

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 881 320**

51 Int. Cl.:

**G06T 7/55** (2007.01)  
**G06T 1/00** (2006.01)  
**G06T 5/50** (2006.01)  
**H04N 5/232** (2006.01)  
**H04N 7/18** (2006.01)  
**G06T 17/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **03.12.2018 PCT/JP2018/044373**

87 Fecha y número de publicación internacional: **20.06.2019 WO19116942**

96 Fecha de presentación y número de la solicitud europea: **03.12.2018 E 18887303 (8)**

97 Fecha y número de publicación de la concesión europea: **09.06.2021 EP 3608873**

54 Título: **Dispositivo de generación, procedimiento de generación y programa para modelo tridimensional**

30 Prioridad:

**14.12.2017 JP 2017239891**  
**07.05.2018 JP 2018089467**  
**06.11.2018 JP 2018209196**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**29.11.2021**

73 Titular/es:

**CANON KABUSHIKI KAISHA (100.0%)**  
**30-2 Shimomaruko 3-chome, Ohta-ku**  
**Tokyo 146-8501, JP**

72 Inventor/es:

**MORISAWA KEISUKE y**  
**KOBAYASHI KIWAMU**

74 Agente/Representante:

**DURAN-CORRETJER, S.L.P**

ES 2 881 320 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**

Dispositivo de generación, procedimiento de generación y programa para modelo tridimensional

5 ESTADO DE LA TÉCNICA ANTERIOR

Sector técnico

La presente invención se refiere a la generación de un modelo tridimensional de un objeto dentro de una imagen.

10 Técnica anterior

15 Convencionalmente, es conocido el procedimiento denominado "procedimiento de intersección de volúmenes visuales" (PTL 1, NPL 1), como un procedimiento de estimación de una forma tridimensional de un objeto utilizando una imagen de múltiples puntos de vista capturada de manera síncrona desde diferentes puntos de vista mediante una serie de cámaras.

20 El documento de Patente US 2009/141941 da a conocer un procedimiento para generar una representación tridimensional de una escena que incluye una serie de objetos (jugadores) dispuestos en un terreno de juego. La publicación "Multi-view Reconstruction of Unknown Objects within a Known Environment", de Stefan Kuhn et al., presenta un procedimiento basado en visión general para reconstruir múltiples objetos desconocidos (por ejemplo, humanos) dentro de un entorno conocido (por ejemplo, mesas, estanterías, robots) que, usualmente, tiene oclusiones.

25 Las figuras 1A a 1C son diagramas que muestran los principios básicos del procedimiento de intersección de volúmenes visuales. A partir de una imagen obtenida capturando un determinado objeto, una imagen de máscara que representa una silueta bidimensional del objeto se obtiene en la superficie de detección de imágenes (figura 1A). A continuación, se considera una forma de pirámide que se propaga a un espacio tridimensional, de manera que cada lado se extiende desde el centro de captura de imágenes de la cámara y atraviesa cada punto del contorno de la imagen de máscara (figura 1B). Esta forma de pirámide es denominada el "volumen visual" objetivo por la cámara relevante. Además, encontrando el área común de una serie de volúmenes visuales, es decir, la intersección de los volúmenes visuales, se encuentra la forma tridimensional (modelo tridimensional) de un objeto (figura 1C). Tal como se ha descrito anteriormente, en la estimación de forma mediante el procedimiento de intersección de volúmenes visuales, la forma tridimensional de un objeto se estima proyectando un punto de muestreo en un espacio que tiene una probabilidad de que exista un objeto sobre la imagen de máscara, y verificando si el punto proyectado está incluido en la imagen de máscara en común con una serie de puntos de vista.

Lista de referencias

40 Bibliografía de patentes

PTL 1 Patente japonesa abierta a inspección pública núm. 2014-10805.

45 Bibliografía no de patentes

NPL 1 Laurentini A: "The Visual Hull Concept for Silhouette-Based Image Understanding", IEEE Transactions Pattern Analysis and machine Intelligence, vol. 16, núm. 2, págs. 150 - 162, febrero de 1994

50 CARACTERÍSTICAS DE LA INVENCION

Problema a resolver por la invención

55 En el procedimiento de intersección de volúmenes visuales descrito anteriormente, es necesario que la imagen de máscara pueda representar correctamente la silueta de un objeto objetivo y, en caso de que la silueta sobre la imagen de máscara sea incorrecta, la forma tridimensional que se genera es asimismo incorrecta. Por ejemplo, en un caso en que un objeto estacionario, tal como una estructura, que existe delante de una persona, impide que sea capturada una parte de la persona, que es un objeto objetivo, y, por lo tanto, se pierde una parte de la silueta de la persona representada por la imagen de máscara, se produce un defecto en el modelo tridimensional generado. Además, en caso de que no se utilice una imagen de máscara cuya parte de la silueta se pierde, se reduce la precisión geométrica del modelo tridimensional que se obtiene. En particular, en el caso en que la parte cuya captura es impedida por la estructura es relativamente pequeña, es deseable utilizar en la medida de lo posible la imagen de máscara cuya parte de la silueta se pierde, para obtener un modelo tridimensional con una alta precisión geométrica utilizando la imagen de máscara.

65 La presente invención se ha realizado en vista de los problemas descritos anteriormente, y un objetivo de la misma es impedir que se produzca un defecto en un modelo tridimensional que se genera, incluso en un caso en el que,

dentro de una escena de captura de imágenes, existe una estructura o similar que impide que se capture una parte de un objeto objetivo.

Medios para resolver el problema

5 Un procedimiento y un dispositivo de generación, según la presente invención, se definen mediante las reivindicaciones independientes y dependientes adjuntas.

Resultado ventajoso de la invención

10 Según la presente invención, se posibilita generar un modelo tridimensional con alta calidad en el que no hay ningún defecto o la magnitud del defecto se reduce, incluso en un caso en el que, dentro de una escena de captura de imágenes, existe una estructura o similar que impide que se capture una parte de un objeto objetivo.

15 Otras características de la presente invención resultarán evidentes a partir de la siguiente descripción de realizaciones a modo de ejemplo, haciendo referencia a los dibujos adjuntos.

#### BREVE DESCRIPCIÓN DE LOS DIBUJOS

20 Las figuras 1A a 1C son diagramas que muestran los principios básicos de un procedimiento de intersección de volúmenes visuales;  
la figura 2A es un diagrama de bloques que muestra una configuración de un sistema de generación de imágenes de punto de vista virtual, y la figura 2B es un diagrama que muestra un ejemplo de disposición de cada cámara que configura un conjunto de cámaras;

25 la figura 3 es un diagrama de bloques funcionales que muestra una configuración interna de un dispositivo de generación de modelo tridimensional, según una primera realización;  
la figura 4 es un diagrama de flujo que muestra un flujo de un proceso de formación de modelo tridimensional, según la primera realización;

30 cada una de las figuras 5A a 5H es un diagrama que muestra un ejemplo de una imagen capturada por cada cámara;  
cada una de las figuras 6A a 6H es un diagrama que muestra un ejemplo de una máscara de estructura;  
cada una de las figuras 7A a 7H es un diagrama que muestra un ejemplo de una máscara del primer plano;  
cada una de las figuras 8A a 8H es un diagrama que muestra un ejemplo de una máscara integrada;

35 la figura 9 es un diagrama que muestra un ejemplo de un modelo tridimensional integrado que se genera en base a la máscara integrada;  
la figura 10 es un diagrama que muestra un ejemplo de un modelo tridimensional que se genera utilizando solamente la máscara del primer plano, mediante un procedimiento convencional;  
la figura 11 es un diagrama de flujo que muestra un flujo del proceso de formación de modelo tridimensional, según una segunda realización;

40 la figura 12A es un diagrama que muestra un modelo tridimensional integrado que se genera basándose en una máscara integrada, la figura 12B es un diagrama que muestra un modelo tridimensional de una estructura, que se genera basándose solamente en una máscara de estructura, y la figura 12C es un diagrama que muestra un modelo tridimensional de solamente un primer plano, obtenido a partir de una diferencia entre el modelo tridimensional integrado de la figura 12A y el modelo tridimensional de la estructura de la figura 12B;

45 cada una de las figuras 13A a 13C es un diagrama que muestra un ejemplo de que configuran un modelo tridimensional, según una tercera realización;  
la figura 14 es un diagrama que muestra un ejemplo de disposición de cámaras de un sistema de generación de imágenes de punto de vista virtual, según la tercera realización;  
la figura 15 es un diagrama que muestra un ejemplo de una configuración funcional de un dispositivo de generación de modelo tridimensional, según la tercera realización;

50 la figura 16 es un diagrama de flujo que muestra un procedimiento de procesamiento llevado a cabo por el dispositivo de generación de modelo tridimensional, según la tercera realización;  
las figuras 17A a 17E son diagramas que muestran ejemplos de imágenes capturadas, capturadas por una serie de cámaras, según la tercera realización;

55 cada una de las figuras 18A a 18E es un diagrama que muestra un ejemplo de una imagen de máscara de estructura, según la tercera realización;  
cada una de las figuras 19A a 19E es un diagrama que muestra un ejemplo de una imagen de máscara del primer plano, según la tercera realización;

60 cada una de las figuras 20A a 20E es un diagrama que muestra un ejemplo de una imagen de máscara integrada, obtenida integrando la imagen de máscara del primer plano y la imagen de máscara de estructura, según la tercera realización;

la figura 21 es un diagrama que muestra un espacio de vóxeles que es un objetivo de generación de un modelo tridimensional de un sistema de estadio deportivo, según la tercera realización;

65 la figura 22 es un diagrama que muestra cómputo verdadero/cómputo falso, según la tercera realización;  
la figura 23 es un diagrama que muestra un ejemplo de un modelo tridimensional que se genera aplicando determinación de cómputo falso basada en umbral, según la tercera realización;

la figura 24 es un diagrama que muestra un ejemplo de un modelo tridimensional que se genera aplicando determinación de cómputo falso basada en umbral, y determinación de cómputo verdadero basada en umbral, según la tercera realización;

la figura 25 es un diagrama que muestra un modelo tridimensional en un caso en el que se produce un defecto;

5 la figura 26 es un diagrama que muestra un ejemplo de una configuración funcional de un dispositivo de generación de modelo tridimensional, según una cuarta realización;

la figura 27 es un diagrama de flujo que muestra un procedimiento de procesamiento llevado a cabo por el dispositivo de generación de modelo tridimensional, según la cuarta realización;

10 la figura 28 es un diagrama que muestra una disposición de cámaras de un sistema de generación de imágenes de punto de vista virtual, y un ejemplo de un primer plano, según la cuarta realización;

la figura 29 es un diagrama que muestra cómputo verdadero/falso, según la cuarta realización;

la figura 30 es un diagrama que muestra un bloque funcional de un dispositivo de generación de modelo tridimensional, según una quinta realización;

15 la figura 31 es un diagrama que muestra un flujo de procesamiento del dispositivo de generación de modelo tridimensional, según la quinta realización;

la figura 32 es un diagrama que muestra el cómputo verdadero/falso sin adición de ponderación y con adición de ponderación, según la quinta realización;

la figura 33 es un diagrama que muestra un bloque funcional de un dispositivo de generación de modelo tridimensional, según una sexta realización;

20 la figura 34 es un diagrama que muestra un flujo de proceso del dispositivo de generación de modelo tridimensional, según la sexta realización; y

la figura 35 es un diagrama que muestra cómputo falso/estructura, según la sexta realización.

#### DESCRIPCIÓN DE LAS REALIZACIONES

25 En lo que sigue, haciendo referencia a los dibujos adjuntos se explicará en detalle la presente invención de acuerdo con las realizaciones. Las configuraciones mostradas en las siguientes realizaciones son tan sólo a modo de ejemplo, y la presente invención no se limita a las configuraciones mostradas esquemáticamente.

30 [Primera realización]

En la presente realización se explica un aspecto en que se genera, para un primer plano, un modelo tridimensional en el que no hay defectos o se reduce la magnitud del defecto, utilizando, además de una silueta bidimensional del primer plano en una escena de captura de imágenes, una imagen de máscara que incluye una silueta bidimensional de una estructura que impide que se capture, por lo menos, una parte de la misma. En este aspecto, se genera un modelo tridimensional que incluye una estructura o similar, que impide que se capture una parte del primer plano. En la presente memoria descriptiva, el "primer plano" se refiere a un objeto en movimiento que existe dentro de una imagen capturada, que se mueve en un caso en el que se lleva a cabo captura de imágenes en una serie temporal desde el mismo ángulo (cuya posición absoluta puede cambiar) y que se puede ver desde un punto de vista virtual. Además, la "estructura" se refiere a un objeto estacionario que existe dentro de una imagen capturada, que no se mueve en un caso en el que la captura de imágenes se lleva a cabo en una serie temporal desde el mismo ángulo (cuya posición absoluta no cambia, es decir, está en reposo) y que tiene una probabilidad de impedir que se capture el primer plano. El modelo tridimensional al que se hace aquí referencia son datos que representan una forma tridimensional.

45 En la siguiente explicación, se supone un caso en que una estructura (objeto estacionario), tal como una portería de fútbol, impide que se capture una parte de un primer plano (objeto de movimiento), tal como un jugador y un balón, en el momento de generar una imagen de punto de vista virtual tomando un partido de fútbol como una escena de captura de imágenes. La imagen de punto de vista virtual es una imagen de vídeo que es generada por un usuario final y/o un operador designado y similar, que controla libremente la posición y orientación de una cámara virtual, y se denomina asimismo una imagen de punto de vista libre, una imagen de punto de vista arbitrario, y similares. Además, la imagen de punto de vista virtual que se genera y la imagen de múltiples puntos de vista que es la fuente de la imagen de punto de vista virtual pueden ser una imagen en movimiento o una imagen estática. En cada realización explicada a continuación, se explica un caso como ejemplo en el que un modelo tridimensional para generar una imagen de punto de vista virtual de una imagen en movimiento se genera utilizando la imagen de múltiples puntos de vista de la imagen en movimiento.

60 En la presente realización, la explicación se da sobre la hipótesis de que se toma el fútbol como escena de captura de imágenes y la portería de fútbol instalada de forma permanente es la estructura, pero esto no es limitativo. Por ejemplo, puede ser posible asimismo tratar un banderín de córner como estructura y, en un caso en el que se tome un estudio interior como la escena de captura de imágenes, es posible asimismo tratar el mobiliario o la utilería como estructura. Es decir, cualquier objeto estacionario se puede tratar como estructura siempre que continúe en su estado estacionario o casi estacionario.

(Configuración del sistema)

La figura 2A es un diagrama de bloques que muestra un ejemplo de una configuración de un sistema de generación de imágenes de punto de vista virtual que incluye un dispositivo de generación de modelo tridimensional, según la presente realización. Un sistema de generación de imágenes de punto de vista virtual 100 incluye un conjunto de cámaras 110 que incluye una serie de cámaras, un dispositivo de control 120, un dispositivo de separación del primer plano 130, un dispositivo de generación de modelo tridimensional 140 y un dispositivo de representación 150. El dispositivo de control 120, el dispositivo de separación del primer plano 130, el dispositivo de generación de modelo tridimensional 140 y el dispositivo de representación 150 están, cada uno, implementados por una CPU que lleva a cabo procesamiento de cálculo o por un ordenador general (aparato de procesamiento de información) que incluye una memoria o similar que almacena resultados de procesamiento de cálculo, programas y similares.

La figura 2B es un diagrama que muestra la disposición de un total de ocho cámaras 211 a 218 que configuran el conjunto de cámaras 110 en un diagrama de vista aérea, en un caso en que el campo 200 se ve directamente desde arriba. Cada una de las cámaras 211 a 218 está instalada a una altura predeterminada desde el suelo, para rodear el campo 200, y adquiere datos de imagen de múltiples puntos de vista desde diferentes puntos de vista capturando la posición frente a una de las porterías desde diversos ángulos. En el campo de césped 200 está dibujada una cancha de fútbol 210 (de hecho, mediante líneas blancas) y en el lado izquierdo de la misma está situada una portería de fútbol 202. Además, una marca x 203 delante de la portería de fútbol 202 indica una dirección de línea de visión común (punto de mira) de las cámaras 211 a 218, y un círculo en línea discontinua 204 indica un área en la que cada una de las cámaras 211 a 218 puede llevar a cabo captura de imágenes con el punto de mira 203 como centro. En la presente realización, una posición se representa en un sistema de coordenadas en el que una esquina del campo 200 se toma como el origen, la dirección del lado largo como un eje x, la dirección de lado corto como un eje y y la dirección de la altura como un eje z. Los datos de la imagen de múltiples puntos de vista obtenida por cada cámara del conjunto de cámaras 110 se envían al dispositivo de control 120 y al dispositivo de separación del primer plano 130. En la figura 2A, cada una de las cámaras 211 a 218, el dispositivo de control 120 y el dispositivo de separación del primer plano 130 están conectados mediante una topología en estrella, pero se puede adoptar asimismo la topología de anillo mediante la conexión en cadena o la topología de bus. Además, en la figura 2B, se describe el ejemplo en que el número de cámaras es de ocho, pero el número de cámaras puede ser menor o mayor de ocho.

El dispositivo de control 120 genera parámetros de cámara y una máscara de estructura, y los suministra al dispositivo de generación de modelo tridimensional 140. Los parámetros de cámara incluyen parámetros externos que representan la posición y orientación (dirección de la línea de visión) de cada cámara, y parámetros internos que representan la longitud focal y el ángulo de visualización (área de captura de imágenes) de una lente incluida en cada cámara, y se obtienen por calibración. La calibración es el proceso para encontrar una relación de correspondencia entre un punto en el sistema de coordenadas mundo tridimensional obtenido utilizando una serie de imágenes en las que se captura un patrón específico, tal como un patrón de cuadros, y un punto bidimensional correspondiente al mismo. La máscara de estructura es una imagen de máscara que representa una silueta bidimensional de la estructura existente en cada imagen capturada adquirida por cada una de las cámaras 211 a 218. La imagen de máscara es una imagen de referencia que especifica cuál es la parte objetivo de extracción dentro de la imagen capturada y una imagen binaria representada por 0 y 1. En la presente realización, la portería de fútbol 202 se trata como una estructura, y la imagen de silueta que indica el área de la portería de fútbol 202 (silueta bidimensional) dentro de la imagen capturada por cada cámara desde una posición predeterminada y un ángulo predeterminado es la máscara de estructura. Como imagen capturada que es la fuente de la máscara de estructura, puede ser posible utilizar una capturada en un tiempo en que no aparece un jugador o similar, que es un primer plano, tal como un tiempo antes o después del partido o un tiempo durante el tiempo de descanso. Sin embargo, existe un caso en el que una imagen capturada previa o posteriormente no es adecuada dado que la captura de imágenes se ve afectada por la variación de la luz solar, por ejemplo, en exteriores. En un caso así, por ejemplo, puede ser posible obtener la imagen capturada que es la fuente de la máscara de estructura eliminando jugadores o similar, de un número predeterminado de cuadros (por ejemplo, cuadros sucesivos correspondientes a diez segundos) de la imagen en movimiento en la que se capturan los jugadores o similar. En este caso, es posible obtener la máscara de estructura en base a la imagen que adopta el promedio de cada valor de píxel en cada cuadro.

El dispositivo de separación del primer plano 130 lleva a cabo un procesamiento para determinar el primer plano correspondiente a los jugadores y al balón sobre el campo 200, diferenciado del área de fondo excepto por el área del primer plano para cada una de las imágenes capturadas a partir de la serie de puntos de vista, que son introducidos. Para la determinación del área del primer plano se utiliza una imagen del fondo preparada previamente (puede ser la misma imagen capturada que es la fuente de la máscara de estructura). Específicamente, se encuentra la diferencia respecto de la imagen de fondo para cada imagen capturada y el área correspondiente a la diferencia se especifica como el área del primer plano. Debido a esto, se genera la máscara del primer plano que indica el área del primer plano para cada imagen capturada. En la presente realización, se genera como la máscara del primer plano una imagen binaria que representa el píxel perteneciente al área del primer plano que representa los jugadores y el balón mediante "0" y el píxel perteneciente al área del fondo excepto para el primer plano mediante "1".

5 El dispositivo de generación de modelo tridimensional 140 genera un modelo tridimensional de un objeto en base a los parámetros de cámara y a la imagen de múltiples puntos de vista. Los detalles del dispositivo de generación de modelo tridimensional 140 se describirán posteriormente. Los datos del modelo tridimensional generado se entregan al dispositivo de representación 150.

10 El dispositivo de representación 150 genera una imagen de punto de vista virtual en base al modelo tridimensional recibido del dispositivo de generación de modelo tridimensional 140, a los parámetros de cámara recibidos del dispositivo de control 120, a la imagen de primer plano recibida del dispositivo de separación del primer plano 130 y a la imagen de fondo preparada previamente. Específicamente, una relación de posiciones entre la imagen del primer plano y el modelo tridimensional se encuentra a partir de los parámetros de cámara y, mapeando la imagen del primer plano correspondiente al modelo tridimensional se genera una imagen de punto de vista virtual en un caso en que un objeto de interés se observa desde un ángulo arbitrario. De este modo, por ejemplo, es posible obtener una imagen de punto de vista virtual de una escena decisiva frente a la portería en la que un jugador ha marcado un gol.

15 La configuración del sistema de generación de imágenes de punto de vista virtual mostrado en la figura 2A es un ejemplo, y la configuración no se limita a esto. Por ejemplo, un ordenador puede incluir las funciones de una serie de dispositivos (por ejemplo, el dispositivo de separación del primer plano 130, el dispositivo de generación de modelo tridimensional 140, y similares). Alternativamente, puede ser posible asimismo diseñar una configuración en la que se hace que el módulo de cada cámara tenga la función del dispositivo de separación del primer plano 130 y, desde cada cámara, se suministran la imagen capturada y los datos de la máscara del primer plano.

20 (Dispositivo de generación de modelo tridimensional)

25 La figura 3 es un diagrama de bloques funcionales que muestra la configuración interna del dispositivo de generación de modelo tridimensional 140, según la presente realización. El dispositivo de generación de modelo tridimensional 140 incluye una unidad de recepción de datos 310, una unidad de almacenamiento de la máscara de estructura 320, una unidad de combinación de máscaras 330, una unidad de conversión de coordenadas 340, una unidad de generación de modelo tridimensional 350 y una unidad de salida de datos 360. A continuación se explica en detalle cada unidad.

30 La unidad de recepción de datos 310 recibe los parámetros de cámara de cada cámara que configura el conjunto de cámaras 110 y la máscara de estructura que representa la silueta bidimensional de la estructura existente dentro de la escena de captura de imágenes, desde el dispositivo de control 120. Además, la unidad de recepción de datos 310 recibe la imagen capturada (imagen de múltiples puntos de vista) obtenida por cada cámara del conjunto de cámaras 110, y los datos de la máscara del primer plano que representan la silueta bidimensional del primer plano existente dentro de cada imagen capturada, desde el dispositivo de separación del primer plano 130. De los datos recibidos, la máscara de estructura se suministra a la unidad de almacenamiento de la máscara de estructura 320, la máscara del primer plano a la unidad de combinación de máscaras 330, la imagen de múltiples puntos de vista a la unidad de conversión de coordenadas 340, y los parámetros de cámara a la unidad de conversión de coordenadas 340 y a la unidad de generación de modelo tridimensional 350, respectivamente.

35 La unidad de almacenamiento de la máscara de estructura 320 almacena la máscara de estructura en la RAM o similar, y suministra la máscara de estructura a la unidad de combinación de máscaras 330, cuando sea necesario.

40 La unidad de combinación de máscaras 330 lee la máscara de estructura desde la unidad de almacenamiento de la máscara de estructura 320 y la combina con la máscara del primer plano recibida desde la unidad de recepción de datos 310, y genera de ese modo una imagen de máscara que integra ambas máscaras en una máscara (en adelante, denominada una "máscara integrada"). La máscara integrada generada es enviada a la unidad de generación de modelo tridimensional 350.

45 La unidad de conversión de coordenadas 340 convierte la imagen de múltiples puntos de vista recibida desde la unidad de recepción de datos 310, del sistema de coordenadas de la cámara al sistema de coordenadas mundo en base a los parámetros de cámara. Mediante esta conversión de coordenadas, cada imagen capturada cuyo punto de vista es diferente de otro se convierte en información que representa qué área indica cada imagen capturada en el espacio tridimensional.

50 La unidad de generación de modelo tridimensional 350 genera un modelo tridimensional de un objeto que incluye una estructura dentro de una escena de captura de imágenes, mediante el procedimiento de intersección de volúmenes visuales, utilizando la imagen de múltiples puntos de vista convertida al sistema de coordenadas mundo y la máscara integrada correspondiente a cada cámara. Los datos del modelo tridimensional generado del objeto son entregados al dispositivo de representación 150 por medio de la unidad de salida de datos 360.

(Proceso de formación de modelo tridimensional)

La figura 4 es un diagrama de flujo que muestra un flujo del proceso de formación de modelo tridimensional, según la presente realización. Esta serie de procesos es implementada por una CPU incluida en el dispositivo de generación de modelo tridimensional 140, que carga en una RAM un programa predeterminado almacenado en un medio de almacenamiento, tal como una ROM y un HDD, y ejecuta el programa. A continuación se proporciona una explicación a lo largo del flujo de la figura 4.

En primer lugar, en la etapa 401, la unidad de recepción de datos 310 recibe la máscara de estructura que representa la silueta bidimensional de la estructura (en este caso, la portería de fútbol 202) en un caso en que la estructura se ve desde cada una de las cámaras 211 a 218, y los parámetros de cámara de cada cámara, desde el dispositivo de control 120. Cada una de las figuras 5A a 5H muestra una imagen capturada por cada una de las cámaras 211 a 218 que configuran el conjunto de cámaras 110. En este caso, un jugador (portero) sale frente a la portería de fútbol 202 en la cancha de fútbol 201. A continuación, en cada imagen capturada en las figuras 5A, 5B y 5H, la portería de fútbol 202 existe entre la cámara y el jugador y, por lo tanto, una parte del jugador está oculta por la portería de fútbol 202. A partir de cada imagen capturada en las figuras 5A a 5H se obtiene respectivamente la máscara de estructura, en la que el área está representada por dos valores, tales como el área de la portería de fútbol 202 es 1 (blanco) y el área excepto el área de la portería de fútbol 202 es 0 (negro). Cada una de las figuras 6A a 6H muestra la máscara de estructura correspondiente a cada imagen capturada de las figuras 5A a 5H.

A continuación, en la etapa 402, la unidad de recepción de datos 310 recibe la máscara del primer plano que indica la silueta bidimensional del primer plano (en este caso, jugadores y balón) en la imagen capturada por cada una de las cámaras 211 a 218, desde el dispositivo de separación del primer plano 130 junto con la imagen de múltiples puntos de vista que es la fuente de la máscara del primer plano. Cada una de las figuras 7A a 7H muestra la máscara del primer plano correspondiente a cada imagen capturada en las figuras 5A a 5H. El dispositivo de separación del primer plano 130 extrae el área que cambia con el tiempo entre las imágenes capturadas desde el mismo ángulo, como primer plano y, por lo tanto, en cada una de las figuras 7A, 7B y 7H, el área de una parte del jugador oculta por la portería de fútbol 202 no se extrae como área del primer plano. Los datos de la máscara del primer plano recibida se envían a la unidad de combinación de máscaras 330.

A continuación, en la etapa 403, la unidad de combinación de máscaras 330 realiza el procesamiento para leer los datos de la máscara de estructura a partir de la unidad de almacenamiento de la máscara de estructura 320 y combinar la máscara de estructura leída y la máscara del primer plano recibida de la unidad de recepción de datos 310. Esta combinación es el proceso de cálculo para encontrar la OR lógica para cada píxel de la máscara del primer plano y la máscara de estructura, ambas representadas por dos valores (blanco y negro). Cada una de las figuras 8A a 8H muestra una máscara integrada obtenida combinando cada máscara de estructura mostrada en las figuras 6A a 6H y cada máscara del primer plano mostrada en las figuras 7A a 7H. En la máscara integrada completa no se ve ningún defecto en la silueta del jugador.

A continuación, en la etapa 404, la unidad de generación de modelo tridimensional 350 genera un modelo tridimensional utilizando el procedimiento de intersección de volúmenes visuales, en base a la máscara integrada obtenida en la etapa 403. Debido a esto, un modelo (en adelante denominado "modelo tridimensional integrado") que representa la forma tridimensional del primer plano y la estructura existente en el área de captura de imágenes común de una serie de imágenes capturadas desde diferentes puntos de vista. En un caso de la presente realización, se genera el modelo tridimensional integrado que incluye la portería de fútbol 202, además del jugador y el balón. La generación del modelo tridimensional integrado se lleva a cabo específicamente mediante el procedimiento, como sigue. En primer lugar, se preparan datos de volumen en los que el espacio tridimensional en el campo 200 se llena con cubos (vóxeles) que tienen un tamaño predeterminado. El valor del vóxel que configura los datos de volumen está representado por 0 y 1, y "1" indica un área que contribuye a la creación de la forma y "0" indica un área que no contribuye a la creación de la forma, respectivamente. A continuación, las coordenadas tridimensionales del vóxel se convierten del sistema de coordenadas mundo al sistema de coordenadas de la cámara utilizando los parámetros de cámara (posición de instalación, dirección de la línea de visión, y similares) de cada una de las cámaras 211 a 218. A continuación, en un caso en que la estructura y el primer plano indicados por la máscara integrada existen en el sistema de coordenadas de la cámara, se genera un modelo de representación de la forma tridimensional de la estructura y el primer plano mediante vóxeles. Puede ser posible asimismo representar una forma tridimensional mediante un conjunto de puntos (nube de puntos) que indican el centro del vóxel, en lugar del propio vóxel. La figura 9 muestra el modelo tridimensional integrado que se genera en base a la máscara integrada mostrada en las figuras 8A a 8H, y el símbolo 901 corresponde a la forma tridimensional del jugador, que es el primer plano, y el símbolo 902 corresponde a la forma tridimensional de la portería de fútbol 202, que es la estructura. Tal como se ha descrito anteriormente, en la máscara integrada no hay ningún defecto en la silueta del jugador, que es el primer plano y, por lo tanto, tampoco hay defectos en el modelo tridimensional integrado completo. La figura 10 muestra un modelo tridimensional que se genera utilizando solamente la máscara del primer plano mediante el procedimiento convencional. Tal como se ha descrito anteriormente, en la máscara del primer plano mostrada en las figuras 7A, 7B y 7H, una parte del jugador no está representada como el área del primer plano y, por lo tanto, la parte se convierte en un defecto en el modelo tridimensional generado. Mediante el procedimiento de la presente realización, utilizando la imagen de máscara que combina la máscara del primer plano

y la máscara de estructura se posibilita evitar que se produzca un defecto en una parte del modelo tridimensional del primer plano.

5 Lo anterior son los contenidos del proceso de formación del modelo tridimensional, según la presente realización. En un caso en el que se genera una imagen de punto de vista virtual de una imagen en movimiento, un modelo tridimensional para cada cuadro se genera realizando repetidamente el proceso de cada etapa descrita anteriormente, en unidades de cuadros. Sin embargo, solamente es necesario realizar la recepción y almacenamiento de la máscara de estructura (etapa 401) inmediatamente después del comienzo del flujo, y es posible omitir los cuadros segundo y subsiguientes. Además, en un caso en que se lleva a cabo captura de imágenes en la misma ubicación de captura de imágenes cambiando la fecha, es posible asimismo realizar la recepción y almacenamiento de la máscara de estructura solamente en la primera ocasión y almacenarla en la RAM o similar, y utilizar en los tiempos posterior y subsiguientes la máscara de estructura almacenada.

15 Tal como se ha descrito anteriormente, de acuerdo con la presente realización, incluso en un caso en el que existe una estructura que oculta un objeto, que es un primer plano, es posible generar un modelo tridimensional extremadamente preciso sin defectos en el primer plano, o en el que se reduce la magnitud del defecto.

[Segunda realización]

20 En la primera realización, se genera un modelo tridimensional de un primer plano sin defectos o en el que se reduce la magnitud del defecto, que incluye una estructura existente dentro de una escena de captura de imágenes. A continuación, se explica un aspecto como segunda realización, en el que se genera un modelo tridimensional con solamente el primer plano, del que se extrae una estructura, y en el que no hay defectos, o en el que se reduce la magnitud del defecto. La explicación de los contenidos comunes con los de la primera realización, tal como la configuración del sistema, se omite o simplifica y en lo que sigue, se explican principalmente los puntos diferentes.

La configuración del dispositivo de generación de modelo tridimensional 140 de la presente realización es asimismo básicamente igual que la de la primera realización (ver la figura 3), pero diferente en los puntos siguientes.

30 En primer lugar, la lectura de la máscara de estructura para la unidad de almacenamiento de la máscara de estructura 320 se lleva a cabo no solamente mediante la unidad de combinación de máscaras 330 sino asimismo mediante la unidad de generación de modelo tridimensional 350. La flecha de línea discontinua de la figura 3 lo indica. A continuación, en la unidad de generación de modelo tridimensional 350, además de la generación de un modelo tridimensional integrado del primer plano y de la estructura utilizando la máscara integrada, se realiza asimismo la generación de un modelo tridimensional de solamente la estructura utilizando la máscara de estructura. Después, hallando la diferencia entre el modelo tridimensional integrado generado en base a la máscara integrada y el modelo tridimensional de la estructura generado en base a la máscara de estructura, se extrae un modelo tridimensional de solamente el primer plano sin defectos, o en el que se reduce la magnitud del defecto.

40 (Proceso de formación de modelo tridimensional)

La figura 11 es un diagrama de flujo que muestra un flujo del proceso de formación de modelo tridimensional, según la presente realización. La serie de procesos es implementada por una CPU incluida en el dispositivo de generación de modelo tridimensional 140, que carga en una RAM un programa predeterminado almacenado en un medio de almacenamiento, tal como una ROM y un HDD, y ejecuta el programa. A continuación se proporciona una explicación a lo largo del flujo de la figura 11.

50 Las etapas 1101 a 1104 corresponden a las etapas 401 a 404, respectivamente, en el flujo de la figura 4 de la primera realización y no hay ninguna diferencia y, por lo tanto, se omite la explicación.

En la siguiente etapa 1105, la unidad de generación de modelo tridimensional 350 lee la máscara de estructura a partir de la unidad de almacenamiento de la máscara de estructura 320 y genera un modelo tridimensional de la estructura mediante el procedimiento de intersección de volúmenes visuales.

55 A continuación, en la etapa 1106, la unidad de generación de modelo tridimensional 350 encuentra la diferencia entre el modelo tridimensional combinado del primer plano y la estructura generado en la etapa 1104 y el modelo tridimensional de la estructura generado en la etapa S1105, y extrae el modelo tridimensional de solamente el primer plano. En este caso, puede ser posible asimismo encontrar la diferencia a partir del modelo tridimensional integrado después de expandir el modelo tridimensional de la estructura mediante, por ejemplo, aproximadamente el 10 % del espacio tridimensional. Debido a esto, es posible eliminar de forma segura la parte correspondiente a la estructura respecto del modelo tridimensional integrado. En este momento, es posible asimismo expandir solamente una parte del modelo tridimensional de la estructura. Por ejemplo, puede ser posible asimismo determinar una parte que se expande de acuerdo con el área, de tal modo que el lateral de la cancha 201 no se expande y solamente el lado opuesto de la cancha 201 se expande en un caso de la portería de fútbol 202 debido a que la probabilidad de que haya un jugador dentro de la cancha de fútbol 201 es alta. Además, puede ser posible asimismo cambiar la relación de expansión (relación de expansión) según cómo de lejos esté el objeto que es el primer plano, tal como un jugador

y un balón, respecto de la estructura. Por ejemplo, en un caso en que el objeto que es el primer plano está situado en una posición lejos de la estructura, la relación de expansión se incrementa de tal modo que el modelo tridimensional de la estructura es extraído de manera segura. Además, al reducir la relación de expansión en un caso en que el objeto que es el primer plano está situado en una posición cerca de la estructura, se impide que la parte del modelo tridimensional del primer plano sea eliminada erróneamente. Puede ser posible asimismo cambiar linealmente la relación de expansión en este momento, en función de la distancia desde el primer plano, o determinar por etapas la relación de expansión, proporcionando como referencia una, o una serie de distancias.

La figura 12A muestra un modelo tridimensional integrado generado en base a la máscara integrada, la misma que en la figura 9 descrita anteriormente. La figura 12B muestra un modelo tridimensional de la estructura generada en base solamente a la máscara de estructura. A continuación, la figura 12C muestra un modelo tridimensional de solamente el primer plano obtenido a partir de la diferencia entre el modelo tridimensional integrado en la figura 12A y el modelo tridimensional de la estructura de la figura 12B.

Lo anterior son los contenidos del proceso de formación del modelo tridimensional, según la presente realización. En un caso en el que se genera una imagen de punto de vista virtual de una imagen en movimiento, se genera un modelo tridimensional para cada cuadro realizando repetidamente el proceso de cada etapa descrita anteriormente, en unidades de cuadros. Sin embargo, solamente es necesario llevar a cabo la recepción y el almacenamiento de la máscara de estructura (etapa 1101) y la generación del modelo tridimensional de la estructura (etapa 1105) inmediatamente después del comienzo del flujo, y es posible omitir el segundo y subsiguientes cuadros. Además, en un caso en que la captura de imágenes se lleva a cabo en la misma ubicación de captura de imágenes cambiando la fecha, puede ser posible asimismo realizar la recepción y almacenamiento de la máscara de estructura y la generación del modelo tridimensional de la estructura solamente en la primera ocasión, y almacenarlos en la RAM o similar, y utilizarlos en la ocasión siguiente y subsiguientes. Igual que antes, según la presente realización, incluso en un caso en que existe una estructura que oculta un objeto que es un primer plano, es posible generar un modelo tridimensional de solamente el primer plano, que no incluye la estructura y es muy preciso.

[Tercera realización]

En la primera y la segunda realizaciones, el modelo tridimensional de solamente el primer plano se genera restando el modelo tridimensional de la estructura respecto del modelo tridimensional integrado del primer plano y de la estructura. A continuación, se explica un aspecto como tercera realización, donde el modelo tridimensional de solamente el primer plano se encuentra contando en qué imagen de máscara está incluido un vóxel, para cada vóxel que configura el modelo tridimensional integrado del primer plano y de la estructura (o para cada área predeterminada), y eliminando del modelo tridimensional integrado la parte cuyo valor de cómputo es menor o igual que un valor umbral.

En la presente realización, en primer lugar, para cada una de una serie de áreas parciales que configuran el espacio tridimensional, se determina si se cumple o no una condición de que el número de cámaras para las que está incluida el área parcial en el área del primer plano que indica el área del objeto objetivo dentro de la imagen capturada de una serie de cámaras es menor o igual que un primer valor umbral. Como primer valor umbral, se ajusta un valor arbitrario menor que el número total de cámaras, teniendo en cuenta la posición de instalación, la dirección de la línea de visión, y similares, de cada cámara. A continuación, se determina un modelo tridimensional de un objeto objetivo que incluye el área parcial para la que no se determina que se cumpla la condición.

(Procedimiento de representación de modelo tridimensional)

La figura 13A muestra un único vóxel cúbico. La figura 13B muestra un conjunto de vóxeles que representan un espacio objetivo de generación de modelo tridimensional. Tal como se muestra en la figura 13B, el vóxel es un área parcial mínima que configura el espacio tridimensional. A continuación, la figura 13C muestra un ejemplo en el que un conjunto de vóxeles de un modelo tridimensional de una pirámide cuadrangular se genera eliminando los vóxeles en el área diferente del área de la pirámide cuadrangular del conjunto de la figura 13B, que es el conjunto de vóxeles en el espacio objetivo. En la presente realización se explica un ejemplo en el que el espacio tridimensional y el modelo tridimensional se configuran mediante vóxeles cúbicos, pero esto no está limitado y el espacio tridimensional y el modelo tridimensional se pueden configurar mediante una nube de puntos o similares.

(Configuración del sistema)

El diagrama de bloques que muestra el ejemplo de configuración del sistema de generación de imágenes de punto de vista virtual que incluye el dispositivo de generación de modelo tridimensional, según la presente realización, es igual que el mostrado en la figura 2A y, por lo tanto, se omite la explicación.

El conjunto de cámaras 110 es un grupo de aparatos de captura de imágenes que incluye una serie de cámaras 110a a 110z y captura un objeto desde varios ángulos y entrega imágenes al dispositivo de separación del primer plano 130 y al dispositivo de control 120. Se supone que las cámaras 110a a 110z, el dispositivo de separación del primer plano 130 y el dispositivo de control 120 están conectados por una topología de estrella, pero pueden estar

conectados por una topología de anillo, bus, o similar mediante una conexión en cadena. El conjunto de cámaras 110 está dispuesto en la periferia del estadio deportivo, por ejemplo, tal como se muestra en la figura 14, y lleva a cabo captura de imágenes en sincronización hacia un punto de mira del campo, común a todas las cámaras, desde varios ángulos. Sin embargo, puede ser posible asimismo ajustar una serie de puntos de mira, tal como un punto de mira hacia el que se hace que miren la de mitad las cámaras incluidas en el conjunto de cámaras 110, y otro punto de mira hacia el que se hace que miren el resto de las cámaras.

En este caso, el primer plano es un objeto objetivo predeterminado (objeto que es un objetivo para el que se genera un modelo tridimensional en base a las imágenes capturadas) que permite la visualización desde un ángulo arbitrario en un punto de vista virtual y, en la presente realización, se refiere a una persona existente en el campo del estadio deportivo. Por otra parte, el fondo es el área excepto el primer plano y, en la presente realización, se refiere a todo el estadio deportivo (campo, tribuna de espectadores, y similares). Sin embargo, el primer plano y el fondo no se limitan a estos ejemplos. Además, se supone que la imagen de punto de vista virtual en la presente realización incluye todas las imágenes que representan el aspecto desde el punto de vista virtual en el que no hay instalada ninguna cámara, no solamente las imágenes que representan el aspecto desde el punto de vista que se puede especificar libremente.

El dispositivo de control 120 calcula parámetros de cámara que indican la posición y la orientación de las cámaras 110a a 110z, a partir de las imágenes capturadas en sincronización mediante el conjunto de cámaras 110, y entrega los parámetros de cámara calculados al dispositivo de generación de modelo tridimensional 140. En este caso, los parámetros de cámara incluyen parámetros externos y parámetros internos. Los parámetros externos incluyen una matriz de rotación y una matriz de traslación, e indican la posición y orientación de la cámara. Los parámetros internos incluyen información sobre la longitud focal, el centro óptico, y similares, de la cámara e indican el ángulo de visualización de la cámara, el tamaño del sensor de formación de imágenes, y similares.

El proceso para calcular el parámetro de cámara se denomina calibración. Es posible encontrar el parámetro de cámara utilizando una relación de correspondencia entre puntos en el sistema de coordenadas mundo tridimensional adquiridos utilizando una serie de imágenes obtenidas capturando un patrón específico, por ejemplo, tal como un damero, y puntos bidimensionales correspondientes al mismo.

El dispositivo de control 120 calcula una imagen de máscara de estructura que indica una área de estructura que tiene una probabilidad de solapar delante del primer plano en las imágenes capturadas por las cámaras 110a a 110z, y entrega la información sobre la imagen de máscara de estructura calculada. En la presente realización, una estructura es un objeto estacionario instalado dentro del espacio objetivo de captura de imágenes y, como ejemplo, la portería de fútbol se trata como estructura y la imagen que indica el área de la portería dentro de la imagen capturada por cada cámara es una imagen de máscara de estructura.

El dispositivo de separación del primer plano 130 identifica el área en la que hay una persona en el campo, como el primer plano, y el área del fondo excepto para el primer plano a partir de las imágenes capturadas por la serie de cámaras, que son introducidas desde el conjunto de cámaras 110, y entrega la imagen de máscara del primer plano que indica el área del primer plano. Como procedimiento de identificación del área del primer plano, es posible utilizar un procedimiento de identificación de un área en la que existe una diferencia entre la imagen de fondo almacenada previamente y la imagen capturada, como la imagen de primer plano, o un procedimiento de identificación de un área de un objeto en movimiento, como el área del primer plano.

En este caso, la imagen de máscara es una imagen de referencia que representa una parte específica que se desea extraer de una imagen capturada y una imagen binaria representada por 0 y 1. Por ejemplo, la imagen de máscara del primer plano indica el área en la que el primer plano, por ejemplo un jugador, existe en la imagen capturada, y una imagen en la que el píxel que indica el área del primer plano está representado por 1 y el píxel diferente del primer plano por 0, a la misma resolución que la del área capturada. Sin embargo, el formato de una imagen de máscara no se limita a esto, y puede ser cualquier información que indique el área del objeto específico dentro de una imagen capturada.

El dispositivo de generación de modelo tridimensional 140 tiene la función de un aparato de procesamiento de información que genera un modelo tridimensional utilizando una serie de imágenes capturadas, capturadas por la serie de cámaras. En primer lugar, el dispositivo de generación de modelo tridimensional 140 recibe los parámetros de cámara y la información sobre la imagen de máscara de estructura desde el dispositivo de control 120, y recibe la imagen de máscara del primer plano desde el dispositivo de separación del primer plano 130. A continuación, el dispositivo de generación de modelo tridimensional 140 genera una imagen de máscara integrada que indica un área integrada, integrando la imagen de máscara de estructura y la imagen de máscara del primer plano. Además, el dispositivo de generación de modelo tridimensional 140 determina si eliminar o no cada vóxel en base al número de cámaras para las que cada vóxel (de hecho, un punto correspondiente a un vóxel, y esto aplica en lo que sigue) dentro del espacio que es el objetivo de generación del modelo tridimensional del primer plano no está incluido en la imagen de máscara integrada, y el número de cámaras para las que cada vóxel está incluido en la imagen de máscara del primer plano. A continuación, en base a los vóxeles restantes después de la eliminación de vóxeles determinados para eliminar, se genera el modelo tridimensional del primer plano, por ejemplo, mediante el

procedimiento de intersección de volúmenes visuales y se entrega al dispositivo de representación 150.

El dispositivo de representación 150 recibe el modelo tridimensional desde el dispositivo de generación de modelo tridimensional 140 y recibe la imagen que indica el primer plano desde el dispositivo de separación del primer plano 130. Además, el dispositivo de representación 150 lleva a cabo coloreado hallando la relación de posiciones entre la imagen que indica el primer plano y el modelo tridimensional a partir de los parámetros de cámara, y pega la imagen de primer plano correspondiente al modelo tridimensional y, de ese modo, genera una imagen de punto de vista virtual en un caso en el que el modelo tridimensional se observa desde un punto de vista arbitrario. En la imagen de punto de vista virtual se puede incluir la imagen del fondo. Es decir, puede ser posible asimismo que el dispositivo de representación 150 genere una imagen de punto de vista virtual en un caso en que el fondo y el primer plano se ven desde un punto de vista de ajuste, ajustando el modelo del fondo, el modelo del primer plano y la posición del punto de vista dentro del espacio tridimensional.

(Configuración funcional del dispositivo de generación de modelo tridimensional)

Después de lo anterior, haciendo referencia a la figura 15, se explica la configuración funcional del dispositivo de generación de modelo tridimensional, según la presente realización. El dispositivo de generación de modelo tridimensional 140 incluye una unidad de recepción 155, una unidad de almacenamiento de la máscara de estructura 101, una unidad de almacenamiento 102 de parámetros de cámara, una unidad de integración de máscaras 103, una unidad de conversión de coordenadas 104, una unidad de determinación interior/exterior de máscara 105, una unidad de ajuste de valores umbral 106, una unidad de generación de modelo de primer plano 107 y una unidad de salida 108.

La unidad de recepción 155 recibe los parámetros de cámara de cada cámara que configura el conjunto de cámaras 110 y la imagen de máscara de estructura que indica el área de la estructura, del dispositivo de control 120. Además, la unidad de recepción 155 recibe la imagen capturada por cada cámara del conjunto de cámaras 110 y la imagen de máscara del primer plano que indica el área del primer plano dentro de la imagen, del dispositivo de separación del primer plano 130 en cada ocasión de captura de imágenes.

La unidad de almacenamiento de la máscara de estructura 101 almacena la imagen de máscara de estructura recibida por la unidad de recepción 155. La imagen de máscara de estructura es una imagen fija, en función de la posición de la cámara.

La unidad de almacenamiento 102 de parámetros de cámara almacena parámetros externos que indican la posición y/o la orientación de cada cámara, que son capturados por el conjunto de cámaras 110, y parámetros internos que indica la longitud focal y/o el tamaño de la imagen, como parámetros de cámara.

La unidad de integración de máscaras 103 genera una imagen de máscara integrada, integrando la imagen de máscara del primer plano recibida del dispositivo de separación del primer plano 130 cada vez que se lleva a cabo captura de imágenes mediante el conjunto de cámaras 110, y la imagen de máscara de estructura almacenada en la unidad de almacenamiento de la máscara de estructura 101. Más adelante se describirán detalles del procedimiento de integración de la imagen de máscara del primer plano y la imagen de máscara de estructura.

La unidad de conversión de coordenadas 104 calcula la posición y el ángulo de visualización de cada imagen capturada en el sistema de coordenadas mundo, en base a los parámetros de cámara almacenados en la unidad de almacenamiento 102 de parámetros de cámara, y los convierte en información que representa qué área en el espacio tridimensional indica cada imagen capturada.

La unidad de determinación interior/exterior de máscara 105 determina, en caso de que el número de cámaras para las que cada vóxel dentro del espacio de vóxeles objetivo está incluido en el interior de la imagen de máscara del primer plano sea menor o igual que un valor umbral, eliminar el vóxel. Además, la unidad de determinación interior/exterior de máscara 105 determina, en caso de que el número de cámaras para las que cada vóxel dentro del espacio de vóxeles objetivo no está incluido en el interior de la imagen de máscara integrada sea mayor o igual que otro valor umbral, eliminar el vóxel.

La unidad de ajuste de valores umbral 106 ajusta cada valor umbral para determinar si eliminar o no un vóxel mediante la unidad de determinación interior/exterior de máscara 105. Este valor umbral se puede ajustar en función de una operación de usuario para el dispositivo de generación de modelo tridimensional 140 o puede ser ajustado automáticamente por la unidad de ajuste de valores umbral 106. La unidad de generación de modelo del primer plano 107 elimina vóxeles que la unidad de determinación interior/exterior de máscara 105 ha determinado eliminar, de los vóxeles dentro del espacio de vóxeles objetivo, y genera un modelo tridimensional en base a los vóxeles restantes. La unidad de salida 108 entrega el modelo tridimensional generado por la unidad de generación de modelo de primer plano 107 al dispositivo de representación 150.

La figura 16 es un diagrama de flujo que muestra un procedimiento de procesamiento llevado a cabo por el dispositivo de generación de modelo tridimensional, según la presente realización.

En S1601, la unidad de recepción 155 recibe del dispositivo de control 120 la imagen de máscara de estructura de cada cámara que configura el conjunto de cámaras 110. En este caso, se explica un ejemplo de la imagen capturada y la imagen de máscara de estructura. Las figuras 17A a 17E muestran ejemplos de cinco imágenes capturadas, capturadas por cinco cámaras que configuran una parte del conjunto de cámaras 110. En este caso hay una persona en el campo y la portería está en el campo como una estructura, y en las figuras 17B, 17C y 17D, la portería, que es una estructura, está frente a la persona y, por lo tanto, una parte de la persona está oculta. Cada una de las figuras 18A a 18E muestra la imagen de máscara de estructura correspondiente a cada imagen capturada mostrada en las figuras 17A a 17E. La imagen de máscara de estructura se muestra como una imagen binaria en la que el área de la portería, que es una estructura, es 1 (blanco) y el área exceptuando la estructura es 0 (negro).

En S1602, la unidad de recepción 155 recibe la imagen de máscara del primer plano que indica el área del primer plano, del dispositivo de separación del primer plano 130. En este caso, se explica un ejemplo de la imagen de máscara del primer plano. Cada una de las figuras 19A a 19E muestra la imagen de máscara del primer plano correspondiente a cada imagen capturada mostrada en las figuras 17A a 17E. El dispositivo de separación del primer plano 130 extrae un área que cambia con el tiempo, como área del primer plano y, por lo tanto, el área de una parte de la persona, que está oculta por la portería, no se extrae como área del primer plano, tal como en la figura 19B, la figura 19C y la figura 19D. Además, en la figura 19E, una parte de la pierna de la persona, que no cambia con el tiempo, no se extrae como área del primer plano.

En S1603, la unidad de integración de máscaras 103 genera una imagen de máscara integrada, integrando la imagen de máscara de estructura y la imagen de máscara del primer plano recibidas en S1601 y S1602. Cada una de las figuras 20A a 20E muestra un ejemplo de la imagen de máscara integrada, que es el resultado de integrar la imagen de máscara de estructura mostrada en cada una de las figuras 18A a 18E y la imagen de máscara del primer plano mostrada en cada una de las figuras 19A a 19E. La imagen de máscara integrada se calcula mediante OR (suma lógica) de la imagen de máscara del primer plano y la imagen de máscara de estructura, ambas representadas por dos valores.

En S1604, la unidad de determinación interior/exterior de máscara 105 selecciona del espacio de vóxeles objetivo un vóxel que aún no ha sido seleccionado.

En S1605, la unidad de determinación interior/exterior de máscara 105 cuenta el número de cámaras para las que el vóxel seleccionado no está incluido dentro del área de máscara de la imagen de máscara integrada de cada cámara (en adelante, denominado cómputo falso).

En S1606, la unidad de determinación interior/exterior de máscara 105 determina si el cómputo falso es o no mayor o igual que un valor umbral. En caso de que el cómputo falso sea mayor o igual que el valor umbral, es posible determinar que el vóxel seleccionado no es primer plano ni estructura y, por lo tanto, el proceso avanza a S1607. Debido a esto, es posible eliminar muchos vóxeles, que obviamente no son primer plano. Por otra parte, en caso de que el cómputo falso sea menor que el valor umbral, es posible determinar que el vóxel seleccionado es un primer plano o una estructura y, por lo tanto, el proceso avanza a S1608.

En S1607, la unidad de generación de modelo de primer plano 107 elimina el vóxel seleccionado del espacio de vóxeles objetivo. En S1608, la unidad de determinación interior/exterior de máscara 105 cuenta el número de cámaras para las que el vóxel seleccionado está incluido en el interior del área de máscara de la imagen de máscara del primer plano de cada cámara (en adelante, denominado cómputo verdadero).

En S1609, la unidad de determinación interior/exterior de máscara 105 determina si el cómputo verdadero es o no menor o igual que otro valor umbral. En el caso en que el cómputo verdadero es menor o igual que otro valor umbral, es posible determinar que el vóxel seleccionado es una estructura y, por lo tanto, el proceso avanza a S1607 y el píxel seleccionado es eliminado del espacio de vóxeles objetivo. Por otra parte, en caso de que el cómputo verdadero exceda otro valor umbral, es posible determinar que el vóxel seleccionado es un primer plano y, por lo tanto, el vóxel seleccionado no se elimina del espacio de vóxeles objetivo.

En S1610, la unidad de determinación interior/exterior de máscara 105 determina si el proceso se ha completado o no para todos los vóxeles dentro del espacio de vóxeles objetivo. En caso de que el proceso se haya completado para todos los vóxeles, el proceso avanza a S1611. Por otra parte, en caso de que el proceso no se haya completado para todos los vóxeles, el proceso vuelve a S1604, y se selecciona el siguiente vóxel de entre los vóxeles no seleccionados aún y a continuación se lleva a cabo el mismo proceso.

En S1611, la unidad de generación de modelo de primer plano 107 genera un modelo tridimensional del primer plano utilizando los vóxeles restantes después de que la determinación de eliminación de vóxeles se lleve a cabo para el espacio de vóxeles objetivo.

En S1612, la unidad de salida 108 entrega el modelo tridimensional del primer plano generado por la unidad de

generación de modelo de primer plano 107 al dispositivo de representación 150. La serie anterior de procesos se lleva a cabo para cada cuadro capturado por cada cámara.

5 En este caso, se explica un ejemplo de generación de un modelo tridimensional tomando como ejemplo el sistema de generación de imágenes de punto de vista virtual que captura el estadio deportivo mediante las 16 cámaras mostradas en la figura 14. La figura 21 es un diagrama que muestra un espacio de vóxeles, que es un objetivo de generación de modelo tridimensional, del sistema del estadio deportivo, según la presente realización, y el área cuboide indicada mediante cuadrículas representa el espacio de vóxeles objetivo.

10 La figura 22 muestra un ejemplo de cómputo falso/cómputo verdadero y resultados de determinación de vóxeles para la persona, la pierna de la persona, la cabeza de la persona, la portería y la otra área, respectivamente, como el primer plano, el primer plano no detectado por una parte de las cámaras, el primer plano oculto por la estructura, la estructura, y el no primer plano, en un caso en el que el estadio deportivo es capturado por las 16 cámaras mostradas en la figura 14. En este caso, se supone que una cámara ha fallado en la extracción del primer plano de  
15 la pierna de la persona, y la cabeza de la persona está oculta por la portería, que es una estructura, en tres cámaras, y estas no se extraen como primer plano por el dispositivo de separación del primer plano 130.

En un caso en que el valor umbral de cómputo falso es un valor fijo de 10 en la determinación en S1606, el cómputo falso del vóxel situado en la otra área es 16 y supera el valor umbral y, por lo tanto, el vóxel es eliminado. Como resultado de esto, por ejemplo, se genera un modelo tridimensional que incluye el primer plano y la estructura que se muestra en la figura 23. En este caso, la figura 23 es un diagrama que muestra un ejemplo de un modelo tridimensional generado aplicando determinación del cómputo falso basada en umbrales.

25 Además, en un caso en que el valor umbral (otro valor umbral) de cómputo verdadero es un valor fijo de 5 en la determinación en S1609, el cómputo verdadero del vóxel situado en el área de la portería, que es una estructura, es 0 y menor o igual que el valor umbral y, por lo tanto, el vóxel es eliminado. Por otra parte, el cómputo verdadero de los vóxeles situados en las áreas de la persona, la pierna de la persona y la cabeza son 16, 15 y 13, respectivamente, superando el segundo valor umbral y, por lo tanto, los vóxeles no son eliminados.

30 Es decir, tal como se muestra en la figura 22, se determina que el primer plano (persona), el primer plano (pierna) cuya parte no se detecta, y el primer plano (cabeza) oculto por la estructura son un vóxel a mantener, y se determina que la estructura (portería) y el no primer plano (la otra área) son un vóxel a eliminar. Por consiguiente, a partir del conjunto de vóxeles del espacio objetivo mostrado en la figura 21, por ejemplo, se genera finalmente un modelo tridimensional de la persona sin defectos, tal como se muestra en la figura 24. En este caso, la figura 24 es un  
35 diagrama que muestra un ejemplo de modelo tridimensional generado aplicando la determinación basada en umbral del cómputo falso y la determinación basada en umbral del cómputo verdadero.

En contraste con esto, la figura 25 muestra un ejemplo en el que un modelo tridimensional es generado por el procedimiento de intersección de volúmenes visuales utilizando solamente la imagen de máscara del primer plano mostrada en las figuras 19A a 19E. En la figura 19A, toda la persona es capturada, pero en las imágenes capturadas mostradas en las figuras 19B, 19C y 19D, una parte de la cabeza de la persona está oculta por la portería, que es una estructura. Además, en la imagen capturada mostrada en la figura 19E, la pierna de la persona no se extrae como primer plano. Debido a esto, también en el modelo tridimensional generado, una parte se convierte en un defecto.  
45

Tal como se ha explicado anteriormente, en la presente realización, para cada vóxel dentro del espacio, que es el objetivo para el que se genera un modelo tridimensional del objeto objetivo (primer plano), se determina si el número de cámaras para las que el vóxel objetivo está incluido en la imagen de máscara del primer plano que indica el área del primer plano es o no menor o igual que el valor umbral (valor umbral de cómputo verdadero) y, en caso de que el número sea menor o igual que el valor umbral, el vóxel es eliminado.  
50

De acuerdo con la presente realización, incluso en un caso en el que hay un defecto en la imagen de máscara del primer plano que indica el área del objeto objetivo (primer plano), es posible evitar que se genere un defecto en un modelo tridimensional del objeto objetivo (primer plano) y mejorar la calidad del modelo tridimensional.  
55

Además, se genera una imagen de máscara integrada integrando la imagen de máscara del primer plano y la imagen de máscara de estructura y, en caso de que el número de cámaras para las que el vóxel objetivo no está incluido en la imagen de máscara integrada sea mayor o igual que el valor umbral (cómputo falso), se determina eliminar el vóxel. Debido a esto, es posible eliminar muchos vóxeles, que son obviamente un no primer plano y, por lo tanto, se posibilita mejorar la velocidad del proceso en la fase posterior.  
60

[Cuarta realización]

65 En la tercera realización descrita anteriormente, no se determina si un vóxel está o no en el interior del rango de captura de imágenes (en el interior del ángulo de visualización) para cada cámara y, por lo tanto, existe una probabilidad de que un vóxel que indica el primer plano sea eliminado erróneamente en un caso en el que el vóxel

está en el exterior del rango de captura de imágenes en muchas cámaras. Por ejemplo, en un caso en que el estadio deportivo es capturado por la disposición de cámaras que se muestra en la figura 14, el número de cámaras para las que el vóxel situado en el área de la persona situada en la proximidad de la portería en el lado opuesto del punto de mira está incluido en el interior del rango de captura de imágenes es de tres y, por lo tanto, el cómputo verdadero es 3. En este momento, en un caso en que el valor umbral del cómputo verdadero es de 5, el vóxel es eliminado debido a que el cómputo verdadero es menor que el umbral. Por consiguiente, como cuarta realización se explica un aspecto en el que el modelo tridimensional se genera de manera que el primer plano situado a parte del punto de mira no es eliminado mediante el ajuste de un valor umbral en base a resultados de la determinación interior/exterior del ángulo de visualización. En la presente realización, al calcular el valor umbral del cómputo verdadero en base al número de cámaras para las que el vóxel está incluido en el interior del rango de captura de imágenes (ángulo de visualización), se evita la eliminación errónea del vóxel que indica el primer plano incluso en un caso en que el vóxel está separado del punto de mira.

(Configuración funcional del dispositivo de generación de modelo tridimensional)

Haciendo referencia a la figura 26, se explica la configuración funcional del dispositivo de generación de modelo tridimensional, según la presente realización. El dispositivo de generación de modelo tridimensional 140, según la presente realización, incluye, además, una unidad de determinación interior/exterior del ángulo de visualización 109 y una unidad de cálculo del valor umbral 260, además de la unidad de recepción 155, la unidad de almacenamiento de la máscara de estructura 101, la unidad de almacenamiento 102 de parámetros de cámara, la unidad de integración de máscaras 103, la unidad de conversión de coordenadas 104, la unidad de determinación interior/exterior de máscara 105, la unidad de ajuste de valores umbral 106, la unidad de generación de modelo de primer plano 107 y la unidad de salida 108. La configuración básica del sistema de generación de imágenes de punto de vista virtual es la misma que la de las realizaciones primera a tercera y, por lo tanto, se omite la explicación. Además, la unidad de recepción 155, la unidad de almacenamiento de la máscara de estructura 101, la unidad de almacenamiento 102 de parámetros de cámara, la unidad de integración de máscaras 103, la unidad de conversión de coordenadas 104, la unidad de determinación interior/exterior de máscara 105, la unidad de ajuste de valores umbral 106, la unidad de generación de modelo de primer plano 107 y la unidad de salida 108 que configuran el dispositivo de generación de modelo tridimensional 140 son las mismas que las de la tercera realización y, por lo tanto, se omite la explicación.

La unidad de determinación interior/exterior del ángulo de visualización 109 determina si cada vóxel en el interior del espacio de vóxeles objetivo está o no dentro del rango de captura de imágenes de cada cámara, en base a los parámetros de cámara de cada cámara.

La unidad de cálculo del valor umbral 260 calcula un valor obtenido multiplicando el número de cámaras para las que se determina que cada vóxel está dentro del rango de captura de imágenes, por una razón predeterminada, como el valor umbral de cómputo verdadero. Por ejemplo, en un caso en el que el número de cámaras para las que un determinado vóxel está dentro del rango de captura de imágenes es de cinco y la razón predeterminada es del 60 %, el valor umbral del cómputo verdadero para el vóxel se calcula como 3. El valor umbral calculado por la unidad de cálculo del valor umbral 260 es entregado a la unidad de ajuste de valores umbral 106, y la unidad de ajuste de valores umbral 106 ajusta el valor umbral introducido desde la unidad de cálculo del valor umbral 260 como el valor umbral del cómputo verdadero.

En un caso en que el número de cámaras para las que un determinado vóxel está dentro del rango de captura de imágenes es menor que un número predeterminado, se considera que se reduce la precisión de un modelo tridimensional a generar y el procesamiento no es necesario y, por lo tanto, puede ser posible asimismo diseñar una configuración en la que el valor umbral se ajusta a un valor predeterminado en un caso en que el número de dichas cámaras es menor que un número predeterminado.

La figura 27 es un diagrama de flujo que muestra un procedimiento de procesamiento llevado a cabo por el dispositivo de generación de modelo tridimensional, según la presente realización. Cada elemento de procesamiento en S2701 a S2704 es igual que cada elemento de procesamiento en S1601 a S1604 en el flujo de la figura 16 de la tercera realización y, por lo tanto, se omite la explicación.

En S2705, la unidad de determinación interior/exterior del ángulo de visualización 109 determina si el vóxel seleccionado en S2704 está o no incluido en el interior del ángulo de visualización de cada cámara, en base a los parámetros de cámara de cada cámara.

En S2706, la unidad de determinación interior/exterior de máscara 105 cuenta el número de cámaras para las que el vóxel seleccionado no está incluido en el interior del área de máscara de la imagen de máscara integrada de cada cámara, y para las que el vóxel seleccionado está incluido en el interior del ángulo de visualización (en adelante, denominado el cómputo falso).

Cada elemento de procesamiento en S2707 a S2709 es igual que cada elemento de procesamiento en S1606 a S1608 en el flujo de la figura 16 descrito anteriormente y, por lo tanto, se omite la explicación.

En S2710, la unidad de cálculo del valor umbral 260 calcula el valor umbral de cómputo verdadero en base al número de cámaras para las que el vóxel seleccionado está incluido en el interior del ángulo de visualización. La unidad de ajuste de valores umbral 106 ajusta el valor umbral de cómputo verdadero calculado por la unidad de cálculo de valor umbral 260.

Cada elemento del proceso en S2711 a S2714 es igual que cada elemento de proceso en S1609 a S1612 en el flujo de la figura 16 descrito anteriormente y, por lo tanto, se omite la explicación. Lo anterior es la serie de procesos en el flujo de la figura 27.

En este caso, la figura 28 muestra la manera mediante la cual el estadio deportivo que incluye un primer plano A indicado por un punto negro en una posición cerca de un punto de mira indicado por la marca x en la figura 28 y un primer plano B indicado por un punto negro en una posición lejos del punto de mira, es capturado por las 16 cámaras de la figura 14. Se supone que el primer plano A está dentro del ángulo de visualización de la totalidad de las 16 cámaras y el primer plano B está dentro del ángulo de visualización de solamente las tres cámaras 110k, 110l y 110m.

Además, la figura 29 muestra un ejemplo de cómputo falso/cómputo verdadero de cada uno del vóxel en la posición del primer plano A cerca del punto de mira y el vóxel en la posición del primer plano B lejos del punto de mira, en la disposición de cámaras mostrada en la figura 28. Se supone que el valor umbral del cómputo falso es un valor fijo de 10 y el valor umbral del cómputo verdadero es el 70 % del número de cámaras para las que el vóxel está incluido dentro del ángulo de visualización.

El vóxel situado en el primer plano A cerca del punto de mira está incluido dentro de la imagen de máscara integrada de la totalidad de las 16 cámaras y, por lo tanto, no hay ninguna cámara para la que el vóxel esté fuera de la imagen de máscara integrada. Por consiguiente, el número de cámaras para las que el vóxel está fuera de la imagen de máscara integrada y para las que el vóxel está dentro del ángulo de visualización es cero y, por lo tanto, el cómputo falso es 0.

Además, el número de cámaras para las que el vóxel situado en el primer plano A cerca del punto de mira está incluido dentro del ángulo de visualización es asimismo 16 y, por lo tanto, el valor umbral del cómputo verdadero es de 11,2, que es el 70 % de 16. Entonces, el vóxel situado en el primer plano A cerca del punto de mira está dentro de la imagen de máscara del primer plano de todas las cámaras, y el cómputo verdadero es 16 y el valor de cómputo es mayor o igual que el valor umbral (11,2) y, por lo tanto, el vóxel no es eliminado.

El vóxel en la posición del primer plano B lejos del punto de mira está fuera del ángulo de visualización de las 13 cámaras (13 cámaras exceptuando las cámaras 110k, 110l y 110m) y dentro del ángulo de visualización de las tres cámaras (las cámaras 110k, 110l y 110m). Además, el vóxel está dentro de la imagen de máscara integrada de las tres cámaras (cámaras 110k, 110l y 110m). Por consiguiente, el número de cámaras para las que el vóxel está fuera de la imagen de máscara integrada y para las que el vóxel está dentro del ángulo de visualización es cero y, por lo tanto, el cómputo falso es 0.

Además, el número de cámaras para las que el vóxel situado en el primer plano B lejos del punto de mira está incluido dentro del ángulo de visualización es de tres y, por lo tanto, el valor umbral del cómputo verdadero es de 2,1, que es el 70 % de 3. A continuación, el vóxel situado en el primer plano B lejos del punto de mira está dentro de la imagen de máscara del primer plano de las tres cámaras, y el cómputo verdadero es 3 y el valor de cómputo es mayor o igual que el valor umbral (2,1) y, por lo tanto, el vóxel no es eliminado.

Tal como se ha descrito anteriormente, ajustando el valor umbral del cómputo verdadero en base al número de cámaras para las que el vóxel objetivo está incluido dentro del ángulo de visualización, es posible generar un modelo tridimensional para un primer plano que está lejos de un punto de mira y en el caso en que el número de cámaras para las que el vóxel objetivo está dentro del ángulo de visualización es pequeño. Por consiguiente, se posibilita generar un modelo tridimensional que suprime la magnitud del defecto incluso para un primer plano lejos de un punto de mira.

En la tercera y cuarta realizaciones descritas anteriormente, se explica el aspecto en el que solamente las cámaras para las que el vóxel está incluido dentro de la imagen de máscara del primer plano se cuentan como cómputo verdadero de cada vóxel. Sin embargo, en ese caso, un vóxel situado en la posición de un primer plano que está oculto por una estructura en muchas cámaras puede ser eliminado debido a que el cómputo verdadero no supere el valor umbral. Por consiguiente, como quinta realización se explica un aspecto en el que se genera un modelo tridimensional sin un defecto, incluso en un caso en el que se impide que un primer plano sea capturado por una estructura en muchas cámaras.

En la presente realización, incluso en un caso en el que un vóxel objetivo está fuera de una imagen de máscara del primer plano, en una situación en la que el vóxel objetivo está incluido dentro de una imagen de máscara de estructura, el vóxel tiene una probabilidad de ser un primer plano y, por lo tanto, se evita un defecto del primer plano

añadiendo un valor obtenido multiplicando el número de cámaras para las que se determina que el vóxel está incluido dentro de la imagen de máscara de estructura, por un valor de ponderación para cómputo verdadero.

Específicamente, en primer lugar, se ajusta un valor de ponderación en base al número de cámaras para las que el vóxel objetivo está incluido en la imagen de máscara de estructura. A continuación, en caso de que la suma del número de cámaras para las que el vóxel objetivo está incluido en la imagen de máscara del primer plano y el valor obtenido multiplicando el número de cámaras para las que el vóxel objetivo está incluido en la imagen de máscara de estructura por el valor de ponderación sea menor o igual que el valor umbral de cómputo verdadero, se determina eliminar el vóxel.

(Configuración funcional del dispositivo de generación de modelo tridimensional)

Haciendo referencia a la figura 30, se explica la configuración funcional del dispositivo de generación de modelo tridimensional, según la presente realización. El dispositivo de generación de modelo tridimensional 140, según la presente realización, incluye también una unidad de ajuste de ponderación 300, además de la configuración del dispositivo de generación de modelo tridimensional de la cuarta realización.

La unidad de ajuste de ponderación 300 ajusta un valor que se añade al cómputo verdadero en un caso en el que se determina que el vóxel objetivo está dentro de la imagen de máscara de estructura, como un valor de ponderación por cámara. Este valor de ponderación es equivalente a un valor que indica la probabilidad de que un vóxel esté situado en el primer plano y, en la presente realización, el valor de ponderación por cámara se ajusta a 0,5. A continuación, se añade al cómputo verdadero un valor obtenido multiplicando el número de cámaras para las que se determina que el vóxel objetivo está dentro de la imagen de máscara de estructura por 0,5, que es el valor de ponderación por cámara.

La figura 31 es un diagrama de flujo que muestra un procedimiento de procesamiento llevado a cabo por el dispositivo de generación de modelo tridimensional, según la presente realización.

Cada elemento de procesamiento en S3101 a S3104 es igual que cada elemento de procesamiento en S2701 a S2704 en el flujo de la figura 27 de la cuarta realización. Además, cada elemento de procesamiento en S3105 a S3108 es igual que cada elemento de procesamiento en S2706 a S2709 en el flujo de la figura 27 descrito anteriormente. Además, cada elemento de procesamiento en S3109 y S3110 es igual que cada elemento de procesamiento en S2705 y S2710 en el flujo de la figura 27 descrito anteriormente.

En S3111, la unidad de determinación interior/externo de máscara 105 cuenta el número de cámaras para las que el vóxel seleccionado está incluido dentro del área de máscara de la imagen de máscara de estructura de cada cámara.

En S3112, la unidad de ajuste de ponderación 300 añade un valor obtenido multiplicando el número de cámaras para las que la máscara seleccionada está incluida dentro del área de máscara de la imagen de máscara de estructura por 0,5, que es el valor de ponderación por cámara, al cómputo verdadero calculado en S3108. Cada elemento de procesamiento en S3113 a S3116 es igual que cada elemento de procesamiento en S2711 a S2714 en el flujo de la figura 27 descrito anteriormente. Lo anterior es la serie de procesos en el flujo de la figura 31.

Aquí, la figura 32 muestra un ejemplo de cómputo verdadero en un caso en el que no hay adición de ponderación, y un ejemplo de cómputo verdadero en un caso en el que hay adición de ponderación, según la presente realización, para un vóxel situado en una determinada área del primer plano.

Se supone que este vóxel está en el interior del ángulo de visualización de la totalidad de las 16 cámaras, el número de cámaras para las que el vóxel objetivo está incluido dentro de la imagen de máscara del primer plano es siete, y el número de cámaras para las que el vóxel objetivo está incluido dentro de la imagen de máscara de estructura es de nueve. En este caso, el número de cámaras para las que el vóxel está fuera de la imagen de máscara integrada es cero (número total de cámaras 16 - 7 - 9). Por consiguiente, el número de cámaras para las que el vóxel está fuera de la imagen de máscara integrada y para las que el vóxel está dentro del ángulo de visualización es cero y, por lo tanto, el cómputo falso es 0.

En un caso en el que no hay adición de ponderación, el número de cámaras para las que el vóxel objetivo está incluido dentro de la imagen de máscara del primer plano es de siete y, por lo tanto, el cómputo verdadero es de 7. Se supone que el valor umbral del cómputo verdadero es el 70 % del número de cámaras para las que el vóxel objetivo está incluido dentro del ángulo de visualización. Entonces, el valor umbral es de 11,2 (16 × 0,7), y el cómputo verdadero (7) < el valor umbral (11,2) y el cómputo verdadero es menor o igual que el valor umbral y, por lo tanto, el vóxel es eliminado.

Por otra parte, en un caso en el que hay adición de ponderación, el número de cámaras para las que el vóxel objetivo está incluido dentro de la imagen de máscara del primer plano es de siete y, por lo tanto, el cómputo verdadero es también de 7 y el valor de ponderación es añadido a este. El número de cámaras para las que el vóxel

objetivo está incluido dentro de la imagen de máscara de estructura es nueve y el valor de ponderación por cámara es de 0,5 y, por lo tanto, se añade  $9 \times 0,5 = 4,5$  como un valor de ponderación. El cómputo verdadero después de que se añada el valor de ponderación es de 11,5, y el cómputo verdadero  $(11,5) >$  el valor umbral  $(11,2)$  y el cómputo verdadero supera el valor umbral y, por lo tanto, el vóxel se considera como primer plano y no es eliminado.

5 En la presente realización, se supone un caso en el que hay una estructura, pero en un caso en el que hay una serie de estructuras diferentes que tienen una probabilidad de solapar con un primer plano, puede ser posible asimismo ajustar un valor de ponderación diferente para cada clase de imagen de máscara de estructura y añadir al cómputo verdadero un valor en base al valor de ponderación. Por ejemplo, para la imagen de máscara de estructura de un  
10 letrero electrónico instalado para rodear el campo deportivo del estadio deportivo, el letrero electrónico es grande y es probable que solape con un primer plano y, por lo tanto, la probabilidad de incluir un primer plano se hace alta y, por lo tanto, el valor de ponderación por cámara se ajusta a 0,5. Además, para la imagen de máscara de estructura de la portería, el valor de ponderación por cámara se ajusta a 0,3. Se considera que la probabilidad de que el letrero electrónico solape con un primer plano (persona) es mayor que la de la portería debido a que el letrero electrónico  
15 es mayor que la portería y no hay espacios vacíos en el letrero electrónico y, por lo tanto, el valor de ponderación para el letrero electrónico se ajusta a un valor mayor que el valor de ponderación para la portería.

Además, puede ser posible asimismo ajustar un valor de ponderación diferente según la posición del vóxel, la escena, el tamaño y la forma del área de máscara, el área del estadio deportivo objetivo de captura de imágenes, y  
20 similares.

Tal como se ha explicado anteriormente, en la presente realización, la determinación basada en umbral se lleva a cabo después de añadir al cómputo verdadero la ponderación en base al número de cámaras para las que el vóxel objetivo está incluido en el interior del área de máscara de la imagen de máscara de estructura. Debido a esto,  
25 incluso en un caso en que se impide que el primer plano sea capturado por una estructura en muchas cámaras, es posible implementar la generación de un modelo tridimensional sin defectos.

Igual que antes, según la primera a quinta realizaciones, incluso en un caso en que existe una estructura que oculta un objeto que es un primer plano, es posible generar un modelo tridimensional muy preciso de solamente el primer  
30 plano sin incluir la estructura.

[Sexta realización]

A continuación se explica un aspecto en una sexta realización en la que se utiliza el número de cámaras para las que el vóxel objetivo está incluido en la imagen de máscara de estructura, en lugar del número de cámaras para las que el vóxel objetivo está incluido en la imagen de máscara del primer plano (cómputo verdadero) utilizado en la  
35 tercera realización. En la tercera realización, para el modelo tridimensional generado en base a la imagen de máscara del primer plano y la imagen de máscara de estructura, la imagen de máscara del primer plano se actualiza siempre que se determine que el vóxel que configura el modelo tridimensional está incluido en la imagen de máscara del primer plano y, por lo tanto, existe un caso en el que el proceso se complica. Por consiguiente, la generación de un modelo tridimensional de un primer plano que no incluye una estructura se lleva a cabo contando el número de  
40 cámaras para las que el vóxel objetivo está incluido en la imagen de máscara de estructura fija para el modelo tridimensional generado en base a la imagen de máscara del primer plano y la imagen de máscara de estructura.

45 (Configuración funcional y configuración de hardware del dispositivo de generación de modelo tridimensional)

La figura 33 es un diagrama que muestra una configuración del dispositivo de generación de modelo tridimensional 140 en la presente realización. La configuración del dispositivo de generación de modelo tridimensional 140 en la presente realización es sustancialmente la misma que la de la tercera realización y, por lo tanto, se omite la explicación del bloque que lleva a cabo el mismo proceso. El dispositivo de generación de modelo tridimensional 140, según la presente realización, incluye una unidad de determinación interior/exterior de máscara 3300 en lugar de la unidad de determinación interior/exterior de máscara 105. La unidad de determinación interior/exterior de máscara 3300 cuenta el número de cámaras para las que cada vóxel dentro del espacio de vóxeles objetivo está incluido en la imagen de máscara integrada y en el área de la imagen de máscara de estructura, determina si  
50 eliminar o no el vóxel objetivo mediante determinación basada en umbrales, y entrega los resultados de la determinación a la unidad de generación de modelo de primer plano 107. La configuración de hardware del dispositivo de generación de modelo tridimensional 140 de la presente realización es igual que la de la figura 15 y, por lo tanto, se omite su explicación.

La figura 34 es un diagrama de flujo que muestra un procedimiento de procesamiento llevado a cabo por el dispositivo de generación de modelo tridimensional 140 en la presente realización. Cada elemento de procesamiento en S3401 a S3407 y S3410 a S3412 es igual que cada elemento de procesamiento en S1601 a S1607 y S1610 a S1612 explicado haciendo referencia la figura 16 en la tercera realización y, por lo tanto, se omite la explicación y se explican principalmente las partes necesarias.  
60

En S3406, la unidad de determinación interior/exterior de máscara 3300 determina si el cómputo falso es o no mayor  
65

o igual que un valor umbral. En caso de que el cómputo falso sea menor que el valor umbral, es posible determinar que el vóxel seleccionado es un primer plano o una estructura y, por lo tanto, el proceso avanza a S3408.

5 En S3408, la unidad de determinación interior/exterior de máscara 3300 cuenta el número de cámaras para las que el píxel o el área correspondiente al vóxel seleccionado está incluida dentro del área de máscara de la imagen de máscara de estructura de cada cámara (en adelante, denominado cómputo de estructura).

10 En S3409, la unidad de determinación interior/exterior de máscara 3300 determina si el cómputo de estructura es mayor o igual que un valor umbral. En caso de que el cómputo de estructura sea mayor o igual que el valor umbral, es posible determinar que el vóxel seleccionado es una estructura y, por lo tanto, el proceso avanza a S3407 y el vóxel seleccionado es eliminado del espacio de vóxeles objetivo. Por otra parte, en caso de que el cómputo de estructura sea menor que el valor umbral, es posible determinar que el vóxel seleccionado es un primer plano y, por lo tanto, el vóxel seleccionado no es eliminado del espacio de vóxeles objetivo.

15 En este caso, se explica un ejemplo de generación de un modelo tridimensional tomando como ejemplo el sistema de generación de imágenes de punto de vista virtual que captura el estadio deportivo mediante las 16 cámaras mostradas en la figura 14. La figura 35 muestra un ejemplo de cómputo falso/cómputo de estructura y resultados de determinación de vóxeles para la persona, la pierna de la persona, la cabeza de la persona, la portería, y la otra área, respectivamente, como el primer plano, el primer plano no detectado por una parte de las cámaras, el primer plano oculto por la estructura, la estructura y el no primer plano en el sistema de generación de imágenes de punto de vista virtual mostrado en la figura 14. En este caso, se supone que una cámara ha fallado en la extracción del primer plano de la pierna de la persona, y la cabeza de la persona está oculta por la portería, que es una estructura, en tres cámaras, y estos no son extraídos como primer plano por el dispositivo de separación del primer plano 130.

25 En un caso en el que el valor umbral del cómputo falso es un valor fijo de 10 en la determinación en S3404, el cómputo falso del vóxel situado en la otra área excepto para la persona, la pierna, la cabeza y el poste de la portería, que es una estructura, es 16 y supera el valor umbral y, por lo tanto, el vóxel es eliminado. El modelo tridimensional generado aplicando determinación basada en umbrales de cómputo falso es igual que el mostrado en la figura 23 descrita anteriormente.

30 Además, en un caso en que el valor umbral del cómputo de estructura es un valor fijo de 3 en la determinación mostrada en S3408, el cómputo de estructura del vóxel situado en el área de la portería, que es una estructura, es 5 y mayor o igual que el valor umbral y, por lo tanto, el vóxel es eliminado. Por otra parte, el cómputo de estructura de cada uno de los vóxeles situados en las áreas de la persona, la pierna de la persona y la cabeza es 0 y menor que el valor umbral y, por lo tanto, el vóxel no es eliminado. Por consiguiente, se genera el modelo tridimensional de la persona sin defectos, tal como se muestra en la figura 24 descrita anteriormente.

35 Mediante el proceso anterior es posible implementar la generación de un modelo tridimensional sin defectos, incluso en un caso en que una estructura impide que se capture un primer plano, mediante determinación basada en umbrales del número de cámaras para las que dicho vóxel objetivo está incluido dentro de la máscara de estructura (cómputo de estructura).

(Otras realizaciones)

45 Es posible asimismo implementar la presente invención mediante procesamiento para suministrar un programa que implemente una o varias funciones de las realizaciones descritas anteriormente, a un sistema o un aparato por medio de una red o un medio de almacenamiento, y hacer que uno o varios procesadores en un ordenador del sistema o del aparato lean y ejecuten el programa. Además, es posible asimismo implementar la presente invención mediante un circuito (por ejemplo, ASIC) que implemente una o varias funciones.

50 La presente invención se ha explicado hasta aquí haciendo referencia a las realizaciones, pero no hace falta decir que la presente invención no se limita a las realizaciones descritas anteriormente. Por consiguiente, para hacer público el alcance de la presente invención se adjuntan las reivindicaciones siguientes.

55 Esta solicitud reivindica el beneficio de la solicitud de patente japonesa núm. 2017-239891, presentada el 14 de diciembre de 2017, la solicitud de patente japonesa núm. 2018-089467, presentada el 7 de mayo de 2018, y la solicitud de patente japonesa núm. 2018-209196, presentada el 6 de noviembre de 2018.

**REIVINDICACIONES**

1. Dispositivo de generación para generar datos de forma tridimensional, que comprende:

5 primeros medios de adquisición para adquirir una serie de imágenes de objeto que indican un área de un objeto dentro de una serie de imágenes obtenidas mediante captura de imágenes desde una serie de direcciones de captura de imágenes;

10 segundos medios de adquisición para adquirir una serie de imágenes de una estructura estacionaria existente dentro de una escena de captura de imágenes que indica un área de la estructura que tiene una probabilidad de impedir que el objeto sea capturado en caso de captura de imágenes desde, por lo menos, una dirección de captura de imágenes de la serie de direcciones de captura de imágenes; y

15 medios de generación para generar datos de forma tridimensional correspondientes al objeto en la escena de captura y a la estructura, en base tanto a la serie de imágenes del objeto que indican el área del objeto adquiridas por los primeros medios de adquisición como a la serie de imágenes de la estructura que indican el área de la estructura adquiridas por los segundos medios de adquisición, y para eliminar datos de forma tridimensional correspondientes a la estructura respecto de los datos de forma tridimensional generados, correspondientes al objeto y a la estructura.

20 2. Dispositivo de generación, según la reivindicación 1, que comprende, además:

medios de combinación para combinar la imagen que indica el área del objeto y la imagen que indica el área de la estructura, donde

25 los medios de generación generan los datos de forma tridimensional correspondientes al objeto y a la estructura, en base a la imagen combinada por los medios de combinación.

30 3. Dispositivo de generación, según la reivindicación 2, en el que

los medios de combinación generan una imagen que indica tanto el área del objeto como el área de la estructura, en base a la imagen que indica el área del objeto y a la imagen que indica el área de la estructura.

35 4. Dispositivo de generación, según la reivindicación 1, en el que

los medios de generación:

40 generan los datos de forma tridimensional correspondientes a la estructura en base a la serie de imágenes de la estructura adquiridas de los segundos medios de adquisición; y

generan los datos de forma tridimensional correspondientes al objeto no incluyendo los datos de forma tridimensional correspondientes a la estructura en base a los datos de forma tridimensional generados correspondientes a la estructura y a los datos de forma tridimensional correspondientes al objeto y a la estructura.

45 5. Dispositivo de generación, según la reivindicación 4, en el que

50 los medios de generación generan los datos de forma tridimensional correspondientes al objeto no incluyendo los datos de forma tridimensional correspondientes a la estructura, en base a los datos de forma tridimensional correspondientes al objeto y a la estructura y a los datos de forma tridimensional correspondientes a la estructura, donde, por lo menos, una parte de la forma tridimensional de la estructura se expande.

6. Dispositivo de generación, según la reivindicación 5, en el que

55 los medios de generación determinan una parte de la estructura, que se expande, para los datos de forma tridimensional correspondientes a la estructura, de acuerdo con un área en un espacio tridimensional en el que existe la estructura.

7. Dispositivo de generación, según la reivindicación 5 o 6, en el que

60 los medios de generación determinan una relación de expansión para los datos de forma tridimensional correspondientes a la estructura, de acuerdo con la distancia entre la estructura y el objeto en un espacio tridimensional en el que existe la estructura.

65 8. Dispositivo de generación, según la reivindicación 7, en el que

los medios de generación aumentan la relación de expansión para los datos de forma tridimensional

correspondientes a la estructura a medida que aumenta la distancia entre la estructura y el objeto.

9. Dispositivo de generación, según cualquiera de las reivindicaciones 1 a 8, en el que

5 los medios de generación determinan un elemento correspondiente a la estructura de entre los elementos que forman los datos de forma tridimensional generados correspondientes al objeto y a la estructura, y

los medios de generación eliminan el elemento determinado correspondiente a la estructura, de los datos de forma tridimensional generados correspondientes al objeto y a la estructura.

10 10. Dispositivo de generación, según cualquiera de las reivindicaciones 1 a 8, en el que  
los medios de generación determinan un elemento correspondiente a la estructura de entre los elementos que forman los datos de forma tridimensional generados correspondientes al objeto y a la estructura, en base al número de imágenes de la estructura que incluyen una zona correspondiente a un elemento que forma los datos de forma tridimensional generados correspondientes al objeto y a la estructura, y

15 los medios de generación eliminan el elemento determinado correspondiente a la estructura, de los datos de forma tridimensional generados correspondientes al objeto y a la estructura.

20 11. Dispositivo de generación, según cualquiera de las reivindicaciones 1 a 10, en el que

el objeto es un objeto en movimiento cuya posición puede cambiar dentro de cada imagen en un caso en que la captura de imágenes se lleva a cabo en una serie temporal a partir de una dirección idéntica de captura de imágenes.

25 12. Dispositivo de generación, según cualquiera de las reivindicaciones 1 a 11, en el que

el objeto es, por lo menos, uno de una persona y un balón.

30 13. Dispositivo de generación, según cualquiera de las reivindicaciones 1 a 12, en el que

la estructura es un objeto cuyo estado estacionario permanece.

35 14. Dispositivo de generación, según cualquiera de las reivindicaciones 1 a 12, en el que

la estructura es, por lo menos, uno de una portería de fútbol y un banderín de córner utilizados en un partido de fútbol.

40 15. Dispositivo de generación, según cualquiera de las reivindicaciones 1 a 14, en el que

la estructura es un objeto instalado en una posición predeterminada.

45 16. Dispositivo de generación, según cualquiera de las reivindicaciones 1 a 15, en el que

por lo menos una parte de la estructura está instalada en un campo en el que una persona, que es un objeto, juega un partido.

50 17. Dispositivo de generación, según cualquiera de las reivindicaciones 1 a 16, en el que

la estructura es un objeto especificado.

18. Procedimiento implementado por ordenador para generar datos de forma tridimensional, que comprende:

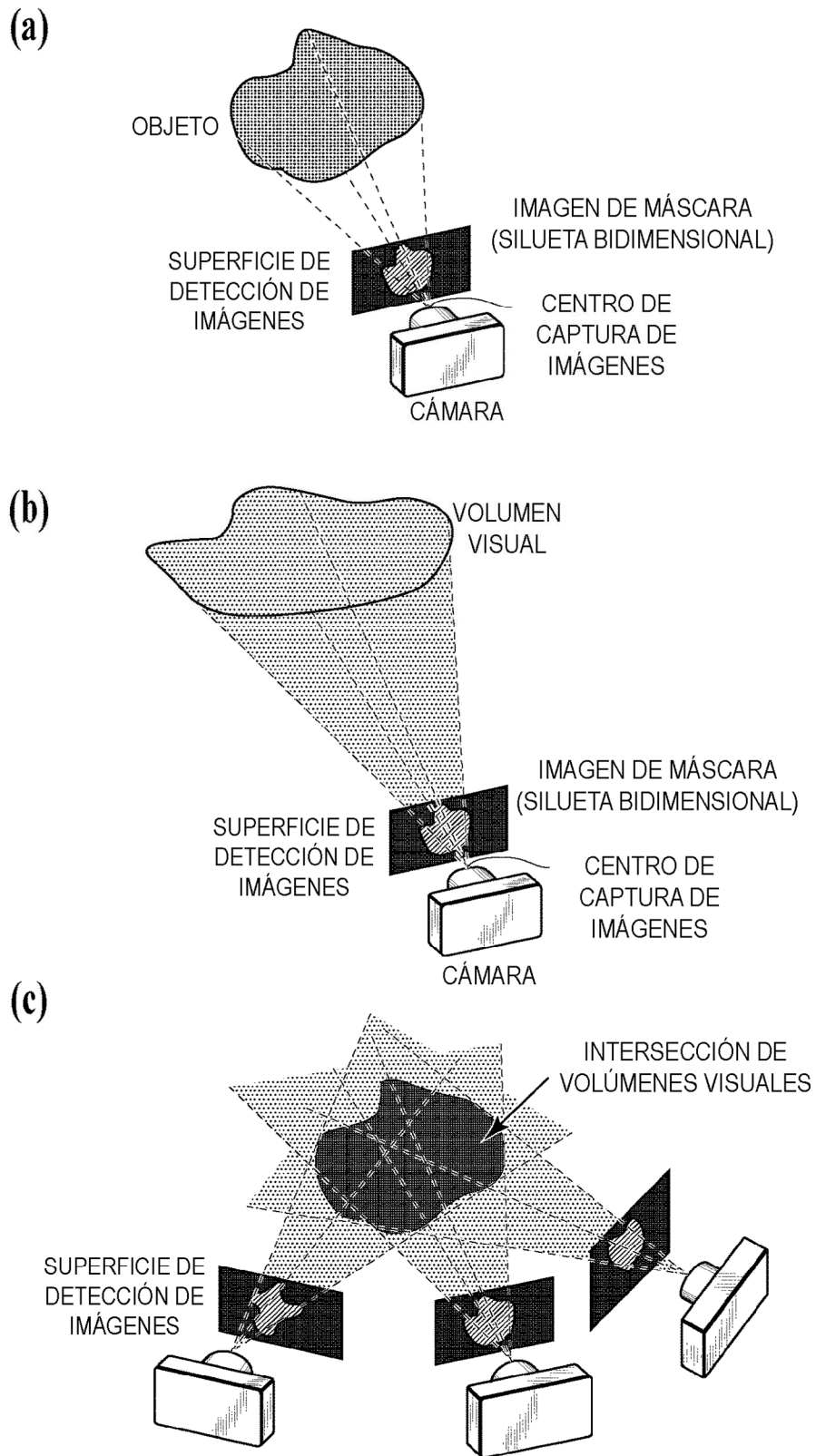
55 una primera etapa de adquisición, para adquirir una serie de imágenes de objeto que indican un área de objeto dentro de una serie de imágenes obtenidas por captura de imágenes desde una serie de direcciones de captura de imágenes;

60 una segunda etapa de adquisición, para adquirir una serie de imágenes de una estructura estacionaria existente dentro de una escena de captura de imágenes que indican un área de la estructura que tiene una probabilidad de impedir que el objeto sea capturado en caso de captura de imágenes desde, por lo menos, una dirección de captura de imágenes de la serie de direcciones de captura de imágenes; y

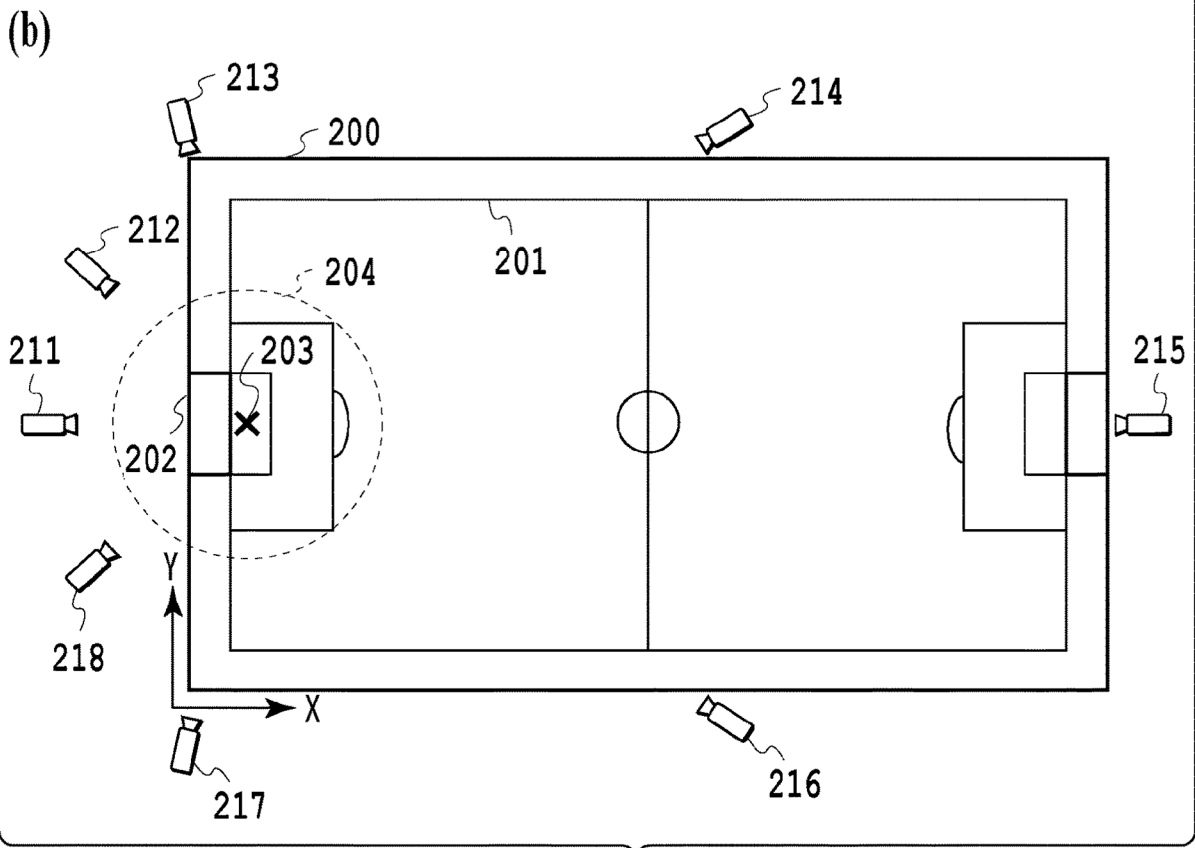
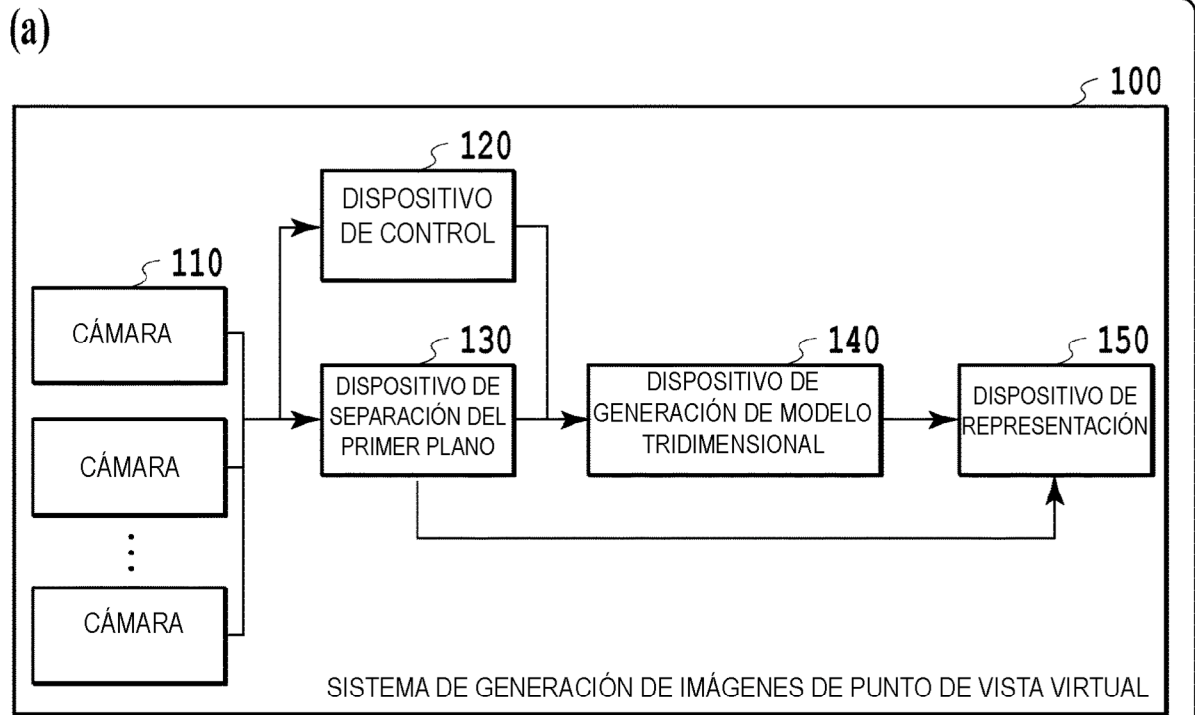
65 una etapa de generación, para generar datos de forma tridimensional correspondientes al objeto en la escena de captura y a la estructura, en base tanto a la serie de imágenes del objeto que indican el área del objeto adquiridas en la primera etapa de adquisición como a la serie de imágenes de la estructura que indican el área de la estructura

adquiridas en la segunda etapa de adquisición, y para eliminar datos de forma tridimensional correspondientes a la estructura respecto de los datos de forma tridimensional generados correspondientes al objeto y a la estructura.

- 5 19. Procedimiento de generación, según la reivindicación 18, que comprende, además:  
una etapa de combinación, para combinar la imagen que indica el área del objeto y la imagen que indica el área de la estructura, donde
- 10 en la etapa de generación, los datos de forma tridimensional correspondientes al objeto y a la estructura se generan en base a la imagen combinada en la etapa de combinación.
20. Programa informático para hacer que un ordenador funcione como dispositivo de generación, según cualquiera de las reivindicaciones 1 a 17.



**FIG. 1**



**FIG. 2**

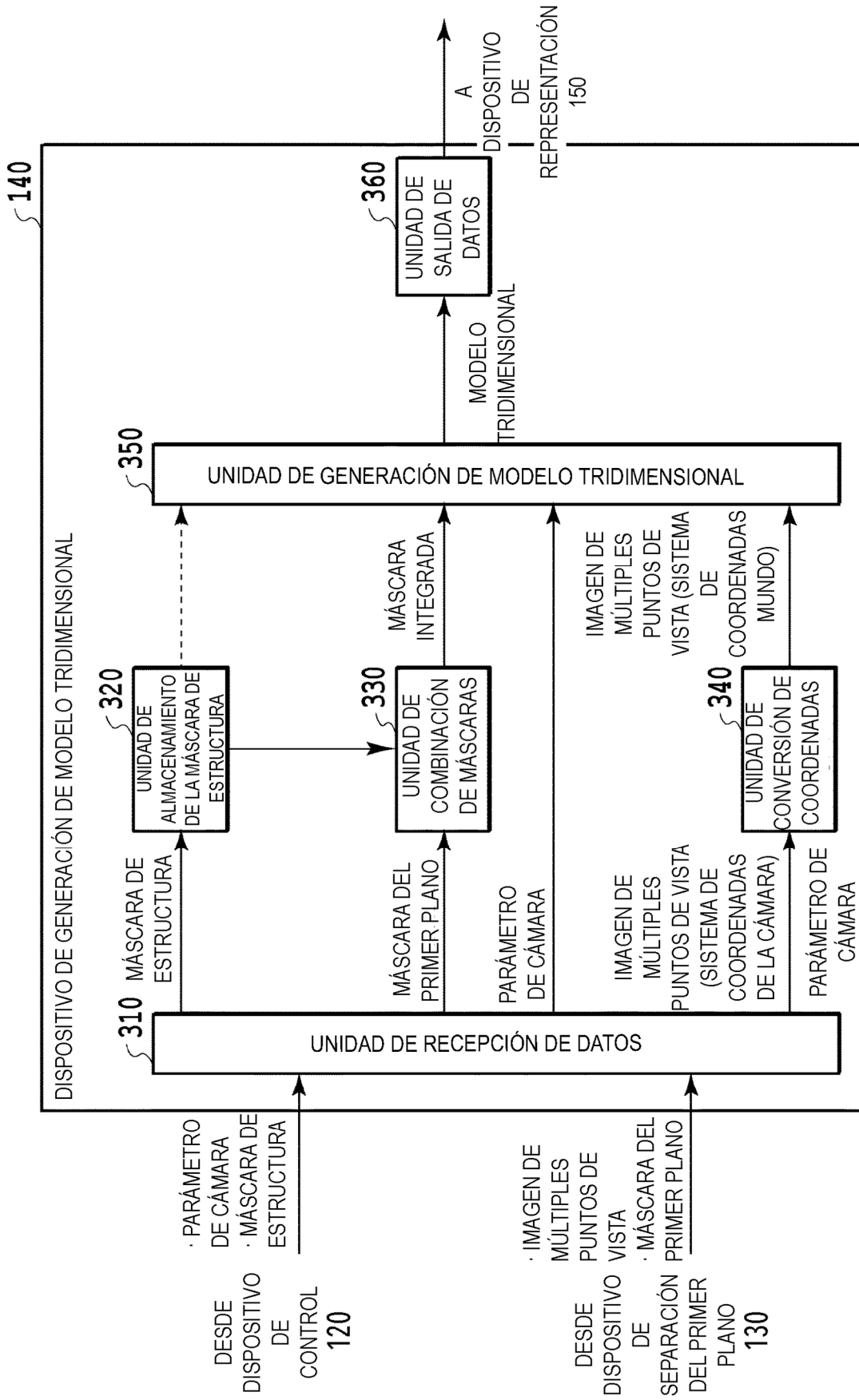
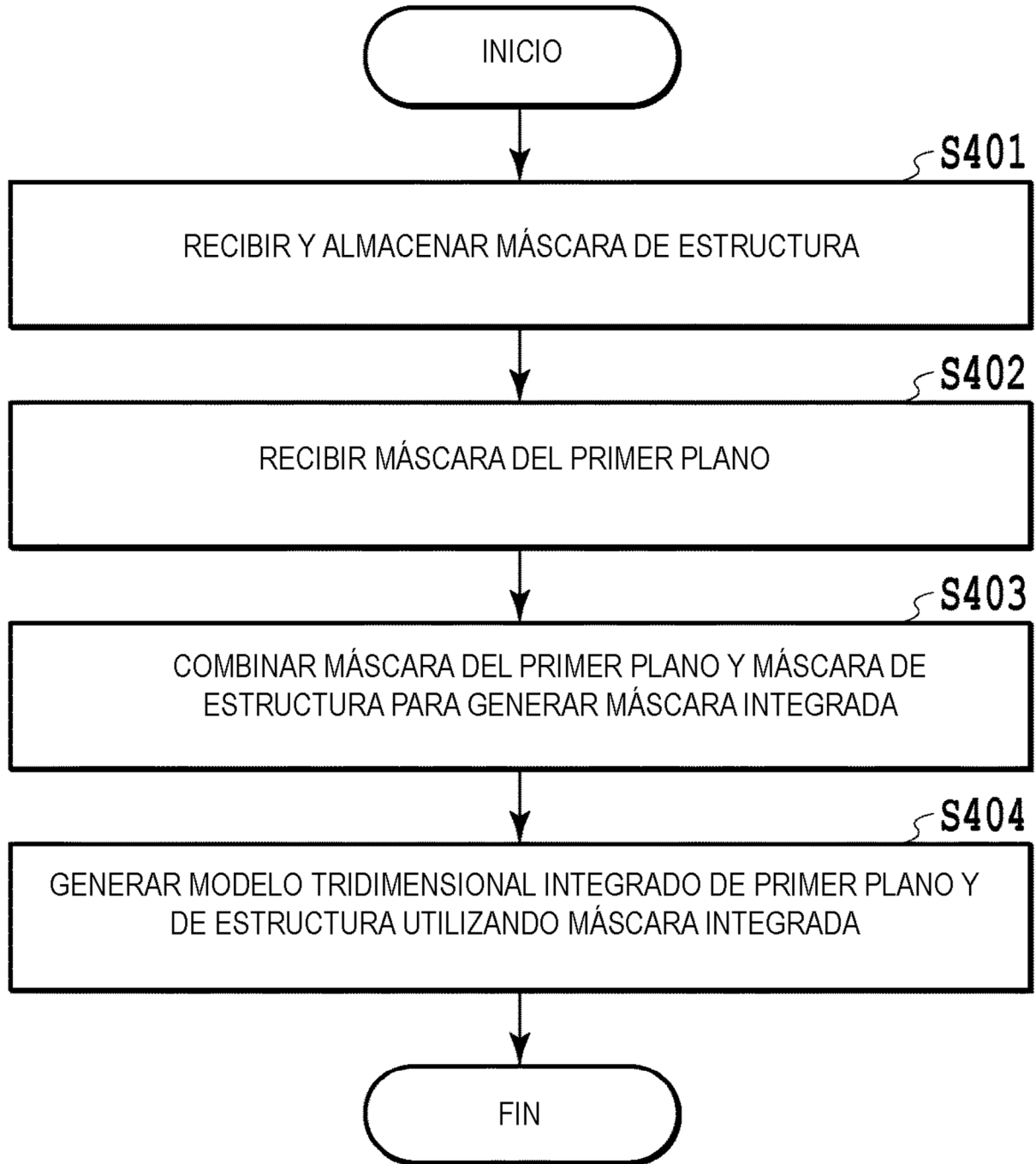
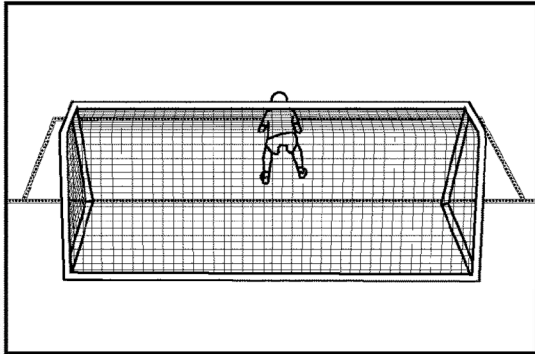


FIG.3

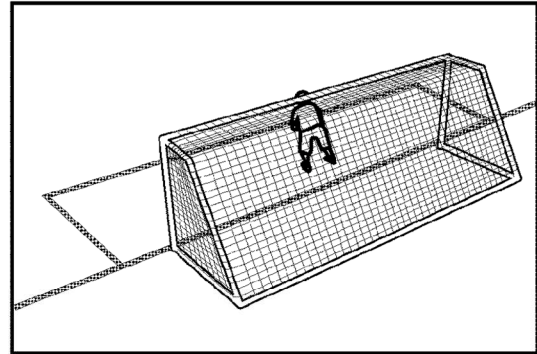


**FIG.4**

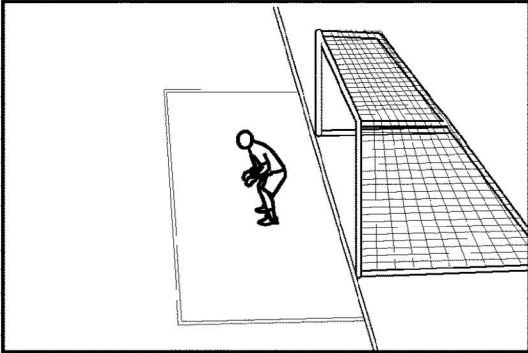
(a) IMAGEN CAPTURADA POR LA CÁMARA 211



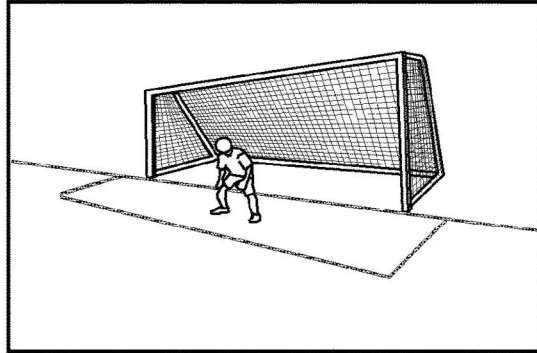
(b) IMAGEN CAPTURADA POR LA CÁMARA 212



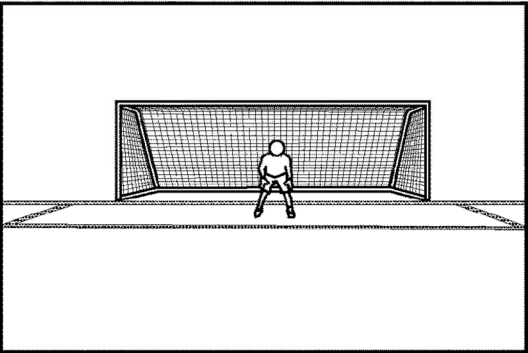
(c) IMAGEN CAPTURADA POR LA CÁMARA 213



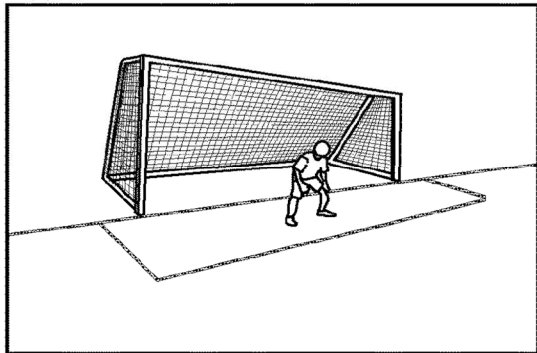
(d) IMAGEN CAPTURADA POR LA CÁMARA 214



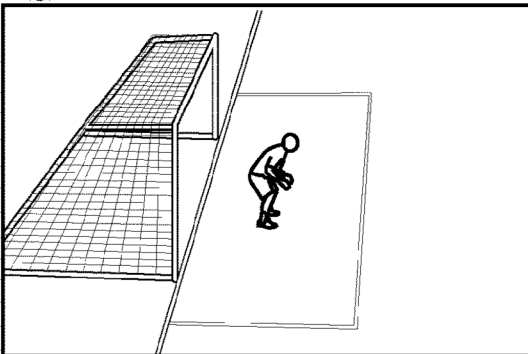
(e) IMAGEN CAPTURADA POR LA CÁMARA 215



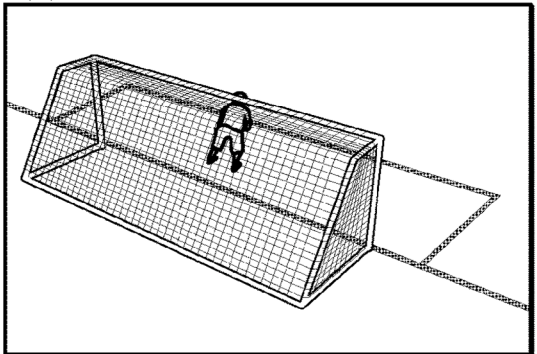
(f) IMAGEN CAPTURADA POR LA CÁMARA 216



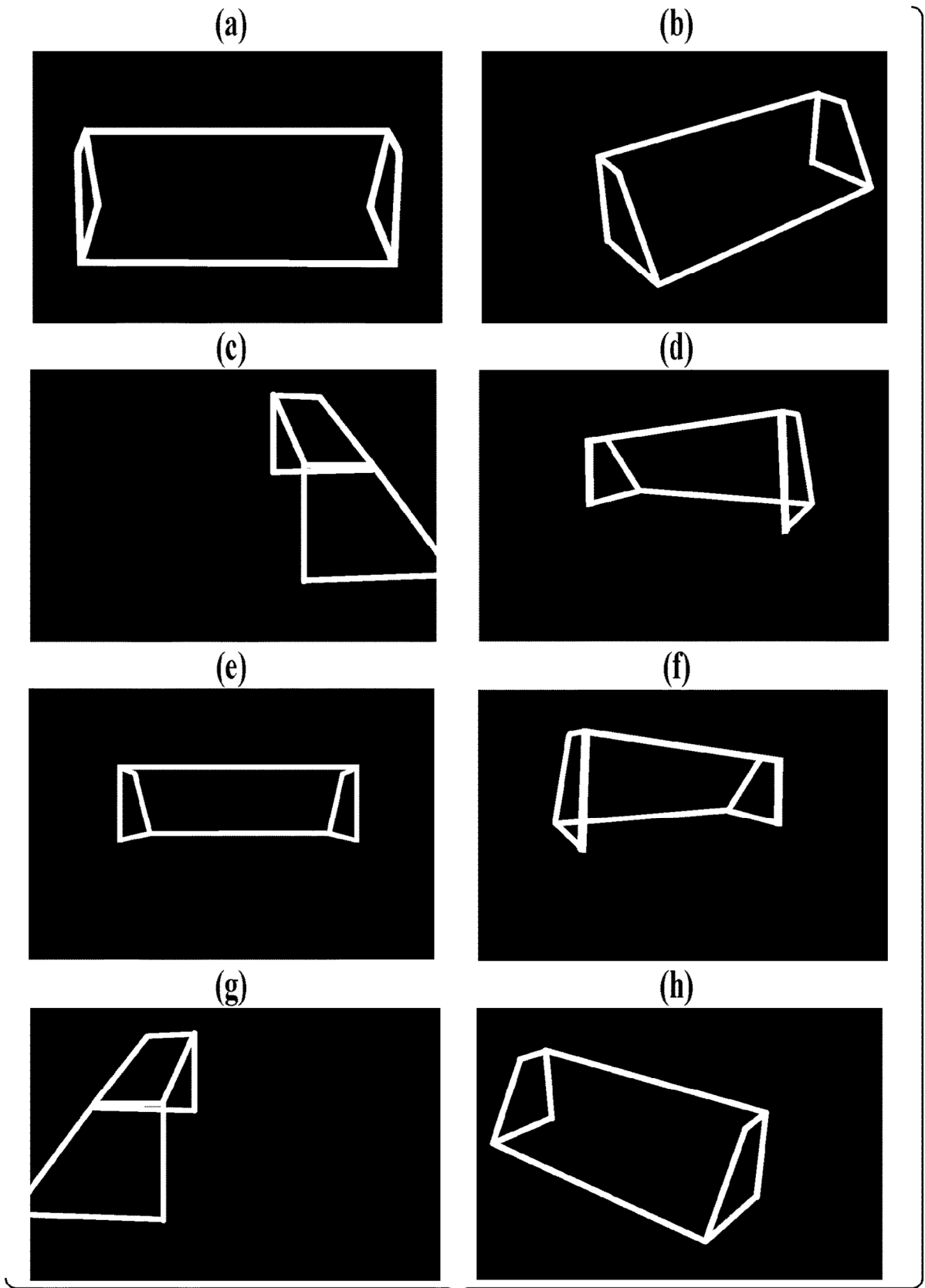
(g) IMAGEN CAPTURADA POR LA CÁMARA 217



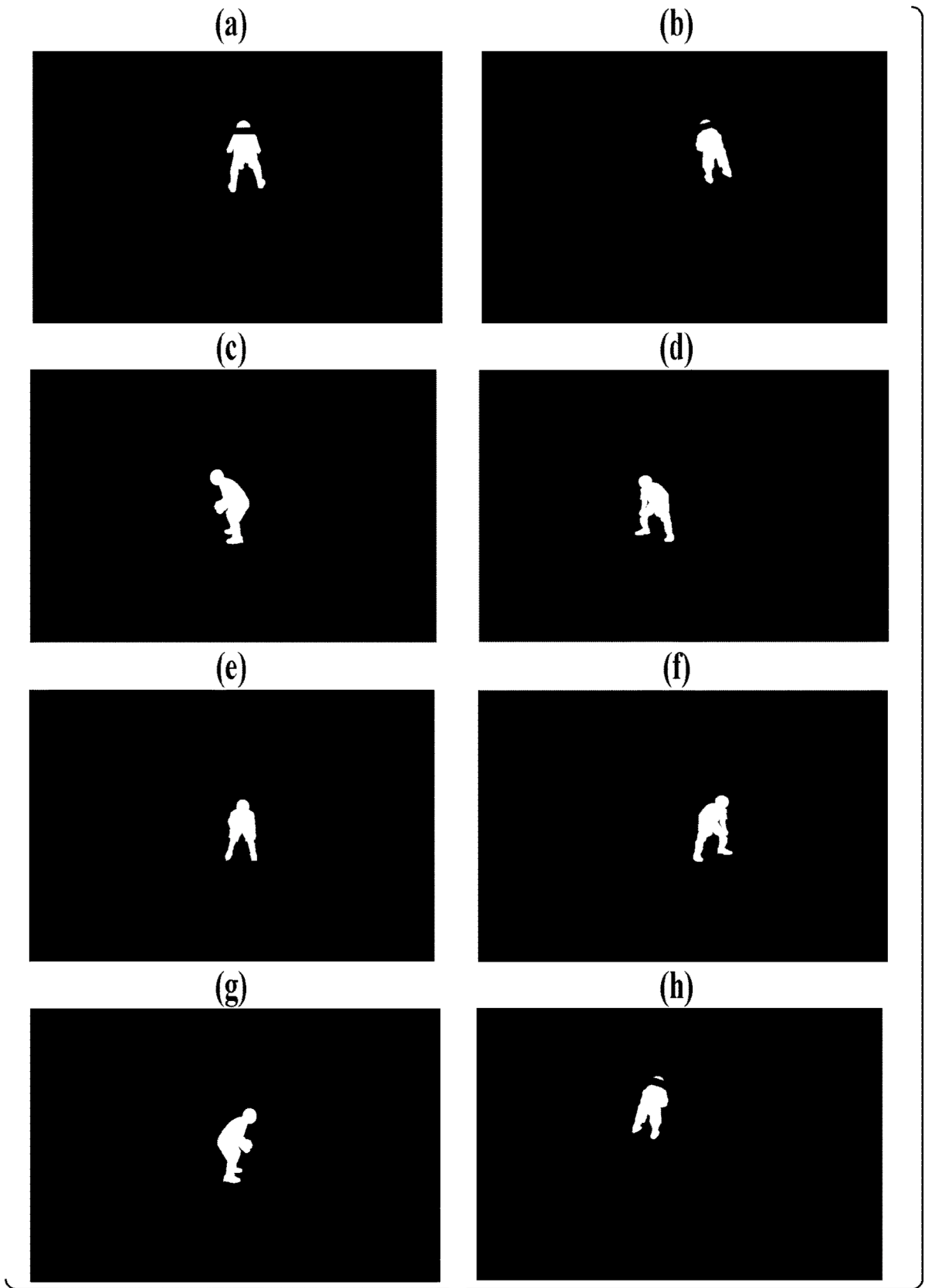
(h) IMAGEN CAPTURADA POR LA CÁMARA 218



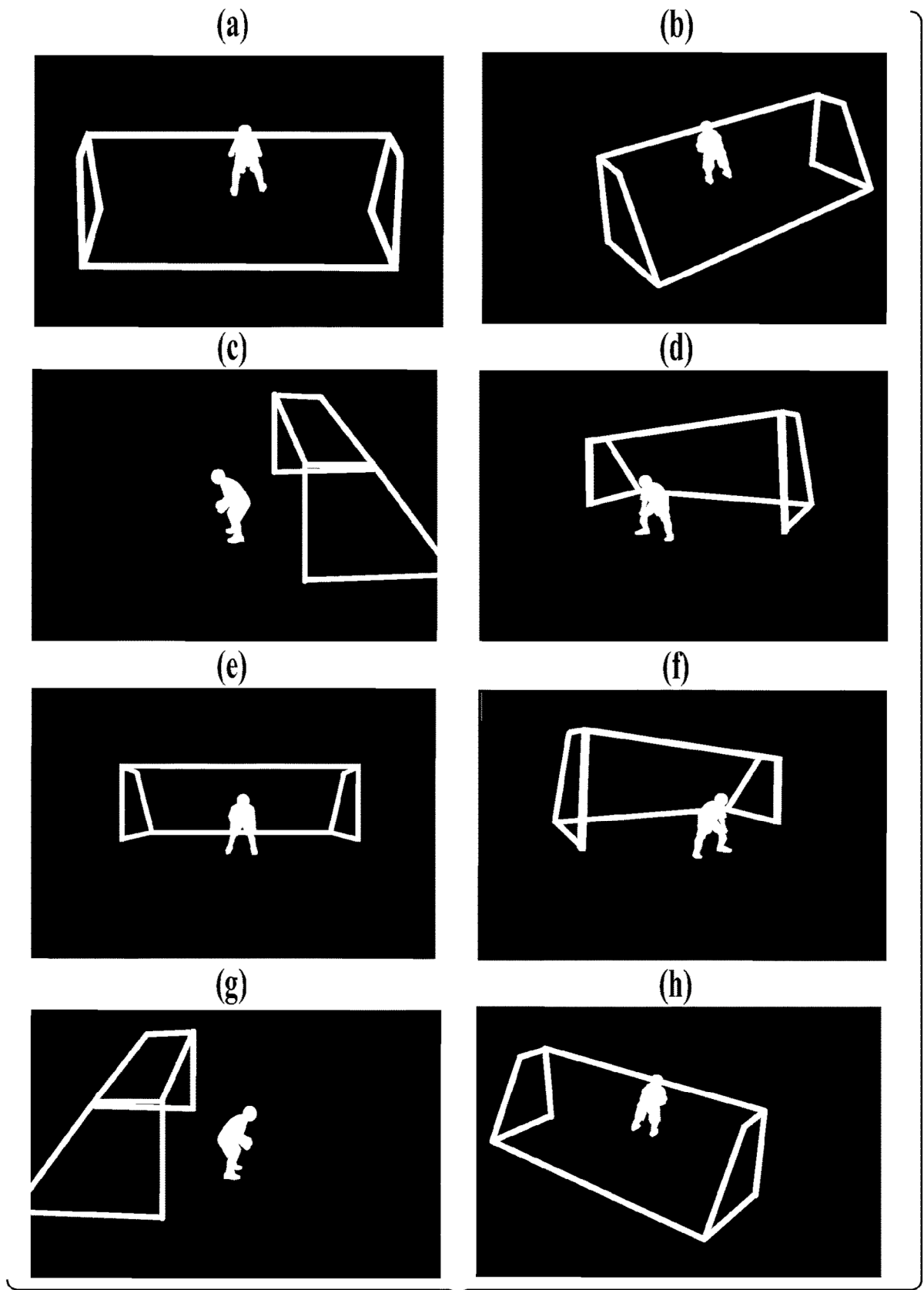
**FIG.5**



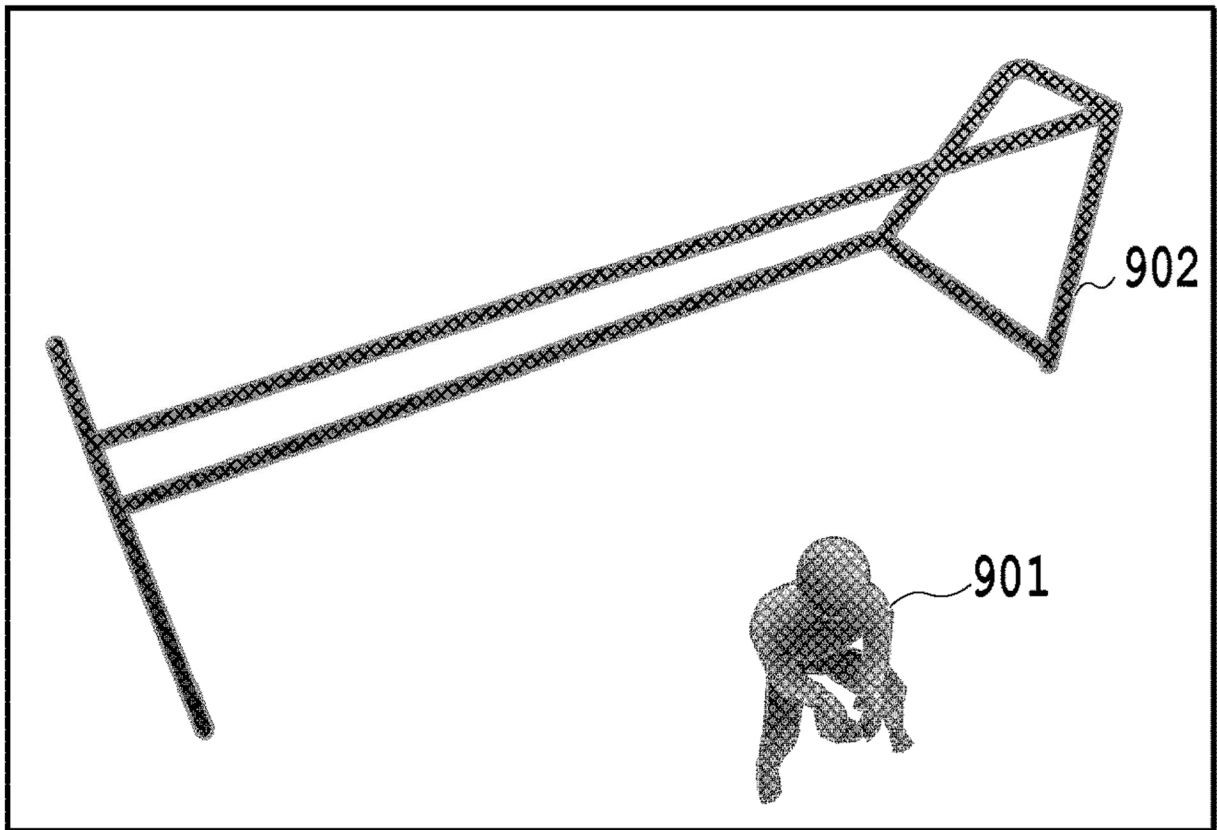
**FIG.6**



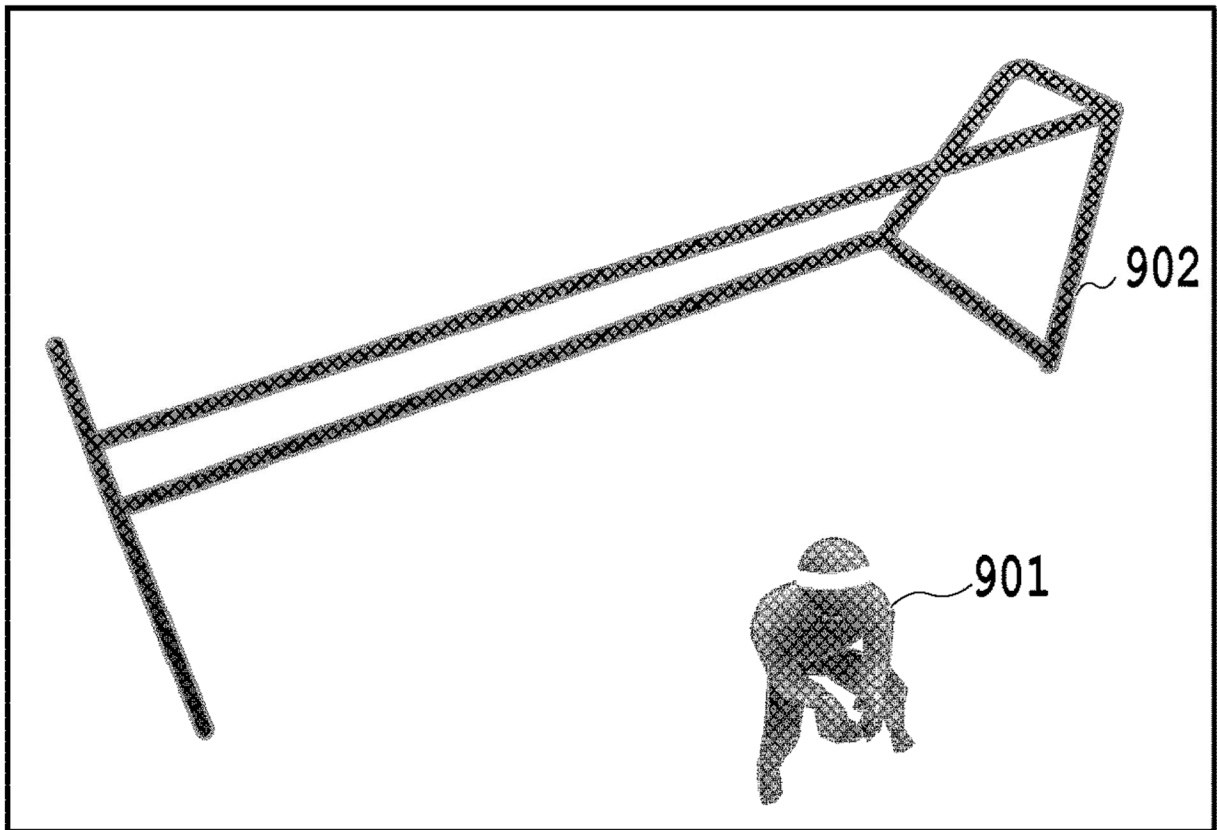
**FIG. 7**



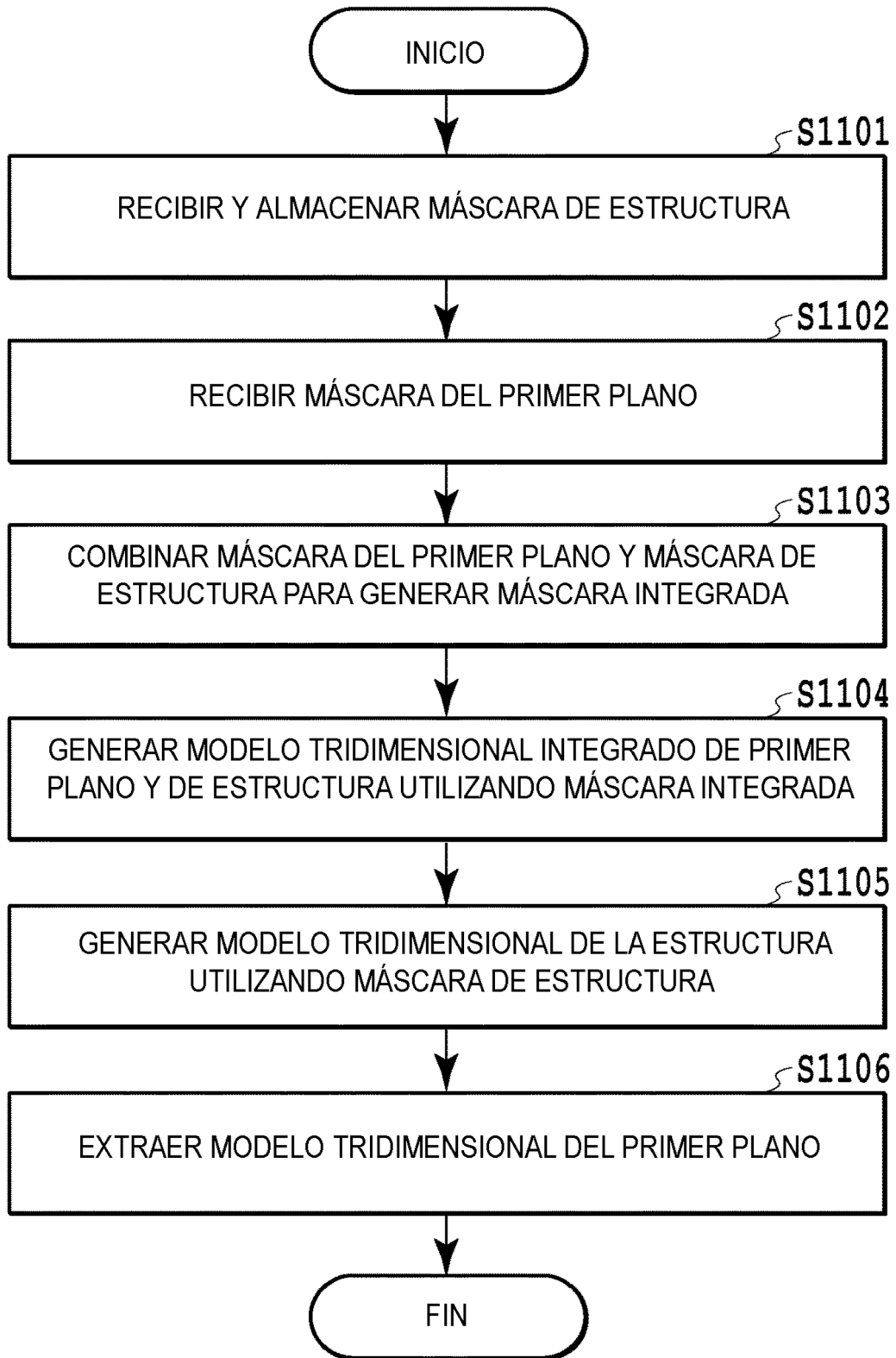
**FIG. 8**



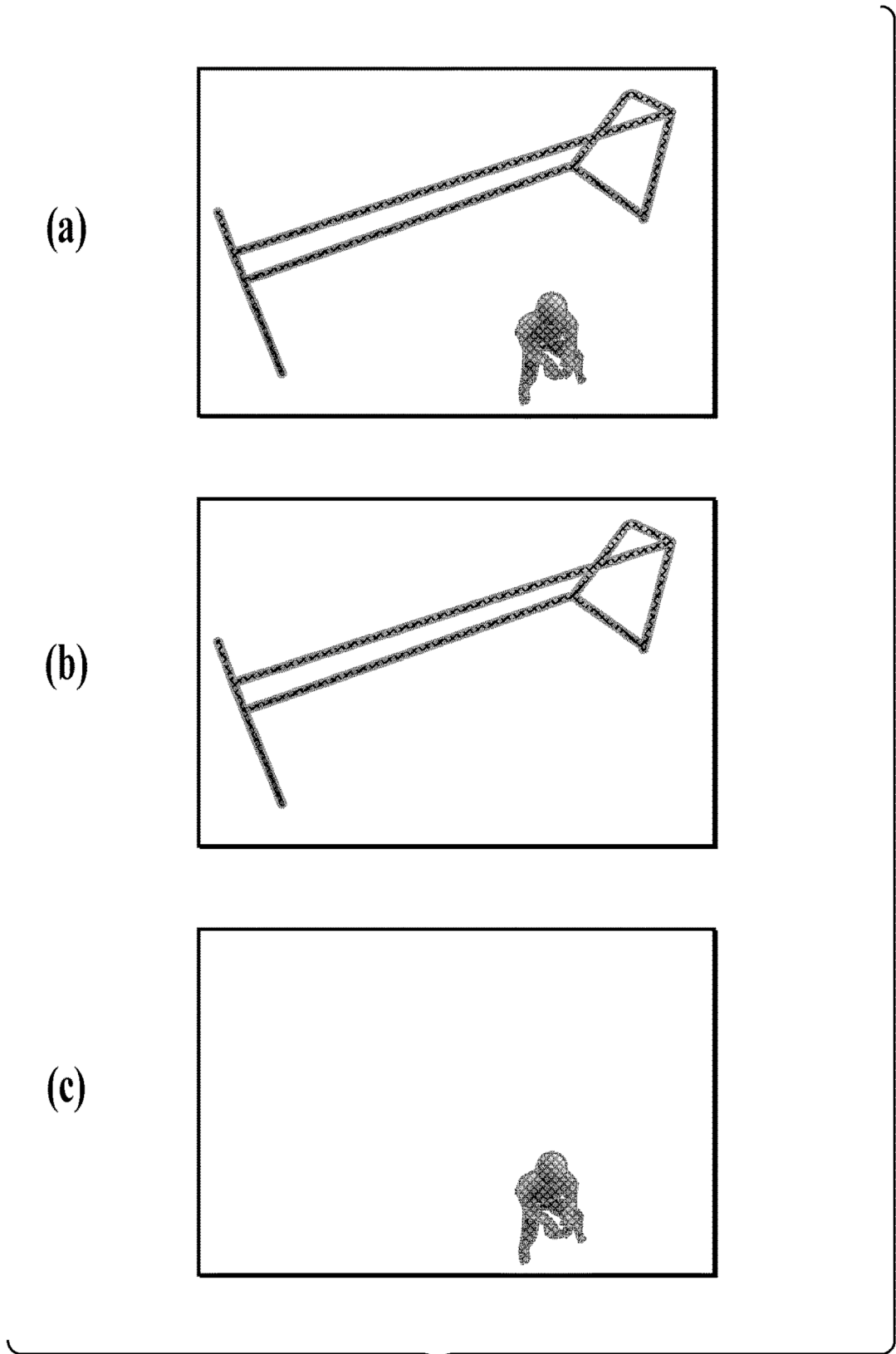
**FIG. 9**



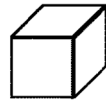
**FIG.10**



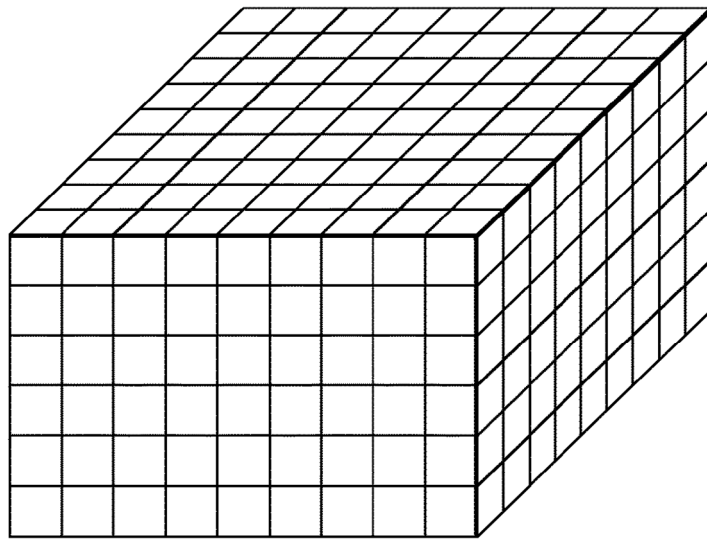
**FIG. 11**



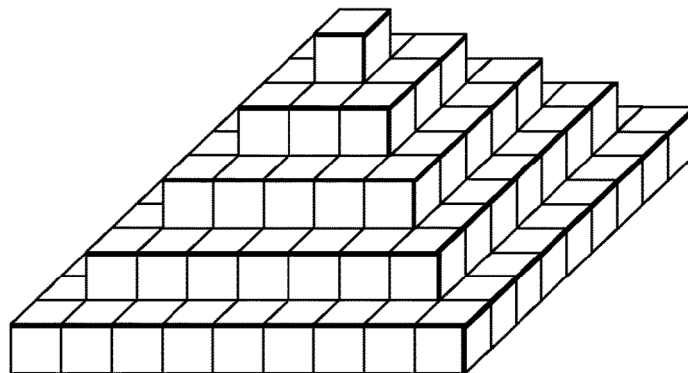
**FIG.12**



(a)

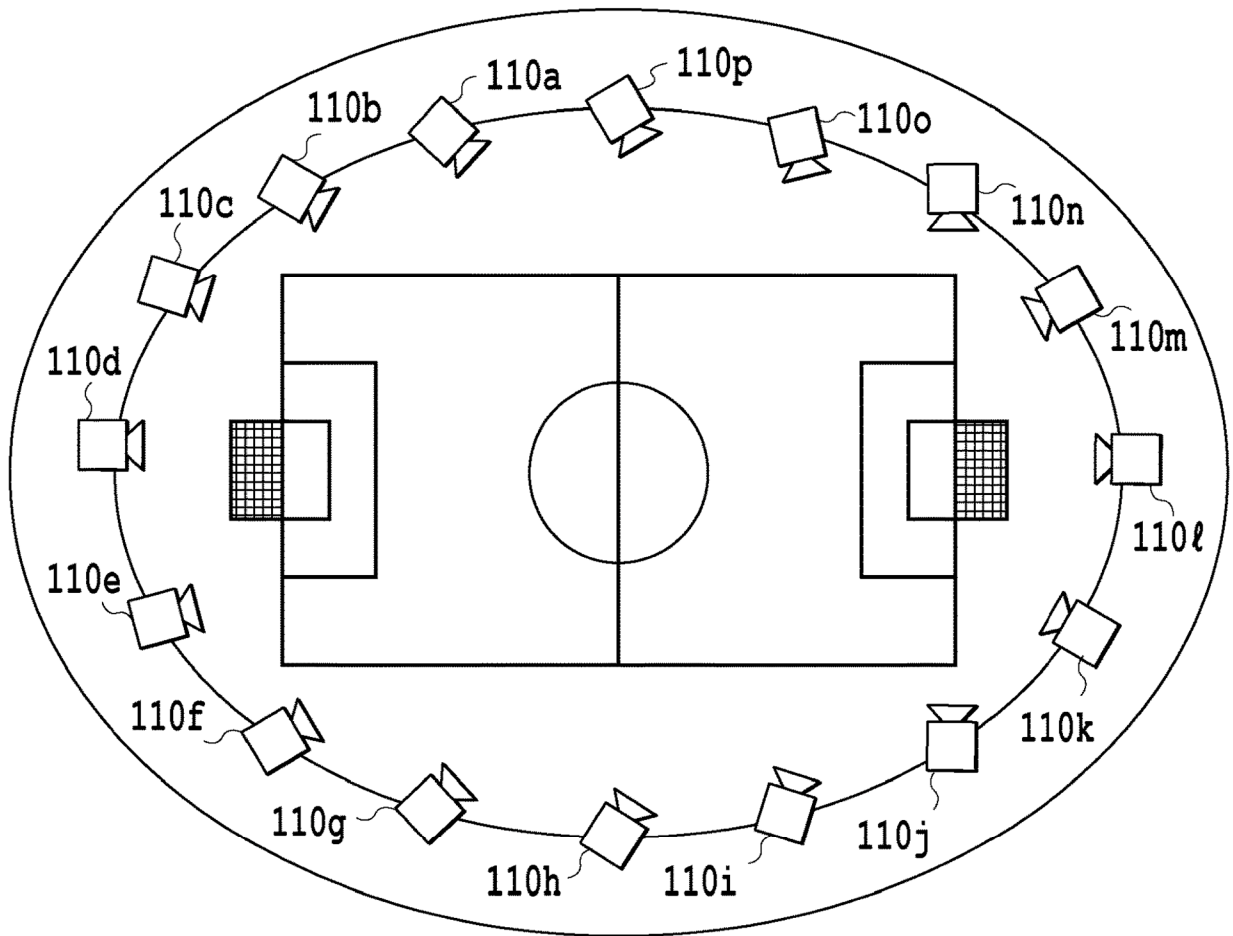


(b)

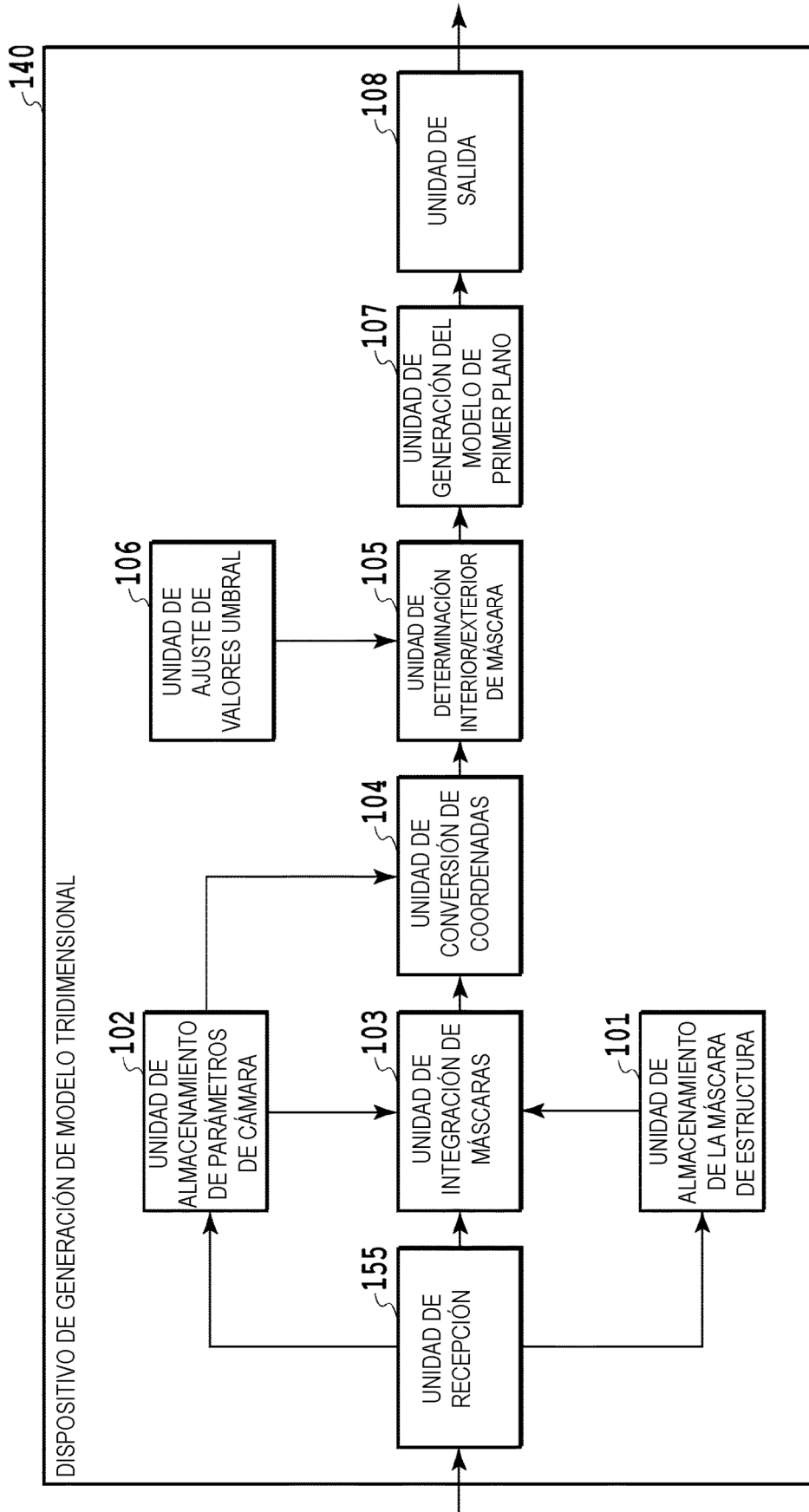


(c)

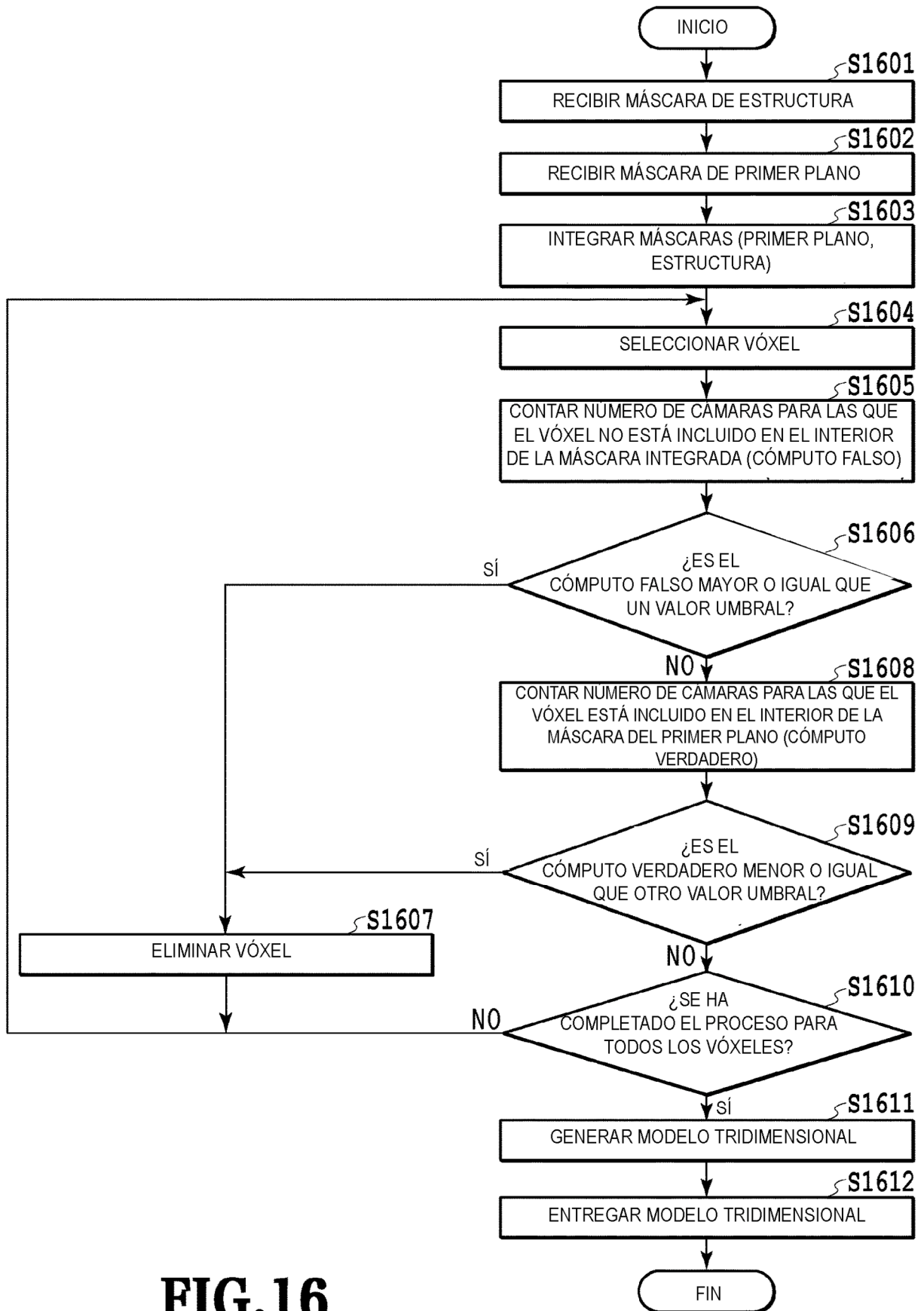
**FIG.13**



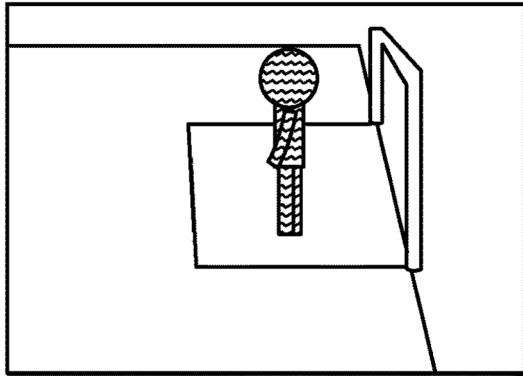
**FIG.14**



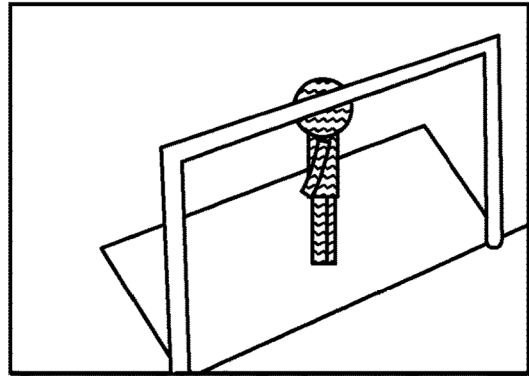
**FIG.15**



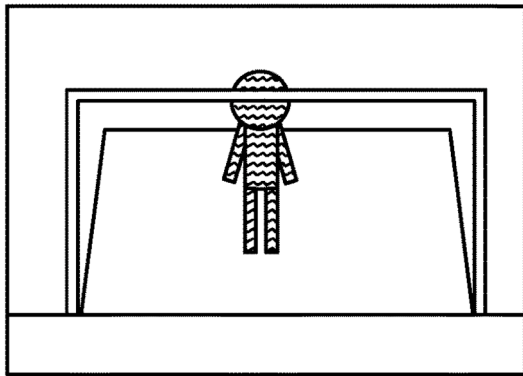
**FIG.16**



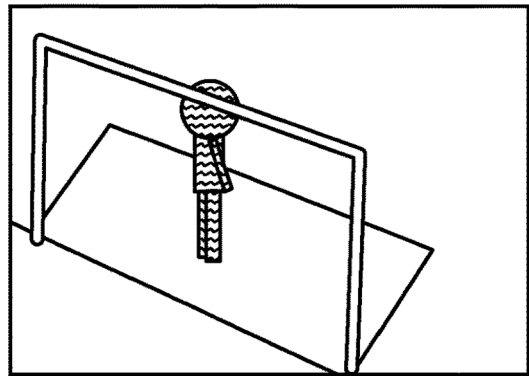
(a)



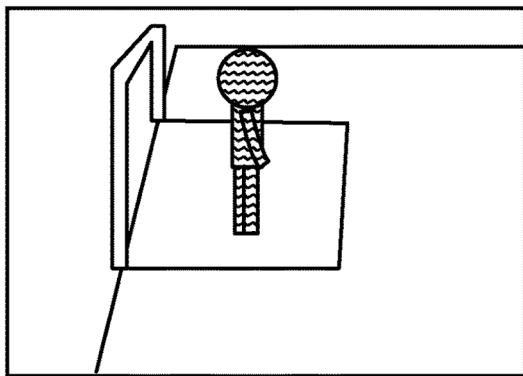
(b)



(c)



(d)

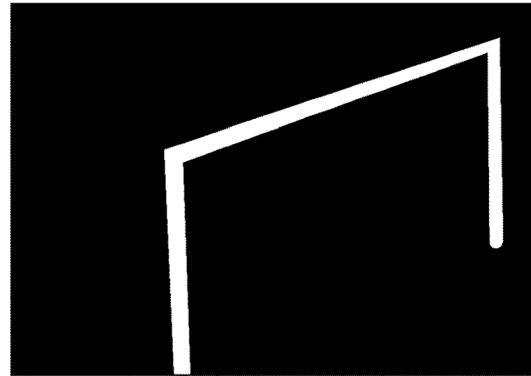


(e)

**FIG.17**



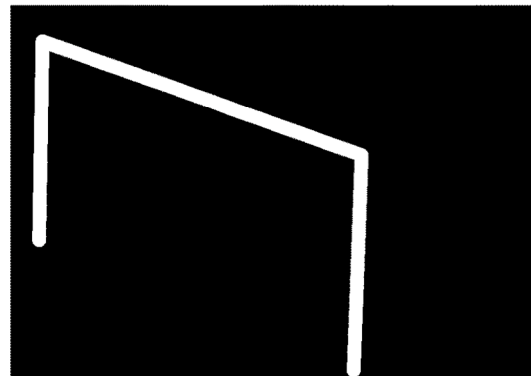
(a)



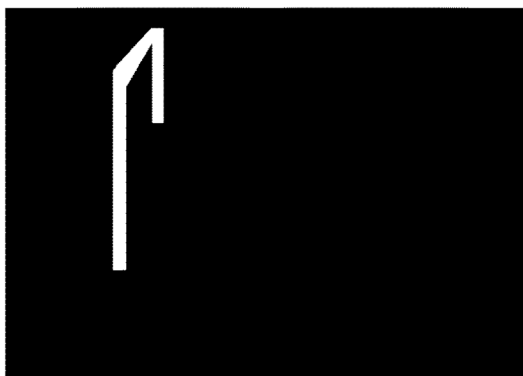
(b)



(c)

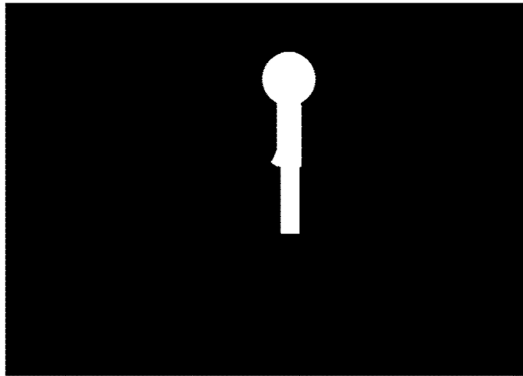


(d)

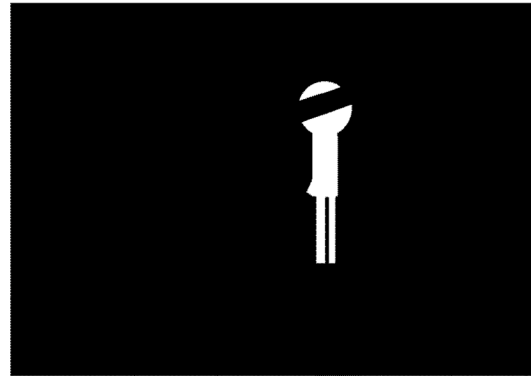


(e)

**FIG.18**



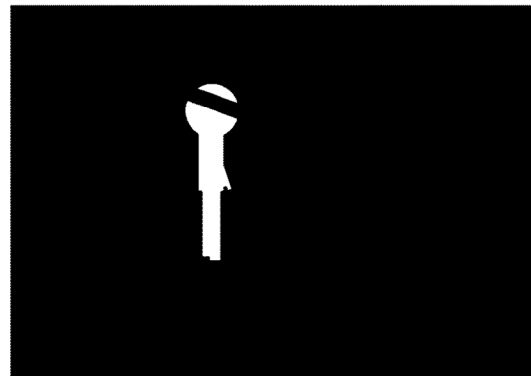
(a)



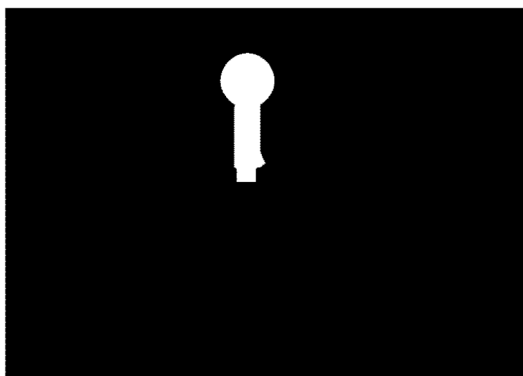
(b)



(c)

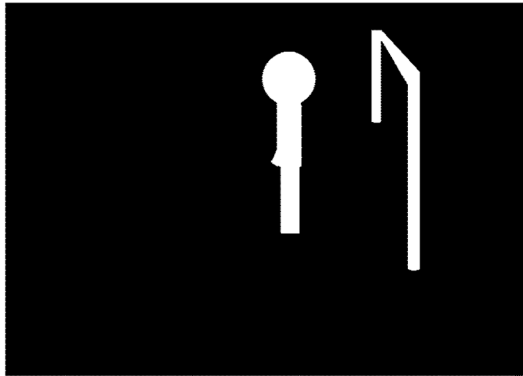


(d)

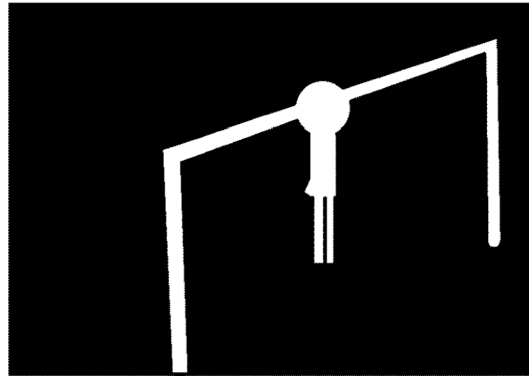


(e)

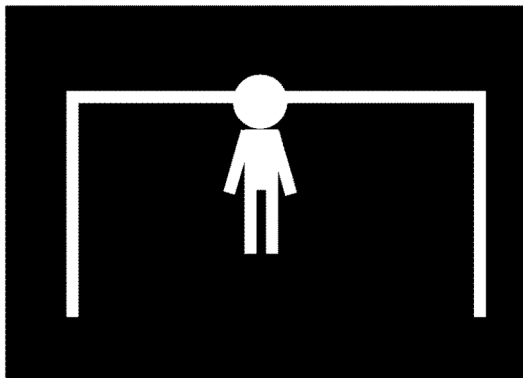
**FIG.19**



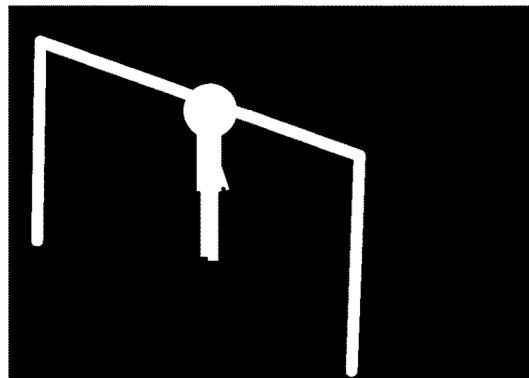
(a)



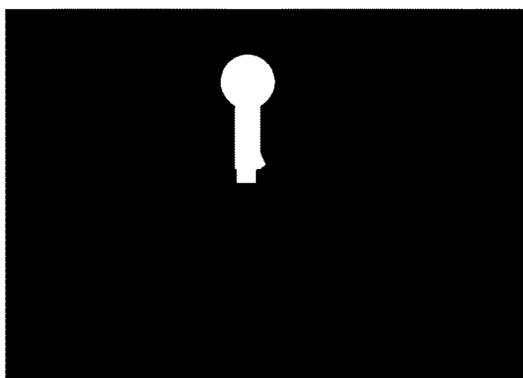
(b)



(c)

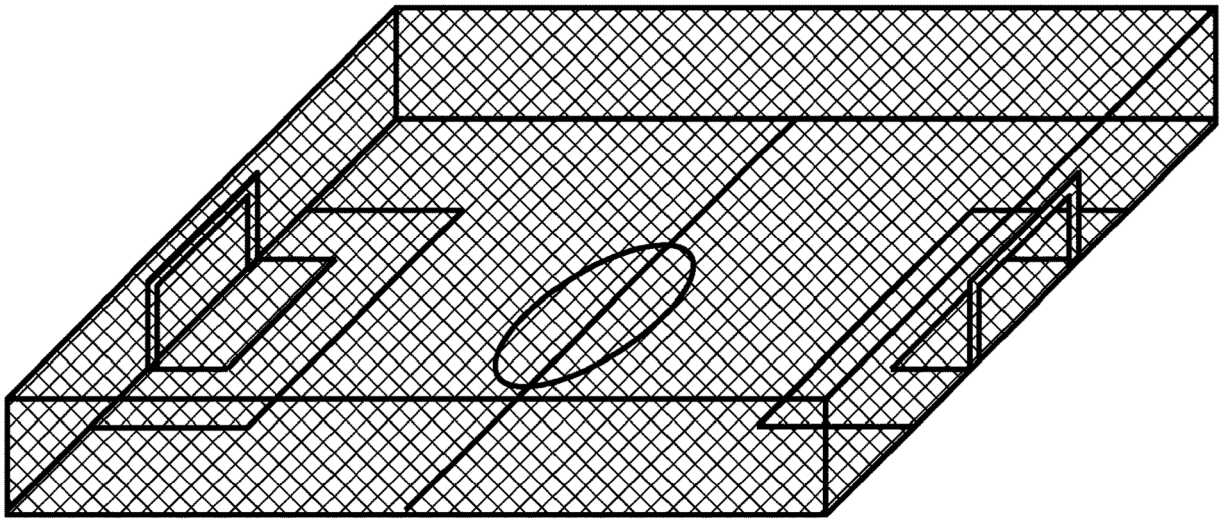


(d)



(e)

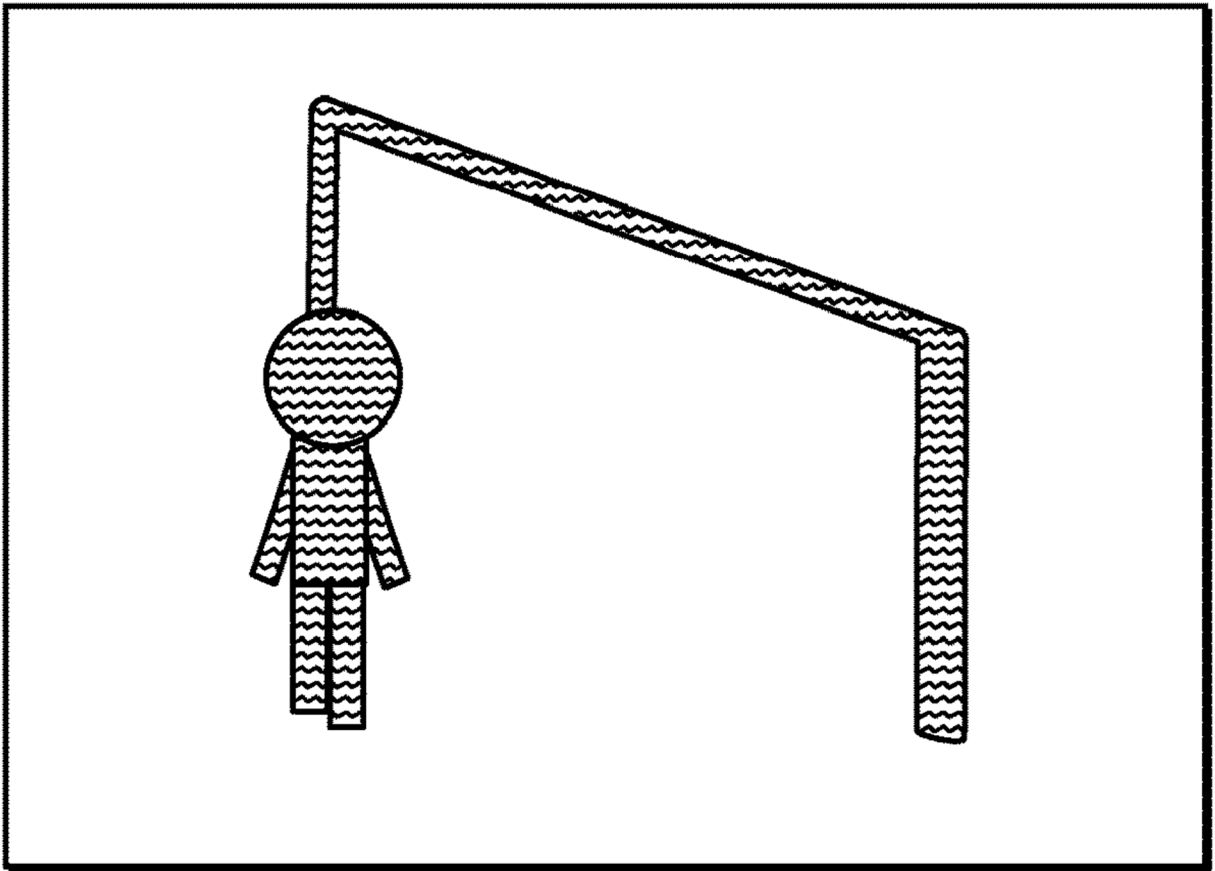
**FIG.20**



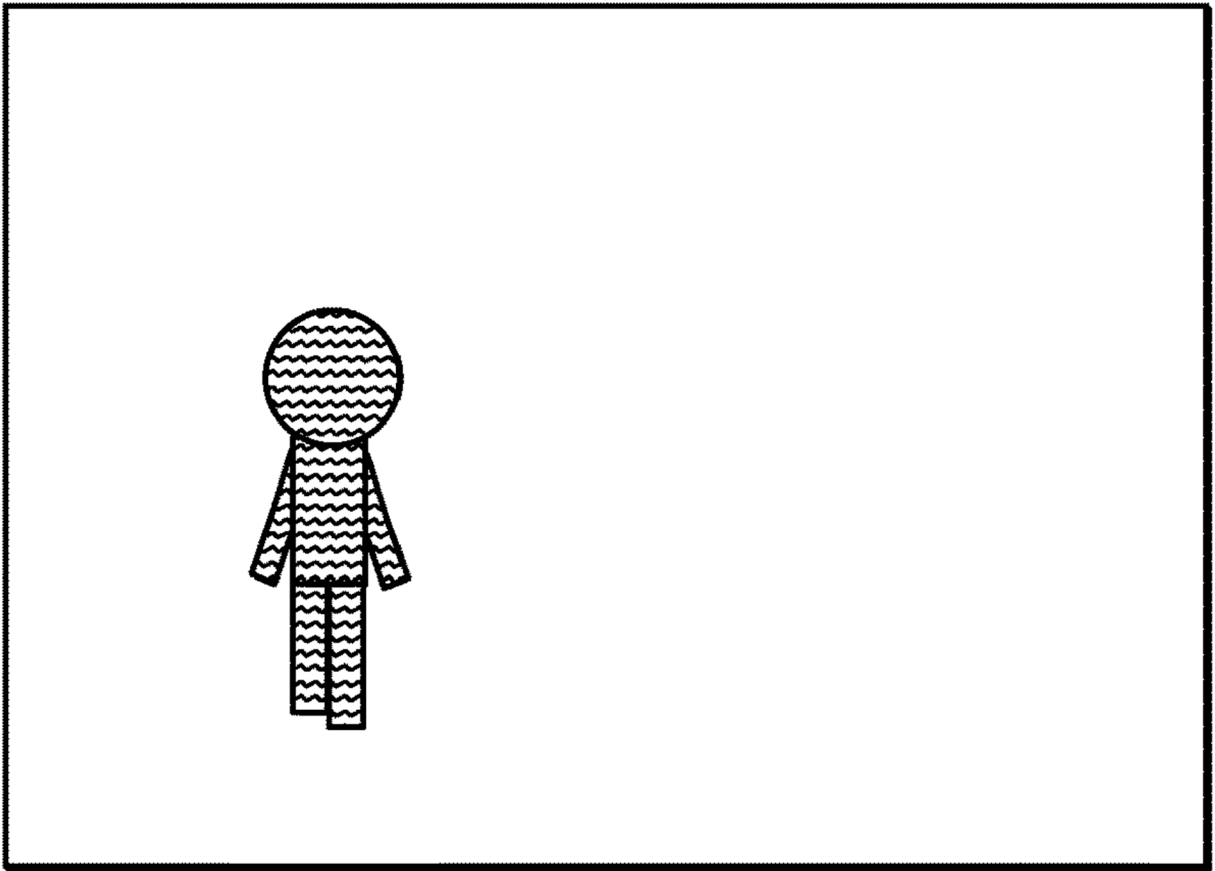
**FIG.21**

	CÓMPUTO FALSO	CÓMPUTO VERDADERO	RESULTADOS DE LA DETERMINACIÓN
PRIMER PLANO (PERSONA)	0	16	VÓXEL A MANTENER
PRIMER PLANO (PIERNA) CUYA PARTE NO SE DETECTA	0	15	VÓXEL A MANTENER
PRIMER PLANO (CABEZA) OCULTO POR ESTRUCTURA	0	13	VÓXEL A MANTENER
ESTRUCTURA (POSTE DE PORTERÍA)	0	0	VÓXEL A ELIMINAR
NO PRIMER PLANO (OTRA ÁREA)	16	-	VÓXEL A ELIMINAR

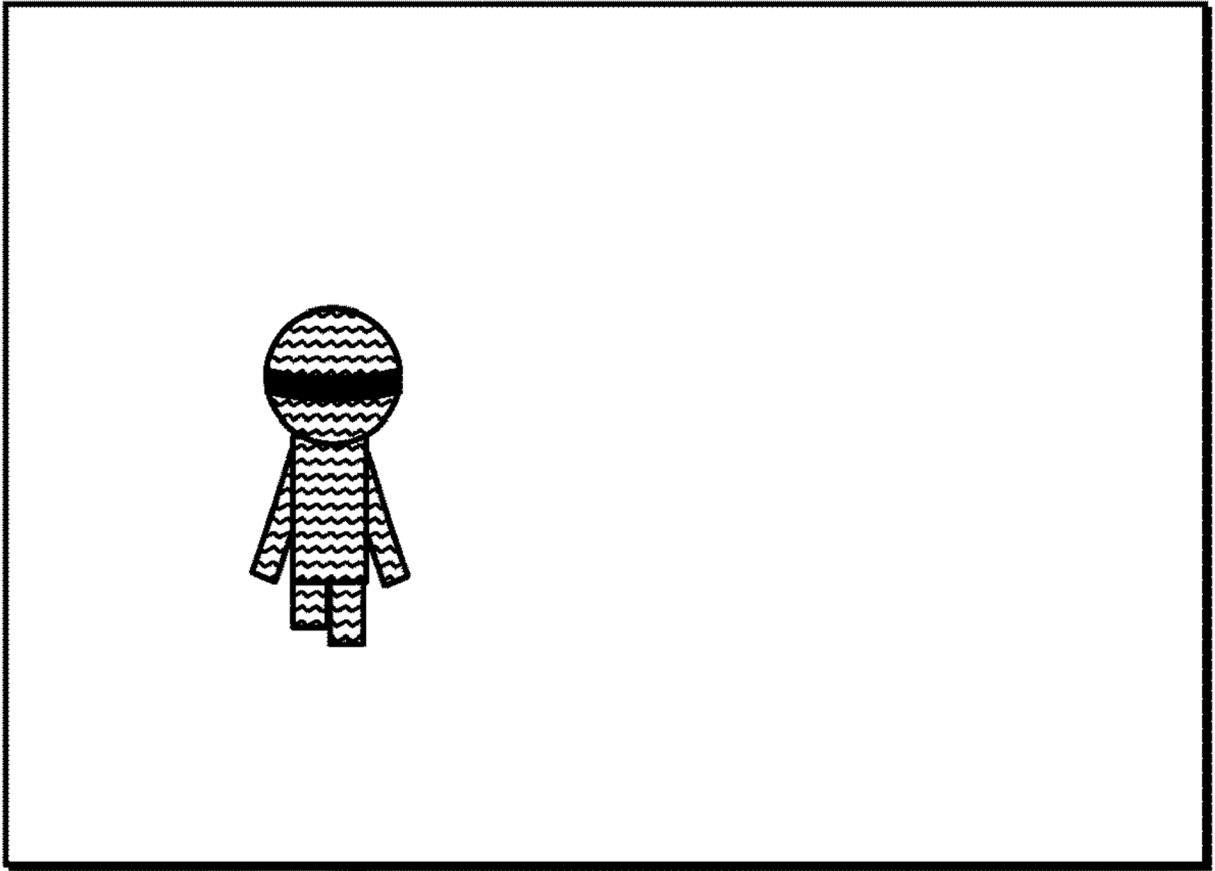
**FIG.22**



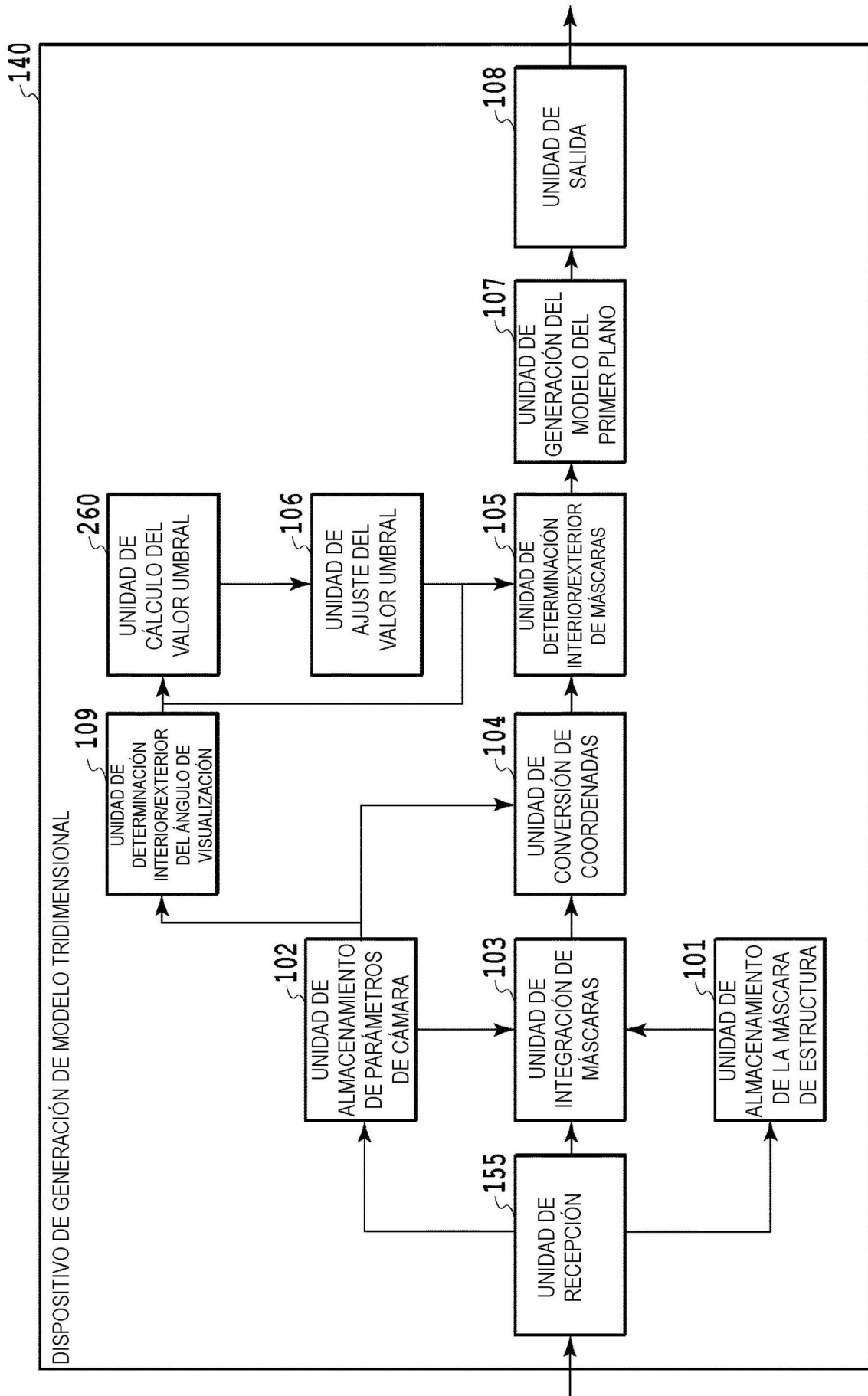
**FIG. 23**



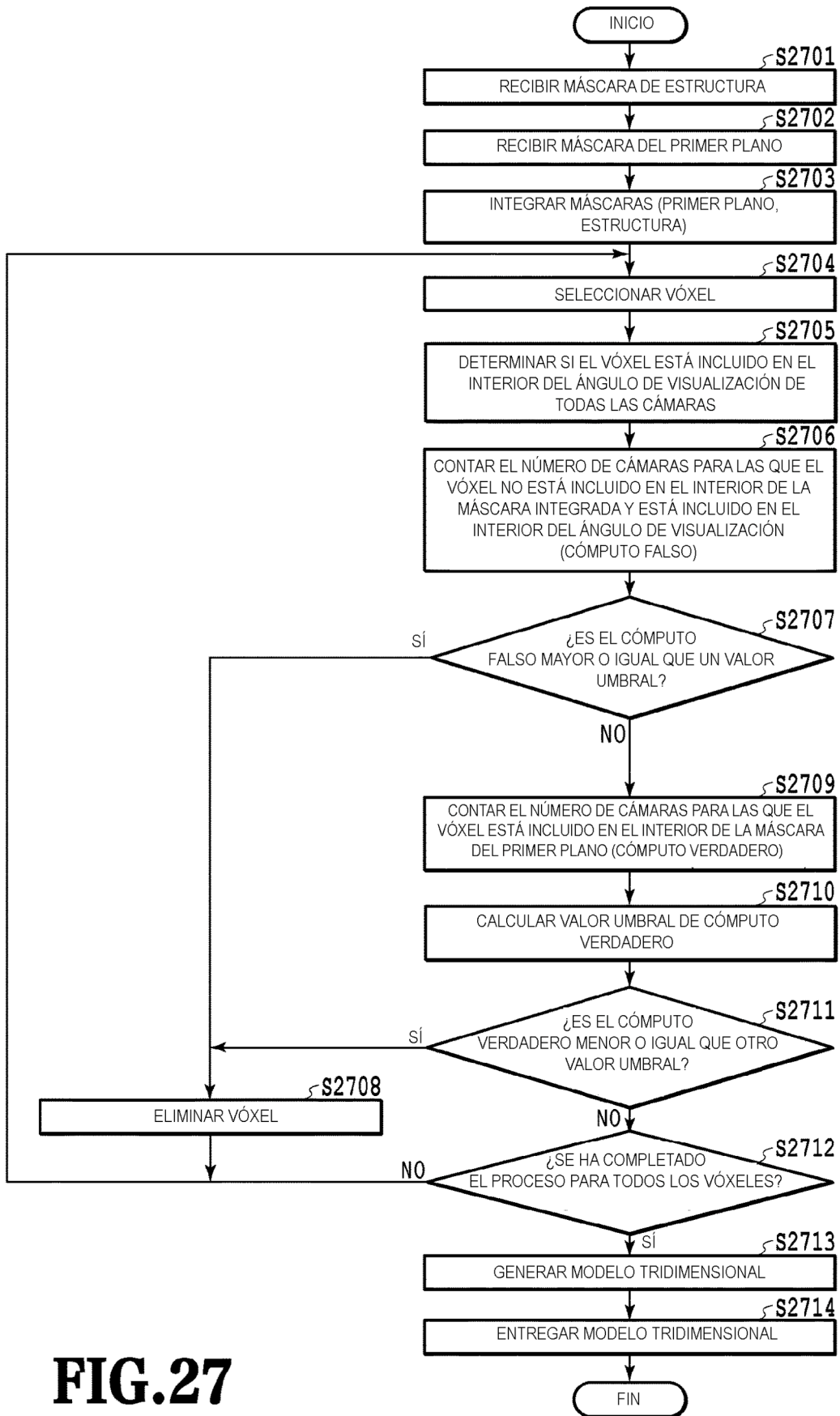
**FIG.24**



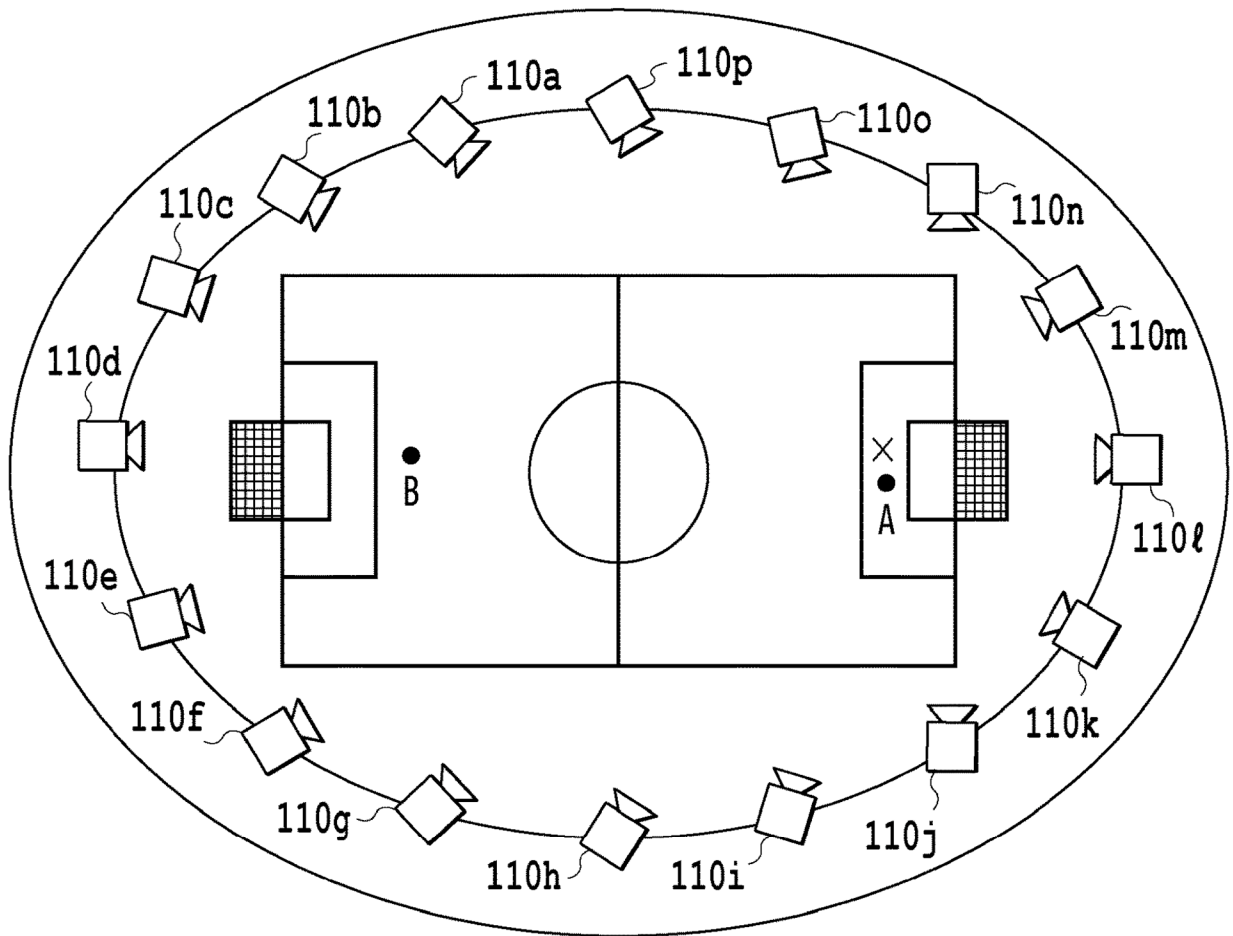
**FIG.25**



**FIG.26**



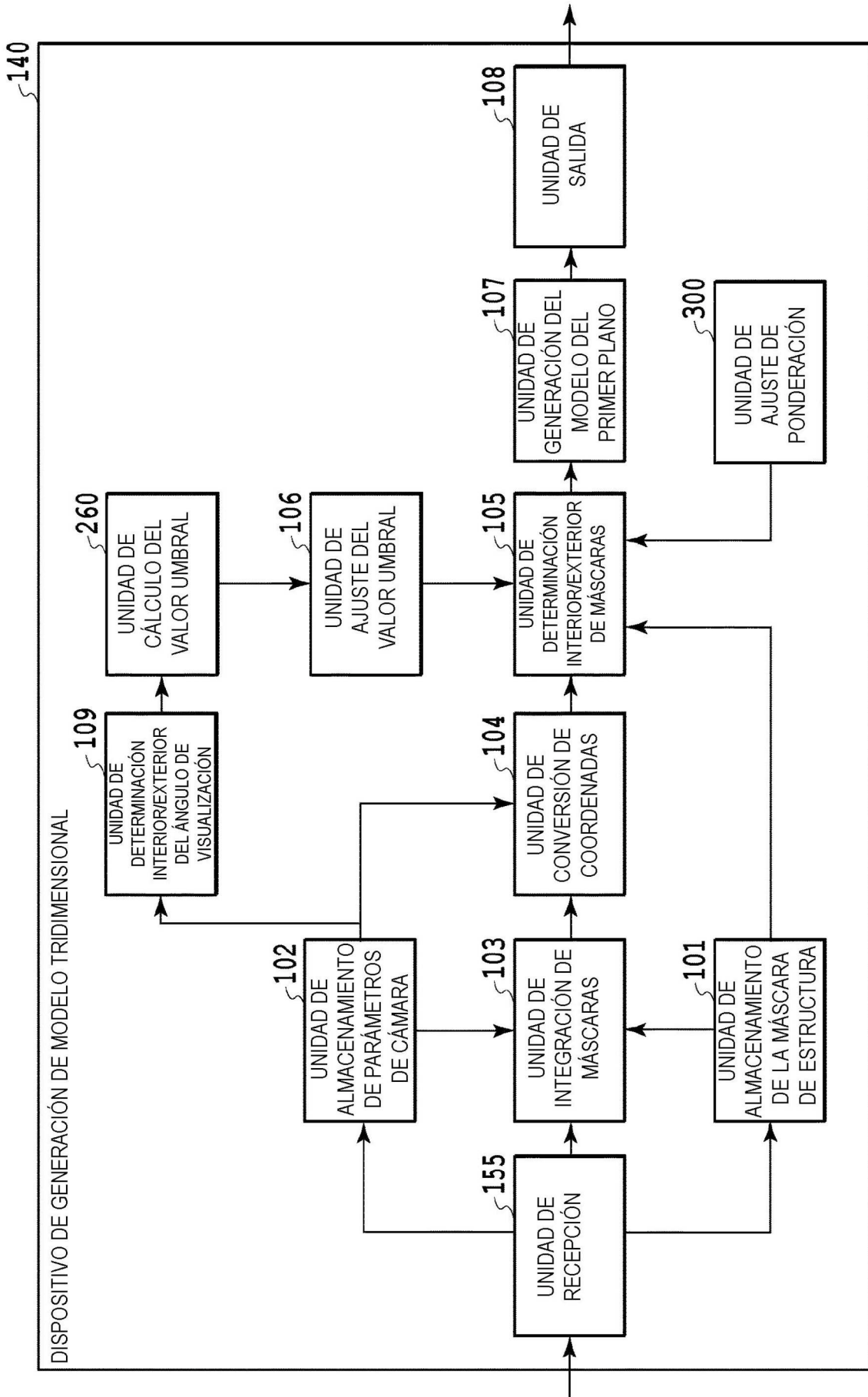
**FIG.27**



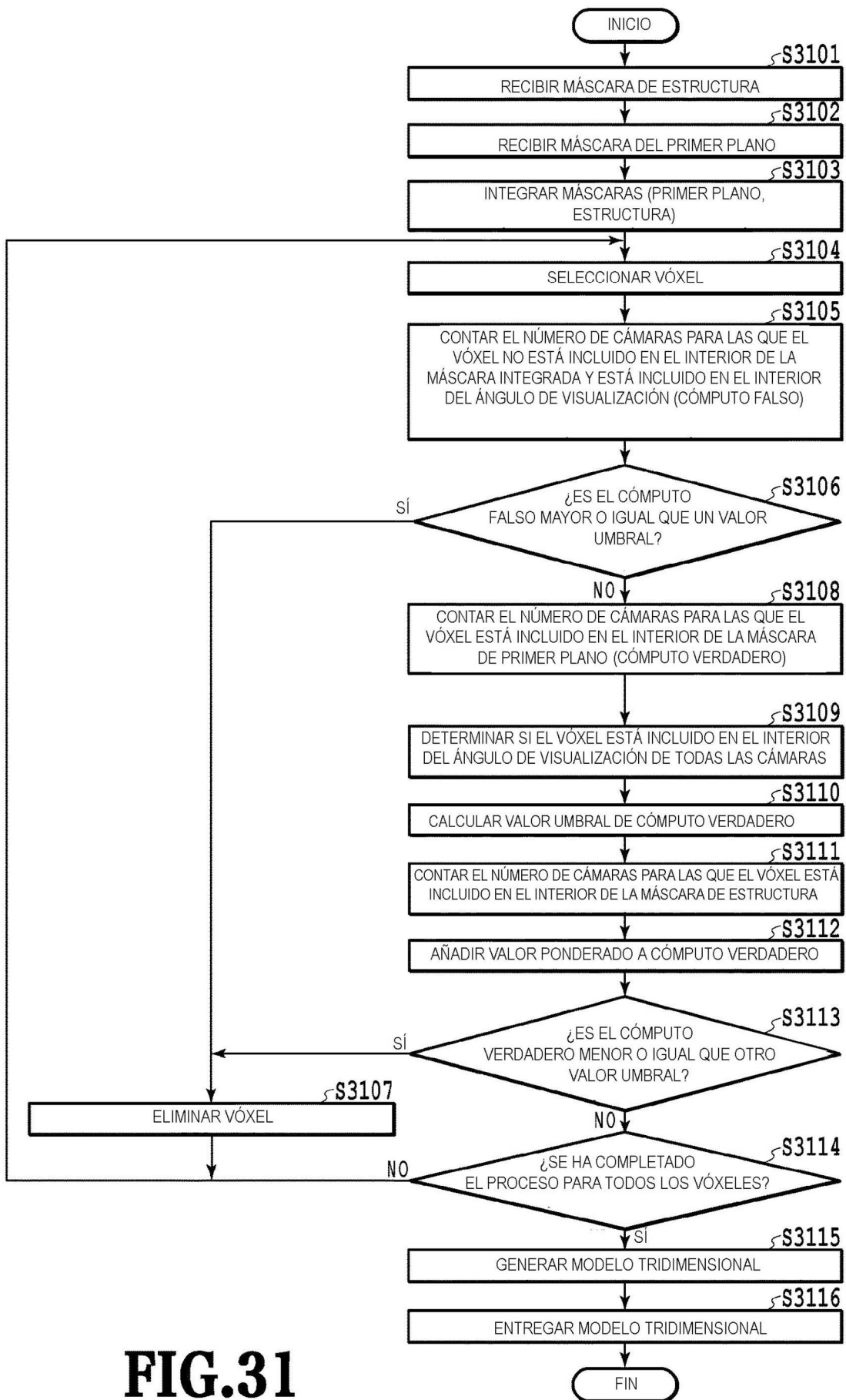
**FIG.28**

	CÓMPUTO FALSO	INTERIOR DEL ÁNGULO DE VISUALIZACIÓN	MAGNITUD RESTANTE DE CÓMPUTO VERDADERO	CÓMPUTO VERDADERO
PRIMER PLANO A	0	16	11,2	16
PRIMER PLANO B	0	3	2,1	3

**FIG.29**



**FIG.30**



**FIG.31**

	CÓMPUTO FALSO	NÚMERO DE CÁMARAS PARA LAS QUE EL VÓXEL ESTÁ INCLUIDO EN LA MÁSCARA DE PRIMER PLANO	NÚMERO DE CÁMARAS PARA LAS QUE EL VÓXEL ESTÁ INCLUIDO EN LA MÁSCARA DE ESTRUCTURA	CÓMPUTO VERDADERO
SIN ADICIÓN DE PONDERACIÓN	0	7	9	7
CON ADICIÓN DE PONDERACIÓN	0	7	9	11,5

**FIG.32**

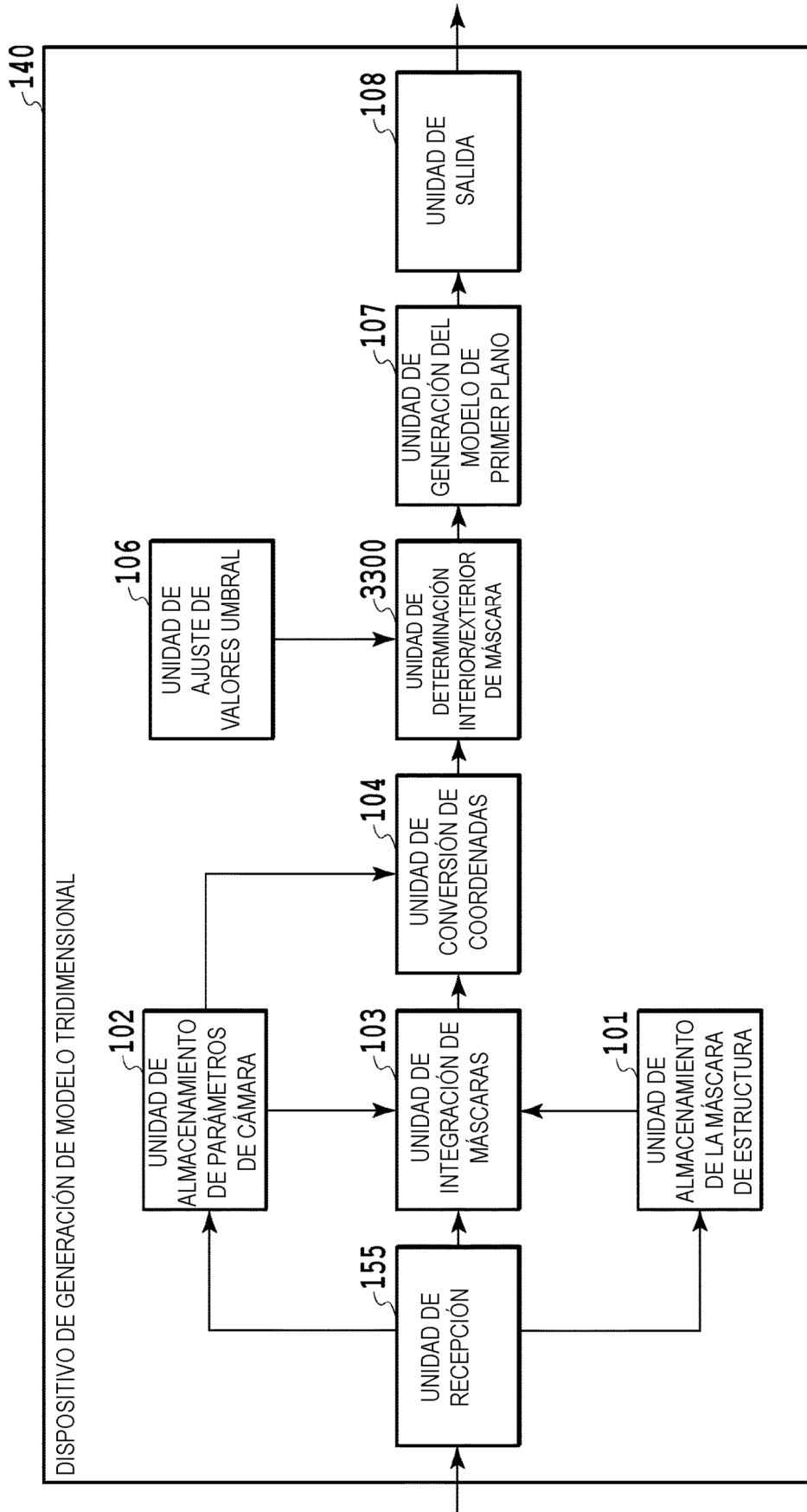
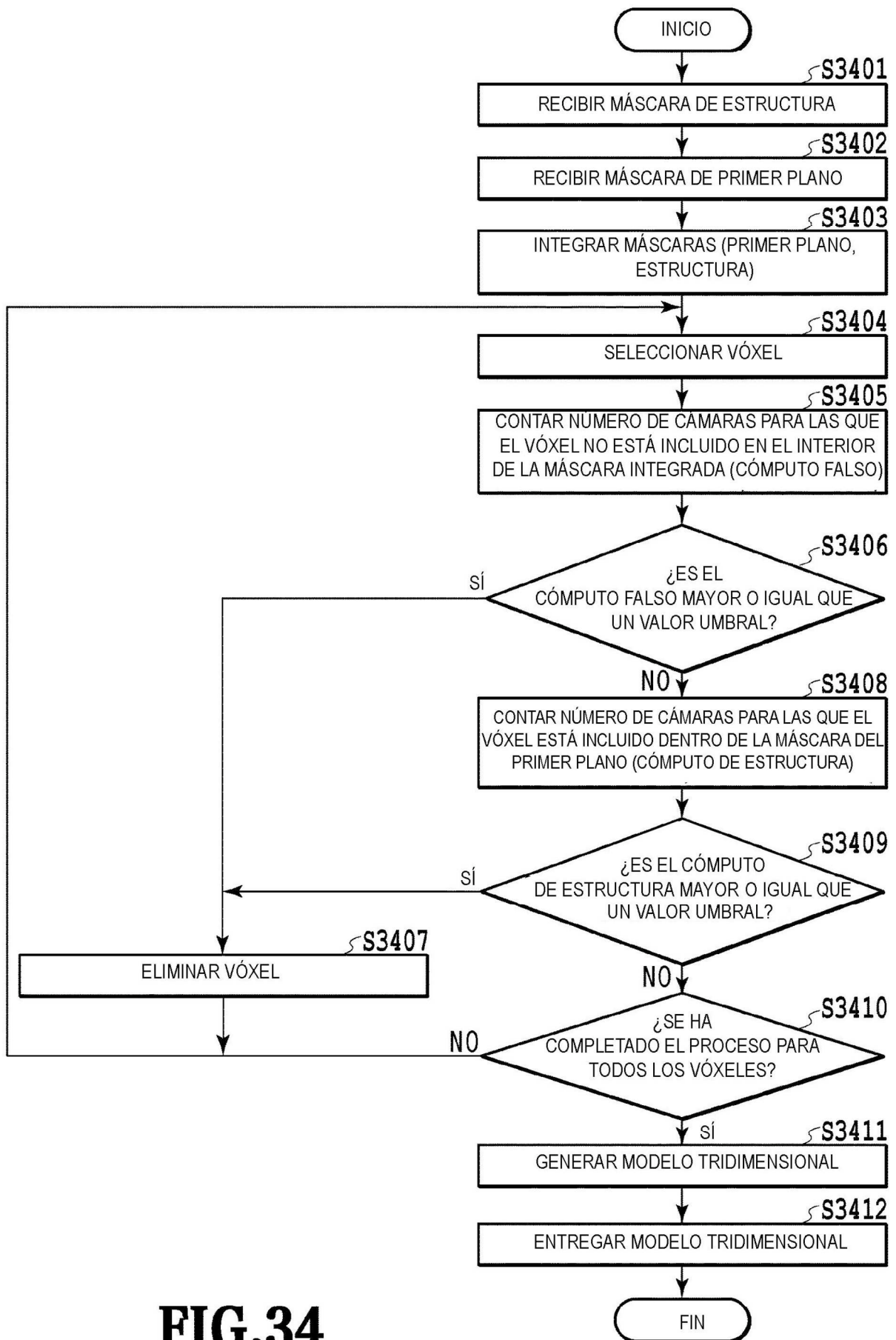


FIG.33



**FIG.34**

	CÓMPUTO FALSO	CÓMPUTO DE ESTRUCTURA	RESULTADOS DE LA DETERMINACIÓN
PRIMER PLANO (PERSONA)	0	0	VÓXEL A MANTENER
PRIMER PLANO (PIERNA) CUYA PARTE NO SE DETECTA	0	0	VÓXEL A MANTENER
PRIMER PLANO (CABEZA) OCULTO POR ESTRUCTURA	0	2	VÓXEL A MANTENER
ESTRUCTURA (POSTE DE PORTERÍA)	0	8	VÓXEL A ELIMINAR
NO PRIMER PLANO (OTRA ÁREA)	16	-	VÓXEL A ELIMINAR

**FIG.35**

**REFERENCIAS CITADAS EN LA DESCRIPCIÓN**

*Esta lista de referencias citada por el solicitante es únicamente para mayor comodidad del lector. No forman parte del documento de la Patente Europea. Incluso teniendo en cuenta que la compilación de las referencias se ha efectuado con gran cuidado, los errores u omisiones no pueden descartarse; la EPO se exime de toda responsabilidad al respecto.*

**Documentos de patentes citados en la descripción**

- US 2009141941 A
- JP 2014010805 A
- JP 2017239891 A
- JP 2018089467 A
- JP 2018209196 A

**Literatura no patente citada en la descripción**

- **STEFAN KUHN.** *Multi-view Reconstruction of Unknown Objects within a Known Environment*
- **LAURENTINI A.** The Visual Hull Concept for Silhouette-Based Image Understanding. *IEEE Transactions Pattern Analysis and machine Intelligence*, February 1994, vol. 16 (2), 150-162