



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



① Número de publicación: **2 334 079**

② Número de solicitud: 200700749

⑤ Int. Cl.:
G06T 11/60 (2006.01)
G06K 9/00 (2006.01)

⑫

SOLICITUD DE PATENTE

A1

② Fecha de presentación: **09.03.2007**

④ Fecha de publicación de la solicitud: **04.03.2010**

④ Fecha de publicación del folleto de la solicitud:
04.03.2010

⑦ Solicitante/s:
Universidad de Las Palmas de Gran Canaria
c/ Juan de Quesada, nº 30
35001 Las Palmas de G. Canaria, Las Palmas, ES

⑦ Inventor/es: **Afonso Suárez, María Dolores;**
Guerra Artal, Cayetano;
Castrillón Santana, Modesto;
Hernández Tejera, Mario;
Lorenzo Navarro, José Javier;
Déniz Suárez, Óscar;
Falcón Martel, Antonio;
Méndez Rodríguez, Juan;
Cabrera Gámez, Jorge;
Hernández Sosa, Daniel;
Isern González, Josep;
Domínguez Brito, Antonio Carlos y
Antón Canalis, Luis

⑦ Agente: **No consta**

⑤ Título: **Expositor virtual.**

⑤ Resumen:

Expositor virtual.

La presente invención hace referencia a un expositor virtual que permite la interacción hombre-máquina con el objetivo de que el usuario obtenga en tiempo real una imagen de sí mismo a la cuál se le pueda añadir de forma virtual complementos de carácter estético.

El expositor virtual está constituido por un soporte multimedia y multimodal que consta de una cámara de video digital de alta resolución, una computadora, un software de reconocimiento y tratamiento de imágenes y una pantalla táctil que permiten la realización de las tareas de detección y localización, en coordenadas de la imagen, de las características faciales del usuario, elección de artículos de la librería de imágenes y modificación del menú de interacción hombre-máquina. Por lo tanto, el principal problema que resuelve este dispositivo es el tiempo que se pierde a la hora de probarse distintos complementos hasta encontrar el adecuado.

La principal aplicación industrial de esta invención es la de promoción y venta de distintos complementos tanto para el hombre como para la mujer, en áreas comerciales y de pública concurrencia.



ES 2 334 079 A1

DESCRIPCIÓN

Expositor virtual.

5 Sector de la técnica

La invención se encuadra en el sector técnico de las tecnologías de la información y las telecomunicaciones, más concretamente en el área de la interacción hombre-máquina.

10 Estado de la técnica

El análisis de imágenes por medio de ordenadores es una ciencia que crece rápidamente. Se encuentra en pleno desarrollo, ya no ofrece solamente soluciones al área de la robótica o del control de calidad, continuamente surgen nuevas aplicaciones como por ejemplo el reconocimiento automático de facciones o del iris.

La detección de caras es un preproceso necesario para cualquier sistema de reconocimiento facial o de análisis de la expresión facial. Sin embargo, ha sido un problema poco tratado, o considerado menor dentro de los sistemas categóricos: reconocimiento y análisis de expresiones. Tal es así que distintos sistemas de reconocimiento asumen que la cara ha sido ya detectada antes de realizar la comparación con los modelos conocidos.

Recientemente el marco de detección descrito en [ViolaOl-cvpr] y ya integrado en la biblioteca OpenCV, permite trabajar a partir de rutinas veloces y efectivas de detección. En este marco se dispone de un detector para la detección de rostros mostrados en vista frontal y boca.

Los sistemas de tiempo real son sistemas informáticos que se encuentran en multitud de aplicaciones, desde la electrónica de consumo hasta el control de procesos industriales complejos. Están presentes en prácticamente todos los aspectos de nuestra sociedad, como teléfonos móviles, automóviles, control del tráfico, procesos automáticos de fabricación, etc... Además cada vez se fabrica un mayor número de máquinas que incluyen sistemas controlados por computador.

El sistema planteado, sobre el que se reivindica la patente, es un sistema de tiempo real empotrado, es decir, el sistema informático se encuentra físicamente incluido en un sistema de ingeniería más complejo. Este sistema presenta una solución basada en la integración de distintos dispositivos entre los cuáles destaca el uso de elementos que forman parte de un sistema de visión por computador y que incluye un módulo de detección de características faciales, un módulo de seguimiento de ojos y un módulo gráfico para la adición de los elementos digitales que forman el cuerpo de esta tecnología. La tecnología planteada permite la interacción hombre-máquina con el objetivo que el usuario obtenga en tiempo real una imagen de sí mismo a la cuál puede añadirle de forma virtual complementos de carácter estético.

40 Descripción detallada de la invención

La presente invención implementada por ordenador hace referencia a un expositor virtual conformado por una serie de dispositivos y por un software que detecta y localiza, en coordenadas de la imagen, las características faciales del usuario mediante técnicas basadas en visión por computador y realiza una adición automática de elementos sintéticos a la imagen del usuario.

El expositor virtual está constituido por un soporte multimedia y multimodal en el cuál el usuario puede probarse virtualmente determinados artículos o complementos. Para ello la imagen es captada por una cámara de vídeo de alta resolución y procesada por un ordenador mediante un software de tratamiento de imágenes.

El software procesa la imagen captada por la cámara mediante la detección y localización, en coordenadas de la imagen, de las características faciales del usuario mediante técnicas basadas en visión por computador. Una vez detectadas, se realiza una adición de elementos sintéticos a la imagen del usuario, de tal manera que este obtenga la ilusión óptica de llevar algún determinado tipo de artículo puesto. La imagen, una vez procesada, es mostrada en el monitor para ser vista en tiempo real por el usuario y alcanzar el efecto de espejo.

Dentro de la gama de artículos, pertenecientes a la librería y a la base de datos, que pueden ser mostrados digitalmente junto con la imagen del usuario, podemos encontrar gafas, pendientes, collares, maquillaje, etc...

60 Modo de realización de la invención

La adquisición de las imágenes comienza con la digitalización de las caras a procesar mediante una cámara de vídeo digital de alta resolución (1), como se muestra en la figura 2. Estas imágenes son transferidas a la computadora (2) en el tamaño y frecuencias adecuadas.

Mediante un software instalado en la computadora se lleva a cabo el procesamiento de las imágenes. Este software realiza la detección de las diferentes características faciales necesarias para la detección precisa de los ojos.

La aproximación basada en visión por computador normalmente divide la detección de los ojos en dos tareas, en primer lugar se localiza una ventana con los ojos, es decir una pequeña área de la imagen donde probablemente se encuentran los ojos. Después de esto se aplican una o más técnicas de aproximación. Nuestra propuesta hace uso de un sistema de detección de la cara que proporciona en tiempo real múltiples detecciones a diferentes resoluciones. Una vez que la cara es detectada, se localizan los ojos. El detector de caras muestra dos etapas, la primera se centra en la búsqueda de la cara y la segunda, una vez que esta ha sido detectada, realiza su seguimiento.

Al comienzo de una sesión de interacción, cuando no hay individuos en el campo de visión, o no se ha detectado presencia en un intervalo de tiempo, la aproximación hace uso de dos detectores de cambio de ventanas que se basan en la detección del marco del objeto. Estos dos detectores, integrados en la versión OpenCV, son los detectores de la vista frontal de la cara y del contexto local basado en el detector de la cara. El último logra mejores ratios de reconocimiento para imágenes con bajas resoluciones, si la cabeza y los hombros están visibles. El tamaño mínimo buscado es de 24x24 y 20x20 pixels. Para no malgastar tiempo de proceso, los detectores se ejecutan alternativamente.

Para cualquier cara que se haya detectado, el sistema intenta detectar los ojos asumiendo que es una vista frontal de la misma, y por tanto su localización verificaría alguna restricción de apariencia y geométrica. Las tareas que se realizan para localizarlos son:

- *Detección de la piel.* Una vez detectada la cara, se modela su color de piel utilizando espacio de color normalizado rojo-verde, considerando justo el centro de la cara que haya sido proporcionado por cualquiera de los detectores basados en los de Viola-Jones. El sistema, de forma heurística, elimina los elementos que no son parte de la cara, por ejemplo el cuello, y le hace corresponder una elipse, para posteriormente poder rotar verticalmente su posición.
- *Localización de los ojos.* En este punto la aproximación busca candidatos a ojos en las áreas en las que hay más probabilidad de encontrarlos dentro de la zona de la cara, considerando que la misma presenta una vista frontal. Se testean varios pares de candidatos de acuerdo a su apariencia y se elige uno de ellos. Las tareas utilizadas para esto son:
 - a) *Áreas negras:* Los ojos son más negros que el resto de piel que se encuentra a su alrededor.
 - b) *Detectores de ojos basados en los detectores de Viola-Jones:* Como la posición de los ojos puede ser más o menos estimada y por tanto limitada, un detector de ojos basado en el de Viola-Jones ofrece resultados rápidos. El detector busca ojos con un tamaño mínimo de 16x12 pixels. Para caras pequeñas, se realiza un aumento a escala antes de realizar la búsqueda.
 - c) *Detectores de pares de ojos basados en los detectores de Viola-Jones:* Si alguna de las anteriores falla, el par de ojos detectados puede dar otra aproximación para la posición de los mismos, y así luego aplicar los pasos a) y b) de nuevo. El tamaño mínimo del patrón buscado es 22x5 pixels.

Cada cara detectada muestra distintas características $X_i = \langle \text{pos, size, color, eyespos, eyespattern, facepattern} \rangle$. Las normas se aplican de acuerdo a estas características y en un orden que tiene en cuenta el coste de cómputo y la fiabilidad. Las normas consideradas son:

- *Seguimiento de ojos:* Un algoritmo de seguimiento de ojos rápido se aplica en el área previamente detectada que los rodea.
- *Detector de cara:* Se aplica un detector de cara de Viola-Jones en un área que cubre la detección anterior.
- *Detector del contexto local de la cara:* Si las técnicas anteriores fallan, el detector de contexto local se aplica en un área que incluye la detección previa.
- *Color de la piel:* Se busca el color de la piel en la ventana de búsqueda que contiene la detección previa, y se testean los nuevos tamaños y posiciones.
- *Seguimiento de la cara:* Si cualquiera de las anteriores falla, el patrón de cara previamente almacenado se busca en un área que cubre la detección previa.

Estas técnicas se aplican siguiendo este orden hasta que una de ellas encuentra la cara que se buscaba. Cuando se detecta la cara, el color de la piel se utiliza para la detección de los ojos, de forma similar al procedimiento ya descrito.

Una vez localizados los ojos se lleva a cabo el seguimiento de estos para mantener sus coordenadas en la imagen continuamente ubicadas, según se muestra en la figura 3.

Por último se realiza una superposición de la imagen digital de un artículo sobre la imagen de su cara, como muestra la figura 1, para lograr la sensación visual de llevarlo puesto.

ES 2 334 079 A1

Mediante un monitor se muestra, al usuario, la imagen ya procesada que será reflejada 90 grados una vez incida sobre el cristal espejo para su correcta visualización.

5 El cristal espejo es un elemento, perteneciente al expositor virtual, que permite lograr el efecto de espejo desde la perspectiva del usuario. Según se muestra en la figura 2 la parte identificada como (4a) se corresponde con la parte del cristal, y la parte designada como (4b) se corresponde con la parte del espejo.

10 Desde su posición, tal y como se muestra en la figura 2, el usuario podrá visualizar su imagen reflejada con el artículo seleccionado superpuesto de forma virtual. Además a través de una pantalla táctil el usuario podrá seleccionar los diferentes modelos de artículos que pertenecen a la librería.

15 Se hace uso de una librería de mezcla multimedia (MML), de desarrollo propio, implementada en DirectShow. DirectShow es una API (Application Programming Interface - Interfaz de programación de aplicaciones) estándar desarrollado por Microsoft para el almacenamiento, codificación y traducción de contenidos multimedia como vídeo y audio que se toman como origen. Se realiza mediante el uso de una estructura de filtros interconectados, que se clasifican en tres categorías: fuente u origen, transformación y filtros de traducción. La librería multimedia implementada está desarrollada con Microsoft DirectShow ampliándolo para ofrecer un solapamiento de diferentes capas y cambiar su orden de acuerdo a algún tipo de prioridad en tiempo de ejecución. Cada capa puede ser una imagen, vídeo previamente almacenado, vídeo capturado on-line desde una cámara, o modelos de animación en 3D. En la figura 4 se muestra un gráfico con los filtros del DirectShow.

20 La arquitectura del software está diseñada como un gráfico más de los filtros del DirectShow, donde el filtro del vídeo fuente corresponde a la imagen de entrada del usuario. La detección de las características de la cara y las rutinas de seguimiento también se encuentran encapsuladas dentro del filtro que proporciona la posición de los ojos en coordenadas (x,y). Con esta localización el filtro MML fija o encaja el artículo seleccionado por el usuario, mediante la pantalla táctil, en el lugar correcto de la imagen.

Descripción de las figuras

30 En la figura 1 se ilustra cómo la imagen de uno de los artículos de la librería es superpuesto a la imagen real del usuario.

35 En la figura 2 se representa el esquema funcional de la invención constituido por un soporte multimedia y multimodal, con una cámara de vídeo digital de alta resolución, una computadora, un software de reconocimiento y tratamiento de imágenes, una pantalla táctil y un cristal espejo; que permiten la realización de las tareas de identificación de caracteres faciales, elección de artículos de la librería de imágenes, modificación del menú para interacción entre el usuario y la computadora y procesado de imágenes.

40 En la figura 3 se describe el procedimiento de seguimiento de los ojos para mantener sus coordenadas en la imagen continuamente ubicadas.

45 Para describir el procedimiento propuesto denotaremos p como un punto que corresponde a la representación visual de un objeto en ϕ en un tiempo t , por ejemplo: p corresponde a la plantilla T en el espacio definido por $R^{m \times n}$. Se establece una distancia d en ϕ . La distancia d entre dos puntos p_1 y p_2 se calcula de acuerdo a la norma L_2 . Esta distancia se utilizará en la imagen de entrada y la plantilla para obtener la mejor correspondencia posible.

Después de aplicar la función distancia entre la imagen y el patrón deslizando la plantilla sobre la ventana de búsqueda, aparecerá un número variable de mínimos locales, entre ellos el que nos interesa.

50 En ϕ el vector p corresponde al patrón de referencia, esto es, la vista del objeto que nos interesa. El vector m_1 corresponderá al mínimo absoluto considerando que es el objeto visual que más se asemeja al objeto de interés. La existencia de más mínimos locales, m_2 y m_3 , implica que hay objetos que se asemejan en cierto grado al objeto de interés. Son los llamados objetos de contexto. Estos objetos, de la misma manera que el objeto de interés, también tienen su propia curva de transformación visual incluida en su copia en el espacio próximo. Aunque por simplicidad estos objetos de contexto permanecerán estáticos.

55 La curva de transformación visual del objeto que centra nuestro interés es el lugar geométrico de los puntos correspondientes a la diferencia mínima después del proceso de correspondencia en imágenes de entrada, en un tiempo determinado. Esta curva se compondrá de los vectores más cercanos m_1, m'_1, m''_1, \dots al patrón de referencia p . Por tanto m_1 corresponde al punto más cercano a p en el momento $t=0$, m'_1 corresponde al punto más cercano a p en $t=1$, etc. Sin embargo si existiera al menos un objeto en el contexto, m_2 , y el patrón de referencia p no se actualizara, esto podría ocurrir después de un determinado número de encuadres, el mínimo absoluto no se correspondería con el objeto real de interés sino con el objeto que más se asemeja del contexto, tal y como muestra la figura 3. De este modo el área de la ventana de búsqueda correspondiente al punto m_2 se tomará como el objeto de interés, dando como resultado un error en el proceso de seguimiento, este es un error muy común de técnicas de actualización que no actualizan el patrón. El origen del problema está en la falta de actualización o actualización con un coeficiente inapropiado del patrón de referencia.

ES 2 334 079 A1

En la figura 4 se muestra un gráfico con los filtros del DirectShow.

Aplicación industrial

5 La principal aplicación industrial de esta invención es la de promoción y venta de distintos complementos tanto para el hombre como para la mujer, en áreas comerciales y de pública concurrencia.

10

15

20

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

5 1. Expositor virtual para servicios en áreas comerciales y con fines de atracción al público en general. Constituido por un soporte multimedia y multimodal, con una cámara de vídeo digital de alta resolución, una computadora, un software de procesado de imágenes y gestión y una pantalla táctil, **caracterizado** porque muestra, en tiempo real, la imagen original del usuario a la que se le han añadido virtualmente y de forma realista artículos elegidos por él mismo desde el menú como: gafas, pendientes, collares, maquillajes, y cualquier clase de complemento estético sobre el busto del usuario.

10 2. Expositor virtual según reivindicación 1 en el cuál el software, de desarrollo propio, cargado y ejecutado en la computadora es el encargado del procesamiento de la imagen del usuario e integrarle a la misma el elemento seleccionado.

15 3. Expositor virtual según reivindicación 1 en el cuál la librería de mezcla multimedia, de desarrollo propio, permite la superposición de diferentes capas y cambia su orden de acuerdo a algún tipo de prioridad en tiempo de ejecución. Cada capa puede ser una imagen, vídeo previamente almacenado, vídeo capturado on-line desde una cámara o modelos de animación en 3D.

20

25

30

35

40

45

50

55

60

65

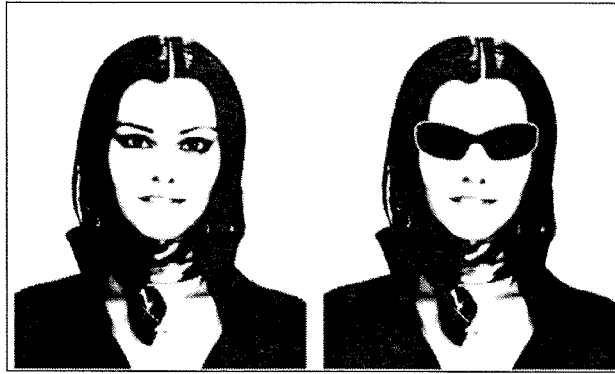


Figura 1

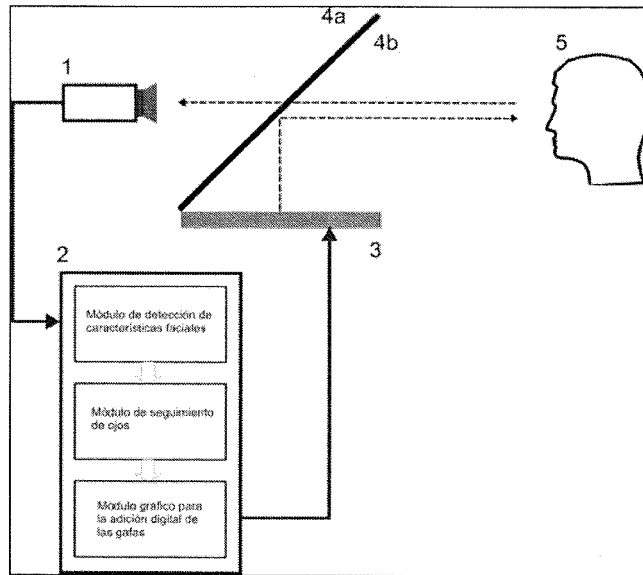


Figura 2

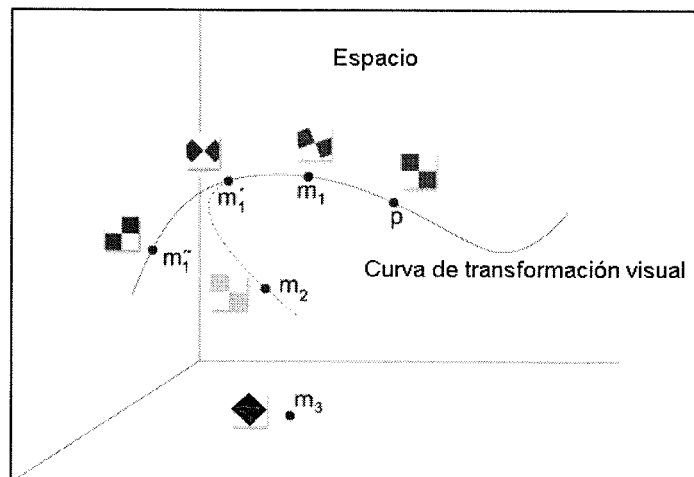


Figura 3

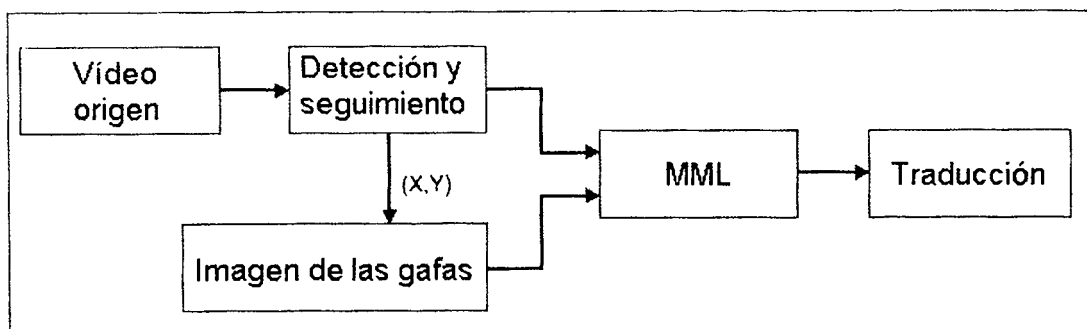


Figura 4



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

① ES 2 334 079

② N° de solicitud: 200700749

③ Fecha de presentación de la solicitud: **09.03.2007**

④ Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑤ Int. Cl.: **G06T 11/60** (2006.01)
G06K 9/00 (2006.01)

DOCUMENTOS RELEVANTES

Categoría	⑥ Documentos citados	Reivindicaciones afectadas
A	EP 1447973 A1 (OMRON TATEISI ELECTRONICS CO) 18.08.2004	1
A	US 2003174869 A1 (SUAREZ et al.) 18.09.2003	1
A	EP 1215618 A2 (EASTMAN KODAK CO) 19.06.2002	1
A	US 2005141771 A1 (YAMAKADO et al.) 30.06.2005	1

Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

El presente informe ha sido realizado

para todas las reivindicaciones

para las reivindicaciones nº:

Fecha de realización del informe

17.02.2010

Examinador

Mª C. González Vasserot

Página

1/4

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

G06T, G06K

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, WPI

Fecha de Realización de la Opinión Escrita: 17.02.2010

Declaración

Novedad (Art. 6.1 LP 11/1986)	Reivindicaciones	1-3	SÍ
	Reivindicaciones		NO
Actividad inventiva (Art. 8.1 LP 11/1986)	Reivindicaciones	1-3	SÍ
	Reivindicaciones		NO

Se considera que la solicitud cumple con el requisito de **aplicación industrial**. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

Base de la Opinión:

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como ha sido publicada.

1. Documentos considerados:

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Número Publicación o Identificación	Fecha Publicación
D01	EP 1447973 A1	18-08-2004
D02	US 2003174869 A1	18-09-2003
D03	EP 1215618 A2	19-06-2002
D04	US 2005141771 A1	30-06-2005

2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración

Los documentos citados sólo muestran el estado general de la técnica y no se consideran de particular relevancia.

Ninguno de los documentos anteriores muestra un procedimiento que detecta y localiza las características faciales del usuario mediante visión artificial y realiza una adición automática de elementos sintéticos a la imagen del usuario como se reivindica en las reivindicaciones 1-2.

Además, no se considera obvio que un experto en la materia conciba dicho procedimiento reivindicado en estas reivindicaciones.

El documento D1 se considera el más cercano al Estado de la Técnica, se diferencia de la solicitud en que no superpone la imagen digital de un artículo sobre la imagen de su cara como en la invención.

Por lo tanto, la invención reivindicada es nueva e implica actividad inventiva.