

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 669 869**

51 Int. Cl.:

G06F 3/03 (2006.01)

G06F 3/14 (2006.01)

G06T 7/20 (2007.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.06.2011 PCT/US2011/040317**

87 Fecha y número de publicación internacional: **29.12.2011 WO11163011**

96 Fecha de presentación y número de la solicitud europea: **14.06.2011 E 11798639 (8)**

97 Fecha y número de publicación de la concesión europea: **28.02.2018 EP 2585896**

54 Título: **Retroalimentación del seguimiento del usuario**

30 Prioridad:

22.06.2010 US 820954

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

29.05.2018

73 Titular/es:

**MICROSOFT TECHNOLOGY LICENSING, LLC
(100.0%)**

**One Microsoft Way
Redmond, Washington 98052, US**

72 Inventor/es:

**YEE, DAWSON y
PEREZ, KATHRYN STONE**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 669 869 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

Retroalimentación del seguimiento del usuario

Antecedentes

5 En un entorno informático típico, un usuario tiene un dispositivo de entrada tal como un teclado, un ratón, una palanca de mando o similar, que se puede conectar al entorno informático mediante una conexión por línea, cable, inalámbrica o similar, o algún otro medio de conexión. Si el control del entorno informático se hubiera de desplazar desde un controlador conectado a un control basado en gestos, particularmente como en una interfaz de usuario natural (NUI), el usuario ya no tiene un dispositivo conectado para informar con una gran consistencia al entorno informático acerca de una instrucción de control para una aplicación.

10 Por ejemplo, cuando un entorno informático tiene una entrada establecida tal como un controlador o un teclado, un usuario puede determinar que tiene un controlador conectado a un puerto, que está presionando teclas o botones y que el sistema está respondiendo. Cuando el control sobre el entorno informático se cambia a los gestos de un usuario, la detección de los gestos se puede inhibir o producir una respuesta subóptima de la aplicación debido a las características visuales o auditivas del área de captura de los movimientos corporales del usuario a diferencia de un controlador. La incapacidad de detectar de manera apropiada los gestos puede frustrar al usuario al interactuar con una aplicación en ejecución. Por ejemplo, se puede frustrar su participación en un juego que es ejecutado por la aplicación.

20 El documento US 2008/0030460 A1 (Hildreth et al.) describe un método para usar visión en estéreo para interactuar con un ordenador. Esto incluye la captura de una imagen en estéreo, y el procesamiento de la imagen en estéreo para determinar la información de la posición de un objeto en la imagen en estéreo. El objetivo es controlado mediante un usuario. En una realización, un detector de imagen o video cámara adquiere las imágenes en estéreo de una región de interés y la escena circundante. Los objetos detectados por el detector de imagen se pueden excluir de consideración si sus posiciones están fuera de la región de interés. El documento US 2009/0228841 A1 (Hildreth) describe un sistema de manipulación de imágenes mejorado basado en los gestos. Este narra también un proceso para orientar o alinear el objeto de control del usuario con una posición objetivo en un espacio libre que facilita o mejora futuros reconocimientos de gestos. Con este fin, un enfoque es presentar inicialmente una representación del usuario fuera de una región central de la interfaz de usuario, y una imagen objetivo en dicha región central, para indicar al usuario que alinee su representación con la imagen objetivo mediante el movimiento de su cuerpo. El documento US 2009/0133051 A1 (Hildreth) se relaciona con las técnicas para controlar de manera automática y pasiva el acceso al contenido de televisión y video.

Compendio

35 Es un objetivo proporcionar retroalimentación a un usuario sobre la capacidad de una aplicación en ejecución para seguir la acción de un usuario para controlar la aplicación en ejecución en un sistema informático. Esto se soluciona mediante el tema de las reivindicaciones independientes. Las realizaciones preferidas se describen mediante las reivindicaciones dependientes. En un ejemplo, un sistema de captura detecta a un usuario en un área de captura. En respuesta a los criterios de seguimiento del usuario que no se satisfacen, se emite una retroalimentación al usuario. En algunos ejemplos, la retroalimentación puede ser un indicador de audio. En otros ejemplos, se proporcionan indicadores visuales como retroalimentación para recomendar al usuario tomar una acción para satisfacer los criterios de seguimiento. En algunas realizaciones, la retroalimentación se proporciona dentro del contexto de una aplicación en ejecución. En una realización, se presenta la tecnología para ayudar al usuario a seleccionar un área de captura. De manera adicional, la selección de una respuesta de retroalimentación se puede determinar según los criterios de algunas realizaciones.

45 Este Compendio se proporciona para introducir una selección de conceptos de una forma simplificada que se describirán de manera adicional a continuación en la Descripción Detallada. Este compendio, no está destinado a identificar las características clave o las características esenciales del tema reivindicado, ni está destinado a ser usado como una ayuda para determinar el alcance del tema reivindicado.

Breve descripción de los dibujos

La Figura 1A ilustra una realización de ejemplo de un sistema de reconocimiento, análisis y seguimiento de objetivos en el que pueden operar las realizaciones tecnológicas para proporcionar la retroalimentación de usuario.

50 La Figura 1B ilustra otro ejemplo de una respuesta de retroalimentación del seguimiento de usuario para sugerir al usuario tomar una acción para satisfacer los criterios de seguimiento.

La Figura 1C ilustra algunos ejemplos de factores de visibilidad.

La Figura 1D ilustra otra realización en la que un usuario que sujeta un objeto interactúa con una realización de un sistema de reconocimiento, análisis, y seguimiento del objetivo.

La Figura 1E ilustra un ejemplo de los indicadores visuales que pueden informar al usuario cuando está demasiado cerca o fuera de un límite.

Las Figuras 1F y 1G ilustran ejemplos adicionales de indicadores visuales que informan a un usuario de que están fuera del campo de visión de un dispositivo de captura.

- 5 La Figura 2 es una realización ilustrativa de una arquitectura funcional implementada por ordenador para un sistema para proporcionar retroalimentación del seguimiento del usuario.

La Figura 3A ilustra un ejemplo detallado de una realización de un entorno informático que se puede usar en una consola de juegos como en las FIG. 1A-1E en las que pueden operar una o más realizaciones para proporcionar la retroalimentación del seguimiento del usuario.

- 10 La Figura 3B ilustra otra realización de ejemplo de un entorno informático tal como un ordenador personal.

La Figura 4 ilustra una realización de ejemplo de un entorno informático en red en el que pueden operar una o más realizaciones para proporcionar la retroalimentación del seguimiento del usuario a un usuario.

La Figura 5 representa un modelo de un usuario humano que se puede crear basado en los datos de imagen capturados y usar para el seguimiento de los movimientos del usuario.

- 15 La Figura 6 representa algunos ejemplos de gestos.

La Figura 7 es un diagrama de flujo de una realización del método para proporcionar retroalimentación a un usuario sobre la capacidad de una aplicación de seguir el movimiento del usuario.

La Figura 8 es un diagrama de flujo de una realización del método para ayudar a un usuario a seleccionar un área de captura.

20 **Descripción detallada**

Se presenta la tecnología para proporcionar retroalimentación a un usuario sobre la capacidad de una aplicación en ejecución de seguir las acciones de un usuario para controlar la aplicación en ejecución en un sistema informático. Un ejemplo de un factor de distinción que puede afectar la capacidad de seguir a un usuario es cuando una parte del cuerpo del usuario que controla un objeto presentado está al menos de manera parcial fuera del campo de visión del sistema de captura de imágenes. Otros factores incluyen los factores ambientales tales como el efecto de la iluminación, ciertos tipos de obstrucciones y factores de audio tales como el volumen y la distinción del habla (por ejemplo sílabas o palabras).

- 25

En algunas realizaciones, la tecnología presenta respuestas de retroalimentación que son sugerencias explícitas a un usuario. En otras realizaciones, la retroalimentación es sutil o implícita siendo proporcionada dentro del contexto de una aplicación en ejecución. Por ejemplo, cuando un usuario está demasiado cerca de un límite del campo de visión de un sistema de captura, un objeto tal como un monstruo aterrador aparece dentro de la escena en el lado de proyección cercano al límite. El usuario está motivado a moverse lejos del límite del campo de visión hacia el centro del campo de visión para escapar del monstruo.

- 30

Ejemplos de factores tras los cuales se puede determinar la respuesta de retroalimentación se discuten de manera adicional a continuación.

- 35

La Figura 1A ilustra una realización de ejemplo de un sistema 10 de reconocimiento, análisis y seguimiento de un objetivo con un usuario 18 jugando a un juego en el que pueden operar las realizaciones tecnológicas para proporcionar la retroalimentación de usuario. En este ejemplo, el sistema 10 de reconocimiento, análisis y seguimiento de objetivos reconoce a los seres humanos en su entorno natural, sin tener acoplados a los sujetos dispositivos de detección especiales. Sin embargo, se puede aplicar la tecnología a otros mecanismos de seguimiento de usuarios tales como un sistema basado en sensores en el que el usuario lleve sensores.

- 40

Según la realización de ejemplo, el objetivo puede ser un objetivo humano (por ejemplo el usuario 18), un objetivo humano con un objeto, dos o más objetivos humanos, o similares que puedan ser escaneados para generar un modelo tal como un modelo esquelético, un modelo humano de malla, o cualquier otra representación adecuada del mismo. El modelo puede ser seguido de manera tal que los movimientos físicos o los gestos del objetivo puedan actuar como una interfaz de usuario en tiempo real que ajuste y/o controle los parámetros de una aplicación. Además, el modelo se puede presentar a aplicaciones como un modelo y ser entregado a ellas en tiempo real. Por ejemplo, los gestos seguidos de un usuario se pueden usar para mover un personaje en pantalla o un avatar en un juego electrónico de rol.

- 45

En un ejemplo en el que el modelo es un modelo esquelético multi punto, el sistema 10 de reconocimientos, análisis, y seguimiento de objetivos sigue de manera eficiente a los humanos y sus movimientos naturales en base a los modelos de mecánica y capacidades naturales del sistema muscular y esquelético humano. El sistema 10 de

- 50

ejemplo también reconoce de manera única a los individuos para permitir a múltiples personar interactuar con el sistema a través de los movimientos naturales de sus extremidades y cuerpos.

5 Los movimientos de un usuario se pueden seguir hasta un avatar que puede ser una imagen generada por ordenador que represente a un usuario que normalmente es un humano. El avatar puede representar una imagen del usuario que sea altamente representativa de como se ve el usuario o puede ser un personaje (por ejemplo humano, fantástico, animal, u objeto animado) con cierto grado de parecido con el usuario o ninguno.

10 Específicamente, la Figura 1A, ilustra una realización de ejemplo de una configuración de un sistema 10 de reconocimiento, análisis y seguimiento de objetivos con un usuario 18 que juega un juego de boxeo. En este ejemplo, el software que se ejecuta en un sistema 12 informático que controla o interactúa con software sobre otros sistemas informáticos del sistema 20 de cámara acoplado de manera comunicativa y la unidad 16 de presentación audiovisual sigue los movimientos del usuario 18 en base a los datos de imagen capturados y los analiza para recibir instrucciones ya que los gestos del usuario controlan de manera directa las acciones de su avatar asociado en tiempo real. Por tanto, en este ejemplo, el usuario 18 puede mover su cuerpo para controla su avatar 24 en la pantalla 14 de presentación en el juego de boxeo contra su avatar 22 oponente.

15 El sistema 16 de presentación audiovisual puede ser un sistema de presentación avanzado tal que una televisión de alta definición (HDTV). En otras realizaciones, el elemento de presentación puede ser un elemento de presentación de baja resolución, algunos ejemplos de los cuales incluyen una televisión, un monitor de ordenador, o un elemento de presentación de un dispositivo móvil- El sistema 16 audiovisual puede recibir las señales audiovisuales del sistema 12 informático sobre una interfaz de comunicación (por ejemplo un cable S-Video, un cable coaxial, un cable HDMI, un cable DVI, un cable VGA) y puede emitir después las imágenes y/o el audio de la aplicación asociadas con las señales audiovisuales al usuario 18.

20 Un gesto comprende un movimiento o postura que actúe como una entrada de usuario para controlar una aplicación en ejecución. A través del movimiento de su cuerpo, un usuario puede crear gestos. Por ejemplo, un usuario puede ser capturado en datos de imágenes. Un gesto identificado del usuario se puede analizar para implicar un control de una aplicación o acción a realizar. Por ejemplo, el usuario 18 lanza un golpe en el juego de boxeo de la Figura 1A. La aplicación del juego lo representa en el combate de boxeo, determina si hace contacto sobre el avatar 22 y aumenta la puntuación del usuario si es así. Un gesto puede ser una postura estática, tal como sostener los antebrazos cruzados en frente de su torso o puede ser uno o más movimientos. Además un gesto puede comprender más de una parte del cuerpo, tal como aplaudir las manos entre sí.

30 Por ejemplo, el sistema 10 de reconocimiento, análisis y seguimiento de objetivos se puede usar para reconocer y analizar un puñetazo del usuario 18 en el área de captura 30 de manera tal que se pueda interpretar el puñetazo como un gesto, en este caso un control de juego de un puñetazo para su avatar 24 de jugador a realizar en el espacio de juego. Otros gestos del usuario 18 se pueden interpretar como otros controles o acciones, tales como controles para sacudir, zigzaguear, arrastras los pies, bloquear, golpear o lanzar una variedad de diferentes puñetazos. Mediante el seguimiento de los puñetazos y los golpes del usuario 18, la aplicación de software del juego de boxeo determina la puntuación del avatar 24 y qué avatar (22 o 24) ganará el combate. Las diferentes aplicaciones reconocerán y seguirán los diferentes gestos. Por ejemplo, se sigue un lanzamiento de un usuario en un juego de béisbol para determinar si es un strike o una bola.

40 Según otras realizaciones de ejemplo, el sistema 10 basado en gestos se puede usar además para interpretar los movimientos del objetivo como los controles de un sistema operativo y/o una aplicación que están fuera del ámbito de los juegos. Por ejemplo, de manera virtual cualquier aspecto controlable de una sistema operativo y/o una aplicación se puede controlar mediante los movimientos de un objetivo tal como el usuario 18.

45 El sistema 20 de cámara captura datos de imágenes del usuario en un área 30 de captura dentro del campo de visión del sistema 20 de cámara. En este ejemplo, el área 30 de captura en el campo de visión del sistema 20 de cámara es un trapecoide 30, que desde una perspectiva de usuario, tiene una línea 30f más corta que el frontal del área de captura, una línea 30b negra (por ejemplo una pared puede formar esta línea) y un lado 30l izquierdo y un lado 30r derecho. El campo de visión puede tener diferentes geometrías. Por ejemplo, los límites y las obstrucciones del área de captura pueden afectar a su geometría. Por ejemplo, si los usuarios están jugando en un gimnasio, la pared trasera puede estar mucho más atrás por lo que el campo de visión tiene más forma de cono que trapecoidal. En otros ejemplos, el tipo de lente de la cámara puede afectar al campo de visión también.

50 En la Figura 1A, el usuario se muestra en una posición 18b previa o anterior en la que parte de su cuerpo está fuera del área 30 de captura en el límite 30l izquierdo. La posición del usuario 18 en su posición actual está más cerca del centro del área 30 de captura y del campo de visión de la cámara. Como se discutirá con un mayor detalle más adelante, la retroalimentación en el contexto de la aplicación se proporciona para sugerir al usuario o motivar al usuario a moverse a su derecha. La posición del otro avatar 22 se puede mover a la derecha de la pantalla para que el usuario 18 necesite moverse para golpearle. En otro ejemplo, la perspectiva de la vista desplegada se puede cambiar desde una vista de perspectiva izquierda a una vista de perspectiva derecha del avatar 22 oponente. En la vista de perspectiva derecha, si el usuario 18 se queda en la posición 18b, su avatar bloqueará la vista de su avatar 22 oponente. Esta acción alienta al usuario 18 para moverse a la derecha para ver mejor a su avatar 22 oponente.

Se pueden usar también uno o más elementos de presentación fuera de pantalla para proporcionar retroalimentación de seguimiento del usuario. En el ejemplo ilustrado, una luz 32 de un elemento de presentación tal como un diodo emisor de luz (LED) puede ser satisfecha con un usuario concreto. Por ejemplo, se pueden usar diferentes colores para mostrar la calidad del seguimiento. El amarillo puede ser una advertencia; el verde puede ser satisfactorio; el rojo puede indicar un problema. En otro ejemplo, los diferentes patrones de iluminación pueden indicar la calidad del seguimiento. Por ejemplo, cada luz se puede asociar con un límite del campo de visión. Si un elemento de luz toma un determinado color, el color es la retroalimentación de que un usuario esté demasiado cerca del límite.

La Figura 1B ilustra otro ejemplo de una respuesta de retroalimentación del seguimiento de usuarios para sugerir al usuario tomar una acción para satisfacer un criterio de seguimiento. El cambio en la apariencia de un avatar se puede usar para informar a un usuario de un problema de seguimiento. El avatar 24 está resaltado en este ejemplo. Además, la retroalimentación explícita en forma de flecha 33 es un ejemplo de un indicador visual. En este ejemplo, la retroalimentación recomienda al usuario asociado con el avatar 24 moverse a la derecha. Estos indicadores visuales son ejemplos de retroalimentación del seguimiento de usuarios independiente de la aplicación ya que no están vinculados al contexto de la aplicación.

El contexto de una aplicación comprende la actividad que es el propósito de la aplicación. Por ejemplo, en una aplicación de interfaz de usuario de un menú, abrir o cerrar un fichero sería contextual a la aplicación. Los avatares y los objetos de la escena que se mueven según la acción de un juego son contextuales al juego. Algunos ejemplos de acciones que son contextuales en una aplicación de juego son lanzar un puñetazo, la llegada de un nuevo enemigo o monstruo así como un obstáculo, donde una bola es lanzada o cogida, un cambio en el escenario según se mueve un avatar o la vista de un usuario a través de un entorno virtual, o el cambio de dirección o de perspectiva de la vista de la acción del juego de un usuario.

La Figura 1C ilustra algunos ejemplos de factores de visibilidad. La luz del sol 45 que entra a través de la ventana 44 puede provocar que la iluminación ambiente sea un problema de visibilidad para el sistema 20 de cámara. Un medidor de luz del sistema 20 de captura de imagen puede indicar que el usuario está siendo arrastrado fuera de los datos de imagen. De manera alternativa, podría ser demasiado oscuro.

La mesa 15 de café es un ejemplo de una obstrucción que puede bloquear una parte del cuerpo del usuario. El juego de boxeo puede tener dificultad detectando el gesto de arrastrar los pies en el boxeo debido a que las piernas 18 del usuario están siendo parcialmente oscurecidas por la mesa de café.

La Figura 1D ilustra otra realización en la que un usuario que sostiene un objeto interactúa con una realización de ejemplo de un sistema 10 de reconocimiento, análisis, y seguimiento de objetivos. En dichas realizaciones, el usuario 19 de un juego electrónico puede estar sosteniendo el objeto de manera tal que los movimientos del usuario 19 y del objeto 21 se pueden usar para ajustar y/o controlar los parámetros del juego, tales como, por ejemplo, golpear una bola 68 en pantalla. En otros ejemplos, el movimiento de un usuario que sostiene una raqueta 21 se puede seguir y utilizar para controlar una raqueta en pantalla en un juego de deportes electrónico. En otra realización de ejemplo, el movimiento de un usuario que sostiene un objeto se puede seguir y utilizar para controlar un arma en pantalla en un juego de combate electrónico. Cualquier otro objeto se puede incluir también tal como uno o más guantes, bolas, bates, garrotes, guitarras, micrófonos, palos, mascotas, animales, tambores y similares.

La Figura 1D ilustra otro ejemplo de uso de retroalimentación en el contexto de una aplicación para proporcionar una sugerencia o una motivación a un usuario para moverse a una posición mejor dentro del campo de visión del sistema 20 de cámara. El usuario en la posición 19b se ha movido demasiado lejos a la izquierda con respecto a la cámara por lo que su parte superior del cuerpo que incluye su brazo derecho y la raqueta 21 están fuera del límite 30l izquierdo del campo de visión. Una bola 68b entrante anterior se hubo de presentar en la esquina superior izquierda de la pantalla. La aplicación en ejecución para este juego coloca la siguiente bola 68 entrante en el medio del elemento de presentación para motivar al usuario 19 a moverse al centro del campo de visión que cubre el área 30 de captura lo cual el usuario 19 ha hecho como se ilustra en su posición actual 19.

Se debería tener en cuenta que la Figura 1D presenta una alternativa en donde no se representa una representación del usuario en pantalla. En la Figura 1D, el usuario puede estar jugando al tenis en primera persona, viendo la bola representada en pantalla y, en un ejemplo, todo o una parte de la raqueta que golpea la bola, pero no se representa el avatar del usuario en pantalla o sólo se representa una parte de las partes del cuerpo del usuario.

La Figura 1E ilustra un ejemplo de indicadores visuales que pueden informar a un usuario cuando está demasiado cerca o fuera de un límite. De manera particular para un usuario que sea un niño muy pequeño, los objetos presentados 28l (para problemas con el límite izquierdo) y 28r (para problemas con el límite derecho) se pueden retroalimentar para informar al usuario de moverse de nuevo al centro del campo de visión. Indicadores visuales similares se pueden usar con los límites frontal y trasero también. De manera opcional, se puede asociar un sonido con cada indicador visual según aparece o después de un cierto número de infracciones de límite.

La Figura 1F ilustra otro ejemplo de un indicador visual donde un usuario 19u está jugando un juego de bolos y un brazo 19-l del usuario se mueve fuera del campo de visión. En la representación 28b en pantalla, el brazo 28A del

usuario está desvanecido para mostrar al usuario que el brazo está fuera del campo de visión de la cámara. Otras alternativas para cambiar la apariencia en pantalla del usuario o de los elementos en los que participa el usuario incluyen el cambio de color de toda o parte de la representación en pantalla del usuario, el cambio del foco de la representación del usuario (por ejemplo haciendo la representación en pantalla difuminada o desvanecida).

5 Aún otro ejemplo que se muestra en la Figura 1G tiene un gráfico 29 de advertencia ilustrado en la esquina superior izquierda del elemento de presentación 16. El gráfico presenta una pequeña representación del usuario 34 y del campo 35 de visión, y puede destellar de manera constante o intermitente cuando el usuario se mueve fuera del campo de visión.

10 El software que proporciona la retroalimentación del seguimiento del usuario puede proporcionar entrenamiento para el usuario en el elemento de presentación para ayudar al usuario a adquirir un sentido de los límites del área de captura y qué significan las diferentes respuestas de retroalimentación. Un sonido determinado se puede identificar con ser centrado en el campo de visión o tener una buena visibilidad, mientras que otro sonido indica que el usuario está demasiado cerca de un límite o que existe una obstrucción u otro efecto o elemento que degrada la calidad del seguimiento.

15 En un ejemplo, la calidad de seguimiento o los criterios de seguimiento se pueden basar en cuántos gestos no fueron capaces de ser identificados en un periodo de tiempo dado mientras se ha establecido la presencia y participación del usuario con el sistema de reconocimiento y seguimiento del objetivo. La calidad de seguimiento o los criterios de seguimiento se pueden basar en detectar la presencia y la participación, pero no ser capaz de reconocer una parte del cuerpo clave en los datos de imagen para la aplicación tal como, por ejemplo, un brazo en un juego de béisbol. Junto con los factores de visibilidad que afectan a la calidad o los criterios de seguimiento, los factores de distinción pueden aplicar también factores de audio tal como algunas aplicaciones dependen de una característica del cuerpo tal como la voz y no sólo de los movimientos de las características del cuerpo que son partes del cuerpo.

20 La Figura 2 es una realización ilustrativa de una arquitectura funcional implementada por ordenador para un sistema 200 para proporcionar retroalimentación del seguimiento de usuarios. Dicho sistema de arquitectura se puede implementar como uno o más módulos que se pueden operar por software que se ejecutan en uno o más procesadores y/o hardware informático o como hardware o firmware.

25 La tecnología se puede realizar en otras formas específicas sin salir del espíritu de las características esenciales de la misma. Igualmente, el nombramiento particular y la división de módulos, rutinas, características, atributos, metodologías y otros aspectos no son obligatorios, y los mecanismos que implementan la tecnología o sus características pueden tener diferentes nombres, divisiones y/o formatos. Además, como será evidente para alguien de conocimiento ordinario en la técnica relevante, los módulos, rutinas, atributos, metodologías y otros aspectos de las realizaciones descritas se pueden implementar como software, hardware, firmware o cualquier combinación de los tres. Por supuesto, donde quiera que un componente, un ejemplo del cual es un módulo, se implemente como software, el componente se puede implementar como un programa autónomo, como parte de un programa mayor, como una pluralidad de programas separados, como una librería enlazada de manera estática o dinámica, como un módulo cargable de núcleo, como un controlador de dispositivo, y/o en todas y cada una de las otras maneras conocidas ahora o en el futuro por aquellos de habilidad ordinaria en el arte de la programación. Por ejemplo, en una realización, el software 213 de retroalimentación del seguimiento discutido anteriormente se puede implementar de manera parcial en una interfaz de programación de aplicaciones (API) para manejar la respuesta independiente de la retroalimentación de la aplicación, y de manera parcial en software o una aplicación específica para manejar la retroalimentación contextual de la aplicación.

30 El sistema 202 de control de visualización del elemento de presentación comprende un módulo 204 de movimientos el cual accede a una memoria intermedia 223 de datos para los datos 205 de imagen de entrada y, de manera opcional, los datos 217 de audio. En la realización de ejemplo mostrada, el sistema 202 de control de visualización recibe los datos 205 de seguimiento del movimiento de manera local desde el sistema 20 de captura de datos audiovisuales. De manera adicional, el sistema 202 de control de visualización del elemento de presentación puede recibir los datos 205i de seguimiento del movimiento de manera remota a través de Internet 203 u otra red. Con respecto a un usuario, los datos de seguimiento del movimiento pueden comprender los datos de imagen en sí o una versión degradada de esos datos. De manera adicional, los datos de profundidad, y diversas formas de datos derivados de la imagen y/o los datos de profundidad, se pueden incluir en los datos de seguimiento del movimiento, algunos ejemplos de los cuales son un modelo para el cuerpo del usuario, los datos de movimiento en matemáticas primitivas que hacen referencia al modelo, o una imagen de máscara de bits derivada para el usuario para la comparación con un estado anterior del modelo. El sistema 202 de control de visualización del elemento de presentación analiza estos datos para reconocer el movimiento del usuario y seguir el movimiento hasta los objetos en el elemento de representación, por ejemplo, hasta el avatar del usuario en pantalla. Un avatar es un tipo de escena u objeto presentado.

55 El módulo 204 de movimientos se acopla de manera comunicativa a un módulo 209 de control de elementos de presentación del avatar, un módulo 211 de control de objetos, el software 213 de retroalimentación del seguimiento, el software y los datos 206 de reconocimiento de gestos, un sistema 20 de captura de datos audiovisuales, e Internet

60

203. El módulo 204 de movimientos también tiene acceso a los almacenes de datos 220 almacenados en la memoria tales como los datos 216 de modelo para al menos uno de cada uno de entre un usuario 216u, un avatar 216a, y un objeto 216o en el elemento de representación o alguno sostenido por el usuario. Los datos del modelo se usan como una referencia para el seguimiento del movimiento mediante el módulo 204 de movimientos de un usuario en un área de captura. o de un avatar o un objeto en una escena o vista de un elemento de presentación. En los sistemas donde los movimientos del cuerpo del usuario se hacen corresponder a los movimientos del avatar (por ejemplo en base a la captura de imagen del usuario, o los sensores en el cuerpo del usuario), pueden haber datos del modelo que representen el usuario 216u y datos del modelo que representen el avatar 216a. Donde el físico del avatar es bastante diferente, el módulo 204 de movimientos realiza la correspondencia entre los dos modelos. Por ejemplo, el usuario 18 muchacho es más bajo y probablemente no tiene el alcance del brazo de su avatar el boxeador 24. En otras palabras, si los modelos esqueléticos se usasen, no serían los mismos para el usuario y para el avatar. En algunas realizaciones, sin embargo, la aplicación usa los mismos datos 216 del modelo para analizar los movimientos del muchacho de un usuario y para dirigir los movimientos del avatar correspondiente. En un ejemplo, el modelo del muchacho se puede implementar como uno o más estructuras de datos que representan las partes del muchacho y sus posiciones en las dimensiones y/o los ángulos de rotación con respecto a una referencia. Los datos 216 del modelo se pueden actualizar en términos de posiciones absolutas o con cambios en las posiciones y rotaciones. Los cambios en las posiciones y las rotaciones se pueden representar como vectores y ángulos.

El módulo 204 de movimientos también tiene acceso a los datos 214 de perfil. En este ejemplo, existen perfiles para los usuarios 214u, los avatares 214a y los objetos 214o. El módulo de movimiento 204 también tiene acceso a los datos 218 de presentación que incluyen los datos 219 de imagen del avatar y los datos 221 de imagen del objeto.

El módulo 209 de control de presentación del avatar actualiza los datos 219 de la imagen del avatar basado en los gestos reconocidos por el software 206 de reconocimiento de gestos y otros movimientos aplicables identificados por el módulo 204 de movimientos. En un ejemplo, los datos 219 de imagen que representan los movimientos o las posturas pueden ser datos de seguimiento del movimiento para el avatar. En un ejemplo, dichos datos de seguimiento del movimiento se pueden almacenar en un archivo de capturas de movimientos que el módulo 204 de movimientos actualiza en el tiempo según se reciben nuevos datos 205 de seguimiento del movimiento.

El módulo 211 de control de objetos actualiza los datos 221 de imagen para los objetos afectados por los gestos reconocidos del usuario. Además, el módulo 209 de control de avatares y el módulo 211 de control de objetos actualizan sus datos (219, 221) de imagen respectivos en respuesta a las instrucciones del módulo 210 de control de acciones. El módulo 210 de control de acciones supervisa la aplicación en ejecución. Por ejemplo, en un entorno de un juego, mantiene la puntuación, determina que es necesario un nuevo fondo ya que el avatar se ha movido a un nuevo nivel del juego, determina como otros avatares u objetos no controlados por el usuario se colocarán en la escena. En un entorno no de juego, identifica que acción está solicitando el usuario. Por ejemplo, si un gesto de usuario es una solicitud para abrir un archivo, puede acceder al archivo en sí o proporcionar instrucciones al módulo 211 de control de objetos para acceder al archivo y presentarlo en el elemento 16 de presentación. El módulo 207 de procesamiento de la presentación combina los datos 218 del elemento de presentación para actualizar la presentación.

Los datos 218 del elemento de presentación proporcionan una escena o un contexto de visión y definen los otros objetos en la escena o vista. Por ejemplo, los datos 218 del elemento de presentación están proporcionando un entorno de contexto de un combate de boxeo en las Figuras 1A y 1B, y las posiciones de los objetos incluyendo las cuerdas del ring de boxeo y la existencia del avatar 22 del otro boxeador oponente. Los movimientos, características y partes del cuerpo más recientes del avatar 22 del otro boxeador, en la escena se pueden almacenar también como otro perfil en el almacén de datos 214 de perfiles de avatar. Las actualizaciones de los datos 218 del elemento de presentación se pueden recibir mediante uno o más módulos tales como el módulo 209 de control de la presentación del avatar o el módulo 204 de movimientos sobre Internet 203 u otra red desde un módulo remoto.

El módulo 209 de control la presentación del avatar y el módulo 211 de control de objetos actualizan de manera periódica o en respuesta a un mensaje desde el módulo 204 de movimientos su perfil o perfiles 214 respectivos y procesan los datos 219, 221 de imagen que representan los movimientos o las posturas para el avatar o el objeto. Los datos 219, 221 de imagen se pueden representar de manera local en un elemento 16 de presentación o se pueden transmitir a través de internet 203 u otra red.

Ejemplos de los datos 214a del perfil del avatar pueden ser los datos del color de la imagen para características del avatar tales como el pelo, los rasgos faciales, el color de la piel, la ropa, la posición del avatar en la escena y otras propiedades asociadas con él.

Ejemplos de información que se puede almacenar en el perfil 214u de un usuario pueden incluir los modos típicos de uso o juego, la edad, la altura, la información del peso, los nombres, discapacidades, puntuaciones más altas o cualquier otra información asociada con un usuario y el uso del sistema.

Un ejemplo de un factor que afecta a la selección de una respuesta de retroalimentación puede incluir la información de perfil. Por ejemplo, la edad o discapacidades de un usuario, físicas o mentales, pueden hacer un tipo de respuestas de retroalimentación más apropiadas que otras. Por ejemplo, una persona de 5 años de edad puede no

ser capaz de leer por lo que una retroalimentación sutil en el contexto de la aplicación puede ser más apropiada que un texto explícito en la pantalla. Un jugador puede ser sordo por lo que una respuesta de audio no es apropiada.

Como se discutió anteriormente, algunos movimiento y posturas, gestos, tienen un significado especial en el contexto de un programa de entretenimiento u otras aplicaciones, y el sistema 202 de control de visualización del elemento de presentación ejecuta las instrucciones para identificarlos. En concreto, el módulo 204 de movimientos tiene que acceder al software 206 de reconocimiento de gestos y los datos para el reconocimiento o la identificación de los gestos basado en un modelo y unos datos de seguimiento del movimiento.

El software 206 de reconocimiento de gestos puede incluir filtros de gestos. En un ejemplo, el módulo 204 de movimientos puede seleccionar uno o más filtros 206 de movimientos basados en un índice de datos asociativo, tal como un índice de partes del cuerpo por ejemplo. Por ejemplo, cuando un conjunto de datos de seguimiento del movimiento actualizado es recibido por el sistema 202 de control de visualización del elemento de presentación y se indican cambios de movimiento para ciertas partes del cuerpo, el módulo 204 de movimientos indexa los filtros de gestos asociados con esas partes del cuerpo determinadas.

Los filtros 206 de gestos ejecutan instrucciones basadas en datos de parámetros que definen criterios para determinar si un gesto en concreto se ha realizado basado en los datos 205 de seguimiento del movimiento. En una realización, cada filtro 206 de gestos es enlazado con un módulo de librería para un gesto particular en una librería de gestos. Cada módulo de librería asociado con un gesto incluye instrucciones ejecutables para realizar el procesamiento de la respuesta al gesto. Este procesamiento a menudo implica la actualización del movimiento o la imagen del avatar para reflejar el gesto de alguna forma.

Por ejemplo, un usuario 18 adolescente en la Figura 1A tiene un físico mucho menor que su avatar boxeador 24. Parte de la diversión de jugar juegos con avatares es que ellos a menudo hacen cosas mejor que el usuario en la vida real. El puñetazo del chico 18 en la Figura 1A en su sala de estar se traduce en un puñetazo mucho más intenso de su avatar 24 en la pantalla 14. El módulo de librería para un "puñetazo" en el juego de boxeo puede determinar la aceleración en la velocidad del puño del usuario que se extiende por su codo, esto es un "puñetazo poderoso" para este usuario 18 en el juego, y el puñetazo 24 de su avatar refleja esto en el elemento de presentación debido a las instrucciones del módulo de puñetazos de la librería. El módulo 210 de control de acciones es notificado también del gesto y responde a este. Por ejemplo, actualiza la puntuación del usuario 18.

Algunos sistemas pueden proporcionar una combinación de sólo un determinado número de movimientos o posturas que el avatar puede realizar para determinadas partes o regiones del cuerpo, por ejemplo las manos y los brazos, mientras que permite un seguimiento directo de otras partes o regiones del cuerpo, por ejemplo las piernas.

El seguimiento de los movimientos del usuario para actualizar la presentación de un avatar u otros objetos visualizados en el elemento de presentación se realiza en tiempo real de manera tal que el usuario pueda interactuar con una aplicación en ejecución en tiempo real. Una presentación en tiempo real se refiere a la presentación de una representación visual en respuesta a un gesto, en donde la presentación se presenta de manera simultánea o casi simultánea con la ejecución del gesto en el espacio físico. Por ejemplo, una tasa de presentación actualizada en la que el sistema puede proporcionar una presentación que refleja a un usuario puede ser una tasa de 20Hz o superior, en donde los retardos de procesamiento insignificantes resultan en un retardo mínimo de la presentación o no son visibles para todos los usuarios. Por tanto, el tiempo real incluye cualquier retardo insignificante correspondiente a la puntualidad de los datos que se han sido retardados por el tiempo requerido para el procesamiento automático de los datos.

El software 213 de retroalimentación del seguimiento recibe un mensaje desde el módulo 204 de movimientos identificando un problema de seguimiento. Un ejemplo de problema de seguimiento es que el usuario se haya movido fuera del campo de visión o que no haya ningún usuario en el campo de visión. Algunos otros ejemplos de problemas de seguimiento son un gesto para una parte del cuerpo que no se puede determinar o que no se satisfaga una probabilidad umbral de que un movimiento o postura corresponda a un gesto particular. Otro ejemplo de problema de seguimiento es que el volumen del sonido no sea suficiente para el software de reconocimiento del habla o del canto para detectar el sonido del usuario. En otro ejemplo, el sonido del usuario tiene sílabas indistinguibles por lo que las palabras no se pueden detectar. Por ejemplo, en una aplicación de entretenimiento musical, la incapacidad para identificar o detectar el canto del usuario puede afectar de manera significativa a la frustración sentida por el jugador usuario.

El módulo 204 de movimientos puede proporcionar también retroalimentación sobre un número de factores de distinción basado en los datos que ha recibido desde el sistema 20 de captura audiovisual o en los mensajes del sistema 20 de captura audiovisual. Un tipo de factor de distinción son los factores de audio como un volumen insuficiente y el carácter distintivo como se mencionó anteriormente. Otro ejemplo de factor de distinción es un factor de visibilidad. Un ejemplo de un factor de visibilidad es el usuario o una parte del cuerpo para controlar la aplicación que está al menos de manera parcial fuera del campo de visión. Otros son problemas de iluminación borrando al usuario de la imagen capturada. Otro ejemplo es una obstrucción. Otro ejemplo de un factor de visibilidad es una obstrucción. Un ejemplo de esto puede ser un mueble, incluso en el campo de visión en sí como la mesa 15 de café en la Figura 1C. Demasiados elementos en los límites del campo de visión o en la visión en sí

- 5 pueden crear un factor de distinción o de visibilidad. Otro ejemplo de una obstrucción es un artículo de ropa o ropa en combinación con otro elemento. Una parte superior de un poncho puede inhibir la detección de los movimientos de los brazos. Una falda larga puede inhibir la detección de los movimientos de las piernas para un juego en que un usuario puede hacer gestos de andar o correr. Pantalones negros contra un sofá negro pueden provocar un problema de visibilidad debido a la ausencia de contraste; el sistema 20 de captura no puede distinguir los movimientos de las piernas cuando el usuario está cerca del sofá negro. El color de la ropa puede provocar problemas de contraste que resultan en un problema de distinción también, como pueden superponerse sus brazos y piernas. Los datos de profundidad pueden ayudar a distinguir al usuario, pero esto puede ser un problema en base a la resolución de profundidad del sistema 20 de captura de imagen.
- 10 El software 213 de retroalimentación del seguimiento puede seleccionar un tipo de retroalimentación que indica un problema de seguimiento para proporcionarlo a un usuario. En algunos ejemplos, la retroalimentación proporciona instrucciones al usuario para mejorar la calidad del seguimiento con un indicador visual o de audio que es independiente de la actividad de la aplicación. Por ejemplo, la superposición explícita 33 de la flecha de pantalla en la Figura 1B. Puede estar acompañada también de un mensaje de texto o de audio para “Moverse a la Derecha”
- 15 mientras apunta al centro del campo 30 de visión.
- Otra retroalimentación puede ser un poco más implícita y sutil. Las características visuales del elemento de presentación se pueden cambiar. Por ejemplo, la retroalimentación puede ser que la nitidez de la visión del elemento de presentación se puede degradar ligeramente según el usuario se aproxima dentro de una distancia de un límite del campo de visión. Se pone borroso por ejemplo. Según el usuario se mueve hacia el centro del campo de visión,
- 20 la nitidez de la escena mejora. El usuario puede incluso no ser consciente del cambio en la nitidez. Junto a la nitidez, otro ejemplo de característica visual que se puede cambiar para alterar la presentación de la escena o la visión es la vitalidad del color. Otro ejemplo es el color en sí. Por ejemplo, según el usuario se mueve fuera del campo de visión, el color en el elemento de presentación se torna en blanco y negro.
- En algunos ejemplos, el software 213 de retroalimentación del seguimiento determina que la retroalimentación en el contexto de la aplicación es apropiada. Algunos de estos ejemplos se han mencionado anteriormente tales como un monstruo en el límite cuando el usuario está demasiado cerca o una bola para que el usuario “golpee” siendo enviada a una dirección a la que el usuario debería moverse. Otros ejemplos incluirían un enemigo al que disparar en una dirección a la que moverse, u otro jugador avatar que entra y golpea al usuario de vuelta hacia el centro u otros personajes que corren en una dirección para volver al campo de visión. En otro ejemplo, se puede dirigir un sonido en una determinada dirección para motivar al usuario a correr hacia este lejos de esta dirección. Si el usuario está demasiado lejos del sistema 20 de captura de imagen como se muestra en los datos de profundidad capturados por el sistema 20, se puede usar un objeto parpadeante u otro objeto en el elemento de presentación que capte nuestra atención para atraer al usuario hacia dentro del campo de visión. Si un usuario se está acercando demasiado a la cámara 20, por ejemplo entre el límite frontal del campo de visión, 30f y la cámara, se puede hacer un objeto de presentación para llenar bastante la pantalla para hacer al usuario tomar unos pasos hacia atrás.
- 25
- 30 Para esta retroalimentación contextual de la aplicación, el módulo 213 de retroalimentación del seguimiento envía una solicitud al módulo 210 de software de control de acciones para la retroalimentación contextual. El módulo 210 de control de acciones o el software 213 de seguimiento pueden seguir qué técnicas de retroalimentación contextual se han usado para evitar la repetición tanto como sea posible. La solicitud de retroalimentación contextual puede incluir el tipo de factor de distinción a ser abordado. Puede incluir además una sugerencia de acción. Por ejemplo, puede incluir también una zona objetivo en el elemento de presentación en la que colocar un objeto para motivar el movimiento de un usuario de vuelta al campo de visión. En otros casos, el módulo 210 de control de acciones puede determinar la acción a tomar.
- 35
- El software 213 de retroalimentación del seguimiento puede basar su selección de una respuesta de retroalimentación en ciertos criterios. Como se mencionó anteriormente, un ejemplo de dichos criterios es la edad del usuario. Para una persona de 3 o 4 años de edad, colocar un elemento en una zona objetivo en la pantalla para hacer que el usuario se mueva puede no ser muy repetitivo para un niño de esa edad. Para una persona de 7 años de edad, la técnica de retroalimentación contextual específica puede necesitar ser un poco más variada.
- 45
- El nivel de competitividad de la aplicación puede ser otro factor. En un juego que se juega contra otro usuario, poner un objetivo al que disparar en una zona del elemento de presentación para animar al movimiento del usuario puede no ser apropiado. Sin embargo, colocar una explosión cerca del límite del campo de visión puede hacer que el usuario se mueva sin aumentar la previsibilidad de los objetivos.
- 50
- Algunas aplicaciones no quieren nada que complique la acción que se está presentando por la aplicación en el elemento de presentación. En estos ejemplos, el elementos 16 de presentación o la consola 12 pueden incluir dispositivos de presentación fuera de pantalla. Por ejemplo, el diodo 32 emisor de luz (LED) en la consola 12 o en la cámara 20 se puede asociar con el usuario, y la capacidad de la aplicación para seguir al usuario se puede indicar mediante una paleta de colores. Por ejemplo, verde es bueno. Amarillo es una advertencia de que la capacidad de seguimiento se está degradando, y rojo indica que la capacidad de seguimiento se ha degradado hasta un nivel inaceptable. Otros dispositivos o vistas de presentación fuera de pantalla se pueden usar tales como gráficos de
- 55

barras u otras luces en la consola 12, el elemento 16 de presentación, o la cámara 20 que indiquen el grado de capacidad de seguimiento.

5 Como compromiso, la aplicación puede permitir que aparezca un pequeño icono en un avatar del usuario para un usuario que esté demasiado cerca de un límite o cuyo seguimiento no esté satisfaciendo los criterios de seguimiento. El usuario puede seleccionar el icono si lo desea. En otro ejemplo, se puede presentar una pequeña caja, imagen en imagen, que muestra al usuario o su avatar indicando un problema de seguimiento e incluso una sugerencia para abordarlo. En otro ejemplo, si el usuario presiona la pausa, se puede mostrar una caja que muestra al usuario o su avatar con el mensaje del problema de seguimiento.

10 En otros ejemplos, el módulo 213 de retroalimentación del seguimiento puede enviar una solicitud para un cambio en la apariencia de un avatar al módulo 209 de control de la presentación del avatar. Un ejemplo de un cambio en la apariencia puede ser destacar el avatar. Otros ejemplos incluyen el cambio de características visuales tales como difuminar la apariencia del avatar, o hacerlo todo negro o todo blanco o en blanco o negro como opuesto al color. El avatar se puede hacer para parecer desvanecido en otro ejemplo. De manera particular si una sesión de entrenamiento ocurre antes del comienzo de la aplicación, el usuario puede recibir instrucciones de que un cambio en la apariencia de su avatar signifique que existen problemas de seguimiento con sus gestos. Para los problemas de visión, puede ser efectivo sin cambiar demasiado la acción competitiva.

20 La retroalimentación puede ser de audio o audiovisual. Por ejemplo, en una sesión de entrenamiento antes del comienzo de la actividad de la aplicación, un usuario puede recibir instrucciones de que un icono en concreto significa que un factor de visibilidad concreto está efectuando el reconocimiento de los gestos del usuario. El icono puede ser acompañado por un sonido concreto para que el usuario sepa que es el quien, por ejemplo, se ha mover hacia dentro.

El no ser capaz de detectar los gestos del usuario de manera apropiada puede afectar de manera significativa a la ejecución de la aplicación, pausar la acción, quizás junto con un cambio en la apariencia del avatar del usuario, puede ser una respuesta de retroalimentación seleccionada también. En otro ejemplo, se puede detener el sonido.

25 En algunas realizaciones, el software de retroalimentación del seguimiento se puede separar de la aplicación concreta. Por ejemplo, el software de retroalimentación del seguimiento puede proporcionar una API para la cual una aplicación puede enviar una solicitud para una respuesta de retroalimentación. Esta puede ser una interfaz conveniente para los desarrolladores de aplicaciones que desean usar respuestas de retroalimentación por defecto. Por supuesto, se pueden usar otras construcciones además de una API. El software de aplicación puede proporcionar tipos adicionales de respuestas de retroalimentación del seguimiento de usuario a aquellas proporcionadas por la API o que los desarrolladores de aplicaciones prefieran usar en lugar de los mecanismos por defecto. En otras realizaciones, el software de retroalimentación del seguimiento se puede manejar enteramente dentro de la aplicación o enteramente mediante software independiente a la aplicación. La tecnología descrita en la presente memoria no se limita a una implementación de nivel de código concreta. El sistema 20 de captura de imagen reconoce objetivos humanos y no humanos en un área de captura (con o sin dispositivos de detección especiales acoplados a los sujetos), únicamente los identifica y los sigue en el espacio tridimensional.

40 Según una realización de ejemplo, el sistema 20 de captura de imagen se puede configurar para capturar video con información de profundidad que incluye una imagen de profundidad que puede incluir valores de profundidad a través de cualquier técnica adecuada que incluya, por ejemplo, tiempo de vuelo, luz estructurada, imagen en estéreo, o similar. Como se muestra en la FIG. 2, según una realización de ejemplo, un componente 70 de cámara de imagen puede incluir un componente 72 de luz IR, una cámara 74 tridimensional (3-D), y una cámara 76 RGB que se puede usar para capturar la imagen de profundidad de un área de captura. Por ejemplo, en el análisis en tiempo de vuelo, el componente 72 de luz IR del sistema 20 de captura puede emitir una luz infrarroja en el área de captura y puede usar entonces sensores para detectar la luz retrodispersada desde la superficie de uno o más objetivos y los objetos en el área de captura usando, por ejemplo, la cámara 74 3-D y/o la cámara 76 RGB. En algunas realizaciones, la luz infrarroja pulsada se puede usar de manera tal que el tiempo entre el pulso de luz de salida y el correspondiente pulso de luz de entrada se pueda medir y usar para determinar la distancia física desde el sistema 20 de captura hasta una ubicación concreta en los objetivos u objetos en el área de captura. De manera adicional, en otras realizaciones de ejemplo, la fase de la onda de luz de salida se puede comparar con la fase de la onda de luz de entrada para determinar el desplazamiento de fase. El desplazamiento de fase se puede usar entonces para determinar la distancia física desde el sistema de captura hasta la ubicación concreta de los objetivos u objetos.

55 Según otra realización de ejemplo, el análisis de tiempo de vuelo se puede usar para determinar de manera indirecta la distancia física desde el sistema 20 de captura hasta una ubicación concreta en los objetivos u objetos analizando la intensidad del haz reflejado de la luz a lo largo del tiempo a través de diversas técnicas que incluyen, por ejemplo, la imagen de pulso con luz obturada.

En otra realización de ejemplo, el sistema 20 de captura puede usar una luz estructurada para capturar la información de profundidad. En dicho análisis, la luz con patrones (esto es, la luz presentada como un patrón conocido tal como un patrón de cuadros o de rayas) se puede proyectar en el área de captura a través de, por

ejemplo, el componente 72 de luz IR. Tras golpear la superficie de uno o más objetivos u objetos en el área de captura, el patrón en respuesta puede resultar deformado. Dicha deformación del patrón puede ser capturada por, por ejemplo, la cámara 3-D 74 y/o la cámara RGB 76 y puede analizarse después para determinar la distancia física desde el sistema de captura hasta la ubicación concreta de los objetivos u objetos.

- 5 Según otra realización, el sistema 20 de captura puede incluir dos o más cámaras físicamente separadas que puedan ver un área de captura desde diferentes ángulos, para obtener datos visuales en estéreo que se pueden resolver para generar la información de profundidad.

Como un ejemplo de la sinergia proporcionada por estos elementos, considere que el componente 72 de luz IR y la cámara 3-D 74 pueden proporcionar una imagen de profundidad de un área de captura, pero en ciertas situaciones la imagen de profundidad sola puede no ser suficiente para discernir la posición o el movimiento de un objetivo humano. En estas situaciones, la cámara RGB 76 puede “tomar el control” o suplementar la información de la cámara 3-D para permitir un reconocimiento más completo del movimiento o la posición de los objetivos humanos. Por ejemplo, la cámara RGB se puede usar para reconocer, entre otras cosas, los colores asociados con uno o más objetivos. Si un usuario está vistiendo una camisa con un patrón en este que la cámara de profundidad no es capaz de detectar, la cámara RGB se puede usar para seguir ese patrón y proporcionar información sobre los movimientos que el usuario está haciendo. Como otro ejemplo, si un usuario gira, la cámara RGB se puede usar para suplementar la información de uno o más otros sensores para determinar el movimiento del usuario. Como un ejemplo adicional, si un usuario está próximo a otro objeto tal como una pared o un segundo objetivo, los datos RGB se pueden usar para distinguir entre los dos objetos. La cámara RGB puede ser capaz también de determinar finas características de un usuario tal como el reconocimiento facial, el color de pelo y similares que se pueden usar para proporcionar información adicional. Por ejemplo, si un usuario se gira hacia atrás, la cámara RGB puede usar el color del pelo y/o la falta de rasgos faciales para determinar que un usuario está de espaldas al sistema de captura.

El sistema 20 de captura puede capturar datos a tasas interactivas, aumentando la fidelidad de los datos y permitiendo a las técnicas descritas procesar los datos de profundidad sin procesar, digitalizar los objetos en la escena, extraer la superficie y la textura del objeto, y realizar cualquiera de estas técnicas en tiempo real de manera tal que el elemento de presentación (por ejemplo 16) pueda proporcionar una representación de la escena en su pantalla de presentación (por ejemplo 54).

En la realización del sistema de la Figura 2, el sistema 20 de captura de imagen se acopla de manera comunicativa con un entorno informático de manera tal que los ejemplos de sistemas informáticos en las Figuras 3A-3B envíen los datos 205I de seguimiento del movimiento y de manera opcional los datos 217 de audio. El acople de comunicación se puede implementar en una o más conexiones por cable o inalámbricas, tales como, una conexión USB, una conexión Firewire, una conexión por cable Ethernet, o similar y/o una conexión inalámbrica tal como una conexión 802.11b, g, a, o n.

El sistema 20 de captura incluye además un componente 82 de memoria para almacenar las instrucciones que pueden ser ejecutadas por el procesador 80, así como los datos de imagen que se pueden capturar en un formato de trama. El componente 82 de memoria puede incluir memoria de acceso aleatorio (RAM), memoria de sólo lectura (ROM), caché, memoria Flash, un disco duro, o cualquier otro componente de almacenamiento adecuado. En una realización, el componente 82 de memoria puede ser un componente separado en comunicación 90 con el componente 70 de captura de imagen y el procesador 80 como se ilustra. Según otra realización, el componente 82 de memoria se puede integrar dentro del procesador 80 y/o el componente 70 de captura de imagen.

El sistema 20 de captura de imagen incluye además un procesador 80 acoplado de manera comunicativa 90 al componente 70 de cámara de imagen para controlarlo y la memoria 82 para almacenar los datos de imagen. El procesador 80 puede incluir un procesador estandarizado, un procesador especializado, un microprocesador, o similar que pueda ejecutar las instrucciones que puedan incluir instrucciones para almacenar perfiles, recibir los datos de una imagen de profundidad, almacenar los datos en un formato específico en la memoria 82, determinar si un objetivo adecuado se puede incluir en la imagen de profundidad, convertir el objetivo adecuado en una representación esquelética u otro tipo de modelo del objetivo, o cualquier otra instrucción adecuada. Además, parte de este procesamiento puede ser ejecutado por otros procesadores en uno o más entornos informáticos acoplados.

La inclusión de capacidades de procesamiento en el sistema 20 de captura de imágenes permite un modelo tal como un modelo esquelético multipunto, de un usuario para ser ofrecido en tiempo real. En una realización, puede existir un procesador separado para cada uno de los múltiples componentes del sistema de captura, o puede existir un procesador central único. Como otro ejemplo, puede existir un procesador central así como al menos otro procesador asociado. Si hay una tarea de elevado coste computacional, los dos o más procesadores pueden compartir las tareas de procesamiento de cualquier manera. El procesador o los procesadores pueden incluir una memoria como se describe anteriormente y la memoria puede almacenar uno o más perfiles de usuario. Estos perfiles pueden almacenar escaneos corporales, modos de uso o juego típicos, la edad, la altura, información del peso, los nombres, los avatares, las puntuaciones máximas o cualquier otra información asociada con un usuario y el uso del sistema.

El sistema 20 de captura puede incluir además un micrófono 78 que se puede usar para recibir señales de audio producidas por el usuario. Por tanto, en esta realización, el sistema 20 de captura de imagen es un sistema de captura de datos audiovisuales. El micrófono o los micrófonos en el sistema de captura se pueden usar para proporcionar información adicional y suplementaria sobre un objetivo para permitir al sistema discernir mejor aspectos de la posición o el movimiento del objetivo. Por ejemplo, el micrófono o los micrófonos pueden comprender un micrófono o micrófonos direccionales o un conjunto de micrófonos direccionales que se pueden usar para discernir más la posición de un objetivo humano o para distinguir entre dos objetivos. Por ejemplo, si dos usuarios son de forma o tamaños similares y están en una área de captura, los micrófonos se pueden usar para proporcionar información sobre los usuarios de manera tal que se puedan distinguir los usuarios el uno del otro, por ejemplo, en base al reconocimiento de sus voces separadas. Como otro ejemplo, los micrófonos se pueden usar para proporcionar información a un perfil de usuario sobre el usuario, o en una realización del tipo 'lenguaje a texto', donde el al menos un micrófono se puede usar para crear texto en el sistema informático.

Los datos de píxel con los valores de profundidad para una imagen son referidos como una imagen de profundidad. Según una realización, la imagen de profundidad puede incluir un área de píxeles bidimensional (2-D) de la escena capturada donde cada píxel en el área de píxeles 2-D tiene un valor de profundidad asociado tal como una longitud o una distancia en, por ejemplo, centímetros, milímetros, o similar de un objeto en la escena capturada a partir de un punto de referencia, por ejemplo, con respecto a algún aspecto del componente 70 de cámara. Por ejemplo, los valores de profundidad para los píxeles se pueden representar en "capas Z", que son capas que pueden ser perpendiculares a un eje Z que se extiende desde la cámara 70 de profundidad a lo largo de su línea de visión. Estos valores de profundidad pueden ser referidos de manera colectiva como mapa de profundidad.

Una imagen de profundidad se puede degradar a una resolución de procesamiento menor de manera tal que la imagen de profundidad se pueda procesar de manera más fácil y/o más rápida con menos sobrecarga informática. Por ejemplo, diversas regiones de la imagen de profundidad observada se pueden separar en regiones de fondo y regiones ocupadas por la imagen del objetivo. Las regiones de fondo se pueden eliminar de la imagen o se pueden identificar para que puedan ser ignoradas durante uno o más de los pasos de procesamiento posteriores. De manera adicional, uno o más valores de profundidad de alta varianza o ruido se pueden eliminar y/o suavizar de la imagen de profundidad. Las partes de la información de profundidad faltantes y/o eliminadas se pueden rellenar y/o reconstruir. Dicho relleno se puede lograr promediando los vecinos más cercanos, filtrando, y/o cualquier otro método adecuado. Otro procesamiento adecuado se puede realizar de manera tal que la información de profundidad se pueda usar para generar un modelo tal como el modelo esquelético.

La Figura 3A ilustra un ejemplo detallado de una realización de un entorno informático que se puede usar en una consola de juegos tal como en las FIG. 1A-1E en las que puede operar una o más realizaciones para proporcionar retroalimentación del seguimiento de usuario. Como se muestra en la Figura 3A, la consola 12 multimedia tiene una unidad 101 central de procesamiento (CPU) que tiene una caché 102 de nivel 1, una caché 104 de nivel 2, y una ROM (Memoria de sólo lectura) 106 flash. La caché 102 de nivel 1, y la caché 104 de nivel 2 almacena de manera temporal los datos y por lo tanto reduce el número de ciclos de acceso a memoria, mejorando de este modo la velocidad de procesamiento y el rendimiento. La CPU 101 se puede proporcionar teniendo más de un núcleo, y así, las cachés 102 y 104 de nivel 1 y nivel 2. La ROM 106 flash puede almacenar el código ejecutable que se carga durante una fase inicial de un proceso de arranque cuando se enciende la consola 12 multimedia.

Una unidad 108 de procesamiento de gráficos (GPU) y un codificador de video/códec (codificador/decodificador) de video 114 crea un canal de procesamiento de video para el procesamiento de gráficos de alta velocidad y alta resolución. Los datos son transportados desde la unidad 108 de procesamiento de gráficos hasta el codificador de video/códec de video 114 a través de un bus. El canal de procesamiento de video emite los datos a un puerto A/V (audio/video) 140 para su transmisión a una televisión u otro elemento de presentación. Un controlador 110 de memoria se conecta a la GPU 108 para facilitar al procesador el acceso a diversos tipos de memoria 112, tales como, pero no limitado a, una RAM (Memoria de Acceso Aleatorio).

La consola 12 multimedia incluye un controlador 120 de I/O, un controlador 122 de gestión de sistema, una unidad 123 de procesamiento de audio, un controlador 124 de la interfaz de red, un primer controlador USB 126, un segundo controlador USB 128 y un subsistema 130 de I/O de panel frontal que son implementados en un módulo 118. Los controladores USB 126 y 128 sirven como organizadores de los controladores 142(1)-142(2) periféricos, un adaptador 148 inalámbrico, y un dispositivo 146 de memoria externa (por ejemplo, una memoria flash, una unidad ROM de CD/DVD externa, un medio extraíble, etc). La interfaz 124 de red y/o el adaptador 148 inalámbrico proporcionan acceso a una red (por ejemplo, Internet, una red doméstica, etc.) y puede ser cualquiera de entre una amplia variedad de diversos componentes adaptadores por cable o inalámbricos que incluyen una tarjeta Ethernet, un módem, un módulo Bluetooth, un módem de cable, y similares.

La memoria 143 de sistema se proporciona para almacenar los datos de aplicación que se cargan durante el proceso de arranque. Una unidad 144 de medios se proporciona y puede comprender una unidad de DVD/CD, un disco duro, u otra unidad de medios extraíble, etc. La unidad 144 de medios puede ser interna o externa a la consola 100 multimedia. Los datos de aplicación se pueden acceder a través de la unidad 144 de medios para su ejecución, reproducción, etc. por la consola 12 multimedia. La unidad 144 de medios se conecta al controlador 120 de I/O a través de un bus, tal como un bus Serial ATA u otra conexión de alta velocidad (por ejemplo, IEEE 1394).

En una realización, se puede almacenar una copia del software y los datos para el sistema 202 de control de visualización del elemento de presentación en la unidad 144 de medios y se puede cargar en la memoria 143 del sistema cuando se ejecuten.

- 5 El controlador 122 de gestión del sistema proporciona una variedad de funciones de servicio relacionadas con asegurar la disponibilidad de la consola 12 multimedia. La unidad 123 de procesamiento de audio y el códec 132 de audio forman el canal de procesamiento de audio correspondiente con procesamiento estéreo y de alta fidelidad. Los datos de audio son transportados entre la unidad 123 de procesamiento de audio y el códec 132 de audio a través de un enlace de comunicación. El canal de procesamiento de audio emite los datos al puerto A/V 140 para su reproducción por un reproductor de audio externo o un dispositivo que tenga capacidades de audio.
- 10 El subsistema 130 de I/O del panel frontal soporta la funcionalidad del botón 150 de encendido y el botón 152 de expulsión, así como todos los LED (diodos emisores de luz) u otros indicadores expuestos en la superficie exterior de la consola 12 multimedia. Un módulo 136 de suministro de la energía del sistema proporciona la energía a los componentes de la consola 12 multimedia. Un ventilador 138 enfría la circuitería dentro de la consola 12 multimedia.
- 15 La CPU 101, la GPU 108, el controlador 110 de memoria, y diversos otros componentes dentro de la consola 100 multimedia se interconectan a través de uno o más buses, incluyendo los buses serie y paralelo, un bus de memoria, un bus periférico, y un procesador o un bus local usando cualquier de entre una variedad de arquitecturas de bus. A modo de ejemplo, dichas arquitecturas pueden incluir un bus de Interconexión de Componentes Periféricos (PCI), un bus PCI- Expres, etc.
- 20 Cuando la consola 12 multimedia se enciende, los datos de aplicación se pueden cargar desde la memoria 143 de sistema en la memoria 112 y/o las cachés 102, 104 y se ejecutan en la CPU 101. La aplicación puede presentar una interfaz de usuario gráfica que proporcione una experiencia de usuario consistente al navegar por los diferentes tipos de medios disponibles en la consola 100 multimedia. En operación, las aplicaciones y/o los otros medios contenidos dentro de la unidad 144 de memoria se pueden lanzar o reproducir desde la unidad 144 de medios para proporcionar funcionalidades adicionales a la consola 12 multimedia.
- 25 La consola 12 multimedia se puede operar como un sistema autónomo mediante simplemente conectar el sistema a una televisión u otro elemento de presentación. En este modo autónomo, la consola 12 multimedia permite a uno o más usuarios interactuar con el sistema, ver películas, o escuchar música. Sin embargo, con la integración de la conectividad de banda ancha que está disponible a través de la interfaz 124 de red o del adaptador 148 inalámbrico, la consola 12 multimedia puede ser operada además como un participante en una comunidad de red mayor.
- 30 Cuando la consola 12 multimedia es encendida, un conjunto de recursos de hardware se reservan para el uso del sistema por el sistema operativo de la consola multimedia. Estos recursos pueden incluir una reserva de memoria (por ejemplo, 16MB), un ciclo de CPU y de GPU (por ejemplo, un 5%), ancho de banda de red (por ejemplo, 8kbs), etc. Ya que estos recursos se reservan en el momento del arranque del sistema, los recursos reservados no existen desde el punto de vista de las aplicaciones.
- 35 En concreto, la reserva de memoria es suficientemente grande para contener el núcleo de lanzamiento, las aplicaciones de sistema concurrentes y los controladores. La reserva de CPU es constante de manera tal que si el uso de CPU reservado no se usa por las aplicaciones de sistema, un hilo inactivo consumirá todos los ciclos sin usar.
- 40 Con respecto a la reserva de la GPU, los mensajes ligeros generados por las aplicaciones de sistema (por ejemplo, las ventanas emergentes) se presentan mediante el uso de un interruptor de GPU para planificar el código para representar la ventana emergente en superposición. La cantidad de memoria requerida para una superposición depende del tamaño del área de las escalas de superposición con la resolución de pantalla. Donde se usa una interfaz de usuario total por la aplicación de sistema concurrente, es preferible usar una resolución independiente de la resolución de la aplicación. Se puede usar un escalador para establecer esta resolución de manera tal que se elimina la necesidad de cambiar la frecuencia y provocar una resincronización de la TV.
- 45 Después de que arranque la consola 12 multimedia y se reserven los recursos multimedia, las aplicaciones de sistema concurrentes se ejecutan para proporcionar las funcionalidades del sistema. Las funcionalidades del sistema se encapsulan en un conjunto de aplicaciones de sistema que se ejecutan dentro de los recursos de sistema reservados descritos anteriormente. El núcleo del sistema operativo identifica los hilos que son hilos de aplicaciones del sistema contra los hilos de aplicaciones de juego. Las aplicaciones de sistema se planifican para ejecutarse en la CPU 101 en momentos e intervalos predeterminados para proporcionar una visión de los recursos de sistema consistente a la aplicación. La planificación es para minimizar la ruptura de la caché para la aplicación de juego que se ejecuta en la consola.
- 50 Cuando una aplicación de sistema concurrente requiere audio, el procesamiento del audio se planifica de manera asíncrona con la aplicación de juego debido a la sensibilidad con el tiempo. Un gestor de aplicaciones de consola multimedia (descrito más adelante) controla el nivel de audio de la aplicación de juego (por ejemplo, en silencio, atenuado) cuando las aplicaciones de sistema están activas.
- 55

Los dispositivos de entrada (por ejemplo, los controladores 142(1) y 142(2)) son compartidos por las aplicaciones de juego y las aplicaciones de sistema. Los dispositivos de entrada no son recursos reservados, sino que se han de conmutar entre las aplicaciones de sistema y la aplicación de juego de manera tal que cada una tendrá el foco de atención del dispositivo. El gestor de aplicaciones controla la conmutación del flujo de entrada, sin el conocimiento de la aplicación de juego y el controlador mantiene la información de estado respecto a las conmutaciones del foco de atención. El sistema 20 de captura de imagen puede definir los dispositivos de entrada adicionales para la consola 12 (por ejemplo para su sistema de cámara).

La Figura 3B ilustra otra realización de ejemplo de un entorno informático tal como un ordenador personal. Con referencia a la Figura 3B, un sistema ejemplar para implementar la tecnología incluye un dispositivo informático de propósito general en forma de un ordenador 310. Los componentes del ordenador 310 pueden incluir, pero no se limitan a, una unidad 320 de procesamiento, una memoria 330 de sistema, y un bus 321 de sistema que acopla diversos componentes de sistema que incluyen la memoria de sistema a la unidad 320 de procesamiento. El bus 321 de sistema puede ser cualquiera de entre varios tipos de estructuras de bus que incluyen un bus de memoria o un controlador de memoria, un bus periférico, y un bus local que usa cualquiera de entre una variedad de arquitecturas de bus. A modo de ejemplo, y no de limitación, dichas arquitecturas incluyen el bus de la Arquitectura Estándar de Industria (ISA), el bus de la Arquitectura de Micro Canal (MCA), el bus de ISA mejorado (EISA), el bus local de la Asociación de Estándares Electrónicos de Vídeo (VESA), y el bus de Interconexión de Componentes Periféricos (PCI) también conocido como bus Mezzanine.

El ordenador 310 normalmente incluye una variedad de medios legibles por ordenador. Los medios legibles por ordenador pueden ser cualquier medio disponible que se pueda acceder por el ordenador 310 e incluye tanto los medios volátiles como los no volátiles, los medios extraíbles como los no extraíbles. A modo de ejemplo, y no de limitación, los medios legibles por ordenador pueden comprender los medios de almacenamiento informático y los medios de comunicación. Los medios de almacenamiento informático incluyen los medios volátiles y los no volátiles, los extraíbles y los no extraíbles implementados en cualquier método o tecnología para el almacenamiento de información tal como las instrucciones legibles por ordenador, las estructuras de datos, los módulos de programa u otros datos. Los medios de almacenamiento informáticos incluyen, pero no se limitan a, una RAM, una ROM, una EEPROM, una memoria flash u otra tecnología de memoria, un CD-ROM, discos digitales versátiles (DVD) u otro almacenamiento en disco óptico, cintas magnéticas, almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio que se pueda usar para almacenar la información deseada y que pueda ser accedido por un ordenador 310. Los medios de comunicación normalmente representan las instrucciones legibles por ordenador. Las estructuras de datos, los módulos de programa u otros datos en una señal de datos modulada tal como una onda portadora u otro mecanismo de transporte e incluye cualquier medio de entrega de información. El término "señal de datos modulada" implica una señal que tiene una o más de sus características establecida o cambiada de tal manera para codificar la información en la señal. A modo de ejemplo, y no de limitación, los medios de comunicación incluyen medios por cable tales como una red por cable o una conexión directa por cable, y los medios inalámbricos tales como los acústicos, de RF, infrarrojos, y otros medios inalámbricos. Las combinaciones de cualquiera de los anteriores se deberían incluir dentro del alcance de los medios legibles por ordenador.

La memoria 330 de sistema incluye un medio de almacenamiento informático en forma de una memoria volátil y/o no volátil tal como una memoria 331 de sólo lectura (ROM) y una memoria 332 de acceso aleatorio (RAM). Un sistema 333 de entrada/salida básico (BIOS), que contiene las rutinas básicas que ayudan a transferir la información entre los elementos dentro de un ordenador 310, tales como durante la puesta en marcha, se almacenan normalmente en la ROM 331. La RAM 332 normalmente contiene los datos y/o los módulos de programa que son accesibles de manera inmediata y/o que actualmente son operados por la unidad 320 de procesamiento. A modo de ejemplo, y no de limitación, la Figura 3B ilustra el sistema 334 operativo, los programas 335 de operación, otros módulos 336 de programa, y los datos 337 de programa.

El ordenador 310 puede incluir también otros medios de almacenamiento informático extraíbles/no extraíbles, volátiles/no volátiles. A modo de ejemplo sólo, la FIG. 3B ilustra una unidad 340 de disco duro que lee desde o escribe a unos medios magnéticos, no extraíbles, no volátiles, una unidad 351 de disco magnético que lee desde o escribe a un disco 352 magnético, extraíble, no volátil, y una unidad 355 de disco óptico que lee desde o escribe a un disco 356 óptico extraíble, no volátil tal como un CR ROM u otro medio óptico. Otros medios de almacenamiento informático extraíbles/no extraíbles, volátiles/no volátiles que se pueden usar en el entorno operativo ejemplar incluyen, pero no se limitan a, cintas magnéticas, tarjetas de memoria flash, discos versátiles digitales, cintas de vídeo digital, RAM de estado sólido, ROM de estado sólido, y similares. La unidad 341 de disco duro se conecta normalmente al bus 321 de sistema a través de una interfaz de memoria no extraíble tal como la interfaz 340, y la unidad 351 de disco magnético y la unidad 355 de disco óptico se conectan normalmente al bus de sistema 321 mediante una interfaz de memoria extraíble, tal como la interfaz 350.

Las unidades y sus medios de almacenamiento informático asociados discutidos anteriormente e ilustrados en la Figura 3B, proporcionan almacenamiento de instrucciones legibles por ordenador, estructuras de datos, módulos de programa y otros datos para el ordenador 310. En la FIG. 3B, por ejemplo, la unidad 341 de disco duro se ilustra como que almacena el sistema 344 operativo, los programas 345 de aplicación, otros módulos 346 de programa, y los datos 347 de programa. Tenga en cuenta que estos componentes pueden bien ser los mismos o diferentes del

sistema 334 operativo, los programas 335 de aplicación, otros módulos 336 de programa, y los datos 337 de programa. Al sistema 344 operativo, los programas 345 de aplicación, otros módulos 346 de programa, y los datos 347 de programa se les da diferentes números en la presente memoria para ilustrar que, como mínimo, son copias diferentes.

- 5 En una realización, se puede almacenar una copia del software y de los datos para el sistema 202 de control de visualización del elemento de presentación en los programas 345 de aplicaciones y los datos 347 de programa almacenados en la unidad 238 de disco duro o de manera remota (por ejemplo 248). Una copia se puede cargar también como un programa 226 de aplicación y datos 228 de programa en la memoria 222 de sistema cuando se ejecute.
- 10 Un usuario puede introducir comandos e información dentro del ordenador 20 a través de dispositivos de entrada tales como un teclado 362 y un dispositivo de apuntamiento 361, comúnmente referido como un ratón, una rueda de desplazamiento o un panel táctil. Otros dispositivos de entrada (no mostrados) pueden incluir un micrófono, una palanca de mando, un panel de juego, una antena parabólica, un escáner, o similares. Estos y otros dispositivos de entrada se conectan a menudo a la unidad 320 de procesamiento a través de una interfaz 360 de entrada que se
- 15 acopla al bus de sistema, pero que se puede conectar mediante otras estructuras de interfaz y bus, tales como un puerto paralelo, un puerto de juegos o un bus serie universal (USB). También se conecta un monitor 391 u otro tipo de dispositivo de presentación al bus 321 de sistema a través de una interfaz, tal como una interfaz 390 de video. Además del monitor, los ordenadores pueden incluir también otros dispositivos periféricos de salida tales como unos altavoces 397 e impresoras 396, que se pueden conectar a través de una interfaz 390 periférica de salida.
- 20 El ordenador 310 puede operar en un entorno de red que usa conexiones lógicas con uno o más ordenadores remotos, tales como el ordenador 380 remoto. El ordenador 380 remoto puede ser un ordenador personal, un servidor, un enrutador, un PC de red, un dispositivo de par u otro nodo de red común, y normalmente incluye muchos o todos los elementos descritos anteriormente en relación con el ordenador 310, aunque sólo se ha ilustrado un dispositivo 381 de almacenamiento de memoria en la Figura 3. Las conexiones lógicas representadas en la
- 25 Figura 3 incluyen una red 371 de área local (LAN) y una red 373 de área amplia (WAN), pero puede incluir también otras redes. Dichos entornos de red son normales en oficinas, redes de ordenadores amplias de empresas, intranets e Internet.

30 Cuando se usa en un entorno de red LAN, el ordenador 310 se conecta a la LAN 371 a través de una interfaz o adaptador de red 370. Cuando se usa en un entorno de red WAN, el ordenador 310 normalmente incluye un módem 372 u otro medio para establecer las comunicaciones a través de la WAN 373, tal como Internet. El módem 372, que puede ser interno o externo, se puede conectar al bus 321 de sistema a través de la interfaz 360 de entrada de usuario, u otro mecanismo apropiado. En un entorno de red, los módulos de programa representados en relación con el ordenador 310, o partes del mismo, se pueden almacenar en el dispositivo de almacenamiento de memoria remota. A modo de ejemplo, y no de limitación, la Figura 3B ilustra los programas 385 de aplicación remota como

35 residiendo en el dispositivo 381 de memoria. Se apreciará que las conexiones de red mostradas son ejemplares y se pueden usar otros medios de establecimiento de un enlace de comunicaciones entre los ordenadores.

40 La Figura 4 ilustra una realización de ejemplo de un entorno informático en red en el que pueden operar una o más realizaciones para proporcionar retroalimentación del seguimiento de usuario a un usuario. Como se muestra en la Fig. 4, las múltiples consolas 400A-400X o dispositivos de procesamiento, tales como los ilustrados en las Figuras 3A y 3B se pueden acoplar a una red 402 y pueden comunicarse las unas con las otras y con un servicio 404 de juego en red que tiene uno o más servidor o servidores 406 a través de la red 402. El servidor o los servidores 406 pueden incluir un componente de comunicación capaz de recibir información desde y transmitir información hasta las consolas 400A-X y puede proporcionar una colección de servicios que las aplicaciones que se ejecutan en las consolas 400A-X pueden invocar y utilizar.

45 La consolas 400A-X pueden invocar el servicio 408 de inicio de sesión de usuario, que se usa para autenticar e identificar a un usuario en las consolas 400A-X. Durante el inicio de sesión, el servicio 408 de inicio de sesión obtiene una etiqueta de jugador (un identificador único asociado con el usuario) y una contraseña del usuario así como un identificador de la consola que identifica de manera única la consola que el usuario está usando y la ruta de red hasta la consola. La etiqueta de jugador y la contraseña se autentican comparándolas con una base 416 de

50 datos de perfiles de usuario, que se puede ubicar en el mismo servidor que el servicio 408 de inicio de sesión o se puede distribuir en diferentes servidores o en una colección de diferentes servidores. Una vez autenticado, el servicio 408 de inicio de sesión de usuario almacena el identificador de la consola y la ruta de red en la base 416 de datos de perfiles de usuario para que los mensajes y la información se pueda enviar a la consola.

55 En una realización, las consola 400A-X pueden incluir un servicio 410 de juego, un servicio 412 compartido, unos datos 428 compartidos de usuario y una base 418 de datos de sustitución. El servicio de juego puede permitir a los usuarios jugar en línea juegos interactivos, crear y compartir entornos de juego para el juego conjunto entre consolas, y proporcionar otros servicios tales como un mercado en línea, un seguimiento de logros centralizado a través de diversos juegos y funciones de experiencia compartida. Un servicio 412 compartido permite a los usuarios compartir los elementos de juego con otros usuarios. Por ejemplo, un usuario en una consola 400x puede crear

60 elementos para usar en juegos y compartirlos o venderlos a otros usuarios. Además, un usuario puede registrar

elementos de la experiencia de juego, tales como un vídeo de una carrera o diversas escenas en un juego, y compartirlas con otros usuarios. La información proporcionada por los usuarios para compartir o vender se puede almacenar en los datos 428 compartidos de usuario.

5 La base 416 de datos de perfiles de usuarios global puede incluir información sobre todos los usuarios en las consolas 400A-X así como la información de las cuentas de los usuarios y un identificador de consola que identifique de manera única la consola concreta que cada usuario esté usando. La base 416 de datos de perfil de usuario puede incluir también la información de las preferencias de usuario asociada con todos los usuarios en las consolas 400A-X. La base 416 de datos de perfiles de usuarios global puede incluir también la información sobre los usuarios tal como los registros del juego una lista de amigos asociada con los usuarios.

10 El software 414 de retroalimentación del seguimiento se puede proporcionar en el servicio 404 de juego. El software de retroalimentación del seguimiento puede responder a los problemas de seguimiento con gestos y movimiento en los elementos del juego subidos al servidor y almacenados en los datos 428 compartidos de usuario.

15 Cualquier número de dispositivos de procesamiento en red se puede proporcionar de acuerdo con un sistema de juego como se proporciona en la Fig. 4. Como tal, la tecnología presentada en la presente memoria puede operar uno o más servidores 406 en conjunción con un servicio 404 de juego o se puede proporcionar en dispositivos de procesamiento individuales en un entorno en red, tal como los dispositivos 400A – 400x.

20 La Figura 5 representa un modelo de un usuario 500 humano que se puede crear usando el sistema 20 de captura y el entorno 12 informático. La correspondencia 500 esquelética de ejemplo de un usuario se ha generado desde los datos 205 de seguimiento del movimiento capturados mediante el sistema 20 de captura de datos audiovisuales. Este modelo se puede usar mediante uno o más aspectos del sistema 10 basado en gestos para determinar los gestos y similares. El modelo puede estar comprendido por articulaciones y huesos. El seguimiento de dichas articulaciones y huesos puede permitir al sistema basado en gestos determinar que gestos está haciendo un usuario. Estos gestos se pueden usar para controlar el sistema basado en gestos. En esta realización, se identifican una variedad de articulaciones y huesos: cada muñeca 502a, 502b, cada antebrazo 504a, 504b, cada codo 506a, 506b, cada bíceps 508a, 508b, cada hombro 510a, 510b, cada cadera 512a, 512b, cada muslo 514a, 514b, cada rodilla 516a, 516b, cada pierna 518a, 518b, cada pie 520a, 520b, una cabeza 522, un torso 524, una parte superior 526 y una parte inferior 528 de la espina, y una cintura 530. Cuantos más puntos se sigan, más características se pueden identificar, tal como las características individuales de la cara, tal como la nariz y los ojos. Sin embargo, cuantos más cambios en los datos se sigan, más se puede ralentizar también la presentación del avatar.

30 La Figura 6 representa una serie de gestos tales como una onda o una mano 602 levantada, haciendo una X con los brazos 604, o un choque de manos 606. Aunque no se limita de ninguna manera por los pocos gestos que se han representado, estos gestos, junto con cualesquiera otros pueden ser comandos para el sistema 10 basado en gestos. En una realización, los gestos pueden ser universales, significando que no deberían estar limitados a aplicaciones de software o hardware concretas. En otra realización los juegos u otros programas operados sobre el entorno 12 informático pueden tener gestos específicos del programa. Por ejemplo, un gesto universal para transferir el control del juego a otro jugador puede ser un apretón de manos; sin embargo, un juego tal como un juego de lucha libre puede tener un gesto específico de programa que realice la transferencia del control si los usuarios realizan un gesto 524 de chocar los cinco.

40 Las realizaciones del método de las Figuras 7 y 8 respectivamente se discuten en el contexto de la realización de la arquitectura funcional implementada por ordenador de la Figura 2 sólo con propósitos ilustrativos y no para ser limitantes del mismo.

45 La Figura 7 es un diagrama de flujo de una realización 700 del método para proporcionar retroalimentación a un usuario sobre la capacidad de una aplicación de seguir el movimiento del usuario. El sistema de captura detecta en 702 a un usuario en un área de captura. El módulo 204 de movimientos recibe los datos 205l del seguimiento del movimiento y de manera opcional, los datos 217 de audio y sigue en 704 al menos una característica corporal del usuario basado en los datos generados por el sistema de captura. En respuesta a los criterios de seguimiento de usuario que no se satisfacen para una aplicación en ejecución, el software 213 de retroalimentación determina en 706 una respuesta de retroalimentación y provoca la salida en 708 de la retroalimentación al usuario. Como se describe anteriormente, la salida de la retroalimentación puede incluir la presentación de indicadores visuales, proporcionando un indicador de audio o proporcionando una acción dentro del contexto de la actividad de la aplicación.

55 La Figura 8 es un diagrama de flujo de una realización 800 del método para ayudar a un usuario en la selección de un área de captura. El software 213 de retroalimentación del seguimiento presenta las instrucciones que instruyen en 802 a un usuario para dirigir el campo de visión de la cámara a un área de captura de prueba, y el sistema 20 de captura, captura datos 804 de imagen del área de captura de prueba. En una realización, el software 213 de retroalimentación del seguimiento puede presentar los datos de imagen del área de captura de prueba en el elemento de presentación para que el usuario pueda verlos por sí mismo. El software 213 de retroalimentación del seguimiento presenta en 806 las instrucciones que dirigen al usuario a hacer al menos un gesto usando al menos una parte del cuerpo usada para controlar la actividad para la aplicación en ejecución. El módulo 204 de

movimientos sigue en 808 el movimiento de al menos una parte del cuerpo que hace el gesto. El software 213 determina en 811 si el criterio de seguimiento está siendo satisfecho. En respuesta a los criterios de seguimiento del usuario que se satisfacen, el software 213 de retroalimentación del seguimiento determina en 812 una puntuación de la calidad del seguimiento para el área de captura de prueba.

5 El módulo 204 de movimientos puede proporcionar las puntuaciones o los pesos o algún valor representativo de la calidad de ciertos factores de visibilidad para el área de captura de prueba para el software 213 de retroalimentación del seguimiento para que este los compare con los criterios de seguimiento basados en los factores de visibilidad. En alguna realización, un algoritmo de pesajes se puede aplicar entonces a estos factores para determinar una puntuación. Algunos ejemplos de los factores incluyen la ubicación de la parte del cuerpo del usuario para el gesto
10 en el campo de visión del sistema de captura, la iluminación en el área de captura, y las obstrucciones de la parte del cuerpo.

Las diferentes áreas de captura se califican o se puntúan, y la mejor se recomienda en 818. En otra alternativa, el sistema 20 de captura de imagen se gira a lo largo de un rango de ángulos capturando diferentes vistas en una habitación, y sugiriendo como área de captura la que proporcione la mejor capacidad de seguimiento. El área de
15 captura con la mejor capacidad de seguimiento se puede presentar en el elemento 16 de presentación a través del módulo 207 de procesamiento de la presentación para identificarlo por el usuario.

El módulo 204 de movimientos puede no ser capaz de hacer coincidir un gesto con el movimiento del usuario que lo hace, y puede enviar un mensaje al software 213 de retroalimentación del seguimiento indicándolo. El módulo 204 de movimientos puede indicar también un factor de visibilidad que contribuya a que los criterios de seguimiento del
20 usuario no se estén satisfaciendo. Por ejemplo, el valor de calidad representativo para el factor de visibilidad de la iluminación puede indicar que es pobre.

En respuesta a los criterios de seguimiento del usuario que no se satisfacen, el software 213 de retroalimentación del seguimiento determina como uno o más de los factores de visibilidad se pueden mejorar para el seguimiento, y emite la retroalimentación en 820 al usuario identificando al menos un cambio en el área de captura para mejorar un
25 factor de visibilidad. La retroalimentación se puede emitir a través de elementos de presentación visuales o como mensajes de audio.

El software 213 de retroalimentación del seguimiento determina en 812 una puntuación de calidad del seguimiento para el área de captura de prueba. El módulo 204 de movimientos puede proporcionar puntuaciones o pesos o algún valor representativo de la calidad de ciertos factores de visibilidad para el área de captura de prueba al software 213
30 de retroalimentación del seguimiento que puede aplicar un algoritmo de pesajes a estos factores para determinar una puntuación.

El software 213 de retroalimentación del seguimiento determina en 814 si existe otra área de captura de prueba a ser probada. Por ejemplo, solicita la entrada del usuario a través del elemento de presentación para ver si hay otra área de captura de prueba. Si hay otra área de captura de prueba, los pasos se repiten para este área de prueba. Si
35 no hay otra área de captura de prueba, el software 213 de retroalimentación del seguimiento presenta una recomendación de área de captura de prueba con la mejor puntuación de visibilidad. Por ejemplo, el área de captura de prueba con la mejor puntuación se presenta en el elemento de presentación. Si hay otra área de captura de prueba, el software 213 presenta en 802 las instrucciones para dirigir el campo de visión de la cámara a la siguiente área de captura y repetir los pasos anteriores.

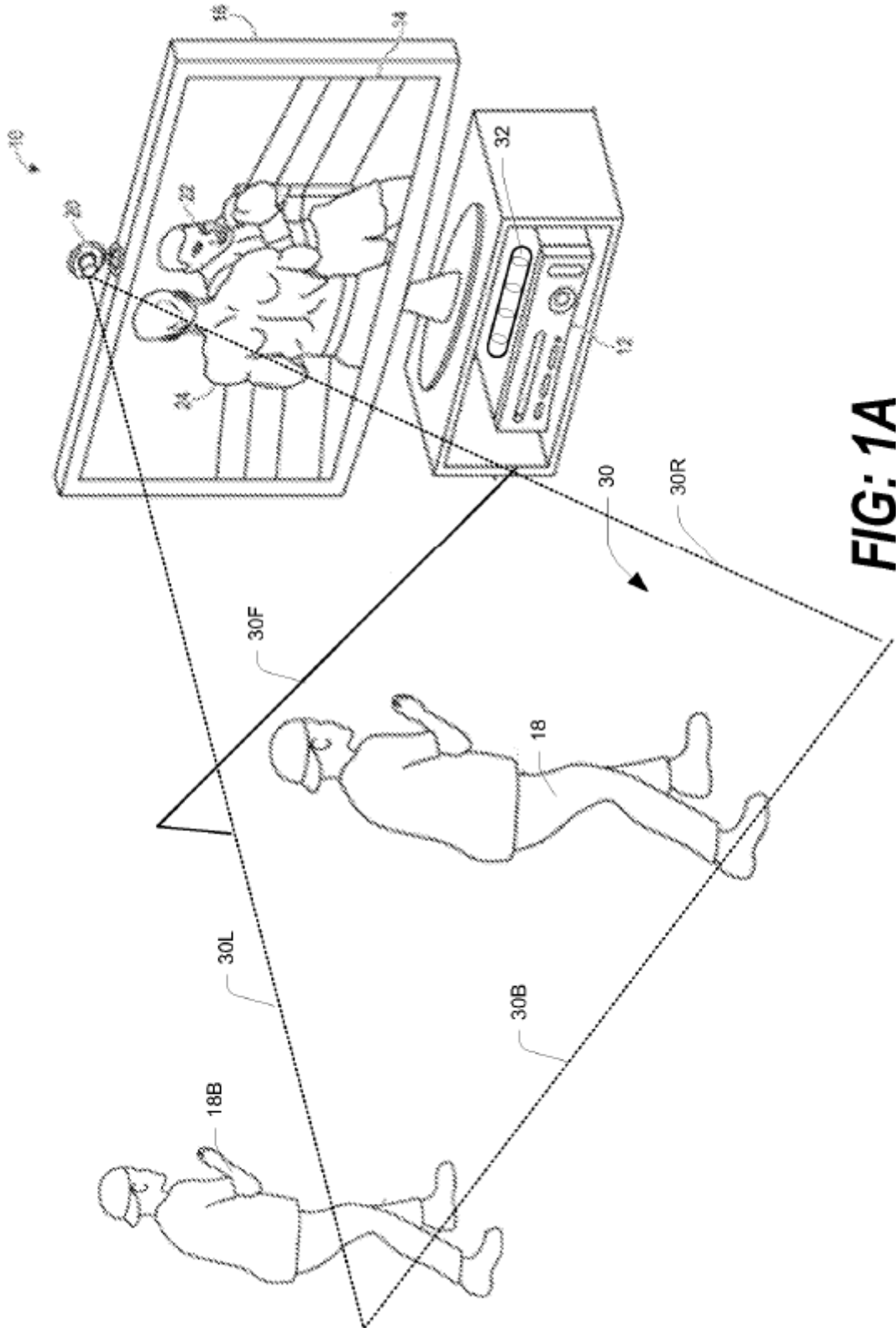
40 La descripción detallada anterior se ha presentado con propósitos de ilustración y descripción. No está destinada a ser exhaustiva o a limitar la tecnología descrita a la forma descrita precisa. Muchas modificaciones y variaciones son posibles a la luz de las enseñanzas anteriores. Las realizaciones descritas se eligieron para explicar mejor los principios de la tecnología y su aplicación práctica para de este modo permitir a otros expertos en la técnica utilizar mejor la tecnología en las diversas realizaciones y con las diversas modificaciones que sean adecuadas al uso
45 concreto contemplado. Además, se debería tener en cuenta que el lenguaje usado en la especificación se ha seleccionado principalmente con el propósito de legibilidad e instrucción.

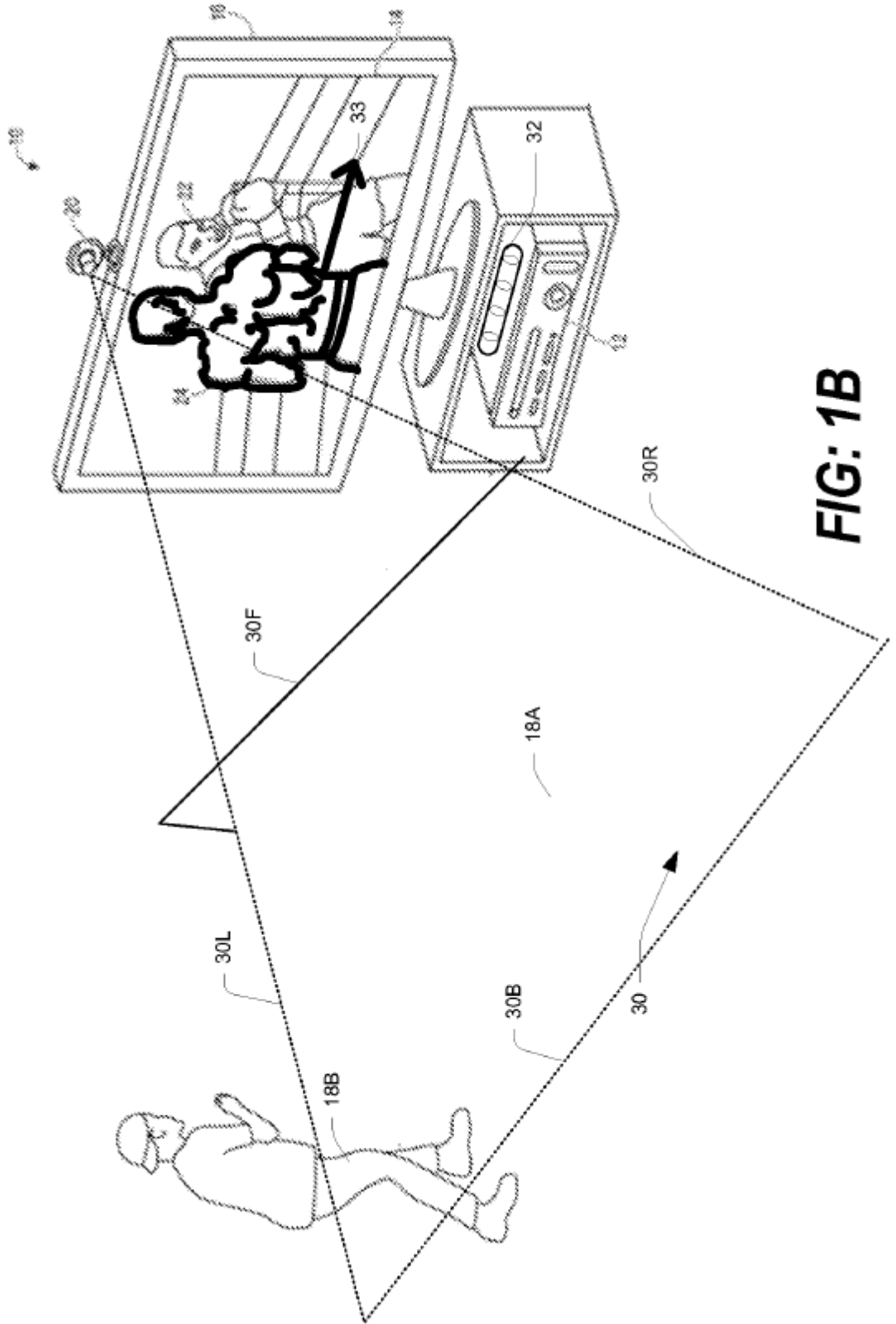
Aunque el tema se ha descrito en un lenguaje específico para las características estructurales y/o los actos metodológicos, se ha de entender que el tema definido en las reivindicaciones adjuntas no está necesariamente limitado a las características o actos específicos descritos anteriormente. Más bien, las características y actos
50 específicos descritos anteriormente se describen como ejemplo de las formas de implementar las reivindicaciones.

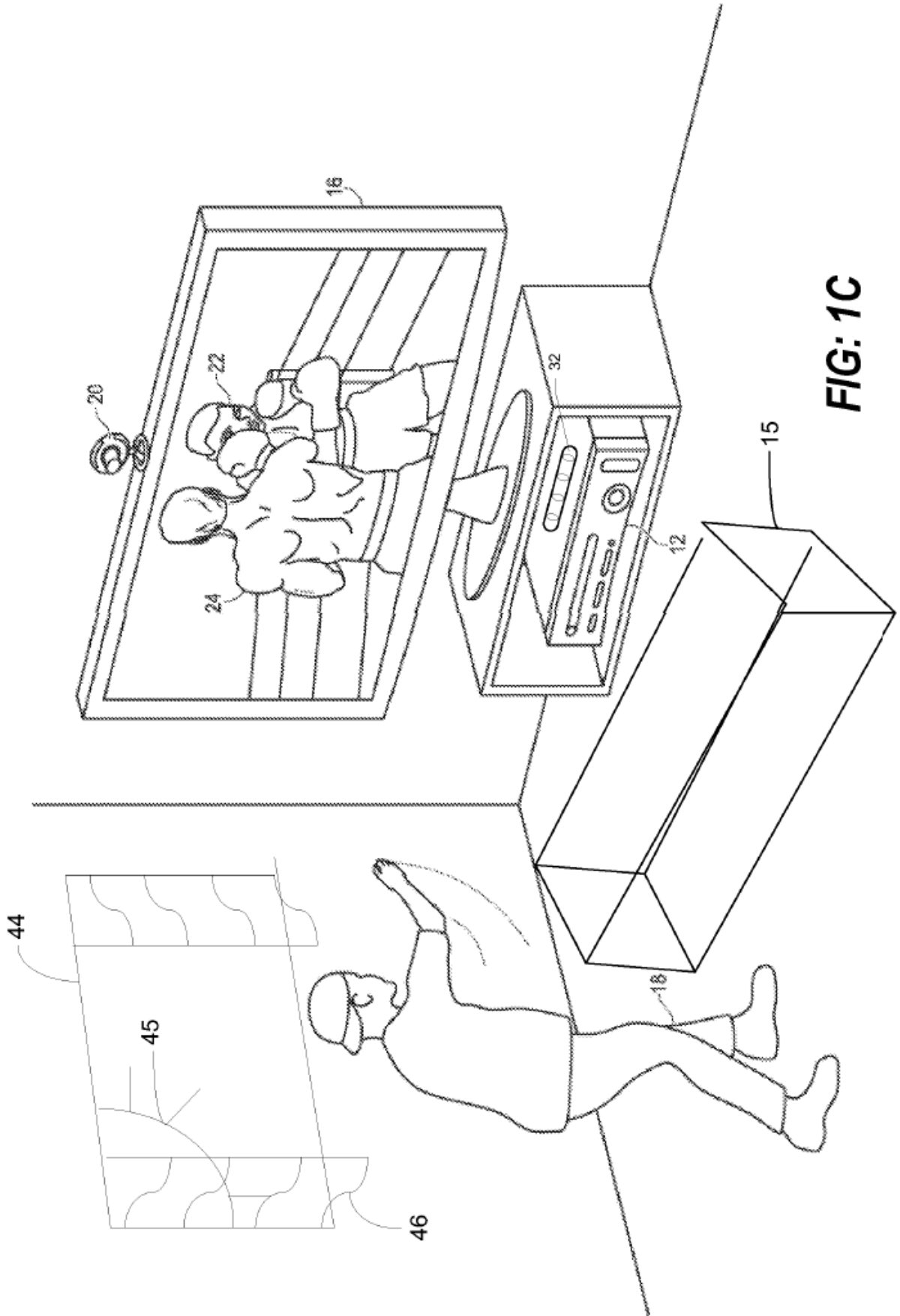
REIVINDICACIONES

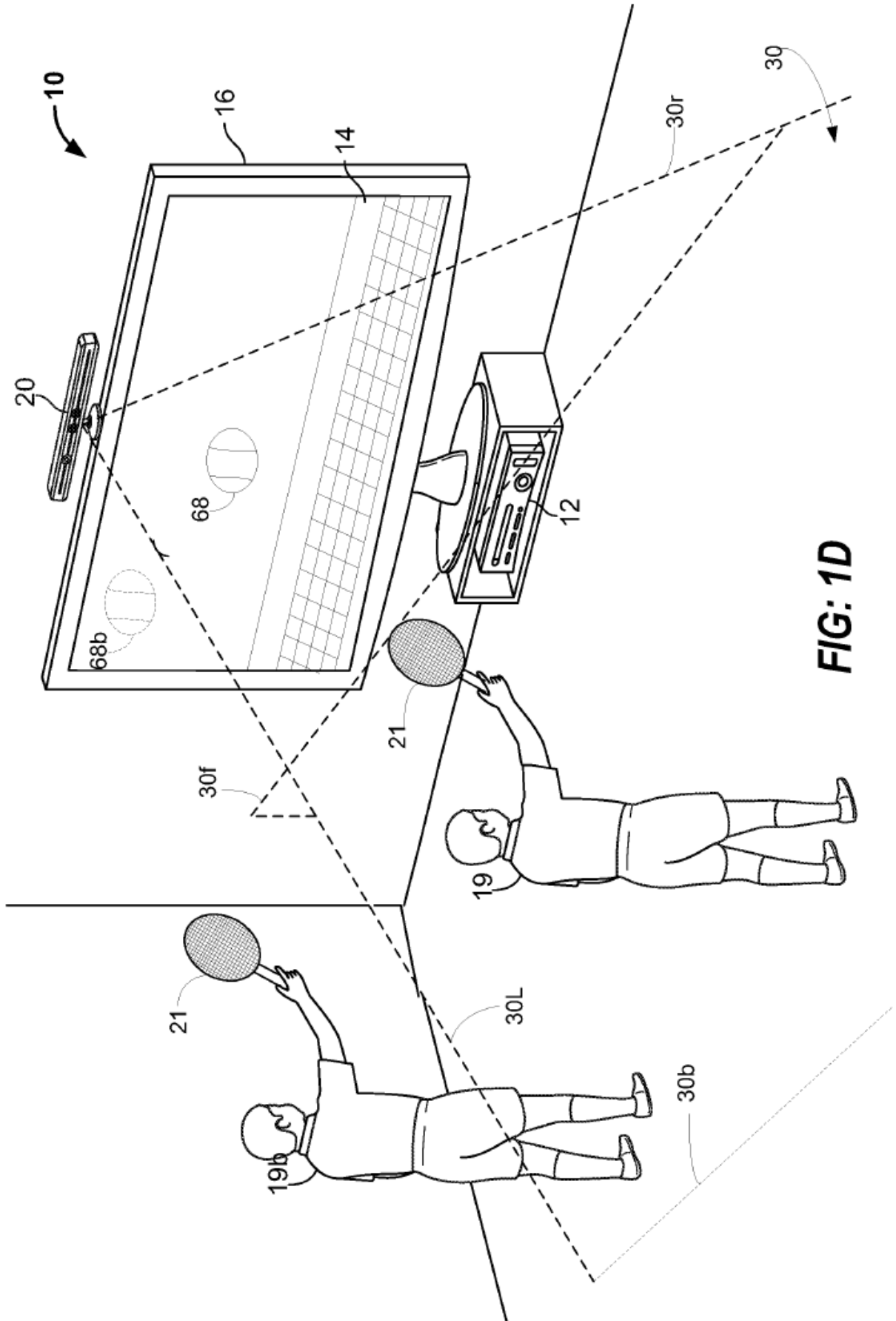
1. Un método (700), realizado por un sistema implementado por ordenador que ejecuta una aplicación de juego que sigue las acciones del usuario para controlar un objeto presentado de un juego, para proporcionar retroalimentación a un usuario sobre la capacidad del sistema para seguir las acciones del usuario que comprende:
- 5 la detección (702) del usuario en un área de captura con un sistema (20) de captura de imagen;
- el seguimiento (704) de al menos una parte del cuerpo del usuario en base a los datos generados por el sistema (20) de captura;
- 10 en respuesta a la determinación de que la parte del cuerpo del usuario que puede controlar el objeto presentado está al menos de manera parcial fuera del campo de visión del sistema de captura de imagen, la determinación (706) de una respuesta contextual de retroalimentación a ser proporcionada dentro del contexto de la aplicación de juego, siendo dicha respuesta de retroalimentación contextual una acción contextual en el juego que cambia al menos una acción de un objeto en el elemento de presentación para proporcionar una motivación al usuario para mover la parte del cuerpo hacia el campo de visión, en donde la respuesta contextual de retroalimentación se determina en base a un perfil del usuario; y
- 15 la emisión (708) de la respuesta contextual de retroalimentación al usuario en el elemento de presentación.
2. El método (700) de la reivindicación 1
- en donde dicha acción contextual en el juego es para enviar el objeto en el elemento de presentación en una dirección en la que el usuario debería moverse para poner la parte del cuerpo de vuelta en el campo de visión.
3. El método (700) de la reivindicación 1
- 20 en donde dicha acción contextual en el juego es un cambio de dirección o de perspectiva de la visión del usuario de la acción del juego.
4. El método (700) de cualquiera de las reivindicaciones precedentes, en donde el perfil del usuario incluye la edad del usuario.
5. El método (700) de cualquiera de las reivindicaciones precedentes, comprendiendo además la emisión de un
- 25 indicador de audio al usuario dirigido en una determinada dirección para motivar al usuario a correr lejos de esa dirección en respuesta a la determinación de que la parte del cuerpo del usuario que puede controlar el objeto presentado está al menos de manera parcial fuera del campo de visión del sistema de captura de imagen.
6. El método (700) de cualquiera de las reivindicaciones precedentes, comprendiendo además la determinación de la respuesta contextual de retroalimentación en respuesta a criterios de seguimiento del usuario adicionales que no se satisfacen, estando dichos criterios de seguimiento de usuario basados en el factor de visibilidad del grupo
- 30 consistente de lo siguiente:
- la iluminación;
- el contraste; y
- una obstrucción de la parte del cuerpo del usuario.
- 35 7. Un medio (112, 143, 330, 341, 352, 356) legible por ordenador que comprende instrucciones ejecutables por ordenador almacenadas en el mismo que, cuando son ejecutadas por una unidad de procesamiento (101, 108, 320), provocan que la unidad (101, 108, 320) de procesamiento realice el método (700) de cualquiera de las reivindicaciones precedentes.
8. Un sistema para proporcionar retroalimentación a un usuario sobre la capacidad del sistema para seguir las
- 40 acciones del usuario que comprende:
- un sistema (20) de captura de imagen para capturar los datos de imagen de un usuario;
- una memoria (112, 143, 330, 341, 352, 356) que almacena instrucciones; y
- uno o más procesadores (101, 108, 320) que se acoplan de manera comunicativa al sistema (20) de captura de imagen y a la memoria (112, 143, 330, 341, 352, 356);
- 45 un elemento (16) de presentación acoplado de manera comunicativa con el uno o más procesadores para presentar la retroalimentación al usuario;

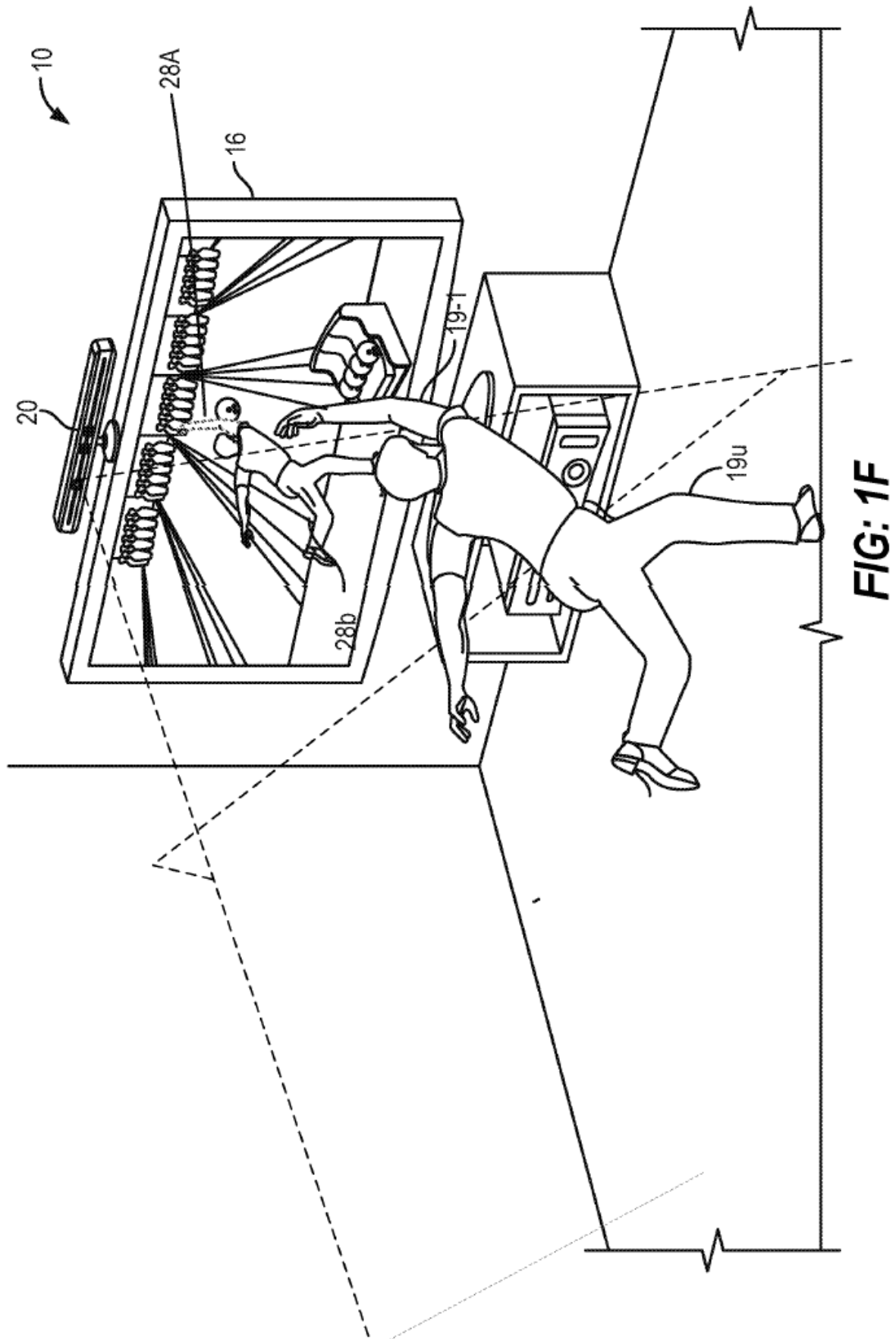
el uno o más procesadores (101, 108, 320) que se configuran para, cuando se ejecutan las instrucciones almacenadas en la memoria (112, 143, 330, 341, 352, 356), realizar el método (700) según cualquiera de las reivindicaciones 1 a 6.

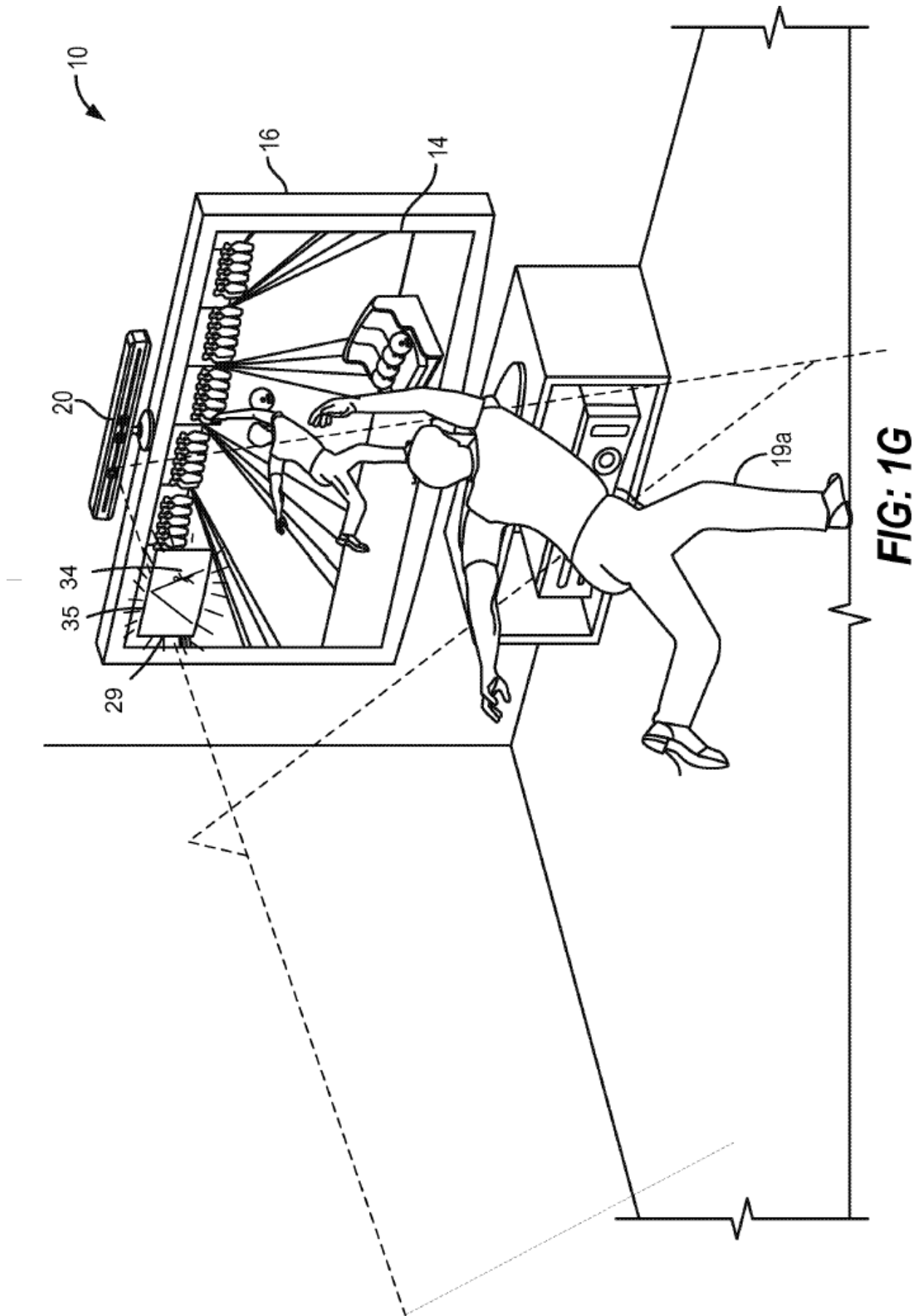












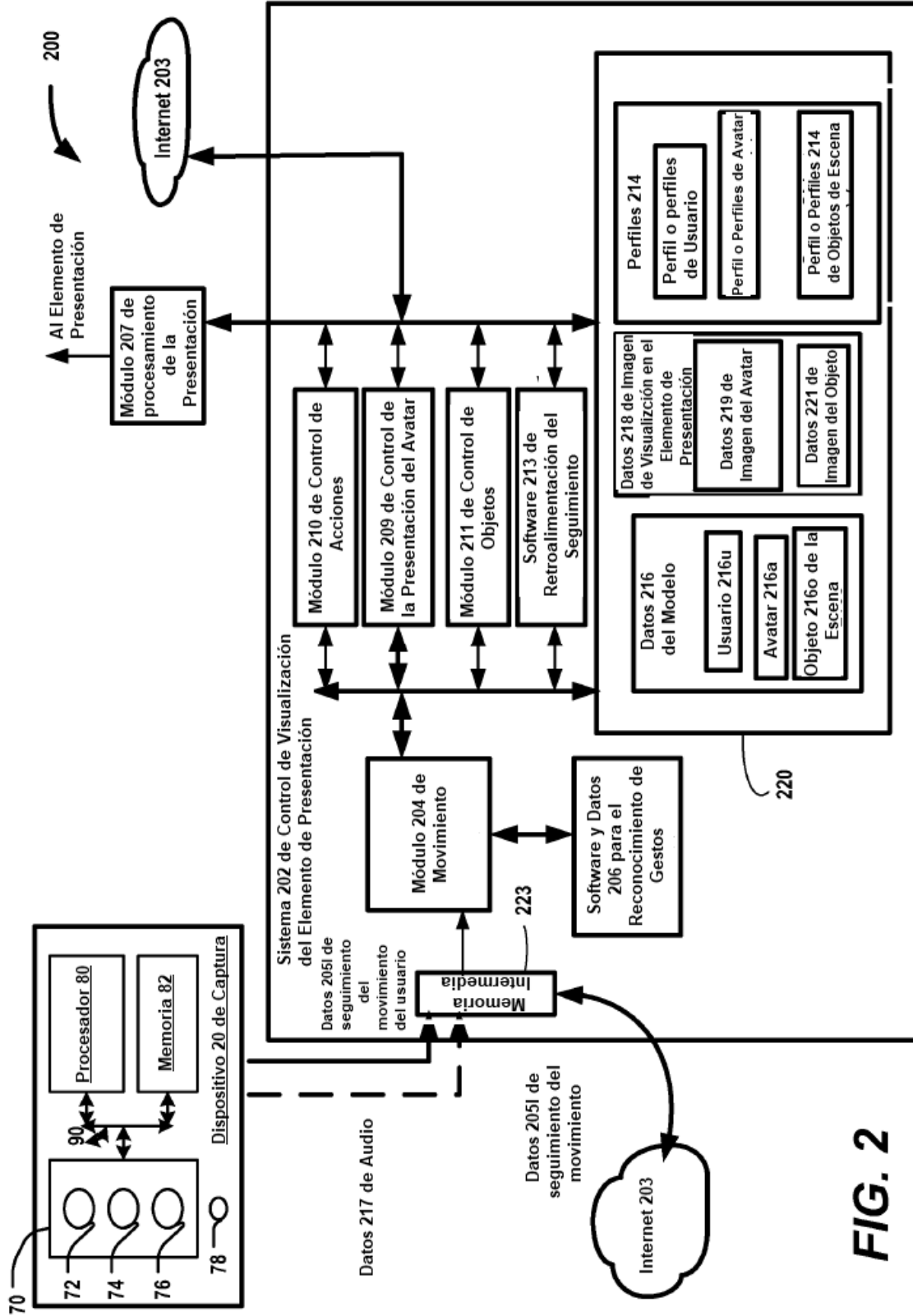


FIG. 2

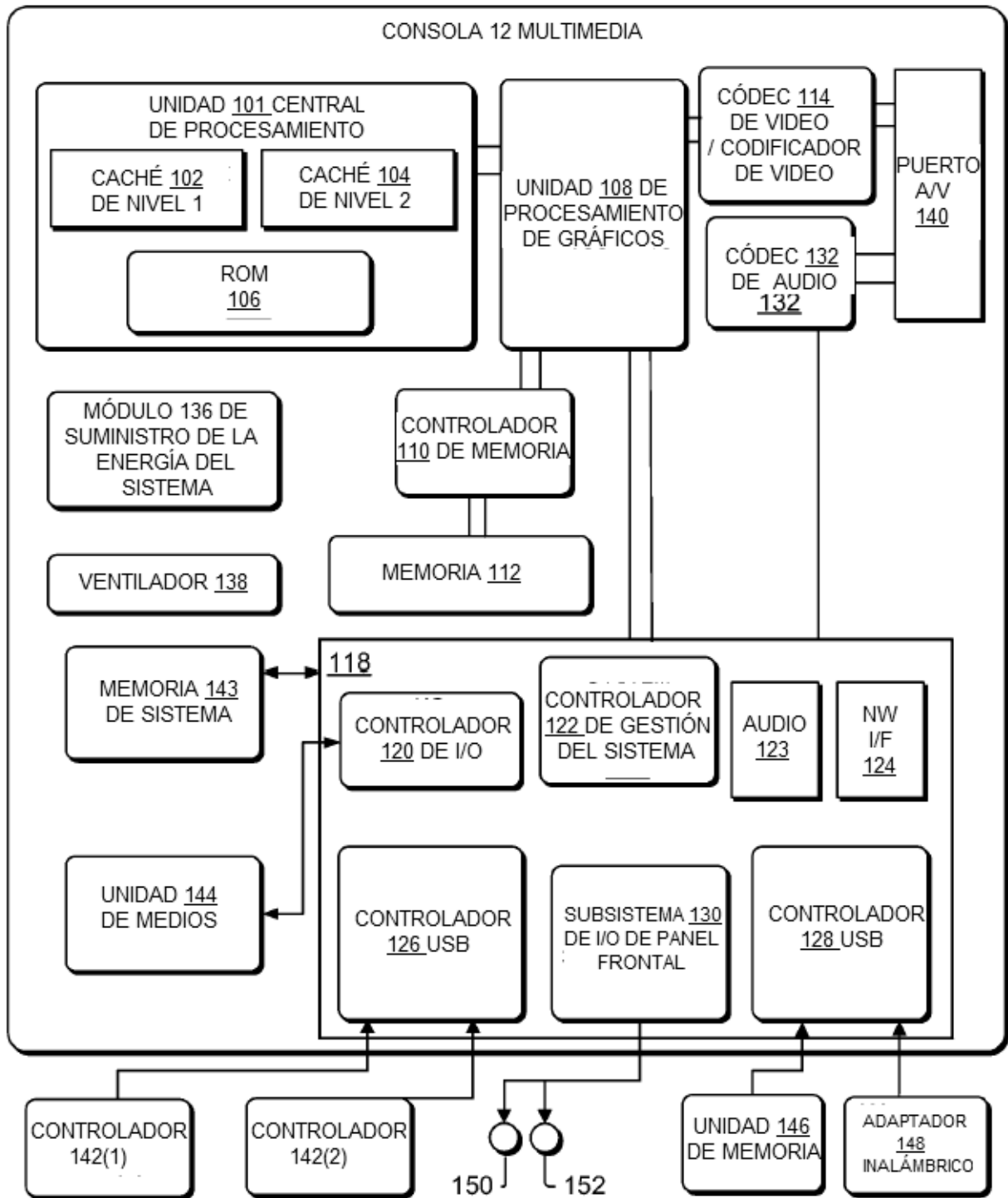


FIG. 3A

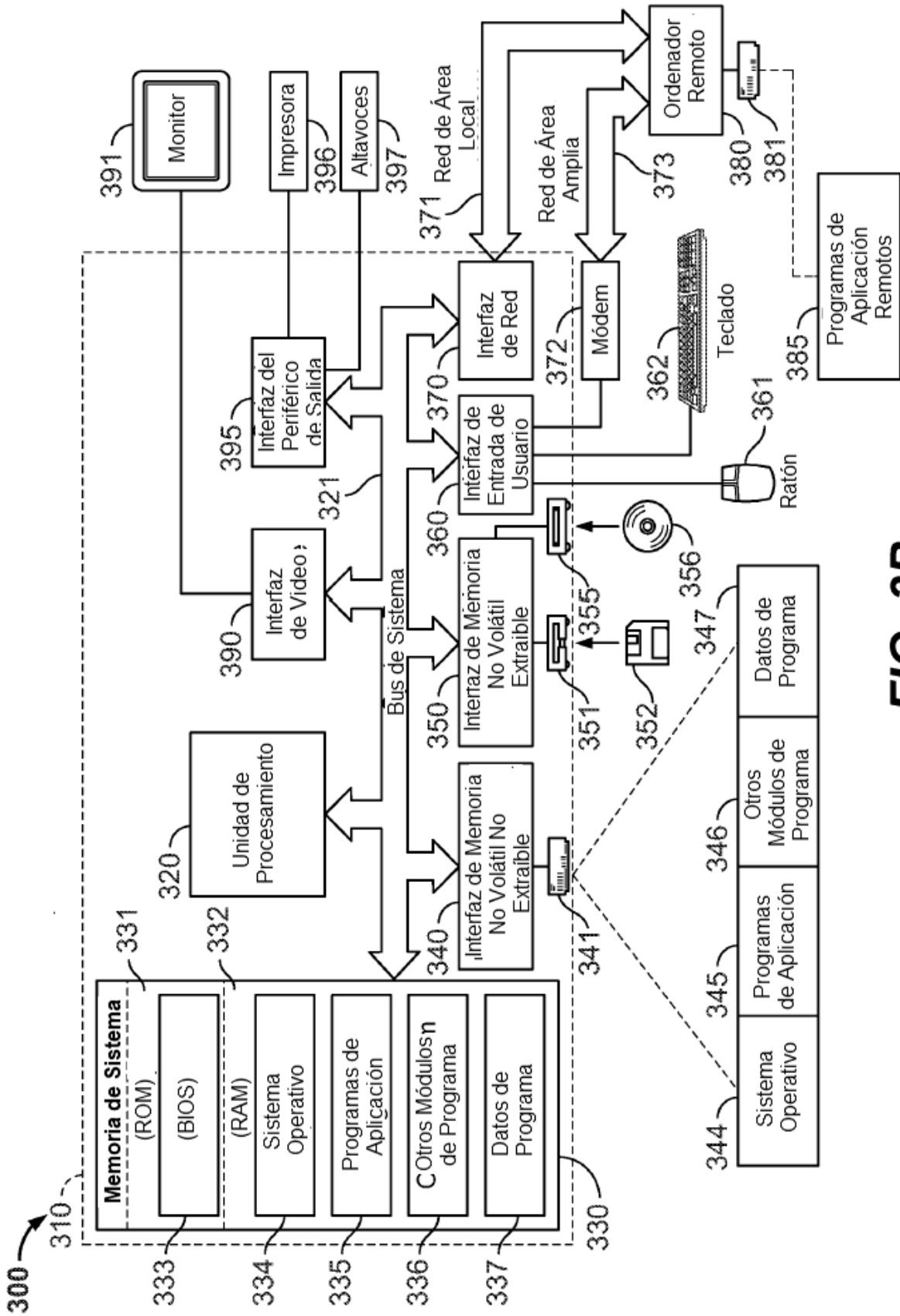


FIG. 3B

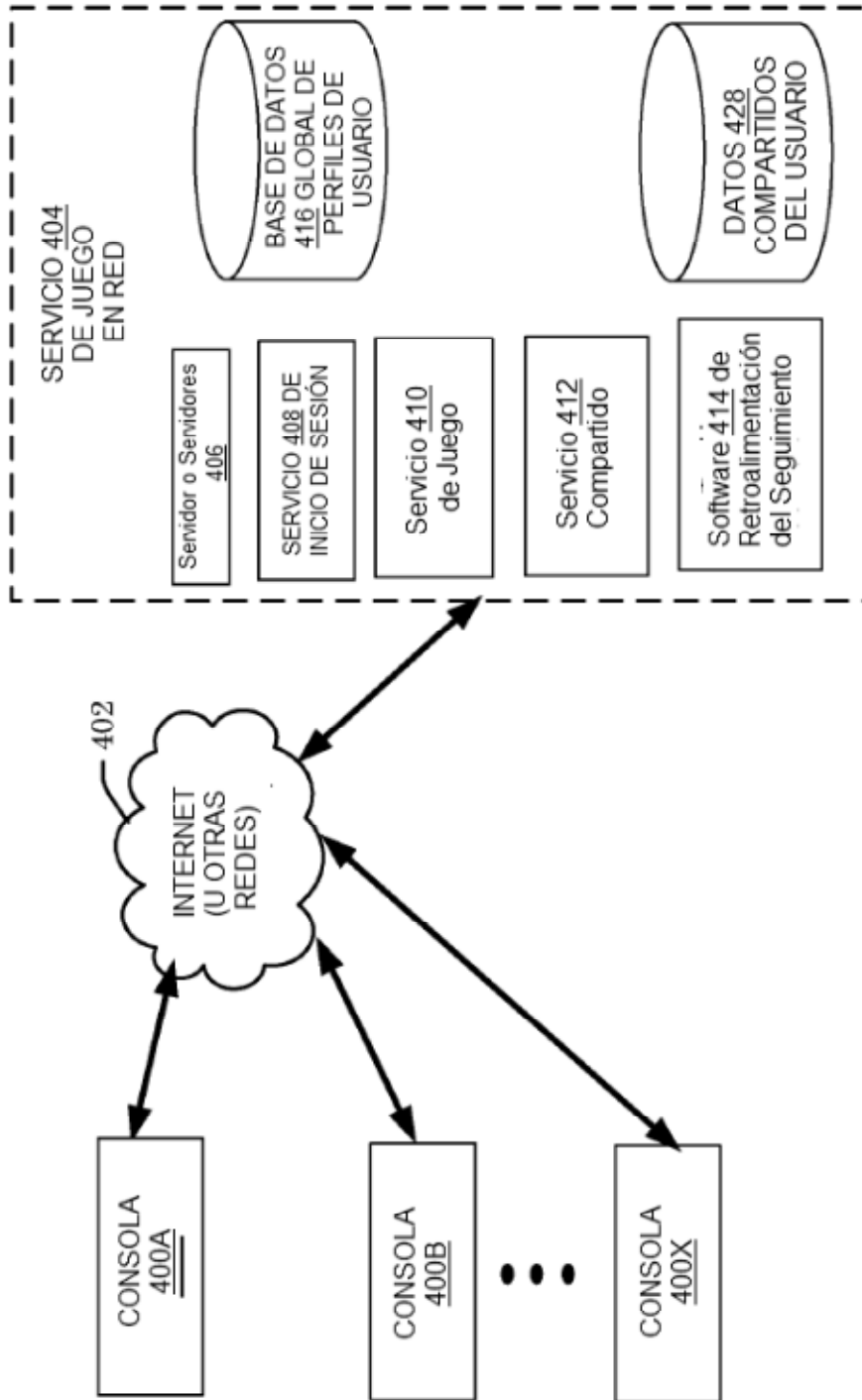


FIG. 4

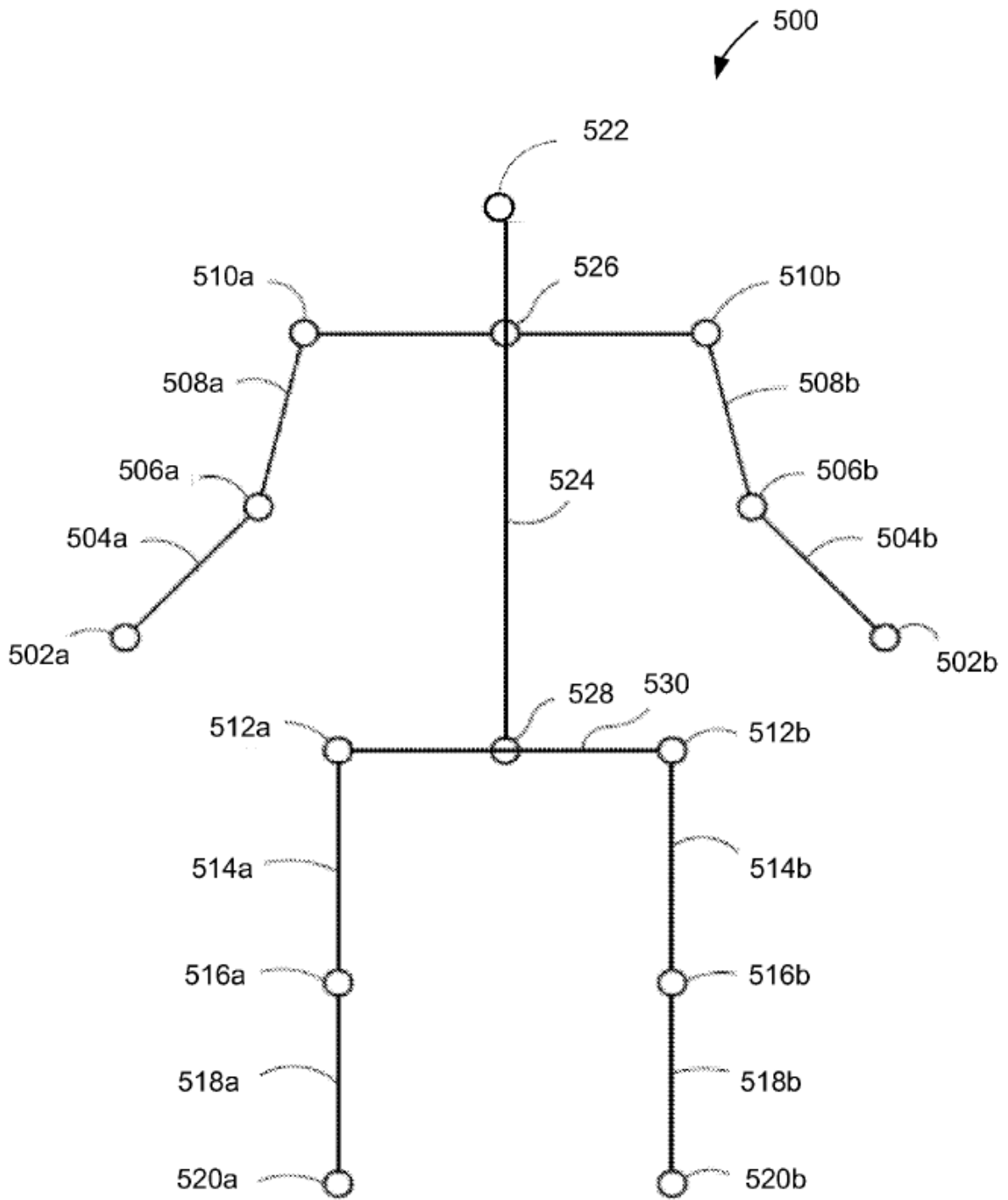


FIG. 5

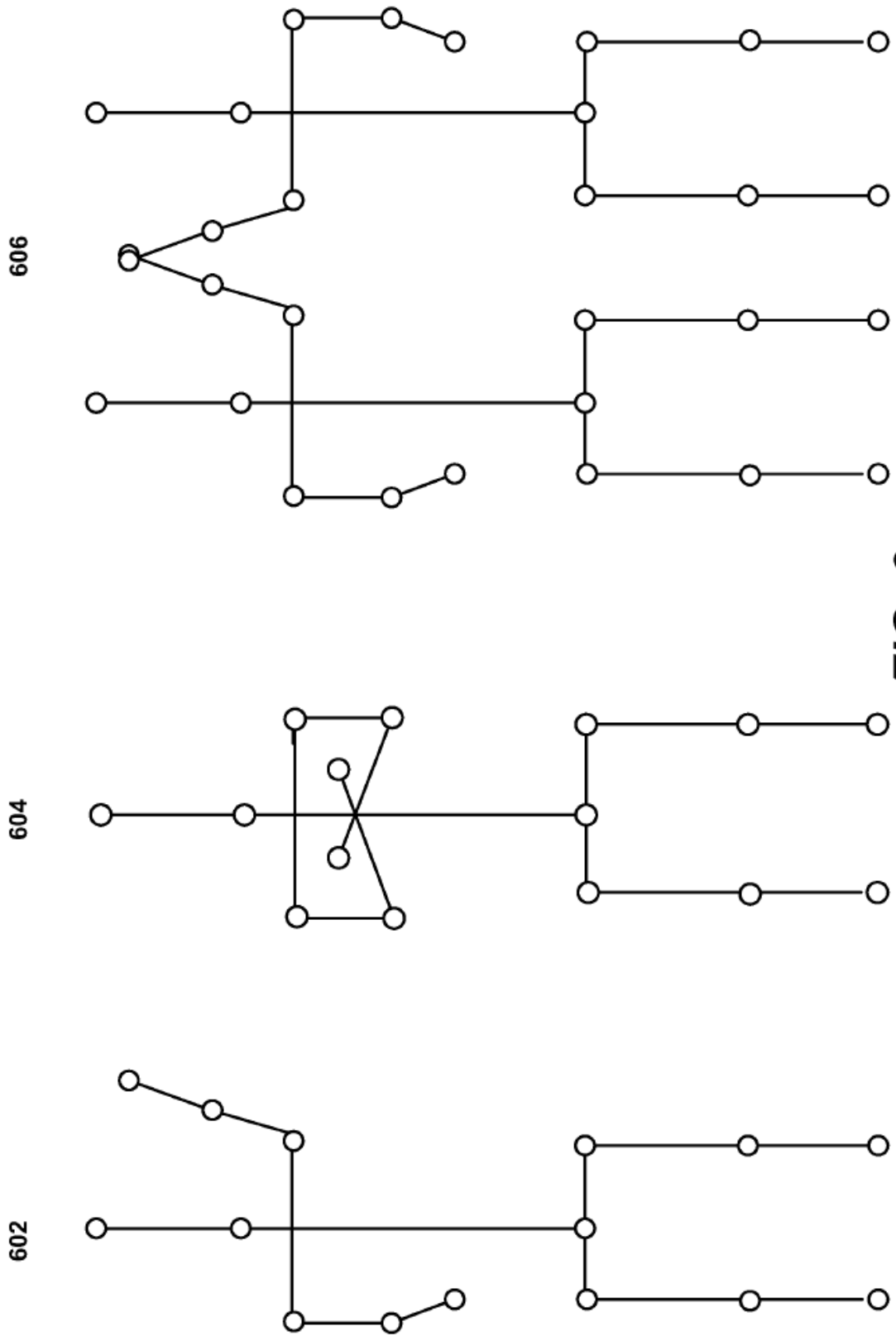


FIG. 6

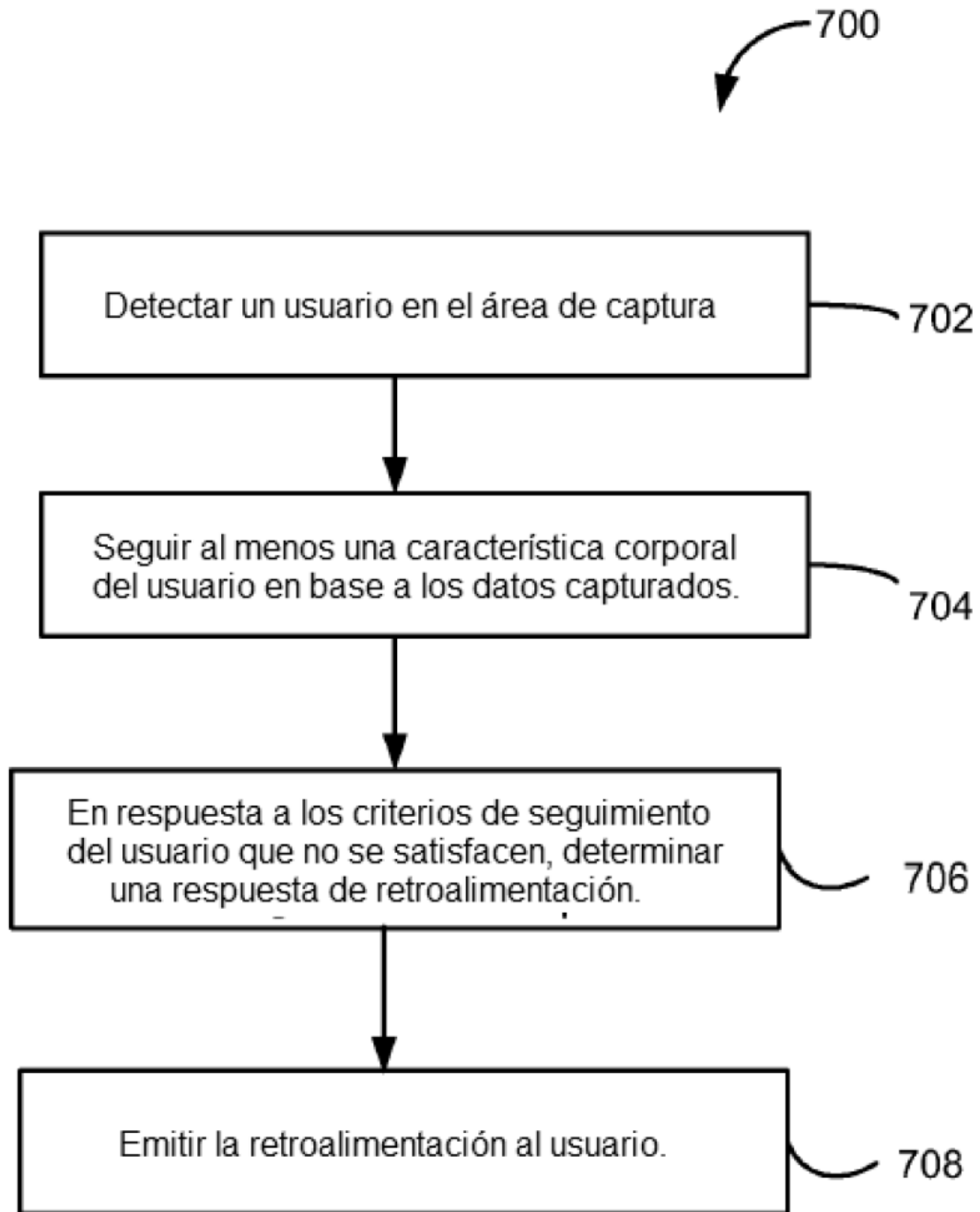


FIG. 7

