

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 545 109**

51 Int. Cl.:

H04N 19/159 (2014.01)
H04N 19/172 (2014.01)
H04N 19/46 (2014.01)
H04N 19/61 (2014.01)
H04N 19/114 (2014.01)
H04N 19/117 (2014.01)
H04N 19/86 (2014.01)
H04N 19/523 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **08.06.1998 E 10012308 (2)**

97 Fecha y número de publicación de la concesión europea: **05.08.2015 EP 2285119**

54 Título: **Procedimiento de decodificación de imágenes**

30 Prioridad:

09.06.1997 JP 15065697

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

08.09.2015

73 Titular/es:

HITACHI, LTD. (100.0%)
6-6, Marunouchi 1-chome Chiyoda-ku
Tokyo 100-8280, JP

72 Inventor/es:

NAKAYA, YUICHIRO

74 Agente/Representante:

CURELL AGUILÁ, Mireia

ES 2 545 109 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 de decodificación de imágenes.

5 Antecedentes de la invención

Campo de la invención

La presente invención se refiere a un procedimiento de decodificación de secuencias de imágenes que lleva a cabo predicción de intertrama utilizando valores cuantificados para intensidad de luminancia o crominancia.

Descripción de la técnica relacionada

En la codificación de secuencias de imágenes de alta eficacia, la técnica de predicción de intertrama (compensación de movimiento), mediante la similitud de las tramas adyacentes a lo largo del tiempo, resulta de gran eficacia para la compresión de datos. El procedimiento de compensación de movimiento utilizado con más frecuencia actualmente es el de la correspondencia de bloques con una precisión de medio píxel, que se utiliza en los estándares internacionales H.263, MPEG1 y MPEG2. En este procedimiento, la imagen que se va a codificar se segmenta en bloques y los componentes horizontal y vertical de los vectores de movimiento de estos bloques se estiman como integrales múltiples de la mitad de la distancia entre píxeles adyacentes. Este procedimiento se describe mediante la siguiente ecuación:

[Ecuación 1]

$$P(x, y) = R(x + u_i, y + v_i) \quad (x, y) \in B_i, 0 \leq i < N \dots (1)$$

en la que $P(x, y)$ y $R(x, y)$ denotan los valores de muestra (intensidad de luminancia o crominancia) de los píxeles situados en las coordenadas (x, y) de la imagen prevista P de la trama actual y la imagen de referencia (imagen decodificada de una trama que se ha codificado antes que la trama actual) R , respectivamente. x e y son enteros y se presupone que todos los píxeles están situados en puntos donde los valores de coordenadas son enteros. Además, se presupone que los valores de muestra de los píxeles se cuantifican como enteros no negativos. N , B_i y (u_i, v_i) denotan el número de bloques de la imagen, el grupo de píxeles incluidos en el i -ésimo bloque de la imagen y los vectores de movimiento del i -ésimo bloque, respectivamente.

Cuando los valores de u_i y v_i no son enteros, es necesario hallar el valor de intensidad en el punto concreto en el que no existen píxeles en la imagen de referencia. Actualmente, la interpolación bilineal mediante los cuatro píxeles adyacentes es el procedimiento utilizado con más frecuencia para este proceso. Este procedimiento de interpolación se describe mediante la siguiente ecuación:

[Ecuación 2]

$$R\left(x + \frac{p}{d}, y + \frac{q}{d}\right) = ((d - q)((d - p)R(x, y)) + ph(x + 1, y)) + q((d - p)R(x, y + 1) + pR(x + 1, y + 1))/d^2 \dots (2)$$

en la que d es un entero positivo y p y q son inferiores a d pero no inferiores a 0. «//» denota una división entera que redondea el resultado de la división normal (división con números reales) al entero más cercano.

En la Fig. 1, se muestra un ejemplo de la estructura de un codificador de vídeo H.263. Como algoritmo de codificación, el H.263 adopta un procedimiento de codificación híbrido (procedimiento de codificación adaptativo de intertrama/intratrama) que es una combinación de la correspondencia de bloques y la DCT (transformada discreta de coseno). Un restador 102 calcula la diferencia entre la imagen de entrada (imagen base de trama actual) 101 y la imagen de salida 113 (descrita más adelante) del selector de codificación de intertrama/intratrama 119 y, a continuación, proporciona una imagen de error 103. Esta imagen de error se cuantifica en un cuantificador 105 tras haber sido convertida en coeficientes DCT en un convertidor DCT 104, generándose coeficientes DCT cuantificados 106. Estos coeficientes DCT cuantificados se transmiten a través del canal de comunicación y, al mismo tiempo, se utilizan para sintetizar la imagen prevista de intertrama en el codificador. El procedimiento para sintetizar la imagen prevista se explicará a continuación. Los coeficientes DCT cuantificados 106 mencionados forman la imagen de error reconstruida 110 (igual que la imagen de error reconstruida de la sección de recepción) tras pasar por un decuantificador 108 y un convertidor DCT inverso 109. Esta imagen de error reconstruida y la imagen de salida 113 del selector de codificación de intertrama/intratrama 119 se suman en el sumador 111 y se obtiene la imagen decodificada 112 de la trama actual (la misma imagen que la imagen decodificada de la trama actual reconstruida en la sección del receptor). Esta imagen se almacena en una memoria de trama 114 y se retarda un tiempo igual al intervalo de la trama. En consecuencia, en el momento actual, la memoria de trama 114 genera la imagen decodificada 115 de la trama anterior. Esta imagen decodificada de la trama anterior y la imagen original 101 de la

trama actual se introducen en la sección de correspondencia de bloques 116, donde se lleva a cabo la correspondencia de bloques entre estas imágenes. En el procedimiento de correspondencia de bloques, la imagen original de la trama actual se segmenta en diversos bloques, y la imagen prevista 117 de la trama actual se sintetiza extrayendo la sección que más se parece a estos bloques de la imagen decodificada de la trama anterior. En este procedimiento, es necesario estimar el movimiento entre la trama anterior y la trama actual para cada bloque. El vector de movimiento de cada bloque estimado en el procedimiento de estimación del movimiento se transmite a la sección del receptor como datos de vector de movimiento 120. En la sección del receptor, se sintetiza la misma imagen de predicción que en la sección del transmisor mediante la información del vector de movimiento y la imagen de decodificación de la trama anterior. La imagen de predicción 117 se introduce junto con una señal «0» 118 en el selector de codificación de intertrama/intratrama 119. Este conmutador 119 selecciona codificación de intertrama o codificación de intratrama seleccionando una de las dos entradas. Se efectúa codificación de intertrama cuando se selecciona la imagen de predicción 117 (es el caso mostrado en la Fig. 2). En cambio, se efectúa codificación de intratrama cuando se selecciona la señal «0», puesto que la propia imagen de entrada se convierte en coeficientes DCT que se proporcionan al canal de comunicación. Para la correcta reconstrucción de la imagen codificada en la sección del receptor, es necesario indicar al receptor si se ha efectuado codificación de intratrama o codificación de intertrama en la sección del transmisor. En consecuencia, se proporciona un indicador de identificación 121 al circuito de comunicación. Por último, se obtiene un tren de bits con codificación H.263 multiplexando los coeficientes DCT cuantificados, los vectores de movimiento y la información del indicador de identificación de intertrama/intratrama en un multiplexor 122.

En la Fig. 2, se muestra la estructura de un decodificador 200 para recibir el tren de bits codificado desde el codificador de la Fig. 1. El tren de bits con codificación H.263 217 recibido se demultiplexa en coeficientes DCT cuantificados 201, datos de vector de movimiento 202 y un indicador de identificación de intertrama/intratrama 203 en el demultiplexor 216. Los coeficientes DCT cuantificados 201 se convierten en una imagen de error decodificada 206 tras ser procesados mediante un cuantificador inverso 204 y un convertidor DCT inverso 205. Esta imagen de error decodificada se suma a la imagen de salida 215 del selector de codificación de intertrama/intratrama 214 en el sumador 207 y la suma de estas imágenes se proporciona como imagen decodificada 208. La salida del selector de codificación de intertrama/intratrama se conmuta según el indicador de identificación de intertrama/intratrama 203. La imagen de predicción 212 utilizada en la codificación de intertrama se sintetiza en el sintetizador de imágenes de predicción 211. En este sintetizador, la posición de los bloques en la imagen decodificada 210 de la trama anterior almacenada en la memoria de trama 209 se desplaza según los datos del vector de movimiento 202. Por otra parte, en la codificación de intratrama, el selector de codificación de intertrama/intratrama proporciona la señal «0» 213 tal cual.

El documento US 5 379 351 A divulga un procedimiento y un sistema de decodificación con las características de la parte del preámbulo de la reivindicación 1.

Sumario de la invención

La imagen con codificación H.263 consiste en un plano de luminancia (plano Y) que contiene información de luminancia, y dos planos de crominancia (plano U y plano V) que contienen información de crominancia. Como es característico, en el momento en que la imagen tiene $2m$ píxeles en la dirección horizontal y $2n$ píxeles en la dirección vertical (siendo m y n enteros positivos), el plano Y tiene $2m$ píxeles en la dirección horizontal y $2n$ píxeles en la dirección vertical y los planos U y V tienen m píxeles en la dirección horizontal y n píxeles en la dirección vertical. La baja resolución del plano de crominancia se debe al hecho de que el sistema visual humano tiene una agudeza visual comparativa reducida respecto de las variaciones espaciales de la crominancia. El H.263, que recibe como entrada dicha imagen, lleva a cabo la codificación y la decodificación en unidades de bloques que se denominan macrobloques. La estructura de un macrobloque se muestra en la Fig. 3. El macrobloque consiste en tres bloques: un bloque Y, un bloque U y un bloque V. El tamaño del bloque Y 301 que contiene la información de luminancia es de 16×16 píxeles, y el tamaño del bloque U 302 y el bloque V 303 que contienen la información de crominancia es de 8×8 píxeles.

En el H. 263, se aplica a cada bloque una correspondencia de bloques con una precisión de medio píxel. En consecuencia, cuando el vector de movimiento estimado se define como (u, v) , u y v son integrales múltiples de la mitad de la distancia entre los píxeles. Dicho de otro modo, $1/2$ se utiliza como unidad mínima. La configuración del procedimiento de interpolación utilizado para los valores de intensidad (de ahora en adelante, los valores de intensidad para «luminancia» y «crominancia» se denominan con el término general de «valor de intensidad») se muestra en la Fig. 4. Cuando se efectúa la interpolación descrita en la Ecuación 2, los cocientes de la división se redondean al entero más cercano y, además, cuando el cociente tiene un valor que es un semientero (es decir, $0,5$ sumado a un entero), el redondeo es al siguiente entero en la dirección que se aleja de cero. Es decir, en la Fig. 4, cuando los valores de intensidad para 401, 402, 403 y 404 son respectivamente L_a , L_b , L_c y L_d (L_a , L_b , L_c y L_d son enteros no negativos), los valores de intensidad interpolados l_a , l_b , l_c e l_d (l_a , l_b , l_c e l_d son enteros no negativos) de las posiciones 405, 406, 407 y 408 se expresan mediante la siguiente ecuación:

[Ecuación 3]

$$\begin{aligned} I_a &= L_a \\ I_b &= \left[(L_a + L_b + 1) / 2 \right] \\ I_c &= \left[(L_a + L_c + 1) / 2 \right] \quad \dots(3) \\ I_d &= \left[(L_a + L_b + L_c + L_d + 2) / 4 \right] \end{aligned}$$

en la que «[]» denota el truncamiento al entero que por defecto es más cercano a cero (es decir la parte fraccionaria se descarta). Las previsiones de los errores ocasionados por este redondeo a enteros se estiman de la forma descrita a continuación. Se presupone que la probabilidad de que se utilice el valor de intensidad de las posiciones 405, 406, 407 y 408 de la Fig. 4 es en todos los casos del 25%. Cuando se halla el valor de intensidad la para la posición 405, el error de redondeo será evidentemente cero. Asimismo, cuando se halla el valor de intensidad Ib para la posición 406, el error es cero cuando La+Lb es un número par, y 1/2 cuando es un número impar. Si la probabilidad de que La+Lb sea un número par y un número impar es en ambos casos del 50%, la previsión de error será $0 \times 1/2 + 1/2 \times 1/2 = 1/4$. Además, cuando se halla el valor de intensidad Ic para la posición 407, la previsión de error es 1/4, igual que para Ib. Cuando se halla el valor de intensidad Id para la posición 408, el error, cuando el residuo de La+Lb+Lc+Ld dividido por cuatro es 0, 1, 2 y 3, es 0, -1/4, 1/2 y 1/4, respectivamente. Si se presupone que la probabilidad de que el residuo sea 0, 1, 2 y 3 es en todos casos la misma (es decir, el 25%), la previsión de error es $0 \times 1/4 - 1/4 \times 1/4 + 1/2 \times 1/4 + 1/4 \times 1/4 = 1/8$. Como se ha indicado anteriormente, si se presupone que la probabilidad de que se utilice el valor de intensidad de las posiciones 405 a 408 es igual en todos los casos, la previsión final de error es $0 \times 1/4 + 1/4 \times 1/4 + 1/4 \times 1/4 + 1/8 \times 1/4 = 5/32$. Esto indica que cada vez que se efectúa la compensación de movimiento mediante la correspondencia de bloques, se produce un error de 5/32 en el valor de intensidad de píxel. En general, en la codificación de baja velocidad, no puede utilizarse un número de bits suficiente para la codificación de la diferencia de error de intertrama y, por consiguiente, la etapa de cuantificación del coeficiente DCT tiende a ser prolongada. En consecuencia, los errores causados por la compensación de movimiento sólo se corrigen si son muy graves. Cuando la codificación de intertrama se lleva a cabo de forma continuada, sin que se lleve a cabo la codificación de intratrama en dicho entorno, los errores tienden a acumularse y a ocasionar defectos en la imagen reconstruida.

Tal como se ha descrito anteriormente, el número de píxeles es de alrededor de la mitad tanto en la dirección vertical como en la horizontal del plano de crominancia. Por consiguiente, para los vectores de movimiento del bloque U y el bloque V, se utiliza la mitad del valor del vector de movimiento del bloque Y para los componentes vertical y horizontal. Puesto que los componentes horizontal y vertical del vector de movimiento del bloque Y son integrales múltiples de 1/2, los componentes del vector de movimiento de los bloques U y V aparecerán como integrales múltiples de 1/4 (precisión de un cuarto de píxel) si se efectúa una división corriente. No obstante, debido a la gran complejidad de los cálculos del procedimiento de interpolación de la intensidad para los vectores de movimiento con una precisión de un cuarto de píxel, los vectores de movimiento para los bloques U y V se redondean con una precisión de medio píxel en el H.263. El procedimiento de redondeo utilizado en el H.263 es el descrito a continuación. Según la definición proporcionada anteriormente, (u, v) denota el vector de movimiento del macrobloque (que es igual al vector de movimiento del bloque Y). Si se presupone que r es un entero y que s es un entero no negativo inferior a 4, u/2 puede reescribirse como $u/2 = r + s/4$. Cuando s es 0 o 2, no se requiere ningún redondeo, puesto que u/2 es una integral múltiple de 1/2. No obstante, cuando s es igual a 1 o 3, el valor de s se redondea a 2. Si se incrementa la probabilidad de que s tome el valor de 2 mediante este procedimiento de redondeo, puede resaltarse el efecto del filtrado de la compensación de movimiento. Cuando la probabilidad de que el valor de s antes del redondeo sea 0, 1, 2 y 3 es en todos casos del 25 por ciento, la probabilidad de que s sea 0 o 2 después del redondeo será, respectivamente, del 25 por ciento y del 75% por ciento. El procedimiento descrito anteriormente relacionado con el componente horizontal u del vector de movimiento también se aplica al componente vertical v. En consecuencia, en el bloque U y el bloque V, la probabilidad de que se utilice el valor de intensidad de la posición 401 es $1/4 \times 1/4 = 1/16$, y la probabilidad de que se utilice el valor de intensidad de las posiciones 402 y 403 es en ambos casos $1/4 \times 3/4 = 3/16$, mientras que la probabilidad de que se utilice el valor de intensidad de la posición 404 es $3/4 \times 3/4 = 9/16$. Si se utiliza el mismo procedimiento que antes, la previsión de error del valor de intensidad es $0 \times 1/16 + 1/4 \times 3/16 + 1/4 \times 3/16 + 1/4 \times 3/16 + 1/8 \times 9/16 = 21/128$. Tal como se ha descrito anteriormente para el bloque Y, cuando la codificación de intertrama se efectúa de forma continuada, se produce el problema de la acumulación de errores.

Como se ha indicado anteriormente, en los procedimientos de codificación y decodificación de secuencias de imágenes en los que se lleva a cabo la predicción de intertrama y la cuantificación de la intensidad de la crominancia, surge el problema de la acumulación de errores de redondeo. Este error de redondeo se genera cuando el valor de intensidad de luminancia o crominancia se cuantifica durante la generación de la imagen de predicción de intertrama.

A la vista de los problemas anteriores, uno de los objetivos de la presente invención es, por lo tanto, mejorar la calidad de la imagen reconstruida impidiendo la acumulación de errores.

Para conseguir este objetivo, los errores de acumulación se previenen limitando la aparición de errores o llevando a cabo una operación para anular los errores que se han producido. La presente invención está definida en la reivindicación 1.

5 Breve descripción de los dibujos

La Figura 1 es un diagrama de bloques que muestra la disposición del codificador de imágenes H.263.

La Figura 2 es un diagrama de bloques que muestra la disposición del decodificador de imágenes H.263.

La Figura 3 es un diagrama de bloques que muestra la estructura del macrobloque.

La Figura 4 es un dibujo que muestra el procedimiento de interpolación de los valores de intensidad para la correspondencia de bloques con una precisión de medio píxel.

La Figura 5 es un dibujo que muestra una secuencia de imágenes codificadas,

La Figura 6 es un diagrama de bloques que muestra un dispositivo de codificación de imágenes de software.

La Figura 7 es un diagrama de bloques que muestra un dispositivo de decodificación de imágenes de software.

La Figura 8 es un diagrama de flujo que muestra un ejemplo de procesamiento en el dispositivo de codificación de imágenes de software.

La Figura 9 es un diagrama de flujo que muestra un ejemplo de procesamiento de decisión de modalidad de codificación para el dispositivo de codificación de imágenes de software.

La Figura 10 es un diagrama de flujo que muestra un ejemplo de procesamiento de estimación de movimiento y de compensación de movimiento en el dispositivo de codificación de imágenes de software.

La Figura 11 es un diagrama de flujo que muestra el procesamiento en el dispositivo de decodificación de imágenes de software.

La Figura 12 es un diagrama de flujo que muestra un ejemplo de procesamiento de compensación de movimiento en el aparato de decodificación de imágenes de software.

La Figura 13 es un dibujo que muestra un ejemplo de medios de memoria en los que se registra un tren de bits codificado generado por un procedimiento de codificación que proporciona trenes de bits que incluyen tramas I, P+ y P-.

La Figura 14 es un grupo de dibujos que muestran ejemplos concretos de dispositivos que utilizan un procedimiento de codificación en el que las tramas P+ y P- coexisten.

La Figura 15 es un dibujo que muestra un ejemplo de medios de memoria en los que se registra un tren de bits codificado generado por un procedimiento de codificación que proporciona trenes de bits que incluyen tramas I, B, P+ y P-.

La Figura 16 es un diagrama de bloques que muestra un ejemplo de unidad de correspondencia de bloques incluida en un dispositivo que utiliza un procedimiento de codificación en el que las tramas P+ y P- coexisten.

La Figura 17 es un diagrama de bloques que muestra el sintetizador de imágenes de predicción, incluido en un dispositivo para decodificar trenes de bits codificados por un procedimiento de codificación en el que las tramas P+ y P- coexisten.

55 Descripción detallada de las formas de realización preferidas

En primer lugar, se considerará en qué circunstancias se produce la acumulación de errores de redondeo en relación con la «técnica anterior». En la Fig. 5, se muestra un ejemplo de una secuencia de imágenes codificada mediante procedimientos de codificación que pueden llevar a cabo tanto la predicción unidireccional como la predicción bidireccional (por ejemplo, MPEG. 1, MPEG.2 y H.263). La imagen 501 es una trama codificada mediante codificación de intratrama y se denomina trama I. En cambio, las imágenes 503, 505, 507 y 509 se denominan tramas P y se codifican mediante codificación de intertrama unidireccional utilizando la trama I o P anterior como imagen de referencia. En consecuencia, por ejemplo, cuando se codifica la imagen 505, la imagen 503 se utiliza como imagen de referencia y se lleva a cabo la predicción de intertrama. Las imágenes 502, 504, 506 y 508 se denominan tramas B y la predicción de intertrama bidireccional se lleva a cabo utilizando la trama I o P anterior y posterior. La trama B se caracteriza porque no se utiliza como imagen de referencia cuando se lleva a cabo la

predicción de intertrama. Puesto que en las tramas I no se lleva a cabo la compensación de movimiento, no se produce el error de redondeo ocasionado por la compensación de movimiento. En cambio, en las tramas P se efectúa compensación de movimiento y, además, las tramas P son utilizadas como imagen de referencia por otras tramas P o B, lo que puede ser origen de una acumulación de errores de redondeo. En las tramas B, por otro lado, se efectúa compensación de movimiento y, por consiguiente, el efecto de la acumulación de errores de redondeo aparece en la imagen reconstruida. Sin embargo, debido al hecho de que las tramas B no son utilizadas como imágenes de referencia, las tramas B no pueden ser el origen de una acumulación de errores de redondeo. Por lo tanto, si se puede prevenir la acumulación de errores de redondeo en la trama P, se podrán mitigar los efectos perjudiciales de los errores de redondeo en la totalidad de la secuencia de imágenes. En el H.263, existe una trama para codificar una trama P y una trama B denominada trama PB. (Por ejemplo, las tramas 503 y 504 pueden codificarse ambas como una trama PB.) Si las dos tramas combinadas se tratan como tramas separadas, entonces puede aplicarse el mismo principio que antes. Dicho de otro modo, si se toman medidas contra los errores de redondeo para la parte de la trama P de una trama PB, podrá prevenirse la acumulación de errores.

Los errores de redondeo se producen durante la interpolación de los valores de intensidad cuando un valor obtenido mediante una división normal (división cuyo resultado es un número real) es un semientero (0,5 sumado a un entero) y este resultado se redondea luego al siguiente entero en la dirección que se aleja de cero. Por ejemplo, cuando se divide por 4 para hallar un valor de intensidad interpolado, los errores de redondeo para los casos en que el residuo es 1 y 3 tienen valores absolutos iguales pero signos diferentes. En consecuencia, los errores de redondeo ocasionados en estos dos casos se anulan cuando se calcula la previsión de errores de redondeo (en términos más generales, cuando se divide por un entero positivo d' , los errores de redondeo ocasionados cuando el residuo es t y $d' - t$ se anulan). No obstante, cuando el residuo es 2, es decir, cuando el resultado de la división normal es un semientero, el error de redondeo no puede anularse, lo que lleva a la acumulación de errores. Para resolver este problema, puede utilizarse un procedimiento que permite la utilización de dos procedimientos de redondeo. Los dos procedimientos de redondeo utilizados aquí son: un procedimiento de redondeo que redondea los semienteros a los que por exceso están más alejados de cero y un procedimiento de redondeo que redondea los semienteros a los que por defecto son más cercanos a cero. Combinando la utilización de estos dos procedimientos de redondeo, podrán anularse los errores de redondeo. De ahora en adelante, el procedimiento de redondeo que redondea el resultado de la división normal al entero más cercano y que redondea los valores de los semienteros a los que por exceso están más alejados de cero se denominará «redondeo positivo». Además, el procedimiento de redondeo que redondea el resultado de la división normal al entero más cercano y redondea los valores de los semienteros a los que por defecto son más cercanos a cero se denominará «redondeo negativo». El procedimiento de redondeo positivo utilizado en la correspondencia de bloques con una precisión de medio píxel se muestra en la Ecuación 3. En cambio, cuando se utiliza redondeo negativo, esta ecuación puede reescribirse de la forma mostrada a continuación.

[Ecuación 4]

$$\begin{aligned} I_a &= L_a \\ I_b &= [(L_a + L_b) / 2] \\ I_c &= [(L_a + L_c) / 2] \quad \dots(4) \\ I_d &= [(L_a + L_b + L_c + L_d + 1) / 4] \end{aligned}$$

De ahora en adelante, los procedimientos de compensación de movimiento que llevan a cabo el redondeo positivo y negativo para la síntesis de imágenes de predicción de intertrama se denominarán «compensación de movimiento mediante redondeo positivo» y «compensación de movimiento mediante redondeo negativo», respectivamente. Además, para las tramas P que utilizan correspondencia de bloques con una precisión de medio píxel para la compensación de movimiento, las tramas que utilizan redondeo positivo se denominarán «tramas P+» y las tramas que utilizan redondeo negativo se denominarán «trama P-» (según esta definición, todas las tramas P del H.263 son tramas P+). La previsión de errores de redondeo en las tramas P+ y P- tiene los mismos valores absolutos pero signos diferentes. En consecuencia, la acumulación de errores de redondeo puede impedirse si se colocan tramas P+ y tramas P- de forma alternada a lo largo del eje del tiempo. En el ejemplo de la Fig. 5, si las tramas 503 y 507 se establecen como tramas P+ y las tramas 505 y 509 se establecen como tramas P-, entonces podrá ejecutarse este procedimiento. La distribución alternada de tramas P+ y tramas P- lleva a la utilización de una trama P+ y una trama P- en la predicción bidireccional para las tramas B. En general, el promedio de la imagen de predicción hacia delante (es decir, la imagen de predicción sintetizada mediante la trama 503 cuando se está codificando la trama 504 de la Fig. 5) y la imagen de predicción hacia atrás (es decir, la imagen de predicción sintetizada mediante la trama 505 cuando se está codificando la trama 504 de la Fig. 5) se utiliza frecuentemente para sintetizar la imagen de predicción para las tramas B. Esto significa que la utilización de una trama P+ (que tiene un valor positivo para la previsión del error de redondeo) y una trama P- (que tiene un valor negativo para la previsión del error de redondeo) en la predicción bidireccional de una trama B anula con eficacia los efectos de los errores de redondeo. Como se ha indicado anteriormente, el procedimiento de redondeo en la trama B no ocasionará acumulación de errores. En consecuencia, tampoco la aplicación del mismo procedimiento de redondeo a todas las tramas B planteará ningún problema. Por ejemplo, aunque la compensación de movimiento mediante redondeo positivo se aplique a todas las tramas B 502, 504, 506 y 508 de la Fig. 5, no se provocará ninguna degradación importante de las imágenes

decodificadas. Preferentemente, sólo se lleva a cabo un tipo de redondeo para una trama B, para simplificar el procedimiento de decodificación de la trama B.

En la Fig. 16, se muestra la sección de correspondencia de bloques 1600 de un codificador de imágenes según el procedimiento de compensación de movimiento descrito anteriormente que utiliza diversos procedimientos de redondeo. En esta Figura, se utilizan números idénticos a los de otros dibujos para indicar partes iguales. Si se sustituye la sección de correspondencia de bloques 116 de la Fig. 1 por 1600, podrán utilizarse diversos procedimientos de redondeo. El procesamiento de estimación del movimiento entre la imagen de entrada 101 y la imagen decodificada de la trama anterior se lleva a cabo en un estimador de movimiento 1601 y el resultado del mismo es la información de movimiento 120. Esta información de movimiento se utiliza en la síntesis de la imagen de predicción llevada a cabo en un sintetizador de imágenes de predicción 1603. El dispositivo de determinación de procedimiento de redondeo 1602 determina si debe emplearse un redondeo positivo o un redondeo negativo como procedimiento de redondeo para la trama que se está codificando actualmente. La información 1604 relativa al procedimiento de redondeo que se determina se introduce en el sintetizador de imágenes de predicción 1603. En este sintetizador de imágenes de predicción 1603, se sintetiza una imagen de predicción 117 y se proporciona basándose en el procedimiento de redondeo determinado mediante la información 1604. En la sección de correspondencia de bloques 116 de la Fig. 1, no existe ningún elemento equivalente a los elementos 1602 y 1604 de la Fig. 16, y la imagen de predicción se sintetiza sólo mediante redondeo positivo. Asimismo, puede proporcionarse el procedimiento de redondeo 1605 determinado en la sección de correspondencia de bloques y esta información puede luego multiplexarse en el tren de bits y transmitirse.

En la Fig. 17, se muestra el sintetizador de imágenes de predicción 1700 de un decodificador de imágenes que puede decodificar trenes de bits generados mediante un procedimiento de codificación en el que se utilizan diversos procedimientos de redondeo. En esta Figura, se utilizan números idénticos a los de otros dibujos para indicar partes iguales. Si se sustituye el sintetizador de imágenes de predicción 211 de la Fig. 2 por 1700, podrán utilizarse diversos procedimientos de redondeo. En el dispositivo de determinación de procedimiento de redondeo 1701, se determina el procedimiento de redondeo adecuado para la síntesis de imágenes de predicción en el procedimiento de decodificación. Para llevar a cabo la decodificación correctamente, el procedimiento de redondeo seleccionado aquí debe ser el mismo procedimiento de redondeo que se seleccionó para la codificación. Por ejemplo, el codificador y el decodificador pueden compartir la siguiente regla: cuando la trama actual es una trama P y el número de tramas P (incluida la trama actual) contadas desde la trama I más reciente es impar, entonces la trama actual es una trama P+. Cuando este número es par, entonces la trama actual es una trama P-. Si el dispositivo de determinación de procedimiento de redondeo de la sección de codificación (por ejemplo, 1602 en la Fig. 16) y el dispositivo de determinación de procedimiento de redondeo 1701 cumplen esta regla común, entonces las imágenes podrán ser correctamente decodificadas. La imagen de predicción se sintetiza en el sintetizador de imágenes de predicción 1703 mediante la información de movimiento 202, la imagen de decodificación 210 de la trama anterior y la información 1702 relativa al procedimiento de redondeo determinado de la forma que se acaba de describir. Una vez que se ha generado esta imagen de predicción 212, ésta se utiliza para la síntesis de la imagen decodificada. Como alternativa al caso mencionado, puede considerarse también el caso en que la información relativa al procedimiento de redondeo se multiplexa en el tren de datos transmitido (dicho tren de bits puede generarse en el codificador proporcionando la información 1605 relativa al procedimiento de redondeo desde la sección de correspondencia de bloques ilustrada en la Fig. 16). En dicho caso, el dispositivo de determinación de procedimiento de redondeo 1701 no se utiliza y la información 1704 relacionada con el procedimiento de redondeo extraída del tren de bits codificado se utiliza en el sintetizador de imágenes de predicción 1703.

Aparte del codificador de imágenes y el decodificador de imágenes que utilizan los circuitos habituales y los chips habituales de la técnica convencional mostrados en la Fig. 1 y la Fig. 2, la presente invención también puede aplicarse a codificadores de imágenes de software y a decodificadores de imágenes de software que utilizan procesadores de uso general. En la Fig. 6 y la Fig. 7, se muestran un codificador de imágenes de software 600 y un decodificador de imágenes de software 700. En el codificador de imágenes de software 600, la imagen de entrada 601 se almacena en primer lugar en la memoria de trama de entrada 602 y, desde ésta, el procesador de uso general 603 obtiene la información y lleva a cabo la codificación. El programa para dirigir este procesador de uso general se copia desde un dispositivo de memoria 608 que puede ser un disco duro, un disquete, etc. y se almacena en una memoria de programas 604. Este procesador de uso general también utiliza una memoria de procesamiento 605 para llevar a cabo la codificación. La información de codificación proporcionada por el procesador de uso general se almacena temporalmente en la memoria tampón de salida 606 y luego se proporciona como un tren de bits codificado 607.

En la Fig. 8, se muestra un diagrama de flujo del software de codificación (medios de registro legibles por ordenador). El procedimiento empieza en 801, y el valor 0 se asigna a la variable N en 802. A continuación, en 803 y 804, se asigna el valor 0 a N cuando el valor de N es 100. N es un contador del número de tramas. Se suma el valor 1 a cada trama cuyo procesamiento ha finalizado, estando permitidos los valores entre 0 y 99 durante la codificación. Cuando el valor de N es 0, la trama actual es una trama I. Cuando N es un número impar, la trama actual es una trama P+ y cuando N es un número par distinto a 0, la trama actual es una trama P-. Cuando el límite superior del valor de N es 99, significa que se codifica una trama I una vez que se han codificado 99 tramas P (tramas P+ o tramas P-). Introduciendo siempre una trama I en un cierto número de tramas codificadas, pueden obtenerse los

siguientes beneficios: (a) la acumulación de errores debidos a un desajuste entre el procesamiento del codificador y el decodificador puede prevenirse (por ejemplo, un desajuste en el cálculo de la DCT) y (b) la carga de procesamiento para obtener la imagen reproducida de la trama deseada a partir de los datos codificados (acceso aleatorio) se reduce. El valor N óptimo varía cuando cambia el rendimiento del codificador o el entorno en el que éste se utiliza. Esto no significa, sin embargo, que el valor de N deba ser siempre 100. El procedimiento para determinar el procedimiento de redondeo y la modalidad de codificación para cada trama se efectúa en 805. El diagrama de flujo que ilustra en detalle esta operación se muestra en la Fig. 9. En primer lugar, se comprueba en 901 si N es 0 o no. Si N es 0, entonces se proporciona «I» como información de distinción de la modalidad de predicción a la memoria tampón de salida en 902. Esto significa que la imagen que se va a codificar se codificará como una trama I. Aquí, «proporcionar a la memoria tampón de salida» significa que, tras haber estado almacenada en la memoria tampón de salida, la información se proporciona a un dispositivo externo como parte del tren de bits codificado. Cuando N no es 0, entonces se comprueba en 904 si N es un número impar o par. Cuando N es un número impar, se proporciona «+» a la memoria tampón de salida como información de distinción para el procedimiento de redondeo en 905 y la imagen se codificará como una trama P+. Por otro lado, cuando N es un número par, se transmite «-» a la memoria tampón de salida como información de distinción para el procedimiento de redondeo en 906, y la imagen que se va a codificar se codificará como una trama P-. El procedimiento vuelve otra vez a la Fig. 8, donde tras determinarse la modalidad de codificación en 805, la imagen de entrada se almacena en la memoria de trama A en 806. En el presente documento, «memoria de trama A» indica una parte de la zona de memoria (por ejemplo, la zona de memoria mantenida en la memoria de 605 en la Fig. 6) del codificador de software. En 807, se comprueba si la trama que se está codificando actualmente es una trama I. Cuando se comprueba que no es una trama I, se lleva a cabo la estimación de movimiento y la compensación de movimiento en 808. El diagrama de flujo de la Fig. 10 muestra detalles de este procedimiento llevado a cabo en 808. En primer lugar, en 1001, se lleva a cabo la estimación del movimiento entre las imágenes almacenadas en las memorias de trama A y B (tal como se indica al final de este párrafo, la imagen decodificada de la trama anterior se almacena en la memoria de trama B). Se halla el vector de movimiento de cada bloque y éste se envía a la memoria tampón de salida. A continuación, en 1002, se comprueba si la trama actual es o no una trama P+. Cuando la trama actual es una trama P+, la imagen de predicción se sintetiza en 1003 utilizando redondeo positivo y esta imagen de predicción se almacena en la memoria de trama C. Por otra parte, cuando la trama actual es una trama P-, la imagen de predicción se sintetiza en 1004 utilizando redondeo negativo y esta imagen de predicción se almacena en la memoria de trama C. Seguidamente, en 1005, se halla la imagen diferencial entre la memoria de trama A y la memoria de trama C y se almacena en la memoria de trama A. Entonces, el procedimiento vuelve otra vez a la Fig. 8. Antes de iniciarse el procesamiento en 809, la imagen de entrada se almacena en la memoria de trama A cuando la trama actual es una trama I, y la imagen diferencial entre la imagen de entrada y la imagen de predicción se almacena en la memoria A cuando la trama actual es una trama P (trama P+ o trama P-). En 809, se aplica una DCT a la imagen almacenada en la memoria de trama A, y los coeficientes DCT calculados se envían a la memoria tampón de salida una vez cuantificados. En 810, se efectúa la cuantificación inversa de los coeficientes DCT y se aplica una DCT inversa. La imagen obtenida aplicando una DCT inversa se almacena en la memoria de trama B. A continuación, en 811, se comprueba de nuevo si la trama actual es una trama I. Cuando la trama actual no es una trama I, las imágenes almacenadas en la memoria de trama B y C se suman y el resultado se almacena en la memoria de trama B. El procedimiento de codificación de una trama finaliza aquí, y la imagen almacenada en la memoria de trama B antes de llegar a 813 es la imagen reconstruida de esta trama (esta imagen es idéntica a la obtenida en la sección de decodificación). En 813, se comprueba si la trama cuya codificación acaba de terminar es la última trama de la secuencia. En caso afirmativo, el procedimiento de codificación finaliza. Si esta trama no es la última trama, se suma 1 a N en 814, el procedimiento vuelve otra vez a 803 y comienza la codificación de la siguiente trama.

En la Fig. 7, se muestra un decodificador de software 700. Una vez que el tren de bits codificado 701 ha sido almacenado temporalmente en la memoria tampón de entrada 702, este tren de bits se carga en el procesador de uso general 703. El programa para dirigir este procesador de uso general se copia desde un dispositivo de memoria 708 que puede ser un disco duro, un disquete, etc. y se almacena en la memoria de programas 704. Este procesador de uso general utiliza también una memoria de procesamiento 605 para efectuar la decodificación. La imagen decodificada obtenida mediante el procedimiento de decodificación se almacena temporalmente en la memoria tampón de salida 706 y luego se envía como imagen de salida 707.

En la Fig. 11, se muestra un diagrama de flujo del software de decodificación para el decodificador de software 700 mostrado en la Fig. 7. El procedimiento empieza en 1101 y, en 1102, se comprueba si se dispone de información de entrada. Cuando no hay información de entrada, el procedimiento de decodificación finaliza en 1103. Cuando hay información de entrada, la información de distinción de la modalidad de predicción se introduce en 1104. La palabra «entrada» utilizada aquí significa que la información almacenada en la memoria tampón de entrada (por ejemplo, 702 de la Fig. 7) es proporcionada por el procesador de uso general. En 1105, se comprueba si la información de distinción de modalidad de codificación es «I». Cuando no es «I», se introduce la información de distinción para el procedimiento de redondeo y se lleva a cabo la síntesis de la imagen de predicción de intertrama en 1107. En la Fig. 12, se muestra un diagrama de flujo que muestra los detalles de la operación efectuada en 1107. En 1201, se introduce un vector de movimiento para cada bloque. A continuación, en 1202, se comprueba si la información de distinción para el procedimiento de redondeo cargado en 1106 es «+». Cuando esta información es «+», la trama que se está decodificando actualmente es una trama P+. En este caso, la imagen de predicción se sintetiza

utilizando redondeo positivo en 1203, y la imagen de predicción se almacena en la memoria de trama D. Aquí, «memoria de trama D» indica una parte de la zona de memoria del decodificador de software (por ejemplo, esta zona de memoria se obtiene en la memoria de procesamiento 705 en la Fig. 7). Cuando la información de distinción del procedimiento de redondeo no es «+», la trama actual que se está decodificando es una trama P-. La imagen de predicción se sintetiza mediante redondeo negativo en 1204 y esta imagen de predicción se almacena en la memoria de trama D. En este momento, si una trama P+ se decodifica como una trama P- debido a algún tipo de error o, a la inversa, si una trama P- se decodifica como una trama P+, no se sintetiza la imagen de predicción correcta en el decodificador y la calidad de la imagen decodificada se deteriora. Tras la sintetización de la imagen de predicción, la operación vuelve a la Fig. 11 y los coeficientes DCT cuantificados se introducen en 1108. A continuación, se aplica cuantificación inversa y DCT inversa a estos coeficientes y la imagen resultante se almacena en la memoria de trama E. En 1109, se comprueba otra vez si la trama que se decodifica actualmente es una trama I. Si la trama actual no es una trama I, las imágenes almacenadas en la memoria de trama D y E se suman en 1110 y la imagen resultante de la suma se almacena en la memoria de trama E. La imagen almacenada en la memoria de trama E antes del inicio del procedimiento en 1111 es la imagen reconstruida. Esta imagen almacenada en la memoria de trama E se proporciona a la memoria de trama de salida (por ejemplo, 706 en la Fig. 7) en 1111 y, a continuación, se obtiene desde el decodificador como la imagen reconstruida. El procedimiento de decodificación de una trama finaliza aquí y el procedimiento de la siguiente trama empieza otra vez desde 1102.

Cuando se ejecuta software basado en el diagrama de flujo mostrado en las Figs. 8 a 12 en los codificadores o decodificadores de imágenes de software, se obtiene el mismo efecto que cuando se utilizan circuitos habituales y chips habituales.

En la Fig. 13, se muestran unos medios de memoria (medios de registro) en los que se registra el tren de bits generado por el codificador de software 601 de la Fig. 6. Se presupone que el algoritmo mostrado en los diagramas de flujo de las Figs. 8 a 10 se utiliza en el codificador de software. La información digital se registra concéntricamente en un disco de registro 1301 capaz de registrar información digital (por ejemplo, un disco magnético, un disco óptico, etc.). Una parte 1302 de la información registrada en este disco digital incluye: información de distinción de modalidad de predicción 1303, 1305, 1308, 1311 y 1314; información de distinción de procedimiento de redondeo 1306, 1309, 1312 y 1315 e información de vector de movimiento y coeficientes DCT 1304, 1307, 1310, 1313 y 1316. La información que representa «I» se registra en 1303, «P» se registra en 1305, 1308, 1311 y 1314, «+» se registra en 1306 y 1312 y «-» se registra en 1309 y 1315. En este caso «I» y «+» pueden representarse mediante un solo bit de 0, y «P» y «-» pueden representarse mediante un solo bit de 1. Mediante esta representación, el decodificador puede interpretar correctamente la información registrada y se sintetiza la imagen reconstruida correcta. Si se almacena un tren de bits codificado en medios de memoria utilizando el procedimiento descrito anteriormente, podrá impedirse la acumulación de errores de redondeo cuando se lee y decodifica el tren de bits.

En la Fig. 15, se muestran unos medios de memoria donde se registra el tren de bits de los datos codificados de la secuencia de imágenes mostrada en la Fig. 5. El tren de bits registrado incluye información relacionada con las tramas P+, P- y B. De la misma forma que en el disco de registro 1301 de la Fig. 13, la información digital se registra concéntricamente en un disco de registro 1501 capaz de registrar información digital (por ejemplo, un disco magnético, un disco óptico, etc.). Una parte 1502 de la información digital registrada en este disco digital incluye: información de distinción de modalidad de predicción 1503, 1505, 1508, 1510 y 1513; información de distinción de procedimiento de redondeo 1506 y 1512 e información de vector de movimiento y coeficientes DCT 1504, 1507, 1509, 1511 y 1514. La información que representa «I» se registra en 1503, «P» se registra en 1505 y 1510, «B» se registra en 1508 y 1513, «+» se registra en 1506 y «-» se registra en 1511. En este caso, «I», «P» y «B» pueden representarse, respectivamente, mediante valores de dos bits 00, 01 y 10, y «+» y «-» pueden representarse, respectivamente, mediante valores de un bit 0 y 1. Mediante esta representación, el decodificador puede interpretar correctamente la información registrada y se sintetiza la imagen reconstruida correcta. En la Fig. 15, la información relacionada con la trama 501 (trama I) de la Fig. 5 es 1503 y 1504, la información relacionada con la trama 502 (trama B) es 1508 y 1509, la información relacionada con la trama 503 (trama P+) es 1505 y 1507, la información relacionada con la trama 504 (trama B) es 1513 y 1514 y la información relacionada con la trama 505 (trama P-) es 1510 y 1512. Cuando la codificación de secuencias de imágenes se efectúa con tramas B, el orden de transmisión y el orden de presentación de las tramas suele ser diferente. Esto es debido a que es necesario codificar la imagen de referencia anterior y posterior antes de sintetizar la imagen de predicción para la trama B. En consecuencia, a pesar de que la trama 502 se presenta antes que la trama 503, la información relacionada con la trama 503 se transmite antes que la información relacionada con la trama 502. Como se ha descrito anteriormente, no es necesario utilizar diversos procedimientos de redondeo para las tramas B, puesto que la compensación de movimiento en las tramas B no ocasiona acumulación de errores de redondeo. Por lo tanto, como se muestra en este ejemplo, la información relativa a los procedimientos de redondeo (p.ej., «+» y «-») no se transmite para las tramas B. Entonces, por ejemplo, aunque sólo se aplique redondeo positivo a las tramas B, no se produce el problema de la acumulación de errores de redondeo. Si se almacenan trenes de bits codificados que contienen información relativa a las tramas B en medios de memoria de la forma descrita anteriormente, la acumulación de errores de redondeo puede impedirse cuando se lee y decodifica este tren de bits.

En la Fig. 14, se muestran ejemplos concretos de codificadores y decodificadores que utilizan el procedimiento

descrito en esta memoria. El procedimiento de codificación y decodificación de imágenes puede utilizarse instalando software de codificación y de decodificación de imágenes en un ordenador 1401. Este software se registra en algún tipo de medios de memoria (CD-ROM, disquete, disco duro, etc.) 1412, se copia en un ordenador y luego se utiliza. Además, el ordenador puede utilizarse como terminal de transmisión de imágenes conectando el ordenador a las líneas de comunicación. También es posible instalar el procedimiento de decodificación descrito en esta memoria en un dispositivo de reproducción 1403 que lee y decodifica el tren de bits codificado registrado en los medios de memoria 1402. En este caso, la señal imagen reconstruida puede presentarse en un monitor de televisión 1404. El dispositivo 1403 puede utilizarse sólo para leer el tren de bits codificado y, en este caso, el dispositivo de decodificación puede instalarse en el monitor de televisión 1404. Como bien se sabe, la transmisión de datos digitales puede realizarse utilizando ondas de satélite y ondas terrestres. También puede instalarse un dispositivo de decodificación en un receptor de televisión 1405 capaz de recibir dichas transmisiones digitales. Asimismo, también puede instalarse un dispositivo de decodificación dentro de una caja de adaptación multimedia 1404 conectada a una antena de ondas de satélite/terrestres, o un cable 1408 de un sistema de televisión por cable, para que las imágenes reconstruidas puedan presentarse en un monitor de televisión 1410. En este caso, el dispositivo de decodificación puede incluirse en el monitor de televisión en vez de en la caja de adaptación multimedia, como en el caso de 1404. La disposición del sistema de difusión por satélite digital se muestra en 1413, 1414 y 1415. La información de vídeo del tren de bits codificado se transmite desde una estación de radiodifusión 1413 hasta un satélite de comunicación o radiodifusión 1414. El satélite recibe esta información, la envía a un domicilio 1415 provisto de equipo para recibir programas de radiodifusión por satélite, y la información de vídeo se reconstruye y presenta en este domicilio en dispositivos como, por ejemplo, un receptor de televisión o una caja de adaptación multimedia. La transmisión de imágenes digitales mediante terminales móviles 1406 ha sido objeto últimamente de una considerable atención, debido a que permite la transmisión de imágenes a muy bajas velocidades binarias. Los terminales portátiles digitales pueden dividirse en tres tipos de categorías: transceptores que incluyen un codificador y un decodificador; transmisores que sólo incluyen un codificador y receptores que sólo incluyen un decodificador. Puede instalarse un dispositivo de codificación en un grabador de vídeo 1407. La cámara también puede utilizarse sólo para capturar la señal de vídeo y esta señal puede proporcionarse a un codificador habitual 1411. Tanto los dispositivos como los sistemas mostrados en este dibujo pueden estar provistos del procedimiento de codificación o decodificación descrito en esta memoria. Utilizando este procedimiento de codificación o decodificación en estos dispositivos o sistemas, pueden obtenerse imágenes de calidad superior a la obtenida con tecnologías convencionales.

Las siguientes variantes están claramente incluidas dentro del alcance de la presente invención.

- (i) Un prerequisite del principio descrito anteriormente es la utilización de correspondencia de bloques como procedimiento de compensación de movimiento. No obstante, la presente invención puede aplicarse además a todos los procedimientos de codificación y decodificación de secuencias de imágenes, en los que la compensación de movimiento se lleva a cabo tomando un valor para los componentes vertical y horizontal del vector de movimiento del píxel que es distinto al entero múltiple del período de muestreo en las direcciones vertical y horizontal del píxel y, a continuación, hallando mediante interpolación el valor de intensidad de una posición en la que no está presente el valor de muestreo. Entonces, por ejemplo, la compensación de movimiento global descrita en la solicitud de patente japonesa n.º Hei 08-060572 y la predicción de distorsión descrita en la solicitud de patente japonesa n.º Hei 08-249601 son aplicables al procedimiento de la presente invención.
- (ii) En la descripción de la presente invención, sólo se ha mencionado el caso en el que se toma un valor de integral múltiple de $1/2$ para los componentes vertical y horizontal del vector de movimiento. Sin embargo, la presente invención también puede aplicarse de forma general a los procedimientos en los que son permitidas las integrales múltiples de $1/d$ (d es un entero positivo y también un número par) para los componentes horizontal y vertical del vector de movimiento. No obstante, cuando d es grande, el divisor para la división en la interpolación bilineal (cuadrado de d , véase la Ecuación 2) también es grande y, en cambio, la probabilidad de que los resultados de la división normal lleguen a un valor de 0,5 es baja. En consecuencia, cuando se efectúa sólo redondeo positivo, el valor absoluto de la previsión de errores de redondeo es pequeña y los efectos perjudiciales ocasionados por la acumulación de errores resultan menos llamativos. Asimismo, también puede aplicarse al procedimiento de la presente invención un procedimiento de compensación de movimiento en el que, por ejemplo, el valor de d es variable, y en el que se utiliza tanto el redondeo positivo como el negativo cuando d es inferior a un valor fijo y sólo se utiliza redondeo positivo o sólo redondeo negativo cuando el valor de d es superior a un valor fijo.
- (iii) Como se menciona en la técnica anterior, cuando se utiliza la DCT como procedimiento de codificación de errores, los efectos adversos de la acumulación de errores de redondeo son propensos a aparecer cuando el tamaño de la etapa de cuantificación del coeficiente DCT es grande. No obstante, también puede aplicarse un procedimiento a la presente invención, en el que, cuando el tamaño de la etapa de cuantificación de los coeficientes DCT es superior a un valor umbral, entonces se utiliza redondeo positivo y redondeo negativo. Cuando el tamaño de la etapa de cuantificación de los coeficientes DCT es inferior al valor umbral, entonces sólo se utiliza redondeo positivo o sólo se utiliza redondeo negativo.

(iv) En los casos en los que se producen acumulaciones de errores en el plano de la luminancia y en los casos en los que se producen acumulaciones de error en el plano de la crominancia, los efectos perjudiciales sobre las imágenes reconstruidas son en general más graves en el caso de las acumulaciones de errores en el plano de la crominancia. Esto es debido al hecho de que los casos en los que se producen cambios globales del color de la imagen son más llamativos que los casos en los que la imagen se oscurece o aclara ligeramente. No obstante, también puede aplicarse a la presente invención un procedimiento en el que se utiliza tanto redondeo positivo como redondeo negativo para la señal de crominancia, y sólo se utiliza redondeo positivo o redondeo negativo para la señal de luminancia.

Como se describe en referencia a las técnicas relacionadas, los vectores de movimiento de 1/4 de píxel de precisión obtenidos dividiendo por dos los vectores de movimiento de 1/2 píxel de precisión se redondean con una precisión de 1/2 píxel en el H.263. No obstante, añadiendo ciertos cambios a este procedimiento, el valor de previsión absoluto de errores de redondeo puede reducirse. En el H.263 mencionado en la técnica anterior, un valor que es la mitad del componente horizontal o vertical del vector de movimiento para el plano de la luminancia se expresa como $r + s/4$ (r es un entero, s es un entero menor que 4 y no inferior a 0), y cuando s es 1 o 3, se efectúa una operación de redondeo para obtener un 2. Esta operación puede cambiarse de la siguiente forma: cuando s es 1, se lleva a cabo una operación de redondeo para obtener un 0, y cuando s es 3, se suma un 1 a r para que s sea 0. Llevando a cabo estas operaciones, el número de veces que se interpolan los valores de intensidad de las posiciones 406 a 408 de la Fig.4 decididamente se reduce (la probabilidad de que los componentes horizontales y verticales del vector de movimiento sean un entero aumenta), de tal forma que el valor de previsión absoluto del error de redondeo disminuye. No obstante, aunque el tamaño del error producido en este procedimiento pueda limitarse, la acumulación de errores no puede evitarse por completo.

(v) La presente invención descrita en esta memoria es aplicable a un procedimiento que obtiene la imagen de predicción de intertrama definitiva calculando el promedio de las imágenes de predicción obtenidas mediante diferentes procedimientos de compensación de movimiento. Por ejemplo, en el procedimiento descrito en la solicitud de patente japonesa n.º Hei 8-2616, se calcula el promedio de las imágenes de predicción de intertrama obtenidas mediante los dos procedimientos siguientes: correspondencia de bloques en la que se asigna un vector de movimiento a cada bloque de 16x16 píxeles y correspondencia de bloques en la que se asigna un vector de movimiento a cada bloque de 8x8 píxeles. En este procedimiento, también se efectúa redondeo cuando se calcula el promedio de las dos imágenes de predicción. Cuando sólo se efectúa redondeo positivo de forma continuada en esta operación de promediación, se acumula un nuevo tipo de error de redondeo. Este problema puede resolverse utilizando diversos procedimientos de redondeo para esta operación de promediación. En este procedimiento, se efectúa redondeo negativo en la operación de promediación cuando se efectúa redondeo positivo en la correspondencia de bloques. Por el contrario, se utiliza redondeo positivo para la promediación cuando se utiliza redondeo negativo para la correspondencia de bloques. Utilizando diferentes procedimientos de redondeo para la promediación y la correspondencia de bloques, pueden anularse los errores de redondeo de dos fuentes diferentes dentro de la misma trama.

(vi) Cuando se utiliza un procedimiento para localizar de forma alternada tramas P+ y tramas P- a lo largo del eje del tiempo, el codificador o decodificador necesita determinar si la trama P procesada actualmente es una trama P+ o una trama P-. A continuación, se proporciona un ejemplo de dicho procedimiento de determinación. Un contador cuenta el número de tramas P tras la última trama I codificada o decodificada, siendo la trama P actual una trama P+ cuando el número es impar y una trama P- cuando el número es par (este procedimiento se denomina sistema implícito). Asimismo, existe un procedimiento, por ejemplo, que escribe, en la sección de encabezamiento de la información de imagen codificada, información para determinar si la trama P codificada actualmente en el codificador es una trama P+ o una trama P- (este procedimiento se denomina sistema explícito). A diferencia del procedimiento implícito, este procedimiento tiene gran capacidad para resistir los errores de transmisión, puesto que no hay necesidad de contar el número de tramas P.

Además, el procedimiento explícito tiene las ventajas indicadas a continuación. Como se describe en el apartado «Descripción de técnicas relacionadas» los estándares de codificación anteriores (como MPEG-1 o MPEG-2) sólo utilizan redondeo positivo para la compensación de movimiento. Esto significa, por ejemplo, que los dispositivos de estimación de movimiento/compensación de movimiento (por ejemplo, equivalentes al 106 de la Fig. 1) para MPEG-1/MPEG-2 disponibles en el mercado no son compatibles con los procedimientos de codificación que utilizan tanto tramas P+ como tramas P-. Se presupone que se dispone de un decodificador que puede decodificar trenes de bits generados mediante un procedimiento de codificación que utiliza tramas P+ y tramas P-. En este caso, si el decodificador se basa en el procedimiento implícito mencionado anteriormente, será difícil crear un codificador que genere trenes de bits que puedan decodificarse correctamente mediante el decodificador mencionado, utilizando el dispositivo de estimación/compensación de movimiento mencionado para MPEG-1/MPEG-2. No obstante, si el decodificador se basa en el procedimiento explícito mencionado, este problema puede resolverse. Un codificador que utiliza un dispositivo de estimación de movimiento/compensación de movimiento MPEG-1/MPEG-2 puede enviar de forma continuada tramas P+, escribiendo de forma continuada información de distinción de procedimiento de redondeo que indica redondeo positivo en el encabezamiento de la información de la trama. Una vez finalizada esta acción, un decodificador basado en el procedimiento explícito puede decodificar correctamente el tren de bits

generado por este codificador. Como es obvio, en dicho caso, será más probable que se produzca una acumulación de errores de redondeo, puesto que sólo están presentes las tramas P+. No obstante, la acumulación de errores no es un problema grave en los casos en que el codificador utiliza sólo pequeños valores como tamaño de etapa de cuantificación para los coeficientes DCT (el codificador habitual utilizado sólo para la codificación de alta velocidad es un ejemplo de dichos codificadores). Además de la interoperabilidad entre estándares anteriores, el procedimiento explícito presenta también las siguientes ventajas: (a) el coste del equipo para los codificadores habituales de alta velocidad y los codificadores no propensos a la acumulación de errores de redondeo debido a la introducción frecuente de tramas I puede reducirse instalando sólo redondeo positivo o redondeo negativo como procedimiento de redondeo de valores de píxel para la compensación de movimiento y (b) los anteriores codificadores no propensos a la acumulación de errores de redondeo tienen la ventaja de que no es necesario decidir si la trama actual debe codificarse como una trama P+ o una trama P-, con lo que el procesamiento se simplifica.

(vii) La presente invención descrita en esta memoria es aplicable a los procedimientos de codificación y decodificación que aplican filtrado junto con redondeo a las imágenes de predicción de intertrama. Por ejemplo, en el estándar internacional H.261 para la codificación de secuencias de imágenes se aplica un filtro pasa baja (denominado filtro de bucle) a las señales de bloque cuyos vectores de movimiento no son 0 en las imágenes de predicción de intertrama. Asimismo, en el estándar H.261, pueden utilizarse filtros para eliminar las discontinuidades en los límites de los bloques (artefacto de bloques). Todos estos filtros llevan a cabo la promediación ponderada de los valores de intensidad de los píxeles y, a continuación, se efectúa el redondeo de los valores de intensidad promediados. Aún en estos casos, la utilización selectiva del redondeo positivo y el redondeo negativo es eficaz para prevenir la acumulación de errores.

(viii) Aparte de IP+P-P+ P-..., existen diversos procedimientos para mezclar tramas P+ y P- (por ejemplo, IP+P+P-P+P+... o IP+P-P+P+...) que son aplicables al procedimiento de la presente invención. Por ejemplo, mediante un generador de números aleatorios que genera 0 y 1 a una probabilidad respectiva del 50%, el codificador puede codificar una trama P+ y una trama P- cuando la salida es 0 y 1, respectivamente. En cualquier caso, cuanto menor sea la diferencia en la probabilidad de que aparezcan tramas P+ y tramas P- durante cierto período de tiempo, menor será la probabilidad de que se produzca una acumulación de errores de redondeo. Además, cuando se permite que el codificador mezcle tramas P+ y tramas P- mediante un procedimiento arbitrario, el codificador y el decodificador deben funcionar basándose en el procedimiento explícito en vez del procedimiento implícito descrito anteriormente. En consecuencia, el procedimiento explícito es superior en la medida en que permite flexibilidad en la configuración del codificador y el decodificador.

(ix) La presente invención descrita en esta memoria no limita el procedimiento de interpolación de los valores de píxel a la interpolación bilineal. Los procedimientos de interpolación para valores de intensidad pueden describirse de forma general mediante la siguiente ecuación:

[Ecuación 5]

$$R(x+r, y+s) = T\left(\sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} h(r-j, s-k) R(x+j, y+k)\right) \quad \dots(5)$$

en la que r y s son números reales, h(r, s) es una función para interpolar los números reales y T(z) es una función para redondear el número real z. Las definiciones de R(x, y), x e y son las mismas que las de la Ecuación 4. La compensación de movimiento mediante redondeo positivo se lleva a cabo cuando T(z) es una función que representa redondeo positivo, y la compensación de movimiento mediante redondeo negativo se lleva a cabo cuando T(z) es una función que representa redondeo negativo. La presente invención es aplicable a los procedimientos de interpolación que pueden describirse mediante la Ecuación 5. Por ejemplo, la interpolación bilineal puede describirse definiendo h(r, s) como se muestra a continuación.

[Ecuación 6]

$$h(r, s) = (1-|r|)(1-|s|), \quad 0 \leq |r| \leq 1, 0 \leq |s| \leq 1, \quad \dots(6)$$

0, en los demás casos.

No obstante, por ejemplo, si h(r, s) se define como:

[Ecuación 7]

$$\begin{aligned} h(r, s) = & 1 - |r| - |s|, \quad 0 \leq |r| + |s| \leq 1, rs < 0, \\ & 1 - |r|, \quad |r| \geq |s|, |r| \leq 1, rs \geq 0, \\ & 1 - |s|, \quad |s| > |r|, |s| \leq 1, rs > 0, \end{aligned} \quad \dots(7)$$

0, en los demás casos

entonces, se ejecuta un procedimiento de interpolación diferente a la interpolación bilineal, pero la presente invención sigue siendo aplicable.

- 5
- (x) La presente invención descrita en esta memoria no se limita al procedimiento de codificación para imágenes de error a la DCT (transformada discreta del coseno). Por ejemplo, también pueden aplicarse a la presente invención la transformada de 'wavelets' (véase el documento «Image Coding Using Wavelet Transform» de M. Antonioni, *et.al.*, IEEE Trans. Image Processing, vol. 1, n.º 2, abril de 1992) y la transformada de Walsh-Hadamard (véase el documento «Digital Pictures» de A. N. Netravalli y B. G. Haskell, Plenum Press, 1998).
- 10

REIVINDICACIONES

1. Procedimiento de decodificación de imágenes que comprende las etapas siguientes:

5 recibir un tren de bits codificado que incluye información de tramas I y P;

determinar si una trama actual del tren de bits codificado es una trama I o una trama P de acuerdo con la información del tipo de trama incluida en el tren de bits codificado; y

10 ejecutar una compensación de movimiento mediante la sintetización de una imagen de predicción de la trama actual utilizando una información de vector de movimiento incluida en el tren de bits codificado y una imagen de referencia que es una imagen previamente decodificada, incluyendo la compensación de movimiento una interpolación de valores de píxel para calcular los valores de intensidad en los puntos en los que no existen píxeles en la imagen de referencia mediante interpolación bilineal;

15 caracterizado por que, únicamente cuando la trama actual es una trama P, comprende

determinar si el redondeo positivo o negativo debe ser utilizado como un procedimiento de redondeo para la interpolación de valores de píxel en la compensación de movimiento, de acuerdo con la información de procedimiento de redondeo incluida en el tren de bits codificado que incluye información relativa a la trama actual;

en el que la interpolación de valores de píxel que utiliza dicho redondeo positivo se calcula según las ecuaciones:

25
$$\begin{aligned} I_b &= [(L_a + L_b + 1)/2], \\ I_c &= [(L_a + L_c + 1)/2], \text{ y} \\ I_d &= [(L_a + L_b + L_c + L_d + 2)/4], \text{ y} \end{aligned}$$

la interpolación de valores de píxel que utiliza dicho redondeo negativo se calcula según las ecuaciones:

30
$$\begin{aligned} I_b &= [(L_a + L_b)/2], \\ I_c &= [(L_a + L_c)/2] \text{ y} \\ I_d &= [(L_a + L_b + L_c + L_d + 1)/4], \end{aligned}$$

35 en las que L_a , L_b , L_c y L_d son, respectivamente, los valores de intensidad de un primer píxel, un segundo píxel que es horizontalmente adyacente a dicho primer píxel, un tercer píxel que es verticalmente adyacente a dicho primer píxel y un cuarto píxel que es verticalmente adyacente a dicho segundo píxel y horizontalmente adyacente a dicho tercer píxel, y I_b , I_c e I_d son, respectivamente, los valores de intensidad interpolados del punto medio entre dicho primer y segundo píxel, el punto medio entre dicho primer y tercer píxel y el punto medio entre dicho primer, segundo, tercer y cuarto píxel.

40

FIG. 1

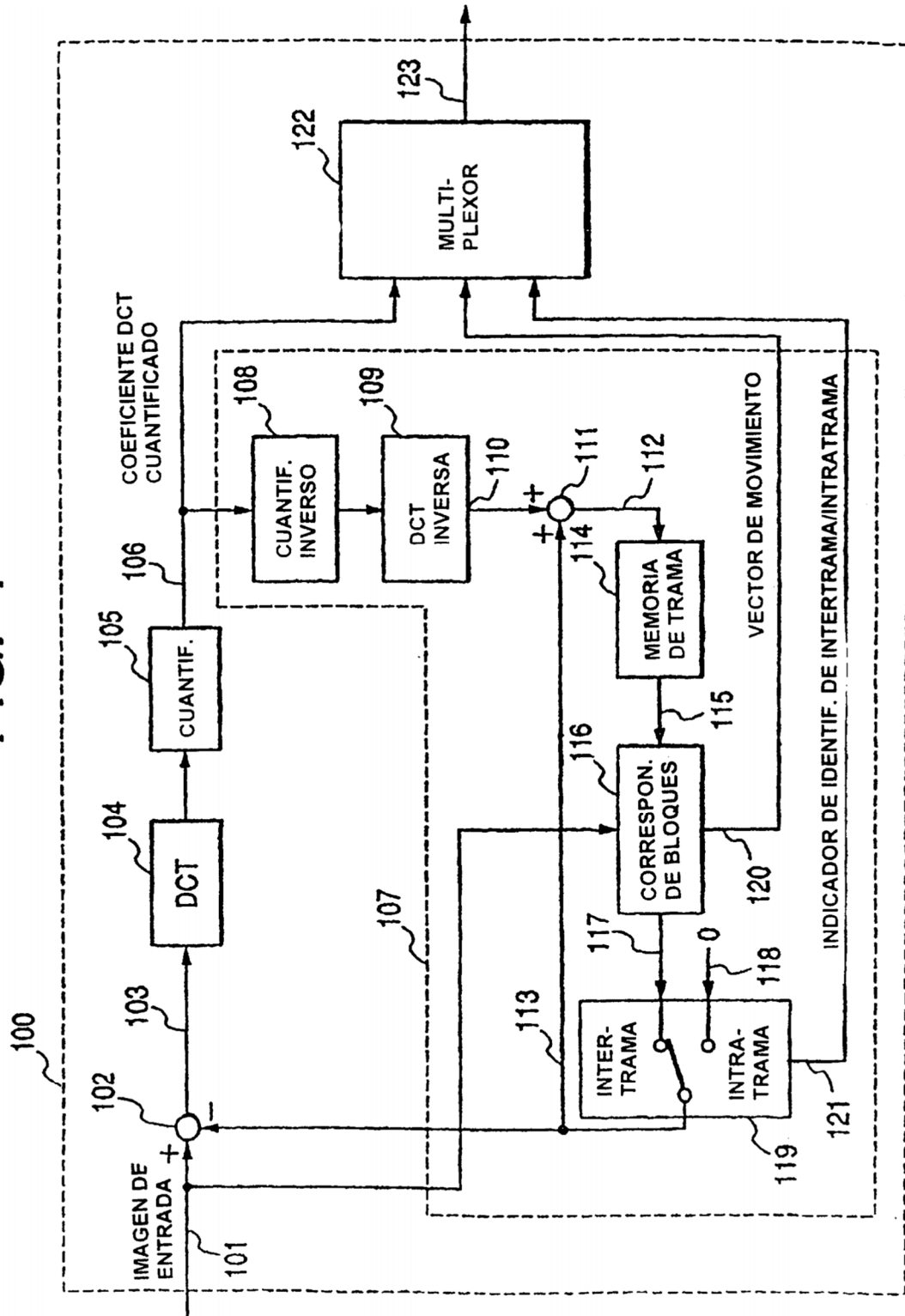


FIG. 2

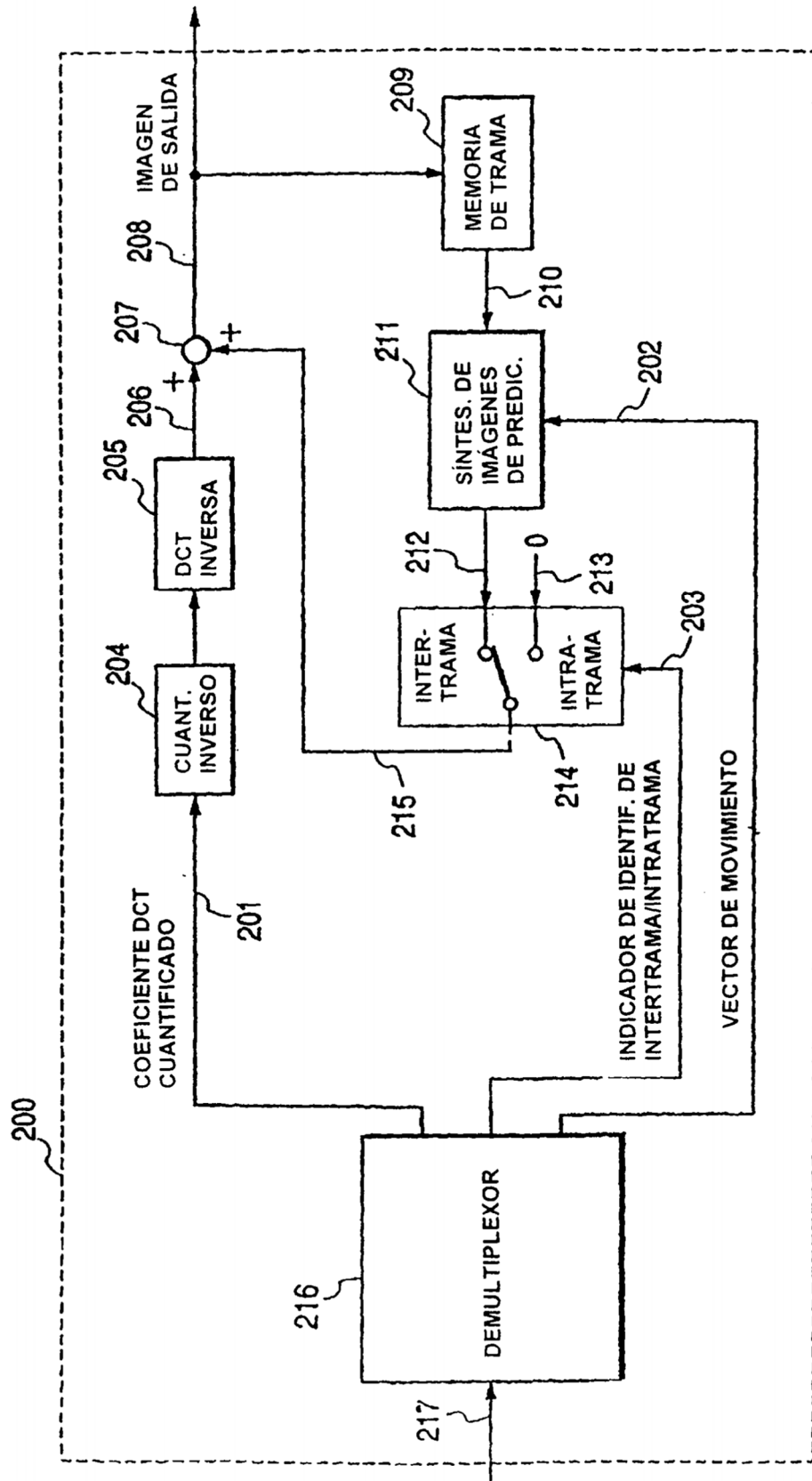


FIG. 3

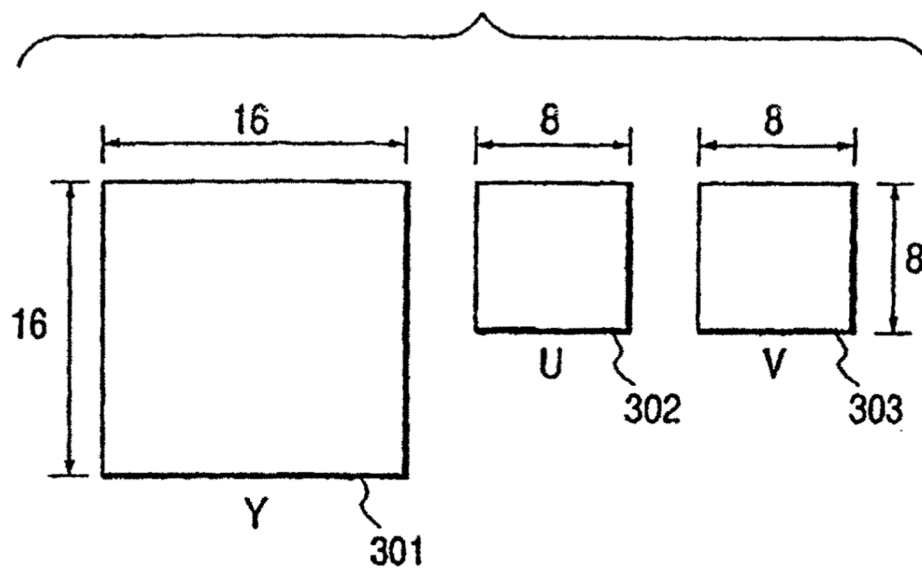


FIG. 4

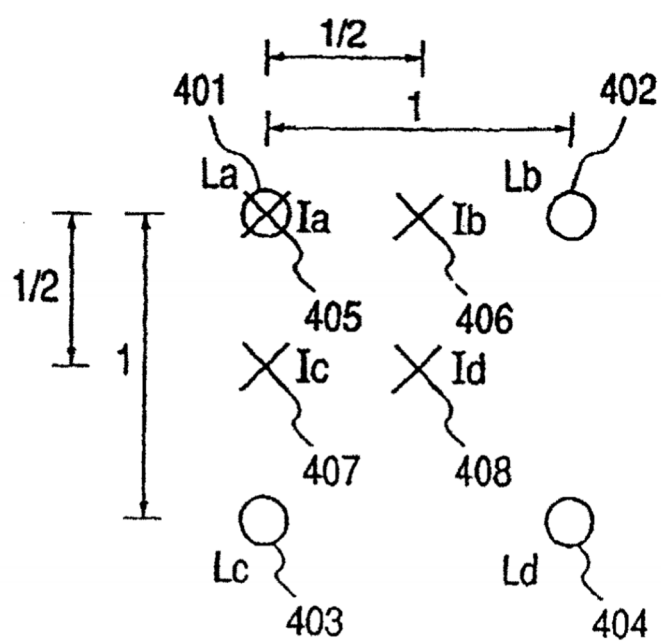


FIG. 5

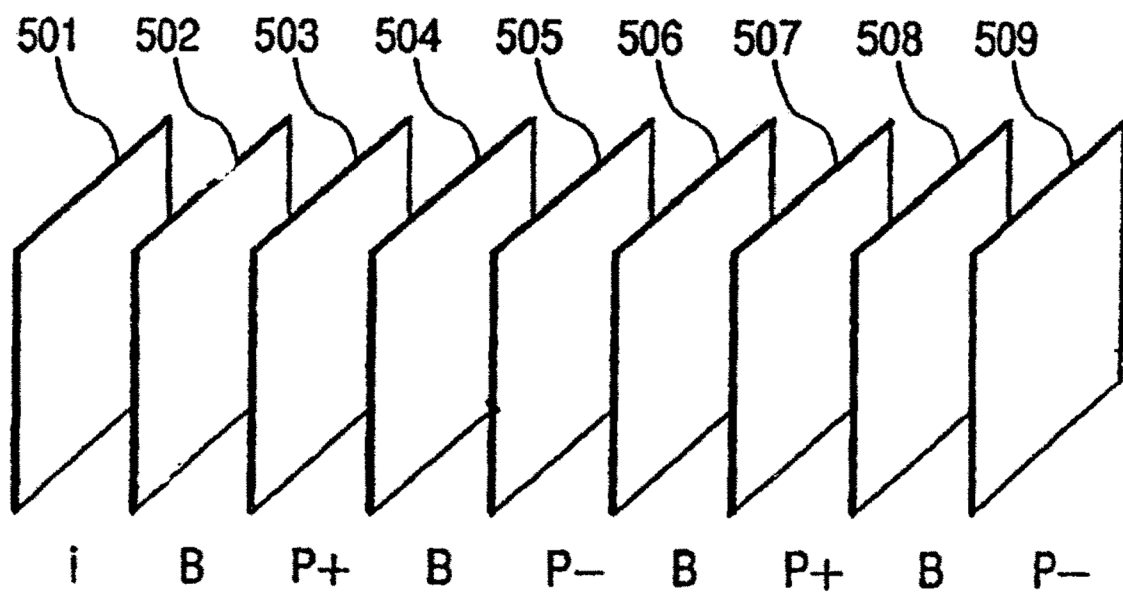


FIG. 6

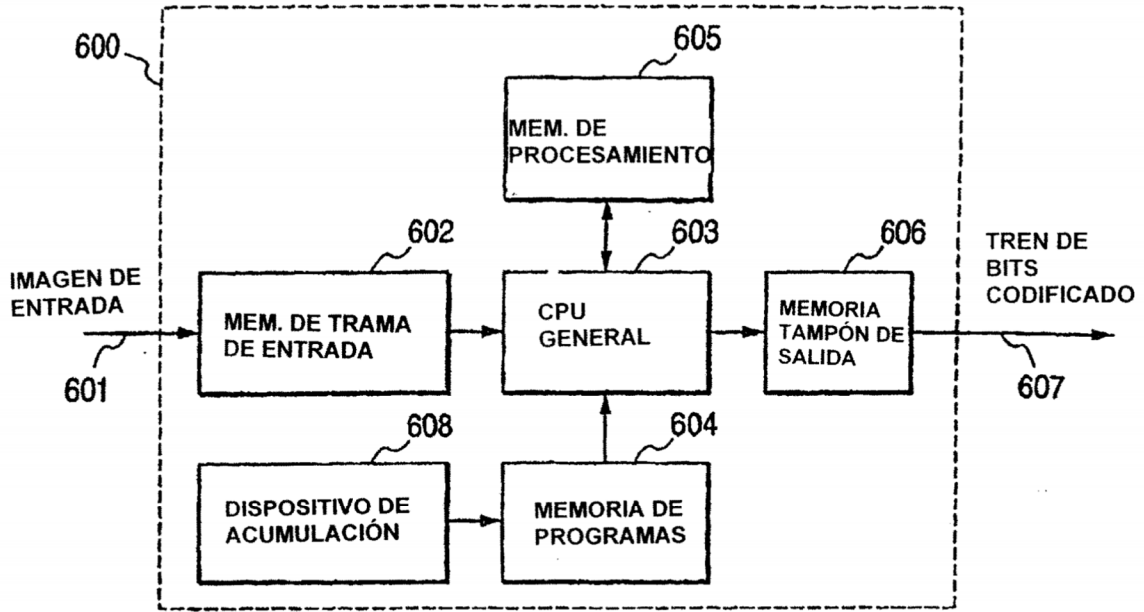


FIG. 7

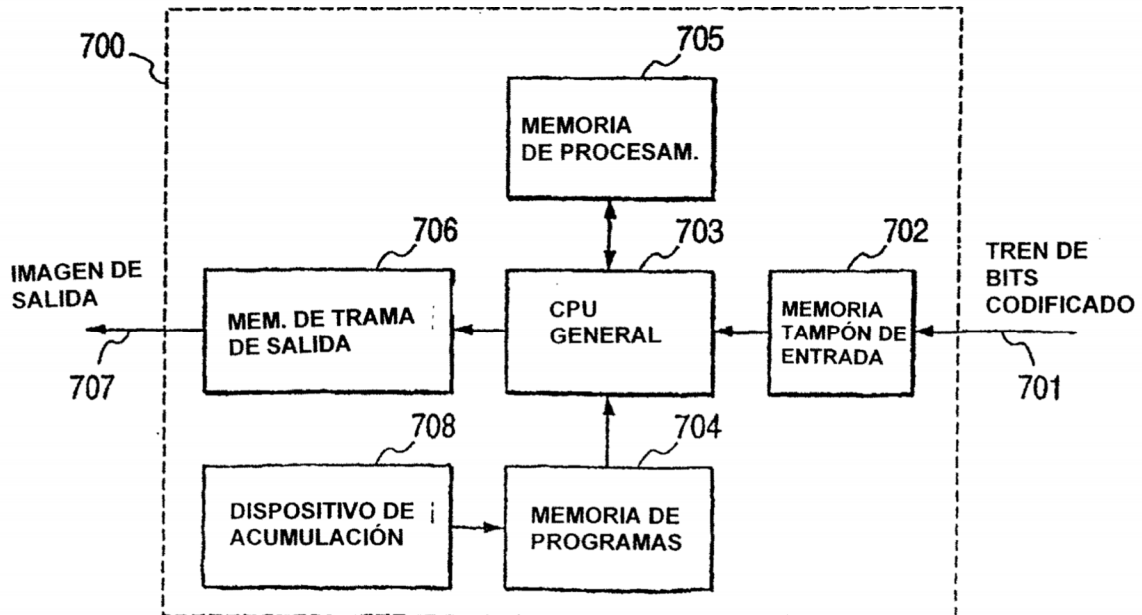


FIG. 8

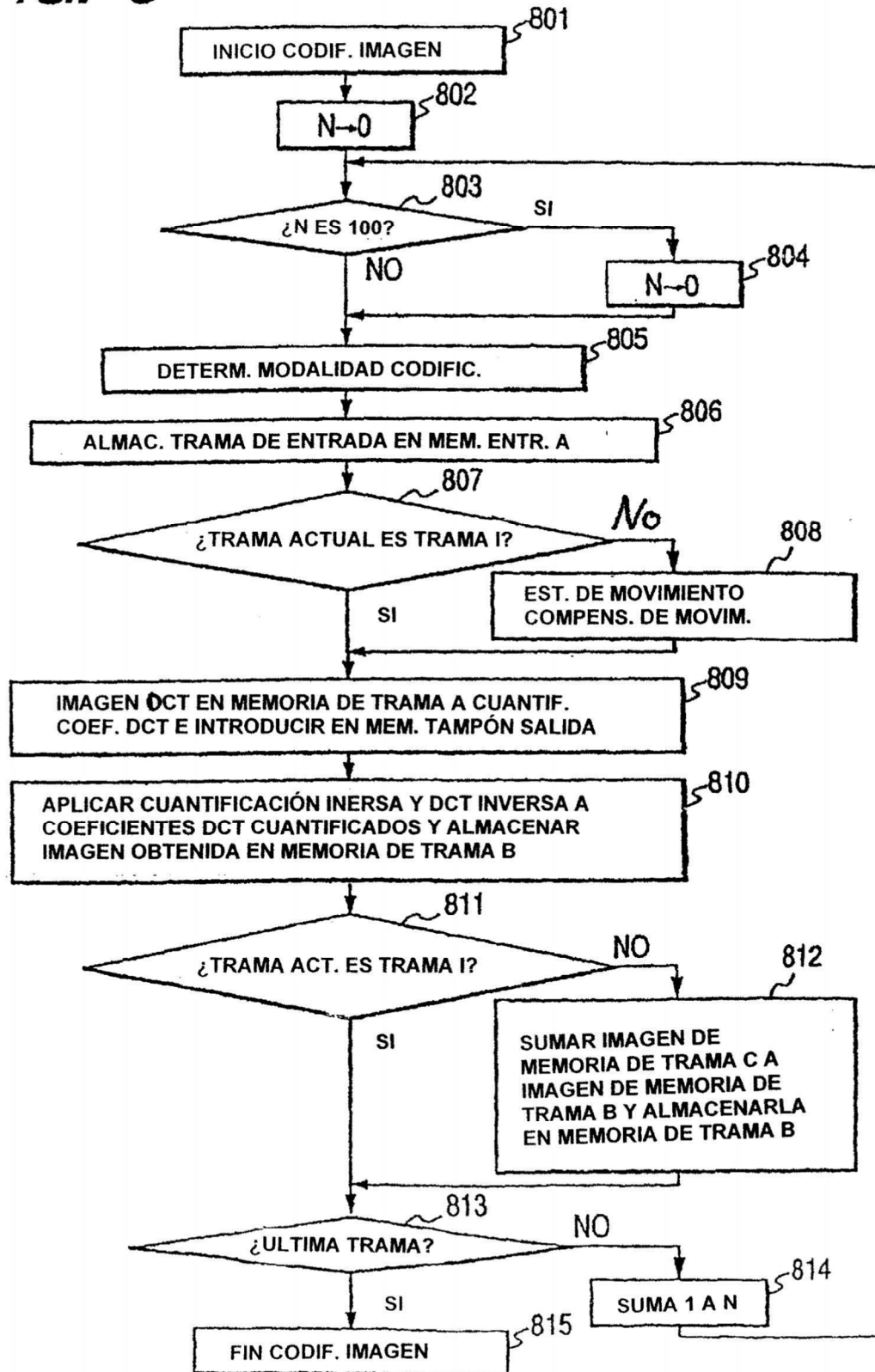


FIG. 9

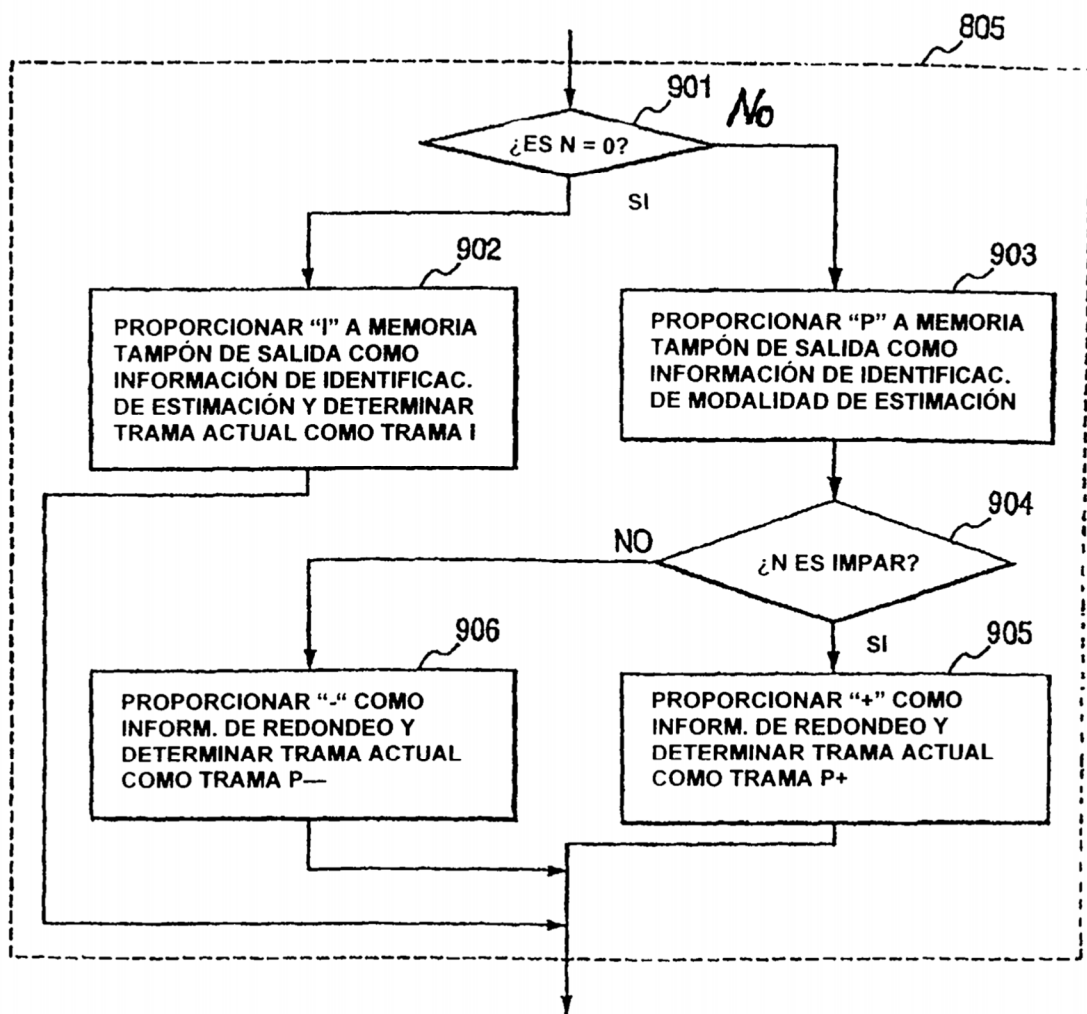


FIG. 10

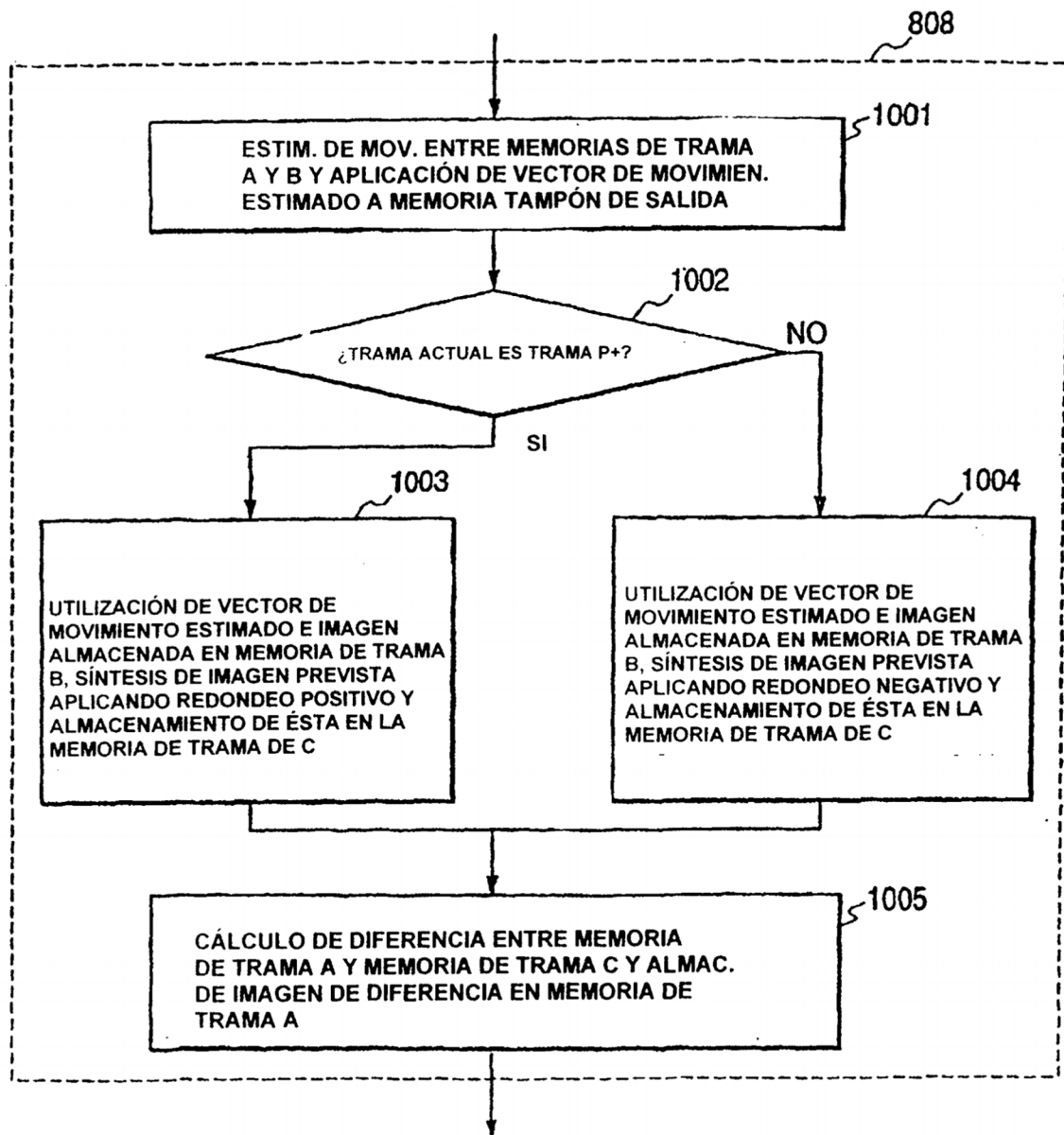


FIG. 11

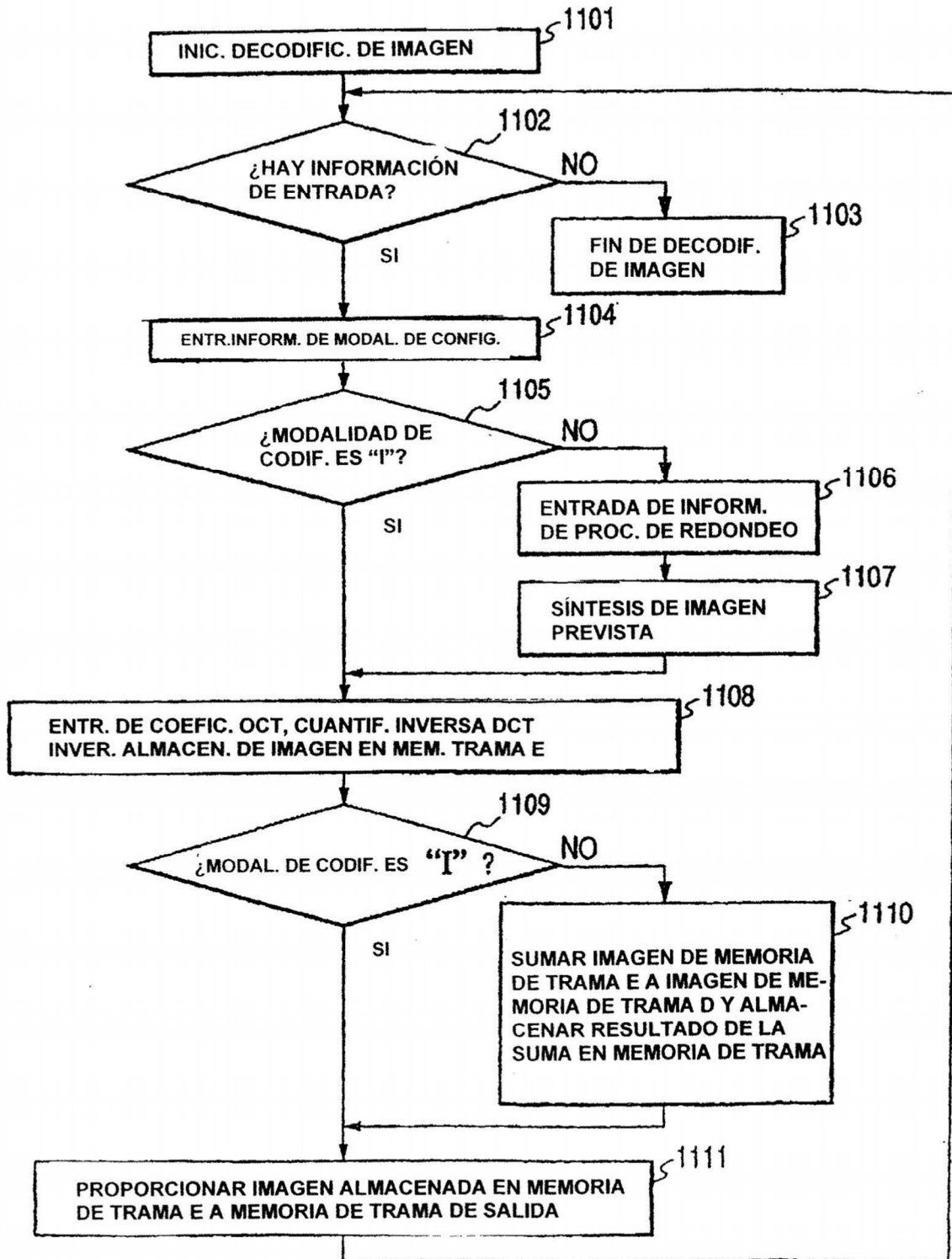


FIG. 12

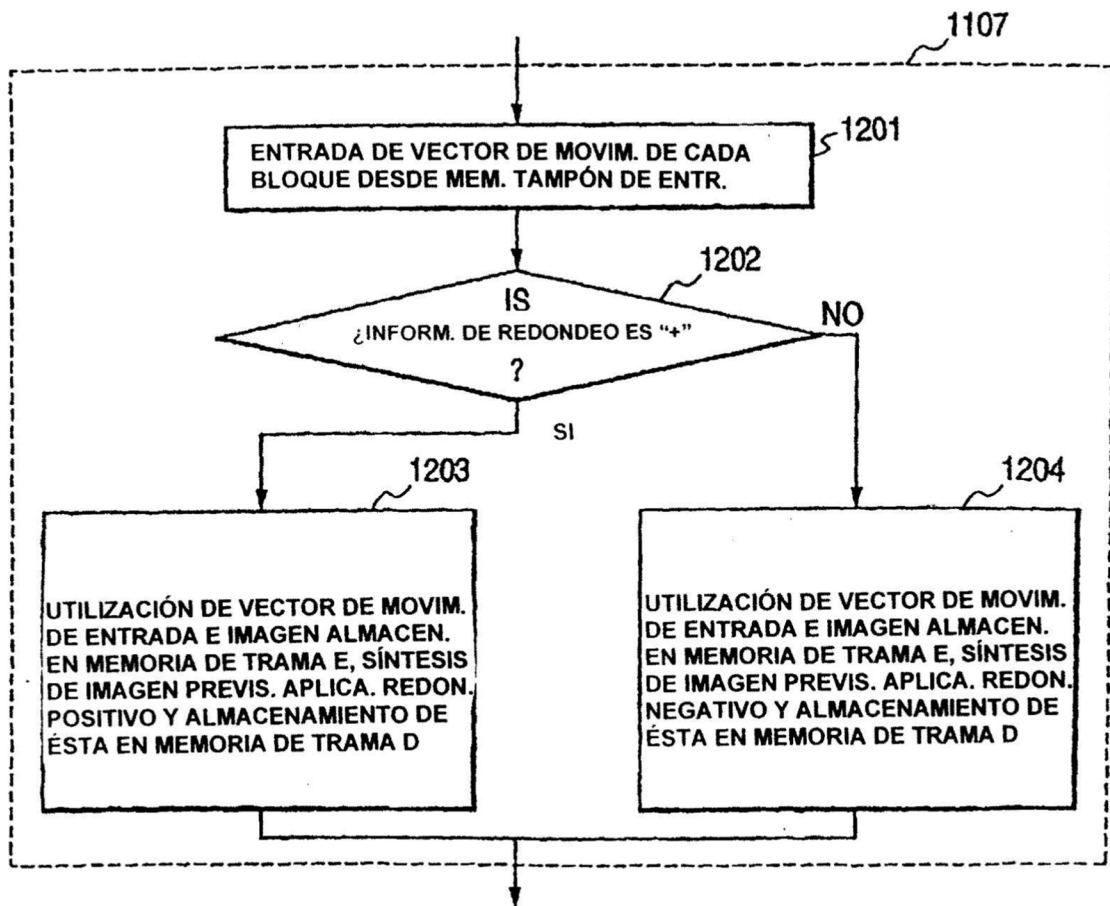


FIG. 13

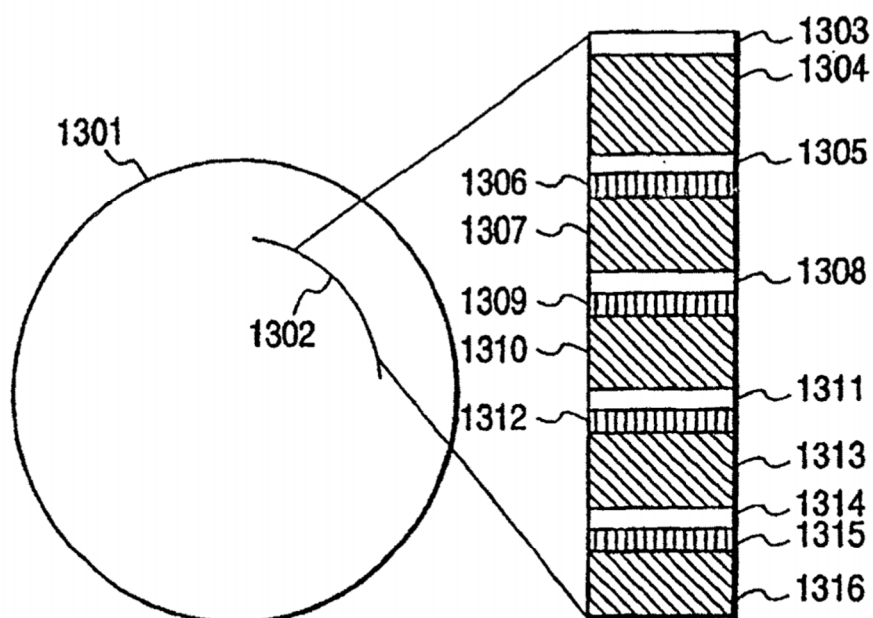


FIG. 15

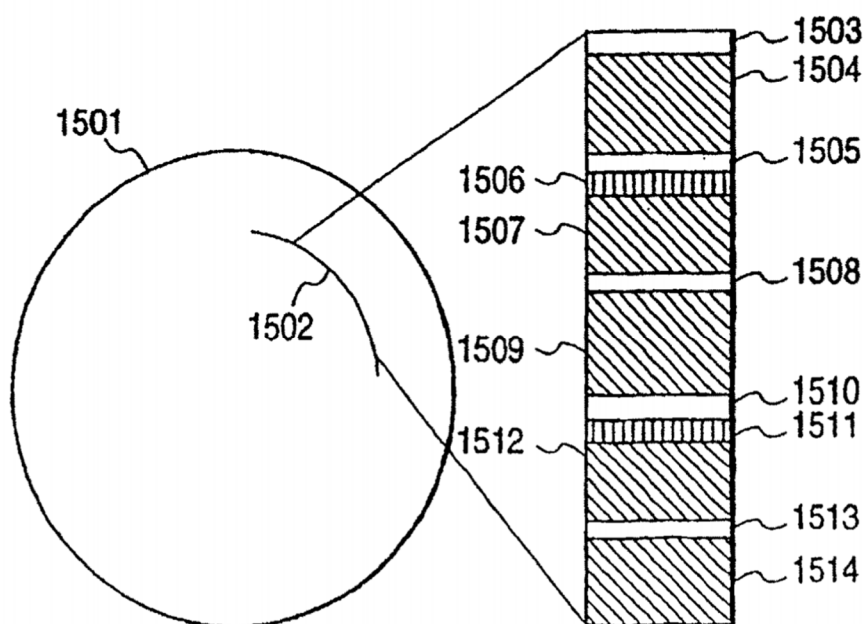


FIG. 14

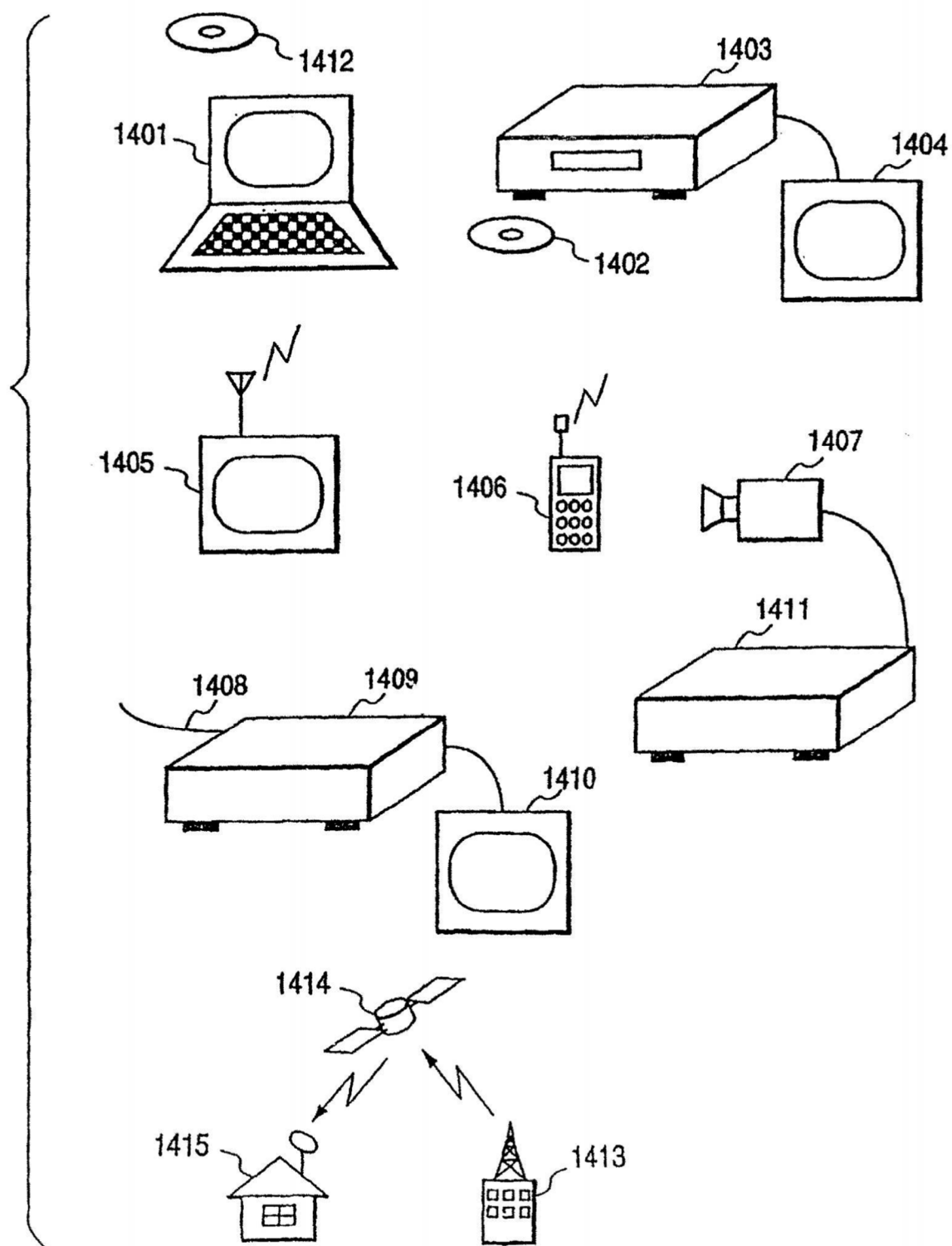


FIG. 16

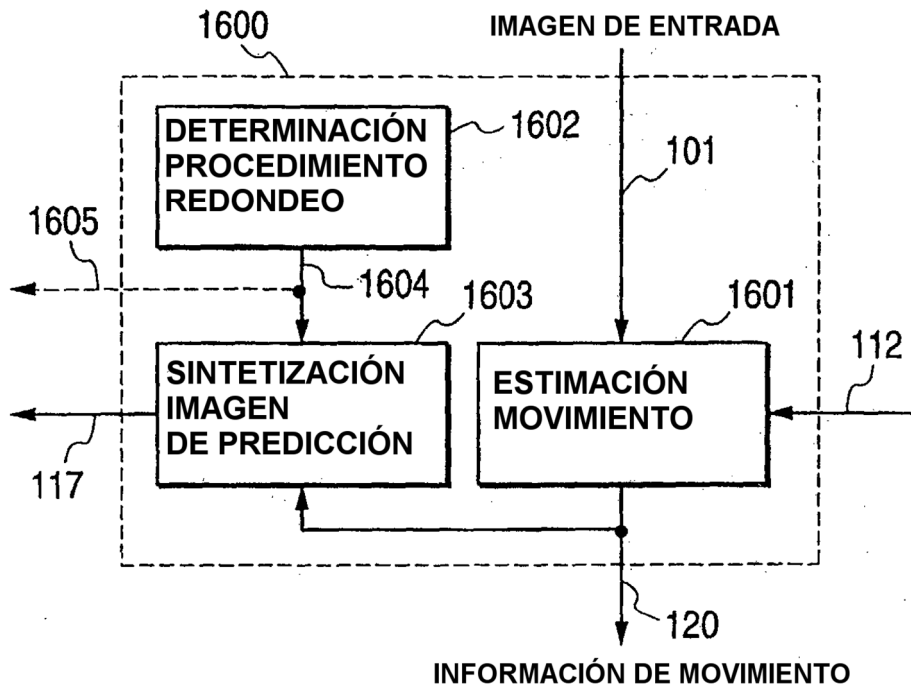


FIG. 17

