(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 704 077**

(51) Int. CI.:

G06T 7/00 (2007.01)
G06T 15/08 (2011.01)
G06T 19/20 (2011.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **17.12.2015 PCT/EP2015/080258**

(87) Fecha y número de publicación internacional: **23.06.2016 WO16097168**

(96) Fecha de presentación y número de la solicitud europea: **17.12.2015 E 15820476 (8)**

(97) Fecha y número de publicación de la concesión europea: **10.10.2018 EP 3234914**

(54) Título: **Procedimiento de discriminación y de identificación de objetos de una escena por procesamiento de imagen en 3D**

(30) Prioridad:

19.12.2014 FR 1402929

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
14.03.2019

(73) Titular/es:

THALES (33.3%)
Tour Carpe Diem, Place des Corolles, Esplanade Nord
92400 Courbevoie, FR;
UNIVERSITÉ DE LORRAINE (33.3%) y
SISPIA (33.3%)

(72) Inventor/es:

BERECHET, STEFAN;
BERECHET, ION;
BELLET, JEAN-BAPTISTE y
BERGINC, GÉRARD

(74) Agente/Representante:

CARPINTERO LÓPEZ, Mario

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

Procedimiento de discriminación y de identificación de objetos de una escena por procesamiento de imagen en 3D

El campo de la invención es el de la discriminación y de la identificación de objetos cualesquiera que pueden estar disimulados en una escena compleja como la ilustrada en la figura 4. Estos objetos pueden enmascararse unos a otros, lo que convierte a la identificación y a la discriminación de los objetos en particularmente difíciles.

Existen diferentes tipos de visualización tridimensional. Ciertas visualizaciones ofrecen una imagen de profundidad, o imagen 2,5 D. Estas técnicas no permiten recuperar el conjunto de los vóxeles ("píxeles volumétricos") tridimensionales, se basan en técnicas estereoscópicas vinculadas a la optimización de una diferencia de recorrido del rayo óptico o electromagnético entre los diferentes puntos de dos imágenes; esta diferencia de recorrido es calculable si los objetos se sitúan a corto alcance del sistema de procesamiento de imagen. La identificación está limitada por tanto a las cortas distancias.

Ciertas técnicas de procesamiento de imagen necesitan un barrido del haz láser o de luz sobre la escena, y la secuencia de medidas da una nube de puntos 3D. Son necesarios varios barridos en varios ángulos de visión para tener una información tridimensional pero el movimiento relativo del portador que contiene el sistema de barrido produce una distorsión de nube de puntos a nube de puntos, el recalibrado es complejo y degrada la resolución y la identificación.

Las técnicas llamadas de siluetas obtenidas a partir de múltiples vistas dan una envolvente exterior poco detallada. Si los objetos contienen una zona de sombra o si el objeto se encuentra en la sombra de otro objeto, una gran parte de la envolvente se pierde por lo que no es realizable la identificación y la discriminación de los dos objetos.

Las tecnologías de tipo Spin Image necesitan unas bases de datos y no pueden aplicarse para objetos disimulados en escenas complejas sin conocimiento a priori.

Las técnicas de reagrupamiento de vóxeles vinculadas a una reconstrucción de la superficie por isodensidad permiten obtener una superficie externa de los objetos pero eliminan las informaciones internas contenidas en el objeto.

Las técnicas de reconstrucción tridimensional vinculadas a una optimización de las bases de conocimientos y a la extracción de señales débiles y de las que se presenta un ejemplo en el documento "Procédé de reconstruction synthétique d'objets illuminés par une onde électromagnétique et/ou excités par une onde élastique" (EP 2 929 421 o US 8 345 960), necesitan bases de datos y que estas estén optimizadas con el fin de realizar la mejor reconstrucción tridimensional posible. Las bases de conocimiento no representan frecuentemente los objetos más que por un conjunto de las superficies tridimensionales externas que pueden ser modelizadas por unas facetas. Las estructuras transparentes (vidrios…) son mal tenidas en cuenta, lo que perjudica grandemente la visualización completa del objeto en 3D. Estas técnicas son también dependientes de la disponibilidad de las bases de conocimiento.

La técnica MIP (Maximum Intensity Projection) permite la visualización en 2D de datos tridimensionales. Esta técnica proyecta los vóxeles sobre un plano de proyección; los vóxeles se determinan por los radios que unen el plano de proyección al punto de observación con un umbral de intensidad impuesto. Los resultados en 2D no permiten obtener un valor de la profundidad y de las distancias. Para crear una ilusión de rotación y una noción de profundidad y mejorar así el rendimiento en 3D, se realizan varios planos de proyección en ángulos de observación sucesivos. Se puede citar como ejemplo de técnica MIP, la descrita en el documento WO 2007/103216.

Las técnicas de restitución de la intensidad de los vóxeles permiten una visualización tridimensional ruidosa, lo que disminuye la discriminación entre diferencias contiguas de la escena observada.

Las nubes de puntos en 3D obtenidas directamente mediante un posicionamiento cartesiano de los vóxeles no permiten más que una reducida discriminación y aportan artificiosidades vinculadas a falsas alarmas.

Las técnicas de finalización superficial y de las que se presenta un ejemplo en el documento "Procédé de reconstruction 3D d'un objet d'une scène" (US2013/0100131), permiten llenar las zonas lagunares por generación de superficies tridimensionales para obtener un conjunto de datos finalizados del objeto, sin recurrir a una base de datos exterior. Dan así una superficie externa completa del objeto sin discontinuidades superficiales; los datos internos situados por detrás de las transparencias (vidrios por ejemplo) no se reconstruyen y se eliminan del proceso de identificación. Un proceso de discriminación de dos objetos próximos es más difícilmente alcanzable, porque la generación de las superficies puede ser matemáticamente compleja.

El objetivo de la invención es paliar estos inconvenientes.

Más precisamente la invención tiene por objeto un procedimiento de discriminación y de identificación por procesamiento de imagen en 3D de un objeto en una escena compleja. Se caracteriza principalmente porque incluye las etapas de la reivindicación 1.

Comprende eventualmente una etapa de visualización del volumen en 3D en vóxeles reducidos y/o del volumen en

vóxeles optimizados.

Según una característica, comprende además una etapa para generar a partir de la secuencia de imágenes en 2D MIP, unas coordenadas y un umbral de intensidad elegido, una nube (8) de puntos en 3D del objeto.

Según otra característica, incluye además las siguientes etapas:

- generar a partir del volumen en 3D de la escena compleja y del umbral de intensidad elegido, una nube de puntos en 3D bruta de la escena compleja,
- generar a partir del volumen en 3D optimizado y del umbral de intensidad optimizado, una nube de puntos en 3D optimizada del objeto,
- generar a partir de una superposición de la nube de puntos en 3D bruta de la escena, de la nube de puntos en 3D optimizada del objeto, y eventualmente de la nube de puntos en 3D del objeto, una nube de puntos en 3D global optimizada del objeto incluido en la escena compleja,
- visualizar la nube en 3D global optimizada.

Estas nubes pueden visualizarse igualmente por el operario.

El procedimiento permite una discriminación, una identificación y eventualmente un posicionamiento del objeto. Se hace posible la precisión de los detalles del objeto por la densificación de la nube de puntos, lo que permite también mejorar la identificación del objeto de la escena. Las nubes de puntos tridimensionales permiten separar y discriminar los diferentes objetos de la escena y discriminar el objeto a identificar de los otros elementos de la escena. El conjunto de las nubes de puntos asociadas al objeto se posiciona en una referencia cartesiana absoluta, lo que permite calcular simplemente las posiciones del objeto y las diferentes distancias entre el objeto y otros elementos de la escena. El objeto es recolocado en su contexto operativo.

La solución aportada subsana los inconvenientes individuales de cada técnica de visualización existente. No es necesario completar los datos (principalmente los datos enmascarados por otros elementos de la escena) porque están implícitamente presentes.

Surgirán otras características y ventajas de la invención con la lectura de la descripción detallada que sigue, realizada a título de ejemplo no limitativo y con referencia a los dibujos adjuntos en los que:

la figura 1 representa esquemáticamente un ejemplo de volumen en 3D con la cuadrícula correspondiente a los diferentes vóxeles, obtenida por reconstrucción en 3D,
la figura 2 muestra un ejemplo de encadenamiento de las diferentes etapas del procedimiento según la invención,
la figura 3 muestra un ejemplo de visualización en 3D de un volumen en 3D de vóxeles de una escena compleja, sin la cuadrícula y obtenida por reconstrucción en 3D,
la figura 4 muestra un ejemplo de imagen en 2D de una escena compleja,
la figura 5 muestra un ejemplo de secuencia de imágenes en 2D MIP de un objeto a identificar,
la figura 6 muestra un ejemplo de visualización en 3D del volumen en 3D reducido del objeto a identificar,
la figura 7 muestra un ejemplo de visualización en 3D del volumen en 3D optimizado del objeto a identificar,
la figura 8 muestra un ejemplo de visualización en 3D de una nube de puntos en 3D MIP del objeto a identificar,
la figura 9 muestra un ejemplo de visualización en 3D de una nube de puntos en 3D bruta de la escena compleja,
la figura 10 muestra un ejemplo de visualización en 3D de una nube de puntos en 3D optimizada del objeto a identificar,
las figuras 11 muestran un ejemplo de visualización en 3D de una nube de puntos en 3D global optimizada del objeto en la escena compleja, teniendo como objeto un vehículo (figura 11a) o dos vehículos (figura 11b).

De una figura a otra, los mismos elementos se referencian mediante las mismas referencias.

El procedimiento según la invención se aplica a un volumen en 3D en vóxeles de una escena compleja. Esta escena se ha transformado previamente por tanto por procesamiento de imagen en 3D en este volumen (1) en 3D del que se muestra un ejemplo en la figura 1.

Este volumen tridimensional puede obtenerse gracias un procedimiento de reconstrucción por transmisión o por fluorescencia (Tomografía de Proyección Óptica, imagen nuclear o tomografía computarizada por rayos X) o por reflexión (reflexión de una onda láser o por reflexión solar en el caso de la banda visible (entre 0,4 µm y 0,7 µm) o próxima al infrarrojo (entre 0,7 µm y 1 µm o SWIR entre 1 µm y 3 µm) o teniendo en cuenta la emisión térmica del objeto (tratamiento de imagen térmica entre 3 µm y 5 µm y entre 8 µm y 12 µm)), estos procesos de reconstrucción tridimensional se describen en la patente “Système optronique et procédé d’élaboration d’images en trois dimensions dédiés à l’identification” (US 8.836.762, EP 2 333 481).

Se utiliza el conjunto de los vóxeles procedentes de una reconstrucción tridimensional con la intensidad asociada, habiéndose obtenido esta reconstrucción preferentemente por reflexión.

El procedimiento de discriminación y de identificación según la invención incluye las siguientes etapas descritas en

relación con la figura 2. Se designa por objeto uno o varios elementos de la escena: un objeto puede ser un vehículo como se muestra en los ejemplos de las figuras 2 a 11a, o dos vehículos como se muestra en el ejemplo de la figura 11b, o incluso un elemento de vehículo como la llanta.

Entre las etapas siguientes, algunas requieren la participación de un operario a partir de una visualización sobre una pantalla, otras se implementan automáticamente.

❖ A) a partir del volumen (1) en 3D en vóxeles de la escena compleja (representada en las figuras 2 y 3 sin la cuadrícula de los vóxeles), se utiliza la técnica MIP para que el operario visualice en tiempo real sobre su pantalla un ejemplo de la escena compleja según diferentes planos de proyección (correspondientes a diferentes ángulos de observación) y según un umbral de intensidad determinado experimentalmente por un operario, y para que en el transcurso de estas diferentes visualizaciones de las que se muestra un ejemplo en la figura 4, esclarezca el volumen a analizar que incluye un objeto que desea discriminar y/o identificar. Como resultado de este proceso iterativo de tipo MIP, se genera una secuencia (2) de imágenes en 2D MIP del objeto a identificar, que corresponde visualmente a un volumen en 3D reducido que contiene el objeto a identificar, pero este volumen reducido no está aún determinado en este estadio. Un ejemplo de una secuencia (2) de tres imágenes en 2D MIP de ese tipo se muestra en las figuras 2 y 5.

❖ B1) extraer automáticamente a partir de esta secuencia (2) de imágenes en 2D MIP, unas coordenadas (3) de este volumen reducido que contiene el objeto a identificar

❖ B2) determinar uno de los umbrales (4) de intensidad que se ha utilizado en los procesos MIP de la etapa A); este umbral se determina experimentalmente por el operario

❖ C) extraer automáticamente a partir del volumen (1) en 3D en vóxeles de la escena compleja, y utilizando las coordenadas (3) y el umbral (4) de intensidad, un volumen (5) en vóxeles correspondiente al objeto a identificar (dicho de otra manera este volumen en 3D reducido se determina en función de las coordenadas (3) y sus vóxeles se umbralizan en intensidad en función del umbral (4) de intensidad): se trata del volumen reducido de la etapa A). El operario puede visualizar este volumen (5) reducido si lo desea, como se muestra en las figuras 2 y 6; esta visualización puede ofrecerle una posibilidad de identificar el objeto.

❖ D) generar automáticamente a partir del volumen (5) en vóxeles, mediante un proceso de optimización por umbralización de intensidad, asociándose cada intensidad a un color:

- un umbral (7) optimizado,
- posteriormente utilizando este umbral optimizado, un volumen (6) 3D optimizado en vóxeles del objeto de identificar.

Se obtiene así de alguna manera el volumen reducido, en color después del umbralizado, que el operario puede visualizar igualmente, como se muestra en las figuras 2 y 7, e identificar. Se obtiene un volumen en 3D en vóxeles que presenta visualmente muchos más detalles que el volumen en 3D de origen.

En el transcurso de las etapas anteriores, A a C, el operario ha tenido la posibilidad de identificar el objeto con una probabilidad creciente. Como resultado de esta etapa D, el operario ha alcanzado la mayor probabilidad de identificar el objeto.

El objeto a identificar puede dividirse a su vez en sub-objetos. Los sub-objetos de un vehículo son por ejemplo las ruedas, el capó,… Las etapas anteriores se aplicarán entonces a cada sub-objeto. Se obtiene para cada sub-objeto, una secuencia (2) de imágenes en 2D MIP, unas coordenadas (3) del volumen en 3D reducido, un umbral (4) de intensidad elegido, un volumen (5) en 3D en vóxeles reducido, un umbral (7) optimizado, y un volumen (6) en 3D optimizado.

Como resultado de esta etapa D, cuando el objeto incluye varios elementos o sub-objetos que presentan un interés para el operario, estos se han podido discriminar.

Las siguientes etapas se dirigen a reponer en situación el objeto identificado.

❖ E) eventualmente generar automáticamente a partir de la secuencia (2) de imágenes en 2D MIP utilizando las coordenadas (3) y el umbral (4) de intensidad, una nube (8) de puntos en 3D del objeto a identificar, que el operario puede visualizar eventualmente, como se muestra en las figuras 2 y 8.

❖ F1) generar automáticamente a partir de la secuencia (2) de imágenes en 2D MIP y utilizando el umbral (4) de intensidad, una nube (9) de puntos en 3D de la escena compleja que incluye por tanto el objeto a identificar, que el operario puede eventualmente visualizar, como se muestra en las figuras 2 y 9.

❖ F2) generar automáticamente a partir del volumen (6) en 3D utilizado en vóxeles y utilizando el umbral (7) de intensidad, una nube (10) de puntos en 3D optimizada del objeto a identificar, que el operario puede eventualmente visualizar, como se muestra en las figuras 2 y 10.

❖ G) generar automáticamente por superposición:

- la nube (9) de puntos en 3D bruta de la escena compleja,
- la nube (10) de puntos en 3D optimizada del objeto, que presenta una gran densidad a nivel de los detalles del objeto que permiten su identificación,
- y la nube (8) de puntos en 3D del objeto si se ha generado,

una nube (11) de puntos en 3D global optimizada del objeto incluido en la escena compleja. Esta nube permite posicionar el objeto en su contexto operativo teniendo en cuenta las coordenadas absolutas, discriminarlo, y ayudar al operario a identificarlo con una gran probabilidad. En efecto, el operario visualiza esta nube (11) global optimizada, de la que se muestra un ejemplo en las figuras 2 y 11a para un objeto que representa un vehículo o en la figura 11b para un objeto que representa dos vehículos; esto le permite visualizar el objeto identificado en su contexto operativo.

Las diferentes extracciones de las nubes de puntos permiten una disminución de las artificiosidades y del ruido.

El conjunto de las nubes de puntos asociadas al objeto se posicionan en una referencia cartesiana absoluta, lo que permite calcular simplemente la posición del objeto y las diferentes distancias entre el objeto y otros elementos de la escena. El objeto se recoloca en su contexto operativo.

Cuando se han identificado unos sub-objetos, cada una de estas etapas de puesta en situación se aplica superponiendo las nubes respectivamente obtenidas a partir de cada sub-objeto.

Cuando el operario considera que el resultado no es satisfactorio, estas etapas A a G pueden iterarse con el fin de obtener la mejor densificación posible de estos detalles.

Como resultado de la etapa G, cuando el objeto incluye varios elementos o sub-objetos discriminados entre sí, estos se resitúan en su contexto y puede medirse su posicionamiento, unos relativamente a los otros, en las coordenadas cartesianas.

Con este procedimiento, es particularmente fácil visualizar en tiempo real diferentes elementos de la escena observada.

Según la invención, la identificación de los objetos de la escena se mejora gracias a una densificación de las nubes de puntos tridimensionales.

Este procedimiento de discriminación y de identificación se aplica principalmente al campo de la seguridad para la identificación y la localización precisa de objetivos, y al campo médico para la identificación y la localización de tumores.

La presente invención puede implementarse a partir de elementos de hardware y/o de software. Puede estar disponible en tanto que producto de programa informático en un soporte legible por ordenador, comprendiendo este programa informático instrucciones de código que permiten efectuar las etapas del procedimiento de discriminación y de identificación. El soporte puede ser electrónico, magnético, óptico, electromagnético o ser un soporte de difusión de tipo infrarrojo. Dichos soportes son por ejemplo, unas memorias de semiconductor (memoria de acceso aleatorio RAM, memoria solo de lectura ROM), unas bandas, unos disquetes o discos magnéticos u ópticos (Disco Compacto - Memoria Solo de Lectura (CD-ROM), disco compacto, de lectura/escritura (CD-R/W) y DVD).

REIVINDICACIONES

1. Procedimiento de discriminación y de identificación por un procesamiento de imagen en 3D de un objeto en una escena compleja, **caracterizado porque** incluye las siguientes etapas:

- generar (A) una secuencia (2) de imágenes del objeto, a partir de un volumen en 3D en vóxeles de la escena compleja, siendo este volumen (1) predeterminado y visualizado por un operario utilizando un proceso iterativo de tipo “Maximum Intensity Projection” a partir de un plano de proyección y de un umbral de intensidad determinados por el operario en cada iteración, designándose la secuencia de imágenes obtenida una secuencia de imágenes en 2D MIP como “Maximum Intensity Projection”,
- extraer (B1) automáticamente a partir de la secuencia (2) de imágenes en 2D MIP, unas coordenadas (3) de un volumen reducido que corresponde a la secuencia (2) de imágenes en 2D MIP,
- elegir (B2) uno (4) de los umbrales de intensidad utilizados en el transcurso de las iteraciones de la etapa A), efectuándose esta elección por el operario,
- a partir del volumen (1) en 3D en vóxeles de la escena compleja, de las coordenadas (3) y del umbral (4) de intensidad, extraer (C) automáticamente un volumen (5) en 3D en vóxeles reducido que contiene el objeto,
- generar (D) automáticamente a partir del volumen (5) reducido, un umbral (7) de intensidad optimizado por optimización de umbralización de intensidad, asociándose a cada intensidad un color, y un volumen (6) en vóxeles optimizado, a partir de este umbral de intensidad optimizado,
- identificar el objeto por visualización.

2. Procedimiento de discriminación y de identificación de un objeto en una escena compleja según la reivindicación anterior, **caracterizado porque** comprende una etapa de visualización del volumen (5) en 3D en vóxeles reducido.

3. Procedimiento de discriminación y de identificación de un objeto en una escena compleja según una de las reivindicaciones anteriores, **caracterizado porque** comprende una etapa de visualización del volumen (6) en vóxeles optimizado.

4. Procedimiento de discriminación y de identificación de un objeto en una escena compleja según una de las reivindicaciones anteriores, **caracterizado porque** comprende además una etapa (E) para generar a partir de la secuencia (2) de imágenes en 2D MIP, de las coordenadas (3) y del umbral (4) de intensidad elegido, una nube (8) de puntos en 3D del objeto.

5. Procedimiento de discriminación y de identificación de un objeto en una escena compleja según una de las reivindicaciones anteriores, **caracterizado porque** incluye además las siguientes etapas:

- generar (F1) a partir del volumen (1) en 3D de la escena compleja y del umbral (4) de intensidad elegido, una nube (9) de puntos en 3D bruta de la escena compleja,
- generar (F2) a partir del volumen (6) en 3D optimizado y del umbral (7) de intensidad optimizado, una nube (10) de puntos en 3D optimizada del objeto,
- generar (G) a partir de una superposición de la nube (9) de puntos en 3D bruta de la escena y de la nube (10) de puntos en 3D optimizada del objeto, una nube (11) de puntos en 3D global optimizada del objeto incluido en la escena compleja,
- visualizar la nube en 3D global optimizada.

6. Procedimiento de discriminación y de identificación de un objeto en una escena compleja según las reivindicaciones 4 y 5 tomadas en combinación, **caracterizado porque** la nube (11) de puntos global optimizada se genera a partir de la superposición de la nube (9) de puntos en 3D bruta de la escena, de la nube (10) de puntos en 3D optimizada del objeto, y de la nube (8) de puntos en 3D del objeto.

7. Procedimiento según la reivindicación 4 o 5, **caracterizado porque** se visualiza por el operario la nube (8) de puntos en 3D del objeto y/o la nube (9) de puntos en 3D de la escena y/o la nube (10) de puntos en 3D optimizada del objeto.

8. Procedimiento según una de las reivindicaciones anteriores, **caracterizado porque** el volumen (1) en 3D en vóxeles se obtiene por reflexión de una onda óptica visible o de IR sobre la escena compleja o por emisión térmica del objeto.

9. Un producto de programa informático, comprendiendo dicho programa informático instrucciones de código que permiten efectuar las etapas del procedimiento según una cualquiera de las reivindicaciones 1 a 8, cuando dicho programa se ejecuta en un ordenador.

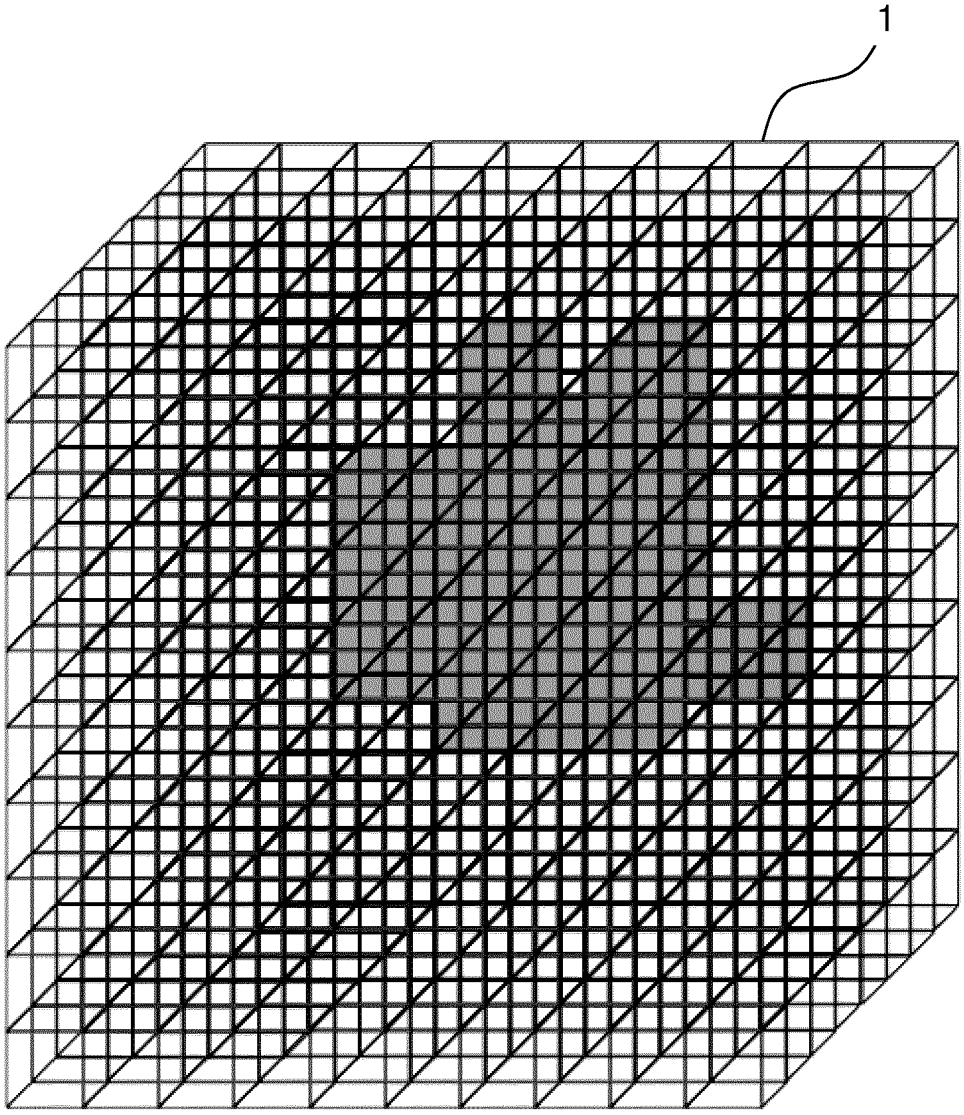

FIG.1

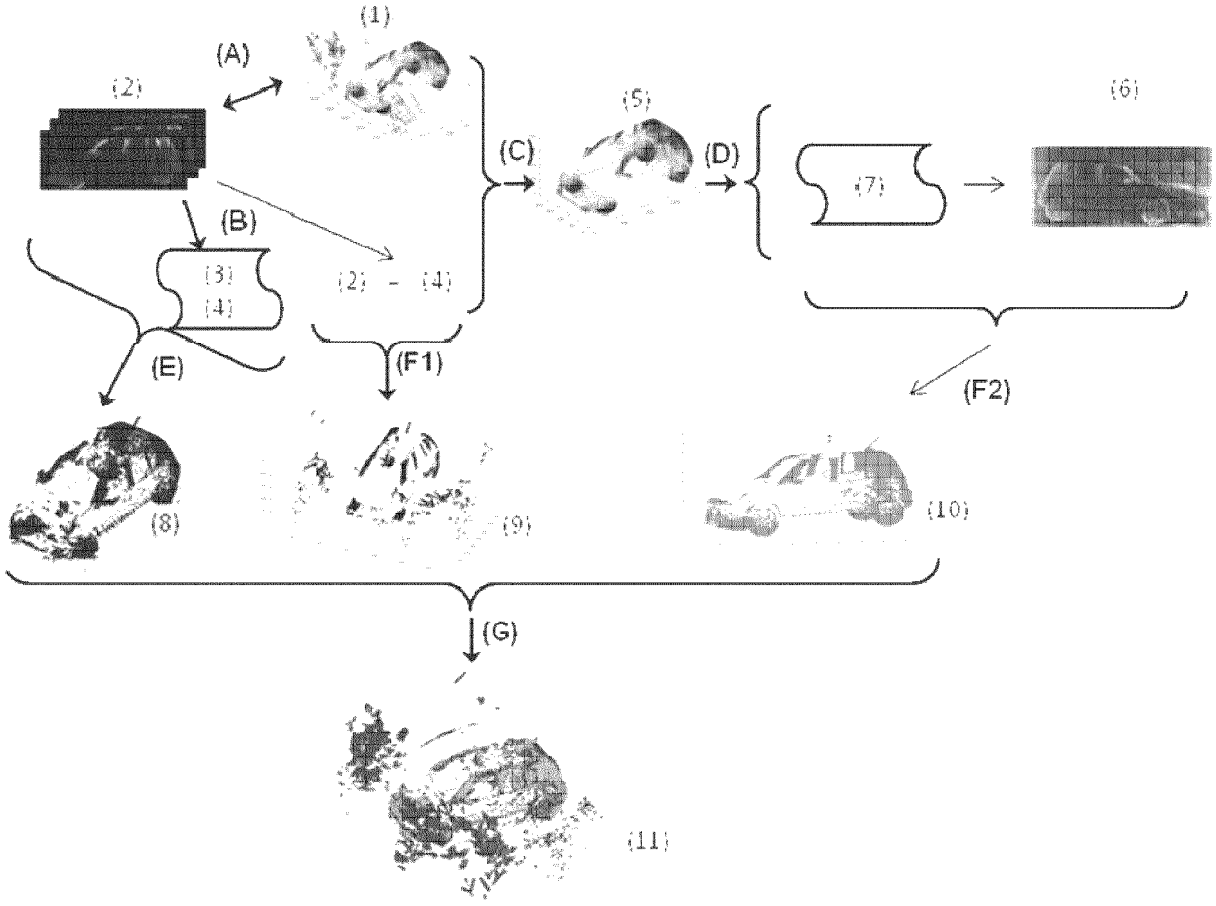

FIG.2

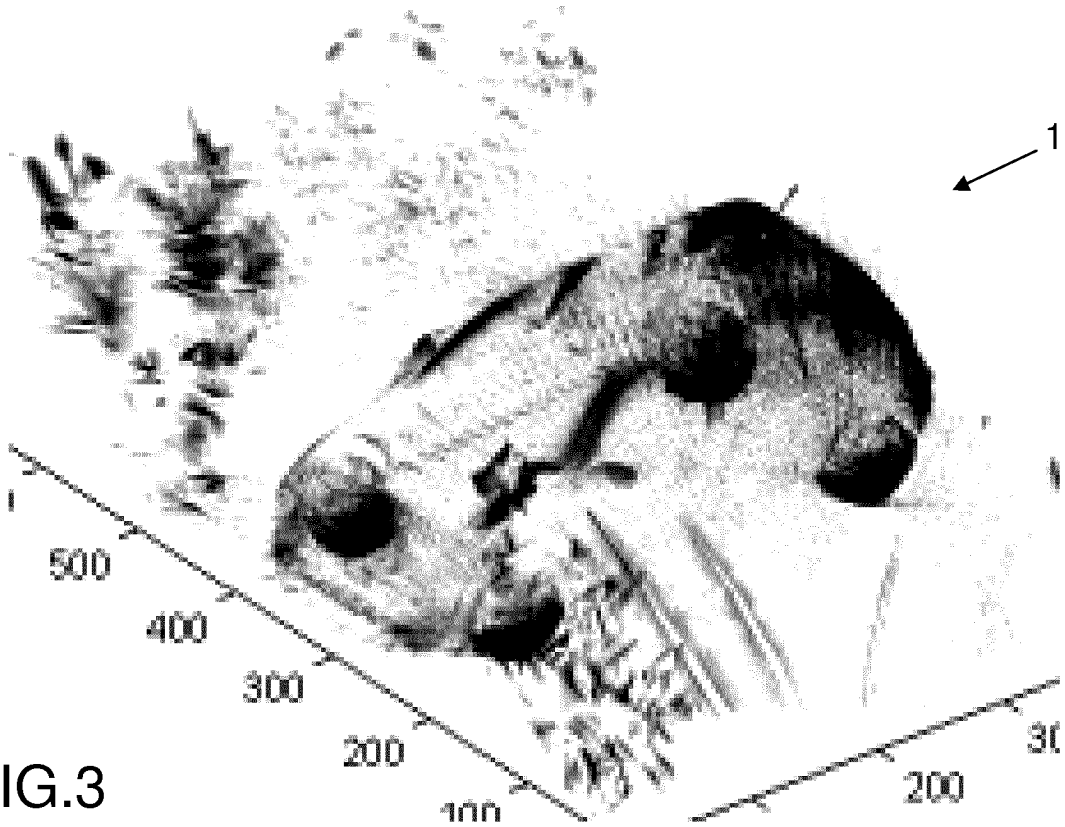

FIG.3

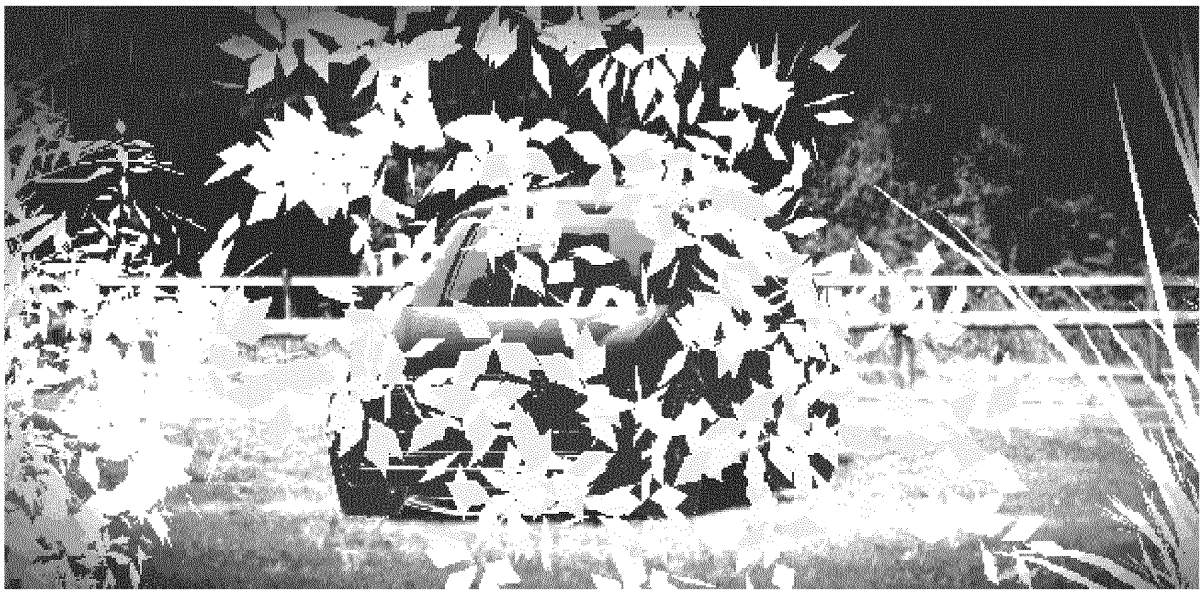

FIG.4

FIG.5

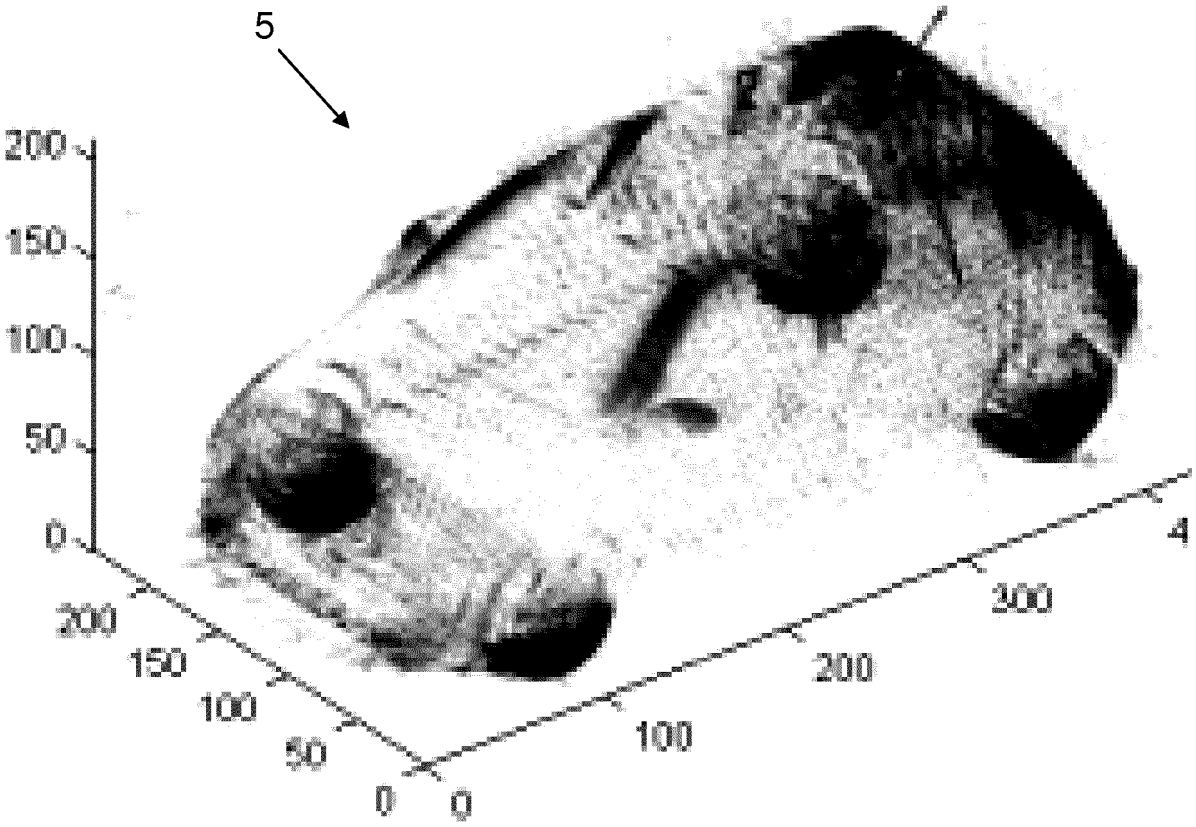

FIG.6

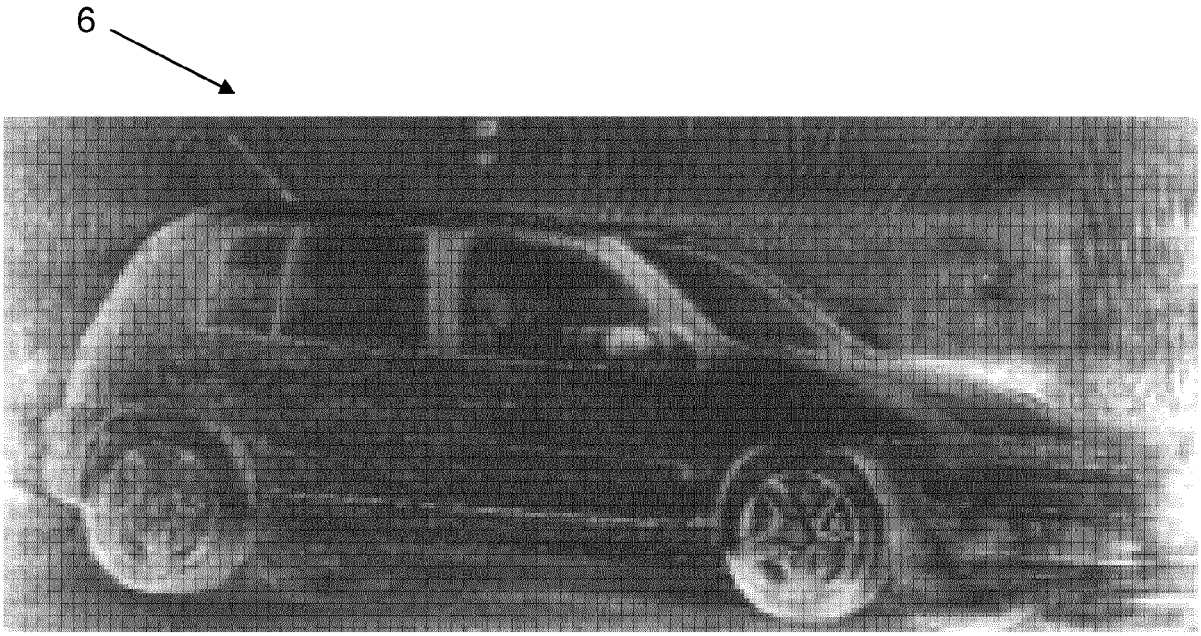

FIG.7

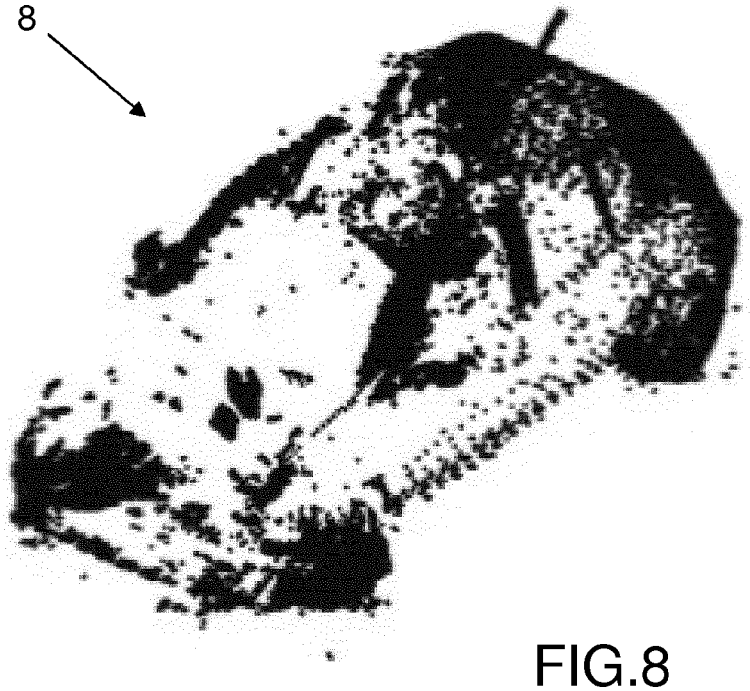

FIG.8

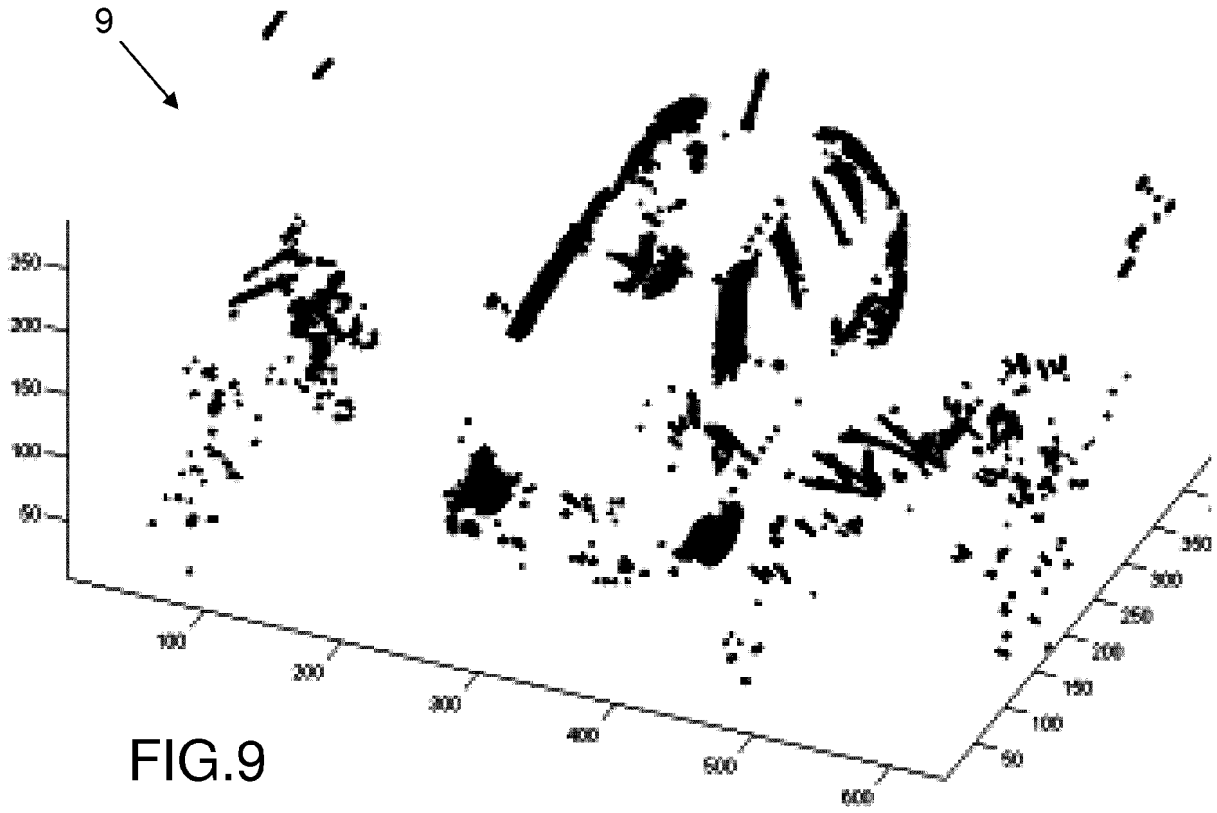
9
350
300
250
200
150
100
50
100
200
300
400
500
600
50
100
150
200
250
300
350

FIG.9

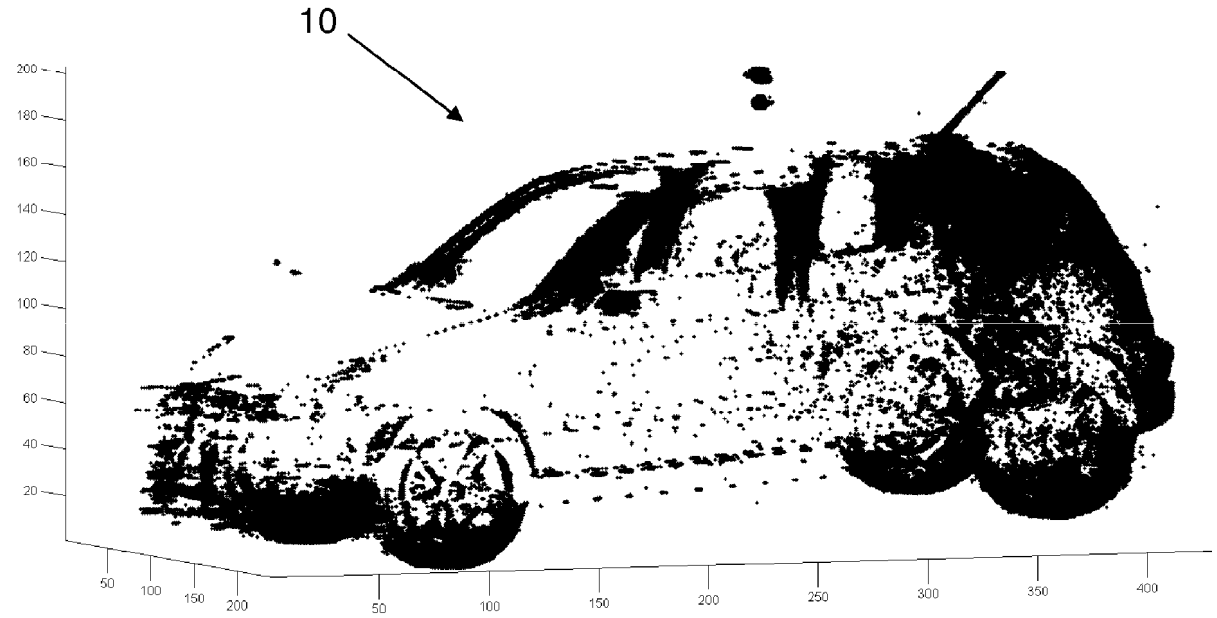
10
200
180
160
140
120
100
80
60
40
20
50
100
150
200
50
100
150
200
250
300
350
400

FIG.10

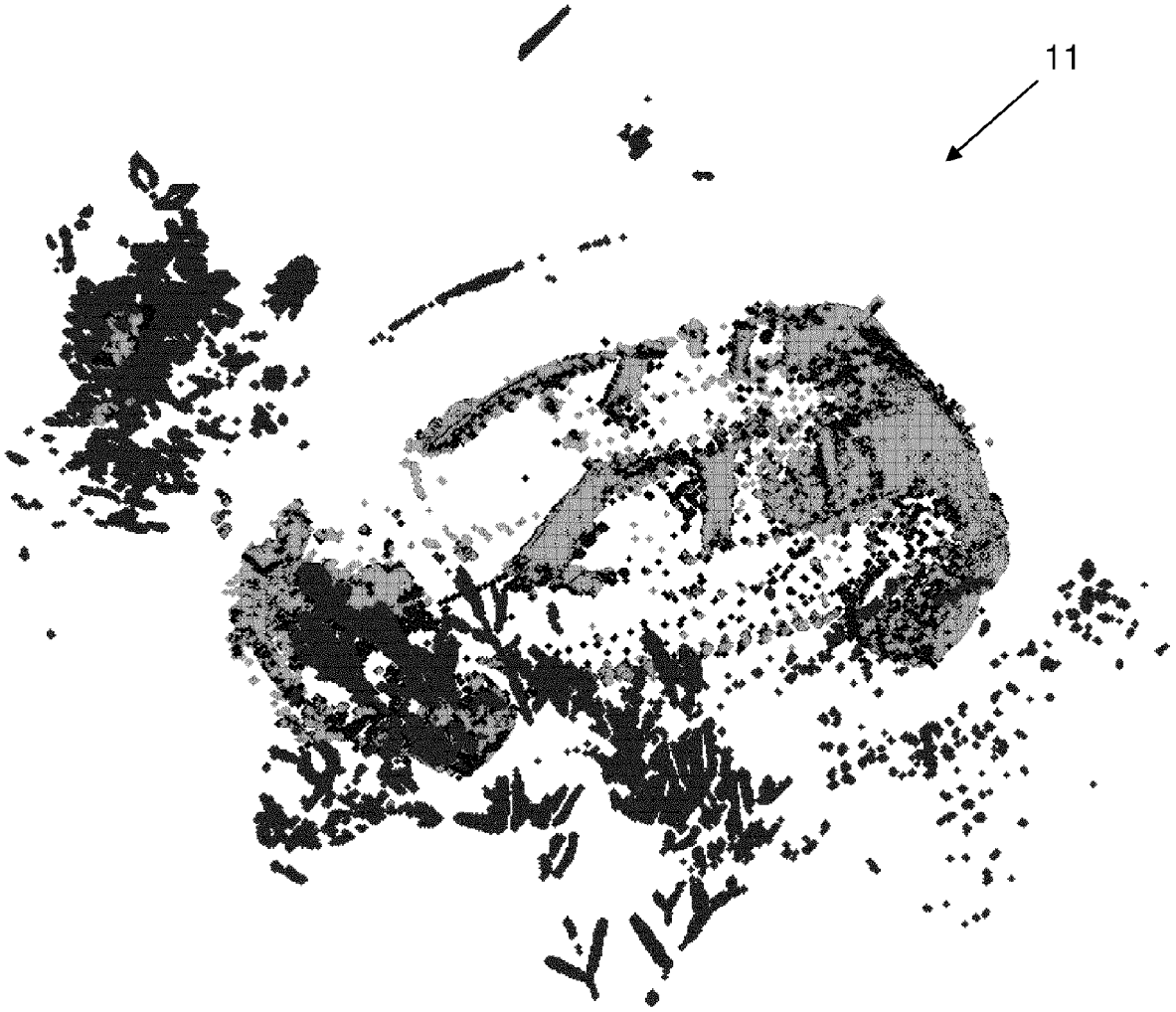

FIG.11a

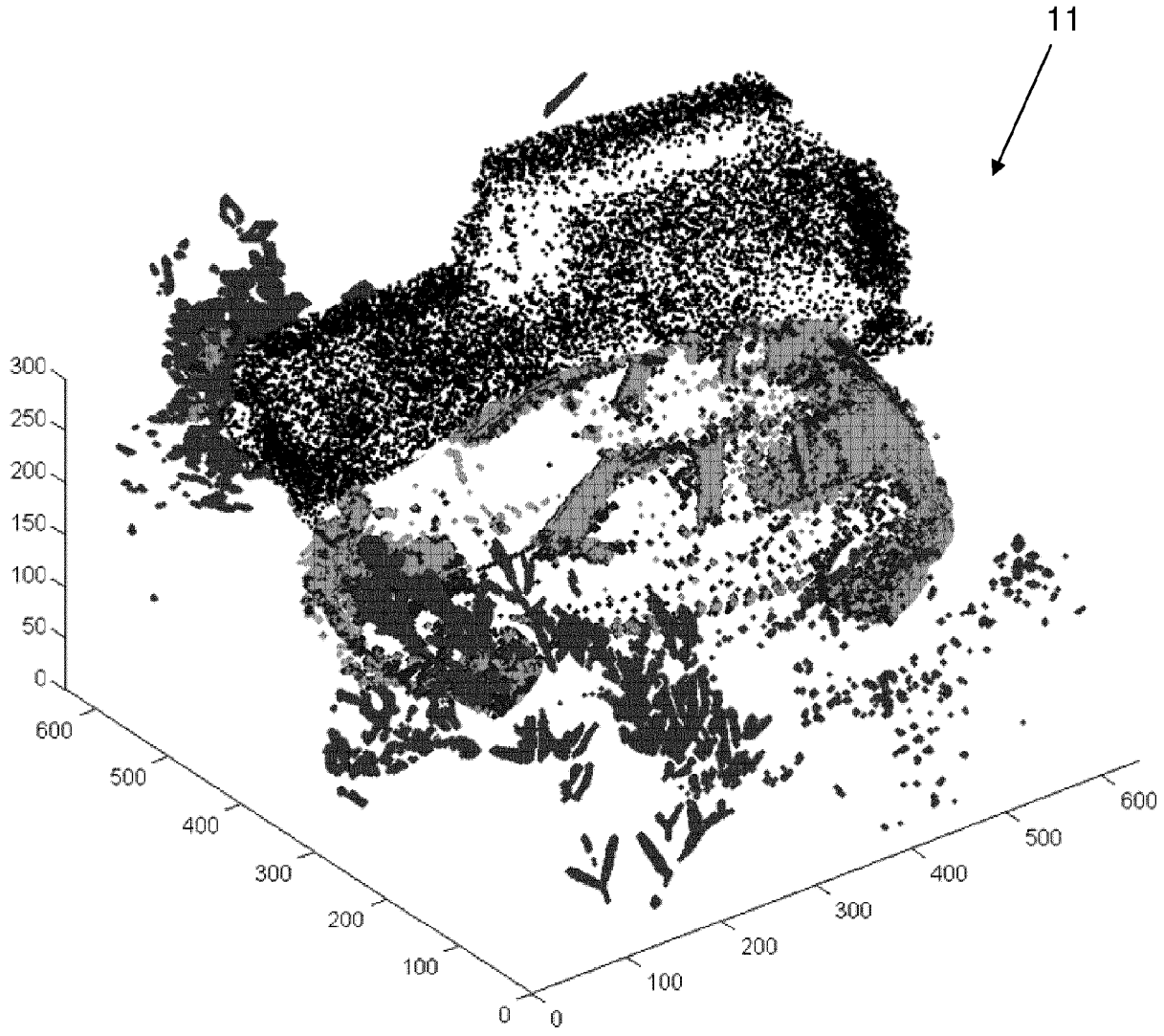
11
300
250
200
150
100
50
0
600
500
400
300
200
100
0
0
100
200
300
400
500
600

FIG.11b