

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 957 257**

51 Int. Cl.:

G01S 5/16	(2006.01)
G06T 7/20	(2007.01)
G06T 11/00	(2006.01)
H04N 1/00	(2006.01)
H04N 1/32	(2006.01)
H04N 5/232	(2006.01)
H04W 4/02	(2008.01)
H04W 64/00	(2009.01)
H04W 88/02	(2009.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **06.10.2011 PCT/IL2011/000787**
- 87 Fecha y número de publicación internacional: **19.04.2012 WO12049674**
- 96 Fecha de presentación y número de la solicitud europea: **06.10.2011 E 11832222 (1)**
- 97 Fecha y número de publicación de la concesión europea: **13.04.2022 EP 2625847**

54 Título: **Realidad aumentada registrada en tiempo real basada en la red para dispositivos móviles**

30 Prioridad:

10.10.2010 IL 20860010

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
16.01.2024

73 Titular/es:

**RAFAEL ADVANCED DEFENSE SYSTEMS LTD.
(100.0%)
P.O. Box 2250
31021 Haifa, IL**

72 Inventor/es:

**BERKOVICH, EREZ;
LITWITZ, ELON y
ADLER, AVIGDOR**

74 Agente/Representante:

DÍAZ DE BUSTAMANTE TERMINEL, Isidro

ES 2 957 257 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

Realidad aumentada registrada en tiempo real basada en la red para dispositivos móviles

5 Campo y antecedentes de la invención

La presente invención se refiere a la realidad aumentada y, en particular, se refiere a un sistema y método que proporciona la realidad aumentada en tiempo real basada en la red para dispositivos móviles.

10 La realidad aumentada (AR) se refiere a una forma de presentar información complementaria o contenido gráfico superpuesto y visualmente combinado con una vista del mundo real, ya sea que se vea directamente o como imágenes en una pantalla. En el caso de una escena vista directamente, la información complementaria generalmente se muestra a través de un arreglo de pantallas transparentes. En el caso de imágenes vistas en una pantalla, el contenido complementario generalmente se combina en la misma pantalla.

15 Una cuestión fundamental en la implementación de cualquier solución AR es la resolución del registro logrado entre la escena del mundo real y la información complementaria. Para la AR de baja resolución, por ejemplo, donde es suficiente proporcionar una etiqueta de texto simple desplegada en la dirección general de las características que ocupan una gran parte del campo de visión, es posible implementar un sistema AR rudimentario mediante el uso de
 20 GPS junto con un sistema de navegación inercial (INS) u otro conjunto básico de sensores de posición. Un ejemplo de tal implementación se describe en "Shared Database of Annotation Information for Wearable Augmented Reality System", Koji Makita y otros, Escuela de Postgrado de Ciencia de la Información, Instituto Nara de Ciencia y Tecnología, 8916-5 Takayama, Ikoma, Nara, 630--0192 Japón, págs. 464-471 (XP-002408355). Una gama de aplicaciones comerciales ha estado disponible recientemente para varios dispositivos de teléfonos inteligentes que
 25 proporciona funcionalidad AR en base a sensores de posición incorporados. El rendimiento de tales aplicaciones se limita por la precisión de los sensores de posición y el nivel de ruido en sus lecturas.

En aplicaciones AR más exigentes, puede ser necesario proporcionar un registro más preciso de la información
 30 complementaria con la vista del mundo real que el que puede lograrse mediante el uso de sensores de posición de bajo costo, por ejemplo, que permite que el sistema distinga entre dos objetos que parecen pequeños y juntos en una imagen muestreada. Para lograr el registro de alta resolución de la información complementaria al mundo real, se han desarrollado técnicas en base al procesamiento de imágenes a través de las cuales las imágenes muestreadas se registran en un sistema de información geográfica (SIG) que incluye datos de ortofotos. Pueden encontrarse ejemplos de esta tecnología en R. P. Wildes y otros, "Video georegistration: algorithm and quantitative evaluation", ICCV 2001,
 35 vol. 2, págs. 343-350; y en R. W. Cannata y otros, "Autonomous video registration using sensor model parameter adjustments", Taller IEEE sobre registro de video (con ICCV 2001).

Las aplicaciones AR basadas en el procesamiento de imágenes, tal como las descritas en estos artículos, ofrecen
 40 potencialmente un registro preciso píxel por píxel a la base de datos geográfica y, por lo tanto, proporcionan un posicionamiento preciso de la información AR complementaria suficiente para identificar todas las características visibles. Sin embargo, este enfoque requiere acceso a una gran base de datos de imágenes de referencia y fuertes capacidades de procesamiento. En términos prácticos, los requisitos de hardware para implementar la AR basada en el procesamiento de imágenes han sido hasta ahora generalmente prohibitivos para los dispositivos electrónicos de
 45 consumidores domésticos y otros dispositivos portátiles. Además, el procesamiento remoto de imágenes típicamente no es una opción efectiva para los dispositivos móviles en red a través de una red de área amplia (WAN) debido al retraso significativo que se introduce al cargar imágenes a través del ancho de banda limitado de la WAN.

El documento US 2010/208033A1 describe un método para acceder a datos ubicados geoméricamente que
 50 representan uno o más elementos virtuales con respecto a un sistema de coordenadas tridimensional; generar un mapa tridimensional basado, al menos en parte, en datos de imágenes reales de un espacio tridimensional adquiridos por una cámara; representar en una pantalla física una escena de realidad mixta que incluye el uno o más elementos virtuales en posiciones tridimensionales respectivas en una imagen real del espacio tridimensional adquirido por la cámara; y volver a representar en la pantalla física la escena de realidad mixta tras un cambio en el campo de visión de la cámara.

55 Por lo tanto, existe la necesidad de un sistema y método que proporcione realidad aumentada en tiempo real basada en la red para dispositivos móviles que permitiría que la AR basada en procesamiento de imágenes se use de manera efectiva dentro de las limitaciones de ancho de banda de procesamiento y comunicación de dispositivos portátiles en red.

60 Resumen de la invención

La presente invención proporciona un método de operar un dispositivo móvil como se define en la reivindicación 1,
 65 para un dispositivo móvil como se define en la reivindicación 10 ' y para un método de operar un servidor como se define en la reivindicación 12. Las características ventajosas adicionales de la presente invención se exponen en las reivindicaciones dependientes.

Breve descripción de los dibujos

La invención se describe en la presente descripción, a manera de ejemplo solamente, con referencia a los dibujos acompañantes, en donde:

- 5 La Figura 1 es un diagrama de bloques esquemático de un sistema de acuerdo con una modalidad de la presente invención para proporcionar realidad aumentada en tiempo real basada en la red para un dispositivo móvil;
- La Figura 2 es un diagrama de flujo esquemático que ilustra la operación del sistema de la Figura 1, de acuerdo con una modalidad de la presente invención;
- 10 La Figura 3 es un diagrama de bloques esquemático del lado del servidor del sistema de la Figura 1, de acuerdo con una modalidad de la presente invención;
- La Figura 4 es una ilustración esquemática de una pantalla de realidad aumentada de acuerdo con una modalidad de la presente invención;
- 15 La Figura 5 es un diagrama de bloques esquemático de una variante de implementación del dispositivo móvil de la Figura 1; y
- La Figura 6 es una vista isométrica esquemática de una implementación del dispositivo de la Figura 5.

Descripción de las modalidades preferidas

20 La presente invención es un sistema y método que proporciona realidad aumentada en tiempo real basada en la red para dispositivos móviles.

Los principios y la operación de los métodos y sistemas de acuerdo con la presente invención pueden entenderse mejor con referencia a los dibujos y la descripción acompañante.

25 Con referencia ahora a los dibujos, la Figura 1 muestra una descripción general de un sistema para proporcionar realidad aumentada en tiempo real basada en la red para dispositivos móviles de acuerdo con una modalidad de la presente invención, mientras que la Figura 2 ilustra un modo de operación del sistema, correspondiente a un método de acuerdo con una modalidad de la presente invención.

30 Volviendo primero a la Figura 1, se muestra un dispositivo móvil, generalmente designado como 10, conectado a través de una red de área amplia (WAN) 12 a un servidor remoto 14, que de esta manera forma un sistema adecuado para implementar una modalidad de la presente invención para proporcionar una visualización de información complementaria alineada con una vista de una escena.

35 En una modalidad como se muestra, el dispositivo móvil 10 incluye una cámara 16, una pantalla 18 y al menos un sensor de posición 20, todos conectados en la comunicación de datos a un sistema de procesamiento 22 (que incluye al menos un procesador 24 y almacenamiento de datos 26). También conectado al sistema de procesamiento 22 está un módulo de comunicaciones inalámbricas 28 para la comunicación en red a través de la WAN 12 al servidor remoto 14. Una interfaz del usuario 30 típicamente se proporciona para controlar las funciones del dispositivo.

40 El dispositivo móvil 10 como se describe puede implementarse mediante el uso de una amplia gama de tipos de dispositivos móviles. El dispositivo móvil 10 puede ser un dispositivo de realidad aumentada dedicado, o puede implementarse en base a varios dispositivos electrónicos portátiles convencionales con la modificación adecuada o la adición de componentes complementarios donde los componentes necesarios aún no están presentes. Los ejemplos de dispositivos adecuados incluyen, pero no se limitan a:

- un teléfono celular o "teléfono inteligente";
- 50 un ordenador portátil;
- un dispositivo de navegación;
- un asistente digital personal;
- una cámara digital;
- un auricular de realidad aumentada;
- monoculares y binoculares;
- 55 un telescopio;
- gafas de visión nocturna;
- pantallas de visualización frontal (HUD);
- pantallas montadas en el casco o en la cabeza (HMD); y
- 60 otros dispositivos que emplean pantallas transparentes o traslúcidas, que pueden implementarse, por ejemplo, mediante el uso de tecnología OLED.

La cámara 16 puede ser cualquier tipo de sensor de imagen que proporcione una imagen de la escena vista por el dispositivo. Los ejemplos incluyen, pero no se limitan a, cámaras de color visibles (VIS) (tal como cámaras basadas en CCD o CMOS), cámaras de infrarrojo cercano (NIR), cámaras térmicas y cámaras multiespectrales o hiperespectrales. Además, en una implementación preferida de la presente invención en la que el registro de las imágenes se realiza con relación a un modelo tridimensional del entorno, también pueden usarse cámaras de

detección de profundidad tales como LADAR. Opcionalmente, puede proporcionarse más de una cámara, tal como por ejemplo una combinación de una cámara a color VIS con una cámara térmica para proporcionar una operación mejorada tanto en operaciones de día como de noche.

5 La pantalla 18 puede ser cualquier pantalla adecuada que pueda usarse para mostrar la información complementaria de la pantalla de realidad aumentada alineada con la escena. Las implementaciones del dispositivo móvil 10 pueden subdividirse ampliamente en dos categorías, ambas representadas esquemáticamente en la Figura 1: dispositivos en los que el usuario ve la escena como una imagen mostrada en la pantalla 18; y dispositivos en los que el usuario ve la escena directamente, por ejemplo, a través de una óptica adecuada. En este último caso, la pantalla 18 se implementa como una pantalla transparente. El término "pantalla transparente" se usa en este contexto para referirse genéricamente a cualquier arreglo de pantallas que permita la visualización directa simultánea de una escena real, que incluye pero sin limitarse a: un panel de pantalla transparente; un arreglo en el que la óptica de división de haz combina dos vistas y sistemas de proyección retiniana que no oscurecen. Donde se proporciona una pantalla transparente, el despliegue de la pantalla se elige para permitir al usuario ver la información complementaria mostrada mientras se enfoca simultáneamente en la escena real más distante, como se conoce bien en la técnica de las pantallas transparentes.

Los sensores de posición 20 son preferentemente una combinación de sensores que proporcionan suficiente información para generar una estimación inicial de la pose de la cámara. El término "sensor de posición" se usa aquí para referirse a cualquier sensor que detecta o deriva la ubicación o la orientación de la cámara o el dispositivo móvil en su conjunto. Los ejemplos de sensores adecuados pueden incluir algunos o todos de los siguientes: un receptor GPS, acelerómetros, sensores de inclinación y una brújula digital. Aunque el tiempo de procesamiento en el servidor se reduce de manera óptima al proporcionar una estimación relativamente buena de la ubicación y la orientación de la cámara, el sistema también puede operar con información parcial, tal como la ubicación GPS sin datos de orientación, o con datos de orientación parcial más la ubicación aproximada de la interpolación de una red celular. En ciertos casos, los sensores de posición 20 pueden omitirse completamente, reemplazándose por datos proporcionados externamente desde otra unidad móvil cercana, desde una red celular o mediante la entrada manual de una ubicación o vecindario aproximado.

30 El módulo de comunicaciones inalámbricas 28 puede ser cualquier interfaz de comunicaciones inalámbricas que proporcione una trayectoria de comunicaciones a la WAN en la que se ubican los servidores. La interfaz más disponible en muchos entornos hoy en día es un sistema de comunicación celular que proporciona comunicación de datos vinculados a Internet. Claramente, para aplicaciones específicas, puede usarse una red de comunicaciones dedicada en base a cualquier tecnología inalámbrica adecuada. Aunque el último tramo de las comunicaciones al dispositivo móvil es preferentemente inalámbrico, todas las demás partes de la cadena de comunicaciones pueden implementarse mediante el uso de cualquier combinación deseada de conexiones por cable e inalámbricas.

El sistema de procesamiento 22 del dispositivo móvil 10 se configura, típicamente mediante un software adecuado cargado en la memoria no volátil y ejecutado por el sistema de procesamiento, lo que representa el dispositivo móvil efectivo para realizar el procesamiento del lado del cliente, correspondiente a un método de acuerdo con una modalidad de la presente invención, como se describirá más abajo con referencia a la Figura 2. Quedará claro que las implementaciones alternativas que emplean hardware dedicado, o combinaciones de hardware y software conocidas como microprograma, también caen dentro del alcance de la presente invención. También cae dentro del alcance de la presente invención el código almacenado en un medio tangible de almacenamiento de datos legible por ordenador y ejecutable por el procesador de un dispositivo móvil para hacer que el dispositivo móvil realice las funciones enumeradas.

En una modalidad como se muestra, el servidor remoto 14 incluye un sistema de procesamiento 32 y una base de datos geográfica 34, que se detallará además más abajo con referencia a la Figura 3.

Volviendo ahora a la Figura 2, esto ilustra un modo de operación del sistema, correspondiente a un método 40 de acuerdo con una modalidad de la presente invención. Específicamente, el lado izquierdo de la Figura muestra las etapas realizadas por el dispositivo móvil, que actúa como el "cliente", mientras que el lado derecho muestra las etapas realizadas por el servidor.

Se debe señalar, aunque aquí se muestra por simplicidad de presentación como un único proceso de cliente que interactúa con el servidor, la arquitectura del sistema es particularmente adecuada para proporcionar un servicio paralelo desde un único servidor o conjunto de servidores a múltiples clientes. En particular, debido a que el procesamiento de registro con gran cantidad de cálculos se realiza solo de manera intermitente, según lo requiera cada cliente, y de una manera que es insensible a retrasos menores, los recursos de procesamiento del servidor pueden reasignarse dinámicamente para atender a múltiples clientes en paralelo.

Con referencia ahora al proceso del lado del cliente de la Figura 2, se obtiene una primera imagen de la cámara en la etapa 42, y se carga al servidor remoto 14 junto con los datos correspondientes de los sensores de posición 20 (etapa 44) o alguna fuente externa, como se describió anteriormente. La imagen cargada será procesada por el servidor (que se describirá por separado más abajo) para obtener el registro en una base de datos geográfica y proporcionar datos

para la alineación correcta de la información AR. Sin embargo, el tiempo combinado de cargar la imagen y procesarla en el servidor introduciría posiblemente grandes errores debido al movimiento de la cámara durante el lapso de tiempo.

5 Es una característica particularmente preferida de un aspecto de la presente invención que el sistema de procesamiento de la unidad móvil realice el procesamiento de imágenes para rastrear el movimiento de la imagen entre la primera imagen y las imágenes subsecuentes obtenidas de la cámara y, por lo tanto, determine un mapeo (etapa 46) que puede usarse para compensar el retraso en obtener los resultados del procesamiento de registro del servidor.

10 La frase "seguimiento del movimiento de la imagen" se usa en este contexto para referirse generalmente a una gama de algoritmos que permiten la asociación de píxeles a píxeles entre imágenes separadas en una secuencia de imágenes. Por medio de ejemplos específicos no limitativos, el seguimiento del movimiento de la imagen realizado por un dispositivo móvil puede implementarse de acuerdo con cualquiera de al menos tres técnicas, que dependen principalmente de los escenarios de uso probables y los recursos de procesamiento disponibles del dispositivo, como sigue.

15 En una implementación más básica, particularmente adecuada para situaciones donde el movimiento de la cámara es principalmente panorámico sin movimiento del ego significativo, el movimiento de la imagen puede modelarse mediante una transformación espacial global, tal como una transformación afín. El término "global" aquí se refiere a una transformación que se aplica uniformemente a toda la imagen como una unidad. El seguimiento para implementar dichas transformaciones puede realizarse al rastrear características relativamente escasas entre imágenes sucesivas, como se conoce bien en la técnica.

20 Una segunda opción de implementación emplea flujo óptico. Esta opción es computacionalmente más pesada, pero proporciona una solución más robusta para los casos donde hay algún movimiento de traslación de la cámara. El procesamiento de flujo óptico para pares o secuencias de imágenes se conoce bien en la técnica y no se describirá aquí en detalle.

25 Una tercera opción, relativamente pesada en su carga computacional, pero aún más robusta para los casos de una plataforma móvil, es emplear el procesamiento monocular de localización y mapeo simultáneo (SLAM). Puede encontrarse una discusión de algoritmos SLAM adecuados para usar con una sola cámara en la técnica y, por ejemplo, en: A. Davison, "Real time simultaneous localisation and mapping with a single camera", ICCV, Niza, Francia, julio de 2003.

30 De acuerdo con cualquiera de estas opciones, el proceso de seguimiento da como resultado una transformación o mapeo que permite asociar ubicaciones de píxeles en una imagen actual con ubicaciones correspondientes en la primera imagen y/o en una imagen intermedia. Se debe señalar en este contexto que el término "imagen actual" se usa para referirse a una imagen muestreada recientemente por la cámara para formar la base para la visualización sustancialmente en tiempo o real de información complementaria de AR. Se debe señalar que la "imagen actual" también puede estar sujeta a retrasos de procesamiento de varios fotogramas, pero preferentemente todavía está dentro del ámbito de lo que se considera aceptable como visualización en tiempo real, típicamente con un retraso significativamente inferior a un segundo.

35 En el caso de que el campo de visión haya cambiado tan drásticamente que los resultados del registro de la imagen cargada previamente ya no sean útiles, el dispositivo móvil puede enviar una notificación de cancelación al servidor y reiniciar el procesamiento desde la etapa 42.

40 Opcionalmente, en la etapa 48, los parámetros resultantes del seguimiento anterior desde el muestreo de la primera imagen pueden cargarse al servidor, que de esta manera proporciona una actualización al servidor del campo de visión actual con relación al campo de visión original de la primera imagen. Esto permite que el servidor seleccione información complementaria de AR que aún es relevante para el campo de visión actual si el campo de visión ha cambiado desde que se muestreó la primera imagen. En muchas aplicaciones, la etapa 48 es innecesaria.

45 En la etapa 50, el dispositivo móvil 10 recibe a través de la red datos indicativos de una ubicación de píxel para la visualización de información complementaria dentro de la primera imagen, o una imagen de referencia intermedia. El mapeo derivado del proceso de seguimiento se emplea luego para determinar una ubicación de píxel correspondiente para la visualización de la información complementaria dentro de la imagen actual (etapa 52). El término "ubicación de píxel" se usa en la presente descripción para referirse a una ubicación dentro de una imagen definida con relación a una cuadrícula de píxeles. La ubicación puede representarse como una dirección de píxel específica, como un grupo o región de píxeles, o como alguna otra relación geométrica con una o más ubicaciones de píxeles. A continuación, la información complementaria se muestra en la pantalla correctamente alineada con la vista de la escena (etapa 54), ya sea al mostrar la información suplementaria superpuesta a una imagen de la vista de la escena en una pantalla no transparente, o al mostrar la información complementaria en una pantalla transparente alineada con una vista directa de la escena.

65

Opcionalmente, en la etapa 50, el servidor puede proporcionar al dispositivo móvil con la transformación de registro calculada, que proporciona una ubicación geográfica precisa para cada píxel de la imagen. Esto puede usarse para habilitar una funcionalidad adicional en el dispositivo móvil, por ejemplo, permitir la identificación de la ubicación geográfica de cualquier punto en el campo de visión para cargarlo al servidor o transferirlo a otro usuario, opcionalmente con anotaciones adicionales. Alternativamente, puede proporcionarse una funcionalidad similar al cargar una designación de ubicación de píxel, con la ubicación geográfica correspondiente derivada por el servidor.

Aunque el proceso antes mencionado puede realizarse opcionalmente con una serie de imágenes fijas, la funcionalidad preferida es mostrar la información complementaria de AR de una manera que se actualiza continuamente en base a imágenes de video continuas muestreadas por la cámara. En este caso, el procesamiento de seguimiento local preferentemente continúa mientras la pantalla AR opera para generar mapeos actualizados para nuevas imágenes actuales a medida que estén disponibles, y se actualiza la ubicación de píxel para mostrar la información complementaria dentro de la imagen actual de acuerdo con el mapeo actualizado (etapa 56). Esto mantiene efectivamente el contenido AR correctamente posicionado dentro de la imagen de video móvil o cambia el campo de visión de un dispositivo de visualización directa.

El procesamiento de seguimiento realizado en el dispositivo móvil proporciona una corrección altamente efectiva para mantener el posicionamiento correcto del contenido AR con relación a la escena vista, pero no proporciona una solución para los cambios completos del campo de visión. En consecuencia, el proceso de cargar una imagen al servidor y recibir nuevos datos para una visualización precisa del contenido AR debe repetirse periódicamente. En algunas aplicaciones, puede ser apropiado simplemente repetir el proceso de carga y registro a intervalos fijos. Sin embargo, para mejorar el uso eficiente de los recursos de procesamiento, el dispositivo móvil evalúa preferentemente un indicador de fiabilidad del seguimiento para evaluar cuándo ha disminuido la fiabilidad del seguimiento (etapa 58), y luego inicia un nuevo proceso de carga y registro (etapas 42-54) según sea necesario. El indicador de fiabilidad del seguimiento puede basarse en uno o más parámetros tales como: tiempo transcurrido desde el muestreo de la primera imagen; grado de superposición entre la imagen actual y la primera imagen; y/o el número de características rastreadas en común entre la imagen actual y la primera imagen. Cuando el indicador de fiabilidad del seguimiento disminuye más abajo de un umbral determinado, que puede predefinirse o variarse dinámicamente en base a los parámetros de operación, la operación vuelve a la etapa 42, preferentemente mientras continúa mostrando el contenido AR actualizado en tiempo real hasta que los datos en base al registro actualizado estén disponibles.

Volviendo ahora al servidor 14, una implementación del servidor se representa esquemáticamente en la Figura 3. Como se describió anteriormente con referencia a la Figura 1, uno o más procesadores 32 se conectan a la WAN 12, que proporcionan comunicación a los dispositivos móviles. En una modalidad preferida pero no limitativa, los procesadores 32 se implementan como recursos de procesamiento asignados dinámicamente desde múltiples ordenadores distribuidos en una red de acuerdo con un modelo de "computación en la nube". Tales servicios se conocen bien en la técnica y están disponibles comercialmente a partir de proveedores de servicios tales como GOOGLE® y AMAZON®. De acuerdo con otra modalidad preferida pero no limitativa, los procesadores 32 son parte de un sistema de servidor dedicado, que puede ser un sistema independiente o puede integrarse con una base de datos geográfica 34.

El servicio AR proporcionado a cada dispositivo móvil se personaliza con la máxima preferencia a los intereses y necesidades de cada usuario y/o mediante la selección de varios modos de operación. Con este fin, el servidor almacena preferentemente una gama de configuraciones en algún dispositivo de almacenamiento de datos adecuado 33 que mantienen y actualizan los perfiles de usuario, las configuraciones individuales y, preferentemente, también identifican grupos de usuarios como pares a quienes se les puede proporcionar información mutua sujeta a los requisitos de privacidad del sistema y del usuario. El sistema también puede incluir módulos para verificar la autorización para usar el sistema y/o recibir pagos, todo como es estándar para los proveedores de servicios en línea, y no se detallará aquí.

La base de datos geográfica 34 en sí misma puede ser un sistema de información geográfica estándar (SIG), o una base de datos dedicada, que almacena capas de información que típicamente todas se anclan a coordenadas de cuadrícula geográfica 80. Las capas típicas de información incluyen, pero no se limitan a, una o más de las siguientes: un mapa digital del terreno (DTM) 82 identifica la topología geográfica y las imágenes correspondientes 84, típicamente de fuentes aéreas o satelitales, que se vincula a la cuadrícula y se ajusta para proporcionar una vista ortogonal en todas las ubicaciones ("ortofoto"); modelos 3D 86 de edificios y/u otras estructuras, preferentemente con la correspondiente información de textura. Además de estos, la base de datos preferentemente también almacena varias capas de otra información estática 88 y opcionalmente también varias informaciones transitorias 90. En este contexto, "información estática" se usa para referirse a cualquier información que típicamente no varía durante el transcurso de una sesión de realidad aumentada, tal como información sobre la identidad, propiedad, historia, propiedades o funciones de objetos o ubicaciones en el modelo 3D. Por el contrario, "información transitoria" se refiere aquí a información, instrucciones o tareas que pueden cambiar en períodos de tiempo relativamente cortos, o que se asocian con una persona, grupo u objeto que puede moverse. Los ejemplos de información transitoria que pueden ser relevantes para ciertas modalidades de la presente invención incluyen, pero no se limitan a, la ubicación y la identidad de (los usuarios de) dispositivos móviles, información de seguimiento sobre vehículos (tal como el transporte público),

información de congestión de tráfico e información meteorológica, y una lista o descripción de las tareas a lograr colectivamente por un grupo.

5 Entre paréntesis, aunque típicamente se implementa en base a coordenadas geográficas, se debe señalar que también pueden usarse implementaciones en base a un modelo que no se ancla geográficamente, y pueden ser suficientes particularmente en entornos cerrados bien definidos, tal como un complejo interior (por ejemplo, dentro de un museo o similar).

10 Volviendo ahora al Lado del Servidor de la Figura 2, el servidor 14 recibe desde el dispositivo móvil 10 la imagen cargada obtenida de la cámara 16 junto con datos indicativos de una posición correspondiente del dispositivo móvil cuando se obtuvo la imagen (etapa 60). Luego, los datos de posición se emplean para generar una estimación de un campo de visión (FOV) de la cámara del dispositivo móvil correspondiente a la primera imagen (etapa 62), que se usa como una base para recuperar de la base de datos³⁴ al menos parte de un modelo tridimensional (3D) anclado geográficamente correspondiente a la estimación del campo de visión (etapa 64).

15 Donde los datos de posición son suficientes, por ejemplo, que incluyen los parámetros estimados para la ubicación, la dirección y la orientación de la cámara, la estimación del FOV puede ser relativamente específica, lo que permite una recuperación altamente selectiva de la base de datos 34 de datos que probablemente se incluirán dentro del FOV. En otros casos, donde solo están disponible datos de posición parciales o muy aproximados, puede necesitarse recuperar una región extendida del modelo para garantizar la disponibilidad de datos suficientes para lograr el registro. Claramente, cuando el registro de imágenes se repite para imágenes subsecuentes de la misma localidad (ya sea desde el mismo dispositivo móvil u otro dispositivo), la etapa de recuperación solo se realiza según sea necesario, y algunos o todos los datos requeridos ya pueden haberse almacenado en caché localmente y estar disponibles para usar.

25 El modelo 3D puede ser un modelo de resolución relativamente baja, tal como un mapa digital del terreno (DTM) de topología geográfica, comúnmente usado como base para el procesamiento en base a imágenes de ortofotos, o puede ser un modelo de mayor resolución que incluye características de edificios específicos y otros puntos de referencia. Opcionalmente, pueden usarse datos de más de un modelo o tipo de modelo. El modelo se complementa preferentemente por imágenes asociadas o datos de textura que también se usan en el proceso de registro subsecuente. Pueden derivarse modelos adecuados mediante una gama de técnicas conocidas que incluyen, pero no se limitan a, procesamiento de imágenes de pares estereoscópicos, procesamiento de estructura a partir de movimiento (SFM) y ajuste de haz para reconstrucción 3D. El modelo puede complementarse o actualizarse opcionalmente sobre la base de las imágenes cargadas desde algunos o todos los dispositivos móviles mediante el uso del sistema.

30 Luego, en la etapa 66, el servidor coprocesa la imagen cargada con la porción potencialmente relevante del modelo tridimensional para derivar un mapeo de registro para mapear píxeles de la primera imagen al modelo tridimensional. El procesamiento para determinar un mapeo de registro entre imágenes 2D y un modelo 3D se conoce bien en la técnica y no se describirá aquí en detalle. Pueden encontrarse ejemplos de técnicas adecuadas no limitativas para determinar tal mapeo en: (A) R.P. Wildes, D.J. Hirvonen, C.S. Hsu, R. Kumar, W.B. Lehman, B. Matei & W.-Y. Zhao (Samoff), "Video georegistration: algorithm and quantitative evaluation", ICCV 2001, vol. 2, págs. 343-350; (B) R.W. Cannata, M. Shah, S.G. Blasik & J.A. Van Workum (Corporación Harris y Universidad de Florida Central), "Autonomous video registration using sensor model parameter adjustments", Taller IEEE sobre registro de video (con ICCV 2001); y (C) "Registration for Tactical Imagery: An Updated Taxonomy", R. Bruce Backman, capítulo 3.3, Informe técnico del Departamento de Defensa del Gobierno de Australia DSTO No. DSTO-TR-1855.

40 Como se mencionó anteriormente en el contexto de la etapa 48 del lado del cliente, para suministrar información complementaria relevante para el FOV actual del dispositivo móvil, en la etapa 68, de acuerdo con ciertas implementaciones de la presente invención, el servidor transmite una solicitud al dispositivo móvil para datos de seguimiento actualizados indicativos de un campo de visión actualizado de la cámara. Adicionalmente o alternativamente, el dispositivo móvil puede proporcionar actualizaciones al servidor, ya sea periódicamente o cuando ocurra un cambio significativo en el FOV. En implementaciones alternativas, se omite esta actualización y se supone que el FOV actual es generalmente similar al FOV de la imagen cargada original, preferentemente con algunos márgenes adicionales a su alrededor.

55 En la etapa 70 de la implementación como se ilustra aquí, el servidor recupera de una base de datos de referencia información complementaria relevante para el FOV actual. El tipo de información complementaria recuperada y el formato de la información pueden tomar muchas formas, en dependencia de la aplicación específica. Con la máxima preferencia, la información complementaria se ancla a un objeto, superficie o posición específicos en el modelo 3D, e incluye datos textuales y/o gráficos y/o un objeto virtual para mostrarse en una posición correspondiente (ubicación y orientación) en la pantalla AR.

60 La información complementaria puede recuperarse de muchas fuentes diferentes. De acuerdo con una opción preferida pero no limitativa, la información complementaria recuperada se recupera de capas 88 y 90 de base de datos 34, descrito anteriormente, o de una o más bases de datos dedicadas de información complementaria de AR.

Alternativamente o adicionalmente, puede obtenerse información complementaria al buscar una amplia gama de fuentes de datos geotiquetados en línea y/o sistemas en red que tienen datos disponibles sobre la presencia geográfica de los usuarios conectados. En ciertos casos, alguna o toda la información complementaria puede precargarse en el dispositivo móvil.

5 La información complementaria se recupera preferentemente de forma selectiva de acuerdo con uno o más de una serie de parámetros. En primer lugar, la información solo se recupera preferentemente si se cae dentro del FOV actual o lo suficientemente cerca del mismo lo que sería útil para mostrar en el caso de un movimiento moderado de la cámara que podría rastrearse por el dispositivo móvil.

10 En segundo lugar, incluso dentro del FOV, en algunas aplicaciones puede desearse suprimir, o permitir la supresión de datos seleccionada por el usuario, de acuerdo con la distancia de los objetos correspondientes desde el dispositivo móvil. Por ejemplo, en algunos casos, es conveniente suprimir el exceso de detalles en el campo lejano de la escena vista donde la resolución de la imagen es insuficiente para resolver visualmente la información. Por el contrario, en otras aplicaciones, puede ser ventajoso dar prioridad a las principales características o ubicaciones visibles en la distancia mientras se evitan las distracciones o el oscurecimiento de la información mediante la visualización de información sobre los objetos próximos.

15 En tercer lugar, el tipo de información seleccionada para mostrar es típicamente una función de la aplicación prevista y/o las opciones seleccionadas por cada usuario.

20 En la etapa 72, el servidor determina preferentemente los parámetros de cómo debería aparecer cada elemento de información complementaria en el FOV actual de la cámara. Principalmente, esto típicamente incluye una ubicación de píxel dentro de la imagen cargada para la visualización de elementos de información complementaria dentro de la imagen.

25 De acuerdo con ciertas modalidades preferidas, la información complementaria también se procesa para fusionarse intuitivamente en la escena real como se ve por el FOV actual del dispositivo móvil. Por tanto, por ejemplo, si la información complementaria es una etiqueta textual o un elemento gráfico asociado con una superficie del modelo 3D, la forma preferentemente se ajusta (deforma) para aparecer como si fuese una superficie en 3D que se ve en el ángulo de la superficie correspondiente en el FOV actual. De manera similar, si la información complementaria es un objeto virtual, se gira para proporcionar el punto de vista apropiado para fusionarse con la escena real de acuerdo con el FOV actual. Preferentemente, la forma también se escala para que aparezca con el tamaño correcto para el rango del objeto en la imagen, aunque la reducción de escala puede limitarse para preservar la legibilidad de las etiquetas en objetos distantes. El servidor puede realizar la manipulación de formas gráficas para generar una forma 2D lista para mostrar para descargar al dispositivo móvil. Por lo tanto, el procesamiento del lado del servidor de acuerdo con ciertas modalidades de la invención permite la visualización de la información complementaria en el dispositivo móvil con el tamaño y la orientación correctos dentro de la vista actual mientras se minimiza cualquier carga de procesamiento adicional en el lado del cliente. Alternativamente, el servidor puede simplemente determinar los parámetros de acuerdo con los cuales la información complementaria debe ajustarse para verse dentro del FOV actual, y la manipulación gráfica real puede realizarse por un procesador de gráficos del sistema de procesamiento 22 del dispositivo móvil.

30 Entonces, en la etapa 74, el servidor transmite al dispositivo móvil datos indicativos de la ubicación de píxel para la visualización de la información complementaria, preferentemente junto con la forma 2D o los parámetros para generar la forma correcta de un elemento gráfico para mostrar por el lado del cliente de acuerdo con las etapas 50-54, como se describió anteriormente. A continuación, el contenido gráfico puede mostrarse de modo que se ajusta sustancialmente a una orientación aparente del objeto al que se refiere la información complementaria.

35 Opcionalmente, si la información actualizada está disponible para el servidor que es relevante para el FOV actual subsecuente a la etapa 72, es decir, entre solicitudes de re-registro sucesivas, el servidor puede configurarse para indicarle al cliente la información actualizada, lo que permite la actualización en tiempo real de información, tal como el etiquetado de un objeto móvil que aparece en el FOV.

40 Las modalidades de la presente invención pueden usarse ventajosamente en una amplia gama de aplicaciones. Una pequeña y no limitada selección de ejemplos es la siguiente:

- 45 • Indicar al usuario la ubicación de un restaurante en el 3er piso de un edificio, e identificar la entrada correspondiente al edificio.
- 50 • Proporcionar un recorrido autoguiado de una ciudad, que incluye recuperar la historia y el relato narrativo relacionado con los edificios o monumentos vistos durante el recorrido.
- 55 • Proporcionar una ayuda de navegación en la que se superponen flechas u otros símbolos que indican una ruta deseada en la escena real.
- 60 • Proporcionar una herramienta de soporte para las fuerzas de rescate, que muestra la ubicación de la entrada a un edificio y la ubicación de las tuberías de agua y otra infraestructura.

- Mostrar una etiqueta que ubica a mi amigo, John, (que también tiene un dispositivo móvil y, por lo tanto, cuya posición se conoce por el servidor) en medio de una multitud de 50 000 personas en un partido de fútbol.

En cada una de estas aplicaciones, la funcionalidad y el rendimiento típicamente se mejoran en gran medida al emplear el registro de imagen a modelo para lograr una precisión de registro que puede acercarse a los límites de la resolución de imágenes disponible.

Las opciones de implementación adicionales incluyen, pero no se limitan a:

- 10 proporcionar contenido de audio hablado o de otro tipo para acompañar la pantalla actual, preferentemente asociado con un elemento gráfico actualmente visualizado que designa el objeto visto correspondiente;
- proporcionar contenido gráfico adicional, más detallado, o recorrer una secuencia de información complementaria diferente, o información sobre diferentes objetos dentro de un campo de visión, si el usuario permanece en una escena determinada por más de una cantidad de tiempo determinada;
- 15 proporcionar flechas u otros indicadores de dirección relacionados con puntos de interés que se encuentran fuera del campo de visión actual como un aviso para que el usuario gire hacia esos puntos de interés.

La Figura 4 ilustra esquemáticamente un ejemplo de una pantalla generada de acuerdo con una modalidad de la presente invención. En este caso, la vista de una calle se complementa con etiquetas numéricas 90 que indican los números de edificio de los edificios dentro del FOV. Se deberá señalar que, en la implementación preferida que se ilustra aquí, las etiquetas 90 se ajustan a la orientación de las superficies frontales de los edificios, de manera que la etiqueta para el edificio "32" parece aplicarse al frontal de un edificio en el lado derecho de la imagen que mira hacia la izquierda mientras que la etiqueta para el edificio "23" aparece aplicada al frontal de un edificio en el lado izquierdo de la imagen que mira hacia la derecha. De manera similar, las etiquetas se escalan de acuerdo con la distancia del objeto a la cámara, con la etiqueta para el edificio distante "8" mucho más pequeña que la del edificio "32" en primer plano.

Como se mencionó anteriormente, el dispositivo móvil preferentemente continúa realizando el seguimiento de movimiento entre fotogramas sucesivos de la imagen y ajusta la posición de las etiquetas 90 para mantener la alineación entre las etiquetas y los objetos correspondientes en la imagen. Cuando la calidad del seguimiento disminuye, el proceso de registro se repite y la información complementaria de AR se actualiza donde es necesario.

Una característica adicional de ciertas implementaciones de la invención ilustrada aquí es la capacidad de mostrar información relacionada con objetos actualmente ocultos a la vista. En esta ilustración, un árbol 94 oculta una serie de edificios en el lado izquierdo de la calle. Sin embargo, el procesamiento de registro de la imagen muestreada en el modelo 3D realizado por el servidor aún permite la recuperación de tal información, que opcionalmente puede mostrarse mediante el uso de etiquetas distintivas, en este caso mostradas como etiquetas de colores invertidos 92 (mediante el uso de negro sobre blanco en lugar de blanco sobre negro), para informar al usuario que se refieren a objetos actualmente ocultos.

Se debe señalar que el registro de las imágenes muestreadas en un modelo 3D puede usarse para proporcionar varias funcionalidades adicionales al dispositivo móvil, además de, o en lugar de, la pantalla de realidad aumentada. Dos ejemplos que se cree que son de particular importancia a este respecto son la operación como telémetro pasivo y como rastreador de ubicación independiente del GPS.

Específicamente, con referencia primero al telémetro pasivo, esto se refiere a la funcionalidad de acuerdo con la cual puede obtenerse una medición del rango desde la posición actual hasta cualquier ubicación dentro de una escena vista sin dirigir ninguna iluminación activa hacia esa ubicación. En cambio, una vez que la vista actual se registra al modelo 3D, la información del rango desde la posición actual hasta cualquier ubicación seleccionada en el FOV puede derivarse directamente del modelo. De acuerdo con una primera implementación de esta característica, puede descargarse un mapa de los rangos para cada píxel de la imagen desde el servidor al dispositivo móvil al final del proceso de registro, que genera esencialmente una imagen LADAR simulada de la escena vista. La información del rango luego está disponible instantáneamente bajo demanda. En una implementación alternativa, el usuario puede ingresar una solicitud para la información del rango en un píxel particular, y el dispositivo móvil luego carga un identificador del píxel correspondiente en la última imagen registrada al servidor. Luego, el servidor emplea la transformación de registro para determinar el punto correspondiente en el modelo y, por lo tanto, determina la distancia desde la ubicación de visualización actual hasta el punto solicitado. Entre paréntesis, como una función adicional de acuerdo con este segundo enfoque, también puede solicitarse al servidor sobre la distancia entre dos ubicaciones cualesquiera dentro de la escena vista. En este caso, ambas ubicaciones se identifican al seleccionar un píxel de imagen correspondiente y la distancia entre las ubicaciones correspondientes se deriva del modelo. En la mayoría de los casos, las funciones del telémetro son más precisas si el dispositivo móvil no se ha movido significativamente desde que se muestreó la última imagen registrada, aunque esta limitación puede relajarse hasta cierto alcance donde hay datos disponibles sobre el movimiento del ego del dispositivo móvil, tal como cuando se usa el procesamiento SLAM.

Volviendo ahora al rastreador de ubicación independiente del GPS, dado que el registro de la imagen cargada en el modelo 3D en la mayoría de los casos también resuelve inherentemente la posición de la cámara, el procesamiento de registro de imágenes puede usarse para proporcionar a un dispositivo móvil con una determinación de posición con relación al modelo 3D anclado geográficamente. Esto permite la verificación, corrección y seguimiento de la posición propia incluso en circunstancias en las que los datos del GPS no están disponibles. De acuerdo con una implementación de esta característica, la ubicación de las coordenadas y el cojinete y/o la orientación de la brújula pueden proporcionarse desde el servidor al dispositivo móvil después de cada registro de imagen exitoso, para mostrar al usuario o para usar en cualquier procesamiento adicional deseado en el dispositivo móvil. Opcionalmente, donde se usan técnicas SLAM para el seguimiento local, el resultado del procesamiento también puede usarse para proporcionar una actualización en tiempo real de los datos de posición hasta que se realice el siguiente registro completo. Alternativamente, o adicionalmente, pueden usarse sensores de inercia a bordo para proporcionar actualizaciones intermedias de los datos de posición durante el tiempo entre registros de imágenes completas.

Volviendo ahora a las Figuras 5 y 6, se ilustra una subcategoría de los dispositivos móviles ilustrados genéricamente en la Figura 1, que se cree que son de particular importancia por derecho propio. Esta subcategoría de dispositivos se relaciona principalmente con dispositivos que emplean una pantalla transparente, y que preferentemente incluyen también arreglos ópticos de aumento, como se ejemplifica en la Figura 6 por un par de binoculares. Los componentes que son funcionalmente equivalentes a los ya descritos anteriormente se etiquetan de manera similar. Estos incluyen la cámara 16, el sistema de procesamiento 22 con al menos un procesador 24 y almacenamiento de datos 26, sensores de posición 20, interfaz de entrada de usuario 30 y módulo de comunicaciones inalámbricas 28.

Esta clase de productos puede definirse alternativamente por el hecho de que su modo de uso es ver la escena por parte del usuario que mira a lo largo de un eje coincidente o paralelo al eje óptico de un arreglo óptico (óptica de visualización 32) dirigido hacia la escena que se ve. Esta clase de dispositivos por lo tanto también incluye varios tipos de gafas de visión nocturna y similares. Los dispositivos típicamente emplean una pantalla transparente 34, como se mencionó anteriormente. La cámara 16 preferentemente se alinea con su eje óptico paralelo al de la óptica de visualización 32, aunque el campo de visión de la cámara 16 no necesita que sea el mismo que el de la óptica de visualización, y ventajosamente puede ser más ancho para facilitar el proceso de registro.

Al implementar un dispositivo móvil de acuerdo con este aspecto de la presente invención, es posible proporcionar dispositivos comunes tales como binoculares, un telescopio (que incluye un monocular), gafas de visión nocturna, etc. con una gama de funciones de valor agregado, que incluyen una o más de: características de visualización de realidad aumentada; funcionalidad de telémetro pasivo; y funcionalidad de seguimiento de posición sin GPS, todo como se describió anteriormente.

También cae dentro del alcance de una modalidad de la presente invención un dispositivo simplificado que no incluye una pantalla transparente, sino solo una pantalla LCD alfanumérica simple. Al realizar el registro de imágenes desde la cámara 16 contra un modelo 3D, la activación de una entrada de usuario (por ejemplo, uno o más botones) 30 inicia la función de telémetro pasivo antes mencionada y/o la funcionalidad de búsqueda de posición sin GPS antes mencionada descrita anteriormente, con los resultados de las mediciones que se muestran como datos en la pantalla alfanumérica. La designación de una ubicación en la escena vista para la función de telémetro se logra preferentemente al alinear un cursor en cruz u otro símbolo de designación definido por la óptica de visualización.

En una variante de implementación, este dispositivo puede implementarse con al menos una parte del modelo 3D almacenado en el dispositivo de almacenamiento de datos local 26 y con el procesamiento de registro de imágenes realizado localmente por el sistema de procesamiento 22. Aunque tal modalidad requiere una energía de procesamiento integrada considerablemente mayor, esto puede justificarse en los casos donde es importante proporcionar un dispositivo autónomo que pueda continuar funcionando sin comunicaciones inalámbricas en curso. En ciertas modalidades, el módulo de comunicaciones inalámbricas 28 puede omitirse completamente y el modelo 3D puede cargarse en un dispositivo de almacenamiento de datos 26 de otra manera, tal como mediante la inserción de un dispositivo de almacenamiento extraíble precargado con los datos del modelo 3D, o mediante una conexión por cable temporal a un ordenador o red.

Las modalidades antes mencionadas pretenden ser ejemplos útiles para comprender la invención. La presente invención se define por las reivindicaciones adjuntas.

REIVINDICACIONES

- 5 1. Un método para operar un dispositivo móvil (10) para proporcionar una visualización de información complementaria alineada con una vista de una escena, el dispositivo móvil que incluye una cámara (16), una pantalla (18), un sistema de procesamiento (22) que incluye al menos un procesador (24) y al menos un sensor de posición (20), el dispositivo móvil que está en comunicación en red a través de una red de área amplia (WAN) (12) con un servidor remoto (14), el método que comprende las etapas de:
- 10 (a) obtener una primera imagen de la cámara (16);
 (b) cargar al menos dicha primera imagen a través de la WAN al servidor remoto (14) junto con los datos correspondientes del al menos un sensor de posición (20);
 (c) realizar el procesamiento de imágenes para rastrear el movimiento de la imagen entre dicha primera imagen y las imágenes subsecuentes obtenidas de la cámara y, por lo tanto, determinar un mapeo entre las características de una imagen de referencia y las características de una imagen actual obtenida de la cámara,
 15 dicha imagen de referencia que se selecciona del grupo que consiste en: dicha primera imagen; y una segunda imagen para la cual se cargaron al servidor parámetros de transformación correspondientes a un mapeo entre dicha primera imagen y dicha segunda imagen, dicho procesamiento que se realiza por el al menos un procesador (24) del dispositivo móvil (10), caracterizado por
 (d) recibir a través de la WAN (12) datos indicativos de una ubicación de píxel para mostrar información complementaria dentro de dicha imagen de referencia, dicha recepción que incluye recibir datos indicativos de una ubicación de píxel para mostrar información complementaria relacionada con un objeto oculto a la vista;
 (e) emplear dicho mapeo para determinar una ubicación de píxel correspondiente para la visualización de la información complementaria dentro de la imagen actual; y
 (f) mostrar la información complementaria en la pantalla_(18) correctamente alineada con la vista de la
 25 escena.
2. El método de la reivindicación 1, en donde dicho mapeo se selecciona del grupo que consiste en: una transformación espacial global que se aproxima a dicho movimiento de imagen; un mapeo de flujo óptico; y un mapeo derivado por procesamiento simultáneo de localización y mapeo (SLAM) de dichas imágenes.
- 30 3. El método de la reivindicación 1 o la reivindicación 2, en donde dicha visualización incluye la visualización de la información complementaria superpuesta a una imagen de la vista de la escena en una pantalla no transparente.
- 35 4. El método de la reivindicación 1 o la reivindicación 2, en donde dicha visualización incluye la visualización de la información complementaria en una pantalla transparente alineada con una vista directa de la escena.
5. El método de cualquier reivindicación anterior, en donde dicha ubicación de píxel para la visualización de información complementaria recibida por el dispositivo móvil se deriva al menos en parte mediante la comparación de dicha primera imagen con un modelo tridimensional de la escena.
- 40 6. El método de cualquier reivindicación anterior, en donde dicho procesamiento de imágenes continúa después de dicha visualización para generar un mapeo actualizado para una nueva imagen actual, y en donde dicha ubicación de píxel para la visualización de la información complementaria dentro de la imagen actual se actualiza de acuerdo con dicho mapeo actualizado.
- 45 7. El método de cualquier reivindicación anterior, que comprende además las etapas de:
- (a) ingresar una solicitud de información de un usuario con respecto a la ubicación de un píxel en la imagen actual;
 50 (b) emplear dicho mapeo para derivar una ubicación de píxel correspondiente en dicha imagen de referencia; y
 (c) transmitir al servidor una solicitud de información con respecto a la ubicación de píxel en dicha imagen de referencia.
- 55 8. El método de la reivindicación 7, en donde dicha solicitud de información es una solicitud de información del rango con relación a una ubicación vista.
9. El método de cualquier reivindicación anterior, que comprende además las etapas de:
- 60 (a) evaluar un indicador de fiabilidad de seguimiento basado en al menos un parámetro seleccionado del grupo que consiste en: tiempo transcurrido desde el muestreo de dicha primera imagen; grado de superposición entre la imagen actual y dicha primera imagen; y número de características rastreadas en común entre la imagen actual y dicha primera imagen; y
 (b) cuando dicho indicador de fiabilidad de seguimiento pasa un umbral, cargar una nueva imagen de referencia obtenida de la cámara al servidor.
 65

10. Un dispositivo móvil (10) para proporcionar una visualización de información complementaria alineada con una vista de una escena, el dispositivo móvil que comprende:
- (a) una cámara (16);
 - (b) una pantalla (18);
 - (c) al menos un sensor de posición (20);
 - (d) un módulo de comunicación (28) para comunicación en red con un servidor remoto; y
 - (e) un sistema de procesamiento (22) que incluye al menos un procesador (24), dicho sistema de procesamiento que está en comunicación de datos con dicha cámara, dicha pantalla, dicho al menos un sensor de posición y dicho módulo de comunicación, dicho sistema de procesamiento que se configura para implementar el método de cualquier reivindicación anterior.
11. El dispositivo móvil de acuerdo con la reivindicación 10, en donde el dispositivo móvil (10) se implementa como un dispositivo seleccionado del grupo que consiste en: un auricular de realidad aumentada; prismáticos; un telescopio; y gafas de visión nocturna.
12. Un método para operar un servidor (14) que tiene al menos un procesador (32) para soportar la visualización en un dispositivo móvil (10) de información complementaria alineada con una vista de una escena, el método que comprende las etapas de:
- (a) recibir desde el dispositivo móvil (10) a través de una red de área amplia (WAN) (12) una primera imagen obtenida de una cámara (16) del dispositivo móvil junto con datos indicativos de una posición del dispositivo móvil cuando dicha primera imagen se obtuvo;
 - (b) generar una estimación de un campo de visión de la cámara del dispositivo móvil correspondiente a dicha primera imagen;
 - (c) recuperar de una base de datos al menos parte de un modelo tridimensional anclado geográficamente correspondiente a dicha estimación de dicho campo de visión;
 - (d) coprocesar dicha primera imagen con al menos dicha parte del modelo tridimensional para derivar un mapeo de registro para mapear píxeles de dicha primera imagen a dicho modelo tridimensional, dicho coprocesamiento que se realiza mediante el uso de al menos un procesador (32) del servidor (14), caracterizado por
 - (e) determinar una ubicación de píxel dentro de dicha primera imagen para mostrar al menos un elemento de información complementaria dentro de dicha primera imagen, dicha determinación que incluye determinar una ubicación de píxel para mostrar información complementaria relacionada con un objeto oculto a la vista; y
 - (f) transmitir a través de la WAN (12) al dispositivo móvil (10) datos indicativos de dicha ubicación de píxel para la visualización de la información complementaria.
13. El método de la reivindicación 12, que comprende además transmitir al dispositivo móvil datos suficientes para la visualización de contenido gráfico en dicha ubicación de píxel, dicho contenido gráfico que se deforma para ajustarse sustancialmente a una orientación aparente de un objeto al que se refiere la información complementaria.
14. El método de la reivindicación 12, que comprende además recibir desde el dispositivo móvil datos indicativos de un campo de visión actualizado de la cámara, y en donde dicha información complementaria se recupera de una base de datos como una función de al menos dicho campo de visión actualizado.

45

Figura 1

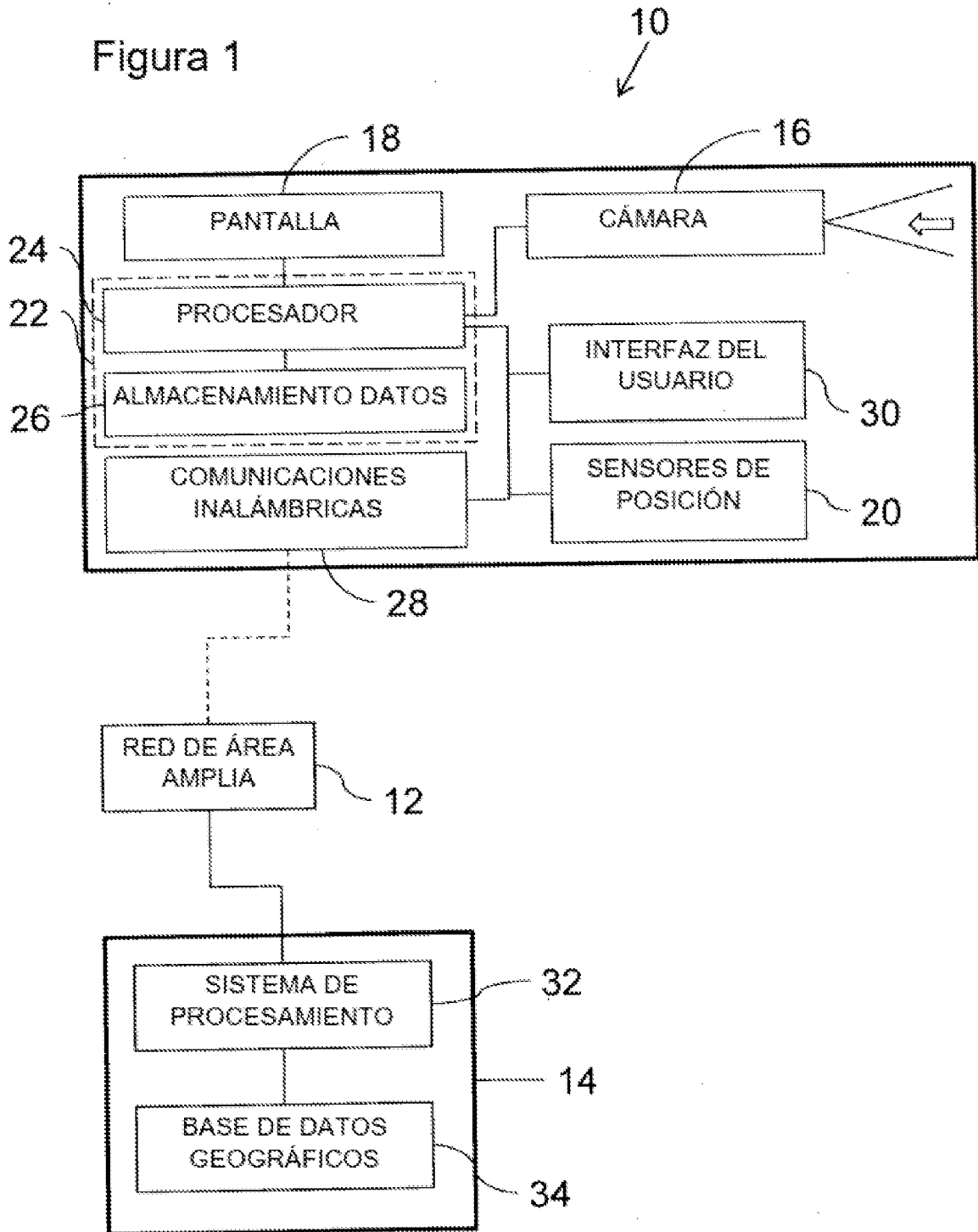


Figura 2

40

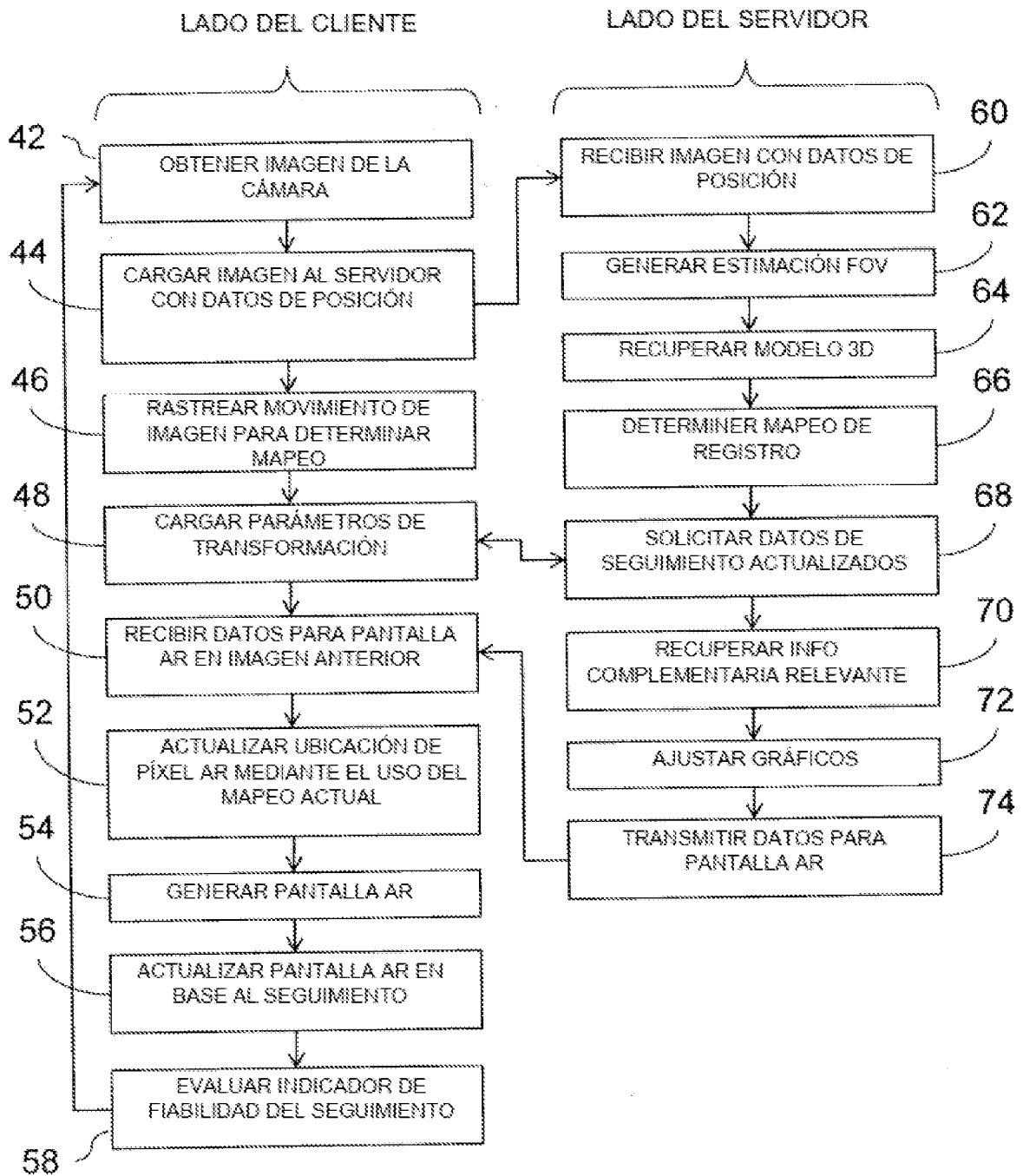


Figura 3

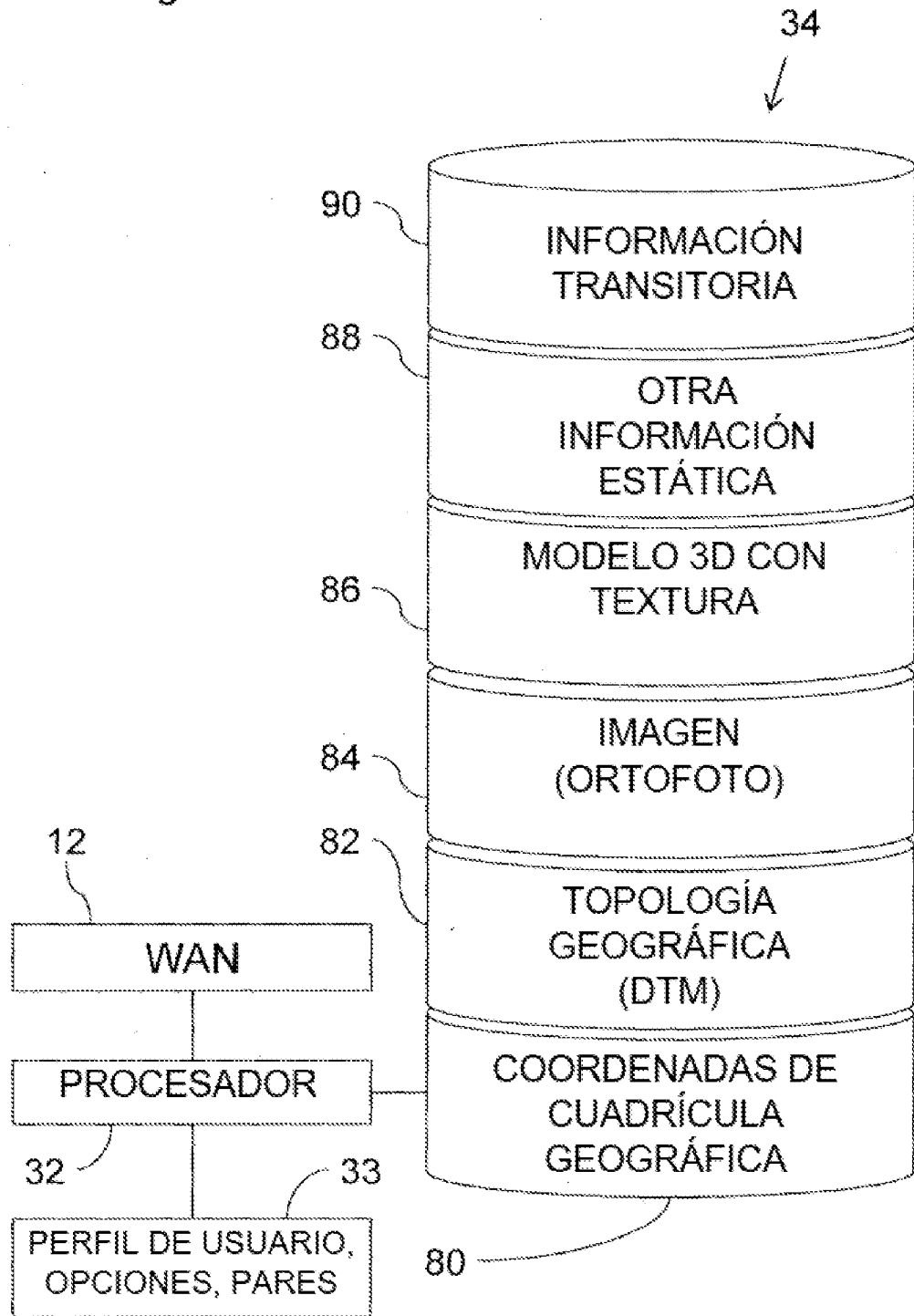


Figura 4

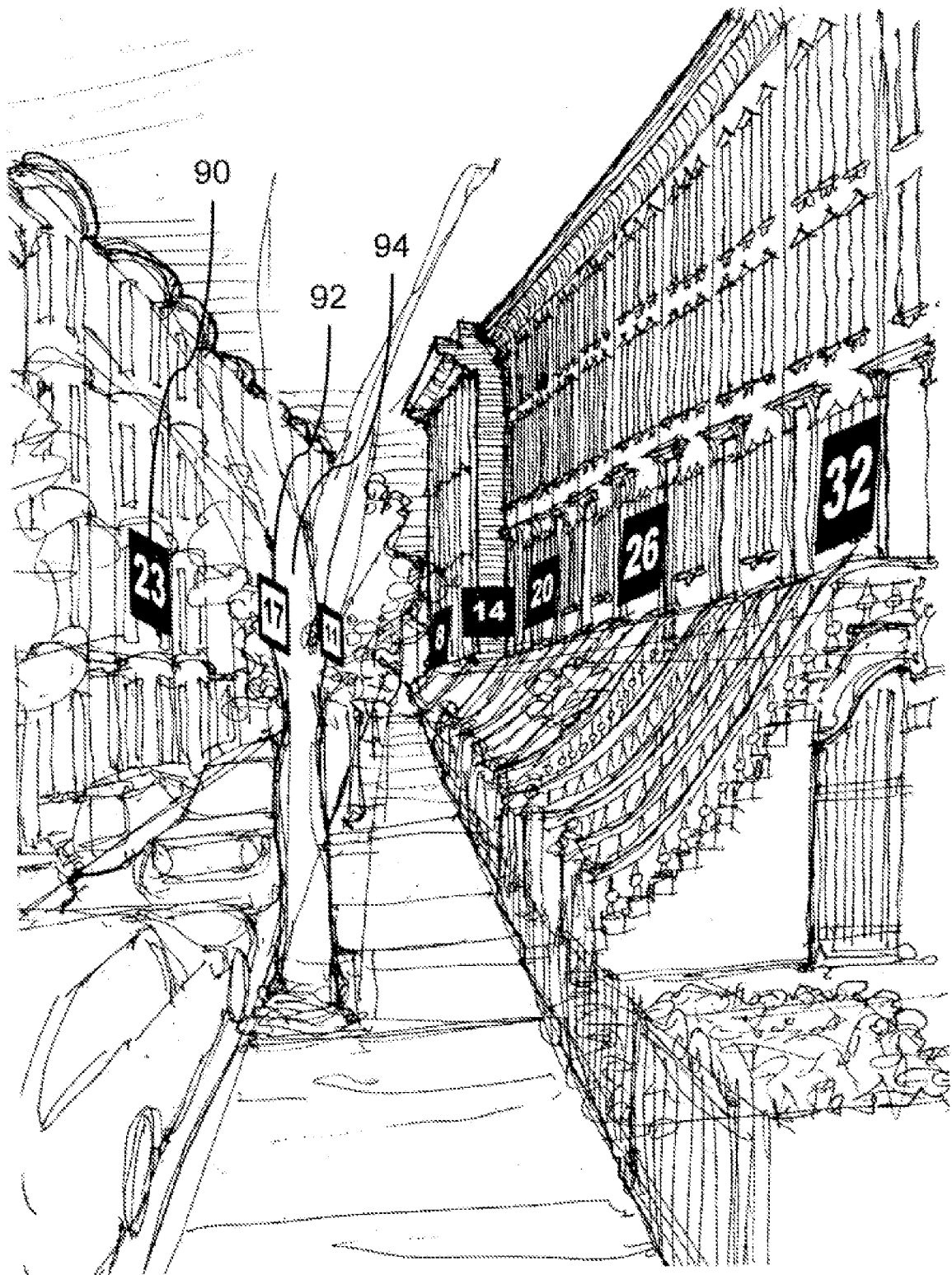


Figura 5

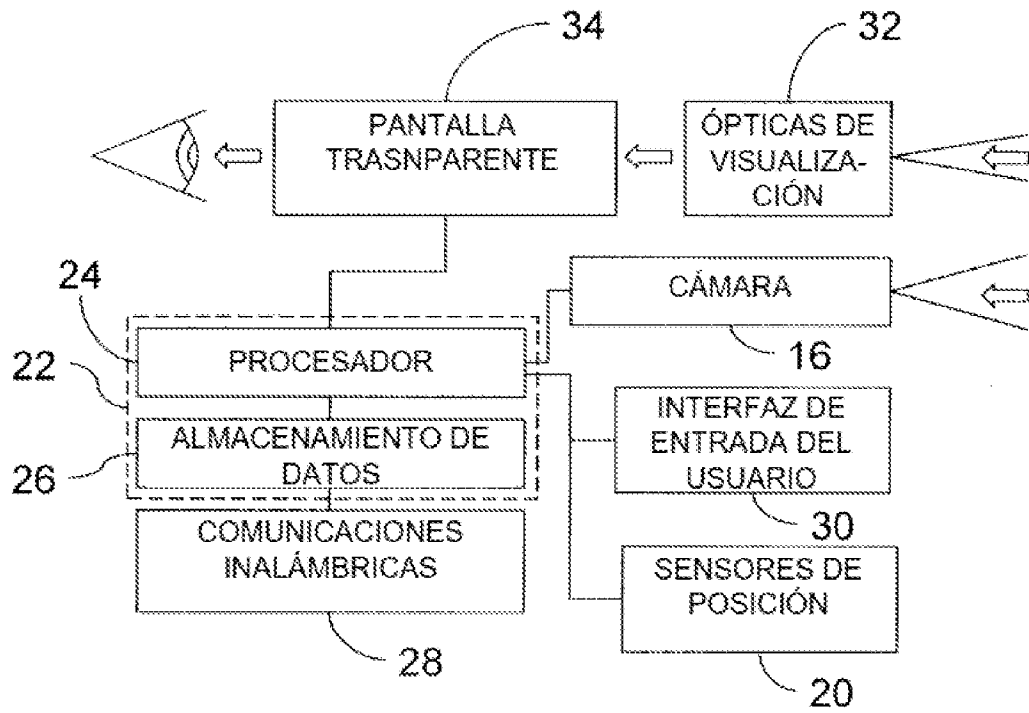


Figura 6

