

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 745 739**

51 Int. Cl.:

**G06T 7/73** (2007.01)

**G06T 7/246** (2007.01)

**G06K 9/00** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.09.2011 PCT/US2011/052135**

87 Fecha y número de publicación internacional: **29.03.2012 WO12040099**

96 Fecha de presentación y número de la solicitud europea: **19.09.2011 E 11767336 (8)**

97 Fecha y número de publicación de la concesión europea: **17.07.2019 EP 2619728**

54 Título: **Un entorno adaptable para realidad aumentada asistida por la nube**

30 Prioridad:

**19.09.2011 US 201113235847**  
**20.09.2010 US 384667 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**03.03.2020**

73 Titular/es:

**QUALCOMM INCORPORATED (100.0%)**  
**International IP Administration, 5775 Morehouse**  
**Drive**  
**San Diego, CA 92121-1714, US**

72 Inventor/es:

**SWAMINATHAN, ASHWIN;**  
**SHARMA, PIYUSH;**  
**JIANG, BOLAN;**  
**CHARI, MURALI, R.;**  
**SPINDOLA, SERAFIN, DIAZ;**  
**BAHETI, PAWAN, KUMAR y**  
**NARAYANAN, VIDYA**

74 Agente/Representante:

**FORTEA LAGUNA, Juan José**

ES 2 745 739 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

Un entorno adaptable para realidad aumentada asistida por la nube

## 5 ANTECEDENTES

[0001] Un sistema de realidad aumentada puede insertar objetos virtuales en la vista del mundo real de un usuario. Puede haber muchos componentes en un sistema AR ("Augmented Reality", realidad aumentada) típico. Estos incluyen: adquisición de datos, procesamiento de datos, detección de objetos, rastreo de objetos, registro, refinamiento y componentes de renderización. Estos componentes pueden interactuar entre sí para proporcionar al usuario una experiencia rica en AR. Sin embargo, varios componentes en la detección y el rastreo en un sistema de AR típico pueden utilizar operaciones computacionalmente intensivas, que pueden interrumpir la experiencia de AR del usuario.

El documento STEPHAN GAMMETER Y OTROS: "Server-side object recognition and client-side object tracking for mobile augmented reality [Reconocimiento de objetos del lado del servidor y rastreo de objetos del lado del cliente para realidad aumentada móvil]", SEMINARIO DE VISIÓN ARTIFICIAL Y RECONOCIMIENTO DE PATRONES (CVPRW), 2010 CONFERENCIA DE IEEE COMPUTER SOCIETY, IEEE, PISCATAWAY, NJ, EE.UU., 13 de junio de 2010 (13-06-2010), páginas 1-8, XP031728435, ISBN: 978-1-4244-7029-7, divulga un sistema para realidad aumentada (AR) móvil basado en el reconocimiento visual. Las tareas de reconocer un objeto y rastrearlo en la pantalla del usuario se dividen en tareas del lado del servidor y del lado del cliente, respectivamente. Las capacidades de este enfoque de cliente-servidor híbrido se demuestran con una aplicación prototipo en la plataforma Android, que puede aumentar tanto los objetos estacionarios (puntos de referencia) como los no estacionarios (cubiertas de medios). La base de datos en el lado del servidor consiste en cientos de miles de puntos de referencia, de los que se realiza un seguimiento utilizando un procedimiento de extracción de datos para las colecciones de fotos de la comunidad.

## SUMARIO

[0002] De acuerdo con la presente invención, se proporcionan un procedimiento como se expone en la reivindicación 1, una plataforma móvil como se expone en la reivindicación 14 y un medio legible por ordenador como se expone en la reivindicación 15. Los modos de realización de la invención se reivindican en las reivindicaciones dependientes.

[0003] Una plataforma móvil procesa de manera eficiente los datos de sensor, incluidos los datos de imagen, utilizando un procesamiento distribuido en el que las operaciones sensibles a la latencia se realizan en la plataforma móvil, mientras que las operaciones no sensibles a la latencia, pero computacionalmente intensivas, se realizan en un servidor remoto. La plataforma móvil adquiere datos de sensor, como los datos de imagen y determina si hay un evento de activación para transmitir los datos de sensor al servidor. El evento de activación es un cambio en los datos de sensor en relación con los datos de sensor adquiridos previamente, por ejemplo, un cambio de escena en la imagen capturada. Cuando hay un cambio presente, los datos de sensor se transmiten al servidor para su procesamiento. El servidor procesa los datos de sensor y devuelve información relacionada con los datos de sensor, como la identificación de un objeto en una imagen. La plataforma móvil puede realizar entonces un rastreo basado en referencias utilizando el objeto identificado.

[0004] En una implementación, un procedimiento incluye adquirir datos de sensor usando una plataforma móvil; determinar si hay un evento de activación que comprende un cambio en los datos de sensor en relación con los datos de sensor adquiridos previamente; transmitir los datos de sensor a un servidor cuando se produce el evento de activación; y recibir información relacionada con los datos de sensor desde el servidor. Los datos de sensor pueden ser una imagen capturada de un objeto, por ejemplo, una foto o una trama de vídeo.

[0005] En otra implementación, una plataforma móvil incluye un sensor adaptado para adquirir datos de sensor y un transceptor inalámbrico. El sensor puede ser, por ejemplo, una cámara para capturar una imagen de un objeto. Un procesador está acoplado al sensor y al transceptor inalámbrico y está adaptado para adquirir datos de sensor a través del sensor, para determinar si hay un evento de activación que comprenda un cambio en los datos de sensor en relación con los datos de sensor previamente adquiridos, para transmitir a través del transceptor inalámbrico los datos de sensor a un procesador externo cuando el evento de activación está presente, y para recibir a través del transceptor inalámbrico la información relacionada con los datos de sensor del procesador externo.

[0006] En otra implementación, una plataforma móvil incluye medios para adquirir datos de sensor; medios para determinar si hay un evento de activación que comprende un cambio en los datos de sensor en relación con los datos de sensor adquiridos previamente; medios para transmitir los datos de sensor a un servidor cuando se produce el evento de activación; y medios para recibir información relacionada con los datos de sensor desde el servidor. Los medios para adquirir datos de sensor son una cámara y los datos de sensor son una imagen capturada de un objeto.

[0007] En otra implementación más, un medio legible por ordenador no transitorio que incluye un código de programa almacenado en el mismo incluye un código de programa para adquirir datos de sensor; código de programa para determinar si hay un evento de activación que comprenda un cambio en los datos de sensor en relación con los datos de sensor adquiridos previamente; código de programa para transmitir los datos de sensor a un procesador externo

cuando el evento de activación está presente, y código de programa para recibir información relacionada con los datos de sensor desde el procesador externo.

## BREVE DESCRIPCIÓN DEL DIBUJO

[0008]

La Fig. 1 ilustra un diagrama de bloques que muestra un sistema para procesamiento distribuido que incluye una plataforma móvil y un servidor remoto.

La Fig. 2 es un diagrama de flujo que ilustra un proceso de procesamiento distribuido con operaciones sensibles a la latencia realizadas por la plataforma móvil y operaciones computacionalmente intensivas y no sensibles a la latencia realizadas por un procesador externo.

La Fig. 3 ilustra un diagrama de bloques de la operación de un sistema para AR asistida por servidor.

La Fig. 4 ilustra un diagrama de flujo de llamadas para la AR asistida por servidor, en la que el servidor remoto proporciona la pose (posición y orientación).

La Fig. 5 ilustra otro diagrama de flujo de llamadas para la AR asistida por servidor, en la que el servidor remoto no proporciona la pose.

La Fig. 6 ilustra un diagrama de flujo del procedimiento realizado por el detector de cambio de escena.

La Fig. 7 es un gráfico que ilustra el rendimiento del sistema de procesamiento distribuido que muestra las transmisiones de red requeridas en función de la brecha de activación mínima.

Las Fig. 8 y 9 ilustran enfoques para el reconocimiento facial mediante el proceso de AR asistida por servidor.

Las Fig. 10 y 11 ilustran enfoques de una búsqueda visual utilizando el proceso de AR asistida por servidor.

Las Fig. 12 y 13 ilustran enfoques para el rastreo basado en referencias utilizando el proceso asistido por servidor.

La Fig. 14 ilustra un enfoque para la creación de modelos 3D utilizando el proceso asistido por servidor.

La Fig. 15 es un diagrama de bloques de una plataforma móvil capaz de procesamiento distribuido utilizando detección basada en servidor.

## DESCRIPCIÓN DETALLADA

[0009] Un sistema de procesamiento distribuido, como se divulga en el presente documento, incluye un dispositivo que puede determinar cuándo proporcionar datos a un servidor a través de una red inalámbrica, o a otro dispositivo a través de la red en un entorno de computación en la nube, para ser procesados. El dispositivo también puede procesar los datos en sí. Por ejemplo, las operaciones sensibles a la latencia pueden elegirse para realizarse en el dispositivo y las operaciones no sensibles a la latencia pueden elegirse para realizarse de forma remota para un procesamiento más eficiente. Los factores para determinar cuándo enviar los datos al servidor para ser procesados pueden incluir si las operaciones que se realizan en los datos son sensibles o no a la latencia, la cantidad de cálculo requerida, velocidad/disponibilidad del procesador en el dispositivo o el servidor, las condiciones de la red o la calidad de servicio, entre otros factores.

[0010] En un modo de realización, se proporciona un sistema que incluye una plataforma móvil y un servidor externo para aplicaciones de Realidad Aumentada (AR), en las que las operaciones sensibles a la latencia se realizan en la plataforma móvil, mientras que las operaciones no sensibles a la latencia, pero computacionalmente intensivas, se realizan de forma remota, por ejemplo, en el servidor, para un procesamiento eficiente. Los resultados pueden ser enviados a continuación por el servidor a la plataforma móvil. Al utilizar el procesamiento distribuido para las aplicaciones de AR, el usuario final puede disfrutar sin problemas de la experiencia de AR.

[0011] Tal y como se utiliza en el presente documento, una plataforma móvil se refiere a cualquier dispositivo electrónico portátil, tal como un dispositivo celular u otro dispositivo de comunicación inalámbrica, un dispositivo de sistema de comunicaciones personal (PCS), un dispositivo de navegación personal (PND), un gestor de información personal (PIM), un asistente digital personal (PDA) u otro dispositivo móvil adecuado. La plataforma móvil puede ser capaz de recibir señales de comunicación y/o de navegación inalámbricas, tales como señales de posicionamiento de navegación. El término "plataforma móvil" también pretende incluir dispositivos que se comunican con un dispositivo de navegación personal (PND), tal como mediante una conexión inalámbrica de corto alcance, una conexión mediante infrarrojos, una conexión por cable u otra conexión, independientemente de si la recepción de señales de satélites, la recepción de datos de asistencia y/o el procesamiento relacionado con la posición se llevan a cabo en el dispositivo o

en el PND. Además, la "plataforma móvil" está destinada a incluir todos los dispositivos electrónicos, incluidos los dispositivos de comunicación inalámbrica, ordenadores, ordenadores portátiles, tabletas, etc., que sean capaces de AR.

**[0012]** La Fig. 1 ilustra un diagrama de bloques que muestra un sistema 100 para procesamiento distribuido utilizando detección e identificación de objetos basados en servidor. El sistema 100 incluye una plataforma móvil 110 que realiza operaciones sensibles a la latencia, como el rastreo, mientras que un servidor remoto 130 realiza operaciones no sensibles a la latencia y computacionalmente intensivas, como la identificación de objetos. La plataforma móvil puede incluir una cámara 112 y una pantalla 114 y/o puede incluir sensores de movimiento 164. La plataforma móvil 110 puede adquirir una imagen 104 de un objeto 102, que puede mostrarse en la pantalla 114. La imagen 104 capturada por la plataforma móvil 110 puede ser una imagen estática, por ejemplo, una fotografía, o una sola trama de un flujo de vídeo, ambas de las cuales se mencionan en el presente documento como una imagen capturada. La plataforma móvil 110 puede adquirir adicionalmente o de forma alternativa otros datos de sensor, incluidos los datos de posición y/o de orientación, de un sensor diferente al de la cámara 112, por ejemplo, utilizando un receptor 166 del sistema de posicionamiento por satélite (SPS) o uno o más sensores de movimiento 164, incluyendo, por ejemplo, acelerómetros, giroscopios, brújula electrónica u otros elementos de detección de movimiento similares. Un SPS puede ser una constelación del Sistema global de navegación por satélite (GNSS), tal como el Sistema de posicionamiento global (GPS) Galileo, Glonass o Compass u otros diversos sistemas regionales, tales como, el Sistema de satélites cuasicenitales (QZSS) en Japón, el Sistema indio de satélites de navegación regional (IRNSS) en la India, Beidou en China, etc., y/o varios sistemas de aumento (por ejemplo, un Sistema de aumento basado en satélites (SBAS)) que pueden estar asociados a, o su uso puede permitirse de otro modo en, uno o más sistemas de satélites de navegación global y/o regional.

**[0013]** La plataforma móvil 110 transmite la información de los datos adquiridos, como la imagen capturada 104 y/o los datos de sensor, como la información de SPS o la información de posición de los sensores de movimiento 164 a bordo, al servidor 130 a través de una red 120. La información de los datos adquiridos puede incluir adicionalmente, o de forma alternativa, datos contextuales, como la identificación de cualquier objeto que actualmente está siendo rastreado por la plataforma móvil 110. La red 120 puede ser cualquier red de comunicación inalámbrica, tal como una red inalámbrica de área extensa (WWAN), una red inalámbrica de área local (WLAN), una red inalámbrica de área personal (WPAN), etc. El servidor 130 procesa la información de los datos proporcionada por la plataforma móvil 110 y genera información relacionada con la información de los datos. Por ejemplo, el servidor 130 puede realizar la detección e identificación de objetos basándose en los datos de imagen proporcionados utilizando una base de datos de objetos 140. El servidor 130 devuelve a la plataforma móvil 110 información relacionada con los datos adquiridos. Por ejemplo, si el servidor 130 identifica un objeto a partir de los datos de imagen proporcionados por la plataforma móvil 110, el servidor 130 puede devolver una identificación del objeto, por ejemplo, incluyendo un identificador como un título o número de identificación o una imagen de referencia 106 del objeto 102, así como cualquier información complementaria deseada, como indicadores destacados, enlaces de información, etc., que puede ser utilizada por la plataforma móvil para la aplicación de realidad aumentada.

**[0014]** Si se desea, el servidor 130 puede determinar y proporcionar a la plataforma móvil 110 una pose (posición y orientación) de la plataforma móvil 110 en el momento en que la imagen 104 fue capturada en relación con el objeto 102 en la imagen de referencia 106, que es, por ejemplo, una imagen del objeto 102 desde una posición y orientación conocidas. La pose devuelta se puede utilizar para arrancar el sistema de rastreo en la plataforma móvil 110. En otras palabras, la plataforma móvil 110 puede rastrear todos los cambios progresivos en su pose, por ejemplo, visualmente o usando sensores de movimiento 164, desde el momento en que captura la imagen 104 hasta el momento en que recibe la imagen de referencia 106 y la pose desde el servidor 130. La plataforma móvil 110 puede entonces usar la pose recibida junto con sus cambios progresivos rastreados en la pose para determinar rápidamente la pose actual con respecto al objeto 102.

**[0015]** En otro modo de realización, el servidor 130 devuelve la imagen de referencia 106, pero no proporciona información de la pose, y la plataforma móvil 110 determina una pose actual con respecto al objeto 102 comparando una imagen capturada actual del objeto 102 con respecto a la imagen de referencia 106 del objeto 102 utilizando un algoritmo de detección de objetos. La pose se puede utilizar como una entrada al sistema de rastreo para que se pueda estimar el movimiento relativo.

**[0016]** En otro modo de realización más, el servidor 130 devuelve solo la información de la pose, pero no proporciona la imagen de referencia. En este caso, la plataforma móvil 110 puede usar la imagen capturada 104 junto con la información de la pose para crear una imagen de referencia que posteriormente puede ser utilizada por el sistema de rastreo. De forma alternativa, la plataforma móvil 110 puede rastrear los cambios progresivos en la posición entre la imagen capturada 104 y una imagen capturada posteriormente (conocida como la imagen actual) y puede calcular la pose de la imagen actual en relación con la imagen de referencia generada por la plataforma móvil usando la pose obtenida del servidor 130 junto con los resultados de rastreo progresivos. En ausencia de la imagen de referencia 102, la imagen actual puede deformarse (o rectificarse) utilizando la pose estimada para obtener una estimación de la imagen de referencia que se puede usar para arrancar el sistema de rastreo.

**[0017]** Además, para minimizar la frecuencia de las solicitudes de detección enviadas por la plataforma móvil 110 al servidor 130, la plataforma móvil 110 puede iniciar una solicitud de detección solo si hay un evento de activación. Un evento de activación puede basarse en un cambio en los datos de imagen o en los datos de sensor de los sensores de movimiento 164 en relación con los datos de imagen o los datos de sensor adquiridos previamente. Por ejemplo, la plataforma móvil 110 puede usar un detector de cambio de escena 304 para determinar si se ha producido un cambio en los datos de la imagen. Por lo tanto, en algunos modos de realización, la plataforma móvil 110 puede comunicarse con el servidor 130 a través de la red para solicitudes de detección solo cuando son activadas por el detector de cambio de escena 304. El detector de cambio de escena 304 activa la comunicación con el servidor para la detección de objetos, por ejemplo, solo cuando hay nueva información presente en la imagen actual.

**[0018]** La Fig. 2 es un diagrama de flujo que ilustra un proceso de procesamiento distribuido con operaciones sensibles a la latencia realizadas por la plataforma móvil 110 y operaciones no sensibles a la latencia y computacionalmente intensivas realizadas por un procesador externo, como el servidor 130. Como se ilustra, los datos de sensor son adquiridos por la plataforma móvil 110 (202). Los datos de sensor pueden ser una imagen adquirida, por ejemplo, una foto capturada o una trama de vídeo, o información obtenida de las mismas, incluido el reconocimiento de caracteres o la extracción de puntos clave. Los datos de sensor pueden incluir adicionalmente o de forma alternativa, por ejemplo, información de SPS, información del sensor de movimiento, reconocimiento de códigos de barras, resultados de detección de texto u otros resultados del procesamiento parcial de la imagen, así como información contextual, como el comportamiento del usuario, las preferencias del usuario, la ubicación, información o datos del usuario (por ejemplo, información de la red social sobre el usuario), hora del día, calidad de la iluminación (natural o artificial) y personas que se encuentran cerca (en la imagen), etc.

**[0019]** La plataforma móvil 110 determina que hay un evento de activación (204), como un cambio en los datos de sensor en relación con los datos de sensor adquiridos previamente. Por ejemplo, el evento de activación puede ser un cambio de escena en el que aparece un objeto nuevo o diferente en la imagen. Los datos de sensor adquiridos se transmiten al servidor 130 después de que se detecta (206) un evento de activación, como un cambio de escena. Por supuesto, si no se detecta ningún cambio de escena, los datos de sensor no necesitan transmitirse al servidor 130, reduciendo así la comunicación y solicitudes de detección.

**[0020]** El servidor 130 procesa la información adquirida, por ejemplo, para realizar el reconocimiento de objetos, que es bien conocido en la técnica. Después de que el servidor 130 procesa la información, la plataforma móvil 110 recibe del servidor 130 información relacionada con los datos de sensor (208). Por ejemplo, la plataforma móvil 110 puede recibir resultados de la identificación del objeto, incluyendo, por ejemplo, una imagen de referencia. La información relacionada con los datos de sensor puede incluir adicionalmente, o de forma alternativa, información tal como los elementos que se encuentran cerca de la plataforma móvil 110 (como edificios, restaurantes, productos disponibles en una tienda, etc.) así como modelos bidimensionales (2D) o tridimensionales (3D) del servidor, o información que se puede utilizar en otros procesos, como en juegos. Si se desea, se puede proporcionar información adicional, incluida la pose de la plataforma móvil 110 con respecto al objeto en la imagen de referencia en el momento en que se capturó la imagen 104, como se analizó anteriormente. Si la plataforma móvil 110 incluye una memoria caché local, entonces la plataforma móvil 110 puede almacenar múltiples imágenes de referencia enviadas por el servidor 130. Estas imágenes de referencia almacenadas pueden usarse, por ejemplo, para nuevas detecciones posteriores que pueden realizarse en la plataforma móvil 110 si se pierde el rastro. En algunos modos de realización, el servidor identifica una pluralidad de objetos del sensor en la imagen. En tales modos de realización, una imagen de referencia u otro identificador de objeto puede enviarse a la plataforma móvil 110 para solo uno de los objetos identificados, o una pluralidad de identificadores de objeto correspondientes a objetos respectivos puede ser transmitida a, y recibida por, la plataforma móvil 110.

**[0021]** Por lo tanto, la información que puede proporcionar el servidor 130 puede incluir un resultado de reconocimiento, información sobre el objeto o los objetos identificados, imágenes de referencia (una o muchas) sobre el objeto o los objetos que se pueden usar para varias funciones, como en el rastreo, el modelo 2D/3D del objeto u objetos reconocidos, la pose absoluta del objeto u objetos reconocidos, la información de aumento que se usará para mostrar, y/o la información destacada sobre el objeto. Además, el servidor 130 puede enviar información relacionada con la coincidencia de objetos que podría mejorar el clasificador en la plataforma móvil 110. Un posible ejemplo es cuando la plataforma móvil 110 está utilizando árboles de decisión para la comparación. En este caso, el servidor 130 podría enviar los valores para los nodos individuales del árbol para facilitar una construcción de árbol más precisa y, por consiguiente, una mejor coincidencia. Los ejemplos de árboles de decisión incluyen, por ejemplo, k-medias, árboles k-d, árboles de vocabulario y otros árboles. En el caso de un árbol de k-medias, el servidor 130 también puede enviar la semilla para inicializar la estructura jerárquica del árbol de k-medias en la plataforma móvil 110, permitiendo así que la plataforma móvil 110 realice una búsqueda para cargar el árbol apropiado.

**[0022]** Opcionalmente, la plataforma móvil 110 puede obtener una pose para la plataforma móvil con respecto al objeto 102 (210). Por ejemplo, la plataforma móvil 110 puede obtener la pose relativa al objeto en la imagen de referencia sin recibir ninguna información de pose del servidor 130 al capturar otra imagen del objeto 102 y comparar la imagen recién capturada con la imagen de referencia. Cuando el servidor 130 proporciona información de la pose, la plataforma móvil puede determinar rápidamente una pose actual, combinando la pose proporcionada por el servidor 130, que es la pose de la plataforma móvil 110 en relación con el objeto en la imagen de referencia en el momento en

que se capturó la imagen inicial 104, con cambios rastreados en la pose de la plataforma móvil 110 desde que se capturó la imagen inicial 104. Cabe señalar que si la pose se obtiene con o sin la asistencia del servidor 130 puede depender de las capacidades de la red 120 y/o de la plataforma móvil 110. Por ejemplo, si el servidor 130 admite la estimación de la pose y si la plataforma móvil 110 y el servidor 130 acuerdan una interfaz de programación de aplicaciones (API) para transmitir la pose, la información de la pose puede transmitirse a la plataforma móvil 110 y usarse para el rastreo. La pose del objeto 102 (210) enviada por el servidor puede estar en forma de matrices de rotación y transformación relativas, una matriz de homografía, una matriz de transformación afín u otra forma.

**[0023]** Opcionalmente, la plataforma móvil 110 puede entonces realizar AR con el objeto, utilizando los datos recibidos del servidor 130, como rastrear el objetivo, estimar la pose del objeto en cada trama e introducir un objeto virtual o aumentar de otro modo la vista o imagen del usuario a través del motor de renderizado utilizando la pose estimada (212).

**[0024]** La figura 3 ilustra un diagrama de bloques de la operación del sistema 100 para la AR 130 asistida por servidor. Como se muestra en la Fig. 3, una nueva imagen capturada 300 se utiliza para iniciar un rastreador sin referencias 302. El rastreador sin referencias 302 realiza un rastreo basado en flujo óptico, correlaciones cruzadas normalizadas (NCC) o cualquier procedimiento similar, conocido en la técnica. El rastreador sin referencias 302 identifica características, tales como puntos, líneas, regiones y similares, en la nueva imagen capturada 300 y sigue estas características de trama a trama, por ejemplo, utilizando vectores de flujo. Los vectores de flujo obtenidos de los resultados de rastreo ayudan a estimar el movimiento relativo entre una imagen capturada anterior y una imagen capturada actual y, a su vez, ayudan a identificar la velocidad del movimiento. La información proporcionada por el rastreador sin referencias 302 es recibida por el detector de cambio de escena 304. El detector de cambio de escena 304 utiliza, por ejemplo, características rastreadas del rastreador sin referencias 302, junto con otros tipos de estadísticas de imagen (como estadísticas de histograma) y otra información disponible de los sensores en la plataforma móvil para estimar el cambio en la escena. Si el detector de cambio de escena 304 no envía ningún evento de activación, el proceso continúa con el rastreador sin referencias 302. Si el detector de cambio de escena 304 identifica un cambio sustancial en la escena, el detector de cambio de escena 304 envía una señal de activación que puede iniciar el proceso de detección en el detector 308 basado en servidor. Si se desea, se puede usar un estimador de calidad de imagen 306 para analizar la calidad de la imagen para controlar aún más la transmisión de solicitudes al detector 308 basado en servidor. El estimador de calidad de la imagen 306 examina la calidad de la imagen y si la calidad es buena, es decir, mayor que un umbral, se activa una solicitud de detección. Si la calidad de la imagen es deficiente, no se activa la detección y la imagen no se transmite al detector 308 basado en servidor. En un modo de realización de la invención, la plataforma móvil 110 puede esperar una imagen de buena calidad durante un período de tiempo finito después de que se haya detectado un cambio de escena antes de enviar la imagen de buena calidad al servidor 130 para el reconocimiento de objetos.

**[0025]** La calidad de la imagen puede basarse en estadísticas de imagen conocidas, medidas de calidad de la imagen y otros enfoques similares. Por ejemplo, el grado de nitidez de una imagen capturada se puede cuantificar mediante el filtrado de paso alto y generando un conjunto de estadísticas que representan, por ejemplo, la intensidad de los bordes y la distribución espacial. La imagen se puede clasificar como una imagen de buena calidad si el valor de la nitidez excede o es comparable a la "nitidez prevaleciente" de la escena, por ejemplo, como el promedio de varias tramas anteriores. En otra implementación, se puede usar un algoritmo de detección rápida de esquinas, como las esquinas FAST (Features from Accelerated Segment Test) o las esquinas de Harris para analizar la imagen. La imagen se puede clasificar como una imagen de buena calidad si hay un número suficiente de esquinas, por ejemplo, el número de esquinas detectadas supera un umbral o es mayor o comparable al "número de esquinas prevaleciente" de la escena, por ejemplo, como el promedio de varias tramas anteriores. En otra implementación, las estadísticas de la imagen, como la media o la desviación estándar de las magnitudes de gradiente de borde, pueden usarse para informar a un clasificador de aprendizaje, que puede usarse para distinguir entre imágenes de buena calidad y de mala calidad.

**[0026]** La calidad de la imagen también se puede medir utilizando las entradas del sensor. Por ejemplo, las imágenes capturadas por la plataforma móvil 110 mientras se mueve rápidamente pueden ser borrosas y, por lo tanto, de menor calidad que si la plataforma móvil 110 fuera estática o se moviera lentamente. Por consiguiente, las estimaciones de movimiento de los datos de sensor, por ejemplo, de los sensores de movimiento 164 o del rastreo visual, pueden compararse con un umbral para determinar si las imágenes de la cámara resultantes son de la calidad suficiente para ser enviadas para la detección de objetos. De manera similar, la calidad de la imagen se puede medir basándose en una cantidad determinada de desenfoque de la imagen.

**[0027]** Además, se puede proporcionar un administrador de tiempo de activación 305 para controlar aún más la cantidad de solicitudes transmitidas al detector 308 basado en servidor. El administrador de tiempo de activación 305 mantiene el estado del sistema y puede estar basado en heurísticas y reglas. Por ejemplo, si el número de imágenes desde la última imagen de activación es mayor que un umbral, por ejemplo, 1000 imágenes, el administrador de tiempo de activación 305 puede generar un evento de activación que puede expirar e iniciar automáticamente el proceso de detección en el detector 308 basado en servidor. Por lo tanto, si no ha habido un evento de activación durante un número extendido de imágenes, el administrador de tiempo de activación 305 puede forzar un evento de activación, lo que es útil para determinar si hay objetos adicionales en el campo de visión de la cámara. Además, el administrador

de tiempo de activación 305 puede programarse para mantener una separación mínima entre dos eventos de activación en un valor elegido de  $\eta$ , es decir, el administrador de tiempo de activación 305 suprime los eventos de activación si está dentro de  $\eta$  imágenes desde la última imagen de activación. La separación de imágenes de activación puede ser útil, por ejemplo, si la escena está cambiando rápidamente. Por lo tanto, si el detector de cambio de escena 304 produce más de un evento de activación dentro de  $\eta$  imágenes, solo se envía una imagen de activación al detector 308 basado en servidor, reduciendo así la cantidad de comunicación al servidor 130 desde la plataforma móvil 110. El administrador de tiempo de activación 305 también puede gestionar las planificaciones de activación. Por ejemplo, si el detector de cambio de escena 304 produce un nuevo evento de activación que es menor que  $\eta$  imágenes y mayor que  $\mu$  imágenes desde el último evento de activación, el nuevo evento de activación puede ser almacenado y pospuesto por el administrador de tiempo de activación 305 hasta un momento en que la brecha de imágenes entre eventos de activación consecutivos es al menos  $\eta$ . A modo de ejemplo,  $\mu$  puede ser 2 imágenes y  $\eta \geq \mu$ , y a modo de ejemplo,  $\eta$  puede variar como 2, 4, 8, 16, 32, 64.

**[0028]** El administrador de tiempo de activación 305 también puede gestionar fallos de detección del servidor 130. Por ejemplo, si un intento anterior de detección basada en un servidor falló, el administrador de tiempo de activación 305 puede producir periódicamente un evento de activación para volver a transmitir una solicitud al detector 308 basado en servidor. Cada uno de estos intentos puede usar una imagen de consulta diferente basada en la imagen capturada más reciente. Por ejemplo, después de un fallo de detección, el administrador de tiempo de activación 305 puede producir un evento de activación periódico con una brecha periódica de  $\eta$ , por ejemplo, si el último intento de detección fallido fue hace más de  $\eta$  imágenes, entonces se envía un evento de activación, donde el valor de  $\eta$  puede ser variable.

**[0029]** Cuando se inicia el detector 308 basado en servidor, el servidor 130 recibe los datos asociados con la nueva imagen capturada 300, que pueden incluir la nueva imagen capturada 300 en sí, la información sobre la nueva imagen capturada 300, así como los datos de sensor asociados con la nueva imagen capturada 300. Si un objeto está identificado por el detector 308 basado en servidor, el objeto encontrado, por ejemplo, una imagen de referencia, un modelo 3D del objeto u otra información relevante se proporciona a la plataforma móvil 110, que actualiza su memoria caché local 310. Si el detector 308 basado en servidor no encuentra ningún objeto, el proceso puede recurrir a la activación periódica, por ejemplo, utilizando el administrador de tiempo de activación 305. Si no se detecta ningún objeto después de  $\Gamma$  intentos, por ejemplo, 4 intentos, se considera que el objeto no está en la base de datos y el sistema se restablece a los eventos de activación basados en el detector de cambio de escena.

**[0030]** Si el objeto encontrado está almacenado en la memoria caché local 310, un detector de objetos 312 que se ejecuta en la plataforma móvil 110 realiza un proceso de detección de objetos para identificar el objeto en la vista actual de la cámara y la pose con respecto al objeto y envía la identidad y la pose del objeto al rastreador basado en referencias 314. La pose y la identidad del objeto enviada por el detector de objetos 312 se pueden utilizar para inicializar y comenzar el rastreador basado en referencias 314. En cada imagen capturada posteriormente (por ejemplo, trama de vídeo), el rastreador basado en referencias 314 puede proporcionar la pose con respecto al objeto a un motor de renderizado en la plataforma móvil 110 que coloca el aumento deseado sobre el objeto mostrado o, de otro modo, dentro de una imagen. En una implementación, el detector 308 basado en servidor puede enviar un modelo 3D del objeto, en lugar de una imagen de referencia. En tales casos, el modelo 3D se almacena en la memoria caché local 310 y posteriormente se utiliza como una entrada al rastreador basado en referencias 314. Una vez que se inicializa el rastreador basado en referencias 314, el rastreador basado en referencias 314 recibe cada nueva imagen capturada 300 e identifica la ubicación del objeto rastreado en cada nueva imagen capturada 300, lo que permite que se muestren datos aumentados con respecto al objeto rastreado. El rastreador basado en referencias 314 se puede usar para muchas aplicaciones, como estimaciones de pose, reconocimiento facial, reconocimiento de edificios u otras aplicaciones.

**[0031]** Además, una vez que se inicializa el rastreador basado en referencias 314, el rastreador basado en referencias 314 identifica las regiones de cada nueva imagen capturada 300 donde el objeto identificado está presente y esta información es almacenada en medios para una máscara de rastreo. Por lo tanto, las regiones en las nuevas imágenes de cámara 300 para las que el sistema tiene información completa se identifican y se proporcionan como una entrada al rastreador sin referencias 302 y al detector de cambio de escena 304. El rastreador sin referencias 302 y el detector de cambio de escena 304 continúan recibiendo cada nueva imagen capturada 300 y usan la máscara de rastreo para operar en las regiones restantes de cada nueva imagen capturada 300, es decir, regiones en las que no hay información completa. El uso de la máscara de rastreo como retroalimentación no solo ayuda a reducir los eventos de activación erróneos del detector de cambios de escena 304 debido a los objetos rastreados, sino que también ayuda a reducir la complejidad computacional del rastreador sin referencias 302 y el detector de cambios de escena 304.

**[0032]** En un modo de realización, ilustrado con líneas de puntos en la Fig. 3, el detector 308 basado en servidor puede proporcionar adicionalmente información de la pose para un objeto en la nueva imagen capturada 300 con respecto al objeto en la imagen de referencia. La información de la pose proporcionada por el detector 308 basado en servidor puede usarse junto con los cambios en la pose, según lo determinado por el rastreador sin referencias 302, por un actualizador de pose 316 para producir una pose actualizada. La pose actualizada se puede proporcionar entonces al rastreador basado en referencias 314.

**[0033]** Además, cuando se pierde el rastro temporalmente, se pueden realizar nuevas detecciones posteriores utilizando un detector local 318 buscando en la memoria caché local 310. Si bien la Fig. 3 ilustra el detector local 318 y el detector de objetos 312 por separado por claridad, si se desea, el detector local 318 puede implementar el detector de objetos 312, es decir, el detector de objetos 312 puede realizar las nuevas detecciones. Si el objeto se encuentra en la memoria caché local, la identidad del objeto se usa para reinicializar e iniciar el rastreador basado en referencias 314.

**[0034]** La Fig. 4 ilustra un diagrama de flujo de llamadas para la AR asistida por servidor, en la cual la pose es proporcionada por el servidor 130, como lo ilustran las líneas discontinuas y el actualizador de poses 316 en la Fig. 3. Cuando el detector de cambio de escena 304 indica que una vista ha cambiado (etapa A), el administrador del sistema 320 inicia el proceso de detección basado en servidor, lo que proporciona al detector 308 basado en servidor, por ejemplo, la nueva imagen, que puede estar en formato JPEG o en otro formato, y una solicitud de detección de objeto (etapa B). También se puede enviar información adicionalmente o de forma alternativa al detector 308, como datos de sensor que incluyen información relacionada con la imagen, información de sensores como SPS, lectura del sensor de orientación, giroscopio, brújula, sensor de presión, altímetro, etc. así como datos de usuario, por ejemplo, datos de uso de la aplicación, perfiles de usuario, información de redes sociales, búsquedas anteriores, información de ubicación/sensor, etc. El administrador del sistema 320 también envía un comando al rastreador sin referencias 302 para rastrear el objeto (etapa C). El detector 308 procesa los datos y devuelve al administrador del sistema 320 una lista de objeto(s), como imágenes de referencia para el objeto u objetos, características tales como características SIFT, líneas con descriptores, etc., metadatos (tal como para el aumento), y la pose de nuevo a la aplicación de AR (etapa D). La imagen de referencia para el objeto se agrega a la memoria caché local 310 (etapa E), que acusa recibo de la adición del objeto (etapa F). El rastreador sin referencias 302 proporciona los cambios en la pose entre la imagen inicial y la imagen actual al detector 312 (etapa G). El detector 312 usa la imagen de referencia para encontrar el objeto en la imagen capturada actualmente, proporcionando el identificador del objeto al administrador del sistema 320 (etapa H). Además, la pose proporcionada por el detector 308 basado en servidor es utilizada por el detector 312 junto con los cambios en la pose del rastreador sin referencias 302 para generar una pose actual, que también se proporciona al administrador del sistema 320 (etapa H). El administrador del sistema 320 da instrucciones al rastreador sin referencias 302 para detener el rastreo de objetos (etapa I) y da instrucciones al rastreador basado en referencias 314 para iniciar el rastreo de objetos (etapa J). El rastreo continúa con el rastreador basado en referencias 314 hasta que se pierde el rastro (etapa K).

**[0035]** La Fig. 5 ilustra otro diagrama de flujo de llamadas para la AR asistida por servidor, en el que el servidor 130 no proporciona la pose. El flujo de llamadas es similar al que se muestra en la Fig. 4, excepto que el detector 308 no proporciona información de la pose al administrador del sistema 320 en la etapa D. Por lo tanto, el detector 312 determina la pose basándose en la imagen actual y la imagen de referencia proporcionada por el detector 308 y proporciona esa pose al administrador del sistema 320 (etapa G).

**[0036]** Como se ha analizado anteriormente, el detector de cambio de escena 304 controla la frecuencia de las solicitudes de detección enviadas al servidor 130 basándose en los cambios en una imagen capturada actual con respecto a las imágenes capturadas anteriores. El detector de cambio de escena 304 se utiliza, ya que es deseable comunicarse con el servidor externo 130, para iniciar la detección de objetos solo cuando hay información nueva significativa en la imagen.

**[0037]** La figura 6 ilustra un diagrama de flujo del procedimiento realizado por el detector de cambio de escena 304. El proceso para la detección del cambio de escena se basa en una combinación de métricas del rastreador sin referencias 302 (Fig. 3) y los histogramas de píxeles de imagen. Como se analizó anteriormente, el rastreador sin referencias 302 usa un enfoque tal como flujo óptico, correlación cruzada normalizada y/o cualquiera de los enfoques tales que siguen el movimiento relativo entre imágenes consecutivas, por ejemplo, como correspondencia de punto, línea o región. Un procedimiento basado en histograma puede funcionar bien para ciertos casos de uso, como pasar las páginas de un libro, en los que se produce un cambio significativo en el contenido de información de la escena en un período de tiempo corto, y por lo tanto puede ser beneficioso para el uso en el proceso de detección de escenas; un proceso de rastreo sin referencias puede detectar de manera eficiente los cambios para otros casos de uso, como la toma de una panorámica, donde hay un cambio gradual en el contenido de la información en la escena.

**[0038]** Por lo tanto, como se ilustra en la Fig. 6, se proporciona una imagen de entrada 402. La imagen de entrada es la imagen capturada actual, que puede ser la trama de vídeo o la foto actual. Si la última imagen no activó la detección de cambio de escena (404), entonces se realiza la inicialización (406) del detector de cambio de escena (406). La inicialización incluye dividir la imagen en bloques (408), por ejemplo, 8x8 bloques para una imagen QVGA, y extraer puntos clave de cada bloque usando, por ejemplo, un detector de esquinas FAST (Features from Accelerated Segment Test), en el que se conservan las M esquinas más fuertes (410), donde M puede ser 2. Por supuesto, de forma alternativa se pueden usar otros procedimientos para extraer puntos clave, como las esquinas de Harris, los puntos de característica de Transformación de característica invariable a la escala (SIFT), las Características de robustez aceleradas (SURF) o cualquier otro procedimiento deseado. Se devuelve una señal de no activación (412).

**[0039]** Si la última imagen activó la detección de cambio de escena (404), las métricas se obtienen del rastreador sin referencias 302 (Fig. 3), ilustrado como proceso de flujo óptico 420, e histogramas de píxeles de imagen, ilustrados como proceso de histograma 430. Si se desea, el rastreador sin referencias 302 puede producir métricas usando procesos distintos al flujo óptico, como la correlación cruzada normalizada. El proceso de flujo óptico 420 rastrea las esquinas de una imagen anterior (422), por ejemplo, utilizando una correlación cruzada normalizada, e identifica sus ubicaciones en la imagen actual. Las esquinas pueden haberse extraído previamente dividiendo la imagen en bloques y seleccionando puntos clave de cada bloque usando, por ejemplo, un detector de esquinas FAST en el que se conservan las M esquinas más fuertes basándose en el umbral de esquinas FAST, como se analizó en la inicialización 406 anterior, o en el caso de las esquinas de Harris, se conservan las M esquinas más fuertes basándose en el umbral de Hesse. El rastreo sin referencias se ejecuta para las esquinas elegidas sobre imágenes consecutivas para determinar la ubicación de las esquinas en la imagen actual y las esquinas que se pierden en el rastreo. La intensidad total de las esquinas perdidas en la iteración actual (d en 424), es decir, entre la imagen actual y la imagen anterior, se calcula como una primera métrica de cambio y la intensidad total de las esquinas perdidas desde la activación anterior (D en 426), es decir, entre la imagen actual y la imagen de activación anterior, se calcula como una segunda métrica de cambio, que se proporciona para un cálculo de estadísticas de vídeo 440. El proceso de histograma 430 divide la imagen de entrada actual (denominada C) en B x B bloques y genera un histograma de color  $H_{i,j}^C$  para cada bloque (432), donde i y j son los índices de bloque en la imagen. Se realiza una comparación por bloques de los histogramas (434) con histogramas correspondiente del bloque de la N-ésima imagen anterior  $H_{i,j}^N$  usando, por ejemplo, el procedimiento de chi-cuadrado. La comparación de los histogramas ayuda a determinar la similitud entre la imagen actual y la N-ésima imagen anterior a fin de identificar si la escena ha cambiado significativamente. Por medio de un ejemplo, B puede ser elegido para ser 10. Para comparar los histogramas de la imagen actual y la N-ésima imagen anterior usando el procedimiento de chi-cuadrado, se lleva a cabo el siguiente cálculo:

$$f_{ij} = d(H_{i,j}^C, H_{i,j}^N) = \sum_{\forall k} \frac{(H_{i,j}^C(k) - H_{i,j}^N(k))^2}{H_{i,j}^C(k) + H_{i,j}^N(k)} \quad \text{ec. 1}$$

**[0040]** La comparación de bloques produce una matriz  $f_{ij}$  de valores de diferencia. La matriz  $f_{ij}$  se clasifica y la métrica de cambio de histograma h se determina, por ejemplo, como la media de la mitad de los elementos en el centro de la matriz clasificada  $f_{ij}$  (436). La métrica de cambio de histograma h también se proporciona para el cálculo de estadísticas de vídeo.

**[0041]** Como se analizó anteriormente, si se desea, se puede usar una máscara de rastreo proporcionada por el rastreador basado en referencias 314 (Fig. 3), durante la detección del cambio de escena para reducir las regiones de la imagen de entrada que se monitorizarán para el cambio de escena. La máscara de rastreo identifica las regiones donde se identifica un objeto y, por lo tanto, se puede omitir la monitorización de los cambios de escena. Así, por ejemplo, cuando la imagen de entrada se divide en bloques, por ejemplo, en 422, 432, la máscara de rastreo se puede usar para identificar bloques que caen dentro de las regiones con objetos identificados y, en consecuencia, esos bloques pueden ignorarse.

**[0042]** El cálculo de estadísticas de vídeo 440 recibe las métricas de flujo óptico d, D y la métrica de cambio de histograma h y produce una determinación de la calidad de la imagen, que se proporciona junto con las métricas d, D y h para determinar si se debe activar la detección. Se calcula una métrica de cambio  $\Delta$  y se compara (458) con un umbral para devolver una señal de activación (460). Por supuesto, si la métrica de cambio  $\Delta$  es menor que el umbral, no se devuelve ninguna señal de activación. La métrica de cambio  $\Delta$  se puede calcular (456) en base a las métricas de flujo óptico d, D y la métrica de cambio de histograma h, por ejemplo, como sigue:

$$\Delta = \alpha d + \beta D + \gamma h. \quad \text{ec. 2}$$

**[0043]** Aquí,  $\alpha$ ,  $\beta$  y  $\gamma$  son ponderaciones que se seleccionan de manera adecuada (452) para proporcionar una importancia relativa a las tres estadísticas, d, D y h. En un modo de realización, los valores de  $\alpha$ ,  $\beta$  y  $\gamma$  pueden establecerse en una constante durante toda la ejecución. En un modo de realización alternativo, los valores de  $\alpha$ ,  $\beta$  y  $\gamma$  pueden adaptarse según la posible retroalimentación recibida sobre el rendimiento del sistema o según el caso de uso objetivo. Por ejemplo, el valor de  $\alpha$  y  $\beta$  puede establecerse relativamente alto en comparación con  $\gamma$  para aplicaciones que involucren detecciones de cambio de escena de tipo panorámico porque las estadísticas d y D pueden ser más confiables en este caso. De forma alternativa, los valores de  $\alpha$  y  $\beta$  pueden configurarse para ser relativamente bajos en comparación con  $\gamma$  para aplicaciones que involucren principalmente el tipo de paso de páginas de libros de casos de uso donde la estadística del histograma h puede ser más informativa. El umbral puede adaptarse (454) basándose en la salida del cálculo de estadísticas de vídeo 440, si se desea.

**[0044]** En un caso, si se desea, el proceso de detección de escena puede basarse en métricas del rastreador sin referencias 302, sin métricas de histogramas, por ejemplo, la métrica de cambio  $\Delta$  de la ecuación 2 se puede usar con  $\gamma = 0$ . En otra implementación, la imagen de entrada se puede dividir en bloques y puntos clave extraídos de cada bloque usando, por ejemplo, un detector de esquinas FAST (Features from Accelerated Segment Test), en el que se

retienen las M esquinas más fuertes, como se analizó anteriormente. Si un número suficiente de bloques ha cambiado entre la imagen actual y la imagen anterior, por ejemplo, en comparación con un umbral, se determina que la escena ha cambiado y se devuelve una señal de activación. Se puede considerar que un bloque ha cambiado, por ejemplo, si el número de esquinas rastreadas es menor que otro umbral.

[0045] Además, si lo desea, el proceso de detección de escena puede basarse simplemente en la intensidad total de las esquinas perdidas desde la activación anterior (D en 426) en relación con la intensidad del número total de esquinas en la imagen, por ejemplo, la métrica de cambio  $\Delta$  de la ecuación 2 se puede usar con  $\alpha = 0$  y  $\gamma = 0$ . La intensidad total de las esquinas perdidas desde la activación anterior se puede determinar como:

$$D_c = \sum_{i=t+1}^c \left( \sum_{j \in Li} s_j \right) \quad \text{ec. 3}$$

[0046] En la ecuación 3,  $s_j$  es la intensidad de la esquina  $j$ ,  $t$  es el número de la última imagen de activación,  $c$  es el número de la imagen actual y  $Li$  es el conjunto que contiene los identificadores de esquinas perdidas de la trama  $i$ . Si se desea, se puede usar una métrica de cambio  $\Delta$  diferente, como:

$$\Delta = \frac{D_c}{\sum_{j=1}^{N_T} s_j} \quad \text{ec. 4}$$

donde  $N_T$  es el número total de esquinas en la imagen de activación. La métrica de cambio  $\Delta$  se puede comparar (458) con un umbral.

[0047] Además, como se analizó anteriormente, la máscara de rastreo puede ser utilizada por el detector de cambio de escena 304 para limitar el área de cada imagen en la que se buscan cambios en la escena. En otras palabras, la métrica relevante es la pérdida de intensidad de las esquinas fuera del área de la máscara de activación. Una reducción en el tamaño del área buscada por el detector de cambio de escena 304 conduce a una reducción correspondiente en el número de esquinas que se puede esperar que se detecten. Por lo tanto, se puede usar un parámetro adicional para compensar la pérdida de esquinas debido a la máscara de rastreo, por ejemplo, como sigue:

$$\lambda = \frac{\text{intensidad de las esquinas en la máscara}}{\text{área de la máscara}} \quad \text{ec. 5}$$

[0048] El parámetro de compensación  $\lambda$  se puede usar para ajustar la métrica de cambio  $\Delta$ . Por ejemplo, si el proceso de detección de escena se basa simplemente en la intensidad total de las esquinas perdidas en el área sin máscara desde la activación anterior (D), la métrica de cambio  $\Delta$  de la ecuación 4 se puede modificar como:

$$\Delta = \frac{D_c + \lambda(A - A_c)}{\sum_{j=1}^{N_T} s_j} \quad \text{ec. 6}$$

donde  $D_c$  es proporcionada por la ecuación 3 (con  $Li$  definido como el conjunto que contiene identificadores de esquinas perdidas en el área sin máscara de la trama  $i$ ),  $A_c$  es el área de la máscara para la imagen  $c$ , y  $A$  se inicializa a  $A_{t+1}$ .

[0049] La Fig. 7 es una tabla que ilustra el rendimiento del sistema para un caso de uso típico de paso de páginas de libros en el que se pasan cinco páginas en 50 segundos. La Fig. 7 ilustra el número de transmisiones de red requeridas para solicitar detecciones de objetos en función de la brecha de activación mínima en segundos. Cuanto menor sea el número de transmisiones de red necesarias para la misma brecha de activación mínima, mayor será el rendimiento. Se ilustran varias curvas, incluida la curva 480 para una activación periódica, la curva 482 para un detector de cambio de escena (SCD) basado en el flujo óptico sin estadísticas de histograma ( $\gamma = 0$ ) y sin el rastreador basado en referencias 314 (Fig. 3), la curva 484 para el detector de cambio de escena (SCD) basado en el flujo óptico sin estadísticas de histograma ( $\gamma = 0$ ), pero con el rastreador basado en referencias 314 y la curva 486 para un detector de cambio de escena (SCD) basado en histograma y flujo óptico combinados (como se describe en la Fig. 6) junto con

el rastreador basado en referencias 314 y el administrador de temporización 305 (Fig. 3). Como puede verse en la Fig. 7, el sistema combinado supera a otros sistemas en el caso de uso de paso de páginas.

**[0050]** La Fig. 8 ilustra un enfoque para el reconocimiento facial mediante el proceso de AR asistido por servidor. Como se ilustra en la Figura 8, una plataforma móvil 110 realiza la adquisición de datos 502, que incluye la adquisición de una imagen de una cara, así como la adquisición de cualquier otra información útil del sensor, como SPS o datos de sensor de posición/movimiento. La plataforma móvil 110 realiza la detección de caras 504 y proporciona los datos de las caras para una o más caras (que pueden ser una imagen de la cara), así como cualquier otra información útil, como SPS o datos de sensores de posición/movimiento al servidor 130, como lo indica la flecha 506. La plataforma móvil 110 rastrea el movimiento 2D de la cara (508). El servidor 130 realiza el reconocimiento facial 510 basándose en los datos faciales proporcionados, por ejemplo, utilizando datos recuperados de una base de datos global 512 y almacenados en una memoria caché local 514. El servidor 130 proporciona datos relacionados con la cara, por ejemplo, la identidad u otra información deseada, a la plataforma móvil 110, que utiliza los datos recibidos para anotar la cara mostrada en la pantalla 114 con el nombre, etc. o proporcionar de otro modo datos aumentados renderizados (516).

**[0051]** La Fig. 9 ilustra otro enfoque para el reconocimiento facial mediante el proceso de AR asistido por servidor. La Fig. 9 es similar al enfoque ilustrado en la Fig. 8, siendo los elementos designados de modo similar los mismos. Sin embargo, como se ilustra en la Fig. 9, la imagen se proporciona al servidor 130 (508') y la detección de caras (504') la realiza el servidor 130.

**[0052]** La Fig. 10 ilustra un enfoque de una búsqueda visual utilizando el proceso de AR asistido por servidor. Como se ilustra en la Fig. 10, una plataforma móvil 110 realiza la adquisición de datos (520), que incluye la adquisición de una imagen del objeto deseado, así como la adquisición de cualquier otra información útil del sensor, como SPS o datos de sensor de posición/movimiento. La plataforma móvil 110 realiza la detección de características (522) y proporciona las características detectadas, así como cualquier otro dato útil, como SPS o datos de sensor de posición/movimiento al servidor 130, como se indica mediante la flecha 526. La plataforma móvil 110 rastrea el movimiento 2D de las características (524). El servidor 130 realiza el reconocimiento de objetos 528 basándose en las características proporcionadas, por ejemplo, utilizando datos recuperados de una base de datos global 530 y almacenados en una memoria caché local 532. El servidor 130 también puede realizar un registro global (534), por ejemplo, para obtener una imagen de referencia, una pose, etc. El servidor 130 proporciona los datos relacionados con el objeto, como una imagen de referencia, una pose, etc., a la plataforma móvil 110, que utiliza los datos recibidos para realizar el registro local (536). La plataforma móvil 110 puede entonces presentar los datos aumentados deseados con respecto al objeto mostrado en la pantalla 114 (538).

**[0053]** La Fig. 11 ilustra otro enfoque de una búsqueda visual utilizando el proceso de AR asistido por servidor. La Fig. 11 es similar al enfoque ilustrado en la Fig. 10, siendo los elementos designados de modo similar los mismos. Sin embargo, como se ilustra en la Fig. 11, la imagen completa se proporciona al servidor 130 (526') y la detección de características (522') la realiza el servidor 130.

**[0054]** La Fig. 12 ilustra un enfoque del rastreo basado en referencias utilizando el proceso asistido por servidor. Como se ilustra en la Fig. 12, una plataforma móvil 110 realiza la adquisición de datos (540), que incluye la adquisición de una imagen del objeto deseado, así como la adquisición de cualquier otra información útil del sensor, como SPS o datos de sensor de posición/movimiento. En algunos modos de realización, la plataforma móvil 110 puede generar información complementaria (541), tal como reconocimiento de texto o la lectura de códigos de barras, etc. La plataforma móvil 110 realiza la detección de características (542) y proporciona las características detectadas, así como cualquier otro dato útil, como SPS o datos de sensor de posición/movimiento al servidor 130, e información complementaria si se genera, como se indica mediante la flecha 546. La plataforma móvil 110 rastrea el movimiento 2D de las características (544), por ejemplo, mediante el rastreo de puntos, líneas o regiones, o un flujo óptico denso. En algunos modos de realización, el servidor 130 puede realizar un reconocimiento de múltiples planos (548) usando las características proporcionadas. Una vez que se han identificado los planos, el reconocimiento de objetos (550) se puede realizar en un plano individual o en un grupo de planos, por ejemplo, usando datos recuperados de una base de datos global 552 y almacenados en una memoria caché local 554. Si se desea, se puede utilizar cualquier otro procedimiento de reconocimiento. En algunos modos de realización, el servidor 130 también puede realizar una estimación de pose (555) si se desea, que se puede proporcionar en seis grados de libertad, con matrices homográficas, afines, rotacionales y traslacionales. El servidor 130 proporciona los datos relacionados con el objeto, como una imagen de referencia, a la plataforma móvil 110, que utiliza los datos recibidos para realizar el registro local (556), que puede ser un registro de homografía local o un registro de matriz esencial local. Como se describió anteriormente, la plataforma móvil 110 puede incluir una memoria caché local 557 para almacenar los datos recibidos, lo que puede ser beneficioso para las nuevas detecciones posteriores que pueden realizarse en la plataforma móvil 110 si se pierde el rastro. La plataforma móvil 110 puede entonces representar los datos aumentados deseados con respecto al objeto mostrado en la pantalla 114 (558).

**[0055]** La Fig. 13 ilustra otro enfoque del rastreo basado en referencias utilizando el proceso asistido por servidor. La Fig. 13 es similar al enfoque ilustrado en la Fig. 12, siendo los elementos designados de modo similar los mismos.

Sin embargo, como se ilustra en la Fig. 13, la imagen completa se proporciona al servidor 130 (546') y la detección de características (542') la realiza el servidor 130.

**[0056]** La Fig. 14 ilustra un enfoque para la creación de modelos 3D utilizando el proceso asistido por servidor. Como se ilustra en la Fig. 14, una plataforma móvil 110 realiza la adquisición de datos (560), que incluye la adquisición de una imagen del objeto deseado, así como la adquisición de cualquier otra información útil del sensor, como SPS o datos de sensor de posición/movimiento. La plataforma móvil 110 realiza un procesamiento de imágenes 2D (562) y rastrea el movimiento (564) utilizando el rastreo sin referencias, por ejemplo, el flujo óptico o los enfoques basados en la correlación cruzada normalizada. La plataforma móvil 110 realiza un registro local de seis grados de libertad (568) para obtener el cálculo aproximado de la pose. Estos datos junto con las imágenes en ciertos modos de realización pueden proporcionarse al servidor 130. El servidor 130 puede entonces realizar un ajuste de paquetes para refinar el registro (570). Dado un conjunto de imágenes y correspondencias de puntos 3D desde diferentes puntos de vista, los algoritmos de ajuste de paquetes ayudan a estimar las coordenadas 3D del punto en un sistema de coordenadas de referencia conocido y ayudan a identificar el movimiento relativo de la cámara entre diferentes puntos de vista. Los algoritmos de ajuste de paquetes son, en general, operaciones computacionalmente intensivas y se pueden realizar de manera eficiente en el lado del servidor pasando información complementaria de la plataforma móvil 110 e información adicional si está disponible desde la memoria caché local 572. Después de estimar la ubicación de los puntos 3D y la pose relativa, se pueden proporcionar directamente a la plataforma móvil 110. De forma alternativa, los modelos 3D del objeto pueden construirse en el servidor basándose en los datos y dichos datos pueden enviarse a la plataforma móvil 110. La plataforma móvil 110 puede entonces presentar los datos aumentados deseados con respecto al objeto mostrado en la pantalla 114 (576) utilizando la información obtenida del servidor 130.

**[0057]** Se debe tener en cuenta que toda la configuración del sistema puede ser adaptable dependiendo de la capacidad de la plataforma móvil 110, el servidor 130 y la interfaz de comunicación, por ejemplo, la red 120. Si la plataforma móvil 110 es un dispositivo de gama baja sin un procesador dedicado, la mayoría de las operaciones pueden estar descargadas en el servidor 130. Por otro lado, si la plataforma móvil 110 es un dispositivo de gama alta que tiene una buena capacidad de cálculo, la plataforma móvil 110 puede seleccionar realizar algunas de las tareas y descargar menos tareas al servidor 130. Además, el sistema puede ser adaptable para manejar diferentes tipos de interfaces de comunicación dependiendo de, por ejemplo, el ancho de banda disponible en la interfaz.

**[0058]** En una implementación, el servidor 130 puede proporcionar retroalimentación a la plataforma móvil 110 sobre la tarea y qué partes de una tarea pueden descargarse en el servidor 130. Dicha retroalimentación puede basarse en las capacidades del servidor 130, el tipo de operaciones a realizar, el ancho de banda disponible en el canal de comunicación, los niveles de potencia de la plataforma móvil 110 y/o el servidor 130, etc. Por ejemplo, el servidor 130 puede recomendar que la plataforma móvil 110 envíe una versión de menor calidad de la imagen si la conexión de red es mala y las velocidades de transferencia de datos son bajas. El servidor 130 también puede sugerir que la plataforma móvil realice un mayor procesamiento de los datos y envíe los datos procesados al servidor 130 si las velocidades de transferencia de datos son bajas. Por ejemplo, la plataforma móvil 110 puede calcular características para la detección de objetos y enviar las características en lugar de enviar la imagen completa si el enlace de comunicación tiene una baja velocidad de transferencia de datos. El servidor 130 puede recomendar de forma alternativa que la plataforma móvil 110 envíe una versión de mayor calidad de la imagen o envíe imágenes con mayor frecuencia (reduciendo así la brecha mínima de trama  $\eta$ ) si la conexión de red es buena o si los intentos anteriores de reconocer un objeto en la imagen han fallado.

**[0059]** Además, la arquitectura de servidor móvil introducida en el presente documento también puede extenderse a escenarios donde se utiliza más de una plataforma móvil 110. Por ejemplo, dos plataformas móviles 110 pueden estar viendo el mismo objeto 3D desde diferentes ángulos y el servidor 130 puede realizar un ajuste de paquetes conjunto de los datos obtenidos de ambas plataformas móviles 110 para crear un buen modelo 3D del objeto. Una aplicación de este tipo puede ser útil para aplicaciones como juegos de varios jugadores o similares.

**[0060]** La Fig. 15 es un diagrama de bloques de una plataforma móvil 110 que puede realizar un procesamiento distribuido usando la detección basada en servidor. La plataforma móvil 110 incluye la cámara 112, así como una interfaz de usuario 150 que incluye la pantalla 114 capaz de mostrar imágenes capturadas por la cámara 112. La interfaz de usuario 150 también puede incluir un teclado 152 u otro dispositivo de entrada a través del cual el usuario puede introducir información en la plataforma móvil 110. Si se desea, el teclado 152 puede obviarse integrando un teclado virtual en la pantalla 114 con un sensor táctil. La interfaz de usuario 150 también puede incluir un micrófono 154 y un altavoz 156, por ejemplo, si la plataforma móvil es un teléfono celular.

**[0061]** La plataforma móvil 110 puede incluir un transceptor inalámbrico 162, que puede usarse para comunicarse con el servidor externo 130 (Fig. 3), como se analizó anteriormente. La plataforma móvil 110 puede incluir opcionalmente características adicionales que pueden ser útiles para aplicaciones de AR, tales como sensores de movimiento 164 que incluyen, por ejemplo, acelerómetros, giroscopios, brújula electrónica u otros elementos similares de detección de movimiento, y un receptor 166 de sistema de posicionamiento por satélite (SPS) capaz de recibir señales de posicionamiento desde un sistema SPS. Por supuesto, la plataforma móvil 110 puede incluir otros elementos no relacionados con la presente divulgación.

**[0062]** La plataforma móvil 110 también incluye una unidad de control 170 que está conectada y se comunica con la cámara 112 y el transceptor inalámbrico 162, junto con otras características, como la interfaz de usuario 150, los sensores de movimiento 164 y el receptor SPS 166, si se utilizan. La unidad de control 170 acepta y procesa los datos de la cámara 112 y controla la comunicación con el servidor externo a través del transceptor inalámbrico 162 en respuesta, como se analizó anteriormente. La unidad de control 170 puede ser proporcionada por un procesador 171 y una memoria asociada 172, que puede incluir un software 173 ejecutado por el procesador 171 para realizar los procedimientos o partes de los procedimientos descritos en el presente documento. La unidad de control 170 puede incluir adicionalmente, o de forma alternativa, hardware 174 y/o firmware 175.

**[0063]** La unidad de control 170 incluye el detector de cambio de escena 304 que activa la comunicación con el servidor externo según se ha analizado anteriormente. También se pueden incluir componentes adicionales, como el administrador de tiempo de activación 305 y el estimador de calidad de imagen 306, ilustrados en la Fig. 3. La unidad de control 170 incluye además el rastreador sin referencias 302, el rastreador basado en referencias 314 y la unidad de detección 312, que se usa para detectar objetos en una imagen actual en base a objetos almacenados en la memoria caché local, por ejemplo, en la memoria 172. La unidad de control 170 incluye además la unidad de realidad aumentada (AR) 178 para generar y mostrar información de AR en la pantalla 114. El detector de cambio de escena 304, el rastreador sin referencias 302, el rastreador basado en referencias 314, la unidad de detección 312 y la unidad de AR 178 se ilustran por separado y están separados del procesador 171 para mayor claridad, pero pueden ser una sola unidad y/o implementarse en el procesador 171 en base a instrucciones en el software 173 que se leen y ejecutan en el procesador 171. Se entenderá como se usa en el presente documento que el procesador 171, así como uno o más del detector de cambio de escena 304, el rastreador sin referencias 302, el rastreador basado en referencias 314, la unidad de detección 312 y la unidad de AR 178 pueden, pero no necesariamente incluyen, uno o más microprocesadores, procesadores integrados, controladores, circuitos integrados de específicos de la aplicación (ASIC), procesadores de señales digitales (DSP) y similares. El término procesador pretende describir las funciones implementadas por el sistema en lugar de hardware específico. Como se usa en el presente documento, el término "memoria" se refiere a cualquier tipo de medio de almacenamiento informático, incluyendo memoria a largo plazo, a corto plazo u otra memoria asociada a la plataforma móvil, y no está limitado a ningún tipo particular de memoria o número de memorias, ni al tipo de medio en el que se almacene la memoria.

**[0064]** Las metodologías descritas en el presente documento pueden implementarse por diversos medios, en función de la aplicación. Por ejemplo, estas metodologías pueden implementarse en hardware 174, firmware 175, software 173 o en cualquier combinación de los mismos. Para una implementación en hardware, las unidades de procesamiento pueden implementarse en uno o más circuitos integrados específicos de la aplicación (ASIC), procesadores de señales digitales (DSP), dispositivos de procesamiento de señales digitales (DSPD), dispositivos lógicos programables (PLD), matrices de puertas programables in situ (FPGA), procesadores, controladores, microcontroladores, microprocesadores, dispositivos electrónicos, otras unidades electrónicas diseñadas para desempeñar las funciones descritas en el presente documento, o en una combinación de los mismos. Por lo tanto, el dispositivo para adquirir datos de sensor puede comprender la cámara 112, el receptor SPS 166 y los sensores de movimiento 164, así como el procesador que puede producir información complementaria, como el reconocimiento de texto o la lectura de códigos de barras, basándose en la imagen producida por la cámara 112 u otros medios para adquirir datos de sensor. El dispositivo para determinar si hay un evento de activación que comprenda un cambio en los datos de sensor en relación con los datos de sensor adquiridos previamente comprende la unidad de detección 312, que puede ser implementada por el procesador 171 realizando instrucciones incorporadas en el software 173, o en el hardware 174 o firmware 175, u otros medios para determinar si hay un evento de activación que comprenda un cambio en los datos de sensor en relación con los datos de sensor adquiridos previamente. El dispositivo para transmitir los datos de sensor a un servidor cuando hay un evento de activación comprende el transceptor inalámbrico 162 u otros medios para transmitir los datos de sensor a un servidor cuando hay un evento de activación. El dispositivo para recibir información relacionada con los datos de sensor desde el servidor comprende el transceptor inalámbrico 162 u otros medios para recibir información relacionada con los datos de sensor desde el servidor. El dispositivo para obtener una pose de la plataforma móvil con respecto al objeto comprende el rastreador sin referencias 302, el transceptor inalámbrico 162 u otros medios para obtener una pose de la plataforma móvil con respecto al objeto. El dispositivo para rastrear el objeto usando la pose y la imagen de referencia del objeto comprende el rastreador basado en referencias 314 u otros medios para rastrear el objeto usando la pose y la imagen de referencia del objeto. El dispositivo para determinar si hay un cambio de escena en la imagen capturada con respecto a una imagen capturada anterior comprende el detector de cambio de escena 304, que puede ser implementado por el procesador 171 realizando instrucciones incorporadas en el software 173, o en el hardware 174 o el firmware 175, u otros medios para determinar si hay un cambio de escena en la imagen capturada con respecto a una imagen capturada anterior.

**[0065]** En una implementación en firmware y/o software, las metodologías pueden implementarse con módulos (por ejemplo, procedimientos, funciones, etc.) que realicen las funciones descritas en el presente documento. Cualquier medio legible por máquina que realice instrucciones de forma tangible puede usarse para implementar las metodologías descritas en el presente documento. Por ejemplo, el software 173 puede incluir códigos de programa almacenados en la memoria 172 y ejecutados por el procesador 171. La memoria puede implementarse dentro del procesador o externa al procesador 171.

- 5 **[0066]** Si se implementan en firmware y/o en software, las funciones pueden almacenarse como una o más instrucciones o código en un medio legible por ordenador. Los ejemplos incluyen medios legibles por ordenador no transitorios codificados con una estructura de datos y medios legibles por ordenador codificados con un programa informático. Los medios legibles por ordenador incluyen medios de almacenamiento informáticos físicos. Un medio de almacenamiento puede ser cualquier medio disponible al que pueda accederse mediante un ordenador. A modo de ejemplo, y no de manera limitativa, tales medios legibles por ordenador pueden comprender RAM, ROM, memoria Flash, EEPROM, CD-ROM u otro almacenamiento de disco óptico, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio que pueda usarse para almacenar código de programa deseado en forma de instrucciones o estructuras de datos y al que puede accederse mediante un ordenador;
- 10 tal y como se usa en el presente documento, un disco incluye un disco compacto (CD), un disco de láser, un disco óptico, un disco versátil digital (DVD), un disco flexible y un disco Blu-ray, donde los discos reproducen normalmente datos de manera magnética o de manera óptica con láser. Las combinaciones de lo anterior se deberían incluir también dentro del alcance de los medios legibles por ordenador.
- 15 **[0067]** Aunque la presente invención se ilustra en relación con modos de realización específicos con fines de instrucción, la presente invención no se limita a los mismos. Sin embargo, se entenderá que se pueden realizar diversas adaptaciones y modificaciones sin apartarse del alcance de la invención.

## REIVINDICACIONES

1. Un procedimiento para el rastreo de objetos, que comprende:
  - 5 adquirir (202) datos de sensor usando una plataforma móvil (110) en el que los datos de sensor comprenden una imagen capturada (104) de un objeto (102);  
rastrear (302) el objeto (102) usando un rastreador sin referencias;
  - 10 determinar (204) si hay un evento de activación que comprenda un cambio en los datos de sensor en relación con los datos de sensor adquiridos previamente;
  - transmitir (206) los datos de sensor a un servidor (130) cuando se produce el evento de activación;
  - 15 recibir (208) información relacionada con los datos de sensor desde el servidor (130), comprendiendo la información una identidad del objeto; e  
inicializar e iniciar un rastreador basado en referencias (314) utilizando la identidad del objeto.
- 20 2. El procedimiento según la reivindicación 1, comprendiendo el procedimiento, además, determinar una calidad de la imagen capturada (104) antes de transmitir (206) los datos de sensor al servidor (130), en el que los datos de sensor se transmiten (206) al servidor (130) solo si la calidad de la imagen capturada (104) es mejor que un umbral.
- 25 3. El procedimiento según la reivindicación 1, en el que determinar la calidad de la imagen capturada (104) comprende al menos uno de analizar un grado de nitidez de la imagen capturada (104), analizar una cantidad de esquinas detectadas en la imagen capturada (104); y utilizar estadísticas obtenidas a partir de la imagen (104) con un clasificador de aprendizaje.
- 30 4. El procedimiento según la reivindicación 1, que comprende, además, representar el aumento con respecto al objeto (102) basándose en la información relacionada con los datos de sensor recibidos desde el servidor (130).
5. El procedimiento según la reivindicación 1, que comprende además:
  - 35 obtener (210) una pose de la plataforma móvil (110) con respecto al objeto (102); y  
rastrear (212) el objeto (102) usando la pose y la información relacionada con los datos de sensor.
- 40 6. El procedimiento según la reivindicación 5, en el que la información relacionada con los datos de sensor comprende una imagen de referencia (106) del objeto (102), y en el que obtener (210) la pose comprende recibir del servidor (130) una primera pose basada en la imagen capturada (104) y la imagen de referencia (106).
- 45 7. El procedimiento según la reivindicación 6, que comprende, además, rastrear sin referencias el objeto (102) hasta que se recibe (210) la primera pose desde el servidor (130).
8. El procedimiento según la reivindicación 6, que comprende además:
  - 50 adquirir una segunda imagen capturada del objeto (102) cuando se recibe (210) la primera pose desde el servidor (130);  
rastrear el objeto (102) entre la imagen capturada (104) y la segunda imagen capturada para determinar un cambio progresivo; y
  - 55 utilizar el cambio progresivo y la primera pose para obtener la pose de la plataforma móvil (110) con respecto al objeto (102).
9. El procedimiento según la reivindicación 6, que comprende además:
  - 60 adquirir una segunda imagen capturada del objeto (102);  
detectar el objeto (102) en la segunda imagen capturada usando la imagen de referencia (106);  
usar el objeto (102) detectado en la segunda imagen capturada y la imagen de referencia (106) para obtener  
65 la pose de la plataforma móvil (110) con respecto al objeto (102); y

utilizar la pose para inicializar el rastreo basado en referencias del objeto (102).

- 5 **10.** El procedimiento según la reivindicación 1, en el que determinar (204) si existe el evento de activación comprende determinar si hay un cambio de escena en la imagen capturada (104) con respecto a una imagen capturada anterior;

en el que determinar si hay un cambio de escena comprende:

10 determinar una primera métrica de cambio utilizando la imagen capturada (104) y la imagen capturada anterior;

determinar una segunda métrica de cambio utilizando la imagen capturada (104) y una segunda imagen capturada anterior de un evento de activación anterior;

15 generar una métrica de cambio de histograma para la imagen capturada (104); y

utilizar la primera métrica de cambio, la segunda métrica de cambio y la métrica de cambio de histograma para determinar el cambio de escena.

- 20 **11.** El procedimiento según la reivindicación 1, comprendiendo además el procedimiento:

adquirir imágenes capturadas adicionales del objeto (102);

identificar el objeto (102) en las imágenes capturadas adicionales utilizando la identificación del objeto;

25 generar una máscara de rastreo para las imágenes capturadas adicionales basándose en la identificación del objeto, indicando la máscara de rastreo las regiones en las imágenes capturadas adicionales donde se identifica el objeto (102);

30 usar la máscara de rastreo con las imágenes capturadas adicionales del objeto (102) para identificar las regiones restantes de las imágenes capturadas adicionales; y

detectar eventos de activación que comprenden cambios de escena en las regiones restantes de las imágenes capturadas adicionales.

- 35 **12.** El procedimiento según la reivindicación 1, en el que los datos de sensor incluyen además uno o más de los siguientes: datos de imagen, datos de sensor de movimiento, datos de posición, reconocimiento de código de barras, resultados de detección de texto e información contextual;

40 en el que la información contextual incluye preferentemente uno o más de los siguientes: comportamiento del usuario, preferencias del usuario, ubicación, información sobre el usuario, hora del día y calidad de iluminación; y

45 en el que los datos de sensor comprenden una imagen de una cara y la información recibida (208) del servidor (130) comprende una identidad asociada con la cara.

- 50 **13.** El procedimiento según la reivindicación 1, en el que los datos de sensor comprenden una pluralidad de imágenes (104) de un objeto (102) capturado con una cámara (112) en diferentes posiciones y una estimación aproximada de una pose de la cámara (112) con respecto al objeto (102) y la información recibida (208) del servidor (130) comprenden al menos uno de un refinamiento de la pose y un modelo tridimensional del objeto (102).

- 14.** Una plataforma móvil, que comprende:

55 medios para adquirir datos de sensor en los que los medios para adquirir datos de sensor son una cámara (112) y los datos de sensor son una imagen capturada (104) de un objeto (102);

medios para rastrear (302) el objeto (102) usando un rastreador sin referencias;

60 medios para determinar si hay un evento de activación (304) que comprende un cambio en los datos de sensor en relación con los datos de sensor adquiridos previamente;

medios para transmitir (162) los datos de sensor a un servidor (130) cuando hay un evento de activación; y

65 medios para recibir información (162) relacionada con los datos de sensor desde el servidor (130), comprendiendo la información una identificación de objeto; y

medios para inicializar e iniciar un rastreador basado en referencias (314) utilizando la identidad del objeto.

- 5      **15.**      Un medio no transitorio legible por ordenador que incluye un código de programa almacenado en él, que comprende:

código de programa para realizar un procedimiento de acuerdo con una de las reivindicaciones 1 a 13.

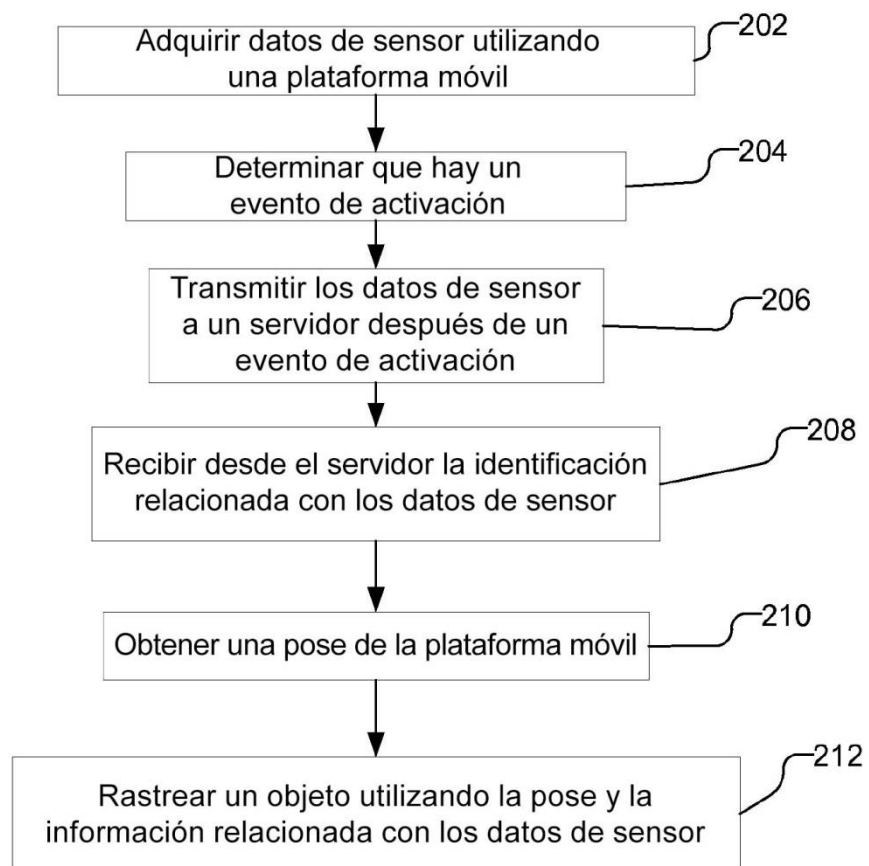
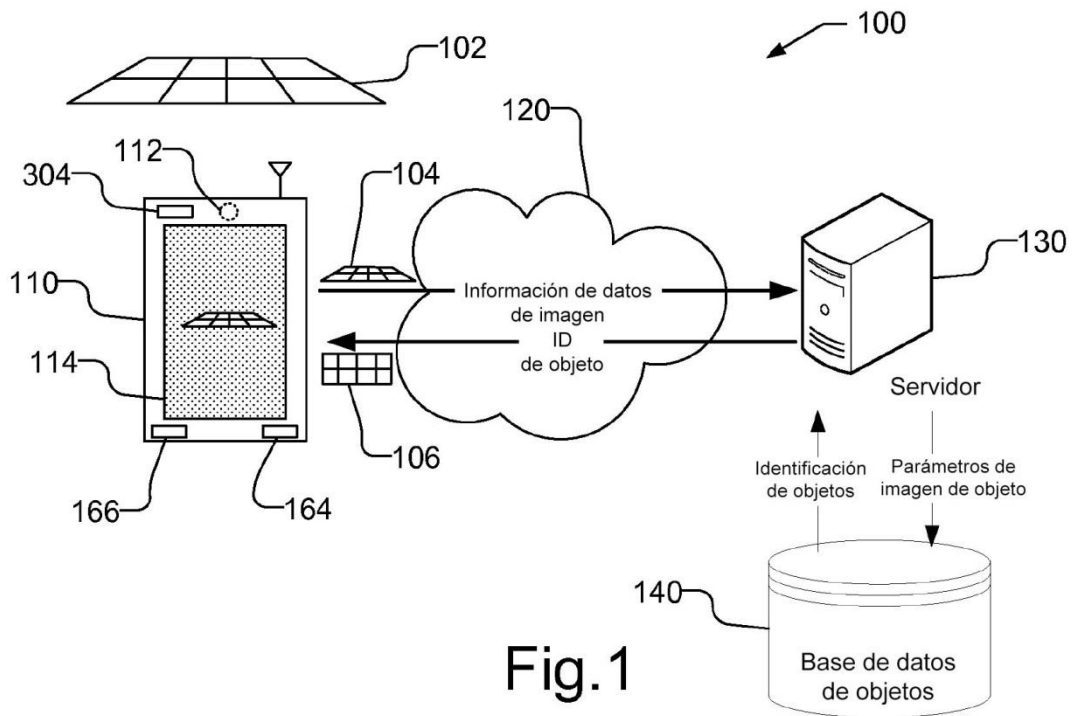
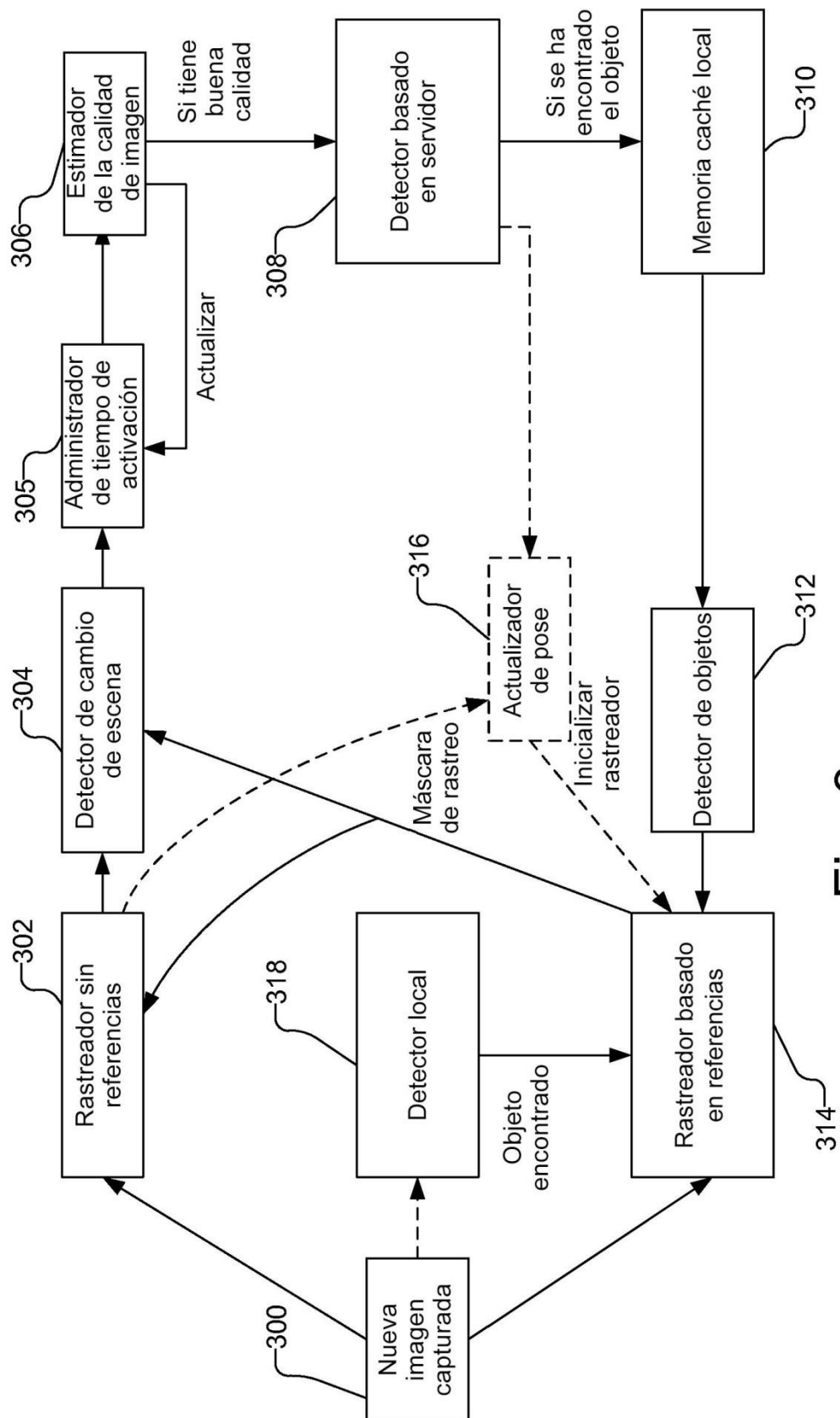


Fig. 2



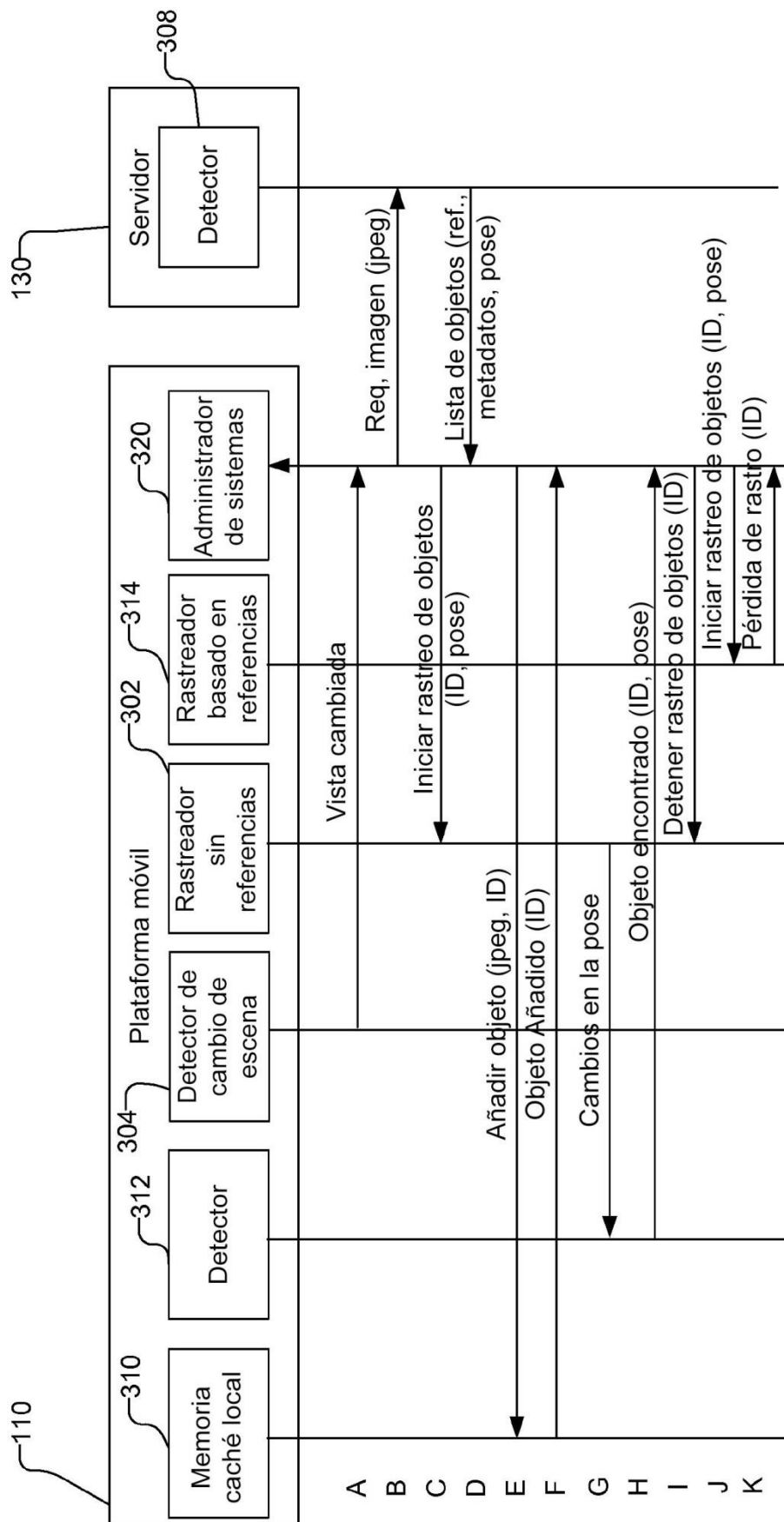


Fig. 4

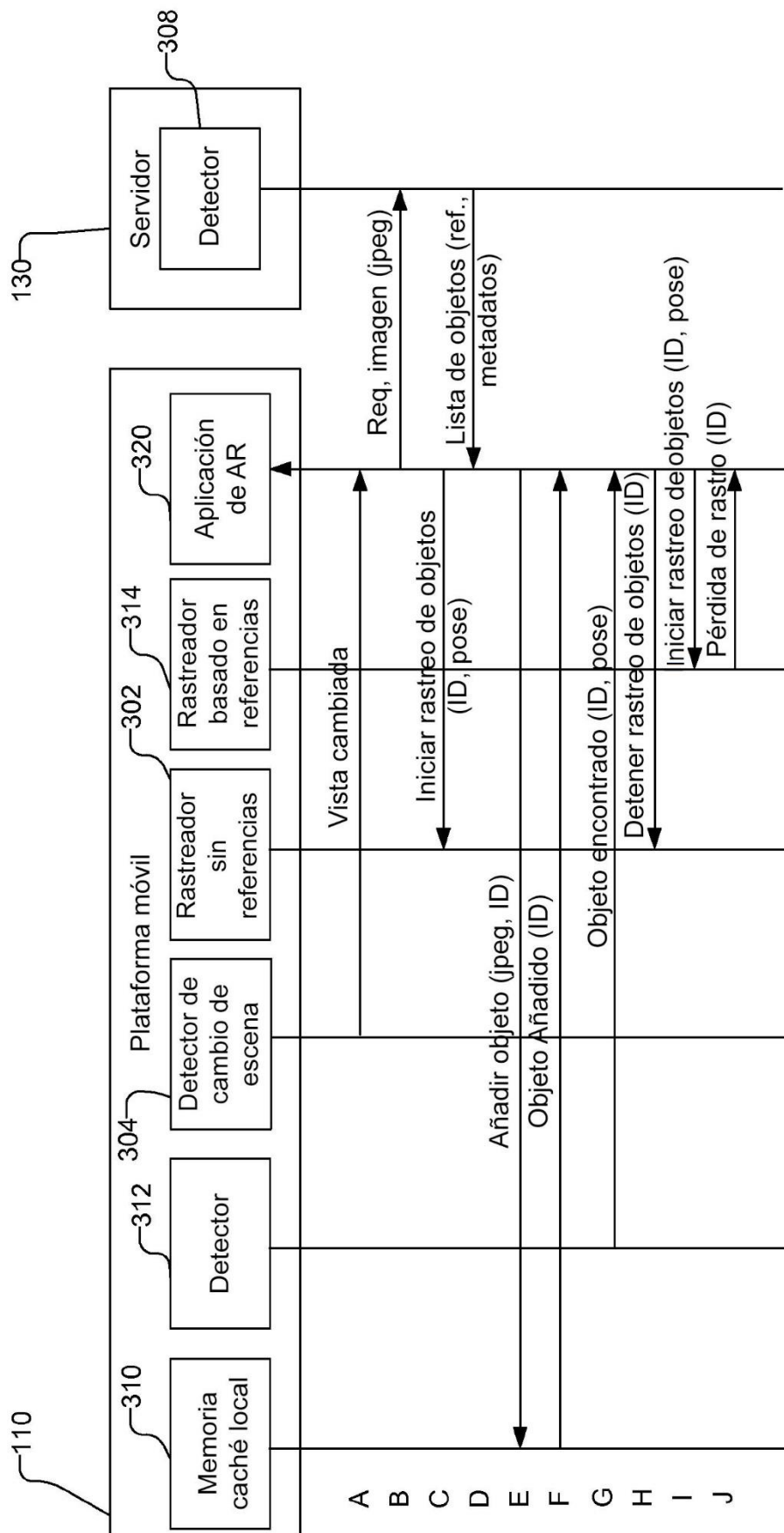


Fig. 5

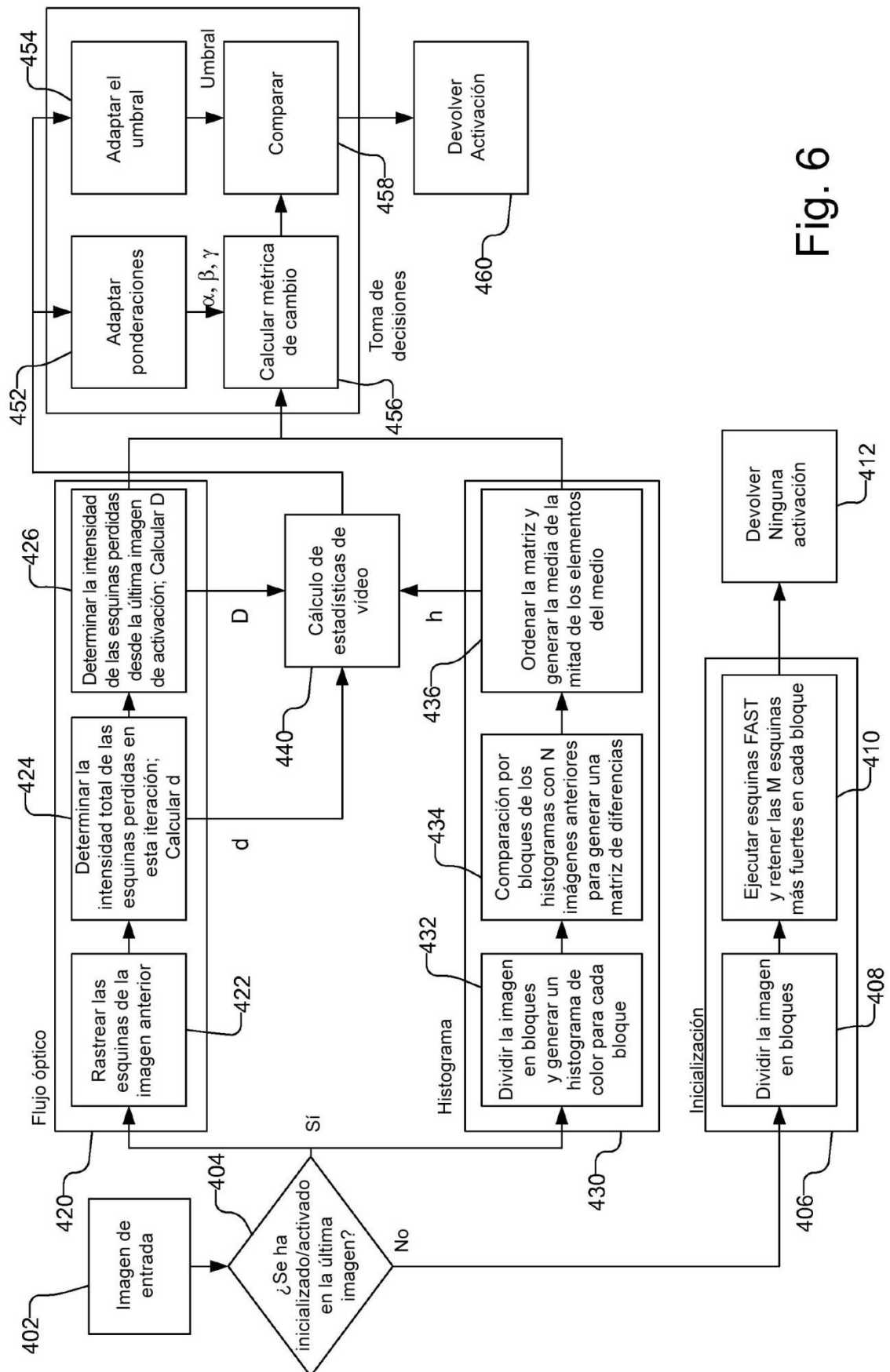


Fig. 6

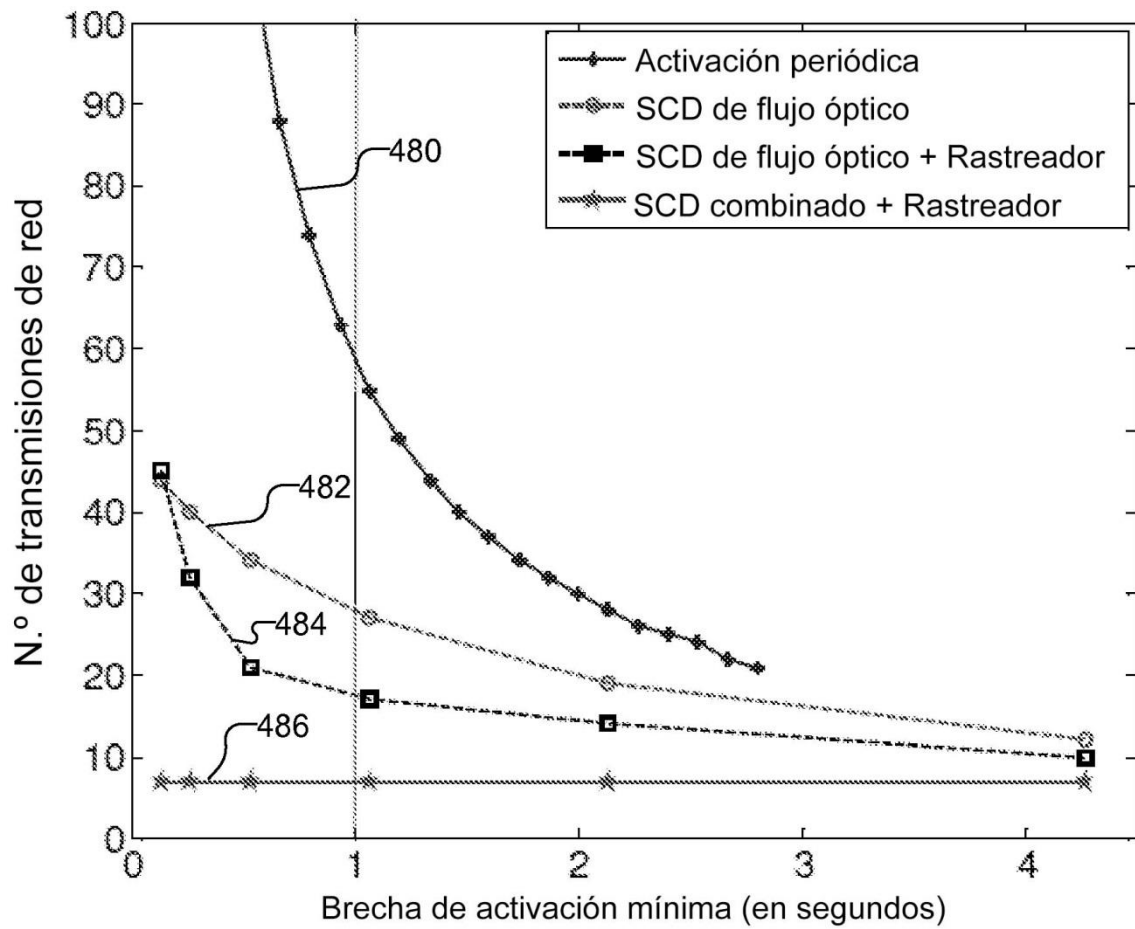


Fig. 7

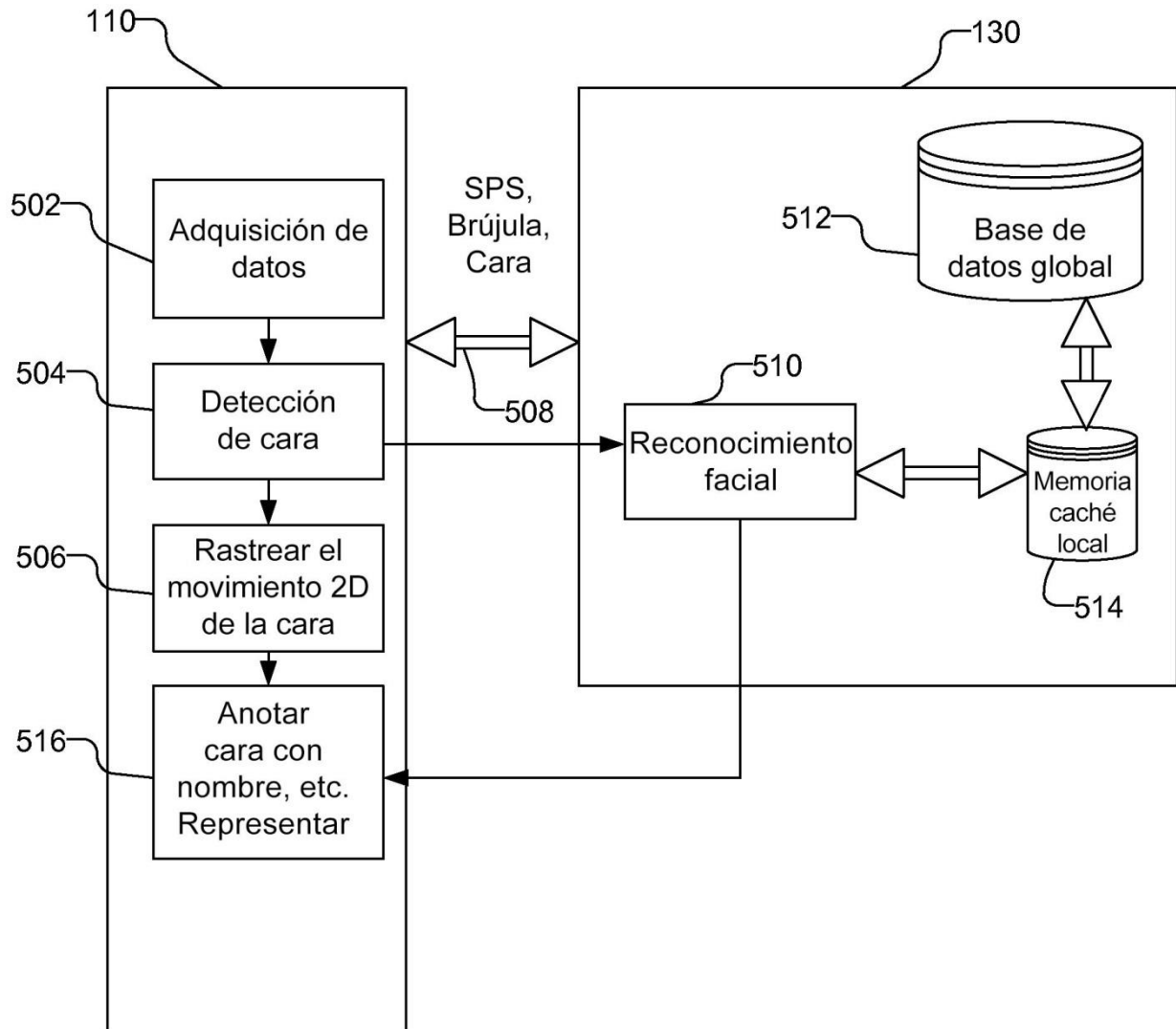


Fig. 8

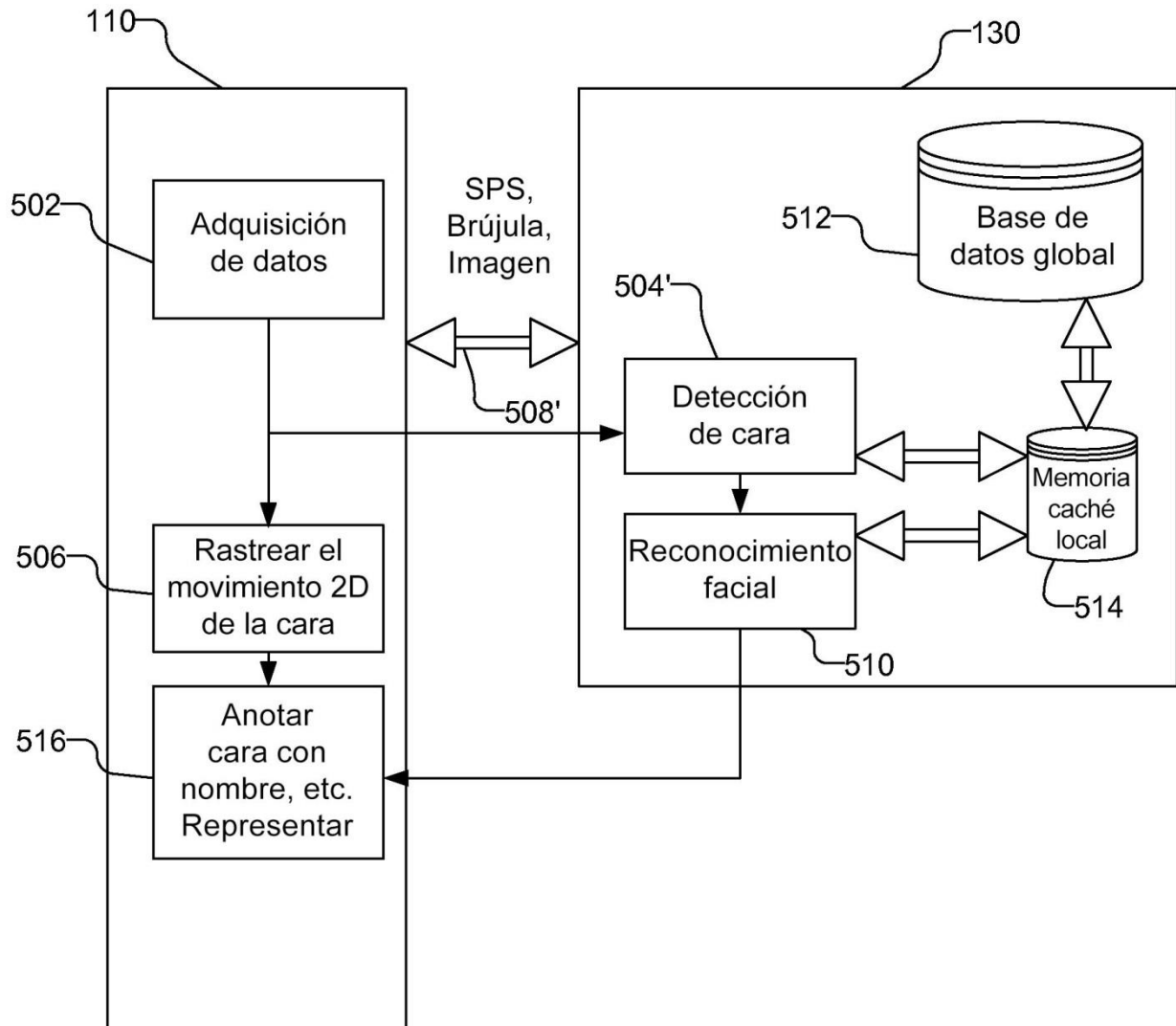


Fig. 9

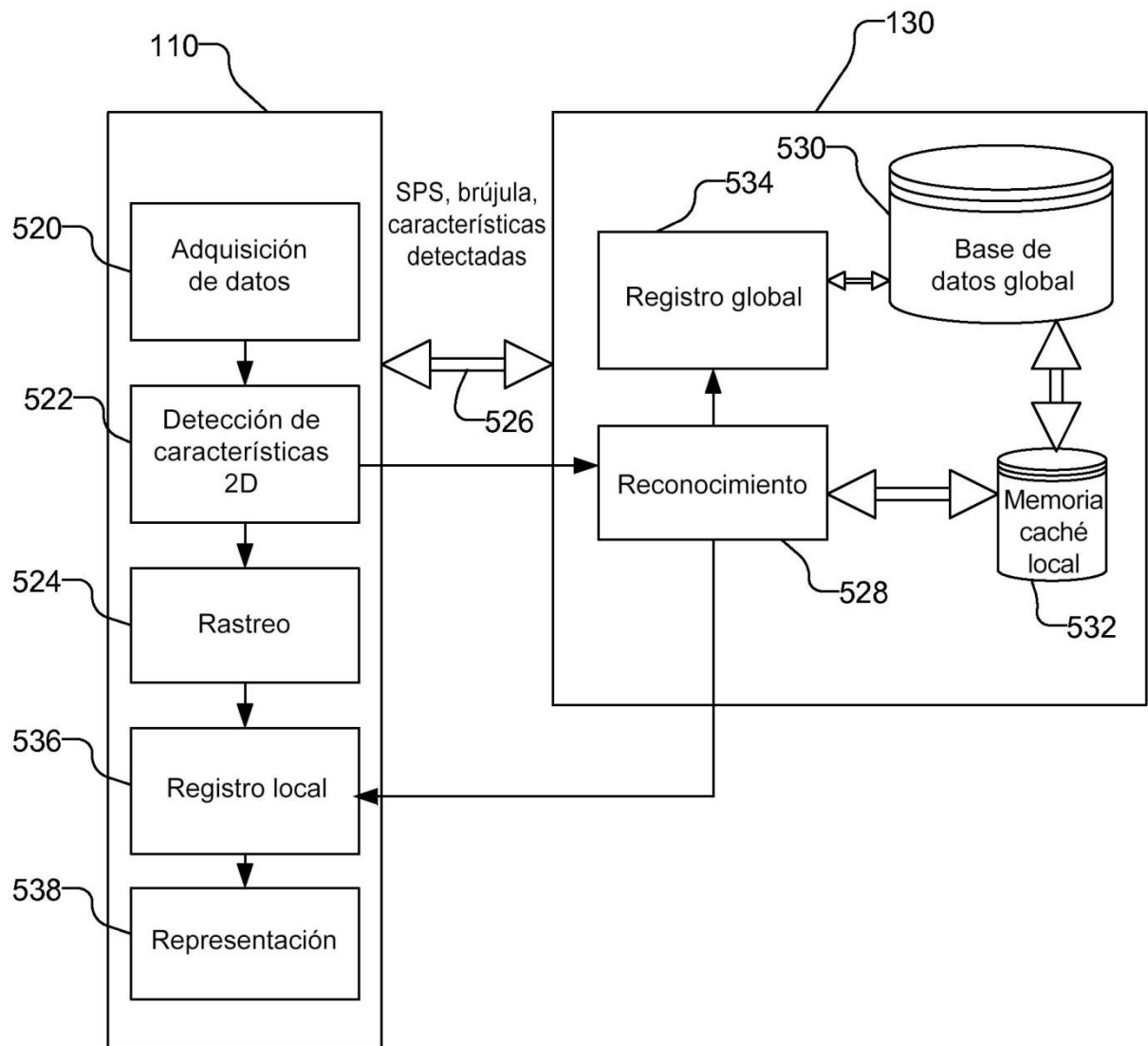


Fig. 10

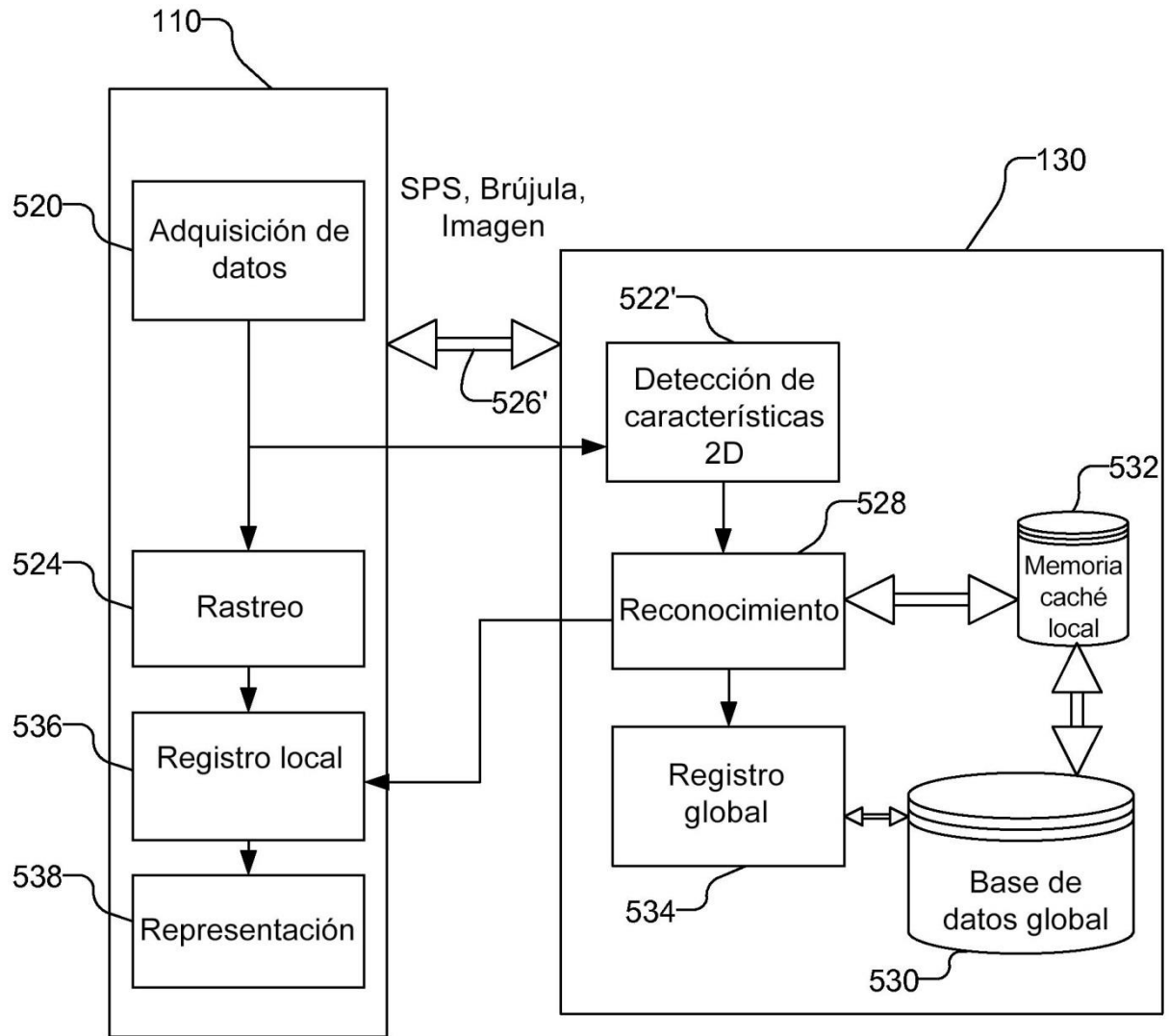


Fig. 11

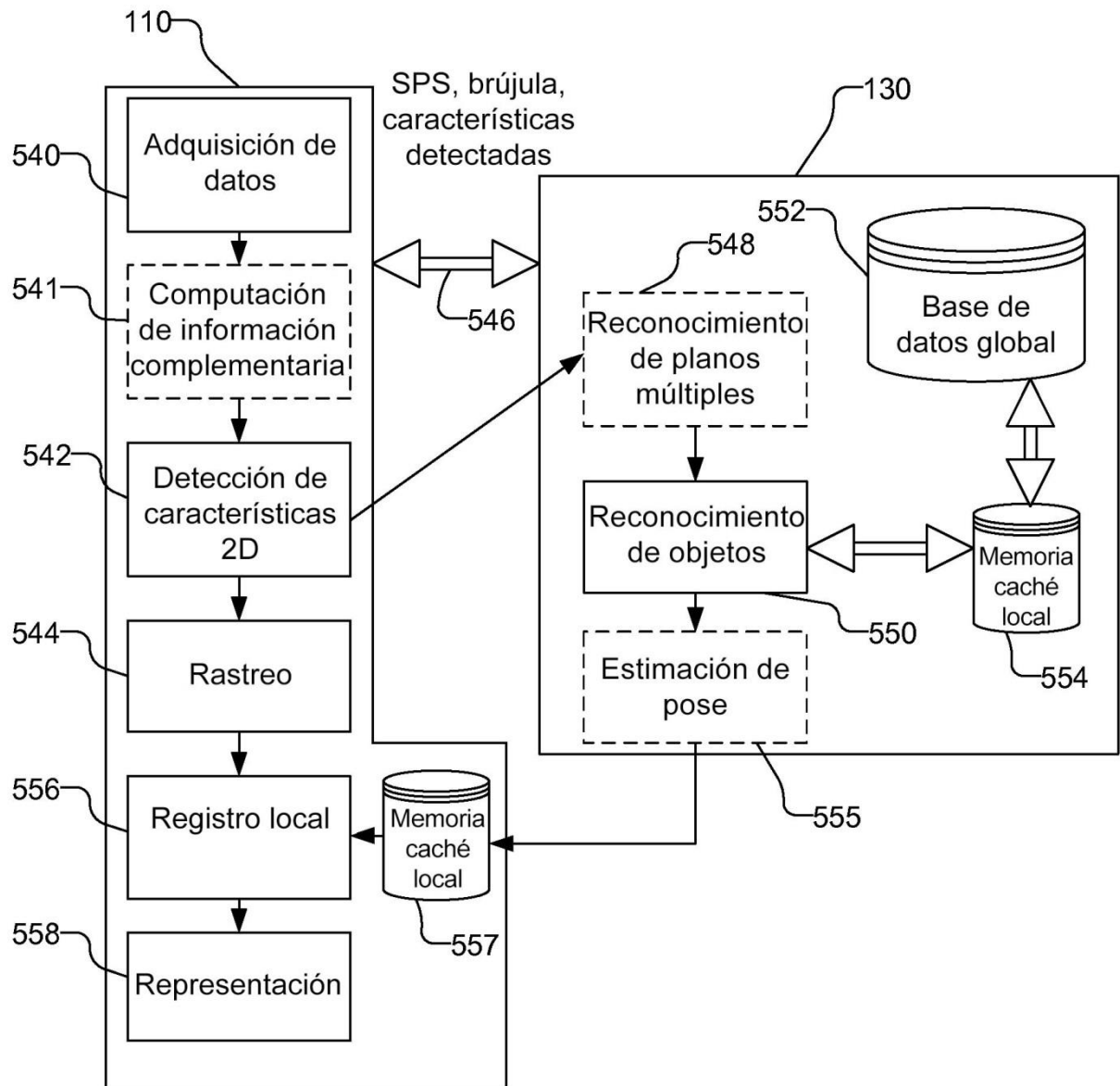


Fig. 12

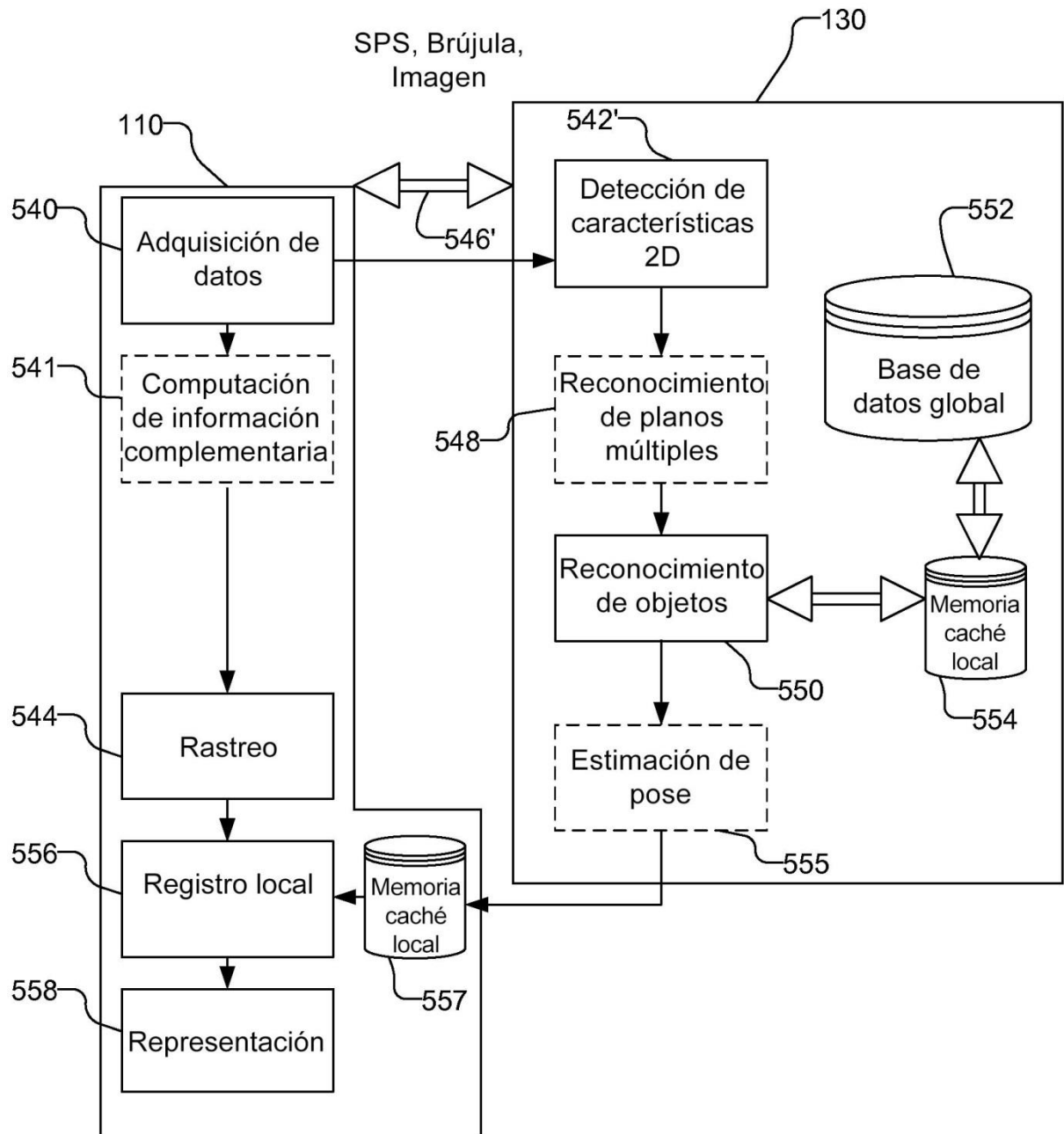


Fig. 13

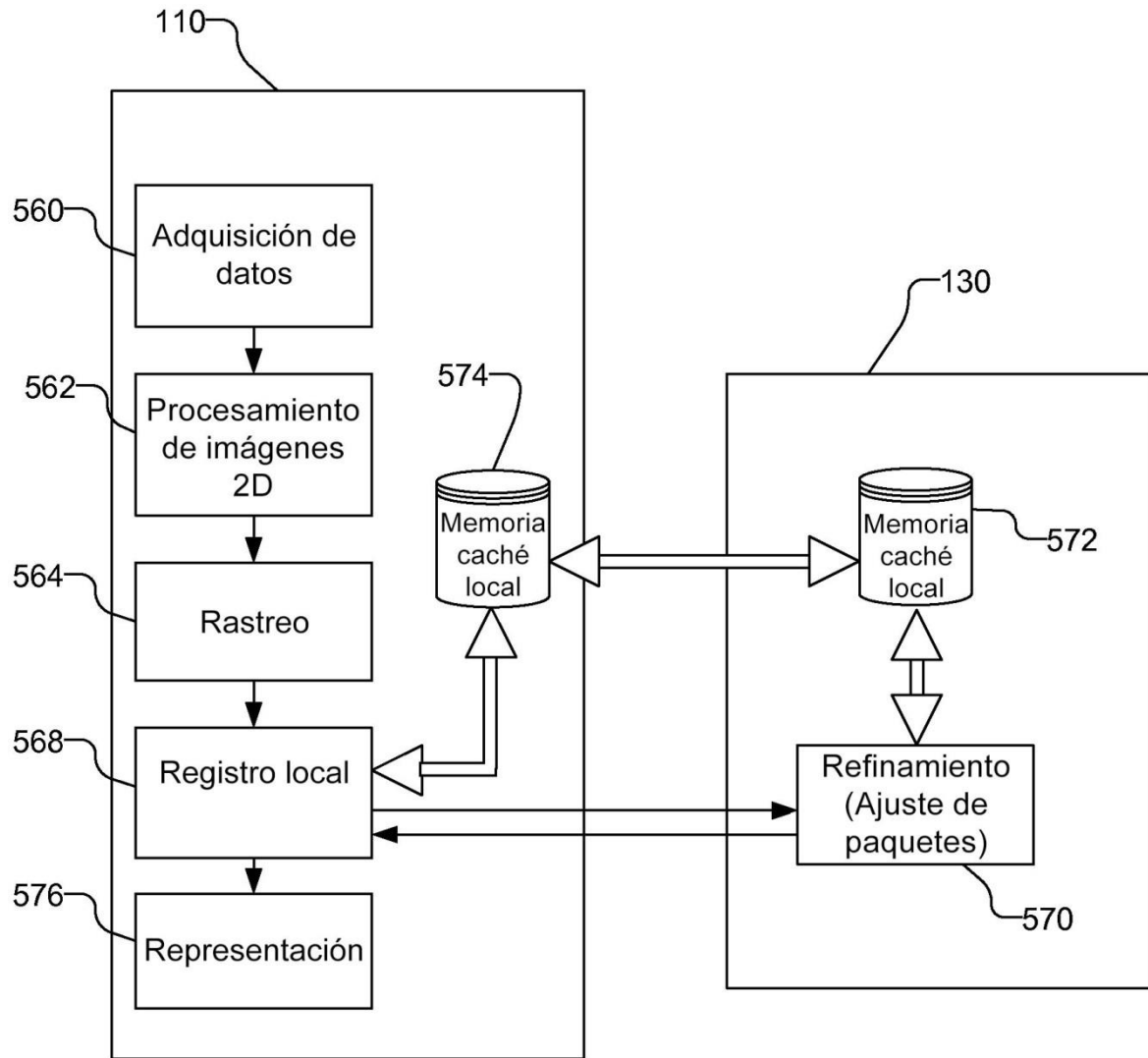


Fig. 14

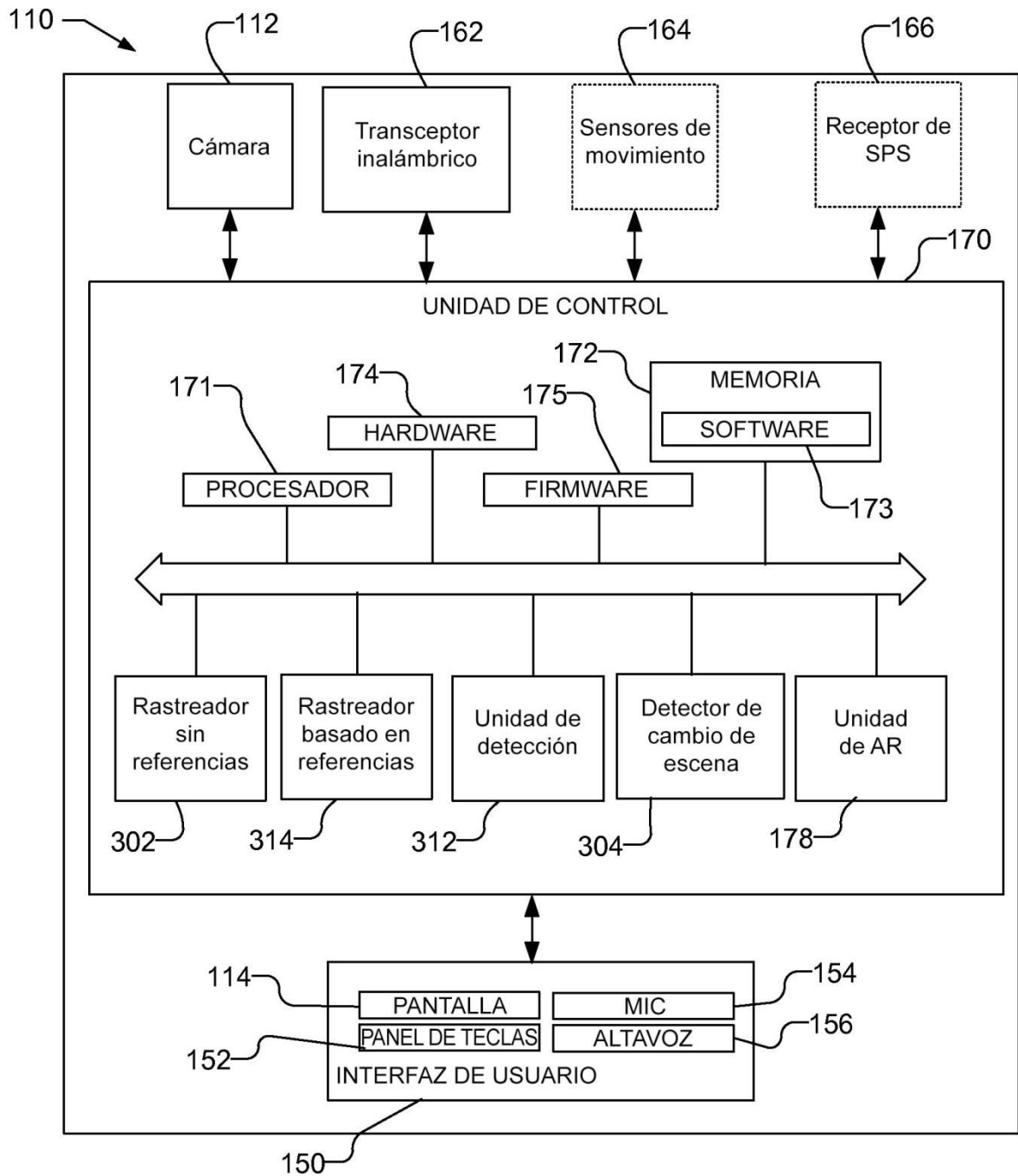


Fig. 15