

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 644 370**

51 Int. Cl.:

**G10L 19/00** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.10.2014 PCT/EP2014/072063**

87 Fecha y número de publicación internacional: **23.04.2015 WO15055683**

96 Fecha de presentación y número de la solicitud europea: **14.10.2014 E 14784448 (4)**

97 Fecha y número de publicación de la concesión europea: **06.09.2017 EP 3044782**

54 Título: **Decodificador de audio, aparato para generar datos de salida de audio codificados y procedimientos que permiten inicializar un decodificador**

30 Prioridad:

**18.10.2013 EP 13189328**

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

**28.11.2017**

73 Titular/es:

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

72 Inventor/es:

**FISCHER, DANIEL;  
CZELHAN, BERND;  
NEUENDORF, MAX;  
RETTTELBACH, NIKOLAUS;  
HOFMANN, INGO;  
FUCHS, HARALD;  
DÖHLA, STEFAN y  
FÄRBER, NIKOLAUS**

74 Agente/Representante:

**SALVA FERRER, Joan**

**Observaciones :**

**Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes**

**ES 2 644 370 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

Decodificador de audio, aparato para generar datos de salida de audio codificados y procedimientos que permiten inicializar un decodificador

5

**[0001]** La presente invención está relacionada con la codificación/decodificación de audio y, en particular, con una estrategia de codificación y decodificación de datos, que permite inicializar un decodificador como el que puede ser necesario al conmutar entre diferentes configuraciones de códecs.

10 **[0002]** Las realizaciones de la invención pueden aplicarse a escenarios, en los que las propiedades de los canales de transmisión pueden variar ampliamente en función de la tecnología de acceso, como DSL, WiFi, 3G, LTE y similares. La recepción del teléfono móvil puede debilitarse en interiores o en zonas rurales. La calidad de las conexiones de Internet inalámbricas depende en gran medida de la distancia a la estación base y la tecnología de acceso, lo que provoca fluctuaciones en la velocidad de transferencia de bits. La velocidad  
15 de transferencia de bits disponible por usuario también puede cambiar con el número de clientes conectados a una estación base.

**[0003]** El documento WO 2010/003563 A1 describe un codificador de audio para codificar muestras de audio, que permite conmutar entre dos dominios de codificación. Un primer dominio de codificación puede ser  
20 AAC y un segundo dominio de codificación puede ser ACELP. Antes de conmutar de AAC a ACELP, un controlador modifica el encuadre de la ACELP de manera que una primera supertrama comprende cinco tramas en lugar de cuatro.

**[0004]** El documento EP 2 581 902 A1 muestra un decodificador de audio híbrido y un nuevo codificador  
25 de audio híbrido que tiene conmutación de bloques para señales de voz y señales de audio. Si la conmutación tiene lugar entre AAC y ACELP, la segunda mitad de la trama anterior i-1 se concatena para formar una trama extendida.

**[0005]** El documento WO 2010/005224 A2 describe un aparato para procesar una señal de audio. Al  
30 conmutar entre diferentes esquemas de codificación, se incluye adicionalmente una señal de salida de cada esquema de codificación antes y después de la conmutación para generar una parte en la que las señales de dos esquemas de codificación se superponen entre sí.

**[0006]** El documento EP 1 396 843 A1 A enseña un esquema de compresión de audio unificado con  
35 pérdidas y sin pérdidas que combina compresión de audio con pérdidas y sin pérdidas dentro de una misma señal de audio. La compresión mixta sin pérdidas codifica una trama de transición entre tramas de codificación con pérdidas y sin pérdidas para producir transiciones sin fisuras. La codificación mixta sin pérdidas ejecuta una transformada con superposición y una transformada con superposición inversa para producir una trama debidamente segmentada en ventanas y doblada en el dominio del pseudotiempo que puede ser  
40 posteriormente codificada sin pérdidas. La codificación mixta sin pérdidas también puede aplicarse a tramas que presentan un mal resultado de compresión con pérdidas.

**[0007]** El objetivo de la invención es proporcionar un concepto que permita suministrar contenido de  
45 audio de una manera flexible.

**[0008]** Según la invención, este objetivo se consigue mediante un decodificador de audio según la  
reivindicación 1, un aparato para generar datos de salida de audio codificados según la reivindicación 7, un procedimiento de decodificación de datos de entrada de audio según la reivindicación 15, un procedimiento para generar datos de audio codificados según la reivindicación 18 y un programa informático según la  
50 reivindicación 20.

**[0009]** Las realizaciones de la invención proporcionan un decodificador de audio para decodificar una corriente de bits de datos de audio codificados, en el que la corriente de bits de datos de audio codificado representa una secuencia de valores de muestra de audio y comprende una pluralidad de tramas, en el que  
55 cada trama incluye valores de muestra de audio codificados asociados, el decodificador de audio que comprende:

un determinador configurado para determinar si una trama de los datos de audio codificados es una trama especial que comprende valores de muestra de audio codificados asociados con la trama especial e

- información adicional, en la que la información adicional comprende valores de muestra de audio codificados de un número de tramas que preceden a la trama especial, en el que los valores de la muestra de audio codificados de las tramas precedentes están codificados utilizando la misma configuración de códec que la trama especial, en el que el número de tramas precedentes, correspondiente a las tramas prerrodadas, se
- 5 corresponde con el número de tramas que el decodificador necesita para construir la señal completa durante la puesta en marcha del decodificador de modo que esté en posición de decodificar los valores de muestra de audio asociados con la trama especial si la trama especial es la primera trama después de la puesta en marcha del decodificador; y
- 10 un inicializador configurado para inicializar el decodificador si el determinador determina que la trama es una trama especial, en el que inicializar el decodificador comprende decodificar los valores de muestra de audio codificados incluidos en la información adicional antes de decodificar los valores de muestra de audio codificados asociados con la trama especial,
- 15 en el que el inicializador está configurado para conmutar el decodificador de audio desde una configuración de códec actual a una configuración de códec diferente si el determinador determina que la trama es una trama especial y si los valores de muestra de audio de la trama especial se han codificado utilizando la configuración de códec diferente, y
- 20 en el que el decodificador está configurado para decodificar la trama especial utilizando la configuración de códec actual y descartar la información adicional si el determinador determina que la trama es una trama especial y si los valores de muestra de audio de la trama especial se han codificado utilizando la configuración de códec actual.
- 25 **[0010]** Las realizaciones de la invención proporcionan un aparato para generar una corriente de bits de datos de audio codificados, que representa una secuencia de valores de muestra de audio de una señal de audio, en el que la corriente de bits de datos de audio codificado comprende una pluralidad de tramas en la que cada trama incluye valores de muestra de audio codificados asociados, en el que el aparato comprende:
- 30 un proveedor de trama especial configurado para proporcionar al menos una de las tramas como una trama especial, la trama especial que comprende valores de muestra de audio codificados asociados con la trama especial e información adicional, en el que la información adicional comprende valores de muestra de audio codificados de un número de tramas que preceden a la trama especial, en el que los valores de la muestra de audio codificada de las tramas precedentes están codificados utilizando la misma configuración de códec que
- 35 la trama especial, y en el que el número de tramas precedentes, correspondiente a tramas prerrodadas, se corresponde con el número de tramas que necesita un decodificador para construir la señal completa durante la puesta en marcha del decodificador de modo que esté en posición de decodificar los valores de muestra de audio asociados con la trama especial si la trama especial es la primera trama después de la puesta en marcha del decodificador, y
- 40 una salida configurada para emitir la corriente de bits de datos de audio codificados,
- en el que los datos de audio codificados comprenden una pluralidad de segmentos, en el que cada segmento está asociado con una de una pluralidad de partes de la secuencia de valores de muestra de audio y
- 45 comprende una pluralidad de tramas, en el que el sumador de tramas especial está configurado para añadir una trama especial en el comienzo de cada segmento independientemente de si la configuración del códec cambia o no.
- [0011]** Las realizaciones de la invención proporcionan un procedimiento de decodificación una corriente
- 50 de bits de datos de audio codificados, en el que la corriente de bits de datos de audio codificados representa una secuencia de valores de muestra de audio y comprende una pluralidad de tramas, en la que cada trama incluye valores de muestra de audio codificados asociados, que comprende:
- determinar si una trama de los datos de audio codificados es una trama especial que comprende valores de
- 55 muestra de audio codificados asociados con la trama especial e información adicional, en la que la información adicional comprende valores de muestra de audio codificados de un número de tramas que preceden a la trama especial, en la que los valores de muestra de las tramas precedentes se codifican utilizando la misma configuración de códec que la trama especial, en la que el número de tramas precedentes, correspondiente a las tramas prerrodadas, se corresponde con el número de tramas que necesita un decodificador para construir

la señal completa durante la puesta en marcha del decodificador de modo que esté en posición de decodificar los valores de muestra de audio asociados con la trama especial si la trama especial es la primera trama después de la puesta en marcha del decodificador;

5 inicializar el decodificador si se determina que la trama es una trama especial, en el que la inicialización comprende decodificar los valores de muestra de audio codificados incluidos en la información adicional antes de decodificar los valores de muestra de audio codificados asociados con la trama especial;

10 conmutar el decodificador de audio desde una configuración de códec actual a una configuración de códec diferente si se determina que la trama es una trama especial y si los valores de muestra de audio de la trama especial se han codificado utilizando la configuración de códec diferente; y

15 decodificar la trama especial utilizando la configuración de códec actual y descartar la información adicional si se determina que la trama es una trama especial y si los valores de muestra de audio de la trama especial se han codificado utilizando la configuración de códec actual.

**[0012]** Las realizaciones de la invención proporcionan un procedimiento para generar una corriente de bits de datos de audio codificados que representan una secuencia de valores de muestra de audio de una señal de audio, en el que la corriente de bits de datos de audio codificados comprende una pluralidad de  
20 tramas, en el que cada trama incluye valores de muestra de audio codificados asociados, que comprende:

proporcionar al menos una de las tramas como una trama especial, la trama especial que comprende valores de muestra de audio codificados asociados con la trama especial e información adicional, en la que la información adicional comprende valores de muestra de audio codificados de un número de tramas que  
25 preceden a la trama especial, en la que los valores de la muestra de audio codificada de las tramas precedentes están codificados utilizando la misma configuración de códec que la trama especial, y en el que el número de tramas precedentes, correspondiente a tramas prerrodadas, se corresponde con el número de tramas que necesita el decodificador para construir la señal completa durante la puesta en marcha del decodificador de modo que esté en posición de decodificar los valores de muestra de audio asociados con la  
30 trama especial si la trama especial es la primera trama después de la puesta en marcha del decodificador, y

generar la corriente de bits concatenando la trama especial y las otras tramas de la pluralidad de tramas,

35 en el que los datos de audio codificados comprenden una pluralidad de segmentos, en el que cada segmento está asociado con una de una pluralidad de partes de la secuencia de valores de muestra de audio y comprende una pluralidad de tramas, en el que se añade una trama especial al comienzo de cada segmento independientemente de si la configuración del códec cambia o no.

**[0013]** Las realizaciones de la invención se basan en el hallazgo de que se puede conseguir la  
40 reproducción inmediata de una corriente de bits de datos de audio codificados, que representan una secuencia de valores de muestra de audio de una señal de audio y que comprende una pluralidad de tramas, si una de las tramas se proporciona como una trama especial que incluye valores de muestra de audio codificados asociados con las tramas precedentes, que son necesarios para iniciar un decodificador que esté en posición de decodificar los valores de muestra de audio codificados asociados con la trama especial. Por consiguiente,  
45 el número de tramas necesarios para iniciar el decodificador depende de la configuración de códec utilizada y es conocido por las configuraciones de códecs. Las realizaciones de la invención se basan en el hallazgo de que la conmutación entre diferentes configuraciones de códecs se puede conseguir de una manera beneficiosa si dicha trama especial está dispuesta en una posición en la cual tendrá lugar la conmutación entre las configuraciones de codificación. La trama especial no solo puede incluir valores de muestra de audio  
50 codificados asociados con la trama especial, sino también información adicional que permite una conmutación entre configuraciones de códecs y una reproducción inmediata después de la conmutación. En las realizaciones de la invención, el aparato y el procedimiento para generar datos de salida de audio codificados y el codificador de audio están configurados para preparar datos de audio codificados de tal manera que puede tener lugar una respuesta inmediata en el lado del decodificador al conmutar entre configuraciones de códecs.  
55 En las realizaciones de la invención, dichos datos de audio generados y emitidos en el lado del codificador se reciben como datos de entrada de audio en el lado del decodificador y permiten una reproducción inmediata en el lado del decodificador. En las realizaciones de la invención, se permite la repetición inmediata en el lado del decodificador al conmutar entre diferentes configuraciones de códecs en el lado del decodificador.

**[0014]** En las realizaciones de la invención, la información adicional comprende información sobre la configuración de códec utilizada para codificar los valores de muestra de audio asociados con la trama especial, en el que el determinador está configurado para determinar si la configuración de códec de la información adicional es diferente de la configuración de códec actual.

5

**[0015]** En las realizaciones de la invención, el decodificador de audio comprende un desvanecedor cruzado configurado para ejecutar un desvanecimiento cruzado entre una pluralidad de valores de muestra de salida obtenidos utilizando la configuración de códec actual y una pluralidad de valores de muestra de salida obtenidos mediante la decodificación de los valores de muestra de audio codificados asociados con la trama especial. En las realizaciones de la invención, el desvanecedor cruzado está configurado para ejecutar el desvanecimiento cruzado de valores de muestra de salida obtenidos mediante el vaciamiento del decodificador en la configuración de códec actual y valores de muestra de salida obtenidos mediante la decodificación de los valores de muestra de audio codificados asociados con la trama especial.

10

**[0016]** En las realizaciones de la invención, una primera trama del número de tramas comprendidas en la información adicional no está codificada de manera diferencial en el tiempo o codificada por entropía en relación con ninguna trama anterior a la trama más temprana y en la que la trama especial no está codificada de manera diferencial en el tiempo o codificada por entropía en relación con cualquier trama anterior a la primera trama del número de tramas que preceden a la trama especial o en relación con cualquier trama anterior a la trama especial.

20

**[0017]** En las realizaciones de la invención, la trama especial comprende la información adicional como una carga de extensión y en la que el determinador está configurado para evaluar la carga de extensión de la trama especial. En las realizaciones de la invención, la información adicional comprende información sobre la configuración de códec utilizada para codificar los valores de muestra de audio asociados con la trama especial.

25

**[0018]** En las realizaciones de la invención, los datos de audio codificados comprenden una pluralidad de segmentos, en los que cada segmento está asociado con una de una pluralidad de partes de la secuencia de valores de muestra de audio y comprende una pluralidad de tramas, en los que el sumador de tramas especiales está configurado para sumar una trama especial al comienzo de cada segmento.

30

**[0019]** En la realización de la invención, los datos de audio codificados comprenden una pluralidad de segmentos, en los que cada segmento está asociado con una de una pluralidad de partes de la secuencia de valores de muestra de audio y comprende una pluralidad de las tramas, en los que el aparato para generar un bit de datos de audio codificados comprende un proveedor de segmentos configurado para proporcionar segmentos asociados con diferentes partes de la secuencia de valores de muestra de audio y codificados por diferentes configuraciones de códec, en los que el proveedor de tramas especiales está configurado para proporcionar una primera trama, de al menos uno de los segmentos, como la trama especial; y un generador configurado para generar los datos de salida de audio disponiendo al menos uno de los segmentos después de otro de los segmentos. En las realizaciones de la invención, el proveedor de segmentos está configurado para seleccionar una configuración de códec para cada segmento basado en una señal de control. En las realizaciones de la invención, el proveedor de segmentos está configurado para proporcionar m versiones codificadas de la secuencia de valores de muestra de audio, con  $m \geq 2$ , en el que las m versiones codificadas se codifican utilizando diferentes configuraciones de códec, en el que cada versión codificada comprende una pluralidad de segmentos que representa la pluralidad de partes de la secuencia de valores de muestra de audio, en el que el proveedor de tramas especiales está configurado para proporcionar una trama especial en el comienzo de cada uno de los segmentos.

35

40

45

**[0020]** En las realizaciones de la invención, el proveedor de segmentos comprende una pluralidad de codificadores, cada uno configurado para codificar, al menos en parte, la señal de audio según una de la pluralidad de configuraciones de códec diferentes. En las realizaciones de la invención, el proveedor de segmentos comprende una memoria que almacena las m versiones codificadas de la secuencia de valores de muestra de audio.

50

55

**[0021]** En las realizaciones de la invención, la información adicional está en forma de una carga de extensión de la trama especial.

**[0022]** En las realizaciones de la invención, la corriente de bits de datos de audio codificados comprende

un primer número de tramas codificadas que utiliza una primera configuración de códec y un segundo número de tramas que sigue al primer número de tramas y codificadas utilizando una segunda configuración de códec, en el que la primera trama del segundo número de tramas es la trama especial.

5 **[0023]** En las realizaciones de la invención, la información adicional comprende información sobre la configuración de códec utilizada para codificar los valores de muestra de audio asociados con la trama especial y el procedimiento comprende determinar si la configuración de códec de la información adicional es diferente de la configuración de códec actual que utiliza los valores de muestra de audio codificados en la corriente de bits, que preceden a la trama especial, que están codificados.

10

**[0024]** Por lo tanto, en las realizaciones de la invención, se realiza el desvanecimiento cruzado para permitir la conmutación sin fisuras entre diferentes configuraciones de códec. En las realizaciones de la invención, la información adicional de la trama especial comprende las tramas prerrodadas necesarias con el fin de inicializar un decodificador para que esté en posición de decodificar la trama especial. En otras palabras, 15 en las realizaciones de la invención, la información adicional comprende una copia de dichas tramas de valores de muestra de audio codificados que preceden a la trama especial y se codifican utilizando la misma configuración de códec que los valores de muestra de audio codificados representados por la trama especial necesaria para inicializar el decodificador y ponerlo en posición de decodificar los valores de muestra de audio asociados con la trama especial.

20

**[0025]** En las realizaciones de la invención, se introducen tramas especiales en los datos de audio codificados a intervalos de tiempo regulares, es decir, de una manera periódica. En las realizaciones de la invención, una primera trama de cada segmento de los datos de audio codificados es una trama especial. En realizaciones, el decodificador de audio está configurado para decodificar las tramas especiales y las tramas 25 siguientes utilizando la configuración de códec indicada en la trama especial hasta que se encuentre una trama especial adicional que indique una configuración de códec diferente.

**[0026]** En las realizaciones de la invención, el decodificador y el procedimiento de decodificación están configurados para realizar un desvanecimiento cruzado cuando cambian de una configuración de códec a otra configuración de códec, con el fin de permitir la conmutación sin fisuras entre múltiples representaciones de 30 audio comprimidas.

**[0027]** En las realizaciones de la invención, las diferentes configuraciones de códec son diferentes configuraciones de códec según la norma AAC (Advanced Audio Coding), es decir, diferentes configuraciones 35 de códec de los códec de la familia AAC. Las realizaciones de la invención pueden estar dirigidas a conmutar entre configuraciones de códec de los códec de la familia AAC y configuraciones de códec de los códec de la familia AMR (Adaptive Multiple Rate).

**[0028]** De este modo, las realizaciones de la invención permiten la reproducción inmediata en el lado del 40 decodificador y la conmutación entre diferentes configuraciones de códec, de modo que la manera en que se suministra el contenido de audio puede adaptarse a las condiciones ambientales, como un canal de transmisión con una velocidad de transferencia de bits variable. De este modo, realizaciones de la invención permiten proporcionar al consumidor la mejor calidad de audio posible para un estado de la red dado.

45 **[0029]** A continuación, se analizan las realizaciones de la invención con respecto a los dibujos adjuntos en los que:

La Fig. 1 muestra una vista esquemática de una realización de un aparato para generar datos de salida de audio codificados;

50 La Fig. 2 muestra una vista esquemática para explicar una realización de una trama especial;

La Fig. 3 muestra una vista esquemática de diferentes representaciones de una señal de audio;

La Fig. 4a y 4b muestran vistas esquemáticas de aparatos para generar datos de salida de audio codificados;

La Fig. 5 muestra una vista esquemática de un decodificador de audio;

55 La Fig. 6 muestra un diagrama de bloques esquemático para explicar una realización de un decodificador de audio y un procedimiento de decodificación;

La Fig. 7 muestra un diagrama de bloques esquemático para explicar la conmutación de un decodificador de audio entre diferentes configuraciones de códec;

La Fig. 8 muestra un diagrama esquemático para explicar el comportamiento del decodificador AAC (Advanced Audio Coding);

La Fig. 9 muestra la conmutación de una primera corriente 1 a una segunda corriente 2; y La Fig. 10 muestra un elemento de sintaxis ejemplar que proporciona información adicional.

**[0030]** En general, las realizaciones de la invención tienen como objetivo el suministro de contenido de audio, posiblemente combinado con el suministro de vídeo, a través de un canal de transmisión con una velocidad de transferencia de bits variable. El objetivo puede ser proporcionar al consumidor la mejor calidad de audio posible para un estado de la red determinado. Las realizaciones de la invención se centran en la implementación de códecs de la familia AAC en un entorno de transmisión adaptativa.

**[0031]** En las realizaciones de la invención, como las que se utilizan en la presente memoria, los valores de muestra de audio que no están codificados representan valores de muestra de audio en el dominio del tiempo, como muestras PCM (modulación por código de impulsos). En las realizaciones de la invención, el término valor de muestra de audio codificado se refiere a valores de muestra en el dominio de la frecuencia obtenidos después de codificar los valores de muestra de audio en el dominio del tiempo. En las realizaciones de la invención, los valores de muestras o muestras de audio codificados se obtienen mediante la conversión de las muestras en el dominio del tiempo en una representación espectral, por ejemplo, mediante una MDCT (transformada discreta del coseno modificada) y codificando el resultado, por ejemplo, mediante cuantificación y codificación Huffman. De acuerdo con ello, en la realización de la invención, codificar significa obtener las muestras en el dominio de la frecuencia de las muestras en el dominio del tiempo y decodificar significa obtener las muestras en el dominio del tiempo de las muestras en el dominio de la frecuencia. Los valores de muestra (muestras) obtenidos, mediante la decodificación de los datos de audio codificados, a veces en este documento se denominan valores de muestra de salida (muestras).

**[0032]** La Fig. 1 muestra una realización de un aparato para generar datos de salida de audio codificados. La Fig. 1 muestra un escenario típico de transmisión de audio adaptativa, en la cual se pueden aplicar las realizaciones de la invención. Una señal de entrada de audio 10 está codificada por varios codificadores de audio 12, 14, 16 y 18, es decir, codificadores 1 a m. Los codificadores 1 a m se pueden configurar para codificar la señal de entrada de audio 10 simultáneamente. Normalmente, los codificadores 1 a m pueden configurarse de tal manera que se pueda lograr un amplio intervalo de tasa de bits. Los codificadores generan diferentes representaciones, es decir, versiones codificadas 22, 24, 26 y 28 de la señal de entrada de audio 10, es decir, representaciones 1 a m. Cada representación incluye una pluralidad de segmentos 1 a k, en el que el segundo segmento de la primera representación ha recibido el número de referencia 30, únicamente con fines ejemplares. Cada segmento comprende una pluralidad de tramas (unidades de acceso) designadas con las letras AU y un índice respectivo 1 a n que indica la posición de la trama en la representación respectiva. La octava trama de la primera representación recibe el número de referencia 40 solo con fines ejemplares.

**[0033]** Los codificadores 12, 14, 16 y 18 están configurados para insertar puntos de acceso a la corriente (SAP) 42 a intervalos de tiempo regulares, que definen los tamaños de los segmentos. De este modo, un segmento, como el segmento 30, consiste en múltiples tramas, como AU5, AU6, AU7 y AU8, en el que la primera trama, AU5, representa un SAP 42. En la Fig. 1, los SAP se indican con un sombreado. Cada representación 1 a m representa una representación de audio comprimida (CAR) para la señal de entrada de audio 10 y consiste en k de dichos segmentos. La conmutación entre diferentes CAR puede tener lugar en los límites de segmento.

**[0034]** En el lado del decodificador, un cliente puede solicitar una de las representaciones que mejor se ajuste a una situación dada, por ejemplo, para estados de la red determinados. Si por alguna razón el estado cambia, el cliente debe poder solicitar un CAR diferente, el aparato para generar los datos de salida codificados debe ser capaz de conmutar entre diferentes CAR en cada límite de segmento, y el decodificador debe ser capaz de conmutar para decodificar el CAR diferente en cada límite de segmento. Por lo tanto, el cliente estaría en una posición de adaptar el promedio de tasa de bits a la tasa de bits disponible en el canal con el fin de maximizar la calidad al mismo tiempo que se minimiza el búfer bajo ejecuciones ("re-buffering"). Si se utiliza HTTP (Hyper Text Transfer Protocol) para descargar los segmentos, dicha arquitectura de transmisión puede denominarse como transmisión adaptativa sobre HTTP. Las implementaciones actuales incluyen Apple HTTP Live Streaming (HLS), Microsoft Smooth Streaming y Adobe Dynamic Streaming, que siguen el principio básico. Recientemente, MPEG ha lanzado un estándar abierto: Dynamic Adaptive Streaming sobre HTTP (MPEG DASH), véase "Guidelines for Implementation: DASH-AVC/264 Interoperability Points", <http://dashif.org/w/2013/08/DASH-AVC-264-v2.00-hd-mca.pdf>. HTTP normalmente utiliza TCP/IP (Protocolo de control de transmisión/Protocolo de Internet) como el protocolo de red subyacente. Las realizaciones de la

invención se pueden aplicar a todos los desarrollos actuales.

**[0035]** Una conmutación entre representaciones (versiones codificadas) debe ser con las mínimas fisuras posible. En otras palabras, no debe haber ningún defecto o clic audible durante la conmutación. Sin otras medidas previstas por las realizaciones de la invención, este requisito solo puede conseguirse con determinadas limitaciones y si se toma especial cuidado durante el procedimiento de codificación.

**[0036]** En la Fig. 1, el respectivo codificador del que procede un segmento se indica mediante una marca respectiva colocada dentro de un círculo. La Fig. 1 muestra además un motor de decisión 50, que decide qué representación hay que descargar para cada segmento. Un generador 52 genera datos de salida de audio codificados 54 a partir de los segmentos seleccionados que reciben los números de referencia 44, 46 y 48 en la Fig. 1 concatenando los segmentos seleccionados. Los datos de salida de audio codificados 54 pueden ser suministrados a un decodificador 60 configurado para decodificar los datos de salida de audio codificados en una señal de salida de audio 62 que comprende las muestras de salida de audio.

**[0037]** En la realización que se muestra en la Fig. 1, los segmentos y, por lo tanto, las tramas, procedentes de diferentes codificadores, se introducen en el mismo decodificador 60, por ejemplo, AU4 desde el codificador 2 y AU5 desde el codificador 3 en el ejemplo de la Fig. 1. En el caso de que se use la misma instancia de decodificador para decodificar estas AU es necesario que ambos codificadores sean compatibles entre sí. En particular, sin medidas adicionales, esta estrategia no puede funcionar si los dos codificadores proceden de una familia de códecs completamente diferente, por ejemplo, AMR para el codificador 2 y G.711 para el codificador 3. Sin embargo, incluso cuando se utiliza el mismo códec en todas las representaciones, se debe tener especial cuidado en restringir el procedimiento de codificación. Esto se debe a que un códec de audio moderno, como Advanced Audio Coding (AAC), son algoritmos flexibles que pueden funcionar en diversas configuraciones utilizando diversas herramientas y modos de codificación. Ejemplos de dichas herramientas de codificación en AAC son Spectral Band Replication (SBR) o Short Blocks (SB). Otros parámetros de configuración importantes son la frecuencia de muestreo (fs, por ejemplo, 48 kHz) o la configuración del canal (mono, estéreo, envolvente). Con el fin de decodificar correctamente las tramas (AU), el decodificador debe saber qué herramientas se utilizan y cómo se configuran (por ejemplo, fs o frecuencia de cruce doble SBR). Por lo tanto, en general, la información requerida se codifica en una cadena de configuración corta y se pone a disposición del decodificador antes de la decodificación. Estos parámetros de configuración pueden denominarse como configuración de códec. En el caso de AAC, esta configuración se conoce como la configuración específica de audio (ASC).

**[0038]** Hasta ahora, a fin de lograr una conmutación sin fisuras, era necesario restringir la configuración del códec para que fuera compatible entre las representaciones (versiones codificadas). Por ejemplo, la frecuencia de muestreo o las herramientas de codificación normalmente deben ser idénticas en todas las representaciones. Si se utilizan configuraciones de códecs incompatibles entre representaciones, entonces el decodificador tiene que ser reconfigurado. Esto significa básicamente que el decodificador viejo tiene que cerrarse y el decodificador nuevo tiene que iniciarse con una nueva configuración. Sin embargo, este procedimiento de reconfiguración no es transparente en todas las circunstancias y puede causar un defecto. Una posible explicación es que el decodificador nuevo no puede producir muestras válidas inmediatamente, sino que requiere varias AU prerrodadas para construir la intensidad de la señal completa. Este comportamiento de puesta en marcha es típico de los códecs que tienen un estado del decodificador, es decir, en los casos en que la decodificación de la AU actual no es completamente independiente de la decodificación de AU anteriores.

**[0039]** Como resultado de este comportamiento, normalmente se requería que la configuración del códec fuera constante en todas las representaciones y el único parámetro cambiante era la tasa de bits. Este es, por ejemplo, el caso del perfil DASH-AVC/264 tal como se define por el DASH Industry Forum.

**[0040]** Esta restricción limitó la flexibilidad del códec y, por lo tanto, la eficacia de la codificación en el intervalo completo de tasa de bits. Por ejemplo, SBR es una valiosa herramienta de codificación para velocidades de transferencia de bits muy bajas, pero limita la calidad de audio a velocidades de transferencia de bits altas. Por lo tanto, si se requiere que la configuración codificada sea constante, es decir, con o sin SBR, se tenía que llegar a una solución de compromiso sobre una velocidad de transferencia de bits alta o baja. Del mismo modo, la eficacia de la codificación podría beneficiarse de cambiar la frecuencia de muestreo a través de las representaciones, pero tendría que mantenerse constante debido a las limitaciones antes mencionadas para una conmutación sin fisuras.

**[0041]** Las realizaciones de la presente invención están dirigidas a una estrategia novedosa que permite la conmutación de audio sin fisuras en un entorno de transmisión adaptativo y, en particular, que permite una conmutación de audio sin fisuras para códecs de audio de la familia AAC en un entorno de transmisión adaptativo. La estrategia de la invención está diseñada para solucionar todas las deficiencias resultantes de las restricciones sobre la configuración del códec tal como se ha descrito anteriormente. El objetivo general es tener más flexibilidad en la configuración a través de las representaciones (versiones codificadas), tales como las herramientas de codificación o la frecuencia de muestreo, al mismo tiempo que se habilita o se garantiza una conmutación sin fisuras.

10

**[0042]** Las realizaciones de la invención se basan en el hallazgo de que pueden superarse las restricciones explicadas anteriormente y se puede conseguir una mayor flexibilidad añadiendo una trama especial que transporta información adicional además de valores de muestra de audio codificados asociados con la trama especial entre otras tramas de datos de audio codificados como una representación de audio comprimido (CAR). Una representación de audio comprimido puede considerarse como una pieza de material de audio (música, voz, ...) después de la compresión mediante un codificador de audio con pérdidas o sin pérdidas, por ejemplo, un codificador de audio de la familia AAC (AAC, HE-AAC, MPEG-D USAC, ...) con una velocidad de transferencia de bits global constante. En particular, la información adicional en la trama especial está diseñada para permitir una reproducción instantánea en el lado del decodificador incluso en el caso de una conmutación entre diferentes configuraciones de códecs. Por lo tanto, la trama especial puede denominarse trama de reproducción instantánea (IPF). La IPF está configurado para compensar el retardo de la puesta en marcha del decodificador y se utiliza para transmitir información de audio en tramas anteriores junto con los datos de la trama actual.

25

**[0043]** Un ejemplo de dicha IPF se muestra en la Fig. 2. La Fig. 2 muestra una serie de tramas (unidades de acceso) 40, numeradas de  $n-4$  a  $n+3$ . Cada trama incluye valores de muestra de audio codificados asociados, es decir, valores de muestra de audio codificados de un número específico de valores de muestra de audio en el dominio del tiempo de una secuencia de valores de muestra de audio en el dominio del tiempo que representan una señal de audio, tal como la señal de entrada de audio 10. Por ejemplo, cada trama puede comprender valores de muestra de audio codificados que representan 1024 valores de muestra de audio en el dominio del tiempo, es decir, valores de muestra de audio de una señal de audio no codificada. En la Fig. 2, la trama  $n$  dispuesta entre la trama anterior  $n-1$  y la trama siguiente  $n+1$  representa la trama especial o IPF 80. La trama especial 80 incluye información adicional 82. La información adicional 82 incluye información 84 sobre la configuración del códec, es decir, información sobre la configuración del códec utilizada para codificar la corriente de datos que incluye las tramas  $n-4$  a  $n+3$  y, por lo tanto, información sobre la configuración del códec utilizada para codificar valores de muestra de audio asociados con la trama especial.

30

35

**[0044]** En la realización que se muestra en la Fig. 2, se supone que un retardo introducido por un decodificador de audio es de tres tramas, es decir, se supone que se necesitan tres tramas denominadas prerrodadas para construir la señal completa durante la puesta en marcha del decodificador de audio. Por lo tanto, suponiendo que la configuración de la corriente (configuración del códec) es conocida por el decodificador, normalmente el decodificador tendría que comenzar a decodificar en la trama  $n-3$  para producir muestras válidas en la trama  $n$ . De este modo, para proporcionar la información necesaria al decodificador, la información adicional 82 comprende un número de tramas de valores de muestra de audio codificados que preceden a la trama especial 80 y se codifican utilizando la configuración de códec 84 indicada en la información adicional 82. Este número de tramas se indica mediante el número de referencia 86 en la Fig. 2. Este número de tramas 86 es necesario para inicializar el decodificador en una posición para decodificar los valores de muestra de audio asociados con la trama especial  $n$ . Por consiguiente, la información de la trama 86 se duplica y se transporta como parte de la trama especial 80. Por lo tanto, esta información está disponible para el decodificador inmediatamente después de conmutar a la corriente de datos que se muestra en la Fig 2 en la trama  $n$ . Sin esta información adicional en la trama  $n$ , ni la configuración de códec 84 ni las tramas  $n-3$  a  $n-1$  estarían disponibles para el decodificador después de una conmutación. Al añadir esta información a la trama especial 80 permite inicializar inmediatamente el decodificador y, por lo tanto, la reproducción inmediata al conmutar a una corriente de datos que comprende la trama especial. El decodificador está configurado de tal manera que dicha inicialización y decodificación de la trama  $n$  se pueden realizar dentro de la ventana de tiempo disponible hasta que las muestras de salida obtenidas por la trama de decodificación  $n$  tienen que ser emitidas.

40

45

50

55

**[0045]** Durante la decodificación normal, es decir, sin conmutar a una configuración de códec diferente,

solo se decodifica la trama  $n$  y se ignoran las tramas  $n-3$  a  $n-1$  incluidas en la información adicional. Sin embargo, después de conmutar a una configuración de códec diferente, se extrae toda la información en la trama especial 80 y se inicializa el decodificador basándose en la configuración de códec incluida y basándose en la decodificación de las tramas prerrodadas ( $n-3$  a  $n-1$ ) antes de decodificar finalmente y reproducir la trama actual  $n$ . La decodificación de las tramas prerrodadas tiene lugar antes de que la trama actual se decodifique y se reproduzca. Las tramas prerrodadas no se vuelven a reproducir, pero el decodificador está configurado para decodificar las tramas prerrodadas dentro de la ventana de tiempo disponible antes de la reproducción de la trama actual  $n$ .

10 **[0046]** El término "configuración del códec" hace referencia a la configuración del códec utilizada para codificar datos de audio o tramas de datos de audio. Por lo tanto, la configuración de la codificación puede indicar diferentes herramientas de codificación y modos utilizados, en la que las herramientas de codificación ejemplares utilizadas en AAC son una replicación de la banda espectral (SBR) o bloques cortos (SB). Un parámetro de configuración puede ser la frecuencia de cruce doble SBR. Otros parámetros de configuración  
15 pueden ser la frecuencia de muestreo o la configuración del canal. Diferentes configuraciones de códec difieren en uno o más de estos parámetros de configuración. En las realizaciones de la invención, diferentes configuraciones de códec pueden comprender también códec completamente diferentes, tales como AAC, AMR o G.711.

20 **[0047]** Por consiguiente, en el ejemplo que se ilustra en la Fig. 2, son necesarias tres tramas, es decir,  $n-3$  a  $n-1$  para compensar el retardo de puesta en marcha del decodificador. Los datos de tramas adicionales pueden ser transmitidos mediante un mecanismo de carga de extensión dentro de la corriente de bits de audio. Por ejemplo, el mecanismo de carga de extensión USAC (UsacExtElement) puede utilizarse para transportar la información adicional. Además, el campo "config" puede utilizarse para transmitir la configuración de la  
25 corriente 94. Esto puede ser útil en caso de conmutación de corrientes de bits o adaptación de la velocidad de transferencia de bits. Ambos, la primera AU prerrodada ( $n-3$ ) y la propia IPF ( $n$ ) pueden ser una trama decodificable de manera independiente. En el contexto de USAC, los codificadores pueden establecer un indicador (usacIndependencyFlag) a "1" para esas tramas. Implementando la estructura de trama como se muestra en la Fig. 2 es posible acceder a la corriente de bits en cada IPF y reproducir muestras PCM válidas  
30 de forma inmediata. El procedimiento de decodificación de una IPF puede incluir las siguientes etapas. Decodificar todas las AU "prerrodadas" ( $n-3$  ...  $n-1$ ) y descartar las muestras PCM de salida resultantes. Los estados del decodificador interno y los búferes se inicializan completamente después de esta etapa. Decodificar la trama  $n$  e iniciar la reproducción regular. Continuar decodificando de la forma habitual con la trama  $n+1$ . La IPF puede utilizarse como un punto de acceso a la corriente de audio (SAP). La reproducción  
35 inmediata de muestras PCM válidas es posible en cada IPF.

**[0048]** Las tramas especiales tal como se definen en el presente documento pueden implementarse en cualquier códec que permita la multiplexión y la transmisión de datos auxiliares o datos de extensión o elementos de corriente de datos o mecanismos similares para transmitir datos externos de códec de audio.  
40 Las realizaciones de la invención se refieren a la implementación de un entramado de códec USAC. Las realizaciones de la invención pueden implementarse en relación con codificadores y decodificadores de audio USAC. USAC significa codificación unificada de voz y audio y se hace referencia a la norma ISO/IEC 23003-3: 2012. En las realizaciones de la invención, la información adicional está contenida en una carga de extensión de la trama correspondiente, tal como la trama  $n$  en la Fig. 2. Por ejemplo, la norma USAC permite la adición  
45 de una carga de extensión arbitraria a datos de audio codificados. La existencia de una carga de extensión es conmutable sobre una base de trama a trama. Por consiguiente, la información adicional puede ser implementada como un nuevo tipo de carga de extensión definido para transportar información de audio adicional de tramas anteriores.

50 **[0049]** Como se ha explicado anteriormente, la trama de reproducción instantánea 80 está diseñada de tal manera que las muestras de salida válidas asociadas con una determinada marca de tiempo (trama  $n$ ) se pueden generar inmediatamente, es decir, sin tener que esperar el número específico de tramas según el retardo del códec de audio. Dicho de otro modo, el retardo del códec de audio puede ser compensado. En la realización que se muestra en la Fig. 2, el retardo del códec de audio es de tres tramas. Además, la IPF está  
55 diseñada de tal manera que se puede decodificar completa e independientemente, es decir, sin ningún conocimiento adicional de la corriente de audio anterior. A este respecto, el primer número de tramas añadidas a la trama especial (es decir, la trama  $n-3$  en la Fig. 2) no está codificado por tiempo diferencial o codificado por entropía en relación con ninguna trama anterior. Así mismo, la trama especial no está codificada por tiempo diferencial o codificada por entropía en relación con ninguna trama anterior al primer número de tramas

contenidas en la información adicional o cualquier trama anterior. Dicho de otro modo, en las tramas n-3 y n de la Fig. 2 pueden eliminarse todas las dependencias con las tramas anteriores, por ejemplo, la codificación diferencial en el tiempo de ciertos parámetros o restablecer la codificación por entropía. Por lo tanto, estas tramas independientes permiten una correcta decodificación y análisis de todos los símbolos, pero no son suficientes en sí mismos para obtener muestras PCM válidas instantáneamente. Mientras que dichas tramas independientes ya están disponibles en códecs de audio comunes, tales como AAC o USAC, dichos códecs de audio no proporcionan tramas especiales, como la trama IPF 80.

**[0050]** En las realizaciones de la invención, se proporciona una trama especial en cada punto de acceso a la corriente de las representaciones que se muestran en la Fig. 1. En la Fig. 1, los puntos de acceso a la corriente son la primera trama en cada segmento y están sombreados. Por consiguiente, la Fig. 1 muestra una realización específica de un aparato para generar datos de salida de audio codificados según la presente invención. Además, cada uno de los codificadores 1 a m ilustrados en la Fig. 1 representa una realización de un codificador de audio según la invención. Según la Fig. 1, los codificadores 12 a 18 representan proveedores configurados para proporcionar segmentos asociados con diferentes partes de la señal de entrada de audio y codificados mediante diferentes configuraciones de códecs. A este respecto, cada uno de los codificadores 12 a 18 utiliza una configuración de códec diferente. La unidad de decisión 50 está configurada para decidir qué representación descargar para cada segmento. De este modo, la unidad de decisión 50 está configurada para seleccionar una configuración de códec (asociada con la representación respectiva) para cada segmento sobre la base de una señal de control. Por ejemplo, la señal de control puede recibirse de un cliente que solicita la representación que mejor se adapta a una situación determinada.

**[0051]** En base a la decisión de la unidad de decisión 50, el bloque 52 genera los datos de salida de audio 54 disponiendo los segmentos uno tras otro, tal como el segmento 46 (segmento 2 de la representación 3) después del segmento 44 (segmento 1 de la representación 2). Así, la trama especial AU5 al comienzo del segmento 2 permite conmutar a la representación 3 y la reproducción inmediata en el límite entre los segmentos 44 y 46 en el lado del decodificador.

**[0052]** De este modo, en la realización que se muestra en la Fig. 1, un proveedor (que comprende los codificadores 1 a m) está configurado para proporcionar m versiones codificadas de la entrada de audio 10, con  $m \geq 2$ , en el que las m versiones codificadas (representaciones) se codifican utilizando diferentes configuraciones de códecs, en el que cada versión codificada incluye una pluralidad de segmentos que representan la pluralidad de partes de la secuencia de valores de muestra de audio, en el que cada uno de los segmentos comprende una trama especial en el comienzo del mismo.

**[0053]** En otras realizaciones de la invención, se pueden almacenar en una memoria representaciones diferentes de la misma entrada de audio, como las representaciones 22 a 28 en la Fig. 1, y se puede acceder a ellas si un usuario solicita el contenido de medios correspondiente.

**[0054]** Los ejemplos de codificador 1 a m que muestran en la Fig. 1 pueden producir un retardo de codificador diferente en función de la configuración del codificador y/o la activación de herramientas en los ejemplos de codificador. En estos casos, se pueden tomar medidas para asegurar que los retardos del codificador se compensen para conseguir una alineación en el tiempo de las m corrientes de salida, es decir, las m representaciones. Esto se puede implementar, por ejemplo, añadiendo en el extremo una cantidad de muestras a la entrada del codificador con el fin de compensar los retardos diferentes del codificador. En otras palabras, los segmentos en las diferentes representaciones tendrán la misma duración con el fin de permitir una conmutación sin fisuras entre representaciones en los límites del segmento. Las duraciones de los segmentos teóricos dependen de las frecuencias de muestreo empleadas y de los tamaños de trama. La Fig. 3 muestra un ejemplo de una posible inserción de IPF en representaciones con diferente encuadre, tal vez debido a diferentes frecuencias de muestreo y/o tamaños de trama. Se pueden añadir muestras cero a segmentos más cortos en una posición apropiada de manera que todas las tramas especiales estén alineadas en el tiempo, como puede verse en la Fig. 3.

**[0055]** La Fig. 4a muestra una vista esquemática de un aparato 90 para generar datos de salida de audio codificados 102; El aparato 90 comprende un proveedor 92 configurado para proporcionar al menos una trama 80 de una pluralidad de tramas 40 como una trama especial tal como se ha definido en este documento. En las realizaciones de la invención, el proveedor 92 puede implementarse como parte de un codificador para codificar valores de muestra de audio, que proporciona las tramas 40 y añade la información adicional a, al menos una, de las tramas al fin de generar la trama especial. Por ejemplo, el proveedor 92 puede configurarse

para añadir la información adicional como una extensión de la carga a una de las tramas 40 a fin de generar la trama especial 80. Las tramas 40, 80 que representan la corriente de bits de datos de audio codificados 102 se emiten a través de una salida 112.

5 **[0056]** La Fig. 4b muestra una vista esquemática de un aparato 100 para generar datos de salida de audio codificados 102; El aparato comprende un proveedor 104 configurado para proporcionar segmentos 106, 108 asociados con diferentes partes de una secuencia de valores de muestra de audio. Una primera trama de, al menos uno, de los segmentos es una trama especial tal como se ha explicado anteriormente. Un generador 110 está configurado para generar los datos de salida de audio disponiendo, al menos uno, de los segmentos  
10 106, 108 después de otro de los segmentos 106, 108. El generador 110 entrega los datos de salida de audio a la salida 112 configurada para emitir los datos de audio codificados 102.

**[0057]** La Fig. 5 muestra una vista esquemática de una realización del decodificador de audio 60 para decodificar los datos de entrada de audio 122. Los datos de entrada de audio pueden ser la salida del bloque  
15 52 que se muestra en la Fig. 1. El decodificador de audio 60 comprende un determinador 130, un inicializador 132 y un núcleo de decodificador 134. El determinador 130 está configurado para determinar si una trama de datos 122 de entrada de audio es una trama especial. El inicializador 132 se configura para inicializar el núcleo del decodificador 134 si la trama es una trama especial y la inicialización es necesaria o deseada. La inicialización comprende decodificar las tramas precedentes incluidas en la información adicional. El núcleo del  
20 decodificador 134 está configurado para decodificar tramas de valores de muestra de audio codificados utilizando la configuración de códec con la que se inicializa.

**[0058]** En el caso de que la trama no sea una trama especial, se suministra directamente al núcleo 134 del decodificador, flecha 136. En el caso de que la trama sea una trama especial y no se requiera la  
25 inicialización del núcleo 134 del decodificador, el determinador 130 puede descartar la información adicional y solo entregar los valores de la muestra de audio codificada de la trama especial (sin las tramas en la información adicional) a el decodificador 134. El determinador 130 puede configurarse para determinar si es necesario inicializar el núcleo del decodificador 134 basándose en la información incluida en la información adicional o basándose en información externa. La información incluida en la información adicional puede ser  
30 información sobre la configuración de códec utilizada para codificar la trama especial, en la que el determinador puede decidir que la inicialización es necesaria si esta información indica que las tramas anteriores están codificadas utilizando una configuración de códec diferente como la trama especial. La información externa puede indicar que el núcleo del decodificador 134 debe ser inicializado o reinicializado al recibir la trama especial siguiente.

35 **[0059]** En las realizaciones de la invención, el decodificador 60 está configurado para iniciar el núcleo del decodificador 134 en una de las diferentes configuraciones de códecs. Por ejemplo, pueden iniciarse diferentes ejemplos de un núcleo de decodificador de software utilizando diferentes configuraciones de códec, es decir, diferentes parámetros de configuración de códec como se ha explicado anteriormente. En las  
40 realizaciones de la invención, la inicialización del decodificador (núcleo) puede comprender cerrar una instancia de decodificador actual y abrir una nueva instancia de decodificador utilizando los parámetros de configuración de códec incluidos en la información adicional (es decir, dentro de la corriente de bits recibida) o suministrados externamente, por ejemplo, externos a la secuencia de bits recibida. El decodificador 60 puede conmutarse a diferentes configuraciones de códecs en función de las configuraciones de códecs utilizadas para codificar los  
45 segmentos respectivos de los datos de audio codificados recibidos.

**[0060]** El decodificador 60 puede configurarse para conmutar desde una configuración del códec actual, es decir, la configuración del códec del decodificador de audio antes de encontrar la trama especial, a una configuración de códec diferente si la información adicional indica una configuración de códec diferente de la  
50 configuración del códec actual.

**[0061]** Otros detalles de una realización de un decodificador de audio que tiene el comportamiento de un decodificador AAC se explican haciendo referencia a las Figs. 6 a 8. La Fig. 8 muestra esquemáticamente el comportamiento de un decodificador AAC. Se hace referencia a la norma ISO/IEC DTR 14496-24, "Audio and  
55 Systems Interaction".

**[0062]** La Fig. 8 muestra el comportamiento del decodificador en un número de estados, un primer estado 200 correspondiente a una o más tramas prerrodadas, un estado asociado con cada una de las tramas AU1, AU2 y AU3, y un estado de "vaciamiento" 202.

**[0063]** Para generar muestras de salida válidas para AU1, tanto la o la(s) trama(s) prerrodada(s) y la trama AU1 tienen que ser decodificadas. Las muestras generadas por la o la(s) trama(s) prerrodada(s) se descartan, es decir, se utilizan solamente para inicializar el decodificador y no se reproducen. Sin embargo, la decodificación de la o la(s) trama(s) prerrodada(s) es obligatoria para configurar los estados internos del decodificador. En las realizaciones de la invención, la información adicional de las tramas especiales incluye la o la(s) trama(s) prerrodada(s). De este modo, el decodificador está en una posición para decodificar la o la(s) trama(s) prerrodada(s) y configurar los estados internos del decodificador de modo que la trama especial pueda ser decodificada y pueda tener lugar la reproducción inmediata de las muestras de salida válidas de la trama especial. El número real de AU (tramas) "prerrodadas" depende del retardo de puesta en marcha del decodificador, en el ejemplo de la Fig. 8 una AU.

**[0064]** En general, para la reproducción de archivos, la reproducción inmediata tal como se ha descrito haciendo referencia a la Fig. 8 se implementa a nivel del sistema. Hasta ahora, solo tiene lugar en la puesta en marcha del decodificador. Sin embargo, una trama especial (IPF) siempre transporta suficiente información para inicializar completamente los estados internos del decodificador y llenar los búferes internos. Por lo tanto, la inserción de tramas especiales permite la reproducción inmediata en posiciones de corriente aleatorias.

**[0065]** El estado de vaciamiento 202 en la Fig. 8 muestra el comportamiento del decodificador si el vaciamiento se realiza después de decodificar la última trama AU3. El vaciamiento significa introducir una hipotética trama de cerro en el descodificador, es decir, una hipotética trama compuesta por todas las muestras de entrada de "ceros digitales". Debido a la suma de superposición de la familia AAC, el vaciamiento da como resultado una salida válida que se logra sin consumir una nueva trama de entrada. Esto es posible puesto que la última trama AU3 incluye información de predicción sobre los valores de la muestra de salida que se obtendrían al decodificar una trama siguiente después de la trama AU3, ya que las tramas se superponen sobre un número de valores de la muestra en el dominio del tiempo. En general, la primera mitad de una trama se superpone con una trama precedente y una segunda mitad de una trama se superpone con una trama siguiente. De este modo, la segunda mitad de los valores de la muestra de salida obtenidos al decodificar una primera trama incluye información sobre la primera mitad de los valores de la muestra de salida obtenida al decodificar una segunda trama después de la primera trama. Esta característica se puede explotar cuando se implementa un desvanecimiento cruzado como se explicará a continuación.

**[0066]** A continuación, se describen detalles adicionales de una realización de un decodificador de audio y un procedimiento de decodificación de datos de entrada de audio, haciendo referencia a la Fig. 6, en la que el decodificador de audio está configurado para llevar a cabo el procedimiento descrito de las Figs. 6 y 7. El procedimiento comienza en 300. El decodificador escanea las tramas entrantes (AU) para una IPF y determina si una trama entrante es una IPF, 302. Si la trama entrante no es una IPF, la trama se decodifica, 304, y el procedimiento salta a la siguiente trama, 306. Si no hay siguiente trama, el procedimiento termina. Las muestras de PCM decodificadas se emiten, como se indica mediante el bloque 308, que puede representar un búfer de salida. Si se determina en 302 que la trama es una IPF, se evalúa la configuración del códec, 310. Por ejemplo, se evalúa el campo "config" que se muestra en la Fig. 2. Se realiza una determinación de si la configuración del códec (configuración de la corriente) ha cambiado, 312. Si la configuración de códec no ha cambiado, es decir, si la información adicional indica una configuración de códec idéntica a la configuración del códec actual, la información adicional, por ejemplo, la carga de extensión, se omite y el procedimiento salta a 304, en cuyo caso la decodificación continúa de la forma habitual.

**[0067]** Si la configuración del códec ha cambiado, se aplican las etapas siguientes. El decodificador se vacía, 314. Las muestras de salida resultantes del vaciamiento del decodificador se almacenan en un búfer de vaciamiento, 316. Estas muestras de salida (o al menos una parte de estas muestras de salida) son una primera entrada a un procedimiento de desvanecimiento cruzado 318. A continuación, el decodificador se reinicializa utilizando la nueva configuración de códec indicada en la información adicional, por ejemplo, en el campo "config" de la Fig. 2, y utilizando las tramas precedentes comprendidas en la trama especial. Tras la reinicialización, el decodificador es capaz de decodificar la trama especial, es decir, los valores de muestra de audio codificados asociados con la trama especial. La trama especial se decodifica, 322. Las muestras de salida (muestras PCM) obtenidas con la decodificación de la trama especial se almacenan como una segunda entrada para el procedimiento de desvanecimiento cruzado 318. Por ejemplo, las muestras de salida de PCM correspondientes pueden almacenarse en un búfer, 324, que puede denominarse búfer IPF. En el procedimiento de desvanecimiento cruzado 318, se calcula un desvanecimiento cruzado basado en las dos señales de entrada procedentes del búfer de vaciamiento y del búfer IPF. El resultado del desvanecimiento

cruzado se emite como muestras de salida PCM, bloque 308. A continuación, el procedimiento salta a la trama siguiente 306 y el procedimiento se repite para la trama siguiente. En caso de que la trama actual sea la última, el procedimiento finaliza.

5 **[0068]** A continuación, se explican más detalles de las etapas realizadas tras detectar un cambio de configuración en 312, haciendo referencia a la Fig. 7. La configuración de códec se recupera de la información adicional de la IPF, 330 y se proporciona para la reinicialización del decodificador 332. Antes de reinicializar el decodificador, el decodificador se vacía, 314, y las muestras de salida resultantes se almacenan en el búfer de vaciamiento, 316. La reinicialización del decodificador puede incluir el cierre de la instancia del decodificador  
10 actual y la apertura de la nueva instancia del decodificador con la configuración nueva. Al volver a abrir la nueva instancia del decodificador, se utiliza la información sobre la configuración de códec contenida en la IPF. Después de abrir la nueva instancia del decodificador, se inicializa decodificando las tramas prerrodadas incluidas en la IPF. Se supone que el número de tramas prerrodadas contenidas en la IPF es  $m$ , como se indica en el bloque 334. Se determina si  $m > 0$ , 336. Si  $m > 0$ , se decodifica la trama prerrodada  $n-m$ , 338, en la  
15 que  $n$  indica la IPF. Las muestras PCM de salida obtenidas se descartan 340.  $m$  se reduce en uno y el procedimiento salta al bloque 336. Repitiendo las etapas 336 a 342 para todas las tramas prerrodadas contenidas en la IPF, se realiza un procedimiento de llenado de los estados decodificadores del decodificador después de volver a abrir el mismo, 344. Si todas las tramas prerrodadas han sido decodificadas, el procedimiento salta al bloque 332, en cuyo caso se decodifica la IPF. Las muestras PCM resultantes se  
20 suministran al búfer PCM 342. El desvanecimiento cruzado 318 se realiza en base a las salidas de los búfers PCM 316 y 324 y la salida del procedimiento de desvanecimiento cruzado 318 se suministra al búfer PCM de salida 308.

**[0069]** En la realización descrita anteriormente, la reinicialización del decodificador incluye cerrar la  
25 instancia del decodificador actual y abrir una nueva instancia del decodificador. En realizaciones alternativas, el decodificador puede incluir una pluralidad de instancias del decodificador en paralelo, de manera que la reinicialización del decodificador puede incluir la conmutación entre diferentes instancias del decodificador. Además, la reinicialización del decodificador incluye llenar estados de decodificador decodificando tramas prerrodadas incluidas en la información adicional de la trama especial.

30 **[0070]** Como se ha explicado anteriormente, aprovechando los estados de memoria interna y los búferes (suma de superposición, estados de filtro) en un decodificador AAC, es posible obtener muestras de salida sin pasar la nueva entrada mediante el procedimiento de vaciamiento. La señal de salida del vaciamiento se parece mucho a la "señal original" para, al menos una parte, de los valores de muestra de salida obtenidos, en  
35 particular la primera parte, véase el estado 202 en la Fig. 8. Los valores obtenidos de la muestra de salida obtenidos con el procedimiento de vaciamiento se utilizan en el procedimiento de desvanecimiento cruzado descrito en detalle a continuación.

**[0071]** Como puede verse en el estado 202 en la Fig. 8, la energía en el búfer de vaciamiento resultante  
40 disminuirá con el tiempo en función de la ventana de transformación y de las herramientas habilitadas de la configuración del códec actual. Por lo tanto, el desvanecimiento cruzado debe aplicarse en la primera parte del búfer de vaciamiento, donde la señal de salida puede considerarse como energía casi completa. Aprovechar el hecho de que los códecs de audio modernos se pueden vaciar para obtener muestras válidas para un desvanecimiento cruzado sucesivo ayuda significativamente a obtener valores de conmutación sin fisuras. De  
45 acuerdo con esto, en realizaciones de la invención, el desvanecimiento cruzado está configurado para realizar el desvanecimiento cruzado entre los valores de salida obtenidos con un procedimiento de vaciamiento de la configuración de códec actual y los valores de muestra de salida obtenidos decodificando la trama especial utilizando la configuración de códec indicada en la información adicional.

50 **[0072]** A continuación, se describe una realización específica del procedimiento de desvanecimiento cruzado. El desvanecimiento cruzado se aplica a las señales de audio tal como se ha descrito anteriormente con el fin de evitar artefactos audibles durante la conmutación de CAR. Un artefacto típico es una caída en la energía de señal de salida. Como se ha explicado anteriormente, la energía de la señal vaciada disminuirá en función de la configuración. Por lo tanto, la longitud del desvanecimiento cruzado debe elegirse con cuidado en  
55 función de la configuración para evitar artefactos. Si la ventana de desvanecimiento cruzado es demasiado corta, entonces el procedimiento de conmutación puede introducir artefactos audibles debido a la diferencia en la forma de onda de audio. Si la ventana de desvanecimiento cruzado es demasiado larga, las muestras de audio vaciadas ya han perdido energía y provocarán una caída en la energía de la señal de salida. En una configuración de códec AAC que utiliza ventanas de transformación cortas de 256 muestras, se puede aplicar

un desvanecimiento cruzado lineal con una longitud de  $n = 128$  muestras (por canal). En otras realizaciones, se puede aplicar un desvanecimiento cruzado lineal con una longitud de, por ejemplo, 64 muestras (por canal).

**[0073]** A continuación, se describe un ejemplo de un procedimiento de desvanecimiento cruzado lineal que utiliza 128 muestras:

El procedimiento de desvanecimiento cruzado puede utilizar las primeras 128 muestras del búfer de vaciamiento. El búfer de vaciamiento se segmenta en ventanas multiplicando las primeras 128 muestras del búfer de vaciamiento  $S_f = S_{f0}, \dots, S_{f127}$  por  $d_1 - \frac{i+1}{128}$ ,  $i$  es el índice de la muestra actual.

10

El resultado puede almacenarse en un búfer interno del desvanecimiento cruzado, es decir  $S_{f,i} = S_{f0} \cdot \left(1 - \frac{1}{128}\right), \dots, S_{f127} \cdot \left(1 - \frac{128}{128}\right)$ .

Además, el búfer IPF  $S_d$  se segmenta en ventanas, donde las primeras 128 muestras de salida IPF decodificadas se multiplican por el factor  $1 - \frac{i+1}{128}$ ,  $i$  es el índice de la muestra actual. El resultado puede

almacenarse en un búfer interno del desvanecimiento cruzado, es decir  $S_{d,i} = S_{d0} \cdot \frac{1}{128}, \dots, S_{d127} \cdot 1, \dots, S_{dn}$ .

**[0074]** Se añaden las primeras 128 muestras de los búferes internos:  $S_0 = S_{d0} + S_{f0}, \dots, S_{d127} + S_{f127}$  y los valores resultantes se envían al búfer de muestras de salida PCM 308.

**[0075]** De este modo, se consigue el desvanecimiento cruzado lineal sobre los primeros 128 valores de muestra de salida del búfer de vaciamiento y los primeros 128 valores de muestra del búfer IPF.

25

**[0076]** En general, el desvanecedor cruzado puede ser configurado para ejecutar un desvanecimiento cruzado entre una pluralidad de valores de muestra de salida obtenidos utilizando la configuración de códec actual y una pluralidad de valores de muestra de salida obtenidos mediante la decodificación de los valores de muestra de audio codificados asociados con la trama especial. En general, en códecs de audio, como los códecs de la familia AAC y los códecs de la familia AMR, los valores de muestra de audio codificados de una trama precedente comprenden implícitamente información sobre la señal de audio codificada en una trama siguiente. Esta propiedad se puede utilizar en la implementación de desvanecimiento cruzado al conmutar entre diferentes configuraciones de códecs. Por ejemplo, si la configuración de códec actual es una configuración de códec AMR, los valores de muestra de salida utilizados en el desvanecimiento cruzado pueden obtenerse basándose en una respuesta a un impulso de cero, es decir, basándose en la respuesta obtenida cuando se aplica una trama cero al núcleo del decodificador después de la última trama de la configuración del códec actual. En las realizaciones de la invención, se pueden utilizar mecanismos adicionales utilizados en la codificación y decodificación de audio en el desvanecimiento cruzado. Por ejemplo, los filtros internos utilizados en SBR (Spectral Band Replication) comprenden retardos y, por lo tanto, tiempos de establecimiento largos que pueden utilizarse en el desvanecimiento cruzado. Por lo tanto, las realizaciones de la invención no se limitan a ningún desvanecimiento cruzado específico con el fin de lograr una conmutación sin fisuras entre configuraciones de códecs. Por ejemplo, el desvanecimiento cruzado puede configurarse para aplicar pesos crecientes a un primer número de valores de muestra de salida de la trama especial y para aplicar pesos decrecientes a un número de valores de muestra de salida obtenidos basándose en la decodificación que utiliza la configuración de códec actual, en el que los pesos pueden aumentar y disminuir linealmente o pueden aumentar y disminuir de manera no lineal.

**[0077]** En las realizaciones de la invención, la inicialización del decodificador comprende inicializar estados de decodificador internos y búfers que utilizan la información adicional de la(s) trama(s) especial(es).

En las realizaciones de la invención, la inicialización del decodificador tiene lugar si cambia la configuración del códec. En otras realizaciones de la invención, la trama especial puede utilizarse para inicializar el decodificador sin cambiar la configuración del códec. Por ejemplo, en las realizaciones de la invención, el decodificador puede configurarse para la reproducción inmediata, en el que los estados internos y los búfers de un decodificador se llenan sin cambiar una configuración de códec, en el que se puede realizar el desvanecimiento cruzado con muestras cero. De este modo, la reproducción inmediata de muestras válidas es posible. En otras realizaciones, puede implementarse una función de avance rápido, en la que la trama especial puede ser decodificada en intervalos predeterminados en función de la velocidad de avance rápido

55

deseada. En las realizaciones de la invención, se puede tomar la decisión de si se llevará a cabo la inicialización utilizando la trama especial, es decir, es necesario o deseado, se puede tomar en base a una señal de control externa proporcionada al decodificador de audio.

5 **[0078]** Como se ha explicado anteriormente, la trama especial (como, por ejemplo, la IPF 80 que se muestra en la Fig. 2) se puede utilizar para la adaptación de la velocidad de transferencia de bits y la conmutación de la corriente de bits, respectivamente. Las siguientes restricciones pueden aplicarse: todas las representaciones (por ejemplo, diferentes velocidades de transferencia de bits, diferentes usos de las herramientas de codificación) están alineadas en el tiempo, las IPFs se insertan en cada representación, las  
 10 IPFs están sincronizadas y el campo IPF "config" de la Fig. 2 contiene la configuración de la corriente, es decir, la activación de herramientas, etc. La Fig. 9 muestra un ejemplo de adopción de la velocidad de transferencia de bits por conmutación de la corriente de bits en un entorno de transmisión adaptativo. La lógica de control (como el sistema que se muestra en la Fig. 1), que a veces se denomina entramado, divide a los datos de audio en segmentos. Un segmento comprende múltiples AU. La configuración de la corriente de audio puede  
 15 cambiar en cada límite de segmento. El decodificador de audio no es consciente de la segmentación, solo está provisto de UA simples mediante la lógica de control. Para habilitar la conmutación de corriente de bits de audio en cada límite de segmento, la primera AU de cada segmento puede ser una IPF como se ha explicado anteriormente. En la Fig. 9, un límite de segmento 400 está indicado por la línea discontinua. En el escenario que se ilustra en la Fig. 9, el decodificador de audio está provisto de AU 40 (AU1 a AU3) de "Corriente 1". La  
 20 lógica de control decide cambiar a "Corriente 2" en el siguiente límite de segmento, es decir, límite 400. Después de decodificar la AU3 de la "Corriente 1", la lógica de control puede pasar la AU4 de la "Corriente 2" al decodificador de audio sin más aviso. La AU4 es una trama especial (IPF) y, por lo tanto, la reproducción inmediata puede tener lugar después de cambiar a la corriente2.

25 **[0079]** Haciendo referencia al escenario que se muestra en la Fig. 9, la conmutación puede tener lugar de la siguiente manera: En AU1 a AU3 de la corriente1, no se detecta IPF y el procedimiento de decodificación se lleva a cabo de la forma habitual. Se detecta una IPF en la AU4 de la corriente2. Además, se detecta un cambio en la configuración de la corriente. El decodificador de audio inicializa el procedimiento de vaciamiento,  
 30 402 en la Fig. 9. Las muestras de salida PCM resultantes se almacenan en un búfer temporal (búfer de vaciamiento) para un uso posterior. El decodificador de audio se reinicializa con la configuración de corriente transportada con la IPF. La carga de IPF ("prerrodada") se decodifica. Se descartan las muestras PCM de salida resultantes. En este punto, los estados internos del decodificador y los búferes se inicializan completamente. AU4 se decodifica. Para evitar la conmutación de artefactos se aplica un desvanecimiento  
 35 PCM resultantes de la decodificación de AU4 y almacenadas en el búfer de vaciamiento se atenúan mientras que, las muestras PCM resultantes de la decodificación de AU4 y almacenadas en el búfer de salida PCM, se resaltan. Se reproduce el resultado del desvanecimiento cruzado.

**[0080]** Por consiguiente, la IPF se puede utilizar para habilitar la conmutación de representaciones de audio comprimidas. El decodificador puede recibir UA simples como entrada, por lo que no se necesita ninguna  
 40 lógica de control adicional.

**[0081]** Los detalles de una realización específica en el contexto de MPEG-D USAC se describen ahora, en el que la sintaxis de la corriente de bits puede ser de la siguiente manera:

45 El elemento de sintaxis AudioPreRoll () se utiliza para transmitir información de audio de tramas anteriores junto con los datos de la trama actual. Los datos de audio adicionales se pueden utilizar para compensar el retardo de la puesta en marcha del decodificador (prerrodado), facilitando de este modo el acceso aleatorio en los puntos de acceso a la corriente que hacen uso de AudioPreRoll (). Un UsacExtElement () se puede utilizar para transmitir el AudioPreRoll (). Para ello se utilizará un nuevo identificador de carga:  
 50

Tabla 1: Identificador de carga para AudioPreRoll ()

Nombre	Valor
ID_EXT_ELE_AUDIOPREROLL	4

55

**[0082]** La sintaxis de AudioPreRoll () se muestra en la Fig. 10 y se explica a continuación:

**configLen** tamaño del elemento de sintaxis de configuración en bytes.  
**Config ()** el elemento de sintaxis de configuración del decodificador. En

el contexto de MPEG-D USAC esta es el Usac-Config() tal como se define en ISO/IEC 23003-3: 2012. El campo Config() se puede transmitir para poder responder a cambios en la configuración de audio (conmutación de corrientes).

**numPreRollFrames** el número de unidades de acceso prerrodadas (AU) transmitidas como datos de audio prerrodados. El número razonable de AU depende del retardo de la puesta en marcha del decodificador.

**AuLen** longitud de AU en bytes.

**AccessUnit ()** la(s) AU(s) prerrodada(s).

**[0083]** Los datos prerrodados transportados en el elemento de extensión se pueden transmitir "fuera de banda", es decir, los requisitos de búfer pueden no ser satisfechos

5 **[0084]** Para usar *AudioPreRoll ()* tanto en el acceso aleatorio como en la adaptación de la velocidad de transferencia de bits, se aplican las siguientes restricciones:

- El primer elemento de cada trama es un elemento de extensión (UsacExtElement) de tipo ID\_EXT\_ELE\_AUDIOPREROLL.

10 - El UsacExtElement () correspondiente será configurado como se describe en la Tabla 2.

- Por consiguiente, si hay datos prerrodados, el UsacFrame() comenzará con la secuencia de bits siguiente:

"1": usacIndependencyFlag.

"1": usacExtElementPresent (en referencia al elemento de extensión de audio prerrodado).

15 "0": usacExtElementUseDefaultLength (en referencia al elemento de extensión de audio prerrodado).

- Si no se transmiten datos prerrodados, la carga de extensión no estará presente (usacExtElementPresent = 0).

20 - Las tramas prerrodadas con índice "0" y "numPreRollFrames-1" deberán ser decodificables independientemente, es decir, *usacIndependencyFlag* se establecerá en "1".

Tabla 2: Configuración de UsacExtElement () para AudioPreRoll ()

usacExtElementType	ID_EXT_ELE_AUDIOPREROLL
usacExtElementConfigLength	0
usacExtElementDefaultLengthPresent	0
usacExtElementPayloadFrag	0

25 **[0085]** El acceso aleatorio y la reproducción inmediata son posibles en cada trama que utiliza la estructura de *AudioPreRoll ()* descrita. El siguiente pseudocódigo describe el procedimiento de decodificación:

```

if(usacIndependencyFlag == 1) {
if(usacExtElementPresent == 1){
30 /* In this case usacExtElementUseDefaultLength must be
0! */
if(usacExtElementUseDefaultLength != 0) goto error;
/* Check for presence of config and reinitialize
if necessary */
35 int configLen = getConfigLen();
if(configLen > 0) {
config c =
getConfig(configLen);
ReConfigureDecoder(c);
40 }
/* Get pre-roll AUs and decode, discard output samples
*/
int numPreRollFrames = getNumPreRollFrames();
for(auldx = 0; auldx < numPreRollFrames; auldx++)
45 int auLen = getAuLen();

```

```

AU nextAU =
getPreRollAU(auLen);
DecodeAU(nextAU);
}
5 }
}
/* Decoder states are initialized at this point. Continue
normal decoding */

```

10 **[0086]** La adaptación de la velocidad de transferencia de bits puede utilizarse conmutando entre diferentes representaciones codificadas del mismo contenido de audio. La estructura de *AudioPreRoll ()* descrita puede ser usada para ese propósito. El procedimiento de decodificación en caso de adaptación de la velocidad de transferencia de bits se describe mediante el siguiente pseudocódigo:

```

15 if (usacIndependencyFlag == 1) {
if(usacExtElementPresent == 1{
/* In this case usacExtElementUseDefaultLength must be
0! */
if (usacExtElementUseDefaultLength != 0) goto error;
20 int configLen = getConfigLen();
if(configLen > 0) {
config newConfig = getConfig(configLen);
/* Configuration did not change, skip
AudioPreRoll and continue decoding as normal */
25 if(newConfig ==
currentConfig) {
SkipAudioPreRoll ();
goto finish;
}
30 /* Configuration changed, prepare for
bitstream switching*/
config c =
getConfig(configLen);
outSamplesFlush =
35 FlushDecoder ();
ReConfigureDecoder(c);
/* Get pre-roll AUs and decode, discard output samples
*/
int numPreRollFrames = getNumPreRollFrames ();
40 for(auldx = 0; auldx < numPreRollFrames; auldx++)
int auLen = getAuLen ();
AU nextAU =
getPreRollAU(auLen);
DecodeAU(nextAU);
45 }
/* Get "regular" AU and decode */
AU au = UsacFrame ();
outSamplesFrame = Decode(au);
/* Apply crossfade */
50 for(i = 0; i < 128; i++) {
outSamples[i] = outSamplesFlush[i] * (1-i/127) +
outSamplesFrame[i] * (i/127)
}
for(i = 128; i < outputFrameLength; i++) {
55 outSamples [i] = outSamplesFrame [i] ;
}
} else {
goto error;
}
}

```

}  
}

**[0087]** Aunque se han descrito algunos aspectos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del procedimiento correspondiente, donde un bloque o dispositivo se corresponde con una etapa del procedimiento o una característica de una etapa del procedimiento. De forma análoga, los aspectos que se describen en el contexto de una etapa del procedimiento también representan una descripción de un bloque correspondiente o un punto o característica del aparato correspondiente. Algunas o todas las etapas del procedimiento se pueden ejecutar con (o utilizando) un aparato de hardware, como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, una o más de las etapas más importantes del procedimiento se pueden ejecutar con dicho aparato. En las realizaciones de la invención, los procedimientos descritos en este documento se implementan con un procesador o se implementan con un ordenador.

**[0088]** En función de ciertos requisitos de implementación, las realizaciones de la invención se pueden implementar en hardware o en software. La implementación puede realizarse utilizando un medio de almacenamiento no transitorio como un medio de almacenamiento digital, por ejemplo, un disquete, un DVD, un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control legibles electrónicamente y almacenadas en el mismo, que coopera (o es capaz de cooperar) con un sistema informático programable de manera que se lleve a cabo el procedimiento respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible mediante ordenador.

**[0089]** Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control legibles electrónicamente y que son capaces de cooperar con un sistema informático programable, de tal manera que se lleve a cabo uno de los procedimientos descritos en este documento.

**[0090]** En general las realizaciones de la presente invención se pueden implementar como un producto de programa informático con un código de programa, siendo el código de programa operativo para llevar a cabo uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa se puede almacenar, por ejemplo, en un soporte legible por máquina.

**[0091]** Otras realizaciones comprenden el programa de ordenador para llevar a cabo uno de los procedimientos descritos en este documento, almacenado en un soporte legible por máquina.

**[0092]** En otras palabras, una realización del procedimiento de la invención es, por lo tanto, un programa informático que tiene un código de programa para realizar uno de los procedimientos descritos en este documento cuando el programa informático se ejecuta en un ordenador.

**[0093]** Una realización adicional del procedimiento de la invención es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para llevar a cabo uno de los procedimientos descritos en este documento. El soporte de datos, el medio de almacenamiento digital o el medio grabado normalmente son tangibles y/o no transitorios.

**[0094]** Una realización adicional del procedimiento de la invención es, por lo tanto, un corriente de datos o una secuencia de señales que representan el programa informático para llevar a cabo uno de los procedimientos descritos en este documento. El corriente de datos o la secuencia de señales pueden, por ejemplo, estar configurados para ser transferidos a través de una conexión de comunicación de datos, por ejemplo, a través de Internet.

**[0095]** Una realización adicional comprende además un medio de procesamiento, por ejemplo, un ordenador o un dispositivo lógico programable, programado, configurado o adaptado para llevar a cabo uno de los procedimientos descritos en este documento.

**[0096]** Una realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para llevar a cabo uno de los procedimientos descritos en este documento.

**[0097]** Una realización adicional según la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa informático con el fin de llevar a

cabo uno de los procedimientos descritos en el presente documento a un receptor. El receptor puede ser, por ejemplo, un ordenador, un dispositivo móvil, un dispositivo de memoria o algo similar. El aparato o sistema puede, por ejemplo, comprender un servidor de archivos para transferir el programa informático al receptor.

- 5 **[0098]** En algunas realizaciones se puede utilizar un dispositivo lógico programable (por ejemplo, una matriz de puertas programables por campo) para llevar a cabo algunas o todas las funcionalidades de los procedimientos descritos en este documento. En algunas realizaciones un campo de matriz de puertas programables podrá cooperar con un microprocesador para llevar a cabo uno de los procedimientos descritos en este documento. En general los procedimientos se llevan a cabo, preferentemente, por cualquier aparato de
- 10 hardware.

## REIVINDICACIONES

1. Un decodificador de audio (60) para decodificar una corriente de bits de datos de audio codificados, en el que la corriente de bits de datos de audio codificado representa una secuencia de valores de muestra de audio y comprende una pluralidad de tramas (40), en el que cada trama (40) incluye valores de muestra de audio codificados asociados, el decodificador de audio (60) que comprende:
- un determinador (130) configurado para determinar si una trama de los datos de audio codificados es una trama especial (42, 80) que comprende valores de muestra de audio codificados asociados con la trama especial (42, 80) e información adicional (82), en la que la información adicional (82) comprende valores de muestra de audio codificados de un número de tramas (86) que preceden a la trama especial, en el que los valores de la muestra de audio codificados de las tramas precedentes están codificados utilizando la misma configuración de códec que la trama especial, en el que el número de tramas precedentes, correspondiente a las tramas prerrodadas, se corresponde con el número de tramas que el decodificador (60) necesita para construir la señal completa durante la puesta en marcha del decodificador de modo que esté en posición de decodificar los valores de muestra de audio asociados con la trama especial (42, 80) si la trama especial es la primera trama después de la puesta en marcha del decodificador; y
- un inicializador configurado para inicializar el decodificador (60) si el determinador determina que la trama es una trama especial, en el que inicializar el decodificador comprende decodificar los valores de muestra de audio codificados incluidos en la información adicional antes de decodificar los valores de muestra de audio codificados asociados con la trama especial (42, 80),
- en el que el inicializador está configurado para conmutar el decodificador de audio (60) desde una configuración de códec actual a una configuración de códec (84) diferente si el determinador (130) determina que la trama es una trama especial (42, 80) y si los valores de muestra de audio de la trama especial se han codificado utilizando la configuración de códec diferente, y
- en el que el decodificador está configurado para decodificar la trama especial (42, 80) utilizando la configuración de códec actual y descartar la información adicional si el determinador determina (130) que la trama es una trama especial (42, 80) y si los valores de muestra de audio de la trama especial se han codificado utilizando la configuración de códec actual.
2. Decodificador de audio de la reivindicación 1, en el que la información adicional comprende información sobre la configuración de códec (84) utilizada para codificar los valores de muestra de audio asociados con la trama especial (42, 80), en el que el determinador está configurado para determinar si la configuración de códec de la información adicional es diferente de la configuración de códec actual.
3. Decodificador de audio (60) según una de las reivindicaciones 1 y 2, que comprende un desvanecedor cruzado (318) configurado para ejecutar un desvanecimiento cruzado entre una pluralidad de valores de muestra de salida obtenidos utilizando la configuración de códec actual y una pluralidad de valores de muestra de salida obtenidos mediante la decodificación de los valores de muestra de audio codificados asociados con la trama especial (42, 80).
4. Decodificador de audio según la reivindicación 3, en el que el desvanecimiento cruzado (318) está configurado para ejecutar el desvanecimiento cruzado de los valores de muestra de salida obtenidos mediante el vaciamiento del decodificador (60) en la configuración de códec actual y valores de muestra de salida obtenidos mediante la decodificación de los valores de muestra de audio codificados asociados con la trama especial (42, 80).
5. Decodificador de audio de una de las reivindicaciones 1 a 4, en el que una primera trama del número de tramas (86) comprendidas en la información adicional (82) no está codificada de manera diferencial en el tiempo o codificada por entropía en relación con ninguna trama anterior a la trama más temprana y en la que la trama especial (42, 80) no está codificada de manera diferencial en el tiempo o codificada por entropía en relación con cualquier trama anterior a la primera trama del número de tramas que preceden a la trama especial (42, 80) o en relación con cualquier trama anterior a la trama especial (42, 80).
6. Decodificador de audio de una de las reivindicaciones 1 a 5, en el que la trama especial (42, 80) comprende la información adicional como una carga de extensión y en el que el determinador está configurado

para evaluar la carga de extensión de la trama especial (42, 80).

7. Aparato (100; 12, 14, 16, 18) para generar una corriente de bits de datos de audio codificados que representan una secuencia de valores de muestra de audio de una señal de audio (10), en el que la corriente de bits de datos de audio codificados comprende una pluralidad de tramas, en el que cada trama incluye valores de muestra de audio codificados asociados, en el que el aparato (100; 12, 14, 16, 18) comprende:

un proveedor de trama especial configurado para proporcionar al menos una de las tramas de una trama especial (42, 80), la trama especial (42, 80) que comprende valores de muestra de audio codificados asociados con la trama especial (42, 80) e información adicional (82), en el que la información adicional (82) comprende valores de muestra de audio codificados de un número de tramas (86) que preceden a la trama especial, en el que los valores de la muestra de audio codificados de las tramas precedentes están codificados utilizando la misma configuración de códec que la trama especial, y en el que el número de tramas precedentes, correspondiente a las tramas prerrodadas, se corresponde con el número de tramas que el decodificador (60) necesita para construir la señal completa durante la puesta en marcha del decodificador de modo que esté en posición de decodificar los valores de muestra de audio asociados con la trama especial (42, 80) si la trama especial es la primera trama después de la puesta en marcha del decodificador; y una salida (112) configurada para emitir la corriente de bits de datos de audio codificados (54, 102),

en el que los datos de audio codificados comprenden una pluralidad de segmentos (30), en el que cada segmento está asociado con una de una pluralidad de partes de la secuencia de valores de muestra de audio y comprende una pluralidad de tramas (40), en el que el sumador de tramas especial está configurado para añadir una trama especial (42, 80) en el comienzo de cada segmento (30) independientemente de si la configuración del códec cambia o no.

8. Aparato (100; 12, 14, 16, 18) según la reivindicación 7, en el que la información adicional comprende información sobre la configuración del códec (84) utilizada para codificar los valores de muestra de audio asociados con la trama especial (42, 80).

9. Aparato (100) según una de las reivindicaciones 7 u 8, el aparato (100) que comprende:

un proveedor de segmentos (104) configurado para proporcionar segmentos (44, 46, 48) asociados con diferentes partes de la secuencia de valores de muestra de audio y codificado por diferentes configuraciones de códec, en el que el proveedor de trama especial está configurado para proporcionar una primera trama (42, 80) de, al menos uno, de los segmentos como la trama especial (42, 80); y

un generador (52, 110) configurado para generar los datos de salida de audio disponiendo, al menos uno, de los segmentos (44, 46, 48) después de otro de los segmentos (44, 46, 48).

10. Aparato de la reivindicación 11, en el que el proveedor de segmentos (100) está configurado para seleccionar una configuración de códec para cada segmento basándose en una señal de control.

11. Aparato de la reivindicación 9 o 10, en el que el proveedor de segmentos (100) está configurado para proporcionar m versiones codificadas (22, 24, 26, 28) de la secuencia de valores de muestra de audio, con  $m \geq 2$ , en el que las m versiones codificadas se codifican utilizando diferentes configuraciones de códec, en el que cada versión codificada comprende una pluralidad de segmentos (30) que representa la pluralidad de partes de la secuencia de valores de muestra de audio, en el que el proveedor de tramas especiales está configurado para proporcionar una trama especial (42, 80) en el comienzo de cada uno de los segmentos.

12. Aparato según la reivindicación 11, en el que el proveedor de segmentos (100) comprende una pluralidad de codificadores (12, 14, 16, 18), cada uno configurado para codificar, al menos en parte, la señal de audio según una de la pluralidad de configuraciones de códec diferentes.

13. Aparato de la reivindicación 12, en el que el proveedor de segmentos comprende una memoria que almacena las m versiones codificadas de la secuencia de valores de muestra de audio.

14. Aparato según una de las reivindicaciones 9 a 13, en el que el proveedor de trama especial (100) está configurado para proporcionar la información adicional como una carga de extensión de la trama especial

(42, 80).

15. Procedimiento para decodificar una corriente de bits de datos de audio codificados, en el que la corriente de bits de datos de audio codificados representa una secuencia de valores de muestra de audio y comprende una pluralidad de tramas (40), en el que cada trama (40) incluye valores de muestra de audio codificados asociados, y que comprende:

determinar si una trama de los datos de audio codificados es una trama especial (42, 80) que comprende valores de muestra de audio codificados asociados con la trama especial (42, 80) e información adicional (82), en el que la información adicional (82) comprende valores de muestra de audio codificados de un número de tramas (86) que preceden a la trama especial, en el que los valores de la muestra de audio codificados de las tramas precedentes están codificados utilizando la misma configuración de códec que la trama especial, en el que el número de tramas precedentes, correspondiente a las tramas prerrodadas, se corresponde con el número de tramas que el decodificador (60) necesita para construir la señal completa durante la puesta en marcha del decodificador de modo que esté en posición de decodificar los valores de muestra de audio asociados con la trama especial (42, 80) si la trama especial es la primera trama después de la puesta en marcha del decodificador.

inicializar el decodificador (60) si se determina que la trama es una trama especial, en el que la inicialización comprende decodificar los valores de muestra de audio codificados incluidos en la información adicional antes de decodificar los valores de muestra de audio codificados asociados con la trama especial (42, 80);

conmutar el decodificador de audio (60) de una configuración de códec actual a una configuración de códec diferente (84) si se determina que la trama es una trama especial (42, 80) y si los valores de muestra de audio de la trama especial han sido codificados utilizando la configuración de códec diferente; y

decodificar la trama especial (42, 80) utilizando la configuración de códec actual y descartar la información adicional si se determina que la trama es una trama especial (42, 80) y si los valores de muestra de audio de la trama especial se han codificado utilizando la configuración de códec actual.

16. Procedimiento según la reivindicación 15, en el que la corriente de bits de datos de audio comprende un primer número de tramas codificadas utilizando una primera configuración de códec y un segundo número de tramas después del primer número de tramas y codificadas utilizando una segunda configuración de códec, en el que la primera trama del segundo número de tramas es la trama especial.

17. Procedimiento según una de las reivindicaciones 15 o 16, en el que la información adicional comprende información sobre la configuración de códec (84) utilizada para codificar los valores de muestra de audio asociados con la trama especial (42, 80), el procedimiento que comprende determinar si la configuración de códec de la información adicional es diferente de la configuración de códec actual que utiliza los valores de muestra de audio codificados en la corriente de bits, que preceden a la trama especial, que están codificados.

18. Procedimiento para generar una corriente de bits de datos de audio codificados que representan una secuencia de valores de muestra de audio de una señal de audio (10), en el que la corriente de bits de datos de audio codificados comprende una pluralidad de tramas, en el que cada trama incluye valores de muestra de audio codificados asociados, que comprende:

proporcionar, al menos una, de las tramas como una trama especial (42, 80), la trama especial (42, 80) que comprende valores de muestra de audio codificados asociados con la trama especial (42, 80) e información adicional (82), en la que la información adicional (82) comprende valores de muestra de audio codificados de un número de tramas (86) que preceden a la trama especial, en la que los valores de la muestra de audio codificados de las tramas precedentes están codificados utilizando la misma configuración de códec que la trama especial, en la que el número de tramas precedentes, correspondiente a las tramas prerrodadas, se corresponde con el número de tramas que el decodificador (60) necesita para construir la señal completa durante la puesta en marcha del decodificador de modo que esté en posición de decodificar los valores de muestra de audio asociados con la trama especial (42, 80) si la trama especial es la primera trama después de la puesta en marcha del decodificador; y

generar la corriente de bits concatenando la trama especial (42, 80) y las otras tramas de la pluralidad de tramas,

en el que los datos de audio codificados comprenden una pluralidad de segmentos (30), en el que cada segmento está asociado con una de una pluralidad de partes de la secuencia de valores de muestra de audio y comprende una pluralidad de tramas (40), en el que se añade una trama especial (42, 80) al comienzo de cada  
5 segmento (30) independientemente de si la configuración del códec cambia o no.

19. Procedimiento según la reivindicación 18, en el que la información adicional comprende información sobre la configuración de códec (84) utilizada para codificar los valores de muestra de audio asociados con la trama especial (42, 80).

10

20. Programa informático para realizar, cuando se ejecuta en un ordenador o un procesador, el procedimiento de una de las reivindicaciones 15 a 19.

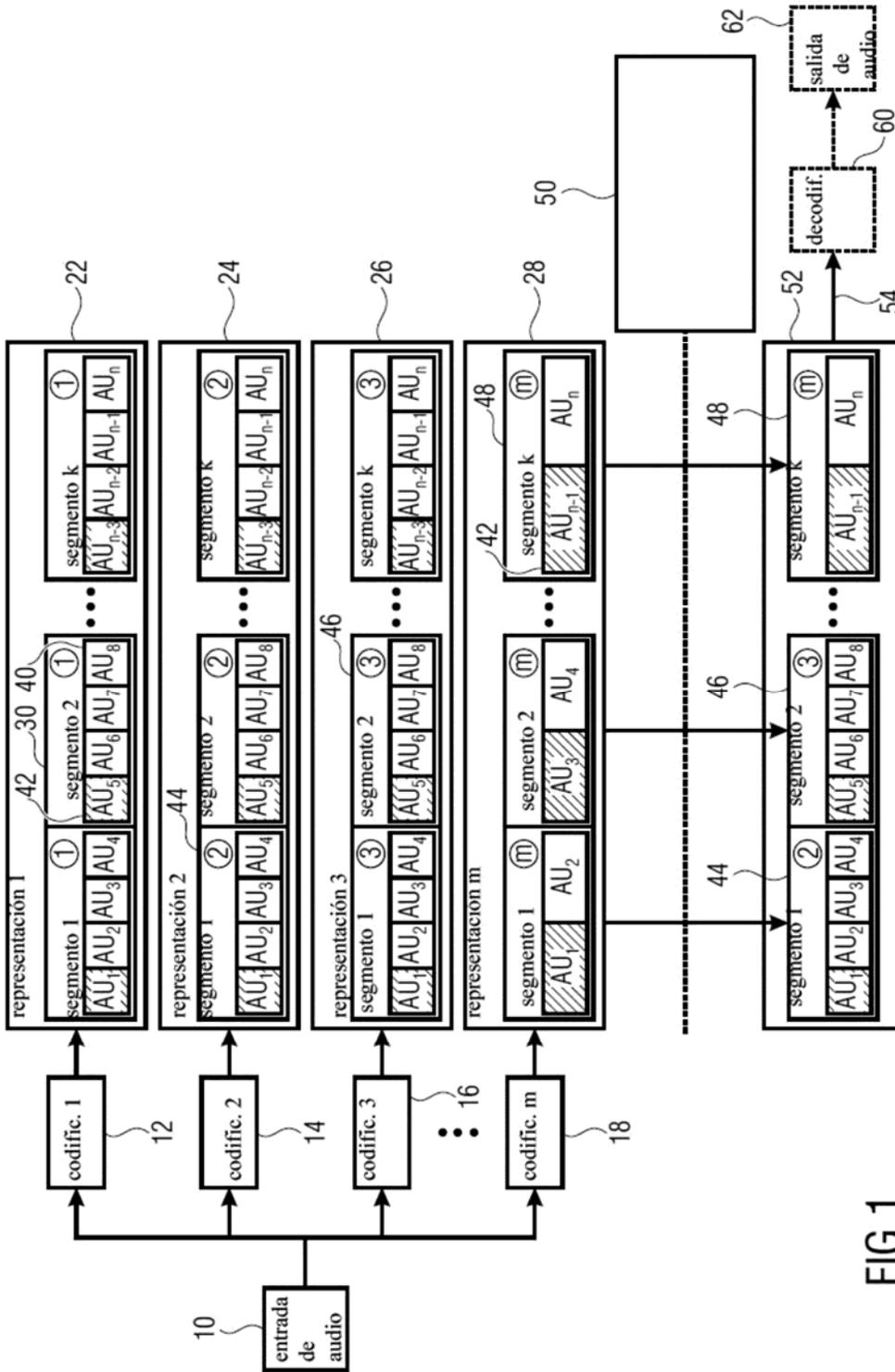


FIG 1

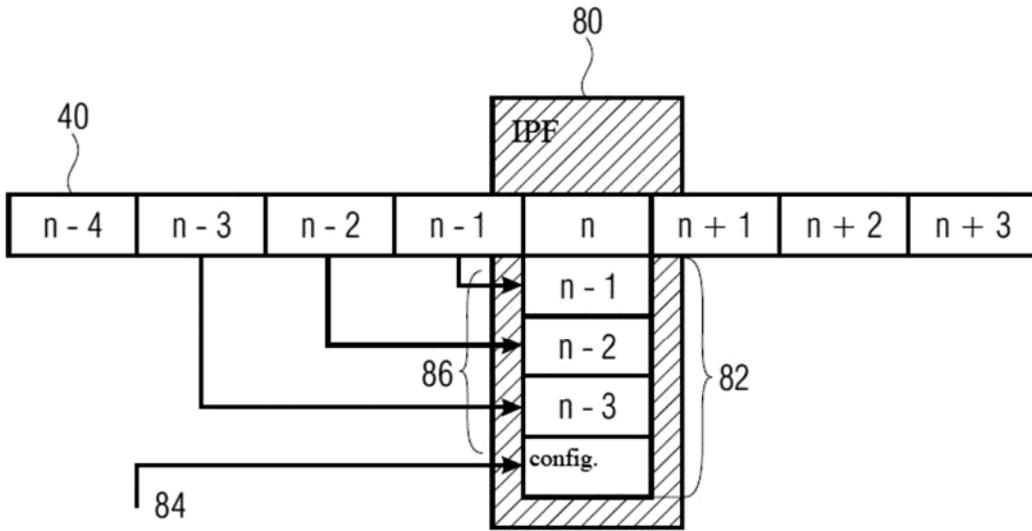


FIG 2

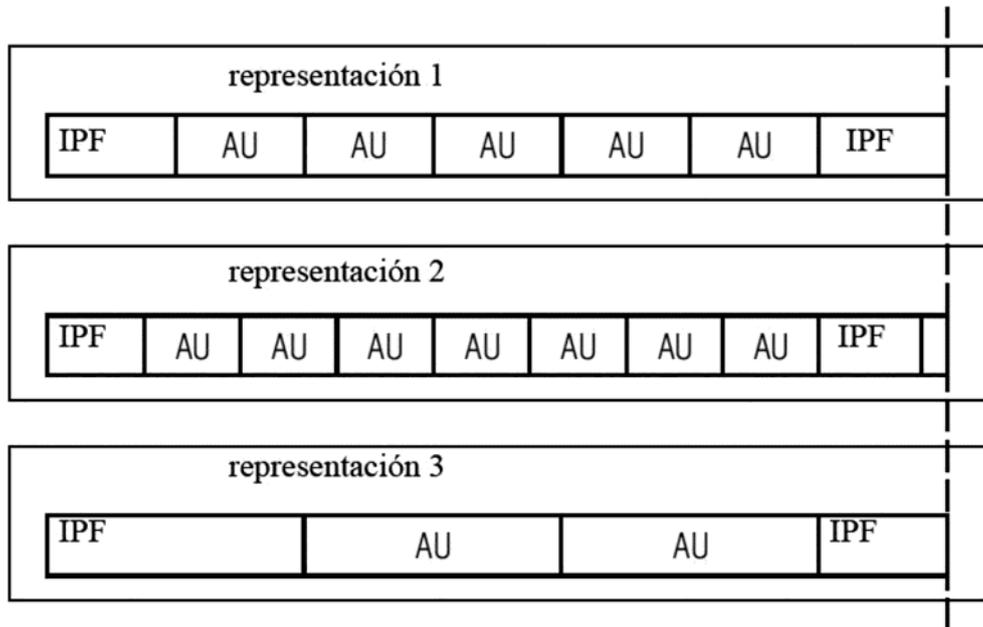


FIG 3

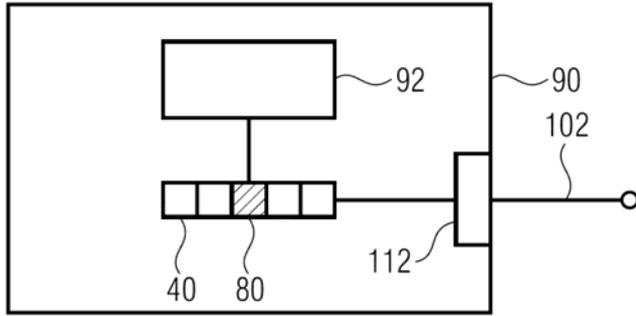


FIG 4A

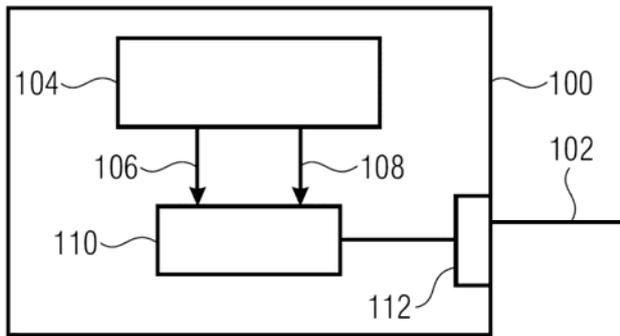


FIG 4B

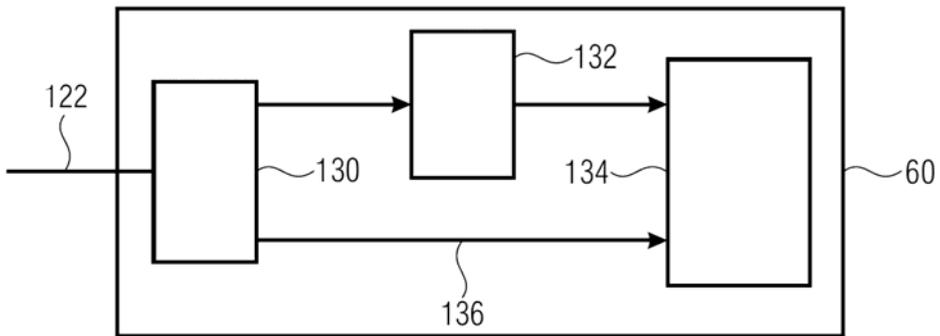


FIG 5

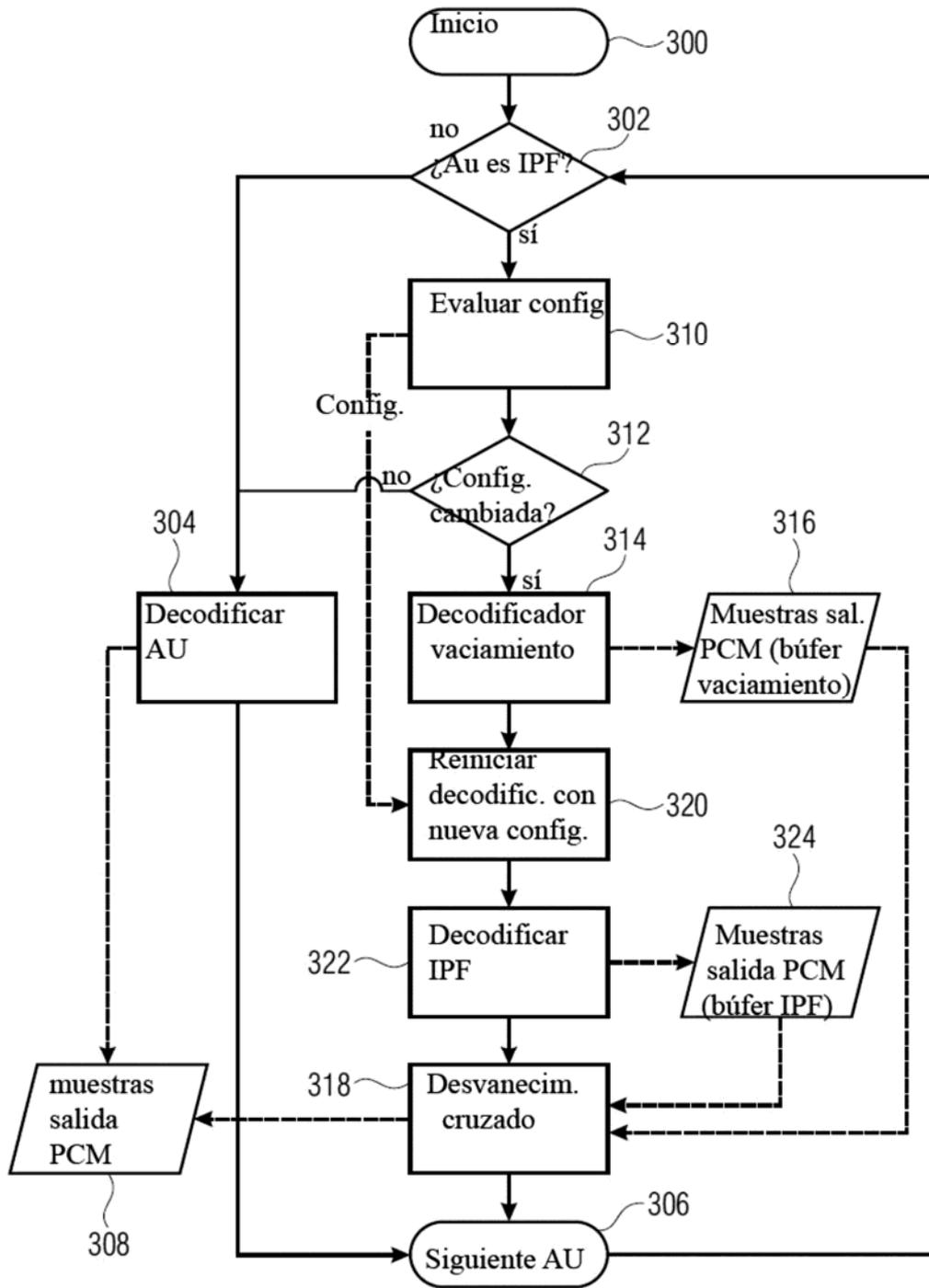


FIG 6

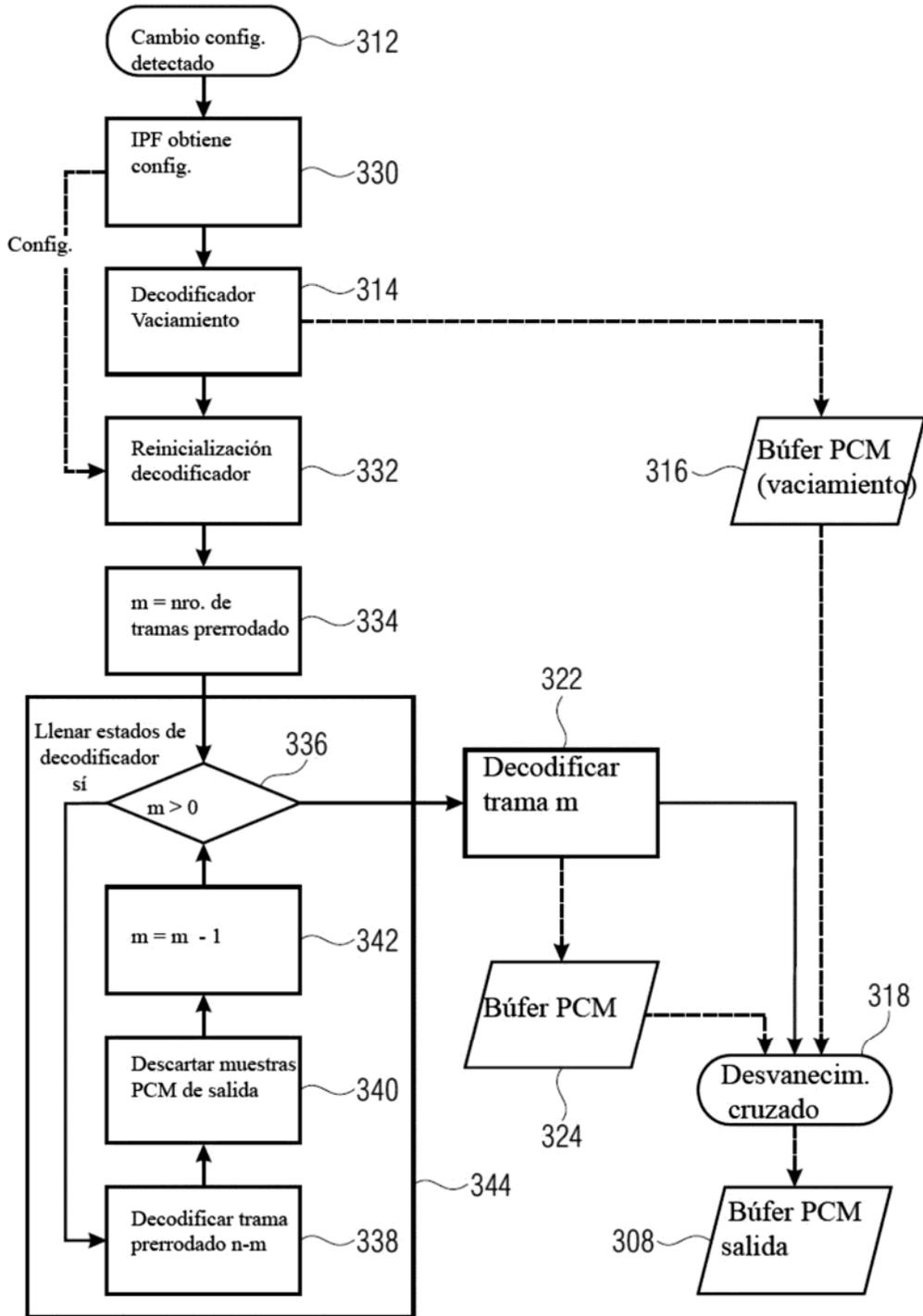


FIG 7

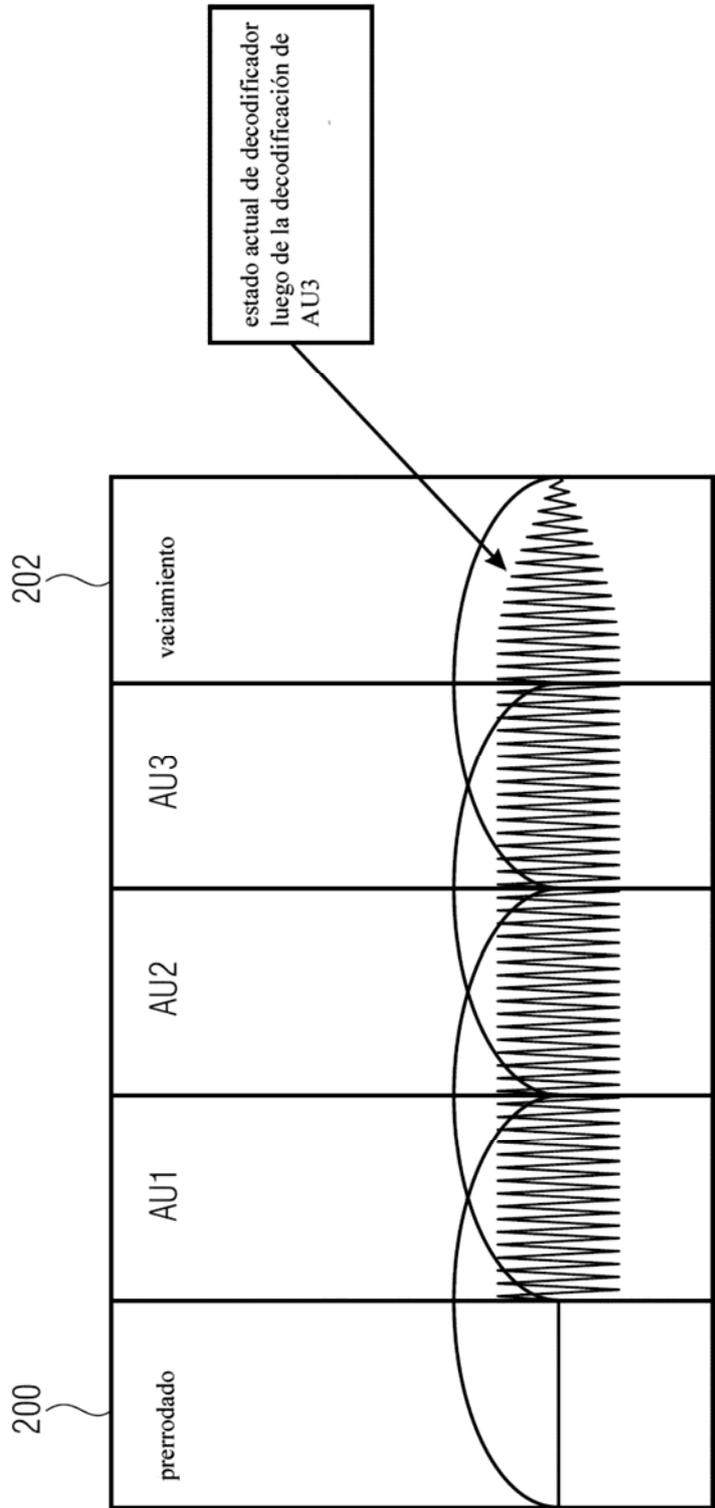


FIG 8

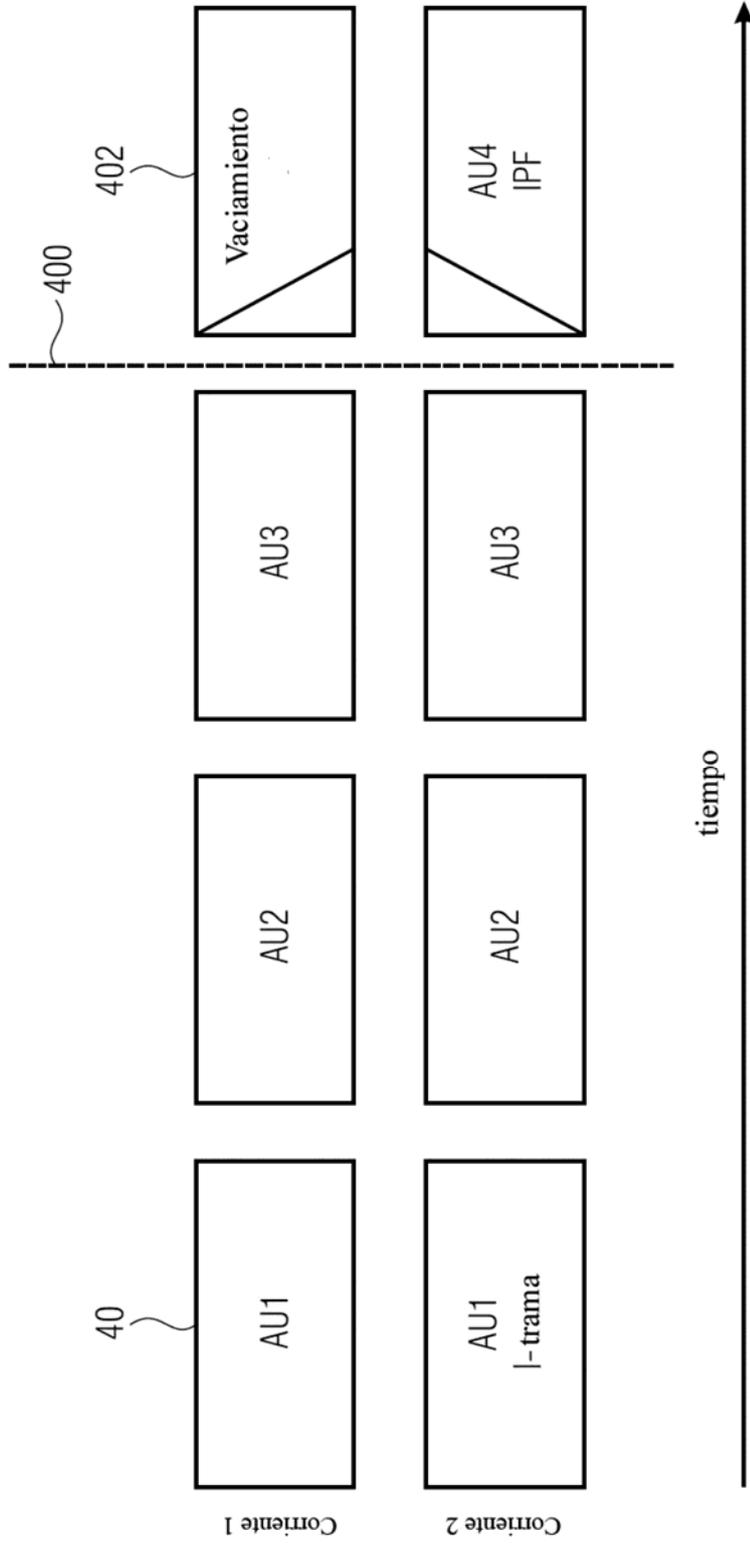


FIG 9

Sintaxis	Nro. de bits	Mnemónico
<pre> AudioPrerod () {   configLon = Valorésape (4,4,8);   Config()   num. TramasPrerod.= Valor escape (2,4,0);   para (trama dx=0; trama dx &lt; numTramasPrerod. ++ trama jx ) {     auLon     UnidadAcceso ()   } } </pre>	<p>4..16 8*configlo n</p> <p>2..6</p> <p>32 8*aulo n</p>	<p>uimsbf</p>

FIG 10