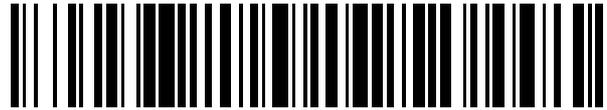


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 684 124**

51 Int. Cl.:

G06T 7/20

(2007.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **18.01.2008 PCT/FR2008/000069**

87 Fecha y número de publicación internacional: **12.09.2008 WO08107554**

96 Fecha de presentación y número de la solicitud europea: **18.01.2008 E 08761787 (4)**

97 Fecha y número de publicación de la concesión europea: **20.06.2018 EP 2111605**

54 Título: **Procedimiento y dispositivo para crear al menos dos imágenes clave correspondientes a un objeto tridimensional**

30 Prioridad:

22.01.2007 FR 0752810

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

01.10.2018

73 Titular/es:

**QUALCOMM INCORPORATED (100.0%)
5775 Morehouse Drive
San Diego, CA 92121 , US**

72 Inventor/es:

**LEFEVRE, VALENTIN;
LIVET, NICOLAS;
CHABI, WOROU, PIERRICK y
QUEMENER, YVES**

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 684 124 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 y dispositivo para crear al menos dos imágenes clave correspondientes a un objeto tridimensional

- 5 **[0001]** La presente invención se refiere a la combinación de imágenes reales y virtuales en tiempo real, también llamada realidad aumentada, y, más en particular, a un procedimiento y un dispositivo para crear imágenes clave correspondientes a un objeto tridimensional.
- 10 **[0002]** El objetivo de la realidad aumentada es insertar uno o más objetos virtuales en las imágenes de un flujo de vídeo. Según el tipo de aplicación, la posición y orientación de estos objetos virtuales pueden ser determinadas por datos externos a la escena representada por las imágenes, por ejemplo, coordenadas obtenidas directamente de un escenario de juego, o por datos vinculados a ciertos elementos de esa escena, por ejemplo, coordenadas de un punto particular de la escena, como la mano de un jugador. Cuando la posición y la orientación están determinadas por datos vinculados a ciertos elementos de esa escena, puede ser necesario rastrear esos elementos en función de los movimientos de la cámara o los movimientos de esos mismos elementos dentro de la escena. Las operaciones de rastreo de elementos e incrustación de objetos virtuales en las imágenes reales pueden ser ejecutadas por computadoras distintas o por una misma computadora.
- 15 **[0003]** Existen varios procedimientos para rastrear elementos en un flujo de imágenes. Generalmente, los algoritmos de rastreo de elementos, también llamados algoritmos de persecución de objetivo, utilizan un marcador que puede ser visual o utilizar otros medios tales como medios basados en radiofrecuencias o infrarrojos. Alternativamente, algunos algoritmos usan un reconocimiento de forma para rastrear un elemento particular en un flujo de imágenes. Por ejemplo, véase el documento D1: XP010644679 "Fusing online and offline information for stable 3D tracking in real-time" de Lucca VACHETTI et al. La Escuela Politécnica Federal de Lausana ha desarrollado un algoritmo de rastreo visual que no utiliza marcadores y cuya originalidad consiste en hacer coincidir puntos particulares entre la imagen actual de un flujo de vídeo y una imagen clave, llamada *keyframe*, proporcionada por el usuario al inicio del sistema y una imagen clave actualizada durante la ejecución del rastreo visual.
- 20 **[0004]** El objetivo de este algoritmo de rastreo visual es encontrar, en una escena real, la postura, es decir, la posición y la orientación, de un objeto cuyo mallado tridimensional está disponible, o encontrar los parámetros extrínsecos de posición y orientación en relación con ese objeto de una cámara inmóvil que filma ese objeto, gracias al análisis de imagen.
- 30 **[0005]** La imagen de vídeo actual se compara con una o varias imágenes claves grabadas para encontrar un número importante de correspondencias entre esos pares de imágenes con el fin de estimar la postura del objeto. Con este fin, una imagen clave se compone de dos elementos: una imagen capturada del flujo de vídeo y una postura (orientación y posición) de un modelo tridimensional que aparece en esa imagen. Conviene distinguir entre las imágenes clave "fuera de línea" u *off line*, y las imágenes clave "en línea" o *on line*. Las imágenes clave fuera de línea son imágenes extraídas del flujo de vídeo en el que el objeto a rastrear se ha colocado manualmente usando un dispositivo señalador, tal como un ratón, o usando una herramienta de ajuste tal como un Pocket Dial comercializado por la compañía Doepfer. Las imágenes clave fuera de línea caracterizan preferiblemente la postura de un mismo objeto en una serie de imágenes. Se crean y almacenan "fuera de línea", es decir, fuera del régimen permanente de la aplicación. Las imágenes clave "en línea" se memorizan dinámicamente durante la ejecución del programa de rastreo. Se calculan cuando el error, es decir, la distancia entre las coincidencias de los puntos de interés, es bajo. Las imágenes clave en línea reemplazan a las imágenes clave fuera de línea utilizadas para inicializar la aplicación. Su uso tiene como objetivo reducir el desfase, también conocido como deriva, que puede volverse importante al alejarse demasiado de la posición relativa inicial entre la cámara y el objeto. El aprendizaje de nuevas imágenes clave en línea también hace que la aplicación se vuelva más robusta ante las variaciones de luz externa y ante las variaciones de colorimetría de las cámaras. Sin embargo, tienen la desventaja de introducir un efecto de "vibración" en la postura del objeto con el tiempo. Al aprender una nueva imagen clave en línea, ésta reemplaza la imagen clave precedente, en línea o fuera de línea. Se usa como imagen clave actual.
- 35 **[0006]** Cada imagen clave fuera, fuera de línea o en línea, incluye una imagen en la que el objeto está presente y una postura para caracterizar la ubicación de ese objeto así como un cierto número de puntos de interés que caracterizan al objeto en la imagen. Por ejemplo, los puntos de interés se construyen a partir de un detector de puntos de Harris y representan ubicaciones en la imagen con fuertes valores de gradientes de direccionalidad.
- 40 **[0007]** Antes de inicializar la aplicación, es necesario determinar una o más imágenes clave fuera de línea. Generalmente se trata de imágenes extraídas del flujo de vídeo que contienen el objeto a rastrear y a las que se asocia una posición y una orientación del modelo tridimensional de ese objeto. Con este fin, un operador lleva a cabo una operación manual que consiste en hacer corresponder visualmente un modelo reticulado con el objeto real. La fase de preparación manual consiste por lo tanto en encontrar una primera estimación de la postura del objeto en una imagen extraída del flujo de vídeo, lo que equivale a formalizar la transformación afín inicial $T_{p \rightarrow c}$ que corresponde a la matriz de paso entre la zona de localización ligada al objeto y la zona de localización ligada a la cámara. La transformación afín inicial se puede descomponer según una primera transformación $T_{o \rightarrow c}$ relacionada con una posición inicial del objeto, por ejemplo, en el centro de la pantalla, es decir, una transformación vinculada al
- 45
50
55
60
65

5 cambio de zona de localización entre la zona de localización de la cámara y la zona de localización del objeto, y una segunda transformación $T_{o \rightarrow c}$ relacionada con el desplazamiento y la rotación del objeto desde su posición inicial en el centro de la pantalla hacia la posición y orientación en las que se encuentra realmente el objeto en la imagen clave, donde $T_{p \rightarrow c} = T_{p \rightarrow o} \cdot T_{o \rightarrow c}$. Si los valores a, b y g corresponden a la traslación del objeto desde su posición inicial en el centro de la imagen hacia su posición en la imagen clave y si los valores q, f y j corresponden a la rotación del objeto desde su posición inicial en el centro de la imagen hacia su posición en la imagen clave según los ejes x, y y z, la transformación $T_{p \rightarrow o}$ puede expresarse entonces en forma de la siguiente matriz:

$$T_p^o = \begin{bmatrix} \cos \varphi \cos \phi + \sin \varphi \sin \theta \sin \phi & \sin \varphi \cos \phi - \cos \varphi \sin \theta \sin \phi & \cos \theta \sin \phi & \alpha \\ -\sin \varphi \cos \theta & \cos \varphi \cos \theta & \sin \theta & \beta \\ \sin \varphi \sin \theta \cos \phi - \cos \varphi \sin \phi & -\cos \varphi \sin \theta \cos \phi - \sin \varphi \sin \phi & \cos \theta \cos \phi & \gamma \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

10 **[0008]** El uso de este modelo permite establecer el vínculo entre las coordenadas de los puntos del modelo tridimensional del objeto en la zona de localización del objeto y las coordenadas de esos puntos en la zona de localización de la cámara.

15 **[0009]** Al inicializar la aplicación, las imágenes clave fuera de línea se procesan para situar puntos de interés en función de los parámetros elegidos al iniciar la aplicación. Esos parámetros se especifican empíricamente para cada tipo de uso de la aplicación y permiten modular el núcleo de detección de coincidencia y obtener una mejor calidad en la estimación de la postura del objeto de acuerdo a las características del entorno real. Entonces, cuando el objeto real en la imagen actual está en una postura que está cerca de la postura de ese mismo objeto en una de las
20 imágenes clave fuera de línea, el número de coincidencias se vuelve importante. Entonces es posible encontrar la transformación afín que permite asociar por claves el modelo tridimensional virtual del objeto al objeto real.

[0010] Cuando se ha encontrado una coincidencia de ese tipo, el algoritmo pasa al régimen permanente. Los desplazamientos del objeto se rastrean de una trama a la otra y las eventuales derivadas se compensan gracias a las informaciones contenidas en la imagen clave fuera de línea retenida en el momento de la inicialización y en la imagen clave en línea calculada durante la ejecución de la aplicación.

30 **[0011]** La aplicación de rastreo combina dos tipos de algoritmo: una detección de puntos de interés, por ejemplo, una versión modificada de la detección de puntos de Harris, y una técnica de reproyección de los puntos de interés situados en el modelo tridimensional hacia la imagen plana. Esta reproyección permite prever el resultado de una transformación espacial de una trama a otra. Estos dos algoritmos combinados proporcionan un rastreo robusto de un objeto de acuerdo a seis grados de libertad.

35 **[0012]** En términos generales, un punto p de la imagen es la proyección de un punto P de la escena real, donde $p \sim P_i \cdot P_E \cdot T_{p \rightarrow c} \cdot P$, donde P_i es la matriz de los parámetros intrínsecos de la cámara, es decir, su valor focal, el centro de la imagen y el desplazamiento, P_E es la matriz de los parámetros extrínsecos de la cámara, es decir, la posición de la cámara en el espacio real, y $T_{p \rightarrow c}$ es la matriz afín de paso entre la zona de localización asociada al objeto rastreado y la zona de localización de la cámara. Aquí sólo se considera la posición relativa del objeto con respecto a la posición relativa de la cámara, lo que equivale a colocar la zona de localización de la escena real en el centro
40 óptico de la cámara. Esto produce la relación $p \sim P_i \cdot T_{p \rightarrow c} \cdot P$, donde $T_{p \rightarrow c}$ es la matriz de la postura del objeto en la zona de localización de la cámara. Como se conoce la matriz P_i , el problema de rastreo consiste pues en determinar la matriz $T_{p \rightarrow c}$.

45 **[0013]** Sin embargo, es importante tener en cuenta que si la medición de errores es demasiado alta, es decir, si el número de coincidencias entre la imagen clave actual y la imagen actual se hace demasiado pequeño, el seguimiento se desincroniza (se considera que la estimación de la postura del objeto ya no es lo suficientemente coherente) y es necesaria una nueva fase de inicialización utilizando las mismas imágenes clave fuera de línea.

50 **[0014]** La postura de un objeto se estima de acuerdo a las coincidencias entre los puntos de interés de la imagen actual del flujo de vídeo, los puntos de interés de la imagen clave actual y los puntos de interés de la imagen anterior del flujo de vídeo. Estas operaciones se llaman fase de coincidencia. A partir de las correlaciones más significativas, el software calcula la postura del objeto que mejor se corresponde con las observaciones.

55 **[0015]** Las figuras 1 y 2 ilustran esta aplicación de rastreo.

[0016] Las soluciones propuestas a menudo se derivan de la investigación y no tienen en cuenta las limitaciones de implementación de sistemas comerciales. En particular, a menudo se ignoran los problemas vinculados a la solidez, a la posibilidad de iniciar la aplicación rápidamente sin necesidad de una fase manual de creación de imágenes clave fuera de línea necesarias para la inicialización del sistema de rastreo, a la detección de errores de

"desincronización" (cuando un objeto a rastrear "se pierde") y a la reinicialización automática y tiempo real después de tales errores.

5 **[0017]** La invención permite resolver al menos uno de los problemas descritos anteriormente.

[0018] Así, la invención tiene por objeto un procedimiento para la creación de al menos dos imágenes clave que comprenden, cada una, una imagen que representa al menos un objeto tridimensional en un entorno tridimensional y la postura del objeto en ese entorno de acuerdo al punto de vista de la imagen asociada, estando caracterizado este procedimiento por que comprende las siguientes etapas:

- 10
- adquisición de una primera imagen que representa el objeto en una posición inicial determinada;
 - creación de una primera imagen clave a partir de la primera imagen adquirida y la postura relativa del objeto en su entorno;
 - 15 - adquisición de al menos una segunda imagen que representa dicho objeto, siendo el punto de vista de dicha al menos una segunda imagen diferente al punto de vista de dicha primera imagen;
 - determinación de la postura relativa del objeto en su entorno de acuerdo a la diferencia de los puntos de vista de la primera imagen y de la dicha al menos una segunda imagen, estando determinado cada uno de dichos puntos de vista con respecto a una posición y una orientación; y
 - 20 - creación de una segunda imagen clave a partir de la al menos una segunda imagen adquirida y la postura relativa del objeto en su entorno.

25 **[0019]** Por lo tanto, el procedimiento de la invención permite automatizar la creación de una pluralidad de imágenes clave, en particular, con vistas a inicializar o reinicializar una aplicación de realidad aumentada utilizando un rastreo automático en tiempo real de objetos geométricos tridimensionales, sin marcador, en un flujo de vídeo. Una multitud de imágenes clave puede en efecto permitir que la aplicación sea inicializada para cualquier tipo de postura relativa entre el objeto a rastrear y la cámara.

30 **[0020]** De acuerdo con una característica particular, la adquisición de al menos una segunda imagen que representa dicho objeto se efectúa por medio de una aplicación de rastreo.

[0021] De acuerdo con otra característica particular, el objeto es, al menos en parte, un objeto real.

[0022] De acuerdo con otra característica particular, el objeto es, al menos en parte, un objeto virtual.

35 **[0023]** De acuerdo con una característica particular, el objeto virtual es una representación de un objeto real de acuerdo con un modelo virtual.

[0024] De acuerdo con otra característica particular, el objeto comprende al menos una parte del entorno.

40 **[0025]** De acuerdo con un modo de realización, los puntos de vista de las imágenes pertenecen a un conjunto de puntos predeterminados.

45 **[0026]** De acuerdo con esta característica, la construcción de imágenes clave se efectúa de acuerdo con un campo de visión determinado.

[0027] De acuerdo con un modo de realización particular, las etapas del procedimiento de creación se repiten para al menos una parte del objeto.

50 **[0028]** La invención también está orientada a un dispositivo para la creación de al menos dos imágenes clave que comprenden, cada una, una imagen que representa al menos un objeto tridimensional en un entorno tridimensional y la postura del objeto en ese entorno de acuerdo con el punto de vista de la imagen asociada, estando caracterizado este dispositivo por que comprende:

- 55
- medios para la adquisición de una primera imagen que representa el objeto en una posición inicial determinada;
 - medios para la creación de una primera imagen clave a partir de la primera imagen adquirida y la postura relativa del objeto en su entorno;
 - medios para la adquisición de al menos una segunda imagen que representa dicho objeto, siendo el punto de vista de dicha al menos una segunda imagen diferente al punto de vista de dicha primera imagen;
 - 60 - medios para determinar la posición relativa del objeto en su entorno de acuerdo con la diferencia de los puntos de vista de la primera imagen y de dicha al menos una segunda imagen, determinándose cada uno de dichos puntos de vista con respecto a una posición y una orientación; y
 - medios para la creación de una segunda imagen clave a partir de dicha al menos una segunda imagen adquirida y de la postura relativa del objeto en su entorno.

65 **[0029]** Este dispositivo tiene las mismas ventajas que el procedimiento brevemente descrito anteriormente, y por consiguiente no se repetirán aquí.

[0030] La presente invención también está orientada a un medio de almacenamiento, eventualmente parcial o totalmente extraíble, legible por un ordenador o un microprocesador, que contiene instrucciones de código de un programa de ordenador para ejecutar etapas del procedimiento como se describe más arriba.

[0031] La presente invención se orienta finalmente a un programa de ordenador que incluye instrucciones adaptadas para ejecutar cada una de las etapas del procedimiento como se describe más arriba.

[0032] Otras ventajas, objetos y características de la presente invención surgen de la siguiente descripción detallada, dada a modo de ejemplo no limitativo, con referencia a los dibujos adjuntos en los cuales:

- la figura 1 representa esquemáticamente los principios esenciales de la aplicación de rastreo de objetos desarrollada por la Escuela Politécnica Federal de Lausana;
- la figura 2 ilustra ciertas etapas del procedimiento para determinar la postura de un objeto en una imagen de un flujo de vídeo a partir de imágenes clave y de la imagen precedente del flujo de vídeo;
- la figura 3 representa un esquema global de la creación de imágenes clave de un objeto tridimensional y de geometría cualquiera, en un entorno que utiliza la invención;
- la figura 4 muestra un ejemplo de dispositivo que permite implementar al menos parte de la invención;
- la figura 5 ilustra un ejemplo de aprendizaje de imágenes clave de un motor de vehículo;
- la figura 6 ilustra un algoritmo para la inicialización automática de una aplicación de rastreo a partir de imágenes clave creadas a partir de objetos tridimensionales reales o virtuales de acuerdo con la invención;
- la figura 7 ilustra un gráfico de estados-transiciones para rastrear un objeto entre una pluralidad de objetos de objetivo;
- la figura 8 ilustra un algoritmo de rastreo en un paisaje;
- la figura 9 ilustra un algoritmo para crear imágenes clave de un paisaje de acuerdo con la invención;
- la figura 10 ilustra un trayecto para aprender un paisaje; y
- la figura 11 ilustra un algoritmo de rastreo en un paisaje de acuerdo con la invención.

[0033] En particular, el procedimiento según la invención tiene como objeto la creación, en especial de manera automática, de imágenes clave de un objeto tridimensional en un entorno con vistas a automatizar las fases de inicialización y reinicialización después de una desincronización de la aplicación de seguimiento de objetos en imágenes de una transmisión de vídeo. Una multitud de imágenes clave pueden permitir que la aplicación se inicialice para cualquier tipo de posturas relativas entre el objeto a rastrear y la cámara. La figura 3 ilustra un esquema global de creación de imágenes clave, en especial imágenes clave "fuera de línea", de un objeto en un entorno que utiliza la invención, para una aplicación de rastreo de objetos.

[0034] Como se muestra en la figura 3, la creación de imágenes clave de un objeto en un entorno y la ejecución de una aplicación de rastreo (300) que utiliza esas imágenes clave comprende cuatro fases: una fase de creación de una primera imagen clave (I), una fase de creación automática de las imágenes clave posteriores (II), una fase de inicialización del rastreo que utiliza la o las imágenes creadas previamente (III), y una fase de rastreo de objetos (IV) que corresponde al régimen permanente de la aplicación.

[0035] La fase de creación una primera imagen clave (I) consiste principalmente en la adquisición de una primera imagen que representa el objeto tridimensional en una posición inicial. Esta adquisición se efectúa, en particular, usando un medio de toma de imágenes tal como una cámara de vídeo o una cámara de fotos. Habiendo adquirido la imagen del objeto tridimensional (etapa 305), se crea una primera imagen clave (etapa 310) que comprende, por un lado, la primera imagen adquirida y, por otro lado, la postura relativa del objeto en el entorno según el punto de vista de la imagen.

[0036] Según la técnica anterior, para construir una imagen clave, conviene colocar manualmente la malla tridimensional correspondiente al objeto sobre éste en la imagen. Sin embargo, esta etapa es fastidiosa. Sin embargo, conocer el tipo de aplicación puede reducir y simplificar la creación de una imagen clave (etapa 310). Por ejemplo, será relevante usar un sensor de orientación para aplicaciones en las que el desplazamiento de la cámara está restringido en rotación, o usar un modelo virtual texturizado del objeto en una posición conocida.

[0037] Para que el algoritmo de rastreo sea más robusto, a veces es importante capturar una serie de imágenes clave, correspondientes a una serie de posturas relativas entre la cámara y el objeto. Durante la fase de creación de esas imágenes clave posteriores (II), una primera etapa consiste en adquirir una nueva imagen que representa el objeto (etapa 315), siendo el punto de vista de la segunda imagen diferente al punto de vista de la primera imagen. Entonces, se determina la postura relativa del objeto en su entorno de acuerdo con la diferencia de los puntos de vista de las imágenes (etapa 320), determinándose cada uno de dichos puntos de vista en relación con una posición y una orientación. Esta etapa puede realizarse de varias maneras. En primer lugar, si está disponible el modelo virtual tridimensional texturizado del objeto a rastrear, es posible crear estas nuevas imágenes clave variando los parámetros de postura del objeto delante de la cámara. También es particularmente interesante utilizar la aplicación de rastreo (335) para generar nuevas imágenes clave. Así, las nuevas imágenes clave creadas en línea podrán reutilizarse para mejorar la calidad de la inicialización del algoritmo de rastreo. Finalmente, se crea una nueva imagen clave (325) a partir de cada nueva imagen y de la postura relativa del objeto en su entorno.

[0038] Las etapas de esta fase se repiten para crear una pluralidad de imágenes clave.

[0039] Durante la fase de inicialización (III), a partir de todas las imágenes clave que se crean durante la fase I y la fase II, la aplicación de rastreo se inicializa buscando una imagen clave que representa el objeto en el flujo de vídeo que contiene el objeto a rastrear (etapa 330) y más cercano a la configuración actual (posición relativa entre la cámara y el objeto).

[0040] Cuando se determina la postura del objeto en la primera imagen y se selecciona la imagen clave actual (imagen clave determinada durante la fase de inicialización) (etapa 330), la aplicación de rastreo puede encontrar el objeto (fase IV) en las imágenes sucesivas del flujo de vídeo, usando un mecanismo de rastreo (etapa 335). En este mecanismo, los desplazamientos del objeto (desplazamiento del objeto en la escena o desplazamiento inducido por el desplazamiento de la cámara en la escena) se siguen por una trama a otra y las eventuales derivas se compensan gracias a las informaciones contenidas en la imagen clave fuera de línea retenida en el momento de la inicialización y, eventualmente, en la imagen clave en línea calculada durante la ejecución de la aplicación (imagen clave que podrá servir posteriormente como imagen clave fuera de línea para inicializar la aplicación automáticamente), con la pequeña diferencia que las nuevas imágenes clave creadas en línea pueden ser utilizadas para esta etapa de reinicialización.

[0041] Cuando la medida del error se vuelve demasiado importante, el rastreo se desincroniza y es necesaria una fase de reinicialización. La fase de reinicialización es similar a la fase de inicialización descrita anteriormente (etapa 330).

[0042] La figura 4 representa esquemáticamente un dispositivo adaptado para implementar la invención. El dispositivo 400 es, por ejemplo, un micro-ordenador, una estación de trabajo o una consola de juegos.

[0043] El dispositivo 400 incluye, preferiblemente, un bus de comunicación 402 al que están conectadas:

- una unidad procesadora central o microprocesador 404 (CPU, *Central Processing Unit*);
- una memoria de sólo lectura 406 (ROM, *Read Only Memory*) que puede contener el sistema operativo y los programas tales como "Prog";
- una memoria de acceso aleatorio o memoria caché 408 (RAM, *Random Access Memory*) que incluye registros adaptados para almacenar variables y parámetros creados y modificados durante la ejecución de los programas anteriormente mencionados;
- una tarjeta de adquisición de vídeo 410 conectada a una cámara 412; y
- una tarjeta gráfica 416 conectada a una pantalla o a un proyector 418.

[0044] Opcionalmente, el dispositivo 400 también puede disponer de los elementos siguientes:

- un disco duro 420 que puede contener los programas "Prog" antes mencionados y datos procesados o que se procesarán de acuerdo con la invención;
- un teclado 422 y un ratón 424 o cualquier otro dispositivo señalador, tal como un lápiz óptico, una pantalla táctil o un control remoto que permite al usuario interactuar con los programas según la invención;
- una interfaz de comunicación 426 conectada a una red de comunicación distribuida 428, por ejemplo la red de Internet, estando adaptada la interfaz para transmitir y recibir datos;
- una tarjeta de adquisición de datos 414 conectada a un sensor (no mostrado); y
- un lector de tarjetas de memoria (no mostrado) adaptado para leer o escribir en las mismas datos procesados o que se procesarán de acuerdo a la invención.

[0045] El bus de comunicación permite la comunicación y la interoperabilidad entre los diferentes elementos incluidos en el dispositivo 400 o conectados al mismo. La representación del bus no es limitativa y, en especial, la unidad central puede comunicar instrucciones a cualquier elemento del dispositivo 400 directamente o mediante otro elemento del dispositivo 400.

[0046] El código ejecutable de cada programa que permite que el dispositivo programable implemente los procedimientos de la invención, se puede almacenar, por ejemplo, en el disco duro 420 o en la memoria de sólo lectura 406.

5 **[0047]** Según una variante, el código ejecutable de los programas podrá ser recibido a través de la red de comunicación 428, a través de la interfaz 426, para ser almacenado de forma idéntica a la descrita anteriormente.

10 **[0048]** Las tarjetas de memoria pueden reemplazarse por cualquier soporte de información tal como, por ejemplo, un disco compacto (CD-ROM o DVD). Como regla general, las tarjetas de memoria se pueden reemplazar por medios de almacenamiento de información, legibles por un ordenador o por un microprocesador, integrados o no en el dispositivo, posiblemente extraíbles, y adaptados para almacenar uno o más programas cuya ejecución permite llevar a cabo el procedimiento según la invención.

15 **[0049]** De manera más general, el o los programas se podrán cargar en uno de los medios de almacenamiento del dispositivo 400 antes de ejecutarse.

20 **[0050]** La unidad procesadora central 404 controla y dirige la ejecución de las instrucciones o partes de código de software del programa o de los programas según la invención, instrucciones que se almacenan en el disco duro 420 o en la memoria de sólo lectura 406 o en los otros elementos de almacenamiento antes citados. Al encenderse, el o los programas que están almacenados en una memoria no volátil, por ejemplo el disco duro 420 o la memoria de sólo lectura 406, se transfieren a la memoria de acceso aleatorio 408 que contiene el código ejecutable del programa o de los programas según la invención, así como registros para almacenar las variables y parámetros necesarios para la implementación de la invención.

25 **[0051]** Conviene notar que el dispositivo de comunicación que incluye el dispositivo según la invención puede ser igualmente un dispositivo programado. Ese dispositivo contiene entonces el código del programa o los programas informáticos, por ejemplo, programado de forma fija en un circuito integrado específico de la aplicación (ASIC).

30 **[0052]** Alternativamente, la imagen de la tarjeta de vídeo 416 se puede transmitir a la pantalla o proyector 418 a través de la interfaz de comunicación 426 y la red de comunicación distribuida 428. Asimismo, la cámara 412 puede conectarse a una tarjeta de adquisición de vídeo 410 independiente del dispositivo 400 de manera que las imágenes de la cámara 412 se transmitan al dispositivo 400 a través de la red de comunicación distribuida 428 y la interfaz de comunicación 426.

35 **[0053]** Debido a la simplificación de la implementación proporcionada por el procedimiento de la invención, la creación de imágenes clave se puede llevar a cabo sin recurrir a un especialista. Después de la creación de un conjunto de imágenes clave, se puede inicializar una aplicación de rastreo sobre la base de ese conjunto y utilizarla de forma estándar para rastrear un objeto en una secuencia de imágenes procedentes de un flujo de vídeo, por ejemplo, para insertar una secuencia de vídeo en un objeto de la escena teniendo en cuenta la posición y la orientación de ese objeto, pero también para determinar el movimiento de una cámara de acuerdo con el análisis de un objeto de la escena. En ese caso, el objeto forma parte del decorado y encontrar la postura de ese objeto en la escena equivale por tanto a encontrar la postura de la cámara con respecto a él. Resulta posible entonces agregar elementos virtuales a la escena siempre y cuando se conozca la transformación geométrica entre el objeto y el modelo geométrico de la escena. Ese es el caso. Este enfoque, por lo tanto, permite aumentar la escena real con objetos virtuales animados que se desplazan en función de la geometría de la escena.

50 **[0054]** En particular, según un primer ejemplo de realización, la aplicación puede consistir en estimar la postura de un objeto tridimensional, por ejemplo, un motor en un vehículo, y agregar informaciones a ese objeto para proporcionar al usuario informaciones relativas a ese objeto, por ejemplo, el montaje y desmontaje del motor.

55 **[0055]** Para ello, la aplicación requiere aprender una serie de imágenes clave que luego le permitirán la inicialización automática de la aplicación de rastreo en la imagen. Teniendo en cuenta que se conoce aproximadamente la posición del usuario, se conoce de paso la posición de la cámara con respecto al objeto tridimensional a rastrear. Por lo tanto, la creación de las imágenes clave (fases I y II) y la inicialización (fase III) se simplifican por el hecho de que se conoce la posición del usuario con respecto al objeto tridimensional y que se requiere un número pequeño de imágenes clave para hacer que la inicialización sea totalmente automática. En efecto, la posición, que corresponde en particular a la altura de una persona, es decir aproximadamente 1,70 metros, y una orientación más o menos encarada al objeto tridimensional, permite el uso de unas pocas imágenes clave pertinentes en esa área de visualización. La figura 5 ilustra esquemáticamente el área de rodaje para producir imágenes clave.

60 **[0056]** En un primer modo de realización, para permitir la inicialización automática del sistema de rastreo de objetos tridimensionales, es necesaria una fase de aprendizaje que consiste en adquirir una serie de imágenes clave que contienen el objeto tridimensional en el área de rodaje del usuario.

65 **[0057]** El área de rodaje se puede adaptar en función del uso.

- 5 [0058] Según un segundo modo de realización, la fase de aprendizaje de las imágenes clave se lleva a cabo mediante modelos sintéticos tridimensionales texturizados del objeto, por ejemplo, el motor y el automóvil. Según este modo de realización, el aprendizaje automático se efectúa haciendo variar los ángulos θ y φ correspondientes a los ángulos vertical y horizontal de la cámara (real o virtual) y la distancia de la cámara con respecto al modelo tridimensional del objeto.
- 10 [0059] En efecto, si se dispone del modelo tridimensional texturizado de la geometría del objeto, por ejemplo, del automóvil y su motor, este modelo se puede desplazar en el campo de visión de la cámara. Se obtiene entonces una serie tridimensional de representaciones del objeto mediante imágenes, que corresponde del objeto visto desde diferentes ángulos. Estas imágenes también están asociadas a la postura conocida del objeto, determinada a partir de los parámetros θ , φ y a partir de la distancia de la cámara con respecto al objeto. Así, se crea un conjunto de imágenes clave de forma totalmente automática (fases I y II).
- 15 [0060] Según este modo de realización, la fase de aprendizaje es por lo tanto poco costosa y permite en especial preparar el mantenimiento de un producto incluso antes de su fabricación. El CAD (“Diseño Asistido por Ordenador”) del producto generalmente se modela incluso antes de que se construya realmente.
- 20 [0061] Además, el modelo de CAD texturizado de un objeto, utilizado en diferentes condiciones de iluminación puede aproximarse a las condiciones del entorno real.
- [0062] La siguiente fase (fase III) es la de la inicialización automática a partir de las imágenes clave.
- 25 [0063] Esta inicialización consiste en una primera estimación de la postura del objeto tridimensional. En el ejemplo considerado aquí, la postura del motor se calcula en el momento en que el usuario abre el capó del automóvil. En ese momento, muchos puntos de interés de la imagen actual están correlacionados con una de las imágenes clave.
- 30 [0064] La Figura 6 ilustra la inicialización automática de las imágenes clave de objetos tridimensionales reales o virtuales.
- [0065] Según un segundo modo de realización, la aplicación puede consistir en encontrar la postura de un objeto en una pluralidad de objetos de destino, que pueden activarse para ser rastreados. Un objeto de destino puede activarse manualmente, en especial mediante una interfaz de tipo OCX (Extensión de control de OLE (*Object linking and Embedding*) en terminología anglosajona).
- 35 [0066] Considérese una máquina herramienta, por ejemplo.
- [0067] En la aplicación actual, por ejemplo, la posición distante de la cámara da la postura del conjunto de la máquina mientras que, de cerca, sólo se rastreará una parte precisa de la máquina herramienta, por ejemplo, un cerramiento eléctrico, primero desde el punto de vista exterior y luego desde el punto de vista interior.
- 40 [0068] Según este modo de realización, el uso de autómatas de tipo transición de estados permite cambiar las imágenes clave en función del estado actual y rastrear la postura de diferentes objetos.
- 45 [0069] La figura 7 ilustra un autómata de estados de transición para el ejemplo considerado.
- [0070] Las transiciones de un estado a otro se desencadenan cuando el usuario se acerca a un nuevo objeto a rastrear y ordena, por ejemplo pulsando un botón, un cambio del objeto de destino a rastrear.
- 50 [0071] La aplicación de rastreo pasa entonces a una modalidad desincronizada intermedia, equivalente a una modalidad de inicialización, que requiere el uso de una nueva serie de imágenes clave correspondientes al objeto de destino seleccionado y asociado al nuevo estado del autómata.
- [0072] Según un tercer modo de realización, la aplicación puede consistir en encontrar la posición del eje de visión de un usuario. En efecto, en una aplicación tal como un simulador de tiro, según este modo de realización, la aplicación puede consistir en encontrar la posición del eje de visión del cañón, en especial haciendo coincidir los puntos de interés entre la imagen actual y el paisaje del simulador de tiro.
- 55 [0073] En este modo de realización, la cámara tiene una posición estática porque está fijada a un trípode fijo y las condiciones exteriores son constantes. En efecto, el paisaje del simulador de tiro no incluye ningún cambio de clima ni la aparición repentina de objetos grandes que se desplazan y enmascaran el paisaje.
- 60 [0074] Según este modo de realización, la aplicación incluye en especial dos etapas que se ilustran en la figura 8, es decir, una etapa de aprendizaje del paisaje (etapa 800) que consiste en la creación de una pluralidad de imágenes clave del paisaje (fases I y II), asociadas a valores de los parámetros de guiñada (“yaw”), cabeceo (“pitch”) y balanceo (“roll”), determinados mediante un sensor, por ejemplo, y una etapa de rastreo en el paisaje
- 65

(etapa 805) en la que las imágenes clave almacenadas previamente se usan para hacer coincidir y estimar la postura actual de la cámara. Este enfoque es posible al conocer el tipo de aplicación. En efecto, en este modo de realización, el movimiento de la cámara está restringido a rotaciones, lo que permite el uso de un sensor.

- 5 **[0075]** Según este modo de realización, la aplicación de rastreo está acoplada a un sensor de movimiento, por ejemplo, un sensor del tipo Mti de la empresa MTI Instrument o un sensor del tipo inerciaCube3 de la empresa Intersense. A pesar de que este sensor no es suficientemente preciso para encontrar una orientación precisa en la imagen, puede corregir significativamente el rastreo en caso de error.
- 10 **[0076]** La fase de aprendizaje consiste en desplazar la cámara alrededor de su eje vertical, definido en especial por el trípode, y adquirir un cierto número de imágenes clave en función de la orientación de la cámara según los parámetros de guiñada (“*yaw*”), cabeceo (“*pitch*”) y balanceo (“*roll*”).
- 15 **[0077]** Esta fase tiene la característica particular de utilizar el algoritmo de rastreo de objetos para crear nuevas imágenes clave (etapas 325).
- [0078]** La figura 9 ilustra este algoritmo.
- 20 **[0079]** El algoritmo comienza en una posición "0" adquiriendo una primera imagen clave y buscando puntos de interés en esa primera imagen. Entonces, la cámara gira según el parámetro de guiñada o de inclinación. La aplicación de rastreo permite entonces encontrar las coincidencias entre los puntos de interés de la primera imagen clave adquirida en la posición "0" y la imagen actual procedente del flujo de vídeo.
- 25 **[0080]** Cuando el número de coincidencias se hace demasiado pequeño, se adquiere una nueva imagen en la posición actual según los parámetros de guiñada y cabeceo. Esta nueva imagen se convierte en la imagen clave.
- [0081]** Para limitar las derivas en este tipo de aprendizaje, la cámara puede regresar regularmente a su posición inicial.
- 30 **[0082]** En la figura 10 se ilustra un ejemplo de trayecto para aprender un paisaje que permite en especial evitar las derivas.
- [0083]** A continuación se describe un algoritmo de rastreo de este modo de realización, ilustrado en la figura 11.
- 35 **[0084]** Durante la ejecución de la aplicación de rastreo, las informaciones previamente aprendidas se utiliza para encontrar la imagen clave fuera de línea para la inicialización automática de la fase de rastreo. Esta inicialización (etapa 1100) puede llevarse a cabo en especial mediante un sensor, por ejemplo un sensor del tipo Mti que proporciona las informaciones de parámetros de orientación de la cámara.
- 40 **[0085]** Después de esta etapa de inicialización, se ejecuta la aplicación de rastreo (etapa 1105).
- [0086]** Sin embargo, el rastreo de un objeto puede desincronizarse (etapa 1110), por ejemplo, si el clima cambia o si el usuario gira la cámara rápidamente o si el usuario sale del área de paisaje aprendida. En este caso, la aplicación de rastreo pasa a la modalidad de respaldo para prolongar la ilusión de rastreo durante el cual, por ejemplo, el sensor MTi decidirá la orientación actual de la cámara.
- 45 **[0087]** El algoritmo de rastreo intenta entonces volver a sincronizarse efectuando, por ejemplo, una inicialización automática utilizando parámetros de orientación enviados por el sensor.
- 50 **[0088]** En la modalidad desincronizada, el rastreo se mantiene pues mediante el uso del sensor, en especial, del sensor Mti. Siendo la precisión de tal sensor de poca importancia, se llevan a cabo intentos de reinicialización.
- [0089]** Durante la ejecución de la aplicación de rastreo, los desplazamientos de la cámara permiten encontrar una nueva imagen clave para el rastreo. En efecto, los desplazamientos se reproducen según los parámetros de guiñada y de cabeceo de las orientaciones de la cámara. Cuando el número de puntos de interés entre la imagen actual y la imagen clave ya no es suficiente, se debe usar una nueva imagen clave de acuerdo con los nuevos parámetros de orientación de ese espacio.
- 55 **[0090]** Naturalmente, para satisfacer requisitos particulares, una persona experta en el campo de la invención podrá hacer modificaciones en la descripción anterior.
- 60

REIVINDICACIONES

- 5 1. Procedimiento de creación automática de al menos dos imágenes clave de inicialización, comprendiendo, cada una, una imagen que representa al menos un objeto tridimensional en un entorno tridimensional y la postura del objeto en ese entorno de acuerdo con el punto de vista de la imagen asociada, caracterizándose este procedimiento por que comprende las siguientes etapas:
 - 10 - adquisición de una primera imagen que representa el objeto en una posición inicial determinada;
 - creación de una primera imagen clave a partir de la primera imagen adquirida y de la postura relativa del objeto en su entorno;
 - adquisición de al menos una segunda imagen que representa dicho objeto, siendo el punto de vista de dicha al menos una segunda imagen diferente del punto de vista de dicha primera imagen;
 - 15 - determinación de la postura relativa del objeto en su entorno de acuerdo con la diferencia de los puntos de vista de la primera imagen y de dicha al menos una segunda imagen, determinándose cada uno de dichos puntos de vista con respecto a parámetros de posición y orientación, estando al menos uno de dichos parámetros determinado independientemente de dichas imágenes; y
 - creación de dicha al menos una segunda imagen clave a partir de la al menos una segunda imagen adquirida y de la postura relativa del objeto en su entorno.
- 20 2. Procedimiento de creación según la reivindicación 1, en el que dicho al menos uno de dichos parámetros se determina de acuerdo con una posición predeterminada.
3. Procedimiento de creación según la reivindicación 1 o la reivindicación 2, en el que el cálculo de dicho al menos uno de dichos parámetros se basa en un valor procedente de un sensor de orientación.
- 25 4. Procedimiento de creación según una cualquiera de las reivindicaciones precedentes, caracterizado por que la adquisición de al menos una segunda imagen que representa dicho objeto se realiza por medio de una aplicación de rastreo.
- 30 5. Procedimiento de creación según una cualquiera de las reivindicaciones precedentes, caracterizado por que dicho objeto es al menos en parte un objeto real.
6. Procedimiento de creación según una cualquiera de las reivindicaciones precedentes, caracterizado por que el objeto es al menos en parte un objeto virtual.
- 35 7. Procedimiento de creación según la reivindicación 6, caracterizado por que el objeto virtual es una representación de un objeto real de acuerdo a un modelo virtual.
8. Procedimiento de creación según la reivindicación 6 o la reivindicación 7, caracterizado por que el objeto comprende al menos una parte del entorno.
- 40 9. Procedimiento de creación según una cualquiera de las reivindicaciones precedentes, caracterizado por que los puntos de vista de las imágenes pertenecen a un conjunto de puntos predeterminados.
- 45 10. Procedimiento de creación según una cualquiera de las reivindicaciones precedentes, caracterizado por que las etapas del procedimiento de creación se repiten para al menos una parte del objeto.
- 50 11. Programa de ordenador que incluye instrucciones adaptadas para ejecutar cada una de las etapas del procedimiento de acuerdo con una cualquiera de las reivindicaciones precedentes.
12. Medio de almacenamiento de información, extraíble o no, parcialmente o totalmente legible por un ordenador o un microprocesador, que contiene instrucciones de código de un programa de ordenador para ejecutar cada una de las etapas del procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 10.
- 55 13. Dispositivo para la creación automática de al menos dos imágenes clave de inicialización, comprendiendo, cada una, una imagen que representa al menos un objeto tridimensional en un entorno tridimensional y la postura del objeto en ese entorno de acuerdo con el punto de vista de la imagen asociada, caracterizándose este dispositivo por que comprende:
 - 60 - medios para la adquisición de una primera imagen que representa el objeto en una posición inicial determinada;
 - medios para la creación de una primera imagen clave a partir de la primera imagen adquirida y de la postura relativa del objeto en su entorno;
 - 65 - medios para la adquisición de al menos una segunda imagen que representa dicho objeto, siendo el punto de vista de dicha al menos una segunda imagen diferente del punto de vista de dicha primera imagen;

- 5
- medios para determinar la postura relativa del objeto en su entorno según la diferencia de los puntos de vista de la primera imagen y de dicha al menos una segunda imagen, determinándose cada uno de dichos puntos de vista con respecto a parámetros de posición y de orientación, determinándose al menos uno de dichos parámetros independientemente de dichas imágenes; y
 - medios para la creación de dicha al menos una segunda imagen clave a partir de la al menos una segunda imagen adquirida y de la postura relativa del objeto en su entorno.
- 10
- 14.** Dispositivo de creación de acuerdo con la reivindicación 13, que comprende además medios para adquirir al menos un valor desde un sensor angular, usándose dicho valor para calcular dicho al menos uno de dichos parámetros.
- 15
- 15.** Dispositivo de creación de acuerdo con la reivindicación 13 o la reivindicación 14, caracterizado por que los medios para la adquisición de al menos una segunda imagen que representa dicho objeto están adaptados para adquirir al menos una segunda imagen por medio de una aplicación de rastreo.
- 16.** Dispositivo de creación de acuerdo con una cualquiera de las reivindicaciones 13 a 15, caracterizado por que el objeto es, al menos en parte, un objeto real o, al menos en parte, un objeto virtual.
- 20
- 17.** Dispositivo de creación de acuerdo con una cualquiera de las reivindicaciones 13 a 16, caracterizado por que los puntos de vista de las imágenes pertenecen a un conjunto de puntos predeterminados.

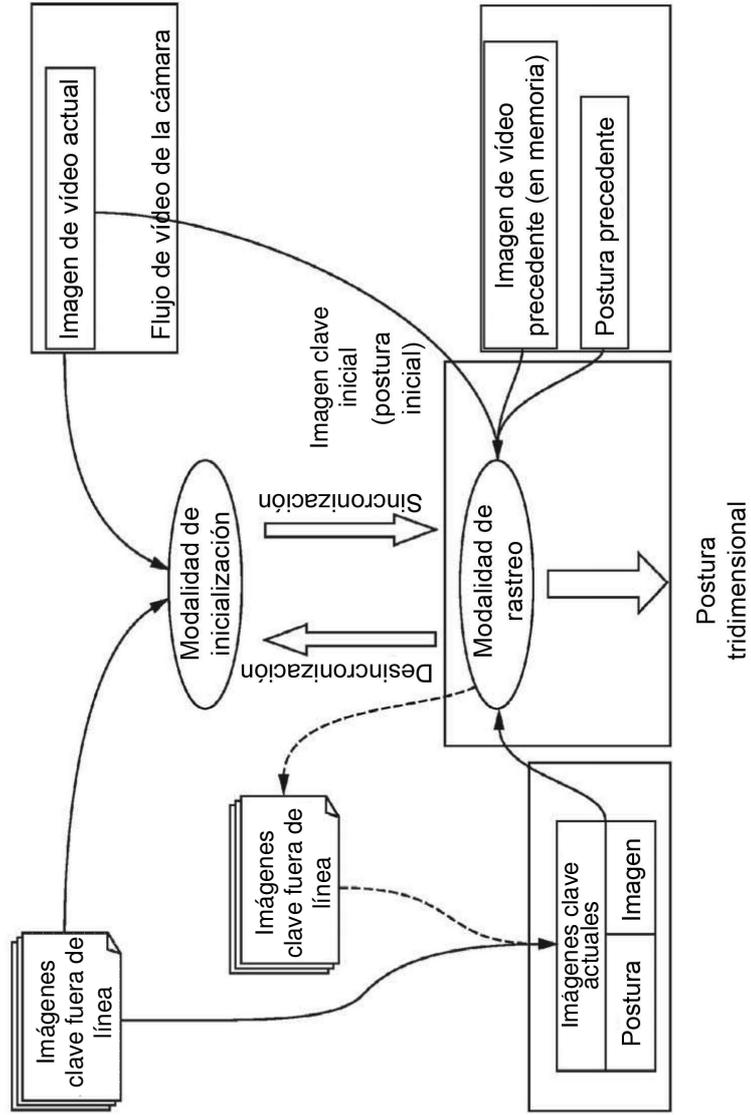


Fig. 1

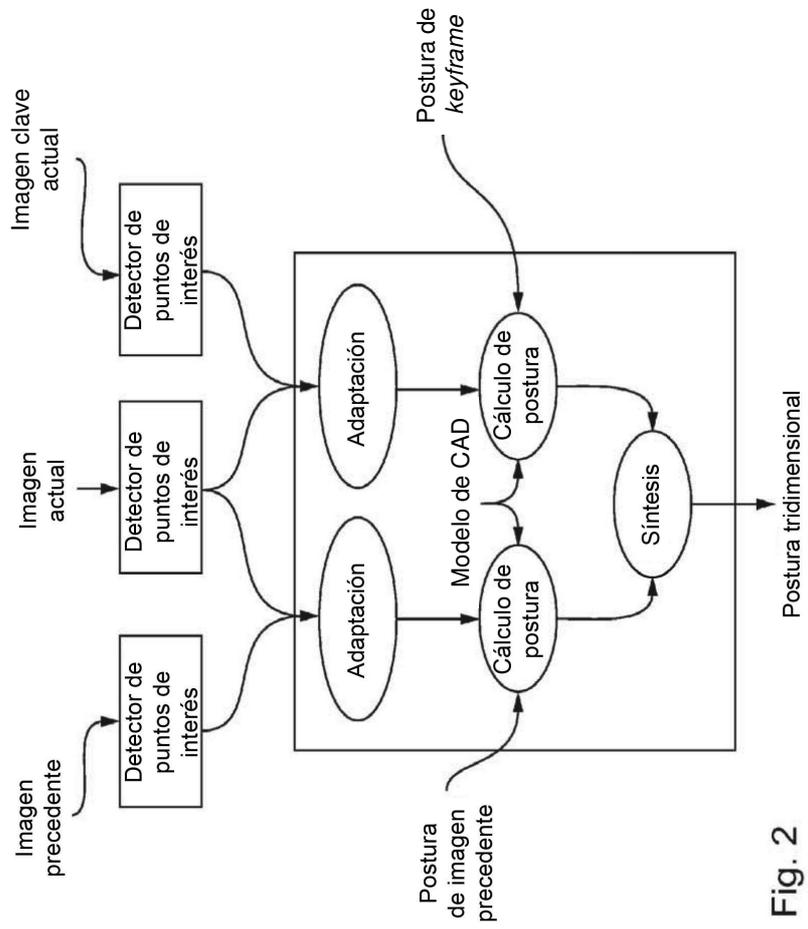


Fig. 2

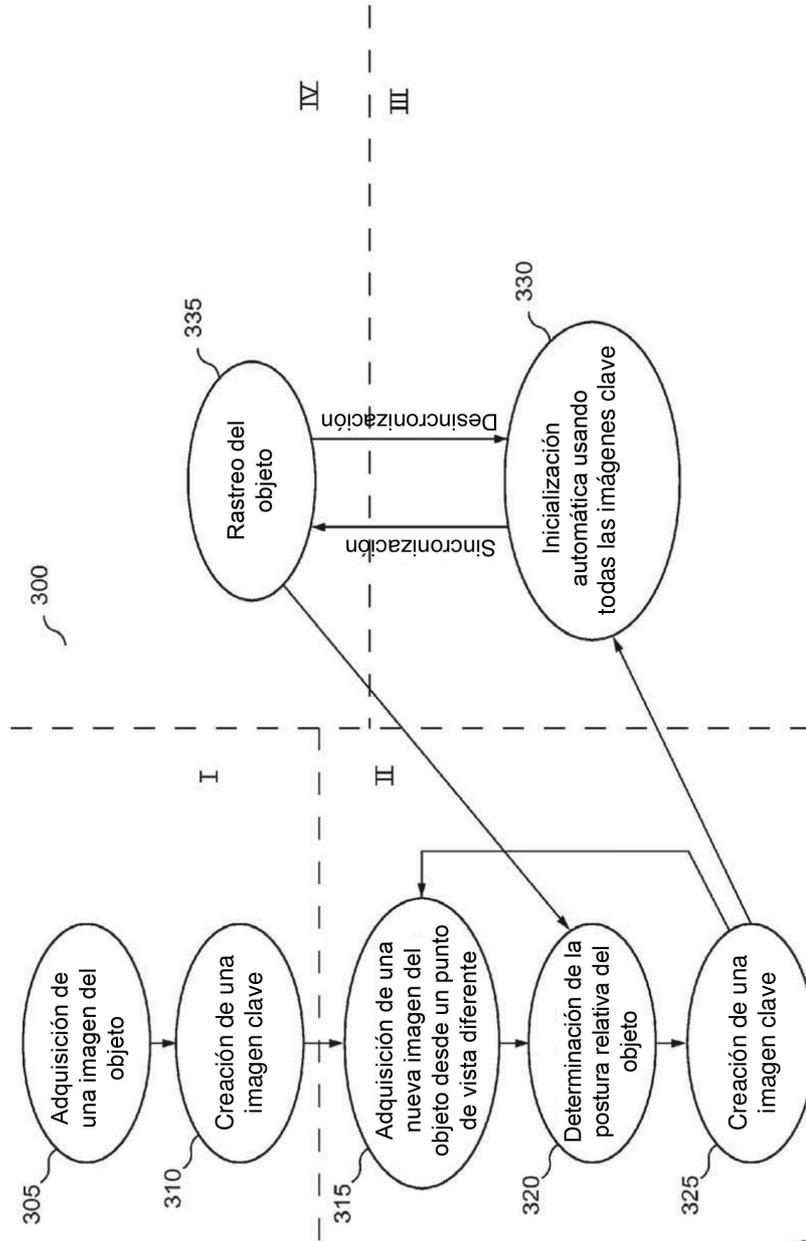


Fig. 3

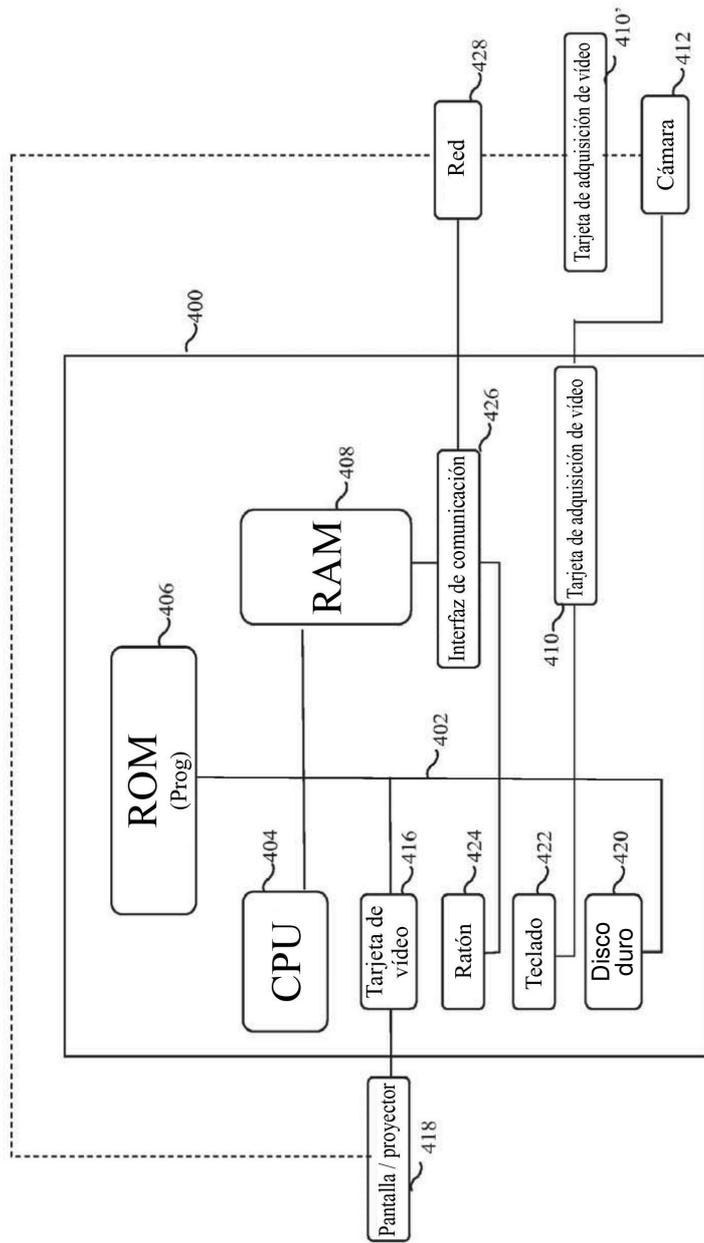


Fig. 4

Fig. 5

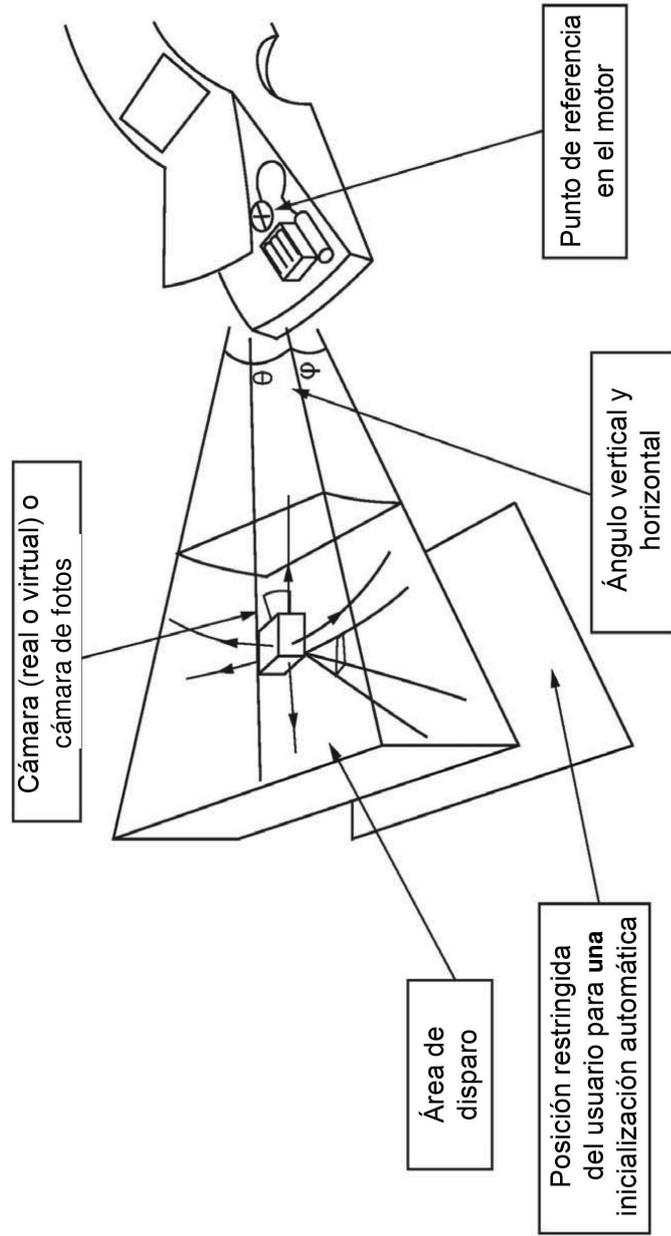
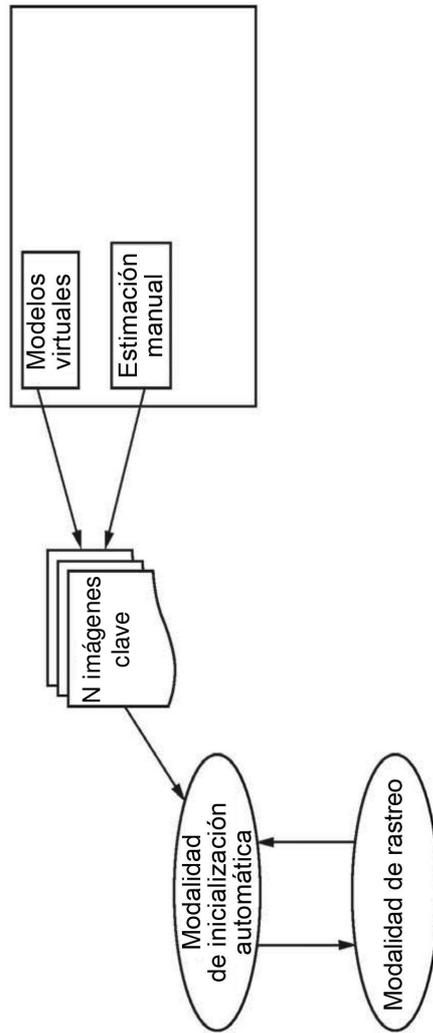


Fig. 6



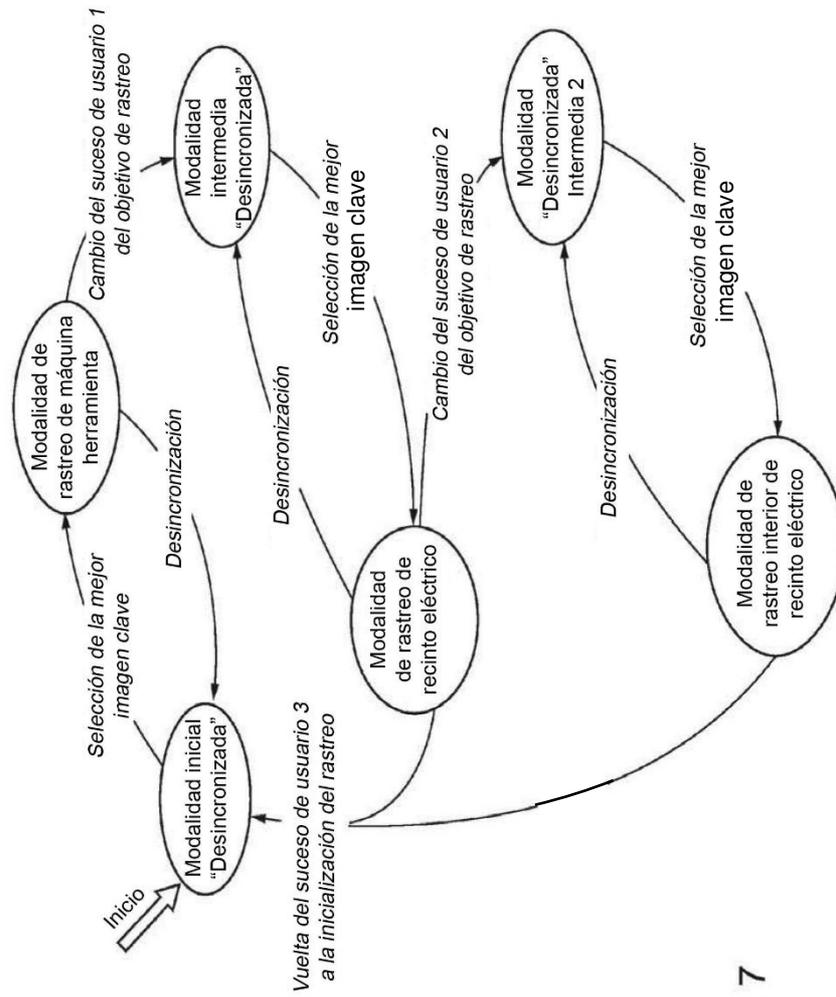


Fig. 7

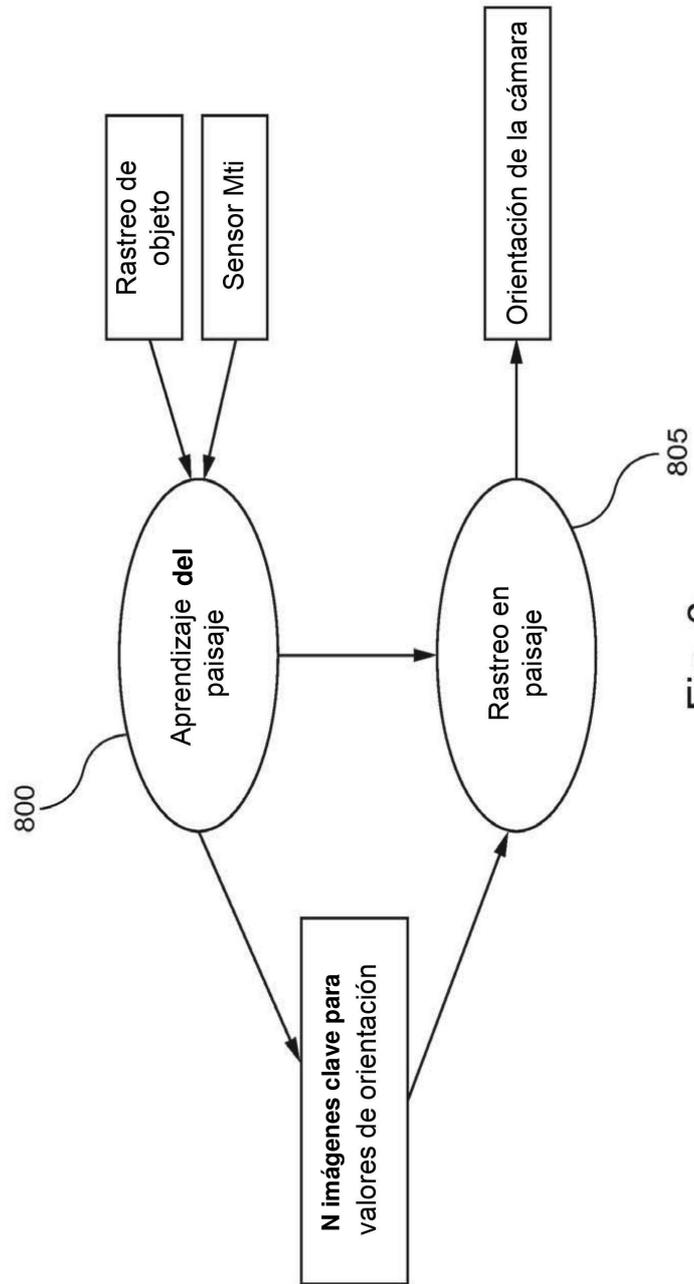


Fig. 8

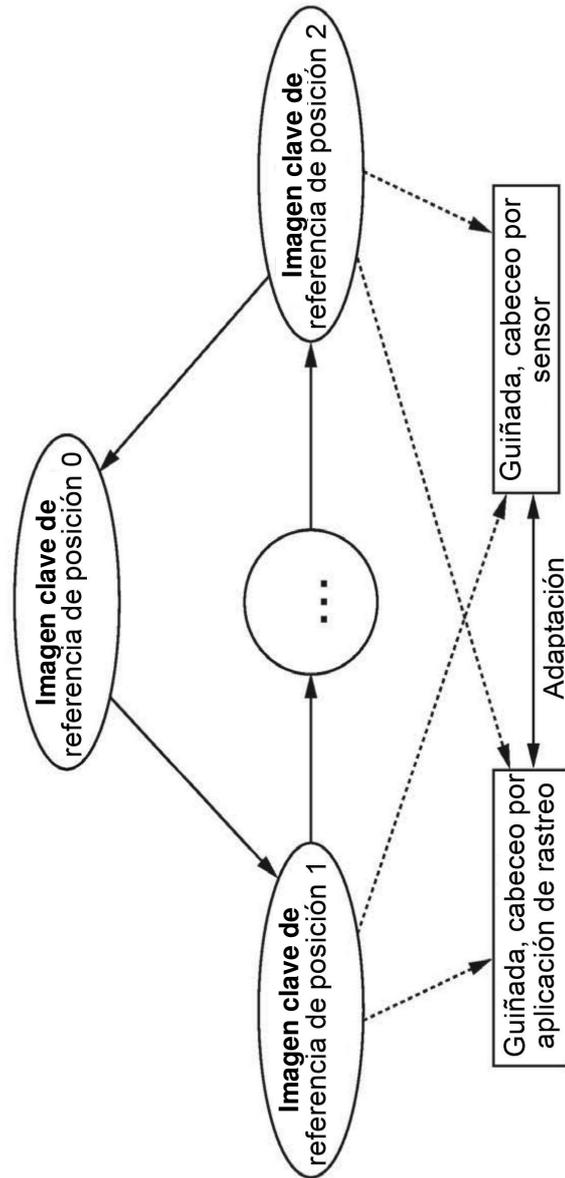


Fig. 9

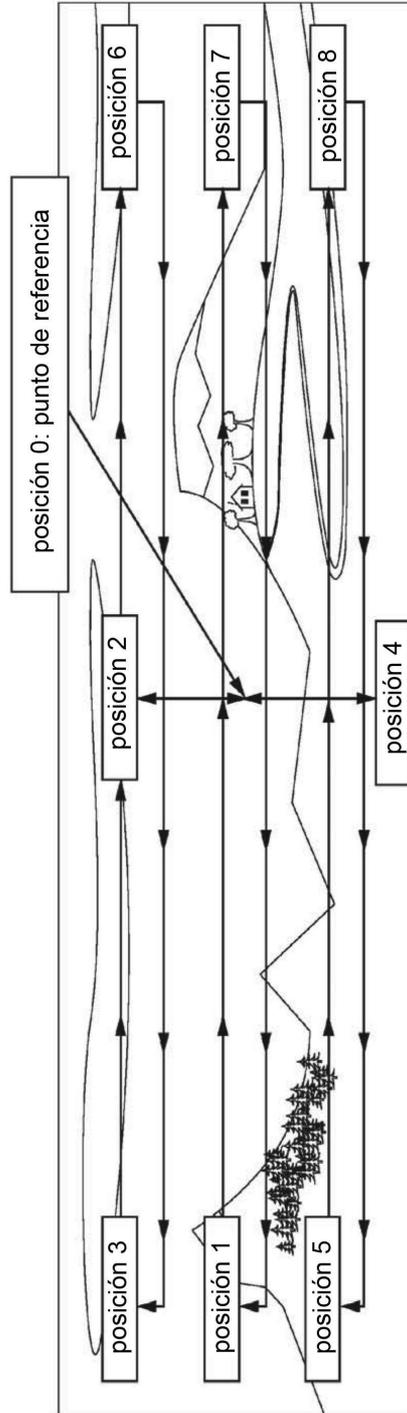


Fig. 10

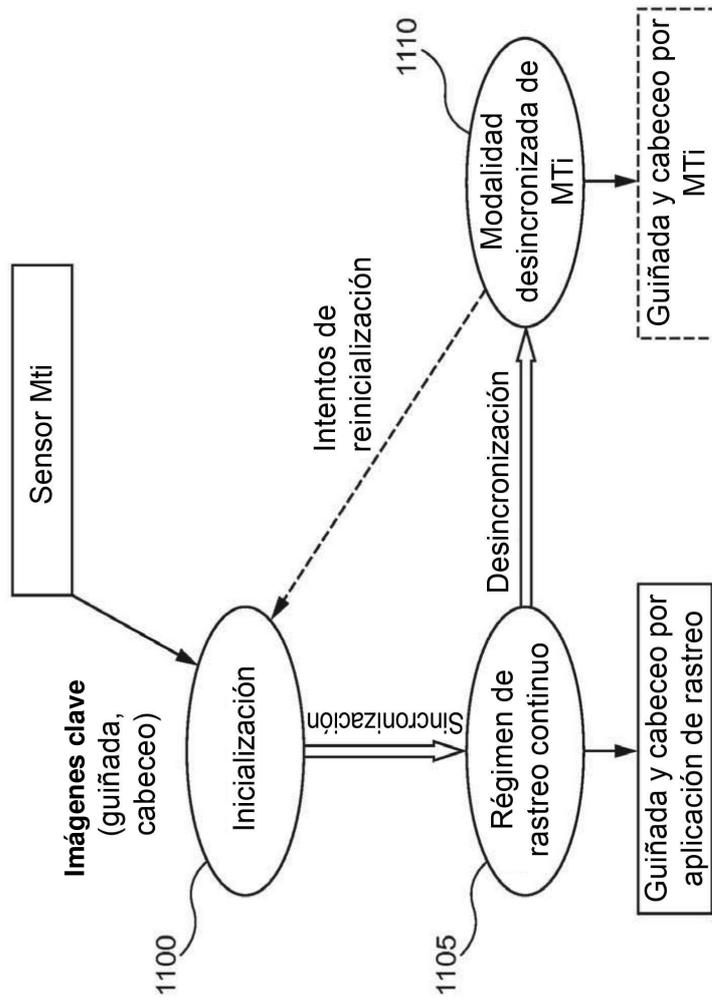


Fig. 11