

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 722 001**

51 Int. Cl.:

H04N 19/196 (2014.01)
H04N 19/105 (2014.01)
H04N 19/176 (2014.01)
H04N 19/172 (2014.01)
H04N 19/46 (2014.01)
H04N 19/513 (2014.01)
H04N 19/149 (2014.01)
H04N 19/61 (2014.01)
H04N 19/573 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **14.07.2003** **E 15161109 (2)**

97 Fecha y número de publicación de la concesión europea: **27.03.2019** **EP 2950537**

54 Título: **Ponderación adaptativa de imágenes de referencia en codificación de vídeo**

30 Prioridad:

15.07.2002 US 395843 P
15.07.2002 US 395874 P
09.04.2003 US 410456

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
06.08.2019

73 Titular/es:

INTERDIGITAL VC HOLDINGS, INC. (100.0%)
200 Bellevue Parkway
Wilmington, DE 19809, US

72 Inventor/es:

BOYCE, JILL, MACDONALD

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 722 001 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

Ponderación adaptativa de imágenes de referencia en codificación de vídeo

Referencia cruzada a solicitudes relacionadas

5 Esta solicitud reivindica el beneficio de la Solicitud de Patente Provisional de EE.UU. Nº de Serie 60/395.843 (expediente del agente Nº PU020340), titulada "Adaptative Weighting Of Reference Pictures In Video CODEC" y presentada el 15 de julio de 2002. Además, esta solicitud reivindica el beneficio de la Solicitud de Patente Provisional de EE.UU. Nº de Serie 60/395.874 (expediente del agente Nº PU020339), titulada "Motion Estimation With Weighting Prediction" también presentada el 15 de julio de 2002.

Campo de la invención

10 La presente invención está dirigida a codificadores de vídeo y, en particular, a ponderación adaptativa de imágenes de referencia en codificadores de vídeo.

Antecedentes de la invención

15 Los datos de vídeo son generalmente procesados y transferidos en forma de flujos de bits. Los codificadores y decodificadores de compresión de vídeo típicos ("CODEC") obtienen mucha de su eficiencia de compresión formando una predicción de imagen de referencia de una imagen a ser codificada y codificando la diferencia entre la imagen actual y la predicción. Cuanto más próximamente este correlacionada la predicción con la imagen actual, menores bits son necesarios para comprimir esa imagen, aumentando por ello la eficiencia del proceso. De este modo, es deseable que sea formada la mejor predicción de imagen de referencia posible.

20 En muchos estándares de compresión de vídeo, incluyendo el Grupo de Expertos de Imagen en Movimiento (Moving Picture Experts Group) ("MPEG")-1, MPEG-2 y MPEG-4, una versión compensada de movimiento de una imagen de referencia previa se usa como predicción para la imagen actual y solamente se codifica la diferencia entre la imagen actual y la predicción. Cuando se usa una única predicción de imagen (imagen "P"), la imagen de referencia no se escala cuando se forma la predicción compensada de movimiento. Cuando se usan predicciones de imagen bi-direccionales (imágenes "B"), las predicciones intermedias se forman a partir de dos imágenes diferentes y entonces
25 las dos predicciones intermedias se promedian juntas, usando factores de ponderación iguales de (1/2, 1/2) para cada una, para formar una única predicción promediada. En estos estándares MPEG, las dos imágenes de referencia son siempre cada una de ellas de la dirección hacia delante y la dirección hacia atrás de las imágenes B.

El documento del estándar JVT-B075 de Kukichi et al, "Improved multiple frame motion compensation using frame interpolation" describe un método de compensación de movimiento mejorado para múltiples cuadros de referencia.

30 El documento del estándar JVT-C066 de Kukichi et al, "Multi-frame interpolative prediction with modified syntax" describe una sintaxis modificada en base a la imagen B que soporta predicción de interpolación multicuadro.

Compendio de la invención

La invención es un aparato de codificación de vídeo como se define en la reivindicación 1, una señal de vídeo como se define en la reivindicación 4, y un medio legible por una máquina como se define en la reivindicación 7.

35 **Breve descripción de los dibujos**

La ponderación adaptativa de imágenes de referencia en codificadores y decodificadores de vídeo según los principios de la presente invención se muestran en las siguientes figuras ejemplares, en las que:

la Figura 1 muestra un diagrama de bloques para un decodificador de vídeo estándar;

la Figura 2 muestra un diagrama de bloques para un decodificador de vídeo con bi-predicción adaptativa;

40 la Figura 3 muestra un diagrama de bloques para un decodificador de vídeo con ponderación de imagen de referencia según los principios de la presente invención;

la Figura 4 muestra un diagrama de bloques de un codificador de vídeo estándar;

la Figura 5 muestra un diagrama de bloques para un codificador de vídeo con ponderación de imagen de referencia según los principios de la presente invención;

45 la Figura 6 muestra una de flujo para un proceso de decodificación según los principios de la presente invención;
y

la Figura 7 muestra un diagrama de flujo para un proceso de codificación según los principios de la presente invención

Descripción detallada de las realizaciones preferidas

5 En algunas secuencias de vídeo, en particular en aquellas con desvanecimiento, la imagen actual o bloque de imágenes a ser codificada está más fuertemente correlacionada con una imagen de referencia escalada por un factor de ponderación que la imagen de referencia en sí misma. Los CODEC de vídeo sin factores de ponderación aplicados a las imágenes de referencia codifican secuencias de desvanecimiento de manera muy ineficiente. Cuando se usan factores de ponderación en codificación, un codificador de vídeo necesita determinar tanto los factores de ponderación como los vectores de movimiento, pero la mejor elección para cada uno de estos depende del otro, con la estimación de movimiento que es típicamente la parte más intensiva desde el punto de vista del cálculo de un codificador de compresión de vídeo digital.

10 En el estándar de compresión de vídeo Joint Video Team ("JVT") propuesto, cada imagen P puede usar múltiples imágenes de referencia para formar una predicción de imagen, pero cada bloque de movimiento individual o región 8x8 de un macrobloque usa solamente una única imagen de referencia para la predicción. Además de codificar y transmitir los vectores de movimiento, un índice de imagen de referencia se transmite para cada bloque de movimiento o región de 8x8, indicando qué imagen de referencia se usa. Un conjunto limitado de posibles imágenes de referencia se almacena tanto en el codificador como en el decodificador y se transmite el número de imágenes de referencia permisible.

20 En el estándar JVT, para imágenes bi-predictivas (también llamadas imágenes "B"), se forman dos predictores para cada bloque de movimiento o región de 8x8, cada uno de los cuales puede ser de una imagen de referencia separada y los dos predictores se promedian juntos para formar un único predictor promediado. Para bloques de movimiento codificados de manera bi-predictiva, las imágenes de referencia pueden ser ambas de la dirección hacia delante, ser ambas de la dirección hacia atrás o cada una de las direcciones hacia delante y hacia atrás. Se mantienen dos listas de las imágenes de referencia disponibles que se pueden usar para la predicción. Las dos imágenes de referencia se conocen como los predictores de la lista 0 y la lista 1. Un índice para cada imagen de referencia se codifica y transmite, ref_idx_l0 y ref_idx_l1 , para las imágenes de referencia de la lista 0 y la lista 1, respectivamente. Las imágenes bi-predictivas o "B" de Joint Video Team ("JVT") permiten una ponderación adaptativa entre las dos predicción, es decir,

$$Pred = [(P0)(Pred0)] + [(P1)(Pred1)] + D,$$

donde P0 y P1 son factores de ponderación, Pred0 y Pred1 son predicciones de la imagen de referencia para la lista 0 y la lista 1 respectivamente y D es un desplazamiento.

30 Se han propuesto dos métodos para la indicación de factores de ponderación. En el primero, los factores de ponderación se determinan por las direcciones que se usan para las imágenes de referencia. En este método, si el índice ref_idx_l0 es menor o igual que ref_idx_l1 , se usan los factores de ponderación ($\frac{1}{2}$, $\frac{1}{2}$), de otro modo se usan (2, -1).

35 En el segundo método ofrecido, cualquier número de factores de ponderación se transmite para cada segmento. Luego se trasmite un índice de factor de ponderación para cada bloque de movimiento o región 8x8 de un macrobloque que utiliza predicción bi-direccional. El decodificador utiliza el índice de factor de ponderación recibido para elegir el factor de ponderación apropiado, a partir del conjunto transmitido, para usar cuando se decodifique el bloque de movimiento o región 8x8. Por ejemplo, si se enviaron tres factores de ponderación en la capa de segmento, corresponderían a los índices de factor de ponderación 0, 1 y 2 respectivamente.

40 Las funciones de los distintos elementos mostrados en las figuras se pueden proporcionar a través del uso de un hardware dedicado así como un hardware capaz de ejecutar un software en asociación con un software apropiado. Cuando se proporcionan por un procesador, las funciones pueden ser proporcionadas por un único procesador dedicado, por un único procesador compartido o por una pluralidad de procesadores individuales, algunos de los cuales pueden ser compartidos. Además, el uso explícito del término "procesador" o "controlador" no se debería interpretar que se refiere exclusivamente a hardware capaz de ejecutar software y puede incluir implícitamente, sin limitación, hardware de procesador de señal digital ("DSP"), memoria de solo lectura ("ROM") para almacenar software, memoria de acceso aleatorio ("RAM") y almacenamiento no volátil. También se puede incluir otro hardware, convencional y/o personalizado. De manera similar, cualesquiera conmutadores mostrados en las figuras son solamente conceptuales. Su función puede ser llevada a cabo a través de la operación de lógica de programa, a través de lógica dedicada, a través de interacción de lógica dedicada y de control de programa o incluso manualmente, siendo la técnica particular seleccionable por el implementador como se entiende más específicamente a partir del contexto.

55 En las reivindicaciones, se pretende que cualquier elemento expresado como medios para llevar a cabo una función específica abarque cualquier modo de llevar a cabo esa función, incluyendo, por ejemplo, a) una combinación de elementos de circuito que llevan a cabo la función o b) software en cualquier forma, incluyendo, por tanto, microprograma, microcódigo o similares, combinados con circuitos apropiados para ejecutar ese software para llevar a cabo la función. La invención como se define en tales reivindicaciones reside en el hecho de que las funcionalidades proporcionadas por los diversos medios indicados se combinan y se ponen en conjunto de la

manera que indican las reivindicaciones. El solicitante, por tanto, considera cualesquiera medios que puedan proporcionar esas funcionalidades, como equivalentes a los mostrados en la presente memoria.

Como se muestra en la Figura 1, un decodificador de vídeo estándar se indica de manera general con el número de referencia 100. El decodificador de vídeo 100 incluye un decodificador de longitud variable ("VLD") 110 conectado en comunicación de señal con un cuantificador inverso 120. El cuantificador inverso 120 está conectado en comunicación de señal con un transformador inverso 130. El transformador inverso 130 está conectado en comunicación de señal con un primer terminal de entrada de un sumador o unión de suma 140, en donde la salida de la unión de suma 140 proporciona la salida del codificador de vídeo 100. La salida de la unión de suma 140 está conectada en comunicación de señal con un almacén de imagen de referencia 150. El almacén de imagen de referencia 150 está conectado en comunicación de señal con un compensador de movimiento 160, que está conectado en comunicación de señal con un segundo terminal de entrada de la unión de suma 140.

Volviendo a la Figura 2, un decodificador de vídeo con bi-predicción adaptativa se indica de manera general con el número de referencia 200. El codificador de vídeo 200 incluye un VLD 210 conectado en comunicación de señal con un cuantificador inverso 220. El cuantificador inverso 220 está conectado en comunicación de señal con un transformador inverso 230. El transformador inverso 230 está conectado en comunicación de señal con un primer terminal de entrada de una unión de suma 240, en el que la salida de la unión de suma 240 proporciona la salida del decodificador de vídeo 200. La salida de la unión de suma 240 está conectada en comunicación de señal con un almacén de imágenes de referencia 250. El almacén de imágenes de referencia 250 está conectado en comunicación de señal con un compensador de movimiento 260, que está conectado en comunicación de señal con una primera entrada de un multiplicador 270.

El VLD 210 además está conectado en comunicación de señal con una búsqueda de factor de ponderación de imagen de referencia 280 para proporcionar un índice de coeficiente de bi-predicción adaptativa ("ABP") a la búsqueda 280. Una primera salida de la búsqueda 280 es para proporcionar un factor de ponderación y está conectada en comunicación de señal a una segunda entrada del multiplicador 270. La salida del multiplicador 270 está conectada en comunicación de señal a una primera entrada de una unión de suma 290. Una segunda salida de búsqueda 280 es para proporcionar un desplazamiento y está conectada en comunicación de señal a una segunda entrada de la unión de suma 290. La salida de la unión de suma 290 está conectada en comunicación de señal con un segundo terminal de entrada de la unión de suma 240.

Volviendo ahora a la Figura 3, un decodificador de vídeo con ponderación de imagen de referencia se indica de manera general con el número de referencia 300. El decodificador de vídeo 300 incluye un VLD 310 conectado en comunicación de señal con un cuantificador inverso 320. El cuantificador inverso 320 está conectado en comunicación de señal con un transformador inverso 330. El transformador inverso 330 está conectado en comunicación de señal con un primer terminal de entrada de una unión de suma 340, en el que la salida de la unión de suma 340 proporciona la salida del decodificador de vídeo 300. La salida de la unión de suma 340 está conectada en comunicación de señal con un almacén de imágenes de referencia 350. El almacén de imágenes de referencia 350 está conectado en comunicación de señal con un compensador de movimiento 360, que está conectado en comunicación de señal con una primera entrada de un multiplicador 370.

El VLD 310 además está conectado en comunicación de señal con una búsqueda de factor de ponderación de imagen de referencia 380 para proporcionar un índice de imagen de referencia a la búsqueda 380. Una primera salida de la búsqueda 380 es para proporcionar un factor de ponderación y está conectada en comunicación de señal a una segunda entrada del multiplicador 370. La salida del multiplicador 370 está conectada en comunicación de señal con la primera entrada de una unión de suma 390. Una segunda salida de la búsqueda 380 es para proporcionar un desplazamiento y está conectada en comunicación de señal a una segunda entrada de la unión de suma 390. La salida de la unión de suma 390 está conectada en comunicación de señal con un segundo terminal de entrada de la unión de suma 340.

Como se muestra en la Figura 4, un codificador de vídeo estándar se indica de manera general por el número de referencia 400. Una entrada al codificador 400 está conectada en comunicación de señal con una entrada de no inversión de una unión de suma 410. La salida de la unión de suma 410 está conectada en comunicación de señal con un transformador de bloques 420. El transformador 420 está conectado en comunicación de señal con un cuantificador 430. La salida del cuantificador 430 está conectada en comunicación de señal con un codificador de longitud variable ("VLC") 440, en donde la salida del VLC 440 es una salida disponible externamente del codificador 400.

La salida del cuantificador 430 está además conectada en comunicación de señal con un cuantificador inverso 450. El cuantificador inverso 450 está conectado en comunicación de señal con un transformador de bloques inverso 460, que, a su vez, está conectado en comunicación de señal con un almacén de imágenes de referencia 470. Una primera salida del almacén de imágenes de referencia 470 está conectada en comunicación de señal con una primera entrada de un estimador de movimiento 480. La entrada al codificador 400 además está conectada en comunicación de señal con una segunda entrada del estimador de movimiento 480. La salida del estimador de movimiento 480 está conectada en comunicación de señal con una primera entrada de un compensador de movimiento 490. Una segunda salida del almacén de imágenes de referencia 470 está conectada en comunicación

de señal con una segunda entrada del compensador de movimiento 490. La salida del compensador de movimiento 490 está conectada en comunicación de señal con una entrada de inversión de la unión de suma 410.

Volviendo a la Figura 5, un codificador de vídeo con ponderación de imagen de referencia se indica de manera general con el número de referencia 500. Una entrada al codificador 500 está conectada en comunicación de señal con una entrada de no inversión de una unión de suma 510. La salida de la unión de suma 510 está conectada en comunicación de señal con un transformador de bloques 520. El transformador de bloques 520 está conectado en comunicación de señal con un cuantificador 530. La salida del cuantificador 530 está conectada en comunicación de señal con un VLC 540, en el que la salida del VLC 440 es una salida disponible externamente del codificador 500.

La salida del cuantificador 530 además está conectada en comunicación de señal con el cuantificador inverso 550. El cuantificador inverso 550 está conectado en comunicación de señal con un transformador de bloques inverso 560, que, a su vez, está conectado en comunicación de señal con un almacén de imágenes de referencia 570. Una primera salida del almacén de imágenes de referencia 570 está conectada en comunicación de señal con una primera entrada de un asignador de factor de ponderación de imagen de referencia 572. La entrada al codificador 500 además está conectada en comunicación de señal con una segunda entrada del asignador de factor de ponderación de imagen de referencia 572. La salida del asignador de factor de ponderación de imagen de referencia 572, que es indicativa de un factor de ponderación, está conectada en comunicación de señal con una primera entrada de un estimador de movimiento 580. Una segunda salida del almacén de imágenes de referencia 570 está conectada en comunicación de señal con una segunda entrada del estimador de movimiento 580.

La entrada al codificador 500 además está conectada en comunicación de señal con una tercera entrada del estimador de movimiento 580. La salida del estimador de movimiento 580, que es indicativa de vectores de movimiento, está conectada en comunicación de señal con una primera entrada de un compensador de movimiento 590. Una tercera salida del almacén de imágenes de referencia 570 está conectada en comunicación de señal con una segunda entrada del compensador de movimiento 590. La salida del compensador de movimiento 590, que es indicativa de una imagen de referencia compensada de movimiento, está conectada en comunicación de señal con una primera entrada de un multiplicador 592. La salida del asignador de factor de ponderación de imagen de referencia 572, que es indicativa de un factor de ponderación, está conectada en comunicación de señal con una segunda entrada del multiplicador 592. La salida del multiplicador 592 está conectada en comunicación de señal con una entrada de inversión de la unión de suma 510.

Volviendo ahora a la Figura 6, un proceso ejemplar para codificar datos de señal de vídeo para un bloque de imágenes se indica de manera general por el número de referencia 600. El proceso incluye un bloque de inicio 610 que pasa el control a un bloque de entrada 612. El bloque de entrada 612 recibe los datos comprimidos del bloque de imágenes y pasa el control a un bloque de entrada 614. El bloque de entrada 614 recibe al menos un índice de imagen de referencia con los datos para el bloque de imagen, cada índice de imagen de referencia que corresponde a una imagen de referencia particular. El bloque de entrada 614 pasa el control a un bloque de función 616, que determina un factor de ponderación correspondiente a cada uno de los índices de imagen de referencia recibidos y pasa el control a un bloque de función opcional 617. El bloque de función opcional 617 determina un desplazamiento correspondiente a cada uno de los índices de imagen de referencia recibidos y pasa el control a un bloque de función 618. El bloque de función 618 recupera una imagen de referencia correspondiente a cada uno de los índices de imagen de referencia recibidos y pasa el control a un bloque de función 620. El bloque de función 620, a su vez, compensa el movimiento de la imagen de referencia recuperada y pasa el control a un bloque de función 622. El bloque de función 622 multiplica la imagen de referencia de movimiento compensado por un factor de ponderación correspondiente y pasa el control a un bloque de función opcional 623. El bloque de función opcional 623 suma la imagen de referencia de movimiento compensado al desplazamiento correspondiente y pasa el control a un bloque de función 624. El bloque de función 624, a su vez, forma una imagen de referencia de movimiento compensado ponderada y pasa el control a un bloque final 626.

Volviendo a la Figura 7, un proceso ejemplar para codificar datos de señales de vídeo para un bloque de imagen se indica de manera general por el número de referencia 700. El proceso incluye un bloque de inicio 710 que pasa el control a un bloque de entrada 712. El bloque de entrada 712 recibe datos de bloque de imágenes sustancialmente no comprimidos y pasa el control al bloque de función 714. El bloque de función 714 asigna un factor de ponderación para el bloque de imagen correspondiente a una imagen de referencia particular que tiene un índice correspondiente. El bloque de función 714 pasa el control a un bloque de función opcional 715. El bloque de función opcional 715 asigna un desplazamiento para el bloque de imagen correspondiente a una imagen de referencia particular que tiene un índice correspondiente. El bloque de función opcional 715 pasa el control a un bloque de función 716, que calcula los vectores de movimiento correspondientes a la diferencia entre el bloque de imagen y la imagen de referencia particular y pasa el control a un bloque de función 718. El bloque de función 718 compensa el movimiento de la imagen de referencia particular en correspondencia con los vectores de movimiento y pasa el control a un bloque de función 720. El bloque de función 720, a su vez, multiplica la imagen de referencia de movimiento compensado por el factor de ponderación asignado para formar una imagen de referencia de movimiento compensado ponderada y pasa el control a un bloque de función opcional 721. El bloque de función opcional 721, a su vez, suma la imagen de referencia de movimiento compensado al desplazamiento asignado para formar una imagen de referencia de movimiento compensado ponderada y pasa el control al bloque de función 722. El bloque de función 722 resta la imagen de referencia de movimiento compensado ponderada del bloque de imagen

sustancialmente no comprimida y pasa el control a un bloque de función 724. El bloque de función 724, a su vez, codifica una señal con la diferencia entre el bloque de imagen sustancialmente no comprimida y la imagen de referencia de movimiento compensado ponderada junto con el índice correspondiente de la imagen de referencia particular y pasa el control a un bloque de final 726.

5 En la presente realización ejemplar, para cada imagen o segmento codificado, se asocia un factor de ponderación con cada imagen de referencia permisible con respecto a cuyos bloques de la imagen actual se puede codificar. Cuando se codifica o decodifica cada bloque individual en la imagen actual, el(los) factor(es) de ponderación y el(los) desplazamiento(s) que corresponde(n) a sus índices de imagen de referencia se aplican a la predicción de referencia para formar un predictor de ponderación. Todos los bloques en el segmento que están codificados con
10 respecto a la misma imagen de referencia aplican el mismo factor de ponderación a la predicción de imagen de referencia.

Si se usa o no ponderación adaptativa cuando se codifica una imagen se puede indicar en el conjunto de parámetros de imagen o conjunto de parámetros de secuencia o en el segmento o cabecera de imagen. Para cada segmento o imagen que usa ponderación adaptativa, se puede transmitir un factor de ponderación para cada una de las
15 imágenes de referencia permisibles que se pueden usar para codificar este segmento o imagen. El número de imágenes de referencia permisibles se transmite en la cabecera del segmento. Por ejemplo, si se pueden usar tres imágenes de referencia para codificar el segmento actual, se transmiten hasta tres factores de ponderación y se asocian con la imagen de referencia con el mismo índice.

Si no se transmiten factores de ponderación, se usan ponderaciones por defecto. En una realización de la invención actual, las ponderaciones por defecto de ($\frac{1}{2}, \frac{1}{2}$) se usan cuando no se transmiten factores de ponderación. Los factores de ponderación se pueden transmitir usando o bien códigos de longitud fija o bien variable.

A diferencia de los sistemas típicos, cada factor de ponderación que se transmite con cada segmento, bloque o imagen corresponde a un índice de imagen de referencia particular. Previamente, cualquier conjunto de factores de ponderación transmitidos con cada segmento o imagen no estaban asociados con ninguna imagen de referencia particular. En su lugar, se transmitió un índice de ponderación de bi-predicción adaptativa para cada bloque de
25 movimiento o región 8x8 para seleccionar cuál de los factores de ponderación del conjunto transmitido iba a ser aplicado para ese bloque de movimiento particular o región 8x8.

En la presente realización, no se transmite explícitamente el índice de factor de ponderación para cada bloque de movimiento o región 8x8. En su lugar, se usa el factor de ponderación que está asociado con el índice de imagen de referencia transmitido. Esto reduce drásticamente la cantidad de sobrecarga en el flujo de bits transmitido para permitir la ponderación adaptativa de imágenes de referencia.

Este sistema y técnica se pueden aplicar o bien a imágenes Predictivas "P", que están codificadas con un único predictor, o bien a imágenes Bi-predictivas "B", que están codificadas con dos predictores. Los procesos de decodificación, que están presentes tanto en codificadores como en decodificadores, se describen más adelante para los casos de imagen P y B. Alternativamente, esta técnica también se puede aplicar a sistemas de codificación que usan conceptos similares a imágenes I, B y P.

Los mismos factores de ponderación se pueden usar para predicción direccional única en imágenes B y para predicción bi-direccional en imágenes B. Cuando se usa un único predictor para un macrobloque, en imágenes P o para predicción direccional única en imágenes B, se transmite un único índice de imagen de referencia para el
40 bloque. Después de que el paso de proceso de decodificación de compensación de movimiento produzca un predictor, el factor de ponderación se aplica al predictor. El predictor ponderado entonces se añade al residual codificado y se realiza el recorte sobre la suma, para formar la imagen decodificada. Para el uso de los bloques en imágenes P o de los bloques en imágenes B que usan solamente predicción de lista 0, el predictor ponderado se forma como:

$$45 \quad \text{Pred} = W0 * \text{Pred0} + D0 \quad (1)$$

donde W0 es el factor de ponderación asociado con la imagen de referencia de lista 0, D0 es el desplazamiento asociado con la imagen de referencia de lista 0 y Pred0 es el bloque de predicción de movimiento compensado a partir de la imagen de referencia de lista 0.

Para el uso de los bloques en imágenes B que usan solamente predicción de lista 0, el predictor ponderado se forma como:

$$50 \quad \text{Pred} = W1 * \text{Pred1} + D1 \quad (2)$$

donde W1 es el factor de ponderación asociado con la imagen de referencia de la lista 1, D1 es el desplazamiento asociado con la imagen de referencia de la lista 1 y Pred1 es el bloque de predicción de movimiento compensado a partir de la imagen de referencia de la lista 1.

Los predictores ponderados pueden ser recortados para garantizar que los valores resultantes estarán dentro del intervalo permisible de valores de píxeles, típicamente de 0 a 255. La precisión de la multiplicación en las fórmulas de ponderación puede estar limitada a cualquier número de bits de resolución predeterminado.

5 En el caso bi-predictivo, los índices de imagen de referencia se transmiten para cada uno de los dos predictores. La compensación de movimiento se realiza para formar los dos predictores. Cada predictor usa el factor de ponderación asociado con su índice de imagen de referencia para formar dos predictores ponderados. Los dos predictores ponderados entonces se promedian juntos para formar un predictor promediado, que entonces se añade al residual codificado.

10 Para el uso de los bloques en imágenes B que usan predicciones de lista 0 y de lista 1, el predictor ponderado se forma como:

$$\text{Pred} = (P0 * \text{Pred0} + D0 + P1 * \text{Pred1} + D1)/2 \quad (3)$$

El recorte se puede aplicar al predictor ponderado o a cualquiera de los valores intermedios en el cálculo del predictor ponderado para garantizar que los valores resultantes estarán dentro del intervalo permisible de valores de píxeles, típicamente de 0 a 255.

15 De este modo el factor de ponderación se aplica a la predicción de imagen de referencia de un codificador y decodificador de compresión de vídeo que usa múltiples imágenes de referencia. El factor de ponderación se adapta a los bloques de movimiento individuales dentro de una imagen, en base al índice de imagen de referencia que se usa para ese bloque de movimiento. Debido a que el índice de imagen de referencia ya se transmitió en el flujo de bits de vídeo comprimidos, la sobrecarga adicional para adaptar el factor de ponderación sobre la base de bloque de movimiento se reduce drásticamente. Todos los bloques de movimiento que están codificados con respecto a la misma imagen de referencia aplican el mismo factor de ponderación a la predicción de imagen de referencia.

20 Estas y otras características y ventajas de la presente invención se pueden averiguar fácilmente por un experto en la técnica pertinente en base a las enseñanzas en la presente memoria. Se ha de entender que las enseñanzas de la presente invención se pueden implementar en diversas formas de hardware, software, microprograma, procesadores de propósito especial o combinaciones de los mismos.

25 Más preferiblemente las enseñanzas de la presente invención se implementan como una combinación de hardware y software. Además, el software se implementa preferiblemente como un programa de aplicaciones incorporado de manera tangible en una unidad de almacenamiento de programa. El programa de aplicaciones se puede descargar a y ser ejecutado por, una máquina que comprenda cualquier arquitectura adecuada. Preferiblemente, la máquina se implementa en una plataforma informática que tiene un hardware tal como una o más unidades centrales de proceso ("CPU"), una memoria de acceso aleatorio ("RAM") e interfaces de entrada/salida ("I/O"). La plataforma informática puede incluir también un sistema operativo y código de microinstrucciones. Los distintos procesos y funciones descritos en la presente memoria pueden ser o bien parte del código de microinstrucciones o bien parte del programa de aplicaciones o cualquier combinación de los mismos, que se pueda ejecutar por una CPU. Además, otras varias unidades periféricas se pueden conectar a la plataforma informática tal como una unidad de almacenamiento de datos adicional y una unidad de impresión.

35 Se ha de entender que, debido a que algunos de los componentes y métodos constituyentes del sistema representados en los dibujos anexos están implementados preferiblemente en software, las conexiones reales entre los componentes del sistema o los bloques de función de proceso puede diferir dependiendo de la manera en la que se programe la presente invención. Dadas las enseñanzas en la presente memoria, un experto en la técnica pertinente será capaz de contemplar estas y similares implementaciones o configuraciones de la presente invención.

REIVINDICACIONES

1. Un aparato de codificación para producir datos de vídeo para una imagen que tiene una pluralidad de bloques de imagen, incluyendo los datos de vídeo datos codificados para un bloque de imagen de la pluralidad de bloques de imagen y que incluyen al menos un índice de imagen de referencia, comprendiendo el aparato:
- 5 medios para acceder (712) al bloque de imagen y a un único índice de imagen de referencia para predecir el bloque de imagen, correspondiendo el único índice de imagen de referencia a una imagen de referencia particular; caracterizado por
- medios para codificar (724) el bloque de imagen usando la imagen de referencia particular que corresponde al único índice de imagen de referencia y usando un factor de ponderación, determinado a partir de un conjunto de factores de ponderación y que corresponde al único índice de imagen de referencia,
- 10 en donde (714) el único índice de imagen de referencia determina tanto la imagen de referencia particular como el factor de ponderación y cada factor de ponderación en el conjunto de factores de ponderación corresponde a una imagen de referencia particular diferente.
2. Un aparato como se define en la reivindicación 1, en el que los medios para codificar comprenden:
- 15 medios para determinar un bloque de movimiento compensado a partir de la imagen de referencia particular;
- medios para ponderar el bloque de movimiento compensado usando el factor de ponderación para producir un bloque de movimiento compensado ponderado;
- medios para determinar un desplazamiento; y
- 20 medios para ajustar el bloque de movimiento compensado ponderado usando el desplazamiento para producir un bloque de movimiento compensado ponderado ajustado.
3. Un aparato como se define en la reivindicación 2, en el que los medios para codificar comprenden adicionalmente medios para restar el bloque de movimiento compensado ponderado ajustado y el bloque de imagen.
4. Una señal de vídeo formateada para incluir datos de vídeo para una imagen que tiene una pluralidad de bloques de imagen, comprendiendo la señal de vídeo:
- 25 datos codificados para un bloque de imagen de la pluralidad de bloques de imagen; y
- un único índice de imagen de referencia para el bloque de imagen, correspondiendo el único índice de imagen de referencia a una imagen de referencia particular usada para predecir el bloque de imagen,
- en donde el bloque de imagen se codifica (724) usando la imagen de referencia particular que corresponde al único índice de imagen de referencia y usando un factor de ponderación, determinado a partir de un conjunto de factores de ponderación y que corresponde al único índice de imagen de referencia, y
- 30 en donde el único índice de imagen de referencia determina (714) tanto la imagen de referencia particular como el factor de ponderación y cada factor de ponderación en el conjunto de factores de ponderación corresponde a una imagen de referencia particular diferente.
5. La señal como se define en la reivindicación 4, en la que el bloque de imagen se codifica:
- 35 determinando un bloque de movimiento compensado a partir de la imagen de referencia particular;
- ponderando el bloque de movimiento compensado usando el factor de ponderación para producir un bloque de movimiento compensado ponderado;
- determinando un desplazamiento; y
- 40 ajustando el bloque de movimiento compensado ponderado usando el desplazamiento para producir un bloque de movimiento compensado ponderado ajustado.
6. La señal como se define en la reivindicación 5, en la que el bloque de imagen se codifica adicionalmente restando el bloque de movimiento compensado ponderado ajustado del bloque de imagen.
7. Un medio legible por una máquina que tiene almacenadas instrucciones ejecutables por una máquina que, cuando se ejecutan, implementan un método para codificar una imagen que tiene una pluralidad de bloques de imagen, comprendiendo el método:
- 45 Acceder (712) al bloque de imagen y un único índice de imagen de referencia para predecir el bloque de imagen, correspondiendo el único índice de imagen de referencia a una imagen de referencia particular; y

codificar (724) el bloque de imagen usando la imagen de referencia particular que corresponde al único índice de imagen de referencia y usando un factor de ponderación, determinado a partir de un conjunto de factores de ponderación y que corresponde al único índice de imagen de referencia,

5 en donde (714) el único índice de imagen de referencia determina tanto la imagen de referencia particular como el factor de ponderación y cada factor de ponderación en el conjunto de factores de ponderación corresponde a una imagen de referencia particular diferente.

8. El medio de la reivindicación 7, en el que la codificación comprende adicionalmente:

determinar un bloque de movimiento compensado a partir de la imagen de referencia particular;

10 ponderar el bloque de movimiento compensado usando el factor de ponderación para producir un bloque de movimiento compensado ponderado;

determinar un desplazamiento; y

ajustar el bloque de movimiento compensado ponderado usando el desplazamiento para producir un bloque de movimiento compensado ponderado ajustado.

15 9. El medio como se define en la reivindicación 8, en donde el bloque de imagen se codifica adicionalmente restando el bloque de movimiento compensado ponderado ajustado del bloque de imagen.

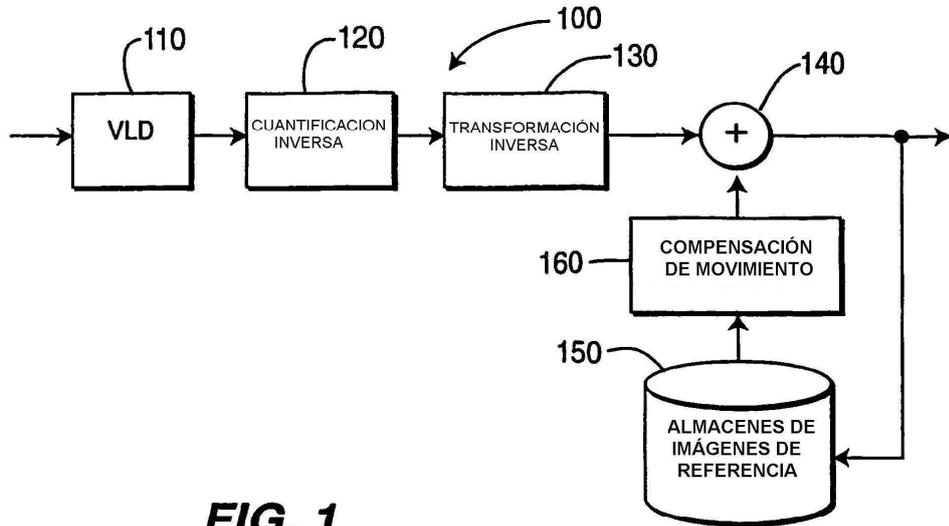


FIG. 1

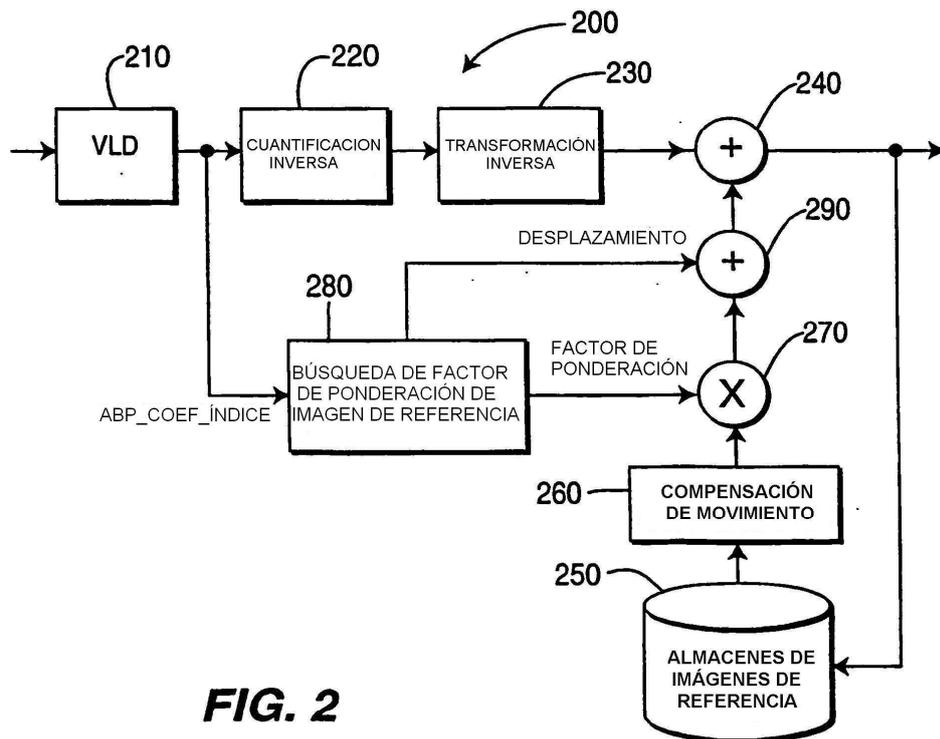


FIG. 2

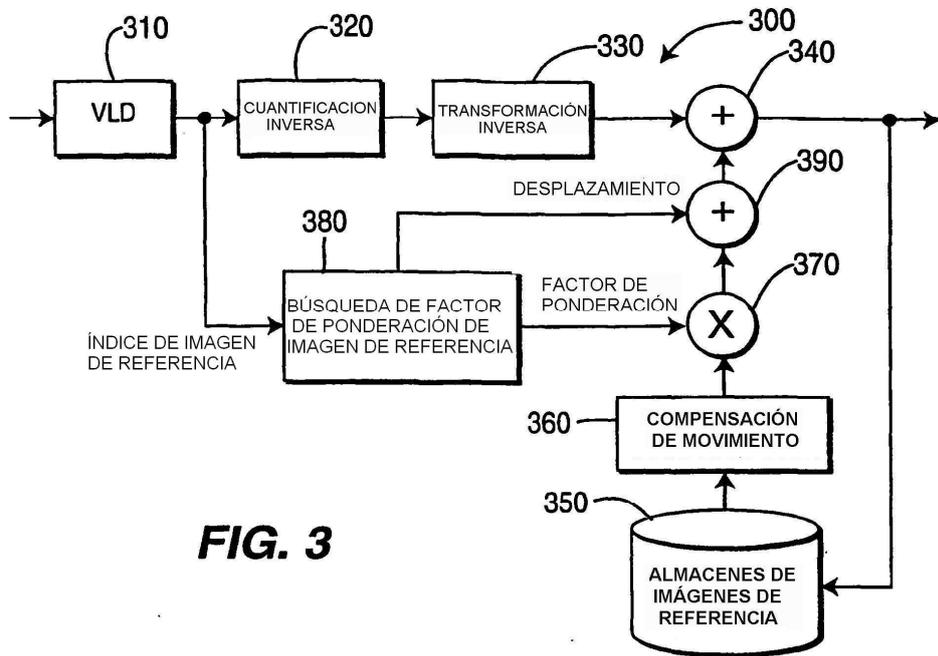


FIG. 3

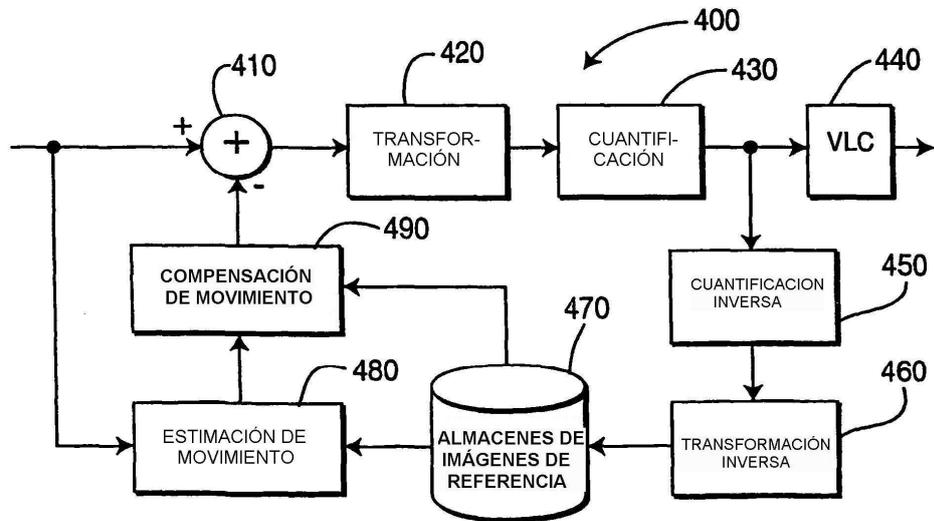


FIG. 4

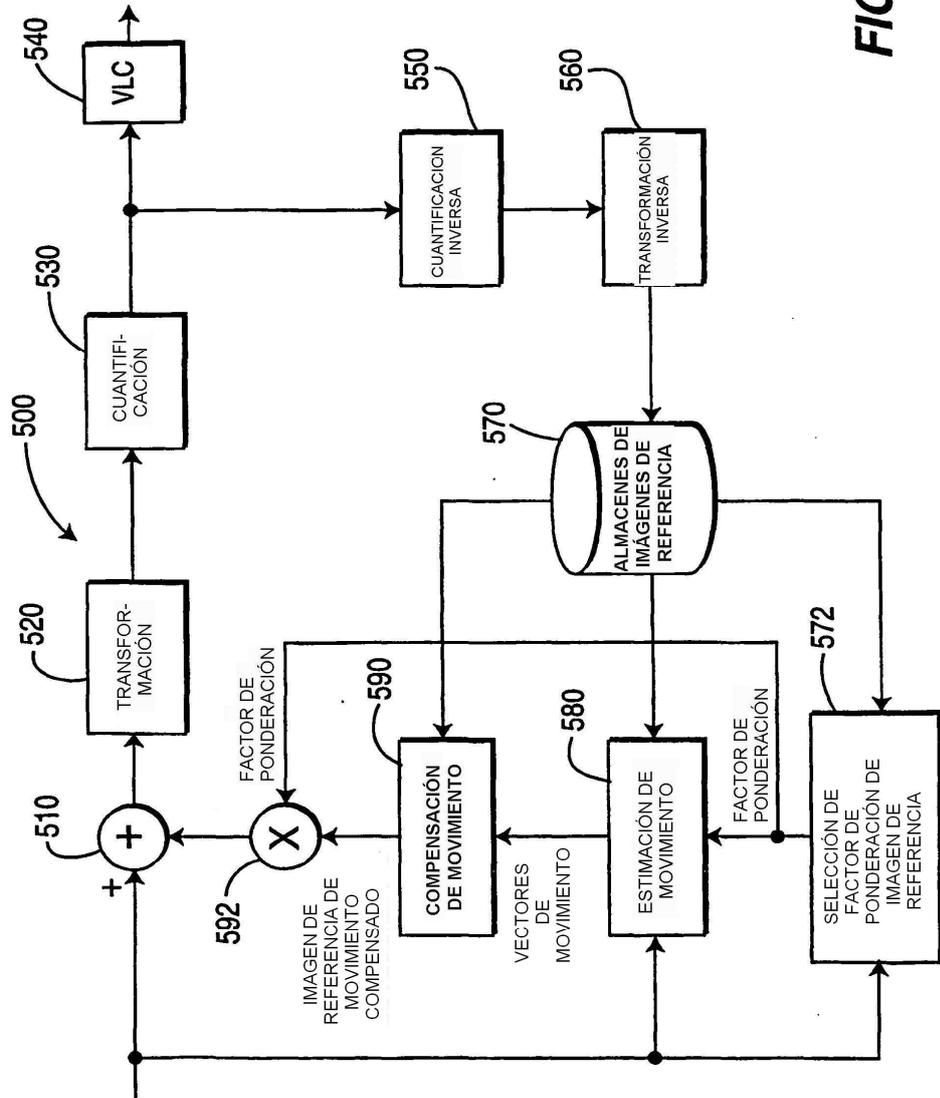


FIG. 5

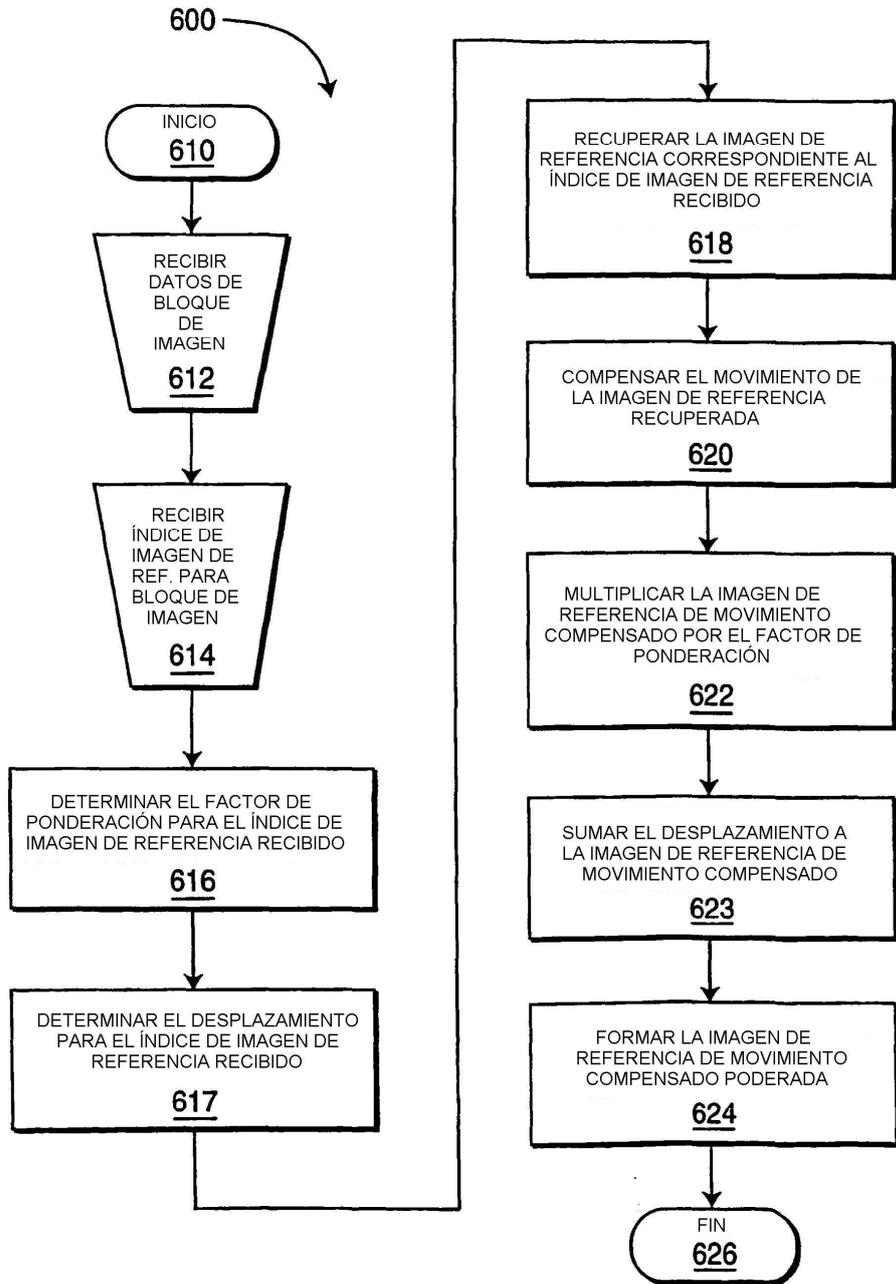


FIG. 6

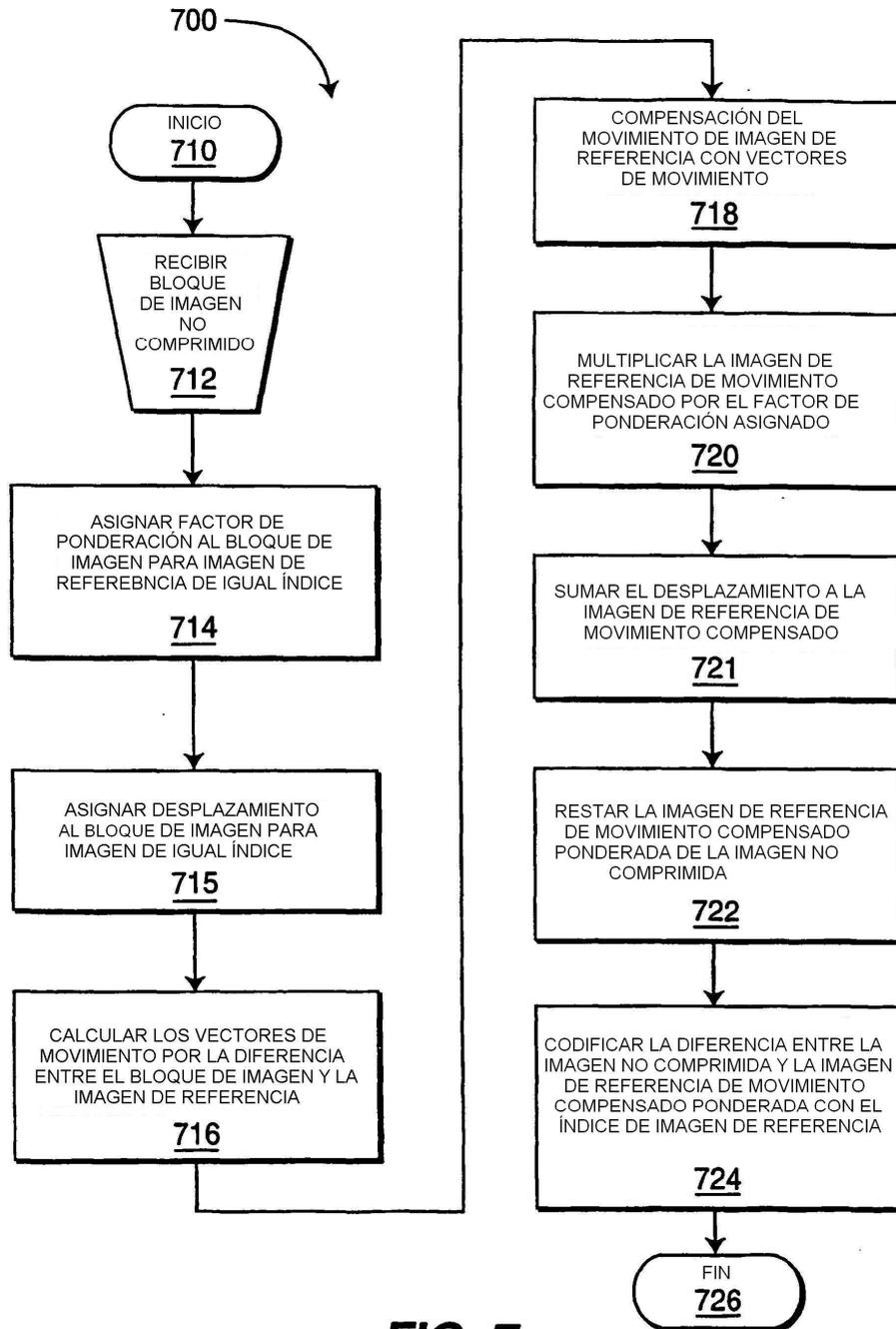


FIG. 7