

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 939 244**

51 Int. Cl.:

G06V 10/46 (2012.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.07.2014 PCT/EP2014/065808**

87 Fecha y número de publicación internacional: **29.01.2015 WO15011185**

96 Fecha de presentación y número de la solicitud europea: **23.07.2014 E 14752573 (7)**

97 Fecha y número de publicación de la concesión europea: **30.11.2022 EP 3025273**

54 Título: **Identificación de puntos clave**

30 Prioridad:

24.07.2013 IT MI20131244

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

20.04.2023

73 Titular/es:

**TELECOM ITALIA S.P.A. (100.0%)
Via Gaetano Negri, 1
20123 Milano, IT**

72 Inventor/es:

**BALESTRI, MASSIMO;
FRANCINI, GIANLUCA y
LEPSØY, SKJALG**

74 Agente/Representante:

PONTI & PARTNERS, S.L.P.

ES 2 939 244 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

Identificación de puntos clave

5 Antecedentes de la invenciónCampo de la invención

10 **[0001]** La presente invención se refiere al campo del análisis de imágenes.

Descripción de la técnica anterior

15 **[0002]** En el campo del análisis de imágenes, antes de someter una imagen formada por una pluralidad de puntos (píxeles), cada uno caracterizado por un valor respectivo de un parámetro físico representativo de la imagen, tal como luminancia, a algunos tipos de procesamiento, tal como comparación con otra imagen, es ventajoso realizar la identificación de la posición y el tamaño de los detalles relevantes representados en esta imagen. En el campo del análisis de imágenes, con "detalle relevante" de una imagen se entiende una parte de un objeto incluido en la imagen que es fácilmente detectable incluso en presencia de cambios en el punto de vista del mismo objeto, en la iluminación y en el tipo de cámara.

20 **[0003]** Hasta hace unos años, era posible identificar la posición de los detalles relevantes de una imagen, pero no su tamaño. Más detalladamente, la identificación de la ubicación de un detalle relevante de una imagen se realiza a través de la identificación de un punto relevante asociado de la imagen, en la jerga, punto clave, que corresponde sustancialmente al centro del detalle relevante. En el caso de un detalle que tiene una forma circular, el punto clave coincide con el centro del detalle, mientras que en el caso de detalles que tienen diferentes formas, la posición del punto clave puede divergir del centro actual del detalle.

25 **[0004]** Recientemente, además de la identificación de puntos clave de una imagen, se han desarrollado procedimientos, gracias a los cuales también es posible determinar el tamaño del detalle relevante asociado con cada punto clave.

30 **[0005]** Actualmente, los procedimientos utilizados para identificar la posición y el tamaño de los detalles relevantes se basan en el concepto de "espacio de escala", que proporciona la aplicación de una serie de filtros gradualmente más intensos a la imagen. Los filtros aplicados a la imagen son típicamente filtros que realizan operaciones diferentes sobre los valores de los parámetros físicos (por ejemplo, luminancia) de los puntos de imagen. Típicamente, tales filtros se basan en la función gaussiana, cuya intensidad de filtrado está regida por un parámetro de filtrado σ (la desviación típica de la función gaussiana): cuanto mayor es el parámetro de filtrado σ , más plana y más ancha es la gaussiana, y más intenso es el efecto de suavizado que tiene la gaussiana. El espacio de escala de una imagen formada por una matriz de píxeles de coordenadas (x, y) es el espacio formado por el conjunto de imágenes filtradas (en términos de luminancia) obtenidas a partir de la imagen de partida aplicando filtros gradualmente más intensos, es decir, con valores gradualmente mayores de σ y es, por lo tanto, un espacio de tres dimensiones (x, y, σ) .

35 **[0006]** La teoría (véase, por ejemplo, T. Lindeberg (1992), "Scale-space behavior of local extremos and blobs", J. of Mathematical Imaging and Vision, 1 (1), páginas 65-99) establece que si se tiene un valor extremo, con respecto a σ , de la imagen filtrada para un punto (x_p, y_p, σ_p) perteneciente al espacio (x, y, σ) , es decir, un máximo o un mínimo, con respecto a σ , en una parte del espacio (x, y, σ) que rodea el punto (x_p, y_p, σ_p) , entonces ese punto está asociado con un detalle relevante, cuyas coordenadas centrales son (x_p, y_p) , y el tamaño es proporcional a σ_p . El tamaño (diámetro) del detalle (en unidades o píxeles) es igual a $2 * \sqrt{2} * \sigma_p$.

40 **[0007]** Identificando todos los puntos extremos en el espacio de escala, se obtiene por lo tanto la posición y el tamaño de los detalles relevantes de la imagen.

45 **[0008]** Para encontrar los puntos extremos en el espacio de escala, los procedimientos conocidos (tales como el procedimiento que usa el descriptor "Transformada de característica invariante de escala", SIFT, descrito en 1999 en el artículo Object recognition from local scale-invariant features of Lowe, David G., Proceedings of the International Conference on Computer Vision 2. páginas 1150 a 1157 y el tema de la Patente de Estados Unidos 6.711.293), considere una secuencia de imágenes filtradas con valores crecientes de σ , y, para cada punto de una imagen filtrada con un σ , compare sus valores con los valores de ocho puntos adyacentes de la misma imagen y los valores de los 18 (9 + 9) puntos adyacentes presentes en las imágenes filtradas correspondientes a los valores anteriores y siguientes de σ en la secuencia. Si este punto es menor o mayor que todos los adyacentes, entonces el punto es un extremo del espacio x, y, σ , y es un candidato para ser un punto clave. Este punto es solo un candidato porque se sabe (véase, por ejemplo, Lowe, DG, "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, 60, 2, páginas 91-110, 2004) que hay que eliminar los puntos correspondientes a partes de la imagen que tienen un bajo contraste y los puntos que se encuentran en estructuras similares a bordes, ya que la ubicación de un

detalle a lo largo de un borde puede variar fácilmente en diferentes imágenes que representan la misma escena. Por lo tanto, el punto no es fiable y, por lo tanto, se descarta.

Resumen de la invención

- 5
- [0009]** El solicitante se ha percatado de que los enfoques conocidos en el estado de la técnica para la identificación de los puntos clave de una imagen usan un subconjunto limitado de valores de σ para filtrar la imagen obteniendo únicamente una representación discreta de la imagen filtrada a medida que σ varía.
- 10 **[0010]** Sin embargo, el solicitante ha observado que para identificar con mayor precisión y eficacia los puntos clave de una imagen, reduciendo al mismo tiempo la cantidad de cálculos necesarios, es posible aproximar la imagen filtrada genérica para representarla con continuidad con respecto a σ , y no solo relativamente a un pequeño conjunto de valores discretos de este parámetro.
- 15 **[0011]** Un aspecto de la presente invención se refiere a un procedimiento para identificar puntos clave en una imagen digital que comprende un conjunto de píxeles. Cada píxel tiene asociado al mismo un valor respectivo de un parámetro representativo de imagen. Dicho procedimiento comprende aproximar una imagen filtrada. Dicha imagen filtrada depende de un parámetro de filtrado y comprende para cada píxel de la imagen una función de filtrado que depende del parámetro de filtrado para calcular un valor filtrado del valor del parámetro representativo del píxel. Dicha
- 20 aproximación comprende:
- a) generar un conjunto de imágenes de base filtradas; cada imagen de base filtrada es la imagen filtrada con un valor respectivo del parámetro de filtrado;
- b) para cada píxel de al menos un subconjunto de dicho conjunto de píxeles, aproximar la función de filtrado
- 25 mediante una función de aproximación respectiva en función de las imágenes de base filtradas; dicha función de aproximación es una función del parámetro de filtrado dentro de un intervalo predefinido del parámetro de filtrado;
- [0012]** El procedimiento comprende además, para cada píxel de dicho subconjunto, identificar dicho píxel como un candidato de punto clave si la función de aproximación tiene un extremo local que también es un extremo general
- 30 con respecto al parámetro de filtrado en un subintervalo respectivo interno a dicho intervalo predefinido.
- [0013]** Para cada píxel identificado como un punto clave candidato, el procedimiento comprende además:
- c) comparar el valor asumido por la función de aproximación en el valor del parámetro de filtrado correspondiente
- 35 al extremo general del píxel con los valores asumidos por las funciones de aproximación de los píxeles adyacentes en la imagen en los valores de los parámetros de filtrado de los respectivos extremos generales de dichos píxeles adyacentes, y
- d) seleccionar dicho píxel en función de esta comparación.
- 40 **[0014]** Según una realización de la presente invención, dicha aproximación de la función de filtrado por medio de una función de aproximación respectiva en función de las imágenes de base filtradas comprende calcular dicha función de aproximación en función de una combinación lineal de dichas imágenes de base filtradas.
- [0015]** Según una realización de la presente invención, dicha función de aproximación se basa en una
- 45 aproximación adicional de dicha combinación lineal de dichas imágenes de base filtradas.
- [0016]** Según una realización de la presente invención, dicha función de aproximación es un polinomio que tiene el parámetro de filtrado como una variable.
- 50 **[0017]** Según una realización de la presente invención, los coeficientes de dicho polinomio se calculan en función de las imágenes de base filtradas y en función de una aproximación de ponderaciones de dicha combinación lineal.
- [0018]** Según una realización de la presente invención, el procedimiento comprende además descartar a partir
- 55 de los píxeles seleccionados los píxeles donde el valor asumido por la función de aproximación en el parámetro de filtrado correspondiente al extremo general del píxel tiene un valor absoluto menor que un primer umbral.
- [0019]** Según una realización de la presente invención, el procedimiento comprende además:
- 60 - para cada píxel seleccionado, calcular la curvatura principal y la curvatura secundaria de la superficie formada por las funciones de filtrado en los píxeles de la imagen contenida en un parche centrado en dicho píxel seleccionado;
- descartar/mantener dicho píxel a partir de/en los píxeles seleccionados en función de la relación entre la curvatura principal y la curvatura secundaria.
- 65 **[0020]** Según una realización de la presente invención, el procedimiento comprende además:

- para cada píxel seleccionado, calcular el valor asumido por la segunda derivada de la función de aproximación con respecto al parámetro de filtrado en el extremo general correspondiente, y
- descartar/mantener dicho píxel a partir de/en los píxeles seleccionados en función de dicho valor asumido por la segunda derivada.

[0021] Según una realización de la presente invención, dicho punto clave de identificación se repite además en al menos una versión escalada de la imagen, usando el mismo intervalo predefinido del parámetro de filtrado.

10 **[0022]** Según una realización de la presente invención:

- al menos uno de los valores del parámetro de filtrado de las imágenes de base filtradas es igual al doble del menor de los valores del parámetro de filtrado de las otras imágenes de base filtradas;
- dicha versión escalada de la imagen se obtiene mediante la aproximación de las imágenes de base filtradas a partir de una versión aproximada de la imagen de base filtrada que tiene el valor más bajo del parámetro de filtrado, dicha versión aproximada de la imagen de base filtrada se aproxima mediante el submuestreo de la imagen de base filtrada con dicho valor del parámetro de filtrado que es el doble del valor más bajo del parámetro de filtrado.

20 **[0023]** Según una realización de la presente invención, dicha imagen filtrada se basa en la aplicación de filtros basados en laplacianos de gaussianos o filtros basados en diferencias de gaussianos, y dicho parámetro de filtrado es la desviación típica de la función gaussiana.

25 **[0024]** Según una realización de la presente invención, dicho polinomio es un polinomio de tercer grado con respecto al parámetro de filtrado.

30 **[0025]** Según una realización de la presente invención, cada píxel de la imagen tiene al menos una coordenada correspondiente que identifica la ubicación de los píxeles en la imagen; dicho procedimiento comprende además, para cada píxel seleccionado, modificar dicha al menos una coordenada de dicho píxel calculando un cambio correspondiente de coordenadas en función de una función de aproximación adicional que aproxima la función de filtrado en el píxel con respecto a dicho cambio de coordenadas; dicha función de aproximación adicional se basa:

- 1) en la función de filtrado de dicho píxel seleccionado en el valor del parámetro de filtrado correspondiente al extremo general del píxel seleccionado, y
- 2) en las funciones de filtrado de los píxeles adyacentes al píxel seleccionado en la imagen en el valor del parámetro de filtrado correspondiente al extremo general del píxel seleccionado.

40 **[0026]** Según una realización de la presente invención, dicho cálculo del cambio de coordenadas comprende identificar puntos máximos o mínimos en la función de aproximación adicional con respecto al cambio de coordenadas y establecer dicho cambio de coordenadas en función de los puntos máximos o mínimos identificados.

Breve descripción de los dibujos

45 **[0027]** Estas y otras características y ventajas de la presente invención resultarán evidentes a partir de la siguiente descripción de algunas realizaciones a modo de ejemplo y no de limitación, que se leerá junto con los dibujos adjuntos, en los que:

- La Figura 1A es un gráfico que muestra una señal de luminancia en función de una coordenada;
- La Figura 1B muestra, para diferentes valores crecientes de σ , un filtro de LoG correspondiente y la señal de la Figura 1A filtrada a través de este filtro de LoG;
- La Figura 2A muestra una imagen bidimensional, cada punto de la cual tiene un valor de luminancia respectivo;
- La Figura 2B muestra, para valores de σ , un filtro de LoG correspondiente y la imagen de la Figura 2A filtrada a través del filtro de LoG;
- La Figura 3A ilustra cuatro filtros de base LoGB;
- La Figura 3B muestra cómo el filtro de LoG aproximado por medio de una combinación lineal de acuerdo con una realización de la presente invención es similar al que se ha calculado explícitamente;
- La Figura 3C ilustra un diagrama que muestra cómo varían las ponderaciones de una combinación lineal de cuatro filtros de base LoG en función de σ para obtener un filtro de LoG genérico;
- La Figura 4A muestra la imagen de la Figura 2A filtrada a través de una convolución con un filtro LoG que tiene un σ igual a 2,5;
- La Figura 4B muestra la imagen de la Figura 2A filtrada que se aproxima al filtro de LoG con σ igual a 2,5 por medio de la función de aproximación de acuerdo con una realización de la presente invención;
- La Figura 4C es la imagen resultante de la diferencia entre la imagen de la Figura 4A y la imagen de la Figura 4B;
- Las Figuras 5A-5B muestran un diagrama de flujo que ilustra en términos de bloques funcionales un proceso para identificar el punto clave de una imagen de acuerdo con una realización de la presente invención;
- La Figura 6A muestra, por medio de una escala de grises, un ejemplo del valor máximo asumido por la función de

aproximación de acuerdo con una realización de la presente invención para cada punto de la imagen ejemplar de la Figura 2A;

La Figura 6B muestra, por medio de una escala de grises, un ejemplo del valor mínimo asumido por la función de aproximación de acuerdo con una realización de la presente invención para cada punto de la imagen de la Figura 2A;

Las Figuras 6C y 6D muestran un ejemplo de cuáles de los puntos de la imagen de la Figura 2A son puntos de máximo y mínimo, respectivamente, que son candidatos para ser puntos clave potenciales;

Las Figuras 7A y 7B muestran, respectivamente, los puntos correspondientes de máximo y mínimo que todavía se consideran puntos clave potenciales después de que se haya llevado a cabo un procedimiento de comparación con los puntos adyacentes de acuerdo con una realización de la presente invención;

La Figura 8A muestra los puntos identificados como punto clave en la primera octava de la imagen en la Figura 2A, y

La Figura 8B muestra los puntos identificados como punto clave en cinco octavas consideradas de la imagen de la Figura 2A.

Descripción detallada de realizaciones ejemplares de la invención

[0028] Típicamente, el filtro basado en la función gaussiana que se aplica a las imágenes puede ser un laplaciano de gaussiano ("Laplacian Of Gaussian", LoG) o una diferencia de gaussianos ("Difference Of Gaussian" DoG). La diferencia de gaussianos se aproxima al laplaciano de gaussiano, pero puede ser conveniente adoptarla por razones computacionales. En consecuencia, aunque en este documento siempre se hará referencia a las operaciones que utilizan filtros de LoG, se aplican consideraciones equivalentes en el caso de los filtros de DoG.

[0029] Con el fin de mostrar el mecanismo que se encuentra en la base de la identificación de detalles relevantes por medio de la aplicación de filtros de LoG, a continuación se presentarán dos ejemplos: en el primer ejemplo, ilustrado en las Figuras 1A y 1B, por simplicidad, en lugar de una imagen bidimensional, se considera una señal de luminancia unidimensional, mientras que el segundo ejemplo, mostrado en las Figuras 2A y 2B, se refiere a una imagen bidimensional.

[0030] Haciendo referencia al primer ejemplo, la Figura 1A es un gráfico que muestra, en función de una única coordenada x, un valor de luminancia; observando el gráfico de la Figura 1A, es posible advertir ya la presencia de dos detalles relevantes, correspondientes a los dos picos de la señal. Para ver cómo se pueden identificar estos dos detalles relevantes mediante un procedimiento de filtros de LoG, que permite no solo identificar la coordenada central, sino también el tamaño, se hará referencia a la Figura 1B, que muestra, para diferentes valores crecientes de σ ($\sigma = 2, \sigma = 6, \sigma = 10, \sigma = 14, \sigma = 18, \sigma = 22$), un filtro de LoG correspondiente (a la izquierda de la figura) y la señal de la Figura 1A filtrada a través de este filtro de LoG (a la derecha en la figura). En el ejemplo considerado, se pueden identificar dos extremos, a saber, un primer extremo en $x = 210$ cuando $\sigma = 6$, y un segundo extremo en $x = 110$ cuando $\sigma = 14$. Estos extremos indican la presencia de dos detalles relevantes cuyos centros están en 210 y 110 puntos (o píxeles si se tratara de una imagen) y cuya anchura es de aproximadamente 16,87 y 39,59 puntos, usando la relación de diámetro de punto relevante = $2 * \text{sqrt}(2) * \sigma$.

[0031] Con referencia al segundo ejemplo, la Figura 2A muestra una imagen bidimensional, cada punto de la cual tiene un valor de luminancia respectivo, mientras que la Figura 2B muestra, para valores crecientes de σ ($\sigma = 2, \sigma = 9, \sigma = 16, \sigma = 23$), un filtro de LoG correspondiente (a la derecha de la figura) y la imagen de la Figura 2A filtrada a través de dicho filtro de LoG (a la izquierda de la figura). La ventana en forma de media luna junto a la palabra "SCUOLA" es un detalle relevante, con una forma distinta fácilmente detectable, que tiene una altura en el centro de aproximadamente 19 píxeles. Esto significa que en el centro de la ventana, el resultado de la aplicación del filtro de LoG a la imagen tiene un valor máximo en un σ igual a $19 / (2 * \text{sqrt}(2)) = 6,46$. De hecho, se puede observar que en el centro de la ventana, el valor más alto (el más claro) obtenido como resultado del filtrado es aquel que corresponde al filtro de LoG que tiene $\sigma = 9$, es decir, el filtro de LoG de los cuatro filtros de LoG empleados que tienen el valor σ más cercano a 6,46.

[0032] Dado que un filtro de LoG tiende a aumentar considerablemente de tamaño a medida que aumenta σ (con $\sigma = 50$ el filtro se puede representar con una matriz de casi 500x500 puntos), el procesamiento descrito anteriormente se puede realizar ventajosamente usando un enfoque de "octava", con el fin de reducir el número de cálculos. El procesamiento de octavas se basa en la observación de que el resultado de un filtro con $\sigma = \sigma^*$ en la imagen original se puede reproducir con un filtro con $\sigma = \sigma^* / 2$ en la imagen escalada al 50 %. En el procesamiento de octavas, se fija un intervalo para σ , las imágenes filtradas se estudian con algunos σ que se encuentran en el intervalo, y luego la imagen se escala al 50 % repitiendo el mismo tipo de análisis en la imagen reducida (realizando los mismos filtrados). El proceso se itera hasta que la imagen escalada tiene un tamaño inferior a un umbral predeterminado. Por ejemplo, a partir de una imagen VGA (640x480) y deteniendo el proceso cuando el lado más corto de la imagen es inferior a 20 píxeles, se obtienen cinco octavas (640x480, 320x240, 160x120, 80x60, 40x30).

[0033] Uno de los conceptos básicos de la solución de acuerdo con las realizaciones de la presente invención se deriva de la observación de que es posible aproximar un filtro de LoG (x, y, σ) (donde x, y son las coordenadas

espaciales de la imagen (es decir, los puntos o píxeles que forman la imagen) y σ es la desviación típica de la gaussiana, con x, y, σ que definen el espacio de escala) como una combinación lineal de n filtros LoGB (x, y, σ_i) calculados previamente con diferentes $n \sigma = \sigma_i$ ($i = 1, 2, \dots, n$), en lo sucesivo denominados filtros básicos:

$$(1): \text{LoG}(x, y, \sigma) \approx p_1(\sigma) \text{LoGB}(x, y, \sigma_1) + p_2(\sigma) \text{LoGB}(x, y, \sigma_2) + p_3(\sigma) \text{LoGB}(x, y, \sigma_3) + \dots + p_n(\sigma) \text{LoGB}(x, y, \sigma_n),$$

5

donde $p_1(\sigma), p_2(\sigma), \dots, p_n(\sigma)$ son ponderaciones cuyo valor es una función de σ , como se mostrará más adelante en esta descripción. La dependencia espacial de x e y se ha omitido por simplicidad.

- 10 **[0034]** Con referencia al ejemplo que se muestra en la Figura 3A, se supone que se han calculado cuatro filtros de base LoGB (σ_1), LoGB (σ_2), LoGB (σ_3), LoGB (σ_4) con $\sigma_1 = 1,8$, $\sigma_2 = 2,846$, $\sigma_3 = 3,6$ y $\sigma_4 = 4,2214$. Haciendo una combinación lineal de estos cuatro filtros de base LoGB, es posible aproximar el filtro de LoG como:

$$(2): \text{LoG}(x, y, \sigma) \approx p_1(\sigma) \text{LoGB}(x, y, 1,8) + P_2(\sigma) \text{LoGB}(x, y, 2846) + P_3(\sigma) \text{LoGB}(x, y, 3,6) + P_4(\sigma) \text{LoGB}(x, y, 4,2214).$$

15

[0035] Usando la relación (2), es posible obtener una buena aproximación del filtro de LoG con σ igual a, por ejemplo, 2,5:

$$(3): \text{LoG}(x, y, 2.5) \approx 0,0161 \text{LoGB}(x, y, 1,8) + 0,2501 \text{LoGB}(x, y, 2,846) - 0,187 \text{LoGB}(x, y, 3,6) + 0,0836 \text{LoGB}(x, y, 4,2214)$$

20

[0036] En la Figura 3B es posible observar cómo el filtro de LoG aproximado por una combinación lineal (a la derecha de la figura) es similar al calculado explícitamente (a la izquierda de la figura).

[0037] Las ponderaciones $p_1(\sigma), p_2(\sigma), \dots, p_n(\sigma)$ se calculan resolviendo el sistema de ecuaciones lineales:

25

$$(4): Ap = b,$$

donde:

- 30 - A es una matriz que tiene un número de columnas igual al número n de filtros de base LoGB (en el ejemplo considerado, cuatro), en la que cada columna representa un filtro de base LoGB correspondiente. Suponiendo que el filtro de LoG genérico se puede representar mediante una matriz cuadrada $m \times m$ (donde cada elemento corresponde a un píxel), cada columna de A se construye dibujando en columnas las columnas de la matriz de cada filtro de base LoGB obteniendo un vector de columna correspondiente de m^2 elementos.

- 35 - b es el vector de columna de m^2 elementos que representa el filtro de LoG a aproximar.
 - p es un vector de n elementos que contienen las ponderaciones $p_1(\sigma), p_2(\sigma), \dots, p_n(\sigma)$ (en el ejemplo considerado, p_1, p_2, p_3, p_4) que se determinan resolviendo el sistema.

- [0038]** Para resolver el sistema, de acuerdo con una realización de la presente invención, es posible usar el procedimiento de mínimos cuadrados conocido o cualquier otro procedimiento que permita reducir la norma de la diferencia entre los valores observados y aproximados, como, por ejemplo, el procedimiento conocido como "recocido simulado" (a este respecto, véase, por ejemplo, Kirkpatrick, S., Gelatt, CD, Vecchi, MP (1983). "Optimization by Simulated Annealing." Science 220 (4598): 671-680).

- 45 **[0039]** Eligiendo un conjunto de q filtros de LoG a aproximar que tengan respectivamente $\sigma = \sigma'_1, \sigma'_2, \dots, \sigma'_q$, y en función de la relación (4), es posible calcular una matriz de ponderación W que tiene una fila para cada uno de los n filtros de base LoGB y una columna para cada uno de los q filtros de LoG a aproximar, y que contiene para cada columna las ponderaciones $p_1(\sigma), p_2(\sigma), \dots, p_n(\sigma)$ para aproximar el filtro de LoG correspondiente a dicha columna según la siguiente relación:

50

$$(5) AW = D,$$

donde D es una matriz que contiene los q filtros de LoG (σ'_j) ($j = 1, 2, \dots, q$).

[0040] Interpolando para cada uno de los n filtros de base LoGB los elementos correspondientes de la matriz de ponderación W es entonces posible determinar cómo varían las ponderaciones $p_1(\sigma)$, $p_2(\sigma)$, ..., $p_n(\sigma)$ con respecto a σ . La precisión con la que se aproxima la tendencia de las ponderaciones $p_1(\sigma)$, $p_2(\sigma)$, ..., $p_n(\sigma)$ con respecto a σ depende del número q de filtros de LoG considerados en relación (5) (cuanto mayor sea q, mejor será la aproximación).

5

[0041] La Figura 3C ilustra un diagrama que muestra cómo variar las ponderaciones $p_1(\sigma)$, $p_2(\sigma)$, $p_3(\sigma)$, $p_4(\sigma)$ del ejemplo considerado anteriormente en función de σ . En este caso, las curvas se generaron interpolando para cada ponderación 13 puntos correspondientes cada uno a 13 diferentes $\sigma = \sigma'_1, \sigma'_2, \dots, \sigma'_q$ (es decir, $q = 13$).

10 **[0042]** Con el fin de filtrar la imagen con un filtro de LoG(σ), la convolución del filtro de LoG se realiza con esta imagen I:

$$(6): L(\sigma) = \text{LoG}(\sigma) * I,$$

15 donde L(σ) es el resultado del filtro de LoG aplicado a la imagen (en lo sucesivo, simplemente denominado "imagen filtrada") y * es el símbolo de convolución.

[0043] Dado que la convolución es un operador lineal, explotando dicha propiedad, es ventajosamente posible obtener la aproximación de cualquier imagen filtrada L(σ) (es decir, para un filtrado correspondiente a cualquier σ) sin tener que calcularla explícitamente. De hecho, al explotar dicha propiedad y sustituir la relación (1) en la relación (6), se obtiene la siguiente relación:

20

$$(7): L(x, y, \sigma) \approx p_1(\sigma) L(x, y, \sigma_1) + p_2(\sigma) L(x, y, \sigma_2) + p_3(\sigma) L(x, y, \sigma_3) + \dots + p_n(\sigma) L(x, y, \sigma_n)$$

25 **[0044]** En otras palabras, gracias a la solución de acuerdo con una realización de la presente invención, es suficiente calcular explícitamente el filtrado para un número reducido de veces (es decir, n) para obtener n imágenes filtradas L(σ_i) ($i = 1, 2, \dots, n$) usando los n filtros de base LoGB(σ_i), y explotar la relación (7) para aproximar la imagen filtrada genérica L(σ) a partir de estas imágenes filtradas L(σ_i).

30 **[0045]** Por lo tanto, para obtener una aproximación a una imagen filtrada L(σ) es suficiente calcular una sola vez la matriz de ponderación W que da el valor de las n ponderaciones $p_i(\sigma)$ para un cierto conjunto de σ suficientemente grande (es decir, considerando una matriz D que contiene un número suficiente q de filtros de LoG) con el fin de satisfacer las necesidades de precisión requeridas.

35 **[0046]** El segundo concepto básico de la solución de acuerdo con las realizaciones de la presente invención proporciona la aproximación de la imagen filtrada genérica L(σ) por medio de una función de aproximación de filtrado que depende de un conjunto continuo de valores de σ .

[0047] De acuerdo con una realización de la presente invención, la función de aproximación es un polinomio de grado r, aunque se aplican consideraciones equivalentes en el caso en el que esta función de aproximación es una función diferente, por ejemplo, la función $a \log(\sigma) + b\sigma^2 + c\sigma + d$. Sin embargo, la elección de un polinomio es ventajosa puesto que los polinomios son fáciles de manejar, ya que son rápidos de calcular, fácilmente derivables y libres de singularidades.

45 **[0048]** Con el fin de calcular la función de aproximación como un polinomio de grado r de acuerdo con una realización de la presente invención, la matriz de ponderación W se aproxima a su vez de la siguiente manera:

$$(8): SF = W^T,$$

50 donde S es la matriz de tamaño q x (r + 1):

$$(9): S = \begin{bmatrix} (\sigma'_1)^r & \dots & (\sigma'_1)^2 & (\sigma'_1)^1 & 1 \\ (\sigma'_2)^r & \dots & (\sigma'_2)^2 & (\sigma'_2)^1 & 1 \\ (\sigma'_3)^r & \dots & (\sigma'_3)^2 & (\sigma'_3)^1 & 1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ (\sigma'_q)^r & \dots & (\sigma'_q)^2 & (\sigma'_{q1})^1 & 1 \end{bmatrix},$$

donde la notación $(\sigma'_1)^r$ significa " σ'_1 elevado a r",

y F es una matriz que contiene los valores de aproximación que sirven para aproximar, mediante polinomios de grado r en $\sigma'_1, \sigma'_2, \dots, \sigma'_q$, los pesos de la matriz de ponderación W a usar para aproximar los filtros de LoG que tienen $\sigma = \sigma'_1, \sigma'_2, \dots, \sigma'_q$, respectivamente. En mayor detalle, la matriz de aproximación F es una matriz de dimensión $(r + 1) \times n$, donde cada columna de F se usa para hacer una combinación lineal de las columnas de S. La matriz S multiplicada por una i-ésima columna F es un vector que se aproxima a las ponderaciones contenidas en la i-ésima columna de W^T . El elemento genérico de la k-ésima columna y la i-ésima fila de F es el valor que se usa en la combinación lineal de la k-ésima columna de S que corresponde a $(\sigma'_i)^{(r-k+1)}$. Con el fin de resolver el sistema (8), de acuerdo con una realización de la presente invención, es posible usar el procedimiento de mínimos cuadrados conocido o cualquier otro procedimiento que permita reducir la norma de la diferencia entre valores observados y aproximados.

[0049] Sustituyendo la relación (8) en (5), se obtiene:

15

$$(10): AF^T S^T \approx D.$$

[0050] Por tanto, sobre la base de la relación (10), de acuerdo con una realización de la presente invención, es posible aproximar un filtro $LoG(\sigma)$ que tiene cualquier σ realizando una combinación lineal de los valores de filtros de base $LoGB(\sigma_i)$ contenidos en la matriz A por medio de la multiplicación con la matriz F^T , y usando el resultado como coeficientes de un polinomio de grado r en σ , como se indica a continuación:

20

$$(11): LoG(\sigma)(:) \approx AF^T \begin{bmatrix} \sigma^r \\ \vdots \\ \sigma^2 \\ \sigma \\ 1 \end{bmatrix},$$

25 donde $(:)$ es una notación que indica que la matriz que precede a la notación se transforma en un vector, obtenido mediante el dibujo en columnas de las diversas columnas de la matriz.

[0051] Cabe señalar que, dada una base formada por los filtros básicos $LoGB(\sigma_i)$ contenidos en la matriz A, la matriz F se calcula solo una vez y se usa para aproximar cualquier filtro $LoG(\sigma)$.

30

[0052] Como se hizo anteriormente, usando la propiedad de linealidad de la convolución, y sustituyendo la relación (11) en la relación (6), se obtiene:

$$(12): [L(x, y, \sigma)(:) \approx L(x, y, \sigma_1)(:) L(x, y, \sigma_2)(:) \dots L(x, y, \sigma_n)(:)] F^T \begin{bmatrix} \sigma^r \\ \vdots \\ \sigma^2 \\ \sigma \\ 1 \end{bmatrix},$$

35

donde $L(\sigma)$ representa una imagen genérica filtrada con un filtro $LoG(\sigma)$, y $L(\sigma_i)$ ($i = 1, 2, \dots, n$) representa imágenes filtradas por medio de los n filtros de base $LoGB(\sigma_i)$.

[0053] En otras palabras, ampliando la relación (12), de acuerdo con una realización de la presente invención, una imagen filtrada genérica $L(x, y, \sigma)$ puede aproximarse con la siguiente función de aproximación:

40

$$(13): L(x, y, \sigma) \approx c_r(x, y)\sigma^r + c_{(r-1)}(x, y)\sigma^{(r-1)} + \dots + c_1(x, y)\sigma + c_0(x, y),$$

donde los $(r + 1)$ coeficientes c_r, \dots, c_0 del polinomio de la función de aproximación son funciones de las imágenes filtradas $L(\sigma_i)$ ($i = 1, 2, \dots, n$) usando los n filtros de base LoGB(σ_i) y de la matriz F , y varían de píxel a píxel en función de las coordenadas x e y . Esta aproximación es válida en el intervalo (cuyos extremos son parámetros que se pueden establecer) en donde σ varía en una sola octava.

[0054] De acuerdo con una realización, el grado r del polinomio de la función de aproximación es ventajosamente igual a 3, ya que está comprobado que es un buen término medio entre la complejidad de los cálculos y la precisión en la aproximación. Específicamente, con $r = 3$, la imagen filtrada genérica $L(x, y, \sigma)$ se puede aproximar con la siguiente función de aproximación:

$$(14): L(x, y, \sigma) \approx c_3(x, y)\sigma^3 + c_2(x, y)\sigma^2 + c_1(x, y)\sigma + c_0(x, y)$$

[0055] Para tener una idea de lo bueno de la aproximación obtenida mediante una función de aproximación como un polinomio de tercer grado, compare la Figura 4A, que muestra la imagen filtrada obtenida a partir de la imagen de la Figura 2A a través de una convolución con un filtro de LoG que tiene un σ igual a 2,5, con la Figura 4B, que representa la imagen filtrada obtenida a partir de la misma imagen de la Figura 2A mediante la aproximación del filtro de LoG con un σ igual a 2,5 por medio de la función de aproximación (14) usando cuatro filtros de base LoGB (σ_i) con $\sigma_i = 1,8, 2,846, 3,6, \text{ y } 4,2214$. La Figura 4C es la imagen resultante de la diferencia entre la imagen de la Figura 4A con la imagen de la Figura 4B. Como puede apreciarse observando la Figura 4C, la diferencia entre la imagen filtrada por medio de una convolución explícita con LoG (Figura 4A) y la imagen filtrada por medio de la función de aproximación (14) (Figura 4B) es cercana a cero.

[0056] Como se describirá en lo sucesivo en el presente documento en detalle, de acuerdo con una realización de la presente invención, la herramienta de la función de aproximación que se acaba de describir se usa ventajosamente para identificar, en cualquier imagen digital I , el conjunto de puntos clave a explotar para realizar el posterior análisis de imágenes.

[0057] El proceso de identificación del punto clave de una imagen digital I de acuerdo con una realización de la presente invención se ilustra en términos de bloques funcionales en el diagrama de flujo 100 mostrado en las Figuras 5A-5B.

[0058] Antes de pasar a describir en detalle los bloques funcionales de este proceso, debe observarse que la construcción de la función de aproximación requiere el uso de la matriz de aproximación F (véase la relación (12)), que se calcula ventajosamente una vez, por ejemplo, durante una fase previa de entrenamiento, y luego se usa para aproximar cualquier filtro LoG(σ) aplicado a cualquier imagen I . Durante esta fase de entrenamiento, seleccionar un conjunto de n filtros de base LoGB (σ_i) ($i = 1, 2, \dots, n$), con un $\sigma_i < \sigma_{i+1}$, y un conjunto de q filtros LoG(σ'_j) ($j = 1, 2, \dots, q$), y calcular la matriz de aproximación F como se describió anteriormente (véase la relación (10)).

[0059] Volviendo ahora a las Figuras 5A-5B, la primera fase del proceso proporciona que, a partir de una imagen genérica I , se calculan las n imágenes correspondientes filtradas por medio de los filtros de base LoGB (σ_i), a saber, se calculan los $L(\sigma_i)$ ($i = 1, 2, \dots, n$) (bloque 102).

[0060] En este punto (bloque 104), se selecciona un intervalo de trabajo en σ , donde se realizan las siguientes operaciones. Como quedará claro en la siguiente descripción, seleccionar como extremo inferior del intervalo de trabajo $\sigma_{i=1}$, y como extremo superior del intervalo de trabajo $\sigma_{i=n}$, es posible evitar hacer algunos cálculos en las etapas posteriores del proceso.

[0061] Entonces se selecciona un punto (x_t, y_t) de la imagen I (bloque 106), tal como el punto de coordinación $(x_t = 0, y_t = 0)$, para realizar en él operaciones relativas a los bloques 108-124.

[0062] La imagen filtrada $L(x_t, y_t, \sigma)$ en el punto seleccionado (x_t, y_t) se aproxima entonces calculando una función de aproximación (por ejemplo, un polinomio de grado r) usando la relación (12) con $x = x_t$ e $y = y_t$ (bloque 108). Por ejemplo, en el caso de $r = 3$, la imagen filtrada $L(x_t, y_t, \sigma)$ se aproxima mediante la siguiente función polinómica de tercer grado de σ (con coeficientes que dependen de (x_t, y_t)): $c^3(x_t, y_t)\sigma^3 + c_2(x_t, y_t)\sigma^2 + c_1(x_t, y_t)\sigma + c_0(x_t, y_t)$.

[0063] Una condición necesaria para que un punto en la imagen sea un punto clave es que ese punto tenga un valor extremo en una parte del espacio de escala (x, y, σ) que rodea este punto. De acuerdo con una realización de la presente invención, gracias al hecho de que las imágenes filtradas $L(x, y, \sigma)$ se aproximan mediante una función de aproximación que depende de σ , la determinación de si un punto tiene un valor extremo puede realizarse ventajosamente comparando la tendencia en σ de la función de aproximación de este punto con la tendencia en σ de

las funciones de aproximación de los puntos adyacentes.

- [0064]** Por esta razón, en la siguiente etapa (bloque 110), se calcula la primera derivada de la función de aproximación con respecto a σ , y se realiza una comprobación de si, y, en el caso afirmativo, dónde, esta derivada asume un valor igual a cero en el intervalo σ considerado (excluyendo los extremos), para identificar posibles puntos máximos o mínimos locales. Usando un polinomio como una función de aproximación, es posible calcular fácilmente la derivada muy rápidamente. Con referencia al ejemplo considerado, la primera derivada de la imagen filtrada en el punto (x_t, y_t) $L(x_t, y_t, \sigma)$ es igual a: $3c_3(x_t, y_t)\sigma^2 + 2c_2(x_t, y_t)\sigma + c_1(x_t, y_t)$.
- 10 **[0065]** Si esta primera derivada asume el valor cero en al menos un punto (σ_m del intervalo σ , excluyendo los extremos de este intervalo, (rama de salida Y del bloque 112), el proceso proporciona el cálculo del valor asumido por la función de aproximación en dicho al menos un σ_m (bloque 114) y la comparación de este valor de la función de aproximación con los valores asumidos por la misma función de aproximación en correspondencia con los extremos del intervalo σ considerado (bloque 116). Si el intervalo σ determinado en el bloque 104 tenía como extremo inferior $\sigma_i = 1$, y como extremo superior $\sigma_i = n$, ni siquiera es necesario tener que calcular los valores de la función de aproximación en los extremos del intervalo, puesto que estos valores ya se han calculado (sin aproximación) en el bloque 102 como imágenes filtradas $L(\sigma_1)$, $L(\sigma_n)$ a través de la base de filtro LoGB (σ_1), LoGB (σ_n).
- 15 **[0066]** A través de la comparación realizada en el bloque 116, se puede determinar si σ_m es también un punto máximo (o mínimo) general de la función de aproximación en el intervalo σ considerado o si es solo un punto máximo (o mínimo) local.
- 20 **[0067]** Si se determina que σ_m es una función de aproximación de punto máximo (o mínimo) general con respecto a σ (rama de salida Y del bloque 118), entonces el punto seleccionado correspondiente (x_t, y_t) que ha determinado los valores de los coeficientes actuales c_r, \dots, c_0 de la función de aproximación es un punto clave potencial. En este caso (bloque 120), las coordenadas (x_t, y_t) del punto, el valor σ_m y el valor de la función de aproximación calculados para σ_m se insertan en un elemento de una primera tabla, identificada como tabla de "puntos clave potenciales". Cabe señalar que para cada uno de los puntos que pertenecen a la primera tabla, también se obtiene una evaluación del diámetro del detalle relevante asociado con ese punto, igual a $2 * \text{sqrt}(2) * \sigma_m$.
- 25 **[0068]** En cambio, si se determina que σ_m no es un punto máximo (o mínimo) general de la función de aproximación con respecto a σ (rama de salida N del bloque 118), o en el caso en el que la derivada de la función de aproximación no asume el valor cero en al menos un punto σ_m en el intervalo σ , excluyendo los extremos de este intervalo, (rama de salida N del bloque 112), entonces el punto seleccionado correspondiente (x_t, y_t) que determinó los valores de los coeficientes actuales c_r, \dots, c_0 de la función de aproximación no puede ser un punto clave potencial. En este caso, (bloque 122), las coordenadas (x_t, y_t) del punto y el valor de la función de aproximación calculados para σ_m se insertan en un elemento de una segunda tabla, identificada como tabla de "puntos descartados".
- 30 **[0069]** De acuerdo con otra realización de la presente invención, con el fin de que un punto se considere un punto clave potencial, y luego que éste sea insertado en la primera tabla, el punto máximo (o mínimo) general correspondiente σ_m debe satisfacer además la condición de estar incluido en un subconjunto del intervalo de trabajo seleccionado en el bloque 104, con dicho subconjunto teniendo un extremo inferior mayor que $\sigma_i = 1$, y un extremo superior menor que $\sigma_i = n$. De esta manera, solo se conocen los puntos máximos o mínimos que ocurren en σ_m de los que se conoce el comportamiento de las funciones de aproximación en un entorno de σ_m que es suficientemente grande, tal como un entorno que tiene un tamaño mínimo de aproximadamente 0,1 (con respecto a σ).
- 35 **[0070]** Además, con el fin de evitar la aparición de artefactos que podrían comprometer la correcta identificación de los puntos clave, los puntos de imagen pertenecientes al borde de la imagen se descartan directamente y, por lo tanto, se insertan en la segunda tabla, independientemente de la presencia de posibles puntos máximos (o mínimos) generales.
- 40 **[0071]** Cabe señalar que para cada punto de coordenadas (x_t, y_t) es posible que haya más puntos máximos y/o mínimos. En este caso, en el caso del punto máximo, se puede considerar solo el punto que tiene el valor $L(x_t, y_t, \sigma)$ más alto, mientras que en el caso del punto mínimo, se puede considerar solo el punto que tiene el valor $L(x_t, y_t, \sigma)$ más bajo.
- 45 **[0072]** De acuerdo con una realización adicional de la presente invención, en lugar de usar para cada punto el mismo intervalo de trabajo en σ , es posible usar un intervalo de trabajo diferente respectivo. Por ejemplo, un punto máximo (o mínimo) local de la función de aproximación puede considerarse como un máximo (o mínimo) general con respecto a un intervalo σ que es un subintervalo del intervalo de trabajo que comprende σ_m y que tiene extremos dependientes de σ_m . En este punto se realiza una comprobación para determinar si el punto seleccionado (x_t, y_t) es el último punto de la imagen I o no (bloque 124).
- 50 **[0073]** En el caso negativo (rama de salida N del bloque 124) se selecciona un nuevo punto (x_t, y_t) de la imagen (bloque 126), y las operaciones descritas anteriormente se repiten en el nuevo punto (volver al bloque 108).

[0074] En el caso afirmativo (rama de salida Y del bloque 124), todos los puntos de la imagen se clasifican en la primera o en la segunda tabla.

5 **[0075]** La Figura 6A muestra, mediante una escala de grises, un ejemplo del valor máximo asumido por la función de aproximación para cada punto de la imagen de la Figura 2A ejemplar, donde un color más claro corresponde a un valor más alto. La Figura 6B muestra, mediante una escala de grises, un ejemplo del valor mínimo asumido por la función de aproximación para cada punto de la imagen de la Figura 2A, donde también en este caso un color más claro corresponde a un valor más alto. Las Figuras 6C y 6D muestran en negro un ejemplo de cuáles de los puntos
10 de la imagen de la Figura 2A son los puntos máximos y mínimos, respectivamente, que son candidatos para ser puntos clave potenciales (es decir, puntos que se incluyeron en la primera tabla).

[0076] De acuerdo con una realización de la presente invención, las operaciones posteriores del proceso de identificación de los puntos clave de las Figuras 5A-5B prevén verificar, para cada punto (x_i, y_i) de la imagen que
15 pertenece a la primera tabla que tiene un máximo en la función de aproximación, si el valor de la función de aproximación de dicho punto en el valor σ_m del máximo identificado también es mayor que los valores máximos asumidos por las funciones de aproximación de los ocho puntos adyacentes a ese punto en la imagen. De manera similar, para cada punto (x_i, y_i) de la imagen que pertenece a la primera tabla que tiene un mínimo en la función de aproximación, se verifica si el valor de la función de aproximación de ese punto en el valor σ_m del mínimo identificado
20 también es menor que los valores mínimos asumidos por las funciones de aproximación de los ocho puntos adyacentes a ese punto en la imagen.

[0077] Considerando los puntos máximos (consideraciones similares también pueden aplicarse a puntos mínimos), de acuerdo con una realización de la presente invención, se selecciona un punto (x_i, y_i) de la primera tabla
25 (bloque 128), y el valor máximo de la función de aproximación del punto, que se puede obtener a partir del elemento correspondiente en la primera tabla, se compara (bloque 130) con los valores máximos de las funciones de aproximación de los ocho puntos adyacentes en la imagen, que se pueden obtener por los elementos correspondientes a esos puntos adyacentes en la primera y/o en la segunda tabla. Se subraya que cada uno de los ocho puntos adyacentes puede ser, a su vez, un punto clave potencial (en este caso, el punto se enumera en la primera tabla) o
30 un punto que ya se ha descartado (en este caso, el punto se enumera en la segunda tabla). Si el valor máximo de la función de aproximación en el punto seleccionado parece ser mayor que todos los valores máximos de las funciones de aproximación de los puntos adyacentes (rama de salida Y del bloque 132), entonces ese punto todavía se considera un punto clave potencial y, por lo tanto, se deja en la primera tabla (bloque 134). Si el valor máximo de la función de aproximación en el punto seleccionado no es mayor que todos los valores máximos de la función de aproximación de
35 los puntos adyacentes (rama de salida N del bloque 132), entonces ese punto ya no debe considerarse un punto clave potencial y, por lo tanto, se elimina de la primera tabla y se inserta en la segunda tabla (bloque 136). A continuación, se realiza una comprobación para determinar si se han comparado o no todos los puntos enumerados en la primera tabla. En el caso negativo (rama de salida N del bloque 138), se selecciona un nuevo punto de la primera tabla (bloque 140), y las operaciones de los bloques 132-136 se llevan a cabo de nuevo en este nuevo punto. En el caso positivo
40 (rama de salida Y del bloque 138), la selección inicial de puntos clave potenciales ha terminado.

[0078] Usando las soluciones de acuerdo con las realizaciones de la presente invención, fue posible evaluar de una manera rápida y eficiente el comportamiento de la imagen filtrada en un punto genérico de la imagen con respecto al comportamiento de la imagen filtrada en los puntos adyacentes, simplemente comparando la tendencia de
45 la función de aproximación de ese punto con la tendencia de las funciones de aproximación de los puntos adyacentes.

[0079] Volviendo al ejemplo ilustrado en las Figuras 6C y 6D, las Figuras 7A y 7B muestran en color negro los puntos máximos y mínimos correspondientes, respectivamente, que han permanecido en la primera tabla (es decir, que todavía son puntos clave potenciales) después de que se haya llevado a cabo el procedimiento de los bloques
50 130-140.

[0080] De acuerdo con una realización de la presente invención, los puntos clave potenciales restantes en la primera tabla se consideran a partir de ahora independientemente del hecho de que son puntos máximos o mínimos.

55 **[0081]** El procedimiento de identificación de puntos clave de acuerdo con una realización de la presente invención comprende además eliminar de la primera tabla de los puntos clave potenciales a aquellos puntos que se supone que tienen una estabilidad deficiente, es decir, los puntos clave que pertenecen a elementos de la escena que, observando la escena de una manera diferente o con diferentes condiciones de iluminación, pueden cambiar la posición con respecto al objeto sobre el que se encuentran, o ya no pueden detectarse. De acuerdo con una realización
60 de la presente invención, la estabilidad se determina llevando a cabo una o más de las siguientes tres pruebas de estabilidad.

[0082] La primera prueba de estabilidad de acuerdo con una realización de la presente invención (bloque 142) proporciona descartar de la primera tabla los puntos con el valor absoluto de la función de aproximación calculado en
65 σ_m correspondiente inferior a un cierto umbral. Estos puntos pertenecen a áreas de la imagen con un contraste inferior

a un contraste mínimo (determinado por el umbral). Esta verificación también permite eliminar posibles puntos que se han identificado como punto clave solo debido a la aproximación realizada mediante la función de aproximación. En la práctica, en correspondencia con un área que tiene un color uniforme (por tanto, un área con un contraste muy bajo), el resultado del filtrado en un punto que pertenece a dicha área como variable debe tener valores casi constantes y cercanos a cero y, por lo tanto, debe tener una tendencia plana, pero la aproximación que explota la función de aproximación tiende a generar (especialmente si la función de aproximación es un polinomio) un máximo o mínimo local cercano a cero solo introducido por la aproximación, lo que puede permitir que el punto se clasifique como un punto clave en lugar de ser descartado.

10 **[0083]** La segunda prueba de estabilidad de acuerdo con una realización de la presente invención (bloque 144) prevé calcular para cada punto de la primera tabla y en el parche de 3x3 píxeles de la imagen centrada en este punto la curvatura principal y la curvatura secundaria (ortogonal a la primera curvatura principal) de la superficie formada por la función $L(x, y, \sigma)$ en los puntos pertenecientes a ese parche, y para comparar esas dos curvaturas, calcular la relación. Si parece que las dos curvas son similares, entonces significa que el punto cae en un área de la imagen donde su posición está bien definida, y el punto se deja en la primera tabla, mientras que si las dos curvas difieren significativamente, entonces significa que el punto cae en un área de la imagen similar a un tablero y, por lo tanto, no es muy fiable ya que su ubicación o existencia varía considerablemente dependiendo de cómo se observe la escena. En este último caso, el punto se elimina de la primera tabla. Esta prueba también se usa en los procedimientos conocidos para la identificación de los puntos clave, pero, a diferencia de este último, en el que el parche de puntos usado para calcular las curvaturas pertenece a una imagen ya filtrada, de acuerdo con las realizaciones de la presente invención, el parche se construye en el momento mediante el cálculo de la imagen filtrada en los puntos en el σ_m del punto considerado, con el fin de tener una imagen más precisa de la superficie, en la escala a la que realmente pertenece el detalle.

25 **[0084]** La tercera prueba de estabilidad de acuerdo con una realización de la presente invención (bloque 146), prevé calcular el valor de la curvatura (en σ) de la función $L(x, y, \sigma)$, dado por la segunda derivada de la función de aproximación calculada en correspondencia con σ_m del punto. Con referencia al ejemplo previamente considerado de función de aproximación correspondiente a un polinomio de tercer grado, la curvatura de la función $L(x_t, y_t, \sigma)$ en el punto σ_m es igual a: $L''(x_t, y_t, \sigma_m) = 6c_3(x_t, y_t) \sigma_m + 2c_2(x_t, y_t)$. Si el valor absoluto de la curvatura es mayor que un umbral, entonces el punto se considera estable y, por lo tanto, se deja en la primera tabla. Si el valor absoluto de la curvatura resulta ser menor que el umbral, entonces el punto se considera inestable y, por lo tanto, se elimina de la primera tabla.

35 **[0085]** Para reducir los cálculos, el proceso para identificar los puntos clave se realiza ventajosamente con el enfoque de octavas, es decir, repitiendo todas las operaciones descritas hasta ahora en versiones de la imagen I que están cada vez más escaladas, usando siempre el mismo intervalo de trabajo de σ .

40 **[0086]** Por esta razón, de acuerdo con una realización de la presente invención, después de realizar las operaciones descritas hasta ahora, se lleva a cabo un refinamiento de las coordenadas de los puntos enumerados en la primera tabla (bloque 148). Hasta este punto, de hecho, las coordenadas (x_t, y_t) de cada punto enumerado en la primera tabla corresponden a las coordenadas reales y enteras de los píxeles de la imagen original I . Si no se llevase a cabo dicho refinamiento, las coordenadas de los puntos identificados en las octavas superiores, en las que la imagen se escala a la mitad, a un cuarto, a un octavo, y así sucesivamente del tamaño original de la imagen, devuelta a resolución completa, provocarían la identificación de puntos clave que no están centrados con los detalles relevantes correspondientes. La fase de refinamiento de las coordenadas está dirigida a determinar con mayor precisión el centro de los detalles relevantes.

50 **[0087]** Con el fin de llevar a cabo este refinamiento, de acuerdo con una realización de la presente invención, un enfoque similar al que se ha expuesto en lo anterior para aproximar con una función de aproximación la imagen filtrada en un punto a medida que σ varía. En este caso, lo que se aproxima es en cambio la imagen filtrada ya que las coordenadas espaciales $x_t - u$ e $y_t - v$ varían en el entorno del punto genérico (x_t, y_t) enumerado en la primera tabla, fijando σ en el valor σ_m correspondiente.

55 **[0088]** Por ejemplo, de acuerdo con una realización de la presente invención, la imagen filtrada al variar x e y puede aproximarse mediante una función de aproximación, por ejemplo, un polinomio de segundo grado en las dos variables u y v :

$$(15): L(x_t - u, y_t - v, \sigma) \approx l_3(x_t, y_t, \sigma)u^2 + l_4(x_t, y_t, \sigma)v^2 + l_3(x_t, y_t, \sigma)uv + l_2(x_t, y_t, \sigma)u + l_1(x_t, y_t, \sigma)v + l_0(x_t, y_t, \sigma)$$

60 **[0089]** De una manera similar a la ya descrita, los coeficientes de la función de aproximación se calculan como una combinación lineal de algunas imágenes filtradas obtenidas por filtrado de LoG. Por ejemplo, de acuerdo con una realización de la presente invención, los coeficientes son una combinación de la imagen filtrada en los puntos 3 x 3

centrados en el punto (x_t, y_t) , con σ en el valor σ_m (es decir, en los valores del parche usado para calcular la relación de las curvaturas principal y secundaria). Generalizando, para obtener los coeficientes, se construye una matriz de aproximación G de la misma manera que la matriz de aproximación F descrita anteriormente, y dicha matriz se multiplica por los filtros de LoG del parche. A continuación, la función de aproximación se somete a operaciones para la identificación de un máximo o mínimo (dependiendo de si el punto (x_t, y_t) se ha identificado como un máximo o un mínimo), correspondiente al punto donde la primera derivada con respecto a u y la primera derivada con respecto a v son iguales a cero. Estando el parche centrado en el punto (x_t, y_t) , la u y la v que resuelven el sistema dado imponiendo la primera derivada con respecto a u y la primera derivada con respecto a v iguales a cero, proporcionan el cambio a aplicar a las coordenadas (x_t, y_t) . De acuerdo con una realización de la presente invención, si se calcula que el cambio es mayor en valor absoluto de un píxel de la imagen al menos a lo largo de u o a lo largo de v , entonces, el punto se descarta de la primera tabla. Este último evento es poco común, pero aún puede ocurrir ya que todo el proceso de identificación de los extremos en el espacio de escala (x, y, σ) ocurrió primero trabajando a lo largo de σ , y luego a lo largo de x e y . De acuerdo con una realización de la presente invención, proporcionada para aumentar los cálculos requeridos y la complejidad del procedimiento, sería posible aproximar la imagen filtrada con una única función de x , y y σ .

[0090] En este punto, todos los puntos que permanecen en la primera tabla se identifican como puntos clave de la imagen I en la octava considerada (bloque 150). Para cada punto clave, se conoce tanto la posición del mismo en la imagen (las coordenadas (x_t, y_t)), posiblemente modificadas de acuerdo con la fase de refinamiento del bloque 148), como el tamaño del detalle relevante asociado (igual a $2 \cdot \sqrt{2} \cdot \sigma_m$).

[0091] La Figura 8A muestra los puntos identificados como puntos clave en la primera octava de la imagen ejemplar mostrada en la Figura 2A. Cada punto clave se identifica con un círculo centrado en la posición del punto clave, y tiene un diámetro proporcional al diámetro del detalle relevante asociado.

[0092] Volviendo a las Figuras 5A-5B, en este punto, se verifica si la octava considerada hasta ahora es la última de un conjunto de octavas seleccionadas (por ejemplo, cinco octavas). En el caso positivo (rama de salida Y del bloque 151), el proceso finaliza, de lo contrario (rama de salida N del bloque 151) se calcula una versión escalada de la imagen (bloque 152) para pasar a la octava siguiente, y luego el proceso de identificación de punto clave se reitera en la nueva octava (retorno al bloque 102). Después de haber reiterado el proceso durante un número suficiente de octavas (por ejemplo, cinco), se termina el proceso de identificación de puntos clave.

[0093] La Figura 8B muestra los puntos identificados como punto clave en todas las octavas consideradas, en el ejemplo considerado, cinco, de la imagen de muestra mostrada en la Figura 2A.

[0094] De acuerdo con una realización de la presente invención, en lugar de calcular directamente la imagen escalada correspondiente a la siguiente octava, la versión escalada de la imagen puede aproximarse eligiendo el σ_i para los filtros de base LoGB (σ_i) de modo que uno de tales σ_i es igual al doble del primer σ_{i-1} (que es el más bajo entre los σ_i considerados), y la imagen filtrada puede submuestrearse con dicho σ_i que es el doble de σ_1 (tomando un píxel cada dos tanto horizontal como verticalmente). De esta manera, se obtiene una buena aproximación de cómo resultaría la imagen escalada al 50 % si se filtra con el filtro de base LoGB (σ_1). Con el submuestreo, por lo tanto, se obtiene la imagen de la siguiente octava filtrada con el primer filtro de base LoGB (σ_1). El filtrado de la imagen escalada al 50 % correspondiente al filtro de base genérico LoGB (σ_i) se obtiene filtrando la imagen escalada al 50 % filtrada con el filtro de base previo LoGB (σ_{i-1}). Las coordenadas x , y y las escalas σ de los puntos clave extraídos en las diversas octavas se comunican posteriormente al tamaño de la imagen original I .

[0095] La descripción anterior muestra y describe en detalle varias realizaciones de la presente invención; sin embargo, hay varias modificaciones posibles a las realizaciones descritas, así como diferentes realizaciones de la invención, sin apartarse del alcance definido por las reivindicaciones adjuntas.

[0096] Por ejemplo, aunque en la presente descripción se hace referencia a un procedimiento para la identificación de un punto clave que prevé realizar operaciones en todos los puntos de imagen (excluyendo los puntos en su borde), se pueden aplicar consideraciones similares en el caso en que solo un subconjunto de los puntos sea sometido a dichas operaciones.

[0097] Es más, aunque en la descripción se ha hecho referencia a filtros basados en LoG o DoG, donde el parámetro de filtrado que determina la intensidad de filtrado de dichos filtros es la desviación típica de la función gaussiana, se aplican consideraciones similares en el caso en el que los filtros se obtienen sobre la base de diferencias de versiones suavizadas de la imagen.

REIVINDICACIONES

1. Un procedimiento para identificar puntos clave en una imagen digital (*I*) que comprende un conjunto de píxeles, teniendo cada píxel, asociado al mismo, un valor respectivo de un parámetro representativo de imagen, 5 comprendiendo dicho procedimiento:

- aproximar (108), sin calcular explícitamente, una imagen filtrada ($L(\sigma)$, 102) de dicha imagen digital (*I*), siendo dicha imagen filtrada obtenida a partir de dicha imagen digital (*I*) por medio de una función de filtrado, siendo dicha función de filtrado un laplaciano de gaussiano o una diferencia de gaussianos y teniendo un parámetro de filtrado 10 que es la desviación típica de la función gaussiana, comprendiendo dicha aproximación:

a) generar (102) un conjunto de imágenes de base filtradas ($L(\sigma_i)$) de dicha imagen digital (*I*), siendo cada imagen de base filtrada ($L(\sigma_i)$) la imagen digital (*I*) filtrada por medio de un filtro de base respectivo, siendo cada filtro de base una matriz cuyos elementos se obtienen calculando dicha función de filtrado con un valor 15 respectivo (σ_i) del parámetro de filtrado;

b) para cada píxel (x_t, y_t) de al menos un subconjunto de dicho conjunto de píxeles de la imagen digital (*I*), aproximar (106) la imagen filtrada ($L(x_t, y_t, \sigma)$) obtenida a partir de dicha imagen digital (*I*) por medio de dicha función de filtrado que tiene un valor de parámetro de filtrado por medio de una función de aproximación respectiva en función de las imágenes de base filtradas, siendo dicha función de aproximación una función del 20 parámetro de filtrado dentro de un intervalo predefinido (104) de valores del parámetro de filtrado (σ), variando dicho intervalo predefinido desde un valor más bajo del parámetro de filtrado hasta un valor más alto del parámetro de filtrado, donde:

donde dicha función de aproximación es un polinomio expresado de la siguiente manera:

25

$$(13): L(x, y, \sigma) \approx c_r(x, y)\sigma^r + c_{(r-1)}(x, y)\sigma^{(r-1)} + \dots + c_1(x, y)\sigma + c_0(x, y),$$

donde:

30 $L(x, y, \sigma)$, es la imagen filtrada a aproximar
 $c_r(x, y)$ es el coeficiente polinómico para σ^r

donde dichos coeficientes polinómicos $c_r(x, y)$ se obtienen como una combinación lineal de dichas imágenes de base filtradas según la siguiente ecuación:

35

$$[c_r(x, y) \ c_{r-1}(x, y) \ \dots \ c_1(x, y) \ c_0(x, y)] \\ = [L(x, y, \sigma_1)(:)\ L(x, y, \sigma_2)(:)\ L(x, y, \sigma_{n-1})(:)\ L(x, y, \sigma_n)(:)]F^T$$

donde F^T se calcula resolviendo la siguiente ecuación:

40

$$SF = W$$

donde:

45 F es la incógnita que, después de ser calculada y transpuesta, da F^T
 S es una matriz formada por valores específicos de σ_i de la siguiente manera:

$$S = \begin{bmatrix} (\sigma'_1)^r & \dots & (\sigma'_1)^2 & (\sigma'_1)^1 & 1 \\ (\sigma'_2)^r & \dots & (\sigma'_2)^2 & (\sigma'_2)^1 & 1 \\ (\sigma'_3)^r & \dots & (\sigma'_3)^2 & (\sigma'_3)^1 & 1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ (\sigma'_q)^r & \dots & (\sigma'_q)^2 & (\sigma'_{q1})^1 & 1 \end{bmatrix}$$

50 donde σ'_i (con i variando de 1 a q) son valores específicos del parámetro de filtrado σ W es una matriz calculada resolviendo la siguiente ecuación:

$$AW = D$$

donde:

5

A es una matriz formada por los filtros de base, donde la matriz A tiene un número de columnas igual al número de filtros de base, donde cada columna de la matriz A representa un filtro de base respectivo, donde cada columna de la matriz A se construye mediante la disposición en columnas de las columnas de cada filtro de base

10

D es una matriz, cada elemento de la cual se obtiene calculando la función de filtrado, para cada punto (x, y) y para cada uno de dichos valores específicos σ^j con j variando de 1 a q
W es la incógnita que, después de ser calculada y transpuesta, da W^T

15

- para cada píxel de dicho al menos un subconjunto, identificar (120) dicho píxel como un punto clave candidato si, en dicho píxel, la función de aproximación tiene un extremo local (110) que también es un extremo general (116) con respecto al parámetro de filtrado para un valor σ_m del parámetro de filtrado en un subintervalo respectivo de valores del parámetro de filtrado interno a dicho intervalo predefinido, teniendo dicho subintervalo respectivo un extremo inferior más grande que dicho valor más bajo del parámetro de filtrado y un extremo superior más bajo que dicho valor más alto del parámetro de filtrado, donde la diferencia entre dicho valor más bajo y dicho extremo inferior, y la diferencia entre dicho valor más alto y dicho extremo superior, son tales que solo se consideran como puntos clave candidatos los puntos máximos o mínimos que ocurren en σ_m de los que se conoce el comportamiento de la función de aproximación en un entorno de σ_m que es suficientemente grande;

20

- para cada píxel identificado como punto clave candidato:

25

c) comparar (130) el valor asumido por la función de aproximación en el valor del parámetro de filtrado correspondiente al extremo general del píxel con los valores asumidos por las funciones de aproximación de los píxeles adyacentes a dicho píxel en la imagen en los valores de los parámetros de filtrado de los respectivos extremos generales de dichos píxeles adyacentes, y

30

d) seleccionar (136) dicho píxel en función de esta comparación.

2. El procedimiento de la reivindicación 1, que comprende además descartar de los píxeles seleccionados los píxeles donde el valor asumido por la función de aproximación en el parámetro de filtrado correspondiente al extremo general del píxel tiene un valor absoluto menor que un primer umbral.

35

3. El procedimiento de cualquiera de las reivindicaciones anteriores, que comprende además:

- para cada píxel seleccionado, calcular la curvatura principal y la curvatura secundaria de la superficie formada por las funciones de filtrado en los píxeles de la imagen contenida en un parche centrado en dicho píxel seleccionado;

40

- descartar/mantener dicho píxel a partir de/en los píxeles seleccionados en función de la relación entre la curvatura principal y la curvatura secundaria.

4. El procedimiento de cualquiera de las reivindicaciones anteriores, que comprende además:

45

- para cada píxel seleccionado, calcular el valor asumido por la segunda derivada de la función de aproximación con respecto al parámetro de filtrado en el extremo general correspondiente, y

- descartar/mantener dicho píxel a partir de/en los píxeles seleccionados en función de dicho valor asumido por la segunda derivada.

50

5. El procedimiento de cualquiera de las reivindicaciones anteriores, donde dicha identificación de puntos clave se repite además en al menos una versión escalada de la imagen, usando el mismo intervalo predefinido del parámetro de filtrado.

55

6. El procedimiento de la reivindicación 5, donde:

- al menos uno de los valores del parámetro de filtrado de las imágenes de base filtradas es igual al doble del menor de los valores del parámetro de filtrado de las otras imágenes de base filtradas;

60

- dicha versión escalada de la imagen se obtiene mediante la aproximación de las imágenes de base filtradas a partir de una versión aproximada de la imagen de base filtrada que tiene el valor más bajo del parámetro de filtrado, siendo dicha versión aproximada de la imagen de base filtrada aproximada mediante el submuestreo de la imagen de base filtrada con dicho valor del parámetro de filtrado que es el doble del valor más bajo del parámetro de filtrado.

7. El procedimiento de cualquiera de las reivindicaciones anteriores, donde dicha función de aproximación es un polinomio de tercer grado con respecto al parámetro de filtrado.

5 8. El procedimiento según cualquiera de las reivindicaciones anteriores, donde cada píxel de la imagen tiene al menos una coordenada (x, y) correspondiente que identifica la ubicación de los píxeles en la imagen, comprendiendo además dicho procedimiento, para cada píxel seleccionado, modificar (148) dicha al menos una coordenada de dicho píxel calculando un cambio correspondiente de coordenadas en función de una función de aproximación adicional que aproxima la imagen filtrada en el píxel con respecto a dicho cambio de coordenadas,
10 basándose dicha función de aproximación adicional:

1) en la función de filtrado de dicho píxel seleccionado en el valor del parámetro de filtrado correspondiente al extremo general del píxel seleccionado, y
15 2) en las funciones de filtrado de los píxeles adyacentes al píxel seleccionado en la imagen en el valor del parámetro de filtrado correspondiente al extremo general del píxel seleccionado.

9. El procedimiento de la reivindicación 8, donde dicho cálculo del cambio de coordenadas comprende identificar puntos máximos o mínimos en la función de aproximación adicional con respecto al cambio de coordenadas y establecer dicho cambio de coordenadas en función de los puntos máximos o mínimos identificados.

20

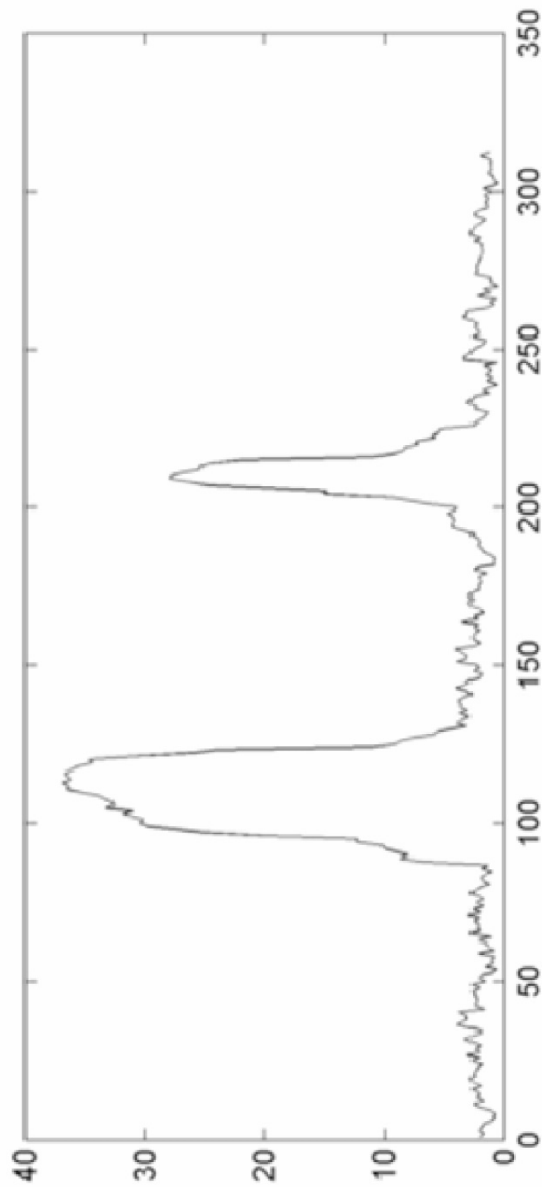


FIG.1A

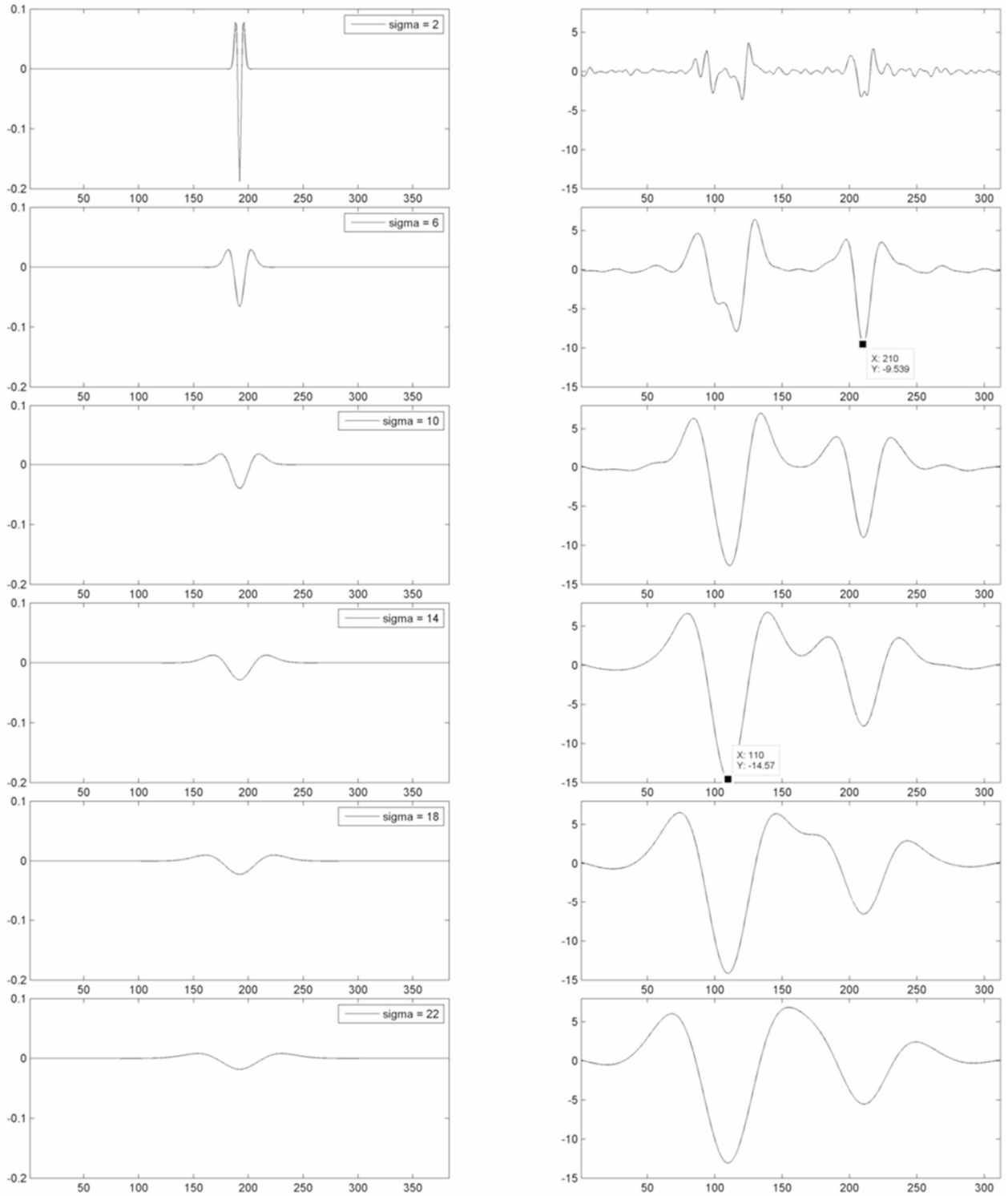


FIG.1B



FIG.2A

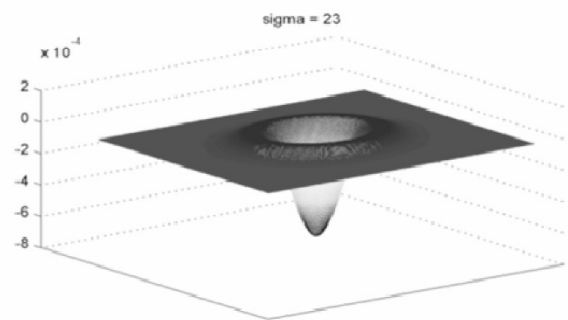
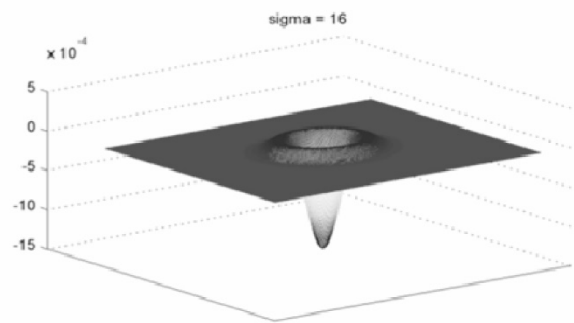
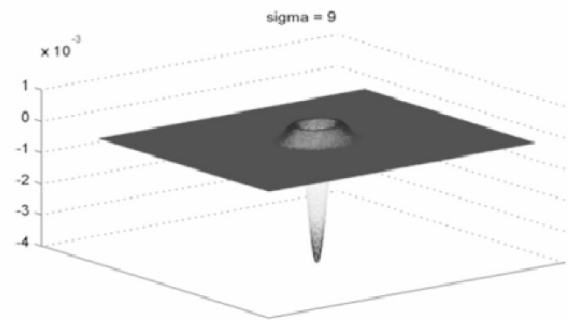
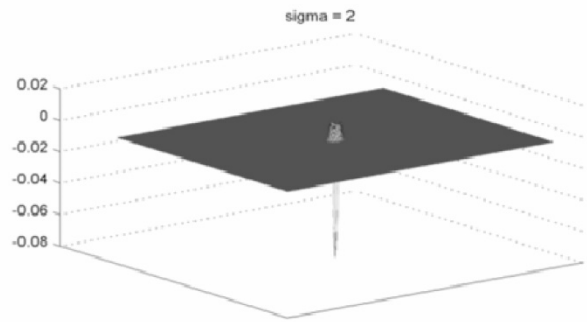
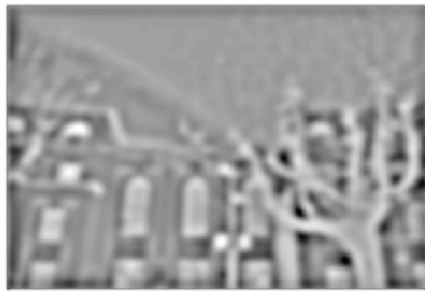


FIG.2B

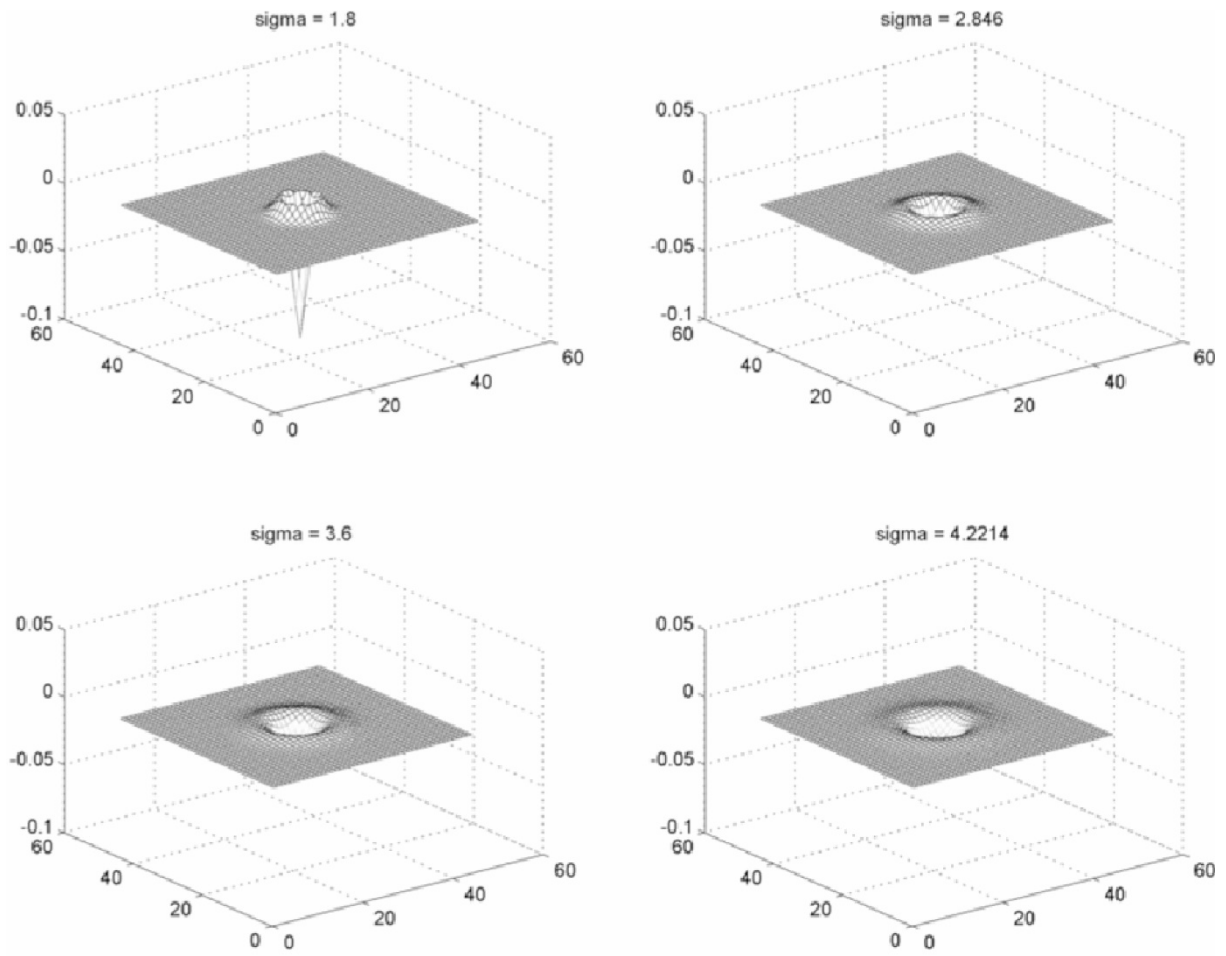


FIG.3A

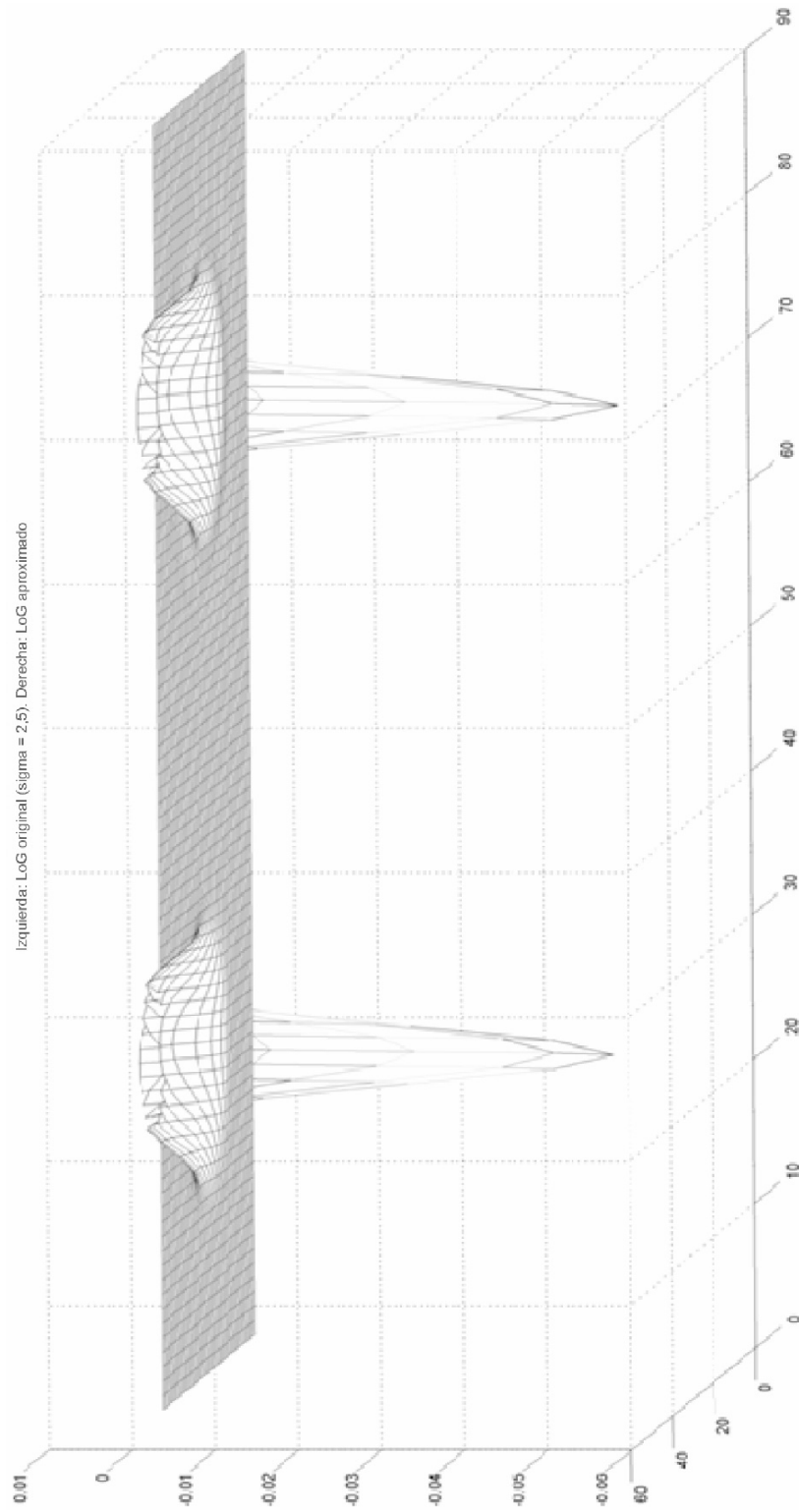


FIG.3B

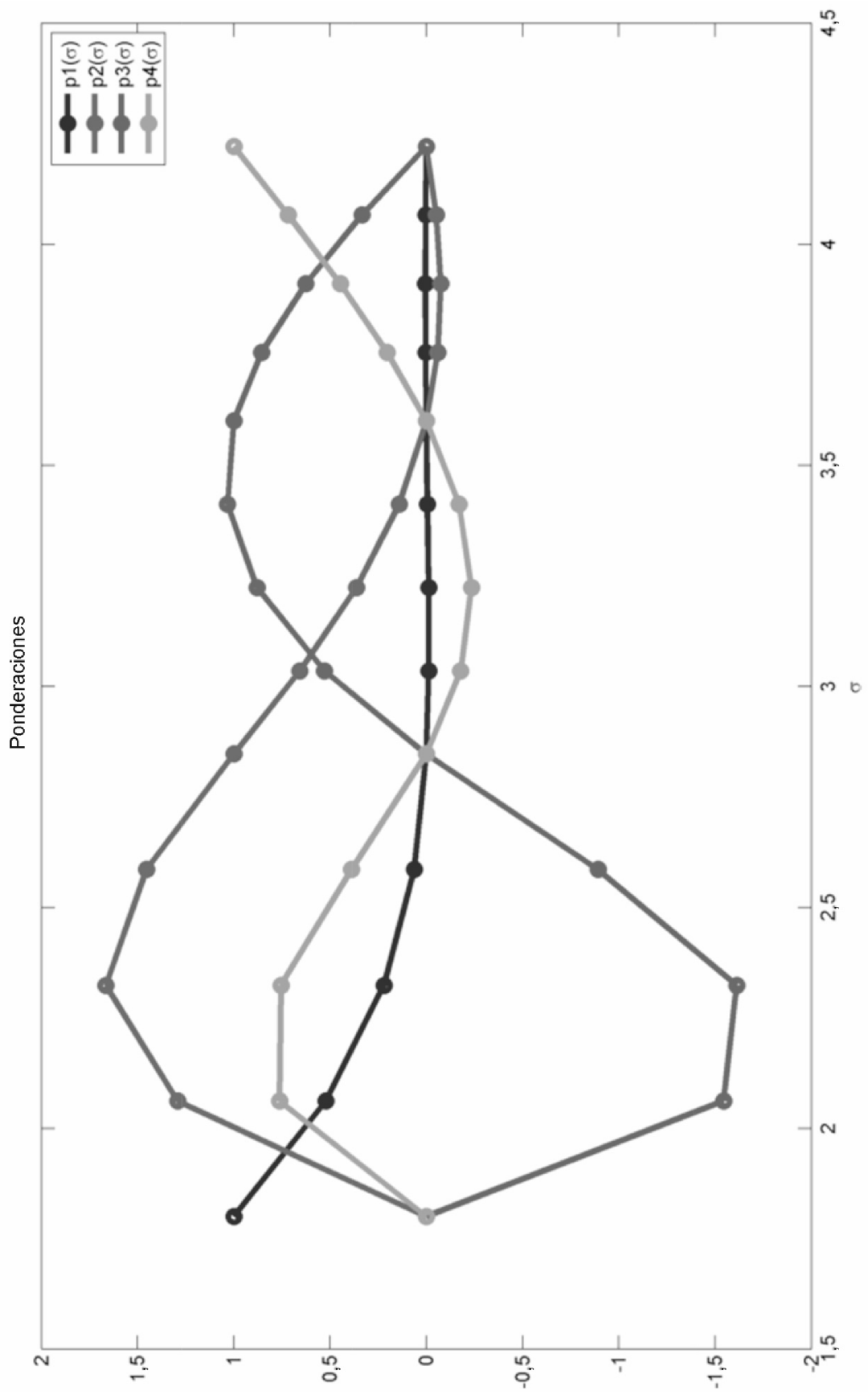


FIG.3C

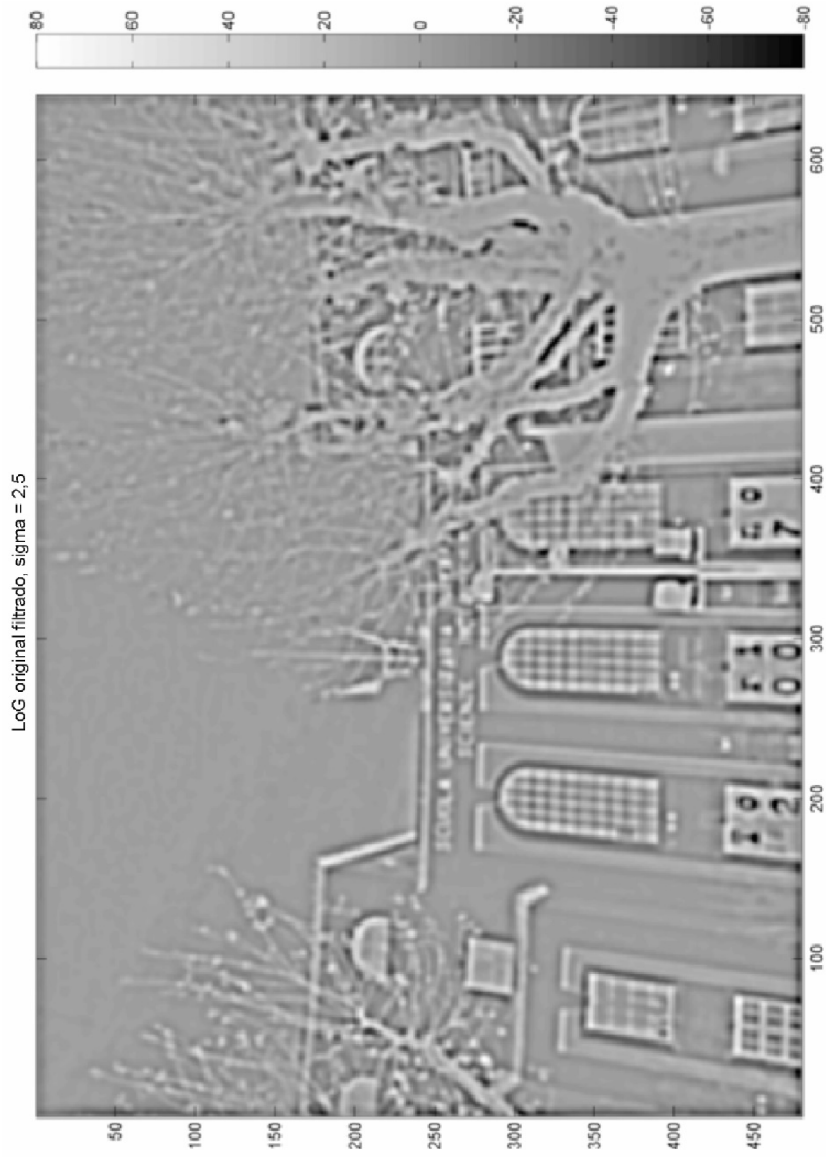


FIG.4A

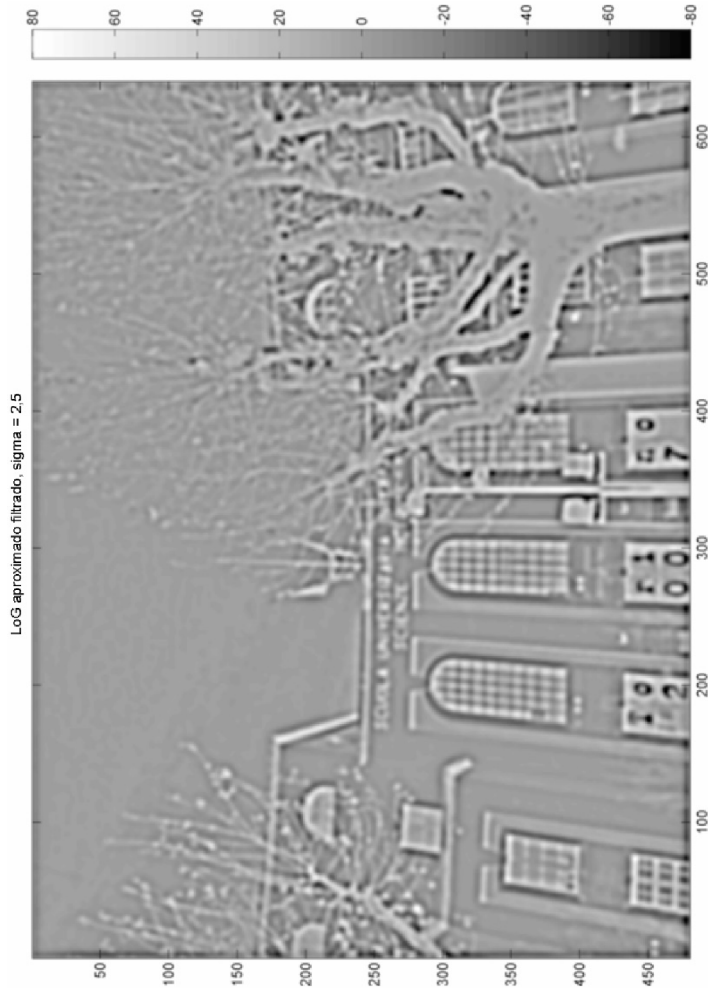


FIG.4B

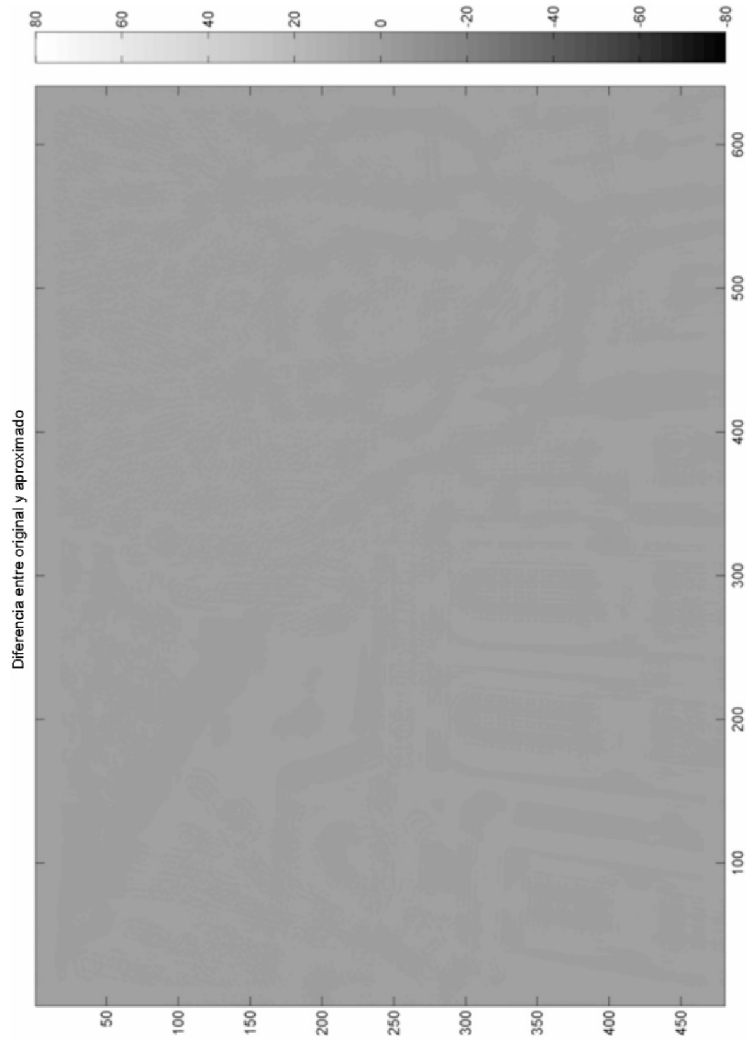


FIG.4C

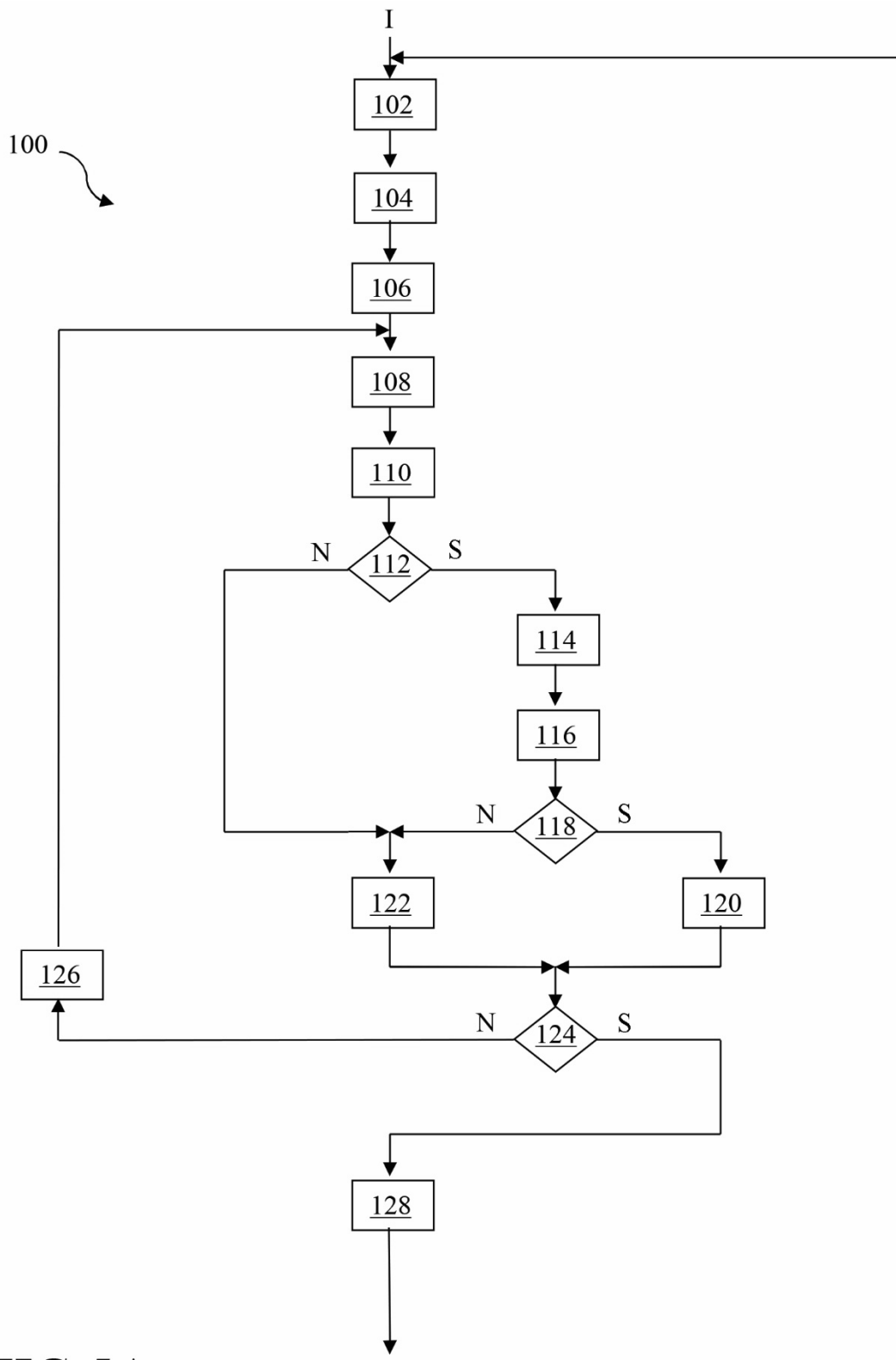


FIG.5A

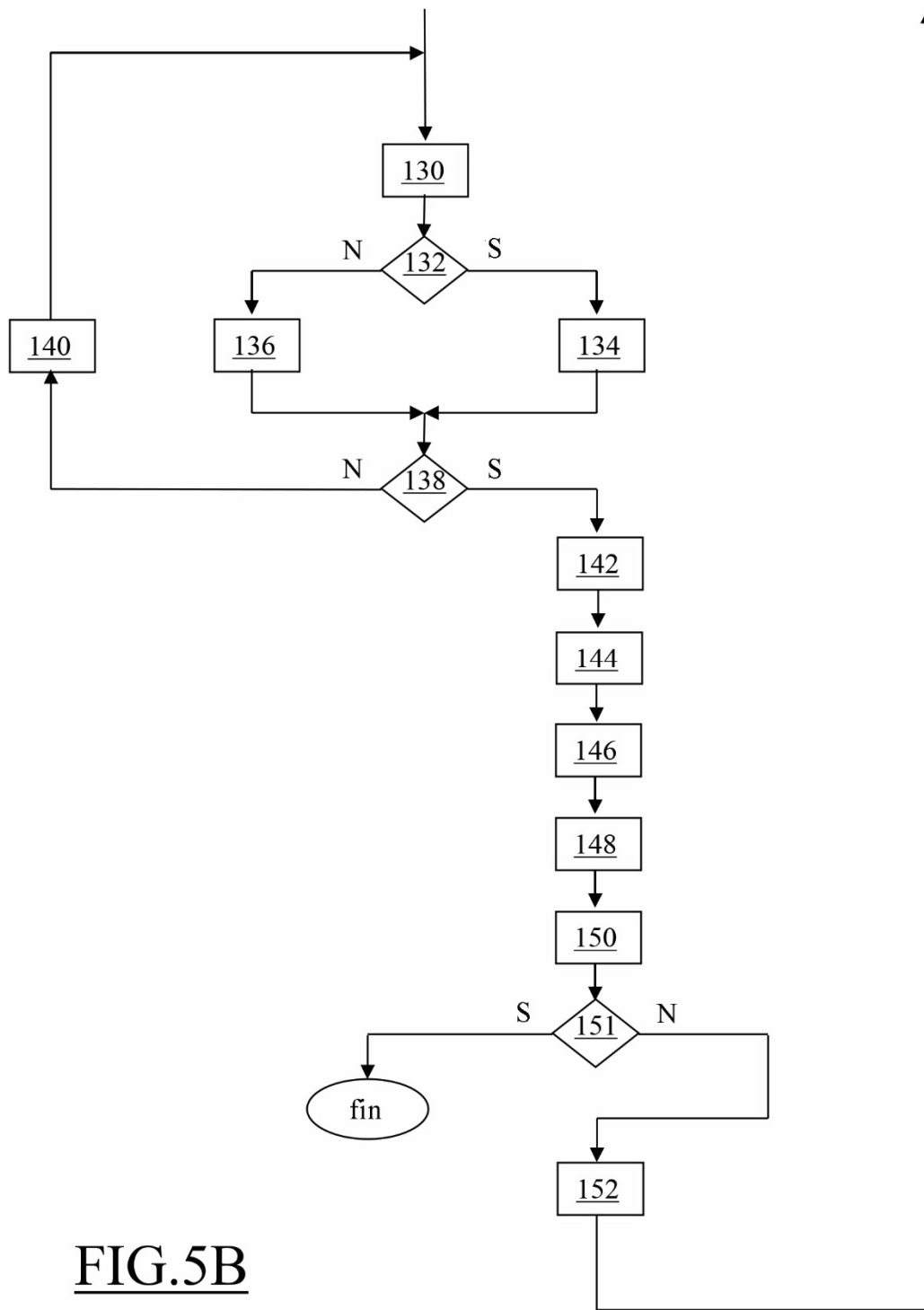


FIG.5B

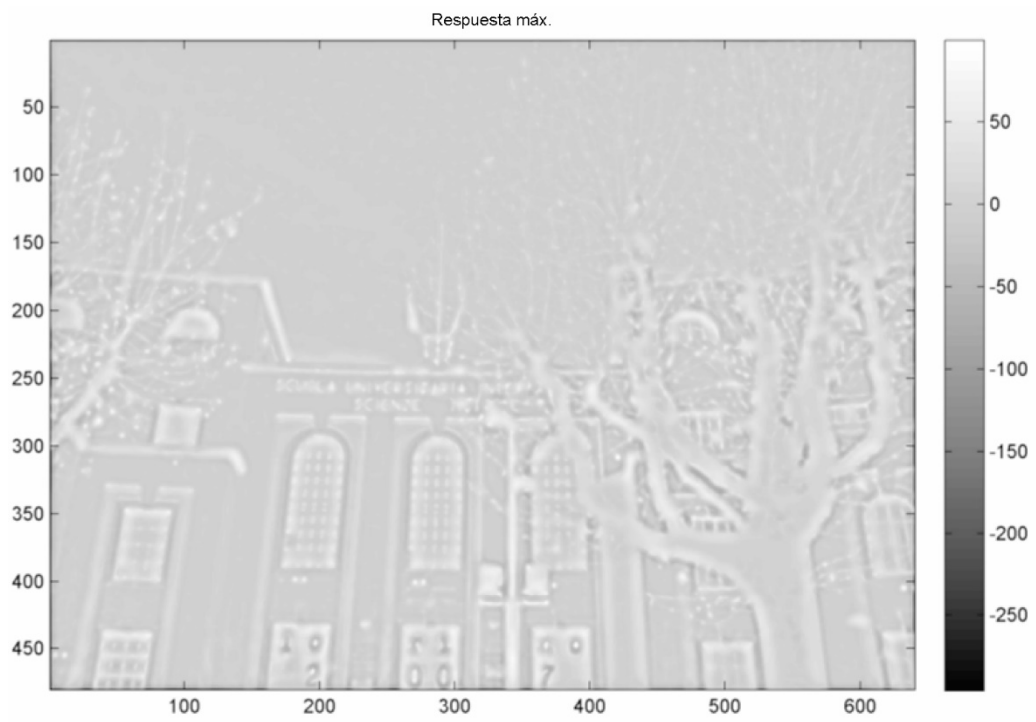


FIG.6A

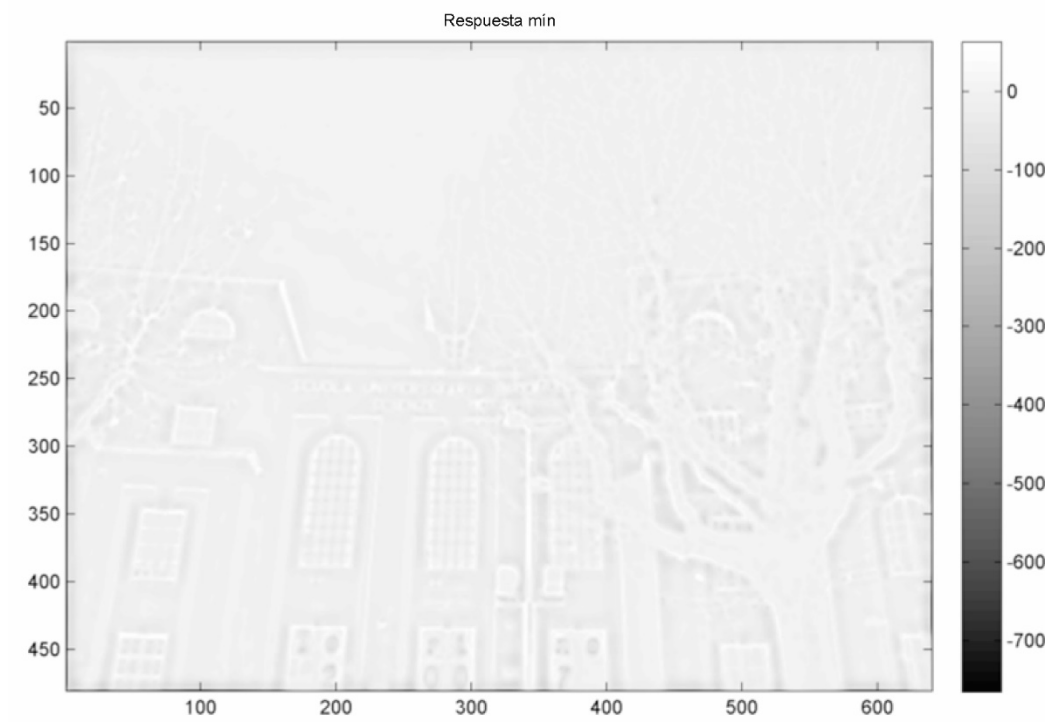


FIG.6B

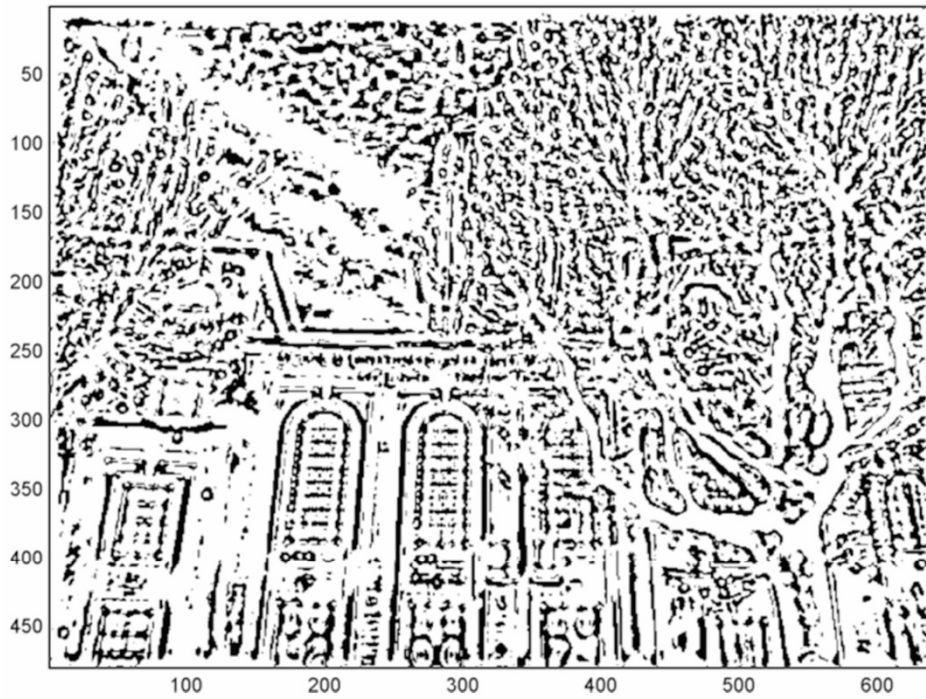


FIG.6C

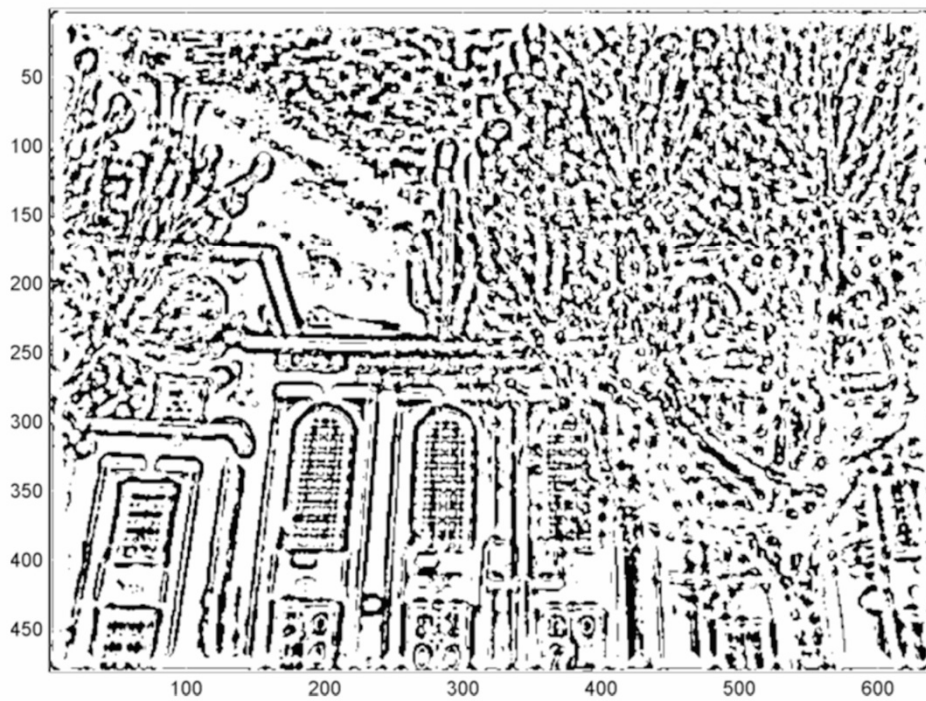


FIG.6D

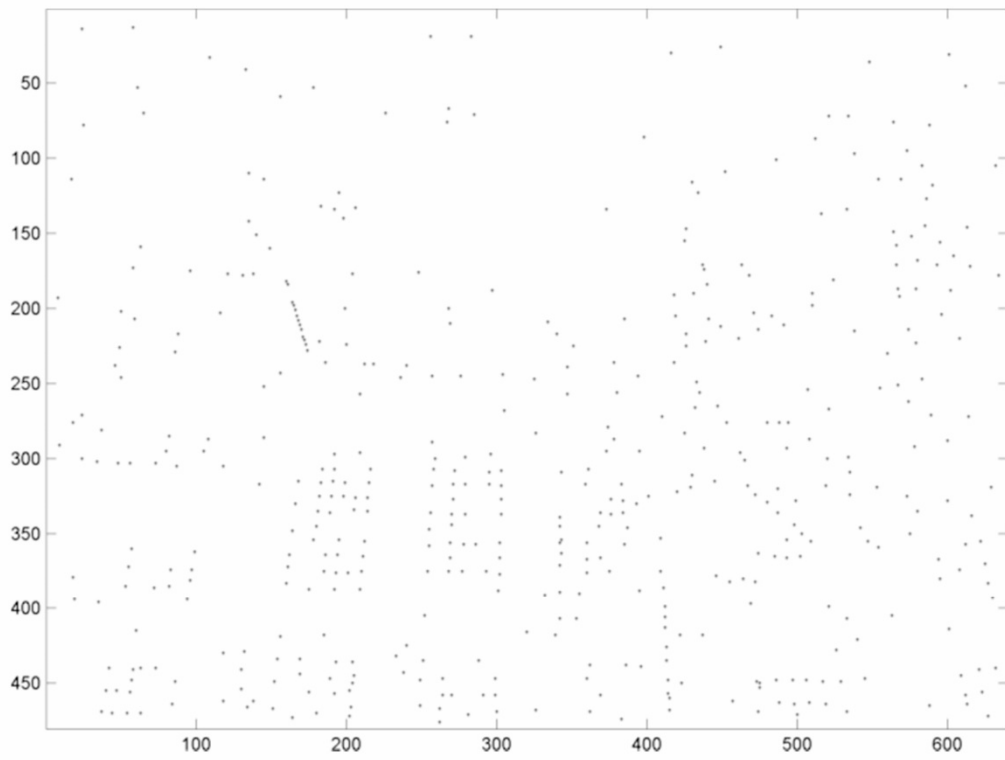


FIG. 7A

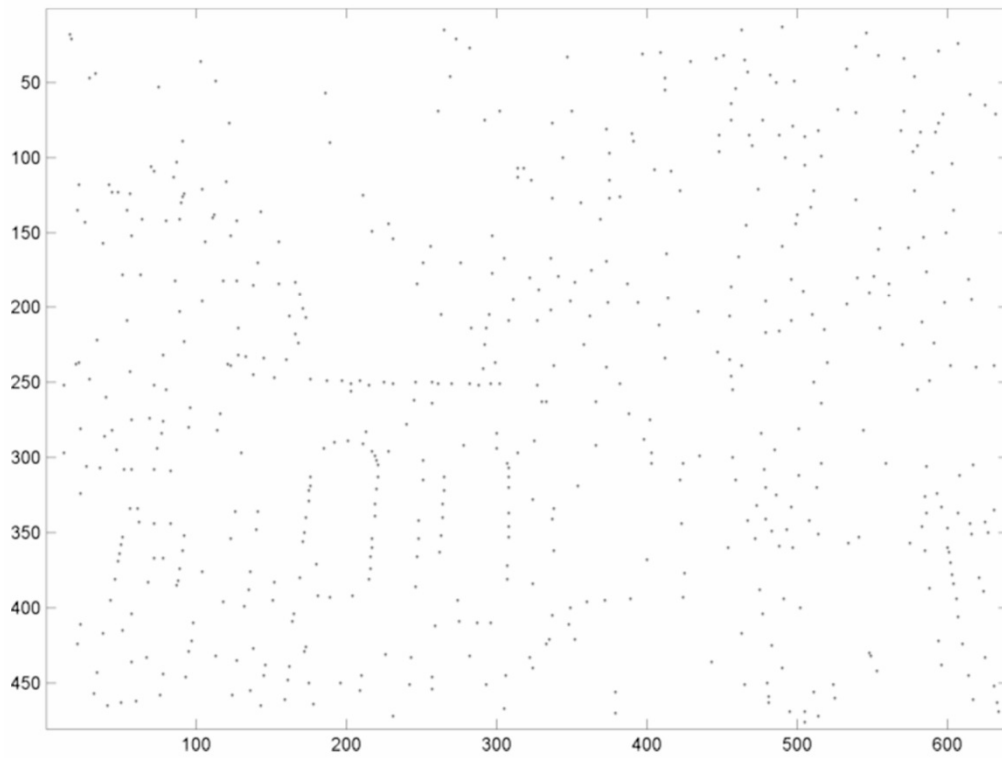


FIG. 7B

Puntos clave de la primera octava



FIG.8A

Puntos clave de todas las octavas

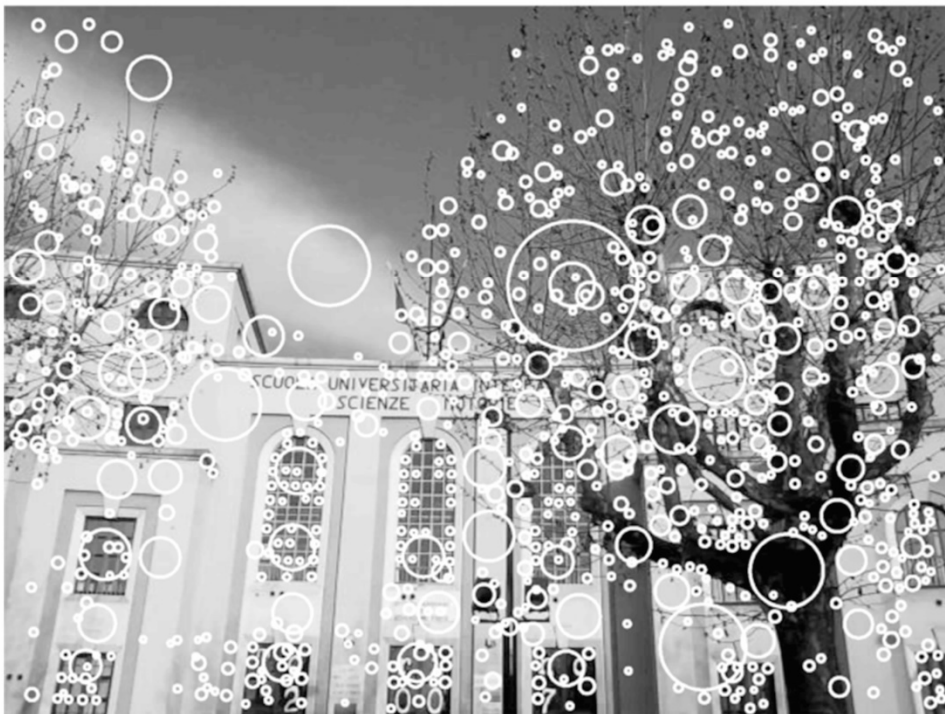


FIG.8B