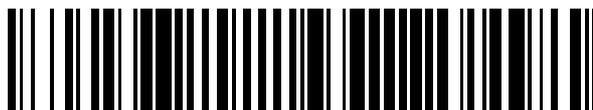


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 402 957**

51 Int. Cl.:

G06K 9/00 (2006.01)

G06T 7/00 (2006.01)

G06T 7/20 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.03.2008 E 08714340 (0)**

97 Fecha y número de publicación de la concesión europea: **02.01.2013 EP 2132680**

54 Título: **Rastreo eficaz y preciso de objetos tridimensionales**

30 Prioridad:

05.03.2007 AU 2007901129 P

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

10.05.2013

73 Titular/es:

**SEEING MACHINES PTY LTD (100.0%)
Innovations Building, Corner Garren & Eggleston
Road
Acton Australian Capital Territory 2601, AU**

72 Inventor/es:

TELL, DENNIS

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 402 957 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

Rastreo eficaz y preciso de objetos tridimensionales

5 Campo de la invención

La presente invención se refiere a procedimientos para la medición de la postura tridimensional (3D) de un objeto tridimensional.

10 La invención ha sido desarrollada principalmente para su uso en el contexto del rastreo del rostro humano y será descrita a continuación en la presente memoria con referencia a esta aplicación. Sin embargo, se apreciará que la invención no está limitada a este campo específico de uso.

Antecedentes de la invención

15 Cualquier exposición de la tecnología anterior a lo largo de la especificación no debería ser considerada en modo alguno como una admisión de que tal tecnología anterior sea ampliamente conocida, o de que forme parte del conocimiento general común en el campo.

El rastreo de rostros y rasgos faciales, tales como los ojos de una persona, ha atraído considerable interés durante la década anterior, según los ordenadores han llegado a ser lo suficientemente potentes como para habilitar soluciones prácticas para este problema.

20 Hay dos enfoques conocidos que han sido propuestos para resolver este problema. El primer enfoque es un enfoque geométrico, utilizando rasgos puntuales tridimensionales en el rostro y un razonamiento geométrico para obtener la postura tridimensional. El segundo es una optimización no lineal de los parámetros de un modelo de apariencia.

25 El primer enfoque, que usa rasgos puntuales, tiene la ventaja de ser determinístico. Los enfoques no iterativos brindan un tiempo breve y predecible para calcular la solución, y han sido populares para sistemas de tiempo real. Edwards et al (patente estadounidense N° 7.043.056) revelan una propuesta típica de esta metodología.

El segundo enfoque es un problema de optimización iterativa, no lineal, que, en general, es cara en términos de cálculo. Los trueques, en términos de precisión predecible de convergencia, son requeridos para lograr tiempos de cálculo predecibles. Según los parámetros de la modalidad de apariencia, una ventaja de este enfoque es que puede lograrse una mejor fidelidad de rastreo. Se entiende que la apariencia del objeto observado puede ser modelada y predicha más precisamente que con el enfoque de rasgos puntuales. *Cootes et al 2001* (T. Cootes, G. Edwards y C. Taylor, "Modelos de apariencia activa", Transacciones del IEEE sobre Análisis de Patrones e Inteligencia de Máquinas, 23(6):681-685, 2001), *Cootes et al 2002* (T. Cootes, G. Wheeler, K. Walker y C. Taylor, "Modelos de apariencia activa basados en vistas", Cálculo de Imagen y Visión, 20:657-664, 2002) y *Matthews et al* (I. Matthews y S. Baker, "Modelos de apariencia activa revisitados", Revista Internacional de Visión por Ordenador, Vol. 60, N° 2, Noviembre, 2004, págs. 135 a 164) revelan una típica implementación de esta metodología.

35 *De la Torre et al* (F. De la Torre y M. Black, "Análisis de componentes robusto y parametrizado: teoría y aplicaciones de modelos bidimensionales de apariencia facial", Visión por Ordenador y Comprensión de Imágenes 91 (2003) 53 a 71) revela el uso de un modelo de apariencia activa bidimensional específico de la persona, que no es capaz de rastrear una persona en tres dimensiones.

40 *Dornaika et al* (F. Dornaika y J. Ahlberg "Adaptación de modelo facial usando correlación robusta y modelos de apariencia activa", Anales del Sexto Taller del IEEE sobre Aplicaciones de la Visión por Ordenador, 2002 (WACV 2002), 3 a 7) y *Hu et al* (C. Hu, R. Feris y M. Turk "Redes de Ondículas Activas para Alineación de Rostros", Anales de la Conferencia Británica de Visión por Máquina, Norwich, 2003) revelan la división de una cara representada en sub-rasgos, pero no logran modelar ni explotar la naturaleza solapada de los rasgos faciales. Un ejemplo típico para estas oclusiones ocurre cuando el rostro es visto en una visión de semi-perfil, donde la cresta de la nariz forma un borde sobre la mejilla opuesta. Los sistemas anteriores están limitados a la vista no ocluida de los rasgos, y fallan en cuanto los rasgos comienzan a solaparse, debido a la proyección de los rasgos faciales en la imagen.

45 Debería apreciarse que los retos técnicos asociados a este problema son considerables, en particular, dado que un objeto como el rostro humano tiene grandes variaciones entre individuos, y el rostro es un objeto sumamente expresivo. Existe la necesidad en la técnica del rastreo automático de rostros humanos en secuencias de vídeo.

Resumen de la invención

55 Es un objeto de la presente invención proporcionar una forma mejorada del rastreo de objetos tridimensionales.

De acuerdo a un primer aspecto de la presente invención, se proporciona un procedimiento de rastreo de un objeto en un flujo de imágenes de entrada, comprendiendo el procedimiento aplicar iterativamente las etapas de: (a) representar un modelo de objeto tridimensional según un vector de estado anteriormente predicho, a partir de un bucle de rastreo previo, o del vector de estado de una etapa de inicialización; (b) extraer una serie de rasgos puntuales del objeto representado; (c) 60 localizar los correspondientes rasgos puntuales en el flujo de imágenes de entrada; (d) obtener un nuevo vector de estado predicho a partir de las ubicaciones de rasgos puntuales en el flujo de imágenes de entrada.

Este procedimiento también puede incluir etapas de inicialización de (i) creación de un modelo tridimensional del objeto a

rastrear; (ii) detección inicial de la posición de los objetos dentro de un flujo de imágenes de entrada.

La etapa (d) incluye preferiblemente la etapa adicional de: seleccionar un conjunto de rasgos congruentes, seguida por un filtrado de Kalman de los rasgos, para proporcionar el nuevo vector de estado.

5 El objeto puede comprender un rostro y el vector de estado puede comprender la postura tridimensional del rostro. En una realización, la selección de un conjunto de rasgos comprende la selección por consenso de muestras aleatorias (RANSAC) de los rasgos. La etapa de localización, preferiblemente, puede incluir utilizar un proceso de correlación cruzada normalizada para localizar los rasgos puntuales.

10 Debería apreciarse que las realizaciones de la presente invención han sido desarrolladas principalmente en el contexto del rastreo del rostro humano. Sin embargo, la presente invención es aplicable a otros dominios, donde la postura tridimensional de un objeto arbitrario ha de ser medida visualmente.

De acuerdo a un segundo aspecto de la presente invención, se proporciona un procedimiento según cualquier procedimiento anteriormente descrito del rastreo de un objeto en un flujo de imágenes de entrada, comprendiendo el procedimiento las etapas de:

- (i) crear un modelo tridimensional del objeto a rastrear;
- (ii) localizar puntos de rasgos iniciales en un flujo inicial de imágenes de entrada; y
- (iii) calcular un vector de estado inicial indicativo de la ubicación del objeto dentro del flujo de imágenes de entrada, en donde el vector de estado inicial se calcula minimizando el error de cuadrados entre los puntos de rasgos localizados iniciales y los correspondientes puntos de rasgos iniciales del modelo tridimensional proyectado en el plano de imágenes.

En el procedimiento, optativamente, en la etapa (a), el modelo del objeto concuerda, bien con el vector de estado predicho calculado en la etapa (d) de una iteración previa, o bien con el vector de estado inicial calculado en la etapa (ii), en donde la representación incluye calcular una máscara para el flujo de imágenes de entrada, para distinguir entre píxeles de fondo y de frente; y / o incluye, en la etapa (b), calcular un número predefinido de rasgos puntuales del objeto representado, en donde el número predefinido de ubicaciones con la mayor angulosidad son seleccionadas como rasgos a partir de la imagen representada de la iteración previa para la siguiente etapa de localización.

La etapa de localización, preferiblemente, usa una correlación de plantilla enmascarada normalizada para calcular los nuevos rasgos puntuales localizados.

El nuevo vector de estado, preferiblemente, es calculado usando la selección por Consenso de Muestras Aleatorias, y seguido por un filtro de Kalman, para estimar el nuevo vector de estado. El filtro de Kalman, preferiblemente, usa un modelo de movimiento de aceleración constante del objeto para predecir el nuevo vector de estado del objeto a partir de vectores de estado previamente calculados y de los nuevos rasgos puntuales localizados.

De acuerdo a un tercer aspecto de la revelación, se proporciona un programa de ordenador o producto de programa de ordenador que comprende un conjunto de instrucciones para llevar a cabo un procedimiento según cualquier procedimiento previamente descrito del rastreo de un objeto en un flujo de imágenes de entrada.

De acuerdo a un cuarto aspecto de la revelación, se proporciona un medio portador legible por ordenador que lleva un conjunto de instrucciones que, cuando son ejecutadas por uno o más procesadores, hacen que uno o más procesadores lleven a cabo un procedimiento de rastreo de un objeto en un flujo de imágenes de entrada, según lo anteriormente descrito.

El medio portador, preferiblemente, puede ser un medio que lleva una señal propagada, detectable por al menos un procesador entre dichos uno o más procesadores, y que representa al conjunto de instrucciones.

De acuerdo a un quinto aspecto de la revelación, se proporciona un sistema para rastrear un objeto en un flujo de imágenes de entrada, comprendiendo el sistema un procesador adaptado para recibir un flujo de imágenes de entrada, y el procesador está adicionalmente adaptado para realizar un procedimiento de rastreo de un objeto en un flujo de imágenes de entrada, según lo anteriormente descrito.

De acuerdo a un sexto aspecto de la presente invención, se proporciona un sistema para rastrear un objeto en un flujo de imágenes de entrada, comprendiendo el sistema un procesador adaptado para recibir un flujo de imágenes de entrada, y el procesador está adicionalmente adaptado para realizar las etapas de:

- (i) crear un modelo tridimensional del objeto a rastrear; (ii) localizar puntos de rasgos iniciales en un flujo inicial de imágenes de entrada; (iii) calcular un vector de estado inicial indicativo de la ubicación del objeto dentro del flujo de imágenes de entrada, en donde el vector de estado inicial se calcula minimizando el error de cuadrados entre los puntos de rasgos localizados iniciales y los correspondientes puntos de rasgos iniciales del modelo tridimensional proyectado en el plano de imágenes;

(a) representar un modelo de objeto tridimensional, en donde el modelo de objeto concuerda, bien con el vector de estado predicho calculado en la etapa (d) de una iteración previa, o bien el vector de estado inicial calculado en la etapa (ii), en

donde la representación incluye calcular una máscara para el flujo de imágenes de entrada, para distinguir entre píxeles de fondo y de frente;

(b) calcular un número predefinido de rasgos puntuales a partir del objeto representado, en donde el número predefinido de ubicaciones con la mayor angulosidad son seleccionadas como rasgos a partir de la imagen representada de la iteración previa para la siguiente etapa de localización;

(c) localizar los correspondientes rasgos puntuales en el flujo de imágenes de entrada;

(d) calcular un nuevo vector de estado a partir de los rasgos puntuales localizados en el flujo de imágenes de entrada; y

(e) realizar iterativamente las etapas (a) a (d) para proporcionar en cada iteración el nuevo vector de estado actualizado a partir de los rasgos puntuales localizados.

El procesador, preferiblemente, está adaptado para realizar una correlación de plantilla normalizada enmascarada, para calcular nuevos rasgos puntuales localizados. El procesador, preferiblemente, está adaptado para aplicar los nuevos rasgos puntuales localizados en un filtro de Kalman, para calcular el nuevo vector de estado.

La invención tiene aplicación específica en la valoración del rendimiento humano, la evaluación de la ergonomía del diseño de interfaces hombre-máquina, en aplicaciones de seguridad para monitorizar la distracción y la fatiga, y en aplicaciones de entretenimiento, donde las mediciones son tomadas a partir de una persona real, para ser luego representadas por un ordenador en un avatar.

Breve descripción de los dibujos

Se describirá ahora una realización preferida de la invención, solamente a modo de ejemplo, con referencia a los dibujos adjuntos, en los cuales:

la FIG. 1 es un diagrama de flujo de un procedimiento según una realización de la invención;

la FIG. 2 es un diagrama de flujo de un procedimiento similar al de la Fig. 1, que muestra etapas de inicializaciones; la FIG. 3 es una vista esquemática de un sistema según una realización de la invención; y

la FIG. 4 es una vista esquemática de un sistema según una realización de la invención, mostrada al rastrear un objeto tridimensional.

Descripción de la realización preferida, y otras

El procedimiento propuesto incorpora las ventajas tanto de la alta precisión en la apariencia predicha (y, por lo tanto, de alta fidelidad) como del cómputo rápido y predecible del cálculo de rasgos puntuales. La realización preferida se refiere a procedimientos para una medición visual sumamente precisa, y a la vez eficaz, de la postura tridimensional de un objeto tridimensional. Debería apreciarse que la postura tridimensional es una medida indicativa de una ubicación y orientación tridimensional, habitualmente representadas como seis parámetros. El procedimiento utiliza componentes de procedimientos de rastreo basados en rasgos, con procedimientos basados en la apariencia, denominados "modelos de apariencia activa". El resultado es un procedimiento sumamente eficaz y preciso para rastrear y medir visualmente la postura de un objeto.

Panorama general de la realización preferida

Con referencia a la FIG. 1, se revela una representación 100 en diagrama de flujo de un procedimiento según una realización de la presente invención, y este procedimiento comprende las etapas de:

(a) representar un modelo de rostro tridimensional según un vector de estado predicho; 120

(b) extraer un cierto número de rasgos puntuales de la imagen del rostro representado; 140

(c) localizar rasgos puntuales; 160 y

(d) obtener un nuevo vector de estado a partir de las ubicaciones 180 de rasgos puntuales.

Con referencia a la FIG. 2, una etapa inicial de un procedimiento 200, según la realización preferida (y otros procedimientos relacionados), incluye crear un modelo 210 tridimensional adecuado de un rostro que ha de ser rastreado – o, equivalentemente, un objeto que ha de ser rastreado. Dentro de este procedimiento, la caracterización del modelo depende del tipo de procedimiento de rastreo. Solamente a modo de ejemplo, el modelo usado en el procedimiento según la realización preferida consiste en un modelo tridimensional texturizado de un rostro. Este modelo permite una representación por ordenador del rostro en cualquier posición y orientación. La localización de un rostro dentro de una imagen 215 proporciona una ubicación facial estimada inicial. Una vez que la ubicación facial está identificada dentro de una imagen, la textura es correlacionada con un modelo tridimensional genérico de un rostro, proyectando la textura sobre los vértices. Solamente a modo de ejemplo, *Loy et al* (Tratado de Cooperación en Patentes, Solicitud N° PCT / AU 2003 / 000345, Número de Publicación WO / 2003 / 081532) revela un procedimiento usado para identificar la ubicación facial dentro de una imagen.

Debería ser apreciado por una persona experta en la técnica que la textura es correlacionada con un modelo tridimensional genérico de un rostro, proyectando la textura sobre los vértices. Solamente a modo de ejemplo, un procedimiento de correlación de textura se describe en *Dornaika et al.*

5 Después de que las etapas anteriores de inicialización se han realizado, se ingresa a un bucle principal de rastreo. Este bucle de rastreo utiliza el modelo tridimensional texturizado y estima el vector de estado del rostro para cada imagen en un flujo de imágenes. Las etapas implicadas en este bucle de rastreo se describen a continuación:

10 (1) Representar un modelo de rostro tridimensional según un vector de estado predicho, 120. Este vector de estado predicho, preferiblemente, se obtiene de un bucle de rastreo previo o de un vector de estado proveniente de las etapas de inicialización.

(2) Extraer un cierto número de rasgos puntuales de la imagen facial representada, 140.

15 (3) Localizar rasgos puntuales, 160. La localización de rasgos puntuales se obtiene con un procedimiento adecuado de localización de rasgos puntuales. En la realización preferida, el procedimiento de localización de rasgos puntuales es la correlación cruzada normalizada.

(4) Obtener un nuevo vector de estado a partir de las ubicaciones 180 de rasgos puntuales. En la realización preferida, la obtención de este nuevo vector de estado a partir de las ubicaciones de rasgos puntuales se realiza mediante la selección por "Consenso de Muestras Aleatorias" (RANSAC) de un conjunto de rasgos congruentes, y es seguida por un filtro de Kalman, para estimar el nuevo vector de estado que, en el caso actual, es la postura tridimensional.

20 Este bucle de rastreo, que comprende las etapas (1) a (4), puede ser iterado indefinidamente, o bien hasta que falle el rastreo del rostro y se requiera una nueva inicialización.

Solamente a modo de ejemplo, se describen a continuación etapas individuales en más detalle.

25 *Etapas (0) – Inicialización*

La inicialización requiere la localización del rostro como un todo y rasgos individuales clave. En la realización preferida, un procedimiento revelado por *Loy et al* es utilizado para la producción de esta localización. Este procedimiento permite la localización del rostro, la ubicación de los cuatro vértices oculares, las fosas nasales, si son visibles, y las dos comisuras de la boca. Estos rasgos son representados como puntos $R = \{r_i \in R^3\}$ de referencia del modelo.

30 Dado que este modelo tridimensional usa un modelo tridimensional rígido, solamente se necesita estimar seis parámetros de postura (por ejemplo, que incluyan tres de traslación y tres de rotación) para ajustar un modelo genérico tridimensional de geometría facial a la imagen. Este modelo genérico 'M' comprende los vértices v_i de esquina y las superficies s_j planas triangulares, y aproxima una geometría facial media. Este modelo genérico puede ser expresado en la siguiente ecuación.

$$35 \quad M = \{v_i \in R^3, s_j\}$$

Debería ser apreciado por una persona experta en la técnica que la representación de una superficie como una pluralidad de triángulos adyacentes es una técnica usada en la representación y modelización gráfica tridimensional. Sin embargo,

40 también debería ser apreciado que otros modelos son adecuados para su uso en la presente invención. Solamente a modo de ejemplo, este modelo captura un rostro desde la frente hasta la barbilla (habitualmente, en la dirección vertical), y el área hasta las orejas (habitualmente en la dirección horizontal). Este modelo excluye el pelo de la persona.

45 La postura tridimensional $P = \{r_p \in R^3, t_p \in R^3\}$, que comprende los componentes de rotación $r_p \in R^3$ y los componentes de traslación $t_p \in R^3$, es estimada alineando el modelo tridimensional con los rasgos de imagen localizados. Solamente a modo de ejemplo, esta alineación puede hacerse de modo que el error de cuadrados entre los puntos de referencia del modelo tridimensional proyectado en el plano $R(p)$ de imagen y las ubicaciones de imágenes de los puntos R_i de imagen sea minimizado. Esta minimización de alineación puede ser representada por la siguiente ecuación.

$$50 \quad P_{inic} = \underset{r, t}{\operatorname{argmin}} \left(\sum \|R(p) - R_i\| \right)$$

Dornaika et al describe un procedimiento que usa una hipótesis de perspectiva débil, que es utilizada en la realización preferida. Debería ser apreciado por una persona experta en la técnica que hay un cierto número de técnicas de optimización que pueden ser usadas para realizar la alineación de modelos.

55 La textura para cada triángulo del modelo puede ser extraída de una imagen y almacenada con un modelo geométrico tridimensional, combinando la geometría tridimensional y la textura para determinar un modelo tridimensional requerido para el rastreo. La postura \bar{P} de cabeza predicha después de la inicialización es equivalente a la postura P_{inic} del modelo tridimensional ajustado.

Etapa (1) – Representación

5 En la etapa de representación, el modelo facial tridimensional texturizado es habitualmente representado según la postura de la cabeza, según lo predicho por la etapa de estimación de estado (o la etapa de inicialización si esta es la primera iteración). Esto puede ser calculado usando procedimientos comunes de gráficos por ordenador, conocidos para una persona experta en la técnica.
 Debería ser apreciado que puede usarse hardware (por ejemplo, una tarjeta de gráficos tridimensionales) y / o software (por ejemplo, OpenGL) para acelerar esta etapa. La persona experta en la técnica está al tanto de estas técnicas.
 10 La salida de esta etapa incluye una imagen facial representada y una máscara para distinguir entre los píxeles que son del fondo y los píxeles que son del frente.

Etapa (2) – Extracción de rasgos

15 Se selecciona un conjunto de rasgos $F = \{f_k\}$ para su uso en las siguientes etapas. En la realización preferida, el conjunto de rasgos son dinámicos y pueden variar en cada ciclo de rastreo para optimizar la robustez.
 Debería apreciarse que, solamente a modo de ejemplo, los procedimientos de extracción de rasgos pueden usar bien una imagen facial representada como imagen entera (por ejemplo, según lo revelado en *Cootes et al 2001* y *Cootes et al 2002*), los vértices extraídos del modelo (por ejemplo, según lo revelado en *Dornaika et al*), o bien conjuntos de rasgos extraídos de una vista frontal sin oclusiones.
 Un inconveniente asociado al uso del rostro representado entero es que el vector de estado solamente puede ser calculado de manera iterativa y cara en términos de cálculo. Según la métrica de distancia usada para establecer la similitud de la imagen representada con la imagen en vivo de la cámara, este enfoque habitualmente tiene una significativa sensibilidad a condiciones de iluminación no homogéneas. Los vértices o rasgos pre-extraídos pueden carecer de la capacidad de modelar y explotar los bordes creados a partir de la auto-oclusión parcial del rostro, debido a la oclusión, por la cresta de la nariz, de la mejilla opuesta cuando la cabeza se gira levemente.
 25 Para superar estas limitaciones, el procedimiento revelado en la presente memoria usa rasgos calculados a partir de la imagen representada en una iteración previa de rastreo. Un número predefinido de ubicaciones con la mayor angulosidad se seleccionan como rasgos provenientes de la imagen representada de la iteración anterior, para la siguiente etapa de localización. La angulosidad 'e' de la ubicación (x, y) en la imagen 'I' se define como la magnitud del gradiente de la ubicación de la imagen, según lo calculado usando la siguiente ecuación.

$$e(I(x,y)) = \left| \frac{I(x,y)}{dxdy} \right|$$

35 Se realiza una búsqueda exhaustiva para hallar los f_k con los mayores valores de gradiente en la imagen. La salida de esta etapa incluye un conjunto de rasgos de rastreo $F = \{f_k\}$.

Etapa (3) – Localización de rasgos

40 Los rasgos seleccionados en la etapa anterior deben ser localizados en una imagen recientemente adquirida del rostro. Debería ser apreciado por una persona experta en la técnica que esta localización podría ser calculada usando un cierto número de procedimientos usados para la localización y rastreo de rasgos. En la realización preferida, se usa una correlación de plantilla enmascarada normalizada para calcular la nueva ubicación de los rasgos.

45 Para la localización, se extrae una plantilla de 16x16 píxeles, centrada en la ubicación f_k del rasgo, de la imagen representada, y se enmascara con la máscara de representación del rostro, para excluir todo píxel de fondo que pudiera ser parte de la región definida por la plantilla de 16x16 píxeles.

Una persona experta en la técnica está al tanto de la formulación, comúnmente usada, de correlación cruzada normalizada, según lo expresado en la siguiente ecuación.

$$C(u,v) = \frac{\sum_{x,y} (I(u+x, v+y) - \bar{I}_{u,v}) \hat{T}(x,y)}{\sum_{x,y} (I(u+x, v+y) - \bar{I}_{u,v})^2 \sum_{x,y} \hat{T}(x,y)^2}$$

50 En esta ecuación, I es la imagen y T es la plantilla. \hat{T} es la plantilla normalizada de brillo $\hat{T} = T(x,y) - \bar{T}$ y \bar{T} es la media de la plantilla. $\bar{I}_{u,v}$ es la media de la imagen local en la ubicación (u,v), la media de la imagen bajo la plantilla cuya esquina superior izquierda está sobre el píxel (u,v). Además, las sumas y la media de la imagen están sujetas al enmascaramiento,

es decir, sin considerar los píxeles que están excluidos por enmascaramiento.

Alrededor de la posición predicha de cada rasgo, se sitúa una ventana de búsqueda. Dentro de esta ventana de búsqueda, se realiza una búsqueda exhaustiva para hallar la ubicación del rasgo con la precitada técnica de correlación. La ubicación de ese punto (o píxel) con el mayor puntaje de correlación es seleccionada como la nueva ubicación f_k de rasgo.

La salida de la etapa de localización de rasgos incluye las nuevas ubicaciones F de rasgos para el conjunto de rasgos F seleccionados.

Etapas (4) – Estimación de Estados

En la realización preferida, solamente a modo de ejemplo, el vector de estado consiste solamente en los seis parámetros tridimensionales de postura. Este vector de estado puede ser calculado aplicando las nuevas ubicaciones de rasgos en un filtro de Kalman, del tipo revelado por *Edwards et al.* Este filtro de Kalman utiliza un modelo de movimiento de aceleración constante del rostro para predecir la nueva posición del rostro a partir de estimaciones anteriores y las mediciones recientemente adquiridas. El filtro de Kalman también puede permitir la predicción del vector \bar{P} de estado de postura de la cabeza para la próxima trama de imagen.

La salida de la etapa de estimación de estados incluye el vector \bar{P} de estado predicho, que es remitido a la etapa de representación en la próxima iteración del algoritmo. Debería ser apreciado que las etapas (1) a (4) pueden ser aplicadas iterativamente para proporcionar un procedimiento relativamente eficaz y preciso para el rastreo de objetos tridimensionales.

Solamente a modo de ejemplo, una realización de un procedimiento de rastreo de un objeto en un flujo de imágenes de entrada comprende las etapas de:

- (i) crear un modelo tridimensional del objeto a rastrear;
- (ii) localizar puntos iniciales de rasgos en un flujo inicial de imágenes de entrada;
- (iii) calcular un vector de estado inicial indicativo de la ubicación del objeto dentro del flujo de imágenes de entrada, en donde el vector de estado inicial es calculado minimizando el error de cuadrados entre los puntos iniciales de rasgos localizados y los correspondientes puntos de rasgos iniciales del modelo tridimensional proyectado en el plano de imágenes;

(a) representar un modelo de objeto tridimensional, en donde el modelo de objeto concuerda, bien con el vector de estado predicho calculado en la etapa (d) de una iteración previa, o bien con el vector de estado inicial calculado en la etapa (ii), en donde la representación incluye calcular una máscara para el flujo de imágenes de entrada, para distinguir entre píxeles de fondo y de frente;

(b) calcular un número predefinido de rasgos puntuales a partir del objeto representado, en donde el número predefinido de ubicaciones con la mayor angulosidad son seleccionadas como rasgos a partir de la imagen representada de la iteración previa, para la siguiente etapa de localización;

(c) localizar los correspondientes rasgos puntuales en el flujo de imágenes de entrada;

(d) calcular un nuevo vector de estado a partir de los rasgos puntuales localizados en el flujo de imágenes de entrada; y

(e) realizar iterativamente las etapas (a) a (d) para proporcionar en cada iteración el nuevo vector de estado actualizado a partir de los rasgos puntuales localizados.

Debería apreciarse adicionalmente que este procedimiento puede ser realizado en tiempo real o sobre datos de vídeo en vivo.

Debería apreciarse que los procedimientos anteriores pueden ser realizados, al menos en parte, por un programa de ordenador. Un programa de ordenador, o producto de programa de ordenador, que comprende un conjunto de instrucciones para llevar a cabo un procedimiento según lo descrito anteriormente. Un medio portador legible por ordenador puede realizar los procedimientos anteriores. Este medio portador legible por ordenador lleva un conjunto de instrucciones que, cuando son ejecutadas por uno o más procesadores, hacen que uno o más procesadores lleven a cabo un procedimiento según un procedimiento anteriormente descrito. Este medio portador es un medio que lleva una señal propagada, detectable por el menos un procesador entre dichos uno o más procesadores, y que representa al conjunto de instrucciones.

La FIG. 3 revela una vista esquemática de un sistema según una realización. Un sistema 310 de ordenador incluye un procesador 312 adaptado para realizar las etapas de:

- (a) representar un modelo facial tridimensional según un vector de estado predicho;
- (b) extraer un cierto número de rasgos puntuales de la imagen facial representada;
- (c) localizar rasgos puntuales; y

(d) obtener un nuevo vector de estado a partir de las ubicaciones de rasgos puntuales.

En una realización, el procesador 312 está adaptado para recibir datos de modelos desde un dispositivo 314 de almacenamiento de datos. Estos datos de modelos son indicativos de un modelo tridimensional adecuado de un rostro (u objeto) que ha de ser rastreado. Debería apreciarse que, en una realización alternativa, los datos de modelos pueden ser recibidos desde un dispositivo externo de almacenamiento de datos.

En una realización, el procesador 312 está adicionalmente adaptado para recibir datos de vídeo en vivo desde un dispositivo 320 de captura de vídeo. Debería apreciarse que, en una realización alternativa, los datos de vídeo pueden ser recibidos desde un dispositivo de almacenamiento de datos (por ejemplo, el dispositivo 314 de almacenamiento de datos). El procesador está adaptado para localizar un rostro (u objeto) dentro de una imagen para proporcionar una ubicación facial estimada inicial. Una vez que la ubicación facial está identificada dentro de una imagen, la textura se correlaciona con un modelo tridimensional genérico de un rostro, proyectando la textura sobre los vértices.

La FIG. 4 revela una vista esquemática de una realización que muestra el rastreo de un objeto tridimensional 410 en tiempo real. En esta realización, un sistema 310 de ordenador está acoplado con un dispositivo de captura de vídeo, en forma de una cámara 320 de vídeo. La cámara de vídeo está adaptada para transmitir datos de vídeo, que son recibidos por el sistema de ordenador. El sistema de ordenador realiza un procedimiento de medición de la postura tridimensional (3D) de un objeto tridimensional, según lo anteriormente revelado. Los datos indicativos de la postura tridimensional medida pueden ser emitidos en un dispositivo 420 de salida. Debería apreciarse que, en una realización alternativa, los datos indicativos de la postura tridimensional medida pueden ser almacenados en un dispositivo de almacenamiento de datos, para su extracción posterior.

Variaciones

Un cierto número de variaciones de la realización anterior pueden incluir el uso de más de una cámara y el uso de procedimientos alternativos empleados en las etapas individuales de cálculo.

A menos que el contexto requiera claramente lo contrario, a lo largo de la descripción y las reivindicaciones, las palabras "comprender", "comprende" y similares han de ser interpretadas en un sentido inclusivo, a diferencia de un sentido exclusivo o exhaustivo; es decir, en el sentido de "incluso, pero no limitado a".

Según se usa en la presente memoria, a menos que se especifique lo contrario, el uso de los adjetivos ordinales "primero", "segundo", "tercero", etc., para describir un objeto común, indica meramente que se están mencionando distintas instancias de objetos iguales, y no están concebidos para implicar que los objetos así descritos deben estar en una secuencia dada, ya sea temporalmente, espacialmente, en clasificación, o de cualquier otra manera.

A menos que se indique específicamente lo contrario, según lo manifiesto a partir de las siguientes exposiciones, ha de apreciarse que, a lo largo de la especificación, las exposiciones que utilizan términos tales como "procesamiento", "cómputo", "cálculo", "determinación", "aplicación", "obtención" o similares, se refieren a la acción y / o a los procesos de un ordenador o sistema informático, o equipo informático electrónico similar, que manipula y / o transforma datos representados como cantidades físicas, tal como electrónicas, en otros datos, similarmente representados como cantidades físicas.

De manera similar, el término "procesador" puede referirse a cualquier dispositivo, o parte de un dispositivo, que procesa datos electrónicos, p. ej., de registros y / o memoria, para transformar esos datos electrónicos en otros datos electrónicos que, p. ej., pueden ser almacenados en registros y / o memoria. Un "ordenador", o un "sistema informático", o una "máquina calculadora", o una "plataforma de cálculo", puede incluir uno o más procesadores.

Se entenderá que las etapas de procedimientos expuestos son realizados en una realización por un procesador (o procesadores) adecuado(s) de un sistema (es decir, un ordenador) de procesamiento que ejecuta instrucciones (código legible por ordenador) almacenadas en un almacenamiento. También se entenderá que la invención no está limitada a ninguna implementación o técnica de programación específica, y que la invención puede ser implementada usando técnicas adecuadas cualesquiera para implementar la funcionalidad descrita en la presente memoria. La invención no está limitada a ningún lenguaje de programación ni sistema operativo específico.

Debería apreciarse que algunas de las realizaciones son descritas en la presente memoria como un procedimiento o combinación de elementos de un procedimiento que pueden ser implementados por uno o más procesadores de un sistema informático, o por otro medio de llevar a cabo la función. De tal modo, un procesador con las instrucciones necesarias para llevar a cabo un tal procedimiento, o elemento de un procedimiento, forma un medio para llevar a cabo el procedimiento, o elemento de un procedimiento. Además, un elemento descrito, en la presente memoria, de una realización de un aparato, es un ejemplo de un medio para llevar a cabo la función realizada por el elemento, con el fin de llevar a cabo la invención.

5 En realizaciones alternativas, el sistema informático que comprende uno o más procesadores funciona como un dispositivo autónomo, o puede ser configurado, p. ej., en red con otro(s) procesador(es), en un despliegue en red. Dicho(s) procesador(es) puede(n) funcionar en capacidad de servidor o máquina cliente en un entorno de red de servidor y cliente, o como una máquina del mismo rango en un entorno de red de igual a igual, o distribuida.

De tal modo, una realización de cada uno de los procedimientos descritos en la presente memoria está en forma de un medio portador legible por ordenador, que lleva un conjunto de instrucciones, p. ej., un programa de ordenador que ha de ejecutarse en uno o más procesadores.

10 La referencia, a lo largo de esta especificación, a “una realización” significa que un rasgo, estructura o característica específica, descrita con respecto a la realización, está incluida en al menos una realización. Así, no todas las apariciones de la frase “en una realización” en diversos lugares a lo largo de esta especificación están refiriéndose necesariamente a la misma realización, pero pueden referirse a la misma realización. Además, los rasgos, estructuras o características específicas pueden combinarse de cualquier manera adecuada, como será evidente para alguien medianamente experto en la técnica a partir de esta revelación, en una o más realizaciones.

20 De manera similar, debería apreciarse que, en la descripción anterior de realizaciones ejemplares de la invención, diversas características de la invención se agrupan a veces entre sí en una única realización, figura o descripción de la misma, con el fin de afinar la revelación y ayudar a la comprensión de uno o más de los diversos aspectos inventivos. Este procedimiento de revelación, sin embargo, no ha de interpretarse como reflejo de una intención de que la invención reivindicada requiere más características que las expresamente indicadas en cada reivindicación. Antes bien, como reflejan las siguientes reivindicaciones, los aspectos inventivos se apoyan en menos de todas las características de una única realización precedente revelada. Así, las reivindicaciones a continuación de la Descripción Detallada son por lo tanto expresamente incorporadas en esta Descripción Detallada, destacándose cada reivindicación como una realización individual de esta invención.

30 Además, si bien algunas realizaciones descritas en la presente memoria incluyen algunas de, pero no todas, las otras características incluidas en otras realizaciones, las combinaciones de características de distintas realizaciones están concebidas para estar dentro del alcance de la invención, y formar distintas realizaciones, como entenderán los expertos en la técnica. Por ejemplo, en las siguientes reivindicaciones, cualquiera entre las realizaciones reivindicadas puede ser usada en cualquier combinación.

35 En la descripción proporcionada en la presente memoria, se estipulan numerosos detalles específicos. Sin embargo, se entiende que las realizaciones de la invención pueden ser puestas en práctica sin estos detalles específicos. En otros casos, procedimientos, estructuras o técnicas bien conocidas no han sido mostrados en detalle, a fin de no oscurecer la comprensión de esta descripción.

40 Aunque la invención ha sido descrita con referencia a ejemplos específicos, los expertos en la técnica apreciarán que la invención puede ser realizada de muchas otras maneras.

REIVINDICACIONES

1. Un procedimiento de rastreo de un objeto en un flujo de imágenes de entrada, comprendiendo dicho procedimiento aplicar iterativamente las etapas de:
- 5 (a) representar un modelo de objeto tridimensional, según un vector de estado predicho previamente a partir de un bucle de rastreo anterior, o un vector de estado proveniente de una etapa de inicialización;
- (b) extraer una serie de rasgos puntuales de dicho objeto representado;
- (c) localizar los correspondientes rasgos puntuales en dicho flujo de imágenes de entrada;
- 10 (d) obtener un nuevo vector de estado predicho a partir de dicho rasgo puntual localizado en el flujo de imágenes de entrada.
2. Un procedimiento según la reivindicación 1, que comprende adicionalmente una etapa de inicialización de
- (i) creación de un modelo tridimensional de dicho objeto a rastrear;
- 15 (ii) detección inicial de la posición de dichos objetos dentro de dicho flujo de imágenes de entrada.
3. Un procedimiento según la reivindicación 1, en el cual dicha etapa (d) incluye adicionalmente la etapa de seleccionar un conjunto de rasgos congruentes, seguida por un filtrado de Kalman de dichos rasgos para obtener dicho nuevo vector de estado.
- 20 4. Un procedimiento según la reivindicación 1, en el cual dicho objeto comprende un rostro.
5. Un procedimiento según la reivindicación 4, en el cual dicho vector de estado comprende la postura tridimensional de dicho rostro.
- 25 6. Un procedimiento según la reivindicación 3, en el cual la selección de un conjunto de rasgos comprende la selección por consenso de muestras aleatorias (RANSAC) de dichos rasgos.
7. Un procedimiento según la reivindicación 1, en el cual dicha etapa de localización incluye utilizar un proceso normalizado de correlación cruzada para localizar dichos rasgos puntuales.
- 30 8. Un procedimiento según la reivindicación 1, que incluye las etapas de inicialización de:
- (i) crear un modelo tridimensional de dicho objeto a rastrear;
- (ii) localizar puntos iniciales de rasgos en dicho flujo inicial de imágenes de entrada; y
- 35 (iii) calcular un vector de estado inicial indicativo de dicha ubicación de objeto dentro de dicho flujo de imágenes de entrada, en donde dicho vector de estado inicial es calculado minimizando el error de cuadrados entre dichos puntos iniciales de rasgos localizados y los correspondientes puntos iniciales de rasgos de dicho modelo tridimensional proyectado en el plano $R(p)$ de imágenes.
- 40 9. Un procedimiento según la reivindicación 8, en el cual, en la etapa (a), dicho modelo de objeto concuerda, bien con dicho vector de estado predicho calculado en la etapa (d) de una iteración anterior, o bien con dicho vector de estado inicial calculado en la etapa (ii), y en el cual dicha representación incluye calcular una máscara para dicho flujo de imágenes de entrada, para distinguir entre píxeles de fondo y de frente.
- 45 10. Un procedimiento según la reivindicación 9, en el cual la etapa (b) incluye calcular un número predefinido de rasgos puntuales a partir de dicho objeto representado, en el cual dicho número predefinido de ubicaciones con la mayor angulosidad son seleccionadas como rasgos de dicha imagen representada de la iteración anterior, para la siguiente etapa de localización.
- 50 11. Un procedimiento según la reivindicación 8, en el cual dicha etapa de localización usa una correlación de plantilla normalizada enmascarada para calcular dichos nuevos rasgos puntuales localizados.
12. Un procedimiento según la reivindicación 8 o la reivindicación 9, en el cual dicho nuevo vector de estado es calculado usando la selección por Consenso de Muestras Aleatorias, y seguida por un filtro de Kalman para estimar dicho nuevo vector de estado.
- 55 13. Un procedimiento según la reivindicación 12, en el cual dicho filtro de Kalman utiliza un modelo de movimiento de aceleración constante de dicho objeto para predecir dicho nuevo vector de estado de dicho objeto a partir de vectores de estado anteriormente calculados, y dichos nuevos rasgos puntuales localizados.
- 60 14. Un sistema para rastrear un objeto en un flujo de imágenes de entrada, comprendiendo el sistema un procesador adaptado para recibir un flujo de imágenes de entrada, y dicho procesador está adicionalmente adaptado para realizar las

etapas de:

- (i) crear un modelo tridimensional del objeto a rastrear;
 - (ii) localizar puntos iniciales de rasgos en dicho flujo inicial de imágenes de entrada;
 - 5 (iii) calcular un vector de estado inicial indicativo de la ubicación de dicho objeto dentro de dicho flujo de imágenes de entrada, en donde dicho vector de estado inicial es calculado minimizando el error de cuadrados entre dichos puntos iniciales de rasgos localizados y los correspondientes puntos de rasgos iniciales de dicho modelo tridimensional proyectado en el plano de imágenes $R(p)$;
 - 10 (a) representar un modelo de objeto tridimensional, en donde dicho modelo de objeto concuerda, bien con dicho vector de estado predicho calculado en la etapa (d) de una iteración previa, o bien con dicho vector de estado inicial calculado en la etapa (ii), en donde dicha representación incluye calcular una máscara para dicho flujo de imágenes de entrada, para distinguir entre píxeles de fondo y de frente;
 - 15 (b) calcular un número predefinido de rasgos puntuales a partir de dicho objeto representado, en donde dicho número predefinido de ubicaciones con la mayor angulosidad son seleccionadas como rasgos a partir de dicha imagen representada de la iteración previa para la siguiente etapa de localización;
 - (c) localizar los correspondientes rasgos puntuales en dicho flujo de imágenes de entrada;
 - (d) calcular un nuevo vector de estado a partir de dichos rasgos puntuales localizados en dicho flujo de imágenes de entrada; y
 - 20 (e) realizar iterativamente las etapas (a) a (d) para proporcionar en cada iteración dicho nuevo vector de estado actualizado a partir de dichos rasgos puntuales localizados.
15. Un sistema según la reivindicación 14, en el cual dicho procesador está adaptado para realizar una correlación de plantilla normalizada enmascarada para calcular dichos nuevos rasgos puntuales localizados.
- 25 16. Un sistema según la reivindicación 14, en el cual dicho procesador está adaptado para aplicar dichos nuevos rasgos puntuales localizados en un filtro de Kalman, para calcular dicho nuevo vector de estado.

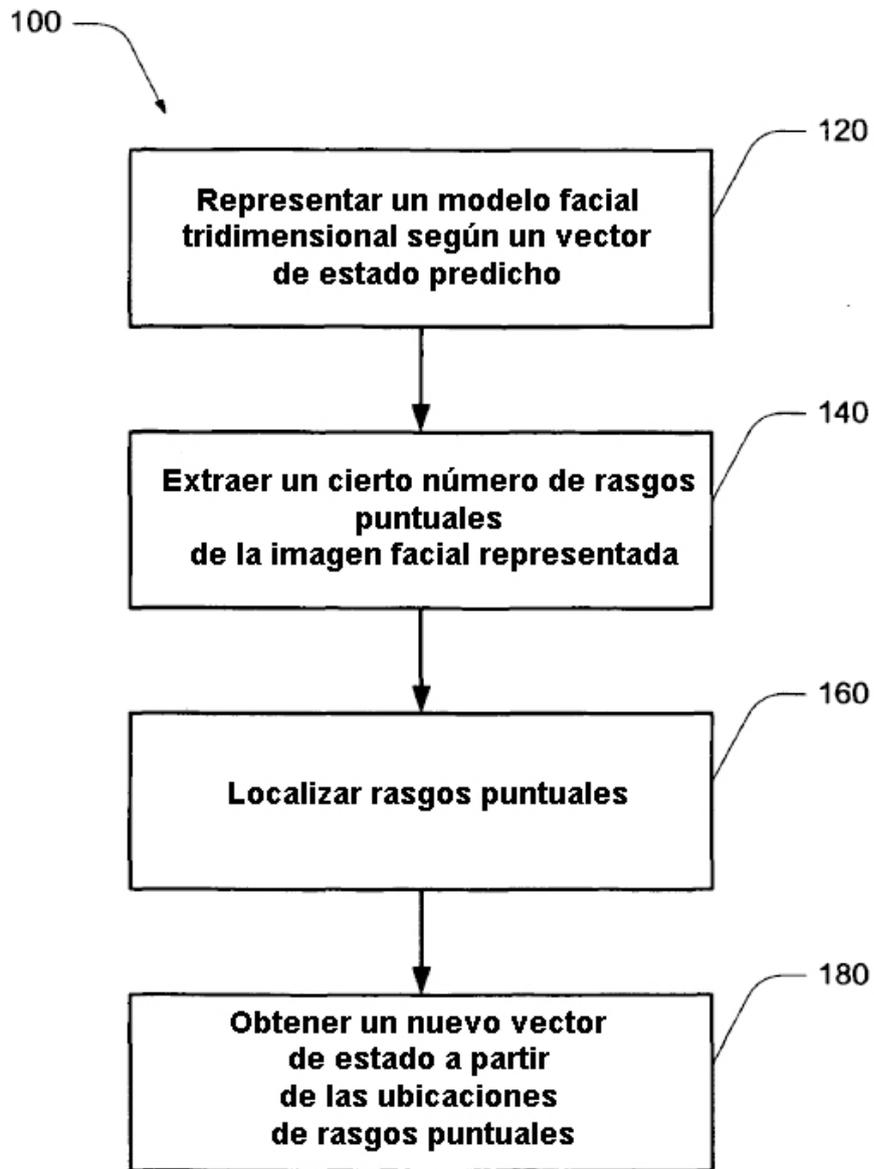


FIG. 1

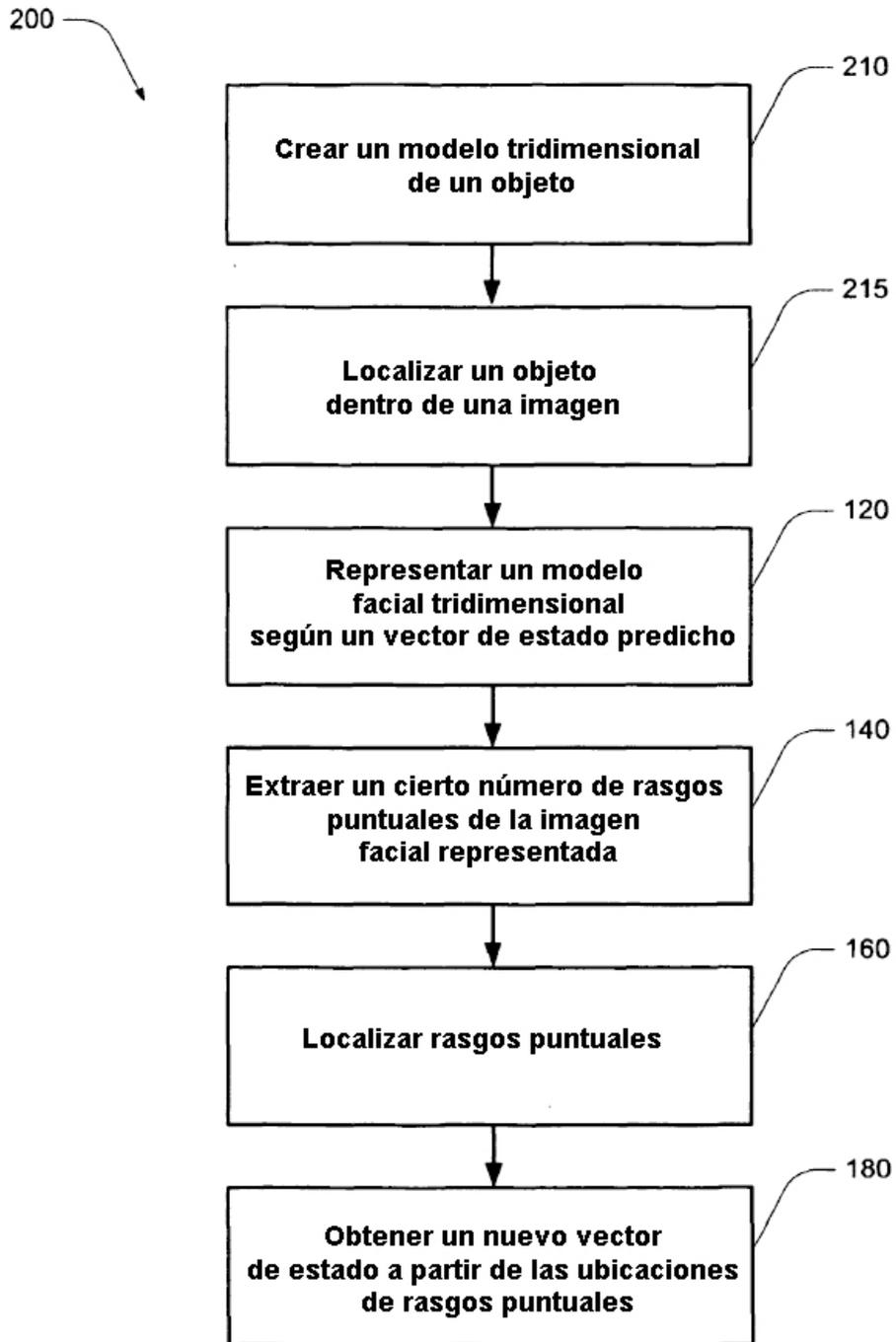


FIG. 2

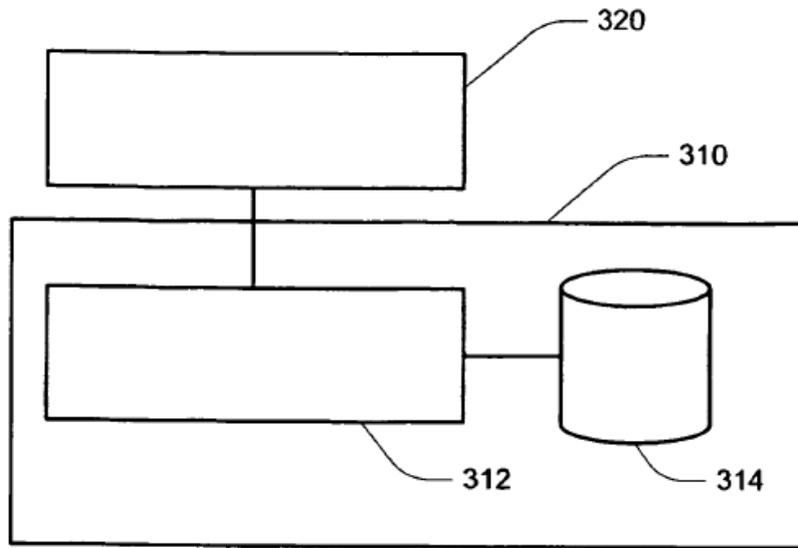


FIG. 3

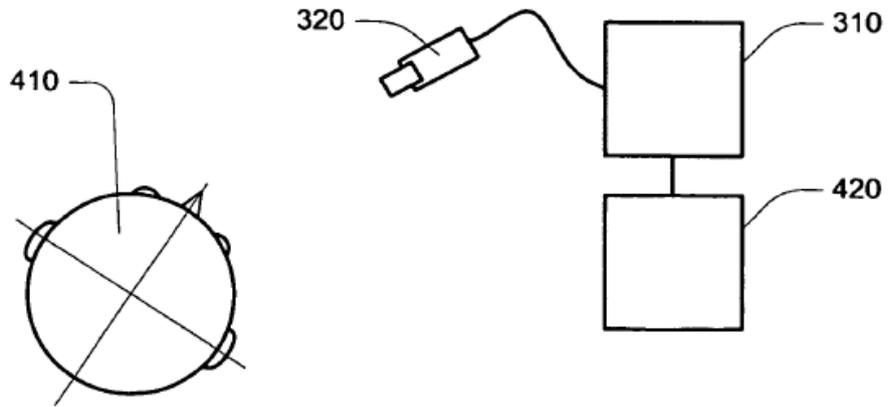


FIG. 4