

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 725 787**

51 Int. Cl.:

**G06F 17/30** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **12.06.2009 PCT/FR2009/000702**

87 Fecha y número de publicación internacional: **23.12.2009 WO09153445**

96 Fecha de presentación y número de la solicitud europea: **12.06.2009 E 09766007 (0)**

97 Fecha y número de publicación de la concesión europea: **27.02.2019 EP 2289009**

54 Título: **Dispositivo de ayuda para el reconocimiento de imágenes mejorado**

30 Prioridad:

**16.06.2008 FR 0803345**

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

**27.09.2019**

73 Titular/es:

**INRIA - INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE (100.0%)  
Domaine de Voluceau, Rocquencourt  
78150 Le Chesnay, FR**

72 Inventor/es:

**JEGOU, HERVÉ;  
SCHMIDT, CORDELIA y  
DOUZE, MATTHIJS**

74 Agente/Representante:

**ELZABURU, S.L.P**

ES 2 725 787 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

Dispositivo de ayuda para el reconocimiento de imágenes mejorado

La invención trata de un dispositivo de ayuda para el reconocimiento de imágenes.

5 Una primera función de tales dispositivos es generalmente buscar imágenes similares en una base de datos, ya sea dos a dos, por ejemplo para suprimir duplicados en esta base, ya sea a partir de una imagen de consulta, por ejemplo para buscar imágenes que representan el mismo tema que la imagen de consulta.

10 Por "imágenes similares", en este punto se entiende imágenes que representan el mismo objeto o la misma escena, en condiciones de toma de imagen potencialmente diferentes. Esta definición abarca especialmente las imágenes modificadas de manera sintética, por ejemplo mediante una operación de compresión o un ataque de filtrado malintencionado.

Cabe contemplar múltiples aplicaciones tales como, por ejemplo, la identificación de objetos robados en sitios de subastas en línea, la selección de lotes de fotografías, la identificación de falsificaciones de modelos o imágenes.

Una segunda función de estos dispositivos es evaluar la semejanza de dos imágenes similares, en particular para ordenar las imágenes resultado según su pertinencia.

15 Existen múltiples tecnologías.

La mayoría de estas tecnologías, al menos las más recientes, se basan en la utilización de descriptores locales, que caracterizan zonas de más particular interés de una imagen.

20 Se puede encontrar información más extensa sobre la detección de estas zonas de interés y sobre la generación de descriptores referidos a ellas en el documento [1]: Lowe, D, "*Distinctive image features from scale-invariant keypoints*", IJCV, 60, (2004) 91-110. La comparación y/o la búsqueda de imágenes equivale entonces a comparar descriptores locales entre sí, lo cual es más fino que una comparación directa de ficheros informáticos entre sí o que una comparación de las imágenes basándose en descriptores globales.

Con el tiempo, las bases de datos memorizan cantidades cada vez mayores de imágenes. Su utilización también se ha desarrollado ampliamente, en particular merced a Internet.

25 En otras palabras, la búsqueda de imágenes similares lleva consigo operaciones de comparación incesantemente más abundantes. Y, en la práctica, esta abundancia es tan cuantiosa que hace imposible la puesta en práctica de dispositivos basados en comparaciones directas de descriptores entre sí.

30 El documento [5]: Sivic, J. and Zisserman, A., "*Video Google: A Text Retrieval Approach to Object Matching in Videos*", in ICCV (2003), propone realizar una correspondencia entre los descriptores y un índice. A cada descriptor se le asocia un valor entero, seleccionado de entre un conjunto finito de valores enteros. Entonces, comparar imágenes entre sí equivale a comparar conjuntos de valores enteros entre sí, lo cual requiere pocos recursos de cálculo y acelera en su tanto la búsqueda.

35 Sin embargo, los dispositivos basados en esta última técnica, llamada del "bag of features", no son plenamente satisfactorios. El número de imágenes que se estiman similares es, muchas veces, demasiado grande, siendo consideradas por el dispositivo imágenes muy diferentes como similares, y/o no es posible distinguir las imágenes más pertinentes.

Si se aumenta el número de valores del índice, se obtiene el efecto contrario, es decir, los dispositivos ya no devuelven prácticamente ninguna imagen.

40 Para paliar esta dificultad, el documento [12]: "*Object retrieval with large vocabularies and fast spatial Matching*", J. Philbin, O. Schum, M. Isard, J. Sivic y A. Zisserman, CPVR'2007, propone completar la técnica del "bag of features" volviendo a evaluar los resultados obtenidos merced a esta última técnica integrando información espacial sobre la posición de los puntos.

45 El artículo "Content based copy retrieval using distortion-based probabilistic similarity search" de Alexis Joly et al. describe una técnica para descubrir copias de un contenido de vídeo en una base de datos utilizando características locales.

50 Por ejemplo, [1] realiza una transformación llamada de "Hough" para determinar cada vez los parámetros de una transformación afín que transforma la imagen de consulta en una de las imágenes resultado. A cada una de las imágenes se le atribuye una nota, en función del número de descriptores que verifican la respectiva transformación afín. El cálculo de una transformación de Hough precisa de recursos tales que esta transformación no puede ser llevada a la práctica para importantes lotes de imágenes.

Sin embargo, por causa de su elevadísimo coste computacional, la fase de reclasificación tan sólo se aplica a un número limitado de imágenes. Así, para bases de imágenes muy voluminosas, se saltan imágenes similares, ya que la técnica del bag-of-features no las ha clasificado suficientemente bien.

El Solicitante propone un dispositivo que viene a mejorar la situación.

5 Este dispositivo puede comprender:

- una memoria, capaz de almacenar una pluralidad de juegos de descriptores ( $d_{ij}$ ), correspondiendo cada juego a una imagen,

- asociándose cada descriptor ( $d_{ij}$ ) a una zona designada de la imagen, y comprendiendo unos primeros datos de tipo vector ( $x_{ij}$ ), unos segundos datos de tipo ángulo ( $a_{ij}$ ) y unos terceros datos de tipo escalar ( $s_{ij}$ ),

10 - un primer operador, capaz de recibir una designación de dos descriptores ( $d_{ij}$ ) y de establecer un booleano que refleja la verificación de un criterio de similitud de los vectores de descriptores ( $x_{ij}$ ), a partir de una comparación entre los primeros datos de estos descriptores,

- un segundo operador, capaz de recibir una designación de dos descriptores ( $d_{ij}$ ) y de establecer un parámetro de ángulo de giro ( $\Delta A$ ) a partir de los segundos datos ( $a_{ij}$ ) respectivos de los dos descriptores ( $d_{ij}$ ),

15 - un tercer operador, capaz de recibir una designación de dos descriptores ( $d_{ij}$ ) y de establecer un parámetro de factor de escala ( $\Delta S$ ) a partir de los terceros datos de estos descriptores ( $s_{ij}$ ),

- un controlador, capaz de recibir las designaciones de un primer y de un segundo juego de descriptores ( $d_{ij}$ ), y establecido para aplicar el primer operador a cada pareja formada a partir de un descriptor ( $d_{ij}$ ) del primer juego y de un descriptor del segundo juego ( $d_{ij}$ ), para inferir una lista de parejas de descriptores similares,

20 - llamar al segundo operador y al tercer operador por cada una de las parejas de descriptores de dicha lista y almacenar en memoria el conjunto de los parámetros obtenidos, y

- un estimador de la semejanza entre las imágenes respectivamente correspondientes al primer juego y al segundo juego de descriptores ( $d_{ij}$ ), a partir de un procesamiento estadístico sobre dicho conjunto de parámetros.

25 Merced a esta configuración, el dispositivo según la invención faculta una comparación más fina y, por tanto, más pertinente, de imágenes entre sí, o de una imagen con una pluralidad de imágenes, al propio tiempo que conserva tiempos de procesamiento cercanos, e incluso inferiores, a los dispositivos del estado de la técnica.

Otras características y ventajas de la invención se irán poniendo de manifiesto conforme se examine la descripción detallada que sigue, y los dibujos que se acompañan, en los cuales:

30 la figura 1 representa esquemáticamente un dispositivo de ayuda para el reconocimiento de imágenes según la invención,

la figura 2 es un diagrama de flujo que ilustra una realización de un dispositivo de firma para el dispositivo de la figura 1,

la figura 3 es un diagrama de flujo que ilustra una parte del dispositivo de firma de la figura 2,

35 la figura 4 es un diagrama de flujo que ilustra una primera realización de un primer operador para el dispositivo de la figura 1,

la figura 5 es análoga a la figura 4, para una variante de realización,

la figura 6 es un diagrama de flujo que ilustra un modo de realización de un segundo operador para el dispositivo de la figura 1,

40 la figura 7 es un diagrama de flujo que ilustra un modo de realización de un tercer operador para el dispositivo de la figura 1,

la figura 8 es un diagrama de flujo que ilustra un modo de realización de un interrogador para el dispositivo de la figura 1,

la figura 9 es análoga a la figura 8 para un segundo modo de realización,

45 la figura 10 es un esquema que ilustra una estructura de listas invertidas para una segunda estructura de almacenamiento para el dispositivo de la figura 1,

la figura 11 es un gráfico que representa una eficiencia del primer operador en función del valor de un parámetro de umbral, y

la figura 12 es un gráfico que representa las prestaciones del primer operador en función de un parámetro de dimensión.

Los dibujos que se acompañan podrán servir no sólo para completar la invención, sino también, en su caso, contribuir a su definición.

5 La figura 1 ilustra un dispositivo de ayuda para el reconocimiento de imágenes según la invención.

El dispositivo 1 comprende memoria 3 y un dispositivo de procesamiento 5, capaz de interacción con esta memoria 3.

10 La memoria 3 comprende una primera estructura de almacenamiento 5 para memorizar, al menos temporalmente, uno o varios juegos de descriptores de imágenes. La primera estructura de almacenamiento 5 memoriza además una correspondencia entre cada uno de los juegos de descriptores y un identificador de la imagen descrita.

En la primera estructura de almacenamiento, cada uno de los descriptores  $d_{ij}$  correspondiente a una imagen  $j$  agrupa:

- unos primeros datos de tipo vector, que representan un vector característico  $x_{ij}$  de una zona de interés  $i$  de la imagen  $j$ ,
- 15 - unos segundos datos de tipo escalar, que representan un ángulo  $a_{ij}$  característico de dicha zona de interés  $i$  de la imagen  $j$ ,
- unos terceros datos de tipo escalar, que representan una escala  $s_{ij}$  característica de dicha zona de interés  $i$  de la imagen  $j$ ,
- 20 - unos cuartos datos, que representan una posición  $c_{ij}$  de un centro de esta zona de interés  $i$  de la imagen  $j$ . Estos cuartos datos también se pueden considerar como un suplemento a los primeros datos de tipo vector, en otras palabras, como primeros datos suplementarios.

Cada uno de los vectores  $x_{ij}$  presenta una dimensión  $d$  escogida, por ejemplo 128.

25 Diferentes herramientas permiten generar un juego de descriptores  $d_{ij}$  a partir de datos de imagen  $j$ . Funcionalmente, estas herramientas ponen en práctica una fase de detección de zonas de interés  $i$  dentro de una imagen  $j$  y, luego, una fase de asignación de un descriptor local  $d_{ij}$  a esta zona. Por ejemplo, [1] da a conocer un método de detección de las zonas de interés y de asignación de descriptores locales.

30 El dispositivo 1 se puede establecer para recibir los juegos de descriptores  $d_{ij}$  de una herramienta de generación. Como variante, el dispositivo 1 puede integrar una herramienta de este tipo. El dispositivo 1 se establece entonces para recibir datos de imágenes. Potestativamente, el dispositivo 1 puede comprender una estructura de almacenamiento suplementaria para memorizar de manera duradera estos datos de imágenes, por ejemplo en forma de una base de datos.

La memoria 3 comprende además una segunda estructura de almacenamiento 7, que memoriza una representación de uno o varios juegos de descriptores de imágenes en una forma particular, diferente de la primera estructura de almacenamiento 5.

35 Por cada uno de los descriptores  $d_{ij}$ , la segunda estructura de almacenamiento 7 memoriza una relación entre este descriptor  $d_{ij}$  y un valor entero  $q(x_{ij})$  seleccionado de entre un conjunto de valores enteros común al conjunto de los descriptores  $d_{ij}$ , denominado índice. Este valor entero  $q(x_{ij})$  en ocasiones se designa por "palabra visual". Este valor entero  $q(x_{ij})$  se establece en función de los datos de vector característico  $x_{ij}$ .

40 El número de índice posible depende de la aplicación que para el dispositivo 1 se contemple. El número de índice puede estar comprendido generalmente entre 1000 y 1 000 000.

Ventajosamente, la segunda estructura de almacenamiento 7 memoriza, por cada uno de los descriptores  $d_{ij}$ , una correspondencia entre un identificador de este descriptor  $d_{ij}$  y su valor entero  $q(x_{ij})$ . Esto permite reducir el tamaño de la segunda estructura de almacenamiento 7.

45 Como variante, la segunda estructura de almacenamiento 7 memoriza cada vez una correspondencia entre un identificador de la imagen  $j$  correspondiente al descriptor  $d_{ij}$  y el valor entero  $q(x_{ij})$  en cuestión. Preferentemente, el identificador de imagen  $j$  es memorizado en sustitución del identificador del descriptor  $d_{ij}$  para reducir todavía más el tamaño de la segunda estructura de almacenamiento 7.

De manera particularmente ventajosa, la segunda estructura de almacenamiento 7 adopta la forma de una estructura de listas invertidas.

50 La segunda estructura de almacenamiento 7 se materializa entonces en forma de un conjunto de listas. A cada uno

- de los valores enteros del índice se le asocia una lista, la cual incluye un identificador de cada uno de los descriptores  $d_{ij}$  al que está asociado este valor entero. Esta estructura de almacenamiento permite memorizar una representación del conjunto de los juegos de descriptores  $d_{ij}$ , con el concurso de valores enteros, de manera muy compacta. Esta forma de realización de la segunda estructura de almacenamiento 5 acelera asimismo las operaciones de interrogación.
- 5
- Como variante, cada una de las listas de la segunda estructura de almacenamiento 7 puede incluir, en sustitución o como complemento del identificador de un propio descriptor  $d_{ij}$ , un identificador  $id_j$  de la imagen a la que corresponde este descriptor  $d_{ij}$  en la primera estructura de almacenamiento 5. La sustitución de un identificador de descriptor  $d_{ij}$  por un identificador de imagen  $id_j$  reduce aún más el tamaño de la segunda estructura de almacenamiento 7.
- 10
- En el caso en que se asocia un mismo valor entero  $q$  a varios descriptores  $d_{ij}$  de una misma imagen  $j$ , la lista correspondiente a este valor entero  $q$  puede incluir, bien un número de aparición en relación con el identificador de la imagen  $id_j$ , o bien varias entradas, correspondiendo cada entrada en la lista a una aparición.
- Se puede dar preferencia a esta segunda variante, ya que tan sólo aumenta muy escasamente el tamaño de la segunda estructura de almacenamiento 7, dado que es poco probable que se asigne un mismo valor entero  $q$  a descriptores diferentes  $d_{ij}$  de una misma imagen  $j$ .
- 15
- El dispositivo de procesamiento 5 comprende un cuantificador 9 capaz de asociar un valor entero  $q(x_{ij})$  a cada uno de los descriptores  $d_{ij}$  memorizados en la primera estructura de almacenamiento 5 en función de los datos de vector característico  $x_{ij}$  correspondientes.
- 20
- El cuantificador 9 se establece para poner en práctica un promedio sobre los  $k$ -vecinos más próximos, o "k-means clustering" en inglés, sobre un conjunto de datos de imágenes testigo. De ello resulta un conjunto de centroides y de celdas de Voronoi. El identificador del centroide más próximo al vector característico  $x_{ij}$  del descriptor  $d_{ij}$  procesado se asigna en cuanto valor entero  $q(x_{ij})$ . La proximidad de un vector característico  $x_{ij}$  a un centroide se puede calcular con el concurso de una distancia euclidiana.
- 25
- El cuantificador 9 es optativo. El dispositivo 1 se puede establecer para recibir el conjunto de los valores enteros  $q(x_{ij})$  que están asociados a los descriptores  $d_{ij}$  de la primera estructura de almacenamiento 5, eventualmente en sustitución de los propios juegos de descriptores  $d_{ij}$ .
- El dispositivo de procesamiento 5 comprende además un dispositivo de firma 11, establecido para recibir un descriptor  $d_{ij}$  y para establecer datos de tipo vector que representan una firma binaria multidimensional  $bi(x_{ij})$  de este descriptor  $d_{ij}$ , a partir de los datos de vector característico  $x_{ij}$ . El vector de firma binaria  $bi(x_{ij})$  presenta una dimensión  $db$  escogida, por ejemplo 64.
- 30
- Cada una de las componentes del vector de firma binaria  $bi(x_{ij})$  puede tomar exclusivamente el valor "0" o el valor "1". En otras palabras, el vector de firma binaria  $bi(x_{ij})$  se compone de booleanos.
- 35
- La segunda estructura de memoria 7 se establece para memorizar una correspondencia entre los datos de firma binaria  $bi(x_{ij})$  de cada uno de los descriptores  $d_{ij}$  memorizados en la primera estructura de almacenamiento 5 y un identificador de este descriptor  $d_{ij}$ , o de la imagen  $j$  correspondiente a este descriptor  $d_{ij}$ .
- Cuando la segunda estructura de almacenamiento 7 está realizada en forma de listas invertidas, cada una de las firmas binarias  $bi(x_{ij})$  se puede almacenar en la entrada relativa al descriptor  $d_{ij}$  correspondiente en la lista. En otras palabras, en esta realización de la segunda estructura de almacenamiento 7, cada una de las firmas binarias  $bi(x_{ij})$  se memoriza en correspondencia con el valor entero  $q(x_{ij})$  asignado al descriptor  $d_{ij}$ , y no con los datos que constituyen este descriptor  $d_{ij}$  (vector, ángulo, escala característicos). Por ejemplo, cada una de las entradas de la estructura en listas invertidas puede presentar un número fijado de bits, por ejemplo 64, para el almacenamiento de una firma binaria. Esta realización permite un almacenamiento particularmente compacto de las firmas binarias, en correspondencia con identificador  $id_j$  de imagen, o de descriptor  $d_{ij}$ .
- 40
- 45
- La figura 2 ilustra una organización del dispositivo de firma 11.
- En una etapa 200, el dispositivo de firma 11 recibe datos de vector característico  $x_{ij}$  relativos a un descriptor  $d_{ij}$  memorizado en la primera estructura de almacenamiento 5.
- En una etapa 202, se calculan unos datos de tipo vector, representativos de una proyección del vector característico  $x_{ij}$ , con el concurso de una matriz de proyección  $P$  de dimensiones  $db \times d$ .
- 50
- La misma matriz de proyección  $P$  es utilizada para el conjunto de los descriptores  $d_{ij}$  memorizados en la primera estructura de almacenamiento 5 y del cual se pretende obtener una representación en la segunda estructura de almacenamiento 7.
- Se obtienen datos representativos de un vector proyectado  $z_{ij}$ , de dimensión  $db$ . Las componentes del vector proyectado  $z_i$  se denotan por  $z_{i1}, z_{i2}, \dots, z_{idb}$ .

En una etapa 204, se asocia un valor entero  $q(x_{ij})$  al vector característico  $x_{ij}$ . El cuantificador 9 puede ser llamado para establecer el valor entero  $q(x_i)$ . El valor entero  $q(x_i)$  puede ser leído asimismo en la segunda estructura de almacenamiento 7, en caso de existir ya, en esta segunda estructura de almacenamiento 7, una representación del descriptor  $d_{ij}$  correspondiente.

5 Cuando esta segunda estructura de almacenamiento 7 está realizada en forma de listas invertidas, la correspondencia entre el valor entero  $q(x_{ij})$  y el vector característico  $x_{ij}$  puede estar ya almacenada, en forma de un identificador de lista, por ejemplo un número de índice, y de una entrada en esta lista que incluye un identificador del descriptor  $d_{ij}$ . Si no figura dicha correspondencia en la segunda estructura de almacenamiento 7, entonces se puede memorizar consecutivamente a la etapa 204.

10 En una etapa 206, se comienza un lazo de cálculo sobre las componentes  $b_{ik}$  del vector de firma  $bi(x_{ij})$ . El índice  $k$  varía de 1 a  $db$ .

En una etapa 208, se compara la  $k$ -ésima componente  $z_{ik}$  del vector proyectado  $z_i$  con un valor mediano  $Tq(x_{ij})k$ , propio del valor entero  $q(x_{ij})$  del vector característico  $x_{ij}$  y con la componente  $k$ .

15 Si la  $k$ -ésima componente  $z_{ik}$  del vector proyectado  $z_i$  es superior al valor mediano  $Tq(x_{ij})k$ , entonces la  $k$ -ésima componente  $b_{ik}(x_{ij})$  del vector de firma binaria  $bi(x_{ij})$  toma el valor "1" (uno), en la etapa 210. Si no, esta  $k$ -ésima componente  $b_{ik}(x_{ij})$  toma el valor "0" (cero), en el curso de la etapa 212.

La figura 3 ilustra la determinación de los valores medianos  $Tq(x_{ij})k$  mediante el dispositivo de firma 11.

20 En una etapa 300, el dispositivo de firma 11 recibe un conjunto  $E$  de descriptores  $e_{ij}$  relativos a datos de imágenes diferenciados de los datos de imágenes representados en la segunda estructura de memoria 7. Cada uno de los descriptores  $e_{ij}$  comprende datos representativos de un vector característico  $w_{ij}$ , de dimensión  $db$ .

En una etapa 302, se proyecta cada uno de los vectores característicos  $w_{ij}$  de los descriptores  $e_{ij}$  del conjunto  $E$  con el concurso de la matriz de proyección  $P$ . Se obtiene un conjunto  $V$  de vectores proyectados  $v_{ij}$ . Las componentes de un vector proyectado  $v_{ij}$  se denotan por  $v_{ij1}, v_{ij2}, \dots, v_{ijdb}$ .

25 En una etapa 304, se asigna un valor entero del índice a cada uno de los vectores característicos  $w_{ij}$ , por ejemplo con el concurso del cuantificador 9. La memoria 3 almacena, al menos temporalmente, una correspondencia entre cada valor entero  $q$  del índice y el conjunto de las componentes de cada uno de los vectores proyectados  $v_{ij}$  a cuyo vector característico  $w_{ij}$  se le ha asignado este valor entero  $q$ . Se puede utilizar una estructura de listas invertidas.

En una etapa 306, se comienza un lazo sobre los valores enteros  $q$  del índice y sobre las componentes  $v_{ijk}$  de los vectores proyectados  $v_{ij}$ . El índice  $k$  varía de 1 a  $db$ .

30 En una etapa 308, se calcula el valor mediano  $Tqk$ , relativo al valor entero  $q$  y a la componente  $k$ , en cuanto valor mediano de las  $k$ -ésimas componentes  $v_{ijk}$  de los vectores proyectados  $v_{ij}$  asociados al valor entero  $q$  en cuestión.

La matriz de proyección  $P$  se obtiene a partir de una matriz de valores gaussianos  $A$  generada aleatoriamente. La matriz  $A$  presenta  $db$  columnas.

35 Sobre la matriz de valores gaussianos  $A$ , se efectúa una factorización, o descomposición, de tipo  $QR$ , donde  $Q$  es una matriz ortogonal y  $R$ , una matriz triangular. Las  $d$  primeras filas de la matriz  $Q$ , que presenta  $db$  columnas, forman la matriz de proyección  $P$ .

40 El dispositivo de procesamiento también comprende un dispositivo de caracterización 10 capaz de interacción con la primera estructura de almacenamiento 5 y la segunda estructura de almacenamiento 7. El dispositivo de caracterización se establece para establecer, por cada uno de los descriptores  $d_{ij}$  memorizados en la primera estructura de almacenamiento 5, unos datos de ángulo característico cuantificado  $q_{aij}$ , a partir de los datos de ángulo característico  $a_{ij}$ , y unos datos de factor de escala cuantificados  $q_{sij}$ , logarítmicos, a partir de los datos de escala característica  $s_{ij}$ . La cuantificación de datos de un ángulo característico comprende la asignación de un valor indexado asociado a un intervalo de valores angulares dentro del cual se sitúa el ángulo característico. Igualmente, la cuantificación de datos de una escala característica comprende la asignación de un valor indexado asociado a un intervalo de valores de factor de escala dentro del cual se sitúa el logaritmo decimal de la escala característica.

45 Los valores de ángulo cuantificado  $q_{aij}$  y de factor de escala cuantificado  $q_{sij}$  de cada uno de los descriptores  $d_{ij}$  memorizados en la primera estructura de almacenamiento 5 se memorizan en la segunda estructura de almacenamiento 7, en correspondencia con un identificador del descriptor  $d_{ij}$  en cuestión, o de la imagen  $j$  para la cual se ha generado este descriptor  $d_{ij}$ .

50 Cuando esta segunda estructura de almacenamiento 7 está realizada en forma de listas invertidas, estos valores de ángulo cuantificado  $q_{aij}$  y de factor de escala cuantificado  $q_{sij}$  se almacenan ventajosamente en la entrada que representa un descriptor  $d_{ij}$  en cuestión. En otras palabras, en la lista particular designada por el valor entero  $q(x_{ij})$ , al identificador del descriptor  $d_{ij}$  o de la imagen  $j$  se le agregan dichos valores cuantificados. Cada entrada de una lista puede comprender cierto número de bits reservados para el almacenamiento de estos valores cuantificados, por

ejemplo 6 bits para el almacenamiento del valor de ángulo cuantificado  $q_{aij}$  y 5 bits para el del valor de factor de escala cuantificado  $q_{sj}$ . Se comprende que el número de bits reservados determina la extensión de los intervalos de cuantificación.

5 El dispositivo de procesamiento 5 también comprende un primer operador 13 capaz de interacción con la segunda estructura de memoria 7. El primer operador 13 se establece para establecer un valor de similitud  $f(xij, zkl)$  entre los vectores característicos  $xij$  e  $ykl$  de dos descriptores de los cuales se memoriza una representación en la segunda estructura de memoria 7.

La determinación de este valor de similitud  $f(xij, ykl)$  consiste esencialmente en verificar, en este punto, si los valores enteros  $q(xij)$  y  $q(ykl)$  de los vectores característicos  $xij$  e  $ykl$  cumplen una condición de similitud prefijada.

10 La figura 4 ilustra un primer modo de realización del primer operador 13.

En una etapa 400, el primer operador 13 recibe un primer vector característico  $xij$  y un segundo vector característico  $ykl$ , o solamente identificadores de estos vectores característicos.

15 En una etapa 402, el primer operador 13 establece el valor entero  $q(xij)$  del índice para el primer vector característico y el valor entero  $q(ykl)$  para el segundo vector característico. Ventajosamente, uno al menos de estos valores enteros  $q(xij)$  y  $q(ykl)$  se obtiene de la segunda estructura de almacenamiento 7. La recuperación de estos valores enteros  $q(xij)$  y  $q(ykl)$ , entonces, es particularmente rápida.

En una etapa 404, el primer operador 13 compara los valores enteros  $q(xij)$  y  $q(ykl)$  obtenidos en la etapa 402.

20 Si estos valores enteros  $q(xij)$  y  $q(ykl)$  son iguales, entonces el primer operador devuelve un booleano de valor "1" en cuanto valor de similitud  $f(xij, ykl)$ , en la etapa 406. Esto indica la similitud de los vectores característicos en el sentido de este primer criterio.

Si no, el valor de similitud  $f(xij, ykl)$  devuelto toma el valor "0", en la etapa 408.

25 Cuando la segunda estructura de almacenamiento 7 está realizada en forma de listas invertidas, y uno de los valores enteros  $q(xij)$  y  $q(ykl)$  es conocido por el primer operador 13, la comparación con el otro de estos valores enteros equivale a recorrer la lista identificada por el valor entero conocido, para verificar si esta lista contiene el identificador del descriptor correspondiente al vector característico del otro valor entero. Un valor entero  $q(xij)$  asignado al vector característico  $xij$  del descriptor  $dij$  es conocido en particular cuando ha sido recibido en la etapa 400, o también cuando ha sido recuperado previamente en la segunda estructura de almacenamiento 7.

30 La memorización de una representación de los vectores característicos de los descriptores en forma de valores enteros en la segunda estructura de almacenamiento 7 permite una comparación particularmente rápida de vectores característicos entre sí. Esta comparación es todavía más rápida cuando se pone en práctica una estructura de listas invertidas para la segunda estructura de almacenamiento 7.

En una variante de realización ilustrada por la figura 5, la etapa 404 se prosigue en las siguientes etapas, en sustitución de la etapa 406.

35 En una etapa 500, el primer operador 13 calcula un producto escalar  $h(bij(xij), bkl(ykl))$  de las firmas binarias  $bij(xij)$  y  $bkl(ykl)$  respectivamente procedentes del primer vector característico  $xij$  y del segundo vector característico  $ykl$ . Los valores de estas firmas binarias  $bij(xij)$  y  $bkl(xkl)$  se recuperan en la segunda estructura de almacenamiento 7, a partir de los valores enteros  $q(xij)$  y  $q(ykl)$  respectivamente correspondientes al primer vector característico  $xij$  y al segundo vector característico  $ykl$ .

En una etapa 502, se compara el producto escalar  $h(bij(xij), bkl(ykl))$  obtenido con un valor umbral  $ht$  predeterminado.

40 Si este producto escalar  $h(bij(xij), bkl(ykl))$  es inferior al valor umbral  $ht$ , entonces el primer operador devuelve un booleano de valor "1" (uno) en cuanto valor de similitud  $f(xij, ykl)$ , en la etapa 504. Si no, el primer operador 13 devuelve un booleano de valor "0" en cuanto valor de similitud  $f(xij, ykl)$ , en la etapa 506.

En un desarrollo ventajoso, el primer operador devuelve, en la etapa 506, un valor ponderado en cuanto valor de similitud  $f(xij, ykl)$  en sustitución de un booleano de valor "1".

45 Como se describe en [5], por ejemplo, se puede utilizar una ponderación de tipo "tf-idf" (por "term frequency - inverse document frequency", a saber: "frecuencia de los términos - frecuencia inversa de los documentos"). Esta ponderación se utiliza con frecuencia dentro del ámbito de la búsqueda de documentos textuales.

50 El dispositivo de procesamiento 5 comprende además un segundo operador 15 capaz de interacción con la segunda estructura de memoria 7 para establecer un parámetro de ángulo de giro  $\Delta A$  para dos descriptores  $dij$  y  $dkl$  de la primera estructura de memoria 5. El segundo operador se establece para establecer este parámetro de ángulo de giro  $\Delta A$  a partir de los datos de valores de ángulo característico cuantificado  $q_{aij}$  y  $q_{akl}$  memorizados en la segunda estructura de almacenamiento 7. Cuando se utiliza una estructura de listas invertidas para esta segunda

estructura de almacenamiento 7, estos valores de ángulo característico cuantificado  $q_{aij}$  y  $q_{akl}$  pueden ser leídos a partir de los valores enteros  $q(xij)$  y  $q(ykl)$  asignados a los descriptores  $dij$  y  $dkl$ .

La figura 6 ilustra un modo de realización de este segundo operador 15.

5 En una etapa 600, el segundo operador 15 recibe un primer vector característico  $xij$  y un segundo vector característico  $ykl$ , o, preferentemente, únicamente identificadores de estos vectores característicos.

10 En una etapa 602, el segundo operador 15 recupera los datos de ángulo característico cuantificado  $q_{aij}$  y  $q_{akl}$  para cada uno de los descriptores  $dij$  y  $dkl$ . Cuando en la etapa precedente son recibidos el primer vector característico  $xij$  y segundo vector característico  $ykl$ , puede ser necesario establecer previamente sus valores enteros  $q(xij)$  y  $q(ykl)$  para interrogar la segunda estructura de almacenamiento 7, en particular cuando la misma está realizada en forma de listas invertidas.

Cuando se utiliza tal estructura de listas invertidas, el conocimiento de uno de estos valores enteros, por ejemplo  $q(xij)$ , permite una búsqueda muy rápida consistente en recorrer la lista identificada por este valor entero  $q(xij)$  para encontrar en ella el identificador del descriptor  $dkl$  al que está agregado el valor de ángulo cuantificado  $q_{akl}$ .

15 En una etapa 604, el segundo operador 15 devuelve una diferencia entre los valores de ángulo característico cuantificado  $q_{aij}$  y  $q_{akl}$  en cuanto parámetro de giro  $\Delta A$ . Preferentemente, el parámetro de giro  $\Delta A$  es devuelto en forma cuantificada.

20 El dispositivo de procesamiento 5 también comprende además un tercer operador 17 capaz de interacción con la segunda estructura de memoria 7 para establecer un parámetro de factor de escala  $\Delta S$  para dos descriptores  $dij$  y  $dkl$  memorizados en la primera estructura de almacenamiento 5 a partir de los datos de factores de escala cuantificados memorizados en la segunda estructura de memoria 7.

La figura 7 ilustra un modo de realización de este tercer operador 17, análogo al modo de realización del segundo operador 15.

25 En una etapa 700, el tercer operador 17 recibe un primer vector característico  $yij$  y un segundo vector característico  $ykl$ , o, preferentemente, únicamente identificadores de los descriptores  $dij$  y  $dkl$  a los que corresponden estos vectores característicos. Se puede recibir asimismo uno de los valores enteros  $q(xij)$  y  $q(ykl)$  asignados a estos vectores característicos.

En una etapa 702, el tercer operador recupera los datos de factor de escala característico cuantificado  $q_{sij}$  y  $q_{skl}$  de cada uno de los dos descriptores  $dij$  y  $dkl$ .

30 En una etapa 704, el tercer operador devuelve la diferencia logarítmica de los factores de escala característicos en cuanto parámetro de escala  $\Delta S$ . Preferentemente, el parámetro de escala  $\Delta S$  es devuelto en una forma cuantificada.

35 El dispositivo de procesamiento también comprende un interrogador 19, capaz de recibir un primer juego de descriptores  $di1$ , relativo a una primera imagen o, preferentemente, únicamente un identificador  $id1$  de esta primera imagen, y un segundo juego de descriptores  $di2$ , relativo a una segunda imagen o, preferentemente, únicamente un identificador  $id2$  de esta segunda imagen.

La figura 8 ilustra un modo de realización del interrogador 19. En la etapa 800, el interrogador 19 recibe el identificador  $id1$  de una primera imagen y el identificador  $id2$  de una segunda imagen.

40 Por cada pareja formada a partir de un descriptor  $di1$  de la primera imagen y de un descriptor  $dj2$  de la segunda imagen, el interrogador 19 llama iterativamente al primer operador 13, en el curso de una etapa 802. Esto se realiza para el conjunto de los juegos de descriptores primero y segundo. El operador 13 es llamado cada vez con el concurso de los identificadores de los descriptores  $di1$  y  $dj2$ .

El primer operador 13 devuelve un valor de similitud  $f(xi1,yk2)$ , en la etapa 804. Como la segunda estructura de almacenamiento 7 optimiza las operaciones de búsqueda de cálculo entre descriptores, merced a la representación indexada, las etapas 802 y 804 se realizan de manera particularmente rápida.

45 El interrogador 19 se establece para interaccionar con una tercera estructura de almacenamiento 20 con el fin de memorizar en ella una correspondencia entre una variable de puntuación  $Sc$  y un identificador de una al menos de las imágenes primera y segunda. La variable de puntuación  $Sc$  suma los valores de similitud  $f(xi1,yk2)$  devueltos en cuanto resultados de las sucesivas llamadas al primer operador 13, en el curso de una etapa 806.

50 Al término de la etapa 806, el valor de la variable de puntuación  $Sc$  resultante da una medida de semejanza, o de similitud, entre la primera imagen y la segunda imagen. El valor de la variable de puntuación  $Sc$  se puede ver asimismo como una medida de pertinencia, o de proximidad, de la segunda imagen con respecto a la primera imagen. Esta primera imagen se puede ver como una imagen de consulta.

El valor final de la variable de puntuación  $Sc$  se puede procesar mediante una función estadística, por ejemplo para ponderar este valor.

5 Cuando la segunda estructura de almacenamiento 7 está realizada en forma de listas invertidas, esta comparación de las imágenes primera y segunda entre sí resulta aún más acelerada. Cuando son conocidos, por ejemplo, los valores enteros  $q(xi1)$  asignados a los descriptores de la primera imagen, las etapas 802, 804 y 806 consisten en recorrer cada una de las listas identificadas por estos valores enteros  $q(xi1)$  en busca, cada vez, de un identificador de un descriptor de la segunda imagen. Cuando se localiza tal identificador, la variable de puntuación  $Sc$  se aumenta en el valor  $f(xi1, yk2)$ . Por ejemplo, la variable de puntuación  $Sc$  se aumenta cada vez en el valor 1 ("uno"). En el caso en que se memoriza un identificador de imagen  $idj$  en sustitución de un identificador de descriptor  $dij$ , lo que se busca es el identificador de la segunda imagen, lo cual acelera aún más la comparación.

El interrogador 19 puede ser llamado repetitivamente para la primera imagen, o el primer juego de descriptores  $di1$ , cada vez con una segunda imagen, o un segundo juego de descriptores  $dij$ , representada en la segunda estructura de almacenamiento 7, en orden a comparar la primera imagen con el conjunto de las imágenes representadas en la segunda estructura de almacenamiento 7.

15 Así, se puede obtener una variable de puntuación  $Scj$  para cada una de las imágenes  $j$  de la base, que representa una medida de proximidad de la imagen  $j$  con respecto a la primera imagen, o imagen de consulta. En otras palabras, la tercera estructura de almacenamiento 20 puede contener un cuadro de puntuaciones  $Sc1j$ , relativo a la primera imagen, que agrupa el conjunto de las variables de puntuaciones  $Scj$  en correspondencia cada vez con un identificador  $idj$  de la imagen  $j$ .

20 Así, se obtiene una clasificación de las imágenes  $j$  en lo que respecta a su semejanza, o similitud, con la primera imagen.

25 Cuando la segunda estructura de almacenamiento 7 está realizada en forma de listas invertidas, la comparación de la primera imagen con el conjunto de las imágenes representadas en esta segunda estructura 7 es particularmente rápida. El interrogador 19 se establece para, a partir de cada uno de los valores enteros asignados a los descriptores de la primera imagen, recorrer la lista identificada por este valor entero en la segunda estructura de memoria 7 y para aumentar la puntuación  $Scj$  de cada una de las imágenes a las que corresponde una entrada de esta lista. Se comprende que, entonces, es de particular interés memorizar en esta entrada el identificador de imagen  $idj$  en sustitución de un identificador de descriptor para aumentar aún más la rapidez de la comparación.

30 Cuando se memorizan firmas binarias  $bi$  en la segunda estructura de almacenamiento 7, la puntuación  $Scj$  de una imagen  $j$  es todavía más pertinente, ya que la comparación de las firmas binarias afina la comparación de dos imágenes entre sí. El Solicitante ha comprobado que las imágenes devueltas por el interrogador 19, en este modo de realización, son más pertinentes que las imágenes identificadas por los dispositivos del estado de la técnica.

35 En estos dispositivos, se debe llevar a cabo una elección de cardinal de índice  $k$ . Un valor de cardinal de índice  $k$  escaso tiene como consecuencia amplias celdas de Voronoi. Esto presenta la ventaja de que tienen una gran probabilidad de encontrarse en la misma celda (mismo valor entero) versiones de los descriptores afectadas de ruido, y el inconveniente de que los descriptores pierdan poder discriminativo, como quiera que pueden encontrarse en la misma celda descriptores muy diferentes. Un valor de cardinal de índice  $k$  bajo tiene como consecuencia celdas de Voronoi reducidas. Esto tiene la ventaja de conservar una buena precisión de los descriptores, pero presenta el inconveniente de una acusada probabilidad de que se encuentre una versión afectada de ruido de un descriptor en una celda diferenciada de aquella del descriptor en cuestión, así como de un mayor consumo de memoria.

Merced a las firmas binarias, el dispositivo según la invención presenta las ventajas de un valor elevado de cardinal de índice  $k$  y de escasos valores de cardinal de índice  $k$ .

45 En un segundo modo de realización, ilustrado por la figura 9, el interrogador 19 se establece para realizar las siguientes operaciones, cuando el valor de similitud devuelto en la etapa 804 es no nulo (etapa 900).

El interrogador 19 llama al segundo operador 13 y al tercer operador 15 para los descriptores  $di1$  y  $dk2$  que están siendo comparados, en la etapa 902. El segundo operador 13 y el tercer operador 15 devuelven respectivamente un valor cuantificado de parámetro de ángulo de giro  $\Delta A$  y un valor cuantificado de parámetro de factor de escala  $\Delta S$ , en la etapa 904.

50 El interrogador 19 se establece para interactuar con la tercera estructura de almacenamiento 20 en orden a memorizar una correspondencia entre el valor cuantificado de ángulo de giro  $\Delta A$ , el valor cuantificado del factor de escala  $\Delta S$  y el valor de similitud  $f(di1, dk2)$ , por una parte, y un identificador de una al menos de las imágenes primera y segunda.

55 En este punto, el interrogador 19 se establece para crear un cuadro de puntuaciones  $Sc2asj$  que reúne, para cada pareja de valor cuantificado de ángulo de giro  $\Delta A$  y de valor cuantificado de factor de escala  $\Delta S$ , una puntuación  $Sc$  aumentada sucesivamente en el valor de similitud  $f(di1, dk2)$  devuelto en esas condiciones de ángulo

- 5 y de escala. En este cuadro, se cuantifican los datos de ángulo y de escala. En otras palabras, el interrogador 19 construye el cuadro de puntuación  $Sc2as[]$  para la segunda imagen un cuadro de dimensiones escogidas, indicativo de las clases de ángulo, por ejemplo en columnas, y de las clases de factor de escala, por ejemplo en ordenadas: el valor de similitud se inscribe en correspondencia con las clases de ángulos y de factor de escala, en su caso, sumado a un valor de similitud anterior.
- 10 El cuadro de puntuación  $Scas[]$  así construido se puede procesar por medio de un estimador 21 capaz de interacción con la tercera estructura de almacenamiento 20 para efectuar procesamientos estadísticos. Por ejemplo, el estimador 21 se puede establecer para establecer, en cuanto valor final de puntuación  $Sc2$  para la imagen 2, el valor máximo de las puntuaciones del conjunto del cuadro de puntuación  $Scas[]$ . En la práctica, el estimador 20 se puede establecer en orden a suavizar los resultados obtenidos reuniendo los resultados de celdas adyacentes en una misma celda, para evitar efectos de borde, o ruido de cuantificación.
- 15 Como variante, el estimador se puede establecer para ponderar los resultados para ciertos valores de ángulos y/o de escala.
- Esta ponderación se elige según la aplicación para dar preferencia a ciertos ángulos de giro y ciertas relaciones de escalas, en particular los ángulos de valor 0, que puedan corresponder a tomas de imagen en modo paisaje, y  $90^\circ$ , que puedan corresponder a una toma de imagen en modo retrato, y al valor de factor de escala 0, que pueda corresponder a la ausencia de efecto de modificación de distancia focal.
- 20 Cuando la segunda estructura de almacenamiento 7 está realizada en forma de listas invertidas, la comparación entre dos imágenes es todavía más rápida, en particular merced al almacenamiento de los valores cuantificados de ángulo y de escala directamente en estas listas. Conociéndose el conjunto de los valores enteros  $q(x1i)$  asignados a los descriptores de la primera imagen, la comparación consiste en recorrer cada una de las listas identificadas por uno de estos valores enteros en busca del identificador  $id2$  de la segunda imagen, o de un identificador de descriptor correspondiente a esta segunda imagen. Cuando se encuentra tal entrada, el valor de similitud se memoriza en correspondencia con los valores cuantificados de ángulo y de factor de escala.
- 25 El interrogador 19 puede realizar el mismo procesamiento con el conjunto de las imágenes  $j$ , o de los juegos de descriptores  $dij$ , una representación del cual está memorizada en la segunda estructura de almacenamiento 7. De ello resulta un cuadro de puntuaciones  $Scas[]$  que memoriza una correspondencia entre un identificador  $idj$  de imagen  $j$  y su cuadro de puntuación particular  $Scjas[]$ . Este cuadro de puntuaciones  $Scas[]$  puede ser procesado por el estimador 21 para que se asocie un solo valor de puntuación, llamado final,  $Scj$ , a cada uno de los identificadores de imagen  $j$ . Así, se puede obtener una clasificación de las imágenes  $j$  en lo que respecta a su similitud con la imagen 1.
- 30 El Solicitante ha comprobado que las imágenes resultados devueltos por el interrogador 19 en este modo de realización son todavía más pertinentes, en particular debido a que se tienen en cuenta características de naturaleza geométrica en las comparaciones. Debido a que la comparación de estas características geométricas se limita al establecimiento de un ángulo de giro y de un factor de escala solamente, el conjunto del proceso de búsqueda no deja de ser, a efectos prácticos, muy rápido. Esta comparación, al obtenerse mediante una simple operación aritmética sobre valores cuantificados y memorizados en la segunda estructura de almacenamiento, es poco consumidora de recursos de cálculo y rápida. En particular, la determinación de los parámetros de ángulo de giro  $\Delta A$  y de factor de escala  $\Delta S$  no hace intervenir los datos de vector característico para evitar cálculos demasiado abundantes que perjudiquen las prestaciones del dispositivo 1.
- 35 El caso de utilizarse una estructura de listas invertidas, la comparación de imágenes es todavía más rápida.
- En particular, se puede utilizar un cuadro de puntuaciones para cada una de las entradas de la estructura de listas invertidas y, luego, los cuadros de puntuación de los descriptores de una misma imagen reunidos en un mismo cuadro. Esto es todavía más rápido y eficiente cuando cada entrada de las listas invertidas contiene un identificador de imagen. En este caso, cada imagen recibe solamente una puntuación única para el primer modo de realización del interrogador.
- 40 En una variante de este segundo modo de realización, el interrogador 19 se establece en orden a mantener para cada imagen  $j$  un cuadro de puntuaciones de ángulo  $Sa$  y un cuadro de puntuación de factor de escala  $Ss$ , que respectivamente reúnen los valores de similitud, eventualmente sumados, en función de los valores cuantificados de parámetro de giro  $\Delta A$  y de factor de escala  $\Delta S$ , clases de ángulo y las puntuaciones en función de las clases de factor de escala.
- 45 En este caso, queda simplificada la tercera estructura de almacenamiento 20, y acelerado el proceso de clasificación.
- 50 El estimador 21 se puede establecer para asignar a una imagen  $j$  el mínimo valor de los mayores valores de puntuaciones de cada uno de los cuadros de puntuación de ángulo  $Sa[]$  y de puntuación de escala  $Scs[]$  en cuanto valor de puntuación final  $Scj$ .
- 55

Cuando se utiliza una estructura de listas invertidas, cada uno de estos vectores se puede almacenar en conjunción con la entrada correspondiente.

Como se ha visto, la segunda estructura de almacenamiento 7 es particularmente compacta cuando está realizada en forma de listas invertidas.

5 En la figura 10, se representa un ejemplo de realización de la segunda estructura de almacenamiento 7 en forma de listas invertidas.

Cada una de las listas  $L_1, L_2, \dots, L_6$  está identificada por un valor de índice, en este punto, 1, 2,... 6.

10 Cada lista  $L_i, i = 1, \dots, 6$ , agrupa cierto número de entradas  $E_{ij}$ . Cada entrada incluye un identificador de imagen  $ID_k$ , un valor cuantificado de ángulo característico  $A_{k,l}$ , un valor cuantificado de factor de escala  $Sk_l$  (valor logarítmico) y un valor de firma binaria  $B_{kl}$ .

Cada entrada  $E_{ij}$  es tal que a un vector característico  $x_{lk}$  de la imagen  $k$  se le ha asignado, en cuanto valor entero  $q(x_{lk})$ , el número de índice  $i$  de la lista. Como variante, cabría almacenar un identificador de descriptor.

15 Por ejemplo, la imagen identificada por  $ID_1$  está representada en esta segunda estructura de almacenamiento 7 mediante un único descriptor (una sola entrada en la estructura de listas invertidas, referenciada con  $E_{61}$ ), al que se ha asignado el valor entero 6. En otras palabras, este descriptor está representado mediante su valor entero en la segunda estructura de almacenamiento 7.

20 La tercera imagen, identificada  $ID_3$ , se caracteriza en la primera estructura de almacenamiento 5 por tres descriptores (3 entradas incluye el identificador  $ID_3$  en la segunda estructura de almacenamiento 7). Estos descriptores están respectivamente representados, en esta estructura de listas invertidas, por los valores enteros 1, 5 y 6, como muestran las entradas  $E_{13}, E_{15}$  y  $E_{62}$ .

25 El interés de una estructura en listas invertidas está en un almacenamiento particularmente compacto de una representación de un número muy grande de imágenes. La cantidad de memoria asignada a cada una de las entradas de las listas se reduce en la medida de lo posible, al propio tiempo que conserva la máxima información para la comparación de descriptores entre sí. El Solicitante ha obtenido buenos resultados en cuanto a compacidad y a pertinencia de los resultados con entradas establecidas como sigue:

- 21 bits de memoria se reservan para la identificación de una imagen ( $ID_k$ ),
- 6 bits de memoria se reservan para el almacenamiento del valor cuantificado de ángulos característicos  $A_{kl}$ ,
- 5 bits de memoria se reservan para el almacenamiento del logaritmo de la escala característica  $Sk_l$ , y
- 64 bits se reservan para el almacenamiento de la firma binaria  $B_{kl}$ .

30 En el primer operador 11, el valor umbral  $ht$  y la dimensión  $db$  se establecen de tal modo que el valor  $ht$  sea inferior o igual al valor de la dimensión  $db$ . Este valor umbral  $ht$  debe ser lo suficientemente elevado para garantizar que los  $n$  vectores más próximos, en el sentido de la distancia euclidiana, de un vector característico  $x_{ij}$  se determinen como similares, es decir, el producto escalar de su firma binaria debe ser inferior a  $ht$ . El valor umbral  $ht$ , al mismo tiempo, debe ser bajo para filtrar los numerosos puntos que se hallan dispuestos en una parte alejada de la correspondiente celda de Voronoi. Por lo tanto, la elección del valor umbral  $ht$  resulta de un compromiso.

Las representaciones gráficas de las figuras 11 y 12 muestran respectivamente la tasa de descriptores recuperados en función del valor umbral  $ht$ , y la tasa de 5-ésimos vectores más próximos recuperados en función de la tasa de puntos recuperados en una celda.

40 El gráfico de la figura 11 se ha obtenido con un valor de dimensión  $db$  fijado en 64. Por tasa de descriptores recuperados, se entiende respectivamente la tasa de descriptores de celda y la tasa de  $i$ -ésimos vectores más próximos, en este punto 5, que son filtrados por el primer operador 13.

45 Los gráficos 11 y 12 se han generado analizando un conjunto de descriptores asignado a un mismo centroide. Dado un vector característico  $x$ , la tasa de descriptores  $m$  filtrados por el primer operador 13, es decir, los descriptores para los cuales el valor de similitud  $f(x, m)$  con el vector  $x$  es nulo, se compara con la tasa de 5 vectores más próximos que son recuperados en una celda que comprende 1000 descriptores.

50 El gráfico de la figura 11 muestra que la elección del valor umbral  $ht$ , escogido en este punto entre 20 y 30, garantiza que se filtren la mayoría de los descriptores de celdas y que los  $n$  vectores más próximos de estos descriptores son preservados con una fuerte probabilidad. Un valor umbral  $ht$  fijado en 22 tiene como consecuencia un filtrado de aproximadamente el 97 % de los descriptores, en tanto que se preserva el 53 % de los cinco vectores más próximos. Un valor umbral  $ht$  más elevado, en este punto fijado en 28, culmina en una conservación al 94 % de los cinco vectores más próximos y un filtrado del 77 % de los descriptores de celdas.

5 El gráfico de la figura 12 representa la desviación entre las tasas de filtrado de estos diferentes descriptores para diferentes valores de dimensiones *db*. La mejor calidad de filtrado se obtiene para los mayores valores de dimensión *db*. Valores grandes de dimensión *db* precisan de un mayor espacio de almacenamiento en memoria. En consecuencia, el valor de la dimensión *db* resulta asimismo de un compromiso entre la precisión del primer operador 13 y la cantidad de memoria utilizada. Una dimensión *db* que valga 64 parece ser un buen compromiso.

10 El dispositivo según la invención permite la construcción de una base compacta con posibilidad de incluir una representación de un número muy grande de imágenes, una búsqueda de imágenes similares más pertinente que los dispositivos del estado de la técnica, teniendo en cuenta elementos geométricos de estas imágenes, y más rápida. Estas ventajas se ven mejoradas más en particular cuando la base en cuestión adopta la forma de una estructura de listas invertidas.

La invención no está limitada al modo de realización antes descrito, a título de ejemplo únicamente.

En particular:

- los operadores primero, segundo y tercero pueden estar realizados en un mismo operador que presente las mismas funcionalidades.
- 15 - Las estructuras de almacenamiento primera, segunda y tercera se han representado como pertenecientes a una misma memoria 3. Se entiende que cada una de estas estructuras puede estar ubicada o repartida en dispositivos físicamente diferenciados, e incluso tecnológicamente diferentes. Por ejemplo, la primera estructura de memoria se puede establecer en un disco duro, en tanto que la segunda estructura de memoria se puede establecer en memoria RAM, en particular cuando está realizada en forma de una estructura de listas invertidas (compacidad).
- 20 - El primer operador y los segundo y tercer operadores pueden ser utilizados independientemente entre sí, en particular en el segundo modo de realización del primer operador.
  - La invención puede ser vista asimismo en forma de procedimientos de ayuda para el reconocimiento de imágenes, cuyas etapas se encuentran en las diferentes formas de realización descritas. Estos procedimientos son susceptibles de generalización de manera acorde con el espíritu de esta descripción.
- 25 - El dispositivo de procesamiento 5 se puede realizar con el concurso de cualesquiera medios de cálculo, en particular un microprocesador capaz de ejecutar productos de programas informáticos.

**REIVINDICACIONES**

1. Dispositivo de ayuda para el reconocimiento de imágenes, que comprende:
    - una memoria (3) capaz de almacenar una pluralidad de juegos de descriptores ( $d_{ij}$ ), correspondiendo cada juego a una imagen,
  - 5 - asociándose cada descriptor ( $d_{ij}$ ) a una zona designada de la imagen, y comprendiendo unos primeros datos de tipo vector ( $x_{ij}$ ), unos segundos datos de tipo ángulo ( $a_{ij}$ ) y unos terceros datos de tipo escalar ( $s_{ij}$ ),
    - un primer operador (13) capaz de recibir una designación de dos descriptores ( $d_{ij}$ ) y de establecer un booleano que refleja la verificación de un criterio de similitud de los vectores de descriptores ( $x_{ij}$ ), a partir de una comparación entre los primeros datos de estos descriptores,
  - 10 - un segundo operador (15) capaz de recibir una designación de dos descriptores ( $d_{ij}$ ) y de establecer un parámetro de ángulo de giro ( $\Delta A$ ) a partir de los segundos datos ( $a_{ij}$ ) respectivos de los dos descriptores ( $d_{ij}$ ),
    - un tercer operador (17) capaz de recibir una designación de dos descriptores ( $d_{ij}$ ) y de establecer un parámetro de factor de escala ( $\Delta S$ ) a partir de los terceros datos de estos descriptores ( $s_{ij}$ ),
  - 15 - un controlador (5) capaz de recibir las designaciones de un primer y de un segundo juego de descriptores ( $d_{ij}$ ), y establecido para aplicar el primer operador (13) a cada pareja formada a partir de un descriptor ( $d_{ij}$ ) del primer juego y de un descriptor del segundo juego ( $d_{ij}$ ), para inferir una lista de parejas de descriptores similares,
    - llamar al segundo operador (15) y al tercer operador (17) por cada una de las parejas de descriptores de dicha lista y almacenar en memoria el conjunto de los parámetros obtenidos, y
  - 20 - un estimador (21) de la semejanza entre las imágenes respectivamente correspondientes al primer juego y al segundo juego de descriptores ( $d_{ij}$ ), a partir de un procesamiento estadístico sobre dicho conjunto de parámetros.
2. Dispositivo según la reivindicación 1, en el que el segundo operador (15) se establece para establecer el parámetro de ángulo de giro ( $\Delta A$ ) a partir de solamente los segundos datos ( $a_{ij}$ ) respectivos de los dos descriptores ( $d_{ij}$ ), y en el que el tercer operador (17) se establece para establecer un parámetro de factor de escala ( $\Delta S$ ) a partir de solamente los terceros datos de estos descriptores ( $s_{ij}$ ).
- 25 3. Dispositivo según una de las reivindicaciones 1 y 2, en el que la memoria comprende una estructura de almacenamiento específica (7), que memoriza una correspondencia entre cada uno de los descriptores ( $d_{ij}$ ) y un dato particular de un conjunto de primeros datos suplementarios de tipo escalar, seleccionados siguiendo una regla predeterminada a partir de los datos de vector característico ( $x_{ij}$ ) del descriptor.
- 30 4. Dispositivo según la reivindicación 3, en el que el primer operador (13) se establece para comparar los primeros datos de los descriptores ( $d_{ij}$ ) designados al menos en parte basándose en una comparación de dicho dato particular ( $q(x_{ij})$ ) correspondiente a dichos descriptores ( $d_{ij}$ ).
- 35 5. Dispositivo según una de las reivindicaciones 3 y 4, en el que la estructura de almacenamiento específica (7) se establece para memorizar una correspondencia entre cada uno de los descriptores ( $d_{ij}$ ), los datos particulares seleccionados para el descriptor ( $d_{ij}$ ) y una representación de algunos al menos de los segundos ( $a_{ij}$ ) y terceros ( $s_{ij}$ ) datos de este descriptor ( $d_{ij}$ ).
6. Dispositivo según la reivindicación 5, en el que uno al menos del segundo operador (15) y del tercer operador (17) se establece para establecer el parámetro de ángulo de giro ( $\Delta A$ ) o el parámetro de factor de escala ( $\Delta S$ ) a partir de las representaciones de los segundos ( $a_{ij}$ ) o terceros ( $s_{ij}$ ) datos en la estructura de memoria específica (7).
- 40 7. Dispositivo según una de las reivindicaciones 3 a 6, en el que la estructura de memoria específica se establece para memorizar una correspondencia entre cada uno de los datos particulares del conjunto de los primeros datos suplementarios y un dato de designación de cada uno de los descriptores ( $d_{ij}$ ) para el cual se selecciona este dato particular, o de designación de este juego de descriptores.
- 45 8. Dispositivo según la reivindicación 7, en el que la estructura de memoria específica (7) se establece en forma de una pluralidad de listas ( $L_i$ ), asociándose cada una de estas listas ( $L_i$ ) a un dato particular ( $q$ ) del conjunto de los primeros datos suplementarios ( $q$ ) e incluyendo, en cuanto entradas ( $E_{kl}$ ), dichos datos de designación.
9. Dispositivo según la reivindicación 8, en el que cada una de las entradas también incluye datos de representación de algunos al menos de los segundos ( $a_{ij}$ ) y terceros ( $s_{ij}$ ) datos del descriptor ( $d_{ij}$ ) designado.

10. Dispositivo según una de las reivindicaciones 3 a 9, en el que el conjunto de los primeros datos suplementarios forma un índice para la estructura de almacenamiento específica.
- 5 11. Dispositivo según una de las reivindicaciones anteriores, en el que los parámetros obtenidos ( $\Delta A$ ,  $\Delta S$ ) se almacenan en forma de un conjunto de parejas formadas cada vez a partir de un valor de parámetro de factor de escala y de un valor de parámetro de ángulo de giro, almacenando además la memoria (3) una correspondencia entre cada una de las parejas así formadas y un identificador del descriptor ( $d_{ij}$ ) del segundo juego a partir del cual se ha establecido esta pareja, o un identificador del juego de este descriptor ( $d_{ij}$ ).
- 10 12. Dispositivo según una de las reivindicaciones 1 a 11, en el que los parámetros obtenidos se almacenan en forma de conjuntos de parámetros de ángulo y de conjuntos de parámetros de factor de escala, almacenando además la memoria una correspondencia entre cada uno de dichos conjuntos y un identificador del descriptor del segundo juego a partir del cual se han establecido dichos valores, o un identificador del juego de este descriptor ( $d_{ij}$ ).
13. Dispositivo según una de las reivindicaciones anteriores, en el que los parámetros obtenidos se asocian respectivamente a clases de valores, y en el que el estimador (21) establece una medida de similitud basándose en un procesamiento estadístico sobre la población de estas clases de valores.
- 15 14. Dispositivo según la reivindicación 13, en el que el procesamiento estadístico comprende la aplicación de un filtro de suavizado sobre las clases de valores, con anterioridad al establecimiento de la medida de similitud.
15. Dispositivo según una de las reivindicaciones anteriores, en el que dicho procesamiento estadístico comprende la aplicación de una función de ponderación para ciertos valores de parámetros de ángulo y/o de parámetros de factor de escala.
- 20 16. Dispositivo según una de las reivindicaciones anteriores, en el que la memoria 3 almacena una correspondencia entre una designación de cada uno de los descriptores ( $d_{ij}$ ) y un valor entero ( $q(x_{ij})$ ) seleccionado para este descriptor ( $d_{ij}$ ) de entre un grupo determinado de valores enteros a partir de los primeros datos ( $x_{ij}$ ), y en el que dicho primer operador (13) se establece de modo que el criterio de similitud de los vectores de descriptores comprenda la identidad de los valores enteros asignados.
- 25 17. Dispositivo según una de las reivindicaciones anteriores, que comprende un dispositivo de firma, capaz de recibir un descriptor ( $d_{ij}$ ), de establecer unos cuartos datos de tipo firma binaria multidimensional, a partir de los primeros datos del descriptor ( $x_{ij}$ ), estableciéndose la memoria (3) para mantener una correspondencia entre un identificador de cada uno de los descriptores, o un identificador del juego de este descriptor, y los cuartos datos establecidos para este descriptor ( $d_{ij}$ ).
- 30 18. Dispositivo según la reivindicación 17 adscrita a la reivindicación 8, en el que cada entrada de la lista ( $i1$ ) incluye además los datos de firma binaria.
- 35 19. Dispositivo según una de las reivindicaciones 17 y 18, en el que el primer operador (13) se establece para obtener las firmas binarias correspondientes a los dos descriptores ( $d_{ij}$ ), y establecer una medida de distancia a partir de estas firmas binarias, y en el que el primer operador (13) se establece de modo que dicho criterio de similitud de los vectores de descriptores comprenda además el hecho de que la medida de distancia obtenida sea inferior a un primer valor umbral ( $ht$ ) determinado.
20. Dispositivo según una de las reivindicaciones 17 a 19, en el que el dispositivo de firma (11) se establece para:
- establecer unos quintos datos de tipo vector, a partir de una matriz de proyección ( $P$ ) determinada y del vector del descriptor ( $x_{ij}$ ),
  - 40 - comparar el valor de cada una de las componentes del vector proyectado con un segundo valor umbral ( $Tq(x_{ij})$ ) determinado,
  - establecer en cuanto firma binaria un conjunto formado a partir de booleanos, reflejando cada booleano el resultado de la comparación para una componente del vector proyectado.
- 45 21. Dispositivo según una de las reivindicaciones anteriores, en el que el conjunto de los descriptores presentes en la memoria (3) se almacenan en forma de una estructura de listas invertidas.

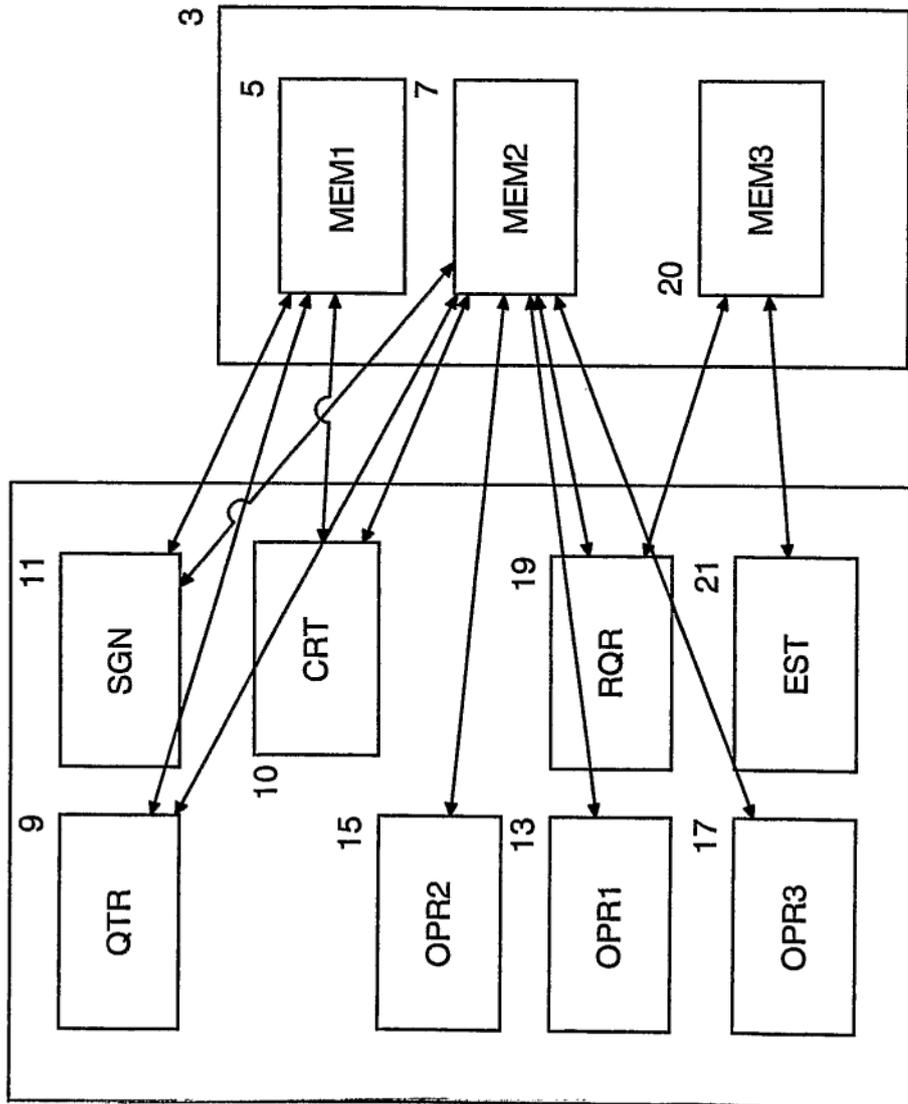


Fig. 1

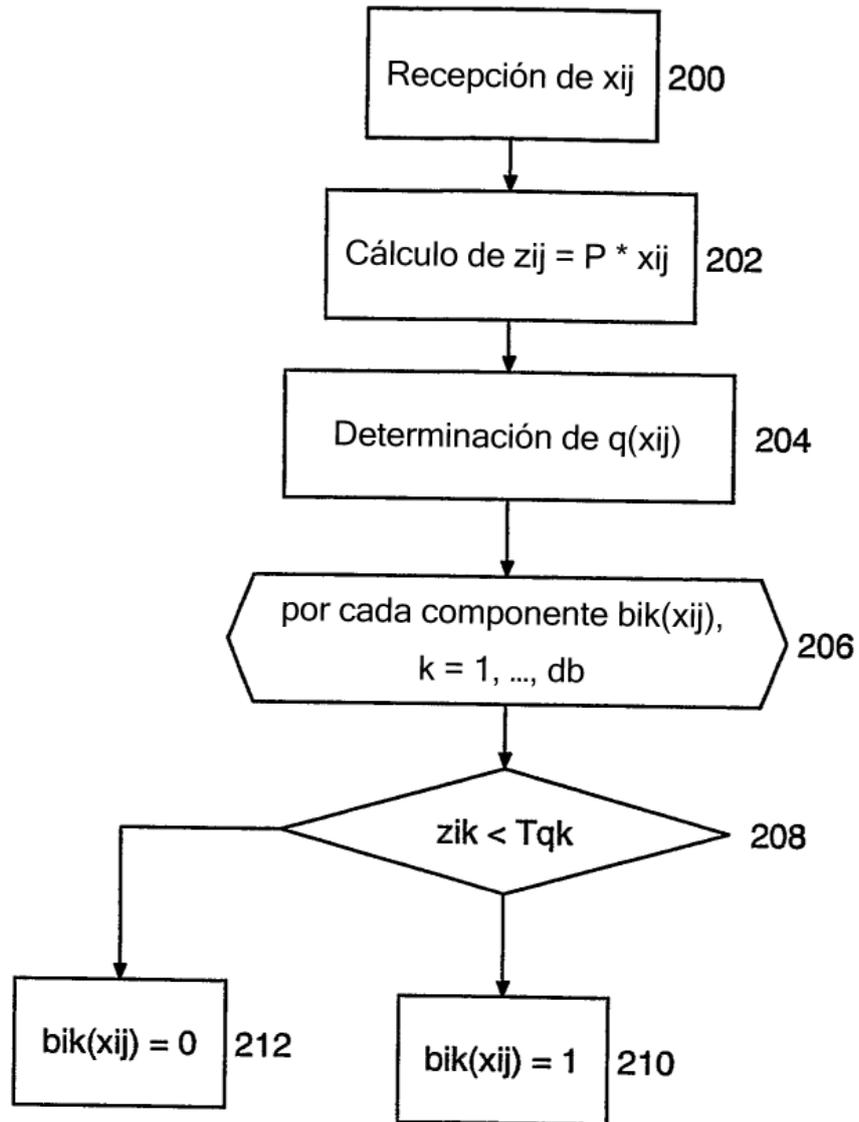


Fig. 2

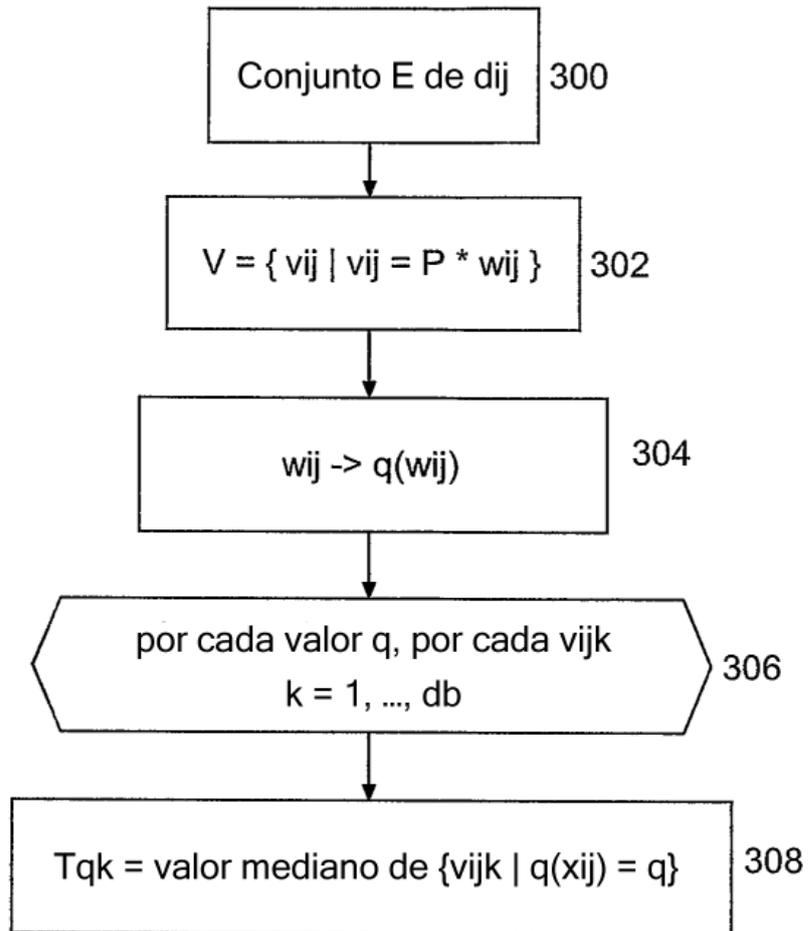


Fig. 3

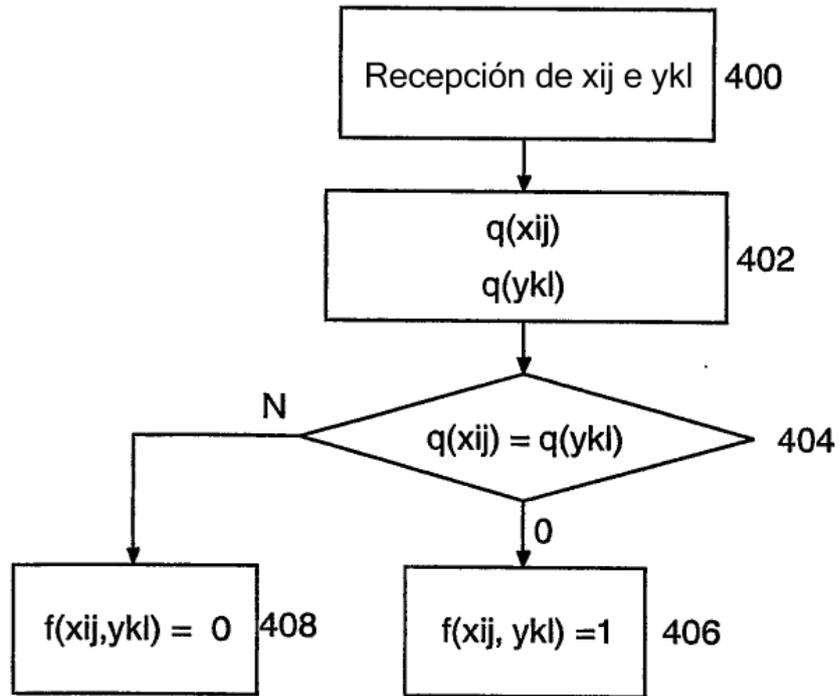


Fig. 4

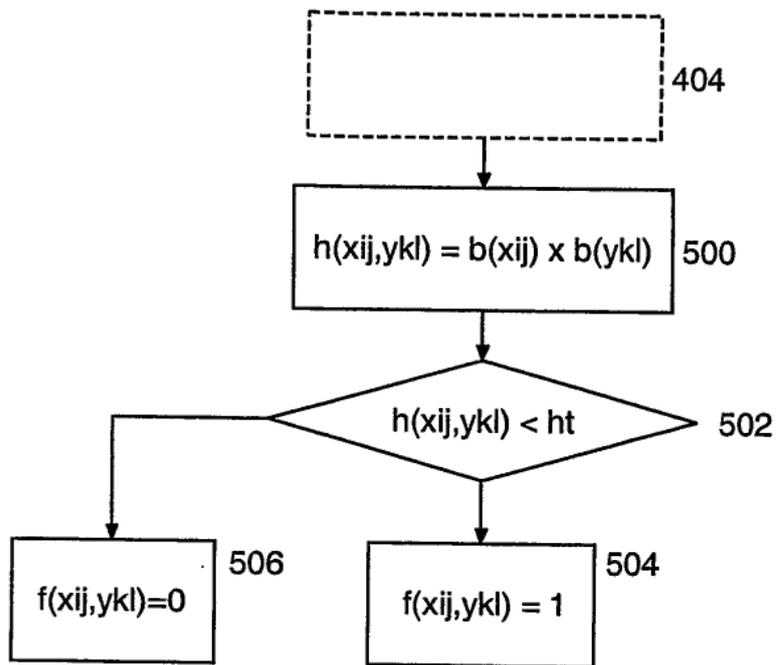
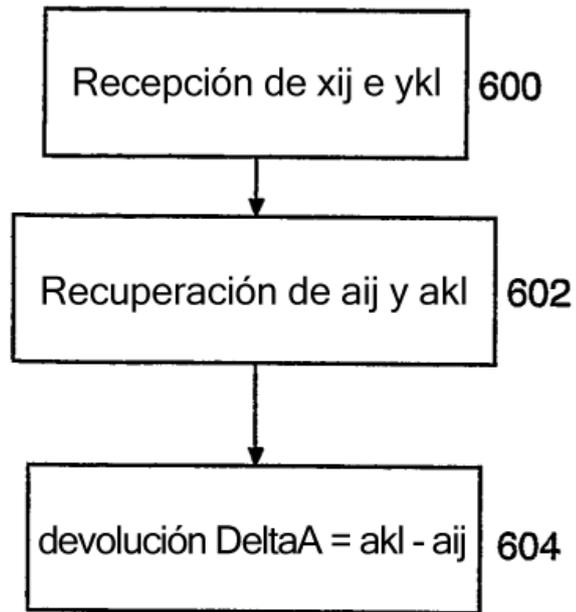


Fig. 5



**Fig. 6**

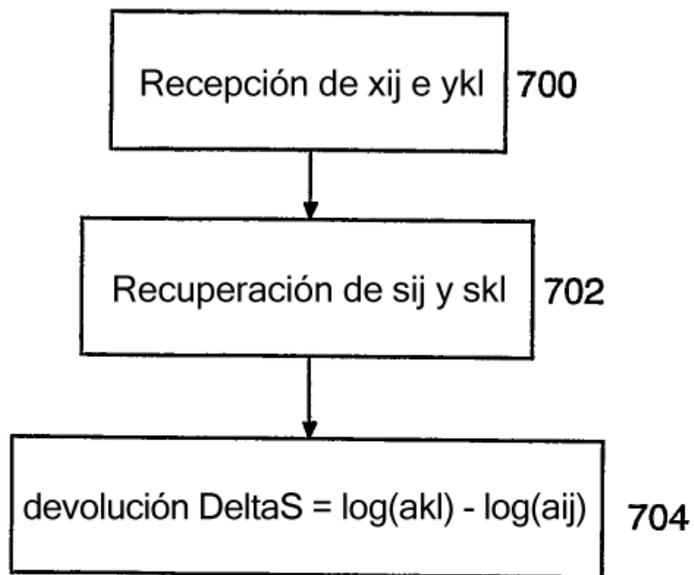


Fig. 7

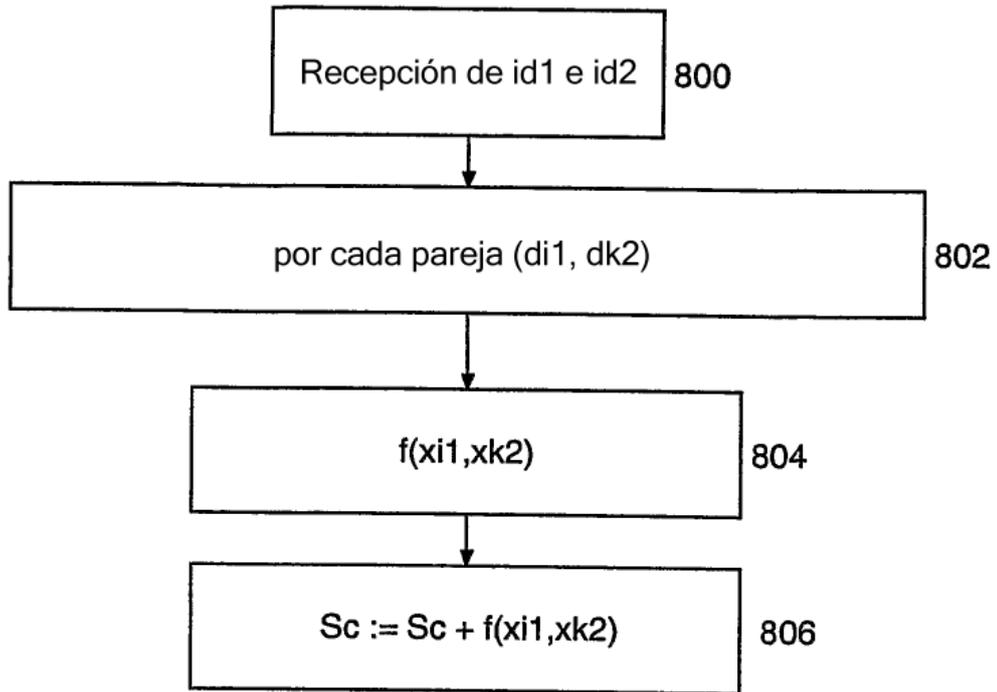
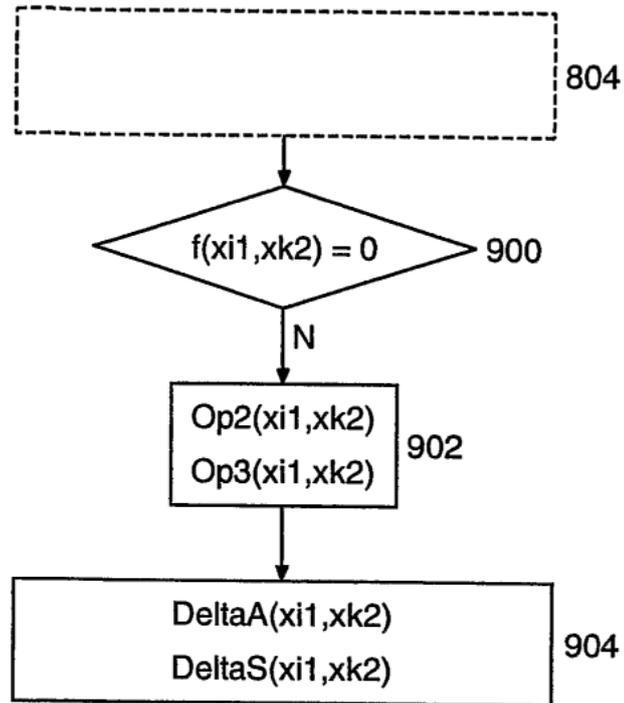


Fig. 8



**Fig. 9**

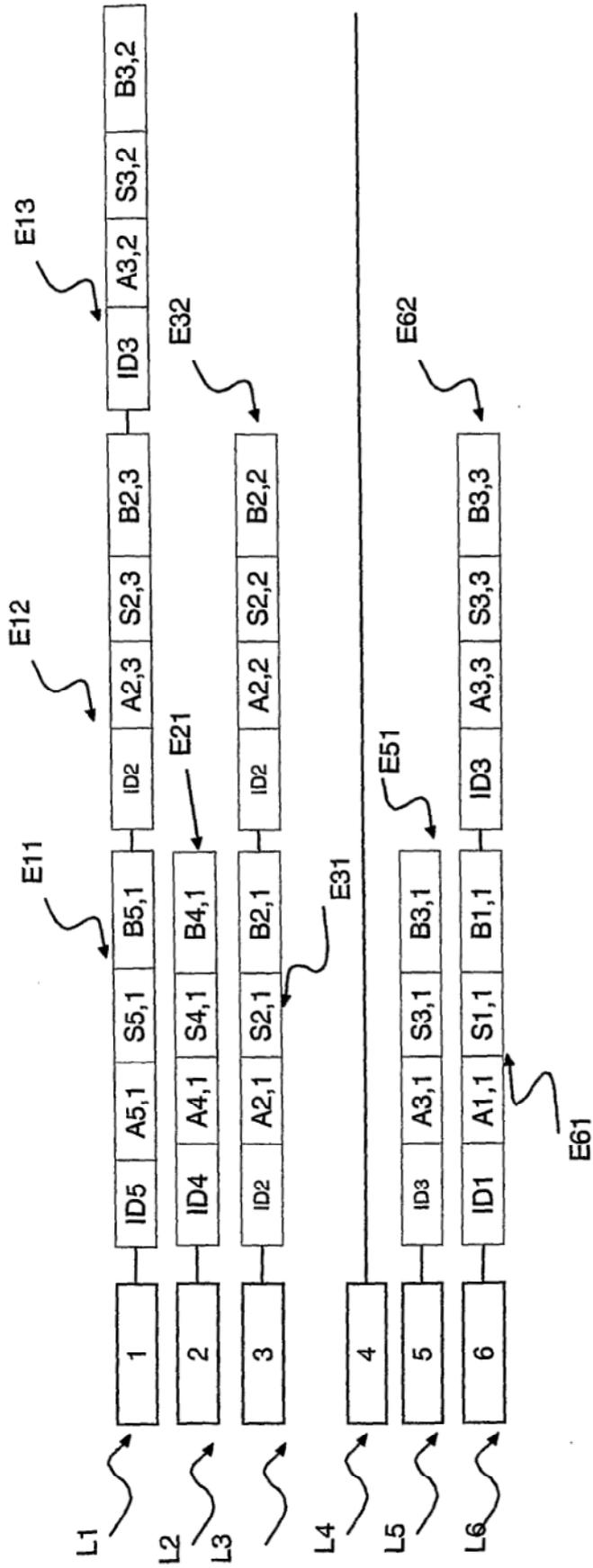


Fig. 10

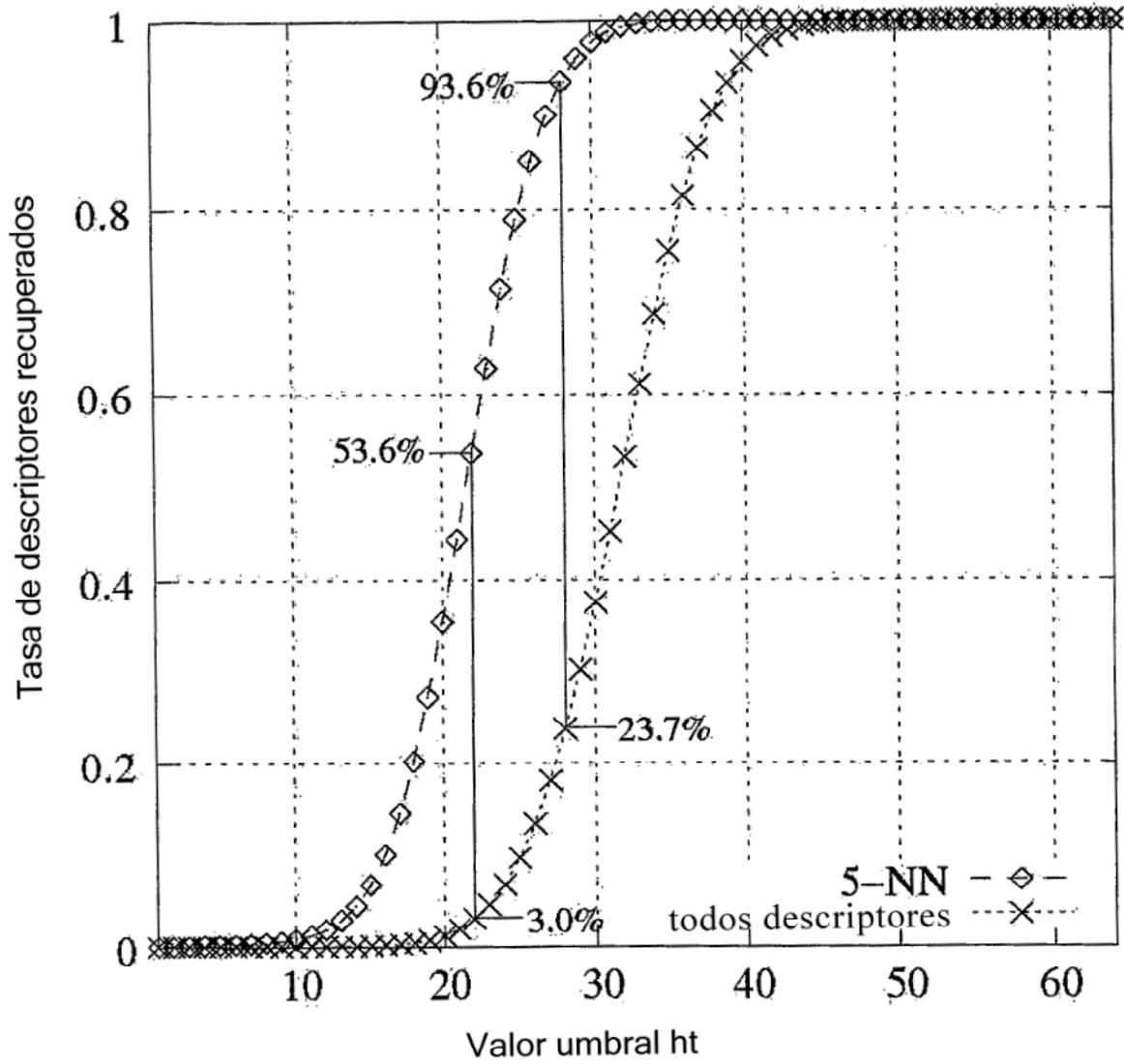


Fig. 11

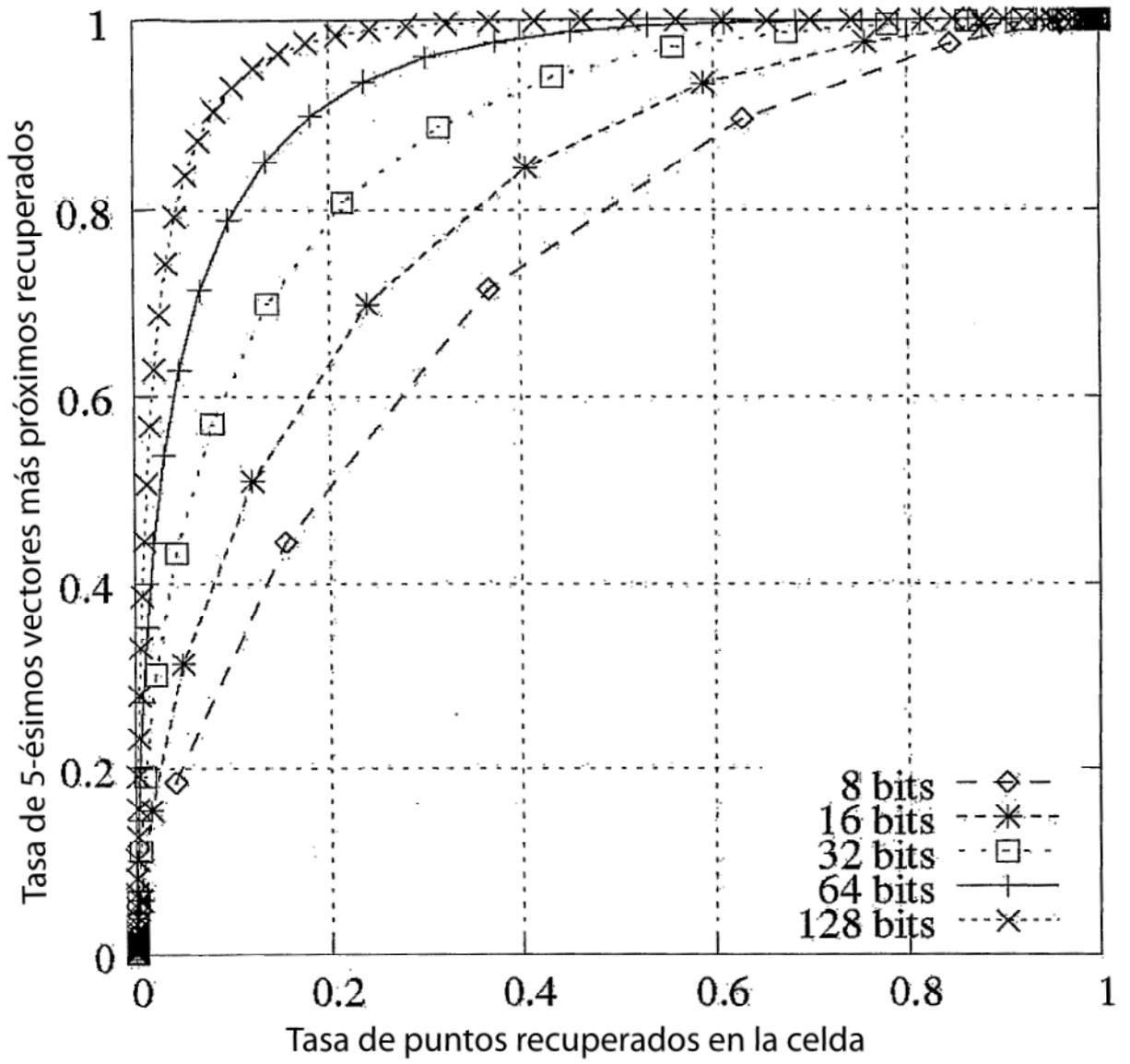


Fig. 12