

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 908 243**

51 Int. Cl.:

G06T 15/20 (2011.01)
G06T 17/00 (2006.01)
H04N 13/00 (2008.01)
H04N 5/222 (2006.01)
H04N 5/232 (2006.01)
H04N 5/247 (2006.01)
H04N 5/262 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **20.10.2017** **PCT/JP2017/037978**
87 Fecha y número de publicación internacional: **03.05.2018** **WO18079430**
96 Fecha de presentación y número de la solicitud europea: **20.10.2017** **E 17864980 (2)**
97 Fecha y número de publicación de la concesión europea: **16.02.2022** **EP 3534337**

54 Título: **Aparato de procesamiento de imágenes, sistema de procesamiento de imágenes, procedimiento de procesamiento de imágenes y programa**

30 Prioridad:

28.10.2016 JP 2016211905

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
28.04.2022

73 Titular/es:

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

72 Inventor/es:

TAKAMA YASUFUMI

74 Agente/Representante:

DURAN-CORRETJER, S.L.P

ES 2 908 243 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

Aparato de procesamiento de imágenes, sistema de procesamiento de imágenes, procedimiento de procesamiento de imágenes y programa

5 Sector técnico

La presente invención se refiere a una tecnología para generar una imagen de punto de vista virtual.

10 Estado de la técnica anterior

Últimamente, ha llamado la atención una tecnología para capturar un sujeto desde múltiples puntos de vista instalando de una pluralidad de cámaras en diferentes posiciones, y generando una imagen de punto de vista virtual o un modelo tridimensional utilizando una pluralidad de imágenes de punto de vista obtenidas mediante la captura. Según la tecnología para generar la imagen de punto de vista virtual a partir de la pluralidad de imágenes de punto de vista descritas anteriormente, por ejemplo, puesto que una escena destacada en fútbol o baloncesto se puede ver desde diversos ángulos, es posible proporcionar una sensación de gran realismo a un usuario en comparación con una imagen normal.

20 De acuerdo con el documento PTL 1, se describe que la calidad de imagen de la imagen de punto de vista virtual se mejora disminuyendo las unidades de representación en una zona límite de un objeto en la imagen en un caso en el que se va a generar una imagen de punto de vista virtual combinando entre sí de las imágenes capturadas desde la pluralidad de puntos de vista.

25 La Patente US2014/270706 describe el procesamiento de vídeos para generar vídeos virtuales que proporcionan diferentes puntos de vista del evento.

La Patente US2013/342647 detalla la generación de una imagen en color de un punto de vista de visualización utilizando una imagen en color y una imagen de profundidad de un punto de vista predeterminado.

30 La Patente WO2004/053591 describe un sistema de visualización y captura de imágenes en el que se genera una imagen sintetizada de un sujeto desde un punto de vista virtual a partir de imágenes tomadas por dos cámaras, y utilizando posicionamiento relativo de las cámaras.

35 Lista de referencias

Bibliografía de patentes

PTL 1: Patente japonesa presentada abierta a inspección pública núm. 2013-223008

40 Características de la invención

Problema técnico

45 Sin embargo, según la tecnología de la técnica relacionada, es concebible que, en algunos casos, no se pueda generar una imagen de punto de vista virtual según una pluralidad de requisitos diferentes con respecto a una calidad de imagen. Por ejemplo, en un caso en el que solo se va a generar la imagen de punto de vista virtual que tiene la calidad de imagen alta, es concebible que se alargue el tiempo de procesamiento relacionado con la generación, y existe el temor de que resulte difícil responder al deseo de un usuario que quisiera observar la imagen de punto de vista virtual en tiempo real, aunque la calidad de imagen sea baja. Por otro lado, en un caso en el que solo se va a generar la imagen de punto de vista virtual que tiene la calidad de imagen baja, existe un temor de que resulte difícil responder a un deseo de un usuario que prioriza la calidad de imagen alta de la imagen de punto de vista virtual sobre la propiedad de ser en tiempo real.

55 La presente invención ha sido realizada a la vista del problema descrito anteriormente, y tiene como objetivo generar una imagen de punto de vista virtual según una pluralidad de requisitos diferentes con respecto a una calidad de imagen.

Solución al problema

60 La presente invención ha sido realizada con el fin de resolver el problema descrito anteriormente. Según un aspecto de la presente invención, se da a conocer un sistema de procesamiento de imágenes tal como se especifica en las reivindicaciones 1 a 11. Según otro aspecto de la presente invención, se da a conocer un procedimiento de procesamiento de imágenes tal como se especifica en las reivindicaciones 12 a 14. Según otro aspecto de la presente invención, se da a conocer un programa informático tal como se especifica en la reivindicación 15.

Efectos ventajosos de la invención

Según la presente invención, es posible generar la imagen de punto de vista virtual según la pluralidad de requisitos diferentes con respecto a la calidad de la imagen.

Breve descripción de los dibujos

[Figura 1] La figura 1 es un diagrama explicativo para describir una configuración de un sistema de procesamiento de imágenes 10.

[Figura 2] La figura 2 es un diagrama explicativo para describir una configuración de hardware de un aparato de procesamiento de imágenes 1.

[Figura 3] La figura 3 es un diagrama de flujo para describir un modo de funcionamiento del aparato de procesamiento de imágenes 1.

[Figura 4] La figura 4 es un diagrama explicativo para describir una configuración de una pantalla de visualización mediante un aparato de visualización 3.

[Figura 5] La figura 5 es un diagrama de flujo para describir un modo de funcionamiento del aparato de procesamiento de imágenes 1.

[Figura 6] La figura 6 es un diagrama de flujo para describir un modo de funcionamiento del aparato de procesamiento de imágenes 1.

Descripción de realizaciones

[Configuración del sistema]

A continuación, se describirán realizaciones de la presente invención haciendo referencia a los dibujos. Primero, se describirá una configuración de un sistema de procesamiento 10 de imágenes que genera y entrega una imagen de punto de vista virtual utilizando la figura 1. El sistema de procesamiento de imágenes 10 según la presente realización incluye un aparato de procesamiento de imágenes 1, un grupo de cámaras 2, un aparato de visualización 3 y un aparato de visualización 4.

Cabe señalar que la imagen de punto de vista virtual, según la presente realización, es una imagen obtenida en un caso en el que un sujeto es capturado desde un punto de vista virtual. En otras palabras, la imagen de punto de vista virtual es una imagen que representa una apariencia en un punto de vista especificado. El punto de vista virtual (punto de vista imaginario) puede ser especificado por un usuario, o también, puede ser especificado de manera automática en base a un resultado de un análisis de imagen o similar. Es decir, la imagen de punto de vista virtual incluye una imagen de punto de vista arbitrario (imagen de punto de vista libre) correspondiente a un punto de vista especificado arbitrariamente por el usuario. Además, en la imagen de punto de vista virtual también se incluyen una imagen correspondiente a un punto de vista especificado por el usuario, de entre una pluralidad de candidatos, y una imagen correspondiente a un punto de vista especificado de manera automática por un aparato. Cabe señalar que, según la presente realización, se describirá principalmente un caso en el que la imagen de punto de vista virtual es una imagen en movimiento, pero la imagen de punto de vista virtual también puede ser una imagen fija.

El grupo de cámaras 2 incluye una pluralidad de cámaras, y cada una de las cámaras captura un sujeto desde direcciones diferentes respectivamente. Según la presente realización, cada una de la pluralidad de cámaras incluidas en el grupo de cámaras 2 está conectada al aparato de procesamiento de imágenes 1 y transmite una imagen capturada, un parámetro de cada cámara y similares al aparato de procesamiento de imágenes 1. Sin embargo, se debe tener en cuenta que la configuración no está limitada a esto, y la pluralidad de cámaras incluidas en el grupo de cámaras 2 pueden comunicar entre sí, y cualquiera de las cámaras incluidas en el grupo de cámaras 2 puede transmitir las imágenes capturadas por la pluralidad de cámaras, los parámetros de la pluralidad de cámaras y similares, al aparato de procesamiento de imágenes 1. Además, en lugar de las imágenes capturadas, cualquiera de las cámaras incluidas en el grupo de cámaras 2 puede transmitir una imagen basada en la captura por parte del grupo de cámaras 2, tal como una imagen generada en base a una diferencia entre las imágenes capturadas por la pluralidad de cámaras.

El aparato de visualización 3 acepta la especificación del punto de vista virtual para generar la imagen de punto de vista virtual, y transmite la información según la especificación, al aparato de procesamiento de imágenes 1. Por ejemplo, el aparato de visualización 3 incluye una unidad de entrada, tal como una palanca de mando, un disco selector, un panel táctil, un teclado o un ratón, y el usuario (operador) que especifica el punto de vista virtual acciona la unidad de entrada para especificar el punto de vista virtual. El usuario, según la presente realización, es un operador que acciona una unidad de entrada del aparato de visualización 3 para especificar el punto de vista virtual, o un espectador que observa la imagen de punto de vista virtual mostrada por el aparato de visualización 4, y el usuario simplemente se menciona en un caso en el que el operador y el espectador no se distinguen particularmente entre sí. Según la presente realización, se describirá principalmente el caso en el que el espectador es diferente del operador, pero la configuración no está limitada a esto, y el espectador y el operador pueden ser el mismo usuario. Cabe señalar que, según la presente realización, la información según la especificación del punto de vista virtual que es transmitida desde el aparato de visualización 3 al aparato de procesamiento de imágenes 1 es la información del

punto de vista virtual que indica la posición y la orientación del punto de vista virtual. Cabe señalar, sin embargo, que la configuración no está limitada a esto, y la información según la especificación del punto de vista virtual puede ser información que indica los contenidos determinados según el punto de vista virtual, tales como la forma o la orientación del sujeto en la imagen de punto de vista virtual, y el aparato de procesamiento de imágenes 1 puede generar la imagen de punto de vista virtual en base a la información descrita anteriormente, según la especificación del punto de vista virtual.

Además, el aparato de visualización 3 muestra la imagen de punto de vista virtual generada y entregada por el aparato de procesamiento de imágenes 1 en base a las imágenes basadas en la captura por parte del grupo de cámaras 2 y de la especificación del punto de vista virtual aceptada por el aparato de visualización 3. Según esto, el operador puede realizar la especificación del punto de vista virtual mientras observa la imagen de punto de vista virtual mostrada en el aparato de visualización 3. Cabe señalar que, según la presente realización, el aparato de visualización 3 que muestra la imagen de punto de vista virtual está configurado para aceptar la especificación del punto de vista virtual, pero la configuración no está limitada a esto. Por ejemplo, el aparato que acepta la especificación del punto de vista virtual y el aparato de visualización que muestra la imagen de punto de vista virtual para que el operador especifique el punto de vista virtual, pueden ser aparatos separados.

El aparato de visualización 3 también realiza una instrucción de generación para iniciar la generación de la imagen de punto de vista virtual con respecto al aparato de procesamiento de imágenes 1 en base a la operación del operador. Cabe señalar que la instrucción de generación no está limitada a esto, y puede ser una instrucción para que el aparato de procesamiento de imágenes 1 reserve la generación de la imagen de punto de vista virtual de tal manera que la generación de la imagen de punto de vista virtual se inicia en un tiempo predeterminado, por ejemplo. Además, la instrucción de generación puede ser una instrucción para una reserva, de tal manera que la generación de la imagen de punto de vista virtual se inicia en un caso en el que ocurre un evento predeterminado, por ejemplo. Cabe señalar que el aparato que realiza la instrucción de generación de la imagen de punto de vista virtual con respecto al aparato de procesamiento de imágenes 1 puede ser un aparato diferente del aparato de visualización 3, y el usuario puede introducir directamente la instrucción de generación con respecto al aparato de procesamiento de imágenes 1.

El aparato de visualización 4 muestra la imagen de punto de vista virtual generada por el aparato de procesamiento de imágenes 1 en base a la especificación del punto de vista virtual por el operador que ha utilizado el aparato de visualización 3 con respecto al usuario (espectador) diferente del operador que especifica el punto de vista virtual. Cabe señalar que el sistema de procesamiento de imágenes 10 puede incluir una pluralidad de aparatos de visualización 4, y la pluralidad de aparatos de visualización 4 puede mostrar, respectivamente, diferentes imágenes de punto de vista virtual. Por ejemplo, el aparato de visualización 4 que muestra la imagen de punto de vista virtual (imagen en directo) que se difundirá en directo y el aparato de visualización 4 que muestra la imagen de punto de vista virtual (imagen no en directo) que será difundida después de la grabación, pueden estar incluidos en el sistema de procesamiento de imágenes 10.

El aparato de procesamiento de imágenes 1 incluye una unidad de obtención de información de la cámara 100, una unidad de obtención de información del punto de vista virtual 110 (en adelante, la unidad de obtención de puntos de vista 110), una unidad de generación de imágenes 120 y una unidad de salida 130. La unidad de obtención de información de la cámara 100 obtiene las imágenes basadas en la captura por parte del grupo de cámaras 2, parámetros externos y parámetros internos de las respectivas cámaras incluidas en el grupo de cámaras 2, y similares, del grupo de cámaras 2, para su entrega a la unidad de generación de imágenes 120. La unidad de obtención de puntos de vista 110 obtiene la información según la especificación del punto de vista virtual por parte del operador desde el aparato de visualización 3 para enviarla a la unidad de generación de imágenes 120. La unidad de obtención de puntos de vista 110 también acepta la instrucción de generación de la imagen de punto de vista virtual por parte del aparato de visualización 3. La unidad de generación de imágenes 120 genera la imagen de punto de vista virtual en base a las imágenes basadas en la captura que son obtenidas por la unidad de obtención de información de la cámara 100, la información según la especificación obtenida por la unidad de obtención de puntos de vista 110, y la instrucción de generación aceptada por la unidad de obtención de puntos de vista 110 para entregar a la unidad de salida 130. La unidad de salida 130 entrega la imagen de punto de vista virtual generada por la unidad de generación de imágenes 120 al aparato externo, como el aparato de visualización 3 o el aparato de visualización 4.

Cabe señalar que, según la presente realización, el aparato de procesamiento de imágenes 1 genera la pluralidad de imágenes de punto de vista virtual que tienen diferentes calidades de imagen, para ser entregadas a los destinos de salida según las respectivas imágenes de punto de vista virtual. Por ejemplo, la imagen de punto de vista virtual que tiene una calidad de imagen baja, en la que el tiempo de procesamiento relacionado con la generación es corto, es entregada al aparato de visualización 4 observado por el espectador que desea la imagen de punto de vista virtual en tiempo real (bajo retardo). Por otro lado, la imagen de punto de vista virtual que tiene la calidad de imagen alta, en la que el tiempo de procesamiento relacionado con la generación es largo, es entregada al aparato de visualización 4 observado por el espectador que desea que la imagen de punto de vista virtual tenga la calidad de imagen alta. Cabe señalar que el retardo según la presente realización corresponde a un período desde que se realiza la captura por parte del grupo de cámaras 2 hasta que se muestra la imagen de punto de vista virtual basada

en la captura. Cabe señalar, sin embargo, que la definición del retardo no está limitada a esto y, por ejemplo, se puede establecer como retardo una diferencia de tiempo entre una hora del mundo real y una hora correspondiente a la imagen visualizada.

5 Posteriormente, se describirá una configuración de hardware del aparato de procesamiento de imágenes 1 utilizando la figura 2. El aparato de procesamiento de imágenes 1 incluye una CPU 201, una ROM 202, una RAM 203, un dispositivo de almacenamiento auxiliar 204, una unidad de visualización 205, una unidad operativa 206, una unidad de comunicación 207 y un bus 208. La CPU 201 controla la totalidad del aparato de procesamiento de imágenes 1 utilizando programas informáticos y datos almacenados en la ROM 202 o la RAM 203. Cabe señalar que el aparato
10 de procesamiento de imágenes 1 puede incluir una GPU (Unidad de procesamiento de gráficos, Graphics Processing Unit), y la GUI puede realizar por lo menos parte del procesamiento mediante la CPU 201. La ROM 202 almacena los programas y parámetros que no requieren cambios. La RAM 203 almacena temporalmente los programas y datos suministrados desde el dispositivo de almacenamiento auxiliar 204, datos suministrados desde el exterior a través de la unidad de comunicación 207 y similares. El dispositivo de almacenamiento auxiliar 204 está
15 constituido, por ejemplo, por una unidad de disco duro o similar, y almacena datos de contenido tales como una imagen fija y una imagen en movimiento.

La unidad de visualización 205 está constituida, por ejemplo, por una pantalla de cristal líquido o similar, y muestra una GUI (interfaz gráfica de usuario, Graphical User Interface) para que el usuario accione el aparato de
20 procesamiento de imágenes 1 y similares. La unidad operativa 206 está constituida, por ejemplo, por un teclado, un ratón o similar, y acepta las operaciones del usuario e introduce varias instrucciones a la CPU 201. La unidad de comunicación 207 realiza una comunicación con un aparato externo tal como el grupo de cámaras 2, el aparato de visualización 3 o el aparato de visualización 4. Por ejemplo, un cable de LAN o similar está conectado a la unidad de comunicación 207 en un caso en el que el aparato de procesamiento de imágenes 1 está conectado al aparato
25 externo de manera cableada. Cabe señalar que, en caso de que el aparato de procesamiento de imágenes 1 incluya una función para comunicarse de manera inalámbrica con un aparato externo, la unidad de comunicación 207 está dotada de una antena. El bus 208 transmite la información conectando las unidades respectivas del aparato de procesamiento de imágenes 1, entre sí.

30 Cabe señalar que, según la presente realización, la unidad de visualización 205 y la unidad operativa 206 existen en el interior del aparato de procesamiento de imágenes 1, pero también se puede adoptar una configuración en la que el aparato de procesamiento de imágenes 1 no esté dotado, por lo menos, de uno de la unidad de visualización 205 y la unidad operativa 206. Además, por lo menos una de la unidad de visualización 205 y la unidad operativa 206 pueden existir fuera del aparato de procesamiento de imágenes 1 como otro aparato, y la CPU 201 puede funcionar
35 como una unidad de control de visualización que controla la unidad de visualización 205, y una unidad de control operativa, que controla la unidad operativa 206.

[Flujo de funcionamiento]

40 A continuación, se describirá un modo de funcionamiento del aparato de procesamiento de imágenes 1 utilizando la figura 3. El procesamiento mostrado en la figura 3 se inicia cuando la unidad de obtención de puntos de vista 110 realiza la aceptación de la instrucción de generación de la imagen de punto de vista virtual y se repite periódicamente (por ejemplo, cada fotograma, en un caso en el que la imagen de punto de vista virtual es una imagen en movimiento). Cabe señalar, sin embargo, que el tiempo de inicio del procesamiento mostrado en la figura
45 3 no está limitado al tiempo descrito anteriormente. El procesamiento mostrado en la Figura 3 se realiza cuando la CPU 201 expande los programas almacenados en la ROM 202 hacia la RAM 203, para ser ejecutados. Cabe señalar que, por lo menos, una parte del procesamiento mostrado en la figura 3 puede ser realizada mediante un hardware de uso exclusivo diferente de la CPU 201.

50 En el flujo mostrado en la figura 3, S2010 y S2020 corresponden al procesamiento para obtener la información, y S2030 a S2050 corresponden al procesamiento para generar y entregar la imagen de punto de vista virtual (imagen de especificación) para que el operador especifique el punto de vista virtual. Además, S2070 a S2100 corresponden al procesamiento para generar y entregar la imagen en directo. S2110 a S2130 corresponden al procesamiento para generar y entregar la imagen no en directo. A continuación, se describirán detalles de los procesos en las etapas
55 respectivas.

En S2010, la unidad de obtención de información de la cámara 100 obtiene las imágenes capturadas de las cámaras respectivas basándose en la captura por parte del grupo de cámaras 2 y en los parámetros externos e internos de las respectivas cámaras. El parámetro externo es información con respecto a una posición y una orientación de la
60 cámara, y el parámetro interno es información con respecto a una distancia focal y un centro de imagen de la cámara.

En S2020, la unidad de obtención de puntos de vista 110 obtiene la información del punto de vista virtual como la información según la especificación del punto de vista virtual, por parte del operador. Según la presente realización,
65 la información del punto de vista virtual corresponde a un parámetro externo y a un parámetro interno de una cámara

virtual que captura el sujeto desde el punto de vista virtual, y se necesita un fragmento de información del punto de vista virtual para generar un fotograma de la imagen de punto de vista virtual.

En S2030, la unidad de generación de imágenes 120 estima una forma tridimensional de un objeto correspondiente al sujeto en base a las imágenes capturadas por el grupo de cámaras 2. El objeto correspondiente al sujeto es, por ejemplo, una persona, un objeto en movimiento o similar, que existe en un rango de captura del grupo de cámaras 2. La unidad de generación de imágenes 120 calcula diferencias entre las imágenes capturadas obtenidas del grupo de cámaras 2 y las imágenes del fondo obtenidas previamente, correspondientes a las respectivas cámaras, para generar imágenes de silueta en las que una parte (zona del primer plano) correspondiente al objeto en la imagen capturada es extraída. A continuación, la unidad de generación de imágenes 120 utiliza las imágenes de silueta correspondientes a las respectivas cámaras y los parámetros de las respectivas cámaras para estimar la forma tridimensional del objeto. Por ejemplo, se utiliza una técnica de Visual Hull para la estimación de la forma tridimensional. Como resultado de este procesamiento se obtiene un grupo de puntos 3D (conjunto de puntos que tienen coordenadas tridimensionales) que representa la forma tridimensional del objeto correspondiente al sujeto. Cabe señalar que el procedimiento para obtener la forma tridimensional del objeto a partir de las imágenes capturadas por el grupo de cámaras 2 no está limitada a esto.

En S2040, la unidad de generación de imágenes 120 realiza la representación del grupo de puntos 3D y un modelo 3D del fondo en base a la información de punto de vista virtual obtenida, y genera la imagen de punto de vista virtual. El modelo 3D del fondo es, por ejemplo, un modelo CG, tal como un circuito de carreras en el que está instalado el grupo de cámaras 2, y se crea previamente para ser guardado en el sistema de procesamiento de imágenes 10. En la imagen de punto de vista virtual generada por el procesamiento hasta ahora, la zona correspondiente al objeto y la zona del fondo se muestran respectivamente en colores predeterminados (por ejemplo, un solo color). Cabe señalar que el procesamiento para realizar la representación del grupo de puntos 3D y el modelo 3D del fondo ya se conoce en el sector de los juegos y el cine, y se conoce un procedimiento para realizar el procesamiento rápidamente, tal como, por ejemplo, un procedimiento para realizar el procesamiento utilizando la GPU. Por esta razón, la imagen de punto de vista virtual generada en el procesamiento hasta S2040 puede ser generada rápidamente según la captura por parte del grupo de cámaras 2 y la especificación del punto de vista virtual por parte del operador.

En S2050, la unidad de salida 130 entrega la imagen de punto de vista virtual generada en S2040 por la unidad de generación de imágenes 120 al aparato de visualización 3 para que el operador especifique el punto de vista virtual. En este caso, se describirá una configuración de pantalla de una pantalla de visualización 30 del aparato de visualización 3 utilizando la figura 4. La pantalla de visualización 30 está constituida por una zona 310, una zona 320 y una zona 330. Por ejemplo, la imagen de punto de vista virtual generada como imagen de la especificación se muestra en la zona 310, la imagen de punto de vista virtual generada como la imagen en directo se muestra en la zona 320, y la imagen de punto de vista virtual generada como imagen no en directo se muestra en la zona 330. Es decir, la imagen de punto de vista virtual generada en S2040 y enviada en S2050 se muestra en la zona 310. A continuación, el operador realiza la especificación del punto de vista virtual mientras observa la pantalla de la zona 310. Cabe señalar que es suficiente cuando el aparato de visualización 3 muestra por lo menos la imagen de la especificación y no necesariamente necesita mostrar la imagen en directo y la imagen no en directo.

En S2060, la unidad de generación de imágenes 120 determina si se realiza o no el procesamiento para generar la imagen de punto de vista virtual que tiene una calidad de imagen superior mayor que la imagen de punto de vista virtual generada en S2040. Por ejemplo, en un caso en el que solo se necesita la imagen que tiene la calidad de imagen baja para especificar el punto de vista virtual, el flujo no avanza hacia S2070, y el procesamiento finaliza. Por otro lado, en el caso de que se necesite una imagen que tenga la calidad de imagen más alta, el flujo avanza hacia S2070 y el procesamiento continúa.

En S2070, la unidad de generación de imágenes 120 aumenta más la precisión del modelo de forma del objeto (grupo de puntos 3D) que se estima en S2030 utilizando una técnica de Photo Hull, por ejemplo. Específicamente, al proyectar los puntos respectivos del grupo de puntos 3D en las imágenes capturadas de las cámaras respectivas, y evaluar las tasas de coincidencia de color en las imágenes capturadas respectivas, se determina si el punto es o no un punto necesario para representar la forma del sujeto. Por ejemplo, con respecto a un determinado punto en el grupo de puntos 3D, una variación de un valor de píxel en un destino de proyección es mayor que un umbral, se determina que el punto no es correcto dado que el punto que representa la forma del sujeto y el punto es eliminado del grupo de puntos 3D. Este procesamiento se realiza con respecto a todos los puntos en el grupo de puntos 3D para realizar el aumento en la precisión del modelo de forma del objeto. Cabe señalar que el procedimiento de aumentar la precisión y del modelo de forma del objeto no está limitado a esto.

En S2080, la unidad de generación de imágenes 120 ejecuta el procesamiento para colorear el grupo de puntos 3D en el que se incrementa la precisión en S2070 y lo proyecta sobre las coordenadas del punto de vista virtual, para generar una imagen de primer plano correspondiente a la zona de primer plano, y el procesamiento para generar una imagen del fondo vista desde el punto de vista virtual. A continuación, la unidad de generación de imágenes 120 superpone la imagen de primer plano sobre la imagen del fondo generada, para generar la imagen de punto de vista virtual como la imagen en directo.

En el presente documento, se describirá un ejemplo del procedimiento de generación de imágenes de primer plano (imagen de la zona correspondiente al objeto) de la imagen de punto de vista virtual. El procesamiento para colorear el grupo de puntos 3D se ejecuta para generar la imagen de primer plano. El procesamiento de coloreado está
 5 constituido por la determinación de la visibilidad del punto y el procesamiento de cálculo del color. En la determinación de la visibilidad, es posible identificar las cámaras que pueden realizar la captura con respecto a los respectivos puntos a partir de relaciones posicionales entre los respectivos puntos en el grupo de puntos 3D y la pluralidad de cámaras incluidas en el grupo de cámaras 2. A continuación, con respecto a los puntos respectivos, se proyecta un punto sobre la imagen capturada de la cámara que puede capturar el punto, y el color de un píxel en el
 10 destino de la proyección se establece como el color del punto. En el caso de que un cierto punto sea capturado por una pluralidad de cámaras, el punto se proyecta sobre las imágenes capturadas de la pluralidad de cámaras y se obtienen los valores de píxel en el destino de la proyección, de modo que el color del punto se decide calculando un promedio de los valores de píxel. Cuando la representación del grupo de puntos 3D coloreado de este modo se realiza mediante una tecnología de representación CG de la técnica relacionada, es posible generar la imagen del
 15 primer plano de la imagen de punto de vista virtual.

A continuación, se describirá un ejemplo del procedimiento para generar la imagen del fondo de la imagen de punto de vista virtual. En primer lugar, se establecen los vértices del modelo 3D del fondo (por ejemplo, los puntos correspondientes a los bordes del circuito de carreras). A continuación, estos vértices son proyectados sobre
 20 sistemas de coordenadas de dos cámaras (establecidas como una primera cámara y una segunda cámara) cerca del punto de vista virtual y un sistema de coordenadas del punto de vista virtual. Además, se calcula una primera matriz de proyección entre el punto de vista virtual y la primera cámara, y una segunda matriz de proyección entre el punto de vista virtual y la segunda cámara, utilizando los puntos correspondientes del punto de vista virtual y la primera cámara y los puntos correspondientes del punto de vista virtual y la segunda cámara. A continuación, la
 25 imagen capturada de la primera cámara y la imagen capturada de la segunda cámara son proyectadas sobre los píxeles respectivos de la imagen del fondo utilizando la primera matriz de proyección y la segunda matriz de proyección, y se calcula el promedio de los valores de píxel en el destino de la proyección, de modo que se deciden los valores de píxel de la imagen del fondo. Cabe señalar que los valores de píxel de la imagen del fondo pueden decidirse a partir de las imágenes capturadas de tres o más cámaras mediante un procedimiento similar.

La imagen de punto de vista virtual coloreada puede ser generada superponiendo la imagen de primer plano sobre la imagen del fondo así obtenida de la imagen de punto de vista virtual. Es decir, la imagen de punto de vista virtual generada en S2080 tiene una calidad de imagen más alta que la imagen de punto de vista virtual generada en
 30 S2040 con respecto al número de gradaciones de los colores. Por el contrario, el número de gradaciones de los colores incluidos en la imagen de punto de vista virtual generada en S2040 es menor que el número de gradaciones de los colores incluidos en la imagen de punto de vista virtual generada en S2080. Cabe señalar que el procedimiento de añadir la información de color a la imagen de punto de vista virtual no está limitado a esto.

En S2090, la unidad de salida 130 entrega la imagen de punto de vista virtual generada en S2080 por la unidad de
 40 generación de imágenes 120 al aparato de visualización 3 y al aparato de visualización 4 como la imagen en directo. La imagen entregada al aparato de visualización 3 se muestra en la zona 320 y puede ser observada por el operador, y el espectador puede observar la imagen entregada al aparato de visualización 4.

En S2100, la unidad de generación de imágenes 120 determina si se realiza o no el procesamiento para generar la
 45 imagen de punto de vista virtual que tiene una calidad de imagen más alta que la imagen de punto de vista virtual generada en S2080. Por ejemplo, en un caso en el que la imagen de punto de vista virtual solo se proporciona para su difusión en directo con respecto al espectador, el flujo no avanza hacia S2110 y el procesamiento finaliza. Por otra parte, en un caso en el que la imagen que tiene la calidad de imagen más alta va a ser difundida hacia el espectador después de la grabación, el flujo avanza hacia S2110, y el procesamiento continúa.

En S2110, la unidad de generación de imágenes 120 aumenta más la precisión del modelo de forma del objeto
 50 generado en S2070. Según la presente realización, el aumento de la precisión se realiza eliminando un punto aislado del modelo de forma. En la eliminación de puntos aislados, en primer lugar, con respecto a un conjunto de vóxeles (grupo de puntos 3D) calculado mediante Photo Hull, se investiga si existe o no otro vóxel en el entorno de los respectivos vóxeles. En un caso en el que el vóxel no existe en el entorno, se determina que el vóxel es el punto
 55 aislado y el vóxel es eliminado del conjunto de vóxeles. Cuando se ejecuta un procesamiento similar a S2080 utilizando el modelo de forma del que se elimina el punto aislado, se genera la imagen de punto de vista virtual en la que se obtiene una mayor precisión de la forma del objeto que la imagen de punto de vista virtual generada en S2080.

En S2120, la unidad de generación de imágenes 120 aplica un procesamiento de suavizado a un límite entre la zona
 60 del primer plano de la imagen de punto de vista virtual generada en S2110 y la zona del fondo, y corrige la imagen de tal manera que una zona de límite se muestra suavemente.

En S2130, la unidad de salida 130 entrega la imagen de punto de vista virtual generada por la unidad de generación de imágenes 120 en S2120 al aparato de visualización 3 y al aparato de visualización 4 como la imagen no en directo. La imagen no en directo entregada al aparato de visualización 3 se visualiza en la zona 330.

Mediante el procesamiento descrito anteriormente, el aparato de procesamiento de imágenes 1 genera la imagen de punto de vista virtual como la imagen de la especificación para que el operador especifique el punto de vista virtual y la imagen en directo correspondiente a la imagen de punto de vista virtual que tiene una calidad de imagen más alta que la imagen de la especificación que se va a mostrar con respecto al espectador, en base a un conjunto de imágenes capturadas y de la información del punto de vista virtual. En el presente documento, la imagen en directo se genera en base a la especificación del punto de vista virtual por parte del operador. Específicamente, la imagen en directo es la imagen de punto de vista virtual correspondiente al punto de vista virtual decidido según la operación de especificación por parte del operador con respecto a la imagen de la especificación. Además, el aparato de procesamiento de imágenes 1 también genera la imagen no en directo correspondiente a la imagen de punto de vista virtual que tiene una calidad de imagen más alta que la imagen en directo. A continuación, el aparato de procesamiento de imágenes 1 entrega la imagen en directo y la imagen no en directo generadas, al aparato de visualización 4, de tal manera que se muestra la imagen en directo antes de que se muestre la imagen no en directo. El aparato de procesamiento de imágenes 1 también entrega la imagen de la especificación generada al aparato de visualización 3, de tal manera que se muestra la imagen de la especificación en el aparato de visualización 3 antes de que se muestre la imagen en directo en el aparato de visualización 4.

Según esto, el aparato de visualización 4 puede mostrar la imagen de la especificación que tiene una calidad de imagen baja, la imagen en directo que se transmitirá en directo que tiene una calidad de imagen más alta que la imagen de la especificación, y la imagen no en directo que se transmitirá después de la grabación que tiene la calidad de imagen aún mayor que la imagen en directo. Cabe señalar que el aparato de visualización 4 también puede mostrar solo una imagen en directo y una imagen no en directo, y en el caso descrito anteriormente, el aparato de procesamiento de imágenes 1 entrega la imagen de punto de vista virtual adecuada para el aparato de visualización 4. Además, el aparato de visualización 3 puede mostrar los tres tipos de imágenes de punto de vista virtual, incluyendo la imagen de punto de vista virtual que tiene la calidad de imagen baja como imagen de la especificación, la imagen de punto de vista virtual que tiene la calidad de imagen media como la imagen en directo, y la imagen de punto de vista virtual que tiene la calidad de imagen alta como la imagen no en directo. Cabe señalar que también es suficiente cuando el aparato de visualización 3 no muestra por lo menos alguna de la imagen en directo y la imagen no en directo.

Es decir, el aparato de procesamiento de imágenes 1 entrega la imagen de la especificación al aparato de visualización 3 para que el usuario especifique el punto de vista virtual. A continuación, el aparato de procesamiento de imágenes 1 entrega por lo menos alguna de la imagen en directo y la imagen no en directo que tiene una calidad de imagen más alta que la imagen de la especificación, al aparato de visualización 4, para mostrar la imagen de punto de vista virtual generada en base a la especificación del punto de vista virtual por el usuario. Según esto, es posible responder tanto a los requisitos del operador que desea mostrar la imagen de punto de vista virtual con un retardo bajo para especificar el punto de vista virtual, como al espectador que desea observar la imagen de punto de vista virtual con una calidad de imagen alta.

Cabe señalar que, en el procesamiento descrito anteriormente, la imagen de punto de vista virtual se genera en base a las imágenes basadas en la captura mediante el grupo de cámaras 2 y en la información según la especificación del punto de vista virtual, y la imagen de punto de vista virtual que tiene la calidad de imagen alta se genera en base al resultado del procesamiento para la generación. Por esta razón, la magnitud total del procesamiento se puede reducir en comparación con un caso en el que la imagen de punto de vista virtual que tiene la calidad de imagen baja y la imagen de punto de vista virtual que tiene la calidad de imagen alta se generan respectivamente mediante procesos independientes. Cabe señalar, sin embargo, que la imagen de punto de vista virtual que tiene la calidad de imagen baja y la imagen de punto de vista virtual que tiene la calidad de imagen alta también pueden ser generadas mediante procesos independientes. Además, en un caso en el que la imagen de punto de vista virtual se muestra en una pantalla instalada en un recinto de competición o en un recinto de un concierto, o se difunde en directo, correspondiente a un caso en el que no es necesario difundir la imagen después de la grabación, el aparato de procesamiento de imágenes 1 no realiza el procesamiento para generar la imagen no en directo. Según esto, es posible reducir la magnitud del procesamiento para generar la imagen no en directo que tiene la calidad de imagen alta.

Además, el aparato de procesamiento de imágenes 1 puede generar una imagen de reproducción para ser mostrada después de la captura, en lugar de la imagen en directo para ser difundida en directo, o además de la imagen en directo. Por ejemplo, la imagen de reproducción se muestra en la pantalla del recinto de la competición durante el intermedio o después del final del partido, en un caso en el que el objetivo de la captura por parte del grupo de cámaras 2 es un partido, tal como fútbol, en el recinto de la competición. La imagen de reproducción tiene una calidad de imagen más alta que la imagen de la especificación y también se genera con una calidad de imagen tal que la generación puede ser completada hasta el final del partido o hasta el intermedio, para ser mostrada.

A continuación, se describirá otro modo de funcionamiento del aparato de procesamiento de imágenes 1 utilizando la figura 5. Según el modo de funcionamiento descrito anteriormente utilizando la figura 3, la imagen de punto de vista virtual que tiene la calidad de imagen alta es generada adicionalmente realizando un procesamiento de un nuevo tipo después de generar la imagen de punto de vista virtual que tiene la calidad de imagen baja. Por otra parte, en el modo de funcionamiento que se describirá a continuación utilizando la figura 5, el aumento de la calidad de imagen de la imagen de punto de vista virtual se realiza aumentando el número de cámaras utilizadas para generar la imagen de punto de vista virtual. En la siguiente explicación, se omitirán las descripciones de la parte similar al procesamiento en la figura 3.

El procesamiento mostrado en la figura 5 se inicia en el tiempo en que la unidad de obtención de puntos de vista 110 realiza la aceptación de la instrucción de generación de la imagen de punto de vista virtual. Cabe señalar, sin embargo, que el tiempo de inicio del procesamiento de la figura 5 no está limitado a esto. En S2010 y S2020, el aparato de procesamiento de imágenes 1 obtiene las imágenes capturadas por las respectivas cámaras del grupo de cámaras 2 y la información del punto de vista virtual mediante un procesamiento similar al descrito en la figura 3.

En S4030, la unidad de generación de imágenes 120 establece el número de cámaras correspondiente a las imágenes capturadas utilizadas para la generación de la imagen de punto de vista virtual. En el presente documento, la unidad de generación de imágenes 120 establece el número de cámaras de tal manera que el procesamiento en S4050 a S4070 se completa en un tiempo de procesamiento menor o igual a un umbral predeterminado (por ejemplo, un tiempo correspondiente a un fotograma en un caso en el que la imagen de punto de vista virtual es una imagen en movimiento). Por ejemplo, se supone que el procesamiento en S4050 a S4070 se ejecuta utilizando las imágenes capturadas de 100 cámaras por adelantado, y el tiempo de procesamiento es de 0,5 segundos. En este caso, cuando se desea que el procesamiento en S4050 a S4070 se complete en 0,016 segundos correspondientes a un fotograma de la imagen de punto de vista virtual en el que la velocidad de fotogramas es de 60 fps (fotograma por segundo), se establece el número de cámaras como 3.

Cabe señalar que, en un caso en el que la continuación de la generación de imágenes se determina en S4080 después de que el procesamiento de S4050 a S4070 entrega la imagen de punto de vista virtual, el flujo vuelve a S4030, y el número de cámaras se ajusta de nuevo. En el presente documento, un tiempo de procesamiento permisible se prolonga de tal manera que se genere la imagen de punto de vista virtual que tiene una calidad de imagen más alta que la imagen de punto de vista virtual previamente entregada, y el número de cámaras se incrementa en consecuencia. Por ejemplo, el número de cámaras a utilizar correspondientes a las imágenes capturadas se ajusta a 20, de tal manera que el procesamiento en S4050 a S4070 se completará en un tiempo de procesamiento menor o igual a 0,1 segundos.

En S4040, la unidad de generación de imágenes 120 selecciona las cámaras correspondientes a las imágenes capturadas a utilizar del grupo de cámaras 2 según el número de cámaras establecido en S4030 para generar la imagen de punto de vista virtual. Por ejemplo, en un caso en el que se seleccionan 3 cámaras de entre 100 cámaras, se selecciona la cámara más cercana al punto de vista virtual, y la cámara 34 y la cámara 67 contadas desde la cámara.

Además, después de generar la imagen de punto de vista virtual una vez, en un caso en el que el procesamiento se realiza en el segundo tiempo aumentando el número de imágenes capturadas a utilizar, puesto que la precisión del modelo de forma estimado en el procesamiento en el primer tiempo se incrementa más, se selecciona una cámara distinta a la cámara seleccionada en el primer procesamiento. Específicamente, en un caso en el que se seleccionan 20 cámaras de entre 100 cámaras, se selecciona primero la cámara más cercana al punto de vista virtual de entre las cámaras que no son seleccionadas en el primer procesamiento, y se selecciona la cámara en intervalos de cada cinco cámaras. En este momento, se salta la cámara ya seleccionada en el primer procesamiento y se selecciona la cámara siguiente. Cabe señalar que, por ejemplo, en un caso en el que se genera la imagen de punto de vista virtual que tiene la calidad de imagen máxima como imagen no en directo, se seleccionan todas las cámaras incluidas en el grupo de cámaras 2, y se ejecuta el procesamiento en S4050 a S4070 utilizando las imágenes capturadas de las respectivas cámaras.

Cabe señalar que un procedimiento de selección de las cámaras correspondientes a las imágenes capturadas a utilizar no está limitado a esto. Por ejemplo, se puede priorizar la selección de la cámara cercana al punto de vista virtual. En este caso, disminuye la precisión de la estimación de la forma de la zona posterior que no se ve desde el punto de vista virtual en la estimación de la forma del objeto correspondiente al sujeto, pero mejora la precisión de la estimación de la forma de la zona del lado delantero, que se ve desde el punto de vista virtual. Es decir, la calidad de imagen en la zona fácilmente observable por el espectador en la imagen de punto de vista virtual puede ser mejorada preferentemente.

En S4050, la unidad de generación de imágenes 120 ejecuta el procesamiento de estimación de la forma del objeto utilizando la imagen capturada por la cámara que se selecciona en S4040. El procesamiento en este momento es, por ejemplo, una combinación del procesamiento en S2030 en la figura 3 (Visual Hull) y el procesamiento en S2070 (Photo Hull). El procesamiento mediante Visual Hull incluye el procesamiento para calcular un producto lógico de volúmenes visuales de la pluralidad de cámaras correspondientes a la pluralidad de imágenes capturadas a utilizar.

Además, el procesamiento de Photo Hull incluye el procesamiento para proyectar los respectivos puntos del modelo de forma sobre la pluralidad de imágenes capturadas y calcular la consistencia de los valores de píxel. Por esta razón, a medida que disminuye el número de cámaras correspondiente a las imágenes capturadas a utilizar, disminuye la precisión de la estimación de la forma y se acorta el tiempo de procesamiento.

En S4060, la unidad de generación de imágenes 120 ejecuta el procesamiento de reproducción. El procesamiento en este caso es similar al procesamiento en S2080 en la figura 3, e incluye el procesamiento de coloreado para el grupo de puntos 3D y el procesamiento de generación para la imagen del fondo. El procesamiento de coloreado para el grupo de puntos 3D y el procesamiento de generación para la imagen del fondo incluyen ambos el procesamiento para decidir el color mediante el cálculo utilizando los valores de píxel de los puntos correspondientes a la pluralidad de imágenes capturadas. Por este motivo, al ser menor el número de cámaras correspondiente a las imágenes capturadas a utilizar, se disminuye la precisión de la representación y se acorta el tiempo de procesamiento.

En S4070, la unidad de salida 130 entrega la imagen de punto de vista virtual generada por la unidad de generación de imágenes 120 en S4060 al aparato de visualización 3 o al aparato de visualización 4.

En S4080, la unidad de generación de imágenes 120 determina si se realiza o no el procesamiento para generar la imagen de punto de vista virtual que tiene una calidad de imagen más alta que la imagen de punto de vista virtual generada en S4060. Por ejemplo, en un caso en el que la imagen de punto de vista virtual generada en S4060 es la imagen para que el operador especifique el punto de vista virtual y la imagen en directo se genera posteriormente, el flujo vuelve a S4030 y se genera la imagen de punto de vista virtual como la imagen en directo, aumentando el número de cámaras utilizadas. Además, en un caso en el que se genera adicionalmente la imagen no en directo después de generar la imagen en directo, y se genera la imagen de punto de vista virtual como imagen no en directo, aumentando más el número de cámaras. Es decir, puesto que el número de cámaras correspondiente a las imágenes capturadas utilizadas para la generación de la imagen de punto de vista virtual como la imagen en directo es mayor que el número de cámaras correspondiente a las imágenes capturadas utilizadas para la generación de la imagen de punto de vista virtual como la imagen de la especificación, la imagen en directo tiene una calidad de imagen más alta que la imagen de la especificación. De manera similar, puesto que el número de cámaras correspondiente a las imágenes capturadas utilizadas para la generación de la imagen de punto de vista virtual como imagen no en directo es mayor que el número de cámaras correspondiente a las imágenes capturadas utilizadas para la generación de la imagen de punto de vista virtual como la imagen en directo, la imagen no en directo tiene la calidad de imagen más alta que la imagen en directo.

Cabe señalar que, en S4080, en un caso en el que se determina que la imagen de punto de vista virtual que tiene la calidad de imagen más alta que la imagen de punto de vista virtual ya generada no necesita ser generada, o en un caso en el que se determina que la imagen de punto de vista virtual que tiene la calidad de imagen más alta no puede ser generada, el procesamiento finaliza.

Mediante el procesamiento descrito anteriormente, el aparato de procesamiento de imágenes 1 puede generar la pluralidad de imágenes de punto de vista virtual en las que la calidad de imagen se mejora gradualmente en los tiempos apropiados respectivos para su entrega. Por ejemplo, es posible generar la imagen de la especificación con poco retardo limitando el número de cámaras a utilizar para la generación de la imagen de punto de vista virtual a un número tal que el procesamiento de generación pueda ser completado en el tiempo de procesamiento establecido. Además, en un caso en el que se van a generar la imagen en directo y la imagen no en directo, es posible generar imágenes de mayor calidad de imagen cuando el procesamiento de generación se realiza aumentando el número de cámaras utilizadas.

A continuación, se describirá otro modo de funcionamiento del aparato de procesamiento de imágenes 1 utilizando la figura 6. Según el modo de funcionamiento descrito anteriormente utilizando la figura 5, cuando se incrementa el número de cámaras utilizadas para generar la imagen de punto de vista virtual, se realiza el aumento en la calidad de imagen de la imagen de punto de vista virtual. Por otra parte, según el modo de funcionamiento que se describirá a continuación utilizando la figura 6, el aumento de la calidad de imagen de la imagen de punto de vista virtual se realiza aumentando gradualmente la resolución de la imagen de punto de vista virtual. En la siguiente explicación, las descripciones de la parte similar al procesamiento en la figura 3 o la figura 5 serán omitidas. Cabe señalar que, según el modo de funcionamiento que se describirá a continuación, el número de píxeles de la imagen de punto de vista virtual a generar se establece regularmente en 4K (3840 x 2160), y la resolución de la imagen de punto de vista virtual es controlada dependiendo de si el cálculo del valor de píxel se realiza para cada bloque de píxeles grande o cada bloque de píxeles pequeño. Cabe señalar, sin embargo, que la configuración no está limitada a esto, y la resolución puede ser controlada cambiando el número de píxeles de la imagen de punto de vista virtual que se va a generar.

El procesamiento mostrado en la figura 6 se inicia en el tiempo en que la unidad de obtención de puntos de vista 110 realiza la aceptación de la instrucción de generación de la imagen de punto de vista virtual. Cabe señalar, sin embargo, que el tiempo de inicio del procesamiento de la figura 6 no está limitado a esto. En S2010 y S2020, el

aparato de procesamiento de imágenes 1 obtiene las imágenes capturadas por las respectivas cámaras del grupo de cámaras 2 y la información del punto de vista virtual mediante un procesamiento similar al descrito en la figura 3.

En S5030, la unidad de generación de imágenes 120 establece una resolución de la imagen de punto de vista virtual a generar. En el presente documento, la unidad de generación de imágenes 120 establece una resolución tal que el procesamiento en S5050 y S4070 se completa en un tiempo de procesamiento menor o igual que un umbral predeterminado. Por ejemplo, se supone que se ejecuta el procesamiento en S5050 y S4070 en un caso en el que la imagen de punto de vista virtual que tiene la resolución 4K es generada previamente, y el tiempo de procesamiento se ajusta a 0,5 segundos. En este caso, cuando se desea que el procesamiento en S5050 y S4070 se complete en 0,016 segundos correspondientes a un fotograma de la imagen de punto de vista virtual en que la velocidad de fotogramas es de 60 fps, la resolución tiene que ser de $0,016/0,5 = 1/31,25$ veces la de 4K o inferior. En vista de lo anterior, cuando las resoluciones vertical y horizontal de la imagen de punto de vista virtual se ajustan respectivamente a 1/8 veces la resolución 4K, el número de bloques de píxeles cuyo valor de píxel se calculará se convierte en 1/64, y el procesamiento se puede completar en menos de 0,016 segundos.

Cabe señalar que, en un caso en el que se determina en S4080 que la generación de imágenes continúa después de que el procesamiento en S5050 y S4070 entrega la imagen de punto de vista virtual, el flujo vuelve a S5030 y la resolución se ajusta nuevamente. En este caso, el tiempo de procesamiento permisible se prolonga de tal manera que se genera la imagen de punto de vista virtual que tiene una calidad de imagen más alta que la imagen de punto de vista virtual emitida anteriormente, y la resolución aumenta de acuerdo con esto. Por ejemplo, cuando las resoluciones vertical y horizontal se ajustan respectivamente a 1/4 de la resolución 4K, el procesamiento en S5050 y S4070 se puede completar en un tiempo de procesamiento menor o igual a 0,1 segundos. En S5040, la unidad de generación de imágenes 120 decide la posición del píxel en la que se calculará el valor de píxel en la imagen de punto de vista virtual según la resolución establecida en S5030. Por ejemplo, en un caso en el que la resolución de la imagen de punto de vista virtual se ajusta a 1/8 de la resolución 4K, los valores de píxel se calculan respectivamente para cada ocho píxeles vertical y horizontalmente. A continuación, se establece el mismo valor de píxel que el píxel (x, y) para los píxeles existentes entre el píxel (x, y) y el píxel (x + 8, y + 8), en los que se calcula el valor de píxel.

Además, después de generar una vez la imagen de punto de vista virtual, en el caso de que el procesamiento en el segundo tiempo se realice aumentando la resolución, el valor de píxel se calcula omitiendo el píxel en el que se calcula el valor de píxel en el primer tiempo. Por ejemplo, en un caso en el que la resolución se ajusta a 1/4 de la resolución 4K, se calcula el valor de píxel del píxel (x + 4, y + 4), y se establece el mismo valor de píxel que el píxel (x + 4, y + 4) para los píxeles existentes entre el píxel (x + 4, y + 4) y el píxel (x + 8, y + 8). De esta manera, cuando se aumenta el número de píxeles en los que se calcula el valor de píxel, la resolución de la imagen de punto de vista virtual se puede aumentar hasta la resolución 4K como máximo.

En S5050, la unidad de generación de imágenes 120 realiza el procesamiento de coloreado para la imagen de punto de vista virtual calculando el valor de píxel del píxel en la posición decidida en S5040. Como procedimiento de cálculo para el valor de píxel, por ejemplo, se puede utilizar un procedimiento de Visual Hull basado en imágenes. Puesto que el valor de píxel se calcula para cada píxel según este procedimiento, a medida que el número de píxeles en los que se calcula el valor de píxel es menor, es decir, a medida que la resolución de la imagen de punto de vista virtual es menor, el tiempo de procesamiento se acorta.

En S4070, la unidad de salida 130 entrega la imagen de punto de vista virtual generada por la unidad de generación de imágenes 120 en S5050 al aparato de visualización 3 o al aparato de visualización 4.

En S4080, la unidad de generación de imágenes 120 determina si se realiza o no el procesamiento para generar la imagen de punto de vista virtual que tiene una calidad de imagen superior a la de la imagen de punto de vista virtual generada en S5050. Por ejemplo, en un caso en el que la imagen de punto de vista virtual generada en S5050 es la imagen para que el operador especifique el punto de vista virtual y la imagen en directo se genera posteriormente, el flujo vuelve a S5030 y se genera la imagen de punto de vista virtual que tiene una resolución mayor. Además, después de que se genere la imagen en directo y se va a generar además la imagen no en directo, se genera la imagen de punto de vista virtual como la imagen no en directo en la que se aumenta más la resolución. Es decir, puesto que la imagen de punto de vista virtual como la imagen en directo tiene la resolución más alta que la imagen de punto de vista virtual como la imagen de la especificación, la imagen en directo tiene la calidad de imagen más alta que la imagen de la especificación. De manera similar, puesto que la imagen de punto de vista virtual como imagen no en directo tiene la resolución más alta que la imagen de punto de vista virtual como la imagen en directo, la imagen no en directo tiene la calidad de imagen más alta que la imagen en directo.

Cabe señalar que, en S4080, en un caso en el que se determina que no es necesario generar la imagen de punto de vista virtual que tiene la calidad de imagen más alta que la imagen de punto de vista virtual ya generada, o en un caso en el que se determina que la imagen de punto de vista virtual que tiene la calidad de imagen más alta no se puede generar, el procesamiento finaliza.

Mediante el procesamiento descrito anteriormente, el aparato de procesamiento de imágenes 1 puede generar la pluralidad de imágenes de punto de vista virtual en las que la resolución se mejora gradualmente en los tiempos apropiados respectivos para ser entregadas. Por ejemplo, es posible generar la imagen de la especificación con poco retardo ajustando la resolución de la imagen de punto de vista virtual a una resolución tal que el procesamiento de generación pueda ser completado en el tiempo de procesamiento establecido. Además, en un caso en el que se van a generar la imagen en directo y la imagen no en directo, es posible generar imágenes de mayor calidad de imagen cuando se realiza el procesamiento de generación aumentando la resolución.

Tal como se ha descrito anteriormente, el aparato de procesamiento de imágenes 1 realiza el procesamiento de imágenes para mejorar la calidad de imagen de la imagen de punto de vista virtual para generar la imagen que tiene la calidad de imagen alta (por ejemplo, la imagen no en directo). El aparato de procesamiento de imágenes 1 también genera la imagen que tiene la calidad de imagen baja (por ejemplo, la imagen en directo) mediante el procesamiento ejecutado en un tiempo de procesamiento menor o igual al umbral predeterminado correspondiente al procesamiento parcial incluido en el procesamiento de imágenes. Según esto, la imagen de punto de vista virtual se mostrará con un retardo menor o igual que el tiempo predeterminado y se puede generar y mostrar la imagen de punto de vista virtual que tiene la calidad de imagen alta.

Cabe señalar que, en la explicación de la figura 6, se supone que se estima el parámetro de generación (resolución) para completar el procesamiento de generación en el tiempo de procesamiento menor o igual al umbral predeterminado, y la imagen de punto de vista virtual se genera mediante el parámetro de generación estimado. Sin embargo, se debe tener en cuenta que la configuración no está limitada a esto, y el aparato de procesamiento de imágenes 1 puede mejorar gradualmente la calidad de imagen de la imagen de punto de vista virtual, y entregar la imagen de punto de vista virtual ya generada, en un momento en que el tiempo de procesamiento alcanza un umbral predeterminado. Por ejemplo, en un caso en el que la imagen de punto de vista virtual en la que la resolución es 1/8 de la resolución 4K ya está generada y la imagen de punto de vista virtual en la que la resolución es 1/4 de la resolución 4K no se ha completado en el momento en que el tiempo de procesamiento alcanza el umbral predeterminado, se puede entregar la imagen de punto de vista virtual en la que la resolución es 1/8. Además, se puede entregar la imagen de punto de vista virtual en la que se realiza el procesamiento para mejorar la resolución de 1/8 de resolución a 1/4 de resolución en la mitad del proceso.

Según la presente realización, se ha descrito principalmente el caso en el que la unidad de generación de imágenes 120 incluida en el aparato de procesamiento de imágenes 1 controla la generación de la imagen de punto de vista virtual en base a la imagen obtenida por la unidad de obtención de información de la cámara 100, y la información de punto de vista virtual obtenida por la unidad de obtención de puntos de vista 110, y genera la pluralidad de imágenes de punto de vista virtual que tienen las calidades de imagen diferentes. Cabe señalar, sin embargo, que la configuración no está limitada a esto, y la función para controlar la generación de la imagen de punto de vista virtual y la función para generar realmente la imagen de punto de vista virtual pueden estar incluidas en aparatos respectivamente diferentes.

Por ejemplo, un aparato de generación (no mostrado) que tiene la función de la unidad de generación de imágenes 120 y genera la imagen de punto de vista virtual también puede existir en el sistema de procesamiento de imágenes 10. Por lo tanto, el aparato de procesamiento de imágenes 1 puede controlar la generación de la imagen de punto de vista virtual mediante el aparato de generación en base a la imagen obtenida por la unidad de obtención de información de la cámara 100 y la información obtenida por la unidad de obtención de puntos de vista 110. Específicamente, el aparato de procesamiento de imágenes 1 transmite las imágenes capturadas y la información del punto de vista virtual al aparato de generación, y realiza la instrucción para controlar la generación de la imagen de punto de vista virtual. A continuación, el aparato de generación genera una primera imagen de punto de vista virtual y una segunda imagen de punto de vista virtual que se visualizará en un tiempo anterior a la visualización de la primera imagen de punto de vista virtual, teniendo la segunda imagen de punto de vista virtual una calidad de imagen inferior a la primera imagen de punto de vista virtual, en base a las imágenes capturadas recibidas y la información de punto de vista virtual. En este caso, la primera imagen de punto de vista virtual es, por ejemplo, la imagen no en directo, y la segunda imagen de punto de vista virtual es, por ejemplo, la imagen en directo. Sin embargo, cabe señalar que el propósito de utilización para la primera imagen de punto de vista virtual y la segunda imagen de punto de vista virtual no está limitado a esto. Cabe señalar que el aparato de procesamiento de imágenes 1 puede realizar el control de tal manera que la primera imagen de punto de vista virtual y la segunda imagen de punto de vista virtual sean generadas por los aparatos de generación respectivamente diferentes. Además, el aparato de procesamiento de imágenes 1 puede realizar un control de salida para controlar el destino de salida de la imagen de punto de vista virtual mediante el aparato de generación y la temporización de salida y similares.

Además, el aparato de generación puede incluir las funciones de la unidad de obtención de puntos de vista 110 y la unidad de generación de imágenes 120, y el aparato de procesamiento de imágenes 1 puede controlar la generación de la imagen de punto de vista virtual por el aparato de generación en base a las imágenes obtenidas por la unidad de obtención de información de la cámara 100. En este caso, las imágenes obtenidas por la unidad de obtención de información de la cámara 100 son imágenes basadas en la captura tales como las imágenes capturadas, capturadas por el grupo de cámaras 2, y las imágenes generadas en base a la diferencia entre la pluralidad de imágenes capturadas. Además, el aparato de generación puede incluir las funciones de la unidad de obtención de información

de la cámara 100 y la unidad de generación de imágenes 120, y el aparato de procesamiento de imágenes 1 puede controlar la generación de la imagen de punto de vista virtual por el aparato de generación. en base a las imágenes obtenidas por la unidad de obtención de puntos de vista 110. En este caso, las imágenes obtenidas por la unidad de obtención de puntos de vista 110 son la información según la especificación del punto de vista virtual, tal como la información que indica los contenidos determinados según el punto de vista virtual, como la forma o la orientación del sujeto en la imagen de punto de vista virtual y la información del punto de vista virtual. Es decir, el aparato de procesamiento de imágenes 1 puede obtener la información relacionada con la generación de la imagen de punto de vista virtual que incluye por lo menos una de las imágenes basadas en la captura y la información según la especificación del punto de vista virtual, y controlar la generación de la imagen de punto de vista virtual en base a la información obtenida.

Además, por ejemplo, el aparato de generación que existe en el sistema de procesamiento de imágenes 10 puede incluir las funciones de la unidad de obtención de información de la cámara 100, la unidad de obtención de puntos de vista 110 y la unidad de generación de imágenes 120, y el aparato de procesamiento de imágenes 1 puede controlar la generación de la imagen de punto de vista virtual por el aparato de generación en base a la información relacionada con la generación de la imagen de punto de vista virtual. La información relacionada con la generación de la imagen de punto de vista virtual en este caso incluye, por ejemplo, por lo menos cualquiera de los parámetros con respecto a la calidad de imagen de la primera imagen de punto de vista virtual y los parámetros con respecto a la calidad de imagen de la segunda imagen de punto de vista virtual que son generados por el aparato de generación. Ejemplos específicos de los parámetros con respecto a la calidad de imagen incluyen el número de cámaras correspondiente a las imágenes capturadas utilizadas para la generación de la imagen de punto de vista virtual, la resolución de la imagen de punto de vista virtual, un tiempo permisible como el tiempo de procesamiento relacionado con la generación de la imagen de punto de vista virtual, y similares. El aparato de procesamiento de imágenes 1 obtiene estos parámetros con respecto a la calidad de imagen en base a la entrada del operador, por ejemplo, y controla el aparato de generación en base a los parámetros obtenidos mediante la transmisión de los parámetros al aparato de generación o similar. Según esto, el operador puede generar la pluralidad de imágenes de punto de vista virtual que tienen las calidades de imagen deseadas mutuamente diferentes.

Tal como se ha descrito anteriormente, el aparato de procesamiento de imágenes 1 acepta la instrucción de generación de la imagen de punto de vista virtual en base a las imágenes basadas en la captura del sujeto desde las direcciones respectivamente diferentes, por la pluralidad de cámaras, y la información según la especificación del punto de vista virtual. A continuación, el aparato de procesamiento de imágenes 1 realiza el control según la aceptación de la instrucción de generación, de tal manera que la primera imagen de punto de vista virtual a entregar a un primer aparato de visualización y la segunda imagen de punto de vista virtual a entregar a un segundo aparato de visualización son generadas a partir de las imágenes en base a la captura y la información según la especificación del punto de vista virtual. En este caso, la segunda imagen de punto de vista virtual es una imagen de punto de vista virtual que tiene la calidad de imagen más alta que la primera imagen de punto de vista virtual. Según esto, por ejemplo, también en el caso de que existan tanto el usuario que desea observar la imagen de punto de vista virtual en tiempo real como el usuario que prioriza la calidad de imagen alta de la imagen de punto de vista virtual sobre la propiedad de ser en tiempo real, es posible generar la imagen de punto de vista virtual adecuada para el tiempo en que se va a realizar la visualización.

Cabe señalar que, según la presente realización, se ha descrito el caso en el que la gradación de color, la resolución y el número de cámaras correspondiente a las imágenes capturadas utilizadas para la generación de la imagen de punto de vista virtual son controladas como la calidad de imagen de la imagen de punto de vista virtual, pero se pueden controlar otros parámetros tales como la calidad de imagen. Además, se pueden controlar al mismo tiempo una pluralidad de parámetros con respecto a la calidad de imagen.

La presente invención también se puede realizar mediante el procesamiento en el que un programa que realiza una o varias funciones de las realizaciones descritas anteriormente se suministra a un sistema o aparato a través de una red o un medio de almacenamiento, y uno o varios procesadores en un ordenador del sistema o el aparato lee y ejecuta el programa. Además, la presente invención se puede realizar mediante un circuito (por ejemplo, un ASIC o similar) que realiza una o varias funciones. Además, el programa puede estar grabado en un medio de grabación legible por ordenador, que se proporcionará.

La presente invención no está limitada a las realizaciones descritas anteriormente. Por supuesto, se comprenderá que esta invención se ha descrito anteriormente solo a modo de ejemplo, y que se pueden realizar modificaciones de detalle dentro del alcance de esta invención, que está definido por las reivindicaciones adjuntas.

REIVINDICACIONES

1. Sistema de procesamiento de imágenes, que comprende:

- 5 un medio de obtención de imágenes (100) para obtener imágenes capturadas desde una pluralidad de direcciones por una pluralidad de cámaras (2);
un medio de obtención de información (110) para obtener información de punto de vista que indica un punto de vista virtual; y
un medio de generación (120) para generar imágenes de punto de vista virtual en base a las imágenes obtenidas
10 por el medio de obtención de imágenes y a la información de puntos de vista obtenida por el medio de obtención de información,
en el que el medio de generación está configurado para generar una primera imagen de punto de vista virtual para ser entregada a un aparato de visualización (3) utilizado por un usuario para especificar un punto de vista virtual y para generar una segunda imagen de punto de vista virtual para ser entregada a un destino de salida diferente del
15 aparato de visualización (3) utilizando por lo menos uno de la primera imagen de punto de vista virtual y los datos generados en un proceso para generar la primera imagen de punto de vista virtual mediante el procesamiento de imágenes utilizando la pluralidad de imágenes obtenidas por el medio de obtención de imágenes, teniendo la segunda imagen de punto de vista virtual una calidad de imagen superior a la de la primera imagen de punto de vista virtual.

20 2. Sistema de procesamiento de imágenes, según la reivindicación 1,
en el que la segunda imagen de punto de vista virtual es entregada a por lo menos otro aparato de visualización (4) diferente del aparato de visualización (3) que muestra la imagen para que el usuario especifique el punto de vista virtual y un dispositivo de almacenamiento.

25 3. Sistema de procesamiento de imágenes, según la reivindicación 1 o 2, que comprende, además:

un medio de salida (130) para entregar la primera imagen de punto de vista virtual y la segunda imagen de punto de vista virtual generadas por el medio de generación (120),
30 en el que el tiempo en el que la primera imagen de punto de vista virtual es entregada por el medio de salida es anterior al tiempo en el que la segunda imagen de punto de vista virtual es entregada por el medio de salida.

4. Sistema de procesamiento de imágenes, según cualquiera de las reivindicaciones 1 a 3,
35 en el que la información del punto de vista obtenida por el medio de obtención de información (110) indica el punto de vista virtual especificado por una operación del usuario según la visualización de la imagen en el aparato de visualización.

5. Sistema de procesamiento de imágenes, según cualquiera de las reivindicaciones 1 a 4,
40 en el que el medio de generación (120) está configurado para generar la segunda imagen de punto de vista virtual realizando un procesamiento de imágenes para mejorar la calidad de imagen con respecto a la primera imagen de punto de vista virtual generada en base a las imágenes obtenidas y a la información del punto de vista obtenida, y para generar la primera imagen de punto de vista virtual realizando un procesamiento ejecutado en un tiempo de procesamiento menor que o igual que un umbral, correspondiendo el procesamiento al procesamiento parcial incluido en el procesamiento para generar la segunda imagen de punto de vista virtual a partir de la primera imagen
45 de punto de vista virtual.

6. Sistema de procesamiento de imágenes, según cualquiera de las reivindicaciones 1 a 5,
50 en el que la primera imagen de punto de vista virtual es una imagen que representa una forma de un objeto capturada por, por lo menos una de la pluralidad de cámaras (2), y
en el que la segunda imagen de punto de vista virtual es una imagen que representa la forma del objeto y también un color del objeto que no está representado en la primera imagen de punto de vista virtual.

7. Sistema de procesamiento de imágenes, según cualquiera de las reivindicaciones 1 a 6,
55 en el que la calidad de imagen relativa a la imagen de punto de vista virtual generada por el medio de generación (120) es el número de gradaciones de un color incluido en la imagen de punto de vista virtual.

8. Sistema de procesamiento de imágenes, según cualquiera de las reivindicaciones 1 a 6,
60 en el que la calidad de imagen relativa a la imagen de punto de vista virtual generada por el medio de generación es una resolución de una imagen de punto de vista virtual.

9. Sistema de procesamiento de imágenes, según cualquiera de las reivindicaciones 1 a 8,
65 en el que el medio de generación (120) está configurado, además, para generar una tercera imagen de punto de vista virtual que tiene una calidad de imagen más alta que la de la segunda imagen de punto de vista virtual, en base a las imágenes obtenidas por el medio de obtención de imágenes y a la información del punto de vista obtenida del medio de obtención de información (110).

- 5 10. Sistema de procesamiento de imágenes, según cualquiera de las reivindicaciones 1 a 9, en el que el medio de generación (120) está configurado para generar la segunda imagen de punto de vista virtual utilizando las imágenes correspondientes a más cámaras que el número de cámaras correspondiente a las imágenes utilizadas para generar la primera imagen de punto de vista virtual.
- 10 11. Sistema de procesamiento de imágenes, según cualquiera de las reivindicaciones 1 a 10, en el que el medio de generación (120) está configurado para generar la segunda imagen de punto de vista virtual utilizando una magnitud de procesamiento mayor que la utilizada para generar la primera imagen de punto de vista virtual.
12. Procedimiento de procesamiento de imágenes, que comprende:
- 15 una etapa de obtención de imágenes, para obtener imágenes basadas en la captura de una pluralidad de direcciones por parte de una pluralidad de cámaras;
- una etapa de obtención de información para obtener información de punto de vista que indica un punto de vista virtual; y
- 20 una etapa de generación para generar imágenes de punto de vista virtual en base a las imágenes obtenidas en la etapa de obtención de imágenes y a la información de punto de vista obtenida en la etapa de obtención de información,
- en el que la etapa de generación incluye generar una primera imagen de punto de vista virtual a entregar a un aparato de visualización utilizado por un usuario para especificar el punto de vista virtual, y generar una segunda imagen de punto de vista virtual a entregar a un destino de salida diferente del aparato de visualización utilizando, por lo menos, uno de la primera imagen de punto de vista virtual y de los datos generados en un proceso para generar la primera imagen de punto de vista virtual mediante procesamiento de imágenes utilizando la pluralidad de
- 25 imágenes obtenidas en la etapa de obtención de imágenes, teniendo la segunda imagen de punto de vista virtual una calidad de imagen más alta que la primera imagen de punto de vista virtual.
13. Procedimiento de procesamiento de imágenes, según la reivindicación 12, en el que la segunda imagen de punto de vista virtual es una imagen de punto de vista virtual para ser mostrada a una audiencia diferente del usuario que especifica el punto de vista virtual.
- 30 14. Sistema de procesamiento de imágenes, según la reivindicación 12 o 13, que comprende:
- una etapa de entrega para entregar la primera imagen de punto de vista virtual y la segunda imagen de punto de vista virtual generadas en la etapa de generación,
- 35 en el que el tiempo en el que la primera imagen de punto de vista virtual es entregada en la etapa de entrega es anterior al tiempo en el que la segunda imagen de punto de vista virtual es entregada en la etapa de entrega.
15. Producto de programa informático que comprende instrucciones que, cuando el programa es ejecutado por un ordenador, hacen que el ordenador ejecute el procedimiento de procesamiento de imágenes, según cualquiera de las reivindicaciones 12 a 14.
- 40

FIG. 1

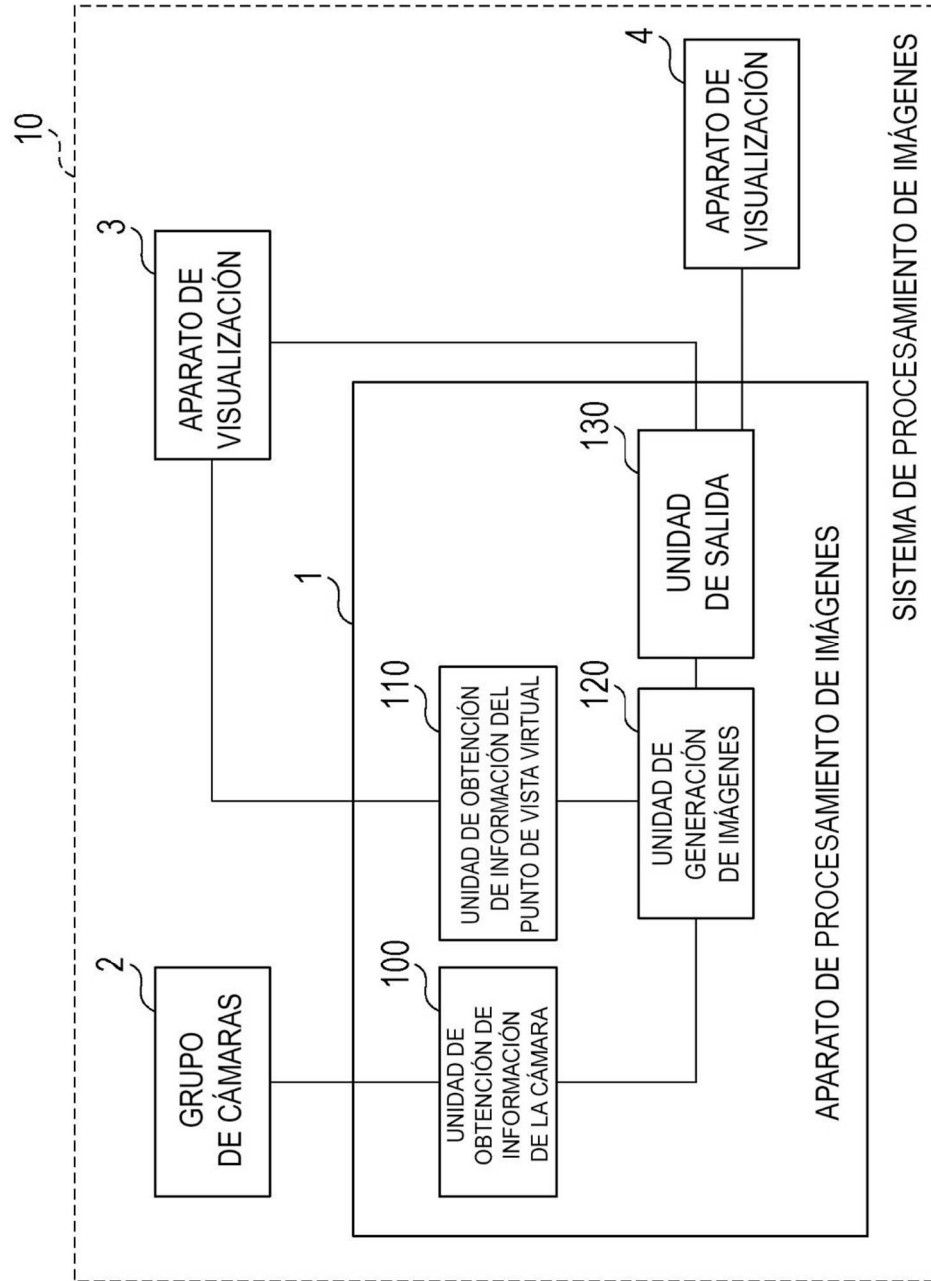


FIG. 2

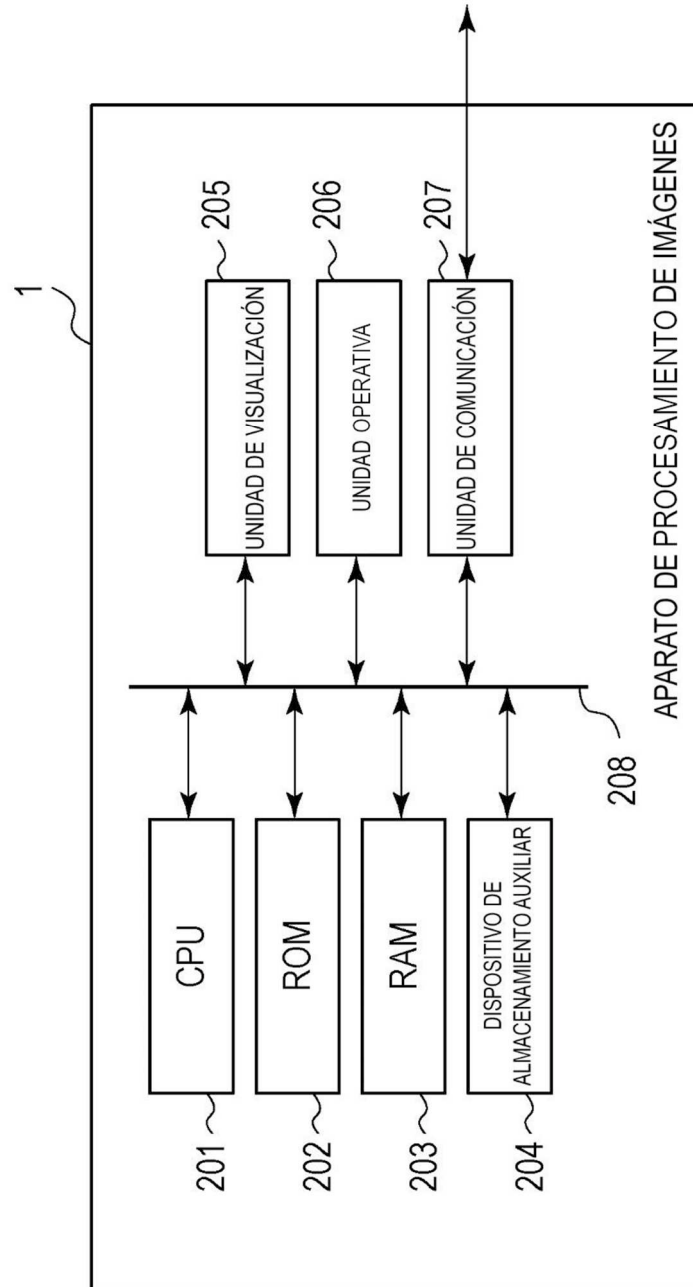


FIG. 3

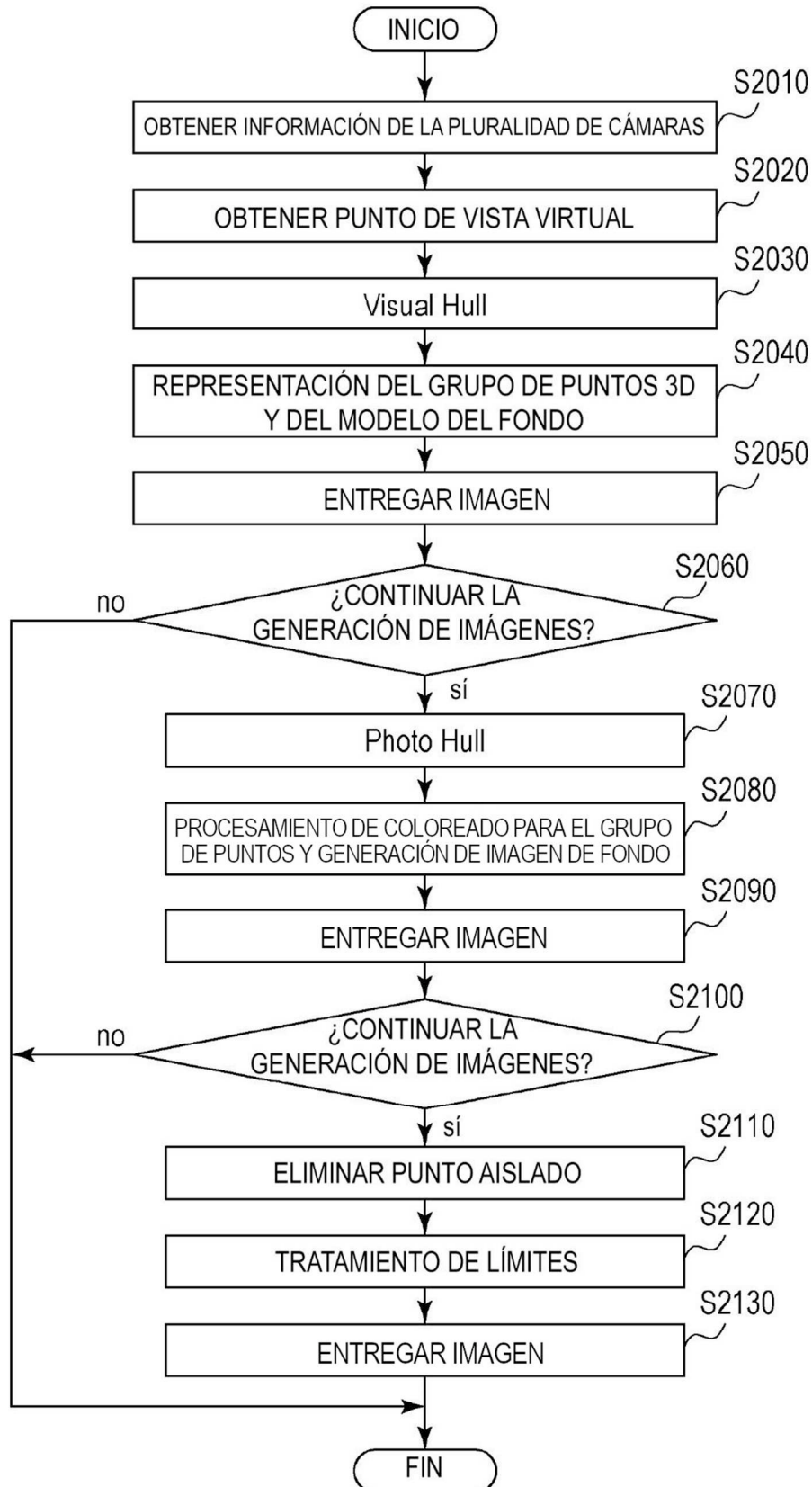


FIG. 4

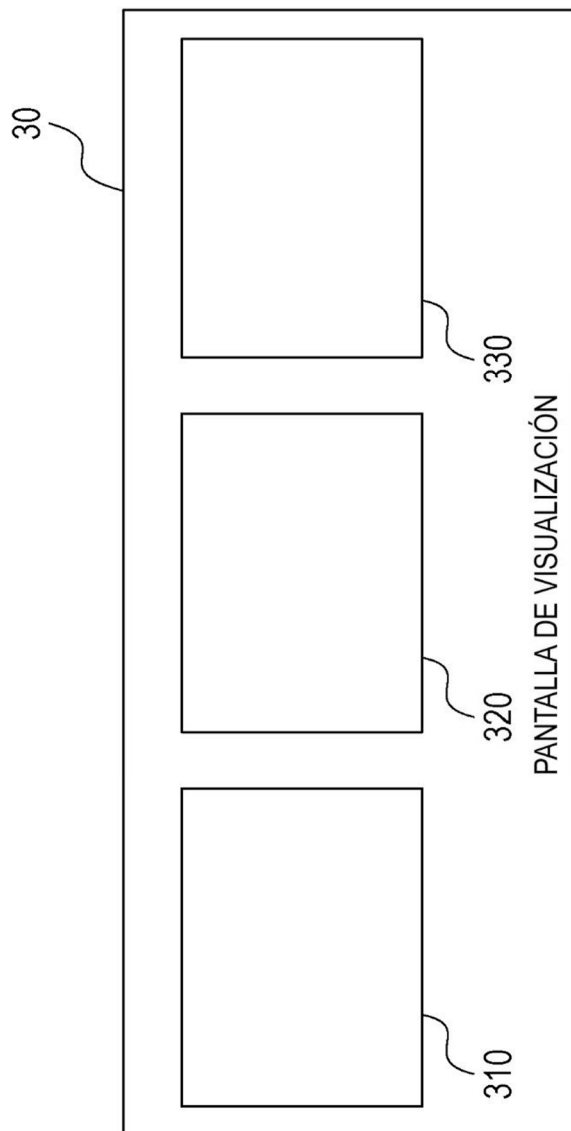


FIG. 5

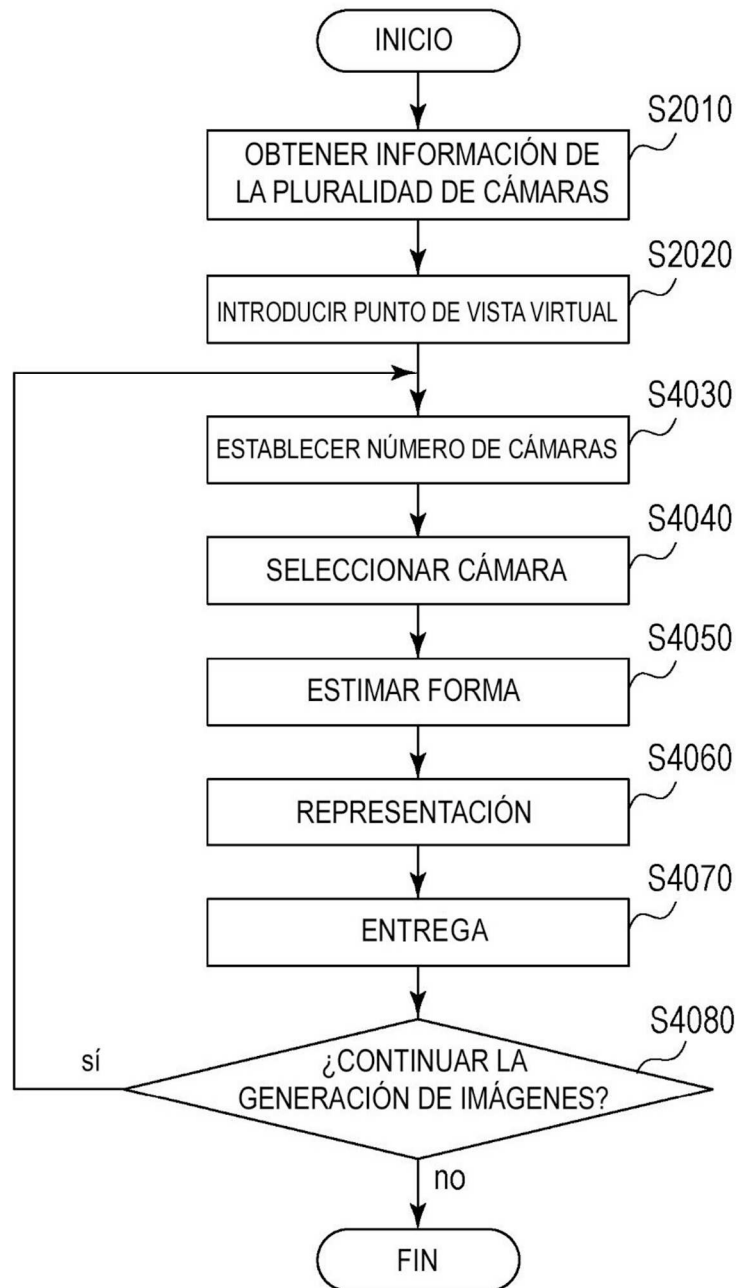
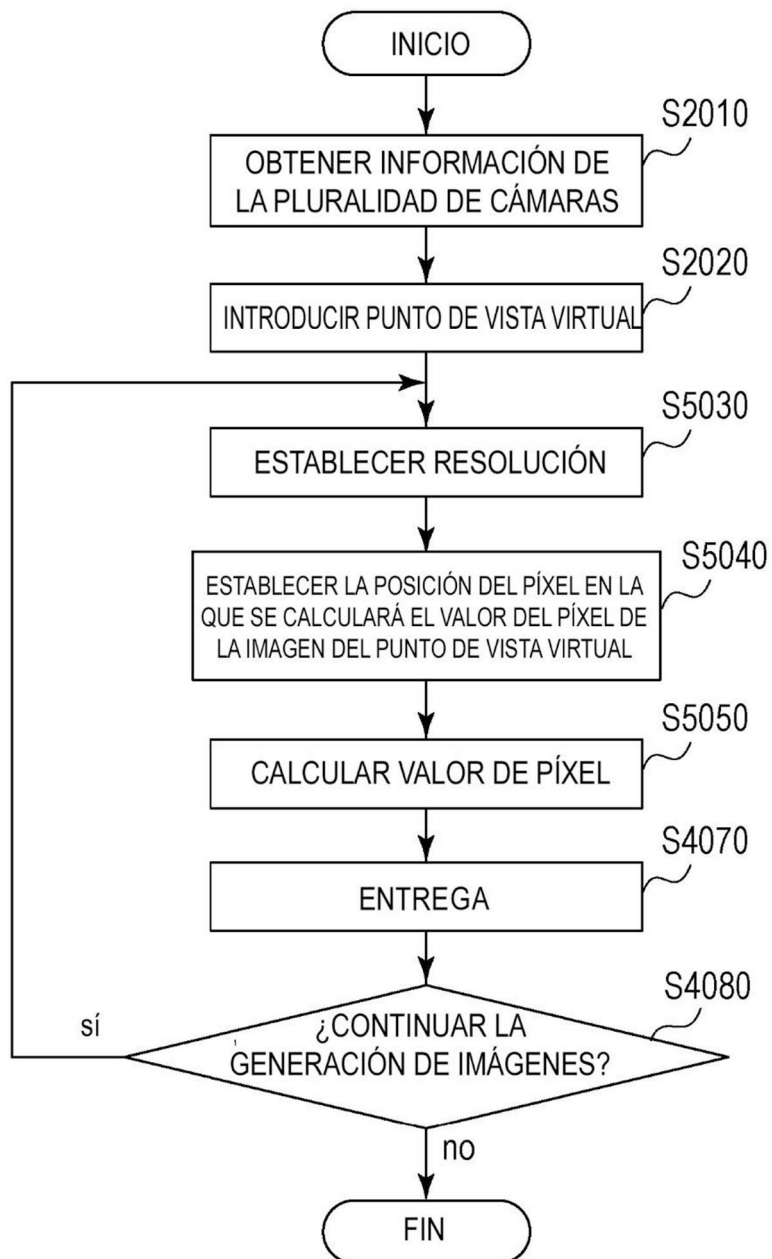


FIG. 6



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 2014270706 A
- US 2013342647 A
- WO 2004053591 A
- JP 2013223008 A