

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 700 506**

51 Int. Cl.:

G06T 7/246 (2007.01)

G06T 5/50 (2006.01)

H04N 5/232 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.07.2013 PCT/CN2013/079852**

87 Fecha y número de publicación internacional: **29.01.2015 WO15010244**

96 Fecha de presentación y número de la solicitud europea: **23.07.2013 E 13889978 (6)**

97 Fecha y número de publicación de la concesión europea: **05.09.2018 EP 3025491**

54 Título: **Suavización de trayectoria adaptativa para estabilización de vídeo**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
18.02.2019

73 Titular/es:
**MICROSOFT TECHNOLOGY LICENSING, LLC
(100.0%)
One Microsoft Way
Redmond, WA 98052, US**

72 Inventor/es:
**YUAN, LU y
SUN, JIAN**

74 Agente/Representante:
ELZABURU, S.L.P

ES 2 700 506 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

Suavización de trayectoria adaptativa para estabilización de vídeo

Antecedentes

5 Un vídeo captado con un dispositivo de mano (por ejemplo, un teléfono celular o una cámara de grabación portátil) a menudo aparece tembloroso e inestable. La calidad del vídeo ha llegado a ser significativamente importante, en parte debido a la ubicua presencia y disponibilidad de los dispositivos móviles (por ejemplo, teléfonos móviles, tabletas, cámaras de grabación y otros dispositivos) capaces de captar vídeo.

10 La estabilización de vídeo digital es capaz de mejorar la calidad de vídeo de semejante vídeo, pero numerosos tipos de estabilización de vídeo digital presentan desventajas diversas. Por ejemplo, la estabilización de vídeo digital puede eliminar el temblor de un vídeo, pero a costa de introducir efectos espurios de vídeo, tales como distorsión de imagen, recorte, pérdida de resolución y otros efectos. De esta forma, existe la necesidad de una técnica de estabilización de vídeo digital que pueda mejorar la calidad del vídeo sin introducir un número excesivo de efectos espurios indeseables.

15 La divulgación "*Video stabilization based on a 3D perspective camera model*" (Estabilización de vídeo basada en un modelo de cámara de perspectiva 3D), por Guofeng Zhang et al., divulga una estabilización de vídeo basada en un modelo de cámara de perspectiva 3D. Se da forma al problema de la estabilización como una función de coste cuadrática en las restricciones de suavidad y similitud, la cual permite que la suavidad sea controlada resolviendo un sistema de pocas ecuaciones lineales. Al aprovecharse de su escaso número, el procedimiento de optimización es muy eficiente. En lugar de recuperar profundidades densas, se utiliza una representación de geometría aproximada y se analizan los errores de deformación resultantes. Limitando apropiadamente el error de deformación, pueden conseguirse resultados visualmente apreciables, incluso utilizando estructuras planas.

20

25 La divulgación "*Auto-directed video stabilization with robust L1 optimal camera paths*" (Estabilización de vídeo autodirigida con trayectorias de cámara óptimas L1 robustas), por Matthias Grundmann et al., divulga un método de estabilización de vídeo que utiliza modelos de movimiento 2D. El método se basa en el seguimiento de rasgos al objeto de determinar una trayectoria de cámara $C(t)$ basándose en la estimación de poses de cámara utilizando diversos modelos de movimiento paramétricos, los cuales incluyen homografías. Se estima una trayectoria de cámara suavizada $P(t)$ basándose en la trayectoria inicial $C(t)$, utilizando la minimización de una función de coste que implica derivadas de $P(t)$.

Compendio

30 Esta divulgación describe, en parte, técnicas y arquitecturas para estabilización de vídeo que pueden ser utilizadas para transformar un vídeo de aspecto tembloroso en un vídeo de aspecto estático. La estabilización de vídeo puede incluir un procedimiento de suavización de trayectoria de cámara que genera una trayectoria de cámara suavizada a partir de una trayectoria de cámara temblorosa original. Utilizando un núcleo de suavización relativamente grande, que comprende un cierto número de tramas de vídeo, un procedimiento de suavización de trayectoria puede eliminar tanto los temblores de alta frecuencia como los rebotes o tumbos de baja frecuencia, y es capaz de preservar los movimientos discontinuos de la cámara (tales como transiciones de encuadre o escena relativamente rápidas), a fin de evitar un recorte o distorsión de geometría excesivos en la trayectoria de cámara suavizada. Puede llevarse a cabo un procedimiento de suavización de trayectoria en una implementación basada en ventana deslizante, que puede ser utilizada para una estabilización en tiempo real.

35

40 Este compendio se ha proporcionado para introducir una selección de conceptos de una forma simplificada, que se describen adicionalmente más adelante en la Descripción detallada. No es la intención que este compendio identifique las características clave o esenciales de la materia objeto reivindicada, ni tampoco se pretende que se utilice como ayuda a la hora de determinar el alcance de la materia objeto reivindicada. En lugar de ello, la invención proporciona un sistema, un método y un medio de almacenamiento legible por computadora, tal como se reivindican más adelante en esta memoria.

45 El término 'técnicas' puede referirse a un sistema(s), método(s), instrucciones legibles por computadora, módulo(s), algoritmos, lógica de hardware (por ejemplo, matrices de puertas programables por efecto de campo (FPGAs –"Field-programmable Gate Arrays"–), circuitos integrados específicos de aplicación (ASICs –"Application-specific Integrated Circuits"–), productos estándar específicos de aplicación (ASSPs –"Application-specific Standard Products"–), sistemas consistentes en sistema sobre chip (SOCs –"System-on-a-chip"–), dispositivos lógicos programables complejos (CPLDs –"Complex Programmable Logic Devices"–)), según lo permita el contexto anterior y todo lo que resta del documento.

50

Breve descripción de los dibujos

55 La descripción detallada se realiza con referencia a las figuras que se acompañan. En las figuras, el (los) dígito(s) más a la izquierda de un número de referencia identifica(n) la figura en la que aparece primeramente el número de referencia. Los mismos números de referencia en diferentes figuras indican elementos similares o idénticos.

La Figura 1 es un diagrama de bloques que representa un entorno proporcionado a modo de ejemplo en el que pueden implementarse las técnicas descritas en la presente memoria.

5 La Figura 2 muestra una representación de una estabilización de vídeo por suavización de la trayectoria de la cámara y suavización de la curva de postratamiento, de acuerdo con diversas realizaciones proporcionadas a modo de ejemplo.

La Figura 3 es un diagrama esquemático que muestra dos tramas de vídeo de una trayectoria de cámara, de acuerdo con diversas realizaciones proporcionadas a modo de ejemplo.

La Figura 4 incluye diagramas esquemáticos que ilustran el recorte y distorsión de trama, de acuerdo con diversas realizaciones proporcionadas a modo de ejemplo.

10 La Figura 5 es un diagrama esquemático que muestra relaciones entre trayectorias de cámara originales y trayectorias de cámara suavizadas, de acuerdo con diversas realizaciones proporcionadas a modo de ejemplo.

La Figura 6 es un diagrama de bloques de un dispositivo de estabilización de vídeo de acuerdo con diversas realizaciones proporcionadas a modo de ejemplo.

15 La Figura 7 es un diagrama esquemático de dos secuencias de vídeo de tramas de vídeo, de acuerdo con diversas realizaciones proporcionadas a modo de ejemplo.

La Figura 8 es un diagrama de bloques de una parte de un dispositivo de estabilización de vídeo de acuerdo con diversas realizaciones proporcionadas a modo de ejemplo.

La Figura 9 es un diagrama de flujo de un procedimiento proporcionado a modo de ejemplo para la suavización de vídeo.

20 Descripción detallada

Generalidades

En diversas realizaciones, las técnicas y dispositivos para la estabilización de vídeo incluyen la suavización de la trayectoria para transformar un vídeo tembloroso en un vídeo de aspecto estático. Un procedimiento de suavización puede generar una trayectoria de cámara modificada a partir de una trayectoria de cámara originalmente temblorosa o errática, correspondiente a un vídeo captado por una cámara de mano, por ejemplo. Un procedimiento de suavización de trayectoria que incluye una implementación basada en ventana deslizante puede eliminar tanto los temblores de alta frecuencia (por ejemplo, ocasionados por la mano temblorosa de un usuario de la cámara) como los rebotes o tumbos de baja frecuencia (por ejemplo, ocasionados por un movimiento al andar o balancearse por parte del usuario) de una trayectoria de la cámara, y es capaz de preservar movimientos de cámara discontinuos (tales como transiciones de encuadre o escena rápidas) con el fin de evitar un recorte o distorsión de la geometría excesivos. Tales técnicas o dispositivos pueden ser utilizados para la estabilización de vídeo en tiempo o real o fuera de conexión.

En algunas realizaciones, un sistema para estabilización de vídeo comprende una puerta de entrada de vídeo para recibir un vídeo que comprende una secuencia de tramas de vídeo. Un módulo de fragmentación de vídeo, conectado comunicativamente con la puerta de entrada de vídeo, fragmenta el vídeo en un cierto número de secuencias de tramas de vídeo. El módulo de fragmentación de vídeo puede ser implementado mediante hardware, software, *firmware*, o software instalado de forma permanente en hardware, o una combinación de los mismos. El módulo de fragmentación de vídeo puede aplicar retardos de trama a secuencias individuales de las secuencias de tramas de vídeo. En consecuencia, secuencias vecinas pueden tener tramas de vídeo duplicadas. En particular, secuencias consecutivas pueden ser idénticas entre sí, a excepción de que una de las secuencias esté retardada en trama con respecto a la otra secuencia. Como parte de la implementación basada en ventana deslizante, el módulo de partición de vídeo puede proporcionar las secuencias individuales retardadas en trama a procesadores de almacenamiento temporal de vídeo individuales que lleven a cabo la suavización de vídeo. En particular, un procesador de almacenamiento temporal de vídeo individual puede incluir un módulo de extracción de rasgos, un módulo de estimación de homografía y un módulo de suavización de trayectoria adaptativa.

45 En algunas implementaciones, cualquiera de entre un procesador de almacenamiento temporal de vídeo, un módulo de extracción de rasgos, un módulo de estimación de homografía y un módulo de suavización de trayectoria adaptativa puede comprender hardware, software, *firmware* o una combinación de los mismos. Por ejemplo, un procesador de almacenamiento temporal de vídeo implementado por hardware es capaz de ejecutar selectivamente software que comprende el módulo de extracción de rasgos, el módulo de estimación de homografía y el módulo de suavización de trayectoria adaptativa. El procesador de almacenamiento temporal de vídeo puede incluir, o puede acceder a, memoria que comprende una pluralidad de registros de almacenamiento temporal destinados a almacenar secuencias de tramas de vídeo. En otro ejemplo, un procesador de almacenamiento temporal de vídeo que comprende código ejecutable, puede, selectivamente, hacer funcionar un módulo de extracción de rasgos implementado en hardware, un módulo de estimación de homografía implementado en hardware y un módulo de suavización de trayectoria adaptativa implementado en hardware. El procesador de almacenamiento temporal de vídeo puede incluir, o puede

acceder a, memoria que comprende una pluralidad de registros de almacenamiento temporal para almacenar secuencias de tramas de vídeo.

El módulo de extracción de rasgos identifica rasgos objeto, o buscados, en tramas de vídeo individuales de una secuencia de las tramas de vídeo. Por ejemplo, tales rasgos objeto pueden comprender puntos, bordes u otros objetos afines de las tramas de vídeo individuales.

Un módulo de estimación de homografía lleva a cabo una estimación homográfica entre, o de entre, tramas de vídeo de la secuencia de tramas de vídeo, a fin de generar una secuencia modificada de tramas de vídeo. La estimación homográfica está basada, al menos en parte, en los rasgos objeto identificados. En algunas implementaciones, la estimación homográfica se lleva a cabo entre tramas de vídeo consecutivas de la secuencia de tramas de vídeo. Un módulo de suavización de trayectoria adaptativa determina errores de estimación de entre tramas de vídeo de la secuencia modificada de tramas de vídeo, y aplica suavización de trayectoria adaptativa a la secuencia modificada de tramas de vídeo para generar una secuencia suavizada de tramas de vídeo. La suavización de trayectoria adaptativa puede estar basada, al menos en parte, en cambios entre las tramas de vídeo individuales de la secuencia de tramas de vídeo, y en los errores de estimación.

En algunas realizaciones, un sistema para estabilización de vídeo puede comprender, adicionalmente, un módulo de agregación de vídeo configurado para combinar la secuencia suavizada de tramas de vídeo con otras secuencias suavizadas de tramas de vídeo, a fin de producir una porción de vídeo agregada, y para aplicar un filtrado de postratamiento para suavizar la porción de vídeo agregada. Estas otras secuencias suavizadas son productos de respectivos procesadores de almacenamiento temporal de vídeo individuales. En diversas implementaciones, las otras secuencias suavizadas de tramas de vídeo procedentes de procesadores de almacenamiento temporal de vídeo individuales, están basadas, al menos en parte, en secuencias de tramas de vídeo desplazadas en trama que están respectivamente descentradas con respecto a la secuencia de tramas de vídeo en un múltiplo entero de n tramas, donde n es un número predeterminado.

Diversas realizaciones se describen adicionalmente con referencia a las Figuras 1-9.

Ejemplo de entorno

El entorno que se describe en lo que sigue no constituye sino un ejemplo y no es la intención que las reivindicaciones estén limitadas por ningún entorno operativo particular. Pueden utilizarse otros entornos sin apartarse del alcance de la materia objeto reivindicada. La Figura 1 muestra un entorno 100 proporcionado a modo de ejemplo, en el que pueden funcionar realizaciones que implican estabilización de vídeo según se describe en la presente memoria. En algunas realizaciones, los diversos dispositivos y/o componentes del entorno 100 incluyen una variedad de dispositivos informáticos 102. En diversas realizaciones, los dispositivos informáticos 102 incluyen los dispositivos 102a-102e. Si bien se ilustran como una variedad diversa de tipos de dispositivo, los dispositivos informáticos 102 pueden ser otros tipos de dispositivo y no están limitados a los tipos de dispositivo que se han ilustrado. Los dispositivos informáticos 102 pueden comprender cualquier tipo de dispositivo con uno o múltiples procesadores 104 conectados operativamente a una interfaz de entrada / salida 106 y a una memoria 108, por ejemplo, a través de un bus 110. Los dispositivos informáticos 102 pueden incluir computadoras personales tales como, por ejemplo, computadoras de sobremesa 102a, computadoras portátiles 102b, computadoras de tableta 102c, dispositivos de telecomunicación 102d, asistentes personales digitales (PDAs –“personal digital assistants”–) 102e, lectores de libros electrónicos, computadoras portables o de mano, computadoras de automóvil, dispositivos de juego, etc. Los dispositivos informáticos 102 pueden también incluir dispositivos orientados a los negocios o a las ventas minoristas, tales como, por ejemplo, computadoras de servidor, clientes delgados, terminales y/o estaciones de trabajo. En algunas realizaciones, los dispositivos informáticos 102 pueden incluir, por ejemplo, componentes para integración en un dispositivo informático, accesorios y otras clases de dispositivos. En algunas realizaciones, algunas o todas las capacidades funcionales descritas como llevadas a cabo por los dispositivos informáticos 102 pueden ser implementadas por uno o más dispositivos informáticos en disposición pareja, por un servidor o servidores distantes, o por un recurso informático radicado en nube. Por ejemplo, los dispositivos informáticos 102 pueden llevar a cabo aplicaciones que están almacenadas a distancia con respecto a los dispositivos informáticos.

En algunas realizaciones, tal como se ha mostrado en relación con el dispositivo 102c, la memoria 108 puede almacenar instrucciones ejecutables por el (los) procesador(es) 104, los cuales incluyen un sistema operativo (OS –“operating system”–) 112, un módulo de gráficos 114 y programas o aplicaciones 116 que son susceptibles de ser cargados y ejecutados por el (los) procesador(es) 104. Los uno o más procesadores 104 pueden incluir unidades centrales de procesamiento (CPUs –“central processing units”–), unidades de tratamiento de gráficos (GPUs –“graphics processing units” –) procesadores de almacenamiento temporal de vídeo, y otros elementos similares. En algunas implementaciones, un módulo de fragmentación de vídeo 120 comprende código ejecutable almacenado en la memoria 108 y que es ejecutable por el (los) procesador(es) 104 y/o los procesadores de almacenamiento temporal de vídeo 118. Un módulo de suavización de trayectoria adaptativa 124 comprende código ejecutable almacenado en la memoria 108 y que es ejecutable por el (los) procesador(es) 104. El módulo de suavización de trayectoria adaptativa 124 determina errores de estimación entre las tramas de vídeo de una secuencia de tramas de vídeo modificada, y aplica suavización de trayectoria adaptativa a la secuencia de tramas de vídeo modificada, a fin de generar una secuencia de tramas de vídeo modificada. La suavización de trayectoria adaptativa está basada, al menos en parte,

en cambios entre tramas de vídeo individuales de una secuencia de tramas de vídeo modificada.

Si bien se han descrito ciertos módulos de manera que llevan a cabo operaciones diversas, los módulos constituyen meramente un ejemplo y las mismas o similares capacidades funcionales pueden ser llevadas a cabo por un número mayor o menor de módulos. Es más, no es necesario que las funciones llevadas a cabo por los módulos representados sean llevadas a cabo localmente por un único dispositivo. En lugar de ello, algunas operaciones pueden ser llevadas a cabo por un dispositivo distante (por ejemplo, un dispositivo en disposición pareja, un servidor, una nube, etc.).

Alternativamente, o además de ello, la capacidad funcional descrita en la presente memoria puede ser llevada a cabo, al menos en parte, por uno o más componentes lógicos de hardware. Por ejemplo, y sin limitación, tipos ilustrativos de componentes lógicos de hardware que pueden ser utilizados incluyen matrices de puertas programables por efecto de campo (FPGAs), circuitos integrados específicos de programa (ASICs), productos estándar específicos de programa (ASSPs), sistemas consistentes en sistema sobre chip (SOCs), dispositivos lógicos programables complejos (CPLDs), etc.

En algunas realizaciones, tal como se muestra con respecto al dispositivo 102d, un dispositivo informático puede estar asociado con una cámara 126 capaz de captar vídeo. Por ejemplo, un dispositivo de mano puede incluir tales cámara y dispositivo de computación 102. La memoria 108 puede incluir un medio legible por computadora o una combinación de estos. El medio legible por computadora puede incluir un medio de almacenamiento informático y/o un medio de comunicación. El medio de almacenamiento informático incluye medios volátiles y no volátiles, extraíbles y no extraíbles, implementados por cualquier método o en cualquier tecnología para el almacenamiento de información, tales como instrucciones legibles por computadora, estructuras de datos, módulos de programa, u otros datos. El medio de almacenamiento informático incluye memoria de cambio de fase (PRAM –“phase change memory”–), memoria de acceso aleatorio estática (SRAM –“static random-access memory”–), memoria de acceso aleatorio dinámica (DRAM –“dynamic random-access memory”–), otros tipos de memoria de acceso aleatorio (RAM), memoria de solo lectura (ROM –“read-only memory”–), memoria de solo lectura programable y borrrable eléctricamente (EEPROM –“electrically erasable programmable read-only memory”–), memoria flash u otra tecnología de memoria, memoria de solo lectura de disco compacto (CD-ROM –“compact disk ROM”–), discos versátiles digitales (DVD –“digital versatile disks”–) u otro almacenamiento óptico, casetes magnéticas, cinta magnética, almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio que no sea de transmisión y que pueda ser utilizado para almacenar información para su acceso por parte de un dispositivo informático, si bien no está limitado por estos.

En contraste, el medio de información puede incorporar instrucciones legibles por computadora, estructuras de datos, módulos de programa u otros datos en una señal de datos modulada, tal como una onda portadora, u otro mecanismo de transmisión. Tal como se define en esta memoria, el medio de almacenamiento informático no incluye medios de comunicación. En diversas realizaciones, la memoria 108 es un ejemplo de un medio de almacenamiento informático que almacena instrucciones ejecutables por computadora que, cuando son ejecutadas por un(os) procesador(es) 104 y/o procesador(es) de almacenamiento temporal de vídeo 118, configuran el (los) procesador(es) para, entre otras cosas, identificar rasgos objeto en tramas de vídeo individuales de una secuencia de tramas de vídeo, llevar a cabo una estimación homográfica entre, o de entre, tramas de vídeo de la secuencia de tramas de vídeo para generar una secuencia modificada de tramas de vídeo, de tal manera que la estimación homográfica está basada, al menos en parte, en los rasgos objeto identificados; determinar errores de estimación de entre tramas de vídeo de la secuencia modificada de tramas de vídeo; y aplicar suavización de trayectoria adaptativa a la secuencia modificada de tramas de vídeo con el fin de generar una secuencia suavizada de tramas de vídeo, de tal modo que la suavización de trayectoria adaptativa está basada, al menos en parte, en los cambios entre las tramas de vídeo individuales de la secuencia de las tramas de vídeo, y en los errores de estimación.

Un dispositivo de entrada puede incluir cualquiera de una variedad de dispositivos que están destinados a proporcionar, y/o implican, movimiento a un objeto presentado visualmente en un dispositivo de salida. Por ejemplo, en diversas realizaciones, un dispositivo de entrada puede ser un dispositivo de entrada táctil directa (por ejemplo, una pantalla táctil), un dispositivo táctil indirecto (por ejemplo, una almohadilla o placa táctil), un dispositivo de entrada indirecta (por ejemplo, un ratón, un teclado, una cámara o conjunto geoméricamente ordenado de cámaras, etc.), u otro tipo de dispositivo no táctil, tal como un dispositivo de entrada de audio.

El (los) dispositivo(s) informático(s) 102 puede(n) incluir una o más interfaces de entrada / salida (E/S –“I/O (input-output)”–) 106 para permitir al dispositivo informático 102 comunicarse con otros dispositivos. Las interfaces de entrada / salida (E/S) 106 pueden también incluir una o más interfaces de red para permitir las comunicaciones entre el dispositivo informático 102 y otros dispositivos conectados por red tales como otro(s) dispositivo(s) 102. Las interfaces de entrada / salida (E/S) 106 pueden permitir a un dispositivo 102 comunicarse con otros dispositivos tales como dispositivos periféricos de introducción por parte de un usuario (por ejemplo, un teclado, un teclado, un ratón, un lápiz, un controlador de juegos, un dispositivo de entrada de voz, un dispositivo de entrada táctil, un dispositivo de entrada por gestos y otros similares) y/o dispositivos periféricos de salida (por ejemplo, un dispositivo de presentación visual, una impresora, altavoces de audio, una salida háptica y otros similares).

La Figura 2 muestra una representación de la estabilización de vídeo por suavización de trayectoria de cámara y suavización de curva de postratamiento, de acuerdo con diversas realizaciones. La trayectoria 202 representa una

trayectoria de cámara original correspondiente al movimiento de una cámara durante la captación de vídeo. El movimiento de la cámara puede incluir movimientos rápidos (sacudidas) correspondientes a una mano inestable, no estática, de un usuario que sostiene la cámara. El movimiento de la cámara puede también incluir movimientos relativamente lentos (tambaleo) correspondientes al andar o al movimiento de oscilación del usuario.

5 Tal movimiento de una cámara puede ponerse de manifiesto como cambios de posición de los objetos de imagen en las tramas de vídeo. Más específicamente, el movimiento de la cámara durante el lapso de tiempo que va de la captación de una primera trama de un vídeo a la de una segunda trama de un vídeo, puede conducir a una traslación de las posiciones de los objetos en la segunda trama con respecto a la primera trama. El movimiento de la cámara puede también conducir a la rotación de los objetos en la segunda trama con respecto a la primera trama. El movimiento de la cámara hacia o en alejamiento del plano de las imágenes captadas puede cambiar la escala de los objetos de la primera trama a la segunda trama. De acuerdo con ello, puede determinarse la trayectoria de cámara original 202 considerando la traslación, la rotación y el cambio de escala entre, o de entre, una pluralidad de tramas de vídeo. Este último parámetro, cuando es relativamente pequeño, puede ser despreciado, de manera que solo se considera el movimiento en un único plano.

15 En diversas realizaciones, las imágenes de las tramas de vídeo de un vídeo original pueden ser modificadas de un modo tal, que un vídeo modificado que comprenda las tramas de vídeo modificadas aparecerá como relativamente suave y estacionario. Las tramas del vídeo original se modifican de tal manera que se eliminen de manera efectiva los cambios de imagen relativamente rápidos (por ejemplo, de contante de tiempos pequeña) entre tramas. Tales cambios rápidos de imagen corresponden a cambios rápidos en la trayectoria de la cámara, ejemplos de los cuales se han indicado por la referencia 204 en la Figura 2. Algunas técnicas para modificar tramas de vídeo incluyen suavizar porciones de un vídeo (por ejemplo, una secuencia de tramas), lo que conduce a una serie de segmentos de trayectoria de cámara suavizados 206. Como se explica en detalle más adelante, los segmentos de trayectoria de cámara suavizados 206 pueden superponerse unos con otros e incluir discontinuidades como consecuencia del procedimiento que generó los segmentos de trayectoria de cámara suavizados 206. De esta forma, como se muestra en la imagen de la derecha de la Figura 2, técnicas de filtrado postratamiento pueden generar una trayectoria suave y continua 208. La salida de vídeo para presentación visual puede estar basada, al menos en parte, en tal trayectoria continua 208. De acuerdo con ello, las sacudidas y el tambaleo en el vídeo original pueden reducirse o eliminarse en el vídeo de salida.

30 La Figura 3 es un diagrama esquemático que muestra dos tramas de vídeo $l(t-1)$ e $l(t)$ de una trayectoria 300 de cámara que incluye una serie de tramas $l(t)$, para $t = 0, 1, 2, \dots$, de acuerdo con diversas realizaciones proporcionadas a modo de ejemplo. La modelización y la definición de términos de esta manera es de utilidad para desarrollar técnicas para la suavización de vídeo, como se describe más adelante. Puede utilizarse homografía plana para representar el movimiento de cámara relativo entre tramas de vídeo consecutivas. Para un vídeo que comprende la secuencia de tramas de vídeo $l(t)$, para $t = 0, 1, 2, \dots$, puede computarse una homografía $F(t)$ entre cada par de tramas vecinas ($l(t-1)$, $l(t)$) a partir de sus puntos de rasgo correspondientes (tales como, por ejemplo, rasgos de imagen afines generados por KLT, SIFT o SURF). Se define la pose de cámara, $C(t)$, en una trama $l(t)$, como:

$$C(t) = C(t-1)F(t-1), \Rightarrow C(t) = F(0)F(1) \dots F(t-1). \quad (1)$$

$\{F(0), F(1), \dots, F(t-1)\}$ son homografías estimadas para cada trama a lo largo de la secuencia.

40 Dada una trayectoria de cámara original $\mathbf{C} = \{C(t)\}$, se determina una nueva trayectoria suavizada $\mathbf{P} = \{P(t)\}$. Condiciones deseables para \mathbf{P} son que \mathbf{P} ha de ser relativamente cercana a \mathbf{C} , al tiempo que está suavizada en todas partes (por ejemplo, sin discontinuidades), a excepción de en los cambios rápidos en \mathbf{C} . Por ejemplo, los cambios rápidos pueden ser intencionados, tal como cuando un usuario hace rotar rápidamente una cámara para calcular una imagen de un objeto que se mueve rápidamente. De esta forma, puede ser deseable mantener algunos cambios rápidos de trayectoria.

45 Un procedimiento de determinación de \mathbf{P} puede implicar minimizar la siguiente función de objetivo:

$$O(\{P(t)\}) = \sum_t (\|P(t) - C(t)\|^2 + \lambda_t \sum_{r \in \Omega_t} w_{t,r}(\mathbf{C}) \cdot \|P(t) - P(r)\|^2), \quad (2)$$

50 donde Ω_t es la vecindad de tramas de vídeo en la trama t . Por ejemplo, tal vecindad de tramas de vídeo puede incluir $l(t-30), l(t-29), \dots, l(t), l(t+1), \dots, l(t+30)$ (por ejemplo, 30 tramas antes y 30 tramas después de la trama t). Un núcleo Ω_t relativamente grande (por ejemplo, 60 tramas) puede suprimir tanto los temblores de alta frecuencia (por ejemplo, ocasionados por las sacudidas de la mano) como los rebotes o tumbos de baja frecuencia (por ejemplo, ocasionados por el movimiento al andar), si bien la materia objeto reivindicada no está limitada por esto.

55 En la Ecuación 2, el término $\|P(t) - C(t)\|^2$ es un término de datos que tiende a constreñir la nueva trayectoria \mathbf{P} de la cámara de manera que sea relativamente cercana a la trayectoria original \mathbf{C} de la cámara, a fin de reducir el recorte y la distorsión de la geometría. El término $\|P(t) - P(r)\|^2$ es un término de suavidad para estabilizar la trayectoria de la cámara. El factor de ponderación adaptativo $w_{t,r}(\mathbf{C})$ es un factor que, entre otras cosas, conserva las discontinuidades del movimiento bajo encuadre / rotación o transición de escena rápidos. El factor λ_t es ajustable para equilibrar el término de datos y el término de suavidad.

El factor de ponderación adaptativo $w_{t,r}(\mathbf{C})$ tiende a preservar la discontinuidad de la trayectoria de la cámara. En algunas implementaciones, $w_{t,r}(\mathbf{C})$ comprende un producto de tres funciones gaussianas:

$$W_{t,r}(\mathbf{C}) = \exp(- (t - r)^2 / 2\sigma^2) * \exp(- (C(t) - C(r))^2 / 2\alpha^2) * \exp(- (I(t) - T_r^\dagger(I(r)))^2 / 2\beta^2) \quad (3)$$

5 La primera función gaussiana, $\exp(- (t - r)^2 / 2\sigma^2)$, proporciona un mayor peso a las tramas de vídeo próximas (con respecto a la trama $I(t)$) y un menor peso a las tramas de vídeo más distantes de la trama $I(t)$. En algunos ejemplos, σ puede establecerse de modo que sea igual a $\Omega_t/3$, si bien la materia objeto reivindicada no está limitada por ello.

10 La segunda función gaussiana, $\exp(- (C(t) - C(r))^2 / 2\alpha^2)$, proporciona una medición del cambio entre dos poses de cámara. Para medir tal cambio de las poses de cámara, puede utilizarse el cambio en los componentes de traslación $x(t)$, $y(t)$ extraídos de la pose de cámara $C(t)$. Tal cambio puede ser expresado como $|x(t) - x(r)| + |y(t) - y(r)|$. La traslación de trama x_t , y_t puede describir movimientos de cámara que no incluyen la rotación en el plano ni el cambio de escala alrededor de un eje principal.

15 La tercera función gaussiana, $\exp(- (I(t) - T_r^\dagger(I(r)))^2 / 2\beta^2)$, computa un error de ajuste de homografía como una suma de diferencia al cuadro (SSD –“sum of squared difference”–) entre dos tramas subsiguientemente al registro basado en homografía. La transformación T_r^\dagger puede ser generada evaluando $T_r^\dagger = C_t^{-1}C_r$. Cuando existen grandes oclusiones o cambios de profundidad en una escena de una trama de vídeo, el modelo de movimiento basado en homografía puede hacerse menos efectivo para registrar las tramas vecinas. En tal caso, el error de SSD puede ser grande, y la resistencia de suavización de $w_{t,r}(\mathbf{C})$ puede verse reducida a fin de conservar deseablemente la discontinuidad de la trayectoria de la cámara, que puede ser una característica intencionada.

20 La Figura 4 incluye diagramas esquemáticos que ilustran aspectos de recorte y distorsión de trama de acuerdo con diversas realizaciones. El dibujo de la izquierda muestra una trama 402 suministrada como entrada, una trama estabilizada 404 y una región solapada 406. El dibujo de la derecha muestra otra trama 408 suministrada como entrada y una trama distorsionada 410. El ángulo α es un parámetro para medir un grado de distorsión. Pueden utilizarse también un conjunto de valores de desplazamiento, \mathbf{d} (que incluye un conjunto d_1, d_2, d_3, \dots) para medir un grado de distorsión. Un procedimiento de estabilización de imagen implica diversas magnitudes de estabilización, solapamiento y distorsión de trama, entre otras cosas, para diferentes tramas, tal como se explica más adelante.

25 El término adaptativo $w_{t,r}(\mathbf{C})$ (Ecuación 3) puede permitir alguna capacidad de controlar el recorte y la distorsión de tramas de vídeo. Sin embargo, el ajuste adaptativo del parámetro λ_t en la función de objetivo (Ecuación 2) para cada trama puede proporcionar un control adicional para la relación de recorte y la distorsión. Por ejemplo, la función de objetivo puede ser evaluada utilizando un valor fijo inicial para λ_t (por ejemplo, 5), y la relación de recorte y la distorsión resultantes de cada trama pueden ser comprobadas para determinar si se satisfacen criterios particulares. Por ejemplo, para cualquier trama que no satisfaga los requisitos particulares del usuario (por ejemplo, la relación de recorte o la distorsión son más pequeñas que un umbral predefinido), el parámetro λ_t puede ser ajustado (por ejemplo, aumentado o reducido) en un escalón (por ejemplo, $1/10\lambda_t$) y la función de objetivo puede ser reevaluada con el nuevo valor de λ_t . Tal procedimiento puede ser iterado hasta que todas las tramas satisfagan los requisitos del usuario.

35 La Figura 5 es un diagrama esquemático que muestra relaciones 500 entre trayectorias de cámara originales y trayectorias de cámara suavizadas, de acuerdo con diversas realizaciones. La notación $C(t - 1)$ de la Figura 5 es diferente de la notación $C(t - 1)$ de la Figura 3 para dar cuenta de que se están considerando celdas de rejilla individuales en la Figura 3, que no se están considerando individualmente en la Figura 5. La relación de recorte y la distorsión utilizadas en una estimación de λ_t pueden ser medidas utilizando una transformación de deformación $B(t) = C(t)^{-1}P(t)$. Por ejemplo, puede utilizarse una regulación en escala anisótropa de $B(t)$ para evaluar la distorsión (por ejemplo, los parámetros α y desplazamientos \mathbf{d} mostrados en la Figura 4), que puede ser computada por una relación de los dos valores propios más grandes de una parte afín de $B(t)$. Haciendo referencia a la Figura 4, $B(t)$ puede también ser utilizada para computar un área de región de solapamiento 406 de una trama de vídeo original 402 y una trama estabilizada 404. Se define la relación de recorte como la relación entre el área de solapamiento y el área de la trama original.

Volviendo a la Ecuación 2, puede utilizarse un resolutor de sistema lineal, tal como un resolutor iterativo basado en Jacobi, para resolver la función de objetivo cuadrática $O(\{P(t)\})$. De acuerdo con ello, la función de objetivo $O(\{P(t)\})$ puede escribirse como:

$$50 \quad P(t)^{(\xi+1)} = (C(t) + \sum_{r \in \Omega_t} \lambda_t w_{t,r} P(r)^{(\xi)}) / (1 + \lambda_t \sum_{r \in \Omega_t} w_{t,r}) \quad (4)$$

En esta ecuación, ξ es un índice de interacción (por ejemplo, $\xi = 20$). Con una trayectoria de solución óptima $\{P(t)\}$, puede computarse la transformación $B(t) = C^{-1}(t)P(t)$ puede ser computada para establecer una relación de correspondencia entre la trayectoria de cámara original $C(t)$ y una trayectoria óptima $P(t)$. Puede obtenerse una trama estabilizada final mediante la aplicación de la transformación $B(t)$ a la trama $I(t)$ suministrada como entrada, por interpolación lineal.

Una trayectoria óptima $P(t)$ puede comprender una serie de secuencia de tramas de vídeo suavizada en solapamiento, ejemplos de la cual son los segmentos de trayectoria de cámara suavizados 206 que se muestran en la Figura 2. Tal solapamiento es beneficioso por cuanto no es necesario que el filtrado o la suavización postratamiento operen sobre segmentos de trayectoria discontinuos, lo que podría llevar consigo pobres resultados. El solapamiento de secuencias de tramas de vídeo suavizadas puede ser generado mediante diversas técnicas. Por ejemplo, los segmentos de vídeo pueden ser dispuestos de manera que se solapen unos con otros si, por ejemplo, ε se selecciona de manera que sea menor que τ . En este caso, cada dos segmentos vecinos tienen $(\tau - \varepsilon)$ tramas solapadas, y cada trama (excepto para las tramas inicial y final) es tratada en un episodio de suavización en un número $(\tau/\varepsilon - 1)$ de veces, en distintos segmentos. Puesto que cada trama tiene diferentes estimaciones que se producen en diferentes segmentos de vídeo, estas estimaciones diferentes pueden ser agregadas unas con otras desde diferentes segmentos, para formar una única trayectoria agregada (por ejemplo, los segmentos de trayectoria de cámara suavizados 206). Puede llevarse a cabo una suavización adaptativa híbrida en una etapa de postratamiento para eliminar adicionalmente discontinuidades en la trayectoria agregada (por ejemplo, la trayectoria continua 208), a la vez que se mantiene el tamaño de recorte y la distorsión por debajo de valores de umbral predefinidos. Por ejemplo, para mantener un tamaño de recorte y una distorsión deseados durante el postratamiento, puede reestimarse λ_t a partir de un procedimiento de optimización anterior para una trayectoria de cámara agregada. λ_t puede ser estimada a partir de una transformación de deformación $B(t) = C^{-1}(t)\hat{P}(t)$, donde \hat{P}_t es una pose de cámara estimada tras agregar todas las estimaciones P_t en diferentes segmentos de vídeo. Puede entonces utilizarse la λ_t estimada en una suavización de trayectoria adaptativa híbrida. Los medios 'híbridos' consistentes en que dos tamaños de ventana diferentes $2\Omega/3$ y $\Omega/3$ son sucesivamente aplicados en un término de suavidad adaptativo $\sum_{r \in \Omega_t} w_{t,r}(\mathbf{C}) \cdot \|P(t) - P(r)\|^2$. En algunos casos, varias iteraciones (por ejemplo, 2 o 4) pueden conducir a convergencia. Tras el postratamiento, la trayectoria agregada es suavizada y proporcionada como vídeo de salida que tiene una trayectoria de cámara estabilizada final. En algunas implementaciones, el postratamiento puede ser implementado en tiempo real. Por ejemplo, las técnicas de estabilización de vídeo que incluyen el postratamiento pueden ser llevadas a cabo sustancialmente en paralelo con un procedimiento de captación (por ejemplo, grabación) o de recepción de vídeo. En otras implementaciones, el postratamiento puede ser implementado fuera de conexión. Por ejemplo, las técnicas de estabilización de vídeo que incluyen postratamiento pueden ser llevadas a cabo sobre archivos de vídeo generados en un tiempo anterior.

La Figura 6 es un diagrama de bloques de un dispositivo de estabilización de vídeo 600 de acuerdo con diversas realizaciones. Si bien el dispositivo de estabilización de vídeo 600 se describe como un dispositivo en diversos ejemplos, los procedimientos y técnicas llevados a cabo por el dispositivo de estabilización de vídeo 600 pueden, alternativamente, ser llevados a cabo por un procesador ejecutando código almacenado en memoria. El dispositivo 600, que puede ser incluido en el dispositivo informático 102, lleva a cabo la estabilización de vídeo utilizando una técnica de filtrado de ventana deslizante aplicada a segmentos de vídeo en proceso de descarga por ventanas temporales deslizantes, tal y como se describe más adelante. En comparación con el tratamiento de toda la secuencia de vídeo, el tratamiento según descarga de vídeos por una ventana deslizante temporal proporciona diversos beneficios. Por ejemplo, para la estabilización fuera de conexión, los usuarios pueden reproducir el vídeo mientras el vídeo es simultáneamente estabilizado en el fondo (por ejemplo, sin una intervención del usuario que no sea la reproducción del vídeo). En otro ejemplo, para la estabilización en tiempo real, los vídeos pueden ser estabilizados en tiempo real durante la captación del vídeo. Para conseguir esta característica, un sistema de estabilización de vídeo puede incluir un procesador de almacenamiento temporal de vídeo que tenga un tamaño concreto. El tamaño τ del registro de almacenamiento temporal puede ser restringido o seleccionarse en consideración a las necesidades de potencia de procesador y de memoria del sistema de estabilización de vídeo (que puede comprender, por ejemplo, una computadora o un dispositivo móvil).

Un módulo de fragmentación de vídeo 602 incluye una puerta de entrada de vídeo para recibir vídeo de entrada que comprende tramas de vídeo. En algunas implementaciones, el módulo de fragmentación de vídeo 602 comprende código ejecutable. En este caso, una puerta de entrada de vídeo puede recibir señales electrónicas representativas del vídeo de entrada. El módulo de fragmentación de vídeo 602 fragmenta el vídeo de entrada en un cierto número de secuencias de tramas de vídeo. Las secuencias individuales se proporcionan, respectivamente, a procesadores de almacenamiento temporal de vídeo 604. Por ejemplo, una primera secuencia de tramas de vídeo se proporciona a un primer procesador de almacenamiento temporal de vídeo, una segunda secuencia de tramas de vídeo se proporciona a un segundo procesador de almacenamiento temporal de vídeo, y así sucesivamente. Las secuencias de vídeo de entrada son tratadas según se descargan en el procesador de almacenamiento temporal de vídeo 604, a fin de someterse a un procedimiento de estabilización de vídeo por suavización de trayectoria adaptativa. Tras tal suavización de trayectoria, ε tramas (por ejemplo, $\varepsilon = 5$) de respectivas secuencias de vídeo se descargan fuera de los procesadores de almacenamiento temporal de vídeo individuales para ser entregados como tramas de salida de vídeo. A las restantes tramas de cada secuencia de vídeo pueden agregárseles nuevas secuencias de tramas de flujo entrante, a fin de ser combinadas formando una nueva secuencia de vídeo agregada, que será, subsiguientemente, suavizada en un episodio de postratamiento llevado a cabo por un módulo de filtro de postratamiento 606, tal y como se explica más adelante. Este procedimiento puede ser repetido hasta que sea tratado todo un vídeo.

La Figura 7 es un diagrama esquemático de dos secuencias de tramas de vídeo 702 y 704 de acuerdo con diversas realizaciones. Las descripciones de las secuencias de tramas de vídeo son de utilidad para explicar ideas y definir términos de una técnica de filtrado de ventana deslizante. Las secuencias de tramas de vídeo comprenden una pluralidad de tramas 706. Las secuencias de tramas de vídeo 702 y 704 están descentradas unas con respecto a otras

por ε tramas. Tal descentramiento corresponde a un retardo temporal de una única secuencia de tramas en relación con la otra secuencia de tramas.

La Figura 8 es un diagrama de bloques de una parte de un dispositivo de estabilización de vídeo 800 de acuerdo con diversas realizaciones proporcionadas a modo de ejemplo. Si bien el dispositivo de estabilización de vídeo 800 se describe como un dispositivo en varios ejemplos, pueden, alternativamente, realizarse procedimientos y técnicas llevados a cabo por el dispositivo de estabilización de vídeo 800, por parte de un procesador que ejecuta código almacenado en memoria. Un módulo de fragmentación de vídeo 802, que puede ser similar al módulo de fragmentación de vídeo 602 representado en la Figura 6, puede ser utilizado para proporcionar secuencias de tramas de vídeo a los procesadores de vídeo 804a, 804b, 804c, y así sucesivamente. El módulo de fragmentación de vídeo 802 puede proporcionar una primera secuencia de tramas de vídeo, a través de un bloque 802a, al procesador de almacenamiento temporal de vídeo 804a, sin descentramiento o retardo relativo, proporcionar una segunda secuencia de tramas de vídeo, a través de un bloque 802b, al procesador de almacenamiento temporal de vídeo 804b, con un descentramiento o retardo de n tramas, proporcionar una tercera secuencia de tramas de vídeo, a través de un bloque 802c, al procesador de almacenamiento temporal de vídeo 804c, con un descentramiento o retardo de $2n$ tramas, proporcionar una cuarta secuencia de tramas de vídeo al siguiente procesador de almacenamiento temporal de vídeo, con un descentramiento o retardo de $3n$ tramas, y así sucesivamente. El módulo de partición de vídeo puede continuar con este patrón para $m + 1$ procesadores de almacenamiento temporal de vídeo. De acuerdo con ello, se proporcionan a sucesivos procesadores de almacenamiento temporal de vídeo secuencias de tramas de vídeo que son sucesivamente retardadas, a modo de una trama, de un procesador de almacenamiento temporal de vídeo al siguiente procesador de almacenamiento temporal de vídeo. Por ejemplo, tal retardo a modo de trama se ha representado en la Figura 7, en la que las secuencias de tramas de vídeo 702 y 704 están descentradas unas con respecto a otras en ε tramas.

Al recibir, o después de recibir, por parte del procesador de almacenamiento temporal de vídeo 804a, una secuencia de tramas de vídeo, un módulo de extracción de rasgos 808a identifica rasgos objeto en tramas de vídeo individuales de la secuencia de tramas de vídeo. Por ejemplo, tales rasgos objeto pueden comprender puntos o bordes en las tramas de vídeo individuales. De forma subsiguiente, el módulo de estimación de homografía 810a lleva a cabo una estimación homográfica entre tramas de vídeo consecutivas de la secuencia de tramas de vídeo, a fin de generar una secuencia de tramas de vídeo modificada. La estimación homográfica se basa, al menos en parte, en los rasgos objeto identificados por el módulo de extracción de rasgos 808a. El módulo 812a de suavización de trayectoria adaptativa determina errores de estimación de entre las tramas de vídeo de la secuencia de tramas de vídeo modificada, y aplica suavización de trayectoria adaptativa a la secuencia de tramas de vídeo modificada, con el fin de generar una secuencia de tramas de vídeo suavizada. La suavización de trayectoria adaptativa está basada, al menos en parte, en cambios entre las tramas de vídeo individuales de la secuencia de tramas de vídeo y en los errores de estimación.

Los procesadores de almacenamiento temporal de vídeo sucesivos 804b y 804c llevan a cabo procedimientos similares al que se ha descrito anteriormente para el procesador de almacenamiento temporal de vídeo 804a. Los segmentos de tramas de vídeo proporcionados a los procesadores de almacenamiento temporal de vídeo sucesivos 804b y 804c, sin embargo, son sucesivamente retardados de un procesador de almacenamiento temporal de vídeo al siguiente procesador de almacenamiento temporal de vídeo. Como se ha explicado anteriormente, el módulo de fragmentación de vídeo 802 proporciona una secuencia de tramas de vídeo al procesador de almacenamiento temporal de vídeo 804b con un descentramiento o retardo de n tramas, y proporciona una tercera secuencia de tramas de vídeo al procesador de almacenamiento temporal de vídeo 804c con un descentramiento o retardo de $2n$ tramas, y así sucesivamente. Los procesadores de almacenamiento temporal de vídeo (804a, 804b y 804c...) pueden operar sobre sus respectivas secuencias de tramas de vídeo en paralelo.

El procesador de almacenamiento temporal de vídeo 804a genera una salida que comprende un segmento de tramas de vídeo suavizado. Un ejemplo de varios de tales segmentos de vídeo suavizados puede ser los segmentos de trayectoria de cámara suavizada 206 que se muestran en la Figura 2. Como se ha explicado anteriormente, los segmentos de vídeo suavizados individualmente generados por los procesadores de almacenamiento intermedio de vídeo (804a, 804b y 804c...), incluyen segmentos en solapamiento. Estos segmentos de vídeo suavizados generados por los procesadores de almacenamiento temporal de vídeo son proporcionados al módulo 814 de agregación de trayectorias, el cual agrega los segmentos de vídeo suavizados individuales. Los segmentos de vídeo suavizados agregados son entonces proporcionados a un módulo de filtrado postratamiento 816, que lleva a cabo una suavización adaptativa híbrida para eliminar adicionalmente la discontinuidad en los segmentos de vídeo suavizados agregados, al tiempo que mantiene el tamaño de recorte y la distorsión por debajo de valores de umbral predefinidos. El módulo de filtrado postratamiento 816 genera, de esta forma, una trayectoria continua suavizada, un ejemplo de la cual es la trayectoria continua 208 de la Figura 2. La trayectoria continua suavizada es entonces proporcionada al módulo de deformación bilineal 818 de tramas de imagen, que suministra entonces como salida un vídeo para ser visionado. A fin de realizar una deformación bilineal, se establece primeramente una relación de correspondencia entre cada coordenada de píxel (x, y) de una trama t y nuevas coordenadas (x', y') de acuerdo con la transformación de deformación inversa $B^{-1}(t)$. A continuación, se determina el valor de píxel $I'_{x,y}$ por medio de la interpolación bilineal de sus cuatro píxeles vecinos $I_{\lfloor x' \rfloor, \lfloor y' \rfloor}, I_{\lfloor x' \rfloor + 1, \lfloor y' \rfloor}, I_{\lfloor x' \rfloor, \lfloor y' \rfloor + 1}, I_{\lfloor x' \rfloor + 1, \lfloor y' \rfloor + 1}$, donde $\lfloor x' \rfloor$ es el máximo entero más allá de x' .

La Figura 9 es un diagrama de flujo de un procedimiento 900 para suavizar vídeo, de acuerdo con diversas realizaciones. En un ejemplo, tal procedimiento puede ser llevado a cabo por el dispositivo de estabilización de vídeo

800. En el bloque 902, una porción de un vídeo es recibida a través de una puerta para vídeo. La porción de vídeo comprende una secuencia de tramas de vídeo. En el bloque 904, rasgos objeto afines de tramas de vídeo individuales de la secuencia de las tramas de vídeo son comparados con rasgos objeto afines de las demás tramas de vídeo individuales de la secuencia de las tramas de vídeo. En el bloque 906, se utiliza una estimación homográfica entre, o de entre, tramas de vídeo de la secuencia de tramas de vídeo, a fin de generar una secuencia modificada de tramas de vídeo. La estimación homográfica está basada, al menos en parte, en el procedimiento de comparación del bloque 904. Por ejemplo, la estimación homográfica puede incluir la generación de relaciones de correspondencia entre tramas de vídeo individuales basada, al menos en parte, en rasgos objeto identificados en el procedimiento de comparación. En el bloque 908, se determinan los errores de estimación entre las tramas de vídeo de la secuencia modificada de tramas de vídeo. En el bloque 910, se aplica una suavización de trayectoria adaptativa a la secuencia modificada de tramas de vídeo para generar una secuencia suavizada de tramas de vídeo. La suavización de trayectoria adaptativa conserva las discontinuidades de movimiento correspondientes a los cambios de imágenes entre la secuencia de tramas de vídeo. La suavización de trayectoria adaptativa está basada, al menos en parte, en estos cambios y en los errores de estimación determinados en el bloque 908.
- 15 Los flujos de operaciones que se ilustran en las Figuras 6, 8 y 9 se han ilustrado como colecciones de bloques y/o flechas que representan secuencias de operaciones que pueden ser implementadas en hardware, software, *firmware* o en una combinación de los mismos. No es la intención que el orden en que se describen los bloques sea interpretado como una limitación, y es posible combinar cualquier número de las operaciones descritas, en cualquier orden, para implementar uno o más métodos, o métodos alternativos. Adicionalmente, pueden omitirse operaciones individuales en el flujo de operaciones sin apartarse del alcance de la materia objeto que se describe en esta memoria. En el contexto del software, los bloques representan instrucciones legibles por computadora que, cuando son ejecutadas por uno o más procesadores, llevan a cabo las operaciones referidas. En el contexto del hardware, los bloques pueden representar uno o más circuitos (por ejemplo, circuitos integrados específicos de aplicación –ASICs–) configurados para llevar a cabo las operaciones referidas.
- 25 Cualesquiera descripciones, elementos o bloques rutinarios de los flujos de operaciones ilustrados en las Figuras 6, 8 y 9 pueden representar módulos, segmentos o partes de código que incluyen una o más instrucciones ejecutables para implementar funciones lógicas o elementos específicos de la rutina.

Conclusión

- 30 Si bien las técnicas se han descrito en un lenguaje específico de características estructurales y/o de acciones metodológicas, ha de entenderse que las reivindicaciones que se acompañan no están necesariamente limitadas a las características o acciones descritas. En lugar de ello, las características y las acciones se describen como implementaciones proporcionadas a modo de ejemplo de tales técnicas.
- 35 Todos los métodos y procedimientos anteriormente descritos pueden ser materializados en, y completamente automatizados por medio de, módulos de código de software ejecutados por una o más computadoras o procesadores de propósito general. Los módulos de código pueden ser almacenados en cualquier tipo de medio de almacenamiento legible por computadora u otro dispositivo de almacenamiento informático. Algunos de los métodos, o todos ellos, pueden, alternativamente, ser materializados en hardware informático especializado.
- 40 Lenguaje condicional tal como, entre otras cosas, 'puede', 'podría', 'es posible' o 'cabe la posibilidad', a menos que se diga específicamente lo contrario, se utiliza para indicar que ciertas realizaciones incluyen, en tanto que otras realizaciones no incluyen, las características, elementos y/o etapas mencionados. Así, pues, a menos que se afirme lo contrario, no es la intención que tal lenguaje condicional implique que características, elementos y/o etapas son en ningún modo necesarios para una o más realizaciones, ni que una o más realizaciones incluyan necesariamente lógica para decidir, con o sin una introducción por parte del usuario o una instancia presentada a este, si estas características, elementos y/o etapas están incluidos o han de ser llevados a cabo en cualquier realización particular.
- 45 Lenguaje combinatorio tal como la frase 'al menos uno de entre X, Y o Z', a menos que se indique específicamente de otro modo, ha de entenderse de manera que establezca que un elemento, término, etc. puede ser el X, o el Y, o el Z, o una combinación de los mismos.
- 50 Es posible realizar muchas variaciones y modificaciones en las realizaciones antes descritas, cuyos elementos han de entenderse de manera que se encuentren entre otros ejemplos aceptables. Es la intención que todas estas modificaciones y variaciones estén incluidas en esta memoria, dentro del alcance de esta invención.

REIVINDICACIONES

- 1.- Un sistema (102) que comprende:
una puerta de entrada de vídeo, para recibir un vídeo que comprende una secuencia de tramas de vídeo;
uno o más procesadores (104); y
- 5 memoria (108) destinada a almacenar instrucciones que, cuando son ejecutadas por uno o más procesadores (104), configuran los uno o más procesadores (104) para llevar a cabo operaciones que comprenden:
identificar (904) rasgos objeto en tramas de vídeo individuales de la secuencia de las tramas de vídeo;
llevar a cabo (906) una estimación homográfica entre tramas de vídeo de la secuencia de tramas de vídeo, a fin de generar una secuencia modificada de tramas de vídeo, de tal manera que la estimación homográfica está basada,
10 al menos en parte, en los rasgos objeto identificados y determina una trayectoria de cámara original constituida por una secuencia de poses de cámara definidas utilizando homografías estimadas entre cada par de tramas de vídeo vecinas de la secuencia de tramas de vídeo;
- caracterizado por que** las operaciones comprenden, adicionalmente: aplicar (910) suavización de trayectoria adaptativa a la secuencia modificada de tramas de vídeo, a fin de generar una secuencia suavizada de tramas de vídeo, de tal modo que la suavización de trayectoria adaptativa comprende determinar una trayectoria de cámara suavizada, dada la trayectoria de cámara original, minimizando una función de objetivo, de tal manera que la función de objetivo comprende un término de datos que tiende a restringir la trayectoria de cámara suavizada de manera que sea cercana a la trayectoria de cámara original, un término de suavización para estabilizar la trayectoria de cámara suavizada, y un factor de ponderación adaptativo, asociado con el término de suavización, que tiende a conservar la discontinuidad en la trayectoria de la cámara, de tal modo que el factor de ponderación adaptativo está basado, al menos en parte, en una función que proporciona una medición de los cambios entre dos poses de cámara de la trayectoria original de la cámara, y en una función que computa un error de ajuste de homografía entre dos tramas de vídeo subsiguientes al registro basado en homografía.
- 15
- 20
- 2.- El sistema (102) de acuerdo con la reivindicación 1, que comprende, de manera adicional, un módulo de fragmentación de vídeo (120), conectado de forma comunicativa con la puerta de entrada de vídeo, de tal manera que el módulo de fragmentación de vídeo (120) está configurado para llevar a cabo operaciones que comprenden:
25 recibir el vídeo desde la puerta de entrada de vídeo;
fraccionar el vídeo en un cierto número de secuencias de tramas de vídeo;
aplicar retardos de trama a secuencias individuales de las secuencias de tramas de vídeo; y
- 30 proporcionar, respectivamente, las secuencias a registros de almacenamiento temporal individuales de los uno o más procesadores, incluyendo proporcionar una primera de dichas secuencias sin retardo de trama.
- 3.- El sistema (102) de acuerdo con la reivindicación 2, en el cual el módulo de fragmentación de vídeo (120) está almacenado en la memoria y es ejecutable por los uno o más procesadores.
- 35 4.- El sistema (102) de acuerdo con la reivindicación 2, en el cual aplicar los retardos de trama comprende, adicionalmente, desplazar en trama las tramas de vídeo de las secuencias de tal modo que secuencias de tramas de vídeo que no son la primera secuencia estén, respectivamente, descentradas con respecto a la primera secuencia de tramas de vídeo en $m \times n$ tramas, donde m es la serie de enteros 1, 2, 3... y n es un número predeterminado.
- 5.- El sistema (102) de acuerdo con la reivindicación 4, que comprende, adicionalmente, un módulo de agregación de vídeo, configurado para llevar a cabo operaciones que comprenden:
- 40 agregar el número de secuencias de tramas de vídeo seguidamente a la suavización de trayectoria adaptativa, a fin de producir una porción de vídeo agregada; y
aplicar filtrado de postratamiento para suavizar la porción de vídeo agregada.
- 6.- El sistema (102) de acuerdo con la reivindicación 5, en el cual el módulo de agregación de vídeo está almacenado en la memoria (108) y es ejecutable por los uno o más procesadores (104).
- 45 7.- El sistema (102) de acuerdo con la reivindicación 1, en el cual los rasgos objeto de las tramas de vídeo individuales de la secuencia de tramas de vídeo comprenden puntos o bordes en las tramas de vídeo individuales.
- 8.- El sistema (102) de acuerdo con la reivindicación 1, en el cual los uno o más procesadores (104) comprenden una o más unidades centrales de procesamiento (CPUs), unidades de tratamiento de gráficos (GPUs) o procesadores de almacenamiento temporal de vídeo.

9.- Un método que comprende:

fragmentar un vídeo en una pluralidad de secuencias de vídeo;

retardar secuencias de vídeo individuales de la pluralidad de las secuencias de vídeo de tal manera que las secuencias de vídeo individuales queden retrasadas en un número particular de tramas unas con respecto a otras; y

5 llevar a cabo, para cada secuencia de vídeo de las secuencias de vídeo individuales, un procedimiento de suavización que comprende:

identificar rasgos objeto en tramas de vídeo individuales de la secuencia de tramas de vídeo;

10 llevar a cabo (906) una estimación homográfica entre tramas de vídeo individuales basándose, al menos en parte, en los rasgos objeto identificados en las tramas de vídeo individuales, a fin de generar una secuencia de vídeo modificada y de determinar una trayectoria de cámara original formada por una secuencia de poses de cámara definida utilizando homografías estimadas entre cada par de tramas de vídeo vecinas de la secuencia de vídeo;

15 **caracterizado por que** el procedimiento de suavización comprende, adicionalmente: aplicar (910) suavización de trayectoria adaptativa a la secuencia de vídeo modificada, a fin de generar una secuencia de vídeo suavizada, de tal manera que la suavización de trayectoria adaptativa comprende determinar una trayectoria de cámara suavizada, dada la trayectoria de cámara original, minimizando una función de objetivo, de tal manera que la función de objetivo comprende un término de datos que tiende a restringir la trayectoria de cámara suavizada de manera que sea cercana a la trayectoria de cámara original, un término de suavización para estabilizar la trayectoria de cámara suavizada, y un factor de ponderación adaptativo, asociado con el término de suavización, que tiende a conservar la discontinuidad en la trayectoria de la cámara, de tal modo que el factor de ponderación adaptativo está basado, al menos en parte, en una función que proporciona una medición de los cambios entre dos poses de cámara de la trayectoria original de la cámara, y en una función que computa un error de ajuste de homografía entre dos tramas de vídeo subsiguientes al registro basado en homografía.

20 10.- El método de acuerdo con la reivindicación 9, en el que las operaciones comprenden, adicionalmente:

25 agregar la pluralidad de secuencias de vídeo seguidamente a la suavización de trayectoria adaptativa, a fin de producir una porción de vídeo agregada; y

aplicar un filtrado de postratamiento (122) para suavizar la porción de vídeo agregada.

11.- El método de acuerdo con la reivindicación 9, en el cual los rasgos objeto de las tramas de vídeo individuales de la secuencia de las tramas de vídeo comprenden puntos o bordes en las tramas de vídeo individuales.

30 12.- El método de acuerdo con la reivindicación 9, en el cual el procedimiento de suavización se lleva a cabo simultáneamente para más de una de las secuencias de vídeo individuales.

13.- Un medio de almacenamiento legible por computadora (108), que almacena instrucciones ejecutables por computadora que, cuando son ejecutadas por uno o más procesadores (104), configuran los uno o más procesadores (104) para llevar a cabo el método de acuerdo con una cualquiera de las reivindicaciones 9 a 12.

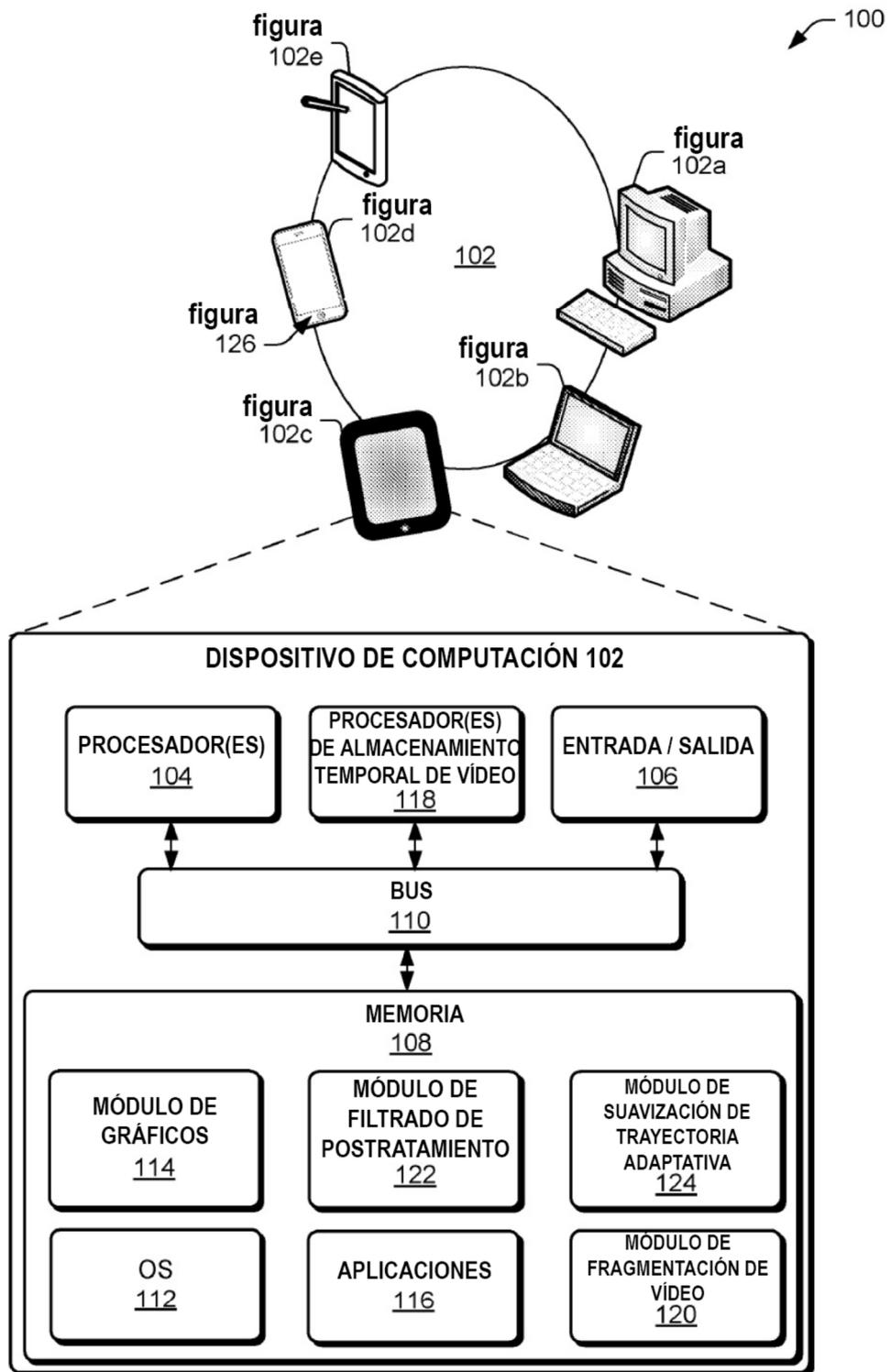


FIG. 1

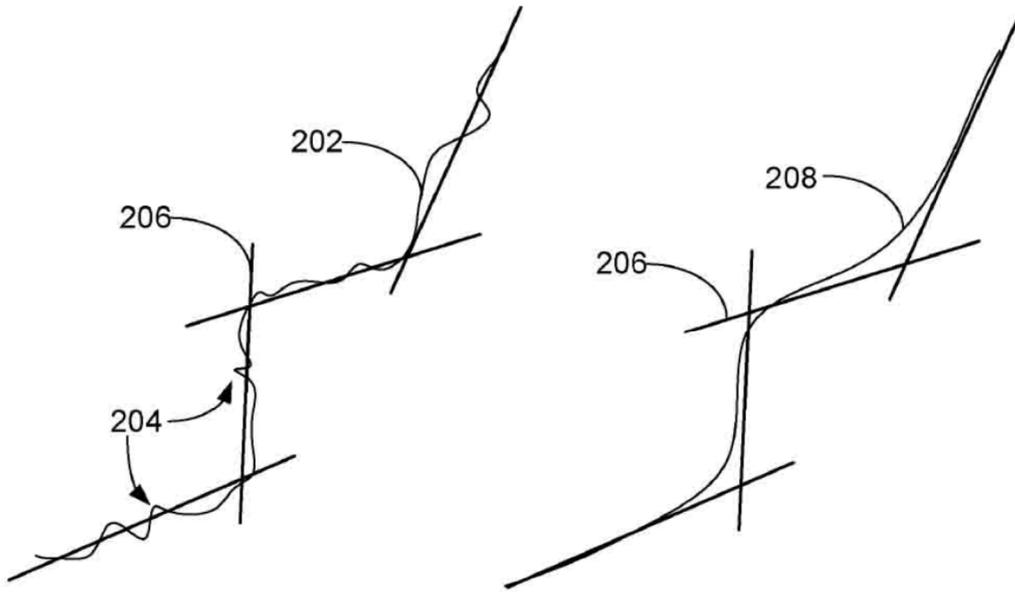


FIG. 2

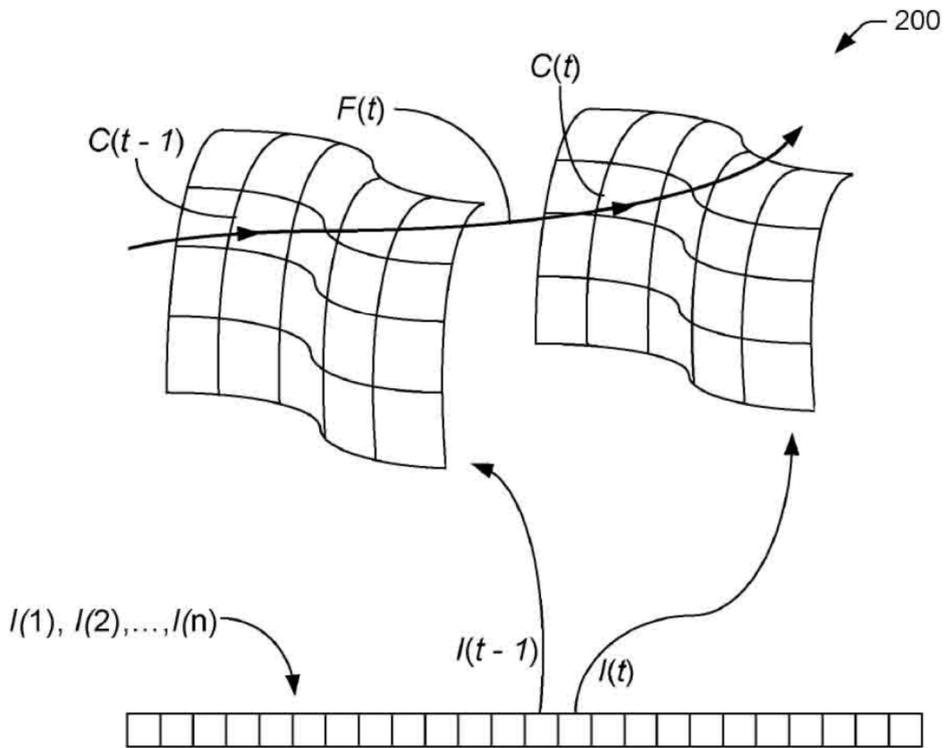


FIG. 3

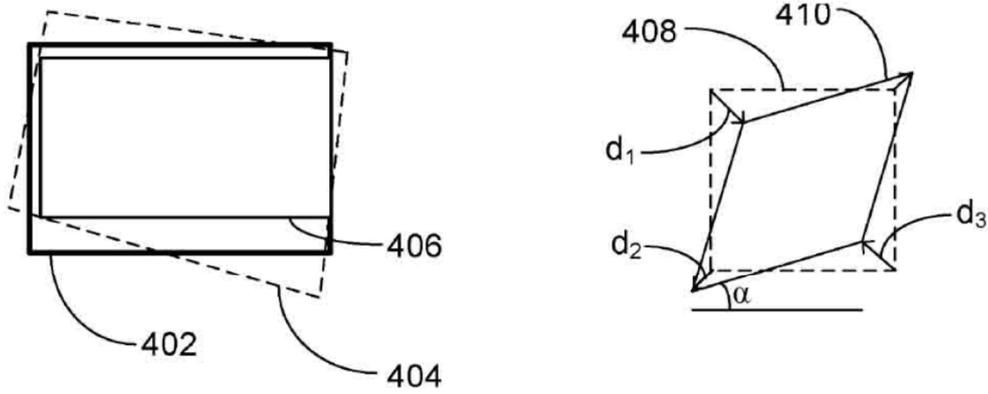


FIG. 4

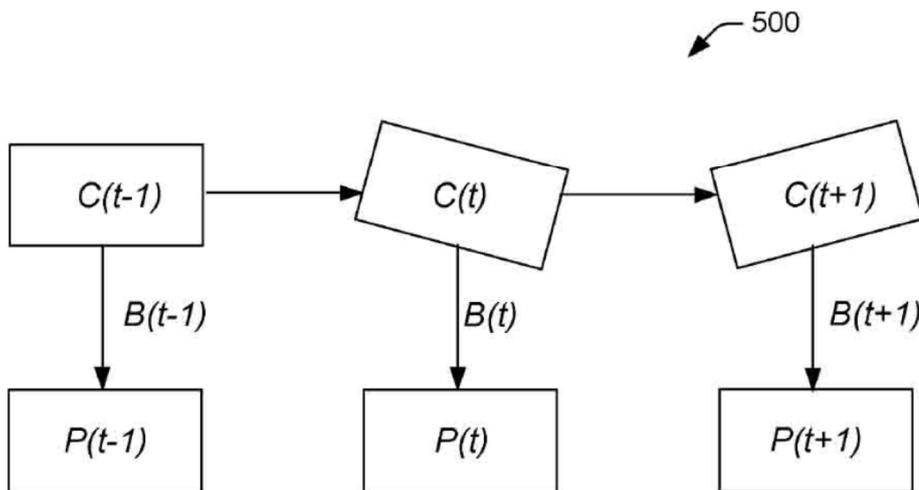


FIG. 5

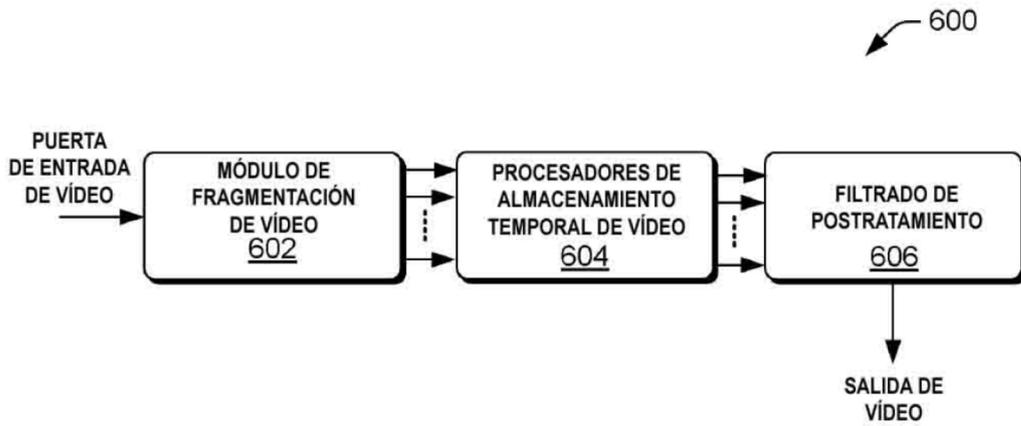


FIG. 6

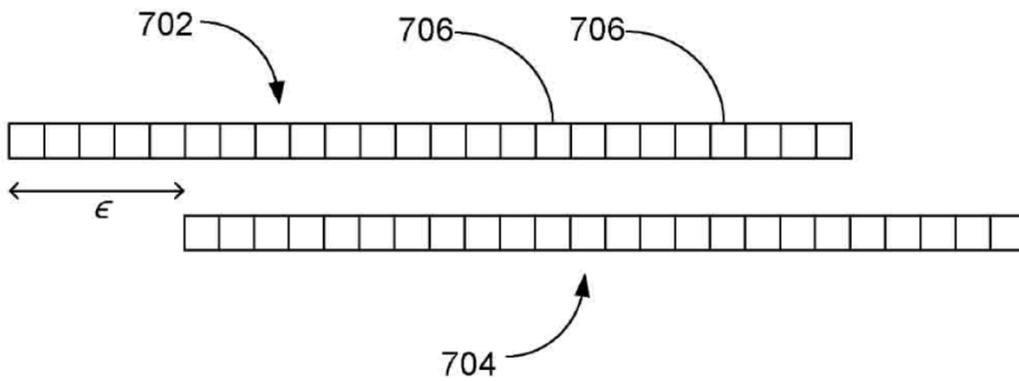


FIG. 7

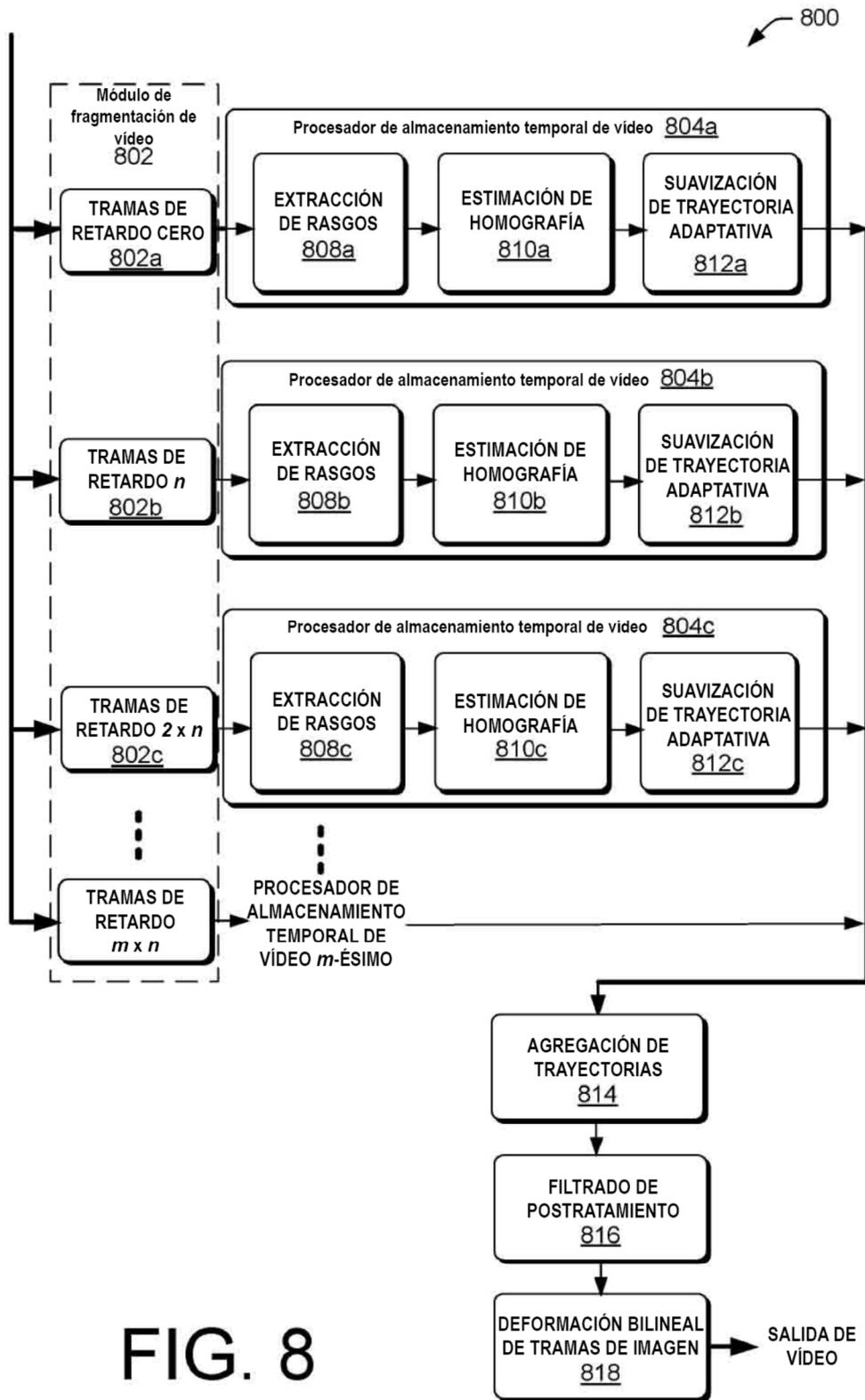


FIG. 8

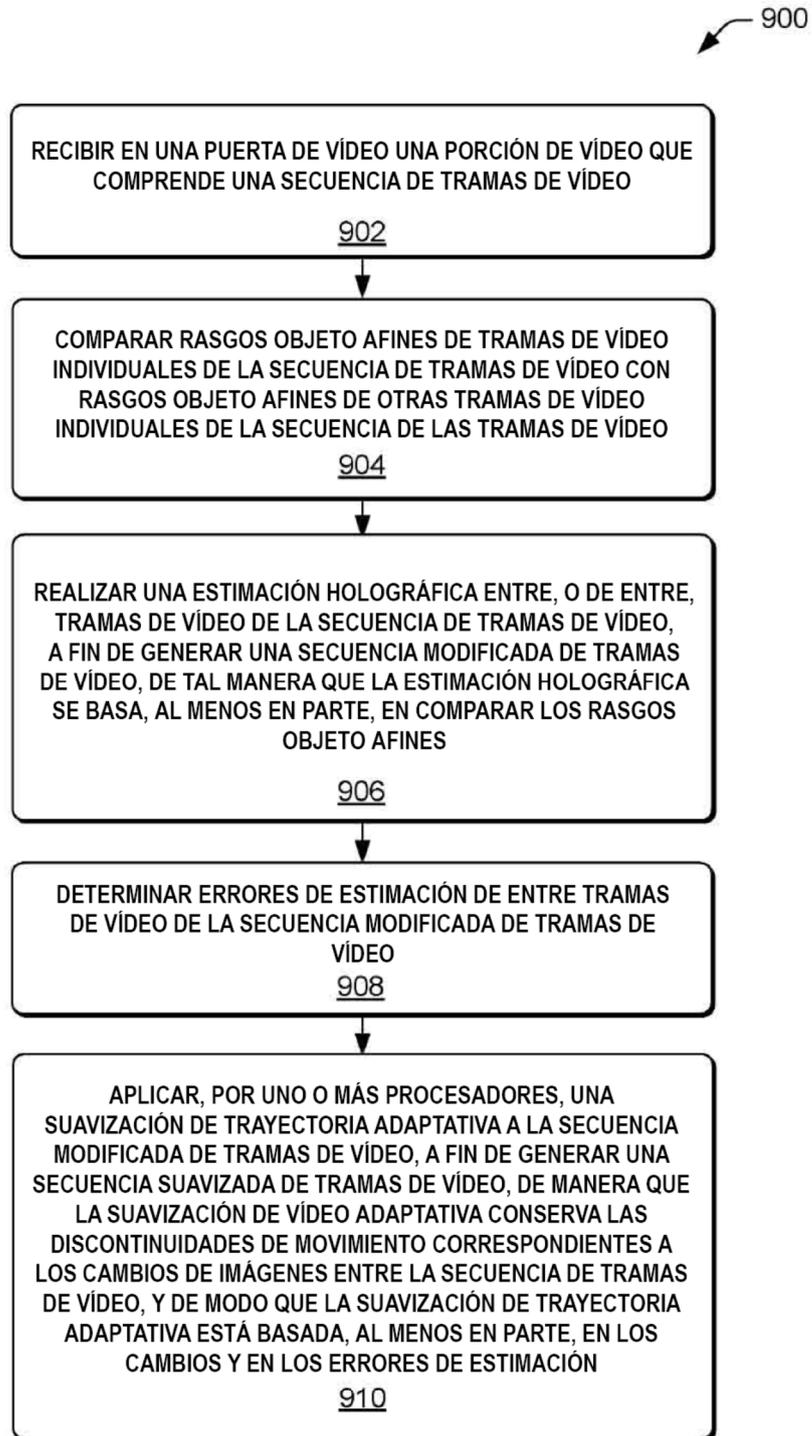


FIG. 9