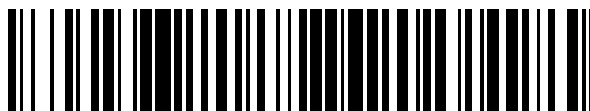


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 811 532**

51 Int. Cl.:

G06T 7/44 (2007.01)

G06K 9/46 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **01.07.2013** **PCT/EP2013/063874**

87 Fecha y número de publicación internacional: **16.01.2014** **WO14009198**

96 Fecha de presentación y número de la solicitud europea: **01.07.2013** **E 13732573 (4)**

97 Fecha y número de publicación de la concesión europea: **24.06.2020** **EP 2870568**

54 Título: **Método para transformar un descriptor de imágenes con base en un histograma de gradientes y un aparato de procesamiento de imágenes relativo**

30 Prioridad:

09.07.2012 IT TO20120602

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

12.03.2021

73 Titular/es:

NEW LUCK GLOBAL LIMITED (100.0%)
P.O. Box 957, Offshore Incorporations Centre
Road Town, Tortola, VG

72 Inventor/es:

PASCHALAKIS, STAVROS y
BOBER, MIROSLAW

74 Agente/Representante:

VIDAL GONZÁLEZ, Maria Ester

ES 2 811 532 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

Método para transformar un descriptor de imágenes con base en un histograma de gradientes y un aparato de procesamiento de imágenes relativo

Esta invención se refiere al cálculo de descriptores de imágenes robustos, discriminativos, escalables y compactos.

Más en particular, la presente invención se refiere a descriptores de imágenes calculados en regiones locales alrededor de puntos de interés de imágenes calculando histogramas de gradientes de subregiones dentro de dichas regiones locales.

Los descriptores de imágenes han encontrado una amplia aplicabilidad en muchas aplicaciones de visión por ordenador, incluido el reconocimiento de objetos, la recuperación de imágenes basadas en contenido y el registro de imágenes, por nombrar algunos. Uno de los ejemplos más conocidos de esta clase de descriptores de imágenes es el descriptor de Transformación de Características Invariantes de Escala (SIFT).

Brevemente, con el método SIFT, los descriptores de imágenes locales se forman de la siguiente manera: primero, se realiza una búsqueda en múltiples escalas y ubicaciones de imágenes para identificar y localizar puntos clave de imagen estables que son invariantes a escala y orientación; luego, para cada punto clave, se determinan una o más orientaciones dominantes en función de los gradientes de imagen locales, lo que permite que el cálculo del descriptor local posterior se realice en relación con la orientación, la escala y la ubicación asignadas de cada punto clave, logrando así la invariancia a estas transformaciones.

Luego, los descriptores de imágenes locales alrededor de los puntos clave se forman de la siguiente manera: primero, la magnitud del gradiente y la información de orientación se calcula en los puntos de muestra de la imagen en una región alrededor del punto clave; luego, estas muestras se acumulan en histogramas de orientación que resumen los contenidos en $n \times n$ subregiones.

Solo a manera de ilustración, se muestra un ejemplo de un descriptor de punto clave en las Figuras 1a y 1b, donde la Figura 1a muestra una subdivisión de la región local R en subregiones 4×4 SR y la Figura 1b muestra una subdivisión del rango de orientaciones de 360° en ocho intervalos h para cada histograma de orientación h , con la longitud de cada flecha correspondiente a la magnitud de esa entrada de histograma.

Por lo tanto, un descriptor de imágenes local como se ilustra en la Figura 1a tiene $4 \times 4 \times 8 = 128$ elementos. El método SIFT se presenta con mayor detalle en el documento de David G. Lowe, "Distinctive image features from scale-invariant keypoints", Revista internacional de visión por ordenador, 60, 2 (2004), págs. 91-110.

Existen varias alternativas y variaciones del método SIFT, que emplean diferentes mecanismos para la detección de puntos clave de imagen estables, diferentes enfoques para la subdivisión de la región local alrededor de puntos clave y diferentes enfoques para el cálculo de histogramas de gradiente de subregión.

Por ejemplo, las Figuras 2a y 2b muestran respectivamente subdivisiones espaciales logarítmicas características de otras técnicas como el Histograma de Orientación de Ubicación de Gradiente (GLOH) descrito en documento de K. Mikolajczyk y C. Schmid, "A performance evaluation of local descriptors", Transacciones IEEE sobre análisis de patrones e inteligencia artificial 27(10): 1615-1630, y el Histograma de Gradientes sin Comprimir (UHoG) descrito en el documento de Chandrasekhar y otros, "Compressed Histogram of Gradients: A Low-Bitrate Descriptor", Revista internacional de visión por ordenador, vol. 94, No. 5, mayo de 2011, como alternativas a la Subdivisión espacial cartesiana empleada en el método SIFT.

Como otro ejemplo, la Figura 3 muestra enfoques para el cálculo de histogramas de gradiente con base en una subdivisión del espacio bidimensional de los componentes x e y de los gradientes en intervalos, característicos de UHoG, como una alternativa a la subdivisión de rango de 360° de orientaciones de gradiente en intervalos que se emplea en el método SIFT.

El documento a nombre de MEHMET FIRAT VURAL y otros: "Registration of multispectral satellite images with Orientation-Restricted SIFT"; Simposio de geociencia y teledetección, 12 de julio de 2009, describe una versión modificada del popular algoritmo SIFT. El algoritmo SIFT restringido de orientación propuesto combina los elementos del vector descriptor SIFT en direcciones opuestas. De esta manera, se logra un mejor rendimiento de coincidencia de características con vectores descriptores más cortos que se traducen en una complejidad de coincidencia más corta. El efecto de la restricción de escala y el estiramiento de contraste se demuestra en imágenes satelitales multiespectrales y se encuentra que la restricción de escala es útil para eliminar características que coinciden incorrectamente.

El documento a nombre de MANJUNATH BS y otros: "Color and Texture Descriptors", Transacciones IEEE en circuitos y sistemas para tecnología de video, 1 de junio de 2001, presenta una descripción general de los descriptores de color y textura que han sido aprobados para el Borrador final del comité del estándar MPEG-7. Los

descriptores de color en el estándar incluyen un descriptor de histograma que se codifica mediante el uso de la transformación de Haar, un histograma de estructura de color, un descriptor de color dominante y un descriptor de diseño de color. Los tres descriptores de textura incluyen uno que caracteriza regiones de textura homogéneas y otro que representa la distribución local del borde. También se define un descriptor compacto que facilita la exploración de texturas. Cada uno de los descriptores se explica en detalle por su semántica, extracción y uso.

Las técnicas de la técnica anterior mencionadas anteriormente se consideran aquí solo a manera de ejemplo de técnicas que producen descriptores de imágenes con base en las cuales la presente invención realiza el cálculo de descriptores de imágenes robustos, discriminativos, escalables y compactos.

Aunque tales descriptores de imágenes han encontrado una amplia aplicabilidad en muchas aplicaciones de visión por ordenador como se describió anteriormente, sus costos de almacenamiento y transmisión, según lo definido por su tamaño en bytes, comúnmente se consideran altos en ciertas áreas de aplicación. Esto se debe a que, aunque el tamaño de un descriptor de imágenes local para un punto clave en una imagen puede ser relativamente bajo, el descriptor de imágenes completo comprenderá cientos de tales puntos clave y descriptores locales asociados, lo que significa que el descriptor de imágenes completo puede tener un tamaño comparable a Versión JPEG comprimida de la imagen real de la que se extrae.

Una de esas áreas de aplicación donde este nivel de tamaño de descriptor se considera problemático es la búsqueda visual mediante el uso de terminales móviles. Aunque las diferentes arquitecturas son factibles en esta área de aplicación, una arquitectura típica implica la captura de una imagen de un objeto de interés por parte de un cliente terminal móvil tal como un teléfono móvil, extracción automática de un descriptor de imágenes por parte del cliente, transmisión del descriptor de imágenes a través de una red de comunicación inalámbrica a un servidor que procesará el descriptor de imágenes y proporcionará una respuesta adecuada, tal como la identidad o información adicional con respecto al objeto de interés, y una devolución de dicha respuesta al cliente. Por lo tanto, es obvio que es conveniente minimizar la cantidad de información transmitida desde el cliente al servidor a través de la red inalámbrica. Para beneficio de tales aplicaciones, ha habido una cantidad significativa de desarrollo en la compresión de tales descriptores de imágenes.

El enfoque más simple para comprimir un histograma de un descriptor de punto clave basado en gradiente es mediante la cuantificación escalar de los valores de intervalo de histograma, lo que significa reducir el número de bits usados en la representación de cada valor del intervalo individualmente. En la práctica, este enfoque no se usa comúnmente porque es difícil lograr tasas de compresión muy altas sin comprometer significativamente el poder discriminativo del descriptor. Por ejemplo, la codificación de los intervalos de histograma del descriptor SIFT con ocho bits por intervalo se usa comúnmente, pero da como resultado descriptores de imágenes cuyo tamaño en bytes se considera demasiado grande para la transmisión a través de redes inalámbricas. Por otro lado, se ha encontrado que la cuantificación escalar a solo unos pocos, por ejemplo, solo uno o dos bits por intervalo compromete el poder discriminativo del descriptor de imágenes.

Por lo tanto, se han propuesto esquemas de compresión más complejos. Una revisión de tales esquemas se presenta en el documento de V. Chandrasekhar y otros, "Survey of SIFT compression schemes", Actas de la Conferencia Internacional sobre Reconocimiento de Patrones (ICPR), Estambul, Turquía, agosto de 2010.

Brevemente, los esquemas que giran en torno a la cuantificación de vectores, mediante los cuales los valores de intervalo se cuantifican conjuntamente al mapearlos en uno de un número finito de centroides de vectores representativos, han sido particularmente populares e investigados en diversas formas, tal como la cuantificación de vectores de productos y estructurada en árbol. El inconveniente de tales enfoques es que conllevan una complejidad computacional relativamente alta y requisitos de memoria bastante significativos, desde cientos de kilobytes hasta varios megabytes o más, para el almacenamiento de los centroides, cuyo número puede variar de miles a millones, y la determinación de lo cual también requiere una fase de entrenamiento computacionalmente compleja.

Los esquemas que giran en torno a la codificación de tipos también se han investigado a fondo, por lo que los valores de intervalo se cuantifican nuevamente formando una red uniforme de tipos dentro del espacio que contiene todos los vectores de entrada posibles y, para cualquier vector de entrada dado, codificándolo por el índice del tipo que está más cerca de eso. Los requisitos de memoria de tales enfoques se reducen en comparación con los enfoques de cuantificación de vectores, pero también se ha encontrado que los descriptores comprimidos resultantes no se comparan bien con los descriptores cuantificados de vectores en términos de rendimiento de reconocimiento a altas tasas de compresión. En general, los costos computacionales asociados con la codificación de tipo son significativamente más altos que para la cuantificación escalar simple.

Otros esquemas de compresión utilizan métodos conocidos de reducción de dimensionalidad, tales como PCA, en descriptores de punto clave, por ejemplo, descriptores de punto clave SIFT de 128 dimensiones, seguidos de cuantificación escalar de las dimensiones resultantes. Un problema clave con tales enfoques es que conllevan una alta complejidad computacional y un alto riesgo de sobreentrenamiento.

En resumen, los enfoques existentes para la compresión del histograma de descriptores basados en gradientes y la

generación de descriptores de imágenes robustos, discriminatorios, escalables y compactos exhiben ciertos inconvenientes.

5 Un enfoque simple tal como la cuantificación escalar de los elementos descriptores tiene el beneficio de una complejidad computacional muy baja y requisitos de memoria, pero se ha encontrado que compromete el poder discriminativo de los descriptores a altas tasas de compresión.

10 Se ha demostrado que los enfoques más complejos logran un mejor rendimiento a altas tasas de compresión, pero sufren diferentes inconvenientes. Los enfoques de cuantificación vectorial han aumentado significativamente la complejidad computacional y los requisitos de memoria. Los enfoques de codificación de tipo implican una mayor complejidad y, aunque no están agobiados por los requisitos de memoria de los enfoques de cuantificación vectorial, también se ha encontrado que tienen un rendimiento inferior en comparación con dichos enfoques. Además, ni la cuantificación del vector ni los enfoques de codificación de tipo son adecuados para la reducción de la dimensionalidad en el dominio comprimido. También se han empleado enfoques con base en técnicas conocidas de reducción de dimensionalidad, tales como PCA, pero también sufren una alta complejidad computacional y un alto riesgo de sobreentrenamiento.

20 Por lo tanto, un objeto de la presente invención es indicar un método para transformar un descriptor de imágenes con base en un histograma de gradiente que permita reducir el tamaño del descriptor de imágenes.

Es un objeto adicional de la presente invención indicar un método para transformar un descriptor de imágenes, con base en un histograma de gradiente, que permita reducir los costos de su almacenamiento y transmisión, en particular a través de una red inalámbrica en aplicaciones de terminales móviles.

25 Es un objeto adicional de la presente invención indicar un método para transformar un descriptor de imágenes con base en un histograma de gradiente que permita reducir el tamaño de un descriptor de imágenes constituido por una pluralidad de histogramas de gradiente relacionados con subregiones que rodean un punto clave de imagen.

30 Estos y otros objetos de la invención se logran mediante un método para transformar un descriptor de imágenes con base en un histograma de gradiente y un aparato de procesamiento de imagen relativo como se reivindica en las reivindicaciones adjuntas, que son una parte integral de la presente descripción.

35 La presente invención proporciona un método para transformar un descriptor de imágenes de acuerdo con la reivindicación 1.

La presente invención también proporciona un aparato de procesamiento de imágenes de acuerdo con la reivindicación 17.

40 En síntesis, la presente invención se refiere al cálculo de descriptores de imágenes robustos, discriminatorios, escalables y compactos a partir de descriptores de imágenes que emplean histogramas de gradientes con base en la transformación de dichos histogramas de gradientes, donde dicha transformación captura la información sobresaliente y robusta contenida en estos en la forma de la forma de las distribuciones y la relación entre sus valores de intervalo. Más específicamente, la presente invención transforma los histogramas de gradientes calculando valores correspondientes a las relaciones entre los intervalos de histogramas.

45 El método de acuerdo con la presente invención también se refiere al cálculo de descriptores de imágenes robustos, discriminatorios, escalables y compactos, con base en diferentes enfoques de manejo del histograma de gradientes relacionados con las subregiones que rodean un punto clave de imagen.

50 Otras características de la invención se exponen en las reivindicaciones adjuntas, que pretenden ser una parte integral de la presente descripción.

55 Los objetos anteriores se harán más evidentes a partir de la siguiente descripción detallada del método para transformar un descriptor de imágenes con base en un histograma de gradiente, con referencia particular a los dibujos anexos, en donde:

- Las Figuras 1a y 1b muestran un ejemplo de un descriptor de punto clave de la técnica anterior;
- Las Figuras 2a y 2b muestran esquemas de subdivisión espacial usados respectivamente en las técnicas de la técnica anterior GLOH y UHoG;
- 60 - La Figura 3 muestra enfoques para el cálculo de histogramas de gradiente de la técnica UHoG;
- Las Figuras 4a y 4b muestran un ejemplo de un descriptor de punto clave de acuerdo con la invención;
- Las Figuras 5 a 7 muestran diferentes formas de manejar histogramas de gradientes correspondientes a subregiones que rodean un punto clave de imagen;
- Las Figuras 8 a 13 representan diagramas de flujo de acuerdo con seis modalidades de la presente invención;
- 65 - La Figura 14 ilustra el procesamiento de un descriptor de imágenes de acuerdo con la invención mediante el uso

del esquema de subdivisión espacial de la Figura 2b.

- La Figura 15 ilustra el procesamiento de un descriptor de imágenes de acuerdo con la invención mediante el uso del esquema de subdivisión de histograma de gradiente de la Figura 3a;
- La Figura 16 ilustra un dispositivo de procesamiento de imágenes adecuado para llevar a cabo el método de acuerdo con la presente invención.

A manera de ejemplo y sin ninguna limitación, se presentan aspectos de la presente invención para el cálculo de descriptores de imágenes robustos, discriminatorios, escalables y compactos a partir de un descriptor de imágenes, en particular un descriptor de imágenes SIFT como se ilustra en la Figura 4, con la Figura 4a ilustrando un ejemplo de subdivisión de la región local R en subregiones 4x4 SR y la Figura 4b que ilustra un ejemplo de histograma de gradiente de 8 intervalos para una subregión.

Deje h denotar el histograma de gradiente de la Figura 4b, con intervalos h_i donde i es el índice de intervalo con $i = 0 \dots n - 1$. Una transformación de histograma de gradiente de acuerdo con la invención calcula los valores v_i correspondientes a las relaciones entre los intervalos de histograma de gradiente.

Los intervalos entre los que se calculan los valores correspondientes a las relaciones pueden seleccionarse de acuerdo con su ubicación o proximidad que, en el caso del histograma de gradiente de la Figura 4b, se traduce en su separación angular.

A manera de ejemplo de los valores correspondientes a las relaciones entre los intervalos de histograma de gradiente, las diferencias de intervalos vecinos, como se muestra en (1) a continuación, capturan la forma general y las ubicaciones extremas del histograma.

$$\begin{aligned} v_0 &= h_0 - h_1 \\ v_1 &= h_1 - h_2 \\ v_2 &= h_2 - h_3 \\ v_3 &= h_3 - h_4 \\ v_4 &= h_4 - h_5 \\ v_5 &= h_5 - h_6 \\ v_6 &= h_6 - h_7 \\ v_7 &= h_7 - h_0 \end{aligned} \quad (1)$$

Las diferencias entre los intervalos que corresponden a gradientes con una diferencia angular de 90° como se muestra en (2) a continuación, o más generalmente tan cerca como 90° como lo permiten las características del histograma de gradiente y la granularidad del intervalo, capturan información sobre la relación entre gradientes a la derecha sustancialmente ángulos el uno al otro.

$$\begin{aligned} v_0 &= h_0 - h_1 \\ v_1 &= h_1 - h_3 \\ v_2 &= h_2 - h_4 \\ v_3 &= h_3 - h_5 \\ v_4 &= h_4 - h_6 \\ v_5 &= h_5 - h_7 \\ v_6 &= h_6 - h_0 \\ v_7 &= h_7 - h_1 \end{aligned} \quad (2)$$

Las diferencias entre intervalos que corresponden a gradientes con una diferencia angular de 180° como se muestra en (3) a continuación, o más generalmente tan cerca como 180° como lo permiten las características del histograma de gradiente y la granularidad de intervalo, capturan información importante de alto nivel sobre la relación entre gradientes dentro de orientaciones, tal como dentro de la orientación horizontal y dentro de la orientación vertical.

$$\begin{aligned} v_0 &= h_0 - h_4 \\ v_1 &= h_1 - h_5 \\ v_2 &= h_2 - h_6 \\ v_3 &= h_3 - h_7 \end{aligned} \quad (3)$$

Las funciones que se muestran en (4) a continuación combinan más de dos intervalos de gradiente y capturan información sobre la relación de gradientes entre orientaciones específicas, tal como los gradientes horizontales en cualquier dirección contra los gradientes verticales en cualquier dirección.

$$\begin{aligned} v_0 &= (h_0 + h_4) - (h_2 + h_6) \\ v_1 &= (h_1 + h_5) - (h_3 + h_7) \end{aligned} \quad (4)$$

Las funciones que se muestran en (5) a continuación combinan todos los intervalos de histograma de gradiente para producir valores discriminatorios que son representativos de la forma de todo el histograma.

$$\begin{aligned} v_0 &= (h_0 + h_2 + h_4 + h_6) - (h_1 + h_3 + h_5 + h_7) \\ v_0 &= (h_0 + h_1 + h_2 + h_3) - (h_4 + h_5 + h_6 + h_7) \end{aligned} \quad (5)$$

5 Los intervalos entre los que se calculan los valores correspondientes a las relaciones también se pueden seleccionar de acuerdo con diferentes criterios. Por ejemplo, los intervalos entre los que se calculan los valores correspondientes a las relaciones pueden seleccionarse para aumentar el poder discriminativo de los valores resultantes. Dichas funciones, por ejemplo, pueden tomar la forma de las funciones que se muestran en (6) a continuación

$$\begin{aligned} v_0 &= (h_0 + h_2 + h_4 + h_6) - (h_1 + h_3 + h_5 + h_7) \\ v_1 &= (h_0 + h_1 + h_2 + h_3) - (h_4 + h_5 + h_6 + h_7) \\ v_2 &= (h_0 + h_4) - (h_2 + h_6) \\ v_2 &= (h_1 + h_5) - (h_3 + h_7) \end{aligned} \quad (6)$$

15 Como otro ejemplo, los intervalos entre los que se calculan los valores correspondientes a las relaciones pueden seleccionarse a través de un proceso de selección aleatorio. Ejemplos de tales funciones se muestran en (7) a continuación

$$\begin{aligned} v_0 &= 2h_0 - h_3 - h_6 \\ v_1 &= h_0 + h_1 + h_5 - 3h_7 \end{aligned} \quad (7)$$

Como otro ejemplo de transformación de histograma, el conjunto de funciones que se muestran en (8) a continuación constituyen una transformación invertible por la cual v_7 es un valor representativo del recuento total de histograma.

$$\begin{aligned} v_0 &= h_0 - h_1 \\ v_1 &= h_1 - h_2 \\ v_2 &= h_2 - h_3 \\ v_3 &= h_3 - h_4 \\ v_4 &= h_4 - h_5 \\ v_5 &= h_5 - h_6 \\ v_6 &= h_6 - h_7 \\ v_7 &= h_0 + h_1 + h_2 + h_3 + h_4 + h_5 + h_6 + h_7 \end{aligned} \quad (8)$$

35 Las funciones de (1)-(8) muestran las relaciones entre intervalos de histograma de gradiente calculados como sumas o diferencias entre intervalos o diferencias entre sumas de intervalos, pero esto no es restrictivo y se pueden emplear diferentes operaciones lineales o no lineales, tal como relaciones entre intervalos o relaciones entre sumas de intervalos o relaciones entre productos de intervalos o diferencias entre productos de intervalos, etc.

40 En general, un histograma de gradiente puede transformarse calculando los valores v_j correspondientes a las relaciones entre los intervalos de histograma de gradiente, la selección de dichos intervalos de histograma de gradiente depende de uno o más criterios, tales como su proximidad o el poder discriminativo de los valores resultantes, o una combinación de criterios de los mismos.

45 Además, cuando los intervalos se seleccionan de acuerdo con su proximidad, un histograma de gradiente puede transformarse calculando los valores v_j correspondientes a las relaciones entre los intervalos de histograma de gradiente que exhiben una separación angular única, por ejemplo, con base en las relaciones entre los intervalos vecinos solo como en (1) o con base en relaciones entre intervalos separados 180° solo como en (3) y así sucesivamente. Alternativamente, un histograma de gradiente también puede transformarse calculando los valores v_j correspondientes a las relaciones entre los intervalos de histograma de gradiente que exhiben una pluralidad de separaciones angulares, por ejemplo, con base en algunas relaciones como se muestra en (1), algunas relaciones como se muestra en (2), y así sucesivamente. En general, el proceso de selección de transformación es bastante importante, ya que equilibra la cantidad de un tipo específico de información con la mezcla de diferentes tipos de información con la dimensionalidad k del descriptor transformado resultante en comparación con la dimensionalidad n del histograma de gradiente, que pueden estar relacionadas entre sí por $k = n$ o $k < n$ o $k > n$.

En lo sucesivo, los valores v_j con $j = 0 \dots k - 1$ calculados a partir de la transformación de un histograma de gradiente h se denominarán colectivamente como el descriptor de histograma de gradiente transformado v de elementos k .

60 Como se vio anteriormente, un único descriptor de imágenes H puede comprender múltiples histogramas de gradientes h_p , cada uno correspondiente a una subregión SR de la región local R alrededor del punto clave de imagen KP. La Figura 4a ilustra un ejemplo de subdivisión de la región local R en subregiones 4×4 SR, cada una descrita por un histograma de gradientes h_p . Tal descriptor de imágenes puede transformarse en un nuevo descriptor de imágenes V mediante la transformación individual del histograma de gradiente de cada subregión SR en un descriptor de histograma de gradiente transformado v_p como se describió anteriormente.

El proceso de transformación del histograma de gradiente descrito anteriormente tiene numerosas ventajas en la producción de un descriptor de imágenes compacto. Una primera ventaja es que, mediante la selección apropiada de la transformación, el proceso descrito anteriormente hace posible codificar características de forma sobresalientes e información discriminativa de un histograma de gradiente mediante el uso de menos elementos que el histograma de gradiente original, por ejemplo, reemplazando ocho intervalos de gradiente con los seis elementos combinados de (3) y (4) mostrados anteriormente, dan como resultado un descriptor de imágenes más pequeño que todavía tiene mucho éxito en establecer o verificar las correspondencias visuales y lograr un alto rendimiento de reconocimiento.

Otra ventaja es que, a diferencia del histograma de gradiente original, el descriptor de histograma de gradiente transformado es adecuado para la compresión mediante cuantificación escalar gruesa simple, por lo que cada elemento del descriptor se cuantifica individualmente a solo unos pocos bits, incluso a uno o dos bits, que en efecto proporciona solo una medida aproximada de la relación entre intervalos de gradiente específicos o el dominio de intervalos específicos sobre otros intervalos específicos. Sin embargo, dicha medida aproximada todavía proporciona un descriptor de imágenes compacto robusto y discriminatorio que tiene éxito en establecer o verificar las correspondencias visuales y logra un mayor rendimiento de reconocimiento que el descriptor con base en el histograma de gradiente original bajo las mismas condiciones de cuantificación escalar y también bajo esquemas de compresión más complejos tales como cuantificación vectorial y codificación de tipo.

Por ejemplo, mediante el uso de la cuantificación escalar, los valores v_j se cuantifican individualmente a niveles q , por ejemplo con $q = 2$ o $q = 3$ o $q = 4$, etc., aunque esto no es restrictivo y se pueden usar diferentes números de niveles de cuantificación para lograr la compacidad deseada en el descriptor resultante. Esta cuantificación puede realizarse comparando el valor de cada valor v_j con un conjunto de umbrales. Existen numerosas opciones para decidir los umbrales de cuantificación. Los umbrales de cuantificación pueden determinarse, por ejemplo, mediante la división uniforme del rango dinámico de cada v_j . Alternativamente, los umbrales de cuantificación pueden establecerse para lograr una distribución específica de valores a largo plazo entre los niveles de cuantificación q , por ejemplo, una distribución uniforme. Además, los umbrales de cuantificación pueden ser los mismos para todos los valores v_j calculados en todas las subregiones SR, o pueden ser los mismos para el mismo índice j en todas las subregiones SR, o pueden ser los mismos para cada subregión SR en todos los j , o pueden ser diferentes para cada subregión SR y cada índice j .

Por lo tanto, el proceso de transformación del histograma de gradiente descrito anteriormente elimina la necesidad de esquemas de compresión complejos como los vistos anteriormente, con la reducción asociada en la complejidad computacional general y los requisitos de memoria, al tiempo que se obtienen características de rendimiento similares o mejoradas.

Una tercera ventaja consecuente es que la dimensionalidad del descriptor de histograma de gradiente transformado es altamente escalable no solo antes de la compresión sino también después de la compresión, y puede reducirse aún más fácilmente si lo requiere una aplicación o las características del canal de transmisión simplemente eliminando uno o más de sus elementos cuantificados individualmente, lo que no se logra fácilmente con técnicas de compresión más complejas, como la cuantificación de vectores o la codificación de tipos.

Mientras tanto, la transformación individual y la compresión del histograma de gradiente de cada subregión como se describió anteriormente es muy ventajoso ya que facilita la escalabilidad de todo el descriptor de imágenes transformado, si es necesario, simplemente eliminando uno o más histogramas de subregión cuantificados y transformados individualmente.

Sin embargo, también debe enfatizarse que el descriptor de histograma de gradiente transformado descrito anteriormente no es incompatible con los esquemas de compresión más complejos descritos anteriormente o con cualquier esquema de compresión que sea aplicable al histograma original del descriptor basado en gradiente y aún puede usarse junto con cualquiera de estos esquemas de compresión si se desea.

Como se vio anteriormente, un solo histograma de descriptor de imágenes con base en gradiente H puede comprender múltiples histogramas de gradientes h_p , cada uno correspondiente a una subregión SR de la región local R alrededor del punto clave de imagen KP y dicho descriptor de imágenes H puede ser transformado por la transformación individual del histograma de gradiente de cada subregión SR como se describió anteriormente. Más específicamente, cada histograma de gradiente de subregión puede transformarse de manera idéntica a los otros histogramas de gradiente de subregión, pero alternativamente, al menos un histograma de gradiente de subregión puede transformarse de manera diferente a otros histogramas de gradiente de subregión.

Por ejemplo, en el contexto de la transformación de algunos histogramas de subregión de manera diferente a otros histogramas de subregión, se puede imponer un requisito en el proceso de transformación del descriptor de imágenes de que las funciones de transformación de cualquier histograma de gradientes deben diferir al menos parcialmente de las funciones de transformación empleadas en al menos algunos de sus histogramas de gradientes espacialmente vecinos. Esto se ilustra en la Figura 5.

Más específicamente, la Figura 5 se refiere a conjuntos de funciones de transformación que son total o parcialmente diferentes entre sí. Por ejemplo, tomar las funciones de (1) como A y las funciones de (2) como B da como resultado conjuntos de funciones completamente diferentes, mientras que toma las funciones de (3) y los elementos pares de (1) como A, y las funciones de (3) y los elementos impares de (1) como B dan como resultado conjuntos de funciones parcialmente diferentes. En la Figura 5, las funciones de transformación A y B se aplican a los histogramas de subregión de modo que para cada histograma de subregión transformado por A (o alternativamente B) sus cuatro histogramas vecinos horizontales y verticales se transforman por B (o alternativamente A). Si bien esta utilización de múltiples conjuntos de funciones de transformación aumenta ligeramente la complejidad de la implementación, tiene una ventaja significativa sobre la utilización de las mismas funciones de transformación en todos los histogramas de subregión. Los expertos en la materia saben que con los descriptores SIFT, así como también con otros histogramas de descriptores de imágenes con base en gradientes, los histogramas de gradiente vecinos, particularmente conectados horizontal o verticalmente, exhiben una correlación significativa, que también se mantiene en los descriptores de histograma de gradiente transformados. La aplicación de la misma transformación en tales histogramas vecinos en combinación con cualquier cuantificación gruesa posterior, da como resultado una mayor probabilidad de que los histogramas vecinos tengan elementos idénticos. Este problema se alivia mediante la utilización de diferentes funciones de transformación como se describió anteriormente, lo que aumenta la entropía y, en consecuencia, el poder discriminativo del descriptor.

Como otro ejemplo en el contexto de la transformación de algunos histogramas de subregiones de manera diferente a otros histogramas de subregiones, los descriptores de histogramas de gradiente transformados de algunas subregiones pueden comprender un número diferente de elementos a los descriptores de histogramas de gradiente transformados de otras subregiones de acuerdo con ciertos factores decisivos, tal como la ubicación de subregión relativa al centro o la región alrededor del punto clave.

La Figura 6 ilustra un ejemplo de esto, donde las cuatro subregiones centrales SR_c , que rodean directamente el punto clave de la imagen KP, se transforman como se describió anteriormente en elementos k_c , mientras que las doce subregiones límite SR_b se transforman como se describe anteriormente en elementos k_b , con $k_c > k_b$. Por lo tanto, esto aumenta el contenido de información de las subregiones centrales SR_c en relación con el descriptor general de imágenes transformadas, siendo dichas subregiones SR_c comúnmente consideradas por los expertos en la materia como más importantes para establecer o verificar las correspondencias visuales.

Como otro ejemplo en el contexto de la transformación de algunos histogramas de subregión de manera diferente a otros histogramas de subregión, los descriptores de histograma de gradiente transformado de algunas subregiones pueden comprimirse de manera diferente a los descriptores de histograma de gradiente transformado de otras subregiones de acuerdo con ciertos factores decisivos, como la ubicación relativa de la subregión al centro de la región alrededor del punto clave.

La Figura 7 ilustra un ejemplo de esto, donde las cuatro subregiones centrales SR_c , que rodean directamente el punto clave de imagen KP, se transforman como se describió anteriormente y se cuantifican a niveles de q_c , mientras que las doce subregiones límite SR_b se transforman como se describe anteriormente y se cuantifican a q_b niveles, con $q_c > q_b$. Por lo tanto, esto aumenta la precisión de representación de las subregiones centrales SR_c dentro del descriptor de imágenes transformado general, siendo dichas subregiones SR_c comúnmente consideradas por los expertos en la técnica como más importantes para establecer o verificar correspondencias visuales.

Además, las diferentes características de compresión también pueden aplicarse solo a elementos específicos de los descriptores de histograma de gradiente transformado de acuerdo con ciertos factores decisivos, como la función de transformación de los elementos específicos o sus estadísticas subyacentes.

En general, factores como la selección de las funciones de transformación apropiadas para diferentes subregiones y la selección de parámetros de compresión apropiados para diferentes subregiones y/o elementos de subregión son bastante importantes, ya que controlan el contenido de la información y el poder discriminativo del descriptor contra su complejidad y costo de almacenamiento/transmisión.

Ahora se describen modalidades y ejemplos de la presente invención con la ayuda de la Figura 8 a la Figura 13.

Primer ejemplo

Un primer ejemplo de la invención se ilustra en la Figura 8, donde un descriptor con base en histograma de gradiente H que comprende histogramas de gradiente h_p , por ejemplo, de acuerdo con la Figura 4, se procesa transformando cada histograma de gradiente h_p para reducir su dimensionalidad calculando valores correspondientes a las relaciones entre los depósitos de histograma de gradiente seleccionados de acuerdo con su proximidad.

Más específicamente, en la Figura 8, cada histograma h_p con $p = 0 \dots N-1$ ($N = 16$) se procesa a su vez. En la etapa S100, p se establece en 0.

Luego, en la etapa S110, el histograma h_p de n intervalos ($n = 8$) se transforma en descriptor v_p de elementos k ($k = 4$), con $k < n$, mediante un conjunto de funciones seleccionadas para capturar las relaciones entre intervalos que tienen una única separación angular, es decir, las funciones de (9), que capturan las relaciones entre los intervalos vecinos.

$$\begin{aligned} v_0 &= h_0 - h_1 \\ v_1 &= h_2 - h_3 \\ v_2 &= h_4 - h_5 \\ v_3 &= h_6 - h_7 \end{aligned} \quad (9)$$

Alternativamente, las funciones de (9) pueden reemplazarse por diferentes funciones que capturan las relaciones entre intervalos que tienen otra separación angular única, tales como funciones que capturan las relaciones entre intervalos separados por 90° , o funciones que dan como resultado un descriptor v_p con un número diferente de elementos k , pero aún con $k < n$.

Luego, en la etapa S120 se examina el valor de p y, si es igual al índice del último histograma en H , el proceso finaliza; de lo contrario, se incrementa en 1 en la etapa S130 y el procesamiento vuelve a la etapa S110.

Por lo tanto, el descriptor resultante V codifica características de forma sobresalientes e información discriminativa de H mediante el uso de significativamente menos elementos que H , lo que resulta en un descriptor más pequeño que todavía tiene mucho éxito en establecer o verificar correspondencias visuales y lograr un alto rendimiento de reconocimiento.

Opcionalmente, los valores de los elementos individuales de V pueden escalarse y desplazarse según corresponda, por ejemplo, para asignarlos a valores no negativos y/o a un rango dinámico específico, por ejemplo, 8 bits.

El procesamiento posterior de los descriptores resultantes, para determinar si dos descriptores se han extraído del mismo punto clave en diferentes imágenes y establecer correspondencias visuales entre imágenes, está más allá del alcance de la invención y puede proceder de manera análoga a la de los descriptores originales de histograma de gradiente, por ejemplo, como se describe en el documento de David G. Lowe, "Distinctive image features from scale-invariant keypoints", Revista internacional de visión por computador, 60, 2 (2004), págs. 91-110, o en cualquier otra manera.

Primera modalidad

Una primera modalidad de la invención se ilustra en la Figura 9, donde un descriptor con base en histograma de gradiente H que comprende histogramas de gradiente h_p , por ejemplo, de acuerdo con la Figura 4, se procesa transformando cada histograma de gradiente h_p para reducir su dimensionalidad calculando valores correspondientes a las relaciones entre los intervalos de histograma de gradiente seleccionados de acuerdo con una pluralidad de criterios con respecto a su proximidad.

Más específicamente, en la Figura 9, cada histograma h_p con $p = 0 \dots N-1$ ($N = 16$) se procesa a su vez. En la etapa S200, p se establece en 0.

Luego, en la etapa S210, el histograma h_p de n intervalos ($n = 8$) se transforma en descriptor v_p de elementos k ($k = 6$), con $k < n$, mediante un conjunto de funciones seleccionadas para capturar las relaciones entre intervalos que tienen una pluralidad de separaciones angulares, específicamente, las funciones de (10), que capturan las relaciones entre intervalos vecinos y las relaciones entre intervalos separados por 180° .

$$\begin{aligned} v_0 &= h_0 - h_1 \\ v_1 &= h_2 - h_3 \\ v_2 &= h_4 - h_5 \\ v_3 &= h_6 - h_7 \\ v_4 &= h_0 - h_4 \\ v_5 &= h_2 - h_6 \end{aligned} \quad (10)$$

Alternativamente, las funciones de (10) pueden ser reemplazadas por diferentes funciones que capturan las relaciones entre intervalos que tienen otra pluralidad de separaciones angulares, tales como funciones que capturan las relaciones entre intervalos separados por 90° y las relaciones entre intervalos separados por 180° , o funciones que resultan en un descriptor v_p con un número diferente de elementos k , pero aún con $k < n$.

Luego, en la etapa S220 se examina el valor de p y, si es igual al índice del último histograma en H , el proceso finaliza; de lo contrario, se incrementa en 1 en la etapa S230 y el procesamiento vuelve a la etapa S210.

Por lo tanto, el descriptor resultante V codifica diferentes tipos de características de forma sobresaliente e información discriminativa de H mediante el uso de menos elementos que H , lo que resulta en un descriptor más

pequeño que todavía tiene mucho éxito en establecer o verificar correspondencias visuales y lograr un alto rendimiento de reconocimiento.

El procesamiento posterior del descriptor resultante puede proceder de manera análoga al primer ejemplo.

Segunda modalidad

Una segunda modalidad de la invención se ilustra en la Figura 10, donde un

descriptor con base en histograma de gradiente H que comprende histogramas de gradiente h_p , por ejemplo, de acuerdo con la Figura 4, se procesa transformando cada histograma de gradiente h_p calculando los valores correspondientes a las relaciones entre los intervalos de histograma de gradiente seleccionados de acuerdo con una pluralidad de criterios con respecto a su proximidad.

Más específicamente, en la Figura 10, cada histograma h_p con $p = 0 \dots N-1$ ($N = 16$) se procesa a su vez. En la etapa S300, p se establece en 0.

Luego, en la etapa S310, el histograma h_p de n intervalos ($n = 8$) se transforma en descriptor v_p de elementos k ($k = 8$), mediante un conjunto de funciones seleccionadas para capturar las relaciones entre intervalos que tienen una pluralidad de separaciones angulares, específicamente, las funciones de (11).

$$\begin{aligned} v_0 &= h_2 - h_6 \\ v_1 &= h_3 - h_7 \\ v_2 &= h_0 - h_1 \\ v_3 &= h_2 - h_3 \\ v_4 &= h_4 - h_5 \\ v_5 &= h_6 - h_7 \\ v_6 &= (h_0 + h_4) - (h_2 + h_6) \\ v_7 &= (h_0 + h_2 + h_4 + h_6) - (h_1 + h_3 + h_5 + h_7) \end{aligned} \quad (11)$$

Alternativamente, las funciones de (11) pueden ser reemplazadas por diferentes funciones que capturan las relaciones entre intervalos que tienen otra pluralidad de separaciones angulares, o funciones que resultan en un descriptor v_p con un número diferente de elementos k , con $k = n$ o $k < n$ o $k > n$.

Luego, en la etapa S320, cada elemento del descriptor v_p de elementos k se cuantifica individualmente a q niveles, dando el descriptor cuantificado $\tilde{v}_p$. En esta modalidad, establecemos $q = 3$, aunque esto no es restrictivo y pueden usarse diferentes números de niveles de cuantificación para lograr la compacidad deseada en el descriptor resultante, por ejemplo, $q = 2$, $q = 4$, y así sucesivamente. Esta cuantificación se puede realizar comparando el valor de cada elemento v_{pj} con un conjunto de umbrales. Existen numerosas opciones para decidir los umbrales de cuantificación. Los umbrales de cuantificación pueden determinarse, por ejemplo, mediante la división uniforme del rango dinámico de v_{pj} . Alternativamente, los umbrales de cuantificación pueden establecerse para lograr una distribución específica de valores a largo plazo entre los niveles de cuantificación q , por ejemplo, una distribución uniforme. Además, los umbrales de cuantificación pueden ser los mismos para todos los elementos v_{pj} , o pueden ser los mismos para el mismo índice j en todos los p , o pueden ser los mismos para el mismo índice p en todos los j , o pueden ser diferentes para cada v_{pj} .

Luego, en la etapa S330 se examina el valor de p y, si es igual al índice del último histograma en H, el proceso finaliza; de lo contrario, se incrementa en 1 en la etapa S340 y el procesamiento vuelve a la etapa S310.

Claramente, las implementaciones alternativas pueden cambiar el orden o ciertas operaciones según sea apropiado en comparación con la Figura 10 sin desviarse del alcance de la invención, por ejemplo calculando primero el descriptor V completo y luego procediendo con la cuantificación para generar el descriptor cuantificado $\tilde{V}$.

Por lo tanto, el descriptor resultante $\tilde{V}$, codifica diferentes tipos de características de forma sobresalientes e información discriminativa de H. La cuantificación escalar gruesa, mediante la cual cada elemento descriptor se cuantifica individualmente a solo unos pocos niveles que proporcionan solo una medida aproximada de la relación entre los intervalos de gradiente específicos o el dominio de los intervalos específicos sobre otros intervalos específicos, resulta en un descriptor de imágenes compacto robusto y discriminatorio que tiene éxito en establecer o verificar correspondencias visuales y logra un mayor rendimiento de reconocimiento que el descriptor con base en histograma de gradiente original bajo las mismas condiciones de cuantificación escalar y también bajo esquemas de compresión más complejos.

El procesamiento posterior del descriptor resultante puede entonces proceder de manera análoga a la modalidad previa.

Tercera modalidad

Una tercera modalidad de la invención se ilustra en la Figura 11, donde un descriptor con base en histograma de gradiente H que comprende histogramas de gradiente h_p , por ejemplo de acuerdo con la Figura 4, se procesa transformando cada histograma de gradiente h_p con un conjunto de funciones de transformación que exhibe diferencias a los conjuntos de funciones de transformación utilizadas en el procesamiento de algunos de los histogramas de gradiente que son vecinos h_p .

Más específicamente, en la Figura 11, cada histograma h_p con $p = 0 \dots N-1$ ($N = 16$) se procesa a su vez. En la etapa S400, p se establece en 0.

Luego, en la etapa S410, se examina el valor de p para la selección de las funciones apropiadas para el procesamiento de h_p . Esta selección puede, por ejemplo, proceder a lo largo de las líneas ilustradas en la Figura 5, donde se usan dos conjuntos de transformaciones, designadas A y B, con A seleccionada cuando $p = 0, 2, 5, 7, 8, 10, 13, 15$ y B seleccionados cuando $p = 1, 3, 4, 6, 9, 11, 12, 14$. Sin embargo, esto no es restrictivo, y se puede emplear un número diferente de conjuntos de funciones de transformación, y con una disposición espacial diferente, de modo que el conjunto de funciones de transformación utilizado en el procesamiento de un histograma de gradiente muestre diferencias con respecto a los conjuntos de transformación. funciones usadas en el procesamiento de algunos de los histogramas de gradiente que son vecinos a dicho histograma.

Luego, en la etapa S420, el histograma h_p de n intervalos ($n = 8$) se transforma en el descriptor v_p de elementos k ($k = 8$), mediante el uso del conjunto de funciones de transformación adecuadamente seleccionado. Los conjuntos de funciones de transformación A y B se pueden definir según (12) y (13), respectivamente, que se muestran a continuación, donde cada conjunto de funciones se selecciona para capturar las relaciones entre los intervalos que tienen una pluralidad de separaciones angulares, y no hay funciones entre A y B.

$$\begin{aligned} v_0 &= h_2 - h_6 \\ v_1 &= h_3 - h_7 \\ v_2 &= h_0 - h_1 \\ v_3 &= h_2 - h_3 \\ v_4 &= h_4 - h_5 \\ v_5 &= h_6 - h_7 \\ v_6 &= (h_0 + h_4) - (h_2 + h_6) \\ v_7 &= (h_0 + h_2 + h_4 + h_6) - (h_1 + h_3 + h_5 + h_7) \end{aligned} \quad (12)$$

$$\begin{aligned} v_0 &= h_0 - h_4 \\ v_1 &= h_1 - h_5 \\ v_2 &= h_1 - h_2 \\ v_3 &= h_3 - h_4 \\ v_4 &= h_5 - h_6 \\ v_5 &= h_7 - h_0 \\ v_6 &= (h_1 + h_5) - (h_3 + h_7) \\ v_7 &= (h_0 + h_1 + h_2 + h_3) - (h_4 + h_5 + h_6 + h_7) \end{aligned} \quad (13)$$

Alternativamente, uno o ambos conjuntos de funciones como se muestra en (12) y (13) pueden ser reemplazados por diferentes funciones que capturan las relaciones entre intervalos que tienen otra pluralidad de separaciones angulares, o funciones que dan como resultado un descriptor con un número diferente de elementos k , con $k = n$ o $k < n$ o $k > n$. Además, los conjuntos de funciones A y B pueden no contener funciones comunes, o pueden contener algunas funciones comunes. Además, los intervalos entre los que se calculan los valores correspondientes a las relaciones pueden seleccionarse de acuerdo con diferentes criterios para uno o ambos conjuntos de funciones A y B. Dichos criterios, como se vio anteriormente, pueden incluir la selección de los intervalos entre los cuales los valores correspondientes a las relaciones deben calcularse para aumentar el poder discriminativo de los valores resultantes, o seleccionar los intervalos entre los cuales los valores correspondientes a las relaciones deben calcularse mediante un proceso de selección aleatorio. Además, cualquier mezcla de dichos criterios también se puede usar en la selección de uno o ambos conjuntos de funciones A y B.

Luego, en la etapa S430, cada elemento del descriptor v_p de elementos k se cuantifica individualmente a q niveles, dando el descriptor cuantificado $\tilde{v}$, p , de manera análoga a la tercera modalidad.

Luego, en la etapa S440 se examina el valor de p y, si es igual al índice del último histograma en H, el proceso finaliza; de lo contrario, se incrementa en 1 en la etapa S450 y el procesamiento vuelve a la etapa S410.

Claramente, implementaciones alternativas pueden cambiar el orden o ciertas operaciones según sea apropiado en comparación con la Figura 11 sin desviarse del alcance de la invención, por ejemplo, calculando primero el descriptor V completo y luego procediendo con la cuantificación para generar el descriptor cuantificado $\tilde{V}$.

Por lo tanto, el descriptor resultante $\tilde{V}$, codifica información discriminativa de H. La utilización de múltiples conjuntos

de funciones de transformación según esta modalidad tiene una ventaja significativa sobre la utilización de las mismas funciones de transformación en todos los histogramas de subregión. Los expertos en la materia saben que con el histograma de descriptores de imágenes basados en gradiente, los histogramas de gradiente vecinos exhiben una correlación significativa, que se mantiene en los descriptores de histograma de gradiente transformados y, en combinación con la cuantificación escalar gruesa, da como resultado una mayor probabilidad de vecino histogramas transformados que tienen elementos idénticos. Este problema se alivia mediante la utilización de diferentes funciones de transformación según esta modalidad, lo que aumenta la entropía y, en consecuencia, el poder discriminativo del descriptor.

El procesamiento posterior del descriptor resultante puede proceder de manera análoga a las modalidades anteriores.

Cuarta Modalidad

Una cuarta modalidad de la invención se ilustra en la Figura 12, donde un descriptor basado en histograma de gradiente H que comprende histogramas de gradiente h_p , por ejemplo de acuerdo con la Figura 4, se procesa transformando cada histograma de gradiente h_p calculando valores correspondientes a las relaciones entre los intervalos de histograma de gradiente y para que los descriptores de histograma de gradiente transformados de algunas subregiones comprendan un número diferente de elementos que los descriptores de histograma de gradiente transformados de otras subregiones.

Más específicamente, en la Figura 12, cada histograma h_p con $p = 0 \dots N-1$ ($N = 16$) se procesa a su vez. En la etapa S500, p se establece en 0.

Luego, en la etapa S510, se examina el valor de p para la selección de la dimensionalidad apropiada para el procesamiento de h_p . Esta selección puede, por ejemplo, proceder a lo largo de las líneas ilustradas en la Figura 6, donde la subregión central de histogramas con $p = 5, 6, 9, 10$ son para transformarse de manera que los descriptores resultantes tienen elementos k_c cada ($k_c = 3$) y los histogramas de subregión límite con $p = 0, 1, 2, 3, 4, 7, 8, 11, 12, 13, 14, 15$ deben transformarse de modo que los descriptores resultantes tengan elementos k_b cada uno ($k_b = 2$). Sin embargo, esto no es restrictivo, y puede emplearse un número diferente de dimensionalidades posibles, y con una disposición espacial diferente.

Luego, en la etapa S520, el histograma h_p de n intervalos ($n = 8$) se transforma en el descriptor v_p calculando los valores correspondientes a las relaciones entre los intervalos de histograma de gradiente de acuerdo con las funciones de (14) que se muestran a continuación

$$\begin{aligned} v_0 &= h_0 - h_4 \text{ (usado para subregiones centrales y de límite)} \\ v_1 &= h_1 - h_5 \text{ (usado para subregiones centrales y de límite)} \\ v_2 &= (h_1 + h_5) - (h_3 + h_7) \text{ (usado solo para subregiones centrales)} \end{aligned} \quad (14)$$

Claramente, esta modalidad puede combinarse con cualquier modalidad anterior, y las funciones de transformación de (14) pueden reemplazarse por funciones de transformación seleccionadas de acuerdo con cualquier combinación de criterios de selección como se vio anteriormente y/o pueden usarse diferentes conjuntos de funciones de transformación para diferentes histogramas de subregiones.

Luego, en la etapa S530, cada elemento del descriptor v_p de elementos k se cuantifica individualmente a q niveles, dando el descriptor cuantificado $\tilde{v}_p$, de manera análoga a la tercera y cuarta modalidad.

Luego, en la etapa S540, se examina el valor de p y, si es igual al índice del último histograma en H, el proceso finaliza; de lo contrario, se incrementa en 1 en la etapa S550 y el procesamiento vuelve a la etapa S510.

Claramente, implementaciones alternativas pueden cambiar el orden o ciertas operaciones según sea apropiado en comparación con la Figura 12 sin desviarse del alcance de la invención, por ejemplo, calculando primero el descriptor V completo y luego procediendo con la cuantificación para generar el descriptor cuantificado $\tilde{V}$.

Por lo tanto, el descriptor resultante $\tilde{V}$, codifica información discriminativa de H de una manera que da más importancia a aquellas partes de H que pueden ser más importantes para establecer o verificar las correspondencias visuales, tales como los histogramas de subregión central versus los histogramas de subregión límite, al permitir una representación de mayor dimensionalidad para esas partes.

Quinta modalidad

Una quinta modalidad de la invención se ilustra en la Figura 13, donde un descriptor basado en histograma de gradiente H que comprende histogramas de gradiente h_p , por ejemplo de acuerdo con la Figura 4, se procesa transformando cada histograma de gradiente h_p calculando valores correspondientes a las relaciones entre los intervalos de histograma de gradiente y para que algunos elementos descriptores de histograma de gradiente

transformados se cuantifiquen a un número diferente de niveles de otros elementos descriptores de histograma de gradiente transformados.

Más específicamente, en la Figura 13, cada histograma h_p con $p = 0 \dots N-1$ ($N = 16$) se procesa a su vez. En la etapa S600, p se establece en 0.

Luego, en la etapa S610, el histograma h_p de n intervalos ($n = 8$) se transforma en descriptor v_p de elementos k ($k = 8$) mediante un conjunto de funciones seleccionadas para capturar las relaciones entre intervalos como se muestra en (15) a continuación.

$$\begin{aligned} v_0 &= h_2 - h_6 \\ v_1 &= h_3 - h_7 \\ v_2 &= h_0 - h_1 \\ v_3 &= h_2 - h_3 \\ v_4 &= h_4 - h_5 \\ v_5 &= h_6 - h_7 \\ v_6 &= (h_0 + h_4) - (h_2 + h_6) \\ v_7 &= (h_0 + h_2 + h_4 + h_6) - (h_1 + h_3 + h_5 + h_7) \end{aligned} \quad (15)$$

Claramente, esta modalidad puede combinarse con cualquier modalidad anterior, y las funciones de transformación de (15) pueden reemplazarse por funciones de transformación seleccionadas de acuerdo con cualquier combinación de criterios de selección como se vio anteriormente y/o pueden usarse diferentes conjuntos de funciones de transformación para diferentes los histogramas de subregión y/o las diferentes dimensiones del descriptor transformado pueden usarse para diferentes histogramas de subregión.

Luego, en la etapa S620, se examina el valor de p para la selección del número apropiado de niveles de cuantificación para la cuantificación de cada elemento v_{pj} . Esta selección puede, por ejemplo, proceder a lo largo de las líneas ilustradas en la Figura 7, donde los descriptores subregión centrales con $p = 5, 6, 9, 10$ han de ser cuantificados a los niveles q_c ($q_c = 4$) y los descriptores subregión de contorno con $p = 0, 1, 2, 3, 4, 7, 8, 11, 12, 13, 14, 15$ han de ser cuantificados a los niveles q_b ($q_b = 2$). Esto, sin embargo, no es restrictivo, y puede emplearse un número diferente de niveles de cuantificación posibles, y con una disposición espacial diferente. Además, aunque esto no se muestra en la Figura 13, el número de niveles de cuantificación para cada elemento v_{pj} puede determinarse de acuerdo con el valor de j , es decir, de acuerdo con el tipo específico de elemento descriptor, en lugar de o además del valor de p , es decir, la ubicación de subregión del elemento.

Luego, en la etapa S630, cada elemento del descriptor v_p de elementos k se cuantifica individualmente al número apropiado de niveles de cuantificación, dando el descriptor cuantificado $\tilde{v}$, p , de manera análoga a la tercera, cuarta y quinta modalidad.

Luego, en la etapa S640 se examina el valor de p y, si es igual al índice del último histograma en H , el proceso finaliza; de lo contrario, se incrementa en 1 en la etapa S650 y el procesamiento vuelve a la etapa S610.

Claramente, implementaciones alternativas pueden cambiar el orden o ciertas operaciones según sea apropiado en comparación con la Figura 13 sin desviarse del alcance de la invención, por ejemplo, calculando primero el descriptor V completo y luego procediendo con la cuantificación para generar el descriptor cuantificado $\tilde{V}$.

Por lo tanto, el descriptor resultante $\tilde{V}$, codifica información discriminativa de H de una manera que da mayor precisión de representación a aquellas partes de H o aquellos elementos de V que pueden ser más importantes para establecer o verificar las correspondencias visuales.

Aunque los aspectos y las modalidades de la presente invención se presentan en detalle para el cálculo de descriptores de imágenes robustos, discriminativos, escalables y compactos a partir de un descriptor de imágenes SIFT como se ilustra en la Figura 4, la invención es aplicable a otros descriptores de imágenes con base en histogramas de gradientes. Por ejemplo, la Figura 14 ilustra el procesamiento de un descriptor basado en histograma de gradiente mediante el uso de una subdivisión espacial logarítmica polar mediante la transformación de cada uno de sus histogramas de gradiente con un conjunto de funciones de transformación que exhibe diferencias con los conjuntos de funciones de transformación usadas en el procesamiento de su vecino histogramas de gradiente, de manera análoga a la cuarta modalidad de la invención y con los símbolos A, B y C en la Figura 14 correspondientes a dichos conjuntos de funciones de transformación. Como otro ejemplo, para un histograma de gradientes con base en una subdivisión del espacio bidimensional de los componentes x e y de los gradientes en intervalos como se ilustra en la Figura 15, un conjunto adecuado de funciones de transformación de una manera análoga a las modalidades anteriores de la invención es

$$\begin{aligned} v_0 &= h_0 - h_1 \\ v_1 &= h_1 - h_2 \\ v_2 &= h_2 - h_3 \end{aligned} \quad (16)$$

$$\begin{aligned}
v_3 &= h_3 - h_4 \\
v_4 &= h_4 - h_5 \\
v_5 &= h_5 - h_6 \\
v_6 &= h_6 - h_7 \\
v_7 &= h_7 - h_0 \\
v_8 &= h_8 - (h_0 + h_1 + h_2 + h_3 + h_4 + h_5 + h_6 + h_7)
\end{aligned}$$

Además, la invención también es aplicable a histograma procesado adecuadamente de descriptores de imágenes basados en gradiente. Tal procesamiento adecuado puede implicar, por ejemplo, la combinación de histogramas de gradiente de subregión antes del procesamiento de acuerdo con la invención.

Solo con fines ilustrativos, refiriéndose a la Figura 4a, para el descriptor de histograma de gradiente H que comprende histogramas de subregión h_p con $p = 0 \dots 15$, los histogramas con $p = 0,1,4,5$ pueden combinarse en un solo histograma de subregión mediante el promedio de sus valores de intervalo, y se puede realizar una combinación análoga para los histogramas con $p = 2,3,6,7$ y $p = 8,9,12,13$ y $p = 10,11,14,15$, lo que da como resultado un descriptor de histograma de gradiente con dimensionalidad reducida que luego puede procesarse de acuerdo con cualquier modalidad previa de la invención. Los histogramas de gradiente también se pueden combinar mediante una función alternativa de sus valores de intervalo, tal como la suma.

Alternativamente, o además, dicho procesamiento adecuado puede implicar, por ejemplo, la fusión de intervalos dentro de histogramas de gradiente de subregión. Con fines ilustrativos, refiriéndose a la Figura 4b, para el histograma de gradiente h, los intervalos vecinos pueden fusionarse en un único intervalo mediante el promedio o la adición o la mediana o cualquier función adecuada, lo que da como resultado un descriptor de histograma de gradiente con dimensionalidad reducida que luego puede procesarse de acuerdo con cualquier modalidad previa de la invención.

Solo a manera de ejemplo, la Figura 16 ilustra un aparato de procesamiento conceptual para llevar a cabo el método de acuerdo con la presente invención. Más específicamente, el aparato de procesamiento 1100 recibe entrada, que puede comprender datos visuales, tales como datos de imagen o video, descriptores precalculados con base en histogramas de gradientes, descriptores compactos precalculados de acuerdo con el método de la presente invención, instrucciones de programación o entrada del usuario, desde el aparato de entrada 1000, que puede tomar la forma de un aparato de entrada del usuario, un lector de medios o un receptor de señales transmitidas. El aparato de procesamiento 1100 comprende los principales bloques de procesamiento de una unidad central de procesamiento 1110 que controla las operaciones de los otros bloques de procesamiento, memoria volátil 1120, memoria no volátil 1130, opcionalmente un bloque extractor de descriptor 1140 configurado para generar descriptores con base en histogramas de gradientes, un bloque extractor de descriptor compacto 1150 configurado para llevar a cabo el método de acuerdo con la presente invención, y opcionalmente un bloque procesador de descriptor compacto 1160 configurado para procesar dichos descriptores compactos, por ejemplo para establecer o verificar correspondencias visuales. El aparato de procesamiento 1100 está conectado al aparato de salida 1900, que puede tomar la forma de una unidad de visualización, un escritor de medios o un transmisor de señales, que proporciona una salida que puede comprender datos visuales anotados, tales como datos de imagen o video, información de procesamiento tales como correspondencias visuales establecidas o verificadas, o descriptores compactos calculados de acuerdo con el método de la presente invención. Debe entenderse que los bloques de procesamiento y la arquitectura mostrados en la Figura 16 son solo conceptuales y pueden no corresponder exactamente a todos los aparatos que implementan el método de acuerdo con la invención.

Por lo tanto, se puede entender fácilmente que la presente invención no se limita a un método para transformar un descriptor de imágenes con base en un histograma de gradiente y un aparato de procesamiento de imagen relativo, sino que puede estar sujeto a muchas modificaciones, mejoras o reemplazos de partes y elementos equivalentes sin partiendo de la idea inventiva, como se especifica claramente en las siguientes reivindicaciones.

REIVINDICACIONES

1. Método para transformar un descriptor de imágenes (H), con base en una pluralidad de histogramas de gradiente (h), cada uno de los cuales comprende una pluralidad de intervalos de histograma (h_i), en un descriptor de imágenes transformado (V) con base en una pluralidad de histogramas de gradiente transformados (v), cada uno de los cuales comprende un conjunto de valores (v_j), en donde, para transformar al menos un histograma de gradiente (h) de dicha pluralidad de histogramas de gradiente en al menos un histograma de gradiente transformado (v) correspondiente, se proporcionan las etapas de:
para cada valor del conjunto de valores (v_i) del al menos un histograma de gradiente transformado (v) correspondiente, seleccionar al menos dos intervalos de histograma de la pluralidad de intervalos de histograma (h_i) del al menos un histograma de gradiente con base en al menos un criterio de ubicación de dicha pluralidad de intervalos de histograma (h_i) y calcular, como el valor, una función de los al menos dos intervalos de histograma seleccionados,
en donde la pluralidad de histogramas de gradiente comprende una pluralidad de histogramas de orientación de gradiente y los intervalos de histograma seleccionados para calcular el conjunto de valores (v_j) del al menos un histograma de gradiente transformado correspondiente exhiben una pluralidad de separaciones angulares entre los al menos dos intervalos seleccionados para cada valor.
2. Método de acuerdo con la reivindicación 1, en donde al menos uno de dichos valores (v_j) se calcula entre los intervalos de histograma adyacentes (h_i).
3. Método de acuerdo con la reivindicación 1 o 2, en donde dichos valores (v_j) se cuantifican individualmente en una pluralidad de niveles (q), en donde dicha pluralidad de niveles puede ser igual para cada valor (v_j) o puede ser diferente para al menos dos valores (v_j).
4. Método de acuerdo con una de las reivindicaciones de la 1 a la 3, en donde dichos valores (v_j) se calculan de acuerdo con una o más de las siguientes operaciones: sumas o diferencias entre los intervalos de histograma (h_i); diferencias entre las sumas de los intervalos de histograma (h_i); operaciones lineales o no lineales entre los intervalos de histograma (h_i); relaciones entre los intervalos de histograma (h_i); relaciones entre las sumas de los intervalos de histograma (h_i); relaciones entre los productos de los intervalos de histograma (h_i); diferencias entre los productos de los intervalos de histograma (h_i).
5. Método de acuerdo con cualquiera de las reivindicaciones de la 1 a la 4, en donde dicho histograma de gradiente (h) tiene una dimensionalidad de un primer número (n) de elementos, dicho conjunto de valores (v_j) tiene una dimensionalidad de un segundo número (k) de elementos, y dicho primer número (n) de elementos es igual a dicho segundo número (k) de elementos.
6. Método de acuerdo con cualquiera de las reivindicaciones de la 1 a la 4, en donde dicho histograma de gradiente (h) tiene una dimensionalidad de un primer número (n) de elementos, dicho conjunto de valores (v_j) tiene una dimensionalidad de un segundo número (k) de elementos, y dicho primer número (n) de elementos es mayor que dicho segundo número (k) de elementos.
7. Método de acuerdo con cualquiera de las reivindicaciones de la 1 a la 4, en donde dicho histograma de gradiente (h) tiene una dimensionalidad de un primer número (n) de elementos, dicho conjunto de valores (v_j) tiene una dimensionalidad de un segundo número (k) de elementos, y dicho primer número (n) de elementos es menor que dicho segundo número (k) de elementos.
8. Método de acuerdo con cualquiera de las reivindicaciones de la 1 a la 7, en donde cada uno de dicha pluralidad de histogramas de gradiente (h) está relacionado con una subregión (SR) de una región local (R) alrededor de un punto clave de imagen (KP).
9. Método de acuerdo con la reivindicación 8, en donde al menos un histograma de gradiente de subregión se transforma de manera diferente a otros histogramas de gradiente de subregión.
10. Método de acuerdo con la reivindicación 8, en donde una pluralidad de histogramas de gradiente de subregión se combina en un solo histograma de subregión, en particular al sumar o promediar sus valores de intervalo.
11. Método de acuerdo con la reivindicación 8, en donde las operaciones empleadas para obtener dicho al menos un histograma de gradiente de subregión transformado difieren al menos parcialmente de las operaciones empleadas para obtener los histogramas de gradiente de subregión transformados de al menos una de sus subregiones vecinas.
12. Método de acuerdo con la reivindicación 8, en donde dicho al menos un histograma de gradiente de subregión transformado comprende un número diferente de elementos que los histogramas de gradiente de subregión transformados de al menos una de sus subregiones vecinas.

13. Método de acuerdo con la reivindicación 12, en donde los histogramas de gradiente de subregión transformados en relación con las subregiones (SR_c) que rodean directamente dicho punto clave de imagen (KP) comprenden más elementos (k_c) que los elementos (k_b) de los histogramas de gradiente transformados de las subregiones restantes (SR_b).
5
14. Método de acuerdo con la reivindicación 13, en donde los elementos de dichos histogramas de gradiente transformados de subregiones (SR_c) que rodean directamente dicho punto clave de imagen (KP) se cuantifican con un número de niveles (q_c) que es mayor que el número de niveles (q_b) con el que se cuantifican los histogramas de gradiente transformados de las subregiones restantes (SR_b).
10
15. Método de acuerdo con la reivindicación 14, en donde dicha cuantificación se aplica solo a elementos específicos de dichos histogramas de gradiente de subregión.
- 15 16. Método de acuerdo con la reivindicación 1, en donde dicho descriptor de imágenes transformado (V) está sujeto a un proceso de cuantificación con el fin de generar un descriptor de imágenes transformado cuantificado ($\tilde{V}$).
20
17. Aparato de procesamiento de imágenes que comprende medios para llevar a cabo las reivindicaciones de la 1 a la 16.

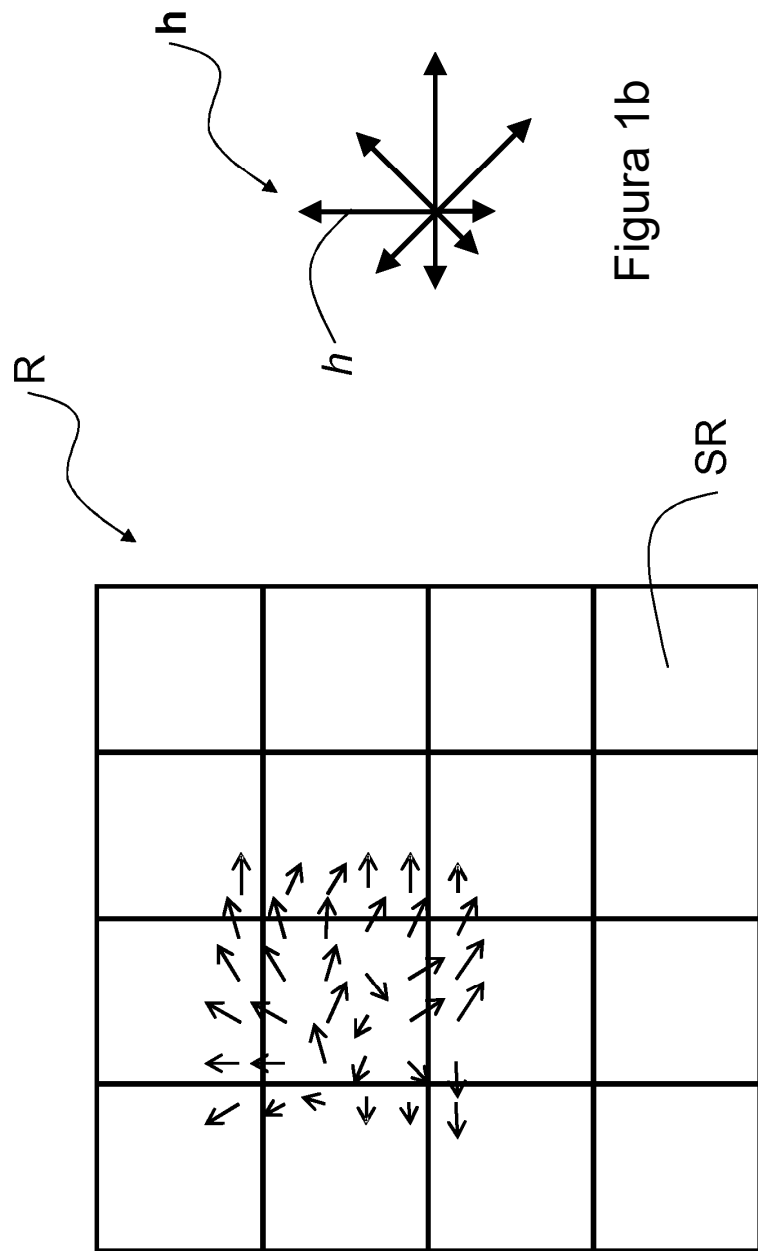


Figura 1a

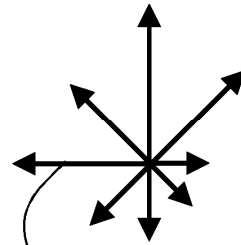


Figura 1b

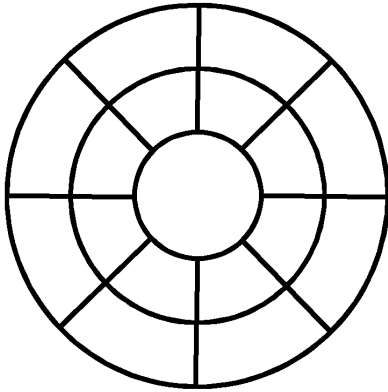


Figura 2a

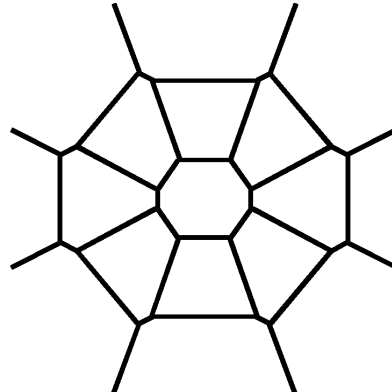


Figura 2b

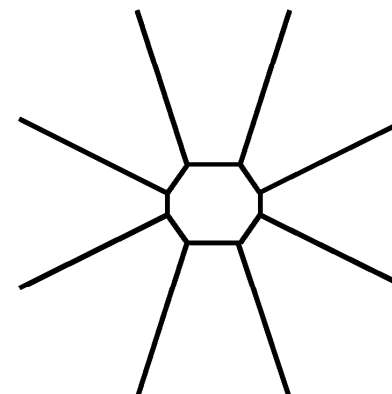


Figura 3a

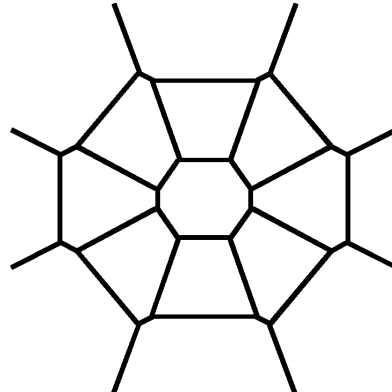


Figura 3b

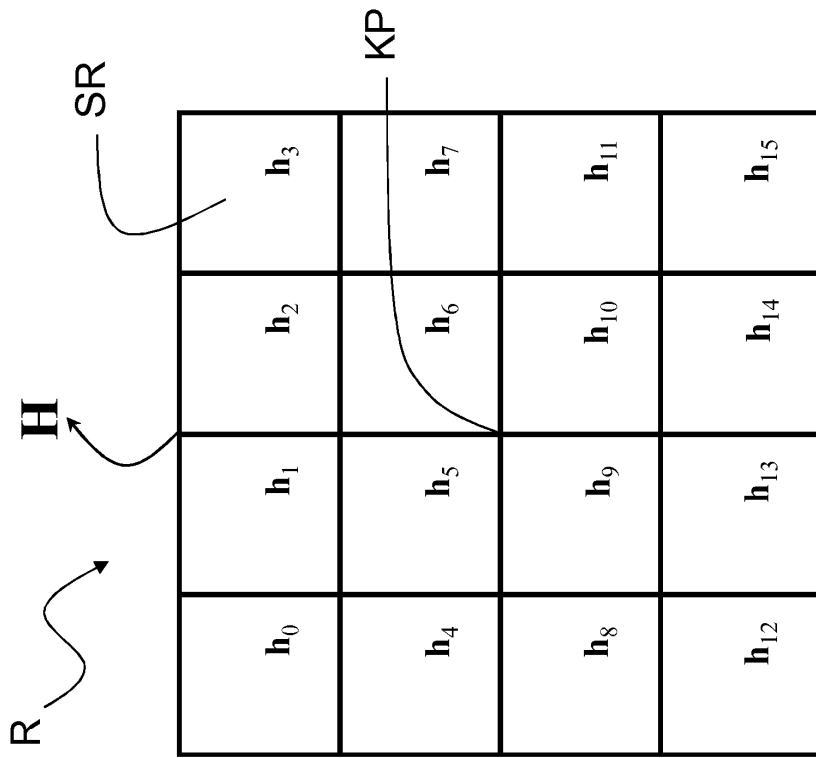


Figura 4a

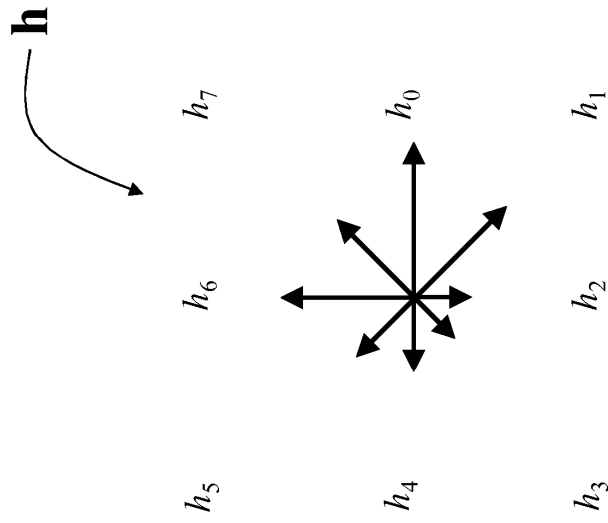


Figura 4b

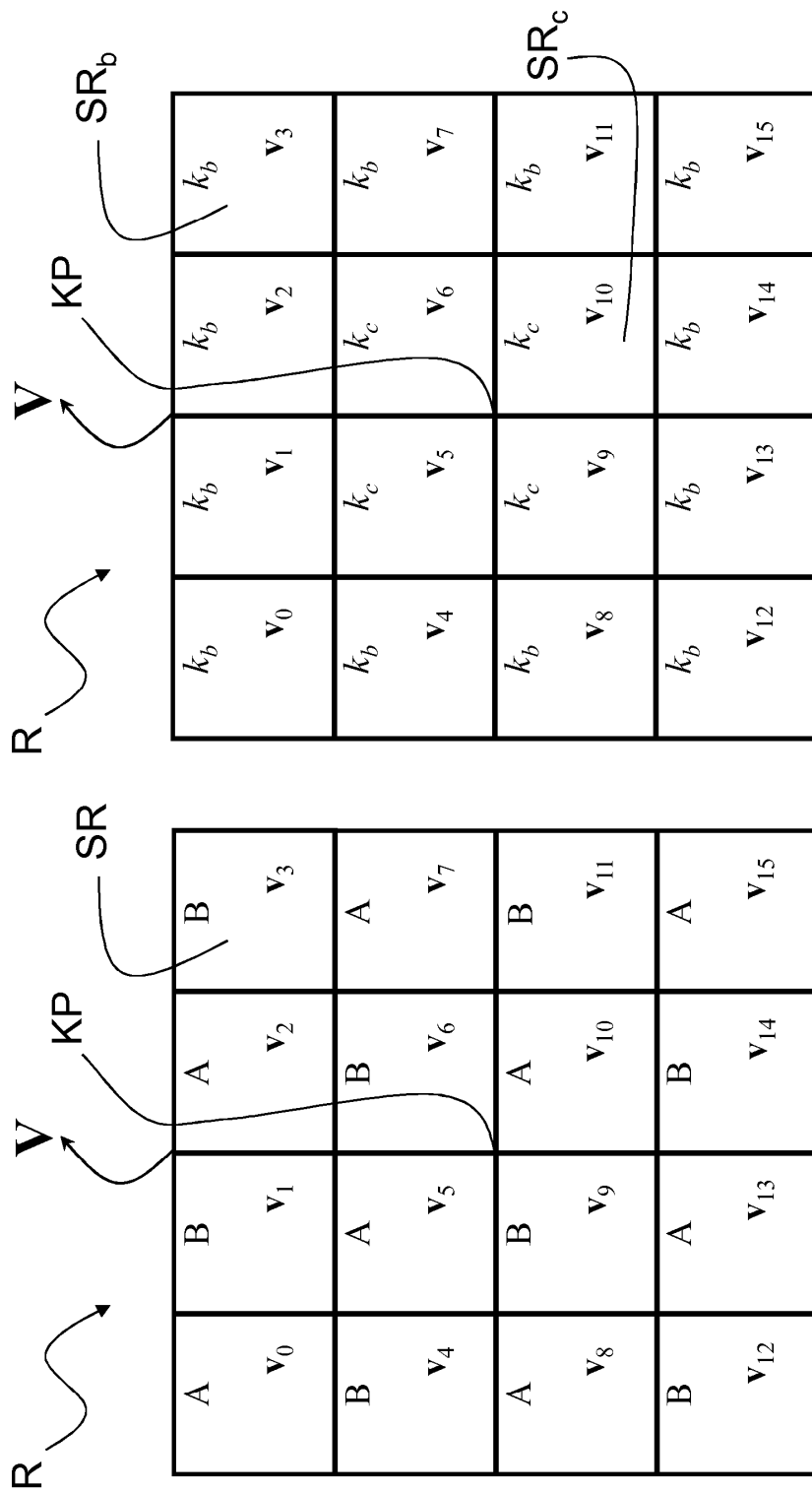


Figura 5

Figura 6

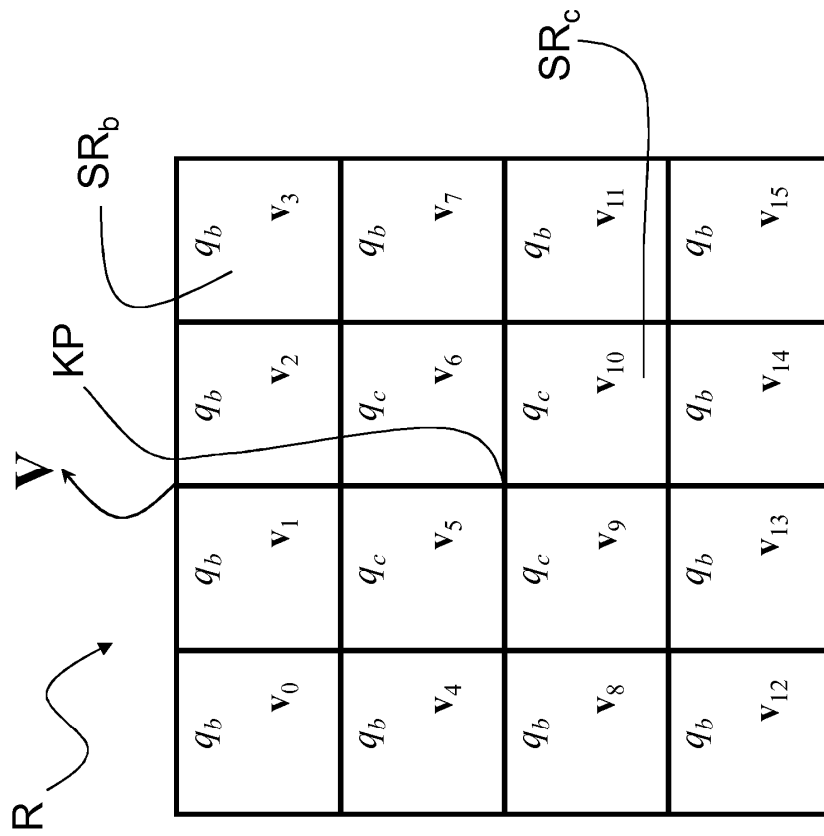


Figura 7

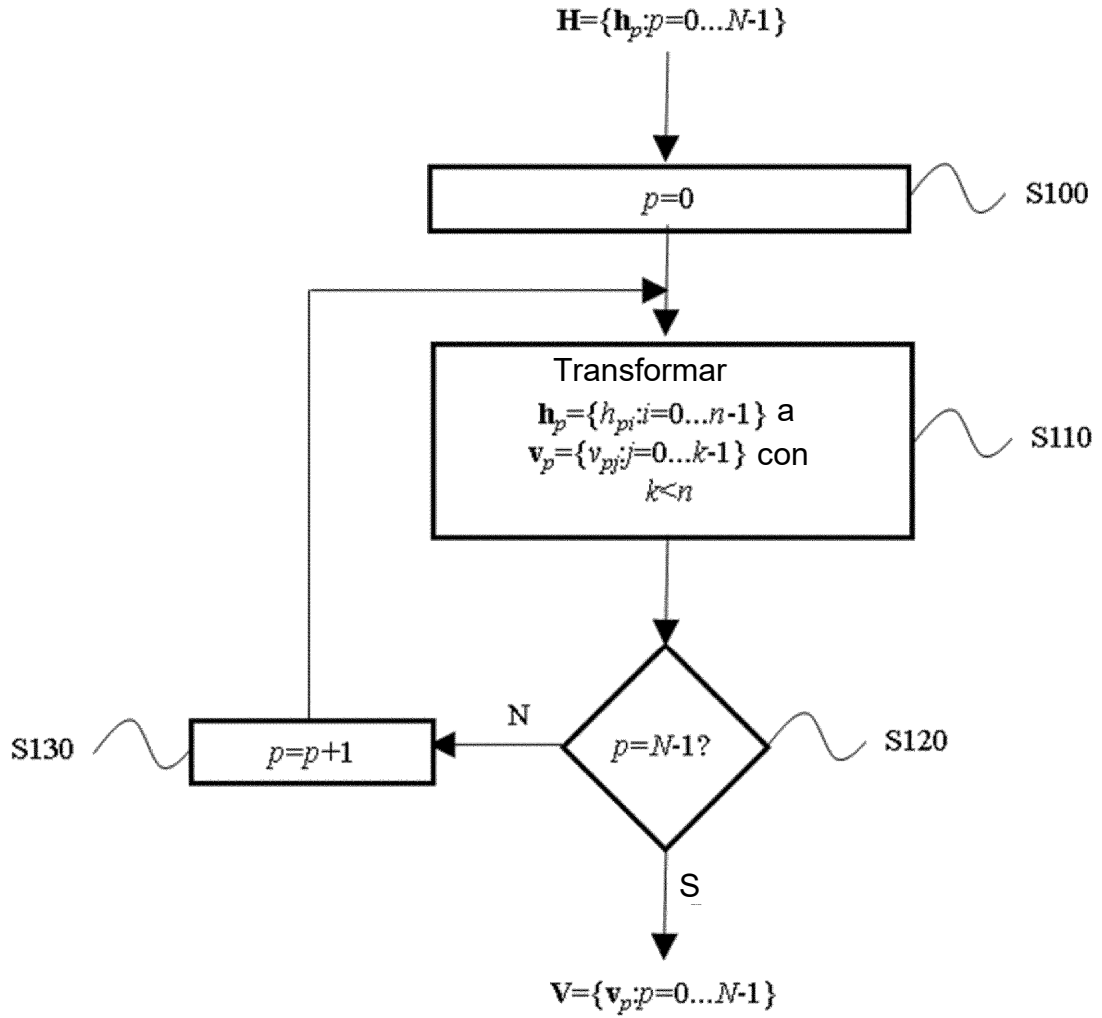


Figura 8

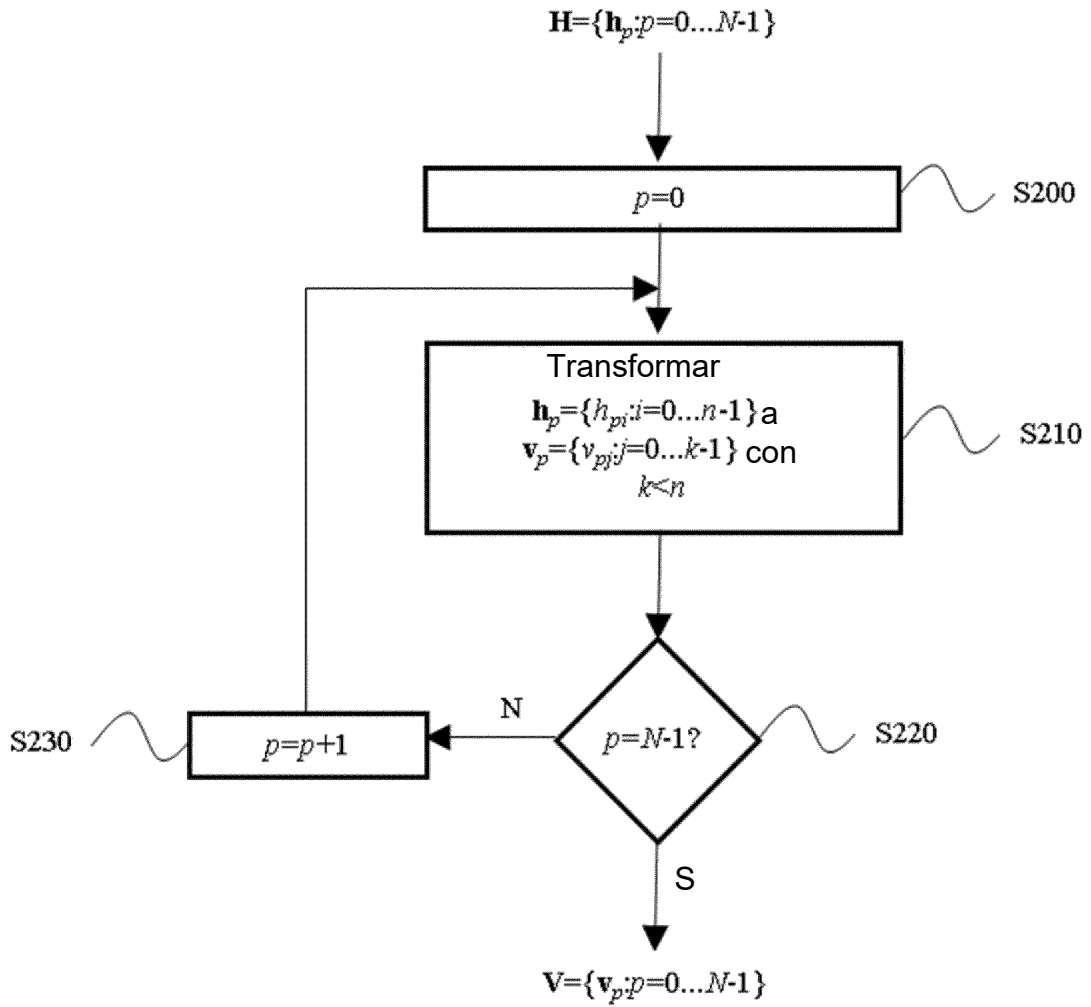


Figura 9

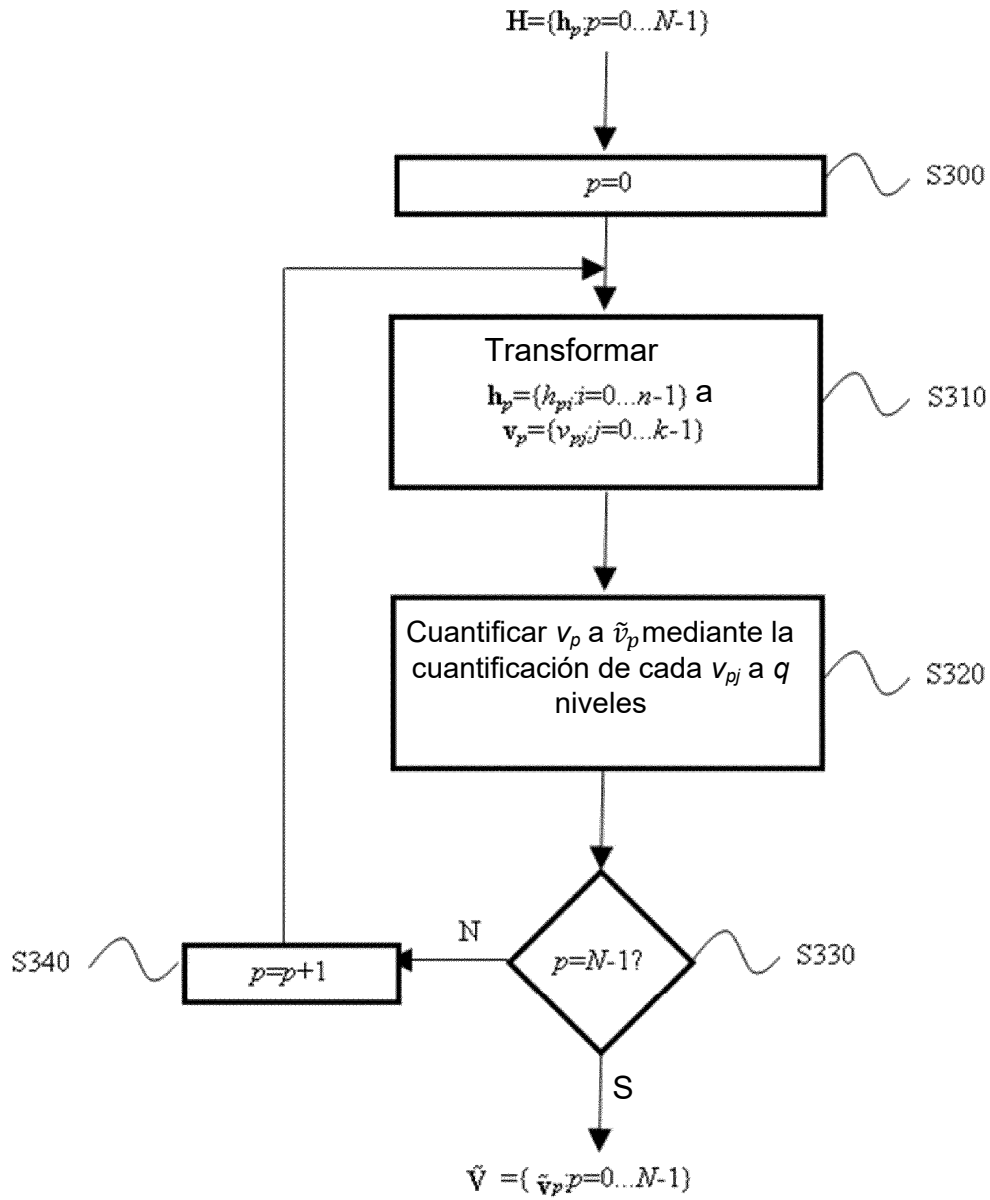


Figura 10

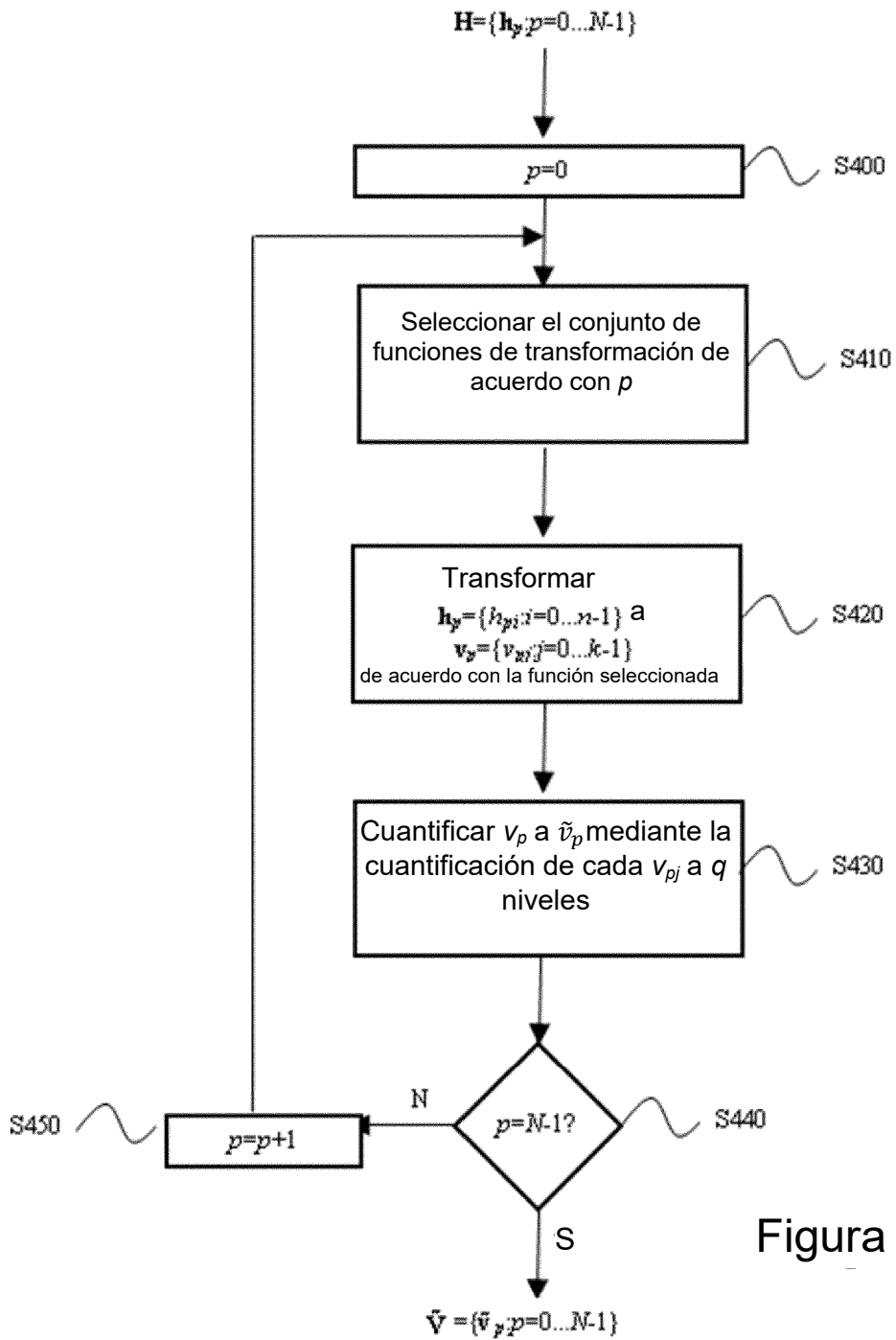


Figura 11

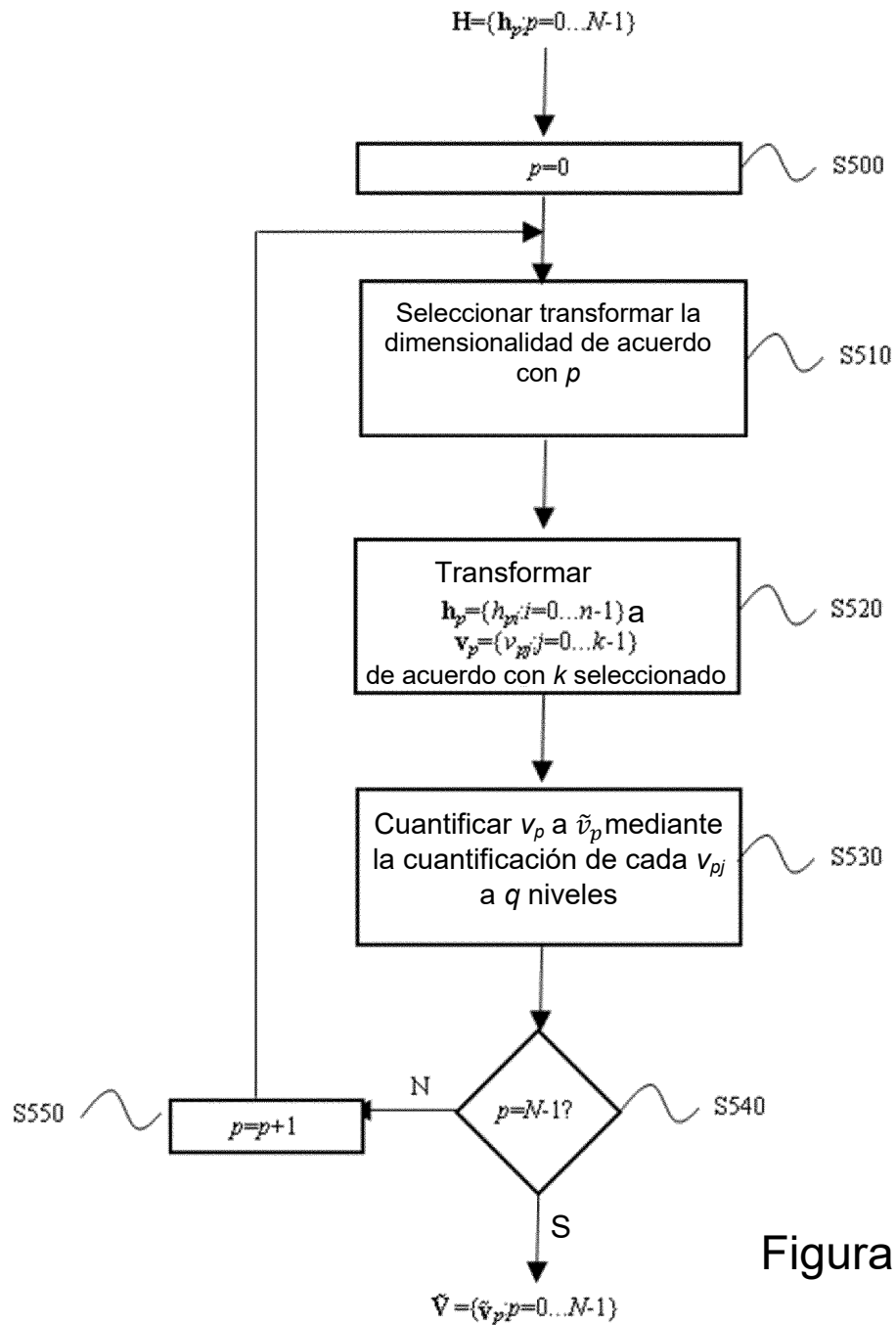


Figura 12

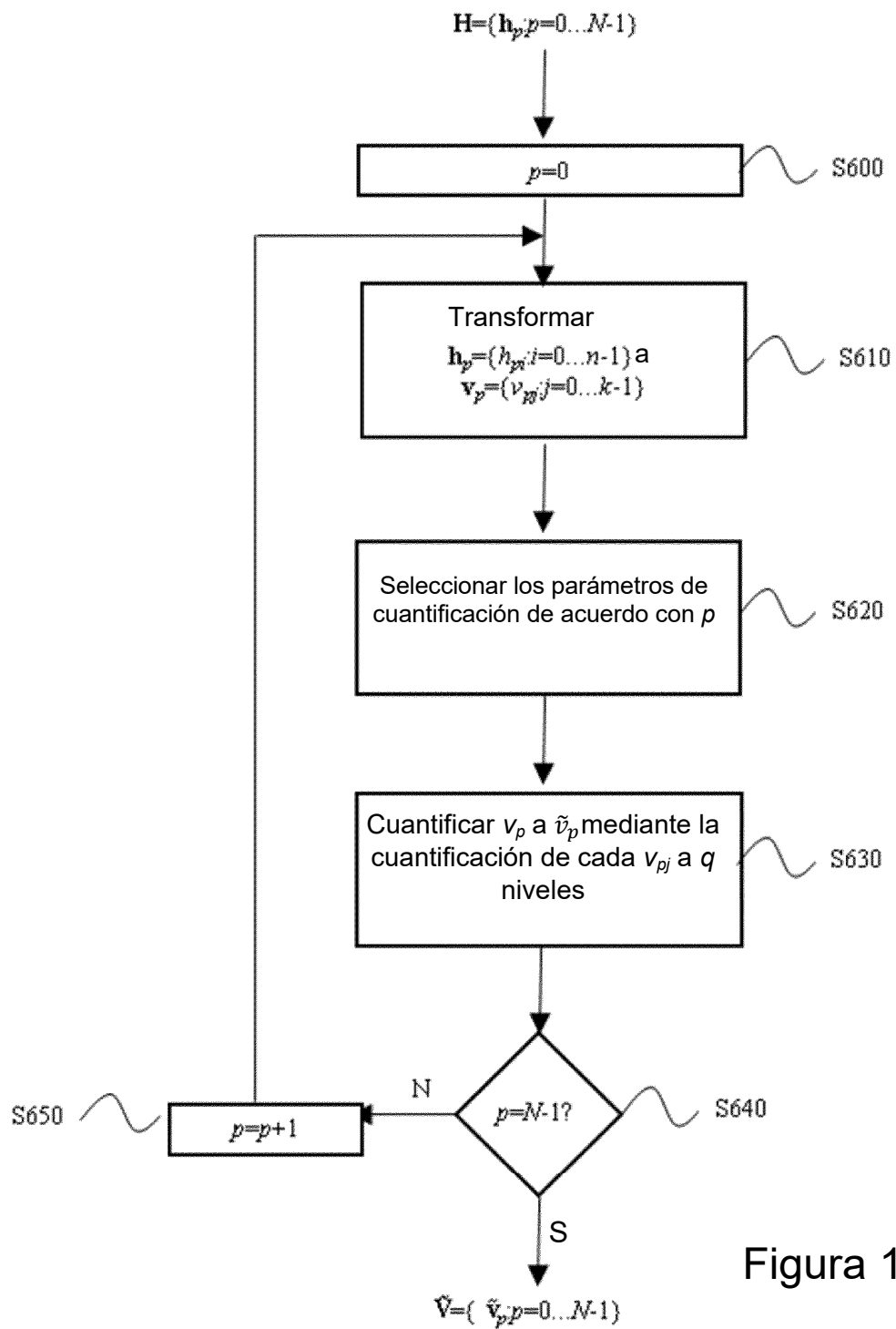


Figura 13

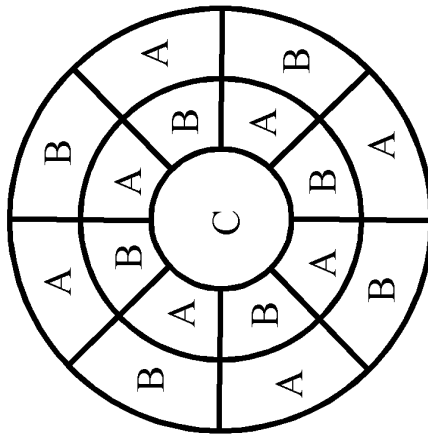


Figura 14

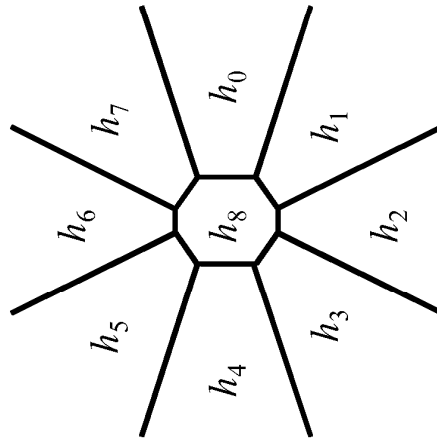


Figura 15

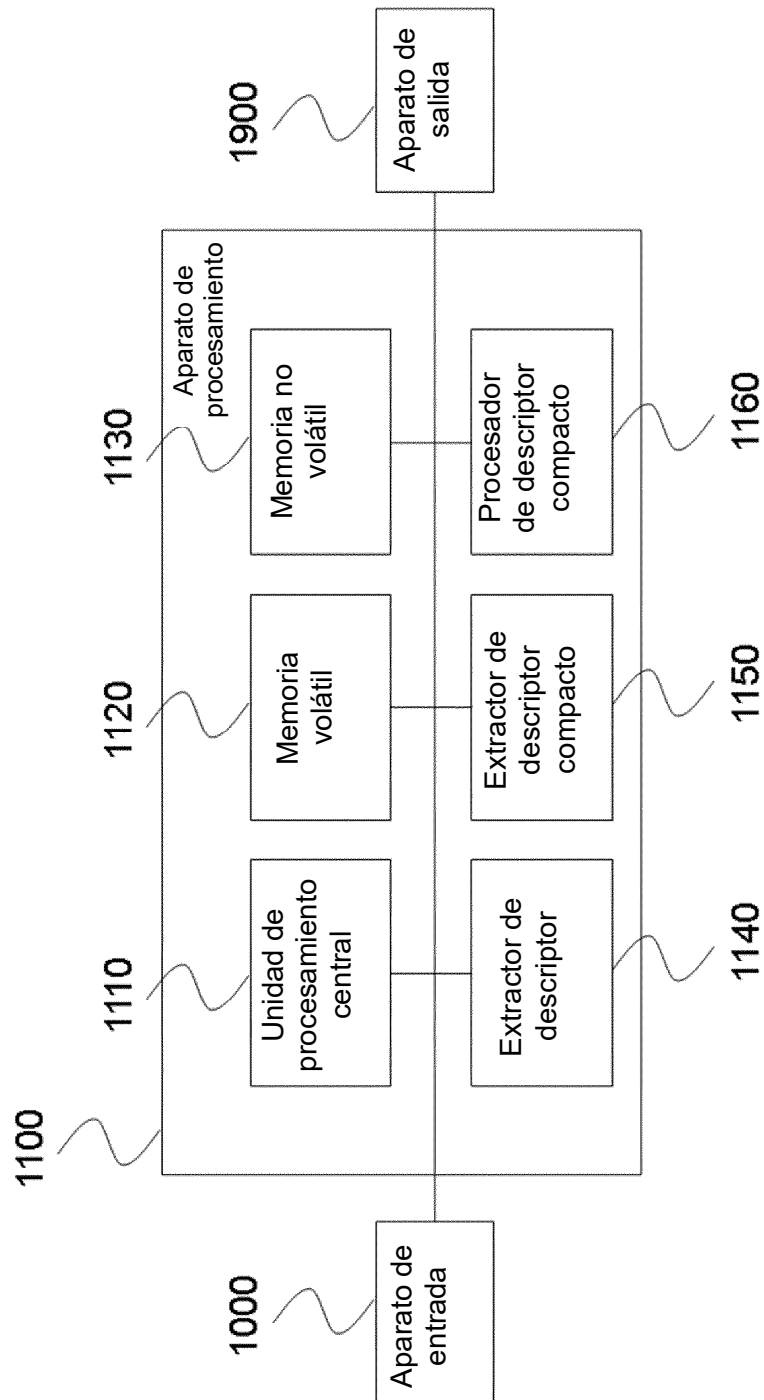


Figura 16