

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 729 193**

51 Int. Cl.:

H04N 19/597 (2014.01)

H04N 13/00 (2008.01)

H04N 19/51 (2014.01)

G06T 7/223 (2007.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.01.2002 PCT/FI2002/00050**

87 Fecha y número de publicación internacional: **01.08.2002 WO02060183**

96 Fecha de presentación y número de la solicitud europea: **23.01.2002 E 02710906 (5)**

97 Fecha y número de publicación de la concesión europea: **06.03.2019 EP 1364535**

54 Título: **Procedimiento de compresión de información de vídeo**

30 Prioridad:

24.01.2001 FI 20010143

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

30.10.2019

73 Titular/es:

**BROADMEDIA GC CORPORATION (100.0%)
8-4-14 Akasaka, Minato-ku
Tokyo 107-0052, JP**

72 Inventor/es:

**PIEHL, ERIK y
SALLINEN, SAMI**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 729 193 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 compresión de información de vídeo

La invención se refiere a un procedimiento para comprimir vídeo digital. El procedimiento es aplicable especialmente para acelerar la compresión de vídeo digital y/o mejorar la calidad del resultado. Además, la invención es especialmente adecuada para casos en los que el material de vídeo se genera sintéticamente a partir de modelos tridimensionales.

El vídeo en movimiento consiste en imágenes fijas individuales que se muestran al espectador sucesivamente a una velocidad lo suficientemente alta como para generar la sensación de movimiento. El objetivo de la compresión de vídeo es disminuir la cantidad de datos necesarios para describir una imagen de vídeo en movimiento, cuando la información sobre la imagen de vídeo en movimiento se almacena o transmite.

Todos los procedimientos de compresión de vídeo conocidos en uso común hoy en día principalmente logra este objetivo mediante la identificación y eliminación de información repetitiva de las imágenes. Prácticamente, esto significa que se identifican los cambios entre las imágenes y se le dice al receptor cómo construir la siguiente imagen utilizando la anterior (o alguna otra imagen comprimida anterior) como referencia.

Un procedimiento para construir una nueva imagen basada en los anteriores es identificar qué partes de la imagen se han movido entre la imagen de referencia y la imagen siguiente. El proceso de encontrar los parámetros de movimiento se llama estimación de movimiento, y generalmente se realiza bloque por bloque buscando bloques similares en áreas cercanas de la imagen de referencia. Las imágenes normalmente se dividen en bloques de una manera fija determinada, pero la partición (es decir, los tamaños de bloques y la disposición de los bloques) se puede hacer alternativamente, por ejemplo, simultáneamente con la estimación de movimiento, teniendo en cuenta los movimientos detectados en la partición. Luego se le dice al descompresor la dirección y la longitud del movimiento de los bloques en la siguiente imagen usando un vector de movimiento. La información sobre la partición de bloques y los vectores de movimiento de los bloques permite al descompresor producir una estimación de la siguiente imagen original utilizando una imagen de referencia reconstruida anteriormente. Esta acción del descompresor se denomina compensación de movimiento, y en la estimación compensada por movimiento de la siguiente imagen se imita el movimiento en la imagen siguiente original.

La compensación de movimiento no suele producir una copia exacta de la imagen siguiente, por lo que una corrección a la imagen con compensación de movimiento se crea por el compresor imagen. Normalmente, el compresor determina la imagen de corrección como la diferencia entre la estimación de movimiento compensado y la imagen siguiente original. Esta imagen de corrección se llama imagen de diferencia.

La estimación de movimiento en el lado del compresor requiere una búsqueda iterativa en un espacio de búsqueda bidimensional. En la mayoría de los algoritmos de compresión de vídeo estandarizados, se toma un bloque de la siguiente imagen y luego se compara con bloques cercanos en la imagen de referencia. Se supone que la diferencia de ubicación entre el bloque más similar encontrado y el bloque buscado es el vector de movimiento. Debido a la gran cantidad de operaciones requeridas en el proceso de búsqueda iterativo, una búsqueda exhaustiva del vector de movimiento ideal no es práctica en los sistemas de compresión de vídeo por hardware u ordenadores actuales.

Los sistemas de compresión de vídeo actualmente disponibles intentan evitar la alta carga informática usando varias técnicas inteligentes y haciendo algunos compromisos. Uno de estos procedimientos es limitar el área desde la cual se realiza la búsqueda, limitando así también la longitud máxima del vector de movimiento que se puede expresar. Otro es estimar el vector de movimiento de un bloque basado en el vector de movimiento de los bloques vecinos. Estos procedimientos, sin embargo, comprometen la eficiencia y la calidad del proceso de compresión.

En la técnica relacionada, el artículo de A. Aydin Alatan y Levent Onural titulada "Estimación de campos de profundidad adecuado para la compresión de vídeo basada en la estructura 3D y movimiento de los objetos" (IEEE Transactions on Image Processing, vol. 7, n.º 6, junio de 1998, páginas 904-908) describe una técnica para la predicción de intensidad a lo largo de las trayectorias de movimiento, cuya técnica tiene como objetivo eliminar la redundancia temporal en los algoritmos de compresión de vídeo. En la codificación de vídeo tridimensional (3D) basada en objetos, se requieren valores de movimiento y profundidad en 3D para la predicción temporal. Los parámetros de movimiento 3D requeridos para cada objeto se encuentran mediante el procedimiento de matriz electrónica basada en la correspondencia. Las estimaciones de la correspondencia - campo de movimiento bidimensional (2D) - entre los cuadros y la segmentación de la escena en objetos se logran simultáneamente al minimizar la energía de Gibbs. El campo de profundidad se calcula minimizando conjuntamente un criterio definido de distorsión y velocidad de bits utilizando los parámetros de movimiento 3D, lo que da como resultado un campo de profundidad que es eficiente en el sentido de velocidad-distorsión.

Un objeto de la presente invención es crear un procedimiento eficaz para la compresión de vídeo. Preferiblemente, el procedimiento permite comprimir el vídeo notablemente más eficientemente que los procedimientos conocidos actualmente sin comprometer la calidad.

Estos y otros objetos de la invención se consiguen mediante la determinación de los vectores de movimiento

causados por los movimientos de la cámara para el uso de estos vectores de movimiento en compresión de vídeo.

Un procedimiento de acuerdo con la invención se caracteriza por lo que se afirma en la parte caracterizadora de la reivindicación independiente dirigida a un procedimiento. Un programa informático según la invención se caracteriza por lo que se indica en la reivindicación independiente dirigida a un programa informático. Un dispositivo para comprimir información de vídeo de acuerdo con la invención se caracteriza por lo que se indica en la porción caracterizadora de la reivindicación independiente dirigida a dicho dispositivo.

Las reivindicaciones dependientes describen realizaciones preferidas de la invención.

El principio básico de la invención es que cuando el vídeo a ser comprimido origina a partir de un modelo sintético conocido o cuando es posible crear un modelo de mapa de profundidad de la escena en las nuevas posibilidades de vídeo que surgen para el proceso de estimación de movimiento.

Cuando un mapa de profundidad de la escena en el vídeo es conocido, esto significa - intuitivamente - que es posible determinar las ubicaciones de los objetos, que están presentes en la escena, en un espacio tridimensional. Con términos más precisos, los puntos en la superficie de la imagen se proyectan al espacio tridimensional. El espacio tridimensional puede corresponder a un modelo sintético o al mundo real. Cuando se conocen los cambios en la ubicación y dirección de la cámara entre una imagen de referencia y la imagen siguiente, es posible determinar a partir de los movimientos de la cámara y el uso de las ubicaciones conocidas de los objetos en el espacio tridimensional (o proyecciones al espacio tridimensional) tales vectores de movimientos para la siguiente imagen, que tienen en cuenta los movimientos de la cámara. Si no hay otros cambios en la escena que los relacionados con los movimientos de la cámara (es decir, los objetos no se mueven entre las instancias de tiempo relacionadas con la imagen de referencia y la siguiente imagen), estos vectores de movimiento describen con precisión la siguiente imagen en relación con la imagen de referencia.

Las imágenes de vídeo que se han de comprimir pueden ser imágenes del mundo real en tres dimensiones. En este caso, es posible determinar o estimar las ubicaciones de los objetos con respecto a la cámara, ya que una imagen en sí misma revela información sobre las ubicaciones de los objetos proyectados en un plano e información sobre la distancia de un objeto desde una cámara puede ser obtenido, por ejemplo, utilizando un sensor de rango. Alternativamente, las imágenes de vídeo pueden originarse a partir de un modelo sintético. En gráficos de ordenador, cuando se determina una vista de un modelo sintético, es necesario averiguar qué objetos del modelo son visibles para un espectador y cuáles no. En general, se utiliza una memoria intermedia z: se almacena en la memoria intermedia z un valor de profundidad (valor z) de cada píxel que se va a mostrar, y solo se dibuja un nuevo píxel, si su valor z es más pequeño que el píxel que ya se encuentra en el mismo lugar. La memoria intermedia z es un mapa de profundidad de la vista, y tal mapa de profundidad se genera normalmente cada vez que se produce una vista de un modelo sintético en gráficos de ordenador.

Los vectores de movimiento relacionados con los movimientos de la cámara pueden determinarse para los bloques y usarse directamente como vectores de movimiento de los bloques en compresión de vídeo. Alternativamente, los vectores de movimiento relacionados con los movimientos de la cámara se pueden usar como valores iniciales y se pueden determinar vectores de movimiento más precisos para los bloques, por ejemplo, buscando un bloque similar cerca del área, al que apunta el vector de movimiento de la cámara.

Además de determinar los vectores de movimiento para los bloques en la imagen siguiente, es posible determinar un campo de vectores de movimiento para la siguiente imagen. Un campo de vector de movimiento generalmente se refiere a la información del vector de movimiento, que está en una escala más fina que los bloques. El campo del vector de movimiento describe típicamente el movimiento en la escala de píxeles. El campo del vector de movimiento se puede determinar directamente en un nivel de píxeles de manera similar a como lo hacen los bloques. Esto es sencillo, ya que el mapa de profundidad generalmente define la información de distancia para los píxeles de la imagen. Alternativamente, es posible determinar un campo de vector de movimiento en un nivel de bloque, y usar estos vectores de movimiento específicos de bloque para interpolar el campo de vector de movimiento específico de píxel.

El campo vector de movimiento se puede usar como un campo de deformación de la imagen de referencia. Esto significa que, además de la traducción simple, una parte de la imagen de referencia también es posible la contracción/expansión y/o la rotación de partes de la imagen de referencia cuando se forma la estimación de movimiento compensado de la siguiente imagen. El uso del campo de vector de movimiento en la compensación de movimiento requiere que el descompresor sea capaz de generar un campo de vector de movimiento a partir de la información que recibe y deformar la imagen de referencia de acuerdo con la imagen de referencia. Un descompresor puede usar, por ejemplo, un procedimiento de interpolación, que el descompresor indica al comienzo del flujo de información comprimida o, alternativamente, el compresor y el descompresor pueden negociar un procedimiento de interpolación soportado por ambos. Normalmente, no es necesario cambiar el procedimiento de interpolación durante la compresión de una secuencia de imágenes, pero esto se puede hacer, siempre que se informe al descompresor sobre el nuevo procedimiento de interpolación.

Para una precisión de la reconstrucción de la imagen, un compresor normalmente no utiliza una imagen anterior

- como una imagen de referencia tal como. Normalmente, un compresor reconstruye una imagen de referencia utilizando la misma información disponible para el descompresor. Esta imagen de referencia reconstruida se utiliza para determinar la estimación de movimiento y la imagen de corrección. A continuación, en la descripción detallada de la invención, el término imagen de referencia puede ser una imagen original anterior como tal o una imagen anterior reconstruida, dependiendo de los detalles de la compresión de la imagen.
- 5 La invención se ilustra a modo de ejemplo en las figuras adjuntas, en las que
- La figura 1a muestra una imagen de vídeo imaginaria,
- La figura 1b muestra el mapa de profundidad asociado con la imagen de la figura 1a,
- La figura 2a muestra una nueva imagen de la misma escena a la que se ha movido la cámara,
- 10 La figura 2b muestra el mapa de profundidad de la nueva imagen de vídeo,
- La figura 3 muestra las posiciones de la cámara para las imágenes de vídeo de las figuras 1a y 2a,
- La figura 4 muestra la geometría relacionada con la proyección de un punto en la imagen de vídeo que se puede proyectar hacia el espacio tridimensional utilizando la información del mapa de profundidad,
- 15 La figura 5 muestra la aplicación de la invención a un juego de ordenador conocido en una situación en la que la cámara está estacionaria.
- La figura 6 muestra la cámara moviéndose aproximadamente en línea recta,
- La figura 7 muestra la cámara girando de derecha a izquierda,
- Las figuras 8a, 8b, 8c muestran un diagrama de flujo de un procedimiento de acuerdo con una realización de la invención, y
- 20 La figura 9 muestra un compresor de vídeo de acuerdo con una realización adicional de la invención.
- La siguiente explicación de la invención se refiere a las figuras 1-4.
- Un procedimiento de acuerdo con la invención requiere que los siguientes datos comiencen tanto con la imagen de referencia y la siguiente imagen que se comprime: la imagen real y los parámetros de la cámara de la imagen. También se necesita un mapa de profundidad correspondiente a la siguiente imagen. Un mapa de profundidad suele ser una matriz bidimensional de números que indica a qué distancia de la cámara está el objeto visible en cada píxel de la imagen real.
- 25 Las figuras 1a y 2a ilustran, como ejemplos, dos imágenes de vídeo: en la figura 1a se ilustra una imagen *I0* de vídeo, y en la figura 2a se ilustra una imagen *I1* de vídeo. Las imágenes *I0* e *I1* se relacionan con una misma escena, pero la ubicación y la dirección de la cámara que mira la escena son diferentes para las imágenes *I0* e *I1*.
- 30 La figura 3 ilustra la ubicación y la dirección de la cámara para las imágenes *I0* y *I1*.
- Los parámetros de la cámara para la imagen que definen la ubicación de la cámara, dirección de la cámara y la distancia de visión de la cámara. La ubicación de la cámara *K* normalmente se presenta como un punto (o un vector que apunta a esta ubicación) en un espacio tridimensional. La dirección de la cámara se presenta normalmente con tres vectores *Kx*, *Ky* y *Kz* de dirección de cámara. La distancia *Dv* de visualización es la distancia de la cámara desde el plano de la imagen. La figura 4 ilustra estos parámetros *K*, *Kx*, *Ky*, *Kz* y *Dv* de cámara.
- 35 Las coordenadas *x* e *y* de los puntos en el plano de imagen se definen típicamente como distancias (en las direcciones definidas por vectores *Kx* y *Ky* de dirección) desde el punto en la superficie de imagen, donde la dirección definida por el vector *Kz* intersecta la superficie de la imagen; normalmente este punto es el punto central de la imagen. La figura 4 ilustra la definición anterior de la coordenada *x*; en la figura 4, la coordenada está marcada con *x1* ya que la figura 4 se relaciona especialmente con la imagen *I1*.
- 40 Las figuras 1b y 2b ilustran mapas *Z0* y *Z1* de profundidad de las imágenes *I0* e *I1*: más oscuro es el tono de gris, mayor es la distancia del objeto en ese punto de la imagen de la cámara. Un mapa *Z* de profundidad especifica la distancia de un objeto *O* desde la cámara en la dirección del vector *Kz*. La figura 4 ilustra, cómo se puede determinar la ubicación del objeto *O*, que se ilustra en la imagen *I1* en el punto (*x1*, *y1*), en un espacio tridimensional utilizando *x1*, *y1*, *Z1* (*x1*, *y1*) y parámetros de la cámara de imagen *I1*. En otras palabras, la figura 4 ilustra la proyección del punto (*x1*, *y1*) en la superficie de la imagen a un punto *P'* del espacio tridimensional.
- 45 Cuando una imagen es una vista generada a partir de un modelo sintético, el mapa *Z* de profundidad normalmente se genera como un producto secundario del proceso de creación de la imagen, como se discutió anteriormente. Cuando la imagen es una imagen del mundo real, es posible determinar el mapa de profundidad utilizando, por ejemplo, sensores de rango.
- 50

Consideremos a continuación, como un ejemplo, una situación, donde la imagen *I0* de vídeo en la figura 1a funciona como una imagen de referencia para la imagen *I1* de vídeo que se ilustra en la figura 2a. Además, como ejemplo, consideremos un bloque de imágenes que contiene la cabeza de la persona en las imágenes *I0* e *I1*.

5 Con el fin de calcular una aproximación útil para el vector de movimiento del bloque en estudio, queremos averiguar donde el píxel *p1* (figura 2a) estaba en la imagen *I0* de referencia, suponiendo que todos los cambios entre las imágenes *I0* y *I1* se deben a cambios en los parámetros de la cámara. La ubicación anterior se llama *p0* (figura 1a). Para encontrar los componentes *x* e *y* de *p0* (es decir, *x0* y *y0*), el punto *p1* de imagen se proyecta primero hacia el espacio *P'* tridimensional utilizando los parámetros de la cámara de la imagen *I1* (figura 4) y luego el punto del espacio tridimensional se proyecta a la superficie de la imagen *I0* de referencia utilizando los parámetros de la cámara de la imagen de referencia. La diferencia en la ubicación del punto en las dos imágenes, generalmente definida como la diferencia con signo *p0-p1*, es el vector de movimiento para usar en la compresión de vídeo.

La proyección de la superficie de la imagen *I* al espacio de tres dimensiones se realiza con la siguiente fórmula:

$$P' = K + \frac{x*Z(x,y)}{D_v} * \bar{K}x + \frac{y*Z(x,y)}{D_v} * \bar{K}y + Z(x,y) * \bar{K}z$$

15 en la que *x* e *y* son coordenadas de un punto en la superficie de la imagen y *Z(x, y)* es el valor del mapa de profundidad del punto (*x, y*). La proyección desde el espacio tridimensional al espacio de la imagen se puede hacer con la fórmula que resulta de resolver *x, y* y *Z(x, y)* de la fórmula anterior. Ambas fórmulas son conocidas fórmulas básicas de gráficos tridimensionales por ordenador. Las coordenadas, los parámetros de la cámara y el mapa de profundidad pueden definirse de una manera ligeramente diferente a las descritas anteriormente, pero es evidente para un experto en la materia modificar la fórmula de proyección en consecuencia.

20 El procedimiento presentado para calcular el vector de movimiento tiene en cuenta todos los cambios en las imágenes de vídeo que dan lugar a cambios en los parámetros de la cámara. Independientemente de los posibles cambios en la escena que ve la cámara, los píxeles/bloques de una imagen siguiente pueden proyectarse primero, utilizando los parámetros de la cámara de la siguiente imagen, al espacio tridimensional y luego, utilizando los parámetros de la cámara de la imagen de referencia, a la superficie de la imagen de referencia. Queremos hacer notar a continuación que el punto relacionado con un determinado objeto en la siguiente imagen con *p1* y en la imagen de referencia con *p_{ref}*. Sin embargo, el valor de *p_{ref}* es, por lo general, desconocido cuando se comprime una imagen siguiente; el objetivo de la compensación de movimiento es determinar la estimación de *p_{ref}*.

30 Queremos hacer notar que un punto *p1* de la siguiente imagen una vez proyectado con *P1'* (esto es un punto del espacio tridimensional) y un punto dos veces proyectada con *P1''* (este es un punto (*x1'', y1''*) en la referencia imagen de la superficie, y tiene un valor *z1''* de profundidad). Si los cambios en la imagen de vídeo en las áreas, donde están *p1* y *p1''*, son causados solo por cambios en los parámetros de la cámara, entonces *p1''* es igual o cercano a *p_{ref}*. Las figuras 1-3 ilustran este caso. Una posible forma de estimar, si *p1''* es igual o cerca de *p_{ref}*, es comparar el valor de profundidad *z1''* con el valor *Z0(x1'', y1'')* de mapa de profundidad de imagen de referencia. Estos valores coinciden estrechamente si los cambios en los parámetros de la cámara provocan cambios en las imágenes. En este caso, normalmente no hay necesidad de tratar de encontrar un vector de movimiento más preciso para *p1* que *p1'' - p1*.

40 Si la escena vista por la cámara cambia, proyectar dos veces un punto de la siguiente imagen como se describe anteriormente produce vectores de movimiento no precisos en esas áreas de la imagen. Una manera informáticamente económica de estimar si existe la necesidad de determinar vectores de movimiento más precisos es, como se describe anteriormente, comparar el valor *z1''* de profundidad con el valor *Z0(x1'', y1'')* del mapa de profundidad. Si esta comparación no produce una coincidencia cercana, entonces se puede determinar un vector de compensación de movimiento más preciso utilizando procedimientos de búsqueda convencionales. El valor *p1'' - p1* se puede usar como un valor inicial para estos procedimientos de búsqueda de vectores de movimiento convencionales. En los procedimientos de compresión de vídeo estándar, los vectores de movimiento generalmente no se calculan para cada píxel por separado, sino para bloques de tamaño fijo de, por ejemplo, 8 x 8 o 16 x 16 píxeles. Debido a esto, el procedimiento de proyección de la presente invención también se aplica preferiblemente a tales bloques para ahorrar esfuerzo informático. Esto puede hacerse, por ejemplo, promediando los valores de profundidad en el interior del bloque, y utilizando el valor de profundidad media y las coordenadas *x* e *y* del centro del bloque como las coordenadas del punto en la proyección.

50 Cada una de las figuras 5, 6 y 7 presenta los campos de vectores de movimiento calculados con el procedimiento de proyección presentada para los fotogramas de vídeo de un juego de ordenador conocido. En este juego de ordenador, la cámara virtual está unida al arma y se puede mover y rotar a voluntad de los jugadores. Las figuras presentan cuadros de vídeo del juego divididos en bloques que se usan en el proceso de estimación de movimiento. La figura 5-7 presenta los cuadros de vídeo para el juego, los cuadros de vídeo se dividen en bloques.

55 En la figura 5 la cámara está estacionaria y los vectores de movimiento calculados usando el procedimiento de proyección por lo tanto no están presentes. La figura 6 muestra una situación en la que el jugador avanza, y la figura 7 muestra al jugador girando de derecha a izquierda. Como puede verse en las figuras 5 - 7, los vectores de

movimiento generados usando el procedimiento de proyección descrito brindan una visión bastante intuitiva del movimiento.

En lugar de ser utilizados como datos de entrada para la compresión de vídeo basada en bloques, los vectores de movimiento calculados usando el procedimiento de proyección descrito pueden, alternativamente, ser utilizados como un campo de deformación de la imagen de referencia. Normalmente, los vectores de movimiento para cada píxel en la imagen de referencia se calculan mediante interpolación (por ejemplo, bilineal) del campo vectorial basado en bloques. Alternativamente, es posible generar el campo del vector de movimiento proyectando los píxeles píxel por píxel. Por supuesto, utilizar este procedimiento requiere la transmisión/almacenamiento de información que describa el campo del vector de movimiento junto con los datos comprimidos. Si la distorsión de la imagen de referencia se realiza tanto en la compresión de la imagen como en la descompresión de la imagen, la parte del algoritmo de compresión de la imagen real es corregir la diferencia entre una imagen de referencia deformada y la imagen siguiente.

El uso del procedimiento de deformación es beneficioso para la relación de compresión y la calidad del vídeo, ya que aumenta el conjunto de operaciones que se pueden expresar con los vectores de movimiento. Cuando en la compresión de vídeo basada en bloque normal, la única operación que se puede usar para construir una imagen siguiente a partir de la imagen de referencia es una copia de bloque traducida, el procedimiento de deformación reproduce correctamente el escalado y la rotación de la imagen como resultado del movimiento de la cámara. Encontrar los parámetros de rotación y escalado con procedimientos de búsqueda iterativos convencionales haría que la compresión de vídeo sea prohibitiva desde el punto de vista informático, ya que agrega efectivamente dos nuevas dimensiones al espacio de búsqueda.

La figura 8 ilustra, como dos ejemplos, diagramas de flujo de los procedimientos de compresión de vídeo 800 y 820 de acuerdo con dos realizaciones de la invención. En la etapa 801, se recibe una imagen siguiente y se establece información sobre los parámetros de la cámara de la imagen siguiente. En la etapa 802 se establece información sobre el mapa de profundidad de la siguiente imagen. En la etapa 803, los bloques de las siguientes imágenes se proyectan al espacio tridimensional utilizando los parámetros de la cámara y el mapa de profundidad de la siguiente imagen. En la etapa 804 se establece información sobre los parámetros de la cámara de la imagen de referencia, y en la etapa 805 los puntos proyectados en el espacio tridimensional se proyectan en la superficie de la imagen de referencia.

En la etapa 806 los vectores de movimiento se determinan usando los resultados de la proyección, como se describe anteriormente. En la etapa 807 se determina si los vectores de movimiento relacionados con los parámetros de la cámara son lo suficientemente precisos. Esto se puede hacer, como se mencionó anteriormente, comparando el valor $Z_0(x_1, y_1)$ del mapa de profundidad de la imagen de referencia con el valor z_1 de profundidad del punto p_1 de la imagen proyectada dos veces. Normalmente, se determina un valor de umbral para la diferencia, y si los valores del mapa de profundidad son lo suficientemente cercanos entre sí, no hay necesidad de vectores de movimiento más precisos. En la etapa 808, si es necesario, se buscan vectores de movimiento más precisos utilizando los parámetros de movimiento relacionados con los parámetros de la cámara como valores iniciales; aquí cualquier procedimiento es aplicable para buscar vectores de compensación de movimiento. La búsqueda se realiza preferiblemente solo en aquellas áreas de la imagen, donde los vectores de movimiento no eran lo suficientemente precisos.

La figura 8b ilustra cómo continúa el procedimiento 800 de compresión de vídeo. En este procedimiento, los vectores de movimiento se utilizan de manera convencional en la compresión de imágenes. En la etapa 809 se determina una estimación de movimiento compensado para la siguiente imagen usando los vectores de movimiento. A partir de entonces, en la etapa 810, se determina una imagen de diferencia utilizando la siguiente imagen y la estimación de movimiento compensado. En la etapa 811, la información que describe los vectores de movimiento se transmite y/o almacena para su uso posterior. En la etapa 812, la información que describe la diferencia de imagen se transmite/almacena. En la etapa 813 se genera una nueva imagen de referencia utilizando la información transmitida/almacenada; se puede generar una nueva imagen de referencia en relación con cada imagen comprimida o se puede usar una misma imagen de referencia para varias imágenes comprimidas. A partir de entonces, el procedimiento 800 continúa desde la etapa 801.

La figura 8c ilustra cómo continúa el procedimiento 820. En el procedimiento 820, un campo de vector de movimiento se determina en la etapa 821. Como se discutió anteriormente, una forma posible de determinar un campo de vector de movimiento es interpolar valores específicos de píxel de vectores de movimiento específicos de bloque. En la etapa 822 se determina una estimación de movimiento de la próxima imagen utilizando el campo de vector de movimiento como campo de deformación. A partir de entonces, el procedimiento 820 continúa de manera similar al procedimiento 800.

En el borde de la imagen, donde el nuevo material entra normalmente a la vista, la estimación de movimiento compensado es incorrecta. Esta situación se puede observar, cuando un vector de movimiento calculado de acuerdo con el procedimiento presentado señala fuera de la imagen de referencia. La diferencia de imagen, sin embargo, se ocupa de la inexactitud en la estimación de movimiento compensado también en este caso.

5 La figura 9 ilustra esquemáticamente, como ejemplo, un diagrama de bloques de un compresor 900 de vídeo de acuerdo con una realización de la invención. El compresor de vídeo tiene medios 901 para almacenar una imagen de referencia, medios 902 para recibir una imagen siguiente y medios 903 para determinar vectores de movimiento para estimar cambios entre la imagen siguiente y la imagen de referencia. Además, comprende medios 904 para establecer información sobre un mapa de profundidad y parámetros de cámara de la siguiente imagen y medios 905 para almacenar parámetros de cámara de la imagen de referencia. Los medios 904 pueden estar dispuestos, por ejemplo, para recibir información de distancia de un sensor de rango o para recibir/obtener información de una memoria de un controlador de pantalla 3D. Los medios 903 para determinar las estimaciones de vectores de movimiento están dispuestos para calcular la proyección de puntos de la siguiente imagen a un espacio tridimensional y la proyección de puntos en un espacio tridimensional a la imagen de referencia.

10 Un compresor de vídeo puede comprender además medios para almacenar información sobre un mapa de profundidad de la imagen de referencia (puede ser implementado usando el mismo medio de memoria 905, que se utilizan en el almacenamiento de los parámetros de la cámara) y medios 906 para la estimación de la exactitud de los vectores de movimiento. Los medios para estimar la precisión de los vectores de movimiento pueden disponerse para determinar, cuando sea necesario, vectores de movimiento más precisos mediante procedimientos de búsqueda convencionales.

15 Un compresor de vídeo típicamente comprende además medios 907 para la construcción de una estimación de movimiento compensado para la siguiente imagen y medios 908 para determinar una imagen de diferencia entre la imagen siguiente y la estimación de movimiento compensado. Los medios 907 para construir una estimación de movimiento compensado pueden estar dispuestos para usar vectores de movimiento o, alternativamente, pueden estar dispuestos para construir un campo de vector de movimiento y para deformar una imagen de referencia de acuerdo con el campo de vector de movimiento.

20 Además, por lo general comprende medios 909 para transmitir/almacenar información que describe la imagen que la diferencia de los vectores de movimiento. Los medios 910 para construir una nueva imagen de referencia utilizando la información transmitida/almacenada también suelen ser parte de un compresor de vídeo.

25 Un programa de dispositivo o equipo según la invención puede estar dispuesto para funcionar de acuerdo con cualquier procedimiento de acuerdo con la invención.

30 El procedimiento descrito abre nuevas posibilidades para el uso de vídeo comprimido, porque el procesador y otros recursos, que normalmente son requeridos por la compresión de vídeo de compensación de movimiento convencional, causan requisitos prohibitivos que deben colocarse en el equipo de compresión de vídeo en tiempo real. La potencia de cómputo que de otro modo sería necesaria para comprimir secuencias de vídeo en tiempo real ahora se puede usar para mejorar la calidad del resultado, para comprimir múltiples secuencias de vídeo simultáneas o, cuando sea apropiado, para realizar otras tareas simultáneamente con la compresión de la imagen.

35 En vista de la descripción anterior, será evidente para una persona experta en la técnica que pueden realizarse diversas modificaciones dentro del ámbito de la invención. Aunque algunas realizaciones preferidas de la invención se han descrito en detalle, debería ser evidente que son posibles muchas modificaciones y variaciones de las mismas, estando todas las cuales dentro del alcance de la invención tal como se define en las reivindicaciones independientes adjuntas.

REIVINDICACIONES

1. Un procedimiento (800, 820) de compresión de información de vídeo, que comprende las etapas de:
 - 5 - proyectar (803) puntos de una imagen siguiente en puntos proyectados en un espacio tridimensional utilizando parámetros de cámara de la imagen siguiente y un mapa de profundidad de la siguiente imagen, en el que el mapa de profundidad se deriva de un mapa de profundidad que reside en una memoria de un controlador de pantalla de un ordenador,
 - proyectar (805) los puntos proyectados desde el espacio tridimensional a una superficie de imagen de referencia utilizando los parámetros de la cámara de la imagen de referencia, obteniendo así los vectores de movimiento para estimar los cambios entre la imagen siguiente y la imagen de referencia,
 - 10 - estimar (807) la precisión de los vectores de movimiento obtenidos comparando el valor de profundidad de un punto proyectado en el espacio tridimensional con el valor del mapa de profundidad en el punto proyectado en la superficie de la imagen de referencia, y
 - usar selectivamente los vectores de movimiento obtenidos como, por ejemplo, vectores de movimiento de compresión de imagen o usar (808) los vectores de movimiento obtenidos como valores iniciales para la búsqueda de vectores de movimiento de compresión de imagen, dependiendo de la precisión estimada de los vectores de movimiento obtenidos.
2. Un procedimiento según la reivindicación 1, en el que los vectores de movimiento obtenidos se aplican para acelerar la compresión.
- 20 3. Un procedimiento según cualquiera de las reivindicaciones anteriores, que comprende además la etapa de:
 - determinar (821) un campo de vector de movimiento para la siguiente imagen utilizando vectores de movimiento obtenidos y/o buscados, y
 - determinar (822) una estimación de movimiento compensado para la siguiente imagen utilizando el campo de vector de movimiento como campo de distorsión de la imagen de referencia.
- 25 4. Un procedimiento según cualquiera de las reivindicaciones anteriores, en el que los vectores de movimiento se determinan para bloques de imágenes de tamaño fijo y en que para cada bloque solo se determina un vector de movimiento.
5. Un procedimiento según la reivindicación 4, en el que para cada bloque de imagen se determina una profundidad promedio utilizando el mapa de profundidad y en el que el punto central de un bloque es el punto proyectado.
- 30 6. Un programa informático que comprende medios de código de programa informático adaptado para realizar todas las etapas de la reivindicación 1 cuando dicho programa se ejecuta en un ordenador.
7. Un programa informático según la reivindicación 6, realizado en un medio legible por ordenador.
8. Un dispositivo para la compresión de información de vídeo, comprendiendo el dispositivo:
 - 35 - medios para proyectar puntos de una imagen siguiente a puntos proyectados en un espacio tridimensional utilizando los parámetros de la cámara de la imagen siguiente y un mapa de profundidad de la imagen siguiente, en el que el mapa de profundidad se deriva de un mapa de profundidad que reside en la memoria de un controlador de pantalla,
 - medios para proyectar los puntos proyectados desde el espacio tridimensional a la superficie de la imagen de referencia utilizando los parámetros de la cámara de la imagen de referencia, obteniendo así vectores de movimiento para estimar los cambios entre la imagen siguiente y la imagen de referencia,
 - 40 - medios para estimar la precisión de los vectores de movimiento obtenidos, dispuestos para comparar el valor de profundidad de un punto proyectado en el espacio tridimensional con el valor del mapa de profundidad en el punto proyectado en la superficie de la imagen de referencia, y
 - medios para usar los vectores de movimiento obtenidos, dispuestos para usar selectivamente los vectores de movimiento obtenidos como, por ejemplo, vectores de movimiento de compresión de imagen o usar los vectores de movimiento obtenidos como valores iniciales para la búsqueda de vectores de movimiento de compresión de imagen, dependiendo de la precisión estimada de los vectores de movimiento obtenidos.
 - 45

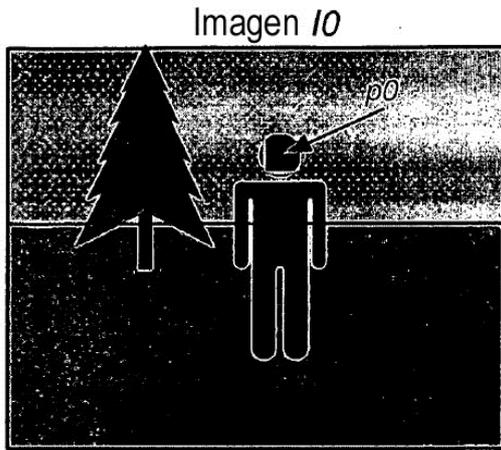


Fig. 1a

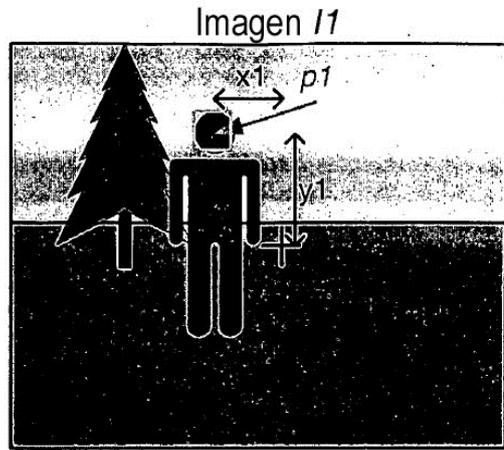


Fig. 2a



Fig. 1b



Fig. 2b

Posición y dirección de la cámara en imágenes *10* ja *11*

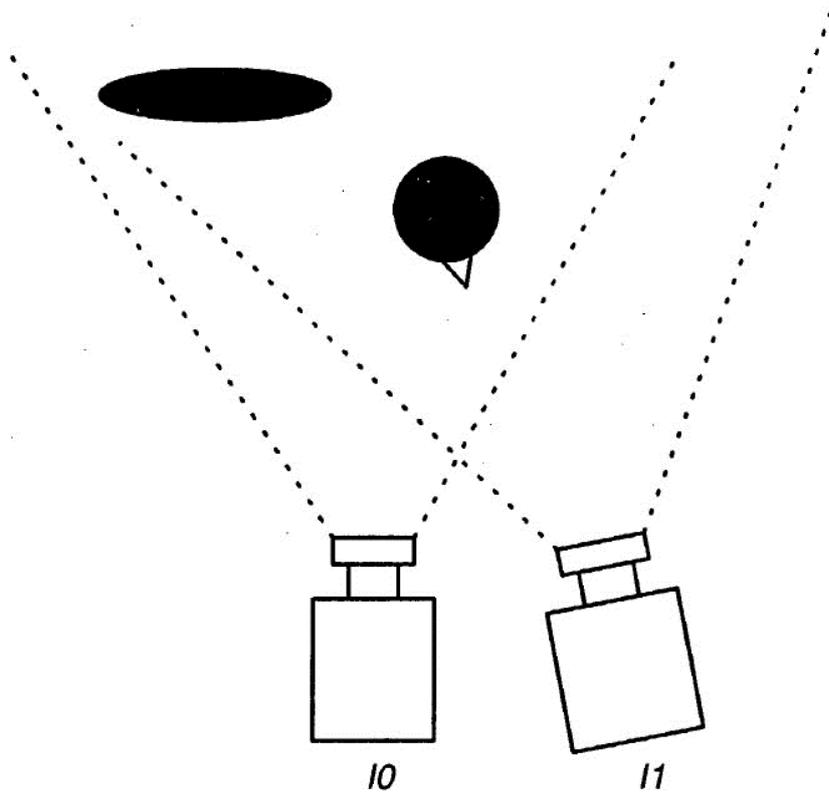


Fig. 3

Proyección de un punto desde la superficie de la imagen al espacio tridimensional

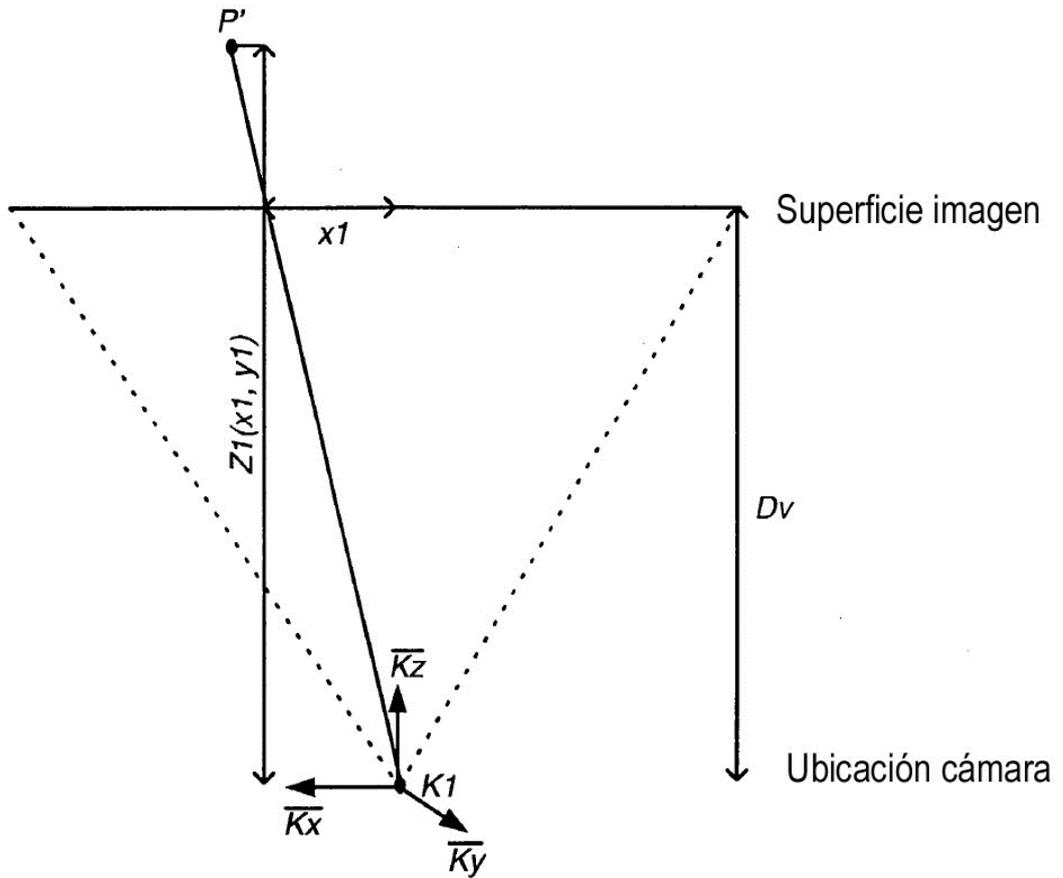


Fig. 4

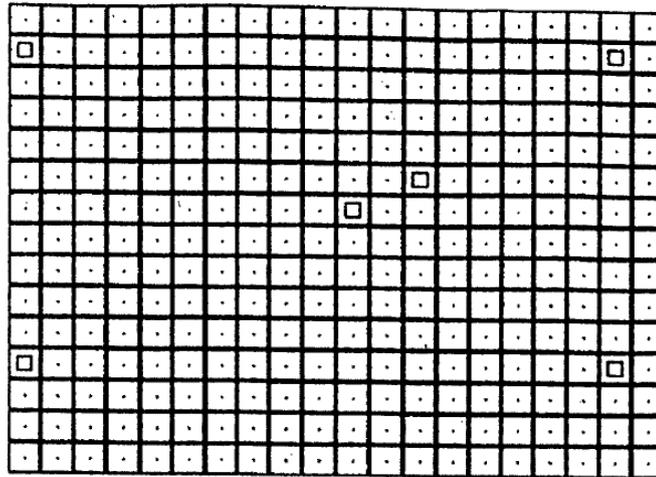


Fig. 5

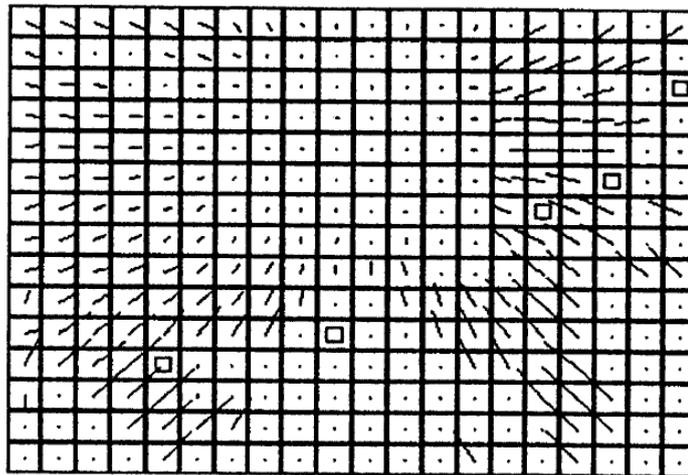


Fig. 6

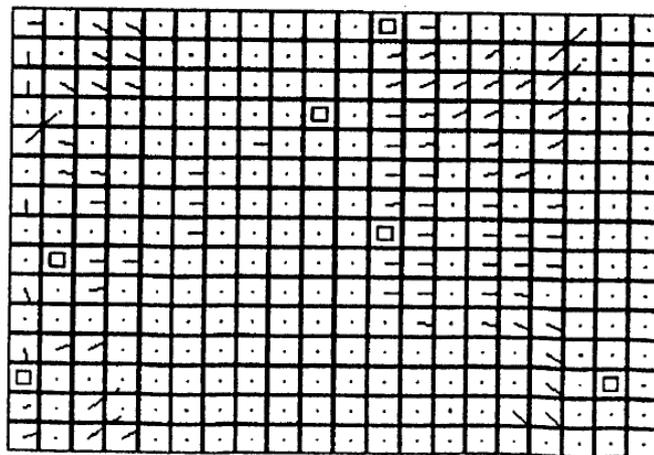


Fig. 7

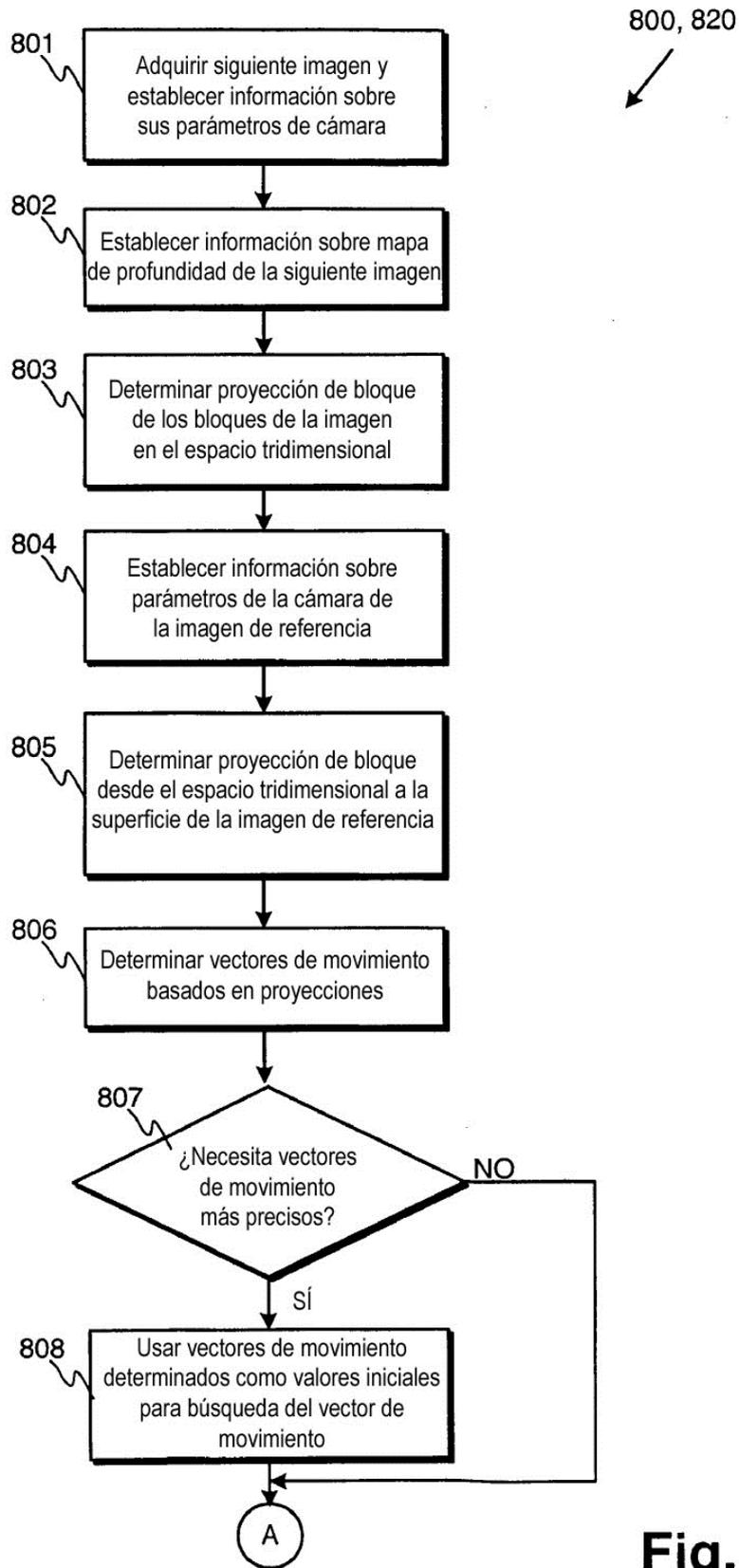


Fig. 8a

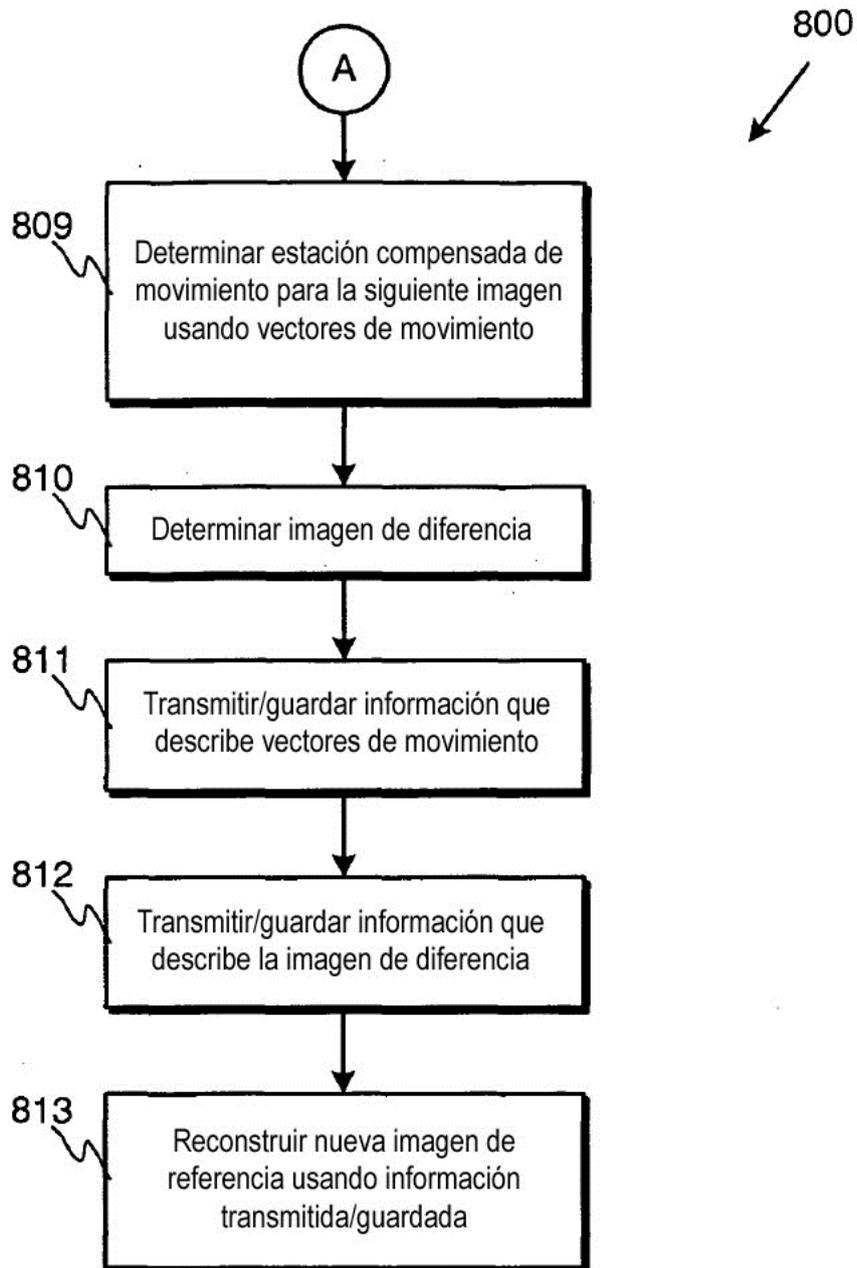


Fig. 8b

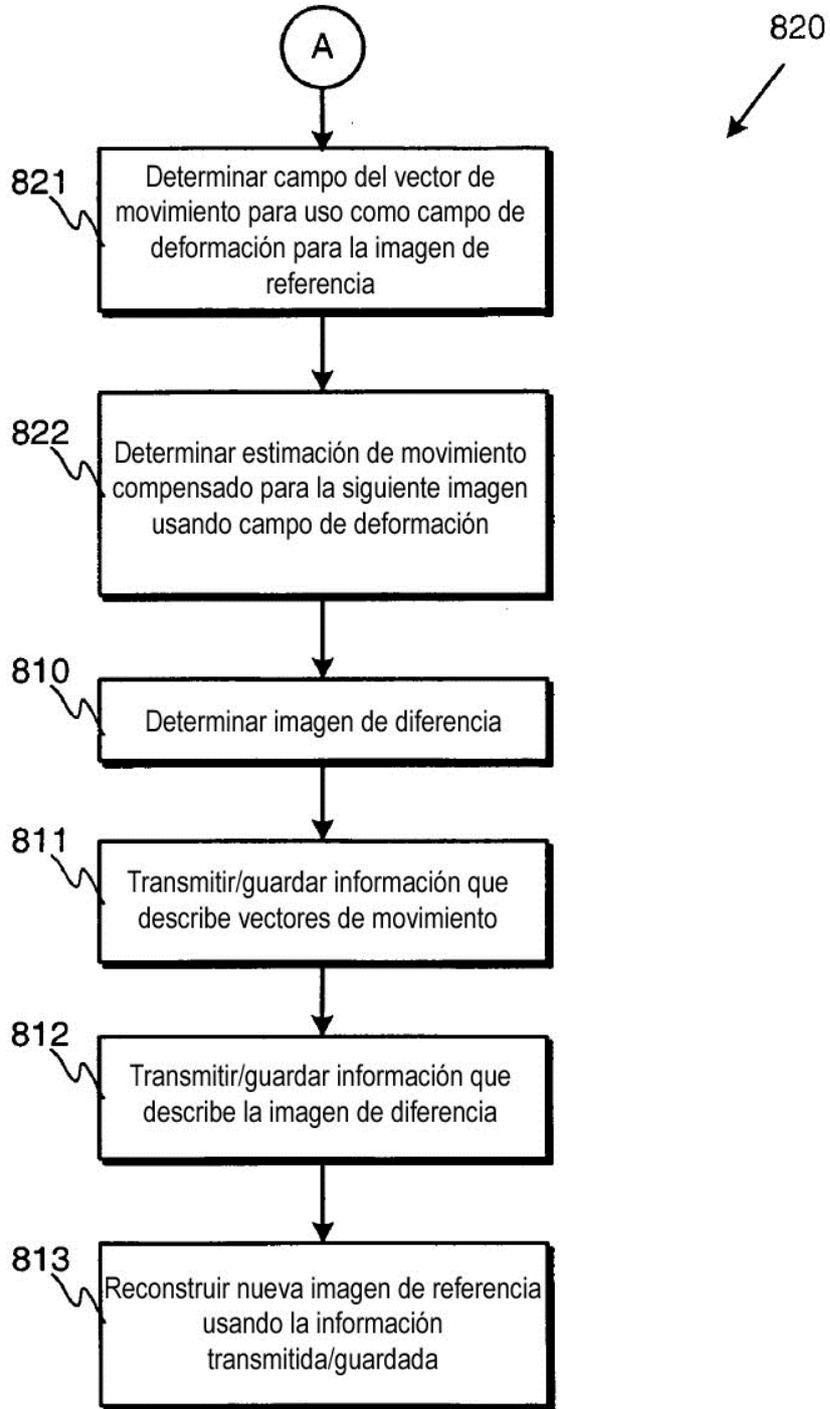


Fig. 8c

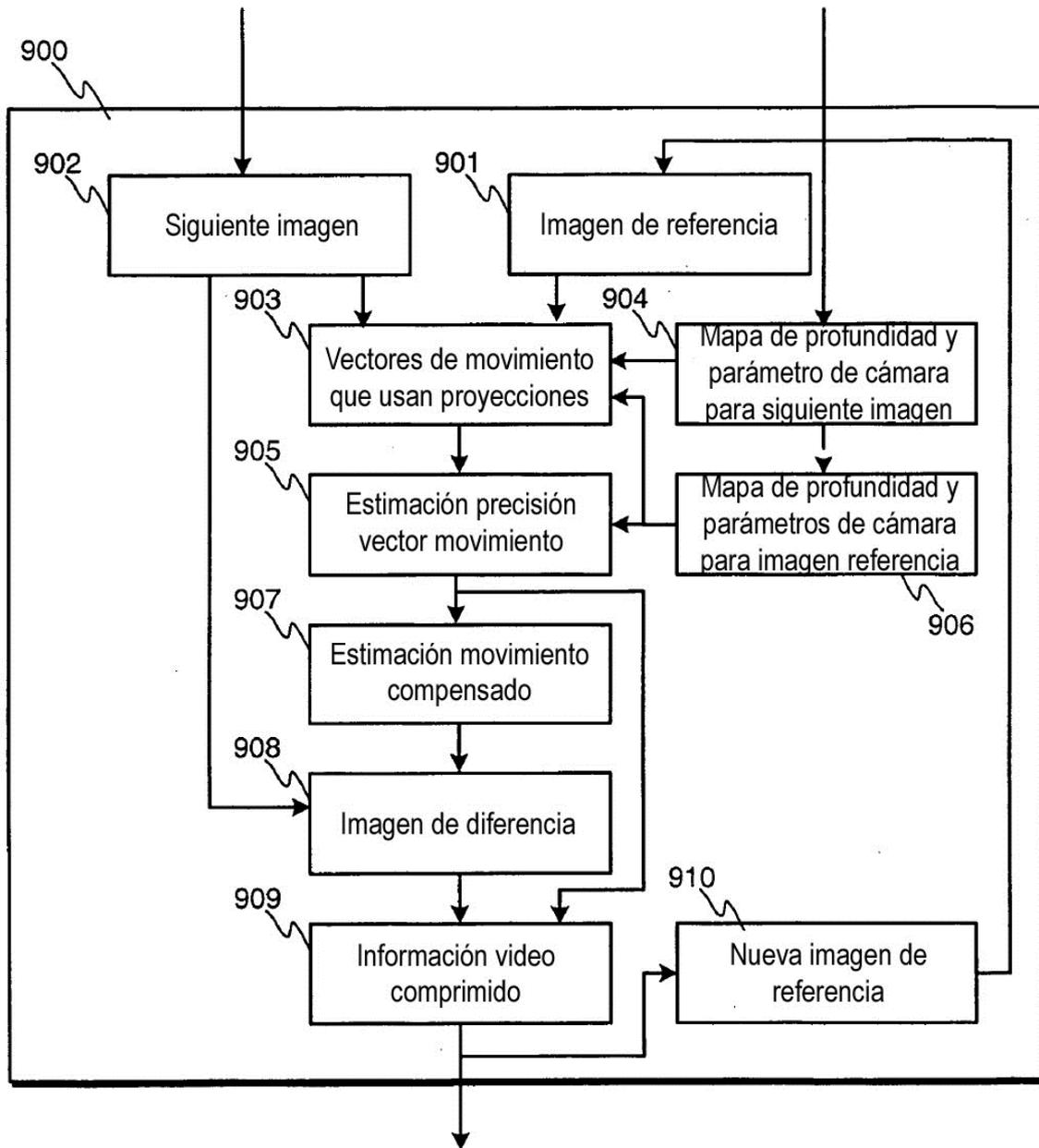


Fig. 9