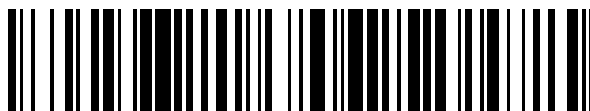


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 870 854**

51 Int. Cl.:

H04N 21/643 (2011.01)

H04N 21/6587 (2011.01)

H04N 21/845 (2011.01)

H04N 21/854 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **29.06.2011** **PCT/US2011/042447**

87 Fecha y número de publicación internacional: **05.01.2012** **WO12003237**

96 Fecha de presentación y número de la solicitud europea: **29.06.2011** **E 11730837 (9)**

97 Fecha y número de publicación de la concesión europea: **24.02.2021** **EP 2589222**

54 Título: **Señalización de muestras de vídeo para representaciones de vídeo en modo truco**

30 Prioridad:

20.10.2010 US 908593

21.07.2010 US 366436 P

13.07.2010 US 363884 P

29.06.2010 US 359748 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
27.10.2021

73 Titular/es:

QUALCOMM INCORPORATED (100.0%)

International IP Administration 5775 Morehouse Drive

San Diego, CA 92121-1714, US

72 Inventor/es:

CHEN, YING y

KARCZEWICZ, MARTA

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 870 854 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

Señalización de muestras de vídeo para representaciones de vídeo en modo truco

5 CAMPO TÉCNICO

[0001] Esta divulgación se refiere al transporte de datos de vídeo codificados.

ANTECEDENTES

10 **[0002]** Las capacidades de vídeo digital se pueden incorporar a una amplia gama de dispositivos, incluidos televisores digitales, sistemas de radiodifusión directa digital, sistemas de radiodifusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, cámaras digitales, dispositivos de grabación digitales, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o de radio por satélite, dispositivos de videoconferencia y similares. Los dispositivos de vídeo digitales implementan técnicas de compresión de vídeo, tales como las descritas en las normas definidas por MPEG-2, MPEG-4, ITU-T H.263 o ITU-T H.264/MPEG-4, parte 10, Codificación de Vídeo Avanzada (AVC) y ampliaciones de dichas normas, para transmitir y recibir información de vídeo digital más eficazmente.

20 **[0003]** Las técnicas de compresión de vídeo llevan a cabo una predicción espacial y/o una predicción temporal para reducir o eliminar la redundancia inherente a las secuencias de vídeo. Para la codificación de vídeo basada en bloques, una trama o un fragmento de vídeo se pueden dividir en macrobloques. Cada macrobloque se puede dividir aún más. Los macrobloques de una trama o un fragmento intracodificados (I) se codifican usando predicción espacial con respecto a macrobloques vecinos. Los macrobloques de una trama o un fragmento intercodificados (P o B) pueden
25 usar predicción espacial con respecto a macrobloques vecinos de la misma trama o fragmento, o predicción temporal con respecto a otras tramas de referencia.

[0004] Se han dedicado esfuerzos para desarrollar nuevas normas de codificación de vídeo basadas en la H.264/AVC. Una de estas normas es la norma de codificación de vídeo escalable (SVC), que es la extensión escalable para la H.264/AVC. Otra norma es la norma de codificación de vídeo multivista (MVC), que se ha convertido en la extensión multivista de H.264/AVC. Una versión de la norma AVC se describe en JVT-AD007, "Editors' draft revision to ITU-T Rec. H.264 | ISO/IEC 14496-10 Advanced Video Coding - in preparation for ITU-T SG 16 AAP Consent (in integrated form)", 30ª reunión de JVT, Ginebra, Suiza, febrero de 2009", disponible en http://wftp3.itu.int/av-arch/jvt-site/2009_01_Geneva/JVT-AD007.zip, que integra la SVC y la MVC integradas en la especificación AVC. El documento US 2006/037057 A1 describe un procedimiento para proporcionar comandos operativos a un programa AV transmitido usando el protocolo HTTP. El procedimiento incluye almacenar un programa AV en un servidor; generar metadatos en el servidor, donde los metadatos incluyen una tabla de mapeo que asocia intervalos de tiempo para el programa AV a intervalos de octetos para el programa AV; transmitir los metadatos y la tabla de mapeo a un cliente/reproductor asociado al servidor; generar y transmitir un comando GET de HTTP desde el cliente/reproductor al servidor en función de un comando operativo deseado; y seleccionar tramas I apropiadas en el servidor y transmitir las tramas I al cliente/reproductor en respuesta al comando GET de HTTP.

BREVE EXPLICACIÓN

45 **[0005]** La invención se define en las reivindicaciones adjuntas a las que ahora debe hacerse referencia. En general, esta divulgación describe técnicas relacionadas con el transporte de datos de medios usando un protocolo de transmisión continua, tal como la transmisión continua del protocolo de transferencia de hipertexto (HTTP). En particular, las técnicas de esta divulgación están dirigidas a señalar características de puntos de acceso aleatorio (RAP) en contenido de medios. Un punto de acceso aleatorio corresponde, en general, a un punto en el que un decodificador puede comenzar a decodificar apropiadamente una representación de vídeo sin referencia a datos anteriores, tales como unidades codificadas que preceden temporalmente al punto de acceso aleatorio. Es decir, los puntos de acceso aleatorio pueden comprender unidades decodificables independientes, por ejemplo, tramas o fragmentos intracodificadas de datos de vídeo.

55 **[0006]** La transmisión continua (por ejemplo, la transmisión continua de HTTP) implica, en general, la transferencia de datos desde un dispositivo de origen a un dispositivo de destino. De acuerdo con las técnicas de esta divulgación, un dispositivo de origen puede señalar características de tramas de actualización de decodificación abierta (ODR), además de tramas de actualización de decodificación instantánea (IDR). Al señalar estas características, el dispositivo de destino puede generar una solicitud para comenzar la reproducción de una representación de vídeo contenida en un flujo de datos de vídeo, en una trama ODR. De esta manera, un usuario del dispositivo de destino puede solicitar comenzar o saltar a una ubicación temporal particular de la representación de vídeo, y el dispositivo de destino puede solicitar la transmisión continua de una parte de la representación de vídeo, que contiene la imagen IDR u ODR más cercana a la ubicación temporal solicitada. Por lo tanto, en lugar de limitarse a iniciar la reproducción y/o transmisión continua de una representación de vídeo en una imagen IDR, un dispositivo de destino de acuerdo
65 con las técnicas de esta divulgación puede iniciar el flujo en una IDR u ODR, lo que puede reducir el consumo de

ancho de banda y/o habilitar puntos de acceso aleatorio y/o puntos de conmutación de vídeo más frecuentes para mejorar la experiencia de visualización del usuario.

[0007] Esta divulgación también proporciona técnicas para señalar características de imágenes intracodificadas, de tal manera que un dispositivo de destino puede recuperar y presentar las imágenes intracodificadas en modo truco. Un modo truco corresponde, en general, a un modo de reproducción de alta velocidad, por ejemplo, avance o rebobinado rápidos. En dichos modos, el dispositivo de destino solo puede reproducir un subconjunto de las tramas disponibles en una representación de vídeo. Por consiguiente, el dispositivo de destino solo necesita recuperar las tramas que se reproducirán en el modo truco. Por lo tanto, de acuerdo con las técnicas de esta divulgación, el dispositivo de destino puede generar solicitudes para las tramas intracodificadas en función de las características señalizadas, lo que puede reducir el consumo de ancho de banda y permitir la reproducción en modo truco cuando los medios se transfieren usando un protocolo de transmisión continua en red.

[0008] Por ejemplo, en el presente documento se describe un procedimiento. El procedimiento incluye recibir, en un dispositivo cliente desde un dispositivo servidor, una indicación de una trama de punto de acceso aleatorio (RAP) de actualización de decodificación abierta (ODR) de una segmentación de medios, donde al menos una trama de la segmentación de medios que sigue a la trama RAP ODR en orden de decodificación puede no estar correctamente decodificada, y donde cada trama de la segmentación de medios que sigue a la trama RAP ODR en orden de visualización puede decodificarse correctamente sin depender del contenido de las tramas antes del RAP ODR en orden de visualización. El procedimiento incluye además enviar, mediante el dispositivo cliente, una solicitud al dispositivo servidor para recuperar un flujo de datos de medios comenzando con la segmentación de medios que contiene la trama RAP ODR indicada.

[0009] De acuerdo con otro ejemplo, en el presente documento se describe un dispositivo. El dispositivo incluye medios para recibir, desde un dispositivo servidor, una indicación de una trama de punto de acceso aleatorio (RAP) de actualización de decodificación abierta (ODR) de una segmentación de medios, donde al menos una trama de la segmentación de medios que sigue a la trama RAP ODR en orden de decodificación puede no estar correctamente decodificada, y donde cada trama de la segmentación de medios que sigue a la trama RAP ODR en orden de visualización puede decodificarse correctamente sin depender del contenido de las tramas antes del RAP ODR en orden de visualización. El dispositivo incluye además medios para enviar una solicitud al dispositivo servidor para recuperar un flujo de datos de medios comenzando con la trama RAP ODR indicada.

[0010] De acuerdo con otro ejemplo, en el presente documento se describe un medio de almacenamiento legible por ordenador. El medio de almacenamiento incluye instrucciones que, cuando se ejecutan, hacen que un procesador de un dispositivo informático cliente reciba, en un dispositivo cliente desde un dispositivo servidor, una indicación de una trama de punto de acceso aleatorio (RAP) de actualización de decodificación abierta (ODR) de una segmentación de medios, donde al menos una trama de la segmentación de medios que sigue a la trama RAP ODR en orden de decodificación puede no estar correctamente decodificada, y donde cada trama de la segmentación de medios que sigue a la trama RAP ODR en orden de visualización puede decodificarse correctamente sin depender del contenido de las tramas antes del RAP ODR en orden de visualización. Las instrucciones hacen además que el dispositivo informático cliente envíe, mediante el dispositivo cliente, una solicitud al dispositivo servidor para recuperar un flujo de datos de medios comenzando con la segmentación de medios que contiene la trama RAP ODR indicada.

[0011] De acuerdo con otro ejemplo, en el presente documento se describe un dispositivo que incluye uno o más procesadores. El uno o más procesadores están configurados para recibir una indicación de una trama de punto de acceso aleatorio (RAP) de actualización de decodificación abierta (ODR) de una segmentación de medios, donde al menos una trama de la segmentación de medios que sigue a la trama RAP ODR en orden de decodificación puede no estar correctamente decodificada, y donde cada trama de la segmentación de medios que sigue a la trama RAP ODR en orden de visualización puede decodificarse correctamente sin depender del contenido de las tramas antes del RAP ODR en orden de visualización. El uno o más procesadores están configurados además para enviar una solicitud al dispositivo servidor para recuperar un flujo de datos de medios comenzando con la segmentación de medios que contiene la trama RAP ODR indicada.

[0012] De acuerdo con otro ejemplo, en el presente documento se describe un procedimiento. El procedimiento incluye recibir una secuencia de tramas de vídeo codificadas que comprenden una trama de punto de acceso aleatorio (RAP) de actualización de decodificación abierta (ODR) de una segmentación de medios, donde al menos una trama de la segmentación de medios que sigue a la trama RAP de ODR en orden de decodificación puede no estar correctamente decodificada, y donde cada trama de la segmentación de medios que sigue a la trama RAP de ODR en orden de visualización puede decodificarse correctamente sin depender del contenido de las tramas antes del RAP de ODR en orden de visualización. El procedimiento incluye además generar una indicación de una ubicación del RAP ODR dentro de la secuencia de tramas de vídeo codificadas. El procedimiento incluye además proporcionar la indicación.

[0013] De acuerdo con otro ejemplo, en el presente documento se describe un dispositivo. El dispositivo incluye medios para recibir una secuencia de tramas de vídeo codificadas que comprenden una trama de punto de acceso aleatorio (RAP) de actualización de decodificación abierta (ODR) de una segmentación de medios, donde al menos

una trama de la segmentación de medios que sigue a la trama RAP de ODR en orden de decodificación puede no estar correctamente decodificada, y donde cada trama de la segmentación de medios que sigue a la trama RAP de ODR en orden de visualización puede decodificarse correctamente sin depender del contenido de las tramas antes del RAP de ODR en orden de visualización. El dispositivo incluye además medios para generar una indicación de una ubicación del RAP ODR dentro de la secuencia de tramas de vídeo codificadas. El dispositivo incluye además medios para proporcionar la indicación.

[0014] De acuerdo con otro ejemplo, en el presente documento se describe un medio de almacenamiento legible por ordenador. El medio de almacenamiento legible por ordenador incluye instrucciones que, cuando se ejecutan, hacen que uno o más procesadores de un dispositivo de origen reciban una secuencia de tramas de vídeo codificadas que comprenden una trama de punto de acceso aleatorio (RAP) de actualización de decodificación abierta (ODR) de una segmentación de medios, donde al menos una trama de la segmentación de medios que sigue a la trama RAP de ODR en orden de decodificación puede no estar correctamente decodificada, y donde cada trama de la segmentación de medios que sigue a la trama RAP de ODR en orden de visualización puede decodificarse correctamente sin depender del contenido de las tramas antes del RAP de ODR en orden de visualización. Las instrucciones hacen además que el dispositivo de origen genere una indicación de una ubicación del RAP ODR dentro de la secuencia de tramas de vídeo codificadas. Las instrucciones hacen además que el dispositivo de origen proporcione la indicación.

[0015] De acuerdo con otro ejemplo, en el presente documento se describe un dispositivo que incluye uno o más procesadores. El uno o más procesadores están configurados para recibir una secuencia de tramas de vídeo codificadas que comprenden una trama de punto de acceso aleatorio (RAP) de actualización de decodificación abierta (ODR) de una segmentación de medios, donde al menos una trama de la segmentación de medios que sigue a la trama RAP ODR en orden de decodificación puede no estar correctamente decodificada, y donde cada trama de la segmentación de medios que sigue a la trama RAP ODR en orden de visualización puede estar correctamente decodificada sin depender del contenido de las tramas antes del RAP ODR en orden de visualización. El uno o más procesadores están configurados además para generar una indicación de una ubicación del RAP ODR dentro de la secuencia de tramas de vídeo codificadas. El uno o más procesadores están configurados además para proporcionar la indicación.

[0016] De acuerdo con otro ejemplo, en el presente documento se describe un procedimiento. El procedimiento incluye recibir, en un dispositivo cliente desde un dispositivo servidor, una indicación de al menos un intervalo de octetos dentro de una segmentación de medios que representa una presentación de vídeo configurada para mostrarse a una primera velocidad de reproducción en la que al menos una trama intracodificada (trama I) se produce dentro del al menos un intervalo de octetos. El procedimiento incluye además enviar, mediante el dispositivo cliente, una solicitud al dispositivo servidor para recuperar datos de medios dentro del al menos un intervalo de octetos. El procedimiento incluye además usar los datos de medios recuperados dentro del al menos un intervalo de octetos para presentar una versión de alta velocidad de la representación de vídeo que incluye la al menos una trama I a una segunda velocidad de reproducción mayor que la primera velocidad de reproducción.

[0017] De acuerdo con otro ejemplo, en el presente documento se describe un dispositivo. El dispositivo incluye medios para recibir, en un dispositivo cliente desde un dispositivo servidor, una indicación de al menos un intervalo de octetos dentro de una segmentación de medios que representa una presentación de vídeo configurada para mostrarse a una primera velocidad de reproducción en la que al menos una trama intracodificada (trama I) se produce dentro del al menos un intervalo de octetos. El dispositivo incluye además medios para enviar, mediante el dispositivo cliente, una solicitud al dispositivo servidor para recuperar datos de medios dentro del al menos un intervalo de octetos. El dispositivo incluye además medios para usar los datos de medios recuperados dentro del al menos un intervalo de octetos para presentar una versión de alta velocidad de la representación de vídeo que incluye la al menos una trama I a una segunda velocidad de reproducción mayor que la primera velocidad de reproducción.

[0018] De acuerdo con otro ejemplo, en el presente documento se describe un medio de almacenamiento legible por ordenador. El medio de almacenamiento legible por ordenador almacena instrucciones configuradas para hacer que un dispositivo informático reciba, en un dispositivo cliente desde un dispositivo servidor, una indicación de al menos un intervalo de octetos dentro de una segmentación de medios que representa una presentación de vídeo configurada para mostrarse a una primera velocidad de reproducción en la que al menos una trama intracodificada (trama I) se produce dentro del al menos un intervalo de octetos. Las instrucciones están configuradas además para hacer que el dispositivo informático envíe, mediante el dispositivo cliente, una solicitud al dispositivo servidor para recuperar datos de medios dentro del al menos un intervalo de octetos. Las instrucciones están configuradas además para hacer que el dispositivo informático use los datos de medios recuperados dentro del al menos un intervalo de octetos para presentar una versión de alta velocidad de la representación de vídeo que incluye la al menos una trama I a una segunda velocidad de reproducción mayor que la primera velocidad de reproducción.

[0019] De acuerdo con otro ejemplo, en el presente documento se describe un dispositivo. El dispositivo incluye al menos un procesador. El al menos un procesador está configurado para recibir una indicación de al menos un intervalo de octetos dentro de una segmentación de medios que representa una presentación de vídeo configurada para mostrarse a una primera velocidad de reproducción en la que al menos una trama intracodificada (trama I) se produce dentro del al menos un intervalo de octetos. El al menos un procesador está configurado además para enviar una

solicitud al dispositivo servidor para recuperar datos de medios dentro del al menos un intervalo de octetos. El al menos un procesador está configurado además para usar los datos de medios recuperados dentro del al menos un intervalo de octetos para presentar una versión de alta velocidad de la representación de vídeo que incluye la al menos una trama I a una segunda velocidad de reproducción mayor que la primera velocidad de reproducción.

[0020] De acuerdo con otro ejemplo, en el presente documento se describe un procedimiento. El procedimiento incluye recibir una secuencia de tramas de vídeo codificadas que comprenden al menos una trama intercodificada (trama I). El procedimiento incluye además generar una indicación de un intervalo de octetos dentro de la secuencia de tramas de vídeo codificadas, donde al menos una trama I se produce dentro del al menos un intervalo de octetos. El procedimiento incluye además proporcionar la indicación de un intervalo de octetos dentro de la secuencia de tramas de vídeo codificadas.

[0021] De acuerdo con otro ejemplo, en el presente documento se describe un dispositivo. El dispositivo incluye medios para recibir una secuencia de tramas de vídeo codificadas que comprenden al menos una trama intercodificada (trama I). El dispositivo incluye además medios para generar una indicación de un intervalo de octetos dentro de la secuencia de tramas de vídeo codificadas, donde al menos una trama I se produce dentro del al menos un intervalo de octetos. El dispositivo incluye además medios para proporcionar la indicación de un intervalo de octetos dentro de la secuencia de tramas de vídeo codificadas.

[0022] De acuerdo con otro ejemplo, en el presente documento se describe un medio de almacenamiento legible por ordenador. El medio de almacenamiento legible por ordenador almacena instrucciones configuradas para hacer que un dispositivo informático reciba una secuencia de tramas de vídeo codificadas que comprenden al menos una trama I. Las instrucciones además hacen que el dispositivo informático genere una indicación de un intervalo de octetos, donde al menos una trama I se produce dentro del al menos un intervalo de octetos. Las instrucciones hacen además que el dispositivo informático proporcione la indicación de un intervalo de octetos dentro de la secuencia de tramas de vídeo codificadas.

[0023] De acuerdo con otro ejemplo, en el presente documento se describe un dispositivo. El dispositivo incluye uno o más procesadores. El uno o más procesadores están configurados para recibir una secuencia de tramas de vídeo codificadas que comprenden al menos una trama I. El uno o más procesadores están configurados además para generar una indicación de un intervalo de octetos dentro de la secuencia de tramas de vídeo codificadas, donde al menos una trama I se produce dentro del al menos un intervalo de octetos. El uno o más procesadores están configurados además para proporcionar la indicación de un intervalo de octetos dentro de la secuencia de tramas de vídeo codificadas.

[0024] Los detalles de uno o más ejemplos se exponen en los dibujos adjuntos y en la descripción siguiente. Otros rasgos característicos, objetivos y ventajas resultarán evidentes a partir de la descripción y los dibujos, y a partir de las reivindicaciones.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0025]

La FIG. 1 es un diagrama de bloques que ilustra un sistema de ejemplo en el que un dispositivo de origen de audio/vídeo (A/V) transporta datos de audio y vídeo a un dispositivo de destino de A/V.

La FIG. 2 es un diagrama conceptual que ilustra un ejemplo de un fragmento de vídeo de conformidad con las técnicas de esta divulgación.

La FIG. 3 es un diagrama conceptual que ilustra un ejemplo de un fragmento de vídeo de conformidad con las técnicas de esta divulgación.

La FIG. 4 es un diagrama de bloques que ilustra un ejemplo de un archivo de segmentación de conformidad con las técnicas de esta divulgación.

La FIG. 5 es un diagrama de bloques que ilustra un ejemplo de una presentación de una película de conformidad con las técnicas de esta divulgación.

La FIG. 6 es un diagrama de bloques que ilustra un ejemplo de una caja de índices de segmento de conformidad con las técnicas de esta divulgación.

La FIG. 7 es un diagrama de flujo que ilustra un ejemplo de un procedimiento para hacer funcionar un dispositivo cliente para recibir una indicación de un punto de acceso aleatorio (RAP) de actualización de decodificación abierta (ODR) de conformidad con las técnicas de esta divulgación.

La FIG. 8 es un diagrama de flujo que ilustra un ejemplo de un procedimiento para hacer funcionar un dispositivo de origen para proporcionar una indicación de un RAP ODR de conformidad con las técnicas de esta divulgación.

5 La FIG. 9 es un diagrama de flujo que ilustra un ejemplo de un procedimiento para hacer funcionar un dispositivo cliente para recibir una indicación de un intervalo de octetos de una trama intracodificada (trama I) de conformidad con las técnicas de esta divulgación.

10 La FIG. 10 es un diagrama de flujo que ilustra un ejemplo de un procedimiento para hacer funcionar un dispositivo de origen para proporcionar una indicación de un intervalo de octetos de una trama intracodificada (trama I) de conformidad con las técnicas de esta divulgación.

La FIG. 11 es un diagrama de flujo que ilustra un ejemplo de etapas que pueden realizar respectivos dispositivos cliente y servidor HTTP de conformidad con las técnicas de esta divulgación.

15 La FIG. 12 es un diagrama de flujo que ilustra un ejemplo de etapas que pueden realizar respectivos dispositivos cliente y servidor HTTP de conformidad con las técnicas de esta divulgación.

DESCRIPCIÓN DETALLADA

20 **[0026]** Las técnicas de esta divulgación están dirigidas, en general, a admitir el transporte de datos de vídeo usando un protocolo de transmisión continua, tal como, por ejemplo, la transmisión continua en el protocolo de transferencia de hipertexto (HTTP). En general, las referencias a HTTP pueden incluir referencias a la transmisión continua de HTTP en esta divulgación.

25 **[0027]** Esta divulgación se refiere a dos tipos específicos de puntos de acceso aleatorio (PAR) de una representación de vídeo. En particular, esta divulgación describe RAP de actualización de decodificación instantánea (IDR) y RAP de actualización de decodificación abierta (ODR). Un punto de acceso IDR es una intratrama decodificable de forma independiente de un GOP cerrado. Un GOP cerrado es una agrupación de tramas de una representación de vídeo que no incluye ninguna trama que dependa de tramas anteriores al GOP, ya sea en el orden de decodificación o de presentación, para que pueda decodificarse correctamente.

30 **[0028]** Un GOP abierto es un GOP en el que las imágenes anteriores a una trama inicial decodificable de forma independiente (trama I) del GOP en orden de visualización (salida), pero posteriores a la trama I en orden de decodificación, pueden no decodificarse correctamente. Por ejemplo, un GOP abierto puede incluir una o más tramas que dependen, al menos en parte, del contenido de al menos una trama fuera del GOP abierto. En algunos ejemplos de un GOP abierto, se puede usar una trama I de un GOP abierto como un punto de acceso aleatorio para el comienzo de la reproducción de una representación de vídeo. Por ejemplo, si todas las tramas que siguen a una trama I inicial en orden de visualización pueden decodificarse correctamente, la trama I puede usarse como un RAP. Dicho RAP puede denominarse RAP ODR.

35 **[0029]** Las tramas de una representación de vídeo pueden disponerse en uno o más GOP. El uno o más GOP se pueden organizar en uno o más fragmentos de vídeo. El uno o más fragmentos pueden disponerse en uno o más segmentos. En algunos ejemplos, una pluralidad de segmentos puede formar una representación de vídeo.

40 **[0030]** Los servicios de transmisión continua, tales como los servicios de transmisión continua de HTTP, pueden permitir que un dispositivo cliente realice una solicitud HTTP para recuperar uno o más GOP, fragmentos, segmentos o una representación de vídeo completa desde un dispositivo servidor. El dispositivo servidor puede contener múltiples representaciones de contenido de medios. En un ejemplo, dicha solicitud puede incluir una solicitud GET de HTTP. Una solicitud GET de HTTP puede proporcionar a un dispositivo cliente acceso a contenido almacenado en un dispositivo servidor al especificar un localizador de recursos uniforme (URL) de una o más segmentaciones de medios de una representación de vídeo (por ejemplo, GOP, fragmento, segmento o archivos de medios a nivel de representación de vídeo). En algunos ejemplos, una solicitud GET de HTTP puede especificar un intervalo de octetos dentro de un archivo de medios en lugar de solicitar un archivo de medios completo. Dicha solicitud GET de HTTP puede denominarse solicitud GET parcial. Como se describe en el presente documento, la expresión "segmentación de medios" se refiere a cualquier archivo, ya sea un fragmento, segmento o segmentación de medios a nivel de representación de vídeo, que se puede recuperar de acuerdo con una solicitud de protocolo de transmisión continua, tal como una solicitud GET o GET parcial de HTTP.

45 **[0031]** En algunos ejemplos, las diversas formas de contenido de medios descritas en el presente documento pueden almacenarse en un servidor en forma de una o más segmentaciones de medios. En algunos ejemplos, un dispositivo servidor se puede configurar para proporcionar información de señalización para facilitar el servicio de transmisión continua de HTTP. De acuerdo con un aspecto de esta divulgación, un dispositivo servidor puede señalar un punto de acceso aleatorio (RAP) de actualización de decodificación abierta (ODR) para una segmentación de medios (por ejemplo, fragmento, segmento, segmentación de medios específica de representación) de una representación de vídeo. De acuerdo con las técnicas de esta divulgación, un dispositivo cliente puede comunicar una solicitud a un dispositivo servidor referente a la transmisión continua de datos de medios en base a la señalización del RAP ODR.

Mediante la señalización de RAP ODR para una representación de vídeo, un dispositivo cliente puede seleccionar entre más posiciones temporales de una representación de vídeo para el comienzo de una transmisión continua, decodificación y/o reproducción de la representación de vídeo.

[0032] En algunos ejemplos, el dispositivo de origen puede señalar si una segmentación de medios contiene, o no, un RAP ODR en caso de que IDR no está contenida en un fragmento o segmento. En consecuencia, la reproducción de una representación de vídeo puede comenzar en una ubicación temporal asociada a una segmentación de medios (por ejemplo, un fragmento o segmento de una representación de vídeo, o una representación de vídeo completa) que incluye un RAP ODR, si no hay segmentación de medios que incluya un RAP IDR disponible cerca de la ubicación temporal deseada.

[0033] En particular, el dispositivo de origen puede señalar un desplazamiento de octetos a un RAP ODR de una segmentación de medios. La segmentación de medios se puede direccionar de forma independiente mediante un localizador de recursos uniforme (URL) o un nombre de recurso uniforme (URN) específico para el fragmento de vídeo. Al señalar el desplazamiento de octetos al RAP ODR en el fragmento de vídeo, el dispositivo cliente puede generar una solicitud GET que comprende un intervalo de octetos que comienza en el RAP ODR y termina al final de la segmentación de medios, en un ejemplo.

[0034] De acuerdo con otro aspecto de esta divulgación, un dispositivo servidor puede señalar un intervalo de octetos para al menos una trama I de una representación de vídeo. Un dispositivo cliente puede comunicar una solicitud a un dispositivo servidor para recuperar la al menos una trama I basándose en el intervalo de octetos señalado, y usar la trama I recuperada para proporcionar una versión de alta velocidad de un contenido de vídeo que incluye la al menos una trama I. Una versión de alta velocidad de un contenido de vídeo puede admitir un modo truco de una representación de vídeo, tal como un avance rápido o un rebobinado rápido. Al señalar directamente un intervalo de octetos para al menos una trama I de una presentación de vídeo, se puede proporcionar una versión de alta velocidad sin solicitar una representación de vídeo completa a la que pertenecen las tramas I, o sin solicitar una versión alternativa de una representación de vídeo creada específicamente para la reproducción a alta velocidad. Además, estas técnicas permiten proporcionar una versión de alta velocidad de una representación de vídeo, al tiempo que minimizan la cantidad de datos que se pueden transmitir de forma continua, decodificar y/o reproducir.

[0035] De acuerdo con este aspecto de la divulgación, un dispositivo servidor puede señalar una indicación de uno o más intervalos de octetos dentro de una o más segmentaciones de medios que representan el contenido de medios. Al menos una trama intracodificada (trama I) puede producirse dentro de cada uno de los uno o más intervalos de octetos señalizados. Un dispositivo cliente puede, basándose en la indicación señalizada del uno o más intervalos de octetos, enviar una solicitud (por ejemplo, una solicitud GET o GET parcial de HTTP) para recuperar datos de medios dentro de los intervalos de octetos. El dispositivo cliente puede presentar consecutivamente dos o más tramas I que se producen con al menos uno de los uno o más intervalos de octetos señalizados que están separados temporalmente por al menos una trama intracodificada, para presentar la presentación de vídeo a una velocidad relativamente alta. Dicha representación puede considerarse una versión de alta velocidad del contenido de medios.

[0036] El Proyecto de Colaboración de Tercera Generación (3GPP) es una colaboración entre grupos de asociaciones de telecomunicaciones para realizar una especificación del sistema de telefonía móvil de tercera generación (3G) aplicable a nivel mundial en el ámbito del proyecto de Telecomunicaciones Móviles Internacionales-2000 de la Unión Internacional de Telecomunicaciones (UIT). Las especificaciones 3GPP se basan en especificaciones evolucionadas del Sistema Global para Comunicaciones Móviles (GSM). La normalización 3GPP abarca la arquitectura de radio, red central y servicio. En algunos ejemplos, las técnicas de esta divulgación se describen en el presente documento usando ejemplos y términos de conformidad con la especificación del protocolo 3GPP. Estos ejemplos y términos se proporcionan simplemente con fines ejemplares; también se contemplan otras especificaciones de protocolo para la transmisión continua de datos de vídeo y de conformidad con esta divulgación.

[0037] En algunos ejemplos de acuerdo con la transmisión continua de HTTP, los archivos de segmentación pueden disponerse en varias representaciones diferentes que corresponden cada una a un programa particular (por ejemplo, película, programa de televisión u otro programa de vídeo). Cada archivo 3GPP puede corresponder a estas diferentes representaciones de un programa y/o segmentos o fragmentos de una presentación de un programa.

[0038] Una representación de vídeo (por ejemplo, un archivo 3GPP) puede estar dispuesto en una pluralidad de fragmentos de película. Cada fragmento de película puede incluir una pluralidad de tramas, o imágenes, de una representación de vídeo. Los fragmentos de una representación de vídeo 3GPP pueden recuperarse individualmente usando una solicitud GET o GET parcial de HTTP que puede hacer referencia a un localizador de recursos uniforme (URL) o un nombre de recurso uniforme (URN) de un archivo 3GPP.

[0039] Un cliente HTTP puede cambiar representaciones durante la reproducción, por ejemplo, cuando cambian las condiciones de red, o cuando se detecta una entrada de control de usuario. Por ejemplo, cuando hay disponible una gran cantidad de ancho de banda, el cliente puede recuperar archivos 3GPP de una presentación de calidad relativamente más alta, mientras que cuando hay disponible una cantidad menor de ancho de banda, el cliente puede recuperar archivos 3GPP de una presentación de calidad relativamente menor. De manera similar, un usuario puede

indicar un cambio deseado en la calidad de vídeo. Un usuario también puede indicarle a un cliente que pause, detenga, busque en (por ejemplo, avance rápido, rebobinado) o explore (indique un punto de inicio deseado) la reproducción de una representación de vídeo.

[0040] Puede ser deseable que un servidor HTTP proporcione un subconjunto de la información disponible para uno o más presentaciones de película para permitir que un cliente HTTP seleccione diferentes representaciones, o porciones de representaciones, para la transmisión continua. Dicha información puede denominarse información de cabecera. La información de cabecera puede ser específica de un programa, específica de una representación del programa o de una parte de la representación, o de un fragmento o segmento de película de una representación.

[0041] En un ejemplo de acuerdo con 3GPP, la información de cabecera puede estar contenida en un archivo de descripción de presentación de medios (MPD). El archivo MPD puede señalar elementos característicos de una pluralidad de representaciones de datos de vídeo tales como, por ejemplo, velocidad de bits, velocidad de tramas, resolución, tipo de exploración entrelazada o progresiva, tipo de codificación (por ejemplo, MPEG-1, MPEG-2, H.263, MPEG-4/H.264 o cualquier otro códec de vídeo), puntos de acceso u otras características.

[0042] Los fragmentos de película pueden incluir una pluralidad de grupos de imágenes (GOP). Un GOP puede incluir, o no, puntos de acceso aleatorio (RAP) para el inicio de la decodificación y/o reproducción.

[0043] Esta divulgación describe técnicas para la señalización de puntos de acceso ODR en el contexto de aplicaciones de transmisión continua de HTTP. Las técnicas de esta divulgación pueden proporcionar una o más ventajas. Por ejemplo, la señalización de puntos de acceso ODR además de los puntos de acceso IDR puede permitir mejoras en la funcionalidad de exploración o búsqueda, al proporcionar una granularidad de control más fina para iniciar flujos de acuerdo con la transmisión continua de HTTP, sobre técnicas conocidas que proporcionan la señalización de puntos de acceso IDR. Las técnicas de esta divulgación también pueden proporcionar ventajas durante la transmisión continua de programas de vídeo más complejos que incluyen un alto nivel de dependencia de contenido entre tramas para la decodificación, ya que se pueden proporcionar más puntos de acceso, mejorándose así la funcionalidad de exploración y/o búsqueda para la transmisión continua de HTTP. Las técnicas de esta divulgación pueden proporcionar ventajas adicionales, porque al señalar uno o más puntos de acceso ODR en la información de cabecera de un archivo de segmentación, un dispositivo cliente puede solicitar un flujo de datos de medios comenzando en un punto de acceso ODR o en un GOP abierto que incluya el punto de acceso ODR, lo que proporciona una granularidad más fina en la selección de un punto de inicio para la transmisión continua/decodificación/reproducción de vídeo representado por un archivo de segmentación. Estas técnicas también pueden ser beneficiosas con respecto a técnicas anteriores al reducir la cantidad de datos a transmitir de manera continua para la reproducción de una representación de vídeo.

[0044] En algunos ejemplos, las técnicas de esta divulgación proporcionan señalización de puntos de acceso ODR y/u otra información en información de cabecera de un archivo de segmentación. En un ejemplo, los puntos de acceso ODR pueden señalizarse en una caja de índices de segmento (SIDX) de una representación de película o segmento. El archivo de segmentación puede formatearse de acuerdo con el formato de archivo de medios base ISO, o cualquier otro tipo de archivo de segmentación, incluidas extensiones del formato de archivo de medios base ISO tal como, por ejemplo, codificación de vídeo escalable (SVC), codificación de vídeo avanzada (AVC) y/o codificación de vídeo multivista (MVC) y similares. Las técnicas de esta divulgación pueden proporcionar además señalización de otra información relacionada con la funcionalidad de exploración y/o búsqueda, que se describen posteriormente con más detalle.

[0045] Las técnicas de esta divulgación proporcionan además señalización mediante un dispositivo servidor HTTP de un intervalo de octetos de una o más tramas intracodificadas (tramas I) para los propósitos de funcionalidad de exploración. Por ejemplo, un dispositivo de destino puede mostrar una versión de alta velocidad de una representación de vídeo con el propósito de avanzar o rebobinar rápidamente, lo que también se conoce como modo truco para la presentación de vídeo. De acuerdo con estas técnicas, un dispositivo cliente puede, basándose en la señalización de intervalos de octetos de tramas I, solicitar y/o decodificar solamente tramas I señalizadas de una representación de vídeo independiente de otras tramas no necesarias para presentar una versión de alta velocidad de una representación de vídeo. Por ejemplo, un dispositivo cliente puede solicitar específicamente (por ejemplo, mediante una o más solicitudes GET de HTTP parciales) una o más tramas I que se producen dentro de intervalos de octetos señalizados. El dispositivo cliente puede presentar consecutivamente dos o más tramas I de los datos multimedia recuperados para presentar el contenido de vídeo a una velocidad relativamente alta. Las dos o más tramas I pueden estar separadas temporalmente por al menos una trama intercodificada (por ejemplo, tramas B o P). Al recuperar, decodificar y visualizar solo las tramas I señalizadas, el dispositivo cliente puede proporcionar a un usuario una versión de alta velocidad de una representación de vídeo con fines de operaciones de avance o rebobinado rápidos. En algunos ejemplos, las tramas I señalizadas pueden incluir puntos de acceso IDR y/u ODR.

[0046] Los intervalos de octetos para tramas I de una representación de vídeo se pueden señalar en la información de cabecera de varias unidades de codificación (por ejemplo, segmentaciones de medios). Por ejemplo, un dispositivo de origen puede señalar los intervalos de octetos de tramas I en el nivel de vídeo (por ejemplo, dentro de un archivo MPD), el nivel de representación, el nivel de segmento, el nivel de fragmento u otro nivel de información de cabecera

de una presentación multimedia. Estas técnicas pueden proporcionar una o más ventajas. Por ejemplo, un dispositivo de destino puede ejecutar una reproducción a alta velocidad de una representación de vídeo en el contexto de un protocolo de transmisión continua, tal como la transmisión continua de HTTP, sin crear, almacenar y/o transmitir una versión de alta velocidad de la representación de vídeo. Además, estas técnicas pueden ser ventajosas, ya que un dispositivo cliente puede presentar a un usuario una reproducción de avance rápido y/o de rebobinado rápido basándose en tramas I señalizadas.

[0047] La FIG. 1 es un diagrama de bloques que ilustra un sistema 10 de ejemplo en el que un dispositivo de origen de audio/vídeo (A/V) 20 transporta datos de audio y vídeo a un dispositivo de destino de A/V 40. Por consiguiente, el dispositivo de origen de A/V 20 puede denominarse "dispositivo de vídeo de origen". El sistema 10 de la FIG. 1 puede corresponder a un sistema de videoconferencia, un sistema cliente/servidor, un sistema radiodifusor/receptor o cualquier otro sistema en el que se envíen datos de vídeo desde un dispositivo de origen, tal como un dispositivo de origen de A/V 20, a un dispositivo de destino, tal como un dispositivo de destino de A/V 40. El dispositivo de destino de A/V 40 también puede denominarse "dispositivo de vídeo de destino" o "dispositivo cliente". En algunos ejemplos, el dispositivo de origen de A/V 20 y el dispositivo de destino de A/V 40 pueden realizar un intercambio de información bidireccional. Es decir, el dispositivo de origen de A/V 20 y el dispositivo de destino de A/V 40 pueden ser capaces tanto de codificar como de decodificar (y de transmitir y recibir) datos de audio y vídeo. En algunos ejemplos, el codificador de audio 26 puede comprender un codificador de voz, también denominado vocodificador.

[0048] El dispositivo de origen de A/V 20, en el ejemplo de la FIG. 1, comprende una fuente de audio 22 y una fuente de vídeo 24. La fuente de audio 22 puede comprender, por ejemplo, un micrófono que produce señales eléctricas representativas de datos de audio captados que el codificador de audio 26 va a codificar. De forma alternativa, la fuente de audio 22 puede comprender un medio de almacenamiento que almacena datos de audio previamente registrados, un generador de datos de audio, tal como un sintetizador informatizado, o cualquier otra fuente de datos de audio. La fuente de vídeo 24 puede comprender una cámara de vídeo que produce datos de vídeo que han de ser codificados por el codificador de vídeo 28, un medio de almacenamiento codificado con datos de vídeo previamente registrados, una unidad de generación de datos de vídeo o cualquier otra fuente de datos de vídeo.

[0049] Los datos de audio y de vídeo sin procesar pueden comprender datos analógicos o digitales. Los datos analógicos se pueden digitalizar antes de que el codificador de audio 26 y/o el codificador de vídeo 28 los codifiquen. La fuente de audio 22 puede obtener datos de audio a partir de un interlocutor mientras el interlocutor está hablando, y la fuente de vídeo 24 puede obtener simultáneamente datos de vídeo del interlocutor. En otros ejemplos, la fuente de audio 22 puede comprender un medio de almacenamiento legible por ordenador que comprende datos de audio almacenados, y la fuente de vídeo 24 puede comprender un medio de almacenamiento legible por ordenador que comprende datos de vídeo almacenados. De esta manera, las técnicas descritas en esta divulgación se pueden aplicar a datos de audio y vídeo que se transmiten en directo, en continuo o en tiempo real, o a datos de audio y vídeo archivados o registrados previamente.

[0050] Las tramas de audio que corresponden a tramas de vídeo son, en general, tramas de audio que contienen datos de audio que se capturaron por la fuente de audio 22 al mismo tiempo que los datos de vídeo capturados por la fuente de vídeo 24, que están contenidos dentro de las tramas de vídeo. Por ejemplo, mientras un interlocutor produce al hablar, en general, datos de audio, la fuente de audio 22 capta los datos de audio y la fuente de vídeo 24 capta los datos de vídeo del interlocutor al mismo tiempo, es decir, mientras la fuente de audio 22 está captando los datos de audio. Así pues, una trama de audio puede corresponder temporalmente a una o más tramas de vídeo en particular. Por consiguiente, una trama de audio correspondiente a una trama de vídeo corresponde, en general, a una situación en la que se han captado datos de audio y datos de vídeo al mismo tiempo, y para la que una trama de audio y una trama de vídeo comprenden, respectivamente, los datos de audio y los datos de vídeo que se han captado al mismo tiempo.

[0051] En algunos ejemplos, el codificador de audio 26 puede codificar una marca de tiempo en cada trama de audio codificada, que representa el momento en que se grabaron los datos de audio para la trama de audio codificada y, de manera similar, el codificador de vídeo 28 puede codificar una marca de tiempo en cada trama de vídeo codificada, que representa el momento en que se grabaron los datos de vídeo para la trama de vídeo codificada. En dichos ejemplos, una trama de audio correspondiente a una trama de vídeo puede comprender una trama de audio que comprende una marca de tiempo y una trama de vídeo que comprende la misma marca de tiempo. El dispositivo de origen de A/V 20 puede incluir un reloj interno a partir del cual el codificador de audio 26 y/o el codificador de vídeo 28 pueden generar las marcas de tiempo, o que la fuente de audio 22 y la fuente de vídeo 24 pueden usar para asociar datos de audio y vídeo, respectivamente, con una marca de tiempo.

[0052] En algunos ejemplos, la fuente de audio 22 puede enviar datos al codificador de audio 26 correspondientes al momento en que se registraron los datos de audio, y la fuente de vídeo 24 puede enviar datos al codificador de vídeo 28, correspondientes al momento en que se registraron los datos de vídeo. En algunos ejemplos, el codificador de audio 26 puede codificar un identificador de secuencia en datos de audio codificados para indicar un orden temporal relativo de los datos de audio codificados, pero sin indicar necesariamente un tiempo absoluto en el que se han registrado los datos de audio y, de forma similar, el codificador de vídeo 28 también puede usar identificadores de

secuencia para indicar un orden temporal relativo de los datos de vídeo codificados. De forma similar, en algunos ejemplos, un identificador de secuencia se puede mapear o correlacionar de otro modo con una marca de tiempo.

[0053] Las técnicas de esta divulgación están dirigidas, en general, al transporte de datos multimedia codificados (por ejemplo, audio y vídeo), y a la recepción y la posterior interpretación y decodificación de los datos multimedia transportados. Las técnicas de esta divulgación se pueden aplicar al transporte de datos de vídeo de varias normas y extensiones tales como, por ejemplo, datos de codificación de vídeo escalable (SVC), de codificación de vídeo avanzada (AVC), de formato de archivo de medios base ISO, de codificación de vídeo multivista (MVC) u otros datos de vídeo. En algunos ejemplos, la fuente de vídeo 24 puede proporcionar una pluralidad de vistas de una escena al codificador de vídeo 28. En algunos ejemplos, múltiples vistas de datos de vídeo pueden ser útiles para generar datos de vídeo tridimensionales que serán utilizados por un dispositivo de visualización tridimensional, tal como una pantalla tridimensional estereoscópica o autoestereoscópica.

[0054] El dispositivo de origen de A/V 20 puede proporcionar un "servicio" al dispositivo de destino de A/V 40. Un servicio puede corresponder, en general, a un subconjunto de vistas disponibles de datos MVC. Por ejemplo, datos de vídeo multivista pueden estar disponibles para ocho vistas, ordenadas de cero a siete. Un servicio puede corresponder a vídeo estéreo que tiene dos vistas, mientras que otro servicio puede corresponder a cuatro vistas y otro servicio más puede corresponder a las ocho vistas. En general, un servicio corresponde a cualquier combinación (es decir, a cualquier subconjunto) de las vistas disponibles. Un servicio también puede corresponder a una combinación de vistas disponibles, así como datos de audio.

[0055] Cada flujo individual de datos (ya sea de audio o vídeo) se menciona como un flujo elemental. Un flujo elemental es un componente único, codificado digitalmente (posiblemente comprimido) de un programa. Por ejemplo, la parte de vídeo o audio codificada del programa puede ser un flujo elemental. Un flujo elemental puede convertirse en un flujo elemental empaquetado (PES) antes de multiplexarse en un flujo de programa o flujo de transporte. Dentro del mismo programa, se utiliza un ID de flujo para distinguir los paquetes PES pertenecientes a un flujo elemental del otro. La unidad básica de datos de un flujo elemental es un paquete de flujo elemental empaquetado (PES). Por tanto, cada vista de datos de vídeo MVC corresponde a flujos elementales respectivos. De forma similar, los datos de audio corresponden a uno o más flujos elementales respectivos.

[0056] En el ejemplo de la FIG. 1, la unidad de encapsulación 30 recibe flujos elementales que comprenden datos de vídeo del codificador de vídeo 28 y flujos elementales que comprenden datos de audio del codificador de audio 26. En algunos ejemplos, el codificador de vídeo 28 y el codificador de audio 26 pueden incluir, cada uno, empaquetadores para formar paquetes PES a partir de datos codificados. En otros ejemplos, tanto el codificador de vídeo 28 como el codificador de audio 26 pueden interactuar con empaquetadores respectivos para formar paquetes PES a partir de datos codificados. En otros ejemplos más, la unidad de encapsulación 30 puede incluir empaquetadores para formar paquetes PES a partir de datos de audio y de vídeo codificados.

[0057] Un "programa", tal como se utiliza en esta divulgación, puede comprender una combinación de datos de audio y de datos de vídeo, por ejemplo, un flujo elemental de audio y un subconjunto de vistas disponibles entregadas por un servicio del dispositivo de origen de A/V 20. Cada paquete PES incluye un `stream_id` que identifica el flujo elemental al que pertenece el paquete PES. La unidad de encapsulación 30 puede ensamblar flujos elementales en flujos de programa o flujos de transporte constituyentes. Un flujo de programa y un flujo de transporte son dos multiplexores alternativos que tienen como objetivo diferentes aplicaciones.

[0058] En general, un flujo de programa incluye datos de un programa, mientras que un flujo de transporte puede incluir datos para uno o más programas. La unidad de encapsulación 30 puede codificar un flujo de programa y/o un flujo de transporte basándose en el servicio que se proporciona, el medio en el que se pasará el flujo, el número de programas que se enviarán u otras consideraciones. Por ejemplo, cuando los datos de vídeo van a codificarse en un medio de almacenamiento, puede ser más probable que la unidad de encapsulación 30 forme un flujo de programa, mientras que, cuando los datos de vídeo van a transmitirse de forma continua a través de una red, radiodifundirse o enviarse como parte de videotelefonía, puede ser más probable que la unidad de encapsulación 30 use un flujo de transporte.

[0059] La unidad de encapsulación 30 recibe paquetes PES para flujos elementales de un programa desde el codificador de audio 26 y el codificador de vídeo 28 y forma las correspondientes unidades de capa de abstracción de red (NAL) a partir de los paquetes PES. En el ejemplo de la norma H.264/AVC (codificación de vídeo avanzada), los segmentos de vídeo codificados están organizados en unidades NAL, que proporcionan una representación de vídeo "apta para redes" dirigida a aplicaciones tales como la videotelefonía, el almacenamiento, la radiodifusión o la transmisión continua. Las unidades NAL se pueden clasificar en unidades NAL de capa de codificación de vídeo (VCL) y unidades NAL no VCL. Las unidades VCL contienen el motor de compresión de núcleo y pueden comprender niveles de bloque, macrobloque y/o fragmento. Otras unidades NAL son unidades NAL no de VCL.

[0060] Una unidad NAL que incluye datos de vídeo en su carga útil puede comprender diversos niveles de granularidad de datos de vídeo. Por ejemplo, una unidad NAL puede comprender un bloque de datos de vídeo, un

macrobloque, una pluralidad de macrobloques, un fragmento de datos de vídeo o una trama completa de datos de vídeo.

[0061] En general, una unidad de acceso puede comprender una o más unidades NAL para representar una trama de datos de vídeo, así como datos de audio correspondientes a la trama cuando dichos datos de audio están disponibles. Una unidad de acceso incluye, en general, todas las unidades NAL para una instancia de tiempo de salida, por ejemplo, todos los datos de audio y vídeo para una instancia de tiempo. En un ejemplo correspondiente a la norma H.264/AVC, una unidad de acceso puede comprender una imagen codificada en una instancia de tiempo, que puede presentarse como una imagen codificada primaria. En consecuencia, una unidad de acceso puede comprender todas las tramas de vídeo de una instancia temporal común, por ejemplo, todas las componentes de vista correspondientes al tiempo X.

[0062] La unidad de encapsulación 30 puede ensamblar unidades NAL en pistas de un archivo de vídeo que se ajuste al formato de archivo de medios base ISO o un derivado del mismo (por ejemplo, SVC, AVC, MVC o 3GPP) y pasar el archivo de vídeo a la interfaz de salida 32. La interfaz de salida 32 puede comprender, por ejemplo, un transmisor, un transceptor, un dispositivo para escribir datos en un medio legible por ordenador tal como, por ejemplo, una unidad óptica, una unidad de medios magnéticos (por ejemplo, una unidad de disquete), un puerto de bus serie universal (USB), una interfaz de red u otra interfaz de salida. La interfaz de salida 32 proporciona la unidad NAL o la unidad de acceso a un medio legible por ordenador (por ejemplo, canal 34), por ejemplo, un medio transitorio tal como una señal de transmisión o una onda portadora, o un medio de almacenamiento legible por ordenador tal como un medio magnético, un medio óptico, una memoria o una unidad flash.

[0063] La interfaz de entrada 36 recupera los datos del medio legible por ordenador (por ejemplo, canal 34). La interfaz de entrada 36 puede comprender, por ejemplo, una unidad óptica, una unidad de medios magnéticos, un puerto USB, un receptor, un transceptor u otra interfaz de medio legible por ordenador. La interfaz de entrada 36 puede proporcionar la unidad NAL o la unidad de acceso a la unidad de desencapsulación 38. La unidad de desencapsulación 38 puede desmultiplexar un flujo de transporte o un flujo de programa en flujos PES constituyentes, desempaquetar los flujos PES para recuperar datos codificados y enviar los datos codificados al decodificador de audio 46 o al decodificador de vídeo 48, dependiendo de si los datos codificados forman parte de un flujo de audio o vídeo, por ejemplo, según lo indicado por las cabeceras de paquetes PES del flujo. La unidad de desencapsulación 38 puede seleccionar inicialmente una de las pistas incluidas en un archivo de vídeo recibido y, a continuación, pasar solamente datos de la pista seleccionada y datos de otras pistas referenciadas por extractores de la pista seleccionada al decodificador de vídeo 48, descartando datos de otras pistas no referenciadas por un extractor de la pista seleccionada. El decodificador de audio 46 decodifica datos de audio codificados y envía los datos de audio decodificados a la salida de audio 42, mientras que el decodificador de vídeo 48 decodifica datos de vídeo codificados y envía los datos de vídeo decodificados, que pueden incluir una pluralidad de vistas de un flujo, a la salida de vídeo 44. La salida de vídeo 44 puede comprender un dispositivo de visualización que usa una pluralidad de vistas de una escena, por ejemplo, una pantalla estereoscópica o autoestereoscópica que presenta cada vista de una escena simultáneamente.

[0064] Como ocurre con la mayoría de normas de codificación de vídeo, la norma H.264/AVC define la sintaxis, la semántica y el proceso de decodificación para flujos de bits libres de errores, cualquiera de los cuales se ajusta a un determinado perfil o nivel. La norma H.264/AVC no especifica el codificador, sino que el codificador tiene la tarea de garantizar que los flujos de bits generados sean compatibles con la norma para un decodificador. En el contexto de la norma de codificación de vídeo, un "perfil" corresponde a un subconjunto de algoritmos, características o herramientas y restricciones que se les aplican. Como se define en la norma H.264, por ejemplo, un "perfil" es un subconjunto de toda la sintaxis de flujo de bits especificada por la norma H.264. Un "nivel" corresponde a las limitaciones del consumo de recursos de decodificador, tales como, por ejemplo, memoria de decodificador y cálculo, que están relacionadas con la resolución de las imágenes, la velocidad binaria y velocidad de procesamiento de macrobloques (MB).

[0065] La norma H.264, por ejemplo, reconoce que, dentro de los límites impuestos por la sintaxis de un perfil dado, todavía es posible requerir una gran variación en el rendimiento de los codificadores y decodificadores dependiendo de los valores tomados por los elementos sintácticos en el flujo de bits, tales como el tamaño especificado de las imágenes decodificadas. La norma H.264 reconoce, además, que, en muchas aplicaciones, no es ni práctico ni económico implementar un decodificador capaz de tratar todos los usos hipotéticos de la sintaxis dentro de un perfil en particular. Por consiguiente, la norma H.264 define un "nivel" como un conjunto especificado de restricciones impuestas a los valores de los elementos de sintaxis en el flujo de bits. Estas restricciones pueden ser simples límites en los valores. De forma alternativa, estas restricciones pueden adoptar la forma de restricciones sobre combinaciones aritméticas de valores (por ejemplo, la anchura de imagen multiplicada por la altura de imagen multiplicada por el número de imágenes decodificadas por segundo). La norma H.264 establece, además, que las implementaciones individuales pueden admitir un nivel diferente para cada perfil soportado.

[0066] Un decodificador conforme a un perfil admite ordinariamente todas las características definidas en el perfil. Por ejemplo, como característica de codificación, la codificación de imágenes B no está soportada en el perfil de línea de base de H.264/AVC, pero está soportada en otros perfiles de H.264/AVC. Un decodificador conforme a un nivel deberá ser capaz de decodificar cualquier flujo de bits que no requiere recursos más allá de las limitaciones definidas en el nivel. Las definiciones de perfiles y niveles pueden ser útiles para la interpretabilidad. Por ejemplo, durante la

transmisión de vídeo, se pueden negociar y acordar un par de definiciones de perfil y nivel para una sesión de transmisión completa. Más específicamente, en la norma H.264/AVC, un nivel puede definir, por ejemplo, limitaciones en el número de macrobloques que necesitan ser procesados, el tamaño del búfer de imágenes decodificadas (DPB), el tamaño del búfer de imágenes codificadas (CPB), el intervalo de vector de movimiento vertical, el número máximo de vectores de movimiento por dos MB consecutivos y si un bloque B puede tener particiones de submacrobloque inferiores a 8x8 píxeles. De esta manera, un decodificador puede determinar si el decodificador es capaz de decodificar apropiadamente el flujo de bits.

[0067] Los conjuntos de parámetros contienen, en general, información de cabecera de capa de secuencia en los conjuntos de parámetros de secuencia (SPS), y la información de cabecera de capa de imagen que cambia ocasionalmente en conjuntos de parámetros de imagen (PPS). Con los conjuntos de parámetros, esta información que cambia ocasionalmente no necesita repetirse para cada secuencia o imagen; de ahí que pueda mejorarse la eficacia de la codificación. Además, el uso de conjuntos de parámetros puede permitir la transmisión fuera de banda de información de cabecera, evitando la necesidad de transmisiones redundantes para lograr resistencia a los errores. En la transmisión fuera de banda, las unidades NAL de conjuntos de parámetros se pueden transmitir en un canal diferente al de otras unidades NAL.

[0068] Los mensajes de información de mejora suplementaria (SEI) pueden contener información que no es necesaria para decodificar muestras de imágenes codificadas a partir de unidades NAL VCL. Los mensajes SEI pueden estar contenidos en unidades NAL no VCL, como se describe anteriormente. En algunos ejemplos, los mensajes SEI son una parte normativa de una especificación estándar (por ejemplo, H.264) y no son obligatorios para la implementación de un decodificador de conformidad con la norma. En algunos ejemplos, los mensajes SEI pueden ayudar en procesos relacionados con la decodificación, visualización, resistencia a errores y otros propósitos. En algunos ejemplos, los mensajes SEI pueden ser mensajes de nivel de secuencia (por ejemplo, una pluralidad de imágenes/tramas consecutivos). En otros ejemplos, los mensajes SEI pueden ser específicos de una o más imágenes/tramas de datos de vídeo.

[0069] En H.264/AVC y otras normas de vídeo puede haber imágenes intracodificadas que son la primera imagen en el orden de decodificación de un GOP abierto. Un GOP abierto es un grupo de imágenes en el que las imágenes que preceden a una intraimagen inicial en el orden de salida pueden no decodificarse correctamente. Un decodificador H.264/AVC puede reconocer una intraimagen que inicia un GOP abierto y comenzar a decodificar las imágenes siguientes. Cabe señalar que dicha intraimagen puede estar asociada a un SEI de punto de recuperación. En algunos ejemplos, una trama que inicia un GOP abierto se denomina en el presente documento trama de actualización de decodificación abierta (ODR). Un GOP cerrado es un grupo de imágenes en el que todas las imágenes se pueden decodificar correctamente. En H.264/AVC, un GOP cerrado puede comenzar desde una unidad de acceso IDR.

[0070] La FIG. 2 es un diagrama conceptual que ilustra un ejemplo de fragmento de vídeo 201 que incluye un grupo de imágenes (GOP) 210. Un GOP puede describirse como una agrupación de tramas secuenciales (en orden de presentación) de una secuencia de vídeo. Un fragmento de vídeo puede incluir cualquier número de GOP similares a GOP 210.

[0071] El ejemplo de la FIG. 2 representa una parte de un fragmento de vídeo 201. El fragmento de vídeo incluye una pluralidad de tramas 220A-220K. El fragmento 201 puede comprender uno o más grupos de imágenes (GOP), tales como GOP 210. El uno o más GOP 210 se pueden describir como una agrupación de tramas de una representación de vídeo.

[0072] El GOP 210 ilustrado en la FIG. 2 es un ejemplo de un GOP cerrado. Es decir, cada una de las tramas 220B-220I se puede decodificar sin referencia a las tramas externas al GOP 210, por ejemplo, las tramas 220A, 220J y 220K. Como se muestra en la FIG. 2, la trama 220A precede (en orden de presentación) a las tramas 220B-220I del GOP 210. Las tramas 220J-220K son posteriores (en orden de presentación) a las tramas 220B-220I del GOP 210. Las tramas 220A y 220J-220K pueden ser parte de otro GOP del fragmento 201.

[0073] Un orden de presentación de las tramas 220A-220K puede ser diferente de un orden de decodificación de las tramas. Por ejemplo, un GOP puede incluir cualquier combinación de tramas I, B o P. Las tramas intracodificadas (tramas I) son aquellas tramas que se pueden decodificar de forma independiente, lo que significa que para decodificar una trama I, un decodificador no necesita depender del contenido de otras tramas. En el ejemplo de la FIG. 2, las tramas 220A y 220I son ejemplos de tramas I. Las tramas P pueden intercodificarse con respecto a una o más tramas en una dirección. En el ejemplo de la FIG. 2, las tramas 220B, 220C y 220E son ejemplos de tramas P. Las tramas B pueden estar intercodificadas con respecto a una o más tramas en dos direcciones. En el ejemplo de la FIG. 2, las tramas 220D, 220F, 220G y 220H son ejemplos de tramas B.

[0074] Como se analizó anteriormente, de acuerdo con el ejemplo de la FIG. 2, un orden de decodificación de las tramas 220A-220K puede ser diferente al orden de presentación de las tramas. Por ejemplo, al decodificar las tramas 220B-220I, la trama 220I (una trama I) puede decodificarse en primer lugar. Como indica la flecha 222E, la trama 220E depende del contenido de la trama 220I para decodificarse correctamente. De este modo, la trama 220E puede decodificarse después de decodificar la trama 220I.

[0075] Como indica la flecha 222C, la trama 220C puede depender del contenido de la trama 220E para decodificarse correctamente. Como indica la flecha 222G, la trama 220G puede depender del contenido de ambas tramas 220E y 220I para decodificarse correctamente. De este modo, en algunos ejemplos, la decodificación de las tramas 220C y 220G puede producirse después de la decodificación de las tramas 220I y 220E. Cada trama 220B, 220D, 220F y 220H depende del contenido de una o más de las tramas 220C, 220E, 220G y 220I, respectivamente, y, por lo tanto, pueden decodificarse después de que las tramas 220C, 220E, 220G y 220I se hayan decodificado.

[0076] Como se ha descrito anteriormente, un punto de acceso de actualización de decodificación instantánea (IDR) se puede describir como un punto de acceso de un GOP cerrado, por ejemplo, el GOP 210 en la FIG. 2. Un GOP que incluye solo tramas que se pueden decodificar correctamente sin depender del contenido de tramas fuera del GOP puede considerarse un GOP cerrado 210.

[0077] La FIG. 2 muestra dos ejemplos de puntos de acceso IDR. Como se muestra en la FIG. 2, por ejemplo, la trama 220A no depende del contenido de ninguna otra trama para decodificarse correctamente, es decir, la trama 220A no incluye ninguna flecha que indique dependencia de otra trama. La trama 220A puede considerarse un GOP en sí misma, porque no hay tramas que precedan a la trama 220A. De este modo, la trama 220A puede considerarse un punto de acceso IDR, ya que la trama 220A es un punto de acceso de un GOP que no depende del contenido de ninguna trama fuera del GOP (que consiste solamente en la trama 220A) para decodificarse correctamente.

[0078] La trama 220I también puede considerarse un punto de acceso IDR del GOP 210 cerrado. Como se muestra en la FIG. 2, por ejemplo, la trama 220I es una trama I que se puede decodificar de forma independiente sin depender del contenido de ninguna otra trama (por ejemplo, tramas 220B-220H) del GOP 210. Aunque cada una de las tramas 220B-220H depende del contenido de otras tramas dentro del GOP 210 para decodificarse correctamente como se describe anteriormente, ninguna de las tramas 220B-220H se basa en el contenido de ninguna trama fuera del GOP 210. De este modo, el GOP 210 puede considerarse un GOP cerrado que incluye un punto de acceso IDR, a saber, la trama 220I.

[0079] La FIG. 3 es un diagrama conceptual que ilustra un ejemplo de al menos una parte de un fragmento de vídeo 301 que incluye un GOP abierto 310 que incluye un punto de acceso de actualización de decodificación abierta (ODR). Similar al ejemplo de la FIG. 2, la trama 320A es una trama I y un punto de acceso IDR. También similar al ejemplo de la FIG. 2, la trama 320I es una trama I correspondiente a un punto de acceso aleatorio. Sin embargo, el ejemplo de la FIG. 3 difiere del de la FIG. 2, en el que tramas de GOP 310 anteriores a la trama I 320I en el orden de visualización depende del contenido de la trama 320A para poder decodificarse correctamente. Por ejemplo, como se indica mediante flechas direccionales, cada una de las tramas 320B, 320C y 320E depende directamente del contenido de la trama 320A. Las tramas 320D y 320F-320H dependen cada una indirectamente del contenido de la trama 320A, ya que cada una depende, al menos en parte, del contenido de una o más de las tramas 320B, 320C y 320E para poder decodificarse correctamente. Sin embargo, como también se muestra en la FIG. 3, las tramas 320J y 320K, que siguen a la trama I 320I en orden de visualización, pueden decodificarse correctamente sin depender de ninguna trama anterior a la trama I 320I. Por tanto, la trama I 320I puede usarse como un RAP ODR.

[0080] Esta divulgación se refiere a técnicas para la mejora de la funcionalidad de acceso aleatorio en el contexto de la transmisión continua de vídeo. Las técnicas de esta divulgación proporcionan la señalización de puntos de acceso ODR, y otra información diversa, para el comienzo de la transmisión continua de HTTP, la decodificación y/o la reproducción de una presentación de vídeo.

[0081] Por ejemplo, un dispositivo cliente puede recibir, desde un dispositivo servidor, una indicación de un punto de acceso ODR de un primer GOP de una representación de vídeo. La información de codificación para al menos una trama del primer GOP indica que al menos una trama de la segmentación de medios que sigue a la trama RAP ODR en orden de decodificación puede no estar correctamente decodificada. Cada trama de la segmentación de medios que sigue a la trama RAP ODR en orden de visualización se puede decodificar correctamente sin depender del contenido de los tramas anteriores al RAP ODR en orden de visualización. El dispositivo cliente puede enviar, al dispositivo servidor, una solicitud (por ejemplo, una solicitud GET o GET parcial) para recuperar un flujo de datos de medios que comienza con el RAP ODR indicado. A continuación, el dispositivo cliente puede comenzar la decodificación/reproducción de los datos multimedia en el RAP ODR.

[0082] En particular, el dispositivo de origen puede señalar un desplazamiento de octetos a un RAP ODR de un fragmento de vídeo. El fragmento de vídeo se puede direccionar de forma independiente mediante un localizador de recursos uniforme (URL) o un nombre de recurso uniforme (URN) específico para el fragmento de vídeo. Al señalar el desplazamiento de octetos al RAP ODR en el fragmento de vídeo, el dispositivo cliente puede generar una solicitud parcial GET que comprende un intervalo de octetos que comienza en el RAP ODR y termina al final del fragmento de vídeo, en un ejemplo.

[0083] Como otro ejemplo, un dispositivo servidor puede recibir una secuencia de tramas de vídeo codificadas. Un archivo de segmentación de la secuencia de tramas de vídeo codificadas incluye al menos un RAP ODR. Es posible que al menos una trama de la segmentación de medios que sigue a la trama RAP ODR en orden de decodificación no

se decodifique correctamente. Cada trama de la segmentación de medios que sigue a la trama RAP ODR en orden de visualización se puede decodificar correctamente sin depender del contenido de las tramas anteriores al RAP ODR en orden de visualización. El dispositivo servidor puede generar una indicación de una ubicación del RAP ODR. El dispositivo servidor puede proporcionar la al menos una indicación a un dispositivo cliente (por ejemplo, en información de cabecera de un archivo de vídeo, presentación o fragmento de una presentación de vídeo). El cliente puede, en base a la al menos una indicación, solicitar la transmisión continua de un segmento o de una parte de un segmento de una presentación de vídeo en base a la al menos una indicación.

[0084] Como otro ejemplo, un dispositivo cliente puede recibir al menos una indicación de un intervalo de octetos para al menos una trama I de un fragmento de vídeo o presentación. El dispositivo cliente puede, en base a la indicación, enviar una solicitud al dispositivo servidor para recuperar la trama I. El dispositivo cliente puede usar la al menos una trama I recuperada para presentar una versión de alta velocidad de una presentación de vídeo o fragmento que incluye la trama I. La trama I puede ser, o no, un punto de acceso IDR u ODR, como se describe arriba. La versión de alta velocidad de la presentación de vídeo puede presentarse en respuesta a que el cliente reciba una solicitud de usuario referente a una operación de avance o rebobinado rápidos.

[0085] La FIG. 4 es un diagrama de bloques que ilustra un ejemplo de un programa 430. El programa 430 puede ser una representación digital de un vídeo (por ejemplo, una película, programa de televisión u otra presentación de imágenes en movimiento). El programa 430 puede estar almacenado en o ser accesible por uno o más dispositivos informáticos configurados para funcionar como un servidor para la transmisión continua de HTTP. De acuerdo con el ejemplo de la FIG. 4, el programa 430 incluye una pluralidad de representaciones 440A-440C y una descripción de presentación de medios (MPD) 432. La pluralidad de representaciones 440A-440C puede incluir una pluralidad de fragmentos de película 442. La pluralidad de fragmentos 442 puede disponerse en orden de reproducción y representar varias porciones temporales (es decir, en el tiempo) de un vídeo.

[0086] Aunque no se representa en la FIG. 4, uno o más fragmentos 442 de una representación de película 440A-440C pueden disponerse en uno o más segmentos de película. Un segmento puede incluir un conjunto de fragmentos de película continuos que se pueden solicitar de acuerdo con un único URL. Por ejemplo, un segmento puede ser un archivo pequeño, que puede contener vídeo y audio. Como otro ejemplo, un segmento puede corresponder a un fragmento de película, que a su vez puede contener un solo fragmento de pista de vídeo y un solo fragmento de pista de audio. Como otro ejemplo más, un segmento puede corresponder a varios fragmentos de película, cualquiera de los cuales puede tener un fragmento de vídeo y un fragmento de audio, y los fragmentos de película pueden ser continuos en tiempo de decodificación. Aunque no se muestra en la FIG. 4, las técnicas para señalar información para la representación de la reproducción de vídeo de acuerdo con la transmisión continua de HTTP pueden ser aplicables a cualquier información de cabecera para un programa de medios, tal como, por ejemplo, información de cabecera para un archivo de segmentación, por ejemplo, una representación de película, un segmento de película, un fragmento de película.

[0087] La MPD 432 puede considerarse información de cabecera para el programa 430. La MPD 432 puede incluir información que describe una o más representaciones 440A-440C. Un dispositivo cliente puede recuperar MPD la 432 con el fin de seleccionar todas o una parte de una o más de las representaciones 440A-440C para la reproducción mediante transmisión continua. Una representación 440 también puede incluir cabeceras de fragmento 443. Las cabeceras de fragmento 443 pueden incluir información específica de un fragmento particular 442 de las representaciones 440A-440C. Como la MPD 432, la cabecera de fragmento 443 puede ser accesible por un dispositivo cliente HTTP, por ejemplo, para la selección de uno o más fragmentos 442 de una o más representaciones 440A-440C.

[0088] Un fragmento 442 de una representación puede incluir, o no, un punto de acceso IDR u ODR. Para algunas representaciones, un fragmento 442 solo puede incluir un único punto de acceso IDR. Se pueden señalar uno o más puntos de acceso IDR en la información de cabecera para el programa 430 (por ejemplo, la MPD 432 de la FIG. 4) o en la información de cabecera 443 para uno o más fragmentos 442 de un programa 430. Un aspecto de esta divulgación está dirigido a señalar, en uno o más de una MPD 432, una cabecera de fragmento 443 u otra información de cabecera de un programa 430, puntos de acceso ODR. Por consiguiente, un cliente HTTP puede estar provisto de un mayor número de opciones temporales para comenzar la transmisión continua, la decodificación y/o la reproducción de un programa 430.

[0089] En algunos ejemplos, una película puede tener una duración de dos horas o más. Un programa 430, o representación que representa la película, puede incluir una pluralidad de fragmentos de película dispuestos temporalmente (en orden de tiempo de reproducción) que pueden representar porciones relativamente cortas de la presentación completa. Por ejemplo, cada fragmento de película puede representar varios segundos o durar varios minutos. De acuerdo con la transmisión continua de HTTP, un dispositivo cliente puede solicitar la reproducción de un archivo de segmentación en un momento específico de la película. Para ello, el dispositivo cliente puede enviar una solicitud GET o GET parcial a un servidor que proporciona la representación. Una solicitud GET de HTTP puede dirigirse a un URL o URN en particular para recuperar un archivo completo asociado al URL o URN. Una solicitud GET parcial de HTTP también se dirige a un URL o URN, pero especifica además un intervalo de octetos del archivo a recuperar.

[0090] De acuerdo con las técnicas de esta divulgación, algunos fragmentos 442 de las representaciones de película 440 pueden incluir información de cabecera que señala si un fragmento particular incluye un punto de acceso IDR. En algunos casos, si un fragmento de película solo incluye una IDR como primera trama de un fragmento de película, un dispositivo cliente puede estar obligado a comenzar la reproducción al principio, o en otra ubicación temporal específica, de un fragmento de vídeo. Es posible que algunos fragmentos de película no incluyan un punto de acceso IDR y solo incluyan uno o más puntos de acceso ODR. En algunos ejemplos, es posible que un dispositivo cliente no sea capaz de iniciar la reproducción en ningún punto del fragmento de vídeo.

[0091] Un fragmento de vídeo puede representar varios segundos, o incluso minutos, de datos de vídeo. Por lo tanto, un cliente puede disponer de opciones limitadas para el comienzo de la reproducción si solo se señalizan puntos de acceso IDR. Por ejemplo, es posible que se requiera que un usuario vea partes ya vistas de una presentación de vídeo para regresar a un instante de tiempo particular de la presentación, por ejemplo si el usuario detiene la reproducción y posteriormente solicita reanudar la reproducción desde una ubicación temporal particular durante una visualización posterior. Como otro ejemplo, un dispositivo cliente puede necesitar descargar más tramas de una presentación de vídeo de las necesarias para la decodificación/reproducción comenzando en una ubicación temporal solicitada por el usuario, o una ubicación temporal iniciada por el dispositivo cliente (por ejemplo, donde el dispositivo cliente ha solicitado una representación alternativa de un programa, tal como cuando las condiciones de ancho de banda han cambiado). Además, algunos fragmentos de vídeo pueden no incluir un punto de acceso IDR.

[0092] Al señalizarse puntos de acceso ODR en la información de cabecera de un archivo de segmentación, puede mejorarse la funcionalidad de exploración y búsqueda de una representación de vídeo, o de un fragmento de vídeo, para la transmisión continua de HTTP. Por ejemplo, donde las técnicas conocidas para la transmisión continua de HTTP solo pueden proporcionar el comienzo de la transmisión/decodificación/reproducción en un pequeño número de fragmentos y/o posiciones dentro de los fragmentos, las técnicas de esta divulgación pueden proporcionar un mayor número de puntos de acceso aleatorio para la reproducción. Las técnicas de esta divulgación también pueden ser útiles para representaciones de vídeo que incluyen una gran cantidad de dependencia entre GOP para la decodificación.

[0093] Otros aspectos de esta divulgación proporcionan la señalización de un intervalo de octetos de tramas I de un archivo de segmentación o presentación con fines de reproducción de alta velocidad (por ejemplo, avance rápido, rebobinado) de una representación de película. De acuerdo con estos aspectos, un intervalo de octetos de una o más tramas I de una presentación de película puede ser señalado a un dispositivo cliente, de modo que el dispositivo cliente pueda solicitar, recibir y/o decodificar tramas I de la representación de película para propósitos de visualización de la representación de vídeo a alta velocidad, por ejemplo, visualizando sólo las tramas I o un subconjunto de las tramas I disponibles de la representación de vídeo.

[0094] Las técnicas de esta divulgación se pueden aplicar a archivos de vídeo que cumplen con cualquiera de entre el formato de archivo de medios base ISO, el formato de archivo de codificación de vídeo escalable (SVC), el formato de archivo de codificación de vídeo avanzada (AVC), el formato de archivo del Proyecto de Asociación de Tercera Generación (3GPP) y/o el formato de archivo de codificación de vídeo multivista (MVC). El formato de archivo de medios base ISO está diseñado para contener información de medios temporizada para una presentación en un formato flexible y extensible que facilite el intercambio, la gestión, la edición y la presentación de los medios. El formato de archivo de medios base ISO (ISO/IEC 14496-12:2004) se especifica en la norma MPEG-4 Parte-12, que define una estructura general para archivos de medios basados en el tiempo. Se usa como base para otros formatos de archivo en la familia, tales como el formato de archivo AVC (ISO/IEC 14496-15) definido para admitir la compresión de vídeo de la norma H.264/MPEG-4 AVC, el formato de archivo 3GPP, el formato de archivo SVC y el formato de archivo MVC. El formato de archivo 3GPP y el formato de archivo MVC son extensiones del formato de archivo AVC. El formato de archivo de medios base ISO contiene la temporización, la estructura y la información de medios para secuencias temporizadas de datos de medios, tales como presentaciones audiovisuales. La estructura de archivo está orientada a objetos. Un archivo puede descomponerse en objetos básicos de manera muy sencilla y la estructura de los objetos está implícita a partir de su tipo.

[0095] Los archivos de conformidad con el formato de archivo multimedia base ISO pueden formarse como una serie de objetos, llamados "cajas". Los datos en el formato de archivo de medios base ISO pueden estar contenidos en cajas y no hay ningún otro dato dentro del archivo. Esto incluye cualquier firma inicial requerida por el formato de archivo específico. La "caja" es un bloque de construcción orientado a objetos, definido por un identificador de tipo único y una longitud. Típicamente, una presentación está contenida en un archivo y la presentación de medios está autocontenida. El contenedor de película (caja de película) contiene los metadatos de los medios, y las tramas de vídeo y audio están contenidas en el contenedor de datos de medios y podrían estar en otros archivos.

[0096] Una presentación (secuencia de movimiento) puede estar contenida en varios archivos. La información de temporización y entramado (posición y tamaño) puede incluirse en un archivo de conformidad con el formato de archivo de medios base ISO, mientras que archivos auxiliares pueden usar fundamentalmente cualquier formato. Esta presentación puede ser "local" al sistema que contenga la presentación, o puede proporcionarse a través de una red u otro mecanismo de suministro de transmisión continua.

[0097] Los archivos pueden tener una estructura lógica, una estructura de tiempo y una estructura física, y no es necesario que estas estructuras estén acopladas. La estructura lógica del archivo puede ser la de una película que contiene a su vez un conjunto de pistas paralelas de tiempo. La estructura temporal del archivo puede ser que las pistas contengan secuencias de muestras en el tiempo, y que esas secuencias se mapeen en la línea de tiempo de toda la película mediante listas de edición opcionales. La estructura física del archivo puede separar los datos necesarios para la descomposición lógica, temporal y estructural de las propias muestras de datos de medios. Esta información estructural puede concentrarse en una caja de película, posiblemente extendida en el tiempo por cajas de fragmentos de película. La caja de película puede documentar las relaciones lógicas y de temporización de las muestras, y también puede contener punteros a donde están ubicadas. Esos punteros pueden estar en el mismo archivo o en otro, por ejemplo, referenciado por un URL.

[0098] Cada flujo de medios puede estar contenido en una pista especializada para ese tipo de medios (audio, vídeo, etc.) y, además, puede estar parametrizado por una entrada de muestra. Las entrada de muestra puede contener el 'nombre' del tipo de medios exacto (el tipo de decodificador necesario para decodificar el flujo) y cualquier parametrización necesaria para ese decodificador. El nombre también puede adoptar la forma de un código de cuatro caracteres, por ejemplo, "moov" o "trak". Existen formatos de entrada de muestra definidos no solo para los medios MPEG-4, sino también para los tipos de medios utilizados por otras organizaciones que usan esta familia de formatos de archivo.

[0099] El soporte para metadatos adopta, en general, dos formas. Primero, los metadatos temporizados pueden almacenarse en una pista apropiada y sincronizarse según se desee con los datos de medios que están describiendo. En segundo lugar, puede haber soporte general para metadatos no temporizados adjuntos a la película o a una pista individual. El soporte estructural es general y permite, como en los datos de medios, el almacenamiento de recursos de metadatos en otra parte del archivo o en otro archivo. Así mismo, a estos recursos se les puede asignar un nombre y pueden protegerse.

[0100] Como se analiza anteriormente, un ejemplo de una extensión de formato de archivo de medios base ISO es la formación de archivo AVC. Un archivo de segmentación creado de acuerdo con el formato de archivo AVC puede incluir unidades NAL que son conjuntos de parámetros, incluidos conjuntos de parámetros de secuencia (SPS) y conjuntos de parámetros de imagen (PPS). Los conjuntos de parámetros pueden asociarse a un flujo elemental de un vídeo, que puede estar incluido en una o más pistas de vídeo. En algunos ejemplos, un conjunto de parámetros puede estar en otra pista (una pista de conjunto de parámetros) que un flujo elemental de un vídeo. La pista de conjunto de parámetros puede ser un flujo elemental de conjunto de parámetros que contiene muestras que son unidades NAL SPS y/o PPS. Las unidades NAL SPS y PPS pueden incluir uno o más mensajes SEI que incluyen información que no es necesaria para decodificar muestras de imágenes codificadas a partir de unidades NAL VCL. La ubicación de unidades NAL que incluyen mensajes SEI no está restringida, siempre y cuando cualquier mensaje SEI esté presente en una muestra con un tiempo de decodificación antes de cualquier imagen que use este mensaje SEI.

[0101] Como también se ha analizado anteriormente, otros ejemplos de extensiones de formato de archivo ISO incluyen formatos de archivo SVC y MVC. SVC y MVC también pueden considerarse extensiones del formato de archivo AVC descrito anteriormente. De acuerdo con los formatos SVC y MVC, múltiples capas/vistas de un vídeo pueden estar contenidas en una sola pista de vídeo o en múltiples pistas de vídeo. Las especificaciones de formato de archivo SVC y MVC pueden proporcionar herramientas que permitan la construcción de puntos de operación cuando las capas/vistas se presentan en varias organizaciones.

[0102] El término "descarga progresiva" se usa para describir la transferencia de segmentos de medios digitales desde un servidor a un cliente, típicamente usando el protocolo HTTP. Cuando se inicia desde un ordenador, el usuario puede iniciar la reproducción de los medios antes de que finalice la descarga. Una diferencia entre la transmisión continua de medios y la descarga progresiva radica en cómo el dispositivo de usuario final que accede a los medios digitales recibe y almacena los datos de medios digitales. Un reproductor de medios que es capaz de una reproducción mediante descarga progresiva se basa en que los metadatos ubicados en la cabecera del archivo estén intactos y en un búfer local de la segmentación de medios digitales a medida que se descarga de un servidor web. En el momento en que una cantidad especificada de datos se vuelve disponible para el dispositivo de reproducción local, comenzará la reproducción de los medios. Esta cantidad especificada de búfer puede incorporarse en el archivo mediante el productor del contenido en los ajustes de codificador y puede reforzarse con ajustes de búfer adicionales impuestos por el reproductor de medios.

[0103] El uso de HTTP para la transmisión continua de vídeo tiene algunas ventajas, y los servicios de transmisión continua de vídeo basados en HTTP se están volviendo populares. Algunas ventajas de la transmisión continua de HTTP incluyen que puedan usarse componentes y protocolos de Internet existentes, de manera que no se necesitan nuevos esfuerzos para desarrollar nuevas técnicas para transportar datos de vídeo a través de una red. Otros protocolos de transporte, por ejemplo, el formato de carga útil RTP, requieren dispositivos de red intermedios, por ejemplo, cajas intermedias, para conocer el formato de medios y el contexto de señalización. Además, la transmisión continua de HTTP puede ser controlada por el cliente, lo que puede evitar problemas de control. Por ejemplo, para aprovechar todas las características para obtener un rendimiento óptimo, el servidor puede realizar un seguimiento

del tamaño y del contenido de paquetes cuya recepción aún no se ha acusado. El servidor también puede analizar la estructura de archivo y reconstruir el estado del búfer de cliente para tomar decisiones óptimas de conmutación/baja resolución de RD. Así mismo, las restricciones en las variaciones de flujo de bits pueden satisfacerse con el fin de seguir cumpliendo con los perfiles negociados. HTTP puede no requerir nuevas implementaciones de hardware o software en un servidor web que tenga implementado HTTP1.1. La transmisión continua de HTTP también puede proporcionar facilidad de manejo de TCP y cruce de cortafuegos. Las técnicas de esta divulgación pueden mejorar la transmisión continua de HTTP de datos de vídeo para resolver problemas relacionados con el ancho de banda, por ejemplo, proporcionando una adaptación de la velocidad binaria.

[0104] Normas de compresión vídeo, tales como ITU-T H.261, H.262, H.263, MPEG-1, MPEG-2 y H.264/MPEG-4 parte 10, hacen uso de la predicción temporal compensada por movimiento para reducir la redundancia temporal. El codificador usa una predicción compensada por movimiento de algunas imágenes previamente codificadas (también mencionadas en el presente documento como tramas) para predecir las imágenes codificadas actuales de acuerdo con los vectores de movimiento. Existen tres tipos de imagen principales en la codificación de vídeo típica. Se trata de imágenes intracodificadas ("imágenes I" o "tramas I"), imágenes predichas ("imágenes P" o "tramas P") e imágenes predichas bidireccionales ("imágenes B" o "tramas B"). Los bloques de imágenes P pueden intracodificarse o predecirse con referencia a otra imagen. En una imagen B, los bloques pueden predecirse a partir de una o dos imágenes de referencia, o pueden estar intracodificados. Estas imágenes de referencia pueden estar ubicadas antes o después de la imagen actual en orden temporal.

[0105] El formato de archivo 3GPP (3GPP) es otro ejemplo de una extensión del formato de archivo base de medios ISO. 3GPP es un formato de contenedor multimedia definido por el Proyecto de Colaboración de Tercera Generación (3GPP) para servicios multimedia UMTS 3G. Se usa típicamente en teléfonos móviles 3G y otros dispositivos con capacidad 3G, pero también se puede reproducir en algunos teléfonos 2G y 4G y otros diversos dispositivos. El formato de archivo 3GPP se basa en el formato de archivo multimedia base ISO. El 3GP más reciente se especifica en 3GPP TS26.244, "Transparent end-to-end packet switched streaming service (PSS); 3GPP file format (3GP)". El formato de archivo 3GPP almacena flujos de vídeo como MPEG-4 Parte 2 o H.263 o MPEG-4 Parte 10 (AVC/H.264). 3GPP permite el uso de códecs AMR y H.263 en el formato de archivo multimedia base ISO (MPEG-4 Parte 12), porque 3GPP especifica el uso de los campos de plantilla y entrada de muestra en el formato de archivo de medios base ISO, así como la definición de nuevas cajas a las que los códecs hacen referencia. Para el almacenamiento de información específica de medios MPEG-4 en archivos 3GP, la especificación 3GP hace referencia a MP4 y al formato de archivo AVC, que también se basan en el formato de archivo de medios base ISO. Las especificaciones de formato de archivo MP4 y AVC describen el uso de contenido MPEG-4 en el formato de archivo de medios base ISO.

[0106] El formato de archivo 3GPP puede admitir la transmisión continua de HTTP. En algunos ejemplos, el formato de archivo 3GPP puede incluir una descripción de presentación de medios MPD (por ejemplo, MPD 432 como se ilustra en la FIG. 4). La MPD puede incluir descripciones de diferentes representaciones alternativas (por ejemplo, servicios de vídeo de diferente calidad) y la descripción puede incluir, por ejemplo, información de código, perfil y nivel. Es posible que la MPD no forme parte de un archivo 3GPP que incluya datos de vídeo.

[0107] La FIG. 5 es un diagrama de bloques que ilustra un ejemplo de una representación 540 para un programa de vídeo (por ejemplo, una de las representaciones 440A-440C ilustradas en la FIG. 4 anterior). El ejemplo de la FIG. 5 ilustra una representación 540 de acuerdo con el formato de archivo 3GPP. La representación 540 de la FIG. 5 proporciona meramente con el propósito de explicar las técnicas de esta divulgación, y no pretende ser limitativa. Las técnicas de esta divulgación pueden aplicarse a cualquier representación digital de datos de vídeo, independientemente de que se mencionen explícitamente o no en el presente documento.

[0108] Como se muestra en la FIG. 5, la representación 540 incluye una caja de película (moov) 544, una caja de índices de segmento (sidx) 570 y uno o más fragmentos de película (moof) 546. Para facilitar la ilustración, solo se representa un fragmento de película 546 en la FIG. 5. Una representación de película típica 540 puede incluir una pluralidad de fragmentos de película 546, como se muestra en la FIG. 4.

[0109] La caja de película 544 puede incluir diversa información específica para la representación 540. Por ejemplo, la caja de película 544 puede indicar información tal como una velocidad binaria deseada para la transmisión continua de la representación de película. La caja de película 544 puede incluir, de forma adicional o alternativa, una o más indicaciones de información específica de los fragmentos 546. Por ejemplo, la caja de película 544 puede incluir información que indique un número de fragmentos 546 contenidos en la representación, un tiempo de reproducción correspondiente al inicio y/o final de uno o más fragmentos 546, y/o una o más indicaciones de intervalos de bits para los fragmentos 546 de la representación 540.

[0110] La representación 540 puede incluir además una o más cajas de índices de segmento 570. Una caja de índices de segmento 570 puede preceder a uno o más fragmentos 546 de la representación 540. En algunos ejemplos, una caja de índices de segmento 570 puede preceder a cada fragmento 546 de la representación 540 (véase, por ejemplo, la FIG. 4). En otros ejemplos, una caja de índices de segmento 570 puede preceder a cada segmento de una representación, donde una pluralidad de fragmentos están dispuestos en un segmento de película.

[0111] La caja de índices de segmento 570 puede configurarse para señalar información relacionada con la funcionalidad de acceso aleatorio para los fragmentos 546. Por ejemplo, de acuerdo con implementaciones conocidas, la caja de índices de segmento 570 puede indicar un punto de acceso IDR de un fragmento 546. La caja de índices de segmento 570 puede indicar además un desplazamiento de octetos (desplazamiento de referencia) de un fragmento que incluye un punto de acceso IDR (un fragmento de referencia). La caja de índices de segmento 570 puede indicar además información de temporización de una primera muestra de un fragmento de referencia (tiempo delta de referencia). La caja de índices de segmento puede indicar además si un fragmento de película actual (por ejemplo, un fragmento posterior a la caja de índices de segmento 570) incluye un punto de acceso aleatorio (contiene RAP). La caja de índices de segmento 570 puede indicar además una temporización diferente entre una primera muestra de un fragmento de película 546 y un primer punto de acceso aleatorio del fragmento de película 546.

[0112] El siguiente pseudocódigo es un ejemplo de código que puede usarse para implementar una caja de índices de segmento 570 de acuerdo con algunas técnicas. El siguiente pseudocódigo describe algunas variables que pueden usarse para una estructura de datos de caja de índices de segmento. Estas variables corresponden, en general, a variables con nombres similares descritas anteriormente.

```
aligned(8) class SegmentIndexBox extends FullBox('sidx', version, 0) {
    unsigned int(32) reference_track_ID;
    unsigned int(16) track_count;
    unsigned int(16) reference_count;
    for (i=1; i<= track_count; i++)
    {
        unsigned int(32) track_ID;
        if (version==0)
        {
            unsigned int(32) decode_time;
        } else
        {
            unsigned int(64) decode_time;
        }
    }
    for(i=1; i <= reference_count; i++)
    {
        bit (1) reference_type;
        unsigned int(31) reference_offset;
        unsigned int(32) reference_delta_time;
        bit(1) contains_rap;
        unsigned int(31) RAP_delta_time;
    }
}
```

Algunas implementaciones de una caja de índices de segmento 570 de una representación de película 540 pueden sufrir varias deficiencias. Por ejemplo, algunas implementaciones de una caja de índices de segmento 570 solo están configuradas para señalar puntos de acceso IDR para un fragmento de vídeo 546. Además, en algunos casos, un fragmento de película 546 puede comenzar con un trama que es un punto de acceso aleatorio IDR. En estos casos, las implementaciones de la caja de índices de segmento 570 no señalizan que la primera trama sea un punto de acceso IDR. En cambio, se sigue señalizando información de temporización de una diferencia entre una primera muestra de un fragmento y un primer punto de acceso aleatorio de la señal (por ejemplo, tiempo delta de referencia). Esto puede no ser deseable, ya que la caja de índices de segmento 570 puede comunicar información innecesaria en algunos casos.

[0113] Algunas implementaciones de una caja de índices de segmento 570 puede ser, además, no deseables debido a que un desplazamiento de octetos de señal para un RAP IDR se presenta con relación al punto en que comienza (primera trama) un fragmento de película. De este modo, en algunos casos, puede ser necesario que todas las tramas de un fragmento 546 sean descargadas por un dispositivo cliente, incluso cuando se produce un RAP IDR en algún momento después de la primera trama del fragmento de vídeo 546 aunque las muestras que preceden al RAP IDR no se puedan decodificar para su reproducción, desperdiciando así un valioso ancho de banda de una conexión de red.

[0114] Esta divulgación se refiere a técnicas mejoradas para la señalización de RAP ODR de una presentación de vídeo, por ejemplo, de un fragmento de vídeo de una representación de vídeo, y/u otra información relevante para escanear y/o buscar funcionalidad para una representación de vídeo. Esta divulgación está dirigida además a señalar un intervalo de octetos de tramas I de un fragmento de vídeo o representación de vídeo. Un dispositivo cliente puede usar un intervalo de octetos señalado para solicitar una o más tramas I señalizadas con el fin de proporcionar una versión de alta velocidad de una presentación de vídeo con fines de operaciones de avance o rebobinado rápidos.

[0115] Con propósitos explicativos, las técnicas de esta divulgación se describen tal como se aplican a una caja de índices de segmento 570 de una representación de vídeo de acuerdo con un formato de archivo 3GPP. Un experto en la técnica reconocerá que las técnicas de esta divulgación son aplicables a cualquier otro formato para la representación digital de una imagen en movimiento para su reproducción. Además, las técnicas de esta divulgación se describen en términos de señalización de puntos de acceso ODR, otra información y/o un intervalo de octetos de tramas I en información de cabecera específica para uno o más fragmentos de un archivo de segmentación. Un experto en la técnica reconocerá que información de cabecera como la descrita en el presente documento puede señalar puntos de acceso ODR, intervalos de octetos de tramas I u otra información en otros niveles de un archivo de segmentación, tal como en un archivo MPD (nivel de archivo de segmentación) general para una representación de película completa, en información de cabecera específica para una representación particular, o en información de cabecera específica para un segmento particular de una representación de película.

[0116] La FIG. 6 es un diagrama de bloques que ilustra diversa información que puede ser señalizada por una caja de índices de segmento 670 de una presentación de película (por ejemplo, la presentación 540 ilustrada en la FIG. 5) de conformidad con las técnicas de esta divulgación. La caja de índices de segmento 670 puede ser una estructura de datos configurada para almacenar información relevante para uno o más archivos de segmentación. En algunos ejemplos, la caja de índices de segmento 670 puede incluir un subconjunto de la información ilustrada en la FIG. 6, o puede incluir toda la información ilustrada en la FIG. 6. A continuación se presenta un ejemplo de pseudocódigo que puede utilizarse para implementar una caja de índices de segmento 670 de conformidad con las técnicas de esta divulgación:

```

aligned(8) class SegmentIndexBox extends FullBox('sidx', version, 0) {
    unsigned int(32) reference_track_ID;
    unsigned int(16) track_count;
    unsigned int(16) reference_count;
    for (i=1; i<= track_count; i++)
    {
        unsigned int(32) track_ID;
        if (version==0)
        {
            unsigned int(32) decode_time;
        } else
        {
            unsigned int(64) decode_time;
        }
    }
    for(i=1; i <= reference_count; i++)
    {
        bit (1) reference_type;
        int (2) contain_RAP;
        int (2) RAP_type;
        bit (2) RAP_data_offset_flags;
        int (2) reserved;
        unsigned int(32) reference_offset;
        unsigned int(32) reference_delta_time;
        if ( contain_RAP > 1 )
            unsigned int(32) RAP_delta_time;
        if ( contain_RAP > 0 && RAP_type !=0 )
        {
            unsigned int(32) delta_DT_PT;
            unsigned int(8) number_skip_samples;
        }
        if ( contain_RAP > 0 && RAP_type == 3)
            unsigned int(32) late_IDR_RAP_delta_time;
        if( RAP_data_offset_flags[0] || RAP_data_offset_flags[1] )
        {
            unsigned int(8) head_offset;
            if (RAP_data_offset_flags[0] && head_offset > 0 )

```

```

        unsigned int(32)    delta_offset;
        if (RAP_data_offset_flags[1] && head_offset > 0)
            unsigned int(32)    late_IDR_delta_offset;
    }
}
}

```

[0117] El pseudocódigo anterior define una pluralidad de variables que se pueden usar para señalar información relacionada con la transmisión continua de medios en una estructura de datos de caja de índices de segmento de un segmento de medios de conformidad con las técnicas de esta divulgación. Las variables anteriores corresponden, en general, a una señalización con nombres similares que se describe posteriormente con referencia a la FIG. 6.

[0118] En un ejemplo, el valor `contains_RAP` del pseudocódigo anterior se puede definir de acuerdo con la siguiente definición semántica: "contains_RAP: cuando la referencia es a un fragmento de película, este valor puede ser 2 si el fragmento de pista dentro de ese fragmento de película para la pista con ID `reference_track_id` contiene al menos un punto de acceso aleatorio; este valor puede ser 1 si el fragmento de pista dentro ese fragmento de película para la pista con ID `reference_track_id` contiene la primera muestra como un punto de acceso aleatorio; de lo contrario, este valor se establece en 0. Cuando la referencia es a un índice de segmento, este valor se establecerá en 1 o 2 solo si alguna de las referencias en ese índice de segmento tiene este valor establecido en 1 o 2, y 0 en caso contrario".

[0119] La siguiente Tabla 1 proporciona un conjunto de valores de ejemplo para `contains_RAP` y las implicaciones semánticas de cada valor.

TABLA 1

Valor de <code>contains_RAP</code>	Significado semántico
0	No hay ningún punto de acceso aleatorio incluido
1	Punto de acceso aleatorio incluido como la primera muestra del fragmento de película o el fragmento de pista
2	Punto de acceso aleatorio incluido, pero no como la primera muestra
3	Reservado

[0120] En un ejemplo, el valor `RAP_type` del pseudocódigo anterior se puede definir de acuerdo con la siguiente definición semántica: "Este valor especifica el tipo de puntos de acceso aleatorio contenidos en la pista referenciada del fragmento de película.

[0121] La siguiente Tabla 2 proporciona un conjunto de valores de ejemplo para `contains_RAP` y las implicaciones semánticas de cada valor.

TABLA 2

Valor de <code>RAP_type</code>	Significado semántico
00 (0)	IDR es el primer punto de acceso aleatorio incluido
01 (1)	ODR es el primer punto de acceso aleatorio incluido e IDR puede estar incluido, o no, después del ODR
10 (2)	ODR es el primer punto de acceso aleatorio incluido, pero IDR también se incluye después del ODR El primer IDR después del ODR se denomina IDR posterior y también se puede utilizar como un punto de acceso aleatorio
11 (3)	Reservado

[0122] En un ejemplo, como se ilustra en la FIG. 6, la caja de índices de segmento 670 incluye al menos una indicación de si un fragmento de vídeo incluye, o no, un punto de acceso aleatorio, 'incluye RAP' (`contains_RAP`) 601. 'Incluye RAP' 601 puede indicar si un fragmento de película incluye un punto de acceso IDR u ODR, o ambos. 'Incluye RAP' 601 puede indicar además si una primera muestra de un fragmento de película es un punto de acceso aleatorio. 'Incluye RAP' 601 puede señalar, en cambio, que un fragmento de película no incluye ningún punto de acceso aleatorio.

[0123] En un ejemplo específico, si un fragmento de película incluye al menos un punto de acceso aleatorio, a 'incluye RAP' 601 se le puede asignar un valor de 2. Si el fragmento de película incluye al menos un punto de acceso aleatorio que es la primera muestra del fragmento, a 'incluye RAP' 601 se le puede asignar un valor de 1. Si el fragmento de película no incluye ningún punto de acceso aleatorio, a 'incluye RAP' se le puede asignar un valor de 0.

[0124] Como también se ilustra en el ejemplo de la FIG. 6, la caja de índices de segmento 670 puede incluir al menos una indicación de un tipo de punto de acceso aleatorio, 'tipo de RAP' (RAP_type) 602. 'Tipo de RAP' 602 puede indicar si un punto de acceso IDR es el primer punto de acceso aleatorio contenido en un fragmento de vídeo. 'Tipo de RAP' 602 puede indicar además si un punto de acceso ODR es el primer punto de acceso aleatorio de un fragmento de vídeo. 'Tipo de RAP' 602 puede indicar además que un ODR es un primer punto de acceso aleatorio de un fragmento de vídeo, y que el fragmento de vídeo también incluye un IDR posterior al punto de acceso ODR. Un punto de acceso IDR subsiguiente a un punto de acceso ODR puede denominarse punto de acceso IDR posterior.

[0125] En un ejemplo específico, a 'tipo de RAP' 602 se le puede asignar un valor de 00 (0) si un punto de acceso IDR es el primer punto de acceso aleatorio contenido en un fragmento de vídeo. De acuerdo con este ejemplo, a 'tipo de RAP' 602 se le puede asignar un valor de 01 (1) si un punto de acceso ODR es el primer punto de acceso aleatorio de un fragmento de vídeo. También de acuerdo con este ejemplo, a 'tipo de RAP' 602 se le puede asignar un valor de 10 (2) si un ODR es un primer punto de acceso aleatorio de un fragmento de vídeo, y el fragmento de vídeo también incluye un IDR subsiguiente al punto de acceso ODR.

[0126] Como también se ilustra en el ejemplo de la FIG. 6, la caja de índices de segmento 670 puede incluir al menos una indicación de un desplazamiento de datos de punto de acceso aleatorio, 'banderas de desplazamiento de datos de RAP' (RAP_data_offset_flags) 603. 'Banderas de desplazamiento de datos de RAP' 603 puede indicar si un desplazamiento de octetos de un RAP debe señalizarse adicionalmente. Por ejemplo, 'banderas de desplazamiento de datos de RAP' 603 puede indicar cuándo un primer punto de acceso aleatorio de un fragmento de vídeo no es una primera muestra de una pista del fragmento. En un ejemplo específico, a 'banderas de desplazamiento de datos de RAP' 603 se les puede asignar un valor de 1 cuando un primer RAP no es la primera muestra de una pista referenciada de un fragmento de película. De acuerdo con este ejemplo, si un primer RAP es una primera muestra de una pista referenciada de un fragmento de película, a 'banderas de desplazamiento de datos de RAP' 603 se le puede asignar un valor de 0. En un ejemplo, a 'banderas de desplazamiento de datos de RAP' 603 solo se le puede asignar un valor de 1 cuando a 'tipo de RAP' 602 se le asigna un valor de 2 (por ejemplo, el fragmento de vídeo incluye un IDR después de un primer RAP del fragmento).

[0127] Como también se ilustra en el ejemplo de la FIG. 6, la caja de índices de segmento 670 puede incluir al menos una indicación de una diferencia entre un tiempo de descomposición (tiempo de decodificación) y un tiempo de presentación de un RAP ODR, 'delta dt pt' 604. En un ejemplo específico, a 'delta dt pt' se le puede asignar un valor numérico que representa un número de minutos, segundos, ciclos de reloj u otra indicación de una diferencia de temporización entre un tiempo de descomposición y un tiempo de presentación de un RAP ODR.

[0128] Como también se ilustra en el ejemplo de la FIG. 6, una caja de índices de segmento 670 puede incluir al menos una indicación de un número de muestras (tramas) de un fragmento de vídeo o presentación con un tiempo de presentación antes de y un tiempo de descomposición (decodificación) después de un RAP ODR, 'número de muestras omitidas' (number_skip_samples) 605. En un ejemplo específico, se puede asignar un valor numérico a 'número de muestras omitidas' que representa un número de muestras con un tiempo de presentación antes de y un tiempo de descomposición después de un RAP ODR.

[0129] Como también se ilustra en el ejemplo de la FIG. 6, una caja de índices de segmento 670 puede incluir al menos una indicación de una diferencia de tiempo de decodificación entre un primer RAP IDR y una primera muestra de un fragmento de vídeo, 'tiempo delta de RAP IDR posterior' (late_IDR_RAP_delta_time) 606. Si a 'tipo de RAP' 602 se le asigna un valor de 2, a 'tiempo delta de RAP IDR posterior' 606 se le puede asignar un valor numérico que indica una diferencia en el tiempo de decodificación (por ejemplo, un número de tramas, intervalo de octetos, número de ciclos de reloj u otra indicación de una diferencia de tiempo de decodificación) entre un primer RAP IDR y una primera muestra de un fragmento de vídeo.

[0130] Como también se ilustra en el ejemplo de la FIG. 6, una caja de índices de segmento 670 puede incluir al menos una indicación de un desplazamiento de un primer octeto de un fragmento de película y un primer octeto de una primera muestra del fragmento de película, 'desplazamiento de cabeza' (head_offset) 607. 'Desplazamiento de cabeza' 607 puede indicar una longitud de octetos de información de cabecera de un fragmento de vídeo. En un ejemplo específico, a 'desplazamiento de cabeza' 607 se le puede asignar un valor numérico que indica un desplazamiento de octetos entre un primer octeto de un fragmento de vídeo y una primera trama del fragmento de vídeo, por ejemplo, un número de octetos de una cabecera o un intervalo de octetos de la cabecera. En un ejemplo, a 'desplazamiento de cabeza' 607 se le puede asignar un valor de 0 si un fragmento de película incluye más de una pista.

[0131] Como también se ilustra en el ejemplo de la FIG. 6, una caja de índices de segmento 670 puede incluir al menos una indicación de un desplazamiento de octetos entre un primer octeto de un punto de acceso aleatorio y el primer octeto de un fragmento de película, 'desplazamiento delta' (delta_offset) 608. En un ejemplo específico, a 'desplazamiento delta' 607 se le puede asignar un valor numérico que indica un desplazamiento de octetos entre un primer octeto de un punto de acceso aleatorio y el primer octeto de un fragmento de película, por ejemplo, un número de octetos o un intervalo de octetos entre un primer octeto de un punto de acceso aleatorio y un primer octeto de un fragmento de película.

[0132] Como también se ilustra en el ejemplo de la FIG. 6, una caja de índices de segmento 670 puede incluir al menos una indicación de un desplazamiento de octetos entre un primer octeto de un RAP IDR posterior y el primer octeto de un fragmento de película, 'desplazamiento delta de IDR posterior' (late_IDR_delta_offset) 609. En un ejemplo específico, a 'desplazamiento delta de IDR posterior' 609 se le puede asignar un valor numérico que indica un número de octetos entre, o un intervalo de octetos, que indica una diferencia entre un primer octeto de un RAP IDR posterior y un primer octeto de un fragmento de película.

[0133] La implementación de 'desplazamiento de cabeza' 607, 'desplazamiento delta' 608 y/o 'desplazamiento delta de IDR posterior' 609 como se muestra en la FIG. 6 puede ser ventajosa, ya que se pueden usar menos bits (por ejemplo, 8 bits) para señalar un desplazamiento de cabeza (por ejemplo, número de octetos de información de cabecera de un fragmento de película o presentación). De acuerdo con técnicas anteriores, se pueden usar 32 bits para señalar un desplazamiento de cabeza para un fragmento de vídeo, a fin de tener en cuenta los casos en los que hay información adicional (por ejemplo, cajas) en un fragmento de película y/o en los que un fragmento de película incluye múltiples pistas.

[0134] Como también se ilustra en el ejemplo de la FIG. 6, una caja de índices de segmento 670 puede incluir además, o en cambio, al menos una indicación de un intervalo de octetos para una o más tramas I de un fragmento, 'intervalo de octetos de tramas I' 610. 'Intervalo de octetos de tramas I' 610 puede indicar un intervalo de octetos de tramas I a un dispositivo cliente de modo que el dispositivo cliente pueda solicitar, recibir y/o decodificar de forma selectiva tramas I para proporcionar una versión de alta velocidad (por ejemplo, avance rápido, rebobinado rápido) de una presentación de vídeo. En algunos ejemplos, un dispositivo cliente puede solicitar, recibir y/o decodificar solamente tramas I indicadas del fragmento. De acuerdo con este ejemplo, un dispositivo cliente puede presentar una versión de alta velocidad de una presentación de vídeo, sin solicitar, recibir y/o decodificar todas las tramas (por ejemplo, tramas P o tramas B) de un fragmento de vídeo, reduciendo así la cantidad de datos transferidos para presentar una versión de alta velocidad del fragmento de vídeo.

[0135] Como también se ilustra en el ejemplo de la FIG. 6, una caja de índices de segmento 670 puede incluir además, o en cambio, al menos una indicación del mayor desplazamiento para tramas I de un fragmento de vídeo, 'mayor desplazamiento de tramas I' 611. 'Mayor desplazamiento de tramas I' 611 puede describir, para uno o más fragmentos de una representación de vídeo, un intervalo de octetos del fragmento para el que el fragmento probablemente incluye al menos una trama I. En un ejemplo, un dispositivo cliente puede recibir el 'mayor desplazamiento de tramas I' 611 y solicitar solamente la porción de un fragmento de vídeo indicado por el 'mayor desplazamiento de tramas I' 611. De acuerdo con este ejemplo, el dispositivo cliente puede recibir tramas del intervalo de octetos solicitado y decodificar al menos las tramas I del intervalo de octetos solicitado para proporcionar una versión de alta velocidad (por ejemplo, avance rápido, rebobinado rápido) de una representación de vídeo que incluye el fragmento.

[0136] La diversa información incluida en una caja de índices de segmento 670 como la descrita con respecto a la FIG. 6 puede ser utilizada por un dispositivo cliente HTTP para adquirir medios con el propósito de una reproducción mediante transmisión continua. Por ejemplo, un dispositivo servidor HTTP puede hacer que una o más cajas de índices de segmento 670 para el conjunto de una presentación de película estén disponibles para su acceso por parte de un dispositivo cliente. El dispositivo cliente puede revisar la información de la caja de índices de segmento 670 y usar esa información para determinar una solicitud referente a una transmisión continua de medios. Por ejemplo, un dispositivo cliente puede determinar, basándose en la información de la caja de índices de segmento 670, en qué ubicaciones temporales (por ejemplo, punto(s) de acceso) de una presentación de vídeo puede comenzar la transmisión continua /decodificación/reproducción. En respuesta a la determinación de uno o más puntos de acceso, el dispositivo cliente puede iniciar una solicitud GET o GET parcial para la transmisión de datos. La solicitud GET o GET parcial puede incluir un fragmento, o una parte de un fragmento, que será transmitido por el dispositivo servidor al dispositivo cliente para su decodificación/reproducción.

[0137] Las técnicas de esta divulgación pueden proporcionar una granularidad mejorada para el inicio de la transmisión continua, la decodificación y/o la reproducción de una presentación de imágenes en movimiento. Por ejemplo, al señalar una identificación de uno o más puntos de acceso ODR para un fragmento de vídeo, un dispositivo cliente puede acceder a más puntos de inicio en una representación que las técnicas anteriores que solo permitían la señalización de puntos de acceso IDR y/o la señalización a nivel de trama (por ejemplo, por medio de mensajes SEI a nivel de imagen como se describe anteriormente). La señalización a nivel de trama de puntos de acceso ODR puede ser desventajosa porque puede ser necesario que un dispositivo cliente adquiera y decodifique un fragmento completo de una presentación de vídeo para determinar la ubicación de un punto de acceso ODR. De este modo, las técnicas de esta divulgación permiten una granularidad mejorada para la selección de un punto de acceso para aplicaciones de vídeo en transmisión continua, sin la transferencia y/o decodificación innecesarias de tramas de presentación de vídeo.

[0138] Además, las técnicas de esta divulgación pueden permitir dispositivos cliente de diferente complejidad elijan un punto de acceso IDR u ODR en un fragmento de vídeo que incluye puntos de acceso IDR y ODR. Por ejemplo, el campo 'tipo de RAP' 602 de una caja de índices de segmento 670 puede indicar a un dispositivo cliente si un fragmento contiene puntos de acceso IDR y ODR. Si un dispositivo cliente está configurado para comenzar la

recepción/decodificación/reproducción en un punto de acceso ODR, el dispositivo cliente puede elegir el punto de acceso ODR para comenzar la decodificación/reproducción. Sin embargo, si el dispositivo cliente solo está configurado para comenzar la decodificación/reproducción desde un punto de acceso IDR, el dispositivo cliente puede omitir el punto de acceso ODR y comenzar en el punto de acceso IDR.

[0139] Las diversas técnicas de esta divulgación pueden permitir además una reducción en la transmisión de datos innecesarios para la decodificación y/o la reproducción. Por ejemplo, al señalar una diferencia de temporización entre un tiempo de presentación y el tiempo de composición de una imagen ODR (por ejemplo, por medio de 'delta dt pt' 604 en el ejemplo de la FIG.6), puede omitirse la transmisión y/o decodificación de muestras antes de la imagen ODR. De este modo, un dispositivo cliente no necesita recibir/decodificar estas muestras, reduciendo así la cantidad de ancho de banda y/o potencia de procesamiento consumida para reproducir una presentación de vídeo.

[0140] Como otro ejemplo, la señalización de un desplazamiento de octetos de diversos aspectos de un fragmento de película (por ejemplo, el desplazamiento de cabeza 607, el desplazamiento delta 608 y/o el desplazamiento delta de IDR posterior 609) puede permitir que se omita la transmisión de datos correspondientes a muestras anteriores a un punto de acceso aleatorio, o de otros datos no necesarios para la decodificación.

[0141] Como otro ejemplo, mediante la señalización de un intervalo de octetos para las tramas-I de un fragmento de vídeo, (por ejemplo, intervalo de octetos de tramas I 610, mayor desplazamiento de tramas I 611), un dispositivo cliente puede solicitar solamente aquellas tramas deseables para proporcionar una versión de alta velocidad de una presentación de vídeo. Esta técnica puede ser beneficiosa en comparación con otras técnicas para proporcionar una versión de alta velocidad de una representación de vídeo ya que no es necesario crear una representación específica con propósitos de reproducción a alta velocidad (por ejemplo, avance rápido, rebobinado rápido) de una presentación de vídeo. En cambio, se pueden utilizar tramas de una presentación existente para proporcionar una representación de alta velocidad. Además, se puede reducir la cantidad de datos necesarios para proporcionar una versión de alta velocidad de una representación de vídeo (por ejemplo, transmitida, recibida, decodificada).

[0142] A continuación se presenta otro ejemplo de pseudocódigo que puede utilizarse para implementar una caja de índices de segmento configurada para señalar información relacionada con puntos de acceso aleatorio de conformidad con las técnicas de esta divulgación:

```
aligned(8) class SegmentIndexBox extends FullBox('sidx', version, 0) {
    unsigned int(32) reference_track_ID;
    unsigned int(16) track_count;
    unsigned int(16) reference_count;
    for(i=1; i <= reference_count; i++)
    {
        bit(1) contains_RAP;
        bit(1) RAP_first_sample;
        unsigned int(1) RAP_type;
        unsigned int(2) RAP_data_offset_flags;
        int (3) reserved;
        if ( contains_RAP && !RAP_first_sample )
            unsigned int(32) RAP_delta_time;
        if ( RAP_type )
        {
            unsigned int(32) delta_PT;
            unsigned int(8) number_skip_samples;
        }
    }
}
```

[0143] De acuerdo con el ejemplo de pseudocódigo anterior, RAP_first_sample puede indicar si una primera muestra de una segmentación de medios es un RAP. Por ejemplo, RAP_first_sample puede tener un valor de 1 para indicar que la primera muestra de la entrada de referencia es un RAP. RAP_first_sample puede tener un valor de 0 para indicar que la primera muestra de la referencia no es un RAP. RAP_type puede indicar un tipo de RAP contenido en una segmentación de medios. Por ejemplo, RAP_type puede tener un valor de 0 para indicar que la segmentación de medios contiene un RAP IDR. RAP_type puede tener un valor de 1 para indicar que una segmentación de medios contiene un RAP ODR. Delta_PT puede indicar una diferencia de tiempo entre un tiempo de presentación (composición) y un tiempo de decodificación de una primera muestra después de un RAP ODR, en orden de presentación. Number_skip_samples puede indicar un número de muestras con un tiempo de presentación antes de y un tiempo de descomposición después de un RAP ODR, que puede ser un primer RAP de una segmentación de medios.

[0144] En otro ejemplo no ilustrado en la FIG. 6, se puede configurar un archivo MPD de una representación de película para señalar información relacionada con ODR y/o RAP IDR. Por ejemplo, un MPD puede incluir un elemento de sintaxis RAPAsODR, que puede comprender un valor de bandera binaria. El valor RAPAsODR puede establecerse para indicar si algunos puntos de acceso aleatorio son ODR o no. Es decir, en algunos ejemplos, el valor RAPAsODR puede establecerse en un valor de "verdadero" para indicar que al menos un punto de acceso aleatorio en la secuencia de video correspondiente a la MPD es una trama ODR, pero puede establecerse en un valor de "falso" para indicar que todos los puntos de acceso aleatorio son tramas IDR.

[0145] La FIG. 7 es un diagrama de flujo que ilustra de forma genérica un ejemplo de un procedimiento para hacer funcionar un dispositivo cliente HTTP de conformidad con las técnicas de esta divulgación. El procedimiento incluye recibir, en un dispositivo cliente (por ejemplo, el dispositivo de destino A/V 40 ilustrado en la FIG. 1) desde un dispositivo servidor (por ejemplo, el dispositivo de origen de A/V 20 ilustrado en la FIG. 1), una indicación de un punto de acceso aleatorio de actualización de decodificación abierta (por ejemplo, la trama 320I ilustrada en la FIG. 3) de una segmentación de medios (por ejemplo, el GOP abierto 310 ilustrado en la FIG. 3) (701). La información de codificación para al menos una trama indica que al menos una trama de la segmentación de medios que sigue a la trama RAP ODR en orden de decodificación puede no estar correctamente decodificada, y que cada trama de la segmentación de medios que sigue a la trama RAP ODR en orden de visualización puede estar correctamente decodificada sin depender del contenido de tramas anteriores al RAP ODR en orden de visualización. El procedimiento incluye además enviar, por parte del dispositivo cliente, al menos una solicitud al dispositivo servidor para recuperar un flujo de datos de medios comenzando con el RAP ODR indicado (702).

[0146] La FIG. 8 es un diagrama de flujo que ilustra un ejemplo de un procedimiento para hacer funcionar un dispositivo servidor HTTP de conformidad con las técnicas de esta divulgación. El procedimiento incluye recibir una secuencia de tramas de video codificadas que comprende una trama de punto de acceso aleatorio (RAP) de actualización de decodificación abierta (ODR) de una segmentación de medios (801). Al menos una trama de la segmentación de medios que sigue a la trama RAP ODR en orden de decodificación puede no estar correctamente decodificada, y cada trama de la segmentación de medios que sigue a la trama RAP ODR en orden de visualización puede estar correctamente decodificada sin depender del contenido de tramas anteriores al RAP ODR en orden de visualización. El procedimiento incluye además generar una indicación de una ubicación del RAP ODR (802). El procedimiento incluye además proporcionar la indicación (803).

[0147] La FIG. 9 es un diagrama de flujo que ilustra un ejemplo de un procedimiento para hacer funcionar un dispositivo cliente HTTP para proporcionar una versión de alta velocidad de una representación de video. El procedimiento incluye recibir una indicación de un intervalo de octetos para al menos una trama intercodificada (trama I) de un fragmento de video (901). El procedimiento incluye además enviar una solicitud a un dispositivo servidor para recibir la al menos una trama I (902). El procedimiento incluye además usar la al menos una trama I recibida para presentar una versión de alta velocidad de una representación de video que incluye la al menos una trama I (903).

[0148] La FIG. 10 es un diagrama de flujo que ilustra un ejemplo de un procedimiento de funcionamiento de un dispositivo servidor HTTP para proporcionar una indicación de un intervalo de octetos de una trama intercodificada (trama I). El procedimiento incluye recibir una secuencia de tramas de video codificadas que comprenden al menos una trama intercodificada (trama I) (1001). El procedimiento incluye además generar una indicación de un intervalo de octetos para la al menos una trama I (1002). El procedimiento incluye además proporcionar la indicación de un intervalo de octetos para la al menos una trama I (1003). El procedimiento puede incluir además recibir una solicitud, por ejemplo, una solicitud GET o GET parcial de HTTP, para transmitir la al menos una trama I en base a la indicación del intervalo de octetos. Un dispositivo cliente puede solicitar la al menos una trama I y usar la al menos una trama I para proporcionar una versión de alta velocidad de una presentación de video.

[0149] La FIG. 11 es un diagrama de flujo que ilustra un ejemplo de etapas que pueden realizar respectivos dispositivos cliente y servidor HTTP de conformidad con las técnicas de esta divulgación. El diagrama de flujo representado en la FIG. 11 se proporciona a modo de ejemplo sólo como un ejemplo de etapas que pueden realizar los respectivos dispositivos cliente y servidor de acuerdo con estas técnicas. Las técnicas de esta divulgación pueden utilizar, de forma adicional o alternativa, más o menos etapas que las representadas en la FIG. 11, o pueden realizar las etapas representadas en otros órdenes que los representados y descritos.

[0150] En general, el procedimiento de la FIG. 11 se describe con respecto a los elementos de la FIG. 1. Por ejemplo, las etapas indicadas como realizadas por el servidor pueden ser realizadas por el dispositivo de origen 20, mientras que las etapas indicadas como realizadas por el cliente pueden ser realizadas por el dispositivo de destino 40. Sin embargo, debe entenderse que estas indicaciones se proporcionan solo con fines de ejemplo. En general, el procedimiento de la FIG. 11 puede ser realizado por cualquier par de dispositivos cliente y servidor.

[0151] Inicialmente, tal como se muestra en la FIG. 11, la unidad de encapsulación 30 del dispositivo de origen 20 puede recibir datos de vídeo codificados (1101). El dispositivo de origen 20 puede recibir los datos de vídeo codificados por medio de una conexión de red, tal como Internet. El dispositivo de origen 20 puede recibir, de forma adicional o alternativa, los datos de vídeo codificados desde uno o más dispositivos de almacenamiento, internos o externos al dispositivo de origen 20. Los datos pueden estar codificados previamente, o el codificador de vídeo 28 puede codificar los datos sobre la marcha. Como otro ejemplo más, el dispositivo de origen 20 puede generar los datos, por ejemplo, en el contexto de información de vídeo generada por ordenador, tal como cuando se ejecuta un videojuego, y el codificador de vídeo 28 puede codificar los datos de vídeo generados. La unidad de encapsulación 30 puede entonces encapsular los datos de vídeo recibidos en fragmentos de vídeo (1102).

[0152] Además, la unidad de encapsulación 30 puede señalar una o más ubicaciones de uno o más RAP IDR y ODR de los fragmentos de vídeo (1103). Por ejemplo, la unidad de encapsulación 30 puede señalar un desplazamiento entre la primera trama de una segmentación de medios (por ejemplo, un fragmento de vídeo) que incluye un RAP ODR, y la trama correspondiente al RAP ODR. Cuando la primera trama de la segmentación de medios es el RAP ODR, la unidad de encapsulación 30 puede establecer una bandera o proporcionar otra información indicativa de que la primera trama es un RAP ODR. La unidad de encapsulación 30 puede señalar además varias muestras (por ejemplo, varias tramas) que tienen tiempos de presentación anteriores al RAP ODR. Asimismo, la unidad de encapsulación 30 puede señalar si la segmentación de medios incluye tanto un IDR como un ODR. Como otro ejemplo más, la unidad de encapsulación 30 puede señalar un desplazamiento de octetos correspondiente al RAP ODR.

[0153] El dispositivo de destino 40 puede solicitar posteriormente la información de señalización de punto de acceso aleatorio, por ejemplo emitiendo una solicitud de información de cabecera de una representación que incluye la segmentación de medios (1104). Por ejemplo, la solicitud puede ser una solicitud de un archivo de descripción de presentación de medios (MPD) como se describe anteriormente. En otro ejemplo, la solicitud puede ser una solicitud de una caja de índices de segmento (SIDX) de uno o más archivos de segmentación. En respuesta a la solicitud, el dispositivo de origen 20 envía las ubicaciones señalizadas de RAP IDR y ODR, como se creó en la etapa 1103 (1105).

[0154] El dispositivo de destino 40 recibe entonces la información de señal (1106). Posteriormente, en este ejemplo, el dispositivo de destino 40 puede recibir una solicitud para saltar a una ubicación temporal dentro de la representación de vídeo (1107). Por ejemplo, el dispositivo de destino puede recibir una indicación desde un usuario para saltar a una ubicación temporal (en el tiempo) particular para la reproducción de vídeo. En otro ejemplo, el dispositivo de destino puede iniciar por sí mismo un salto a una ubicación temporal, por ejemplo, donde el dispositivo de destino ha detectado un cambio en el ancho de banda para la transmisión continua u otro cambio de condición, e iniciar una solicitud para recibir una representación alternativa de un programa de vídeo en respuesta al cambio detectado en el ancho de banda u otro cambio de condición.

[0155] En respuesta a la solicitud de saltar a una ubicación temporal, el dispositivo de destino 40 puede determinar el punto de acceso aleatorio más cercano a la ubicación temporal (1108). Es decir, el dispositivo de destino 40 puede usar la información de señal RAP para identificar el RAP más cercano a la ubicación temporal solicitada. Después de seleccionar este RAP, el dispositivo de destino 40 puede generar una solicitud para iniciar un flujo de datos comenzando en el RAP determinado. Por ejemplo, el dispositivo de destino 40 puede generar una solicitud GET de HTTP para recuperar el fragmento completo, incluido el RAP, por ejemplo cuando el RAP es la primera trama del fragmento. Como otro ejemplo, el dispositivo de destino 40 puede usar la información señalizada para determinar una ubicación de octeto inicial del RAP dentro de una segmentación de medios. Supóngase, por ejemplo, que la segmentación de medios tiene una longitud de N octetos y que los datos del RAP comienzan en el octeto X, donde $0 < X < N$. El dispositivo de destino 40 puede generar una solicitud GET parcial de HTTP para la segmentación de medios especificando un intervalo de octetos de (X, N) para recuperar datos correspondientes al RAP hasta el final del archivo.

[0156] Además, varios dispositivos de destino pueden estar configurados para usar de diferente manera la información señalizada. Por ejemplo, algunos dispositivos de destino pueden ser lo suficientemente sofisticados como para seleccionar entre RAP IDR y ODR. Por otro lado, algunos dispositivos de destino solo pueden configurarse para admitir la generación de solicitudes para RAP IDR. Al señalar RAP IDR y ODR, las técnicas de esta divulgación proporcionan dispositivos más sofisticados con la capacidad de implementar una granularidad de selección más fina de un punto de inicio de un flujo, sin aislar o excluir dispositivos heredados o dispositivos que no están equipados para utilizar RAP ODR.

[0157] Después de generar la solicitud, el dispositivo de destino 40 puede enviar la solicitud al dispositivo de origen 20 (1109). Después de recibir la solicitud (1110), el dispositivo de origen 20 puede responder enviando datos de vídeo en base a la solicitud en forma de flujo al dispositivo de destino 40 (1111). Aunque se ilustra secuencialmente para facilitar la ilustración y el entendimiento, debe entenderse que el dispositivo de destino 40 puede generar múltiples solicitudes GET y/o GET parcial de HTTP para recibir el flujo completo de datos de vídeo. En cualquier caso, después de recibir los datos de vídeo, la unidad de desencapsulación 38 puede desencapsular los datos de vídeo, el decodificador de vídeo 48 puede decodificar los datos y la salida de vídeo 44 puede mostrar los datos de vídeo (1112).

[0158] La FIG. 12 es un diagrama de flujo que ilustra un ejemplo de etapas que pueden realizar respectivos dispositivos cliente y servidor HTTP de conformidad con las técnicas de esta divulgación. El diagrama de flujo representado en la FIG. 12 se proporciona a modo de ejemplo sólo como un ejemplo de etapas que pueden realizar los respectivos dispositivos cliente y servidor de acuerdo con estas técnicas. Las técnicas de esta divulgación pueden utilizar, de forma adicional o alternativa, más o menos etapas que las representadas en la FIG. 12, o pueden realizar las etapas representadas en otros órdenes que los representados y descritos.

[0159] En general, el procedimiento de la FIG. 12 se describe con respecto a los elementos de la FIG. 1. Por ejemplo, las etapas indicadas como realizadas por el servidor pueden ser realizadas por el dispositivo de origen 20, mientras que las etapas indicadas como realizadas por el cliente pueden ser realizadas por el dispositivo de destino 40. Sin embargo, debe entenderse que estas indicaciones se proporcionan solo con fines de ejemplo. En general, el procedimiento de la FIG. 12 puede ser realizado por cualquier par de dispositivos cliente y servidor.

[0160] Inicialmente, tal como se muestra en la FIG. 12, la unidad de encapsulación 30 del dispositivo de origen 20 puede recibir datos de vídeo codificados (1201). El dispositivo de origen 20 puede recibir los datos de vídeo codificados usando una conexión de red tal como Internet. El dispositivo de origen 20 puede recibir, de forma adicional o alternativa, los datos de vídeo codificados desde uno o más dispositivos de almacenamiento, internos o externos al dispositivo servidor. Los datos pueden estar codificados previamente, o el codificador de vídeo 28 puede codificar los datos sobre la marcha. Como otro ejemplo más, el dispositivo de origen 20 puede generar los datos, por ejemplo, en el contexto de información de vídeo generada por ordenador, tal como cuando se ejecuta un videojuego, y el codificador de vídeo 28 puede codificar los datos de vídeo generados. La unidad de encapsulación 30 puede entonces encapsular los datos de vídeo recibidos en fragmentos de vídeo (1202).

[0161] Además, la unidad de encapsulación 30 puede señalar al menos un intervalo de octetos dentro de los datos de vídeo codificados que incluye al menos una trama intracodificada (trama I) (1203). Por ejemplo, la unidad de encapsulación 30 puede señalar un intervalo de octetos para una o más tramas I (por ejemplo, un intervalo de octetos de la trama I). En uno de tales ejemplos, la unidad de encapsulación 30 puede señalar un punto de inicio para una trama I señalizando un desplazamiento de octetos desde un primer octeto de la presentación de vídeo codificada. De acuerdo con este ejemplo, la unidad de encapsulación 30 puede señalar además un punto de extremo para una trama I señalizando un desplazamiento de octetos desde un primer octeto de la presentación de vídeo codificada, o señalizando un desplazamiento de octetos desde el punto de inicio de la trama I. En otro ejemplo, la unidad de encapsulación 30 puede señalar un intervalo de octetos de los datos de vídeo codificados que incluye la al menos una trama I, así como otras tramas intercodificadas (por ejemplo, tramas B y P).

[0162] El dispositivo de destino 40 puede enviar una solicitud al dispositivo de origen 20 referente a información de señal de trama I (1204). La solicitud de información de señal de trama I puede incluir una solicitud referente a una parte de o toda la información de cabecera para uno o más archivos de segmentación. Por ejemplo, la solicitud puede ser una solicitud de un archivo de descripción de presentación de medios (MPD) como se describe anteriormente. En otro ejemplo, la solicitud puede ser una solicitud de una caja de índices de segmento (SIDX) de uno o más archivos de segmentación. En respuesta a la solicitud, el dispositivo de origen 20 puede enviar el uno o más intervalos de octetos señalizados que incluyen al menos una trama I de la etapa 1203 (1205). El dispositivo de destino 40 puede recibir la información de señal de trama I (1206). El dispositivo de destino 40 puede recibir (por ejemplo, desde un usuario o desde uno o más programas de software) una solicitud para reproducir una versión a alta velocidad de una presentación de vídeo (por ejemplo, una versión de avance rápido, de rebobinado o de rebobinado rápido de la presentación de vídeo) (1207). En un ejemplo, la solicitud puede incluir una indicación de un usuario para modificar la reproducción de una representación de vídeo para presentar una versión a alta velocidad de la presentación de vídeo.

[0163] El dispositivo de destino puede determinar tramas I para la reproducción a alta velocidad (1208). Por ejemplo, si el dispositivo de origen 20 señala el al menos un intervalo de octetos en la etapa 1203 mediante la señalización de un intervalo de octetos para una o más tramas I (por ejemplo, un intervalo de octetos de la trama I), el dispositivo de destino 40 puede determinar una o más más tramas I para una reproducción a alta velocidad determinando cuál de una pluralidad de tramas I para las que se señala un intervalo de octetos se usará (por ejemplo, transmitirá, decodificará) para su reproducción a alta velocidad. En otro ejemplo, si el dispositivo de origen señaló un intervalo de octetos que incluye al menos una trama I, así como otras tramas intercodificadas (por ejemplo, tramas B y P), el dispositivo de destino puede determinar a partir del uno o más intervalos de octetos señalizados tramas I que se pueden usar para la reproducción a alta velocidad. El dispositivo de destino 40 puede enviar a un dispositivo de origen 20 una solicitud para transmitir en continuo datos de medios que incluyen la una o más tramas I determinadas en la etapa 1208 (1209). La solicitud puede ser una solicitud GET parcial que identifica un archivo de segmentación (por

ejemplo, fragmento de película, segmento y/o representación) y un intervalo de octetos deseado del archivo de segmentación que se transmitirá en continuo en respuesta a la solicitud.

[0164] El dispositivo de origen 20 puede recibir la solicitud para transmitir en continuo datos de medios (1210). En respuesta a la solicitud, el dispositivo de origen 20 puede transmitir en continuo datos de medios en base a la solicitud de flujo (1211). El dispositivo de destino 40 puede decodificar los datos de vídeo recibidos en respuesta a la solicitud y mostrar los datos de vídeo recibidos (1212) como una versión a alta velocidad (por ejemplo, avance rápido, rebobinado, rebobinado rápido) de la representación de vídeo. En algunos ejemplos, el dispositivo de origen puede decodificar solamente datos de medios determinados en la etapa 1208 para su uso para proporcionar una versión a alta velocidad de una representación de vídeo. En un ejemplo, el dispositivo de destino 40 puede presentar la versión a alta velocidad de la representación de vídeo presentando consecutivamente dos o más tramas l de datos de medios recuperados que están separados temporalmente por al menos una trama intercodificada en los datos de medios codificados recibidos por el servidor en la etapa 1201.

[0165] En uno o más ejemplos, las funciones descritas se pueden implementar en hardware, software, firmware o cualquier combinación de los mismos. Si se implementan en software, las funciones pueden almacenarse en, o transmitirse por, un medio legible por ordenador como una o más instrucciones o código, y ejecutarse mediante una unidad de procesamiento basada en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador que correspondan a un medio tangible, tales como medios de almacenamiento de datos, o medios de comunicación que incluyan cualquier medio que facilite la transferencia de un programa informático de un lugar a otro, por ejemplo, de acuerdo con un protocolo de comunicación. De esta manera, los medios legibles por ordenador pueden corresponder, en general, a (1) medios de almacenamiento tangibles legibles por ordenador que sean no transitorios o a (2) un medio de comunicación tal como una señal o una onda portadora. Los medios de almacenamiento de datos pueden ser cualquier medio disponible al que se pueda acceder mediante uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en esta divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

[0166] A modo de ejemplo, y no de limitación, dichos medios de almacenamiento legibles por ordenador pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento de disco óptico, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que se pueda usar para almacenar un código de programa deseado en forma de instrucciones o estructuras de datos y al que se pueda acceder mediante un ordenador. Además, cualquier conexión recibe apropiadamente la denominación de medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde un sitio web, un servidor u otra fuente remota usando un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas tales como infrarrojos, radio y microondas están incluidos en la definición de medio. Sin embargo, se debe entender que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales ni otros medios transitorios, sino que, en cambio, se refieren a medios de almacenamiento tangibles no transitorios. Los discos, como se usan en el presente documento, incluyen el disco compacto (CD), el disco láser, el disco óptico, el disco versátil digital (DVD), el disco flexible y el disco Blu-ray, donde algunos discos reproducen habitualmente datos magnéticamente y otros discos reproducen datos ópticamente con láseres. Las combinaciones de lo anterior también se deben incluir dentro del alcance de los medios legibles por ordenador.

[0167] Las instrucciones pueden ser ejecutadas por uno o más procesadores, tales como una o más unidades centrales de procesamiento (CPU), procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados específicos de la aplicación (ASIC), formaciones lógicas programables *in situ* (FPGA) u otros circuitos lógicos equivalentes, integrados o discretos. Por consiguiente, el término "procesador", como se usa en el presente documento, se puede referir a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento se puede proporcionar dentro de módulos de hardware y/o de software dedicados configurados para la codificación y la decodificación, o incorporarse en un códec combinado. Además, las técnicas se podrían implementar por completo en uno o más circuitos o elementos lógicos.

[0168] Las técnicas de esta divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, incluidos un teléfono inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). En esta divulgación se describen diversos componentes, módulos o unidades para destacar aspectos funcionales de dispositivos configurados para realizar las técnicas divulgadas, pero no se requiere necesariamente su realización mediante diferentes unidades de hardware. En cambio, como se describe anteriormente, diversas unidades se pueden combinar en una unidad de hardware de códec o proporcionar mediante un grupo de unidades de hardware interoperativas, que incluyen uno o más procesadores descritos anteriormente, junto con software y/o firmware adecuados.

REIVINDICACIONES

1. Un procedimiento, que comprende:

recibir (901, 1206), en un dispositivo cliente desde un dispositivo servidor, un desplazamiento máximo de octetos, indicando el desplazamiento máximo de octetos un intervalo de octetos para el cual una pluralidad de segmentaciones de medios de una presentación de vídeo incluye al menos una trama intracodificada, trama I, donde la presentación de vídeo está configurada para mostrarse a una primera velocidad de reproducción, donde el desplazamiento máximo de octetos comprende un desplazamiento desde el comienzo de una segmentación de medios, describiendo el desplazamiento el intervalo de octetos, y donde una segmentación de medios comprende un archivo de medios que se puede recuperar de acuerdo con una solicitud GET del protocolo de transferencia de hipertexto, HTTP;

enviar (902, 1209), mediante el dispositivo cliente, para la pluralidad de segmentaciones de medios, solicitudes GET de HTTP al dispositivo servidor, incluyendo cada solicitud GET de HTTP un localizador de recursos uniforme, URL, de la segmentación de medios y el intervalo de octetos, con el fin de recuperar datos de medios dentro del intervalo de octetos para la pluralidad de segmentaciones de medios; y

usar (903, 1212) los datos de medios recuperados que incluyen al menos una trama intracodificada por segmentación de medios para presentar una versión a alta velocidad de la presentación de vídeo que incluye solamente tramas I a una segunda velocidad de reproducción superior a la primera velocidad de reproducción.

2. El procedimiento de la reivindicación 1, en el que usar los datos de medios recuperados para presentar una versión a alta velocidad de una presentación de vídeo comprende presentar de forma consecutiva dos o más tramas I de los datos de medios recuperados, donde las dos o más tramas I están separadas temporalmente por al menos una trama intercodificada.

3. El procedimiento de la reivindicación 1, en el que usar los datos de medios recuperados para presentar una versión a alta velocidad de una presentación de vídeo comprende presentar una versión de avance rápido o de rebobinado rápido de la presentación de vídeo.

4. Un dispositivo, que comprende:

medios para recibir, en un dispositivo cliente desde un dispositivo servidor, un desplazamiento máximo de octetos, indicando el desplazamiento máximo de octetos un intervalo de octetos para el cual una pluralidad de segmentaciones de medios de una presentación de vídeo incluye al menos una trama intracodificada, trama I, donde la presentación de vídeo está configurada para mostrarse a una primera velocidad de reproducción, donde el desplazamiento máximo de octetos comprende un desplazamiento desde el comienzo de una segmentación de medios, describiendo el desplazamiento el intervalo de octetos, y donde una segmentación de medios comprende un archivo de medios que se puede recuperar de acuerdo con una solicitud GET del protocolo de transferencia de hipertexto, HTTP;

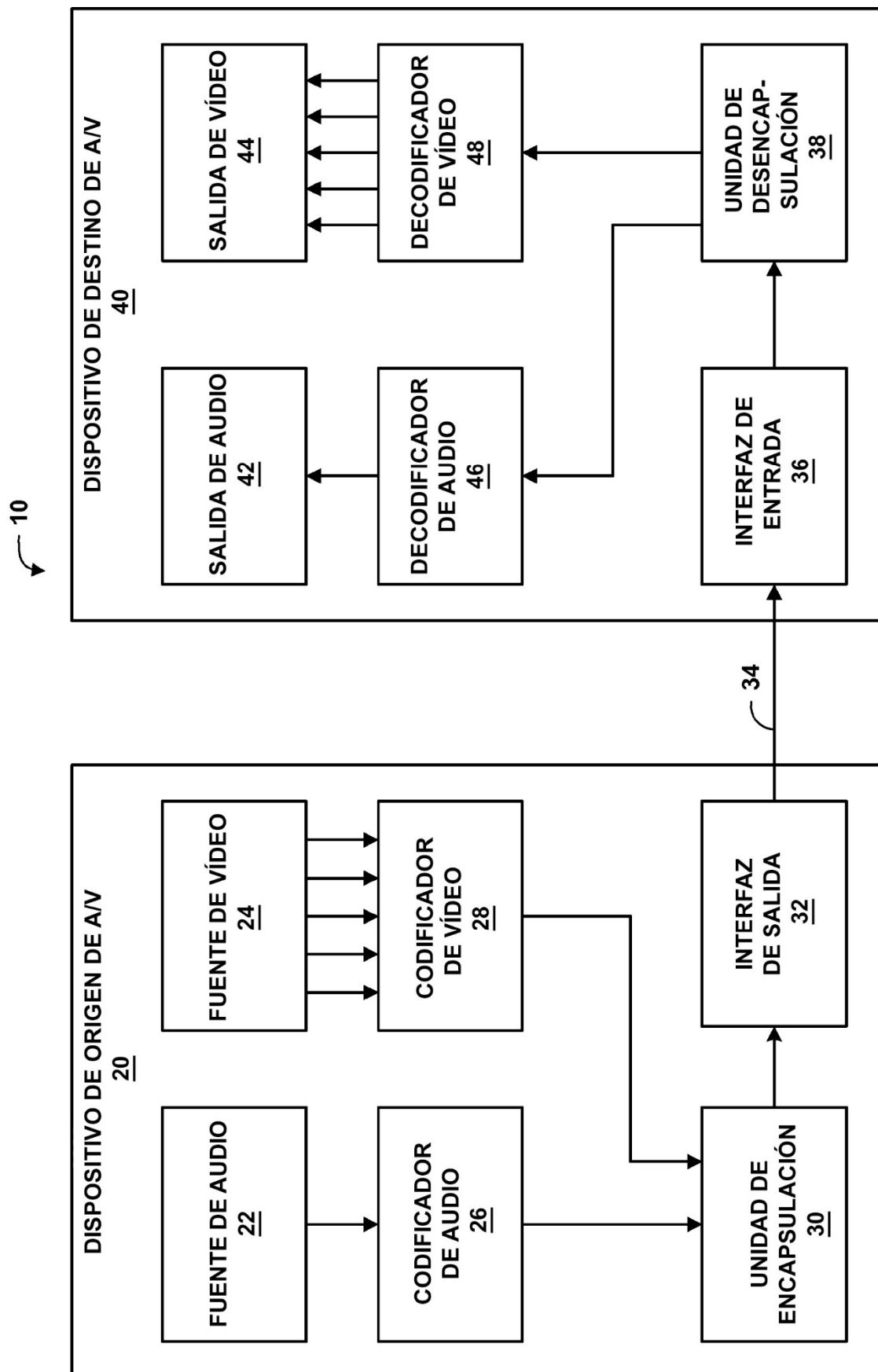
medios para enviar, mediante el dispositivo cliente, para la pluralidad de segmentaciones de medios, solicitudes GET del protocolo de transferencia de hipertexto, HTTP, al dispositivo servidor, incluyendo cada una un localizador de recursos uniforme, URL, de la segmentación de medios y el intervalo de octetos, con el fin de recuperar datos de medios dentro del intervalo de octetos para la pluralidad de segmentaciones de medios; y

medios para usar los datos de medios recuperados que incluyen al menos una trama intracodificada por segmentación de medios para presentar una versión a alta velocidad de la presentación de vídeo a una segunda velocidad de reproducción superior a la primera velocidad de reproducción.

5. El dispositivo de la reivindicación 4, en el que usar los datos de medios recuperados para presentar una versión a alta velocidad de una presentación de vídeo comprende presentar de forma consecutiva dos o más tramas I de los datos de medios recuperados, donde las dos o más tramas I están separadas temporalmente por al menos una trama intercodificada.

6. El dispositivo de la reivindicación 4, en el que usar los datos de medios recuperados para presentar una versión a alta velocidad de una presentación de vídeo comprende presentar una versión de avance rápido o de rebobinado rápido de la presentación de vídeo.

7. Un medio de almacenamiento legible por ordenador que almacena instrucciones configuradas para hacer que un dispositivo servidor realice el procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 3.



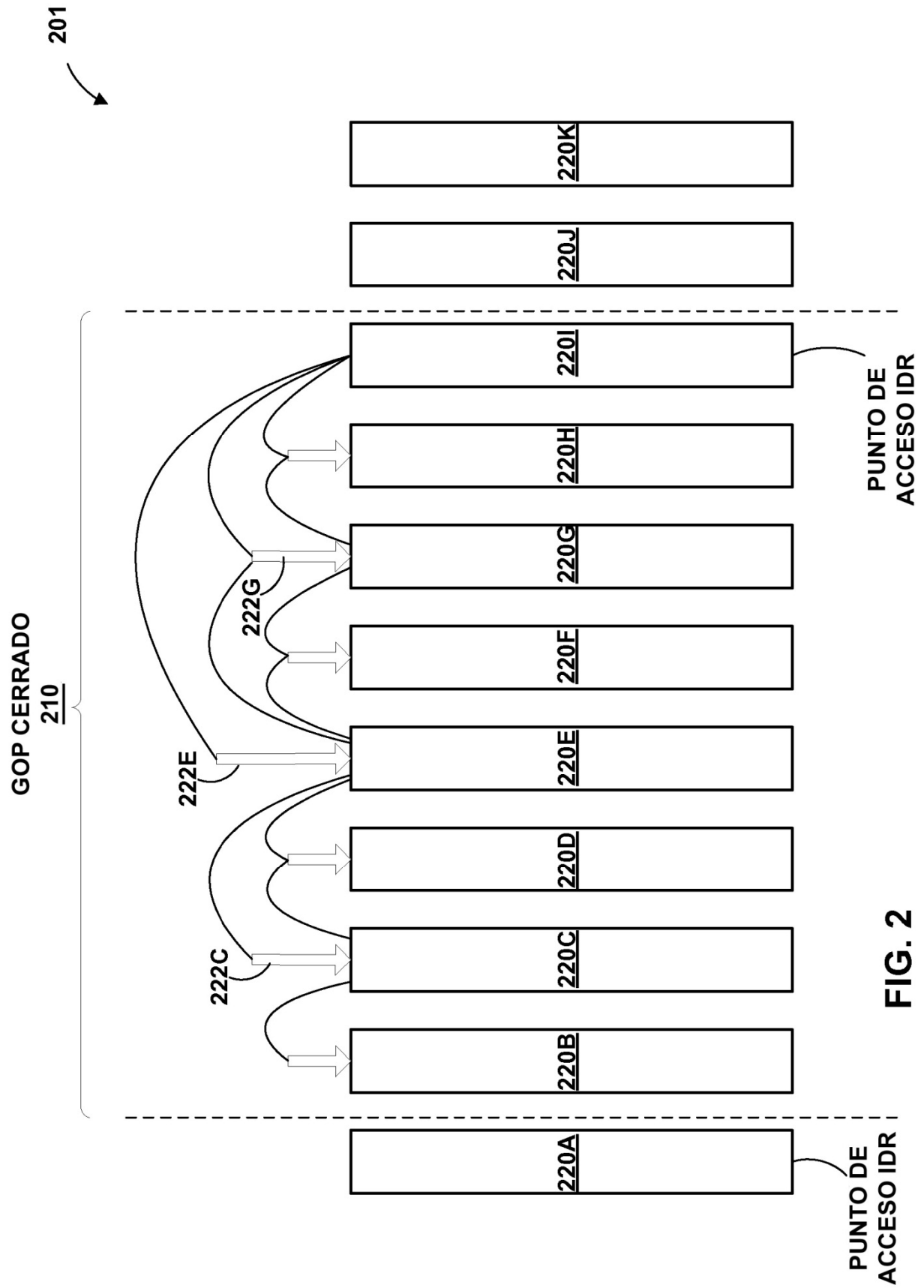


FIG. 2

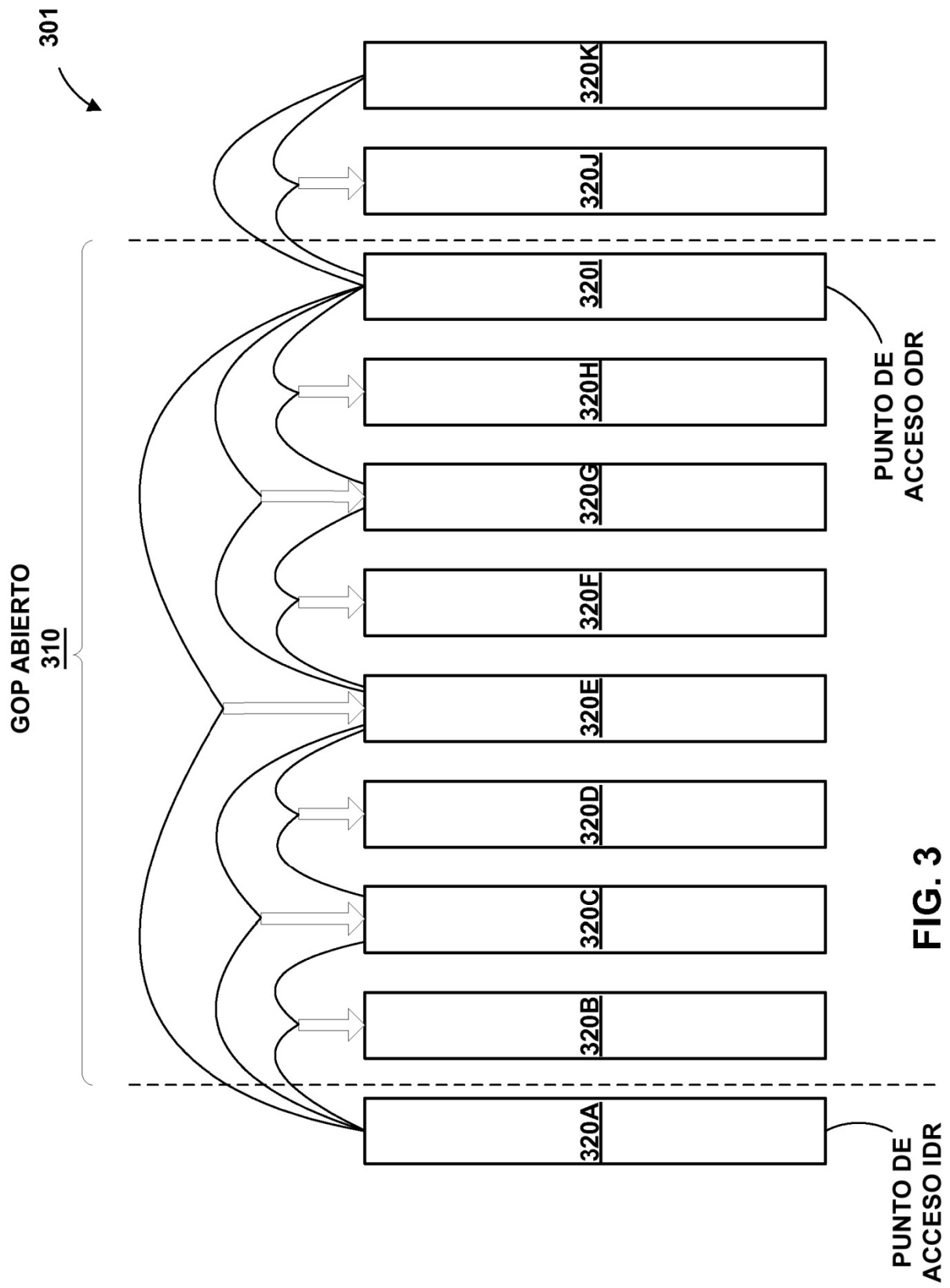


FIG. 3

PROGRAMA
430

**DESCRIPCIÓN DE
PRESENTACIÓN DE
MEDIOS (MPD)**
432

PRESENTACIÓN 1
440A

CABECERA DE
FRAGMENTO
443

FRAGMENTO 1
442

CABECERA DE
FRAGMENTO
443

FRAGMENTO 2
442

CABECERA DE
FRAGMENTO
443

FRAGMENTO N
442

PRESENTACIÓN 2
440B

CABECERA DE
FRAGMENTO
443

FRAGMENTO 1
442

CABECERA DE
FRAGMENTO
443

FRAGMENTO 2
442

CABECERA DE
FRAGMENTO
443

FRAGMENTO N
442

PRESENTACIÓN N
440C

CABECERA DE
FRAGMENTO
443

FRAGMENTO 1
442

CABECERA DE
FRAGMENTO
443

FRAGMENTO 2
442

CABECERA DE
FRAGMENTO
443

FRAGMENTO N
442

FIG. 4

PRESENTACIÓN
540

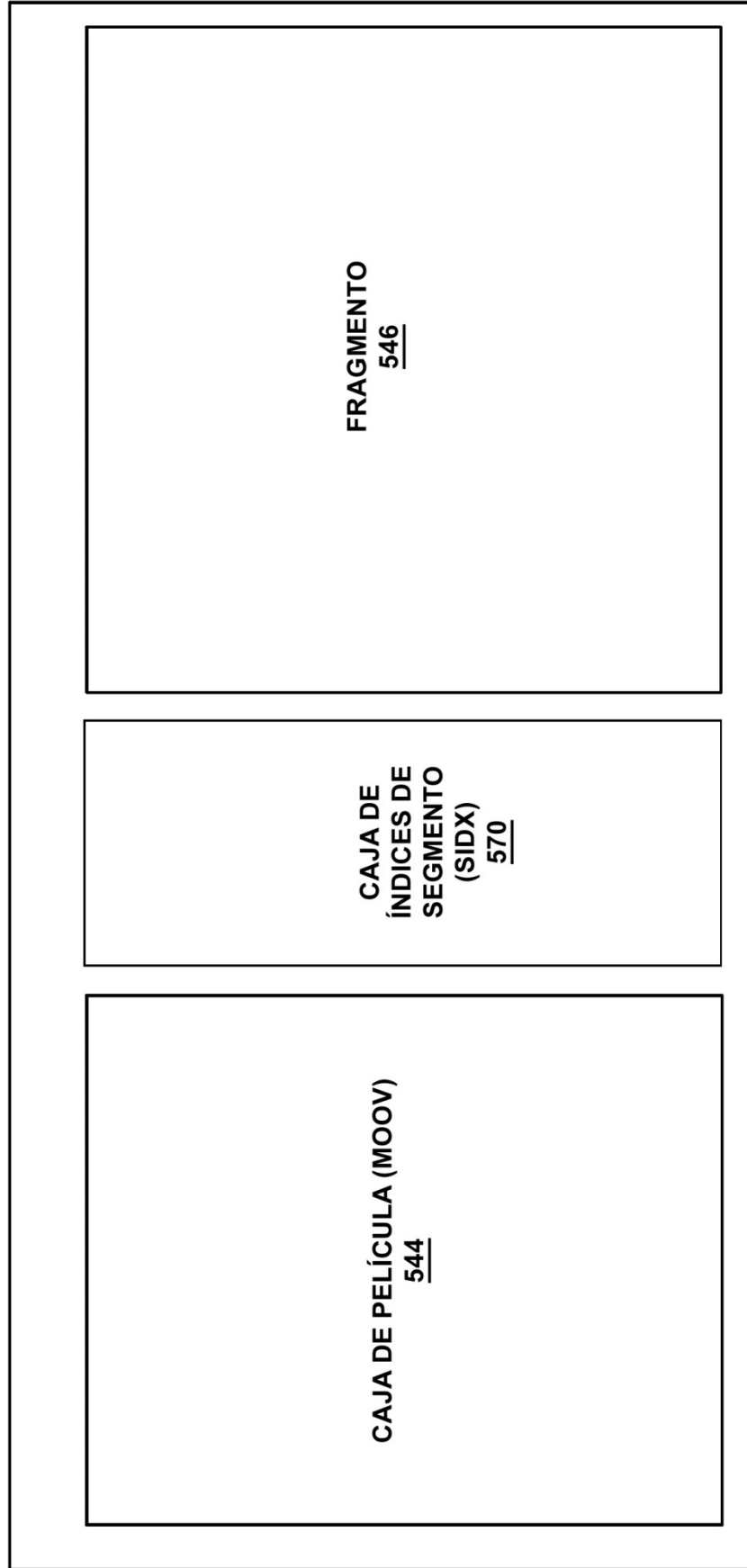


FIG. 5

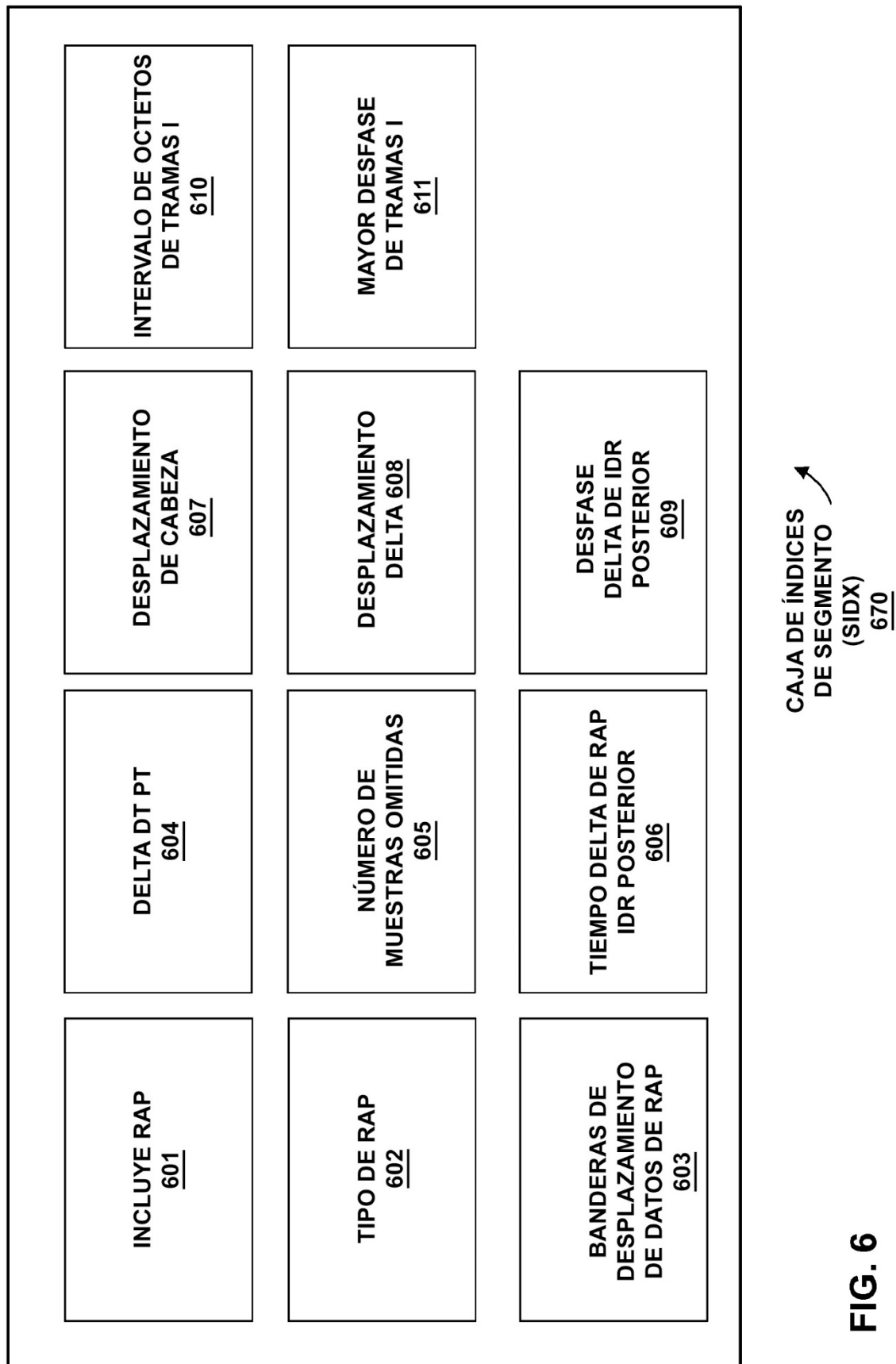


FIG. 6



FIG. 7

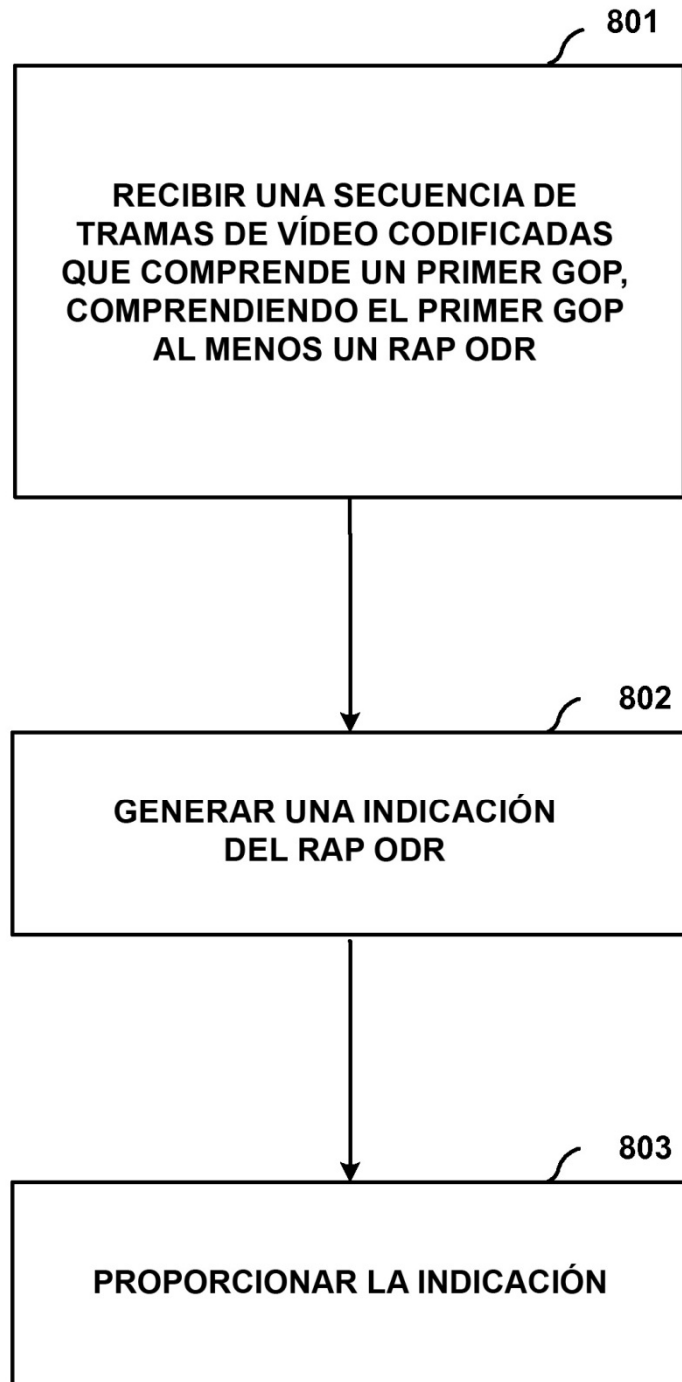


FIG. 8

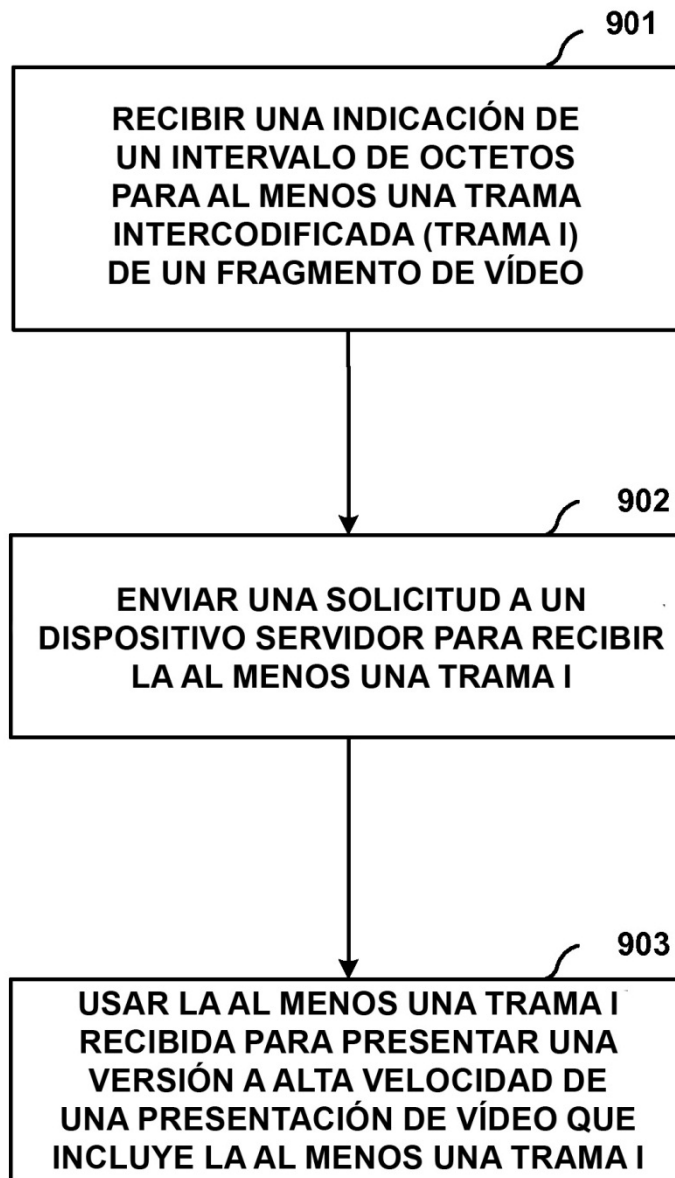


FIG. 9

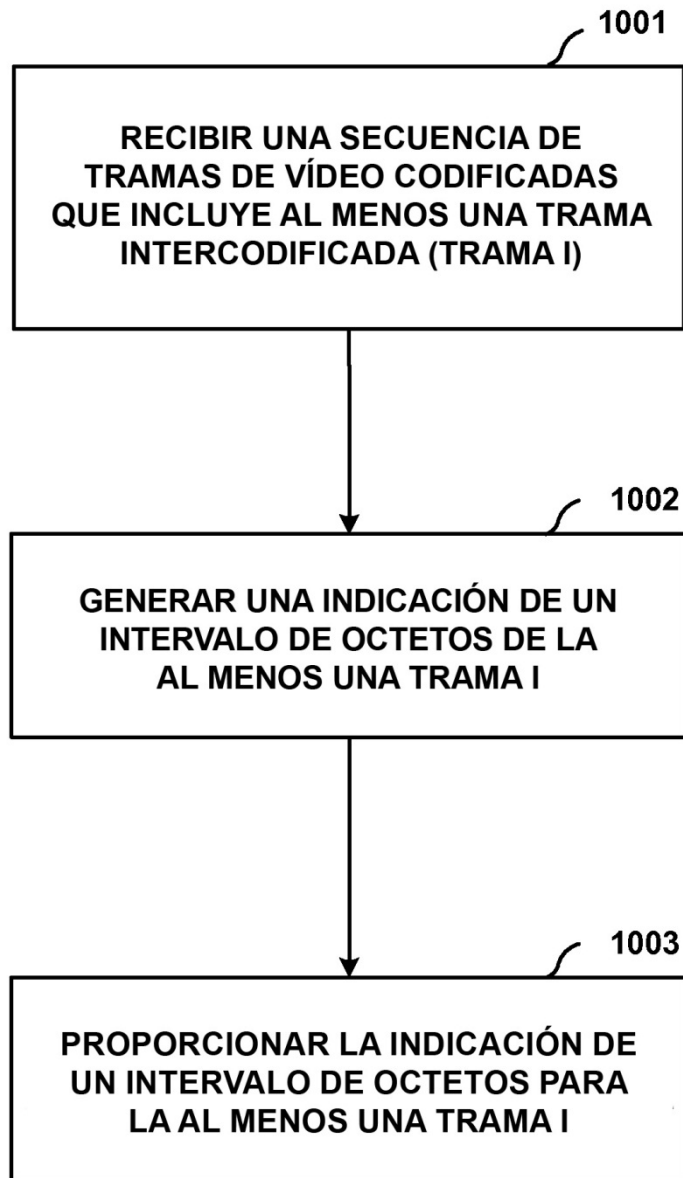


FIG. 10

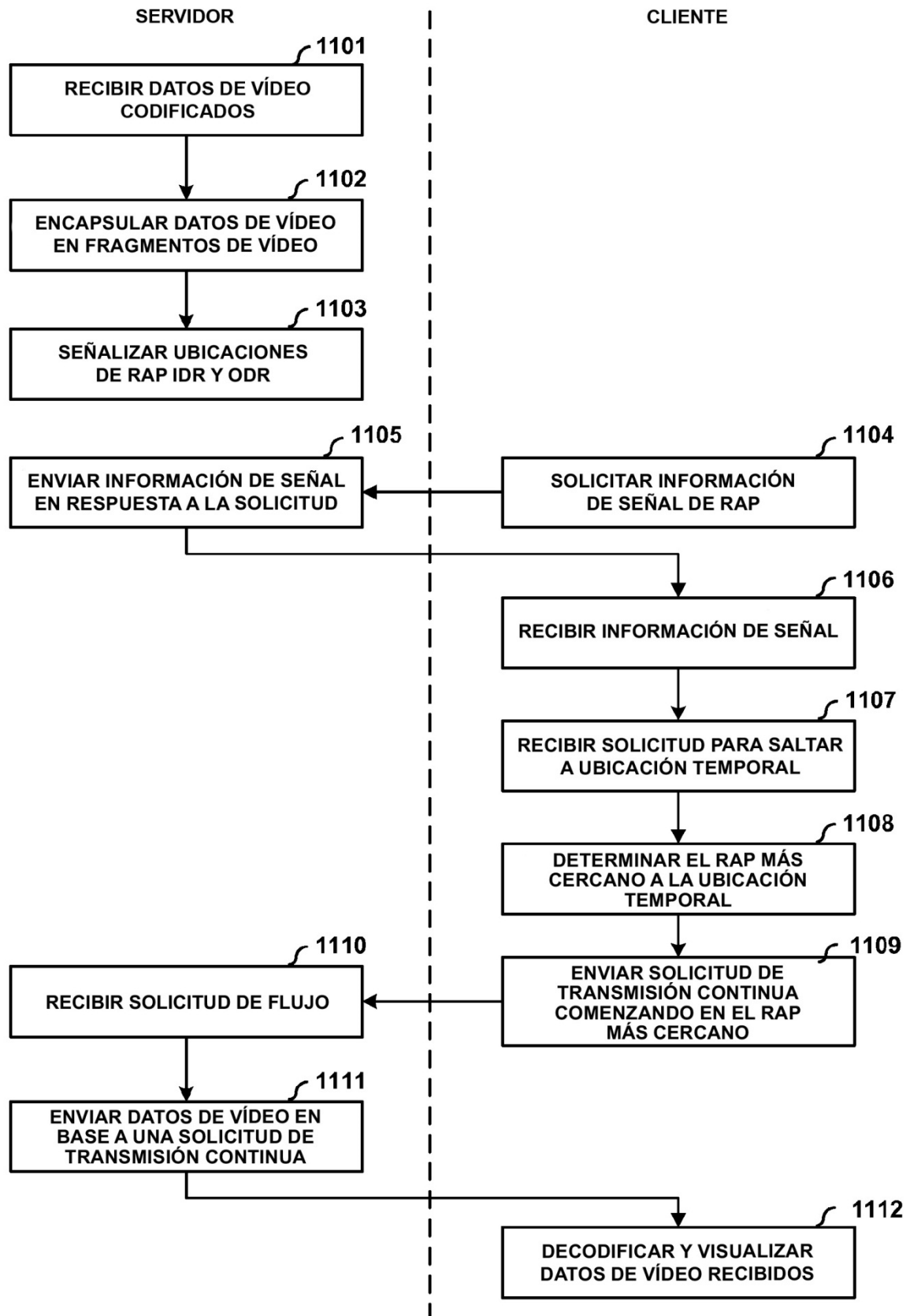


FIG. 11

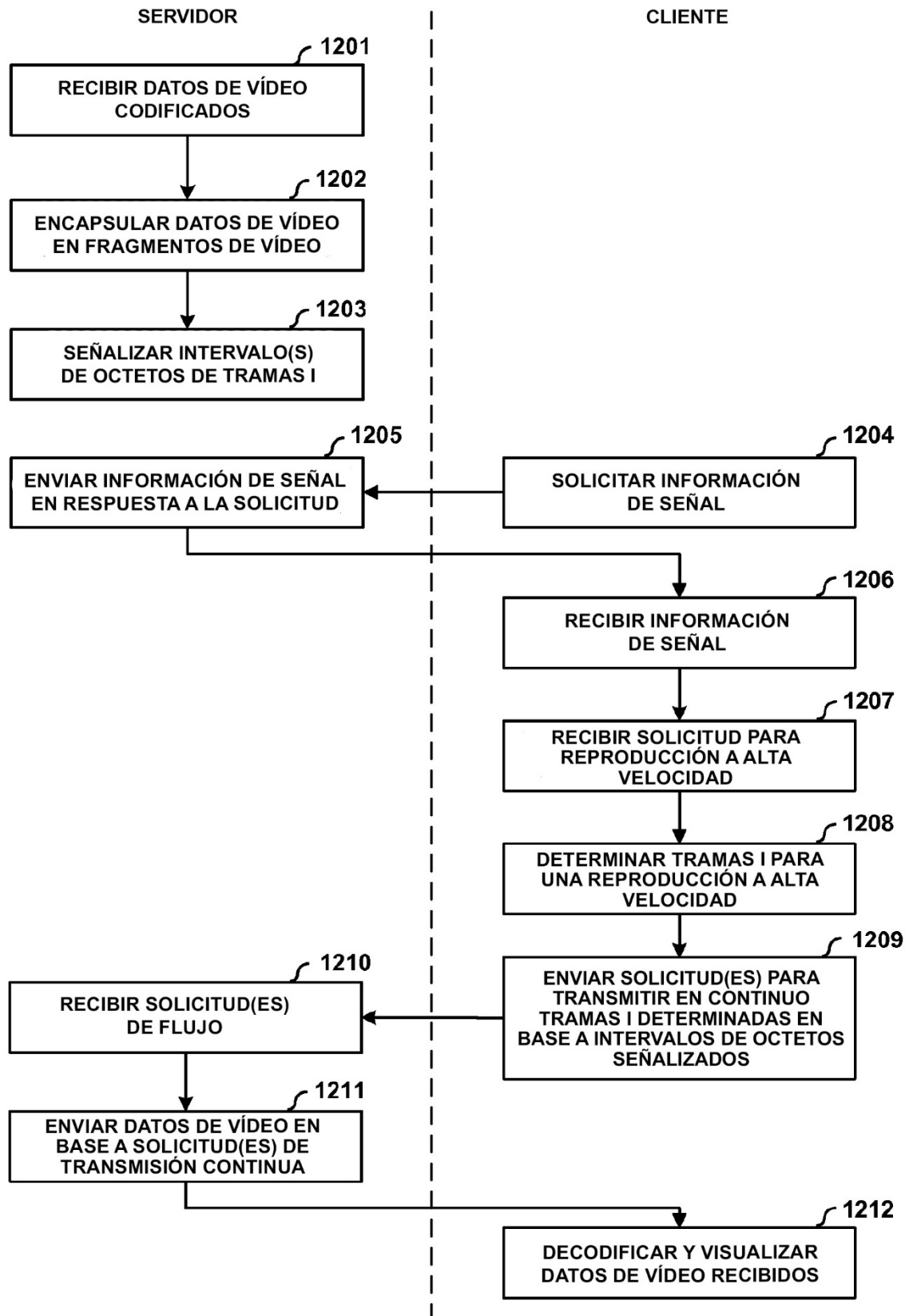


FIG. 12