

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 996 894**

51 Int. Cl.:

**H04N 13/111** (2008.01)  
**G06T 7/593** (2007.01)  
**G06T 17/05** (2011.01)  
**H04N 13/239** (2008.01)  
**B64C 39/02** (2013.01)  
**G01C 11/02** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **22.07.2019** **PCT/SE2019/050707**

87 Fecha y número de publicación internacional: **27.02.2020** **WO20040679**

96 Fecha de presentación y número de la solicitud europea: **22.07.2019** **E 19850915 (0)**

97 Fecha y número de publicación de la concesión europea: **07.08.2024** **EP 3841744**

54 Título: **Un método y sistema correspondiente para generar modelos basados en vídeo de un objetivo tal como un evento dinámico**

30 Prioridad:

**22.08.2018 US 201862720982 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**13.02.2025**

73 Titular/es:

**I-CONIC VISION AB (100.00%)**  
**Drottninggatan 33**  
**111 51 Stockholm, SE**

72 Inventor/es:

**HÅKAN, WIMAN y**  
**MIKAEL, STERN**

74 Agente/Representante:

**CARVAJAL Y URQUIJO, Isabel**

ES 2 996 894 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 método y sistema correspondiente para generar modelos basados en vídeo de un objetivo tal como un evento dinámico

## CAMPO TÉCNICO

La tecnología propuesta se relaciona en general con un método y un sistema correspondiente para generar uno o más modelos basados en vídeo de un objetivo tal como un evento dinámico, un sistema de sensor dinámico, un sistema terrestre configurado para generar uno o más modelos basados en vídeo de un objetivo, así como un programa informático correspondiente y producto de programa informático.

## ANTECEDENTES

Los vídeos tomados desde vehículos en movimiento o que pueden moverse, como vehículos aéreos y/o vehículos de superficie y/o vehículos submarinos, incluidos vehículos tripulados o no tripulados, como drones o vehículos aéreos no tripulados (UAV), pueden proporcionar hoy en día a las autoridades de emergencia y seguridad y a las fuerzas armadas apoyo para tomar decisiones rápidas en situaciones de crisis. Estas situaciones pueden incluir, por ejemplo, incendios, accidentes, inundaciones y/u otras operaciones de seguridad o militares y similares. Otros escenarios objetivo pueden incluir eventos deportivos, producciones cinematográficas y similares.

Sin embargo, existe una necesidad general de modelos mejorados basados en vídeo de escenarios objetivo, como eventos dinámicos, y/o soporte de decisiones mejorado basado en dichos modelos, por ejemplo, para ampliar los modelos 3D basados en imágenes para incluir también eventos dinámicos para una mejor visualización, interacción del usuario y/o soporte de decisiones.

El documento EP 2 728 308 A2 se refiere a la fotogrametría aérea. La divulgación se refiere específicamente a un sistema fotogramétrico aéreo que utiliza dos o más vehículos voladores, cada uno de ellos equipado con un dispositivo GPS y una unidad de captación de imágenes. El procedimiento principal implica una etapa de configuración de dos o más puntos de fotografía y de configuración de un área de puntos de fotografía respectivamente con cada uno de los puntos de fotografía como centro, una etapa de medición de una posición del vehículo volador por dicho dispositivo GPS, una etapa donde cada uno de los vehículos voladores alcanza cada área de puntos de fotografía correspondiente y mantiene la posición del área de puntos de fotografía, una etapa de adquisición de un tiempo cuando el vehículo volador finalmente alcanza el área de puntos de fotografía, una etapa de configuración de un tiempo de sincronización del obturador después de un tiempo predeterminado a partir del momento cuando el vehículo volador finalmente ha alcanzado el área de puntos de fotografía, y una etapa de toma de fotografías aéreas por los dos o más vehículos voladores en el tiempo de sincronización del obturador.

El documento "Numerisation conjointe image-laser pour la modelisation 3D des environnements complexes ou habites" de Craciun, Daniela, 7 de julio de 2012 se refiere a la fusión de imagen-láser para el modelado 3D de entornos complejos.

El documento evalúa el potencial de un sistema híbrido de imagen-láser para generar modelos 3D completos y fotorrealistas in situ de entornos desafiantes, sin requerir la intervención de un operador humano. Este documento se centra específicamente en dos aspectos principales: (i) la automatización del proceso de modelado 3D, que apunta a la correspondencia automática de datos en áreas sin características y sin GPS para el modelado del mundo in situ y (ii) la explotación de los modelos 3D generados junto con procedimientos de servicio visual para proporcionar sistemas móviles con capacidades autónomas de digitalización y exploración del sitio, y menciona brevemente una fusión de imagen-láser complementaria y cooperativa que conduce a un prototipo de sensor de mosaico 4D.

## SUMARIO

Un objeto es proporcionar un método para generar uno o más modelos basados en vídeo de un objetivo según la reivindicación 1.

También es un objeto proporcionar un sistema configurado para generar uno o más modelos basados en vídeo de un objetivo según la reivindicación 10.

Otro objeto más es proporcionar un sistema terrestre configurado para generar uno o más modelos basados en vídeo de un objetivo según la reivindicación 13.

Otro objeto más es proporcionar un programa informático para generar, cuando es ejecutado por una pluralidad de procesadores, uno o más modelos basados en vídeo de un objetivo según la reivindicación 14.

Según un primer aspecto, se proporciona un método para generar uno o más modelos basados en vídeo de un objetivo. El método comprende:

- 5                   • proporcionar transmisiones de vídeo desde al menos dos vehículos en movimiento o que pueden moverse equipados con cámaras para obtener imágenes simultáneamente del objetivo desde diferentes puntos de vista,
- 10                   o en donde se proporciona una sincronización de posición de los vehículos en movimiento o que pueden moverse para crear una base de imagen estable, que representa la distancia entre los vehículos en movimiento o que pueden moverse,
- 15                   o en donde se proporciona una sincronización de apuntamiento de las cámaras para cubrir el (los) mismo(s) objeto(s) y/o evento(s) dinámico(s), y
- 20                   o en donde se proporciona una sincronización temporal de los fotogramas de vídeo de las transmisiones de vídeo para obtener, para al menos un punto en el tiempo, un conjunto de uno o más fotogramas de vídeo registrados simultáneamente; y
- 20                   • generar, para dicho al menos un punto en el tiempo, al menos un modelo tridimensional, 3D, del objetivo basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente.

25 Según un segundo aspecto, se proporciona un sistema configurado para generar uno o más modelos basados en vídeo de un objetivo. El sistema está configurado para proporcionar transmisiones de vídeo desde al menos dos vehículos en movimiento o que pueden moverse equipados con cámaras para obtener imágenes simultáneamente del objetivo desde diferentes puntos de vista. Se proporciona una sincronización de posición de los vehículos en movimiento o que pueden moverse para crear una base de imagen estable. Se proporciona una sincronización de apuntamiento de las cámaras para cubrir el (los) mismo(s) objeto(s) y/o evento(s) dinámico(s), y se habilita la sincronización temporal de los fotogramas de vídeo de las transmisiones de vídeo para obtener, para al menos un punto en el tiempo, un conjunto de fotogramas de vídeo registrados simultáneamente. El sistema también está configurado para generar, para dicho al menos un punto en el tiempo, al menos un modelo tridimensional, 3D, del objetivo basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente.

35 Según un tercer aspecto, se proporciona un sistema de sensor dinámico. El sistema de sensor dinámico comprende un conjunto de vehículos en movimiento o que pueden moverse que funcionan sincrónicamente, equipados con cámaras para permitir transmisiones de vídeo para obtener imágenes simultáneamente de un objetivo desde diferentes puntos de vista. El sistema de sensor dinámico está configurado para funcionar basándose en la sincronización de posición de los vehículos en movimiento o que pueden moverse para crear una base de imagen estable, y el sistema de sensor dinámico está configurado para funcionar basándose en la sincronización de apuntamiento de las cámaras para cubrir el (los) mismo(s) objeto(s) y/o evento(s) dinámico(s). El sistema de sensor dinámico también está configurado para permitir la sincronización temporal de los fotogramas de vídeo de las transmisiones de vídeo para permitir, para al menos un punto en el tiempo, la generación y/o extracción de un conjunto de fotogramas de vídeo registrados simultáneamente. El sistema de sensor dinámico está configurado además para transmitir datos de imagen de las transmisiones de vídeo a un segmento de tierra desde al menos uno de los vehículos en movimiento o que pueden moverse para permitir la generación de uno o más modelos tridimensionales/tetradimensionales, 3D/4D, del objetivo basados en los fotogramas de vídeo registrados simultáneamente en el segmento de tierra.

50 Según un cuarto aspecto, se proporciona un sistema terrestre configurado para generar uno o más modelos basados en vídeo de un objetivo. El sistema está configurado para recibir transmisiones de vídeo desde al menos dos vehículos en movimiento o que pueden moverse equipados con cámaras para obtener imágenes simultáneamente del objetivo desde diferentes puntos de vista. Los vehículos en movimiento o que pueden moverse se hacen funcionar basándose en la sincronización de posición para una base de imagen estable, y las cámaras se hacen funcionar basándose en la sincronización de apuntamiento para cubrir el (los) mismo(s) objeto(s) y/o evento(s) dinámico(s), y la sincronización de tiempo de los fotogramas de vídeo de las transmisiones de vídeo se habilita para permitir, para al menos un punto en el tiempo, la generación y/o extracción de un conjunto de fotogramas de vídeo registrados simultáneamente. El sistema también está configurado para generar, para dicho al menos un punto en el tiempo, al menos un modelo tridimensional, 3D, del objetivo basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente.

60 Según un quinto aspecto, se proporciona un programa informático para generar, cuando es ejecutado por un procesador, uno o más modelos basados en vídeo de un objetivo. El programa informático comprende instrucciones que, cuando son ejecutadas por el procesador, hacen que el procesador:

- 65                   • reciba transmisiones de vídeo de al menos dos vehículos en movimiento o que pueden moverse equipados con cámaras para obtener imágenes simultáneamente del objetivo desde diferentes puntos de vista,

o los vehículos en movimiento o que pueden moverse se operan basándose en la sincronización de posición para una base de imagen estable,

o las cámaras se operan basándose en la sincronización de apuntamiento para cubrir el (los) mismo(s) objeto(s) y/o evento(s) dinámico(s), y

o se proporciona una sincronización temporal de los fotogramas de vídeo de las transmisiones de vídeo para obtener, para al menos un punto en el tiempo, un conjunto de fotogramas de vídeo registrados simultáneamente; y

- genere, para dicho al menos un punto en el tiempo, al menos un modelo tridimensional, 3D, o modelo de cuatro dimensiones, 4D, del objetivo basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente.

Según un sexto aspecto, se proporciona un producto de programa informático que comprende un medio legible por ordenador no transitorio que contiene el programa informático mencionado anteriormente.

Otras ventajas de la invención se apreciarán al leer la descripción detallada a continuación.

## BREVE DESCRIPCIÓN DE LOS DIBUJOS

La invención, junto con otros objetos y ventajas de la misma, se puede entender mejor haciendo referencia a la siguiente descripción tomada junto con los dibujos adjuntos, en los que:

La FIG. 1 es un diagrama esquemático que ilustra un ejemplo de una vista general del sistema que incluye componentes básicos del sistema.

La FIG. 2 es un diagrama esquemático que ilustra un ejemplo de una vista general del sistema que incluye componentes básicos del sistema y funciones básicas.

La FIG. 3 es un diagrama esquemático que ilustra un ejemplo de un sistema de sensor dinámico, aquí ejemplificado como un segmento aéreo basado en drones equipados con cámaras a bordo.

La FIG. 4 es un diagrama esquemático que ilustra un ejemplo de cómo se pueden generar modelos 3D momentáneos, por ejemplo, a partir de pares de fotogramas de vídeo sincronizados en el tiempo o, más generalmente, n-tuplas para visualización y exploración.

La FIG. 5 es un diagrama esquemático que ilustra un ejemplo de cómo los modelos 3D momentáneos juntos pueden usarse para construir modelos 4D casi en tiempo real para visualización y exploración.

La FIG. 6 es un diagrama esquemático que ilustra un ejemplo de una implementación informática según una realización.

La FIG. 7 es un diagrama de flujo esquemático que ilustra un ejemplo de un método para generar uno o más modelos basados en vídeo de un objetivo.

## DESCRIPCIÓN DETALLADA

Como se ha mencionado, los vídeos tomados desde vehículos en movimiento o que pueden moverse, como vehículos aéreos y/o vehículos de superficie (por ejemplo, vehículos acuáticos y/o terrestres) y/o vehículos submarinos, incluidos vehículos tripulados y/o no tripulados, como drones y/o vehículos aéreos no tripulados (UAV), ya hoy pueden proporcionar a las autoridades de emergencia y seguridad y a las fuerzas armadas apoyo para tomar decisiones rápidas en situaciones de crisis. La innovación supone un gran paso adelante en el soporte de decisiones, ya que permite la creación de modelos 3D en tiempo real, así como modelos 4D, es decir, una serie de modelos 3D dependientes del tiempo.

La tecnología propuesta también es aplicable para proporcionar modelos basados en vídeo de otros eventos, como eventos deportivos, y para producciones cinematográficas y similares.

Los modelos avanzados de la presente invención no sólo permiten la visualización de objetos fijos, sino también de eventos dinámicos como vehículos en movimiento, columnas de humo, inundaciones, accidentes y multitudes, utilizando fotogrametría. Esto le da al usuario la oportunidad de estudiar eventos dinámicos de la misma manera que estudiamos modelos 3D estáticos, por ejemplo, desde los ángulos y con el zoom deseado por el usuario. La innovación también permite realizar mediciones manuales y automáticas de las dimensiones (por ejemplo, volumen, área, longitud, altura), direcciones y velocidad de objetos en movimiento.

Los objetos en movimiento, por ejemplo, los vehículos, pueden *sin embargo* sólo reconstruirse en tres dimensiones cuando se hayan obtenido imágenes de los mismos desde diferentes puntos de vista en al menos dos imágenes simultáneamente. Al utilizar dos o más imágenes en una sola secuencia de vídeo o imagen fija de un dron, el objeto en movimiento habrá cambiado de posición y/o forma, lo que deshabilitará la reconstrucción 3D mediante técnicas fotogramétricas y de visión por ordenador. Esta innovación hace uso de dos o más drones o vehículos en movimiento o que pueden moverse similares con transmisiones de vídeo sincronizadas o al menos sincronizables para permitir la obtención de imágenes simultánea desde diferentes puntos de vista y, por tanto, la reconstrucción en 3D de un objetivo. Por ejemplo, cada par o, más generalmente,  $m$ -tuple ( $m \geq 2$ , donde  $m$  representa el número de cámaras) de imágenes o fotogramas de vídeo registrados simultáneamente permite la generación de un modelo 3D para un punto particular en el tiempo,  $t_i$ . La unión de todos los modelos 3D para un intervalo de tiempo  $t_0$  a  $t_n$ , por ejemplo, la duración de las grabaciones de vídeo simultáneas o un subconjunto seleccionado de las mismas, constituyen un modelo 4D. Aquí,  $n$  representa el número de puntos en el tiempo. En un modelo 4D, el usuario puede navegar tanto en las tres dimensiones espaciales como en la dimensión temporal.

Para simplificar, a partir de ahora se utilizará el término vídeo para representar cualquier tipo de secuencia de imágenes.

Para simplificar, a partir de ahora se utilizará el término dron para representar cualquier vehículo adecuado en movimiento o que puede moverse de manera controlable desde el cual se puedan tomar secuencias de vídeo de eventos y utilizarlas para modelado 3D/4D.

Sin embargo, debe entenderse que la presente innovación puede utilizar otros tipos de vehículos, incluidos vehículos aéreos y/o de superficie, así como vehículos tripulados y/o no tripulados. También debe entenderse que los vehículos pueden llevar pasajeros y/o incluso conductores y aun así ser más o menos autónomos, como los coches de autoconducción. En realizaciones particulares, se pueden utilizar vehículos no tripulados tales como drones. En algunas aplicaciones, puede ser deseable utilizar vehículos aéreos no tripulados, mientras que en otras aplicaciones pueden preferirse vehículos de superficie no tripulados. El término altura del terreno se interpretará como distancia del objeto para los vehículos y aplicaciones aplicables, por ejemplo, para vehículos que no son aéreos.

También debe entenderse que los vehículos en movimiento o que pueden moverse pueden estar realizando movimientos correctivos y/o adaptativos muy pequeños para mantener una determinada posición, al menos durante un período de tiempo determinado, por ejemplo, cuando los drones están sobrevolando temporalmente un área objetivo.

El término objetivo se utilizará para representar cualquier área objetivo general y/o situación de interés.

A modo de ejemplo, el objetivo puede representar uno o más eventos dinámicos e incluir uno o más objetos en movimiento. Esto representa una aplicación particularmente interesante.

Sin embargo, debe entenderse que, además de modelar eventos dinámicos, la tecnología propuesta también puede ser útil para permitir una cobertura más rápida de un área geográfica determinada y/o una precisión mejorada para el procesamiento en tiempo real.

Según un primer aspecto, se proporciona un método para generar uno o más modelos basados en vídeo de un objetivo. Por ejemplo, se puede hacer referencia al diagrama de flujo esquemático de la FIG. 7.

Básicamente el método comprende:

S1: proporcionar transmisiones de vídeo sincronizadas desde al menos dos vehículos en movimiento o que pueden moverse equipados con cámaras para obtener imágenes simultáneamente del objetivo desde diferentes puntos de vista,

- en donde se proporciona una sincronización de posición de los vehículos en movimiento o que pueden moverse para crear una base de imagen estable, que representa la distancia entre los vehículos en movimiento o que pueden moverse,

- en donde se proporciona una sincronización de apuntamiento de las cámaras para cubrir el (los) mismo(s) objeto(s) y/o evento(s) dinámico(s), y

- en donde se proporciona una sincronización temporal de los fotogramas de vídeo de los flujos de vídeo para obtener, para al menos un punto en el tiempo, un conjunto de fotogramas de vídeo registrados simultáneamente; y

S2: generar, para dicho al menos un punto en el tiempo, al menos un modelo tridimensional, 3D, del objetivo basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente.

En un ejemplo particular, se proporciona una sincronización temporal de los fotogramas de vídeo de las transmisiones de vídeo para obtener, para cada uno de una pluralidad de puntos en el tiempo, un conjunto de fotogramas de vídeo registrados simultáneamente. En este caso, puede generarse al menos un modelo 3D del objetivo, para cada uno de un conjunto seleccionado de la pluralidad de puntos en el tiempo, basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente.

En consecuencia, el método puede incluir la siguiente etapa opcional:

S3: combinar al menos un subconjunto de los modelos 3D generados para generar un modelo de cuatro dimensiones, 4D, del objetivo para permitir que un usuario navegue a través del modelo 4D en las tres dimensiones espaciales y en la dimensión temporal.

En general, los vehículos en movimiento o que pueden moverse representan cualquier vehículo en movimiento o que puede moverse de manera controlable desde el cual se pueden tomar secuencias de vídeo de eventos y utilizarlas para modelado 3D/4D.

Según la invención, los datos de imagen basados en las transmisiones de vídeo se transmiten por enlace descendente a un segmento de tierra desde al menos uno de los vehículos en movimiento o que pueden moverse y/o se opera un enlace ascendente de control para controlar los vehículos en movimiento o que pueden moverse y/o las cámaras desde el segmento de tierra.

Por ejemplo, los vehículos en movimiento o que pueden moverse pueden incluir vehículos aéreos y/o vehículos de superficie, incluidos vehículos tripulados o no tripulados.

En un ejemplo particular, los vehículos en movimiento o que pueden moverse incluyen drones o vehículos aéreos no tripulados (UAV), cada uno de los cuales tiene una cámara a bordo.

A modo de ejemplo, el vuelo sincrónico de un dron, incluido el apuntamiento de la cámara, puede ser controlado desde un segmento de tierra por un oficial de primera línea, utilizando parámetros preestablecidos para el vuelo, o el vuelo puede ser controlado de manera autónoma.

En un ejemplo particular, los drones están configurados para volar uno al lado del otro, cubriendo un área de interés a lo largo de la pista. En otro ejemplo, los drones están configurados para volar uno tras otro a lo largo de una pista. En otro ejemplo, los drones están configurados para pasar sobre un área de interés.

En general, la sincronización de posicionamiento y/o apuntamiento corresponden a la orientación relativa. A modo de ejemplo, la orientación relativa se puede calcular de manera continua en tiempo real basándose en el contenido de las transmisiones de vídeo y se puede usar para controlar los drones y/o las cámaras de forma que obtengan y/o mantengan la sincronización de posicionamiento y/o apuntamiento deseada.

Por ejemplo, se puede implementar un control de vuelo más o menos autónomo utilizando las transmisiones de vídeo transmitidos por enlace descendente para determinar y ajustar las posiciones y actitudes relativas de los drones. En consecuencia, se pueden generar instrucciones y/o comandos correspondientes y transmitirlos por enlace ascendente desde una estación terrestre a los drones para mantener la base y las sincronizaciones de posicionamiento y/o apuntamiento relevantes.

Por ejemplo, la sincronización de posición y/o la sincronización de apuntamiento puede realizarse mediante un enlace de radio entre los drones, directo o a través de una estación de radio terrestre, en la que se pueden comunicar las posiciones de los drones de manera que se pueda mantener una relación base-altura correcta, o la sincronización de posición puede realizarse mediante un plan de vuelo, que se crea antes de la misión.

Básicamente, se puede realizar un apuntamiento automático o manual de las cámaras para permitir cubrir el objetivo de manera óptima para el procesamiento fotogramétrico de los datos de vídeo.

Por ejemplo, las cámaras pueden apuntar de manera convergente al mismo objeto, que se muestra en el centro de cada imagen, o pueden apuntar estrictamente hacia abajo en el nadir para relajar los requisitos de sincronización de apuntamiento.

En un conjunto particular de realizaciones, la cámara de cada vehículo en movimiento o que puede moverse toma un fotograma de vídeo en el mismo momento que la cámara de un vehículo cooperativo para proporcionar tomas de vídeo sincrónicas.

A modo de ejemplo, la sincronización temporal puede realizarse haciendo coincidir una imagen de un objeto en movimiento desde un primer vehículo en el momento  $t_i$  con una secuencia de imágenes en momentos  $[t_i - n \dots t_i + n]$  de un segundo vehículo para encontrar la co-ocurrencia óptima del objeto en movimiento en el momento  $t_j$ ;  $-n \leq j \leq +n$ .

Opcionalmente, una diferencia en el tiempo,  $\Delta t = t_j - t_i$ , puede ser equivalente a una diferencia en los índices de fotogramas en las secuencias de vídeo, lo que permite la sincronización temporal con la precisión de la velocidad de fotogramas de vídeo.

También es posible realizar la sincronización temporal basándose en la detección del índice del primer fotograma en el que una luz de color distintivo, encendida mientras las cámaras de los vehículos en movimiento o que pueden moverse están captando imágenes, es visible en cada vídeo, y la definición del desfase temporal entre los vídeos como la diferencia en los índices de fotogramas.

Alternativamente, la sincronización temporal se puede realizar iniciando las cámaras exactamente en el mismo momento.

En una realización particular, se pueden generar imágenes estéreo, listas para ser mostradas estereoscópicamente en un casco de Realidad Virtual o con gafas 3D, basándose en los fotogramas de vídeo registrados.

Por ejemplo, se puede realizar una correspondencia de imágenes densas en las imágenes estéreo para generar mapas de profundidad, donde cada valor de píxel representa la distancia relativa entre la cámara y el objeto del que se han obtenido imágenes.

Por ejemplo, se pueden identificar áreas con objetos en movimiento utilizando los mapas de profundidad.

En un ejemplo particular, los objetos estáticos de fondo se separan de los objetos dinámicos en movimiento, y se genera un modelo 3D coherente sobre el objetivo que incluye solo terreno estático y objetos estáticos. Se puede crear una base de datos de objetos dinámicos en movimiento en la que se clasifican los objetos sólidos en movimiento y los objetos no sólidos en movimiento y se proporcionan los modelos 3D correspondientes de los objetos sólidos dinámicos en movimiento y los modelos 3D de los objetos no sólidos dinámicos en movimiento. El modelo 4D puede entonces basarse en el modelo 3D del terreno estático y de los objetos estáticos y en los modelos 3D de los objetos sólidos dinámicos en movimiento y en los modelos 3D de los objetos no sólidos dinámicos en movimiento.

De esta manera se puede proporcionar visualización y/o soporte de decisiones basado en los modelos 3D/4D generados.

Por ejemplo, el punto de vista y el zoom dentro del modelo 4D se pueden adaptar a las preferencias del usuario.

Para obtener una descripción general del sistema, se puede hacer referencia a los ejemplos esquemáticos ilustrados en la FIG. 1 y la FIG. 2.

Según un segundo aspecto, se proporciona un sistema 100 configurado para generar uno o más modelos basados en vídeo de un objetivo. El sistema 100 está configurado para proporcionar transmisiones de vídeo desde al menos dos vehículos en movimiento o que pueden moverse 10-1, 10-2, ..., 10-m ( $m \geq 2$ ) equipados con cámaras 20-1, 20-2, ..., 20-m para obtener imágenes simultáneamente del objetivo 50 desde diferentes puntos de vista. Se proporciona una sincronización de posición de los vehículos en movimiento o que pueden moverse 10-1, 10-2, ..., 10-m para crear una base de imagen estable. Se proporciona una sincronización de apuntamiento de las cámaras 20-1, 20-2, ..., 20-m para cubrir el (los) mismo(s) objeto(s) y/o evento(s) dinámico(s), y se habilita la sincronización temporal de los fotogramas de vídeo de las transmisiones de vídeo para obtener, para al menos un punto en el tiempo, un conjunto de fotogramas de vídeo registrados simultáneamente. El sistema 100 también está configurado para generar, para dicho al menos un punto en el tiempo, al menos un modelo tridimensional, 3D, del objetivo basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente.

En un ejemplo particular, el sistema 100 está configurado para permitir la sincronización temporal de los fotogramas de vídeo de las transmisiones de vídeo para obtener, para cada uno de una pluralidad de puntos en el tiempo, un conjunto de fotogramas de vídeo registrados simultáneamente. El sistema 100 puede estar configurado para generar, para cada uno de un conjunto seleccionado de la pluralidad de puntos en el tiempo, al menos un modelo 3D del objetivo basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente.

Opcionalmente, el sistema 100 está configurado además para combinar al menos un subconjunto de los modelos 3D generados para generar un modelo de cuatro dimensiones, 4D del objetivo para permitir que un usuario navegue a través del modelo 4D en las tres dimensiones espaciales y en la dimensión temporal.

Por ejemplo, el sistema puede estar configurado para descargar o transmitir por enlace descendente datos de imágenes basados en las transmisiones de vídeo a un segmento de tierra 30 desde al menos uno de los vehículos en movimiento o que pueden moverse y/o para hacer funcionar un enlace ascendente de control para controlar los vehículos en movimiento o que pueden moverse 10-1, 10-2, ..., 10-m y/o las cámaras 20-1, 20-2, ..., 20-m desde el segmento de tierra 30.

En un ejemplo particular, el sistema está configurado para separar los objetos estáticos de fondo de los objetos dinámicos en movimiento, y generar un modelo 3D coherente sobre el objetivo que incluye solo terreno estático y objetos estáticos, y crear una base de datos de objetos dinámicos en movimiento en la que se clasifican los objetos sólidos en movimiento y los objetos no sólidos en movimiento y los modelos 3D correspondientes de los objetos sólidos dinámicos en movimiento y los modelos 3D de los objetos no sólidos dinámicos en movimiento. El sistema puede configurarse entonces para generar el modelo 4D basado en el modelo 3D del terreno estático y los objetos estáticos y los modelos 3D de los objetos sólidos dinámicos en movimiento y los modelos 3D de objetos no sólidos dinámicos en movimiento.

A modo de ejemplo, el sistema 100 (véase la FIG. 6) comprende un circuito de procesamiento 110 y una memoria 120, comprendiendo la memoria 120 instrucciones que, cuando son ejecutadas por el circuito de procesamiento 110, hacen que el circuito de procesamiento 110 genere los modelos basados en vídeo de eventos dinámicos, o circuitos electrónicos configurados para generar los modelos basados en vídeo de eventos dinámicos.

Según un tercer aspecto, se proporciona un sistema de sensor dinámico 25. El sistema de sensor dinámico 25 comprende un conjunto de vehículos en movimiento o que pueden moverse 10-1, 10-2, ..., 10-m que funcionan sincrónicamente, equipados con cámaras 20-1, 20-2, ..., 20-m para permitir transmisiones de vídeo para obtener imágenes simultáneamente de un objetivo desde diferentes puntos de vista. El sistema de sensor dinámico 25 está configurado para funcionar basándose en la sincronización de posición de los vehículos en movimiento o que pueden moverse 10-1, 10-2, ..., 10-m para crear una base de imagen estable, y el sistema de sensor dinámico 25 está configurado para funcionar basándose en la sincronización de apuntamiento de las cámaras 20-1, 20-2, ..., 20-m para cubrir el (los) mismo(s) objeto(s) y/o evento(s) dinámico(s). El sistema de sensor dinámico 25 también está configurado para permitir la sincronización temporal de los fotogramas de vídeo de las transmisiones de vídeo para permitir, para al menos un punto en el tiempo, la generación y/o extracción de un conjunto de fotogramas de vídeo registrados simultáneamente. El sistema de sensor dinámico 25 está configurado además para transmitir datos de imagen de las transmisiones de vídeo a un segmento de tierra 30 desde al menos uno de los vehículos en movimiento o que pueden moverse 10-1, 10-2, ..., 10-m para permitir la generación de uno o más modelos tridimensionales/tetradimensionales, 3D/4D, del objetivo basados en los fotogramas de vídeo registrados simultáneamente en el segmento de tierra 30.

A modo de ejemplo, el sistema de sensor dinámico 25 permite la formación de un segmento aéreo de drones 10-1, 10-2, ..., 10-m operativos sincrónicamente que tienen cámaras a bordo 20-1, 20-2, ..., 20-m para permitir que los flujos de vídeo sincronizados tomen imágenes simultáneas del objetivo desde diferentes puntos de vista.

Según un cuarto aspecto, se proporciona un sistema terrestre 30 configurado para generar uno o más modelos basados en vídeo de un objetivo. El sistema 30 está configurado para recibir transmisiones de vídeo desde al menos dos vehículos en movimiento o que pueden moverse 10-1, 10-2, ..., 10-m equipados con cámaras 20-1, 20-2, ..., 20-m para obtener imágenes simultáneamente del objetivo 50 desde diferentes puntos de vista. Los vehículos en movimiento o que pueden moverse 10-1, 10-2, ..., 10-m se hacen funcionar basándose en la sincronización de posición para una base de imagen estable, y las cámaras 20-1, 20-2, ..., 20-m se hacen funcionar basándose en la sincronización de apuntamiento para cubrir el (los) mismo(s) objeto(s) y/o evento(s) dinámico(s), y la sincronización de tiempo de los fotogramas de vídeo de las transmisiones de vídeo se habilita para permitir, para al menos un punto en el tiempo, la generación y/o extracción de un conjunto de fotogramas de vídeo registrados simultáneamente. El sistema 30 también está configurado para generar, para dicho al menos un punto en el tiempo, al menos un modelo tridimensional, 3D, del objetivo basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente.

En un ejemplo particular, el sistema 30 está configurado para permitir la sincronización temporal de los fotogramas de vídeo de las transmisiones de vídeo para obtener, para cada uno de una pluralidad de puntos en el tiempo, un conjunto de fotogramas de vídeo registrados simultáneamente. El sistema 30 puede estar configurado entonces para generar, para cada uno de un conjunto seleccionado de la pluralidad de puntos en el tiempo, al menos un modelo 3D del objetivo basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente.

Además, el sistema 30 puede estar configurado opcionalmente para combinar al menos un subconjunto de los modelos 3D generados para generar un modelo de cuatro dimensiones, 4D del objetivo para permitir que un usuario navegue a través del modelo 4D en las tres dimensiones espaciales y en la dimensión temporal.

Por ejemplo, el sistema 30 está configurado para hacer funcionar un enlace ascendente de control para controlar los vehículos en movimiento o que pueden moverse 10-1, 10-2, ..., 10-m y/o las cámaras 20-1, 20-2, ..., 20-m.



A modo de ejemplo, el sistema 30 puede estar configurado para determinar la sincronización de posición y la sincronización de apuntamiento y enviar un comando en el enlace ascendente de control para adaptar la sincronización de posición y la sincronización de apuntamiento de los vehículos en movimiento o que pueden moverse 10 y/o las cámaras 20.

A modo de ejemplo, el sistema 30 (véase la FIG. 6) comprende un circuito de procesamiento 110 y una memoria 120, comprendiendo la memoria 120 instrucciones que, cuando son ejecutadas por el circuito de procesamiento 110, hacen que el circuito de procesamiento 110 genere los modelos basados en vídeo de eventos dinámicos, o circuitos electrónicos configurados para generar los modelos basados en vídeo de eventos dinámicos.

A continuación, se describirá la tecnología propuesta con referencia a diversos ejemplos no limitativos con fines ilustrativos.

Para simplificar, la innovación se describirá ahora principalmente en el contexto de dos drones con cámaras de vídeo. Sin embargo, la innovación funciona igualmente bien con más drones simultáneos y algunos aspectos de calidad serían incluso mejores con más drones.

En realizaciones particulares, la tecnología propuesta se basa en tres tipos de sincronización:

- **Sincronización de posición** de los drones para crear una base de imagen precisa, estable y conocida;
- **Sincronización de apuntamiento** de las cámaras para cubrir el mismo objeto;
- **Sincronización temporal** de los fotogramas en la transmisión de vídeo para obtener imágenes simultáneas del mismo objeto.

Para muchas aplicaciones de usuario, existe otra parte importante de la innovación, es decir, el aspecto de tiempo real:

- Enlace descendente **en tiempo real** de la transmisión de vídeo
- Procesamiento **en tiempo real**
- Visualización **en tiempo real** del modelo 4D

Para una mejor comprensión de la presente invención puede ser útil comenzar con una breve descripción general del sistema. Aunque en este conjunto particular de ejemplos los drones son aéreos, debe entenderse que se pueden emplear vehículos de superficie o incluso submarinos para capturar las secuencias de vídeo.

### Ejemplo de sistema general

En este ejemplo, el sistema comprende (véase FIG. 1 y/o FIG. 2):

- Segmento Aéreo
  - Una multitud (dos o más) de drones controlados desde el Segmento de Tierra;
  - Al menos uno de los drones puede descargar los datos de imagen de la multitud al Segmento de Tierra.
- Sistema de sensor/vídeo/cámara
  - Normalmente una cámara de fotogramas a bordo de cada uno de los drones;
  - La velocidad de fotogramas de la cámara puede depender de la aplicación.
- Segmento de Tierra
  - Normalmente se ubica en la primera línea y mantiene el control general del sistema, incluyendo uno o más de los siguientes:
  - Control de vuelo de drones,

- Recepción de datos de la multitud,
- Procesamiento de los datos,
- Herramienta de visualización y toma de decisiones para el oficial de primera línea,
- Comunicación externa potencial, compartición de datos con back-office, etc.

- Objetivo

o Cualquier tipo de objetivo, como un evento dinámico donde los modelos 3D/4D en tiempo real facilitarían la toma de decisiones, por ejemplo, incendios, accidentes, inundaciones u otras operaciones militares o de seguridad. Otros escenarios objetivo pueden incluir eventos deportivos, producciones cinematográficas y similares.

### Ejemplos de sincronización de posición de los drones

La base de imagen (véase FIG. 2 y FIG. 3), es decir, la distancia entre los drones, está optimizada para una reconstrucción automática precisa de puntos 3D y depende de la altura sobre el terreno y de los parámetros intrínsecos de las cámaras.

Con referencia particular a la FIG. 2, el vuelo sincrónico del dron 1 (incluido el apuntamiento de sensor 2) puede ser controlado desde el segmento de tierra por un oficial de primera línea, principalmente utilizando parámetros preestablecidos para el vuelo. Como alternativa, el vuelo podría controlarse de forma autónoma.

A modo de ejemplo, el área de superposición entre fotogramas de vídeo sincronizados se puede determinar como una etapa integrada en la generación de modelos 4D. Esta superposición puede usarse en un ejemplo particular para determinar cómo deben reconfigurarse los drones para mantener una sincronización de posicionamiento y de apuntamiento óptima. Los comandos necesarios para actualizar la sincronización de posicionamiento y de apuntamiento de los drones, en este ejemplo, pueden calcularse y enviarse a los drones de manera autónoma por el sistema.

Como se mencionó, la sincronización de posicionamiento y/o apuntamiento corresponden a la orientación relativa, que si se desea se puede calcular de manera continua en tiempo real en función del contenido de las transmisiones de vídeo y se puede utilizar para controlar los drones y/o cámaras de manera que obtengan y/o mantengan la sincronización de posicionamiento y/o apuntamiento deseada.

Por ejemplo, la sincronización de posición puede realizarse mediante un enlace de radio entre los drones (directo o a través de una estación de radio terrestre), en el que se comunican las posiciones de los drones para que se pueda mantener una relación base-altura correcta.

Como otro ejemplo, la sincronización de posición puede realizarse mediante un plan de vuelo, que se crea antes de la misión, en el que la distancia entre los drones se establece en un valor constante basado en la altura de vuelo, el campo de visión y otros parámetros.

### Ejemplos de sincronización de apuntamiento de las cámaras

Para garantizar que el objetivo esté cubierto de forma óptima para el procesamiento fotogramétrico de los datos de vídeo, el sistema puede utilizar el apuntamiento automático o manual de las cámaras.

A modo de ejemplo, para maximizar el área de superposición cubierta por ambos drones (o todos), las cámaras están apuntando de manera convergente al mismo objeto, que aparece en el centro de cada imagen.

Como otro ejemplo, al relajar los requisitos de sincronización de apuntamiento, las cámaras pueden apuntar estrictamente hacia abajo en el nadir. El área de superposición se determina entonces mediante la relación base-altura y los parámetros intrínsecos de la cámara, como el campo de visión.

A modo de ejemplo, la sincronización de apuntamiento se puede realizar mediante un enlace de radio entre los drones (posiblemente a través de una estación de radio terrestre), en el que se comunican las posiciones de los drones para que se pueda calcular un objeto objetivo correcto en el que centrarse.

### Ejemplos de sincronización temporal de los fotogramas

Para que sea posible crear modelos 4D de objetivos en movimiento, el sensor/cámara de cada dron debe tomar un fotograma de vídeo en el mismo momento que el dron cooperativo (el vídeo sincrónico toma 3), de lo contrario el objetivo se moverá o cambiará de forma entre las tomas de fotogramas.

La precisión de la sincronización, así como la velocidad de fotogramas de vídeo, dependen de los requisitos de cada aplicación específica. Además, cada fotograma puede estar asociado a un etiquetado de tiempo preciso.

Por ejemplo, la sincronización temporal se puede lograr haciendo coincidir una imagen de un objeto en movimiento de un dron en el momento  $t_i$  con una secuencia de imágenes en los momentos  $[t_{i-n}, t_{i+n}]$  de un segundo dron para encontrar la co-ocurrencia óptima del objeto en movimiento en el momento  $t_{j_i}$ ,  $-n \leq j \leq +n$ . La diferencia en el tiempo,  $\Delta t = t_j - t_i$ , es equivalente a una diferencia en los índices de fotogramas en las secuencias de vídeo, lo que permite la sincronización temporal con la precisión de la velocidad de fotogramas de vídeo, por ejemplo, 1/30 segundos. Esta calibración temporal se puede iterar durante el vuelo para evitar desviaciones de tiempo debido, por ejemplo, a diferentes relojes internos de la cámara.

Como otro ejemplo, la sincronización temporal se puede realizar utilizando un único obturador electrónico, conectado a ambos drones (todos) antes de la misión, iniciando así las cámaras exactamente al mismo tiempo.

Como otro ejemplo, se enciende una luz de un color distintivo mientras ambos drones (todos) están tomando imágenes. El índice del primer fotograma en el que es visible la luz se detecta automáticamente en cada vídeo. La diferencia en los índices de fotogramas define el desfase temporal entre los vídeos.

## Ejemplos de enlace(s) de comunicación

Se puede establecer un enlace de comunicación 4 entre el segmento aéreo y el segmento de tierra, que incluye al menos un enlace descendente de vídeo en tiempo real y, opcionalmente, un enlace de control (enlace ascendente) para controlar el segmento aéreo, por ejemplo, los drones y/o los sensores.

## Ejemplos de procesamiento para modelos 3D

Ejemplos de características para el procesamiento 5 para modelos 3D en tiempo real:

- Las orientaciones interiores (es decir, las características físicas de las cámaras) pueden determinarse antes o durante la misión, utilizando objetivos u objetos naturales.
- Cada conjunto/par de imágenes de diferentes vídeos/cámaras, pero del mismo momento  $t$ , pueden orientarse entre sí a través de la coincidencia de imágenes y la orientación del sensor.
- Se pueden generar imágenes estéreo, listas para ser mostradas estereoscópicamente en cascos de realidad virtual (VR) o con gafas 3D, preferiblemente en cada uno de una pluralidad de puntos en el tiempo (de modo que un vídeo pueda verse estereoscópicamente).
- Se puede realizar una coincidencia de imágenes densas en las imágenes estéreo para generar mapas de profundidad, es decir, imágenes donde el valor del píxel es la distancia relativa entre la cámara y el objeto del que se han obtenido imágenes.
- Los datos del GPS (Sistema de Posicionamiento Global) y de la IMU (Unidad de Medición de Inercia) se pueden utilizar para calcular las coordenadas de los objetos (X, Y, Z) a partir de los mapas de profundidad y para calcular la orientación exterior (es decir, la posición y las orientaciones de las cámaras) -> nubes de puntos y modelos digitales de superficie (DSM), lo que permite la producción de modelos 3D georreferenciados o nubes de puntos a partir de una correspondencia densa.
- Los datos de orientación y los mapas de profundidad de una secuencia de fotogramas ya calculados se pueden utilizar para predecir los datos en el fotograma procesado actualmente para acelerar los cálculos.
- Se pueden guardar mapas de profundidad y/o DSM para permitir la reproducción del escenario en tiempo real y proporcionar información para el proceso casi en tiempo real.

## Ejemplos de procesamiento para modelos 4D

Ejemplos de características para el procesamiento 5 para modelos 4D casi en tiempo real:

- Para todas las transmisiones de vídeo sincronizadas (dos o más) o para un período de tiempo determinado, se puede realizar una coincidencia entre fotogramas dentro de cada transmisión de vídeo.
- Junto con los mapas de profundidad entre las transmisiones de vídeo de la solución en tiempo real, obtenemos observaciones para un bloque de imágenes.

- El bloque de imágenes se puede ajustar para estimar datos de orientación globalmente óptimos.
  - Esto puede hacerse inicialmente sin datos de GPS e IMU, en cuyo caso la orientación absoluta del bloque se realiza después del ajuste,
  - o las observaciones de GPS e IMU se pueden incluir en el ajuste.
- La coincidencia de imágenes densas se puede realizar utilizando los parámetros del sensor ajustados y los mapas de profundidad existentes del procesamiento en tiempo real.
- Las áreas con objetos en movimiento se pueden identificar utilizando los mapas de profundidad, flujo óptico y, por ejemplo, factorización afin y proyectiva.
- Se puede crear un modelo 3D coherente que abarque toda el área y que incluya únicamente terreno y objetos estáticos (que no se mueven).
- Se puede crear una base de datos con objetos en movimiento en la que se clasifiquen objetos en movimiento (i) sólidos y (ii) deformables. La base de datos puede incluir modelos 3D dependientes del tiempo.
- Juntos, los modelos 3D de objetos estáticos y en movimiento definen un modelo 4D dependiente del tiempo de toda el área de imágenes.

## Ejemplos de visualización

La tecnología propuesta puede opcionalmente incluir también una herramienta 6 de visualización y/o toma de decisiones/soprote basada en los modelos 3D/4D generados de los eventos estáticos y/o dinámicos, tal y como se indica en el presente documento.

La FIG. 4 es un diagrama esquemático que ilustra un ejemplo de cómo se pueden generar modelos 3D momentáneos, por ejemplo, a partir de pares de fotogramas de vídeo sincronizados en el tiempo o, más generalmente, n-tuples para visualización y exploración.

La FIG. 5 ilustra un ejemplo de cómo los modelos 3D momentáneos juntos pueden usarse para construir modelos 4D en tiempo real y casi en tiempo real para visualización y exploración.

En un ejemplo particular no limitativo, ilustrado por simplicidad con referencia a dos sensores en la FIG. 4, el procesamiento en tiempo real puede brindar al usuario una vista 4D inmediata, pero algo limitada, del objetivo:

1. Los sensores 1 y 2 toman una transmisión de fotogramas de vídeo y crean un par de imágenes en cada momento.
2. A partir de cada par de imágenes, el sistema procesa un modelo 3D momentáneo en tiempo real,
3. Cada modelo 3D se exporta como una trama (en "espacio de imagen") al módulo de visualización. Juntos forman un modelo 4D del objetivo.

Para simplificar, los gráficos muestran solo dos sensores voladores, pero el sistema podría manejar más sensores, lo que mejoraría la calidad del modelo 4D en tiempo real.

Por ejemplo, la visualización en tiempo real puede ser un modelo 4D que ofrece al usuario una impresión inmediata y mejorada del objetivo. Esta vista es mucho mejor que simplemente ver el vídeo en tiempo real desde un dron, ya que el punto de vista y el zum se pueden adaptar a las preferencias del usuario.

Además, en los modelos 4D se pueden realizar mediciones manuales y automáticas como altura y volumen.

En un ejemplo particular no limitativo, para simplificar, ilustrado con referencia a dos sensores en la FIG. 5, el procesamiento casi en tiempo real puede brindar al usuario una vista 4D algo retrasada (en segundos), pero a escala completa del objetivo, que también se puede estudiar en el dominio del tiempo (retroceder para ver lo que sucedió antes):

1. La creación de modelos 3D momentáneos a partir de los sensores 1 y 2 es similar, pero paralela, al procesamiento en tiempo real.

2. El sistema separa los objetos estáticos de fondo de los objetos dinámicos en movimiento.

3. El sistema separa los objetos en movimiento sólidos de los no sólidos.

4. Los modelos 3D de objetos estáticos y los objetos sólidos dinámicos se exportan como TIN (en "espacio de objetos") al módulo de exploración. Una TIN (red irregular triangulada) es una representación de una superficie continua que consiste enteramente de facetas triangulares.

5. Los modelos 3D de objetos dinámicos no sólidos se exportan como una nube de puntos (en "espacio de objetos") al módulo de exploración.

6. Juntos, todos los modelos 3D componen un modelo 4D del objetivo.

Las estructuras de datos TIN digitales se pueden utilizar en una variedad de aplicaciones, incluidos los sistemas de información geográfica (GIS) y el dibujo asistido por computadora (CAD) para la representación visual de una superficie topográfica. Una TIN es normalmente una representación basada en vectores de la superficie física de la tierra o del fondo del mar, formada por nodos y líneas distribuidos irregularmente con coordenadas tridimensionales (x, y, z) que están dispuestos en una red de triángulos no superpuestos.

Para simplificar, los gráficos muestran solo dos sensores voladores, pero el sistema podría manejar más sensores, lo que mejoraría la calidad del modelo 4D en tiempo real.

Por ejemplo, la exploración casi en tiempo real puede implicar una herramienta para un análisis más profundo. Por ejemplo, puede almacenar los modelos 4D generados y puede visualizar los modelos simultáneamente, también en el dominio del tiempo. La herramienta también puede permitir análisis más avanzados como, por ejemplo:

- Seguimiento de objetos
- Cambios en el volumen objetivo
- Modelos estáticos de objetos en movimiento
- Identificación de objetos.

Al menos algunas de las etapas, funciones, procedimientos, módulos y/o bloques descritos en el presente documento pueden implementarse en software tal como un programa informático para su ejecución mediante circuitos de procesamiento adecuados tales como uno o más procesadores o unidades de procesamiento. Los ejemplos de circuitos de procesamiento incluyen, entre otros, uno o más microprocesadores, una o más unidades de procesamiento de gráficos (GPU), uno o más procesadores de señales digitales (DSP), una o más unidades centrales de procesamiento (CPU), hardware de aceleración de vídeo y/o cualquier circuito lógico programable adecuado, como una o más matrices de puertas programables en campo (FPGA) o uno o más controladores lógicos programables (PLC).

Alternativamente, o como complemento, al menos algunos de las etapas, funciones, procedimientos, módulos y/o bloques descritos en el presente documento pueden implementarse en hardware utilizando cualquier tecnología convencional, tal como tecnología de circuitos discretos o circuitos integrados, incluyendo tanto circuitos electrónicos de propósito general como circuitos específicos de la aplicación. También debe entenderse que puede ser posible reutilizar las capacidades generales de procesamiento de cualquier dispositivo o unidad convencional en el que se implemente la tecnología propuesta. También es posible reutilizar el software existente, por ejemplo, reprogramando el software existente o añadiendo nuevos componentes de software.

La funcionalidad general también puede dividirse entre software programado, SW, para su ejecución en uno o más procesadores, y uno o más circuitos de hardware preconfigurados o posiblemente reconfigurables, como ASIC y/o FPGA. La partición real de hardware y software puede ser decidida por un diseñador de sistemas basándose en una serie de factores, entre ellos la velocidad de procesamiento, el costo de implementación y otros requisitos.

La FIG. 6 es un diagrama esquemático que ilustra un ejemplo de una implementación informática según una realización. En este ejemplo particular, al menos algunos de las etapas, funciones, procedimientos, módulos y/o bloques descritos en el presente documento se implementan en un programa informático, que se carga en la memoria para su ejecución mediante un circuito de procesamiento 110 que incluye uno o más procesadores. El (los) procesador(es) 110 y la memoria 120 están interconectados entre sí para permitir la ejecución normal del software. Un dispositivo de entrada/salida 130 opcional también puede interconectarse al (a los) procesador(es) 110 y/o a la memoria 120 para permitir la entrada y/o salida de datos relevantes tales como parámetro(s) de entrada y/o parámetro(s) de salida resultante(s).

En una realización particular, el procesamiento y la visualización pueden compartir recursos de memoria a través de la interoperabilidad de GPU.

El término "procesador" debe interpretarse en un sentido general como cualquier sistema o dispositivo capaz de ejecutar código de programa o instrucciones de programa informático para realizar una tarea particular de procesamiento, determinación o cálculo.

El circuito de procesamiento que incluye uno o más procesadores está así configurado para realizar, al ejecutar el programa informático, tareas de procesamiento bien definidas como las descritas en el presente documento.

El circuito de procesamiento no tiene que estar dedicado únicamente a ejecutar las etapas, funciones, procedimientos y/o bloques descritos anteriormente, sino que también puede ejecutar otras tareas.

En una realización particular, el programa informático 125; 145 comprende instrucciones que, cuando son ejecutadas por al menos un procesador 110, hacen que el (los) procesador(es) 110 realice(n) al menos parte de las etapas y/o acciones o tareas descritas en el presente documento.

En un ejemplo particular, se proporciona un programa informático 125; 145 para generar, cuando es ejecutado por un procesador 110, uno o más modelos basados en vídeo de un objetivo. El programa informático 125; 145 comprende instrucciones que, cuando son ejecutadas por el procesador 110, hacen que el procesador 110:

- reciba transmisiones de vídeo de al menos dos vehículos en movimiento o que pueden moverse equipados con cámaras para obtener imágenes simultáneamente del objetivo desde diferentes puntos de vista,

- o los vehículos en movimiento o que pueden moverse se operan basándose en la sincronización de posición para una base de imagen estable,

- o las cámaras se operan basándose en la sincronización de apuntamiento para cubrir el (los) mismo(s) objeto(s) y/o evento(s) dinámico(s), y

- o se proporciona una sincronización temporal de los fotogramas de vídeo de las transmisiones de vídeo para obtener, para al menos un punto en el tiempo, un conjunto de fotogramas de vídeo registrados simultáneamente; y

- genere, para dicho al menos un punto en el tiempo, al menos un modelo tridimensional, 3D, o modelo de cuatro dimensiones, 4D, del objetivo basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente.

Opcionalmente, al generarse varios modelos 3D sobre una pluralidad de puntos en el tiempo, es posible combinar al menos un subconjunto de los modelos 3D generados sobre la pluralidad de puntos en el tiempo para generar un modelo de cuatro dimensiones, 4D, del objetivo para permitir que un usuario navegue a través del modelo 4D en las tres dimensiones espaciales y en la dimensión temporal.

A modo de ejemplo, el software o programa informático 125; 145 puede realizarse como un producto de programa informático, que normalmente se transporta o almacena en un medio legible por ordenador 120; 140, en particular un medio no volátil. El programa informático 125; 145 puede así cargarse en la memoria 120 operativa de un ordenador o dispositivo de procesamiento equivalente para su ejecución por el circuito 110 de procesamiento del mismo.

El diagrama o diagramas de flujo presentados en el presente documento pueden considerarse como un diagrama o diagramas de flujo de computadora, cuando son realizados por uno o más procesadores. Un aparato correspondiente puede definirse como un grupo de módulos de función, donde cada etapa realizada por el procesador corresponde a un módulo de función. En este caso, los módulos de función se implementan como un programa informático que se ejecuta en el procesador.

El programa informático que reside en la memoria puede así organizarse como módulos de función apropiados configurados para realizar, cuando son ejecutados por el procesador, al menos parte de las etapas y/o tareas descritas en el presente documento.

Las realizaciones descritas anteriormente deben entenderse como algunos ejemplos ilustrativos de la presente invención. Los expertos en la materia entenderán que se pueden realizar diversas modificaciones, combinaciones y cambios en las realizaciones sin alejarse del alcance de la presente invención tal como se define en las reivindicaciones.

## REIVINDICACIONES

1. Un método para generar uno o más modelos basados en vídeo de un objetivo, en donde el método comprende:
  - proporcionar (S1) transmisiones de vídeo desde al menos dos vehículos (10) en movimiento o que pueden moverse equipados con cámaras (20) para obtener imágenes simultáneamente del objetivo (50) desde diferentes puntos de vista,
  - en donde se proporciona una sincronización de posición de los vehículos (10) en movimiento o que pueden moverse para crear una base de imagen estable, que representa la distancia entre los vehículos en movimiento o que pueden moverse,
  - en donde se proporciona una sincronización de apuntamiento de las cámaras (20) para cubrir el (los) mismo(s) objeto(s) y/o evento(s) dinámico(s), y
  - en donde se proporciona una sincronización temporal de los fotogramas de vídeo de las transmisiones de vídeo para obtener, para cada uno de una pluralidad de puntos en el tiempo, un conjunto de fotogramas de vídeo registrados simultáneamente, y
  - en donde se genera al menos un modelo 3D del objetivo, para cada uno de un conjunto seleccionado de la pluralidad de puntos en el tiempo, basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente utilizando fotogrametría,**caracterizado por que** el método comprende, además
  - combinar (S3) al menos un subconjunto de los modelos 3D generados para generar un modelo de cuatro dimensiones, 4D, como una serie de modelos 3D dependientes del tiempo, del objetivo para permitir que un usuario navegue a través del modelo 4D en las tres dimensiones espaciales y en la dimensión temporal, y descargar las transmisiones de vídeo desde dichos al menos dos vehículos (10) en movimiento o que pueden moverse a un sistema terrestre (30), que realiza el procesamiento de datos de vídeo para generar los modelos 3D y el modelo 4D.
2. El método de la reivindicación 1, en donde los datos de imagen basados en las transmisiones de vídeo se transmiten en enlace descendente a un segmento de tierra (30) desde al menos uno de los vehículos (10) en movimiento o que pueden moverse y/o se opera un enlace ascendente de control para controlar los vehículos (10) en movimiento o que pueden moverse y/o las cámaras (20) desde el segmento de tierra (30).
3. El método de la reivindicación 1 o 2, en donde los vehículos (10) en movimiento o que pueden moverse incluyen drones o vehículos aéreos no tripulados (UAV), cada uno de los cuales tiene una cámara (20) a bordo.
4. El método de cualquiera de las reivindicaciones 1 a 3, en donde la cámara (20) en cada vehículo (10) en movimiento o que puede moverse toma un fotograma de vídeo en el mismo momento que la cámara (20) de un vehículo (10) en movimiento o que puede moverse cooperativo para proporcionar tomas de vídeo sincrónicas.
5. El método de cualquiera de las reivindicaciones 1 a 4, en donde la sincronización temporal se realiza haciendo coincidir una imagen de un objeto en movimiento desde un primer vehículo en el momento  $t_i$  con una secuencia de imágenes en momentos  $[t_i, \dots, t_{i+n}]$  desde un segundo vehículo para encontrar la coocurrencia óptima del objeto en movimiento en el momento  $t_j$ ;  $-n \leq j \leq +n$ , y una diferencia en el tiempo,  $\Delta t = t_j - t_i$ , es equivalente a una diferencia en los índices de fotogramas en las secuencias de vídeo, lo que permite la sincronización temporal con la precisión de la velocidad de fotogramas de vídeo, o
  - en donde la sincronización temporal se basa en detectar el índice del primer fotograma en el que una luz de un color distinto, encendida mientras las cámaras de los vehículos en movimiento o que pueden moverse están captando imágenes, es visible en cada vídeo, y definir el desfase temporal entre los vídeos como la diferencia en los índices de fotogramas, o
  - en donde la sincronización temporal se realiza iniciando las cámaras (20) exactamente al mismo tiempo.
6. El método de cualquiera de las reivindicaciones 1 a 5, en donde se generan imágenes estéreo, listas para ser mostradas estereoscópicamente en un casco de Realidad Virtual o con gafas 3D, basándose en los fotogramas de vídeo registrados.
7. El método de la reivindicación 6, en donde se realiza una correspondencia de imágenes densas en las imágenes estéreo para generar mapas de profundidad, donde cada valor de píxel representa la distancia relativa entre la cámara y el objeto del que se han obtenido imágenes, y se identifican áreas con objetos en movimiento utilizando los mapas de profundidad.
8. El método de cualquiera de las reivindicaciones 1 a 7, en donde se crea un modelo 3D coherente sobre un área completa y que incluye solo terreno estático y objetos que no se mueven, y se crea una base de datos con objetos en movimiento que incluye modelos 3D dependientes del tiempo, y en donde el modelo 3D coherente y los modelos 3D de objetos en movimiento definen juntos un modelo 4D dependiente del tiempo, y/o
  - en donde los objetos estáticos de fondo se separan de los objetos dinámicos en movimiento, y se genera un modelo 3D coherente sobre el objetivo que incluye solo terreno estático y objetos estáticos, y se crea una base de datos de objetos dinámicos en movimiento en la que se clasifican los objetos sólidos en movimiento y los objetos no sólidos en movimiento y se proporcionan modelos 3D correspondientes de los objetos sólidos dinámicos en

movimiento y modelos 3D de objetos no sólidos dinámicos en movimiento, en donde el modelo 4D se basa en el modelo 3D de terreno estático y objetos estáticos y los modelos 3D de los objetos sólidos dinámicos en movimiento y los modelos 3D de objetos no sólidos dinámicos en movimiento.

5 9. El método de la reivindicación 8, en donde los modelos 3D de objetos estáticos y los objetos sólidos dinámicos se extraen como estructuras de datos de red irregular triangulada (TIN).

10. Un sistema (100) configurado para generar uno o más modelos basados en vídeo de un objetivo,  
 - en donde el sistema (100) está configurado para proporcionar transmisiones de vídeo desde al menos dos  
 10 vehículos (10) en movimiento o que pueden moverse equipados con cámaras (20) para obtener imágenes simultáneamente del objetivo (50) desde diferentes puntos de vista,  
 • en donde se proporciona una sincronización de posición de los vehículos (10) en movimiento o que pueden moverse para crear una base de imagen estable,  
 • en donde se proporciona una sincronización de apuntamiento de las cámaras (20) para cubrir el (los) mismo(s)  
 15 objeto(s) y/o evento(s) dinámico(s), y  
 - en donde el sistema (100) está configurado para permitir la sincronización temporal de los fotogramas de vídeo de las transmisiones de vídeo para obtener, para cada uno de una pluralidad de puntos en el tiempo, un conjunto de fotogramas de vídeo registrados simultáneamente, y  
 - en donde el sistema (100) está configurado para generar, para cada uno de un conjunto seleccionado de la  
 20 pluralidad de puntos en el tiempo, al menos un modelo 3D del objetivo basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente utilizando fotogrametría, **caracterizado por que**  
 - el sistema (100) está configurado para combinar al menos un subconjunto de los modelos 3D generados para generar un modelo de cuatro dimensiones, 4D, como una serie de modelos 3D dependientes del tiempo, del  
 25 objetivo para permitir que un usuario navegue a través del modelo 4D en las tres dimensiones espaciales y en la dimensión temporal,  
 - y en donde el sistema (100) comprende un sistema terrestre (30) configurado para recibir las transmisiones de vídeo de dichos al menos dos vehículos (10) en movimiento o que pueden moverse y configurado para realizar el procesamiento de datos de vídeo para generar los modelos 3D y el modelo 4D.

30 11. El sistema de la reivindicación 10, en donde el sistema (100) está configurado para descargar o transmitir en enlace descendente datos de imágenes basados en las transmisiones de vídeo a un segmento de tierra (30) desde al menos uno de los vehículos (10) en movimiento o que pueden moverse y/o para operar un enlace ascendente de control para controlar los vehículos (10) en movimiento o que pueden moverse y/o las cámaras (20) desde el  
 35 segmento de tierra (30).

12. El sistema de la reivindicación 10 u 11, en donde el sistema está configurado para crear un modelo 3D coherente sobre un área entera y que incluye solo terreno estático y objetos que no se mueven, y para crear una base de datos con objetos en movimiento que incluyen modelos 3D dependientes del tiempo, y para combinar el  
 40 modelo 3D coherente y los modelos 3D de objetos en movimiento para definir un modelo 4D dependiente del tiempo.

13. Un sistema terrestre (30) configurado para generar uno o más modelos basados en vídeo de un objetivo,  
 - en donde el sistema (30) está configurado para recibir transmisiones de vídeo desde al menos dos vehículos (10)  
 45 en movimiento o que pueden moverse equipados con cámaras (20) para obtener imágenes simultáneamente del objetivo (50) desde diferentes puntos de vista,  
 • los vehículos (10) en movimiento o que pueden moverse se operan basándose en la sincronización de posición para una base de imagen estable,  
 • las cámaras (20) se operan basándose en la sincronización de apuntamiento para cubrir el (los) mismo(s)  
 50 objeto(s) y/o evento(s) dinámico(s), y  
 - en donde el sistema (30) está configurado para operar un enlace ascendente de control para controlar los vehículos (10) en movimiento o que pueden moverse y/o las cámaras (20);  
 - en donde el sistema (30) está configurado para permitir la sincronización temporal de los fotogramas de vídeo de las transmisiones de vídeo para obtener, para cada uno de una pluralidad de puntos en el tiempo, un conjunto de  
 55 fotogramas de vídeo registrados simultáneamente, y  
 - en donde el sistema (30) está configurado para generar, para cada uno de un conjunto seleccionado de la pluralidad de puntos en el tiempo, al menos un modelo 3D del objetivo basado en el conjunto correspondiente de fotogramas de vídeo registrados simultáneamente utilizando fotogrametría, **caracterizado por que**  
 - el sistema (30) está configurado para combinar al menos un subconjunto de los modelos 3D generados para  
 60 generar un modelo de cuatro dimensiones, 4D del objetivo para permitir que un usuario navegue a través del modelo 4D en las tres dimensiones espaciales y en la dimensión temporal.

14. Un programa informático (125; 145) para generar, cuando es ejecutado por una pluralidad de procesadores (110), uno o más modelos basados en vídeo de un objetivo, en donde el programa informático (125; 145) comprende instrucciones, que cuando son ejecutadas por los procesadores (110), hacen que los procesadores  
 65 (110) realicen el método de cualquiera de las reivindicaciones 1 a 9.



DIBUJOS

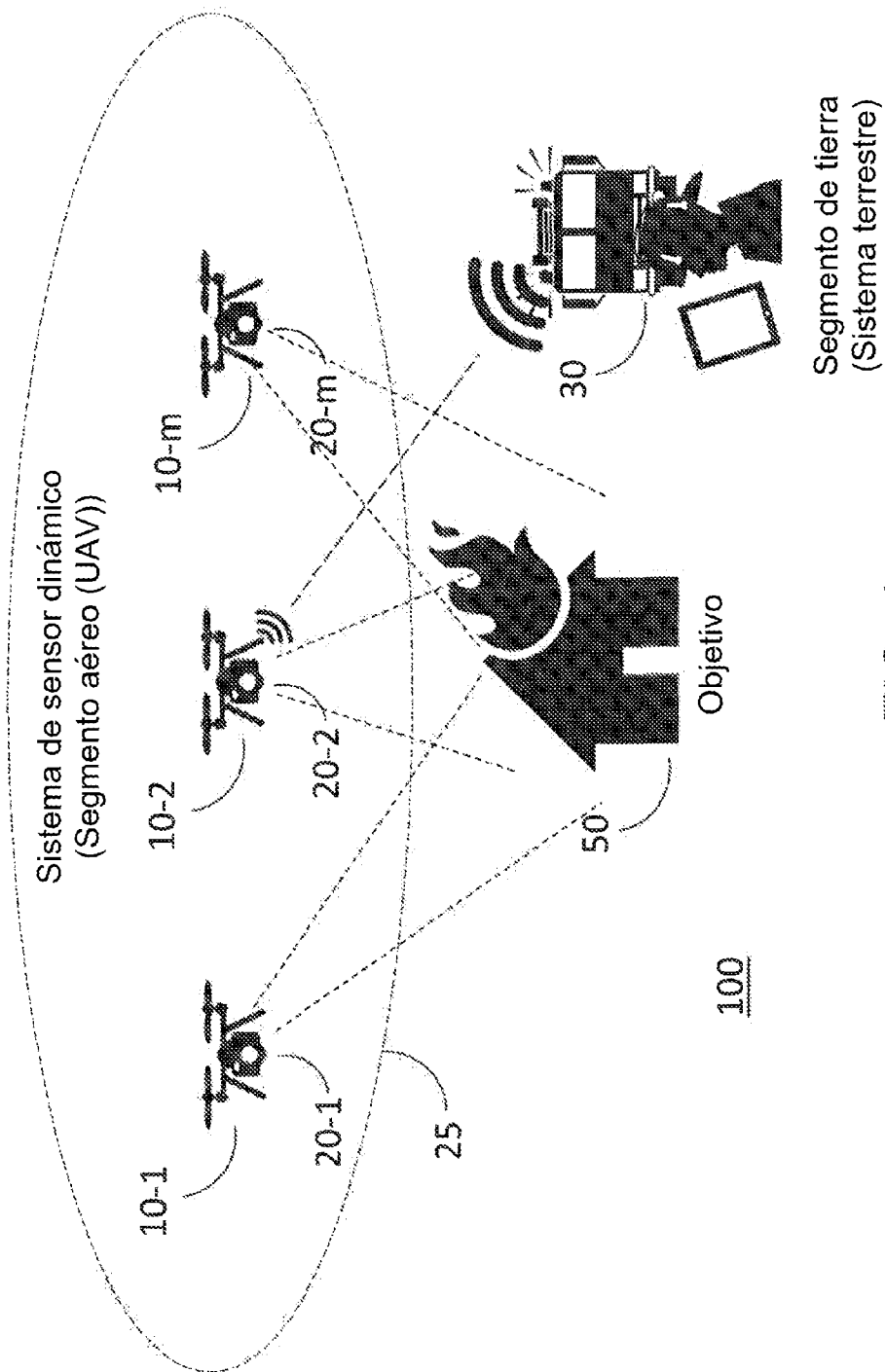


FIG. 1

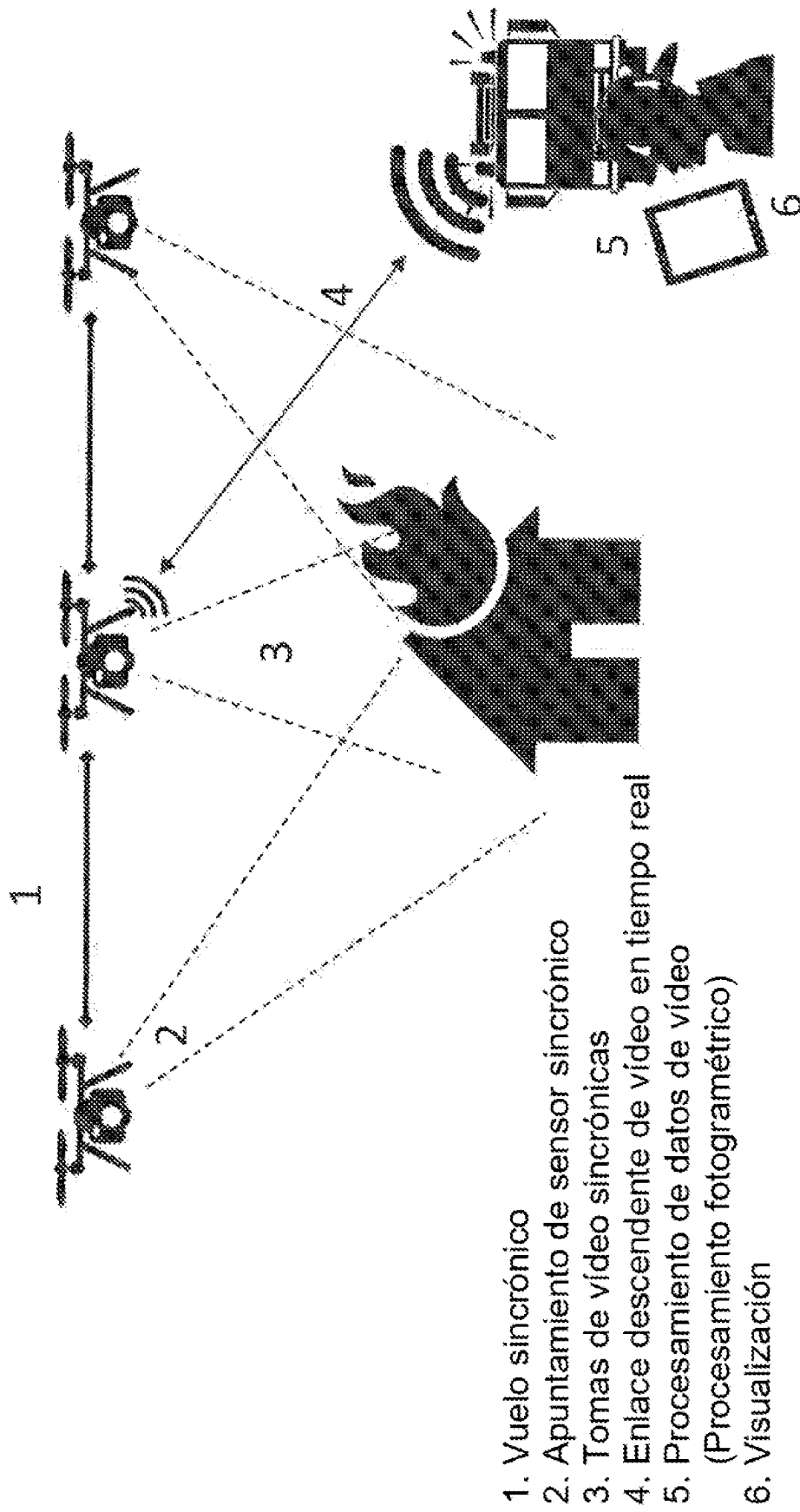


FIG. 2

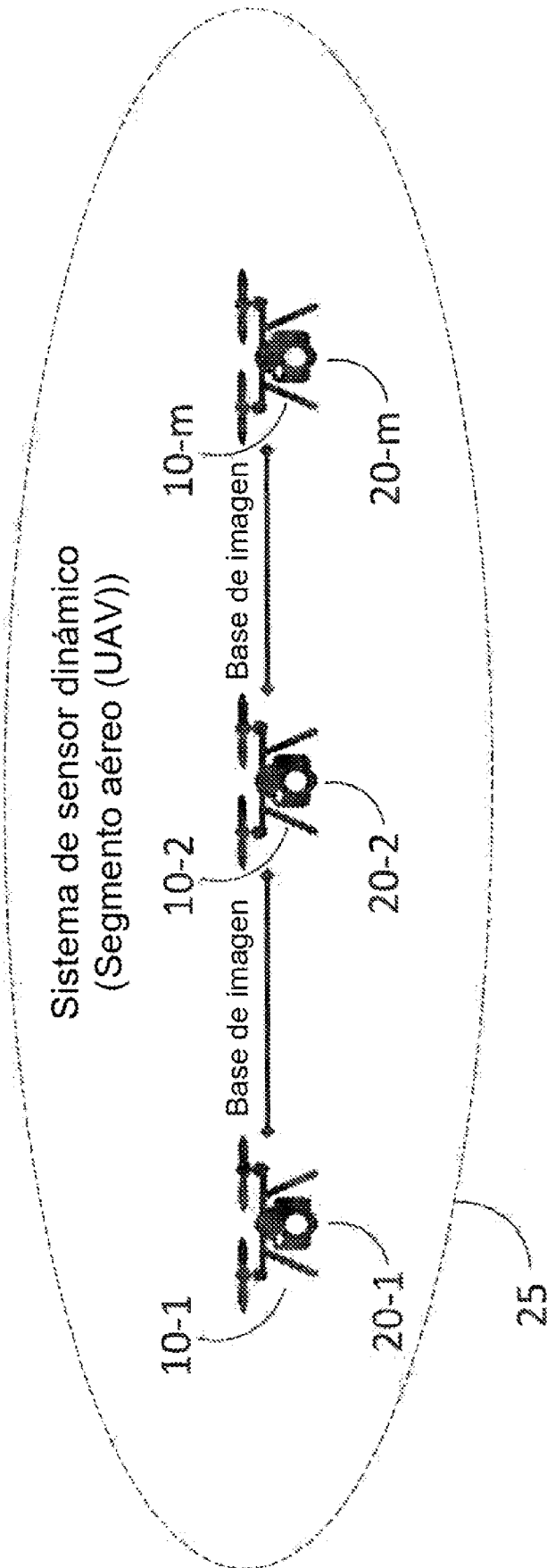


FIG. 3

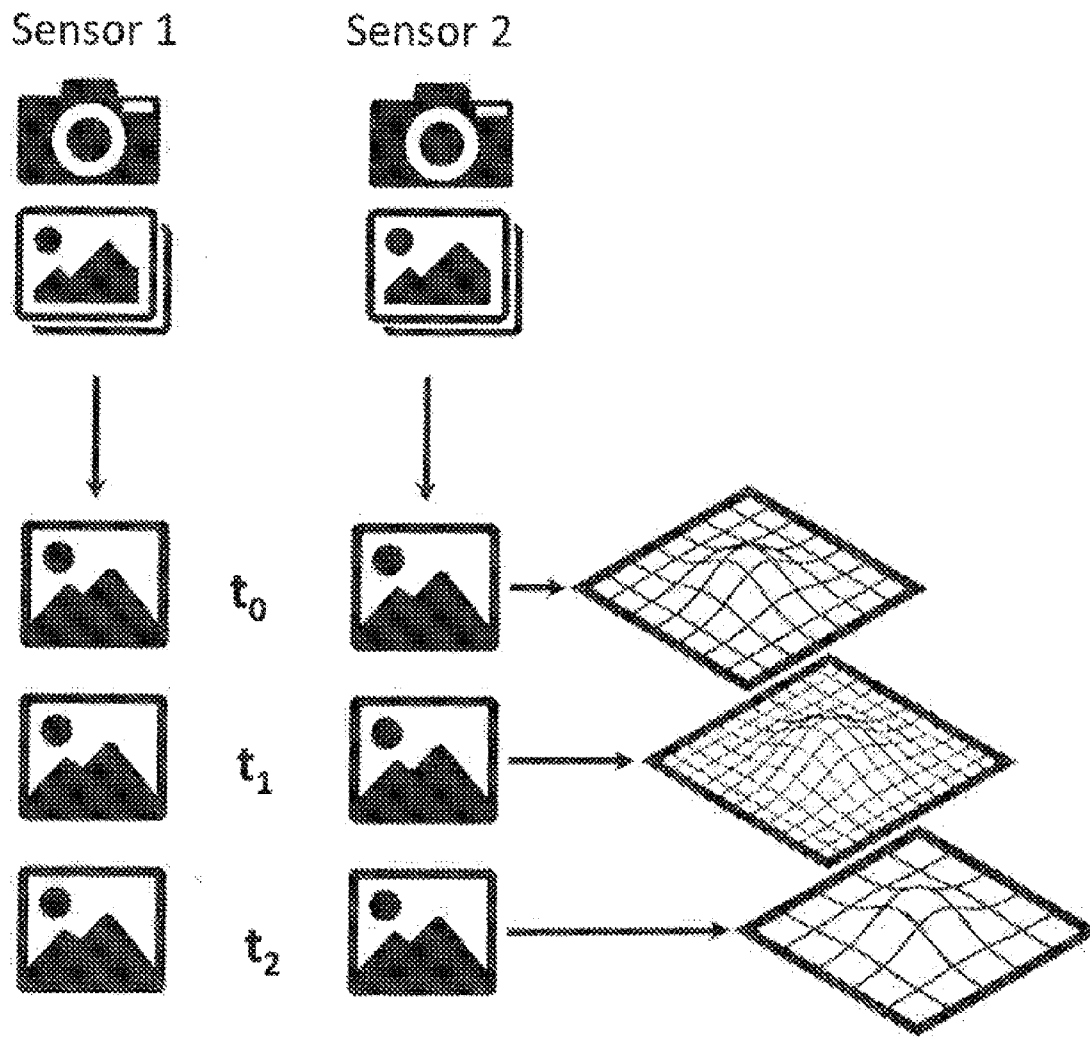


FIG. 4

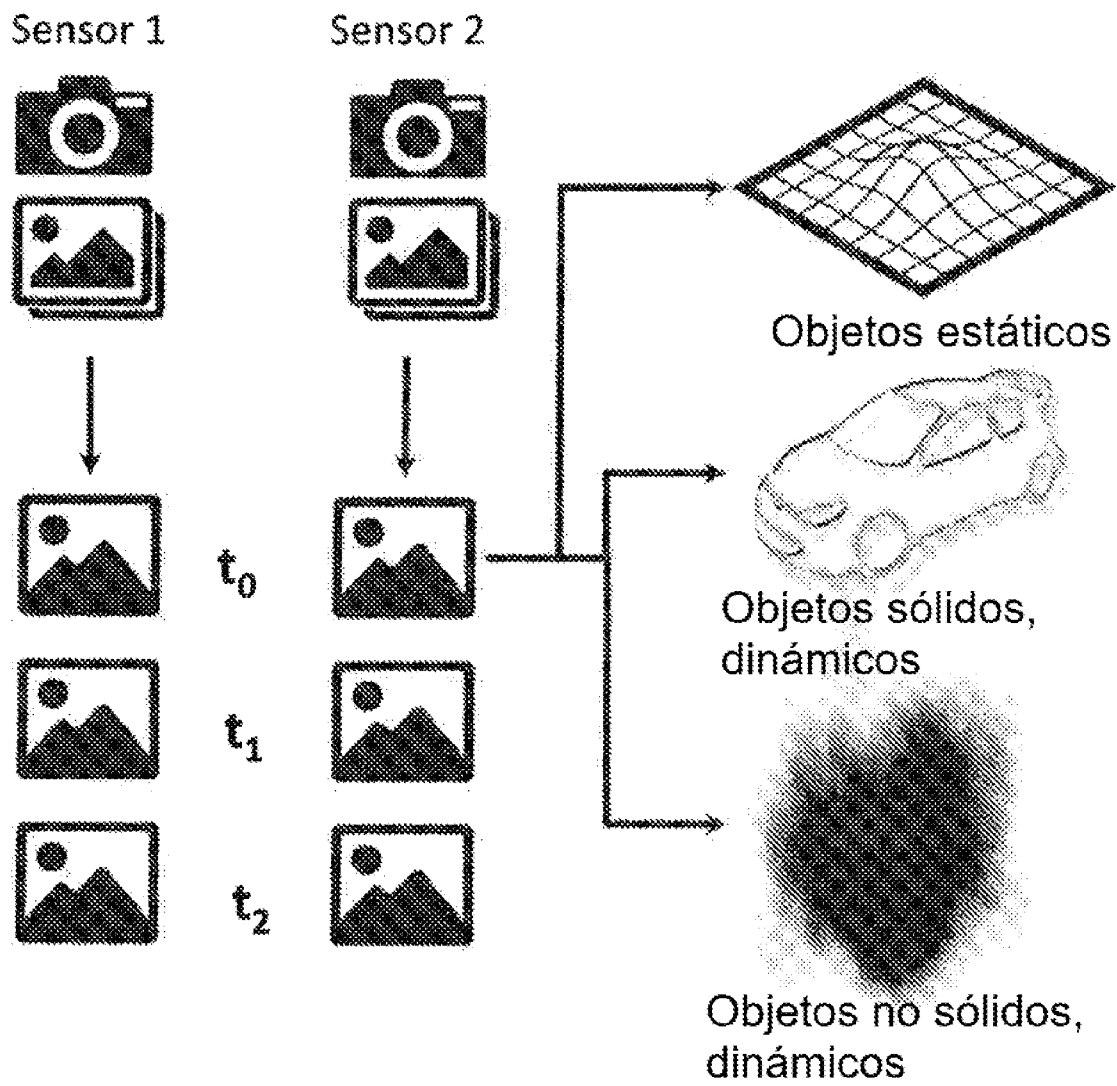


FIG. 5

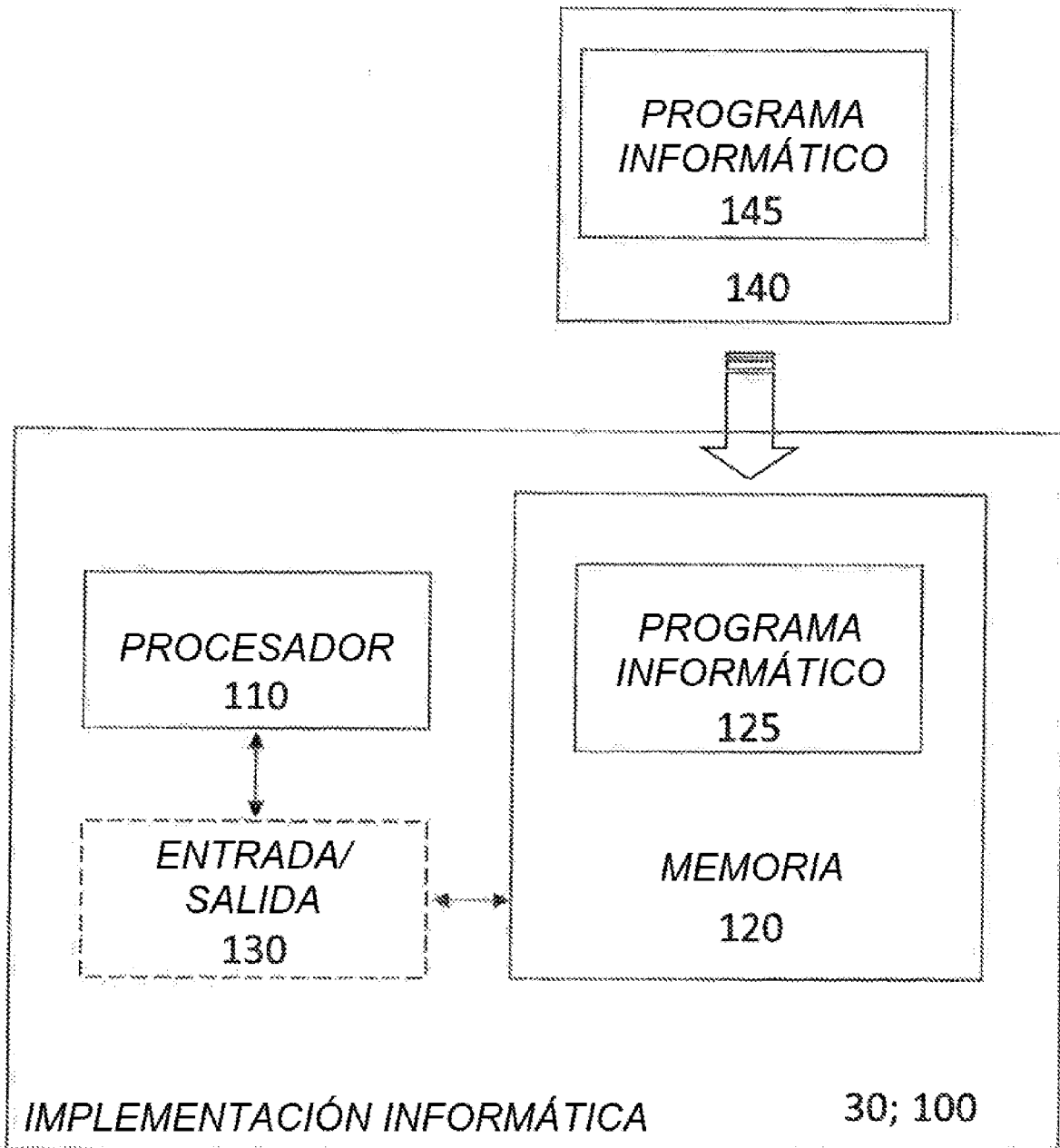


FIG. 6

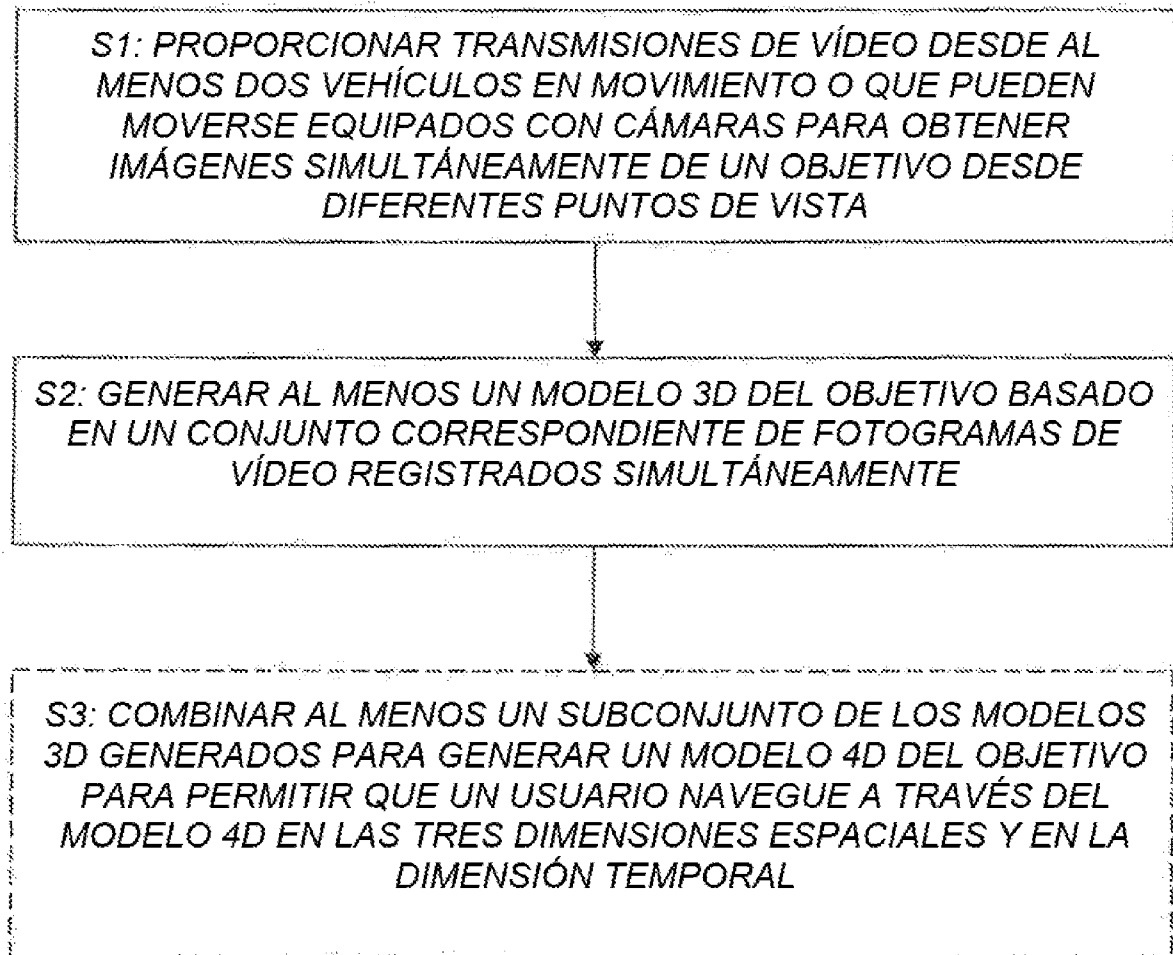


FIG. 7