



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 339 100**

21 Número de solicitud: 201030025

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

12

PATENTE DE INVENCION CON EXAMEN PREVIO

B2

22 Fecha de presentación: **14.01.2010**

43 Fecha de publicación de la solicitud: **14.05.2010**

Fecha de la concesión: **07.09.2010**

45 Fecha de anuncio de la concesión: **21.09.2010**

45 Fecha de publicación del folleto de la patente:
21.09.2010

73 Titular/es: **Universidad Politécnica de Madrid
c/ Ramiro de Maeztu, 7
28040 Madrid, ES
Universidad Rey Juan Carlos**

72 Inventor/es: **Bekios Calfa, Juan;
Baumela Molina, Luis y
Buenaposada Biencinto, José Miguel**

74 Agente: **Carvajal y Urquijo, Isabel**

54 Título: **Dispositivo y procedimiento automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes.**

57 Resumen:

Dispositivo y procedimiento automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes que comprende una etapa de entrenamiento y una de clasificación. El entrenamiento comprende: preprocesar imágenes de entrenamiento; construir un clasificador biclase, hombre y mujer, en el espacio unidimensional resultado de un procedimiento de Análisis de componentes Principales (PCA) seguido de un procedimiento de Análisis Discriminante Lineal de Fisher (LDA), siendo el paso clave la elección del conjunto de los autovectores de PCA que mejor discriminan mediante validación cruzada. La etapa de clasificación comprende: preprocesar una imagen; proyectar la imagen sobre el espacio de proyección PCA y LDA utilizando la proyección más discriminante obtenida en el entrenamiento; y clasificar el valor unidimensional resultante de la proyección mediante el clasificador del entrenamiento.

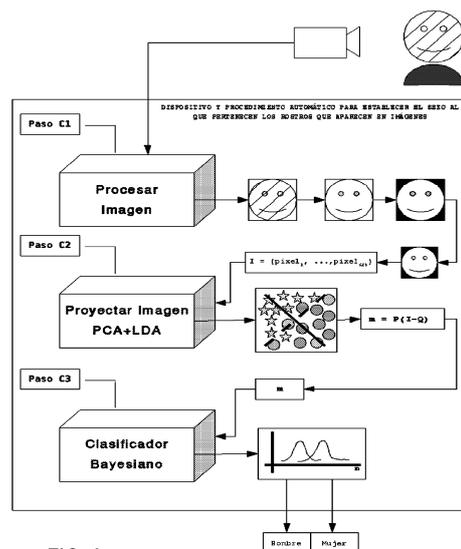


FIG. 6

ES 2 339 100 B2

Aviso: Se puede realizar consulta prevista por el art. 40.2.8 LP.

DESCRIPCIÓN

Dispositivo y procedimiento automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes.

Campo de la invención

La presente invención pertenece al campo técnico de la inteligencia artificial, concretamente al campo de la visión por computador la cual tiene como objeto el reconocimiento de imágenes o patrones de una escena o imagen.

Estado de la técnica

La clasificación de imágenes faciales según su género ha atraído la atención de los investigadores en Visión por Computadora y Reconocimiento de Patrones durante años, siendo SEXNET de Golomb en 1991 el primer intento de reconocer el sexo al que pertenece una cara. A partir de entonces se han desarrollado numerosos trabajos relacionados con este tema, concretamente, se han producido importantes avances en las áreas de la detección automática de rostros en imágenes y el seguimiento automático de los mismos en una secuencia de vídeo, lo que ha multiplicado la aparición de aplicaciones prácticas del análisis facial automático como la detección de rostros, sonrisas o parpadeos en cámaras digitales de fotografía. Los sistemas empotrados de procesamiento de caras como esos poseen recursos de cómputo muy limitados y es crucial para su éxito el desarrollo de procedimientos eficientes de análisis facial.

La patente US6990217 presenta un método global de clasificación de rostros en imágenes según el sexo al que pertenezcan. En ella el entrenamiento se realiza mediante SVM (Support Vector Machine) con un kernel de Funciones de Base Radial los cuales definen un hiperplano con el fin de separar los grupos correspondientes a hombres y mujeres. Tras el entrenamiento, se realiza la clasificación de imágenes mediante el hiperplano definido por SVM.

La patente US7505621 presenta un sistema y un método para la extracción automática de información demográfica a partir de imágenes. El sistema detecta la cara en una imagen, localiza las diferentes componentes, extrae características de las componentes y clasifica las componentes para identificar la edad, el sexo o el origen étnico de la persona o personas de la imagen. En esta patente se demuestra que usar componentes para la clasificación demográfica ofrece mejores resultados que los obtenidos por las técnicas actualmente conocidas. Además, el sistema y el método descritos se pueden utilizar para extraer información demográfica de manera más robusta en entornos donde existe un alto grado de variabilidad de tamaño, forma, color, textura, posición y oclusión. Esta invención realiza una fusión de clasificadores a nivel de componentes utilizando un clasificador multinivel para fusionar resultados de varios clasificadores de componentes demográficas.

En general, el problema de la clasificación de género en imágenes se realiza de dos formas:

- 1) global, utilizando la imagen de la cara completa como atributo de clasificación después de haberla recortado, normalizado su tamaño y normalizada su iluminación,
- 2) basada en características, extrayendo una serie de elementos discriminantes de la imagen facial.

Hasta la fecha, los mejores resultados con clasificadores globales se han obtenido en la patente US6990217 utilizando un clasificador de Máquinas de Vectores Soporte (SVM) y un kernel de Funciones de Base Radial (RBF), a este clasificador se le denominará SVM+RBF. Utilizando SVM+RBF se reportó un 96,6% de acierto en la clasificación utilizando 1775 imágenes de la base de datos FERET utilizando imágenes normalizadas y recortadas y validación cruzada con 5 grupos. Sin embargo, Baluja y Rowley ("Boosting sex identification performance", International Journal of Computer Vision, 71(1)) detectaron un sesgo en la estimación de la patente US6990217 causado por la existencia de imágenes con la misma persona en diferentes grupos (mismo individuo presente en algún grupo de entrenamiento y en el de prueba). Baluja y Rowley consiguieron un 93,5% de acierto en un experimento similar al de dicha patente con alineación manual de las imágenes y una validación cruzada realizada con la presencia de cada individuo únicamente en uno de los grupos.

Las aproximaciones basadas en características utilizan, desde diferencias de niveles de gris de dos píxeles, rectángulos parecidos a wavelets de Haar, bancos de filtros multiescala, hasta Patrones Binarios Locales (LBP) para reconocer el sexo al que pertenece una cara. Shakhnarovich, Viola y Moghaddam ("A unified learning framework for real time face detection and classification", Actas de la IEEE International Conference on Automatic Face and Gesture Recognition, FG 2002) obtuvieron un 79% y un 79,2% de tasa de acierto en clasificación de sexo y etnia en un conjunto de imágenes difíciles descargadas de Internet. Para ello utilizaron rectángulos parecidos a wavelets de Haar y una aproximación basada en AdaBoost para seleccionar las características más discriminantes que permite clasificar la imagen de una cara 50 veces más rápido que las aproximaciones basadas en SVM. Baluja y Rowley utilizan comparaciones directas de los niveles de gris de dos píxeles como clasificadores básicos en un esquema de aprendizaje basado en AdaBoost. Utilizando imágenes de caras alineadas manualmente sacadas de los conjuntos "fa" y "fb" de la base de datos FERET obtienen una tasa de acierto del 94%. Su clasificador es aproximadamente 50 veces más rápido que la solución basada en SVM+RBF de la patente US6990217.

Una conclusión importante de los resultados aparecidos en la literatura es que es bastante difícil compararlos, dado que los investigadores no utilizan la misma base de datos y, cuando la utilizan, frecuentemente no usan las mismas imágenes. Esta es la razón por la que Mäkinen y Raisamo realizaron un gran número de experimentos sobre reconocimiento de sexo ("Evaluation of gender classification methods with automatically detected and aligned faces", IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(3)). Sin embargo, sus pruebas se realizaron con muy pocas imágenes (411), con lo que no se pueden extraer conclusiones definitivas de su trabajo.

Por otro lado, sí que podemos asumir que los procedimientos con mejor tasa de clasificación se encuentran en una tasa de 93% de acierto de la patente US6990217 con SVM+RBF, y de Baluja y Rowley con AdaBoost sobre la base de datos FERET.

La tasa de acierto no es la única variable importante a considerar. La eficiencia computacional es también un aspecto importante, dado que la implantación masiva de dispositivos de análisis demográfico con visión por computadora necesitará procedimientos de clasificación que puedan funcionar, por ejemplo, en cámaras en red o en teléfonos móviles. Hasta ahora, sólo los procedimientos basados en AdaBoost de Baluja y Shakhnarovich para la clasificación de sexo podían cubrir este requisito de eficiencia computacional.

Sin embargo, era deseable un dispositivo y un procedimiento automático para establecer el sexo al que pertenecen los rostros en las imágenes, que se caracterizasen por su sencillez y que utilizaran menos recursos de cómputo, obteniendo resultados capaces de competir con los ofrecidos por los ya existentes.

Descripción de la invención

La presente invención resuelve los problemas existentes en el estado de la técnica mediante el uso de un clasificador basado en uno de los métodos más sencillos de clasificación lineal: Análisis de Componentes Principales (PCA) seguido del Análisis Discriminante Lineal de Fisher (LDA) utilizando un clasificador Bayesiano en el espacio de dimensión reducida que se obtiene del LDA siendo la selección de las características de PCA más discriminantes el paso crucial para el aumento del rendimiento.

A continuación se explicará el funcionamiento del mismo.

Análisis Discriminante Lineal de Fisher (LDA)

Dado un problema de clasificación multiclase con c clases y p vectores ejemplo de dimensión n , el LDA permite obtener una matriz de proyección lineal de los p vectores en un subespacio de dimensión como mucho $d=c-1$ en la que el ratio entre la separación dentro de cada clase y la separación entre las clases es máximo. Los vectores de la base del subespacio transformado, w_i , se obtienen maximizando la siguiente función objetivo:

$$J(w) = \sum_{i=1}^d \frac{w_i^T S_B w_i}{w_i^T S_W w_i} \quad [1]$$

donde S_B y S_W son, respectivamente, las matrices de dispersión entre clases y dentro de las clases y se definen como:

$$S_B = \frac{1}{n} \sum_{j=1}^c n_j (m_j - m)(m_j - m)^T, \quad S_W = \frac{1}{n} \sum_{j=1}^c n_j \sum_{i \in j} (x_i - m_j)(x_i - m_j)^T \quad [2]$$

siendo n_j el número de ejemplos que pertenecen a la clase j , n es el número de ejemplos (suma de todos los n_j), m_j es el vector media de todos los ejemplos pertenecientes a la clase j , m es el vector media de todos los ejemplos y x_i es el vector de ejemplo i .

El máximo de $J(w)$ viene dado por el siguiente problema de autovalores generalizado:

$$S_B W = S_W W D \quad [3]$$

donde W es una matriz cuyas columnas son los vectores w_i y D es la matriz diagonal con los autovalores. La matriz S_B tiene como mucho rango $c-1$ y, generalmente este es el rango de la matriz de proyección LDA, W . En el caso de una matriz S_W no singular, la solución anterior se transforma en:

$$S_W^{-1} S_B W = W D \quad [4]$$

ES 2 339 100 B2

Es importante destacar que el resultado del LDA es la matriz W de dimensión $(c-1) \times n$ y al ser ortonormal la proyección de una imagen I (vector $n \times 1$) sobre el espacio LDA vendrá dado por $m=WI$. En el caso biclase del género y con imágenes de 25×25 píxeles la matriz W tendrá dimensiones 1×625 , esto es será un único vector fila.

5
LDA en el espacio transformado de PCA (PCA+LDA)

10 Cuando se trata con problemas de clasificación de imágenes es muy común tener menos imágenes de ejemplo que características (píxeles). En esos casos la matriz de dispersión dentro de la clase, S_w , es singular y la matriz de proyección LDA, W , no se puede calcular utilizando la ecuación [3].

15 Debido a que la matriz de covarianzas global del conjunto de entrenamiento es $S_M = S_B + S_w$, la solución a la ecuación [2] no cambiará si utilizamos S_M en lugar de S_w . En este caso, realizar Análisis de Componentes Principales (PCA) reteniendo sólo los autovectores asociados con los autovalores no nulos y realizando a continuación LDA en el subespacio transformado mediante PCA es equivalente a realizar sólo LDA en el subespacio original. A partir de ahora, se llamará PCA+LDA a este procedimiento independientemente de los autovectores elegidos en el paso de PCA. Así que, PCA+LDA, en general, no será estrictamente equivalente a la versión de LDA global cuando se descarten autovectores asociados a autovalores no nulos.

20 Por último, si la matriz de la base y el vector media provenientes del PCA son, respectivamente M y Q , entonces la proyección de una imagen I sobre el subespacio de PCA será $L = M^T(I-Q)$. Así que después del PCA+LDA la proyección de una imagen I sobre el subespacio LDA vendrá dada por $L = WM^T(I-Q) = P(I-Q)$. Con lo que la matriz P (en realidad vector $1 \times n$) junto con el vector media Q , será todo lo que se necesitará para proyectar una nueva imagen sobre el espacio de LDA. Es importante destacar que en ese caso las dimensiones de la matriz P son independientes del número de autovectores de PCA retenidos, aunque se tendrán diferentes matrices P cuando los autovectores de PCA sean diferentes.

30 *Selección de Componentes Principales en LDA en dos pasos (PCA-M+LDA)*

30 Cuando se realiza PCA+LDA el punto crucial es la elección de qué autovectores del paso PCA retener para no perder información discriminante. No hay información a priori de qué autovalores distintos de cero son producidos por ruido y cuales son altamente discriminantes.

35 Si se utiliza la matriz de covarianzas global al conjunto de entrenamiento, S_M , en lugar de S_w , LDA resolverá el problema de autovalores dado por $S_M^{-1}S_B W = W D$. En este caso se define un criterio de ordenación de los autovectores de la matriz S_M en términos de su concordancia con la matriz S_B . Para ser discriminante, un autovector de S_M debería encontrarse correlado con al menos un autovector, w_i de S_B . Esto significa que se dará mayor importancia a los autovectores de S_M que sean paralelos al subespacio LDA.

40 La importancia de un autovector de S_M , u_i , vendrá dada por:

$$I_j = \sum_{i=1}^q (u_j^T w_i)^2, q = \text{rango}(S_B) \quad [5]$$

45 donde w_i son los autovectores de S_B .

50 Con PCA-M+LDA se denotará el procedimiento que realiza PCA, ordena los autovectores del PCA de mayor a menor valor de I_j , retiene los k primeros autovectores en el nuevo orden y después se realiza LDA. La proyección de un vector imagen sobre el subespacio LDA definido por PCA-M+LDA tendrá el mismo coste computacional que el procedimiento PCA+LDA. Siendo la única diferencia el cómo se eligen los autovectores de PCA que se utilizan.

55 Como se ha mencionado, la presente invención se refiere a un nuevo procedimiento para la clasificación de imágenes de caras en dos clases: hombre o mujer. El procedimiento se aplicará a imágenes con la cara enmarcada (mediante un procedimiento automático de localización de caras), de tamaño reducido (por ejemplo 25×25 puntos de imagen o píxeles), con una máscara que permite eliminar el fondo y la mayor parte del pelo y ecualizadas para normalizar frente a los cambios de iluminación.

60 El procedimiento tiene dos partes: entrenamiento y clasificación. El entrenamiento consistirá en:

- Paso E1: Tomar un conjunto de imágenes de caras como las descritas en el párrafo anterior (cara centrada, tamaño reducido, máscara colocada y niveles de gris ecualizados) de las que se conoce el sexo con anterioridad.

ES 2 339 100 B2

- Paso E2: Validación cruzada. Dividir el conjunto de imágenes de entrenamiento en N grupos (por ejemplo, 5) y utilizar $N-1$ grupos como entrenamiento y el grupo restante para probar el clasificador.
- 5 • Paso E3: Realizar Análisis de Componentes Principales (PCA) sobre las imágenes de los $N-1$ grupos de entrenamiento. Elegir K autovectores del PCA, con $k=1$ en la primera iteración, y proyectar las *imágenes de entrenamiento* sobre el subespacio resultante, obteniendo un vector de coeficientes correspondientes a los autovectores seleccionados, por cada imagen de entrenamiento. El resultado será una matriz de proyección de PCA, M^T , y el vector media de las imágenes de entrada, Q .
- 10 • Paso E4: A continuación se realiza LDA sobre el conjunto de vectores de coeficientes y se proyecta cada imagen de las imágenes de entrenamiento sobre el espacio unidimensional obtenido en el LDA. El resultado de este paso será la matriz de proyección LDA, W y la matriz de proyección PCA y LDA, $P=WM^T$.
- 15 • Paso E5: Construir un clasificador bayesiano de dos clases gaussianas sobre el espacio unidimensional resultante de la proyección de los vectores de PCA sobre el espacio de género definido por el LDA.
- Paso E6: Clasificar el grupo de imágenes de prueba con el clasificador del paso E5.
- 20 • Paso E7: Si no se han utilizado todos los grupos como grupo de prueba en la validación cruzada ir al Paso E3. En otro caso, la tasa de acierto con K autovectores del PCA es la media de las tasas de acierto de las N pruebas de la validación cruzada.
- Paso E8: Hacer $K = K+1$ hasta K igual al número de píxeles de las imágenes de entrada y volver entonces al Paso E2.
- 25 • Paso E9: Elegir las matrices de proyección P y Q para las que se obtuvo la mejor tasa de acierto. Utilizando estas matrices, la proyección de una imagen de entrada Im sobre el subespacio PCA+LDA vendrá dada por $L=WM^T(Im-Q)=P(Im-Q)$.

30 La parte más importante del entrenamiento, y principal resultado de la invención, es el Paso E3 en el que se elige el mejor conjunto de vectores de la base del subespacio de PCA (autovectores) para la clasificación de género. En el procedimiento se puede utilizar para la elección de los autovectores cualquiera de los siguientes dos métodos:

- 35 ○ Método 1 (PCA+LDA): Ordenar los autovectores del PCA de mayor a menor autovalor asociado. En un proceso iterativo, se tomará primero el autovector con mayor autovalor y a continuación se irá añadiendo un autovector cada vez, siguiendo el orden de mayor a menor autovalor. Es importante destacar que se prueba el clasificador con diferente número de autovectores de PCA y no se retiene un número fijo de autovectores (por ejemplo, los que representen el 95% de la variabilidad) como en otras aproximaciones. Éste es el detalle crucial para obtener
- 40 tasas de acierto comparables con los mejores clasificadores de género.
- Método 2 (PCA-M+PCA): Ordenar los autovectores del PCA de mayor a menor importancia según I_j en la ecuación [5]. En un proceso iterativo se tomará primero el autovector de mayor importancia y a continuación se añadirán autovectores de uno en uno siguiendo el orden de mayor a menor importancia.
- 45

La clasificación de cualquier imagen, una vez realizado el entrenamiento, consistirá en:

- 50 • Paso C1: Tomar una imagen, I , de una cara centrada, tamaño reducido, máscara colocada y niveles de gris ecualizados.
- Paso C2: Proyectar la imagen sobre el espacio LDA utilizando la matriz de proyección P (PCA+LDA) y el vector media Q provenientes del entrenamiento, $m=P(I-Q)$.
- 55 • Paso C3: Utilizar el valor de la proyección de la imagen I sobre el espacio LDA, m , en el clasificador bayesiano biclase cuyos parámetros se calcularon en el entrenamiento.

60 Utilizando el procedimiento que se presenta es posible obtener resultados comparables a los mejores clasificadores de género no lineales (SVM+RBF o AdaBoost) utilizando menos recursos de cómputo tanto en el entrenamiento como en la clasificación. El procedimiento que se presenta permite calcular de una forma simple y rápida si una cara es de un hombre o de una mujer. El procedimiento de entrenamiento es muy sencillo y la tasa de acierto es comparable a aproximaciones mucho más elaboradas en las que se emplean Máquinas de Vector Soporte como en la patente US6990217 o en el AdaBoost de Baluja y Rowley.

65

ES 2 339 100 B2

El procedimiento automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes, comprende:

- una etapa de entrenamiento que a su vez comprende:

- a) preprocesar una pluralidad de imágenes de entrada para adecuarlas a su posterior procesamiento;
- b) dividir las imágenes de entrada en N grupos para realizar un procedimiento de validación cruzada obteniendo $N-1$ grupos de imágenes de entrenamiento para entrenar un clasificador y un grupo de imágenes de prueba para probar dicho clasificador;
- c) realizar Análisis de Componentes Principales (PCA) sobre las imágenes de entrenamiento, I , correspondientes a los $N-1$ grupos obtenidos en la etapa b);
- d) elegir K autovectores del PCA, con $K=1$ en la primera iteración de la etapa de entrenamiento;
- e) obtener M , matriz de la base del PCA, cuyas columnas son los K autovectores elegidos en la etapa d);
- f) obtener Q , matriz media de las imágenes de entrenamiento I ;
- g) proyectar las imágenes de entrenamiento I sobre el espacio PCA obteniendo un vector de coeficientes i igual a $i=M^T(I-Q)$ por cada imagen de entrenamiento I ;
- h) ejecutar el Análisis Discriminante Lineal de Fisher (LDA) sobre el conjunto de vectores de coeficientes i provenientes de la etapa g), obteniendo como resultado la matriz de proyección, W , sobre el espacio LDA;
- i) proyectar cada imagen de las imágenes de entrenamiento, I , sobre el espacio unidimensional de LDA obtenido según $m=WM^T(I-Q)=P(I-Q)$, donde $P=WM^T$ es la matriz de proyección PCA y LDA;
- j) construir un clasificador biclase utilizando los vectores de m correspondientes a cada imagen de las imágenes de entrenamiento I ;
- k) clasificar las imágenes del grupo de prueba con el clasificador construido en la etapa j) calculando y almacenando la tasa de acierto del grupo de prueba correspondiente;
- l) volver al paso c) hasta utilizar cada uno de los N grupos una vez como grupo de prueba para la validación cruzada;
- m) calcular la media de las tasas de acierto de las N pruebas realizadas;
- n) actualizar K sumándole una unidad hasta K igual al número de píxeles de las imágenes de entrada y volver al paso b);
- o) elegir la matriz de proyección PCA y LDA, P , y la matriz media de las imágenes de entrenamiento, Q , con mayor tasa de acierto;
- p) construir un segundo clasificador biclase utilizando la totalidad de imágenes de entrada y las matrices P y Q elegidas en la etapa o);

- una etapa de clasificación que a su vez comprende:

- q) preprocesar una imagen Im ;
- r) proyectar la imagen Im , utilizando la matriz de proyección PCA y LDA, P , y la matriz media de las imágenes de entrenamiento, Q , elegidas en la etapa o) del entrenamiento, mediante $m'=P(Im-Q)$;
- s) clasificar m' con el segundo clasificador construido en la etapa p).

Preferentemente, la elección de los K autovectores del PCA se realiza seleccionando entre:

- un método basado en autovalores que comprende:
 - ordenar los autovectores del PCA de mayor a menor autovalor asociado;

ES 2 339 100 B2

- tomar iterativamente K autovectores, incrementando en cada iteración el valor de K en una unidad, para obtener el número de autovectores del PCA que mayor tasa de acierto obtenga en la clasificación;
- 5 • un método basado en la discriminación entre las clases del clasificador de cada autovector del PCA que comprende:
 - ordenar los autovectores del PCA de mayor a menor valor asociado según un criterio de discriminación entre las clases respecto a las que se clasifica;
 - 10 • tomar iterativamente K autovectores, incrementando en cada iteración el valor de K en una unidad, para obtener el número de autovectores del PCA que mayor tasa de acierto obtenga en la clasificación.

15 De manera preferente, la elección de los K autovectores del PCA a partir de la discriminación entre las clases del clasificador, se calcula sumando las correlaciones al cuadrado de cada uno de los autovectores del PCA elegidos en el paso d), con cada uno de los autovectores de la matriz de covarianzas de los vectores media de cada clase.

La etapa de preprocesamiento de las imágenes de entrada comprende preferentemente las siguientes etapas:

- 20 • ejecutar un procedimiento de detección de caras localizando el área que contiene puntos de interés;
- normalizar el tamaño de las imágenes de entrada siendo las dimensiones preferentes de 25x25 píxeles;
- 25 • ecualizar el histograma asociado a las imágenes de entrada normalizadas;
- aplicar una máscara para eliminar el fondo de las imágenes normalizadas y ecualizadas.

30 El procedimiento de detección de caras puede ser automático o manual, en cuyo caso se localiza el área rectangular donde se encuentra la cara o alternativamente posición de los ojos y centro de la boca.

El dispositivo automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes, comprende:

- 35 • unos medios de adquisición de imágenes configurados para capturar imágenes y conectados a unos medios de procesamiento;
- los medios de procesamiento configurados para la realización de dos etapas:
 - 40 • una etapa de entrenamiento que a su vez comprende:
 - t) preprocesar una pluralidad de imágenes de entrada para adecuarlas a su posterior procesamiento;
 - 45 u) dividir las imágenes de entrada en N grupos para realizar un procedimiento de validación cruzada obteniendo $N-1$ grupos de imágenes de entrenamiento para entrenar un clasificador y un grupo de imágenes de prueba para probar dicho clasificador;
 - v) realizar Análisis de Componentes Principales (PCA) sobre las imágenes de entrenamiento, I , correspondientes a los $N-1$ grupos obtenidos en la etapa u);
 - 50 w) elegir K autovectores del PCA, con $K=1$ en la primera iteración de la etapa de entrenamiento;
 - x) obtener M , matriz de la base del PCA, cuyas columnas son los K autovectores elegidos en la etapa w);
 - 55 y) obtener Q , matriz media de las imágenes de entrenamiento I ;
 - z) proyectar las imágenes de entrenamiento I sobre el espacio PCA obteniendo un vector de coeficientes i igual a $i=M^T(I-Q)$ por cada imagen de entrenamiento I ;
 - 60 aa) ejecutar el Análisis Discriminante Lineal de Fisher (LDA) sobre el conjunto de vectores de coeficientes i provenientes de la etapa g), obteniendo como resultado la matriz de proyección, W , sobre el espacio LDA;
 - 65

ES 2 339 100 B2

- bb) proyectar cada imagen de las imágenes de entrenamiento, I , sobre el espacio unidimensional de LDA obtenido según $m=WM^T(I-Q)=P(I-Q)$, donde $P=WM^T$ es la matriz de proyección PCA y LDA;
- 5 cc) construir un clasificador biclase utilizando los vectores de m correspondientes a cada imagen de las imágenes de entrenamiento I ;
- dd) clasificar las imágenes del grupo de prueba con el clasificador construido en la etapa cc) calculando y almacenando la tasa de acierto del grupo de prueba correspondiente;
- 10 ee) volver al paso v) hasta utilizar cada uno de los N grupos una vez como grupo de prueba para la validación cruzada;
- ff) calcular la media de las tasas de acierto de las N pruebas realizadas;
- 15 gg) actualizar K sumándole una unidad hasta K igual al número de píxeles de las imágenes de entrada y volver al paso u);
- hh) elegir la matriz de proyección PCA y LDA, P , y la matriz media de las imágenes de entrenamiento, Q , con mayor tasa de acierto;
- 20 ii) construir un segundo clasificador biclase utilizando la totalidad de imágenes de entrada y las matrices P y Q elegidas en la etapa hh);
- 25 • una etapa de clasificación que a su vez comprende:
- jj) preprocesar una imagen Im ;
- kk) proyectar la imagen Im , utilizando la matriz de proyección PCA y LDA, P , y la matriz media de las imágenes de entrenamiento, Q , elegidas en la etapa hh) del entrenamiento, mediante $m=P(Im-Q)$;
- 30 ll) clasificar m con el segundo clasificador construido en la etapa ii).

35 Preferentemente, los medios de procesamiento están configurados para elegir los K autovectores del PCA, seleccionando entre:

- 40 • un método basado en autovalores que comprende:
- ordenar los autovectores del PCA de mayor a menor autovalor asociado;
 - tomar iterativamente K autovectores, incrementando en cada iteración el valor de K en una unidad, para obtener el número de autovectores del PCA que mayor tasa de acierto obtenga en la clasificación;
- 45 • un método basado en la discriminación entre las clases del clasificador de cada autovector del PCA que comprende:
- ordenar los autovectores del PCA de mayor a menor valor asociado según un criterio de discriminación entre las clases respecto a las que se clasifica;
 - tomar iterativamente K autovectores, incrementando en cada iteración el valor de K en una unidad, para obtener el número de autovectores del PCA que mayor tasa de acierto obtenga en la clasificación.
- 50

55 De manera preferente, los medios de procesamiento encargados de elegir los K autovectores del PCA a partir de la discriminación entre las clases del clasificador, están configurados para calcular dicha discriminación sumando las correlaciones al cuadrado de cada uno de los autovectores del PCA elegidos en el paso w), con cada uno de los autovectores de la matriz de covarianzas de los vectores media de cada clase.

60

Los medios de procesamiento configurados para realizar el preprocesamiento de las imágenes de entrada realizarán preferentemente las siguientes etapas:

- 65 • ejecutar un procedimiento de detección de caras localizando el área que contiene puntos de interés;
- normalizar el tamaño de las imágenes de entrada;

ES 2 339 100 B2

- ecualizar el histograma asociado a las imágenes de entrada normalizadas;
- aplicar una máscara para eliminar el fondo de las imágenes normalizadas y ecualizadas.

5

La detección de caras puede ser automática o manual, en cuyo caso los medios de procesamiento localizan el área rectangular donde se encuentra la cara o alternativamente posición de los ojos y centro de la boca.

Breve descripción de los dibujos

10

A continuación, para facilitar la comprensión de la invención, a modo ilustrativo pero no limitativo se describirá una realización de la invención que hace referencia a una serie de figuras.

15

La figura 1 muestra posibles imágenes de entrada al procedimiento de clasificación descrito. En la primera fila se muestran las imágenes originales y en la segunda, el resultado de aplicar a dichas imágenes una máscara elíptica para eliminar el fondo y ecualizarlas.

20

La figura 2 muestra los resultados del procedimiento de entrenamiento con validación cruzada y el método 1 como selección de autovectores de PCA antes del LDA. El eje horizontal muestra el índice de autovectores de PCA ordenados, de izquierda a derecha, de mayor a menor variabilidad y el eje vertical, muestra el porcentaje de acierto después de añadir todos los autovectores de mayor variabilidad hasta ese punto.

25

La figura 3 muestra los resultados del procedimiento de entrenamiento con validación cruzada y el método 2 como selección de autovectores de PCA antes del LDA. El eje horizontal muestra el índice de autovectores de PCA ordenados, de izquierda a derecha, de mayor a menor importancia y el eje vertical, muestra el porcentaje de acierto después de añadir todos los autovectores de mayor importancia hasta ese punto.

30

La figura 4 muestra la proyección de todas las imágenes de la base de datos UCN sobre el espacio PCA+LDA (entrenamiento con método 1) obtenido con la base de datos UCN y 430 autovectores de PCA retenidos.

35

La figura 5 muestra la proyección de todas las imágenes de la galería "fa" de la base de datos FERET sobre el espacio PCA+LDA (entrenamiento con método 1) obtenido con la base de datos UCN y 430 autovectores de PCA retenidos.

40

La figura 6 muestra un esquema del procedimiento automático propuesto para establecer el sexo al que pertenecen los rostros que aparecen en imágenes.

La figura 7 muestra un esquema del procedimiento propuesto para entrenar el dispositivo de reconocimiento de sexo.

Descripción detallada de un modo de realización

45

Para demostrar la presente invención, se han evaluado las dos aproximaciones basadas en LDA explicadas con anterioridad y se han comparado con el clasificador SVM+RBF de la patente US6990217 debido a que es uno de los dos mejores procedimientos presentados hasta la fecha, siendo el otro el de Baluja y Rowley, y teniendo ambos la misma tasa de acierto en la clasificación.

50

En los experimentos realizados se ha utilizado una base de datos no pública de la Universidad Católica del Norte (Chile), a la que llamaremos UCN, la base de datos estándar Color FERET y la base de datos Productive Aging Lab Face (PAL) de la Universidad de Tejas en Dallas.

55

- La base de datos UCN consiste en imágenes frontales (una por cada individuo) de estudiantes y profesores de la UCN. Fueron tomadas con cámaras con diferente calidad de imagen. Hay 10.700 individuos, 5646 hombres y 5054 mujeres. En los experimentos realizados se han utilizado 5628 imágenes de hombres y 5041 imágenes de mujeres dado que el detector de caras utilizado no fue capaz de detectar algunas caras.

60

- La base de datos Color FERET es un recurso disponible públicamente para la investigación en análisis facial. Consiste en múltiples imágenes de 994 individuos, 591 hombres y 403 mujeres. En este caso, se utilizó una única imagen por sujeto de la galería "fa" de la base de datos. En los experimentos realizados sólo se utilizaron 402 imágenes de mujeres dado que el detector de caras perdió una cara de mujer.

65

- La base de datos PAL se encuentra disponible únicamente para investigación. Consiste en imágenes frontales de 576 individuos. Imágenes del perfil derecho y de algunas expresiones faciales están disponibles pero únicamente para algunos sujetos. Hay 219 hombres y 357 mujeres divididos en cuatro grupos dependiendo de su edad: 18-29, 30-49, 50-69 y 70-93.

Antes de clasificar se recortan y redimensionan las imágenes a 25x25 píxeles utilizando un procedimiento automático de detección de caras. Cuando se utiliza alineación manual también se utilizó la información sobre la posición de los ojos y del centro de la boca. Adicionalmente, se realizó la ecualización del histograma con el fin de obtener cierta independencia a los cambios de iluminación. Finalmente se aplicó una máscara con forma elíptica para que eliminase los píxeles del fondo y únicamente permitiese visualizar la textura de la cara.

En el primer experimento se realizaron pruebas sobre diferentes bases de datos, utilizando un esquema de validación cruzada con 5 grupos (para SVM+RBF y LDA) y el procedimiento de entrenamiento anteriormente descrito para PCA+LDA y para PCA-M+LDA. En la Tabla 1 se muestran los mejores resultados obtenidos donde entre paréntesis se indica el número de vectores soporte para SVM y el número de autovectores de PCA retenidos antes de realizar LDA.

TABLA 1

Base de datos	Clasificador			
	SVM+RBF	PCA+LDA	PCA-M+LDA	LDA
FERET	90,93% (415)	91,82% (120)	92,12% (120)	77,88%
FERET alineada	91,85% (435)	93,57% (120)	93,47% (60)	76,78%
UCN	93,12% (3104)	92,30% (430)	92,27% (320)	92,15%

Las condiciones experimentales de la realización de “FERET alineada” (imágenes alineadas manualmente) son similares a las de Baluja y Rowley salvo porque se mantienen los sujetos con la misma identidad en un único grupo dado que se utiliza sólo una imagen por individuo. Los mejores resultados con LDA sobre “FERET alineada”, utilizando 794 imágenes de entrenamiento y 199 de prueba, es 76,78% de acierto. Este experimento confirma el mal resultado en el uso de LDA obtenido por la patente US6990217. Dicha patente reportó un porcentaje de acierto del 86,07%, claramente mejor que el ahora presentado, pero la razón de la diferencia es que, para algunas de las identidades, la evaluación realizada en la patente situaba a sujetos con las mismas identidades en diferentes grupos de validación, además, utilizaba un mayor número de imágenes (alrededor de 2000).

Por otro lado, los resultados para el LDA en la base de datos UCN, con alrededor de 10.000 imágenes, es 92,15%. Con ello, se puede concluir que 794 imágenes no aportan suficiente información para que LDA pueda encontrar la dirección de proyección adecuada en un espacio de 625 dimensiones. Sin embargo, incrementando el número de imágenes de entrenamiento a 10.000 se tiene suficientes datos para que el LDA llegue a ser competitivo incluso con el clasificador SVM+RBF con 93,12% de acierto y 92,15% para el LDA en caso de la base de datos UCN.

Por último, se analizan los resultados para las aproximaciones PCA+LDA de la presente invención. En cualquiera de los dos casos, PCA+LDA (Método 1) o PCA-M+LDA (Método 2), se ha realizado validación cruzada con 5 grupos incrementando de forma iterativa el número de autovectores retenidos en el PCA. En la Tabla 1 se presentan los mejores resultados obtenidos de este procedimiento iterativo. Los resultados más interesantes son los de los experimentos con la base de datos FERET en la Tabla 1. Mientras que el LDA global (sin PCA previo) no es capaz de aprender la mejor dirección de proyección utilizando sólo 794 imágenes, sin embargo, tanto PCA+LDA como PCA-M+LDA son capaces de conseguirlo, incluso logran una tasa de acierto ligeramente mejor que SVM+RBF.

Por tanto, ambas versiones del procedimiento presentado en esta invención (PCA+LDA o PCA-M-LDA) permiten lograr una tasa de acierto de 93% comparable con los mejores procedimientos de clasificación de género.

En el segundo experimento se realizó el cruce de diferentes bases de datos, se entrenó con todas las imágenes de una base de datos y se probó con todas las imágenes de otra completamente diferente. Con este experimento se muestran las capacidades de generalización conseguidas por los diferentes procedimientos de clasificación probadas. En la Tabla 2 se muestran los resultados conseguidos y para el caso de SVM+RBF se escribe entre paréntesis el número de vectores soporte empleados.

TABLA 2

Entrenamiento/ Prueba	Clasificador			
	SVM+RBF	PCA+LDA	PCA-M+LDA	LDA
FERET/UCN	81,11% (415)	80,63%	78,33%	58,34%
FERET/PAL	59,24% (415)	71,50%	74,61%	63,55%
UCN/FERET	90,13% (3104)	88,41%	88,31%	87,31%
UCN/PAL	70,46% (3104)	71,84%	71,15%	70,46%

En el segundo experimento se utilizó el mejor número de vectores soporte (para el SVM+RBF) o autovectores de PCA obtenidos en los experimentos anteriores (para PCA+LDA y PCA-M+LDA). De nuevo, cuando se entrena con la base de datos FERET (en este caso con las imágenes recortadas utilizando únicamente el detector facial) y se prueba con UCN (“FERET/UCN” en la Tabla 2) los resultados son malos para el clasificador LDA. Por contra, cuando se entrena con UCN y se prueba con FERET (“UCN/FERET”), la aproximación LDA compite favorablemente con el resto de procedimientos de clasificación. Más aun, aunque el mejor resultado cuando se entrena con FERET y se prueba con UCN (“FERET/UCN”) se consigue para el clasificador SVM+RBF (81,11%), el segundo mejor es el PCA+LDA con una tasa marginalmente inferior de 80,63%. Cuando entrenamos con UCN y probamos con FERET (“UCN/FERET”) también se logra un resultado similar, es decir, el mejor resultado es para el clasificador SVM+RBF con un 90,13% pero el segundo mejor es el PCA+LDA con un 88,41%, un resultado muy cercano. Un resultado obvio de este experimento es que cuantos más datos utilizamos en el entrenamiento (UCN tiene alrededor de 10.000 imágenes) mejor es la capacidad de generalización que exhibirán los clasificadores así entrenados. En el resultado “FERET/PAL” cabe destacar que mientras el clasificador SVM+RBF no generaliza bien, ya que entrenando con FERET presenta una tasa de acierto muy baja sobre la base de datos PAL, el procedimiento propuesto obtiene resultados mucho mejores en ambas versiones: PCA+LDA y PCA-M+LDA.

Si se consideran los tiempos de ejecución, el procedimiento de Baluja y Rowley es el más rápido publicado hasta la fecha. En él, se utilizan diferencias de niveles de gris de dos píxeles, una característica que es muy rápida de calcular. Su procedimiento, es capaz de obtener la misma tasa de acierto que el SVM+RBF, con únicamente 1000 comparaciones de píxeles (19,53 microsegundos de tiempo de ejecución para imágenes de 20x20 píxeles). Observando el número de vectores soporte utilizados en el caso “FERET alineada” de la Tabla 1 (435 vectores soporte), necesitaremos realizar $435 \times 25 \times 25 = 271.875$ operaciones a nivel de píxel para clasificar una imagen. En el caso PCA+LDA o PCA-M+LDA el tamaño de la matriz de proyección es independiente del número de autovectores de PCA utilizados. Las matrices de proyección PCA+LDA y PCA-M+LDA son, de hecho, un vector fila con tantos componentes como el número de píxeles en la imagen, lo que resulta en $1 \times 25 \times 25 = 625$ operaciones a nivel de píxel. Por otro lado, el número de operaciones a nivel de píxel necesarias utilizando SVM+RBF en la base de datos UCN sería de $3104 \times 25 \times 25 = 1.940.000$.

Por tanto, el tiempo de ejecución del procedimiento propuesto (PCA+LDA o PCA-M+LDA) es comparable en requisitos de cómputo al de Baluja y Rowley, sin embargo, el procedimiento propuesto presenta, como ventaja adicional, una considerable reducción de la complejidad tanto en el entrenamiento como en la clasificación.

La realización preferida de la presente invención consistiría en un dispositivo en una cámara de adquisición de imágenes digitales conectada a un sistema de procesamiento de imágenes en el que para cada imagen capturada se detectarían los posibles rostros, se recortarían al tamaño de 25x25 píxeles, se colocaría una máscara elíptica, se ecualizaría y se clasificaría mediante el procedimiento presentado utilizando el método 1 de elección de componentes principales en el entrenamiento.

Como aplicaciones industriales posibles de la invención, entre otras, se pueden destacar:

- Dispositivos de medición de audiencia de televisión que necesitan obtener información demográfica sobre los televidentes.
- Dispositivos de presentación de vídeos publicitarios (“marketing dinámico”) que miden la audiencia incluyendo información demográfica.
- Estudios de mercado en centros comerciales, tiendas, bancos o cualquier otro comercio mediante dispositivos que utilizan una cámara para contar personas y extraer información demográfica.

ES 2 339 100 B2

- Quioscos interactivos de información que presenten una interfaz de diálogo avanzada (un vendedor virtual, por ejemplo) utilizando procesamiento de lenguaje natural y que necesiten de información sobre el usuario extraída automáticamente, como puede ser el género de la persona, para mejorar la interacción.
- 5
- Reconocimiento de caras para el desbloqueo de sistemas o terminales de acceso, autenticación en servicios ofrecidos por operadores telefónicos, etc.
 - Mejora disponible para los sistemas de recuperación de información.

10 Una vez descrita de forma clara la invención, se hace constar que las realizaciones particulares anteriormente descritas son susceptibles de modificaciones de detalle siempre que no alteren el principio fundamental y la esencia de la invención.

15

20

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. Procedimiento automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes, **caracterizado** porque comprende:

- una etapa de entrenamiento que a su vez comprende:
 - a) preprocesar una pluralidad de imágenes de entrada para adecuarlas a su posterior procesamiento;
 - b) dividir las imágenes de entrada en N grupos para realizar un procedimiento de validación cruzada obteniendo N-1 grupos de imágenes de entrenamiento para entrenar un clasificador y un grupo de imágenes de prueba para probar dicho clasificador;
 - c) realizar Análisis de Componentes Principales (PCA) sobre las imágenes de entrenamiento, I , correspondientes a los N-1 grupos obtenidos en la etapa b);
 - d) elegir K autovectores del PCA, con $K=1$ en la primera iteración de la etapa de entrenamiento;
 - e) obtener M, matriz de la base del PCA, cuyas columnas son los K autovectores elegidos en la etapa d);
 - f) obtener Q, matriz media de las imágenes de entrenamiento I ;
 - g) proyectar las imágenes de entrenamiento I sobre el espacio PCA obteniendo un vector de coeficientes i igual a $i=M^T(I-Q)$ por cada imagen de entrenamiento I ;
 - h) ejecutar el Análisis Discriminante Lineal de Fisher (LDA) sobre el conjunto de vectores de coeficientes i provenientes de la etapa g), obteniendo como resultado la matriz de proyección, W, sobre el espacio LDA;
 - i) proyectar cada imagen de las imágenes de entrenamiento, I , sobre el espacio unidimensional de LDA obtenido según $m=WM^T(I-Q)=P(I-Q)$, donde $P=WM^T$ es la matriz de proyección PCA y LDA;
 - j) construir un clasificador biclase utilizando los vectores de m correspondientes a cada imagen de las imágenes de entrenamiento I ;
 - k) clasificar las imágenes del grupo de prueba con el clasificador construido en la etapa j) calculando y almacenando la tasa de acierto del grupo de prueba correspondiente;
 - l) volver al paso c) hasta utilizar cada uno de los N grupos una vez como grupo de prueba para la validación cruzada;
 - m) calcular la media de las tasas de acierto de las N pruebas realizadas;
 - n) actualizar K sumándole una unidad hasta K igual al número de píxeles de las imágenes de entrada y volver al paso b);
 - o) elegir la matriz de proyección PCA y LDA, P , y la matriz media de las imágenes de entrenamiento, Q , con mayor tasa de acierto;
 - p) construir un segundo clasificador biclase utilizando la totalidad de imágenes de entrada y las matrices P y Q elegidas en la etapa o);
- una etapa de clasificación que a su vez comprende:
 - q) preprocesar una imagen Im ;
 - r) proyectar la imagen Im , utilizando la matriz de proyección PCA y LDA, P , y la matriz media de las imágenes de entrenamiento, Q , elegidas en la etapa o) del entrenamiento, mediante $m'=P(Im-Q)$;
 - s) clasificar m' con el segundo clasificador construido en la etapa p).

ES 2 339 100 B2

2. Procedimiento automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes, según la reivindicación anterior, **caracterizado** porque la elección de los K autovectores del PCA se realiza seleccionando entre:

- 5 • un método basado en autovalores que comprende:
 - ordenar los autovectores del PCA de mayor a menor autovalor asociado;
 - tomar iterativamente K autovectores, incrementando en cada iteración el valor de K en una unidad, para obtener el número de autovectores del PCA que mayor tasa de acierto obtenga en la clasificación;
- 10 • un método basado en la discriminación entre las clases del clasificador de cada autovector del PCA que comprende:
 - 15 • ordenar los autovectores del PCA de mayor a menor valor asociado según un criterio de discriminación entre las clases respecto a las que se clasifica;
 - tomar iterativamente K autovectores, incrementando en cada iteración el valor de K en una unidad, para obtener el número de autovectores del PCA que mayor tasa de acierto obtenga en la clasificación.
- 20

3. Procedimiento automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes, según la reivindicación 2, **caracterizado** porque la elección de los K autovectores del PCA a partir de la discriminación entre las clases del clasificador, se calcula sumando las correlaciones al cuadrado de cada uno de los autovectores del PCA elegidos en el paso d), con cada uno de los autovectores de la matriz de covarianzas de los vectores media de cada clase.

4. Procedimiento automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes, según cualquiera de las reivindicaciones anteriores, **caracterizado** porque el preprocesamiento de las imágenes de entrada comprende las siguientes etapas:

- ejecutar un procedimiento de detección de caras localizando el área que contiene puntos de interés;
- 35 • normalizar el tamaño de las imágenes de entrada;
- ecualizar el histograma asociado a las imágenes de entrada normalizadas;
- 40 • aplicar una máscara para eliminar el fondo de las imágenes normalizadas y ecualizadas.

5. Dispositivo automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes, **caracterizado** porque comprende:

- 45 • unos medios de adquisición de imágenes configurados para capturar imágenes y conectados a unos medios de procesamiento;
- los medios de procesamiento configurados para la realización de dos etapas:
 - 50 • una etapa de entrenamiento que a su vez comprende:
 - t) preprocesar una pluralidad de imágenes de entrada para adecuarlas a su posterior procesamiento;
 - 55 u) dividir las imágenes de entrada en N grupos para realizar un procedimiento de validación cruzada obteniendo N-1 grupos de imágenes de entrenamiento para entrenar un clasificador y un grupo de imágenes de prueba para probar dicho clasificador;
 - v) realizar Análisis de Componentes Principales (PCA) sobre las imágenes de entrenamiento, *I*, correspondientes a los N-1 grupos obtenidos en la etapa u);
 - 60 w) elegir K autovectores del PCA, con $K=1$ en la primera iteración de la etapa de entrenamiento;
 - x) obtener *M*, matriz de la base del PCA, cuyas columnas son los K autovectores elegidos en la etapa w);
 - 65 y) obtener *Q*, matriz media de las imágenes de entrenamiento *I*;

ES 2 339 100 B2

- z) proyectar las imágenes de entrenamiento I sobre el espacio PCA obteniendo un vector de coeficientes i igual a $i=M^T(I-Q)$ por cada imagen de entrenamiento I ;
- 5 aa) ejecutar el Análisis Discriminante Lineal de Fisher (LDA) sobre el conjunto de vectores de coeficientes i provenientes de la etapa g), obteniendo como resultado la matriz de proyección, W , sobre el espacio LDA;
- 10 bb) proyectar cada imagen de las imágenes de entrenamiento, I , sobre el espacio unidimensional de LDA obtenido según $m=WM^T(I-Q)=P(I-Q)$, donde $P=WM^T$ es la matriz de proyección PCA y LDA;
- 15 cc) construir un clasificador biclase utilizando los vectores de m correspondientes a cada imagen de las imágenes de entrenamiento I ;
- 20 dd) clasificar las imágenes del grupo de prueba con el clasificador construido en la etapa cc) calculando y almacenando la tasa de acierto del grupo de prueba correspondiente;
- 25 ee) volver al paso v) hasta utilizar cada uno de los N grupos una vez como grupo de prueba para la validación cruzada;
- ff) calcular la media de las tasas de acierto de las N pruebas realizadas;
- gg) actualizar K sumándole una unidad hasta K igual al número de píxeles de las imágenes de entrada y volver al paso u);
- 30 hh) elegir la matriz de proyección PCA y LDA, P , y la matriz media de las imágenes de entrenamiento, Q , con mayor tasa de acierto;
- ii) construir un segundo clasificador biclase utilizando la totalidad de imágenes de entrada y las matrices P y Q elegidas en la etapa hh);
- una etapa de clasificación que a su vez comprende:
 - 35 jj) preprocesar una imagen Im ;
 - kk) proyectar la imagen Im , utilizando la matriz de proyección PCA y LDA, P , y la matriz media de las imágenes de entrenamiento, Q , elegidas en la etapa hh) del entrenamiento, mediante $m=P(Im-Q)$;
 - 40 ll) clasificar m con el segundo clasificador construido en la etapa ii).

6. Dispositivo automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes, según la reivindicación 5, **caracterizado** porque los medios de procesamiento están configurados para elegir los K autovectores del PCA, seleccionando entre:

- un método basado en autovalores que comprende:
 - ordenar los autovectores del PCA de mayor a menor autovalor asociado;
 - 50 • tomar iterativamente K autovectores, incrementando en cada iteración el valor de K en una unidad, para obtener el número de autovectores del PCA que mayor tasa de acierto obtenga en la clasificación;
- 55 • un método basado en la discriminación entre las clases del clasificador de cada autovector del PCA que comprende:
 - ordenar los autovectores del PCA de mayor a menor valor asociado según un criterio de discriminación entre las clases respecto a las que se clasifica;
 - 60 • tomar iterativamente K autovectores, incrementando en cada iteración el valor de K en una unidad, para obtener el número de autovectores del PCA que mayor tasa de acierto obtenga en la clasificación.

65 7. Dispositivo automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes, según la reivindicación 6, **caracterizado** porque los medios de procesamiento encargados de elegir los K autovectores del PCA a partir de la discriminación entre las clases del clasificador, están configurados para calcular dicha discriminación

ES 2 339 100 B2

sumando las correlaciones al cuadrado de cada uno de los autovectores del PCA elegidos en el paso w), con cada uno de los autovectores de la matriz de covarianzas de los vectores media de cada clase.

5 8. Dispositivo automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes, según cualquiera de las reivindicaciones 5-7, **caracterizado** porque los medios de procesamiento están configurados para realizar el preprocesamiento de las imágenes de entrada comprendiendo dicho preprocesamiento las siguientes etapas:

- ejecutar un procedimiento de detección de caras localizando el área que contiene puntos de interés;
- 10 • normalizar el tamaño de las imágenes de entrada;
- ecualizar el histograma asociado a las imágenes de entrada normalizadas;
- 15 • aplicar una máscara para eliminar el fondo de las imágenes normalizadas y ecualizadas.

15

20

25

30

35

40

45

50

55

60

65



FIG. 1

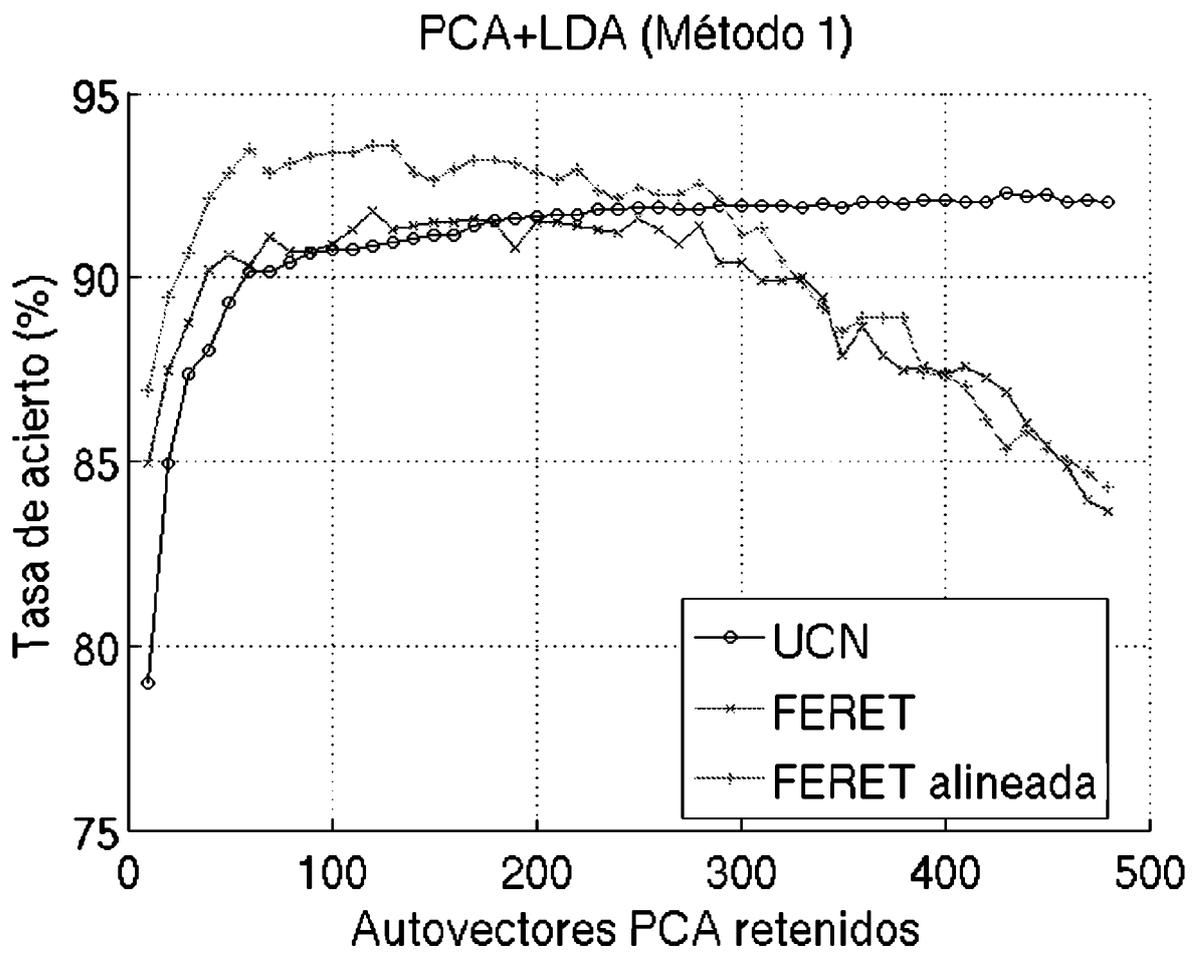


FIG. 2

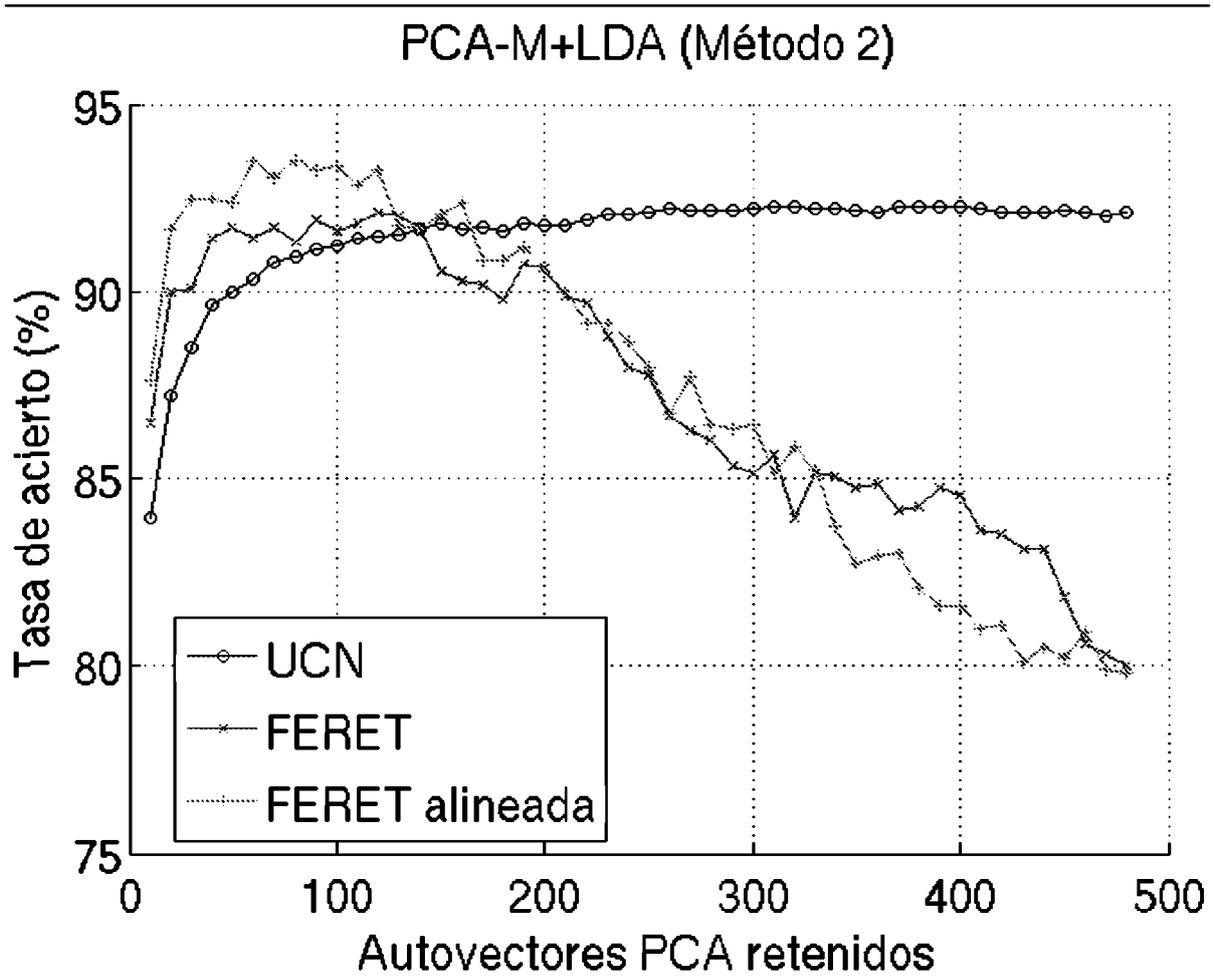
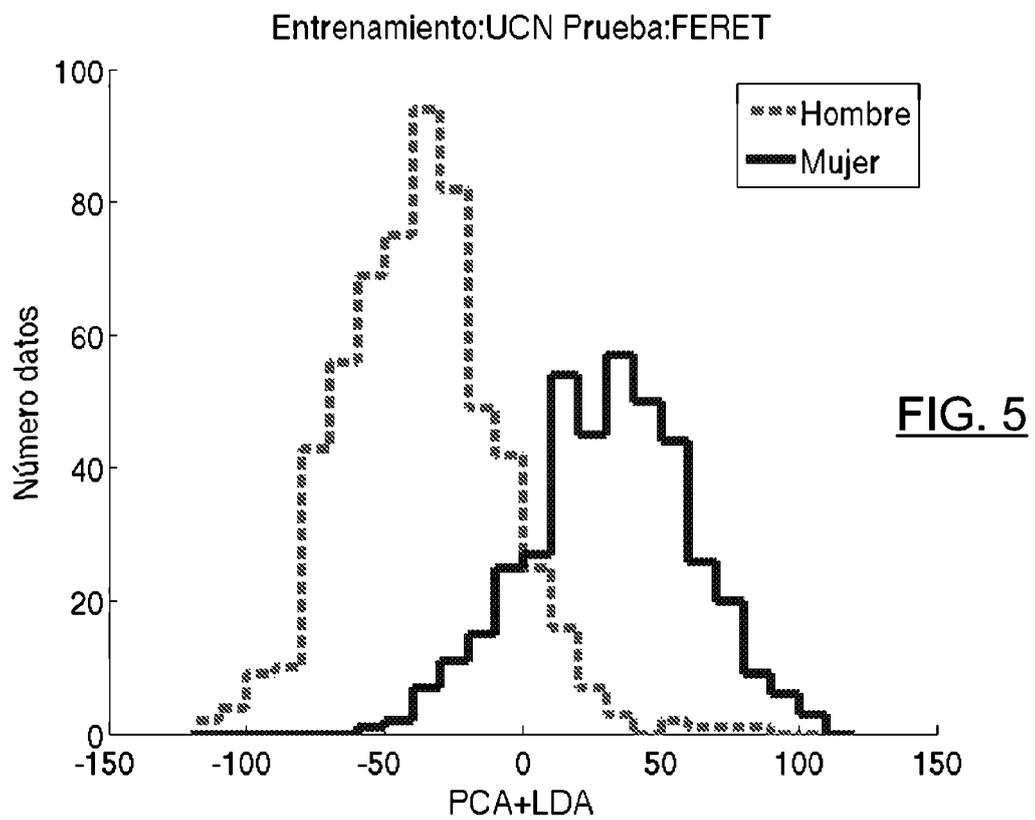
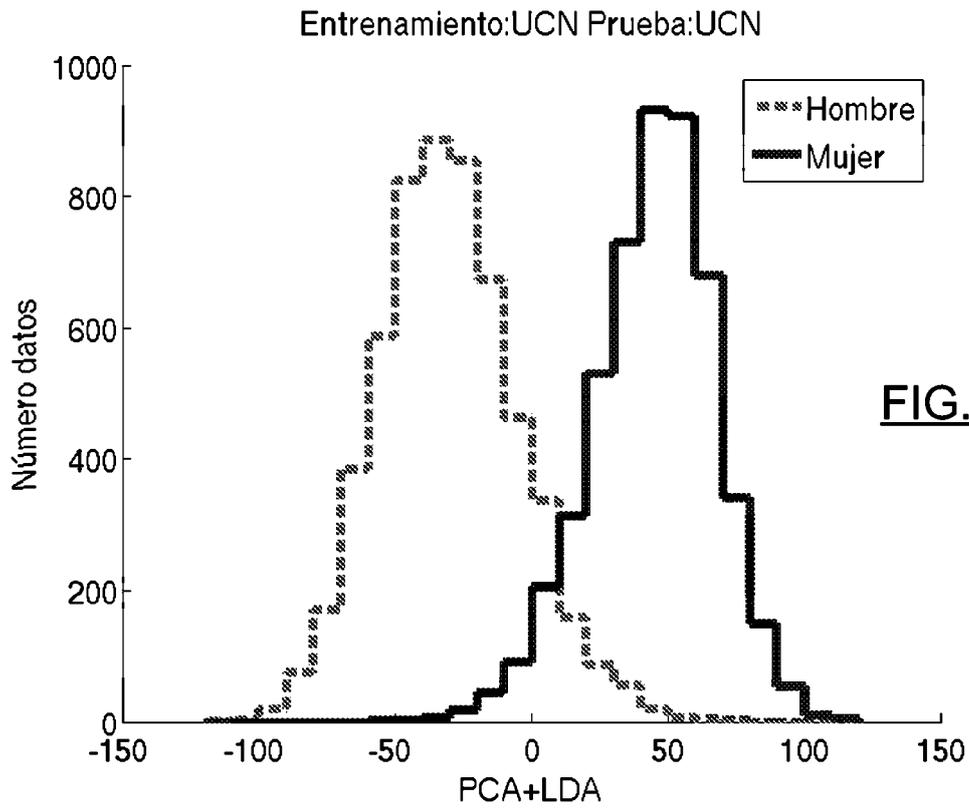


FIG. 3



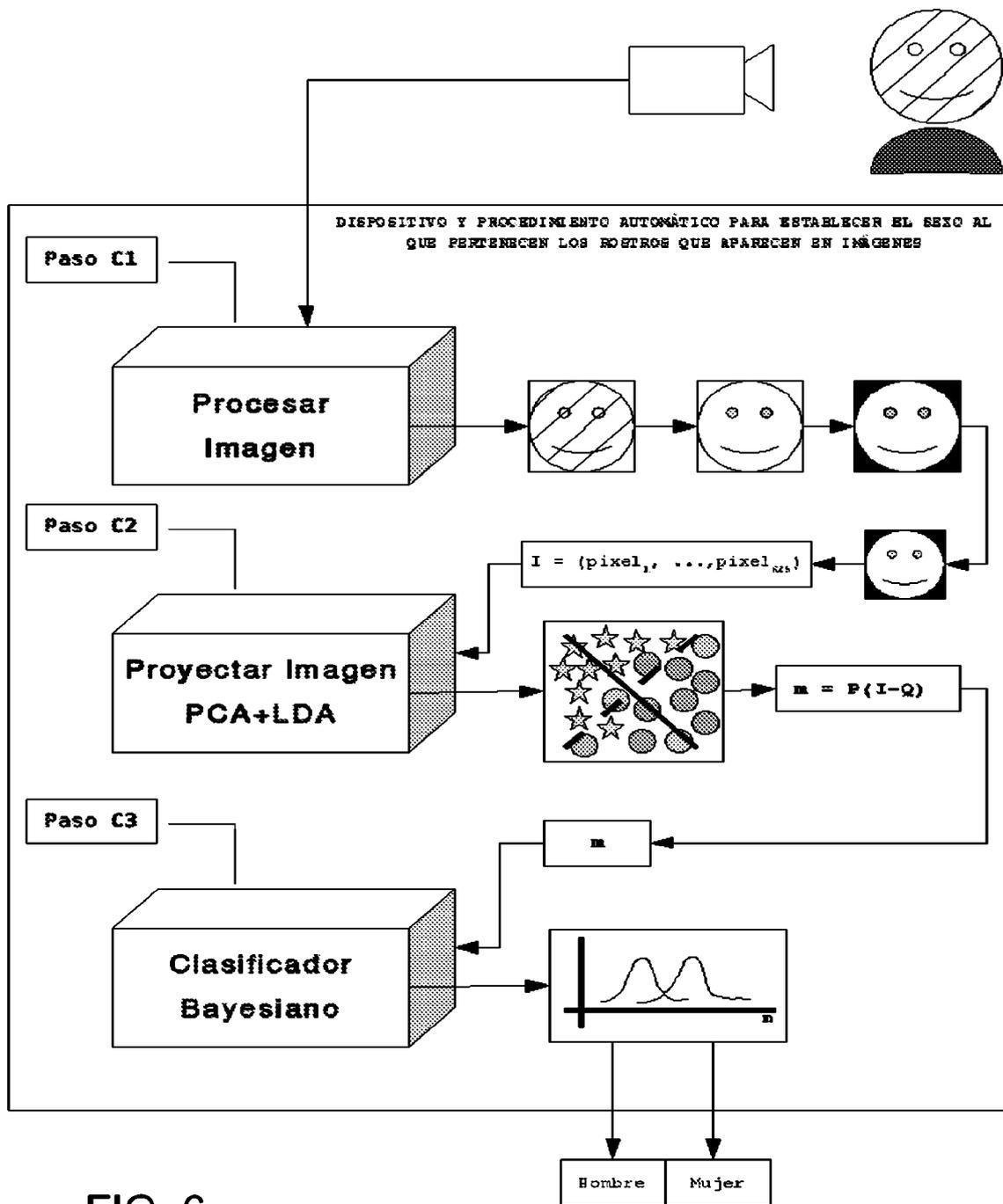


FIG. 6

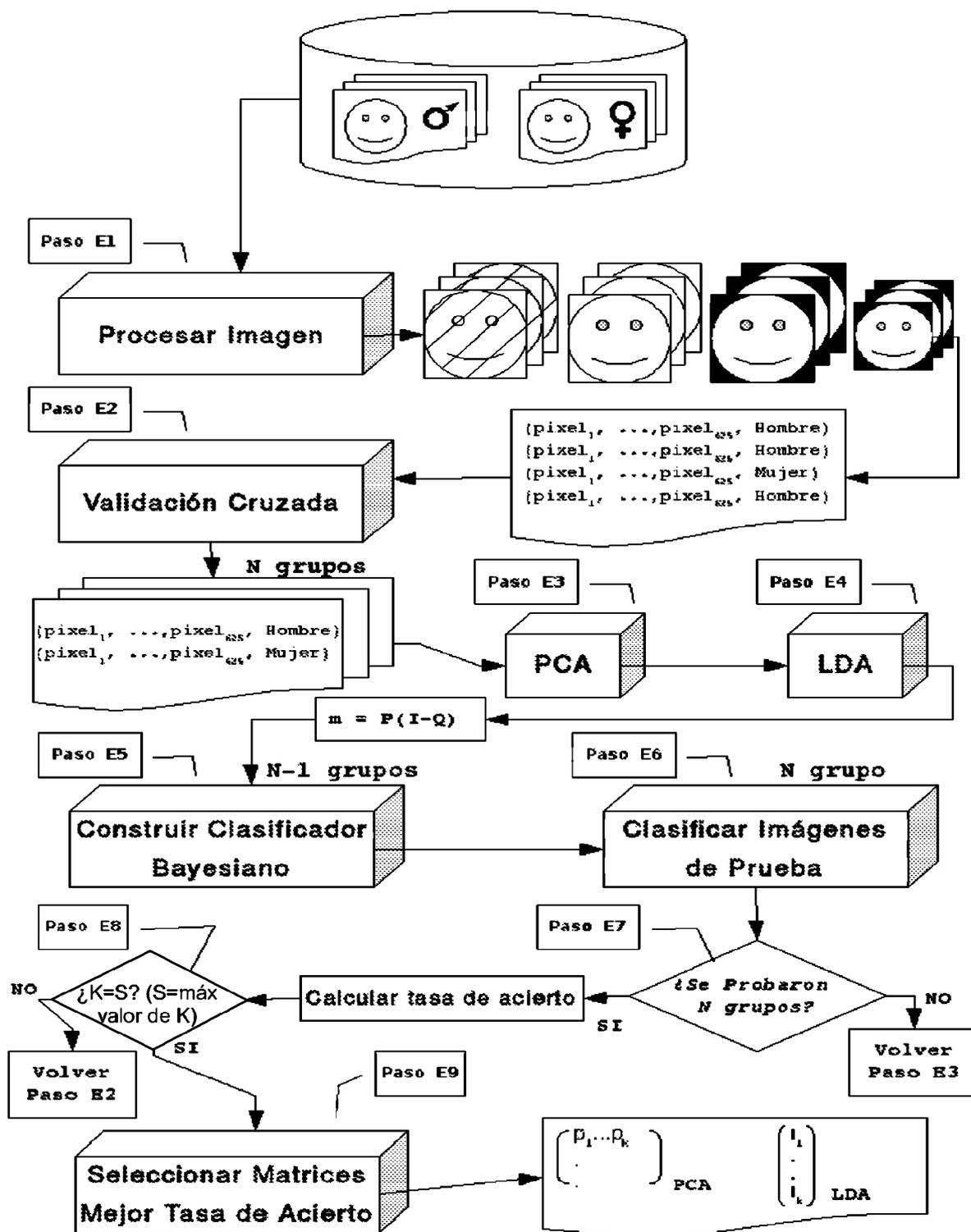


FIG. 7



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

① ES 2 339 100

② Nº de solicitud: 201030025

③ Fecha de presentación de la solicitud: 14.01.2010

④ Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑤ Int. Cl.: G06K 9/00 (2006.01)

DOCUMENTOS RELEVANTES

Categoría	⑥ Documentos citados	Reivindicaciones afectadas
A	US 2005123202 A1 (HWANG et al.) 09.06.2005	1
A	WO 0191041 A1 (UNIV SURREY; KITTLER JOSEF) 29.11.2001	1
A	US 6137909 A (GREINER et al.) 24.10.2000	1
A	US 2007160296 A1 (LEE et al.) 12.07.2007	1

Categoría de los documentos citados

X: de particular relevancia

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

A: refleja el estado de la técnica

O: referido a divulgación no escrita

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

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

El presente informe ha sido realizado

para todas las reivindicaciones

para las reivindicaciones nº:

Fecha de realización del informe

29.04.2010

Examinador

M^a C. González Vasserot

Página

1/4

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

G06K

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

INVENES, EPODOC,WPI

Fecha de Realización de la Opinión Escrita: 29.04.2010

Declaración

Novedad (Art. 6.1 LP 11/1986)	Reivindicaciones 1-8	SÍ
	Reivindicaciones	NO
Actividad inventiva (Art. 8.1 LP 11/1986)	Reivindicaciones 1-8	SÍ
	Reivindicaciones	NO

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

Base de la Opinión:

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

1. Documentos considerados:

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

Documento	Número Publicación o Identificación	Fecha Publicación
D01	US 2005123202 A1	09.06.2005
D02	WO 0191041 A1	29.11.2001
D03	US 6137909 A	24.10.2000
D04	US 2007160296 A1	12.07.2007

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

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

Ninguno de los documentos anteriores muestra un dispositivo ni un procedimiento automático para establecer el sexo al que pertenecen los rostros que aparecen en imágenes que comprende una etapa de entrenamiento y una de clasificación como se reivindican en las reivindicaciones 1-8.

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

El documento D1 se considera el más cercano al Estado de la Técnica, se diferencia de la solicitud en que no se establece el sexo al que pertenecen los rostros de la imagen, pero sí reconoce el rostro de la imagen y además no sigue el procedimiento de la solicitud, no hay validación cruzada de la etapa b, ni se repite el paso c hasta utilizar cada uno de los N grupos una vez como grupo de prueba.

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