

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 995 983**

51 Int. Cl.:

G06T 11/00 (2006.01)

G06F 3/01 (2006.01)

G06T 3/00 (2014.01)

H04N 7/18 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **16.04.2014** **PCT/US2014/034333**

87 Fecha y número de publicación internacional: **12.02.2015** **WO15020703**

96 Fecha de presentación y número de la solicitud europea: **16.04.2014** **E 14833667 (0)**

97 Fecha y número de publicación de la concesión europea: **10.07.2024** **EP 3028177**

54 Título: **Dispositivos, sistemas y métodos de virtualización de un espejo**

30 Prioridad:

04.08.2013 US 201361862025 P
17.10.2013 US 201361892368 P

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
11.02.2025

73 Titular/es:

EYESMATCH LTD (100.00%)
19 Waterfront Drive, P.O. Box 3540
Road Town, Tortola VG1110, VG

72 Inventor/es:

VILCOVSKY, NISSI y
SABAN, OFER

74 Agente/Representante:

BERTRÁN VALLS, Silvia

ES 2 995 983 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

Dispositivos, sistemas y métodos de virtualización de un espejo

- 5 La invención se refiere a un sistema informático programado para realizar un cambio de color en un objeto dentro de un vídeo en tiempo real.

10 El espejo convencional (es decir, superficie reflectante) es la herramienta común y más fiable para que un individuo explore su propio aspecto real, en tiempo real. Se han propuesto unas pocas alternativas referentes a la combinación de una cámara y una pantalla para sustituir al espejo convencional. Sin embargo, estas técnicas no son convincentes y aún no se aceptan como imagen fiable del individuo como si se estuviera mirando a sí mismo en un espejo convencional. Esto se debe principalmente a que la imagen generada por una cámara es muy diferente de una imagen generada por un espejo.

- 15 En la publicación de solicitud de patente estadounidense n.º 2013/0229482 A1, se propone una nueva tecnología (método informatizado) para abordar las dificultades de convertir y transformar una imagen instantánea o vídeo en 2D/3D creado mediante una o más cámaras, con o sin otros sensores, en una experiencia de espejo o conferencia de vídeo.

20 El documento US 8 174 539 B1 divulga métodos, sistemas y aparatos, que incluyen productos de programa informático, para generar una imagen de diseño para fabricar un objeto diseñado. En un aspecto, un método incluye recibir una entrada que especifica un diseño visual de un objeto en una primera imagen, generar una segunda imagen basándose en la entrada recibida, mapear la segunda imagen a una superficie que tiene uno o más contornos basándose en una o más líneas de flujo para generar una tercera imagen, presentar una composición de la tercera imagen con la primera imagen, y proporcionar la segunda imagen a un procedimiento automatizado para fabricar el objeto diseñado.

30 En el artículo "Human Friendly Interface Design for Virtual Fitting Room Applications on Android Based Mobile Devices", Gracia Martin Cecilia *ET AL*, Journal of Signal and Information Processing, 2012, 3, 481-490, se describe un flujo de diseño de procesamiento de imágenes para aplicaciones de probadores virtuales (VFR), que se dirige tanto a ordenadores personales como a dispositivos móviles. La interfaz fácil de manejar propuesta se implementa mediante un algoritmo de tres etapas: detección y dimensionamiento en el cuerpo del usuario, detección de puntos de referencia basándose en la detección facial y marcadores de realidad aumentada, y superposición de prendas de ropa sobre la imagen del usuario.

35 El problema técnico objetivo es proporcionar un sistema informático programado para realizar un cambio de color en un objeto dentro de un vídeo en tiempo real que garantice que un objeto mostrado en un vídeo y que desea manipular un usuario se identifique de manera rápida y precisa, en el que el objeto no se conoce con anterioridad y que proporcione realidad aumentada en tiempo real a través de una alimentación de vídeo.

40 Este problema técnico objetivo se logra mediante el sistema informático según la reivindicación 1.

45 El siguiente resumen de la divulgación se incluye para proporcionar una comprensión básica de algunos aspectos y características de la invención. Este resumen no es una visión general extensa de la invención y, como tal, no pretende identificar particularmente elementos clave o críticos de la invención o delimitar el alcance de la invención. Su único propósito es presentar algunos conceptos de la invención de una forma simplificada como prelude de la descripción más detallada que se presenta a continuación.

50 Esta divulgación se refiere a una técnica informatizada que toma un vídeo, una imagen instantánea, o un grupo de imágenes instantáneas, antes o después de la transformación descrita en el documento US 2013/0229482 A1, y continúa con el método informatizado para abordar funcionalidades adicionales, tales como la interfaz de usuario, la realidad aumentada, el cambio de color, el cambio de textura, la manipulación de la forma de un objeto, cuerpo o fondo, y más.

55 Para las capacidades de realidad aumentada, en primer lugar se creará una máscara o un modelo en 2D y/o 3D para uno o más artículos o elementos (por ejemplo, camisa, vestimenta, pantalones, cabeza, zapato, gafas, cuerpo entero, y similares), luego la manipulación de la máscara o el modelo puede permitir las características de realidad aumentada como color, sustitución y manipulación de textura, o medida geométrica, o sustitución completa de objeto. La técnica divulgada se diferencia de otras técnicas que usan solo una cámara en 3D (cámara de IR o dual).

60 En las siguientes técnicas divulgadas, puede construirse el modelo a partir de una única cámara regular y pueden mejorarse los resultados con múltiples cámaras e información adicional de cámara de IR, 3D u otros sensores e información con respecto al objeto que trata de manipularse.

65 Los dibujos adjuntos, que se incorporan en esta memoria descriptiva, ilustran una o más realizaciones a modo de ejemplo divulgadas en el presente documento y, junto con la descripción detallada, sirven para explicar los principios e implementaciones a modo de ejemplo.

En los dibujos, en los que números de referencia iguales se refieren a una referencia igual en la memoria descriptiva:

- 5 La figura 1 ilustra una realización para realidad aumentada para simular un aspecto en el espejo en un elemento de visualización de pantalla plana;
- la figura 2 ilustra una realización de un módulo de realidad aumentada;
- 10 la figura 3 ilustra una realización de un módulo de realidad aumentada que puede sustituir una parte del cuerpo, color, orientación y textura de un artículo u objeto en primer plano o fondo de la imagen;
- la figura 4 ilustra una realización para métodos informáticos para crear un modelo para el intercambio de color y textura;
- 15 la figura 5 ilustra un ejemplo de un método para extraer una medida corporal de una imagen, múltiples imágenes o un vídeo corto;
- la figura 6 ilustra un ejemplo de una aplicación de sistema de seguridad para aprendizaje y autenticación de usuario de múltiples niveles;
- 20 la figura 7 ilustra un ejemplo de una aplicación de sistema de seguridad para manipulación de cámaras en paralelo o en banda;
- 25 las figuras 8a-8c ilustran esquemas para compartir una experiencia de espejo en una conferencia de vídeo/voz;
- la figura 9 ilustra un ejemplo para una entrada en 2D al generador de modelos;
- la figura 10 es un ejemplo de un modelo en 2D de la camisa superior a partir de la imagen de la figura 9;
- 30 la figura 11 es un esquema que ilustra la clasificación de tipos de cuerpo masculinos;
- la figura 12 es un esquema que ilustra la clasificación de tipos de cuerpo femeninos;
- 35 la figura 13 representa un ejemplo del resultado de procedimientos adicionales que eliminan la mano del cuerpo y apuntan a diferentes intervalos de dónde mirar para medidas de cuerpo particulares;
- la figura 14 representa un ejemplo de transformación en un flujo de cámara para crear un efecto de espejo en el entorno de fondo;
- 40 la figura 15 ilustra un ejemplo de espejo virtual con reconocimiento y autenticación de usuario y con interfaz de usuario;
- la figura 16 ilustra un ejemplo de un espejo virtual con modo de pantalla dividida y miniaturas; y
- 45 la figura 17 ilustra una realización de flujo de uso de usuario.

Los siguientes ejemplos ilustran algunas realizaciones y aspectos de la invención. Para facilitar la comprensión, la descripción incluye la descripción de diversas características con títulos específicos.

Visión general

- Las realizaciones de la presente invención utilizan una cámara y un elemento de visualización de pantalla plana para proporcionar al usuario la experiencia de mirarse al espejo. Es decir, el flujo de imágenes de la cámara se manipula y transforma de tal manera que cuando se visualiza en la pantalla, parece imitar un reflejo en un espejo, en lugar de un flujo de vídeo grabado. Puesto que la imagen de "espejo virtual" se produce de manera digital, puede almacenarse y manipularse para proporcionar características de mejora adicionales. Por ejemplo, la imagen puede manipularse para cambiar los colores, para permitir la identificación y autenticación de usuario de múltiples niveles, y para permitir realizar medidas corporales. El sistema incluye una cámara calibrada, de tal manera que el sistema es capaz de proporcionar características de realidad aumentada, por ejemplo, color, textura, manipulación de la forma de elementos del cuerpo o el fondo, por ejemplo, en un espejo virtual o aplicaciones de conferencia de vídeo. También es capaz de realizar medidas corporales y monitorización corporal para aplicaciones comerciales, de seguridad, y de asistencia sanitaria.
- 65 Con respecto al hardware, diversas realizaciones pueden incluir una única cámara o múltiples cámaras de todos los tipos de sensores, incluyendo, pero sin limitarse a, CCD, CMOS, IR CMOS, y similares, y protocolos de formato que

incluyen HDMI, Firewire, USB, IP, HDbaseT inalámbrico, y similares, y cualquier resolución disponible incluyendo HD, SIF regular, 4k, 8k, y similares. La pantalla puede incluir una única pantalla o múltiples pantallas o pantalla virtual, y pantallas regulares en todos los tamaños, formas y proporciones. El elemento de visualización puede incluir pantallas similares a espejos, teléfonos inteligentes o tabletas, proyectores, hologramas, pantallas en 3D, dispositivo Google glass, dispositivo de visualización montado en la cabeza (HMD) de proyector, gafas (con capa oscura conmutable), y similares.

En las realizaciones divulgadas, la cámara puede estar ubicada en cualquier lugar. La mejor práctica es proporcionar la cámara por encima de la pantalla orientada hacia el usuario. Ubicaciones adicionales pueden incluir la parte inferior de la pantalla, los lados de la pantalla o detrás de la pantalla si la pantalla es una pantalla bidireccional. En el caso del 3D, pueden proporcionarse dos cámaras, por ejemplo, en la parte superior y una al lado de otra, lo cual puede proporcionar resultados superiores y facilitar el método informatizado. Además, puede manipularse 3D a partir de una única cámara una vez que el usuario esté moviéndose y/o girando.

Las realizaciones pueden incluir un método informatizado para cubrir un cuerpo completo o un cuerpo parcial, y ajustes a escala diferente según la selección del usuario (por ejemplo, corrección de la cabeza y/u ojos, dirección de visión/punto de vista, y similares).

Cada una de las realizaciones divulgadas puede proporcionarse sobre una imagen instantánea impresa en 2D o 3D, una imagen instantánea digital en 2D o 3D, sobre video analógico/digital que se grabó por una cámara analógica/digital o por un captador de tramas HDMI (o equivalente), IP, USB, Firewire (enlace cableado o inalámbrico), y similares. Cada una de las realizaciones divulgadas puede proporcionarse a través de flujo continuo de video digital que puede suministrarse a un dispositivo informático a través de cualquier método de comunicación adecuado tal como USB, HDMI, IP cat-5/fibra, WLAN, Firewire, HDbaseT, cualquier combinación de los anteriores en un único cable o en múltiples cables. Cada una de las realizaciones divulgadas puede proporcionarse cuando el dispositivo informático reside en la cámara, en la pantalla, en la nube, o en un ordenador que incluye una estación de trabajo, un servidor, y similares.

Cada una de las realizaciones divulgadas puede proporcionarse cuando el dispositivo informático está distribuido entre los elementos del sistema, por ejemplo, el método informatizado puede residir parcialmente en la cámara, parcialmente en el elemento de adquisición de video, parcialmente en la unidad de procesamiento gráfico (GPU) de la pantalla, parcialmente en la nube, parcialmente en el dispositivo de teléfono inteligente del usuario o cualquier combinación adecuada de los mismos. Cada una de las realizaciones divulgadas puede proporcionarse cuando el dispositivo informático reside en un teléfono inteligente, tableta, ordenador portátil, pantalla, cámara, HMD, dispositivo Google glass, y similares.

Cada una de las realizaciones divulgadas puede proporcionarse cuando el video se procesa previamente para extraer un modelo en 3D de un artículo seleccionado, en el que la selección previa puede realizarse manualmente por el usuario o automáticamente aplicando reglas para la selección. La extracción de modelo puede realizarse de manera local mediante DSP/CPU junto a la cámara, en la nube o en la aplicación del usuario en el dispositivo del usuario. Resulta más eficiente extraer el modelo en una GPU (unidad de procesamiento gráfico) para un único objeto o una imagen instantánea; en este caso, una CPU regular puede ser suficiente.

Cada una de las realizaciones divulgadas puede proporcionarse cuando está realizándose la representación gráfica de manera local en la estación de realidad aumentada, en la nube, o en el dispositivo móvil del usuario. La representación gráfica puede realizarse en la GPU para una única imagen o un video muy corto; en estos casos, una CPU puede proporcionar un rendimiento suficiente.

Cada una de las realizaciones divulgadas puede proporcionarse cuando se realizan técnicas de representación gráfica potenciadas para mejorar la calidad de video, por ejemplo, interpolando el objeto y la imagen a una resolución superior y decimando después de combinar para suavizar los bordes y eliminar un efecto de solapamiento. La presente invención puede eliminar los bordes con parpadeo, por ejemplo, suavizando el modelo trama a trama, eliminando los cambios permitidos por trama en los bordes, suavizando los bordes por cada trama, aplicando filtros de cálculo de promedio en una única trama o múltiples tramas, y similares. Pueden lograrse mejoras adicionales aumentando la tasa de tramas y aplicando una técnica de suavizado adicional en la trama añadida para suavizar el efecto que se produciría en la tasa de tramas original.

Cada una de las realizaciones divulgadas puede proporcionarse independientemente de si la pantalla reside cerca de la cámara. Por ejemplo, la pantalla puede estar ubicada en una sala de control. Puede accederse a un flujo principal de cámara o información grabada a través de control remoto, a través de otra dirección directamente con infraestructura física, o a través de la nube.

En una realización, el control sobre la funcionalidad de la estación puede ser a través de un gesto, movimiento ocular, movimiento de la palma de la mano, movimiento de un dedo o control cerebral. La presente invención puede controlarse con accesorios adicionales tales como un puntero de haz láser, control de voz, rastreo inalámbrico, o de ultrasonidos.

Segmentación

Una segmentación precisa de objetos y/o elementos de una imagen o vídeo es una de las mayores dificultades en el presente campo. Las técnicas conocidas para resolver este problema implican un claro compromiso entre precisión y velocidad. Antes de la presente invención, no hay ninguna solución satisfactoria al problema. El uso de una cámara de infrarrojos (IR) con medida de profundidad tridimensional (3D) es un enfoque en la industria; sin embargo, la calidad del enfoque es reducida. Por ejemplo, bordes, distorsión de imagen, y/o mala alineación de baja resolución entre las cámaras IR y regular evitan una segmentación satisfactoria y convincente. Las técnicas de separación cromática y de detección de borde son muy demandantes desde un punto de vista de recursos informáticos y no proporcionan un rendimiento sistemático. Algunas técnicas para extraer un usuario u objeto suponen, por ejemplo, un fondo verde que es fácil de distinguir, aunque no son aplicables en aplicaciones en las que no se dispone del fondo verde.

Una vez obtenido un modelo del artículo o artículos que van a manipularse, un segundo problema es crear un flujo de flujo de representación gráfica suave que actualice el modelo y lo represente gráficamente de manera suave de vuelta en el vídeo original o en otro vídeo de realidad aumentada con la nueva información. Las siguientes realizaciones divulgadas logran un flujo de representación gráfica suave relativamente rápido con alta calidad, de modo que la manipulación puede realizarse en tiempo real, en semi-tiempo real, o de manera relativamente rápida, dependiendo del requisito de la aplicación particular. Además, las realizaciones están adaptadas para manipular los artículos (color, textura, forma) en tiempo real, en una alimentación de flujo principal de cámara, o fuera de línea en vídeos grabados del usuario.

Basándose en el modelo, la máscara, y/o el canal alfa por cada objeto de un vídeo, se aprende el movimiento del usuario y pueden instalarse diferentes objetos para ajustarse o no ajustarse con la forma y conformación de objeto de una manera convincente. Las realizaciones divulgadas pueden distinguirse de técnicas anteriores, que intentan realizar todo de manera virtual. Las realizaciones incluyen la manipulación con una máscara para cubrir otros cambios tales como tela diferente, el cambio de la cara de un usuario, un cambio del tipo de cuerpo del usuario, eliminación de un usuario a partir de una imagen y similares.

Otra característica divulgada es el reconocimiento de tela, que puede identificar los artículos y la ropa que lleva puesta el usuario. Por ejemplo, puede incluir un código oculto o información en el material textil.

Otras características y realizaciones pueden incluir: un escáner de microondas, en el que el sensor de microondas está integrado en la pantalla; características de medida corporal incluyendo, por ejemplo, detección de líneas corporales en 2D y 3D o generación de modelo/máscara en 2D/3D a partir de un vídeo; conclusión analítica a partir de las medidas, por ejemplo, tipo de cuerpo, origen, sexo, edad, tallas sugeridas de ropa, dieta o tratamiento sugerido, y similares; seguimiento de líneas corporales a lo largo del tiempo para diagnóstico, identificación de cambios, identificación de usuario, y similares.

En algunas realizaciones, el flujo de cámara principal puede manipularse en tiempo real para cambiar el aspecto del usuario, o para hacer que desaparezca completamente del vídeo o la imagen. En esta situación, el DSP de cámara o dispositivo con capacidad de DSP o GPU independiente necesita obtener el flujo de cámara y procesar el flujo de cámara mediante un centro de control remoto y empaquetarlo de nuevo como flujo principal de cámara. Un cálculo paralelo al flujo principal de cámara (analizador) también es una opción en una realización.

A continuación se proporcionan varios ejemplos de aplicaciones adicionales que pueden aprovechar la tecnología de la presente invención. Específicamente, puede usarse transformación de vídeo de base para crear vestimenta virtual fiable.

Puede usarse transformación de vídeo de base para crear una conferencia de vídeo fiable cuando el/los usuario(s) en el otro lado de la línea puede(n) ver al participante como si el participante le(s) estuviera mirando directamente a él/ellos, en vez de tener una imagen desde el punto de vista de la cámara. Alternativamente, puede usarse una transformación para crear una imagen como si un usuario estuviera mirando a los ojos de otro usuario con el fin de proporcionar una experiencia en directo convincente.

Puede proporcionarse transformación de vídeo de base para realidad aumentada para múltiples propósitos incluyendo, aunque sin limitarse a: manipulación y redimensionamiento corporal, medidas corporales, diagnóstico médico/monitorización/entrenamiento y/o seguimiento, juegos, vestimenta, citas, reuniones, seguridad, simuladores de cuerpo completo, entrenamiento de baile, peluquería, maquillaje, entrenamiento general, viajes, y similares.

Las realizaciones pueden aplicarse para conferencia de vídeo o juegos en los que el usuario querrá cambiar su aspecto (por ejemplo, tez de la piel, vestimenta, cara, y similares). Según la presente invención, el modelo puede permitir un cambio suave para dar un aspecto deseado o incluso crear una imagen mediante la cual el usuario desaparece del vídeo.

Las realizaciones pueden usarse en una aplicación de seguridad, con reconocimiento facial preciso sobre un dispositivo de Eyes-Match calibrado (véase la solicitud estadounidense n.º US 2013/0229482 A1) cuando la cámara está ubicada con perspectiva de proyección hacia los ojos del usuario, con reconocimiento de código QR dinámico (respuesta rápida) (por ejemplo, puede generarse autorización y gestión de acceso de usuario mediante código QR de aplicación móvil, o tarjeta de identificación, para la aplicación móvil a partir de la nube y puede verificarse mediante autenticación de usuario en la aplicación), con reconocimiento de cuerpo completo, escaneo, y aprendizaje, con reconocimiento de artículos (por ejemplo, ropa, uniforme, tarjeta de identificación, y similares), con un analizador inalámbrico (por ejemplo, WiFi, celular, ultrasonidos, y similares), con API para una base de datos de seguridad, con sistemas de localización, con información pública, y similares. Las realizaciones pueden usarse con una acción de prevención de amenazas basándose en reconocimiento de una actividad especificada (por ejemplo, bloquear puerta, encender/apagar una luz, liberar niebla, máquina de disparo, por ejemplo, descarga eléctrica, balas, papel, y similares). Debido a la geometría de la imagen calibrada, la ubicación del objetivo en la imagen puede medirse de manera muy precisa y puede usarse para realizar un ataque preciso en un objetivo de manera eficiente sin poner a otras personas en peligro.

Puede proporcionarse grabación y comparación de aspecto en tiempo real y fuera de línea de modo que un usuario puede realizar un seguimiento de cambios de aspecto a lo largo del tiempo y a lo largo de los años. Estos procedimientos pueden proporcionarse con múltiples propósitos incluyendo registros médicos, que puede ser a lo largo de un periodo de tiempo dado, entre múltiples exploraciones, usando sensores múltiples/diferentes, por ejemplo, para detectar la temperatura, tensión arterial, y similares.

En algunas realizaciones, la cámara/múltiples cámaras calibradas con/sin información complementaria a partir de otros sensores pueden usarse para diagnóstico de pacientes y monitorización de pacientes. Por ejemplo, aunque sin limitación, la presente invención puede usarse para monitorizar y alertar por problemas en la piel tales como melanoma de cáncer de piel, cambios de mama que pueden sugerir pruebas adicionales para detectar cáncer de mama, medida de frecuencia cardíaca y tensión arterial, temperatura del paciente, halo del usuario, monitorización de grabación de piel y uñas para detectar cambios, problemas oculares, giro, huesos, músculos y tendencias y cambios de flexibilidad corporal.

Plataforma de realidad aumentada

La figura 1 es un diagrama de bloques de sistema para una plataforma de realidad aumentada que soporta un procesamiento de vídeo/imagen grabado o en tiempo real. El sistema puede incluir uno o una pluralidad (1:n) de dispositivos 101 de entrada incluyendo una cámara de vídeo, una cámara fotográfica, una cámara de IR, una cámara en 2D o una cámara en 3D. El dispositivo 101 de entrada puede estar adaptado para enviar información a uno o más módulos 102, 103, 104, 105, 107, 108 y 109 de realidad aumentada de visión artificial. El uno o más módulos 102, 103, 104, 105, 107, 108 y 109 de realidad aumentada de visión artificial pueden estar adaptados para enviar información a una o una pluralidad (1:m) de pantallas 106. El uno o más módulos 102, 103, 104, 105, 107, 108 y 109 de realidad aumentada de visión artificial pueden estar adaptados para enviar/recibir información a/desde una interfaz o módulo 110 de interfaz de usuario. La interfaz 110 puede estar adaptada para enviar/recibir información a/desde uno o más de una nube, una web/tienda o un dispositivo de usuario, por ejemplo, teléfono inteligente o tableta.

El uno o más módulos 102, 103, 104, 105, 107, 108 y 109 de realidad aumentada de visión artificial pueden incluir un módulo 102 de captación de imágenes, un módulo 103 de transformación de coincidencia de ojos, un módulo 104 de realidad aumentada, un módulo 105 de grabación de vídeo/instantáneas, un módulo 107 de acontecimiento desencadenante, un módulo 108 de elemento de control, y un módulo 109 de calibración en fábrica.

El módulo 102 de captación de imágenes puede incluir una o más de las siguientes características: filtros de potenciación, conversión de formato, separación de tramas de vídeo, recorte de imagen, redimensionamiento de imagen, combinación de imágenes, y similares. El módulo 102 de captación de imágenes puede estar adaptado para enviar información al módulo 103 de transformación de coincidencia de ojos. El módulo 102 de captación de imágenes puede estar adaptado para enviar/recibir información a/desde el módulo 107 de acontecimiento desencadenante.

El módulo 103 de transformación de coincidencia de ojos puede estar adaptado para aplicar en la imagen el mapeo correcto para hacer coincidir el punto de vista de la cámara con el punto de vista del espejo teórico (reflejo de ojos de usuario) y rellenar los píxeles ciegos si los hay después del mapeo. El módulo 103 de transformación de coincidencia de ojos puede estar adaptado para enviar información al módulo 104 de realidad aumentada y/o al módulo 105 de grabación de vídeo/instantáneas. Además, el módulo 103 de transformación de coincidencia de ojos puede estar adaptado para enviar/recibir información a/desde el módulo 108 de elemento de control. Además, el módulo 103 de transformación de coincidencia de ojos puede estar adaptado para enviar información a la una o una pluralidad de pantallas 106, para visualizar una imagen que imita un reflejo en un espejo.

El módulo 104 de realidad aumentada puede estar adaptado, por ejemplo, para proporcionar sustitución de color y textura virtual, vestimenta virtual, inserción de objetos, y similares. El módulo 104 de realidad aumentada puede

estar adaptado para enviar/recibir información a/desde el módulo 108 de elemento de control y/o el módulo 105 de grabación de vídeo/instantáneas. Además, el módulo 104 de realidad aumentada puede estar adaptado para enviar información a la una o una pluralidad de pantallas 106, para visualizar una imagen que imita un reflejo en un espejo que se modifica por los cambios producidos por el módulo 104 de realidad aumentada.

El módulo 105 de grabación de vídeo/instantáneas puede estar adaptado para grabar una única imagen (trama) o una toma corta basándose en control de software. El módulo 105 de grabación de vídeo/instantáneas puede estar adaptado para enviar/recibir información a/desde el módulo 108 de elemento de control. Además, el módulo 105 de grabación de vídeo/instantáneas puede estar adaptado para enviar información a la una o una pluralidad de pantallas 106.

El módulo 107 de acontecimiento desencadenante puede incluir una o más de las siguientes características: reconocimiento de un usuario delante del espejo, reconocimiento facial, reconocimiento de comandos de gestos de usuario, reconocimiento de artículos, medida de distancia, medidas corporales/evaluaciones de usuario (incluyendo, por ejemplo, altura, edad, peso, grupo étnico, sexo, y similares) y cálculo de punto de vista teórico de usuario en espejo teórico. El módulo 107 de acontecimiento desencadenante puede estar adaptado para enviar/recibir información a/desde el módulo 108 de elemento de control.

El módulo 108 de elemento de control puede incluir una o más de las siguientes características: control y gestión para establecer la cámara para optimizar la calidad, para establecer otros elementos de hardware, una interfaz entre módulos de algoritmo y código superior/aplicación/interfases de usuario, y emitir datos calibrados desde la fábrica a los elementos de algoritmo. El módulo de elemento de control puede estar adaptado para enviar/recibir información a/desde el módulo 109 de calibración en fábrica.

El módulo 109 de calibración en fábrica puede estar adaptado para definir la transformación de mapeo entre la cámara y el punto de vista de usuario delante de la pantalla. Además, el módulo 109 de calibración en fábrica puede estar adaptado para calibrar la imagen basándose en distancia, una ubicación especial (por ejemplo, marcas en el suelo), altura de usuario o cualquier combinación de los mismos.

La figura 1 y la siguiente descripción representan tan sólo un ejemplo de una realización de la presente invención; pueden asignarse otros flujos o funcionalidades entre los módulos, que representan realizaciones adicionales que son parte de la invención. Los presentes inventores proponen dos métodos para habilitar las capacidades de realidad aumentada (tiempo real y fuera de línea). Ambos métodos envuelven el módulo 104 de realidad aumentada con datos de imagen o vídeo reales que son en tiempo real o se tomaron tras un procesamiento, por ejemplo, mediante el módulo 103 de transformación de coincidencia de ojos. Es decir, la imagen sobre la que se hace funcionar el módulo 104 de realidad aumentada puede ser una imagen transformada que representa un reflejo en el espejo de un usuario.

Una característica es que un usuario puede definir de manera manual o automática (por ejemplo, mediante la interfaz 110) reglas referentes a qué artículos le gustaría procesar y manipular al usuario y cuál es el resultado final esperado, por ejemplo, una regla automatizada puede ser algo como buscar la identificación automática de una camisa de un usuario, que puede cambiarse automática o manualmente a un color diferente, o que un usuario seleccione una camisa mediante selección manual con una pantalla táctil y selección de color manual. Después, puede procesarse el objeto seleccionado y extraerse/separarse por segmentación y grabarse en la base de datos vinculada al vídeo o imagen grabado original. Entonces, el módulo 104 de realidad aumentada procesa el modelo/máscara en tiempo real a una tasa de tramas dada, que puede ser a una tasa de tramas inferior o superior a la original, y a un tamaño igual o diferente del original. Algunas aplicaciones no requieren (pero, no obstante, pueden incluir) realidad aumentada en directo, como probarse ropa cuando el usuario quiere verse a sí mismo con la modificación (una o más opciones). Una vez guardado el objeto extraído de la escena en directo, es más fácil representar gráficamente múltiples cambios (color, textura, tamaño, y similares). Además, es más fácil realizar un procedimiento más largo, de manera mucho más precisa, con calidad superior y usando un procedimiento que proporciona más información, por ejemplo, movimiento de usuario, medidas corporales, y calidad basándose en integración de tramas, y similares. Para otras aplicaciones que requieren procedimientos en tiempo real, como una conferencia de vídeo, es mejor usar la segunda opción. Con la segunda opción, debe adaptarse el rendimiento para ser lo más próximo posible a la tasa de tramas real de la cámara con el fin de evitar la introducción de retardo/retraso que puede producir caídas de tramas y malos rendimientos. Si se encuentran retardos/retrasos, entonces tales retardos/retrasos pueden gestionarse de una manera que produce un vídeo suave sin retraso de las tramas. Tal gestión puede lograrse con memorias intermedias, sellos de tiempo de imagen más algún retardo, o con sellos de tiempo y replanificaciones de tramas para su representación gráfica.

Para la entrada de vídeo, se recomienda encarecidamente que el procedimiento de representación gráfica se realice en un dispositivo con DSP o GPU con el fin de evitar la introducción de retardo en el vídeo. La creación de la máscara/modelo también puede realizarse para la primera opción (no en tiempo real) en una CPU. Para la segunda opción, la máscara/modelo también puede realizarse y calcularse en la GPU/DSP.

En el módulo 107 de acontecimiento desencadenante, parte de la funcionalidad de factor desencadenante puede

automatizarse completamente, por ejemplo, puede iniciarse un procedimiento si se realiza una detección facial o detección de presencia. Algunos de los factores desencadenantes pueden realizarse de una manera semiautomatizada a partir del módulo 110 de interfaz de usuario, que puede incluir cualquier manera de controlar el dispositivo informatizado. Parte de la funcionalidad del acontecimiento desencadenante es calcular la transformación de imagen basándose en información geométrica, calibración, y/o seguimiento de usuario en tiempo real, por ejemplo, ubicación del usuario, de los ojos, cabeza, manos, posición, movimiento, y similares. El seguimiento puede realizarse usando una o más técnicas tales como sustracción de fondo, reconocimiento de patrón, segmentación de color, y similares. La funcionalidad de cálculo de seguimiento de transformación también puede implementarse en los demás módulos.

El módulo 108 de elemento de control puede estar adaptado para configurar el ajuste del sistema, autenticación de dispositivo de cámara, y similares, y también puede proporcionar información a partir de la función de transformación de seguimiento al módulo de transformación de geometría real o módulo de realidad aumentada, y similares.

Con el módulo 109 de calibración en fábrica, parte de la información necesaria para calcular la transformación que va a aplicarse en la imagen/vídeo puede generarse durante la calibración en fábrica o puede calcularse basándose en información adicional sobre la orientación real de la cámara en el campo, por ejemplo, altura por encima del suelo o mesa o similar, punto de vista en 3D, campo de visión de lente (FOV), y similares. La información de fábrica más la geometría real de la implementación pueden procesarse y suministrarse al elemento correcto en el sistema que usará la información para una mejor calibración y precisión.

Además de la aplicación a una pantalla local para una aplicación de espejo, la presente invención puede usarse para enviar el vídeo más los medios para transformarlo o el vídeo manipulado y transformado real a otra dirección en la nube o de manera local como una conferencia de vídeo individual o bidireccional.

La figura 2 representa un ejemplo de un módulo de realidad aumentada, que puede corresponder con el módulo 104 de realidad aumentada descrito anteriormente. Específicamente, el módulo de realidad aumentada puede tener una función de permitir que un usuario se vista de manera virtual, cambie de aspecto, tal como el color, accesorios, etc. En esta realización, el sistema obtiene imagen o vídeo de entrada a partir, por ejemplo, del método 201 informatizado de EyesMatch o a partir de cualquier otra fuente de imagen/vídeo, por ejemplo, teléfono inteligente de usuario, cámara de seguridad, dispositivo Google glass, cámara móvil o cámara estacionaria. Realizaciones adicionales pueden incluir información geométrica adicional que ayudará a calcular una proporción tal como altura de usuario, mirada, y similares. Si el vídeo o imagen de usuario procede del módulo de EyesMatch (imagen/vídeo calibrado), puede crearse un modelo más exhaustivo que permite medidas corporales, postura de objetos, tamaño, detección de orientación altamente precisa, y similares. La información adicional que puede calcularse a partir del objeto o vídeo calibrado puede permitir ajuste de objetos, sustitución de objetos e inserción de nuevos objetos en la trama/vídeo, puesto que cualquier distorsión introducida por la ubicación y el campo de visión de la cámara se ha tenido en cuenta y corregido. Estas correcciones permiten medidas altamente precisas de la altura, cintura, etc. del usuario, y ajuste del cuerpo del usuario a tipos de cuerpo clasificados generalmente.

El módulo 202 de elección puede obtener información de elección a partir de la interfaz 206 de manera manual a partir del usuario (X,Y o nombre de objeto) o de manera automática a partir de un método de elección, por ejemplo, un mecanismo que puede detectar automáticamente objetos predefinidos tales como pantalones, vestidos, camisas, y similares. Por ejemplo, la interfaz 206 puede permitir al usuario seleccionar una prenda de ropa que va a cambiarse, por ejemplo, cambiar el color del tipo de material textil. Esta elección se envía al módulo 202 de elección, de modo que se extrae el elemento apropiado que va a modificarse del resto de la imagen.

La extracción del objeto es una tarea bastante difícil. Por ejemplo, si el usuario lleva puesta una camisa roja, y en el fondo hay una silla roja y el usuario está de pie sobre una alfombra roja, el ojo humano puede distinguir fácilmente qué partes rojas pertenecen a la camisa y cuáles no lo hacen. Sin embargo, realizar tal determinación automáticamente con un ordenador es una tarea muy difícil. Además, dependiendo de la iluminación y los pliegues de la camisa, diferentes píxeles pertenecientes a la camisa aparecerán como colores diferentes. Por tanto, si se va eligiendo simplemente todos los píxeles rojos, algunos píxeles oscuros pertenecientes a la camisa no se incluirán, mientras que se seleccionarán algunos píxeles pertenecientes a la silla y a la alfombra.

El módulo 203 de extracción puede determinar la ubicación del objeto (es decir, todos los píxeles que pertenecen al objeto) y muestras del color (o el color promedio del objeto, que puede tener más de un color). El módulo 203 puede usar esta información para crear una máscara en blanco y negro que se usa en primer lugar para generar una máscara en color y sombreada con texturas en 2D o 3D.

La técnica para extraer el objeto se basa en una correlación de color en 3D o cualquier otra técnica tal como la distancia euclidiana más próxima entre el color promedio de objeto y el color de píxel para separar los píxeles del objeto con respecto a toda la imagen. La decisión sobre si el píxel está en el objeto o no puede realizarse en múltiples niveles y no se limita a los siguientes ejemplos:

1. Correlación de color y la primera decisión puede basarse en un umbral de distancia euclidiana, en el que el umbral de distancia euclidiana está en el espacio de color de RGB o el espacio de color cromático. Es decir, para cada píxel, se mide una distancia euclidiana desde el color de píxel al color de objeto según se determina tomando un único color o color promedio del objeto. Si el color de píxel está a una distancia mayor que el umbral, se determina que no pertenece al objeto.

2. Filtrar ruido aplicando operadores morfológicos tales como dilución y erosión, lo cual puede mejorar la decisión con respecto a píxeles que están "erróneamente etiquetados" como que forman parte o que no forman parte del objeto. En la mayoría de las implementaciones, la forma del objeto se conocería o puede estimarse. Por ejemplo, si el usuario selecciona los pantalones, el sistema tendría una estimación de la forma de los pantalones. Por tanto, los resultados de la selección en la etapa (1) anterior pueden ajustarse adicionalmente comprobando si cada píxel excluido está dentro de la forma estimada y debe incluirse en el objeto y, a la inversa, si un píxel incluido está realmente fuera de la forma estimada de los pantalones y debe excluirse.

3. Decisión basada en información a partir de tramas anteriores o posteriores, o a partir de píxeles vecinos en una fila o alrededor del píxel. Esta etapa representa una decisión principal en el procedimiento. En la mayoría de la implementación, sería inusual que se seleccionaran alternativamente píxeles consecutivos como pertenecientes y no pertenecientes al objeto. Los objetos físicos son más grandes y están cubiertos por más de un único píxel. Por tanto, los píxeles se agruparían como pertenecientes o no, y puede usarse una comprobación de píxeles vecinos para confirmar la selección.

4. Distancia de objeto a partir de la elección original, que puede usarse como umbral. Por ejemplo, para cada píxel se calcula su distancia cartesiana al píxel seleccionado original y, si está fuera de la distancia esperada, se considera que el píxel está fuera del objeto, y viceversa.

5. Continuación de superficie de objeto, mediante lo cual, si se sabe que una superficie de objeto continúa, entonces puede eliminarse mediante filtrado parte del ruido. Por ejemplo, a veces la imagen puede tener reflexiones o sombreado que provocarían que el color de píxel parezca drásticamente diferente del color real y, por tanto, se indique que no pertenece al objeto. Sin embargo, un objeto es generalmente una superficie continua. De ese modo, puede usarse la continuidad para eliminar tal mal etiquetado.

6. Bordes de objeto, mediante lo cual puede mejorarse la decisión alrededor de bordes mediante detección de borde que puede realizarse mediante filtros de paso alto (HP) u otras técnicas.

7. Decisión basada en energía de color. Uno de los problemas de la separación de color es que el color en condiciones de poca iluminación puede verse como negro, y el intervalo dinámico de la decisión se reduce significativamente. Pueden aislarse píxeles oscuros/negros y pueden aplicarse otras técnicas para decidir si los píxeles oscuros/negros pertenecen al objeto o no, por ejemplo, la presente invención puede determinar si el píxel está ubicado dentro del límite de objeto, o la distancia de la energía desde el color de desviación estándar (STD) de objeto cambia.

8. Usar información previa sobre la forma de objeto esperada para obtener mejores resultados.

9. En el caso en el que el objeto es una combinación de múltiples colores o formas o tiene logotipos u otras imágenes, puede usarse correlación y combinación de múltiples colores. Adicionalmente, puede usarse cualquiera de los métodos de múltiples niveles especificados anteriormente para obtener una decisión de nivel superior referente al objeto.

10. La decisión también puede basarse en una mayoría o una decisión referente a un píxel/imagen vecino como factor ponderado en la decisión. En el caso en el que se procesa la decisión sobre la imagen como vector, puede resultar más fácil analizar los vecinos en la misma fila o misma columna dependiendo de cómo vuelve a conformarse la matriz de imagen para dar un vector.

11. La estimación de la desviación estándar (STD) de color de material/textura de artículo también puede añadir información significativa para la segmentación de objeto. En alguna implementación, se almacena una base de datos de diferentes materiales textiles y su textura para ayudar en tal determinación.

12. Cualquier combinación de una o más de las etapas 1-11 anteriores.

La máscara puede usarse para representar gráficamente una máscara simple en blanco y negro. Sin embargo, con el fin de crear una sensación convincente de un objeto, puede mantenerse información adicional a partir de la textura o aspecto de un objeto. Con el fin de obtener la información importante adicional, se aplica la máscara sobre la trama o vídeo original, y puede obtenerse la escala de brillo o sombreado de textura en RGB o escala de grises sobre el objeto. Esta información es mucho más precisa y convincente para cambios de color dado que guarda la textura de arrugas del objeto original, sombreado, reflejo de luz, firma de material y similares.

La máscara de modelo se construye en capas para una manipulación mejorada. Los ejemplos de una posible estructura de capas pueden ser los siguientes:

1. Máscara en blanco y negro (para separar por segmentación el objeto). La máscara en blanco y negro puede ser muy importante para distinguir entre el objeto y el fondo o entre el objeto y otro elemento alrededor del objeto. Pueden usarse múltiples técnicas para optimizar la decisión de máscara/límites de objeto.

2. Máscara de borde de objeto – que representa el borde o contorno del objeto.

3. Máscara de color rojo – que representa zonas rojas del objeto.

4. Máscara de color verde – que representa zonas verdes del objeto.

5. Máscara de color azul – que representa zonas azules del objeto.

6. Texturas que se aplican a todas las máscaras de color – que representan el aspecto de textura del objeto.

7. Máscara de sombra o brillo – que representa zonas sombreadas o brillantes del objeto. Esto puede ser un mapa de intensidad de todos los píxeles pertenecientes al objeto.

8. Máscara de reflejo de luz de material– que representa reflejo de luz del objeto.

9. Máscara de absorción de luz de material– que representa zonas de absorción de luz del objeto.

10. Máscara de otros sensores tales como IR, microondas, profundidad, ultrasonidos, ultrabanda, y similares.

11. Capas similares a las descritas anteriormente.

Una vez que el modelo de máscara tiene la información requerida, con el fin de cambiar el color o la textura, puede usarse el módulo 204 de representación gráfica para modificar la capa/capas específicas y regenerar el objeto a partir de las múltiples capas dando como resultado un vídeo 205 representado gráficamente que es extremadamente realista. Por ejemplo, si el usuario llevaba puesta una camisa roja, puede sustituirse la máscara roja por una máscara azul para representar gráficamente al usuario llevando puesta una camisa azul. Dado que todas las demás máscaras permanecen iguales, la camisa azul se representará gráficamente con todo el sombreado, brillo, reflejo, textura, etc., de la camisa roja, representando por tanto gráficamente una camisa azul muy realista que tiene los mismos pliegues y que sigue los mismos contornos del cuerpo que la camisa roja. El efecto de algunas capas puede introducirse multiplicando o añadiendo la capa modificada a la trama. La sustracción y división también pueden definir relaciones entre capas. Técnicas adicionales que permiten una manipulación de artículos más compleja incluyen una técnica de registro, que puede basarse en unos pocos puntos, puede estirar/transformar un objeto o material para ajustarse dentro de los límites del objeto manipulado.

El módulo puede obtener el vídeo/imagen original, la máscara modelada de múltiples canales y el cambio requerido. El cambio requerido puede ser cualquier combinación de color, brillo, material, textura, icono/logotipo, impresión, material textil, y similares.

En una realización, el cambio requerido puede estar fuera o dentro de los límites de objeto original y puede crearse una máscara modificada para los límites de nuevo objeto para sustituir al modelo de máscara original.

En una realización, el cambio requerido es material textil con textura orientada específica y color que puede insertarse en una orientación específica. En esta realización, la orientación de material puede modificarse y aplicarse en consecuencia. Por ejemplo, una base de datos puede almacenar máscaras correspondientes a diferentes materiales textiles, de tal manera que un usuario puede cambiar el material textil de un artículo que lleva el usuario simplemente seleccionando un material textil diferente en la interfaz de usuario. El sistema sustituiría entonces la máscara correspondiente al artículo real que lleva el usuario, por una máscara de material textil correspondiente al material textil seleccionado.

En una realización, el cambio requerido puede ser otra máscara de prueba del usuario. Usando una técnica de registro, puede aplicarse la otra máscara de prueba del usuario para ajustar el nuevo objeto sobre el objeto similar del usuario para probarlo. Por ejemplo, se mide a un usuario para una camisa y se mide a otro usuario para una chaqueta. Con el fin de mostrar al primer usuario cómo queda la chaqueta del segundo usuario, el método puede usar una técnica de registro que ajusta la chaqueta del segundo usuario para obtener y aplicar la forma de líneas corporales correcta del primer usuario, lo cual puede crear un ajuste más convincente. Un beneficio de este método es que la chaqueta parecerá mucho más real ya que se representa gráficamente para conservar toda la textura del segundo usuario. Para producir estos resultados, el método puede registrar la orientación, ubicación y tamaño del primer usuario. El punto de registro puede incluir, por ejemplo, el punto medio entre los ojos, los bordes de los

hombros, los bordes de la cintura, las rodillas, etc. Estirando o contrayendo el punto de registro de un usuario a otro usuario, la prenda de ropa de un usuario puede mostrarse como si la llevase puesta otro usuario.

5 En una realización, la máscara puede usarse como punteros para registro de objetos virtuales. En una realización, puede usarse la grabación de máscara más fondo en la transformación correcta como usuario para eliminar al usuario/artículo a partir de la imagen/vídeo.

10 En una realización, puede obtenerse la elección del objeto individual o los múltiples (1:n) objetos que se necesita modelar. A partir del vídeo, se crea una máscara por cada trama. Si el usuario está girando, puede crearse un modelo en 3D o parcialmente en 3D trama a trama. A partir de este modelo de trama a trama, pueden obtenerse diferentes perspectivas y usarse para crear un modelo en 3D que incluye algunos o la totalidad de los movimientos del usuario. Después, puede usarse esta información para crear una vestimenta virtual más convincente. Es decir, el presente método puede usar los propios movimientos del usuario en la formación del modelo.

15 En una realización, la representación gráfica puede realizarse en la GPU, CPU, GPU en la nube o CPU en la nube. Los elementos introducidos que van a representarse gráficamente pueden proceder de la CPU, de la base de datos de usuario en la nube, o de un enlace activo con el inventario/cualquier otra base de datos/impresión en 3D, base de datos de comercio electrónico, base de datos social, y similares.

20 Los cambios en el color de artículo pueden basarse en inventario disponible y de prueba real o impresión en 3D opcional en el sitio o en la fábrica para personalización por el usuario. Una ventaja con respecto a la vestimenta virtual regular es que el artículo se muestra envuelto sobre el cuerpo como lo haría en realidad, dado que todos los pliegues, sombreado, etc., se transferirán con las máscaras apropiadas. Esto es una característica muy importante desde muchos aspectos. El usuario puede percibir cómo se siente el artículo sobre su cuerpo, cómo puede afectar a y cambiar la forma de su cuerpo, y similares.

25 En una realización, puede añadirse un accesorio o cualquier otro artículo aprendiendo el movimiento dinámico y el modelo de máscara del objeto relevante. Además, puede aumentarse el fondo para cambiar o crear un entorno diferente mediante la misma técnica. Una vez etiquetados todos los objetos requeridos, pueden enmascarse los objetos requeridos y puede usarse la máscara combinada para cambiar el fondo.

30 En una realización, el módulo de representación gráfica puede representar gráficamente el objeto con una técnica de representación gráfica potenciada de interpolación del objeto y la trama para dar una resolución superior, puede combinar el objeto a alta resolución, puede suavizar los bordes, y puede decimar el objeto de vuelta a la resolución requerida con mejor calidad de integración en la trama. Las técnicas adicionales incluyen trabajar directamente en el borde del objeto calculando el promedio con algún factor de ponderación del valor del píxel con el fin de combinar mejor el objeto con el color de fondo.

35 La figura 3 representa un módulo de realidad aumentada que puede sustituir una parte del cuerpo, color, orientación y textura de un artículo u objeto en el primer plano o fondo de la imagen, por ejemplo, el módulo puede añadir pelo al usuario, cambiar el color de pelo, piel y ojos del usuario, puede cambiar la postura de los ojos, y similares.

40 Los módulos 301, 302, 303 y 306 pueden funcionar de una manera similar en comparación con los módulos 201, 202, 203 y 206 del módulo de realidad aumentada general anteriormente descrito. El módulo 304 puede tener la capacidad de calcular u obtener información adicional tal como postura de la cabeza o dirección de movimiento del cuerpo directamente a partir del módulo de EyesMatch o a partir del módulo 307 usando un detector dedicado para el movimiento y la orientación en 3D del artículo y puede usar esta información para modificar la parte requerida del cuerpo, por ejemplo, obtener la postura de la cabeza permitirá una corrección de la dirección de los ojos modificando los ojos de la máscara/modelo a la dirección requerida. Además, la detección de cabeza puede permitir añadir pelo en la orientación correcta, un sombrero, y similares.

45 En una realización en la que la máscara requerida en algunas dimensiones es menor que la máscara original, el módulo 307 puede realizar o suministrar transformación de EyesMatch en tiempo real o fuera de línea de la imagen o el vídeo de fondo grabado. La imagen o el vídeo de fondo transformado puede ser capaz de representar gráficamente una parte de la parte del cuerpo o artículo usando manipulación de primer plano o fondo, por ejemplo, en caso de que se desee eliminar la cabeza o el torso de un usuario del vídeo. La imagen o el vídeo de fondo transformado puede aplicarse sobre la máscara que captura la cabeza y torso, y el resultado puede representarse gráficamente con el vídeo original transformado o no transformado. El resultado puede ser un vídeo 305 representado gráficamente sin la cabeza y el torso del usuario.

50 Por ejemplo, en un caso más complejo, puede desearse representar una longitud más corta de un vestido dado. Puede requerirse la manipulación de máscara en el módulo 304 para crear una máscara más corta para el nuevo vestido, y la diferencia entre la máscara original y la máscara tras la manipulación puede ser una nueva máscara para su manipulación. En la nueva máscara, alguna parte será la estimación de las piernas del usuario y alguna parte representará el fondo que será visible ahora con la longitud de vestido más corta. La nueva máscara puede dividirse en piernas y fondo, y el nuevo objeto representado gráficamente puede adoptar la combinación de imagen

de fondo y piernas predicha para crear un nuevo artículo representado gráficamente. El resultado tras la representación gráfica del artículo modificado en el vídeo es un usuario con un vestido más corto. Pueden usarse múltiples técnicas para predecir el aspecto que deben tener las piernas en la zona que puede usarse.

5 Colograma

Para manipular el color o aspecto de un elemento en la trama, es necesario identificar los píxeles que pertenecen a ese elemento. Esto se realiza normalmente comparando colores y suponiendo que los píxeles del mismo color pertenecen al mismo objeto. Sin embargo, tal procedimiento no es preciso. En las realizaciones divulgadas, se realiza la comparación usando variables distintas del color. El método intenta simular el ojo humano, que puede distinguir objetos incluso cuando partes de los objetos pueden estar sombreadas y otras partes pueden estar muy iluminadas o incluso ser reflectantes. Por ejemplo, en una realización, el objeto se reconoce como un objeto conocido con límites predecibles.

Cuando se implementan como espejo virtual, los objetos pueden incluir prendas de vestir, tales como camisas, pantalones, chaquetas, etc. Así, cuando un usuario selecciona una camisa para su manipulación, el sistema utiliza una asociación predictiva para predecir la ubicación de los píxeles que deben incluirse en el objeto. Por ejemplo, todos los píxeles por encima de la cabeza del usuario y todos los píxeles por debajo de la cintura del usuario no deben estar relacionados con la camisa. Por tanto, se reduce el espacio de búsqueda. Este algoritmo predictivo se mejora cuando se aplica a una trama que ha experimentado el procedimiento de calibración de EyesMatch. Para proporcionar un ejemplo sorprendente de la combinación de estos métodos, si el usuario elige cambiar solo el color de los zapatos del usuario, cuando se calibra la trama usando EyesMatch, se conoce la ubicación aproximada de los zapatos, y todos los píxeles por encima de los pies del usuario pueden ignorarse cuando se buscan los píxeles pertenecientes a los zapatos. Esto limita drásticamente el campo de búsqueda.

En resumen, para identificar el objeto que va a manipularse, además de la comparación de colores, también pueden evaluarse los píxeles usando la ubicación predictiva del artículo, la geometría conocida o aproximada del objeto, la textura del objeto, una desviación estándar de la textura, desviación estándar de un color representativo, etc. A cada una de estas variables se le puede asignar un peso específico en una evaluación estadística de cada píxel. Además, a cada parámetro se le puede asignar un error o valor de desviación estadística diferente.

La figura 4 representa métodos informáticos para crear un modelo para intercambio de color y textura, que puede denominarse colograma. Esta técnica es la más beneficiosa para el cálculo en paralelo que puede soportar un gran número de usuarios o un gran número de tramas/vídeo a diferencia de técnicas de intercambio de color de superalta calidad que pueden encontrarse en programas de software tales como Photoshop. Tales métodos pueden requerir mucho tiempo y pueden no resultar prácticos de llevar a cabo en un gran número de imágenes o vídeo de usuario. La descripción de la figura 4 es tan sólo un ejemplo y cualquier derivado del flujo representado forma parte de la presente invención.

Una dificultad para cambiar un color de un objeto en un vídeo o imagen es identificar con precisión los píxeles relevantes del objeto. En un archivo de vídeo, la velocidad es un factor limitante para una transformación aplicable. En la figura 4, se representa un ejemplo simplificado de un método para segmentar/extraer un objeto a partir de un vídeo. El método de la figura 4 puede ejecutarse por los sistemas descritos con respecto a las figuras 2 y 3.

La imagen o el vídeo que va a modificarse se recibe en 401. En 402, se convierte la trama de la imagen o el vídeo a color en un vector lineal, que es opcional, aunque la vectorización de la imagen puede acelerar drásticamente el tiempo de procedimiento. El tamaño del vector es de 1 por n, donde n es el número total de píxeles en la trama, es decir, el número total de píxeles en la anchura por el número total de píxeles en la altura de la trama, 3 veces para cada uno de los colores de RGB. Además, en 403, se elimina el efecto del brillo. Hay muchas técnicas para eliminar el efecto del brillo. En este ejemplo, se usa un cálculo de promedio de energía por cada píxel en el espacio cromático XYZ, dividiendo cada píxel entre la suma de XYZ. El espacio de color XYZ de CIE abarca todas las sensaciones de color que una persona media puede experimentar. Y se define como la luminancia, y para cualquier valor de Y dado, el plano XZ contendrá todas las cromaticidades posibles en esa luminancia. Por ejemplo, puede usarse una matriz de 3 X 3 para convertir RGB en XYZ, usando las coordenadas de cromaticidad de un sistema de RGB (x_r, y_r), (x_g, y_g) y (x_b, y_b) y su blanco de referencia (X_w, Y_w, Z_w).

En paralelo, en 404 se realiza la selección de objeto determinando en primer lugar un color de comparación que puede usarse para examinar todos los píxeles para la determinación de si cada píxel pertenece o no al objeto. El color de comparación se determina usando análisis estadístico del color de varios K píxeles que se cree que están dentro del objeto. Esto puede realizarse seleccionando en primer lugar varios puntos K(x,y) que se cree que pertenecen al objeto que va a transformarse mediante, por ejemplo, el conocimiento de la geometría o la ubicación aproximada del objeto dentro de la imagen. K es el número de ubicaciones/zonas con color distinguido que pueden separarse por segmentación del fondo o de otros objetos. En alguna implementación, para cada ubicación, se selecciona una ventana, o zona, alrededor del punto específico, para garantizar que el punto o píxel no sea una anomalía, sino más bien representativo de la zona particular. Después, en 405 se hace pasar cada punto K(x,y) a través de la misma transformación que la realizada en el módulo 403. En 406, se realizan k iteraciones para

encontrar el color de cada píxel y a partir de eso para encontrar el mejor color representativo del objeto. Este color puede usarse para la investigación de seguimiento para encontrar todos los píxeles que pertenecen al objeto. K es igual a o mayor de 2 en esta técnica. Para cada k, se calcula la distancia euclidiana 2D o 3D. Se guardan la distancia mínima y el valor de K y se usan como el color de comparación o representativo. Esta operación puede realizarse con todos los píxeles a la vez en un procedimiento relativamente rápido.

$$\text{dist} = \sqrt{(X - x_i(k))^2 + (Y - y_i(k))^2 + (Z - z_i(k))^2}$$

Después de K iteraciones, puede obtenerse la imagen etiquetada. La distancia euclidiana "dist" es tan sólo un ejemplo de un método computacional para distinguir entre colores; hay otros métodos para calcular la distancia entre colores, por ejemplo, un modelo de distancia color basado en percepción humana del color (cromática, saturación y brillo), técnicas de calibración avanzadas para hacer coincidir la sensibilidad y la capacidad para separar colores con el ojo humano tal como en CIE76, CIE94, CIEDE2000, y similares, o cualquier combinación con cámara de profundidad IR/3D con estiramiento de histograma, integración de color a lo largo del tiempo o cualquier otro método para mejorar la sensibilidad de la detección de color (módulo 411). La aplicación o el cruce de la información adicional a partir del módulo 411 puede suceder al nivel 406 de comparación de distancia, justo al final de la creación 409 de modelo, o cualquier combinación dependiendo de la naturaleza de la información adicional (determinista, estadística, variante con el tiempo, y similares).

Por ejemplo, cuando se usa una cámara de IR, la imagen de la cámara de IR puede registrarse con la imagen de la cámara convencional. Luego, la imagen de la cámara de IR puede usarse para separar fácilmente al usuario del fondo. Esta información puede usarse para limitar los píxeles de la cámara convencional que va a evaluarse si pertenecen o no al objeto. Por ejemplo, cuando la comparación se realiza usando un vector, la información de la cámara de IR puede usarse para reducir el número de entradas en el vector que deben evaluarse. De manera similar, puede usarse una máscara de textura para eliminar píxeles del vector de evaluación. Por ejemplo, la textura de la alfombra puede almacenarse y cualquier sección de la imagen que tenga la misma textura puede eliminarse de la evaluación dado que se sabe que el usuario no llevaría puesta una alfombra. Así, pueden usarse varias capas o máscaras para operar sobre el vector y aumentar la precisión de los resultados de aislar el objeto y asignar todos los píxeles que pertenecen al objeto.

Además de la diferencia de color, también pueden usarse otras técnicas que pueden añadir información sobre el objeto para mejorar la decisión tales como: probabilidad de zona (un píxel dado necesita tener vecinos o alguna masa de píxeles), característica de zona, filtros de límite para aislar el límite de objeto antes de realizar una decisión final, información de profundidad (que normalmente necesita hacer coincidir el contorno de la información de profundidad con la imagen final del objeto en 2D o 3D), integración en el tiempo para determinar si el píxel está en la zona de objeto a lo largo de múltiples tramas, y similares.

El módulo 407 es un ejemplo de una realización de cómo distinguir entre los colores requeridos y el otro espacio de color. En el módulo 407, se ponen a cero todos los píxeles con una distancia mayor que un umbral al no ser relevantes (un píxel con un color diferente de uno cualquiera de los colores 1 a k), y se asigna 1 a todos los píxeles relevantes, generando de ese modo una máscara binaria que identifica todos los píxeles que pertenecen al objeto.

El módulo 407 es un ejemplo en el que se desea separar un color o colores específicos. En este caso, pueden ponerse a cero todos los índices excepto por el requerido. El procedimiento avanza de la siguiente manera: poner a cero todos los índices no relevantes, obtener un fondo y valor de colores no relevantes = 0, y elegir el objeto de color requerido etiquetado = 1. Si hay más de un color en el objeto, puede asignarse 1 a cualquier índice elegido de 2 a k+1 y cero a todos los demás.

En 408, puede usarse un filtro en blanco y negro para limpiar ruido y suavizar la forma del objeto. Pueden usarse otras técnicas para mejorar la decisión de qué píxeles pertenecen al objeto. Como resultado, un índice para todos los colores relevantes comienza de 2 a K+1.

En el módulo 409, se aplica la máscara en blanco y negro obtenida sobre la imagen a color original y se obtiene el modelo en 3D para cambios de color y textura. El modelo puede ser un canal alfa en 2D de escala de grises o 3D en el espacio de color. El módulo 410 puede obtener un modelo en 2D o 3D del objeto. En caso de vídeo a partir de una única cámara, es posible obtener un modelo en 3D aunque el usuario esté moviéndose delante de la cámara, por ejemplo, girando delante de la cámara. En este caso, también es posible obtener medida de objeto en múltiples cortes para estimar las curvas del cuerpo en 3D del usuario.

A continuación se proporciona un ejemplo de una entrada en 2D en el generador de modelo, en el que se desea crear un modelo de la camisa azul del usuario mostrada en la figura 9. La figura 10 es un ejemplo de un modelo o máscara en 2D de la camisa, sin la información de color. En su lugar, se genera una máscara de escala de grises del objeto seleccionado, en este caso la camisa, y puede usarse posteriormente con cualquier color aplicado. De esta manera se conserva la textura de la camisa, de modo que es relativamente fácil manipular el color o la textura o incluso cambiar el límite del modelo para crear un objeto diferente.

El modelo basado tan sólo en diferencia de color no es perfecto en cuanto a calidad, por tanto puede usarse información y técnicas adicionales para mejorar la calidad del modelo de objeto (véase el módulo 411). Pueden aplicarse técnicas de información adicionales como interpolación y decimación o suavizado de bordes después del procesamiento mediante el módulo 410 con el fin de mejorar la calidad del modelo.

Medidas corporales

Las realizaciones referentes a "Medida corporal" no son según la invención y están presentes únicamente con propósitos de ilustración.

La figura 5 representa un ejemplo de un método para extraer una medida corporal a partir de una imagen, múltiples imágenes o un video corto. Las medidas corporales pueden usarse para diversas aplicaciones y servicios, tales como, por ejemplo, estimación de peso, determinación de ropa de mejor ajuste, monitorización de la forma corporal y el peso a lo largo del tiempo, etc. Dado que el sistema adquiere muchas imágenes del usuario en diferentes posturas y orientaciones, puede ser beneficioso seleccionar la mejor imagen para realizar las medidas corporales. La determinación de qué imagen seleccionar puede realizarse antes o como parte del procedimiento ilustrado en la figura 5. Específicamente, puede obtenerse una buena medida del cuerpo cuando la imagen es del cuerpo del usuario en alineación perpendicular al eje óptico de la cámara. Esto puede determinarse determinando en primer lugar si la cara parece simétrica a lo largo de una línea vertical que pasa a medio camino entre los ojos. Si la cara parece simétrica a lo largo de esta línea, puede usarse una línea vertical a lo largo del centro del cuerpo para confirmar que el cuerpo es simétrico a lo largo de esa línea. Para esta etapa, puede ser beneficioso eliminar en primer lugar las manos de la imagen, de tal manera que sólo se considere el centro del cuerpo a partir de la determinación simétrica. El proceso de retirar las manos de la imagen se describe a continuación con más detalle.

En el módulo 501, se analiza la imagen usando una línea corporal y análisis de indicador de masa corporal (IMC) usando en primer lugar un procedimiento similar al enfoque de EyesMatch descrito en el documento US2013/0229482. En este enfoque, puede reordenarse la imagen y transformarse para compensar la distorsión debida a la geometría de ajuste de la cámara y distorsión óptica. Los ejemplos de factores de distorsión de ajuste incluyen la distancia del usuario a la cámara, la altura de cámara, ángulo de proyección de cámara, la óptica del FOV de la cámara y otras distorsiones ópticas. Tras la corrección de la imagen, la representación de área de cada píxel puede ser aproximadamente uniforme en área (cm²). Puede realizarse el procesamiento en CPU, GPU, DSP en la cámara, ordenador local cerca de la cámara, o servidor remoto.

El módulo 502 es un elemento opcional, que puede transformar una imagen en 2D en una conversión de vector para acelerar el procesamiento. El tamaño del vector es la anchura de la imagen, multiplicada por la altura, multiplicada por tres píxeles (para RGB). La imagen introducida en el módulo 501 también puede ser ya un vector. Puede ser más fácil realizar algunos tipos de procesamiento de imágenes en una matriz en 2D mientras que otros tipos de procesamiento de imágenes son más adecuados para la imagen vectorial.

En el módulo 503, puede ser más fácil realizar la manipulación de color para eliminar el efecto de luz y saturación en la imagen vectorial. Esto se realiza de una manera similar a lo que se describió anteriormente con respecto a 403.

En el módulo 506, puede hacerse que la imagen pase por un filtro de detección de borde que puede enfatizar cambios de color y potenciar la visibilidad de la línea corporal. La detección de borde puede soportar bordes de la línea corporal de modo que puede soportar múltiples direcciones de borde. Por ejemplo, un filtro de borde con capacidades vertical, horizontal y +/- 45° puede producir una buena detección de borde de la línea corporal.

Los módulos 504 y 505 pueden proporcionar información adicional en el algoritmo que puede ayudar con el análisis de imagen, por ejemplo, sexo, edad, raza, y similares, del usuario y proporciones estadísticas asociadas con la información. Esta información puede usarse posteriormente para centrar la búsqueda de partes del cuerpo. Puede usarse la medida de altura con información de proporción corporal complementaria para centrar la búsqueda de partes específicas del cuerpo, por ejemplo, el busto, ombligo, cintura, caderas, y similares.

El módulo 507 puede proporcionar detección de cabeza. Hay múltiples técnicas para la detección de cabeza. Por ejemplo, una puede tomar la imagen con bordes enfatizados, voltearla a lo largo de la longitud del cuerpo, y realizar una correlación entre las imágenes. El pico de la correlación puede indicar la masa central del cuerpo. Otro método alternativo es un cálculo de masa de centro de centroide únicamente de los bordes. El cálculo de masa de centro de centroide puede ser menos preciso si la luz en el usuario no es lo suficientemente uniforme; sin embargo, esta técnica puede ser más rápida. Otras técnicas pueden basarse en reconocimiento de patrón, detección de ojos, nariz, y forma de la cabeza. Una vez encontrada la masa central, puede recortarse una ventana relevante para identificar el cuerpo a partir de la imagen. Además, el borde enfatizado de la punta de la cabeza puede producir la altura del usuario en píxeles. Dado que la imagen tiene un peso uniforme de longitud por píxel, la altura total puede calcularse en consecuencia. La suposición es que la transformación del módulo 501 producirá imágenes manipuladas de manera sistemática a lo largo de disposiciones, por ejemplo, 1-5 metros (m), en las que la altura real del usuario en la pantalla en píxeles será proporcional a la altura de usuario real. Pueden añadirse factores adicionales si esta

suposición no es exactamente el caso para todas las distancias.

El módulo 508 puede medir el tamaño de partes del cuerpo en píxeles. La técnica puede ser un procedimiento adicional en una zona de enfoque para enfatizar adicionalmente la parte del cuerpo. La figura 13 representa un ejemplo del resultado de procedimientos adicionales que eliminan la mano del cuerpo y apuntan a diferentes intervalos de dónde buscar medidas corporales particulares. Los intervalos pueden calcularse generando una línea central vertical, 1301, que puede pasar a través del centro de gravedad o el eje de simetría vertical de la imagen del usuario. Luego, pueden añadirse líneas de sección horizontal para generar zonas adaptadas para las medidas corporales. Por ejemplo, la línea 1303 delimita la ubicación en la que las piernas se conectan al torso, mientras que las líneas 1304 y 1305 delimitan el área más estrecha de la cintura.

Las medidas en píxeles pueden traducirse en centímetros (cm) según la curva predicha, por ejemplo, el cuello puede modelarse como un cilindro de modo que la anchura medida del cuello en píxeles puede representar 2 x radiales en píxeles, después los radiales en píxeles pueden convertirse en cm según la calibración. El busto para un hombre tiene un modelo más elíptico, de modo que la traslación será un poco diferente. Si hay medidas secundarias adicionales disponibles, pueden añadirse tales medidas para proporcionar información más precisa al modelo.

En el módulo 510, puede deducirse la medida a partir del tipo de cuerpo del usuario, por ejemplo, una forma de "pera" si las caderas son más anchas que el busto y la cintura (véanse las figuras 11 y 12). Además, algún análisis puede sugerir lo siguiente:

1. riesgo de salud por condición corporal para problema cardíaco, y similar;
2. el tipo de ropa que mejor se ajusta al tipo de cuerpo del usuario;
3. tendencias en el cuerpo del usuario para comparación histórica;
4. publicidad dirigida basándose en el tipo de cuerpo del usuario (por ejemplo, mesomorfo/endomorfo puede obtener cupones dirigidos para la mejor nutrición que se adapta a su tipo de cuerpo);
5. reconocimiento corporal; y
6. diagnóstico corporal puede monitorizar cambios corporales a lo largo del tiempo (condición de giro, flexibilidad, posibles tumores, y similares).

Puede aprenderse o estimarse información sobre el IMC del usuario a partir del modelo de objeto descrito en la figura 4 o directamente a partir del análisis de IMC de imagen tal como se describe en la figura 5.

En diversas aplicaciones es beneficioso "aprender" la textura de artículos específicos en la imagen. Según diversas realizaciones, para medir la textura con una buena resolución, en primer lugar se aísla el objeto de la imagen. En una realización, después de separar por segmentación un elemento o parte del cuerpo, puede implementarse una reconstrucción o medida de textura fina adicional. La medida de textura puede realizarse con un sensor 2D o 3D (cámara).

En realizaciones que utilizan un sensor 2D, pueden implementarse los dos ejemplos siguientes. Dado que el usuario se moverá delante del espejo, el sistema puede capturar dos imágenes (por ejemplo, imágenes consecutivas) que se usarán para calcular la profundidad o textura. Existen múltiples maneras de hacerlo y, como ejemplo, el sistema puede extraer dos o más imágenes del vídeo y, basándose en la detección de masa central, puede calcular la distancia y usar la distancia para el cálculo estereoscópico de la textura. Es decir, cuando se toma una imagen estereoscópica usando dos cámaras, se conoce la distancia entre las cámaras y su ángulo de eje óptico. Pero si sólo hay una cámara, entonces el sistema toma dos imágenes usando la misma cámara, pero separadas en el tiempo. En consecuencia, el usuario se habría movido un poco entre las dos imágenes. El sistema calcula entonces la diferencia entre las dos imágenes y realiza triangulación inversa. Esto puede hacerse calculando, por ejemplo, la masa central de cada imagen y luego calculando la distancia x-y entre las dos masas centrales (por ejemplo, en píxeles). Esto puede usarse para realizar triangulación inversa como si las dos imágenes se tomaran usando dos cámaras. Pueden usarse otros punteros para calcular la distancia, por ejemplo, el procesador puede identificar los ojos en las dos imágenes y calcular la distancia entre los ojos en las dos imágenes. Otra opción es determinar el giro del usuario, es decir, (r, θ) y usarlo para realizar la triangulación inversa. En cualquiera de estos métodos, dado que la cámara puede colocarse a una altura que está por encima de la masa central del usuario, por ejemplo, en la parte superior del monitor, puede usarse la triangulación inversa para calcular la profundidad, actuando de ese modo como un sensor de profundidad.

Alternativamente, el sistema puede encontrar punteros perceptibles en ambas imágenes e intentar hacerlos coincidir, eliminando entonces todos los punteros sin identificación clara. Esto puede realizarse usando, por ejemplo, la técnica de Coincidencia de Muestras Aleatorias RANSAM para encontrar a partir de un conjunto aleatorio de

punteros el grupo que tiene un comportamiento similar y usar esos punteros para encontrar la distancia en píxeles para la medida de textura o construcción en 3D de un elemento.

Otra opción es estimar la distancia entre el espejo basándose en la distancia y el giro del cuerpo. Por ejemplo, si el sistema puede medir la postura del cuerpo del usuario, puede estimar la textura.

Con el sensor 3D, el sistema puede hacer lo mismo que en 2D, aunque para mejorar el intervalo dinámico y la precisión el sistema necesita aislar el elemento que va a medirse. La sustracción de fondo puede mejorarse mediante un umbral aproximado de medida de profundidad. El sistema puede estimar la textura cada 1 o más tramas y simplemente suavizar los cambios de textura entre tramas mediante el filtro de Kalman o cualquier otra técnica de interpolación-extrapolación.

En algunas realizaciones, solo un sensor 3D generará una medida de profundidad que se usará para controlar un modelo paramétrico del avatar del usuario. El usuario puede vestirse de manera virtual y explorarse en 360°. Además, el sistema puede controlarse para adaptar el movimiento de posición hasta el usuario para aplicaciones en el espejo o conferencias de vídeo. Para hacer coincidir el avatar con el movimiento del usuario, puede usarse una técnica de registro en la que los punteros de registro pueden asignarse de múltiples maneras, por ejemplo, detección automática de elementos perceptibles y técnica de muestreo y coincidencia aleatorios de RANSAM. Una vez que el sistema tiene los punteros, es más fácil mover el avatar basándose en los movimientos reales del usuario. Los datos procedentes de un sensor de profundidad o cámara 3D pueden pasar por un motor de transformación de mapeo similar y el resultado puede ser la base de un avatar en 3D más preciso del usuario. El avatar puede usarse total o parcialmente como base de modelo para presentar elementos virtuales al usuario.

Otra técnica para ajustar el mapeo de registro es identificar el elemento en la imagen como la cabeza, las manos, las piernas del cuerpo superior y hacer coincidir el movimiento del avatar con el movimiento de estos elementos. El movimiento puede medirse y estimarse cada 1 o más tramas y la aproximación suave del movimiento del avatar puede realizarse mediante la técnica de filtro de Kalman. El movimiento del avatar puede realizarse en tiempo real, semi-tiempo real o fuera de línea.

En algunas realizaciones, el sistema recibe el flujo de imágenes (ya sea de una cámara, de dos cámaras o de una cámara 3D), y usa el flujo de imágenes para generar realidad virtual en tiempo real. Las imágenes originales de la cámara se descartan, y en su lugar, el mundo virtual se presenta en el monitor. Dado que las imágenes en tiempo real se usan para generar la realidad virtual, las imágenes de realidad virtual representan fielmente el mundo real tal como se toman imágenes mediante las cámaras. Por tanto, tal como se ha indicado anteriormente, las imágenes en tiempo real pueden usarse para generar un avatar en tiempo real, representando fielmente el avatar al usuario capturado en las imágenes. Por otro lado, pueden cambiarse el fondo y los artículos que lleva puestos el usuario, de tal manera que el avatar del usuario puede visualizarse como si estuviera en una ubicación diferente llevando los mismos o diferentes artículos que el usuario. Por ejemplo, si el usuario se prueba un abrigo para esquiar, las imágenes en tiempo real pueden usarse para generar un avatar que corresponda al usuario y que lleve puesto el mismo abrigo para esquiar, sin embargo, el fondo puede cambiarse para visualizar una pista de esquí, de modo que el usuario pueda obtener imágenes llevando el abrigo en una pista de esquí.

Por otro lado, también es posible almacenar las imágenes en tiempo real, y luego operar en las imágenes fuera de línea, de modo que esté disponible una mayor capacidad de procesamiento, para generar una película usando el avatar basado en el usuario. La película puede enviarse entonces al usuario, por ejemplo, a través de Internet.

Seguridad y otras aplicaciones

Las realizaciones referentes a "Seguridad y otras aplicaciones" no son según la invención y están presentes únicamente con propósitos de ilustración.

La figura 6 es un ejemplo de una aplicación de sistema de seguridad para aprendizaje y autenticación de usuario de múltiples niveles. Esto es tan sólo un ejemplo de implementación de flujo posible; cualquier combinación del flujo o separación funcional diferente entre los módulos de diagrama forma parte de la invención.

Como en las figuras anteriores, pueden introducirse instantáneas o vídeo a partir del dispositivo 601 en un módulo 602 de captación de imágenes. El módulo 602 de captación de imágenes puede controlarse mediante un acontecimiento desencadenante tal como los anteriormente descritos y procesarse mediante el módulo 607 de acontecimiento desencadenante, que puede proporcionar información de seguimiento y reglas para iniciar un procedimiento de captación de imágenes. Específicamente, el módulo 602 de captación de imágenes puede captar la imagen a partir de la entrada 601 y emitirla al módulo 603 de transformación de EyesMatch con la información de seguimiento adicional (medida geométrica en tiempo real). Basándose en la información de seguimiento y el requisito de acontecimiento, el módulo 603 de transformación de EyesMatch puede calcular una transformación para manipular con el punto de vista de cámara, ángulo de visión, y similares para crear una imagen calibrada de la escena. Además, la transformación de EyesMatch puede realizarse de manera local en GPU/CPU/DSP, en la nube y similares. En una realización, puede alimentarse información avanzada sobre el usuario o sobre la estructura

geométrica de la escena al algoritmo de seguimiento y puede participar en la calibración. La información adicional puede proporcionarse como información geométrica de una vez, por ejemplo, altura de usuario, distancia entre sus ojos, y similares. En ese caso, la información geométrica adicional puede usarse para calcular la transformación requerida. La diferencia entre la imagen distorsionada y la geometría conocida puede usarse para calcular y calibrar la transformación.

Entonces la salida calibrada a partir del módulo 603 de transformación de EyesMatch puede alimentarse en el módulo 604 de escáner de seguridad, que también puede denominarse módulo analítico de escena. El módulo 604 de escáner de seguridad puede estar adaptado para aprender uno o más de lo siguiente: exploración de líneas corporales o curvas del cuerpo basándose en múltiples técnicas, por ejemplo, detección de borde, detección de borde estadística, cámara de IR, sensor de microondas, cámara en 3D, cámara individual y múltiples cortes cuando un usuario está girando. También puede tomarse un modelo en 3D completo del cuerpo cuando el usuario está girando delante de la cámara de vídeo y puede extraerse la ropa del usuario a un canal alfa tal como se describe en la técnica representada en la figura 4, es decir, colograma, y similares.

Otros métodos de aprendizaje o autenticación incluyen, por ejemplo, detección de ropa, mediante lo cual se crea una firma en un artículo y se clasifica en una base de datos. La base de datos puede actualizarse a partir de Internet, explorando ropa a partir de múltiples bases de datos y tiendas de comercio electrónico o escaneando activamente y actualizando información por el operario, por ejemplo, puede introducirse un uniforme de personal de seguridad, el tipo de material textil, y similares.

Además, la combinación de EyesMatch con reconocimiento facial puede permitir que el operario instale una cámara por encima de la altura del usuario, por ejemplo, por encima de una puerta orientada hacia abajo a aproximadamente 30-45 grados; tal instalación permite al usuario moverse libremente por debajo de la cámara. En esta situación, por ejemplo, el intervalo de reconocimiento facial detectable es de aproximadamente 1-5 metros por delante de la cámara. Esto resulta ventajoso ya que los sistemas de reconocimiento facial conocidos no son muy precisos si la cabeza del usuario se escanea a un ángulo de +/- 15 grados.

Los sensores de autenticación corporal adicionales pueden incluir voz, olor, palma, huellas dactilares, ojos, ADN, esqueleto con sensores de rayos X o ultrasonidos, dientes, pelo, impresión, color, ojos, sangre, halo, temperatura, marcas de la piel, orejas, y similares.

El vídeo, instantáneas o los datos analizados pueden grabarse con un módulo 605 de grabación de vídeo/instantáneas directamente desde la cámara, desde el módulo 604 de transformación de EyesMatch, desde la imagen procesada o desde datos a partir de los sensores.

El vídeo/imagen visual o los datos analizados pueden visualizarse delante de la pantalla de usuario (cuerpo completo o más pequeña), suministrarse por la nube o suministrarse directamente a un centro de control para visualizarse en una pantalla 606 o a un módulo para su procesamiento y análisis adicional.

Los módulos 607, 608, 609 y 610 son similares a los módulos 107, 108, 109 y 110, respectivamente (véase la figura 1 y descripciones asociadas de la misma).

La figura 7 representa un ejemplo de una aplicación de sistema de seguridad para manipulación de cámaras en paralelo o en banda.

La capacidad del dispositivo en la figura 7 puede residir en un dispositivo autónomo tal como se representa en la figura 7 o como parte de la capacidad de DSP de cámara; el control de la activación de características puede realizarse de manera inalámbrica o a través de una infraestructura cableada. Además, parte de la capacidad puede soportarse en una ubicación remota, por ejemplo, medida y seguimiento. Además, el dispositivo puede desplegarse delante de la cámara y puede proyectar una imagen manipulada en una pantalla pequeña a la cámara.

El flujo principal de cámara puede alimentar el dispositivo o el dispositivo puede estar adaptado para intervenir en, y analizar, el flujo principal de cámara. Tras el procesamiento, el dispositivo puede estar adaptado para enviar datos en paralelo al operario.

El dispositivo puede tener capacidad de cálculo de múltiples imágenes, capacidad geométrica de EyesMatch, por ejemplo, en el módulo 703 de transformación de EyesMatch, medida corporal y autenticación de usuario, por ejemplo, en el módulo 704 de exploración de seguridad, y capacidades de realidad aumentada, por ejemplo, en el módulo 706 de realidad aumentada, en el que puede manipularse la imagen en el flujo principal o en paralelo al flujo principal, por ejemplo, las manipulaciones pueden incluir un cambio de un tipo de cuerpo de usuario, color, artículos sujetos por el usuario, peinado, desaparición completa del usuario, y similares. Estas capacidades son muy importantes para su uso en seguridad.

Además, el flujo de vídeo puede alimentar otros dispositivo tales como un robot, una dispositivo de visualización montado en la cabeza, una aplicación de usuario, y similares.

La interfaz para controlar la funcionalidad de dispositivo puede ser a través de la infraestructura cableada o inalámbrica o a una ubicación local o remota. El módulo de medida puede residir en el dispositivo o en la nube.

5 En una realización, la capacidad, por ejemplo, puede realizarse en tiempo real o fuera de línea. La activación del dispositivo puede realizarse, por ejemplo, de manera periódica, en respuesta a un acontecimiento desencadenante o de manera manual, según se necesite o se desee.

10 El dispositivo puede soportar control en tiempo real y factores desencadenantes para otros dispositivos, por ejemplo, un proyector para cambiar la imagen que puede ver el usuario (camuflaje), o un acontecimiento desencadenante para un dispositivo de eliminación de amenaza preciso que necesita una coordinación precisa delante de la cámara.

Puede compartirse vídeo en tiempo real con estaciones adicionales. Hay múltiples maneras de compartir el vídeo. A continuación se comentan en detalle algunas situaciones a modo de ejemplo.

15 Compartir la experiencia de espejo con uno o más de otros usuarios de manera remota, en la que el usuario se ve a sí mismo en una estación local y el usuario remoto ve al usuario en el espejo y comparte con él la experiencia. El usuario remoto puede tener cualquier dispositivo adecuado; la experiencia principal es ver al usuario que está de pie delante del espejo y está haciendo algo para que se revise. El usuario podrá ver una imagen de espejo de sí mismo basándose en EyesMatch o cualquier otra tecnología. La figura 8a representa esta situación. El usuario remoto puede simplemente ver la radiodifusión del usuario, enviar mensajes de texto al usuario, hablar con el usuario y/o que se le vea en una ventana pequeña en la pantalla de espejo.

20 La figura 8a representa una experiencia de espejo compartido en una conferencia de vídeo/voz. Tal como se muestra en la figura 8a, la experiencia de espejo puede compartirse en directo con un(os) usuario(s) en la nube o a través de otra estación de espejo. En el primer caso, el usuario puede verse a sí mismo en un modo de espejo aunque el flujo de vídeo que va a la nube necesita pasar por un volteo de espejo adicional para corregir la orientación de mirada en el otro lado. Por tanto, cuando el usuario (en modo de espejo) mira de derecha a izquierda, sus ojos/mirada en el otro lado se moverán en la dirección correcta como si estuviera mirando en la dirección correcta.

25 La figura 8b representa compartir una conferencia de vídeo de cuerpo completo. A diferencia de una conferencia de vídeo regular en la que la cámara puede emitirse en flujo "tal cual" al otro lado, en la estación de espejo, la imagen que está enviándose al otro lado necesita voltearse de izquierda-derecha en la ubicación remota. Esto puede realizarse volteando el espejo capturado de manera local o en el lado remoto.

30 En la figura 8b, cuando dos o más usuarios se comunican entre sí en las estaciones de cuerpo completo, el/los usuario(s) en una estación también verá(n) al/a los usuario(s) desde la otra dirección en la orientación correcta. Tal como se muestra en la figura 8b, el usuario puede verse a sí mismo de pie en el lado izquierdo de la pantalla local, y el usuario puede ver al otro usuario en el lado derecho de la pantalla local. En este caso, el vídeo local ya reflejado de manera especular, se necesitará un volteo/reflejo especular adicional cuando se transmite en flujo el vídeo a la ubicación/ubicaciones remota(s) (cuando se emite por radiodifusión).

35 Cualquier combinación de reflejo especular, volteo, rotación, y similares del vídeo local o transmitido en flujo para hacer coincidir esta experiencia del espejo local y corrección de mirada en la otra ubicación remota forma parte de la invención.

40 En una realización, ambas estaciones presentan tecnología EyesMatch/están calibradas con respecto a las mismas geometrías y dimensiones; todas las estaciones crean un espejo de cuerpo completo y lo envían a la(s) ubicación/ubicaciones remota(s). La experiencia de cuerpo completo y EyesMatch también se producirá en la conferencia de vídeo, y el usuario tendrá la sensación de estar mirándose a los ojos.

45 En una realización, cuando hay más de un usuario en una ubicación, la tecnología EyesMatch o corrección de distorsión de cuerpo puede realizarse por cada usuario con capacidad de seguimiento por cada usuario o para ambos simultáneamente. Con el fin de eliminar el efecto de línea, en este caso, la capacidad de realidad aumentada puede sustituir el fondo tal como se describió anteriormente. En el modo de realidad aumentada, puede usarse colograma u otra técnica para sustituir el fondo por un fondo liso. La línea de conexión puede colocarse entre los usuarios para eliminar líneas o discontinuidad con respecto a los propios usuarios.

50 En una realización, en la que hay más de dos estaciones, la pantalla puede dividirse para permitir visualizar múltiples usuarios desde múltiples ubicaciones al mismo tiempo, o pueden colocarse múltiples pantallas unas junto a otras tal como se representa, por ejemplo, en la figura 8c. Específicamente, la figura 8c representa múltiples espejos/estaciones de cuerpo completo con pantalla dividida o múltiples pantallas. Tal como se muestra en la figura 8c, el orden en el que se representan los usuarios en el espejo permite una comunicación de mirada correcta. Cuando el usuario está mirando a la derecha, por ejemplo, en la pantalla remota, tras el volteo, estará viendo al usuario mirando a la izquierda hacia el usuario remoto.

Esto es tan sólo un ejemplo y debe organizarse cualquier número de usuarios cerca de la pantalla o en la ubicación remota para permitir un enfoque de mirada correcto. En una realización, puede aplicarse una mejora adicional de mirada y contacto visual para mejorar la experiencia. En una realización, la pantalla puede incluir capacidad de 3D, y capacidad de 3D por cada usuario en la pantalla.

En una realización, además de EyesMatch o cualquier procesamiento de imágenes informatizado en un lado, la ubicación remota puede tener capacidad informatizada adicional para manipular vídeo en el extremo alejado basándose en el seguimiento de usuario local, por ejemplo, si el usuario se acerca a la pantalla, el vídeo o imagen de extremo alejado se procesará para parecer un poco más grande y se proyectará para proporcionar una mejor sensación de una sesión en directo. Otro ejemplo es cuando un usuario se acerca en un extremo, verá un FOV más ancho en el extremo alejado.

Interfaz de usuario

Las realizaciones referentes a "Interfaz de usuario" no son según la invención y están presentes únicamente con propósitos de ilustración.

La presente invención incluye un control de espejo virtual y características de sistema virtual para facilidad de uso del espejo virtual. El control de espejo virtual y sistema virtual pueden incluir diversos modos de funcionamiento de estación y modos de funcionamiento que incluyen, por ejemplo, encender, apagar, reiniciar, espejo, visualización, inactivo, y similares. El control de espejo virtual y sistema virtual pueden incluir autenticación y registro de usuarios, que pueden ser automáticos, manuales o una combinación de automáticos y manuales. El control de espejo virtual y sistema virtual pueden incluir un flujo de funcionamiento óptimo para facilitar y mejorar la experiencia del usuario.

El control de espejo virtual y sistema virtual pueden incluir una o más interfaces para controlar el espejo, por ejemplo, gesto automático, voz, apuntar con los ojos, aplicación móvil, experto remoto, asistencia local, y similares. El control de espejo virtual y sistema virtual pueden incluir una combinación exhaustiva de interfaces, incluyendo las siguientes: reproducir un vídeo grabado (automático/manual); efectos avanzados (manipulación de color de artículo, características de realidad aumentada, características de juegos, cambio de fondo, efecto en 3D, efecto de iluminación, y similares); modo de pantalla (orientación/tamaño completo o pantalla dividida); técnicas de compartir y métodos para compartir la experiencia con amigos/expertos en tiempo real y sobre la marcha; aplicación de usuario móvil para controlar la visualización y compartición, por ejemplo, compartir vídeos e imágenes personales tomados a partir del espejo virtual o el dispositivo móvil del usuario; venta al por menor o empresa en remoto para gestión general (proporcionar información de venta al por menor sobre el usuario en tiempo real con capacidad analítica y de control avanzada); y un módulo para enseñar al usuario cómo hacer funcionar el espejo (animación, vídeo, voz, indicaciones visuales, y similares).

Las figuras 14, 15 y 16 representan ejemplos de un flujo de interfaz de usuario de espejo virtual (UI) que puede secuenciarse en conjunto para simular una experiencia de espejo. Fuera de las horas de funcionamiento, la pantalla de espejo y/o el dispositivo informático que alimenta la pantalla puede apagarse o permanecer encendido con un salvapantallas para eliminar un envejecimiento rápido de la electrónica de visualización. La estación de espejo puede estar habilitada para encenderse automáticamente. La estación de espejo puede controlarse mediante control remoto de manera local o a través de la nube para reiniciar y entrar en funcionamiento normal.

Cuando el sistema se enciende, antes de que un usuario se encuentre delante del espejo, el sistema funciona en un modo inactivo, durante el cual el controlador puede proporcionar o no una transformación para imitar un reflejo del espejo. Una vez que el usuario entra en una zona especificada delante del campo de visión de la cámara, el procesador es capaz de realizar un seguimiento del usuario en la imagen grabada. Basándose en el seguimiento, el motor de vídeo del procesador puede calcular una transformación de la imagen para imitar el comportamiento en el espejo.

En el modo inactivo, que puede funcionar normalmente cuando no hay ningún usuario delante del espejo, el monitor puede visualizar un salvapantallas, una imagen estática, fragmentos de vídeo, y similares. Todos estos ejemplos pueden denominarse imagen inactiva, en el que el término "imagen" incluye vídeo. Durante el modo inactivo, el monitor también puede visualizar la imagen de fondo grabada en ese caso particular por la cámara. Para visualizar adecuadamente la imagen de fondo, de modo que aparezca como un reflejo en el espejo, el motor de vídeo puede adoptar un ajuste por defecto (por ejemplo, 2 metros desde el monitor) y aplicar la transformación de 2 metros al flujo de cámara para crear un efecto de espejo en el entorno de fondo tal como se representa, por ejemplo, en la figura 14.

En diversas realizaciones, el sistema es capaz de reconocer de manera autónoma la presencia de un usuario delante de la cámara e iniciar automáticamente el procedimiento que imita un espejo, incluyendo identificación de usuario, coincidencia de cuentas, etc. En una realización, la presencia de un usuario se realiza analizando de manera continua las imágenes captadas por la cámara para detectar cambios en la imagen e identificar un usuario. En un ejemplo, una zona de desencadenamiento puede designarse colocando una estera o usando una marca

específica en la alfombra delante y dentro del campo de visión de la cámara. La alfombra puede estar designada con un patrón y color específicos que mejoran el seguimiento y la detección del usuario. Además, los colores de la alfombra pueden mejorar la calidad del vídeo, por ejemplo, mejorando el contraste específico. Por ejemplo, pueden usarse colores más claros para mejorar las condiciones de iluminación y mejorar la calidad del vídeo resultante. En una realización, el patrón en la alfombra puede alternarse a aproximadamente 5-10 cm con colores más claros y más oscuros de modo que, cuando un usuario está de pie con calzado de color oscuro o claro, el calzado del usuario se detectará y se seguirá fácilmente, y el error de distancia máximo no será de más que la mitad de la separación de color, es decir de 5-10/2. En una realización, la alfombra puede incluir instrucciones para los usuarios que informan al usuario sobre cómo hacer funcionar el espejo. Alternativamente, en una realización, puede proporcionarse un dispositivo de funcionamiento en el que un usuario puede tocar con su pie una almohadilla de recepción para hacer funcionar el espejo con sus pies en vez de con sus manos o con la voz. En una realización, el espejo puede pedir al usuario que permanezca de pie en la ubicación correcta, por ejemplo, cuando se acerca un usuario, el espejo puede conmutar de vuelta a modo inactivo o el espejo puede cortar las piernas del usuario en la imagen, y similares.

El sistema puede estar configurado de modo que, cuando el usuario entra en la zona de seguimiento o registro delante del espejo, el motor de vídeo del controlador reacciona y empieza a realizar un seguimiento del objeto. Basándose en la ubicación del objeto, el motor de vídeo puede ajustar la transformación de vídeo para imitar el comportamiento de espejo. Una entrada adicional en el seguimiento puede incluir la altura del usuario, ubicación espacial, postura, y similares, tal como se describe en el procedimiento de EyesMatch. En este punto, el usuario todavía no está registrado en el sistema, de modo que puede realizarse un procedimiento adicional de identificación de usuario. Por ejemplo, la autenticación puede incluir reconocimiento facial, una tarjeta especial que puede presentar un usuario al sistema para su exploración, un dispositivo móvil de usuario (audio, inalámbrico, código QR (respuesta rápida)), u otra característica de registro biomédica.

En una realización, dado que el reconocimiento facial no es completamente fiable al 100% del tiempo, pueden visualizarse características de reconocimiento adicionales para saltar en la pantalla durante una cantidad de tiempo apropiada, tal como unos pocos segundos. En el ejemplo actual, se abre un soporte 1505 de QR (figura 15) y el usuario puede mostrar al espejo una tarjeta de identificación de QR o QR que se le envió a su aplicación única, por ejemplo, en un teléfono inteligente, y comunicar eficazmente al sistema que la presente imagen facial está asociada con el usuario autenticado. Esta técnica puede usarse para permitir una fusión de usuario semiautomática en el caso en el que un usuario particular tenga más de una cuenta facial. Además, una vez que se fusiona información facial, pueden realizarse mejoras y/o medidas faciales adicionales para mejorar el reconocimiento del usuario a lo largo del tiempo. Es decir, cuando el usuario se autentica usando el código QR, se usa toda la información facial asociada con este código específico, de las sesiones actuales y todas las anteriores, para actualizar la información de reconocimiento facial para este usuario.

En un ejemplo específico, ilustrado en la figura 15, la imagen presentada en el monitor comprende un marco 1505 superpuesto a la imagen de vídeo. Por tanto, el usuario que mira al monitor ve su "reflejo" artificial en el monitor y el marco o soporte 1505 superpuesto. El usuario puede presentar entonces un código, de tal manera que el código encaje dentro del marco en el monitor. En un ejemplo, el código es un código de barras impreso en una tarjeta. Según otro ejemplo, el usuario puede descargar una aplicación en el dispositivo móvil del usuario, por ejemplo, un teléfono inteligente. La aplicación incluye un código de barras, tal como un código QR, que corresponde específicamente al usuario o al teléfono inteligente del usuario. Cuando el usuario ve el marco, el usuario abre la aplicación y presenta el teléfono inteligente de tal manera que el teléfono inteligente encaja dentro del marco. El controlador identifica entonces el código dentro del marco y, por tanto, identifica al usuario.

Cuando se identifica al usuario, puede abrirse la cuenta del usuario y pueden visualizarse las últimas grabaciones, por ejemplo, en una realización, puede visualizarse una configuración 1510 de miniatura, tal como la representada en la figura 15. Alternativamente, puede visualizarse cualquier otra barra de control de imagen. Si no se identifica al usuario, puede comenzar un procedimiento de registro de usuario, entonces, después de unos pocos segundos, puede abrirse una nueva cuenta y puede configurarse el espejo para empezar a grabar automáticamente. Por ejemplo, si no se identifica al usuario, puede visualizarse un código, tal como un código QR en el monitor, de tal manera que el usuario puede escanearlo con un dispositivo móvil para descargar la aplicación. Cuando se descarga la aplicación y el usuario completa el proceso de registro, la aplicación en el dispositivo del usuario incluiría un código que puede presentarse al marco en futuras visitas.

En una realización, el sistema puede grabar y almacenar un fragmento de vídeo durante N segundos, por ejemplo, 10-13 segundos es tiempo suficiente para permitir que el usuario tenga una buena impresión sobre cómo le queda un artículo particular y puede incluir, por ejemplo, que el usuario se dé la vuelta y similares. Así, mientras que el usuario puede continuar inspeccionando el elemento actual en el espejo, y el controlador puede continuar proporcionando transformación para imitar el espejo, el sistema registra sólo una subsección de toda la sesión. Esta subsesión puede almacenarse en la base de datos, por ejemplo, en la nube, y estar disponible durante futuras sesiones del mismo usuario, estar disponible para descargarse en el dispositivo del usuario, y/o puede estar disponible para el envío por el usuario a otros usuarios o dispositivos. Por ejemplo, el usuario puede ser capaz de subir la sesión a redes sociales u otras aplicaciones.

En una realización, aunque el espejo puede estar configurado para empezar a grabar automáticamente cuando reconoce al usuario, si el usuario no desea tal reconocimiento automático, el espejo puede estar configurado para guardar de manera local el vídeo, el usuario puede activar una de las miniaturas anteriores, el espejo puede estar configurado para conmutar a un modo de visualización y el espejo puede estar configurado para reproducir el vídeo. Es decir, puede pedirse al usuario que ejecute un procedimiento mediante el cual se corta una escena particular por el usuario y no se guarda en el sistema.

En una realización, puede añadirse una o más miniaturas 1510 en el espejo, lo cual puede ser útil cuando la pantalla está en una configuración estrecha, por ejemplo, 21:9, o cualquier otra configuración de pantalla ancha. En una realización, la miniatura 1510 puede visualizarse para aparecer en una barra independiente junto al flujo de espejo o bajo el flujo de vídeo imitado. Esta característica puede ser útil cuando la pantalla es más ancha de lo requerido o proporcionada para el flujo de espejo imitado. En una realización, se presentan al usuario seis (6) miniaturas 1510, aunque puede presentarse cualquier número adecuado de miniaturas. El tamaño de las miniaturas puede configurarse para soportar una proporción razonable para la pantalla. Cada miniatura 1510 incluye un hipervínculo a una sesión anterior del usuario, a la que puede accederse activando el hipervínculo. El hipervínculo puede activarse mediante diversos métodos, tales como, por ejemplo, movimiento de la mano, ratón, control remoto, etc. En un ejemplo, el controlador identifica la mano del usuario en la imagen. A medida que el usuario mueve la mano hacia arriba y hacia abajo, el controlador sigue el movimiento y destaca una miniatura correspondiente. Cuando el usuario cierra la mano del usuario para formar un puño, el controlador activa el hipervínculo correspondiente a la miniatura que se destacó en ese momento particular.

En una realización, tal como puede observarse, por ejemplo, en la figura 15, también puede visualizarse un indicador 1520 de grabación. El indicador de grabación puede incluir visualización de una expresión tal como "REC", un círculo de color rojo, y similares.

Además, el texto o el nombre del usuario pueden mostrarse tal como se muestra en 1515 en la figura 15. En una realización, una contraseña adicional puede asociarse con el usuario y visualizarse como texto 1515, y el usuario puede tomar esta autenticación basada en contraseña y aplicarla al procedimiento de registro. En una realización, un empleado o el propio usuario puede introducir el teléfono, correo electrónico, señalización de NFC o cualquier otra información de identificación y obtener un enlace a la nube directamente o con posterioridad.

En una realización, también puede añadirse una capacidad de control adicional, por ejemplo, empezar la grabación, borrar vídeo, añadir efectos de vídeo tales como iluminación, cambios de color, selección de fondo, y similares. En una realización, todos los vídeos pueden subirse a la nube y pueden borrarse de la estación local después de algún periodo de tiempo preprogramado. En una realización, los vídeos pueden procesarse después de grabarse con efectos adicionales que mejoran el vídeo o cambian el vídeo para reflejar efectos de realidad aumentada o para análisis de IMC/tomar medidas/medidas de líneas corporales. En una realización, la grabación de vídeo puede incluir la grabación de audio del entorno para su análisis adicional. En una realización, la grabación de vídeo puede incluir grabación de WLAN del entorno para la grabación de la MAC de los usuarios y posteriormente, con correlación adicional, puede estar adaptada para asociar la MAC del dispositivo móvil al usuario.

La presente invención incluye un sistema para facilitar el control por un usuario del espejo. En una realización, si el usuario todavía está en la zona predefinida y se reconoció una vez por el espejo, mientras todavía permanezca de pie ahí y no esté reconociéndose a nadie más, el usuario será capaz de controlar el espejo, por ejemplo, reproducir, iniciar y detener imágenes/vídeos, eliminar imágenes/vídeos, añadir características de realidad aumentada, y similares. El usuario también puede controlar el espejo mediante control por gestos o desde su dispositivo, por ejemplo, teléfono inteligente, a través de una aplicación dedicada o características de control adicionales como parte de la aplicación de venta al por menor. En una realización, el control por gestos puede permitir alguna funcionalidad básica y puede permitirse la aplicación del usuario para permitir mucha más funcionalidad con respecto al control por gestos. En una realización, el asistente en la tienda o un asistente disponible de manera virtual/remota en la nube puede ayudar al usuario a hacer funcionar el espejo. En una realización, el usuario puede establecer sus propias preferencias para la UI desde su aplicación o desde la web. Todos los ajustes dedicados del usuario pueden añadirse a la cuenta del usuario, por ejemplo, el usuario puede cambiar la duración de tiempo de grabación de una sesión, el número de miniaturas, y similares.

En un modo por defecto, el usuario puede ver sus últimos n vídeos. Los vídeos pueden haberse tomado en la ubicación de espejo específica o no, pero puede accederse a los mismos en una ubicación de almacenamiento central, por ejemplo, desde la nube. El usuario puede establecer el aspecto de las miniaturas de los vídeos de prueba locales. Desde su aplicación, el usuario puede ver todas las miniaturas y puede ser capaz de tocar las mismas o usar cualquier otro método de elección para activar la reproducción del vídeo asociado con la miniatura específica.

En una realización, las miniaturas 1510 están configuradas para aparecer encima del flujo de espejo imitado tal como puede observarse, por ejemplo, en la figura 15. Las miniaturas pueden enviarse al fondo cuando está reproduciéndose el vídeo que imita el espejo en tiempo real o en modo dividido cuando no hay ninguna identificación

de control por gestos tal como se muestra, por ejemplo, en la figura 16. El usuario también puede establecer la pantalla a un modo dividido tal como se muestra, por ejemplo, en la figura 16. Específicamente, la figura 16 representa un ajuste lado a lado en el que el usuario todavía puede verse a sí mismo en tiempo real en un lado de la pantalla y puede reproducir vídeos anteriormente grabados en la otra mitad.

5 En una realización, el control remoto del usuario puede obtener automáticamente la información de usuario identificado de modo que el asistente de la tienda puede enviar al usuario un enlace a su cuenta para el registro o para ayudar al usuario con capacidad adicional tal como cambiar colores de ropa o cambiar la ropa en general.

10 En una realización, el usuario puede usar su dispositivo móvil para iniciar una videollamada en el espejo con un experto remoto o un amigo que puede aconsejarle en tiempo real. La llamada puede realizarse desde una aplicación dedicada o cualquier otra aplicación de tercero tal como Skype. El flujo de vídeo de espejo puede alimentarse a Skype local y el usuario remoto puede obtener el flujo de espejo en tiempo real. Por ejemplo, el vídeo que imita el espejo almacenado o en tiempo real puede enviarse al dispositivo del usuario usando, por ejemplo, una conexión
15 WiFi. La aplicación permite entonces acoplar el vídeo descargado o transmitido en flujo a la aplicación de comunicación, tal como Skype. Por otro lado, la aplicación puede permitir al usuario enviar un enlace a través de, por ejemplo, correo electrónico o SMS, para permitir un acceso a la nube de terceros al vídeo que imita el espejo de transmisión en flujo en tiempo real.

20 La figura 17 representa una realización de un flujo de uso sugerido. Específicamente, la figura 17 representa una realización de flujo de uso de usuario de las realizaciones adicionales representadas en las figuras 14, 15 y 16 y sus descripciones asociadas. En la etapa 1701 el espejo está en posición inactiva, tiempo durante el cual se presenta una imagen inactiva en el monitor. La imagen inactiva puede ser, por ejemplo, un salvapantallas, publicidad, presentación con diapositivas o sólo una imagen del campo de visión delante de la cámara y transpuesta por el
25 controlador para imitar un reflejo en el espejo del campo de visión.

En la etapa 1702, un usuario se aproxima al espejo y se detecta la presencia del usuario, por ejemplo, por un sensor de movimiento o detectando un cambio de la imagen observada por la cámara, el sistema inicia el funcionamiento en modo de espejo. Es decir, el controlador realiza una operación de transformación en la imagen del usuario, de tal
30 manera que la imagen presentada en el monitor imita el reflejo del usuario en un espejo. En la etapa 1703, el sistema inicia un procedimiento de autenticación para identificar y autenticar al usuario. Por ejemplo, en una realización el sistema usa reconocimiento facial para identificar y autenticar al usuario, mientras que en otra realización puede autenticarse al usuario usando un dispositivo tal como un teléfono inteligente con capacidades tales como WiFi, Bluetooth, NFC, etc. Por ejemplo, en una realización, se usa reconocimiento facial para identificar
35 al usuario, pero se usa una segunda etapa, por ejemplo, el uso de un código QR para autenticar al usuario. Si no se ha identificado al usuario, por ejemplo, es un nuevo usuario, puede pedirse al usuario que abra una cuenta. Por ejemplo, puede presentarse un código QR en el monitor, permitiendo al usuario escanear el código QR usando un dispositivo móvil, para descargar de ese modo una aplicación al dispositivo del usuario. De lo contrario, o si el usuario declina abrir una cuenta, el usuario puede proceder como invitado, aunque algunas características, tales como acceso remoto a imágenes almacenadas, pueden no estar disponibles mediante acceso como invitado.
40

En la etapa 1704, se da al usuario control de usuario sobre el dispositivo de visualización. En una realización, se guardan las preferencias de control específicas para cada usuario particular y se activan una vez que se ha reconocido al usuario. De lo contrario, se habilita una interfaz de usuario general, por ejemplo, una interfaz activada
45 por gesto manual. En una realización, cuando se ha reconocido a un usuario, o se ha probado varios conjuntos en la misma sesión, se presentan pruebas anteriores como miniaturas en el lado del dispositivo de visualización principal. Cuando el usuario levanta la mano correspondiente al lado de las miniaturas, dependiendo de la altura de la mano, se destaca una miniatura correspondiente para su selección. A medida que el usuario levanta o baja la mano, se destaca otra miniatura correspondiente a la altura de la mano para su selección. Entonces, cuando el usuario cierra
50 su puño, se selecciona la miniatura destacada y se visualiza la imagen o el vídeo correspondiente en la pantalla principal.

Tal como se muestra en 1705, las imágenes y vídeo de cada sesión pueden almacenarse en la nube, y pueden ser accesibles de manera remota usando, por ejemplo, teléfonos inteligentes, tabletas, etc. Por tanto, un usuario puede
55 conversar con amigos de manera remota y recibir su opinión sobre el conjunto que está probándose el usuario. Por ejemplo, el usuario puede enviar a un amigo un enlace a la prueba grabada o a un vídeo en directo en tiempo real a partir del sistema. Por tanto, el usuario puede compartir la experiencia de compra con personas ubicadas de manera remota.

60 Además, tal como también se muestra en 1705, dado que el sistema es capaz de identificar al usuario y también calcular parámetros del usuario, por ejemplo, peso, altura, etc., el sistema puede ser capaz de acceder a una base de datos de artículos disponibles que se recomendarán al usuario basándose en estos parámetros. Más específicamente, si el usuario ha grabado dos pruebas de dos camisas diferentes dentro de la misma sesión, el sistema puede descifrar que el usuario está interesado en adquirir una camisa y realizar o bien recomendaciones
65 alternativas, es decir, camisas diferentes, o bien recomendaciones complementarias, por ejemplo, pantalones específicos que van bien con las camisas que se ha probado. Además, dado que el sistema puede identificar la

camisa y la marca de la camisa, puede ser capaz de ofrecer incentivos específicos de ese fabricante, tal como se muestra a modo de ejemplo en 1706.

Además, en 1706 al usuario se le puede ofrecer la capacidad de cambiar colores sin tener que cambiar realmente la ropa. Tal como se explicó anteriormente, mediante el uso de múltiples máscaras o capas, el sistema es capaz de cambiar el color del artículo, al tiempo que mantiene el efecto visual realista del material textil según se lleva puesto por el usuario. Dado que el sistema puede mantener una capa de sombreado, una capa de textura, una capa de reflejo, etc., la capa de color puede cambiarse mientras se mantienen todas las demás capas, de tal manera que la imagen representada gráficamente conserva todas las características de la imagen anterior, cambiando tan sólo el color.

Con el fin de evitar la complicación del flujo de la experiencia delante del espejo, una aplicación complementaria puede permitir características y ajustes adicionales, por ejemplo, un usuario avanzado que se siente cómodo con el control por gestos puede habilitar más gestos que ayudarán a hacer funcionar características avanzadas sin el uso de una aplicación de interfaz táctil. Para usuarios que no se sienten cómodos con los gestos manuales, a tales usuarios se les pueden proporcionar características automatizadas básicas y puede usarse la aplicación móvil para hacer funcionar la otra funcionalidad avanzada.

En una realización, la estación de espejo puede estar configurada para identificar el comportamiento del usuario y adaptar el control por gestos a la manera en la que el usuario intenta realizarlo, por ejemplo, algunos usuarios apuntarán, algunos usuarios agarrarán, y algunos usuarios empujarán para accionar una miniatura. El sistema puede estar adaptado para aprender y actualizar un control de perfil por cada usuario.

En una realización, la estación de espejo puede soportar instrucciones para el usuario, de modo que, cuando el sistema identifica a un usuario que no está controlando el espejo de la manera correcta, puede aparecer una breve explicación y guiar al usuario por el procedimiento.

En una realización, el espejo puede estar adaptado para demostrar características predeterminadas basándose en información referente a la familiaridad del usuario con el sistema, por ejemplo, basándose en si el usuario es un usuario por primera vez, basándose en el número de vídeos en el sistema para el usuario, basándose en la última vez que el usuario requirió que el sistema habilitara instrucciones, puede estar adaptado para proporcionar una asistencia por voz de funcionalidad relativamente más sencilla, puede estar adaptado para proporcionar asistencia de experto remoto y similares.

En una realización, el concepto de un número predefinido de miniaturas, tal como se muestra en las figuras 15 y 16, puede sustituirse por un menú deslizante de miniaturas.

En el modo de funcionamiento normal tal como se representa, por ejemplo, en las figuras 15 y 16, el usuario puede mover sus manos arriba y abajo para seleccionar miniaturas, una vez que el usuario detiene su mano, puede elegirse la miniatura más cercana y puede designarse de tal manera que el usuario obtendrá realimentación para la selección. Por ejemplo, puede proporcionarse un sombreado azul alrededor de la miniatura seleccionada para indicar su elección. Cuando el usuario agarra, empuja o apunta con sus manos/dedos a la miniatura, el vídeo puede empezar a reproducirse basándose en el ajuste de modo de visualización, y el usuario también puede detener el vídeo y realizar otras operaciones, y similares. Una vez que el usuario, por ejemplo, detiene el vídeo reproducido, puede visualizarse una barra de selección en la miniatura para designar el estado del vídeo reproducido y devolver acuse de recibo al usuario de que su comando se aceptó por la estación de espejo.

En una realización, para simplificar adicionalmente el funcionamiento, el control por gestos con la mano puede estar configurado para detectar un salto entre una miniatura aunque el usuario esté moviendo las manos a la izquierda y a la derecha en lugar de arriba y abajo o cualquier otra dirección. El sistema puede estar configurado para adaptarse al alcance y la velocidad de movimientos de manos del usuario, de modo que el usuario puede ver la elección de modo que puede también adaptar su velocidad y alcance de movimiento.

Las realizaciones pueden implementarse en un medio de almacenamiento legible por ordenador no transitorio para hacer funcionar un monitor, una cámara, y un procesador, para visualizar una imagen que imita un espejo en el monitor, que comprende: en un dispositivo que tiene el procesador y una memoria que almacena un programa para su ejecución por el procesador, incluyendo el programa instrucciones para: detectar a un usuario; iniciar un modo que imita un espejo para visualizar la imagen que imita un espejo en el monitor; iniciar un procedimiento de autenticación; y pedir al usuario que controle el monitor.

Comercialización aumentada

Las realizaciones referentes a la "Comercialización aumentada" no son según la invención y están presentes únicamente con propósitos de ilustración.

El sistema de espejos puede usarse para proporcionar una comercialización aumentada permitiendo un mejor servicio a los usuarios, fomentando la comercialización, y proporcionando realimentación a los comerciantes. El espejo puede estar configurado para complementar y aumentar el procedimiento de negocio. Puesto que la estación de espejo según la presente invención es un nuevo dispositivo en el mercado, los usos del espejo para crear un flujo de ingresos también son únicos. Las siguientes realizaciones incluyen detalles que proporcionan ejemplos en cuanto a cómo puede usarse un espejo activo para aumentar el negocio y para crear nuevos flujos de ingresos.

Por ejemplo, en una realización, puede añadirse el logotipo de una marca específica o socio o cualquier otro logotipo al vídeo grabado, de modo que el usuario obtendrá, por ejemplo, vídeo con el logotipo de la marca, y el usuario puede compartir el vídeo resultante con sus amigos. Por consiguiente, puede usarse una única sesión delante del sistema para promocionar un determinado artículo y/o marca. Por ejemplo, si el usuario sube la sesión en un sitio de redes sociales, muchas personas pueden ver la sesión, todas las cuales estarían expuestas al logotipo superpuesto.

En una realización, el espejo puede estar configurado para proporcionar capacidad de realidad aumentada convincente mediante un artículo que se probó anteriormente otro usuario con un tipo de cuerpo similar. Esta técnica de coincidencia proporciona una experiencia mucho más convincente y proporciona mejores recomendaciones al usuario.

El espejo puede insertar publicidad, cupones y logotipos en el espejo en modo inactivo, dentro de una porción especificada del espejo, o mediante modo de pantalla dividida. El espejo también puede introducir publicidad, cupones y logotipos en vídeos grabados de modo que un usuario puede ver el vídeo con la publicidad/cupón/logotipo. Estas visualizaciones pueden venderse como publicidad. El usuario puede compartir la publicidad/cupón/logotipo subiendo, por ejemplo, una sesión en sitios de redes sociales. Al usuario se le puede ofrecer un incentivo a cambio de compartir la publicidad/cupón/logotipo.

Puede analizarse la sesión de prueba de ropa por parte del usuario para determinar medidas y/o características tales como IMC, sexo, origen, edad, medidas corporales, expresión facial, expresión vocal, tallas recomendadas, y similares. Estos datos pueden incorporarse en aplicaciones de comercio electrónico. Estos datos tienen un alto valor y pueden compartirse con marcas, el usuario, un tercero basándose en un acuerdo con el usuario y las marcas, etc. Usando la presente invención, puede generarse un flujo de ingresos a partir de los datos analizados referentes al usuario, por ejemplo, presentando al usuario publicidad dirigida y precisa según las sesiones grabadas.

El espejo puede estar configurado como un servicio que permite que un usuario interaccione con expertos, un amigo o un asesor cuando contempla el usuario una compra particular. Pueden generarse ingresos adicionales usando la presente invención fomentando oportunidades de venta relacionadas con el consejo del experto remoto. Los expertos pueden puntuarse por el usuario, y un usuario puede escoger su propio experto favorito o seleccionar un experto designado por la marca.

Al usuario se le puede presentar una visualización que incluye una sugerencia y/o incentivo para un artículo particular que se ha probado el usuario, por ejemplo, con diferentes colores. Este método presenta una oportunidad adicional para fomentar un aumento de ventas. Puede usarse realidad aumentada para fomentar un aumento de ventas. Específicamente, basándose en información obtenida por el espejo y otras pruebas de usuario, puede generarse una visualización de realidad aumentada más convincente basándose en una combinación de información sobre el usuario y usuarios similares que tienen un mismo tipo de cuerpo.

En una realización, la red social del usuario puede integrarse para ayudar al usuario y/o comentar sobre la contemplación del artículo por parte del usuario. El sistema puede estar configurado para compartir los vídeos del usuario. Las características de red social pueden usarse para ampliar la base de datos de clientes abordables.

En una realización, el usuario puede subir sus propias fotografías, los motores de vídeo pueden estar configurados para procesar las imágenes y proporcionar un análisis similar de IMC, origen de cara, y similares. El sistema puede proporcionar tallas recomendadas para aplicaciones de comercio electrónico. El sistema puede mostrar el artículo que se compró en diferentes colores. De esta manera, el espejo puede usarse para obtener información de mercadotecnia para una base de datos apropiada aunque el usuario no esté físicamente de pie delante del propio espejo.

En una realización, el usuario puede ver su talla predicha en la aplicación y actualizar las medidas, lo cual puede usarse para mejorar el modelo de predicción de talla e incluir ajustes para marcas particulares.

Diseño físico

La presente invención incluye el diseño mecánico y el diseño de aspecto del equipo. La pantalla puede montarse para permanecer en la pared o bien en vertical o bien en horizontal o puede ser conmutable entre vertical y horizontal (mediante una solución mecánica más grande o similar, la pantalla puede inclinarse, hacerse rotar, y similares). La pantalla puede montarse en un soporte dedicado, montarse en una pared o montarse detrás de la pared. Cuando la pantalla está dentro de la pared, debe proporcionarse un conducto de ventilación térmica para

soportar la pantalla y el dispositivo informático.

En una realización, la pantalla puede tener una proporción de un espejo, por ejemplo, 21:9.

- 5 En una realización, la pantalla puede tener una proporción regular, es decir, 16:9, y el ajuste de la visualización puede ser pantalla dividida o barras oscuras en los lados para proporcionar la proporción de espejo, es decir, 21:9.

10 En una realización, la pantalla puede presentarse en un acabado mate para eliminar reflejos que pueden verse sobre colores oscuros. Dicho de otro modo, cuando el usuario lleva puesto algo negro, y la pantalla no tiene acabado mate, el usuario puede ver su propio reflejo de espejo verdadero; este efecto no se desea y puede reducir, distorsionar o arruinar completamente el efecto de espejo virtual.

15 La trama puede presentarse de una manera tal como se representa en las figuras 14, 15 y 16. Puede usarse una trama mínima o la trama puede ocultarse detrás de una pared.

20 En una realización, las cámaras o las cámaras controladas pueden estar ubicadas en la trama. La cámara puede estar montada por encima del marco de la pantalla, puede estar oculta y puede estar montada en la pared. La cámara puede estar dotada de un sombreado que puede eliminarse mediante filtrado de la imagen. Dado que la cámara está orientada hacia abajo, una cubierta que ocultará la mayor parte del cuerpo de cámara puede bloquear la vista desde una distancia.

La pantalla puede ser de LED, LCD, plasma, vidrio, proyector y similares.

25 Puede usarse un fondo que permite un mejor equilibrado de blancos que da como resultado vídeos de calidad superior.

La iluminación puede incluir una combinación de blanco, amarillo o cualquier combinación en un punto o una combinación de proyector y puede estar configurada para mejorar la calidad de vídeo y los colores.

30 Puede usarse una alfombra dedicada para permitir cambios de fondo, para definir una zona de usuario, para mejorar la detección de usuario y el seguimiento delante del espejo, para indicar al usuario dónde ponerse de pie e indicar al usuario cómo hacer funcionar el espejo.

35 Conclusión

40 En particular, y con respecto a las diversas funciones realizadas por los componentes, dispositivos, circuitos, sistemas y similares anteriormente descritos, se pretende que los términos usados para describir tales componentes correspondan, a menos que se indique lo contrario, a cualquier componente que realice la función especificada del componente descrito (por ejemplo, un equivalente funcional), aunque no sea estructuralmente equivalente a la estructura divulgada, que realice la función en los aspectos a modo de ejemplo ilustrados en el presente documento del objeto reivindicado. Con respecto a esto, también se reconocerá que la innovación incluye un sistema así como un medio de almacenamiento legible por ordenador que tiene instrucciones ejecutables por ordenador para realizar las acciones y/o acontecimientos de los diversos métodos del objeto reivindicado.

45 Los sistemas/circuitos/módulos anteriormente mencionados se han descrito con respecto a interacción entre varios componentes/bloques. Puede apreciarse que tales sistemas/circuitos y componentes/bloques pueden incluir los componentes o subcomponentes especificados, algunos de los componentes o subcomponentes especificados, y/o componentes adicionales, y según diversas permutaciones y combinaciones de los anteriores. También pueden implementarse subcomponentes como componentes acoplados en comunicación con otros componentes en vez de incluidos dentro de componentes originales (jerárquicos). Adicionalmente, debe observarse que uno o más componentes pueden combinarse para dar un único componente proporcionando funcionalidad agregada o dividirse para dar varios subcomponentes independientes, y puede proporcionarse una cualquiera o más capas centrales, tales como una capa de gestión, para acoplarse en comunicación con tales subcomponentes con el fin de proporcionar funcionalidad integrada. Cualquier componente descrito en el presente documento también puede interactuar con uno o más de otros componentes no descritos específicamente en el presente documento pero conocidos por los expertos en la técnica.

60 Además, aunque una característica particular de la innovación objeto puede haberse divulgado con respecto a tan sólo una de varias implementaciones, tal característica puede combinarse con una o más de otras características de las otras implementaciones según pueda desearse y resultar ventajoso para cualquier aplicación dada o particular. Además, el alcance de los términos "incluye", "que incluye", "tiene", "contiene", variantes de los mismos, y otros términos similares se usan o bien en la descripción detallada o bien en las reivindicaciones, se pretende que esos términos sean inclusivos de una manera similar al término "que comprende" como término de transición abierto sin excluir ningún elemento adicional u otro.

65 Tal como se usan en esta solicitud, generalmente se pretende que los términos "componente", "módulo", "sistema", o

similares se refieran a una entidad relacionada con ordenadores, ya sea de hardware (por ejemplo, un circuito), una combinación de hardware y software, software, o cualquier entidad relacionada con una máquina operacional con una o más funcionalidades específicas. Por ejemplo, un componente puede ser, pero no se limita a ser, un procedimiento que se ejecuta en un procesador (por ejemplo, procesador de señales digitales), un procesador, un objeto, un elemento ejecutable, un hilo de ejecución, un programa, y/o un ordenador. A modo de ilustración, tanto una aplicación que se ejecuta en un controlador como el controlador pueden ser un componente. Uno o más componentes pueden residir dentro de un procedimiento y/o hilo de ejecución y un componente puede estar ubicado en un ordenador y/o distribuido entre dos o más ordenadores. Además, un "dispositivo" puede presentarse en forma de hardware especialmente diseñado; hardware generalizado que se hace que sea especializado mediante la ejecución de software en el mismo que permite que el hardware realice una función específica; software almacenado en un medio legible por ordenador; o una combinación de los mismos.

Los dispositivos informáticos incluyen normalmente una variedad de medios, que pueden incluir medios de almacenamiento legibles por ordenador y/o medios de comunicaciones. Los medios de almacenamiento legibles por ordenador pueden ser generalmente cualquier medio de almacenamiento disponible al que puede acceder el ordenador, son normalmente de naturaleza no transitoria, y pueden incluir medios tanto volátiles como no volátiles, medios tanto extraíbles como no extraíbles. A modo de ejemplo, y no de limitación, los medios de almacenamiento legibles por ordenador pueden implementarse en relación con cualquier método o tecnología para almacenamiento de información tal como instrucciones legibles por ordenador, módulos de programa, datos estructurados, o datos no estructurados. Los medios de almacenamiento legibles por ordenador pueden incluir, pero no se limitan a, RAM, ROM, EEPROM, memoria flash u otra tecnología de memoria, CD-ROM, disco versátil digital (DVD) u otro almacenamiento en disco óptico, casetes magnéticos, cinta magnética, almacenamiento en disco magnético u otros dispositivos de almacenamiento magnéticos, u otros medios tangibles y/o no transitorios que pueden usarse para almacenar información deseada. Puede accederse a medios de almacenamiento legibles por ordenador mediante uno o más dispositivos informáticos locales o remotos, por ejemplo, mediante peticiones de acceso, consultas u otros protocolos de recuperación de datos, para una variedad de operaciones con respecto a la información almacenada por el medio.

Los medios de comunicaciones implementan normalmente instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos estructurados o no estructurados en una señal de datos que puede ser transitoria tal como una señal de datos modulada, por ejemplo, una onda portadora u otro mecanismo de transporte, e incluye cualquier medio de suministro o transporte de información. El término "señal de datos modulada" o señales se refiere a una señal que tiene una o más de sus características establecidas o cambiadas de tal manera que codifican información en una o más señales. A modo de ejemplo, y no de limitación, los medios de comunicación incluyen medios cableados, tales como una conexión cableada directa o de red cableada, y medios inalámbricos tales como medios acústicos, RF, infrarrojos y otros medios inalámbricos.

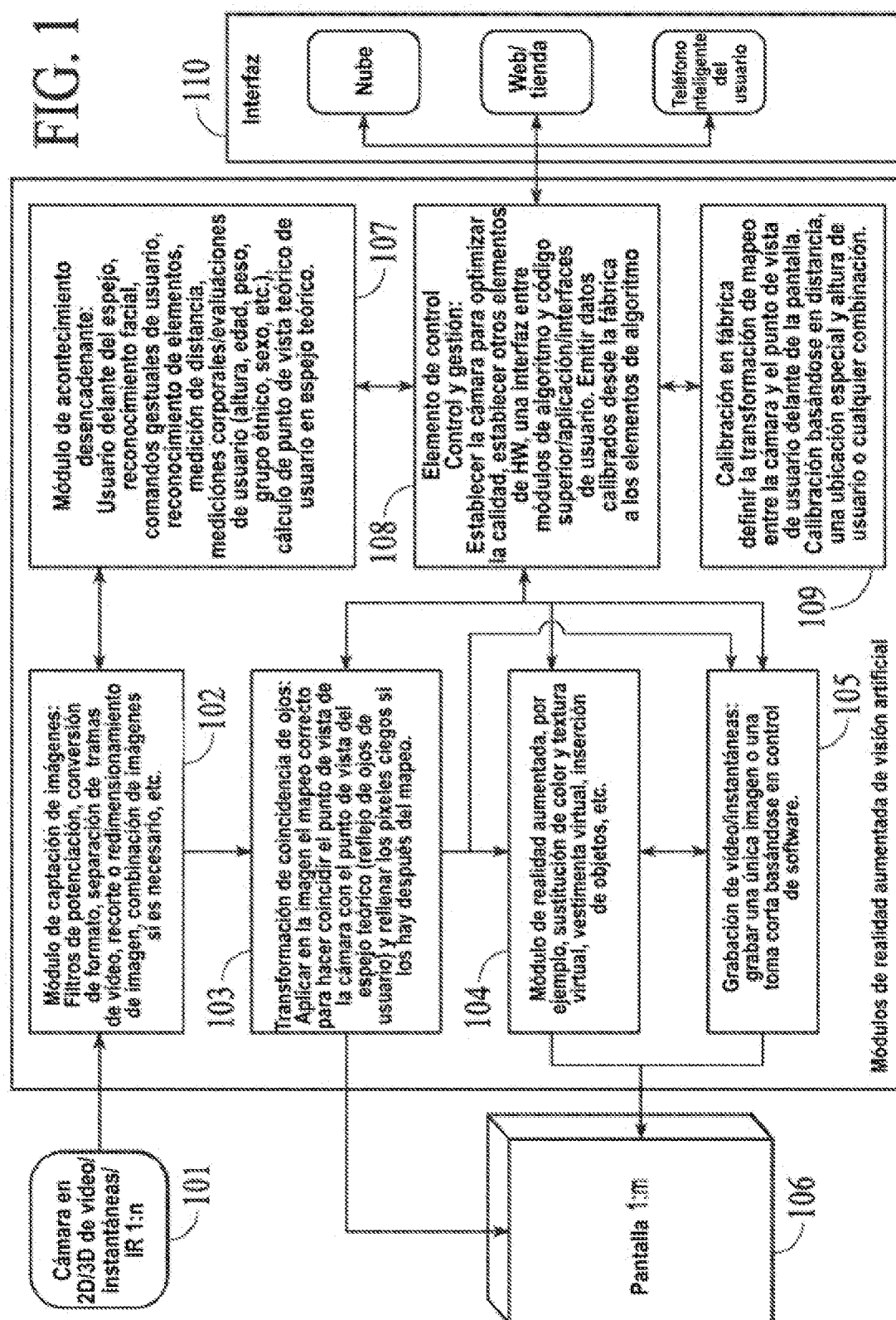
A la vista de los sistemas a modo de ejemplo descritos anteriormente, las metodologías que pueden implementarse según el objeto descrito se apreciarán mejor con referencia a los diagramas de flujo de las diversas figuras. Por simplicidad de la explicación, las metodologías se representan y describen como una serie de acciones. Sin embargo, las acciones según esta divulgación pueden realizarse en diversos órdenes y/o de manera simultánea, y con otras acciones no presentadas y descritas en el presente documento. Además, no todas las acciones ilustradas pueden requerir implementar las metodologías según el objeto divulgado. Además, los expertos en la técnica entenderán y apreciarán que las metodologías pueden representarse alternativamente como una serie de estados interrelacionados mediante un diagrama de estados o acontecimientos. Adicionalmente, debe apreciarse que las metodologías en esta memoria descriptiva son capaces de almacenarse en un artículo de fabricación para facilitar el transporte y la transferencia de tales metodologías a dispositivos informáticos. Se pretende que el término artículo de fabricación, tal como se usa en el presente documento, abarque un programa informático accesible desde cualquier medio de almacenamiento o dispositivo legible por ordenador, por ejemplo, almacenado en un servidor acoplado a la nube.

REIVINDICACIONES

1. Sistema informático programado para realizar un cambio de color en un objeto dentro de un vídeo en tiempo real,
que comprende una memoria que almacena un programa para la ejecución por un procesador, incluyendo el programa instrucciones para:
recibir una selección de objeto en el vídeo que va a volver a colorear;
utilizar una ubicación predictiva y una forma geométrica predictiva del objeto para seleccionar píxeles que van a ignorarse en la búsqueda de píxeles pertenecientes al objeto y reducir de ese modo el espacio de búsqueda de píxeles que deben incluirse en el objeto;
realizar una selección de elementos para obtener una colección de píxeles que pertenecen al objeto seleccionado;
realizar una comprobación de píxeles vecinos para confirmar la colección de píxeles;
agrupar píxeles pertenecientes al objeto usando información de tramas anteriores;
realizar una comprobación de continuidad de superficie para eliminar píxeles mal etiquetados debido a reflexiones o sombreados;
generar una pluralidad de máscaras, almacenando cada máscara diferentes valores de propiedad de la colección de píxeles, en el que la pluralidad de máscaras comprende: una máscara en blanco y negro que separa por segmentación el objeto, una máscara de escala de grises que almacena valores de escala de grises de píxeles pertenecientes al objeto, una máscara de color rojo que representa zonas rojas del objeto, una máscara de color verde que representa zonas verdes del objeto, una máscara de color azul que representa zonas azules del objeto, una máscara de brillo de mapa de intensidad de todos los píxeles pertenecientes al objeto, y una máscara de reflejo de luz que almacena valores de reflejo de píxeles pertenecientes al objeto;
seleccionar un nuevo color;
sustituir el color de objeto mediante manipulación de máscara, en el que la manipulación de máscara incluye sustituir una de la pluralidad de máscaras por una nueva máscara de color, y/o multiplicar o añadir una nueva máscara de color a la pluralidad de máscaras y/o sustraer y dividir una nueva máscara de color de la pluralidad de máscaras;
aplicar la máscara/máscaras manipulada/s sobre el vídeo original para generar un vídeo nuevo; y,
visualizar el nuevo vídeo.
2. Sistema informático según la reivindicación 1, en el que para encontrar todos los píxeles que pertenecen al objeto, se calcula la distancia euclidiana "dis" entre colores basándose en un modelo de distancia de color en la percepción humana del color, es decir, cromática, saturación y brillo.
3. Sistema informático según la reivindicación 1, que comprende además instrucciones para la calibración de tramas del vídeo y
realizar una transformación de distorsión en las tramas antes de generar la pluralidad de máscaras.
4. Sistema informático según la reivindicación 1, en el que el programa incluye instrucciones para la selección del elemento por un usuario.
5. Sistema informático según la reivindicación 1, en el que el programa incluye instrucciones para cambiar el aspecto del elemento modificando al menos uno de material textil, color, textura, reflejo, absorción, brillo, y tonos.
6. Sistema informático según la reivindicación 1, en el que el programa incluye instrucciones para su representación gráfica en un monitor de realidad aumentada, que comprende al menos uno de: cambio de fondo, vestimenta virtual del usuario; cambio de color del artículo en la imagen, y adición de accesorio.
7. Sistema informático según la reivindicación 1, en el que el programa incluye instrucciones para usar la imagen de infrarrojos para obtener la línea corporal de orientación del usuario.

8. Sistema informático según la reivindicación 1, en el que el programa incluye instrucciones para establecer puntos de registro en una imagen de un usuario.
- 5 9. Sistema informático según la reivindicación 8, en el que el programa incluye instrucciones para captar un objeto almacenado de una base de datos, y ajustar el objeto en la imagen del usuario estirando o contrayendo la imagen del objeto almacenado según los puntos de registro.
- 10 10. Sistema informático según la reivindicación 8, en el que el programa incluye instrucciones para modificar la imagen para representar gráficamente una imagen reflejada correspondiente a una imagen obtenida desde un punto de vista del usuario, para imitar de ese modo una imagen de un espejo.
11. Sistema informático según la reivindicación 8, en el que el programa incluye instrucciones para la detección de bordes
- 15 procesar, mediante una o más de las siguientes técnicas de acentuación de bordes: separación de colores, información de profundidad, detección de gradientes.
12. Sistema informático según la reivindicación 8, en el que el programa incluye instrucciones para crear la firma de medida del cuerpo o la parte del cuerpo.
- 20 13. Sistema informático según la reivindicación 8, en el que el programa incluye instrucciones para la calibración de la imagen basándose en la transformación proyectiva de una cámara usada para obtener el vídeo y basándose en una distancia de usuario a la cámara.
- 25 14. Sistema informático según la reivindicación 8, en el que el programa incluye instrucciones para promediar con el valor del factor de ponderación de los píxeles para combinar mejor el objeto con el color de fondo.

10



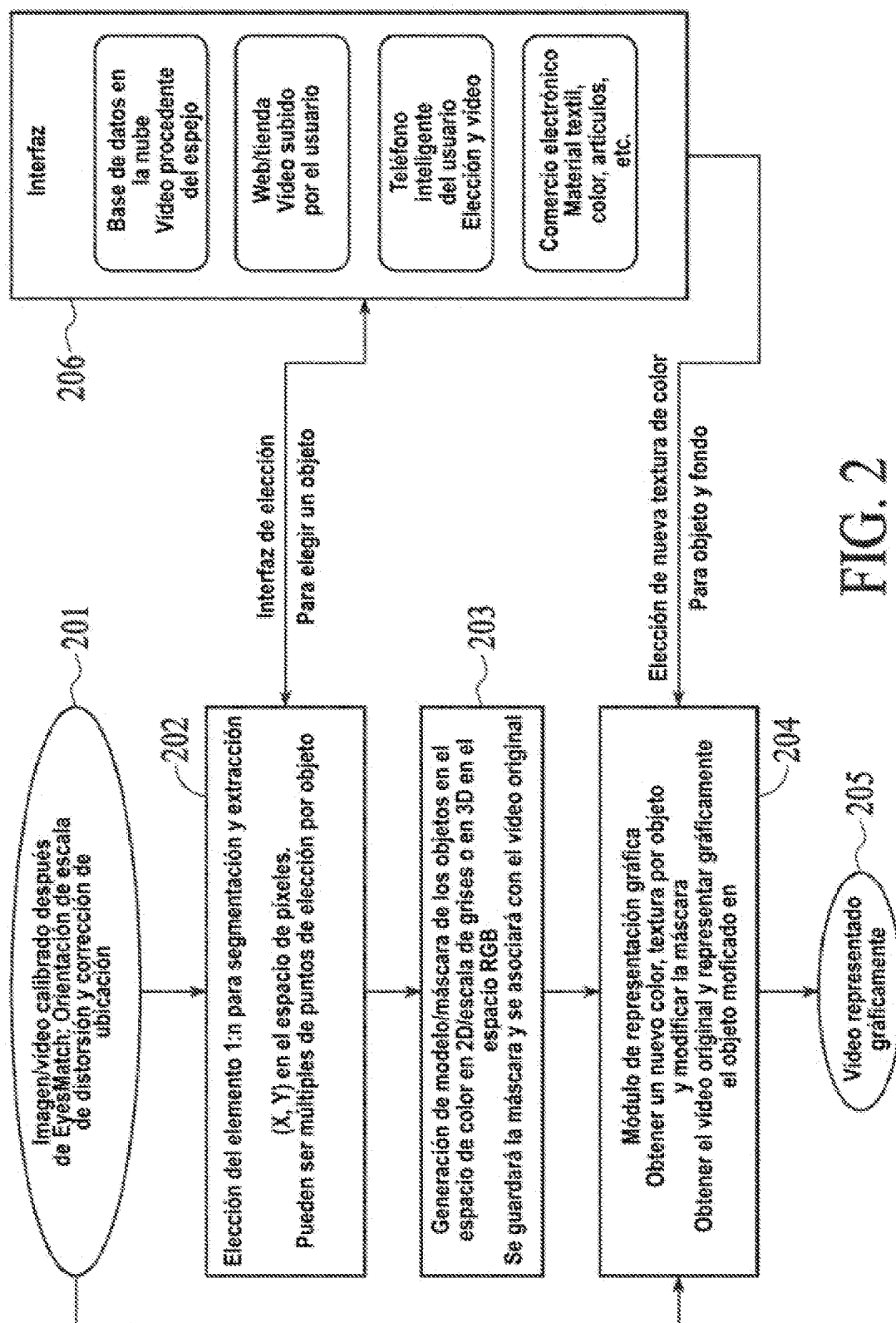


FIG. 2

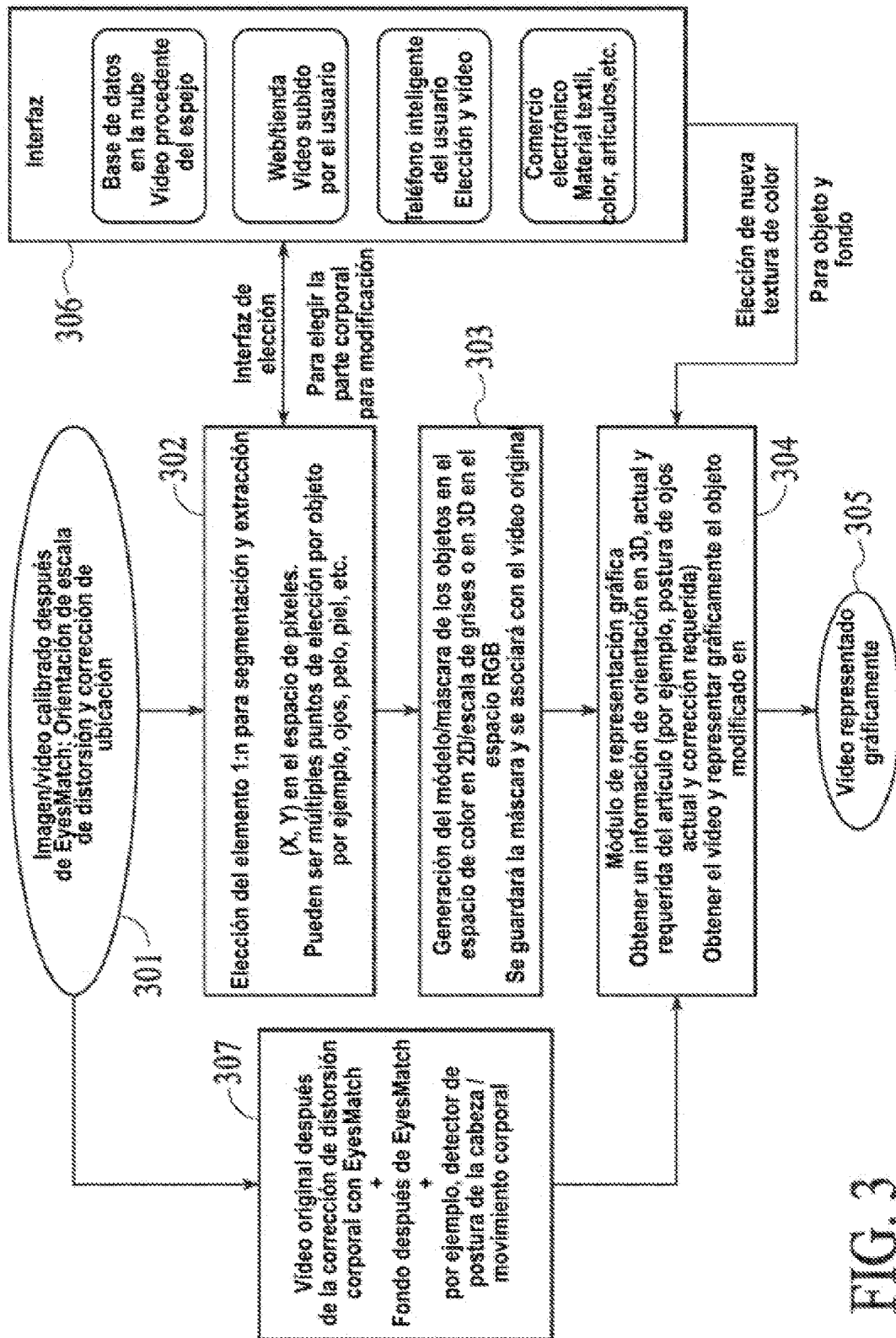


FIG. 3

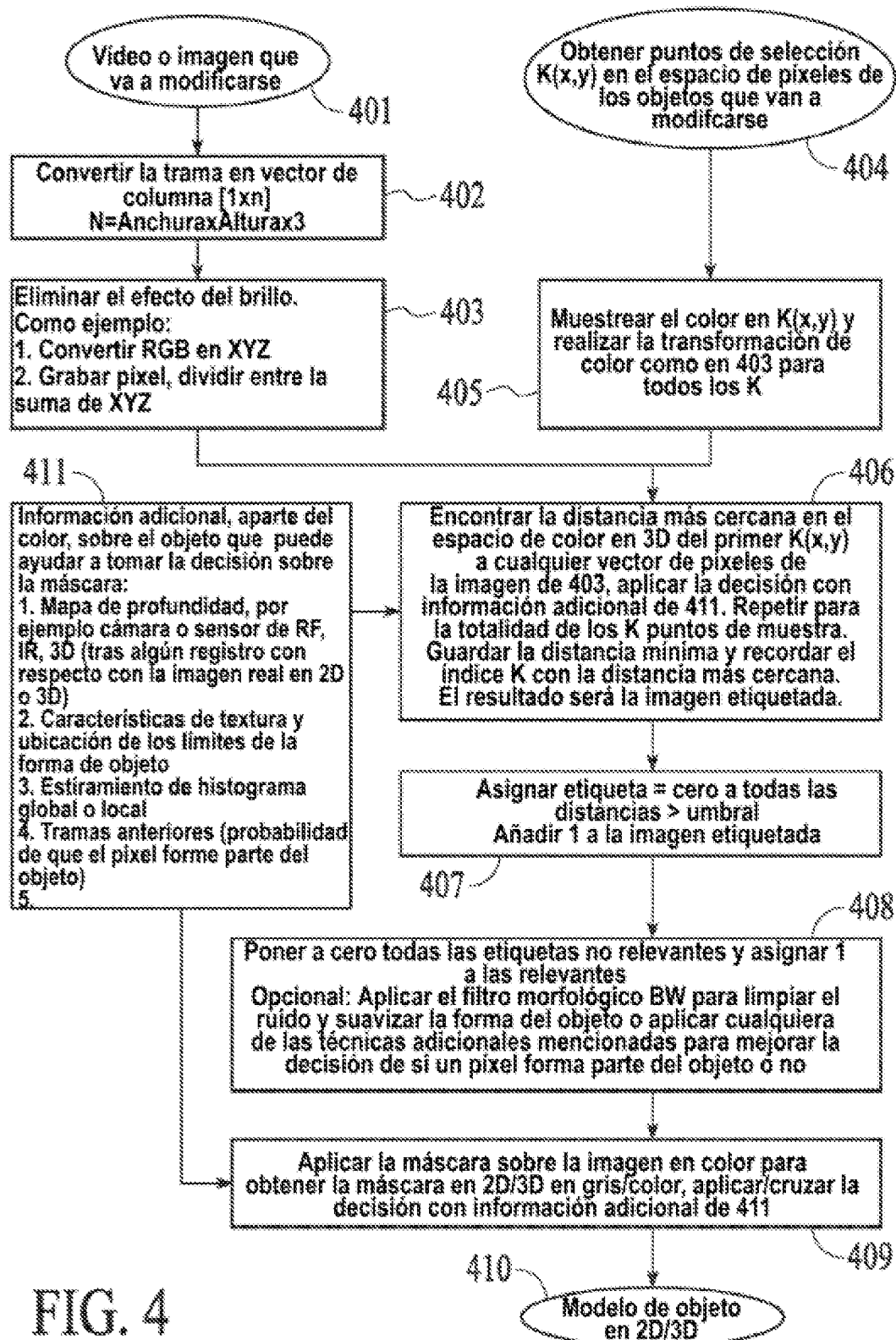
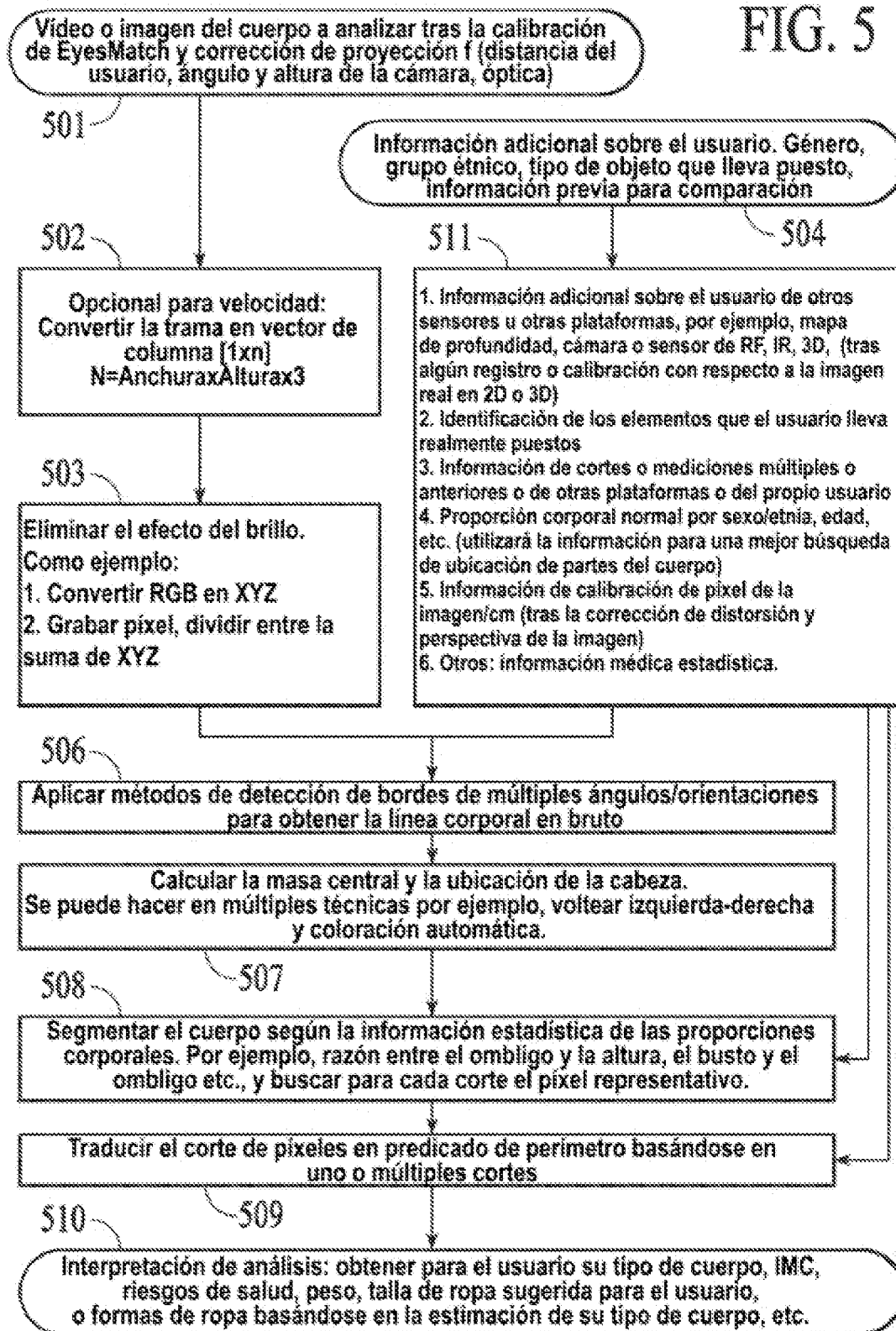


FIG. 5



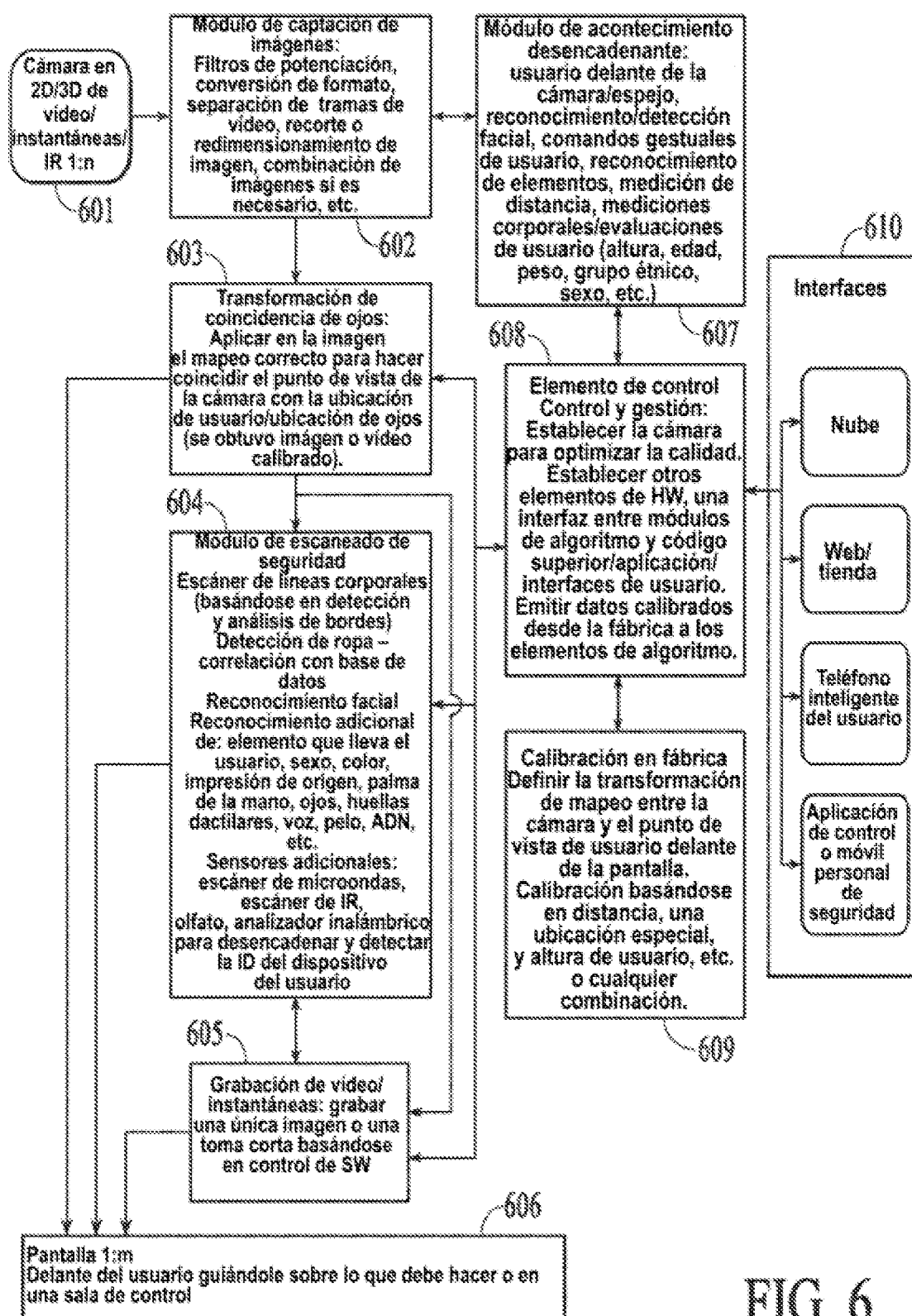
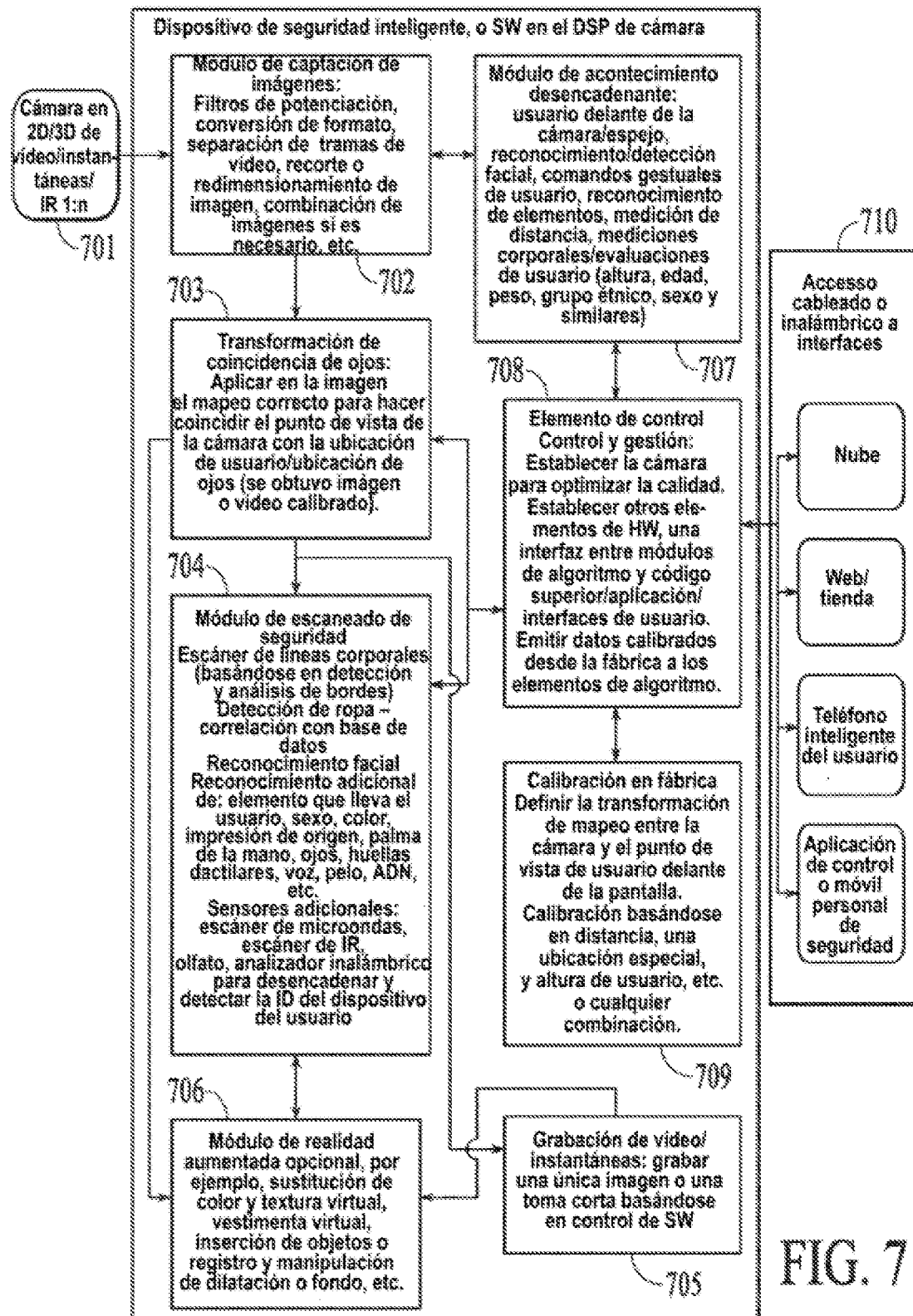


FIG. 6



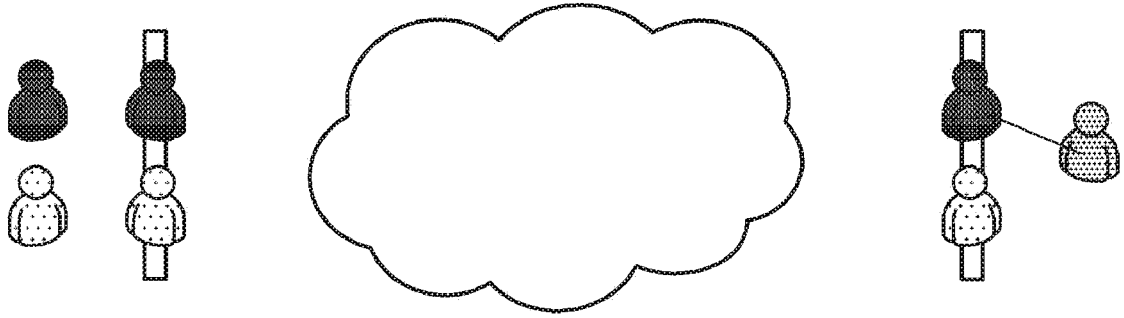


FIG. 8a

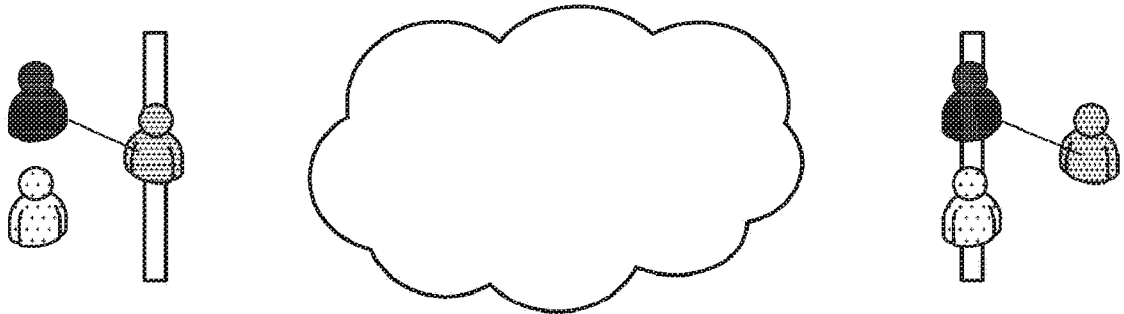


FIG. 8b

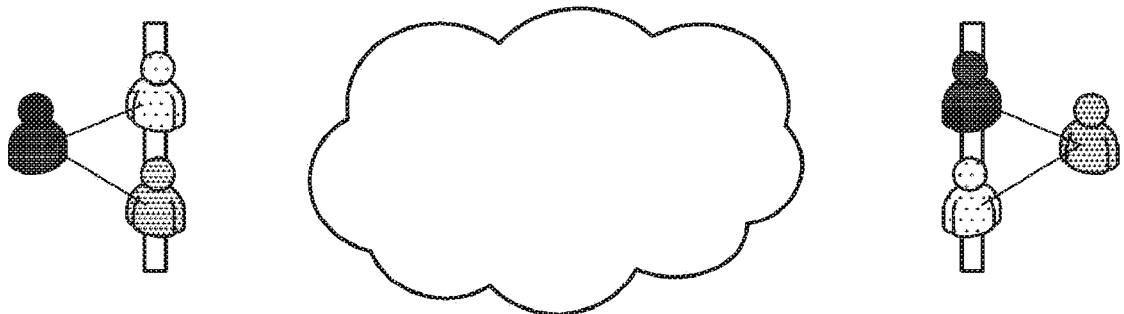
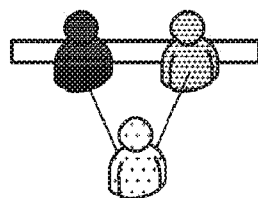


FIG. 8c



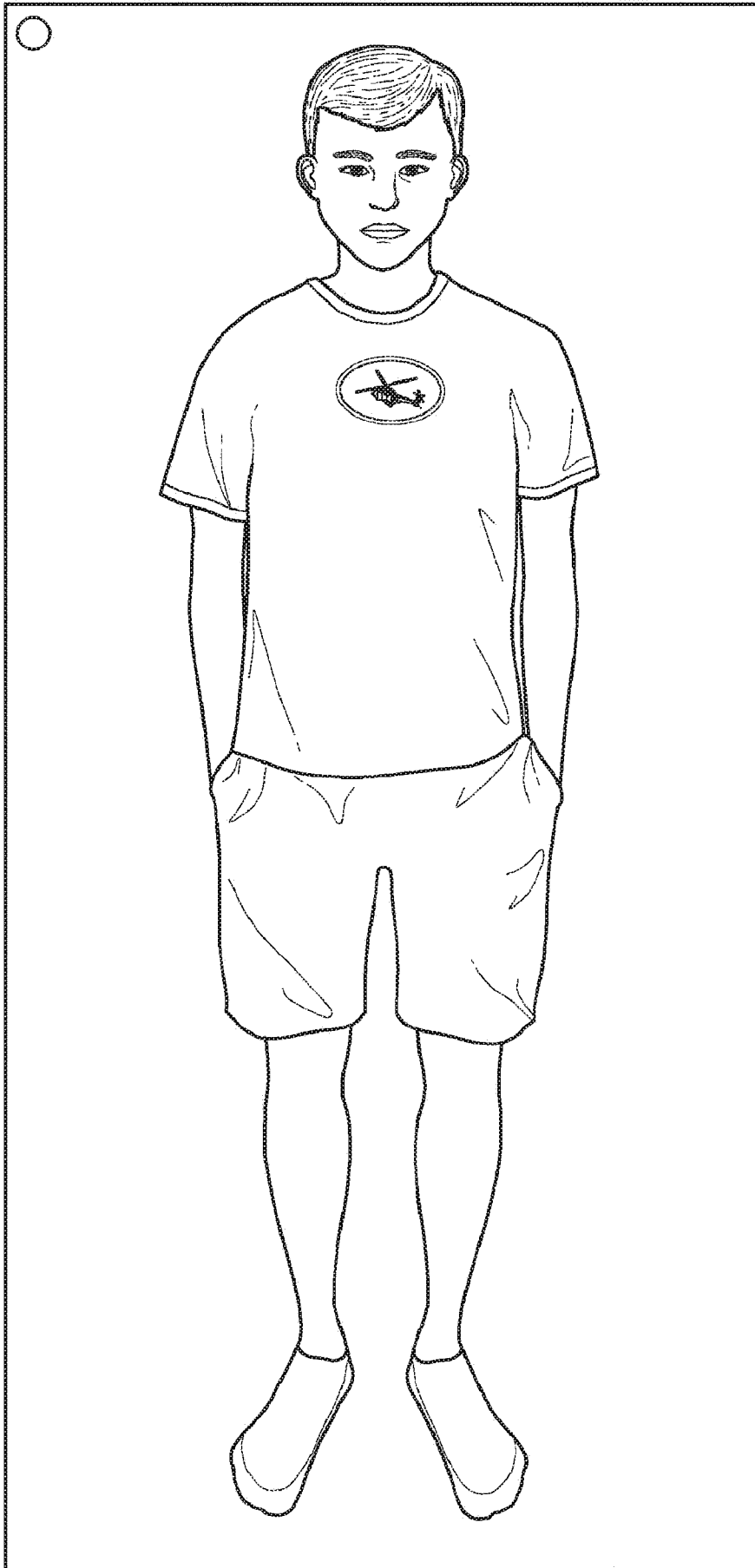


FIG. 9

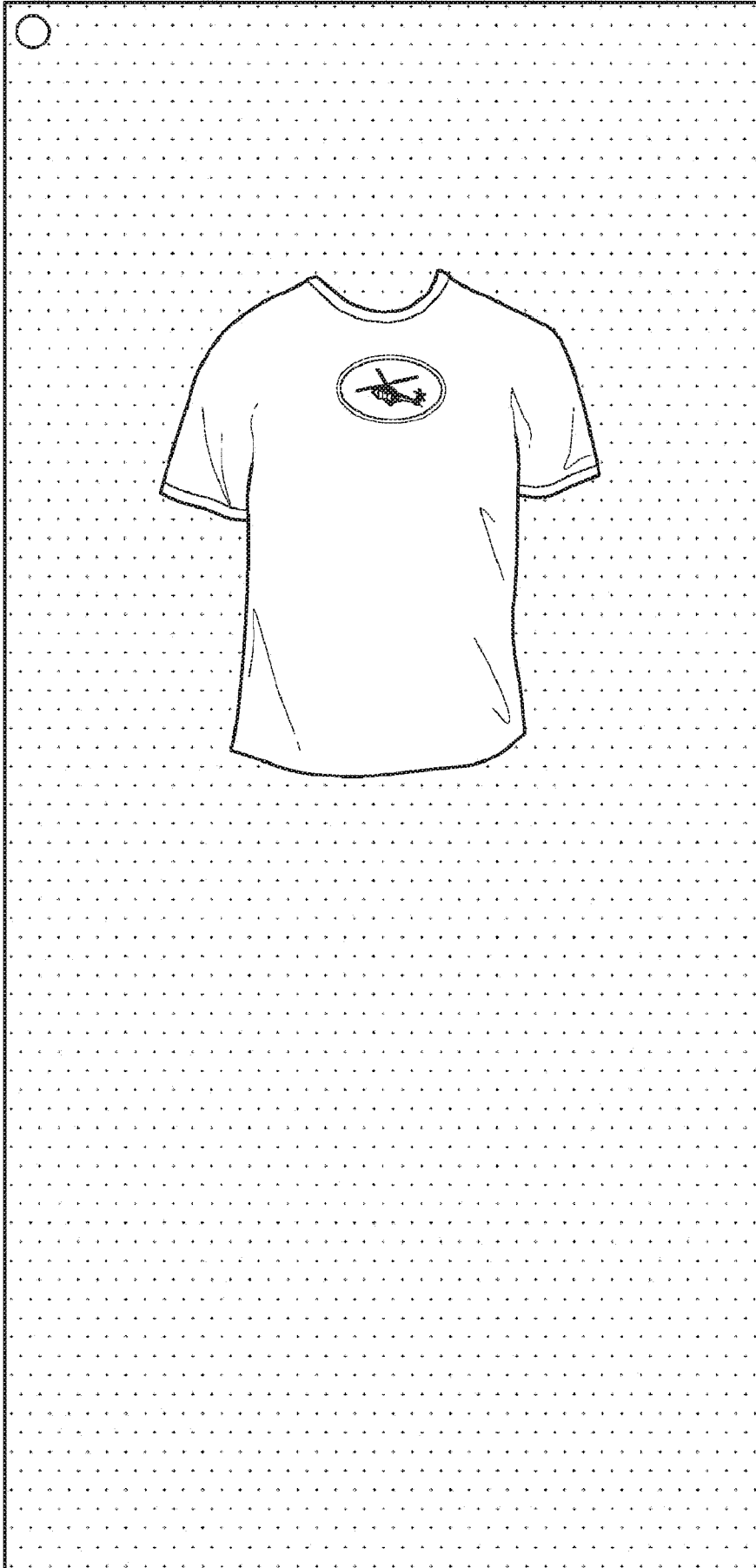


FIG. 10

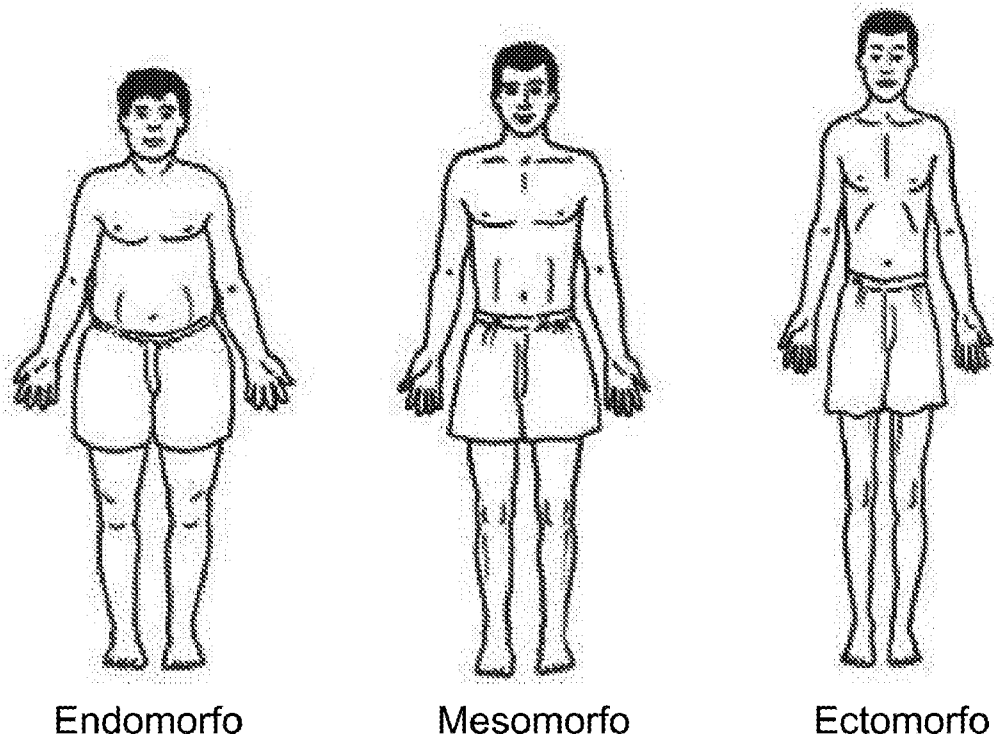


FIG. 11

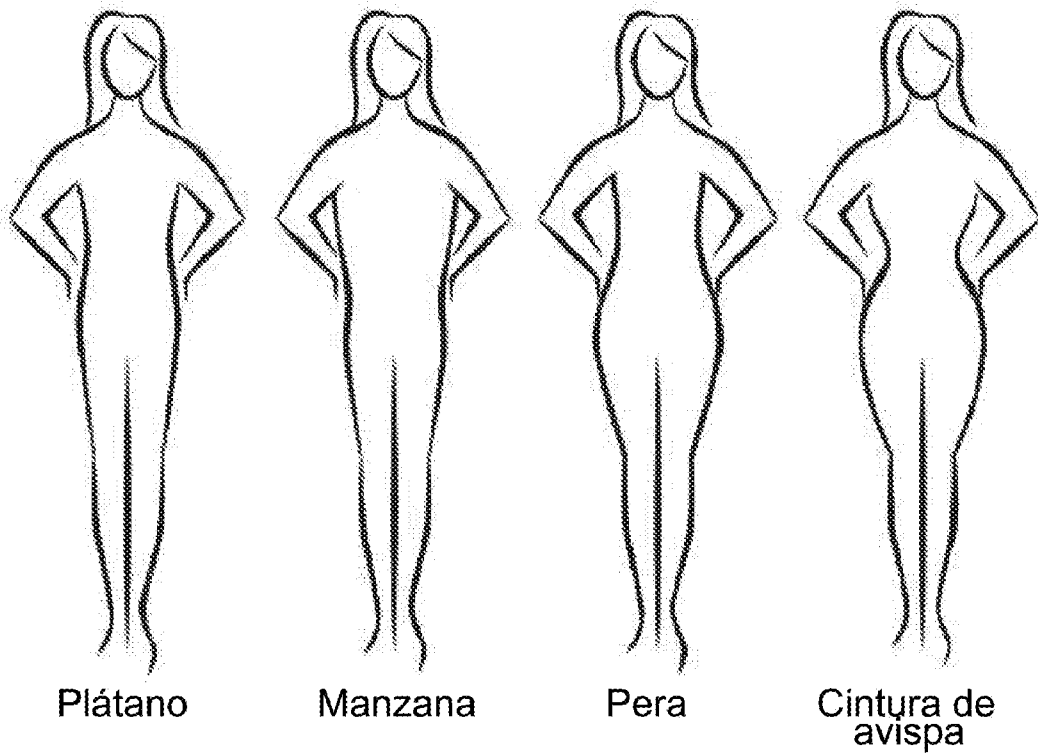


FIG. 12

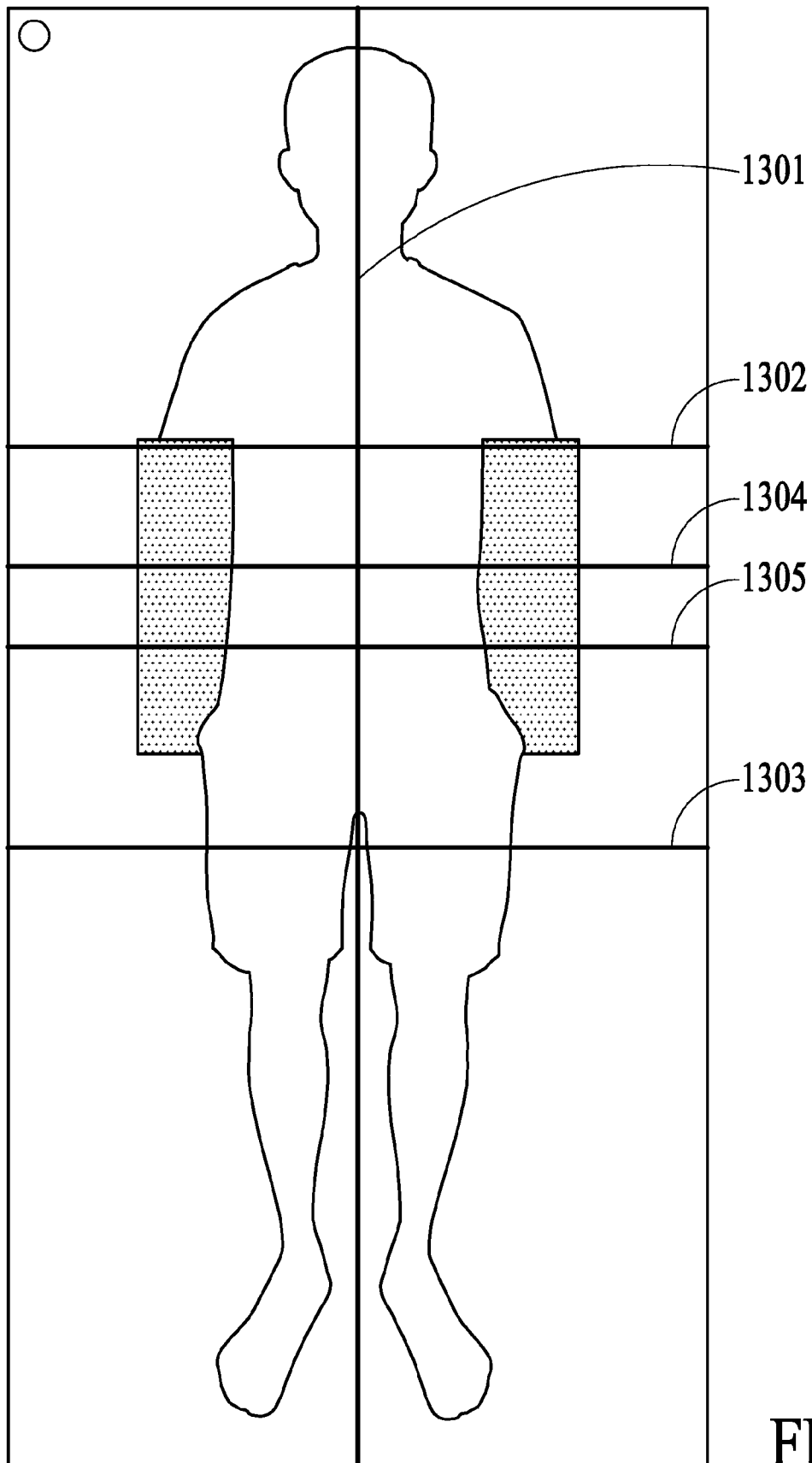


FIG. 13

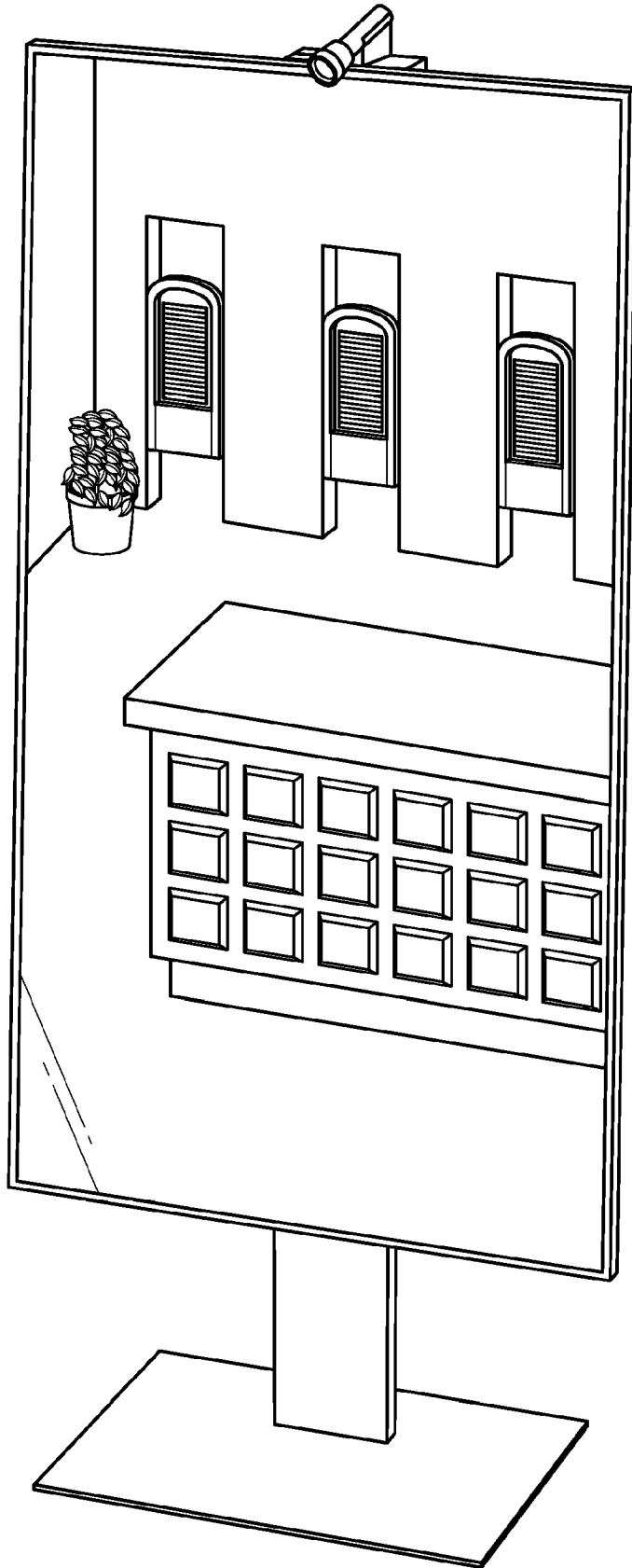


FIG. 14

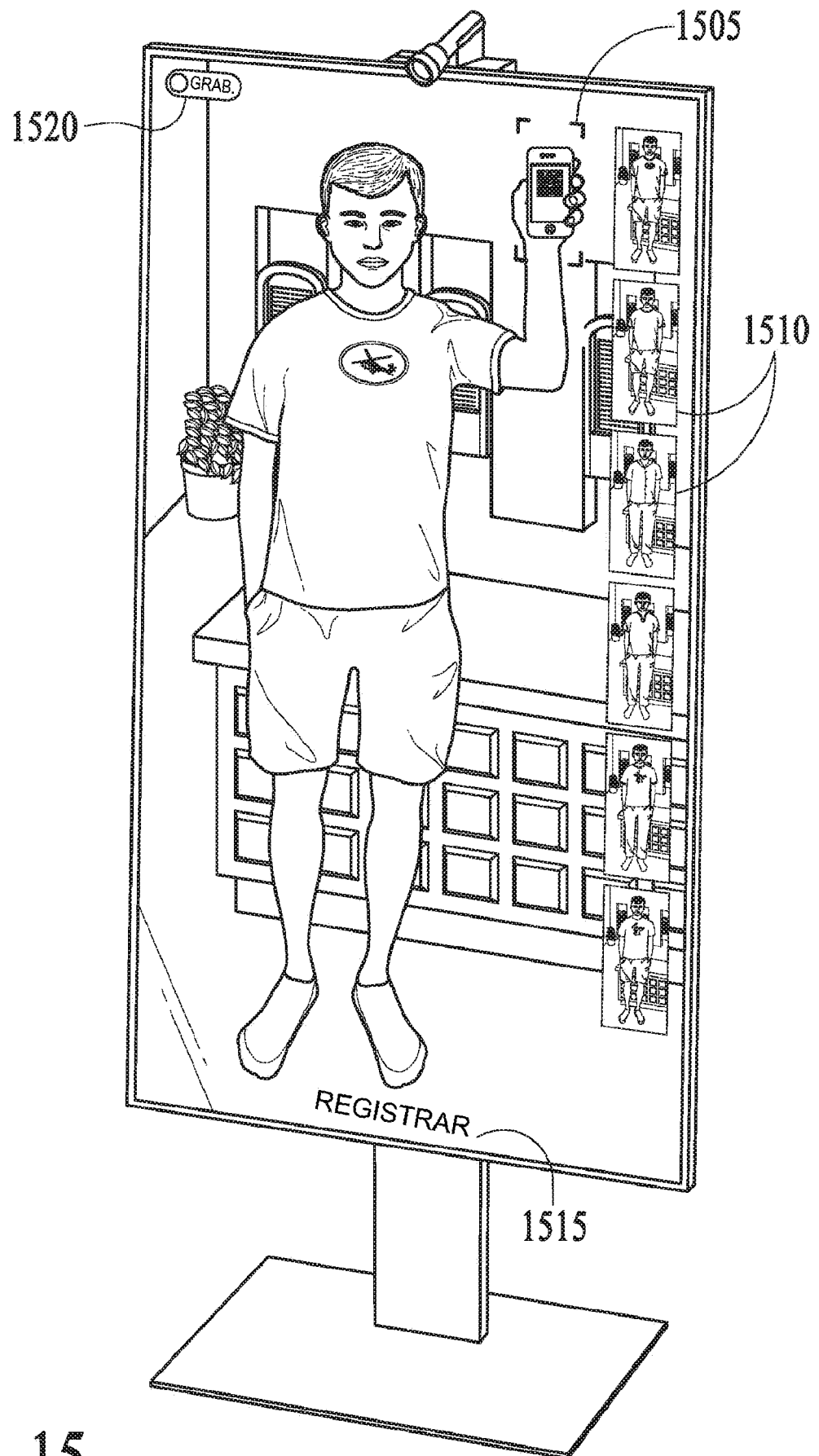


FIG. 15

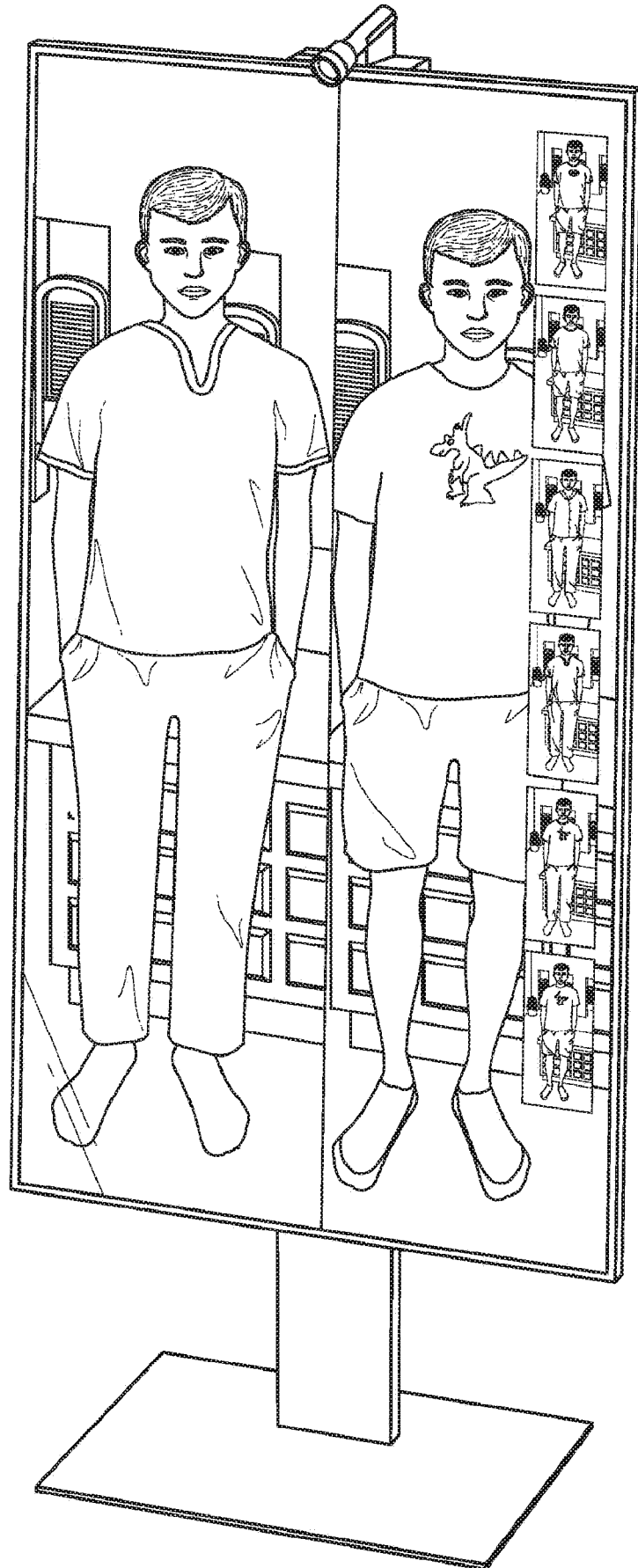


FIG. 16

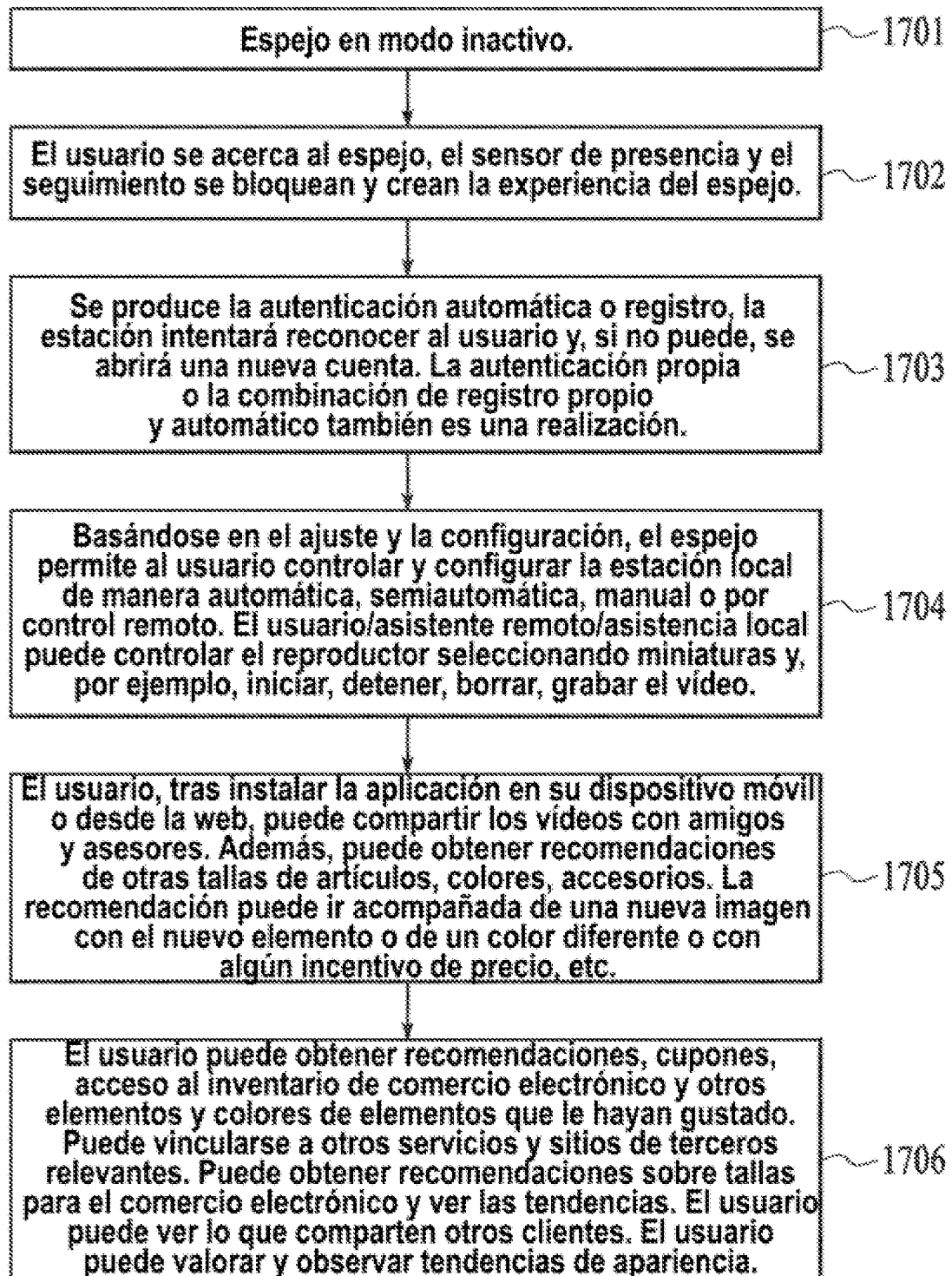


FIG. 17