

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 404 132**

51 Int. Cl.:

G10L 19/00 (2013.01)

G10L 19/02 (2013.01)

G10L 21/04 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.07.2009 E 09776910 (3)**

97 Fecha y número de publicación de la concesión europea: **03.04.2013 EP 2260485**

54 Título: **Descodificador de señal de audio, codificador de señal de audio, representación de señal de audio de multi-canal modificada, procedimientos y programa de computadora**

30 Prioridad:

11.07.2008 US 79873 P
08.10.2008 US 103820 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
24.05.2013

73 Titular/es:

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

72 Inventor/es:

BAYER, STEFAN;
DISCH, SASCHA;
GEIGER, RALF;
FUCHS, GUILLAUME;
NEUENDORF, MAX;
SCHULLER, GERALD y
EDLER, BERND

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 404 132 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

Descodificador de señal de audio, codificador de señal de audio, representación de señal de audio de multi-canal modificada, procedimientos y programa de computadora

ANTECEDENTES DE LA INVENCION

5 **[0001]** Las realizaciones de acuerdo con la invención se refieren a un descodificador de señal de audio. Realizaciones adicionales de acuerdo con la invención se refieren a un codificador de señal de audio. Realizaciones adicionales de acuerdo con la invención se refieren a una representación de señal de audio de multi-canal codificada. Realizaciones adicionales de acuerdo con la invención se refieren a un procedimiento para proveer una
10 representación de señal de audio de multi-canal descodificada, con un procedimiento para proveer una representación codificada de una señal de audio de multi-canal y con un programa de computadora para implementar dichos procedimientos.

[0002] Algunas realizaciones de acuerdo con la invención se refieren a procedimientos para un codificador de transformada de MDCT distorsionada en el tiempo.

15 **[0003]** En lo siguiente, se dará una breve introducción al campo de codificación de audio distorsionada en el tiempo, conceptos de los cuales pueden ser aplicados en conjunción con algunas realizaciones de la invención.

[0004] En años recientes, se han desarrollado técnicas para transformar una señal de audio a una representación de dominio de frecuencia y para codificar eficientemente esta representación de dominio de frecuencia, por ejemplo, tomando en cuenta umbrales de enmascaramiento perceptual. Este concepto de codificación de señal de audio es particularmente eficiente si las longitudes de bloque, para los cuales un conjunto de coeficientes espectrales codificados son transmitidos, son largos y si solamente un número comparativamente pequeño de coeficientes espectrales está por encima del umbral de enmascaramiento global mientras que un número grande de coeficientes espectrales están cerca o debajo del umbral de enmascaramiento global y pueden así ser despreciados (o codificados con longitud de código mínima).
20

[0005] Por ejemplo, las transformadas dobladas moduladas a base de coseno o a base de seno son frecuentemente usadas en aplicaciones para codificación fuente debido a sus propiedades de compactación de energía. Esto es, para tonos armónicos con frecuencias fundamentales constantes (altura), concentran la energía de señal a un número bajo de componentes espectrales (sub-bandas), lo que conduce a una representación de señal eficiente.
25

[0006] En general, la altura (fundamental) de una señal se comprenderá que es la frecuencia dominante más baja distinguible del espectro de la señal. En el modelo de habla común, la altura es la frecuencia de la señal de excitación modulada por la garganta humana. Si solamente una frecuencia fundamental estuviera presente, el espectro sería extremadamente simple, comprendiendo la frecuencia fundamental y los sobretonos solamente. Tal espectro podría ser codificado altamente de manera eficiente. Para señales con altura variable, sin embargo, la energía correspondiente a cada componente armónico es esparcida sobre varios coeficientes de transformada, conduciendo así a una reducción de eficiencia de codificación.
30

[0007] Con el fin de superar esta recepción de la eficiencia de codificación, la señal de audio a ser codificada es re-muestreada efectivamente en una rejilla temporal no uniforme. En el procesamiento subsecuente, las posiciones de muestra obtenidas por el re-muestreo uniforme son procesadas como si representaran valores en una rejilla temporal uniforme. Esta operación es denotada comúnmente por la parte "formación de tiempo". Los tiempos de muestra pueden ser escogidos ventajosamente en dependencia de la variación temporal de la altura, de tal manera que una variación de altura en la versión distorsionada en el tiempo de la señal de audio es más pequeña que una variación de altura en la variación original de la señal de audio (antes de la distorsión de tiempo). Después de la distorsión en el tiempo de la señal de audio, la versión distorsionada en el tiempo de la señal de audio es convertida al dominio de frecuencia. La distorsión de tiempo altura-dependiente tiene el efecto de que la representación del dominio de frecuencia de la señal de audio distorsionada en el tiempo es comúnmente concentrada a un número mucho más pequeño de componentes espectrales que una representación de dominio de frecuencia de la señal de audio original (no distorsionada en el tiempo).
35
40
45

[0008] En el lado del descodificador, la representación de frecuencia-dominio de la señal de audio distorsionada en el tiempo es convertida de regreso al dominio de tiempo, de tal manera que una representación de tiempo-dominio de la señal de audio distorsionada en el tiempo está disponible en el lado del descodificador. Sin embargo, en la representación del dominio de tiempo de la señal de audio distorsionada en el tiempo reconstruida del lado del descodificador, las variaciones de altura originales de la señal de audio de entrada del lado del codificador no son incluidas. Así, todavía otra distorsión de tiempo mediante re-muestreo de la representación del dominio de tiempo reconstruida del lado del descodificador de la señal de audio distorsionada en el tiempo es aplicada. Con el fin de obtener una buena reconstrucción de la señal de audio concentrada del lado del codificador en el descodificador, es deseable que la distorsión en el tiempo del lado del descodificador sea por lo menos aproximadamente la operación inversa con respecto a la distorsión de tiempo del lado del codificador. Con el fin de obtener una distorsión de tiempo apropiada, es deseable tener información disponible en el lado del descodificador que permite el ajuste de distorsión de tiempo del lado del descodificador.
50
55

[0009] Ya que es comúnmente requerido transferir tal información del decodificador de señal de audio al decodificador de señal de audio, es deseable mantener una velocidad de bits requerida para esta transmisión pequeña en tanto que todavía permite una reconstrucción confiable de la información de distorsión de tiempo requerida en el lado del decodificador.

5 **[0010]** En vista de la discusión anterior, hay un deseo de tener un concepto que permita un almacenamiento eficiente en velocidad de bits y/o transmisión de una señal de audio de multi-canal.

RESUMEN DE LA INVENCION

10 **[0011]** Una realización de acuerdo con la invención crea un decodificador de señal de audio para proveer una representación de señal de audio de multi-canal descodificada en base a una representación de señal de audio de multi-canal descodificada. El decodificador de señal de audio comprende un decodificador de distorsión del tiempo configurado para usar selectivamente contornos de distorsión del tiempo específicos de canal de audio individuales o un contorno de distorsión del tiempo de multi-canal adjunto para una reconstrucción de distorsión del tiempo de una pluralidad de canales de audio representados por la representación de señal de audio de multi-canal codificada.

15 **[0012]** Esta realización de acuerdo con la invención está basada en el descubrimiento de que una codificación eficiente de diferentes tipos de señales de audio de multi-canal puede ser obtenida al conmutar entre un almacenamiento y/o transmisión de contornos de distorsión del tiempo específicos del canal de audio y contornos de distorsión del tiempo de multi-canal adjuntos. Se ha encontrado que, en algunos casos, una variación de altura es significativamente diferente en los canales de una señal de audio de multi-canal. También, se ha encontrado que en otros casos, la variación de altura es aproximadamente igual para múltiples canales de una señal de audio de multi-canal. En vista de estos tipos diferentes de señales (o porciones de señal de una sola señal de audio), se ha encontrado que la exigencia de codificación puede ser mejorada si el decodificador es apto de derivar flexiblemente (conmutable o selectivamente) los contornos de distorsión del tiempo para la reconstrucción de los diferentes canales de la señal de audio de multi-canal a partir de representaciones de contorno de distorsión del tiempo específicas de canal de audio individuales o de una representación de contorno de distorsión del tiempo de multi-canal adjunta.

20

25

[0013] En una realización preferida, el decodificador de distorsión del tiempo está configurado para usar selectivamente un contorno de distorsión del tiempo de multi-canal adjunto para una reconstrucción de distorsión del tiempo de una pluralidad de canales de audio para los cuales la información del dominio espectral codificada individual está disponible. De acuerdo con un aspecto de la invención, se ha encontrado que el uso de un contorno de distorsión del tiempo de multi-canal adjunto para una reconstrucción de distorsión del tiempo de una pluralidad de canales de audio no es solamente aplicable si los diferentes canales de audio representan un contenido de audio similar, si no aún si diferentes canales de audio representan un contenido de audio significativamente diferente. Así, se ha encontrado que es útil combinar el concepto de usar un contorno de distorsión del tiempo de multi-canal adjunto para la evaluación de información de dominio espectral codificada individual para diferentes canales de audio. Por ejemplo, este concepto es particularmente útil si un primer canal de audio representa una primera parte de una pieza polifónica de música, mientras que un segundo canal de audio representa una segunda parte de la pieza polifónica de música. La primera señal de audio y la segunda señal de audio pueden por ejemplo, representar el sonido producido por diferentes cantantes o por diferentes instrumentos. Así, una representación de dominio espectral del primer canal de audio puede ser significativamente diferente de una representación del dominio espectral de dicho canal de audio. Por ejemplo, las frecuencias fundamentales de los diferentes canales de audio pueden ser diferentes. También, los diferentes canales de audio pueden comprender diferentes características con respecto a las armónicas de la frecuencia fundamental. No obstante, puede haber una tendencia significativa de que las alturas de los diferentes canales de audio varían aproximadamente en paralelo. En este caso, es muy eficiente aplicar una distorsión en tiempo común (descrita por el contorno de distorsión del tiempo de multi-canal adjunto) a los diferentes canales de audio, aunque los diferentes canales de audio comprenden contenidos de audio significativamente diferentes (por ejemplo, que tienen diferentes frecuencias fundamentales y diferentes espectros armónicos. No obstante, en otros casos, es naturalmente deseable aplicar diferentes distorsiones de tiempo a diferentes canales de audio.

30

35

40

45

[0014] En una realización preferida de la invención, el decodificador de distorsión del tiempo está configurado para recibir una primera información de dominio espectral codificada asociada con un primero de los canales de audio y para proveer, en base a lo mismo, una representación del dominio de tiempo distorsionada del primer canal de audio utilizando una transformación de dominio de frecuencia a dominio de tiempo. También, el decodificador de distorsión del tiempo está configurado además para recibir una segunda información de dominio espectral codificada, asociada con un segundo de los canales de audio y para proveer, en base a lo mismo, una representación del dominio de tiempo distorsionada del segundo canal de audio utilizando una transformación de dominio de frecuencia a dominio de tiempo. En este caso, la segunda información de dominio espectral codificada puede ser diferente de la primera información de dominio espectral. También, el decodificador de distorsión del tiempo está configurado para re-muestrear de manera variable en el tiempo, en base al contorno de distorsión del tiempo de multi-canal adjunto, la representación del dominio del tiempo distorsionada del primer canal de audio o una versión procesada del mismo, para obtener una representación muestreada regularmente con el primer canal de audio y para re-muestrear de manera variable en el tiempo, también en base al contorno de distorsión del tiempo de

50

55

60

multi-canal adjunto, la representación del dominio de tiempo distorsionada del segundo canal de audio o una versión procesada de la misma, para obtener una representación muestreada regularmente del segundo canal de audio.

5 **[0015]** En otra realización preferida, el descodificador de distorsión del tiempo está configurado para derivar un contorno de tiempo de multi-canal adjunto a partir de la información de contorno de distorsión del tiempo de multi-canal adjunta. Además, el descodificador de distorsión del tiempo está configurado para derivar una primera forma de ventana específica de canal individual asociada con el primero de los canales de audio en base a la primera información de forma de ventana codificada y para derivar una segunda forma de ventana específica de canal individual asociada con el segundo de los canales de audio en base a la segunda información de forma de ventana codificada. El descodificador de distorsión del tiempo está configurado además para aplicar la primera forma de ventana a la representación del dominio de tiempo distorsionada del primer canal de audio, para obtener una versión procesada de la representación del dominio de tiempo distorsionada del primer canal de audio y para aplicar la segunda forma de ventana a la representación del dominio de tiempo distorsionada del segundo canal de audio, para obtener una versión procesada de la representación del dominio de tiempo distorsionada del segundo canal de audio. En este caso, el descodificador de distorsión del tiempo es apto de aplicar diferentes formas de ventana a las representaciones de dominio de tiempo distorsionadas del primero y segundo canal de audio dependiendo de la formación de forma de ventana específica de canal individual.

10 **[0016]** Se ha encontrado que es en algunos casos recomendable aplicar ventanas de diferentes formas a diferentes señales de audio en preparación de una operación de distorsión de tiempo, aún si la separación de distorsión de tiempo está basada en un contorno de distorsión del tiempo común. Por ejemplo, puede haber una transición entre un cuadro, en el cual hay un contorno de distorsión del tiempo común para dos canales de audio y un cuadro subsecuente en el cual hay diferentes contornos de distorsión del tiempo para los dos canales de audio. Sin embargo, el contorno de distorsión del tiempo de uno de los dos canales de audio en el cuadro subsecuente puede ser una continuación no variable del contorno de distorsión del tiempo común en el cuadro presente, mientras que el contorno de distorsión del tiempo del otro canal de audio en el cuadro subsecuente puede ser variable con respecto al contorno de distorsión del tiempo común del cuadro presente. Así, una forma de ventana que está adaptada a una evolución no variable del contorno de distorsión del tiempo es representada para uno de los canales de audio, mientras que una forma de ventana adaptada a una evolución variable del contorno de distorsión del tiempo puede ser aplicada para el otro canal de audio. Así, la evolución diferente de los canales de audio puede ser tomada en consideración.

20 **[0017]** En otra realización de acuerdo con la invención, el descodificador de distorsión del tiempo puede estar configurado para aplicar un escalamiento de tiempo común, que es determinado por el contorno de distorsión del tiempo del multi-canal adjunto y diferentes formas de ventana cuando se forman ventanas de las representaciones de dominio del tiempo de los primeros y segundos canales de audio. Se ha encontrado que aún si se usan diferentes formas de ventana para la formación de ventanas diferentes canales de audio antes de la distorsión del tiempo respectiva, el escalamiento en el tiempo del contorno de distorsión debe ser adaptado en paralelo con el fin de evitar la degradación de la impresión de audición.

25 **[0018]** Otra realización de acuerdo con la invención crea un codificador de señal de audio para proveer una representación codificada de una señal de audio de multi-canal. El codificador de señal de audio comprende un proveedor de representación de audio codificada configurado para proveer selectivamente una representación de audio que comprende una información de contorno de distorsión del tiempo común, asociada comúnmente con una pluralidad de canales de audio de la señal de audio de multi-canal o una representación de audio codificada que comprende información de contorno de distorsión del tiempo individual, asociada individualmente con los diferentes canales de audio de la pluralidad de canales de audio, dependiendo de una información que describe una similitud de diferencias entre los contornos de distorsión del tiempo asociados con canales de audio de la pluralidad de canales de audio. Esta realización de acuerdo con la invención está basada en el descubrimiento de que en muchos casos, múltiples canales de una señal de audio de multi-canal comprenden características de variación de altura similares. Así, es en algunos casos eficiente incluir a la representación codificada de la señal de audio de multi-canal una información de contorno de distorsión del tiempo común, asociada comúnmente con una pluralidad de los canales de audio. De esta manera, la eficiencia de codificación puede ser mejorada para muchas señales. Sin embargo, se ha encontrado que para otros tipos de señales (o aún para otras porciones de una señal) no es recomendable usar tal información de distorsión del tiempo común. Así, una codificación de señal es suficiente y puede ser obtenida si el codificador de señal de audio determina la similitud o diferencia entre contornos de distorsión asociados con los diferentes canales de audio bajo consideración. Sin embargo, se ha encontrado que es por supuesto valioso tener una mirada en los contornos de distorsión del tiempo individuales, debido a que hay muchas señales que comprenden una representación del dominio del tiempo significativamente diferentes o representación del dominio de frecuencia significativamente diferente, aunque tienen contornos de distorsión del tiempo muy similares. Así, se ha encontrado que la evaluación del contorno de distorsión del tiempo es un nuevo criterio para la determinación de la similitud en las señales, que proporciona una información extra cuando se compara con una sola evaluación de las representaciones del dominio de tiempo de múltiples señales de audio o de las representaciones del dominio de frecuencia de las señales de audio.

30 **[0019]** En una realización preferida, el proveedor de representación de audio codificada está configurado para aplicar selectivamente la información de contorno de distorsión del tiempo común para obtener una versión

distorsionada en el tiempo de un primero de los canales de audio y para obtener una versión distorsionada en el tiempo de un segundo de los canales de audio. El proveedor de representación de audio codificada está configurado además para proveer una primera información del dominio espectral codificada individual asociada con el primero de los canales de audio en base a la versión distorsionada en el tiempo del canal de audio y para proveer una segunda información del dominio espectral codificado individual asociada con el segundo canal de audio en base a la versión distorsionada en el tiempo del segundo de los canales de audio. Esta realización está basada en el descubrimiento mencionado anteriormente de que los canales de audio pueden tener contenidos de audio significativamente diferentes, aún si tienen un contorno de distorsión del tiempo muy similar. Así, es frecuentemente recomendable proveer diferente información del dominio espectral asociado con diferentes canales de audio, aún si los canales de audio son distorsionados en el tiempo de acuerdo con una información de distorsión del tiempo común. En otras palabras, la realización está basada en el descubrimiento de que no hay ninguna interrelación estricta entre la similitud de los contornos de distorsión del tiempo y la similitud de las representaciones del dominio de frecuencia de los diferentes canales de audio.

[0020] En otra realización preferida, el codificador está configurado para obtener la información de contorno de distorsión común de tal manera que el contorno de distorsión común representa un promedio de contornos de distorsión individuales asociados con el primer canal de señal de audio y con el segundo canal de señal de audio.

[0021] En otra realización preferida, el proveedor de representación de audio codificada está configurado para proveer una información lateral dentro de la representación codificada de la señal de audio de multi-canal, de tal manera que la información lateral indica, en una base según el cuadro de audio, si los datos de distorsión del tiempo están presentes para un cuadro si una información del contorno de distorsión del tiempo común está presente para un cuadro. Al proveer información de si datos de distorsión del tiempo están presentes para un cuadro, es posible reducir la velocidad de bits preferida para la transmisión de la información de distorsión del tiempo. Se ha encontrado que es comúnmente referido transmitir una información que describe una pluralidad de valores de contorno de distorsión del tiempo dentro de un cuadro, si se usa la información del tiempo para tal cuadro. Sin embargo, también se ha encontrado que hay muchos cuadros para los cuales la aplicación de una distorsión del tiempo no efectúa una ventaja significativa. Todavía, se ha encontrado que es más eficiente indicar, utilizando por ejemplo un bit de información adicional, si los datos de distorsión del tiempo para un cuadro están disponibles. Al usar tal señalización, la transmisión de la información de distorsión del tiempo extensa (que comprende comúnmente información con respecto a una pluralidad de valores de contorno de distorsión del tiempo) puede ser omitida, ahorrando mediante esto bits.

[0022] Una realización adicional de acuerdo con la invención crea una representación de señal de audio de multi-canal codificada que representa una señal de audio de multi-canal. La representación de señal de audio de multi-canal comprende una representación del dominio de frecuencia codificada que representa una pluralidad de canales de audio distorsionados en el tiempo, distorsionándose selectivamente en el tiempo de acuerdo con una distorsión del tiempo común, en función de una información que describe una similitud o diferencia entre contornos de tiempo asociados con los canales de audio de la señal de audio multi-canal. La representación de señal de audio de multi-canal también comprende una representación codificada de una información de contorno de distorsión del tiempo común, asociada comúnmente con los canales de audio y que representan la distorsión del tiempo común.

[0023] En una realización preferida, la representación del dominio de frecuencia codificada comprende información del dominio de frecuencia codificada de múltiples canales de audio que tienen diferente contenido de audio. También, la representación codificada de la información de contorno de distorsión común está asociada con los múltiples canales de audio que tienen diferentes contenidos de audio.

[0024] Otra realización de acuerdo con la invención crea un procedimiento para proveer una representación de señal de audio multi-canal codificada en base a una representación de señal de audio de multi-canal codificada. Este procedimiento puede ser complementado por cualquiera de los elementos y funcionalidades descritos en la presente también para el aparato de la invención.

[0025] Todavía otra realización de acuerdo con la invención crea un procedimiento para proveer una representación codificada de una señal de audio de multi-canal. Este procedimiento puede ser complementado por cualquiera de los elementos y funcionalidades descritos en la presente también para el aparato de la invención.

[0026] Todavía otra realización de acuerdo con la invención crea un programa de computadora para implementar los procedimientos mencionados anteriormente.

BREVE DESCRIPCIÓN DE LAS FIGURAS

[0027] Las realizaciones de acuerdo con la invención serán descritas subsecuentemente tomando referencias a las figuras adjuntas en las cuales:

La figura 1 muestra un diagrama de bloques esquemático de un codificador de audio de distorsión de tiempo;

La figura 2 muestra un diagrama de bloques esquemático de un decodificador de audio de distorsión de tiempo;

La figura 3 muestra un diagrama de bloques esquemático de un descodificador de señal de audio de acuerdo con una realización de la invención;

La figura 4 muestra un diagrama de flujo de un procedimiento para proveer una representación de señal de audio descodificada de acuerdo con una realización de la invención;

5 La figura 5 muestra un extracto detallado de un diagrama de bloques esquemático de un descodificador de señal de audio de acuerdo con una realización de la invención;

La figura 6 muestra un extracto detallado de un diagrama de flujo de un procedimiento para proveer una representación de señal de audio descodificada de acuerdo con una realización de la invención;

10 Las figuras 7a, 7b muestran una representación gráfica de una reconstrucción del contorno de distorsión de tiempo de acuerdo con una realización de la invención;

La figura 8 muestra otra representación gráfica de una reconstrucción de un contorno de distorsión de tiempo de acuerdo con una realización de la invención;

Las figuras 9a y 9b muestran algoritmos para el cálculo del contorno de distorsión de tiempo;

15 La figura 9c muestra una tabla de un mapeo de un índice de proporción de distorsión de tiempo a un valor de proporción de distorsión de tiempo;

Las figuras 10a y 10b muestran representaciones de algoritmos para el cálculo de un contorno de tiempo, una posición de muestra, una longitud de transición, una "primera posición" y una "última posición";

La figura 10c muestra una representación de algoritmos para un cálculo de forma de ventana;

Las figuras 10d y 10e muestran una representación de algoritmos para una aplicación de una ventana;

20 La figura 10f muestra una representación de algoritmos para un re – muestreo variable en el tiempo;

La figura 10g muestra una representación gráfica de algoritmos para un procesamiento de cuadro de distorsión post – tiempo y para superposición y adición;

Las figuras 11a y 11b muestran una leyenda;

25 La figura 12 muestra una representación gráfica de un contorno de tiempo que puede ser extraído de un contorno de distorsión de tiempo;

La figura 13 muestra un diagrama de bloques detallado esquemático de un aparato para proveer un contorno de distorsión, de acuerdo con una realización de la invención;

La figura 14 muestra un diagrama de bloques esquemático de un descodificador de señal de audio de acuerdo con otra realización de la invención;

30 La figura 15 muestra un diagrama de bloques esquemático de otro calculador de contorno de distorsión de tiempo, de acuerdo con una realización de la invención;

Las figuras 16a, 16b muestran una representación gráfica del cálculo de valores de nodo de distorsión de tiempo, de acuerdo con una realización de la invención;

35 La figura 17 muestra un diagrama de bloques esquemático de otro codificador de señal de audio de acuerdo con una realización de la invención;

La figura 18 muestra un diagrama de bloques esquemático de otro descodificador de señal de audio de acuerdo con una realización de la invención y

Las figuras 19a – 19f muestran representaciones de elementos de sintaxis de una corriente de audio, de acuerdo con una realización de la invención.

40 **DESCRIPCIÓN DETALLADA DE LAS REALIZACIONES**

1. Codificador de audio de distorsión de tiempo de acuerdo con la figura 1

45 **[0028]** Ya que la presente invención está relacionada con la codificación de audio de formación de tiempo y descodificación de audio de distorsión de tiempo, se dará una vista general breve de un prototipo de codificador de audio de formación de tiempo y un descodificador de audio de distorsión de tiempo en los cuales se puede aplicar la presente invención.

[0029] La figura 1 muestra un diagrama de bloques esquemático de un codificador de audio de distorsión de tiempo, al cual se pueden integrar algunos aspectos y realizaciones de la invención. El codificador de señal de audio 100 de la figura 1 está configurado para recibir una señal de audio de entrada 110 y proveer una representación codificada de la señal de audio de entrada 110 en una secuencia de cuadros. El codificador de audio 100 comprende un muestreador 104, que está adaptado para tomar muestras de la señal de audio 110 (señal de entrada) para derivar bloques de señales (representaciones muestreadas) 105 usadas como base para una transformada de dominio de frecuencia. El codificador de audio 100 comprende además un calculador de ventana de transformada 106, adaptado para derivar ventanas de escalamiento para las representaciones muestreadas 105 emitidas del muestreador 104. Estas son introducidas a un formador de ventana 108 que es apto para aplicar las ventanas de escalamiento a las representaciones muestreadas 105 derivadas del muestreador 104. En algunas realizaciones, el codificador de audio 100 puede comprender adicionalmente un transformador de dominio de frecuencia 108a con el fin de derivar una representación de dominio de frecuencia (por ejemplo, en forma de coeficiente de transformada) de las representaciones muestreadas y escaladas 105. Las representaciones de dominio de frecuencia pueden ser procesadas o transmitidas adicionalmente como una representación codificada de la señal de audio 110.

[0030] El codificador de audio 100 usa además un contorno de altura 112 de la señal de audio 110, que puede ser provisto al codificador de audio 100 o que puede ser derivado por el codificador de audio 100. El codificador de audio 100 puede por consiguiente comprender opcionalmente un estimador de altura para derivar el contorno de altura 112. El muestreador 104 puede operar sobre una representación continua de la señal de audio de entrada 110. Alternativamente, el muestreador 104 puede operar sobre una representación ya muestreada de la señal de audio de entrada 110. En el último caso, el muestreador 104 puede volver a tomar muestras de la señal de audio 110. El muestreador 104 puede por ejemplo estar adaptado a la superposición de bloques de audio vecinos de distorsión de tiempo de tal manera que la porción de superposición tiene una altura constante o variación de altura reducida dentro de cada uno de los bloques de entrada después de la toma de muestras.

[0031] El calculador de ventana de transformada 106 deriva las ventanas de escalamiento para los bloques de audio dependiendo de la distorsión de tiempo efectuada por el muestreador 104. Para este fin, un bloque de ajuste de velocidad de muestras opcional 114 puede estar presente con el fin de definir una regla de distorsión de tiempo usada por el muestreador, que es luego también provista al especulador de ventana de transformada 106. En una realización alternativa, el bloque de ajuste de velocidad de muestreo 114 puede ser emitido y el contorno de altura 112 puede ser provisto directamente al calculador de ventana de transformada 106, que puede por sí mismo efectuar los cálculos apropiados. Además, el muestreador 104 puede comunicar el muestreo aplicado al calculador de ventana de transformada 106 con el fin de permitir el cálculo de ventanas de escalamiento apropiadas.

[0032] La distorsión de tiempo es efectuada de tal manera que un contorno de altura de los bloques de audio muestreados distorsionados en tiempo y muestreados por el muestreador 104 es más constante que el contorno de altura de la señal de audio original 110 dentro del bloque de entrada.

2. Descodificador de audio de distorsión de tiempo de acuerdo con la figura 2.

[0033] La figura 2 muestra un diagrama de bloques esquemático de un descodificador de audio de distorsión de tiempo 200 para el procesamiento de una primera representación distorsionada en el tiempo y muestreada o simplemente representación distorsionada en el tiempo de un primero y segundo cuadro de una señal de audio que tiene una señal de audio en los cuales el segundo cuadro sigue el primer cuadro y para el procesamiento adicional de una segunda representación distorsionada en el tiempo del segundo cuadro y de un tercer cuadro en seguida del segundo cuadro en la secuencia de cuadros. El descodificador de audio 200 comprende un calculador de ventana de transformada 210 apto para derivar una primera ventana de escalamiento para la primera representación distorsionada en el tiempo 211a utilizando información en cuanto a un contorno de altura 212 del primero y el segundo cuadro y para derivar una segunda ventana de escalamiento para la segunda representación distorsionada en el tiempo 211b utilizando información en cuanto a un contorno de altura del segundo y tercer cuadro, en donde las ventanas de escalamiento pueden tener números idénticos de muestras y en donde el primer número de muestras usadas para desvanecer la primera ventana de escalamiento puede diferir del segundo número de muestras usadas para desvanecer la segunda ventana de escalamiento. El descodificador de audio 200 comprende además un formador de ventana 216 apto para aplicar la primera ventana de escalamiento a la primera representación de distorsión de tiempo y para aplicar la segunda ventana de escalamiento a la segunda representación distorsionada en el tiempo. El descodificador de audio 200 comprende además un re – muestreador 218 adaptado para distorsionar inversamente el tiempo la primera representación distorsionada en el tiempo escalada para derivar una primera representación muestreada utilizando la información en cuanto al contorno de altura del primero y el segundo cuadro y para distorsionar inversamente en el tiempo la segunda representación escalada para derivar una segunda representación muestreada utilizando la información en cuanto al contorno de altura del segundo y el tercer cuadro, de tal manera que una porción de la primera representación muestreada correspondiente al segundo cuadro comprende un contorno de altura que es igual, dentro del intervalo de tolerancia predeterminado, a un contorno de altura de la porción de la segunda representación muestreada correspondiente al segundo cuadro. Con el fin de derivar la ventana de escalamiento, el calculador de ventana de transformada 210 puede ya sea recibir el contorno de altura 212 directamente o recibir información en cuanto a la distorsión de tiempo desde un ajustador de velocidad de muestras opcional 220, que recibe el contorno de altura 212 y que deriva una estrategia de distorsión de tiempo inversa, de tal manera que las posiciones de muestra en una escala de tiempo

lineal para las muestras de superposición son idénticas o caso idénticas y espaciadas regularmente, de tal manera que la altura se convierte en la misma en las regiones de superposición y opcionalmente las longitudes de desvanecimiento diferentes de partes de ventana traslapantes antes de la distorsión de tiempo inversa se vuelven de la misma longitud después de la distorsión de tiempo inversa.

5 **[0034]** El descodificador de audio 200 comprende además un adicionador opcional 230, que está adaptado para agregar la porción de la primera representación muestreada correspondiente al segundo cuadro y la porción de la segunda representación muestreada correspondiente al segundo cuadro para derivar una representación de construida de segundo cuadro de la señal de audio como una señal de salida 242. La primera representación distorsionada en el tiempo y la segunda representación distorsionada en el tiempo podrían ser provistas, en una
10 realización, como una entrada al descodificador de audio 200. En una realización adicional, el descodificador de audio 200 puede comprender opcionalmente un transformador de dominio de frecuencia inverso 240, que puede derivar las primeras y segundas representaciones distorsionadas en el tiempo de representaciones de dominio de frecuencia de las primeras y segundas representaciones distorsionadas en el tiempo provistas a la entrada del transformador de dominio de frecuencia inverso 240.

15 **3. Descodificador de señal de audio de distorsión de tiempo de acuerdo con la Figura 3**

[0035] En lo siguiente, se describirá un descodificador de señal de audio simplificado. La Figura 3 muestra un diagrama de bloques esquemático de este descodificador de señal audio simplificado 300. El descodificador de señal de audio 300 está configurado para recibir la representación de señal de audio codificada 310, y proveer, en base a la misma, una representación de señal de audio descodificada 312, en donde la representación de señal de
20 audio codificada 310 comprende una información de evaluación del contorno de distorsión de tiempo. El descodificador de señal de audio 300 comprende un calculador de contorno de formación de tiempo 320 configurado para generar datos de contorno de distorsión de tiempo 322 en base a la información de evolución de contorno de distorsión de tiempo, tal información de evolución del contorno de distorsión de tiempo describe una evolución temporal del contorno de distorsión de tiempo, y tal información de evolución del contorno de distorsión de tiempo está comprendida por la representación de señal de audio codificada 310. Cuando se derivan los datos de contorno
25 de distorsión de tiempo 322 de la información de evolución del contorno de distorsión de tiempo 312, el calculador de contorno de distorsión de tiempo 320 se reinicia repetidamente desde un valor de inicio de contorno de distorsión de tiempo predeterminada, como se describirá en detalle en lo siguiente. El reinicio puede tener la consecuencia de que el contorno de distorsión de tiempo comprende discontinuidades (cambios graduales que son más grandes que las etapas codificadas por la información de evolución del contorno de formación de tiempo 312). El decodificador de
30 señal de audio 300 comprende además un re-escalador de datos de contorno de formación de tiempo 330 que está configurado para reescalar por lo menos una porción de los datos de contorno de distorsión de tiempo 322, de tal manera que se evita una discontinuidad en el reinicio del cálculo del contorno de distorsión de tiempo, es reducido o eliminado, en una versión reescalada 332 del contorno de distorsión de tiempo.

35 **[0036]** El descodificador de señal de audio 300 también comprende un descodificador de distorsión 340 configurado para proveer una representación de señal de audio descodificada 312 en base a la representación de señal de audio codificada 310 y que utiliza la versión reescalada 332 del contorno de distorsión de tiempo.

[0037] Para poner el descodificador de señal de audio 300 en el contexto de decodificación de audio de distorsión de tiempo, se debe notar que la representación de señal de audio codificada 310 puede comprender una
40 representación codificada de los coeficientes de transformada 211 y también una representación codificada del contorno de altura 212 (también designado como el contorno de distorsión de tiempo). El calculador de contorno de distorsión de tiempo 320 y el re-escalador de datos de contorno de distorsión de tiempo 330 pueden estar configurados para proveer una representación reconstruida del contorno de altura 212 en forma de la versión reescalada 332 del contorno de distorsión de tiempo. El descodificador de distorsión 340 por ejemplo, puede tomar
45 la funcionalidad de la formación de ventana 216, la retoma de muestras 218, el ajuste de velocidad de muestreo 220 y el ajuste de la forma de ventana 210. Además, el descodificador de distorsión 340 puede por ejemplo comprender opcionalmente la funcionalidad de la transformada inversa 240 y el traslape/adición 230, de tal manera que la representación de señal de audio descodificada 312 puede ser equivalente a la señal de audio de salida 232 del descodificador de audio de distorsión de tiempo 200.

50 **[0038]** Al aplicar el re-escalamiento a los datos de contorno de distorsión de tiempo 322, se puede obtener una versión reescalada continua (o por lo menos aproximadamente continua) 332 del contorno de distorsión de tiempo, asegurando mediante esto que se evite un sobreflujo o subflujo numérico aún cuando se usa una información de evolución de contorno de distorsión de tiempo de variación relativa eficiente de codificar.

55 **4. Procedimiento para proveer una representación de señal de audio descodificada de acuerdo con la Figura 4**

[0039] La Figura 4 muestra un diagrama de flujo de un procedimiento para proveer una representación de señal de audio descodificada en base a una representación de señal de audio codificada que comprende una información de evolución del contorno de distorsión de tiempo, que puede ser efectuada por el aparato 300 de acuerdo con la Figura 3. El procedimiento 400 comprende una primera etapa 410 para generar los datos de contorno de distorsión

de tiempo, reiniciándose repetidamente desde un valor de inicio de contorno de distorsión de tiempo predeterminado, en base a la información de evolución de contorno de distorsión de tiempo que describe una evolución temporal del contorno de distorsión de tiempo.

5 [0040] El procedimiento 400 comprende además una etapa 420 de re-escalamiento de lo menos una porción de los datos de control de distorsión de tiempo, de tal manera que una discontinuidad en uno de los reinicios es evitada, reducida o eliminada en una versión reescalada del contorno de distorsión de tiempo.

[0041] El procedimiento 400 comprende además una etapa 430 de proveer una representación de señal de audio descodificada en base a la representación de señal de audio codificada utilizando la versión reescalada del contorno de distorsión de tiempo.

10 **5. Descripción detallada de una realización de acuerdo con la invención con referencia a las Figuras 5-9**

[0042] En lo siguiente, se describirá una realización de acuerdo con la invención en detalle con referencia a las Figuras 5-9.

15 [0043] La Figura 5 muestra un diagrama de bloques esquemático de un aparato 500 para proveer información de control de distorsión de tiempo 512 en base a la información de evolución del contorno de distorsión de tiempo 510. El aparato 500 comprende medios 520 para proveer una información de contorno de distorsión de tiempo reconstruida 522 en base a la información de evolución de contorno de distorsión de tiempo 510, y un calculador de información de control de distorsión de tiempo 530 para proveer la información de control de distorsión de tiempo 512 en base a la información de contorno de distorsión de tiempo reconstruida 522.

Medios 520 para proveer la información de contorno de distorsión de tiempo reconstruida

20 [0044] En lo siguiente, se describirá la estructura y funcionalidad de los medios 520. Los medios 520 comprenden un calculador de contorno de distorsión de tiempo 540, que está configurado para recibir la información de evolución de contorno de distorsión de tiempo 510 y para proveer, en base a la misma, una nueva información de porción de contorno de distorsión 542. Por ejemplo, un conjunto de información de evolución del contorno de distorsión de tiempo puede ser transmitido al aparato 500 para cada cuadro de la señal de audio a ser reconstruido. No obstante,
25 el conjunto de información de evolución de contorno de distorsión de tiempo 510 asociado con el cuadro de la señal de audio a ser reconstruido puede ser usado para la reconstrucción de una pluralidad de cuadros de la señal de audio. Similarmente, una pluralidad de conjuntos de información de evolución de contorno de distorsión de tiempo puede ser usado para la reconstrucción del contenido de audio de un solo cuadro de la señal de audio, como se discutirá en detalle en lo siguiente. Como conclusión, se puede afirmar que en algunas realizaciones, la información de evolución del contorno de distorsión de tiempo 510 puede ser actualizada a la misma velocidad a la cual
30 conjuntos del coeficiente de dominio de transformada de la señal de audio a ser reconstruida o actualizada (una porción de contorno de distorsión de tiempo por cuadro de la señal de audio).

[0045] El calculador de contorno de distorsión de tiempo 540 comprende un calculador de valor de modo de distorsión 544, que está configurado para calcular una pluralidad (o secuencia temporal) de valores de nodo de contorno de distorsión en base a la pluralidad (o secuencia temporal) de valores de proporción de contorno de distorsión de tiempo (o índices de proporción de distorsión de tiempo), en donde los valores de proporción de distorsión de tiempo (o índices) están comprendidos por la información de evolución del contorno de distorsión de tiempo 510. Para este propósito, el calculador de valor de nodo de distorsión 544 está configurado para iniciar la provisión de los valores de nodo de contorno de distorsión de tiempo a un valor de partida predeterminado (por ejemplo 1) y para calcular valores de nodo de contorno de distorsión de tiempo subsecuentes que utilizan los valores de proporción de contorno de distorsión de tiempo, como se discutirá posteriormente en la presente.

[0046] Además, el calculador de contorno de distorsión de tiempo 540 comprende opcionalmente un interpolador 548 que está configurado para interpolar entre valores de nodo de contorno de distorsión de tiempo subsecuentes. Así, se obtiene la descripción 542 de la nueva porción de contorno de distorsión de tiempo, en donde la nueva porción de contorno de distorsión de tiempo se inicia comúnmente a partir del valor de partida predeterminado usado por el calculador de valor de nodo de distorsión 524. Además, los medios 520 están configurados para considerar porciones de contorno de distorsión de tiempo adicionales, es decir una llamada "última porción de contorno de distorsión de tiempo" y una llamada "porción de contorno de distorsión de tiempo actual" para la provisión de una plena sección de contorno de distorsión de tiempo. Para este propósito, los medios 520 están configurados para
45 guardar la llamada "última porción de contorno de distorsión de tiempo" y la llamada "porción de contorno de distorsión de tiempo actual" en una memoria no mostrada en la Figura 5.

[0047] Sin embargo, los medios 520 también comprenden un re-escalador 550, que está configurado para reescalar la "última porción de contorno de distorsión de tiempo" y la "porción de contorno de distorsión de tiempo actual" para evitar (o reducir o eliminar) cualesquier discontinuidades en la plena sección de contorno de distorsión de tiempo, que está basada en la "última porción de contorno de distorsión de tiempo", la "porción de contorno de distorsión de tiempo actual" y la "nueva porción de contorno de distorsión de tiempo". Para este propósito, el re-escalador 550 está configurado para recibir la descripción almacenada de la "última porción de contorno de distorsión de tiempo" y de la "porción de contorno de distorsión de tiempo actual" y para reescalar conjuntamente la "última porción de
55

contorno de distorsión de tiempo” y la “porción de contorno de distorsión de tiempo actual”, para obtener versiones reescaladas de la “última porción de contorno de distorsión de tiempo” y la “porción de contorno de distorsión de tiempo actual”. Detalles concernientes con el re-escalamiento efectuado por el re-escalador 550 serán discutidos a continuación, con referencia a las Figuras 7a, 7b y 8.

5 **[0048]** Además, el re-escalador 550 puede también estar configurado para recibir, por ejemplo de una memoria no mostrada en la Figura 5, un valor de suma asociado con la “última porción de contorno de distorsión de tiempo” y otro valor de suma asociado con la “porción de contorno de distorsión de tiempo actual”. Estos valores de suma son algunas veces designado con “last_warp_sum” y “cur_warp_sum”, respectivamente. El re-escalador 550 está configurado para reescalar los valores de suma asociados con las porciones de contorno de distorsión de tiempo que utilizan el mismo factor de re-escalamiento con los cuales las porciones de contorno de distorsión de tiempo correspondientes son reescaladas. Así, se tienen valores de suma reescalados.

10 **[0049]** En algunos casos, los medios 520 pueden comprender un actualizador 560, que está configurado para actualizar repetidamente las porciones de contorno de distorsión de tiempo introducidas al re-escalador 550 y también los valores de suma introducidos al re-escalador 550. Por ejemplo, el actualizador 560 puede estar configurado para actualizar dicha información a la velocidad de cuadro. Por ejemplo, la “nueva porción de contorno de distorsión de tiempo” del presente ciclo de cuadros puede servir como la “porción de contorno de distorsión de tiempo actual” en un siguiente ciclo de cuadro. Similarmente, la “porción de contorno de distorsión de tiempo actual” reescalada del ciclo de cuadros actual puede servir como la “última porción de contorno de distorsión de tiempo” en un siguiente ciclo de cuadros. Así, se crea una implementación eficiente en memoria, debido a que la “última porción de contorno de distorsión de tiempo” del ciclo de cuadros actual puede ser descartada después de la consumación del ciclo de cuadros actual.

15 **[0050]** Para resumir lo anterior, los medios 520 están configurados para proveer, para cada ciclo de cuadros (con la excepción de algunos ciclo de cuadros especiales, por ejemplo al comienzo de una secuencia de cuadros, o al final de una secuencia de cuadros, o en un cuadro en el cual la distorsión de tiempo es inactiva) una descripción de una sección de contorno de distorsión de tiempo que comprende una descripción de una “nueva porción de contorno de distorsión de tiempo”, de una “porción de contorno de distorsión de tiempo actual reescalada” y de una “última porción de contorno de distorsión de tiempo reescalada”. Además, los medios 520 pueden proveer, para cada ciclo de cuadro (con la excepción del ciclo de cuadros especial mencionado anteriormente) una representación de valores de suma de contorno de distorsión, por ejemplo, que comprenden un “nuevo valor de toma de porción de contorno de distorsión de tiempo”, un “valor de suma de contorno de distorsión de tiempo actual reescalado” y un “último valor de suma de contorno de distorsión de tiempo reescalado”.

20 **[0051]** El calculador de información de control de distorsión de tiempo 530 está configurado para calcular la información de control de distorsión de tiempo 512 en base a la información de contorno de distorsión de tiempo reconstruida provista por los medios 520. Por ejemplo, el calculador de información de control de distorsión de tiempo comprende un calculador de contorno de tiempo 570, que está configurado para calcular un contorno de tiempo 572 en base a la información de control de distorsión de tiempo reconstruida. Además, el calculador de información de contorno de distorsión de tiempo 530 comprende un calculador de posición de muestras 574, que está configurado para recibir el contorno de tiempo 572 y proveer, en base a lo mismo, una información de posición de muestra, por ejemplo en forma de un vector de posición de muestra 576. El vector de posición de muestra 576 describe la distorsión de tiempo efectuada, por ejemplo por el re-muestreador 218.

25 **[0052]** El calculador de información de control de distorsión de tiempo 530 también comprende un calculador de longitud de transición, que está configurado para derivar una información de longitud de transición a partir de la información de control de distorsión de tiempo reconstruida. La información de longitud de transición 582 puede comprender por ejemplo una información que describe una longitud de transición izquierda y una información que describe una longitud de transición derecha. La longitud de transición puede depender por ejemplo de la longitud de segmento de tiempo descritos por la “última porción de contorno de distorsión de tiempo”, la “porción de contorno de distorsión de tiempo actual” y la “nueva porción de contorno de distorsión de tiempo”. Por ejemplo, la longitud de transición puede ser acertada (cuando se compara con una longitud de transición predeterminada) si la extensión temporal de un segmento de tiempo descrito por la “última porción de contorno de distorsión de tiempo” es más corta que una extensión temporal del segmento de tiempo descrito por la “porción de contorno de distorsión de tiempo actual”, o si la extensión temporal de un segmento de tiempo descrita por la “nueva porción de contorno de distorsión de tiempo” es más corta que la extensión temporal del segmento de tiempo descrito por la “porción de contorno de distorsión de tiempo actual”. Además, el calculador de información de control de distorsión de tiempo 530 puede comprender además un calculador de primera y última posición 584, que está configurado para calcular una llamada “primera posición” y una llamada “última posición” en base a la longitud de transición izquierda y derecha. La “primera posición” y la “última posición” incrementan la eficiencia del re-muestreador, ya que las regiones fuera de estas posiciones son idénticas a cero después de la formación de ventanas y por consiguiente no necesitan ser tomados en cuenta para la distorsión de tiempo. Se debe notar aquí que el vector de posición de muestra 576 comprende, por ejemplo, información requerida por la distorsión de tiempo efectuada por el re-muestreador 280. Además, la longitud de transición izquierda y derecha 582 y la “primera posición” y “última posición” 586 constituyen información, que es por ejemplo, requerida por el formador de ventanas 216.

[0053] Así, se puede decir que los medios 520 y el calculador de información de control de distorsión de tiempo 530 pueden emprender conjuntamente la funcionalidad del ajuste de velocidad de muestras 220, del ajuste de forma de ventana 210 y del cálculo de posición de toma de muestras 219.

5 [0054] En lo siguiente, la funcionalidad de un descodificador de audio que comprende los medios 520 y el calculador de información de control de distorsión de tiempo 530 serán descritos con referencia a las Figuras. 6, 7a, 7b, 8, 9a-9c, 10a-10g, 11a, 11b y 12.

[0055] La Figura 6 muestra un diagrama de flujo de un procedimiento para descodificar una representación codificada de una señal de audio, de acuerdo con una realización de la invención. El procedimiento 600 comprende proveer una información de contorno de distorsión de tiempo reconstruida, en donde la provisión de la información de contorno de distorsión de tiempo reconstruida comprende calcular 610 valores de nodo de distorsión, interpolar 620 entre los valores de nodo de distorsión y re-escalamiento 630 de una o más porciones de contorno de distorsión previamente calculadas y uno o más valores de suma de contorno de distorsión previamente calculados. El procedimiento 600 comprende además calcular 640 información de control de distorsión de tiempo utilizando una "nueva porción de contorno de distorsión de tiempo" obtenida en las etapas 610 y 620, las porciones de contorno de distorsión de tiempo previamente calculadas reescaladas ("porción de contorno de distorsión de tiempo actual" y "última porción de contorno de distorsión de tiempo") y también, opcionalmente, utilizando los valores de suma de contorno de distorsión previamente calculados reescalados. Como resultado, una información de contorno de tiempo y/o una información de posición de muestra y/o una información de longitud de transición y/o una primera porción y información de última posición puede ser obtenida en la etapa 640.

20 [0056] El procedimiento 600 comprende además efectuar 650 la reconstrucción de señal distorsionada en el tiempo utilizando la información de control de distorsión de tiempo obtenida en la etapa 640. detalles concernientes con la reconstrucción de señal de distorsión de tiempo serán descritos subsecuentemente.

[0057] El procedimiento 600 también comprende una etapa 660 de actualizar una memoria como se describirá posteriormente en la presente.

25 Cálculo de las porciones de contorno de distorsión de tiempo

[0058] En lo siguiente, detalles concernientes con el cálculo de las porciones de contorno de distorsión de tiempo serán descritos, con referencia a las Figuras 7a, 7b, 8, 9a, 9b, 9c.

[0059] Se supondrá que un estado inicial está presente, que es ilustrado en una representación gráfica 710 de la Figura 7a. Como se puede ver, una primera porción de contorno de distorsión 716 (porción de contorno de distorsión 1) y una segunda porción de contorno de distorsión 718 (porción de contorno de distorsión 2) están presentes. Cada una de las porciones de contorno de distorsión comprenden comúnmente una pluralidad de valores de datos de contorno de distorsión discretos, que son almacenados comúnmente en una memoria. Los diferentes valores de datos del contorno de distorsión están asociados con los valores de tiempo, en donde un tiempo es mostrado en una absisa 712. La magnitud de los valores de datos de contorno de distorsión son mostrados en las ordenadas 714. Como se puede ver, la primera porción de contorno de distorsión tiene un valor final de uno y la segunda porción de contorno de distorsión tiene un valor de partida de uno, en donde el valor de uno puede ser considerado como un "valor predeterminado". Se debe notar que la primera porción de contorno de distorsión 716 puede ser considerada como la "última porción de distorsión de tiempo" (también designada como "last_warp_contour"), mientras que la segunda porción de contorno de distorsión 718 puede ser considerada como una "porción de contorno de distorsión de tiempo actual" (también denominada como "cur_warp-contour").

[0060] Iniciando del estado inicial, se calcula una nueva porción de contorno de distorsión, por ejemplo, en las etapas 610, 620 del procedimiento 600. Así, los valores de datos de contorno de distorsión de la tercera porción de contorno de distorsión (también designada como "porción de contorno de distorsión 3" o "nueva porción de contorno de distorsión de tiempo" o "new_warp-contour") es calculada. El cálculo puede ser separado, por ejemplo, en un cálculo de valores de nodo de distorsión, de acuerdo con un algoritmo 910 mostrado en la Figuras 9a y una interpolación 620 entre los valores de nodo de distorsión de acuerdo con el algoritmo 920 mostrado en la Figura 9a. Así, se obtiene una nueva porción de contorno de distorsión 722, que inicia a partir del valor predeterminado (por ejemplo, uno) y que es mostrado en una representación gráfica 720 de la Figura 7a. Como se puede ver, la primera porción de contorno de distorsión de tiempo 716, la segunda porción de contorno de distorsión de tiempo 718 y la tercera porción de contorno de distorsión de tiempo nueva están asociadas con intervalos de tiempo subsecuentes y contiguos. Además, se puede ver que hay una discontinuidad 724 entre un punto final 718b de la segunda porción de contorno de distorsión de tiempo 718 y un punto de inicio 722a de la tercera porción de contorno de distorsión de tiempo.

[0061] Se debe notar aquí que la discontinuidad 724 comprende comúnmente una magnitud que es más grande que una variación entre cualesquier dos valores de datos de contorno de distorsión temporalmente adyacentes del contorno de distorsión de tiempo dentro de una porción de contorno de distorsión de tiempo. Esto es debido al hecho de que el valor de inicio 722a de la tercera porción de contorno de distorsión de tiempo 722 es forzada a un valor predeterminado (por ejemplo, uno), independiente del valor final 718b de la segunda porción de contorno de

distorsión de tiempo 718. Se debe notar que la discontinuidad 724 es por consiguiente más grande que la variación inevitable entre dos valores de datos de contorno de distorsión discretos adyacentes.

5 **[0062]** No obstante, esta discontinuidad entre la segunda porción de contorno de distorsión de tiempo 718 y la tercera porción de contorno de distorsión de tiempo 722 sería perjudicial para el uso adicional de los valores de datos de contorno de distorsión de tiempo.

10 **[0063]** Así, la primera porción de contorno de distorsión de tiempo y la segunda porción de contorno de distorsión de tiempo son reescaladas conjuntamente en la etapa 630 del procedimiento 600. Por ejemplo, los valores de datos del contorno de distorsión de tiempo de la primera porción del contorno de distorsión de tiempo 716 y los valores de datos del contorno de distorsión de tiempo de la segunda porción del contorno de distorsión de tiempo 718 son reescalados por multiplicación con un factor de re-escalamiento (también designado como "norm_fac"). Así, se obtiene una versión reescalada 716' de la primera porción de contorno de distorsión de tiempo 716 y también una versión reescalada 718' de la segunda porción de contorno de distorsión de tiempo 718. En contraste, la tercera porción de contorno de distorsión de tiempo es comúnmente dejada sin aceptar en esta etapa de re-escalamiento, como se puede ver en una representación grafica 730 de la Figura 7a. El re-escalamiento puede ser efectuado de tal manera que el punto final reescalado 718b comprende por lo menos aproximadamente el mismo valor de datos como el punto de partida 722a de la tercera porción de contorno de distorsión de tiempo 722. Así, la versión reescalada 716' de la primera porción de contorno de distorsión de tiempo, la versión reescalada 718' de la segunda porción de contorno de distorsión de tiempo y la tercera porción de contorno de distorsión de tiempo 722 forman conjuntamente una sección de contorno de distorsión de tiempo (aproximadamente) continua. En particular, el escalamiento puede ser efectuado de tal manera que la diferencia entre el valor de datos del punto final reescalado 718b y el punto de partida 722 no es mayor que un máximo de la diferencia entre cualesquier dos valores de datos adyacentes de las porciones de contorno de distorsión de tiempo 716', 718', 722.

25 **[0064]** Así, la sección de contorno de distorsión de tiempo aproximadamente continua que comprende las porciones de contorno de distorsión de tiempo escaladas 716', 718' y la porción de contorno de distorsión de tiempo original 722 es usada para el cálculo de la información de control de distorsión de tiempo que es efectuada en la etapa 640. Por ejemplo, la información de control de distorsión de tiempo puede ser calculada para un cuadro de audio asociado temporalmente con la segunda porción de contorno de distorsión de tiempo 718.

30 **[0065]** Sin embargo, después del cálculo de la información de control de distorsión de tiempo en la etapa 640, se puede efectuar una reconstrucción de señal distorsionada en el tiempo en la etapa 650, que será explicada en más detalle posteriormente en la presente.

35 **[0066]** Subsecuentemente, se requiere obtener información de control de distorsión de tiempo para un siguiente cuadro de audio. Para este propósito, la versión reescalada 716' de la primera porción del contorno de distorsión de tiempo puede ser descartada para ahorrar memoria, debido a que ya no se necesita. Sin embargo, la versión reescalada 716' puede naturalmente también ser guardada para cualquier propósito. Además, la versión reescalada 718' de la segunda porción de contorno de distorsión de tiempo toma el lugar de la "última porción de contorno de distorsión de tiempo" para el nuevo cálculo, como se puede ver en una representación gráfica 740 de la Figura 7b. Además, la tercera porción de contorno de distorsión de tiempo 722, que tomó el lugar de la "nueva porción de contorno de distorsión de tiempo" en el cálculo previo, toma el papel de la "porción de contorno de distorsión de tiempo actual" para un siguiente cálculo. La asociación es mostrada en la representación gráfica 740.

40 **[0067]** Subsecuentemente a esta actualización de la memoria (etapa 660 del procedimiento 600), se calcula una nueva porción del contorno de distorsión de tiempo 752 como se puede ver en la representación gráfica 750. Para este propósito, las etapas 610 y 620 del procedimiento 600 pueden ser re-ejecutadas con nuevos datos introducidos. La cuarta porción de contorno de distorsión de tiempo 752 toma el papel de la "nueva porción de del tiempo" por ahora. Como se puede ver, hay comúnmente una discontinuidad entre un punto final 722b de la tercera porción de contorno de distorsión de tiempo y un punto de partida 750a de la cuarta porción de contorno de distorsión de tiempo 752. Esta discontinuidad 754 es reducida o eliminada por un re-escalamiento subsecuente (etapa 630 del procedimiento 600) de la versión reescalada 718' de la segunda porción de contorno de distorsión de tiempo y de la versión original de la tercera porción de contorno de distorsión de tiempo 722. Así, una versión reescalada dos veces 718" de la segunda porción de contorno de distorsión de tiempo y una versión reescalada una vez 722' de la tercera porción de contorno de distorsión de tiempo son obtenidas, como se puede ver de una representación gráfica 760 de la Figura 7b. Como se puede ver, las porciones de contorno de distorsión de tiempo 718", 722', 752 forman una sección de contorno de distorsión de tiempo por lo menos aproximadamente continua, que puede ser usada para el cálculo de la información de control de distorsión de tiempo en la red-ejecución de la etapa 640.

55 **[0068]** Por ejemplo, la información de control de distorsión de tiempo puede ser calculada en base a las porciones de contorno de distorsión de tiempo 718", 722', 752, tal información de control de distorsión de tiempo está asociada con un cuadro de tiempo de señal de audio centrado sobre la segunda porción de contorno de distorsión de tiempo.

[0069] Se debe notar que en algunos casos es deseable tener un valor de suma de contorno de distorsión asociado para cada una de las porciones de contorno de distorsión de tiempo. Por ejemplo, un primer valor de suma de contorno de distorsión puede ser asociado con la primera porción de contorno de distorsión de tiempo, un segundo

valor de suma de contorno de distorsión puede ser asociado con la segunda porción de contorno de distorsión de tiempo y así sucesivamente. Los valores de suma de contorno de distorsión pueden ser usados por ejemplo, para el cálculo de la información de control de distorsión de tiempo en la etapa 640.

5 **[0070]** Por ejemplo, el valor de suma de contorno de distorsión puede representar una suma de los valores de datos de contorno de distorsión de una porción de contorno de distorsión de tiempo respectiva. Sin embargo, ya que las porciones de contorno de distorsión de tiempo son escaladas, es algunas veces deseable también escalar el valor de suma de contorno de distorsión de tiempo, de tal manera que el valor de suma del contorno de distorsión de tiempo sigue la característica de su porción de contorno de distorsión de tiempo asociada. Así, un valor de suma de contorno de distorsión asociado con la segunda porción de contorno de distorsión de tiempo 718 puede ser escalado (por ejemplo, por el mismo factor de escalamiento) cuando la segunda porción de contorno de distorsión de tiempo 718 es escalada para obtener la porción escalada 718' de la misma. Similarmente, el valor de suma de contorno de distorsión asociado con la primera porción de contorno de distorsión de tiempo 716 puede ser escalado (por ejemplo, con el mismo factor de escalamiento) cuando la primera porción de contorno de distorsión de tiempo 716 es escalada para obtener la porción escalada 716' de la misma, si se desea.

15 **[0071]** Además, se puede efectuar una re-asociación (o re-asignación de memoria) cuando se procede a la consideración de una nueva porción de contorno de distorsión de tiempo. Por ejemplo, el valor de suma de contorno de distorsión asociado con la versión escalada 718' de la segunda porción de contorno de distorsión de tiempo, que toma el papel de un "valor de suma de contorno de distorsión de tiempo actual" para el cálculo de la información de control de distorsión de tiempo asociada con las porciones de contorno de distorsión de tiempo 716', 718', 722 puede ser considerado como un "último valor de suma de distorsión de tiempo" para el cálculo de una información de control de distorsión de tiempo asociada con las porciones de contorno de distorsión de tiempo 718", 722', 752. Similarmente, el valor de suma del contorno de distorsión asociado con la tercera porción de contorno de distorsión de tiempo 722 puede ser considerado como un "nuevo valor de suma de contorno de distorsión" para el cálculo de la información de control de distorsión de tiempo asociada con las porciones de contorno de distorsión de tiempo 716', 718', 722 y puede ser mapeada para actuar como "valor de suma de contorno de distorsión actual" para el cálculo de la información de control de distorsión de tiempo asociada con las porciones de contorno de distorsión de tiempo 718", 722', 752. Además, el valor de suma de contorno de distorsión recién calculado de la cuarta porción de contorno de distorsión de tiempo 752 puede tomar el papel del "nuevo valor de suma del contorno de distorsión" para el cálculo de la información del control de distorsión de tiempo asociada con las porciones de contorno de distorsión de tiempo 718", 722', 752.

Ejemplo de acuerdo con la Figura 8

35 **[0072]** La Figura 8 muestra una representación gráfica que ilustra un problema que es resuelto por las realizaciones de acuerdo con la invención. Una primera representación gráfica 810 muestra una evolución temporal de una altura relativa reconstruida con el tiempo, que es obtenida en algunas realizaciones convencionales. Las abcisas 812 describen el tiempo, las ordenadas 814 describen la altura relativa. La curva 816 muestra la evolución temporal de la altura relativa con el paso del tiempo, que podría ser reconstruida a partir de una información de altura relativa. Concerniente con la reconstrucción del contorno de altura relativo, se debe notar que para la aplicación de la transformada de coseno discreta modificada distorsionada en el tiempo (MDCT) solamente el conocimiento de la variación relativa de la altura dentro del cuadro actual es necesaria. Con el fin de entender esto, se hace referencia a las etapas de cálculo para obtener el contorno de tiempo a partir del contorno de altura relativa, que conduce a un contorno de tiempo idéntico para versiones escaladas del mismo contorno de altura relativo. Por consiguiente, es suficiente codificar solamente el valor de altura relativo en lugar de un valor de altura absoluto, lo que incrementa la eficiencia de codificación. Para incrementar adicionalmente la eficiencia, el valor cuantificado real no es la altura relativa, si no el cambio relativo en altura, esto es, la proporción de la altura relativa actual con respecto a la altura relativa previa (como se discutirá en detalle en lo siguiente). En algunos cuadros, en donde por ejemplo, la señal exhibe ninguna estructura armónica, ninguna distorsión de tiempo podría ser deseable. En tales casos, una bandera adicional puede indicar opcionalmente un contorno de altura plano en lugar de la codificación de este contorno plano con el procedimiento mencionado anteriormente. Puesto que en las señales del mundo real la cantidad de tales cuadros es comúnmente suficientemente alta, el intercambio entre los bits adicionales agregados en todos los tiempos y los bits ahorrados para el cuadro sin distorsión están a favor del ahorro de bits.

55 **[0073]** El valor de partida para el cálculo de la variación de altura (contorno de altura relativo o contorno de distorsión de tiempo) puede ser elegido arbitrariamente y aún diferir en el codificador y descodificador. Debido a la naturaleza de los valores de partida diferentes de MDCT distorsionados en el tiempo (TW-MDCT) de la variación de altura todavía produce las mismas posiciones de muestra y formas de ventana adaptadas para efectuar el TW-MDCT.

[0074] Por ejemplo, un codificador (audio) obtiene un contorno de altura para cada nodo que es expresado como retraso de altura real en muestras en conjunción con una especificación con voz/sin voz opcional que fue por

ejemplo, obtenida al aplicar una estimación de altura y decisión con voz/sin voz conocida de la codificación de habla. Si para el nodo actual la clasificación es ajustada a con voz o no está disponible ninguna una decisión de voz/sin voz, el codificador calcula la proporción entre el retraso de altura real y lo cuantifica o solo ajusta la proporción a uno y es sin voz. Otro ejemplo, podría ser que la variación de altura es estimada directamente mediante un procedimiento apropiado (por ejemplo, estimación de variación de señal).

[0075] En el descodificador, el valor de inicio para la primera altura relativa al inicio del audio codificado es ajustado a un valor arbitrario, por ejemplo, a uno. Por consiguiente, el contorno de altura relativo descodificado ya no está en el mismo intervalo absoluto del contorno de altura del codificador, si no una versión escalada del mismo. Todavía, como se describe anteriormente, el algoritmo de TW-MDCT conduce a las mismas porciones de muestra y formas de ventana. Además, el codificador podría decidir si las proporciones de altura codificadas producirían un contorno de altura plano, no enviar el contorno plenamente codificado, si no ajustarla bandera de activePitchData a cero en lugar de esto, ahorrando bits en este cuadro (por ejemplo, ahorrando $\text{numPitchbits} * \text{numPitches}$ bits en este cuadro).

[0076] En lo siguiente, los problemas serán discutidos que ocurren en ausencia de la re-normalización del contorno de altura de la invención. Como se menciona anteriormente, para el TW-MDCT, solamente el cambio de altura relativo dentro de una cierta extensión de tiempo limitado alrededor del bloque actual es necesario para el cálculo de la distorsión de tiempo y la adaptación de forma de ventana correcta (véanse las especificaciones anteriores). La distorsión de tiempo sigue el contorno descodificado para segmentos en donde se ha detectado un cambio de altura y sigue siendo constante en todos los otros casos (véase la representación gráfica 810 de la Figura 8). Para el cálculo de las posiciones de ventana y toma de muestras de un bloque, tres segmentos de contorno de altura relativos consecutivos (por ejemplo, tres porciones de contorno de distorsión de tiempo) son necesarios, en donde el tercero es el recién transmitido en el cuadro (designado como “nueva porción de contorno de distorsión de tiempo”) y los otros dos son guardados en la memoria del pasado (por ejemplo, designados como “última porción del contorno de distorsión de tiempo” y “porción de contorno de distorsión de tiempo actual”).

[0077] Para tener un ejemplo, se hace referencia por ejemplo, a las explicaciones que se hicieron con referencia a las Figuras 7a y 7b y también a las representaciones gráficas 810, 860 de la Figura 8. Para calcular, por ejemplo, las posiciones de toma de muestras de la ventana para (o asociada con) el cuadro 1, que se extiende desde el cuadro 0 al cuadro 2, se necesitan los contornos de altura de (o asociados con) el cuadro 0, 1 y 2. En la corriente de bits, solamente la información de altura para el cuadro 2 es enviada en el cuadro actual y los otros dos son tomados del pasado. Como se explica en la presente, el contorno de altura puede ser conseguido mediante la aplicación de la primera proporción de altura relativa descodificada a la última altura del cuadro 1 para obtener el cuadro en el primer nodo del cuadro 2 y así sucesivamente. Es ahora posible, debido a la naturaleza de la señal, que si el contorno de altura es simplemente proseguido (esto es, si la parte recién transmitida del contorno es anexada a las dos partes existentes sin ninguna modificación) que un traslape del intervalo en el formato del número interno del codificador se presente después de un cierto tiempo. Por ejemplo, una señal podría iniciar con un segmento de características armónicas fuertes y un valor de altura alto al comienzo que es descendente en todo el segmento, conduciendo a una altura relativa decreciente. Luego, un segmento y ninguna información de altura puede seguir, de tal manera que la altura relativa se mantiene constante. Luego, otra vez, una sección armónica puede iniciar con una altura absoluta que es más alta que la última altura absoluta al segmento previo y otra vez yendo hacia abajo. Sin embargo, simplemente se continúa la altura relativa, es la misma como al final del último segmento armónico y avanzará además hacia abajo y así sucesivamente. Si la señal es suficiente fuerte y tiene en sus segmentos armónicos una tendencia global al ir ya sea hacia arriba o hacia abajo (como se muestra en la representación gráfica 810 de la Figura 8), más pronto o más tarde la altura relativa llega a la frontera de un intervalo del formato del número interno. Es bien conocido de la codificación del habla que las señales de habla exhiben por supuesto tal característica. Por consiguiente, no es sorpresa que la codificación de un conjunto concatenado de señales del mundo real que incluyen habla excedan realmente el intervalo de los valores flotantes usados para la altura relativa después de una cantidad relativamente corta de tiempo cuando se usa el procedimiento convencional descrito anteriormente.

[0078] Para resumir, para un segmento de señal de audio (o cuadro) para el cual se puede diseñar una altura, una evolución apropiada del contorno de altura relativo (o contorno de distorsión de tiempo) podría ser determinado. Para segmentos de señal de audio (o cuadro de señal de audio) para los cuales no se puede determinar una altura (por ejemplo, debido a que los segmentos de señal de audio son semejantes a ruido) el contorno de altura relativo (o contorno de distorsión de tiempo) podría ser mantenido constante. Así, si hubiera un equilibrio entre los segmentos de audio con altura incrementada y altura decreciente, el contorno de altura relativo (o contorno de distorsión de tiempo) avanzaría ya sea a un sub-flujo numérico o un desbordamiento numérico.

[0079] Por ejemplo, en la representación gráfica 810, se muestra un contorno de altura relativa para el caso en que hay una pluralidad de porciones de contorno de altura relativa 820a, 820b, 820c, 820d con altura decreciente y algunos segmentos de audio 822a, 822b sin altura, pero ningún segmento de audio con altura incrementada. Así, se puede ver que el contorno de altura relativo 810 avanza a un subflujo numérico (por lo menos bajo circunstancias muy adversas).

[0080] En lo siguiente, se describirá una solución para este problema. Para impedir los problemas mencionados anteriormente, en particular el sub-flujo o desbordamiento numérico, una normalización de contorno de altura relativa

periódica ha sido introducida de acuerdo con un aspecto de la invención. Puesto que el cálculo del contorno de tiempo distorsionado y la forma de ventana solamente dependen del cambio relativo con respecto a los tres segmentos de contorno de altura relativos mencionados anteriormente (también designadas como “porciones de contorno de distorsión de tiempo”), como se explica en la presente, es posible normalizar este contorno (por ejemplo, el contorno de distorsión de tiempo que puede estar compuesto de tres piezas de “porciones de contorno de distorsión de tiempo”) para cada cuadro (por ejemplo, de la señal de audio) nuevamente con el mismo resultado.

[0081] Para esto, la referencia fue por ejemplo, escogida para ser la última muestra del segundo segmento de contorno (también designado como “porción de contorno de distorsión de tiempo”) y el contorno es ahora normalizado (por ejemplo, multiplicativamente en el dominio lineal) de tal manera que esta muestra tiene un valor de 1.0 (véase la representación gráfica 860 de la Figura 8).

[0082] La representación gráfica 860 de la Figura 8 representa la normalización de contorno de altura relativa. Las abscisas 862 muestran en tiempo, subdividido en cuadros (cuadros 0, 1, 2). Las ordenadas 864 describen el valor de contorno de altura relativa. Un contorno de altura relativa antes de la normalización es diseñado con 870 y cubre dos cuadros (por ejemplo, cuadro número 0 y cuadro número 1). Un nuevo segmento de contorno de altura relativa (también designada como “porción de contorno de distorsión de tiempo”) que inicia del valor de partida de contorno de altura relativa predeterminado (o valor de partida del contorno de distorsión de tiempo) es designado con 874. Como se puede ver, el reinicio del nuevo segmento de contorno de altura relativo 874 del valor de partida del contorno de altura relativa predeterminado (por ejemplo, uno) efectúa una discontinuidad entre el segmento de contorno de altura relativa 870 al punto en el tiempo de inicio y el nuevo segmento de contorno de altura relativa 874, que es designado con 878. Esta discontinuidad efectuaría un problema severo para la derivación de cualquier información de control de distorsión de tiempo del contorno y dará como resultado posiblemente distorsiones de audio. Por consiguiente, un segmento de contorno de altura relativa obtenido previamente 870 precedente al reinicio de punto en el tiempo de reinicio es reescalado (o normalizado), para obtener un segmento de contorno de altura relativa reescalado 870'. Las normalizaciones efectuadas de tal manera que en la última muestra del segmento de contorno de altura relativa 870 es escalada al valor de partida de contorno de altura relativa predeterminada (por ejemplo, de 1.0).

Descripción detallada del algoritmo

[0083] En lo siguiente, algunos de los algoritmos efectuados por el decodificador de audio de conformidad con una realización de la invención serán descritos en detalle. Para este propósito, se hará referencia a las Figuras 5, 6, 9a, 9b, 9c y 10a-10g. Además, se hace referencia a la leyenda de elementos de datos, elementos de ayuda y constantes de las Figuras 11a y 11b.

[0084] Generalmente hablando, se puede decir que el procedimiento descrito en la presente puede ser usado para descodificar una corriente de audio que es codificada de acuerdo con una transformada de coseno discreta modificada de distorsión de tiempo. Así, cuando la TW-MDCT es habilitada para la corriente de audio (que puede ser indicada por una bandera, por ejemplo, denominada como bandera “twMdct”, que puede estar comprendida en una información de configuración específica), un banco de filtros distorsionados en el tiempo y conmutación de bloqueo para reemplazar un banco de filtros estándar y conmutación de bloques. Adicionalmente, a la transformada de coseno discreta modificada inversa (IMDCT) el banco de filtros distorsionado en el tiempo y conmutación de bloque contiene un mapeo de dominio de tiempo a dominio de tiempo desde una red en el tiempo espaciada arbitrariamente a la rejilla de tiempo espaciada regularmente de manera normal y adaptación correspondiente de formas de ventana.

[0085] En lo siguiente, el proceso de descodificación será descrito. En una primera etapa, el contorno de distorsiones descodificado. El contorno de distorsión puede por ejemplo, puede ser verificado utilizando índices del libro de códigos del nodo del contorno de distorsión. Los índices del libro de códigos transformados de contorno de distorsión son descodificados, por ejemplo utilizando el algoritmo mostrado en una representación gráfica 910 de la Figura 9a. De acuerdo con el algoritmo, se derivan valores de proporción de distorsión (warp_value_tbl) a partir de los índices de libro de códigos de proporción de distorsión (tw_proporción), por ejemplo utilizando un mapeo definido por la tabla de mapeo 990 de la Figura 9c. Como se puede ver del algoritmo mostrado como el número de referencia 910, los valores de nodo de distorsión pueden ser ajustados a un valor predeterminado constante, si la bandera (tw_data_present) indica que los datos de distorsión de tiempo no están presentes. En contraste, si la bandera indica que datos de distorsión en tiempo están presentes, un primer valor de nodo de distorsión puede ser ajustado al valor de partida del contorno de distorsión de tiempo predeterminado (por ejemplo, uno). Los valores de nodo de distorsión ejemplares (de una porción de contorno de distorsión de tiempo) pueden ser combinados en base a la formación de un producto de valores de proporción de distorsiones de tiempo múltiples. Por ejemplo, un nodo de distorsión de un nodo inmediatamente siguiente al primer nodo de distorsión $i = 0$) puede ser igual a un primer valor de proporción de distorsión (si el valor de partida es uno) o igual a un producto del primer valor de proporción de distorsión y el valor de partida. Valores de nodo de distorsión de tiempo subsecuentes ($i = 2, 3, \dots, \text{num_tw_nodos}$) son calculados al formar un producto de múltiples valores de proporción de distorsión de tiempo (opcionalmente tomando en consideración el valor de partida, si el valor de partida difiere de uno). Naturalmente, el orden de la formación del producto es arbitrario. Sin embargo, es ventajoso derivar un valor de modo de distorsión ($i + 1$)-ésimo a partir de un valor de nodo de distorsión i -ésimo al multiplicar el valor del nodo de distorsión i -ésimo con un solo valor de proporción de distorsión que describe la proporción entre dos valores del nodo subsecuente del contorno de

distorsión de tiempo.

[0086] Como se puede ver del algoritmo mostrado en el número de referencia 910, pueden haber múltiples índices de código de libros de proporción de distorsión para una sola porción de contorno de distorsión de tiempo en un solo cuadro de audio (en donde puede haber una correspondencia de 1 a 1 entre porciones de contorno de distorsión de tiempo y cuadros de audio).

[0087] Para resumir, una pluralidad de valores de nodos de distorsión de tiempo puede ser obtenida para una porción del contorno de distorsión de tiempo dada (o un cuadro de audio dado) en la etapa 610, por ejemplo, utilizando el calculador de valor de nodo de distorsión 544. Subsecuentemente, se puede efectuar una interpolación lineal entre los valores de nodo de distorsión de tiempo (`warp_node_values[i]`). Por ejemplo, para obtener los valores de datos de contorno de distorsión de tiempo de la “nueva porción de contorno de distorsión de tiempo” (`new_warp_contour`) se pueden usar el algoritmo mostrado con el número de referencia 920 en la Figura 9a pueden ser usados. Por ejemplo, el número de muestras de la nueva porción de contorno de distorsión de tiempo es igual a la mitad del número de las muestras de dominio de tiempo de una transformada de coseno discreta modificada inversa. Con respecto a esta cuestión, se debe notar que cuadros de señal de audio adyacentes son desplazados comúnmente (por lo menos aproximadamente) por la mitad del número de las muestras de dominio de tiempo del MDCT o IMDCT. En otras palabras, para obtener el nuevo_ contorno de distorsión [], el `warp_node_values[]` de muestra en muestra (`N_muestras` de largo), los valores de `nodo_de distorsión` son interpolados linealmente entre los nodos espaciados igualmente (`Inter._dist` aparte) utilizando el algoritmo mostrado con el número de referencia 920.

[0088] La interpolación puede ser efectuada por ejemplo, por el interpolador 548 del aparato de la Figura 5 o en la etapa 620 del procedimiento 600.

[0089] Antes de obtener el contorno completo de distorsión para este cuadro (esto es, para el cuadro actualmente bajo consideración) los valores almacenados en memoria del pasado son reescalados, de tal manera que el último valor de distorsión del `past_warp_contour[]` es igual a uno (o cualquier otro valor predeterminado que es preferiblemente igual al valor de partida de la nueva porción de contorno de distorsión de tiempo).

[0090] Se debe notar aquí que el término de “contorno de distorsión del pasado” comprende preferiblemente la “última porción del contorno de distorsión de tiempo” descrita anteriormente y la “porción de contorno de distorsión de tiempo actual” descrita anteriormente. También se debe notar que el “contorno de distorsión del pasado” comprende comúnmente una longitud que es igual al número de muestras de dominio del tiempo del IMDCT, de tal manera que valores del “contorno de distorsión del pasado” son designados con índices entre cero y $2*n_largo-1$. Así, `past_warp_contour[2*n_largo-1]` designa un último valor de designación del “contorno de distorsión del pasado”. Así, un factor de normalización “`norm_fac`” puede ser calculado de acuerdo con la ecuación mostrada con el número de referencia 930 en la Figura 9a. Así, el contorno de distorsión del pasado (que comprende la “última porción de contorno de distorsión de tiempo” y la “porción de contorno de distorsión de tiempo actual”) pueden ser reescalados multiplicativamente de acuerdo con la ecuación mostrada con el número de referencia 932 en la Figura 9a. Además, el “último valor de suma del contorno de distorsión” (`last_warp_sum`) y el “valor de suma de contorno de distorsión actual” (`cur_warp_sum`) pueden ser reescalados multiplicativamente, como se muestra con los números de referencia 934 y 936 de la Figura 9a. El re-escalamiento puede ser efectuado por el re-escalador 550 de la Figura 5 o en la etapa 630 del procedimiento 600 de la Figura 6.

[0091] Se debe notar que la normalización descrita en la presente, por ejemplo en el número de referencia 930, podría luego ser modificada, por ejemplo, al reemplazar el valor de partida “1” por cualquier otro valor predeterminado deseado.

[0092] Al aplicar la normalización, un “`warp_contour[]` completo” también designado como una “sección de contorno de distorsión de tiempo” es obtenida al concatenar el “`past_warp_contour`” y el “`new_warp_contour`”. Así, tres porciones de contorno de distorsión de tiempo (“última porción de contorno de distorsión de tiempo” y “porción de contorno de distorsión de tiempo actual” y “nueva porción de contorno de distorsión de tiempo”) del “contorno de distorsión pleno” que pueden ser aplicados en etapas adicionales de cálculo.

[0093] Además, se calcula un valor de suma de contorno de distorsión (`new_warp_sum`) por ejemplo, como una suma sobre todos los valores de “`new_warp_contour[]`”. Por ejemplo, un nuevo valor de suma de contorno de distorsión puede ser calculado de acuerdo con los algoritmos mostrados en el número de referencia 940 en la Figura 9a.

[0094] En seguida de los cálculos descritos anteriormente, la información de entrada requerida por el calculador de información de control de distorsión de tiempo 330 o por la etapa 640 del procedimiento 600 está disponible. Así, el cálculo 640 de la información de control de distorsión de tiempo puede ser efectuado, por ejemplo, por el calculador de información de control de distorsión de tiempo 530. También, la reconstrucción de señal distorsionada en el tiempo 650 puede ser efectuada por el descodificador de audio. Tanto el cálculo 640 como la reconstrucción de señal distorsionada en el tiempo 650 serán explicados en más detalle posteriormente en la presente.

[0095] Sin embargo, es importante notar que el algoritmo presente procede iterativamente. Es por consiguiente, computacionalmente eficiente actualizar una memoria. Por ejemplo, es posible descartar información acerca de la

última porción de contorno de distorsión de tiempo. Además, es recomendable usar la “porción de contorno de distorsión de tiempo actual” presente como una “última porción de contorno de distorsión de tiempo” en un siguiente ciclo de cálculos. Además, es recomendable usar la presente “nueva porción de contorno de distorsión de tiempo” como una “porción de contorno de distorsión de tiempo actual” en un siguiente signo del cálculo. Esta asignación se puede hacer utilizando la ecuación ilustrada con el número de referencia en la presente 950 en la Figura 9b (en donde $\text{warp_contour}[n]$ describe en la presente la “nueva porción de contorno de distorsión de tiempo presente” para $2 * n_largo \leq 3.n_largo$).

[0096] Asignaciones apropiadas se pueden ver en los números de referencia 952 y 954 en la Figura 9b.

[0097] En otras palabras, las memorias temporales usadas para descodificar el siguiente cuadro pueden ser actualizadas de acuerdo con las ecuaciones mostradas en los números de referencia 950, 952 y 954.

[0098] Se debe notar que la actualización de acuerdo con las ecuaciones 950, 952 y 954 no proporciona un resultado razonable si la información apropiada no es generada para un cuadro previo- así, antes de descodificar el primer cuadro o si el último cuadro fue codificado con un tipo diferente de codificador (por ejemplo, un codificador de dominio de LPC) en el contexto de un codificador cambiado, los estados de memoria pueden ser ajustados de acuerdo con las ecuaciones mostradas con los números de referencia 960, 962 y 964 de la Figura 9b.

Cálculo de información de control de distorsión de tiempo

[0099] En lo siguiente, se describirá brevemente cómo se puede calcular la información de contorno de distorsión de tiempo en base al contorno de distorsión de tiempo (que comprende, por ejemplo, tres porciones de contorno de distorsión de tiempo) y en base a los valores de suma de contorno de distorsión.

[0100] Por ejemplo, es deseable reconstruir un contorno de tiempo utilizando el contorno de distorsión de tiempo. Para este propósito, se puede usar un algoritmo que es mostrado con los números de referencia 1010, 1012 en la Figura 10a. Como se puede ver, el contorno del tiempo mapea un índice i ($0 \leq i < 3.n_largo$) sobre un valor de contorno de tiempo correspondiente. Un ejemplo de tal mapeo es mostrado en la Figura 12.

[0101] En base al cálculo del contorno de tiempo, es comúnmente requerido calcular una posición de muestra ($\text{simple_pos}[]$), que describe posiciones de las muestras distorsionadas en el tiempo en una escala de tiempo lineal. Tal cálculo puede ser efectuado utilizando un algoritmo que es mostrado con el número de referencia 1030 en la Figura 10b. En el algoritmo 1030, se pueden usar funciones auxiliares, que son mostradas con los números de referencia 1020 y 1022 en la Figura 10a. Así, se puede obtener una información acerca del tiempo de muestra.

[0102] Además, algunas longitudes de las transiciones distorsionadas en el tiempo ($\text{warped_trans_len_left}$; $\text{warped_trans_len_right}$) son calculadas, por ejemplo utilizando un algoritmo 1032 mostrado en la Figura 10b. Opcionalmente, las longitudes de transición de distorsión de tiempo pueden ser adaptadas dependiendo del tipo de ventana o una longitud de transformada, por ejemplo utilizando un algoritmo mostrado con el número de referencia 1034 en la Figura 10b. Además, una llamada “primera posición” y una llamada “última posición” pueden ser calculadas en base a las informaciones de longitudes de transición, por ejemplo utilizando un algoritmo mostrado con el número de referencia 1036 en la Figura 10b. Para resumir, las posiciones de muestra y ajuste de longitudes de ventana, que pueden ser efectuadas por el aparato 530 en la etapa 640 del procedimiento 600 serán efectuadas. Del “ $\text{warp_contour}[]$ ” un vector de las posiciones de muestra (“ $\text{simple_pos}[]$ ”) de las muestras distorsionadas en el tiempo sobre una escala de tiempo lineal pueden ser calculadas. Para esto, primero el contorno de tiempo puede ser generado utilizando el algoritmo mostrado en los números de referencia 1010, 1012. Con las funciones auxiliares “ $\text{warp_in_vec}()$ ” y “ $\text{warp_time_inv}()$ ”, mostradas con los números de referencia 1020 y 1022, el vector de posición de muestra (“ $\text{sample_pos}[]$ ”) y las longitudes de transición de (“ $\text{warped_trans_len_left}$ ” y “ $\text{warped_trans_len_right}$ ”) son calculadas, por ejemplo utilizando los algoritmos mostrados en los números de referencia 1030, 1032, 1034 y 1036. Así, se obtiene la información de control de distorsión de tiempo 512.

Reconstrucción de señal distorsionada en el tiempo

[0103] En lo siguiente, la reconstrucción de señal distorsionada en el tiempo, que puede ser efectuada en base a la información de control de distorsión de tiempo será discutida brevemente para poner el cálculo del contorno de distorsión de tiempo en el contexto apropiado.

[0104] La reconstrucción de una señal de audio comprende la ejecución de una transformada de coseno discreta modificada inversa, que no es descrita en la presente en detalle, debido a que es bien conocido para cualquiera experimentado en el arte. La ejecución de la transformada de coseno discreta modificada inversa permite reconstruir muestras de dominio de tiempo distorsionadas en base a un conjunto de coeficientes de dominio de frecuencia. La ejecución del IMDCT puede ser efectuada por ejemplo, de cuadro en cuadro, lo que significa, por ejemplo un cuadro de 2048 muestras de dominio de tiempo distorsionadas es reconstruido en base a un conjunto de 1024 coeficientes de dominio de frecuencia. Para la reconstrucción correcta es necesario que no más de dos ventanas subsecuentes se traslapen. Debido a la naturaleza de TW-MDCT podría ocurrir que una porción distorsionada en el tiempo inversamente de un cuadro se extienda a un cuadro no vecino, violando así el prerrequisito afirmado anteriormente. Por consiguiente, la longitud de desvanecimiento de la forma de ventana necesita ser acortada al calcular los valores

de warped_trans_len_left y warped_trans_len_right apropiados mencionados anteriormente.

[0105] Una operación de formación de ventana y cambio de bloques 650b es luego aplicada a las muestras de dominio de tiempo obtenidas del IMDCT. La formación de ventana y cambio de bloques pueden ser aplicados a las muestras de dominio de tiempo distorsionados provistos por el IMDCT 650a dependiendo de la información de control de distorsión de tiempo, para obtener muestras de dominio de tiempo distorsionadas de ventana. Por ejemplo, dependiendo de la información de "forma_ventana" o elemento, se pueden usar diferentes prototipos de ventana de transformada sobre-muestreadas, en donde la longitud de las ventanas sobre-mestreadas puede ser dada por la ecuación mostrada en el número de referencia 1040 en la Figura 10c. Por ejemplo, para un primer tipo de forma de ventana (por ejemplo, forma de ventana = 1) los coeficientes de ventana son dados por una ventana derivada de "kaiser-Bessel" (KBD) de acuerdo con la definición mostrada en el número de referencia 1042 de la Figura 10c, en donde W' , la "función de ventana de centro de Kaiser-Bessel" es definida como se muestra en el número de referencia 1044 en la Figura 10c.

[0106] De otra manera, cuando se usa una forma de ventana diferente (por ejemplo, si la forma de ventana = 0), se puede emplear una ventana de seno de acuerdo con la definición del número de referencia 1046. Para todas las clases de secuencias de ventana ("secuencias_ventana"), el prototipo usado para la parte de ventana izquierda es determinada por la forma de ventana del bloque previo. La fórmula mostrada con el número de referencia 1048 de la Figura 10c expresa este hecho. Asimismo, el prototipo para la forma de ventana derecha es determinado por la fórmula mostrada con el número de referencia 1050 en la Figura 10c.

[0107] En lo siguiente, se describirá la aplicación de las ventanas descritas anteriormente a las muestras de dominio de tiempo distorsionadas provistas por el IMDCT. En algunas realizaciones, la información para un cuadro puede ser provista por una pluralidad de secuencias cortas (por ejemplo, ocho secuencias cortas). En otras realizaciones, la información para el cuadro puede ser provista utilizando bloques de diferentes longitudes, en donde se puede requerir un tratamiento especial para secuencias de partida, secuencias separadas y/o secuencias de longitudes no estándar. Sin embargo, puesto que la longitud transicional puede ser determinada como se describe anteriormente, puede ser suficiente diferenciar entre cuadros codificados utilizando ocho secuencias cortas (indicadas por una información del tipo de cuadro apropiado "eight_short_sequence") y todos los otros cuadros.

[0108] Por ejemplo, en un cuadro descrito por ocho secuencias cortas, un algoritmo mostrado con el número de referencia 1060 en la Figura 10d puede ser aplicado para la formación de ventanas. En contraste, para cuadros codificados utilizando otra información, un algoritmo es mostrado con el número de referencia 1064 en la Figura 10e puede ser aplicado. En otras palabras, la porción semejante al código C mostrada con el número de referencia 1060 en la Figura 10d describe la formación de ventana y superposición-adición interna de una llamada "ocho secuencias cortas". En contraste, la porción semejante al código C mostrada en el número de referencia 1064 en la Figura 10d describe la formación de ventanas en otros casos.

Re-muestreo

[0109] En lo siguiente, la distorsión de tiempo inversa 650c de las muestras de dominio de tiempo distorsionadas en ventana en dependencia de la información de control de distorsión de tiempo será descrita, mediante lo cual muestras de dominio de tiempo muestreadas regularmente o simplemente muestras de dominio de tiempo son obtenidas mediante un re-muestreo o re-toma de muestras variable en el tiempo. En el re-muestreo variable en el tiempo, el bloque de ventana $z[]$ es re-muestreado de acuerdo con las posiciones muestreadas, por ejemplo utilizando una respuesta de impulso mostrada con el número de referencia 1070 en la Figura 10f. Antes del re-muestreo, el bloque de ventana puede ser relleno con ceros sobre ambos extremos, como se muestra con el número de referencia 1072 en la Figura 10f. El re-muestreo por sí mismo es descrito por la sección de pseudocódigos mostrada con el número de referencia 1074 en la Figura 10f.

Procesamiento de cuadro del post-remuestreador

[0110] En lo siguiente, un post-procesamiento opcional 650d de las muestras de dominio de tiempo será descrito. En algunas realizaciones, el procesamiento de cuadro de post-re-toma de muestras puede ser efectuado dependiendo del tipo de la secuencia de ventana. Dependiendo del parámetro "secuencia_ventana", ciertas etapas de procesamiento adicionales pueden ser aplicadas.

[0111] Por ejemplo, si la secuencia de ventana es una llamada "OCHO_SECUENCIAS_CORTAS", una llamada "SECUENCIA DE INICIO LARGA", una llamada "SECUENCIA DE INICIO DE PARADA", una llamada "SECUENCIA 1152 DE INICIO DE PARADA" seguida por una llamada SECUENCIA_LPD, un post-procesamiento como se muestra en los números de referencia 1080a, 1080b, 1082 puede ser efectuado.

[0112] Por ejemplo, si la siguiente secuencia de ventana es una llamada "SECUENCIA_LPD", se puede calcular una ventana de corrección $W_{corr}(n)$ puede ser calculada como se muestra en el número de referencia 1080a, tomando en cuenta las definiciones mostradas en el número de referencia 1080b. También, la ventana de corrección $W_{corr}(n)$ puede ser aplicada como se muestra con el número de referencia 1082 en la Figura 10g.

[0113] Para todos los otros casos, nada se puede hacer, como se puede ver con el número de referencia 1084 en la

Figura 10g.

Traslape y adición con secuencias de ventana previas

5 [0114] Además, un traslape y adición 650e de las muestras de dominio de tiempo actuales con una o más muestras de dominio de tiempo previas puede ser efectuada. El traslape y adición pueden ser los mismos para todas las secuencias y pueden ser descritos matemáticamente como se muestra con el número de referencia 1086 en la Figura 10g.

Leyenda

10 [0115] Con respecto a las explicaciones dadas, también se hace referencia a la leyenda que es mostrada en las Figuras 11a y 11d. En particular, la longitud de ventana de síntesis N para la transformada inversa es comúnmente una función del elemento de sintaxis “ventana_secuencia” y el contexto algorítmico. Puede por ejemplo ser definida como se muestra en el número de 1190 de la Figura 11b.

Realización de acuerdo con la Figura 13

15 [0116] La Figura 13 muestra un diagrama de bloques esquemático de medios 1300 para proveer una información de contorno de distorsión de tiempo reconstruida que emprende la funcionalidad de los medios 520 descritos con referencia a la Figura 5. Sin embargo, la trayectoria de datos y la memoria temporales son mostrados en más detalle. Los medios 1300 comprenden un calculador de valor de nodo de distorsión 1344, que toma la función del calculador de valor de nodo distorsionado 544. El calculador de valor de nodo de distorsión 1344 recibe un índice de libro de códigos “tw_ratio[]” de la proporción de distorsión como información de proporción de distorsión codificada. El calculador de valor de nodo de distorsión comprende una tabla de valores de distorsión que representan, por ejemplo, el mapeo de un índice de proporción de distorsión de tiempo sobre un valor de proporción de distorsión en el tiempo representado en la Figura 9c. El calculador de valor de nodo de distorsión 1344 puede comprender además un multiplicador para efectuar el algoritmo representado con el número de referencia 910 de la Figura 9a. Así, el calculador del valor de nodo de distorsión proporciona valores de nodo de distorsión “warp_node_values[i]”. Además, los medios 1300 comprenden un interpolador de contorno de distorsión 1348, que toma la función del interpolador 540a, y que puede ser configurado para efectuar el algoritmo mostrado en el numeral de referencia 920 de la Figura 9a, obteniendo mediante esto valores del nuevo contorno de distorsión (“new_warp_contour”). Los medios 1300 comprenden además una nueva memoria temporal de contorno de distorsión 1350, que guarda los valores del nuevo contorno de distorsión (esto es warp_contour [i], con $2 \cdot n_long \leq i < 3 \cdot n_long$). Los medios 1300 comprenden además una memoria temporal/actualizador de contorno de distorsión del pasado 1360, que guarda la “última porción de contorno de distorsión de tiempo” y la “porción de contorno de distorsión de tiempo actual” y actualiza el contenido de memoria en respuesta a un re-escalamiento y en respuesta a la consumación del procesamiento del cuadro actual. Así, la memoria temporal/actualizador de contorno de distorsión del pasado 1360 puede estar en cooperación con el re-escalador de contorno de distorsión del pasado 1370, de tal manera que la memoria temporal/actualizador de contorno de distorsión del pasado y el re-escalador de contorno de distorsión del pasado satisfacen conjuntamente la funcionalidad de los algoritmos 930, 932, 934, 936, 950, 960. Opcionalmente, la memoria temporal/actualizador de contorno de distorsión del pasado 1360 puede también tomar la funcionalidad de los algoritmos 932, 936, 952, 954, 962, 964.

30 [0117] Así, los medios 1300 proveen el contorno de distorsión (“warp_contour”) y opcionalmente también proveen los valores de suma de contorno de distorsión.

Codificador de Señal de Audio de acuerdo con la Figura 14

40 [0118] En lo siguiente, se describirá un codificador de señal de audio de acuerdo con un aspecto de la invención. El codificador de señal de audio de la Figura 14 es designado en su totalidad con 1400. El codificador de señal de audio 1400 está configurado para recibir una señal de audio 1410 y opcionalmente, una información de contorno de distorsión provista externamente 1412 asociada con la señal de audio 1410. Además, el codificador de señal de audio 1400 está configurado para proveer una representación codificada 1440 de la señal de audio 1410.

45 [0119] El codificador de señal de audio 1400 comprende un codificador de contorno de distorsión de tiempo 1420, configurado para recibir una información de contorno de distorsión de tiempo 1422 asociada con la señal de audio 1410 y para proveer una información de contorno de distorsión de tiempo codificada 1424 en base a la misma.

50 [0120] El codificador de señal de audio 1400 comprende además aun procesador de señal de distorsión de tiempo (o codificador de señal de distorsión de tiempo) 1430 que está configurado para recibir la señal de audio 1410 y para proveer, en base a la misma, una representación codificada distorsionada en el tiempo 1432 de la señal de audio 1410, tomando en cuenta la formado de tiempo descrita por la información de distorsión en el tiempo 1422. La representación codificada 1414 de la señal de audio 1410 comprende la información de contorno de distorsión de tiempo codificada 1424 y la representación codificada 1432 del espectro de la señal de audio 1410.

55 [0121] Opcionalmente, el codificador de señal de audio 1400 comprende un calculador de información de contorno de distorsión 1440, que está configurado para proveer la información de contorno de distorsión en el tiempo 1422 en

base a la señal de audio 1410. Alternativamente, sin embargo, la información de contorno de distorsión en el tiempo 1422 puede ser provista en base a la información de contorno de distorsión provista externamente 1412.

[0122] El codificador de contorno de distorsión en el tiempo 1420 puede estar configurado para calcular la proporción entre valores de nodo subsecuentes del contorno de distorsión de tiempo descrito por la información de contorno de distorsión en el tiempo 1422. Por ejemplo, los valores de nodo pueden ser valores de muestra del contorno de distorsión de tiempo representado por la información de contorno de distorsión en el tiempo. Por ejemplo, si la información de contorno de distorsión en el tiempo comprende una pluralidad de valores para cada cuadro de la señal de audio 1410, los valores de nodo de distorsión en el tiempo puede ser un sub-conjunto verdadero de esta información de contorno de distorsión de tiempo. Por ejemplo, los valores de nodo de distorsión en el tiempo pueden ser un sub-conjunto verdadero periódico de los valores de contorno de distorsión de tiempo. Un valor de nodo de contorno de distorsión en el tiempo puede estar presente por N de las muestras de audio, en donde N puede ser mayor o igual a 2.

[0123] El calculador de proporción de valor de nodo de contorno de tiempo puede estar configurado para calcular una proporción entre valores de nodo de distorsión en el tiempo subsecuentes del contorno de distorsión de tiempo, proporcionando así una información que describe una proporción entre valores de nodo subsecuentes del contorno de distorsión de tiempo. Un codificador de proporción del codificador de contorno de distorsión de tiempo puede estar configurado para codificar la proporción entre valores de nodo subsecuentes del contorno de distorsión de tiempo. Por ejemplo, el codificador de proporción puede mapear diferentes proporciones a diferentes índices de libro de códigos diferentes. Por ejemplo, se puede escoger un mapeo de tal manera que las proporciones provistas por el calculador de proporción de valor de distorsión de contorno en el tiempo están dentro de un intervalo de entre 0.9 y 1.1, o aún entre 0.95 y 1.05. Así, el codificador de proporción puede estar configurado para mapear este intervalo a diferentes índices de libro de códigos. Por ejemplo, correspondencias mostradas en la tabla de la Figura 9c pueden actuar como puntos de soporte en este mapeo, de tal manera que, por ejemplo, una proporción de 1 es mapeada sobre un índice de libro de códigos de 3, mientras que una proporción de 1.0057 es mapeada a un índice de libro de código de 4, y así sucesivamente (compárese Figura 9c). Los valores de proporción entre aquellos mostrados en la tabla de la Figura 9c pueden ser mapeados a índices de libro de códigos apropiados, por ejemplo al índice de libro de códigos del valor de proporción más cercano para el cual el índice de libro de códigos es dado en la tabla de la Figura 9c.

[0124] Naturalmente, diferentes codificaciones pueden ser usadas, de tal manera que, por ejemplo, un número de índices de libro de código disponibles puede ser escogidos más grandes o más pequeños que los mostrados en la presente. También, la asociación entre valores de nodo de contorno de distorsión e índices de valores de libro de códigos pueden ser escogidos apropiadamente. También, los índices de libro de códigos pueden ser codificados, por ejemplo utilizando una codificación binaria, utilizando opcionalmente una codificación de entropía.

[0125] Así, se obtienen las proporciones codificadas 1424.

[0126] El procesador de señal de distorsión de tiempo 1430 comprende un convertidor de dominio de tiempo de distorsión de tiempo a dominio de frecuencia 1434, que está configurado para recibir la señal de audio 1410 y una información de contorno de distorsión de tiempo 1422a asociada con la señal de audio (o una versión codificada de la misma), y para proveer, en base a la misma, una representación de dominio espectral (dominio de frecuencia) 1436.

[0127] La información de contorno de distorsión de tiempo 1422a puede ser derivada preferiblemente de la información codificada 1424 provista por y el codificador de contorno de distorsión de tiempo 1420 utilizando un descodificador de distorsión 1425. De esta manera, es puede obtener que el codificador (en particular el procesador de señal de distorsión de tiempo 1430 del mismo) y el descodificador (que recibe la representación codificada 1414 de la señal de audio) operen en los mismos contornos de distorsión, es decir, el contorno de distorsión descodificado (tiempo). Sin embargo, en una realización simplificada, la información de contorno de distorsión de tiempo 1422a usada por el procesador de señal de distorsión de tiempo 1430 puede ser idéntica a la información de contorno de distorsión de tiempo 1422 introducida al codificador de contorno de distorsión de tiempo 1420.

[0128] El convertidor de dominio de tiempo de distorsión de tiempo a dominio de frecuencia 1434 puede considerar por ejemplo una distorsión de tiempo cuando forma la representación de dominio espectral 1436, por ejemplo utilizando una operación de re-muestreo variable en el tiempo de la señal de audio 1410. Alternativamente, sin embargo, el re-muestreo variable en el tiempo y la conversión de dominio en el tiempo a dominio de frecuencia pueden ser integrados en una sola etapa de procesamiento. El procesador de señal de distorsión de tiempo también comprende un codificador de valor espectral 1438, que está configurado para codificar la representación de dominio espectral 1346. El codificador de valor espectral 1438 puede estar configurado por ejemplo para tomar en consideración el enmascaramiento perceptual. También, el codificador de valor espectral 1438 puede estar configurado para adaptar la exactitud de codificación a la relevancia perceptual de las bandas de frecuencia y para aplicar una codificación de entropía. Así, se obtenido la representación codificada 1432 de la señal de audio 1410.

Calculador de contorno de distorsión de tiempo de acuerdo con la Figura 15

[0129] La Figura 15 muestra un diagrama de bloques esquemático de un calculador de contorno de distorsión de tiempo, de acuerdo con otra realización de la invención. El calculador de contorno de distorsión de tiempo 1500 está configurado para recibir una información de proporción de distorsión codificada 1510 para proveer, en base a la misma, una pluralidad de valores de nodo de distorsión 1512. El calculador de contorno de distorsión de tiempo 1500 comprende, por ejemplo, un descodificador de proporción de distorsión 1520, que está configurado para derivar una secuencia de valores de proporción de distorsión 1522 a partir de la información de proporción de distorsión codificada 1510. El calculador de contorno de distorsión de tiempo 1500 también comprende un calculador de contorno de distorsión 1530, que está configurado para derivar la secuencia de valores de nodo de distorsión 1512 a partir de la secuencia de los valores de proporción de distorsión 1522. Por ejemplo, el calculador de contorno de distorsión puede estar configurado para obtener los valores de nodo de contorno de distorsión partiendo de un valor de inicio de contorno de distorsión, en donde las proporciones entre el valor de partida del contorno de distorsión, asociado con un nodo de partida de contorno de distorsión, y los valores de nodo de contorno de distorsión son determinados por los valores de proporción de distorsión 1522. El calculador de valor de nodo de distorsión también está configurado para calcular un valor del nodo de contorno de distorsión 1512 de un nodo de contorno de distorsión dado que está espaciado del nodo de inicio de contorno de distorsión por un nodo de contorno de distorsión intermedio, en base a una formación de producto que comprende una proporción entre el valor de partida de contorno de distorsión (por ejemplo 1) y el valor de nodo de contorno de distorsión del nodo de contorno de distorsión intermedio y una proporción entre el valor de nodo de contorno de distorsión del nodo de contorno de distorsión intermedio y el valor de nodo de contorno de distorsión del nodo de contorno de distorsión dado como factores.

[0130] En lo siguiente, la operación del calculador de contorno de distorsión en el tiempo 1500 será discutida brevemente con referencia a las Figuras 16a y 16b.

[0131] La Figura 16a muestra una representación gráfica de un cálculo sucesivo de un contorno de distorsión de tiempo. Una primera representación gráfica 1610 muestra una secuencia de índices de libro de códigos de proporción de distorsión en el tiempo 1510 (índice=0, índice=1, índice=2, índice=3, índice=7). Además, la representación gráfica 1610 muestra una secuencia de virus de proporción de distorsión (0.983, 0.988, 0.994, 1.000, 1.023) asociados con los índices de libro de códigos. Además, se puede ver que un primer valor de nodo distorsionado 1621 ($i=0$) es escogido como 1 (en donde 1 es un valor de partida). Como se puede ver, un segundo valor de nodo de distorsión 1622 ($i=1$) es obtenido al multiplicar el valor de partida de 1 con el primer valor de proporción de 0.983 (asociado con el primer índice 0). Se puede ver además que el tercer valor de nodo de distorsión 1623 es obtenido al multiplicar el segundo valor de nodo de distorsión 1622 de 0.983 con el segundo valor de proporción de distorsión de 0.988 (asociado con el segundo índice de 1). De la misma manera, el cuarto valor de nodo de formación 1624 es obtenido al multiplicar el tercer valor de nodo de distorsión 1623 con el tercer valor de proporción de distorsión de 0.994 (asociado con un tercer índice de 2).

[0132] Así, se obtiene una secuencia de valores de nodo de distorsión 1621, 1622, 1623, 1624, 1625, 1626.

[0133] Un valor de nodo de distorsión respetivo es obtenido efectivamente de tal manera que es un producto del valor de partida (por ejemplo 1) y todos los valores de proporción de distorsión intermedios que caen entre los nodos de distorsión de partida 1621 y el respectivo valor de nodo de distorsión 1622 a 1626.

[0134] Una representación gráfica 1640 ilustra una interpolación lineal entre los valores de nodo de distorsión. Por ejemplo, los valores interpolados 1621a, 1621b, 1621c podrían ser obtenidos en un descodificador de señal de audio entre dos valores de nodo de distorsión de tiempo adyacentes 1621, 1622, por ejemplo haciendo uso de una interpolación lineal.

[0135] La Figura 16b muestra una representación gráfica de una reconstrucción de contorno de distorsión de tiempo utilizando un reinicio periódico de un valor de partida predeterminado, que puede opcionalmente ser implementado en el calculador de contorno de distorsión de tiempo 1500. En otras palabras, el reinicio repetido o periódico no es un elemento esencial, a condición de que un desbordamiento numérico pueda ser evitado mediante cualquier otra medida apropiada en el lado del codificador o en el lado del descodificador. Como se puede ver, una porción de contorno de distorsión puede iniciar desde un nodo de partida 1660, en donde se pueden determinar los nodos de contorno de distorsión 1661, 1662, 1663, 1664 para este propósito, valores de proporción de distorsión (0.983, 0.988, 0.965, 1.000) pueden ser considerados, de tal manera que los nodos de contorno de distorsión adyacentes 1661 a 1664 de la primera porción de contorno de distorsión de tiempo están separados por proporciones determinadas por estos valores de proporción de distorsión. Sin embargo, una segunda porción de contorno de distorsión en el tiempo adicional puede ser iniciada después de un nodo final 1664 de la primera porción de contorno de distorsión de tiempo (que comprende los nodos 1660-1664) ha sido alcanzado. La segunda porción de contorno de distorsión de tiempo puede iniciar desde un nuevo nodo de partida 1665, que puede tomar el valor de partida predeterminado, independientemente de cualesquier valores de proporción de distorsión. Así, los valores de nodo de distorsión de la segunda porción de contorno de distorsión de tiempo pueden ser calculados partiendo del nodo de partida 1665 de la segunda porción de contorno de distorsión de tiempo en base a los valores de proporción de distorsión de la segunda porción de contorno de distorsión de tiempo. Más tarde, una tercera porción de contorno de

distorsión en el tiempo puede iniciar de un nodo de partida correspondiente 1670, que puede tomar otra vez el valor de partida predeterminado independientemente de cualesquier valores de proporción de distorsión. Así, se tiene un reinicio periódico de las porciones de contorno de distorsión de tiempo. Opcionalmente, una re-normalización repetida puede ser aplicada, como se describe en detalle anteriormente.

5

El codificador de señal de audio de acuerdo con la Figura 17

[0136] En lo siguiente, un codificador de señal de audio de acuerdo con otra realización de la invención será descrito brevemente con referencia a la Figura 17. El codificador de señal de audio 1700 está configurado para recibir una señal de audio de multi-canal 1710 y proveer una representación codificada 1712 de la señal de audio de multi-canal 1710. El codificador de señal de audio 1700 comprende un proveedor de representación de audio codificado 1720, que está configurado para proveer selectivamente una representación de audio que comprende una información de contorno de distorsión común, asociada comúnmente con una pluralidad de canales de audio de la señal de audio de multi-canal, o una representación de audio codificada que comprende información de contorno de distorsión individual, asociada individualmente con los diferentes canales de audio de la pluralidad de canales de audio, dependiendo de la información que describe una similitud o diferencia entre contornos de distorsión asociados con los canales de audio de la pluralidad de canales de audio.

[0137] Por ejemplo, el codificador de señal de audio 1700 comprende un calculador de similitud de contorno de distorsión o calculador de diferencia de contorno de distorsión 1730 configurado para proveer la información 1732 que describe la similitud o diferencia entre contorno de distorsión asociado con los canales de audio. El proveedor de representación de audio codificado comprende, por ejemplo, un codificador de contorno de distorsión de tiempo selectivo 1722 configurado para recibir información de contorno de distorsión de tiempo 1724 (que puede ser provista externamente o que puede ser provista por un calculador de información de contorno de distorsión de tiempo opcional 1734) y la información 1732. Si la información 1732 indica que los contornos de distorsión de tiempo de dos o más canales de audio son suficientemente similares, el codificador de contorno de distorsión de tiempo selectivo 1722 puede estar configurado para proveer una información de contorno de distorsión de tiempo codificada adjunta. La información de contorno de distorsión adjunta puede por ejemplo estar basada en un promedio de la información de contorno de distorsión de dos o más canales. Sin embargo, alternativamente, la información de contorno de distorsión adjunta puede estar basada en una sola información de contorno de distorsión de un solo canal de audio, pero asociada conjuntamente con una pluralidad de canales.

[0138] Sin embargo, si la información 1732 indica que los contornos de distorsión de múltiples canales de audio no son suficientemente similares, el codificador de contorno de distorsión de tiempo selectivo 1722 puede proveer información codificada separada de los diferentes contornos de distorsión de tiempo.

[0139] El proveedor de representación de audio codificado 1720 también comprende un procesador de señal de distorsión de tiempo 1726, que también está configurado para recibir la información de contorno de distorsión de tiempo 1724 y la señal de audio de multi-canal 1710. El procesador de señal de distorsión de tiempo 1726 está configurado para codificar los múltiples canales de la señal de audio 1710. El procesador de señal de distorsión de tiempo 1726 puede comprender diferentes modos de operación. Por ejemplo, el procesador de señal de distorsión de tiempo 1726 puede estar configurado para codificar selectivamente canales de audio individual o conjuntamente codificarlos, aprovechando similitudes de inter-canal. En algunos casos, es preferido que el procesador de señal de distorsión de tiempo 1726 sea capaz de codificar comúnmente múltiples canales de audio que tienen una información de contorno de distorsión en el tiempo común. Hay casos en los cuales un canal de audio izquierdo y un canal de audio derecho exhiben la misma evolución de altura relativa pero tienen de otra manera diferentes características de señal, por ejemplo diferentes frecuencias fundamentales absolutas o diferentes envolventes espectrales. En este caso, no es deseable codificar el canal de audio izquierdo y el canal de audio derecho conjuntamente, debido a la diferencia significativa entre el canal de audio izquierdo y el canal de audio derecho. No obstante, la evolución de altura relativa en el canal de audio izquierdo y el canal de audio derecho puede ser paralela, de tal manera que la aplicación de una distorsión de tiempo común es una solución muy eficiente. Un ejemplo de tal señal de audio es una música polifónica, en donde el contenido de múltiples canales de audio exhiben una diferencia significativa (por ejemplo, son dominadas por diferentes cantantes o instrumentos musicales), pero exhiben una variación de altura similar. Así, la eficiencia de codificación puede ser mejorada significativamente al proveer la posibilidad de tener una codificación conjunta de los contornos de distorsión de tiempo para múltiples canales de audio en tanto que se mantiene la opción de codificar separadamente los espectros de frecuencia de los diferentes canales de audio para los cuales se provee una información de contorno de altura común.

[0140] El proveedor de representación de audio codificado 1720 comprende opcionalmente un codificador de información lateral 1728, que está configurado para recibir la información 1732 y proveer una información lateral que indica si un contorno de distorsión codificado común es provisto para múltiples canales de audio o si contornos de distorsión codificados individuales son provistos para los múltiples canales de audio. Por ejemplo, tal información lateral puede ser provista en forma de una bandera de 1 bit llamada "common_tw".

[0141] Para resumir, el codificador de contorno de distorsión de tiempo selectivo 1722 provee selectivamente representaciones codificadas individuales de los contornos de audio de distorsión de tiempo asociados con múltiples señales de audio, o una representación de contorno de distorsión de tiempo codificada adjunta que representa un solo contorno de distorsión de tiempo adjunto asociado con múltiples canales de audio. El codificador de información lateral 1728 provee opcionalmente una información lateral que indica si las representaciones de contorno de distorsión de tiempo individuales o una representación de contorno de distorsión de tiempo adjunta son provistas. El procesador de señal de distorsión de tiempo 1726 proporciona representaciones codificadas de los múltiples canales de audio. Opcionalmente, una información codificada común puede ser provista para múltiples canales de audio. Sin embargo, comúnmente es aún posible proveer representaciones codificadas individuales de múltiples canales de audio, para los cuales una representación de contorno de distorsión de tiempo común está disponible, de tal manera que diferentes canales de audio que tienen contenido de audio diferente, pero idéntico distorsión de tiempo son representados apropiadamente. Consecuentemente, la representación codificada 1712 comprende información codificada provista por el codificador de contorno de distorsión de tiempo selectivo 1722, y el procesador de señal de distorsión de tiempo 1726 y opcionalmente, el codificador de información lateral 1728.

El decodificador de señal de audio de acuerdo con la Figura 18

[0142] La Figura 18 muestra un diagrama de bloques esquemático de un decodificador de señal de audio de acuerdo con una realización de la invención. El decodificador de señal de audio 1800 está configurado para recibir una representación de señal de audio codificada 1810 (por ejemplo, la representación codificada 1712) y proveer, en base a la misma, una representación descodificada 1812 de la señal de audio de multi-canal. El decodificador de señal de audio 1800 comprende un extractor de información lateral 1820 y un decodificador de distorsión de tiempo 1830. El extractor de información lateral 1820 está configurado para extraer una información de aplicación de contorno de distorsión de tiempo 1822 y una información de contorno de distorsión 1824 de la representación de señal de audio codificada 1810. Por ejemplo, el extractor de información lateral 1820 puede estar configurado para reconocer si una sola información de contorno de distorsión de tiempo común está disponible para múltiples canales de la señal de audio codificada, o si la información de contorno de distorsión de tiempo separada está disponible para múltiples canales. Así, el extractor de información lateral puede proveer tal información de aplicación de contorno de distorsión en el tiempo 1822 (que indica si información de contorno de distorsión de tiempo conjunta o individual está disponible) y la información de contorno de distorsión en el tiempo 1824 (que describe la evolución temporal del contorno de distorsión en el tiempo común (adjunta) o de los contornos de distorsión de tiempo individuales). El decodificador de distorsión de tiempo 1830 puede estar configurado para reconstruir la representación descodificada de la señal de audio de multi-canal en base a la representación de señal de audio codificada 1810, tomando en consideración la distorsión de tiempo descrita por la información 1822, 1824. Por ejemplo, el decodificador de distorsión de tiempo 1830 puede estar configurado para aplicar un contorno de distorsión de tiempo común para descodificar diferentes canales de audio, para los cuales información de dominio de frecuencia codificada individual está disponible. Así, el decodificador de distorsión de tiempo 1830 puede reconstruir por ejemplo diferentes canales de la señal de audio de multi-canal, que comprenden distorsión de tiempo similar o idéntica pero diferente altura.

Corriente de audio de acuerdo con las Figuras 19a a 19e

[0143] En lo siguiente, una corriente de audio será descrita, que comprende una representación codificada de uno o más señal de canales de audio y uno o más contornos de distorsión de tiempo.

[0144] La Figura 19a muestra una representación gráfica de un llamado elemento de corriente de datos "USAC_raw_data_block" que puede comprender un solo elemento de canal (SCE), un par de elementos de canal (CPE) o una combinación de uno o más elementos de canal individuales y/o uno o más pares de elementos de canal.

[0145] El "USAC_raw_data_block" puede comprender comúnmente un bloque de datos de audio codificados, mientras que la información de contorno de distorsión de tiempo individual puede ser provista en un elemento de corriente de datos separado. No obstante, es usualmente posible codificar algo de los datos de contorno de distorsión de tiempo al "USAC_raw_data_block".

[0146] Como se puede ver de la Figura 19b, un solo elemento de canal comprende comúnmente una corriente de canal de dominio de frecuencia ("fd_channel_stream"), que será explicado en detalle con referencia a la Figura 9d.

[0147] Como se puede ver de la Figura 19c, un par de elementos de canal ("channel_pair_element") comprende comúnmente una pluralidad de corrientes de canal de dominio de frecuencia. También, el par de elementos de canal puede comprender información de distorsión de tiempo. Por ejemplo, una bandera de activación de distorsión de tiempo ("tw_MDCT") que puede ser transmitida en un elemento de corriente de datos de configuración o en el "USAC_saw_data_block" determina si la información de distorsión en el tiempo está incluida en el par de elementos de canal. Por ejemplo, si la bandera "tw_MDCT" indica que la distorsión de tiempo está activa, el par de elementos de canal puede comprender una bandera ("common_tw") que indica si hay una distorsión de tiempo común para los canales de audio del par de elementos de canal. Si la bandera (common_tw) indica que hay una distorsión de tiempo común para múltiples de los canales de audio, entonces una información de distorsión en el tiempo común (tw_data)

está incluida en el par de elementos de canal, por ejemplo, separados de las corrientes de canal de dominio de frecuencia.

[0148] Con referencia ahora a la Figura 19d, se describe la corriente de canal de dominio de frecuencia. Como se puede ver de la Figura 19d, la corriente de canal de dominio de frecuencia, por ejemplo, comprende una información de ganancia global. También, la corriente de canal de dominio de frecuencia comprende los datos de distorsión de tiempo, si la distorsión de tiempo está activa (bandera "tw_MDCT" activa) y si no hay una información de distorsión en el tiempo común para múltiples canales de señal de audio (bandera "common_tw" está inactiva).

[0149] Además, una corriente de canal de dominio de frecuencia también comprende datos de factor de escala ("scale_factor_data") y datos espectrales codificados (para datos espectrales codificados por ejemplo aritméticamente "ac_espectral_data").

[0150] Con referencia ahora a la Figura 19e, la sintaxis de los datos de distorsión de tiempo son discutidos brevemente. Los datos de distorsión de tiempo pueden comprender por ejemplo, opcionalmente una bandera (por ejemplo, "tw_data_present" o "active Pitch Data") que indica si los datos de distorsión de tiempo están presentes. Si los datos de distorsión de tiempo están presentes, (esto es, el contorno de distorsión de tiempo no es plano) los datos de distorsión de tiempo pueden comprender una secuencia de una pluralidad de valores de proporción de distorsión de tiempo codificados (por ejemplo, "tw_ratio [i]" o "pitchIdx[i]"), que pueden por ejemplo ser codificados de acuerdo con la tabla de libro de códigos de la Figura 9c.

[0151] Así, los datos de distorsión de tiempo pueden comprender una bandera que indica que no hay datos de distorsión de tiempo disponibles, que pueden ser ajustados por un codificador de señal de audio, si el contorno de distorsión de tiempo es constante (las proporciones de distorsión de tiempo son aproximadamente iguales a 1.000). En contraste, si el contorno de distorsión de tiempo es variable, las proporciones entre nodos de contorno de distorsión de tiempo subsecuentes pueden ser codificados utilizando los índices de libro de códigos que compone la información de "tw_ratio".

Conclusión

[0152] Resumiendo lo anterior, las realizaciones de acuerdo con la invención efectúan diferentes mejoras en el campo de distorsión de tiempo.

[0153] Los aspectos de la invención descritos en la presente están en el contexto de un codificador de transformada de MDCT distorsionada en el tiempo (véase, por ejemplo, referencia [1]). Las realizaciones de acuerdo con la invención proveen procedimientos para un desempeño mejorado de un codificador de transformada de MDCT distorsionada en el tiempo.

[0154] De acuerdo con un aspecto de la invención, se provee un formato de corriente de bits particularmente eficiente. La descripción de formato de corriente de bits está basada sobre y mejora la sintaxis de corriente de bits AAC de MPEG-2 (véase, por ejemplo, referencia [2]), pero es por supuesto aplicable a todos los formatos de corrientes de bits con un encabezado de descripción general al inicio de una corriente y una sintaxis de información de cuadro en cuadro individual.

[0155] Por ejemplo, la siguiente información lateral puede ser transmitida en la corriente de bits:

[0156] En general, una bandera de un bit (por ejemplo, llamada "tw_MDCT") puede estar presente en la configuración específica de audio general (GASC), que indica si la distorsión de tiempo está activa o no. Los datos de altura pueden ser transmitidos utilizando la sintaxis mostrada en la Figura 19e o la sintaxis mostrada en la Figura 19f. En la sintaxis mostrada en la Figura 19f, el número de alturas ("numPitches") puede ser igual a 16, y el número de bits de altura en ("numPitchBits") puede ser igual a 3. En otras palabras, pueden haber 16 valores de proporción de distorsión codificados por porción de contorno de distorsión de tiempo (o por cuadro de señal de audio), y cada valor de proporción de contorno de distorsión puede ser codificado utilizando 3 bits.

[0157] Además, en un solo elemento de canal (SCE) los datos de altura (pitch_data[]) pueden estar ubicados antes de los datos de sección en el canal individual, si la distorsión está activa.

[0158] En un par de elementos de canal (CPE), una bandera de altura común señala si hay datos de altura comunes para ambos canales, lo que sigue después que, si no, los contornos de altura individuales son encontrados en los canales individuales.

[0159] En lo siguiente, se dará un ejemplo para un par de elementos de canal. Un ejemplo podría ser una señal de una sola fuente de sonido armónica, colocada dentro del panorama estéreo. En este caso, los contornos de altura relativos para el primer canal y el segundo canal serán iguales o diferirían solamente ligeramente debido a algunos errores pequeños en la estimación de la variación. En este caso, el codificador puede decidir que en lugar de enviar dos contornos de altura codificados separados para cada canal, enviar solamente un contorno de altura que es el promedio del contorno de altura del primero y segundo canal y usar el mismo contorno en la aplicación del TW-MDCT sobre ambos canales. Por otra parte, podría haber una señal en donde la estimación del contorno de altura

produce resultado diferentes para el primero y segundo canal respectivamente. En este caso, los contornos de altura codificados individualmente son enviados dentro del canal correspondiente.

5 **[0160]** En lo siguiente, se describirá una decodificación ventajosa de datos de contorno de altura de acuerdo con un aspecto de la invención. Por ejemplo, si la bandera "active PitchData" es 0, el contorno de altura es ajustado a 1 para todas las muestras en el cuadro, de otra manera los nodos de contorno de altura individuales son calculados como sigue:

- hay numPitches + 1 nodos,

- nodo [0] es siempre 1.0;

10 - nodo [i]=nodo[i-1]•relChange[i] (i=1..numPitches+1), en donde el relChange es obtenido por la cuantificación inversa del pitchIdx[i].

[0161] El contorno de altura es luego generado por la interpolación lineal entre los nodos, en donde las posiciones de muestra de nodo son 0:frameLen/numPitches:frameLen.

Alternativas de Implementación

15 **[0162]** Dependiendo de ciertos requerimientos de implementación, las realizaciones de la invención pueden ser implementadas en elementos físicos o en elementos de programación. La implementación puede ser efectuada utilizando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un CD, un ROM, un PROM, un EPROM, un EEPROM o una memoria instantánea, que tiene señales de control que se pueden leer electrónicamente guardadas en el mismo, que cooperan (o son aptas de cooperar) con un sistema de computadora programable de tal manera que el procedimiento respectivo es ejecutado.

20 **[0163]** Algunas realizaciones de acuerdo con la invención comprenden un portador de datos que tiene señales de control que se pueden leer electrónicamente, que son aptas de cooperar con un sistema de computadora programable, de tal manera que uno de los procedimientos descritos en la presente es efectuado.

25 **[0164]** En general, las realizaciones de la presente invención pueden ser implementadas como un producto de programa de computadora con un código de programa, el código de programa es operativo para efectuar uno de los procedimientos cuando el producto de programa de computadora se ejecuta en una computadora. El código de programa puede por ejemplo ser guardado en un portador que se puede leer por una máquina.

[0165] Otras realizaciones comprenden el programa de computadora para efectuar uno de los procedimientos descritos en la presente, guardados sobre un portador que se puede leer por una máquina.

30 **[0166]** En otras palabras, una realización del procedimiento de la invención es, por consiguiente, un programa de computadora que tiene códigos de programa para efectuar uno de los procedimientos descritos en la presente, cuando el programa de computadora se ejecuta en una computadora.

[0167] Una realización adicional del procedimiento de la invención es, por consiguiente, un portador de datos (o un medio de almacenamiento digital o un medio que se puede leer por computadora) que comprende, grabado sobre el mismo, el programa de computadora para efectuar uno de los procedimientos descritos en la presente.

35 **[0168]** Una realización adicional del procedimiento de la invención es, por consiguiente, una corriente de datos o una secuencia de señales que representan el programa de computadora para efectuar uno de los procedimientos descritos en la presente. La corriente de datos o la secuencia de señales pueden por ejemplo estar configuradas conexión de comunicación de datos, por ejemplo vía Internet.

40 **[0169]** Una realización adicional comprende medios de procesamiento, por ejemplo una computadora, o un dispositivo lógico programable, configurado para ser transferidas vía una o apto para efectuar uno de los procedimientos descritos en la presente.

[0170] Una realización adicional comprende una computadora que tiene instalado en la misma el programa de computadora para efectuar uno de los procedimientos descritos en la presente.

45 **[0171]** En algunas realizaciones, se puede usar un dispositivo lógico programable (por ejemplo un arreglo de compuertas programable en el campo) para efectuar algunas o toda las funcionalidades de los procedimientos descritos en la presente. En algunas realizaciones, un arreglo de compuerta programable en el campo puede cooperar con un microprocesador con el fin de efectuar uno de los procedimientos descritos en la presente.

Referencias

[0172]

50 [1] L. Villemoes, "Time Warped Transform Coding of Audio Signals", PCT/EP2006/010246, Int. patent application,

November 2005 [2] Generic Coding of Moving Pictures and Associated Audio: Advanced Audio Coding. International Standard 13818-7, ISO/IECJTC1/SC29/WG11 Moving Pictures Expert Group, 1997

REIVINDICACIONES

1. Un descodificador de señal de audio (200; 300; 1500; 1800) para proveer una representación de señal de audio de multi-canal descodificada (232;312;1812) en base a una representación de señal de audio de multi-canal codificada (211,212;310;1810), comprendiendo el descodificador de señal de audio:
- 5 un descodificador de distorsión del tiempo (210,216,218,219,220, 230,240;340;1830) configurado para usar selectivamente contornos de distorsión del tiempo específicos de canal de audio individuales (332;1824) o un contorno de distorsión del tiempo de multi-canal adjunto (332;1824) para la reconstrucción de una pluralidad de canales de audio representados por la representación de señal de audio de multi-canal codificada.
- 10 **2.** El descodificador de señal de audio (200;300;1800) de conformidad con la reivindicación 1, en el que el descodificador de distorsión del tiempo (210,216,218,219,220, 230,240;340;1830) está configurado para usar selectivamente un contorno de distorsión del tiempo de multi-canal adjunto (332;1824) para una reconstrucción de distorsión en el tiempo de una pluralidad de canales de audio representados por la representación de señal de audio de multi-canal codificada para la cual información del dominio espectral codificada individual (211) está disponible.
- 15 **3.** El descodificador de señal de audio (200;300;1800) de conformidad con la reivindicación 2, caracterizado porque el descodificador de distorsión del tiempo (210,216,218,219,220, 230,240;340;1830) está configurado para recibir una primera información del dominio espectral asociada con un primero de los canales de audio y para proveer, a partir de esta, una representación del dominio de tiempo del primer canal de audio utilizando una transformación del dominio de frecuencia a dominio de tiempo distorsionado;
- 20 en donde el descodificador de distorsión del tiempo está configurado además para recibir una segunda información del dominio espectral codificada, asociada con un segundo de los canales de audio y para proveer, en base a lo mismo, una representación del dominio de tiempo distorsionada del segundo canal de audio utilizando una transformación del dominio de frecuencia a dominio de tiempo;
- 25 en donde la segunda información del dominio espectral es diferente de la primera información del dominio espectral y
- en donde el descodificador de distorsión del tiempo está configurado para re-muestrear variablemente en el tiempo, en base al contorno de distorsión del tiempo de multi-canal adjunto, la representación del dominio de tiempo distorsionada del primer canal de audio o una versión procesada de la misma, para obtener una representación muestreada regularmente del primer canal de audio y
- 30 para re-muestrear de manera variable en el tiempo, en base al contorno de distorsión del tiempo de multi-canal adjunto, la representación del dominio de tiempo distorsionada del segundo canal de audio o una versión procesada del mismo, para obtener una representación muestreada regularmente del segundo canal de audio.
- 35 **4.** El descodificador de señal de audio (200;300;1800) de conformidad con cualquiera de las reivindicaciones 1 a 3, en el que el descodificador de distorsión del tiempo está configurado para derivar un contorno de tiempo de multi-canal adjunto a partir de la información de contorno de distorsión del tiempo de multi-canal adjunta y
- derivar una primera forma de ventana específica de canal individual asociada con el primero de los canales de audio en base a la primera información de forma de ventana codificada y
- derivar una segunda forma de ventana específica de canal individual asociada con el segundo de los canales de audio en base a la segunda información de forma de ventana codificada y
- 40 aplicar la primera forma de ventana a la representación del dominio de tiempo distorsionada del primer canal de audio, para obtener una versión procesada de la representación del dominio de tiempo distorsionada del primer canal de audio y
- aplicar la segunda forma de ventana a la representación del dominio de tiempo distorsionada del segundo canal de audio, para obtener una versión procesada de la representación del dominio de tiempo distorsionada del segundo canal de audio;
- 45 en donde el descodificador de distorsión del tiempo es apto de aplicar diferentes formas de ventana a las representaciones del dominio de tiempo distorsionadas del primero y segundo canal de audio de un cuadro dado dependiendo de la información de forma de ventana específica de canal individual.
- 50 **5.** El descodificador de señal de audio (200;300;1800) de conformidad con la reivindicación 4, en el que el descodificador de distorsión del tiempo está configurado para aplicar un escalamiento de tiempo común, que es determinado por el contorno de tiempo multi-canal adjunto a diferentes formas de ventana cuando se forman en ventanas las representaciones del dominio de tiempo distorsionadas de los primeros y segundos canales de

audio.

6. Un codificador de señal de audio (100;1700) para proveer una representación codificada (150,152;1712) de una señal de audio de multi-canal, comprendiendo el codificador de señal de audio:

5 un proveedor de representación de audio codificada configurado (104,106,108,108a,114;1720) para proveer selectivamente una representación de audio codificada (150,152;1712) que comprende una información de contorno de distorsión del tiempo de multi-canal común, asociada comúnmente con una pluralidad de canales de audio de la señal de audio de multi-canal o una representación de audio codificada que comprende información del contorno de distorsión del tiempo individual, asociada individualmente con los diferentes canales de audio de la pluralidad de canales de audio, dependiendo de la información que describe una similitud de diferencias entre los contornos de distorsión del tiempo asociados con los canales de audio de la pluralidad de canales de audio.

10 **7.** El descodificador de señal de audio (100;1700) de conformidad con la reivindicación 6, en el que el proveedor de representación de audio codificada (104, 106, 108, 108a, 114; 1720) está configurado para aplicar selectivamente la información de contorno de distorsión del tiempo de multi-canal común para obtener una versión distorsionada en el tiempo de un primero de los canales de audio y para obtener una versión distorsionada en el tiempo de un segundo de los canales de audio y proveer una primera información de dominio espectral codificada individual asociada con un primero de los canales de audio, en base a la versión distorsionada en el tiempo del primer canal de audio y proveer una segunda información del dominio espectral codificada individual asociada con un segundo de los canales de audio, en base a la versión distorsionada en el tiempo del segundo canal de audio.

15 **8.** El descodificador de señal de audio (100;1700) de conformidad con la reivindicación 6 o 7, en el que el proveedor de representación de audio codificado (104,106,108,108a,114;1720) está configurado para proveer la representación codificada (150,152;1712) de la señal de audio de multi-canal, de tal manera que la representación codificada de la señal de multi-canal comprende la información de contorno de distorsión del tiempo de multi-canal común, una representación espectral codificada de una versión distorsionada en el tiempo de una primera señal de canal de audio, distorsionada en el tiempo de acuerdo con la información de contorno de distorsión del tiempo de multi-canal común y una representación espectral codificada de una versión distorsionada en el tiempo de una segunda señal de audio de canal, distorsionada en el tiempo de acuerdo con la información de contorno de distorsión del tiempo de multi-canal común.

20 **9.** El descodificador de señal de audio (100;1700) de conformidad con cualquiera de las reivindicaciones 6 a 8, en el porque el codificador de señal de audio está configurado para obtener la información de contorno de distorsión en el tiempo de multi-canal común, de tal manera que la información de contorno de distorsión del tiempo de multi-canal común representa un promedio de contornos de distorsión individuales asociados con el primer canal de señal de audio y el segundo canal de señal de audio.

25 **10.** El descodificador de señal de audio (100;1700) de conformidad con cualquiera de las reivindicaciones 6 a 9, en el que el proveedor de representación de audio codificado está configurado para proveer una información lateral (tw_data_present; common_tw) dentro de la representación codificada de la señal (150;152;1712) de audio de multi-canal, la información lateral indica, en una base por cuadro de audio, si datos de distorsión del tiempo están presentes para un cuadro de audio dado y si una información de contorno de distorsión del tiempo común está presente para el cuadro de audio dado.

30 **11.** Una representación de señal de audio de multi-canal codificada (usac_raw_data_block) que representa una señal de audio de multi-canal, comprendiendo la representación de señal de audio de multi-canal:

35 una representación del dominio de frecuencia codificada (fd_channel_stream) que representa una pluralidad de canales de audio distorsionados en el tiempo, selectivamente distorsionados en el tiempo, de acuerdo con una distorsión del tiempo común en función de una información que describe una similitud o diferencia entre contornos de distorsión en el tiempo asociados con los canales de audio de la señal de audio multi-canal; y

40 una representación codificada (tw data) de una información de contorno de distorsión del tiempo de multi-canal común, asociada comúnmente con los canales de audio y que representa la distorsión del tiempo común.

45 **12.** La representación de canal de audio codificada (usac_raw_data_block) de conformidad con la reivindicación 11, en la que la representación del dominio de frecuencia codificada comprende información del dominio de frecuencia codificada individual (fd_channel_stream) de múltiples canales de audio que tienen diferentes contenidos de audio y en donde la representación codificada (tw data) de la información de contorno de distorsión del tiempo de multi-canal común está asociada con los múltiples canales de audio que tienen diferente contenido de audio.

13. Un procedimiento para proveer una representación de señal de audio de multi-canal descodificada (232;300;1500;1800) en base a una representación de señal de audio de multi-canal codificada (211,212;310;1810), comprendiendo el procedimiento:

5 utilizar selectivamente contornos de distorsión de tiempo específico de canal de audio individuales o un contorno de distorsión del tiempo de multi-canal adjunto para una reconstrucción de la pluralidad de canales de audio representados por la representación de señal de audio de multi-canal codificada.

14. Un procedimiento para proveer una representación codificada (150,152;1712) de una señal de audio de multi-canal, comprendiendo el procedimiento: proveer selectivamente una representación de audio codificada que comprende una información de contorno de distorsión del tiempo de multi-canal común, asociada comúnmente con una pluralidad de canales de audio con la señal de audio de multi-canal o una representación de audio codificada que comprende información de contorno de distorsión del tiempo individual, asociada individualmente con los diferentes canales de audio de la pluralidad de canales de audio, dependiendo de la información que describe la similitud o diferencia entre contornos de distorsión del tiempo asociados con los canales de audio de la pluralidad de señales de audio.

15 **15.** Un programa de computadora adaptado para efectuar el procedimiento de conformidad con la reivindicación 13 o la reivindicación 14, en el que el programa de computadora se ejecuta en una computadora.

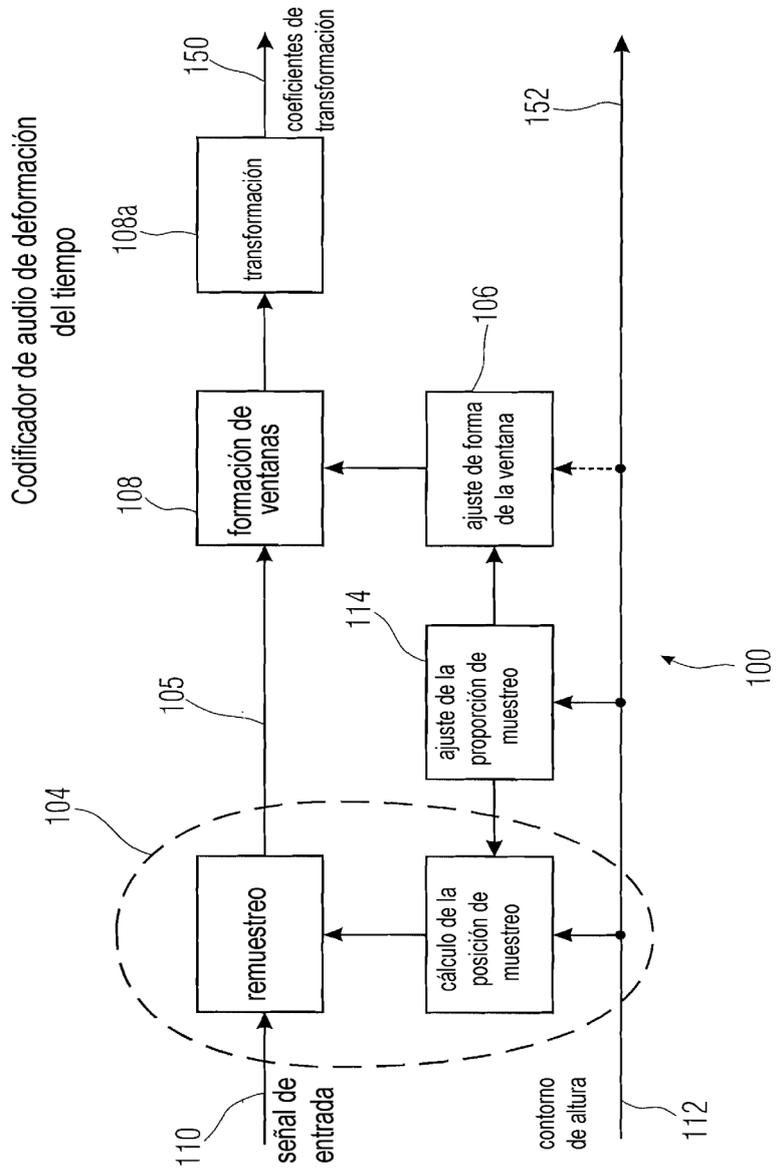


FIG 1

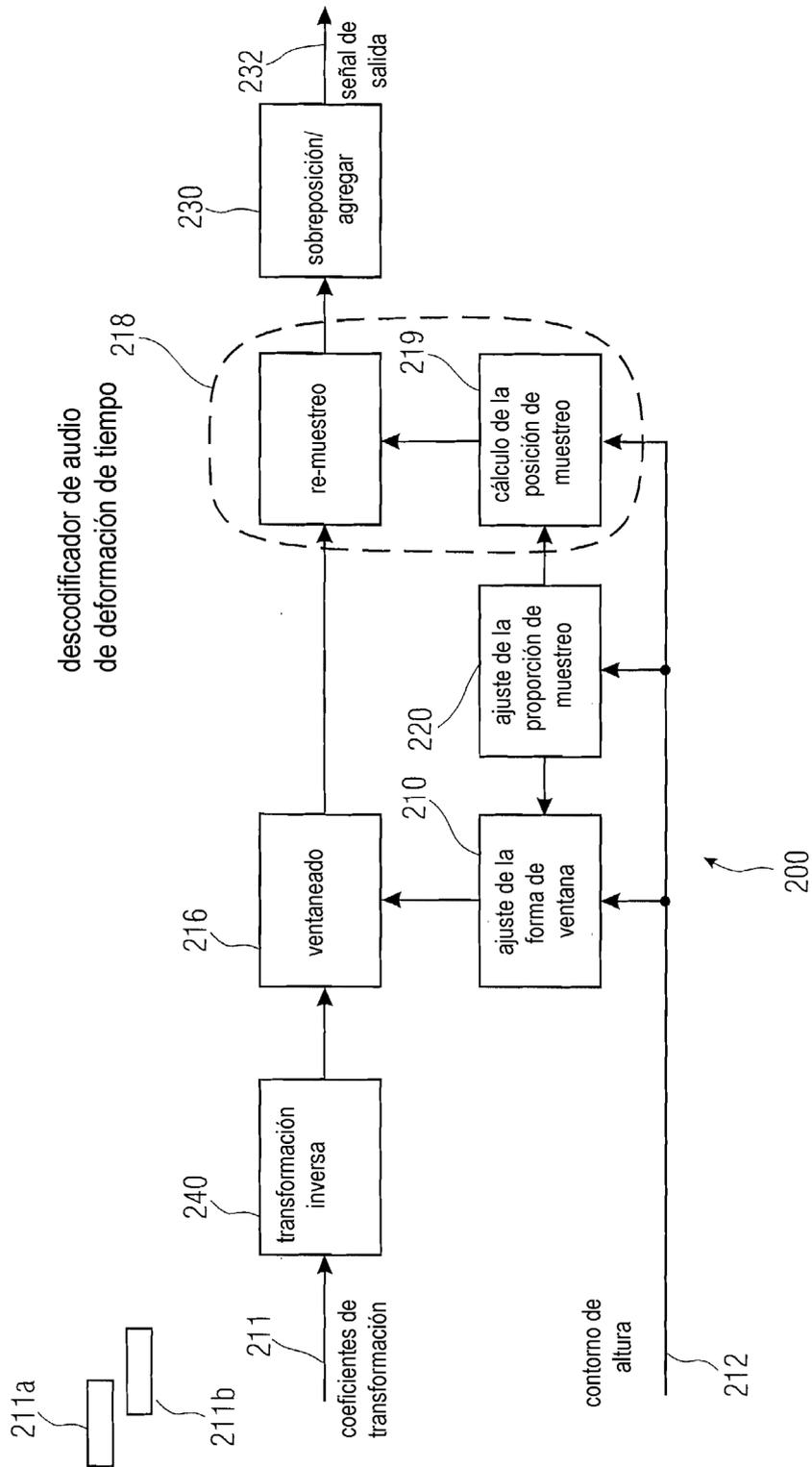


FIG 2

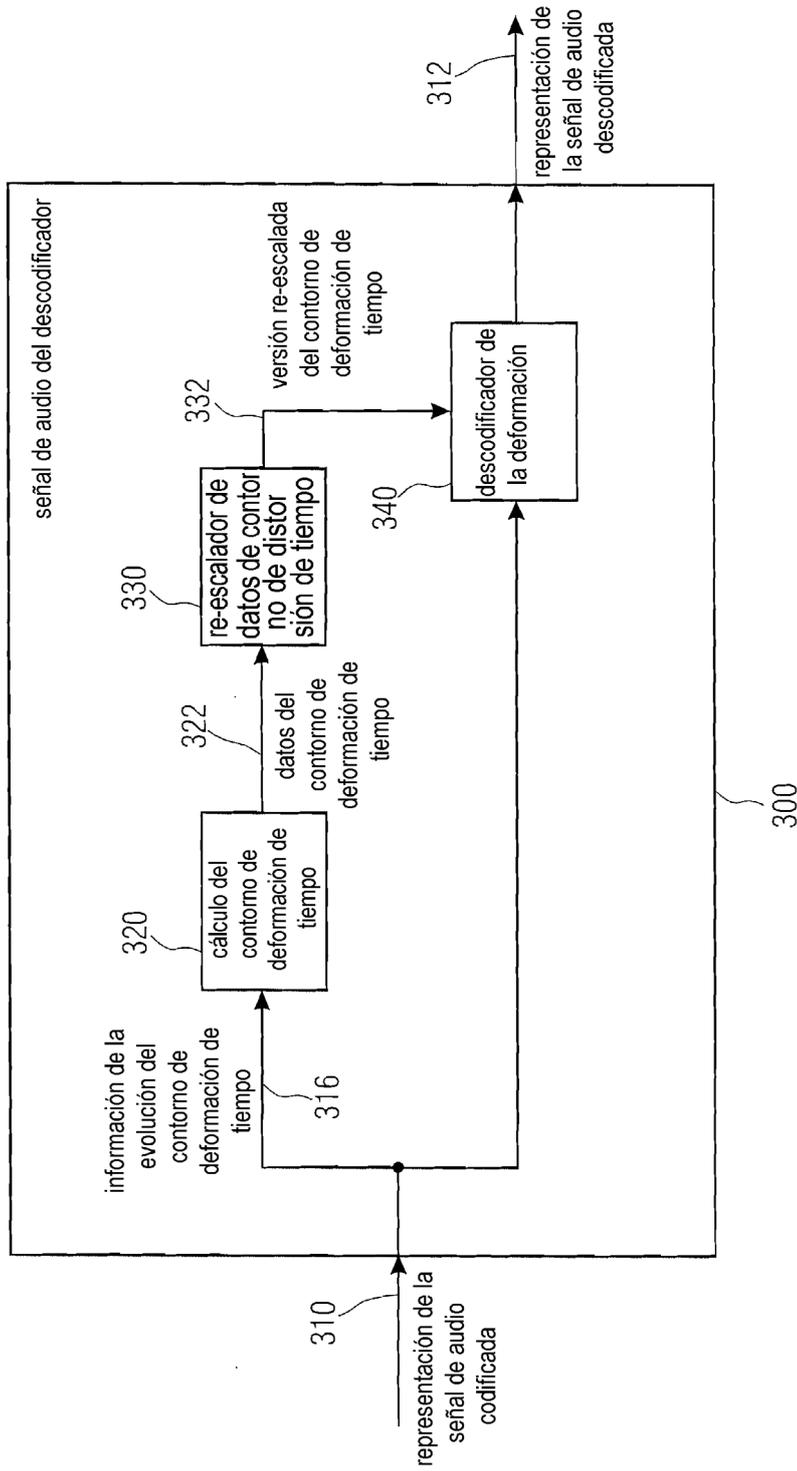


FIG 3

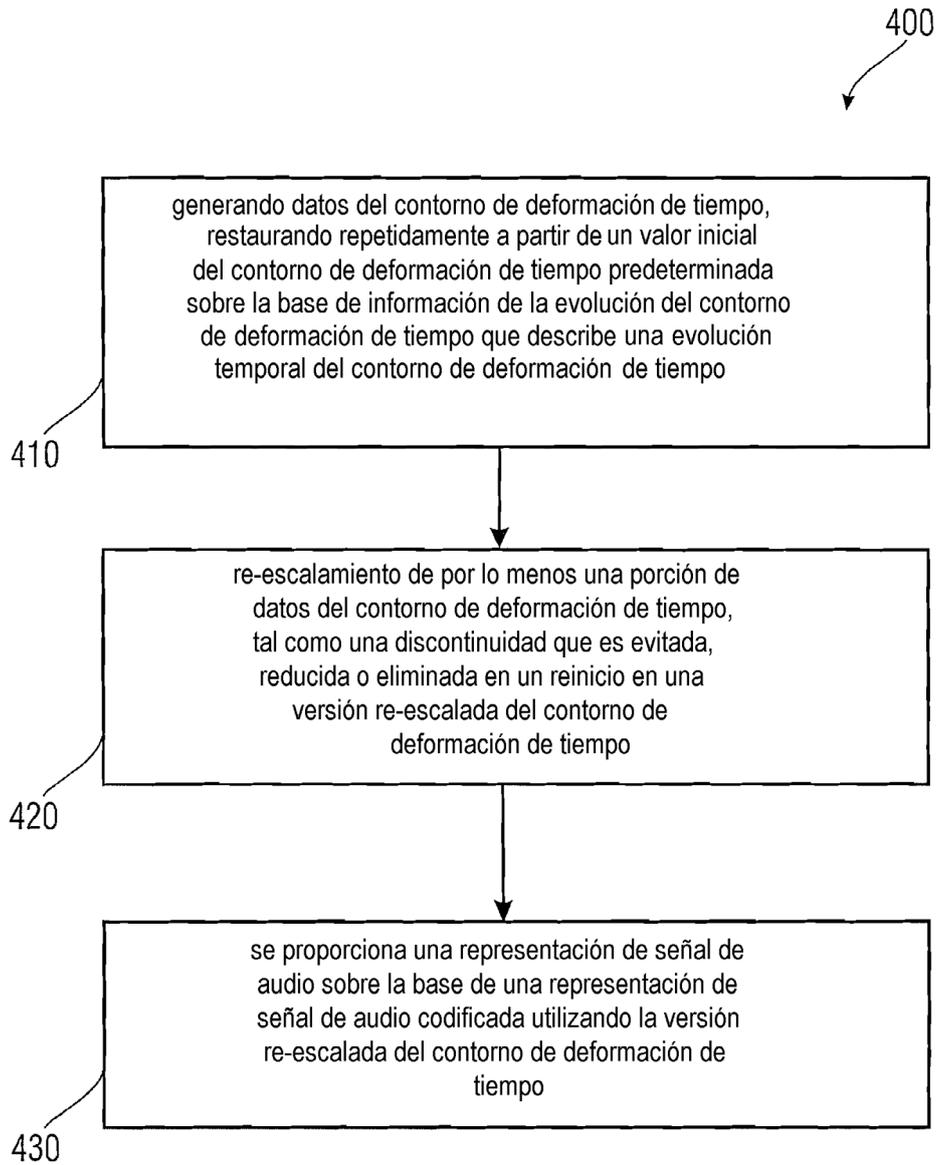


FIG 4

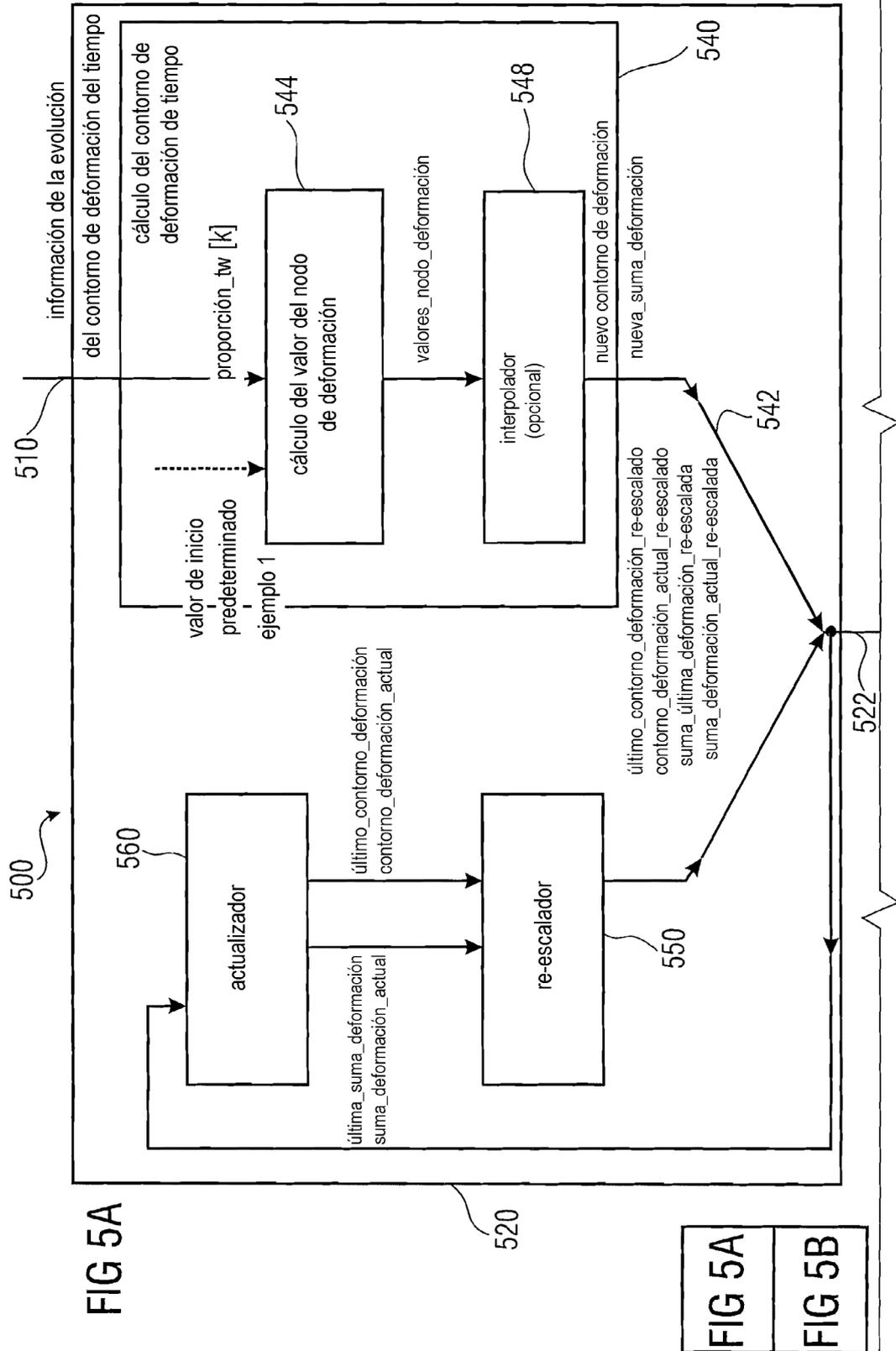


FIG 5A

FIG 5A

FIG 5B

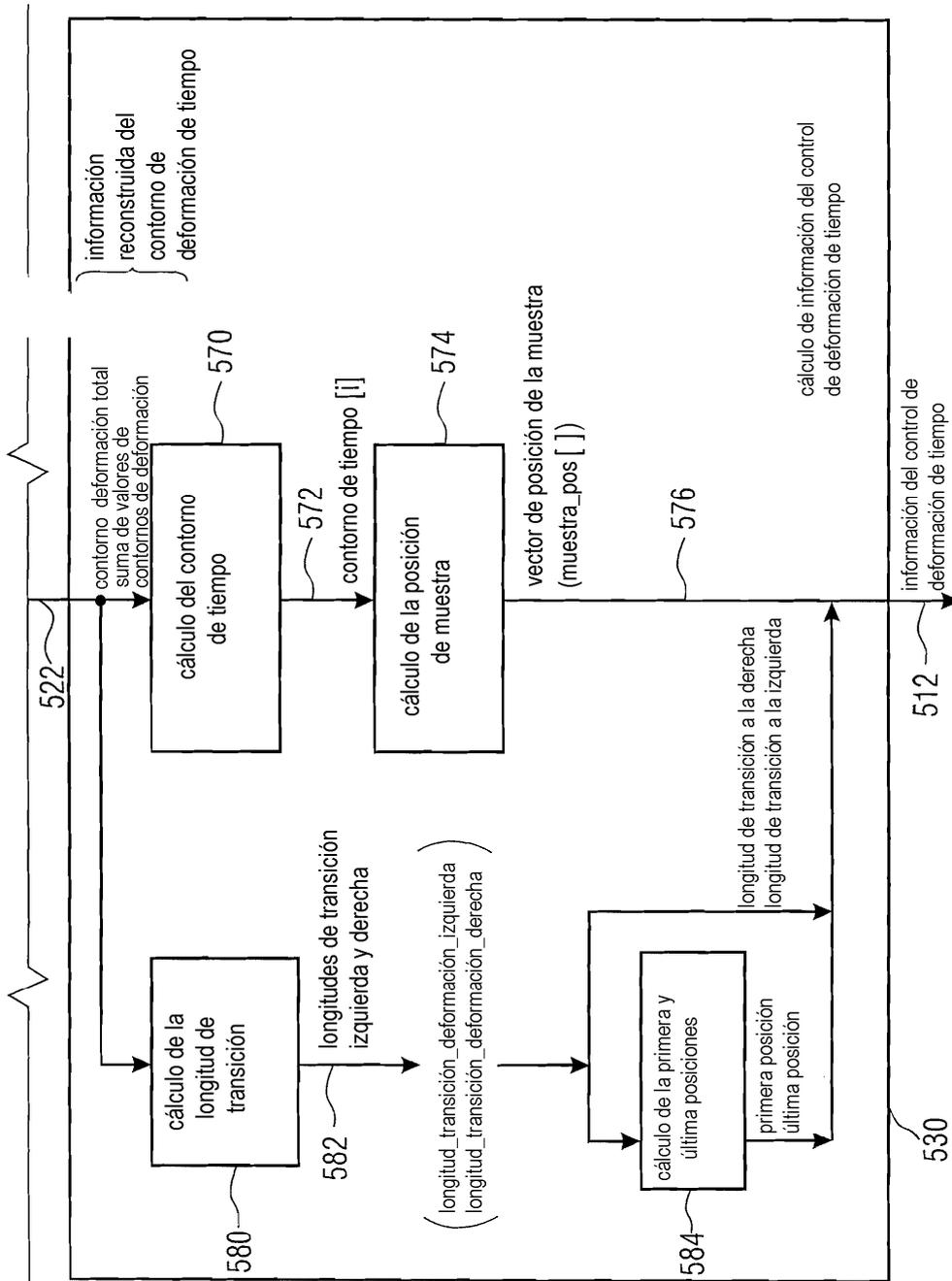
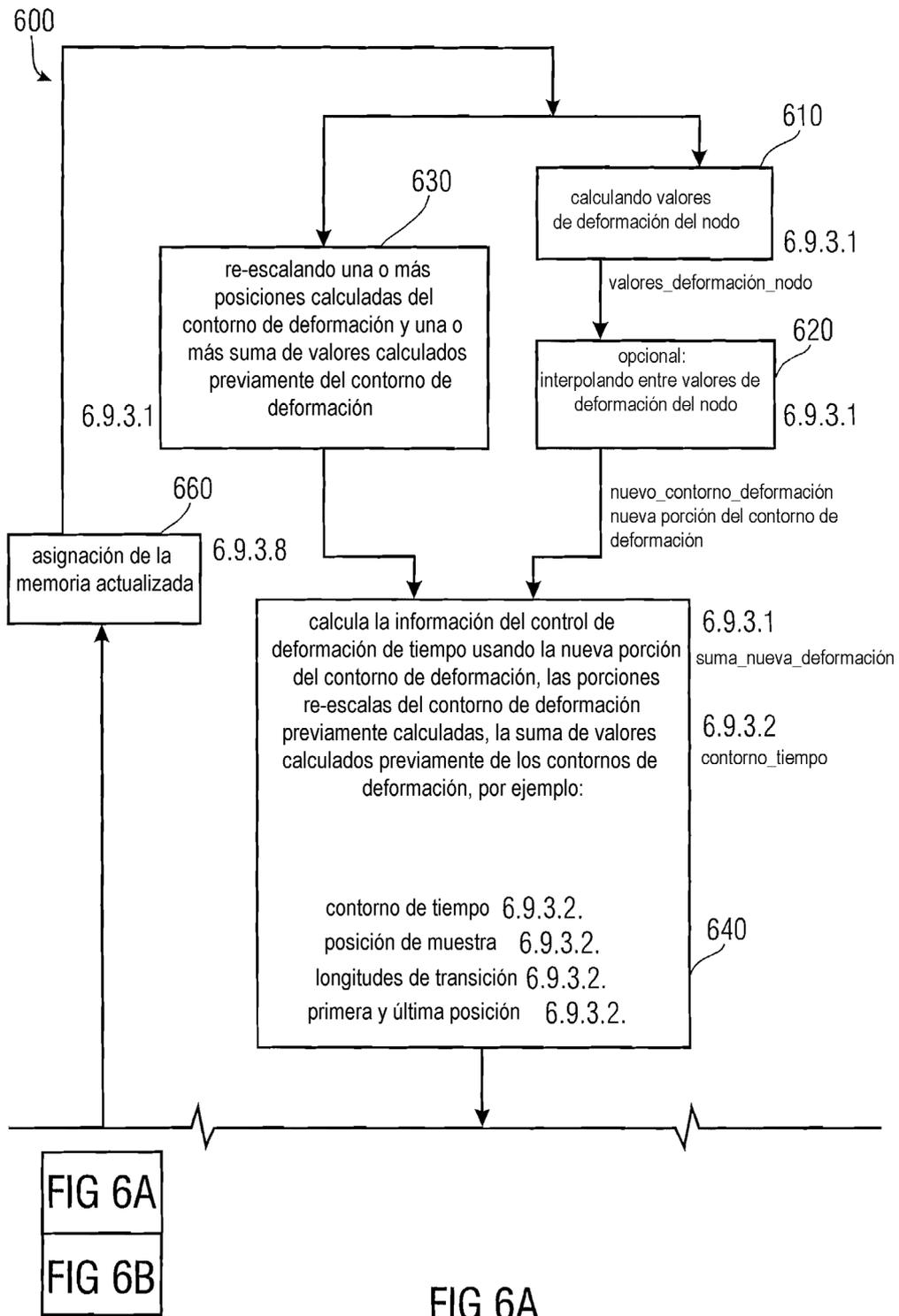


FIG 5B

FIG 5A
FIG 5B



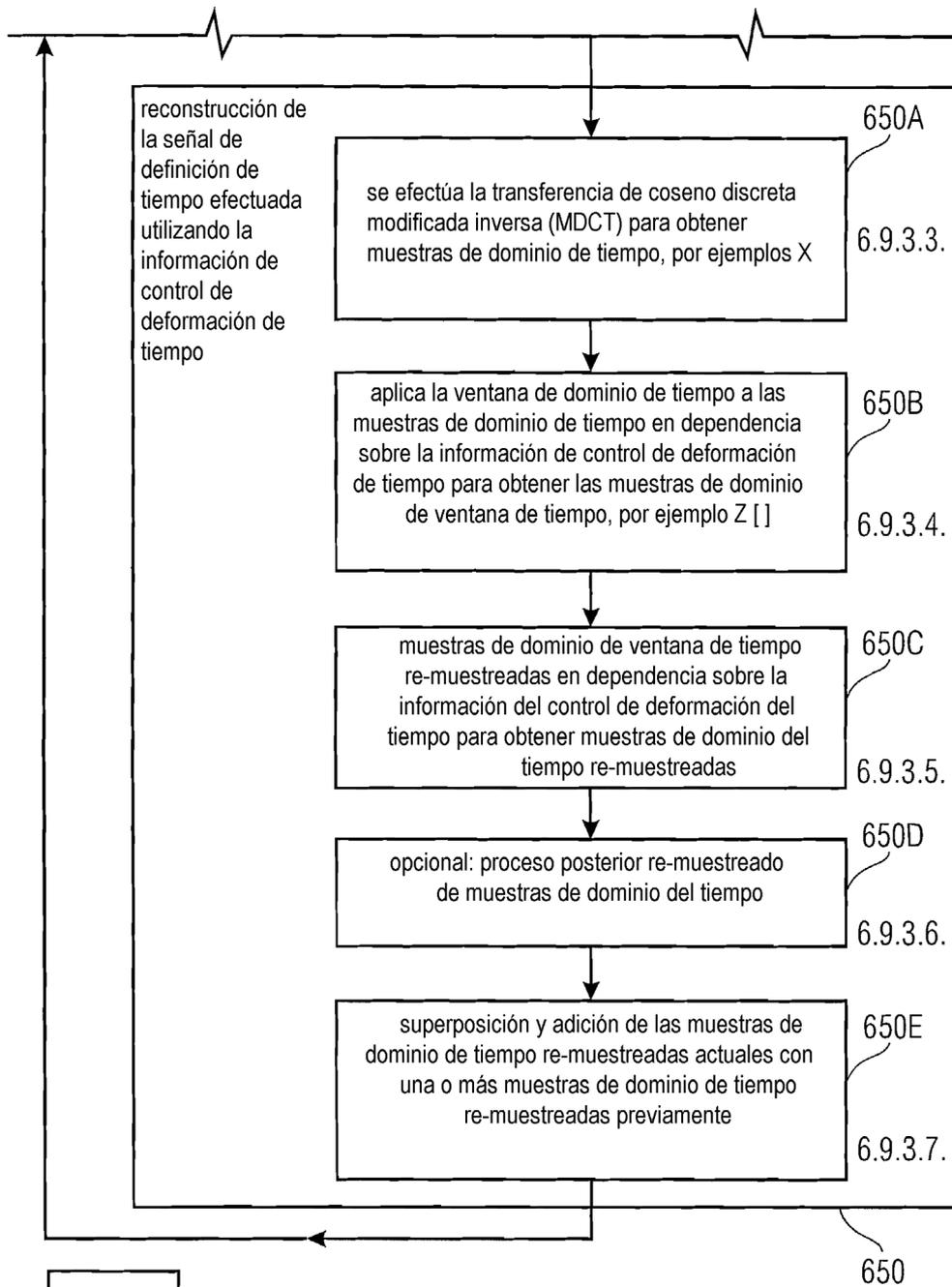


FIG 6A

FIG 6B

FIG 6B

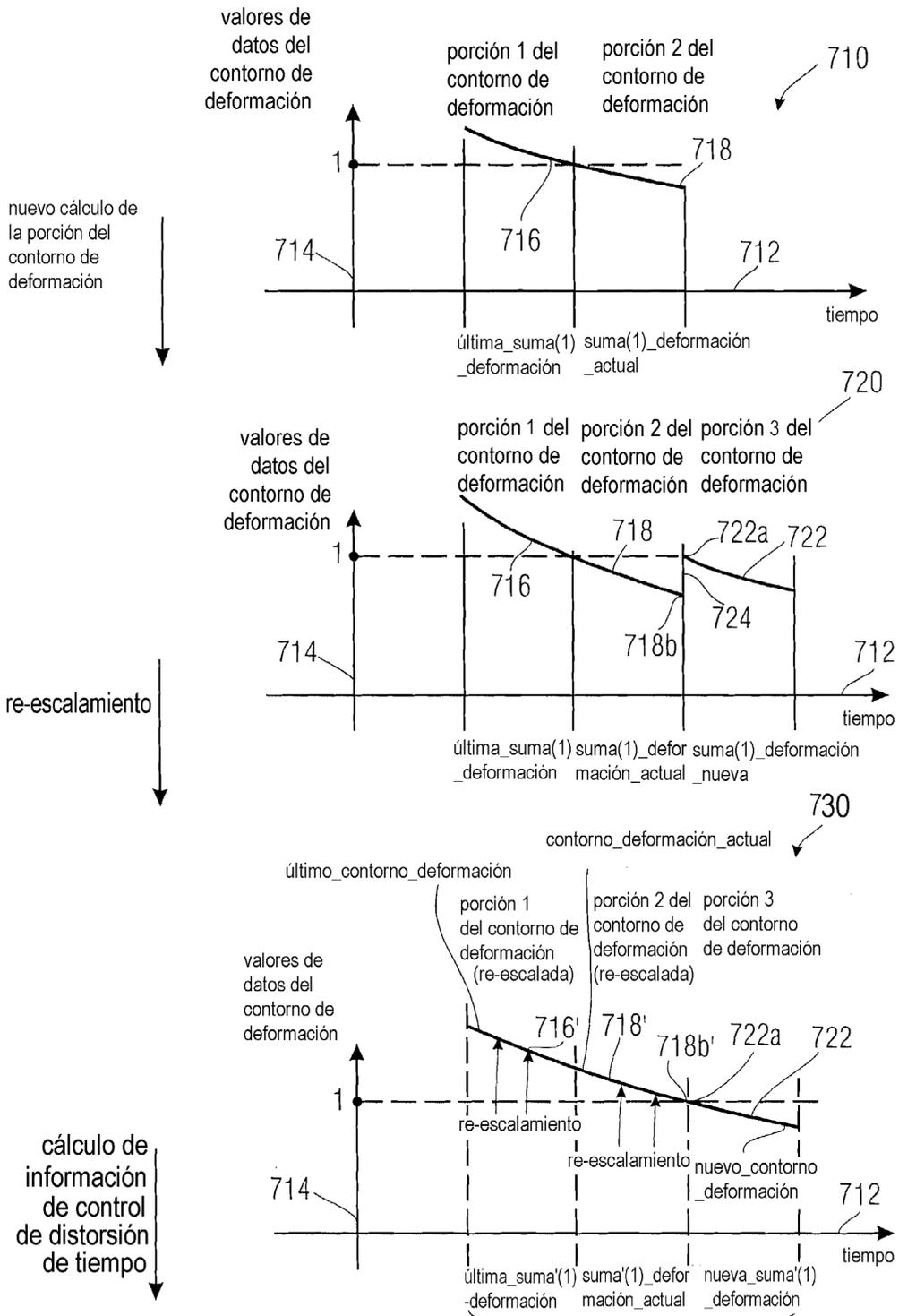


FIG 7A

La sección de contorno de deformación del tiempo es utilizada para el cálculo de la información del control de deformación del tiempo

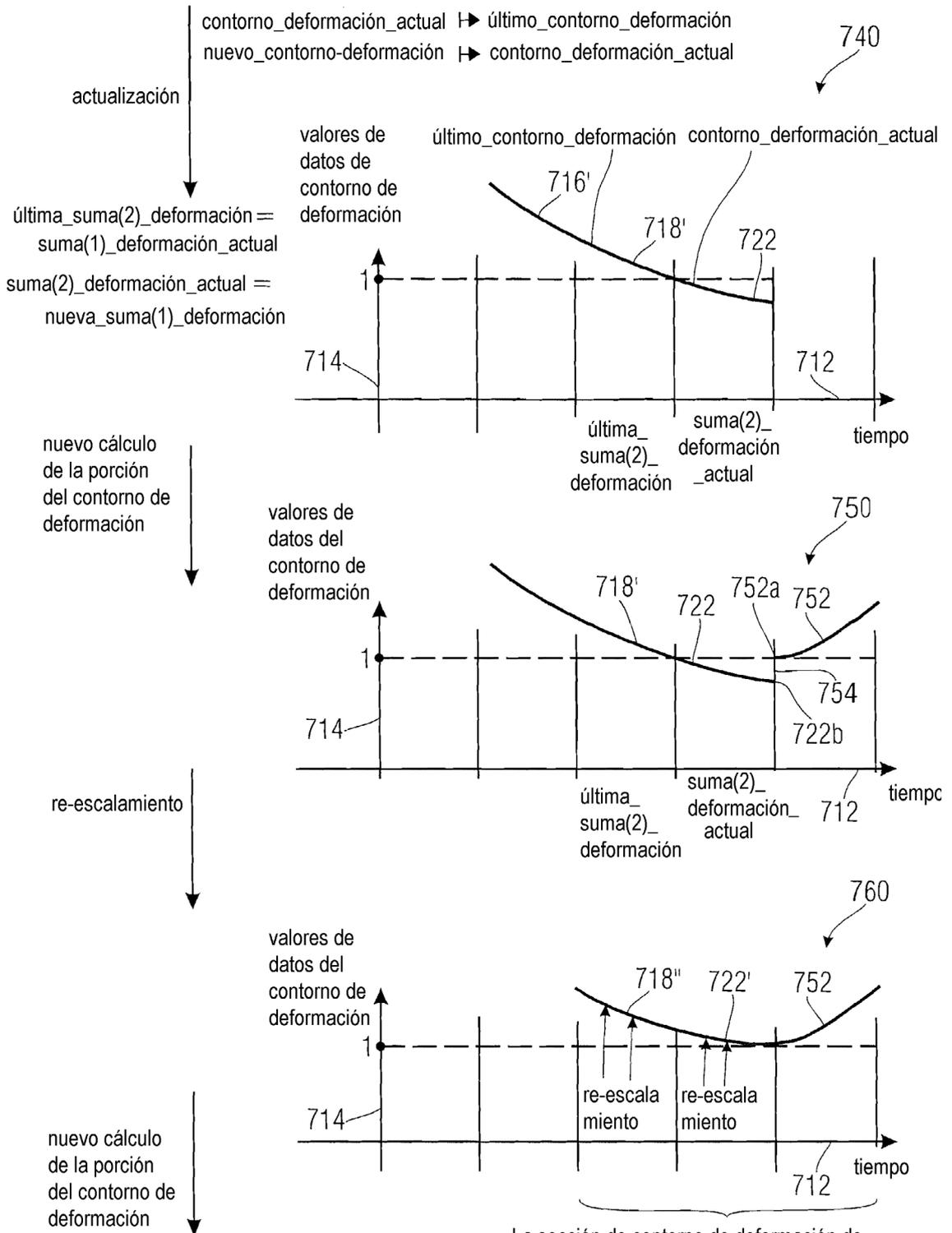


FIG 7B

La sección de contorno de deformación de tiempo es utilizada para el cálculo de la información del control de deformación de tiempo

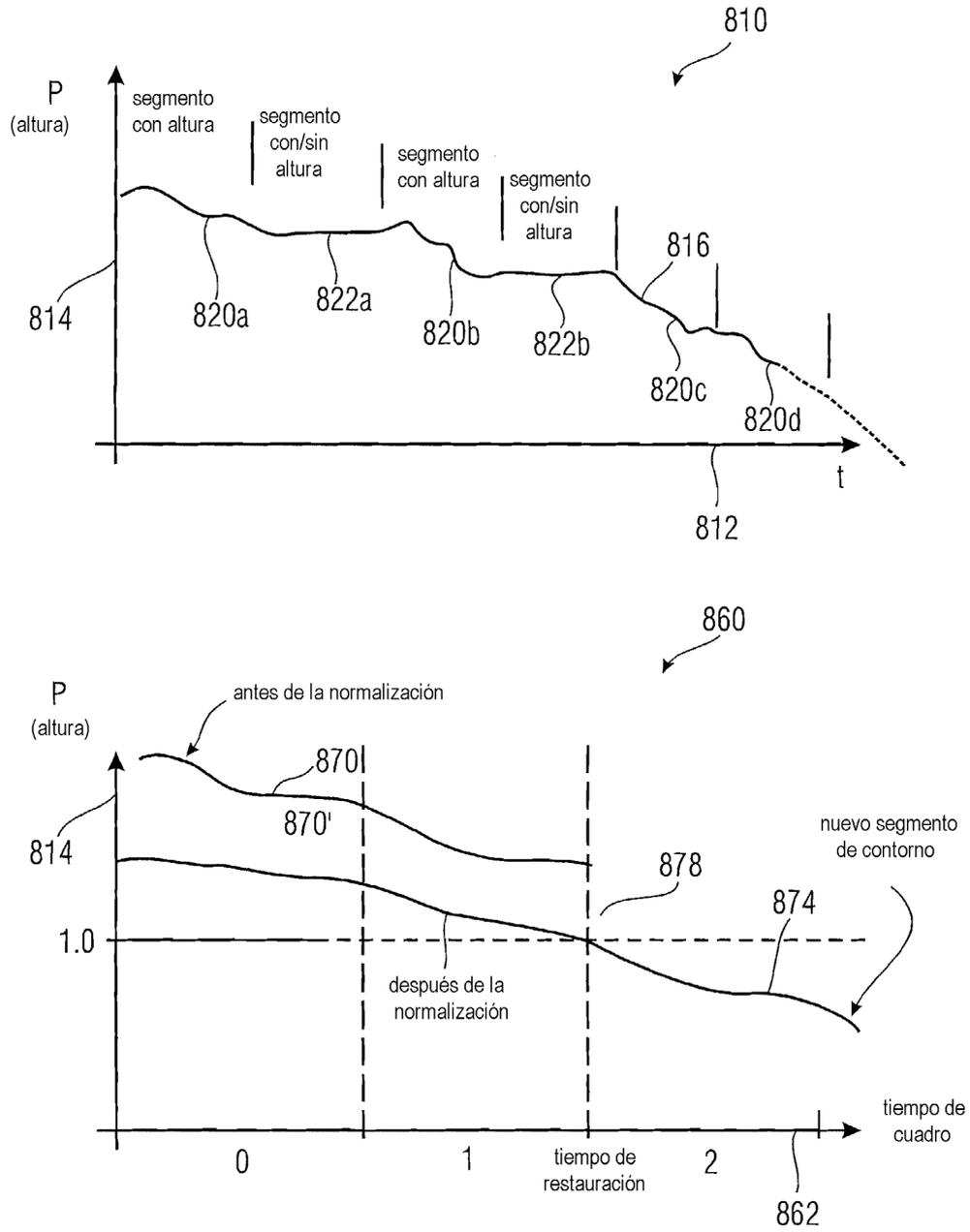


FIG 8

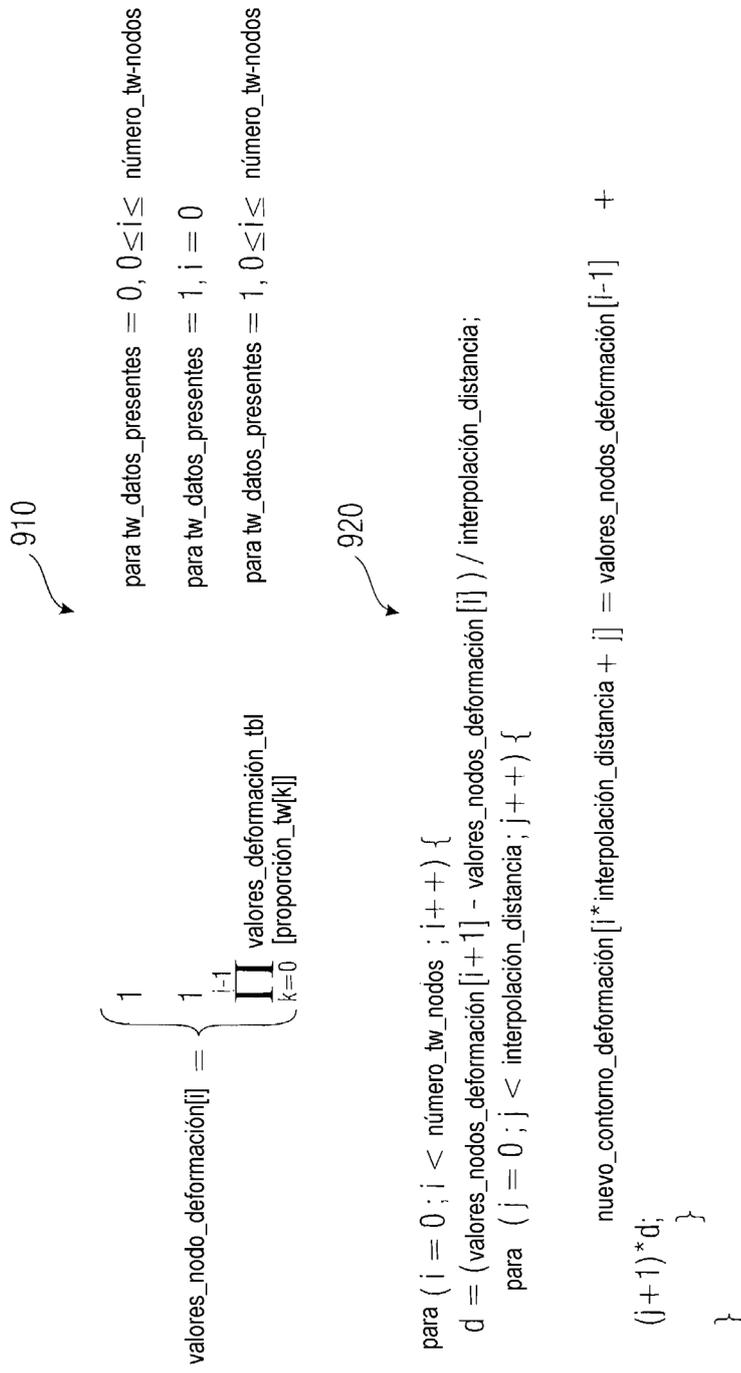


FIG 9A-1

$$\text{norm_fac} = \frac{1}{\text{contorno_deformación_pasado}[2 \cdot n_longitud - 1]}$$

930

$$\text{contorno_deformación_pasado}[i] = \text{contorno_deformación_pasado}[i] \cdot \text{norm_fac} \text{ para } 0 \leq i < 2 \cdot n_longitud$$

932

$$\text{última_suma_deformación} = \text{última_suma_deformación} \cdot \text{norm_fac}$$

934

$$\text{suma_deformación_actual} = \text{suma_deformación_actual} \cdot \text{norm_fac}$$

936

$$\text{new_warp_sum} = \sum_{i=0}^{n_long-1} \text{new_warp_contour}[i]$$

940

FIG 9A-2

contorno_deformación_pasado [n] = contorno_deformación [n + n_longitud], para $0 \leq n < 2 \cdot n_longitud$ 950
suma_deformación_actual = nueva_suma_deformación 952
última_suma_deformación = suma_deformación_actual 954

contorno[n]_deformación_pasado = 1, para $0 \leq n < 2 \cdot n_longitud$ 960
suma_deformación_actual = n_longitud 962
suma_deformación_pasado = n_longitud 964

FIG 9B

990

tabla_valor_deformación

Índice (por ejemplo, tw_proporción)	valor (por ejemplo, tabla_valor_deformación [tw_proporción])
0	0.982857168
1	0.988571405
2	0.994285703
3	1
4	1.0057143
5	1.01142859
6	1.01714289
7	1.02285719

FIG 9C

cálculo del contorno de tiempo

$$\text{contorno}[i]_{\text{tiempo}} = \begin{cases} -w_{\text{res}} \cdot \text{última_suma_deformación} & \text{para } i = 0 \\ w_{\text{res}} \left(-\text{última_suma_deformación} + \sum_{k=0}^{i-1} \text{contorno_deformación}[k] \right) & \text{para } 0 < i \leq 3 \cdot n_{\text{longitud}} \end{cases}$$

en donde $w_{\text{res}} = \frac{n_{\text{longitud}}}{\text{suma_deformación_actual}}$

funciones auxiliares

```

tiempo_deformación_inv (contorno[]_tiempo, t_deformación) {
    i = 0;
    si (t_deformación < contorno [0] ) {
        retorno SIN TIEMPO;
    }
    mientras (t_deformación > tiempo_contorno [i+1] ) {
        i++;
    }
    retorno (i + (t_deformación-tiempo_contorno [i]) / (tiempo_contorno [i+1] -
tiempo_contorno[i]));
}
    
```

```

deformación_inv_vec (tiempo_contorno[] t_inicio, n_muestras, muestra_pos[]) {
    t_deformación = t_inicio;
    j = 0;
    mientras ((i = piso (tiempo_deformación_inv (tiempo_contorno, t_deformación-0.5))) ==
SIN TIEMPO) {
        t_warp += 1;
        j++;
    }
    mientras (j < n_muestras && (t_deformación + 0.5) < contorno_tiempo [3*n_longitud] ) {
        mientras (t_deformación < contorno_tiempo [i+1] ) {
            i++;
        }
        muestra_pos[j] =
            i + (t_deformación- contorno_tiempo / (contorno_tiempo [i+1] -
contorno_tiempo[i]));
        j++;
        t_deformación += 1;
    }
}
    
```

FIG 10A

cálculo de forma de ventana

$$N_{os} = 2 \cdot n_{long} \cdot os_factor_ganancia$$

← 1040

$$W_{KBD} \left(n - \frac{N_{os}}{2} \right) = \sqrt{\frac{\sum_{p=0}^{N_{os}-n-1} [W(p,a)]}{\sum_{p=0}^{N_{os}/2} [W(p,a)]}} \quad \text{para } \frac{N_{os}}{2} \leq n < N_{os}$$

← 1042

en donde:

$$W(n,a) = \frac{I_0 \left[\pi \alpha \sqrt{1.0 - \left(\frac{n - N_{os}/4}{N_{os}/4} \right)^2} \right]}{I_0 [\pi \alpha]} \quad \text{para } 0 \leq n \leq \frac{N_{os}}{2}$$

$$I_0 [x] = \sum_{k=0}^{\infty} \left[\frac{x^k}{k!} \right]^2$$

α = factor alfa de ventana kernel, $\alpha = 4$

} ← 1044

$$W_{SENO} \left(n - \frac{N_{os}}{2} \right) = \text{seno} \left(\frac{\pi}{N_{os}} \left(n + \frac{1}{2} \right) \right) \quad \text{para } \frac{N_{os}}{2} \leq n < N_{os}$$

← 1046

$$\text{forma_ventana_izquierda}[n] = \begin{cases} W_{KBD}[n], & \text{si bloque_previo_forma_ventana} = 1 \\ W_{SIN}[n], & \text{si bloque_previo_forma_ventana} = 0 \end{cases}$$

← 1048

$$\text{forma_ventana_derecha}[n] = \begin{cases} W_{KBD}[n] & \text{si forma_ventana} = 1 \\ W_{SIN}[n], & \text{si forma_ventana} = 0 \end{cases}$$

← 1050

FIG 10C

ventana ("SECUENCIA CORTA OCHO")

```

tw_ventana_corta(X [], z [], primera_pos, última_pos, deformación_trans_len_izquierda, deformación_
trans_len_derecha, izquierda_ventana_forma[], derecha_ventana_forma[]) {

    compensación = n_longitud - 4 * n_corto n_longitud/2

    tr_escalal = 0.5 * n_longitud/deformación_trans_len_izquierda * os_factor_ventana;
    tr_pos_l = deformación_trans_len_izquierda + (primera_pos - n_longitud/2) + 0.5 * tr_escalal;
    tr_escalar = 8 * os_factor_ventana;
    tr_pos_r = tr_escalar/2

    para ( i = 0 ; i < n_toma ; i++ ) {
        z [i] = X [0] [i];
    }

    para ( i = 0 ; i < primera_pos ; i++ )
        z [i] = 0. ;

    para ( i = n_longitud - 1 - primera_pos ; i >= primera_pos ; i-- ) {
        z [i] * = forma_ventana_izquierda [piso (tr_pos_l)];
        tr_pos_l + = tr_escalal;
    }

    para ( i = 0 ; i < n_corto ; i++ ) {
        z [compensación + i + n_corto] =
            X[0] [i + n_corto] * forma_ventana_derecha [piso (tr_pos_r)];
        tr_pos_r + = tr_escalar;
    }

    compensación + = n_corto,

    para ( k = 1 ; k < 7 ; k++ ) {
        tr_escalal = n_corto * os_factor_ventana;
        tr_pos_l = tr_escalal/2;
        tr_pos_r = os_factor_ventana * n_longitud - tr_pos_l;
        for ( i = 0 ; i < n_corto ; i++ ) {
            z [i + compensación] + = X [k] [i] * forma_ventana_derecha [piso (tr_pos_r)];
            z [compensación + n_corto + i] =
                X[k] [n_corto + i] * forma_ventana_derecha [piso (tr_pos_l)];
            tr_pos_l + = tr_escalal;
            tr_pos_r - = tr_escalal;
        }
        compensación + = n_corto;
    }
}

```

1060 →

FIG 10D-1

```

tr_escal_a_l = n_corto*os_factor_ganancia;
tr_pos_l = tr_escal_a_l/2;

para ( i = n_corto - 1; i >= 0; i-- ) {
z [i + compensación] += X [7][i]*forma_ventana_derecha [ (int) piso (tr_pos_l)];
tr_pos_l += tr_escal_a_l,
}

para ( i = 0 ; i < n_corto ; i++ ) {
z [compensación+ n_corto + i] = X [7] [n_corto + i];
}

tr_escal_a_r = 0.5*n_longitud/deformaciónTransLenRight*os_factor_ganancia;
tr_pos_r = 0.5*tr_escal_a_r+.5;

tr_pos_r = (1.5*n_longitud-(flotador)wEnd - 0.5 + deformaciónTransLenRight)*tr_escal_a_r;

para ( i = 3*n_longitud-1- última_pos ; i <= wEnd; i++ ) {
[i] *= forma_ventana_derecha[piso(tr_pos_r)];
tr_pos_r += tr_escal_a_r;
}

para (i=lsat_pos+1; i<2*n_longitud; i++)
z [i] = 0. ;

```

FIG 10D-2

Ventana)TODOS LOS OTROS)

Tw_ventana_larga (X[0][], z [], primera_pos, última_pos,
deformación_trans_len_izquierda, deformación_trans_len_derecha,
izquierda_ventana_forma [], derecha_ventana_forma []) {

Para (i = 0 ; i < primera_pos; i++)

z [i]=0.;

Para (i = última_pos+1 ; i < N_f; i++)

z [i]=0.;

tr_escala =

0.5*n_longitud/deformación_trans_len_izquierda*os_factor_ventana;

tr_pos = (deformación_trans_len_izquierda+ primera_pos-N_f/4)+0.5)*

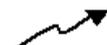
tr_escala;

Para (i = N_f/2-1- primera_pos ; i >= primera_pos; i--) {

z [i] = X[0] [i] * izquierda_ventana_forma [piso(tr_pos)];

tr_pos += tr_escala;

}

1064 

tr_escala = 0.5*n_longitud/deformación_trans_len_derecha*os_factor_ventana;

tr_pos = (3*N_f/4- última_pos-0.5+deformación_trans_len_derecha)* tr_escala;

Para (i = 3*N_f/2-1- última_pos ; i <= última_pos; i++) {

z [i] = X[0] [i] * derecha_ventana_forma [piso(tr_pos)];

tr_pos += tr_escala;

}

}

FIG 10E

tiempo de variación del re-muestreo

$$b[n] = l_0 [\alpha]^{-1} l_0 \left[\alpha \sqrt{1 - \frac{n^2}{ip_len_2^2}} \right] \cdot \frac{\text{seno} \left(\frac{\pi n}{os_factor_resamp} \right)}{\frac{\pi n}{os_factor_resamp}} \cdot \text{para } 0 \leq n < ip_tamaño-1 \quad \left. \vphantom{b[n]} \right\} \leftarrow 1070$$

$\alpha = 8$

$$zp[n] = \begin{cases} 0, & \text{para } 0 \leq n < ip_len_2s \\ z[n-ip_len_2s], & \text{para } ip_len_2s \leq n < N_f + ip_len_2s \\ 0, & \text{para } 2 \cdot N_f + ip_len_2s \leq n < N_f + 2 \cdot ip_len_2s \end{cases} \quad \left. \vphantom{zp[n]} \right\} \leftarrow 1072$$

```

compensación_pos = 0.5;
número_muestras_entrada = N_f + 2 * ip_len_2s;
número_muestras_salida = 3 * n_longitud;
j_centro = 0;
para (i = 0; i < numMuestrasFuera; i++) {
en tanto que (j_centro < num_muestras_in && muestra_pos [j_centro] - compensación_pos <= i)
    j_centro++;
    j_centro--;
    y[i] = 0;
    si (j_centro < num_muestras_in-1 && j_centro > 0) {
        frac_tiempo = piso ((i - (muestra_pos [j_centro] - compensación_pos))
            / (muestra_pos [j_centro+1] - muestra_pos [j_centro])
            * os_factor);
        j = ip_len_2s * os_factor + frac_tiempo;

        para (k = j_centro - ip_len_2s; k <= j_centro + ip_len_2s; k++) {
            si (k >= 0 && k < num_muestras_entrada)

                y[i] += b[abs(j)] * zp[k];
                j -= os_factor;
            }
        }
    si (j_centro < 0)
        j_centro++;
}
    
```

FIG 10F

ES 2 404 132 T3

a) SECUENCIA_OCHO_CORTOS;, SECUENCIA_LONGITUD_INICIO,
SECUENCIA_INICIO_PARADA, SECUENCIA_1152_INICIO_PARADA seguida por una SECUENCIA LPD

$$W_{\text{corr}}(n) = \frac{W_{\text{corto}}(n)}{W_{\text{FD_LPD}}(n)}, \text{ para } 0 \leq n < \frac{n_{\text{corto}}}{2} \quad \leftarrow 1080a$$

$$W_{\text{corto}}(n) = \begin{cases} W_{\text{SENO_DERECHA}} \cdot n_{\text{corto}}(n), & \text{si forma_ventana} = 0 \\ W_{\text{KBD_DERECHA}} \cdot n_{\text{corto}}(n), & \text{si forma_ventana} = 1 \end{cases}$$

$$W_{\text{FD_LPD}}(n) = \begin{cases} W_{\text{SENO_DERECHA}} \cdot n_{\text{corto}/2}(n), & \text{si forma_ventana} = 0 \\ W_{\text{KBD_DERECHA}} \cdot n_{\text{corto}/2}(n), & \text{si forma_ventana} = 1 \end{cases} \quad \leftarrow 1080b$$

es aplicado:

$$y'_{i,n} = \begin{cases} y[n] & , \text{para } 0 \leq n < \frac{5n_{\text{long}}}{2} - \frac{n_{\text{corto}}}{2} \\ y[n] \cdot W_{\text{corr}}\left(n - \left(\frac{5n_{\text{long}}}{2} - \frac{n_{\text{corto}}}{2}\right)\right) & , \frac{5n_{\text{long}}}{2} - \frac{n_{\text{corto}}}{2} \leq n < \frac{5n_{\text{long}}}{2} \\ 0 & , \text{para } \frac{5n_{\text{long}}}{2} \leq n < 3n_{\text{long}} \end{cases} \quad \leftarrow 1082$$

b) otras clases completas:

ninguna se ha realizado

$$y'_{i,n} = y[n] \text{ para } 0 \leq n < 3 \cdot n_{\text{long}} \quad \leftarrow 1084$$

sobreposición y adición:

$$\text{salida}_{i,n} = \begin{cases} y'_{i,n} + y'_{i-1, n+n_{\text{long}}} + y'_{i-2, n+2 \cdot n_{\text{long}}} & \text{para } 0 \leq n < n_{\text{long}}/2 \\ y'_{i,n} + y'_{i-1, n+n_{\text{long}}} & \text{para } n_{\text{long}}/2 \leq n < n_{\text{long}} \end{cases} \quad \leftarrow 1086$$

FIG 10G

ES 2 404 132 T3

<p>Datos de elementos</p> <p>tw_dato()</p> <p>tw_dato_presente</p> <p>tw_proporción[]</p> <p>secuencia_ventana</p> <p>forma de ventana</p> <p>elementos auxiliares</p> <p>valores_nodo_deformación[]</p> <p>tabla_valor_deformación</p> <p>tabla_valor_deformación[]</p> <p>nuevo_contorno_deformación[]</p> <p>contorno_deformación_pasado[]</p> <p>normal_fac</p> <p>contorno[]_deformación</p> <p>última_suma_deformación</p> <p>suma_deformación_actual</p> <p>siguiente_suma_deformación</p> <p>tiempo_contorno[]</p> <p>muestra_pos[]</p> <p>X[w]{}[]</p>	<p>contiene la información lateral necesaria para descodificar y aplicar el tiempo de deformación MDCT en una corriente_canal()_fd para los elementos SCE y CPE. Las corrientes canal fd de un elemento()_canal_pareado pueden compartir un tw_dato() común.</p> <p>1 sonido indica que un contorno de deformación no plano es transmitido en este cuadro</p> <p>índice del libro de códigos de la proporción de deformación para el nodo i.</p> <p>2 sonidos indican que la secuencia de ventana está en uso (esto es, tamaño de bloque)</p> <p>1 sonido indica que la ventana de función es seleccionada.</p> <p>valores del nodo de contorno de deformación descodificados</p> <p>véase tabla "tabla_valor_deformación"</p> <p>tabla de cuantificación para la proporción de valores del nodo de deformación, por favor, véase Tabla 9C</p> <p>contorno de deformación interpolado y descodificado para este cuadro (n_long muestras)</p> <p>contorno de deformación pasado (2*n_long muestras), comprende contorno_última_deformación y contorno_deformación_actual</p> <p>normalización del contorno_deformación pasado</p> <p>contorno de deformación completo (3*n_long muestras)</p> <p>suma de la primera parte del contorno de deformación</p> <p>suma de la parte media del contorno de deformación</p> <p>suma de la última parte del contorno de deformación</p> <p>tiempo de contorno completo (3*n_long + 1_muestras)</p> <p>posiciones de las muestras deformadas sobre una escala de tiempo lineal (2*n_long muestras + 2*ip_len_2s)</p> <p>salida del MDCT a la ventana w</p>
--	--

FIG 11A

ES 2 404 132 T3

Elementos auxiliares continuación)

z[]	ventana y opcionalmente) cuadro en el dominio deformado de tiempo	vector for one
zp[]	z[] con acolchado cero	
y[]	vector de tiempo para un cuadro en el dominio de tiempo lineal después del re-muestreo	
y' _{i,n}	vector de tiempo del cuadro i después del reprocesamiento	
salida[]	vector de salida para un cuadro	
b[]	impulso de respuesta para el filtro re-muestreador	
N	longitud de la ventana de síntesis. véase más adelante	
N_f	longitud de estructura, ya sea 2304 en el caso de SECUENCIA_1152_PARADA, SECUENCIA_1152_INICIO_PARADA o 2048 para todas las otras secuencias de ventana	

Constantes

número_tw_nodos	16
factor_ventana_os	16
re-muestreo_factor_os	128
ip_len_2s	64
ip_len_2	os_factor_re-muestreo * ip_len_2s + 1
ip_tamaño	ip_len_2 + os_factor_re-muestreo
n_longitud	1024 (960)
n_corto	128 (120)
interp_dist	n_longitud/número_tw_nodos
SIN TIEMPO	-100000

FIG 11B-1

Longitud de ventana 2304:

$$N = \begin{cases} 2304, & \text{si SECUENCIA_1152_PARO} \\ 2304, & \text{si SECUENCIA_1152_INICIO_PARO} \end{cases}$$

Longitud de ventana 2048:

$$N = \begin{cases} 2048, & \text{si SECUENCIA_LONGITUD_ÚNICA} \\ 2048, & \text{si SECUENCIA_INICIO_LONGITUD} \\ 256, & \text{si SECUENCIA_CORTO_OCHO} \\ 2048, & \text{si SECUENCIA_PARO_LONGITUD} \\ 2048, & \text{si SECUENCIA_INICIO_PARO} \end{cases}$$

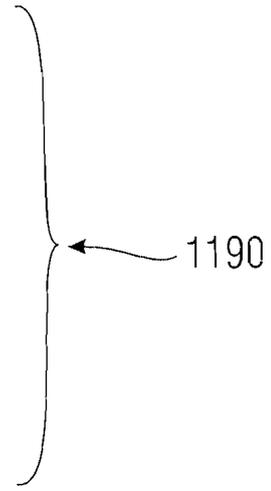


FIG 11B-2

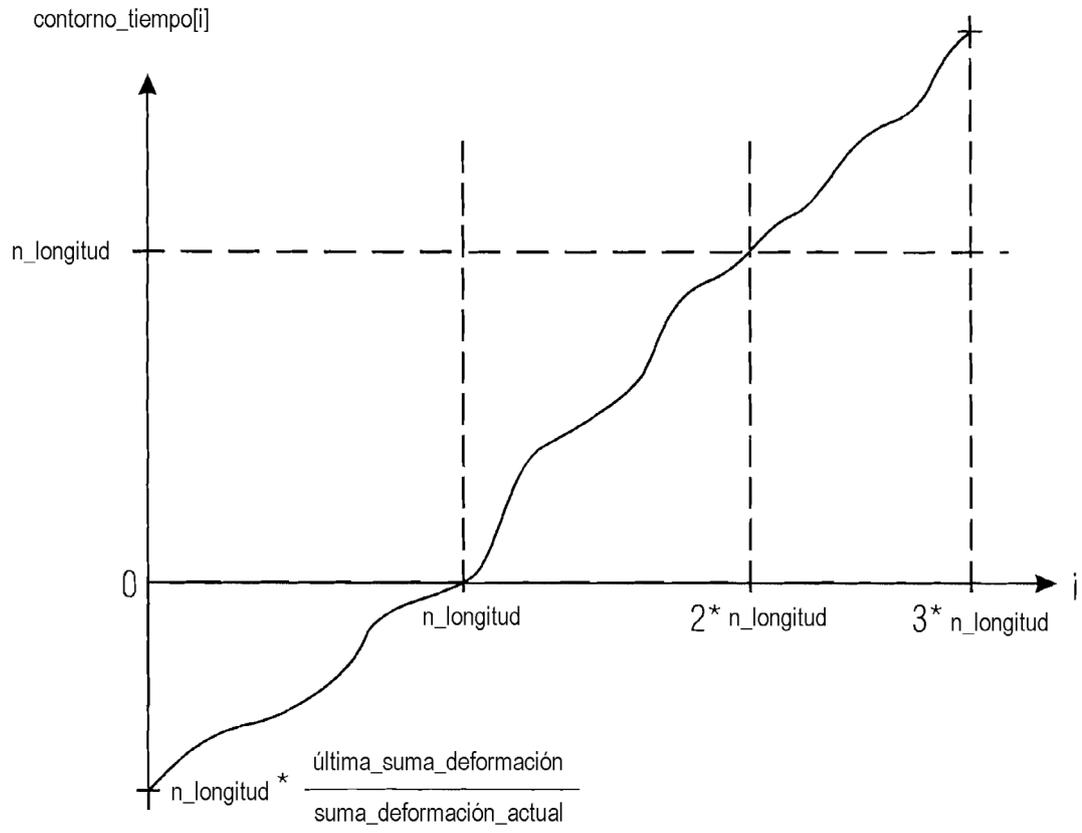


FIG 12

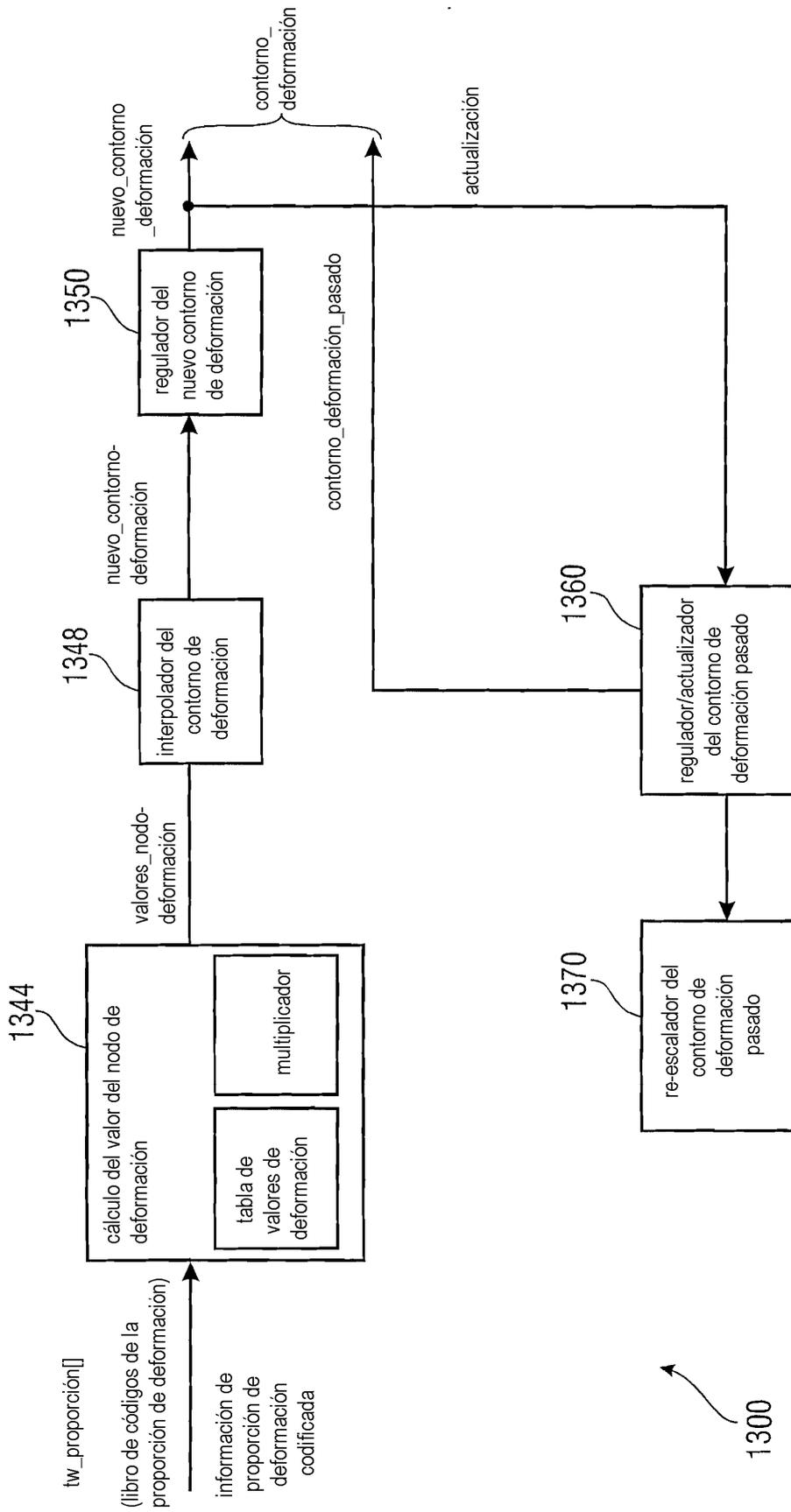


FIG 13

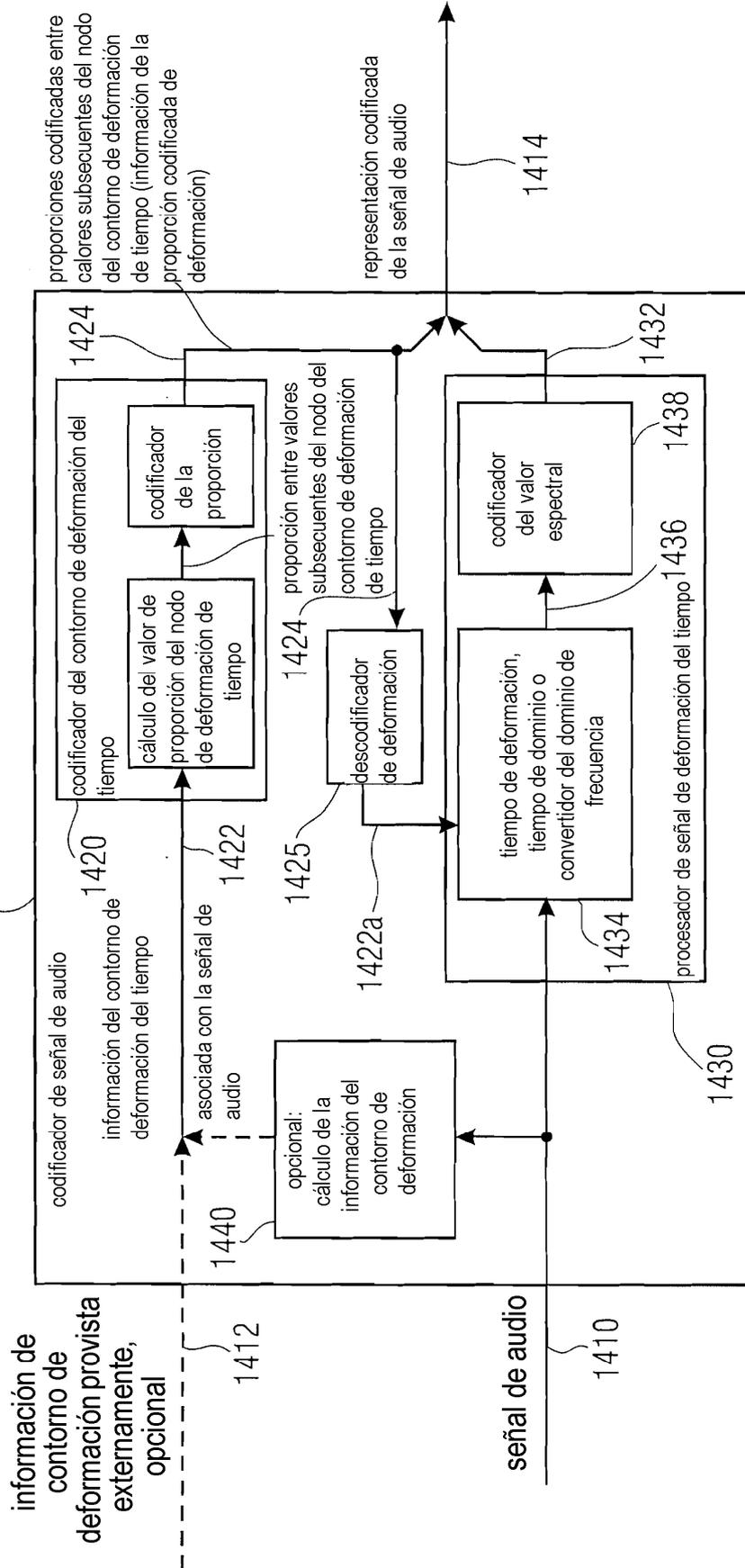


FIG 14

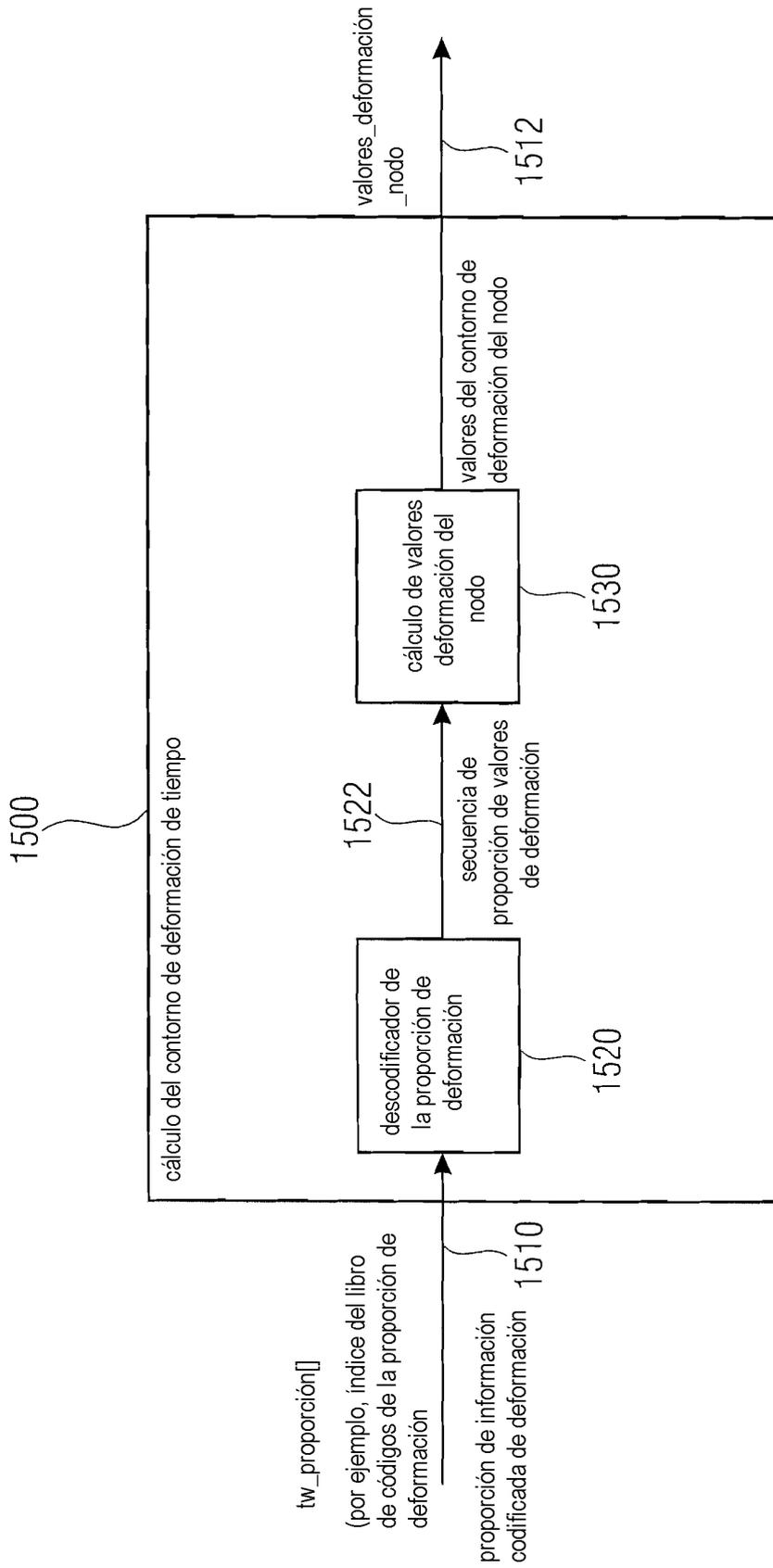
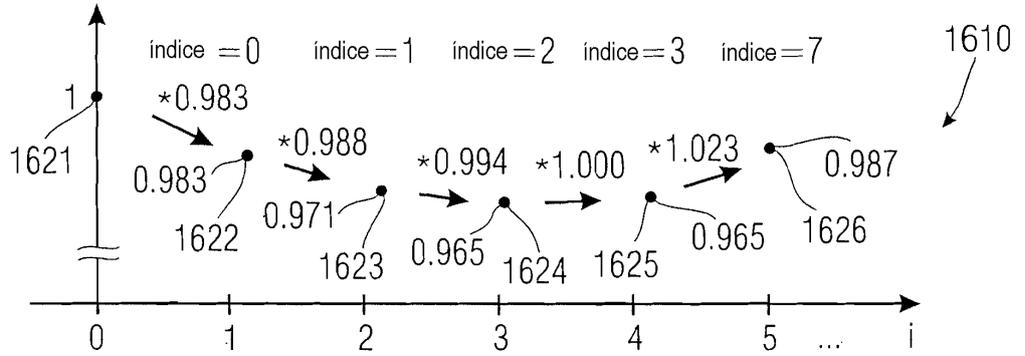


FIG 15

valores_nodo_deformación



interpolación lineal

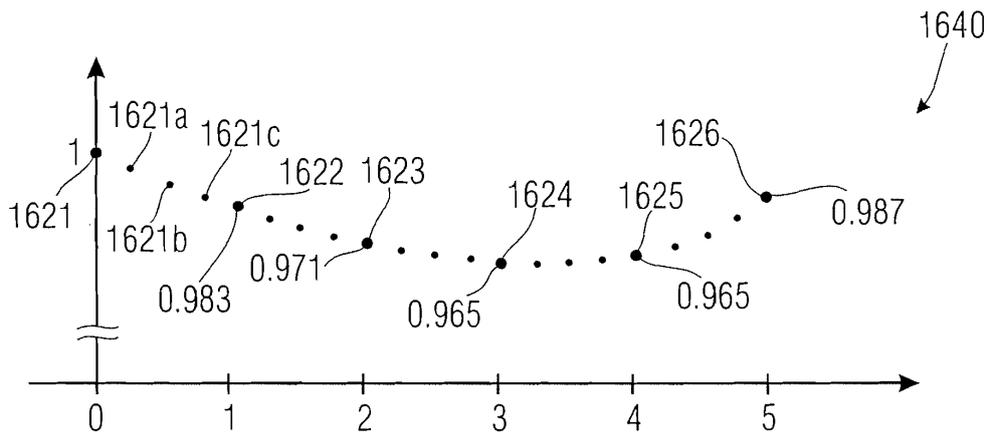


FIG 16A

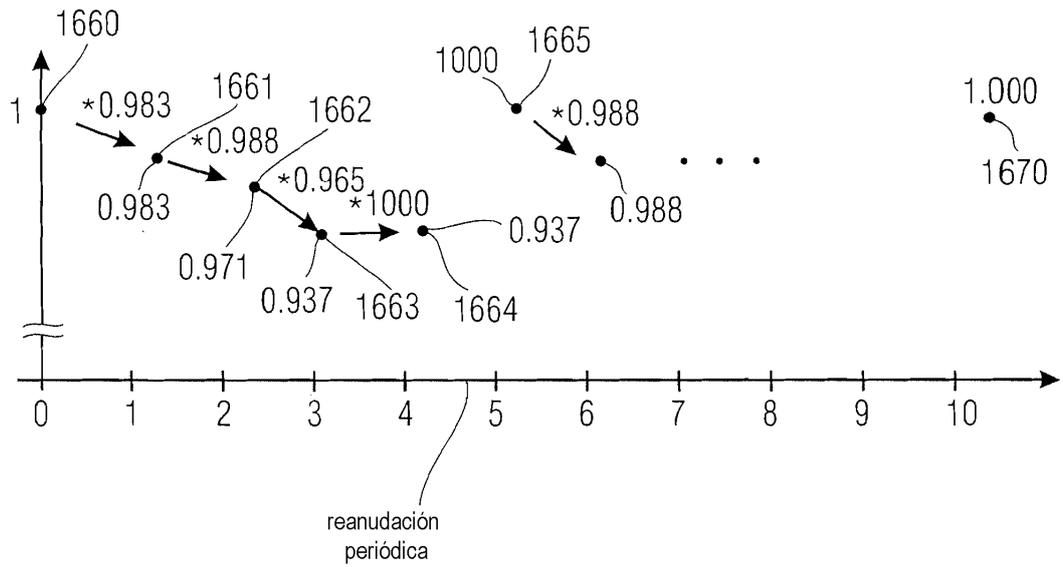


FIG 16B

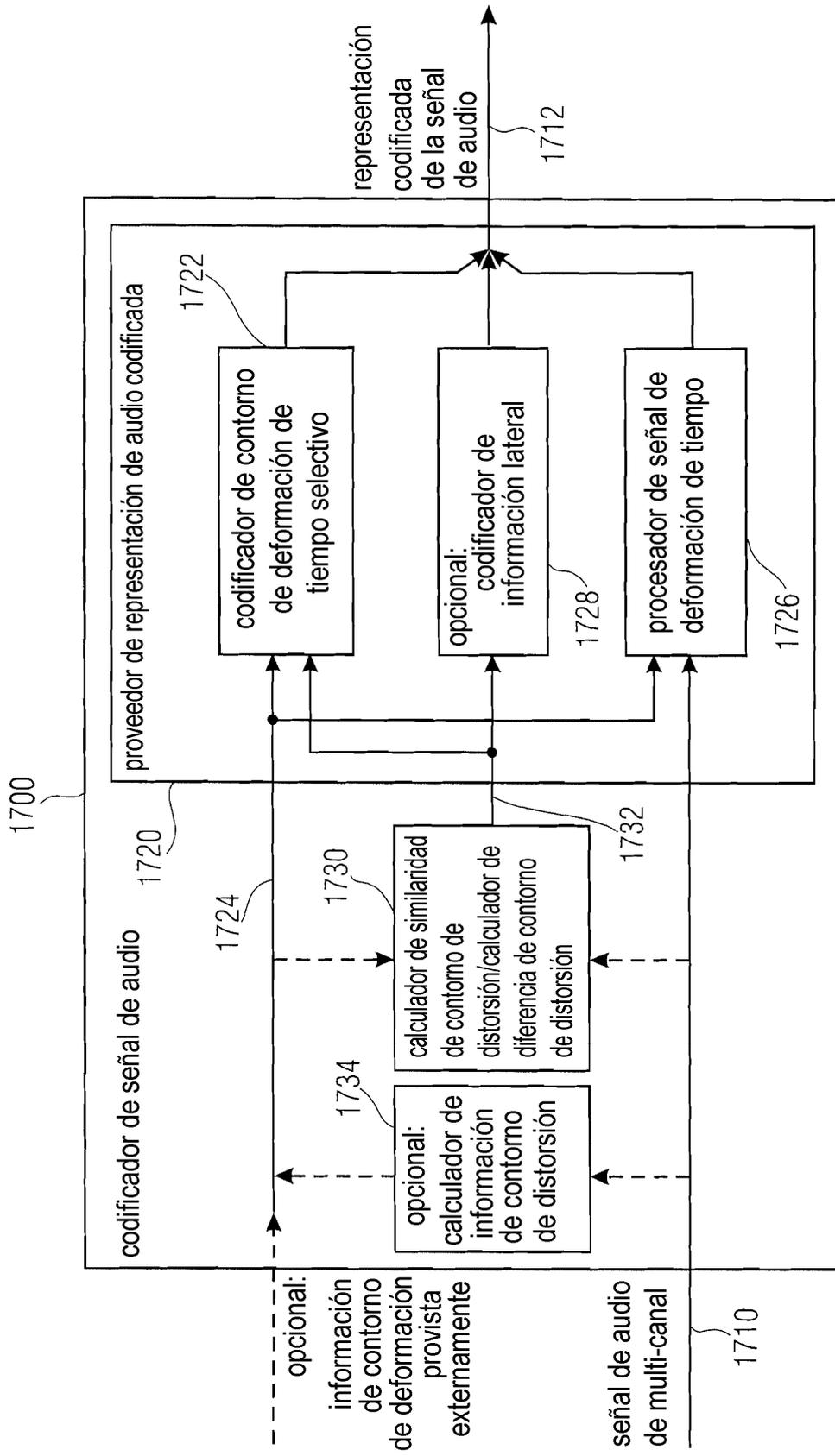


FIG 17

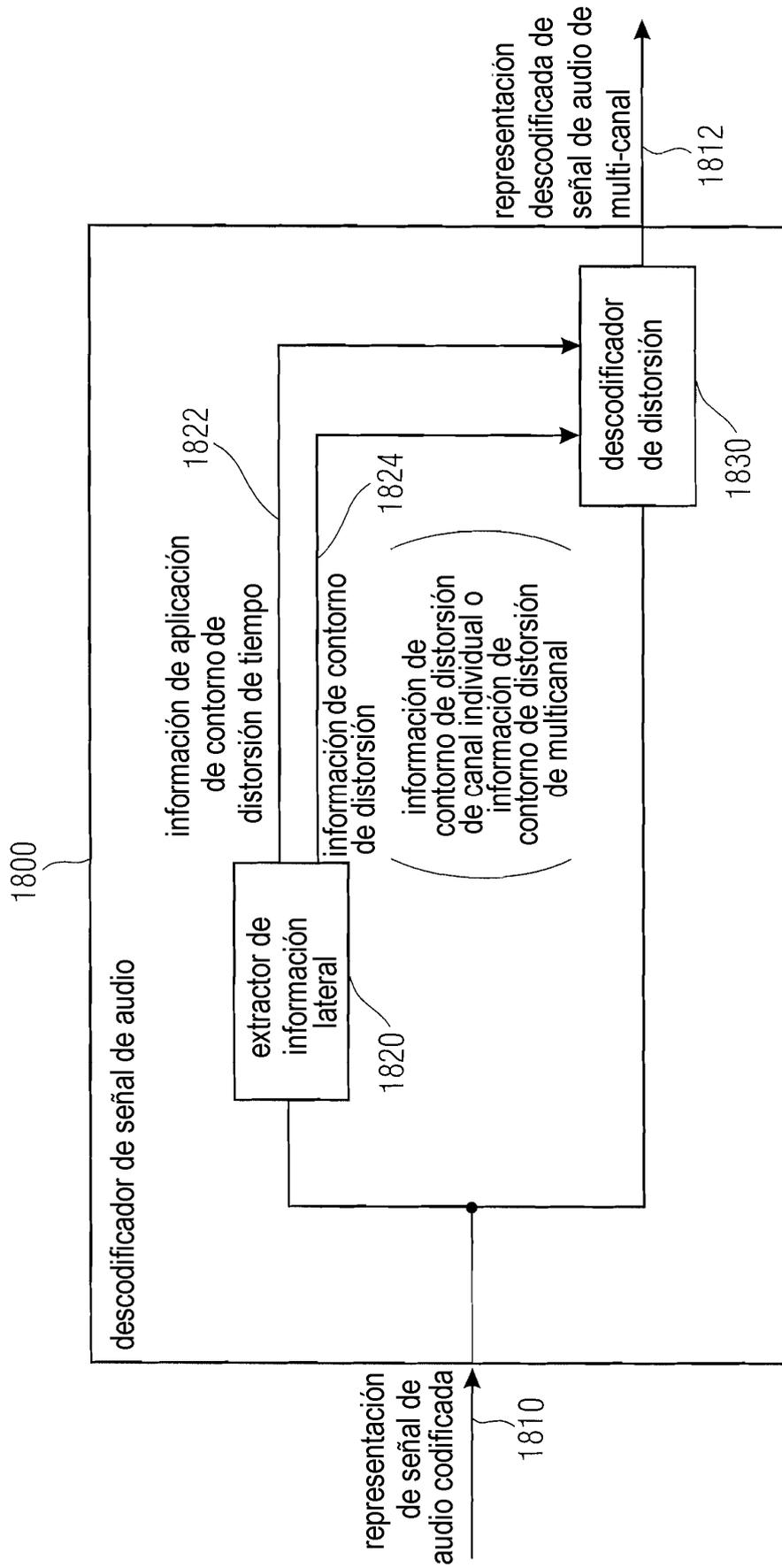


FIG 18

```
usac_dato_bloque()_aleatorio
{
    elemento()_canal_sencillo;
    0
    elemento()_canal_pareado;
    0
    elemento()_canal_sencillo;
    y
    elemento()_canal_pareado;
}
```

FIG 19A

```
elemento()_canal_sencillo
{
    corriente_canal_fd (*, *, *);
}
```

FIG 19B

```
elemento_canal_pareado
{
  si (tw_mdct) {
    común_tw;
    si ( común_tw) {
      tw_datos();
    }
  }
  corriente_canal_fd (*, *, *);
  corriente_canal_fd (*, *, *);
}
```

FIG 19C

```

corriente_canal_fd (*, *, *);
{
ganancia_global;
si (tw_mdct) {
    si (no_comú_tw) {
        tw_datos();
    }
}
factor_escalado_datos();
ac_espectral_datos();
}.

```

FIG 19D

Sintaxis	No. de sonidos	Mnemonic
tw_datos() {		
tw_datos_presentes	1	ulmsfb
si (tw_datos_presentes == 1) {		
para (i= 1 ; i < num_tw_nodos ; i++) {		
tw_proporción[i];	3	ulmsbf
}		
}		
}		

FIG 19E

Sintaxis de datos()._altura

Sintaxis	Número de tonos	Mnemon
datos()._altura		
{		
DatosAlturaActivos	1	uimsfb
si (DatosAlturaActivos == 1) {		
para (i= 1 ; i < numAlturas ; i++) {	1	uimsbf
índicealtura[i];	numAlturaSonidos	uimsbf
}		
}		
}		

FIG 19F