

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 713 521**

51 Int. Cl.:

G02B 27/01 (2006.01)

G06F 3/01 (2006.01)

G06T 19/00 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **31.10.2011 PCT/US2011/058604**

87 Fecha y número de publicación internacional: **18.05.2012 WO12064546**

96 Fecha de presentación y número de la solicitud europea: **31.10.2011 E 11840164 (5)**

97 Fecha y número de publicación de la concesión europea: **28.11.2018 EP 2638693**

54 Título: **Foco virtual variable automático para presentaciones de realidad aumentada**

30 Prioridad:

08.11.2010 US 941825

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

22.05.2019

73 Titular/es:

**MICROSOFT TECHNOLOGY LICENSING, LLC
(100.0%)
One Microsoft Way
Redmond, WA 98052, US**

72 Inventor/es:

**BAR-ZEEV, AVI y
LEWIS, JOHN**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 713 521 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

Foco virtual variable automático para presentaciones de realidad aumentada

Antecedentes

5 La realidad aumentada es una tecnología que permite que las imágenes virtuales sean mezcladas con un entorno físico del mundo real. Típicamente, sistemas de presentación o pantallas, transparentes, próximas a los ojos son llevadas por los usuarios para ver las imágenes mezcladas de objetos virtuales y reales. Las pantallas próximas a los ojos típicamente utilizan una combinación de elementos ópticos y la estereopsis para enfocar imágenes virtuales a una distancia fija. Sin embargo, con el objeto virtual a una distancia fija, y la ubicación del usuario cambiando, el objeto virtual no se mueve dentro y fuera del foco para un usuario como si el usuario viera el objeto virtual con luz natural. "A Novel Prototype for an Optical See-Through Head-Mounted Display with Addressable Focus Cues" por Sheng Liu et al, describe el uso de una lente líquida en una presentación montada en la cabeza tanto en el modo de plano focal variable en el que un señal de alojamiento es abordable, y un modo de plano multifocal de tiempo multiplexado en el que tanto las señales de alojamiento como de señales de borrosidad retinales pueden ser realizados.

Compendio

15 La invención proporciona un sistema de realidad aumentada y un método como el reivindicado, que incluye las realizaciones para conseguir focos variables para pantallas de realidad mezclada o de realidad aumentada. Un usuario ve una escena a través de un dispositivo de pantalla transparente. La escena incluye uno o más objetos reales en un entorno físico que el usuario observa directamente con sus ojos a través de lentes claras o transparentes del dispositivo de presentación. Uno o más objetos virtuales son proyectados por el dispositivo de presentación en al menos una de las ubicaciones del ojo del usuario. De este modo, la presentación de la escena real está aumentada con objetos virtuales.

20 En una realización, la tecnología proporciona variación de la distancia focal de los objetos virtuales proyectados por el dispositivo de presentación. Un campo de visión de tres dimensiones es determinado para el usuario, y las ubicaciones en tres dimensiones de uno o más objetos virtuales dentro del campo de visión del usuario son determinadas también. Una región focal en tres dimensiones actual del usuario está determinada dentro del campo de visión. Uno o más objetos virtuales en la región focal actual del usuario son identificados en base a sus ubicaciones.

25 Para presentar un objeto virtual en la región focal actual del usuario en la presentación, el objeto virtual es movido a una región de la imagen en la región focal actual. En una realización, esto se realiza cambiando la región focal de un conjunto de micropresentación del dispositivo de presentación transparente. Un conjunto de micropresentación puede incluir elementos tales como al menos un elemento óptico, una unidad de micropresentación alineada a lo largo de una trayectoria óptica, así como un ajustador de foco virtual variable. La unidad de micropresentación genera imágenes para la presentación, y la luz de las imágenes se desplaza a lo largo de la trayectoria óptica a través de al menos un elemento óptico, por ejemplo una lente de colimación, hasta un elemento reflectante.

30 El ajustador cambia el desplazamiento a lo largo de la trayectoria óptica entre al menos dos elementos del conjunto de micropresentación para cambiar la región focal de un objeto virtual en una imagen. En otra realización, adicionalmente una longitud focal de un elemento óptico puede ser ajustada para obtener la región focal deseada. Por ejemplo, la polarización de al menos una lente birrefringente puede ser cambiada o puede ser ajustado un radio de curvatura de una lente de fluido o líquida.

35 En una realización, una profundidad artificial de la técnica de campo puede ser aplicada a un objeto virtual fuera de la región focal del usuario pero dentro del campo de visión del usuario como una función de la distancia desde la región focal del usuario.

40 El sistema puede incluir además un generador de sincronización en los circuitos de control para controlar la sincronización del ajustador de foco virtual variable que barre el conjunto de micropresentación a través de diferentes regiones focales. En la mayoría de las realizaciones, la velocidad de barrido es ajustada para ser lo suficientemente rápida, de manera que la fusión de imagen temporal humana hará que las imágenes generadas en las diferentes regiones focales parezcan estar presentes a la vez. Un usuario que observa una presentación de imágenes generadas en diferentes regiones focales observa los sujetos virtuales de forma natural dentro y fuera del foco como si los objetos virtuales fueran objetos reales en el entorno físico.

45 La tecnología también puede proporcionar un método para generar objetos virtuales en diferentes regiones focales en una presentación de realidad aumentada. Un campo de visión de tres dimensiones está determinado para un usuario que ve una escena que incluye objetos reales y virtuales utilizando un dispositivo de presentación transparente. Una región focal de tres dimensiones está identificada para uno o más objetos virtuales dentro del campo de visión del usuario. El conjunto de micropresentación está ajustado para barrer a través de un cierto número de regiones focales. Las regiones focales pueden ser seleccionadas para incluir una región focal que el usuario está actualmente viendo y aquellas regiones que incluyen un objeto virtual. Una imagen es generada para

cada región focal.

En un ejemplo, cada una de estas imágenes puede ser presentada al usuario a una velocidad lo suficientemente rápida, de manera que la fusión de imagen temporal humana hace que parezcan estar presentes a la vez para el ojo humano. En otro ejemplo, es presentada una imagen compuesta, de partes en foco de imágenes virtuales generadas en diferentes regiones focales.

5

Este Compendio se proporciona para introducir una selección de conceptos de una manera simplificada que serán descritos con más detalle en la Descripción Detallada. Este Compendio no está destinado a identificar características claves o características esenciales de la materia objeto reivindicada, ni está destinado a ser utilizado como una ayuda en la determinación del alcance de la materia objeto reivindicada.

10 **Breve descripción de los dibujos**

La Fig. 1 es un diagrama de bloques que muestra componentes a modo de ejemplo de una realización del sistema para proporcionar foco variable de objetos virtuales.

La Fig. 2A es una vista superior de una parte de una realización de una unidad de presentación montada en la cabeza.

15 La Fig. 2B es una vista superior de una parte de otra realización de una unidad de presentación montada en la cabeza.

La Fig. 3A es un ejemplo de un sistema de lentes giratorias para utilizar como parte de un conjunto de micropresentación de presentación próxima al ojo.

20 La Fig. 3B1 y la Fig. 3B2 son ejemplos de una lente de fluido que presentaban diferentes radios de curvatura, siendo la lente de fluido para utilizar como parte de un conjunto de micropresentación.

La Fig. 3C es un ejemplo de un sistema de lente birrefringente para utilizar como parte de un conjunto de micropresentación de presentación próxima a los ojos.

La Fig. 3D es un ejemplo de un sistema de lente de para utilizar como parte de un conjunto de micropresentación.

25 La Fig. 4 es un diagrama de bloques de una realización de los componentes de la unidad de presentación montada en la cabeza.

En la Fig. 5 es un diagrama de bloques de una realización de los componentes de una unidad de procesamiento asociada con una unidad de presentación montada en la cabeza.

La Fig. 6 es un diagrama de bloques de una realización de los componentes de un sistema de ordenador concentrador utilizado con la unidad de presentación montada en la cabeza.

30 La Fig. 7 es un diagrama de bloques de una realización del sistema de ordenador que puede ser utilizado para implementar el sistema de ordenador concentrador descrito en la presente memoria.

La Fig. 8 es un diagrama de bloques que muestra un sistema de múltiples usuarios que puede variar el foco del contenido virtual para estar en foco para un usuario.

35 La Fig. 9 es un diagrama de flujo que describe una realización de un proceso para presentar un objeto virtual para que esté en foco cuando sea visto por el usuario.

La Fig. 10 es un diagrama de flujo que describe una realización de un proceso para crear un modelo de un espacio.

La Fig. 11 es un diagrama de flujo que describe una realización de un proceso para segmentar un modelo en objetos.

La Fig. 12 es un diagrama de flujo que el describe una realización de un proceso para identificar objetos.

40 La Fig. 13 es un diagrama de flujo que describe una realización de un proceso para presentar un objeto virtual para estar en foco cuando es visto por un usuario que observa una presentación transparente.

La Fig. 14 es un diagrama de flujo que describe una realización de un proceso para el concentrador que rastrea la posición y la orientación de un usuario y/o la unidad de presentación montada en la cabeza.

La Fig. 15 es un diagrama de flujo que describe una realización de un proceso para rastrear la posición de un ojo.

45 La Fig. 16 es un diagrama de flujo que describe una realización de un proceso para determinar el campo de visión de un usuario.

La Fig. 17A es un diagrama de flujo que describe una realización de un proceso para cambiar una región focal de un conjunto de micropresentación desplazando al menos una lente del conjunto.

La Fig. 17B es un diagrama de flujo que describe una realización de un proceso para cambiar una región focal de un conjunto de micropresentación cambiando una polarización de al menos una lente del conjunto.

- 5 La Fig. 17C es un diagrama de flujo que describe una realización de un proceso para cambiar una región focal de un conjunto de micropresentación cambiando un radio de curvatura de al menos una lente de fluido del conjunto.

La Fig. 18A es un diagrama de flujo que describe una realización de un proceso para generar imágenes de objetos virtuales en diferentes regiones focales en una presentación de realidad aumentada.

- 10 La Fig. 18B es un diagrama de flujo que describe otra realización de un proceso para generar imágenes de objetos virtuales en diferentes regiones focales en una presentación de realidad aumentada.

Las Figuras 19A y 19B ilustran esquemáticamente ejemplos de las diferentes formas de presentar objetos virtuales de múltiples focos en una presentación de realidad aumentada.

Descripción detallada

- 15 La tecnología proporciona realizaciones para conseguir focos variables para presentaciones de realidad mezclada o de realidad aumentada. En una realización, un sistema incluye una presentación próxima al ojo para realidad aumentada un dispositivo de presentación transparente y una unidad de procesamiento en comunicación con el dispositivo de presentación transparente. En las realizaciones descritas más adelante, la presentación transparente está en un conjunto de anteojos pero otros formatos HMD y soportes de presentación próxima a los ojos también pueden ser utilizados.

- 20 La visión natural del usuario es utilizada para ver realmente y directamente objetos reales en una escena vista con la presentación transparente. Basada en el contexto de ejecución de software, por ejemplo, una aplicación de juego, el sistema puede proyectar imágenes virtuales en la presentación que son visibles por la persona que lleva el dispositivo de presentación transparente a la vez que la persona que está también viendo objetos del mundo real a través de la presentación.

- 25 Uno o más sensores son utilizados para escanear el entorno físico que el usuario ve proporcionando de este modo datos sobre los cuales puede ser construido un modelo tridimensional del entorno escaneado. El modelo es segmentado en objetos reales, y como se explica más adelante, es aumentado con las ubicaciones de las imágenes de objetos virtuales.

- 30 Adicionalmente, son utilizados distintos sensores para detectar la posición y orientación de la cabeza del usuario y la posición del ojo con el fin de determinar dónde proyectar las imágenes virtuales. El sistema rastrea automáticamente dónde está mirando el usuario, de manera que el sistema puede calcular el campo de visión del usuario a través de la presentación del dispositivo de presentación transparente. El usuario puede ser rastreado utilizando cámaras de profundidad y cualquiera de los distintos sensores que incluyen sensores de profundidad, sensores de imagen, sensores de inercia, sensores de posición del ojo, etc. Además del campo de visión del usuario, el sistema también determina en qué ubicación del campo de visión el usuario está enfocando o está mirando, referida a menudo como región focal del usuario.

- 40 En algunas realizaciones, la región focal del usuario es un volumen conocido como área fusional de Panum, en la que los ojos humanos ven objetos con visión sencilla. Los seres humanos tienen visión binocular o visión estereoscópica. Cada ojo produce una imagen desde un punto de vista diferente. Sólo en este pequeño volumen del área fusional de Panum los humanos ven un objeto con visión sencilla. Esto es generalmente lo que significa cuando se dice que un objeto está en foco. Fuera de este área, los objetos pueden aparecer borrosos o incluso aparecer como imágenes dobles. Dentro del centro del área fusional de Panum hay un Horóptero que incluye el punto focal de los ojos del usuario. Cuando un usuario está enfocando en un punto en el espacio, en lo que sigue el punto focal, ese punto focal está situado en una línea curva. Los objetos en esta línea curvada en el espacio caen sobre la retinas de los ojos en la fóvea. La línea curvada está a veces referida como el horóptero horizontal. También existe un horóptero vertical que es una línea a través de la línea curvada que se inclina alejándose de los ojos por encima del punto focal y hacia los ojos debajo del punto focal sobre la curva. El término Horóptero como se utiliza en lo que sigue se refiere tanto a sus componentes verticales como horizontales.

- 50 Una vez que el sistema conoce el campo de visión del usuario, y la región focal del usuario, el sistema puede calcular dónde van a ser proyectadas una o más imágenes de objetos virtuales por la presentación del dispositivo de presentación transparente. La generación de una presentación de un objeto virtual en una ubicación de región focal seleccionada en la presentación proyectada se puede realizar ajustando una longitud de trayectoria óptica en el conjunto de micropresentación del dispositivo de presentación transparente.

- 55 El paralaje se genera desplazando las imágenes colocadas en las micropresentaciones izquierda y derecha. Ajustando una cantidad de paralaje para un objeto virtual, su distancia virtual es supuesta por la distancia a la que la

línea de vista al objeto virtual generado por la presentación de ojo izquierdo intersecta la correspondiente línea de vista de la presentación de ojo derecho. Las presentaciones convencionales pueden ajustar esta distancia implicada de paralaje. Independiente del paralaje hay una distancia implicada por la curvatura del frente de onda que sale de la presentación. Este radio de curvatura es simplemente la distancia a un punto sobre el objeto. Los objetos cercanos tienen frentes de onda fuertemente curvados debido a que el radio de curvatura es por tanto pequeño. Los objetos distantes tienen frentes de onda mucho más planos debido a que el radio de curvatura es correspondientemente grande. En el límite de los objetos muy distantes el frente de onda llega a ser el plano. Las presentaciones montadas en la cabeza convencionales tienen una curvatura de frente de onda fija debido a que no tienen elementos ópticos que puedan ser variados en base al contenido de la escena.

En el caso en el que la ubicación del usuario cambie a una ubicación más lejos de la ubicación actual del objeto virtual, el paralaje y la escala para el objeto virtual pueden ser cambiados colocando las imágenes apropiadas en las micro presentaciones izquierda y derecha. Un Horóptero, un área fusional de Panum para la nueva ubicación del usuario puede ser definida. Las presentaciones convencionales no pueden ajustar el radio de curvatura del frente de ondas para llevar el objeto virtual al interior del área fusional de Panum. La tecnología descrita más adelante puede establecer la curvatura del frente de ondas a una distancia, que también es acorde con otras señales establecidas por la escala y en paralaje, y que llevan el objeto virtual dentro del área fusional de Panum. De esta manera las imágenes parecen naturales y reales. El mismo argumento se aplica si el usuario permanece fijo en una ubicación, y el objeto virtual se mueve más cerca o más lejos con relación a objetos naturales en la escena.

Un conjunto de micropresentación comprende elementos de procesamiento de luz y un ajustador de foco variable. Algunos ejemplos de elemento de procesamiento de luz son una unidad de micropresentación, uno o más elementos ópticos, por ejemplo, lentes de un sistema de lentes y un elemento reflectante, por ejemplo una superficie reflectante o una superficie parcialmente reflectante. La unidad de micropresentación incluye una fuente de luz y genera una imagen de un objeto virtual. La unidad de micropresentación está ópticamente alineada con uno o más elementos ópticos y con el elemento reflectante. La alineación óptica puede ser a lo largo de un eje óptico o una trayectoria óptica que incluye uno o más ejes ópticos. La luz de imagen puede ser colimada y dirigida por uno o más elementos ópticos hasta el elemento reflectante. La luz reflejada desde un elemento parcialmente reflectante puede ser dirigida hasta las ubicaciones del ojo del usuario en una realización. Dado que la superficie es parcialmente reflectante, puede permitir que la luz procedente de la escena natural la atraviese y se vea. En otras realizaciones la luz procedente de la unidad de micropresentación reflejada por el elemento reflectante viaja al interior de otro elemento óptico que proyecta imágenes para ser visualizadas por el usuario, y también permite que la luz natural sea vista. El ajustador de foco variable cambia el desplazamiento entre uno o más elementos de procesamiento de luz en la trayectoria óptica del conjunto de micropresentación o una potencia óptica de un elemento en el conjunto de micropresentación. La potencia óptica de la lente está definida como el recíproco de su longitud focal, por ejemplo $1/\text{longitud focal}$ de manera que un cambio en uno afecta al otro. El cambio da lugar a un cambio en la región del campo de visión que está en foco para una imagen generada por el conjunto de micropresentación con el desplazamiento o la potencia óptica cambiados. Como se expone para las realizaciones siguientes, puede haber un conjunto de micropresentación para cada ojo. Cada uno de los conjuntos de micropresentación realiza el procesamiento para la perspectiva de su respectivo ojo.

En un ejemplo, una profundidad artificial de la técnica de campo tal como una borrosidad artificial se aplica a cualquier objeto virtual en el campo de visión y fuera de la región focal en proporción a su distancia desde la región focal. En otra realización, el ajustador barre a través de un rango de longitudes focales correspondientes a un rango de regiones focales a una velocidad o frecuencia a la vez que se presenta objetos virtuales dentro de las respectivas regiones focales. La velocidad o frecuencia puede ser igual o más rápida que una velocidad de fotograma para el dispositivo de presentación. En una realización, las imágenes de objeto virtual generadas en las diferentes regiones focales son presentadas como imágenes en capas, y la velocidad de presentación es lo suficientemente rápida para que las imágenes parezcan presentes al mismo tiempo. En otra realización, se presenta una imagen compuesta de las partes en foco de las imágenes generadas en las diferentes regiones focales. Cuando el usuario cambia su foco actual a otra región focal, los objetos virtuales en las diferentes regiones focales entran y salen del foco como cuando son vistos con la visión natural.

Las imágenes son entonces reproducidas mediante el dimensionamiento y la orientación de la imagen virtual y reproducida esa imagen dimensionada/orientada en una presentación transparente.

La Fig. 1 es un diagrama de bloques que muestra componentes a modo de ejemplo de una realización de un sistema 10 para proporcionar foco variable de objetos virtuales. El sistema 10 incluye un dispositivo de presentación transparente como dispositivo de presentación montado en la cabeza 2 en comunicación con la unidad de procesamiento 4 a través de un cable 6. En otras realizaciones, el dispositivo de presentación montado en la cabeza 2 se comunica con la unidad de procesamiento 4 a través de una comunicación sin cable. El dispositivo de presentación montado en la cabeza 2, que en una realización tiene forma de gafas con una montura 115 proporciona un soporte para sujetar los elementos del sistema en su sitio y también un conducto para las conexiones eléctricas, es llevado en la cabeza del usuario de manera que el usuario puede ver a través de una presentación o pantalla y con ello tener una visión directa y real del espacio delante del usuario. El uso de la expresión "en visión real y directa" se refiere a la capacidad de ver los objetos del mundo real directamente con el ojo humano, en lugar de ver representaciones de los objetos. Por ejemplo, mirar a través de las gafas en una habitación permite que el

usuario tenga una visión real directa de la habitación, mientras que el visionado de un video de una habitación en una televisión no es una imagen real directa de la habitación. Más adelante se proporcionan más detalles del dispositivo de presentación montado en la cabeza 2.

5 En una realización, la unidad de procesamiento 4 es llevada en la muñeca del usuario e incluye mucha de la potencia de ordenador utilizada para operar el dispositivo de presentación montado en la cabeza 2. La unidad de procesamiento 4 se comunica de manera inalámbrica (por ejemplo, mediante WiFi, Bluetooth, infrarrojos, u otros medios de comunicación inalámbrica) con uno o más sistemas de ordenador concentrador 12.

10 El sistema de ordenador concentrador 12 puede ser un ordenador, un sistema de juego o consola, o similar. De acuerdo con una realización a modo de ejemplo, el sistema de ordenador concentrador 12 puede incluir componentes de hardware y/o componentes de software, de manera que el sistema de ordenador concentrador 12 puede ser utilizado para ejecutar aplicaciones tales como aplicaciones de juego, aplicaciones que no son de juego, o similares. En una realización, el sistema de ordenador concentrador 12 puede incluir un procesador tal como un procesador estandarizado, un procesador especializado, un microprocesador, o similar, que puede ejecutar instrucciones almacenadas en un dispositivo de almacenamiento leíble por procesador para realizar los procesos descritos en la presente memoria.

15 El sistema de ordenador concentrador 12 incluye además uno o más dispositivos de captura, tal como dispositivos de captura 20A y 20B. En otras realizaciones, se pueden utilizar más de dos o menos de dos dispositivos de captura. En una implementación a modo de ejemplo, los dispositivos de captura 20A y 20B apuntan en direcciones diferentes de manera que capturan diferentes partes de la habitación. Puede resultar ventajoso que el campo de visión de los dos dispositivos de captura se superponga ligeramente de manera que el sistema de ordenador concentrador 12 puede entender cómo se relacionan entre sí los campos de visión de los dispositivos de captura. De esta manera, pueden ser utilizados múltiples dispositivos de captura para ver una habitación completa (u otro espacio). Alternativamente, un dispositivo de captura puede ser utilizado si el dispositivo de captura puede ser separado durante la operación de manera que en el tiempo el espacio relevante completo es visto por el dispositivo de captura.

Los dispositivos de captura 20A y 20B pueden ser, por ejemplo, cámaras que monitorizan visualmente uno o más usuarios y el espacio circundante de manera que los gestos y/o movimientos realizados por uno o más usuarios, así como la estructura del espacio circundante, pueden ser capturados, analizados, y rastreados para realizar uno o más controles o acciones dentro de la aplicación y/o animar un avatar o un carácter en la pantalla.

30 El sistema de ordenador concentrador 12 puede estar conectado a un dispositivo audiovisual 16 tal como una televisión, un monitor, una televisión de alta definición (HDTV), o similar que puede proporcionar elementos visuales de juego o de aplicación. Por ejemplo, el sistema de ordenador concentrador 12 puede incluir un adaptador de video tal como una tarjeta gráfica y/o un adaptador de audio tal como una tarjeta de sonido que puede proporcionar señales audiovisuales asociadas con la aplicación de juego, con la aplicación que no es un juego, etc. El dispositivo audiovisual 16 puede recibir las señales audiovisuales procedentes del sistema de ordenador concentrador 12 y puede entonces extraer los elementos visuales de juego o de aplicación y/o el audio asociado con las señales audiovisuales. De acuerdo con una realización, el dispositivo audiovisual 16 puede estar conectado a un sistema de ordenador concentrador 12 a través de, por ejemplo, un cable de video S, un cable coaxial, un cable HDMI, un cable DVI, un cable VGA, un cable de video componente, cables RCA, etc. En un ejemplo, un dispositivo audiovisual 16 incluye altavoces internos. En otras realizaciones, el dispositivo audiovisual 16, un estéreo separado o sistema de ordenador concentrador 12 está conectado a los altavoces externos 22.

35 El dispositivo de ordenador concentrador 10, con los dispositivos de captura 20A y 20B, pueden ser utilizados para reconocer, analizar, y/o rastrear objetivos humanos (y de otros tipos). Por ejemplo, un usuario que lleva un dispositivo de presentación montado en la cabeza 2 puede ser rastreado utilizando los dispositivos de captura 20A y 20B, de manera que los gestos y/o movimientos del usuario pueden ser capturados para animar un avatar o un carácter en la pantalla y/o pueden ser interpretados como controles que pueden ser utilizados para acercar la aplicación que está siendo ejecutada por el sistema de ordenador concentrador 12.

40 La Fig. 2A muestra una vista superior de una parte del dispositivo de presentación montado en la cabeza 2, que incluye una parte de la montura que incluye la patilla 102 y el puente de nariz 104. Sólo se muestra el lado derecho del dispositivo de presentación montado en la cabeza 2. Integrado en el puente de nariz 104 hay un micrófono 110 para grabar sonidos y transmitir esos datos de audio a la unidad de procesamiento 4, como se describe más adelante. En la parte frontal del dispositivo de presentación montado en la cabeza 2 hay una cámara de vídeo 113 vuelta hacia el entorno físico que puede capturar vídeo e imágenes estáticas. Esas imágenes son transmitidas a la unidad de procesamiento 4, como se describe más adelante.

45 Una parte de la montura 115 del dispositivo de presentación montado en la cabeza 2 rodeará una presentación (que incluye uno o más elementos ópticos). Para mostrar los componentes del dispositivo de presentación montado en la cabeza 2, una parte de la montura 115 que rodea la presentación no está mostrada. La presentación incluye un elemento óptico de guiado de luz 112, un filtro de opacidad 114, lentes transparentes 116 y lentes transparentes 118. En una realización, el filtro de opacidad 114 está detrás y alineado con la lente transparente 116, el elemento

5 óptico de guiado de luz 112 está detrás y alineado con el filtro de opacidad 114, y la lente transparente 118 está detrás y alineada con el elemento óptico de guiado de luz 112. Las lentes transparentes 116 y 118 son lentes estándar utilizadas en gafas y pueden estar hechas de acuerdo con cualquier prescripción (incluyendo sin prescripción). En una realización, las lentes transparentes 116 y 118 pueden ser sustituidas por una lente de prescripción variable. En algunas realizaciones, el dispositivo de presentación montado en la cabeza 2 incluirá solo una lente a través transparente o ninguna lente transparente. En otra alternativa, una lente de prescripción puede ir dentro del elemento óptico de guiado de luz 112. El filtro de opacidad 114 filtra la luz natural (o bien en base a los píxeles o bien uniformemente) para aumentar el contraste de la imagen virtual. El elemento óptico de guiado de luz 112 canaliza la luz artificial hasta el ojo. Se proporcionan más adelante más detalles del filtro de opacidad 114 y del elemento óptico de guiado de luz 112.

10 Montada en o dentro de la patilla 102 hay una fuente de imagen, que en una o más realizaciones incluye la micropresentación 120 que proyecta una imagen a través de uno o más elementos ópticos, por ejemplo un sistema de lentes 122, sobre un elemento reflectante que en esta realización es una superficie reflectante 124 que guía la imagen al interior del elemento óptico de guiado de luz 112. Existen diferentes tecnologías de generación de imágenes que pueden ser utilizadas para implementar la micropresentación 120. Por ejemplo la micropresentación 120 puede ser incrementada utilizando una tecnología de proyección transmisiva en donde la fuente de luz es modulada mediante material ópticamente activo, retroiluminación con luz blanca. Estas tecnologías normalmente son implementadas utilizando presentaciones de tipo LCD con potentes retroluminaciones y densidades de energía óptica elevadas. La micropresentación 120 también puede ser implementada utilizando una tecnología reflectante para la cual la luz externa es reflejada y modulada por un material ópticamente activo. La iluminación es iluminación delantera mediante o bien una fuente blanca o bien una fuente RGB, dependiendo de la tecnología. El procesamiento de luz digital (DLP), tecnología de presentación de cristal líquido en silicio (LCOS) y Mirasol® de Qualcomm, Inc., son ejemplos de tecnologías reflectantes que son eficientes cuando la mayoría de la energía es reflejada desde la estructura modulada y puede ser utilizada en el sistema descrito en la presente memoria. Adicionalmente, la micropresentación 120 puede ser implementada utilizando una tecnología de emisión en donde la luz es generada por la presentación. Por ejemplo, un motor PicoP™ de Microvision, Inc., emite una señal láser con un micro espejo que gira o bien sobre una pantalla diminuta que actúa como elemento de transmisión o bien es emitida directamente dentro del ojo (por ejemplo, láser).

15 En la realización ilustrada, la micropresentación 120 es parte de un conjunto de micropresentación 173 que comprende elementos de procesamiento de luz para transferir una imagen a una presentación transparente. El conjunto de micropresentación en este ejemplo comprende la micropresentación 120, uno o más elementos ópticos encarnados en un sistema de lentes 122, y la superficie reflectante 124. El sistema de lentes 122 puede comprender una única lente o una pluralidad de lentes. El sistema de lentes 122, la unidad de micropresentación 120 y la superficie reflectante 124 (por ejemplo, un espejo u otra superficie) están alineados en una trayectoria óptica, en este ejemplo, alineados a lo largo del eje óptico 133. La luz de imagen puede ser colimada y dirigida por el sistema de lentes 122 hasta la superficie reflectante 124.

20 El conjunto de micropresentación 173 comprende además un ajustador de foco virtual variable 135 que controla un desplazamiento entre sistema de lentes 122 y la unidad de micropresentación 120, o un desplazamiento entre sistema de lentes 122 y la superficie reflectante 124 o ambos a lo largo de la trayectoria óptica 133. Diferentes desplazamientos entre los elementos de procesamiento de luz del conjunto de micropresentación corresponden a diferentes regiones focales en el campo de visión de tres dimensiones del usuario en el que los elementos virtuales pueden ser proyectados. En este ejemplo, los cambios de desplazamiento son guiados dentro de una armadura 137 que soporta al menos un elemento de procesamiento de luz, tal como el sistema de lentes 122 y la micropresentación 120 en este ejemplo. La armadura 137 ayuda a estabilizar la alineación a lo largo de la trayectoria óptica 133 durante el movimiento físico de los elementos para conseguir un desplazamiento seleccionado o una potencia óptica seleccionada. El rango de desplazamiento es típicamente del orden de unos pocos milímetros (mm). En un ejemplo, el rango desde 1-2 mm.

25 En un ejemplo, el ajustador 135 puede ser un actuador tal como un motor piezoeléctrico. También pueden ser utilizadas otras tecnologías para el actuador y algunos ejemplos de tales tecnologías son una bobina de voz formada de una bobina y un imán permanente, un elemento de magnetostricción y un elemento de electrostricción.

30 El elemento óptico de guiado de luz 112 transmite luz desde la micropresentación 120 hasta el ojo 140 del usuario que lleva el dispositivo de presentación montado en la cabeza 2. El elemento óptico de guiado de luz 112 permite también que la luz procedente de la parte delantera del dispositivo de presentación montado en la cabeza 2 sea transmitida a través del elemento óptico de guiado de luz 112 hasta el ojo 140, como se representa mediante la flecha 142, lo que permite que el usuario tenga una visión real directa del espacio delante del dispositivo de presentación montado en la cabeza 2 además de recibir una imagen visual desde la micropresentación 120. De este modo, las paredes del elemento óptico de guiado de luz 112 son transparentes. El elemento óptico de guiado de luz 112 incluye una primera superficie reflectante 124. La luz procedente de la micropresentación 120 pasa a través del sistema de lentes 122 y se hace incidente sobre la superficie reflectante 124. La superficie reflectante 124 refleja la luz incidente procedente de la micropresentación 120 de manera que la luz es atrapada dentro de un "planer", sustrato que comprende el elemento óptico de guiado de luz 112 mediante reflexión interna. Después de varias reflexiones de las superficies del sustrato, las ondas de luz atrapadas alcanzan una disposición de superficies

selectivamente reflectantes 126. Nótese que solo una de las cinco superficies está etiquetada 126 para evitar sobrecargar el dibujo. Las superficies reflectantes 126 se asocian con las ondas de luz incidente sobre esas superficies reflectantes fuera del sustrato en el ojo 140 del usuario. Cuando diferentes rayos de luz se desplazan y rebotan fuera del interior del sustrato en diferentes ángulos, los diferentes rayos chocarán con las distintas superficies reflectantes 126 en diferentes ángulos. Por lo tanto, serán reflejados diferentes rayos de luz fuera del sustrato por diferentes superficies reflectantes. La selección de qué los rayos de luz serán reflejados fuera del sustrato mediante qué superficie 126 está diseñado seleccionando un ángulo apropiado de las superficies 126. Se pueden encontrar más detalles de un elemento óptico de guiado de luz en la Publicación de Solicitud de Patente de Estados Unidos 2008/0285140, N° de Serie 12/214.366, publicada el 20 de noviembre de 2008, "Substrate-Guided Optical Devices". En una realización, cada ojo tendrá su propio elemento óptico de guiado de luz 112. Cuando el dispositivo de presentación montada en la cabeza tiene dos elementos ópticos de guiado de luz, cada ojo puede tener su propia micropresentación 120 que puede presentar la misma imagen en ambos ojos o diferentes imágenes en los dos ojos. En otra realización, puede haber un elemento óptico de guiado de luz que refleje luz en ambos ojos.

El filtro de opacidad 114, que está alineado con el elemento óptico de guiado de luz 112, bloquea selectivamente la luz natural, o bien uniformemente o bien en base a los píxeles, para que no pase a través del elemento óptico de guiado de luz 112. En una realización, el filtro de opacidad puede ser un panel LCD transparente, una película electrocrómica, o un dispositivo similar que sea capaz de servir como filtro de opacidad. Tal panel LCD transparente se puede obtener retirando varias capas de sustrato, retroiluminación y difusores de un LCD convencional. El panel LCD puede incluir uno o más chips LCD transmisores de luz que permiten que la luz pase a través del cristal líquido. Tales chips son utilizados en proyectores LCD, por ejemplo.

El filtro de opacidad 114 puede incluir una red densa de píxeles, en donde la transmisividad de luz de cada pixel se puede controlar individualmente entre transmisividades mínima y máxima. Aunque un rango de transmisividad de 0-100% es ideal, también son aceptables rangos más limitados. Como ejemplo, un panel LCD monocromo con no más de dos filtros de polarización es suficiente para proporcionar un rango de opacidad de aproximadamente 50% a 90% por píxel, hasta la resolución del LCD. En la mínima del 50%, la lente puede tener una apariencia ligeramente tintada, lo cual es tolerable. La transmisividad del 100% representa una lente perfectamente clara. Una escala "alfa" pueden estar definida de 0-100%, en donde 0% no permite el paso de luz y 100% permite que toda la luz pase. El valor que alfa puede ser establecido para cada píxel por el circuito de control de filtro de opacidad 224 descrito más adelante.

Más detalles de un filtro de opacidad se proporcionan en la Solicitud de Patente de Estados Unidos N° 12/887.426, "Opacity Filter For See-Through Mounted Display," presentada el 21 de septiembre de 2010, y la Solicitud de Patente de Estados Unidos N° 12/905.952 presentada el 15 de octubre de 2010 titulada "Fusing Virtual Content Into Real Content".

El dispositivo de presentación montado en la cabeza 2 incluye también un sistema para rastrear la posición de los ojos del usuario. Como se explicará más adelante, el sistema rastreará la posición y orientación del usuario de manera que el sistema puede determinar el campo de visión del usuario. Sin embargo, un humano no percibirá nada delante de él. En su lugar, los ojos del usuario serán dirigidos a un subconjunto del entorno. Por lo tanto, en una realización, el sistema incluirá tecnología para rastrear la posición de los ojos del usuario con el fin de mejorar la medida del campo de visión del usuario. Por ejemplo, el dispositivo de presentación montado en la cabeza 2 incluye un conjunto de rastreo de ojos 134 (véase la Fig. 2A), que incluirá un dispositivo de iluminación de rastreo de ojos 134A y la cámara de rastreo de ojos 134B (véase la Fig. 4). En una realización, la fuente de iluminación de rastreo de ojos 134A incluye uno o más emisores infrarrojos (IR), que emiten luz infrarroja hacia el ojo. La cámara de rastreo de ojos 134B incluye una o más cámaras que detectan la luz infrarroja reflejada.

La posición de la pupila puede ser identificada mediante técnicas de generación de imagen conocidas que detectan la reflexión de la córnea. Por ejemplo, véase la patente de Estados Unidos 7.401.920, titulada "Head Mounted Eye Tracking and Display System", expedida el 22 de julio de 2008 a Ophir et al., incorporada en la presente memoria como referencia. Tal técnica puede situar una posición del centro del ojo con relación a la cámara de rastreo. Generalmente, el rastreo del ojo implica obtener una imagen del ojo y utilizando técnicas de visión de ordenador determinar la ubicación de la pupila dentro del globo ocular. En una realización, es suficiente rastrear la ubicación de un ojo dado que los ojos normalmente se mueven al unísono. Sin embargo, es posible rastrear cada ojo de manera separada.

En una realización, el sistema utilizará cuatro LEDs de infrarrojos (IR) y cuatro fotodetectores de infrarrojos en una configuración rectangular, de manera que hay un LED de infrarrojos y un fotodetector de infrarrojos en cada esquina de la lente del dispositivo de presentación montado en la cabeza 2. La luz procedente de los LEDs se refleja en los ojos. La cantidad de luz infrarroja detectada en cada uno de los cuatro fotodetectores de infrarrojos determina la dirección de la pupila. Esto es, la cantidad de blanco frente al negro en el ojo determinará la cantidad de luz reflejada fuera del ojo para ese fotodetector particular. Esto es, el fotodetector tendrá una medida de la cantidad de blanco o negro en el ojo. A partir de las cuatro muestras, el sistema puede determinar la dirección del ojo.

Otra alternativa es utilizar cuatro LEDs infrarrojos como se ha descrito anteriormente, pero sólo un dispositivo de generación de imágenes infrarrojas en el lado de la lente del dispositivo de presentación montado en la cabeza 2. El

dispositivo de generación de imágenes utilizará un pequeño espejo y/o lente (ojo de pez) de manera que el dispositivo de generación de imágenes puede generar imágenes de hasta el 75% del ojo visible desde la montura de las gafas. El dispositivo de generación de imágenes detectará entonces una imagen y utilizará visión de ordenador para encontrar la imagen, como se ha descrito anteriormente. De este modo, aunque la Fig. 2A muestra un conjunto con un emisor de infrarrojos, la estructura de la Fig. 2A puede ser ajustada para tener cuatro transmisores de infrarrojos y/o cuatro sensores de infrarrojos. También se pueden utilizar más de cuatro o menos de cuatro transmisores de infrarrojos y/o cuatro sensores de infrarrojos.

Otra realización para rastrear la dirección de los ojos está basada en el rastreo de carga. Este concepto está basado en la observación de que una retina lleva una carga positiva medible y la córnea tiene una carga negativa. Los sensores están montados en los oídos del usuario (cerca de los auriculares 130) para detectar el potencial eléctrico mientras los ojos se mueven alrededor y leen de forma efectiva lo que los ojos están haciendo en tiempo real. También se pueden utilizar otras realizaciones para el rastreo de los ojos.

Los circuitos de control 136 proporcionan diversos elementos electrónicos que soportan los otros componentes del dispositivo de presentación montado en la cabeza 2. Más detalles de los circuitos de control 136 se proporcionan más adelante con respecto a la Fig. 4. Dentro de o montados en la patilla 102, hay auriculares 130, sensores de inercia 132 y un sensor de temperatura 138. En una realización, los sensores de inercia 132 incluyen un magnetómetro de tres ejes 132A, un giroscopio de tres ejes 132B y un acelerómetro de tres ejes 132C (véase la Fig. 4). Los sensores de energía son para detectar la posición, orientación, aceleraciones repentinas del dispositivo de presentación montado en la cabeza 2.

La Fig. 2A sólo muestra la mitad del dispositivo de presentación montado en la cabeza 2. Un dispositivo de presentación montado en la cabeza, completo incluiría otro conjunto de lentes transparentes, otro filtro de opacidad, otro elemento óptico de guiado de luz, otra micropresentación 120, otro sistema de lentes 122, la cámara vuelta hacia la habitación 113, el conjunto de rastreo de ojo 134, los auriculares 130, y el sensor de temperatura 138.

La Fig. 2B es una vista superior de una parte de otra realización de una unidad de presentación montada en la cabeza. La montura 115 se muestra en líneas de puntos para exponer la disposición de los elementos ópticos soportados en la misma. En esta realización, la luz procedente de un elemento reflectante 124a es dirigida a un elemento parcialmente reflector 124b que combina la vista de imagen virtual que viaja a lo largo de la trayectoria óptica 133 con la vista natural o real y directa 142. La vista natural 142 puede estar afectada por el filtro de opacidad 114. La combinación de vistas es dirigida al ojo del usuario 140 en lugar de al interior de otro elemento óptico tal como el elemento óptico de guiado de luz 112. En esta realización, la montura 115 proporciona una montura de gafas conveniente como soporte para los elementos del sistema tales como que el conjunto de micropresentación 173 que incluye sus elementos de procesamiento de luz 122 y 120, el ajustador de foco variable 135 y la armadura 137. En este ejemplo, la cámara de rastreo de ojos 134r, 1341 para cada ojo está situada sobre el puente 104. En otras realizaciones, se pueden utilizar otras estructuras de soporte además de una montura de gafas. Un ejemplo de tal estructura es un visor.

Como se ha mencionado anteriormente, la configuración de los elementos de procesamiento de luz del conjunto de micropresentación 173 crea una distancia focal o región focal en la que un objeto virtual aparece en una imagen. Cambiando la configuración se cambia la región focal para la imagen de objeto virtual. La región focal determinada por los elementos de procesamiento de luz puede ser determinada y cambiada en base a la ecuación $1/S_1 + 1/S_2 = 1/f$. El símbolo f representa la longitud focal de una lente tal como el sistema de lentes 122 en el conjunto de micropresentación 173. El sistema de lentes 122 tiene un punto nodal delantero y un punto nodal trasero. Si los rayos de luz son dirigidos hacia un punto nodal en un ángulo dado con relación al eje óptico, los rayos de luz surgirán desde el otro punto nodal en un ángulo equivalente con relación al eje óptico. En un ejemplo, el punto nodal trasero del sistema de lentes 122 estaría entre él mismo y el elemento reflectante 124 en la Fig. 2A o el elemento reflectante 124a en la Fig. 2B. La distancia desde el punto nodal trasero al elemento reflectante 124, 124a puede ser designada como S_2 . El punto nodal delantero estaría entre sistema de lentes 122 y una ubicación objetivo de la imagen virtual generada por la micropresentación 120 en un modelo tridimensional del espacio físico del usuario. (Véase la exposición de las Figuras 10-12 más adelante para más detalles acerca de la creación del modelo.) La distancia desde el punto nodal delantero a la ubicación objetivo de la imagen virtual puede ser designada como S_1 .

Si la longitud focal de la lente es fija, S_1 y S_2 son modificados para enfocar objetos virtuales en diferentes profundidades. Por ejemplo, una posición inicial puede tener S_1 establecido para el infinito, y S_2 ser igual a la longitud focal del sistema de lentes 122. Asumiendo que el sistema de lentes 122 tiene una longitud focal de 10 mm, se considera un ejemplo en el que el objeto virtual va ser colocado aproximadamente un pie o 30 cm dentro del campo de visión del usuario. S_1 es ahora aproximadamente 30 cm o 300 mm, f es 10 mm y S_2 es establecido actualmente en la posición inicial de la longitud focal, 10 mm, lo que significa que el punto nodal trasero del sistema de lentes 122 está a 10 mm del elemento reflectante 124, 124a. La nueva distancia o el nuevo desplazamiento entre la lente 122 y el elemento reflectante 124, 124a es determinado en base a $1/300 + 1/S_2 = 1/10$ con todas las unidades en mm. El resultado es de aproximadamente 10,3 mm para S_2 .

En un ejemplo, la unidad de procesamiento 4 puede calcular los valores del desplazamiento S_1 y S_2 , dejando la longitud focal f fija y haciendo que los circuitos de control 136 hagan que el accionador de ajustador variable 237

envíe señales de accionamiento para hacer que el ajustador de foco virtual variable 135 mueva el sistema de lentes 122 a lo largo de la trayectoria óptica 133 por ejemplo. En algunos ejemplos, el ajustador 135 puede mover uno o más elementos ópticos 122 dentro de la armadura 137. En otros ejemplos, la armadura puede tener ranuras o espacio en el área alrededor de un elemento de procesamiento de luz de manera que se desliza sobre elemento, por ejemplo la micropresentación 120, sin mover el elemento de procesamiento de luz. Otro elemento en la armadura tal como el uno o más elementos ópticos 122 están unidos de manera que se deslizan o se mueven con la armadura en movimiento 237. En otras realizaciones, la unidad de micropresentación 120 o el elemento reflectante 124, 124a o ambos se pueden mover en lugar de, o además de, moviendo el sistema de lentes 122.

En otras realizaciones, la longitud focal de al menos una lente en el sistema de lentes 122 puede ser cambiada con cambios en el desplazamiento a lo largo de la trayectoria óptica 133 también. Algunas realizaciones de conjuntos de micropresentación se presentan en las Figuras 3A a 3D. El número específico de lentes mostrado es sólo a modo de ejemplo. Se pueden utilizar otros números y configuraciones de lentes que funcionan basadas en los mismos principios.

La Fig. 3A es un ejemplo de un sistema de lentes giratorio para utilizar como parte de un conjunto de la micropresentación de una presentación próxima al ojo. Cada una de las lentes 122a a 122d tiene una longitud focal diferente y está soportada dentro de un soporte de disco 160 que puede girar por el ajustador de foco virtual variable 135. La unidad de procesamiento 4 determina una región focal y selecciona una de las lentes de longitud focal para obtener esa región focal. Como se muestra con respecto a la Fig. 3, un accionador ajustador variable 237 de los circuitos de control 136 envía al menos una señal de control al ajustador de foco virtual variable 135 para girar el disco, de manera que la lente seleccionada es alineada en la trayectoria óptica 133 del conjunto.

La Fig. 3B1 y la Fig. 3B2 son ejemplos de una lente de fluido que presentan diferentes radios de curvatura como se muestra en Hongwen et al., "Tunable-focus liquid lens controlled using a servo motor", OPTICS EXPRESS, 4 septiembre 2006, Vol. 14, No. 18, pp. 8031-8036. La lente de fluido puede ser utilizada como parte de un conjunto de micropresentación. La longitud focal de la lente puede ser cambiada cambiando el radio de curvatura. La relación entre el radio de curvatura, R, y una longitud focal f viene dada por $f = R / (n_{\text{líquido}} - 1)$. El índice de refracción del líquido o del fluido de la lente es $n_{\text{líquido}}$.

Esta realización comprende un soporte 137 tal como un anillo de obturación anular que tiene una membrana externa flexible 152, en un ejemplo una membrana de goma, como una parte o conectada para él. La membrana externa 153 está en contacto con un depósito de líquido 156. La membrana de lente 150 se apoya sobre la parte superior de, o forma un lado flexible de una célula de lente de líquido 158 que puede recibir desde, y liberar líquido en el, depósito 156. En el ejemplo citado, la membrana de lente flexible es una membrana elástica tal como una membrana elastomérica de polidimetilsiloxano (PDMS). La placa de vidrio 154 detrás de la célula de líquido proporciona soporte. El ajustador de foco virtual variable 135 es controlado para empujar contra la membrana externa flexible 152 como se muestra en la Fig. 3B2 y liberar la membrana 152 como en la Fig. 3B1 para hacer que el volumen de agua del depósito 156 entre y salga de la lente de líquido 158 haciendo, de este modo, convexa la membrana elástica 150 y relajando la membrana elástica 150 debido a los cambios en el volumen de líquido. Los cambios en el volumen de líquido producen cambios en el radio de curvatura de la membrana de lente 150 y de este modo es la longitud focal de la lente de líquido 158. La relación entre el radio de curvatura y el cambio de volumen ΔV se puede expresar como sigue:

$$\Delta V = (1/3) \pi (2R^2 - r_0^2 - 2R\sqrt{R^2 - r_0^2}) (2R + \sqrt{R^2 - r_0^2})$$

en donde r_0 es el radio de la apertura de lente.

La Fig. 3C es un ejemplo que un sistema de lentes birrefringentes para utilizar como parte de un conjunto de micropresentación de una presentación próxima al ojo. Un material birrefringente es anisótropo o direccionalmente dependiente. Describir la luz como un rayo es una construcción ilustrativa, una lente birrefringente descompone la luz en un rayo ordinario y en un rayo extraordinario. Para un único eje de anisotropía o eje óptico, diferentes pueden existir índices de refracción, y por tanto diferentes longitudes focales, para diferentes polarizaciones, una paralela y una perpendicular al eje. En un ejemplo de la Fig. 3C, las lentes 122a y 122b son lentes hechas de un material birrefringente con diferentes polarizaciones como se indica mediante las flechas. En este ejemplo de dos lentes, pueden ser predeterminados por selección cuatro índices de refracción diferentes o longitudes focales. Cada longitud focal diferente puede estar asociada con una región focal diferente para la selección mediante la unidad de procesamiento 4. Las combinaciones de polarización pueden ser polarizaciones perpendiculares para las lentes 122a y 122b como se muestra en la Fig. 3C, las polarizaciones perpendiculares opuestas a las mostradas en la Fig. 3C, las dos lentes que tienen las mismas polarizaciones en una dirección, y las dos lentes tienen las mismas polarizaciones en la otra dirección de polarización. El ajustador de foco virtual variable puede aplicar un voltaje a cada lente para efectuar una polarización seleccionada en una realización. En otra realización, se puede aplicar un esfuerzo físico para cambiar la polarización de una lente.

La Fig. 3D es un ejemplo de un sistema de lentes insertable para utilizar como parte de un conjunto de micropresentación. En esta realización, cada una de un número de lentes 122 está unida a un brazo respectivo 123

unido a la armadura 137. Cada brazo 123 mueve su uno o más elementos ópticos 122, por ejemplo una lente o un sistema de lentes 122, en oposición en un desplazamiento en la trayectoria óptica 133 del conjunto de micropresentación bajo el control del ajustador de foco virtual variable 135. Por ejemplo, si están siendo utilizados desplazamientos predeterminados para preestablecer regiones focales, cada lente 122 puede ser establecida a una cierta distancia separada de sus vecinas, por ejemplo separada 0,1 milímetros (mm). También se puede utilizar separación no uniforme y desplazamientos ajustables.

En cada uno de los ejemplos anteriores que ajusta la longitud focal de una lente, también se realiza el desplazamiento a lo largo de la trayectoria óptica 133.

De nuevo como se ha mencionado anteriormente, cada uno de los conjuntos de micropresentación realiza el procesamiento para la perspectiva de su respectivo ojo, de manera que la imagen virtual aparece en la región focal del usuario, por ejemplo, el área fusional de Panum, el área de visión sencilla para los seres humanos.

La Fig. 4 es un diagrama de bloques que muestra los diversos componentes del dispositivo de presentación montado en la cabeza 2. La Fig. 5 es un diagrama de bloques que describe los distintos componentes de la unidad de procesamiento 4. El dispositivo de presentación montado en la cabeza 12, cuyos componentes se muestran en la Fig. 4, son utilizados para proporcionar una imagen virtual en foco con la vista del usuario del mundo real. Adicionalmente, los componentes del dispositivo de presentación montado en la cabeza de la Fig. 4 incluyen muchos sensores que rastrean diversas condiciones. El dispositivo de presentación montado en la cabeza recibirá instrucciones acerca de la imagen virtual procedente de la unidad de procesamiento 4 y proporcionará la información de sensor de nuevo a la unidad de procesamiento 4. La unidad de procesamiento 4, cuyos componentes se muestran en la Fig. 5 recibirán la información de sensor procedente del dispositivo de presentación montado en la cabeza 3 y también procedente del dispositivo de ordenador concentrador 12 (véase la Fig. 1). En base a esa información, la unidad de procesamiento 4 determinará dónde y cuándo proporcionar una imagen virtual en foco al usuario y enviar instrucciones acordes con el dispositivo de presentación montado en la cabeza de la Fig. 4.

Nótese que algunos de los componentes de la Fig. 4 (por ejemplo, la cámara vuelta hacia el entorno físico 113, la cámara de rastreo de ojos 134B, el ajustador de foco virtual variable 135, la micropresentación 120, el filtro de opacidad 114, la iluminación de rastreo de ojos 134A, los auriculares 130, y el sensor de temperatura 138) se muestran sombreados para indicar que hay dos de cada uno de esos dispositivos, uno en el lado izquierdo y uno en el lado derecho del dispositivo de presentación montado en la cabeza 2. La Fig. 4 muestra el circuito de control 200 en comunicación con el circuito de gestión de potencia 202. El circuito de control 200 incluye el microprocesador 210, el controlador de memoria 212 en comunicación con la memoria 214 (por ejemplo, D-RAM), la interfaz de cámara 216, el búfer de cámara 218, el accionador de presentación 220, el formateador de presentación 222, el generador de sincronización 226, la interfaz de salida de presentación 228, la interfaz de entrada de presentación 230. En una realización, todos los componentes del circuito de control 220 están en comunicación entre sí a través de líneas dedicadas uno o más buses. En otra realización, cada uno de los componentes del circuito de control 200 está en comunicación con el procesador 210. La interfaz de cámara 216 proporciona una interfaz a las dos cámaras vueltas hacia el entorno físico 113 y almacena imágenes recibidas desde las cámaras vueltas hacia el entorno físico en el búfer de cámara 218. El accionador de presentación 220 accionará la micropresentación 120. El formateador de presentación 222 proporciona información acerca de la imagen virtual que está siendo presentada en la micropresentación 120, al circuito de control de opacidad 224, que controla el filtro de opacidad 114. El generador de sincronización 226 se utiliza para proporcionar datos de sincronización para el sistema. La salida de presentación 228 es un búfer que proporciona imágenes procedentes de las cámaras vueltas hacia el entorno físico 113 a la unidad de procesamiento 4. La entrada de presentación 230 es un búfer para recibir imágenes tales como una imagen virtual que va ser presentada en la micropresentación 120. La salida de presentación 228 y la entrada de presentación 230 comunican con la interfaz de banda 232 que es una interfaz para la unidad de procesamiento 4.

El circuito de gestión de potencia 202 incluye el regulador de voltaje 234, el accionador de iluminación de rastreo de ojo 236, el accionador ajustador variable 237, el DAC de audio y el amplificador 238, el preamplificador de micrófono y el ADC de audio 240, la interfaz de sensor de temperatura 242, el generador de reloj 244. El regulador de voltaje 234 recibe potencia desde la unidad de procesamiento 4 a través de la interfaz de banda 232 y proporciona esa potencia a los otros componentes del dispositivo de presentación montado en la cabeza 2. El accionador de iluminación de rastreo de ojos 236 proporciona la fuente de luz infrarroja para la iluminación de rastreo de ojos 134A, como se ha descrito anteriormente. El accionador ajustador variable 237 proporciona una señal de control, por ejemplo una corriente de accionamiento o un voltaje de accionamiento, al ajustador 135 para mover uno o más elementos del conjunto de micropresentación para conseguir un desplazamiento para una región focal calculada mediante el software que se ejecuta en la unidad de procesamiento 4 o el ordenador concentrador 12 ambos. En las realizaciones de barrido a través de un rango de desplazamientos y, por tanto, un rango de regiones focales, el accionador ajustador variable 237 recibe señales de sincronización procedentes del generador de sincronización 226, o alternativamente, el generador de reloj 244 para operar a una velocidad o frecuencia preprogramadas. El DAC de audio y el amplificador 238 reciben la información de audio procedente de los auriculares 130. El preamplificador de micrófono y el ADC de audio 240 proporcionan una interfaz para el micrófono 110. La interfaz de sensor de temperatura 242 es una interfaz para el sensor de temperatura 138. La unidad de gestión de potencia 202 proporciona también potencia y recibe datos procedentes del magnetómetro de tres ejes 132A, el giroscopio que tres ejes 132B y el acelerómetro de tres ejes 132C.

La Fig. 5 es un diagrama de bloques que describe los distintos componentes de la unidad de procesamiento 4. La Fig. 5 muestra el circuito de control 304 en comunicación con el circuito de gestión de potencia 306. El circuito de control 304 incluye una unidad de procesamiento central 320, la unidad de procesamiento de gráficos 322, el caché 324, la RAM 326, el control de memoria 328 en comunicación con la memoria 330 (por ejemplo, D-RAM), el controlador de memoria temporal 332 en comunicación con la memoria temporal 334 (u otro tipo de almacenamiento no volátil), el búfer de salida de presentación 336 en comunicación con el dispositivo de presentación montado en la cabeza 2 a través de la interfaz de banda 302 y la interfaz de banda 232, el búfer de entrada de presentación 338 en comunicación con el dispositivo de presentación montado en la cabeza 2 a través de la interfaz de banda 302 y la interfaz de banda 232, la interfaz de micrófono 340 en comunicación con un conector de micrófono externo 342 para conectar a un micrófono, la interfaz exprés de PCI para conectar al a un dispositivo de comunicación sin cable 346, y el(los) puerto(s) USB 348. En una realización, el componente de comunicación sin cable 346 puede incluir un dispositivo de comunicación habilitado para Wi-Fi, un dispositivo de comunicación Bluetooth, dispositivo de comunicación de infrarrojos, etc. El puerto USB puede ser utilizado para conectar la unidad de procesamiento 4 al dispositivo de ordenador concentrador 12 con el fin de cargar datos o el software en la unidad de procesamiento 4, así como cargar la unidad de procesamiento 4. En una realización, la CPU 320 y la GPU 322 son los componentes principales para determinar dónde, cuándo y cómo insertar imágenes virtuales en la visión del usuario. Se proporcionan más detalles más adelante.

El circuito la gestión de potencia 306 incluye el generador de reloj 360, el convertor de analógico a digital 362, el cargador de batería 364, el regulador de voltaje 366, la fuente de alimentación de presentación montada en la cabeza 376, y la interfaz de sensor de temperatura 372 en comunicación con el sensor de temperatura 374 (situado en la banda de muñeca de la unidad de procesamiento 4). Una corriente alterna para dirigir el convertidor de corriente 362 está conectada a un enchufe de carga 370 para recibir un suministro de corriente alterna y crear un suministro de corriente continua para el sistema. El regulador de voltaje 366 está en comunicación con la batería 368 para suministrar potencial sistema. El cargador de batería 364 se utiliza para cargar la batería 368 (a través del regulador de voltaje 366) después de recibir potencia desde el enchufe de carga 370. La interfaz de potencia HMD 376 proporciona potencia al dispositivo de presentación montado en la cabeza 2.

El sistema descrito anteriormente estará configurado para insertar una imagen virtual en el campo de visión de un usuario de manera que la imagen virtual aparece en foco o fuera de foco de forma natural para el usuario. En varias realizaciones, la imagen virtual será ajustada para coincidir con la orientación apropiada, tamaño y forma en base al entorno para el que la imagen está siendo insertada. En una realización, el dispositivo de presentación montado en la cabeza 2, la unidad de procesamiento 4 y el dispositivo de ordenador concentrador 12 trabajan juntos ya que cada uno de los dispositivos incluye un subconjunto de sensores que son utilizados para obtener los datos para determinar dónde cuándo y cómo insertar las imágenes virtuales. En una realización, los cálculos que determinan dónde, cómo y cuándo insertar las imágenes virtuales son realizados por el dispositivo de ordenador concentrador 12. En otra realización, estos cálculos son realizados mediante la unidad de procesamiento 4. En otra realización algunos de los cálculos son realizados por el dispositivo de ordenador concentrador 12 mientras que otros cálculos son realizados por la unidad de procesamiento 4. En otras realizaciones, los cálculos pueden ser realizados por el dispositivo de presentación montado en la cabeza 2.

En una realización a modo de ejemplo, el dispositivo de ordenador concentrador 12 creará un modelo del entorno en el que el usuario está y rastreará diversos objetos en movimiento en ese entorno. Además, el dispositivo de ordenador concentrador 12 rastrea el campo de visión del dispositivo de presentación montado en la cabeza 2 rastreando la posición y la orientación del dispositivo de presentación montado en la cabeza 2. El modelo y la información de rastreo son proporcionados desde el dispositivo de ordenador concentrador 12 a la unidad de procesamiento 4. La información de sensor obtenida por el dispositivo de presentación montado en la cabeza 2 es transmitida a la unidad de procesamiento 4. La unidad de procesamiento 4 entonces utiliza la información de sensor adicional que recibe desde el dispositivo de presentación montado en la cabeza 2 para redefinir el campo de visión del usuario y proporcionar instrucciones al dispositivo de presentación montado en la cabeza 2 acerca de cómo, dónde y cuando insertar la imagen virtual.

La Fig. 6 ilustra una realización a modo de ejemplo del sistema de ordenador concentrador 12 con un dispositivo de captura. En una realización, los dispositivos de captura 20A y 20B son la misma estructura, por tanto, la Fig. 6 solo muestra dispositivo de captura 20A.

De acuerdo con una realización a modo de ejemplo, el dispositivo de captura 20A puede estar configurado para capturar vídeo con información de profundidad que incluye una imagen de profundidad que puede incluir valores de profundidad mediante cualquier técnica apropiada, por ejemplo, tiempo de vuelo, luz estructurada, estéreo imagen, o similar. De acuerdo con una realización, el dispositivo de captura 20A puede organizar la información de profundidad en "capas Z", o capas que pueden ser perpendiculares a un eje Z que se extiende desde la cámara de profundidad a lo largo de su línea de visión.

Como se muestra en la Fig. 6, el dispositivo de captura 20A puede incluir un componente de cámara 423. De acuerdo con una realización a modo de ejemplo, el componente de cámara 423 puede ser o puede incluir una cámara de profundidad que puede capturar una imagen de profundidad de una escena. La imagen de profundidad puede incluir un área de píxeles en dos dimensiones (2-D) de la escena capturada en donde cada píxel en el área

de píxeles 2D puede representar un valor de profundidad tal como una distancia en, por ejemplo, centímetros, milímetros, o similares, de un objeto en la escena capturada desde la cámara.

El componente de cámara 23 puede incluir un componente de luz infrarroja (IR) 425, una cámara tridimensional (3-D) 426, y una cámara RGB (imagen visual) 428 que se puede utilizar para capturar la imagen de profundidad de una escena. Por ejemplo, en el análisis de tiempo de vuelo, el componente de luz infrarroja 425 del dispositivo de captura 20A puede emitir una luz infrarroja sobre la escena y puede entonces utilizar sensores (en algunas realizaciones, incluyendo sensores no mostrados) para detectar la luz retrodispersada procedente de la superficie de uno o más objetivos y objetos en la escena utilizando, por ejemplo, la cámara 3-D 326 y/o la cámara RGB 428. En algunas realizaciones, puede ser utilizada luz infrarroja de pulsos de manera que el tiempo entre un pulso de luz que sale y un correspondiente pulso de luz que entra puede ser medido y utilizado para determinar una distancia física desde el dispositivo de captura 20A hasta una ubicación particular sobre los objetivos u objetos en la escena. Adicionalmente, en otras realizaciones a modo de ejemplo, la fase de la onda de luz saliente puede ser comparada con la fase de la onda de luz entrante para determinar un desplazamiento de fases. El desplazamiento de fases puede ser entonces utilizado para determinar una distancia física desde el dispositivo de captura a una ubicación particular en los objetivos u objetos.

De acuerdo con otra realización a modo de ejemplo, el análisis de tiempo de vuelo puede ser utilizado para determinar de forma indirecta una distancia física desde el dispositivo de captura 20A a una ubicación particular en los objetivos u objetos analizando la intensidad del rayo de luz reflejado en el tiempo a través de distintas técnicas que incluyen por ejemplo generación de imágenes de pulsos de luz obturada.

En otra realización a modo de ejemplo, el dispositivo de captura 20A puede utilizar una luz estructurada para capturar información de profundidad. En tal análisis, la luz como patrón (es decir, la luz presentada como un patrón conocido tal como un patrón reticular, un patrón de rayas, o un patrón diferente) puede ser proyectada sobre la escena mediante, por ejemplo, el componente de luz infrarroja 424. Después de chocar con la superficie de uno o más objetivos u objetos en la escena, el patrón se puede deformar como respuesta. Tal deformación del patrón puede ser capturada mediante, por ejemplo, la cámara 3-D 426 y/o la cámara RGB 428 (y/u otro sensor) y puede entonces ser analizada para determinar una distancia física desde el dispositivo de captura hasta una ubicación particular sobre los objetivos u objetos. En algunas implementaciones, el componente de luz infrarroja 425 está desplazado de las cámaras 425 y 426 de manera que se puede utilizar triangulación para determinar la distancia desde las cámaras 425 y 426. En algunas implementaciones, el dispositivo de captura 20A incluirá un sensor de infrarrojos dedicado para detectar la luz infrarroja, o un sensor con un filtro de infrarrojos.

De acuerdo con otra realización, el dispositivo de captura 20A puede incluir dos o más cámaras físicamente separadas que pueden ver una escena desde diferentes ángulos para obtener datos de estéreo visual que pueden ser resueltos para generar información de profundidad. También se pueden utilizar otros tipos de sensores de imagen de profundidad para crear una imagen en profundidad.

El dispositivo de captura 20A puede incluir además un micrófono 430, que incluye un transductor o sensor que puede recibir y convertir un sonido en una señal eléctrica. El micrófono 430 puede ser utilizado para recibir señales de audio que también pueden ser proporcionadas por el sistema de ordenador concentrador 12.

En una realización a modo de ejemplo, el dispositivo de captura 20A puede incluir además un procesador 432 que puede estar en comunicación con el componente de cámara de imagen 423. El procesador 432 puede incluir un procesador estandarizado, un procesador especializado, un microprocesador, o similar, que puede ejecutar instrucciones que incluyen, por ejemplo, instrucciones para recibir una imagen de profundidad, generar el formato de datos apropiado (por ejemplo, fotograma) y transmitir los datos al sistema de ordenador concentrador 12.

El dispositivo de captura 20A también puede incluir una memoria 434 que puede almacenar las instrucciones que son ejecutadas por el procesador 432, imágenes o fotogramas de imágenes capturados por la cámara 3-D y/o la cámara RGB, o cualquier otra información adecuada, imágenes, o similares. De acuerdo con una realización a modo de ejemplo, la memoria 434 puede incluir memoria de acceso aleatorio (RAM), una memoria de solo lectura (ROM), caché, memoria temporal, un disco duro, o cualquier otro componente de almacenamiento adecuado. Como se muestra en la Fig. 6, en una realización, la memoria 434 puede ser un componente separado en comunicación con el componente de captura de imagen 423 y el procesador 432. De acuerdo con otra realización, la memoria 434 puede estar integrada en el procesador 432 y/o el componente de captura de imagen 422.

Los dispositivos de captura 20A y 20B están en comunicación con el sistema de ordenador concentrador 12 a través de un enlace de comunicación 436. El enlace de comunicación 436 puede ser una conexión por cable que incluye, por ejemplo, una conexión USB, una conexión Firewire, una conexión de cable Ethernet, o similar y/o una conexión sin cable tal como una conexión inalámbrica 802.11b, g, a, o n. De acuerdo con una realización, el sistema de ordenador concentrador 12 puede proporcionar un reloj para el dispositivo de captura 20A que puede ser utilizado para determinar cuándo capturar, por ejemplo, una escena mediante el enlace de comunicación 436. De manera adicional, el dispositivo de captura 20A proporciona las imágenes de información de profundidad y visuales (por ejemplo, RGB) capturadas por, por ejemplo, la cámara 3-D 426 y/o la cámara RGB 428 para el sistema de ordenador concentrador 12 a través del enlace de comunicación 436. En una realización, las imágenes de

profundidad y las imágenes visuales son transmitidas a 30 fotogramas por segundo; sin embargo, se pueden utilizar otras velocidades de fotograma. El sistema de ordenador concentrador 12 puede entonces crear y utilizar un modelo, información de profundidad, e imágenes capturadas para, por ejemplo, controlar una aplicación tal como un juego o un procesador de palabras y/o animar un avatar o un carácter que en la pantalla.

5 El sistema de ordenador concentrador 12 incluye un módulo que procesamiento de imagen de profundidad y de rastreo óseo 450, que utiliza las imágenes de profundidad para rastrear una o más personas separables por la función de cámara de profundidad del dispositivo de captura 20A. El módulo de procesamiento de imagen de profundidad y de rastreo óseo 450 proporciona la información de rastreo para la aplicación 452, que puede ser un juego de vídeo, una aplicación de productividad, una aplicación de comunicaciones u otra aplicación de software, etc. Los datos de audio y los datos de imagen visual también son proporcionados a la aplicación 452 y al módulo de procesamiento de imagen de profundidad y rastreo óseo 450. La aplicación 452 proporciona la información de rastreo, los datos de audio y los datos de imagen visual al motor reconocedor 454. En otra realización, el motor reconocedor 454 recibe la información de rastreo directamente desde el módulo de procesamiento de imagen de profundidad y de rastreo óseo 450 y recibe los datos de audio y los datos de imagen visual directamente desde los dispositivos de captura 20A y 20B.

El motor reconocedor 454 está asociado con una colección de filtros 460, 462, 464, ..., 466 que comprende cada uno, información referente a los gestos, acción o condición que pueden ser realizados por cualquier persona u objeto detectable por el dispositivo de captura 20A o 20B. Por ejemplo, los datos procedentes del dispositivo de captura 20A pueden ser procesados por los filtros 460, 462, 464, ..., 466 para identificar cuándo un usuario o grupo de usuarios ha realizado uno o más gestos u otras acciones. Esos gestos pueden estar asociados con varios controles, objetos o condiciones de la aplicación 452. De este modo, el sistema de ordenador concentrador 12 puede utilizar el motor reconocedor 454, con los filtros, para interpretar y rastrear el movimiento de objetos (incluyendo personas).

Los dispositivos de captura 20A y 20B proporcionan imágenes RGB (o imágenes visuales en otros formatos o espacios de color) e imágenes de profundidad al sistema de ordenador concentrador 12. La imagen de profundidad puede ser una pluralidad de píxeles observados en donde cada píxel observado tiene un valor de profundidad observado. Por ejemplo, la imagen de profundidad puede incluir un área de píxeles de dos dimensiones (2-D) de la escena capturada en donde cada píxel en el área de píxeles 2-D puede tener un valor de profundidad tal como la distancia de un objeto en la escena capturada desde el dispositivo de captura. El sistema de ordenador concentrador 12 utilizará las imágenes RGB y las imágenes de profundidad para rastrear los movimientos de un usuario o de un objeto. Por ejemplo el sistema rastreará el esqueleto de una persona utilizando las imágenes de profundidad. Existen muchos métodos que pueden ser utilizados para rastrear el esqueleto de una persona utilizando imágenes de profundidad. Un ejemplo adecuado de rastreo de un esqueleto utilizando imágenes de profundidad se proporciona en la Solicitud de Patente de Estados Unidos 12/603.437, "Pose Tracking Pipeline" presentada el 21 de octubre de 2009, Craig, et al. (en lo que sigue referida como Solicitud '437). El proceso de la Solicitud '437 incluye adquirir una imagen de profundidad, muestrear los datos, eliminar y/o suavizar los datos de ruido de varianza elevada, identificar y retirar el fondo, y asignar cada uno de los píxeles de primer plano a diferentes partes del cuerpo. En base a estas etapas, el sistema encajará un modelo con los datos y creará un esqueleto. El esqueleto incluirá una serie de articulaciones y conexiones entre las articulaciones. También se pueden utilizar otros métodos para el rastreo. Tecnologías de rastreo adecuadas están también descritas en las siguientes Solicitudes de Patente de Estados Unidos: Solicitud de Patente de Estados Unidos 12/475.308, "Device for Identifying and Tracking Multiple Humans Over Time", presentada el 29 de Mayo de 2009; Solicitud de Patente de Estados Unidos 12/696.282, "Visual Based Identity Tracking", presentada el 29 de enero de 2010; Solicitud de Patente de Estados Unidos 12/641.788, "Motion Detection Using Depth Images," presentada el 18 de diciembre de 2009; and Solicitud de Patente de Estados Unidos 12/575.388, "Human Tracking System", presentada el 7 de octubre de 2009.

45 El motor reconocedor 454 incluye múltiples filtros 460, 462, 464, ..., 466 para determinar un gesto o acción. Un filtro comprende información que define un gesto, acción o condición junto con los parámetros, o metadatos, para ese gesto, acción o condición. Por ejemplo, un lanzamiento, que comprende movimiento de una de las manos desde detrás de la parte trasera del cuerpo para pasar la parte delantera del cuerpo, puede ser implementado como un gesto que comprende información que representa el movimiento de una de las manos del usuario desde detrás de la parte trasera del cuerpo para pasar la parte delantera del cuerpo, como ese movimiento sería capturado por la cámara de profundidad. Los parámetros pueden entonces ser establecidos para ese gesto. Cuando el gesto es un lanzamiento, un parámetro puede ser una velocidad umbral que la mano tiene que alcanzar, una distancia que la mano recorre (o bien absoluta o bien relativa al tamaño del usuario como un conjunto), y un índice de confianza por el motor reconocedor de que el gesto se ha producido. Estos parámetros del gesto pueden variar entre aplicaciones, entre contextos de una única aplicación, o dentro de un contexto de una aplicación a lo largo del tiempo.

Los filtros pueden ser modulares o intercambiables. En una realización, un filtro tiene un número de entradas (teniendo cada una de estas entradas un tipo) y un número de salidas (teniendo cada una de estas salidas un tipo). Un primer filtro puede ser remplazado por un segundo filtro que tenga el mismo número de tipos de entradas y salidas que el primer filtro sin alterar ningún otro aspecto de la arquitectura del motor reconocedor. Por ejemplo, puede haber un primer filtro para accionar que tiene como datos de esqueleto de entrada y salida una confianza de que el gesto asociado con el filtro está ocurriendo y un ángulo de giro. Cuando se desea sustituir este primer filtro de

accionamiento por un segundo filtro accionamiento -quizás debido a que el segundo filtro de accionamiento es más eficiente y requiere menos recursos de procesamiento -se puede hacer simplemente sustituyendo el primer filtro por el segundo filtro siempre y cuando segundo filtro tenga las mismas entradas y salidas - una entrada de tipo de datos de esqueleto, y dos salidas de tipo de confianza y tipo de ángulo.

- 5 Un filtro no necesita tener un parámetro. Por ejemplo, un filtro de "altura de usuario" que devuelve la altura del usuario puede no permitir que pueda ser devuelto ningún parámetro. Un filtro de "altura de usuario" alternativo puede tener parámetros ajustables - tales como tener en cuenta el calzado del usuario, el peinado, las prendas de la cabeza y postura en la determinación de la altura del usuario.

- 10 Las entradas a un filtro pueden comprender cosas tales como datos de articulación acerca de una posición de la articulación del usuario, ángulos formados por los huesos que se unen en la articulación, datos de color RGB procedentes de la escena, y la velocidad de cambio de un aspecto del usuario. La salida de un filtro puede comprender cosas tales como la confianza de que un gesto dado está siendo realizado, la velocidad a la que se hace el movimiento del gesto, y un momento en el que el movimiento de gesto se hace.

- 15 El motor reconocedor 454 puede tener un motor reconocedor base que proporciona funcionalidad a los filtros. En una realización, la funcionalidad que el motor que reconocedor 454 implementa incluye un archivo de entrada en el tiempo que rastrea gestos reconocidos y otra entrada, una implementación de Modelo de Hidden Markov (en donde el sistema modelado se supone que es un proceso Markov - en donde un estado presente encapsula cualquier información de estado pasado necesaria para determinar un estado futuro, de manera que ninguna otra información de estado pasado puede ser mantenida para este fin - con parámetros desconocidos, y parámetros ocultos son determinados a partir de datos observables), así como otra funcionalidad para resolver casos particulares de reconocimiento de gestos.

- 20 Los filtros 460, 462, 464, ..., 466 son cargados e implementados en la parte superior del motor reconocedor 454 y pueden utilizar servicios proporcionados por el motor reconocedor 454 a todos los filtros 460, 462, 464, ..., 466. En una realización, el motor reconocedor 454 recibe datos para determinar si cumple los requisitos de cualquier filtro 460, 462, 464, ..., 466. Dado que estos servicios proporcionados, tales como analizar sintácticamente la entrada, son proporcionados una vez por el motor reconocedor 454 en lugar de por cada filtro 460, 462, 464, ..., 466, tal servicio sólo necesita ser procesado una vez en un periodo de tiempo en lugar de una vez por filtro durante ese periodo, de manera que el procesamiento utilizado para determinar gestos se reduce.

- 25 La aplicación 452 puede utilizar los filtros 460, 462, 464, ..., 466 provistos del motor reconocedor 454, o puede proporcionar su propio filtro que se enchufa en el motor reconocedor 454. En una realización, todos los filtros tienen una interfaz común para habilitar esta característica de enchufe. Además, todos los filtros pueden utilizar parámetros, de manera que una única herramienta de gestos debajo puede ser utilizada para depurar y ajustar todo el sistema de filtro.

- 30 Se puede encontrar más información acerca del motor reconocedor 454 en la Solicitud de Patente de Estados Unidos 12/422,661, "Gesture Recognizer System Architecture", presentada el 13 de abril de 2009. Se puede encontrar más información acerca del reconocimiento de gestos en la Solicitud de Patente de Estados Unidos 12/391,150, "Standard Gestures," presentada el 23 de febrero de 2009; y en la Solicitud de Patente de Estados Unidos 12/474,655, "Gesture Tool" presentada el 29 de mayo de 2009.

- 35 La Fig. 7 ilustra una realización a modo de ejemplo de un sistema de ordenador que puede ser utilizado para implementar el sistema de ordenador concentrador 12. Como se muestra en la Fig. 7, la consola multimedia 500 tiene una unidad de procesamiento central (CPU) 501 que tiene un caché 502 de nivel 1, caché 504 de nivel 2, una ROM temporal (memoria de solo lectura) 506. El caché 502 de nivel 1 y el caché 504 de nivel 2 almacenan temporalmente datos y por tanto reducen el número de ciclos de acceso de memoria, mejorando con ello la velocidad y la producción de procesamiento. La CPU 501 puede estar dispuesta teniendo más de un núcleo, y de este modo, los cachés 502 y 504 de nivel 1 y de nivel 2 adicionales. La ROM temporal 506 puede almacenar un código ejecutable que es cargado durante una fase inicial de un proceso de arranque cuando se enciende la consola multimedia 500.

- 40 Una unidad de procesamiento de gráficos (GPU) 508 y un codificador de vídeo/códec de vídeo (codificador/descodificador) 514 forman una línea de procesamiento de vídeo para procesar gráficos de alta resolución y alta velocidad. Los datos son llevados desde la unidad de procesamiento de gráficos 508 hasta el codificador de vídeo/códec de vídeo 514 a través de un bus. La línea de procesamiento de vídeo envía los datos a un puerto A/V (audio/vídeo) 540 para la transmisión a una televisión u otra presentación. Un controlador de memoria 510 está conectado a la GPU 508 para facilitar acceso de procesador a varios tipos de memoria 512, tales como, pero no limitados a, RAM (memoria de acceso aleatorio).

- 45 La consola multimedia 500 incluye un controlador I/O 520, un controlador de gestión de sistema 522, una unidad de procesamiento de audio 523, un controlador de interfaz de red 524, un primer controlador de ordenador principal USB 526, un segundo controlador USB 528 y un subconjunto I/O de panel frontal 530 que están preferiblemente implementados en un módulo 518. Los controladores USB 526 y 528 sirven como ordenadores principales para los

5 controladores periféricos 542(1)-542(2), un adaptador inalámbrico 548, y un dispositivo de memoria externo 546 (por ejemplo, memoria temporal, accionador CD/DVD ROM externo, medios retirables, etc.). La interfaz de red 524 y/o el adaptador inalámbrico 548 proporcionan acceso a una red (por ejemplo, internet, una red domestica, etc.) y pueden ser cualquiera de una amplia variedad de diversos componentes que adaptador con cable o inalámbricos que incluyen una tarjeta de Ethernet, un módulo de Bluetooth, un módem de cable, y similares.

10 La memoria del sistema 543 está dispuesta para almacenar datos de aplicación que son cargados durante el proceso de arranque. Un accionador de medios 544 está dispuesto y puede comprender un accionador DVD/CD, un accionador Blu-Ray, un accionador de disco duro, u otro accionador de medios retirable, etc. El accionador de medios 144 puede ser interno o externo respecto a la consola multimedia 500. Se puede acceder a los datos de aplicación a través del accionador de medios 544 para la ejecución, reproducción, etc. por la consola multimedia 500. El accionador de medios 544 está conectado al controlador I/O 520 a través de un bus, tal como un bus de la serie ATA u otra conexión de alta velocidad (por ejemplo, IEEE 1394).

15 El controlador de gestión del sistema 522 proporciona una variedad de funciones de servicio relacionadas con la capacidad de asegurar de la consola multimedia 500. La unidad de procesamiento de audio 523 y el códec de audio 532 forman una correspondiente línea de procesamiento de audio con procesamiento estéreo y de alta fidelidad. Los datos de audio son llevados entre la unidad de procesamiento de audio 523 y el códec de audio 532 a través de un enlace de comunicación. La línea de procesamiento de audio envía los datos al puerto A/V 540 para la reproducción por un usuario o dispositivo de audio externo que tenga capacidades de audio.

20 El subconjunto I/O de pared delantero 530 soporta la funcionalidad del botón de potencia 550 y el botón de expulsión 552, así como cualesquiera LEDs (diodos de emisión de luz) u otros indicadores expuestos en la superficie exterior de la consola multimedia 100. El módulo de suministro de potencia de sistema 536 proporciona potencia a los componentes de la consola multimedia 100. Un ventilador 538 enfría los circuitos dentro de la consola multimedia 500.

25 La CPU 501, la GPU 508, el controlador de memoria 510, y diversos otros componentes dentro de la consola multimedia 500 están interconectados a través de uno o más buses, que incluyen buses en serie y paralelos, un bus de memoria, un bus periférico, un procesador o bus local que utilizan a cualquiera de una variedad de arquitecturas de bus. A modo de ejemplo, tales arquitecturas pueden incluir un bus de Interconexión de Componente Periférico (PCI), un bus PCI-Express, etc.

30 Cuando se enciende la consola multimedia 500, los datos de aplicación pueden ser cargados desde la memoria del sistema 543 en la memoria 512 y/o los cachés 502, 504 y ejecutados en la CPU 501. La aplicación puede presentar una interfaz de usuario gráfica que proporciona una experiencia de usuario consistente cuando se navega a diferentes tipos de medios disponibles en la consola multimedia 500. En funcionamiento, las aplicaciones y/u otros medios contenidos dentro del accionador de medios 544 pueden ser lanzados o reproducidos desde el accionador de medios 544 para proporcionar funcionalidades adicionales a la consola multimedia 500.

35 La consola multimedia 500 puede ser operada como un sistema independiente simplemente conectando el sistema a una televisión otro medio de presentación. En este modo independiente, la consola multimedia 500 permite que uno o más usuarios interactúen con el sistema, vean películas, o escuchen música. Sin embargo, con la integración de la conectividad de banda ancha hecha disponible a través del interfaz de red 524 o el adaptador inalámbrico 548, la consola multimedia 500 puede ser además operada como un participante en una comunidad de red más grande. Adicionalmente, la consola multimedia 500 puede comunicarse con la unidad de procesamiento 4 a través de un adaptador inalámbrico 548.

45 Cuando se enciende la consola multimedia 500, una cantidad de conjuntos de recursos de hardware son reservados para el uso del sistema por el sistema de operación de consola multimedia. Esto recursos pueden incluir una reserva de memoria, un ciclo CPU y GPU, anchura de banda de vez, etc. Dado que estos recursos son reservados en el momento de arranque del sistema, los recursos reservados no existen desde el punto de vista de la aplicación. En particular, la reserva de memoria preferiblemente es lo suficientemente grande para contener la cuestión del lanzamiento, aplicaciones de sistema concurrentes y accionadores. La reserva de CPU es preferiblemente constante de manera que si el uso de CPU reservada no se utiliza por las aplicaciones del sistema, un hilo inactivo consumirá cualesquiera ciclos no utilizados.

50 Con respecto a la reserva de GPU, los mensajes de peso ligero generados por las aplicaciones del sistema (por ejemplo, mensajes surgentes) son presentados utilizando una interrupción GPU para programar un código para realizar mensajes surgentes en una transparencia. La cantidad de memoria requerida para una superposición depende del tamaño del área de superposición y la superposición preferiblemente se escala con la resolución de la pantalla. Cuando una interfaz de usuario total es utilizada por la aplicación del sistema concurrente, es preferible utilizar una resolución independiente de la resolución de aplicación. Se puede utilizar un escalador para establecer esta resolución de manera que se elimine la necesidad de cambiar la frecuencia y se produzca una resincronización de TV.

Después de que la consola multimedia 500 arranque y los recursos del sistema sean reservados, las aplicaciones

del sistema concurrentes se ejecutan para proporcionar funcionalidades del sistema. Las funcionalidades del sistema son encapsuladas en un conjunto de aplicaciones de sistema que se ejecutan dentro de los recursos del sistema reservados anteriormente descritos. El núcleo del sistema de operación identifica hilos que son hilos de aplicación del sistema frente a hilos de aplicación de juego. Las aplicaciones del sistema son preferiblemente programadas para ser ejecutadas en la CPU 501 en unos momentos e intervalos predeterminados con el fin de proporcionar una vista del recurso de sistema consistente a la aplicación. La programación es para minimizar la perturbación para la aplicación de juego que se ejecuta en la consola.

Cuando una aplicación del sistema concurrente requiere audio, el procesamiento del audio se programa de manera asíncrona respecto a la aplicación de juego debido a la sensibilidad de tiempo. Un gestor de aplicación de consola multimedia (descrito más adelante) controla el nivel de audio de aplicación de juego (por ejemplo, silencio, atenuación) cuando las aplicaciones del sistema están activas.

Los dispositivos de entrada opcionales (por ejemplo, los controladores 542(1) y 542(2)) son compartidos por las aplicaciones de juego y por las aplicaciones del sistema. Los dispositivos de entrada no son recursos reservados, pero pueden ser conmutados entre aplicaciones de sistema y aplicaciones de juego, de manera que cada uno tendrá un foco del dispositivo. El gestor de aplicación preferiblemente controla la conmutación de la cadena de entrada, sin saber el conocimiento de la aplicación de juego y el accionador mantiene la información de estado referente a los conmutadores de foco. Los dispositivos de captura 20A y 20B pueden definir dispositivos de entrada para la consola 500 a través del controlador USB 526 u otra interfaz. En otras realizaciones, el sistema de ordenador concentrador 12 puede ser implementado utilizando otras arquitecturas de hardware. No se requiere ninguna arquitectura de hardware.

El sistema de ordenador a modo de ejemplo ilustrado en las Figuras 4, 5, 6 y 7 incluye ejemplos de medios de almacenamiento leíbles por ordenador. Tales medios pueden incluir medios volátiles y no volátiles, retirables y no retirables implementados en cualquier método o tecnología para el almacenamiento de información tal como instrucciones leíbles por ordenador, estructuras de datos, módulos de programa u otros datos. El medio de almacenamiento de ordenador incluye, pero no se limita a, RAM, ROM, EEPROM, memoria temporal u otra tecnología de memoria, CD-ROM, discos versátiles digitales (DVD) u otro almacenamiento de disco óptico, tarjetas o pinchos de memoria, casetes magnéticos, cintas magnéticas, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnéticos, o cualquier otro medio que pueda ser utilizado para almacenar la información deseada y al que se pueda acceder mediante un ordenador. Algunos ejemplos de medios de almacenamiento leíbles por ordenador en las Figuras incluyen la memoria 214 de la Fig. 4, la memoria temporal 334, la memoria 330, la RAM 326 y el caché 324 de la Fig. 5, la memoria 434 de la Fig. 6, y la ROM 506, los cachés 502, 504, la memoria 512, la memoria del sistema 543, la unidad de memoria 546 y el accionador de medios 544.

La Fig. 1 muestra un dispositivo de presentación montado en la cabeza 2 y una unidad de procesamiento 4 (en conjunto referidos como dispositivo de presentación móvil) en comunicación con el dispositivo de ordenador concentrador 12 (referido como un concentrador). En otra realización, múltiples dispositivos de presentación móvil pueden estar en comunicación con un único concentrador. Cada uno de los dispositivos de presentación móvil se comunicará con el concentrador utilizando comunicación inalámbrica, como se ha descrito anteriormente. En tal realización, se contempla que mucha de la información que es útil para todos los dispositivos de presentación móviles será calculada y almacenada en el concentrador y transmitida a cada uno de los dispositivos de presentación móviles. Por ejemplo, el concentrador generará el modelo del entorno y proporcionará ese modelo a todos los dispositivos de presentación móviles en comunicación con el concentrador. Adicionalmente, el concentrador puede rastrear la ubicación y orientación de los dispositivos de presentación móviles y de los objetos en movimiento en la habitación, y después transferir esa información a cada uno de los dispositivos de presentación móviles.

En otra realización, un sistema de múltiples usuarios que puede variar el foco del contenido virtual que va a estar en foco para un usuario podría incluir múltiples concentradores, incluyendo cada concentrador uno o más dispositivos de presentación móviles. Los concentradores se pueden comunicar entre sí o a través de internet (u otras redes). Por ejemplo, la Fig. 8 muestra concentradores 560, 562 y 564. El concentrador 560 se comunica directamente con el concentrador 562. El concentrador 560 se comunica con el concentrador 564 a través de Internet. El concentrador 560 está comunicado con los dispositivos de presentación móviles 570, 572, ..., 574. El concentrador 562 se comunica con el dispositivo de presentación móvil 578, 580, ..., 582. El concentrador 564 se comunica con el dispositivo de presentación móvil 584, 586, ..., 588. Cada uno de los dispositivos de presentación móviles se comunica con su respectivo concentrador mediante comunicación inalámbrica como se ha descrito anteriormente. Si estos concentradores están en un entorno común, después cada uno de los concentradores puede proporcionar una parte del modelo de los entornos, o un concentrador puede crear el modelo para los otros concentradores. Cada uno de los concentradores rastrea a un subconjunto de objetos en movimiento y compartirá esa información con los otros concentradores, que a su vez que compartirán la información con los dispositivos de presentación móviles apropiados. La información de sensor de los dispositivos de presentación móviles será proporcionada a sus respectivos concentradores y después compartida con otros concentradores para compartir finalmente con otros dispositivos de presentación móviles. De este modo, la información compartida entre los concentradores puede incluir rastreo de esqueleto, información acerca de modelos, diversos estados de aplicaciones, y otros rastreos. La información comunicada entre los concentradores y sus respectivos dispositivos de presentación móviles incluye información de rastreo de objetos en movimiento, actualizaciones de estado y físicas para los modelos reales,

información de geometría y textura, vídeo y audio, y otra información utilizada para realizar las operaciones descritas en la presente memoria.

La Fig. 9 es un diagrama de flujo que describe una realización de un proceso para presentar un objeto virtual que va a estar en foco cuando es visto por el usuario que está realizado mediante los componentes descritos anteriormente. En la etapa 602, es configurado el sistema. Por ejemplo, una aplicación (por ejemplo, la aplicación 452 de la Fig. 6) puede configurar el sistema para indicar que una imagen virtual particular (que representa un objeto virtual) va ser insertada en el modelo de tres dimensiones de la escena, en una ubicación deseada. Se considera un ejemplo en el que una habitación vacía está poblada con objetos virtuales para convertir la habitación vacía en una realidad virtual. En otro ejemplo, la aplicación que se ejecuta en el sistema de ordenador concentrador 12 indicará que una imagen virtual particular (que representa un objeto virtual) va ser insertada en la escena como parte de un videojuego o de otro proceso.

En la etapa 604, el sistema creará un modelo volumétrico del espacio para el que está situado el dispositivo de presentación montado en la cabeza 2. En una realización, por ejemplo, el dispositivo de ordenador concentrador 12 utilizará imágenes de profundidad procedentes de una o más cámaras de profundidad para crear un modelo tridimensional del entorno o del espacio en el que el dispositivo de presentación montado en la cabeza 2 está situado. En las etapas 606, ese modelo es segmentado en uno o más objetos. Por ejemplo, si el dispositivo de ordenador concentrador 12 crea un modelo tridimensional de una habitación, es probable que esa habitación tenga múltiples objetos en ella. Ejemplos de objetos que pueden estar en una habitación incluyen personas, sillas, mesas, sofás, etc. La etapa 606 incluye distinguir unos objetos de otros. En la etapa 608, el sistema identificará los objetos. Por ejemplo, el dispositivo de ordenador concentrador 12 puede identificar que un objeto particular es una mesa y otro objeto que es una silla. En la etapa 610, el sistema presentará un objeto virtual para estar en foco cuando es visto por un usuario. El objeto virtual aparece en foco como si fuera un objeto real que se ve realmente y directamente a través de las lentes transparentes, por ejemplo 116, 118s. Esto es, el sistema presenta una imagen virtual en el campo de visión del usuario a la profundidad focal que el usuario está viendo mientras el usuario está mirando a través del dispositivo de presentación montado en la cabeza 2. De este modo, la imagen virtual aparece en su ubicación del mundo real en un dispositivo de presentación que permite visualizar de forma real y directa al menos una parte del entorno físico a través de dicha presentación. El objeto virtual puede ser un objeto estacionario o un objeto en movimiento. En la etapa 612, el usuario del dispositivo de presentación montado en la cabeza 2 interactuará con una aplicación que se ejecuta en un dispositivo de ordenador concentrador 12 (u otro dispositivo de ordenador) en base al objeto virtual que está siendo presentado en el dispositivo de presentación montado en la cabeza 2. Cada una de las etapas 604-610 será descrita con más detalle más adelante.

La Fig. 10 es un diagrama de flujo que describe una realización de un proceso para crear un modelo tridimensional del espacio. Por ejemplo, en proceso de la Fig. 10 es una implementación a modo de ejemplo de la etapa 604 de la Fig. 9. En la etapa 640, el sistema de ordenador concentrador 12 recibe una o más imágenes de profundidad para múltiples perspectivas del espacio en el que está el dispositivo de presentación montado en la cabeza 2. Por ejemplo, el dispositivo de ordenador concentrador 12 puede obtener imágenes de profundidad procedentes de las múltiples cámaras de profundidad, o múltiples imágenes de profundidad procedentes de la misma cámara apuntando la cámara en diferentes direcciones o utilizando una cámara de profundidad con una lente que permita una visión total del espacio para el cual el modelo será construido. En la etapa 642, los datos de profundidad procedentes de las distintas imágenes de profundidad son combinados en base a un sistema de coordenadas común. Por ejemplo, si este sistema recibe imágenes de profundidad procedentes de múltiples cámaras, el sistema correlacionará las dos imágenes que tengan un sistema de coordenadas común (por ejemplo alineando las imágenes). En la etapa 644, se crea una descripción volumétrica del espacio utilizando los datos de profundidad. En la etapa 646, el sistema detectará y rastreará objetos en movimiento tales como personas que se mueven en la habitación. Más detalles acerca del rastreo de personas se han descrito anteriormente. En la etapa 648, la descripción volumétrica del espacio es actualizada en base al rastreo de los objetos en movimiento. Las etapas 646 y 648 se pueden realizar de forma continua durante el funcionamiento del sistema, de manera que el modelo es actualizado de forma continua teniendo en cuenta los objetos en movimiento. En otra realización, el modelo sólo incluirá información acerca de objetos estacionarios y los objetos en movimiento son rastreados de forma separada.

La Fig. 11 es un diagrama de flujo que describe una realización de un proceso para segmentar el modelo del espacio que en objetos. Por ejemplo, el proceso de la Fig. 11 es una implementación a modo de ejemplo de la etapa 606 de la Fig. 9. En la etapa 680 de la Fig. 11, el sistema recibirá una o más imágenes de profundidad desde una o más cámaras de profundidad como se ha descrito anteriormente. Alternativamente, el sistema puede acceder a una o más imágenes de profundidad que ya ha recibido. En la etapa 682, el sistema recibirá una o más imágenes visuales procedentes de las cámaras descritas anteriormente. Alternativamente, el sistema puede acceder a una o más imágenes visuales ya recibidas. En la etapa 684, el sistema de ordenador concentrador detectará una o más personas en base a las imágenes de profundidad y/o a las imágenes visuales. Por ejemplo, el sistema reconocerá uno o más esqueletos. En la etapa 686, el dispositivo de ordenador concentrador detectará bordes dentro del modelo en base a las imágenes de profundidad y/o a las imágenes visuales. En la etapa 688, el dispositivo de ordenador concentrador utilizará los bordes detectados para distinguir unos objetos de otros. Por ejemplo, se supone que los bordes son límites entre objetos. En la etapa 690, el modelo creado utilizando el proceso de la Fig. 10 será actualizado para mostrar qué partes del modelo están asociadas con los diferentes objetos.

La Fig. 12 es un diagrama de flujo que describe una realización de un proceso para identificar objetos. Por ejemplo el proceso de la Fig. 12 es una implementación a modo de ejemplo de la etapa 608 de la Fig. 9. En la etapa 710, el dispositivo de ordenador concentrador 12 hará coincidir personas identificadas con identidades de usuario. Por ejemplo, el sistema puede tener perfiles de usuario que tengan imágenes visuales que puedan ser hechas coincidir con las imágenes detectadas de los objetos. Alternativamente, un perfil de usuario puede describir características de la persona que pueden corresponder en base a las imágenes de profundidad o imágenes visuales. En otra realización, los usuarios pueden registrarse en el sistema y un dispositivo de ordenador concentrador 12 puede utilizar el proceso de acceso para identificar un usuario particular y rastrear ese usuario a través de la interacción descrita la presente memoria. En la etapa 712, el dispositivo de ordenador concentrador 12 accederá a la base de datos de las formas. En la etapa 714, el dispositivo de ordenador concentrador hará coincidir tantos objetos en el modelo con las formas en la base de datos. En la etapa 716, aquellas formas que no coincidan serán destacadas y presentadas al usuario (por ejemplo, utilizando el monitor 16). En la etapa 718, el dispositivo de ordenador concentrador 12 recibirá la entrada del usuario que identifica cada (o un subconjunto) de las formas destacadas. Por ejemplo, el usuario puede utilizar un teclado, un ratón, una entrada de voz, u otro tipo de entrada para indicar qué forma no identificada es cada una. En la etapa 720, la base de datos de las formas es actualizada en base a la entrada del usuario en la etapa 718. En la etapa 722, el modelo de entorno creado en la etapa 604, y actualizado en la etapa 606, es actualizado adicionalmente añadiendo metadatos para cada uno de los objetos. Los metadatos identifican el objeto. Por ejemplo, los metadatos pueden indicar que el objeto particular es una mesa brillante redonda, John Doe, un sofá de cuero verde, etc.

La Fig. 13 es un diagrama de flujo que describe una realización de un proceso para presentar un objeto virtual que va a estar en foco cuando es visto por un usuario que ve una presentación transparente. Por ejemplo, el proceso de la Fig. 13 es una implementación a modo de ejemplo de la etapa 610 de la Fig. 9. En la etapa 950 de la Fig. 13, el sistema determina el campo de visión del usuario. Esto es, el sistema determina a qué parte del entorno por espacio está mirando el usuario. En una realización, es un esfuerzo colaborativo utilizando el dispositivo de ordenador concentrador 12, la unidad de procesamiento 4 y el dispositivo de presentación montado en la cabeza 2. En una implementación a modo de ejemplo, el dispositivo de ordenador concentrador 12 rastreará el usuario y el dispositivo de presentación montado en la cabeza 2 con el fin de proporcionar una determinación preliminar de la localización y orientación del dispositivo de presentación montado en la cabeza 2. Los sensores sobre el dispositivo de presentación montado en la cabeza 2 serán utilizados para perfeccionar la orientación determinada. Por ejemplo, los sensores de inercia 34, descritos anteriormente, se pueden usar para perfeccionar la orientación del dispositivo de presentación montado en la cabeza 2. Adicionalmente, el proceso de rastreo de ojos descrito anteriormente puede ser utilizado para identificar un subconjunto del campo de visión inicialmente determinado que corresponde a donde en particular un usuario está mirando conocido de otro modo como la región focal del usuario o foco de profundidad en el campo de visión. Más detalles serán descritos más adelante respecto a las Figuras 14, 15 y 16.

Los tipos de objetos virtuales y su ubicación en un campo de visión del usuario se determinan mediante la aplicación 452 basada en parámetros por defecto, de entrada de usuario, o una combinación de ambos. Por ejemplo, el usuario puede haber seleccionado un objeto del mundo real de una mesa de café en la habitación en donde se está ejecutando la aplicación 452 para que aparezca como una roca. La ubicación objetivo de la roca virtual será relacionada con la ubicación de la mesa de café en el modelo tridimensional. En otro ejemplo, un delfín virtual puede ser seleccionado por un usuario para nadar alrededor de la habitación. El movimiento del delfín virtualmente alrededor de la habitación puede ser implementado por recorrido de trayectoria del objeto del delfín en que el que una ubicación objetivo para el delfín es actualizada en cada fotograma de presentación. Incluso si el usuario mira fijamente a la roca, el delfín probablemente entrará y saldrá de la región focal del usuario y del campo de visión generalmente. A medida que el usuario mueve sus ojos o cabeza, el campo de visión del usuario y la región focal actual son actualizadas con estos movimientos también. Uno o más objetos virtuales pueden estar situados en la región focal del usuario y en el campo de visión del usuario en cualquier momento dado.

Después de determinar el campo de visión en la etapa 950 anterior, el sistema determina si hay alguna ubicación objetivo de objetos virtuales en el campo de visión actual del usuario en la etapa 952. El software que se ejecuta en uno o más sistemas de ordenador tal como el dispositivo de ordenador concentrador 12 o en la unidad de procesamiento 4 identificarán la ubicación del objetivo en el campo de visión. En una realización, el dispositivo de ordenador concentrador 12 proporcionará el modelo a la unidad de procesamiento 4. Como parte de la etapa 952, la unidad de procesamiento 4 utilizará el modelo del entorno y el conocimiento de la posición y orientación del usuario para determinar si la ubicación objetivo de cualquier objeto virtual está dentro del campo de visión del usuario.

Si no hay objetos virtuales en el campo de visión del usuario actual, el procesamiento en la etapa 966 vuelve para determinar y actualizar el campo de visión del usuario en la etapa 950.

Si hay al menos un objeto virtual que tenga una ubicación objetivo en el campo de visión del usuario, entonces en la etapa 954, el sistema tal como el software que se ejecuta en la unidad de procesamiento 4, determina la región focal actual del usuario dentro del campo de visión del usuario. Como se describe con más detalle más adelante en las Figuras 15 y 16, el procesamiento de rastreo de ojos basado en los datos capturados por la cámara de rastreo de ojos 134 para cada ojo, puede proporcionar la región focal actual del usuario. Por ejemplo, la convergencia entre las pupilas con los datos que indican la posición de la cara del usuario pueden ser utilizados para triangular a un punto focal en una curva focal, el Horóptero, desde el que la región focal, el área fusional de Panum se pueden calcular. El

área fusional de Panum es el área de visión sencilla para estereopsis binocular utilizada por los ojos humanos.

En la etapa 956, el software que se ejecuta en el sistema de ordenador concentrador 12, la unidad de procesamiento 4 o ambos identifican cuál de los objetos virtuales en el campo de visión del usuario está en la región focal del usuario actual en base a la ubicación objetivo del objeto virtual en el modelo.

- 5 En la etapa 958, la unidad de procesamiento 4 escalará y orientará los objetos virtuales para una imagen que va ser insertada en la visión del usuario. El escalado y la orientación de la imagen virtual estarán basados en la ubicación del objetivo en el campo de visión y la forma conocida del objeto virtual en base al modelo.

- 10 En la etapa 960, la unidad de procesamiento 4, los circuitos de control 136 o ambos que comparten procesamiento para colocar cada objeto virtual identificado en la región focal del usuario actual cambiando la región focal del conjunto de micropresentación. En la etapa 962, una profundidad artificial de la técnica de campo se aplica a objetos virtuales cuyas ubicaciones objetivo están dentro del campo de visión del usuario pero fuera de la región focal actual del usuario como una función de la distancia desde la región focal actual. Un ejemplo de una profundidad artificial de la técnica de campo es una técnica de borrosidad artificial.

- 15 La borrosidad artificial se puede conseguir aplicando una profundidad de oscurecedor de campo u otro filtro de borrosidad Gaussiano para simular que el objeto está fuera del foco como función de la distancia desde la región focal. La técnica se puede realizar total o parcialmente mediante un software que se ejecuta en un ordenador concentrador 12, en la unidad de procesamiento 4 o en ambos. Para facilitar la restricción, se hará referencia a la unidad de procesamiento 4. A partir de un valor de profundidad como parte de la ubicación objetivo de un objeto virtual, la distancia focal del objeto en la imagen se determina mediante la unidad de procesamiento 4 que también determina qué píxeles en un elemento de presentación 112 como ese que encierra el elemento óptico de guiado de luz 112 mapearán los objetos virtuales en una imagen. En base a la distancia focal, uno o más filtros de borrosidad Gaussianos ponderados son aplicados para cubrir al menos dos dimensiones convolucionando la imagen con una función Gaussiana tal como las transformaciones aplicadas a los píxeles desde el punto focal o punto de fijación de la imagen recibirá más efectos de borrosidad. En un ejemplo, el filtro de borrosidad Gaussiano actúa como un filtro del paso bajo que elimina la información de elevada frecuencia.
- 20
- 25

- En la etapa 964, se presenta una imagen virtual que incluye los objetos virtuales en el campo de visión del usuario. En un ejemplo, la unidad de procesamiento 4 envía instrucciones al accionador de presentación 220 de los circuitos de control 136 para la presentación sobre la micropresentación 120 de la imagen virtual que incluye los objetos virtuales en el campo de visión del usuario. El sistema de lentes 122 entonces proyecta la imagen virtual recibida de la micropresentación 120 sobre la superficie reflectante 124 y hacia los ojos del usuario o dentro del elemento óptico de guiado de luz 112 para el visionado por el usuario. En una implementación, la presentación en la que el usuario está mirando a través del dispositivo de presentación montado en la cabeza (por ejemplo, el elemento óptico de guiado de luz 112) es dividida en píxeles. La etapa 964 puede incluir determinar qué píxeles corresponden a la ubicación objetivo, y esos píxeles presentarán las imágenes virtuales en la etapa 964. El filtro de opacidad 114 puede ser utilizado para evitar efectos visuales no reales. Por ejemplo, el filtro de opacidad 114 puede modificar la luz a los píxeles en el elemento óptico de guiado de luz 112, de manera que un objeto virtual de fondo no se ve a través de un objeto virtual en primer plano. El procesamiento vuelve a la etapa 950 y determina el campo de visión del usuario en este sistema de presentación en tiempo real. Las etapas de procesamiento de la Fig. 13 pueden ser realizadas de forma continua durante el funcionamiento del sistema, de manera que el campo de visión del usuario y la región focal del usuario son actualizados cuando el usuario mueve su cabeza, y los objetos virtuales pueden ser presentados, por consiguiente, con un movimiento natural que entra y sale del foco para el usuario.
- 30
- 35
- 40

- La Fig. 16 es un diagrama de flujo que describe una realización de un proceso para determinar el campo de visión del usuario, que es una implementación a modo de ejemplo de la etapa 950 de la Fig. 13. El proceso de la Fig. 16 se basa en la información procedente del dispositivo de ordenador concentrador 12 y la tecnología de rastreo de ojos descrita anteriormente. La Fig. 14 es un diagrama de flujo que describe una realización de un proceso realizado por el sistema de ordenador concentrador para proporcionar información de rastreo utilizada en el proceso de la Fig. 16. La Fig. 15 es un diagrama de flujo que describe una realización de un proceso para rastrear un ojo, cuyos resultados son utilizados por el proceso de la Fig. 16.
- 45

- En la etapa 810 de la Fig. 14, el dispositivo de ordenador concentrador 12 rastreará la posición del usuario. Por ejemplo, el dispositivo de ordenador concentrador 12 utilizará una o más imágenes de profundidad y una o más imágenes visuales para rastrear a un usuario (por ejemplo, utilizando rastreo de esqueleto). Una o más imágenes de profundidad y una o más imágenes visuales pueden ser utilizadas para determinar la posición del dispositivo de presentación montado en la cabeza 2 y la orientación del dispositivo de presentación montado en la cabeza 2 en la etapa 812. En la etapa 814, la posición y orientación del usuario y del dispositivo de presentación montado en la cabeza 2 son transmitidas desde el dispositivo de ordenador concentrador 12 a la unidad de procesamiento 4. En la etapa 816, la información de posición y orientación es recibida en la unidad de procesamiento 4. Las etapas de procesamiento de la Fig. 14 pueden ser realizadas de forma continua durante el funcionamiento del sistema de manera que el usuario es rastreado de forma continua.
- 50
- 55

La Fig. 15 es un diagrama de flujo que describe una realización para rastrear un ojo utilizando la tecnología descrita

anteriormente. En la etapa 860, el ojo es iluminado. Por ejemplo, el ojo puede ser iluminado utilizando luz infrarroja procedente de la iluminación de rastreo de ojos 144A. En la etapa 862, la reflexión desde el ojo es detectada utilizando una o más cámaras de rastreo de ojos 134B. En la etapa 864, los datos de reflexión son enviados desde el dispositivo de presentación montado en la cabeza 2 a la unidad de procesamiento 4. En la etapa 866, la unidad de procesamiento 4 determinará la posición del ojo en base a los datos de reflexión, como se ha descrito anteriormente. En la etapa 868, la unidad de procesamiento 4 también determinará la ubicación de profundidad focal o región focal que los ojos del usuario están viendo en base a los datos de reflexión. Las etapas de procesamiento de la Fig. 15 pueden ser realizadas de forma continua durante el funcionamiento del sistema, de manera que los ojos del usuario son rastreados de forma continua proporcionando datos para rastrear la región focal actual del usuario.

La Fig. 16 es un diagrama de flujo que describe una realización de un proceso para determinar el campo de visión (por ejemplo, la etapa 950 de la Fig. 13). En la etapa 902, la unidad de procesamiento 4 accederá a la información de última posición y orientación recibida desde el concentrador. El proceso de la Figura 14 puede ser realizado de forma continua como se muestra mediante la flecha desde la etapa 814 a la etapa 810, por lo tanto, la unidad de procesamiento 4 recibirá de forma periódica la información de posición y orientación actualizada procedente del dispositivo de ordenador concentrador 12. Sin embargo, la unidad de procesamiento 4 necesitará extraer la imagen virtual de forma más frecuente de la que recibe la información actualizada desde el dispositivo de ordenador concentrador 12. Por lo tanto, la unidad de procesamiento 4 necesitará basarse en la información detectada localmente (por ejemplo, desde el dispositivo montado en la cabeza 2) para proporcionar actualizaciones de la orientación entre muestras procedentes del dispositivo de ordenador concentrador 12. En la etapa 904, la unidad de procesamiento 4 accederá a datos procedentes del giroscopio de tres ejes 132B. En la etapa 906, la unidad de procesamiento 4 accederá a datos procedentes del acelerómetro de tres ejes 132C. En la etapa 908, la unidad de procesamiento 4 accederá a datos procedentes del magnetómetro de tres ejes 132A. En la etapa 910, la unidad de procesamiento 4 perfeccionará (o actualizará de otro modo), los datos de posición y orientación procedentes del dispositivo de ordenador concentrador 12 con los datos procedentes del giroscopio, acelerómetro y magnetómetro. En la etapa 912, la unidad de procesamiento 4 determinará el campo de visión potencial en base a la localización y orientación del dispositivo de presentación montado en la cabeza.

En la etapa 914, la unidad de procesamiento 4 accederá a la última información de posición de ojos. En la etapa 916, la unidad de procesamiento 4 determinará una parte del modelo que está siendo visto por el usuario, como un subconjunto del campo de visión potencial en base a la posición del ojo. Por ejemplo, el usuario puede estar vuelto hacia una pared, y por tanto, el campo de visión de la presentación montada en la cabeza podría incluir cualquier lugar a lo largo de la pared. Si embargo, si los ojos del usuario están apuntando la derecha, entonces la etapa 916 concluirá en que el campo de visión del usuario es sólo la parte derecha de la pared. Al final de la etapa 916, la unidad de procesamiento 4 ha determinado el campo de visión del usuario a través de la presentación montada en la cabeza 2. La unidad de procesamiento 4 puede identificar una ubicación objetivo de un objeto virtual dentro de ese campo de visión. En la etapa 918, la unidad de procesamiento 4 determina la región focal del usuario actual con la parte del modelo basada en la posición del ojo. La unidad de procesamiento 4 puede identificar una ubicación objetivo en el modelo de un objeto virtual que está dentro de la región focal del usuario actual. Las etapas de procesamiento de la Fig. 16 se pueden realizar de forma continua durante el funcionamiento del sistema de manera que el campo de visión del usuario y la región focal son actualizadas de forma continua cuando el usuario mueve la cabeza, y por consiguiente, los objetos virtuales pueden ser presentados el movimiento natural entrando y saliendo del foco para el usuario.

La Fig. 17A es un diagrama de flujo que describe una realización de un proceso para cambiar una región focal de un conjunto de micropresentación desplazando al menos una lente del conjunto. Esta realización puede ser utilizada para implementar la etapa 960 en la Fig. 13. Para un contexto ilustrativo, está referenciado un conjunto de micropresentación como se ha descrito anteriormente con respecto a las Figuras 2A y 2B. El conjunto de micropresentación comprende la unidad de micropresentación 120 alineada en una trayectoria óptica 133 con un sistema de lentes 122 que colima la luz de la imagen procedente de la unidad de micropresentación 120 hasta el elemento de reflexión 124 o los elementos de reflexión 124a, 124b y o bien en los ojos del usuario o bien en otro elemento óptico 112. El conjunto de micropresentación también incluye el ajustador de foco virtual variable 135 para mover físicamente uno de los elementos de procesamiento de luz para obtener una región focal seleccionada o determinada.

Antes de las exposiciones de las Figuras 3A-3D, los desplazamientos del conjunto de micropresentación y las longitudes focales están referidas por la siguiente ecuación $1/S_1 + 1/S_2 = 1/f$. En la etapa 980, la unidad de procesamiento 4 determina un valor de desplazamiento S_1 , entre un punto nodal delantero del sistema de lentes 122 y la ubicación objetivo del objeto virtual. Adicionalmente, en la etapa 982, la unidad de procesamiento 4 determina un valor de desplazamiento, S_2 , entre el elemento reflectante 124, 124a y el punto nodal trasero del sistema de lentes para una longitud focal del sistema de lentes, f , en base a la ecuación. La unidad de procesamiento 4 hace, en la etapa 984, que el procesador de los circuitos de control 136 haga que el accionador ajustador variable 237 aplique al menos una señal de accionamiento al ajustador de foco virtual variable 135 para mover físicamente al menos una lente del sistema de lentes para generar los valores del desplazamiento determinados S_1 y S_2 . Además del conjunto de micropresentación mostrado en las Figuras 2A y 2B, otra implementación es un conjunto de micropresentación que utiliza el ejemplo de sistema de lentes insertable de la Fig. 3D. El movimiento es el de liberar un brazo 123 para una lente si colocando en su sitio y moviendo un brazo 123 para la lente cuando el

desplazamiento determinado bloquea la lente en la trayectoria óptica 133.

La Fig. 17B es un diagrama de flujo que describe una realización de un proceso para cambiar una región focal de un conjunto de micropresentación cambiando la polarización de al menos una lente del conjunto. La realización puede ser utilizada para implementar la etapa 960 en la Fig. 13 y es adecuada para utilizar con la realización de un conjunto de micropresentación tal como el mostrado en la Fig. 3C. En esta realización, el sistema de lente incluye al menos una lente birrefringente en la trayectoria óptica entre la micropresentación 120 y el elemento reflectante 124 (Fig. 2A), 124a (Fig. 2B). En la etapa 986, la unidad de procesamiento 4 selecciona un valor de desplazamiento para S_1 en base a una ubicación objetivo de un objeto virtual y la ecuación $1/S_1 + 1/S_2 = 1/f$. En la etapa 988, el accionador ajustador variable 237 aplica al menos una señal de accionamiento para hacer que el ajustador de foco variable 135 cambie la polarización de al menos una lente birrefringente para cambiar la longitud focal, f , de la lente birrefringente para generar el valor seleccionado S_1 . Dado que cada lente birrefringente tiene dos longitudes focales discretas correspondientes a dos ejes de polarización, una combinación de los letales lentes proporciona una selección de cuatro longitudes focales discretas. De este modo, bajo el control del software, la unidad de procesamiento 4 o el procesador 210 de los circuitos de control 136 seleccionan la longitud focal disponible más próxima para aproximar el valor de f . Para cada lente birrefringente añadida, el número de longitudes focales discretas se duplica.

La Fig. 17C es un diagrama de flujo que describe una realización de un proceso para cambiar una región focal de un conjunto de micropresentación cambiando un radio de curvatura de al menos una lente de fluido del conjunto. Esta realización puede ser utilizada para implementar la etapa 960 de la Fig. 13 y es adecuada para utilizar con la realización de un conjunto de micropresentación que emplea al menos una lente líquida en su sistema de lentes 122 tal como el mostrado en las Figuras 3B1 y 3B2. En la etapa 990, la unidad de procesamiento 4 selecciona un valor de desplazamiento S_1 en base a una ubicación objetivo de un objeto virtual y la ecuación $1/S_1 + 1/S_2 = 1/f$. En la etapa 992, el accionador ajustador variable 237 aplica al menos una señal de accionamiento para hacer que el ajustador de foco variable 135 produzca un cambio de volumen en el fluido de la lente de fluido produciendo un cambio en su radio de curvatura para cambiar su longitud focal, f , para generar un valor seleccionado S_1 .

Como se ha mencionado anteriormente la reivindicación de la Fig. 13, para los objetos virtuales en el campo de visión del usuario pero no en la región focal del usuario, pueden ser aplicadas técnicas de borrosidad artificial. En otras realizaciones, puede ser generada una serie de imágenes en diferentes regiones focales. Moviéndose a través de un rango de focos o de regiones focales y presentando una imagen en cada región focal, se puede hacer que el usuario vea una imagen compuesta por capas de diferentes imágenes de región focal. Cuando el usuario reajusta su foco, el usuario se establecerá en una de estas regiones, y los objetos virtuales en el resto a las regiones se volverán borrosos de forma natural. El barrido a través de un rango de regiones focales se puede hacer a una velocidad o frecuencia predeterminadas. Esto se hace lo suficientemente rápido, para que la fusión de imagen temporal humana haga que parezcan estar presentes a la vez. La necesidad de profundidad artificial de las técnicas de campo tal como la borrosidad artificial será reducida, aunque la carga de realización puede aumentar significativamente.

El movimiento a través de un rango de regiones focal se puede implementar en una realización cambiando el desplazamiento entre los elementos de procesamiento de luz del conjunto de micropresentación o la potencia óptica de un elemento de procesamiento de luz en el conjunto a un régimen velocidad. El régimen de velocidad puede ser al menos tanto como un régimen de fotogramas de 30 fotogramas por segundo (fps), pero puede ser en algunas realizaciones a un régimen de velocidad tal como 60, 120 o 180 Hz. Una lente que es desplazada a lo largo de un eje óptico a un régimen de velocidad elevado proporcionando imágenes en diferentes regiones focales a veces referida como una lente vibrante o una lente oscilante. En algunos casos, la región de imagen, por ejemplo la superficie reflectante 124, 124a, se mueve en lugar de un sistema de lentes 122 o la fuente de imágenes 120, pero lo principal es lo mismo.

La Fig. 18A es un diagrama de flujo que describe una realización de un proceso para generar imágenes de objetos virtuales en diferentes regiones focales de una presentación de realidad aumentada que puede ser utilizada para implementar la etapa 610 en la Fig. 9. En la etapa 1002 de la Fig. 18A, el dispositivo de ordenador concentrador 12 o la unidad de procesamiento 4, o ambos, bajo el control del software determinan el campo de visión del usuario como en la etapa 950 de la Fig. 13, y determinan en la etapa 1004 si hay ubicaciones objetivo de objetos virtuales en el campo de visión actual del usuario como en la etapa 952. Similar a la etapa 966, si no hay objetos virtuales en el campo de visión actual del usuario, el procesamiento en la etapa 1006 vuelve para determinar y actualizar el campo de visión del usuario de la etapa 1002.

Como en la etapa 958, la unidad de procesamiento 4 escalará y orientará los objetos virtuales para una imagen que va ser insertada en la visión del usuario en la etapa 1012. El escalado y la orientación de la imagen virtual estarán basados en la ubicación del objetivo en el campo de visión y la en forma conocida del objeto virtual en base al modelo.

En la etapa 1016, el periodo de régimen de barrido se establece mediante el generador de sincronización 226 o el generador de reloj 244 de los circuitos de control. En la etapa 1018, se inicia un contador para ir a través de un número de regiones focales en el rango durante el periodo de barrido. En algunos casos, las regiones focales son

predeterminadas. En cada etapa de tiempo del periodo de barrido, la región focal de los elementos del conjunto de micropresentación es ajustada, y una imagen es presentada para cada región focal en su tiempo de barrido o etapa de tiempo al usuario en la etapa 1022. La siguiente región focal es seleccionada en la etapa 1024 incrementando el contador, y el procesamiento de las etapas 1020 a 1024 es repetido hasta que el contador indica que se ha completado el barrido de rango. El periodo de barrido finaliza en la etapa 1028. En otros ejemplos, un extremo del fotograma para la presentación puede interrumpir un barrido a través del rango de barrido, y otro barrido puede empezar con el siguiente fotograma.

En una realización, la carga de realización puede ser reducida realizando un subconjunto de regiones focales y utilizando rastreo de la región focal del usuario para optimizar aquéllas regiones focales que son seleccionadas para la realización. En otros ejemplos, los datos de rastreo de ojos no proporcionan suficiente precisión para determinar la profundidad de la ubicación del foco que el usuario está viendo. La lógica de deducción puede deducir un objeto en el que el usuario está enfocado. Una vez que un objeto seleccionado como el objeto del foco, se pueden utilizar los tres modelos tridimensionales o el mapeo de la escena actual se puede utilizar para determinar la distancia al objeto del foco.

La Fig. 18B ilustra otra realización de un proceso para presentar un objeto virtual que va a estar en foco cuando sea visto por un usuario que ve una presentación de realidad aumentada que puede ser utilizada para implementar la etapa 610 en la Fig. 9. La Fig. 18B es un diagrama de flujo que describe una realización de un proceso para generar imágenes de objeto virtual en diferentes regiones focales en una presentación de realidad aumentada. En la etapa 1002 de la Fig. 18B, el dispositivo de ordenador concentrador 12 o la unidad de procesamiento 4, o ambos, bajo el control del software determinan el campo de visión del usuario igual que en la etapa 950 de la Fig. 13 y determinan en 1004 si hay alguna ubicación objetivo de objetos virtuales en el campo de visión actual del usuario como en la etapa 952. Similar a la etapa 966, si no hay objetos virtuales en el campo de visión actual del usuario, el procesamiento en la etapa 1006 vuelve a determinar y actualizar el campo de visión del usuario en la etapa 1002.

Si hay al menos un objeto virtual que tiene una ubicación objetivo en el campo de visión del usuario, entonces en la etapa 1008 como en la etapa 954, el sistema tal como un software que se ejecuta en la unidad de procesamiento 4, determina la región focal actual del usuario dentro del campo de visión del usuario, y como en la etapa 956, el software que se ejecuta en el sistema de ordenador concentrador 12, la unidad de procesamiento 4 ambos identifica en la etapa 1010 cuál de los objetos virtuales en el campo de visión del usuario están en la región focal del usuario actual en base a la ubicación objetivo del objeto virtual en el modelo. Como en la etapa 958, la unidad de procesamiento 4 escalará y orientará los objetos virtuales para una imagen que va ser insertada en la visión del usuario en la etapa 1012. El escalado y la orientación de la imagen virtual estarán basados en la ubicación del objetivo en el campo de visión y de la forma conocida del objeto virtual basada en el modelo.

En la etapa 1014, la unidad de procesamiento 4 selecciona un rango de regiones focales que incluye la región focal actual del usuario. La unidad de procesamiento 4 puede seleccionar el rango de regiones focales en base a criterios tales como el contexto de la aplicación que se ejecuta, por ejemplo 452. Una aplicación puede utilizar una selección de objetos virtuales con trayectorias de movimiento y eventos predeterminados que desencadenan su aparición. Dado que el modelo de los objetos es actualizado con el movimiento de los objetos, la unidad de procesamiento 4 recibe estas actualizaciones como por la exposición de la Fig. 10. Entre las actualizaciones, la unidad de procesamiento 4 puede utilizar los datos del sensor sobre la posición y orientación de la cabeza del usuario para determinar qué objetos virtuales es más probable que enfoque en ese momento. La unidad de procesamiento 4 puede seleccionar un número de regiones focales en las que la trayectoria del objeto virtual se desplazará en base al modelo tridimensional de la escena. Por lo tanto, un criterio a modo de ejemplo es incluir cada región focal en la que un objeto virtual está localizado. Además, las regiones focales también pueden ser seleccionadas en las que la regiones la trayectoria del objeto virtual el movimiento estará en un fotograma de tiempo predeterminado. El fotograma de tiempo predeterminado puede ser hasta la siguiente actualización de los datos de modelo indicando las ubicaciones exactas de los objetos virtuales en un ejemplo.

El punto de región focal de inicio para el barrido a través del rango de regiones focales puede ser el más cercano al infinito en el campo de visión del usuario desde los ojos del usuario. Otras ubicaciones de inicio pueden ser utilizadas, sin embargo, empezar en el infinito puede simplificar la aplicación del filtro de opacidad 114. En la etapa 1016, el periodo de velocidad de barrido es iniciado por el generador de sincronización 226 o por el generador de reloj 244 de los circuitos de control. En la etapa 1018 se inicia un contador para ir a través de un número de regiones focales en el rango durante el periodo de barrido. En algunos casos, las regiones focales son predeterminadas. En cada etapa de tiempo del periodo de barrido, la región focal de los elementos del conjunto de micropresentación son ajustados, y la imagen es presentada para cada región focal en su tiempo de barrido o etapa de tiempo al usuario en la etapa 1022. La siguiente región focal en seleccionada en la etapa 1024 incrementando el contador, y el procesamiento de las etapas 1020 a 1024 es repetido hasta que el contador indica que está completo del barrido del rango. El periodo de barrido termina en la etapa 1028. En otros ejemplos, un extremo del fotograma para una presentación puede interrumpir un barrido a través del rango de barrido, y empezar otro barrido con el siguiente fotograma.

Las Figuras 19A y 19B ilustran esquemáticamente ejemplos de las diferentes formas de presentar objetos virtuales de múltiples focos en una presentación de realidad aumentada. En la Fig. 19A, un ejemplo de 4 imágenes 1050a a

5 1050d son generadas en diferentes regiones focales generadas en diferentes desplazamientos en la trayectoria óptica 133 del conjunto de micropresentación, y todas las imágenes son enviadas para la presentación en el elemento óptico de guiado de luz 112. En la Fig. 19B, la unidad de procesamiento 4 hace que una sección o parte de cada imagen que está en foco en esa región focal o de imagen sea incluida en una imagen compuesta 1055 para la ser presentada en la presentación 112. La sección y parte de cada imagen que está en el foco en un ejemplo puede ser la parte en donde el objeto virtual está situado en la imagen en el momento en el que fue generada.

REIVINDICACIONES

1. Un método para presentar objetos virtuales con foco variable mediante un sistema de realidad aumentada (10), comprendiendo el método:
 - 5 determinar mediante uno o más procesadores (210, 320, 322) un campo de visión del usuario tridimensional de un dispositivo de presentación próximo a los ojos (2) de un sistema de realidad aumentada;
 - identificar mediante uno o más procesadores una ubicación objetivo tridimensional de uno o más objetos virtuales dentro del campo de visión del usuario;
 - determinar mediante uno o más procesadores una región focal actual tridimensional del usuario que lleva el dispositivo de presentación próximo a los ojos dentro del campo de visión del usuario;
 - 10 identificar cuál de los uno o más objetos virtuales está en la región focal actual del usuario en base a las ubicaciones objetivo o tridimensionales del uno o más objetos virtuales dentro del campo de visión del usuario; y
 - presentar mediante el dispositivo de presentación próximo a los ojos el uno o más objetos virtuales identificados como que están en la región focal actual del usuario ajustando una región focal de un conjunto de micropresentación (173) que incluye elementos de procesamiento de luz alineados en una trayectoria óptica (133) dirigida sobre una presentación (112) del dispositivo de presentación próximo a los ojos moviendo los elementos de procesamiento de luz en la trayectoria óptica bajo el control de uno o más procesadores; en donde:
 - 15 los elementos de procesamiento de luz del conjunto de micropresentación comprenden al menos un elemento óptico (122) y una unidad de micropresentación (120) alineada en la trayectoria óptica con un elemento reflectante (124); y
 - 20 mover los elementos de procesamiento de luz en la trayectoria óptica del conjunto de micropresentación bajo el control de los uno o más procesadores comprende además:
 - determinar mediante uno o más procesadores un valor de desplazamiento, S1, entre un punto nodal delantero del al menos un elemento óptico y la ubicación objetivo de al menos un objeto virtual, estando el al menos un elemento óptico situado en una trayectoria óptica entre la unidad de micropresentación y el elemento reflectante,
 - 25 determinar mediante uno o más procesadores un valor de desplazamiento, S2, entre el elemento reflectante y un punto nodal trasero del al menos un elemento óptico para una longitud focal de al menos un elemento óptico, f, en base a la ecuación $1/S1 + 1/S2 = 1/f$; y
 - 30 mover el al menos un elemento óptico bajo el control de uno o más procesadores para generar los valores del desplazamiento determinados S1 y S2.
2. El método de la reivindicación 1, que comprende además:
 - 35 aplicar al menos una profundidad artificial de la técnica de campo a un objeto virtual que está dentro del campo de visión del usuario pero fuera de la región focal actual del usuario como una función de la distancia desde la región focal actual.
3. Un medio leible con ordenador que comprende instrucciones ejecutables por un ordenador almacenadas en el mismo que, cuando se ejecutan mediante una unidad de procesamiento, hacen que la unidad de procesamiento realice el método de una cualquiera de las reivindicaciones precedentes.
4. Un sistema de realidad aumentada que proporciona foco variable de objetos virtuales que comprende:
 - una unidad de presentación transparente (2) que comprende:
 - 40 un conjunto de micropresentación (173) para generar imágenes para una presentación de una unidad de presentación transparente, incluyendo el conjunto de micropresentación un ajustador de foco virtual variable (135) y elementos de procesamiento de luz que incluyen al menos un elemento óptico (122) y una unidad de micropresentación (120), estando el al menos un elemento óptico y la unidad de micropresentación alineados en una trayectoria óptica (133) con un elemento reflectante (124) de una unidad de presentación transparente, cambiando el
 - 45 ajustador de foco virtual variable la región focal del conjunto de la micropresentación para colocar uno o más objetos virtuales en una ubicación objetivo en una región focal actual del usuario que lleva la unidad de presentación transparente (2) moviendo el uno o más elementos de procesamiento de luz en la trayectoria óptica bajo el control del uno o más procesadores (210, 320, 322), en donde
 - 50 el uno o más procesadores que controlan el ajustador de foco virtual variable mueven los elementos de procesamiento de luz en la trayectoria óptica del conjunto de micropresentación comprenden además:

- determinar mediante el uno o más procesadores un valor de desplazamiento, S1, entre un punto nodal delantero del al menos un elemento óptico y la ubicación objetivo del al menos un objeto virtual, estando el al menos un elemento óptico situado en una trayectoria óptica entre la unidad de micropresentación y el elemento reflectante,
- 5 determinar mediante el uno o más procesadores un valor de desplazamiento, S2, entre el elemento reflectante y el punto nodal trasero del al menos un elemento óptico para una longitud focal del al menos un elemento óptico, f, basado en la ecuación $1/S1 + 1/S2 = 1/f$; y
- mover el al menos un elemento óptico bajo el control del uno o más procesadores para generar los valores del desplazamiento determinados S1 y S2;
- 10 incluyendo los circuitos de control (136) un dicho procesador (210) del uno o más procesadores conectados en comunicación con el conjunto del micropresentación, incluyendo los circuitos de control un accionador (237) para controlar el ajustador de foco virtual variable (135), el procesador bajo el control del software que dirige la unidad de micropresentación para incluir al menos un objeto virtual en una región focal de usuario actual, y
- una presentación (112) conectada ópticamente al conjunto del micropresentación para proyectar imágenes virtuales tridimensionales dirigidas a al menos una ubicación del ojo.
- 15 5. El sistema de realidad aumentada de la reivindicación 4, que comprende además:
- una armadura (137) para soportar los elementos de procesamiento de luz del conjunto de micropresentación (173) y estando la armadura bajo el control del ajustador de foco virtual variable (135).
6. El sistema de realidad aumentada de la reivindicación 4 o 5, en el que:
- 20 los circuitos de control (136) comprenden además un generador de sincronización (226) para proporcionar señales de sincronización al accionador (237) que controla el ajustador de foco virtual variable (135) para operar a un régimen programado;
- moviendo el ajustador de foco virtual variable los elementos de procesamiento de luz al régimen programado a través de un rango de desplazamientos en la trayectoria óptica (133) del conjunto de micropresentación (173), correspondiendo el rango de desplazamientos a un rango de regiones focales;
- 25 incluyendo el rango de regiones focales la región focal actual del usuario; y
- dirigiendo el uno o más procesadores (210, 320, 322) la unidad de micropresentación (120) para generar una imagen que incluye cualquier objeto virtual identificado que está situado en una respectiva región focal del rango cuando se alcanza el desplazamiento respectivo en el rango.
7. El sistema de realidad aumentada para reivindicación 6, en el que
- 30 la unidad de micropresentación (120) genera una imagen compuesta que incluye partes en foco de las diferentes imágenes generadas por el rango de desplazamientos, y
- la presentación (112) proyecta la imagen compuesta en al menos una ubicación del ojo.
8. El sistema de realidad aumentada de la reivindicación 4, en el que el conjunto de micropresentación (173) comprende además un sistema de lentes giratorio, que comprende una pluralidad de lentes (122a-d) teniendo cada una, una longitud focal diferente y estando soportadas dentro de un soporte de disco (160) giratorio mediante el ajustador de foco virtual variable (135), estando el ajustador de foco virtual variable (135) configurado para girar el soporte de disco (160), de manera que una lente seleccionada (122a-d) es alineada en la trayectoria óptica (133) del conjunto de micropresentación.
- 35 9. El sistema de realidad aumentada de la reivindicación 4, en el que el conjunto del micropresentación (173) comprende además un sistema de lentes insertable, en el que cada uno de un número de elementos ópticos (122) está unido a un respectivo brazo (123) unido a la armadura (137), estando cada brazo (123) configurado para mover su uno o más elementos ópticos (122) en posición en un desplazamiento de la trayectoria óptica (133) del conjunto de micropresentación bajo el control del ajustador de foco virtual variable (135).
- 40

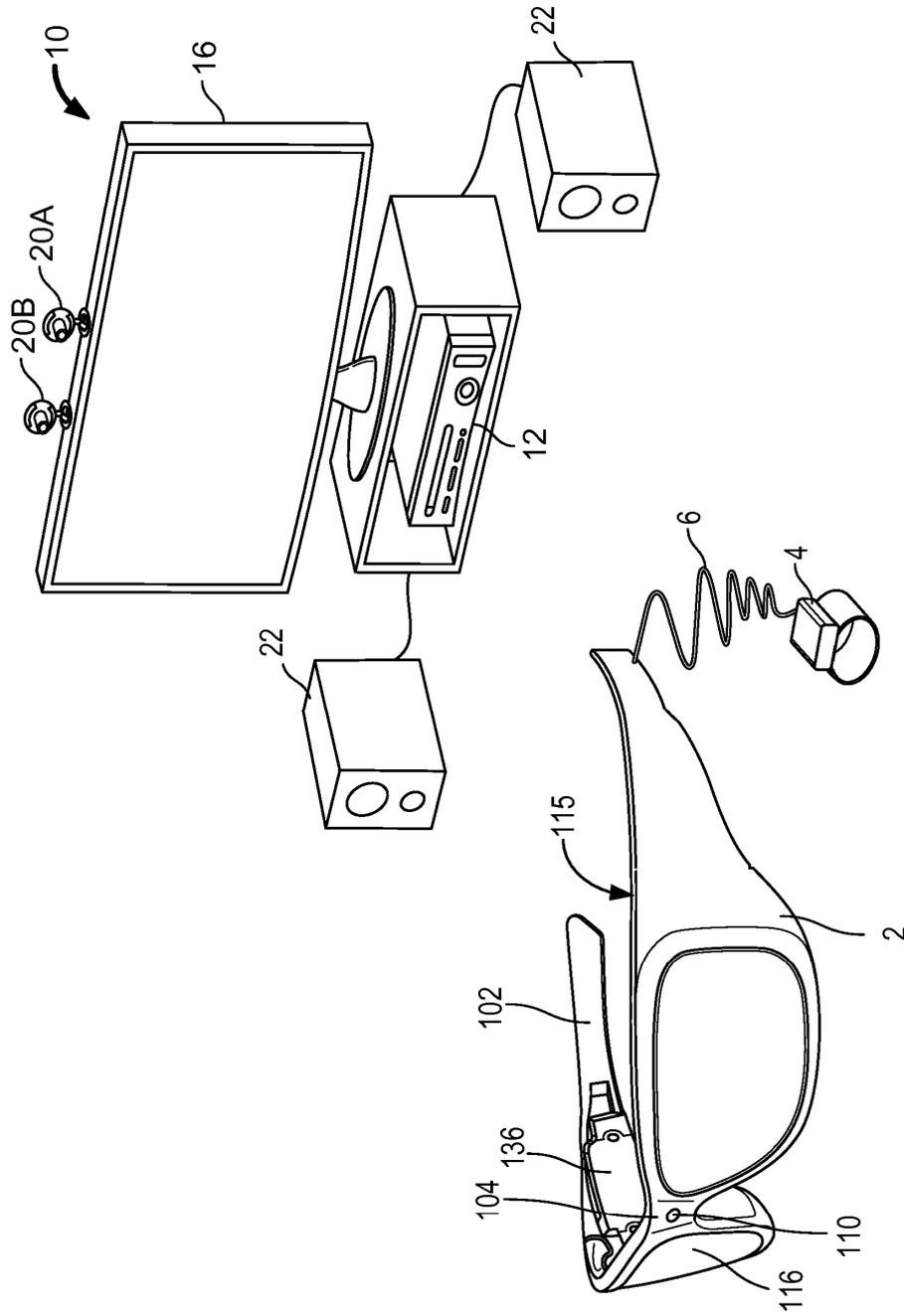


FIG. 1

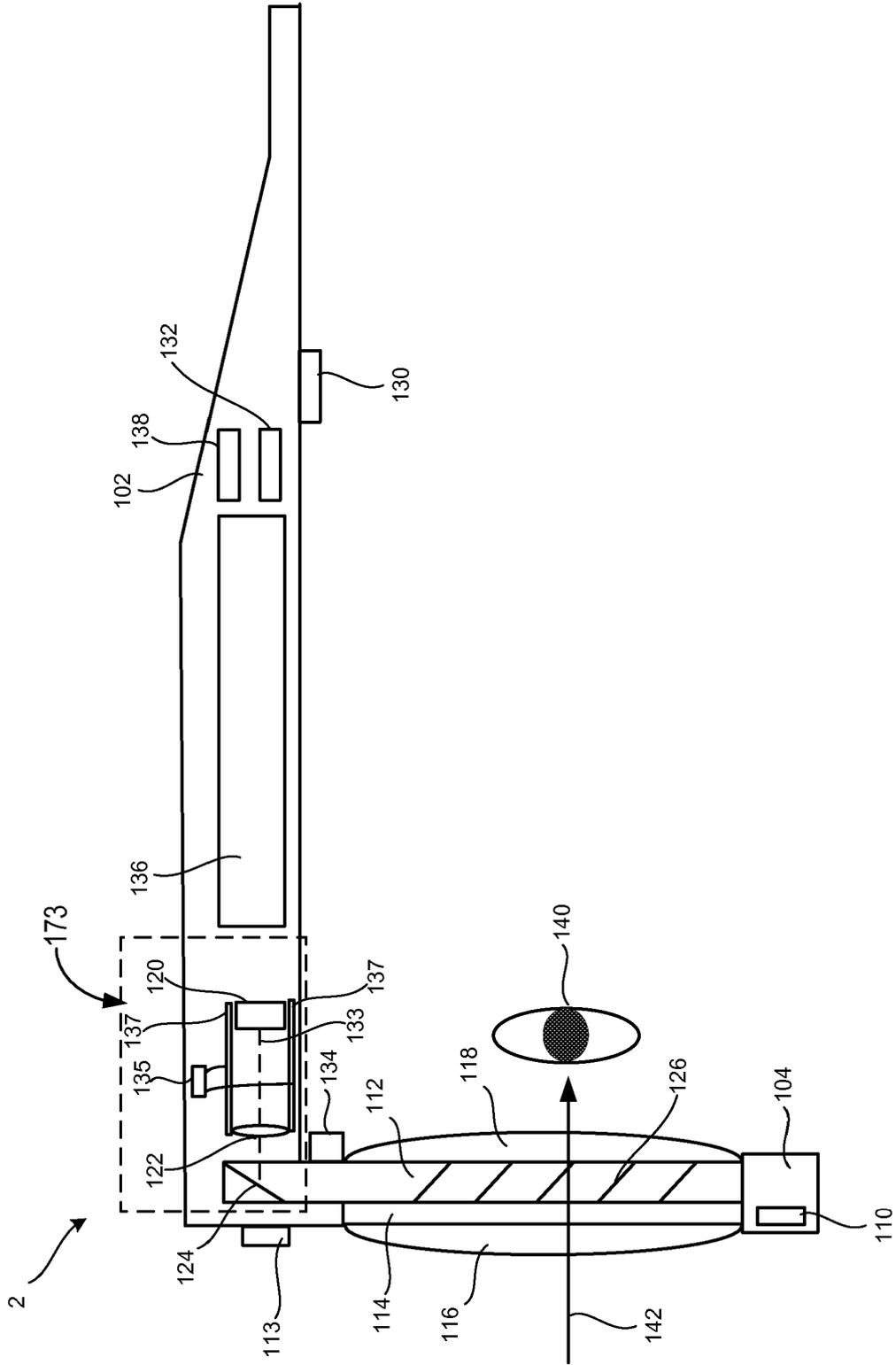


FIG. 2A

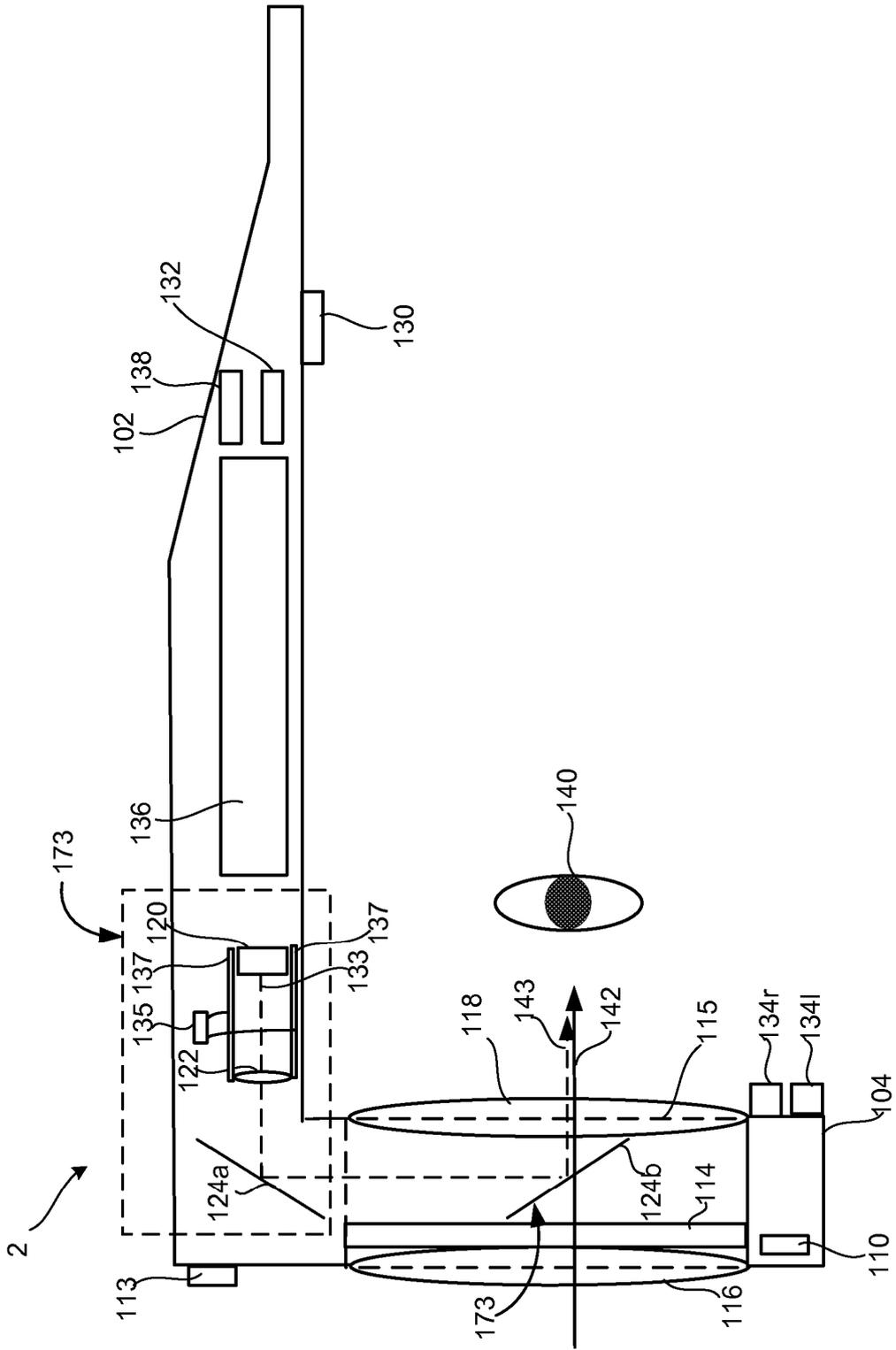


FIG. 2B

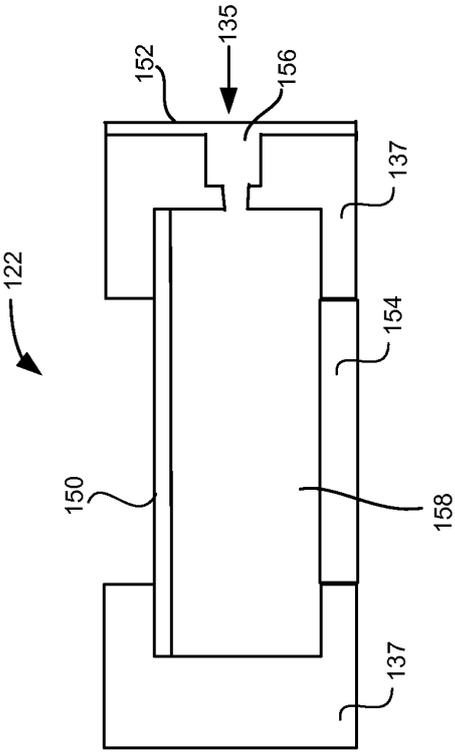


FIG. 3B1

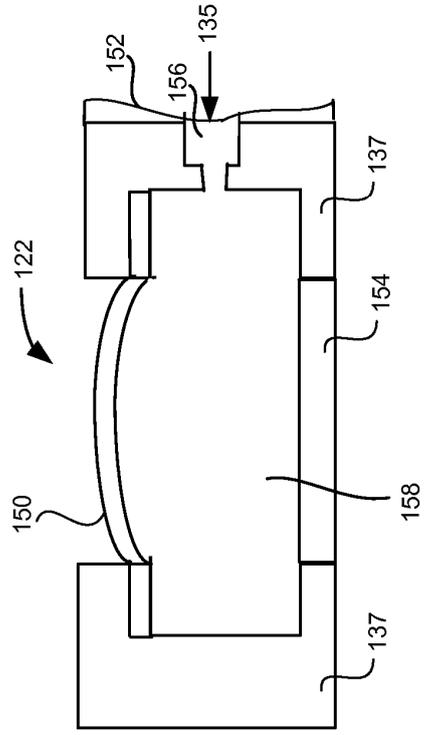


FIG. 3B2

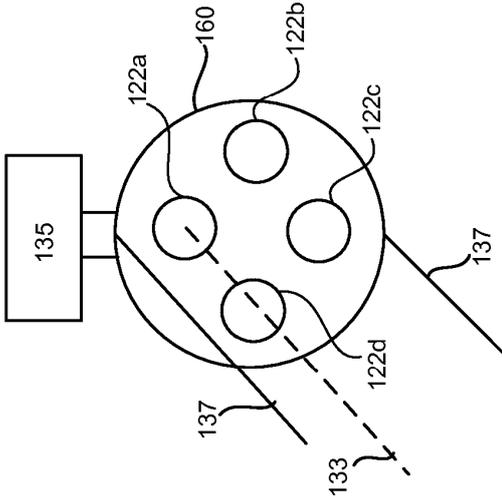


FIG. 3A

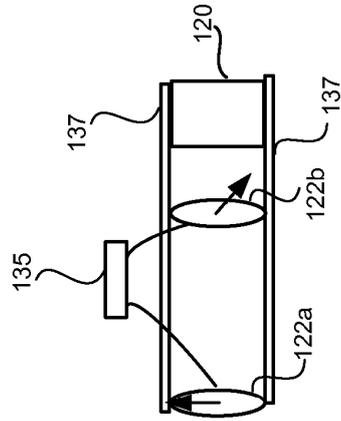


FIG. 3C

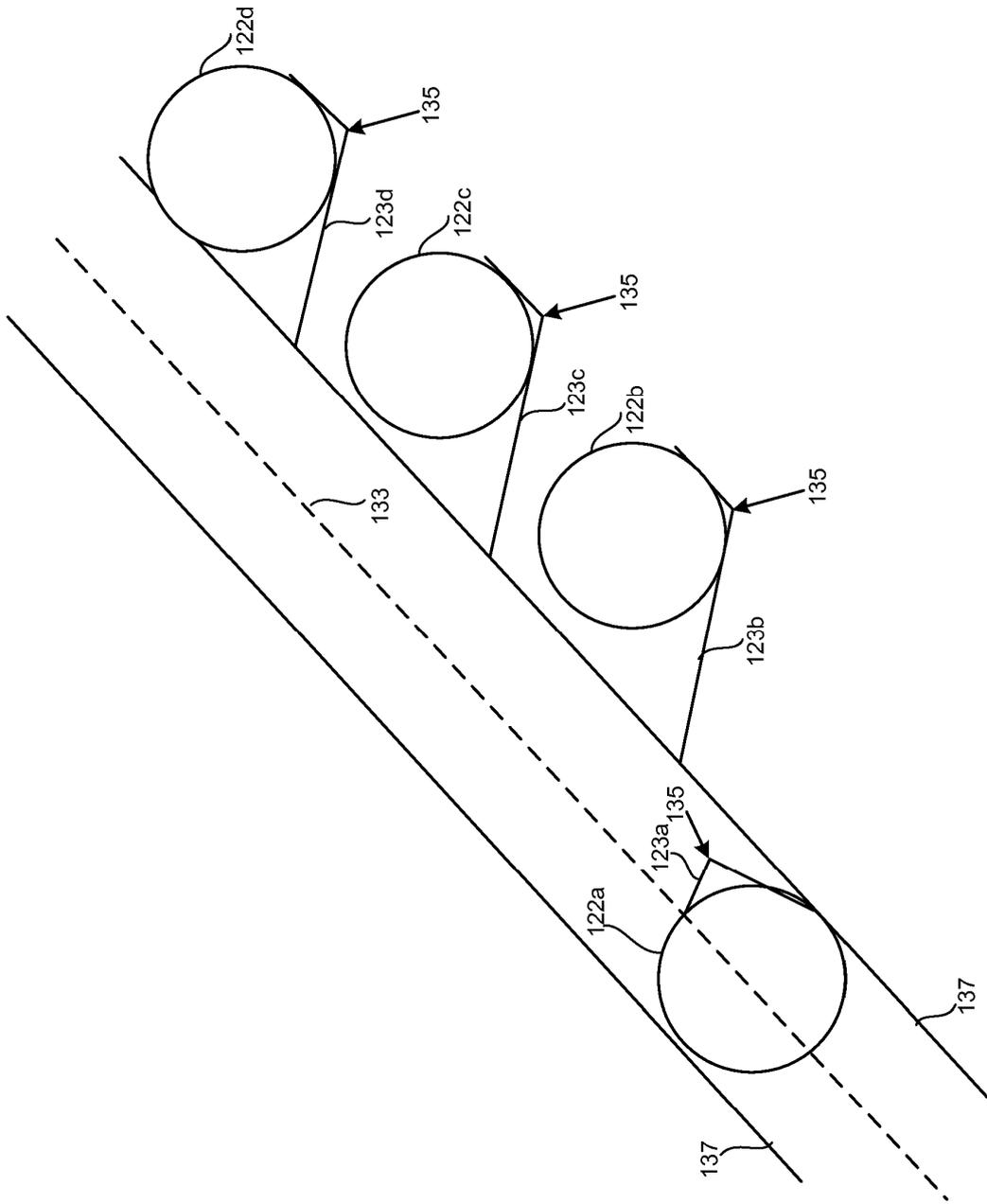


FIG. 3D

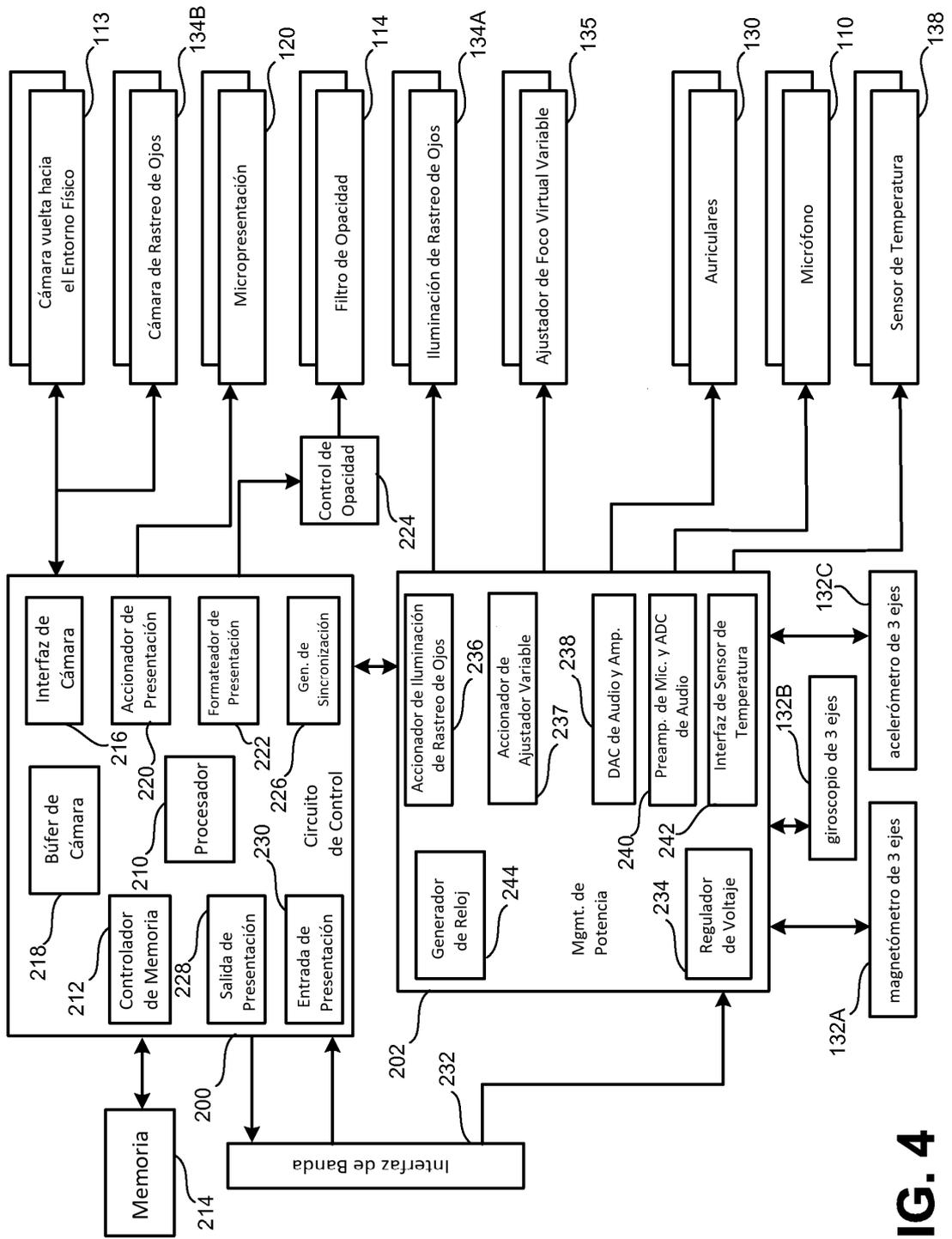


FIG. 4

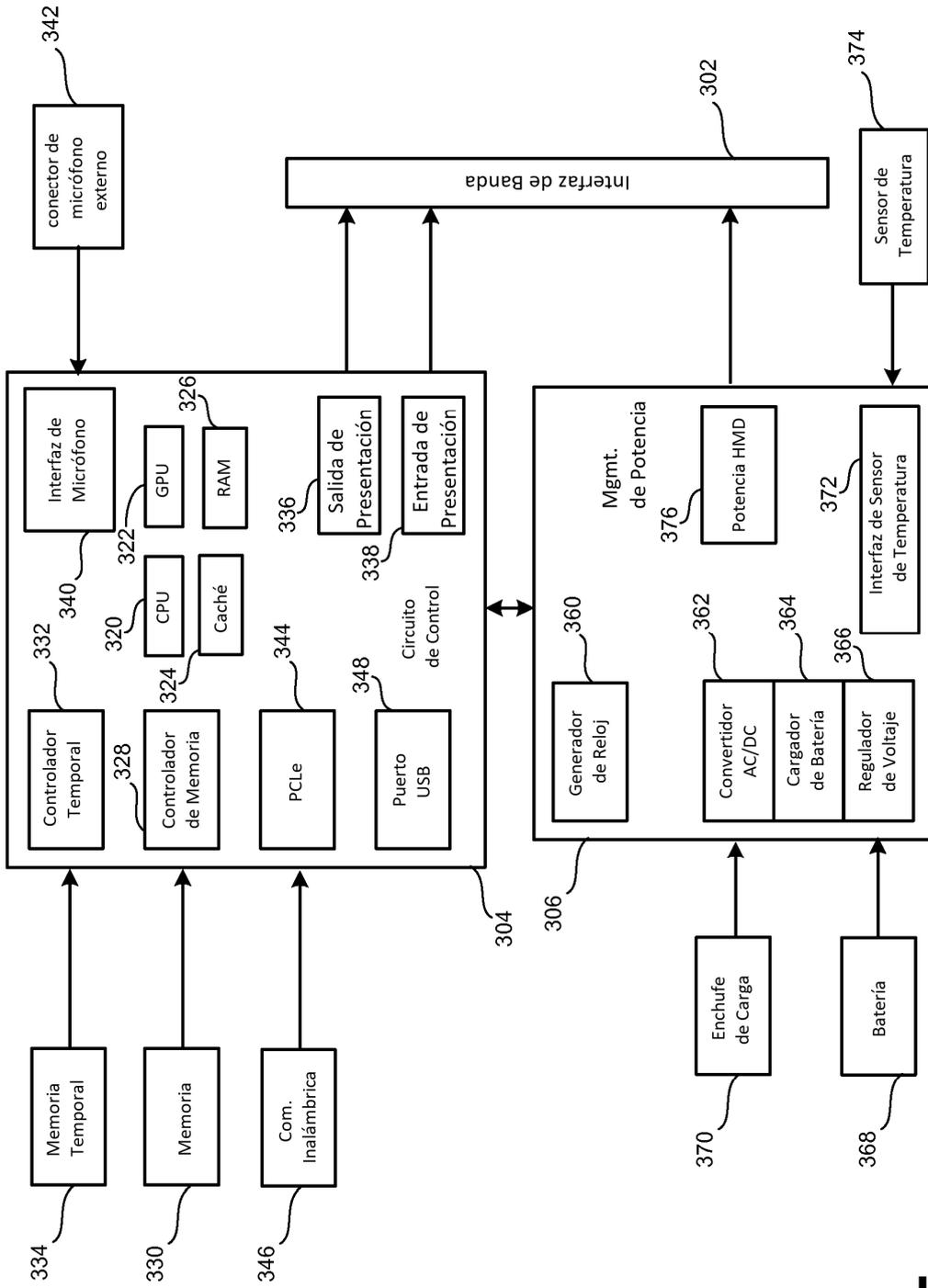


FIG. 5

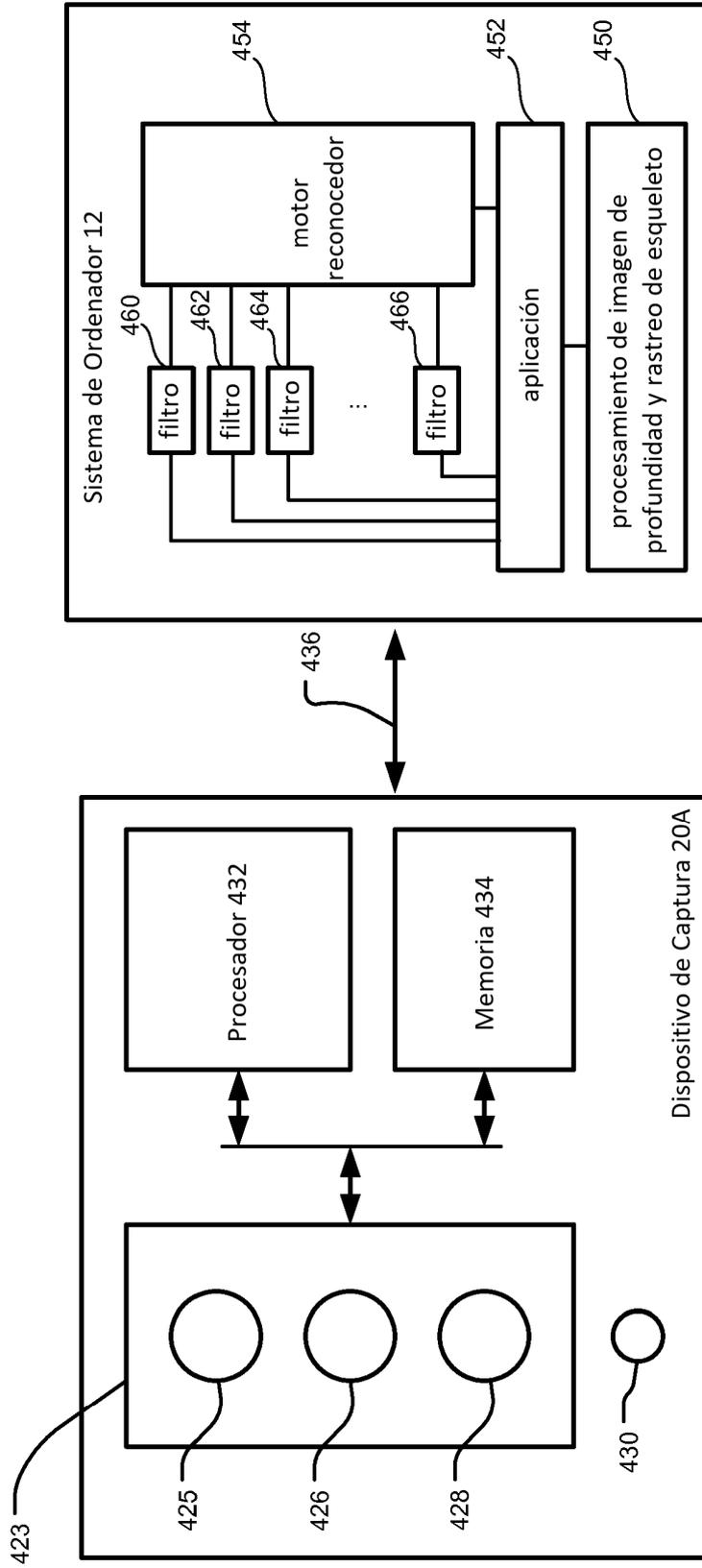


FIG. 6

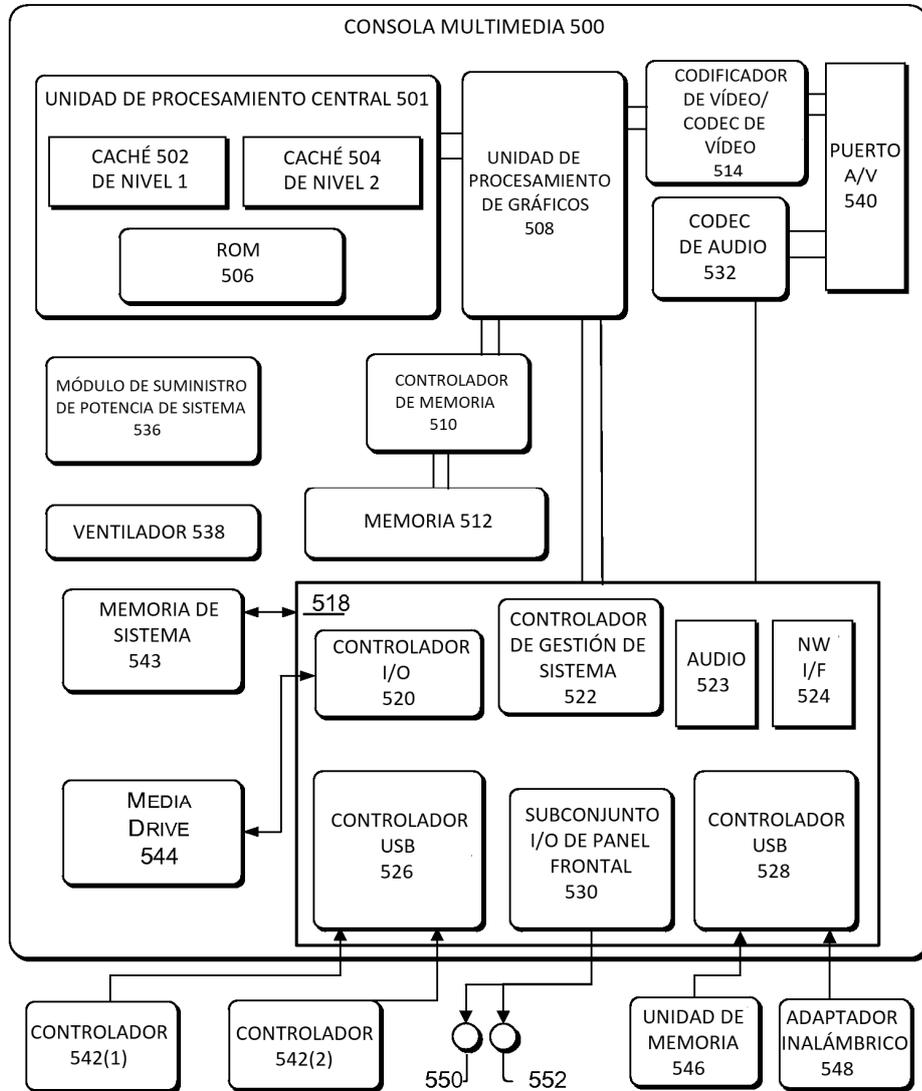


FIG. 7

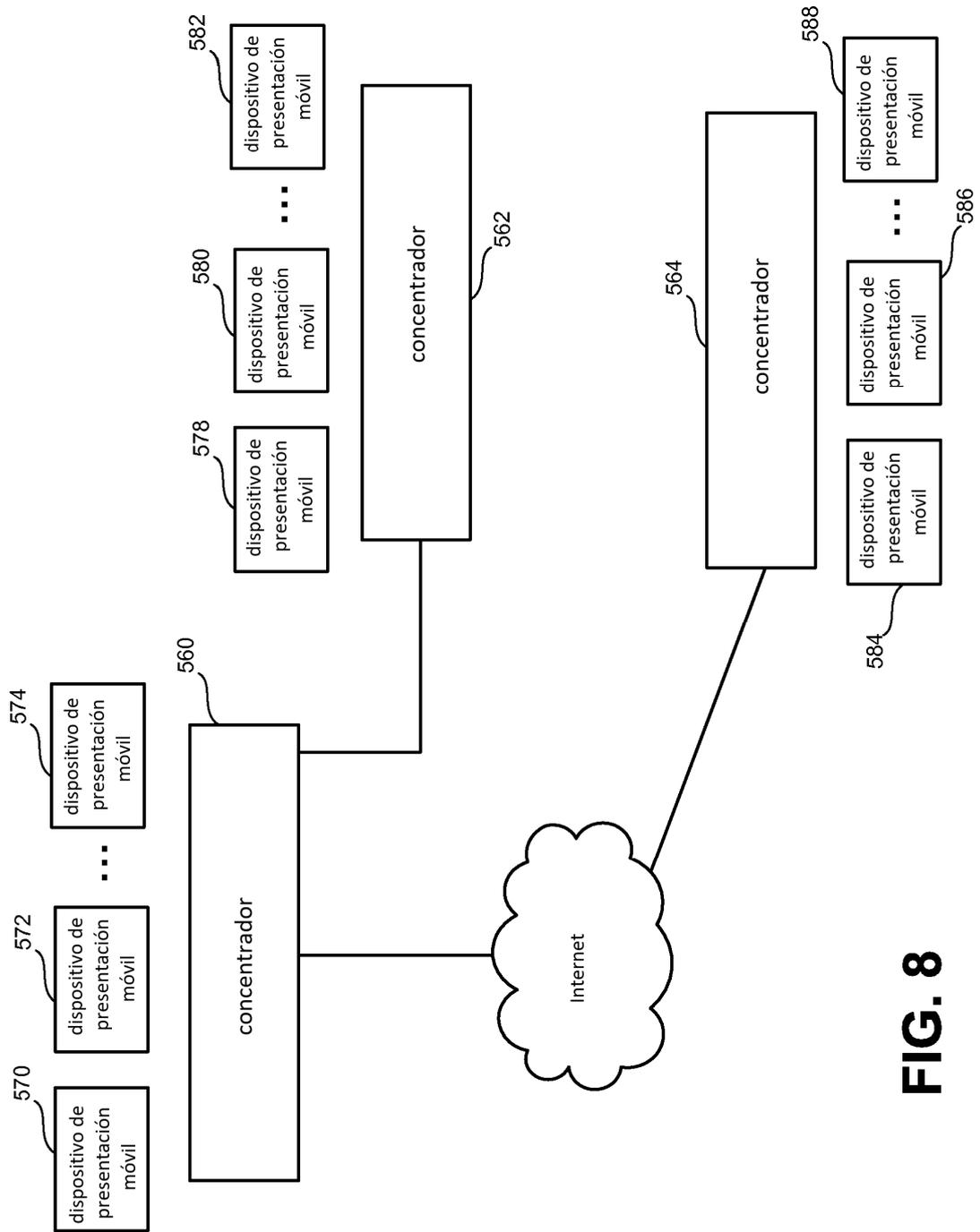


FIG. 8

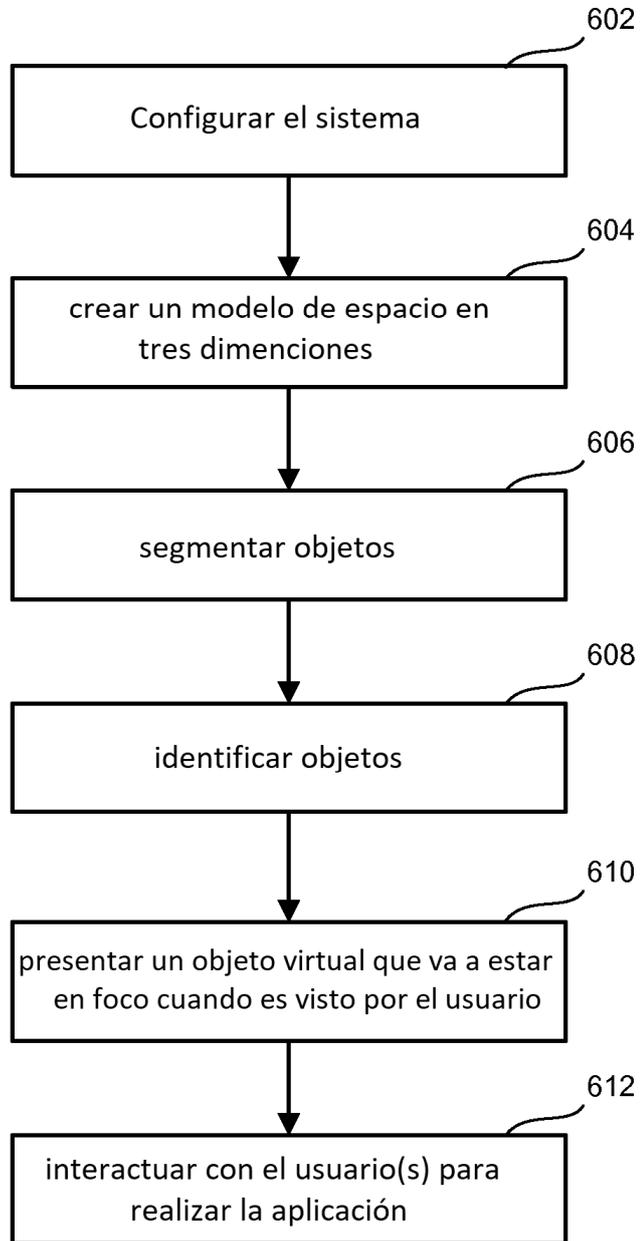


FIG. 9

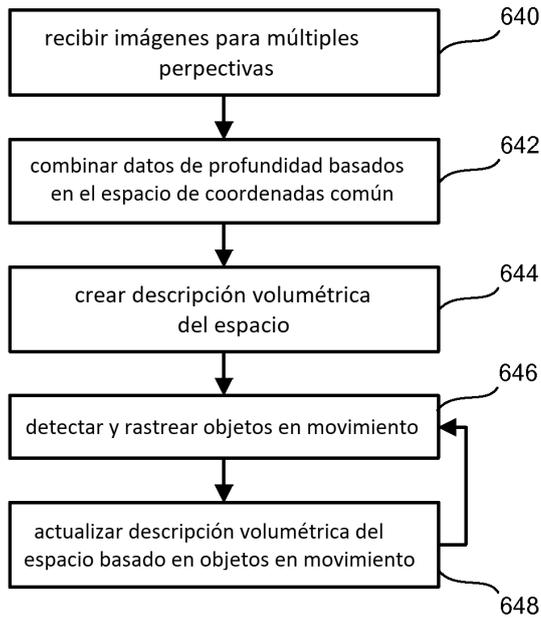


FIG. 10

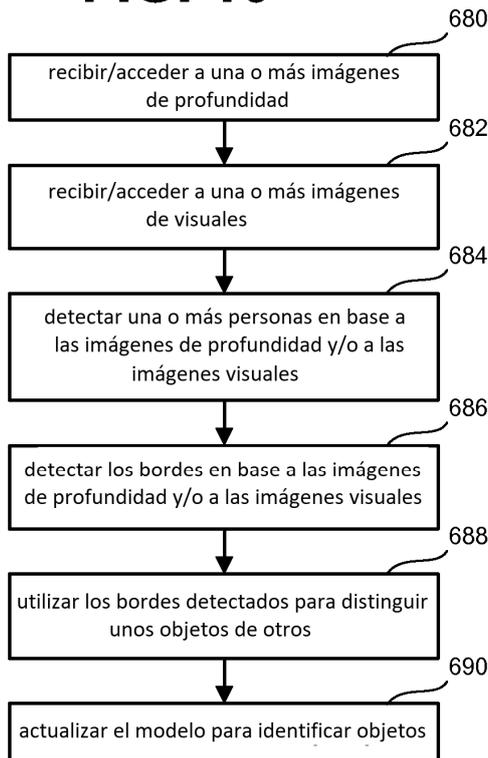


FIG. 11

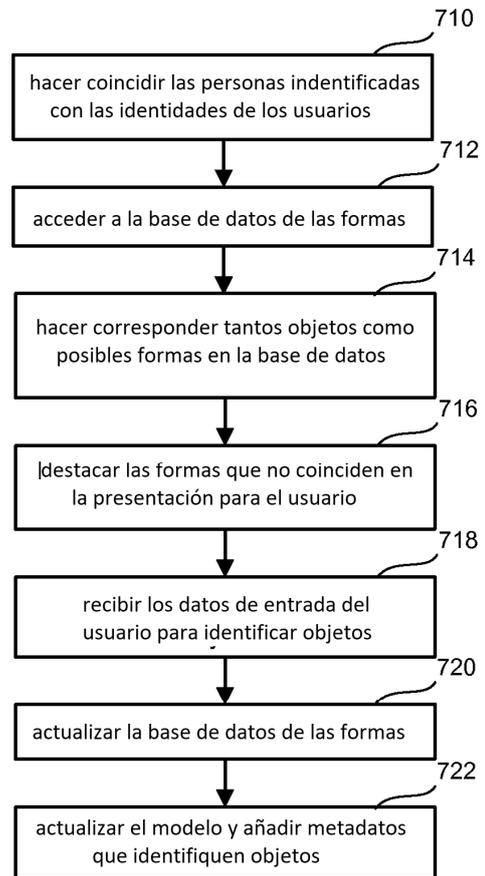


FIG. 12

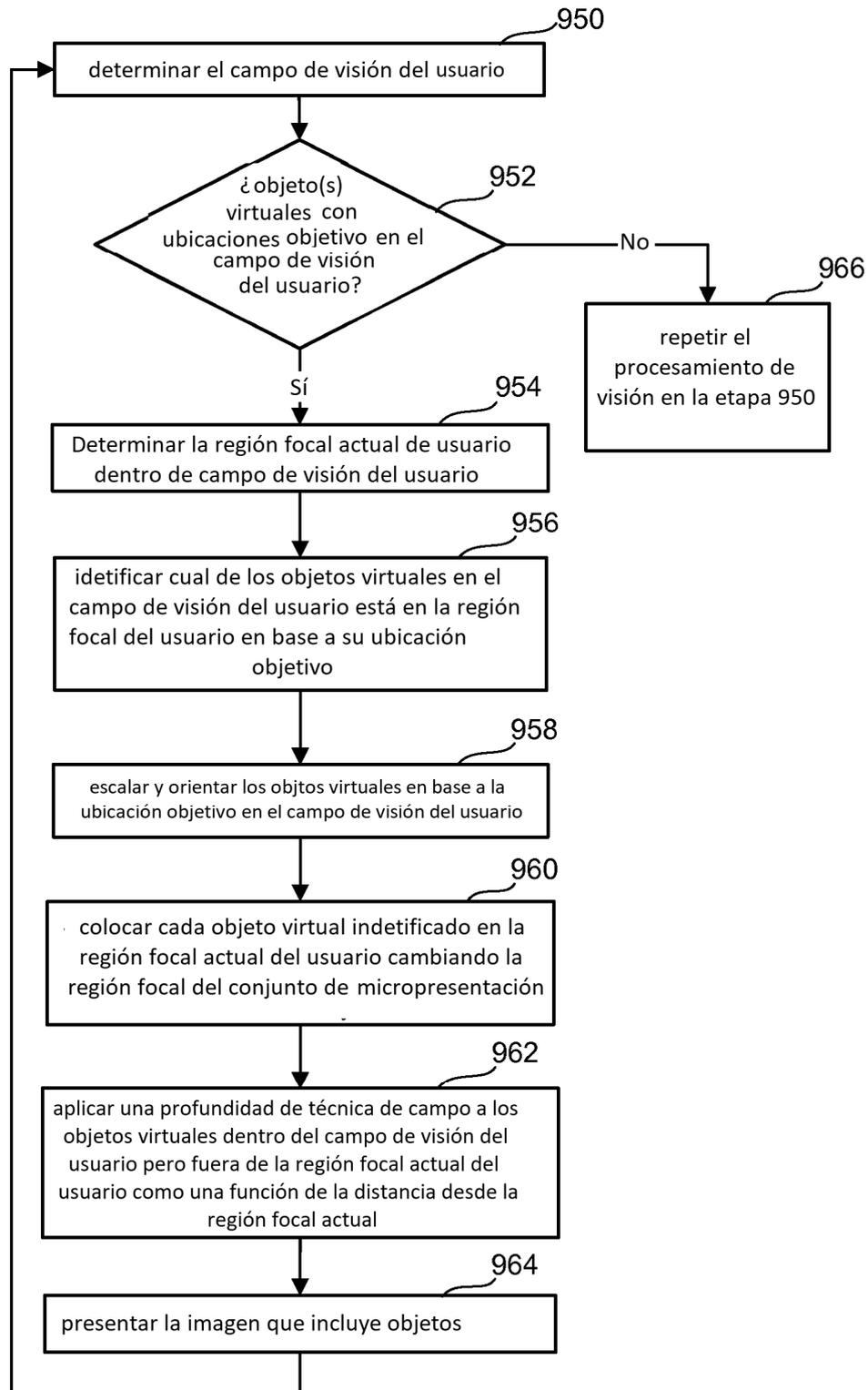


FIG. 13

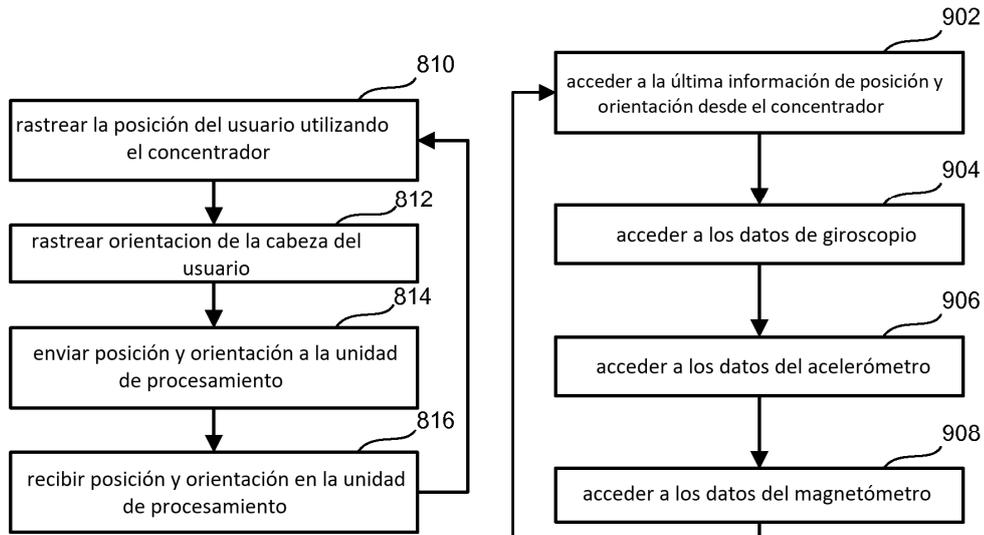


FIG. 14

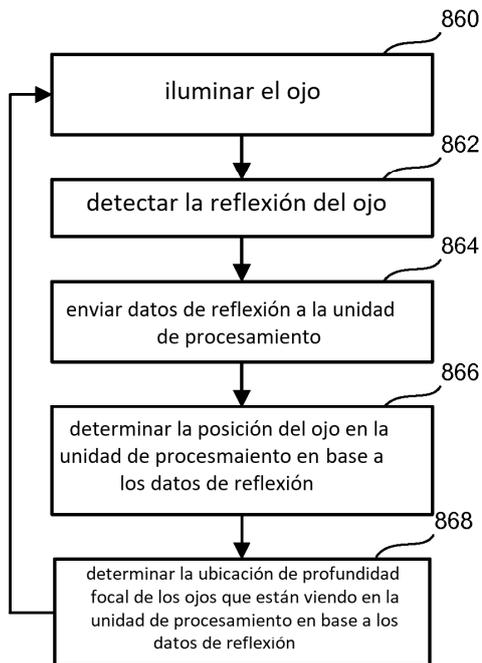


FIG. 15

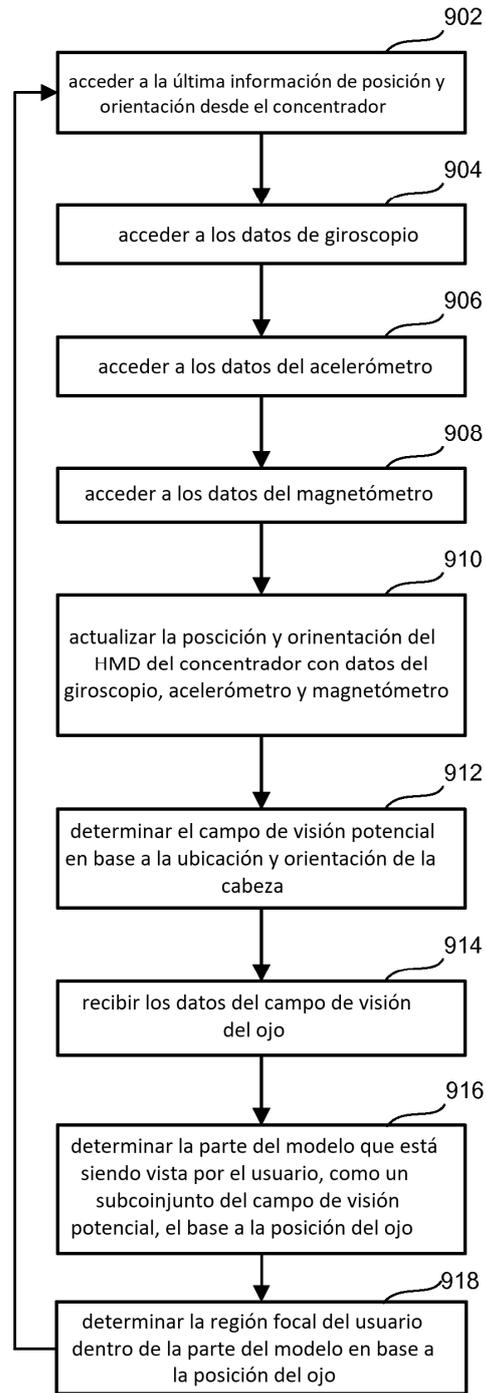


FIG. 16

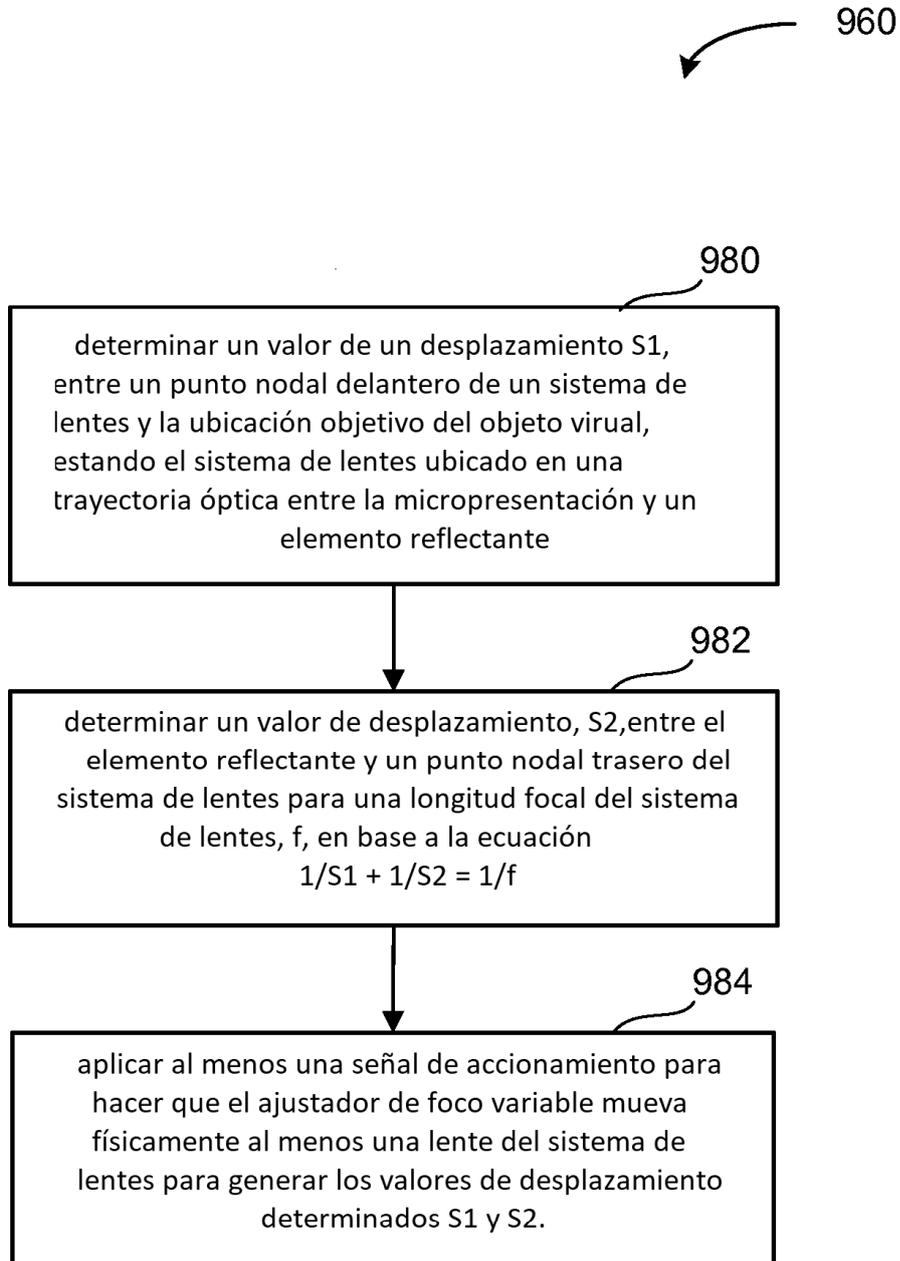


FIG. 17A

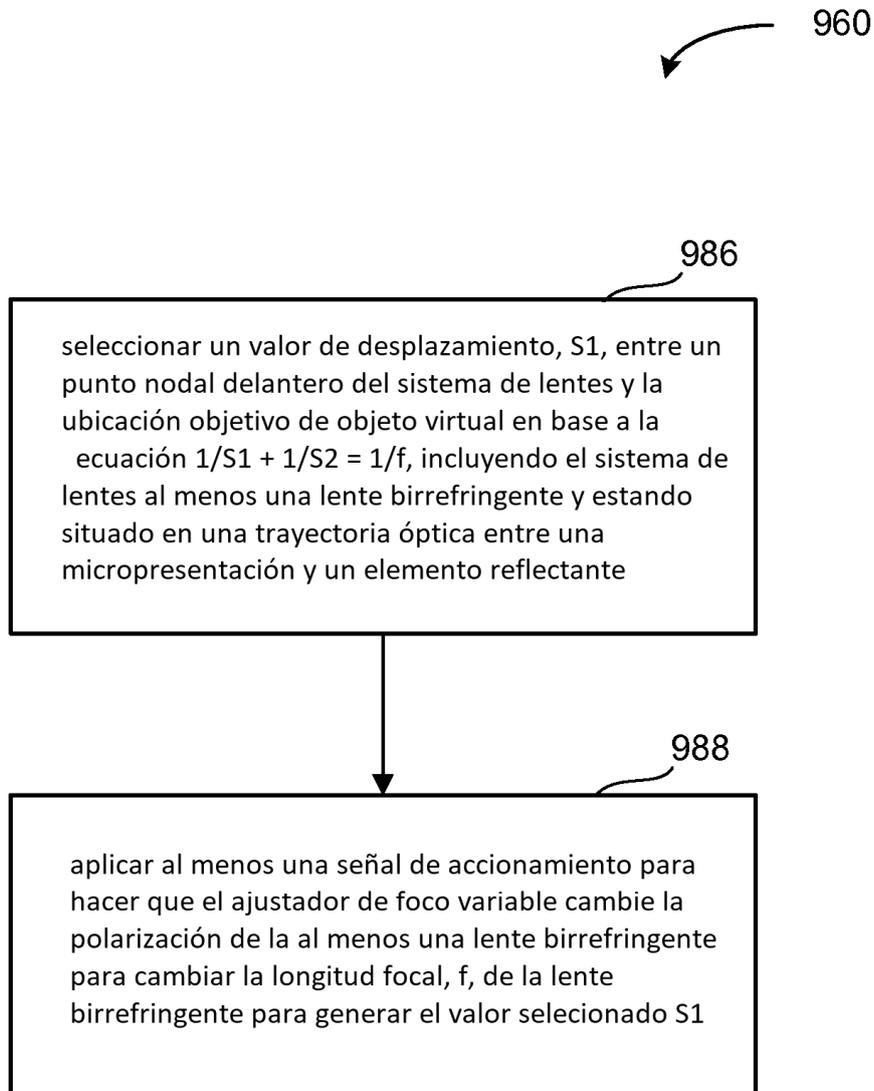
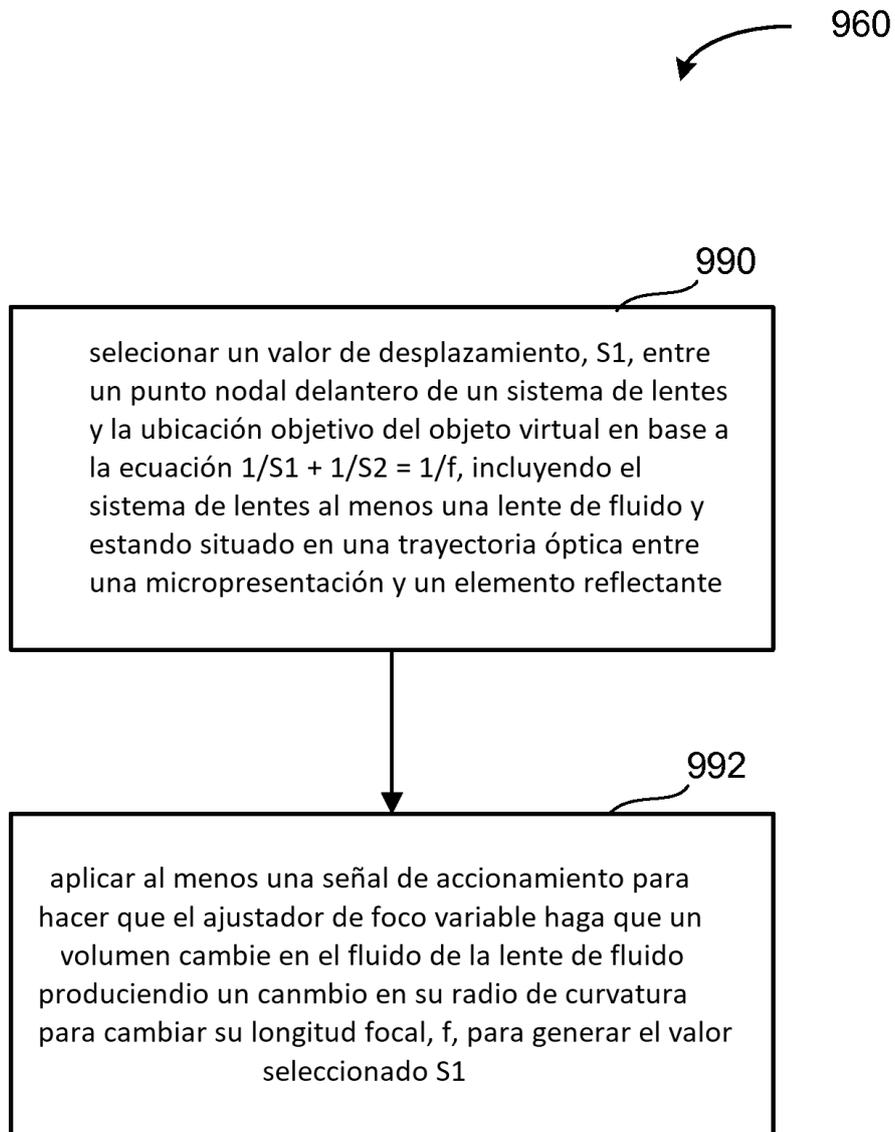


FIG. 17B

**FIG. 17C**

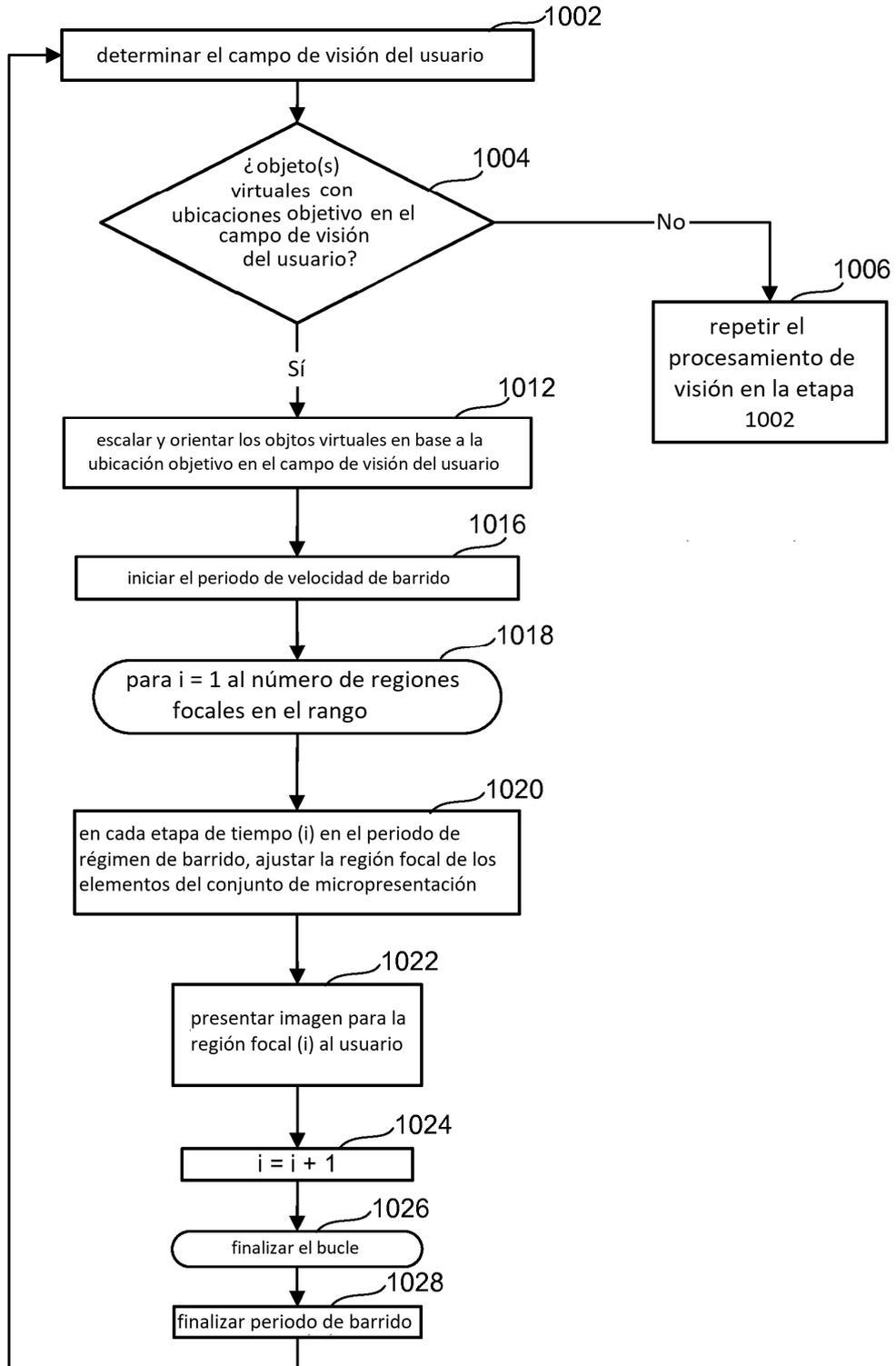


FIG. 18A

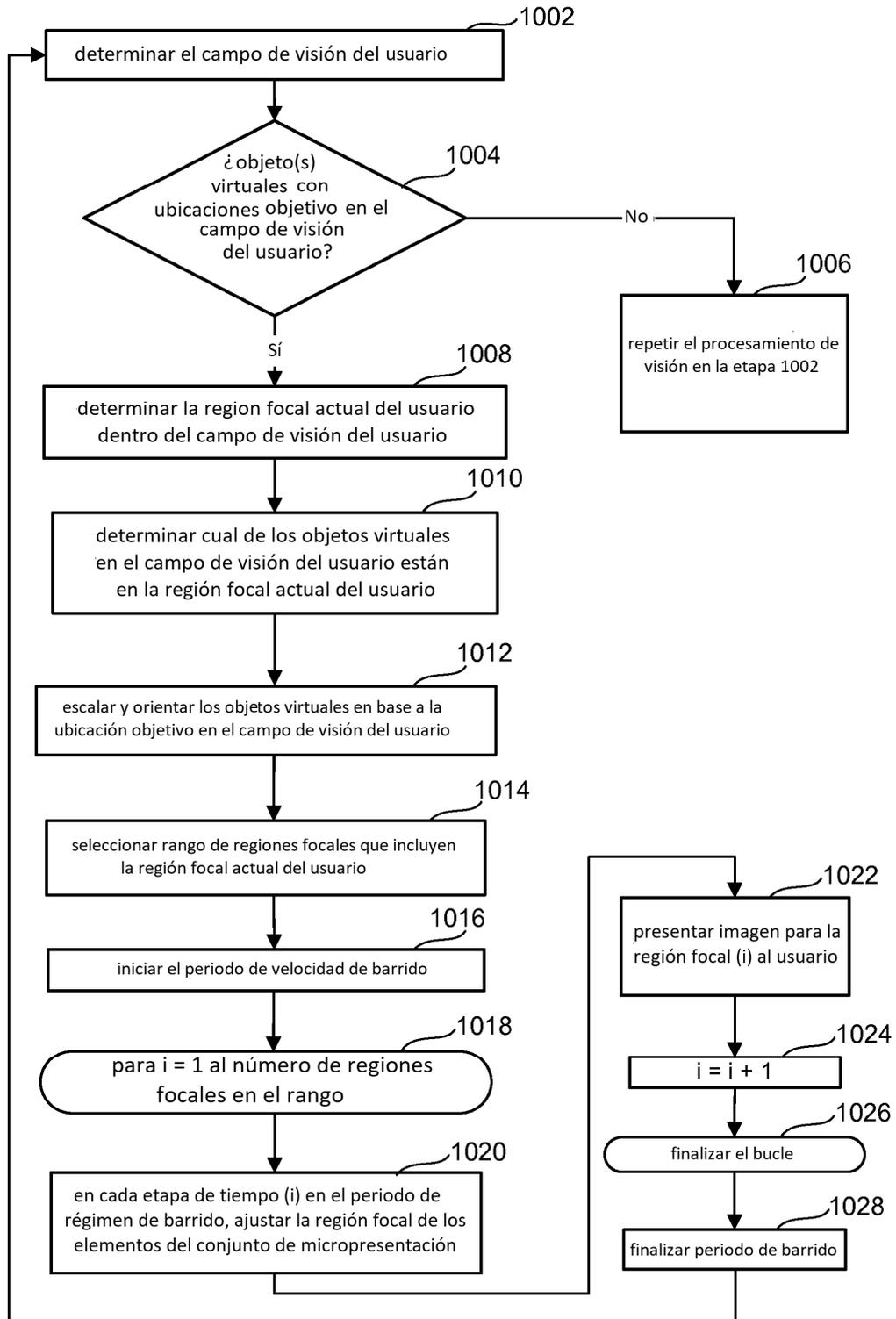


FIG. 18B

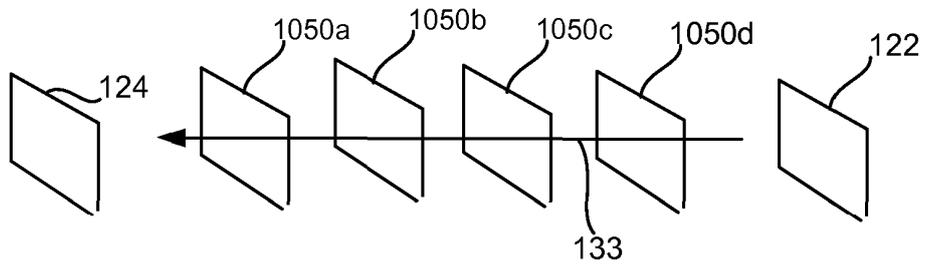


FIG. 19A

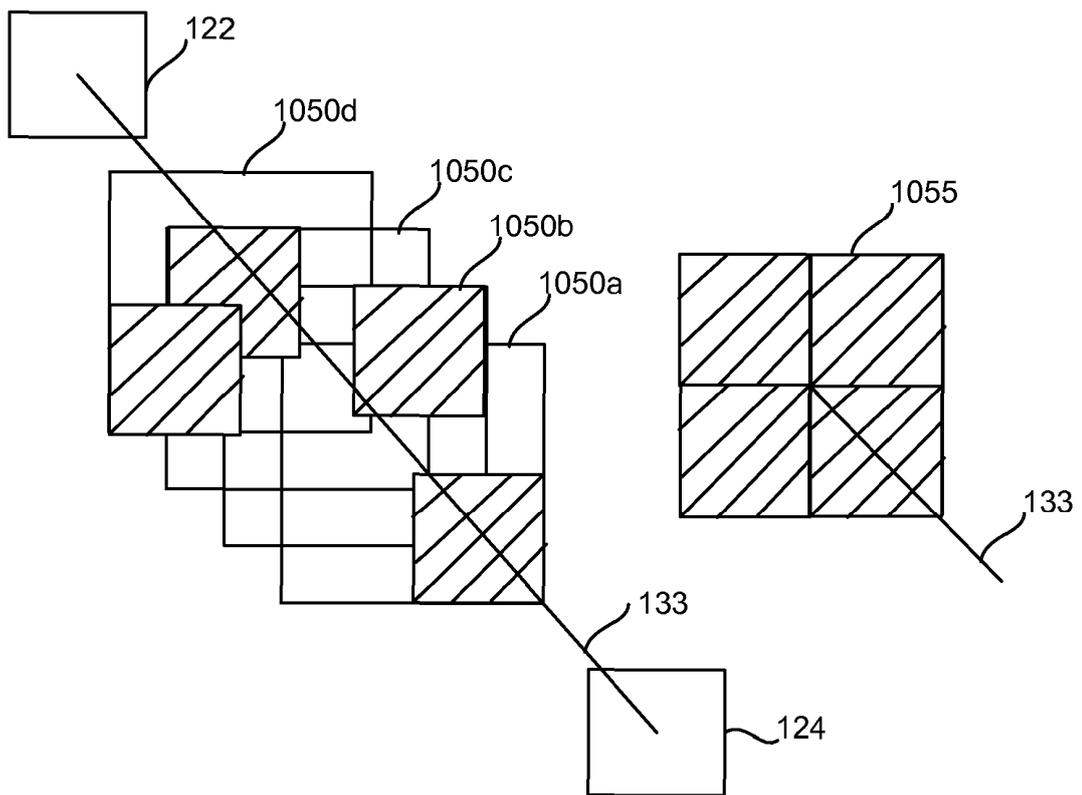


FIG. 19B