

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 388 080**

51 Int. Cl.:  
**H04N 7/32**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **07706352 .7**

96 Fecha de presentación: **04.01.2007**

97 Número de publicación de la solicitud: **1971155**

97 Fecha de publicación de la solicitud: **17.09.2008**

54 Título: **Método de codificación de vídeo y método de decodificación, dispositivo para los mismos, programas para los mismos y medios de almacenamiento para almacenar los programas**

30 Prioridad:  
**05.01.2006 JP 2006000393**

45 Fecha de publicación de la mención BOPI:  
**08.10.2012**

45 Fecha de la publicación del folleto de la patente:  
**08.10.2012**

73 Titular/es:  
**NIPPON TELEGRAPH AND TELEPHONE CORPORATION  
3-1, OTEMACHI 2-CHOME, CHIYODA-KU  
TOKYO 100-8116, JP y  
NATIONAL UNIVERSITY CORPORATION  
NAGOYA UNIVERSITY**

72 Inventor/es:  
**TANIMOTO, Masayuki;  
FUJII, Toshiaki;  
YAMAMOTO, Kenji;  
KITAHARA, Masaki;  
KIMATA, Hideaki;  
SHIMIZU, Shinya;  
KAMIKURA, Kazuto y  
YASHIMA, Yoshiyuki**

74 Agente/Representante:  
**Linage González, Rafael**

ES 2 388 080 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

Método de codificación de vídeo y método de decodificación, dispositivo para los mismos, programas para los mismos y medios de almacenamiento para almacenar los programas

5 **Campo técnico**  
La presente invención está relacionada con técnicas de codificación y decodificación de imágenes de vídeo de puntos de vista múltiples.

10 **Técnica anterior**  
Las imágenes de vídeo de puntos de vista múltiples son imágenes de vídeo obtenidas fotografiando el mismo objeto y fondo del mismo, utilizando una pluralidad de cámaras en posiciones diferentes. En lo que sigue, una imagen de vídeo obtenida por una sola cámara es denominada "imagen de vídeo bidimensional", y un conjunto de imágenes de vídeo bidimensionales, obtenidas fotografiando el mismo objeto y fondo del mismo, es denominada "imagen de vídeo de puntos de vista múltiples". Hay una fuerte correlación entre las imágenes de vídeo bidimensionales (de diferentes cámaras) incluidas en la imagen de vídeo de puntos de vista múltiples. Si las cámaras están sincronizadas entre sí, los fotogramas (de las cámaras) correspondientes al mismo momento han capturado el objeto y el fondo del mismo enteramente en el mismo estado, de manera que hay una fuerte correlación entre las cámaras.

25 En primer lugar, se ilustrarán las técnicas convencionales relativas a la codificación de imágenes de vídeo bidimensionales. En muchos métodos conocidos de codificación de imágenes de vídeo bidimensionales, tal como el H.264, MPEG-4, MPEG-2 (que son estándares internacionales de codificación), y similares, se efectúa una codificación altamente eficiente por medio de la compensación del movimiento, la transformación ortogonal, la cuantificación, la codificación de longitud variable o similares.

30 Por ejemplo, en el H.264, cada fotograma I puede ser codificado por medio de una correlación dentro de los fotogramas (intrafotograma); cada fotograma P puede ser codificado por medio de una correlación entre fotogramas (interfotograma), junto con una pluralidad de fotogramas transcurridos; y cada fotograma B puede ser codificado por medio de una correlación interfotograma junto con una pluralidad de fotogramas transcurridos o futuros.

35 Aunque el documento no-patente 1 divulga las técnicas H.264 en detalle, a continuación se describirá el perfil de las mismas. En cada fotograma I, el fotograma se divide en bloques (denominados "macrobloques", el tamaño de cada bloque es 16 x 16 (píxeles)), y la predicción intrafotograma (intra-predicción) se efectúa en cada macrobloque. En la intra-predicción, cada macrobloque se divide además en bloques más pequeños (denominados a continuación "sub-bloques"), y se puede aplicar un método individual de intra-codificación a cada sub-bloque.

40 En cada fotograma P, la intra-predicción o la inter-predicción (predicción interfotograma) puede ser realizada en cada macrobloque. La intra-predicción aplicada a un fotograma P es similar a la aplicada a un fotograma I. En la inter-predicción, se realiza la compensación del movimiento. También en la compensación del movimiento, cada macrobloque se divide en bloques más pequeños, y cada sub-bloque puede tener un vector de movimiento individual y una imagen de referencia individual.

45 También en cada fotograma B se puede efectuar la intra-predicción o la inter-predicción. En la inter-predicción de un fotograma B, además de un fotograma transcurrido, se puede hacer referencia a un fotograma futuro como imagen de referencia en la compensación del movimiento. Por ejemplo, cuando se codifica una secuencia de fotogramas de "I→B→B→P", los fotogramas pueden ser codificados en el orden de "I→P→B→B". Además, en cada fotograma B, se puede realizar la compensación del movimiento haciendo referencia a un fotograma I o P. Adicionalmente, de forma similar al fotograma P, cada sub-bloque (obtenido dividiendo un macrobloque) puede tener un vector de movimiento individual.

50 Cuando se realiza una intra o inter-predicción, se obtiene una predicción residual. En cada macrobloque, un bloque de predicción residual está sometido a una DCT (transformada discreta del coseno), para realizar la cuantificación. Los valores cuantificados obtenidos de los coeficientes DCT son sometidos después a una codificación de longitud variable.

60 En un método de codificación conocido para imágenes de vídeo de puntos de vista múltiples, las imágenes de vídeo de puntos de vista múltiples son codificadas con una alta eficiencia por medio de la "compensación del paralaje" en la cual se aplica la compensación del movimiento a imágenes obtenidas por cámaras diferentes al mismo tiempo. En este caso, el "paralaje" es la diferencia entre posiciones, a las cuales se proyecta el mismo punto en un objeto, sobre un plano de imágenes de cámaras que están dispuestas en posiciones diferentes.

65 La figura 9 es una vista esquemática que muestra el concepto del paralaje generado entre tales cámaras. En la vista esquemática, un plano de imágenes de cámaras, cuyos ejes ópticos son paralelos entre sí, está visto hacia abajo

verticalmente. Generalmente, tales puntos, a los cuales se proyecta el mismo punto en un objeto, en un plano de imágenes de cámaras diferentes, son denominados “puntos correspondientes”. Como el paralaje puede ser representado como una diferencia posicional en el plano de imagen relevante, puede ser representado como datos de un vector bidimensional.

5 En la compensación del paralaje, el punto correspondiente de la imagen de una cámara de referencia, que se corresponde con un pixel objetivo en una imagen de cámara objetivo para la codificación relevante, se estima utilizando una imagen de referencia, y el valor del pixel de referencia se predice utilizando un valor de pixel asignado al punto correspondiente. A continuación, tal “paralaje estimado” es denominado también “paralaje” por conveniencia de la explicación.

10 El documento no-patente 2 divulga un método de codificación utilizando una compensación del paralaje y, en tal método, los datos de paralaje y cada predicción residual son codificados con respecto a los pixeles de una imagen objetivo a codificar. Más específicamente, en el método relevante, la compensación del paralaje se efectúa para cada bloque como una unidad, donde tal paralaje para cada bloque unitario se representa utilizando un vector bidimensional. La figura 10 es una vista esquemática que ilustra un vector de paralaje. Esto es, en este método, se codifican los datos de paralaje como vector bidimensional y la predicción residual relevante. Como este método no utiliza parámetros de la cámara en la codificación, es eficaz cuando los parámetros de la cámara son desconocidos.

15 Además, cuando hay una pluralidad de imágenes de referencia obtenidas por cámaras diferentes, la compensación del paralaje puede efectuarse utilizando una técnica de imágenes de puntos de vista arbitrarios. El documento no-patente 3 divulga la compensación del paralaje utilizando una técnica de imágenes de puntos de vista arbitrarios. Más específicamente, cada valor de pixel de una imagen obtenida por una cámara objetivo para la codificación relevante, se predice por medio de la interpolación utilizando los valores de pixel de los puntos correspondientes (pertenecientes a cámaras diferentes) que se corresponden con el pixel relevante. La figura 11 es una vista esquemática que muestra tal interpolación. En la interpolación, el valor del pixel  $m$  en una imagen objetivo a codificar se predice efectuando la interpolación entre los pixeles  $m'$  y  $m''$  de las imágenes de referencia 1 y 2, donde los pixeles  $m'$  y  $m''$  se corresponden con el pixel  $m$ .

20 Cuando hay dos o más imágenes de referencia obtenidas por cámaras diferentes (como se divulga en el documento de patente 3), el paralaje de cada pixel de una imagen objetivo (a codificar) para cada imagen de referencia, puede ser estimado sin utilizar la imagen objetivo. La figura 12 es una vista esquemática para ilustrar el concepto de tal estimación del paralaje.

25 Como se ilustra en la figura 12, en el verdadero paralaje, los valores de pixel de los puntos correspondientes entre las imágenes de referencia relevantes deben ser casi idénticos entre sí. Por tanto, en muchos métodos de estimación del paralaje, con respecto a cada una de las diversas profundidades, se comparan entre sí los valores de pixel de los puntos correspondientes entre las imágenes de referencia, y el paralaje se puede estimar basándose en la profundidad que contiene los valores de pixel más cercanos. Tal proceso puede ser aplicado a cada pixel de una imagen objetivo a codificar.

30 Como se ha descrito anteriormente, cuando hay dos o más imágenes de referencia obtenidas por cámaras diferentes, y es posible la estimación del paralaje en el lado de la descodificación, se puede efectuar entonces la compensación del paralaje en el lado de la descodificación utilizando los datos del paralaje para cada pixel, sin proporcionar datos del paralaje, que se codifican explícitamente en el lado de la codificación, al lado de la descodificación.

35 Documento no-patente 1: ITU-T Rec.H.264/ISO/IEC 11496-10, “Codificación avanzada de vídeo”, Borrador Final del Comité, Documento JVT-E022, Septiembre de 2002.

40 Documento no-patente 2: Hideaki Kimata y Masaki Kitahara, “Resultados preliminares sobre la codificación de vídeo de vistas múltiples (3DAV)”, documento M10976, Reunión de Redmont del MPEG, Julio de 2004.

45 Documento no-patente 3: Masayuki Tanimoto, Toshiaki Fujii, “Respuesta a la Llamada a la Evidencia en la Codificación de Vídeo de Vistas Múltiples”, documento Mxxxxx, Reunión de Hong Kong del MPEG, Enero de 2005.

50 En las técnicas convencionales, cuando hay dos o más imágenes de referencia obtenidas por cámaras diferentes, y es posible la estimación del paralaje en el lado de la descodificación, la compensación del paralaje puede efectuarse entonces en el lado de la descodificación utilizando datos del paralaje para cada pixel, sin proporcionar datos de paralaje, que son codificados explícitamente en el lado de la codificación, al lado de la descodificación. Tal paralaje con respecto a una imagen objetivo a codificar o descodificar, que puede ser estimado en el lado de la codificación o descodificación sin utilizar la imagen objetivo (para la descodificación, sin descodificar la imagen relevante), es denominado “paralaje de referencia”.

55 Sin embargo, el paralaje de referencia, que es estimado en el lado de la descodificación, no es un paralaje óptimo considerando la eficiencia de la predicción, y la cantidad de código asignada a la predicción residual puede ser

incrementada. Cuando se calcula el paralaje para maximizar la eficiencia de la predicción en el lado de la codificación, y se codifica la diferencia (denominada en lo que sigue “desplazamiento del paralaje”) entre el paralaje calculado y el paralaje de referencia para cada pixel, la eficiencia de la predicción puede ser mejorada, mejorando con ello la eficiencia de la codificación con respecto a la predicción residual.

5 Sin embargo, en tal técnica que puede ser fácilmente anticipada, como se codifica el desplazamiento del paralaje para cada pixel, la cantidad de código de datos de paralaje aumenta, de manera que la eficiencia total de la codificación no puede ser alta.

## 10 **Divulgación de la invención**

Con el fin de resolver el problema anterior, un objeto de la presente invención es disminuir la cantidad de código de datos de paralaje sin hacer un sacrificio considerable de la precisión de compensación del paralaje.

15 La característica más distintiva de la presente invención en comparación con las técnicas convencionales, es que con el fin de resolver los problemas anteriormente descritos, se efectúa un proceso de determinación para cada bloque de la división de bloques, y un proceso de codificación de los datos de la división de bloques determinada y de los datos de desplazamiento del paralaje, y tales datos son manejados como datos codificados.

20 En un primer modo de los métodos de codificación y decodificación de vídeo de la presente invención, basándose en la división del área, que se fija de acuerdo con la característica de una imagen objetivo a codificar, se pueden codificar los datos de desplazamiento del paralaje para cada área dividida.

25 Generalmente, el desplazamiento del paralaje tiene una correlación espacial en un fotograma de imagen. Por tanto, cuando se fija la división de área apropiada, y se codifica el desplazamiento del paralaje para cada área dividida, la cantidad de código de datos del paralaje puede reducirse entonces sin degradar la eficiencia de la predicción de la compensación del paralaje.

30 En el lado de la codificación del vídeo, en un paso de ajuste del paralaje de referencia, se fija el paralaje para una imagen objetivo a codificar (es decir, el paralaje de referencia) utilizando los datos ya codificados (es decir, la imagen de referencia), y en un paso de ajuste de la división del área, se fija la división del área en un fotograma de la imagen objetivo. En un paso de ajuste del desplazamiento del paralaje, se fija el desplazamiento del paralaje para cada área dividida que se ha fijado en la división del área. Adicionalmente, en el paso de codificación de datos de la división de área, se codifican los datos de la división de área para indicar la división de área, y en un paso de  
35 codificación de los datos de desplazamiento del paralaje, se codifican los datos de desplazamiento del paralaje.

40 Como el paralaje de referencia fijado en el paso de ajuste del paralaje de referencia, (i) el paralaje estimado utilizando imágenes de referencia basándose en el principio ilustrado en la figura 12 (se hace referencia a un tercer modo que se explicará más adelante); (ii) el paralaje basado en una imagen de paralaje o un modelo tridimensional, que es codificado por otro dispositivo, por ejemplo para generar una imagen de punto de vista arbitrario, y es enviado al lado de la decodificación; o (iii) el paralaje proporcionado por otro método.

45 En el lado de decodificación, en un paso de ajuste del paralaje de referencia, similar al lado de la codificación, se fija el paralaje de referencia para una imagen objetivo a decodificar, utilizando datos ya decodificados (es decir, la imagen de referencia). En un paso de decodificación de los datos de división del área, se decodifican los datos de división del área, y en un paso de decodificación de los datos del desplazamiento del paralaje, se decodifican los datos del desplazamiento del paralaje para cada área, de acuerdo con los datos de división del área.

50 Un segundo modo de los métodos de codificación y decodificación de la presente invención es básicamente similar al primer modo anteriormente descrito. Sin embargo, en el segundo modo, se efectúa la división del área para generar bloques rectangulares. Más específicamente, se define la división de bloques rectangulares (se hace referencia a la figura 10 que se explica más adelante) para cada macrobloque. Los datos para tal división de bloques pueden ser codificados eficientemente utilizando una técnica conocida de codificación de entropía (por ejemplo, codificación de entropía en H.264).

55 Un tercer modo de los métodos de codificación y decodificación de la presente invención es básicamente similar a los primero y segundo modos anteriormente descritos. Sin embargo, en el tercer modo, se utiliza el paralaje de referencia, que es estimado por una pluralidad de imágenes de referencia sin utilizar una imagen objetivo a codificar (es decir, de acuerdo con el principio anteriormente descrito como se ilustra en la figura 12). Cuando se estima el paralaje de referencia utilizando imágenes de referencia como se ha explicado anteriormente, puede ser estimado utilizando solamente datos que han sido ya enviados al lado de la decodificación. Por tanto, es posible no generar datos adicionales a codificar, para estimar el paralaje de referencia.

60 De acuerdo con la presente invención, en los métodos de codificación y decodificación de vídeo para estimar el paralaje utilizando imágenes de referencia, se utiliza la correlación espacial de los datos de desplazamiento del paralaje, de manera que la cantidad de código de los datos del desplazamiento del paralaje puede reducirse sin

degradar la eficiencia de la predicción en la compensación del paralaje, mejorando con ello la eficiencia total de la codificación.

**Breve descripción de los dibujos**

5 La figura 1 es un diagrama que muestra un dispositivo de codificación de vídeo como modo de realización de la presente invención.

La figura 2 es un diagrama que ilustra relaciones de referencia entre cámaras en el modo de realización.

10 La figura 3 es un diagrama que ilustra la disposición de las cámaras en el modo de realización.

La figura 4 es un diagrama de flujo de la codificación en el modo de realización.

15 La figura 5 es un diagrama de flujo detallado del paso S106 de la figura 4.

La figura 6 es un diagrama que muestra ejemplos de la división de bloques en cada macrobloque.

20 La figura 7 es un diagrama que muestra un dispositivo de descodificación de vídeo en el modo de realización.

La figura 8 es un diagrama de flujo de la descodificación en el modo de realización.

La figura 9 es una vista esquemática que muestra el concepto de paralaje generado entre cámaras.

25 La figura 10 es una vista esquemática que muestra un vector de paralaje.

La figura 11 es una vista esquemática que muestra la interpolación de valores de pixel.

30 La figura 12 es una vista esquemática que muestra el concepto de estimación del paralaje.

**Mejor modo de llevar a cabo la invención**

35 La figura 1 es un diagrama que muestra la estructura de un dispositivo de codificación de vídeo, como un modo de realización de la presente invención.

El dispositivo 100 de codificación de vídeo incluye una parte 101 de entrada de imágenes en la cual se introduce cada imagen original de la cámara C (es decir, la imagen objetivo a codificar); una parte 102 de entrada de la imagen de referencia en la cual se introducen imágenes descodificadas (como imágenes de referencia) de las cámaras A y B; una memoria 103 de imágenes de referencia para almacenar cada imagen de referencia; una parte 40 104 de ajuste del paralaje de referencia para obtener el paralaje de referencia utilizando imágenes de referencia; una parte 105 del ajuste del desplazamiento del paralaje para obtener el desplazamiento del paralaje; una parte 106 de ajuste de la división de bloques para ajustar el estado de la división del bloques; una parte 107 de codificación de los datos de la división de bloques para codificar los datos de la división de bloques; una parte 108 de codificación de los datos del desplazamiento del paralaje para codificar los datos del desplazamiento del paralaje; y una parte 45 109 de codificación de la predicción residual para codificar la predicción residual relevante.

50 La figura 2 es un diagrama que muestra las relaciones de referencia entre cámaras en el presente modo de realización. Como se ilustra en la figura 2, en el presente modo de realización para codificar las imágenes de vídeo de puntos de vista múltiples, obtenidas por tres cámaras, las imágenes de vídeo de la cámara C son codificadas utilizando imágenes descodificadas de las cámaras A y B, como imágenes de referencia.

55 Las flechas de la figura 2 representan relaciones de referencia en la compensación del paralaje. Con el fin de codificar cada imagen de la cámara C, las imágenes descodificadas (de las cámaras A y B) que tienen el mismo tiempo de presentación, son codificadas como imágenes de referencia. En el proceso relevante, se genera una imagen de predicción calculando el valor medio entre los valores de pixel de los correspondientes puntos que pertenecen a las cámaras A y B (cada punto correspondiente es un pixel indicado por un vector que es proporcionado por la suma del vector de desplazamiento del paralaje relevante y del vector del paralaje de referencia).

60 La figura 3 es un diagrama que muestra la disposición de las cámaras en el presente modo de realización. En el presente modo de realización, como se ilustra en la figura 3, los puntos de vista de las tres cámaras se alinean a lo largo de una línea recta a intervalos regulares, y los ejes ópticos de las cámaras son perpendiculares a la línea recta. Es decir, los ejes ópticos de las tres cámaras son paralelos entre sí.

65 Además, el sistema de coordenadas x - y en el plano de imágenes relevantes puede ser obtenido mediante una traslación paralela (no se efectúa rotación o similar) con respecto a la línea recta sobre la cual se disponen las

cámaras, y se forman los píxeles dividiendo cada uno de los ejes x e y del plano de la imagen a intervalos regulares para cada cámara. Es decir, cada cámara tiene la misma resolución, y un paralaje de P píxeles entre las cámaras C y A se corresponde con un paralaje de P píxeles entre las cámaras C y B.

5 La figura 4 es un diagrama de flujo de la codificación del presente modo de realización. La figura 5 es un diagrama de flujo detallado del paso S106 de la figura 4.

En el presente modo de realización, se utiliza como unidad un macrobloque consistente en 16 x 16 píxeles (en las direcciones vertical y horizontal), y se designa la división de bloques (estado) de cada macrobloque. Se obtienen los  
10 datos del desplazamiento del paralaje para cada bloque definido en la división de bloques (denominado simplemente "bloque" por conveniencia) y se codifican.

El paralaje de referencia se representa utilizando un vector bidimensional. Por tanto, en el lado de descodificación (y en el lado de codificación), se obtiene un vector bidimensional (es decir, el paralaje de referencia) para cada imagen de referencia y para cada píxel, utilizando la imagen de referencia.  
15

Por otra parte, se codifica para cada bloque un solo desplazamiento del paralaje (para la cámara A) representado por un vector bidimensional. En este caso, de forma similar al cálculo del paralaje de referencia, se supone una condición de restricción física en la cual cada vector (de paralaje) (utilizado para la compensación del paralaje), que es proporcionada como la suma del vector del paralaje de referencia y un vector del paralaje de referencia para cada  
20 píxel de cada una de las imágenes de referencia, designa la misma posición en el objeto. En este caso, si se conoce el vector del paralaje de referencia para la cámara A, el vector del desplazamiento del paralaje para otra cámara se determina unívocamente.

25 En lugar de eso, el desplazamiento del paralaje para cada imagen de referencia puede ser calculado independientemente y codificado sin suponer tal condición de restricción. Sin embargo, tal método puede ser fácilmente implementado como una variación del presente modo de realización, y se omiten explicaciones específicas del mismo.

30 Hay muchos estados posibles de división de bloques aplicables a cada macrobloque. La figura 6 muestra ejemplos de los mismos. Como se ilustra en la figura 6, "blkMode" es un índice para indicar el tipo de división de bloques, y el número de bloques para cada "blkMode" está indicado como "maxBlk[blkMode]".

Bajo las condiciones anteriores, la operación de codificación será explicada con referencia al diagrama de flujo de la  
35 figura 4.

En primer lugar, se introduce una imagen de la cámara C en la parte 101 de entrada de imágenes (véase el paso S101) donde las imágenes descodificadas, que pertenecen a las cámaras A y B y tienen el mismo tiempo de presentación, han sido almacenadas en la memoria 103 de imágenes de referencia a través de la parte 102 de  
40 entrada de imágenes de referencia.

A continuación, se introducen las imágenes relevantes descodificadas de las cámaras A y B desde la memoria 103 de imágenes de referencia (véase el paso S102), y se obtiene el paralaje de referencia para cada píxel de una imagen objetivo a codificar, para cada imagen de referencia de entrada utilizando las imágenes de referencia (véase  
45 el paso S103). Es decir, dos imágenes de referencia, que pertenecen a las cámaras A y B, son introducidas en la parte 104 de ajuste del paralaje de referencia, y se obtiene el paralaje de referencia de cada píxel de la imagen objetivo de la cámara C.

Consecuentemente, se obtienen dos "vectores bidimensionales" para cada píxel de la imagen objetivo de la cámara C. Para las coordenadas (x, y) en el plano de la imagen de la cámara C (x e y son cada una de ellas un entero que indica un valor de coordenada del píxel relevante), los vectores del paralaje de referencia para las cámaras A y B están indicados respectivamente por  $d_A[x, y]$  y  $d_B[x, y]$ .  
50

Además, "MBBlk" es un índice para indicar cada macrobloque, y el número de macrobloques está indicado por "maxMBBlk".  
55

Después de que el índice MBBlk para cada macrobloque se haya inicializado en cero (véase el paso S104), se aplican repetidamente los pasos siguientes (S105 a S111) a cada macrobloque, mientras que se añade "1" al índice MBBlk (véase el paso S110).  
60

En primer lugar, después de que el blkMode de la división de bloques (índice) se haya inicializado en cero (véase el paso S105), se calcula un coste de distorsión de la velocidad para el macrobloque relevante MBBlk y el blkMode de la división de bloques (véase el paso S106) mientras que se añade "1" al blkMode de la división de bloques (véase el paso S107), hasta que el blkMode de la división de bloques alcanza el valor máximo del índice "maxBlkMode" (véase el paso S108). Es decir, en la parte 105 de ajuste del desplazamiento del paralaje, se calcula el coste de la distorsión de la velocidad "MBCost" para cada blkMode de la división de bloques.  
65

El coste de la distorsión de la velocidad MBCost para cada blkMode de la división de bloques se obtiene calculando el coste de la distorsión de la velocidad blkCost para cada bloque (el número total de bloques se indica con "maxBlk[blkMode]"), y calculando la suma total de los costes de distorsión de la velocidad blkCost.

5 El coste de la distorsión de la velocidad blkCost para cualquier bloque puede obtenerse como valor mínimo del coste de la distorsión de la velocidad "cost" obtenido cuando se utiliza en vector de desplazamiento del paralaje. Por tanto, para cada bloque, debe calcularse el vector del desplazamiento del paralaje para minimizar el "cost" y el propio "cost".

10 En este caso, se efectúa la búsqueda del desplazamiento del paralaje utilizado para la codificación, en la proximidad del paralaje de referencia relevante. Es decir, se definen " $e_0, e_1, \dots, e_{N-1}$ " como candidatos para el vector de desplazamiento del paralaje, y se repite un proceso de cálculo del coste de la distorsión de la velocidad para el vector de paralaje " $d_A[x, y] + e_n$ " para " $n = 0, 1, \dots, N-1$ ", para codificar el vector de paralaje óptimo para el bloque presente.

15 Con el fin de calcular el "cost", en primer lugar se calcula la suma total "SAD[ $e_n$ ]" de los valores absolutos de la predicción residual para el bloque relevante, que se obtiene cuando se utiliza el vector  $e_n$  del desplazamiento del paralaje. Además, se calcula también una cantidad estimada R[ $e_n$ ] de código del vector de desplazamiento del paralaje, que se obtiene cuando se codifica el vector  $e_n$  del desplazamiento del paralaje. El "cost" se calcula con la siguiente fórmula:

$$\text{Cost} = \text{SAD}[e_n] + \lambda R[e_n]$$

25 (1)

A continuación, en la parte 106 de ajuste de la división de bloques, se calcula el "bestBlkMode" de la división de bloques para minimizar el coste MBCost de la distorsión de velocidad, para cada macrobloque MBBIk (el correspondiente vector de desplazamiento del paralaje ha sido ya obtenido).

30 La figura 5 muestra el proceso detallado del paso S106.

En primer lugar, se inicializa en cero el coste MBCost de la distorsión de la velocidad del macrobloque relevante y el valor índice del bloque "blk" (véase el paso S1061), y se obtiene el vector del desplazamiento del paralaje para minimizar el coste blkCost de la distorsión de velocidad del bloque "blk" (véase el paso S1062).

Después de que el coste blkCost obtenido de la distorsión de la velocidad es añadido al coste MBCost de distorsión de la velocidad del macrobloque (véase el paso S1063), se añade "1" al bloque "blk", y se repiten los pasos S1062 y S1064 hasta que el bloque "

40 blk" alcanza el número total maxBlk[blkMode] de bloques (véase el paso S1065).

45 Cuando el bloque "blk" alcanza el número total maxBlk[blkMode] de bloques, se determina si el coste MBCost de la distorsión de la velocidad del macrobloque actualmente calcula es menor o no que el coste actualmente mínimo de distorsión de la velocidad minMBCost del macrobloque (véase el paso S1066). Si es menor, se almacena el coste MBCost de distorsión de la velocidad como el coste mínimo de distorsión de la velocidad minMBCost, y se almacena el correspondiente blkMode de división de bloques como el óptimo bestBlkMode de la división de bloques (véase el paso S1067).

50 De acuerdo con la operación anterior, se obtienen los datos "bestBlkMode" de la división de bloques, los datos del desplazamiento del paralaje (es decir, el vector de desplazamiento del paralaje) y la predicción residual generada por la compensación del paralaje, que se utilizan para la compensación del paralaje del macrobloque MBBIk. A continuación, se codifican los datos "bestBlkMode" de la división de bloques en la parte 107 de codificación de datos de la división de bloques, y se codifican los datos del desplazamiento del paralaje correspondientes al bestBlkMode en la parte 108 de codificación de datos del desplazamiento del paralaje (véase el paso S109).

Además, la predicción residual correspondiente a los datos del desplazamiento del paralaje son codificados en la parte 109 de codificación de la predicción residual.

60 La figura 7 muestra un dispositivo de descodificación de vídeo utilizado en el presente modo de realización. El dispositivo 200 de descodificación de vídeo incluye una parte 201 de descodificación de los datos de la división de bloques; una parte 202 de descodificación de los datos del desplazamiento del paralaje; una parte 203 de descodificación de la predicción residual; una parte 204 de compensación del paralaje; y una memoria 205 de imágenes de referencia.

65 La figura 8 es un diagrama de flujo de la descodificación utilizando el dispositivo 200 de descodificación de vídeo en

el presente modo de realización. Este diagrama de flujo muestra la operación de descodificar un fotograma de la cámara C, y será explicado con detalle a continuación. En este caso, los fotogramas de las cámaras A y B del mismo tiempo han sido descodificados por adelantado, y las imágenes descodificadas han sido almacenadas en la memoria 205 de imágenes de referencia.

5 En primer lugar, las imágenes descodificadas relevantes de las cámaras A y B son introducidas desde la memoria 205 de imágenes de referencia (véase el paso S201), y se obtiene el paralaje de referencia para cada pixel de la imagen objetivo a codificar, utilizando cada imagen de referencia (véase el paso S202). Es decir, se introducen dos imágenes de referencia que pertenecen a las cámaras A y B en la parte 204 de compensación del paralaje, y se obtiene la compensación del paralaje para cada pixel de la imagen relevante de la cámara C. Consecuentemente, se obtienen dos "vectores bidimensionales" para cada pixel en la imagen de la cámara C.

10 A continuación, después de que se haya inicializado en cero el índice MBBIk del macrobloque (véase el paso S203), se repiten los pasos siguientes (S204 a S212) para cada macrobloque del fotograma relevante, mientras que se añade "1" al MBBIk (véase el paso S211) ("maxMBBik" indica el número de macrobloques de un fotograma). Consecuentemente, se descodifica el fotograma relevante de la cámara C.

15 En la descodificación de cada macrobloque, en primer lugar, se descodifican los datos de la división de bloques bestBlkMode del macrobloque MBBIk por medio de la parte 201 de descodificación de los datos de la división de bloques (véase el paso S204). A continuación, después de haber inicializado a cero el valor del índice del bloque "blk" (véase el paso S205), se repiten los pasos siguientes (S206 a S209) para cada bloque "blk" ("maxBlk[blkMode]" indica el número máximo de bloques).

20 En primer lugar, en la parte 202 de descodificación de los datos del desplazamiento del paralaje, se descodifican los datos del desplazamiento del paralaje para el bloque "blk" (véase el paso S206), y se genera una imagen de predicción para el bloque "blk", utilizando los datos del desplazamiento del paralaje y del paralaje de referencia (es decir, efectuando la interpolación entre los valores de pixel de las cámaras A y B) (véase el paso S207).

25 De acuerdo con el proceso de repetición anterior para el bloque "blk", se genera una imagen de predicción para el macrobloque MBBIk. Por tanto, la predicción residual para el macrobloque MBBIk se descodifica en la parte 203 de descodificación de la predicción residual, y la parte 204 de compensación del paralaje calcula la suma de la imagen de predicción y de la predicción residual, obteniendo con ello una imagen descodificada del macrobloque MBBIk (véase el paso S210).

30 La operación anterior se repite para cada macrobloque del fotograma relevante (véanse los pasos S211 y S212), descodificando con ello el fotograma de la cámara C.

35 En el presente modo de realización, se codifica una imagen de la cámara C haciendo referencia a las otras cámaras (A y B). Sin embargo, tal caso de efectuar la codificación refiriéndose a las otras cámaras y otro caso de efectuar la compensación del movimiento haciendo referencia a una imagen descodificada, pueden ser conmutados adaptativamente en un fotograma de imagen (es decir, para una imagen).

40 Más específicamente, se puede efectuar la conmutación entre la compensación del paralaje y la compensación del movimiento para cada bloque dividido ilustrado en la figura 6. En este caso, el lado de codificación de vídeo debe codificar los datos para indicar qué compensación de paralaje y compensación de movimiento han sido utilizadas, y el lado de la descodificación de vídeo debe descodificar los datos.

45 Los procesos de codificación y descodificación de vídeo anteriormente descritos pueden ser implementados utilizando un ordenador y un programa de software. Tal programa puede ser proporcionado almacenándolo en un medio de almacenamiento legible por ordenador, o a través de una red.

### Aplicación industrial

50 En los métodos de codificación y descodificación de vídeo para estimar el paralaje utilizando imágenes de referencia, se utiliza la correlación espacial de los datos de desplazamiento del paralaje, de manera que la cantidad de código de los datos de desplazamiento del paralaje pueden ser reducidos sin degradar la eficiencia de la predicción en la compensación del paralaje, mejorando con ello la eficiencia total de la codificación.

**REIVINDICACIONES**

1. Un método de codificación de vídeo, para codificar imágenes de vídeo como una imagen de vídeo de puntos de vista múltiples, utilizando la compensación del paralaje que efectúa la predicción utilizando el paralaje espacial entre las imágenes de vídeo, comprendiendo el método:
- 5 un paso (S103) de ajuste del paralaje de referencia, para ajustar el paralaje de referencia para cada pixel de una imagen objetivo a codificar, donde se estima el paralaje de referencia utilizando imágenes de referencia sin usar la imagen objetivo;
- 10 un paso (S1067) de ajuste de la división del área, para ajustar la división del área en un fotograma de una imagen;
- un paso (S1062) de ajuste del desplazamiento del paralaje, para ajustar el desplazamiento del paralaje para cada área dividida fijada en el paso de ajuste de la división del área, donde se define el desplazamiento del paralaje como la diferencia entre el paralaje de referencia y el paralaje real utilizado para generar una imagen de predicción para la compensación del paralaje, y tiene un valor común dentro de cada área dividida;
- 15 un paso de generación de la imagen de predicción, para generar una imagen de predicción para la compensación del paralaje, para cada pixel de la imagen objetivo, utilizando el paralaje real obtenido para el pixel como una unidad añadiendo el paralaje de referencia ajustado al desplazamiento del paralaje común fijado para el área a la cual pertenece el pixel;
- 20 un paso (S109) de codificación de los datos de la división del área, para codificar los datos de la división del área para indicar la división del área que ha sido ajustada en el paso de ajuste de la división del área; y
- 25 un paso (S109) de codificación de los datos de desplazamiento del paralaje, para codificar los datos de desplazamiento del paralaje para indicar el desplazamiento del paralaje que ha sido fijado en el paso de ajuste del desplazamiento del paralaje.
- 30 2. El método de codificación de vídeo, de acuerdo con la reivindicación 1, en el que:
- el ajuste de la división del área en el paso de ajuste de la división del área, se efectúa por medio de un método seleccionado entre una pluralidad de métodos de división del área para realizar la división del área en bloques rectangulares.
- 35 3. Un método de descodificación de vídeo, para descodificar imágenes de vídeo como una imagen de vídeo de puntos de vista múltiples, utilizando la compensación del paralaje que realiza la predicción, utilizando el paralaje espacial entre las imágenes de vídeo, comprendiendo el método:
- 40 un paso (S202) de ajuste del paralaje de referencia, para fijar el paralaje de referencia para cada pixel en una imagen objetivo a descodificar, donde el paralaje de referencia se estima utilizando imágenes de referencia sin utilizar la imagen objetivo;
- un paso (S204) de descodificación de datos de la división de área, para descodificar los datos de división del área para indicar la división del área, donde los datos de división de área están incluidos en los datos codificados;
- 45 un paso (S206) de descodificación de los datos de desplazamiento del paralaje, para descodificar los datos del desplazamiento del paralaje para cada área indicada por los datos de división del área, que son descodificados en el paso de descodificación de datos de la división del área, donde los datos del desplazamiento del paralaje están incluidos en los datos codificados, y el desplazamiento del paralaje se define como la diferencia entre el paralaje de referencia y el paralaje real utilizado para generar una imagen de predicción para la compensación del paralaje y tiene un valor común dentro de cada área indicada por los datos de división del área; y
- 50 un paso (S207) de generación de la imagen de predicción, para generar la imagen de predicción para la compensación del paralaje para cada pixel de la imagen objetivo, utilizando el paralaje real obtenido para el pixel como una unidad, añadiendo el paralaje de referencia ajustado al desplazamiento del paralaje común que se ha fijado para el área a la cual pertenece el pixel.
- 55 4. El método de descodificación de vídeo, de acuerdo con la reivindicación 3, en el que:
- 60 los datos de división del área descodificados en el paso de descodificación de los datos de la división del área indican un método seleccionado entre una pluralidad de métodos de división de área para efectuar la división del área en bloques rectangulares.
- 65 5. Un dispositivo de codificación de vídeo para codificar imágenes de vídeo como una imagen de vídeo de puntos de vista múltiples, utilizando la compensación del paralaje que realiza la predicción utilizando el paralaje espacial entre

imágenes de vídeo, comprendiendo el dispositivo:

5 un dispositivo (104) de ajuste del paralaje de referencia, para fijar el paralaje de referencia para cada pixel de una imagen objetivo a codificar, donde el paralaje de referencia se estima utilizando imágenes de referencia sin utilizar la imagen de referencia;

un dispositivo (106) de ajuste de la división del área, para fijar la división del área en un fotograma de una imagen;

10 un dispositivo (105) de ajuste del desplazamiento del paralaje, para fijar el desplazamiento del paralaje para cada área dividida fijada por el dispositivo de ajuste de división del área, donde el desplazamiento del paralaje está definido como la diferencia entre el paralaje de referencia y el paralaje real utilizado para generar una imagen de predicción para la compensación del paralaje, y tiene un valor común dentro de cada área dividida;

15 un dispositivo (109) de generación de imágenes de predicción, para generar la imagen de predicción para la compensación del paralaje, para cada pixel de la imagen objetivo, utilizando el paralaje real que es obtenido para el pixel como unidad, añadiendo el paralaje de referencia fijado al desplazamiento del paralaje común que es fijado por el área a la cual pertenece el pixel;

20 un dispositivo (107) de codificación de datos de división del área, para codificar los datos de división del área para indicar la división de área que ha sido fijada por el dispositivo de ajuste de división del área; y

25 un dispositivo (108) de codificación de los datos del desplazamiento del paralaje, para codificar los datos de desplazamiento del paralaje para indicar el desplazamiento del paralaje que se ha fijado por el dispositivo de ajuste de desplazamiento del paralaje.

6. Un dispositivo de descodificación de vídeo, para descodificar imágenes de vídeo como una imagen de vídeo de puntos de vista múltiples, utilizando la compensación del paralaje que efectúa la predicción utilizando el paralaje espacial entre las imágenes de vídeo, comprendiendo el dispositivo:

30 un dispositivo (204) de ajuste del paralaje de referencia, para ajustar el paralaje de referencia para cada pixel en una imagen objetivo a descodificar, donde el paralaje de referencia se estima utilizando imágenes de referencia sin utilizar la imagen objetivo;

35 un dispositivo (201) de descodificación de los datos de división del área, para descodificar los datos de división del área para indicar la división del área, donde los datos de división del área están incluidos en los datos codificados;

40 un dispositivo (202) de descodificación de los datos de desplazamiento del paralaje, para descodificar datos del desplazamiento del paralaje para cada área indicada por los datos de división del área, que son descodificados por el dispositivo de descodificación de datos de división del área, donde los datos del desplazamiento del paralaje están incluidos en los datos codificados, y el desplazamiento del paralaje se define como la diferencia entre el paralaje de referencia y el paralaje real utilizado para generar una imagen de predicción para la compensación del paralaje, y tiene un valor común dentro de cada área indicada por los datos de división del área; y

45 un dispositivo (204) de generación de imágenes de predicción, para generar la imagen de predicción para la compensación del paralaje, para cada pixel de la imagen objetivo, utilizando el paralaje real que se obtiene para el pixel como una unidad, añadiendo el paralaje de referencia fijado al desplazamiento del paralaje común, que está fijado para el área a la cual pertenece el pixel.

50 7. Un programa de codificación de vídeo, para hacer que un ordenador ejecute el método de codificación de vídeo de acuerdo con la reivindicación 1.

8. Un programa de descodificación de vídeo, para hacer que un ordenador ejecute el método de codificación de vídeo de acuerdo con la reivindicación 3.

55 9. Un medio de almacenamiento legible por ordenador, que almacena un programa de descodificación de vídeo para hacer que un ordenador ejecute el método de codificación de vídeo de acuerdo con la reivindicación 1.

10. Un medio de almacenamiento legible por ordenador, que almacena un programa de descodificación de vídeo para hacer que un ordenador ejecute el método de descodificación de vídeo de acuerdo con la reivindicación 3.

60 11. El método de codificación de vídeo, de acuerdo con la reivindicación 1, en el que:

en el paso de generación de imágenes de predicción, la imagen de predicción se genera basándose en los valores de los píxeles de las imágenes de referencia.

65 12. El método de codificación de vídeo, de acuerdo con la reivindicación 11, en el que:

en el paso de generación de imágenes de predicción, la imagen de predicción se genera calculando un promedio entre los valores de los píxeles de las imágenes de referencia.

5 13. El método de decodificación de vídeo, de acuerdo con la reivindicación 3, en el que:

en el paso de generación de imágenes de predicción, la imagen de predicción se genera basándose en los valores de los píxeles de las imágenes de referencia.

10 14. El método de decodificación de vídeo, de acuerdo con la reivindicación 13, en el que:

en el paso de generación de imágenes de predicción, la imagen de predicción se genera calculando un promedio entre los valores de los píxeles de las imágenes de referencia.

15 15. El dispositivo de codificación de vídeo, de acuerdo con la reivindicación 6, en el que:

el dispositivo de generación de imágenes de predicción genera la imagen de predicción basándose en los valores de los píxeles de las imágenes de referencia.

20 16. El dispositivo de codificación de vídeo, de acuerdo con la reivindicación 15, en el que:

el dispositivo de generación de imágenes de predicción genera la imagen de predicción calculando el promedio entre los valores de los píxeles de las imágenes de referencia.

25 17. El dispositivo de decodificación de vídeo, de acuerdo con la reivindicación 7, en el que:

el dispositivo de generación de imágenes de predicción genera la imagen de predicción basándose en los valores de los píxeles de las imágenes de referencia.

30 18. El dispositivo de decodificación de vídeo, de acuerdo con la reivindicación 17, en el que:

el dispositivo de generación de imágenes de predicción genera la imagen de predicción calculando el promedio entre los valores de los píxeles de las imágenes de referencia.

FIG. 1

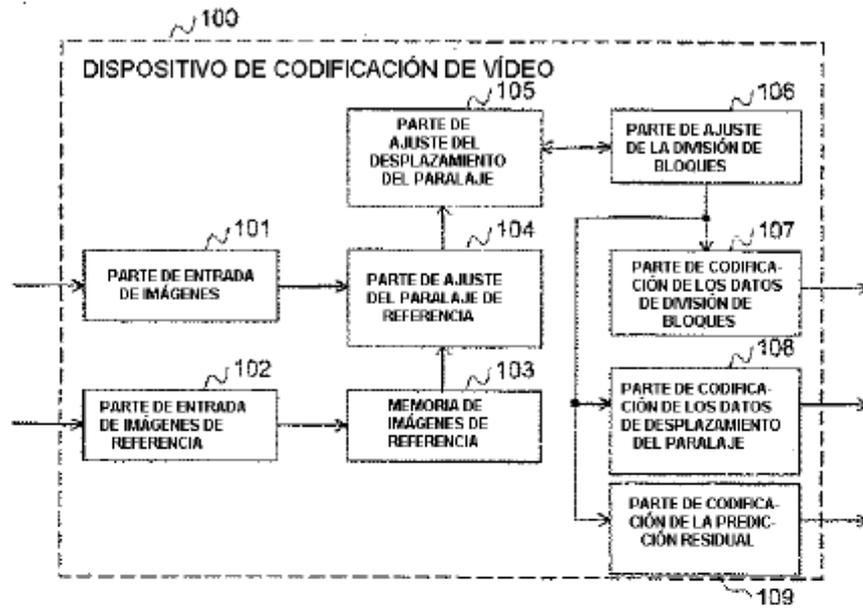


FIG. 2

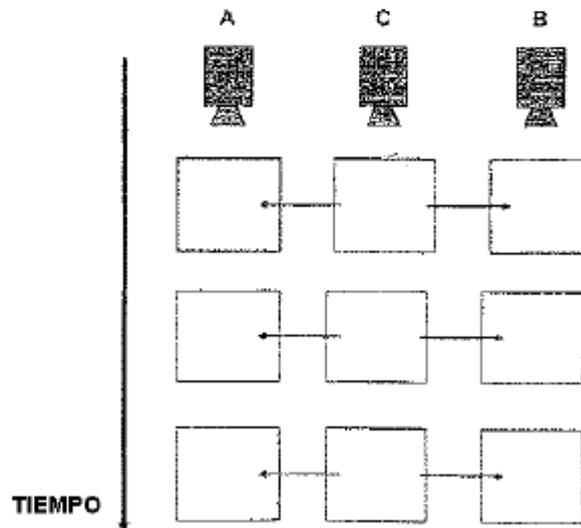


FIG. 3

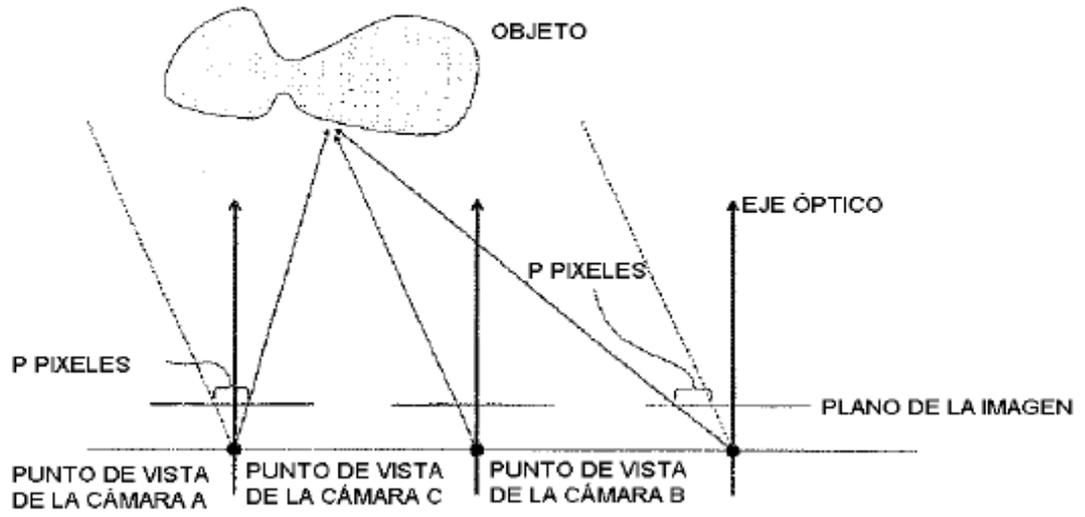


FIG. 4

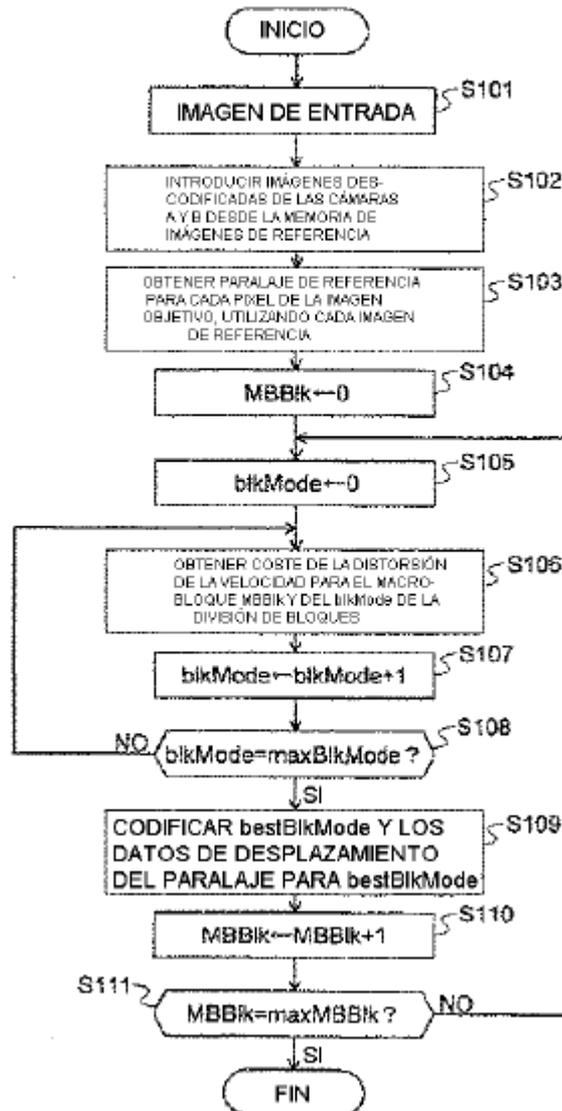


FIG. 5

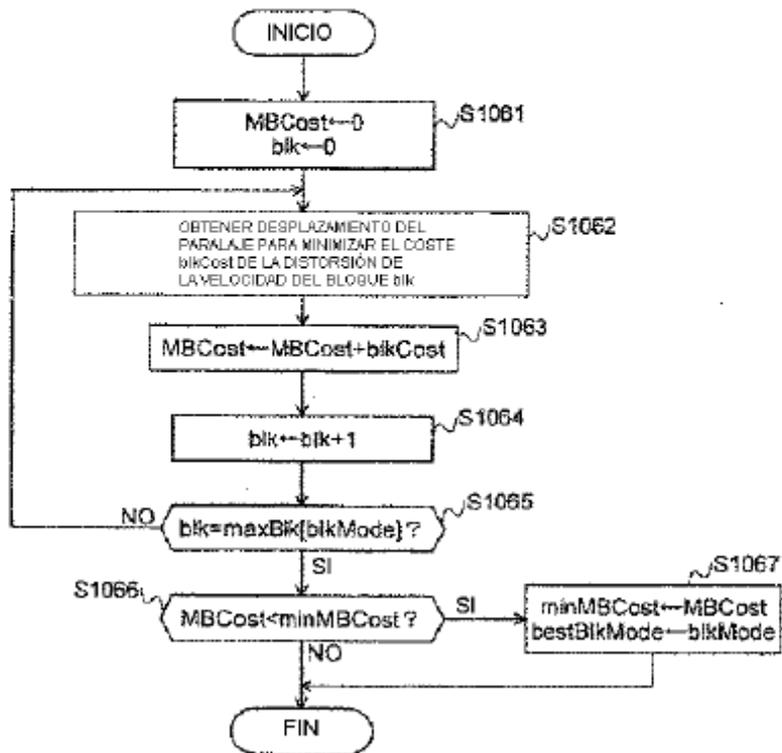


FIG. 6

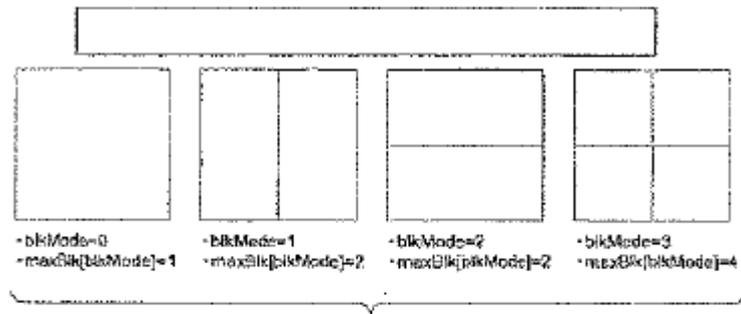


FIG. 7

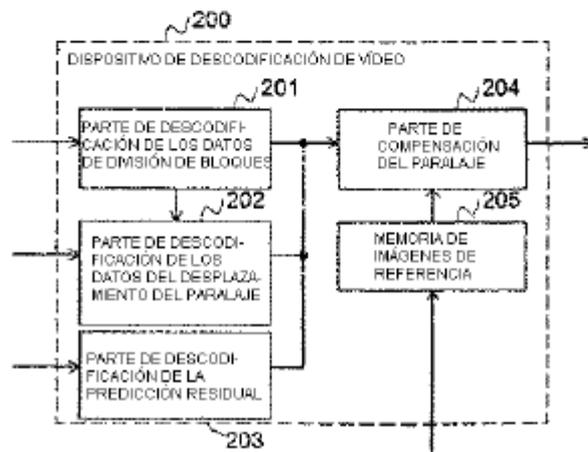


FIG. 8

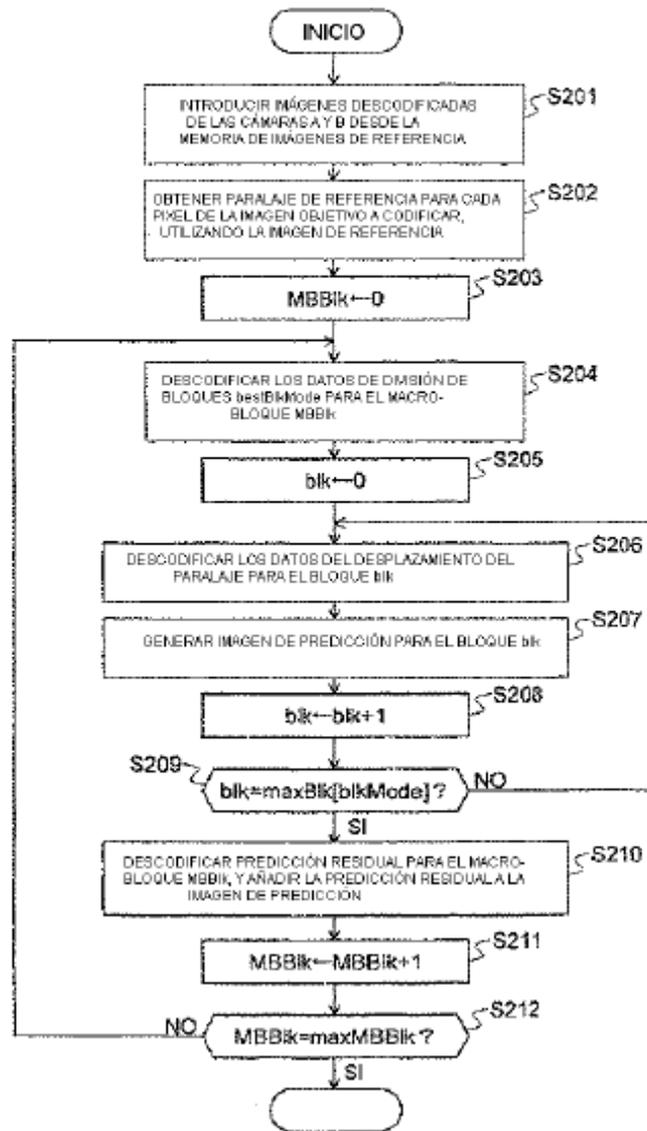


FIG. 9

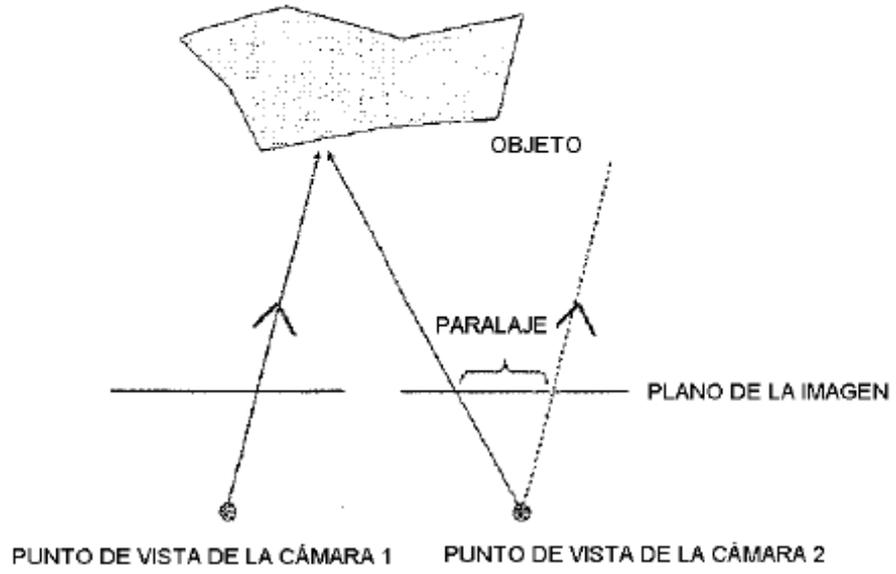


FIG. 10

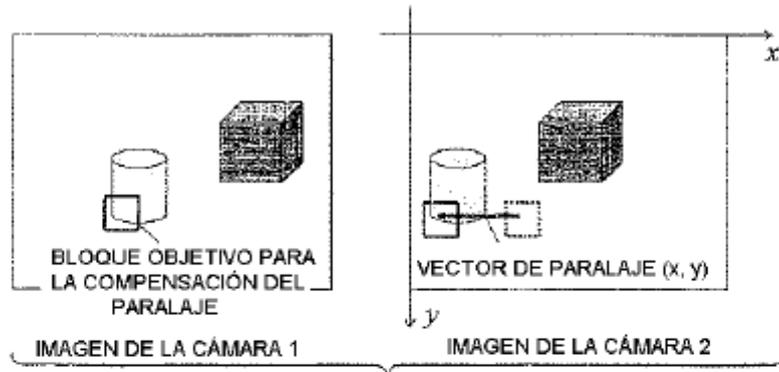


FIG. 11

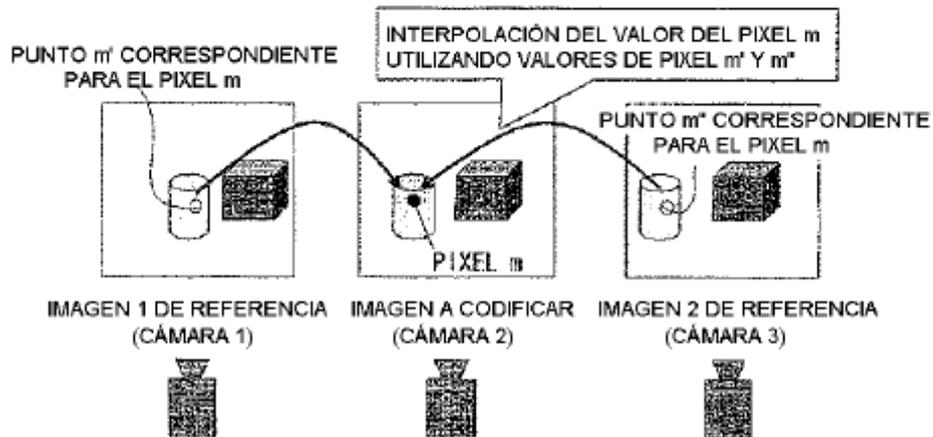


FIG. 12

