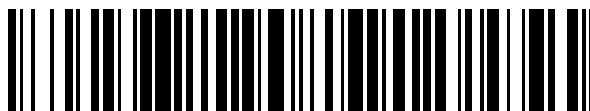


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 385 041**

51 Int. Cl.:  
**G06K 9/00** (2006.01)  
**G06K 9/64** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **05771849 .6**
- 96 Fecha de presentación: **11.08.2005**
- 97 Número de publicación de la solicitud: **1810216**
- 97 Fecha de publicación de la solicitud: **25.07.2007**

54 Título: **Reconocimiento de objetos 3D**

30 Prioridad:  
**19.08.2004 SE 0402048**  
**23.08.2004 US 603266 P**

45 Fecha de publicación de la mención BOPI:  
**17.07.2012**

45 Fecha de la publicación del folleto de la patente:  
**17.07.2012**

73 Titular/es:  
**Apple Inc.**  
**1 Infinite Loop**  
**Cupertino, CA 95014, US**

72 Inventor/es:  
**SOLEM, Jan Erik y**  
**KAHL, Fredrik**

74 Agente/Representante:  
**Fàbrega Sabaté, Xavier**

**ES 2 385 041 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

Reconocimiento de objetos 3D

5 **Campo de la invención**

La presente invención se refiere a reconocimiento automático de objetos y en particular a reconocimiento automático de objetos 3D utilizando información estadística de forma.

10 **Antecedentes de la invención**

Existen procedimientos extremadamente fiables para identificación personal utilizando datos biométricos tales como, por ejemplo, huellas digitales, patrones retinales o características únicas similares del sujeto que dependen de la colaboración del sujeto. El reconocimiento facial puede ser una manera efectiva de identificar a una persona sin la colaboración o conocimiento de la persona. Existen dos problemas generales principales para un sistema de reconocimiento facial: identificar a una persona, es decir, determinar la identidad a partir de imágenes, y verificar la identidad de una persona, es decir, certificar que una persona es quién dice ser. Aplicaciones específicas son, por ejemplo, inmigración, documentos de identidad, pasaportes, acceso a un computador, seguridad en intranet, vigilancia mediante video y sistemas de acceso. La presente invención pretende incrementar las prestaciones y eficiencia de tales sistemas utilizando información geométrica disponible a través del uso de modelos estadísticos de forma.

En el área de modelos estadísticos de forma, la invención se refiere a los Modelos Activos de Forma (ASM) introducidos por Cootes y Taylor, ([1]: Cootes T.F. y Taylor C.J., "Active Shape model Search using Grey-level Models: A Quantitative Evaluation", Conferencia Británica de Visión Artificial, páginas 639-648, 1993). Una distinción es que los ASM han sido utilizados para inferir formas 2D a partir de observaciones 2D o formas 3D a partir de informaciones 3D mientras que la invención utiliza observaciones 2D, es decir, imágenes, para inferir formas 3D. Además, las observaciones son desde múltiples puntos de vista (uno o más dispositivos de imagen) algo que el ASM estándar no trata. Cootes y Taylor tienen un número de patentes en este área, de las cuales las más relevantes son (WO02103618A1 – Modelo Estadístico) en donde se trata la parametrización de formas 2D ó 3D, (WO0135326A1 – Identificación de Clase de Objeto, Verificación o Síntesis de Imagen de Objeto) en donde una clase de objeto se identifica en imágenes y (WO02097720A1 – Identificación de Objeto) en la que los objetos se identifican utilizando versiones modificadas de ASM y técnicas relacionadas. También relacionado está Cootes et al. ([2]: Cootes T.F., Wheeler G.V., Walker K.N y Taylor C.J., "View-based Active Appearance Models", Image and Vision Computing, 20(9-10), páginas 657-664; 2002) en donde se utilizan los modelos multi-perspectiva pero el modelo no contiene datos 3D explícitos o consistentes. Existen también procedimientos para deformar un modelo 3D del objeto para encajar las proyecciones 2D del objeto en las imágenes tal como en Blanz y Vetter ([3] Blanz V. y Vetter T., "Face Recognition based on Fitting a 3D Morphable Model", IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(9), páginas 1063-1073, 2003.). Estos procedimientos son computacionalmente muy costosos y a menudo requieren intervención manual. Las patentes relacionadas son US6556196/EP1039417 (Procedimientos y aparatos para el procesamiento de imágenes) que describe un procedimiento de transformar un modelo 3D de forma que sea una representación 3D del objeto en la imagen mediante la minimización del error de proyección en la imagen.

Un problema común para el reconocimiento basado en imagen es detectar la forma 2D del objeto en la imagen, es decir, encontrar la región relevante de la imagen. Los procedimientos recientes para detectar objetos en imágenes incluyen habitualmente escanear la imagen a distintas escalas en busca de patrones de imagen específicos del objeto y luego utilizar un clasificador para decidir si la región es relevante o no. Los últimos desarrollos sugieren el uso de Máquinas de Vectores de Soporte (SVM) para esta tarea. Un elemento clave es la extracción de características de imagen, es decir, partes de la imagen tales como esquinas, bordes u otros puntos interesantes. Esto se hace habitualmente utilizando esquemas basados en correlación que usan plantillas o procedimientos basados en los bordes que utilizan gradientes. Para una visión general de procedimientos para reconocimiento facial y extracción de características de imagen, véase Zhao y Chellapa ([4]: Zhao W., Chellapa R., Rosenfeld A y Phillips P.J., "Face Recognition: A Literature Survey", informe técnico CAR-TR-948, 2000.) y las referencias en este documento. En [4], se presenta también un resumen de los procedimientos actuales basados en imagen para reconocimiento facial.

Al usar procedimientos basados en imagen para identificación y verificación existen dos problemas principales, variación de la iluminación y variación de la pose. La variación de la iluminación afectará a todos los procedimientos basados en correlación en los que se comparan partes de la imagen ya que los valores de los píxeles pueden variar con iluminación cambiante. Las reflexiones especulares pueden llevar a grandes cambios en la intensidad de pixel. La variación de pose ocurre debido a que la proyección en la imagen puede cambiar dramáticamente al rotar el objeto. Estos dos problemas han sido documentados en muchos sistemas de reconocimiento facial y son inevitables cuando las imágenes se adquieren en entornos no controlados. La mayoría de los procedimientos conocidos fallan en resolver estos problemas de forma robusta.

El problema de iluminación es gestionado por la invención ya que no se lleva a cabo correlación de imagen o comparación de partes de imagen. En vez, se computan características de imagen como esquinas que son robustos a cambios de intensidad, lo que hace la reconstrucción de forma, en gran parte, insensible a la iluminación y a las

reflexiones especulares. La invención gestiona el problema de la pose usando cualquier número de imágenes con pose diferentes para entrenar el modelo estadístico. Cualquier subconjunto de las imágenes, tan pequeño como una única imagen, puede utilizarse para inferir la forma 3D del objeto.

5 El artículo titulado "Inferring 3D structure with a Statistical Image-based Shape Model" de Grauman K et al. (Proceedings of the Eight IEEE International Conference on Computer Vision, Niza, Francia, Octubre 13-16, 2003, volumen 2, páginas 641-648, XP010662422, ISBN:0-7695-1950-4) describe un enfoque basado en imagen para inferir parámetros de la estructura 3D utilizando un modelo probabilístico "forma + estructura". La forma 3D de una clase de objetos se representa mediante conjuntos de contornos a partir de vistas de silueta observados simultáneamente a partir de  
10 múltiples cámaras calibradas. Las reconstrucciones bayesianas de nuevas formas son entonces estimadas utilizando una densidad anterior construida con un modelo de mezclas y un análisis probabilístico de componentes principales. El modelo de forma se expande para incorporar características estructurales de interés; ejemplos novedosos con falta de parámetros de estructura se reconstruyen entonces para obtener estimaciones de estos parámetros. Hacer coincidir el modelo y la inferencia de parámetros, se hace completamente en el dominio de la imagen y no requiere una construcción  
15 3D explícita. El modelo de forma permite una estimación precisa de la estructura a pesar de errores de segmentación o falta de vistas en las siluetas de entrada, y funciona incluso con solo una única vista de entrada. Utilizando un conjunto de datos de imágenes de peatones generadas con un modelo sintético, las localizaciones 3D de 19 articulaciones en el cuerpo basadas en los contornos de silueta de imágenes reales puede inferirse de forma precisa.

## 20 RESUMEN DE LA INVENCIÓN

La invención consiste en un modelo estadístico de las variaciones de forma en una clase de objetos que relaciona la proyección en dos dimensiones (2D) en imágenes a la forma en tres dimensiones (3D) del objeto y el uso de la información de forma 3D para identificación o verificación del objeto. Además, la presente invención se refiere a un  
25 dispositivo o sistema de procesamiento de imagen para implementar tal procedimiento. El procedimiento es completamente automático y puede usarse para, por ejemplo, identificación biométrica a partir de imágenes de caras o identificación de objetos en por ejemplo imágenes de Rayos X en seguridad aeroportuaria. La forma 3D recuperada es la forma más probable consistente con las proyecciones 2D, es decir, las imágenes. El modelo estadístico necesita un banco de datos, denominado datos de entrenamiento, en el que las posiciones 3D de las características de imagen son conocidas, para aprender los parámetros del modelo. Tal muestreo de datos puede hacerse utilizando, por ejemplo, escáneres binoculares, estéreo de múltiple perspectiva o de rango. Una vez que se han aprendido los parámetros del modelo, la forma 3D puede computarse utilizando una o varias imágenes. Se usa entonces la forma 3D, por medio de la presente invención junto con los datos de imagen 2D, para identificar o  
30 verificar el objeto como una instancia particular de la clase de objetos, por ejemplo, la cara que pertenece a un individuo particular. Una identificación positiva (o negativa) inicia las acciones adecuadas por medio de la presente innovación.

En una realización de la invención un procedimiento de reconocimiento de objetos de un objeto tridimensional, 3D, objeto (607), comprende las etapas de: obtener al menos una representación bidimensional (606), 2D, de dicho  
40 objeto; detectar características de imagen en dicha representación 2D obtenida (606); inferir una forma 3D dispersa correspondiente a las características de imagen utilizando una representación vectorial latente de dicha forma 3D dispersa y utilizar un modelo de forma estadístico multi-perspectiva; caracterizado por comprender además las etapas de extender la forma 3D dispersa a una forma 3D densa encajando una superficie a dicha forma 3D dispersa utilizando un modelo estadístico aprendido de forma consistente en una superficie media completa del objeto para regularizar dicha forma de superficie de una manera específica para una clase de objetos; comparar la forma 3D  
45 densa con formas 3D de referencia (605) obtenidas previamente; y por lo tanto identificar un objeto individual de la clase de objetos.

Todavía en el procedimiento, la forma 3D dispersa puede inferirse de las características 2D.

50 En otro aspecto del procedimiento según la presente invención, la clase de objetos puede contener objetos no rígidos y el modelo de forma estadística puede ser aprendido utilizando datos 2D y 3D específicos para posibles deformaciones de los objetos en la clase de objetos no rígidos.

55 El procedimiento puede además comprender la etapa de identificar un objeto individual de una clase de objetos o ayudar a la identificación de un objeto individual utilizando la forma 3D densa.

El procedimiento puede además comprender la etapa de verificar la identidad de un objeto individual de una clase de objetos o ayudar a la verificación de un objeto individual utilizando la forma 3D densa.

60 En el procedimiento el objeto puede ser uno de; una cara humana, un cuerpo humano, órgano(s) interno(s) de un ser humano, vasos sanguíneos, animales, órganos internos de un animal, un tumor, producto(s) manufacturado(s) de un proceso industrial, un vehículo, una aeronave, un barco, objeto(s) militar(es).

65 En el procedimiento, las formas de referencia 3D pueden almacenarse en al menos uno de una memoria no volátil, servidor de base de datos y tarjeta de identificación personal.

En otra realización de la presente invención un dispositivo (400) para reconocimiento de objetos de un objeto (607) tridimensional, 3D, que comprende: medios para obtener (404), al menos una representación bidimensional, 2D, de dicho objeto (607) ; medios para detectar (401) características de imagen en dicha representación 2D obtenida;

5 medios para inferir una forma 3D dispersa correspondiente a las características de imagen utilizando una representación vectorial latente de dicha forma 3D dispersa utilizando un modelo estadístico aprendido de forma consistente en una superficie media completa del objeto para regularizar dicha forma de superficie de una forma específica para una clase de objetos; comparar la forma 3D densa con las formas de referencia previamente obtenidas, identificando por lo tanto un objeto individual de la clase de objetos.

10 En el dispositivo la forma 3D dispersa puede inferirse de las características de imagen 2D.

En el dispositivo la clase de objetos puede contener objetos no rígidos y el modelo de forma estadística puede ser aprendido utilizando datos 2D y 3D específicos para posibles deformaciones de los objetos en la clase de objetos no rígidos.

15 El dispositivo puede además comprender medios para identificar un objeto individual de una clase de objetos o ayudar en la identificación de un objeto individual utilizando la forma 3D densa.

20 El dispositivo puede además comprender medios para verificar la identidad de un objeto individual de una clase de objetos o ayudar en la verificación de identidad de un objeto individual utilizando la forma 3D densa.

En el dispositivo el objeto puede ser uno o varios de: una cara humana, un cuerpo humano, órgano(s) interno(s) de un ser humano, vasos sanguíneos, animales, órganos internos de un animal, un tumor, producto(s) manufacturado(s) de un proceso industrial, un vehículo, una aeronave, un barco, objeto(s) militar(es).

25 En el dispositivo las formas 3D densas de los vasos sanguíneos o órganos recuperados a partir de proyecciones 2D, por ejemplo, utilizando imágenes de Rayos X puede usarse para navegar catéteres dirigibles o ayudar a los médicos a mostrar la forma 3D densa.

30 Las formas 3D densas de características faciales pueden usarse en el dispositivo para identificar o verificar la identidad de un individuo en un sistema de control de acceso o de seguridad, resultando en la aceptación o rechazo del individuo.

35 El dispositivo puede además comprender una interfaz para comunicarse con una tarjeta de identificación personal en la que están almacenadas las formas 3D de referencia.

- En otra realización de la presente invención un programa de ordenador almacenado en un medio de almacenamiento legible por ordenador (402) que comprende una pluralidad de instrucciones que cuando son ejecutadas por una unidad de computación (901) hace que la unidad lleve a cabo el procedimiento de cualquiera de las reivindicaciones 1 a 6.
- En otra realización de la presente invención un sistema de reconocimiento de objetos de un objeto tridimensional, 3D, que comprende: un dispositivo (400) según la reivindicación 7; y medios para responder (604) a un resultado de dichos medios de comparación.

40 El sistema puede además incluir medios para identificar y/o verificar un objeto individual de una clase de objetos o ayudar en la identificación y/o verificación del objeto individual utilizando la forma 3D recuperada.

45 En el sistema las formas 3D de referencia pueden almacenarse en al menos uno de una memoria no volátil, servidor de base de datos y una tarjeta de identificación personal.

### Breve descripción de los dibujos

50 De aquí en adelante la invención se describe de forma no limitante y con más detalle con referencia a las realizaciones de ejemplo ilustradas en los dibujos contenidos, en los que:

La Figura 1 ilustra un procedimiento de dos etapas para recuperar datos 3D de una imagen de entrada.

La Figura 2 ilustra un proceso de encaje de superficie en una superficie 3D recuperada.

La Figura 3 es un diagrama de bloques de un dispositivo según la presente invención.

La Figura 4 ilustra un diagrama de bloques esquemático de las etapas de un procedimiento según la presente invención.

La Figura 5 es una ilustración esquemática de un sistema según la presente invención.

**Descripción detallada de la invención**

5 La invención consiste en un sistema de procesamiento de imagen para recuperación automática de formas 3D a partir de imágenes de objetos que pertenecen a una cierta clase. Esta reconstrucción 3D se hace estableciendo un modelo de forma estadístico, denotado modelo de características, que relaciona las características 2D, por ejemplo, puntos de interés o curvas, con sus correspondientes posiciones 3D. Tal modelo es aprendido, es decir, se estiman los parámetros del modelo, a partir de datos de entrenamiento en los que la correspondencia 2D-3D es conocida. Esta fase de aprendizaje puede hacerse utilizando cualquier sistema apropiado para obtener tal correspondencia 2D-3D, 10 incluyendo, pero no limitado a sistemas de adquisición de imágenes binoculares o de múltiple perspectiva, escáneres de rango o disposiciones similares. En este proceso se mide el objeto de interés y se obtiene un modelo de referencia del objeto que puede utilizarse en un análisis de imagen subsiguiente tal y como se describe a continuación.

15 Dada una imagen de entrada, el proceso de recuperar la forma 3D es un procedimiento de dos etapas. Primero las características como puntos, curvas y contornos son encontrados en las imágenes, por ejemplo, utilizando técnicas como, por ejemplo, ASM [1] o procedimientos o clasificadores basados en gradiente tales como SVM. Entonces se infiere la forma 3D utilizando el modelo aprendido de características. Esto se ilustra en la Figura 1. La Figura 1a ilustra una imagen de una cara a analizar, la Figura 1b ilustra la detección de características de objeto para su uso en el análisis y proceso de información de forma y la Figura 1c ilustra la forma 3D inferida para su uso en el proceso de reconocimiento. 20 de reconocimiento.

También existe la opción de extender la representación de forma 3D desde curvas y puntos a un modelo de superficie completo encajando una superficie en los datos 3D. Esto se ilustra en la Figura 2, en donde la Figura 2a, 25 ilustra la forma 3D inferida, 2b ilustra una superficie encajada a los datos 3D, y la Figura 2c ilustra un modelo de superficie 3D renderizada del superficie encajada.

*El modelo de características*

30 Supóngase que tenemos un número de elementos en un vector  $d$ -dimensional  $\mathbf{t}$ , por ejemplo, una colección de puntos 3D en algún sistema normalizado de coordenadas. El punto de inicio para la derivación del modelo es que los elementos en  $\mathbf{t}$  pueden estar relacionados con algún vector latente  $\mathbf{u}$  de dimensión  $q$  en donde la relación es lineal:

$$\mathbf{t} = W\mathbf{u} + \mu \tag{1}$$

35 en donde  $W$  es una matriz de tamaño  $d \times q$  y  $\mu$  es un vector- $d$  que permite una media no nula. Una vez que los parámetros del modelo  $W$  y  $\mu$  han sido aprendidos a partir de los ejemplos, se mantienen fijos. Sin embargo, nuestras medidas tienen lugar en las imágenes, lo cual es normalmente una función no lineal de las características 3D según el modelo de proyección para el dispositivo de imagen relevante. 40

Denótese la función de proyección con  $f: R^d \rightarrow R^e$ , que proyecta todas las características de imagen 3D a características de imagen 2D, para una o más imágenes. Además, necesitamos cambiar el sistema de coordenadas de las características 3D para que se adapte a la actual función de proyección. Denótese este mapeo mediante  $T: R^d \rightarrow R^d$ . Típicamente,  $T$  es una transformación de similitud del sistema de coordenadas mundial. Por lo tanto,  $f(T(\mathbf{t}))$  proyectará todos los datos 3D normalizados en todas las imágenes. Finalmente, necesita especificarse un modelo de ruido. Asumimos que las medidas de imagen son independientes y normalmente distribuidas, de igual forma, las variables latentes se suponen Gaussianas con varianza unitaria  $\mathbf{u} \sim N(0,1)$ . Por lo tanto, en resumen: 45

$$\mathbf{t}_{2D} = f(T(\mathbf{t})) + \varepsilon = f(T(W\mathbf{u} + \mu)) + \varepsilon \tag{2}$$

50 en donde  $\varepsilon \sim N(\sigma^2 \mathbf{1})$  para algún escalar  $\sigma$ . El modelo está relacionado con PPCA, véase Tipping y Bishop ([5]: Tipping M.E., Bishop C.M., "Probabilistic Principal Component Analysis", Phil. Trans. Royal Soc. London B, 61(3), p. 611-622, 1999.), pero también existen diferencias debido a la no linealidad de  $f(\cdot)$ . Antes de poder usar el modelo, sus parámetros deben ser estimados a partir de datos de entrenamiento. Dado que es un modelo probabilístico, esto 55

se hace de la mejor forma posible con máxima probabilidad (ML). Supóngase que se nos dan  $n$  ejemplos  $\{\mathbf{t}_{2D,i}\}_{i=1}^n$ , la estimación ML para  $W$  y  $\mu$  se obtiene minimizando:

$$\sum_{i=1}^n \left( \frac{1}{\sigma^2} \|\mathbf{t}_{2D} - f(T_i(\mathbf{u}_i))\|^2 + \|\mathbf{u}_i\|^2 \right) \quad (3)$$

5

sobre todos los desconocidos. La desviación estándar  $\sigma$  se estima a priori a partir de los datos. Una vez que los parámetros del modelo  $W$  y  $\mu$  han sido aprendidos a partir de ejemplos, se mantienen fijos. En la práctica, para

10 minimizar (3) optimizamos alternativamente sobre  $(W, \mu)$  y  $\{\mathbf{u}_i\}_{i=1}^n$  utilizando descenso de gradiente. Los estimadores iniciales pueden obtenerse intersectando la estructura 3D de cada grupo de imágenes y luego aplicando algoritmos PPCA para la parte lineal. La normalización  $T_i(\cdot)$  se escoge de tal forma que cada muestra 3D normalizada tiene media cero y varianza unitaria.

15 Existen tres tipos diferentes de características geométricas empotradas en el modelo.

**Puntos:** Un punto 3D que es visible en  $m > 1$  imágenes se representará en el vector  $\mathbf{t}$  con sus coordenadas 3D (X,Y,Z). Para puntos visibles en solo una imagen,  $m = 1$ , no existe información de profundidad, y tales puntos se representan similarmente a puntos aparentes de contorno.

20 **Curvas:** Una curva se representará en el modelo mediante un número de puntos a lo largo de la curva. En el entrenamiento del modelo, es importante parametrizar cada curva 3D de forma que cada punto en la curva corresponde aproximadamente al mismo punto en la curva correspondiente en los otros ejemplos.

25 **Contornos aparentes:** Como para las curvas, muestreamos los contornos aparentes (en las imágenes). Sin embargo, no existe información 3D disponible para los contornos aparentes ya que dependen de la perspectiva. Una forma sencilla es tratar los puntos de los contornos aparentes como puntos 3D con una estimación de profundidad constante, aproximada (pero gruesa).

30 *Encontrar características de la imagen*

En el caso en línea de una nueva muestra de entrada, queremos encontrar automáticamente las variables latentes  $\mathbf{u}$  y, a su vez, computar estimaciones de las características 3D  $\mathbf{t}$ . El componente que falta en el modelo es la relación entre las características de imagen 2D y los valores en escala de grises (o color) subyacentes en esos píxeles. Existen varias maneras de resolver esto, por ejemplo, utilizando un ASM (denominado el *modelo de escala de grises*) o enfoques basados en detectores.

*El modelo de escala de grises*

40 De nuevo, adoptamos un modelo lineal (PPCA). Utilizando la misma notación que en (1), pero ahora con el subíndice  $gl$  para escala de grises, el modelo puede escribirse

$$\mathbf{t}_{gl} = W_{gl} \mathbf{u}_{gl} + \mu_{gl} + \epsilon_{gl} \quad (4)$$

45 en donde  $\mathbf{t}_{gl}$  es un vector que contiene los valores en escala de grises de todas las características de imagen 2D y  $\epsilon_{gl}$  es ruido gaussiano en las mediciones. En la fase de entrenamiento, cada muestra de datos de escala de grises se normaliza sustrayendo la media y escalando la varianza unitaria. La estimación de  $G_{gl}$  y  $\mu_{gl}$  se computa con el algoritmo EM [5].

50 *Procedimientos basados en detector*

Los puntos y curvas de interés de imagen pueden encontrarse analizando el gradiente de imagen utilizando, por ejemplo, el detector de esquinas Harris. Además, filtros diseñados específicamente pueden utilizarse como

detectores de características de imagen. Diseñando los filtros de forma que la respuesta a ciertas estructuras de imagen locales es alta, se pueden encontrar las características de imagen utilizando una convolución 2D.

*Procedimientos de clasificación*

5 Utilizando clasificadores como SVM, se pueden clasificar las regiones de imagen como correspondientes a una cierta característica o no. Combinando una serie de tales clasificadores, uno para cada tipo de característica (puntos, curvas, contornos, etc.) y escaneando la imagen a todas las escalas apropiadas se pueden extraer las características. Ejemplos pueden ser, por ejemplo, un detector de ojos para imágenes faciales.

10 *Modelos deformables*

Utilizar un modelo deformable como los Modelos de Contorno Activo, también denominados serpientes, de un cierta característica es muy común en el campo de la segmentación. Normalmente las características son curvas. El proceso es iterativo y trata de optimizar una función de energía. Una curva inicial se deforma gradualmente para encajar de la mejor forma en una función de energía que puede contener términos que regulan la suavidad del encaje así como otras propiedades de la curva.

20 **Encaje de superficie a los datos 3D**

Una vez que se han recuperado los datos 3D, se puede encajar un modelo de superficie a la estructura 3D. Esto puede ser deseable en el caso de que el procedimiento anterior de dos etapas produzca solamente un conjunto disperso de características en el espacio 3D tal como, por ejemplo, puntos y curvas espaciales. Incluso si estas indicaciones son características para una muestra en particular (o individual), es a menudo no suficiente para inferir un modelo de superficie completo, y en particular, esto es difícil en las regiones en las que las características están dispersas. Por lo tanto, se introduce un modelo de superficie 3D consistente en la superficie media completa. Esto servirá como un regularizador de dominio específico, es decir, específico para una cierta clase de objetos. Este enfoque requiere que haya una información densa de forma 3D disponible para algunos ejemplos de entrenamiento en los datos de entrenamiento de la clase de objetos obtenida a partir de, por ejemplo, escáneres láser, o en el caso de imágenes médicas de, por ejemplo, MRI o tomografía computerizada. A partir de estas formas 3D densas, se puede construir un modelo separado del modelo de características anterior. Esto significa que, dada una forma recuperada 3D, en forma de puntos y curvas, a partir del modelo de características, puede calcularse la mejor forma densa según la forma 3D recuperada. Esta información de forma densa puede usarse para mejorar el encaje de superficie.

35 Para ilustrarlo con un ejemplo, considérese el caso de que una clase de objetos sean caras. El modelo es entonces aprendido utilizando, por ejemplo, curvas y contornos en imágenes junto con la forma real 3D correspondientes a estas características obtenidas de, por ejemplo, técnicas estéreo de múltiples perspectivas. Se crea y aprende entonces un segundo modelo utilizando, por ejemplo, escáneres láser de caras, resultando en un conjunto de superficies faciales. Este segundo modelo puede utilizarse para encontrar la superficie facial media (según el segundo modelo) más probable (o al menos altamente probable) correspondiente a las características o la forma 3D recuperada. Puede encajarse entonces una superficie 3D con la condición adicional de que cuando no haya una forma 3D recuperada, la superficie debería parecerse a la superficie facial media más probable.

45 Como segundo ejemplo, considérese el caso de que la clase de objetos sea un vaso sanguíneo en particular, por ejemplo, la aorta. Entonces, se aprende al modelo utilizando, por ejemplo, curvas y contornos en imágenes junto con la verdadera forma 3D obtenida como, por ejemplo, una imagen MRI 3D. A partir de las formas 3D se aprende a un segundo modelo que comprende la superficie de la aorta. Entonces, la superficie de aorta más probable (o altamente probable) puede recuperarse a partir de las características de imagen o a partir de la forma 3D recuperada por el modelo de forma primario.

El procedimiento proporciona la forma más probable o al menos altamente probable, en muchas aplicaciones esto es suficiente y el proceso de identificación y/o verificación no es necesario para la aplicación final.

55 Hemos ahora descrito el procedimiento subyacente utilizado con fines de verificación y/o identificación. En referencia ahora la Figura 3 se proporciona una descripción de un dispositivo 400 que implementa el procedimiento preferido según la presente invención. Tal dispositivo 400 puede ser cualquier tipo apropiado de dispositivo computacional tal como, pero no limitado a, un ordenador personal (PC), estación de trabajo, computador empotrado, o un dispositivo autónomo con una unidad de computación 401, tal como un microprocesador, DSP (procesador digital de señal), FPGA (matriz programable de puertas), o ASIC (circuito integrado de aplicación específica). El dispositivo 400 tiene algunos medios de entrada 404 para obtener imágenes para análisis y/o verificación o identificación final. Los medios de entrada 404 pueden ser de cualquier interfaz de comunicación adecuada dependiendo del tipo de imagen e incluyen, pero no se limitan a, USB (bus serie universal), digitalizador, Ethernet o Firewire. Los datos de imagen se transfieren a una unidad de computación 401 en la que reside software para la ejecución del procedimiento descrito anteriormente según la presente invención. El dispositivo 400 puede además comprender alguna memoria volátil o no volátil 402 que contiene información relativa a un material de referencia con objetivos de comparación y/o

análisis, por ejemplo, relaciones conocidas 2D-3D de objetos de interés. El dispositivo 400 puede además comprender medios de comunicación para comunicarse con otros dispositivos de comunicación sobre, por ejemplo, un protocolo de red (como Ethernet o protocolos similares) y medios de salida 405 para entregar resultados a por ejemplo una pantalla para su visionado conveniente a un dispositivo de control (no mostrado) para controlar un proceso externo del que los objetos de interés son parte. Tales procesos pueden incluir, pero no se limitan a, procesos de producción industrial en los que los objetos pueden ser seleccionados o deseleccionados dependiendo del resultado del procedimiento de identificación y/o verificación según la presente invención, procesos de seguridad otra vez con objetivos de selección o desección en, por ejemplo, sistemas de seguridad aeroportuaria para examen de los contenidos de las maletas, bolsas u otros aparatos de equipaje, o aplicaciones médicas en las que la forma recuperada 3D puede usarse para, por ejemplo, navegación de instrumentos o dispositivos médicos.

El procedimiento de reconocimiento de objetos según la presente invención puede ilustrarse usando la Figura 4. El procedimiento puede comprender las siguientes etapas:

1. Obtener al menos una imagen de un objeto para ser identificado y/o verificado (501).
2. Detectar características de imagen, tales como curvas, puntos y contornos aparentes (502).
3. Analizar la imagen obtenida e inferir la forma 3D correspondiente a las características de imagen, utilizando un modelo de forma estadístico (503);
4. Comparar el análisis con imágenes de referencia previamente obtenidos y comparar la forma 3D en una forma dispersa o densa con la forma 3D previamente obtenida (504); y
5. Responder a una salida del proceso de comparación (505).

En otra realización de la presente invención se utiliza un sistema para obtener imágenes, analizar y responder a resultados del proceso de identificación y/o verificación, como puede verse en la Figura 5. Tal sistema puede incluir por lo menos un dispositivo de adquisición de imagen 601, un dispositivo de computación 400, 603 como se describió anteriormente, y algún tipo de equipo receptivo tal como, por ejemplo, el equipo de proceso industrial o el equipo de proceso de seguridad descrito anteriormente. Al menos un dispositivo de adquisición de imagen 601 se utiliza para adquirir una o más imágenes que se transfieren 602 a un dispositivo de computación 603 para análisis y verificación y/o identificación. El resultado de este proceso se transmite a un sistema de control o a un sistema de visualización 604. En el caso de un sistema de detección de caras se obtiene al menos una imagen de una persona, por ejemplo la imagen de la cara de la persona, y la imagen o imágenes se transmiten al dispositivo de computación 603, utilizando cualquier medio de comunicación adecuado 602 (cableado o inalámbrico), para el análisis y comparación de la imagen adquirida o imágenes con datos obtenidos a partir de medidas de referencia por ejemplo con relaciones 2D-3D conocidas; sin embargo, se puede hacer una comparación entre una forma 3D inferida con datos de referencia 3D almacenados o entre un modelo de superficie 2D con un modelo de superficie de referencia 2D almacenado. El resultado puede hacerse disponible mediante por ejemplo una unidad de visualización 604 y puede mostrarse con propósitos ilustrativos tanto con una imagen de referencia 605 y la imagen obtenida 606 o imágenes renderizadas a partir de la forma 3D recuperada como se muestra en la Figura 5. Debe ser apreciado por el experto en la materia que el sistema de adquisición de imagen y/o sistema de visualización/control puede estar incorporado al dispositivo de computación formando una unidad integral y que el resultado puede mostrarse de una forma adecuada y no se limita al ejemplo descrito con anterioridad. En vez de transferir el resultado a una unidad de visualización 604 puede usarse en cualquier proceso de control adecuado para controlar, por ejemplo, una alarma, un sistema de entrada, puerta de control, puerta de peaje, etc.

Algunos de los beneficios que aporta la presente invención al campo técnico, pueden ilustrarse con la siguiente lista:

- Cualquier número de imágenes, incluso tan poco como una única imagen, puede utilizarse para recuperar automáticamente la forma 3D de un objeto en la clase de objetos.
- Un modelo estadístico multi-perspectiva que representa datos 2D y 3D consistentemente.
- El proceso es automático y computacionalmente eficiente.
- El proceso es robusto a iluminación y reflexiones especulares, lo que es un problema para procedimientos de reconstrucción 3D basados en correlación de imágenes o fotoconsistencia.
- Se pueden encajar superficies a la estructura 3D utilizando regularizadores de dominio específico aprendidos a partir de modelos estadísticos de forma.

La flexibilidad de la presente invención puede ilustrarse con la siguiente lista:

- El modelo estadístico de forma puede usarse para cualquier clase de objetos y la proyección de estos objetos en imágenes.
- El enfoque puede utilizarse para cualquier tipo de dispositivos de imágenes (cámara, Rayos X, multispectral, termal, etc.).
- La invención puede usarse con cualquier número de dispositivos de imagen (uno o más).
- La invención incluye la posibilidad de combinar muchas técnicas diferentes para establecer



correspondencia 2D-3D (sistemas de adquisición de imagen, escáneres de rango, sistema de imagen estéreo, sistema de imagen estéreo de múltiple perspectiva. Rayos X, etc.)

- La invención incluye la posibilidad de utilizar diferentes técnicas, tales como ASM, procedimientos basados en gradiente, o modelos deformables para encontrar las características de imagen.
- Si la clase de objetos incluye objetos no rígidos, la invención incluye la posibilidad de establecer modelos de 2D a 3D para diferentes deformaciones del objeto (por ejemplo, diferentes expresiones faciales).
- La invención incluye la posibilidad de utilizar un modelo estadístico de forma para encajar la superficie a la forma 3D recuperada.

10 Las representaciones de referencia de objetos pueden almacenarse en varias ubicaciones diferentes y con diferentes tipos de sistemas, tales como, pero no limitados a, localmente en alguna memoria no volátil en un dispositivo que utiliza el reconocimiento de objetos según la presente invención, en un servidor centralizado, por ejemplo, un servidor de base de datos, o una tarjeta de identificación personal que contiene una representación de referencia de un objeto tal como una persona y esta tarjeta de identificación puede usarse por ejemplo en un sistema de acceso. Comunicación entre un sistema de reconocimiento de objetos y un sistema de almacenamiento de representación de referencia puede utilizarse con diferentes de niveles de seguridad y/o esquemas, tales como RADIO, DIAMÉTRO, SSL, SSH o cualquier otro sistema de comunicación encriptado como lo entendería el experto en la materia.

20 Las posibles áreas de aplicación para la invención descrita anteriormente van desde identificación de objetos y verificación en procesos industriales, determinar y/o identificar objetos por motivos de seguridad, reconocimiento de imágenes por motivos militares, por ejemplo, determinación automática de vehículos militares, barcos militares, aeronaves y similares, sistemas de reconocimiento facial para muchas aplicaciones diferentes, por ejemplo, biometría, seguridad de información, aplicación de leyes, tarjetas inteligentes, control de acceso y similares.

25 Las realizaciones mencionadas y descritas anteriormente se dan solo como ejemplos y no deberían limitar la presente invención. Otras soluciones, usos, objetivos y funciones dentro del alcance de la invención según se reivindica en las reivindicaciones descritas a continuación deberían ser aparentes para cualquier experto en la materia.

30

**REIVINDICACIONES**

1. Un procedimiento para reconocimiento de objetos de un objeto tridimensional (307), 3D, que comprende las etapas de:
  - obtener al menos una representación bidimensional (606), 2D, de dicho objeto;
  - detectar características de imagen en dicha representación 2D (606) obtenida;
  - inferir una forma 3D dispersa correspondiente a las características de imagen utilizando una representación vectorial latente de dicha forma 3D dispersa y utilizando un modelo de forma estadístico multi-perspectiva.

**caracterizado por** comprender además las etapas de

  - extender la forma 3D dispersa a una forma 3D densa encajando la superficie a dicha forma 3D dispersa utilizando un modelo de forma estadístico aprendido que consiste en una superficie media completa del objeto para regularizar dicha forma de superficie de forma específica para una clase de objetos;
  - comparar la forma 3D densa con formas 3D de referencia (605) obtenidas previamente; y por lo tanto
  - identificar un objeto individual de la clase de objetos.
2. El procedimiento según la reivindicación 1, en el que dicha forma 3D dispersa se infiere a partir de características 2D.
3. El procedimiento según la reivindicaciones 1 a 2, en el que dicha clase de objetos contiene objetos no rígidos y dicho modelo de forma estadístico es aprendido utilizando datos 2D y 3D específicos para posibles deformaciones de los objetos en dicha clase de objetos no rígidos.
4. El procedimiento según cualquiera de las reivindicaciones 1 a 3, que comprende además la etapa de verificar la identidad de un objeto individual de una clase de objetos o ayudar en la verificación de la identidad de un objeto individual utilizando dicha forma 3D densa.
5. El procedimiento según las reivindicaciones 1 a 4, en el que dicho objeto (607) puede ser uno o varios de: una cara humana, un cuerpo humano, órgano(s) interno(s) de un cuerpo humano, vasos sanguíneos, animales, órganos internos de un animal, un tumor, producto(s) manufacturado(s) de un proceso industrial, un vehículo, una aeronave, un barco, objeto(s) militar(es).
6. El procedimiento según la reivindicación 1, en el que dichas formas 3D de referencia (605) se almacenan en al menos uno de una memoria no volátil, servidor de base de datos y tarjeta de identificación personal.
7. Un dispositivo (400) para reconocimiento de objetos de un objeto tridimensional (607), 3D, que comprende:
  - medios para obtener al menos una representación bidimensional (606), 2D, de dicho objeto;
  - medios para detectar características de imagen en dicha representación 2D (606) obtenida;
  - medios para inferir una forma 3D dispersa correspondiente a las características de imagen utilizando una representación vectorial latente de dicha forma 3D dispersa y utilizando un modelo de forma estadístico multi-perspectiva.

**caracterizado por** comprender además medios para

  - extender la forma 3D dispersa a una forma 3D densa encajando la superficie a dicha forma 3D dispersa utilizando un modelo de forma estadístico aprendido que consiste en una superficie media completa del objeto para regularizar dicha forma de superficie de forma específica para una clase de objetos;
  - comparar la forma 3D densa con formas 3D de referencia (605) obtenidas previamente, y por lo tanto

identificar un objeto individual de la clase de objetos.
8. El dispositivo (400) según la reivindicación 7, en el que dicha forma 3D dispersa se infiere a partir de características 2D.
9. El dispositivo (400) según la reivindicaciones 7 a 8, en el que dicha clase de objetos contiene objetos no rígidos y dicho modelo de forma estadístico es aprendido utilizando datos 2D y 3D específicos para posibles deformaciones de los objetos en dicha clase de objetos no rígidos.
10. El dispositivo (400) según cualquiera de las reivindicaciones 7 a 9, que comprende además la etapa de

verificar la identidad de un objeto individual de una clase de objetos o ayudar en la verificación de la identidad de un objeto individual utilizando dicha forma 3D densa.

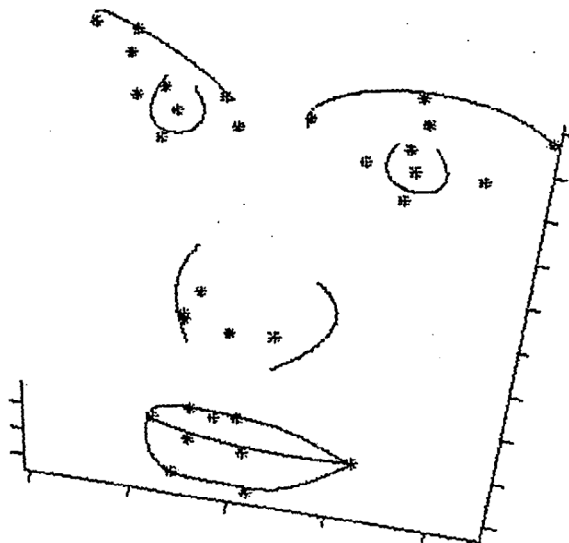
- 5      **11.** El dispositivo (400) según las reivindicaciones 7 a 10, en el que dicho objeto (607) puede ser uno o varios de: una cara humana, un cuerpo humano, órgano(s) interno(s) de un cuerpo humano, vasos sanguíneos, animales, órganos internos de un animal, un tumor, producto(s) manufacturado(s) de un proceso industrial, un vehículo, una aeronave, un barco, objeto(s) militar(es).
- 10     **12.** El dispositivo (400) según las reivindicación 11, en el que dichas formas 3D densas de vasos sanguíneos o órganos recuperadas a partir de proyecciones 2D, por ejemplo, utilizando imágenes de Rayos X, se utilizan para navegar catéteres dirigibles o ayudar a médicos mostrando dicha forma 3D densa.
- 15     **13.** El dispositivo (400) según las reivindicación 11, en el que dichas formas 3D densas de características faciales se utilizan para identificar o para verificar la identidad de un individuo en un sistema de control de acceso o sistema de seguridad, resultando en la aceptación o rechazo de dicho individuo.
- 20     **14.** El dispositivo (400) según las reivindicación 8, que comprende además una interfaz para comunicarse con una tarjeta de identificación personal en la que se almacenan dichas formas 3D de referencia.
- 25     **15.** Un programa de ordenador almacenado en un medio legible por ordenador (402) que comprende una pluralidad de instrucciones que cuando son ejecutadas por una unidad de computación (901) hace que la unidad de computación lleve a cabo el procedimiento de cualquiera de las reivindicaciones 1 a 6.
- 30     **16.** Un sistema para reconocimiento de objetos de un objeto tridimensional, 3D, que comprende:  
       - un dispositivo (400) según la reivindicación 7; y  
       - medios para responder (604) a un resultado de dichos medios de comparación.
- 35     **17.** El sistema según la reivindicación 16, que comprende además medios para identificar y/o verificar un objeto individual de una clase de objetos o ayudar en dicha identificación y/o verificación de dicho individuo utilizando dicha forma 3D densa.
- 18.** El sistema según la reivindicación 16, en el que dichas formas de referencia 3D se almacenan en al menos uno de una memoria no volátil, servidor de base de datos y tarjeta de identificación personal.



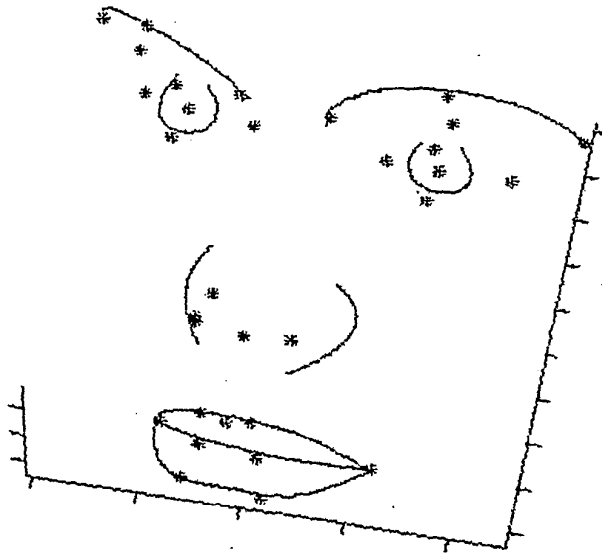
*Fig. 1a*



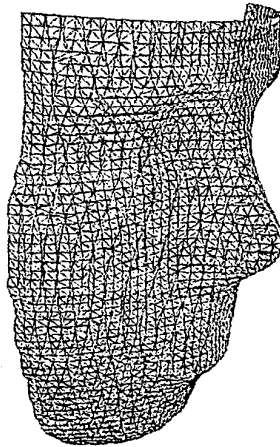
*Fig. 1b*



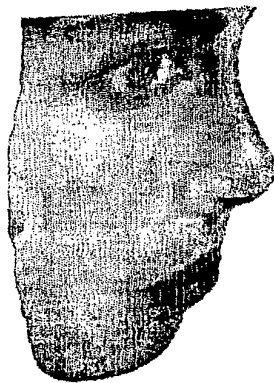
*Fig. 1c*



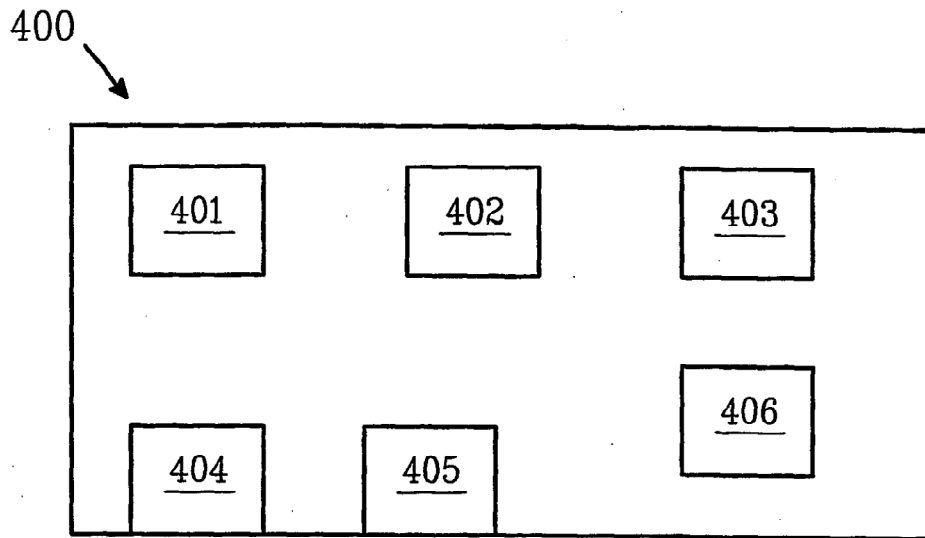
*Fig.2a*



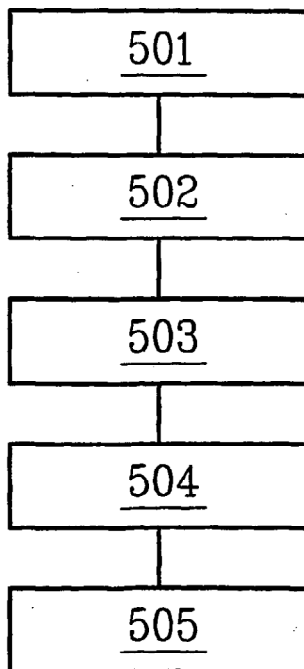
*Fig.2b*



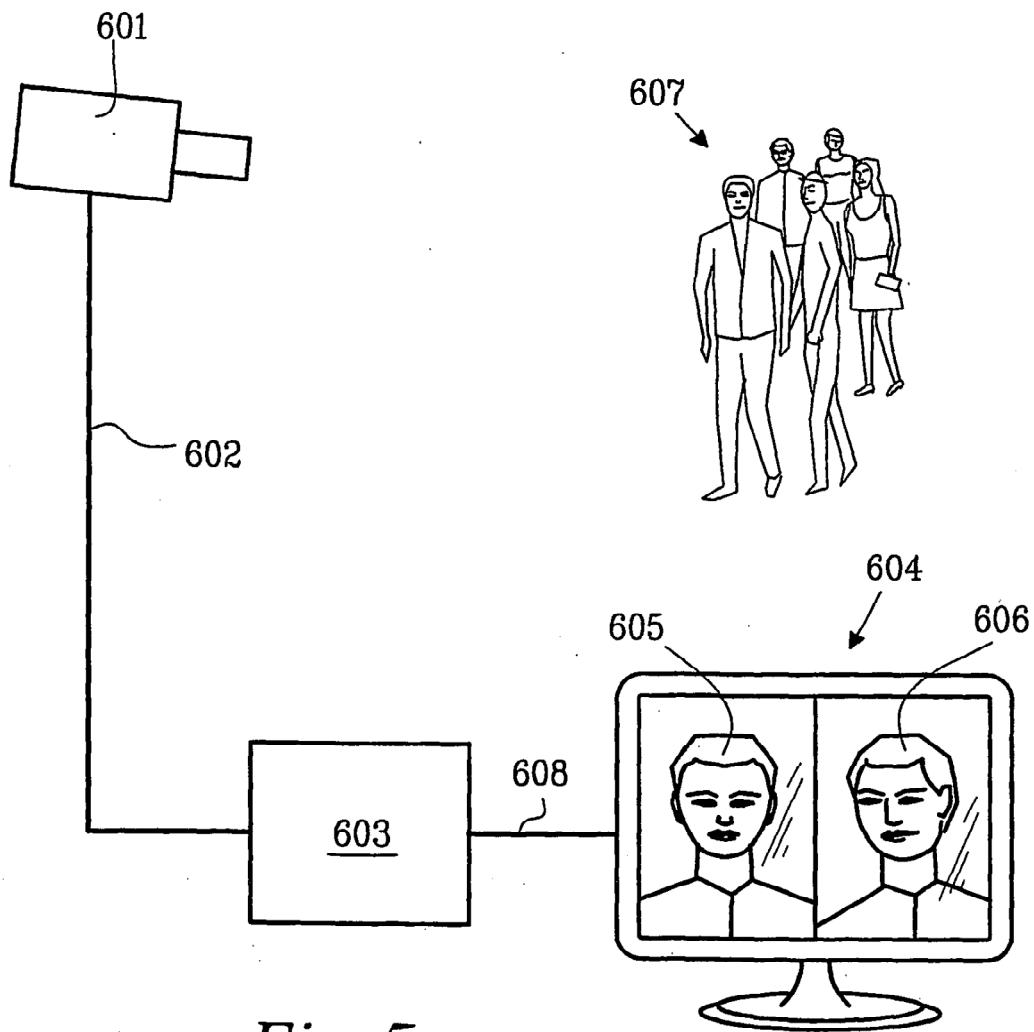
*Fig.2c*



*Fig. 3*



*Fig. 4*



*Fig. 5*