

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 961 825**

51 Int. Cl.:

G06T 7/73 (2007.01)

G06V 10/143 (2012.01)

G06V 20/10 (2012.01)

G06V 20/64 (2012.01)

G06V 40/10 (2012.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.04.2020** **PCT/GB2020/051005**

87 Fecha y número de publicación internacional: **05.11.2020** **WO20221989**

96 Fecha de presentación y número de la solicitud europea: **23.04.2020** **E 20721289 (5)**

97 Fecha y número de publicación de la concesión europea: **06.09.2023** **EP 3963548**

54 Título: **Un sistema y un método de localización usando huellas**

30 Prioridad:

29.04.2019 GB 201905945

29.04.2019 EP 19275060

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

14.03.2024

73 Titular/es:

BAE SYSTEMS PLC (100.0%)

6 Carlton Gardens

London SW1Y 5AD, GB

72 Inventor/es:

ASHCROFT, JENNIFER y

JACKSON, EDDIE

74 Agente/Representante:

DEL VALLE VALIENTE, Sonia

ES 2 961 825 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 sistema y un método de localización usando huellas

5 **Antecedentes**

La presente invención se refiere a un sistema y método para la localización, en general, y más específicamente a un sistema y método para usar una cámara para realizar una localización tridimensional.

10 La localización en los campos de detección de imágenes y visión informática indica generalmente la capacidad del dispositivo para establecer su propia posición y orientación dentro de un fotograma de referencia.

Es bien sabido que el sistema de posicionamiento global (GPS) puede usarse para localizar un dispositivo habilitado para GPS. Sin embargo, los sistemas de localización basados en GPS son indeseables en áreas donde las señales de GPS son débiles o no están disponibles. Este es particularmente el caso en áreas cerradas, tales como interiores y subterráneos, y también en condiciones meteorológicas extremas. Por lo tanto, sería beneficioso tener un sistema que pueda realizar la localización sin usar señales GPS.

20 Los sistemas de localización basados en sensores de imágenes pueden evitar la necesidad de usar señales de GPS. Sin embargo, para lograr una localización tridimensional, los sistemas de localización basados en sensores de imágenes convencionales necesitan implicar cámaras de visión estéreo, en las que un par de cámaras se desplazan horizontalmente entre sí para obtener dos vistas diferentes de la escena. La información de profundidad relativa se puede obtener en forma de un mapa de disparidad. Sigue siendo difícil usar una cámara monocular, es decir, una cámara única, para realizar una localización tridimensional.

25 En los sistemas de localización basados en sensores de imágenes, al menos una característica necesitará seleccionarse de un entorno circundante para actuar como un fotograma de referencia, de manera que la posición de la o las cámaras pueda determinarse con respecto a la característica. Sin embargo, sigue siendo difícil identificar una característica adecuada para actuar como un fotograma de referencia en una variedad de entornos, ya que la mayoría de las características no están presentes en una amplia gama de entornos. Sigue existiendo la necesidad de un sistema y método que pueda usar la misma característica en una variedad de entornos como fotograma de referencia para la localización.

30 Por consiguiente, un objeto de la presente invención es superar los problemas de los sistemas y métodos de localización convencionales. Las realizaciones descritas a continuación no se limitan a implementaciones que resuelven alguna o todas las desventajas de los sistemas conocidos.

35 HUANG SHIYAO Y COL: El documento "Camera Calibration from Periodic Motion of a Pedestrian", CONFERENCIA IEEE 2016 SOBRE VISIÓN POR ORDENADOR Y RECONOCIMIENTO DE PATRONES (CVPR), IEEE, 27 de junio de 2016 (27-06-2016), páginas 3025-3033, describe la calibración de la cámara de secuencias de imágenes de un peatón.

Resumen de la invención

45 A continuación, se presenta una breve descripción de la invención para proporcionar una comprensión básica al lector. El presente sumario no pretende identificar características clave o características esenciales de la materia reivindicada, ni pretende usarse como una ayuda para determinar el alcance de la materia reivindicada. Su único propósito es presentar una selección de conceptos desvelados en la presente memoria de forma simplificada como preludio a la descripción más detallada que se presenta más adelante.

50 Según un primer aspecto, se proporciona un sistema, que comprende: una cámara para capturar una imagen, al menos un procesador y al menos una memoria, almacenado la al menos una memoria instrucciones configuradas para hacer que el procesador: obtenga una imagen de una escena capturada usando la cámara, comprendiendo la escena una huella de un usuario del sistema; identifique la huella en la imagen; identifique una primera pluralidad de puntos de referencia, teniendo cada uno una posición fija con respecto a la huella en la imagen; identifique una segunda pluralidad de puntos de referencia teniendo cada uno una posición fija con respecto a un modelo tridimensional del pie del usuario construido basándose en al menos una dimensión conocida del pie del usuario; determine una matriz de transformación mapeando la segunda pluralidad de puntos de referencia del modelo tridimensional a la primera pluralidad de puntos de referencia; y determine, basándose en la matriz de transformación, una posición tridimensional de la cámara con respecto a la huella en la escena en el momento de capturar la imagen.

Preferiblemente, la al menos una memoria almacena instrucciones configuradas para hacer que el procesador determine una orientación tridimensional de la cámara con respecto a la huella en la escena en el momento de capturar la imagen.

65

Preferiblemente, dicha imagen de la escena es una primera imagen de la escena capturada en un primer punto en el tiempo, dicha matriz de transformación es una primera matriz de transformación, y la al menos una memoria almacena instrucciones configuradas para hacer que el procesador: obtenga una segunda imagen de la escena capturada usando la cámara en un segundo punto en el tiempo; identifique la huella en la segunda imagen; identifique una tercera pluralidad de puntos de referencia, teniendo cada uno una posición fija con respecto a la huella en la segunda imagen; identifique una cuarta pluralidad de puntos de referencia teniendo cada uno una posición fija con respecto al modelo tridimensional del pie del usuario construido basándose en la al menos una dimensión conocida del pie del usuario; y determine una segunda matriz de transformación mapeando la cuarta pluralidad de puntos de referencia del modelo tridimensional a la tercera pluralidad de puntos de referencia en el plano de la segunda imagen.

Preferiblemente, la posición tridimensional de la cámara es una primera posición tridimensional de la cámara con respecto a la huella en la escena en el primer punto en el tiempo, y la al menos una memoria almacena instrucciones configuradas para hacer que el procesador: determine una matriz de transformación entre fotogramas basándose en la primera matriz de transformación y la segunda matriz de transformación, y basándose en la matriz de transformación entre fotogramas y la primera posición tridimensional de la cámara con respecto a la huella en la escena, una segunda posición tridimensional de la cámara con respecto a la huella en la escena en el segundo punto en el tiempo.

Preferiblemente, la al menos una memoria almacena instrucciones configuradas para hacer que el procesador calcule una ubicación del mundo real actualizada de la cámara basándose en una ubicación del mundo real inicial de la cámara y la primera y segunda posiciones tridimensionales determinadas de la cámara con respecto a la huella en la escena en el momento de capturar la imagen.

Según un segundo aspecto, se proporciona un método implementado por ordenador, que comprende: recibir una representación digital de una imagen de una escena capturada usando una cámara, comprendiendo la escena una huella de un usuario; identificar la huella en la imagen; identificar una primera pluralidad de puntos de referencia, teniendo cada uno una posición fija con respecto a la huella en la imagen; identificar una segunda pluralidad de puntos de referencia, teniendo cada uno una posición fija con respecto a un modelo tridimensional del pie del usuario construido basándose en al menos una dimensión conocida del pie del usuario; determinar una matriz de transformación mapeando la segunda pluralidad de puntos de referencia del modelo tridimensional a la primera pluralidad de puntos de referencia; y determinar, basándose en la matriz de transformación, una posición tridimensional de la cámara con respecto a la huella en la escena en el momento de capturar la imagen.

Según un tercer aspecto, se proporciona un medio legible por ordenador que almacena instrucciones ejecutables por ordenador configuradas para realizar el método anterior.

Breve descripción de los dibujos

La Figura 1 ilustra un entorno en el que puede usarse un sistema según una realización de la presente invención.

La Figura 2 es un diagrama de bloques de los componentes de un sistema según una realización de la presente invención.

La Figura 3 es un histograma que muestra una distribución estadística ilustrativa de intensidades de calor de píxeles en una imagen térmica según una realización.

La Figura 4 muestra resultados ilustrativos de filtrado de una imagen térmica capturada en cuatro umbrales de intensidad térmica diferentes según una realización.

La Figura 5 muestra un mapa de probabilidad ilustrativo que ilustra la probabilidad de que se presenten huellas térmicas en diversas ubicaciones de una imagen térmica según una realización.

La Figura 6 ilustra un método ilustrativo para identificar las huellas situadas más cerca de una cámara en el momento en que se captura la imagen por la cámara según una realización.

La Figura 7 ilustra un conjunto ilustrativo de puntos de referencia para su uso en la identificación de una pose de una cámara con respecto a una huella según una realización.

La Figura 8 ilustra líneas que indican orientaciones de huellas identificadas en una imagen según una realización.

La Figura 9 ilustra un segmento de línea ilustrativo que indica una dirección del movimiento de un usuario según una realización.

La Figura 10 ilustra un método ilustrativo para identificar las huellas generadas inmediatamente antes de las ubicadas más cerca de una cámara en el momento en que se captura la imagen según una realización.

La Figura 11 ilustra una segmentación de todas las huellas identificadas a partir de una imagen térmica según una realización.

Descripción detallada

La descripción detallada proporcionada a continuación con respecto a los dibujos adjuntos se pretende como una descripción de los presentes ejemplos y no pretende representar las únicas formas en las que se construyen o utilizan los presentes ejemplos. La descripción expone las funciones del ejemplo y la secuencia de operaciones para construir y operar el ejemplo. Sin embargo, las mismas funciones o secuencias equivalentes pueden lograrse mediante diferentes ejemplos.

En general, la presente invención se refiere a una técnica novedosa para usar las huellas de un usuario como una característica de localización y rastreo. En algunas realizaciones, las huellas son huellas térmicas de un usuario, que se generan debido a una diferencia térmica entre las áreas de tierra, que han estado en contacto térmico con los pies del usuario, y las áreas de tierra circundantes, que no han estado en contacto térmico con los pies del usuario y, por lo tanto, no se ven afectadas por la temperatura corporal del usuario. El uso de huellas térmicas como una característica de localización y rastreo son ventajas, ya que son una característica generada por el usuario y están disponibles en una variedad de entornos siempre que haya una diferencia de temperatura entre el cuerpo de una persona y el entorno circundante.

Según diversas realizaciones, las imágenes de huellas en una escena son capturadas por una cámara usada por un usuario que podría montarse en cualquier posición adecuada en un usuario para facilitar su uso, por ejemplo, en un casco o pantalla montada en la cabeza (HMD) o en la espalda del usuario. En algunas realizaciones, la cámara es monocular, es decir, una sola cámara. El uso de una sola cámara como sensor de imágenes en un sistema de localización es ventajoso ya que reduce los costes de fabricación y reduce el tamaño y el peso del sistema en comparación con las cámaras de visión estéreo convencionales.

Una imagen capturada por la cámara monocular se procesa para identificar y localizar huellas en las imágenes. Un modelo tridimensional (3D) de la huella del usuario se construye basándose en al menos una dimensión del pie del usuario y un conjunto de puntos de referencia 3D que se identifican en el modelo 3D. A continuación, se encuentra un conjunto correspondiente de puntos de referencia bidimensionales (2D) en una huella identificada a partir de la imagen. A partir de las correspondencias conocidas entre los conjuntos de puntos 2D y 3D, los algoritmos estándar, tales como para resolver el problema de perspectiva desde n puntos (PnP), pueden usarse para estimar la posición y orientación, o pose, de la cámara con respecto a la huella en la escena en el momento de capturar la imagen. La pose de la cámara puede usarse para estimar la pose del usuario con respecto a la huella en la escena en el momento de capturar la imagen debido a la posición y orientación conocidas de la cámara con respecto al usuario.

En algunas realizaciones, al capturar una pluralidad de imágenes de las huellas del usuario en diferentes puntos de tiempo, es posible rastrear el movimiento del usuario con respecto al suelo en el que camina el usuario. El rastreo puede inicializarse desde una ubicación inicial con una coordenada del mundo real conocida. A medida que el usuario se aleja de la ubicación inicial, el usuario genera huellas. Al capturar y realizar el procesamiento de estas huellas, puede estimarse la nueva posición del usuario con respecto a la ubicación inicial y puede usarse para actualizar la ubicación del usuario en el mundo real.

Las realizaciones de la presente invención pueden usarse en cualquier entorno aplicable, incluyendo entornos en los que el rastreo basado en GPS y Wi-Fi no está disponible, tal como interiores, subterráneos y áreas con condiciones meteorológicas extremas. También es un sistema ventajoso para rastrear áreas anchas o en emergencias donde no es posible configurar cualquier infraestructura para facilitar el rastreo de antemano, por ejemplo, en situaciones que involucran soldados enviados en misiones críticas y primeros respondedores en la escena de una emergencia. A medida que se generan las huellas térmicas, y la cámara es usada por el usuario, el propio entorno no necesita contener ningún contenido distintivo y no hay necesidad de montar ninguna cámara o marcadores en el entorno antes del rastreo. Además, a medida que se generan las huellas por la diferencia de calor entre el usuario y su entorno, las realizaciones de la invención podrían aplicarse independientemente de si las huellas térmicas están más frías o más calientes que el entorno del usuario.

En algunas realizaciones, el sistema de rastreo puede ampliarse con sensores adicionales para localización, por ejemplo, sensores inerciales, transceptores de señal GPS y transceptores Wi-Fi. Estos sensores adicionales podrían usarse en combinación con la cámara en el sistema para mejorar la precisión de la localización. En algunas realizaciones, se puede proporcionar opcionalmente un equipo para ajustar la firma de calor de los pies del usuario, por ejemplo, para generar más calor y proporcionar una huella térmica más pronunciada.

La operación de algunas realizaciones de la invención se describirá en la presente memoria con referencia a las Figuras 1-11.

La Figura 1 ilustra un entorno en el que puede usarse una realización de la presente invención. Mientras que un usuario 102 camina sobre un suelo, se generan varias huellas 104. En algunas realizaciones, las huellas 104 pueden

ser huellas térmicas. Las huellas térmicas se generan como resultado de una diferencia térmica entre las áreas de tierra, que tienen temperaturas afectadas por la temperatura corporal y fricción del usuario, y sus áreas de tierra circundantes, que no han sido afectadas térmicamente por la temperatura corporal del usuario. En otras realizaciones, las huellas 104 pueden ser hendiduras formadas por los pies del usuario en un suelo arenoso, embarrado, húmedo o herboso, en lugar de huellas térmicas.

El usuario 102 puede usar un sistema 100, que comprende un dispositivo 106 de detección de imágenes que apunta generalmente al suelo para capturar imágenes del suelo. En una realización, como se muestra en la Figura 1, el dispositivo 106 de detección de imágenes puede montarse en la parte posterior de la cabeza del usuario y orientarse para capturar las huellas 104 generadas en el suelo detrás del usuario. En otras realizaciones, el dispositivo 106 de detección de imágenes puede colocarse en la espalda del usuario o cualquier otra posición en el cuerpo del usuario adecuado para capturar las huellas 104 del usuario. En algunas realizaciones, el dispositivo 106 de detección de imágenes puede ser una cámara térmica. En otras realizaciones, el dispositivo 106 de detección de imágenes puede ser una cámara óptica. En varias realizaciones, el dispositivo de detección de imágenes puede ser una cámara monocular.

En algunas realizaciones, el sistema 100 puede comprender opcionalmente una interfaz 108 de usuario, que está acoplada operativamente al dispositivo 106 de detección de imágenes. En algunas realizaciones, la interfaz 108 de usuario y el dispositivo 106 de detección de imágenes pueden formar un solo dispositivo. Por ejemplo, estos dispositivos pueden formar partes de una pantalla montada en la cabeza o un casco. En algunas realizaciones, la interfaz 108 de usuario y el dispositivo 106 de detección de imágenes pueden ser dos dispositivos separados. Por ejemplo, la interfaz 108 de usuario puede ser un dispositivo informático portátil (no mostrado en la Figura 1), tal como un teléfono móvil, una tableta o un asistente personal digital, que está en comunicación operativa con el dispositivo 106 de detección de imágenes.

La Figura 2 ilustra un conjunto ilustrativo de componentes para el sistema 100. El sistema 100 puede implementarse como uno o más dispositivos informáticos y/o electrónicos. El sistema 100 comprende uno o más procesadores 202 que pueden ser microprocesadores, controladores o cualquier otro tipo adecuado de procesadores para procesar instrucciones ejecutables por ordenador para controlar la operación del sistema 100. El software de plataforma que comprende un sistema operativo 206 o cualquier otro software de plataforma adecuado puede proporcionarse en el sistema para permitir que el software 208 de aplicación se ejecute en el sistema. En algunas realizaciones, el software 208 de aplicación puede comprender un programa de software para procesar imágenes, obtener datos de las imágenes y procesar los datos obtenidos de las imágenes según diversos métodos descritos en la presente memoria. Los componentes del sistema 100 descritos en la presente memoria pueden estar encerrados en una carcasa 210 como se muestra en la Figura 2 o en más de una carcasa (no mostradas en la Figura 2).

Las instrucciones ejecutables por ordenador pueden proporcionarse usando cualquier medio legible por ordenador que sea accesible por el sistema 100. Los medios legibles por ordenador pueden incluir, por ejemplo, medios de almacenamiento informáticos tales como una memoria 204 y medios de comunicación. Los medios de almacenamiento de ordenador, tales como una memoria 204, incluyen medios volátiles y no volátiles, extraíbles y no extraíbles implementados en cualquier método o tecnología para el almacenamiento de información tal como instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos. Los medios de almacenamiento de ordenador incluyen, pero no se limitan a, RAM, ROM, EPROM, EEPROM, memoria flash u otra tecnología de memoria, CD-ROM, discos versátiles digitales (DVD) u otro almacenamiento óptico, casetes magnéticos, cinta magnética, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio no de transmisión que pueda usarse para almacenar información para el acceso por un dispositivo informático. Por el contrario, los medios de comunicación pueden incorporar instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos en una señal de datos modulada, tal como una onda portadora u otro mecanismo de transporte. Aunque el medio de almacenamiento informático (la memoria 204) se muestra dentro del sistema 100, un experto en la técnica apreciará que al menos una parte del almacenamiento puede distribuirse o ubicarse de forma remota y accederse a través de una red u otro enlace de comunicación (por ejemplo, usando una interfaz 212 de comunicación).

El sistema 100 puede comprender un controlador 214 de entrada/salida dispuesto para recibir y procesar una entrada de uno o más dispositivos 218 de entrada que pueden estar separados o ser integrales al sistema 100 y dispuestos para emitir información a uno o más dispositivos 216 de salida que pueden estar separados o ser integrales al sistema 100. En algunas realizaciones, los dispositivos 218 de entrada pueden comprender dispositivos de entrada para controlar la operación del sistema 100, tal como un conjunto de botones o claves. Por ejemplo, los dispositivos 218 de entrada pueden comprender claves para controlar una cámara, tal como ajustar una orientación y/o un zoom de una cámara, y/o para manipular una imagen que se muestra en una pantalla. Los dispositivos 218 de entrada también pueden incluir medios adecuados para recibir un valor de un usuario. En una realización, el valor es una dimensión y/o tamaño del pie del usuario. En algunas realizaciones, los dispositivos 218 de entrada pueden incluir un control remoto.

Los dispositivos 218 de entrada pueden comprender además al menos un dispositivo de detección. El al menos un dispositivo de detección puede incluir un dispositivo de detección de imágenes, tal como una cámara. En algunas

realizaciones, la cámara es una cámara de formación de imágenes térmicas. En otras realizaciones, la cámara puede ser una cámara óptica. En diversas realizaciones, la cámara es monocular. En algunas realizaciones, el al menos un dispositivo de detección puede comprender adicionalmente un sensor inercial.

Los dispositivos 216 de salida pueden incluir una pantalla de visualización. En una realización, el dispositivo 216 de salida también puede actuar como un dispositivo de entrada, por ejemplo, cuando el dispositivo 216 de salida es una pantalla táctil. El controlador 214 de entrada/salida también puede emitir datos a dispositivos distintos del dispositivo de salida, por ejemplo, a un dispositivo informático conectado localmente, tal como un teléfono inteligente. Según una realización, el procesamiento de imágenes y los cálculos basados en datos obtenidos de imágenes capturadas por el dispositivo 218 de entrada y/o cualquier otra funcionalidad tal como se describe en las realizaciones, pueden implementarse por software o firmware, por ejemplo, el sistema operativo 206 y el software 208 de aplicación trabajan juntos y/o independientemente, y son ejecutados por el procesador 202.

La interfaz 212 de comunicación permite que el sistema 100 se comuniquen con otros dispositivos y sistemas. La interfaz 212 de comunicación puede incluir transceptores de señal GPS, transceptores de señal Wi-Fi y/o transceptores Bluetooth™.

La funcionalidad descrita en la presente memoria en las realizaciones puede realizarse, al menos en parte, por uno o más componentes lógicos de hardware. Según una realización, el dispositivo informático 100 está configurado por los programas 206, 208 almacenados en la memoria 204 cuando son ejecutados por el procesador 202 para ejecutar las realizaciones de las operaciones y la funcionalidad descritas. Alternativamente, o además, la funcionalidad descrita en la presente memoria puede realizarse, al menos en parte, por uno o más componentes lógicos de hardware. Por ejemplo, y sin limitación, los tipos ilustrativos de componentes lógicos de hardware que pueden usarse incluyen matrices de puertas programables en campo (FPGA), circuitos integrados de aplicación específica (ASIC), productos estándar de aplicación específica (ASSP), sistemas de sistema en chip (SOC), dispositivos lógicos programables complejos (CPLD), unidades de procesamiento gráfico (GPU).

Identificar huellas en una imagen

Según diversas realizaciones, que se describirán con referencia a las Figuras 3 a 11, los dispositivos 218 de entrada del sistema 100 comprenden una cámara 106 de formación de imágenes térmicas para capturar imágenes térmicas. La cámara 106 de formación de imágenes térmicas puede ser una cámara monocular. En operación, la cámara 106 de formación de imágenes térmicas puede colocarse y orientarse en el usuario 102 para capturar huellas generadas en el suelo detrás del usuario 102.

La operación del sistema 100 comienza con la cámara de formación de imágenes térmicas que captura una imagen térmica. La imagen térmica puede comprender al menos una huella térmica 104 del usuario 100. Las huellas térmicas 104 pueden identificarse mediante cualquier método adecuado. A continuación, se describe un proceso ilustrativo para identificar posibles huellas a partir de una imagen térmica.

Una imagen térmica comprende una matriz bidimensional de píxeles, cada uno de los cuales tiene un valor de píxel que representa una intensidad térmica en el píxel respectivo. La Figura 3 muestra un histograma, que representa una distribución estadística de píxeles en un intervalo de intensidades térmicas para una imagen térmica capturada. El eje x del gráfico de la Figura 3 representa la intensidad térmica, y el eje y del gráfico representa el número de píxeles en la imagen térmica capturada. A continuación, la intensidad térmica en la que está presente el mayor número de píxeles se puede encontrar a partir del histograma. Esta intensidad térmica es de aproximadamente $2,5 \times 10^4$ en el caso de la Figura 3. A continuación, se elige un número de umbrales de intensidad térmica de manera que generalmente se distribuyen uniformemente alrededor de la intensidad térmica a intervalos iguales. En el caso de la Figura 3, los umbrales de intensidad térmica se eligen para que sean 1×10^4 , 2×10^4 , 3×10^4 y 4×10^4 . A continuación, en cada umbral de intensidad térmica, la imagen térmica se filtra de manera que todos los píxeles con una intensidad térmica por debajo del umbral respectivo se filtran como ruido. Las imágenes filtradas en cada uno de los cuatro umbrales de intensidad térmica se muestran en la mitad superior de la Figura 4.

A continuación, en cada una de las imágenes filtradas, se puede ajustar un contorno a cada grupo de píxeles vecinos. Los contornos pueden llenarse usando cualquier método adecuado. A continuación, los contornos pueden filtrarse por sus áreas, es decir, por el número de píxeles dentro de cada contorno, de manera que los píxeles dentro de los contornos, que tienen un área fuera de un rango predeterminado (por ejemplo, 500-10000 píxeles en el caso de la Figura 4), se filtran como ruido. Esta etapa filtra los contornos que son demasiado grandes o demasiado pequeños para calificar como posibles huellas.

Debido a la forma de las huellas y su similitud con las formas de elipse, las elipses se ajustan a los contornos restantes como se muestra en la Figura 4. Aunque aquí se elige una forma de elipse, también se pueden usar otras formas que pueden aproximar la forma real del pie de un usuario.

Como resultado, se ha proporcionado cada uno de los grupos restantes de píxeles en una imagen en un umbral con un contorno correspondiente y una elipse correspondiente. El contorno es una aproximación más cercana de la forma exacta de un grupo de píxeles vecinos que la elipse correspondiente.

5 A continuación, puede calcularse un conjunto de parámetros para evaluar la probabilidad de que un grupo de píxeles represente una huella en cada umbral. En un ejemplo, los parámetros pueden incluir uno o más de los siguientes: 1) la cantidad y/o proporción de que un contorno del grupo de píxeles se llene en una elipse correspondiente, 2) el tamaño de las hendiduras formadas entre un contorno y su elipse correspondiente, y 3) la extensión en la que un contorno coincide con una elipse correspondiente. A continuación, para una imagen en cada umbral de intensidad térmica, los grupos de píxeles se filtran adicionalmente basándose en uno o más de los parámetros anteriores, de modo que los grupos de píxeles, que mejor coinciden con una forma de elipse (es decir, la forma de una huella), se eligen y los otros grupos de píxeles se filtran. El resultado es que, para cada umbral de intensidad térmica, se produce una imagen filtrada que tiene los grupos de píxeles que representan ubicaciones de huellas más probables.

15 Un mapa de probabilidad puede obtenerse entonces de una suma (o superposición) de las imágenes filtradas en los diversos umbrales. Como se muestra en el mapa de probabilidad en la Figura 5, las áreas más brillantes indican una mayor probabilidad de la presencia de una huella y áreas más oscuras indican una menor probabilidad de que una huella esté presente.

20 Lo anterior solo representa un proceso ilustrativo para identificar posibles huellas a partir de una imagen capturada. También pueden usarse otros métodos adecuados para identificar huellas de una imagen o una imagen térmica.

Opcionalmente, en la siguiente etapa, se identifican la o las huellas más recientes, es decir, la o las huellas situadas más cerca de la cámara en el momento de capturar la imagen.

25 En la realización de esta invención, como se conoce la ubicación y orientación de la cámara con respecto al usuario, se pueden usar para identificar la huella o huellas situadas más cerca de la cámara en el momento de capturar la imagen. Al operar la cámara y orientarse para capturar imágenes en el suelo detrás del usuario, si se sabe que el lado superior de la imagen de las Figuras 4-6 es proximal al usuario cuando se capturó la imagen, puede obtenerse que la o las huellas más recientes deben ser las derechas situadas debajo del centro superior de la imagen capturada.

30 En consecuencia, se puede crear un área en el mapa de probabilidad en el que probablemente se ubique la o las huellas más recientes. El lado superior de esta área debe alinearse con el lado superior de la imagen. La forma y el tamaño de esta área pueden predefinirse de cualquier manera adecuada. En el ejemplo de la Figura 6, el área tiene una forma rectangular y las dimensiones del rectángulo pueden ser figuras empíricas obtenidas experimentalmente. A continuación, los grupos de píxeles que tienen la mayor probabilidad de representar huellas, es decir, los grupos de píxeles más brillantes, se identifican en el área rectangular para ser las huellas más recientes en el momento de capturar la imagen. Las huellas más recientes pueden segmentarse como se muestra en las Figuras 7 y 8.

40 Opcionalmente, las huellas anteriores, que son las huellas formadas inmediatamente antes de las huellas más recientes, también pueden identificarse basándose en la dirección del movimiento de un usuario obtenido a partir de las orientaciones de las huellas identificadas más cerca de la cámara cuando se capturó la imagen.

45 La orientación de una huella puede identificarse ajustando la huella a una forma geométrica, por ejemplo, un rectángulo o una elipse. En algunas realizaciones, el par más reciente de huellas segmentadas a partir de la imagen se ajustan a rectángulos. Según muestra la Figura 8, se calcula un eje longitudinal para el rectángulo ajustado a cada huella más reciente y se considera que representa la orientación de la huella respectiva.

50 A continuación, se calcula una línea biseccionada de los dos ejes longitudinales de las huellas. En una realización, la línea bisectada es un bisector de ángulo interno del ángulo formado por los dos ejes longitudinales de las huellas más recientes. La línea biseccionada representa la dirección del movimiento del usuario como se muestra en la Figura 9.

55 La línea bisectriz de los dos ejes longitudinales del par más reciente de huellas identificadas en una imagen puede considerarse para indicar una orientación o pose de un usuario en el momento de capturar la imagen. También puede usarse para estimar la dirección de movimiento del usuario y predecir dónde va a continuación el usuario.

A medida que se puede considerar que el usuario se mueve generalmente a lo largo de la dirección de la línea bisectada, la línea bisectriz puede usarse para identificar las huellas anteriores formadas antes de las identificadas más cerca de la cámara en el momento de capturar la imagen. Se puede extraer un segmento de línea de una longitud predeterminada a lo largo de la línea bisectada, como se muestra en las Figuras 9 y 10. La longitud predeterminada puede ser igual a un tamaño de paso de un usuario. Puede ser un valor empírico o un valor proporcionado o medido por un usuario. En una realización, la longitud predeterminada es tres veces la longitud de la huella identificada más próxima a la cámara mientras se capturó la imagen. El extremo del segmento de línea proximal a la posición de la cámara puede estar ubicado entre el par de huellas de los pies más reciente, por ejemplo, en el punto medio de un segmento de línea que une los centros del par de huellas más recientes. El extremo del segmento de línea distal a la posición de la cámara puede usarse para centrar una segunda área, tal como una segunda área de rectángulo como

se muestra en la Figura 10. La forma y las dimensiones de la segunda área pueden predefinirse o calcularse basándose en el tamaño del pie del usuario o de cualquier manera adecuada. A continuación, las huellas anteriores pueden identificarse como aquellas que tienen la mayor probabilidad en el mapa de probabilidad (distinto de las ya identificadas como huellas más recientes) en el segundo rectángulo.

Opcionalmente, el proceso anterior puede repetirse (usando la línea bisectriz de los ejes longitudinales del segundo par de huellas) para encontrar más huellas en la imagen térmica, si hay más huellas en la imagen. La Figura 11 muestra la segmentación de todas las huellas identificadas en la imagen capturada.

Determinar una posición 3D de la cámara con respecto a la posición de una huella en la escena

La posición y orientación, o pose, de la cámara con respecto a una huella en la escena en el momento de la captura de imágenes se puede estimar resolviendo un problema de perspectiva desde n puntos (PnP), que es el problema de estimar la pose de una cámara dada un conjunto de puntos 3D en el mundo y sus correspondientes proyecciones 2D en una imagen.

Más específicamente, el problema PnP se define como un conjunto de n puntos 3D en un fotograma de referencia mundial y sus correspondientes proyecciones de imágenes 2D, así como los parámetros de cámara intrínseca calibrados, que determinan las 6 poses de grados de libertad de la cámara en forma de rotación y traslación con respecto al mundo. Matemáticamente, el problema PnP puede definirse como:

$$s p_c = K [R | T] p_w \quad (1)$$

donde $p_w = [x \ y \ z]^T$ son las coordenadas 3D de los puntos de referencia 3D, $p_c = [u \ v]^T$ son las coordenadas correspondientes de los puntos de referencia 2D, K es la matriz de parámetros de cámara intrínseca, s es un factor de escala para el punto de imagen, y R y T son la rotación 3D y la traslación 3D de la cámara que se están calculando. Entonces la ecuación (1) se puede expandir como:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & \gamma & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2)$$

donde f_x y f_y son las longitudes focales escaladas, γ es el parámetro de sesgo que a veces se supone que es 0, y $(u_0 \ v_0)$ es el punto principal.

La huella identificada en la imagen capturada puede considerarse como una proyección bidimensional de la huella del usuario. Se puede construir un modelo tridimensional de la huella del usuario basándose en las dimensiones reales del pie del usuario. Puede proporcionarse un grupo de n puntos de referencia en el modelo tridimensional. Si un grupo de n puntos de referencia correspondientes se identifican en la proyección bidimensional (es decir, la huella identificada en la imagen), entonces el problema PnP puede resolverse para dar una posición estimada de la cámara con respecto a la huella en la escena en el momento de la captura de imágenes.

El modelo tridimensional para la huella del usuario puede construirse según las dimensiones reales del pie del usuario. En algunas realizaciones, el modelo tridimensional es una aproximación de la forma del pie de un usuario. Puede elegirse cualquier forma que se aproxime a una huella humana para el modelo tridimensional. En una realización, se elige un cuboide rectangular para el modelo tridimensional de la huella del usuario. La longitud y el ancho del cuboide rectangular se pueden establecer para que sean la longitud y el ancho del pie del usuario. Como el cuboide rectangular representa un modelo tridimensional de la huella del usuario, puede establecerse para tener una altura muy pequeña e insignificante. En un ejemplo, la altura del modelo tridimensional se establece para ser una unidad en la ecuación (2) anterior. En un ejemplo, se eligen cuatro esquinas de un rectángulo formado por la longitud y el ancho del cuboide rectangular en el modelo tridimensional como puntos de referencia. Sin embargo, sería posible establecer otros puntos en el modelo tridimensional como los puntos de referencia.

A continuación, los puntos de referencia correspondientes se identifican en una huella en la imagen capturada, que es una proyección bidimensional de la huella. Un rectángulo A se ajusta a la huella identificada en la imagen capturada. Según muestra la Figura 7, se eligen las cuatro esquinas del rectángulo para que sean los puntos de referencia de la proyección bidimensional de la huella, ya que corresponden a las cuatro esquinas del modelo tridimensional de la huella. El conjunto de puntos de referencia de la proyección bidimensional tiene que corresponder al conjunto de puntos de referencia en el modelo tridimensional, de modo que el conjunto anterior de puntos es un resultado de una proyección bidimensional del último conjunto de puntos.

La matriz de cámara también es un parámetro conocido. Puede predeterminarse mediante calibración de la cámara. Puede usarse cualquier método de calibración de cámara conocido adecuado para determinar la matriz de cámara. En un ejemplo, puede obtenerse la matriz de la cámara calibrando la cámara usando un tablero de ajedrez.

Por lo tanto, en la ecuación (1) anterior, p_c puede obtenerse a partir de los puntos de referencia en el modelo 3D de la huella, p_w puede obtenerse de los puntos de referencia correspondientes en la proyección 2D de la huella en la imagen capturada, y K puede determinarse mediante calibración de cámara. En un ejemplo, el factor de escala s en la ecuación (1) se puede establecer en 1. Después, la matriz de rotación y traslación $[R | T]$, que representa la posición y orientación de la cámara con respecto a la huella en la escena en el momento de la captura de imágenes, puede calcularse usando la ecuación (1).

En las realizaciones descritas anteriormente, tanto la longitud como el ancho del pie del usuario son conocidos y se usan para construir el modelo tridimensional del pie del usuario. Sin embargo, en otras realizaciones, podría ser que solo se conozca una dimensión, tal como la longitud, del pie de un usuario, y la otra dimensión del pie del usuario puede estimarse mediante el uso de una relación predeterminada entre la longitud y el ancho de un pie humano. La relación puede ser una relación empírica o un valor estadístico obtenido de la medición de las relaciones de longitud a anchura de un número de pies humanos.

Para mejorar la precisión en situaciones cuando el sistema de localización basado en la formación de imágenes puede no funcionar con precisión, por ejemplo, en una situación de hacinamiento donde muchas huellas de diferentes personas se superponen entre sí, el sistema puede ampliarse con otros sensores y métodos de rastreo. Por ejemplo, la pose obtenida por el sistema de huella térmica podría combinarse con al menos uno de una pose obtenida por un sistema de localización y un mapeo simultáneo (SLAM), la información de sensores inerciales, tales como acelerómetros, giroscopios y/o magnetómetros, y las posiciones estimadas a partir de señales de balizas de radio o de banda ultra ancha (UWB) o señales de comunicación inalámbricas existentes (por ejemplo, señales Wi-Fi) en el entorno.

Ya que se conoce la posición y orientación de la cámara con respecto al usuario, entonces es posible estimar la posición y orientación del usuario con respecto a una huella en una imagen en el momento de capturar la imagen. Por ejemplo, es posible proporcionar una transformación de coordenadas entre las coordenadas de la cámara y las coordenadas del usuario, de modo que la posición y orientación estimadas de la cámara desde el sistema de localización pueden transformarse en una posición y orientación estimadas para el usuario.

Rastreo de la posición de la cámara con respecto a una huella en la escena a lo largo del tiempo

En un primer punto en el tiempo, resolviendo el problema PnP como se ha descrito anteriormente, puede obtenerse la matriz de rotación y traslación $[R | T]$ para indicar la posición y orientación de la cámara con respecto a una huella particular en la escena. A medida que la cámara se mueve con el usuario a lo largo del tiempo, la posición y orientación de la cámara con respecto a la huella particular identificada en la escena también cambia. A continuación, posteriormente, en un segundo punto en el tiempo, puede obtenerse una segunda matriz de rotación y traslación $[R | T]$ para indicar la posición y orientación de la cámara con respecto a la huella particular en la escena. Si T_1 designa la matriz de rotación y traslación $[R | T]$ en el primer momento y T_2 designa la matriz de rotación y traslación $[R | T]$ en el segundo punto en el tiempo, entonces puede calcularse la matriz de transformación entre fotogramas $T_{i(1, 2)}$, que transforma la posición y orientación de la cámara con respecto a la huella particular en la escena en el primer punto en el tiempo T_1 a la posición y orientación de la cámara en el segundo momento en el tiempo T_2 , de la siguiente manera:

ya que

se deduce que

$$\begin{aligned} T_1 \times T_{i(1, 2)} &= T_2, \\ T_{i(1, 2)} &= T_1^{-1} \times T_2 \end{aligned} \quad (3)$$

A medida que el usuario y la cámara se mueven más a lo largo del tiempo, la matriz de rotación y traslación $[R | T]$ puede obtenerse en cada uno de los puntos posteriores en el tiempo, y una matriz de transformación entre fotogramas T_i puede obtenerse para cada par de fotogramas de imagen capturados en puntos consecutivos en el tiempo. Después la matriz de transformación total $T_{i-total}$ puede calcularse como:

$$T_{i-total} = T_{i(1, 2)} \times T_{i(2, 3)} \times \dots \times T_{i(n, n+1)} \quad (4)$$

donde $T_{i(n, n+1)}$ es la matriz de transformación entre fotogramas entre las imágenes capturadas en el n -ésimo y los $(n+1)$ -ésimos puntos en el tiempo.

Para determinar la transformación entre fotogramas entre imágenes capturadas en puntos consecutivos, se pueden usar métodos de flujo óptico para rastrear las huellas, de modo que la matriz de rotación y traslación $[R | T]$ puede obtenerse con respecto a la misma huella en imágenes capturadas en puntos consecutivos en el tiempo. El método Lucas-Kanade es un ejemplo de un método de flujo óptico para detectar y rastrear un objeto presente en múltiples imágenes tomadas en puntos consecutivos en el tiempo.

El proceso anterior puede usarse junto con la navegación a estima para rastrear la ubicación en el mundo real de un usuario a lo largo del tiempo. El rastreo puede inicializarse a partir de la ubicación en el mundo real inicial del usuario, que tiene una coordenada del mundo real conocida. La posición inicial y la orientación pueden obtenerse a partir de, por ejemplo, posicionamiento GPS, un punto caliente Wi-Fi cercano y/o una aplicación que podría ser utilizada por la persona rastreada para “verificar” con una ubicación conocida.

A medida que el usuario se aleja de la ubicación inicial, el usuario genera huellas. La ubicación instantánea real del usuario puede actualizarse mediante el uso de la posición y orientación relativas obtenidas a partir del sistema de localización de la huella como se ha descrito anteriormente. Más específicamente, esto implica capturar imágenes del suelo detrás del usuario en diferentes puntos en el tiempo, identificar las huellas de las imágenes, calcular la posición y orientación de la cámara con respecto a una huella en la imagen capturada en cada uno de los diferentes puntos en el tiempo, determinar una matriz de transformación entre fotogramas T_i para cada par de imágenes capturadas en puntos consecutivos en el tiempo y, por lo tanto, determinar la matriz de transformación total $T_{i-total}$ durante un período de tiempo según la realización descrita anteriormente.

Con el tiempo, los errores pueden acumularse en el sistema a medida que se usa la información posicional estimada de las huellas para actualizar la posición inicial conocida. Para anular errores, se pueden proporcionar correcciones regulares al sistema mediante GPS (cuando está disponible), comprobaciones de usuario u otra información de sensor de localización. Esto podría usarse para sobrescribir la información posicional del sistema de rastreo de huella o podría combinarse con la posición y orientación obtenidas a partir del sistema de localización de la huella mediante el uso de un algoritmo estadístico tal como un filtro Kalman.

En algunas realizaciones donde hay conocimiento existente del área (por ejemplo, está disponible un plano o mapa de piso de construcción), esta información podría usarse para proporcionar correcciones a la salida del sistema de localización de la huella térmica. Las restricciones conocidas sobre el movimiento, tales como las ubicaciones de las paredes en el área, podrían usarse para descartar ciertos movimientos y proporcionar ajustes si se devuelven movimientos improbables o imposibles del rastreador, tal como una persona que camina a través de una pared.

En situaciones donde la diferencia de calor entre la huella térmica del usuario y el entorno no es lo suficientemente significativa para un rastreo fiable, el sistema podría ampliarse opcionalmente a través de la inclusión de un dispositivo de ajuste de calor en el zapato del usuario que ajuste la firma térmica de sus pies. El dispositivo de ajuste de calor puede generar calor para proporcionar una huella térmica más pronunciada y/o para proporcionar una huella térmica en un patrón particular. Al proporcionar calor al zapato del usuario en un patrón particular, puede aplicarse una identidad a la huella térmica para permitir la identificación de huellas térmicas del usuario en una imagen rellena con muchas otras huellas térmicas, por ejemplo, en una situación cuando el usuario está en una multitud con otras personas.

En algunas realizaciones, se puede usar información de sensores inerciales, tales como acelerómetros, giroscopios y/o magnetómetros, para detectar cuándo el usuario ha detenido su movimiento, de modo que el sistema pueda mantener una posición detectada actual hasta que el usuario reanude el movimiento. Esto ayuda a reducir la sobrecarga de procesamiento en la localización de las huellas y puede conducir a resultados más precisos que continuar calculando nuevas estimaciones de pose de cámara, lo que puede introducir errores adicionales mientras el usuario no genera ninguna nueva huella.

Como el sistema de localización basado en la cámara descrita anteriormente depende de que haya una línea clara de visión (LOS) entre la cámara y las huellas del usuario, el sistema ocasionalmente puede sufrir una oclusión si la LOS entre la cámara y las huellas se bloquea temporalmente, por ejemplo, por otra persona que camina a través del campo de visión de la cámara. Opcionalmente, un uso de métodos de rastreo distintos de los que usan una cámara, como se han descrito anteriormente, permitiría la conmutación a uno o más de estos otros métodos, mientras que la obstrucción está presente. Por ejemplo, en el caso de oclusión entre la cámara y el suelo cuando se captura una imagen, el sistema puede conmutar a un método de rastreo basado en sensores inerciales.

Aunque la presente invención se ha descrito en relación con algunas realizaciones, no se pretende que esté limitada a la forma específica expuesta en la presente memoria. Más bien, el ámbito de la presente invención está limitado únicamente por las reivindicaciones adjuntas. Además, aunque pueda parecer que una característica se ha descrito en relación con realizaciones particulares, un experto en la técnica reconocerá que diversas características de las realizaciones descritas pueden combinarse según la invención. En las reivindicaciones, la expresión “que comprende” no excluye la presencia de otros elementos o etapas.

Además, el orden de las características en las reivindicaciones no implica ningún orden específico en el que deben realizarse las características y, en particular, el orden de etapas individuales en una reivindicación de método no implica que las etapas deban realizarse en este orden. Más bien, las etapas pueden realizarse en cualquier orden adecuado. Además, las referencias en singular no excluyen una pluralidad. Por lo tanto, las referencias a “uno”, “una”, “primero”, “segundo”, etc. no excluyen una pluralidad. En las reivindicaciones, la expresión “que comprende” o “que incluye” no excluye la presencia de otros elementos.

REIVINDICACIONES

1. Un sistema (100) que comprende
5 una cámara (106) para capturar una imagen,
al menos un procesador (202) y al menos una memoria (204), almacenando la al menos una memoria instrucciones configuradas para hacer que el procesador:
10 obtenga una imagen de una escena capturada usando la cámara, comprendiendo la escena una huella (104) de un usuario del sistema;
identifique la huella (104) en la imagen;
identifique una primera pluralidad de puntos de referencia, teniendo cada uno una posición fija con respecto a la huella (104) en la imagen;
15 identifique una segunda pluralidad de puntos de referencia, teniendo cada uno una posición fija con respecto a un modelo tridimensional del pie del usuario construido basándose en al menos una dimensión conocida del pie del usuario;
determine una matriz de transformación mapeando la segunda pluralidad de puntos de referencia del modelo tridimensional a la primera pluralidad de puntos de referencia; y
20 determine, basándose en la matriz de transformación, una posición tridimensional de la cámara (106) con respecto a la huella (104) en la escena en el momento de capturar la imagen.
2. El sistema de la reivindicación 1, en donde la al menos una memoria almacena instrucciones configuradas para hacer que el procesador determine una orientación tridimensional de la cámara con respecto a la huella en la escena en el momento de capturar la imagen.
25
3. El sistema de la reivindicación 1 o 2, en donde dicha imagen de la escena es una primera imagen de la escena capturada en un primer punto en el tiempo, en donde dicha matriz de transformación es una primera matriz de transformación, y en donde la al menos una memoria almacena instrucciones configuradas para hacer que el procesador, además:
30 obtenga una segunda imagen de la escena capturada usando la cámara en un segundo punto en el tiempo;
identifique la huella en la segunda imagen;
35 identifique una tercera pluralidad de puntos de referencia, teniendo cada uno una posición fija con respecto a la huella en la segunda imagen;
identifique una cuarta pluralidad de puntos de referencia teniendo cada uno una posición fija con respecto al modelo tridimensional del pie del usuario construido basándose en la al menos una dimensión conocida del pie del usuario; y
40 determine una segunda matriz de transformación mapeando la cuarta pluralidad de puntos de referencia del modelo tridimensional a la tercera pluralidad de puntos de referencia en el plano de la segunda imagen.
4. El sistema de la reivindicación 3, en donde la posición tridimensional de la cámara es una primera posición tridimensional de la cámara con respecto a la huella en la escena en el primer punto en el tiempo, y en donde la al menos una memoria almacena instrucciones configuradas para hacer que el procesador, además:
45 determine una matriz de transformación entre fotogramas basándose en la primera matriz de transformación y la segunda matriz de transformación, y
50 determine, basándose en la matriz de transformación entre fotogramas y la primera posición tridimensional de la cámara con respecto a la huella en la escena, una segunda posición tridimensional de la cámara con respecto a la huella en la escena en el segundo punto en el tiempo.
5. El sistema de la reivindicación 4, en donde la al menos una memoria almacena instrucciones configuradas para hacer que el procesador calcule una ubicación en el mundo real actualizada de la cámara basándose en una ubicación en el mundo real inicial de la cámara y las posiciones tridimensionales determinadas primera y segunda de la cámara con respecto a la huella en la escena en el momento de capturar la imagen.
55
6. El sistema de la reivindicación 5, en donde la al menos una memoria almacena instrucciones configuradas para hacer que el procesador calcule la ubicación del mundo real actualizada de la cámara basándose, además, en al menos uno de: tener conocimiento del área donde se ubica el usuario, un sistema de localización y mapeo simultáneo (SLAM), al menos un sensor inercial, un sensor magnético, balizas que emiten señales, señales de comunicación inalámbrica, señales de oportunidad, posición designada por el usuario y señales GPS.
60
65
7. El sistema de cualquier reivindicación anterior, en donde la cámara es monocular.

8. El sistema de cualquier reivindicación anterior, en donde la al menos una memoria almacena instrucciones configuradas para hacer que el procesador determine una dirección de la al menos una huella en la imagen determinando un eje longitudinal de la al menos una huella.
9. El sistema de cualquier reivindicación anterior, en donde la cámara es una cámara de formación de imágenes térmicas, y en donde la imagen es una imagen térmica.
10. El sistema de la reivindicación 9, en donde el sistema comprende además un dispositivo de ajuste de calor en el zapato del usuario para ajustar una huella térmica del usuario.
11. El sistema de cualquiera de las reivindicaciones 1-8, en donde la cámara es una cámara para capturar luz visible.
12. El sistema de cualquier reivindicación anterior, en donde la al menos una dimensión conocida del pie del usuario comprende al menos la longitud o el ancho del pie del usuario.
13. El sistema de cualquier reivindicación anterior, que comprende además una pantalla montada en la cabeza, a la que está acoplada operativamente la cámara, o un casco, en el que se configura la cámara para colocarse.
14. Un método implementado por ordenador, que comprende:
 - recibir una representación digital de una imagen de una escena capturada usando una cámara, comprendiendo la escena una huella de un usuario;
 - identificar la huella en la imagen;
 - identificar una primera pluralidad de puntos de referencia, teniendo cada uno una posición fija con respecto a la huella en la imagen;
 - identificar una segunda pluralidad de puntos de referencia, teniendo cada uno una posición fija con respecto a un modelo tridimensional del pie del usuario construido basándose en al menos una dimensión conocida del pie del usuario;
 - determinar una matriz de transformación mapeando la segunda pluralidad de puntos de referencia del modelo tridimensional a la primera pluralidad de puntos de referencia; y
 - determinar, basándose en la matriz de transformación, una posición tridimensional de la cámara con respecto a la huella en la escena en el momento de capturar la imagen.
15. El método de la reivindicación 14, que comprende además determinar una orientación tridimensional de la cámara con respecto a la huella en la escena en el momento de capturar la imagen.
16. El método de la reivindicación 14 o 15, en donde dicha imagen de la escena es una primera imagen de la escena capturada en un primer punto en el tiempo, en donde dicha matriz de transformación es una primera matriz de transformación, y en donde el método comprende además:
 - obtener una segunda imagen de la escena capturada usando la cámara en un segundo punto en el tiempo;
 - identificar la huella en la segunda imagen;
 - identificar una tercera pluralidad de puntos de referencia, teniendo cada uno una posición fija con respecto a la huella en la segunda imagen;
 - identificar una cuarta pluralidad de puntos de referencia, teniendo cada uno una posición fija con respecto al modelo tridimensional del pie del usuario construido basándose en la al menos una dimensión conocida del pie del usuario; y
 - determinar una segunda matriz de transformación mapeando la cuarta pluralidad de puntos de referencia del modelo tridimensional a la tercera pluralidad de puntos de referencia en el plano de la segunda imagen.
17. El método de la reivindicación 16, en donde la posición tridimensional de la cámara es una primera posición tridimensional de la cámara en relación con la huella en la escena en el primer punto en el tiempo, y en donde el método comprende además: determinar una matriz de transformación entre fotogramas basándose en la primera matriz de transformación y la segunda matriz de transformación, y
 - determinar, basándose en la matriz de transformación entre fotogramas y la primera posición tridimensional de la cámara con respecto a la huella en la escena, una segunda posición tridimensional de la cámara con respecto a la huella en la escena en el segundo punto en el tiempo.
18. El método de la reivindicación 17, que comprende además calcular una ubicación en el mundo real actualizada de la cámara basándose en una ubicación en el mundo real inicial de la cámara y las posiciones

tridimensionales determinadas primera y segunda de la cámara con respecto a la huella en la escena en el momento de capturar la imagen.

- 5 19. Un medio de almacenamiento legible por ordenador que comprende instrucciones que, cuando son ejecutadas por un ordenador, hacen que el ordenador lleve a cabo el método de cualquiera de las reivindicaciones 14-18.

Figura 1

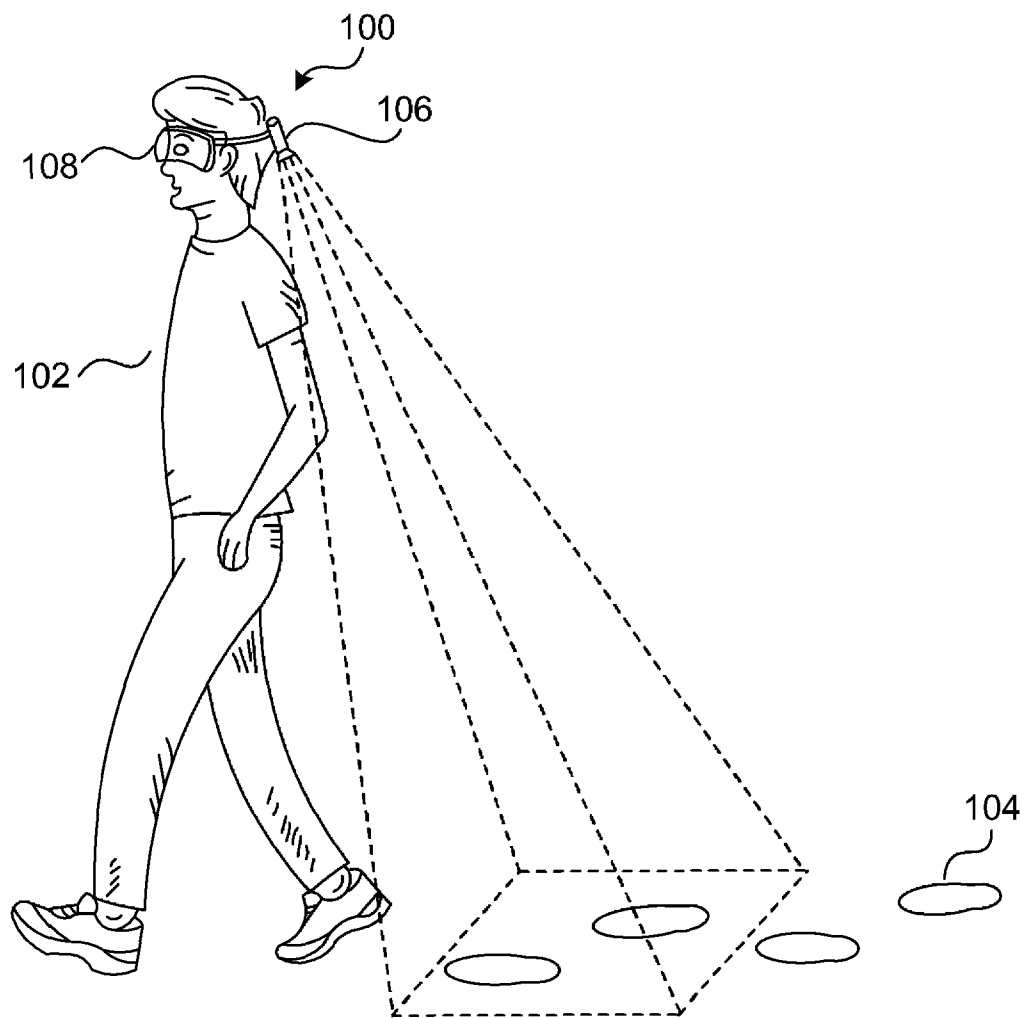
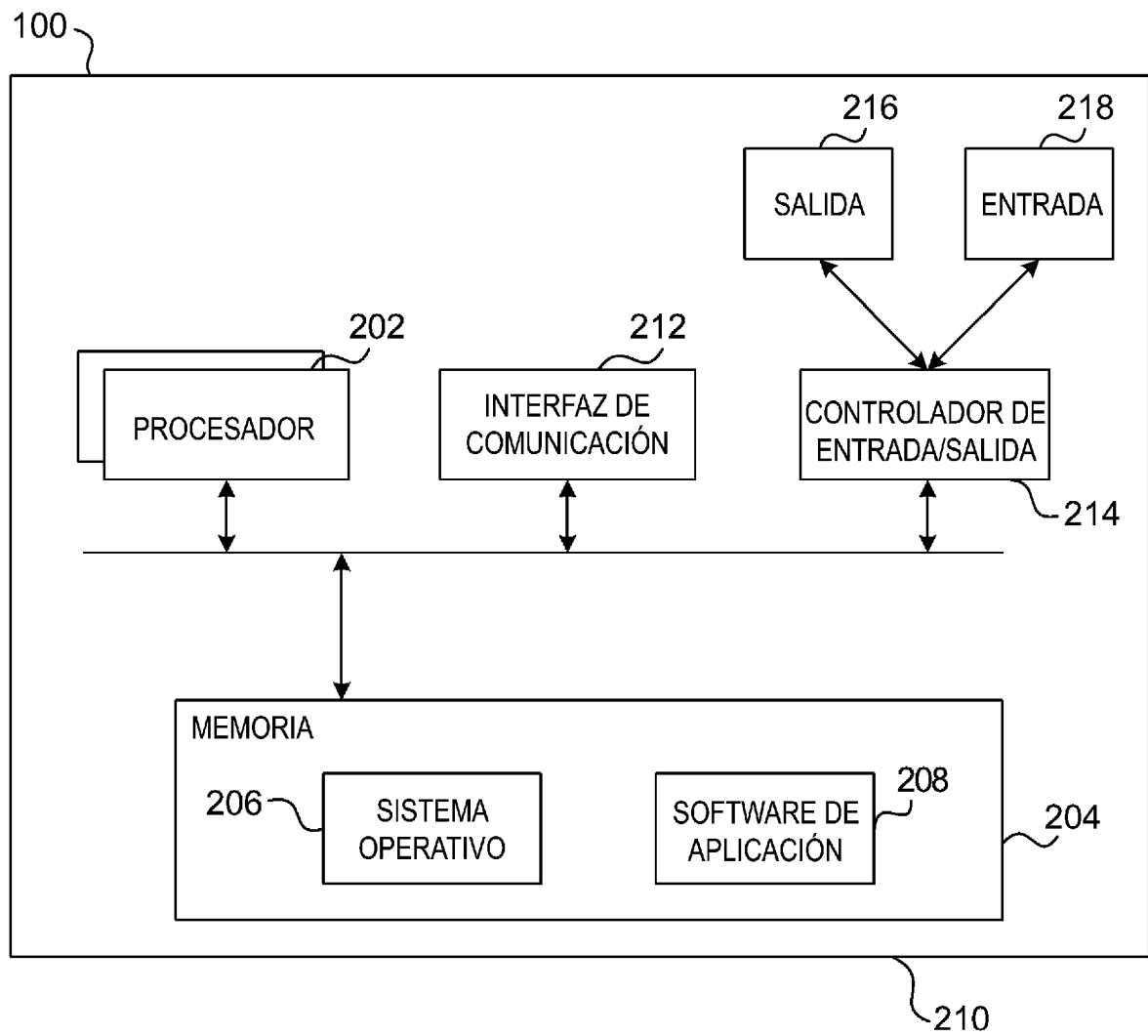


Figura 2



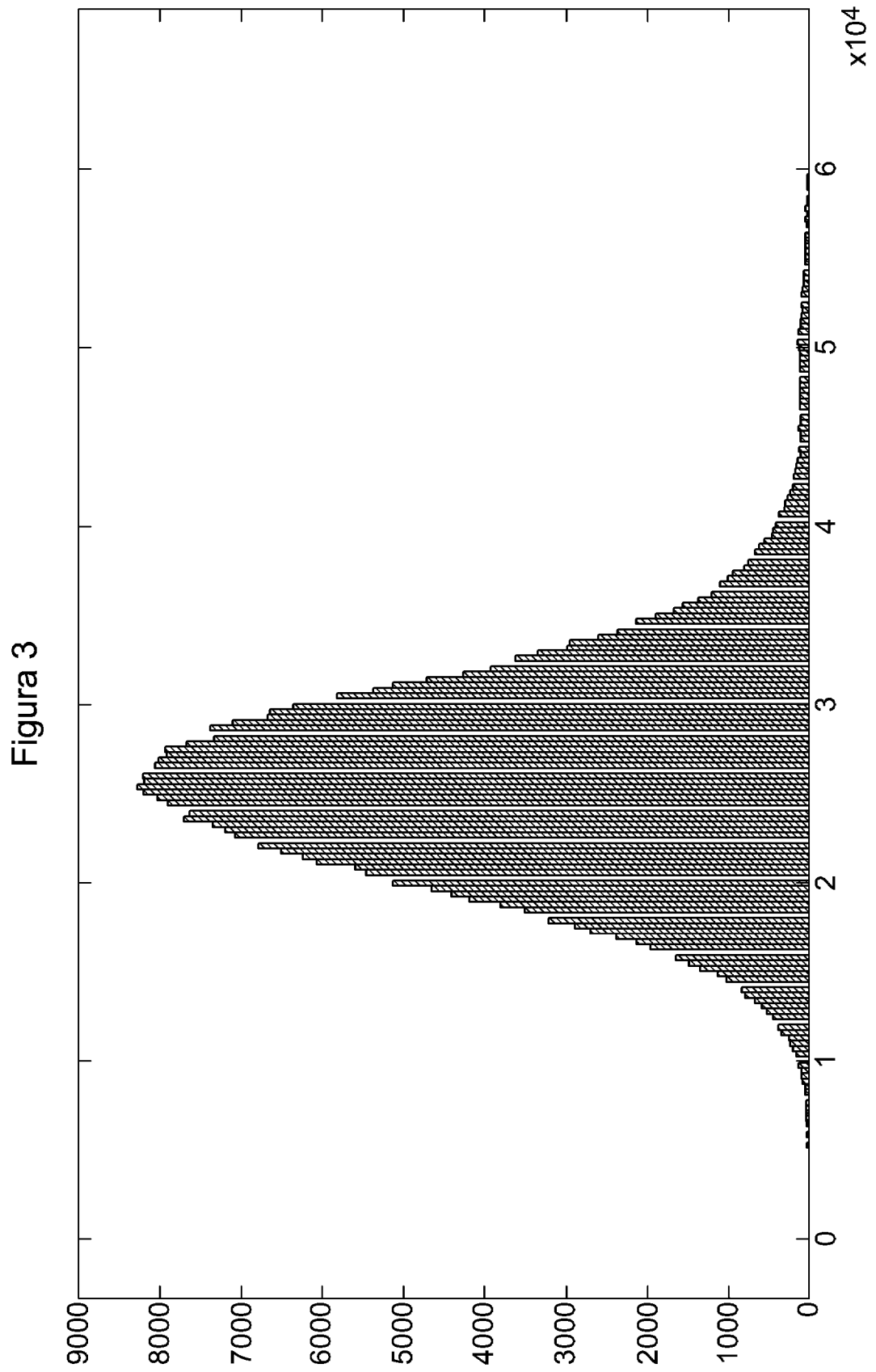


Figura 4

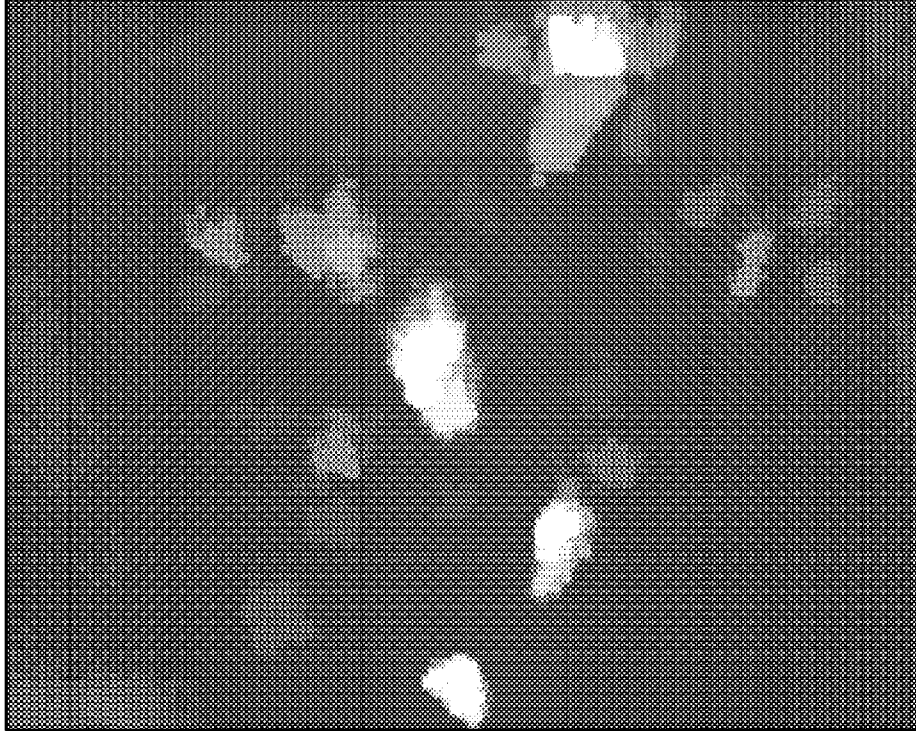


Figura 5

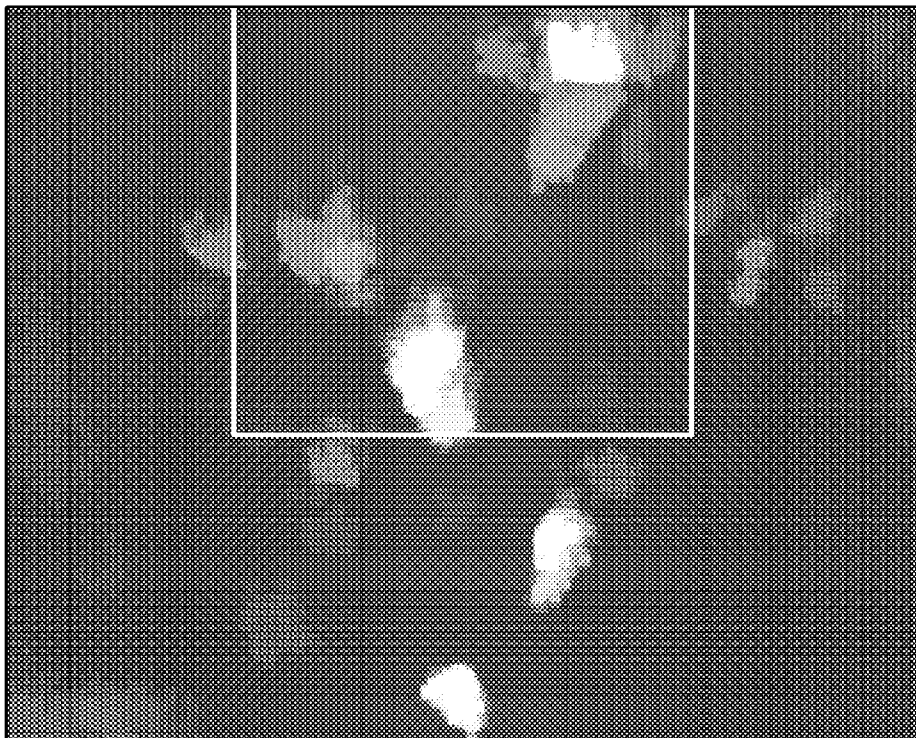


Figura 6

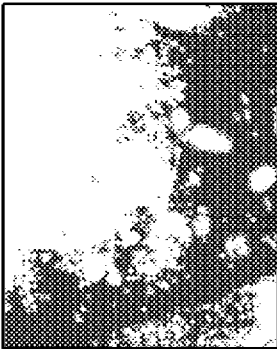
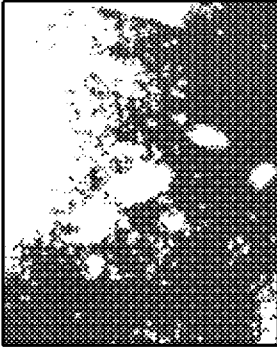
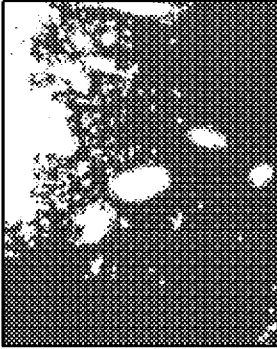
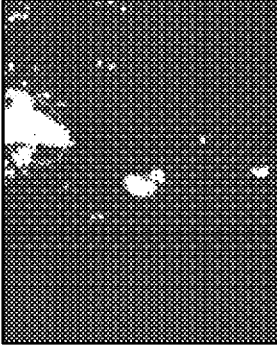
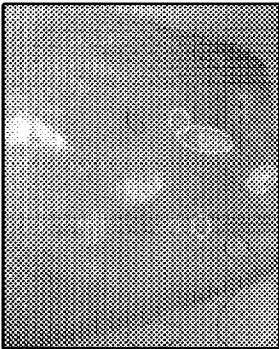
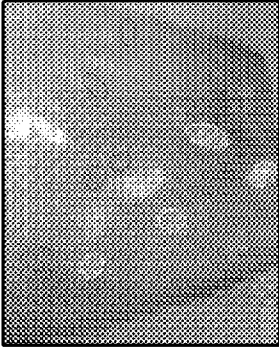
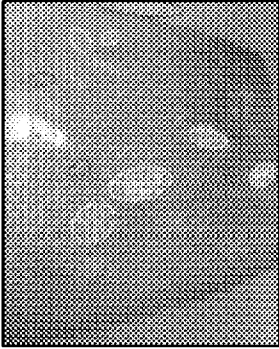
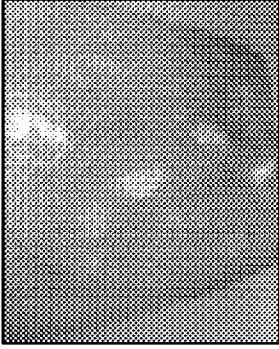
Tipo	Umbral 1	Umbral 2	Umbral 3	Umbral 4
Umbral				
Imagen Original				

Figura 7

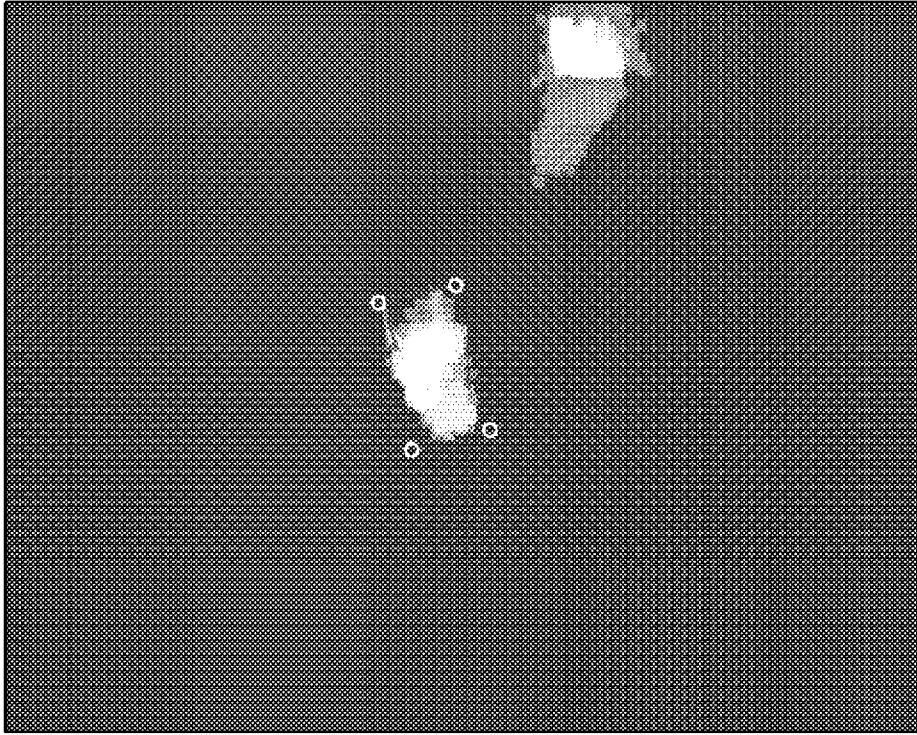


Figura 8

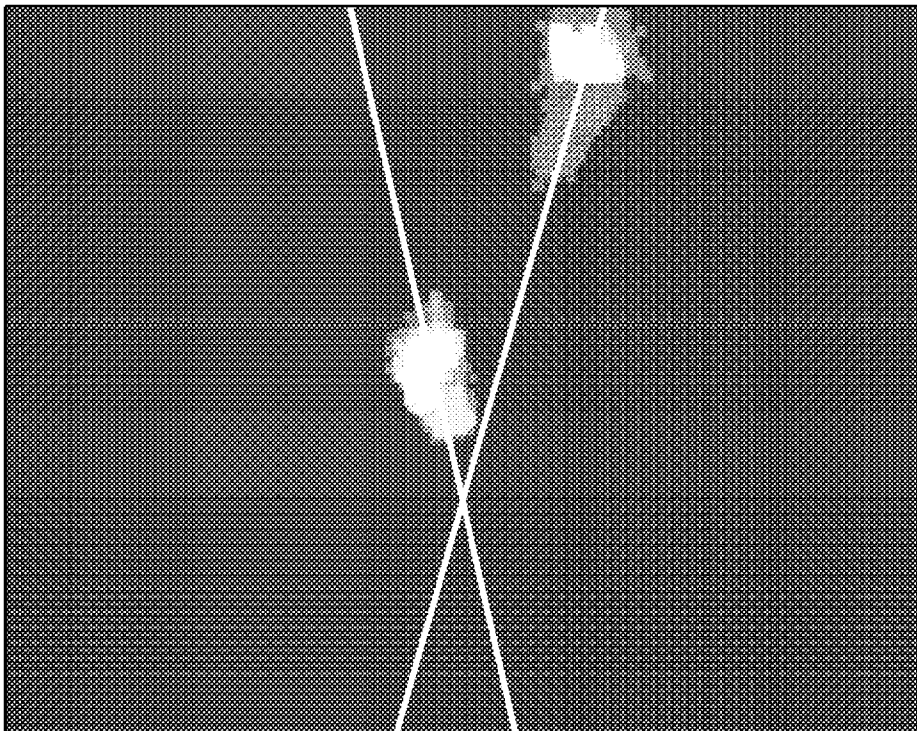


Figura 9

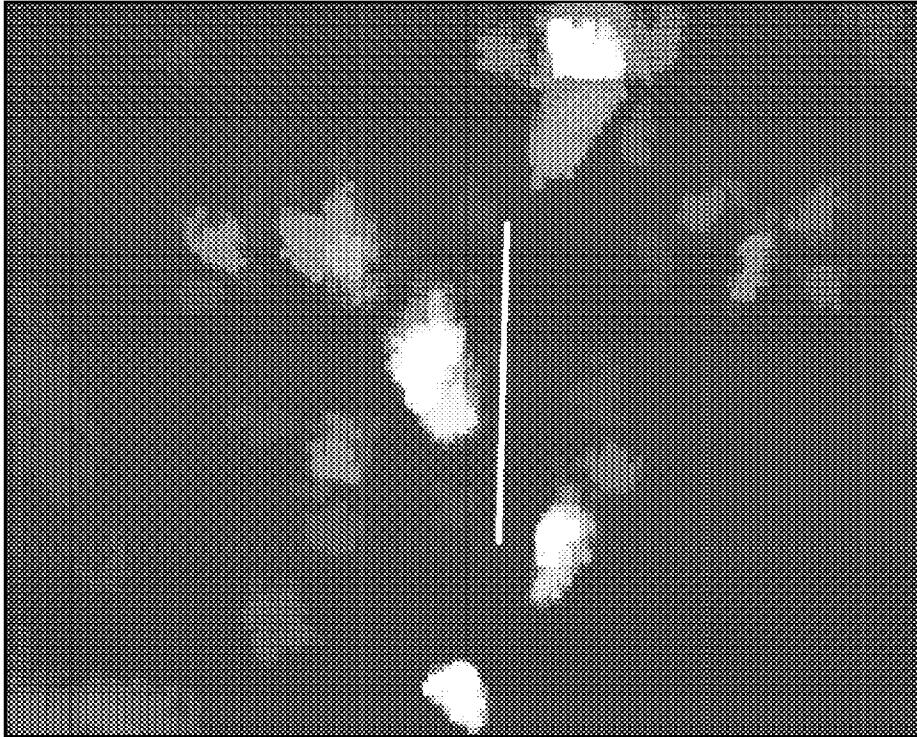


Figura 10

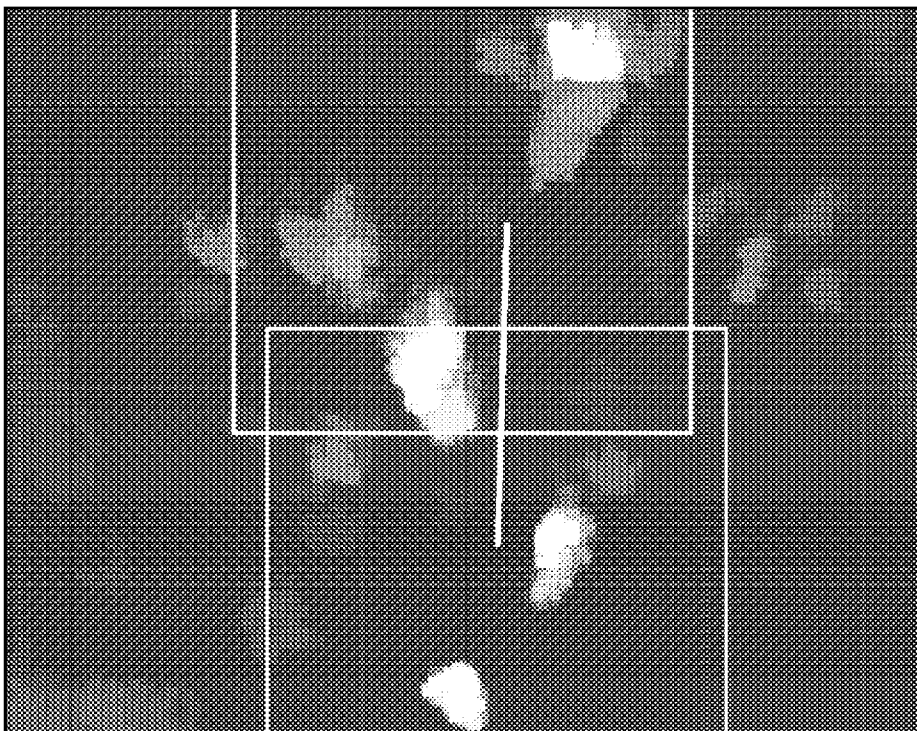


Figura 11

