

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 398 032**

21 Número de solicitud: 200901210

51 Int. Cl.:

**G06T 19/00** (2011.01)

**G06T 7/20** (2006.01)

**G06K 9/00** (2006.01)

12

SOLICITUD DE PATENTE

A1

22 Fecha de presentación:

**07.05.2009**

43 Fecha de publicación de la solicitud:

**13.03.2013**

71 Solicitantes:

**UNIVERSIDAD DE LA LAGUNA (100.0%)  
OTRI. EDIFICIO CENTRAL, C/ DELGADO  
BARRETO S/N  
38201 LA LAGUNA (Tenerife) ES**

72 Inventor/es:

**GONZÁLEZ GONZÁLEZ, Carina Soledad;  
CABRERA PRIMO, David;  
RIVERO GUTIÉRREZ, Melvin y  
SIGUT SAAVEDRA, José Francisco**

54 Título: **SISTEMA QUE PERMITE LA APLICACIÓN DE EFECTOS VISUALES EN OBJETOS O SUJETOS ESTÁTICOS O EN MOVIMIENTO EN TIEMPO REAL SIN NECESIDAD DE MARCAS.**

57 Resumen:

La presente invención se basa en un sistema que permite la aplicación de efectos visuales en objetos o sujetos estáticos o en movimiento en tiempo real sin necesidad de marcas logrando una interacción natural entre el objeto o el sujeto y los dispositivos de entrada y de salida para lo que se utilizan técnicas de visión por computador.

El sistema comprende un dispositivo de entrada que permita la captura de imágenes, en cualquier formato y en tiempo real, un escenario configurado de tal forma que la iluminación y el fondo estén controlados, un módulo de procesamiento de imágenes en tiempo real y detección de movimiento a través de la implementación de filtros y un dispositivo de salida donde se proyectan las imágenes procesadas con el efecto visual correspondiente.

ES 2 398 032 A1

**DESCRIPCION**

**Sistema que permite la aplicación de efectos visuales en objetos o sujetos estáticos o en movimiento en tiempo real sin necesidad de marcas**

5 **Sector de la técnica**

Implementaciones en ordenador. Visión por computador.

**Introducción**

10 El presente documento describe un sistema de interacción entre objetos o sujetos y determinados efectos visuales preprogramados que son proyectados en el propio objeto o sujeto simulando el funcionamiento de un espejo. El usuario puede apreciar en tiempo real dichos efectos visuales proyectados sobre cuerpos estáticos o en movimiento.

15 Este sistema está compuesto por un escenario con luces y fondo controlado, sensores de entrada como puede ser una cámara de vídeo, una CPU, como puede ser un ordenador, una serie de filtros de detección y proyección basados en visión por computador, así como un dispositivo de salida que puede ser un proyector, una televisión, una pantalla de ordenador, etc.

20 La ventaja fundamental del sistema se basa en la detección del movimiento del sujeto u objeto sin la necesidad de marcas como colores especiales, sensores, etc. en dicho cuerpo, ni de condiciones especiales de entrenamiento en comportamientos o en movimientos predeterminados.

25 Tanto la detección del movimiento como la proyección de determinados efectos visuales sobre un sujeto u objeto se realizan mediante una serie de filtros implementados en computador utilizando técnicas de visión por computador y análisis y tratamiento de imágenes en tiempo real. La solución presentada trabaja con estos filtros identificando zonas de análisis y zonas de aplicación de dichos efectos visuales. Asimismo, el sistema permite el reconocimiento de acciones o eventos predefinidos y la aplicación de filtros diseñados y relacionados a los mismos.

30

**Estado de la técnica**

Existen dispositivos y procesos para la generación de contenido, animaciones o efectos visuales en el contorno interior delimitado por la silueta de un objeto o sujeto.

5 Estos dispositivos y procesos permiten que el usuario conozca, por ejemplo, la apariencia que tiene ese objeto o sujeto cuando se le cambian determinadas condiciones.

Son utilizados en moda, de tal forma que un sujeto puede comprobar cómo le sientan, por ejemplo, unas determinadas gafas sin necesidad de disponer físicamente de las mismas, sino que se muestra en una pantalla el aspecto que tendría si las utilizase.

10 De igual forma, se utilizan para la interacción con distintos tipos de peinado, o de prendas de vestir. La utilización de este tipo de sistemas tiene usos diversos, en automoción, por ejemplo, para observar la apariencia externa o interna de un vehículo, color, tapizado, disponiendo sólo de un único elemento físico, etc.

Sin embargo, estos dispositivos presentan el inconveniente que los efectos visuales mostrados o bien no se ajustan a sujetos u objetos cuando estos se encuentran en movimiento o bien para la detección del movimiento es necesario que el sujeto u objeto disponga de determinadas marcas como prendas de un color determinado, sensores en el cuerpo, etc.

Se conocen algunos documentos donde se describen procedimientos o sistemas que tratan de abordar esta cuestión.

En el documento US200212609A1 se tiene como objetivo la navegación en la pantalla a través de un cursor dependiendo del estado de un objeto detectado por una cámara. El sistema de control controla en movimiento del cursor en la pantalla y determina cuando este cursor fue seleccionado mediante el estado de un objeto. De esa forma se pretende mover el cursor y navegar en un rango de movimientos equiparados a acciones del ratón. Sin embargo, solo se aplica a una parte detectada como objeto, cogen un punto de referencia y fijan los límites, y desde ahí calculan el movimiento realizando proyecciones basadas en los puntos de referencia. No aplican efectos especiales ni utilizan la técnica del análisis y procesamiento de imágenes con filtros de optimización en tiempo real, como ocurre en US598666OA, donde se describe un sistema de este tipo pero que tampoco trabaja en tiempo real, realizando las operaciones sobre datos pregrabados. En esta caso se utilizan 3,4 o más dimensiones y las convierten a distintas perspectivas dimensionales.

Por otro lado, otros sistemas que utilizan las mencionadas marcas de referencia se observan en US2003076293A1, donde se describe un sistema que utiliza marcas de referencia activas y pasivas por identificadores como guates, pintar el cuerpo. Además, los movimientos detectados se realizan con respecto al punto de referencia en matrices de 4x4 dimensiones utilizando varias cámaras de entrada y reconociendo los movimientos con una base de datos de gestos.

De forma similar en W0200497612A2 se describe un dispositivo que utiliza para la detección del movimiento al menos 2 imágenes provenientes desde distintos ángulos para determinar las posiciones en 3D del cuerpo de una persona y se realiza una representación de su volumen. La estructura de la solución implica la utilización de contenido 3D en los sensores, la adquisición, procesamiento, modelado, extracción de parámetros, calibración, etc. La solución propuesta en el presente documento es una solución 2D que requiere únicamente un solo dispositivo de entrada.

En otros campos de la técnica se conocen sistemas como el descrito en US2006184003A donde se remite a una base de datos de imágenes provenientes de resonancia magnética que proveen de los puntos de referencia asociadas al sujeto. Se describe la utilización de dispositivos externos como marcas al sujeto que permiten identificar la zona donde se actuará. Se utiliza para la visualización de intervenciones, mediante el empleo de realidad aumentada. En este caso, además de la visión estereoscópica se necesita un pre-procesamiento de las imágenes de referencia y nuevamente las marcas sobre el objeto o sujeto para el reconocimiento del movimiento.

Respecto a la proyección de los efectos visuales se conoce el sistema descrito en JP2001034381 que utiliza dibujos asociados a movimientos del cuerpo humano, que a su vez son empleados con el fin de controlar la interacción con el sistema. La diferencia principal con la propuesta actual es que en este caso se plantea la posibilidad de disparar y calcular los efectos visuales sobre la imagen del usuario objeto o usuario detectado por el dispositivo de entrada calculando su posición y utilizando máscaras como parámetros de referencia.

Finalmente, se conoce el sistema descrito en JP2005339100 que pretende el análisis de movimiento del cuerpo humano y plantea un método para realizar este análisis. De las imágenes detectadas por el dispositivo de entrada se extrae el cuerpo y esta silueta se transforma en un esqueleto. La diferencia entre el citado documento y la presente invención se sitúa en que el procesamiento de la identificación del movimiento se realiza a partir de este esqueleto, no sobre las imágenes o silueta detectada. Dado que pretende analizar el ritmo del movimiento plantea la separación en partes de ese esqueleto y el análisis de la cantidad y frecuencia de movimiento que existe en cada parte. Para lograrlo se crea un vector que analiza la cantidad de movimiento mediante la utilización de técnicas como la transformación lineal de Hough, filtro de Kalman para el seguimiento del tiempo, transformada de Fourier para la extracción de las características del movimiento y frecuencia, etc., difiriendo conceptos y planteamiento con nuestra invención.

## 15 **Descripción de la invención**

La presente invención se basa en un sistema que permite la aplicación de efectos visuales en objetos o sujetos estáticos o en movimiento en tiempo real sin necesidad de marcas, logrando una interacción natural entre el objeto o el sujeto y los dispositivos de entrada y de salida para lo que se utilizan técnicas de visión por computador.

El sistema comprende:

- 25 - Un dispositivo de entrada que permita la captura de imágenes, en cualquier formato y en tiempo real, que disponga de una interfaz de acceso al flujo de datos capturado. Dependiendo del tipo de dispositivo, se puede controlar y realizar ajustes en tiempo real sobre el flujo de datos capturado, tales como: resolución, tamaño y calidad de la entrada, así como propiedades como, brillo, contraste, saturación, etc.
- 30 - Un escenario configurado de tal forma que la iluminación y el fondo estén controlados. El escenario debe contener la iluminación y el fondo que serán utilizados para la configuración de los parámetros iniciales de los filtros de procesamiento de imágenes. Este escenario no debe contener "falsos" objetos en el fondo, es decir,

sombras u objetos en movimiento diferentes de aquél en el que queremos aplicar los efectos visuales.

- Un módulo de procesamiento de imágenes en tiempo real y detección de movimiento a través de filtros. El módulo de procesamiento de imágenes contendrá los  
5 filtros que se utilizarán para detectar, analizar y aplicar los distintos efectos visuales en las imágenes en tiempo real.

- Un dispositivo de salida donde se proyectan las imágenes procesadas. El dispositivo de salida puede ser cualquier visualizador de imágenes: pantallas de ordenador, proyector, televisión, etc.

10

Como ventaja técnica se debe mencionar que la interacción entre el objeto o sujeto y los dispositivos de entrada y salida, se realiza sin marcas en el objeto o sujeto, por lo que es no invasiva y no intrusiva. De esta forma el objeto o el sujeto interactúa con su propio cuerpo y con sus movimientos naturales si este se encuentra en  
15 movimiento.

Esta tecnología puede aplicarse para generar contenidos visuales donde un observador externo, o el propio objeto o sujeto aprecien cómo actúan agentes externos o acciones realizadas en/con/sobre el cuerpo. Por ejemplo se pueden generar efectos  
20 visuales relacionados con las prendas, de tal forma que un sujeto aprecie el aspecto de una prenda sobre su figura y cómo estas prendas se ajustan al movimiento que el sujeto realiza. También pueden mostrarse aspectos relacionados con el cabello, con la anatomía interior del cuerpo humano o cualquier otro evento, efecto gráfico o animación que se desee. Dichos eventos, efectos gráficos o animaciones son aplicados  
25 en el cuerpo en tiempo real.

Para el funcionamiento del sistema se necesita la aplicación de diferentes algoritmos que forman parte de cada uno de los filtros implementados en el módulo de procesamiento de imágenes en tiempo real y detección de movimiento. Se tienen dos tipos de filtros fundamentales: filtros de detección y filtros de proyección. Donde:

30

Los filtros de “detección” consisten en:

-Un filtro objeto que permite la detección de la zona ocupada por el sujeto

-Un filtro de movimiento que permite la detección de movimientos/acciones del sujeto.

Los filtros de proyección considerados consisten, por ejemplo, en:

- 5 - Efecto halo
- Efecto desaturado
- Efecto reproducción animación en zona dentro del objeto o sujeto y seguimiento de movimiento.

Los términos que se manejan se definen de la forma siguiente:

10

-Fotograma: Definiremos como  $F_i$  al fotograma capturado en el instante de tiempo  $i$

-Escenario: Espacio físico de dimensión variable, donde se situará el individuo u objeto sobre el que se realizará el estudio.

15

-Fotograma clave B: fotograma de referencia con el que se comparan los fotogramas del flujo de datos capturados con el objetivo de encontrar diferencias en atributos de color respecto al el mismo.

-Fotograma máscara  $M_i$ . Imagen que contiene la máscara para el instante de tiempo  $i$ , que determina que parte del fotograma  $F_i$  es parte del sujeto u objeto.

20

-Máscara es una matriz de valores reales entre 0 y 1, que se corresponden a un píxel de la imagen obtenida por la cámara o dispositivo de entrada. En este caso los almacenamos como valores de gris. La máscara se utiliza en los filtros de detección y proyección. En los filtros de detección se genera la máscara. En los filtros de proyección lo utilizamos para el cálculo de aplicación de efectos visuales.

25

Cada uno de los filtros tiene el siguiente funcionamiento:

30

-Filtro objeto.

Se utiliza para determinar, en primer lugar, a través del estudio del flujo de datos de entrada del dispositivo de captura el espacio ocupado por el sujeto u objeto

dentro del escenario utilizado. El segundo paso consiste en comparar de manera sucesiva y en tiempo real, el fotograma actual capturado y el fotograma clave. La operación de comparación entre fotograma actual y fotograma entrenado, consiste en restar, a nivel de píxel, el fotograma actual menos el entrenado. El resultado obtenido de esta operación, conformará el "fotograma máscara", que consiste en una imagen en la que se muestra el espacio ocupado por el objeto o sujeto en color blanco y el resto de la imagen en color negro, logrando la acotación de la región de la imagen sobre la que se aplicarán los demás efectos. Se utilizan los valores máximos y mínimos de las tres componentes de color R,G,B para aplicar los filtros. El blanco es un vector (1.0, 1.0, 1.0) compuesto de los valores máximos de las tres componentes de color. El negro es un vector (0.0, 0.0, 0.0) compuesto por los valores mínimos de las tres componentes. Además, se utilizan cualquiera de los tres componentes como factor de mezcla, donde 1.0 significa aplicar el efecto completamente y 0.0 significa no aplicarlo.

La detección se produce en dos pasos, el primero consiste en el estudio de un conjunto de fotogramas. El resultado de este estudio generará fotograma clave y lo denotaremos como B. Se tomarán tantos fotogramas como sean necesarios hasta eliminar el ruido en el fotograma clave. El ruido de los fotogramas capturados dependerá del tipo de sensor utilizado, de las condiciones de iluminación y de los ajustes internos de la cámara o sensor (autoiluminación, balance de blanco, ganancia, etc.), del escenario sin ningún sujeto u objeto en pantalla, etc. A este proceso se denomina "Entrenamiento". Una vez finalizado obtenemos el "fotograma entrenado". El fotograma entrenado se compone de varias capturas realizadas en la inicialización y que contiene el fondo original.

Calculamos el fotograma clave  $B_i$  para  $n$  fotogramas como:

$$B_i = \text{Average} ( F_{i-u}, F_{i-(u-1)}, \dots, F_{i-1} )$$

A continuación, para conseguir aislar la silueta del sujeto del resto de la imagen capturada nos ayudamos del fotograma clave  $B_i$ . El objetivo es determinar que zona del fotograma actual capturado ( $F_i$ ) ha cambiado con respecto al fotograma clave  $B_i$  al introducir un elemento sobre el mismo.

Para determinar que píxeles de la imagen capturada han cambiado con respecto al fondo original, procesamos cada píxel del fotograma  $F_i$  con su correspondiente píxel en fotograma clave  $B_i$  de la siguiente manera:

5 Sean  $F_{iR}$ ,  $F_{iG}$ ,  $F_{iB}$  las componentes R G y B del fotograma capturado en el instante  $i$ .

Sean  $B_{iR}$ ,  $B_{iG}$ ,  $B_{iB}$  las componentes R G y B del fotograma clave.

Sean  $\delta R$ ,  $\delta G$ ,  $\delta B$  los umbrales de aceptación de cada componente RGB.

Calculamos la máscara del sujeto para el instante  $i$  como:

10  $1.0$  si  $|B_{iR} - F_{iR}| > \delta R$  ó  $|B_{iG} - F_{iG}| > \delta G$  ó  $|B_{iB} - F_{iB}| > \delta B$

$M_i = 0.0$  en caso contrario.

Siendo la entrada:

A: Píxel de la imagen capturada

15 B: Píxel correspondiente en la imagen original.

Salida:

C: Píxel que determina si el píxel actual ha cambiado:

Dado que el procesamiento se realiza fotograma a fotograma y una imagen  
20 estática no tiene movimiento, determinamos que hay movimiento en las zonas que no son comunes entre el fotograma clave y el fotograma capturado en un cualquier instante. De esta forma queda definido el movimiento en un fotograma con respecto al fotograma clave como negro en las zonas comunes y manchas blancas en las zonas que son diferentes entre ambos.

25

Hay que tener en cuenta el umbral de aceptación para cada componente del espacio de color. Por ejemplo si el espacio es RGB, necesitaremos umbrales para las componentes:  $R_{th}$ ,  $G_{th}$ ,  $B_{th}$ .

30 Por lo tanto, para cada componente del espacio de color se hace la diferencia de los píxeles de las componentes significativas y se toma la mayor de las diferencias que dependen del espacio de color utilizado. En este caso se ilustra el procedimiento con el

RGB donde se hace la diferencia entre las tres componentes, ya que todas proporcionan la misma cantidad de información sobre el color.

Luego, para calcular la citada diferencia:

5  $|B - A|$

A continuación, se compara cada diferencia con el umbral correspondiente a su componente, previamente definido, y determinado según el grado de aceptación que queramos sobre las manchas generadas al hacer las diferencias.

Si alguna de las diferencias es mayor que su umbral de aceptación:

10  $C = \text{Sujeto.}$

O si no,  $C = \text{Fondo.}$

-Filtro movimiento

15 La implementación de este filtro permite detectar el movimiento del objeto o sujeto y asociar este movimiento a una acción predeterminada e interpretada por el sistema.

El reconocimiento de la acción disparará el efecto de reproducción de la animación dentro del objeto o sujeto y el seguimiento del movimiento del objeto o  
20 sujeto.

Para reconocer una acción se debe detectar un movimiento continuo en una determinada dirección y a una velocidad determinada entre dos o más posiciones de la imagen.

25 El procedimiento es el siguiente:

En primer lugar calcularíamos la velocidad del desplazamiento, que puede ser en distintos sentidos, y en vertical u horizontal entre dos fotogramas entre los instantes  $i$  e  $i+1$ . Por ejemplo, en el caso del desplazamiento vertical se calcula de la siguiente forma:

30 
$$V_i = \frac{Y_{i+1} - Y_i}{t_{i+1} - t_i}$$

Siendo  $Y_i$  la altura detectada en el instante  $i$  y siendo  $t$  el tiempo.

Por otro lado, se calcula el desplazamiento ( $h$ ) entre dos fotogramas entre los instantes  $i$  e  $i+1$  como:

5

$$h_i = |C_{i+1} - C_i|$$

Se define una zona de captura sobre la imagen para realizar el cálculo del movimiento, de esta manera, se aíslan "falsos movimientos" que pueda realizar el objeto o sujeto y que no tengan relación con la acción que estamos detectando en una zona determinada de interés. La zona de captura depende del tamaño del objeto o sujeto a realizar el movimiento, por tanto, se deben definir en tiempo real una serie de parámetros, tales como origen y final, velocidad y distancia.

15 Las acciones también se han de definir. Por ejemplo, puede implementarse la acción de beber que consiste en que el sujeto levanta un objeto desde la cintura hasta la mitad de la cabeza. Esta acción se describirá posteriormente.

El filtro actuaría inicializando los siguientes parámetros a partir del objeto o sujeto detectado:

20

-Posición inicial: coordenadas de origen donde comenzará a detectarse el movimiento.

-Posición final: coordenadas de destino donde se dará por concluido el movimiento.

25

Una vez detectado que se inicia un movimiento, se debe calcular el centro de la zona capturada donde se está detectando el movimiento o acción. Para ello obtenemos el rectángulo que lo contiene y calculamos su centro. El valor obtenido será considerado el centro del objeto en movimiento.

30

Se calcula, a continuación, la distancia medida en píxeles entre dos fotogramas consecutivos, valor que indica el desplazamiento que se ha realizado en línea recta entre

los centros del movimiento de dos fotogramas consecutivos. Con este valor podemos estimar la velocidad con la que se mueve el objeto.

Para ilustrar el funcionamiento de este filtro se explica el funcionamiento de la  
5 detección una acción específica del sujeto: beber.

Suponiendo que se desea conocer si el sujeto capturado por la cámara está  
bebiendo, se debe detectar entonces un movimiento continuo y ascendente y a una  
velocidad adecuada entre dos posiciones de la imagen.

10

La implementación será de la forma siguiente:

15

Para determinar en que zona de la imagen capturada hay movimiento,  
realizamos una diferencia del fotograma actual con el fotograma anterior, almacenado  
en una imagen auxiliar. De esta manera, la zona que en ambas imágenes es similar  
quedará negra y la zona con diferencias notables en colores claros. Dado que la  
diferencia entre los fotogramas contendrá un número elevado de píxeles claros aislados  
debido a movimientos leves e inevitables, se procesa cada píxel de la imagen generada  
por la diferencia de fotogramas de la siguiente manera:

20

Entrada: Imagen generada a partir de la diferencia entre dos fotogramas

Salida: Máscara que determina en que zona de la imagen hay movimiento.

Parámetros: Umbral de aceptación para cada componente del espacio de color.

Distancia: Distancia a la que se encuentran los píxeles cercanos.

25

30

A continuación se compara cada componente con su correspondiente umbral de  
aceptación. Este umbral de aceptación es dependiente de las condiciones externas de  
iluminación, cantidad de colores que hay en cada momento en la imagen, precisión que  
se desee obtener y el ajuste que se desee entre el objeto que se mueve y la zona de  
movimiento que se dese capturar. Con esto se consigue despreciar los movimientos que  
no nos interesan.

Para los píxeles en los que se haya detectado movimiento se debe analizar si una zona contiene píxeles de movimiento:

Para ello, se comprueba si los píxeles que se encuentran a cierta distancia (1 a n píxeles) del píxel en el que se detectó el movimiento, también son de movimiento.

5 Si todos los píxeles comprobados son de movimiento: se considera definitivamente de movimiento.

En caso contrario no se considera movimiento en ese píxel.

La implementación de este código puede realizarse por ejemplo en Adobe Píxel  
10 Blender Toolkit.

Partiendo de una imagen generada a partir de la diferencia entre dos fotogramas consecutivos, se procesa cada píxel de la siguiente manera:

```

15 void evaluatePíxel() {
    float threshold = 0.1;
    float2 coord = outCoord();
    float2 hOffsetx = float2(píxelSize(foreImage).x, 0.0);
    float2 hOffsety = float2(0.0, píxelSize(foreImage).y);
20
    float4 forePíxel = sampleNearest(foreImage, outCoord());
    // Asignamos a la variable WhiteColor un valor 1.0 si las tres
    // componentes superan el umbral de color y 0.0 si no es así.
    float whiteColor = (forePíxel.r > thresold && forePíxel.g > thresold
25 && forePíxel.b > thresold) ? 1.0: 0.0;
    // Si la variable es mayor que cero tomamos el valor de los píxeles
    // Que se encuentran exactamente en un radio de 5 píxeles.
    // En este caso, utilizamos solo los píxeles que se encuentran en la //
    horizontal y en la vertical del píxel que estamos estudiando.
30 if (whiteColor > 0.0) {
    píxel4 up = sampleNearest(foreImage, coord - 5.0 * hOffsety);
    píxel4 down = sampleNearest(foreImage, coord + 5.0 *
hOffsety);
    píxel4 left = sampleNearest(foreImage, coord - 5.0 * hOffsetx);
35 píxel4 right = sampleNearest(foreImage, coord + 5.0 * hOffsetx);
    // Volvemos a asignarle a la variable whiteColor el valor 1.0 // o 0.0 si las
    componentes de los 4 píxeles estudiados superan
    // el umbral establecido.
    whiteColor = (
40 up.r > threshold && up.g > threshold && up.b > threshold

```

```

    && down.r > threshold && down.g > threshold  && down.b >
threshold
    && left.r > threshold && left.g > threshold  && left.b > threshold
    && right.r > threshold && right.g > threshold  && right.b >
5 threshold)
    ? 1.0: 0.0;
    }
    // Finalmente asignamos a las componentes del píxel de destino el
    // valor de la variable whiteColor
10     dst = float4 (whiteColor, whiteColor, whiteColor, 1.0);
    }

```

De esta forma se obtiene una máscara que determina donde hay movimiento en alguna zona de la imagen.

15 A continuación inicializamos las siguientes variables:

- Ultima\_Posición: altura donde queremos capturar el comienzo del movimiento ascendente.

20 -Máximo\_Tamaño\_Rectángulo: Dimensiones máximas que debe tener el rectángulo para considerar el movimiento y evitar realizar cálculos sobre movimientos de todo el cuerpo. Se utiliza un rectángulo dado que nos proporciona los valores de alto, ancho y posición horizontal y vertical de una de sus esquinas. Esta figura es la más simple para el cálculo computacional de la detección del movimiento.

- Posición\_Actual: Altura actual del rectángulo.

- Altura\_Inferior, Altura\_Superior: Alturas origen y destino del movimiento.

25 - Distancia\_Máxima: Se utiliza para limitar la velocidad del movimiento.

30 Seguidamente obtenemos el rectángulo que contiene a la zona blanca de esta máscara mediante la función Bounding Box, una función que nos devuelve el rectángulo más pequeño que contendría a todos los píxeles de determinadas características dentro de una imagen. Particularmente todos los de un color determinado. En este caso buscamos el que contenga a los píxeles de color blanco.

En cada fotograma tomamos el borde superior (coordenada Y de la esquina superior derecha) del rectángulo: Posición\_Actual.

35

Si se cumplen las siguientes condiciones,

- Tamaño del rectángulo es menor a `Máximo_Tamaño_Rectángulo`.
- `Posición_Actual < Ultima_Posición`
- `Posición_Actual < Altura_Inferior`
- `Última_Posición - Posición_Actual < Distancia_Máxima`.

5

Entonces se realizan lo siguiente:

- Se almacena el valor de `Posición_Actual` en `Ultima_Posición`.

- Si `Posición_Actual > Altura_Superior` entonces la persona está bebiendo.

10

Un ejemplo de la implementación del filtro podría ser el siguiente:

```

public function detectMovement (mixLevel: Number): void {
  // Obtenemos la parte superior de la silueta del sujeto: Calculamos el rectángulo
15 // que contiene su silueta y obtenemos la coordenada y del mismo.
  // A partir de este valor definimos la altura donde comienza y termina el
  // movimiento.
  var top: int = subjectMask.getColorBoundsRect(0xFFFFFFFF, 0xFFFFFFFF, true).y + 40;
  var bottom: int = top + 90;
20 // Obtenemos la diferencia entre el fotograma actual y el anterior.
  lastFrame.draw(_src, null, null, BlendMode.SUBTRACT);
  // Aplicamos el filtro que elimina los píxeles aislados.
  lastFrame.applyFilter(lastFrame, imageRect, new Point(),
    detectMovementShaderFilter);
25 // Obtenemos el rectángulo que contiene la zona de la imagen donde hay movimiento.
  movementRectangle = lastFrame.getColorBoundsRect(0xFFFFFFFF, 0xFFFFFFFF,
    true);
  // Primeros requisitos del rectángulo:
  // Utilizamos las dimensiones del rectángulo obtenido para determinar si hay
30 // movimiento en la imagen y también para rechazar imágenes con movimientos de
  // objetos mayores al que se quiere detectar
  if ((movementRectangle.width > 0) && (movementRectangle.width < 100)) {
  // Segundos requisitos del rectángulo
  // Comprobamos que la altura del rectángulo esté entre las alturas
35 // previamente calculadas, y que la altura en el fotograma actual es mayor

```

```

// que la detectada en el fotograma anterior. Con esto limitamos la
// detección a un movimiento ascendente.
// También comprobamos que la diferencia entre la altura del rectángulo del
// fotograma actual y el anterior (en píxeles) no supere cierto valor.
5 // En este caso lo hemos ajustado a 50 píxeles.
if ((movementRectangle.height < 120)
    && (movementRectangle.y < lastArmPosition)
    && (movementRectangle.y < bottom)
    && (lastArmPosition - movementRectangle.y < 50)) {
10 // Dado que este fotogramas cumple con los requisitos de un
// Movimiento correcto (Condiciones anteriores), Iniciamos a 0 el
// contador de fotogramas con fallo.
frameCount = 0;
// Actualizamos la altura anterior con la actual.
15 lastArmPosition = movementRectangle.y;
// Si hemos alcanzado la altura final, consideramos que hemos
// detectado el movimiento
if (lastArmPosition < top)
    bebiendo = true;
20 // Para los fotogramas que cumplen los primeros requisitos (hay movimiento
// y de dimensiones adecuadas) pero no los segundos (no se realiza
// movimiento a la velocidad y en el sentido adecuados, incrementamos el
// Contador de fotogramas con fallo.
} else {
25 frameCount++;
}
}
// Al superar el numero de fotogramas con fallo reponemos el contador y asigamos a
// la posicione anterior del brazo la posición inicial.
30 if (frameCount > 15) {
lastArmPosition = bottom;
frameCount = 0;
}
}
35

```

Descritos los filtros de detección, pasemos a describir los filtros de proyección.

-Filtro Halo

Se utiliza para proyectar la superposición de un “halo” coloreado alrededor del contorno del sujeto u objeto. Este efecto añade a la imagen original un efecto luminoso del color deseado alrededor de la figura de un sujeto u objeto mostrado ante la cámara o dispositivo de entrada. Se emplea la máscara con la silueta del sujeto, generada en etapas anteriores, para dibujar el halo en la zona definida como fondo (negro) más próxima al sujeto.

El funcionamiento del mismo consiste en ( véase Figura 1):

Aplicar, en primer lugar, a la máscara del sujeto u objeto un efecto de difuminado que agregará un degradado de blanco a negro entre la silueta del sujeto u objeto y el fondo, generando una nueva máscara en diferentes niveles de gris. El efecto de difuminado se consigue asignando a cada píxel el valor promedio de todos los píxeles contenidos dentro de una ventana de  $n \times n$  píxeles teniendo en su centro al píxel que queremos modificar. La distancia de separación con respecto al píxel determinará la intensidad del difuminado.

Se considera zona válida para dibujar el halo a aquella zona externa a la silueta del sujeto pero que no sea negra en la máscara difuminada.

A continuación, se mezcla la imagen original con el color deseado en la proporción determinada por la intensidad de gris de la zona válida.

La descripción algorítmica es la siguiente (véase Figura2):

Primero: Sea A la máscara del sujeto, realizamos una copia B de la misma.

Segundo: Aplicamos sobre A un efecto de difuminado.

Tercero: Para todos los píxeles (i,j) de A, Si  $B(i,j) = 1.0$ ,  $A(i,j) = 0.0$ , o lo que es lo mismo, se procede a eliminar la zona ocupada por el sujeto.

Cuarto: Para todos los píxeles (i,j) de A,  $A(i,j) = A(i,j) * 2.0$ ; o lo que es lo mismo, se procede a duplicar la intensidad del efecto.

5. Quinto: Aplicamos sobre A un efecto de difuminado.

-Filtro desaturado

Permite eliminar la diferencia entre las componentes de la imagen en la zona  
 5 determinada por la máscara de sujeto, creando un efecto “Blanco y Negro” sobre el  
 sujeto detectado.

El funcionamiento del mismo consiste en que para cada píxel de la imagen se  
 calcula su valor desaturado realizando la media entre sus componentes R, G y B. Esto  
 10 es, para cada píxel de la imagen, obtenemos el valor de sus tres componentes como  

$$R = B = G = (R_{original} + G_{original} + B_{original}) / 3.0;$$

A continuación, la imagen desaturada resultante se mezclará con la original  
 según los niveles de la máscara del sujeto de la forma siguiente:

15

$$R = R_{original} * (1 - \text{valor de la máscara}) + R_{desaturado} * \text{Valor de la máscara}$$

$$G = G_{original} * (1 - \text{valor de la máscara}) + G_{desaturado} * \text{Valor de la máscara}$$

$$B = B_{original} * (1 - \text{valor de la máscara}) + B_{desaturado} * \text{Valor de la máscara}$$

20

Para conseguir un suavizado entre la zona saturada y el fondo se puede aplicar  
 también un efecto de difuminado de la intensidad deseada, definiendo niveles  
 intermedios entre el blanco y el negro de la máscara. Esto es así, dado que la máscara  
 obtenida del sujeto solo contiene valores 1.0 o 0.0 por lo que los contornos presentan los  
 denominados “dientes de sierra”. Aplicar un efecto de difuminado suaviza los contornos  
 25 proporcionando un aspecto mas suave y natural.

Para cada par de píxeles correspondientes entre la imagen original y la imagen  
 resultado se tiene como entrada

A: Píxel de la imagen original.

30

B: Píxel de la máscara. Valores comprendidos entre 0 y 1.

Y como salida:

S: Píxel de la imagen resultado.

A continuación se calcula:

$C = \text{Píxel desaturado a partir del píxel de la imagen original.}$

5           Donde, por ejemplo, para RGB se calcula como  $(R+G+B) / 3$

Para HSV, un espacio de color en el que las componentes son: tonalidad (una de las frecuencias del espectro lumínico), intensidad de iluminación y saturación (cantidad de color respecto de su valor de iluminación) se calcula haciendo la componente  $S = 0$  (Saturación = 0), aunque como alternativa, es suficiente multiplicar la componente de saturación por B.

10

Finalmente se aplica la máscara (B) para determinar que cantidad de A y C se almacenará en S.

15           Lo mencionado equivale a aplicar si se trabaja en el espacio RGB, a cada componente de color:  $S = A * (1 - B) + C * B$ , mientras que para HSV se multiplica la componente de saturación por B, por lo que no es necesario calcular C (Píxel desaturado).

20           Como ejemplo de código implementado se tiene:

```
// Evaluación de cada píxel en lenguaje Adobe Pixel Blender Toolkit.
void evaluatePixel() {
// Obtención de los valores del píxel de ambas máscaras
float4 frontPixel = sampleNearest(frontImage, outCoord());
25 float4 maskPixel = sampleNearest(maskImage, outCoord());
// Obtención de la componente r de cada píxel.
float B = frontPixel.r;
float A = maskPixel.r;
// Definimos el valor inicial a partir de los valores
30 // anteriores
float C = 2.0 * (1.0 - A) * B;
// Asignamos el valor obtenido a todas las componentes
// del píxel de destino (Mascara compuesta resultante)
// Obteniendo una imagen es diferentes niveles de gris.
35 dst = float4(C, C, C, 1.0);
}
```

- Filtro de reproducción animación en zona dentro del objeto sujeto y seguimiento de movimiento.

5            En este filtro se reproduce una animación dentro de una zona del sujeto u objeto detectado (zona en máscara objeto), realizándose tanto el procesamiento como la reproducción por cada frame analizado de tal forma que se sigue el movimiento del objeto o sujeto.

10           El funcionamiento de este filtro consiste en aplicar para cada frame saturado las siguientes etapas:

          En primer lugar, generar la máscara con el contorno del sujeto u objeto.

15           A continuación, tomar el rectángulo mínimo que contiene al sujeto u objeto en la máscara. Este rectángulo definirá la posición y dimensiones del sujeto u objeto dentro del fotograma actual, por lo que se dispone del ancho, alto y la posición horizontal y vertical de uno de sus vértices.

          Acto seguido se fijan las coordenadas donde desea reproducir el vídeo, efecto visual o animación deseada, así como su tamaño en posiciones y dimensiones relativas al rectángulo detectado.

20           Finalmente se realizan las operaciones aritméticas sobre los valores de posición y tamaño, tanto del rectángulo como de la imagen a superponer, para calcular la posición y tamaño finales de la imagen a reproducir.

25           Suponiendo que se desea reproducir una animación en el centro del sujeto, se procedería de la forma siguiente:

          Primero se define:

$X_r, Y_r, W_r, H_r$ : Posición horizontal, posición vertical ancho y alto del rectángulo.

$X_v, Y_v, W_v, H_v$ : Posición horizontal, posición vertical ancho y alto del vídeo.

30           Segundo se define un factor de escala, compuesto por un número entero que indica la escala del video con respecto al tamaño original

          Tercero se calcula:

$$X_v = X_r + (W_r / 2) + (W_v / 2)$$

$$Y_v = Y_r + (H_r / 2) + (H_v / 2)$$

$$H_v = H_r * \text{Factor de escala.}$$

$$W_v = W_r * \text{Factor de escala.}$$

5

En cuarto y último lugar, se proyecta y reproduce la animación que se desee.

De forma general este sistema que permite la aplicación de efectos visuales en objetos o sujetos estáticos o en movimiento en tiempo real sin necesidad de marcas  
10 funcionaría de la siguiente forma ( véase Figura 2):

Un sujeto u objeto entra en el escenario controlado y se posiciona frente al dispositivo de entrada, por ejemplo, una cámara estática a una distancia mínima que permita su visualización de su cuerpo entero o parte del cuerpo que desee verse en el  
15 dispositivo de salida.

El sujeto no necesita llevar ropas, guantes, ni ningún objeto o color que para su identificación, detección y o tratamiento por el sistema, es decir, no necesita ninguna  
20 marca.

El fondo y la iluminación deben estar controlados, evitando las sombras sobre el fondo o sujeto. Se puede seleccionar por ejemplo:

-Fondo: color primario, uniforme y mate, cubriendo todo el angular de la  
25 cámara.

-Iluminación: tipo de iluminación “luz de día”, iluminando fondo y sujeto sin proyección de sombra en el angular de la cámara ni iluminación directa.

El dispositivo de visualización o de salida se ubica frente al sujeto, de forma que simule un espejo, aunque este dispositivo también podrá ubicarse en otras posiciones,  
30 de forma que el sujeto pueda observarse en él (por ejemplo, a un costado o arriba).

El dispositivo de procesamiento controlará y procesará las imágenes desde los dispositivos de entrada hacia los dispositivos de salida que se encuentra conectado a los mismos.

5 Con el sujeto en posición e iluminación y fondo controlado el proceso comienza de la siguiente forma (Véase Figura 2):

Primero: se reciben desde el buffer de datos del dispositivo de entrada (cámara estática) imágenes secuencialmente en formato nativo (bitmap). Dependiendo del  
10 dispositivo de entrada se necesitarán codec diferente para decodificar las imágenes. En el caso de bitmap, no hay codificación y se hace automático.

Segundo: se toman las imágenes en la resolución nativa del dispositivo y se redimensiona a una resolución menor para el cálculo. Los fotogramas son capturadas a la resolución nativa del dispositivo de entrada, para garantizar así la calidad de los  
15 mismos. Durante el proceso de estudio, estos son redimensionados a una resolución cercana a 320x240 ( o proporcionalmente, en caso de que la resolución óptima del dispositivo de entrada sea mayor) para garantizar que el proceso de cálculo se desarrolla a un nivel óptimo en tiempo real. La capacidad de cómputo es directamente proporcional a la potencia de la máquina en la que se ejecute el sistema.

20 Tercero: se realiza un análisis y procesamiento secuencial de imágenes en tiempo real analizando en cada una de ellas sus componentes de color para detectar sus diferencias con el flujo de datos de entrada (vector de mapas de bits).

Cuarto: se optimiza el procesamiento y la comparación de los frames mediante los denominados “filtros de detección siguientes:

25 - Filtro objeto: compara A con B, donde A es fondo entrenado y B es fotograma actual

-Filtro movimiento: compara A con B, donde A es el frame anterior y B es el frame actual

30 El proceso de comparación general funciona de la siguiente forma:

Una vez se recibe el flujo de datos de entrada (frame-bitmap), se descompone dicho frame de entrada en sus componentes según su espacio de color. En el caso de RGB, se descompone en 3 canales de color, se procede a la realización de la diferencia

de componentes a nivel de píxel y se toma el que más diferencia tenga en cada píxel por cada canal de color, en función de un umbral predeterminado (que se pide como parámetro). Los umbrales se fijan en función de la iluminación y fondo.

5           Quinto: A la imagen original se le aplican filtros de proyección, que dependerán de los valores de color, límites y puntos de referencia (parámetros) de la máscara detectada y del dispositivo de salida.

10           Todos los filtros son inicializados con una serie de parámetros que pueden variar según condiciones de iluminación, tamaño y efecto a conseguir. Se pueden incluir 1 a n filtros y configurarlos.

            Para la visualización, se redimensionan nuevamente a la resolución adecuada al dispositivo de salida utilizado en cada caso.

15

### **Descripción de las figuras**

En la figura 1 se describen los pasos para la crear el efecto de halo sobre el sujeto u objeto detectado. Estos pasos son:

- 20    1.     Se captura un fotograma  
       2.     Se obtiene la máscara correspondiente al fotograma capturado A y se crea una copia B.  
       3.     Se aplica un efecto de difuminado a A.  
       4.     Se multiplica A por 2 y se le resta B  
 25    5.     Se aplica a A un segundo difuminado.  
       6.     A determina donde dibujar el halo y con que intensidad alrededor de la silueta del sujeto.

30           En los pasos 2, 3, 4 y 5 se observa las variaciones de valores en el rango 0.0-1.0 del fondo (negro) y de la silueta del sujeto/objeto (blanco) con las que se obtiene este efecto de halo sobre la silueta.

            La Figura 2 representa un esquema que describe en forma resumida como se estructura el sistema en etapas dentro de un proceso secuencial y los elementos claves que componen cada etapa.

### **Modos de realización de la invención**

A modo de ejemplo, un sistema que permite la aplicación de efectos visuales en objetos o sujetos estáticos o en movimiento en tiempo real sin necesidad de marcas está  
5 constituido por:

- Dispositivo de entrada: Webcam Creative Labs
- Escenario controlado: Habitación de 3 m. de largo por 2 m. de ancho. Fondo cromático (color azul ó verde) e iluminación “luz de día” (temperatura de color de entre 5.600 K y 6.000 K).
- 10 - Dispositivo de salida: Proyector LCD haz de luz rectangular.
- Dispositivo de procesamiento CPU de un ordenador portátil en el que se han implementado los filtros de detección y los filtros de proyección descritos anteriormente.

15 El objetivo se basa en mostrar como se encuentra el aparato digestivo de una persona antes de beber y después de beber una determinada sustancia líquida, de tal forma que la persona observe en el dispositivo de salida un efecto visual animado en el interior de su cuerpo.

20 Como se ha indicado, en primer lugar, se ha de fijar el fondo, la iluminación, y establecer los umbrales y parámetros iniciales para los filtros de procesamiento de imágenes. Los umbrales dependen de los colores y contrastes capturados por el dispositivo de entrada. La correcta definición de los umbrales influye en la precisión de la detección.

25 Se debe considerar que unos umbrales con valores más altos que los establecidos como correctos harán que la silueta del sujeto obtenida presente manchas en el interior del contorno definido por la silueta, estas manchas se interpretan como fondo, mientras que unos umbrales con valores más bajos presentarán manchas en el fondo de la imagen  
30 confundiéndolas como parte del sujeto. Esto afecta especialmente en la reproducción del efecto visual en el interior del contorno definido por la silueta del sujeto y en el filtro de detección de movimiento.

A continuación se sitúa al sujeto en el escenario anterior y se le hace posicionar frente a la cámara tomando como referencia la visualización frente al proyector LCD, que el sujeto interpreta como si fuese un espejo. La imagen del sujeto es proyectada.

5           Acto seguido, se aplica el filtro de detección de sujeto descrito en apartados anteriores, realizando el recorte de la figura del fondo. Esta máscara del sujeto se utiliza como referencia para aplicar un filtro de efecto visual “desaturado”, también descrito, creando la impresión en el sujeto de una pérdida de color, que es interpretado de manera imaginaria como una posible deshidratación.

10           A continuación el sujeto toma líquido, para lo cual toma una botella y se aproxima la misma a la boca. Esta acción se detecta mediante el filtro de máscara de movimiento. La acción beber es interpretada por las posiciones prefijadas de punto inicial y final en los frames, tal como se indicó anteriormente en la explicación del funcionamiento de este filtro. Al detectarse la acción, se dispara la aplicación del filtro  
15 de efecto visual de reproducción animación en zona dentro del sujeto y seguimiento de movimiento.

Dentro de la zona de reproducción, se muestra la animación deseada, por ejemplo, de unas bolitas cayendo por el cuerpo y su expansión dentro del cuerpo (caen por la garganta hasta el intestino y allí se expanden).

20           El sujeto puede moverse en el escenario y la animación seguirá su movimiento y se reajustará a su tamaño gracias al filtro anterior.

Cuando termina esta animación, se dispara el siguiente filtro de efecto visual de halo, en donde se dibuja un halo de color sobre la silueta del sujeto con las características predefinidas.

25           En todo momento el sujeto puede moverse en el escenario y los efectos se ajustarán a sus movimientos. Esto se debe a que el análisis y procesamiento de imágenes se hacen frames a frames en tiempo real.

**REIVINDICACIONES**

1. Sistema que permite la aplicación de efectos visuales en objetos o sujetos estáticos o en movimiento en tiempo real sin necesidad de marcas caracterizado por:
- 5
- Un dispositivo de entrada que permite la captura de imágenes
  - Un escenario configurado con iluminación y fondo controlado
  - Un dispositivo de procesamiento de imágenes en tiempo real y detección de movimiento.
- 10
- Un conjunto procedimientos utilizados para detectar, analizar y aplicar efectos visuales en los objetos o sujetos en tiempo real que comprende la aplicación de:
    - i. Filtros de detección,
    - ii. Filtros de proyección.
- 15
- Un dispositivo de salida donde se proyectan las imágenes procesadas.
2. Sistema que permite la aplicación de efectos visuales en objetos o sujetos estáticos o en movimiento en tiempo real sin necesidad de marcas según reivindicación 1 donde los filtros de detección comprenden:
- 20
- Un filtro objeto que permite la detección de la zona ocupada por el objeto.
  - Un filtro de movimiento que permite la detección de acciones o movimientos desarrollados por el objeto.
- 25
3. Sistema que permite la aplicación de efectos visuales en objetos o sujetos estáticos o en movimiento en tiempo real sin necesidad de marcas según reivindicación 2 donde se emplea un filtro de movimiento para detectar la acción de beber realizada por un sujeto.
- 30
4. Sistema que permite la aplicación de efectos visuales en objetos o sujetos estáticos o en movimiento en tiempo real sin necesidad de marcas según reivindicación 1 donde los filtros de proyección comprenden:
- Un filtro para la aplicación del efecto halo.

- Un filtro para la aplicación del efecto desaturado.
- Un filtro para la reproducción de una animación en una zona interior del contorno del objeto que acompaña al movimiento.

5

5. Procedimiento para la aplicación de efectos visuales en cuerpos o sujetos estáticos o en movimiento en tiempo real sin necesidad de marcas mediante el uso de un sistema según la reivindicación 1 caracterizado por las siguientes etapas:

10

-Etapa 0: Control y definición de las condiciones iniciales de: iluminación, resolución del dispositivo de entrada, fondo y situación del objeto en el lugar predefinido.

15 -Etapa 1: Detección en tiempo real de la imagen captada por el dispositivo de entrada, considerando color y/o movimiento mediante la aplicación de filtros de proyección y detección.

- Etapa 2: Análisis en tiempo real de la imagen para la separación de la silueta del sujeto/objeto del fondo mediante la aplicación de filtros de proyección y detección.

20 -Etapa 3: Reconocimiento y almacenamiento de la silueta del sujeto/objeto.

-Etapa 4: Aplicación de máscaras sobre la silueta del sujeto/objeto mediante la aplicación de filtros de proyección y detección.

-Etapa 5: Detección de las secuencias de eventos/acciones preprogramadas.

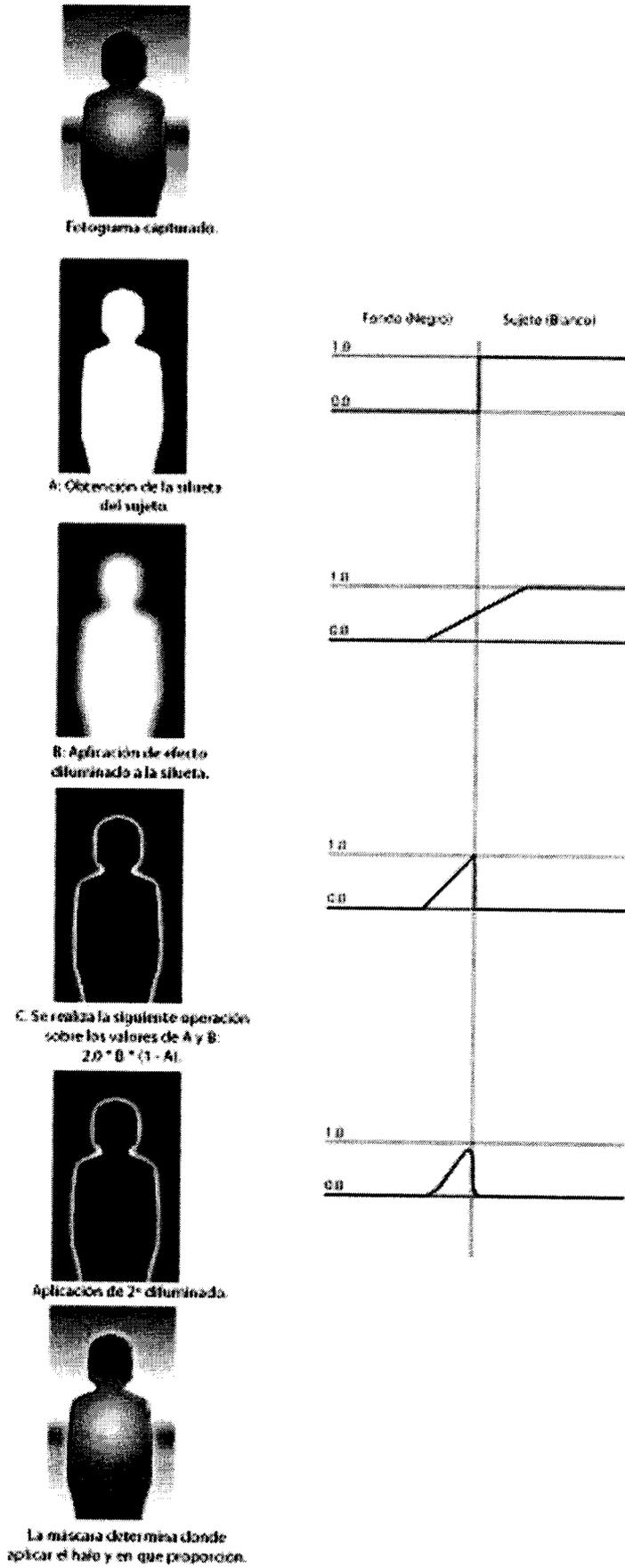
25 -Etapa 6: Según las acciones/eventos que se hayan detectado, determinación de los efectos visuales y activación de los mismos.

-Etapa 7: Aplicación de efectos visuales en la silueta que permite la generación de efectos visuales sobre la imagen del sujeto/objeto.

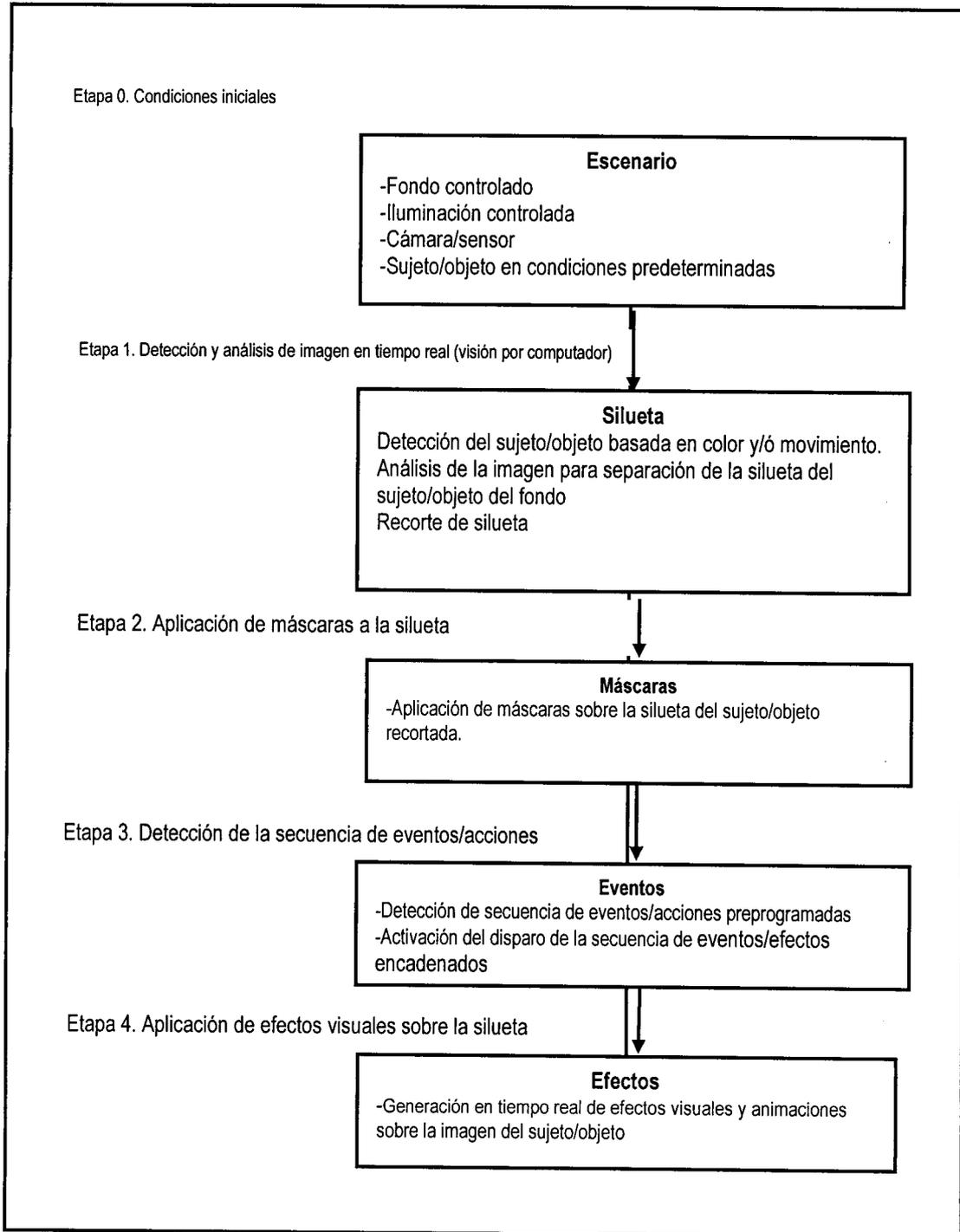
- 30 6. Procedimiento para la aplicación de efectos visuales en cuerpos o sujetos estáticos o en movimiento en tiempo real sin necesidad de marcas mediante el uso de un sistema según la reivindicación 2 caracterizado por las siguientes etapas:

- Etapa 0: Control y definición de las condiciones iniciales de: iluminación, resolución del dispositivo de entrada, fondo y situación del objeto en el lugar predefinido.
- 5 -Etapa 1: Detección en tiempo real de la imagen captada por el dispositivo de entrada, considerando color y/o movimiento mediante la aplicación de filtros de proyección y detección.
- Etapa 2: Análisis en tiempo real de la imagen para la separación de la silueta del sujeto/objeto del fondo mediante la aplicación de filtros de proyección y detección
- 10 -Etapa 3: Reconocimiento y almacenamiento de la silueta del sujeto/objeto.
- Etapa 4: Aplicación de máscaras sobre la silueta del sujeto/objeto mediante la aplicación de filtros de proyección y detección.
- Etapa 5: Detección de las secuencias de eventos/acciones preprogramadas.
- Etapa 6: Según las acciones/eventos que se hayan detectado, determinación de  
15 los efectos visuales y activación de los mismos.
- Etapa 7: Aplicación de efectos visuales en la silueta que permite la generación de efectos visuales sobre la imagen del sujeto/objeto.

Figura 1



**Figura 2**





- ②① N.º solicitud: 200901210  
 ②② Fecha de presentación de la solicitud: 07.05.2009  
 ③② Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TECNICA

⑤① Int. Cl.: Ver Hoja Adicional

DOCUMENTOS RELEVANTES

Categoría	⑤⑥ Documentos citados	Reivindicaciones afectadas
X A	EISERT P et al. 3-D Tracking of shoes for Virtual Mirror applications. Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, 23/06/2008 VOL: Pags: 1-6 ISBN 978-1-4244-2242-5; ISBN 1-4244-2242-6 ZHEN LEI; QINQUN BAI; RAN HE; LI S Z.	1-3 4-6
A	US 2006202986 A1 (OKADA RYUZO et al.) 14.09.2006, párrafos [31-44]; figuras 1-2.	1-6
A	US 2009116766 A1 (MATSUMOTO TAKASHI et al.) 07.05.2009, párrafos [7,10,15,37-39].	1-6
A	WO 03098922 A1 (RHODE ISLAND EDUCATION et al.) 27.11.2003, páginas 2-5; figura 2.	1-2

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  
27.02.2013

Examinador  
A. Astudillo Lizaga

Página  
1/4

CLASIFICACIÓN OBJETO DE LA SOLICITUD

**G06T19/00** (2011.01)

**G06T7/20** (2006.01)

**G06K9/00** (2006.01)

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, NPL, INSPEC, COMPENDEX

Fecha de Realización de la Opinión Escrita: 27.02.2013

**Declaración**

<b>Novedad (Art. 6.1 LP 11/1986)</b>	Reivindicaciones 4-6	<b>SI</b>
	Reivindicaciones 1-3	<b>NO</b>
<b>Actividad inventiva (Art. 8.1 LP11/1986)</b>	Reivindicaciones 4-6	<b>SI</b>
	Reivindicaciones 1-3	<b>NO</b>

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 se publica.

**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	EISERT P et al. 3-D Tracking of shoes for Virtual Mirror applications. Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, 23/06/2008 VOL: Pags: 1-6 ISBN 978-1-4244-2242-5; ISBN 1-4244-2242-6 ZHEN LEI; QINQUN BAI; RAN HE; LI S Z	23.06.2008
D02	US 2006202986 A1 (OKADA RYUZO et al.)	14.09.2006
D03	US 2009116766 A1 (MATSUMOTO TAKASHI et al.)	07.05.2009
D04	WO 03098922 A1 (RHODE ISLAND EDUCATION et al.)	27.11.2003

**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**

La solicitud de patente en estudio presenta seis reivindicaciones en las que las cuatro primeras describen un "sistema que permite la aplicación de efectos visuales en objetos o sujetos estáticos o en movimiento en tiempo real sin necesidad de marcas" que se caracteriza por: a) un dispositivo de entrada de captura de imágenes, b) un escenario de iluminación y fondo controlado, c) un dispositivo de procesamiento de imágenes y detección de movimiento en tiempo real, d) un conjunto de procedimientos para detectar, analizar y aplicar efectos visuales en objetos o sujetos en tiempo real mediante la aplicación de filtros de detección (filtro objeto y filtro de movimiento) y filtros de proyección (efecto halo, efecto desaturado, y reproducción de animación en el interior del contorno del objeto/sujeto). Las reivindicaciones 5 y 6 corresponden a un procedimiento de aplicación de efectos visuales en cuerpos o sujetos estáticos o en movimiento en tiempo real sin marcas que hace uso del sistema anterior.

Se considera como el documento del estado de la técnica más próximo al sistema y procedimiento reivindicados el documento D01.

Así, el documento D01 corresponde a un sistema que permite la visualización en tiempo real y de forma virtual de la imagen de un usuario mostrando diferentes calzados deportivos de diseño personalizado. **No requiere marcas** y dado que los algoritmos implementados **operan en tiempo real**, el usuario puede moverse libremente y observarse en una pantalla luciendo diferentes modelos de calzado recién diseñados (ver resumen e introducción, página 1). El sistema (ver descripción, página 2) está constituido por **una cámara y una pantalla de visualización**. Para simplificar la segmentación entre el usuario y el fondo en un entorno de iluminación cambiante, la parte del suelo frente a la cámara se pinta de verde para utilizar técnicas de "chroma keying" así como se incluye una luz adicional bajo la cámara para reducir los efectos de las sombras (**iluminación y fondo controlados**). Los distintos **algoritmos de procesamiento, detección de usuario** (considerando los "pixels" de color), **detección e interpretación de movimiento** (basados en la información de la silueta), y **aplicación de efectos visuales** (inserción de calzado virtual) **se ejecutan en un PC**. Este documento es relevante y afecta a la novedad de las reivindicaciones de sistema 1, 2 y 3 dado que incluye todos sus elementos técnicos (tal como se describe anteriormente). Dado que no se han localizado referencias al uso de filtros para la aplicación de efecto halo o desaturado, las reivindicaciones 4 de sistema y 5-6 de procedimiento sí tendrían novedad y actividad inventiva según lo especificado en los artículos 6 y 8 de la Ley de Patentes.

El documento D02 corresponde a un sistema, dispositivo y método para el modelado y prueba virtual de prendas de vestir y complementos, basado en una simulación tridimensional del usuario y una estimación de diferentes posturas y su interacción con la ropa o complementos, que permite visualizar en tiempo real una imagen lo más fidedigna posible de su aspecto personal (ver resumen, párrafos 0031-0032, 0037, 0041 y 0044). Este documento constituye estado de la técnica previo para las reivindicaciones 1-6 si bien no afecta a su novedad y/o actividad inventiva.

El documento D03 introduce un método y aparato que facilitan a un usuario una prueba virtual de prendas de vestir mediante un "espejo interactivo" así como una comparación simultánea (a través de varias pantallas adyacentes) con otras alternativas. El sistema, mediante técnicas de visión por ordenador, y el uso de un par de cámaras, detecta y almacena imágenes del usuario como posturas de partida para que posteriormente sean recuperadas de forma sincronizada con los movimientos del usuario (ver párrafos 0007, 0010, 0015 y 0037-0039). Este documento constituye estado de la técnica previo para las reivindicaciones 1-6 si bien no afecta a su novedad y/o actividad inventiva.

El documento D04 corresponde a un sistema de procesamiento de imágenes para la detección y seguimiento de un sujeto en movimiento mediante la substracción de imágenes consecutivas definida por un mapa de "pixels". Se determina un umbral y se calcula un centroide de los "pixels" de la imagen de substracción para obtener un vector de movimiento (ver páginas 2-5, figura 2). Este documento constituye estado de la técnica previo para las reivindicaciones 1 y 2 (filtro de movimiento) si bien no afecta a su novedad y/o actividad inventiva (su uso está dirigido al sector de la vigilancia).