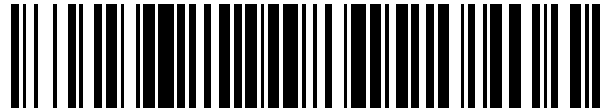


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 399 030**

51 Int. Cl.:

G06K 9/00

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **12.07.2006 E 06787260 (6)**

97 Fecha y número de publicación de la concesión europea: **02.01.2013 EP 1907980**

54 Título: **Indexación automatizada asistida de forma manual de imágenes usando reconocimiento facial**

30 Prioridad:

18.07.2005 US 700282 P

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

25.03.2013

73 Titular/es:

**HYSTERICAL SUNSET LIMITED (100.0%)
P.O. Box 10310 APO
Grand Cayman, KY1-1003, KY**

72 Inventor/es:

**GOLDBERG, DAVID A.;
GRAY, ROBERT;
ANGELL, JOE y
ROSE, JEFFREY**

74 Agente/Representante:

UNGRÍA LÓPEZ, Javier

ES 2 399 030 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

Indexación automatizada asistida de forma manual de imágenes usando reconocimiento facial

5 Campo de la técnica

La presente invención se refiere a la indexación automatizada de fotografías de acuerdo con las personas representadas dentro de la fotografía, usando reconocimiento facial automatizado con asistencia manual.

10 Antecedentes

El uso del reconocimiento facial para la indexación de imágenes está en la actualidad bien establecido en la técnica anterior. Un ejemplo de tal técnica anterior se da en la patente de los EE. UU. 6.526.158 a nombre de Goldberg, y que se proporciona a modo de ejemplo adicionalmente en la patente de los EE. UU. 6.819.783 a nombre de Goldberg, y col.

Ha de observarse, no obstante, que el reconocimiento facial completamente automatizado es imperfecto, y está caracterizado por dos tipos de errores. En el primer caso se encuentran los errores por falsos positivos, en los que los rostros que proceden de dos personas diferentes se asignan como procedentes de la misma persona. En el segundo caso se encuentran los errores por falsos negativos, en los que los rostros que proceden de la misma persona se asignan como procedentes de dos personas separadas. Usualmente puede compensarse un tipo de error con el otro - es decir, es posible, en general, reducir los errores de falsos negativos aumentando los errores por falsos positivos y, a la inversa, reducir los errores por falsos positivos aumentando los errores por falsos negativos. Es posible eliminar en conjunto los errores por falsos positivos, no asignando nunca dos rostros como procedentes de la misma persona y es posible, de forma similar, eliminar los errores por falsos negativos asignando todos los rostros como procedentes de la misma persona. Estos casos extremos no son, no obstante, de importancia práctica y, en general, todos los procedimientos de reconocimiento facial automatizado existirán, generalmente, con errores tanto por falsos positivos como por falsos negativos.

Para conseguir el objetivo de una colección indexada completamente de imágenes, se requiere asistencia manual. Esta asistencia manual puede proceder de una persona que conoce las identidades reales de las personas representadas en la colección de imágenes, tal como en la indexación de una colección de imágenes privada. Como alternativa, como podría tener lugar en la fotografía de acontecimientos (por ejemplo, con la generación de imágenes en cruceros), las fases finales de la indexación podrían estar, en su lugar, asistidas por un empleado de la compañía de generación de imágenes en cruceros.

La dificultad en tal procesamiento manual puede apreciarse cuando se consideran los números de imágenes que pueden estar presentes dentro de una colección. Por ejemplo, en un crucero de una semana de duración de una embarcación con más de 3.000 pasajeros, pueden tomarse más de 25.000 imágenes, comprendiendo 60.000 rostros o más (un promedio de 2-3 personas por imagen). El número de posibles coincidencias de rostro a rostro puede ser entonces superior a 3.000 millones. El reconocimiento facial automatizado es imperfecto y, dependiendo de si son aceptables más falsos positivos o más falsos negativos, el número de conjuntos de rostros que han de revisarse con el fin de establecer un conjunto perfectamente o casi perfectamente indexado puede ser de tantos como decenas de miles, conllevando cientos de horas de trabajo. Incluso una colección personal de pocos miles de imágenes puede conllevar una cantidad de tiempo, reduciendo el atractivo del reconocimiento facial en la indexación de imágenes.

Se pretende que los procedimientos y las composiciones de la presente invención superen estas y otras deficiencias, tal como se describe en las realizaciones a continuación.

El documento DAS M Y COL: "Automatic face-based image grouping for albuming." SYSTEMS, MAN AND CYBERNETICS, 2003. IEEE INTERNATIONAL CONFERENCE ON; [IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN, AND CYBERNETICS], IEEE, PISCATAWAY, NJ, EE. UU., vol. 4, 5 de octubre de 2003 (05-10-2003), páginas 3726-3731, XP010668196 ISBN: 978-0-7803-7952-7 da a conocer un procedimiento para indexar imágenes faciales en una colección de imágenes de acuerdo con las personas que están representadas por las imágenes faciales, que comprende indexar de forma automática las imágenes faciales en la colección con el fin de crear una multiplicidad de conjuntos de imágenes de rostros, en el que cada conjunto comprende unas imágenes faciales que se determina que representan la misma persona, y en el que los conjuntos comprenden unos errores que se eligen del grupo que consiste en errores por falsos positivos y errores por falsos negativos, y presentar a un operador un conjunto, en el que el operador determina de forma manual si la totalidad de las imágenes faciales en el conjunto representan la misma persona.

El documento WO 2004/072897 A da a conocer un sistema semiautomatizado para agrupar imágenes faciales, de acuerdo con el cual el usuario tiene la posibilidad de fusionar de forma manual diferentes agrupaciones que - de acuerdo con su juicio - muestran la misma persona.

El documento FotoFile: A consumer multimedia organization and retrieval system, PROCEEDINGS OF CHI: ACM CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 1999 da a conocer un procedimiento mediante el cual un reconocimiento de rostros se realiza de tal modo que, para cada fotografía en la que el sistema automatizado no está absolutamente seguro de si la persona se ha identificado correctamente, se solicita que el usuario confirme las coincidencias de rostros antes de que las anotaciones asociadas con estos rostros se validen. Por lo tanto, el usuario tiene que comprobar cada una de las fotografías y accionar una información sobre herramientas mediante el ratón con el fin de proceder con esta fotografía específica. Esta etapa tiene que repetirse hasta que se haya confirmado o denegado la última fotografía.

Girgensohn y colaboradores, en el documento "Leveraging Face Recognition Technology to Find and Organize Photos" (2004) describen unos procedimientos para usar el reconocimiento de rostros para mejorar la velocidad y la precisión del hallazgo y la organización de fotografías. Estos indican el uso de paneles de vistas en miniatura faciales, que se clasifican en función de la similitud con una imagen objetivo que representa una persona objetivo. A medida que se añaden vistas en miniatura al conjunto de vistas en miniatura etiquetado con la persona objetivo, mejora la precisión del reconocimiento de rostros, y las vistas en miniatura adicionales se etiquetan como relacionadas con la persona objetivo, la cual puede añadirse entonces a la colección de imágenes que representan la persona.

Sumario de la invención

Un objeto de la presente invención es reducir la cantidad de tiempo que se requiere para asistir de forma manual la indexación automatizada de imágenes usando reconocimiento facial, para reducir el número de comparaciones manuales que se requieren para completar la indexación automatizada de imágenes usando reconocimiento facial y para mejorar la precisión de la indexación automatizada asistida de forma manual de imágenes usando reconocimiento facial.

De acuerdo con la invención, el objeto se soluciona mediante las características de las reivindicaciones independientes, las reivindicaciones secundarias respectivas contienen desarrollos preferidos adicionales de la invención.

Para conseguir los objetos anteriores y otros, y de acuerdo con los fines de la presente invención, tal como se incorpora y se describe ampliamente el presente documento, la presente invención se dirige a un procedimiento para indexar imágenes faciales en una colección de imágenes de acuerdo con las personas que están representadas por las imágenes faciales. El procedimiento puede comprender indexar de forma automática las imágenes faciales en la colección con el fin de crear una multiplicidad de conjuntos de imágenes faciales, en el que cada conjunto comprende unas imágenes faciales que se determina que representan la misma persona, y en el que los conjuntos comprenden unos errores que se eligen del grupo que consiste en errores por falsos positivos y errores por falsos negativos. Este puede comprender también presentar a un operador un conjunto, en el que el operador determina de forma manual si la totalidad de las imágenes faciales en el conjunto representan la misma persona y escindir de forma manual un conjunto que representa más de una persona para dar una multiplicidad de conjuntos, cada uno de los cuales consiste en unas imágenes faciales que representan una única persona. Este puede comprender además presentar al operador por lo menos una imagen facial a partir de un primer conjunto y por lo menos una imagen facial a partir de un segundo conjunto, en el que el operador determina de forma manual si el primer conjunto y el segundo conjunto comprenden unas imágenes faciales que representan la misma persona; y fusionar de forma manual el primer conjunto y el segundo conjunto si los conjuntos comprenden unas imágenes faciales que representan la misma persona. A la conclusión del procedimiento, los conjuntos resultantes de imágenes faciales ya no contienen errores por falsos positivos o por falsos negativos.

La etapa de presentación puede comprender adicionalmente mostrar al operador una primera imagen fuente a partir de la cual se obtuvo un primer rostro seleccionado en el conjunto y una segunda imagen fuente a partir de la cual se obtuvo un segundo rostro seleccionado en el conjunto, en el que la primera imagen fuente y la segunda imagen fuente asisten al operador en la determinación manual de si el primer rostro y el segundo rostro representan la misma persona. La etapa de presentación puede comprender adicionalmente mostrar al operador una primera imagen fuente a partir de la cual se obtuvo un primer rostro elegido del primer conjunto y una segunda imagen fuente a partir de la cual se obtuvo un segundo rostro elegido del segundo conjunto, en el que la primera imagen fuente y la segunda imagen fuente asisten al operador en la determinación manual de si el primer rostro y el segundo rostro representan la misma persona.

La etapa de presentación puede comprender la generación de una imagen en miniatura de por lo menos un rostro en el conjunto, lo que puede comprender la determinación de las ubicaciones de los ojos dentro de la imagen que comprende el rostro, poner a escala la imagen de tal modo que los ojos están separados por un número predeterminado de píxeles, y recortar la imagen con respecto a las ubicaciones de los ojos.

La etapa de presentación puede comprender la generación de una imagen en miniatura de por lo menos un rostro que procede del primer conjunto y por lo menos un rostro que procede del segundo conjunto, lo que puede comprender la determinación de las ubicaciones de los ojos dentro de la imagen fuente a partir de la cual se obtuvo

el rostro, poner a escala la imagen de tal modo que los ojos están separados por un número predeterminado de píxeles, y recortar la imagen con respecto a las ubicaciones de los ojos.

5 La etapa de presentación puede comprender adicionalmente presentar al operador una o más imágenes faciales a partir de un tercer conjunto.

10 El procedimiento puede comprender adicionalmente una segunda indexación automatizada que se realiza sobre la colección a la cual se han añadido imágenes incrementales, utilizando los conjuntos resultantes de imágenes faciales, en el que cada conjunto resultante individual es un subconjunto de los conjuntos que se producen en la segunda indexación automatizada. Las imágenes faciales a partir de las imágenes incrementales pueden añadirse, como máximo, a uno de los conjuntos resultantes.

15 La presente invención se dirige además a eliminar las asociaciones de falsos positivos entre las imágenes faciales que se han indexado de forma automática a partir de una colección de imágenes en un conjunto representativas de una única persona. El presente procedimiento puede comprender presentar a un operador el conjunto de imágenes faciales, seleccionar el operador de forma manual una o más imágenes faciales que son asociaciones de falsos positivos con otras imágenes faciales dentro del conjunto, y eliminar las imágenes faciales seleccionadas a partir del conjunto. El conjunto resultante de imágenes faciales puede ahora no contener errores por falsos positivos.

20 Las imágenes faciales pueden ordenarse por el momento en el que se capturaron las imágenes. Así mismo, las imágenes faciales pueden ordenarse por la similitud de las imágenes faciales según se determine por medios automatizados.

25 El procedimiento puede comprender también crear un nuevo conjunto a partir de las imágenes faciales que se han eliminado, en el que el nuevo conjunto de imágenes faciales es representativo de una segunda persona.

30 La presente invención se dirige también a un sistema para indexar imágenes faciales en una colección de imágenes de acuerdo con las personas que están representadas por las imágenes faciales. El sistema puede comprender un indizador automatizado que asocia imágenes faciales para dar una multiplicidad de conjuntos, determinándose que cada conjunto de la cual comprende unas imágenes faciales representativas de la misma persona, en el que los conjuntos comprenden unos errores que se eligen del grupo que consiste en errores por falsos positivos y errores por falsos negativos. El sistema se ejecutará por un operador. El sistema puede comprender además una pantalla de escisión que presenta imágenes faciales a partir de un primer conjunto, una herramienta de selección que permite que el operador seleccione una o más imágenes faciales a partir del primer conjunto que son asociaciones de falsos positivos con otras imágenes faciales dentro del primer conjunto, y una función de escisión que permite que el operador elimine del primer conjunto las imágenes faciales seleccionadas. El sistema también puede comprender una pantalla de fusión que presenta imágenes faciales a partir de un segundo conjunto e imágenes faciales a partir de un tercer conjunto, en el que una o más imágenes faciales a partir del segundo conjunto tienen similitud con una o más imágenes faciales a partir del tercer conjunto según algún criterio del indizador automatizado, y una función de fusión que permite que el operador fusione las imágenes faciales a partir del tercer conjunto con el segundo conjunto/ La aplicación por el operador de la función de escisión a unas imágenes faciales que se seleccionan con la herramienta de selección en la pantalla de escisión elimina las asociaciones de falsos positivos, y la aplicación por el operador de la función de fusión a las imágenes faciales que se presentan en la pantalla de fusión elimina las asociaciones de falsos negativos, de tal modo que las imágenes faciales indexadas resultantes con contendrán asociaciones de falsos positivos y de falsos negativos.

50 Las imágenes faciales presentadas en la pantalla de escisión pueden ordenarse por el momento en el que se capturaron las imágenes a partir de las cuales se obtuvieron las imágenes faciales. Las imágenes faciales presentadas en la pantalla de escisión pueden ordenarse por la similitud de las imágenes faciales según se determine por el indizador automatizado. Las imágenes faciales presentadas en la pantalla de escisión pueden ser un subconjunto de las imágenes fuente a partir de las cuales se obtuvieron las imágenes faciales. La pantalla de escisión puede comprender además una o más imágenes fuente a partir de las cuales se obtuvieron las imágenes faciales. Las imágenes faciales que se presentan en la pantalla de fusión son un subconjunto de las imágenes fuente a partir de las cuales se obtuvieron las imágenes faciales.

55 **Breve descripción de los dibujos**

60 La figura 1 es un diagrama de flujo de proceso de la indexación automatizada asistida de forma manual de la presente invención.

La figura 2A es un diagrama esquemático de una pantalla de ordenador que comprende una visualización de unas vistas en miniatura de rostros que pueden usarse por un operador para realizar una escisión.

La figura 2B es un diagrama esquemático de una pantalla de ordenador que puede usarse por un operador para realizar una escisión, al igual que en la figura 2A, que además comprende unas imágenes fuente completas de los rostros para un subconjunto de las vistas en miniatura de rostros.

65 La figura 3 es un diagrama esquemático de una pantalla de ordenador que comprende unas vistas en miniatura de rostros que pueden usarse por un operador para realizar una fusión.

La figura 4 es un diagrama esquemático de una pantalla de ordenador que comprende unas filas de conjuntos de rostros que pueden usarse por un operador para realizar con rapidez una escisión o una fusión sobre un número de conjuntos de una sola vez.

5 Mejores modos para llevar a cabo la invención

Terminología

10 Una “fotografía” significa una representación física de una persona, tal como podría imprimirse un proceso con plata, sublimación de tinta, u otro proceso sobre un sustrato de papel.

Una “imagen” de una persona significa una representación de una persona, la cual puede ser electrónica (por ejemplo, un archivo JPG) o física (por ejemplo, una fotografía).

15 Una “persona” es una persona que se representa en una imagen. El plural de persona se indica como “personas” en la presente descripción.

Un “rostro” es una representación de una persona dentro de una imagen.

20 Una “colección” significa un conjunto de imágenes.

Un “operador” es la persona que realiza la asistencia manual en la indexación de las imágenes.

25 La expresión “coincidencia” en relación con dos rostros significa que los dos rostros son representaciones de la misma persona.

La expresión “no coincidencia” en relación con dos rostros significa que los dos rostros no son representaciones de la misma persona.

30 Una “asociación” es la identificación, ya sea por medios automatizados o manuales, de dos o más rostros como pertenecientes a la misma persona. Una asociación se realiza o bien “asociando” o bien “asignando” rostros.

Una “asociación de falsos positivos” es una asociación que se realiza entre dos rostros que no son representativos de la misma persona.

35 Una “asociación de falsos negativos” es la falta de asociación entre dos rostros que son representativos de la misma persona.

Un “conjunto” es un grupo de rostros que están asociados, o asignados, entre sí.

40 Un “índice” es un agrupamiento de rostros dentro de una colección de imágenes en conjuntos.

45 Ha de apreciarse que las breves definiciones anteriores son solo por conveniencia, y no engloban completamente los significados de las expresiones tal como se usan en la memoria descriptiva a continuación. Otros significados dentro de la presente invención se descubrirán en el siguiente análisis.

Visión de conjunto

50 La figura 1 es un diagrama de flujo de proceso de la indexación automatizada asistida de forma manual de la presente invención. En una primera etapa 100 de la indexación automatizada, una indexación preliminar de rostros se realiza de una forma automatizada sin una entrada humana directa en la coincidencia de rostros, formando un índice de los conjuntos (es decir, unos rostros que están asociados uno con otro). En la etapa 100, las ubicaciones de rostros se extraen de las imágenes. Estos rostros se inscriben a continuación, de tal modo que los valores de píxeles se expresan de una forma que permite el establecimiento de coincidencias de rostros. Esta inscripción puede implicar el uso de codificación en el dominio del tiempo o de la frecuencia, el uso de análisis de componentes principales, codificación del rostro usando algoritmos de redes neuronales u otros medios de este tipo. Se establecen entonces coincidencias de los rostros inscritos de este modo entre sí, y se obtiene una puntuación que indica su similitud. Para los fines del presente análisis, se supondrá que una puntuación mayor indica una similitud más alta, a pesar de que los análisis a continuación de la puntuación funcionarían a la inversa con el mismo efecto si una puntuación inferior indicara una similitud más alta.

65 Dadas las puntuaciones, las personas se ensamblan entonces en conjuntos de acuerdo con las áreas de una similitud más alta. En general, existen dos procedimientos de realización de la presente etapa, a pesar de que otros procedimientos son consistentes con la presente invención. En un primer procedimiento, los rostros con unas puntuaciones muy altas se asocian entre sí como conjuntos “semilla”. Entonces, los rostros restantes con unas puntuaciones inferiores se asocian entonces con los conjuntos, en general con el conjunto que comprende los

rostros para los cuales los rostros restantes tienen sus puntuaciones más altas.

En un segundo procedimiento, los rostros se almacenan en un "grafo" (tal como se conoce en la técnica de la ciencia informática) que comprende los rostros en los nodos, y las puntuaciones entre dos rostros en los bordes. En el presente caso, los bordes de puntuaciones inferiores pueden recortarse, hasta que quedan los rostros que están asociados con el mismo conjunto.

Existen numerosos procedimientos de indexación automatizada de rostros además de los anteriores, y cada uno de estos tiene diferentes éxitos dependiendo de los procedimientos empleados, la calidad de las imágenes y la distribución de personas dentro de la colección de imágenes. Por ejemplo, una colección con un pequeño número de personas (por ejemplo, a partir de una colección de fotografías personales) tendrá, en general, diferentes errores de indexación asociados con diferentes procedimientos que una colección con un gran número de personas (por ejemplo, a partir de la generación de imágenes de acontecimientos, tal como la fotografía en cruceros). La totalidad de estos procedimientos tienen errores que pueden beneficiarse de la presente invención.

En ambos de los procedimientos anteriores, existen criterios mediante los cuales el sistema decide si un rostro es un miembro de un conjunto - es decir, es representativo de la misma persona. Este criterio, expresándose en términos generales, comprende uno o más umbrales de decisión, a un lado de los cuales el rostro no está asociado con la persona, y al otro lado de los cuales se considera que el rostro está asociado con la persona.

Ha de observarse que las puntuaciones pueden incluir información además de las puntuaciones de reconocimiento facial, incluyendo información tal como la concerniente a si los rostros provienen de la misma imagen, provienen de unas imágenes tomadas casi al mismo tiempo, o provienen de unas imágenes que tienen las mismas otras personas. Un ejemplo de tales puntuaciones compuestas se da en la patente de los EE. UU. 6.819.783 a nombre de Goldberg, y col.

Una vez que los rostros se han ensamblado en conjuntos en la primera etapa 100 de la forma que se describe anteriormente, o por otros medios automatizados, ahora quedan tanto errores por falsos positivos como por falsos negativos que deberían eliminarse. En general, en la presente invención se eliminan estos errores en un procedimiento en dos etapas que se describe a continuación.

En una segunda etapa 200, se presentan al operador unas imágenes a partir de un conjunto cada vez. El operador "escinde" de ese conjunto los rostros que no son representativos de la misma persona. De esta forma, las coincidencias de falsos positivos se eliminan. El presente proceso se denomina "división". Los rostros que se escinden de un conjunto pueden o bien convertirse en su propio conjunto o bien, como alternativa, se permite que entonces se establezcan potencialmente, de forma automática, coincidencias de los rostros que se escinden del conjunto con otro conjunto. Por ejemplo, si un rostro pudiera coincidir potencialmente con dos conjuntos diferentes (un primer conjunto y un segundo conjunto), con una mejor puntuación de coincidencia con el primer conjunto, este se encontrará colocado, en general, con el primer conjunto de acuerdo con el criterio de puntuación en la etapa de indexación automatizada 100. Si el rostro se escinde entonces de su coincidencia original con el primer conjunto en la etapa de escisión 200, este puede entonces asociarse potencialmente de forma automática con el segundo conjunto.

En una tercera etapa 300, los conjuntos que tienen un cierto grado de similitud pueden compararse uno con otro, para que el operador decida de forma manual si los dos conjuntos diferentes son representativos de la misma persona. De esta forma, los errores de falsos negativos se eliminan. La razón para que los rostros en los dos conjuntos no estuvieran asociados con anterioridad en un único conjunto podría ser que el grado de similitud no fuera lo bastante alto para permitir una indexación automatizada. Como alternativa, los rostros en un conjunto podrían haberse asociado con otro conjunto de una similitud más alta, del cual estos estaban separados en la etapa de la escisión 200. Si el operador considera los dos conjuntos como representativos de la misma persona, los rostros que proceden de los dos conjuntos se combinan en un único conjunto. El presente proceso se denomina "fusión".

La etapa de escisión 200 y la etapa de fusión 300 se describirán a continuación con más detalle.

División 200

En la etapa de escisión 200, los rostros que se asociaran de forma inadecuada en la indexación automatizada para dar un conjunto (es decir, falsos positivos) se escinden de forma manual en conjuntos diferentes. La figura 2A es un diagrama esquemático de una pantalla de ordenador 210 que comprende unas vistas en miniatura de rostros que pueden usarse por un operador para realizar la escisión 200. Cada uno de los recuadros 212 comprende una imagen del rostro específico dentro de la imagen a partir de la cual el rostro se recuperó, y cada una de los rostros se ha asociado mediante la etapa de indexación automatizada 100 para encontrarse en el mismo conjunto (es decir, estos representan la misma persona). Este conjunto indicaba la "persona actual" en la etapa de escisión 200. O bien cada conjunto a partir de la etapa de indexación automatizada 100 puede presentarse al operador, o de otro modo solo aquellos conjuntos en los que los umbrales de decisión para las asociaciones en el conjunto estuvieran por

debajo de un cierto umbral de decisión (tal como se describe anteriormente), siendo por lo tanto inciertos.

Los rostros pueden colocarse en la pantalla como la totalidad de la imagen, en la que el rostro se resalta, por ejemplo rodeando el rostro con un color que se distinga bien, colocando puntos en las ubicaciones de los ojos del rostro, mostrando en gris u oscureciendo las partes de la imagen que no son el rostro, o por otros medios que indican de forma no ambigua el rostro en la imagen que se ha asignado a la persona actual.

Un procedimiento preferido es la creación de una imagen "en miniatura" del rostro. La fase inicial del reconocimiento facial automatizado es, en general, el hallazgo de un rostro, lo que implica en general la determinación de la ubicación de los ojos del rostro. Dada la ubicación de los ojos, puede calcularse una región aproximadamente rectangular que comprende la cabeza. Esta región tiene, preferiblemente, entre 2 y 4 veces la anchura de la distancia entre los ojos medidos de este modo, y preferiblemente los ojos se encuentran a una distancia de entre un 50 % y 75 % aquella entre la parte de debajo y la parte de arriba de la vista en miniatura creada. Las vistas en miniatura se construyen del mismo tamaño, y los rostros se ponen a escala para ser de aproximadamente el mismo dentro de las vistas en miniatura. Los rostros que se asignan a la persona actual se centran dentro de las vistas en miniatura. En el análisis a continuación, se hará referencia al recuadro 212 como su realización preferida, la vista en miniatura 212. En general, la generación de una vista en miniatura comprende localizar los ojos, y a continuación poner a escala y recortar la imagen de tal modo que los ojos se encuentran en la misma ubicación dentro de la imagen en miniatura.

Debido a que solo hay un rostro centrado dentro de la vista en miniatura 212 (otros rostros se encontrarán hacia los bordes de la vista en miniatura), la identificación del rostro en la imagen que se asignó a la persona actual es instantánea y, al ser los rostros del mismo tamaño, son más fáciles de comparar.

En la figura 2A, los identificadores dentro de las vistas en miniatura 212 indican las personas concretas a partir de las cuales se obtuvieron los rostros en las vistas en miniatura 212. Por ejemplo, "A1" hace referencia al primer rostro que procede de la persona "A", mientras que "B2" hace referencia al segundo rostro que procede de la persona "B". En el ejemplo en la figura 2A, el conjunto que se está presentando comprende doce rostros que proceden de cuatro personas diferentes ("A", "B", "C" y "D"). A pesar de que el número de rostros que proceden de la persona A es el más grande, no es necesario considerar que la persona actual es la persona A, sino que puede ser, como alternativa, cualquiera de las personas cuyos rostros se han asignado a la persona actual, y que las otras sean las asignaciones de falsos positivos.

El operador a continuación escinde los rostros que proceden de la misma persona de aquellos que no proceden de la misma persona. El operador selecciona una o más vistas en miniatura 212 usando, por ejemplo, técnicas de selección múltiple de sistema operativo Windows convencionales. En el presente modo, mantener pulsadas las teclas de Control permitirá que se alterne un conjunto de vistas en miniatura 212 como parte de la selección. Por ejemplo, en la figura 2A, se seleccionan las vistas en miniatura 212 B1 y B2, y la selección se indica mediante un límite grueso. Como alternativa, arrastrar el cursor seleccionaría las vistas en miniatura 212 dentro del área de arrastre. Hacer clic sobre una vista en miniatura 212, mantener presionada la tecla de mayúsculas, y a continuación hacer clic sobre otra vista en miniatura 212 permitiría la selección de la totalidad de las vistas en miniatura 212 sobre las que se hizo clic, así como todas las vistas en miniatura 212 participantes. Una segunda acción de operador, tal como presionar la tecla "S" (para una escisión), hacer clic sobre el botón medio, desplegar hacia abajo un menú y seleccionar "división" o hacer clic derecho y a continuación elegir "división" a partir de un menú secundario, u otra acción de operador, daría lugar a que los rostros seleccionados se escindan de la persona actual. Preferiblemente, los rostros escindidos se asignarían como procedentes de otra persona, reteniendo su asociación entre sí, tal como se describirá a continuación.

El proceso de la escisión 200 puede realizarse sobre el mismo conjunto de rostros en una variedad de diferentes órdenes. En el ejemplo de la figura 2A, pueden escindirse los rostros A1-A8, y a continuación los rostros B1-B2, y a continuación el rostro C1, dejando solo el rostro D1 como parte del conjunto actual. Como alternativa, pueden escindirse los rostros B1-B2, y a continuación el rostro C1 y a continuación el rostro D1, dejando los ocho rostros A1-A8 como procedentes del mismo conjunto. Como alternativa, los rostros B1-B2, C1 y D1 pueden escindirse al mismo tiempo, creando un nuevo conjunto con los rostros que proceden de las tres personas ("B", "C", y "D"), los cuales se escindirían entonces uno de otro en una etapa posterior.

El orden de los rostros que se presentan al operador, y así mismo su manera de presentación, pueden facilitar la etapa de la escisión 200. Por ejemplo, ha de apreciarse que las etapas de la indexación pueden realizarse múltiples veces sobre la misma colección. Por ejemplo, en el caso de generación de imágenes de acontecimientos en embarcaciones de crucero, la indexación se realizará cada día del crucero, a medida que se añaden nuevas imágenes a la colección. Los rostros que están asociados uno con otro a través de la indexación automatizada asistida de forma manual de la presente invención pueden pasar a través de la etapa de escisión 200 a medida que se asocian nuevos rostros con una persona. Si los rostros que se habían indexado entre sí con anterioridad se presentan de forma consecutiva en la pantalla 210, esto acelerará el proceso de seleccionar nuevos rostros que se han asociado falsamente con la persona actual, de tal modo que los rostros asociados por falsos positivos pueden dividirse con la mayor facilidad. Además, el marcado de los rostros que se habían asociado con anterioridad de

forma manual uno con otro, por ejemplo con un límite o punto de color, notifica al operador qué rostros no deberían escindirse uno de otro. Ha de apreciarse que, por lo demás, se puede marcar aquellos rostros que no se han asociado con anterioridad, con el fin de centrar la atención del operador en aquellos rostros.

5 Otra ordenación preferida de rostros es clasificar los rostros de acuerdo con la similitud de los rostros según se determine por algunos medios automatizados. Por ejemplo, las puntuaciones entre los rostros pueden usarse por un algoritmo de agrupamiento (por ejemplo, agrupamiento jerárquico o de K-medias), y disponerse entonces de tal modo que los rostros que están más lo próximamente relacionados se encontrarán cerca uno de otro. Esta disposición es de un valor particular cuando un gran número de rostros se indexan a la vez, sin una indexación
10 previa, y se presentan al operador. Cuando se comparan y se seleccionan los rostros individuales a escindir, encontrarse cerca en la pantalla 210 facilita las operaciones.

Las vistas en miniatura 212 eliminan parte del contexto de las imágenes, tal como las otras personas en las imágenes, lo que puede ser útil en la determinación de si los rostros en dos vistas en miniatura diferentes 212 están
15 asociados uno con otro. Para facilitar la indexación manual, es conveniente permitir la recuperación de la imagen completa mediante alguna acción de operador, tal como hacer doble clic sobre una vista en miniatura 212, o haciendo clic derecho seguido por la selección a partir de un menú secundario. De hecho, es lo más preferible que el operador sea capaz de seleccionar dos vistas en miniatura diferentes 212 a partir de las cuales se recuperan las imágenes completas 214, de tal modo que el contexto a partir de ambas de las imágenes (por ejemplo, las otras
20 personas en las imágenes) puede compararse de la mejor forma.

La figura 2B es un diagrama esquemático de una pantalla de ordenador 210 que puede usarse por un operador para realizar la escisión 200, al igual que en la figura 2A, que además comprende unas imágenes fuente completas 214 a partir de las cuales se obtuvo un subconjunto de las vistas en miniatura de rostros 212. En el presente caso, la
25 pantalla se escinde en vertical para dar una región izquierda y una derecha, en el que la región izquierda comprende una disposición ordenada de vistas en miniatura 212 similar a la que se muestra en la figura 2A. La región derecha comprende un área para dos imágenes completas 214, a partir de la cual se obtuvieron los rostros en la región izquierda (en el presente caso A5 y D1). Esta región derecha permite con más facilidad que el operador determine si los rostros A5 y D1 (indicados mediante elipses) proceden de la misma persona. Es preferible que aquellas vistas en
30 miniatura 212 que se corresponden con las imágenes fuente 214 se marquen de una forma tal que la correspondencia sea evidente para el operador. Por ejemplo, un punto rojo puede colocarse en la vista en miniatura 212 que se corresponde con la imagen fuente superior, que también tiene un punto rojo, mientras que un punto azul puede colocarse en la vista en miniatura que se corresponde con la imagen fuente superior, que también tiene un punto azul, permitiendo que el operador establezca con facilidad la coincidencia de la vista en miniatura 212 con su
35 imagen fuente 214.

Ha de observarse que, con la escisión de uno o más rostros que proceden del conjunto de personas actual, puede considerarse que ese rostro o bien es ahora un nuevo conjunto de personas compuesto por los rostros escindidos o bien, como alternativa, los rostros pueden asociarse ahora con otro conjunto. Por ejemplo, en la indexación
40 automatizada, un rostro (o grupo de rostros) puede tener una alta similitud tanto con un conjunto primario como con un conjunto secundario, con una similitud más alta de acuerdo con el conjunto primario. El sistema automatizado puede asignar los rostros a un conjunto, pero si se determinara por indexación manual que esta asociación se ha realizado con error, la similitud con el conjunto secundario puede ser lo bastante alta de tal modo que, durante el proceso de escisión manual 200, se dará al operador el conjunto secundario para su revisión. Este conjunto
45 secundario se asocia ahora con los rostros que se escinden del conjunto primario, y el operador puede determinar si la asociación se realizó correctamente.

A la conclusión de la presente etapa de la escisión 200, no debería haber asociaciones de falsos positivos de los rostros dentro de conjunto alguno. No obstante, pueden existir asociaciones de falsos negativos, en las que los rostros que proceden de la misma persona no están asociados uno con otro, sino con conjuntos diferentes.

Fusión 300

La figura 3 es un diagrama esquemático de una pantalla de ordenador 310 que comprende unas vistas en miniatura de rostros 212 que pueden usarse por un operador para realizar la fusión 300. En la fusión 300, la pantalla 310 se divide en regiones mediante unas barras verticales 216. Una región de personas actuales 320 comprende un número de vistas en miniatura 212 de la persona actual. La persona actual se corresponde con el conjunto para el cual el operador está buscando otras personas, que comprende rostros que se han asignado de forma incorrecta como no representativos de la persona actual (es decir, esta es una asociación de falsos negativos).

Una región de personas similares 340 comprende un número de vistas en miniatura 212 de los rostros que son representativos de los conjuntos que son potencialmente representativos de la persona actual. En la figura 3, existe una única vista en miniatura 212 para cada conjunto de personas similares, a pesar de que puede ser conveniente también tener múltiples vistas en miniatura 212 a partir de cada conjunto de una persona similar. Por ejemplo, la
65 región 340 puede comprender, como alternativa, unas filas (o columnas) de unas vistas en miniatura 212 en las que cada fila (o columna) comprendería unas vistas en miniatura 212 a partir del mismo conjunto, permitiendo que el

operador explore muchos rostros tanto en el conjunto de personas actual como en los conjuntos de personas similares.

5 El operador puede seleccionar una de las personas a partir de las personas similares seleccionando la vista en miniatura correspondiente, que se indica entonces mediante un límite grueso. Esta persona se considera entonces la persona candidata. En la figura 3, la persona candidata es "C1", lo que se indica mediante el límite grueso.

10 La región de personas candidatas 330 comprende un número de vistas en miniatura de los rostros que proceden de la persona candidata seleccionada en la región de personas similares. Debido a que la persona seleccionada es C1, las vistas en miniatura presentadas en la región de personas candidatas 330 son unos rostros tomados del conjunto de esa persona y, en el presente caso, se designan C1 a C4 (y pueden continuar a través de desplazamiento a un número mayor de rostros).

15 La visibilidad concurrente de múltiples rostros que proceden del conjunto de personas actual (A1 a A5) y el conjunto de personas candidato (C1 a C4) permite que el operador compare de forma muy eficiente la persona actual con la persona candidata para decidir si estas deberían fusionarse en la persona actual. Con el fin de asistir adicionalmente al operador, las imágenes a partir de las cuales se obtuvieron ciertas vistas en miniatura 212 pueden presentarse en una región de imágenes fuente 350 hacia el extremo derecho. En el presente caso, el operador ha seleccionado la vista en miniatura de persona actual 212 A5 y la vista en miniatura de persona candidata 212 C2, y las imágenes fuente 214 para estas vistas en miniatura 212 se muestran en la región de imágenes fuente 350.

25 Es conveniente ordenar los conjuntos en la región de personas similares de acuerdo con la similitud de cada conjunto de personas similares con el conjunto de personas actual, de tal modo que las personas similares que coinciden con la mayor probabilidad con las personas actuales deberían encontrarse en la fila de arriba de la región 340. Puede ser conveniente también ordenar de forma similar las vistas en miniatura individuales 212 en la región de personas candidatas 330 de tal modo que los rostros que proceden del conjunto candidato que son los más similares a los rostros en el conjunto de personas actual sean los primeros rostros que se presentan en la región 330.

30 Es preferible que se haga que cada una de las regiones 320, 330 y 340 se desplace en vertical, de tal modo que si hay más vistas en miniatura 212 de las que pueden caber de una sola vez en la pantalla, el operador puede realizar un desplazamiento hacia abajo para ver más de las vistas en miniatura 212.

35 Ha de observarse que deberían hacerse eliminado las asociaciones de falsos positivos tanto de la persona actual como de los conjuntos de personas similares, por medio de la etapa de la escisión 200 antes de la presente etapa de fusión 200. A pesar de que es posible indexar rostros mediante, en primer lugar, la fusión 300 y a continuación la escisión 200, esto puede ser confuso para el operador. Considérese que el operador encuentra un conjunto de personas actual y un conjunto de personas similares en la pantalla 310 de la fusión 300, en la que ambos de los conjuntos de personas comprenden falsos positivos, de los cuales hay algunos rostros en ambos conjuntos representativos de la misma persona, y otros que no son representativos (por ejemplo, el conjunto de personas actual está compuesto por los rostros A1, A2, A3 y B1, mientras que el conjunto de personas similares comprende los rostros A4, A5, A6 y C1). No quedaría claro para el operador si los dos conjuntos deberían fusionarse, debido a que esto introduciría también nuevas asociaciones de falsos positivos con el conjunto fusionado (es decir, los rostros B1 y C1 con el rostro A).

45 **División y fusión de múltiples conjuntos de forma simultánea**

50 Los procedimientos anteriores están dispuestos óptimamente para la escisión y la fusión de conjuntos actuales individuales. Dada una gran colección, tomada especialmente para la generación de imágenes de acontecimientos - tal como para un crucero o un parque temático - el número de conjuntos puede ser muy grande. Además, puede ser necesario indexar la colección múltiples veces de una forma incremental a medida que se recogen nuevas imágenes. La carga sobre el operador puede ser muy grande, debido a que este ha de inspeccionar visualmente grandes números de conjuntos para determinar si hay o bien errores por falsos positivos o bien por falsos negativos.

55 Con el fin de dar cabida a estos ejemplos, el operador puede inspeccionar muchos conjuntos de una sola vez. Esto se ilustra en la figura 4, que es un diagrama esquemático de una pantalla de ordenador 410 que comprende unas filas de conjuntos de rostros que pueden usarse por un operador para realizar con rapidez una escisión o una fusión sobre un número de conjuntos de una sola vez. La pantalla 410 se escinde mediante un número de separadores horizontales 218, en la que cada fila comprende unas vistas en miniatura de rostros 212 o bien a partir del mismo conjunto (al igual que en la pantalla de escisión 210) o bien, como alternativa, los rostros que proceden de dos conjuntos que se considera por medios automatizados (por ejemplo, puntuaciones de reconocimiento facial) que tienen una alta similitud (al igual que en la pantalla de fusión 310).

65 Las vistas en miniatura de rostros para cada fila se dividen en dos grupos, indicado cada grupo por alguna marca visual que se distinga con facilidad. En el ejemplo dado en la figura 4, las vistas en miniatura 212 a la izquierda se distinguen por la letra N en un círculo que está relleno con un color brillante. Existen muchas otras marcas que

pueden usarse de forma conveniente, incluyendo diferentes límites de color, separando las vistas en miniatura a la derecha y a la izquierda 212 con una línea, u otros medios. Las vistas en miniatura a la derecha provendrán, en general, de un único conjunto que ha pasado con anterioridad a través de la etapa de escisión 200 y, por lo tanto, no contiene asociaciones de falsos positivos.

5 Las vistas en miniatura a la izquierda comprenden unos rostros supuestos para su asociación. Estos rostros supuestos pueden proceder de un conjunto que tiene una alta similitud con el conjunto a la derecha, y presentarse para su fusión. Como alternativa, si existen imágenes y rostros que se están añadiendo de forma incremental a una colección indexada completamente y con precisión, las vistas en miniatura 212 a la izquierda pueden comprender
10 “nuevos” rostros que se han añadido de forma automática al conjunto a través de la etapa de indexación automatizada 100. La designación con la letra “N” hace referencia al hecho de que estos rostros son “nuevos” para la colección.

El operador puede pasar a continuación a través de las filas y asignar cada fila a una de cuatro acciones diferentes:

- 15 1) Permitir que el conjunto permanezca “tal como está” - en el presente caso, se determina que la totalidad de los rostros a la izquierda coincide con los rostros a la derecha y, por lo tanto, el conjunto “tal como está” no contiene asociaciones de falsos positivos y debería mantenerse con su composición actual.
- 20 2) Realizar una “división” - en el presente caso, se determina que ninguno de los rostros a la izquierda coincide con los rostros a la derecha, pero que la totalidad de los mismos sí coincide entre sí. En el presente caso, los rostros a la izquierda se escinden en su propio conjunto.
- 25 3) Realizar un “despiece” - en el presente caso, se determina que ninguno de los rostros a la izquierda coincide con los rostros a la derecha y, además, que estos no están asociados uno con otro. En el presente caso, los rostros a la izquierda se escinden en un número de conjuntos individuales, comprendiendo cada uno un único rostro.
- 30 4) Reservar para una inspección manual más detallada - en el presente caso, o bien no es posible determinar si los rostros a la izquierda coinciden con los rostros a la derecha, o bien existe una relación compleja que no permite o bien una escisión o bien un despiece (por ejemplo, hay tres rostros a la izquierda, con dos que coinciden entre sí y no con el tercer rostro).

Estos conjuntos se presentan al usuario, en general, en una pantalla para su escisión al igual que en la figura 2A o la figura 2B.

Con una pantalla tal como esta, el operador puede examinar múltiples conjuntos de una sola vez.

Adición incremental de rostros

Tal como se indica anteriormente, en muchos casos las imágenes se añadirán de forma incremental a lo largo de un periodo de tiempo. En estos casos, el operador indexará la colección de acuerdo con la presente invención, dando
40 como resultado ninguna asociación de falsos positivos o de falsos negativos. Entonces, se añadirán a la colección nuevas imágenes con nuevos rostros. En tales casos, existen medios para garantizar que la máxima cantidad de la información anterior se usa en el proceso de indexación para minimizar el esfuerzo subsiguiente y para proporcionar la máxima precisión.

45 En primer lugar, en la visualización escindida 210 (y también en la pantalla 410) es útil indicar, en general, aquellos rostros que se ha determinado con anterioridad que proceden de la misma persona. Es decir, si se está mirando a un conjunto de 10 rostros, de las cuales 8 se habían indexado correctamente con anterioridad, y 2 son “nuevos” rostros, indicando cuales son los nuevos rostros y cuales son los indexados con anterioridad, el operador sabe que no tiene que examinar los rostros indexados con anterioridad, debido a que puede suponerse que estos con
50 correctos. El etiquetado puede ser el de los nuevos rostros (por ejemplo, al igual que con la “N” en el círculo de la figura 4), el de los rostros indexados con anterioridad, o el de ambos (es decir, unas marcas mutuamente distinguibles).

Adicionalmente, la etapa de indexación automatizada 100 debería tener la información a partir de la indexación
55 previa con el fin de retener la identidad de los conjuntos establecidos con anterioridad. Es decir, ningunos dos conjuntos deberían fusionarse entre sí a partir de una indexación previa (debido a que todas las asociaciones de falsos negativos se habían tratado con anterioridad), y ningún conjunto a partir de una indexación previa debería escindirse entre, o en, dos conjuntos (debido a que todas las asociaciones de falsos positivos se habían tratado con anterioridad). Es decir, cada conjunto en una indexación previa de acuerdo con la presente invención debería ser un subconjunto de un conjunto en la siguiente etapa de la indexación automatizada. Una forma de garantizar esto es
60 una disposición tal que la indexación automatizada con las imágenes incrementales simplemente añade rostros dentro de las imágenes incrementales a los conjuntos anteriores, excepto en los casos en los que las coincidencias con rostros en rostros previos no pueden establecerse, caso en el que se forman nuevos conjuntos.

65 Además, solo deberían darse al operador decisiones a realizar en relación con los nuevos rostros, y no debería presentarse al operador conjunto alguno que no tenga nuevos rostros.

5 Debería ser evidente para un experto en la técnica que las realizaciones que se mencionan anteriormente son meramente ilustraciones de unas pocas de las muchas realizaciones específicas posibles de la presente invención. También debería apreciarse que los procedimientos de la presente invención proporcionan un número casi incontable de disposiciones de indicadores, etiquetas, detectores, medios de mezclado, medios de aplicación de fuerza y más.

10 Además, se pretende que todas las declaraciones en el presente documento que enumeren principios, aspectos y realizaciones de la presente invención, así como ejemplos específicos de los mismos, engloben equivalentes tanto estructurales como funcionales de la misma. Adicionalmente, se pretende que tales equivalentes incluyan tanto los equivalentes conocidos en la actualidad como los equivalentes que se desarrollen en el futuro, es decir, cualquier elemento desarrollado que realice la misma función, con independencia de su estructura.

15 Ha de apreciarse adicionalmente que las definiciones en la sección "Terminología" no son limitantes, sino que más bien reflejan una breve descripción de las expresiones para el beneficio del lector. Típicamente, tales expresiones deberían leerse con referencia a la técnica anterior y al uso convencional de las expresiones. Unos significados más amplios de las expresiones se proporcionan o bien directa o bien indirectamente en el análisis más extenso que sigue a esa sección.

20 Ha de observarse que, a pesar de que los medios de identificación que se describen anteriormente comprenden el reconocimiento facial, también es posible usar otros medios para la determinación de las identidades de los patrones en las imágenes. Así mismo, ha de observarse que la ubicación, los tamaños, las orientaciones (por ejemplo, los límites horizontales entre capas frente a los vertical) y las formas de las regiones de las pantallas 210 y 310 (por ejemplo, las regiones de imagen completa, o las regiones de personas actuales) pueden variarse sin cambiar las operaciones fundamentales de la presente invención. Además, debido a que todos los motores de reconocimiento facial automatizado tienen un cierto número de errores (o bien errores por falsos positivos y/o bien por falsos negativos), los procedimientos de la presente invención servirán para permitir la reducción manual de esos errores. De forma similar, los algoritmos para asociar rostros uno con otro en personas tendrán errores, que pueden reducirse de forma similar.

30 En la memoria descriptiva de la presente invención, se pretende que cualquier elemento expresado como unos medios para realizar una función especificada englobe cualquier forma de realización de esa función. La invención, tal como se define por tal memoria descriptiva, se basa en el hecho de que las funcionalidades que se proporcionan mediante los diversos medios enumerados se combinan y se reúnen de la forma que requiere la memoria descriptiva. El solicitante de la presente invención considera, por lo tanto, cualquier medio que pueda proporcionar aquellas funcionalidades como equivalente a los que se muestran en el presente documento.

35

REIVINDICACIONES

1. Un procedimiento para indexar imágenes faciales en una colección de imágenes de acuerdo con las personas que están representadas en la colección, que comprende:
- 5
- indexar de forma automática las imágenes faciales para dar una multiplicidad de conjuntos de imágenes usando reconocimiento de rostros, determinándose que cada conjunto de la cual comprende unas imágenes faciales representativas de la misma persona;
 - presentar los conjuntos de imágenes a un operador;
- 10
- recibir una acción procedente del operador para un conjunto de imágenes para crear un conjunto resultante de imágenes, en el que la acción se elige del grupo que consiste en: i) permanecer tal como está, ii) escindir, iii) despiezar, y iv) reservar para una inspección manual;
- 15
- en el que las imágenes que se escinden forman un nuevo conjunto de imágenes y las imágenes escindidas se asignan como procedentes de otra persona reteniendo su asociación entre sí, y
en el que las imágenes de un conjunto despiezado se escinden en conjuntos individuales, comprendiendo cada uno una única imagen,
- 20
- presentar al operador un primer conjunto resultante de imágenes con otro conjunto resultante de imágenes en función de la similitud, en el que el operador determina de forma manual si el primer conjunto de imágenes y el segundo conjunto de imágenes comprenden unas imágenes faciales que representan la misma persona; y
 - fusionar los conjuntos de imágenes resultantes si el operador determina que las imágenes representan la misma persona.
- 25
2. El procedimiento de acuerdo con la reivindicación 1, en el que las imágenes que se despiezan se asocian con otro conjunto de imágenes en función de la similitud a partir de una indexación automatizada.

Fig. 1

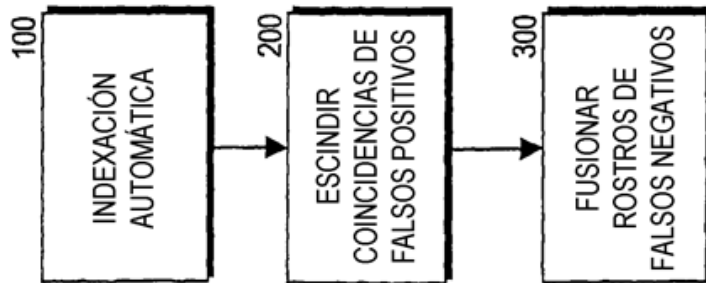
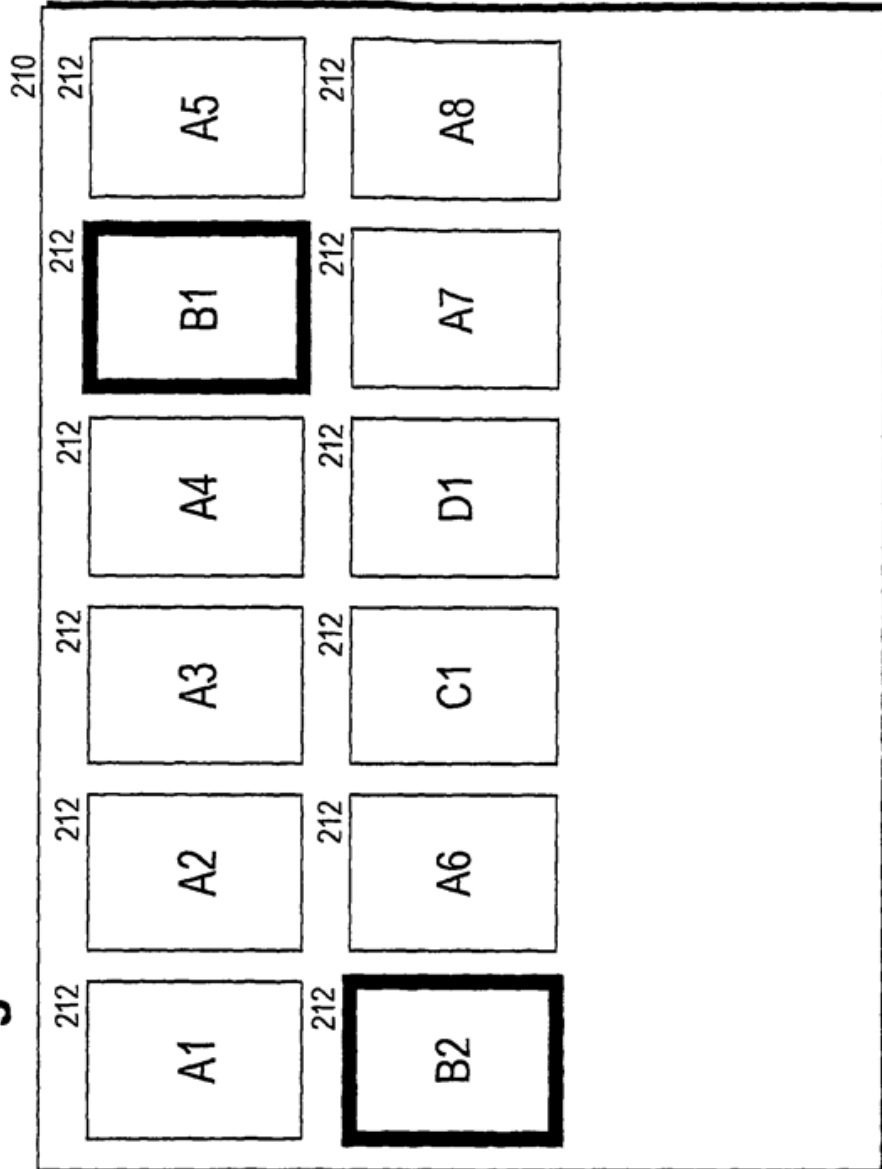


Fig. 2A



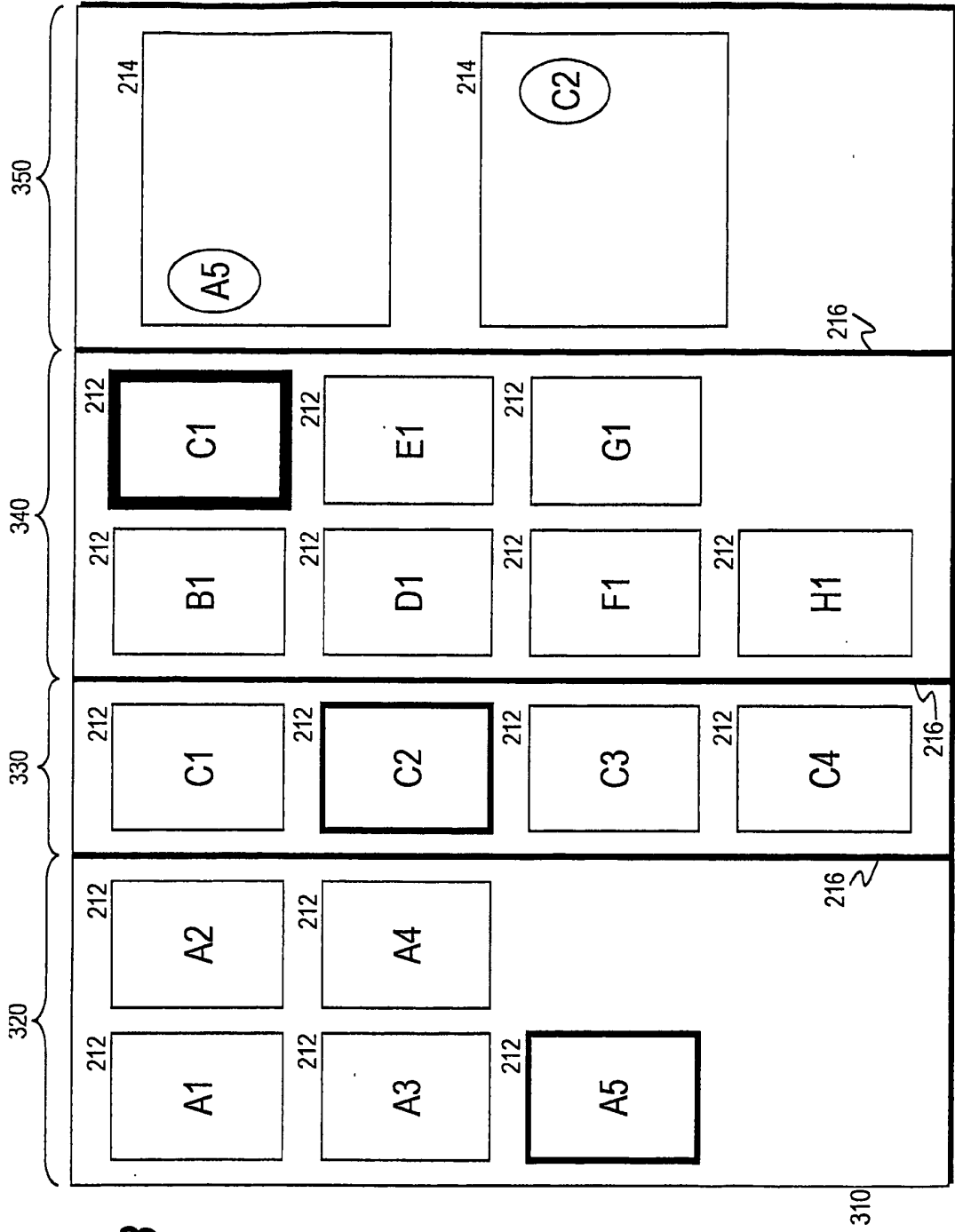


Fig. 3

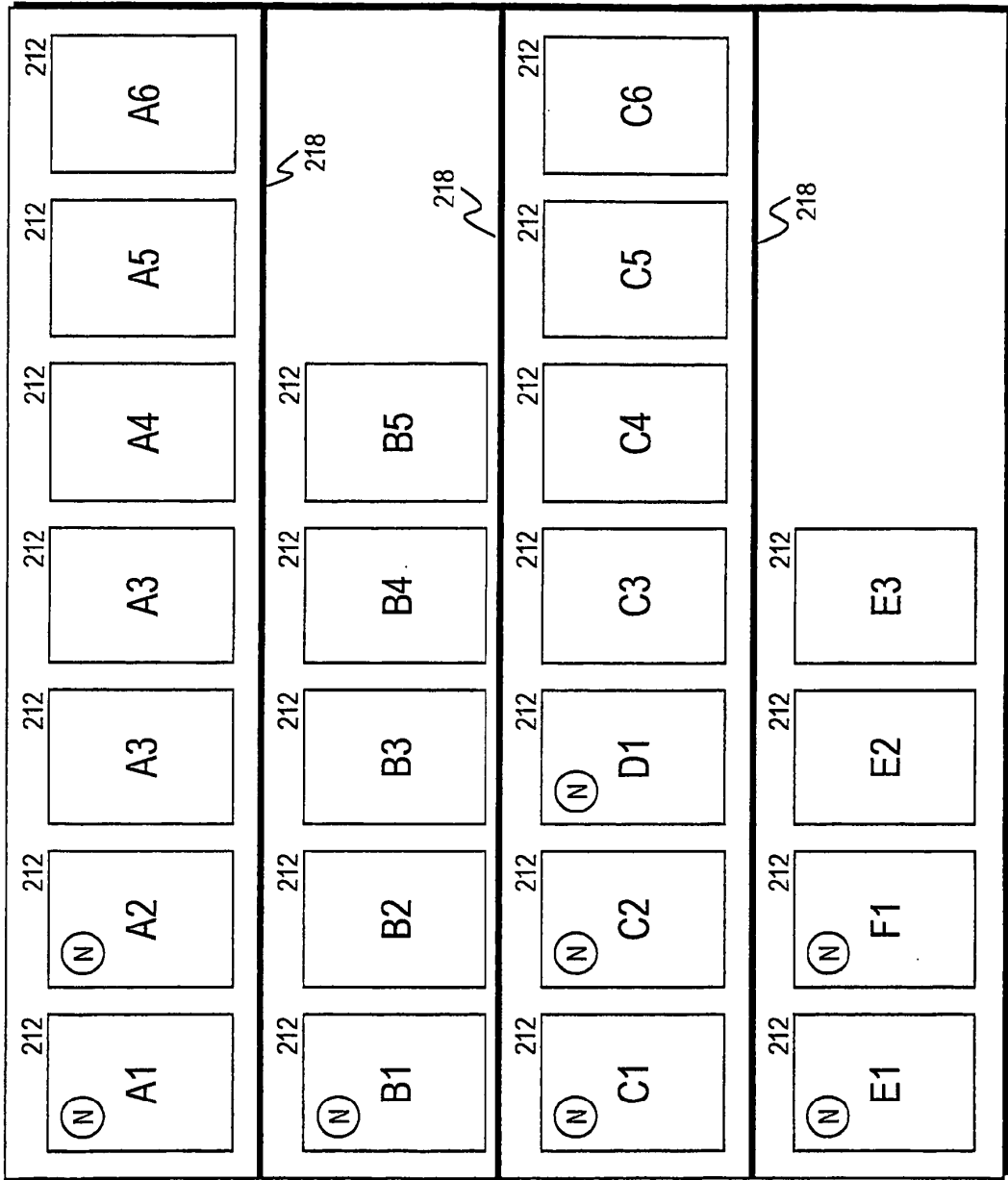


Fig. 4