

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 986 126**

51 Int. Cl.:

G10L 19/16

(2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **12.04.2011** **E 19154231 (5)**

97 Fecha y número de publicación de la concesión europea: **03.07.2024** **EP 3499503**

54 Título: **Decodificación de la representación exacta de una señal de audio**

30 Prioridad:

13.04.2010 US 32344010 P

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
08.11.2024

73 Titular/es:

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

72 Inventor/es:

**DÖHLA, STEFAN y
SPERSCHNEIDER, RALPH**

74 Agente/Representante:

PONTI & PARTNERS, S.L.P.

ES 2 986 126 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

Decodificación de la representación exacta de una señal de audio

5 Campo Técnico

[0001] Las formas de realización de la invención se refieren al campo de codificación de fuente de una señal de audio. Más específicamente, las formas de realización de la invención se refieren a un procedimiento para codificar información en los datos de audio válidos originales y un decodificador asociado. Más específicamente, las formas de realización de la invención proporcionan la recuperación de los datos de audio con su duración original.

Antecedentes De La Invención

[0002] Los codificadores de audio se emplean de forma característica para comprimir una señal de audio para transmisión o almacenamiento. Dependiendo del codificador utilizado, la señal puede ser codificada sin pérdidas (permitiendo una reconstrucción perfecta) o con pérdidas (para una reconstrucción imperfecta pero suficiente). El decodificador asociado invierte la operación de codificación y crea la señal de audio perfecta o imperfecta. Cuando la bibliografía menciona artefactos, entonces se pretende de forma característica la pérdida de información, que es típica para la codificación con pérdidas. Estas incluyen un ancho de banda de audio limitado, artefactos de eco y llamada y otra información que puede ser audible o enmascarada debido a las propiedades de la audición humana.

[0003] Otra fuente de distorsión audible durante la reproducción de archivos multimedia comprimidos se menciona en la Publicación de solicitud de patente de Estados Unidos número US 2008/0065393 A1 titulada "Playback of compressed media files without quantization gaps" ("Reproducción de archivos multimedia comprimidos sin espacios de cuantificación"). En particular, el documento US 2008/0065393 A1 menciona la aparición de distorsiones cuando se realiza la llamada "unión a tope". Una "unión a tope" es la conexión abrupta de un archivo de audio a otro archivo de audio para que se conviertan en un archivo de audio continuo. Según el documento US 2008/0065393 A1, el archivo multimedia digital está formado por un número de muestras de audio agrupadas en un número correspondiente de tramas de audio. Cuando se implementa como un procedimiento, se lleva a cabo una identificación de un codificador utilizado. Se obtienen un valor de retraso del codificador para el codificador identificado y un valor de retraso del decodificador para el decodificador. Se determina un recuento de muestras de audio correspondiente a una última muestra de audio válida y se configura una resincronización después de la búsqueda de la marca N de las tramas de audio de la última muestra de audio válida. Solo las muestras de audio válidas se decodifican mediante el uso del valor de retraso del codificador, el valor de retraso del decodificador y el recuento de muestras correspondiente a la última muestra de audio válida.

[0004] Para archivos Mp3, una página de Internet bajo la dirección gabriel.mp3-tech.org/mp3infotag.html describe la llamada etiqueta de información Mp3. Esta etiqueta contiene una sección de 3 bytes titulada "retrasos del codificador" con la siguiente estructura: [xxxxxxx][xxx xxxxyyyy][yyyyyyyy]. Los dos valores de 12 bits X e Y indican cuántas muestras se agregaron al inicio (retraso del codificador) en X y cuántas muestras cero se rellenaron al final en Y para completar la última trama.

Resumen de la Invención

[0005] El problema abordado por esta invención se refiere a otro conjunto de artefactos, que de forma característica no están cubiertos en la bibliografía de codificación de audio: periodos de silencio adicionales al inicio y al fin de una codificación. Existen soluciones para estos artefactos, que a menudo se refieren como procedimientos de reproducción sin interrupciones. Las fuentes para estos artefactos son al principio la granularidad gruesa de los datos de audio codificados donde, por ejemplo, una unidad de datos de audio codificados siempre contiene información para 1024 muestras de audio no codificadas originales. En segundo, el procesamiento de señal digital a menudo solo es posible con retrasos algorítmicos debido a los filtros digitales y bancos de filtros involucrados.

[0006] Muchas aplicaciones no requieren la recuperación de las muestras válidas originalmente. La difusión de radio, por ejemplo, normalmente no es problemática, ya que la transmisión de audio codificada es continua y no sucede una concatenación de codificaciones separadas. Las difusiones de TV también se configuran a menudo estáticamente, y un solo codificador se emplea antes de la transmisión. Los periodos de silencio extra se vuelven sin embargo un problema, cuando varias transmisiones previamente codificadas se unen (como se emplea para inserción de anuncios), cuando una sincronización de audio-vídeo se vuelve una consideración, para el almacenamiento de datos comprimidos, donde la decodificación no deberá exhibir las muestras de audio extra al inicio y al final (en especial para codificación sin interrupciones que requiere una reconstrucción exacta de bits de los datos de audio no comprimidos originales), y para editar en el dominio comprimido.

[0007] Mientras que muchos usuarios ya se han adaptado a estos periodos de silencio extra, otros usuarios se quejan por el silencio extra, lo que es especialmente problemático cuando se concatenan varias codificaciones y

datos de audio sin interrupciones anteriormente no comprimidos se interrumpen cuando son codificados y decodificados. Un objeto de la invención es proporcionar una estrategia mejorada que permite la eliminación de silencio no deseado al inicio y al final de las codificaciones.

- 5 **[0008]** La codificación de vídeo mediante el uso de mecanismos de codificación diferencial, que utiliza tramas I, tramas P y tramas B, no introduce ninguna trama extra al inicio o al final. En contraste, el codificador de audio tiene de forma característica muestras previamente pendientes adicionales. Dependiendo de su número, pueden llevar a una pérdida perceptible de sincronización de audio-vídeo. Esto a menudo es referido como el problema de postsincronización, o doblaje, la desigualdad o desfase entre el movimiento experimentado de la boca de un interlocutor y el sonido que se escucha. Muchas aplicaciones abordan este problema al hacer un ajuste para postsincronización, que debe realizarse por el usuario ya que es altamente variable, dependiendo del códec en uso y sus ajustes. Un objeto de la invención es proporcionar una estrategia mejorada que permite una reproducción sincronizada de audio y vídeo.
- 10
- 15 **[0009]** En el pasado, las difusiones digitales se volvieron más heterogéneas, con diferencias regionales y programas personalizados y anuncios. Una transmisión de difusión principal es reemplazada y unida, por lo tanto, con un contenido local o específico de usuario, que pueden ser datos de transmisión en vivo o datos previamente codificados. La unión de estas transmisiones depende primordialmente del sistema de transmisión; sin embargo, el audio a menudo no puede ser perfectamente unido, como se desea, debido a los periodos de silencio desconocidos.
- 20 Un procedimiento actual, es a menudo dejar los periodos de silencio en la señal, aunque estos espacios en la señal de audio pueden ser percibidos. Un objeto de la invención es proporcionar una estrategia mejorada que permite la unión de dos transmisiones de audio comprimidas.

- [0010]** La edición normalmente se realiza en el dominio no comprimido, donde las operaciones de edición son bien conocidas. Si el material fuente sin embargo ya es una señal de audio codificada con pérdida, entonces incluso operaciones de corte simple requieren una codificación completa nueva, resultando en artefactos de codificación en tándem. Por lo tanto, las operaciones de decodificación y codificación en tándem deberán ser evitadas. Un objeto de la invención es proporcionar una estrategia mejorada que permite el corte de una transmisión de audio comprimida.
- 25

- 30 **[0011]** Un aspecto diferente es el borrado de muestras de audio no válidas en sistemas que requieren una ruta de datos protegida. La ruta multimedia protegida se emplea para hacer valer la administración o gestión de derechos digitales y asegurar la integridad de datos al utilizar comunicación cifrada entre los componentes de un sistema. En estos sistemas, este requisito puede ser cumplido solo si duraciones no constantes de una unidad de datos de audio se vuelven posibles, ya que pueden aplicarse solo en elementos fiables dentro de las operaciones de edición de audio de ruta multimedia protegida. Estos elementos fiables solo son de forma característica los decodificadores y los elementos de renderizado.
- 35

[0012] El alcance de la invención queda definido por las reivindicaciones independientes.

40 Breve Descripción de los Dibujos

[0013] Formas de realización según la invención se describirán posteriormente haciendo referencia a las figuras anexas en las cuales:

- 45 La Figura 1 ilustra un comportamiento de decodificador HE AAC: modo de velocidad dual;
La Figura 2 ilustra un intercambio de información entre una entidad de Capa de sistemas y un decodificador de audio;
La Figura 3 muestra un diagrama de flujo esquemático de un procedimiento para proporcionar información respecto a la validez de datos de audio codificados según una primera forma de realización posible;
- 50 La Figura 4 muestra un diagrama de flujo esquemático de un procedimiento para proporcionar información respecto a la validez de datos de audio codificados según una segunda forma de realización posible de las enseñanzas descritas en esta invención;
La Figura 5 muestra un diagrama de flujo esquemático de un procedimiento para proporcionar información sobre la validez de datos de audio codificados según una tercera forma de realización posible de las enseñanzas descritas en esta invención;
- 55 La Figura 6 muestra un diagrama de flujo esquemático de un procedimiento para recibir datos codificados que incluyen la información respecto a la validez de datos según una forma de realización de las enseñanzas descritas en esta invención;
La Figura 7 muestra un diagrama de flujo esquemático del procedimiento para recibir datos codificados según otra forma de realización de las enseñanzas descritas en esta invención;
- 60 La Figura 8 muestra un diagrama de entrada/salida de un codificador según una forma de realización de las enseñanzas descritas en esta invención;
La Figura 9 muestra un diagrama de entrada/salida esquemático de un codificador según otra forma de realización de las enseñanzas descritas en esta invención;
- 65 La Figura 10 muestra un diagrama de bloques esquemático de un decodificador según una forma de realización de

las enseñanzas descritas en esta invención; y

La Figura 11 muestra un diagrama de bloques esquemático de un decodificador según otra forma de realización de las enseñanzas descritas en esta invención.

5 Descripción detallada de formas de realización ilustrativas

[0014] La Figura 1 muestra el comportamiento de un decodificador con respecto a unidades de acceso (AU = Access Units) y unidades de composición (CU = Composition Units) asociadas. El decodificador se conecta a una entidad denominada "Sistemas" que recibe una salida generada por el decodificador. Como un ejemplo, el decodificador deberá considerarse que funciona bajo la norma de Codificación de Audio Avanzado – Alta Eficiencia (HE-AAC = High Efficiency – Advanced Audio Coding). Un decodificador HE-AAC esencialmente es un decodificador AAC seguido por una etapa de "posprocesamiento" de Reducción de Banda Espectral (SBR = Spectral Band Reduction). El retraso adicional impuesto por la herramienta SBR se debe al banco QMF y a los amortiguadores de datos dentro de la herramienta SBR. Puede derivarse por la siguiente fórmula:

$$\text{Retraso}_{\text{Herramienta de SBR}} = L_{\text{Filtro de análisis}} - N_{\text{Canales de Análisis}} + 1 + \text{Retraso}_{\text{amortiguador}}$$

donde

$N_{\text{Canales de análisis}} = 32$, $L_{\text{Filtro de análisis}} = 320$ y $\text{Retraso}_{\text{amortiguador}} = 6 \times 32$.

[0015] Esto significa que el retraso impuesto por la herramienta SBR (a la velocidad de muestreo de alimentación, es decir, la velocidad de muestreo de salida de AAC) es

$\text{Retraso}_{\text{Herramienta de SBR}} = 320 - 32 + 1 + 6 \times 32 = 481$

muestras.

[0016] De forma característica, la herramienta SBR se ejecuta en el modo "incrementar la velocidad de muestreo" (o "velocidad dual"), en el cual, el retraso de 481 muestras a la velocidad de muestreo AAC se traduce en un retraso de 962 muestras a la velocidad de salida SBR. También puede operar a la misma velocidad de muestreo que la salida AAC (denotada como "modo SBR con velocidad de muestreo reducida"), en el cual, el retraso adicional es de solo 481 muestras a la velocidad de salida de SBR. Hay un modo "compatible hacia atrás" en el cual la herramienta SBR es despreciable y la salida AAC es la salida del decodificador. En este caso, no hay retraso adicional.

[0017] La Figura 1 muestra el comportamiento del decodificador para el caso más común en el cual la herramienta SBR se ejecuta en el modo de incremento de velocidad de muestreo y el retraso adicional es de 962 muestras de salida. Este retraso corresponde aproximadamente al 47% de la longitud de la trama AAC con incremento de velocidad de muestreo (después de procesamiento SBR). Cabe destacar que T1 es el sello de tiempo asociado con CU 1 después del retraso de 962 muestras, es decir, el sello de tiempo para la primera muestra válida de la salida HE AAC. Cabe destacar adicionalmente que si HE AAC se ejecuta en el "modo SBR con reducción en velocidad de muestreo" o modo "de una sola velocidad", el retraso sería 481 muestras pero el sello de tiempo será idéntico ya que en el modo de una sola velocidad, las CU son la mitad del número de muestras de tal manera que serían idénticas ya que en el modo de una sola velocidad las CU son la mitad del número de muestras de modo que el retraso es todavía el 47% de la duración CU.

[0018] Para todos los mecanismos de señalización disponibles (es decir, señalización implícita, señalización explícita compatible hacia atrás o señalización explícita jerárquica) si el decodificador es HE-AAC, entonces debe transportar a Sistemas cualquier retraso adicional incurrido por procesamiento de SBR, de otra forma la falta de una indicación desde el decodificador señala que el decodificador es AAC. Por lo tanto, Sistemas puede ajustar el sello de tiempo para compensar el retraso SBR adicional.

[0019] La siguiente sección describe como un codificador y decodificador para un códec de audio basado en transformada se refiere a Sistemas MPEG, y propone un mecanismo adicional para asegurar la identidad de la señal después de un viaje redondo de codificador-decodificador excepto "artefactos de codificación" – especialmente en la presencia de extensiones códec. Emplear las técnicas descritas asegura una operación pronosticable desde un punto de vista de Sistemas y también retira la necesidad por señalización sin "interrupciones" de propiedad adicional, que es normalmente necesaria para describir el comportamiento del codificador.

[0020] En esta sección, se hace referencia a las siguientes normas:

[1] ISO/IEC TR 14496-24:2007: Tecnología de Información – Codificación de objetos audio-visuales – Parte 24: Interacción de audio y sistemas

[2] ISO/IEC 14496-3:2009 Tecnología de Información – Codificación de objetos audio-visuales – Parte 3: Audio

[3] ISO/IEC 14496-12:2008 Tecnología de Información – Codificación de objetos audio-visuales – Parte 12: ISO formato de archivo multimedia base.

[0021] Brevemente [1] se describe en esta sección. Básicamente, la Codificación de Audio Avanzada (AAC, *Advanced Audio Coding*) y sus sucesores HE AAC, HE AAC v2 son códecs que no tienen una correspondencia 1:1 entre datos comprimidos y no comprimidos. El codificador agrega muestras de audio adicionales al inicio y al final de los datos no comprimidos y también produce Unidades de Acceso con datos comprimidos para éstas, además de las Unidades de Acceso que cubren los datos originales no comprimidos. Un decodificador que cumple con normas o estándares generará entonces una transmisión de datos no comprimida que contiene las muestras adicionales, que son añadidas por el codificador.

[0022] [1] describe como herramientas existentes del formato de archivo multimedia basado en ISO [3] pueden reutilizarse para marcar el intervalo válido de los datos descomprimidos de tal manera que (además de artefactos códec) la transmisión no comprimida original puede recuperarse. La marca se logra al utilizar una lista de adición con una entrada, que contiene el intervalo válido después de la operación de decodificación.

[0023] Ya que esta solución no está lista en tiempo, soluciones de propiedad para marcar el periodo válido no son amplias en uso (por nombrar solo dos: Apple iTunes y Ahead Nero). Puede argumentarse que el procedimiento propuesto en [1] no es muy práctico y padece el problema de que las listas de edición originalmente estaban pensadas para un propósito diferente – potencialmente complejo – para el cual solo están disponibles unas cuantas implementaciones.

[0024] Además, [1] muestra cómo pueden manejarse datos antes de vídeo o reproducción “(pre-roll)” al utilizar grupos de muestra Formato de Archivo ISO (ISO FF = ISO File Format) [3]. Antes de vídeo “(pre-roll)” no marca qué datos son válidos sino cuántas Unidades de Acceso (o muestras en la nomenclatura ISO FF) se van a decodificar antes de que el decodificador saque un punto arbitrario en tiempo. Para AAC esto siempre es una muestra (es decir, una Unidad de Acceso) con anticipación debido a las ventanas de superposición en el dominio MDCT, por lo tanto, el valor para antes de vídeo es -1 para todas las Unidades de Acceso.

[0025] Otro aspecto se refiere a consulta anticipada adicional de muchos codificadores. La consulta anticipada adicional depende por ejemplo del procesamiento de señal interna dentro del codificador que intenta crear salida en tiempo real. Una opción para tener en cuenta la consulta anticipada adicional puede ser utilizar la lista de edición también para el retraso de consulta anticipada del codificador.

[0026] Como se ha mencionado anteriormente, es cuestionable si el propósito original de la herramienta de lista de edición era marcar los intervalos válidos originalmente dentro de un medio. [1] no hace mención de las implicaciones de edición adicional del archivo con listas de edición, por lo tanto, puede considerarse que utilizar la lista de edición para el propósito [1] agrega cierta fragilidad.

[0027] Como una nota secundaria, las soluciones de propiedad y soluciones para audio MP3 definen todas el retraso de extremo-a-extremo adicional y la longitud de los datos de audio no comprimidos originales, muy similar a las soluciones de Nero y de iTunes mencionadas anteriormente para las que se emplea la lista de edición en [1].

[0028] En general, [1] es silencioso respecto al comportamiento correcto de aplicaciones de transmisión en tiempo real, que no utilizan el formato de archivo MP4, pero requieren sellos de tiempo para la correcta sincronización de audio vídeo y a menudo operan en un modo muy silencioso o mudo. Estos sellos de tiempo se ajustan a menudo de manera incorrecta y, por lo tanto, se requiere un botón en el dispositivo de decodificación para poner todo en sincronía de nuevo.

[0029] La interfaz entre Audio MPEG-4 y Sistemas MPEG-4 se describe con más detalle en los siguientes párrafos.

[0030] Toda unidad de acceso suministrada al decodificador de audio de la interfaz de Sistemas resultará en una unidad de composición correspondiente suministrada desde el decodificador de audio a la interfaz de sistemas, es decir, el compositor. Esto incluirá condiciones de arranque y parada, es decir cuando la Unidad de acceso es la primera o la última en una secuencia finita de unidades de acceso.

[0031] Para una unidad de composición de audio, ISO/IEC 14496-1 subclase 7.1.3.5 *Composition Time Stamp (CTS)* especifica que el tiempo de composición aplica a la muestra de audio n -ésima con la unidad de composición. El valor de n es 1 a menos que se especifique diferente en el resto de esta subcláusula.

[0032] Para datos comprimidos, como el audio codificado HE-AAC, que pueden decodificarse por diferentes configuraciones de decodificador, se requiere atención especial. En este caso, la decodificación puede realizarse de una forma compatible hacia atrás (solo AAC) así como de una forma mejorada (AAC+SBR). Con el fin de asegurar que los sellos de tiempo de composición se manejen correctamente (de tal manera que el audio permanezca

sincronizado con otros medios), aplica lo siguiente:

- Si los datos comprimidos permiten tanto la decodificación mejorada como compatible hacia atrás, y si el decodificador opera de una forma compatible hacia atrás, entonces el decodificador no tiene que realizar ninguna acción especial. En este caso, el valor de n es 1.
- Si los datos comprimidos permiten tanto la decodificación mejorada como compatible hacia atrás, y si el decodificador opera de una forma mejorada, tal que utiliza un posprocesador que inserta cierto retraso adicional (por ejemplo, un posprocesador SBR en HE-AAC), entonces debe asegurarse que este retraso de tiempo adicional incurrido respecto al modo compatible hacia atrás, como se describió por un valor correspondiente de n , se tiene en cuenta cuando presenta la unidad de composición. El valor de n se especifica en la siguiente tabla.

Valor de n	Retraso adicional (Nota 1)	Modo de operación del decodificador
1	0	A) Todos los modos de operación no citados en otra parte de esta tabla.
963	962	B1) Decodificador HE-AAC o HE-AAC v2 con SBR operado en modo de velocidad dual; decodificación HE-AAC o HE-AAC v2 audio comprimido.
482	481	B2) mismo que B1), pero con SBR operado en modo de reducción en la velocidad de muestreo.
Nota 1: el retraso introducido por el posprocesamiento se da en el número de muestras (por canal de audio) a la velocidad de muestra de salida para el modo de operación del decodificador determinado.		

[0033] La descripción de la Interfaz entre Audio y Sistemas ha demostrado que funciona de manera fiable, cubriendo la mayoría de los casos de uso actual. Si se busca cuidadosamente, sin embargo, no se mencionan dos aspectos:

- En muchos sistemas, el origen del sello de tiempo es el valor cero. AUs antes de vídeo no se considera que existen, aunque por ejemplo AAC tiene un retraso de codificador mínimo inherente de una Unidad de Acceso que requiere una Unidad de Acceso frente a la Unidad de Acceso en el sello de tiempo cero. Para el formato de archivo de MP4 una solución para este problema se describe en [1].
- Duraciones no-entero del tamaño de trama no están cubiertas. La estructura AudioSpecificConfig() permite la señalización de un pequeño conjunto de tamaños de trama que describen las longitudes de banco de filtros, por ejemplo 960 y 1024 para AAC. Datos del mundo real, sin embargo, de forma característica no caben en una rejilla de tamaños de trama fijos y, por lo tanto, un codificador debe rellenar la última trama.

[0034] Estos dos aspectos que faltan se han convertido en un problema recientemente, con la aparición de aplicaciones multimedia avanzadas que requieren la unión de dos transmisiones AAC o la recuperación del intervalo de muestras válidas después del viaje redondo codificador-decodificador – especialmente en la ausencia del formato de archivo MP4 y los procedimientos descritos en [1].

[0035] Para superar los problemas mencionados anteriormente, antes de vídeo “(pre-roll)”, después de vídeo (post-roll) y todas las demás fuentes deben describirse de forma adecuada. Además, un mecanismo para múltiples no enteros del tamaño de trama se requiere que tenga representaciones de audio precisas de muestra.

[0036] Antes de vídeo se requiere inicialmente para un decodificador, de tal manera que sea capaz de decodificar completamente los datos. Como un ejemplo, AAC requiere antes de vídeo de 1024 muestras (una Unidad de Acceso) antes de la decodificación de una Unidad de Acceso de tal manera que las muestras de salida de la operación de superposición – adición representa la señal original deseada, como se ilustra en [1]. Otros códecs de audio pueden tener diferentes requisitos antes de vídeo.

[0037] Después de vídeo (post-roll) es equivalente a antes de vídeo con la diferencia de que se van a alimentar más datos después de la decodificación de una Unidad de Acceso al decodificador. La causa para después de vídeo (post-roll) es la extensión de códec que desarrolla una eficiencia códec en intercambio por retraso algorítmico, tal como se ha citado en la tabla anterior. Dado que a menudo se desea una operación de modo dual, antes de vídeo permanece constante de tal manera que un decodificador sin las extensiones implementadas puede utilizar completamente los datos codificados. Por lo tanto, antes de vídeo y sellos de tiempo se refieren a las capacidades del decodificador anterior. Después de vídeo, se requiere entonces además de un decodificador que soporte estas extensiones, dado que la línea de retraso internamente existente debe descargarse para recuperar toda la representación de la señal original. Desafortunadamente, después de vídeo es dependiente del decodificador. Sin embargo, es posible manejar antes de vídeo y después de vídeo independiente del decodificador si los valores antes de vídeo y después de vídeo son conocidos para la capa de sistemas y la salida del decodificador de antes de vídeo y después de vídeo se puede colocar allí.

[0038] Con respecto a un tamaño de trama de audio variable, dado que los códecs de audio codifican

siempre bloques de datos con un número fijo de muestras, una representación precisa de muestras se vuelve solo posible mediante la señalización adicional en el nivel de Sistemas. Dado que es más fácil que un decodificador maneje el recorte preciso de muestra, parece conveniente hacer que el decodificador corte una señal. Por lo tanto, se propone un mecanismo de extensión opcional que permite el recorte de las muestras de salida por el
5 decodificador.

[0039] Respecto a un retraso de codificador específico de proveedor, MPEG sólo especifica la operación de un decodificador, mientras que solo se proporcionan codificadores de manera informal. Esta es una de las ventajas de las tecnologías MPEG, donde los codificadores pueden mejorar con el tiempo para utilizar completamente las
10 capacidades de un códec. La flexibilidad al diseñar un codificador sin embargo ha llevado a problemas de interoperabilidad de retraso. Dado que los codificadores requieren de forma característica una vista previa de la señal de audio para realizar decisiones de codificación más inteligentes, esto es altamente específico del distribuidor. Razones para este retraso del codificador son, por ejemplo, decisiones conmutadas de bloque, que requieren un retraso de las superposiciones de ventana posibles y otras optimizaciones, que son primordialmente
15 relevantes para codificadores de tiempo real.

[0040] La codificación basada en archivo de contenido disponible fuera de línea no requiere este retraso que es solo relevante cuando se codifican datos de tiempo real, sin embargo, la mayoría de los codificadores agregan al principio silencio también al inicio de codificaciones fuera de línea.
20

[0041] Una parte de la solución para este problema es el ajuste correcto de sellos de tiempo en la capa de sistemas, de tal manera que estos retrasos sean irrelevantes y tengan por ejemplo valores de sello de tiempo negativos. Esto también puede lograrse con la lista de edición, como se propone en [1].

[0042] La otra parte de la solución es un alineamiento del retraso del codificador a las fronteras de la trama, de tal manera que un número entero de Unidades de Acceso por ejemplo con sellos de tiempo negativos puede saltarse inicialmente (además de las Unidades de Acceso antes de vídeo).
25

[0043] Las enseñanzas descritas en esta invención se refieren también a la norma industrial ISO/IEC 14496-3:2009, subparte 4, sección 4.1.1.2. Según las enseñanzas descritas en esta invención, se propone lo siguiente: cuando está presente, una herramienta de recorte posterior al decodificador selecciona una porción de la señal de audio reconstruida, de tal manera que se puedan unir dos transmisiones en conjunto en el dominio codificado y la reconstrucción precisa de muestras se vuelve posible dentro de la capa de Audio.
30

[0044] La entrada a la herramienta de recorte posdecodificador es:

- La señal de audio reconstruida en dominio de tiempo
- La información de control posterior a recorte

[0045] La salida de la herramienta de recorte posdecodificador es:

- La señal de audio reconstruida en dominio de tiempo

[0046] Si la herramienta de recorte posdecodificador no está activa, la señal de audio reconstruida en
45 dominio de tiempo se pasa directamente a la salida del decodificador. Esta herramienta se aplica después de cualquier herramienta de codificación de audio anterior.

[0047] La siguiente tabla ilustra una sintaxis propuesta de una estructura de datos `extension_payload()` que puede emplearse para implementar las enseñanzas descritas en esta invención.
50

Sintaxis	N. ° de bits	Mnemónica
extension_payload(cnt)		
{		
extension_type;	4	uimbsf
align = 4;		
switch(extension_type) {		
case EXT_TRIM:		

```

    return trim_info();

case EXT_DYNAMIC_RANGE:

    return dynamic_range_info();

case EXT_SAC_DATA:

    return sac_extension_data(cnt);

case EXT_SBR_DATA:

    return sbr_extension_data(id_aac, 0);                                Nota 1

case EXT_SBR_DATA_CRC:

    return sbr_extension_data(id_aac, 1);                                Nota 1

case EXT_FILL_DATA:

    fill_nibble; /* must be '0000' */                                4           uimsbf
    for (i=0; i<cnt-1; i++) {

        fill_byte[i]; /* must be '10100101' */                        8           uimsbf
    }

    return cnt;

case EXT_DATA_ELEMENT:

    data_element_version;                                            4           uimsbf

    switch( data_element_version ) {

        case ANC_DATA:

            loopCounter = 0;

            dataElementLength = 0;

            do {

                dataElementLengthPart;                                8           uimsbf

                dataElementLength += dataElementLengthPart;

                loopCounter++;

            } while (dataElementLengthPart == 255);

            for (i=0; i<dataElementLength; i++) {

                data_element_byte[i];                                    8           uimsbf
            }
    }

```

```

        return (dataElementLength+loopCounter+1);
    default:
        align = 0;
    }

case EXT_FIL:
default:
    for (i=0; i<8*(cnt-1)+align; i++) {
        other_bits[i];
        1          uimsbf
    }

    return cnt;
}
}

Nota 1: id_aac es el id_syn_ele del elemento AAC correspondiente (ID_SCE o ID_CPE) o ID_SCE en caso de CCÉ.

```

[0048] La siguiente tabla ilustra una sintaxis propuesta de una estructura de datos trim_infoo que puede utilizarse para implementar las enseñanzas descritas en esta invención.

Sintaxis	No. de bits	Mnemónica
trim_infoo()		
{		
custom_resolution_present;	1	uimsbf
trim_resolution = samplingFrequency;		
if (custom_resolution_present == 1) {		
custom_resolution;	19	uimsbf
trim_resolution = custom_resolution;		
}		
trim_from_beginning;	12	uimsbf
trim_from_end;	12	uimsbf
}		

5 con las siguientes definiciones respecto a Recorte posdecodificador:

custom_resolution_present Bandera que indica si está presente custom_resolution.

10 **custom_resolution** Una resolución a la medida en Hz que se utiliza para la operación de recorte. Se recomienda establecer una resolución a la medida cuando es posible el procesamiento de múltiples velocidades de la señal de audio y la operación de recorte requiere ser realizada con la resolución adecuada más alta.

trim_resolution El intervalo predefinido es la frecuencia de muestreo nominal como se indica en la Tabla 1.16 de ISO/IEC 14496-3:2009 por samplingFrequency o samplingFrequencyIdx. Si se ajusta la bandera custom_resolution_present entonces la resolución para la herramienta de recorte posdecodificador es el valor de

custom_resolution.

trim_from_beginning (N_B) Número de muestras PCM que se van a retirar desde el inicio de la Unidad de Composición. El valor solo es válido para una señal de audio con la velocidad de trim_resolution. Si trim_resolution no es igual a la frecuencia de muestreo de la señal de alimentación de dominio-tiempo, el valor debe ajustarse en

5 escala de forma apropiada según la siguiente ecuación:

$N_B = \text{suelo} (N_B \cdot \text{sampling_frequency} / \text{trim_resolution})$

trim_from_end (N_E) Número de muestras PCM que se van a retirar desde el final de la Unidad de Composición. Si trim_resolution no es igual a la frecuencia de muestreo de la señal de alimentación de dominio de tiempo, el valor debe ajustarse en escala apropiadamente según la siguiente ecuación:

10 $N_E = \text{suelo} (N_E \cdot \text{sampling_frequency} / \text{trim_resolution})$

[0049] Otro algoritmo para mezclado de transmisión posible puede tener en cuenta la unión sin juntas (sin la posibilidad de discontinuidades de la señal). Este aspecto también es válido para datos PCM sin comprimir y es ortogonal a las enseñanzas descritas en esta invención.

15

[0050] En lugar de una resolución a la medida, también puede ser apropiado un porcentaje. Alternativamente, la velocidad de muestreo más alta puede emplearse, pero esto puede entrar en conflicto con el procesamiento de velocidad dual y los decodificadores que soportan recorte, pero no procesamiento de velocidad dual, por lo tanto, se prefiere una solución independiente de implementación de decodificador y parece lógica una resolución de recorte a

20 la medida.

[0051] Respecto al proceso de decodificación, se aplica el recorte posdecodificador después de que se procesan todos los datos de una Unidad de Acceso (es decir, después de que se han aplicado extensiones como DRC, SBR, PS, etc.). El recorte no se realiza en la capa de Sistemas MPEG-4; sin embargo, los sellos de tiempo y

25 valores de duración de una Unidad de Acceso deberán corresponder a la consideración de que se aplica el recorte.

[0052] El recorte se aplica para la Unidad de Acceso que transporta la información solo si no se ha introducido retraso extra debido a extensiones opcionales (por ejemplo, SBR). Si estas extensiones están en su sitio y se utilizan dentro del decodificador, entonces la aplicación de la operación de recorte se retrasa por el retraso de extensiones opcional. Por lo tanto, la información de recorte requiere ser almacenada dentro del decodificador y

30 deben proporcionarse Unidades de Acceso adicionales por la capa de Sistemas.

[0053] Si el decodificador puede operar a más de una velocidad, se recomienda utilizar una resolución a la medida para la operación de recorte con la más alta velocidad.

35

[0054] El recorte puede llevar a discontinuidades de la señal, lo que puede provocar distorsión de señal. Por lo tanto, información de recorte solo deberá insertarse en la transmisión de bits al inicio o al final de toda la codificación. Si se unen en conjunto dos transmisiones, estas discontinuidades no pueden evitarse excepto por un codificador que establece cuidadosamente los valores de trim_from_end y trim_from_beginning de tal manera que

40 las dos señales de dominio de tiempo de salida se ajustan en conjunto sin discontinuidades.

[0055] Recortar Unidades de Acceso puede llevar a requisitos de computación inesperados. Muchas implementaciones tienen en cuenta el tiempo de procesamiento constante para Unidades de Acceso con duración constante, que no es más válido si la duración cambia debido al recorte, pero los requisitos computacionales para una Unidad de Acceso permanecen. Por tanto, deberán ser considerados los decodificadores con recursos computacionales restringidos y, por lo tanto, el recorte deberá ser utilizado raramente, de preferencia al codificar

45 datos de una forma que se alinean con las fronteras de Unidad de Acceso y el recorte solo se utiliza al final de una codificación, como se describe en [ISO/IEC 14496-24:2007 Anexo B.2].

[0056] Las enseñanzas descritas en esta invención se refieren también a la norma industrial ISO/IEC 14496-24:2007. Según las enseñanzas descritas en esta invención, se propone lo siguiente respecto a una interfaz de decodificador de audio para Acceso preciso de muestra: un decodificador de audio creará siempre una Unidad de Composición (CU) de una Unidad de Acceso (AU). La cantidad requerida de antes de vídeo "(pre-roll)" y después de vídeo "(post-roll)" de AUs es constante para un ajuste en serie de AUs por un codificador.

55

[0057] Cuando empieza la operación de decodificación, el decodificador se inicializa con un AudioSpecificConfig (ASC). Después de que el decodificador ha procesado esta estructura, los parámetros más relevantes pueden solicitarse a partir del decodificador. Además, la capa de Sistemas transporta parámetros que en general son independientes del tipo de transmisión, sean audio o vídeo u otros datos. Esto incluye información de sincronización, datos antes de vídeo y después de vídeo. En general, el decodificador requiere r_{pre} AUs antes de vídeo antes de la AU, que contiene la muestra solicitada. Además, se requieren r_{post} después de vídeo, esto depende sin embargo del modo de decodificación (decodificación de una extensión puede requerir AUs después de vídeo mientras que la operación de decodificación básica se define como que no requiere una AU después de vídeo).

60

[0058] Cada AU deberá ser marcado por el decodificador si es una AU antes de vídeo o después de vídeo,

65

para permitir que el decodificador cree la información de estado interno requerida para la posterior decodificación o para descargar datos restantes dentro del decodificador, respectivamente.

[0059] Según la invención, la comunicación entre la capa de sistemas y el decodificador de audio se ilustra en la Figura 2.

[0060] Según la invención, el decodificador de audio se inicializa por la capa de Sistemas con una estructura `AudioSpecificConfig()`, que resulta en una configuración de salida del decodificador a la capa de Sistemas, que contiene información en la frecuencia de muestra, la configuración de canal (por ejemplo 2 para estéreo), el tamaño de trama n (por ejemplo 1024 en el caso de AAC LC) y un retraso extra d para extensiones de códec explícitamente señaladas, tales como SBR. En particular, la Figura 2 muestra las siguientes acciones:

1. Según la invención, las primeras Unidades de Acceso antes de vídeo r_{pre} se proporcionan al decodificador y se descartan silenciosamente después de la decodificación por la Capa de Sistemas.

2. Según la invención, la primera Unidad de Acceso no-antes de vídeo contiene información `trim_from_beginning` en una carga útil de extensión del tipo `EXT_TRIM` de tal manera que el decodificador solo envía salida a muestras PCM. Además, las muestras extra d PCM, generadas por una extensión de códec opcional, deben ser borradas.

Dependiendo de la implementación, esto puede suceder al retrasar todas las demás transmisiones paralelas por d o al marcar las primeras muestras d como no válidas y tomar acción apropiada tal como el borrado de las muestras no válidas al tiempo de procesamiento o reproducción, o de preferencia dentro del decodificador.

Si el borrado de las d muestras sucede dentro del decodificador, como se recomienda, entonces la capa de sistemas requiere estar al tanto de que la primera Unidad de Composición que contiene a muestras solo puede proporcionarse por el decodificador después del consumo de r_{post} Unidades de Acceso, como se establece en la 6ª etapa.

3. Entonces todas las Unidades de Acceso con la duración constante n se decodifican y las Unidades de Composición se proporcionan a la capa de Sistemas.

4. Según la invención, la Unidad de Acceso antes de las Unidades de Acceso después de vídeo puede contener información opcional de `trim_from_end`, de tal manera que el decodificador solo genera b muestras PCM.

5. Según la invención, se proporcionan las últimas Unidades de Acceso r_{post} después de vídeo al decodificador de audio de tal manera que las muestras d PCM que faltan puedan generarse. Dependiendo del valor d (que puede ser cero) esto puede resultar en Unidades de Composición sin ninguna muestra. Se recomienda proporcionar todas las Unidades de Acceso después de vídeo al decodificador, de tal manera que puede des-inicializar por completo, independientemente del valor del retraso extra d .

[0061] Los codificadores deberán tener comportamiento de sincronización consistente. Un codificador deberá alinear las señales de alimentación de tal manera que después de decodificar las AUs r_{pre} antes de vídeo la señal de alimentación original resulte, sin pérdida inicial y sin las muestras de encabezado. En especial para operaciones de codificador basadas en archivo esto requerirá que las muestras de consulta anticipada adicional del codificador y adicionalmente muestras de silencio insertado son un múltiplo entero del tamaño de trama de audio y de esta manera pueden descartarse en la salida del codificador.

[0062] En escenarios donde este alineamiento no es posible, por ejemplo, codificación en tiempo real de audio, el codificador deberá insertar información de recorte de tal manera que el decodificador se habilita para borrar muestras de consulta anticipada accidentalmente insertadas con la herramienta de recorte posdecodificador.

Similarmente, los codificadores deberán insertar información de recorte posdecodificador para las muestras subsecuentes. Estas deberán ser señalizadas en la Unidad de Acceso que precede a las últimas AUs r_{post} después de vídeo.

[0063] La información de sincronización ajustada en el codificador deberá ajustarse considerando que está disponible la herramienta de recorte posterior al decodificador.

[0064] La Figura 3 muestra un diagrama de flujo esquemático de un procedimiento para proporcionar información en la validez de datos de audio codificados según una primera forma de realización posible. El procedimiento comprende una acción 302 de acuerdo a la cual se proporciona información que describe la cantidad de datos al inicio de una unidad de datos de audio que no son válidos. La información proporcionada puede entonces insertarse en o combinarse con la unidad de datos de audio codificada y que está involucrada. La cantidad de datos puede ser expresada como un número de muestras (por ejemplo, muestras PCM), microsegundos, milisegundos o un por ciento de una longitud de una sección de señal de audio que se proporciona por la unidad de datos de audio codificados.

[0065] La Figura 4 muestra un diagrama de flujo esquemático de un procedimiento para proporcionar información de la validez de los datos de audio codificados, según una segunda forma de realización posible de las enseñanzas descritas en esta invención. El procedimiento comprende una acción 402, de acuerdo a la cual se proporciona información que describe la cantidad de datos al final de una unidad de datos de audio que no es válida.

[0066] La Figura 5 muestra un diagrama de flujo esquemático de un procedimiento para proporcionar información de la validez de datos de audio codificados según una tercera forma de realización posible de las enseñanzas descritas en esta invención. El procedimiento comprende una acción 502 de acuerdo a la cual se proporciona información que describe tanto la cantidad de datos al inicio como al final de una unidad de datos de audio que no son válidos.

[0067] En las formas de realización ilustradas en las Figuras 3 a 5, la información que describe la cantidad de datos dentro de la unidad de datos de audio que no son válidos, puede obtenerse de un proceso de codificación que genera los datos de audio codificados. Durante la codificación de datos de audio, un algoritmo de codificación puede considerar un intervalo de alimentación de muestras de audio que se extiende sobre una frontera (inicio o fin) de una señal de audio que se va a codificar. Procesos de codificación típicos obtienen una pluralidad de muestras de audio en "bloques" o "tramas" de tal manera que un bloque o trama que no se llena por completo con muestras de audio actuales puede llenarse con muestras de audio "de relleno o falsas" que tienen de forma característica una amplitud cero. Para el algoritmo de codificación, esto ofrece la ventaja de que los datos de alimentación siempre se realizan de la misma manera, de tal manera que el procesamiento de datos dentro del algoritmo no tiene que ser modificado dependiendo de los datos de audio procesados que contienen una frontera (inicio o fin). En otras palabras, los datos introducidos se acondicionan, respecto a organización y dimensión de datos, a los requisitos del algoritmo de codificación. De forma característica, el acondicionamiento de los datos de entrada lleva inherentemente una estructura correspondiente de los datos de salida, es decir los datos de salida reflejan el acondicionamiento de los datos de entrada. Por lo tanto, los datos de salida difieren de los datos de entrada original (antes del acondicionamiento). Esta diferencia es inaudible de forma característica debido a que solo las muestras que tienen una amplitud cero se han agregado a los datos de audio originales. Sin embargo, el acondicionamiento puede modificar la duración de los datos de audio originales, alargando de forma característica los datos de audio originales por segmentos silenciosos.

[0068] La Figura 6 muestra un diagrama de flujo esquemático de un procedimiento para recibir datos codificados incluyendo la información de la validez de datos según una forma de realización de las enseñanzas descritas en esta invención. El procedimiento comprende una acción 602 de recepción de los datos codificados. Los datos codificados contienen información que describe la cantidad de datos que no son válidos. Al menos tres casos pueden distinguirse: la información puede describir la cantidad de datos al inicio de una unidad de datos de audio que no son válidos, la cantidad de datos al final de una unidad de datos de audio que no son válidos, y la cantidad de datos al inicio y final de una unidad de datos de audio que no son válidos.

[0069] En una acción 604 del procedimiento para la recepción de datos codificados, se proporcionan datos de salida decodificados que solo contienen las muestras no marcadas como no válidas. Un consumidor de los datos de salida decodificados posteriores de un elemento que ejecuta el procedimiento para recibir datos codificados, puede emplear los datos de salida decodificados proporcionados, sin tener que tratar con el aspecto de la validez de porciones de los datos de salida, tales como muestras únicas.

[0070] La Figura 7 muestra un diagrama de flujo esquemático del procedimiento para la recepción de datos codificados según otra forma de realización de las enseñanzas descritas en esta invención. Los datos codificados se reciben en una acción 702. En una acción 704, los datos de salida decodificados que contienen todas las muestras de audio de una unidad de datos de audio codificados se proporcionan, por ejemplo, a una aplicación posterior que consume los datos de salida decodificados. Además, se proporciona información mediante una acción 706, cuya parte de los datos de salida decodificados es válida. La aplicación que consume los datos de salida decodificados puede entonces desprender datos no válidos y concatenar segmentos sucesivos de datos válidos, por ejemplo. De esta manera, los datos de salida decodificados pueden ser procesados por la aplicación para no contener silencios artificiales.

[0071] La Figura 8 muestra un diagrama de entrada/salida de un codificador 800 según una forma de realización de las enseñanzas descritas en esta invención. El codificador 800 recibe datos de audio, por ejemplo, una transmisión de muestras PCM. Los datos de audio se codifican entonces mediante el uso de un algoritmo de codificación sin pérdida o un algoritmo de codificación con pérdida. Durante ejecución, el algoritmo de codificación puede tener que modificar los datos de audio que se proporcionan en una entrada del codificador 800. Una razón para esta modificación puede ser hacer el ajuste de datos de audio originales a los requisitos del algoritmo de codificación. Como se ha mencionado anteriormente, una modificación típica de los datos de audio originales es la inserción de muestras de audio extra de tal manera que los datos de audio originales se ajustan en un número entero de tramas o bloques y/o de tal manera que el algoritmo de codificación se inicializa adecuadamente antes que se procese la primera muestra de audio real. Información respecto a la modificación realizada puede ser obtenida a partir del algoritmo de codificación o una entidad del codificador 800 que realiza el acondicionamiento de los datos de audio de entrada. De esta información de modificación, puede derivarse una información que describe la cantidad de información al inicio y/o al final de una unidad de datos de audio que no son válidos. El codificador 800 puede comprender por ejemplo un contador para contar muestras marcadas como no válidas por el algoritmo de codificación o la entidad de acondicionamiento de datos de audio de entrada. La información que describe la cantidad de información al inicio y/o al final de la unidad de datos de audio, que no es válida se proporciona en una

salida del codificador 800 junto con los datos de audio codificados.

[0072] La Figura 9 muestra un diagrama de entrada/salida esquemático de un codificador 900 según otra forma de realización de las enseñanzas descritas en esta invención. En comparación con el codificador 800
5 mostrado en la Figura 8, la salida del codificador 900 mostrada en la Figura 9 sigue un formato diferente. Los datos de audio codificados enviados por el codificador 900 se formatean como una transmisión o serie de unidades de datos de audio codificados 922. Junto con cada unidad de datos de audio codificados 922, una información de validez 924 está contenida en la transmisión. Una unidad de datos de audio codificados 922 y su información de validez correspondiente 924 puede considerarse como una unidad de datos de audio codificados mejorada 920.
10 Utilizando la información de validez 924, un receptor de la transmisión de unidades de datos de audio mejorada 920 puede decodificar las unidades de datos de audio codificados 922 y utilizar solo aquellas partes que están marcadas como datos válidos. Cabe destacar que la expresión "unidad de datos de audio codificados mejorada" no necesariamente implica que su formato es diferente de unidades de datos de audio codificados no mejorados. Por ejemplo, la información de validez puede almacenarse en un campo de datos actualmente no usado de una unidad
15 de datos de audio codificada.

[0073] La Figura 10 muestra un diagrama de bloques esquemático de un decodificador 1000 según una forma de realización de las enseñanzas descritas en esta invención. El decodificador 1000 recibe datos codificados en una entrada 1002 que envía unidades de datos de audio codificados a una porción de decodificación 1004. Los
20 datos codificados comprenden información de la validez de datos, como se ha descrito anteriormente respecto a la descripción del procedimiento para proporcionar información de la validez de datos de audio codificados o el codificador correspondiente. La entrada 1002 del decodificador 1000 puede configurarse para recibir información de la validez de datos. Esta característica es opcional como se indica por la flecha discontinua que lleva a la entrada 1002. Además, la entrada 1002 puede ser configurada para proporcionar la información de la validez de datos a la
25 porción de decodificación 1004. De nuevo, esta característica es opcional. La entrada 1002 puede simplemente enviar la información respecto a la validez de datos a la porción de decodificación 1004, o la entrada 1002 puede extraer la información respecto a la validez de datos desde los datos codificados en los cuales está contenida la información en la validez de datos. Como una alternativa a la entrada 1002 que maneja la información respecto a la
30 validez de datos, la porción de decodificación 1004 puede extraer esta información y utilizarla para filtrar datos no válidos. La porción de decodificación 1004 se conecta a una salida 1006 del decodificador 1000. Muestras de audio decodificadas válidas se transmiten o envían por la porción de decodificación 1004 a la salida 1006 que proporciona muestras de audio válidas a una entidad consumidora posterior de las muestras de audio válidas, tal como un procesador o reproductor de audio. El procesamiento de la información de la validez de datos es transparente a la entidad de consumo posterior. Al menos una de la porción de decodificación 1004 y la salida 1006 puede
35 configurarse para disponer las muestras de audio decodificadas válidas de tal manera que no se produzca espacio, incluso si se han retirado muestras de audio no válidas de una transmisión de muestras de audio que se van a presentar en la entidad de consumo posterior.

[0074] La Figura 11 muestra un diagrama de bloques esquemático de un decodificador 1100 según otra forma de realización de las enseñanzas descritas en esta invención. El decodificador 1100 comprende una entrada
40 1102, la porción de decodificación 1104 y una salida 1106. La entrada 1102 recibe datos codificados y proporciona unidades de datos de audio codificados a la porción de decodificación 1104. Como se ha explicado anteriormente en conexión con el decodificador 1000 mostrado en la Figura 10, la entrada 1102 puede, como una opción, recibir información de validez separada que puede entonces enviarse a la porción de decodificación 1104. La porción de decodificación 1104 convierte las unidades de datos de audio codificados en muestras de audio decodificadas y las envía a la salida 1106. Además, la porción de decodificación también envía la información respecto a la validez de
45 datos a la salida 1106. En caso de que la información respecto a la validez de datos no sea proporcionada por la entrada 1102 a la porción de decodificación 1104, la porción de decodificación 1104 puede determinar la información respecto a la validez de los propios datos. La salida 1106 proporciona las muestras de audio decodificadas y la información respecto a la validez de los datos a una entidad de consumo posterior.
50

[0075] La entidad de consumo posterior puede explotar entonces la información respecto a la validez de los propios datos. Las muestras de audio decodificadas generadas por la porción de decodificación 1104 y proporcionadas por la salida 1106 contienen en general todas las muestras de audio decodificadas, es decir,
55 muestras de audio válidas y muestras de audio no válidas.

[0076] El procedimiento para proporcionar la información respecto a la validez de datos de audio codificados, puede utilizar diversas piezas de información para determinar la cantidad de datos de una unidad de datos de audio que no son válidos. También, el codificador puede utilizar estas piezas de información. Las siguientes secciones
60 describen una cantidad de piezas de información que pueden emplearse para este objetivo: cantidad de datos antes de vídeo, cantidad de datos artificiales extra agregados por el codificador, longitud de los datos de entrada no comprimidos originales, y cantidad de después de vídeo.

[0077] Una pieza de información importante es la cantidad de datos antes de vídeo, que es la cantidad de
65 datos comprimidos que deben ser decodificados antes que la unidad de datos comprimidos correspondiente al inicio

de datos no comprimidos original. De manera ejemplar, una codificación y decodificación de un conjunto de unidades de datos no comprimidos se explica. Dado un tamaño de trama de 1024 muestras y la cantidad de 1024 muestras antes de vídeo también, un conjunto de datos de audio PCM sin comprimir original que consiste en 2000 muestras será codificado como tres unidades de datos codificados. La primera unidad de datos codificados será la
5 unidad de datos antes de vídeo con una duración de 1024 muestras. La segunda unidad de datos codificados resultará en las originales 1024 muestras de la señal fuente (sin dar otros artefactos de codificación). La tercera unidad de datos codificados resultará en 1024 muestras, que consisten en las restantes 976 muestras de la señal fuente y 48 muestras siguientes posteriores introducidas por la granularidad de trama. Debido a las propiedades de los procedimientos de codificación, tales como transformada coseno discreta modificada (MDCT = Modified Discrete
10 Cosine Transform) o un filtro espejo de cuadratura (QMF = Quadrature Mirror Filter) involucrado, no puede evitarse antes de vídeo y es esencial para el decodificador que reconstruye toda la señal original. Por lo tanto, para el ejemplo anterior siempre se requiere una unidad de datos comprimida más que lo esperado por un no experto. La cantidad de datos antes de vídeo es dependiente de codificación y fija para un modo de codificación y constante a lo largo del tiempo. Por lo tanto, esto se requiere también para el acceso de forma aleatoria a unidades de datos
15 comprimidos. Antes de vídeo también se requiere para obtener los datos de salida no comprimidos decodificados correspondientes a los datos de entrada no comprimidos.

[0078] Otra pieza de información es la cantidad de datos artificiales extra agregados por el codificador. Estos datos extra resultan de forma característica de una vista previa de muestras futuras dentro del codificador, de tal
20 manera que puedan realizarse decisiones más inteligentes de codificación, como conmutar bancos de filtro cortos a bancos de filtro largos. Solo el codificador conoce este valor de consulta anticipada y es diferente entre implementaciones de codificador de un distribuidor específico para el mismo modo de codificación, aunque constante a lo largo del tiempo. La longitud de estos datos extra es difícil de detectar por un decodificador y a menudo se aplican heurísticas, por ejemplo, la cantidad de silencio al inicio se considera como retraso de codificador
25 extra o un valor mágico si un cierto codificador es detectado por algunas otras heurísticas.

[0079] La siguiente pieza de información solo disponible para el codificador es la longitud de los datos de alimentación no comprimidos originales. En el ejemplo anterior, 48 muestras posteriores son creadas por el
30 decodificador que no han estado presentes en los datos no comprimidos de entrada originales. La razón es la granularidad de trama, que se fija en un valor dependiente de códec. Un valor típico es 1024 o 960 para MPEG-4 AAC, por lo tanto, el codificador rellena siempre los datos originales para ajustarse en la rejilla de tamaño de trama. Soluciones existentes agregan de forma característica meta datos al nivel de sistema que contienen la suma de todas las muestras extra de encabezado, que resultan de datos antes de vídeo y extra artificiales, y la longitud de los datos de audio fuente. Este procedimiento sin embargo funciona para operaciones basadas en archivo solamente,
35 donde se conoce la duración antes de codificar. También tiene cierta fragilidad cuando se hacen ediciones al archivo; entonces también los meta datos requieren ser actualizados. Una estrategia alternativa es el uso de sellos de tiempo o duraciones a nivel de sistema. Utilizar esto no define desafortunadamente claramente la mitad de los datos que es válida. Además, el recorte puede no ser realizado de forma característica a nivel de sistema.

[0080] Finalmente, otra pieza de información se vuelve cada vez más importante, que es la cantidad de información después de vídeo. Después de vídeo define que tantos datos deben darse a un decodificador después
40 de la unidad de datos codificados, de tal manera que el decodificador puede proporcionar los datos no comprimidos correspondientes a los datos originales no comprimidos. En general, después de vídeo puede intercambiarse con antes de vídeo y viceversa. Sin embargo, la suma de después de vídeo y antes de vídeo no es constante para todos
45 los modos de decodificador. Actuales especificaciones tales como [ISO/IEC 14496-24:2007] consideran un antes de vídeo fijo para todos los modos del decodificador e ignora mencionar después de vídeo en favor de definir el retraso adicional que tiene un valor equivalente a después de vídeo. Aunque se ilustra en la Figura 4 de [ISO/IEC 14496-24:2007], no se menciona que la última unidad de datos codificados (una Unidad de Acceso, AU, en la terminología MPEG) es opcional y actualmente es una AU después de vídeo que solo se requiere para procesamiento de
50 velocidad dual de un decodificador con una baja velocidad y una extensión con velocidad duplicada. Es una forma de realización de la invención definir también un procedimiento para eliminar datos no válidos en la presencia de después de vídeo.

[0081] La información anterior se utiliza, por ejemplo, parcialmente en [ISO/IEC 14496-24:2007] para MPEG-
55 4 AAC en el Formato de Archivo MP4 [ISO/IEC 14496-14]. Ahí, se utiliza una lista de edición denominada de este modo para marcar la porción válida de los datos codificados al definir un desplazamiento y un periodo de validez para los datos codificados en una edición denominada de este modo. También, la cantidad de antes de vídeo puede definirse en una granularidad de trama. Una desventaja de esta solución es el uso de la lista de edición para superar problemas específicos de codificación de audio. Esto entra en contradicción con el uso previo de listas de edición
60 para definir edición no lineal genérica sin modificación de datos. Por lo tanto, se vuelve difícil o incluso imposible distinguir entre las ediciones específicas de audio y ediciones genéricas.

[0082] Otra solución potencial es el procedimiento para recuperar longitud de archivo original en mp3 y mp3Pro. Ahí, el retraso de códec y la duración total del archivo se proporcionan en la primera unidad de datos de
65 audio codificada. Esto desafortunadamente tiene el inconveniente de que solo funciona para operaciones basadas

en archivo o transmisiones con toda la longitud ya conocida cuando el codificador crea la primera unidad de datos de audio codificados, ya que la información está contenida ahí.

[0083] Para superar las desventajas de soluciones existentes, las formas de realización de la invención proporcionan información respecto a la validez de los datos a la salida del codificador dentro de los datos de audio codificados. Las piezas de información se conectan a las unidades de datos de audio codificados que son afectadas. Por lo tanto, datos extras artificiales al inicio se marcan como datos no válidos y datos posteriores empleados para llenar una trama, también se marcan como datos no válidos que deben ser recortados. La marca, según las formas de realización de la invención, permite la distinción de datos válidos contra no válidos dentro de una unidad de datos codificados, de tal manera que un decodificador puede borrar los datos no válidos antes de que proporcione datos a la salida o puede marcar los datos de forma alternativa, por ejemplo, de una forma similar a la representación dentro de la unidad de datos codificados, de tal manera que pueden suceder acciones apropiadas en otros elementos de procesamiento. Los demás datos relevantes, que son antes de vídeo y después de vídeo, se definen dentro del sistema y están comprendidos tanto por el codificador como el decodificador, de tal manera que los valores se conocen para un modo de decodificador dado.

[0084] Por lo tanto, un aspecto de las enseñanzas descritas propone la separación de datos que varían con el tiempo y datos que no varían con el tiempo. Los datos que varían con el tiempo consisten en la información en datos extras artificiales que solo está presente al inicio y los datos siguientes empleados para rellenar una trama. Los datos que no varían con el tiempo consisten en datos antes de vídeo y después de vídeo y no requieren de esta manera ser transmitidos en unidades de datos de audio codificados, sino que deberán ser transmitidos por el contrario fuera-de-banda o se conocen con anticipación por el modo de decodificación, que puede derivarse del registro de configuración del decodificador para un esquema de codificación de audio dado.

[0085] Además, se recomienda ajustar los sellos de tiempo de datos de audio codificados según la información que representa una unidad de datos de audio codificada. Por lo tanto, una muestra de audio no comprimida original con sello de tiempo t se considera que se recupera por la operación de decodificación de la unidad de datos de audio codificados con sello de tiempo t . Esto no incluye unidades de datos antes de vídeo y después de vídeo, que se requieren, además. Por ejemplo, una señal de audio original determinada con 1.500 muestras y un sello de tiempo inicial con valor 1, se codificará como tres unidades de datos de audio codificados de tamaño de trama 1024, antes de vídeo 1024 y retraso artificial extra de 200 muestras. La primera unidad de datos de audio codificados tiene un sello de tiempo de $1-1024 = -1023$ y solamente se emplea para antes de vídeo. La segunda unidad de datos de audio codificados tiene un sello de tiempo de 1 e incluye información dentro de la unidad de datos de audio codificados para recortar las primeras 200 muestras. Aunque el resultado de decodificación consistirá normalmente en 1024 muestras, las primeras 200 muestras se retiran de la salida y solo 824 muestras permanecen. La tercera unidad de datos de audio codificada tiene un sello de tiempo de 825 y también contiene información dentro de la unidad de datos de audio codificados para recortar las muestras de salida de audio resultantes con longitud 1024 a las restantes 676 muestras. Por lo tanto, la información de que las últimas muestras $1024-676=348$ no son válidas se almacena dentro de las unidades de datos de audio codificados.

[0086] En la presencia de por ejemplo 1.000 muestras después de vídeo debido a un modo decodificador diferente, la salida del codificador cambiará a cuatro unidades de datos de audio codificados. Las tres primeras unidades de datos de audio codificados permanecen constantes pero otros datos de audio codificados se agregan. Cuando se decodifican, la operación para la primera Unidad de Acceso antes de vídeo permanece como en el ejemplo anterior. La decodificación para la segunda Unidad de Acceso, sin embargo, debe tener en cuenta el retraso extra para el modo decodificador alterno. Se presentan tres soluciones básicas dentro de este documento para manejar de forma correcta el retraso del decodificador extra.

1. el retraso del decodificador se transmite desde el decodificador al sistema, que retrasa entonces todas las demás transmisiones paralelas para conservar la sincronización de audio-vídeo.

2. el retraso del decodificador se transmite desde el decodificador al sistema, que puede retirar entonces las muestras no válidas en un elemento de procesamiento de audio, por ejemplo, el elemento de procesamiento o reproducción.

3. el retraso del decodificador se retira dentro del decodificador. Esto resulta en una unidad de datos descomprimida con un tamaño más pequeño inicialmente, debido a la eliminación del retraso extra o un retraso de la salida de datos hasta que el número señalado de unidades de datos codificados después de vídeo se proporciona al decodificador. Este último procedimiento se recomienda y se asume para el resto del documento.

[0087] Según la invención, ya sea el decodificador o la capa de sistema incrustada descartará toda la salida que se proporciona por el decodificador para cualquier unidad de datos codificada antes de vídeo y/o después de vídeo. Para las unidades de datos de audio codificados con información de recorte extra incluida, ya sea el decodificador o la capa de incrustación, guiados por el decodificador de audio con información adicional, puede retirar muestras. Existen tres soluciones básicas para manejar correctamente el recorte:

1. la información de recorte se transmite desde el decodificador al sistema, que para el recorte inicial retrasa

todas las demás transmisiones paralelas para conservar la sincronización de audio-vídeo. El recorte al final no se aplica.

2. la información de recorte se transmite desde el decodificador al sistema junto con las unidades de datos descomprimidas, que pueden aplicarse entonces para retirar las muestras no válidas en un elemento de procesamiento de audio, por ejemplo, el elemento de reproducción.

3. la información de recorte se aplica dentro del decodificador y muestras no válidas se retiran desde el inicio o el final de una unidad de datos descomprimidos antes de que se proporcione al sistema. Esto resulta en unidades de datos descomprimidos con una duración más corta que la duración de trama común. Se recomienda que un sistema asuma un decodificador que aplica el recorte y los sellos de tiempo y la duración dentro del sistema deberá reflejar, por lo tanto, el recorte que se va a aplicar.

[0088] Para operaciones de decodificación de múltiples velocidades, la resolución de la operación de recorte deberá relacionarse con la frecuencia de muestreo original, que se codifica de forma característica como el componente de velocidad más alta. Varias resoluciones para la operación de recorte son imaginables, por ejemplo, una resolución fija en microsegundos, la frecuencia de muestreo de velocidad más baja o la frecuencia de muestreo de velocidad más alta. Para igualar la frecuencia de muestreo original, una forma de realización de la invención es proporcionar la resolución de la operación de recorte junto con los valores de recorte como una resolución a la medida. Por lo tanto, el formato de la información de recorte puede ser representado como una sintaxis como lo siguiente:

```
typedef struct trim {
    unsigned int resolution;
    unsigned short remove_from_begin;
    unsigned short remove_from_end;
```

```
};
```

[0089] Cabe destacar que la sintaxis presentada es solo un ejemplo de cómo puede estar contenida la información de recorte dentro de una unidad de datos de audio codificada. Otras variantes modificadas son abarcadas por la invención, asumiendo que permiten la distinción entre muestras válidas y no válidas.

[0090] Aunque algunos aspectos de la invención se describieron en el contexto de un aparato, cabe destacar que estos aspectos representan también una descripción del procedimiento correspondiente, es decir, un bloque o dispositivo corresponde a una etapa de procedimiento o una característica de una etapa de procedimiento. De forma análoga, aspectos descritos en el contexto de una etapa de procedimiento representan también una descripción de un bloque o elemento o característica correspondientes de un aparato correspondiente.

REIVINDICACIONES

1. Un procedimiento para recibir datos codificados que incluyen información respecto a la validez de datos y proporcionar datos de salida decodificados, comprendiendo el procedimiento:

- 5 recibir (602; 702) datos codificados con cualquiera de
información (924) a un nivel de trama que describe la cantidad de datos al inicio de una trama (920) que no es válida,
10 o información (924) a un nivel de trama que describe la cantidad de datos al final de la trama (920) que no es válida,
o información (924) a un nivel de trama que describe tanto la cantidad de datos al inicio como al final de la trama (920) que no es válida,
15 y recibir información en un nivel de trama que describe, si la trama correspondiente (920) es una Unidad de acceso antes de vídeo "(pre-roll)" o una Unidad de acceso después de vídeo "(post-roll)",
proporcionar (604; 704), mediante el uso de un decodificador de audio (1000; 1100), datos de salida decodificados que solo contienen las muestras no marcadas como no válidas, o que contienen todas las muestras de audio de la trama (920) y proporcionar (706) información a la aplicación sobre qué parte de los
20 datos es válida; y,
en caso de que la trama sea una Unidad de acceso antes de vídeo "(pre-roll)" o una Unidad de acceso después de vídeo "(post-roll)", proporcionar la Unidad de acceso antes de vídeo "(pre-roll)" o después de vídeo "(post-roll)" al decodificador de audio (1000; 1100) y descartar una salida correspondiente del decodificador de audio (1000; 1100) después de la decodificación por una capa de Sistemas.
25 donde el decodificador de audio es inicializado por la capa de sistemas con una estructura de configuración específica de audio, que da lugar a una configuración de salida del decodificador de audio a la capa de sistemas, que contiene información sobre una frecuencia de muestra, una configuración de canal, un tamaño de trama y un retardo adicional para las extensiones de códec señaladas explícitamente.

30 2. Un decodificador (1000; 1100) para recibir datos codificados y proporcionar datos de salida decodificados, siendo el decodificador (1000; 1100) un decodificador de audio, comprendiendo el decodificador:

- una entrada (1002; 1102) para recibir una serie de tramas codificadas (920) con una pluralidad de muestras de audio codificadas (922), donde algunas tramas (920) contienen información (924) respecto a la validez de datos,
35 de tal manera que se puedan recortar los datos no válidos causados por un retraso del codificador o el relleno de datos, siendo la información formateada como se describe en el procedimiento para recibir datos de audio codificados que incluyen información respecto a la validez de datos según la reivindicación 1,
una porción de decodificación (1004; 1104) acoplada a la entrada (1002; 1104) y configurada para aplicar la información (924) respecto a la validez de datos,
40 una salida (1006; 1106) para proporcionar muestras de audio decodificadas, donde ya sea solo se proporcionan las muestras de audio válidas,
o donde se proporciona la información respecto a la validez de las muestras de audio decodificadas; y
una interfaz a una capa de Sistemas configurada para proporcionar información a nivel de trama que describe, si la trama correspondiente es una Unidad de Acceso de antes de vídeo "(pre-roll)" o una Unidad de Acceso de
45 después de vídeo "(post-roll)", causando dicha información que la capa de Sistemas proporcione una Unidad de acceso de antes de vídeo "(pre-roll)" o una Unidad de acceso de después de vídeo "(post-roll)" al decodificador y descartar una salida correspondiente del decodificador después de la decodificación,
donde el decodificador de audio es inicializado por la capa de sistemas con una estructura de configuración específica de audio, que da lugar a una configuración de salida del decodificador de audio a la capa de sistemas,
50 que contiene información sobre una frecuencia de muestra, una configuración de canal, un tamaño de trama y un retardo adicional para las extensiones de códec señaladas explícitamente.

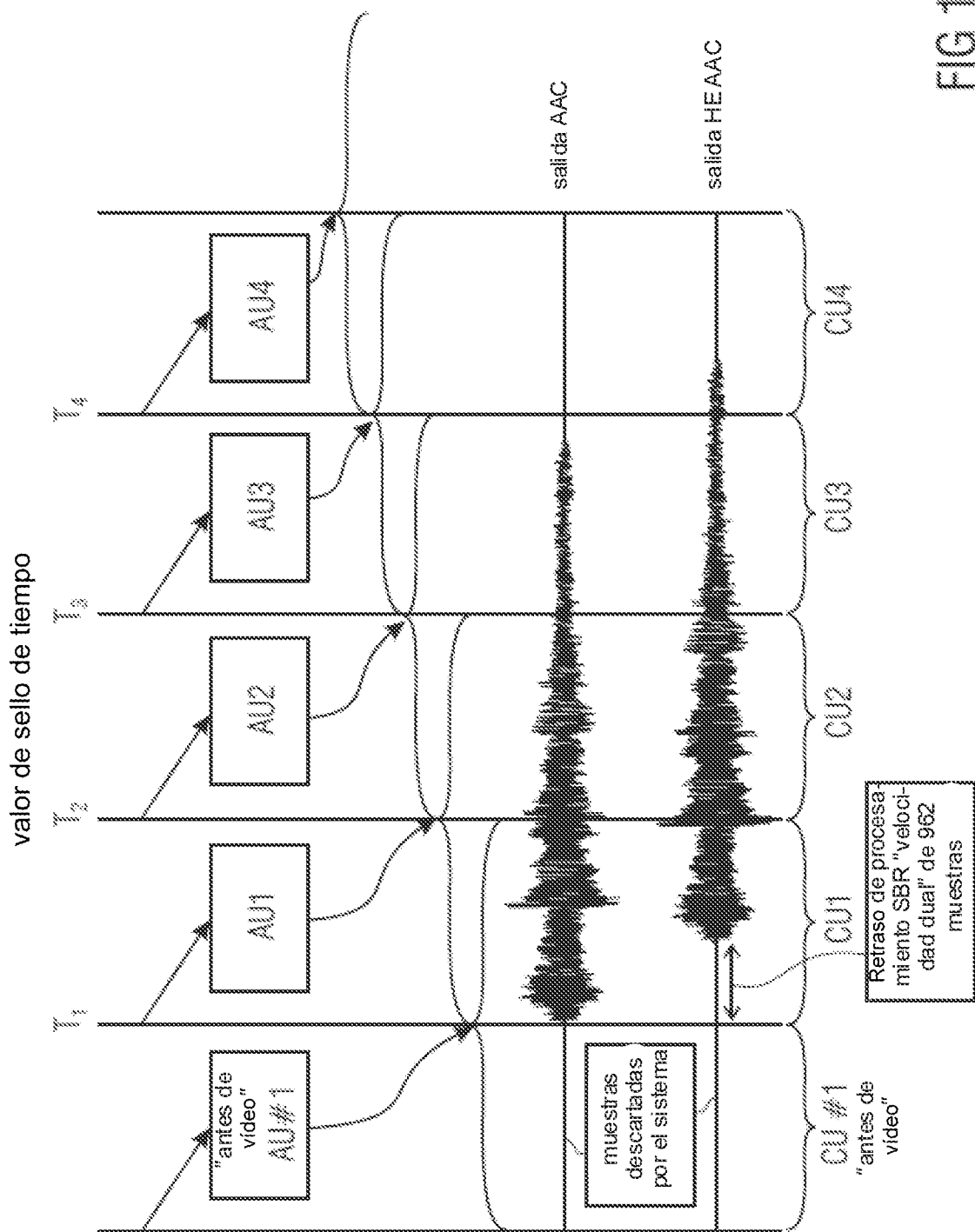


FIG 1

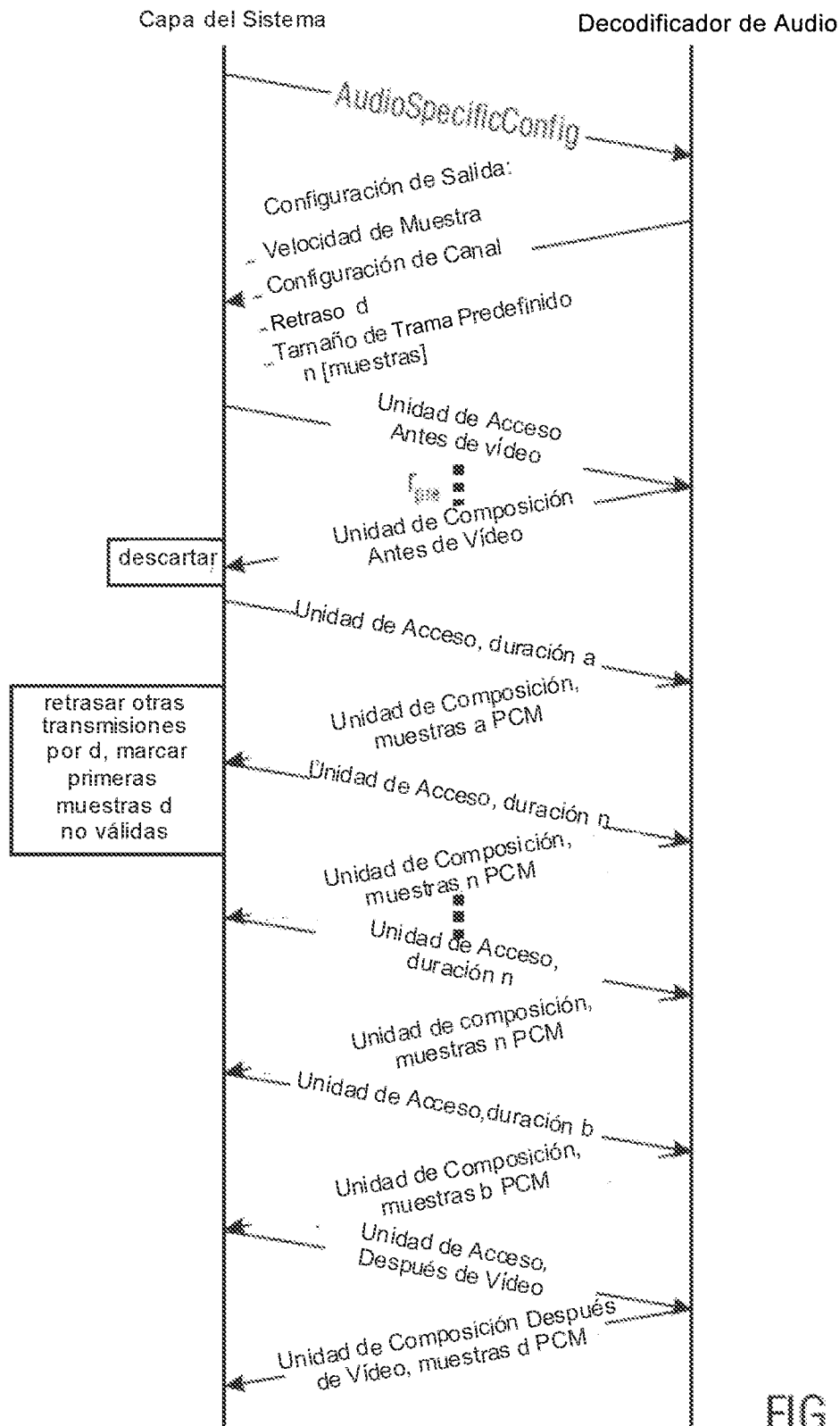


FIG 2

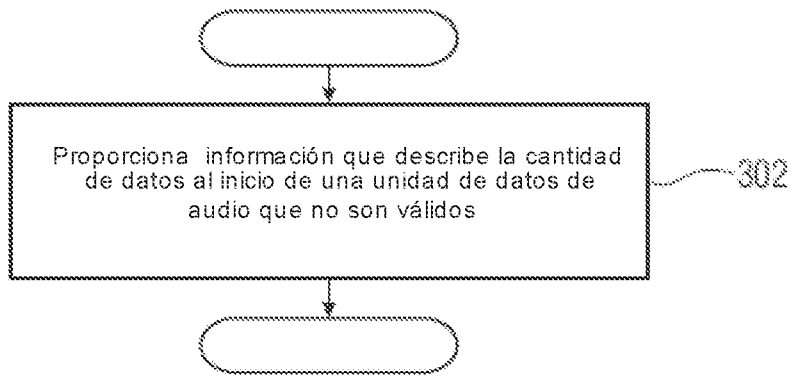


FIG 3

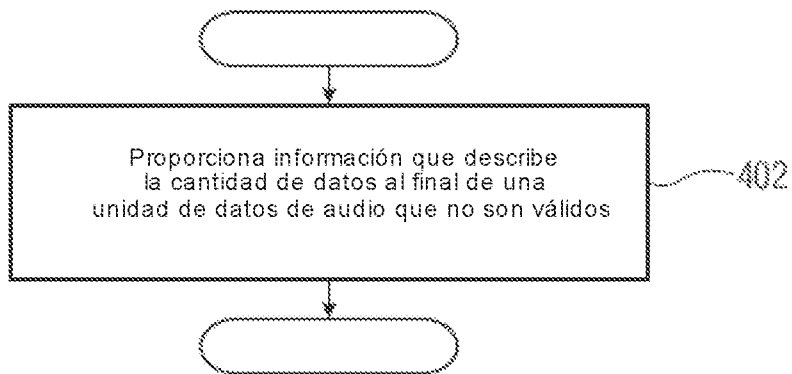


FIG 4

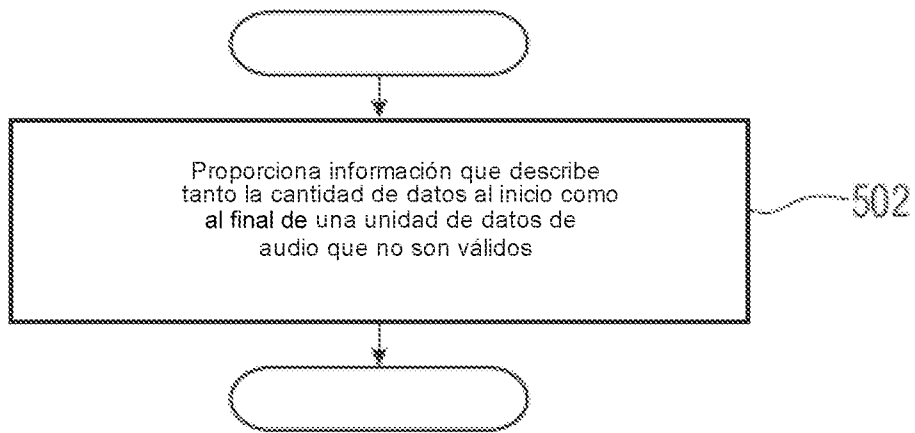


FIG 5

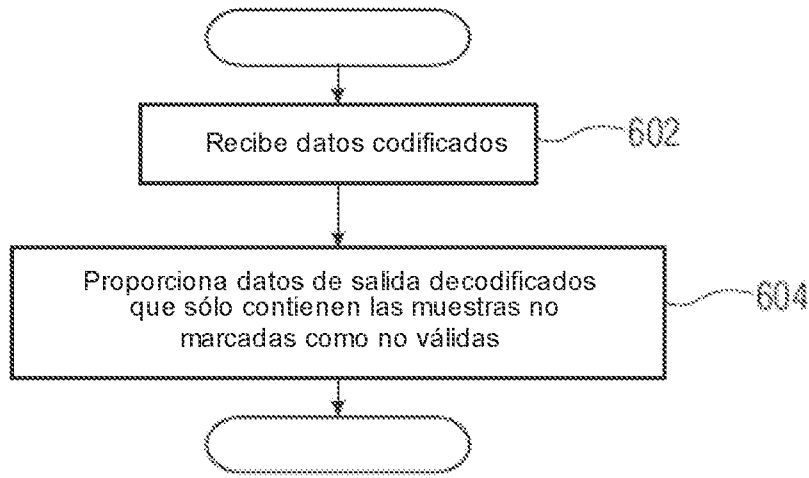


FIG 6

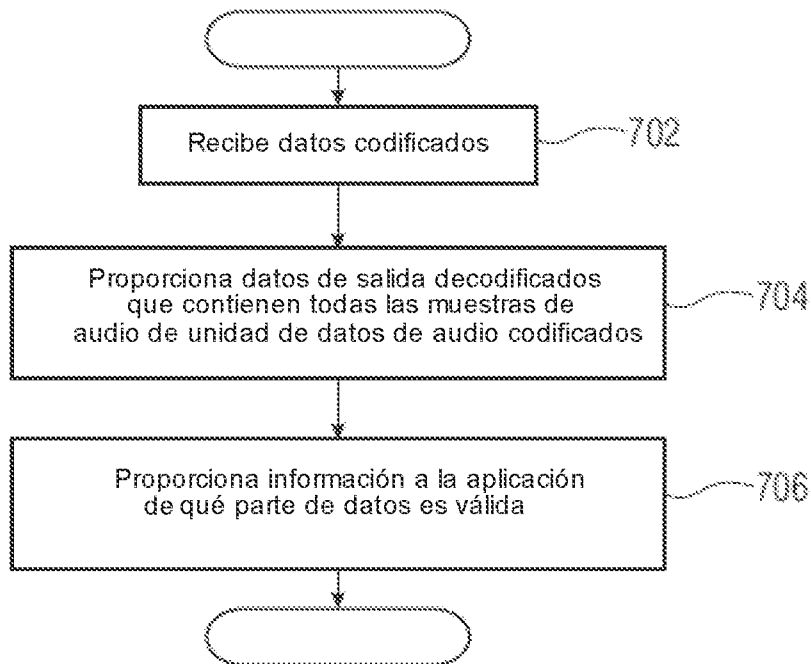


FIG 7

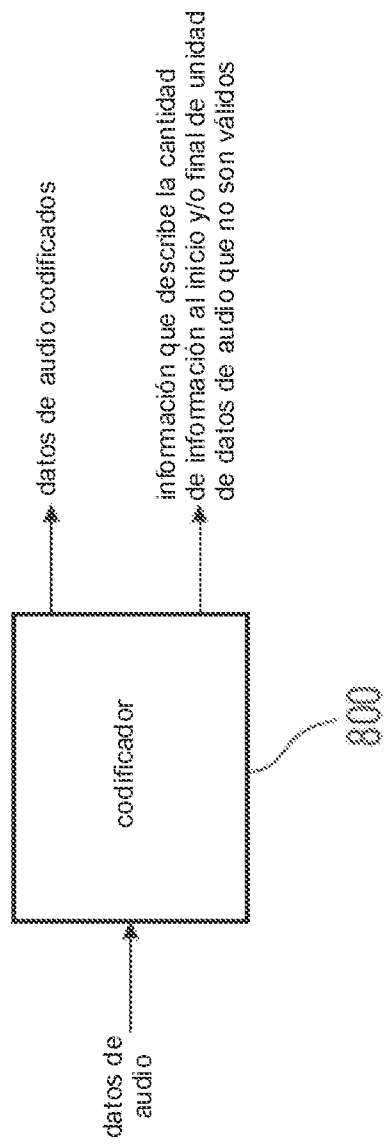


FIG 8

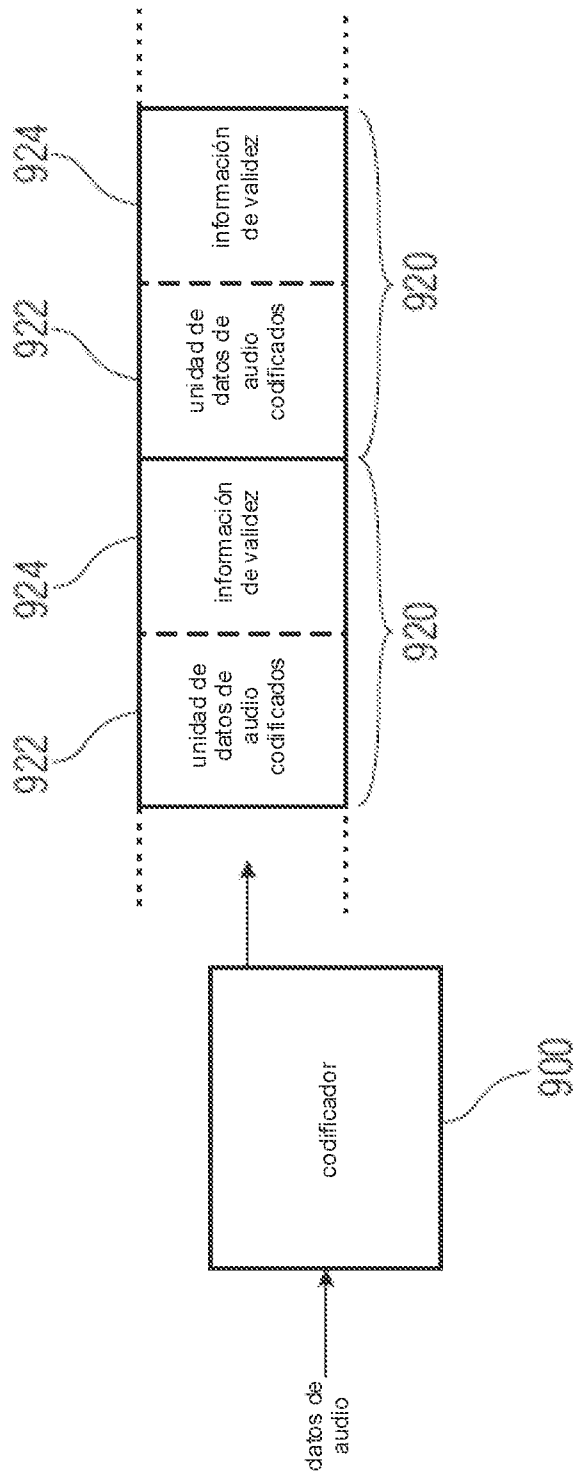


FIG 9

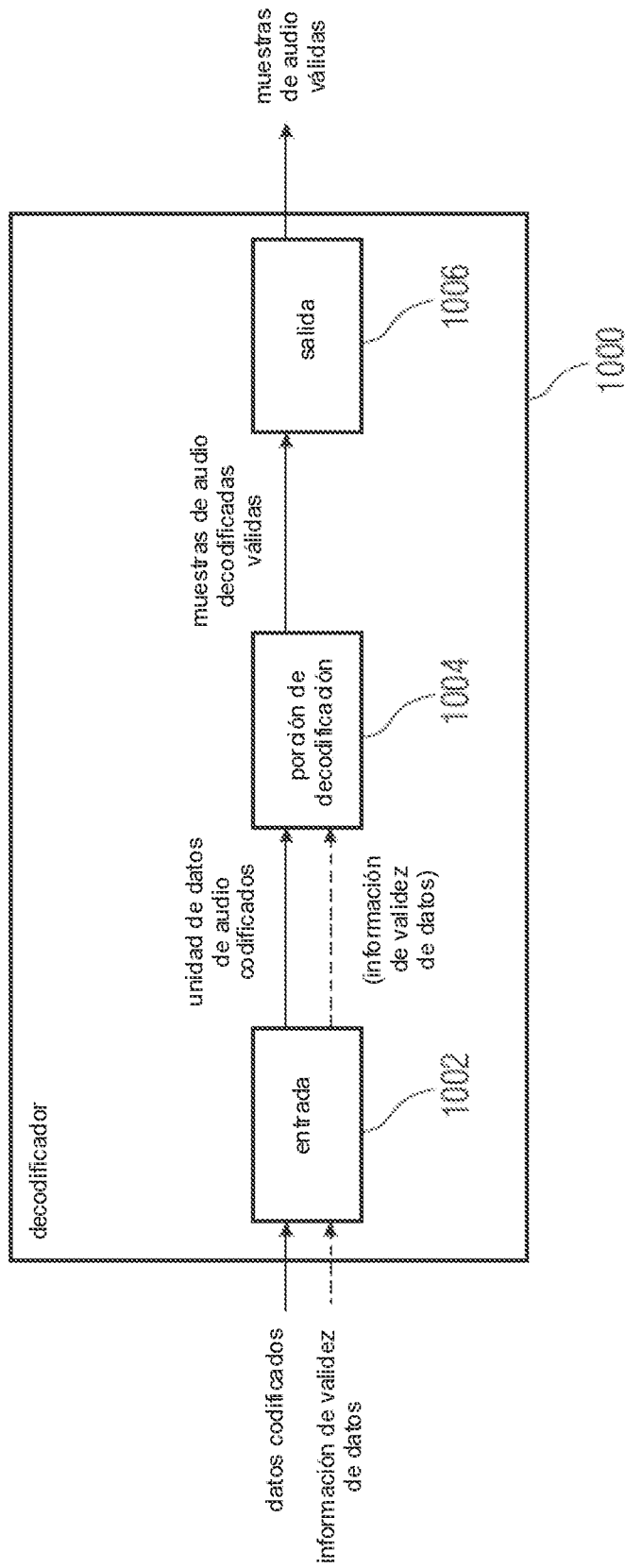


FIG 10

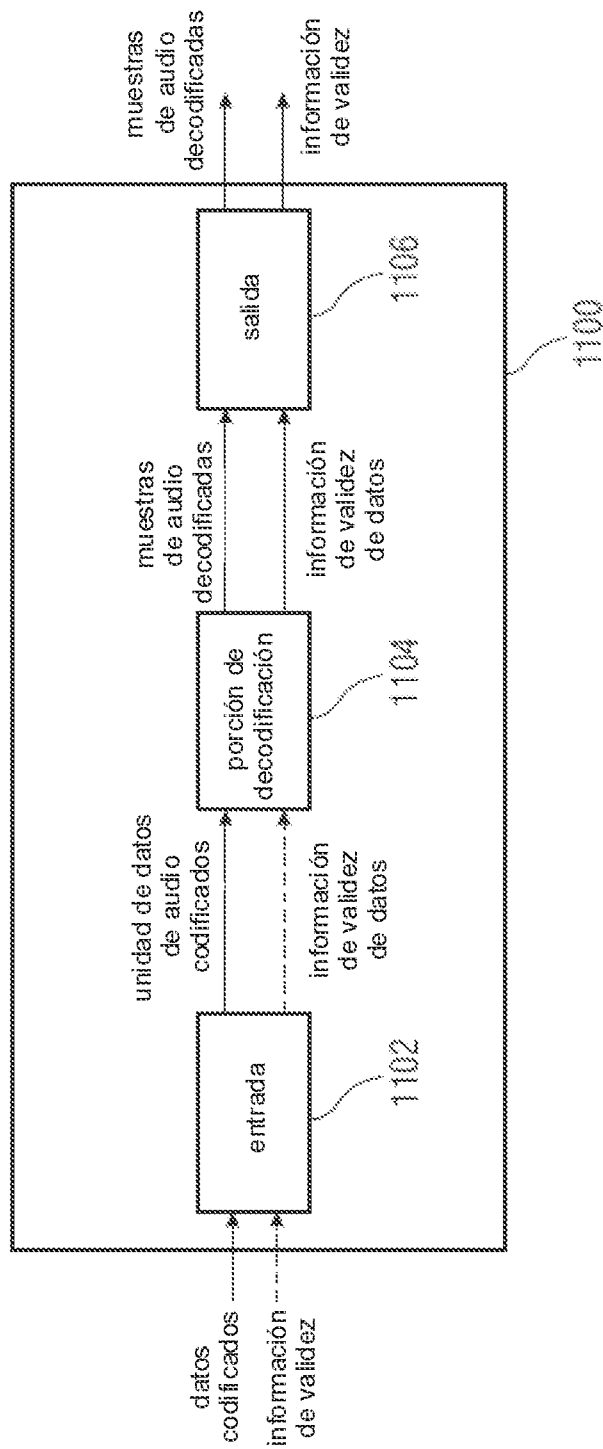


FIG 11

FIGURA PARA EL RESUMEN

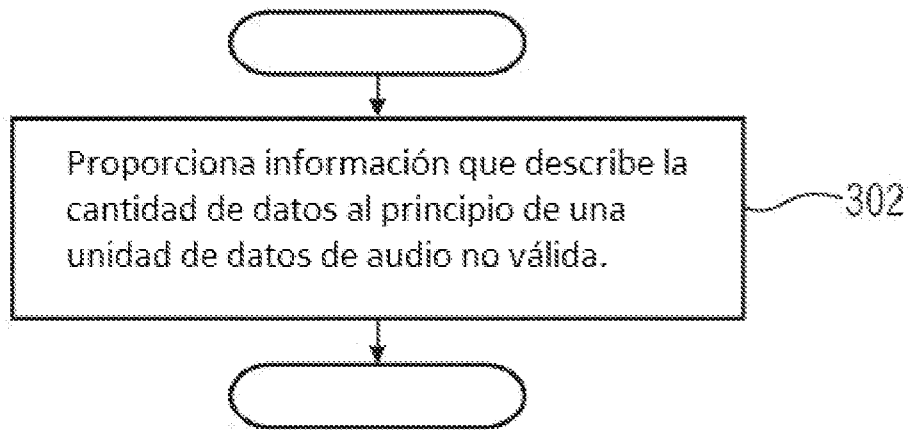


FIG 3