

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 726 349**

51 Int. Cl.:

H04N 5/232 (2006.01)

G06T 19/00 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.08.2017** E 17188202 (0)

97 Fecha y número de publicación de la concesión europea: **10.04.2019** EP 3291536

54 Título: **Métodos, dispositivos y sistemas para hacer zoom automático cuando se reproduce una escena de realidad aumentada**

30 Prioridad:

06.09.2016 EP 16306115

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

03.10.2019

73 Titular/es:

**INTERDIGITAL CE PATENT HOLDINGS (100.0%)
3 rue du Colonel Moll
75017 Paris, FR**

72 Inventor/es:

**LAURENT, ANTHONY;
JOUET, PIERRICK y
BAILLARD, CAROLINE**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 726 349 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

Métodos, dispositivos y sistemas para hacer zoom automático cuando se reproduce una escena de realidad aumentada

1. Campo técnico

- 5 La presente descripción se relaciona generalmente con el campo de la Realidad Aumentada, y más particularmente con una aplicación de realidad aumentada que se ejecuta en un dispositivo de paso de video.

2. Antecedentes

10 Una aplicación de Realidad Aumentada (AR) superpone objetos artificiales de tres dimensiones (por ejemplo, imágenes generadas por ordenador, también llamadas CGI) encima del mundo real. El realismo y el rendimiento en tiempo real son criterios clave para evaluar aplicaciones de AR. Hay dos tipos de dispositivos de AR: “dispositivos ópticos de ver a través”, como por ejemplo gafas inteligentes y dispositivos de “paso de video”, como teléfonos inteligentes o tabletas por ejemplo. La presente descripción se relaciona con este segundo tipo de dispositivos de AR.

15 Los dispositivos de paso de video son aparatos equipados con una pantalla de presentación, una cámara y medios para que el usuario controle la pose (ubicación y dirección de orientación) de la cámara en tiempo real. Por ejemplo, los teléfonos inteligentes y tabletas son dispositivos de AR que ejecutan una aplicación de AR, mediante el uso de su pantalla frontal y su cámara trasera. En este caso, el usuario controla la pose de la cámara mediante el manejo del dispositivo. El video capturado por la cámara es reproducido sobre la marcha en la pantalla. La aplicación de AR superpone objetos virtuales de una escena de AR en las imágenes del video.

20 En ciertos ángulos de cámara óptimos, el usuario puede ver una escena de AR completa. Sin embargo, si la cámara está demasiado cerca de la ubicación de la escena de AR o si la dirección de orientación de la cámara no está en la dirección óptima, algunos de los objetos de AR quedan fuera del alcance de la cámara. Si la cámara está demasiado lejos de la ubicación de la escena de AR, los objetos de AR son pequeños y los detalles pueden no ser visibles en la pantalla. Para resolver este problema, el usuario tiene que buscar la pose óptima él mismo. Tal solución no es satisfactoria porque el usuario no siempre es consciente de que se está perdiendo objetos o detalles y porque la búsqueda de la pose óptima es una tarea tediosa que molesta el uso agradable de la aplicación de AR.

El documento US-A1-2015 254855 describe un método para proporcionar un nivel de zoom óptimo automático mediante la sugerencia a un usuario para capturar partes que faltan dentro de un límite de capturas de imágenes.

3. Compendio

30 La presente descripción se relaciona con un método para adaptar de manera automática el zoom y la dirección de orientación de la cámara de un dispositivo de paso de video cuando ejecuta una aplicación de Realidad Aumentada. Según la estimación de la pose de la cámara del dispositivo de paso de video y el cálculo de límites de la escena de AR, el zoom óptico o digital de la cámara es adaptado para optimizar la visión de la AR sin que el usuario se mueva. Una posición óptima puede ser indicada al usuario para incitarle a moverse.

35 La presente descripción se relaciona con un método para reproducir imágenes en un dispositivo de presentación cuando se reproduce una escena de realidad aumentada desde el punto de vista de una cámara, el método que comprende:

- determinar límites de la escena de realidad aumentada;
- determinar un coeficiente de zoom según una estimación de una pose de la cámara y los límites de la escena; y
- 40 - reproducir imágenes en el dispositivo de presentación desde la cámara según el coeficiente de zoom.

Según una característica particular, el método además comprende la indicación de una posición de la cámara al usuario mediante el uso de efectos seleccionados a partir de un grupo que comprende efectos visuales, de audio o hápticos.

45 Según una realización específica, la posición es determinada según una dirección y un peso asociado con al menos un elemento de la escena de realidad aumentada.

Según una realización específica, la escena de realidad aumentada comprende elementos no representados correspondientes a objetos de una escena real capturada por la cámara, dichos elementos tomados en cuenta cuando se calculan dichos límites de la escena de realidad aumentada.

Según una característica particular, la cámara está motorizada y el método además comprende controlar la dirección de orientación de la cámara según los límites de la escena de realidad aumentada y el coeficiente de zoom calculado.

- 5 Según una realización específica, la cámara comprende una función de zoom óptico y el método además comprende controlar la función de zoom óptico según el coeficiente de zoom calculado. En otra realización, un zoom digital es operado en las imágenes según el coeficiente de zoom.

La presente descripción también se relaciona con un dispositivo que tiene un elemento de presentación configurado para reproducir imágenes cuando reproduce una escena de realidad aumentada desde el punto de vista de la cámara, el dispositivo comprende una memoria asociada con al menos un procesador configurado para:

- 10 - determinar límites de la escena de realidad aumentada;
 - determinar un coeficiente de zoom según una estimación de una pose de la cámara y los límites de la escena;
 - reproducir imágenes desde la cámara delantera según el coeficiente de zoom.

La presente descripción también se relaciona con un dispositivo que tiene un elemento de presentación configurado para reproducir imágenes desde una cámara cuando reproduce una escena de realidad aumentada, el dispositivo que comprende:

- 15 - medios para determinar los límites de la escena de realidad aumentada;
 - medios para determinar un coeficiente de zoom según una estimación de una pose de la cámara y los límites de la escena;
 - medios para reproducir imágenes desde la cámara delantera según el coeficiente de zoom.

- 20 La presente descripción también se relaciona con un producto de programa informático descargable desde una red de comunicación y/o grabado en un medio legible por un ordenador y/o ejecutable por un procesador, que comprende instrucciones código de programa para implementar los pasos del método anterior para representar imágenes desde una cámara en un dispositivo de presentación cuando reproduce una escena de realidad aumentada cuando es ejecutado en un procesador.

25 **4. Lista de figuras**

La presente descripción se comprenderá mejor, y otras características y ventajas específicas emergerán tras la lectura de la siguiente descripción, la descripción que hace referencia a los dibujos anexados donde:

- La Figura 1 ilustra un ejemplo de una tableta usada como un dispositivo de paso de video, según una realización específica de los presentes principios;
- 30 - La Figura 2 ilustra un ejemplo de un dispositivo de presentación montado en la cabeza (HMD) de ocultación usado como dispositivo de paso de video que ejecuta una aplicación de AR como la tableta de la Figura 1, según una realización específica de los presente principios;
- La Figura 3 ilustra una realización ejemplar de un dispositivo de paso de video que ejecuta una aplicación de AR como en las Figuras 1 y 2, en la cual la cámara está motorizada y lejos del dispositivo de presentación de control, según una realización específica de los presentes principios;
- 35 - La Figura 4 ilustra cálculos basados en el tronco de la cámara de las Figuras 1, 2 y 3 y los límites de la escena de realidad aumentada de las Figuras 2 y 3, según una realización específica de los presentes principios;
- La Figura 5a ilustra un cálculo ejemplar de un coeficiente de disminución de zoom según los límites de la Figura 4, según una realización específica de los presentes principios;
- 40 - La Figura 5b ilustra un cálculo ejemplar de un coeficiente de aumento de zoom según los límites de la Figura 4, según una realización específicas de los presentes principios;
- La Figura 6 ilustra un cálculo ejemplar de una posición 64 óptima para la cámara de las Figuras 1, 2, y 3, según una realización específica de los presentes principios;
- 45 - La Figura 7 muestra una realización de hardware de un aparato configurado para reproducir imágenes desde la cámara de las Figuras 1, 2 y 3 cuando reproduce una escena de realidad aumentada, según una realización específica de los presentes principios;
- La Figura 8 muestra diagramáticamente una realización de un método como se implementa en un dispositivo de procesamiento tal como el dispositivo de la Figura 7 según una realización ventajosa no restrictiva.

5. Descripción detallada de realizaciones

La materia es ahora descrita con referencia a los dibujos, donde numerales de referencia iguales son usados para referirse a elementos iguales a lo largo del documento. En la siguiente descripción, con propósito de explicación, numerosos detalles específicos son expuestos para proporcionar una comprensión completa de la materia. Se comprende que las realizaciones de la materia pueden ser practicadas sin estos detalles específicos.

Según una realización no limitante de la presente descripción, un método y un dispositivo para adaptar de manera automática el zoom y la dirección de orientación de la cámara de un dispositivo de paso de video cuando ejecuta una aplicación de Realidad Aumentada son descritos.

Buscar la pose óptima de la cámara de un dispositivo de paso de video que ejecuta una aplicación de AR consiste en modificar la pose de la cámara. La pose es la combinación de la ubicación en el espacio tridimensional y la dirección de orientación de la cámara. Controlar de manera automática la ubicación de la cámara requiere mecanismos motorizados para desplazar la cámara. Los dispositivos de paso de video no están normalmente equipados con tales mecanismos.

La Figura 1 ilustra un ejemplo de una tableta 10a usada como un dispositivo de paso de video. El dispositivo 10a constituye un equipo que comprende una pantalla 102 y una cámara 101. La cámara 101 filma una escena del mundo que comprende objetos 11 reales, por ejemplo un televisor en la Figura 1a. La escena es la parte del mundo incluida en el tronco 12 de la cámara 101. La escena capturada es presentada en tiempo real en la pantalla 102 de la tableta 10a. Esta funcionalidad hace de la tableta 10a un dispositivo de paso de video. Mediante el manejo de la tableta 10a, un usuario controla la pose de la cámara 101. En efecto, la ubicación y la dirección de orientación de la cámara 101 están directamente relacionadas con la ubicación y la orientación del dispositivo 10a. En una realización particular, la cámara 101 de la tableta 10a es motorizada y la aplicación de AR puede controlar parcialmente la dirección de orientación de la cámara 101. Una aplicación de AR requiere medios para determinar la pose de la cámara. En la Figura 1, un marcador 13 de referencia es presentado en la pantalla del objeto 11 real como un ejemplo de los medios de estimación de pose. En este ejemplo, la aplicación de AR procesa las imágenes capturadas por la cámara 101, reconoce el marcador 13 de referencia y según la forma y el tamaño del marcador, estima la pose (esto es ubicación en un marco de referencia relacionado con el marcador y la dirección de orientación) de la cámara 101 (y como consecuencia del dispositivo 10a). En una variante, la aplicación de AR estima la pose de la cámara según el reconocimiento de los objetos, por ejemplo muebles específicos que se sabe que deben estar en una ubicación particular, en las imágenes. En otra realización, los medios de estimación de pose en un sistema de cámaras de infrarrojos ubicadas en la habitación y que siguen marcadores de infrarrojos punteados en la parte trasera de la tableta 10a. En otra realización, la aplicación usa información a partir de la unidad de mediciones inercial (por ejemplo acelerómetros y giroscopios) de la tableta para estimar la pose del dispositivo 10a. La presente descripción no está limitada a estos ejemplos de sistemas de estimación de pose.

La Figura 2 ilustra un ejemplo de un dispositivo 10b de presentación montado en la cabeza (HMD) de ocultación usado como un dispositivo de paso de video que ejecuta una aplicación de AR. El HMD 10b constituye un equipo que comprende una o más pantallas en frente de los ojos del usuario y al menos una cámara 101 que captura la escena en frente del usuario. El video capturado por la cámara 101 es representado en tiempo real en las pantallas del HMD, que hace del HMD 10a un dispositivo de paso de video. Un sistema de estimación de pose es usado por la aplicación para determinar la ubicación de la dirección de orientación de la cámara. Mientras el usuario mira lo que la cámara 101 está filmando, la pose de la cámara 101 se corresponde con la pose de la mirada del usuario. En la Figura 2, un marcador 13 es presentado en la pantalla de un objeto 11 real visto por la cámara. Como para la tableta 10a de la Figura 1, cualquier tipo de sistema de estimación de pose puede ser empleado. Una vez que la pose de la cámara es determinada, la aplicación de AR añade objetos artificiales en la escena de tres dimensiones. Una imagen de parte de estos objetos incluidos en el tronco 12 de la cámara 101 es superpuesta en la imagen de la escena capturada por la cámara y su composición es representada en las pantallas del dispositivo 10b de paso de video. La ocultación entre objetos (objetos reales y artificiales) es gestionada por la aplicación de AR. Por ejemplo, en la Figura 2, la aplicación de AR añade una Torre Eiffel 21 artificial y un par de dados 22 artificiales en un marco de referencia definido por el marcador y, así, alrededor del televisor 11. El usuario ve que está comprendido en el tronco 12. Así él está viendo el televisor 11, está viendo un parte (y solo una parte) de la Torre Eiffel 21 y no está viendo el par de dados 22 en absoluto ya que están fuera del alcance de la cámara 101.

En la realización ilustrada en la Figura 2, el usuario controla la pose de la cámara mediante el movimiento de su cabeza. En otra realización, la cámara 101 es motorizada y la aplicación de AR puede controlar parcialmente la dirección de orientación de la cámara 101. El sistema de estimación de pose toma la orientación de la cámara 101 en el marco de referencia del HMD en cuenta para estimar la pose.

La Figura 3 ilustra una realización ejemplar de un dispositivo de paso de video que ejecuta una aplicación de AR. Un teléfono inteligente 10c controla una cámara 103 motorizada a través de un enlace 31. El enlace 31 puede ser una interfaz cableada (por ejemplo, una interfaz de bus, una interfaz de red de área ancha, una interfaz de red de área local) o una interfaz inalámbrica (tal como una interfaz IEEE 802.11 o una interfaz Bluetooth®). Por ejemplo, el teléfono inteligente está equipado con una unidad de medición inercial (IMU) y la aplicación de AR sigue los

movimientos del teléfono inteligente y controla la rotación de la cámara según los movimientos detectados. En otra realización, una interfaz de usuario se ejecuta en el teléfono inteligente que permite al usuario controlar la rotación de la cámara. La cámara captura la escena en frente de ella (esto es la parte del mundo real incluido en su tronco) que contiene objetos 11 reales. Las imágenes capturadas por la cámara son transferidas al teléfono inteligente y representadas en tiempo real en la pantalla del teléfono inteligente 10c. Tal sistema constituye un dispositivo de paso de video. La pose de la cámara es conocida en cada momento según sus motores. En una variante, un sistema de estimación de pose exterior, por ejemplo basado en marcadores de referencia, es usado además, por ejemplo si la cámara está equipada con dispositivos de movimiento (por ejemplo ruedas u orugas).

Cuando se ejecuta una aplicación de AR, el dispositivo 10c superpone total o parcialmente objetos 21 artificiales en las imágenes capturadas por la cámara 103. Algunos objetos artificiales de la escena de AR pueden estar fuera del alcance de la cámara, como, por ejemplo, el par de dados 22 en el ejemplo de la Figura 3. Según una realización particular de la presente invención, los límites 32 de la escena de realidad aumentada son calculados. La escena de AR engloba cada objeto 21 y 22 artificial que constituye el contenido de AR de la aplicación de AR. Esto es lo que el usuario tiene que ver. En una variación la escena también contiene objetos artificiales invisibles. Un objeto invisible es un elemento de la escena de AR que no es representado (esto es no se superpone una imagen artificial en la imagen capturada por la cámara para este elemento). Tal objeto invisible puede, por ejemplo, envolver la forma de un objeto real como el televisor 11 de las Figuras 1, 2 y 3. Así, un objeto real puede ser incluido dentro de los límites de la escena sin modificar la imagen capturada por la cámara. Un objeto invisible se corresponde con un volumen del mundo real que el creador de la aplicación de AR quiere incluir en los límites de la escena de AR. Los límites abarcan los elementos del contenido de AR lo más cerca posible. Los límites 32 son determinados de un modo que engloba cada elemento del contenido de AR, que pueden ser objetos artificiales y reales, preferiblemente de un modo que abarque esos elementos lo más cerca posible. En la Figura 3, los límites de la escena de AR son cajas delimitadoras en tres dimensiones del conjunto de elementos del contenido de AR. En variantes, los límites pueden ser una elipse delimitadora del casco convexo de la escena de AR. En una realización particular, los elementos de la escena de AR pueden moverse o cambiar la forma en el tiempo. En una variante, el número de elementos de la escena puede variar en el tiempo. En tal realización, los límites son calculados cada vez que se detecta un cambio en la escena. En una variante, los límites son calculados en una frecuencia regular, por ejemplo cada décima de segundo o cada medio segundo.

La Figura 4 ilustra cálculos basados en el tronco de la cámara y los límites. En base al cálculo de los límites según la ubicación y la dirección de orientación de la cámara en el mundo virtual, determinado según la estimación de pose de la cámara 101, dos rectángulos notables son calculados en el marco virtual de referencia. Se comprende que la cámara está ubicada dentro del mundo virtual de la aplicación de AR según la estimación de pose. Una cámara es reducible, en una escena virtual, a un punto correspondiente al punto focal de la cámara real. Cuando se estima la pose de la cámara 101, el sistema estima, primero, la ubicación del punto focal de la cámara en el marco de referencia de la escena virtual, y segundo, la dirección 41 de orientación de la cámara en el mismo marco de referencia. El tronco 12 referenciado en las Figuras 1, 2 y 4 se basan en la estimación de pose. Una pirámide 42 es determinada, la pirámide que apunta al punto focal de la cámara y engloba todos los límites 32. En el ejemplo de la Figura 4, los límites 32 tienen forma de una caja limitadora y la pirámide 42 es una pirámide de cuatro lados. En variantes de límites, la pirámide 42 puede ser un núcleo o un volumen puntiagudo de cualquier forma. En base a los límites 32 y la pirámide 42, un rectángulo 43 llamado el rectángulo 43 de cara cercana de los límites es calculado. La cara 43 cercana es el rectángulo definido por la unión de las cinco condiciones siguientes:

- El rectángulo 43 es perpendicular a la dirección 41 de orientación;
- El rectángulo 43 es un rectángulo delimitador de la pirámide 42;
- El rectángulo 43 engloba la totalidad de los límites;
- Los bordes del rectángulo 43 son paralelos a los bordes del rectángulo 44 del tronco;
- El rectángulo 43 es el más cercano a la cámara 101 de los rectángulos que cumple las condiciones anteriores.

El rectángulo 44 del tronco es el rectángulo definido por el tronco 12 de la cámara que está en el plano del rectángulo 43 de cara cercana.

La Figura 5a ilustra un cálculo ejemplar de un coeficiente de disminución de zoom. Para una pose dada de la cámara 101, la escena de AR es representada con tamaño máximo si, el rectángulo 43 de cara cercana está incluido en el rectángulo 44 del tronco y si al menos un borde del rectángulo 43 de cara cercana es superpuesto en un borde del rectángulo 44 del tronco. En el ejemplo de la Figura 5a, una parte 51 del rectángulo 43 está dentro del rectángulo 44 del tronco y la parte 52 restante está fuera. Para maximizar la representación de la escena de AR sin mover la cámara 101 (esto es sin mover el centro del rectángulo 44 del tronco), el rectángulo del tronco es alargado a lo largo de su diagonal para hacer que uno de los bordes se superponga al borde externo del rectángulo 43 de cara cercana para obtener un nuevo rectángulo 54 del tronco. Esta operación se corresponde con una disminución de zoom para la cámara 101. El coeficiente de zoom es determinado, por ejemplo, mediante la división de la longitud de la

diagonal del nuevo rectángulo 54 del tronco por la longitud de la diagonal del rectángulo 44 del tronco real. En este ejemplo, el coeficiente del zoom es mayor que uno, así que se corresponde con una disminución de zoom.

La Figura 5b ilustra un cálculo ejemplar de un coeficiente de aumento de zoom. Cuando el rectángulo 43 de cara cercana está totalmente incluido en el rectángulo 44 del tronco, la representación de la escena de AR no es óptima pues puede ser representada en un modo más grande. El rectángulo 44 del tronco es reducido a lo largo de su diagonal 53 para hacer que al menos uno de sus bordes se superponga con al menos uno de los bordes del rectángulo 43 de cara cercana. En este ejemplo, el coeficiente de zoom es menor que uno así que se corresponde con un aumento del zoom. El coeficiente de zoom es determinado según el rectángulo 43 de cara cercana y el rectángulo 44 del tronco.

La Figura 6 ilustra un cálculo ejemplar de una posición 64 óptima para la cámara 101. La cámara 101 está equipada con un aparato de zoom óptico que modifica el focal de su sistema de lentes según las instrucciones de control desde la aplicación de AR. El focal de una cámara está directamente enlazado con su campo de visión. Debido a las restricciones físicas, la cámara 101 tiene un campo máximo de visión sobre el cual no puedo operar. De manera recíproca, un campo mínimo de visión puede ser determinado bajo el cual las condiciones para filmar ya no son satisfactorias. Cuando la aplicación de AR aplica el coeficiente de zoom calculado según el rectángulo 43 de cara cercana y el rectángulo 44 del tronco a la cámara 101, modifica el campo de visión de la cámara 101 hasta uno de sus máximos. Cuando el campo de visión mínimo es alcanzado, significa que la cámara 101 está demasiado lejos del rectángulo de cara cercana para capturar la escena en una dimensión satisfactoria para el usuario. De manera recíproca, cuando el campo de visión máximo es alcanzado, la cámara 101 está demasiado cerca del rectángulo de cara cercana y no puede capturarlo completamente como se ilustra en la Figura 6. En la Figura 6, vista desde arriba, los rectángulos 43, 44 y 54 han sido desplazados por el bien de la claridad. Se comprende que estos tres rectángulos pertenecen a un mismo plano. En este ejemplo, el ángulo 61 se corresponde con la máxima apertura del campo de visión de la cámara 101. Considerando el ancho del rectángulo 43 de cara cercana, según una realización particular de los presentes principios, la aplicación de AR ordena la máxima disminución del zoom a la cámara 101. Sin embargo, la cámara 101 está demasiado cerca del rectángulo de cara cercana para capturar la escena virtual completa. La aplicación de AR no desplaza la cámara 101. Sin embargo, la aplicación de AR puede indicar al usuario donde está la posición óptima para tener una representación óptima de la escena virtual. Por ejemplo, la aplicación de AR puede presentar flechas en la pantalla del dispositivo 10a, 10b, o 10c para indicar al usuario hacia dónde moverse. En una variación, el sistema usa mensajes de audio o efectos hápticos (por ejemplo vibraciones) para indicar la posición óptima al usuario. Primero, para una representación óptima, la dirección de orientación de la cámara 101 tiene que cruzar el centro del rectángulo 43 de cara cercana. En el ejemplo de la Figura 6, el usuario tiene que moverse a su izquierda. Segundo, el ángulo 61 máximo, llamado α , es el que permite un desplazamiento hacia atrás mínimo. La distancia 63, llamada z_2 , entre el rectángulo 43 de cara cercana y la posición 64 óptima, es calculada según la distancia 62, llamada z_1 , entre la cámara 101 y el plano de los tres rectángulos, el ancho, llamado w_1 , del nuevo rectángulo 54 del tronco y el ancho, llamado w_2 de la cara 43 cercana según la siguiente ecuación [E1]:

$$\tan(\alpha) = \frac{w_1}{2z_1} = \frac{w_2}{2z_2} \Rightarrow z_2 = \frac{z_1 w_2}{w_1}$$

En el ejemplo de la Figura 7, el usuario es incitado a moverse hacia tras en una distancia de $z_2 - z_1$. Los mismos principios son usados para el campo de visión mínimo de la cámara 101.

En otra realización, la cámara 101 no está equipada con un aparato de zoom óptico. En tal realización, la aplicación de AR usa una función de zoom digital para controlar el campo de visión de la imagen representada. El máximo campo de visión se corresponde con el campo de visión de la cámara 101. Como un aumento de zoom digital disminuye la calidad de la imagen, el campo de visión mínimo es determinado para mantener una calidad mínima. Los mismos principios son aplicados en la realización. En una variante, el contenido de AR viene con una estimación del tamaño máximo de los límites durante la experiencia. Se le puede pedir al usuario que se mueva hacia una posición óptima solo una vez, por ejemplo, al inicio de la experiencia. En una variante, al menos uno de los elementos de la escena virtual está asociado con una dirección principal. Por ejemplo, en el ejemplo de las Figuras 1, 2 y 3, el televisor 11, aunque constituya un elemento no representado de la escena virtual, puede ser asociado con una dirección principal (por ejemplo la normal a su pantalla). La posición óptima es calculada según estas direcciones principales, así el usuario es incitado a moverse en frente de los elementos principales. Los elementos de la escena virtual con una dirección principal pueden asociarse con un peso (por ejemplo un entero entre 0 y 100 o un número real entre 0 y 1) para calcular la posición óptima de la cámara 101. Un algoritmo de optimización es usado para encontrar una compensación óptima entre las direcciones principales pesadas.

En algunas realizaciones, la cámara 101 es motorizada y es capaz de rotar. Una orientación óptima es calculada según el campo de visión mínimo y máximo de la cámara 101. La orientación óptima puede ser calculada según el rectángulo 43 de cara cercana como calculado para la dirección de orientación real de la cámara. En una variante, como el rectángulo de cara cercana depende de la dirección de orientación de la cámara, un rectángulo de cara

cercana es calculado para orientaciones posibles. Un algoritmo de optimización es usado para calcular la orientación óptima.

La Figura 7 muestra una realización de hardware de un aparato 70 configurado para representar imágenes desde una cámara 101 en un dispositivo 79 de presentación cuando reproduce una escena de realidad aumentada. La Figura 7 es una realización de hardware ejemplar de sistemas como los descritos con las figuras 1 a 3. En este ejemplo, el dispositivo 70 comprende los siguientes elementos, conectados entre sí mediante un bus 73 de direcciones y datos que también transporta una señal de reloj:

- un microprocesador 71 (o CPU),

- una tarjeta 76 gráfica,

- una memoria 74 no volátil de tipo ROM (Memoria de Solo Lectura),

- una Memoria (75) de Acceso Aleatorio o RAM, la tarjeta 76 gráfica puede incluir registros de memoria de acceso aleatorio,

- un sistema 720 de estimación de pose interno opcional, por ejemplo en la forma de una unidad de medición inercial que comprende giroscopios, acelerómetros y/o brújulas,

- un conjunto opcional de dispositivos 721 de I/O (Entrada/Salida) tales como por ejemplo un ratón, un joystick, una cámara web, etc., y

- una fuente 77 de alimentación.

El dispositivo 70 puede estar conectado a un sistema 722 de estimación de pose externo que transmite datos a la CPU 71 a través del bus de direcciones 73.

El dispositivo 70 está conectado a una cámara 101. En algunas realizaciones, la cámara 101 está motorizada. En realizaciones particulares, la cámara 101 está equipada con un aparato de zoom óptico que modifica el focal de su sistema de lentes según las instrucciones de control recibidas mediante el bus 73.

De manera ventajosa, el dispositivo 70 está conectado a uno o más dispositivos 79 de presentación del tipo de la pantalla de presentación directamente a la tarjeta 76 gráfica para presentar imágenes calculadas en la tarjeta gráfica. En una variante, el uno o más dispositivos 79 de presentación están conectados a la tarjeta 76 gráfica a través del bus 73. En una realización particular, el sistema 722 de estimación de pose y/o el uno o más dispositivos 79 de presentación están integrados con el dispositivo 70 tal como Dispositivos Montados en la Cabeza, tabletas o teléfonos inteligentes como se describe con las Figuras 1 a 3.

Se observa que la palabra "registro" usada en la descripción de las memorias 74, 75 y 76 designa en cada una de las memorias mencionadas, tanto una zona de memoria de baja capacidad (algunos datos binarios) como una zona de memoria de gran capacidad (que permite que un programa completo sea almacenado o todo o parte de los datos representativos de datos calculados o a ser presentados).

Cuando está encendido, el microprocesador 71, según el programa en un registro de la ROM 74, carga y ejecuta las instrucciones del programa en registros de la RAM 75.

Según una realización particular, los algoritmos que implementan los pasos del método específicos a la presente descripción y descritos en adelante son de manera ventajosa almacenados en una memoria GRAM de la tarjeta 76 gráfica asociada con el dispositivo 70 que implementa estos pasos.

Según una variante, la fuente 77 de alimentación es externa al dispositivo 70.

La Figura 8 muestra diagramáticamente una realización de un método 80 como se implementa en un dispositivo de procesamiento tal como el dispositivo 70 según la realización ventajosa no restrictiva.

En un paso 801 de inicialización, el dispositivo obtiene datos necesarios para reproducir la aplicación de AR. Los programas asociados y la escena de AR son cargados en la RAM 75, en una variante, al menos para una parte, en la GRAM de la tarjeta 76 gráfica. La pose inicial de la cámara 101 es estimada.

En un paso 802, la pose de la cámara 101 dentro del marco de referencia de la escena de AR es estimada. En una variante, este paso es ejecutado en paralelo con otros pasos e iterados de manera regular, por ejemplo a veinte hertzios o a setenta hertzios. La estimación de pose es calculada según la información de al menos una fuente. Por ejemplo, la cámara 101 está integrada en el dispositivo 70 que está equipado con una IMU. La aplicación de AR sigue la pose diferencial del dispositivo, y así, de la cámara. En una variante, el sistema está equipado con un conjunto de cámaras de infrarrojos que filman trazos de infrarrojo en la cámara 101. Un procesamiento de imagen es requerido, por ejemplo por la CPU 71 del dispositivo 70, para estimar la pose de la cámara según parámetros

intrínsecos y extrínsecos de cámaras de infrarrojos. En otra variante, la cámara 101 filma al menos un marcador de referencia (como en el ejemplo de la Figura 1) y la aplicación de AR estima la pose de la cámara 101, mediante la ejecución de un procesamiento de imagen para ejemplo en la CPU 71 del dispositivo 70, según la pose conocida de las marcas de referencia reconocidas.

5 En un paso 803, los límites de la escena de AR son calculados como se describe en la Figura 3. Un rectángulo 43 de cara cercana y un rectángulo 44 de tronco son calculados según la pose estimada de la cámara 101 (esto es su ubicación y su dirección de orientación), su campo de visión y los límites son descritos en la Figura 4. La barra 820
 10 doble en la Figura 8 indica que el siguiente paso del método 80 depende de la realización del sistema. El método puede consistir, en un paso 804, de prueba de si la posición real de la cámara 101 es óptima (o no muy lejos de la posición óptima). Se comprende que la palabra "posición", así como "pose", se refieren a la ubicación y dirección de
 15 orientación de la cámara. Por ejemplo, la prueba puede consistir en estimar si el rectángulo de cara cercana puede ser incluido en el campo de visión de la cámara cuando el factor de zoom es mínimo (esto es el campo de visión de la cámara es máximo). Si no, la posición real no es óptima porque para cualquier factor de zoom, la cámara está demasiado cerca de la escena de AR para englobarlo todo en su tronco. Los mismos principios son aplicables
 20 cuando la cámara está demasiado lejos de la escena (esto es el rectángulo de cara cercana ocupa una parte pequeña del rectángulo del tronco mínimo). En una variante, la prueba puede consistir en calcular si la cámara 101 está realmente en frente de elementos de la escena asociados con las direcciones principales. Los pesos asociados a estas direcciones son tomados en consideración para calcular esta prueba. Si la cámara 101 no está en una posición óptima, el paso 805 es ejecutado. Si la cámara 101 pasa la prueba 804 de posición óptima, el paso 807 es
 25 ejecutado. En otra realización, el paso 805 es ejecutado de manera regular, en paralelo con el paso 807. En otra realización, el paso 805 es ejecutado sistemáticamente después del paso 803. En tal realización, la prueba 804 nunca es realizada.

En un paso 805, una posición óptima es calculada para la cámara 101. La posición óptima calculada es un punto de
 25 vista desde el cual la visión de la escena de AR es optimizada. Este cálculo toma los límites y los parámetros de la cámara en cuenta como se describe en la Figura 6. En una variante, este cálculo también toma direcciones principales pesadas asociadas con al menos un elemento de la escena de AR para determinar la posición óptima. En un paso 806, la posición óptima calculada es indicada al usuario a través de efectos visuales, por ejemplo mediante la superposición de flechas en las imágenes reproducidas por la aplicación de AR. En una variante, la posición óptima es indicada al usuario a través de efectos hápticos, por ejemplo mediante la activación de vibración
 30 por ejemplo incluida en el dispositivo 70 o llevada por el usuario como una pulsera o un chaleco hápticos. En otra variación, la posición óptima es indicada gracias a mensajes de audio, por ejemplo, pitidos en estéreo o expresiones orales.

En un paso 807, un coeficiente de zoom es calculado según el rectángulo 43 de cara cercana, la pose de la cámara
 35 101 y los parámetros de la cámara 101 que comprende su campo de visión máximo y mínimo como se describe en las Figuras 5a y 5b. La línea 821 doble en la Figura 8 indica que el siguiente paso depende de la realización del sistema. Si la cámara 101 está equipada con un aparato de zoom óptico, un paso 808 puede ser ejecutado. El coeficiente de zoom calculado en el paso 807 es usado por la CPU 71 del dispositivo 70 para transmitir instrucciones al aparato de zoom óptico de la cámara 101. Si la cámara 101 no está equipada con un aparato de zoom óptico, un
 40 paso 809 es ejecutado en vez del paso 808. En este paso, un zoom digital es operado según el coeficiente de zoom calculado. Un recorte es realizado en las imágenes capturadas por la cámara 101 hasta un tamaño que se corresponde con el campo de visión mínimo bajo el cual la calidad de las imágenes ya no se estima suficiente. Si el coeficiente de zoom se corresponde con el campo de visión máximo, no se realiza un recorte. En una variante, el paso 809 puede ser realizado antes, después o en paralelo con el paso 808. Si la cámara 101 está motorizada, un
 45 paso 810 es ejecutado. La aplicación de AR calcula una dirección de orientación óptima para la cámara 101 para optimizar las condiciones de visión de la escena de AR y transmitir instrucciones a la cámara 101 para hacerla rotar. En una variante, el paso 810 es realizado antes, después o en paralelo con el paso 808 y/o 809 según la realización del sistema.

Cuando el zoom y la dirección de orientación de la cámara 101 han sido ajustados, un paso 811 es ejecutado. Este
 50 paso consiste en el procesamiento clásico de la aplicación de AR, esto es superposición de la parte de la escena de RA vista desde el punto de vista ajustado de la cámara 101 en las imágenes capturadas por la cámara 101. El método 80 es ejecutado de manera iterativa desde el paso 802 para tener en cuenta los cambios de la pose de la cámara y/o cambios de la escena de AR y actualizar las imágenes reproducidas.

Naturalmente, la presente descripción no está limitada a las realizaciones descritas anteriormente.

55 En particular, la presente descripción no está limitada a un método de reproducción de imágenes desde una cámara en un dispositivo de presentación cuando reproduce una escena de realidad aumentada sino que también se extiende a cualquier método de transmisión de instrucciones a motores que controlan la cámara o a cualquier método de transmisión de efectos visuales, de audio o hápticos para indicar una posición óptima al usuario. La implementación de cálculos necesarios para generar los límites y los rectángulos usados para el cálculo del coeficiente de zoom no está limitada tampoco a una implementación en microprogramas de tipo sombreado, sino

que también se extiende a una implementación en cualquier tipo de programa, por ejemplo programas que pueden ser ejecutados por un microprocesador de tipo CPU.

Las implementaciones descritas en este documento pueden ser implantadas en, por ejemplo, un método o un proceso, un aparato, un programa de software, un flujo de datos, o una señal. Incluso si solo se discute el contexto de una única forma de implementación (por ejemplo, discutido solo como un método o un dispositivo), la implementación de características descritas puede también ser implementada en otras formas (por ejemplo un programa). Un aparato puede ser implementado, por ejemplo, en hardware, software y firmware apropiados. Los métodos pueden ser implementados, por ejemplo, en aparatos tales como, por ejemplo, un procesador, que se refiere a dispositivos de procesamiento en general, que incluyen, por ejemplo, un ordenador, un microprocesador, un circuito integrado, o un dispositivo de lógica programable. Los procesadores también incluyen dispositivos de comunicación, tales como, por ejemplo, teléfonos inteligentes, tabletas, ordenadores, teléfonos móviles, asistentes digitales portátiles/personales ("PDA"), y otros dispositivos que faciliten la comunicación de información entre usuarios finales.

Las implementaciones de varios procesos y características descritas en este documento pueden ser realizadas en una variedad de equipos o aplicaciones diferentes, particularmente, por ejemplo, equipos o aplicaciones asociados con codificación de datos, decodificación de datos, generación de vistas, procesamiento de texturas, y otro procesamiento de imágenes y texturas relacionadas y/o información de profundidad. Ejemplos de tales equipos incluyen un codificador, un decodificador, un postprocesador que procesa una salida de un decodificador, un preprocesador que proporciona una entrada a un codificador, un codificador de video, un decodificador de video, un códec de video, un servidor web, un decodificador de televisión, un portátil, un ordenador personal, un teléfono móvil, un PDA, y otros dispositivos de comunicación. Debería quedar claro, el equipo puede ser móvil e incluye instalado en un vehículo móvil.

De manera adicional, los métodos pueden ser implementados mediante instrucciones que son realizadas por un procesador, y tales instrucciones (y/o valores de datos producidos por una implementación) pueden ser almacenadas en un medio legible por un ordenador tal como, por ejemplo, un circuito integrado, un portador de software u otro dispositivo de almacenamiento tal como, por ejemplo, un disco duro, un disco compacto ("CD"), un disco óptico (tal como, por ejemplo, un DVD, a menudo referido como un disco versátil digital o un disco de video digital), una memoria de acceso aleatorio ("RAM"), o una memoria de solo lectura ("ROM"). Las instrucciones pueden formar un programa de aplicación realizado de manera tangible en un medio legible por un ordenador. Las instrucciones pueden ser, por ejemplo, en hardware, firmware, software, o una combinación. Las instrucciones pueden encontrarse en, por ejemplo, un sistema operativo, una aplicación separada, o una combinación de las dos. Un procesador puede ser caracterizado, por lo tanto, como, por ejemplo, tanto un dispositivo configurado para llevar a cabo un proceso y un dispositivo que incluye un medio legible por un ordenador (tal como un dispositivo de almacenamiento) que tiene instrucciones para llevar a cabo un proceso. Además, un medio legible por un ordenador puede almacenar, además de o en lugar de instrucciones, valores de datos producidos por una implementación.

Como será evidente a un experto en la técnica, las implementaciones pueden producir una variedad de señales formateadas para llevar información que puede ser, por ejemplo, almacenada o transmitida. La información puede incluir, por ejemplo, instrucciones para realizar un método, o datos producidos por una de las implementaciones descritas. Por ejemplo, una señal puede ser formateada para llevar como datos las reglas para escribir o leer la sintaxis de una realización descrita, o para llevar como datos los valores de sintaxis reales escritos por una realización descrita. Tal señal puede ser formateada, por ejemplo, como una onda electromagnética (por ejemplo, mediante el uso de una parte de frecuencia de radio del espectro) o como una señal en banda base. El formateo puede incluir, por ejemplo, codificar un flujo de datos y modular una portadora con el flujo de datos codificado. La información que la señal lleva puede ser, por ejemplo, información analógica o digital. La señal puede ser transmitida sobre una variedad de enlaces cableados o inalámbricos diferentes, como se conoce. La señal puede ser almacenada en un medio legible por un ordenador.

Varias implementaciones han sido descritas. Sin embargo, se comprenderá que varias modificaciones pueden hacerse. Por ejemplo, elementos de diferentes implementaciones pueden combinarse, suplementarse, modificarse, o eliminarse para producir otras implementaciones. De manera adicional, alguien moderadamente experto comprenderá que otras estructuras y procesos pueden ser sustituidos por los descritos y las implementaciones resultantes realizarán al menos sustancialmente la o las mismas funciones, en uno o unos modos sustancialmente iguales, para alcanzar al menos sustancialmente el o los mismos resultados que los contemplados por esta solicitud.

REIVINDICACIONES

1. Un método (80) de reproducción de imágenes en un dispositivo (79) de presentación cuando reproduce una escena de realidad aumentada desde el punto de vista de una cámara (101, 103), el método comprende:
- 5 determinar límites (32) de la escena de realidad aumentada, dichos límites que abarcan cada elemento de la escena de AR;
- determinar un coeficiente de zoom según una estimación de una pose de la cámara (101, 103) y los límites (32) de la escena; y
- reproducir imágenes en el dispositivo (79) desde la cámara (101, 103) según el coeficiente de zoom.
- 10 2. El método de la reivindicación 1, que comprende indicar una posición (64) de la cámara (101, 103) al usuario mediante el uso de efectos desde un grupo que comprende efectos visuales, de audio o hápticos.
3. El método de la reivindicación 2, donde dicha posición (64) es determinada según una dirección o un peso asociado con al menos un elemento (11, 21, 22) de la escena de realidad aumentada.
4. El método de la reivindicación 1 o 3, donde la escena de realidad aumentada comprende elementos (11) no reproducidos correspondientes a objetos de una escena real capturada por la cámara (101, 103), dichos elementos que son tomados en cuenta cuando se determinan dichos límites (32) de la escena de realidad aumentada.
- 15 5. El método de una de las reivindicaciones 1 a 4, que comprende controlar la dirección de orientación de dicha cámara (103) según los límites (32) de la escena de realidad aumentada y dicho coeficiente de zoom.
6. El método de una de las reivindicaciones 1 a 5, donde dicha cámara (101, 103) comprende una función de zoom óptico, el método además comprende controlar dicha función de zoom óptico según dicho coeficiente de zoom.
- 20 7. El método de una de las reivindicaciones 1 a 6, que comprende operar un zoom digital en dichas imágenes según dicho coeficiente de zoom.
8. Un dispositivo que tiene un elemento de presentación (79) configurado para reproducir imágenes cuando reproduce una escena de realidad aumentada desde un punto de vista de una cámara (101, 103), el dispositivo que comprende una memoria asociada con al menos un procesador configurado para:
- 25 determinar límites (32) de la escena de realidad aumentada, dichos límites que abarcan cada elemento de la escena de AR;
- determinar un coeficiente de zoom según una estimación de una pose de la cámara (101, 103) y los límites (32) de la escena; y
- reproducir imágenes desde la cámara (101, 103) según el coeficiente de zoom.
- 30 9. El dispositivo de la reivindicación 8, que además comprende un transmisor configurado para transmitir una indicación de una posición (64) de la cámara (101, 103) al usuario mediante el uso de efectos desde un grupo que comprende efectos visuales, de audio o hápticos.
10. El dispositivo de la reivindicación 9, donde dicha posición (64) es determinada según una dirección y un peso asociados con al menos un elemento (11, 21, 22) de la escena de realidad aumentada.
- 35 11. El dispositivo de la reivindicación 8 o 10, donde la escena de realidad aumentada comprende elementos (11) no reproducidos correspondientes a objetos de una escena real capturada por la cámara (101, 103), dichos elementos que son tomados en cuenta cuando se determinan dichos límites (32) de la escena de realidad aumentada.
12. El dispositivo de una de las reivindicaciones 8 a 11, donde dicho procesador está además configurado para controlar la dirección de orientación de dicha cámara (103) según los límites (32) de la escena de realidad aumentada y dicho coeficiente de zoom.
- 40 13. El dispositivo de una de las reivindicaciones 8 a 12, donde dicha cámara (101, 103) comprende una función de zoom óptico, dicho procesador está además configurado para controlar dicha función de zoom óptico según dicho coeficiente de zoom.
14. El dispositivo de una de las reivindicaciones 8 a 13, donde dicho procesador está además configurado para operar un zoom digital en dichas imágenes según dicho coeficiente de zoom.
- 45

- 5 15. Un producto de programa informático descargable desde una red de comunicación y/o grabado en un medio legible por un ordenador y/o ejecutable por un procesador, que comprende instrucciones de código de programa para implementar los pasos del método para reproducir imágenes desde una cámara (101, 103) en un dispositivo (79) de presentación cuando reproduce una escena de realidad aumentada según una de las reivindicaciones 1 a 8 cuando es ejecutado por un procesador.

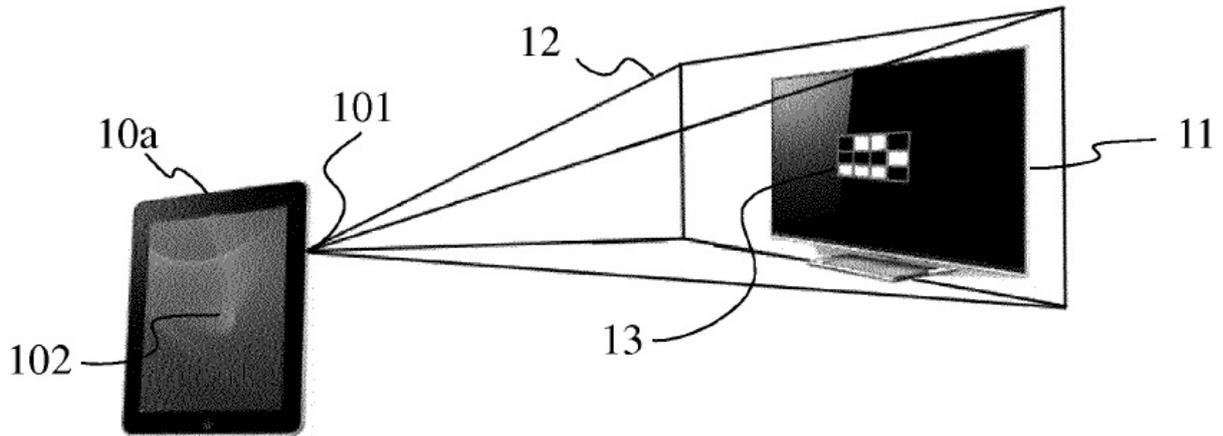


Figura 1

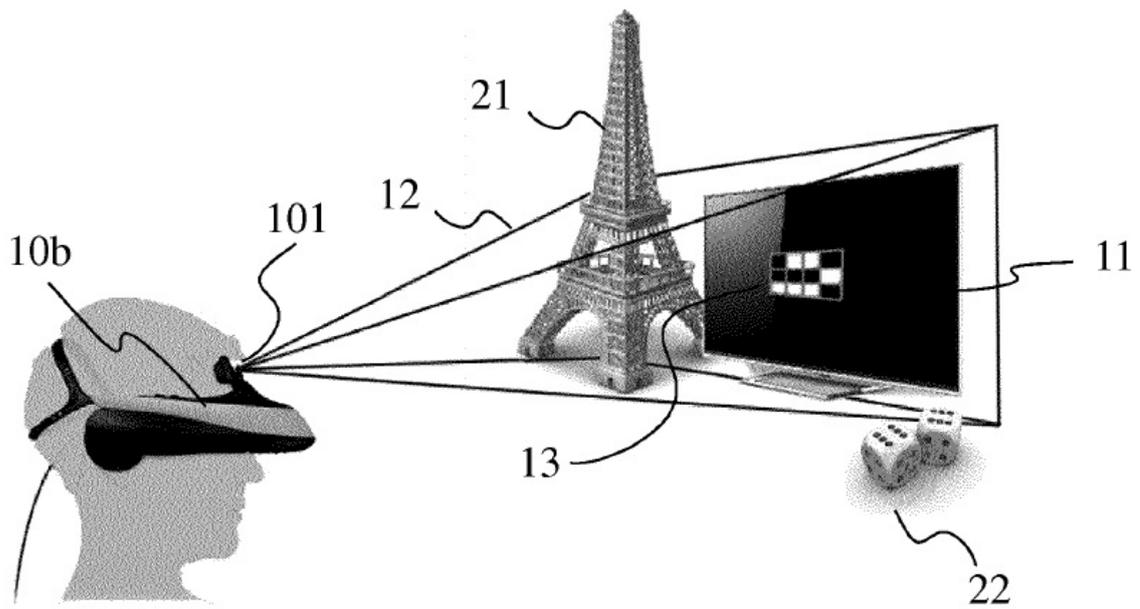


Figura 2

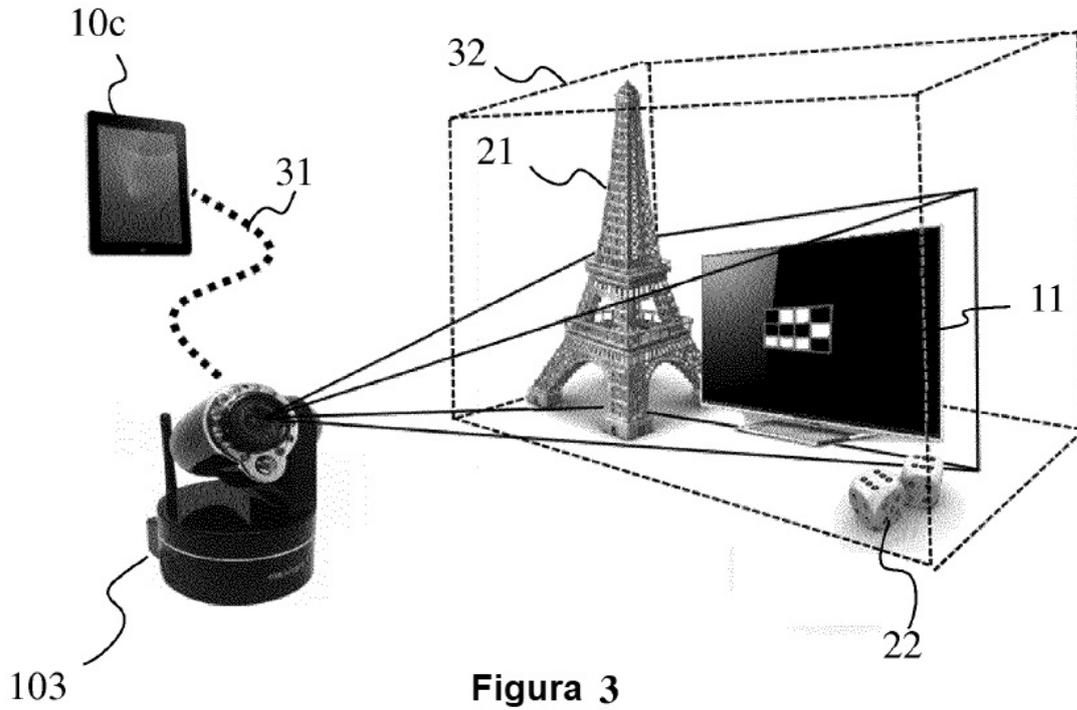


Figura 3

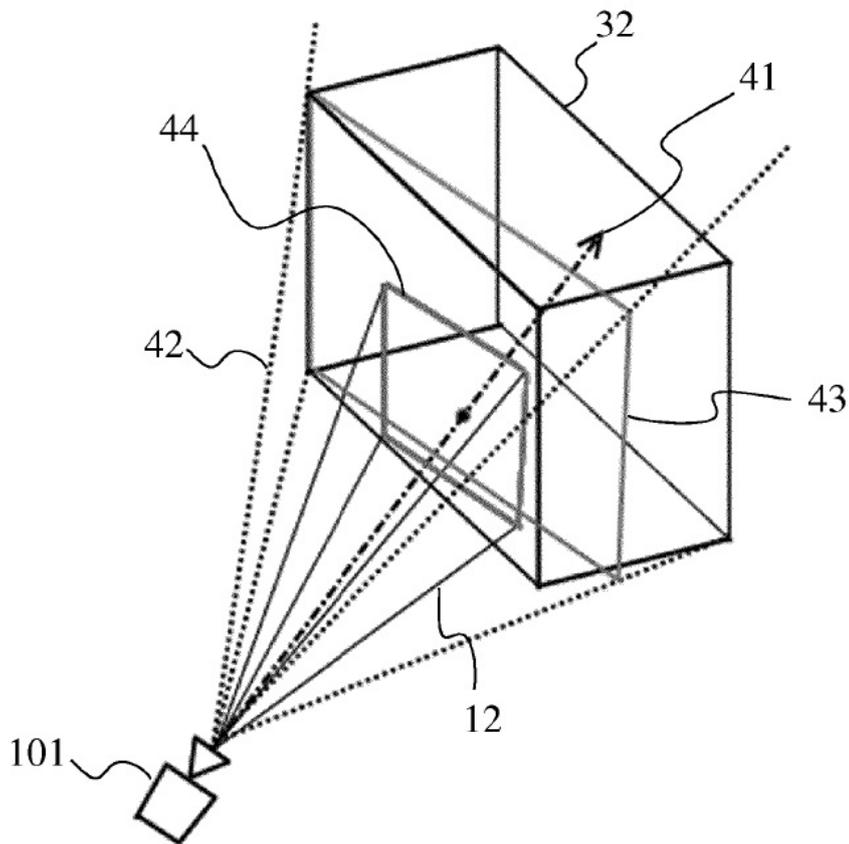


Figura 4

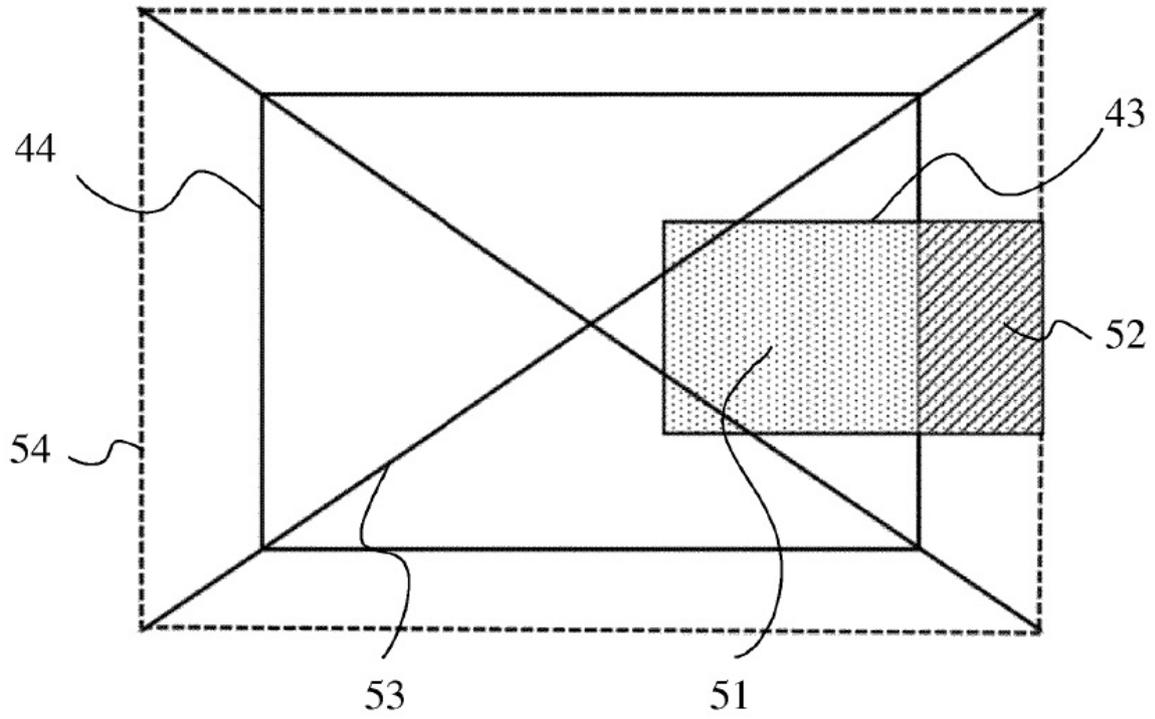


Figura 5a

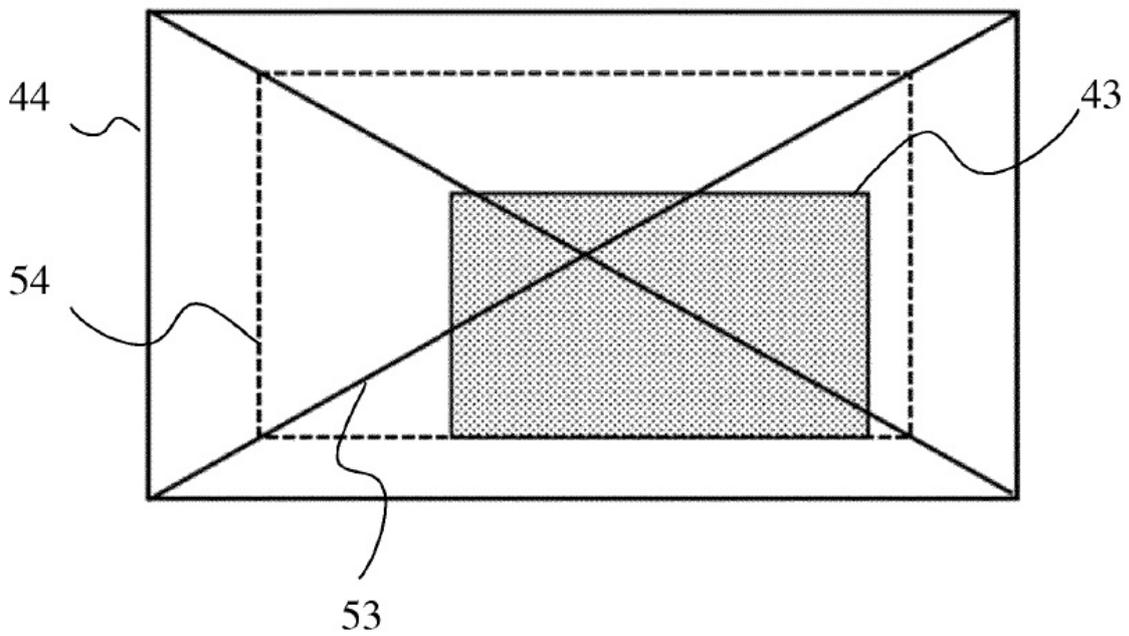


Figura 5b

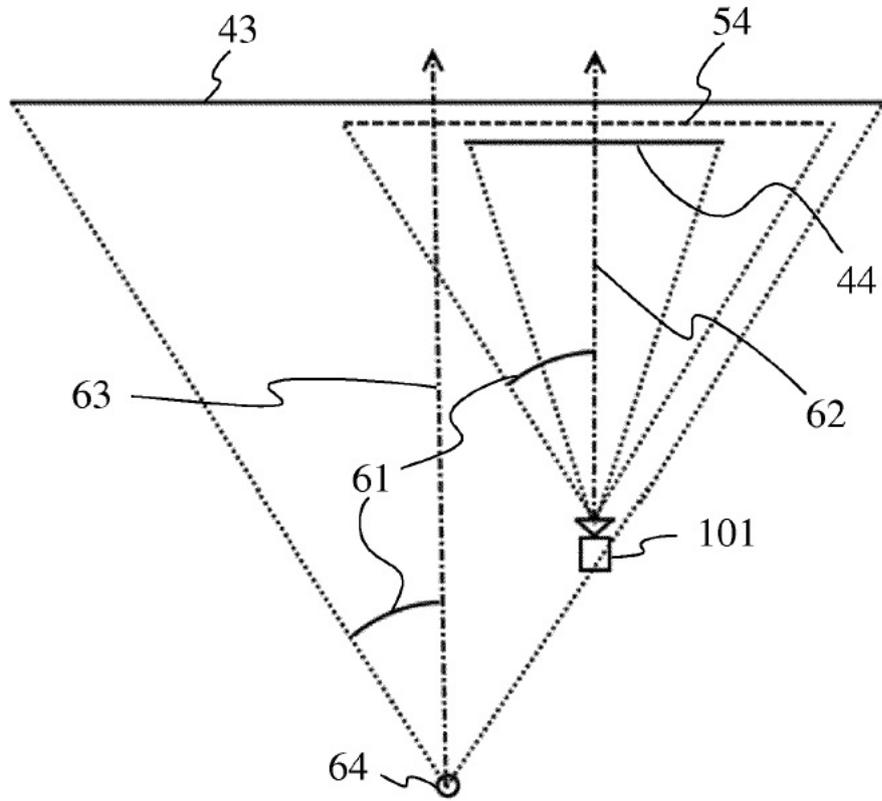


Figura 6

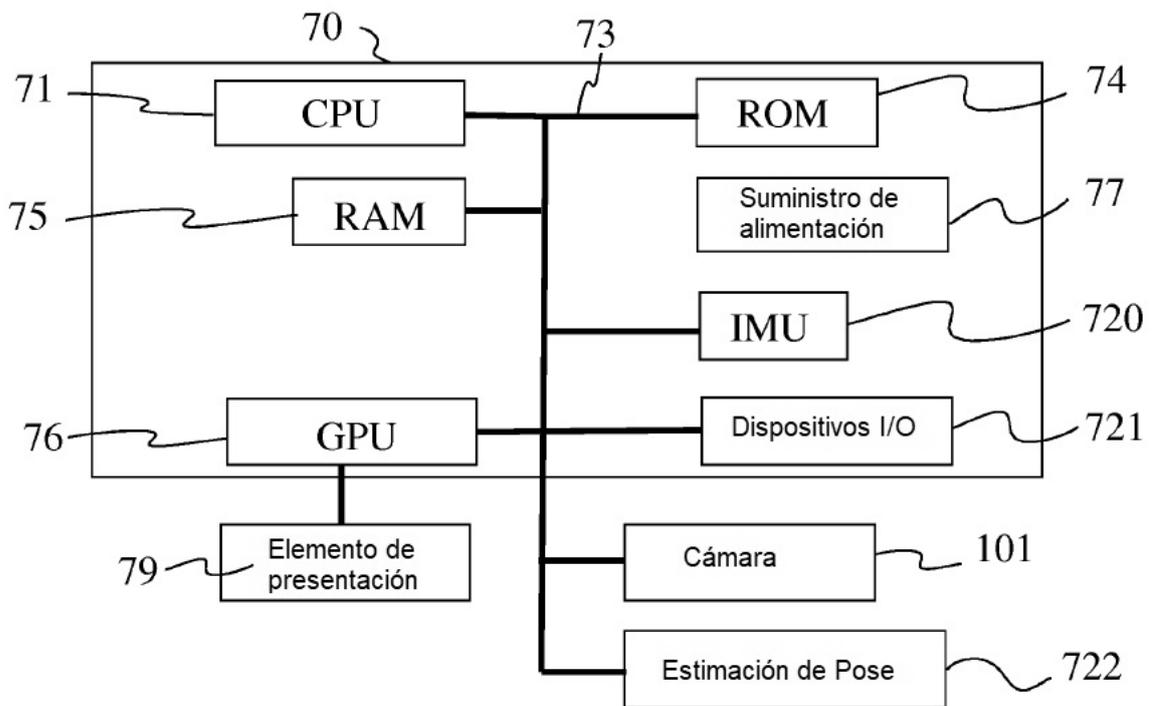


Figura 7

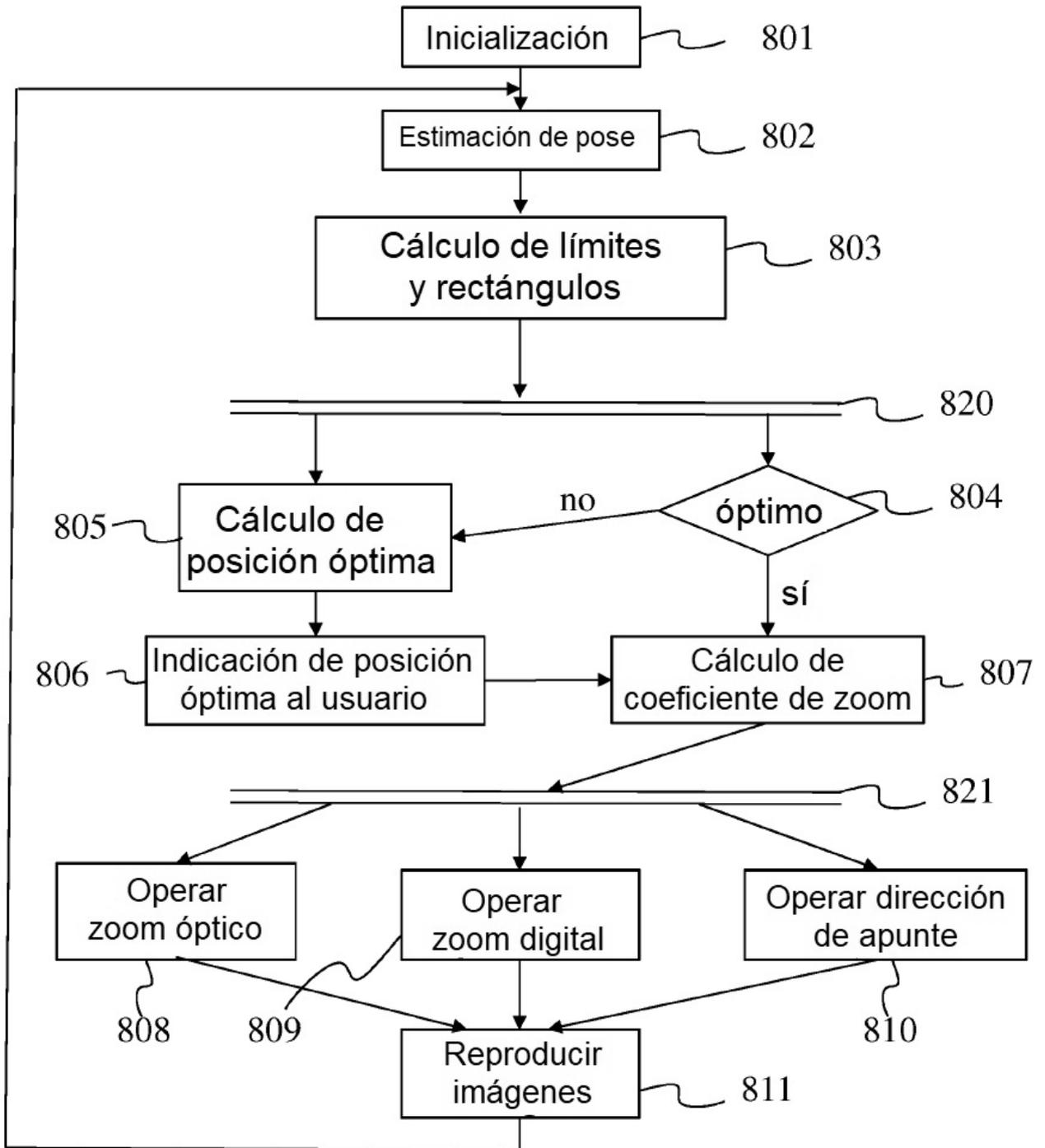


Figura 8