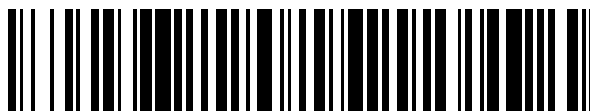


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 787 226**

51 Int. Cl.:

G06T 3/00 (2006.01)

G06T 3/40 (2006.01)

G06T 11/60 (2006.01)

G06T 7/00 (2007.01)

G06T 7/50 (2007.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.08.2007** **PCT/US2007/077242**

87 Fecha y número de publicación internacional: **06.03.2008** **WO08028040**

96 Fecha de presentación y número de la solicitud europea: **30.08.2007** **E 07814579 (4)**

97 Fecha y número de publicación de la concesión europea: **26.02.2020** **EP 2057585**

54 Título: **Imágenes oblicuas en mosaico y métodos para elaborar y usar las mismas**

30 Prioridad:

30.08.2006 US 840992 P

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

15.10.2020

73 Titular/es:

PICTOMETRY INTERNATIONAL CORP. (100.0%)
25 Methodist Hill Drive
Rochester, NY 14623, US

72 Inventor/es:

SCHULTZ, STEPHEN;
GIUFFRIDA, FRANK y
GRAY, ROBERT

74 Agente/Representante:

IZQUIERDO BLANCO, María Alicia

ES 2 787 226 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

Imágenes oblicuas en mosaico y métodos para elaborar y usar las mismas

5 ANTECEDENTES DE LA INVENCION

1. *Campo de la invención*

10 Las invenciones actualmente reivindicadas y divulgadas se refieren a imágenes oblicuas en mosaico y a métodos para elaborar y usar las mismas. Más particularmente, las invenciones actualmente reivindicadas y divulgadas usan una metodología mediante la cual se combinan imágenes aéreas capturadas de manera oblicua por separado en por lo menos una única imagen de mosaico oblicua. La por lo menos una sola imagen de mosaico oblicua es visualmente agradable y geográficamente precisa.

15 2. Antecedentes de la técnica

En la industria de teledetección/obtención de imágenes aéreas, las imágenes se usan para capturar vistas de un área geográfica y poder medir objetos y estructuras dentro de las imágenes, así como para poder determinar las ubicaciones geográficas de puntos dentro de la imagen. Generalmente, estas se denominan "imágenes georreferenciadas" y se dividen en dos categorías básicas:

1. Conjunto de imágenes capturadas-estas imágenes tienen la apariencia de que han sido capturadas por la cámara o el sensor empleado.
- 25 2. Conjunto de imágenes proyectadas-estas imágenes se han procesado y convertido de tal manera que se ajusten a una proyección matemática.

Toda conjunto de imágenes comienza como un conjunto de imágenes capturado, pero como la mayoría de software no puede geo-referenciar conjuntos de imágenes capturados, ese conjunto de imágenes se reprocesa luego para crear un conjunto de imágenes proyectado. La forma más común forma de conjunto de imágenes proyectado es la imagen orto-rectificadas. Este proceso alinea la imagen con una cuadrícula ortogonal o rectilínea (compuesta de rectángulos). La imagen de entrada usada para crear una imagen orto-rectificada es una imagen nadir – es decir, una imagen capturada con la cámara apuntando directamente hacia abajo.

35 A menudo es bastante deseable combinar múltiples imágenes en una imagen compuesta más grande de tal manera que la imagen cubra un área geográfica más grande en el suelo. La forma más común de esta imagen compuesta es la "imagen orto-mosaica", que es una imagen creada a partir de una serie de imágenes nadir superpuestas o adyacentes que se combinan matemáticamente en una única imagen orto-rectificada.

40 Cada imagen nadir de entrada, así como la imagen orto-mosaico de salida, se compone de píxeles discretos (elementos de imagen individuales) de información o datos. Como parte del proceso para crear una imagen orto-rectificada y, por lo tanto, una imagen de orto-mosaico, se hace un intento de reproyectar (mover dentro de un modelo matemático) cada píxel dentro de la imagen de tal manera que la imagen resultante aparezca como si cada píxel en la imagen fuese un píxel de nadir, es decir, que la cámara estuviese directamente encima de cada píxel en la imagen.

La razón por la que es necesario este proceso de orto-rectificación es que actualmente no es posible capturar una imagen en la que cada píxel esté en el nadir (directamente debajo) de la cámara a menos que: (1) la cámara usada sea tan grande como el área de captura, o (2) la cámara esté colocada a una distancia infinita sobre el área de captura de tal manera que el ángulo desde la cámara al píxel esté tan cercano directamente abajo que puede considerarse nadir. El proceso de orto-rectificación crea una imagen que se aproxima a la apariencia de ser capturada con una cámara donde el área en el suelo que captura cada píxel se considera nadir para ese píxel, es decir, directamente debajo de ese píxel. Este proceso se realiza creando un modelo matemático del suelo, generalmente en una cuadrícula rectilínea (una cuadrícula formada por rectángulos), y reproyectando de la imagen de la cámara capturada individual en esta cuadrícula rectilínea. Este proceso mueve los píxeles desde su localización no nadir relativa dentro de las imágenes individuales a sus posiciones nadir dentro de la cuadrícula rectilínea, es decir, la imagen se deforma para alinearse con la cuadrícula.

60 Cuando se crea un orto-mosaico, se usa este mismo proceso de orto-rectificación, sin embargo, en lugar de usar una única imagen nadir de entrada, se usa una colección de imágenes nadir superpuestas o adyacentes y se combinan para formar una única imagen orto-rectificada compuesta conocida como orto-mosaico. En general, el proceso de orto-mosaico implica los pasos siguientes:

- Se crea una cuadrícula rectilínea, que da como resultado una imagen de orto-mosaico donde cada píxel de la cuadrícula cubre la misma cantidad de área en el suelo.

- La localización de cada píxel de la cuadrícula se determina a partir de la definición matemática de la cuadrícula. Generalmente, esto significa que a la cuadrícula se le da una localización de inicio o de origen X e Y y un tamaño X e Y para los píxeles de la cuadrícula. Por tanto, la localización de cualquier píxel es simplemente la localización de origen más el número de píxeles multiplicado por el tamaño de cada píxel. En términos matemáticos: $X_{\text{píxel}} = X_{\text{origen}} + X_{\text{tamaño}} \times \text{Columna}_{\text{píxel}}$ e $Y_{\text{píxel}} = Y_{\text{origen}} + Y_{\text{tamaño}} \times \text{fila}_{\text{píxel}}$.
- Las imágenes nadir disponibles se comprueban para ver si cubren el mismo punto en el suelo que el píxel de la cuadrícula que se está rellenando. Si es así, se usa una fórmula matemática para determinar dónde se proyecta ese punto en el suelo sobre el mapa de imágenes de píxeles de la cámara y ese valor de píxel resultante se transfiere al píxel de la cuadrícula. Durante este proceso de selección, se realizan dos pasos importantes:
 - Al seleccionar la imagen que se usará para proporcionar el valor de píxel, se usa una fórmula matemática para seleccionar una imagen que minimice la inclinación del edificio, el efecto en el que los edificios parecen alejarse de la cámara. Esto se logra de varias maneras, pero la más común es elegir la imagen donde el píxel de la cuadrícula se reproyecta lo más cerca posible del centro de la cámara y, por lo tanto, lo más cerca posible del punto nadir de esa cámara.
 - Al determinar el valor de píxel de origen a usar, se tiene en cuenta la elevación del suelo para garantizar que se seleccione el valor de píxel correcto. Los cambios en la elevación hacen que la localización aparente del píxel cambie cuando es capturada por la cámara. Un punto en el suelo que está más alto aparecerá más lejos del centro de la imagen que un punto en el suelo en la misma localización que está más abajo. Por ejemplo, la parte superior de un edificio aparecerá más lejana del centro de una imagen que la parte inferior de un edificio. Teniendo en cuenta la elevación del suelo al determinar el valor del píxel de origen, el efecto neto es "aplanar" la imagen de manera que se eliminen los cambios en la localización del píxel debido a la elevación del suelo.

Como las cuadrículas rectilíneas usadas para el orto-mosaico son generalmente las mismas cuadrículas usadas para crear mapas, las imágenes de orto-mosaico tienen una sorprendente similitud con los mapas y, como tales, son generalmente muy fáciles de usar desde el punto de vista de la dirección y la orientación. Sin embargo, como tienen una apariencia dictada por proyecciones matemáticas en lugar de la apariencia normal que captura una sola cámara y porque se capturan mirando directamente hacia abajo, esto crea una visión del mundo a la que no estamos acostumbrados. Como resultado, muchas personas tienen dificultades para determinar qué es lo que están mirando en la imagen. Por ejemplo, podrían ver un rectángulo amarillo en la imagen y no darse cuenta de que lo que están viendo es la parte superior de un autobús escolar. O podrían tener dificultades para distinguir entre dos propiedades comerciales, ya que lo único que pueden ver de las propiedades en el orto-mosaico son sus tejados, donde la mayoría de las propiedades distintivas se encuentran a los lados de los edificios. Ha surgido toda una profesión, el intérprete fotográfico, para abordar estas dificultades, ya que estas personas tienen años de capacitación y experiencia específicamente en la interpretación de lo que están viendo en conjuntos de imágenes nadir o de orto-mosaico.

Como una imagen oblicua, por definición, se captura en ángulo, presenta una apariencia más natural ya que muestra los lados de los objetos y las estructuras, lo que estamos más acostumbrados a ver. Además, como las imágenes oblicuas generalmente no están orto-rectificadas, todavía mantienen el aspecto natural que captura la cámara en lugar de la construcción matemática de la imagen orto-mosaica. Esta combinación hace que sea muy fácil para las personas mirar algo en una imagen oblicua y darse cuenta de qué es ese objeto. No se requieren habilidades de interpretación fotográfica cuando se trabaja con imágenes oblicuas.

Las imágenes oblicuas, sin embargo, presentan otro problema. Como las personas han aprendido habilidades de navegación en los mapas, el hecho de que las imágenes oblicuas no estén alineadas con una cuadrícula del mapa, como las imágenes de orto-mosaico, las hace mucho menos intuitivas al intentar navegar o determinar la dirección de una imagen. Cuando se crea un orto-mosaico, como se crea en una cuadrícula rectilínea que generalmente es una cuadrícula de mapa, la parte superior de la imagen del orto-mosaico es Norte, el lado derecho es Este, la parte inferior es Sur y el lado izquierdo es Oeste. Así es como las personas generalmente están acostumbradas a orientarse y navegar en un mapa. Pero una imagen oblicua puede capturarse desde cualquier dirección y la parte superior de la imagen generalmente está "hacia arriba y hacia atrás", lo que significa que las estructuras verticales apuntan hacia la parte superior de la imagen, pero que la parte superior de la imagen también está más cercana al horizonte. Sin embargo, como la imagen puede capturarse desde cualquier dirección, el horizonte puede estar en cualquier dirección, Norte, Sur, Este, Oeste o cualquier punto intermedio. Si la imagen se captura de tal manera que la cámara apunta hacia el Norte, entonces el lado derecho de la imagen es el Este y el lado izquierdo de la imagen es el Oeste. Sin embargo, si la imagen se captura de tal manera que la cámara apunte hacia el Sur, entonces el lado derecho de la imagen es el Oeste y el lado izquierdo de la imagen es el Este. Esto puede causar confusión para alguien que intenta navegar dentro de la imagen.

Adicionalmente, como la cuadrícula de orto-mosaico es generalmente una cuadrícula rectilínea, por definición matemática, las cuatro direcciones cardinales de la brújula se encuentran en ángulo recto (90 grados). Pero con una imagen oblicua, como todavía está en la forma original que capturó la cámara y no ha sido reproyectada en un modelo matemático, no es necesariamente cierto que las direcciones de la brújula se encuentren en ángulo recto dentro de la imagen. Como en la perspectiva oblicua, te estás moviendo hacia el horizonte a medida que te mueves hacia arriba en la imagen, la imagen cubre un área más amplia en el suelo cerca de la parte superior de la imagen en comparación con el área en el suelo cubierta cerca de la parte inferior de la imagen. Si tuviese que pintar una cuadrícula rectangular en el suelo y capturarla con una imagen oblicua, las líneas a lo largo de la dirección en que apunta la cámara parecerían converger en la distancia y las líneas a través de la dirección en la que apunta la cámara parecerían estar más espaciadas en el frente de la imagen que en la parte posterior de la imagen. Esta es la vista en perspectiva que todos estamos acostumbrados a ver - las cosas son más pequeñas en la distancia que de cerca y las líneas paralelas, como las vías del ferrocarril, parecen converger en la distancia. Al contrario, si se crease una imagen de orto-mosaico sobre esta misma cuadrícula rectangular pintada, aparecería como una cuadrícula rectangular en la imagen de orto-mosaico ya que toda perspectiva se elimina como una parte incidental del proceso de orto-mosaico.

Debido a estas diferencias fundamentales en perspectiva y apariencia, la creación de una imagen de orto-mosaico mediante el proceso descrito anteriormente no funciona bien para imágenes oblicuas. Debido a que el eje óptico de la cámara (una línea imaginaria a través del centro de la lente o la óptica que sigue el objetivo de la cámara) apunta generalmente en un ángulo de 45 grados o más desde el nadir (apunta 45 grados o más hacia arriba desde abajo), los efectos de la inclinación de los edificios, las diferencias de elevación y los píxeles no cuadrados son todos exagerados, efectos que se consideran cualidades negativas en una imagen de orto-mosaico. En la industria del orto-mosaico, los requisitos se colocan generalmente en el proceso de captura de imágenes de tal manera que limitan la cantidad de oblicuidad a tan poco como 5 grados desde el nadir para minimizar cada uno de estos efectos negativos.

Además, si se deben mantener las propiedades admirables de una imagen oblicua, concretamente, ver los lados de las estructuras y la apariencia natural de las imágenes, entonces claramente un proceso que intenta eliminar los desplazamientos verticales y, por lo tanto, los lados de los edificios, y una que deforma la imagen para que se ajuste a una cuadrícula rectilínea no es una opción viable. Es necesario un nuevo proceso, uno que cumpla con las siguientes cualidades deseables en un esfuerzo por preservar las propiedades admirables de la imagen oblicua:

- Si se debe mantener la perspectiva oblicua, los píxeles no se pueden alinear a una cuadrícula rectilínea, ni siquiera a una cuadrícula trapezoidal. En cambio, los píxeles se alinean preferiblemente a la perspectiva natural que captura una cámara.
- Como parte de la perspectiva oblicua, los píxeles de la imagen no pueden medir el mismo tamaño en el suelo, ya que los píxeles en primer plano de la imagen cubren un área mucho más pequeña en el suelo que los píxeles en el fondo de la imagen - es decir por definición parte de la perspectiva natural de una cámara.
- Debido a que los píxeles están muy lejos del nadir, los efectos de la inclinación de los edificios se vuelven extremos y las soluciones estándar empleadas en el proceso de orto-mosaico no hacen un trabajo lo suficientemente adecuado para compensar este efecto - deben desarrollarse nuevas técnicas para compensar mejor este efecto.
- Si se retroceden los efectos de los cambios en la elevación, la imagen resultante tiene una apariencia muy poco natural - los lados verticales de los edificios pueden deformarse y retorcerse, algo que no estamos acostumbrados a ver y, por lo tanto, al mirar dicha imagen, tendemos a "rechazarla". Por tanto, para mantener los edificios, estructuras y objetos dentro de una imagen con un aspecto natural, es preferible dejar los efectos de elevación en la perspectiva de la imagen y, en su lugar, exponerlos de otra manera.

"Desarrollo de un sistema de asistencia de estacionamiento usando una síntesis de imagen de punto de vista virtual" por Okamoto et al. Congreso Mundial ITS, (1 de noviembre de 2000), XP002418339, describe un sistema de asistencia de estacionamiento que usa tecnología de síntesis de imágenes de punto de vista virtual. Este documento describe métodos para ayudar a un conductor proporcionando una vista aérea del vehículo y el área circundante, como un método para sintetizar una imagen virtual vista desde un punto de vista elevado arbitrario convirtiendo e integrando múltiples imágenes tomadas de cámaras montadas en el vehículo.

La EP 1223750 divulga un dispositivo de cámara para usar en la medición y síntesis de una imagen que puede usarse para una medición y síntesis precisas incluso sin suprimir la variación en las características del dispositivo. Un medio de almacenamiento de parámetros de cámara almacena un parámetro de cámara que indica características exclusivas para el dispositivo de cámara. Un medio de detección de estado detecta un estado del dispositivo de cámara como una temperatura. Una salida de parámetro significa que externamente emite un parámetro de cámara de acuerdo con la información de estado.

Debido a estos problemas, la práctica común en la industria es proporcionar conjuntos de imágenes oblicuas como una serie de imágenes individuales. Sin embargo, algunos de los mismos beneficios del orto-mosaico también se aplican a un mosaico oblicuo (una imagen creada a partir de una colección de imágenes oblicuas superpuestas o adyacentes), concretamente el hecho de que el mosaico cubre un área geográfica más grande que cada una de las imágenes individuales que se usaron para crearlo. Esta invención detalla un medio mediante el cual se puede crear un mosaico oblicuo de calidad, superando las limitaciones anteriores.

SUMARIO DE LA INVENCION

Esta invención permite la creación de una imagen de mosaico oblicua que tiene tanto una apariencia natural y además está preferiblemente georreferenciada para mantener la capacidad de medir y determinar coordenadas geográficas. Aunque la realización preferida aplica esta invención a las imágenes oblicuas aéreas, la invención también funcionará con imágenes oblicuas no aéreas capturadas de varias maneras, incluyendo, pero no limitado a, cámaras montadas oblicuamente en un poste vertical, cámaras de mano apuntadas oblicuamente, y cámaras montadas en ángulos oblicuos en una sonda submarina.

Se divulga un método para crear una imagen de mosaico oblicua a partir de una pluralidad de imágenes oblicuas de la fuente de un área geográfica. Las imágenes oblicuas de la fuente son preferiblemente imágenes oblicuas capturadas. En este método, se identifica un área deseada para ser fotografiada y recogida en una imagen de mosaico oblicua, y se crea un modelo matemático de un sensor de una cámara virtual. La cámara virtual tiene una elevación mayor que una elevación del área a ser fotografiada. El modelo matemático tiene un mapa de píxeles de mosaico oblicuo para el sensor del área deseada. Se determina una localización de superficie para cada píxel incluida en el mapa de píxeles de mosaico oblicuo. La presente invención comprende además el paso de seleccionar imágenes oblicuas de la fuente de la localización de la superficie capturada en un ángulo oblicuo y una dirección de la brújula similar al ángulo oblicuo y la dirección de la brújula de la cámara virtual. Luego, se reproyecta por lo menos un píxel de imagen oblicua de la fuente del área a ser fotografiada para cada píxel incluido en el mapa de píxeles de mosaico oblicuo para crear de este modo una imagen de mosaico oblicua del área geográfica deseada. Por tanto, la imagen de mosaico oblicua es una imagen compuesta formada a partir de píxeles reproyectados a partir de múltiples imágenes oblicuas de la fuente.

En una versión de la invención, las múltiples imágenes oblicuas de la fuente representan la misma localización de superficie. En esta realización, el método puede comprender además pasos para seleccionar qué píxeles y/o imágenes oblicuas de la fuente utilizar para formar la imagen de mosaico oblicua. En otra versión, los píxeles de cada imagen oblicua de la fuente que representan la misma localización de superficie se comparan para determinar qué píxel de la fuente es el más representativo de la localización de la superficie, el píxel más representativo que se incluirá en la imagen de mosaico oblicua.

En una versión de la invención, el paso de asignación se define además mediante la proyección de cada píxel a través de la perspectiva de la cámara virtual para determinar una localización de superficie correspondiente para cada píxel en el mapa de píxeles de mosaico oblicuo.

En otra versión de la invención, el paso de la reproyección se define además reproyectando el píxel para que coincida con el tamaño y la forma de la localización de la superficie representada, tomada desde la elevación, la dirección de la brújula, y el ángulo oblicuo de la cámara virtual.

En otra versión más de la invención, el paso de la reproyección se define adicionalmente eliminando los efectos de elevación de las imágenes oblicuas de la fuente antes de la reproyección y luego añadiendo los efectos de elevación a la imagen de mosaico oblicua después de la reproyección.

En otra versión más de la invención, el método incluye el paso de minimizar la inclinación de las estructuras verticales dentro de la imagen de mosaico oblicua. Esto puede lograrse de varias maneras, como creando un modelo de elevación a partir de las imágenes oblicuas de la fuente teniendo en cuenta las estructuras verticales, y reproyectar por lo menos un píxel de la imagen oblicua de la fuente del área de la que se va a captar la imagen utilizando el modelo de elevación. Alternativamente, puede minimizarse la inclinación de las estructuras verticales haciendo coincidir las estructuras verticales en múltiples imágenes oblicuas de la fuente y cambiando la localización aparente de los píxeles en por lo menos una de las imágenes oblicuas de la fuente en una altura relativa sobre un modelo de suelo.

BREVE DESCRIPCIÓN DE VARIAS VISTAS DE LOS DIBUJOS

La Fig. 1 es una vista esquemática de una escena capturada por una cámara real reproyectada a una "cámara virtual", es decir, un modelo matemático que describe una cámara más grande que la cámara real y de elevación más alta que la cámara real.

La Fig. 2 es una ilustración de un diagrama de flujo de software/función ejemplar de un método para elaborar una imagen de mosaico oblicua usando una metodología mediante la cual las imágenes aéreas capturadas oblicuamente separadas se combinan en la imagen de mosaico oblicua.

La Fig. 3a es una representación pictórica de una imagen de mosaico oblicua del estado de la técnica.

La Fig. 3b es una vista ampliada de la representación pictórica de la imagen de mosaico oblicua del estado de la técnica de la Fig. 3a.

La Fig. 4a es una representación pictórica de una imagen de mosaico oblicua ejemplar creada de acuerdo con la presente invención.

La Fig. 4b es una vista ampliada de la representación pictórica de la imagen de mosaico oblicua ejemplar creada de acuerdo con la presente invención de la Fig. 4a.

DESCRIPCIÓN DETALLADA DE LA INVENCION RECIENTEMENTE REVELADA Y RECLAMADA

Antes de explicar por lo menos una realización de la invención con detalle, debe entenderse que la invención no está limitada en su aplicación a los detalles de construcción, experimentos, datos ejemplares y/o la disposición de los componentes expuestos en la siguiente descripción o ilustrados en los dibujos. La invención es capaz de otras realizaciones o de ser puesta en práctica o llevada a cabo de varias maneras. Además, debe entenderse que la fraseología y la terminología empleadas en la presente tienen fines descriptivos y no deben considerarse limitativas.

Las invenciones actualmente reivindicadas y divulgadas se refieren a imágenes de mosaico oblicuo y a métodos para elaborar y usar las mismas. Más particularmente, las invenciones actualmente reivindicadas y divulgadas usan una metodología por la cual las imágenes aéreas capturadas oblicuamente separadas se combinan en por lo menos una única imagen de mosaico oblicua. La por lo menos una única imagen de mosaico oblicua es visualmente agradable y geográficamente precisa.

En referencia ahora a las Figuras y en particular a las Figs. 1 y 2, en la Fig. 1 se muestra y se designa por un número de referencia 10 una vista esquemática de un método para crear una imagen de mosaico oblicua 12 (ver Figuras 4A y 4B). La Fig. 2 es una ilustración de un diagrama de flujo de software/función ejemplar 14 de un método para elaborar la imagen de mosaico oblicua 12 usando una metodología mediante la cual las imágenes aéreas oblicuas capturadas se combinan en la imagen de mosaico oblicua 12. La Fig. 4a es una representación pictórica de una imagen de mosaico oblicua 12 ejemplar creada de acuerdo con la presente invención. La Fig. 4b es una vista ampliada de la representación pictórica de la imagen de mosaico oblicua 12 ejemplar creada de acuerdo con la presente invención.

En general, el método identifica un área deseada 15 para ser capturada en imágenes y recogida en la imagen de mosaico oblicua 12. Las imágenes de la fuente se obtienen utilizando una cámara real 14 que captura una escena 16 como se indica mediante un bloque 18 en la Fig. 2. Luego, se crea una cámara virtual 20 como se indica mediante un bloque 22 en la Fig. 2. La cámara virtual 20 es un modelo matemático que describe una cámara más grande que la cámara real 14 y más elevada que la cámara real 14 como se muestra en la Fig. 1. El modelo matemático tiene un mapa de píxeles de mosaico oblicuo para el sensor del área deseada. Se asigna una localización de superficie a cada píxel incluido en el mapa de píxeles oblicuo, y luego, como se indica mediante un bloque 24, se seleccionan las imágenes oblicuas de la fuente para las localizaciones de suelo y luego por lo menos un píxel de imagen oblicua de la fuente del área de la que se van a capturar imágenes para cada píxel incluido en el mapa de píxeles del mosaico oblicuo se re proyecta para crear de este modo la imagen de mosaico oblicua del área geográfica deseada (como se indica mediante un bloque 26). Opcionalmente, el método incluye además uno o más pasos para minimizar el efecto de la inclinación de la estructura vertical como se indica mediante un bloque 28.

Como se ha descrito anteriormente en la presente, el uso de una metodología de reproyección de cuadrícula para crear imágenes de mosaico oblicuas está plagado de problemas y es poco probable que proporcione una imagen utilizable. Por lo tanto, para producir la imagen de mosaico oblicua de calidad 12, debe realizarse un proceso mejorado. Tal proceso mejorado y único se describe y reivindica en la presente y preferiblemente usa, generalmente, las siguientes consideraciones:

- Primero, en lugar de proyectar sobre una cuadrícula rectilínea, los píxeles de la imagen de entrada se proyectan sobre la "cámara virtual" 20. La cámara virtual 20 es un modelo matemático que describe una cámara muy grande, en lo alto del cielo. Como es una creación matemática, esta cámara 20 no está sujeta a las limitaciones actuales de la fabricación de cámaras.
- En segundo lugar, los píxeles no se re proyectan para que sean cuadrados o incluso del mismo tamaño en la imagen de salida, como es el caso en el proceso estándar de orto-rectificación usado cuando se crea un

orto-mosaico. En cambio, se re proyectan para que coincidan con el tamaño que proyectarían sobre el suelo desde esta cámara virtual 20, es decir, la perspectiva natural que captura una cámara.

- En tercer lugar, para alinear mejor el conjunto de imágenes combinadas, los efectos de los cambios en la elevación se eliminan primero del conjunto de imágenes de entrada durante la proyección desde la localización de la cámara de la imagen de entrada original hasta el suelo, y luego los efectos de elevación se reintroducen durante la proyección desde el suelo hasta la localización de la cámara virtual. El resultado es una imagen de aspecto natural que muestra apropiadamente los contornos de la tierra como se ve desde una perspectiva oblicua natural.
- Finalmente, los efectos de la inclinación de los edificios deseablemente se minimizan. Hay varias maneras diferentes en las que puede minimizarse la inclinación de los edificios:
 - "Dirigiendo" las líneas de corte entre las imágenes de entrada por las calles de tal manera que las líneas de corte no pasen por un edificio, lo que puede hacer que la inclinación del edificio en conflicto distorsione en gran medida la apariencia de un edificio.
 - El ángulo de inclinación del edificio puede calcularse y las imágenes deformarse para compensar la inclinación del edificio.
 - Los edificios pueden hacerse coincidir en imágenes adyacentes y alinearse entre sí.

En la práctica, la metodología divulgada y reivindicada en la presente consiste de múltiples pasos y transformaciones de datos que pueden ser realizados por un experto en la técnica dada la presente especificación. Hay una serie de algoritmos ya conocidos en la técnica que dirigen líneas de corte para orto-mosaicos y podrían ser adaptados fácilmente para su uso con imágenes oblicuas. Además, el trabajo de seguimiento podría crear nuevos algoritmos diseñados específicamente para tratar las complejidades de las imágenes oblicuas.

El primer paso para crear la imagen de mosaico oblicua 12 de acuerdo con la invención actualmente divulgada y reivindicada requiere la selección de un área de la que se van a captar imágenes. En general, el área de la que se van a captar imágenes sería una localización geográfica específica. Sin embargo, también pueden seleccionarse otras áreas para captar imágenes en la imagen de mosaico oblicua 12, como los laterales de edificios, paredes, paisajes, laderas de montañas y similares. La creación de una "cámara virtual" 20 que sea capaz de cubrir todo el paisaje o área de interés de la que se van a captar imágenes.

Una vez que se ha determinado un área deseada para ser captada en imágenes y recogida en la imagen de mosaico oblicua 12, el usuario u operador crea la "cámara virtual" 20, es decir, una construcción matemática que es capaz de cubrir o capturar una parte o la totalidad del área deseada. La cámara virtual 20 es un modelo matemático de una cámara con los parámetros de geometría de cámara necesarios (los valores matemáticos que definen el modelo de la cámara, por ejemplo, el número de filas y columnas del plano del sensor, el tamaño del plano del sensor en milímetros, la distancia focal en milímetros, altura sobre el suelo, guiñada, inclinación y alabeo del eje óptico) que le permiten "capturar" preferiblemente la escena deseada. Por ejemplo, puede concebirse una cámara virtual que tenga un sensor muy grande (por ejemplo, columnas de 20.000 píxeles y filas de 10.000 píxeles), un campo de visión estándar (plano del sensor de 36 mm por 24 mm y una distancia focal de 60 mm), y "colocarse" a una altitud relativamente alta (por ejemplo, 30.000 pies) mirando hacia abajo en un ángulo oblicuo hacia el norte (guiñada y alabeo de 0 y cabeceo de -40 grados). En una realización preferida, se usa un modelo de sensor de una cámara real y el usuario simplemente modifica los parámetros de tal manera que cumpla los requisitos para "capturar" el área deseada.

El segundo paso crea el mapa de píxeles oblicuo resultante para la cámara virtual 20. El mapa de píxeles corresponde al sensor de la cámara virtual y, por tanto, típicamente pero no necesariamente, tiene el mismo número de filas y columnas que el sensor de la cámara virtual. Luego, para cada píxel en la imagen del mapa de píxeles, se usan ecuaciones proyectivas para la cámara virtual 20 para proyectar cada píxel hacia abajo y lejos de la cámara virtual 20 y sobre el suelo, teniendo en cuenta la elevación al hacerlo (generalmente a través del uso de un modelo de elevación matemático de la superficie del suelo). Esto da como resultado una localización de suelo correspondiente para el píxel de esa cámara virtual.

Una vez que se ha encontrado la localización del suelo correspondiente, puede usarse para seleccionar qué imágenes capturadas previamente contienen datos de imagen para esa localización del suelo. Esto se realiza generalmente comprobando si la localización del terreno se encuentra dentro de los límites de una imagen capturada previamente.

Cuando se seleccionan imágenes oblicuas de la fuente, por ejemplo, imágenes capturadas de entrada, para lograr una imagen de salida deseable, es importante usar imágenes oblicuas de la fuente que se capturaron en la misma orientación relativa, o casi la misma, de la cámara virtual 20, en términos de ángulo hacia abajo oblicuo y dirección de la brújula del eje óptico. Aunque generalmente no es un problema usar imágenes de entrada de una

cámara cuyo modelo es diferente al modelo de la cámara virtual, si ese modelo es radicalmente diferente (por ejemplo, un escáner de línea frente a un dispositivo de captura de fotograma completo), puede dar como resultado una imagen resultante indeseable.

Aunque esta invención analiza el uso de imágenes capturadas como entrada al mosaico oblicuo 12, en realidad no es necesario. Es posible usar una imagen proyectada como entrada para este proceso o incluso usar otro mosaico oblicuo como entrada para este proceso. Sin embargo, como este proceso re proyecta las imágenes de entrada, es deseable usar imágenes de entrada no proyectadas, es decir, imágenes capturadas. La razón es que la reproyección de datos ya proyectados a menudo puede llevar a artefactos, algo así como errores de redondeo en los cálculos matemáticos. Estos artefactos pueden crear un mosaico oblicuo resultante indeseable.

En general, es deseable crear el mosaico oblicuo continuo 12. Para hacerlo, debe haber datos de imágenes capturadas para toda el área "capturada" por la cámara virtual 20. Esto significa que si se combinan múltiples imágenes capturadas para crear el mosaico oblicuo 12, esas imágenes de entrada deben ser adyacentes o, más comúnmente, estar superpuestas. Como resultado de esta superposición, es común que haya múltiples imágenes capturadas que cubran la misma área en el suelo. Si hay múltiples imágenes capturadas disponibles para la selección, se elige una imagen capturada preferida de acuerdo con los criterios de selección que se describen a continuación.

Cuando varias imágenes de cámaras reales 14 cubren el mismo punto en el suelo, puede usarse un proceso de selección para determinar qué imagen de cámara real debe usarse como entrada para la creación de la imagen de mapa de píxeles de la cámara virtual. Este proceso de selección puede realizarse asignando ponderaciones (valores numéricos asignados) a los siguientes criterios de entrada, multiplicando esas ponderaciones por el criterio normalizado (un valor que se ha escalado entre 0 y 1), y luego seleccionando la imagen con la mayor suma de estos productos de ponderación/criterio. Aunque puede usar cualquier número de criterios, se han usado los siguientes tres criterios en el desarrollo de esta invención:

Criterio de selección: distancia al eje óptico

La distancia entre el punto en el suelo que se selecciona y el punto donde el eje óptico de la cámara de entrada interseca el suelo. Este valor puede normalizarse dividiendo la distancia por la distancia máxima que se puede medir en la escena.

Criterio de selección: diferencia angular al eje óptico

La diferencia entre los siguientes dos ángulos: el ángulo del eje óptico de la cámara de entrada (generalmente medido con respecto a la perpendicular) y el ángulo del rayo que se proyecta desde la cámara virtual al punto en el suelo que se selecciona (de nuevo, medido generalmente con respecto a la perpendicular). Este valor puede normalizarse dividiendo por 180 grados.

Criterio de selección: distancia a la línea central de la calle más cercana

La distancia entre el punto que se selecciona en el suelo y la línea central de la calle más cercana. Las líneas centrales de la calle pueden obtenerse de archivos de datos vectoriales como archivos TIGER u otros archivos del Sistema de Información Geográfica. Este valor puede normalizarse dividiendo por la distancia máxima que se puede medir en la escena.

Una vez que se ha seleccionado la imagen capturada preferida, se usan las ecuaciones proyectivas para la cámara de la imagen capturada para proyectar desde el suelo hasta la cámara, teniendo en cuenta la elevación del suelo al hacerlo. Esta proyección a través del punto focal y sobre el plano del sensor de la cámara encontrará una fila de píxeles y una columna correspondiente al punto en el suelo. Como esto normalmente no cae en una fila o columna entera, se usa la interpolación bilineal (una fórmula matemática estándar de la industria para encontrar un valor individual desde la proximidad proporcional a los cuatro píxeles circundantes) para encontrar el valor de píxel para el punto correspondiente en el plano del sensor de la cámara.

Este valor de píxel se usa para llenar el píxel en la imagen que corresponde al plano del sensor de la cámara virtual desde el cual se proyectó el rayo original sobre el suelo. Este proceso se repite para algunos o todos los píxeles restantes en el plano del sensor de la cámara virtual, lo que da como resultado una imagen que cubre algo del área o todo el área en el suelo que la cámara virtual 20 puede "ver". Preferiblemente, este proceso se repite para todos los píxeles restantes en el plano del sensor de la cámara virtual, lo que da como resultado una imagen completa que cubre toda el área en el suelo que la cámara virtual 20 puede "ver".

Luego se almacenan la imagen resultante y sus ecuaciones proyectivas correspondientes. La imagen resultante puede almacenarse en cualquier formato, incluyendo uno de los muchos formatos de imagen estándar de la industria, como TIFF, JFIF, TARGA, Windows Bitmap File, PNG o cualquier otro formato estándar de la industria.

Para las ecuaciones proyectivas correspondientes, los siguientes datos deberían, en una realización preferida, almacenarse como metadatos con la imagen resultante, adjuntada al mismo archivo o en otro archivo fácilmente accesible (una práctica estándar de la industria es usar el mismo nombre de archivo pero con una extensión de archivo diferente):

1. La localización de la cámara - en general, la localización de la cámara está especificada por una localización geográfica y altitud (o altura sobre el suelo).

2. La orientación de la cámara - generalmente especificada por tres ángulos: guiñada, cabeceo y alabeo (u omega, phi y kappa).

3. El tamaño del sensor - generalmente especificado como el tamaño total del sensor en milímetros (o por el tamaño de un elemento sensor individual).

4. La distancia focal de la lente - generalmente especificada en milímetros.

5. Opcionalmente, información de distorsión de la lente - generalmente especificada como un desplazamiento del punto principal (desplazamiento entre el eje óptico y el centro del sensor) y como términos de distorsión radial (coeficientes polinómicos que describen la cantidad de distorsión que se produce radialmente fuera del eje óptico).

6. Opcionalmente, un modelo de elevación del suelo - generalmente especificado como una cuadrícula de valores de elevación.

Como se ha tratado anteriormente, la perspectiva relativa de la cámara provoca un efecto conocido como "inclinación de edificios". Aunque la inclinación de edificios se aplica más comúnmente y, por lo tanto se analiza, a edificios, también se aplica a cualquier estructura vertical en el objeto, como torres eléctricas, árboles, automóviles, cabinas telefónicas, buzones, letreros de calles, etc. La inclinación de edificios es un efecto que hace que parezca que los edificios que no están a lo largo del eje óptico de la cámara se "inclinan" lejos del eje óptico - cuanto más lejos del eje óptico, mayor es la inclinación. Esta inclinación es el resultado de la perspectiva, que hace que los objetos que se levantan del suelo aparezcan más "atrás" en la imagen, lejos de la cámara. Por tanto, la parte superior de un edificio aparece más atrás que la parte inferior de un edificio. Cuando esta inclinación corresponde a la perspectiva de la cámara, parece normal. Sin embargo, como parte del proceso de mosaico oblicuo, las imágenes capturadas, cada una con su propia perspectiva, se combinan en la perspectiva de una única cámara virtual.

Esta combinación de diferentes perspectivas se vuelve problemática, especialmente cuando dos imágenes capturadas diferentes desde diferentes puntos de vista aportan valores de píxeles a áreas adyacentes en la imagen del mapa de píxeles de la cámara virtual. Por ejemplo: si una cámara a la izquierda de un edificio proporciona el lado izquierdo de un edificio, entonces esa parte del edificio parecerá inclinarse hacia la derecha (lejos de la cámara). Si una cámara que está localizada a la derecha del edificio proporciona el lado derecho del edificio, entonces esa parte del edificio parecerá inclinarse hacia la izquierda. Como las dos mitades del edificio "se inclinan" entre sí, la imagen combinada resultante tiene un edificio que tiene una apariencia triangular, en lugar de rectangular.

Como la inclinación del edificio solo afecta a las superficies por encima de la superficie del suelo, generalmente es bastante difícil tener en cuenta o corregir estos efectos ya que, para hacerlo, el usuario debe tener conocimiento de la presencia del edificio o estructura. Las características que se encuentran en el suelo no experimentan esta inclinación de edificios ya que el cambio en la perspectiva relativa se revierte cuando se tiene en cuenta la elevación del suelo en las ecuaciones proyectivas. El modelo matemático usado para definir la superficie del suelo durante el proceso proyectivo asegura que se seleccione la localización del suelo correcta. Sin embargo, para los objetos que se elevan por encima de la superficie del suelo y no están representados en el modelo matemático usado para definir la superficie del suelo, este cambio de perspectiva relativo hace que la cámara virtual "vea" la parte superior del edificio en el lugar equivocado, es decir, demasiado atrás de la cámara.

Un método para minimizar los efectos de la inclinación de edificios, como se contempla en la presente, es hacer la transición entre una imagen de cámara de entrada y la siguiente imagen de cámara de entrada sobre un área donde no hay estructuras por encima o por debajo del modelo de elevación del suelo. En una realización, esto se logra colocando la transición en el medio de una calle. Por tanto, al tener un criterio de selección apropiadamente ponderado para la distancia a la línea central de la calle, si hay una calle en el área donde se superponen dos imágenes capturadas, se producirá la transición de píxeles contribuyentes de una imagen capturada a los píxeles contribuyentes de la segunda imagen capturada a lo largo de esta calle, minimizando de este modo los efectos de la inclinación de edificios.

Un método para eliminar la inclinación de edificios por completo, como se contempla en la presente, es proporcionar un modelo de elevación del suelo preciso que tenga en cuenta los edificios y otras estructuras verticales. Por tanto, cada píxel que comprende la imagen del edificio se representa en el modelo matemático del

modelo de elevación del suelo y, por lo tanto, se tienen en cuenta el cambio en la perspectiva relativa para el proceso proyectivo. Sin embargo, para que esto funcione bien, el modelo de elevación debe estar altamente correlacionado con las imágenes de entrada. Si hay algún cambio en la localización entre el conjunto de imágenes y el modelo de elevación, entonces los edificios no se proyectarán apropiadamente cuando se crea la imagen de mosaico oblicua 12.

Para superar esta limitación, la metodología preferida es crear el modelo de elevación a partir del propio conjunto de imágenes. Esto se puede hacer usando un proceso estándar de la industria conocido como aerotriangulación, que encuentra la elevación de un punto en el suelo comparando su localización en dos imágenes capturadas superpuestas y usando las ecuaciones proyectivas de sus cámaras correspondientes para triangular su localización y elevación. La repetición de este proceso en toda el área de superposición puede producir un modelo matemático preciso de la superficie no solo del suelo, sino también de la superficie de las estructuras y los objetos dentro de las imágenes. De modo más importante, debido a que este modelo se deriva del propio conjunto de imágenes está, por definición, altamente correlacionado con la imagen de entrada.

Otro método para eliminar la inclinación de los edificios, contemplado para su uso en la presente, es intentar identificar estructuras verticales mediante el uso de un proceso de coincidencia de bordes entre el conjunto de imágenes oblicuo y el correspondientes al nadir. Las estructuras verticales no aparecen en una imagen verdaderamente nadir, y apenas aparecen en una imagen ligeramente fuera del nadir. Por tanto, al comparar una imagen oblicua con su imagen de nadir correspondiente, la diferencia principal entre las estructuras que aparecen en las dos imágenes serán las estructuras verticales. Usando uno o más algoritmos de detección de bordes (como un filtro laplaciano estándar de la industria), es posible identificar las varias estructuras dentro de las dos imágenes y luego aislar los bordes verticales en la imagen oblicua restando las estructuras no verticales que también aparecen en la imagen nadir. Una vez que se han encontrado estas estructuras verticales, los píxeles de esas estructuras verticales pueden desplazarse para eliminar los efectos del cambio en perspectiva relativa. Al cambiar la localización aparente del píxel en la imagen oblicua capturada por la altura relativa sobre el modelo de suelo que se encuentra a través de la medición de los bordes verticales, puede determinarse su localización de suelo adecuada, negando de este modo los efectos de la inclinación del edificio.

Debe entenderse que los procesos descritos anteriormente pueden realizarse con la ayuda de un sistema informático que ejecuta un software de procesamiento de imágenes adaptado para realizar las funciones descritas anteriormente, y las imágenes y los datos resultantes se almacenan en uno o más medios legibles por ordenador. Los ejemplos de un medio legible por ordenador incluyen un dispositivo de almacenamiento óptico, un dispositivo de almacenamiento magnético, un dispositivo de almacenamiento electrónico o similar. El término "sistema informático", como se usa en la presente, significa un sistema o sistemas que pueden incorporar y/o ejecutar la lógica de los procesos descritos en la presente. La lógica incorporada en forma de instrucciones de software o firmware puede ejecutarse en cualquier hardware apropiado que puede ser un sistema o sistemas dedicados, o un sistema informático de propósito general, o un sistema informático de procesamiento distribuido, todo lo cual es bien entendido en la técnica, y una descripción detallada de cómo hacer o usar tales ordenadores no se considera necesaria en la presente. Cuando el sistema informático se usa para ejecutar la lógica de los procesos descritos en la presente, tales ordenadores y/o ejecución pueden realizarse en una misma localización geográfica o en múltiples localizaciones geográficas diferentes. Además, la ejecución de la lógica puede realizarse de forma continua o en múltiples momentos discretos. Además, dicha lógica puede realizarse aproximadamente simultáneamente con la captura de las imágenes, o posteriormente o combinaciones de las mismas.

Aunque la invención anterior se ha descrito con cierto detalle a modo de ilustración y ejemplo con propósitos de claridad de comprensión, será obvio para los expertos en la técnica que pueden ponerse en práctica ciertos cambios y modificaciones sin apartarse del alcance de la misma, como se describe en esta especificación y como se define en las reivindicaciones adjuntas a continuación.

REIVINDICACIONES

1. Un método para crear una imagen de mosaico oblicua (12) a partir de una pluralidad de imágenes oblicuas de la fuente (16) tomadas a una primera elevación por una cámara real, que comprende los pasos de:

identificar un área geográfica deseada (15) para ser captada en imágenes y recogida en una imagen de mosaico oblicua (12);
 crear un modelo matemático de un sensor de una cámara virtual (20), donde la cámara virtual (20) tiene una segunda elevación mayor que una elevación del área a ser captada en imágenes (15), la segunda elevación más alta que la primera elevación, el modelo matemático teniendo un mapa de píxeles de mosaico oblicuo para el sensor del área geográfica deseada (15);
 determinar una localización de superficie para cada píxel incluido en el mapa de píxeles de mosaico oblicuo; seleccionar imágenes oblicuas de fuente (16) de la localización de la superficie capturadas en un ángulo oblicuo y dirección de la brújula que son iguales, o casi iguales al ángulo oblicuo y la dirección de la brújula de la cámara virtual (20); y
 reproyectar por lo menos un píxel de la imagen oblicua de la fuente del área a ser captada en imágenes (16) para cada píxel incluido en el mapa de píxeles de mosaico oblicuo para crear de este modo una imagen de mosaico oblicua (12) del área geográfica deseada (15).

2. El método de la reivindicación 1, en el que el paso de determinar se define adicionalmente proyectando cada píxel a través de la perspectiva de la cámara virtual (20) para determinar una localización de superficie correspondiente para cada píxel en el mapa de píxeles de mosaico oblicuo.

3. El método de la reivindicación 1, en el que las imágenes oblicuas de múltiples fuentes (16) representan la misma localización de superficie.

4. El método de la reivindicación 3, en el que los píxeles de cada imagen oblicua de la fuente (16) que representan la misma localización de superficie se comparan para determinar qué píxel de la fuente es más representativo de la localización de superficie, el píxel más representativo se incluirá en la imagen de mosaico oblicua (12).

5. El método de la reivindicación 1, en el que el paso de reproyección se define además reproyectando el píxel para que coincida con el tamaño y la forma de la localización de la superficie representada, tomada desde la elevación, la dirección de la brújula y el ángulo oblicuo de la cámara virtual (20).

6. El método de la reivindicación 1, en el que el paso de reproyección se define adicionalmente eliminando los efectos de elevación de las imágenes oblicuas de la fuente (16) antes de la reproyección y luego añadiendo los efectos de elevación a la imagen de mosaico oblicua (12) después de la reproyección.

7. El método de la reivindicación 1, que comprende además el paso de minimizar la inclinación de las estructuras verticales dentro de la imagen de mosaico oblicua (12).

8. El método de la reivindicación 7, en el que el paso de minimizar la inclinación de las estructuras verticales dentro de la imagen de mosaico oblicua (12) se define además como los pasos de crear un modelo de elevación a partir de las imágenes oblicuas de la fuente (16) teniendo en cuenta las estructuras verticales, y reproyectando por lo menos un píxel de la imagen oblicua de la fuente del área a ser captada en imágenes (15).

9. El método de la reivindicación 7, en el que el paso de minimizar la inclinación de las estructuras verticales se define además haciendo coincidir estructuras verticales en múltiples imágenes oblicuas de la fuente (16) y desplazando la localización aparente de píxeles en por lo menos una de las imágenes oblicuas de la fuente (16) por una altura relativa sobre un modelo de suelo.

10. El método de la reivindicación 1, en el que los metadatos se almacenan con la imagen de mosaico oblicua (12).

11. El método de la reivindicación 10, en el que los metadatos comprenden la localización de la cámara, la orientación de la cámara, el tamaño del sensor y la distancia focal de la lente de la cámara.

12. El método de cualquiera de las reivindicaciones 1-11, en el que la imagen de mosaico oblicua (12) está georreferenciada.

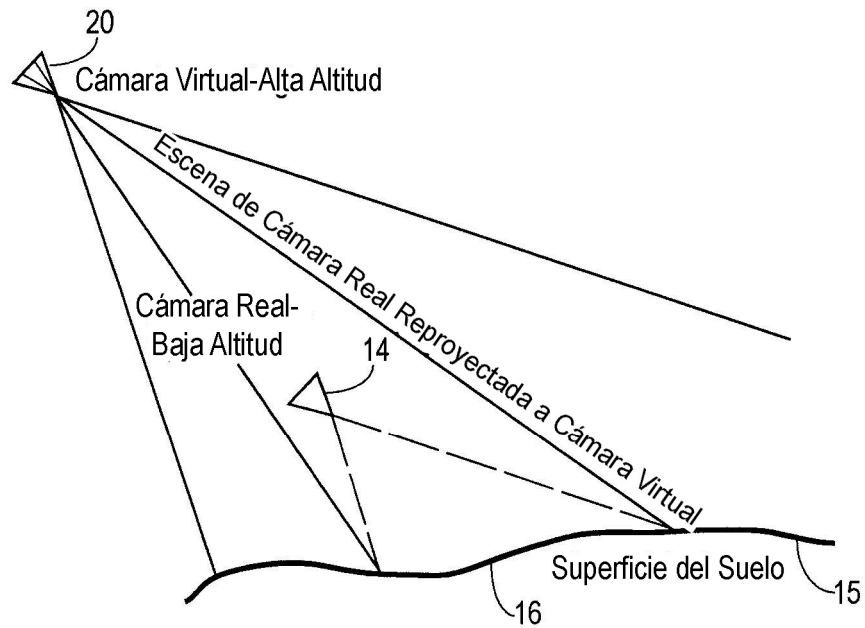


Fig. 1

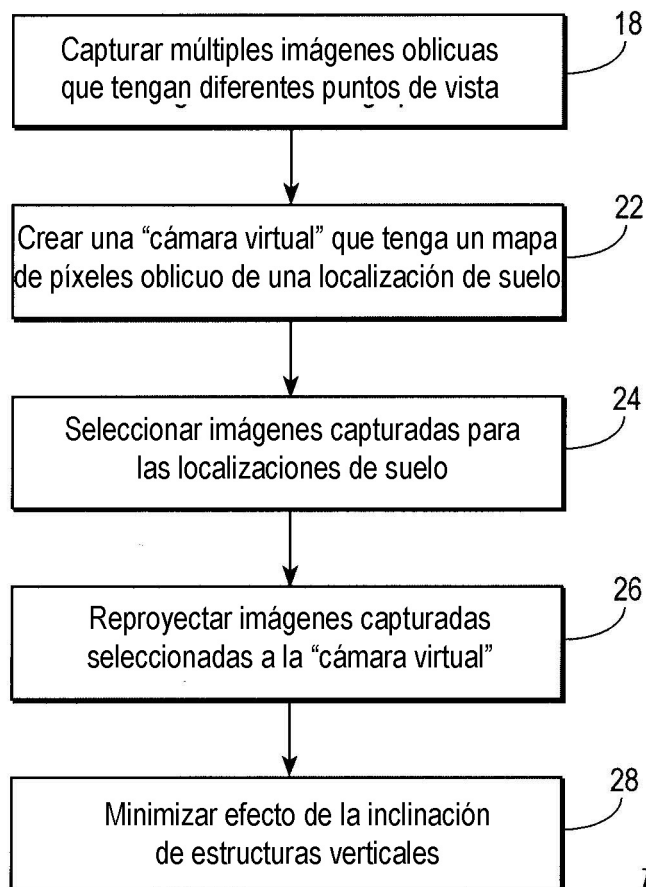


Fig. 2

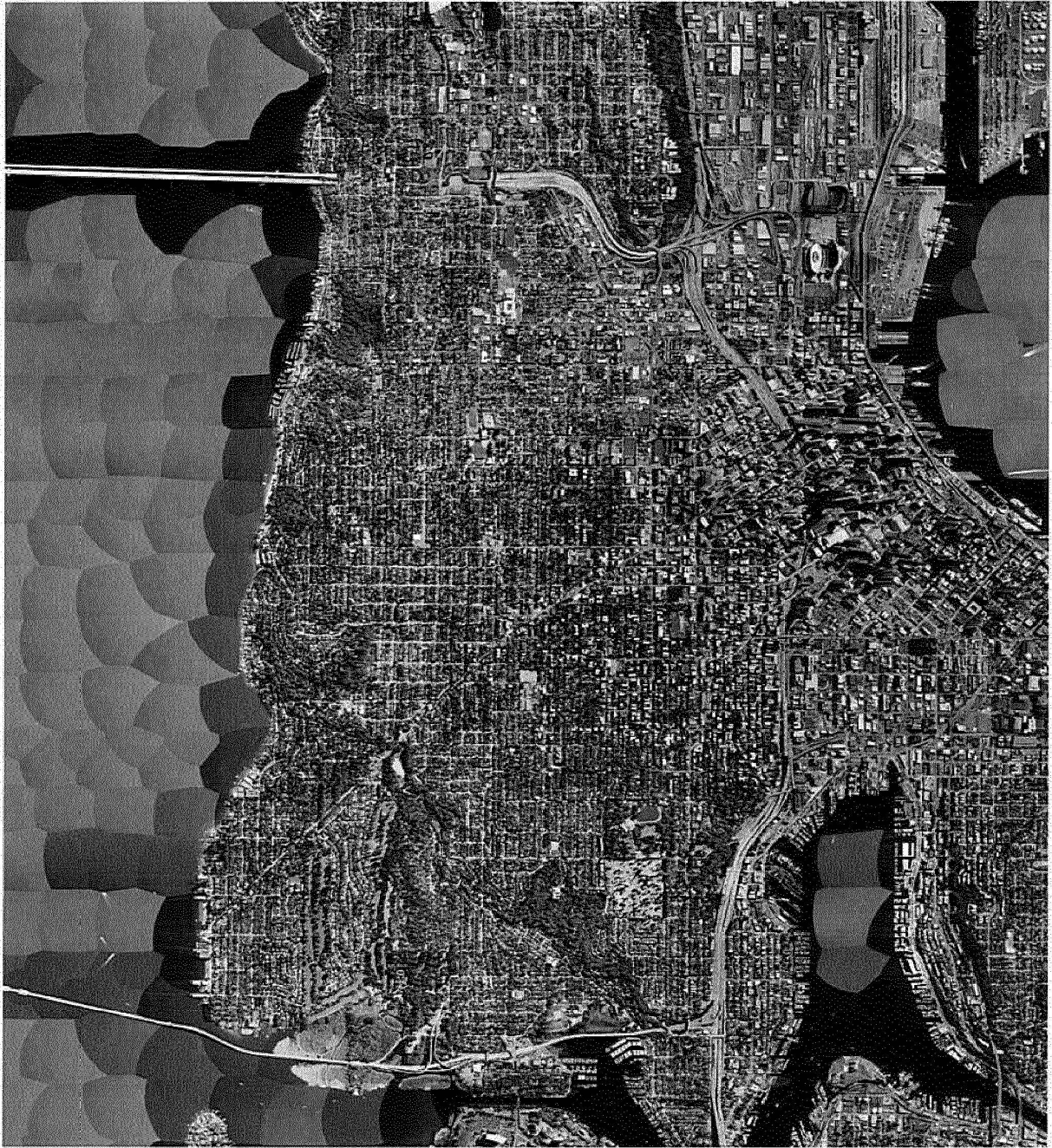


FIG. 3A

Oblicua-Rectificada

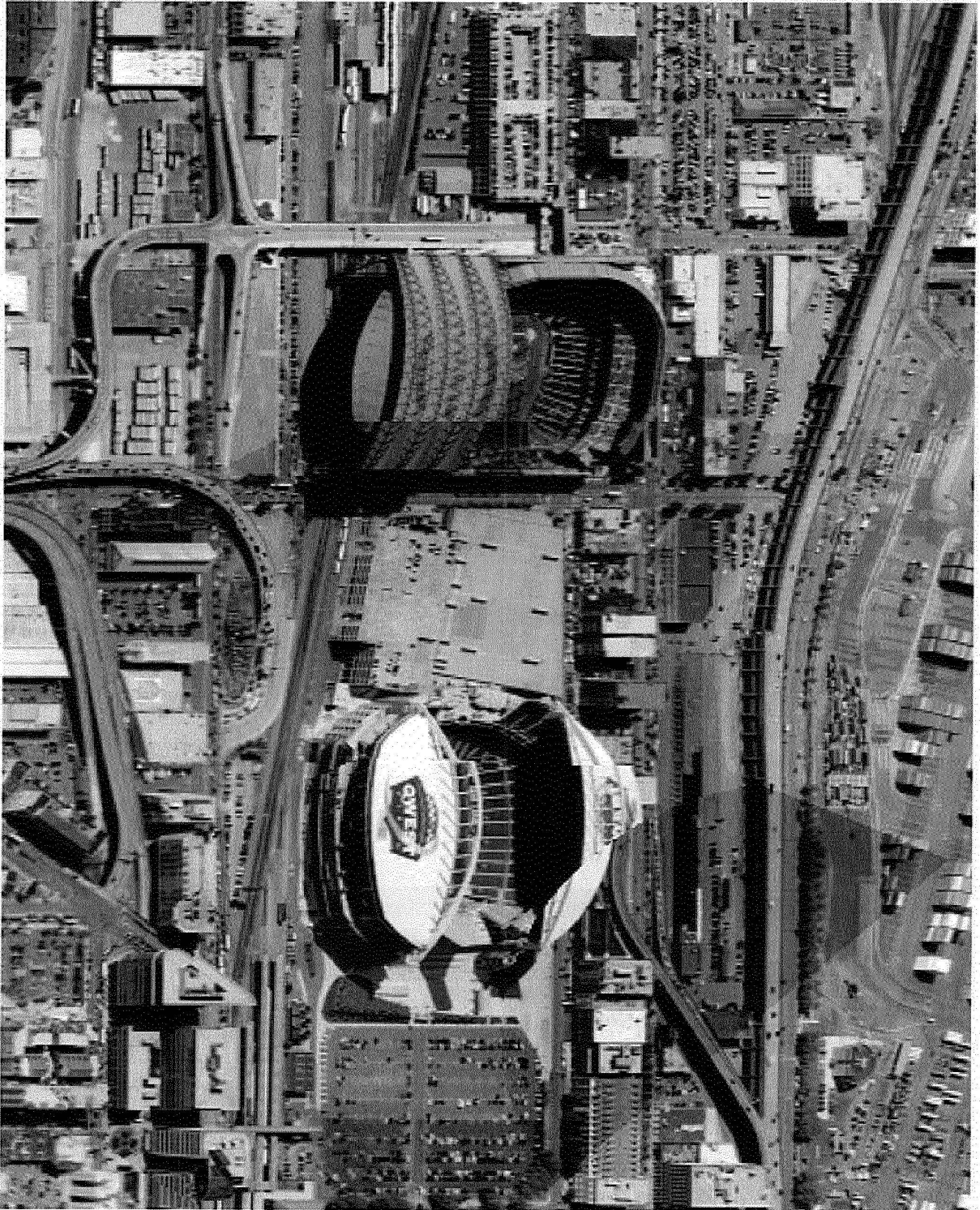


FIG. 3B

Oblicua-Rectificada Primer plano

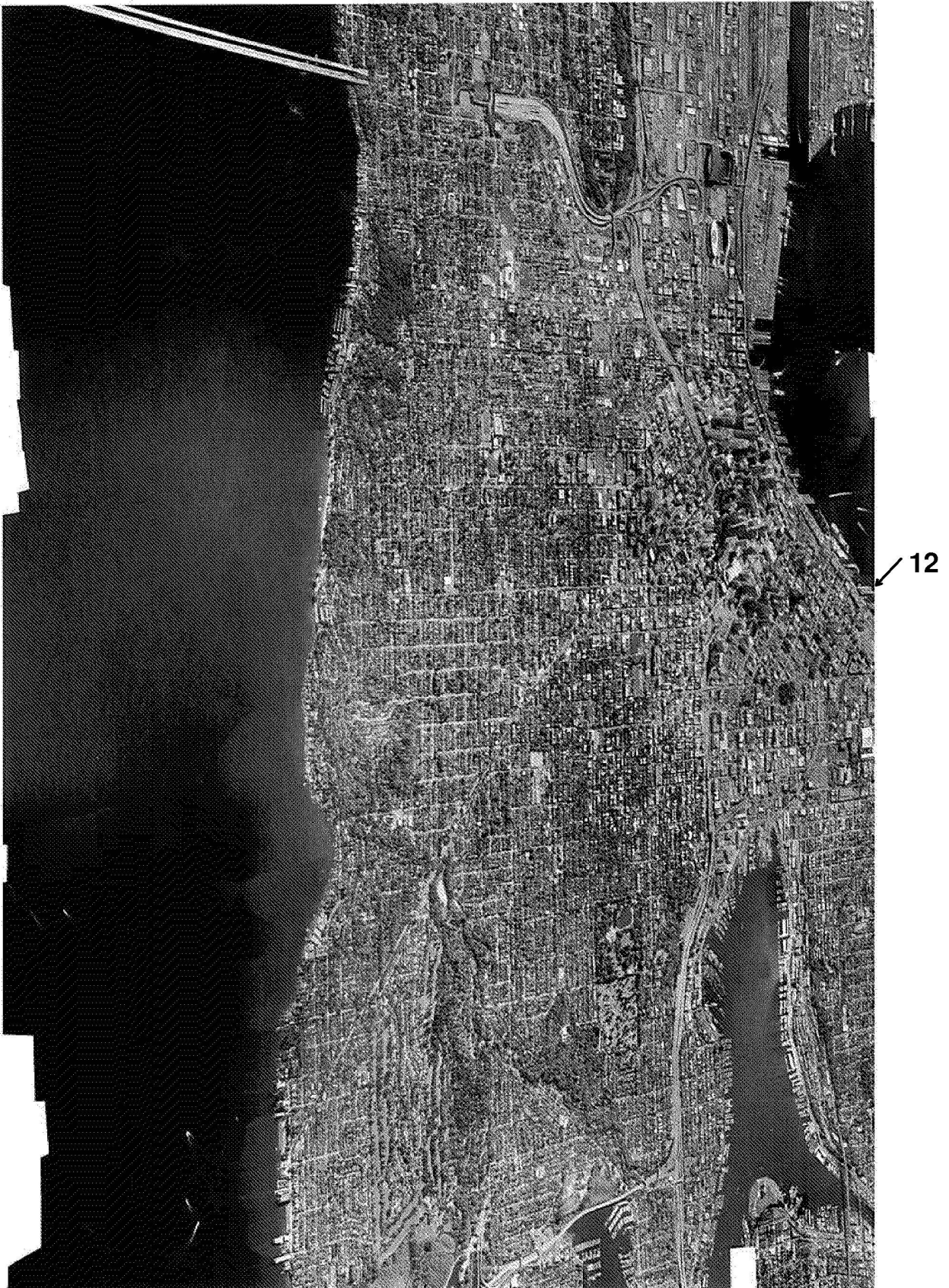


Fig. 4A

Mosaico Oblicuo



FIG. 4B

Mosaico Oblicuo Primer plano