

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 393 544**

51 Int. Cl.:

H04B 1/66 (2006.01)

H04N 7/26 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **11155441 .6**

96 Fecha de presentación: **11.07.2003**

97 Número de publicación de la solicitud: **2343829**

97 Fecha de publicación de la solicitud: **13.07.2011**

54 Título: **Procedimiento y aparato de especificación de temporización entre imágenes de precisión variable para la codificación de video digital**

30 Prioridad:

08.11.2002 US 291320

15.07.2002 US 396363 P

45 Fecha de publicación de la mención BOPI:

26.12.2012

45 Fecha de la publicación del folleto de la patente:

26.12.2012

73 Titular/es:

APPLE INC. (100.0%)

1 Infinite Loop

Cupertino, CA 95014, US

72 Inventor/es:

HASKELL, BARIN G.;

SINGER, DAVID W.;

DUMITRAS, ADRIANA y

PURI, ATUL

74 Agente/Representante:

DURÁN MOYA, Luis Alfonso

ES 2 393 544 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 especificación de temporización entre imágenes de precisión variable para la codificación de vídeo digital

5

CAMPO DE LA INVENCION

La presente invención se refiere al campo de los sistemas de compresión multimedia. En particular, la presente invención divulga procedimientos y sistemas para especificar temporización entre imágenes de precisión variable.

10

ANTECEDENTES DE LA INVENCION

Los formatos de medios electrónicos de base digital están finalmente a punto de reemplazar ampliamente los formatos de medios electrónicos analógicos. Los discos compactos digitales (CD) reemplazaron los registros analógicos de vinilo hace mucho tiempo. Las cintas de casete magnético analógico están volviéndose cada vez más raras. Los sistemas de audio digital de segunda y tercera generación, tales como los Minidiscos y el MP3 (Audio MPEG - capa 3) están quitando ahora una porción del mercado al formato de audio digital de primera generación de los discos compactos.

15

Los medios de vídeo han sido más lentos para avanzar hacia formatos de almacenamiento y transmisión digitales que el audio. Esto se ha debido en gran medida a las enormes cantidades de información digital requeridas para representar con precisión el vídeo en forma digital. Las enormes cantidades de información digital necesarias para representar con precisión el vídeo requieren sistemas de almacenamiento digital de muy alta capacidad y sistemas de transmisión de gran ancho de banda.

20

Sin embargo, el vídeo está avanzando ahora rápidamente hacia los formatos digitales de almacenamiento y transmisión. Los procesadores de ordenador más rápidos, los sistemas de almacenamiento de alta densidad y los nuevos algoritmos eficaces de compresión y codificación han hecho finalmente que el vídeo digital sea práctico en los puntos de venta al consumidor. El DVD (Disco Versátil Digital), un sistema de vídeo digital, ha sido uno de los productos electrónicos de consumo de mayores ventas en varios años. Los DVD han estado suplantando rápidamente a los Grabadores de Videocasetes (VCR) como el sistema favorito de reproducción de vídeo pregrabado, debido a su alta calidad de vídeo, muy alta calidad de audio, comodidad y características adicionales. El anticuado sistema analógico NTSC (Comité de Estándares Nacionales de Televisión) de transmisión de vídeo está actualmente en proceso de reemplazo por el sistema digital de transmisión de vídeo ATSC (Comité de Estándares Avanzados de Televisión).

25

30

35

Los sistemas informáticos han estado usando diversos formatos de codificación de vídeo digital durante un buen número de años. Entre los mejores sistemas de compresión y codificación de vídeo digital usados por los sistemas informáticos han estado los sistemas de vídeo digital respaldados por el Grupo de Expertos en Películas, popularmente conocido por el acrónimo MPEG. Los tres formatos de vídeo digital mejor conocidos y más frecuentemente usados del MPEG se conocen sencillamente como MPEG-1, MPEG-2 y MPEG-4. Los CD de vídeo (VCD) y los primeros sistemas de edición de vídeo digital para el nivel de consumo usan el primer formato de codificación de vídeo digital MPEG-1. Los Discos Versátiles Digitales (DVD) y el sistema de difusión de televisión del Satélite de Difusión Directa (DBS) de marca Dish Network usan el sistema de compresión y codificación de vídeo digital MPEG-2, de mayor calidad. El sistema de codificación MPEG-4 está siendo adoptado rápidamente por los más recientes codificadores de vídeo digital basados en ordenadores y los reproductores asociados de vídeo digital.

40

45

50

Los estándares MPEG-2 y MPEG-4 comprimen una serie de tramas de vídeo o campos de vídeo y luego codifican las tramas o campos comprimidos en un flujo de bits digital. Al codificar una trama o campo de vídeo con los sistemas MPEG-2 y MPEG-4, la trama o campo de vídeo se divide en una rejilla rectangular de macrobloques. Cada macrobloque se comprime y codifica independientemente.

55

Al comprimir una trama o campo de vídeo, el estándar MPEG-4 puede comprimir la trama o campo en uno entre tres tipos de tramas o campos comprimidos: Intratramas (I-tramas), tramas Predichas Unidireccionales (P-tramas) o tramas Predichas Bidireccionales (B-tramas). Las intratramas codifican independientemente por completo una trama de vídeo independiente, sin ninguna referencia a otras tramas de vídeo. Las P-tramas definen una trama de vídeo con referencia a una única trama de vídeo previamente exhibida. Las B-tramas definen una trama de vídeo con referencia tanto a una trama de vídeo exhibida antes de la trama actual como a una trama de vídeo a exhibir después de la trama actual. Debido a su utilización eficaz de información de vídeo redundante, las P-tramas y las B-tramas brindan generalmente la mejor compresión.

60

RESUMEN DE LA INVENCION

Se revela un procedimiento y aparato para la especificación de temporización entre imágenes de precisión variable para la codificación de vídeo digital. Específicamente, la presente invención divulga un sistema que permite que la temporización relativa de imágenes de vídeo contiguas se codifique de manera muy eficaz. En una realización, se

65

determina la diferencia de tiempo de exhibición entre una imagen de vídeo actual y una imagen de vídeo contigua. La diferencia de tiempo de exhibición se codifica luego en una representación digital de la imagen de vídeo. En una realización preferente, la imagen de vídeo contigua es la imagen almacenada más recientemente transmitida.

5 Para la eficacia de la codificación, la diferencia de tiempo de exhibición puede codificarse usando un sistema de codificación de longitud variable, o codificación aritmética. En una realización alternativa, la diferencia de tiempo de exhibición se codifica como una potencia de dos, a fin de reducir el número de bits transmitidos.

10 Otros objetos, características y ventajas de la presente invención serán evidentes a partir de los dibujos acompañantes y de la siguiente descripción detallada.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

15 Los objetos, características y ventajas de la presente invención serán evidentes a alguien versado en la tecnología, a la vista de la siguiente descripción detallada, en la cual:

La **figura 1** ilustra un diagrama en bloques de alto nivel de un posible sistema codificador de vídeo digital.

20 La **figura 2** ilustra una serie de imágenes de vídeo en el orden en que deberían exhibirse las imágenes, en donde las flechas que conectan las distintas imágenes indican la dependencia entre imágenes creada usando la compensación de movimiento.

25 La **figura 3** ilustra las imágenes de vídeo de la **figura 2** enumeradas en un orden de transmisión preferente de las imágenes, en donde las flechas que conectan distintas imágenes indican la dependencia entre imágenes creada usando la compensación de movimiento.

La **figura 4** ilustra gráficamente una serie de imágenes de vídeo en la cual las distancias entre imágenes de vídeo que mantienen referencias mutuas se escogen como potencias de dos.

30 DESCRIPCIÓN DETALLADA DE LA REALIZACIÓN PREFERENTE

Se divulga un procedimiento y sistema para especificar Temporización Entre Imágenes de Precisión Variable en un sistema de compresión y codificación multimedia. En la siguiente descripción, con fines explicativos, se estipula una nomenclatura específica para proporcionar una comprensión exhaustiva de la presente invención. Sin embargo, será evidente a alguien versado en la tecnología que estos detalles específicos no se requieren a fin de poner en práctica la presente invención. Por ejemplo, la presente invención ha sido descrita con referencia al sistema MPEG-4 de compresión y codificación multimedia. Sin embargo, las mismas técnicas pueden aplicarse fácilmente a otros tipos de sistemas de compresión y codificación.

40 **Panorama general de la compresión y codificación multimedia**

La **figura 1** ilustra un diagrama en bloques de alto nivel de un típico codificador **-100-** de vídeo digital, bien conocido en la tecnología. El codificador **-100-** de vídeo digital recibe un flujo entrante de vídeo de tramas **-105-** de vídeo a la izquierda del diagrama en bloques. Cada trama de vídeo es procesada por una unidad **-110-** de Transformación Discreta de Coseno (DCT). La trama puede procesarse independientemente (una intratrama) o con referencia a información de otras tramas recibidas desde la unidad de compensación de movimiento (una intertrama). Luego, una unidad **-120-** Cuantizadora (Q) cuantiza la información de la unidad **-110-** de Transformación Discreta de Coseno. Finalmente, la trama de vídeo cuantizada se codifica luego con una unidad **-180-** codificadora de entropía (H) para producir un flujo de bits codificado. La unidad **-180-** codificadora de entropía (H) puede usar un sistema de codificación de longitud variable (VLC).

55 Dado que una trama de vídeo codificada como intertrama está definida con referencia a otras tramas de vídeo contiguas, el codificador **-100-** de vídeo digital necesita crear una copia del aspecto en que aparecerá cada trama descodificada dentro de un descodificador de vídeo digital, de forma tal que las intertramas puedan codificarse. Así, la porción inferior del codificador **-100-** de vídeo digital es efectivamente un sistema descodificador de vídeo digital. Específicamente, una unidad **-130-** cuantizadora inversa (Q^{-1}) invierte la cuantización de la información de trama de vídeo y una unidad **-140-** de Transformación inversa Discreta de Coseno (DCT^{-1}) invierte la Transformación Discreta de Coseno de la información de la trama de vídeo. Una vez que todos los coeficientes de la DCT son reconstruidos a partir de la iDCT, la unidad de compensación de movimiento usará la información, junto con los vectores de movimiento, para reconstruir la trama codificada, que se usa luego como la trama de referencia para la estimación de movimiento de la próxima trama.

60 La trama de vídeo descodificada puede usarse luego para codificar intertramas (P-tramas o B-tramas) que se definen con respecto a información en la trama de vídeo descodificada. Específicamente, se usan una unidad **-150-** de compensación de movimiento (MC) y una unidad **-160-** de estimación de movimiento (ME) para determinar vectores de movimiento y generar valores diferenciales, usados para codificar intertramas.

Un controlador **-190-** de velocidad recibe información desde muchos componentes distintos en un codificador **-100-** de vídeo digital, y usa la información para adjudicar un presupuesto de bits para cada trama de vídeo. El controlador **-190-** de velocidad debería adjudicar el presupuesto de bits de manera tal que genere el flujo de bits de vídeo digital de máxima calidad que cumpla un conjunto especificado de restricciones. Específicamente, el controlador **-190-** de velocidad intenta generar el flujo de vídeo comprimido de máxima calidad sin desbordar los almacenes temporales (superar la cantidad de memoria disponible en un descodificador enviando más información de la que puede almacenarse) ni infranquear los almacenes temporales (no enviar tramas de vídeo con la rapidez suficiente, de manera que a un descodificador se le acaben las tramas de vídeo para exhibir).

Panorama general de la compresión y codificación multimedia

En algunas señales de vídeo el tiempo entre imágenes (tramas o campos) sucesivas de vídeo puede no ser constante. (Observación: Este documento usará el término imágenes de vídeo para referirse genéricamente a tramas de vídeo o a campos de vídeo). Por ejemplo, algunas imágenes de vídeo pueden perderse debido a restricciones del ancho de banda de transmisión. Además, la temporización de vídeo también puede variar debido a la irregularidad de la cámara o a efectos especiales, tales como la cámara lenta o la cámara rápida. En algunos flujos de vídeo, la fuente de vídeo original puede sencillamente tener tiempos no uniformes entre imágenes, por diseño. Por ejemplo, el vídeo sintetizado, tal como las animaciones gráficas por ordenador, pueden tener temporización no uniforme, ya que ninguna temporización arbitraria de vídeo es creada por un sistema de captura uniforme de vídeo tal como un sistema de cámara de vídeo. Un sistema flexible de codificación de vídeo digital debería ser capaz de asimilar la temporización no uniforme.

Muchos sistemas de codificación de vídeo digital dividen las imágenes de vídeo en una rejilla rectangular de macrobloques. Cada macrobloque individual de la imagen de vídeo se comprime y codifica independientemente. En algunas realizaciones, se usan subbloques de los macrobloques, conocidos como "bloques de píxeles". Tales bloques de píxeles pueden tener sus propios vectores de movimiento, que pueden interpolarse. Este documento se referirá a macrobloques, aunque las revelaciones de la presente invención pueden aplicarse igualmente tanto a macrobloques como a bloques de píxeles.

Algunos estándares de codificación de vídeo, p. ej., los estándares ISO MPEG o el estándar ITU H.264, usan distintos tipos de macrobloques predichos para codificar imágenes de vídeo. En un escenario, un macrobloque puede ser uno de entre tres tipos:

1. I-macrobloque - Un Intra (I) macrobloque no usa ninguna información de ninguna otra imagen de vídeo en su codificación (es completamente autodefinido);
2. P-macrobloque - Un macrobloque unidireccionalmente predicho (P) se refiere a información de imagen de una imagen precedente de vídeo; o
3. B-macrobloque - Un macrobloque bidireccional (B) predicho usa información de una imagen precedente y de una imagen futura de vídeo.

Si todos los macrobloques en una imagen de vídeo son Intra-macrobloques, entonces la imagen de vídeo es una Intratrama. Si una imagen de vídeo sólo incluye macrobloques predichos unidireccionales o intra-macrobloques, entonces la imagen de vídeo se conoce como una P-trama. Si la imagen de vídeo contiene macrobloques bidireccionales predichos cualesquiera, entonces la imagen de vídeo se conoce como una B-trama. Para simplificar, este documento considerará el caso en que todos los macrobloques dentro de una imagen dada son del mismo tipo.

Un ejemplo de secuencia de imágenes de vídeo a codificar podría representarse como

I₁ B₂ B₃ B₄ P₅ B₆ B₇ B₈ B₉ P₁₀ B₁₁ P₁₂ B₁₃ I₁₄...

donde la letra (I, P o B) representa si la imagen de vídeo es una I-trama, P-trama o B-trama, y el número representa el orden de cámara de la imagen de vídeo en la secuencia de imágenes de vídeo. El orden de cámara es el orden en el cual una cámara registró las imágenes de vídeo y, por ello, también es el orden en el cual las imágenes de vídeo deberían exhibirse (el orden de exhibición).

El ejemplo previo de serie de imágenes de vídeo se ilustra gráficamente en la **figura 2**. Con referencia a la **figura 2**, las flechas indican que los macrobloques de una imagen almacenada (I-trama o P-trama en este caso) se usan en la predicción compensada de movimiento de otras imágenes.

En el escenario de la **figura 2**, no se usa ninguna información de otras imágenes en la codificación de la imagen I₁ de vídeo de intratrama. La imagen P₅ de vídeo es una P-trama que usa información de vídeo de la imagen I₁ de vídeo previa en su codificación, por lo que se traza una flecha desde la imagen I₁ de vídeo a la imagen P₅ de vídeo. La imagen B₂ de vídeo, la imagen B₃ de vídeo y la imagen B₄ de vídeo usan todas información tanto de la imagen I₁ de vídeo como de la imagen P₅ de vídeo en su codificación, por lo que se trazan flechas desde la imagen I₁ de vídeo

y la imagen P₅ de vídeo a la imagen B₂ de vídeo, la imagen B₃ de vídeo y la imagen B₄ de vídeo. Como se ha afirmado anteriormente, los tiempos entre imágenes, en general, no son los mismos.

5 Como las imágenes B usan información de imágenes futuras (imágenes que se exhibirán más tarde), el orden de transmisión es usualmente distinto al orden de exhibición. Específicamente, las imágenes de vídeo que se necesitan para construir otras imágenes de vídeo deberían transmitirse en primer lugar. Para la secuencia anterior, el orden de transmisión podría ser

10 I₁ P₅ B₂ B₃ B₄ P₁₀ B₆ B₇ B₈ B₉ P₁₂ B₁₁ I₁₄ B₁₃...

La **figura 3** ilustra gráficamente el orden de transmisión anterior de las imágenes de vídeo de la **figura 2**. Nuevamente, las flechas en la figura indican que los macrobloques de una imagen de vídeo almacenada (I o P en este caso) se usan en la predicción compensada de movimiento de otras imágenes de vídeo.

15 Con referencia a la **figura 3**, el sistema transmite primero la I-trama I₁, que no depende de ninguna otra trama. Luego, el sistema transmite la imagen P₅ de vídeo de P-trama, que depende de la imagen I₁ de vídeo. Luego, el sistema transmite la imagen B₂ de vídeo de B-trama después de la imagen P₅ de vídeo, incluso aunque la imagen B₂ de vídeo se exhibirá antes de la imagen P₅ de vídeo. La razón para esto es que cuando llega el momento de descodificar B₂, el descodificador ya habrá recibido y almacenado la información en las imágenes I₁ y P₅ de vídeo, necesarias para descodificar la imagen B₂ de vídeo. De manera similar, las imágenes I₁ y P₅ de vídeo están listas para ser usadas a fin de descodificar las siguientes imagen B₃ de vídeo e imagen B₄ de vídeo. El receptor / descodificador reordena la secuencia de imágenes de vídeo para su visualización adecuada. En esta operación, las imágenes I y P se denominan a menudo imágenes almacenadas.

25 La codificación de las imágenes de P-tramas utiliza habitualmente Compensación de Movimiento, en donde se calcula un Vector de Movimiento para cada macrobloque en la imagen. Usando el vector de movimiento calculado, puede formarse un macrobloque de predicción (P-macrobloque) por traslación de píxeles en la precitada imagen anterior. La diferencia entre el macrobloque real en la imagen de P-trama y el bloque de predicción se codifica luego para su transmisión.

30 Cada vector de movimiento también puede transmitirse mediante la codificación predictiva. Por ejemplo, una predicción de vector de movimiento puede formarse usando vectores de movimiento contiguos. En tal caso, luego la diferencia entre el vector de movimiento real y la predicción del vector de movimiento se codifica para su transmisión.

35 Cada B-macrobloque usa dos vectores de movimiento: un primer vector de movimiento que hace referencia a la precitada imagen anterior de vídeo y un segundo vector de movimiento que hace referencia a la imagen futura de vídeo. A partir de estos dos vectores de movimiento se calculan dos macrobloques de predicción. Los dos macrobloques predichos se combinan luego entre sí, usando alguna función, para formar un macrobloque predicho final. Como antes, la diferencia entre el macrobloque efectivo en la imagen de B-trama y el macrobloque predicho final se codifica entonces para su transmisión.

45 Como con los P-macrobloques, cada vector de movimiento (MV) de un B-macrobloque puede transmitirse mediante codificación predictiva. Específicamente, un vector de movimiento predicho se forma usando vectores de movimiento contiguos. Luego, la diferencia entre el vector de movimiento efectivo y el predicho se codifica para su transmisión.

50 Sin embargo, con los B-macrobloques existe la oportunidad de interpolar vectores de movimiento a partir de vectores de movimiento en el macrobloque de imagen almacenada más cercano. Tal interpolación se lleva a cabo tanto en el codificador de vídeo digital como en el descodificador de vídeo digital.

55 Esta interpolación de vectores de movimiento funciona especialmente bien en imágenes de vídeo de una secuencia de vídeo donde una cámara está desplazándose horizontalmente con lentitud sobre un fondo estático. De hecho, tal interpolación de vectores de movimiento puede ser bastante buena como para ser usada por sí sola. Específicamente, esto significa que no se necesita calcular o transmitir ninguna información diferencial para estos vectores de movimiento de B- macrobloque, codificados usando interpolación.

60 Para ilustrar adicionalmente, en el escenario anterior representemos el tiempo de exhibición entre imágenes entre las imágenes i y j como D_{i,j}, es decir, si los tiempos de exhibición de las imágenes son T_i y T_j, respectivamente, entonces

$$D_{i,j} = T_i - T_j$$

de lo que se deduce que

$$65 \quad \begin{aligned} D_{i,k} &= D_{i,j} + D_{j,k} \\ D_{i,k} &= -D_{k,i} \end{aligned}$$

Obsérvese que $D_{i,j}$ puede ser negativo en algunos casos.

5 Así, si $MV_{5,1}$ es un vector de movimiento para un macrobloque P_5 con referencia desde I_1 , entonces, para los correspondientes macrobloques en B_2 , B_3 y B_4 , los vectores de movimiento con referencia desde I_1 y P_5 , respectivamente, se interpolarían por

$$\begin{aligned}
 10 \quad & MV_{2,1} = MV_{5,1} * D_{2,1} / D_{5,1} \\
 & MV_{3,1} = MV_{5,1} * D_{3,1} / D_{5,1} \\
 & MV_{4,1} = MV_{5,1} * D_{4,1} / D_{5,1}
 \end{aligned}$$

15 Obsérvese que, dado que se usan ratios de tiempos de exhibición para la predicción de vectores de movimiento, no se necesitan tiempos de exhibición absolutos. Por ello, pueden usarse tiempos de exhibición relativos para valores $D_{i,j}$ del tiempo de exhibición.

20 Este escenario puede generalizarse, como, por ejemplo, en el estándar H.264. En la generalización, una imagen P o B puede usar cualquier imagen previamente transmitida para su predicción de vector de movimiento. Así, en el caso anterior, la imagen B_3 puede usar la imagen I_1 y la imagen B_2 en su predicción. Además, los vectores de movimiento pueden extrapolarse, y no sólo interpolarse. Así, en este caso tendríamos.

$$25 \quad MV_{3,1} = MV_{2,1} * D_{3,1} / D_{2,1}$$

Tal extrapolación (o interpolación) de vectores de movimiento también puede usarse en el proceso de predicción para la codificación predictiva de vectores de movimiento.

30 En cualquier caso, el problema en el caso de tiempos no uniformes entre imágenes es transmitir los valores relativos del tiempo de exhibición $D_{i,j}$ al receptor, y ese el objeto de la presente invención. En una realización de la presente invención, para cada imagen después de la primera imagen transmitimos la diferencia de tiempo de exhibición entre la imagen actual y la imagen almacenada más recientemente transmitida. Para disponer de flexibilidad ante errores, la transmisión podría repetirse varias veces dentro de la imagen, p. ej., en las llamadas cabeceras de tajada de los estándares MPEG o H.264. Si se pierden todas las cabeceras de tajada, entonces, supuestamente, otras imágenes que dependen de la imagen perdida para descodificar información tampoco pueden descodificarse.

35 Así, en el escenario anterior transmitiríamos lo siguiente:

$$40 \quad D_{5,1} \ D_{2,5} \ D_{3,5} \ D_{4,5} \ D_{10,5} \ D_{6,10} \ D_{7,10} \ D_{8,10} \ D_{9,10} \ D_{12,10} \ D_{11,12} \ D_{14,12} \ D_{13,14} \dots$$

Con fines de estimación de vectores de movimiento, los requisitos de precisión para $D_{i,j}$ pueden variar de imagen a imagen. Por ejemplo, si hay sólo una única imagen B_6 de B-trama a medio camino entre las imágenes P_5 y P_7 de P-trama, entonces basta enviar sólo:

$$45 \quad D_{7,5} = 2 \text{ y } D_{6,7} = -1$$

donde los valores $D_{i,j}$ del tiempo de exhibición son valores de tiempos relativos. Si, en cambio, la imagen B_6 de vídeo está sólo a un cuarto de la distancia entre la imagen P_5 de vídeo y la imagen P_7 de vídeo, entonces los valores $D_{i,j}$ adecuados del tiempo de exhibición a enviar serían:

$$50 \quad D_{7,5} = 4 \text{ y } D_{6,7} = -1$$

Obsérvese que en ambos ejemplos precedentes, el tiempo de exhibición entre la imagen B_6 de vídeo y la imagen P_7 de vídeo se está usando como la "unidad" de tiempo de exhibición y que la diferencia de tiempo de exhibición entre la imagen P_5 de vídeo y la imagen P_7 de vídeo es de cuatro "unidades" de tiempo de exhibición.

60 En general, la estimación de vectores de movimiento es menos compleja si los divisores son potencias de dos. Esto se logra fácilmente en nuestra realización si el valor de $D_{i,j}$ (el tiempo entre imágenes) entre dos imágenes almacenadas se escoge como una potencia de dos, según se ilustra gráficamente en la **figura 4**. Alternativamente, el procedimiento de estimación podría definirse para truncar o redondear todos los divisores a una potencia de dos.

65 En el caso en que un tiempo entre imágenes es una potencia de dos, el número de bits de datos puede reducirse si sólo se transmite la potencia entera (de dos) en lugar del valor completo del tiempo entre imágenes. La **figura 4** ilustra gráficamente un caso en el cual las distancias entre imágenes se escogen como potencias de dos. En tal caso, el valor 2 del tiempo de exhibición $D_{3,1}$ entre la imagen P_1 de vídeo y la imagen P_3 de vídeo se transmite como

REIVINDICACIONES

1. Procedimiento que comprende:

5 la recepción de un flujo de bits que comprende una pluralidad de imágenes de vídeo codificadas y una pluralidad de valores de orden codificados, teniendo dicha pluralidad de imágenes de vídeo un primer orden secuencial separado con respecto a un segundo orden secuencial de las imágenes de vídeo del flujo de bits, siendo cada valor de orden un valor no temporal asociado con una imagen de vídeo y representativo de una posición de la imagen de vídeo en el primer orden secuencial de las imágenes de vídeo, comprendiendo la pluralidad de valores de orden una
10 secuencia de valores de orden no uniforme que comprende un primer valor de orden para una primera imagen, un segundo valor de orden para una segunda imagen, y un tercer valor de orden para una tercera imagen, en el que en el primer orden secuencial, la primera imagen es seguida por la segunda imagen, que es seguida por la tercera imagen, en el que la pluralidad de imágenes de vídeo codificadas no incluye imágenes que se encuentren entre la primera y tercera imágenes del primer orden secuencial distintas de la segunda imagen de vídeo, en el que la diferencia entre el segundo y tercer valores de orden es distinta que la diferencia entre el tercer y segundo valores de orden; y

decodificar las imágenes de vídeo codificadas utilizando dichos valores de orden.

20 2. Procedimiento, según la reivindicación 1, en el que como mínimo un valor de orden está comprimido por utilización de codificación de longitud variable.

3. Procedimiento, según la reivindicación 1, en el que el primer orden secuencial es una secuencia para visualizar las imágenes de vídeo.

25 4. Procedimiento, según la reivindicación 1, en el que la primera imagen de vídeo es una imagen de vídeo I que no comprende macrobloques que referencian otras imágenes de vídeo.

5. Procedimiento, según la reivindicación 1, que comprende además la decodificación de la tercera imagen de vídeo utilizando la segunda imagen de vídeo, en el que la segunda imagen de vídeo codificada comprende como mínimo un macrobloque de predicción bidireccional y la tercera imagen de vídeo codificada no comprende macrobloques de predicción bidireccional y como mínimo un macrobloque de predicción unidireccional que está codificado por referencia a un macrobloque de la segunda imagen de vídeo.

35 6. Decodificador que comprende:

medios para recibir un flujo de bits que comprende una pluralidad de imágenes de vídeo codificadas y una pluralidad de valores de orden codificados, teniendo dicha pluralidad de imágenes de vídeo un primer orden secuencial separado con respecto a un segundo orden secuencial de las imágenes de vídeo en el flujo de bits, siendo cada
40 valor de orden un valor no temporal asociado con una imagen de vídeo y representativo de una posición de la imagen de vídeo en el primer orden secuencial de las imágenes de vídeo, comprendiendo la pluralidad de valores de orden una secuencia de valores de orden no uniformes que comprende un primer valor de orden para una primera imagen, un segundo valor de orden para una segunda imagen, y un tercer valor de orden para la tercera imagen, en el que en el primer orden secuencial, la primera imagen es seguida por la segunda imagen, que es seguida por la tercera imagen, en el que la pluralidad de imágenes de vídeo codificadas no incluye ninguna imagen que se encuentre entre la primera y tercera imágenes en el primer orden secuencial distintas de la segunda imagen de vídeo, en el que la diferencia entre el segundo y primer valores de orden es distinta que la diferencia entre el tercer y segundo valores de orden; y

50 medios para decodificar las imágenes de vídeo codificadas utilizando dichos valores de orden.

7. Decodificador, según la reivindicación 6, en el que la decodificación comprende el cálculo de un vector de movimiento para decodificar como mínimo una imagen de vídeo utilizando los valores de orden.

55 8. Decodificador, según la reivindicación 6, en el que el cálculo del vector de movimiento comprende la realización de una operación de interpolación.

9. Decodificador, según la reivindicación 8, en el que el cálculo del vector de movimiento comprende la realización de una operación de extrapolación.

60 10 . Decodificador, según la reivindicación 6, en el que como mínimo un valor de orden es comprimido utilizando codificación de longitud variable.

11. Decodificador, según la reivindicación 6, en el que el primer orden secuencial es una secuencia para visualizar las imágenes de vídeo.

65

12. Decodificador, según la reivindicación 6, en el que la primera imagen de vídeo es una imagen de vídeo I que no comprende macrobloques que referencien otras imágenes de vídeo.
- 5 13. Decodificador, según la reivindicación 6, que comprende además medios para decodificar la tercera imagen de vídeo utilizando la segunda imagen de vídeo, en el que la segunda imagen de vídeo codificada comprende como mínimo un macrobloque de predicción bidireccional y la tercera imagen de vídeo codificada no comprende macrobloques bidireccionales de predicción y como mínimo un macrobloque de predicción unidireccional que está codificado por referencia a un macrobloque de la segunda imagen de vídeo.
- 10 14. Decodificador, según la reivindicación 6, en el que el primer orden secuencial es un orden de visualización de la pluralidad de imágenes de vídeo y el segundo orden secuencial es una orden de transmisión de la pluralidad de imágenes de vídeo.
- 15 15. Procedimiento, según la reivindicación 1, en el que el primer orden secuencial es un orden de visualización de la pluralidad de imágenes de vídeo y el segundo orden secuencial es un orden de transmisión de la pluralidad de imágenes de vídeo.

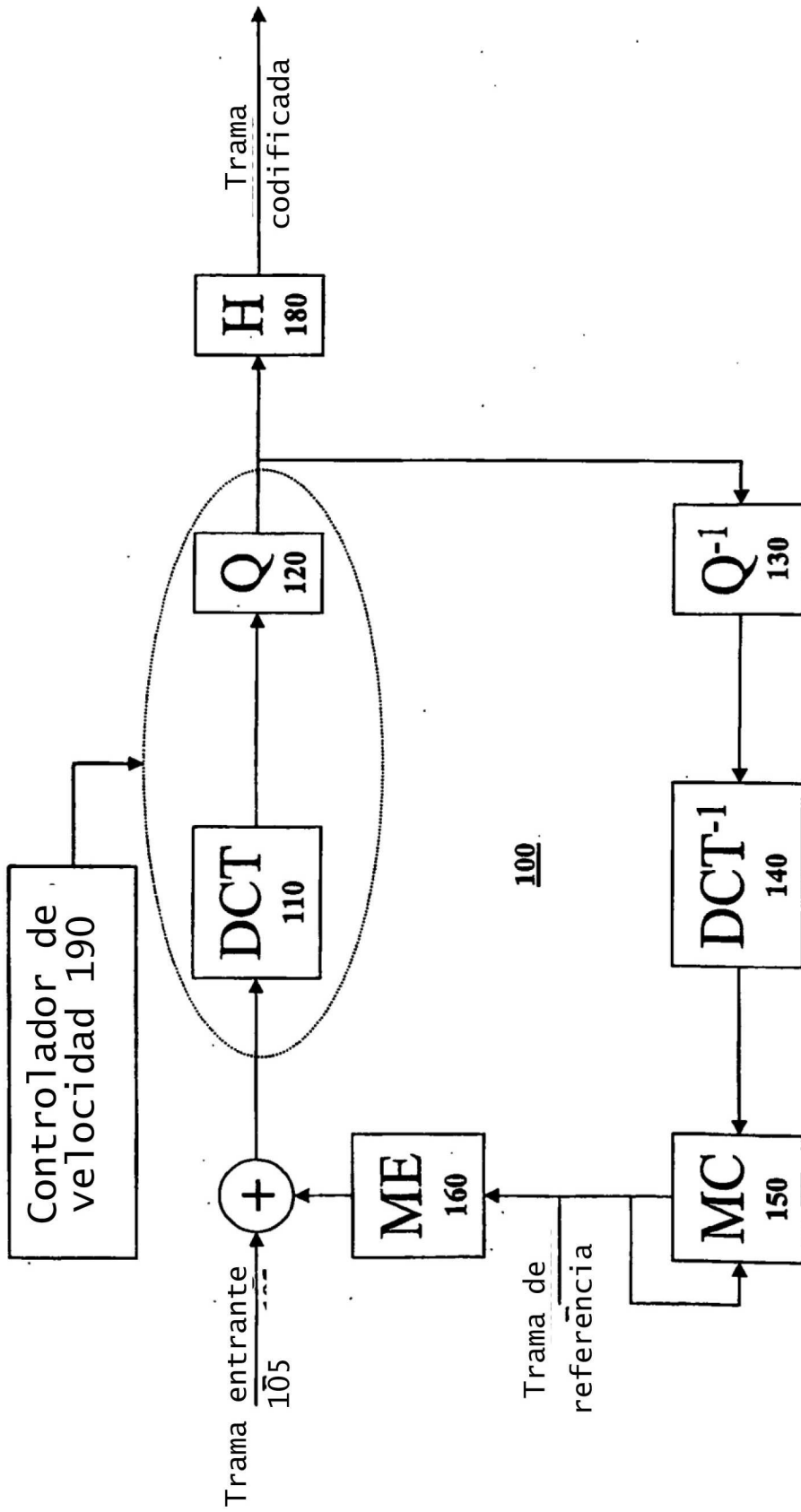


Figura 1

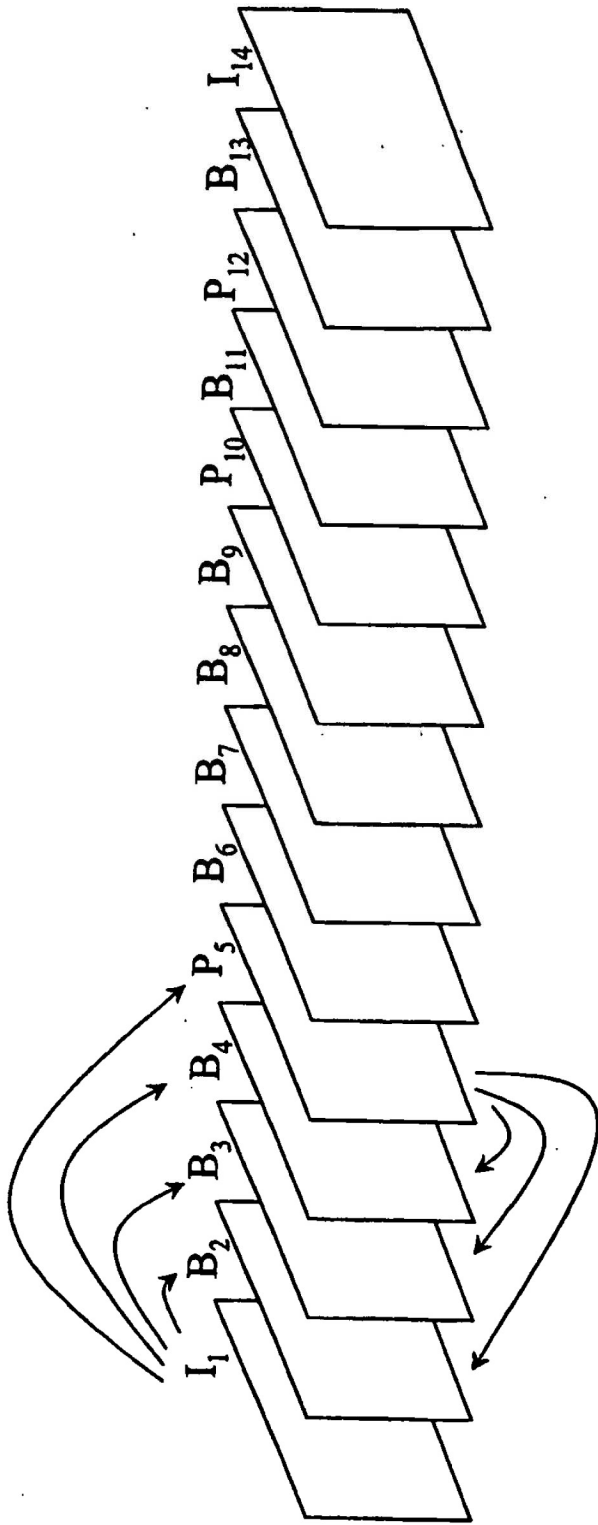


Figura 2

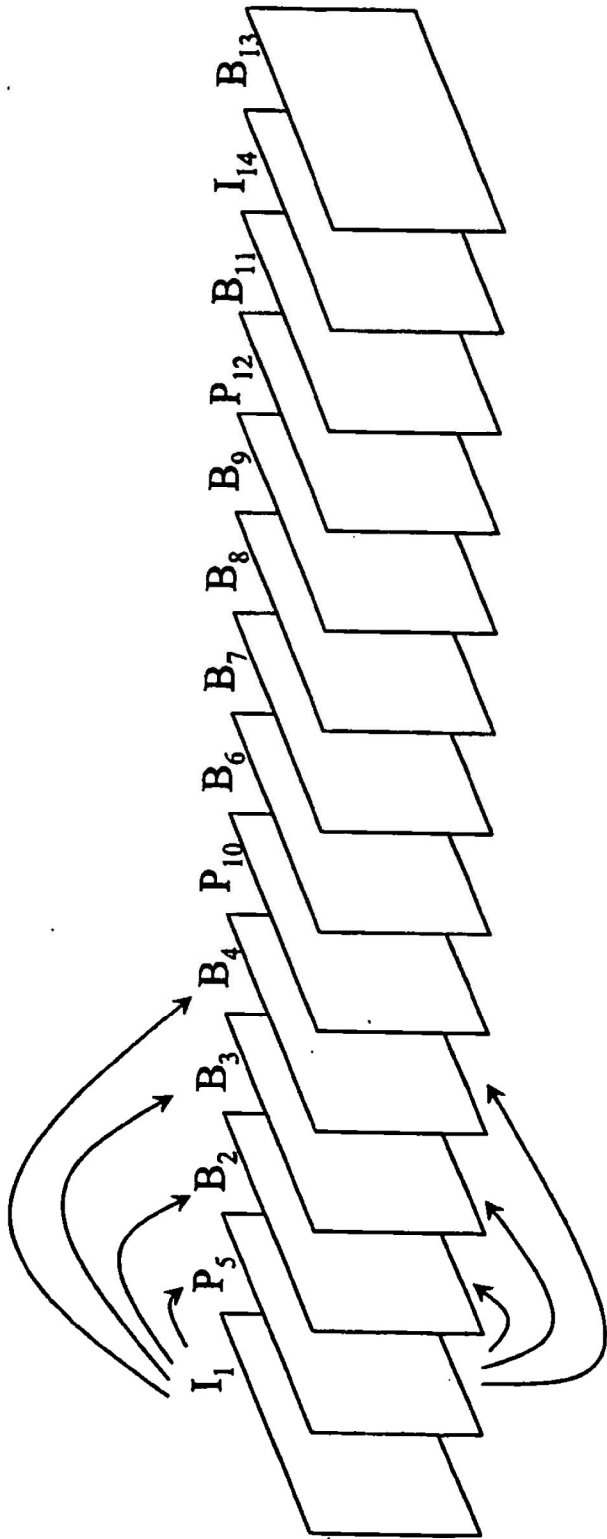


Figura 3

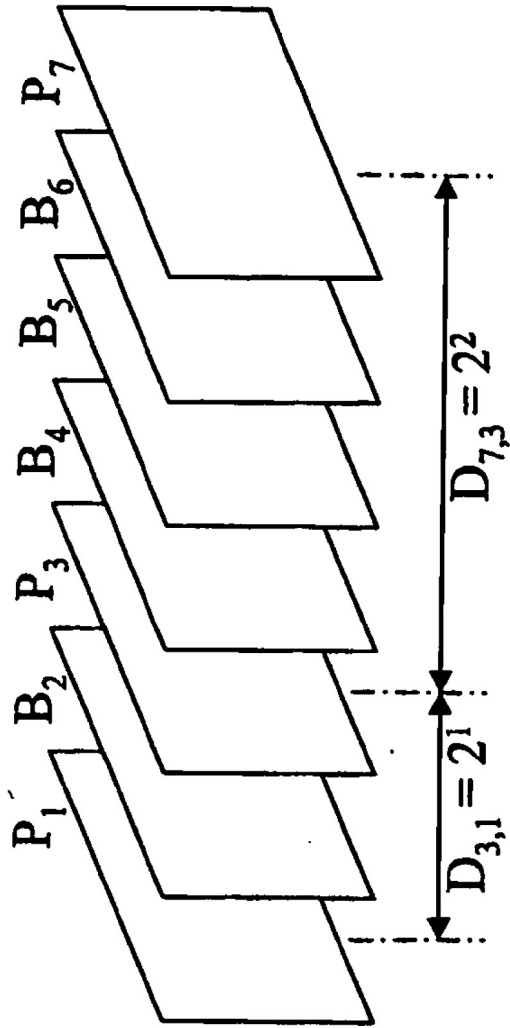


Figura 4