

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 638 295**

51 Int. Cl.:

**H04N 19/109** (2014.01)  
**H04N 19/147** (2014.01)  
**H04N 19/176** (2014.01)  
**H04N 19/513** (2014.01)  
**H04N 19/52** (2014.01)  
**H04N 19/56** (2014.01)  
**H04N 19/573** (2014.01)  
**H04N 19/577** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **14.01.2003 E 10016061 (3)**

97 Fecha y número de publicación de la concesión europea: **31.05.2017 EP 2323402**

54 Título: **Procedimiento y aparato de codificación de vídeo mejorado**

30 Prioridad:

**25.01.2002 US 352127 P**  
**26.04.2002 US 376005 P**  
**27.06.2002 US 186284**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**19.10.2017**

73 Titular/es:

**MICROSOFT TECHNOLOGY LICENSING, LLC**  
**(100.0%)**  
**One Microsoft Way**  
**Redmond, WA 98052, US**

72 Inventor/es:

**TOURAPIS, ALEXANDROS;**  
**WU, FENG y**  
**LI, SHIPENG**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

**ES 2 638 295 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Procedimiento y aparato de codificación de vídeo mejorado

### Campo técnico

5 Esta invención se refiere a la codificación de vídeo, y más específicamente a procedimientos y aparatos para proporcionar técnicas de codificación y/o de predicción mejorados asociados con diferentes tipos de datos de vídeo.

### Antecedentes

10 La motivación para una mayor eficacia de codificación en la codificación de vídeo ha llevado a la adopción en el equipo mixto de vídeo (JVT) (un organismo de normalización) de modelos y modos más refinados y complicados que describen la información de movimiento para un macrobloque dado. Estos modelos y modos tienden a aprovechar mejor las ventajas de las redundancias temporales que pueden existir dentro de una secuencia de vídeo. Véase, por ejemplo, el Grupo de expertos en codificación de vídeo (VCEG) de la UIT-T, "VT Coding - (ITU-T H.26L & ISO/IEC JTC1 Standard) - Working Draft Number 2 (WD-2)", ITU-T JVT-B1118, Marzo de 2002; y/o Heiko Schwarz y Thomas Wiegand, "Tree-structured macroblock partition", Doc. VCEG-N17, Diciembre de 2001.

15 En el documento WIEGAND T: "H.26L TEST MODEL LONG-TERM NUMBER 9 (TML-9) DRAFT0", ITU-T SECTOR DE NORMALIZACIÓN DE LAS TELECOMUNICACIONES DE LA UIT, Ginebra, CH, 21 de diciembre de 2001 (21/12/2001), páginas 1, 3-75, xP001086625, desvela la codificación de un fotograma P predictivo actual usando un modo de codificación de una pluralidad de modos de codificación que incluyen un intra modo, un primer inter modo en el que el macrobloque actual se copia del bloque colocado en el fotograma anterior, y un segundo inter modo en el que el macrobloque actual está codificado con una compensación de movimiento en relación con uno de una pluralidad de fotogramas de referencia usando una información de vector de movimiento que está codificada en el flujo de bits, en el que, de acuerdo con el segundo inter modo, la información de vector de movimiento para el macrobloque actual está codificado en el flujo de bits como una información de vector de movimiento diferencial. Los modelos recientes incluyen, por ejemplo, la indexación de múltiples fotogramas de los vectores de movimiento, el aumento de la precisión del subpíxel, el multi-referenciamiento y el macrobloque estructurado en árbol y la asignación de movimiento, de acuerdo con el que diferentes subáreas de un macrobloque están asignadas a diferentes informaciones de movimiento. Desafortunadamente, estos modelos tienden también a aumentar significativamente el porcentaje de bits necesario para la codificación de la información de movimiento dentro de la secuencia. De este modo, en algunos casos los modelos tienden a reducir la eficacia de tales procedimientos de codificación.

30 A pesar de que, en algunos casos, los vectores de movimiento se codifican diferencialmente frente a un predictor espacial, o incluso se omiten en el caso de movimiento cero mientras que no tienen una imagen residual para transmitir, esto no parece ser suficiente para mejorar la eficacia.

35 Por lo tanto, sería ventajoso reducir aún más los bits necesarios para la codificación de la información de movimiento, y por lo tanto de toda la secuencia, mientras que al mismo tiempo no afectan significativamente la calidad.

40 Otro problema que también se introduce por la adopción de tales modelos y modos es el de determinar el mejor modo de entre todas las opciones posibles, por ejemplo, dada una tasa de bits objetivo, unos parámetros de codificación/cuantificación, etc. En la actualidad, este problema puede resolverse parcialmente por el uso de medidas/penalizaciones de coste en función del modo y/o de la cuantificación a usar, o incluso empleando técnicas de optimización de tasa distorsión con el objetivo de minimizar una función lagrangiana.

Tales problemas y otros se vuelven aún más significativos, sin embargo, en el caso de los fotogramas bidireccionalmente predictivos (B) donde puede predecirse un macrobloque a partir de tanto los fotogramas futuros como pasados. Esto significa esencialmente que puede necesitarse un porcentaje aún mayor de bits para la codificación de los vectores de movimiento.

45 Por lo tanto, existe una necesidad de un procedimiento y aparatos mejorados para su uso en la codificación de datos de vídeo (por ejemplo, codificación y/o decodificación).

### Sumario

50 Se proporcionan procedimientos y aparatos de codificación de vídeo que hacen uso de diversos modelos y/o modos para mejorar significativamente la eficacia de codificación especialmente para secuencias de movimiento altas/complejas. Los procedimientos y aparatos se aprovechan de las correlaciones temporales y/o espaciales que pueden existir dentro de las partes de los fotogramas, por ejemplo, a nivel del macrobloque, etc. Los procedimientos y aparatos tienden a reducir significativamente la cantidad de datos necesarios para codificar una información de movimiento mientras que se mantiene o incluso se mejora la calidad de imagen de vídeo.

Por lo tanto, por medio de un ejemplo, de acuerdo con ciertas implementaciones de la presente invención, se

- proporciona un procedimiento para su uso en la codificación de datos de vídeo dentro de una secuencia de fotogramas de vídeo. El procedimiento incluye codificar al menos una parte de un fotograma de referencia para incluir una información de movimiento asociada con la parte del fotograma de referencia. El procedimiento incluye además definir al menos una parte de al menos un fotograma predecible que incluye unos datos de vídeo correlacionados predictivamente con la parte del fotograma de referencia basándose en la información de movimiento y codificar al menos la parte del fotograma predecible sin incluir una información de movimiento correspondiente, pero incluyendo unos datos de identificación de modo que identifiquen que la parte del fotograma predecible puede obtenerse directamente usando la información de movimiento asociada con la parte del fotograma de referencia.
- 5
- 10 También se proporciona un aparato para su uso en la codificación de datos de vídeo para una secuencia de fotogramas de vídeo en una pluralidad de fotogramas de vídeo que incluye al menos un fotograma predecible. En este caso, por ejemplo, el aparato incluye una memoria y una lógica, en el que la lógica está configurada para codificar al menos una parte de al menos un fotograma de referencia para incluir una información de movimiento asociada con la parte del fotograma de referencia. La lógica determina también al menos una parte de al menos un
- 15 fotograma predecible que incluye unos datos de vídeo correlacionados de manera predictiva con la parte del fotograma de referencia basándose en la información de movimiento y codifica al menos la parte del fotograma predecible de tal manera que se proporcionan los datos de identificación de modo para especificar que la parte del fotograma predecible puede obtenerse usando la información de movimiento asociada con la parte del fotograma de referencia.
- 20 De acuerdo con todavía otras implementaciones a modo de ejemplo, se proporciona un procedimiento para su uso en la decodificación de los datos de vídeo codificados que incluye al menos un fotograma de vídeo predecible. El procedimiento incluye determinar la información de movimiento asociada con al menos una parte de al menos un fotograma de referencia y almacenar en la memoria intermedia la información de movimiento. El procedimiento también incluye determinar unos datos de identificación de modo que identifiquen que al menos una parte de un
- 25 fotograma predecible puede obtenerse directamente usando al menos la información de movimiento almacenada en una memoria intermedia, y generar la parte del fotograma predecible usando la información de movimiento almacenada en una memoria intermedia.
- También se proporciona un aparato para decodificar datos de vídeo. El aparato incluye una memoria y una lógica, en el que la lógica está configurada para almacenar en la memoria intermedia la información de movimiento asociada con al menos una parte de al menos un fotograma de referencia, comprobar que los datos de identificación de modo identifican que al menos una parte de un fotograma predecible puede obtenerse directamente usando al menos la información de movimiento almacenada en una memoria intermedia, y generar la parte del fotograma predecible usando la información de movimiento almacenada en una memoria intermedia .
- 30

### **Breve descripción de los dibujos**

- 35 La presente invención se ilustra a modo de ejemplo y no de limitación en las figuras de los dibujos adjuntos. Los mismos números se usan a lo largo de las figuras para referirse como componentes y/o características.
- La figura 1 es un diagrama de bloques que representa un entorno de cálculo a modo de ejemplo que es adecuado para su uso con ciertas implementaciones de la presente invención.
- 40 La figura 2 es un diagrama de bloques que representa un dispositivo representativo a modo de ejemplo que es adecuado para su uso con ciertas implementaciones de la presente invención.
- La figura 3 es un diagrama ilustrativo que representa una técnica de proyección de movimiento directo adecuada para su uso en la codificación de fotograma B, de acuerdo con ciertas implementaciones proporcionadas a modo de ejemplo.
- 45 La figura 4 es un diagrama ilustrativo que representa unas técnicas de codificación P y B directa dentro de una secuencia de fotogramas de vídeo, de acuerdo con ciertas implementaciones proporcionadas a modo de ejemplo.
- La figura 5 es un diagrama ilustrativo que representa una predicción de movimiento directo para macrobloques colocados que tienen información de movimiento idéntica, de acuerdo con ciertas implementaciones a modo de ejemplo de la presente invención.
- 50 La figura 6 es un diagrama ilustrativo que representa una técnica de proyección de movimiento directo adecuada para su uso en la codificación de fotograma P, de acuerdo con ciertas implementaciones proporcionadas a modo de ejemplo.
- La figura 7 es un diagrama ilustrativo que representa el uso de la información de aceleración en la proyección de movimiento directo, de acuerdo con ciertas implementaciones proporcionadas a modo de ejemplo.
- 55 La figura 8 es un diagrama ilustrativo que representa una técnica de proyección de píxeles directa adecuada para

su uso en la codificación de fotograma B, de acuerdo con ciertas implementaciones proporcionadas a modo de ejemplo.

5 La figura 9 es un diagrama ilustrativo que representa una técnica de proyección de píxeles directa adecuada para su uso en la codificación de fotogramas P, de acuerdo con ciertas implementaciones proporcionadas a modo de ejemplo.

La figura 10 es un diagrama de bloques que representa un codificador de vídeo convencional a modo de ejemplo.

La figura 11 es un diagrama de bloques que representa un decodificador de vídeo convencional a modo de ejemplo.

10 La figura 12 es un diagrama de bloques que representa un codificador de vídeo mejorado a modo de ejemplo que usa la predicción directa, de acuerdo con ciertas implementaciones a modo de ejemplo de la presente invención.

La figura 13 es un diagrama de bloques que representa un decodificador de vídeo mejorado a modo de ejemplo que usa la predicción directa, de acuerdo con ciertas implementaciones a modo de ejemplo de la presente invención.

15 La figura 14 es un diagrama ilustrativo que representa una técnica de proyección de píxeles/bloques directa, de acuerdo con ciertas implementaciones proporcionadas a modo de ejemplo.

La figura 15 es un diagrama ilustrativo que representa una técnica de proyección de movimiento directo adecuada para su uso en la codificación de fotograma B, de acuerdo con ciertas implementaciones proporcionadas a modo de ejemplo.

20 La figura 16 es un diagrama ilustrativo que representa unas predicciones de vector de movimiento, de acuerdo con ciertas implementaciones proporcionadas a modo de ejemplo.

La figura 17 es un diagrama ilustrativo que representa unas técnicas de codificación entrelazadas para los fotogramas P, de acuerdo con ciertas implementaciones proporcionadas a modo de ejemplo.

25 La figura 18 es un diagrama ilustrativo que representa unas técnicas de codificación entrelazadas para los fotogramas B, de acuerdo con ciertas implementaciones proporcionadas a modo de ejemplo.

La figura 19 es un diagrama ilustrativo que representa un esquema para codificar imágenes de campo/fotograma mixtas, de acuerdo con ciertas implementaciones proporcionadas a modo de ejemplo.

### **Descripción detallada**

30 De acuerdo con ciertos aspectos de la presente invención, se proporcionan procedimientos y aparatos para la codificación de datos de vídeo (por ejemplo, codificación y/o decodificación). Los procedimientos y aparatos pueden configurarse para mejorar la eficacia de codificación de las tecnologías de transmisión de codificación de vídeo progresivo o "entrelazadas". En ciertas implementaciones, por ejemplo, con respecto a la norma H.26L actual, los llamados "fotogramas P" se han mejorado significativamente introduciendo varios modos de macrobloque adicionales. En algunos casos, ahora puede ser necesario transmitir hasta 16 vectores de movimiento por macrobloque. Ciertos aspectos de la presente invención proporcionan una forma de codificar estos vectores de movimiento. Por ejemplo, tal como se describe a continuación, pueden usarse técnicas de predicción P directas para seleccionar los vectores de movimiento de los píxeles colocados en el fotograma anterior.

40 Si bien se describen estos y otros procedimientos y aparatos a modo de ejemplo, debería tenerse en cuenta que las técnicas de la presente invención no se limitan a los ejemplos descritos y representados en los dibujos adjuntos, sino que también pueden adaptarse claramente a otros esquemas de codificación de vídeo similares existentes y futuros, etc.

Antes de la introducción de tales procedimientos y aparatos a modo de ejemplo, se proporciona una introducción en la siguiente sección para entornos operativos a modo de ejemplo adecuados, por ejemplo, en la forma de un dispositivo informático y otros tipos de dispositivos/aparatos.

### 45 **Entornos operacionales a modo de ejemplo:**

Volviendo a los dibujos, en los que los números de referencia similares se refieren a elementos similares, la invención se ilustra como implementada en un entorno informático adecuado. Aunque no se requiere, la invención se describirá en el contexto general de las instrucciones ejecutables por ordenador, tales como unos módulos de programa, que se ejecutan por un ordenador personal.

50 En general, los módulos de programa incluyen rutinas, programas, objetos, componentes, estructuras de datos, etc., que realizan tareas específicas o implementan tipos de datos abstractos específicos. Los expertos en la materia

apreciarán que la invención puede practicarse con otras configuraciones de sistema informático, que incluyen dispositivos portátiles, sistemas multiprocesador, electrónica de consumo programable o basada en microprocesadores, PC en red, minicomputadoras, ordenadores de sistema central, dispositivos de comunicación portátiles y similares.

- 5 La invención también puede ponerse en práctica en entornos informáticos distribuidos donde las tareas se realizan por dispositivos de procesamiento remoto que están enlazados a través de una red de comunicaciones. En un entorno informático distribuido, los módulos de programa pueden estar localizados tanto en dispositivos de almacenamiento de memoria local como remota.

10 La figura 1 ilustra un ejemplo de un entorno 120 informático adecuado en el que pueden implementarse los sistemas, aparatos y procedimientos descritos a continuación. El entorno 120 informático a modo de ejemplo es solo un ejemplo de un entorno informático adecuado y no pretende sugerir ninguna limitación en cuanto al ámbito de uso o funcionalidad de los procedimientos y sistemas mejorados descritos en el presente documento. Tampoco debería interpretarse el entorno 120 informático como si tuviera cualquier dependencia o requisito relacionado con uno cualquiera o una combinación de componentes ilustrados en el entorno 120 informático.

15 Los procedimientos y sistemas mejorados en el presente documento pueden funcionar con numerosos otros entornos o configuraciones de sistema informático de fin general o de fin especial. Ejemplos de sistemas informáticos, entornos y/o configuraciones bien conocidos que pueden ser adecuados incluyen, pero no se limitan a, ordenadores personales, ordenadores servidores, clientes ligeros, clientes pesados, dispositivos de mano o portátiles, sistemas multiprocesador, sistemas basados en microprocesador, decodificadores, electrónica de  
20 consumo programable, PC en red, minicomputadoras, ordenadores de sistema central, entornos informáticos distribuidos que incluyen cualquiera de los sistemas o dispositivos anteriores, y similares.

Como se muestra en la figura 1, el entorno 120 informático incluye un dispositivo informático de fin general en forma de un ordenador 130. Los componentes del ordenador 130 pueden incluir uno o más procesadores o unidades 132  
25 de procesamiento, una memoria 134 de sistema y un bus 136 que acopla diversos componentes de sistema incluyendo la memoria 134 de sistema al procesador 132.

El bus 136 representa uno o más de cualquiera de varios tipos de estructuras de bus, incluyendo un bus de memoria o controlador de memoria, un bus periférico, un puerto de gráficos acelerados, y un bus de procesador o local que usa cualquiera de una variedad de arquitecturas de bus. A modo de ejemplo, y no de limitación, tales arquitecturas  
30 incluyen el bus de arquitectura estándar de la industria (ISA), el bus de arquitectura de micro canal (MCA), el bus ISA mejorado (EISA), el bus local de la asociación de normas electrónicas de vídeo (VESA) y el bus de interconexiones de componentes periféricos (PCI) también conocido como bus Mezzanine.

El ordenador 130 incluye normalmente una variedad de medios legibles por ordenador. Tales medios pueden ser cualquier medio disponible que sea accesible por el ordenador 130, e incluye tanto medios volátiles como no volátiles y tanto medios removibles como no removibles.

35 En la figura 1, la memoria 134 de sistema incluye unos medios legibles por ordenador en la forma de memoria volátil, tal como una memoria 140 de acceso aleatorio (RAM) y/o una memoria no volátil, tal como una memoria 138 de solo lectura (ROM). Un sistema 142 básico de entrada/salida (BIOS), que contiene las rutinas básicas que ayudan a transferir información entre los elementos dentro del ordenador 130, tal como durante el arranque, se almacena en la ROM 138. La RAM 140 contiene normalmente unos módulos de datos y/o un programa que pueden  
40 accederse inmediatamente y/o se operan actualmente por el procesador 132.

El ordenador 130 puede incluir además otro medio de almacenamiento informático extraíble/no extraíble, volátil/no volátil. Por ejemplo, la figura 1 ilustra una unidad 144 de disco duro para leer y escribir en un medio magnético no volátil, no removible (no mostrado y normalmente llamado "disco duro"), una unidad 146 de disco magnético para leer y escribir en un disco 148 magnético extraíble, no volátil (por ejemplo, un "disquete") y una unidad 150 de disco  
45 óptico para leer o escribir en un disco 152 óptico extraíble, no volátil tal como un CD-ROM/R/RW, un DVD-ROM/R/RW/+R/RAM u otros medios ópticos. La unidad 144 de disco duro, la unidad 146 de disco magnético y la unidad 150 de disco óptico están conectadas cada una al bus 136 por una o más interfaces 154.

Las unidades y medios legibles por ordenador asociados proporcionan un almacenamiento no volátil de instrucciones legibles por ordenador, estructuras de datos, módulos de programa y otros datos para el ordenador  
50 130. Aunque el entorno a modo de ejemplo descrito en el presente documento emplea un disco duro, un disco 148 magnético extraíble y un disco 152 óptico extraíble, debería apreciarse por los expertos en la materia que otros tipos de medios legibles por ordenador que pueden almacenar datos accesibles por un ordenador, tales como cintas magnéticas, tarjetas de memoria flash, discos de vídeo digitales, memorias de acceso aleatorio (RAM), memorias de solo lectura (ROM), y similares, también pueden usarse en el entorno operativo a modo de ejemplo.

55 Un número de módulos de programa pueden almacenarse en el disco duro, el disco 148 magnético, el disco 152 óptico, la ROM 138, o la RAM 140, incluyendo, por ejemplo, un sistema 158 operativo, uno o más programas 160 de aplicación, otros módulos 162 de programa, y los datos 164 de programa.

Los procedimientos y sistemas mejorados descritos en el presente documento pueden implementarse dentro del sistema 158 operativo, uno o más programas 160 de aplicación, otros módulos 162 de programa, y/o los datos 164 de programa.

5 Un usuario puede proporcionar comandos e información al ordenador 130 a través de unos dispositivos de entrada tales como el teclado 166 y el dispositivo 168 de señalización (tal como un "ratón"). Otros dispositivos de entrada (no mostrados) pueden incluir un micrófono, una palanca de mando, un controlador para juegos, una antena parabólica, un puerto serie, un escáner, una cámara, etc. Estos y otros dispositivos de entrada están conectados a la unidad 10 132 de procesamiento a través de una interfaz 170 de entrada de usuario que está acoplada al bus 136, pero pueden estar conectados por otras estructuras de interfaz y de bus, tales como un puerto paralelo, un puerto de juego o un bus serie universal (USB).

Un monitor 172 u otro tipo de dispositivo de visualización también están conectados al bus 136 a través de una interfaz, tal como un adaptador 174 de vídeo. Además del monitor 172, los ordenadores personales incluyen normalmente otros dispositivos de salida periféricos (no mostrados), tales como altavoces e impresoras, que pueden conectarse a través de la interfaz 175 periférica de salida.

15 El ordenador 130 puede funcionar en un entorno de red usando conexiones lógicas a uno o más ordenadores remotos, tal como un ordenador 182 remoto. El ordenador 182 remoto puede incluir muchos o todos los elementos y características descritas en el presente documento con respecto al ordenador 130.

Las conexiones lógicas mostradas en la figura 1 son una red 177 de área local (LAN) y una red 179 de área amplia (WAN). Tales entornos de red son comunes en oficinas, redes informáticas de toda la empresa, intranets e Internet.

20 Cuando se usa en un entorno de red LAN, el ordenador 130 está conectado a la LAN 177 a través de una interfaz de red o adaptador 186. Cuando se usa en un entorno de red WAN, el ordenador normalmente incluye un módem 178 u otro medio para establecer las comunicaciones a través de la WAN 179. El módem 178, que puede ser interno o externo, puede estar conectado al bus de sistema 136 a través de la interfaz 170 de entrada de usuario u otro mecanismo apropiado.

25 En la figura 1, se representa una implementación específica de una WAN a través de Internet. En este caso, el ordenador 130 emplea el módem 178 para establecer comunicaciones con al menos un ordenador 182 remoto a través de Internet 180.

30 En un entorno de red, los módulos de programa representados con respecto al ordenador 130, o partes de los mismos, pueden almacenarse en un dispositivo de almacenamiento de memoria remoto. De este modo, por ejemplo, como se representa en la figura 1, los programas 189 de aplicación remota pueden residir en un dispositivo de memoria del ordenador 182 remoto. Se apreciará que las conexiones de red mostradas y descritas son a modo de ejemplo y pueden usarse otros medios para establecer un enlace de comunicaciones entre los ordenadores.

35 A continuación, se llama la atención sobre la figura 2, que es un diagrama de bloques que representa otro dispositivo 200 a modo de ejemplo que también es capaz de beneficiarse de los procedimientos y aparatos desvelados en el presente documento. El dispositivo 200 es representativo de uno cualquiera o más dispositivos o aparatos que están configurados operativamente para procesar vídeo y/o cualquier tipo de datos relacionado de acuerdo con todos o parte de los procedimientos y aparatos descritos en el presente documento y sus equivalentes. De este modo, el dispositivo 200 puede adoptar la forma de un dispositivo informático como el de la figura 1, o de alguna otra forma, tal como, por ejemplo, un dispositivo inalámbrico, un dispositivo de comunicación portátil, un asistente digital personal, un reproductor de vídeo, un reproductor de DVD, un reproductor de CD, una máquina de karaoke, un quiosco, un proyector de vídeo digital, un mecanismo de pantalla de vídeo de panel plano, un decodificador, una máquina de videojuegos, etc. En este ejemplo, el dispositivo 200 incluye la lógica 202 configurada para procesar datos de vídeo, una fuente 204 de datos de vídeo configurada para proporcionar datos de vídeo a la lógica 202 y al menos un módulo 206 de visualización capaz de mostrar al menos una parte de los datos de vídeo para que los vea 40 un usuario. La lógica 202 es representativa de un hardware, firmware, software y/o cualquier combinación de los mismos. En ciertas implementaciones, por ejemplo, la lógica 202 incluye un compresor/descompresor (codec), o similar. La fuente 204 de datos de vídeo es representativa de cualquier mecanismo que pueda proporcionar, comunicar, emitir y/o al menos almacenar momentáneamente datos de vídeo adecuados para su procesamiento por la lógica 202. La fuente de reproducción de vídeo se muestra ilustrativamente como que está dentro y/o fuera del dispositivo 200. El módulo 206 de visualización es representativo de cualquier mecanismo que un usuario pueda ver 45 directa o indirectamente y ver los resultados visuales de los datos de vídeo presentados en el mismo. Adicionalmente, en ciertas implementaciones, el dispositivo 200 también puede incluir alguna forma o capacidad para reproducir o manejar de otra manera datos de audio asociados con los datos de vídeo. Por lo tanto, se muestra un módulo 208 de reproducción de audio.

55 Con los ejemplos de las figuras 1 y 2 en mente, y otros como los mismos, las siguientes secciones se centran en ciertos procedimientos y aparatos a modo de ejemplo que pueden practicarse al menos parcialmente usando tales entornos y tales dispositivos.

Predicción directa para fotogramas predictivos (P) y bidireccionalmente predictivos (B) en codificación de vídeo:

Esta sección presenta un nuevo tipo de inter macrobloque altamente eficiente que puede mejorar significativamente la eficacia de codificación especialmente para secuencias de movimiento altas/complejas. Este nuevo tipo de inter macrobloque aprovecha las correlaciones temporales y espaciales que pueden existir dentro de los fotogramas en el nivel de macrobloque, y como resultado puede reducir significativamente los bits necesarios para codificar la información de movimiento al tiempo que mantiene o incluso mejora la calidad.

Predicción Directa

Los problemas anteriormente mencionados y/u otros se resuelven al menos parcialmente en el presente documento mediante la introducción de un "modo de predicción directa", en el que, en lugar de codificar la información de movimiento actual, los dos vectores de movimiento hacia delante y/o hacia atrás se obtienen directamente de los vectores de movimiento usados en el macrobloque correlacionado del fotograma de referencia posterior.

Esto se ilustra, por ejemplo, en la figura 3, que muestra tres fotogramas de vídeo, a saber, un fotograma 300 P, un fotograma 302 B y un fotograma 304 P, que corresponden respectivamente a los tiempos  $t$ ,  $t+1$ , y  $t+2$ . En la figura 3 También se ilustran unos macrobloques dentro de los fotogramas 300, 302 y 304 y una información de vector de movimiento (MV) a modo de ejemplo. En el presente documento, los fotogramas tienen asociadas con los mismos unas coordenadas  $x$  e  $y$ . La información de vector de movimiento para el fotograma 302 B se predice (en este caso, por ejemplo, interpolada) a partir de la información de vector de movimiento codificada para los fotogramas 300 y 304 P. La técnica a modo de ejemplo se obtiene de la suposición de que un objeto se mueve con velocidad constante, y por lo tanto es posible predecir su posición actual dentro del fotograma 302 B sin tener que transmitir ningún vector de movimiento. Si bien esta técnica puede reducir la tasa de flujo significativamente para una calidad dada, no siempre puede aplicarse.

Se introduce en el presente documento, de acuerdo con ciertas implementaciones proporcionadas, un nuevo tipo de inter macrobloque siempre que pueda explotar eficazmente las correlaciones espaciales y temporales que puedan existir en el nivel de macrobloque y en particular con respecto a la información de vector de movimiento del macrobloque. De acuerdo con este nuevo modo, es posible que un macrobloque actual pueda tener un movimiento que pueda obtenerse directamente de la información decodificada anteriormente (por ejemplo, una proyección de movimiento). Por lo tanto, como se muestra ilustrativamente en la figura 4, puede no ser necesaria la transmisión de vectores de movimiento para un macrobloque, sino incluso para un fotograma entero. En este caso, se representa una secuencia 400 de fotogramas de vídeo con flechas sólidas que indican relaciones codificadas entre fotogramas y con líneas discontinuas que indican relaciones de macrobloque predecibles. El fotograma 402 de vídeo es un fotograma I, los fotogramas 404, 406, 410 y 412 de vídeo son fotogramas B, y los fotogramas 408 y 414 de vídeo son fotogramas P. En este ejemplo, si el fotograma 408 P tiene un campo de movimiento descrito por  $\overrightarrow{MF}_{406}$ , el movimiento de los macrobloques colocados en las imágenes 404, 406 y 414 también está altamente correlacionado. En particular, suponiendo que la velocidad es en general constante en todo el fotograma y que los fotogramas 404 y 406 están igualmente separados en el tiempo entre los fotogramas 402 y 408, y considerando también que para los fotogramas B pueden usarse tanto vectores de movimiento hacia delante como hacia atrás, los campos de movimiento en el fotograma 404 podrían ser iguales a  $\overrightarrow{MF}_{404}^{fw} = \frac{1}{3} \times \overrightarrow{MF}_{406}$  y  $\overrightarrow{MF}_{404}^{bw} = -\frac{1}{3} \times \overrightarrow{MF}_{406}$  para los campos de movimiento hacia delante y hacia atrás respectivamente. Del mismo modo, para el fotograma 408 los campos de movimiento podrían ser  $\overrightarrow{MF}_{408}^{fw} = \frac{2}{3} \times \overrightarrow{MF}_{406}$  y  $\overrightarrow{MF}_{408}^{bw} = -\frac{2}{3} \times \overrightarrow{MF}_{406}$  para los vectores de movimiento hacia delante y hacia atrás respectivamente. Ya que 414 y 406 están separados igualmente, entonces, usando la misma suposición, el macrobloque colocado podría tener unos vectores de movimiento  $\overrightarrow{MF}_{414} = \overrightarrow{MF}_{406}$ .

Al igual que en el modo directo en los fotogramas B, suponiendo de nuevo que la velocidad es constante, el movimiento para un macrobloque puede obtenerse directamente del macrobloque correlacionado del fotograma de referencia. Esto se ilustra adicionalmente en la figura 6, por ejemplo, que muestra tres fotogramas de vídeo, a saber, un fotograma 600 P, un fotograma 602 B y un fotograma 604 P, que corresponden respectivamente a los tiempos  $t$ ,  $t+1$ , y  $t+2$ . En este caso, los macrobloques colocados ilustrados tienen una información de movimiento similar, si no idéntica.

Incluso es posible considerar la aceleración para refinar dichos parámetros de movimiento, por ejemplo, véase la figura 7. En este caso, por ejemplo, se muestran tres fotogramas, a saber, un fotograma 704 actual en el tiempo  $t$ , y los fotogramas 702 (tiempo  $t-1$ ) y 700 (tiempo  $t-2$ ) anteriores, con diferentes informaciones de aceleración ilustradas por diferentes vectores de movimiento de longitud.

El procedimiento también puede mejorarse significativamente, en lugar de considerar la proyección de movimiento a nivel de macrobloque, teniendo en cuenta que los píxeles en el interior de la imagen anterior se mueven posiblemente con una velocidad constante o una aceleración constante (por ejemplo, la proyección de píxeles). Como tal, puede generar una predicción significativamente más precisa del fotograma actual para la codificación de fotogramas B, como se ilustra, por ejemplo, en la figura 8, y para la codificación de fotogramas P como se ilustra, por ejemplo, en la figura 9. La figura 8 muestra, por ejemplo, tres fotogramas de vídeo, a saber, un fotograma 800 P, un

fotograma 802 B y un fotograma 804 P, que corresponden respectivamente a los tiempos  $t$ ,  $t+1$ , y  $t+2$ . La figura 9 muestra, por ejemplo, tres fotogramas de video, a saber, un fotograma 900 P, un fotograma 902 B y un fotograma 904 P, que corresponden respectivamente a los tiempos  $t$ ,  $t+1$ , y  $t+2$ .

5 En ciertas implementaciones, también es posible combinar entre sí ambos procedimientos para un rendimiento aún mejor.

De acuerdo con ciertas implementaciones adicionales, el movimiento puede obtenerse también a partir de la información espacial, por ejemplo, usando técnicas de predicción empleadas para la codificación de vectores de movimiento a partir de la información de movimiento de los macrobloques circundantes. Además, puede mejorarse también el rendimiento combinando estos dos procedimientos diferentes en una arquitectura de predicción de múltiples hipótesis que no necesita que se transmita la información de movimiento. En consecuencia, tales nuevos tipos de macrobloques pueden lograr unas reducciones significativas de la tasa de bits, al tiempo que se consigue una calidad similar o mejorada.

Procedimientos de codificación a modo de ejemplo:

15 La figura 10 ilustra un entorno 1000 de codificación a modo de ejemplo, que tiene un codificador 1002 de video convencional basado en bloques, en el que se proporcionan datos 1004 de video al codificador 1002 y se emite un flujo de bits de datos de video codificado correspondiente.

Los datos 1004 de video se proporcionan a un módulo 1006 de suma, que también recibe como entrada, la salida de un módulo 1022 de compensación de movimiento (MC). La salida del módulo 1006 de suma se proporciona a un módulo 1010 de transformada de coseno discreta (DCT). La salida del módulo 1010 de DCT se proporciona como una entrada a un módulo 1012 de cuantificación (QP). La salida del módulo 1012 de QP se proporciona como una entrada a un módulo 1014 de cuantificación inversa ( $QP^{-1}$ ) y como una entrada a un módulo 1016 de codificación de longitud variable (VLC). El módulo 1016 de VLC también recibe como una entrada, una salida de un módulo 1008 de estimación de movimiento (ME). La salida del módulo de VLC 1016 es un flujo 1210 de bits de video codificado.

25 La salida del módulo 1014 de  $QP^{-1}$  se proporciona como una entrada a un módulo 1018 de transformada de coseno discreta (DCT). La salida de 1018 se proporciona como una entrada a un módulo 1020 de suma, que tiene como otra entrada, la salida del módulo 1022 de MC. La salida del módulo 1020 de suma se proporciona como una entrada a un módulo 1024 de filtro de bucle. La salida del módulo 1024 de filtro de bucle se proporciona como una entrada a un módulo 1026 de memoria intermedia de fotograma. Una salida del módulo 1026 de memoria intermedia de fotograma se proporciona como una entrada al módulo 1008 de ME y otra salida se proporciona como una entrada al módulo 1022 de MC. El módulo 1008 de ME también recibe como entrada los datos 1004 de video. Una salida del ME 1008 se prueba como una entrada para el módulo 1022 de MC.

35 En este ejemplo, el módulo 1022 de MC recibe unas entradas del módulo 1008 de ME. En este caso, el ME se realiza en un fotograma actual contra un fotograma de referencia. El ME puede realizarse usando diversos tamaños de bloques e intervalos de búsqueda, después de lo cual se codifica y transmite (INTER codificación) un "mejor" parámetro usando, por ejemplo algún criterio predefinido. La información residual también se codifica después de realizar el DCT y el QP. También es posible que en algunos casos el rendimiento del ME no produzca un resultado satisfactorio, y por lo tanto un macrobloque, o incluso un subbloque, podría INTRA codificarse.

40 Considerando que la información de movimiento podría ser bastante costosa, el procedimiento de codificación puede modificarse como en la figura 12, de acuerdo con ciertas implementaciones proporcionadas, a considerar también en un procedimiento adicional la posibilidad de que los vectores de movimiento para un macrobloque podrían predecirse temporal y/o espacialmente a partir de la información de movimiento anteriormente codificada. Tales decisiones, por ejemplo, pueden realizarse usando unas técnicas de optimización de tasa distorsión u otras medidas de coste. El uso de tales técnicas/modos puede que no sea necesario para transmitir una información de movimiento detallada, debido a que esta puede reemplazarse con un modo de predicción directa (P directa), por ejemplo, como se ilustra en la figura 5.

45 El movimiento puede modelarse, por ejemplo, en cualquiera de los siguientes modelos o sus combinaciones: (1) proyección de movimiento (por ejemplo, como se ilustra en la figura 3 para los fotogramas B y en la figura 6 para los fotogramas P); (2) proyección de píxeles (por ejemplo, como se ilustra en la figura 8 para los fotogramas B y en la figura 9 para los fotogramas P); (3) predicción de MV espacial (por ejemplo, valor medio de los vectores de movimiento de los macrobloques colocados); (4) Promedio ponderado de la proyección de movimiento y la predicción espacial; (5) u otras técnicas similares.

También pueden usarse otros modelos de predicción (por ejemplo, aceleración, filtrado, etc.). Si solo se va a usar uno de estos modelos, entonces esto debería ser común tanto en el codificador como en el decodificador. De lo contrario, pueden usarse submodos que guiarán inmediatamente el decodificador en cuanto a qué modelo debería usarse. Los expertos en la materia también reconocerán que también es posible hacer multi-referencia a un bloque o macrobloque usando cualquier combinación de los modelos anteriores.

55 En la figura 12, un entorno 1200 de codificación de vídeo mejorado incluye un codificador 1202 de vídeo que recibe

datos 1004 de video y emite un flujo de bits de datos de vídeo codificado correspondiente.

En este caso, se ha modificado el codificador 1202 de vídeo para incluir la mejora 1204. La mejora 1204 incluye un módulo 1206 de memoria intermedia de vector de movimiento (MV) adicional y un módulo 1208 de decisión DIRECTA. Más específicamente, como se muestra, el módulo 1206 de memoria intermedia de MV está configurado para recibir como entradas, la salida del módulo 1026 de memoria intermedia de fotograma y la salida del módulo 1008 de ME. Se proporciona la salida del módulo 1206 de memoria intermedia de MV, junto con la salida del módulo 1008 de ME, como una entrada al módulo 1208 de decisión DIRECTA. La salida del módulo 1208 de decisión DIRECTA se proporciona a continuación como una entrada al módulo 1022 de MC junto con la salida del módulo 1026 de memoria intermedia de fotograma.

Para que la arquitectura a modo de ejemplo trabaje con éxito, la información de movimiento del fotograma codificado anteriormente se almacena intacta, que es el fin de añadir el módulo 1206 de memoria intermedia de MV. El módulo 1206 de memoria intermedia de MV puede usarse para almacenar vectores de movimiento. En ciertas implementaciones, el módulo 1206 de memoria intermedia de MV también puede almacenar la información sobre el fotograma de referencia usado y el modo de movimiento usado. Por ejemplo, en el caso de la aceleración, el almacenamiento en una memoria intermedia adicional puede ser útil para almacenar la información de movimiento del 2º o incluso los N fotogramas anteriores cuando, por ejemplo, se emplea un modelo más complicado para la aceleración.

Si un macrobloque, subbloque, o píxel no está asociado con un vector de movimiento (es decir, un macrobloque está intracodificado), entonces, se supone que para este bloque el vector de movimiento usado es (0, 0) y que solo el fotograma anterior se usa como referencia.

Si se usa la referencia de múltiples fotogramas, puede seleccionarse el uso de la información de movimiento tal como está, y/o para interpolar la información de movimiento haciendo referencia al fotograma codificado anteriormente. Esto es esencialmente para el diseño, pero también en la práctica parece que, especialmente para el caso de los vectores de movimiento (0, 0), es menos probable que el bloque actual todavía se referencie a partir de un fotograma mucho más antiguo.

Puede combinarse la predicción directa con un conjunto adicional de información de movimiento que está, a diferencia de antes, codificado como parte de la predicción directa. En tal caso, la predicción puede, por ejemplo, ser una predicción de múltiples hipótesis tanto de la predicción directa como de la información de movimiento.

Ya que hay varios sub-modos de predicción directa posibles que pueden combinarse, podrían combinarse también dentro de un marco de trabajo de múltiples hipótesis. Por ejemplo, la predicción de la proyección de movimiento podría combinarse con la de la proyección de píxeles y/o la predicción de MV espacial.

La predicción directa también puede usarse en el nivel de subbloque dentro de un macrobloque. Esto se hace ya para fotogramas B en el interior del códec H.26L actual, pero actualmente solo se está usando la proyección de movimiento y no la proyección de píxeles o sus combinaciones.

Para la codificación de fotograma B, puede realizarse una predicción directa de una sola dirección (hacia adelante o hacia atrás) y no siempre necesariamente de ambos lados. También puede usarse la predicción directa en el modo bidireccional de los fotogramas B, donde una de las predicciones usa la predicción directa.

Por ejemplo, en el caso de imágenes de múltiples hipótesis, es posible que un fotograma P haga referencia a un fotograma futuro. En este caso, la escala adecuada, y/o la inversión de la información de movimiento puede realizarse de manera similar a la interpolación de movimiento similar del fotograma B.

Por ejemplo, la codificación de longitud de ejecución, puede usarse también de acuerdo con que, si se usan unos modos P directos "equivalentes" posteriores en la codificación de un fotograma o diapositiva, entonces estos pueden codificarse usando una representación de longitud de ejecución.

El módulo 1208 de decisión DIRECTA realiza esencialmente la decisión de si el modo de predicción directa debería usarse en lugar de los inter o intra modos preexistentes. A modo de ejemplo, la decisión puede basarse en los criterios de optimización de tasa distorsión mixta, y/o también en los requisitos o restricciones de tasa de bits o de distorsión por separado.

También es posible, en implementaciones alternativas, que el módulo 1208 de predicción directa preceda al módulo 1008 de ME. En tal caso, si puede proporcionarse la predicción directa inmediatamente con una buena estimación suficiente, basándose en algunas condiciones predefinidas, para los parámetros de movimiento, el módulo 1008 de ME podría pasarse por alto completamente, por lo que también reduce considerablemente el cálculo de la codificación.

Procedimientos de decodificación a modo de ejemplo:

Se hace referencia ahora a la figura 11, que representa un entorno 1100 de decodificación convencional a modo de

ejemplo que tiene un decodificador 1102 de vídeo que recibe un flujo 1104 de bits de datos de vídeo codificado y unas salidas (decodificadas) que corresponden a los datos 1120 de vídeo.

El flujo 1104 de bits de datos de vídeo codificado se proporciona como una entrada a un módulo 1106 de decodificación de longitud variable (VLD). La salida del módulo 1106 de VLD se proporciona como una entrada a un módulo 1108 de QP<sup>-1</sup>, y como una entrada a un módulo 1110 de MC. La salida del módulo 1108 de QP<sup>-1</sup> se proporciona como una entrada a un módulo 1112 de IDCT. La salida del módulo 1112 de IDCT se proporciona como una entrada a un módulo 1114 de suma, que también recibe como una entrada la salida del módulo 1110 de MC. La salida del módulo 1114 de suma se proporciona como una entrada a un módulo 1116 de filtro de bucle. La salida del módulo 1116 de filtro de bucle se proporciona a un módulo 1118 de memoria intermedia de fotograma. Una salida del módulo 1118 de memoria intermedia de fotograma se proporciona como una entrada al módulo 1110 de MC. El módulo 1118 de memoria intermedia de fotograma emite también los datos 1120 de vídeo (decodificados).

Un decodificador 1302 mejorado a modo de ejemplo para su uso en un entorno 1300 de predicción directa incluye, además, una mejora 1306. En este caso, como se muestra en la figura 13, el decodificador 1302 mejorado recibe un flujo 1210 de bits de datos de vídeo codificado, por ejemplo, como salida del codificador 1202 de vídeo mejorado de la figura 12, y emite los datos 1304 de vídeo correspondientes (decodificados).

La mejora 1306, en este ejemplo, se inserta operativamente entre el módulo 1110 de MC y un módulo 1106' de VLD. La mejora 1306 incluye un módulo 1308 de memoria intermedia de MV que recibe como una entrada, una salida del módulo 1106' de VLD. La salida del módulo 1308 de memoria intermedia de MV se proporciona como una entrada seleccionable a un módulo 1312 de selección de la mejora 1306. Un módulo 1310 de modo de bloque se proporciona también en la mejora de 1306. El módulo 1310 de modo de bloque recibe como una entrada, una salida del módulo 1106' de VLD. Una salida del módulo 1310 de modo de bloque se proporciona como una entrada al módulo 1106' de VLD, y también como una entrada de control al módulo 1312 de selección. Una salida del módulo 1106' de VLD se proporciona como una entrada seleccionable al módulo 1312 de selección. El módulo 1312 de selección está configurado para proporcionar selectivamente o una salida del módulo 1308 de memoria intermedia de MV o del módulo 1106' de VLD como una entrada del módulo 1110 de MC.

Por ejemplo, con la mejora 1306 puede almacenarse la información de movimiento para cada píxel, y si el modo de un macrobloque se identifica como el modo de predicción directo, entonces se seleccionan y se usan la información de movimiento almacenada y el procedimiento de proyección o de predicción apropiado. Debería observarse que si solo se usa la proyección de movimiento, entonces los cambios en un decodificador existente son menores, y la complejidad adicional que se añade en el decodificador podría considerarse insignificante.

Si se usan submodos, entonces un decodificador 1302 mejorado puede, por ejemplo, configurarse para realizar las etapas opuestas a las etapas de predicción que realiza el codificador 1202 mejorado, con el fin de decodificar correctamente el macrobloque actual.

Una vez más los píxeles no referenciados (tales como los intra bloques) pueden considerarse como que tienen un movimiento cero para el almacenamiento de movimiento.

Algunos esquemas a modo de ejemplo

Teniendo en cuenta que existen varios predictores posibles que pueden usarse inmediatamente con la predicción directa, para fines de brevedad en esta descripción un subconjunto más pequeño de casos, que no solo son más bien eficientes, sino que también son más fáciles de implementar, se describen con mayor detalle. En particular, los siguientes modelos se examinan con mayor detalle demostrativo:

(A) En este ejemplo, el movimiento de proyección es el único modo usado. No se usa la codificación de longitud de ejecución de los modos directos, donde se transmite también como información residual. Se realiza una modificación especial de los parámetros de movimiento en el caso de que se use un vector de movimiento cero. En tal situación, el fotograma de referencia para la predicción directa se establece siempre en cero (por ejemplo, el fotograma codificado anterior). Además, los bloques intracodificados se consideran como que tienen unos parámetros de fotograma de movimiento y de referencia cero.

(B) Este ejemplo es como el ejemplo (A) excepto que no se transmite ningún residuo.

(C) Este ejemplo es básicamente una combinación de los ejemplos (A) y (B), en que si  $QP < n$  (por ejemplo,  $n = 24$ ) entonces el residuo se codifica también, de lo contrario el residuo no se transmite.

(D) Este ejemplo es un esquema de predicción directa mejorado que combina tres submodos, a saber:

- (1) Proyección de movimiento ( $\overrightarrow{MV}_{MP}$ );
- (2) Predicción de MV espacial ( $\overrightarrow{MV}_{SP}$ ); y
- (3) Una media ponderada de estos dos casos

$$\left( \frac{[\overrightarrow{MV}_{MP} + 2 * \overrightarrow{MV}_{SP}]}{3} \right).$$

En la que, el residuo no se transmite para  $QP < n$  (por ejemplo,  $n = 24$ ). En este caso, no se usa la codificación de longitud de ejecución. La partición de los submodos puede establecerse de la siguiente manera:

Submodos	Código
Predictor espacial	0
Movimiento de proyección	1
Promedio Ponderado	2

5 El mejor submodo podría seleccionarse usando un procedimiento de optimización de tasa distorsión (mejor compromiso entre la tasa de bits y la calidad).

(E) Una combinación del ejemplo (C) con la proyección de píxel. En este caso, por ejemplo, un promedio de dos predicciones para el modo de predicción directa.

10 (F) Esta es una combinación del ejemplo (C) con Motion\_Copy R2 (véase, por ejemplo, Jani Lainema y Marta Karczewicz, "Skip mode motion compensation", Doc. JVT-C027, mayo de 2002) o similares. Este caso puede verse como una alternativa del uso del predictor de MV espacial usado en el ejemplo (D), con una diferencia que es que el predictor espacial, bajo ciertas condiciones, reemplaza completamente el modo de salto de ceros, y que este ejemplo (F) puede codificarse por longitud de ejecución para poder conseguir de este modo un rendimiento más eficiente.

15 Predicción de vectores de movimiento en fotogramas (B) bidireccionalmente predictivos con respecto al modo directo:

20 La norma JVT actual parece ser bastante confusa sobre cómo deberían considerarse un macrobloque o un bloque codificado de modo directo en la predicción de vectores de movimiento dentro de los fotogramas (B) bidireccionalmente predictivos. En cambio, parece que el software actual considera un macrobloque o sub-bloque de modo directo como que tiene un "fotograma de referencia diferente" y por lo tanto no se usan en la predicción. Por desgracia, teniendo en cuenta que aún puede haber una alta correlación entre los vectores de movimiento de un bloque predicho directo con sus vecinos, tal condición podría dificultar considerablemente el rendimiento de los fotogramas B y reducir su eficacia. Esto también podría reducir la eficacia de los algoritmos de ocultación de errores cuando se aplican a los fotogramas B.

25 En esta sección, se presentan unos enfoques alternativos a modo de ejemplo, que pueden mejorar la eficacia de codificación aumentando, por ejemplo, la correlación de los vectores de movimiento dentro de los fotogramas B. Esto se hace considerando un bloque codificado en modo directo esencialmente equivalente a un bloque bidireccionalmente predicho dentro de la fase de predicción de movimiento.

30 Los macrobloques o bloques de modo Directo (por ejemplo, en el caso de las sub-particiones de 8x8) podrían mejorar considerablemente la eficacia de los fotogramas (B) bidireccionalmente predichos ya que pueden explotar de manera efectiva las correlaciones temporales de la información de vector de movimiento de los fotogramas adyacentes. La idea se obtiene esencialmente de técnicas de interpolación temporal, donde se hace la suposición de que si un bloque se ha movido desde una posición  $(x + dx, y + dy)$  en el tiempo  $t$  a una posición  $(x, y)$  en el tiempo  $t+2$ , entonces, usando la interpolación temporal, en el tiempo  $t+1$  el mismo bloque debe haber estado esencialmente en la posición:

$$\left(x + \frac{dx}{2}, y + \frac{dy}{2}\right)$$

40 Esto se ilustra, por ejemplo, en la figura 14, que muestra tres fotogramas, a saber, un fotograma 1400 P, un fotograma 1402 B y fotograma 1404 P, correspondientes a los tiempos  $t$ ,  $t+1$ , y  $t+2$ , respectivamente. El enfoque, aunque usado más a menudo en las normas de codificación actuales asume en su lugar que el bloque en la posición  $(x, y)$  del fotograma en el tiempo  $t+1$  puede encontrarse más probablemente en las posiciones:

$$\left(x + \frac{dx}{2}, y + \frac{dy}{2}\right) \text{ en el tiempo } t$$

y

$$\left(x - \frac{dx}{2}, y - \frac{dy}{2}\right) \text{ en el tiempo } t+2.$$

45 Lo último se ilustra en la figura 15, que muestra tres fotogramas, es decir, un fotograma 1500 P, un fotograma 1502 B y fotograma P 1504, que corresponden respectivamente a los tiempos  $t$ ,  $t+1$ , y  $t+2$ . Ya que el número de bloques codificados en modo directo dentro de una secuencia puede ser significativa, mientras que no se transmite información residual y de movimiento para un caso tal, puede aumentarse considerablemente la eficacia de los fotogramas B. La codificación por longitud de ejecución (por ejemplo, si se usa la codificación de entropía por código de longitud variable universal (UVLC)) puede usarse también para mejorar aún más el rendimiento.

5 Por desgracia, la norma JVT actual no aclara cómo debería realizarse la predicción de vectores de movimiento de los bloques adyacentes para los bloques de modo directo. Como se desprende del software actual, los bloques de modo directo se consideran actualmente como que tienen un “fotograma de referencia diferente” por lo tanto no se explota ninguna correlación espacial en tal caso. Esto podría reducir considerablemente la eficacia de la predicción, pero también podría afectar potencialmente al rendimiento de los algoritmos de ocultación de errores aplicados en los fotogramas B en el caso de que se necesite.

A modo de ejemplo, si se quisiera predecir el vector de movimiento de E en el codec actual, si A, B, C, y D están todos codificados en el modo directo, entonces el predictor se establecerá como (0,0) lo que no sería una buena decisión.

10 En la figura 16, por ejemplo, E se predice a partir de A, B, C y D. Por lo tanto, si A, B, C o D están codificados en el modo directo, entonces sus valores reales no se usan actualmente en la predicción. Sin embargo, esto puede modificarse. De este modo, por ejemplo, si A, B, C o D están codificados en el modo directo, entonces los valores reales de los vectores de movimiento y los fotogramas de referencia pueden usarse en la predicción. Esto proporciona dos opciones seleccionables: (1) si un macrobloque/bloque colocado en el fotograma P posterior está intracodificado, entonces un fotograma de referencia se establece en -1; (2) si un macrobloque/bloque colocado en el fotograma P posterior está intracodificado, entonces se asume que el fotograma de referencia es 0.

15 De acuerdo con ciertos aspectos proporcionados, en su lugar puede usarse la información de movimiento real disponible de los bloques codificados en modo directo, para realizar la predicción de vectores de movimiento. Esto permitirá una mayor correlación de los vectores de movimiento dentro de una secuencia de fotogramas B, y por lo tanto puede conducir a la mejora de la eficacia.

20 Un posible problema es cómo manejar apropiadamente los macrobloques de modo directo para los que, el bloque/macrobloque colocado en el fotograma posterior está intracodificado. Por ejemplo, en este caso, son posibles dos opciones:

- 25 (1) Considerar este macrobloque/bloque como que tiene un fotograma de referencia diferente, por lo que no lo usa en la predicción de vectores de movimiento; y  
(2) Considerar este macrobloque como que tiene un vector de movimiento (0, 0) y el fotograma de referencia 0.

30 De acuerdo con ciertas otras implementaciones a modo de ejemplo proporcionadas, puede hacerse una modificación adicional en el procedimiento de filtro de desbloqueo. Para el caso del modo directo, puede configurarse un procedimiento de filtro de desbloqueo para comparar la información de vector de movimiento almacenada que se toma de los bloques codificados en modo directo, de otro modo, por lo general se considerarían como cero. Sin embargo, en otra modificación en su lugar puede configurarse el procedimiento de filtro de desbloqueo para comparar los vectores de movimiento (exactos) independientemente del tipo de bloque que se use. Por lo tanto, en ciertas implementaciones, si para los bloques de codificación directa no se transmite ningún residuo, un filtro de desbloqueo “fuerte” puede proporcionar un rendimiento mejorado.

35 Además, en ciertas otras implementaciones, puede rediseñarse la decisión de tasa distorsión para los fotogramas B ya que es muy probable que para ciertas implementaciones del esquema de predicción de vectores de movimiento, un parámetro lagrangiano diferente usado en decisiones de optimización de tasa distorsión, puede conducir a una eficacia de codificación. Tal  $\lambda$  pueden tomarse, por ejemplo, como:

$$\lambda = 0.85 \times 2^{\frac{Qp}{3}}$$

#### 40 Refinamiento de decisión de inter modo:

45 La norma JVT tiene actualmente una ventaja abrumadora de rendimiento en comparación con otras normas de codificación basadas en bloque actuales. Parte de este rendimiento puede atribuirse a la posibilidad de usar tamaños de bloque variables que varían de 16x16 a 4x4 (píxeles), en lugar de tener tamaños de bloque fijos. Haciendo esto se permite, por ejemplo, una explotación más eficaz de la correlación temporal. Por desgracia, se ha descubierto que, debido a las técnicas de decisión de modo existentes actualmente en la lógica de codificación convencional (por ejemplo, hardware, firmware y/o software), las decisiones de modo no podrían realizarse de manera óptima, por lo tanto, se perderían bits que podrían asignarse mejor.

50 En esta sección, se proporcionan otros procedimientos y aparatos que al menos en parte resuelven este problema y/u otros. En este caso, los procedimientos y aparatos a modo de ejemplo se han configurado para su uso con al menos los modos de bloque 16x8 y 8x16 (píxel). Además, usando una solución relativamente simple donde se introduce al menos un criterio adicional, se proporciona un ahorro de entre aproximadamente el 5 % y el 10 % en la complejidad del codificador.

55 Dos características clave de la norma JVT son la selección de modo de macrobloque variable y la optimización de tasa distorsión. Un macrobloque 16x16 (píxel) puede codificarse usando diferentes modos de partición para los que también se transmite la información de movimiento. La selección del modo a usarse puede realizarse en la fase de

optimización de tasa distorsión de la codificación, donde se intenta una decisión mixta de la mejor calidad posible a la mejor tasa de bits posible. Desafortunadamente, ya que las asignaciones de la mejor información de movimiento posible para cada subpartición se hace en un procedimiento completamente diferente de la codificación, es posible en algunos casos, que un modo 16x16 (por ejemplo, 16x8 u 8x16 (pixel)) no lleve información de movimiento que sea equivalente a un macrobloque 16x16. Ya que los predictores de movimiento usados para cada modo también podrían ser diferentes, es bastante posible en muchos casos que tal información de movimiento de tipo 16x16 pueda ser diferente de la asignada al modo 16x16. Además, bajo ciertas condiciones, la optimización de tasa distorsión puede que al final decida no usar el tipo de macrobloque 16x16, a pesar de que continúa la información de movimiento de 16x16, sin examinar si esta podría ser mejor si se codifica usando un modo 16x16.

- 5 Reconociendo esto, puede configurarse un sistema a modo de ejemplo para determinar cuándo se produce un caso de este tipo, de tal manera que puede alcanzarse una mejora de rendimiento. De acuerdo con ciertas implementaciones proporcionadas a modo de ejemplo, dos modos adicionales, por ejemplo, denominados como P2to1 y P3to1, se ponen a disposición dentro del procedimiento/fase de decisión de modo. Los modos P2to1 y P3to1 se activan cuando la información de movimiento de unas subparticiones 16x8 y 8x16, respectivamente, es equivalente a la de un modo 16x16.

En ciertas implementaciones todo el fotograma de vector de movimiento y de referencia asignado a cada partición puede ser igual. Como tal, el modo equivalente puede activarse y examinarse durante un procedimiento/fase de tasa distorsión. Ya que la información residual y de distorsión no es probable que cambien en comparación con el caso de subpartición, pueden reutilizarse sin aumentar el cálculo significativamente.

- 20 Teniendo en cuenta sin embargo que la decisión de modo de tasa distorsión no es perfecta, es posible que la suma y la consideración de estos dos modos adicionales, con independencia del mejor modo actual, puede reducir, en algunos casos limitados, la eficacia en lugar de mejorarla. Como una alternativa, estos modos pueden permitirse solo cuando el modo subpartición correspondiente sea también el mejor posible de acuerdo con la decisión de modo empleada. Haciendo esto, pueden producirse mejoras (por ejemplo, una reducción de la tasa de bits) frente a la otra lógica (por ejemplo, códec, etc.), mientras que no afecte al PSNR.

Si la información de movimiento de las subparticiones 16x8 u 8x16 es equivalente a la del modo 16x16, entonces puede ser innecesario la realización de la decisión de modo para un modo de este tipo. Por ejemplo, si el predictor de vectores de movimiento de la primera subpartición es exactamente el mismo que el predictor de vectores de movimiento del modo 16x16 es innecesario realizar la decisión de modo. Si tal condición se cumple, puede omitirse por completo este modo durante el procedimiento de decisión de modo. Si lo hace, puede reducir significativamente la complejidad, ya que no sería necesario, para este modo, realizar la DCT, la cuantificación, y/u otros procedimientos/mediciones de tasa distorsión similares, que tienden a ser bastante costosos durante el procedimiento de codificación.

- 30 En ciertas otras implementaciones a modo de ejemplo, todo el procedimiento puede extenderse también además a una partición de macrobloque con estructura de árbol. Véase, por ejemplo, Heiko Schwarz y Thomas Wiegand, "Tree-structured macroblock partition", Doc. VCEG-N17, Diciembre de 2001.

Un algoritmo a modo de ejemplo

A continuación, se presentan algunas acciones que pueden realizarse para proporcionar un refinamiento en un codec a modo de ejemplo u otra lógica similar (Obsérvese que en ciertas otras implementaciones, el orden de las acciones puede modificarse y/o que ciertas acciones pueden realizarse juntas):

- 40 Acción 1: Establecer  $Valid[P2to1] = Valid[P3to1] = 0$ .
- Acción 2: Realizar decisión de vector de movimiento y fotograma de referencia para cada posible inter modo. Permitir que  $\overrightarrow{MV}_{16x16}$ ,  $\overrightarrow{MVP}_{16x16}$ , y  $reframe_{16x16}$  sean el vector de movimiento, el predictor de vectores de movimiento, y el fotograma de referencia del modo 16x16,  $\{\overrightarrow{MV}_{16x8}^a, \overrightarrow{MV}_{16x8}^b\}$ ,  $\{\overrightarrow{MVP}_{16x8}^a, \overrightarrow{MVP}_{16x8}^b\}$  y  $\{reframe_{16x8}^a, reframe_{16x8}^b\}$  la información correspondiente para el modo de 16x8, y  $\{\overrightarrow{MV}_{8x16}^a, \overrightarrow{MV}_{8x16}^b\}$ ,  $\{\overrightarrow{MVP}_{8x16}^a, \overrightarrow{MVP}_{8x16}^b\}$ , y  $\{reframe_{8x16}^a, reframe_{8x16}^b\}$  para el modo de 8x16.

Acción 3: Si  $(\overrightarrow{MV}_{16x8}^a \neq \overrightarrow{MV}_{16x8}^b)$  o  $(reframe_{16x8}^a \neq reframe_{16x8}^b)$  e ir a la Acción 7.

Acción 4: Si  $(\overrightarrow{MV}_{16x8}^a \neq \overrightarrow{MV}_{16x16})$  o  $(\overrightarrow{MVP}_{16x8}^a \neq \overrightarrow{MVP}_{16x16})$  o  $(reframe_{16x8}^a \neq reframe_{16x16})$ , entonces ir a la Acción 6.

- 50 Acción 5:  $Valid[16x8] = 0$ ; ir a la Acción 7 (por ejemplo., desactivar el modo 16x8 si es idéntico al 16x16. Reducción de complejidad).

Acción 6:  $Valid[P2to1] = 1$ ; (por ejemplo, habilitar el modo de refinamiento para 16x8)

$$\overrightarrow{MV}_{P2to1} = \overrightarrow{MV}_{16x8}^a; refframe_{P2to1} = refframe_{16x8}^a;$$

Acción 7: Si

$$(\overrightarrow{MV}_{8x16}^a \neq \overrightarrow{MV}_{8x16}^b) \text{ o } (refframe_{8x16}^a \neq refframe_{8x16}^b),$$

5 entonces, ir a la Acción 11.

Acción 8: Si

$$(\overrightarrow{MV}_{8x16}^a \neq \overrightarrow{MV}_{16x16}) \text{ o } (\overrightarrow{MVP}_{8x16}^a \neq \overrightarrow{MVP}_{16x16}) \text{ o } (refframe_{8x16}^a \neq refframe_{16x16}),$$

entonces, ir a la Acción 10.

10 Acción 9:  $Valid[8x16] = 0$ ; ir a la Acción 11 (por ejemplo, deshabilitar el modo de 8x16 si es idéntico al 16x16 para reducir la complejidad)

Acción 10:  $Valid[P3to1] = 1$  (Por ejemplo, activar el modo de refinamiento para 8x16)

$$\overrightarrow{MV}_{P3to1} = \overrightarrow{MV}_{8x16}^a; refframe_{P3to1} = refframe_{8x16}^a;$$

15 Acción 11: Realizar la optimización de tasa distorsión para todos los inter & intra modos si ( $Valid[MODO] = 1$ ) donde  $MODO \in \{INTRA\ 4x4, INTRA\ 16x16, OMITIR, 16x16, 16x8, 8x16, P8x8\}$ , usando el funcional lagrangiano:  $J(s, c, MODO|QP, \lambda_{MODO}) = SSD(s, c, MODO|QP) + \lambda_{MODO} \cdot R(s, c, MODO|QP)$  Establecer Acción mejor modo a *Mejormodo*

Acción 12: Si ( $Mejormodo! = 16x8$ ), entonces  $Valid[P3to1] = 0$  (obsérvese que esta acción es opcional).

Acción 13: Si ( $Mejormodo! = 8x16$ ) entonces,  $Valid[P2to1] = 0$  (obsérvese que esta acción es opcional).

20 Acción 14: Realizar la optimización de tasa distorsión para los dos modos adicionales si ( $Valid[MODO] = 1$ ) donde  $MODO \in \{P2to1, P3to1\}$  (por ejemplo, los modos se consideran equivalentes a los modos 16x16).

Acción 15: Establecer *Mejormodo* al mejor modo general descubierto.

Aplicar las técnicas de predicción directa a modo de ejemplo para la codificación de entrelazado:

25 Debido al creciente interés de la codificación de vídeo entrelazado dentro de la norma H.26L, se han presentado varias propuestas para mejorar el rendimiento de codificación de las secuencias entrelazadas. En esta sección, se presentan las técnicas que pueden implementarse en la sintaxis actual de H.26L, y/u otros sistemas similares. Estas técnicas a modo de ejemplo pueden proporcionar una mejora de rendimiento. Además, se introduce la tecnología de predicción P directa, similar a la predicción B directa, que puede aplicarse tanto en la codificación de vídeo entrelazado como progresivo.

Información adicional sobre técnicas de predicción P directa a modo de ejemplo:

30 El modo directo de los vectores de movimiento dentro de los fotogramas B puede beneficiarse significativamente del rendimiento de codificación, ya que puede reducir considerablemente los bits necesarios para la codificación de vectores de movimiento, considerando especialmente que tienen que transmitirse hasta dos vectores de movimiento. Sin embargo, si un bloque se codifica usando el modo directo, no son necesarios vectores de movimiento cuando, en su lugar, estos se calculan como interpolaciones temporales de los vectores de movimiento de los bloques colocados en la primera imagen de referencia posterior. Parece no haberse considerado un enfoque similar para los

35 fotogramas P ya que la estructura de los fotogramas P y de su macrobloque correspondiente era mucho más simple, mientras que cada macrobloque requiere solo un vector de movimiento. Lo más probable es que con la adición de un modo de este tipo se habría incurrido en una sobrecarga significativa, negando posiblemente de este modo cualquier ganancia posible.

40 Por otro lado, en H.26L los fotogramas P se han mejorado significativamente introduciendo varios modos de macrobloque adicionales. Como se ha descrito anteriormente, en muchos casos incluso podría necesitarse transmitir hasta 16 vectores de movimiento por macrobloque. Considerando esta sobrecarga de modo adicional que pueden contener los fotogramas P en H.26L, podría ser viable una implementación de predicción directa de los vectores de movimiento. De tal manera, todos los bits para los vectores de movimiento y para el fotograma de referencia usado

45 pueden ahorrarse en solo el coste del modo adicional, por ejemplo, véase la figura 4.

A pesar de que un procedimiento más sencillo de predicción P directa es seleccionar los vectores de movimiento de

los píxeles localizados en el fotograma anterior, en otras implementaciones puede también considerarse la aceleración de movimiento como una solución alternativa. Esto viene del hecho de que el movimiento podría cambiarse fotograma por fotograma, que no es constante, y usando la aceleración podrían obtenerse mejores resultados, por ejemplo, véase la figura 7.

5 Tales técnicas pueden aplicarse además a la codificación de vídeo progresiva. Sin embargo, considerando la correlación que los campos pueden tener en algunos casos dentro de las secuencias de entrelazado, tales como por ejemplo las regiones con un único movimiento horizontal constante, este enfoque puede también ayudar a mejorar la eficacia de codificación para la codificación de secuencia de entrelazado. Esto es en particular beneficioso para los fotogramas de tipo de campo conocidos, por ejemplo, si se supone que el movimiento de los campos adyacentes es el mismo. En este tipo de disposición, pueden considerarse los mismos campos de paridad como nuevos fotogramas y se codifican secuencialmente sin tener la consideración de la función de entrelazado. Todo esto está a la izquierda del decodificador. Sin embargo, usando este modo P directo a modo de ejemplo, puede usarse un conjunto de vectores de movimiento para el primer macrobloque de campo a codificar (por ejemplo, del tamaño de 16x16 píxeles), en la medida que el segundo campo en la misma localización está reutilizando la misma información de movimiento. La única otra información necesaria a enviarse es la imagen residual codificada. En otras implementaciones, es posible mejorar aún más en estas técnicas considerando las correlaciones entre las imágenes residuales de los dos bloques de campo colocados.

Con el fin de permitir que el modo directo en los fotogramas P, es básicamente necesario añadir un inter modo adicional en el sistema. Por lo tanto, en lugar de tener solo 8 inter modos, en un ejemplo, ahora pueden usarse 9, que se muestran a continuación:

INTER MODOS	DESCRIPCIÓN
COPY_MB	0 Modo de macrobloque omitido
M16x16_MB	1 Uno de 16x16 bloques
M16x8_MB	2 Dos de 16x8 bloques
M8x16_MB	3 Dos de 8x16 bloques
M8x8_MB	4 Cuatro de 8x8 bloques
M8x4_MB	5 Ocho de 8x4 bloques
M4x8_MB	6 Ocho de 4x8 bloques
M4x4_MB	7 Dieciséis 16x8 bloques
PDIRECT_MB	8 Modo de copia y vectores de movimiento de macrobloque colocados en el fotograma previo

En general, tales modos directos a modo de ejemplo para los fotogramas P pueden aparecer si el macrobloque colocado también es de tipo INTER, excepto el macrobloque omitido, pero incluyendo el modo directo, ya que en otros casos no hay información de movimiento que pueda usarse. En el caso de que el macrobloque anterior también esté codificado en el modo P directo, los vectores de movimiento más recientes y el modo para este macrobloque se consideran en su lugar. Para manejar más eficientemente los casos en que este modo no aparecerá lógicamente, y en particular si se usa el INTRA modo, puede seleccionarse permitir que este modo aparezca también en estos casos con el modo significando ahora un segundo modo de macrobloque omitido donde una copia de la información no es del fotograma anterior, pero si a partir del anterior. En este caso, no se codifica la información residual. Esto es específicamente útil para las secuencias de entrelazado, ya que es más probable que un macrobloque pueda encontrarse con mayor precisión a partir del mismo fotograma de campo de paridad, y no a partir del fotograma de campo anteriormente codificado como se presentó en las técnicas anteriores.

Para mejorar más la eficacia, si se usa un conjunto de dos fotogramas de tipo campo cuando se codifican las imágenes entrelazadas, puede configurarse el modo de macrobloque omitido para usar las mismas imágenes de campo de paridad. Por ejemplo, si el modo directo P se usa como una bandera de omisión, entonces se usa en su lugar una paridad diferente. Un beneficio adicional del modo P directo, es que puede permitir una reducción significativa de complejidad en el codificador, ya que es posible permitir que el sistema realice una pre-comprobación de si el modo P directo proporciona una solución lo suficientemente satisfactoria, y si es así, puede no necesitarse ningún cálculo adicional para la decisión de modo y la estimación de movimiento de ese bloque específico. Para abordar también la cuestión de la codificación de vectores de movimiento, los vectores de movimiento usados para la codificación P directa pueden usarse "tal cual" para el cálculo de un predictor MEDIANO.

Técnica de primer mejor campo y reorganización de campo:

La codificación de secuencia de entrelazado que permite soportar tanto el material de fotograma entrelazado como las imágenes de campo entrelazado separado dentro del mismo flujo proporcionaría probablemente una solución mucho mejor que la codificación usando solo uno de los dos procedimientos. La técnica de campo entrelazado separado tiene algunos beneficios adicionales, tales como, por ejemplo, el desbloqueo, y en particular puede proporcionar una resistencia a errores mejorada. Si ocurre un error de imagen dentro de un campo, por ejemplo, el error puede consumarse fácilmente usando la información de la segunda imagen.

Esto no es el caso de la técnica basada en fotogramas, donde especialmente cuando se considera el tamaño a menudo grande y los bits usados por tales fotogramas, pueden producirse errores dentro un fotograma de este tipo con mucha mayor probabilidad. La correlación reducida entre los píxeles/bloques no puede promover la recuperación de errores.

- 5 En este caso, puede mejorarse aún más el concepto de codificación de campo/fotograma, permitiendo que el codificador seleccione qué campo debería codificarse en primer lugar, sin tener en cuenta qué campo se va a visualizar primero. Esto podría manejarse de manera automática en un decodificador, donde será necesario una mayor memoria intermedia para almacenar un fotograma de campo futuro antes de mostrarle. Por ejemplo, a pesar de que el campo superior precede al campo inferior en términos de tiempo, la eficacia de codificación podría ser mayor si el campo inferior se codifica y se transmite primero, seguido por el fotograma de campo superior. La decisión puede tomarse, por ejemplo, en el procedimiento/fase de optimización de tasa distorsión, donde se examina primero cuál será el rendimiento si el campo impar se codifica en primer lugar seguido por el campo par, y del rendimiento si el campo par se codifica en su lugar y se usa como una referencia para el campo impar. Un procedimiento de este tipo implica que tanto el codificador como el decodificador deberían saber qué campo debería mostrarse primero, y cualquier reorganización hecha sin problemas. También es importante que, a pesar de que el campo impar se ha codificado primero, tanto el codificador como el decodificador son conscientes de este cambio al indexar el fotograma con el fin de una inter/intra predicción. Los ejemplos ilustrativos de un esquema de predicción de este tipo, que usa 4 fotogramas de referencia, se representan en la figura 17 y en la figura 18. En la figura 17, la codificación de entrelazado se muestra usando un esquema de primer mejor campo a modo de ejemplo en los fotogramas P. En la figura 18, la codificación de entrelazado se muestra usando un esquema de primer mejor campo en los fotogramas B.

En el caso de codificación de imágenes de campo/fotograma mixtas, puede emplearse el esquema representado de manera ilustrativa en la figura 19. En este caso, se muestra una implementación a modo de ejemplo de un esquema de primer mejor campo con el fotograma y la codificación basada en campos. Si se usan dos fotogramas para la estimación de movimiento basada en fotogramas, entonces al menos pueden usarse cinco fotogramas de campo para la estimación de movimiento de los campos, especialmente si se produce el intercambio de campo. Esto permite hacer referencia a al menos dos fotogramas de campo de la misma paridad. En general, deberían almacenarse  $2xN+1$  fotogramas de campo si se usan N fotogramas completos. Los fotogramas también podrían fácilmente intercalarse y desintercalarse en el codificador y en el decodificador para tales procedimientos.

30 Aunque lo que se reivindica es la invención, la siguiente es una lista de características adicionales:

Punto 1: Un procedimiento para su uso en la codificación de datos de vídeo dentro de una secuencia de fotogramas de vídeo, comprendiendo el procedimiento;  
 codificar al menos una parte de al menos un fotograma de referencia para incluir la información de movimiento asociada con dicha parte de dicho fotograma de referencia;  
 35 definir al menos una parte de al menos un fotograma predecible que incluye unos datos de vídeo de manera predictiva correlacionados con dicha parte de dicho fotograma de referencia basándose en dicha información de movimiento; y  
 codificar al menos dicha parte de dicho fotograma predecible sin incluir una información de movimiento correspondiente e incluyendo datos de identificación de modo que identifican que dicha parte de dicho fotograma predecible puede obtenerse directamente usando al menos dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia.

Punto 2: El procedimiento de acuerdo con el punto 1, en el que dichos datos de identificación de modo definen un tipo de modelo de predicción necesaria para decodificar dicha parte codificada de dicho fotograma predecible.

Punto 3: El procedimiento de acuerdo con el punto 2, en el que dicho tipo de modelo de predicción incluye un modelo de predicción directa mejorado que incluye al menos un submodo seleccionado de un grupo que comprende un submodo de proyección de movimiento, un submodo de predicción de vectores de movimiento espacial, y un submodo de promedio ponderado.

Punto 4: El procedimiento de acuerdo con el punto 3, en el que dichos datos de identificación de modo identifican dicho al menos un submodo.

50 Punto 5: El procedimiento de acuerdo con el punto 1, en el que dicho procedimiento genera una pluralidad de fotogramas de vídeo que comprenden al menos un fotograma predecible seleccionado de un grupo de fotogramas predecibles que comprende un fotograma P y un fotograma B.

Punto 6: El procedimiento de acuerdo con el punto 1, en el que dicha parte de dicho fotograma de referencia incluye unos datos para al menos un píxel dentro de dicho fotograma de referencia, y dicha parte de dicho fotograma predecible incluye unos datos para al menos un píxel dentro de dicho fotograma predecible.

Punto 7: El procedimiento de acuerdo con el punto 6, en el que dicha parte de dicho fotograma de referencia incluye unos datos de al menos una parte de un macrobloque dentro de dicho fotograma de referencia, y dicha parte de dicho fotograma predecible incluye unos datos para al menos una parte de un macrobloque dentro de

dicho fotograma predecible.

Punto 8: El procedimiento de acuerdo con el punto 1, en el que dicho fotograma de referencia precede temporalmente a dicho fotograma predecible dentro de dicha secuencia de fotogramas de vídeo.

5 Punto 9: El procedimiento de acuerdo con el punto 1, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información de velocidad.

Punto 10: El procedimiento de acuerdo con el punto 1, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información de aceleración

Punto 11: El procedimiento de acuerdo con el punto 1, en el que dicha parte de dicho fotograma de referencia, y dicha parte de dicho fotograma predecible están correlacionadas espacialmente.

10 Punto 12: El procedimiento de acuerdo con el punto 1, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información de proyección de píxel necesaria para decodificar dicha parte codificada de dicho fotograma predecible.

15 Punto 13: El procedimiento de acuerdo con el punto 1, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia, incluye una información de predicción de vectores de movimiento espacial necesaria para decodificar dicha parte codificada de dicho fotograma predecible.

Punto 14: El procedimiento de acuerdo con el punto 1, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información combinada de proyección de píxel y de predicción de vectores de movimiento espacial necesaria para decodificar dicha parte codificada de dicho fotograma predecible.

20 Punto 15: El procedimiento de acuerdo con el punto 1, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información de predicción de múltiples hipótesis necesaria para decodificar dicha parte codificada de dicho fotograma predecible.

25 Punto 16: El procedimiento de acuerdo con el punto 1, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia es nula y dichos datos de identificación de modo identifican que dicha parte de dicho fotograma de referencia incluye dicha parte de dicha predecible.

Punto 17: El procedimiento de acuerdo con el punto 1, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información residual correspondiente.

30 Punto 18: El procedimiento de acuerdo con el punto 1, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información residual correspondiente solo si un parámetro de cuantificación (QP) satisface al menos una condición definida.

Punto 19: El procedimiento de acuerdo con el punto 18, en el que dicha al menos una condición definida incluye un valor umbral.

Punto 20: El procedimiento de acuerdo con el punto 18, en el que dicho un valor umbral es de aproximadamente  $QP > \text{veintitrés}$ .

35 Punto 21: El procedimiento de acuerdo con el punto 1, en el que dicho al menos un fotograma predecible y un fotograma de referencia son parte de una secuencia de entrelazado de campos de vídeo.

Punto 22: El procedimiento de acuerdo con el punto 21, en el que la información de movimiento está asociada con al menos un píxel colocado en dicho fotograma de referencia.

40 Punto 23: El procedimiento de acuerdo con el punto 22, en el que codificar al menos una parte de dicho al menos un fotograma de referencia incluye además codificar basándose en una correlación entre las imágenes residuales de dos bloques de campo colocados.

Punto 23a: El procedimiento de acuerdo con el punto 21, que comprende además para cada uno de dicho fotograma de referencia y dicho fotograma predecible seleccionar un orden en el que deben codificarse los campos dentro de dicha secuencia de entrelazado de campos de vídeo.

45 Punto 24: El procedimiento de acuerdo con el punto 21, en el que dicho al menos un fotograma predecible y un fotograma de referencia tienen cada uno al menos dos campos asociados con los mismos.

Punto 25: El procedimiento de acuerdo con el punto 1, que comprende además:

50       determinar selectivamente si se usa un modo de predicción directa en lugar de un modo de pre-existencia durante dicha codificación de dicha al menos dicha parte de dicho fotograma predecible basándose en al menos un factor.

Punto 26: Un medio legible por ordenador que tiene unas instrucciones que pueden implementarse en ordenador para realizar acciones que comprenden:

5           codificar datos de vídeo para una secuencia de fotogramas de vídeo en al menos un fotograma predecible seleccionado de un grupo de fotogramas predecibles que comprenden un fotograma P y un fotograma B, mediante:

                  codificar al menos una parte de al menos un fotograma de referencia para incluir la información de movimiento asociada con dicha parte de dicho fotograma de referencia;  
 10           definir al menos una parte de al menos un fotograma predecible que incluye unos datos de vídeo correlacionados de manera predictiva con dicha parte de dicho fotograma de referencia basándose en dicha información de movimiento; y  
 codificar al menos dicha parte de dicho fotograma predecible sin incluir una información de movimiento correspondiente e incluyendo unos datos de identificación de modo que identifican que dicha parte de dicho fotograma predecible puede obtenerse directamente usando al menos dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia.

15           Punto 27: El medio legible por ordenador de acuerdo con el punto 26, en el que dichos datos de identificación de modo definen un tipo de modelo de predicción necesaria para decodificar dicha parte codificada de dicho fotograma predecible.

20           Punto 28: El medio legible por ordenador de acuerdo con el punto 27, en el que dicho tipo de modelo de predicción incluye un modelo de predicción directa mejorado que incluye al menos un submodo seleccionado de un grupo que comprende un submodo de proyección de movimiento, un submodo de predicción de vectores de movimiento espacial, y un submodo de promedio ponderado, y en el que dichos datos de identificación de modo identifican dicho al menos un submodo.

25           Punto 29: El medio legible por ordenador de acuerdo con el punto 26, en el que dicha parte de dicho fotograma de referencia incluye unos datos para al menos un píxel dentro de dicho fotograma de referencia, y dicha parte de dicho fotograma predecible incluye unos datos para al menos un píxel dentro de dicho fotograma predecible.

                  Punto 30: El medio legible por ordenador de acuerdo con el punto 26, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información seleccionada de un grupo que comprende una información de velocidad y una información de aceleración.

30           Punto 31: El medio legible por ordenador de acuerdo con el punto 26, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información necesaria para decodificar dicha parte de dicho fotograma predecible que se selecciona de un grupo que comprende:

- una información de proyección de píxeles;
- una información de predicción de vectores de movimiento espacial;
- una información de proyección de píxeles y de predicción de vectores de movimiento espacial ponderadas; y

35           una información de predicción de múltiples hipótesis.

                  Punto 32: El medio legible por ordenador de acuerdo con el punto 26, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia es nula y dichos datos de identificación de modo identifican que dicha parte de dicho fotograma de referencia incluye dicha parte de dicho predecible.

40           Punto 33: El medio legible por ordenador de acuerdo con el punto 26, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información residual correspondiente.

                  Punto 34: El medio legible por ordenador de acuerdo con el punto 26, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información residual correspondiente solo si un parámetro de cuantificación (QP) satisface al menos una condición definida.

45           Punto 35: El medio legible por ordenador de acuerdo con el punto 34, en el que dicha al menos una condición definida incluye un valor umbral.

                  Punto 36: El medio legible por ordenador de acuerdo con el punto 26, en el que dicho al menos un fotograma predecible y un fotograma de referencia son parte de una secuencia de entrelazado de campos de vídeo.

                  Punto 37: El medio legible por ordenador de acuerdo con el punto 36, en el que la información de movimiento está asociada con al menos un píxel colocado en dicho fotograma de referencia.

50           Punto 38: El medio legible por ordenador de acuerdo con el punto 37, en el que codificar al menos una parte de dicho al menos un fotograma de referencia incluye además codificar basándose en una correlación entre las

imágenes residuales de dos bloques de campo colocados.

Punto 39: El medio legible por ordenador de acuerdo con el punto 36, que comprende además para cada uno de dicho fotograma de referencia y dicho fotograma predecible seleccionar un orden en el que deben codificarse los campos dentro de dicha secuencia de entrelazado de campos de vídeo.

5 Punto 40: El medio legible por ordenador de acuerdo con el punto 36, en el que dicho al menos un fotograma predecible y un fotograma de referencia tienen cada uno al menos dos campos asociados con los mismos.

Punto 41: El medio legible por ordenador de acuerdo con el punto 26, que tiene unas instrucciones que pueden implementarse por ordenador para realizar unas acciones adicionales que comprenden:

10 determinar selectivamente si se usa un modo de predicción directa en lugar de un modo de pre-existencia durante dicha codificación de dicha al menos dicha parte de dicho fotograma predecible basándose en al menos un factor.

Punto 42: Un aparato para su uso en codificar datos de vídeo para una secuencia de fotogramas de vídeo en una pluralidad de fotogramas de vídeo que incluye al menos un fotograma predecible seleccionado de un grupo de fotogramas predecibles que comprenden un fotograma P y un fotograma B, comprendiendo dicho aparato:

15 una memoria para almacenar la información de movimiento; y

una lógica acoplada de manera operativa a dicha memoria y configurada para codificar al menos una parte de al menos un fotograma de referencia para incluir la información de movimiento asociada con dicha parte de dicho fotograma de referencia, determinar al menos una parte de al menos un fotograma predecible que incluye unos datos de vídeo correlacionados predictivamente con dicha parte de dicho fotograma de referencia basándose en dicha información de movimiento, y codificar al menos dicha parte de dicho fotograma predecible sin incluir la información de movimiento correspondiente e incluyendo unos datos de identificación de modo que identifican que dicha parte de dicho fotograma predecible puede obtenerse directamente usando al menos dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia.

20

25 Punto 43: El aparato de acuerdo con el punto 42, en el que dichos datos de identificación de modo definen un tipo de modelo de predicción necesaria para decodificar dicha parte codificada de dicho fotograma predecible.

Punto 44: El aparato de acuerdo con el punto 43, en el que dicho tipo de modelo de predicción incluye un modelo de predicción directa mejorado que incluye al menos un submodo seleccionado de un grupo que comprende un submodo de proyección de movimiento, un submodo de predicción de vectores de movimiento espacial, y un submodo de promedio ponderado, y en el que dichos datos de identificación de modo identifican dicho al menos un submodo.

30

Punto 45: El aparato de acuerdo con el punto 42, en el que dicha parte de dicho fotograma de referencia incluye unos datos para al menos un píxel dentro de dicho fotograma de referencia, y dicha parte de dicho fotograma predecible incluye unos datos para al menos un píxel dentro de dicho fotograma predecible.

35 Punto 46: El aparato de acuerdo con el punto 42, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información seleccionada de un grupo que comprende una información de velocidad y una información de aceleración.

Punto 47: El aparato de acuerdo con el punto 42, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información necesaria para decodificar dicha parte codificada de dicho fotograma predecible que se selecciona de un grupo que comprende:

40

una información de proyección de píxeles;

una información de predicción de vectores de movimiento espacial;

una información de proyección de píxeles y de predicción de vectores de movimiento espacial ponderadas; y

una información de predicción de múltiples hipótesis.

45 Punto 48: El aparato de acuerdo con el punto 42, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia es nula y dichos datos de identificación de modo identifican que dicha parte de dicho fotograma de referencia, incluye dicha parte de dicha predecible.

Punto 49: El aparato de acuerdo con el punto 42, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información residual correspondiente.

50 Punto 50: El aparato de acuerdo con el punto 42, en el que dicha información de movimiento asociada con dicha

parte de dicho fotograma de referencia incluye una información residual correspondiente solo si un parámetro de cuantificación (QP) satisface al menos una condición definida.

Punto 51: El aparato de acuerdo con el punto 42, en el que dicho al menos un fotograma predecible y un fotograma de referencia son parte de una secuencia de entrelazado de campos de vídeo.

5 Punto 52: El aparato de acuerdo con el punto 51, en el que la información de movimiento está asociada con al menos un píxel colocado en dicho fotograma de referencia.

Punto 53: El aparato de acuerdo con el punto 52, en el que dicha lógica codifica dicha al menos una parte de dicho al menos un fotograma de referencia basándose en una correlación entre las imágenes residuales de dos bloques de campo colocados.

10 Punto 54: El aparato de acuerdo con el punto 51, en el que dicha lógica se configura además para seleccionar, para cada uno de dicho fotograma de referencia y dicho fotograma predecible, un orden en el que deben codificarse los campos dentro de dicha secuencia de entrelazado de campos de vídeo.

Punto 55: El aparato de acuerdo con el punto 51, en el que dicho al menos un fotograma predecible y un fotograma de referencia, tiene cada uno al menos dos campos asociados con los mismos.

15 Punto 56: El aparato de acuerdo con el punto 42, en el que dicha lógica está configurada además para determinar selectivamente si se usa un modo de predicción directa en lugar de un modo de pre-existencia cuando se codifica dicha al menos dicha parte de dicho fotograma predecible basándose en al menos un factor.

20 Elemento 57: Un método para su uso en decodificar los datos de vídeo codificados que incluye una pluralidad de fotogramas de vídeo que comprenden al menos un fotograma predecible seleccionado de un grupo de fotogramas predecibles que comprenden un fotograma P y un fotograma B, comprendiendo el método;

25 determinar la información de movimiento asociada con al menos una parte de al menos un fotograma de referencia;  
 almacenar en una memoria intermedia dicha información de movimiento;  
 determinar los datos de identificación de modo que identifican al menos una parte de un fotograma predecible que puede obtenerse directamente usando al menos dicha información de movimiento almacenada en una memoria intermedia; y  
 generar dicha parte de dicho fotograma predecible usando dicha información de movimiento almacenada en una memoria intermedia.

30 Punto 58: El método de acuerdo con el punto 57, en el que dichos datos de identificación de modo definen un tipo de modelo de predicción necesaria para decodificar dicha parte codificada de dicho fotograma predecible.

Punto 59: El método de acuerdo con el punto 58, en el que dicho tipo de modelo de predicción incluye un modelo de predicción directa mejorado que incluye al menos un submodo seleccionado de un grupo que comprende un submodo de proyección de movimiento, un submodo de predicción de vectores de movimiento espacial, y un submodo de promedio ponderado.

35 Punto 60: El método de acuerdo con el punto 59, en el que dichos datos de identificación de modo identifican dicho al menos un submodo.

Punto 61: El método de acuerdo con el punto 57, en el que dicha parte de dicho fotograma de referencia incluye unos datos para al menos un píxel dentro de dicho fotograma de referencia, y dicha parte de dicho fotograma predecible incluye unos datos para al menos un píxel dentro de dicho fotograma predecible.

40 Punto 62: El método de acuerdo con el punto 61, en el que dicha parte de dicho fotograma de referencia incluye los datos de al menos una parte de un macrobloque dentro de dicho fotograma de referencia, y dicha parte de dicho fotograma predecible incluye los datos para al menos una parte de un macrobloque dentro de dicho fotograma predecible.

45 Punto 63: El método de acuerdo con el punto 57, en el que dicho fotograma de referencia precede temporalmente a dicho fotograma predecible dentro de dicha secuencia de fotogramas de vídeo.

Punto 64: El método de acuerdo con el punto 57, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información seleccionada de un grupo que comprende una información de velocidad y una información de aceleración.

50 Punto 65: El método de acuerdo con el punto 57, en el que dicha parte de dicho fotograma de referencia y dicha parte de dicho fotograma predecible están correlacionadas espacialmente.

Punto 66: El método de acuerdo con el punto 57, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información seleccionada de un grupo que comprende una información de proyección de píxeles, una información de predicción de vectores de movimiento espacial, una

información de proyección de píxeles y de predicción de vectores de movimiento espacial combinadas y una información de predicción de múltiples hipótesis.

5 Punto 67: El procedimiento de acuerdo en el punto 57, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia es nula y dichos datos de identificación de modo identifican que dicha parte de dicho fotograma de referencia incluye dicha parte de dicho fotograma predecible.

Punto 68: Un medio legible por ordenador que tiene instrucciones que pueden implementarse en ordenador para realizar acciones que comprenden:

10 decodificar los datos de vídeo codificados que incluyen una pluralidad de fotogramas de vídeo que comprenden al menos un fotograma predecible seleccionado de un grupo de fotogramas predecibles que comprenden un fotograma P y un fotograma B, mediante:

almacenar en una memoria intermedia una información de movimiento asociada con al menos una parte de al menos un fotograma de referencia;

15 determinar los datos de identificación de modo que identifican que al menos una parte de un fotograma predecible puede obtenerse directamente usando al menos dicha información de movimiento almacenada en una memoria intermedia ; y

generar dicha parte de dicho fotograma predecible usando dicha información de movimiento almacenada en una memoria intermedia.

20 Punto 69: El medio legible por ordenador de acuerdo con el punto 68, en el que dichos datos de identificación de modo definen un tipo de modelo de predicción necesaria para decodificar dicha parte codificada de dicho fotograma predecible.

Punto 70: El medio legible por ordenador de acuerdo con el punto 69, en el que dicho tipo de modelo de predicción incluye un modelo de predicción directa mejorado que incluye al menos un submodo seleccionado de un grupo que comprende un submodo de proyección de movimiento, un submodo de predicción de vectores de movimiento espacial, y un submodo de promedio ponderado

25 Punto 71: El medio legible por ordenador de acuerdo con el punto 70, en el que dichos datos de identificación de modo identifican dicho al menos un submodo.

Punto 72: El medio legible por ordenador de acuerdo con el punto 68, en el que dicha parte de dicho fotograma de referencia incluye unos datos para al menos un píxel dentro de dicho fotograma de referencia, y dicha parte de dicho fotograma predecible incluye unos datos para al menos un píxel dentro de dicho fotograma predecible.

30 Punto 73: El medio legible por ordenador de acuerdo con el punto 72, en el que dicha parte de dicho fotograma de referencia incluye unos datos de al menos una parte de un macrobloque dentro de dicho fotograma de referencia, y dicha parte de dicho fotograma predecible incluye unos datos para al menos una parte de un macrobloque dentro de dicho fotograma predecible.

35 Punto 74: El medio legible por ordenador de acuerdo con el punto 68, en el que dicho fotograma de referencia precede temporalmente a dicho fotograma predecible dentro de dicha secuencia de fotogramas de vídeo.

Punto 75: El medio legible por ordenador de acuerdo con el punto 68, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información seleccionada de un grupo que comprende una información de velocidad y una información de aceleración.

40 Punto 76: El medio legible por ordenador de acuerdo con el punto 68, en el que dicha parte de dicho fotograma de referencia y dicha parte de dicho fotograma predecible están correlacionadas espacialmente.

45 Punto 77: El medio legible por ordenador de acuerdo con el punto 68, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información seleccionada de un grupo que comprende una información de proyección de píxeles, una predicción de vectores de movimiento espacial, una información de proyección de píxeles y de predicción de vectores de movimiento espacial combinadas y una información de predicción de múltiples hipótesis.

Punto 78: El medio legible por ordenador de acuerdo con el punto 68, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia es nula y dichos datos de identificación de modo identifican que dicha parte de dicho fotograma de referencia incluye dicha parte de dicho fotograma predecible.

50 Elemento 79: Un aparato para su uso en decodificar los datos de vídeo para una secuencia de fotogramas de vídeo en una pluralidad de fotogramas de vídeo que incluye al menos un fotograma predecible seleccionado de un grupo de fotogramas predecibles que comprenden un fotograma P y un fotograma B, comprendiendo dicho aparato:

una memoria para almacenar la información de movimiento; y

5 una lógica acoplada operativamente a dicha memoria y configurada para almacenar en dicha memoria la información de movimiento asociada con al menos una parte de al menos un fotograma de referencia, comprobar que los datos de identificación de modo identifican que al menos una parte de un fotograma predecible puede obtenerse directamente usando al menos dicha información de movimiento almacenada en una memoria intermedia, y generar dicha parte de dicho fotograma predecible usando dicha información de movimiento almacenada en una memoria intermedia.

Punto 80: El aparato de acuerdo con el punto 79, en el que dichos datos de identificación de modo definen un tipo de modelo de predicción necesaria para decodificar dicha parte codificada de dicho fotograma predecible.

10 Punto 81: El aparato de acuerdo con el punto 80, en el que dicho tipo de modelo de predicción incluye un modelo de predicción directa mejorado que incluye al menos un submodo seleccionado de un grupo que comprende un submodo de proyección de movimiento, un submodo de predicción de vectores de movimiento espacial, y un submodo de promedio ponderado

15 Punto 82: El aparato de acuerdo con el punto 81, en el que dichos datos de identificación de modo identifican dicho al menos un submodo.

Punto 83: El aparato de acuerdo con el punto 79, en el que dicha parte de dicho fotograma de referencia incluye unos datos para al menos un píxel dentro de dicho fotograma de referencia, y dicha parte de dicho fotograma predecible incluye unos datos para al menos un píxel dentro de dicho fotograma predecible.

20 Punto 84: El aparato de acuerdo con el punto 83, en el que dicha parte de dicho fotograma de referencia incluye unos datos de al menos una parte de un macrobloque dentro de dicho fotograma de referencia, y dicha parte de dicho fotograma predecible incluye unos datos para al menos una parte de un macrobloque dentro de dicho fotograma predecible.

Punto 85: El aparato de acuerdo con el punto 79, en el que dicho fotograma de referencia precede temporalmente a dicho fotograma predecible dentro de dicha secuencia de fotogramas de vídeo.

25 Punto 86: El aparato de acuerdo con el punto 79, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información seleccionada de un grupo que comprende una información de velocidad y una información de aceleración.

Punto 87: El aparato de acuerdo con el punto 79, en el que dicha parte de dicho fotograma de referencia y dicha parte de dicho fotograma predecible están correlacionados espacialmente.

30 Punto 88: El aparato de acuerdo con el punto 79, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia incluye una información seleccionada de un grupo que comprende una información de proyección de píxeles, una predicción de vectores de movimiento espacial, una información de proyección de píxeles y de predicción de vectores de movimiento espacial combinadas y una información de predicción de múltiples hipótesis.

35 Punto 89: El aparato de acuerdo con el punto 79, en el que dicha información de movimiento asociada con dicha parte de dicho fotograma de referencia es nula y dichos datos de identificación de modo identifican que dicha parte de dicho fotograma de referencia incluye dicha parte de dicho fotograma predecible.

Punto 90: El aparato de acuerdo con el punto 79, en el que dicha lógica incluye un codec.

**REIVINDICACIONES**

1. Un procedimiento de codificación de vídeo, comprendiendo el procedimiento:

codificar un fotograma P de vídeo predictivo actual con un codificador (1202) de vídeo, para producir un vídeo (1210) codificado en un flujo de bits, en el que la codificación incluye:

5 seleccionar un modo de macrobloque de entre varios modos de macrobloque posibles para el fotograma P actual, en el que el fotograma P actual tiene varios fotogramas de referencia, y en el que los varios modos de macrobloque incluyen:

- un INTRA modo en el que se intracodifica un macrobloque actual del fotograma P actual;
- un primer INTER modo en el que el macrobloque actual se codifica en relación con uno de los varios fotogramas de referencia usando un movimiento cero, en el que el fotograma P actual tiene un fotograma anterior entre los varios fotogramas de referencia, y en el que, de acuerdo con el primer INTER modo, el macrobloque actual se copia del fotograma anterior usando un movimiento cero;
- un segundo INTER modo en el que el macrobloque actual se codifica con una compensación de movimiento en relación con uno de los varios fotogramas de referencia usando la información de vector de movimiento que se codifica en el flujo de bits, en el que, de acuerdo con el segundo INTER modo, la información de vector de movimiento para el macrobloque actual se codifica en el flujo de bits como una información de vector de movimiento diferencial;
- un tercer INTER modo en el que el macrobloque actual se codifica con una compensación de movimiento usando el modo y la información de vector de movimiento del macrobloque colocado en el fotograma anterior, en el que si dicho macrobloque colocado también se codifica en el tercer INTER modo, se usa la información de vector de movimiento para dicho macrobloque colocado; y

codificar el macrobloque actual de acuerdo con el modo de macrobloque seleccionado; y

emitir, desde el codificador de vídeo, el vídeo codificado en el flujo de bits.

25 2. El procedimiento de la reivindicación 1, en el que de acuerdo con el segundo INTER modo, la codificación de la información de vector de movimiento para el macrobloque actual incluye:

- predecir la información de vector de movimiento para el macrobloque actual;
- determinar la información de vector de movimiento diferencial para el macrobloque actual basándose en las diferencias entre la información de vector de movimiento y la información de vector de movimiento predicha para el macrobloque actual; y
- 30 codificar mediante entropía la información de vector de movimiento diferencial para el macrobloque actual.

3. Un procedimiento de decodificación de vídeo, comprendiendo el procedimiento:

recibir un vídeo (1210) codificado en un flujo de bits; y  
 decodificar un fotograma P predictivo actual usando el vídeo codificado con un decodificador (1302) de vídeo, en el que la decodificación incluye:

35 seleccionar un modo de macrobloque de entre varios modos de macrobloque posibles para el fotograma P actual, en el que el fotograma P actual tiene varios fotogramas de referencia, y en el que los varios modos de macrobloque incluyen:

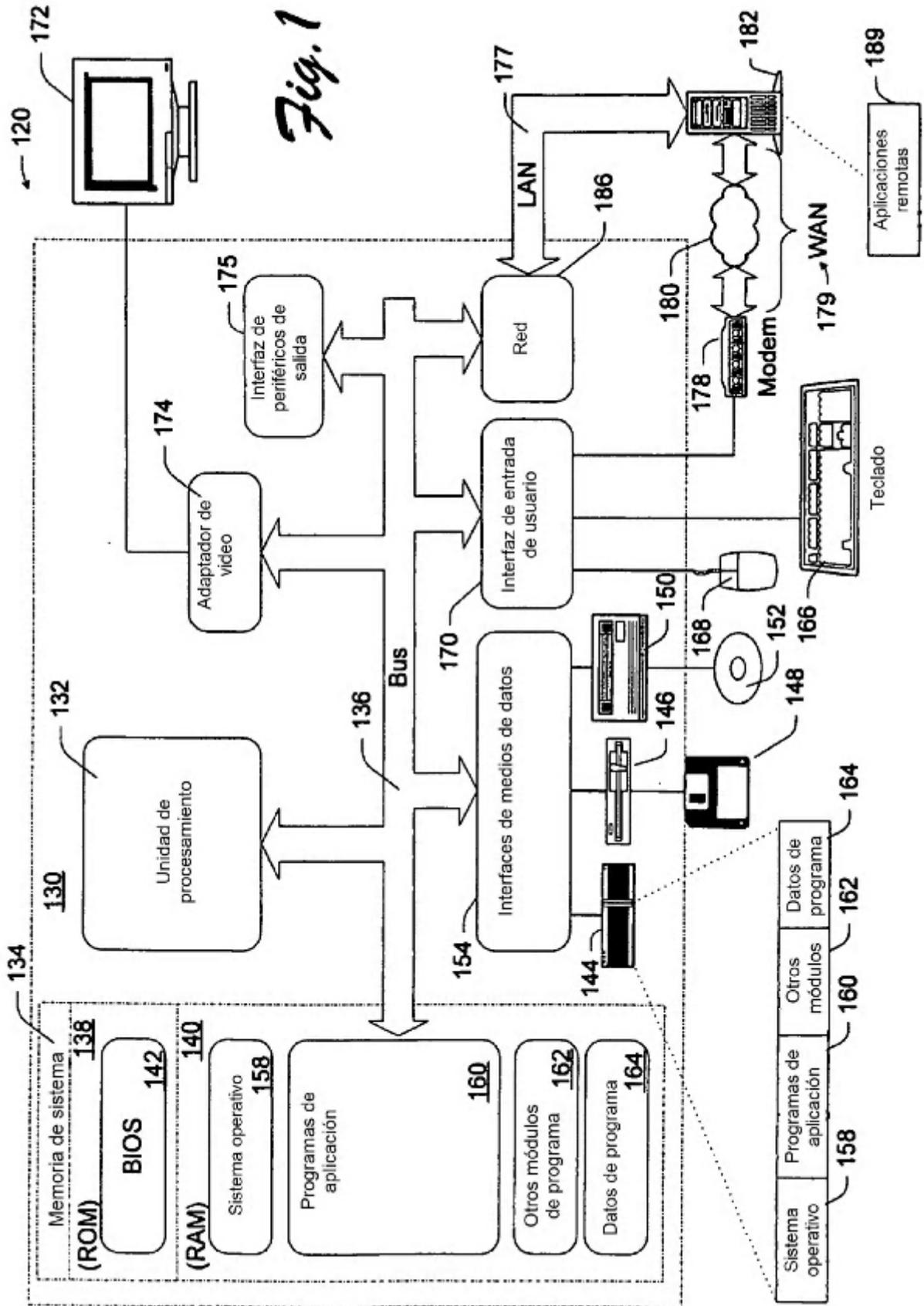
- un INTRA modo en el que se intradecodifica un macrobloque actual del fotograma P actual;
- un primer INTER modo en el que el macrobloque actual se decodifica en relación con uno de los varios fotogramas de referencia usando un movimiento cero, en el que el fotograma P actual tiene un fotograma anterior entre los varios fotogramas de referencia, y en el que, de acuerdo con el primer INTER modo, el macrobloque actual se copia del fotograma anterior usando un movimiento cero;
- un segundo INTER modo en el que el macrobloque actual se decodifica con una compensación de movimiento en relación con uno de los varios fotogramas de referencia usando la información de vector de movimiento que se decodifica del flujo de bits, en el que, de acuerdo con el segundo INTER modo, la información de vector de movimiento para el macrobloque actual se decodifica en el flujo de bits como una información de vector de movimiento diferencial; y
- un tercer INTER modo en el que el macrobloque actual se decodifica con una compensación de movimiento usando el modo y la información de vector de movimiento del macrobloque colocado en el fotograma anterior, en el que si dicho macrobloque colocado también se codifica en el tercer INTER modo, se usa la información de vector de movimiento para dicho macrobloque colocado; y

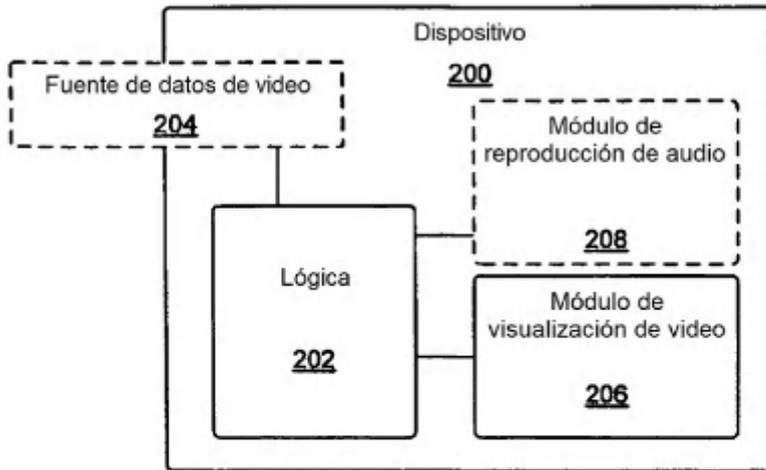
decodificar el macrobloque actual de acuerdo con el modo de macrobloque seleccionado.

4. El procedimiento de la reivindicación 3, en el que, de acuerdo con el segundo INTER modo, la decodificación de la información de vector de movimiento para el macrobloque actual incluye:

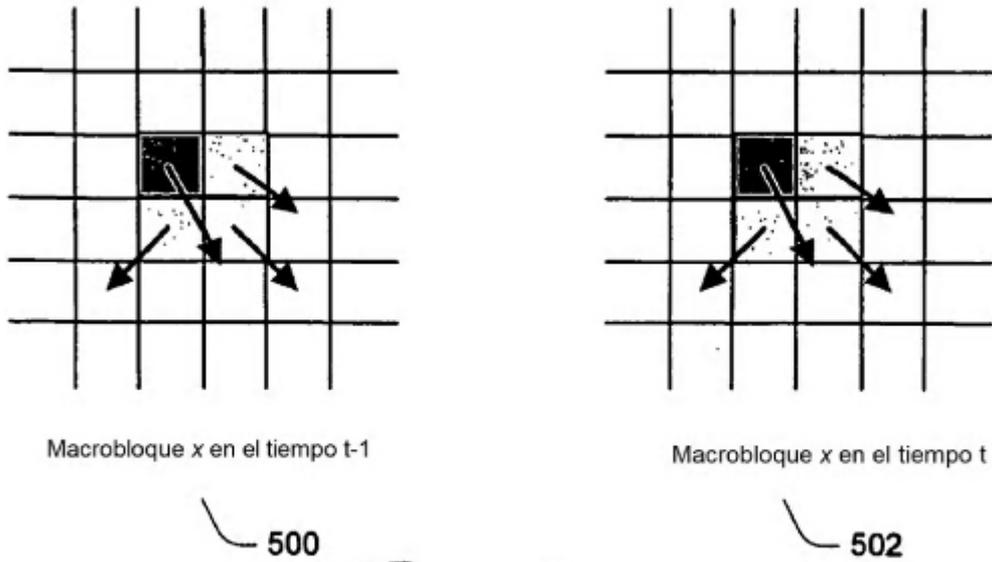
predecir la información de vector de movimiento para el macrobloque actual;  
 codificar mediante entropía la información de vector de movimiento diferencial para el macrobloque actual; y  
 combinar la información de vector de movimiento diferencial para el macrobloque actual con la información de  
 vector de movimiento predicha para el macrobloque actual.

- 5 5. El procedimiento de cualquiera de las reivindicaciones 1-4, en el que el vídeo codificado en el flujo de bits incluye una información de modo de macrobloque que indica el modo de macrobloque seleccionado.
6. El procedimiento de la reivindicación 5, en el que el vídeo codificado en el flujo de bits incluye además una información de submodos.
- 10 7. El procedimiento de acuerdo con la reivindicación 1, en el que la selección del modo de macrobloque para el macrobloque actual del fotograma P actual se basa en la optimización tasa-distorsión.
8. El procedimiento de acuerdo con la reivindicación 1, en el que la selección del modo de macrobloque para el macrobloque actual del fotograma P actual se basa en los requisitos de tasa de bits.
9. El procedimiento de acuerdo con la reivindicación 1, en el que la selección del modo de macrobloque para el macrobloque actual del fotograma P actual se basa en los requisitos de distorsión.
- 15 10. El procedimiento de acuerdo con la reivindicación 1, en el que la selección del modo de macrobloque incluye pasar por alto selectivamente la estimación de movimiento para el macrobloque actual del fotograma P actual.
11. Un aparato (200) de procesamiento de vídeo adaptado para realizar las etapas del procedimiento de cualquiera de las reivindicaciones 1-10.
- 20 12. El aparato de acuerdo con la reivindicación 11, en el que el aparato comprende un procesador, una memoria, una pantalla, un sistema operativo, un micrófono, una cámara y una conexión de red.
13. El aparato de acuerdo con la reivindicación 12, en el que el aparato comprende una lógica (202) implementada en hardware para realizar el procedimiento.
14. El aparato de acuerdo con la reivindicación 12, en el que el aparato comprende una lógica (202) implementada en software para realizar el procedimiento.
- 25 15. El aparato de acuerdo con la reivindicación 11, en el que el aparato comprende una lógica (202) que implementa unos módulos que incluyen un módulo (1014, 1108) de cuantificación inversa, un módulo (1018, 1112) de transformada de coseno discreta inversa, un filtro (1024, 1116) de bucle, una memoria (1026, 1118) intermedia de fotograma, un módulo (1022, 1110) de compensación de movimiento y una memoria (1206, 1308) intermedia de vector de movimiento.
- 30 16. Uno o más medios legibles por ordenador que tienen almacenados en los mismos unas instrucciones ejecutables por ordenador que, cuando se ejecutan en un dispositivo informático, hacen que el dispositivo informático realice las etapas del procedimiento de cualquiera de las reivindicaciones 1-10.





*Fig. 2*



*Fig. 5*

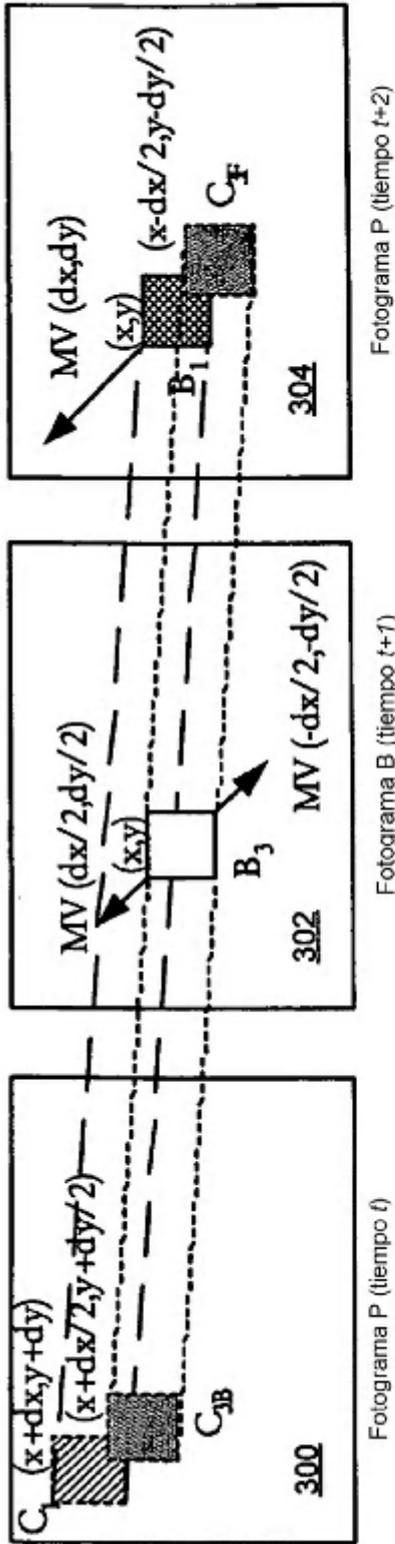


Fig. 3

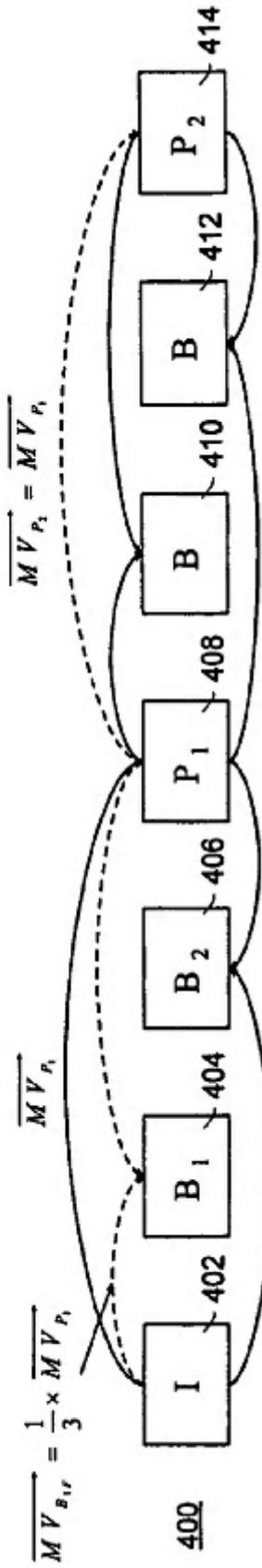
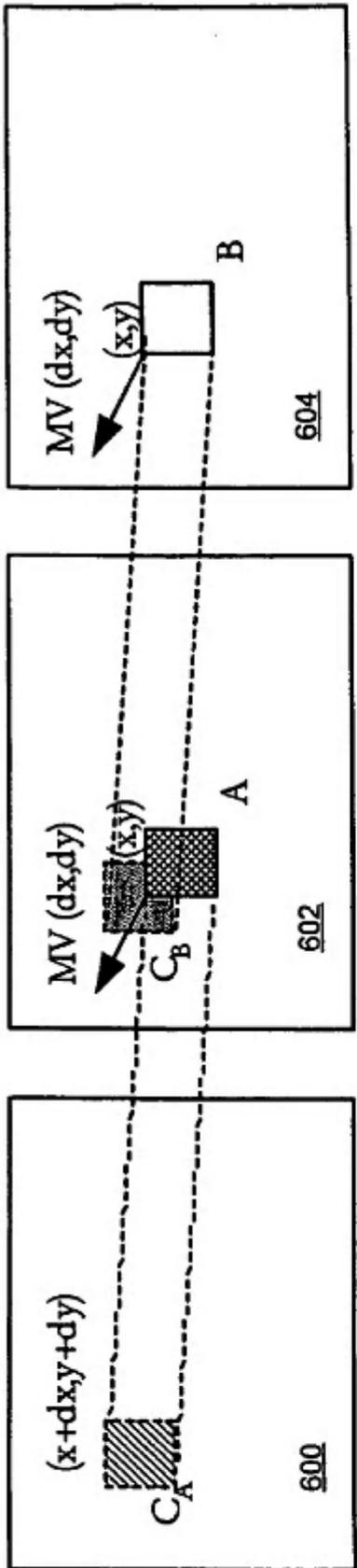


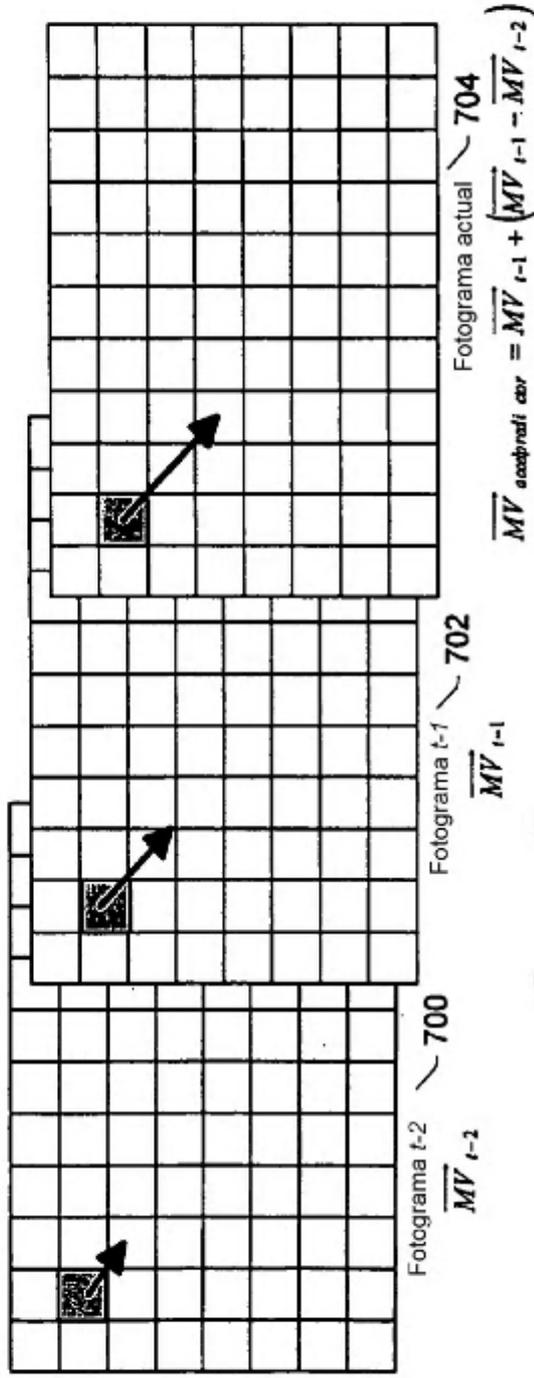
Fig. 4



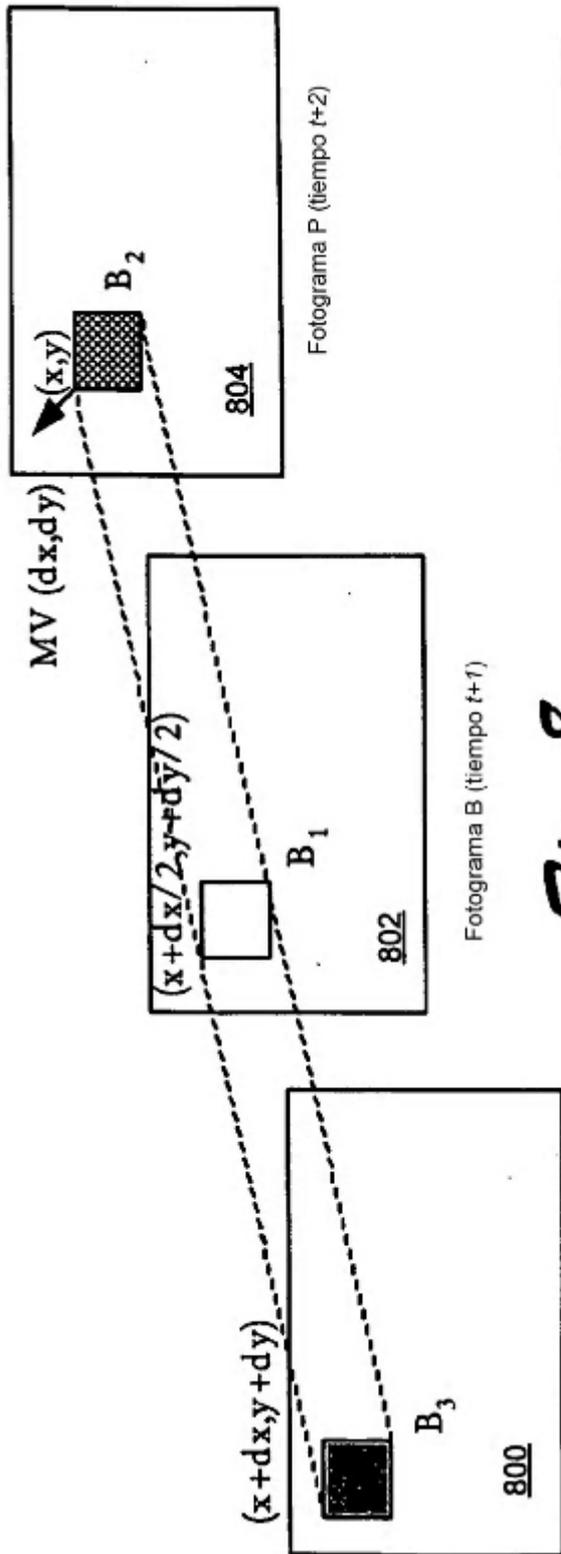
Fotograma P (tiempo t+2)

Fotograma P (tiempo t+1)

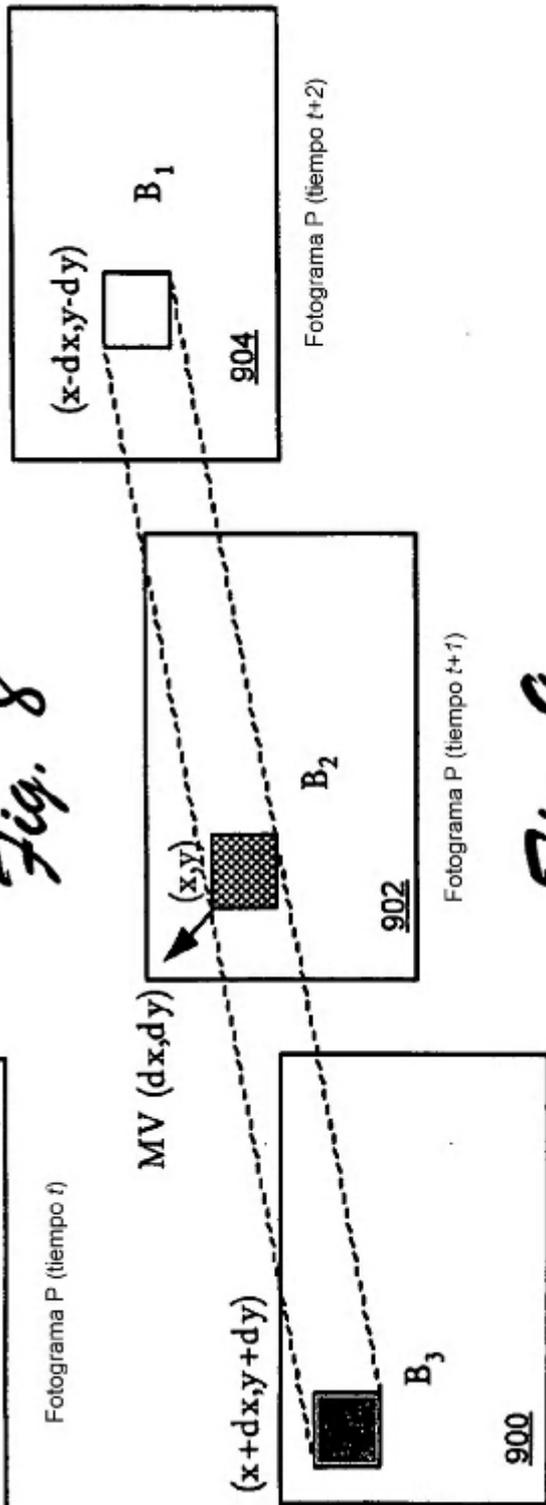
Fotograma P (tiempo t)



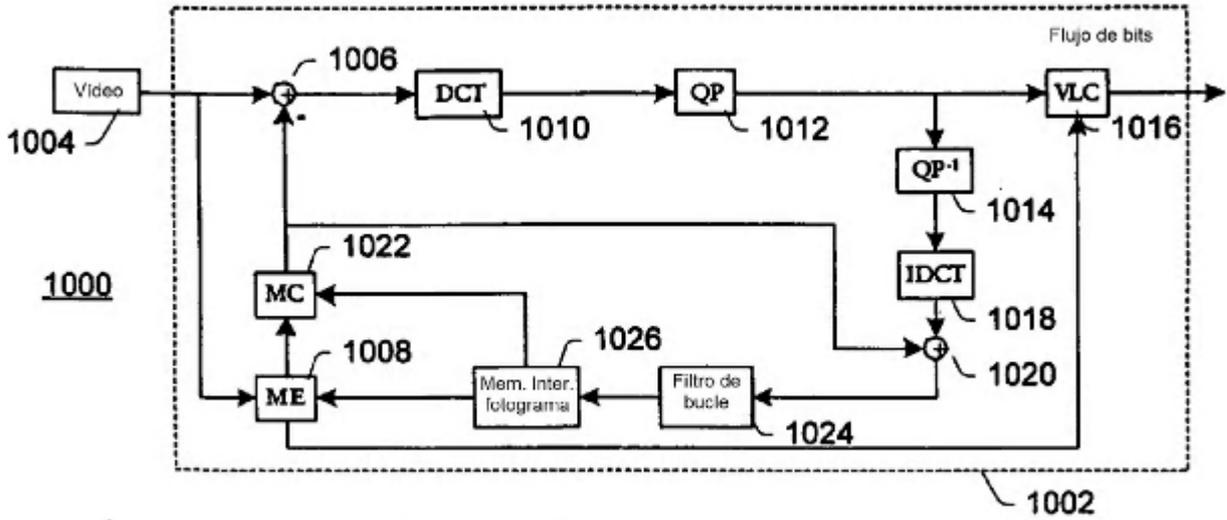
$$\overline{MV}_{actuali\ cabr} = \overline{MV}_{t-1} + (\overline{MV}_{t-1} - \overline{MV}_{t-2})$$



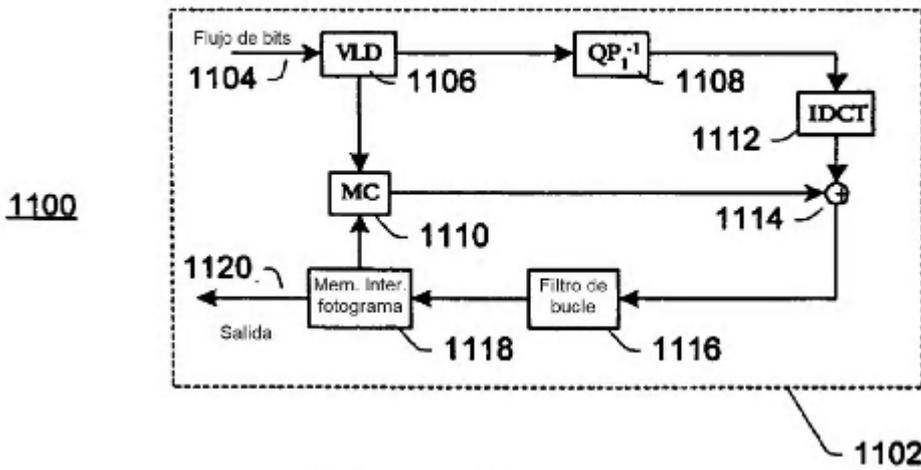
*Fig. 8*



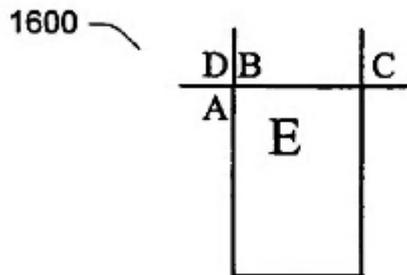
*Fig. 9*



*Fig. 10 (Técnica anterior)*

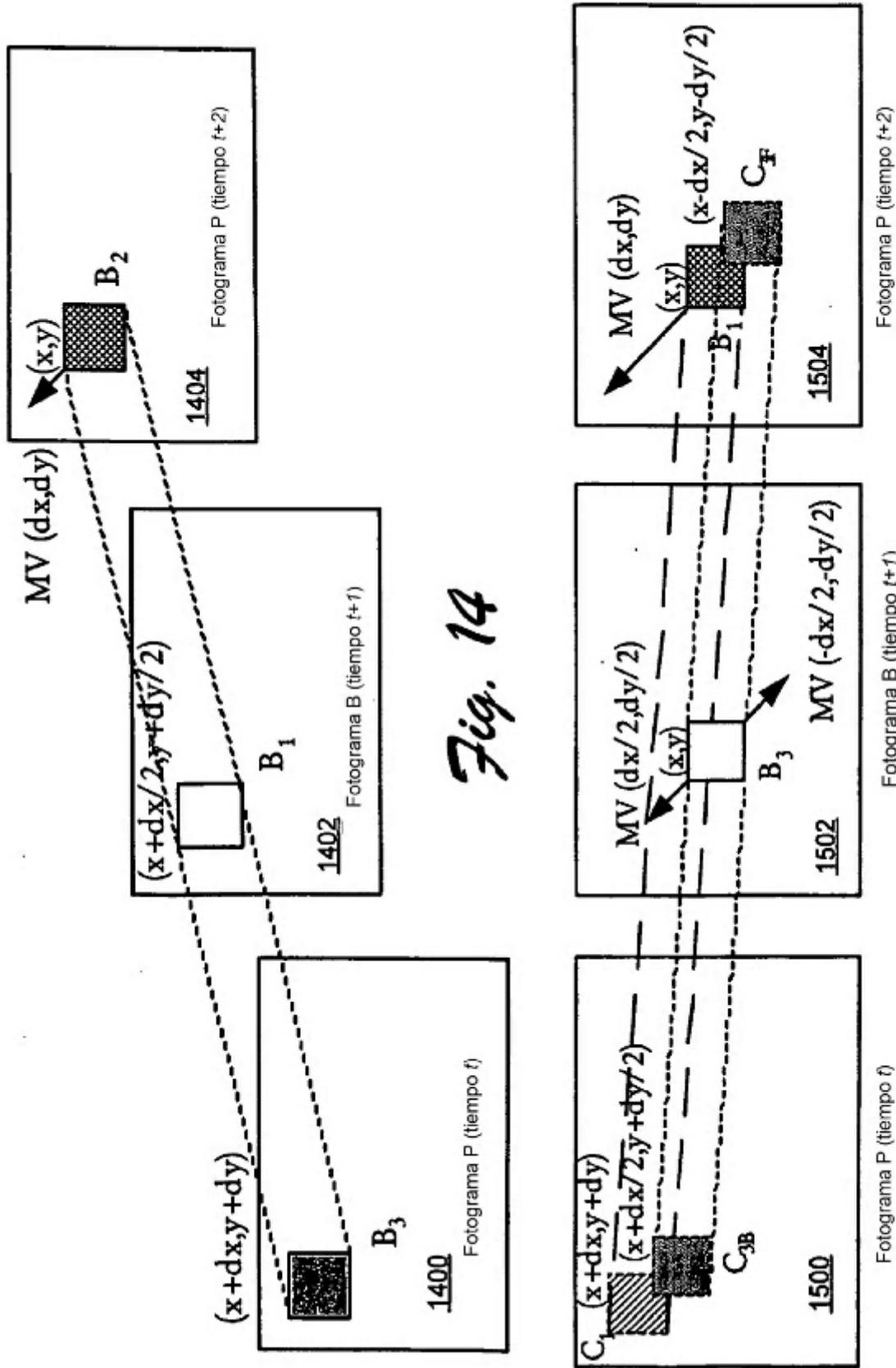


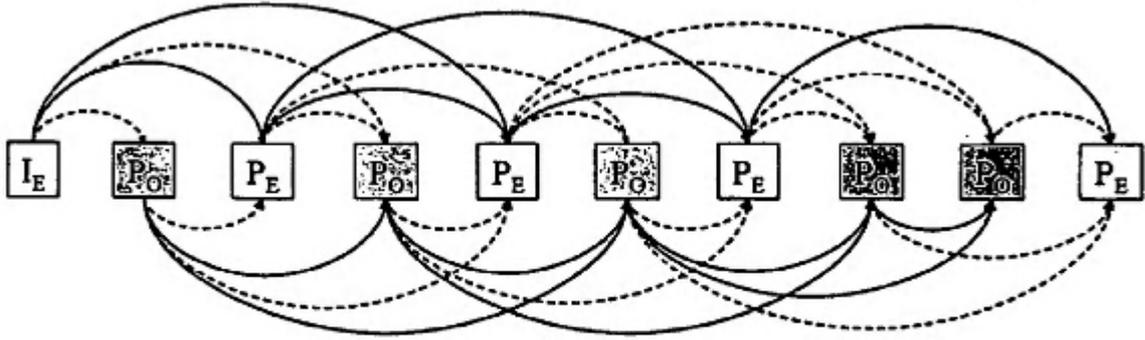
*Fig. 11 (Técnica anterior)*



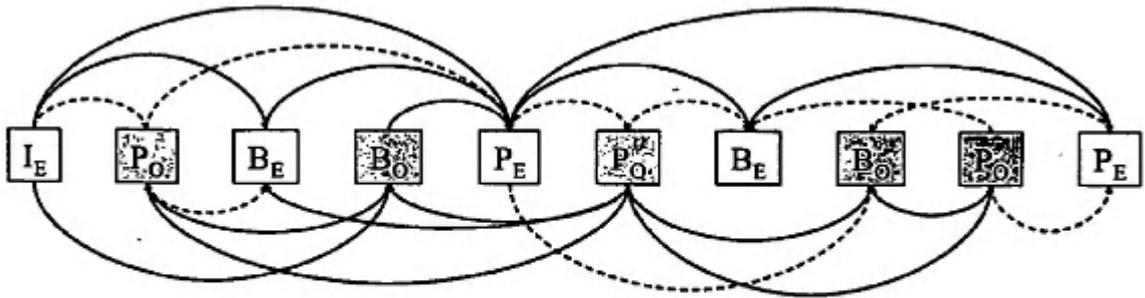
*Fig. 16*



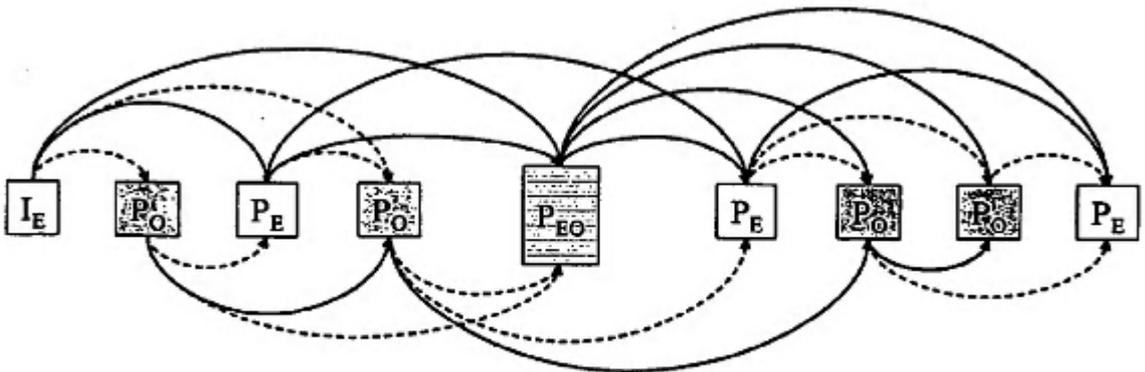




*Fig. 17*



*Fig. 18*



*Fig. 19*