

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 698 438**

51 Int. Cl.:

H04N 19/70 (2014.01)

H04N 19/46 (2014.01)

H04N 19/61 (2014.01)

H04N 19/117 (2014.01)

H04N 19/82 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.12.2011 PCT/CA2011/001412**

87 Fecha y número de publicación internacional: **05.07.2012 WO12088595**

96 Fecha de presentación y número de la solicitud europea: **28.12.2011 E 11852303 (4)**

97 Fecha y número de publicación de la concesión europea: **03.10.2018 EP 2659679**

54 Título: **Procedimiento para interrumpir de manera selectiva la predicción en la codificación de vídeo**

30 Prioridad:

28.12.2010 US 201061427569 P

23.12.2011 US 201113336475

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

04.02.2019

73 Titular/es:

**DOLBY INTERNATIONAL AB (100.0%)
Apollo Building, 3E, Herikerbergweg 1-35
1101 CN Amsterdam Zuid-Oost, NL**

72 Inventor/es:

HOROWITZ, MICHAEL

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 698 438 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

Procedimiento para interrumpir de manera selectiva la predicción en la codificación de vídeo

5 Esta solicitud reivindica la prioridad de la solicitud de patente provisional estadounidense con n.º de serie 61/427.569, presentada el 28 de diciembre de 2010, titulada "PICTURE SEGMENTATION USING GENERALIZED SLICES", y de la solicitud de patente estadounidense con n.º de serie 13/336.475, presentada el 23 de diciembre de 2011, titulada "METHOD AND SYSTEM FOR SELECTIVELY BREAKING PREDICTION IN VIDEO CODING".

10 **Campo**

Las formas de realización de la invención se refieren a la compresión de vídeo y, más específicamente, al uso selectivo de mecanismos de predicción y de filtrado en bucle en límites de segmento de imagen de imágenes de vídeo.

15 **Antecedentes**

Las capacidades del vídeo digital pueden incorporarse en una amplia gama de dispositivos, incluidos televisores digitales, sistemas digitales de radiodifusión directa, sistemas de radiodifusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, cámaras de vídeo, dispositivos de grabación digital, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o de radio por satélite, y similares. Los dispositivos de vídeo digital pueden implementar técnicas de compresión de vídeo, tales como las descritas en normas como MPEG-2, MPEG-4, ambas disponibles por la Organización Internacional de Normalización ("ISO") 1, ch. de la Voie-Creuse, Case postale 56, CH-1211 Ginebra 20, Suiza, o www.iso.org, o ITU-T H.264/MPEG-4, Parte 10, Codificación de Vídeo Avanzada ("AVC"), disponible por la Unión Internacional de Telecomunicaciones ("ITU"), Place de Nations, CH-1211 Ginebra 20, Suiza o www.itu.int, o según otras especificaciones estándar o no estándar, para codificar y/o descodificar información de vídeo digital de manera eficiente. Otras técnicas de compresión pueden desarrollarse en el futuro o estar actualmente en desarrollo. Por ejemplo, el comité JCT-VC está desarrollando una nueva norma de compresión de vídeo conocida como HEVC/H.265. El borrador de trabajo HEVC/H.265 se describe en el documento de Wiegand *et. al.*, "WD3: Working Draft 3 of High-Efficiency Video Coding, JCTVC-E603", marzo de 2011, denominado en lo sucesivo "WD3".

Un codificador de vídeo puede recibir información de vídeo sin codificar para su procesamiento en cualquier formato adecuado, que puede ser un formato digital conforme a ITU-R BT 601 (disponible por la Unión Internacional de Telecomunicaciones, Place des Nations, 1211 Ginebra 20, Suiza, www.itu.int), o en algún otro formato digital. El vídeo sin codificar puede organizarse tanto espacialmente, en valores de píxel dispuestos en una o más matrices bidimensionales, como temporalmente, en una serie de imágenes sin codificar, donde cada imagen sin codificar comprende una o más de las matrices bidimensionales de valores de píxel mencionadas anteriormente. Además, cada píxel puede comprender una pluralidad de componentes independientes utilizados para representar el color en formato digital. Un formato común de vídeo sin codificar que se introduce en un codificador de vídeo tiene, para cada grupo de cuatro píxeles, cuatro muestras de luminancia que contienen información referente al brillo/luminosidad u oscuridad de los píxeles, y dos muestras de crominancia que contienen información de color (p.ej., YCrCb 4:2:0).

Una función de los codificadores de vídeo es convertir (más generalmente, "transformar") imágenes sin codificar en un flujo de bits, un flujo de paquetes, un flujo de unidades NAL o en otro formato de transmisión adecuado (todos ellos denominados en lo sucesivo "flujo de bits"), con objetivos tales como reducir la cantidad de redundancia codificada en el flujo de bits para aumentar así las velocidades de transmisión, aumentar la resistencia del flujo de bits para suprimir errores de bits o borrados de paquetes que puedan producirse durante la transmisión (lo que se conoce colectivamente como "resistencia a errores"), u otros objetivos específicos de aplicación. Las formas de realización de la presente invención proporcionan al menos uno de entre la supresión o reducción de la redundancia, el aumento de la resistencia a errores y la capacidad de implementar codificadores de vídeo y/o descodificadores asociados en arquitecturas de procesamiento en paralelo.

Una función de los descodificadores de vídeo es recibir como entrada un vídeo codificado en forma de flujo de bits que puede haber sido producido por un codificador de vídeo conforme a la misma norma de compresión de vídeo. Después, el codificador de vídeo convierte (más generalmente, "transforma") el flujo de bits codificado recibido en información de vídeo sin codificar que puede mostrarse, almacenarse o manipularse de otra manera.

Tanto los codificadores de vídeo como los descodificadores de vídeo pueden implementarse usando configuraciones de hardware y/o de software, incluidas combinaciones de hardware y software. Implementaciones de uno o ambos pueden incluir el uso de componentes de hardware programables tales como unidades de procesamiento central (CPU) de propósito general, como los encontrados en ordenadores personales (PC), procesadores integrados, procesadores de tarjetas gráficas, procesadores de señales digitales (DSP), matrices de puertas programables *in situ* (FPGA) u otros. Para implementar al menos partes de la codificación o descodificación de vídeo pueden necesitarse instrucciones, y esas instrucciones pueden almacenarse y distribuirse usando uno o más medios legibles por ordenador no transitorios. Las opciones de medios legibles por ordenador incluyen memoria de sólo

lectura de disco compacto (CD-ROM), memoria de sólo lectura de disco de vídeo digital (DVD-ROM), memorias USB, ROM integradas, u otras.

5 La compresión y descompresión de vídeo se refieren a ciertas operaciones realizadas en un codificador y/o un descodificador de vídeo. Un descodificador de vídeo puede realizar todas o un subconjunto de operaciones inversas a las operaciones de codificación. A menos que se indique lo contrario, las técnicas de codificación de vídeo descritas en el presente documento también abarcan técnicas inversas a las técnicas de codificación de vídeo descritas (en concreto, técnicas de descodificación de vídeo asociadas).

10 La representación digital sin comprimir de vídeo puede considerarse como un flujo de muestras, donde las muestras pueden procesarse por el dispositivo de visualización de vídeo en el orden de exploración. Un tipo de límite que suele producirse en este flujo de muestras es el límite entre imágenes en el flujo de muestras. Muchas normas de compresión de vídeo reconocen este límite y suelen dividir el flujo de bits codificado en estos límites, por ejemplo mediante la inserción de una cabecera de imagen u otros metadatos al comienzo de cada imagen sin codificar.
15 Otros límites que pueden producirse en el flujo de muestras incluyen límites de sección y límites de mosaico que pueden producirse dentro de una imagen sin codificar, como se describe más adelante.

La predicción en la codificación de vídeo puede producirse en muchos niveles.

20 Un nivel se denominará en lo sucesivo "nivel de codificación por entropía" y la predicción en ese nivel se denomina "predicción de codificación". En este nivel, la descodificación de un símbolo codificado por entropía puede requerir la descodificación satisfactoria de símbolos previos codificados por entropía. Todas o casi todas las normas actuales de compresión de vídeo interrumpen la predicción de codificación tanto a nivel de imagen como a nivel de sección. Es decir, cuando se detecta una cabecera de imagen o de sección en el flujo de bits (o equivalente), los estados relacionados con la codificación por entropía utilizados en la codificación por entropía se restablecen a un estado de inicialización. Un ejemplo de predicción codificada por entropía es el restablecimiento de los estados CABAC en ITU-T Rec. H.264.
25

30 Además, puede haber mecanismos de codificación no incluidos en lo que se entiende habitualmente como predicción relacionada con la codificación por entropía, definida anteriormente, pero que aun así están relacionados con la información de control de reconstrucción asociada al flujo de bits, en lugar de a valores de píxel. Como un ejemplo, incluso algunas normas antiguas, tales como la ITU-T Rec. H.261, permiten la codificación de vectores de movimiento con respecto a uno o más vectores de movimiento previamente codificados. La detección de una cabecera de grupo de bloques (GOB), de sección o de imagen restablece este vector de predicción a (0, 0).
35

También hay mecanismos de predicción que abarcan múltiples imágenes. Por ejemplo, la compensación de movimiento puede utilizar valores de píxel (posiblemente compensados en movimiento) de una o más imágenes de referencia para la predicción. Este tipo de predicción se interrumpe por el tipo de macrobloque (o equivalente). Por ejemplo, los intra-macrobloques no suelen usar predicción a partir de imágenes de referencia, mientras que los inter-macrobloques pueden usarla. En este sentido, las intra- e inter-secciones son simplemente acumulaciones de macrobloques que pertenecen a esos tipos de macrobloque diferentes.
40

También hay niveles de predicción que incluyen predicción basada en valores de píxel que ya han sido reconstruidos durante el proceso de reconstrucción de la imagen que está codificándose. Un ejemplo son los mecanismos de intrapredicción, tales como los descritos en el anexo I de la ITU-T Rec. H.263. (Mecanismos similares también están disponibles en otras normas de codificación de vídeo).
45

Además de mecanismos de predicción, varias normas de codificación de vídeo especifican filtros para realizar un filtrado en bucle. Un ejemplo es el filtro en bucle especificado en el anexo J de la ITU-T Rec. H.263.
50

En algunas aplicaciones, puede ser ventajoso segmentar la imagen que está codificándose en bloques de datos más pequeños, segmentación que puede producirse antes de, o durante, la codificación. A continuación se describen dos casos de uso en los que la segmentación de imágenes puede ser ventajosa.

55 El primer caso de uso implica un procesamiento en paralelo. Anteriormente, el vídeo de definición estándar (p.ej., de 720x480 o 720x576 píxeles) fue el formato más grande de mayor difusión comercial. Más recientemente, formatos HD (de hasta 1920x1080 píxeles), así como de 4k (4096x2048 píxeles), 8k (8192x4096 píxeles), y formatos aún mayores, están surgiendo y encontrando utilidad en diversos espacios de aplicación. A pesar del aumento de la potencia informática que puede conseguirse a lo largo de los años, como resultado de los elevados tamaños de imagen asociados a algunos de estos nuevos y mayores formatos, a menudo es conveniente aprovechar la eficacia del procesamiento en paralelo en los procesos de codificación y descodificación. La codificación y descodificación en paralelo pueden producirse, por ejemplo, a nivel de instrucción (p.ej., utilizando SIMD), en una cadena de procesamiento en la que varias unidades de codificación de vídeo pueden procesarse en diferentes fases de manera simultánea, o en una gran estructura base en la que colecciones de subunidades de codificación de vídeo son procesadas por motores de cálculo independientes como entidades independientes (p.ej., un procesador de propósito general de múltiples núcleos). La última forma de procesamiento en paralelo puede requerir segmentación
60
65

de imágenes.

El segundo caso de uso implica segmentación de imágenes para crear un flujo de bits adecuado para su transporte eficiente a través de redes de paquetes. Códecs cuyo vídeo codificado se transporta a través de redes IP y otras redes de paquetes pueden estar sujetos a una restricción de tamaño de unidad de transmisión máxima ("MTU"). A veces es ventajoso que el tamaño de sección codificada sea tal que el paquete resultante que contiene la sección codificada se aproxime lo más posible al tamaño de MTU sin superar ese tamaño, para mantener alta la relación de carga útil/sobrecarga de paquetización, al tiempo que se evita la fragmentación (y la consiguiente mayor probabilidad de pérdida) por la red.

El tamaño de MTU difiere ampliamente de una red a otra. Por ejemplo, el tamaño de MTU de muchas conexiones de Internet puede fijarse por el tamaño de MTU más pequeño de infraestructura de red que se utiliza habitualmente en Internet, que corresponde a menudo a limitaciones en Ethernet y puede ser aproximadamente de 1500 octetos.

El número de bits en una imagen codificada depende de muchos factores, tales como las dimensiones de la imagen fuente, la calidad deseada, la complejidad del contenido en lo que respecta a la idoneidad para la predicción, y otros factores. Sin embargo, incluso con ajustes de calidad y una complejidad de contenido moderados, para secuencias de resolución HD y superiores, el tamaño de una imagen codificada promedio supera fácilmente el tamaño de MTU. Por ejemplo, un codificador de videoconferencia puede requerir cerca de 2 MBit/s para codificar una secuencia de vídeo de 720p60. Esto da como resultado un promedio de tamaño de imagen codificada de aproximadamente 33333 bits o 4167 octetos, que es considerablemente mayor que los 1500 octetos del tamaño de MTU de Internet. Con resoluciones más altas, el promedio de tamaño de imagen aumenta a valores significativamente por encima del tamaño de MTU de Internet. Suponiendo una relación de compresión similar como en el ejemplo anterior de 720p60, un vídeo de 4096x2048 (4k) a 60 fps (4kp60) puede requerir más de 300.000 bits o 25 paquetes de tamaño MTU para cada imagen de vídeo codificada.

En muchas normas de codificación de vídeo anteriores (por ejemplo, hasta e incluyendo WD3), un segmento de imagen (o, al menos, una forma de un segmento de imagen) se conoce como "sección". En la siguiente descripción, cualquier tipo de fragmentación de imagen (p. ej. basada en codificación de vídeo) que interrumpe al menos una forma de predicción en imagen, un filtrado en bucle u otro mecanismo de codificación, puede denominarse, de manera genérica, "sección". De este modo, estructuras tales como el grupo de bloques ("GOB") en la ITU.T Rec. H.261 o ITU Rec. H.263 (disponibles por la ITU; véase más arriba para H.264), o secciones en H.264 o la familia de normas MPEG, pueden constituir una "sección", en la manera en que este término se utiliza a lo largo del presente documento. Sin embargo, las unidades de fragmentación de RFC3984 o las particiones de datos de H.264 pueden no constituir una "sección", en la manera en que este término se utiliza en el presente documento, ya que subdividen el flujo de bits de una imagen codificada y no interrumpen la predicción en imagen, el filtrado en bucle u otro mecanismo de codificación.

Haciendo referencia a la FIG. 1, se muestra un ejemplo **100** de segmentación de imagen usando secciones. Una imagen **101** se divide en dos secciones **102**, **103** en el orden de exploración. El límite de sección se muestra como una línea en negrita **104**. El primer macrobloque **105** de la segunda sección **103** tiene la dirección 11. El flujo de bits **106** correspondiente para transmitir la imagen **101**, por ejemplo cuando se genera usando la norma H.264, puede contener uno o más conjuntos de parámetros **107** que no contienen información acerca de los límites de sección, seguidos por las cabeceras de sección **108**, **110** y los datos de sección **109**, **111** de las dos secciones **102**, **103**. La cabecera de sección **110** de la segunda sección **103** se muestra ampliada. Las dimensiones de la sección sin codificar **103**, por ejemplo, se determinan por un decodificador mediante una combinación de al menos dos factores. En primer lugar, la cabecera de sección **110** contiene la dirección del primer macrobloque **105** de la sección **103**. En segundo lugar, se determina el final de la sección, por ejemplo mediante la detección de una nueva cabecera de sección en el flujo de bits o, en el ejemplo descrito, mediante el final de la imagen codificada en el flujo de bits **112**, es decir, después del macrobloque 24. Todos los macrobloques entre el primer macrobloque y el final de la sección forman la sección. Cabe señalar que modificaciones en el orden de exploración, tal como la ordenación de macrobloques flexible de H.264, puede cambiar el número de macrobloques en la sección mediante la creación de espacios.

Una ventaja de utilizar secciones en lugar de mecanismos de segmentación que no tienen en cuenta los medios, tales como, por ejemplo, los mecanismos proporcionados mediante IP en la capa de encaminamiento, es que las secciones pueden descodificarse de manera independiente, al menos hasta cierto punto (como se explica posteriormente en mayor detalle), al interrumpir ciertos tipos de predicción en los límites entre secciones. Por lo tanto, la pérdida de una sección no implica necesariamente que las otras secciones de una imagen codificada no puedan usarse o descodificarse. Dependiendo de la implementación de un mecanismo de fragmentación, la pérdida de un fragmento, en cambio, puede hacer que otros muchos fragmentos no puedan utilizarse debido a que la fragmentación, tal y como se utiliza este término a lo largo del presente documento, no interrumpe ninguna forma de predicción.

WD4 (B. Bross *et. al.*, "WD4: Working Draft 4 of High-Efficiency Video Coding", disponible en http://wftp3.itu.int/av-arch/jctvc-site/2011_07_F_Torino/) es el borrador de una especificación relativa a una norma de codificación de

vídeo digital en desarrollo, que puede denominarse codificación de vídeo de alta eficiencia (HEVC) o H.265. Además de secciones, WD4 también incluye un mecanismo de segmentación de imágenes conocido como "mosaicos". Según WD4, una imagen fuente puede dividirse en unidades rectangulares denominadas mosaicos, de tal manera que cada píxel de la imagen fuente es parte de un mosaico (también pueden aplicarse otras restricciones). Un mosaico es, por tanto, una parte rectangular de una imagen. Los límites del mosaico se determinan mediante coordenadas disponibles en estructuras sintácticas de alto nivel, que se conocen en WD4 como conjuntos de parámetros. Los mosaicos se describen posteriormente en mayor detalle.

Con la posible excepción de la predicción inter-imagen, cada uno de los mecanismos de predicción en imagen o mecanismos de codificación descritos anteriormente pueden interrumpirse por la descodificación de una cabecera de imagen (o equivalente, tal como la descodificación de una sección con un número de cuadro diferente al de la sección anterior). El que esos mecanismos de predicción se interrumpan a través de los límites de sección o de mosaico depende de la norma de compresión de vídeo y del tipo de sección en uso.

En H.264, las secciones pueden descodificarse de manera independiente con respecto a la predicción de vector de movimiento, la intra-predicción, los estados CA-VLC y CABAC, y otros aspectos de la norma H.264. Sólo está permitida la predicción inter-imagen (incluida la importación de datos de píxel fuera de los límites de sección a través de la compensación de movimiento). Aunque esta independencia de descodificación aumenta la resistencia a errores, inhabilitar la predicción antes mencionada a través de los límites de sección reduce la eficacia de la codificación.

En H.263, un codificador de vídeo tiene más flexibilidad a la hora de seleccionar los mecanismos de predicción que se interrumpen a través el uso de secciones o GOB con cabeceras de GOB no vacías. Por ejemplo, hay un bit incluido en la cabecera de imagen, que puede seleccionarse cuando se usa el anexo R, que indica al descodificador que ninguna predicción o filtrado se produce a través de límites de sección/GOB (con cabeceras no vacías). Ciertos mecanismos de predicción, tales como predicción de vector de movimiento, se interrumpen a través de GOB con cabeceras no vacías y a través de límites de sección, independientemente del estado del Anexo R. Otros son controlados por el Anexo R. Por ejemplo, si el bit no está fijado, los vectores de movimiento pueden apuntar fuera del área espacial ubicada conjuntamente con la sección/el GOB actual con cabecera no vacía en la(s) imagen(es) de referencia, lo cual puede "importar" posiblemente los valores de muestra que se utilizan para la compensación de movimiento en la sección actual de un área que no está dentro del área geométrica de la sección/del GOB en la imagen de referencia. Además, a menos que el anexo R esté activo, el filtrado en bucle puede incorporar valores de muestra fuera de la sección/del GOB. Asimismo, hay otro bit en la cabecera de imagen que habilita o inhabilita la intra-predicción.

Sin embargo, en la mayoría de normas, la decisión de interrumpir la predicción en imagen se hace, al menos, con granularidad de imagen y, en algunos casos, con granularidad de secuencia. De otro modo, utilizando H.263 como un ejemplo, no es posible mezclar secciones en una imagen dada que tiene el filtro de desbloqueo habilitado o inhabilitado (respectivamente), ni es posible habilitar/inhabilitar la intra-predicción a nivel de sección.

Como ya se ha descrito, la segmentación de imágenes permite interrumpir una imagen en áreas espaciales más pequeñas que una imagen completa. Aunque las aplicaciones más comunes de la segmentación de imágenes, tal como se describe, parece que son la adaptación al tamaño de MTU y la paralelización, la segmentación de imágenes también puede utilizarse con otros muchos fines, incluidos aquellos que adaptan el tamaño y la forma del segmento al contenido. La región de codificación de interés es uno de varios ejemplos. En tales casos, es posible que ciertas partes de una imagen puedan codificarse con mayor eficacia que otras (en el sentido de que utilizar un menor número de bits para la codificación da como resultado una experiencia visual comparable) cuando se aplican diferentes herramientas de codificación, incluidos diferentes mecanismos de predicción. Por ejemplo, algunos contenidos pueden beneficiarse del filtrado de desbloqueo y pueden no responder bien a la intra-predicción, mientras que otro contenido de la misma imagen puede codificarse mejor sin filtrado de desbloqueo pero podría beneficiarse de la intra-predicción. Un tercer contenido puede codificarse mejor con el filtrado de desbloqueo y la intra-predicción habilitados. Todo este contenido puede estar ubicado en la misma imagen cuando la imagen es un mosaico, lo que ocurre, por ejemplo, cuando se realiza una entrevista o una videoconferencia.

Un inconveniente de los mecanismos existentes para la interrupción de la predicción en límites de segmento es que la habilitación y/o inhabilitación de la interrupción de la predicción está generalmente codificada en las normas actuales de codificación de vídeo, lo que hace difícil o imposible interrumpir de manera selectiva los mecanismos de predicción en los límites de segmento en función de, por ejemplo, las características del contenido que va a codificarse.

Por lo tanto, existe la necesidad de un procedimiento y un sistema mejorados para habilitar o inhabilitar, sección a sección, mecanismos de predicción y de filtrado en bucle de manera individual o como un grupo. En consecuencia, se desea una solución que resuelva, al menos en parte, los anteriores y otros inconvenientes.

Además, existe la necesidad de habilitar o inhabilitar mecanismos de predicción y/o de filtrado en bucle imagen a imagen (o por grupo de imágenes, secuencias, etc.), a través de límites de segmento de imagen sin cabecera (o

equivalentes) (tales como límites de mosaico), de manera individual o como un grupo. En consecuencia, se desea una solución que resuelva, al menos en parte, los anteriores y otros inconvenientes.

El documento US 2010/0189181 describe un procedimiento y dispositivo para la codificación y decodificación. El procedimiento incluye: dividir una imagen que va a codificarse en varias secciones, cada una conteniendo macrobloques continuos en una secuencia de exploración designada en la imagen; dividir secciones de la imagen en uno o más conjuntos de secciones según información de atributo de las secciones, conteniendo cada conjunto de secciones una o más secciones; y codificar las secciones de los conjuntos de secciones según información de división de sección y de conjunto de secciones para obtener un flujo de bits codificado de la imagen. El procedimiento de decodificación incluye: obtener información de división de sección y de conjunto de secciones a partir de un flujo de bits que va a descodificarse y descodificar el flujo de bits según la información obtenida de división de sección y de conjunto de secciones.

El documento de Y-K: Wang *et al.* "AHG4: Dependency and loop filtering control over tile boundaries", 7.JCT-VC MEETING; 98. MPEG MEETING; 21/11/2011 - 30/11/2011; Ginebra; (EQUIPO DE COLABORACIÓN CONJUNTA SOBRE CODIFICACIÓN DE VÍDEO DE ISO/IEC JTC1/SC29/WG11 e ITU-T SG.16); URL: [HTTP://WFTP3.ITU.INT/AV-ARCH/JCTVC-SITE/](http://wftp3.itu.int/av-arch/jctvc-site/), n.º JCTVC-G317, 9 de noviembre de 2011 (09/11/2011) da a conocer un procedimiento para inhabilitar el filtrado en bucle a través de límites de mosaico.

Resumen

Las formas de realización de la presente invención proporcionan procedimiento(s) y sistema(s) para la codificación y/o decodificación de imágenes de vídeo en las que una pluralidad de herramientas de predicción y de filtrado en bucle para un segmento de imagen pueden habilitarse o inhabilitarse de manera selectiva.

Según un aspecto de la invención, un codificador puede indicar a una o más herramientas de predicción si esa herramienta puede tomar información de fuera del segmento de imagen que se está procesando actualmente como información de referencia para el procesamiento dentro de ese segmento de imagen. El codificador puede proporcionar tal indicación a una única herramienta de predicción (por ejemplo, predicción por entropía, intra-predicción, predicción compensada por movimiento, predicción de vector de movimiento, denominadas en lo sucesivo herramientas de predicción) y/o una sola herramienta de filtrado (por ejemplo, filtrado de interpolación adaptativo, filtrado en bucle adaptativo, filtrado de desbloqueo, desfase adaptativo de muestra, denominados en lo sucesivo herramientas de filtro en bucle), así como otras. Como alternativa, el codificador puede proporcionar tal indicación a una pluralidad de herramientas predefinidas o un grupo predefinido de herramientas que puede incluir cualquiera de las herramientas de predicción y de filtro en bucle antes mencionadas, así como otras. Esto puede ser útil para dar soporte a la paralelización de los codificadores y decodificadores, así como en determinados escenarios de aplicación, tales como software de presencia continua (unión de imágenes codificadas en el dominio comprimido).

Según un aspecto de la invención, cuando se utiliza la segmentación de imágenes sin cabecera (como los mosaicos), el codificador puede indicar a la herramienta de predicción, la herramienta de filtro en bucle o la pluralidad de grupos predefinidos de herramientas, si esa herramienta puede tomar información a través de límites de mosaico horizontales, verticales o tanto horizontales como verticales que se utilizarán como información de referencia.

En el caso particular de H.264 o HEVC, como un ejemplo, el codificador puede establecer valores para banderas de "indicación de interrupción de codificación" utilizadas en herramientas de predicción y de filtrado en bucle tales como: intra-predicción que hace referencia a valores de muestra fuera del límite de sección/mosaico; vectores que hacen referencia a valores de muestra fuera del límite de sección/mosaico (es decir, a través de la compensación de movimiento); uso de estados CABAC fuera del límite de sección/mosaico; uso de estados CA-VLC fuera del límite de sección/mosaico; uso de estados de codificación por entropía PIPE o V2V similares fuera del límite de sección/mosaico (sólo HEVC); y uso de estados y valores de muestra fuera del límite de sección/mosaico de filtros en bucle tales como filtros de interpolación adaptativos, filtros en bucle adaptativos, filtros en bucle de desbloqueo o desfase adaptativo de muestras.

Según un aspecto de la invención, la utilización u otra habilitación de herramientas de codificación puede no estar indicada en forma de banderas, sino a través de la manipulación de otras estructuras de datos, tales como enteros de "indicación de interrupción de codificación" que pueden, en algunos casos, combinar una pluralidad de las banderas antes mencionadas, o permutaciones preferidas de esas banderas, en un único símbolo.

Según un aspecto de la invención, la longitud máxima de un vector de movimiento que apunta hacia fuera de un límite de sección puede codificarse en una representación de codificación por entropía apropiada de un entero, lo que indica no sólo la no utilización de la compensación de movimiento hasta la distancia permitida por el nivel en uso, sino también el máximo permitido, que puede, por ejemplo, ayudar en la asignación de recursos en la implementación de un decodificador.

Según un aspecto de la invención, al menos una de las banderas de indicación de interrupción de codificación antes mencionadas u otras estructuras de indicación de interrupción de codificación de datos pueden almacenarse en la cabecera de sección, la cabecera de imagen, un conjunto de parámetros o estructuras equivalentes.

5 Según un aspecto de la invención, un descodificador puede reaccionar a la presencia de las banderas o de otras estructuras de datos interrumpiendo las herramientas de predicción indicadas a través de límites de sección/mosaico en oposición a otros límites potencialmente adecuados.

10 En un sentido amplio, se proporciona un procedimiento para descodificar una imagen de vídeo codificada según la reivindicación 1.

En otro sentido amplio, se proporciona un procedimiento para codificar una imagen de vídeo según la reivindicación 2.

15 **Breve descripción de los dibujos**

Otras características y ventajas de las formas de realización de la presente invención resultarán evidentes a partir de la siguiente descripción detallada, tomada en combinación con los dibujos adjuntos, en los que:

20 la FIG. 1 es un diagrama que ilustra una imagen a modo de ejemplo que presenta secciones en el orden de exploración, y un flujo de bits que representa la imagen codificada según una forma de realización de la invención;

la FIG. 2 es un diagrama que ilustra mosaicos y secciones según una forma de realización de la invención;

25 la FIG. 3 es un diagrama de bloques que ilustra un flujo de bits codificado según una forma de realización de la invención;

la FIG. 4 es un diagrama de bloques que ilustra un flujo de bits codificado según una forma de realización de la invención;

la FIG. 5 es un diagrama de bloques que ilustra un flujo de bits codificado según una forma de realización de la invención;

30 la FIG. 6 es un diagrama de flujo que ilustra el funcionamiento de un descodificador a modo de ejemplo según una forma de realización de la invención;

la FIG. 7 es un diagrama de flujo que ilustra el funcionamiento de un descodificador a modo de ejemplo cuando se descodifica una sección según una forma de realización de la invención; y,

35 la FIG. 8 es un diagrama de bloques que ilustra un sistema de procesamiento de datos (p.ej., un ordenador personal ("PC") basado en una implementación según una forma de realización de la invención.

Cabe señalar que a lo largo de los dibujos adjuntos, características similares se identifican mediante números de referencia similares.

40 **Descripción detallada de las formas de realización**

En la siguiente descripción se exponen detalles para poder entender la invención. En algunos casos no se describen o muestran den detalle cierto software, sistemas de circuitos, estructuras ni procedimientos con el fin de no oscurecer la invención. El término "sistema de procesamiento de datos" se utiliza en el presente documento para hacer referencia a cualquier máquina para el procesamiento de datos, incluidos sistemas informáticos, dispositivos inalámbricos y disposiciones de red descritos en el presente documento. Las formas de realización de la presente invención pueden implementarse en cualquier lenguaje de programación informático siempre que el sistema operativo del sistema de procesamiento de datos proporcione herramientas que puedan dar soporte a las necesidades de estas formas de realización. Las formas de realización de la presente invención también pueden implementarse en hardware o en una combinación de hardware y software.

55 Al menos algunas formas de realización de la presente invención se refieren a la interrupción selectiva de mecanismos de predicción y/o a la inhabilitación selectiva de mecanismos de filtrado en bucle junto con la segmentación de imágenes en la compresión de vídeo.

Términos tales como "segmento" o "segmento de imagen" se utilizan en lo sucesivo para hacer referencia a uno cualquiera o más macrobloques o estructuras equivalentes (por ejemplo, bloques de árbol en WD4) que son más pequeños que la imagen completa, y en cuyos límites se interrumpe al menos una forma de predicción y/o se inhabilita al menos una forma de filtrado en bucle. Secciones como las de H.264, así como mosaicos como los de WD4, descritos más adelante (con *tile_boundary_independence_idc* igual a 1) son ejemplos no limitativos de segmentos.

65 La FIG. 2 muestra un ejemplo **200**, donde una imagen **201** está dividida en dos mosaicos **202**, **203** mediante un límite de mosaico de división vertical **204**, representado como un línea continua en negrita. Los mosaicos pueden coexistir con secciones en la imagen **201**. Por ejemplo, la imagen **201** está dividida en dos secciones por el límite de sección **205** al mismo tiempo que está dividida en los mosaicos **202**, **203** mediante el límite de mosaico **204**. Mosaicos

(con *tile_boundary_independence_idc* igual a 1), como se describe en WD4, pueden, en uno o más aspectos, generalizar otro tipo de segmento de imagen, que recibe el nombre de columna, que se describe en detalle en la solicitud de patente estadounidense en tramitación junto con la presente con n.º de serie 13/336.675, presentada el 23 de diciembre de 2011, titulada "METHOD AND SYSTEM FOR PICTURE SEGMENTATION USING COLUMNS".

El flujo de bits **206** correspondiente a la transmisión de la imagen **201** puede contener, por ejemplo, conjuntos de parámetros **207** u otros elementos sintácticos de capa alta que incluyen información de límite de mosaico **208** para identificar el límite de mosaico **204**. Sin embargo, las partes del flujo de bits distintas de los conjuntos de parámetros **207** no incluyen ninguna información acerca de los límites de los mosaicos. Un decodificador puede identificar el mosaico al que pertenece un macrobloque codificado (también conocido como mayor unidad de codificación (LCU) o (en WD4) bloque de árbol) al asociar su información de estado interno del macrobloque que está procesándose actualmente y la información de las dimensiones de mosaico conocidas a partir de los conjuntos de parámetros **207**.

Una diferencia entre los mosaicos y otros mecanismos de segmentación de imágenes rectangulares tales como secciones rectangulares (un submodo del anexo K de la ITU-T Rec. H.263) es que un mosaico (a diferencia de las secciones rectangulares) no requiere una cabecera. Las dimensiones físicas de un mosaico pueden definirse, en cambio, en el conjunto de parámetros sin la inclusión de una cabecera. En algunos casos (con *tile_boundary_independence_idc* igual a 1), los límites de mosaico según WD4 interrumpen todos los mecanismos de predicción en imagen, pero permiten la referencia de muestras en imágenes de referencia que no están ubicadas conjuntamente con muestras asociadas al mosaico para el cual se está realizando la compensación de movimiento. Además, los límites de mosaico no interrumpen el filtrado en bucle, incluido el filtro de desbloqueo, el filtro de desfase adaptativo de muestras y el filtro de bucle adaptativo.

Sin embargo, también puede ser conveniente o deseable que un codificador o un decodificador interrumpieran un conjunto diferente de mecanismos de predicción usando mosaicos. Por ejemplo, a una resolución muy elevada, puede ser ventajoso dividir las imágenes de vídeo en mosaicos sujetos a requisitos tales como que los vectores de movimiento no puedan apuntar fuera de los límites de mosaico, y/o de que un límite de mosaico sea tratado por el codificador y el decodificador como un límite de imagen (semejante al anexo R de H.263), o similar, por lo que, por ejemplo, no sólo se evita la compensación de movimiento a través de límites de mosaico, sino también un filtrado en bucle.

En otros casos, puede ser conveniente o deseable que el codificador o el decodificador puedan manejar una codificación de vídeo a máxima resolución, excepto para la codificación por entropía de símbolos. Tal codificador o decodificador de este tipo podrían necesitar, por ejemplo, hardware de procesamiento de señales dedicado para el procesamiento basado en muestras, pero pueden usar una CPU multinúcleo de propósito general para la codificación y/o descodificación por entropía, donde un único núcleo no sería capaz de manejar la carga (en HEVC se sabe que, en especial, la codificación por entropía CABAC es exigente desde un punto de vista computacional). Por consiguiente, con el fin de dar soporte a este caso de uso, puede ser necesario interrumpir la codificación por entropía en los límites de mosaico, mientras que otros mecanismos de predicción en imagen o inter-imagen pueden ser capaces de cruzar límites de sección y/o de mosaico.

En otros casos, puede ser conveniente o deseable que el codificador o el decodificador permitan una coordinación limitada entre procesadores a través de límites de mosaico. En estos casos, las referencias a los valores de píxel pueden no ser posibles, mientras que la referencia a la información de control (tal como la información necesaria para la predicción de vector de movimiento) puede estar disponible en el canal de comunicación entre los procesadores. En tal caso, la intra-predicción puede no ser posible, pero puede usarse la predicción de vector de movimiento.

Puede haber herramientas de codificación que no estén directamente relacionadas con la predicción, pero aún pueden interrumpirse ventajosamente a través de los límites de sección o de mosaico. Por ejemplo, solicitud de patente estadounidense en tramitación junto con la presente con n.º de serie 13/286.828, presentada el 1 de noviembre de 2011, titulada "ADAPTIVE INTERPOLATION IN DIGITAL VIDEO CODING", da a conocer un filtro de interpolación adaptativo cuya naturaleza y coeficientes pueden seleccionarse por el codificador. Puede ser ventajoso limitar el uso de muestras fuera de una sección para el filtrado de interpolación. Asimismo, WD4 incluye un filtro de interpolación adaptativo, cuyo control se obtiene, al menos en parte, de ciertos píxeles. Puede ser ventajoso limitar esta obtención solamente a píxeles que estén dentro del límite de sección o de mosaico. También puede ser ventajoso limitar el filtrado en sí (a diferencia de la obtención de información de control de filtro) a píxeles dentro de los límites de sección o de mosaico. Además, WD4 incluye otros filtros de bucle tal como el filtro en bucle adaptativo (relacionado con el filtrado de todas las muestras), un filtro de desbloqueo (relacionado con el filtrado de límites de bloque) y un mecanismo de filtrado conocido como desfase adaptativo de muestras. Todos estos filtros pueden compartir propiedades similares a AIF. Por ejemplo, en el caso del filtro de bucle adaptativo, tal como se especifica en WD4, puede ser ventajoso (posiblemente de manera independiente) inhabilitar el acceso a información utilizada para obtener las etapas de filtro a través de los límites de mosaico, así como inhabilitar el filtrado a través de los propios límites de mosaico.

Los límites de los segmentos pueden definirse mediante estructuras sintácticas a nivel de imagen (o superiores),

tales como, por ejemplo, conjuntos de parámetros cuando se utilizan mosaicos WD4, mediante información de cabecera de segmento (tales como, por ejemplo, secciones rectangulares del anexo K de H.263), mediante una combinación de la colocación de una cabecera de segmento en el flujo de bits y un estado de codificador/descodificador (tal como, por ejemplo, secciones de H.264 cuando no se utiliza la ordenación de macrobloque flexible (FMO), o una combinación de dos o más de los mecanismos anteriormente mencionados (es decir, la FMO define grupos de secciones, y el segmento de imagen se define dentro del grupo de secciones a través de una combinación de colocaciones de cabecera de sección en el flujo de bits (identificándose el primer macrobloque de la sección por su dirección), y la ascendencia implícita de direcciones de macrobloque dentro del grupo de secciones hasta el final de la sección se detecta a través del análisis sintáctico del flujo de bits u otros medios).

A continuación se describe en primer lugar mecanismos que permiten la selección de herramientas de predicción para los límites de mosaico, seguidos de mecanismos que permiten la selección de herramientas de predicción para los límites de secciones. Finalmente se describirá el interfuncionamiento de los dos mecanismos.

Haciendo referencia a un ejemplo **300** en la FIG. 3, se muestra un flujo de bits codificado **301** que comprende un conjunto de parámetros **302** y dos secciones codificadas **304**, **305**. Las secciones codificadas **304**, **305** pueden pertenecer a una o dos imágenes codificadas. En WD4, un límite de imagen puede identificarse mediante una cabecera de sección con una dirección LCU de 0. El conjunto de parámetros **302** puede incluir información de control de mosaico **303** (tal como límites de mosaico), y, en este ejemplo, se supone que la información del conjunto de parámetros **302** pertenece a ambas secciones codificadas (es decir, la referencia al conjunto de parámetros en la cabecera de sección contiene el mismo índice). En muchos WD4 y sistemas basados en H.264, un conjunto de parámetros pertenece a decenas de, cientos de, o más, secciones.

Según una forma de realización, el conjunto de parámetros **302** puede contener una pluralidad de banderas de indicación de herramienta de predicción (PTI). Cuando las PTI están establecidas (es decir, habilitadas), por ejemplo, puede permitirse la predicción a través de límites de segmento cuando la herramienta de codificación o descodificación está asociada a la bandera; en otro caso, cuando las PTI no están establecidas (es decir, están inhabilitadas), esa predicción puede estar prohibida. Las banderas pueden definirse, por ejemplo, para la predicción de codificación por entropía **306**, la intra-predicción **307**, la predicción de vector de movimiento **308**, la predicción compensada por movimiento **309**, el filtrado de bucle adaptativo **310**, el filtrado de interpolación adaptativo **311**, el filtrado de desbloqueo **312**, el desfase adaptativo de muestras **313** y posiblemente otras herramientas de predicción y de filtrado en bucle definidas en el mecanismo de codificación de vídeo.

La inclusión de PTI en mecanismos individuales de predicción y de filtrado en bucle, que pertenecen a todas las secciones e imágenes que hacen referencia al conjunto de parámetros, puede ayudar a adaptar el flujo de bits al entorno de codificación y/o descodificación, tal como la arquitectura de hardware del codificador o del descodificador. Puesto que las banderas pueden ser parte de un conjunto de parámetros que pueden aplicarse a muchas secciones o imágenes, la sobrecarga de PTI en el conjunto de parámetros puede ser insignificante en comparación con los beneficios que brindan.

Haciendo referencia al ejemplo **400** ilustrado en la FIG. 4, se muestra un flujo de bits codificado **401** que comprende un conjunto de parámetros **402**, y una imagen codificada que contiene dos secciones **403**, **404**. Cada sección comienza con una cabecera de sección **405**, **406**. La cabecera de sección **405** está ampliada para mostrar partes de su información.

Según una forma de realización, la cabecera de sección **405** puede contener una pluralidad de banderas de indicación de herramienta de predicción (PTI). Cuando una o más de las PTI están establecidas, por ejemplo, puede permitirse la predicción y/o el filtrado en bucle a través de límites de segmento cuando la herramienta de codificación o descodificación está asociada a la bandera; en otro caso, cuando las PTI no están establecidas, esa predicción puede estar prohibida. Las PTI pueden definirse, por ejemplo, para la predicción por entropía **407**, la intra-predicción **408**, la predicción de vector de movimiento **409**, la predicción compensada por movimiento **410**, el filtrado en bucle adaptativo **411**, el filtrado de interpolación adaptativo **412**, el filtrado de desbloqueo **413**, el desfase adaptativo de muestras **414** y, posiblemente, otras herramientas de predicción y de filtrado en bucle definidas en el mecanismo de codificación de vídeo.

La inclusión de PTI en mecanismos individuales de predicción y de filtrado en bucle que pertenecen a una sección dada, puede contribuir a adaptar el flujo de bits al contenido, mejorando así la eficacia de la codificación.

A continuación se describe el modo en que los dos mecanismos antes descritos pueden interactuar.

Haciendo referencia a un ejemplo **500** mostrado en la FIG. 5, se muestra un flujo de bits codificado **501** que contiene un conjunto de parámetros **502** y dos secciones **503**, **504**, cada una comenzando con una cabecera de sección **505**, **506** correspondiente.

El conjunto de parámetros **502**, que se muestra ampliado en **507**, por ejemplo, incluye la información de control de

mosaico **508** u otra información relativa a límites de segmento sin cabecera que, por ejemplo, podrían indicar un límite de mosaico vertical **204**, como el mostrado en la FIG. 2. Además, el conjunto de parámetros **502** puede contener una o más PTI. Aquí se muestran tres PTI, una asociada a la predicción por entropía **509**, una a la intra-predicción **510**, y una a la compensación de movimiento **511**. Estas banderas pueden controlar la predicción del descodificador en el límite de mosaico **204**. El límite de mosaico **204** puede, por ejemplo, fijarse por la información de control de mosaico **508** de tal forma que la imagen **201** se divide verticalmente en dos mosaicos **202**, **203**. El mecanismo aquí descrito también puede funcionar con otras disposiciones de los límites de mosaico, incluida una combinación de límites verticales y horizontales.

La imagen codificada también puede contener, por ejemplo, dos secciones codificadas **503**, **504**, cada una comenzando con una cabecera de sección **505**, **506** correspondiente. Como se muestra en la FIG. 2, las secciones (sin codificar) correspondientes a las secciones codificadas **503**, **504** pueden abarcar, por ejemplo, el área espacial de las direcciones de macrobloque 1 a 14, y 15 a 24, respectivamente. La cabecera de sección **506** se muestra ampliada en **512** y puede contener una pluralidad de PTI. Se muestran dos PTI, una asociada a intra-predicción **513**, y la otra a filtrado en bucle adaptativo (ALF) **514**. Sin embargo, cabe señalar que puede haber, aunque no es un requisito, una superposición entre las PTI del conjunto de parámetros **502** o la cabecera de sección **506**.

Según una forma de realización, las PTI **509**, **510**, **511** del conjunto de parámetros **502** controlan la predicción y el filtrado en bucle a través del límite de mosaico **204** según lo definido por la información de control de mosaico **508**.

Según una forma de realización, las PTI **513**, **514** de la cabecera de sección **512** controlan la predicción y el filtrado en bucle a través de los límites entre las secciones **503**, **504**. Por ejemplo, los límites de sección de la sección **504** tienen un límite distinto al límite de imagen, que está marcado por la línea discontinua y en negrita de límite de sección **205**.

Como resultado, en el ejemplo **200**, algunos mecanismos de predicción y de filtrado en bucle son interrumpidos por límites de mosaico (a fin de habilitar la distribución del esfuerzo de codificación de imágenes entre varios procesadores), mientras que otros mecanismos de predicción y de filtrado en bucle son interrumpidos de manera selectiva, y bajo el control de la cabecera de sección **506**, en límites de sección (dando por tanto al codificador pleno control sobre la interrupción de los mecanismos de predicción y de filtrado en bucle para poder seleccionar cualquier combinación particular de mecanismos de predicción y de filtrado en bucle para el contenido que está codificándose, incluidas combinaciones que pueden ser deseables o convenientes para una aplicación o uso dados).

Si una PTI perteneciente al mismo mecanismo de predicción o de filtrado en bucle está presente tanto en el conjunto de parámetros **502** como en una cabecera de sección **506**, y en la situación en la que los límites de sección y de mosaico correspondientes están alineados, el descodificador puede reaccionar de al menos dos maneras posibles. Tales opciones pueden especificarse de manera estática en la norma, a través de una selección de nivel/perfil, o de manera dinámica en función de información de control en un conjunto de parámetros u otros elementos sintácticos de capa alta.

Una opción es que las PTI del conjunto de parámetros **502** sobrescriban información contradictoria en la cabecera de sección **506**. Esta opción puede tener la ventaja de hacer que un descodificador tenga la certeza de que puede distribuir segmentos a varios procesadores o núcleos, sin tener que implementar mecanismos que permitan el intercambio de información entre esos segmentos.

Otra opción es que las PTI de la cabecera de sección **508** sobrescriban información contradictoria en el conjunto de parámetros **502**. Esta elección puede permitir una mayor flexibilidad en el codificador a la hora de elegir sus herramientas. Otras reacciones pueden ser posibles.

A fin de optimizar la codificación de las banderas (independientemente de que estén ubicadas en la cabecera de sección **508** o en el conjunto de parámetros **502**), en algunos casos, puede ser beneficioso especificar en una norma cualquiera de lo siguiente:

(1) algunas PTI pueden no ser parte de un conjunto de parámetros o de una cabecera de sección si un determinado perfil y/o nivel está indicado, ya que la herramienta de predicción o de filtrado en bucle no está disponible en este perfil/nivel.

(2) dos o más PTI pueden "agruparse" en una única combinación de PTI, si, por ejemplo, en un perfil determinado, se determina que la flexibilidad de activar/desactivar esas PTI de manera individual es innecesaria o incluso indeseable.

(3), en algunos casos, una PTI puede no codificarse de la mejor manera como un parámetro lógico (es decir, binario). Por ejemplo, la necesidad de una coordinación entre procesadores en el caso de la compensación de movimiento puede determinarse, al menos en parte, por la longitud del vector de movimiento que apunta hacia fuera del área espacial ubicada conjuntamente y cubierta por la sección o el mosaico. En consecuencia, en una forma de realización, la información de PTI también puede codificarse como un parámetro entero u otro parámetro que no sea un parámetro lógico, para indicar así intervalos de valores adecuados para la

predicción, tal como la longitud máxima de un vector de movimiento que apunta hacia fuera de un límite de segmento.

(4) En algunos casos, los valores de PTI pueden no estar físicamente presentes en el flujo de bits, ya que sus valores pueden obtenerse a partir de otras propiedades del flujo de bits. Por ejemplo, una intra-sección puede no necesitar incluir una PTI relacionada con la compensación de movimiento, ya que la compensación de movimiento puede, por el diseño de la norma, no producirse en una intra-sección.

A continuación se describe el funcionamiento de un codificador que, según formas de realización, puede ser adecuado para su uso con cualquiera de las configuraciones de información de PTI descritas anteriormente.

Haciendo referencia a la FIG. 6, en una forma de realización, un codificador puede funcionar según el diagrama de flujo **600**. Antes de codificar una primera sección de una secuencia de vídeo, el codificador puede determinar (**601**) los ajustes relacionados con secuencias para las PTI, así como un diseño de mosaico de las imágenes de vídeo en la secuencia de vídeo. Esta determinación puede tener en cuenta la arquitectura de hardware del codificador, la arquitectura de hardware de un descodificador, un posible diseño de mosaico sugerido o dictaminado por las arquitecturas de hardware, conocimientos acerca de una red de transmisión (si la hubiera), tales como el tamaño de MTU, etc. En algunos casos, los valores de PTI pueden ser establecidos por una norma a nivel de sistema, que pueden tenerse en cuenta en la determinación realizada por el codificador. Por ejemplo, una futura norma de televisión digital puede requerir, posiblemente, que un determinado diseño de mosaico y ciertas configuraciones de PTI que controlan la predicción y el filtrado en bucle a través de los límites de mosaico se utilicen para ciertas resoluciones (altas), a fin de permitir implementaciones multinúcleo/multiprocesador rentables. Es posible que sólo un subconjunto de todas las PTI tenga que fijarse a nivel de secuencia.

Varias opciones para estas configuraciones ya se han descrito antes.

Después de la determinación, el codificador puede codificar (**602**) las PTI relacionadas con secuencias en una estructura sintáctica de alto nivel apropiada, tal como una secuencia o un conjunto de parámetros de imagen, secuencia, GOP, o cabecera de imagen. El codificador puede tener también la opción (a través de la estructura sintáctica de la norma de codificación de vídeo) de dejar la(s) PTI sin definir durante esta codificación.

Las PTI relacionadas con secuencias pueden permanecer constantes al menos para una imagen de vídeo completa (a no ser que se sobrescriban por PTI basadas en cabecera de sección, como se describe más adelante), pero, en muchos casos, pueden permanecer constantes durante al menos una "secuencia" (por ejemplo, todas las imágenes entre dos imágenes de IDR y la primera imagen de IDR en el flujo de vídeo), y quizás durante toda una sesión de codificación de vídeo. Por ejemplo, las PTI relacionadas con secuencias pueden tener, al menos en parte, limitaciones de hardware que, con poca probabilidad, cambiarán durante una sesión. De ahora en adelante se utilizará, por comodidad, este último caso.

El codificador continúa con la codificación de secciones. Para ello, el codificador puede determinar (**603**) PTI de nivel de sección, que pueden interactuar con PTI relacionadas con secuencias, como ya se ha descrito. Las PTI a nivel de sección pueden codificarse (**604**) como parte de la codificación de la cabecera de sección.

Después, las secciones pueden codificarse (**605**) según cualquier norma de codificación que esté aplicándose, tal como WD4 o H.264, al tiempo que se tiene en cuenta la interrupción de mecanismos de predicción y/o de filtrado en bucle a través de límites tanto de sección como de mosaico según lo indicado por las PTI.

La codificación continúa (**606**) con la siguiente sección.

A continuación se describe el funcionamiento de un descodificador que, según formas de realización, puede ser adecuado para su uso con cualquiera de las configuraciones de información de PTI descritas anteriormente.

La FIG. 7 es un diagrama de flujo **700** de un descodificador que puede utilizarse en una forma de realización de la invención. El descodificador puede recibir (**701**) una unidad NAL desde el flujo de bits y determinar su tipo. Si el tipo de unidad NAL indica un conjunto de parámetros (**702**), el descodificador puede analizar sintácticamente y almacenar (**703**) el conjunto de parámetros según la norma de codificación de vídeo empleada. (Otras estructuras sintácticas de alto nivel, tales como secuencias, GOP, o cabeceras de imagen también pueden utilizarse con este fin).

Si el tipo de unidad NAL indica los datos de sección (**704**) (no se ilustran otros casos) el descodificador puede analizar sintácticamente la cabecera de sección (**705**) y, a continuación, responder según la información, tal como la información de PTI, codificada en la misma. Por ejemplo, la cabecera de sección puede contener una referencia al conjunto de parámetros, y este conjunto de parámetros puede "activarse" (**706**) como se describe en la norma de codificación de vídeo; es decir, los valores del conjunto de parámetros con la referencia indicada pasan a ser válidos. Puesto que las PTI pueden ser parte del conjunto de parámetros, a través de la activación (**706**), sus valores pueden ser válidos también.

La cabecera de sección puede incluir además sus propias PTI, como ya se ha descrito, que pueden ser diferentes de las PTI incluidas en el conjunto de parámetros. Ya se han descrito opciones de cómo arbitrar entre información de PTI codificada en la cabecera de sección e información de PTI codificada en el conjunto de parámetros. Por ejemplo, correlacionando las PTI basadas en cabecera de sección (si las hubiera) con PTI de cabecera de conjunto de parámetros (si las hubiera), y teniendo en cuenta cualquier restricción que pueda estar presente en otras partes de las normas de codificación de vídeo (como, por ejemplo, restricciones y/o configuraciones por defecto de PTI mediante perfiles y niveles), el descodificador puede determinar **(707)** los ajustes de PTI finales que se utilizarán en la descodificación de la sección en cuestión. Cabe señalar que las PTI pueden ser diferentes para diferentes bordes de la sección, dependiendo de los ajustes de PTI del conjunto de parámetros y de los ajustes de PTI de la cabecera de sección, incluido el caso particular de cuando un límite de sección está alineado con un límite de mosaico.

Teniendo en cuenta los ajustes finales de PTI, el descodificador puede decodificar **(708)** la sección usando técnicas de predicción y/o de filtrado en bucle a través de límites de sección o de mosaico según lo indicado por la información codificada en las PTI.

El proceso continúa **(709)** con la siguiente unidad de NAL.

En la FIG. 7 no se muestra el procesamiento de unidades NAL distintas de unidades NAL de sección o de conjunto de parámetros.

La FIG. 8 es un diagrama de bloques que ilustra una implementación basada en un sistema de procesamiento de datos (p.ej., un ordenador personal ("PC")) **800** según una forma de realización de la invención. Hasta este punto, por comodidad, la descripción no ha se referido explícitamente en detalle a posibles implementaciones físicas del codificador y/o del descodificador. Muchas implementaciones físicas diferentes basadas en combinaciones de software y/o componentes son posibles. Por ejemplo, en algunas formas de realización, el/los codificador(es) y/o descodificador(es) de vídeo puede(n) implementarse usando circuitos integrados personalizados o de matrices de puertas, en muchos casos por razones relativas a la rentabilidad y/o un consumo de energía eficiente.

Además, implementaciones basadas en software son posibles utilizando arquitecturas de procesamiento de propósito general, un ejemplo de las cuales es el sistema de procesamiento de datos **800**. Por ejemplo, utilizando un ordenador personal o un dispositivo similar (p.ej., un descodificador, un ordenador portátil, un dispositivo móvil), una estrategia de implementación de este tipo puede ser posible, como se describe a continuación. Como se muestra en la FIG. 8, según las formas de realización descritas, el codificador y/o el descodificador para un PC o dispositivo similar **800** pueden proporcionarse en forma de un medio legible por ordenador **801** (p.ej., CD-ROM, ROM de semiconductor, memorias USB) que contenga instrucciones configuradas para activar un procesador **802**, solo o en combinación con hardware acelerador (p.ej., un procesador de gráficos) **803**, junto con memoria **804** acoplada al procesador **802** y/o el hardware de acelerador **803** para llevar a cabo la codificación o descodificación. El procesador **802**, la memoria **804** y el hardware acelerador **803** pueden estar acoplados a un bus **805** que puede utilizarse para proporcionar el flujo de bits y el vídeo sin comprimir hacia/desde los dispositivos antes mencionados. Dependiendo de la aplicación, pueden acoplarse dispositivos periféricos al bus **805** para la entrada/salida del flujo de bits o el vídeo sin comprimir. Por ejemplo, un cámara **806** puede conectarse a través de una interfaz adecuada, tal como un registrador/visualizador de fotogramas **807** o un enlace USB **808**, al bus **805** para introducir en tiempo real vídeo no comprimido. Una interfaz similar puede usarse en dispositivos de almacenamiento de vídeo sin comprimir, tales como VTR. Puede proporcionarse vídeo sin comprimir a través de un dispositivo de visualización, tal como un monitor de ordenador o una pantalla de televisión **809**. Una unidad DVD RW o equivalente (p.ej., CD-ROM, CD-RW, Blue Ray, memorias USB) **810** puede utilizarse para la entrada y/o salida del flujo de bits. Finalmente, para la transmisión en tiempo real a través de una red **812**, una interfaz de red **811** puede utilizarse para transmitir el flujo de bits y/o el vídeo sin comprimir, dependiendo de la capacidad del enlace de acceso a la red **812** y de la propia red **812**.

Según varias formas de realización, el/los procedimiento(s) descrito(s) anteriormente puede(n) implementarse mediante un módulo de software respectivo. Según otras formas de realización, el/los procedimiento(s) descrito(s) anteriormente puede(n) implementarse mediante un módulo de hardware respectivo. Según otras formas de realización adicionales, el/los procedimiento(s) descrito(s) anteriormente puede(n) implementarse mediante una combinación de módulos de software y/o de hardware.

Aunque, por comodidad, las formas de realización se han descrito principalmente con referencia a un procedimiento de ejemplo, el aparato descrito anteriormente con referencia a un sistema de procesamiento de datos **800** puede, según las formas de realización descritas, programarse para poder llevar a la práctica el/los procedimiento(s) descrito(s). Además, un artículo de fabricación para su uso con un sistema de procesamiento de datos **800**, tal como un dispositivo de almacenamiento pregrabado u otro medio legible por ordenador similar o producto que incluye instrucciones de programa grabadas en el mismo, puede controlar el sistema de procesamiento de datos **800** con el fin de facilitar el llevar a la práctica el/los procedimiento(s) descrito(s). Debe entenderse que tales aparatos y artículos de fabricación, además de los procedimientos descritos, están dentro del alcance de las formas de realización descritas.

5 En particular, las secuencias de instrucciones que cuando se ejecutan hacen que el procedimiento descrito en el presente documento se lleve a cabo por el sistema de procesamiento de datos **800** pueden estar contenidas en un producto portador de datos según una forma de realización de la invención. Este producto portador de datos puede cargarse en y ejecutarse por el sistema de procesamiento de datos **800**. Además, las secuencias de instrucciones que cuando se ejecutan hacen que el procedimiento descrito en el presente documento se lleve a cabo por el sistema de procesamiento de datos **800** pueden estar contenidas en un programa informático o producto de software según una forma de realización de la invención. Este programa informático o producto de software puede cargarse en y ejecutarse por el sistema de procesamiento de datos **800**. Además, las secuencias de instrucciones que cuando se ejecutan hacen que el procedimiento descrito en el presente documento se lleve a cabo por el sistema de procesamiento de datos **800** pueden estar contenidas en un producto de circuito integrado (por ejemplo, un módulo o módulos de hardware) que puede incluir un coprocesador o memoria según una forma de realización de la invención. Este producto de circuito integrado puede instalarse en el sistema de procesamiento de datos **800**.

15 Las anteriores formas de realización pueden contribuir a un sistema y procedimiento mejorados para interrumpir de manera selectiva la predicción y/o el filtrado en bucle de la codificación de vídeo y pueden proporcionar una o más ventajas. Por ejemplo, la inclusión de PTI en mecanismos individuales de predicción y/o de filtrado en bucle, que pertenecen a todas las secciones e imágenes que hacen referencia al conjunto de parámetros, puede ayudar a adaptar el flujo de bits al entorno de codificación y/o decodificación, tal como la arquitectura de hardware del codificador o del decodificador. Además, la inclusión de PTI en mecanismos individuales de predicción y/o de filtrado en bucle que pertenecen a una sección dada puede contribuir a adaptar el flujo de bits al contenido, mejorándose así la eficacia de la codificación.

REIVINDICACIONES

1. Un procedimiento para descodificar un flujo de bits (206, 301, 401, 501) correspondiente a una imagen de vídeo codificada (201) que comprende una pluralidad de segmentos de mosaico (202, 203) que coexisten con secciones de la imagen de vídeo codificada de tal manera que la imagen de vídeo codificada (201, 301) se divide en los segmentos de mosaico mediante límites de segmento de mosaico (204) y en secciones mediante límites de sección, comprendiendo el procedimiento:

en relación con al menos un segmento de mosaico de la imagen de vídeo codificada que no tiene una cabecera de segmento de mosaico asociada, obtener, a partir de un conjunto de parámetros (207) contenido en el flujo de bits (206, 301, 401, 501) una indicación (313, 414) de que una operación de desfase adaptativo de muestras va a aplicarse a través de límites de segmento de mosaico de la imagen de vídeo codificada, donde la indicación del conjunto de parámetros se define de manera individual para la operación de desfase adaptativo de muestras;

en relación con al menos una sección, obtener a partir de una cabecera de sección (405) una indicación de que una operación de desfase adaptativo de muestras va a aplicarse a través de un límite de sección, donde la indicación de la cabecera de sección se define de manera individual para la operación de desfase adaptativo de muestras; y

controlar la operación de desfase adaptativo de muestras a través de un límite de segmento de mosaico entre segmentos de mosaico en respuesta a la indicación del conjunto de parámetros (207) y controlar la operación de desfase adaptativo de muestras a través del límite de sección en respuesta a la indicación de la cabecera de sección (405),

donde si un límite de segmento de mosaico y un límite de sección están alineados y la indicación del conjunto de parámetros contradice la indicación de la cabecera de sección, la indicación del conjunto de parámetros sobrescribe la indicación de la cabecera de sección,

donde la imagen de vídeo codificada contiene al menos dos segmentos de mosaico sin una cabecera de segmento asociada.

2. Un procedimiento para codificar, en un flujo de bits (206, 301, 401, 501) correspondiente, una imagen de vídeo (201) que comprende una pluralidad de segmentos de mosaico que coexisten con secciones de la imagen de vídeo de tal manera que la imagen de vídeo (201, 301) se divide en los segmentos de mosaico mediante límites de segmento de mosaico (204) y en secciones mediante límites de sección, comprendiendo el procedimiento:

en relación con al menos un segmento de mosaico de la imagen de vídeo que no tiene una cabecera de segmento de mosaico asociada, obtener una primera indicación (313, 414) de que una operación de desfase adaptativo de muestras va a aplicarse al al menos un segmento de mosaico que no tiene una cabecera de segmento asociada;

en relación con al menos una sección, obtener una segunda indicación de que una operación de desfase adaptativo de muestras va a aplicarse a través de un límite de sección;

controlar la operación de desfase adaptativo de muestras durante la codificación de la imagen de vídeo a través de un límite entre segmentos de mosaico en respuesta a la primera indicación y controlar la operación de desfase adaptativo de muestras durante la codificación de la imagen de vídeo a través de un límite de sección entre segmentos en respuesta a la segunda indicación; y

codificar la primera indicación en un conjunto de parámetros (207) contenido en el flujo de bits (206, 301, 401, 501) y la segunda indicación en una cabecera de sección (405) del flujo de bits (206, 301, 401, 501), donde la primera indicación codificada en el conjunto de parámetros se define de manera individual para la operación de desfase adaptativo de muestra y la segunda indicación codificada en la cabecera de sección se define de manera individual para la operación de desfase adaptativo de muestras,

donde si un límite de segmento de mosaico y un límite de sección están alineados y la primera indicación contradice la segunda indicación, la primera indicación sobrescribe la segunda indicación,

donde la imagen de vídeo contiene al menos dos segmentos de mosaico sin una cabecera de segmento asociada.

FIG. 1 100

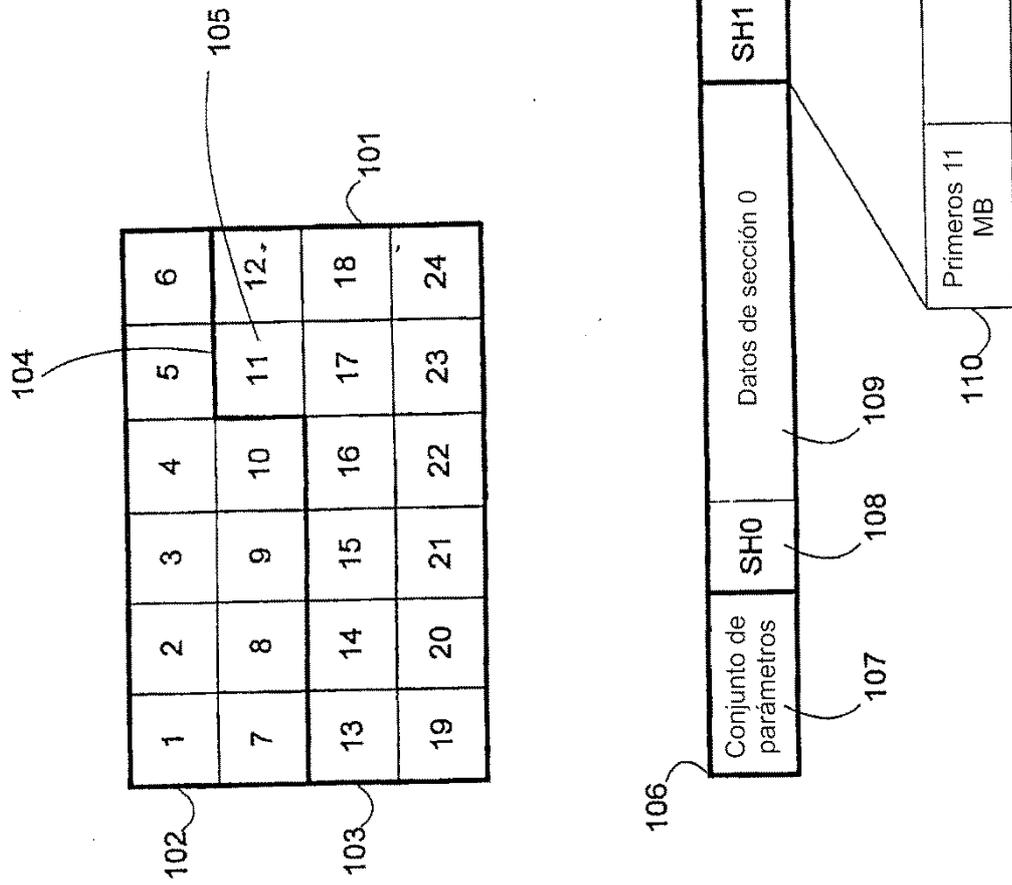


FIG. 2 200

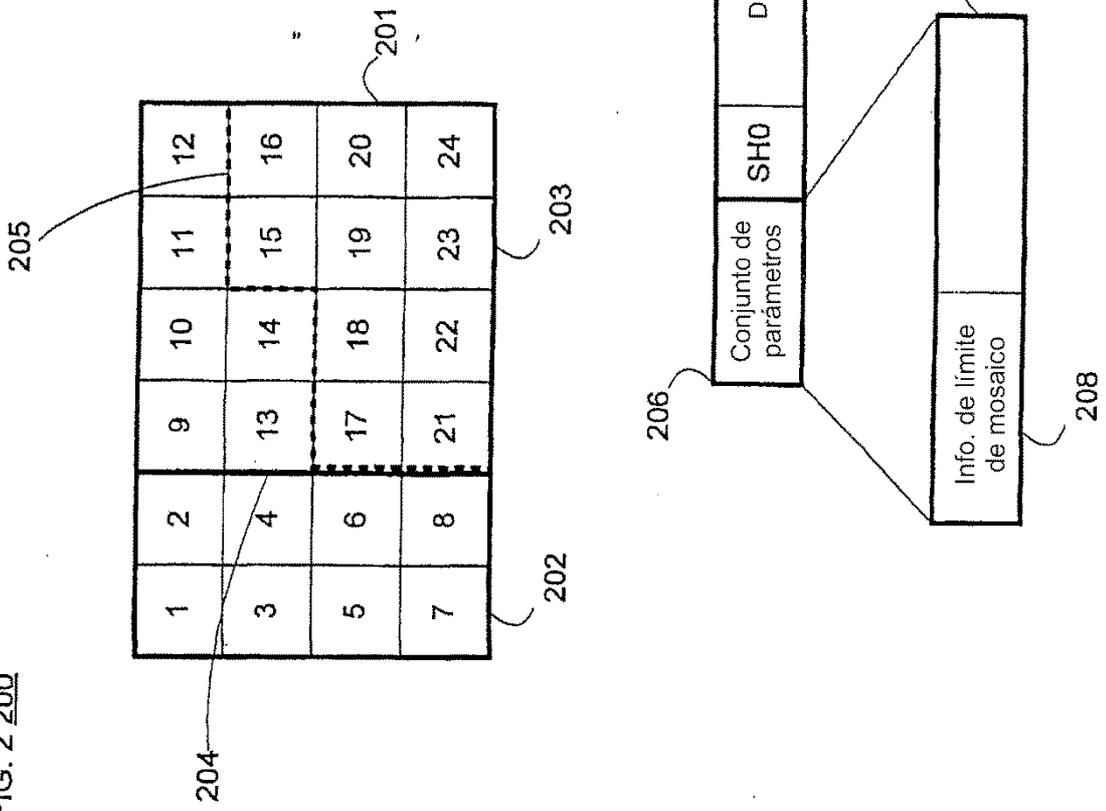


FIG 3 300

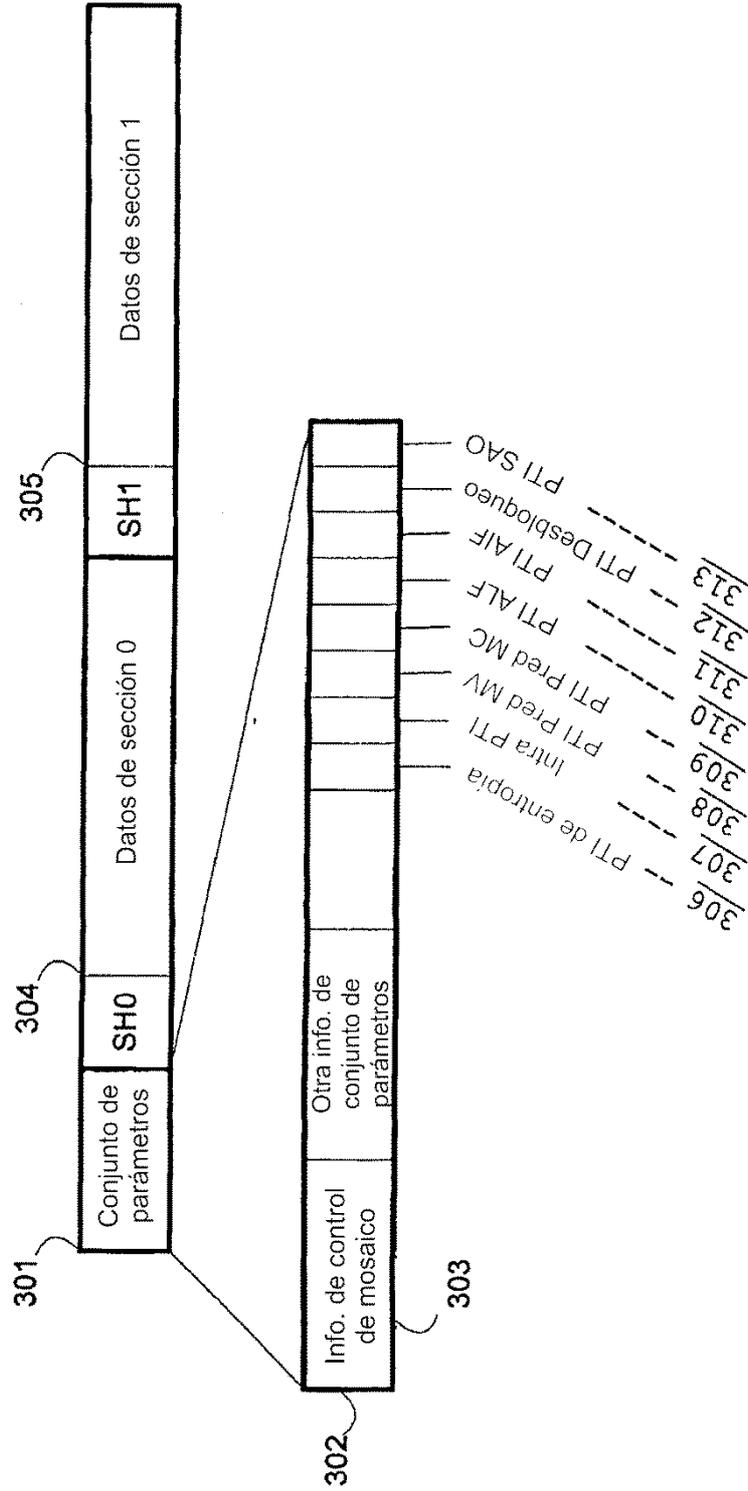


FIG. 4 400

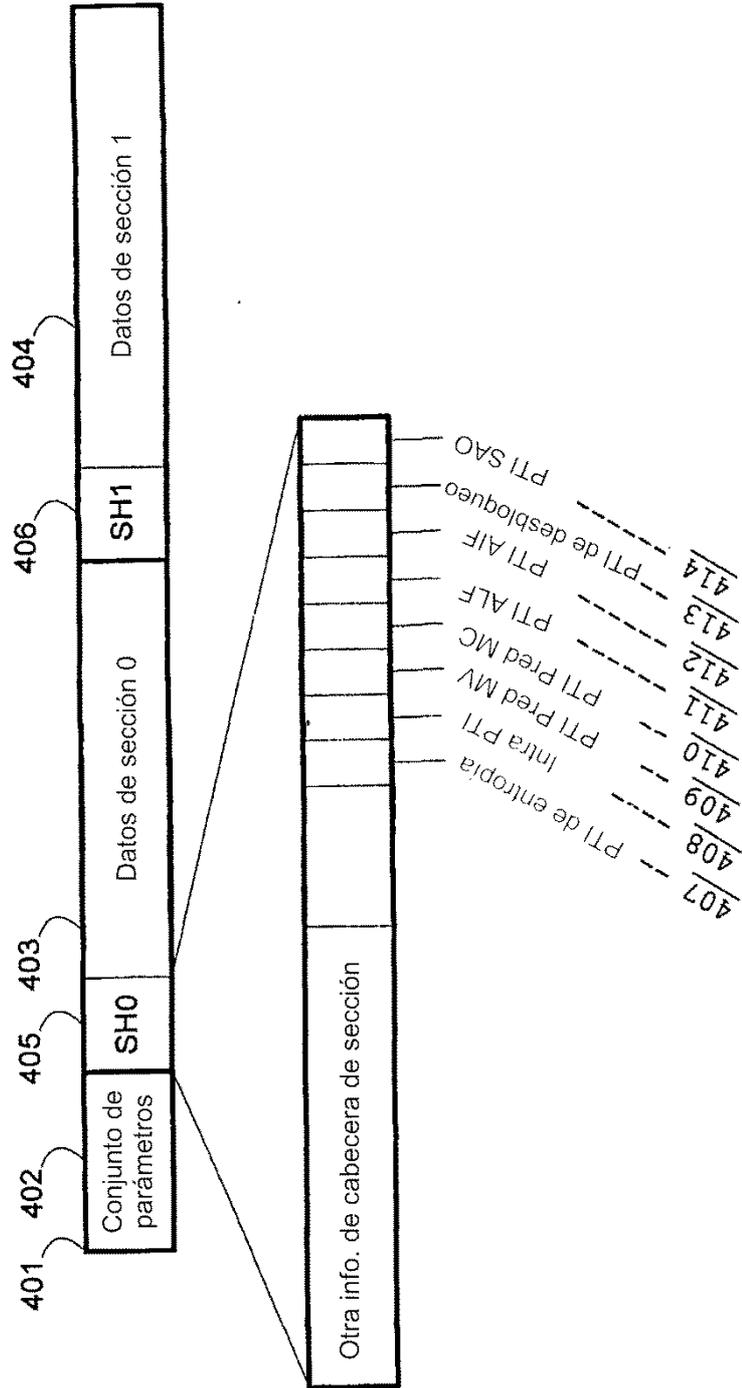


FIG. 5 500

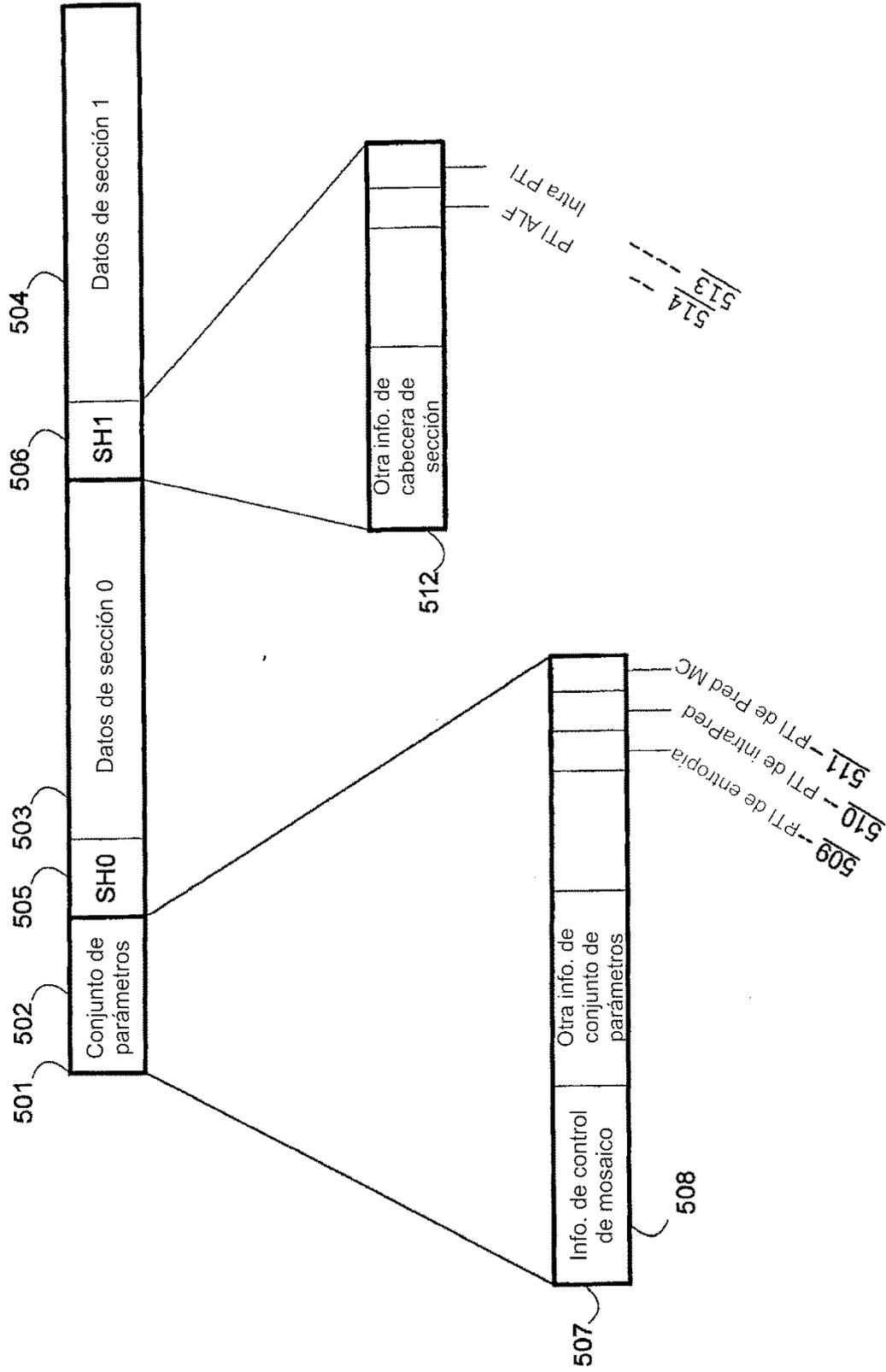


FIG. 6 600

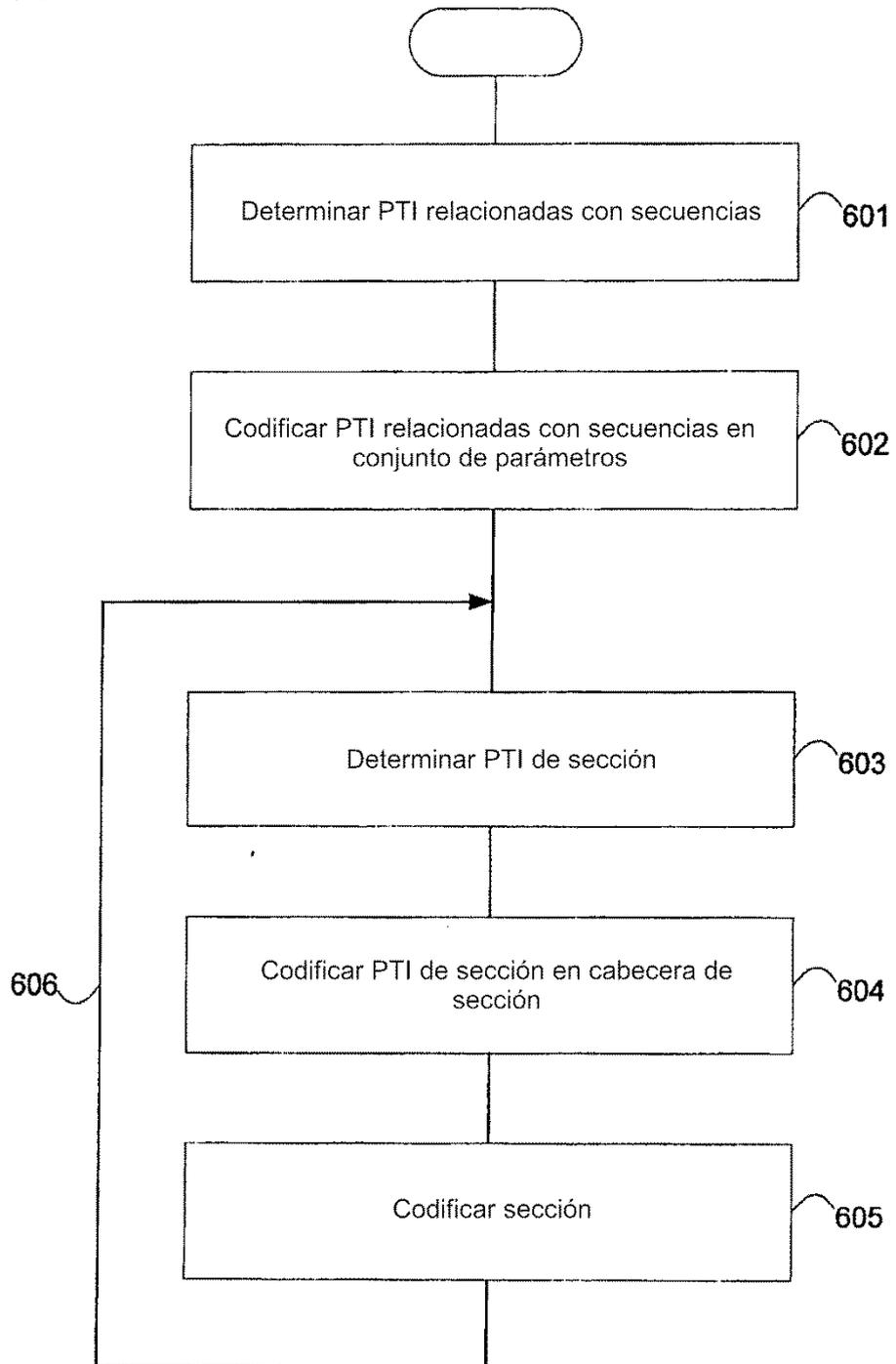


FIG. 7 700

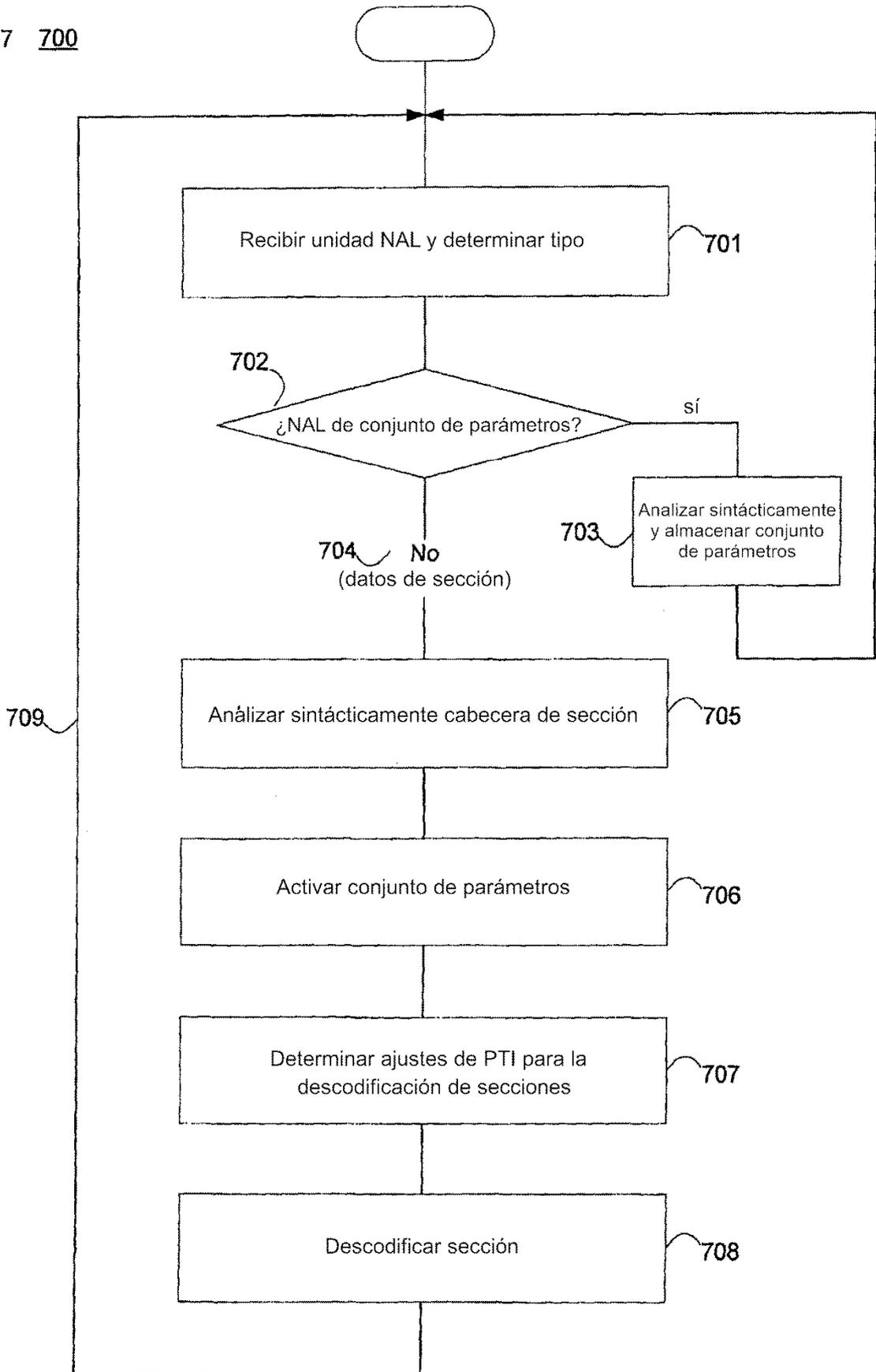


FIG. 8 800

