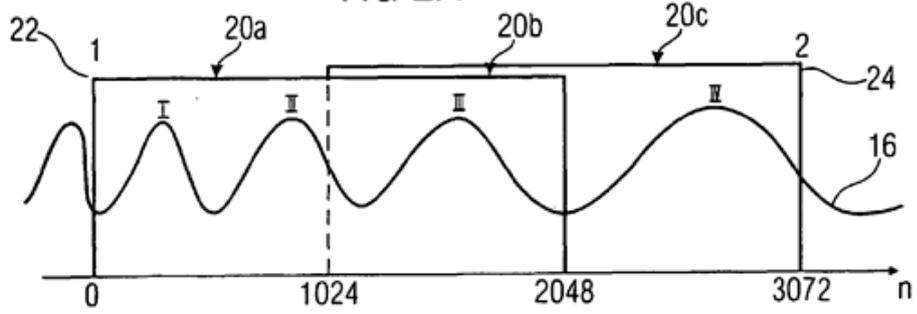


FIG 2B

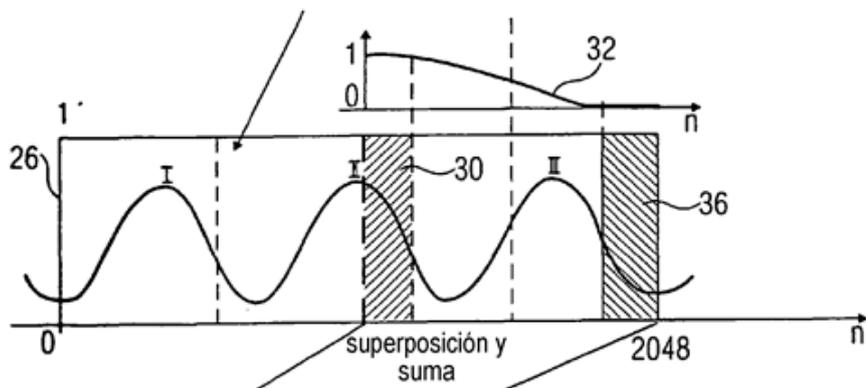


FIG 2C

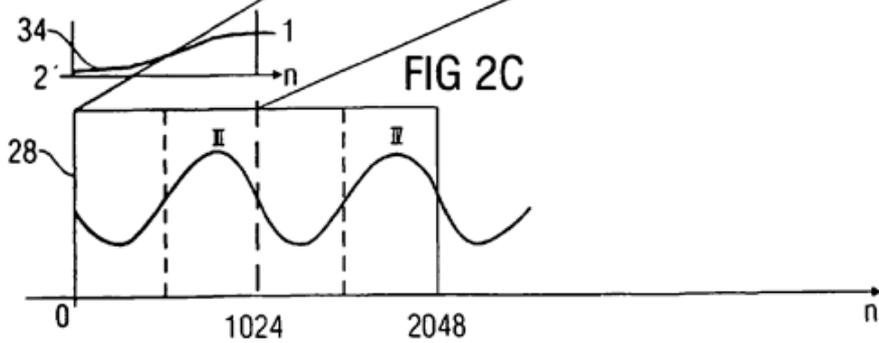


FIG 2D

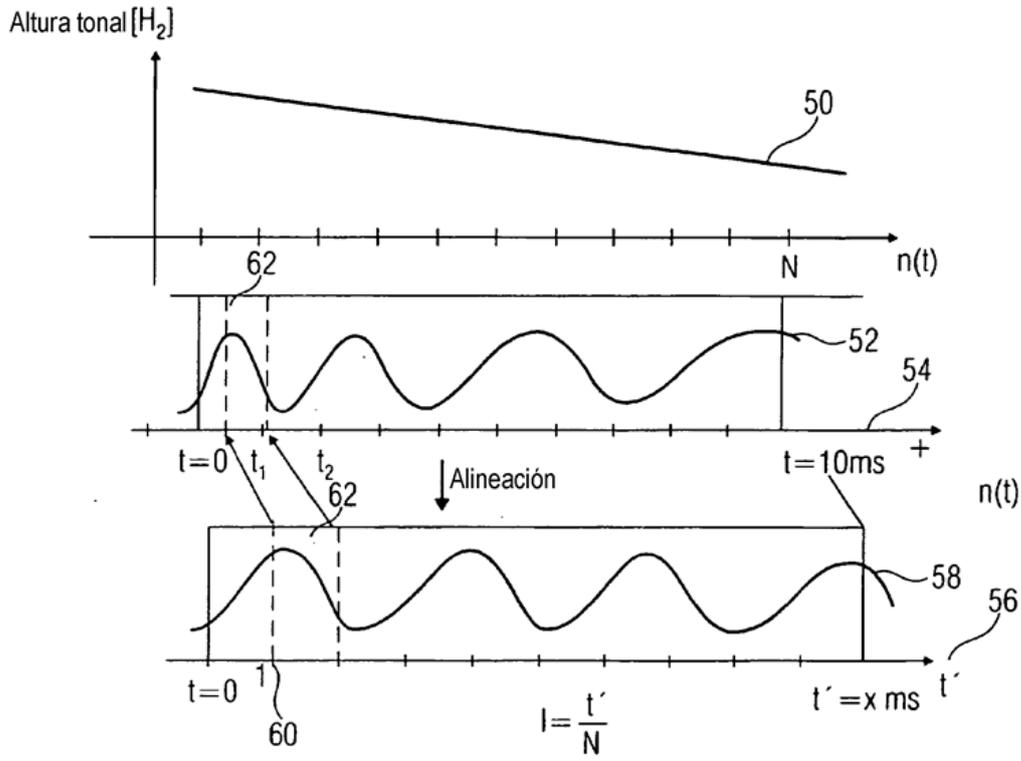


FIG 3

Contorno de tiempo

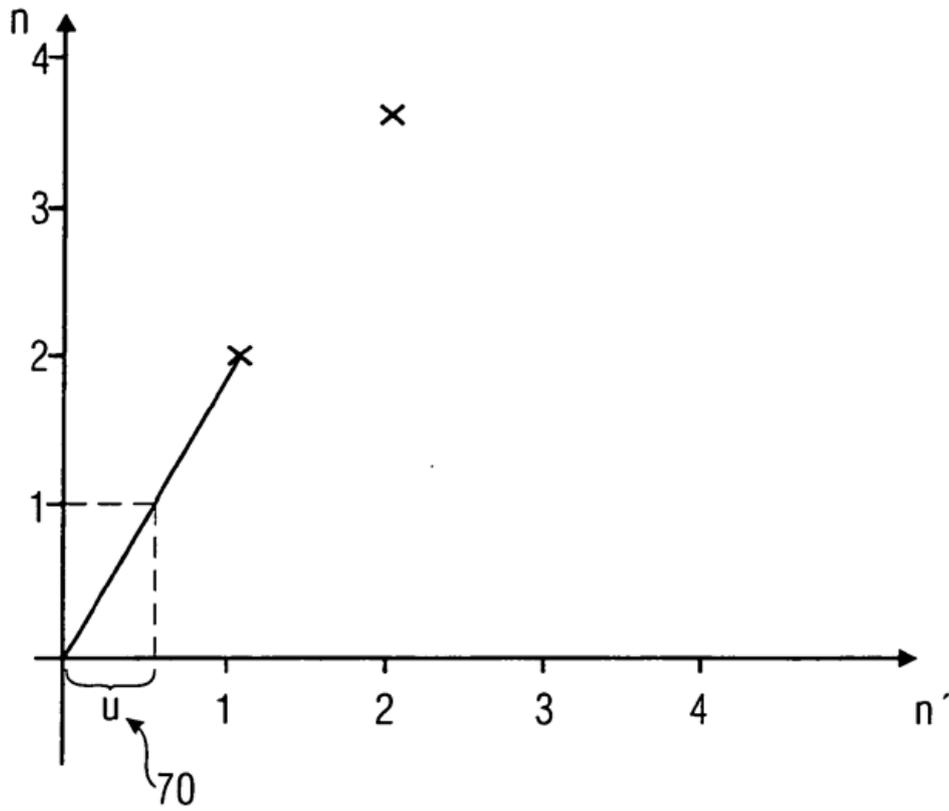


FIG 4

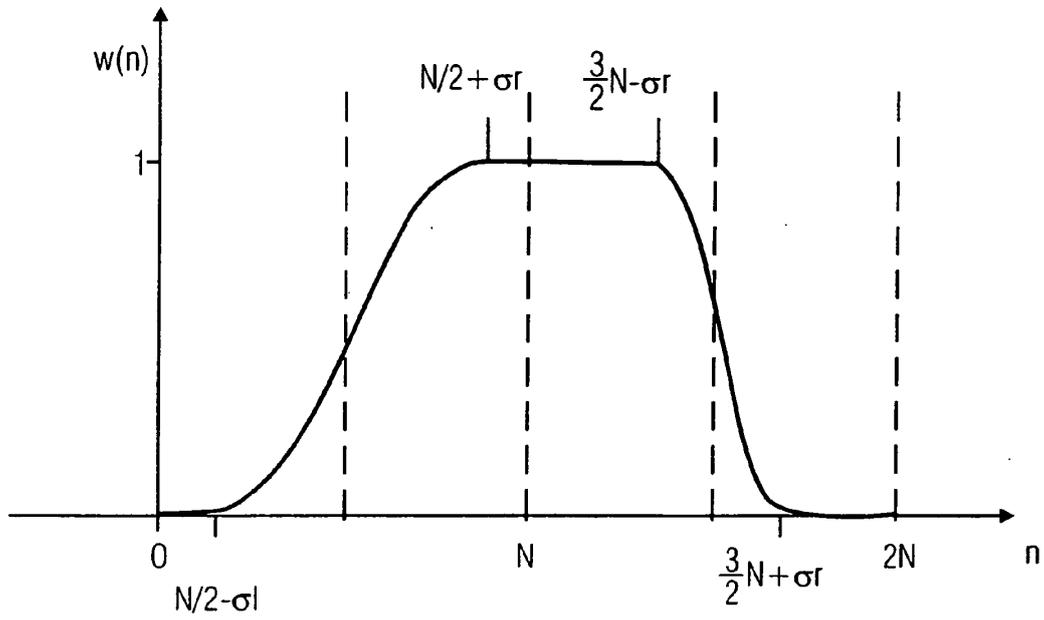


FIG 5

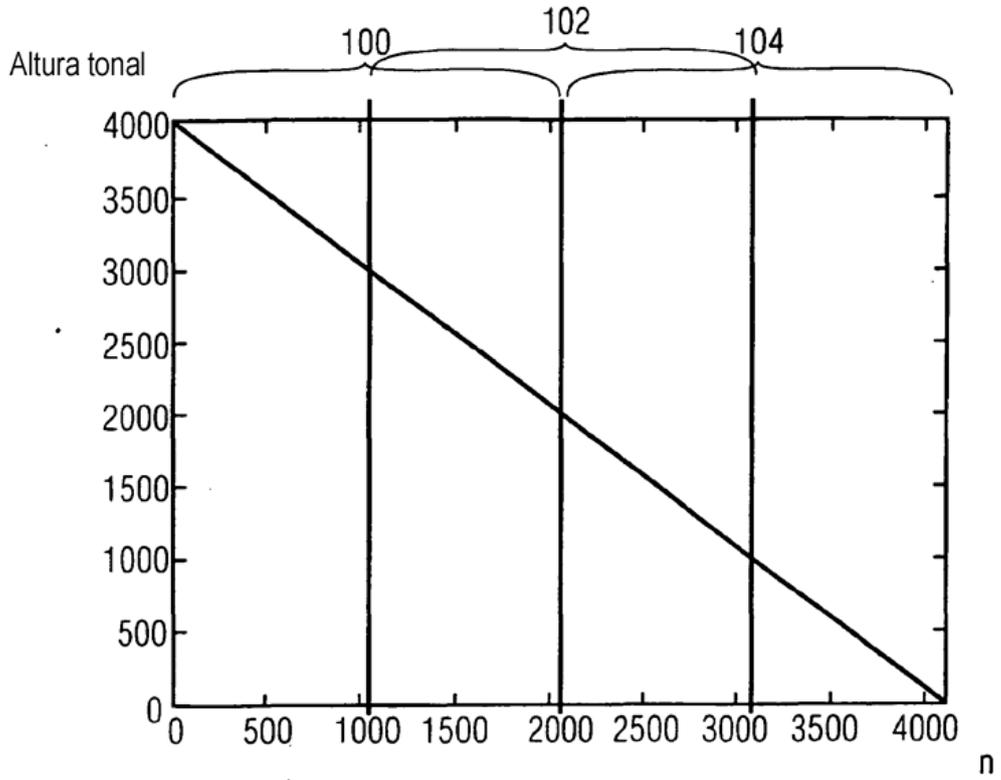


FIG 6

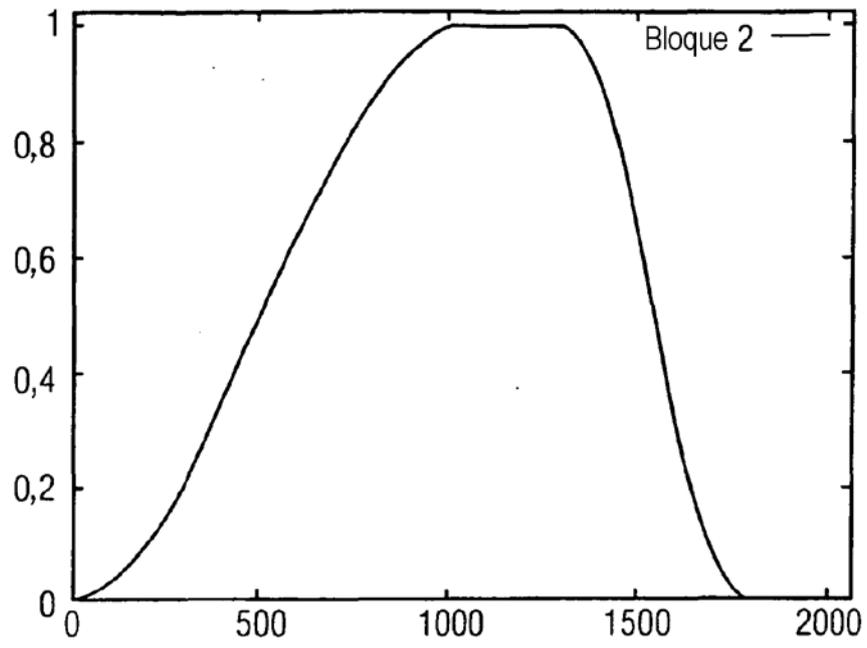


FIG 7

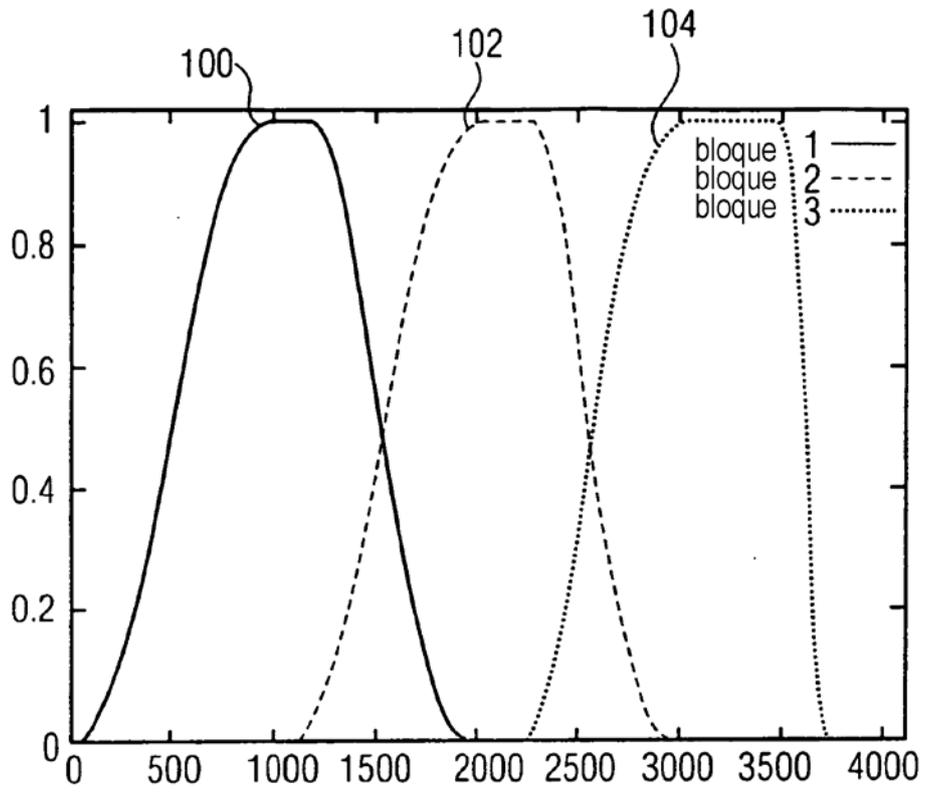


FIG 8

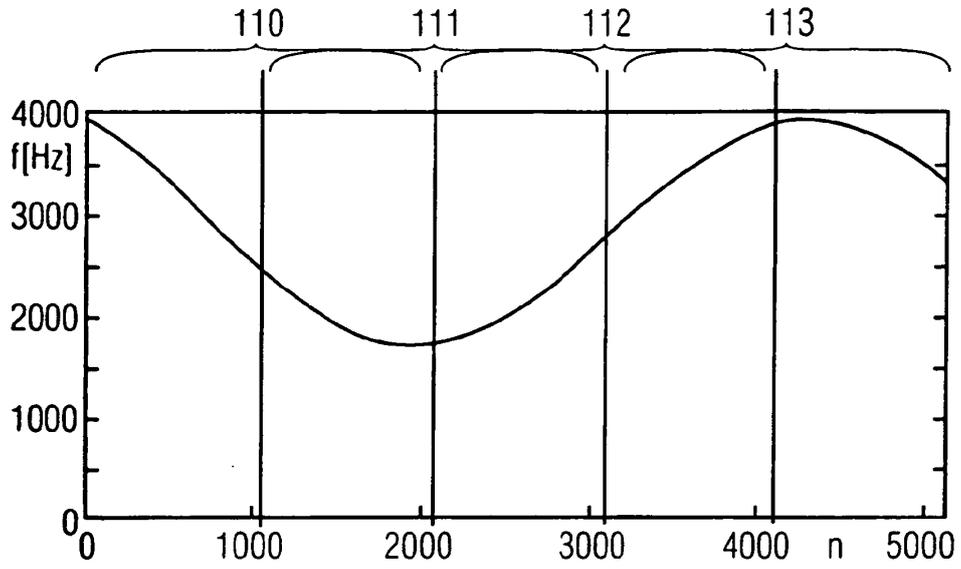


FIG 9

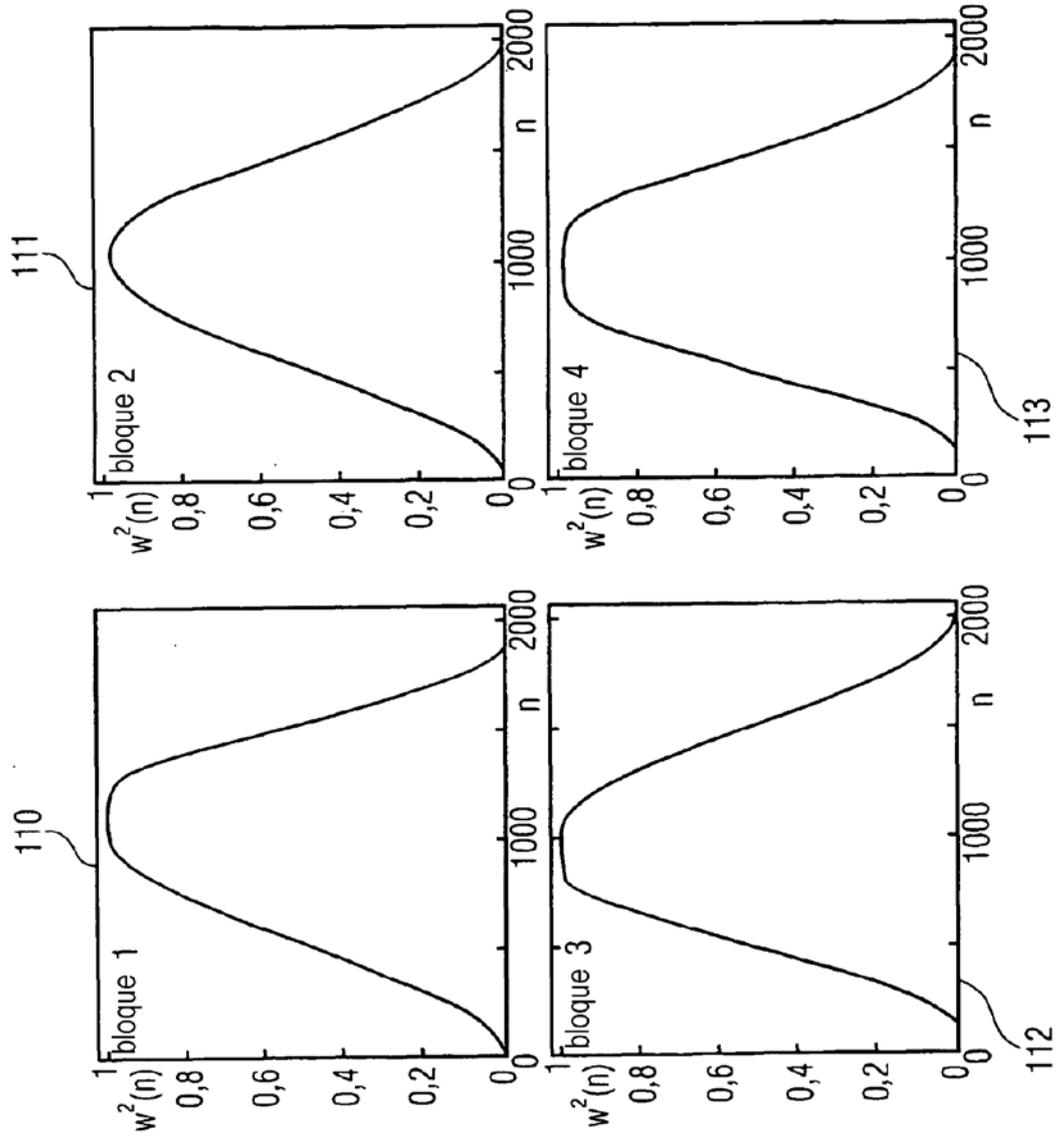


FIG 10

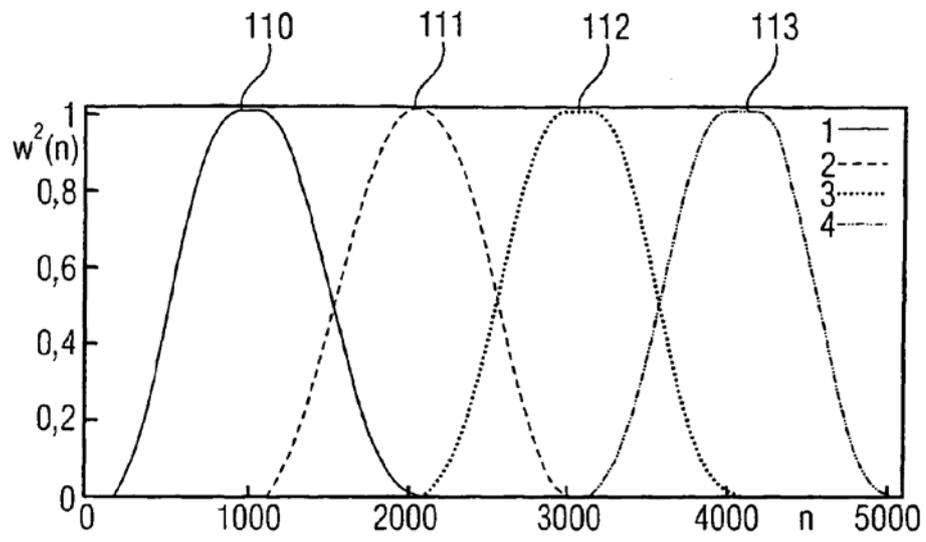


FIG 11

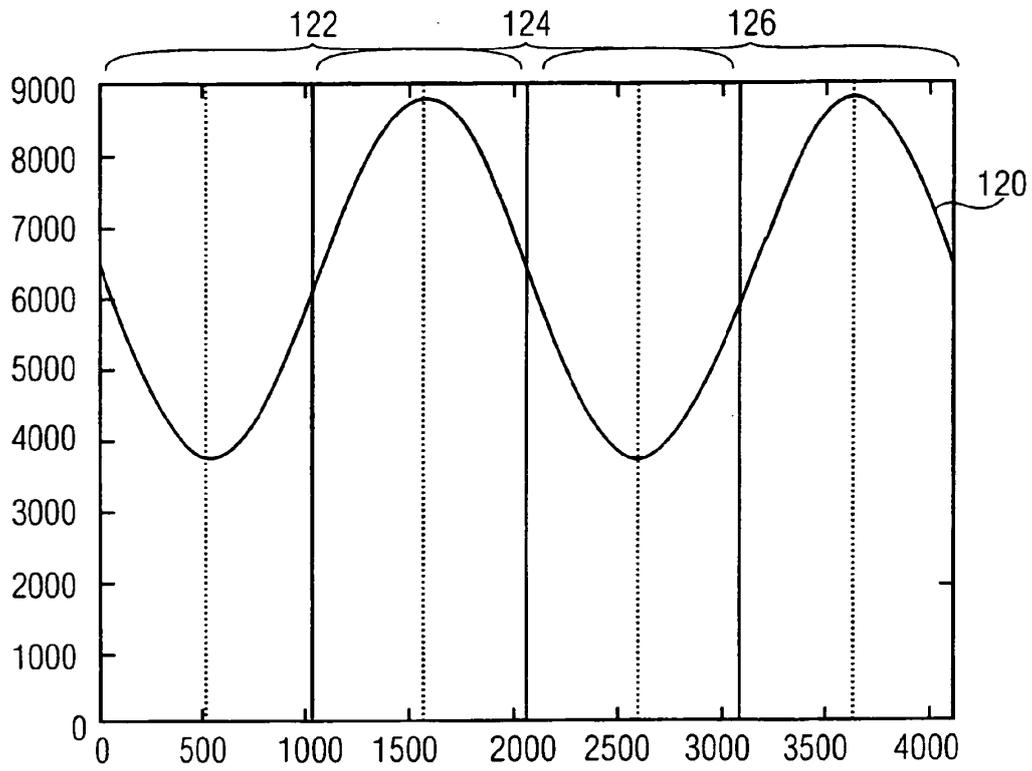


FIG 11A

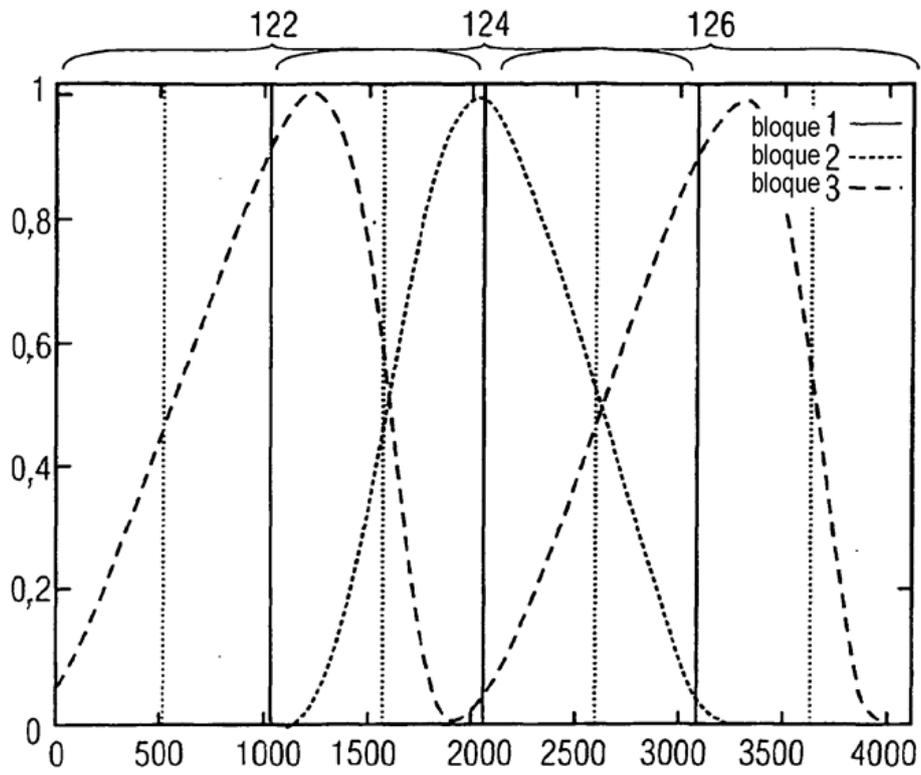


FIG 11B

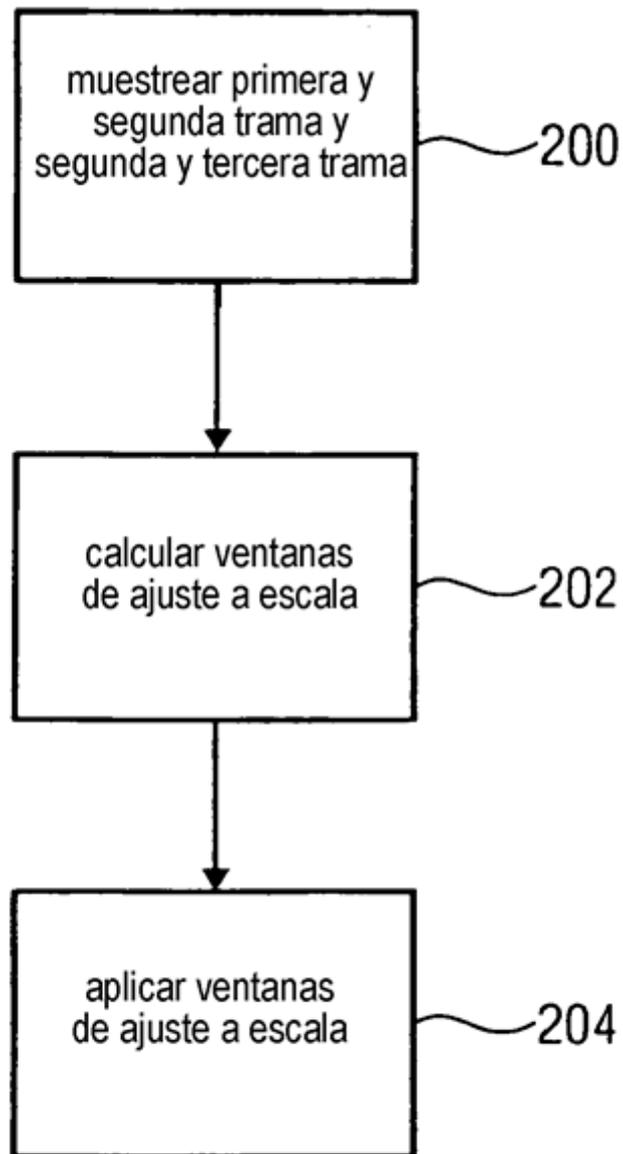


FIG 12

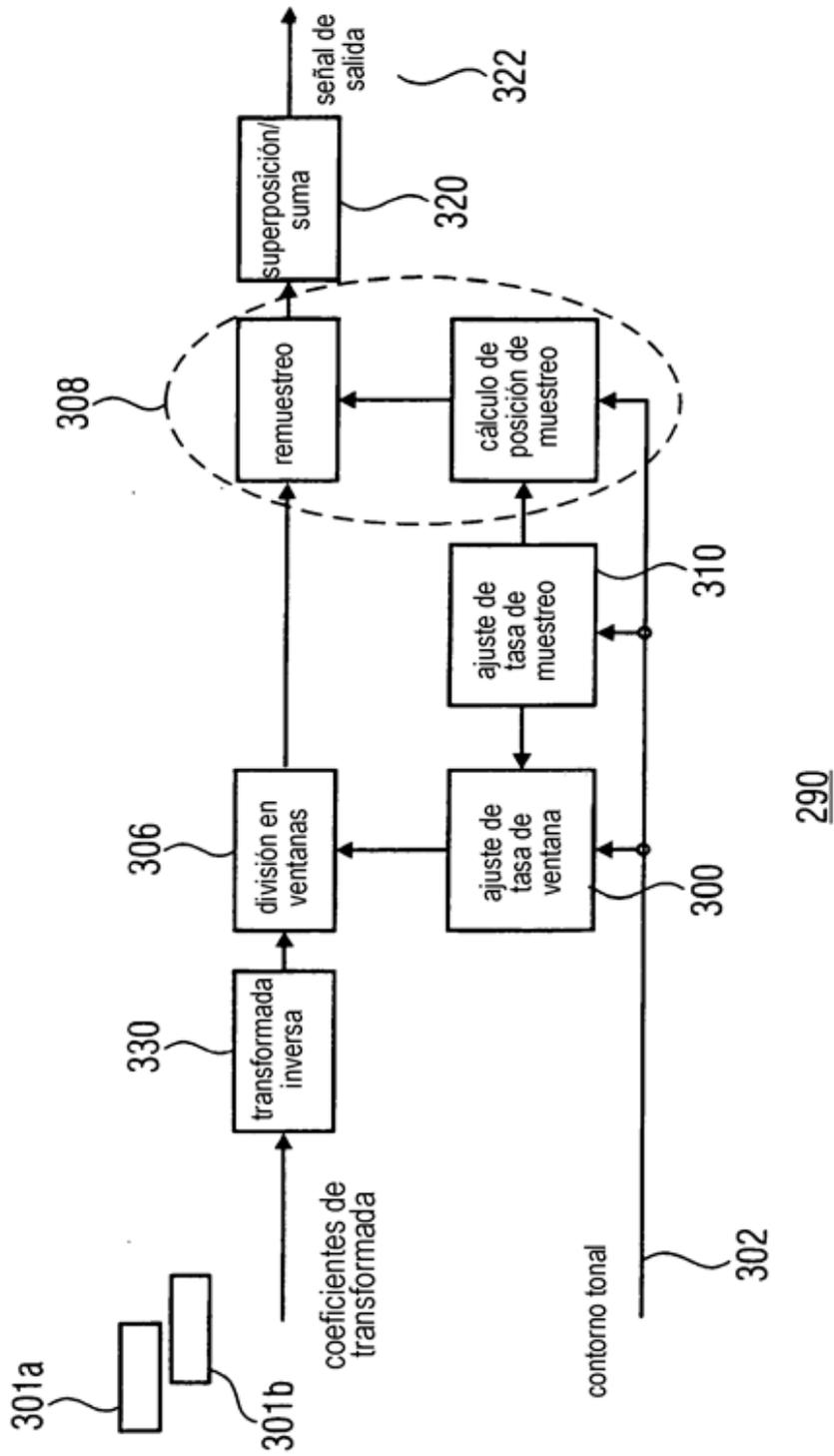


FIG 13

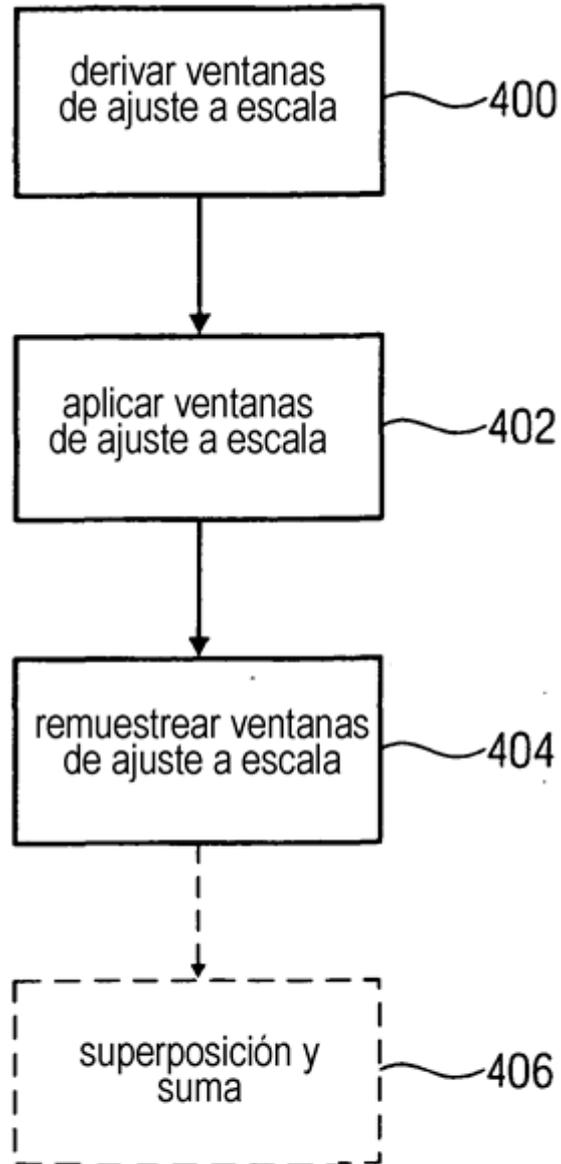


FIG 14