

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 728 146**

51 Int. Cl.:

**H04N 19/52** (2014.01)  
**H04N 19/51** (2014.01)  
**H04N 19/31** (2014.01)  
**H04N 19/61** (2014.01)  
**H04N 19/159** (2014.01)  
**H04N 19/70** (2014.01)  
**H04N 19/577** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **16.01.2013 PCT/JP2013/000150**
- 87 Fecha y número de publicación internacional: **25.07.2013 WO13108616**
- 96 Fecha de presentación y número de la solicitud europea: **16.01.2013 E 13738409 (5)**
- 97 Fecha y número de publicación de la concesión europea: **06.03.2019 EP 2805511**

54 Título: **Procedimientos y aparato de codificación y decodificación de vídeo utilizando predicción temporal de vector de movimiento**

30 Prioridad:

**20.01.2012 US 201261589261 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**22.10.2019**

73 Titular/es:

**SUN PATENT TRUST (100.0%)  
450 Lexington Avenue, 38th Floor  
New York, NY 10017, US**

72 Inventor/es:

**NAING, SUE MON THET;  
WAHADANIAH, VIKTOR;  
LIM, CHONG SOON;  
SUN, HAI WEI;  
WIN, KYAW KYAW;  
NISHI, TAKAHIRO;  
SASAI, HISAO;  
SHIBAHARA, YOUJI;  
SUGIO, TOSHIYASU;  
TANIKAWA, KYOKO;  
MATSUNOBU, TORU y  
TERADA, KENGO**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

ES 2 728 146 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**

Procedimientos y aparato de codificación y decodificación de vídeo utilizando predicción temporal de vector de movimiento

5 La presente invención se refiere a un procedimiento de codificación de un vídeo y un procedimiento para descodificar un vídeo usando la predicción temporal de vectores de movimiento, y aparatos de los mismos. La presente invención se puede aplicar en cualquier codificación de datos multimedia y, más particularmente, en la codificación de contenidos de imágenes y vídeos utilizando la predicción temporal de vectores de movimiento para la predicción inter-imágenes.

10 Los esquemas de codificación de vídeo, tales como H.264/MPEG4 AVC y el próximo HEVC (codificación de vídeo de alta eficiencia) realizan la codificación/descodificación del contenido de imagen/vídeo usando la predicción inter-imágenes (o simplemente "inter") a partir de imágenes de referencia previamente codificadas/descodificadas para explotar la redundancia de información a través de imágenes consecutivas en el tiempo.

15 En una secuencia de bits de vídeo codificada, una imagen de referencia utilizada para el proceso de predicción inter de una unidad de predicción (tal como un bloque MxN de muestras) se identifica o se conoce usando un índice de referencia. Un índice de referencia es un índice en una lista ordenada (conocida como lista de imágenes de referencia) que comprende una o más imágenes de referencia. Cada índice de referencia está asociado de forma única con una imagen de referencia en la lista de imágenes de referencia. Es decir, el índice de referencia es un valor que se utiliza para distinguir varias imágenes de referencia entre sí.

20 Los esquemas de codificación mencionados anteriormente admiten la predicción temporal de vectores de movimiento (es decir, predicción de vector de movimiento o MVP), conforme a la cual los vectores de movimiento de un bloque objetivo de muestras se predicen a partir de los vectores de movimiento de uno o más bloques de muestras previamente codificados en una imagen de referencia co-localizada. La predicción temporal del vector de movimiento reduce aún más la tasa de bits asociada con los vectores de movimiento al explotar temporalmente la redundancia de información entre los vectores de movimiento vecinos. La imagen de referencia co-localizada se  
25 seleccionando la primera imagen de referencia en una lista predeterminada de imágenes de referencia (tal como la lista 0 de imágenes de referencia) como imagen de referencia co-localizada.

30 En aplicaciones que requieren transmisión de vídeo a través de un entorno con pérdidas, la predicción temporal del vector de movimiento es susceptible de una predicción errónea del vector de movimiento cuando la imagen de referencia co-localizada se pierde o contiene errores. En el estándar HEVC en desarrollo, se describió una técnica para deshabilitar la predicción temporal del vector de movimiento de ciertas unidades de sub-imágenes (por ejemplo, un corte). JCTVC-G398, "High-level Syntax: Marking process for non-TMVP pictures", Equipo de colaboración conjunto sobre codificación de vídeo (JCT-VC) de ITU-T SG 16 WP3 e ISO/CEI JTC1/SC29/WG11 7ª reunión, Ginebra, CH, noviembre de 2011. En esta técnica, es necesario introducir un indicador de marcado en un conjunto  
35 de parámetros de imagen (PPS) utilizado para marcar una imagen en la memoria intermedia de imagen del descodificador (DPB) como "no utilizado para la predicción temporal de vector de movimiento". Este proceso de marcado se realiza mediante un descodificador cuando una unidad de sub-imagen se refiere a un PPS que tiene un indicador de marcado igual a "VERDADERO". Otros ejemplos de indicadores que indican la predicción temporal del vector de movimiento se describen en las NPL 3 y 4.

40 Bibliografía que no es de patentes

NPL 1: ISO/IEC 14496-10, "MPEG-4 Part 10 Advanced Video Coding"

NPL 2: JCTVC-G398, "High-level Syntax: Marking process for non-TMVP pictures", Equipo de colaboración conjunto sobre codificación de vídeo (JCT-VC) de ITU-T SG 16 WP3 e ISO/CEI JTC1/SC29/WG11 7ª reunión, Ginebra, CH, noviembre de 2011.

45 NPL 3: ALEXIS MICHAEL TOURAPIS, FENG WU, SHIPENG LI: "Direct Mode Coding for Bipredictive Slices in the H.264 Standard", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE SERVICE CENTER, PISCATAWAY, NJ, EE. UU., vol. 15, n° 1, 1 de enero de 2005 (2005-01-01), páginas 119-126

50 NPL 4: LI (USTC) B ET AL: "Constrained temporal motion vector prediction for error resilience",4. REUNIÓN JCTVC; 95. REUNIÓN MPEG; 20-1-2011 - 28-1-2011; DAEGU; (Equipo de colaboración conjunto sobre codificación de vídeo de ISO/IEC JTC1/SC29/WG11 e ITU-T SG.16); URL: HTTP:// WFTP3.ITU.INT/AV-ARCH/JCTVC-SITE/,n°. JCTVC-D139, 15 de enero de 2011 (2011-1-15)

55 Como se mencionó en los antecedentes, en una técnica descrita para deshabilitar la predicción temporal del vector de movimiento de ciertos cortes, es necesario introducir un indicador de marcado en un conjunto de parámetros de imagen (PPS) utilizado para marcar una imagen en la memoria intermedia de imagen del descodificador (DPB) como "no utilizado para la predicción temporal de vectores de movimiento". Un problema importante asociado con esta

técnica es que cuando un corte que invoca el proceso de marcado se pierde o contiene un error, el descodificador no puede realizar el proceso de marcado deseado. Como resultado, se pierde la sincronización posterior entre el codificador y el descodificador. Por lo tanto, la técnica mencionada anteriormente para deshabilitar la predicción temporal del vector de movimiento no es robusta.

5 La presente invención busca proporcionar procedimientos y aparatos para codificar y descodificar vídeo usando la predicción temporal de vectores de movimiento con una robustez de error mejorada. En particular, la predicción temporal del vector de movimiento para una unidad de sub-imagen (por ejemplo, un corte) se habilita/inhabilita de una manera que es menos susceptible de errores. Por ejemplo, según las realizaciones de la presente invención, se elimina el proceso de marcado mencionado anteriormente (es decir, para marcar imágenes de referencia como "no  
10 utilizado para la predicción temporal del vector de movimiento") que debe realizar el descodificador.

Las realizaciones de la presente invención proporcionan procedimientos y aparatos para codificar y descodificar vídeo usando la predicción temporal de vectores de movimiento con una robustez de error mejorada de la predicción entre imágenes. Por ejemplo, las realizaciones también pueden dar como resultado una flexibilidad y una eficiencia de codificación mejoradas de la predicción inter-imágenes, ya que la predicción temporal del vector de movimiento  
15 puede habilitarse e inhabilitarse independientemente para una pluralidad de unidades de sub-imagen dentro de la misma imagen.

La figura 1 representa una representación del diagrama de despiece de una secuencia de bits de vídeo codificado a modo de ejemplo según una realización de la presente invención;

20 la figura 2 muestra un diagrama de flujo que ilustra un procedimiento para codificar un vídeo según una realización de la presente invención;

la figura 3 muestra un diagrama de bloques esquemático de un aparato de ejemplo para codificar una secuencia de bits de vídeo/imagen de entrada;

la figura 4 representa un diagrama de flujo que ilustra un procedimiento para descodificar un vídeo codificado según una realización de la presente invención;

25 la figura 5 representa un diagrama de bloques esquemático de un aparato de ejemplo para descodificar una secuencia de bits codificada de entrada;

la figura 6 representa un diagrama que muestra las diferentes capas temporales para un grupo de imágenes de ejemplo;

30 la figura 7 representa un diagrama de flujo que ilustra un procedimiento para determinar el valor del indicador de uso de predicción temporal de vector de movimiento según una primera realización;

la figura 8 representa un diagrama de flujo que ilustra un procedimiento para determinar el valor del indicador de uso de predicción temporal del vector de movimiento según una segunda realización;

la figura 9 representa un diagrama de flujo que ilustra un procedimiento para determinar el valor del indicador de uso de predicción temporal de vector de movimiento según una tercera realización;

35 la figura 10 representa una representación en diagrama de un flujo de unidad NAL, es decir, una serie de unidades NAL para una secuencia de bits de vídeo codificado;

la figura 11 representa una representación en diagrama de una imagen RAP de ejemplo que contiene múltiples vistas/capas con múltiples cortes;

40 la figura 12 representa un diagrama de flujo que ilustra un procedimiento para determinar el valor del indicador de uso de predicción temporal de vectores de movimiento según una cuarta realización;

la figura 13 muestra una configuración general de un sistema que proporciona contenido para implementar servicios de distribución de contenido;

la figura 14 muestra una configuración general de un sistema de transmisión digital;

la figura 15 muestra un diagrama de bloques que ilustra un ejemplo de una configuración de un televisor;

45 la figura 16 muestra un diagrama de bloques que ilustra un ejemplo de una configuración de una unidad de reproducción/grabación de información que lee y escribe información desde y sobre un medio de grabación que es un disco óptico;

la figura 17 muestra un ejemplo de una configuración de un medio de grabación que es un disco óptico;

la figura 18A muestra un ejemplo de un teléfono móvil;

- la figura 18B es un diagrama de bloques que muestra un ejemplo de una configuración de un teléfono móvil;
- la figura 19 ilustra una estructura de datos multiplexados;
- la figura 20 muestra esquemáticamente cómo se multiplexa cada secuencia en datos multiplexados;
- 5 la figura 21 muestra cómo una secuencia de vídeo se almacena en una secuencia de paquetes PES con más detalle;
- la figura 22 muestra una estructura de paquetes TS y paquetes fuente en los datos multiplexados;
- la figura 23 muestra una estructura de datos de un PMT;
- la figura 24 muestra una estructura interna de información de datos multiplexados;
- la figura 25 muestra una estructura interna de información de atributos de secuencia;
- 10 la figura 26 muestra los pasos para identificar datos de vídeo;
- la figura 27 muestra un ejemplo de una configuración de un circuito integrado para implementar el procedimiento de codificación de imágenes en movimiento y el procedimiento de descodificación de imágenes en movimiento según cada una de las realizaciones;
- la figura 28 muestra una configuración para conmutar entre las frecuencias de control;
- 15 la figura 29 muestra los pasos para identificar datos de vídeo y conmutar entre las frecuencias de control;
- la figura 30 muestra un ejemplo de una tabla de consulta en la que los estándares de datos de vídeo están asociados con las frecuencias de control;
- la figura 31A es un diagrama que muestra un ejemplo de una configuración para compartir un módulo de una unidad de procesamiento de señal;
- 20 la figura 31B es un diagrama que muestra otro ejemplo de una configuración para compartir un módulo de la unidad de procesamiento de señal.

Según realizaciones de ejemplo de la presente invención, se proporciona un procedimiento para codificar vídeo y un procedimiento para descodificar vídeo usando predicción temporal de vectores de movimiento (TMVP), y aparatos de los mismos. En particular, la predicción temporal del vector de movimiento para una unidad de sub-imagen (por ejemplo, corte) se habilita/inhabilita de una manera que es menos susceptible de errores. Para lograr esto, según una realización preferida de la presente invención, se introduce un indicador en una cabecera de una imagen o, más preferiblemente, en una cabecera de una unidad de sub-imagen para indicar si la predicción temporal del vector de movimiento se usa o no para la predicción inter-imágenes (o simplemente "inter") de la unidad de sub-imagen. Este indicador también se puede denominar indicador de uso de predicción temporal de vector de movimiento. En aspectos adicionales de la presente invención, las técnicas preferidas para determinar/decidir el valor del indicador se describen en diversas realizaciones.

25

30

Para mayor claridad y simplicidad, ahora se describirán realizaciones de ejemplo de la presente invención con detalles adicionales en los cuales la unidad de sub-imagen es un corte de una imagen. Un experto en la técnica apreciará que la división en cortes es simplemente un procedimiento posible para dividir una imagen en varias particiones de sub-imágenes. Por lo tanto, las realizaciones de la presente invención descritas a continuación no están limitadas a que la unidad de sub-imagen sea un corte. Por ejemplo, otros procedimientos de división en sub-imágenes, tales como mosaicos, cortes de entropía y unidades de división de frente de onda, están todos dentro del ámbito de la presente invención.

35

La figura 1 es una representación esquemática del diagrama de despiece de una secuencia 100 de bits de vídeo codificado de ejemplo según una realización de la presente invención. La secuencia 100 de bits de vídeo codificado comprende una cabecera 110 y una pluralidad de imágenes 112 asociadas con la cabecera 110. La imagen 112 se divide típicamente en una pluralidad de unidades 114 de subimagen (por ejemplo, cortes). Cada corte 114 comprende una cabecera 116 del corte y datos 118 de un corte asociada con la cabecera 116 del corte. Los datos 118 del corte comprenden una pluralidad de unidades 120 de predicción de tipo de predicción inter.

40

En la realización de ejemplo como se ilustra en la figura 1, el indicador 122 para indicar si la predicción temporal del vector de movimiento se usa o no para la inter-predicción del corte 114 se localiza preferiblemente en la cabecera 116 del corte. Como resultado, la predicción temporal del vector de movimiento para cada corte 114 se puede habilitar e inhabilitar independientemente de otras cortes 114 de la misma imagen 112. La cabecera 116 del corte comprende además los parámetros 124 de ordenación de listas de imágenes de referencia para especificar el orden de las imágenes de referencia en una o más listas de imágenes de referencia. Estos parámetros 124 determinan el orden efectivo o final de las imágenes de referencia en las listas de imágenes de referencia utilizadas para la

45

50

predicción inter del corte 114 asociado con la cabecera 116 del corte. Estos parámetros 124 pueden especificar que se realice un proceso de reordenación en una o más de las listas de imágenes de referencia iniciales, o puede especificar que las listas de imágenes de referencia iniciales se utilicen sin reordenar. Como se muestra en la figura 1, el indicador 122 se localiza preferiblemente en la misma cabecera 116 de corte que los parámetros 124 de ordenación de listas de imágenes de referencia. Un parámetro 126 de selección de predictor de vector de movimiento está localizado en cada unidad 120 de predicción para seleccionar un predictor de vector de movimiento de entre la pluralidad de predictores de vectores de movimiento disponibles para la predicción inter de la unidad 120 de predicción.

En otra realización, los parámetros 124 de ordenación de listas de imágenes de referencia y el indicador 122 de uso del predictor temporal de vector de movimiento se encuentran en una cabecera (no mostrada) compartida por una pluralidad de cortes 114 dentro de la misma imagen 112. Por ejemplo, la cabecera 110 de nivel de la imagen puede ser el conjunto de parámetros de adaptación (APS) o el segmento de cabecera común de corte en el esquema de codificación HEVC.

Como se explicó anteriormente, la división en cortes es meramente un procedimiento posible para dividir una imagen en varias particiones de sub-imágenes. Se pueden usar otros posibles procedimientos de división en sub-imágenes, por ejemplo, mosaicos, cortes de entropía y unidades de división de frente de onda. En dichos otros procedimientos de división en sub-imágenes, los parámetros 124 y el indicador 122 localizados en la cabecera 116 del corte como se describe anteriormente pueden localizarse en cambio en una cabecera de una unidad de sub-imagen.

La figura 2 representa un diagrama de flujo que ilustra un procedimiento 200 para codificar un vídeo según una realización de la presente invención. En el paso S202, uno o más parámetros 124 (es decir, parámetros de ordenación de listas de imágenes de referencia) se escriben en una cabecera 116 de un corte 114 para especificar el orden de las imágenes de referencia en una o más listas de imágenes de referencia utilizadas para la predicción inter del corte 124. Una localización predeterminada (tal como la primera imagen) en una de las listas de imágenes de referencia (tal como la lista 0 de imágenes de referencia) indica la imagen de referencia co-localizada. En el paso S204, se determina el valor de un indicador 122 que indica si se usa o no se usa el vector de movimiento temporal para la predicción inter del corte 124. Una serie de técnicas para determinar el valor del indicador 122 se describirán más adelante según diversas realizaciones de la presente invención. Posteriormente, en el paso 206, el indicador 122 se escribe en la cabecera 116 del corte 114. En el paso S208, el valor del indicador 122 se analiza o juzga para determinar si el indicador 122 indica que la predicción temporal del vector de movimiento se usa o no se usa. Por ejemplo, un indicador 122 que tiene un valor "0" puede indicar que no se usa la predicción temporal del vector de movimiento y un indicador 122 que tiene un valor "1" puede indicar que se usa la predicción temporal del vector de movimiento, o viceversa.

Si el indicador 122 indica que se utiliza la predicción temporal del vector de movimiento, en el paso S210, se crea una lista de predictores de vectores de movimiento (por ejemplo, una primera lista) que comprende una pluralidad de predictores de vectores de movimiento que incluyen al menos un predictor temporal de vector de movimiento temporal de al menos un vector de movimiento de la imagen de referencia co-localizada. Solo a modo de ejemplo, la pluralidad de vectores de movimiento puede incluir al menos un predictor temporal de vector de movimiento, uno o más vectores de movimiento derivados de unidades/bloques de predicción espacial vecinos (es decir, predictores espaciales de vector de movimiento) y un vector de movimiento cero. En el paso S212, se selecciona un predictor de vector de movimiento de la lista de predictores de vector de movimiento para un bloque objetivo de muestras (es decir, unidad de predicción) 120 en el corte 124. En el paso 214, un parámetro 126 (es decir, un parámetro de selección de predictor de vector de movimiento) (por ejemplo, un primer parámetro) se escribe en la secuencia 100 de bits de vídeo codificado (es decir, en la unidad 120 de predicción del corte 114) para indicar el predictor de vector de movimiento seleccionado de entre la lista de predictores de vector de movimiento.

Por otra parte, si el indicador 122 indica que no se usa la predicción temporal del vector de movimiento, en el paso S216, se crea una lista de predictores de vector de movimiento (por ejemplo, una segunda lista) que comprende una pluralidad de predictores de vector de movimiento sin ningún predictor temporal de vector de movimiento. En el paso S218, se selecciona un predictor de vector de movimiento de la lista de predictores de vector de movimiento para un bloque objetivo de muestras (es decir, unidad de predicción) en el corte 124. En el paso S220, se escribe un parámetro (es decir, un parámetro de selección de predictor de vector de movimiento) (por ejemplo, un segundo parámetro) en la secuencia 100 de bits de vídeo codificado (es decir, en cada unidad 120 de predicción de los datos 118 de corte asociados con la cabecera del corte 116) para indicar el predictor de vector de movimiento seleccionado de entre la lista predictores de vectores de movimiento.

Después del paso S214 o del paso S220, se realiza una predicción inter compensada por movimiento para el corte 214 utilizando el predictor de vector de movimiento seleccionado para producir un bloque de muestras de predicción. Posteriormente, en el paso S226, el bloque de muestras de predicción se resta de un bloque de muestras originales para producir un bloque de muestras residuales. Posteriormente, en el paso S226, el bloque de muestras residuales correspondientes al bloque objetivo se codifica en la secuencia 100 de bits de vídeo codificado.

Por consiguiente, en la realización descrita anteriormente de la presente invención, el indicador 122 para indicar si se usa o no se usa la predicción temporal del vector de movimiento es capaz de controlar un corte 114

independientemente de otras cortes 114 en la misma imagen 112. Por lo tanto, el indicador 122 correspondiente a una primera corte 114 no determina si la predicción temporal del vector de movimiento se usa o no se usa en una segunda u otra corte de la misma imagen 112. Además, en la realización descrita anteriormente, el proceso de marcado en las imágenes de referencia en la memoria intermedia de la imagen del descodificador (DPB) como se describe en los antecedentes se ha eliminado. Esto da como resultado una mayor flexibilidad y eficiencia de codificación de la predicción inter.

En una realización de la presente invención, la primera y la segunda listas de predictores de vectores de movimiento comprenden diferente número de predictores de vectores de movimiento. Preferiblemente, la segunda lista comprende un predictor de vector de movimiento menos que la primera lista. Tanto en la primera como en la segunda lista, los predictores de vector de movimiento distintos del predictor temporal de vector de movimiento pueden ser iguales o equivalentes. Esto puede aumentar la eficiencia de la codificación, ya que el codificador tiene más opciones para seleccionar el mejor candidato de una lista que incluye un predictor temporal de vector de movimiento (es decir, de la primera lista). La segunda lista puede proporcionar una mejor resiliencia al error porque no se utiliza la predicción temporal del vector de movimiento. En la secuencia 100 de bits de vídeo codificado, el primer y segundo parámetros que representan el predictor de vector de movimiento seleccionado pueden usar diferentes representaciones de bits, por ejemplo, usando representaciones unarias truncadas que tienen diferentes valores máximos en la binarización de codificación aritmética o en el código de longitud variable.

En otra realización de la presente invención, la primera y la segunda lista comprenden el mismo número predeterminado de predictores de vectores de movimiento. En lugar del predictor temporal de vector de movimiento, la segunda lista comprende otro predictor de vector de movimiento predeterminado único que no está presente en la primera lista. Esto puede aumentar la eficiencia de la codificación, ya que el codificador tiene más opciones para seleccionar el mejor candidato de una lista que incluye un único predictor de vector de movimiento predeterminado (es decir, de la segunda lista). Dado que el número máximo de predictores temporales de vectores de movimiento candidatos es el mismo tanto para la primera como para la segunda lista, esto reduce la complejidad en el proceso de análisis del parámetro del índice utilizado para indicar el predictor de vector de movimiento seleccionado. El predictor de vector de movimiento único se deriva sin dependencia temporal, es decir, sin usar vectores de movimiento de ninguna imagen de referencia. Solo a modo de ejemplo, el predictor único de vector de movimiento puede ser un predictor espacial de vector de movimiento de una localización vecina predeterminada. Como otro ejemplo, el predictor único de vector de movimiento puede ser un predictor de vector de movimiento cero.

A continuación, se describirá un aparato 300 de ejemplo para codificar un vídeo según una realización de la presente invención.

La figura 3 representa un diagrama de bloques esquemático del aparato 300 de ejemplo para codificar una secuencia 302 de bits de vídeo/imagen de entrada bloque a bloque para generar una secuencia 304 de bits de vídeo codificado. El aparato 300 comprende una unidad 306 de transformación operable para transformar un dato de entrada en coeficientes de frecuencia, una unidad 308 de cuantificación operable para realizar la cuantificación de unos datos de entrada, una unidad 310 de cuantificación inversa operable para realizar una cuantificación inversa de unos datos de entrada, una unidad 312 de transformación inversa operable para realizar la transformación de frecuencia inversa sobre unos datos de entrada, una memoria 314 de bloque y una memoria 316 de imagen operables para almacenar datos tales como vídeos e imágenes, una unidad 318 de predicción intra operable para realizar predicción intra, una unidad 320 de predicción inter operable para realizar predicción inter, una unidad 322 de codificación de entropía operable para codificar unos datos de entrada en una secuencia 304 de bits de vídeo codificado, una unidad 324 de control operable para decidir si la predicción temporal del vector de movimiento es usada o no es usada para la predicción inter de un corte objetivo, una unidad 330 de predicción de vector de movimiento y una unidad 328 de escritura operable para escribir datos en la secuencia 304 de bits de vídeo codificado.

Para mayor claridad, ahora se describirá un flujo de datos de ejemplo a través del aparato 300 como se muestra en la figura 3. Un vídeo 302 de entrada se introduce en un sumador, y el valor 305 agregado se envía a la unidad 306 de transformación. La unidad 306 de transformación transforma los valores 305 agregados en coeficientes de frecuencia, y envía los coeficientes 307 de frecuencia resultantes a la unidad 308 de cuantificación. La unidad 308 de cuantificación cuantifica los coeficientes 307 de frecuencia ingresados, y envía los valores 309 cuantificados resultantes a la unidad 310 de cuantificación inversa y a la unidad 322 de codificación de entropía. La unidad 322 de codificación de entropía codifica la salida de los valores 309 cuantificados de la unidad 308 de cuantificación, y envía una secuencia 304 de bits de vídeo codificado.

La unidad 310 de cuantificación inversa cuantifica de manera inversa la salida de los valores 309 cuantificados de la unidad 308 de cuantificación, y envía los coeficientes 311 de frecuencia a la unidad 312 de transformación inversa. La unidad 312 de transformación inversa realiza la transformada de frecuencia inversa sobre los coeficientes 311 de frecuencia, con el fin de transformar los coeficientes de frecuencia en valores de muestra de la secuencia de bits, y envía los valores 313 de muestra resultantes a un sumador. El sumador agrega los valores 313 de muestra de la salida de la secuencia de bits de la unidad 314 de transformación inversa a la salida de los valores 319 de vídeo/imagen predichos de la unidad 318 o 320 de predicción intra o inter, y envía los valores 315 agregados resultantes a la memoria 105 de bloque o a la memoria 106 de imagen para una predicción adicional. La unidad 318

o 320 de predicción intra o inter busca dentro de los vídeos/imágenes reconstruidos almacenados en la memoria 314 de bloque o en la memoria 316 de imágenes, y estima un área de vídeo/imagen que es, por ejemplo, lo más similar a los vídeos/imágenes de entrada para la predicción.

5 La unidad 324 de control toma una decisión sobre si la predicción temporal del vector de movimiento se usa o no se usa para la predicción inter de un corte objetivo y emite una señal 325 que indica la decisión a la unidad 330 de predicción del vector de movimiento y a la unidad 322 de escritura. Una serie de técnicas para decidir/determinar si la predicción temporal del vector de movimiento se usa o no se usa (es decir, la determinación del valor del indicador 122) se describirá más adelante según varias realizaciones de la presente invención. En función de esta decisión, la unidad 320 de predicción inter realiza la predicción inter con o sin el uso del predictor temporal de vector de movimiento. En particular, la unidad 330 de predicción del vector de movimiento está configurada para recibir el indicador 122 y si el indicador tiene un primer valor (por ejemplo, "1"), la unidad 330 de predicción del vector de movimiento es operable para crear la primera lista de predictores de vector de movimiento que comprende una pluralidad de predictores de vector de movimiento que incluyen al menos un predictor temporal de vector de movimiento derivado de al menos un vector de movimiento de una imagen de referencia co-localizada, y selecciona un predictor de vector de movimiento de la primera lista para una unidad de predicción en la unidad de sub-imagen. La unidad 328 de escritura también es operable para escribir el primer parámetro en la secuencia de bits de vídeo codificado para indicar el predictor 331 de vector de movimiento seleccionado de la primera lista. Por otro lado, si el indicador 122 tiene un segundo valor (por ejemplo, "0"), la unidad 330 de predicción del vector de movimiento es operable para crear la segunda lista de predictores de vectores de movimiento que comprenden una pluralidad de predictores de vectores de movimiento sin ningún predictor temporal de vector de movimiento, y selecciona un predictor de vector de movimiento de la segunda lista para una unidad de predicción en la unidad de sub-imagen. En este caso, la unidad 328 de escritura también es operable para escribir el segundo parámetro en la secuencia 304 de bits de vídeo codificado para indicar el predictor 331 de vector de movimiento seleccionado de la segunda lista. La unidad 328 de escritura también es operable para escribir los datos 326 representativos del indicador 122 que tiene bien un primer o bien un segundo valor (por ejemplo, "0" o "1") que indica si la predicción temporal del vector de movimiento se usa o no se usa en la secuencia 304 de bits de vídeo codificado (por ejemplo, una cabecera de la unidad de sub-imagen o una cabecera de la imagen).

La figura 4 representa un diagrama de flujo que ilustra un procedimiento 400 de descodificación de un vídeo codificado según una realización de la presente invención. En particular, el procedimiento 400 es operable para descodificar una secuencia 100 de bits de vídeo codificado según el procedimiento descrito anteriormente para codificar un vídeo como se muestra en la figura 2. En el paso S402, uno o más parámetros (es decir, parámetros de orden de listas de imágenes de referencia) se analizan desde una cabecera 116 de un corte 114 para especificar el orden de las imágenes de referencia en una o más listas de imágenes de referencia utilizadas para la predicción inter del corte 114. Como se mencionó anteriormente, una localización predeterminada (tal como la primera imagen) en una de las listas de imágenes de referencia (tal como la lista 0 de imágenes de referencia) indica la imagen de referencia co-localizada. En el paso S404, un indicador 122 (es decir, el indicador de predicción temporal del vector de movimiento) se analiza desde la cabecera 116 que indica si la predicción temporal del vector de movimiento se usa o no se usa para la predicción inter del corte 118. En el paso S406, el valor del indicador 122 se analiza o juzga para determinar si el indicador 122 indica que la predicción temporal del vector de movimiento se usa o no se usa.

Si el indicador 122 indica que se usa la predicción temporal del vector de movimiento, en el paso S408, se crea una lista de predictores de vector de movimiento (por ejemplo, una primera lista) que comprende una pluralidad de predictores de vectores de movimiento que incluyen al menos un predictor temporal de vector de movimiento derivado de al menos un vector de movimiento de la imagen de referencia co-localizada. Solo a modo de ejemplo, la pluralidad de vectores de movimiento puede incluir al menos un predictor temporal de vector de movimiento, uno o más vectores de movimiento derivados de unidades / bloques de predicción espacial vecinos (es decir, predictores espaciales de vector de movimiento) y un vector de movimiento cero. En el paso S410, se analiza un parámetro 126 (es decir, un parámetro de selección de predictor de vector de movimiento) (por ejemplo, un primer parámetro) de la secuencia 100 de bits de vídeo codificado (es decir, de una unidad 120 de predicción del corte 114) que indica un predictor de vector de movimiento seleccionado de entre la lista de predictores de vector de movimiento para un bloque objetivo de muestras (es decir, una unidad 120 de predicción) en el corte 114.

Por otro lado, si el indicador 122 indica que no se utiliza la predicción temporal del vector de movimiento, en el paso S412, se crea una lista de predictores de vector de movimiento (por ejemplo, una segunda lista) que comprende una pluralidad de predictores de vector de movimiento sin ningún predictor de vector de movimiento temporal. En el paso S414, se analiza un parámetro (es decir, un parámetro de selección de predictor de vector de movimiento) (por ejemplo, un segundo parámetro) de la secuencia 100 de bits de vídeo codificado (es decir, de una unidad 120 de predicción del corte 114) que indica un predictor de vector de movimiento seleccionado de entre la lista de predictores de vector de movimiento para un bloque objetivo de muestras (es decir, una unidad 120 de predicción) en el corte 114.

Después del paso S410 o del paso S414, se realiza una predicción inter compensada por movimiento utilizando el predictor de vector de movimiento seleccionado para producir un bloque de muestras de predicción en el paso S416. Posteriormente, en el paso S418, se descodifica un bloque de muestras residuales de la secuencia 100 de bits de

vídeo codificado. Posteriormente, en el paso S420, el bloque de muestras de predicción y el bloque de muestras residuales se suman para producir un bloque de muestras reconstruidas correspondientes al bloque objetivo.

A continuación se describirá un aparato 500 de ejemplo para descodificar un vídeo codificado según una realización de la presente invención.

5 La figura 5 representa un diagrama de bloques esquemático del aparato 500 de ejemplo para descodificar una secuencia 502 de bits codificada de entrada bloque a bloque y enviar vídeos/imágenes 504, por ejemplo, a una pantalla. El aparato 500 comprende una unidad 506 de descodificación de entropía operable para descodificar una secuencia 502 de bits codificada de entrada, una unidad 508 de cuantificación inversa operable para realizar una cuantificación inversa sobre unos datos de entrada, una unidad 510 de transformación inversa operable para realizar una transformación de frecuencia inversa sobre unos datos de entrada, una memoria 512 de bloque y una 514 memoria de imagen operable para almacenar datos tales como vídeos e imágenes, una unidad 516 de predicción intra para realizar la predicción intra, una unidad 518 de predicción inter para realizar la predicción inter, una unidad 522 de predicción de vector de movimiento y una unidad 503 de análisis operable para analizar el secuencia 502 de bits codificada de entrada y emitir varios parámetros 520, 521.

15 Para mayor claridad, ahora se describirá un flujo de datos de ejemplo a través del aparato 500 como se muestra en la figura 5. Se ingresa una secuencia 502 de bits codificada de entrada a la unidad 506 de descodificación de entropía. Después de que la secuencia 502 de bits codificada de entrada se ingresa a la unidad 506 de descodificación de entropía, la unidad 506 de descodificación de entropía descodifica la secuencia 502 de bits codificada de entrada, y envía los valores descodificados 507 a la unidad 508 de cuantificación inversa. La unidad 508 de cuantificación inversa cuantifica de manera inversa los valores 507 descodificados, y envía los coeficientes 509 de frecuencia a la unidad 510 de transformación inversa. La unidad 510 de transformación inversa realiza la transformación de frecuencia inversa sobre los coeficientes 509 de frecuencia para transformar los coeficientes 509 de frecuencia en valores 511 de muestra y envía los valores 511 de muestra resultantes a un sumador. El sumador agrega los valores 511 de muestra resultantes a la salida de los valores 519 de vídeo/imagen predichos que salen de la unidad 516 o 518 de predicción intra o inter, y envía los valores 504 resultantes a, por ejemplo, una pantalla, y a la memoria 512 de bloque o a la memoria 514 de imagen para predicción adicional. Además, la unidad 516 o 518 de predicción intra o inter busca dentro de los vídeos/imágenes almacenados en la memoria 512 de bloque o en la memoria 514 de imágenes, y estima un área de vídeo/imagen que es, por ejemplo, lo más similar a los vídeos/imágenes descodificados para la predicción.

30 Además, la unidad 506 de análisis analiza un indicador 122 de una cabecera del corte o de la imagen que indica si la predicción temporal del vector de movimiento se usa o no se usa para la predicción inter de un corte objetivo, y envía los datos 520 analizados a la unidad 522 de predicción de vector de movimiento. La unidad 518 de predicción inter es operable para realizar predicción inter con o sin usar el predictor temporal de vector de movimiento basado en el valor del indicador 122 y el predictor de vector de movimiento seleccionado de la unidad 522 de predicción de vector de movimiento. En particular, la unidad 522 de predicción de vector de movimiento está configurada para recibir los datos 520 que contienen el indicador 122 y si el indicador tiene un primer valor (por ejemplo, "1"), la unidad 522 de predicción de vector de movimiento es operable para crear la primera lista de predictores de vector de movimiento que comprende una pluralidad de predictores de vector de movimiento que incluyen al menos un predictor de vector de movimiento temporal derivado de al menos un vector de movimiento de una imagen de referencia co-localizada. Si el indicador tiene un segundo valor (por ejemplo, "0"), la unidad 522 de vector de movimiento es operable para crear una segunda lista de predictores de vector de movimiento que comprenden una pluralidad de predictores de vector de movimiento sin ningún predictor temporal de vector de movimiento. La unidad 503 de análisis es además operable para analizar el primer o el segundo parámetro de la secuencia 502 de bits de vídeo codificado que indica un predictor de vector de movimiento seleccionado de la segunda lista para una unidad de predicción en la unidad de sub-imagen, y envía los datos 521 analizados a la unidad 522 de predicción de vectores de movimiento.

Como se mencionó anteriormente, varias técnicas para decidir/determinar si se usa o no se usa la predicción temporal de vector de movimiento (es decir, determinar el valor del indicador 122) ahora se describirán según varias realizaciones de la presente invención.

50 Según una primera realización, el valor del indicador 122 se determina en función de una capa temporal de la imagen actual. La figura 6 representa un diagrama que muestra las diferentes capas temporales para un grupo de imágenes cuando el tamaño/estructura del grupo se configura a 4, por ejemplo. En el ejemplo, hay tres capas temporales, a saber, la capa 602 temporal "0", la capa 604 temporal "1" y la capa 606 temporal "2". Las imágenes con valores de conteo de orden de imágenes (POC) de 0, 4 y 8 están localizadas en la capa 602 temporal "0", las imágenes con valores POC de 2 y 6 están localizadas en la capa 604 temporal "1", y las imágenes con valores POC de 1, 3, 5 y 7 están localizadas en la capa 606 temporal "2". Las capas temporales "0", "1" y "2" están asociadas respectivamente con o representadas por la identificación temporal ID 0, 1 y 2. En consecuencia, las imágenes en la capa 602 temporal "0" tienen asociada la identificación temporal ID "0", las imágenes en la capa 604 temporal "1" tienen asociada la ID temporal 1, y las imágenes en la capa 606 temporal "2" tienen asociada la ID temporal 2.

La figura 7 representa un diagrama de flujo que ilustra un procedimiento 700 para determinar el valor del indicador 122 según la primera realización. En el paso S702, la capa temporal de la imagen actual se determina en función de la ID temporal asociada a la imagen actual. Posteriormente, en el paso S704, se analiza o se juzga si la capa temporal determinada es la capa más baja o la capa base (es decir, si la ID temporal = 0). Si la capa temporal es la capa más baja, en el paso S706, el indicador 122 se establece en un valor (por ejemplo, "0") que indica que no se usa la predicción temporal del vector de movimiento. Por otro lado, si la capa temporal no es la capa más baja, en el paso S708, el indicador 122 se establece en un valor (por ejemplo, "1") que indica que se usa la predicción temporal del vector de movimiento. Esto se debe a que en una estructura de codificación típica, las imágenes con una ID temporal = 0 son referenciadas a menudo por las imágenes con una ID temporal mayor. En el caso de que la imagen con ID temporal = 0 se pierda o contenga un error, el error se propagará a cualquier imagen que haga referencia a esa imagen con ID temporal = 0. Esta propagación de error puede continuar y efectuar la reconstrucción de todas las imágenes subsiguientes que usan la imagen de vector de movimiento temporal con ID temporal = 0. Por lo tanto, esta realización mejora la resiliencia al error al no usar imagen de vector de movimiento temporal con ID temporal = 0.

Según una segunda realización, el valor del indicador 122 se determina en función del valor POC de la imagen actual. La figura 8 representa un diagrama de flujo que ilustra un procedimiento 800 para determinar el valor del indicador 122 según la segunda realización. En el paso S802, se obtienen o determinan un valor POC para la imagen actual y los valores POC de todas las imágenes de referencia en el DPB. En el paso S804, se analiza o se juzga si el valor POC de la imagen actual es mayor que cualquiera de los valores POC de la imagen de referencia en el DPB. Si es así, en el paso S806, el indicador 122 se establece en un valor (por ejemplo, "0") que indica que no se usa la predicción temporal de vector de movimiento. De lo contrario, en el paso S808, el indicador 122 se establece en un valor (por ejemplo, "1") que indica que se usa la predicción temporal de vector de movimiento. Esto se debe a que una imagen de mayor calidad (por ejemplo, imágenes de la capa temporal 0) solo hace referencia a imágenes de la misma calidad o superior. En esta realización, se identifica una imagen de mayor calidad en vista de los valores POC de las imágenes de referencia contenidas en la memoria intermedia de imágenes descodificadas que almacena una pluralidad de imágenes de referencia. Por una razón similar a la primera realización descrita anteriormente, las imágenes de mayor calidad a menudo son referenciadas por las imágenes subsiguientes. Por consiguiente, para prevenir o minimizar la propagación de errores y mejorar la resiliencia a los errores, el indicador 122 se inhabilita en imágenes de mayor calidad.

Según una tercera realización, el valor del indicador 122 se determina en función del tipo de corte de un corte inter en la imagen actual. Un corte inter es un corte que se codifica o descodifica utilizando la predicción inter. La figura 9 representa un diagrama de flujo que ilustra un procedimiento 900 para determinar el valor del indicador 122 según la tercera realización. En el paso 902, se determina un tipo de corte de un corte inter de la imagen actual. Posteriormente, se analiza o se juzga si el tipo de corte es un corte P (es decir, un corte predictiva). Si es así, en el paso S906, el indicador 122 se establece en un valor (por ejemplo, "0") que indica que no se usa la predicción temporal de vector de movimiento. Por otro lado, si el tipo de corte determinado no es un corte P (por ejemplo, es un corte B o bi-predictiva), en el paso S908, el indicador 122 se establece en un valor (por ejemplo, "1") que indica que se utiliza la predicción temporal de vector de movimiento. Una razón para esto es porque el corte P utiliza una predicción unidireccional hacia delante. Por lo tanto, para evitar o minimizar la propagación de errores y para mejorar la resiliencia a los errores, el indicador 122 se inhabilita en el corte P.

Según una cuarta realización, el valor del indicador 122 se determina en función de si la imagen es una imagen de punto de acceso aleatorio (RAP). Una imagen RAP es una imagen que, ella misma y todas las imágenes subsiguientes en el orden de descodificación, puede descodificarse correctamente sin tener que realizar el proceso de descodificación de cualquier imagen que preceda a la imagen RAP en el orden de descodificación. Por ejemplo, la especificación HEVC especifica una imagen RAP como una imagen codificada para la cual cada segmento de corte tiene un tipo de unidad NAL (es decir, `nal_unit_type`) en el intervalo de 7 a 12, inclusive. La figura 10 representa una representación del diagrama de un flujo de unidad NAL, es decir, una serie de unidades 102 NAL para una secuencia de bits de vídeo codificado. Como un experto en la técnica sabe, la NAL (capa de abstracción de red) formatea la representación de la capa de codificación de vídeo (VCL) de un vídeo codificado y proporciona información de cabecera de manera adecuada para su transmisión por una variedad de capas de transporte o medios de almacenamiento. Cada unidad 102 NAL comprende una cabecera 104 seguida de una sección 106 de datos. La cabecera 104 incluye un parámetro que indica el tipo de datos en la unidad 102 NAL y la sección 106 de datos contiene los datos indicados por la cabecera 104. Por ejemplo, la figura 10 muestra tres unidades NAL, una primera unidad NAL que contiene un conjunto de parámetros (según lo indicado por el tipo 108 de unidad NAL), una segunda unidad NAL que contiene una vista/capa base (según lo indicado por el tipo 110 de unidad NAL) y una tercera unidad NAL que contiene una vista/capa no base (como lo indica el tipo 112 de unidad NAL). La cabecera 104 de cada unidad NAL comprende además una ID temporal como se describió en la primera realización que se muestra en la figura 7.

La figura 11 representa una representación del diagrama de una imagen 1100 RAP de ejemplo que contiene varias vistas/capas con múltiples cortes. Como se muestra, la imagen 1100 RAP comprende varios cortes 1102 en la capa 1104 base (vista intra) y varios cortes 1106 en la capa 1110 no base (vista inter).

La figura 12 representa un diagrama de flujo que ilustra un procedimiento 1200 para determinar el valor del indicador 122 según la cuarta realización. En el paso S1202, la imagen se analiza para determinar u obtener un parámetro de cada corte de la imagen que especifica el tipo de unidad NAL de los cortes. Posteriormente, en el paso S1204, se determina o se juzga si la imagen que contiene el corte actual es una imagen RAP basada en el (los) parámetro(s) obtenido(s) y si el corte actual pertenece a una capa/vista no base de la imagen. Si la imagen es una imagen 1100 RAP se puede determinar analizando el valor del tipo 1008, 1010, 1012 de unidad NAL en la cabecera 1004 de cada unidad NAL o corte 1002 de la imagen. Como se mencionó anteriormente, una imagen 1100 RAP es una imagen que, ella misma y todas las imágenes subsiguientes en el orden de descodificación, pueden descodificarse correctamente sin realizar el proceso de descodificación de ninguna imagen que preceda a la imagen 1100 RAP en el orden de descodificación. Por ejemplo, la especificación HEVC especifica una imagen RAP como una imagen codificada para la cual cada segmento de corte tiene un tipo de unidad NAL en el intervalo de 7 a 12, inclusive. En consecuencia, en este ejemplo, si el tipo 1008, 1010, 1012 de unidad NAL de cada unidad 1002 NAL de la imagen está en el intervalo de 7 a 12, inclusive, se determina que la imagen es una imagen 1100 RAP. Si el corte actual es una capa no base de la imagen se puede determinar al examinar el tipo 1008, 1010, 1012 de unidad NAL del corte actual. Por ejemplo, el tipo 1012 de unidad NAL indica que el corte 1006 asociado pertenece a una capa no base y el tipo 1010 de unidad NAL indica que el corte 1006 asociado pertenece a una capa base. Sin embargo, los expertos en la técnica apreciarán que la capa no base puede identificarse en función de otro(s) parámetro(s) dependiendo del esquema de codificación de vídeo. Por ejemplo, en el borrador de trabajo de HEVC de vista múltiple de HEVC actual, si el corte actual es una capa no base de la imagen lo determina la ID de la capa. Si la imagen es una imagen 1100 RAP y el corte actual pertenece a una capa no base de la imagen, en el paso S1206, el indicador 122 se establece en un valor (por ejemplo, "0") que indica que no se usa la predicción temporal de vector de movimiento. De lo contrario, en el paso S1208, el indicador 122 se establece en un valor (por ejemplo, "1") que indica que se utiliza la predicción temporal de vector de movimiento. Una razón para esto es porque el beneficio de usar la predicción temporal de vector de movimiento es mejorar la predicción de vector de movimiento temporalmente, es decir, la predicción a partir de otras imágenes que son diferentes en el tiempo. Sin embargo, si las imágenes intra e inter se encuentran dentro del mismo tiempo en la imagen actual, no hay ningún beneficio en usar la predicción temporal de vector de movimiento. Por lo tanto, para mejorar la eficiencia de codificación/descodificación, el indicador 122 se inhabilita para los cortes 1106 que pertenecen a la capa no base (o vista inter) de una imagen RAP 1100.

(Realización A)

El procesamiento descrito en cada una de las realizaciones se puede implementar simplemente en un sistema informático independiente, mediante la grabación, en un medio de grabación, de un programa para implementar las configuraciones del procedimiento de codificación de imágenes en movimiento (procedimiento de codificación de imágenes) y el procedimiento de descodificación de imágenes en movimiento (procedimiento de descodificación de imágenes) descrito en cada una de las realizaciones. Los medios de grabación pueden ser cualesquier medios de grabación siempre que se pueda grabar el programa, tales como un disco magnético, un disco óptico, un disco óptico magnético, una tarjeta IC y una memoria de semiconductores.

En lo sucesivo, se describirán las aplicaciones del procedimiento de codificación de imágenes en movimiento (procedimiento de codificación de imágenes) y del procedimiento de descodificación de imágenes en movimiento (procedimiento de descodificación de imágenes) descritas en cada una de las realizaciones y sistemas que las utilizan. El sistema tiene la característica de tener un aparato de codificación y descodificación de imágenes que incluye un aparato de codificación de imágenes que utiliza el procedimiento de codificación de imágenes y un aparato de descodificación de imágenes que utiliza el procedimiento de descodificación de imágenes. Otras configuraciones del sistema se pueden cambiar según corresponda dependiendo de los casos.

La figura 13 ilustra una configuración general de un sistema ex100 que proporciona contenido para implementar servicios de distribución de contenido. El área para proporcionar servicios de comunicación se divide en celdas del tamaño deseado, y las estaciones ex106, ex107, ex108, ex109 y ex110 base que son estaciones inalámbricas fijas se colocan en cada una de las celdas.

El sistema ex100 que proporciona contenido está conectado a dispositivos tales como un ordenador ex111, un asistente personal digital (PDA) ex112, una cámara ex113, un teléfono móvil ex114 y una máquina de juegos ex115, a través de Internet ex101, de un servicio proveedor de Internet ex102, de una red telefónica ex104, así como a las estaciones ex106 a ex110 base, respectivamente.

Sin embargo, la configuración del sistema ex100 que proporciona el contenido no se limita a la configuración mostrada en la figura 13, y una combinación en la cual cualquiera de los elementos estén conectados es aceptable. Además, cada dispositivo puede estar conectado directamente a la red telefónica ex104, en lugar de a través de las estaciones ex106 a ex110 base, que son las estaciones inalámbricas fijas. Además, los dispositivos pueden estar interconectados entre sí a través de una comunicación inalámbrica de corta distancia y otros.

La cámara ex113, tal como una cámara de vídeo digital, es capaz de capturar vídeo. Una cámara ex116, tal como una cámara digital, es capaz de capturar imágenes fijas y vídeo. Además, el teléfono móvil ex114 puede ser el que cumpla con cualquiera de los estándares, tales como el sistema global para comunicaciones móviles (GSM) (marca registrada), acceso múltiple por división de código (CDMA), acceso múltiple por división de código de banda ancha

(W-CDMA) , evolución a largo plazo (LTE) y acceso a paquetes de alta velocidad (HSPA). Alternativamente, el teléfono móvil ex114 puede ser un sistema de teléfono personal (PHS).

En el sistema ex100 que proporciona el contenido, un servidor ex103 de retransmisión está conectado a la cámara ex113 y a otras a través de la red telefónica ex104 y a la estación ex109 base, lo que permite la distribución de imágenes de un programa en directo y otras. En dicha distribución, el contenido (por ejemplo, el vídeo de un programa de música en directo) capturado por el usuario que utiliza la cámara ex113 se codifica como se describió anteriormente en cada una de las realizaciones (es decir, la cámara funciona como el aparato de codificación de imágenes según un aspecto de la presente invención), y el contenido codificado se transmite al servidor ex103 de retransmisión. Por otro lado, el servidor ex103 de retransmisión lleva a cabo la distribución en secuencia de los datos de contenido transmitidos a los clientes al solicitarla. Los clientes incluyen el ordenador ex111, la PDA ex112, la cámara ex113, el teléfono móvil ex114 y la máquina de juegos ex115 que son capaces de descodificar los datos codificados mencionados anteriormente. Cada uno de los dispositivos que han recibido los datos distribuidos descodifica y reproduce los datos codificados (es decir, funciona como el aparato de descodificación de imágenes según un aspecto de la presente invención).

Los datos capturados pueden ser codificados por la cámara ex113 o el servidor ex103 de retransmisión que transmite los datos, o los procesos de codificación pueden ser compartidos entre la cámara ex113 y el servidor ex103 de retransmisión. De forma similar, los datos distribuidos pueden ser descodificados por los clientes o por el servidor ex103 de retransmisión, o los procesos de descodificación pueden ser compartidos entre los clientes y el servidor ex103 de retransmisión. Además, los datos de las imágenes fijas y de vídeo capturados no solo por la cámara ex113, sino también por la cámara ex116 pueden transmitirse al servidor ex103 de retransmisión a través del ordenador ex111. Los procesos de codificación pueden ser realizados por la cámara ex116, el ordenador ex111 o el servidor ex103 de retransmisión, o pueden repartirse entre ellos.

Además, los procesos de codificación y descodificación pueden realizarse mediante un LSI ex500 generalmente incluido en cada uno de los ordenadores ex111 y en los dispositivos. El LSI ex500 puede estar configurado por un solo chip o por una pluralidad de chips. El software para codificar y descodificar vídeo puede integrarse en algún tipo de medio de grabación (tal como un CD-ROM, un disco flexible y un disco duro) que sea legible por el ordenador ex111 y otros, y los procesos de codificación y descodificación pueden ser realizados utilizando el software. Además, cuando el teléfono móvil ex114 está equipado con una cámara, los datos de vídeo obtenidos por la cámara pueden ser transmitidos. Los datos de vídeo son datos codificados por el LSI ex500 incluido en el teléfono móvil ex114.

Además, el servidor ex103 de retransmisión puede estar compuesto por servidores y ordenadores, y puede descentralizar datos y procesar los datos descentralizados, registrar o distribuir datos.

Como se describió anteriormente, los clientes pueden recibir y reproducir los datos codificados en el sistema ex100 que proporciona el contenido. En otras palabras, los clientes pueden recibir y descodificar la información transmitida por el usuario y reproducir los datos descodificados en tiempo real en el sistema ex100 que proporciona el contenido, de modo que el usuario que no tiene ningún derecho en particular y el equipo puede implementar la difusión personal.

Aparte del ejemplo del sistema ex100 que proporciona el contenido, al menos uno de los aparatos de codificación de imágenes en movimiento (aparatos de codificación de imágenes) y el aparato de descodificación de imágenes en movimiento (aparatos de descodificación de imágenes) descritos en cada una de las realizaciones pueden implementarse en el sistema ex200 digital de radiodifusión ilustrado en la figura 14. Más específicamente, una estación ex201 de difusión comunica o transmite, a través de ondas de radio a un satélite ex202 de difusión, datos multiplexados obtenidos multiplexando datos de audio y otros sobre los datos de vídeo. Los datos de vídeo son datos codificados por el procedimiento de codificación de imágenes en movimiento descrito en cada una de las realizaciones (es decir, datos codificados por el aparato de codificación de imágenes según un aspecto de la presente invención). Al recibir los datos multiplexados, el satélite ex202 de difusión transmite ondas de radio para su difusión. Luego, una antena ex204 de uso doméstico con una función de recepción de difusión por satélite recibe las ondas de radio. A continuación, un dispositivo como un televisor ex300 (receptor) y un dispositivo de recepción y descodificación ex217 (STB) de señal de televisión descodifica los datos multiplexados recibidos y reproduce los datos descodificados (es decir, funciona como el aparato de descodificación de imágenes según un aspecto de la presente invención).

Además, un lector/grabador ex218 (i) lee y descodifica los datos multiplexados grabados en un medio de grabación ex215, como un DVD y un BD, o (i) codifica señales de vídeo en el medio de grabación ex215, y en algunos casos, escribe los datos obtenidos multiplexando una señal de audio sobre los datos codificados. El lector/grabador ex218 puede incluir el aparato de descodificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento como se muestra en cada una de las realizaciones. En este caso, las señales de vídeo reproducidas se muestran en el monitor ex219, y pueden ser reproducidas por otro dispositivo o sistema utilizando el medio de grabación ex215 en el que se graban los datos multiplexados. También es posible implementar el aparato de descodificación de imágenes en movimiento en el dispositivo de recepción y descodificación ex217 de señal de televisión conectado al cable ex203 para una televisión por cable o a la antena ex204 para la difusión por satélite y/o

terrestre, a fin de mostrar las señales de vídeo en el monitor ex219 de la televisión ex300. El aparato de descodificación de imágenes en movimiento puede implementarse no en el descodificador sino en el televisor ex300.

La figura 15 ilustra la televisión ex300 (receptor) que utiliza el procedimiento de codificación de imágenes en movimiento y el procedimiento de descodificación de imágenes en movimiento descrito en cada una de las realizaciones. El televisor ex300 incluye: un sintonizador ex301 que obtiene o proporciona datos multiplexados obtenidos multiplexando datos de audio sobre datos de vídeo, a través de la antena ex204 o del cable ex203, etc. que recibe una transmisión; una unidad ex302 de modulación/demodulación que demodula los datos multiplexados recibidos o modula los datos en datos multiplexados para ser suministrados al exterior; y una unidad ex303 de multiplexación/demultiplexación que demultiplexa los datos multiplexados modulados en datos de vídeo y datos de audio, o multiplexa datos de vídeo y datos de audio codificados por una unidad ex306 de procesamiento de señales en datos.

El televisor ex300 incluye además: una unidad ex306 de procesamiento de señales que incluye una unidad ex304 de procesamiento de señales de audio y una unidad ex305 de procesamiento de señales de vídeo que descodifican datos de audio y de vídeo y codifican datos de audio y de vídeo, respectivamente (que funcionan como el aparato de codificación de imágenes y el aparato de descodificación de imágenes según los aspectos de la presente invención); y una unidad ex309 de salida que incluye un altavoz ex307 que proporciona la señal de audio descodificada, y una unidad ex308 de pantalla que muestra la señal de vídeo descodificada, tal como una pantalla. Además, el televisor ex300 incluye una unidad ex317 de interfaz que incluye una unidad ex312 de entrada de operación que recibe una entrada de una operación de usuario. Además, el televisor ex300 incluye una unidad de control ex310 que controla en general cada elemento constitutivo del televisor ex300, y una unidad ex311 de circuito de fuente de alimentación que suministra energía a cada uno de los elementos. Aparte de la unidad ex312 de entrada de operación, la unidad ex317 de interfaz puede incluir: un puente ex313 que está conectado a un dispositivo externo, tal como el lector/grabador ex218; una unidad ex314 de ranura para permitir la conexión del medio ex216 de grabación, tal como una tarjeta SD; un controlador ex315 para conectarse a un medio de grabación externo, tal como un disco duro; y un módem ex316 para ser conectado a una red telefónica. Aquí, el medio ex216 de grabación puede registrar eléctricamente información utilizando un elemento semiconductor de memoria no volátil/volátil para almacenamiento. Los elementos constitutivos de la televisión ex300 están conectados entre sí a través de un bus síncrono.

En primer lugar, se describirá la configuración en la que la televisión ex300 descodifica los datos multiplexados obtenidos desde el exterior a través de la antena ex204 y demás y reproduce los datos descodificados. En la televisión ex300, tras una operación del usuario a través de un controlador remoto ex220 y demás, la unidad ex303 de multiplexación/demultiplexación demultiplexa los datos multiplexados demodulados por la unidad ex302 de modulación / demodulación, bajo el control de la unidad ex310 de control, incluida una CPU. Además, la unidad ex304 de procesamiento de señales de audio descodifica los datos de audio demultiplexados, y la unidad ex305 de procesamiento de señales de vídeo descodifica los datos de vídeo demultiplexados, utilizando el procedimiento de descodificación descrito en cada una de las realizaciones, en el televisor ex300. La unidad ex309 de salida proporciona la señal de vídeo descodificada y la señal de audio al exterior, respectivamente. Cuando la unidad ex309 de salida proporciona la señal de vídeo y la señal de audio, las señales pueden almacenarse temporalmente en las memorias intermedias ex318 y ex319, y demás para que las señales se reproduzcan con sincronización entre sí. Además, la televisión ex300 puede leer datos multiplexados no a través de una transmisión y demás, sino de los medios de grabación ex215 y ex216, tales como un disco magnético, un disco óptico y una tarjeta SD. A continuación, se describirá una configuración en la que el televisor ex300 codifica una señal de audio y una señal de vídeo, y transmite los datos al exterior o escribe los datos en un medio de grabación. En la televisión ex300, tras una operación del usuario a través del controlador remoto ex220 y demás, la unidad ex304 de procesamiento de señales de audio codifica una señal de audio, y la unidad ex305 de procesamiento de señales de vídeo codifica una señal de vídeo, bajo el control de la unidad ex310 de control utilizando el procedimiento de codificación descrito en cada una de las realizaciones. La unidad ex303 de multiplexación/demultiplexación multiplexa la señal de vídeo codificada y la señal de audio, y proporciona la señal resultante al exterior. Cuando la unidad ex303 de multiplexación/demultiplexación multiplexa la señal de vídeo y la señal de audio, las señales pueden almacenarse temporalmente en las memorias intermedias ex320 y ex321, y demás para que las señales se reproduzcan con sincronización entre sí. Aquí, las memorias intermedias ex318, ex319, ex320 y ex321 pueden ser múltiples como se ilustra, o al menos una memoria intermedia puede ser compartida en la televisión ex300. Además, los datos se pueden almacenar en una memoria intermedia para evitar el desbordamiento y el subdesbordamiento del sistema entre la unidad ex302 de modulación/demodulación y la unidad ex303 de multiplexación/demultiplexación, por ejemplo.

Además, el televisor ex300 puede incluir una configuración para recibir una entrada AV desde un micrófono o una cámara distinta de la configuración para obtener datos de audio y vídeo desde una transmisión o desde un medio de grabación, y puede codificar los datos obtenidos. Aunque el televisor ex300 puede codificar, multiplexar y proporcionar datos externos en la descripción, puede ser capaz de solo recibir, descodificar y proporcionar datos externos, pero no de codificar, multiplexar y proporcionar datos externos.

Además, cuando el lector/grabador ex218 lee o escribe datos multiplexados desde o sobre un medio de grabación, uno de los televisores ex300 y el lector/grabador ex218 puede descodificar o codificar los datos multiplexados, y el televisor ex300 y el lector/registrator ex218 puede compartir la descodificación o la codificación.

5 Como ejemplo, la figura 16 ilustra una configuración de una unidad ex400 de reproducción/grabación de información cuando los datos se leen o se escriben desde o en un disco óptico. La unidad ex400 de reproducción / grabación de información incluye los elementos constituyentes ex401, ex402, ex403, ex404, ex405, ex406 y ex407 que se describirán a continuación. El cabezal óptico ex401 irradia un punto láser en una superficie de grabación del medio de grabación ex215 que es un disco óptico para escribir información, y detecta la luz reflejada desde la superficie de grabación del medio de grabación ex215 para leer la información. La unidad ex402 de registro de modulación controla eléctricamente un láser semiconductor incluido en el cabezal óptico ex401, y modula la luz del láser según los datos registrados. La unidad ex403 de demodulación de reproducción amplifica una señal de reproducción obtenida detectando eléctricamente la luz reflejada desde la superficie de grabación utilizando un fotodetector incluido en el cabezal óptico ex401, y demodula la señal de reproducción al separar un componente de señal grabado en el medio de grabación ex215 para reproducir la información necesaria. La memoria intermedia ex404 retiene temporalmente la información a grabar en el medio de grabación ex215 y la información reproducida desde el medio de grabación ex215. El motor de disco ex405 hace girar el medio de grabación ex215. La unidad ex406 de control de servo mueve el cabezal óptico ex401 a una pista de información predeterminada mientras controla el impulso de rotación del motor de disco ex405 a fin de seguir el punto del láser. La unidad ex407 de control del sistema controla en general la unidad ex400 de reproducción/grabación de información. Los procesos de lectura y escritura pueden ser implementados por la unidad ex407 de control del sistema utilizando información diversa almacenada en la memoria intermedia ex404 y generando y agregando nueva información según sea necesario, y por la unidad ex402 de registro de modulación, la unidad ex403 de demodulación de reproducción y la unidad ex406 de control de servo que registra y reproduce información a través del cabezal óptico ex401 mientras se opera de manera coordinada. La unidad ex407 de control del sistema incluye, por ejemplo, un microprocesador, y ejecuta el procesamiento al hacer que un ordenador ejecute un programa para leer y escribir.

Aunque el cabezal óptico ex401 irradia un punto de láser en la descripción, puede realizar grabaciones de alta densidad utilizando luz de campo cercano.

La figura 17 ilustra el medio de grabación ex215 que es el disco óptico. Sobre la superficie de grabación del medio de grabación ex215, los surcos de guía están formados en espiral, y una pista ex230 de información registra, por adelantado, información de dirección que indica una posición absoluta en el disco según el cambio en la forma de los surcos guía. La información de la dirección incluye información para determinar las posiciones de los bloques de grabación ex231 que son una unidad para registrar datos. La reproducción de la pista ex230 de información y la lectura de la información de dirección en un aparato que registra y reproduce datos puede llevar a la determinación de las posiciones de los bloques de grabación. Además, el medio de grabación ex215 incluye un área ex233 de grabación de datos, un área ex232 de circunferencia interior y un área ex234 de circunferencia exterior. El área ex233 de registro de datos es un área que se usa para registrar los datos del usuario. El área ex232 de la circunferencia interior y el área ex234 de la circunferencia exterior que están dentro y fuera del área ex233 de grabación de datos, respectivamente, son para uso específico, excepto para registrar los datos del usuario. La unidad 400 de reproducción/grabación de información lee y escribe audio codificado, datos de vídeo codificados o datos multiplexados obtenidos multiplexando los datos codificados de audio y vídeo, desde y sobre el área ex233 de grabación de datos del medio de grabación ex215.

Aunque un disco óptico que tiene una capa, tal como un DVD y un BD, se describe como un ejemplo en la descripción, el disco óptico no se limita a los mismos, y puede ser un disco óptico que tiene una estructura multicapa y es capaz de ser grabado en una parte que no sea la superficie. Además, el disco óptico puede tener una estructura para la grabación/reproducción multidimensional, tal como la grabación de información usando luz de colores con diferentes longitudes de onda en la misma porción del disco óptico y para grabar información que tenga diferentes capas desde varios ángulos.

Además, un automóvil ex210 que tiene una antena ex205 puede recibir datos desde el satélite ex202 y demás, y reproducir vídeo en un dispositivo de visualización tal como un sistema ex211 de navegación de automóviles instalado en el automóvil ex210, en el sistema ex200 de transmisión digital. Aquí, una configuración del sistema ex211 de navegación para automóviles será una configuración, por ejemplo, que incluye una unidad de recepción de GPS de la configuración ilustrada en la figura 15. Lo mismo ocurrirá con la configuración del ordenador ex111, del teléfono móvil ex114 y demás.

La figura 18A ilustra el teléfono móvil ex114 que utiliza el procedimiento de codificación de imágenes en movimiento y el procedimiento de descodificación de imágenes en movimiento descrito en las realizaciones. El teléfono móvil ex114 incluye: una antena ex350 para transmitir y recibir ondas de radio a través de la estación ex110 base; una unidad ex365 de cámara capaz de capturar imágenes fijas y en movimiento; y una unidad ex358 de visualización tal como una pantalla de cristal líquido para mostrar los datos tales como el vídeo descodificado capturado por la unidad ex365 de cámara o recibido por la antena ex350. El teléfono móvil ex114 incluye además: una unidad del cuerpo principal que incluye una unidad ex366 de tecla de operación; una unidad ex357 de salida de audio, tal como un altavoz para la salida de audio; una unidad ex356 de entrada de audio, tal como un micrófono para la entrada de

audio; una unidad ex367 de memoria para almacenar vídeo capturado o imágenes fijas, audio grabado, datos codificados o decodificados del vídeo recibido, imágenes fijas, correos electrónicos o demás; y una unidad ex364 de ranura que es una unidad de interfaz para un medio de grabación que almacena datos de la misma manera que la unidad ex367 de memoria.

5 A continuación, se describirá un ejemplo de una configuración del teléfono móvil ex114 con referencia a la figura 18B. En el teléfono móvil ex114, una unidad ex360 de control principal diseñada para controlar en conjunto cada unidad del cuerpo principal, incluida la unidad ex358 de pantalla, así como la unidad ex366 de tecla de operación, se conecta mutuamente, a través de un bus ex370 síncrono, a una unidad ex361 de circuito de alimentación, a una  
10 unidad ex362 de control de entrada de operación, a una unidad ex355 de procesamiento de señal de vídeo, a una unidad ex363 de interfaz de cámara, a una unidad ex359 de control de pantalla de cristal líquido (LCD), a una unidad ex352 de modulación/demodulación, a una unidad ex353 de multiplexación/demultiplexación, a una unidad ex354 de procesamiento de señal de audio, a la unidad ex364 de ranura y a la unidad ex367 de memoria.

15 Cuando una tecla de fin de llamada o una tecla de encendido se activa mediante la operación de un usuario, la unidad ex361 del circuito de fuente de alimentación suministra energía a las unidades respectivas desde una batería a fin de activar el teléfono móvil ex114.

20 En el teléfono móvil ex114, la unidad ex354 de procesamiento de señal de audio convierte las señales de audio recogidas por la unidad ex356 de entrada de audio en el modo de conversación de voz en señales de audio digital bajo el control de la unidad ex360 de control principal que incluye una CPU, ROM y RAM. Luego, la unidad ex352 de modulación/demodulación realiza el procesamiento de espectro expandido sobre las señales de audio digital, y la  
25 unidad ex351 transmisora y receptora realiza la conversión de digital a analógica y la conversión de frecuencia sobre los datos, a fin de transmitir los datos resultantes a través de la antena ex350. Además, en el teléfono móvil ex114, la unidad ex351 transmisora y receptora amplifica los datos recibidos por la antena ex350 en el modo de conversación de voz y realiza la conversión de frecuencia y la conversión analógica a digital sobre los datos. Luego, la unidad ex352 de modulación/demodulación realiza un procesamiento de espectro de dispersión inverso sobre los  
30 datos, y la unidad ex354 de procesamiento de señal de audio lo convierte en señales de audio analógicas, a fin de emitirlos a través de la unidad ex357 de salida de audio.

35 Además, cuando se transmite un correo electrónico en el modo de comunicación de datos, los datos de texto del correo electrónico ingresado al operar la unidad ex366 de tecla de operación y demás del cuerpo principal se envían a la unidad ex360 de control principal a través de la unidad ex362 de control de entrada de operación. La unidad ex360 de control principal hace que la unidad ex352 de modulación/demodulación realice el procesamiento de espectro expandido sobre los datos de texto, y la unidad ex351 transmisora y receptora realiza la conversión de digital a analógica y la conversión de frecuencia sobre los datos resultantes a fin de transmitir los datos a la estación ex110 base a través de la antena ex350. Cuando se recibe un correo electrónico, el procesamiento que es aproximadamente inverso al procesamiento para transmitir un correo electrónico se realiza sobre los datos recibidos, y los datos resultantes se proporcionan a la unidad ex358 de pantalla.

40 Cuando se transmite vídeo, imágenes fijas o vídeo y audio en el modo de comunicación de datos, la unidad ex355 de procesamiento de señal de vídeo comprime y codifica las señales de vídeo suministradas desde la unidad ex365 de cámara usando el procedimiento de codificación de imágenes en movimiento que se muestra en cada una de las realizaciones (es decir, funciona como el aparato de codificación de imágenes según el aspecto de la presente invención) y transmite los datos de vídeo codificados a la unidad ex353 de multiplexación/demultiplexación. En cambio, cuando la unidad ex365 de cámara captura vídeo, imágenes fijas y demás, la unidad ex354 de procesamiento de señales de audio codifica las señales de audio recogidas por la unidad ex356 de entrada de audio y transmite los datos de audio codificados a la unidad ex353 de multiplexación/demultiplexación.

45 La unidad ex353 de multiplexación/demultiplexación multiplexa los datos de vídeo codificados suministrados desde la unidad ex355 de procesamiento de señal de vídeo y los datos de audio codificados suministrados desde la unidad ex354 de procesamiento de señal de audio, utilizando un procedimiento predeterminado. Luego, la unidad ex352 de modulación/demodulación (unidad de circuito de modulación/demodulación) realiza el procesamiento de espectro extendido sobre los datos multiplexados, y la unidad ex351 transmisora y receptora realiza la conversión digital a analógica y la conversión de frecuencia sobre los datos a fin de transmitir los datos resultantes a través de la antena  
50 ex350.

55 Cuando se reciben datos de un archivo de vídeo que está vinculado a una página web y demás en el modo de comunicación de datos o cuando se recibe un correo electrónico con vídeo y/o audio adjunto, para decodificar los datos multiplexados recibidos a través de la antena ex350, la unidad ex353 de multiplexación/demultiplexación demultiplexa los datos multiplexados en una secuencia de bits de datos de vídeo y en una secuencia de bits de datos de audio, y suministra a la unidad ex355 de procesamiento de señal de vídeo los datos de vídeo codificados y a la unidad ex354 de procesamiento de señal de audio los datos de audio codificados, a través del bus ex370 síncrono. La unidad ex355 de procesamiento de señal de vídeo decodifica la señal de vídeo usando un procedimiento de decodificación de imágenes en movimiento correspondiente al procedimiento de codificación de imágenes en movimiento que se muestra en cada una de las realizaciones (es decir, funciona como el aparato de decodificación de imágenes según el aspecto de la presente invención), y luego la unidad ex358 de pantalla  
60 decodifica de imágenes según el aspecto de la presente invención), y luego la unidad ex358 de pantalla

muestra, por ejemplo, el vídeo y las imágenes fijas incluidas en el archivo de vídeo vinculado a la página web a través de la unidad ex359 de control LCD. Además, la unidad ex354 de procesamiento de señal de audio descodifica la señal de audio, y la unidad ex357 de salida de audio proporciona el audio.

5 Además, de manera similar a la televisión ex300, un terminal como el teléfono móvil ex114 probablemente tenga 3 tipos de configuraciones de implementación que incluyen no solo (i) un terminal de transmisión y recepción que incluye tanto un aparato de codificación como un aparato de descodificación, sino también (ii) un terminal de transmisión que incluye solo un aparato de codificación y (iii) un terminal de recepción que incluye solo un aparato de descodificación. Aunque el sistema ex200 de transmisión digital recibe y transmite los datos multiplexados obtenidos multiplexando datos de audio sobre datos de vídeo en la descripción, los datos multiplexados pueden ser  
10 datos obtenidos multiplexando datos no de audio sino datos de caracteres relacionados con el vídeo sobre datos de vídeo, y pueden no ser datos multiplexados sino datos de vídeo en sí.

Como tal, el procedimiento de codificación de imágenes en movimiento y el procedimiento de descodificación de imágenes en movimiento en cada una de las realizaciones se pueden usar en cualquiera de los dispositivos y sistemas descritos. Por lo tanto, se pueden obtener las ventajas descritas en cada una de las realizaciones.

15 Además, la presente invención no está limitada a realizaciones, y son posibles diversas modificaciones y revisiones sin apartarse del ámbito de la presente invención.

(Realización B)

Los datos de vídeo pueden generarse cambiando, según sea necesario, entre (i) el procedimiento de codificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento que se muestra en cada una de  
20 las realizaciones y (ii) un procedimiento de codificación de imágenes en movimiento o un aparato de codificación de imágenes en movimiento en conformidad con un estándar diferente, tal como MPEG-2, MPEG-4 AVC y VC-1.

Aquí, cuando se genera una pluralidad de datos de vídeo que se ajustan a los diferentes estándares y luego se descodifica, los procedimientos de descodificación deben seleccionarse para que se ajusten a los diferentes estándares. Sin embargo, dado que no se puede detectar a qué estándar se ajusta cada uno de la pluralidad de  
25 datos de vídeo a descodificar, existe el problema de que no se puede seleccionar un procedimiento de descodificación apropiado.

Para resolver el problema, los datos multiplexados obtenidos multiplexando datos de audio y otros sobre datos de vídeo tienen una estructura que incluye información de identificación que indica a qué estándar se ajustan los datos de vídeo. A continuación se describirá la estructura específica de los datos multiplexados que incluyen los datos de  
30 vídeo generados en el procedimiento de codificación de imágenes en movimiento y mediante el aparato de codificación de imágenes en movimiento que se muestra en cada una de las realizaciones. Los datos multiplexados son una secuencia digital en el formato de secuencia de transporte MPEG-2.

La figura 19 ilustra una estructura de los datos multiplexados. Como se ilustra en la figura 19, los datos multiplexados pueden obtenerse multiplexando al menos uno de entre una secuencia de vídeo, una secuencia de  
35 audio, una secuencia de gráficos de presentación (PG) y una secuencia de gráficos interactivos. La secuencia de vídeo representa el vídeo principal y el vídeo secundario de una película, la secuencia de audio (IG) representa una parte de audio principal y una parte de audio secundaria que se mezcla con la parte de audio principal, y la secuencia de gráficos de presentación representa los subtítulos de la película. Aquí, el vídeo principal es el vídeo normal que se muestra en una pantalla, y el vídeo secundario es el vídeo que se muestra en una ventana más pequeña en el vídeo principal. Además, la secuencia de gráficos interactivos representa una pantalla interactiva que se genera al organizar los componentes GUI en una pantalla. La secuencia de vídeo se codifica en el procedimiento de codificación de imágenes en movimiento o mediante el aparato de codificación de imágenes en movimiento que se muestra en cada una de las realizaciones, o en un procedimiento de codificación de imágenes en movimiento o mediante un aparato de codificación de imágenes en movimiento de conformidad con un estándar convencional, tal  
40 como MPEG-2, MPEG-4 AVC, y VC-1. La secuencia de audio se codifica según un estándar, tal como Dolby-AC-3, Dolby Digital Plus, MLP, DTS, DTS-HD y PCM lineal.

Cada secuencia incluida en los datos multiplexados se identifica mediante PID. Por ejemplo, 0x1011 se asigna a la secuencia de vídeo que se usará para el vídeo de una película, 0x1100 a 0x111F se asignan a las secuencias de audio, 0x1200 a 0x121F se asignan a las secuencias de gráficos de presentación, 0x1400 a 0x141F se asignan a las  
45 secuencias de gráficos interactivos, 0x1B00 a 0x1B1F se asignan a las secuencias de vídeo que se utilizarán para el vídeo secundario de la película, y 0x1A00 a 0x1A1F se asignan a las secuencias de audio que se utilizarán para que el audio secundario se mezcle con el audio primario.

La figura 20 ilustra esquemáticamente cómo se multiplexan los datos. Primero, una secuencia ex235 de vídeo compuesta de fotogramas de vídeo y una secuencia ex238 de audio compuesta de muestras de audio se transforman en una secuencia ex236 de paquetes PES y en una secuencia ex239 de paquetes PES, y luego en  
50 paquetes ex237 TS y en paquetes ex240 TS, respectivamente. De manera similar, los datos de una secuencia ex241 de gráficos de presentación y los datos de una secuencia ex244 de gráficos interactivos se transforman en una secuencia ex242 de paquetes PES y en una secuencia ex245 de paquetes PES, y luego en paquetes ex243 TS

y en paquetes ex246 TS, respectivamente. Estos paquetes TS se multiplexan una secuencia para obtener datos ex247 multiplexados.

La figura 21 ilustra cómo una secuencia de vídeo se almacena en una secuencia de paquetes PES con mayor detalle. La primera barra en la figura 21 muestra una secuencia de fotogramas de vídeo en una secuencia de vídeo. La segunda barra muestra a secuencia de paquetes PES. Como lo indican las flechas designadas como yy1, yy2, yy3, e yy4 en la figura 21, la secuencia de vídeo se divide en imágenes como imágenes I, imágenes B e imágenes P, cada una de las cuales es una unidad de presentación de vídeo, y las imágenes se almacenan en un cuerpo de cada uno de los paquetes PES. Cada uno de los paquetes PES tiene una cabecera PES, y la cabecera PES almacena un sello de tiempo de presentación (PTS) que indica un tiempo de visualización de la imagen, y un sello de tiempo de descodificación (DTS) que indica un tiempo de descodificación de la imagen.

La figura 22 ilustra un formato de paquetes TS que se escribirán finalmente en los datos multiplexados. Cada uno de los paquetes TS es un paquete de longitud fija de 188 bytes que incluye una cabecera TS de 4 bytes que tiene información, tal como un PID para identificar una secuencia y un cuerpo TS de 184 bytes para almacenar datos. Los paquetes de PES se dividen y se almacenan en los cuerpos TS, respectivamente. Cuando se utiliza una ROM de BD, a cada uno de los paquetes TS se le asigna una cabecera extra TP de 4 bytes, lo que da como resultado paquetes de origen de 192 bytes. Los paquetes de origen se escriben en los datos multiplexados. La cabecera extra TP almacena información tal como un sello de tiempo de llegada (ATS). El ATS muestra un tiempo de inicio de transferencia en el que cada uno de los paquetes TS se transferirá a un filtro PID. Los paquetes de origen están dispuestos en los datos multiplexados como se muestra en la parte inferior de la figura 22. Los números que se incrementan desde la cabecera de los datos multiplexados se denominan números de paquetes de origen (SPN).

Cada uno de los paquetes TS incluidos en los datos multiplexados incluye no solo secuencias de audio, vídeo, subtítulos y demás, sino también una tabla de asociación de programa (PAT), una tabla de mapas de programa (PMT) y una referencia de reloj del programa (PCR). El PAT muestra lo que indica un PID en un PMT utilizado en los datos multiplexados, y un PID del PAT mismo se registra como cero. El PMT almacena los PID de las secuencias de vídeo, audio, subtítulos y demás incluidos en los datos multiplexados, y la información de los atributos de las secuencias correspondientes a los PID. El PMT también tiene varios descriptores relacionados con los datos multiplexados. Los descriptores tienen información tal como información de control de copia que muestra si se permite o no la copia de los datos multiplexados. El PCR almacena la información de tiempo de STC correspondiente a un ATS que se muestra cuando el paquete de PCR se transfiere a un descodificador, para lograr la sincronización entre un reloj de hora de llegada (ATC) que es un eje de tiempo de ATS y un reloj de tiempo del sistema (STC) que es un eje de tiempo de los PTS y DTS.

La figura 23 ilustra en detalle la estructura de datos del PMT. Una cabecera PMT está dispuesto en la parte superior de la PMT. La cabecera de PMT describe la longitud de los datos incluidos en el PMT y demás. Una pluralidad de descriptores relacionados con los datos multiplexados se dispone después de la cabecera de PMT. La información tal como la información de control de copia se describe en los descriptores. Después de los descriptores, se dispone una pluralidad de cortes de información de secuencias relacionadas con las secuencias incluidas en los datos multiplexados. Cada corte de información de secuencia incluye descriptores de secuencia, cada uno de los cuales describe información, tal como un tipo de secuencia para identificar un codificador/descodificador de compresión de una secuencia, una secuencia de PID e información de atributos de secuencia (tales como una velocidad de fotograma o una relación de aspecto). Los descriptores de secuencia son iguales en número al número de secuencias en los datos multiplexados.

Cuando los datos multiplexados se graban en un medio de grabación y demás, se graban junto con archivos de información de datos multiplexados.

Cada uno de los archivos de información de datos multiplexados es información de gestión de los datos multiplexados como se muestra en la figura 24. Los archivos de información de datos multiplexados tienen una correspondencia de uno a uno con los datos multiplexados, y cada uno de los archivos incluye información de datos multiplexados, información de atributos de secuencia y un mapa de entrada.

Como se ilustra en la figura 24, la información de datos multiplexados incluye una velocidad del sistema, una hora de inicio de reproducción y una hora de finalización de reproducción. La velocidad del sistema indica la velocidad de transferencia máxima a la que un descodificador objetivo del sistema que se describirá más adelante transfiere los datos multiplexados a un filtro PID. Los intervalos de los ATS incluidos en los datos multiplexados se establecen en no mayores que una velocidad del sistema. El tiempo de inicio de la reproducción indica un PTS en un fotograma de vídeo en la cabecera de los datos multiplexados. Se agrega un intervalo de un fotograma a un PTS en un fotograma de vídeo al final de los datos multiplexados, y el PTS se establece en el tiempo de finalización de la reproducción.

Como se muestra en la figura 25, se registra un corte de información de atributo en la información de atributo de secuencia, para cada PID de cada secuencia incluida en los datos multiplexados. Cada corte de información de atributo tiene información diferente dependiendo de si la secuencia correspondiente es una secuencia de vídeo, una secuencia de audio, una secuencia de gráficos de presentación o una secuencia de gráficos interactivos. Cada corte de información de atributos de secuencia de vídeo contiene información que incluye qué tipo de

codificador/descodificador de compresión se usa para comprimir la secuencia de vídeo, y la resolución, la relación de aspecto y la velocidad de fotogramas de los cortes de datos de imagen que se incluyen en la secuencia de vídeo. Cada corte de información de atributos de secuencia de audio contiene información que incluye qué tipo de codificador/descodificador de compresión se utiliza para comprimir la secuencia de audio, cuántos canales se incluyen en la secuencia de audio, qué idioma admite la secuencia de audio y cuán alta es la frecuencia de muestreo. La información de atributos de secuencia de vídeo y la información de atributos de secuencia de audio se usan para inicializar un descodificador antes de que el reproductor reproduzca la información.

En la presente realización, los datos multiplexados a usar son de un tipo de secuencia incluida en el PMT. Además, cuando los datos multiplexados se graban en un medio de grabación, se utiliza la información de atributos de secuencia de vídeo incluida en la información de datos multiplexados. Más específicamente, el procedimiento de codificación de imágenes en movimiento, o el aparato de codificación de imágenes en movimiento, descritos en cada una de las realizaciones, incluye un paso o una unidad para asignar información única que indica datos de vídeo generados por el procedimiento de codificación de imágenes en movimiento, o el aparato de codificación de imágenes en movimiento en cada una de las realizaciones, al tipo de secuencia incluida en la PMT o a la información de atributo de secuencia de vídeo. Con la configuración, los datos de vídeo generados por el procedimiento de codificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento descritos en cada una de las realizaciones pueden distinguirse de los datos de vídeo que se ajustan a otra norma.

Además, la figura 26 ilustra los pasos del procedimiento de descodificación de imágenes en movimiento según la presente realización. En el paso exS100, el tipo de secuencia incluida en la PMT o la información de atributo de flujo de vídeo incluida en la información de datos multiplexados se obtiene de los datos multiplexados. A continuación, en el paso exS101, se determina si el tipo de secuencia o la información del atributo de secuencia de vídeo indica que los datos multiplexados se generan mediante el procedimiento de codificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento en cada una de las realizaciones. Cuando se determina que el tipo de secuencia o la información del atributo de secuencia de vídeo indica que los datos multiplexados se generan mediante el procedimiento de codificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento en cada una de las realizaciones, en el paso exS102, la descodificación se realiza mediante el procedimiento de descodificación de imágenes en movimiento en cada una de las realizaciones. Además, cuando el tipo de flujo o la información de atributo de secuencia de vídeo indica conformidad con los estándares convencionales, tales como MPEG-2, MPEG-4 AVC y VC-1, en el paso exS103, la descodificación se realiza mediante un procedimiento de descodificación de imágenes en movimiento de conformidad con los estándares convencionales.

Como tal, la asignación de un nuevo valor único al tipo de secuencia o la información de atributo de secuencia de vídeo permite determinar si el procedimiento de descodificación de imágenes en movimiento o el aparato de descodificación de imágenes en movimiento que se describe en cada una de las realizaciones puede realizar la descodificación. Incluso cuando se ingresan datos multiplexados que se ajustan a un estándar diferente, se puede seleccionar un procedimiento o aparato de descodificación apropiado. Por lo tanto, es posible descodificar información sin ningún error. Además, el procedimiento o aparato de codificación de imágenes en movimiento, o el procedimiento o aparato de descodificación de imágenes en movimiento de la presente realización se puede usar en los dispositivos y sistemas descritos anteriormente.

(Realización C)

Cada uno de los procedimientos de codificación de imágenes en movimiento, el aparato de codificación de imágenes en movimiento, el procedimiento de descodificación de imágenes en movimiento y el aparato de descodificación de imágenes en movimiento de cada una de las realizaciones se obtienen generalmente en forma de un circuito integrado o de un circuito integrado a gran escala (LSI). Como ejemplo del LSI, la figura 27 ilustra una configuración del LSI ex500 que está hecha en un chip. El LSI ex500 incluye elementos ex501, ex502, ex503, ex504, ex505, ex506, ex507, ex508 y ex509 que se describen a continuación, y los elementos están conectados entre sí a través de un bus ex510. La unidad ex505 del circuito de la fuente de alimentación se activa al suministrar energía a cada uno de los elementos cuando la unidad ex505 del circuito de la fuente de alimentación está encendida.

Por ejemplo, cuando se realiza la codificación, el LSI ex500 recibe una señal AV de un micrófono ex117, de una cámara ex113 y demás a través de una AV IO ex509 bajo el control de una unidad ex501 de control que incluye una CPU ex502, un controlador ex503 de memoria, un controlador ex504 de secuencia y una unidad ex512 de control de frecuencia de accionamiento. La señal AV recibida se almacena temporalmente en una memoria ex511 externa, tal como una SDRAM. Bajo el control de la unidad ex501 de control, los datos almacenados se segmentan en cortes de datos según la cantidad de procesamiento y la velocidad a transmitir a una unidad ex507 de procesamiento de señal. Luego, la unidad ex507 de procesamiento de señal codifica una señal de audio y/o una señal de vídeo. Aquí, la codificación de la señal de vídeo es la codificación descrita en cada una de las realizaciones. Además, la unidad ex507 de procesamiento de señales a veces multiplexa los datos de audio codificados y los datos de vídeo codificados, y un flujo ex506 IO proporciona los datos multiplexados al exterior. Los datos multiplexados proporcionados se transmiten a la estación ex107 base o se escriben sobre el medio ex215 de grabación. Cuando los conjuntos de datos se multiplexan, los datos deben almacenarse temporalmente en la memoria intermedia ex508 a fin de que los conjuntos de datos se sincronicen entre sí.

Aunque la memoria ex511 es un elemento fuera del LSI ex500, puede incluirse en el LSI ex500. La memoria intermedia ex508 no está limitada a una memoria intermedia, sino que puede estar compuesto de memorias intermedias. Además, el LSI ex500 se puede hacer en un chip o en una pluralidad de chips.

5 Además, aunque la unidad ex501 de control incluye la CPU ex502, el controlador ex503 de memoria, el controlador ex504 de flujo, la unidad ex512 de control de frecuencia de accionamiento, la configuración de la unidad ex501 de control no está limitada a los mismos. Por ejemplo, la unidad ex507 de procesamiento de señales puede incluir además una CPU. La inclusión de otra CPU en la unidad ex507 de procesamiento de señal puede mejorar la velocidad de procesamiento. Además, como otro ejemplo, la CPU ex502 puede servir o ser parte de la unidad ex507 de procesamiento de señal y, por ejemplo, puede incluir una unidad de procesamiento de señal de audio. En tal caso, la unidad ex501 de control incluye la unidad ex507 de procesamiento de señales o la CPU ex502 que incluye una parte de la unidad ex507 de procesamiento de señales.

El nombre utilizado aquí es LSI, pero también puede llamarse IC, sistema LSI, super LSI o ultra LSI en función del grado de integración.

15 Además, las formas de lograr la integración no se limitan al LSI, y un circuito especial o un procesador de propósito general, etc., también pueden lograr la integración. La matriz de puerta programable de campo (FPGA) que se puede programar después de la fabricación de los LSI o un procesador reconfigurable que permita la reconfiguración de la conexión o la configuración de una LSI se puede usar con el mismo fin.

20 En el futuro, con el avance en la tecnología de semiconductores, una tecnología completamente nueva puede reemplazar al LSI. Los bloques funcionales pueden integrarse utilizando dicha tecnología. La posibilidad es que la presente invención se aplique a la biotecnología.

(Realización D)

25 Cuando los datos de vídeo generados en el procedimiento de codificación de imágenes en movimiento o por el aparato de codificación de imágenes en movimiento descritos en cada una de las realizaciones se descodifican, en comparación con cuando los datos de vídeo se ajustan a un estándar convencional, como MPEG-2, MPEG-4 AVC, y VC-1 se descodifican, la cantidad de procesamiento probablemente aumenta. Por lo tanto, el LSI ex500 debe configurarse a una frecuencia de accionamiento mayor que la de la CPU ex502 para usarse cuando se descodifican datos de vídeo de conformidad con el estándar convencional. Sin embargo, cuando la frecuencia de accionamiento se fija más alta, existe el problema de que aumenta el consumo de energía.

30 Para resolver el problema, el aparato de descodificación de imágenes en movimiento, tal como el televisor ex300 y el LSI ex500, está configurado para determinar a qué estándar se ajustan los datos de vídeo y cambiar entre las frecuencias de accionamiento según el estándar determinado. La figura 28 ilustra una configuración ex800 de la presente realización. Una unidad ex803 de conmutación de frecuencia de accionamiento establece una frecuencia de accionamiento a una frecuencia de accionamiento mayor cuando los datos de vídeo se generan mediante el procedimiento de codificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento descritos en cada una de las realizaciones. Luego, la unidad ex803 de conmutación de frecuencia de accionamiento instruye a una unidad ex801 de procesamiento de descodificación que ejecuta el procedimiento de descodificación de imágenes en movimiento descrito en cada una de las realizaciones para descodificar los datos de vídeo. Cuando los datos de vídeo se ajustan a la norma convencional, la unidad ex803 de conmutación de frecuencia de accionamiento establece una frecuencia de accionamiento a una frecuencia de accionamiento menor que la de los datos de vídeo generados por el procedimiento de codificación de imágenes en movimiento o por el aparato de codificación de imágenes en movimiento descritos en cada una de las realizaciones. Luego, la unidad ex803 de conmutación de frecuencia de accionamiento instruye a la unidad ex802 de procesamiento de descodificación que se ajusta a la norma convencional para descodificar los datos de vídeo.

45 Más específicamente, la unidad ex803 de conmutación de frecuencia de accionamiento incluye la CPU ex502 y la unidad ex512 de control de frecuencia de accionamiento de la figura 27. Aquí, cada una de las unidades ex801 de procesamiento de descodificación que ejecuta el procedimiento de descodificación de imágenes en movimiento descrito en cada una de las realizaciones y la unidad ex802 de procesamiento de descodificación que se ajusta a la norma convencional corresponde a la unidad ex507 de procesamiento de señal de la figura 27. La CPU ex502 determina a qué norma se ajustan los datos de vídeo. Luego, la unidad ex512 de control de frecuencia de accionamiento determina una frecuencia de accionamiento basada en una señal de la CPU ex502. Además, la unidad ex507 de procesamiento de señales descodifica los datos de vídeo en función de la señal de la CPU ex502. Por ejemplo, la información de identificación descrita en la Realización B se usa probablemente para identificar los datos de vídeo. La información de identificación no se limita a la descrita en la Realización B, sino que puede ser cualquier información siempre que la información indique a qué norma se ajustan los datos de vídeo. Por ejemplo, cuándo y a qué norma se ajustan los datos de vídeo se puede determinar en función de una señal externa para determinar que los datos de vídeo se utilizan para un televisor o un disco, etc., la determinación se puede hacer en base a dicha señal externa. Además, la CPU ex502 selecciona una frecuencia de accionamiento basada, por ejemplo, en una tabla de consulta en la que las normas de los datos de vídeo están asociados con las frecuencias de accionamiento como se muestra en la figura 30. La frecuencia de accionamiento puede seleccionarse

almacenando la tabla de consulta en la memoria intermedia ex508 y en una memoria interna de un LSI, y con referencia a la tabla de consulta a través de la CPU ex502.

La figura 29 ilustra los pasos para ejecutar un procedimiento de la presente realización. Primero, en el paso exS200, la unidad ex507 de procesamiento de señal obtiene información de identificación de los datos multiplexados. A continuación, en el paso exS201, la CPU ex502 determina si los datos de vídeo se generan o no mediante el procedimiento de codificación y el aparato de codificación descritos en cada una de las realizaciones, en función de la información de identificación. Cuando los datos de vídeo se generan mediante el procedimiento de codificación de imágenes en movimiento y el aparato de codificación de imágenes en movimiento descritos en cada una de las realizaciones, en el paso exS202, la CPU ex502 transmite una señal para fijar la frecuencia de accionamiento a una frecuencia de accionamiento mayor que la unidad ex512 de control de frecuencia de accionamiento. Luego, la unidad ex512 de control de frecuencia de accionamiento ajusta la frecuencia de accionamiento a la frecuencia de accionamiento más alta. Por otro lado, cuando la información de identificación indica que los datos de vídeo se ajustan a la norma convencional, tal como MPEG-2, MPEG-4 AVC y VC-1, en el paso exS203, la CPU ex502 transmite una señal para fijar la frecuencia de accionamiento a una frecuencia de accionamiento inferior a la unidad ex512 de control de frecuencia de accionamiento. Luego, la unidad ex512 de control de frecuencia de accionamiento establece la frecuencia de accionamiento a la frecuencia de accionamiento más baja que la del caso en el que los datos de vídeo se generan mediante el procedimiento de codificación de imágenes en movimiento y el aparato de codificación de imágenes en movimiento descritos en cada una de las realizaciones.

Además, junto con la conmutación de las frecuencias de accionamiento, el efecto de ahorro de energía puede mejorarse cambiando la tensión que se aplicará al LSI ex500 o a un aparato que incluya el LSI ex500. Por ejemplo, cuando la frecuencia de accionamiento se establece más baja, la tensión que se aplicará al LSI ex500 o al aparato que incluye el LSI ex500 probablemente esté configurada a una tensión inferior a la del caso en el que la frecuencia de accionamiento se establece más alta.

Además, cuando la cantidad de procesamiento para la descodificación es mayor, la frecuencia de accionamiento puede establecerse más alta, y cuando la cantidad de procesamiento para la descodificación es menor, la frecuencia de accionamiento puede establecerse más baja como procedimiento para establecer la frecuencia de accionamiento. Por lo tanto, el procedimiento de configuración no se limita a los descritos anteriormente. Por ejemplo, cuando la cantidad de procesamiento para descodificar datos de vídeo de conformidad con MPEG-4 AVC es mayor que la cantidad de procesamiento para descodificar datos de vídeo generados por el procedimiento de codificación de imágenes en movimiento y el aparato de codificación de imágenes en movimiento descrito en cada una de las realizaciones, la frecuencia de accionamiento es probable que se establezca en orden inverso a la configuración descrita anteriormente.

Además, el procedimiento para establecer la frecuencia de accionamiento no se limita al procedimiento para establecer una frecuencia de accionamiento menor. Por ejemplo, cuando la información de identificación indica que los datos de vídeo son generados por el procedimiento de codificación de imágenes en movimiento y el aparato de codificación de imágenes en movimiento descrito en cada una de las realizaciones, la tensión a aplicar al LSI ex500 o al aparato que incluye el LSI ex500 probablemente se establezca más alta. Cuando la información de identificación indica que los datos de vídeo se ajustan a la norma convencional, tal como MPEG2, MPEG-4 AVC y VC-1, la tensión a aplicar al LSI ex500 o al aparato que incluye el LSI ex500 probablemente se establezca más baja. Como otro ejemplo, cuando la información de identificación indica que los datos de vídeo son generados por el procedimiento de codificación de imágenes en movimiento y el aparato de codificación de imágenes en movimiento descritos en cada una de las realizaciones, la activación de la CPU ex502 probablemente no tiene que suspenderse. Cuando la información de identificación indica que los datos de vídeo cumplen con la norma convencional, tal como MPEG-2, MPEG-4 AVC y VC-1, la activación de la CPU ex502 probablemente se suspenda en un momento dado porque la CPU ex502 tiene una capacidad de procesamiento adicional. Incluso cuando la información de identificación indica que los datos de vídeo son generados por el procedimiento de codificación de imágenes en movimiento y el aparato de codificación de imágenes en movimiento descritos en cada una de las realizaciones, en el caso en el que la CPU ex502 tenga capacidad de procesamiento adicional, la activación de la CPU ex502 probablemente se suspenda en un momento dado. En tal caso, el tiempo de suspensión probablemente se establezca más corto que en el caso en que cuando la información de identificación indique que los datos de vídeo se ajustan a la norma convencional, tal como MPEG-2, MPEG-4 AVC y VC-1.

Por consiguiente, el efecto de ahorro de energía puede mejorarse cambiando entre las frecuencias de activación según la norma a la que se ajustan los datos de vídeo. Además, cuando el LSI ex500 o el aparato que incluye el LSI ex500 se activa con una batería, la vida de la batería se puede extender con el efecto de ahorro de energía.

(Realización E)

Hay casos en los que se proporciona a los dispositivos y sistemas una pluralidad de datos de vídeo que se ajustan a diferentes normas, tales como un televisor y un teléfono móvil. Para poder descodificar la pluralidad de datos de vídeo que se ajustan a las diferentes normas, la unidad ex507 de procesamiento de señal del LSI ex500 debe ajustarse a las diferentes normas. Sin embargo, los problemas de aumento en la escala del circuito del LSI ex500 y

el aumento en el coste surgen con el uso individual de las unidades ex507 de procesamiento de señal ex507 que se ajustan a las normas respectivas.

Para resolver el problema, lo que se concibe es una configuración en la que la unidad de procesamiento de descodificación para implementar el procedimiento de descodificación de imágenes en movimiento descrito en cada una de las realizaciones y la unidad de procesamiento de descodificación que se ajusta a la norma convencional, tal como AVC MPEG 2, MPEG-4 y VC-1 se comparten parcialmente. El Ex900 en la figura 31A muestra un ejemplo de la configuración. Por ejemplo, el procedimiento de descodificación de imágenes en movimiento que se describe en cada una de las realizaciones y el procedimiento de descodificación de imágenes en movimiento que se ajusta al AVC MPEG-4 tienen parcialmente en común los detalles del procesamiento, tales como la codificación de entropía, la cuantificación inversa, el filtrado de desbloqueo y la predicción compensada de movimiento. Los detalles del procesamiento que se compartirá probablemente incluyen el uso de una unidad ex902 de procesamiento de descodificación que se ajusta al AVC MPEG-4. En cambio, una unidad ex901 de procesamiento de descodificación dedicada se usa probablemente para otro procesamiento exclusivo de un aspecto de la presente invención. Dado que el aspecto de la presente invención se caracteriza por una cuantificación inversa en particular, por ejemplo, la unidad ex901 de procesamiento de descodificación dedicada se utiliza para la cuantificación inversa. De lo contrario, la unidad de procesamiento de descodificación probablemente se comparte para una de las descodificaciones de entropía, para el filtrado de desbloqueo y para la compensación de movimiento, o para todo el procesamiento. La unidad de procesamiento de descodificación para implementar el procedimiento de descodificación de imágenes en movimiento que se describe en cada una de las realizaciones puede repartirse para compartir el procesamiento, y una unidad de procesamiento de descodificación dedicada puede usarse para el procesamiento exclusivo del AVC MPEG-4.

Además, el ex1000 de la figura 31B muestra otro ejemplo en el que el procesamiento está parcialmente compartido. Este ejemplo utiliza una configuración que incluye una unidad ex1001 de procesamiento de descodificación dedicada que admite el procesamiento exclusivo de un aspecto de la presente invención, una unidad ex1002 de procesamiento de descodificación dedicada que admite el procesamiento exclusivo a otra norma convencional y una unidad ex1003 de procesamiento de descodificación que admite que el procesamiento se comparta entre el procedimiento de descodificación de imágenes en movimiento según el aspecto de la presente invención y el procedimiento de descodificación de imágenes en movimiento convencional. En este caso, las unidades ex1001 y ex1002 de procesamiento de descodificación dedicadas no están necesariamente especializadas para el procesamiento según el aspecto de la presente invención y el procesamiento de la norma convencional, respectivamente, y pueden ser capaces de implementar el procesamiento general. Además, la configuración de la presente realización puede implementarse mediante el LSI ex500.

Como tal, es posible reducir la escala del circuito de un LSI y reducir el coste compartiendo la unidad de procesamiento de descodificación para que el procesamiento sea compartido entre el procedimiento de descodificación de imágenes en movimiento según el aspecto de la presente invención y el procedimiento de descodificación de imágenes en movimiento en conformidad con la norma convencional.

El experto en la técnica apreciará que pueden realizarse numerosas variaciones y/o modificaciones a la presente invención tal como se muestra en las realizaciones específicas sin apartarse del alcance de la invención tal como se define en las reivindicaciones. Las presentes realizaciones, por lo tanto, deben considerarse en todos los aspectos como ilustrativas y no restrictivas.

#### Aplicabilidad industrial

La presente invención es aplicable a un aparato de codificación que codifica audio, imágenes fijas y vídeo, y a un aparato de descodificación que descodifica datos codificados por el aparato de codificación. Por ejemplo, la presente invención es aplicable a diversos dispositivos audiovisuales, tales como dispositivos de audio, teléfonos móviles, cámaras digitales, grabadoras BD y televisores digitales.

**REIVINDICACIONES**

1. Un procedimiento de codificación de un vídeo en una secuencia de bits de vídeo codificado con predicción temporal de vector de movimiento, comprendiendo el procedimiento:
  - 5           determinar (S204) un valor de un indicador para indicar si se utiliza o no se utiliza la predicción temporal de vector de movimiento para la predicción entre imágenes de una unidad de sub-imagen de una imagen;
  - escribir (s206) el indicador que tiene dicho valor en una cabecera de la unidad de sub-imagen o en una cabecera de la imagen; y
  - en donde si el indicador indica que se utiliza la predicción temporal de vector de movimiento, el procedimiento comprende, además:
    - 10           crear (S210) una primera lista de predictores de vectores de movimiento que comprende una pluralidad de predictores de vector de movimiento que incluyen al menos un predictor temporal de vector de movimiento derivado de al menos un vector de movimiento de una imagen de referencia co-localizada;
    - seleccionar (S212) un predictor de vector de movimiento de la primera lista para una unidad de predicción en la unidad de sub-imagen; y
    - 15           escribir (S214) un primer parámetro en la secuencia de bits de vídeo codificado para indicar el predictor de vector de movimiento seleccionado de la primera lista;
    - caracterizado porque**
      - 20           el valor del indicador se establece para indicar que la predicción temporal de vector de movimiento no se utiliza si se determina que la capa temporal de la imagen es la capa inferior o base y, de lo contrario, el valor del indicador se establece para indicar que se utiliza la predicción temporal de vector de movimiento.
  2. El procedimiento según la reivindicación 1, en el que si el indicador indica que no se utiliza la predicción temporal de vector de movimiento, el procedimiento comprende además:
    - crear (S216) una segunda lista de predictores de vector de movimiento que comprende una pluralidad de predictores de vector de movimiento sin ningún predictor temporal de vector de movimiento;
    - 25           seleccionar (S218) un predictor de vector de movimiento de la segunda lista para una unidad de predicción en la unidad de sub-imagen; y
    - escribir (S220) un segundo parámetro en la secuencia de bits de vídeo codificado para indicar el predictor de vector de movimiento seleccionado de la segunda lista.
  3. El procedimiento según la reivindicación 1 o la 2, en el que el valor del indicador se determina adicionalmente en función de un valor de conteo de orden de imágenes (POC) de la imagen.
  - 30           4. El procedimiento según la reivindicación 3, en el que sí se determina que el valor POC de la imagen es mayor que cualquier valor POC de imágenes de referencia de una memoria intermedia de imágenes del descodificador (DPB), el valor del indicador se establece para indicar que no se utiliza la predicción temporal de vector de movimiento; de lo contrario, el valor del indicador se establece para indicar que se utiliza la predicción temporal de vector de movimiento.
  - 35           5. El procedimiento según la reivindicación 1 o la 2, en el que el valor del indicador se determina adicionalmente en función de un tipo de unidad de sub-imagen de una unidad de sub-imagen de inter-imagen en la imagen.
  6. El procedimiento según la reivindicación 5, en el que si el tipo de unidad de sub-imagen es un tipo predictivo (P), el valor del indicador se establece para indicar que no se utiliza la predicción temporal de vector de movimiento; de lo contrario, el valor del indicador se establece para indicar que se utiliza la predicción temporal de vector de movimiento.
  - 40           7. El procedimiento según la reivindicación 1 o la 2, en el que el valor del indicador se determina adicionalmente en función de si la imagen que contiene la unidad de sub-imagen es una imagen de punto de acceso aleatorio (RAP).
  8. El procedimiento según la reivindicación 7, en el que si la imagen es la imagen RAP y la unidad de sub-imagen pertenece a una capa no base de la imagen, el valor del indicador se establece para indicar que no se utiliza la predicción temporal de vector de movimiento; de lo contrario, el valor del indicador se establece para indicar que se utiliza la predicción temporal de vector de movimiento.
  - 45           9. El procedimiento según una cualquiera de las reivindicaciones 1 a 8, en el que el indicador se escribe en la cabecera de la unidad de sub-imágenes.

10. El procedimiento según una cualquiera de las reivindicaciones 1 a 9, en el que el procedimiento comprende además escribir uno o más parámetros en la cabecera de la unidad de sub-imagen para especificar el orden de las imágenes de referencia en una o más listas de imágenes de referencia utilizadas para la predicción inter de la unidad de sub-imagen.

5 11. El procedimiento según una cualquiera de las reivindicaciones 1 a 10, en el que el procedimiento comprende además:

realizar (S222) una predicción inter-imagen compensada por movimiento utilizando el predictor de vector de movimiento seleccionado para producir la unidad de predicción;

10 restar (S224) la unidad de predicción de un bloque de muestras originales para producir un bloque de muestras residuales; y

codificar (S226) el bloque de muestras residuales correspondientes a la unidad de predicción en la secuencia de bits de vídeo codificado.

12. Un procedimiento de descodificación de una secuencia de bits de vídeo codificado con predicción temporal de vector de movimiento, comprendiendo el procedimiento:

15 analizar (S402) un indicador de una cabecera de una unidad de sub-imágenes o de una cabecera de una imagen del vídeo codificado; y

determinar (S406) si el indicador indica que la predicción temporal de vector de movimiento se usa o no se usa;

en el que si el indicador indica que se usa la predicción temporal de vector de movimiento, el procedimiento comprende, además:

20 crear (S408) una primera lista de predictores de vector de movimiento que comprende una pluralidad de predictores de vector de movimiento que incluyen al menos un predictor temporal de vector de movimiento derivado de al menos un vector de movimiento de una imagen de referencia co-localizada; y

25 analizar (S410) un primer parámetro de la secuencia de bits de vídeo codificado que indica un predictor de vector de movimiento seleccionado de la primera lista para una unidad de predicción en la unidad de sub-imagen;

**caracterizado porque**

30 el valor del indicador se establece durante la codificación para indicar que no se utiliza la predicción temporal de vector de movimiento si se determina que la capa temporal de la imagen es la capa inferior o base y, de lo contrario, el valor del indicador se establece durante la codificación para indicar que se utiliza la predicción de vector de movimiento.

13. Un aparato para codificar un vídeo en una secuencia de bits de vídeo codificado con predicción temporal de vector de movimiento, comprendiendo el aparato:

35 una unidad (324) de control operable para determinar un valor de un indicador para indicar si se utiliza o no se utiliza la predicción temporal de vector de movimiento para la predicción inter-imagen de una unidad de sub-imagen de una imagen;

una unidad (328) de escritura operable para escribir el indicador que tiene dicho valor en una cabecera de la unidad de sub-imagen o en una cabecera de la imagen;

una unidad (330) de predicción de vector de movimiento; y

40 una unidad (320) de predicción inter-imagen para realizar la predicción inter-imagen en función de un predictor de vector de movimiento seleccionado de la unidad (330) de predicción de vectores de movimiento,

45 en el que la unidad (330) de predicción del vector de movimiento está configurada para recibir el indicador y en función de que el indicador sea un primer valor, la unidad (330) de predicción de vector de movimiento es operable para crear una primera lista de predictores de vector de movimiento que comprende una pluralidad de predictores de vector de movimiento que incluyen al menos un predictor temporal de vector de movimiento derivado de al menos un vector de movimiento de una imagen de referencia co-localizada, y seleccionar un predictor de vector de movimiento de la primera lista para una unidad de predicción en la unidad de sub-imagen; y

la unidad (328) de escritura además es operable para escribir un primer parámetro en la secuencia de bits de vídeo codificado para indicar el predictor de vector de movimiento seleccionado de la primera lista;

50 **caracterizado porque**

el valor del indicador se establece para indicar que no se utiliza la predicción temporal de vector de movimiento si se determina que la capa temporal de la imagen es la capa inferior o base y, de lo contrario, el valor del indicador se establece para indicar que se utiliza la predicción temporal de vector de movimiento.

5 14. Un aparato para decodificar una secuencia de bits de vídeo codificado con predicción temporal de vector de movimiento, comprendiendo el procedimiento:

una unidad (503) de análisis operable para analizar un indicador de una cabecera de una unidad de sub-imagen o de una cabecera de una imagen del vídeo codificado; y

una unidad (522) de predicción de vector de movimiento operable para determinar si el indicador indica que se utiliza o no se utiliza la predicción temporal de vector de movimiento;

10 en el que si el indicador indica que se utiliza la predicción temporal de vector de movimiento,

la unidad (522) de predicción de vector de movimiento también es operable para crear una primera lista de predictores de vector de movimiento que comprende una pluralidad de predictores de vector de movimiento que incluyen al menos un predictor temporal de vector de movimiento derivado de al menos un vector de movimiento de una imagen de referencia co-localizada; y

15 la unidad (503) de análisis es operable además para analizar un primer parámetro de la secuencia de bits de vídeo codificado que indica un predictor de vector de movimiento seleccionado de la primera lista para una unidad de predicción en la unidad de sub-imagen;

**caracterizado porque**

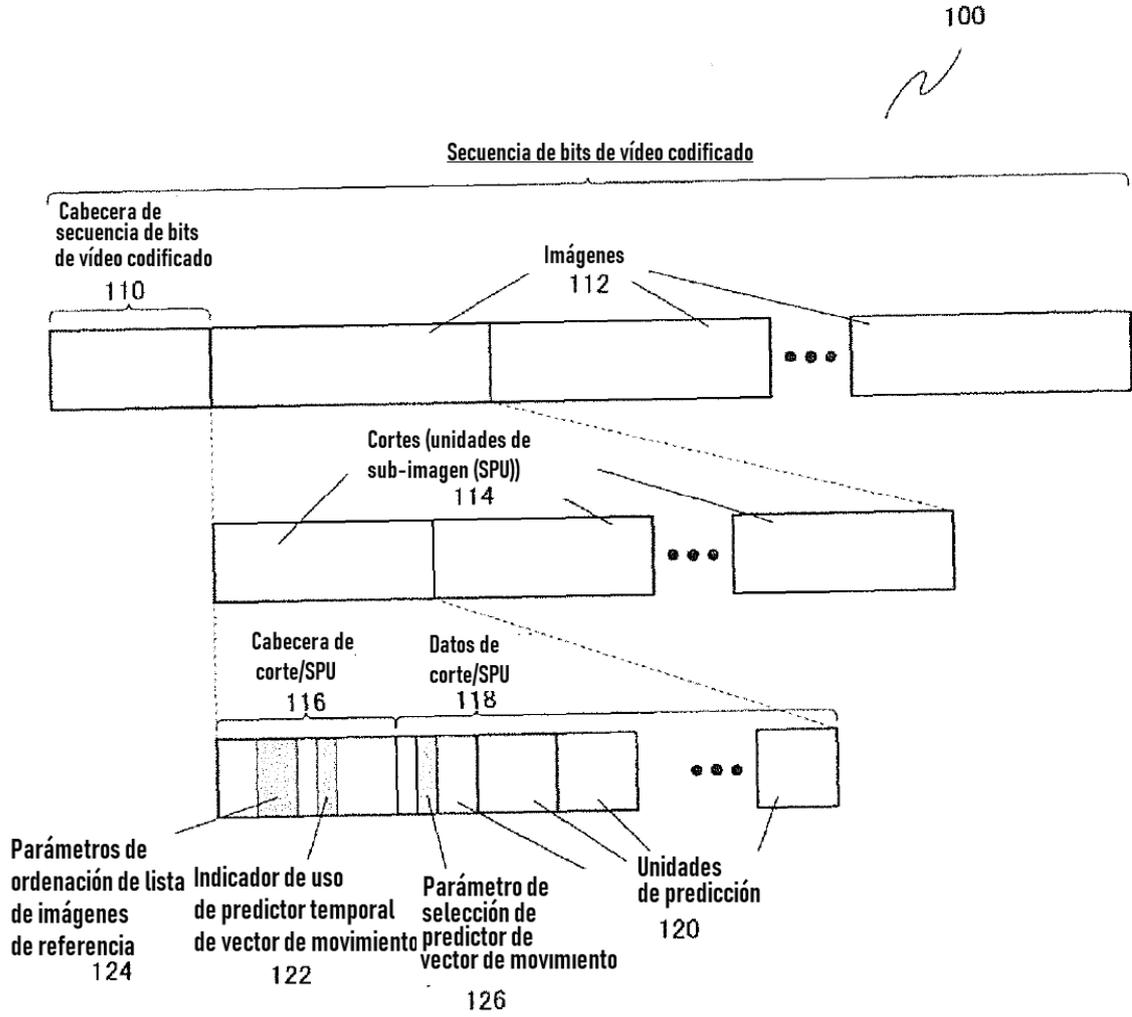
20 el valor del indicador se establece durante la codificación para indicar que no se utiliza la predicción temporal de vector de movimiento si se determina que la capa temporal de la imagen es la capa inferior o base y, de lo contrario, el valor del indicador se establece durante la codificación para indicar que se utiliza la predicción temporal de vector de movimiento.

15. Un sistema para codificar y decodificar un vídeo con predicción temporal de vector de movimiento, comprendiendo el sistema:

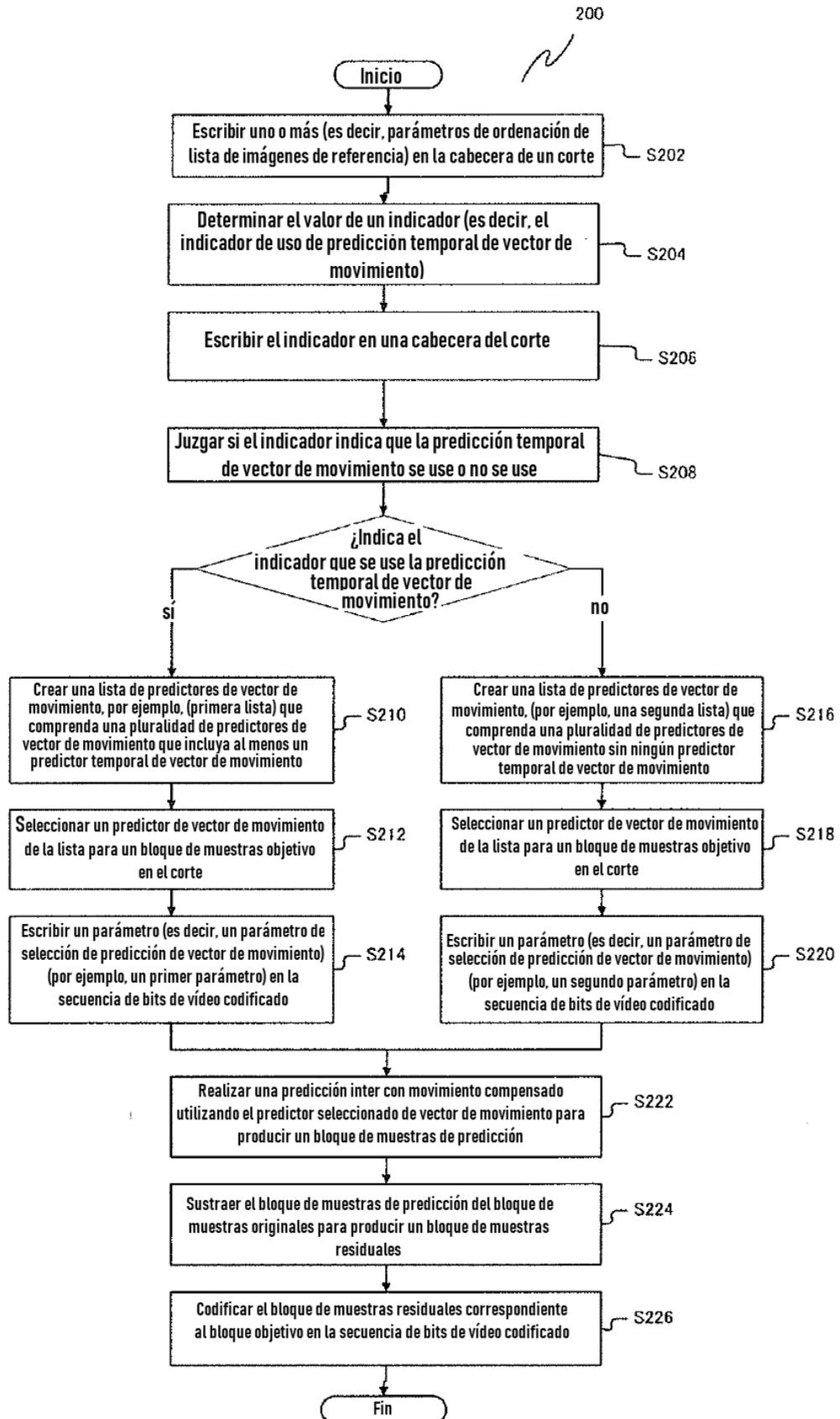
25 el aparato según la reivindicación 13; y

el aparato según la reivindicación 14.

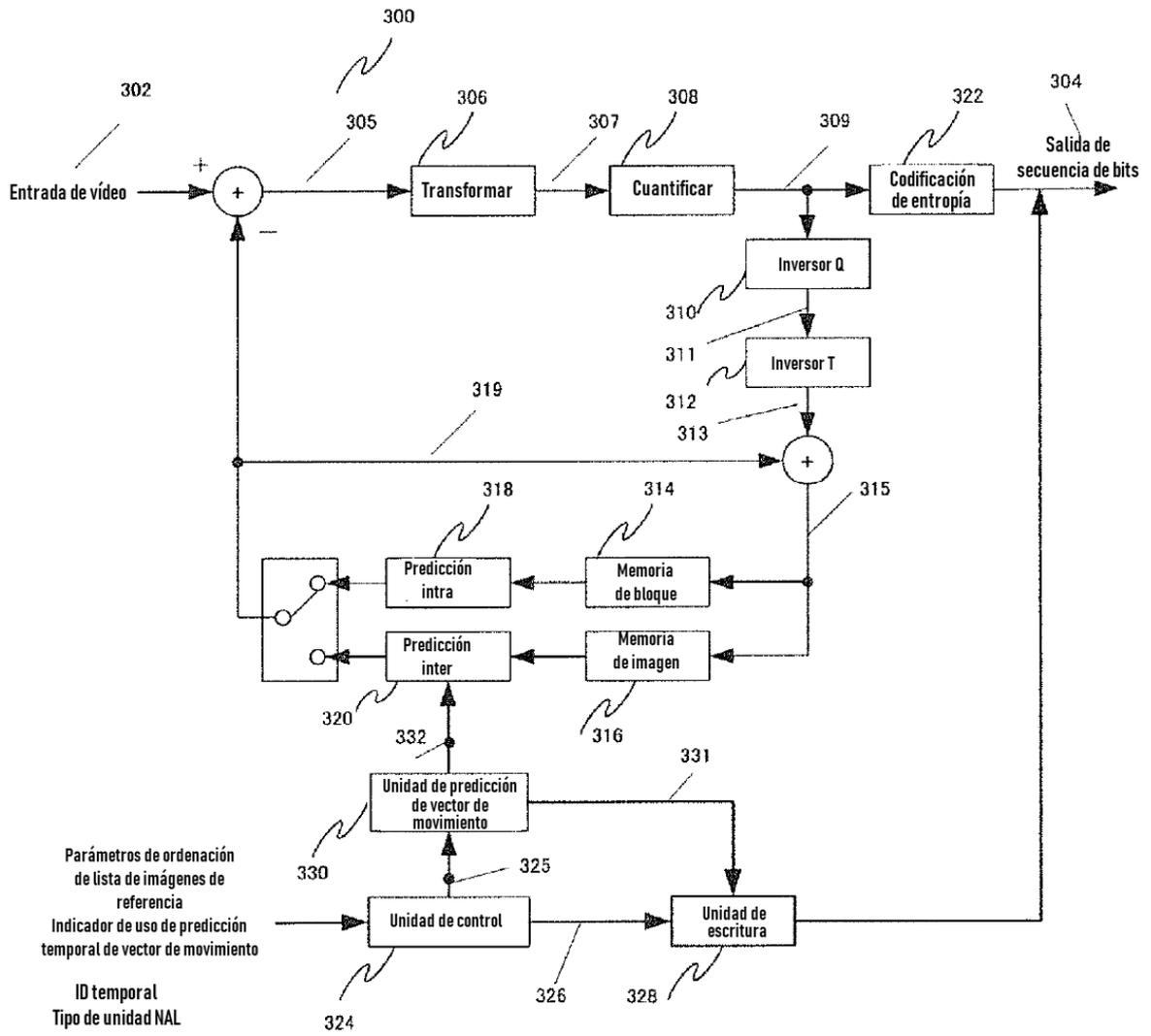
[Fig. 1]



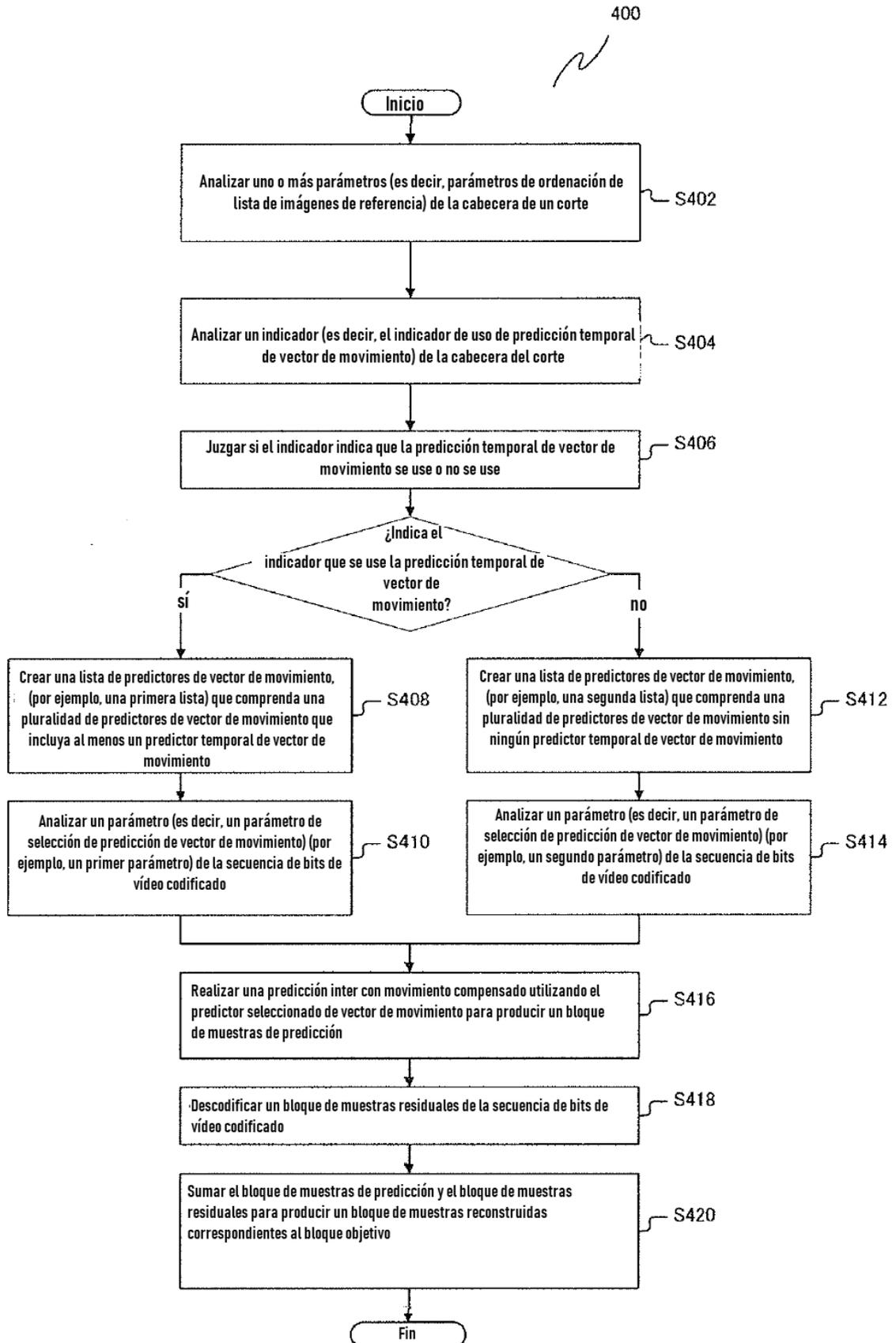
[Fig. 2]



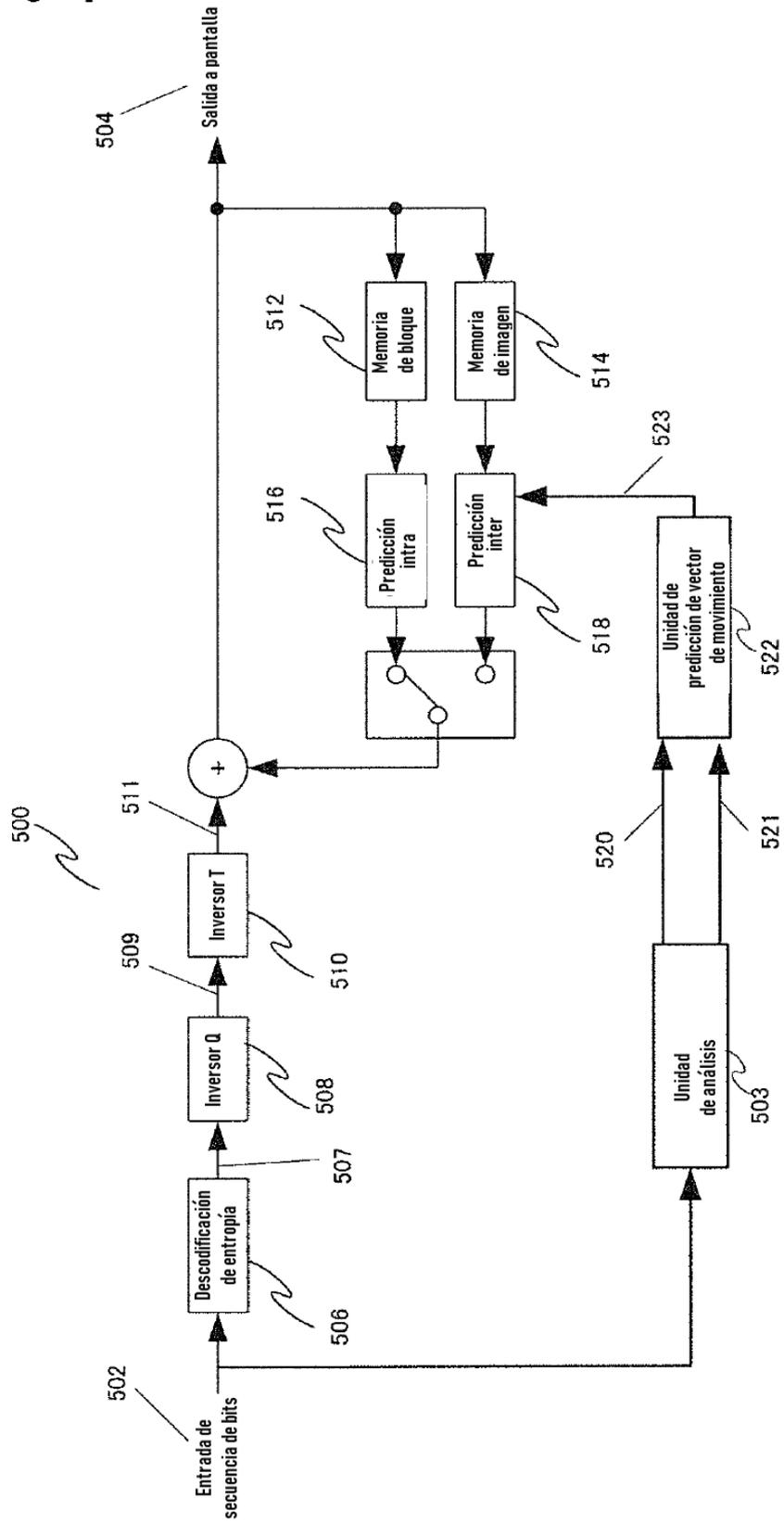
[Fig. 3]



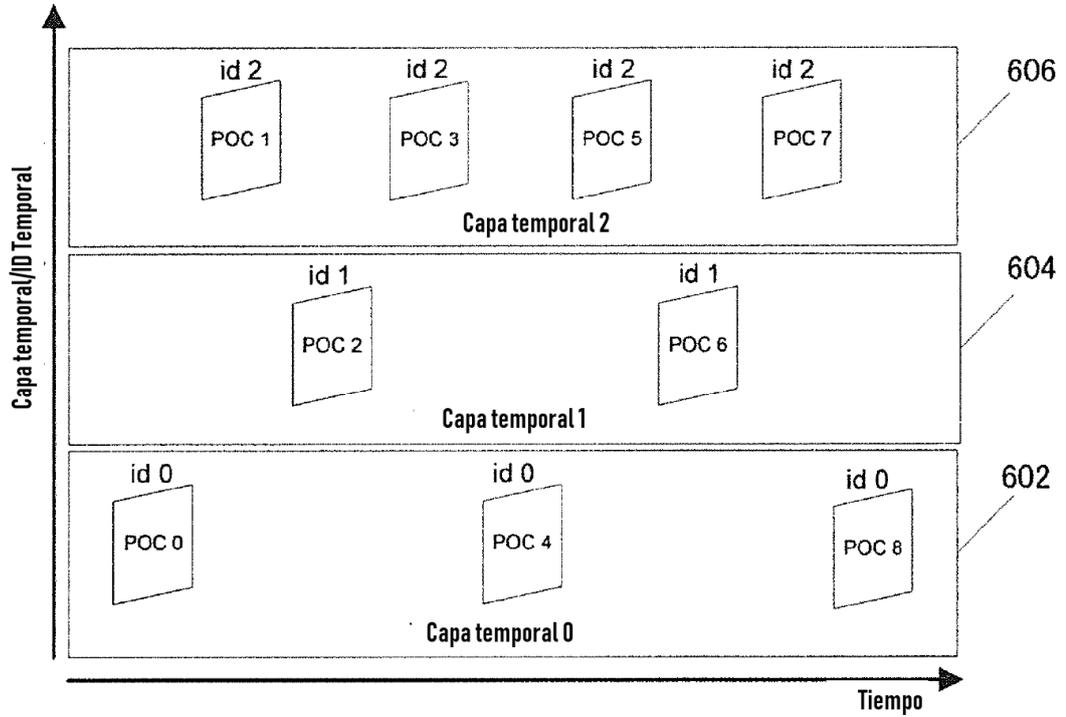
[Fig. 4]



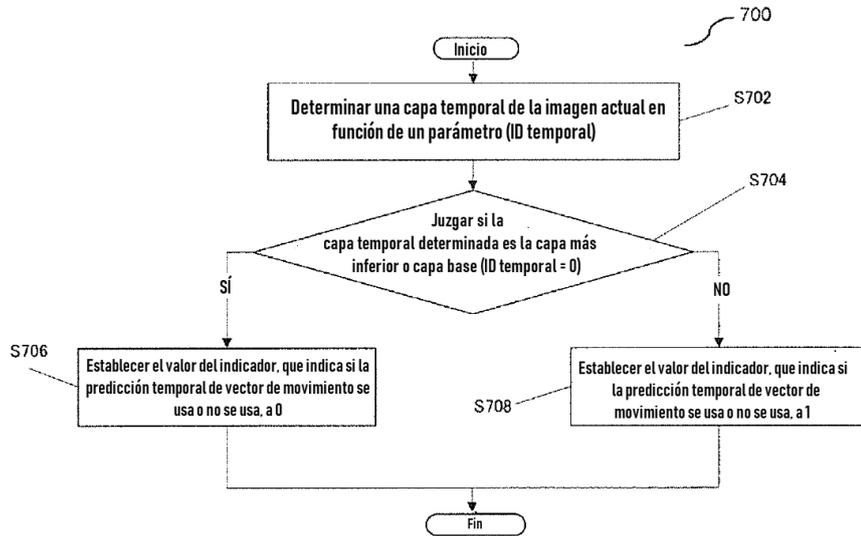
[Fig. 5]



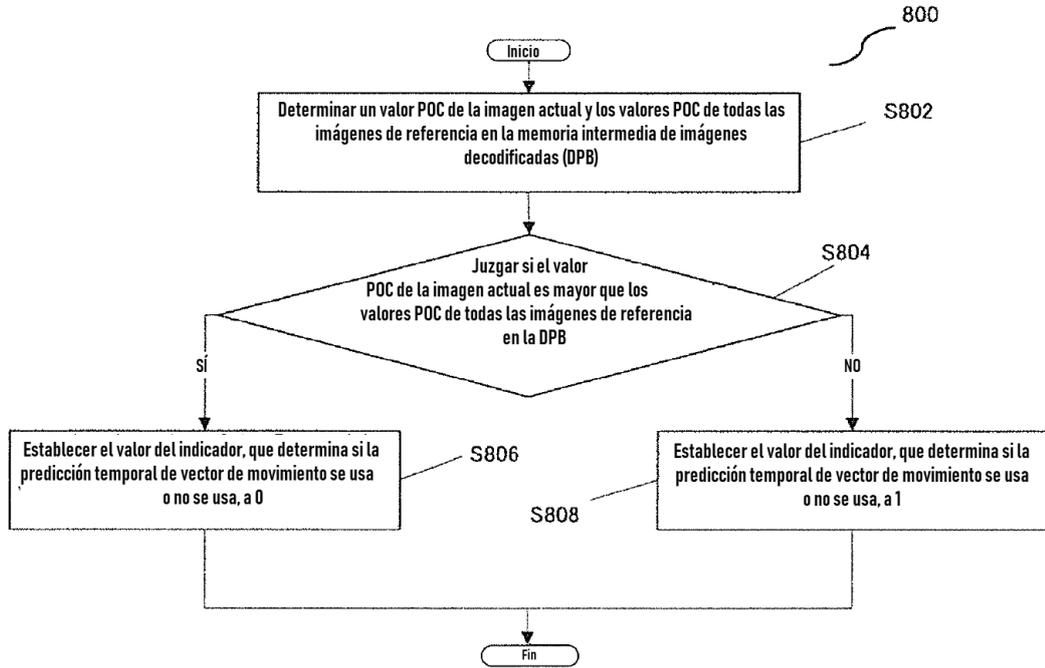
[Fig. 6]



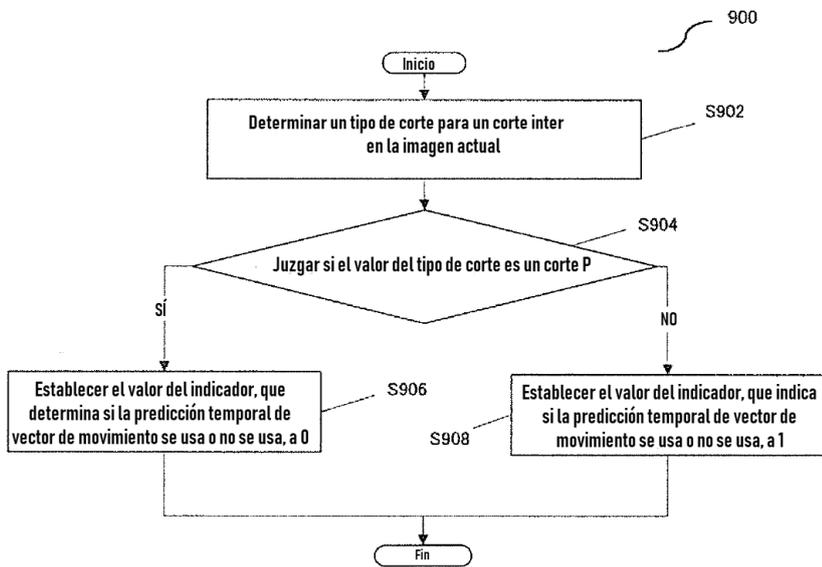
[Fig. 7]



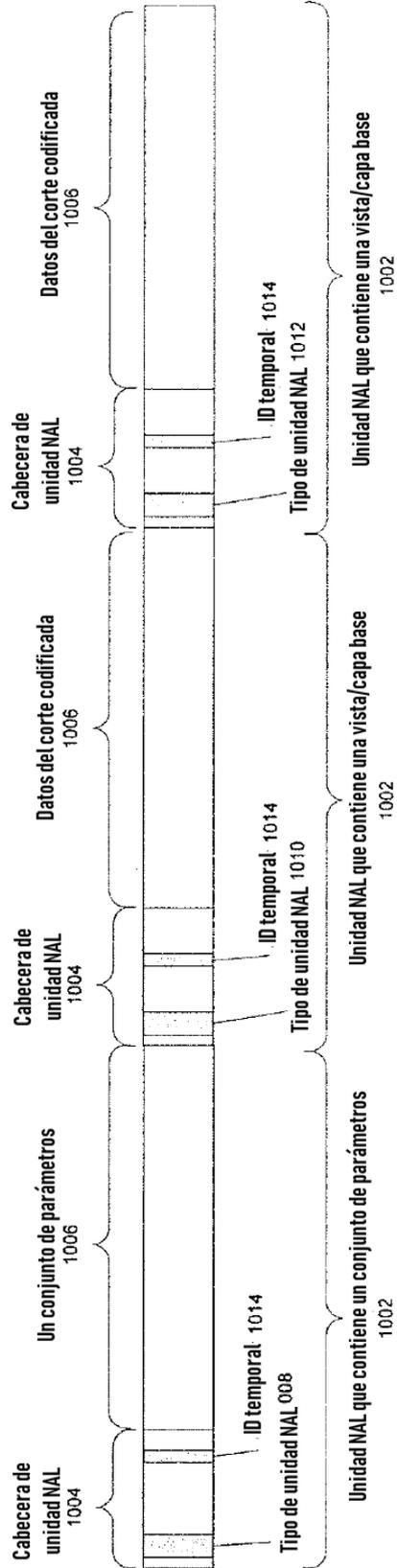
[Fig. 8]



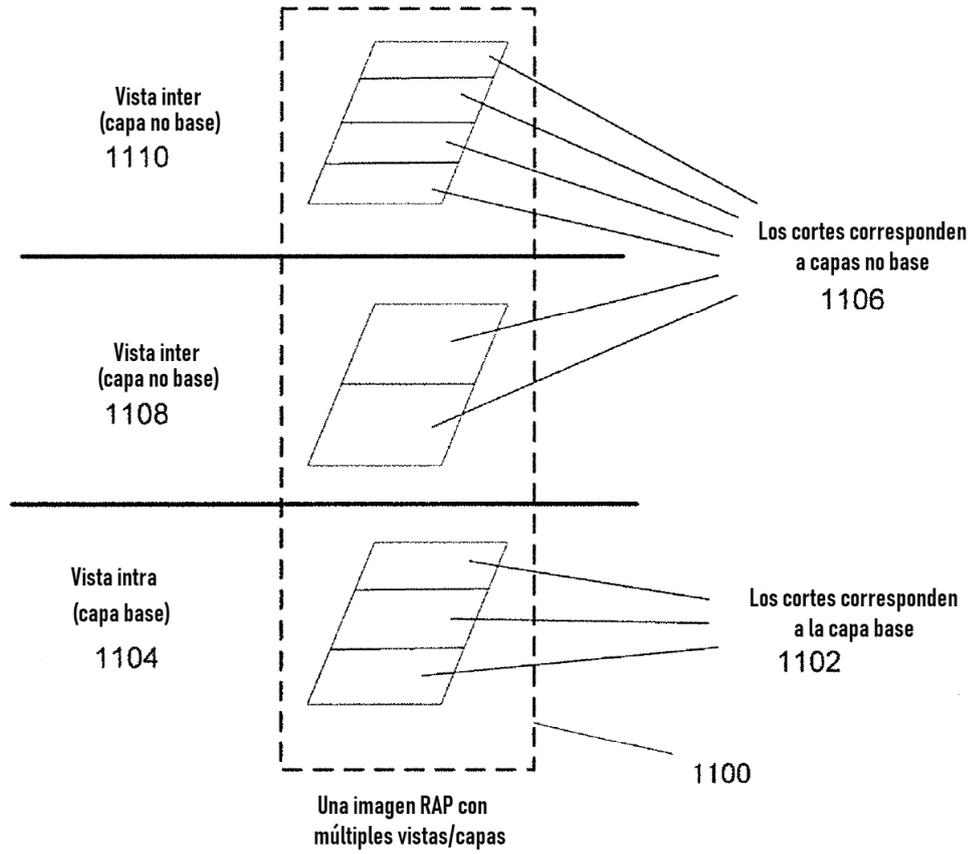
[Fig. 9]



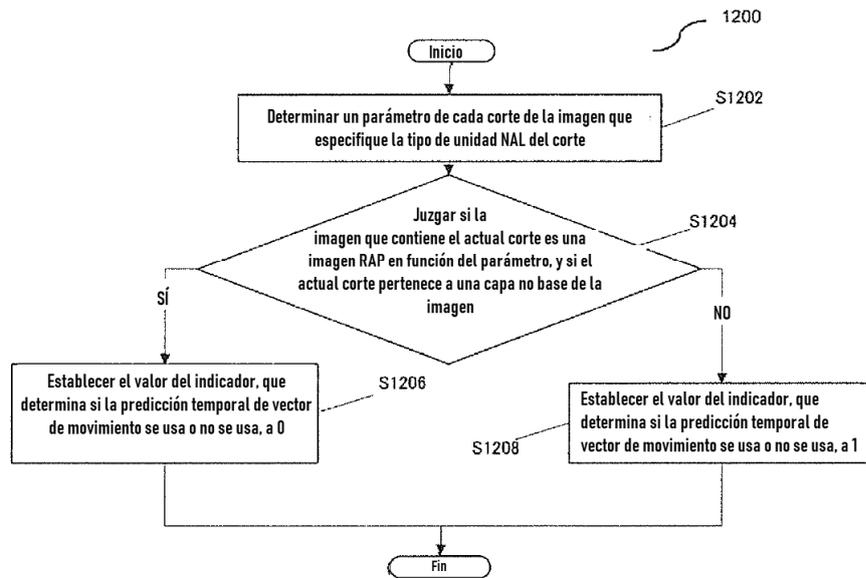
[Fig. 10]



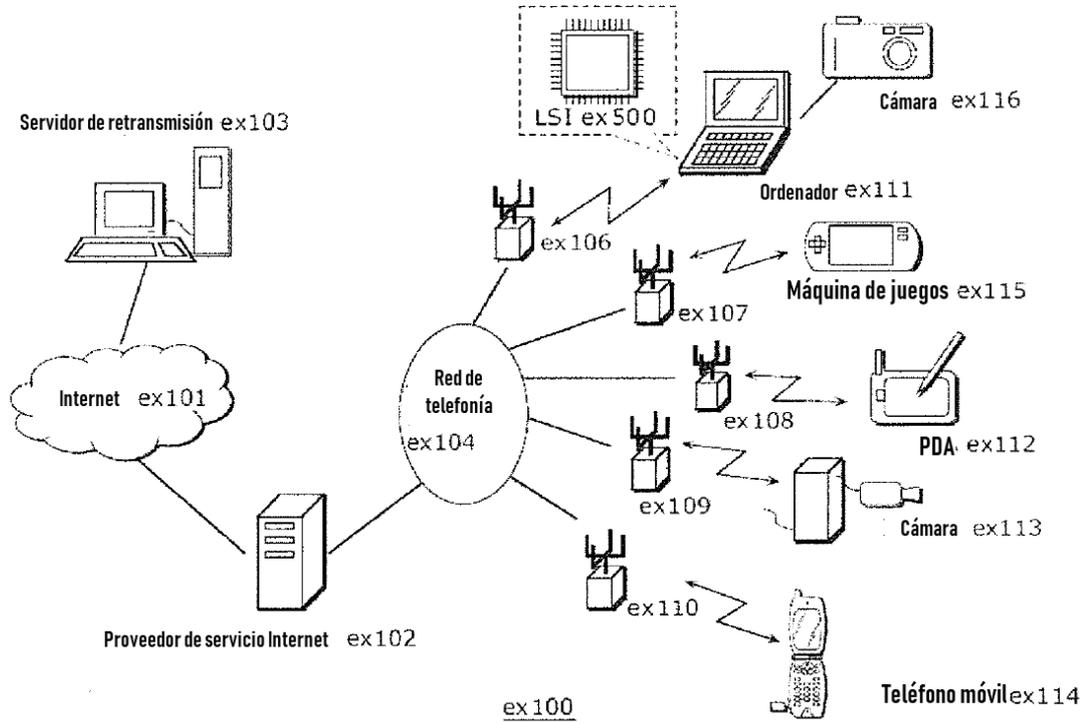
[Fig. 11]



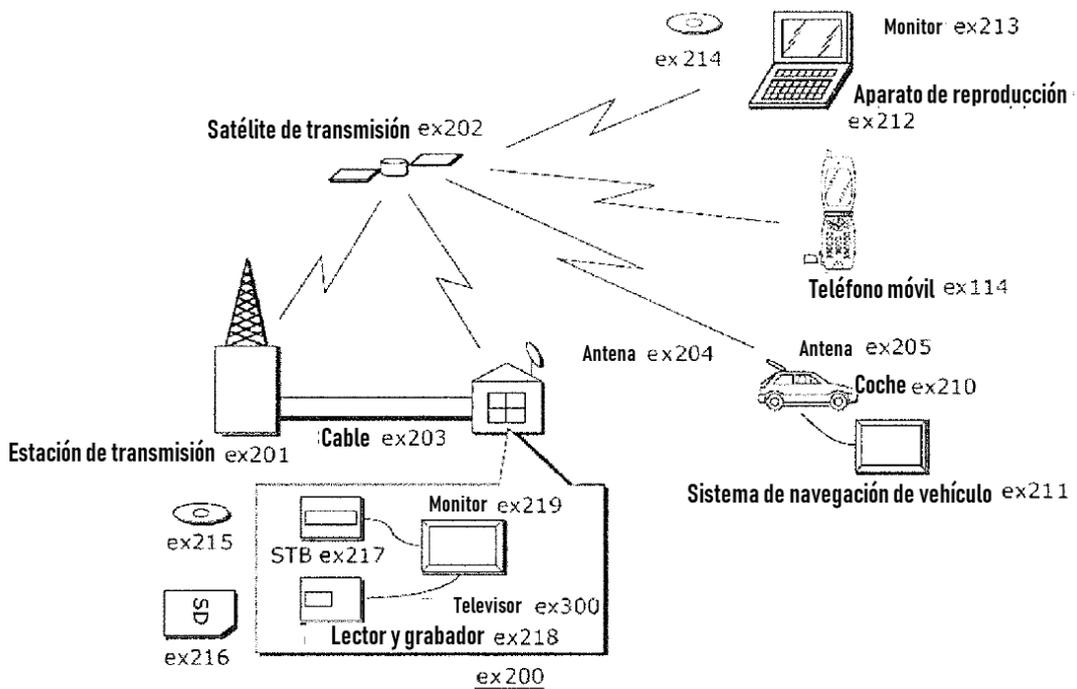
[Fig. 12]



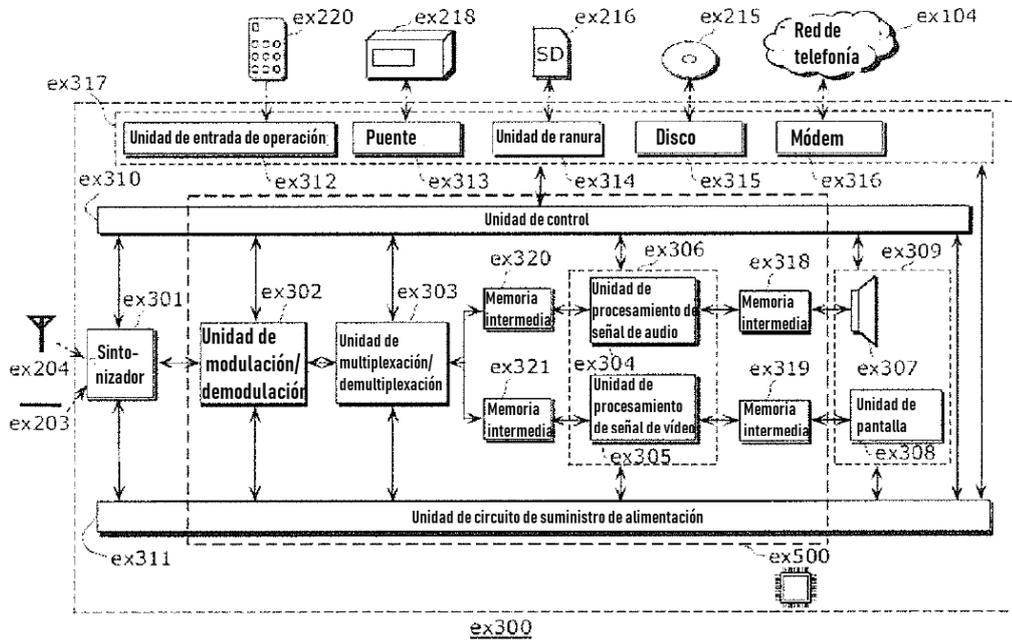
[Fig. 13]



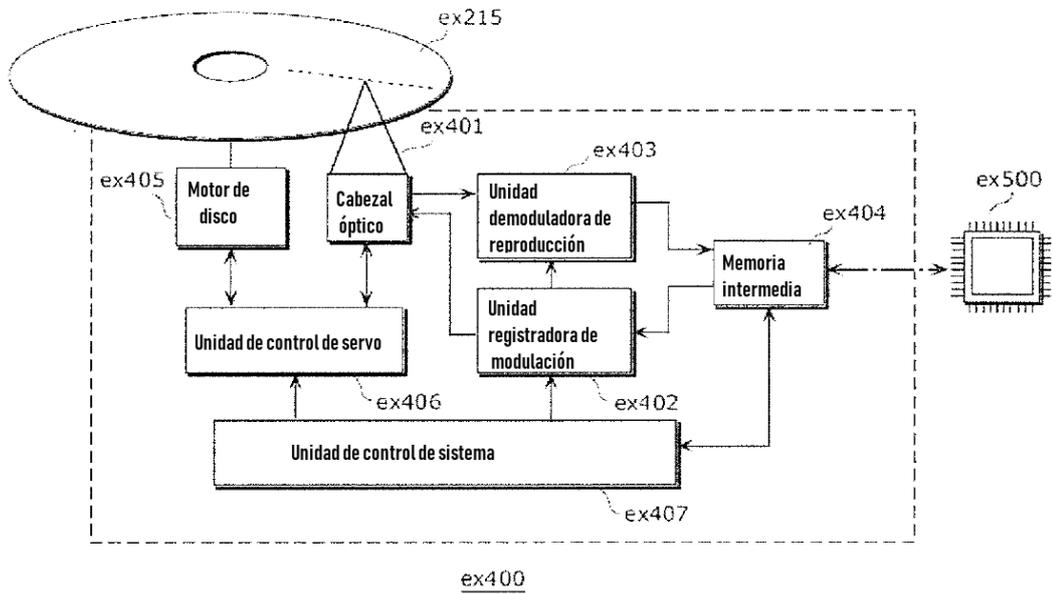
[Fig. 14]

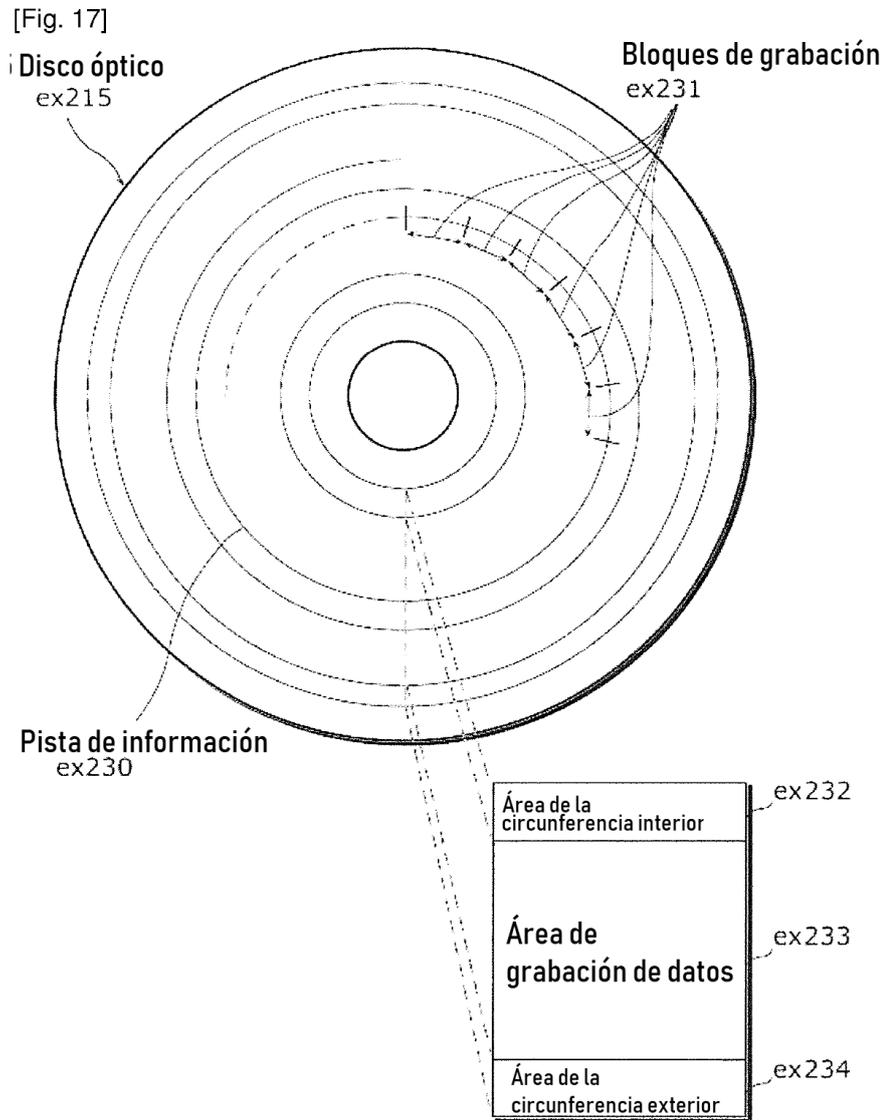


[Fig. 15]

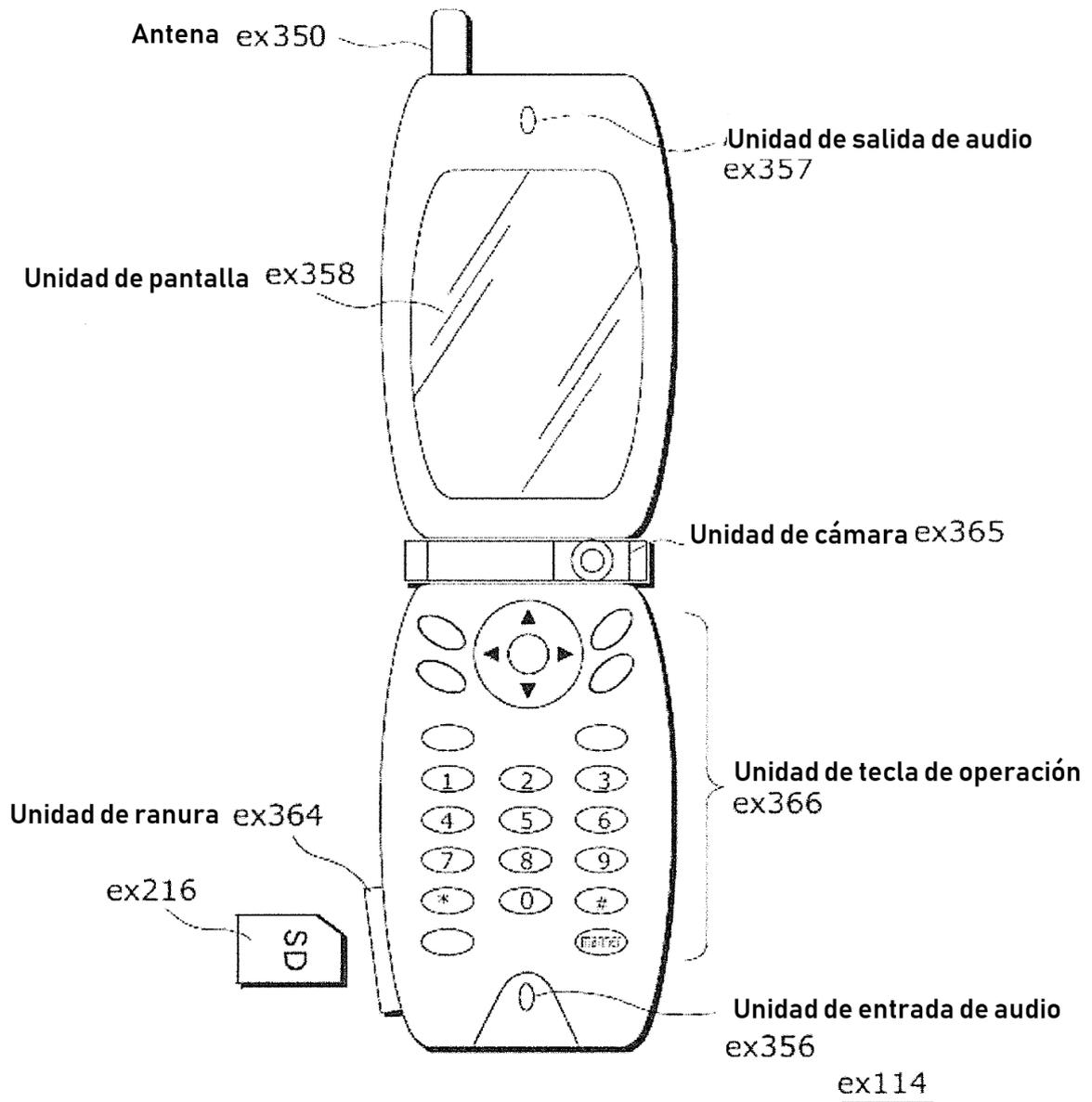


[Fig. 16]

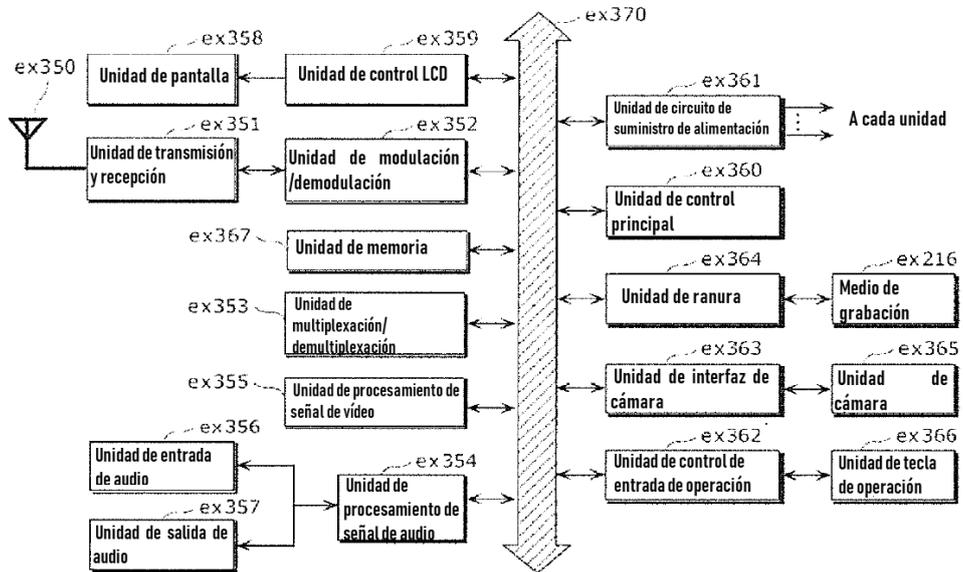




[Fig. 18A]



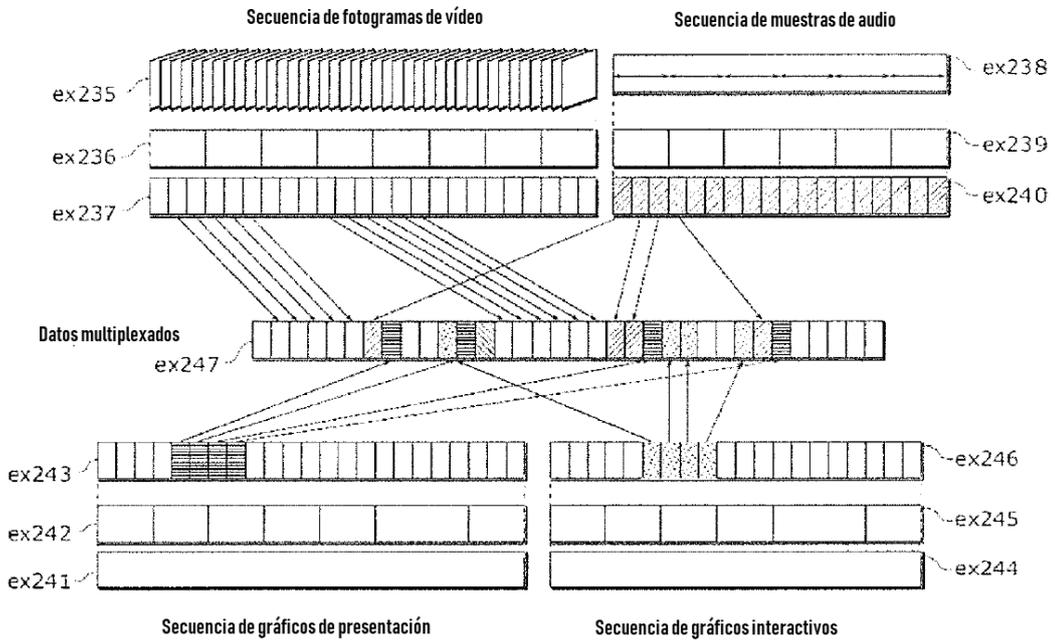
[Fig. 18B]



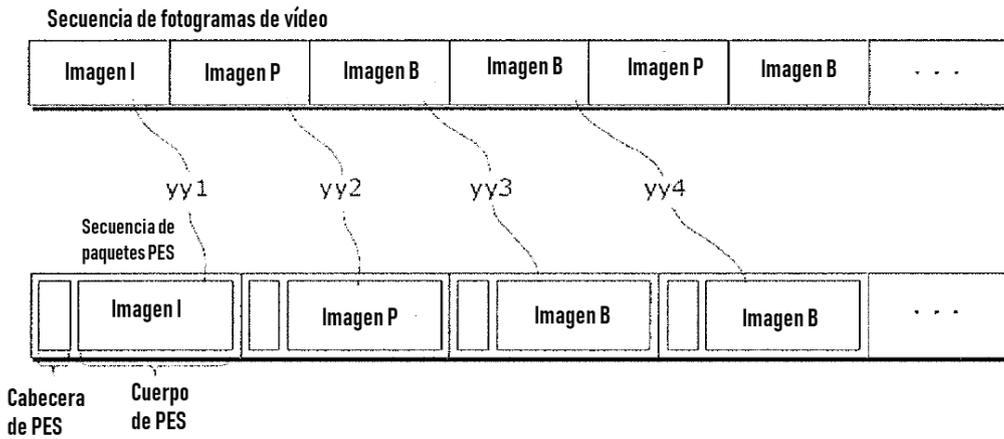
[Fig. 19]

Secuencia de vídeo (PID = 0x1011, vídeo primario)
Secuencia de audio (PID = 0x1100)
Secuencia de audio (PID = 0x1101)
Secuencia de gráficos de presentación (PID = 0x1200)
Secuencia de gráficos de presentación (PID = 0x1201)
Secuencia de gráficos de interactivos (PID = 0x1400)
Secuencia de vídeo (PID = 0x1B00, vídeo secundario)
Secuencia de vídeo (PID = 0x1B01, vídeo secundario)

[Fig. 20]

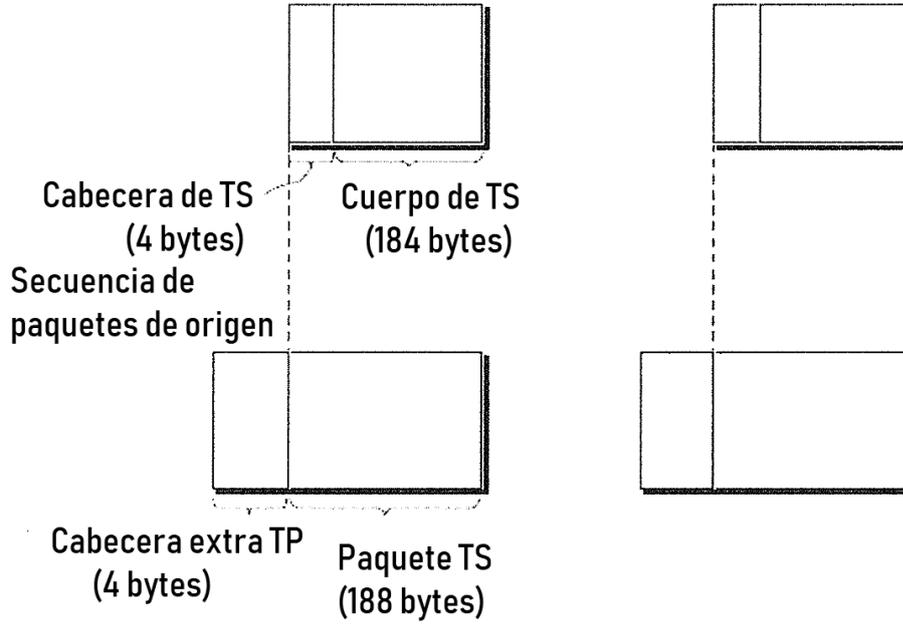


[Fig. 21]



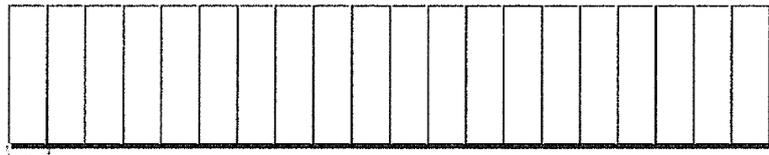
[Fig. 22]

Secuencia de paquetes TS



Datos multiplexados

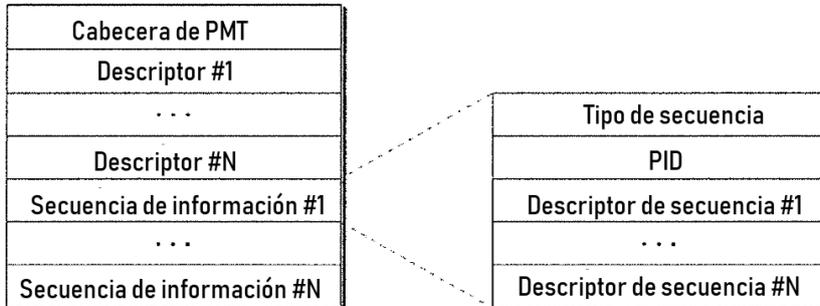
SPN 0 1 2 3 4 5 6 7 ...



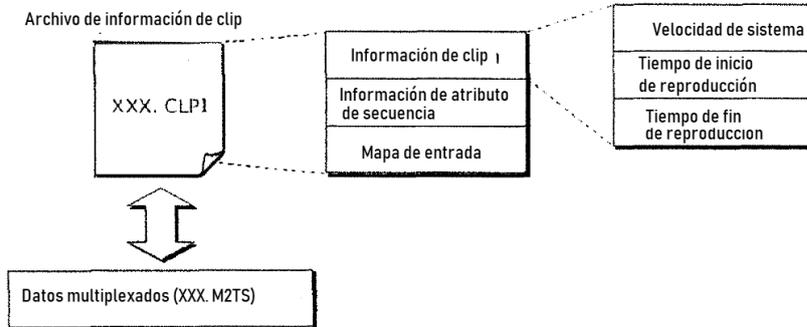
Paquete de origen

[Fig. 23]

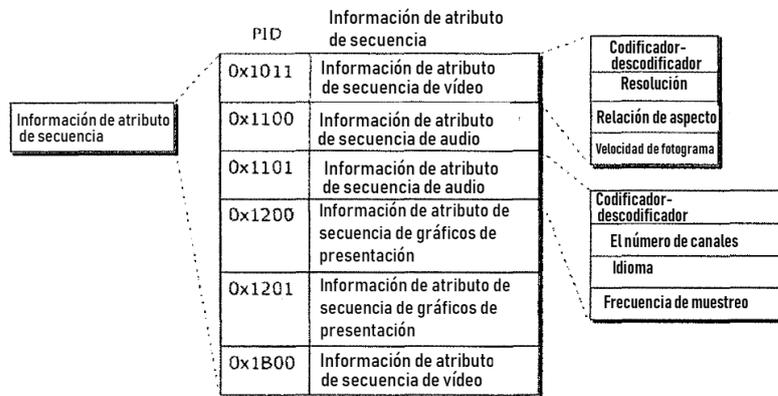
Estructura de datos de PMT



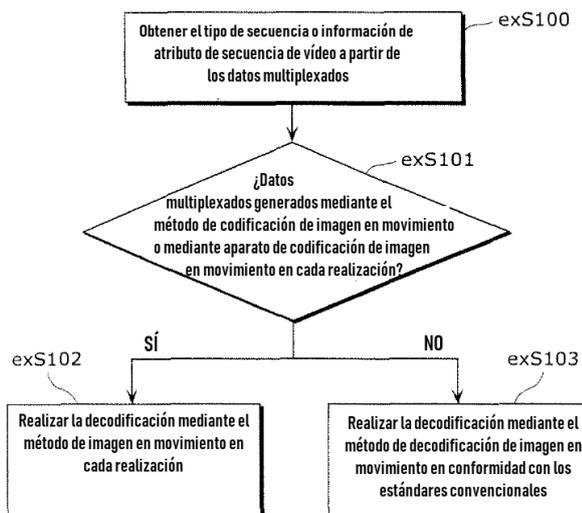
[Fig. 24]



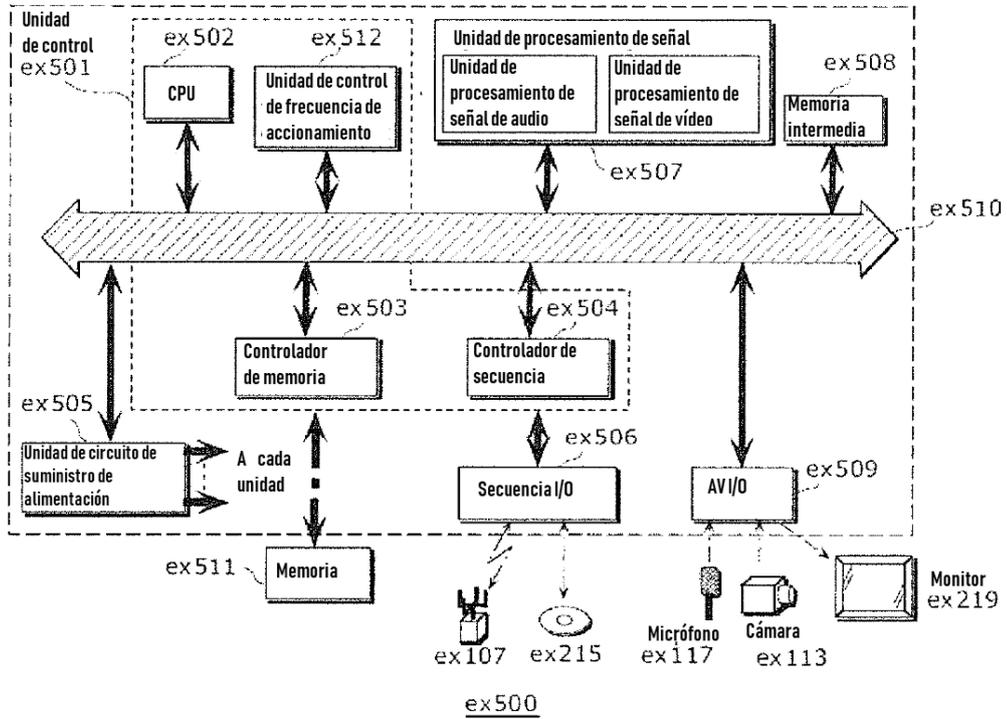
[Fig. 25]



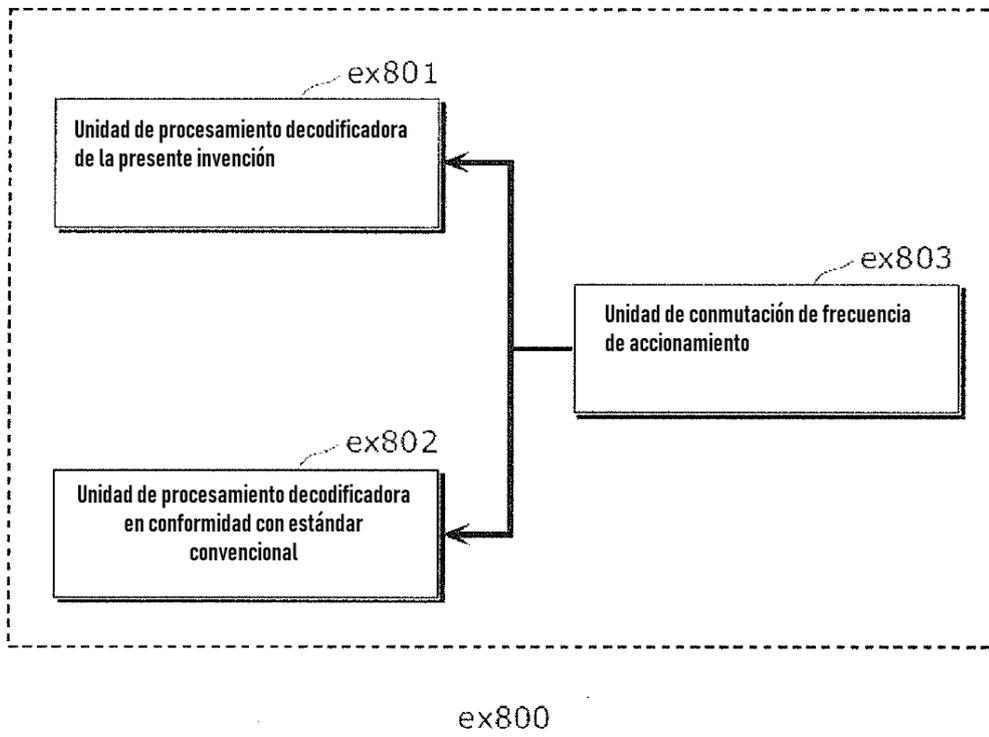
[FIG. 26]



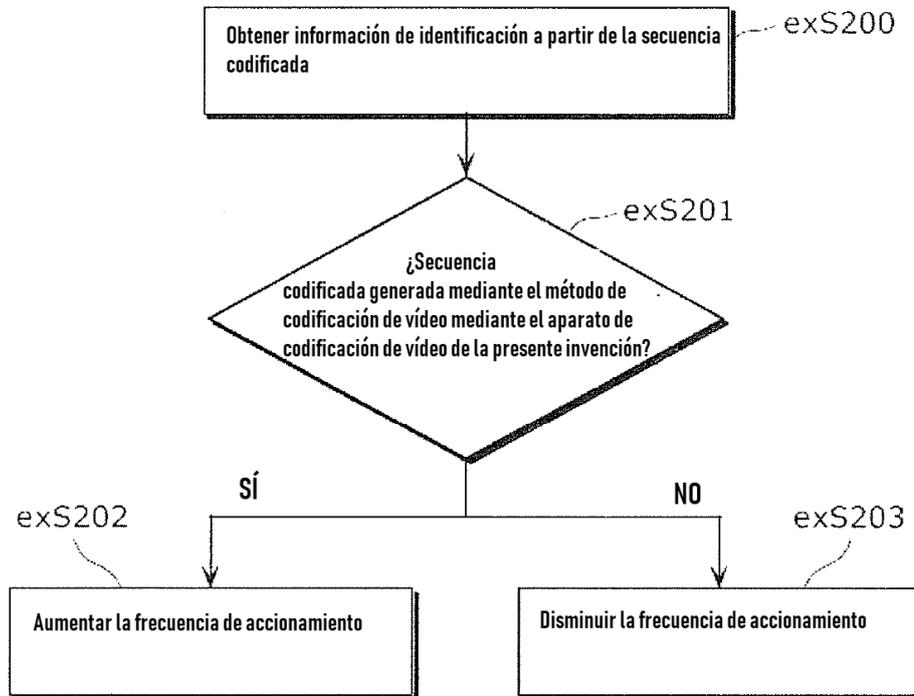
[Fig. 27]



[Fig. 28]



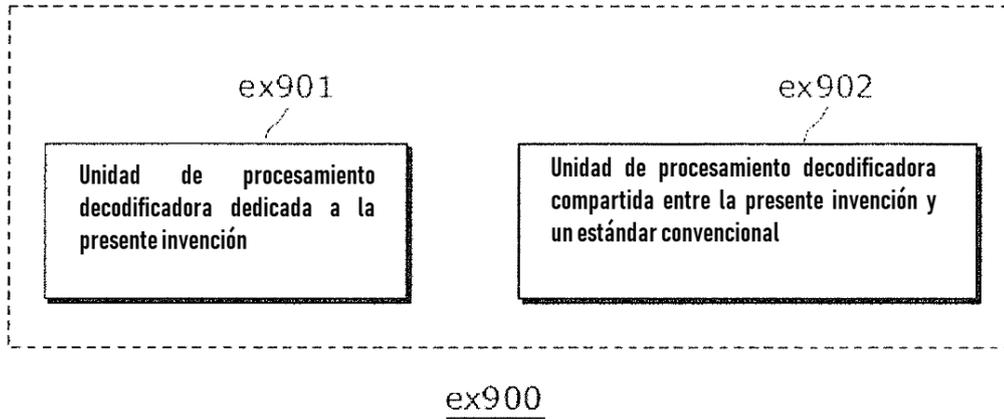
[Fig. 29]



[Fig. 30]

Estándar correspondiente	Frecuencia de accionamiento
MPEG-4 AVC	500 MHz
MPEG-2	350 MHz
⋮	⋮

[Fig. 31A]



[Fig. 31B]

