

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 928 631**

51 Int. Cl.:

G06T 7/00 (2007.01)

G06T 7/90 (2007.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.06.2019** **PCT/EP2019/067180**

87 Fecha y número de publicación internacional: **16.01.2020** **WO20011549**

96 Fecha de presentación y número de la solicitud europea: **27.06.2019** **E 19732690 (3)**

97 Fecha y número de publicación de la concesión europea: **17.08.2022** **EP 3821396**

54 Título: **Proceso implementado por ordenador sobre una imagen de una muestra biológica**

30 Prioridad:

11.07.2018 EP 18182961

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

21.11.2022

73 Titular/es:

VILNIAUS UNIVERSITETAS (25.0%)

Universiteto g. 3

01513 Vilnius, LT;

UNIVERSITÉ DE CAEN NORMANDIE (25.0%);

INSTITUT NATIONAL DE LA SANTÉ ET DE LA

RECHERCHE MÉDICALE (25.0%) y

CENTRE REGIONAL FRANCOIS BACLESSE

(25.0%)

72 Inventor/es:

POULAIN, LAURENT;

LAURINAVICIUS, ARVYDAS;

PLANCOULAINE, BENOÎT y

ELIE, NICOLAS

74 Agente/Representante:

DURAN-CORRETJER, S.L.P

ES 2 928 631 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

Proceso implementado por ordenador sobre una imagen de una muestra biológica

5 **Sector técnico de la invención**

La presente invención se refiere a un proceso implementado por ordenador para clasificar píxeles de una imagen digital de una muestra biológica.

10 **Estado de la técnica anterior**

Se conocen procesos para el tratamiento automatizado de imágenes de una muestra biológica. Sin embargo, la clasificación automatizada de los píxeles, por ejemplo, en función de su color, habitualmente es insatisfactoria. Por ejemplo, muy frecuentemente, los píxeles se clasifican erróneamente como teñidos mientras que están en un área que realmente está sin teñir.

El documento de MEAS-YEDID V et al.: "Automatic color space selection for biological image segmentation", PATTERN RECOGNITION, 2004. ICPR 2004. PROCEEDINGS OF THE 17TH INTERNATIONAL CONFERENCE ON CAMBRIDGE, REINO UNIDO, 23 A 26 DE AGOSTO, 2004, PISCATAWAY, NUEVA JERSEY, EE. UU. IEEE, LOS ALAMITOS, CALIFORNIA, EE. UU., vol. 3, 23 de agosto de 2004 (23/08/2004), páginas 514 a 517, describe una selección de espacio de color automática para segmentación de imágenes biológicas.

25 **Características de la invención**

Por lo tanto, un objetivo de la presente invención es dar a conocer un proceso para el tratamiento automatizado de imágenes de una muestra biológica con mayor precisión en la clasificación de los píxeles.

Según un primer aspecto, un objetivo de la invención es un proceso para clasificar píxeles de una imagen digital de una muestra biológica, y que comprende las etapas de:

- (a) recibir una primera imagen digital que representa la muestra biológica, con el fin de obtener un valor X_1 asociado con cada píxel de la primera imagen digital, y obtener por lo menos otro valor X_i para cada píxel, de la primera imagen digital y/o de una imagen digital adicional, siendo i un entero tal que $2 \leq i \leq n$ y siendo n un entero tal que $n \geq 2$;
- (b) cambiar de espacio con la multiplicación matricial

$$I = A \times X$$

para obtener una segunda imagen digital, siendo X una matriz de n filas y 1 columna que incluye los valores X_1 a X_n , siendo I una matriz de m filas y 1 columna que incluye los valores I_1 a I_m , siendo m un entero igual a 2 o superior, siendo I_1 a I_m valores asociados con cada píxel de la segunda imagen digital, siendo A una matriz de m filas y n columnas;

- (c) determinar una condición de clasificación basada en los valores en I_1 y en los valores en I_j de por lo menos algunos píxeles en la segunda imagen digital con $j > 1$ y $j \leq m$; y
- (d) clasificar los píxeles de la segunda imagen digital que cumplan la condición de clasificación en una primera clase;

en el que la etapa (c) comprende:

- determinar un conjunto de una o varias distribuciones, comprendiendo el conjunto por lo menos una distribución de un número de píxeles en la segunda imagen digital en función de I_1 , y comprendiendo el conjunto por lo menos una distribución de un número de píxeles en la segunda imagen digital en función de I_j , con $j > 1$ y $j \leq m$,
- ajustar la o las distribuciones del conjunto para determinar por lo menos un parámetro de ajuste, y
- determinar la condición de clasificación a partir del por lo menos un parámetro de ajuste.

En el proceso según la invención, varios colores de la primera imagen digital o por lo menos un color de la primera imagen digital y por lo menos un color de la imagen digital adicional se combinan en una matriz común X, para cada píxel de la primera imagen digital. Por tanto, esta matriz sintetiza información relevante en una herramienta que es matemáticamente sencilla de utilizar, en un primer espacio. La matriz X se transforma entonces para obtener la información en un segundo espacio donde la información es más fácil de manejar. La matriz I se obtiene a partir de esa transformación. A continuación, se realiza un análisis de frecuencia, que implica uno o varios ajustes de una o varias distribuciones en por lo menos dos coeficientes de la matriz I (es decir, dos dimensiones en el segundo espacio), con el fin de determinar la condición de clasificación. La condición de clasificación se comprueba para los píxeles de la segunda imagen digital, por ejemplo, considerados uno tras otro, para clasificarlos (o no) en la primera clase.

En la invención, es muy interesante que la determinación de la condición de clasificación utiliza información que proviene de varias dimensiones en el segundo espacio. De hecho, este cruce de diferentes fuentes de información hace posible obtener una condición de clasificación que es especialmente fiable.

Los valores X_i e I_j de los píxeles se pueden denominar "colores", incluso si no son colores reales en todas las realizaciones de la invención. Se puede considerar que la matriz X define el píxel en un primer espacio de color y la matriz I define el píxel en un segundo espacio de color. El cambio de espacio entre los valores X_i de los píxeles y los valores I_j de los píxeles se puede denominar un cambio de espacio de color.

Cuando se obtiene por lo menos un X_i para $i \geq 2$ a partir de la imagen digital adicional, los píxeles de la primera imagen digital y de la imagen digital adicional se corresponden de tal modo que hay un solapamiento entre las estructuras u objetos de la primera imagen digital y de la imagen digital adicional. Aplica lo mismo cuando se utilizan más de dos imágenes como entrada para el proceso. Cuando la primera imagen digital y la imagen digital adicional se utilizan para determinar los valores de X, la segunda imagen digital es una combinación de información que se origina a partir de la primera imagen digital y la imagen digital adicional.

En el procedimiento según la invención, la determinación de la condición de clasificación tiene en cuenta por lo menos una distribución en función de por lo menos uno de los elementos de la matriz I. La o las distribuciones están comprendidas en el conjunto de una o varias distribuciones. Una distribución del conjunto es función de I_1 , $f(I_1)$. Una distribución del conjunto es función de I_j ($m \geq j > 1$), donde dicha distribución puede ser la misma que la distribución que es función de I_1 [y se puede escribir $f(I_1, I_2)$] o dicha distribución puede ser otra distribución, que no es función de I_1 , $g(I_j)$ con $m \geq j > 1$. Dichas una o varias distribuciones pueden ser función de otros coeficientes de la matriz I.

La utilización de una distribución hace posible tener en cuenta información de todo el intervalo de valores de color. Por ejemplo, se puede utilizar una distribución unidimensional en I_1 y una distribución unidimensional en I_2 o I_3 o en cualquier I_j para $m \geq j > 1$; o se puede utilizar una distribución bidimensional en I_1 e I_2 o I_3 o en cualquier I_j para $m \geq j > 1$.

La condición de clasificación implica preferentemente los valores del píxel considerado en varias dimensiones de I y valores de referencia en dichas dimensiones de I.

Preferentemente, la condición de clasificación comprende una condición que se refiere a los valores del píxel considerado en I_1 e I_2 conjuntamente. Por ejemplo, la condición de clasificación puede ser que

$$(I_1 - I_1^0)^2 + (I_2 - I_2^0)^2 < R^2$$

donde I_1^0 , I_2^0 y R son valores de referencia obtenidos en la etapa (c) y $X = (I_1, I_2)$ es el píxel considerado.

El ajuste hace posible habitualmente determinar aproximadamente intervalos de color correspondientes al color o a la tinción de los píxeles. El parámetro de ajuste cuantifica estos intervalos de color. La condición de clasificación preferentemente refleja la separación entre los intervalos de color.

En el marco del presente documento, la muestra biológica es preferentemente una sección histológica o una pluralidad de secciones histológicas en serie, por ejemplo, separadas por 2 a 5 mm. También puede ser, por

ejemplo, un tejido biológico, una muestra de sangre, una o varias células, etc. El proceso según la invención se puede utilizar en histología, citología, hematología, ...

5 La determinación de un conjunto de una o varias distribuciones, comprendiendo el conjunto por lo menos una distribución de un número de píxeles en la segunda imagen digital en función de I_1 , y comprendiendo el conjunto por lo menos una distribución de un número de píxeles en la segunda imagen digital en función de I_j , con $m \geq j > 1$, se puede denominar análisis de frecuencia.

10 El proceso según la invención es especialmente sencillo de implementar en un software. También es especialmente rápido cuando se realiza por un dispositivo electrónico, tal como un ordenador.

15 El procedimiento puede incluir otra etapa adicional, por ejemplo, cualquier etapa adicional de tratamiento de imagen, especialmente entre las etapas (a) y (b). Es posible realizar una etapa de aumentar o disminuir la resolución entre las etapas (a) y (b). Las etapas (b) a (d) se pueden realizar entonces sobre píxeles de otro tamaño diferente al de los píxeles de la imagen tomada originalmente.

Las etapas del proceso se realizan preferentemente en el siguiente orden: (a), (b), (c) y (d). Sin embargo, es posible que, por ejemplo, parte de la etapa (d) se realice antes que parte de la etapa (c).

20 El procedimiento según la invención puede incluir, en cualquier etapa, una selección de por lo menos una parte de la primera imagen digital, la imagen digital adicional o la segunda imagen digital como la o las "partes de interés". El proceso preferentemente continúa solo con la o las partes seleccionadas. La selección se realiza preferentemente entre las etapas (a) y (b). Por ejemplo, en una realización de la invención, una etapa entre las etapas (a) y (b) o antes de la etapa (a) incluye que un biólogo o un software seleccione en la primera imagen digital una o varias áreas de interés. Esto puede ser útil para eliminar las áreas que son claramente sanas y los artefactos (pliegues, puntos, ...). Preferentemente, las etapas (b), (c) y (d) se realizan entonces solo en el o las áreas de interés.

30 En una realización de la invención, la selección de por lo menos una parte de la primera o de la segunda imagen digital como la o las "partes de interés" entre las etapas (a) y (b) se basa en los valores de color RGB de los píxeles. Esta puede incluir, por ejemplo, una determinación de umbrales en función del valor de color de los píxeles en azul, verde o rojo. Esto es especialmente interesante para un carcinoma en el que las áreas de interés, que corresponden al epitelio, se pueden seleccionar por un proceso de determinación de umbrales en una distribución de los valores de color azul.

35 La condición de clasificación se determina preferentemente basándose en los valores en I_1 y en los valores en I_j con $m \geq j > 1$ de todos los píxeles en la segunda imagen digital.

40 En el marco del presente documento, una imagen digital es una matriz de píxeles. Algunas imágenes digitales pueden ser imágenes digitales en color en las que tres valores de color están asociados con cada píxel.

Preferentemente, cada píxel está presente en la primera imagen digital y en la segunda imagen digital pero tiene diferentes valores en la primera y la segunda imágenes digitales.

45 En la primera imagen, tres valores están asociados con cada píxel: un valor en rojo, un valor en verde y un valor en azul. Se puede decir que la primera imagen está en un primer espacio de color que es un espacio de color rojo-verde-azul (RGB). En la segunda imagen, tres valores están asociados con cada píxel: un valor en I_1 , un valor en I_2 y un valor en I_3 . Se puede decir que la segunda imagen está en un segundo espacio de color. El segundo espacio de color no es un espacio de color RGB. El segundo espacio de color se puede denominar un espacio de color Ohta.

50 El valor de color I_1 se puede denominar "intensidad". El valor de color I_2 se puede denominar "rojo contra azul". El valor de color I_3 se puede denominar "magenta contra verde".

55 En el marco del presente documento, la condición de clasificación puede comprender varias condiciones específicas. Preferentemente, la condición de clasificación se cumple si se cumplen todas las condiciones específicas.

En el marco del presente documento, una distribución es preferentemente un histograma.

En una realización de la invención, las etapas (a), (b), (c) y (d) están completamente automatizadas.

5

El proceso según la invención es especialmente apropiado para muestras biológicas que están "teñidas", es decir, que están coloreadas por cualquier proceso de tinción o que están coloreadas naturalmente. El proceso de tinción se puede basar, por ejemplo, en tinción de proteínas, ADN, ARN, diferenciación de células, ... El proceso de tinción puede implicar coloración por tinte y/o puede implicar fluorescencia y/o luz visible, por ejemplo, para hematoxilina, eosina, azafrán. El proceso de tinción puede implicar compuestos cromógenos. El proceso de tinción puede implicar reacciones de anticuerpos/antígenos.

10

La tinción se puede realizar con cualquier proceso de tinción de biología o histología, en particular, histoquímica. La tinción se realiza preferentemente con un aparato de tinción. Por ejemplo, la tinción puede ser en azul (núcleos contrateñidos con hematoxilina) y marrón (proteínas de varios orígenes inmunodetectadas con sistemas de peroxidasa/DAB convencionales). Por ejemplo, la tinción podría permitir la detección de células proliferantes (detección de ki67 o ciclina A), células apoptóticas (caspasa 3 escindida, citoqueratina 18 escindida, ...), antígenos de membrana (receptores de factor de crecimiento, receptores de adhesión, receptores de muerte, ...), proteínas citoplásmicas tales como proteínas de la familia Bcl-2, antígenos nucleares (factores de transcripción, ...), etc.

15

20

En una realización de la invención, la muestra biológica es una sección histológica tomada en una porción de tejido donde se sospecha cáncer.

25

En la etapa (a) se puede utilizar cualquier proceso fotográfico digital, por ejemplo, utilizando una cámara o un escáner. La imagen representa preferentemente la muestra biológica completa.

30

En una realización de la invención, la primera clase incluye píxeles clasificados como representando un primer tipo de estructura o un primer tipo de objeto, por ejemplo, un compartimento tumoral. Se considera que los píxeles que no están clasificados en la primera clase no representan el primer tipo de estructura u objeto. Por ejemplo, la primera clase de píxeles puede corresponder a píxeles "etiquetados", "coloreados" o "teñidos", mientras que los píxeles que no pertenecen a la primera clase se consideran como píxeles "no coloreados", píxeles "no teñidos" o píxeles "de fondo".

35

También es posible que la primera clase de píxeles corresponda a píxeles coloreados o teñidos en un primer color, mientras que los píxeles que no pertenecen a la primera clase pueden ser no coloreados, no teñidos o de otro color, por ejemplo, teñidos en otro color.

40

En una realización de la invención, también se realiza un proceso de separación de la primera imagen digital en un compartimento estromal y un compartimento tumoral, por ejemplo, en paralelo con el proceso para clasificar píxeles según la invención. Dicho proceso de separación se describe, por ejemplo, en la Patente WO2004084135. Una realización de la invención incluye el proceso para clasificar píxeles según la invención y una etapa de determinar los píxeles de la primera clase que pertenecen al compartimento tumoral.

45

Después del proceso según la invención, los píxeles de la primera clase se pueden utilizar para un análisis adicional. Este análisis adicional puede comprender una asignación de los píxeles, preferentemente los píxeles de la primera clase, a macropíxeles. Los macropíxeles son preferentemente hexagonales. La asignación es preferentemente una teselación. En una realización de la invención, el análisis adicional comprende una asignación de los píxeles de la primera clase y que pertenecen al compartimento tumoral a macropíxeles.

50

Entonces se puede determinar un histograma de macropíxeles. Este se puede utilizar, por ejemplo, para la clasificación de tumores o para el estudio de heterogeneidad de tumores.

55

En una realización de la invención, la etapa (c) comprende determinar una condición de clasificación adicional y la etapa (d) comprende clasificar en una segunda clase píxeles de la segunda imagen digital que cumplen la condición de clasificación adicional. Por ejemplo, la condición de clasificación puede corresponder a un primer intervalo de valores de color en el espacio de color de la segunda imagen y la condición de clasificación

adicional puede corresponder a un segundo intervalo de valores de color en el espacio de color de la segunda imagen.

5 En una realización de la invención, el conjunto de una o varias distribuciones comprende una distribución que es función de I_1 y función de I_j con $m \geq j > 1$. Esta distribución tiene preferentemente por lo menos dos dimensiones. Por ejemplo, esta distribución puede estar ajustada por una suma de gaussianas en m dimensiones.

10 En una primera realización de la invención, el conjunto de una o varias distribuciones comprende varias distribuciones que son función de I_1 y función de I_j con $m \geq j > 1$. Por ejemplo, el conjunto puede comprender cuatro distribuciones que dependen de I_1 e I_2 o de I_1 , I_2 e I_3 .

15 En una segunda realización de la invención, el conjunto de una o varias distribuciones comprende una primera distribución que es función de I_1 y una segunda distribución que es función de I_j con $m \geq j > 1$. La primera y la segunda distribuciones pueden ser unidimensionales.

En una realización de la invención, la condición de clasificación para un píxel considerado se refiere al valor del píxel considerado en por lo menos uno de I_1 , I_2 e I_3 .

20 Esto hace posible clasificar fácilmente los píxeles a partir de sus valores en el segundo espacio de color. El píxel "considerado" es el píxel que se clasifica actualmente.

25 Preferentemente, la condición de clasificación implica un umbral en por lo menos uno de I_1 , I_2 e I_3 o un umbral en combinación de por lo menos dos de I_1 , I_2 e I_3 . En una realización de la invención, la condición de clasificación es que los píxeles tienen que tener un valor en I_1 mayor que un umbral, habiéndose determinado dicho umbral a partir de una distribución en I_1 .

30 En una realización de la invención, la condición de clasificación se refiere al valor del píxel considerado en I_1 y al valor del píxel considerado en el por lo menos un I_j con $m \geq j > 1$. En otras palabras, se utilizan las mismas dimensiones para los uno o varios ajustes y para la clasificación.

35 Preferentemente, la condición de clasificación implica un umbral en I_1 y un umbral en por lo menos uno de I_2 e I_3 , o un umbral en combinación de I_1 , I_2 e I_3 . En una realización de la invención, la condición de clasificación es que los píxeles tienen que tener un valor en I_1 mayor que un primer umbral y un valor en I_2 mayor que un segundo umbral, habiéndose determinado el primero a partir de una distribución en I_1 y habiéndose determinado dicho primero a partir de una distribución en I_2 .

40 En una realización de la invención, se utiliza un parámetro de calibración empírico en la determinación de la condición de clasificación, habiéndose determinado dicho parámetro de calibración empírico mostrando muestras biológicas de calibración, preparadas del mismo modo que la muestra biológica, a profesionales de la biología.

45 Puede haber varios parámetros de calibración empíricos. Preferentemente hay dos parámetros de calibración empíricos. Más preferentemente, hay un parámetro de calibración correspondiente a I_1 y un parámetro de calibración correspondiente a I_j , siendo I_j el valor de color utilizado para determinar la condición de clasificación. La determinación del parámetro de calibración empírico se puede haber realizado mucho antes de la utilización del proceso reivindicado. Preferentemente, las condiciones de preparación, incluyendo, por ejemplo, una herramienta de tinción y un tinte de tinción, son idénticas en la preparación de las muestras biológicas de calibración y en la preparación de la muestra biológica analizada por el proceso reivindicado. Las "muestras biológicas de calibración" también se pueden denominar "muestras biológicas ficticias". La utilización de un parámetro de calibración empírico hace posible tener en cuenta la experiencia de profesionales de la biología sobre qué píxeles consideran "coloreados" y "no coloreados". Por lo tanto, esto hace posible calibrar el proceso de clasificación según la invención. Por ejemplo, si la condición de clasificación implica un umbral, la utilización de un parámetro de calibración empírico puede afectar al valor de este umbral. Preferentemente, hay un parámetro de calibración por valor de color utilizado en la etapa (c). Más preferentemente, hay un parámetro de calibración en I_1 y un parámetro de calibración en I_2 o I_3 .

Para una sección histológica teñida, por ejemplo, la utilización de un parámetro de calibración empírico hace posible tener en cuenta la experiencia de profesionales de histología sobre qué píxeles consideran "teñidos" y "no teñidos". En dicho caso, la "preparación" preferentemente incluye una tinción.

5 Si se utiliza un parámetro de calibración empírico en la determinación de la condición de clasificación, preferentemente este se tiene en cuenta en la determinación de la condición de clasificación a partir del por lo menos un parámetro de ajuste.

10 En una realización de la invención, la o las distribuciones del conjunto se ajustan con una o varias distribuciones gaussianas.

15 El por lo menos un parámetro de ajuste preferentemente comprende la o las medias de la o las distribuciones gaussianas. El por lo menos un parámetro de ajuste preferentemente comprende la o las medias y la desviación típica de la o las distribuciones gaussianas. La determinación de la condición de clasificación a partir del por lo menos un parámetro de ajuste preferentemente implica una suma ponderada de la media y la desviación típica. Si se utiliza un parámetro de calibración empírico en la determinación de la condición de clasificación, preferentemente este se tiene en cuenta en por lo menos una ponderación de dicha suma ponderada.

20 En una realización de la invención, el ajuste utiliza un modelo de mezcla gaussiana. Este modelo es capaz de adaptarse especialmente bien a cambios de histograma.

En una realización de la invención, el ajuste utiliza un modelo de máxima verosimilitud.

25 El ajuste también se puede realizar con otro procedimiento, como el procedimiento de media cuadrática. Los inventores han descubierto que utilizar el procedimiento de máxima verosimilitud proporciona mejores resultados que otros procedimientos porque las estimaciones de parámetros se realizan a partir de un conjunto de píxeles submuestreados. Este está especialmente bien adaptado para el portaobjetos virtual amplio (WSI).

30 Por ejemplo, el ajuste se puede realizar con una suma de funciones gaussianas bidimensionales.

En una realización de la invención, el proceso comprende clasificar en una segunda clase píxeles de la segunda imagen digital que no cumplen la condición de clasificación.

35 Por ejemplo, estos píxeles se pueden considerar como no teñidos o se pueden considerar como teñidos en otro color o se puede considerar que tienen otro color. En una realización de la invención, los píxeles que no cumplen la condición de clasificación se clasifican entre una pluralidad de clases, por ejemplo, basándose en una condición de clasificación adicional. Por tanto, es posible clasificar los píxeles en la primera clase (primer color o primer color de tinción), una segunda clase (segundo color o segundo color de tinción) y una tercera clase (tercer color o sin tinción).

40 En una realización de la invención, el proceso comprende una determinación de la matriz A a partir de un análisis de componentes principales, a partir de un análisis de componentes independientes o a partir de un análisis factorial sobre la matriz X. De hecho, es interesante centrarse en las dimensiones que representan información relevante, con el fin de realizar los uno o varios ajustes sobre el coeficiente de la matriz I que focalice la información más relevante. Otra posibilidad es que la matriz A sea conocida antes del proceso.

45 En una realización de la invención, denominada la realización de Ohta, $n=3$, $m=3$, X_1 es un valor rojo R en la primera imagen digital, X_2 es un valor verde G en la primera imagen digital y X_3 es un valor azul B en la primera imagen digital, y

50

$$A = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \end{bmatrix}$$

En la realización de Ohta de la invención, la ecuación utilizada para el cambio de espacio de color de la etapa (b) se basa en un análisis de componentes principales, tal como mostró Ohta en sus publicaciones. Por lo tanto, la separación entre colores es mucho más fuerte en el espacio de color de la segunda imagen digital, que se puede denominar "segundo espacio de color" o "espacio de color de Ohta". Por lo tanto, el proceso según la invención es especialmente apropiado para discriminar píxeles de diferentes colores, lo que aumenta su precisión de clasificación.

Además, debido a que la condición de clasificación se basa en la intensidad de los píxeles en el segundo espacio de color, I_1 , y en por lo menos otro valor de color, I_2 , I_3 , de los píxeles. La utilización de I_1 es especialmente apropiada para discriminar entre señal y fondo. La utilización de I_2 o I_3 es especialmente apropiada para discriminar entre colores. Por lo tanto, la utilización de I_1 e I_2 o I_3 ofrece informaciones complementarias. Por tanto, la combinación de las utilidades de I_1 con I_2 o I_3 conduce a una fuerte mejora de la precisión del proceso.

Los inventores han mostrado que la combinación de I_1 con I_2 y/o I_3 en el espacio de Ohta es, por ejemplo, especialmente eficiente para diferenciar entre colores marrón y azul, y puede ser muy eficiente para diferenciar también entre otros colores.

En la primera imagen, tres valores están asociados con cada píxel: un valor en rojo, un valor en verde y un valor en azul. Se puede decir que la primera imagen está en un primer espacio de color que es un espacio de color rojo-verde-azul (RGB). En la segunda imagen, tres valores están asociados con cada píxel: un valor en I_1 , un valor en I_2 y un valor en I_3 . Se puede decir que la segunda imagen está en un segundo espacio de color. El segundo espacio de color no es un espacio de color RGB. El segundo espacio de color se puede denominar un espacio de color Ohta.

El valor de color I_1 se puede denominar "intensidad". El valor de color I_2 se puede denominar "rojo contra azul". El valor de color I_3 se puede denominar "magenta contra verde".

En una realización de la invención, la primera imagen digital muestra por lo menos parte de la muestra biológica teñida según una primera técnica de tinción. Esta parte de la muestra biológica es preferentemente una primera sección histológica. La primera imagen digital está preferentemente en RGB. Es posible que la primera imagen digital sea nivel de grises, por ejemplo, si es una imagen en fluorescencia o una imagen en polarización.

En el marco del presente documento, una primera técnica de tinción preferentemente difiere de una segunda técnica de tinción en los productos químicos utilizados para la coloración.

En una realización de la invención, la imagen digital adicional muestra por lo menos parte de la muestra biológica teñida según una segunda técnica de tinción. En dicho caso, el proceso según la invención hace posible combinar información proporcionada por varias técnicas de tinción para obtener una condición o criterio común para la clasificación de píxeles. La parte de la muestra biológica visible en la imagen digital adicional puede ser una segunda sección histológica, o la primera sección histológica que ha sido no teñida, y teñida según la segunda técnica de tinción.

En una realización de la invención, la primera imagen digital se obtiene a partir de una imagen óptica, una imagen en fluorescencia o una imagen en polarización. Cualquier otra técnica de formación de imágenes también es posible dentro del marco de la invención. La imagen digital adicional se puede obtener a partir de una imagen óptica, una imagen en fluorescencia o una imagen en polarización. Con "obtenida a partir de" se quiere decir que la imagen digital puede ser la imagen óptica, la imagen en fluorescencia, la imagen en polarización, una combinación de por lo menos dos de estas imágenes, o una imagen que contiene información que proviene de por lo menos una de estas imágenes. Una "imagen óptica" preferentemente significa una imagen en luz visible, más preferentemente en color.

En una realización de la invención, X_2 se obtiene a partir de la imagen digital adicional y por lo menos un X_i con $i > 2$ se obtiene a partir de otra imagen digital. Por ejemplo, X_2 , X_3 y X_4 pueden basarse en la imagen digital adicional y X_5 , X_6 y X_7 pueden basarse en la otra imagen digital.

En una realización de la invención, X_1 se obtiene a partir de una combinación de la primera imagen digital con

otra imagen digital y por lo menos un X_i para $2 \leq i \leq n$ se obtiene a partir de una combinación de la imagen digital adicional con la otra imagen digital. Por ejemplo, la información de la primera imagen digital y la otra imagen digital se puede combinar para proporcionar X_1 , X_2 y X_3 , y la información de la imagen digital adicional y la otra imagen digital se puede combinar para proporcionar X_4 , X_5 y X_6 .

5

La condición de clasificación determinada en la etapa (c) se puede reutilizar posteriormente para otras muestras biológicas.

10

La invención también se refiere a un dispositivo de procesamiento de datos que comprende medios para llevar a cabo un proceso según cualquier realización de la invención.

15

La invención también se refiere a un programa informático que comprende instrucciones que, cuando el programa es ejecutado por un dispositivo de procesamiento de datos, hacen que el dispositivo de procesamiento de datos lleve a cabo un proceso según cualquier realización de la invención. El dispositivo de procesamiento de datos puede ser un ordenador, una tableta, un teléfono inteligente, un controlador, un procesador y puede estar incluido en cualquier tipo de dispositivo.

20

La invención también se refiere a un medio legible por ordenador que comprende instrucciones que, cuando son ejecutadas por un dispositivo de procesamiento de datos, hacen que el dispositivo de procesamiento de datos lleve a cabo un proceso según cualquier realización de la invención.

Breve descripción de las figuras

25

Para una mejor comprensión de la presente invención, a continuación se hará referencia, a modo de ejemplo, a los dibujos adjuntos, en los que:

30

- la figura 1a ilustra un diagrama de flujo que ilustra un primer ejemplo de un proceso para clasificar píxeles según una realización de la invención;

- la figura 1b ilustra un diagrama de flujo que ilustra un segundo ejemplo de un proceso para clasificar píxeles según una realización de la invención;

- la figura 2 es un diagrama de flujo que ilustra una realización ejemplar de la determinación de la condición de clasificación según la invención;

- la figura 3 es un diagrama de flujo que ilustra una realización ejemplar de una determinación de por lo menos un parámetro de calibración empírico según la invención;

35

- la figura 4 es un diagrama de flujo que ilustra una realización ejemplar de las etapas de determinación de la condición de clasificación y clasificar los píxeles en una primera clase según la invención;

- la figura 5 ilustra una primera imagen digital en un primer ejemplo de la primera realización de la presente invención;

- la figura 6a es una ampliación de una parte de la primera imagen digital de la figura 5;

40

- la figura 6b ilustra una parte de la segunda imagen digital que corresponde a la parte de la primera imagen digital ilustrada en la figura 6a;

- la figura 7 ilustra el primer histograma en el primer ejemplo de la primera realización de la presente invención;

- la figura 8 ilustra el segundo histograma en el primer ejemplo de la primera realización de la presente invención;

45

- la figura 9a ilustra una imagen enviada a un profesional de histología en el primer ejemplo de la primera realización de la presente invención;

- la figura 9b ilustra la misma imagen que la figura 9a, en la que los diferentes tipos de puntos indican si el profesional de histología ha considerado el punto indicado como teñido o no teñido;

- la figura 10a ilustra límites entre áreas "teñidas" y "no teñidas" que se obtienen sin los parámetros de calibración empíricos;

50

- la figura 10b ilustra límites entre áreas "teñidas" y "no teñidas" que se obtienen con la utilización de los parámetros de calibración empíricos tal como se describe en la presente memoria;

- la figura 11 ilustra una imagen resultante de una clasificación según la presente invención, en el primer ejemplo de la primera realización de la presente invención;

55

- la figura 12 ilustra una máscara binaria en el primer ejemplo de la primera realización de la presente invención;

- la figura 13 ilustra el resultado de la aplicación de la máscara binaria de la figura 12 en la imagen resultante de la clasificación ilustrada en la figura 11;

- la figura 14 ilustra un teselado de la figura 13 con macropíxeles hexagonales;
- la figura 15 ilustra el resultado de un cálculo realizado sobre los macropíxeles de la figura 14;
- la figura 16 ilustra un histograma que proporciona el número de macropíxeles en función del factor de transmisión;
- 5 - la figura 17 ilustra una primera imagen digital en un segundo ejemplo de la primera realización de la presente invención;
- la figura 18a es una ampliación de una parte de la primera imagen digital de la figura 17;
- la figura 18b ilustra una parte de la segunda imagen digital que corresponde a la parte de la primera imagen digital ilustrada en la figura 18a;
- 10 - la figura 19 ilustra el primer histograma en el segundo ejemplo de la primera realización de la presente invención;
- la figura 20 ilustra el segundo histograma en el segundo ejemplo de la primera realización de la presente invención;
- la figura 21a ilustra una imagen enviada a un profesional de histología en el segundo ejemplo de la primera realización de la presente invención;
- 15 - la figura 21b ilustra la misma imagen que la figura 21a, en la que los diferentes tipos de puntos indican si el profesional de histología ha considerado el punto indicado como teñido o no teñido;
- la figura 22a ilustra límites entre áreas "teñidas" y "no teñidas" que se obtienen sin los parámetros de calibración empíricos;
- 20 - la figura 22b ilustra límites entre áreas "teñidas" y "no teñidas" que se obtienen con la utilización de los parámetros de calibración empíricos tal como se describe en la presente memoria;
- la figura 23 ilustra una imagen resultante de la clasificación según la presente invención, en el segundo ejemplo de la primera realización de la presente invención;
- la figura 24 ilustra una máscara binaria en el segundo ejemplo de la primera realización de la presente invención;
- 25 - la figura 25 ilustra el resultado de la aplicación de la máscara binaria de la figura 24 en la imagen resultante de la clasificación ilustrada en la figura 23;
- la figura 26 ilustra un teselado de la figura 25 con macropíxeles hexagonales;
- la figura 27 ilustra el resultado de un cálculo realizado sobre los macropíxeles de la figura 26;
- 30 - la figura 28 ilustra un histograma que proporciona el número de macropíxeles en función del factor de transmisión;
- la figura 29 ilustra una primera imagen digital en un ejemplo de la segunda realización de la presente invención;
- la figura 30a es una ampliación de una parte de la primera imagen digital de la figura 29;
- 35 - la figura 30b ilustra la parte de la segunda imagen digital que corresponde a la parte de la primera imagen digital ilustrada en la figura 30a;
- la figura 31a ilustra una distribución bidimensional en el ejemplo de la segunda realización de la presente invención;
- la figura 31b ilustra cuatro curvas gaussianas que ajustan la distribución bidimensional de la figura 31a;
- 40 - la figura 32 ilustra el resultado de un análisis con las cuatro elipses de la figura 31b;
- la figura 33a es una ampliación de un área determinada en la figura 29;
- la figura 33b es una ampliación de la misma área de la figura 33a, en la figura 32;
- la figura 34 ilustra un teselado de la figura 32 con macropíxeles hexagonales;
- la figura 35 ilustra los macropíxeles con un color que corresponde a su cantidad de colágeno;
- 45 - la figura 36 ilustra los macropíxeles con un color que corresponde a su cantidad de citoplasma;
- la figura 37 ilustra una distribución de los macropíxeles en función de su cantidad de colágeno;
- la figura 38 ilustra una distribución de los macropíxeles en función de su cantidad de citoplasma;
- la figura 39 ilustra una primera imagen digital en un ejemplo de la presente invención; y
- 50 - la figura 40 ilustra una segunda imagen digital que corresponde a la parte de la primera imagen digital ilustrada en la figura 39.

Descripción de la invención

- La presente invención se describirá con respecto a realizaciones particulares y haciendo referencia a ciertos dibujos, pero la invención no se limita a esto. Los dibujos descritos son solo esquemáticos y son no limitativos. En los dibujos, el tamaño de algunos de los elementos puede estar exagerado y no dibujado a escala, con fines ilustrativos.

Asimismo, los términos primero, segundo, tercero y similares en la descripción y en las reivindicaciones, se utilizan para distinguir entre elementos similares y no necesariamente para describir un orden secuencial o cronológico. Los términos son intercambiables bajo circunstancias apropiadas y las realizaciones de la invención pueden funcionar en otras secuencias distintas a las descritas o ilustradas en la presente memoria.

Asimismo, las diversas realizaciones, aunque se denominan "preferentes" deben interpretarse como maneras ejemplares en las que se puede implementar la invención, en lugar de limitar el alcance de la invención.

El término "comprende", utilizado en las reivindicaciones, no debe interpretarse como limitado a los elementos o etapas enumerados después del mismo; no excluye otros elementos o etapas. Debe interpretarse que este especifica la presencia de las características, números enteros, etapas o componentes indicados, tal como se mencionan, pero que no impide la presencia o adición de una o varias características, números enteros, etapas o componentes adicionales, o grupos de los mismos. Por tanto, el alcance de la expresión "un dispositivo que comprende A y B" no debe estar limitado a dispositivos que consisten solo en los componentes A y B, sino que, con respecto a la presente invención, los únicos componentes enumerados del dispositivo son A y B, y, además, debe interpretarse que la reivindicación incluye equivalentes de esos componentes.

En las figuras, elementos idénticos o análogos pueden denominarse por un mismo número.

La figura 1a ilustra un diagrama de flujo que ilustra un ejemplo de un proceso 100 para clasificar píxeles según una realización de la invención en la que no se utiliza ninguna imagen digital adicional. Preferentemente, una muestra biológica 101 se tiñe 201. El resultado es una muestra biológica teñida 102. A continuación, se crea una imagen de por lo menos parte de la muestra biológica 101 o de por lo menos parte de la muestra biológica teñida 102 para proporcionar 202 a un dispositivo que comprende medios de procesamiento de datos una primera imagen digital 103 que representa por lo menos parte de la muestra biológica 101 con o sin tinción.

La primera imagen digital 103 comprende píxeles. La primera imagen digital 103 está en un primer espacio con n dimensiones ($n \geq 2$), y cada píxel tiene n valores denominados X_1 a X_n . En la realización de la invención en la que solo se utiliza la primera imagen digital 103 para determinar X, los valores X_2 a X_n se obtienen 202b junto con el valor X_1 , en la etapa 202, porque estos valores están incluidos en la primera imagen digital 103.

El primer espacio es preferentemente un espacio de color, y, más específicamente, un espacio de color RGB. Por tanto, tres valores de color están asociados con cada píxel de la primera imagen digital 103: un valor en rojo R ($=X_1$), un valor en verde G ($=X_2$) y un valor en azul B ($=X_3$).

Según el proceso 100, a continuación, se realiza un cambio 203 de espacio (o de espacio de color), por ejemplo, comenzando a partir de los valores RGB de los píxeles de la primera imagen digital 103. El cambio 203 de espacio de color comprende determinar una segunda imagen digital 104 en un segundo espacio de color realizando, para cada píxel, la siguiente multiplicación matricial

$$I=A \times X$$

La matriz A es la misma para todos los píxeles. Los componentes I_1 a I_m de la matriz I son los valores de los píxeles en la segunda imagen digital 104, con $m \geq 2$.

En la realización de Ohta de la invención, los valores de color de los píxeles en la segunda imagen digital 104 son I_1 , I_2 e I_3 . Estos se determinan por:

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

A continuación, se determina 204 una condición de clasificación 105. La determinación 204 se basa en:

- los valores en I_1 de por lo menos algunos píxeles de la segunda imagen digital 104, junto con
- los valores de por lo menos un I_j (con $m \geq j > 1$) de por lo menos algunos píxeles de la segunda imagen digital 104.

En la realización de Ohta de la invención, esta determinación 204 se basa en:

- los valores en I_1 de por lo menos algunos píxeles de la segunda imagen digital 104, junto con
- los valores en I_2 y/o los valores en I_3 de por lo menos algunos píxeles de la segunda imagen digital 104.

Los valores en I_1 son especialmente útiles para distinguir los píxeles de interés. Los valores en I_2 o I_3 son especialmente útiles para discriminar el color o la tinción. La combinación de I_1 con I_2 o I_3 hace posible determinar con precisión los píxeles de interés coloreados o teñidos.

La condición de clasificación 105 preferentemente se determina para representar la condición que solo cumplen los píxeles coloreados o teñidos, mientras que los píxeles sin teñir no la cumplen.

A continuación, los píxeles de la segunda imagen digital 104 que cumplen la condición de clasificación 105 se clasifican entonces en una primera clase 106. En otras palabras, estos píxeles se seleccionan como píxeles que pertenecen a la primera clase 106.

En una realización de la invención, los píxeles de la primera clase 106 se utilizan para formar una tercera imagen digital, que es una máscara, preferentemente una imagen binaria o una imagen en escala de grises. Un primer color en esta imagen binaria corresponde a píxeles de la primera clase 106 y un segundo color en esta imagen binaria corresponde a píxeles que no son de la primera clase 106.

La invención también se refiere a un proceso para el tratamiento de imágenes biológicas que incluye el proceso de clasificación 100 y por lo menos uno de los siguientes procesos:

- una determinación de por lo menos un parámetro de calibración empírico 113 (véase, por ejemplo, la figura 3);
- un proceso de separación de la primera imagen digital en un compartimento estromal y un compartimento tumoral;
- una asignación de los píxeles, preferentemente de los píxeles de la primera clase 106, a macropíxeles; y
- un análisis adicional basado en los píxeles de la primera clase 106.

En una realización de la invención, el proceso de separación de la primera imagen digital en un compartimento estromal y un compartimento tumoral determina una máscara binaria que se puede superponer sobre la tercera imagen binaria y solo se seleccionan los píxeles que son del compartimento tumoral y de la primera clase.

Preferentemente, solo aquellos píxeles seleccionados se distribuyen en macropíxeles y/o se utilizan para cualquier análisis adicional.

El análisis adicional preferentemente implica un análisis estadístico en función del número de macropíxeles que tienen unas propiedades determinadas.

La figura 1b ilustra un diagrama de flujo que ilustra un ejemplo de un proceso 100 para clasificar píxeles según una realización de la invención en la que se utiliza una imagen digital adicional 103b. Preferentemente, una primera sección histológica de una muestra biológica 101 se tiñe 201 según una primera técnica de tinción. El resultado es una primera sección histológica teñida de una muestra biológica 102. A continuación, se crea una imagen de por lo menos parte de la primera histológica para proporcionar 202 a un dispositivo que comprende medios de procesamiento de datos una primera imagen digital 103 que representa por lo menos parte de la primera sección histológica. La primera imagen digital 103 comprende píxeles. Cada píxel tiene, por lo menos, en la primera imagen digital 103, un valor, denominado X_1 .

A continuación, o en paralelo, una segunda sección histológica de la muestra biológica 101 se tiñe 201b según una segunda técnica de tinción. El resultado es una segunda sección histológica teñida de una muestra biológica 102b. A continuación, se crea una imagen de por lo menos parte de la segunda sección histológica para proporcionar al dispositivo que comprende medios de procesamiento de datos la imagen digital adicional 103b que representa por lo menos parte de la segunda sección histológica. La imagen digital adicional 103b incluye valores asociados con sus píxeles. Hay por lo menos un valor X_i ($i > 1$) asociado con cada píxel de la imagen digital adicional 103b que es obtenido 202b por el dispositivo cuando el dispositivo recibe la imagen digital adicional 103b. Preferentemente, cada píxel de la imagen digital adicional 103b corresponde a un solo píxel en la primera imagen digital 103. Para cada píxel, hay una matriz de valores de dicho píxel $X = (X_1, \dots, X_n)^T$ que combina el o los valores obtenidos de la primera imagen digital 103 y el o los valores obtenidos de la imagen digital adicional 103b.

El proceso utiliza entonces la matriz X para proceder a las etapas 203 a 205 de una forma similar a la descrita haciendo referencia a la figura 1a.

Es posible, dentro del marco de la presente invención, que la primera imagen digital 103 y/o la imagen digital adicional 103b se obtengan a partir de una imagen óptica (en RGB, por ejemplo), una imagen en fluorescencia o en polarización. Realmente, cualquier tipo de imagen sería apropiada.

El proceso de la figura 1a considera solo la primera imagen digital 103 como entrada para el tratamiento digital. El proceso de la figura 1b considera la primera imagen digital 103 y la imagen digital adicional 103b como entradas para el tratamiento digital. Es posible, dentro del marco de la presente invención, que también se utilice por lo menos otra imagen digital para obtener por lo menos un componente de X . Por ejemplo, la primera imagen digital 103 y la otra imagen digital se pueden combinar para obtener X_1 , mientras que la imagen digital adicional 103b y la otra imagen digital se pueden combinar para obtener X_2 . Dicha combinación de imágenes se puede realizar realmente a través de un análisis de componentes principales con el fin de extraer directamente los fragmentos de información más relevantes de las imágenes.

La figura 2 es un diagrama de flujo que ilustra una realización ejemplar de la determinación 204 de la condición de clasificación 105 según la invención.

Según esta realización ejemplar, se determina 211 un conjunto 111 de una o varias distribuciones. El conjunto 111 comprende por lo menos una distribución de un número de píxeles en la segunda imagen digital 104 en función de I_1 . El conjunto 111 comprende por lo menos una distribución de un número de píxeles en la segunda imagen digital 104 en función de I_j con $m \geq j > 1$, por ejemplo, I_2 o I_3 .

Por ejemplo, el conjunto 111 puede comprender un primer histograma que representa el número de píxeles en la segunda imagen digital en función del valor de I_1 y un segundo histograma que representa el número de píxeles en la segunda imagen digital en función del valor de I_2 .

A continuación, en la realización ilustrada en la figura 2, la o las distribuciones del conjunto 111 se ajustan 212 con funciones paramétricas con el fin de extraer el o los parámetros 112 de estas funciones paramétricas.

A continuación, en la realización ilustrada en la figura 2, además del o de los parámetros 112, la determinación 204 tiene en cuenta 213 por lo menos un parámetro de calibración empírico 113 para calcular la condición de clasificación 105.

En una realización de la invención, se realizan las etapas 211 y 212, pero no se realiza la etapa 213. En otras palabras, la condición de clasificación 105 se determina sin la utilización de un parámetro de calibración empírico 113 pero con un ajuste 212 de una distribución.

La figura 3 es un diagrama de flujo que ilustra una realización ejemplar de una determinación 300 de por lo menos un parámetro de calibración empírico 113 según la invención.

Según esta realización ejemplar, las muestras biológicas de calibración 301 se tiñen o preparan 401 del mismo modo que la muestra biológica 102 analizada por el proceso de clasificación 100. Estas se convierten en muestras biológicas de calibración teñidas o preparadas 302. En 402 se toman imágenes de estas, en RGB, y

se colocan puntos en algunos píxeles de las imágenes. Las imágenes con puntos 303 se muestran 403 entonces a profesionales de la biología, tales como médicos habituados a interpretar muestras biológicas. Los profesionales de la biología proporcionan una indicación de tinción. La indicación de tinción puede, por ejemplo, decir qué puntos consideran los profesionales de la biología como teñidos/no teñidos o teñidos en un primer color/teñidos en un segundo color/.../no teñidos. El resultado 304 puede tener el formato de una tabla con valores de RGB e indicaciones de tinción o color. Los parámetros de calibración empíricos 113 se determinan 404 entonces basándose en los valores de RGB y en las indicaciones correspondientes de tinción o color. Los puntos de los profesionales se cruzan con las dos máscaras (o imágenes binarias), una para la imagen etiquetada (o imagen teñida/coloreada) y la otra para la imagen no etiquetada. Cuando un punto etiquetado está asociado a su máscara (o imagen binaria) se cuenta como verdadero, de otro modo como falso. Se calcula una puntuación determinando el primer y/o el segundo parámetros de calibración.

La figura 4 es un diagrama de flujo que ilustra una realización ejemplar de las etapas de determinación 204 de la condición de clasificación 105 y clasificación de los píxeles 205 en una primera clase 106 según la invención. En la descripción de la figura 4, se tiene que comprender que el índice J es un entero estrictamente mayor que 1, y menor o igual que m, siendo m el número de elementos en la matriz I.

Un experto comprenderá que algunas de las etapas ilustradas en la figura 4 en paralelo se pueden realizar secuencialmente, o lo contrario.

Según esta realización ejemplar, la determinación 211 del conjunto 111 de una o varias distribuciones comprende una determinación 501 de un primer histograma 601, que es un histograma en I_1 , y una determinación 502 de un segundo histograma 602, que es un histograma en I_j . El conjunto 111 comprende el primer 601 y el segundo 602 histogramas.

A continuación, en la etapa 212, el primer histograma 601 se ajusta utilizando un modelo de mezcla gaussiana con la suma de una primera curva gaussiana de I_1 y una segunda curva gaussiana de I_1 . La primera curva gaussiana de I_1 corresponde preferentemente a una señal teñida/coloreada en la escala de I_1 , y la segunda curva gaussiana de I_1 corresponde al fondo en la escala de I_1 . La primera curva gaussiana de I_1 habitualmente alcanza un valor mucho mayor que la segunda curva gaussiana de I_1 . La primera curva gaussiana de I_1 se escribe como

$$f(I_1) = A_1 \exp\left(-\frac{(I_1 - \mu_1)^2}{2\sigma_1^2}\right)$$

donde μ_1 y σ_1 603 son algunos de los parámetros de ajuste 112.

En la etapa 212, el segundo histograma se ajusta utilizando un modelo de mezcla gaussiana con la suma de una primera curva gaussiana de I_j y una segunda curva gaussiana de I_j . La primera curva gaussiana de I_j corresponde preferentemente a una señal teñida/coloreada en la escala de I_j , y la segunda curva gaussiana de I_j corresponde al fondo en la escala de I_j . La primera curva gaussiana de I_j habitualmente alcanza un valor mucho mayor que la segunda curva gaussiana de I_j . La primera curva gaussiana de I_j se escribe como

$$f(I_j) = A_2 \exp\left(-\frac{(I_j - \mu_j)^2}{2\sigma_j^2}\right)$$

donde μ_j y σ_j 604 son algunos de los parámetros de ajuste 112.

En la etapa 213, un primer parámetro de calibración ITC_1 605 se tiene en cuenta 505 para determinar un primer umbral IT_1 como $IT_1 = \mu_1 + ITC_1 * \sigma_1$, y un segundo parámetro de calibración ITC_j se tiene en cuenta 506 para determinar un segundo umbral como $IT_j = \mu_j + ITC_j * \sigma_j$.

La condición de clasificación 105 consiste en:

- una primera condición específica 607, que es $I_1 > IT_1$, y
- una segunda condición específica 608, que es $I_j > IT_j$.

La etapa de clasificación 205 comprende:

- una comparación 507 de los píxeles de la segunda imagen digital 104 con el primer umbral IT_1 , seleccionando solo los píxeles 609 en los que $I_1 > IT_1$; y
- una comparación 508 de los píxeles seleccionados con el segundo umbral IT_j , clasificando en la primera clase 106 solo los píxeles en los que $I_j > IT_j$.

La figura 4 es un proceso ejemplar de una primera realización de la invención. En la primera realización de la invención, el conjunto de una o varias distribuciones comprende una primera distribución que proporciona el número de píxeles en la segunda imagen digital en función de su valor en I_1 y una segunda distribución que proporciona el número de píxeles en la segunda imagen digital en función de su valor en I_j , con $j > 1$. La primera realización de la invención se ilustra mediante los ejemplos 1 y 2.

Según una segunda realización de la invención, el conjunto de una o varias distribuciones comprende una distribución que proporciona el número de píxeles en la segunda imagen digital en función de su valor en I_1 y como una función de por lo menos uno de sus valores en I_j , con $2 \leq j \leq m$. Esta distribución multidimensional se ajusta preferentemente con una suma de por lo menos dos o tres gaussianas. El modelo de mezcla gaussiana se puede utilizar para el ajuste.

En dos o más dimensiones, las gaussianas resultantes del ajuste se pueden representar por elipses. Pueden ser, por ejemplo, una primera elipse correspondiente a los píxeles teñidos/coloreados, una segunda elipse correspondiente a los píxeles no teñidos/coloreados y una tercera elipse correspondiente al fondo. Las elipses se pueden clasificar por su superficie: la elipse con la mayor superficie tiene la menor prioridad y la elipse con la menor superficie tiene la mayor prioridad. Los parámetros de calibración empíricos se pueden tener en cuenta aumentando o disminuyendo la longitud de los ejes mayor y menor de las elipses mediante una multiplicación. La segunda realización de la invención se ilustra mediante el ejemplo 3.

Ejemplo 1

Las figuras 5 a 16 ilustran un primer ejemplo de la primera realización de la presente invención. En este ejemplo, la tinción de una sección histológica fue realizado con BIM.

La figura 5 ilustra la primera imagen digital 103 (figura 1). La figura 6a es una ampliación de una parte de la primera imagen digital 103 de la figura 5. La figura 6b ilustra la parte de la segunda imagen digital 104 (figura 1) que corresponde a la parte de la primera imagen digital 103 ilustrada en la figura 6a.

La figura 7 ilustra el primer histograma 601 (figura 4) con el ajuste como una suma de la primera curva gaussiana de I_1 701 y la segunda curva gaussiana de I_1 702. La figura 8 ilustra el segundo histograma 602 (figura 4) con el ajuste como una suma de la primera curva gaussiana de I_2 703 y la segunda curva gaussiana de I_2 704.

La figura 9a ilustra la imagen 303 con puntos 705 enviada 403 a un profesional de histología (figura 3). La figura 9b ilustra la imagen 303 con puntos, en la que el tipo de los puntos indica si el profesional de histología ha considerado el punto indicado como teñido (punto gris claro 706) o como no teñido (punto gris oscuro 707). Los puntos que todavía son blancos corresponden a píxeles en los que el profesional de histología no pudo decidir si estaban teñidos o no.

La figura 10a ilustra límites 809a entre áreas "teñidas" y "no teñidas" que se obtienen sin los parámetros de calibración empíricos. La figura 10b ilustra límites 809b entre áreas "teñidas" y "no teñidas" que se obtienen con la utilización de los parámetros de calibración empíricos tal como se describe en la presente memoria. Es especialmente visible en las cajas 708a, 708b, 709a, 709b, que la utilización de los parámetros de calibración empíricos mejora la clasificación.

La figura 11 ilustra una imagen resultante de la clasificación 205: los píxeles en rojo son los píxeles que pertenecen a la primera clase 106, y los píxeles en verde son los píxeles que no pertenecen a la primera clase 106. El proceso que fue aplicado en este ejemplo 1 implicaba una selección de áreas de interés y los píxeles

negros son aquellos que no fueron considerados como parte de las áreas de interés.

La figura 12 ilustra una máscara binaria en la que los píxeles blancos corresponden al compartimento tumoral.

- 5 La figura 13 ilustra el resultado de la aplicación de la máscara binaria de la figura 12 sobre la imagen resultante de la clasificación 205 ilustrada en la figura 11.

La figura 14 ilustra un teselado de la figura 13 con macropíxeles hexagonales.

- 10 La figura 15 ilustra el resultado de un cálculo realizado sobre los macropíxeles de la figura 14. Los colores indican los valores obtenidos por este cálculo.

La figura 16 ilustra un histograma que proporciona el número de macropíxeles en función del factor de transmisión.

15 Ejemplo 2

Las figuras 17 a 28 ilustran un segundo ejemplo de la primera realización de la presente invención. En este ejemplo, la tinción de una sección histológica fue realizada con P-ERK.

- 20 La figura 17 ilustra la primera imagen digital 103 (figura 1). La figura 18a es una ampliación de una parte de la primera imagen digital 103 de la figura 17. La figura 18b ilustra la parte de la segunda imagen digital 104 (figura 1) que corresponde a la parte de la primera imagen digital 103 ilustrada en la figura 18a.

- 25 La figura 19 ilustra el primer histograma 601 (figura 4) con el ajuste como una suma de la primera curva gaussiana de I_1 701 y la segunda curva gaussiana de I_1 702. La figura 20 ilustra el segundo histograma 602 (figura 4) con el ajuste como una suma de la primera curva gaussiana de I_2 703 y la segunda curva gaussiana de I_2 704.

- 30 La figura 21a ilustra la imagen 303 con puntos 705 enviada 403 a un profesional de histología (figura 3). La figura 21b ilustra la imagen 303 con puntos, donde el tipo de los puntos indica si el profesional de histología ha considerado el punto indicado como teñido (punto gris claro 706) o como no teñido (punto gris oscuro 707). Los puntos que todavía son blancos corresponden a píxeles en los que el profesional de biología no pudo decidir si estaban teñidos o no.

- 35 La figura 22a ilustra límites 809a entre áreas "teñidas" y "no teñidas" que se obtienen sin los parámetros de calibración empíricos. La figura 22b ilustra límites 809b entre áreas "teñidas" y "no teñidas" que se obtienen con la utilización de los parámetros de calibración empíricos tal como se describe en la presente memoria. Es especialmente visible en las cajas 708a, 708b, 709a, 709b que la utilización de los parámetros de calibración empíricos mejora la clasificación.
- 40

- La figura 23 ilustra una imagen resultante de la clasificación 205: los píxeles en rojo son los píxeles que pertenecen a la primera clase 106 y los píxeles en verde son los píxeles que no pertenecen a la primera clase 106. El proceso que fue aplicado en este ejemplo 2 implicaba una selección de áreas de interés y los píxeles negros son aquellos que no fueron considerados como parte de las áreas de interés.
- 45

La figura 24 ilustra una máscara binaria en la que los píxeles blancos corresponden al compartimento tumoral.

- 50 La figura 25 ilustra el resultado de la aplicación de la máscara binaria de la figura 24 sobre la imagen resultante de la clasificación 205 ilustrada en la figura 23.

La figura 26 ilustra un teselado de la figura 25 con macropíxeles hexagonales.

- 55 La figura 27 ilustra el resultado de un cálculo realizado sobre los macropíxeles de la figura 26. Los colores indican los valores obtenidos por este cálculo.

La figura 28 ilustra un histograma que proporciona el número de macropíxeles en función del factor de

transmisión.

Ejemplo 3

- 5 Las figuras 29 a 38 ilustran un ejemplo de la segunda realización de la presente invención. En este ejemplo, fue realizada con RGB3 una tinción en una sección histológica, con un tricroma en el que el rojo corresponde a fibras de colágeno, el cian corresponde a citoplasma y el púrpura corresponde a núcleos.
- 10 La figura 29 ilustra la primera imagen digital 103 (figura 1). La figura 30a es una ampliación de una parte de la primera imagen digital 103 de la figura 29. La figura 30b ilustra la parte de la segunda imagen digital 104 (figura 1) que corresponde a la parte de la primera imagen digital 103 ilustrada en la figura 30a.
- 15 La figura 31a ilustra una distribución bidimensional en función de I_1 e I_2 . Esta distribución bidimensional forma un conjunto de una o varias distribuciones. Cada punto 710 representa el valor de un píxel en el segundo espacio de color. La figura 31b ilustra cuatro curvas gaussianas que ajustan la distribución bidimensional de la figura 31a. Una primera elipse 711 corresponde a rojo en el espacio RGB. Una segunda elipse 712 corresponde a azul en el espacio RGB. Una tercera elipse 713 corresponde a verde en el espacio RGB. Una cuarta elipse 714 corresponde a píxeles de fondo.
- 20 La figura 32 ilustra el resultado del análisis con las cuatro elipses con el rojo (primera elipse - núcleos), el azul (segunda elipse - citoplasma), el verde (tercera elipse - colágeno) y el negro para los píxeles de fondo. Esta corresponde a la misma área que la figura 29.
- 25 La figura 33a es una ampliación de un área determinada en la figura 29 y la figura 33b es una ampliación de la misma área en la figura 32.
- La figura 34 ilustra un teselado de la figura 32 con macropíxeles hexagonales.
- 30 La figura 35 ilustra los macropíxeles con un color que corresponde a su cantidad de colágeno.
- La figura 36 ilustra los macropíxeles con un color que corresponde a su cantidad de citoplasma.
- La figura 37 ilustra una distribución de los macropíxeles en función de su cantidad de colágeno.
- 35 La figura 38 ilustra una distribución de los macropíxeles en función de su cantidad de citoplasma.

Ejemplo 4

- 40 La figura 39 ilustra la primera imagen digital 103 (figura 1) mostrando todas las células inmunoteñidas utilizando un anticuerpo antiluciferasa, 48 horas después de transfección con un plásmido de luciferasa.
- La figura 40 ilustra la segunda imagen digital 104 (figura 1) que corresponde a la parte de la primera imagen digital 103 ilustrada en la figura 39.
- 45 Una posible realización de la invención es un proceso (100) para clasificar píxeles de una imagen digital de una muestra biológica (101) y que comprende las etapas de:
- (a) recibir (202) una primera imagen digital (103) que representa una muestra biológica (101), estando asociados un valor rojo R, un valor verde G y un valor azul B con cada píxel de la primera imagen digital (103);
- 50 (b) cambiar (203) de espacio de color con la ecuación

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

para obtener una segunda imagen digital (104), siendo I_1 , I_2 e I_3 valores de color asociados con cada píxel de la segunda imagen digital (104);

(c) determinar (204) una condición de clasificación (105) basada en los valores en I_1 y en los valores en I_2 y/o I_3 de por lo menos algunos píxeles en la segunda imagen digital (104); y

(d) clasificar (205) los píxeles de la segunda imagen digital (104) que cumplan la condición de clasificación (105) en una primera clase (106).

En este proceso, se prefiere que la etapa (c) comprenda:

- determinar (211) un conjunto (111) de una o varias distribuciones, comprendiendo el conjunto (111) por lo menos una distribución de un número de píxeles en la segunda imagen digital (104) en función de I_1 , y comprendiendo el conjunto (111) por lo menos una distribución de un número de píxeles en la segunda imagen digital (104) en función de I_j , con $j = 2$ o 3 ,
- ajustar (212) la o las distribuciones del conjunto (111) para determinar por lo menos un parámetro de ajuste (112), y
- determinar la condición de clasificación (105) a partir del por lo menos un parámetro de ajuste (112).

En otras palabras, la invención se refiere a un proceso implementado por ordenador para analizar automáticamente una imagen digital 103 de una muestra biológica 101. El proceso implica un cambio 203 de un primer espacio de color (por ejemplo, RGB) a un segundo espacio de color (por ejemplo, Ohta). A continuación, se realizan ajustes teniendo en cuenta varias dimensiones del segundo espacio de color para clasificar píxeles. Por ejemplo, los valores de color en intensidad y en por lo menos uno de "rojo contra azul" y "magenta contra verde" se utilizan para seleccionar píxeles teñidos o coloreados.

Aunque la presente invención se ha descrito anteriormente con respecto a realizaciones particulares, se apreciará fácilmente que también son posibles otras realizaciones.

REIVINDICACIONES

1. Proceso implementado por ordenador (100) para clasificar píxeles de una imagen digital de una muestra biológica (101), y que comprende las etapas de:

- (a) recibir (202) una primera imagen digital (103) que representa la muestra biológica (101), con el fin de obtener un valor X_1 asociado con cada píxel de la primera imagen digital (103), y obtener (202b) por lo menos otro valor X_i para cada píxel de la primera imagen digital (103) y/o de una imagen digital adicional (103b), siendo i un entero tal que $2 \leq i \leq n$ y siendo n un entero tal que $n \geq 2$;
- (b) cambiar (203) de espacio con la multiplicación matricial

$$I = A \times X$$

para obtener una segunda imagen digital (104), siendo X una matriz de n filas y 1 columna que incluye los valores X_1 a X_n , siendo I una matriz de m filas y 1 columna que incluye los valores I_1 a I_m , siendo m un entero igual a 2 o superior, siendo I_1 a I_m valores asociados con cada píxel de la segunda imagen digital (104), siendo A una matriz de m filas y n columnas;

(c) determinar (204) una condición de clasificación (105) basada en los valores en I_1 y en los valores en I_j de por lo menos algunos píxeles en la segunda imagen digital (104) con $j > 1$ y $j \leq m$; y

(d) clasificar (205) los píxeles de la segunda imagen digital (104) que cumplan la condición de clasificación (105) en una primera clase (106);

en el que la etapa (c) comprende:

- determinar (211) un conjunto (111) de una o varias distribuciones, comprendiendo el conjunto (111) por lo menos una distribución de un número de píxeles en la segunda imagen digital (104) en función de I_1 y comprendiendo el conjunto (111) por lo menos una distribución de un número de píxeles en la segunda imagen digital (104) en función de I_j , con $j > 1$ y $j \leq m$,
- ajustar (212) la o las distribuciones del conjunto (111) para determinar por lo menos un parámetro de ajuste (112), y
- determinar la condición de clasificación (105) a partir del por lo menos un parámetro de ajuste (112).

2. Proceso, según la reivindicación 1, en el que el conjunto (111) de una o varias distribuciones comprende una distribución que es función de I_1 y función de I_j con $j > 1$ y $j \leq m$.

3. Proceso, según la reivindicación 1, en el que el conjunto (111) de una o varias distribuciones comprende una primera distribución que es función de I_1 y una segunda distribución que es función de I_j con $j > 1$ y $j \leq m$.

4. Proceso, según cualquiera de las reivindicaciones anteriores, en el que la condición de clasificación para un píxel considerado se refiere al valor del píxel considerado en I_1 y en el por lo menos un I_j , con $j > 1$ y $j \leq m$.

5. Proceso, según cualquiera de las reivindicaciones anteriores, en el que se utiliza un parámetro de calibración empírico (113) en la determinación (204) de la condición de clasificación (105), habiéndose determinado dicho parámetro de calibración empírico (113) mostrando (403) muestras biológicas de calibración (302), preparadas del mismo modo que la muestra biológica (101), a profesionales de la biología.

6. Proceso, según cualquiera de las reivindicaciones anteriores, que comprende una determinación de la matriz A a partir de un análisis de componentes principales, a partir de un análisis de componentes independientes o a partir de un análisis factorial sobre la matriz X .

7. Proceso, según cualquiera de las reivindicaciones anteriores, en el que $n=3$, $m=3$, X_1 es un valor rojo R en la primera imagen digital (103), X_2 es un valor verde G en la primera imagen digital (103) y X_3 es un valor azul B en la primera imagen digital (103), y en el que

$$A = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \end{bmatrix}$$

8. Proceso, según cualquiera de las reivindicaciones anteriores, en el que la primera imagen digital (103) muestra por lo menos parte de la muestra biológica (101) teñida según una primera técnica de tinción.
- 5 9. Proceso, según la reivindicación anterior, en el que la imagen digital adicional (103b) muestra por lo menos parte de la muestra biológica (101) teñida según una segunda técnica de tinción.
- 10 10. Proceso, según cualquiera de las reivindicaciones anteriores, en el que la primera imagen digital (103) se obtiene a partir de una imagen óptica, una imagen en fluorescencia o una imagen en polarización.
- 11 11. Proceso, según cualquiera de las reivindicaciones anteriores, en el que X_2 se obtiene de la imagen digital adicional (103b) y por lo menos un X_i con $i > 2$ se obtiene de otra imagen digital.
- 15 12. Proceso, según cualquiera de las reivindicaciones anteriores, en el que X_1 se obtiene a partir de una combinación de la primera imagen digital (103) con otra imagen digital y por lo menos un X_i para $2 \leq i \leq n$ se obtiene a partir de una combinación de la imagen digital adicional (103b) con la otra imagen digital.
- 20 13. Dispositivo de procesamiento de datos que comprende medios para llevar a cabo el proceso según cualquiera de las reivindicaciones anteriores.
14. Programa informático que comprende instrucciones que, cuando el programa es ejecutado por un dispositivo de procesamiento de datos, hacen que el dispositivo de procesamiento de datos lleve a cabo el proceso según cualquiera de las reivindicaciones 1 a 12.
- 25 15. Medio legible por ordenador que comprende instrucciones que, cuando son ejecutadas por un dispositivo de procesamiento de datos, hacen que el dispositivo de procesamiento de datos lleve a cabo el proceso según cualquiera de las reivindicaciones 1 a 12.

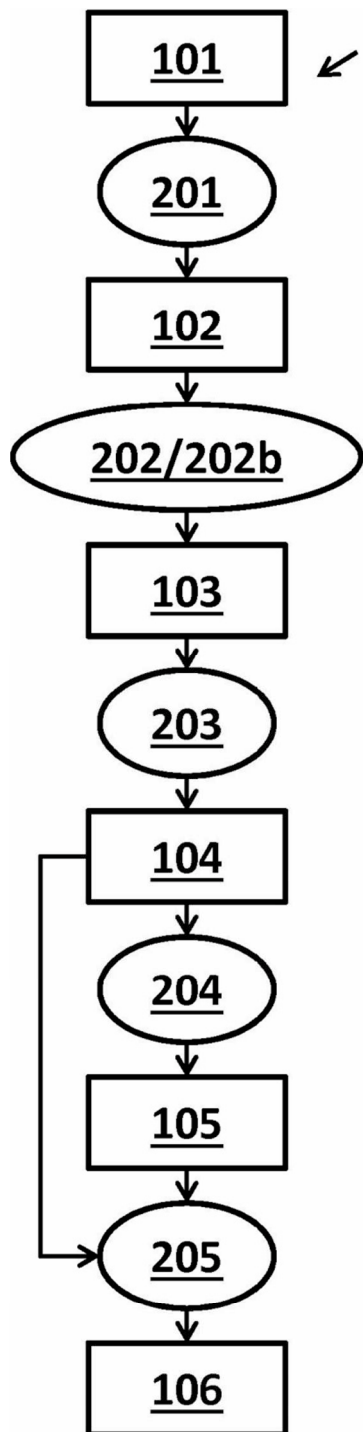


Figura 1a

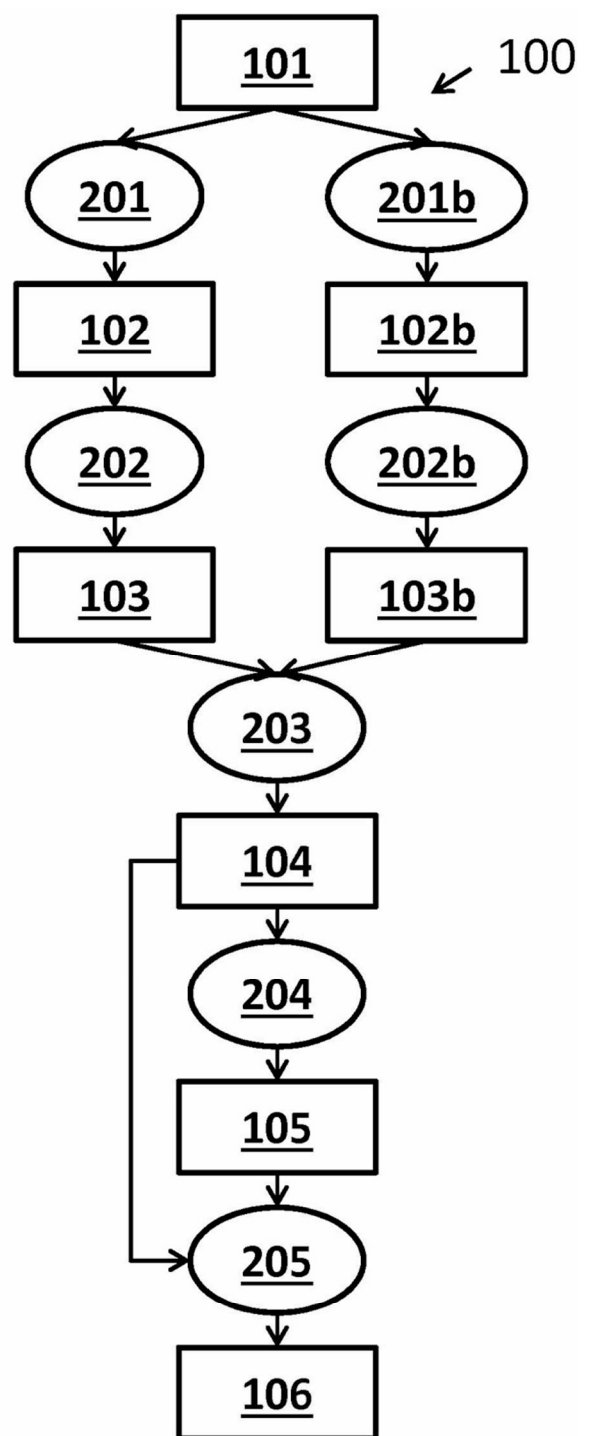


Figura 1b

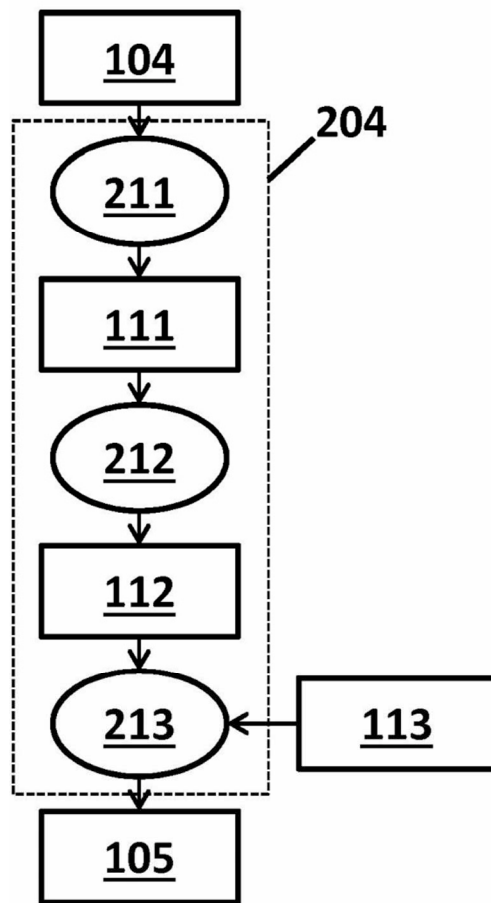


Figura 2

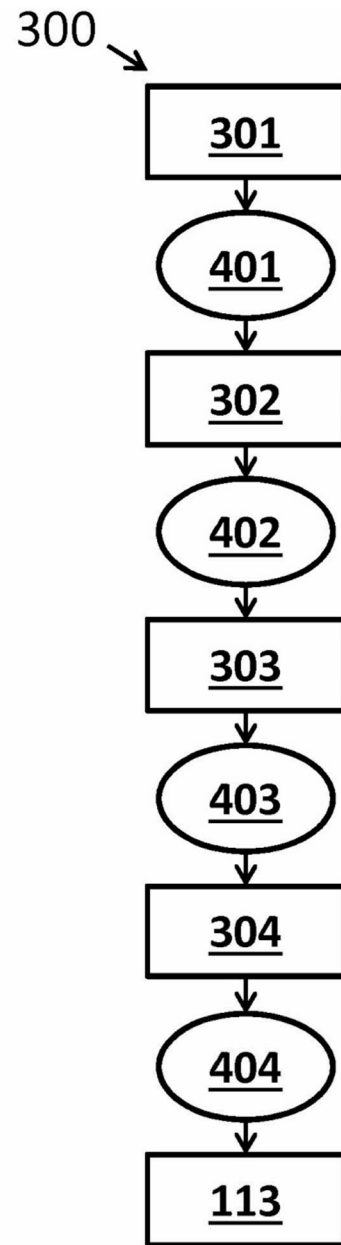


Figura 3

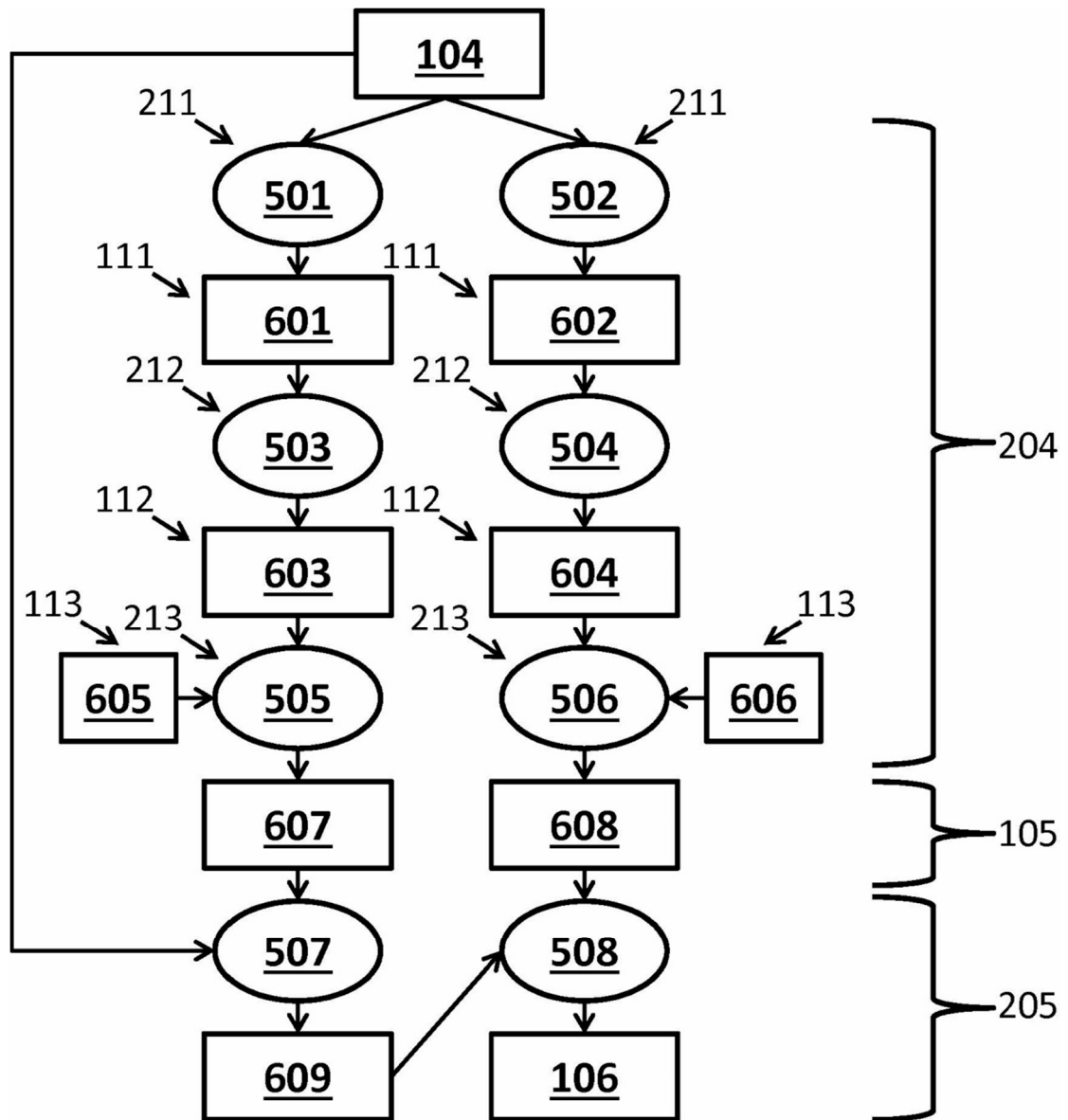


Figura 4

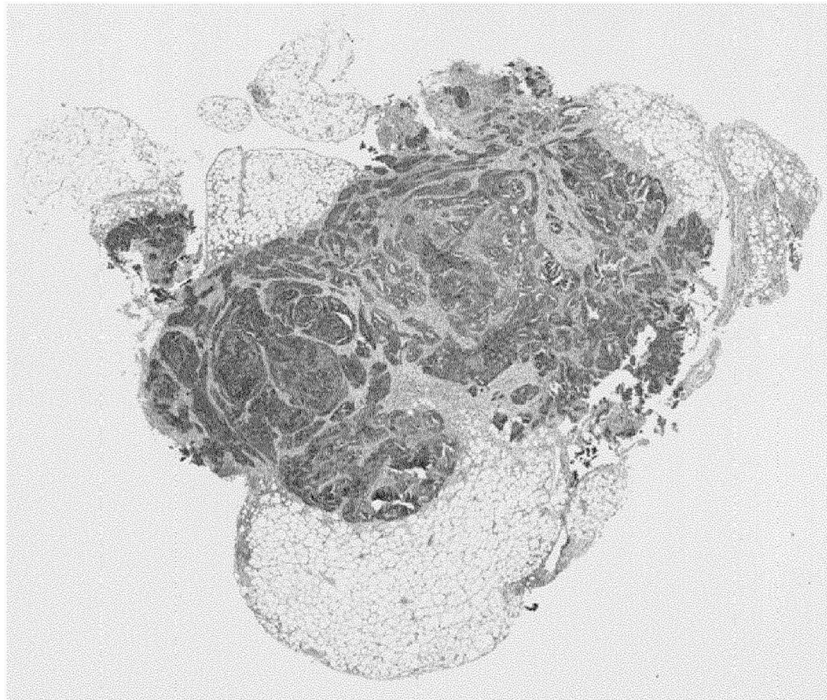


Figura 5

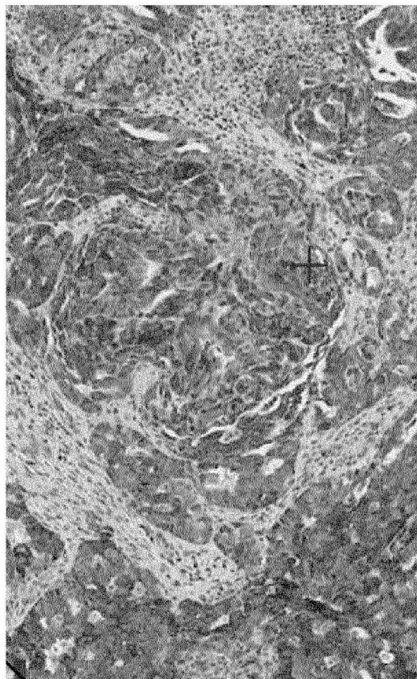


Figura 6a

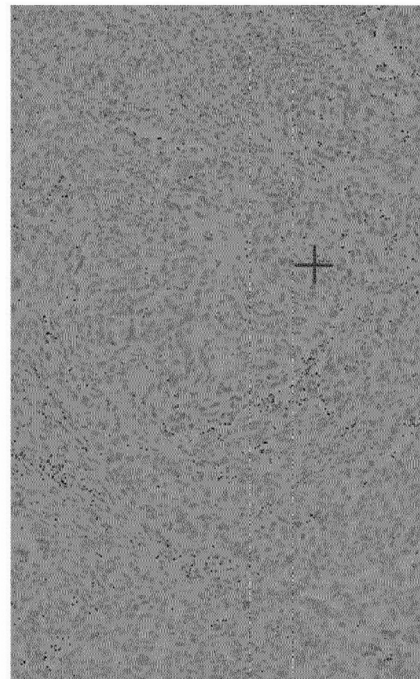


Figura 6b

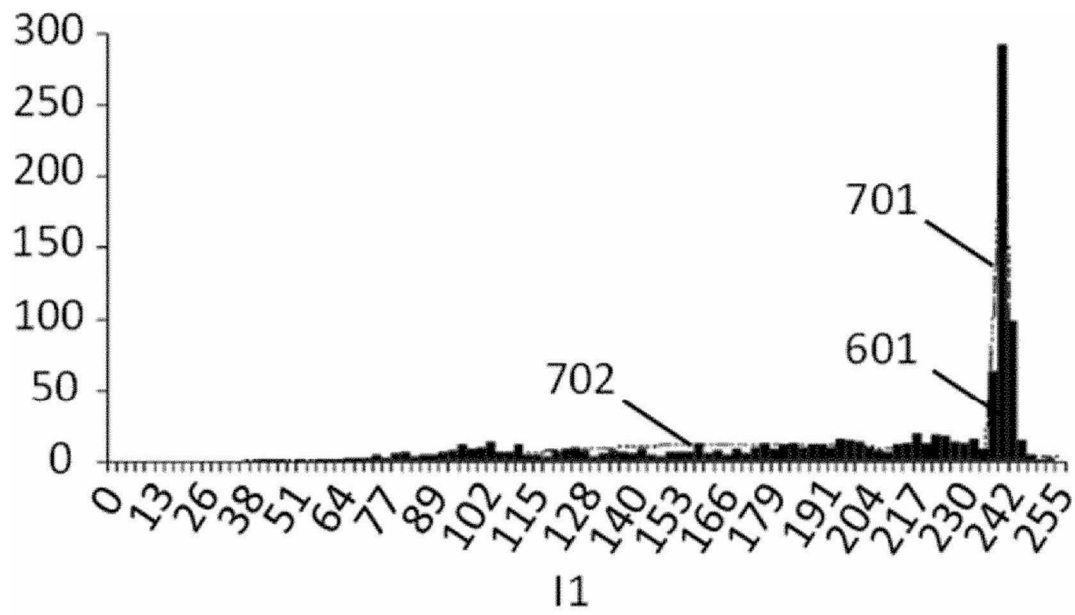


Figura 7

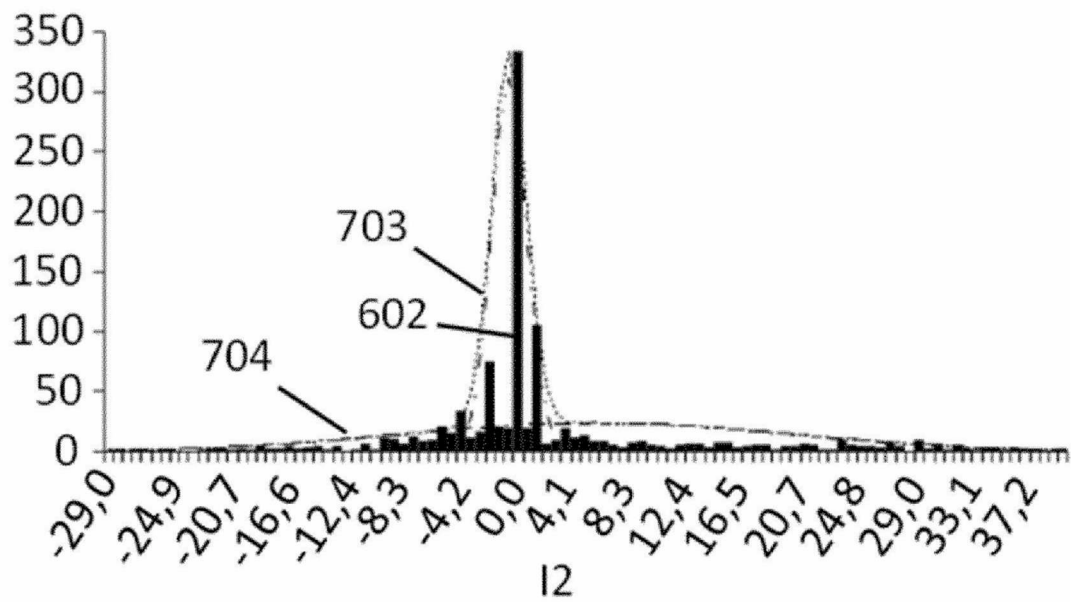


Figura 8

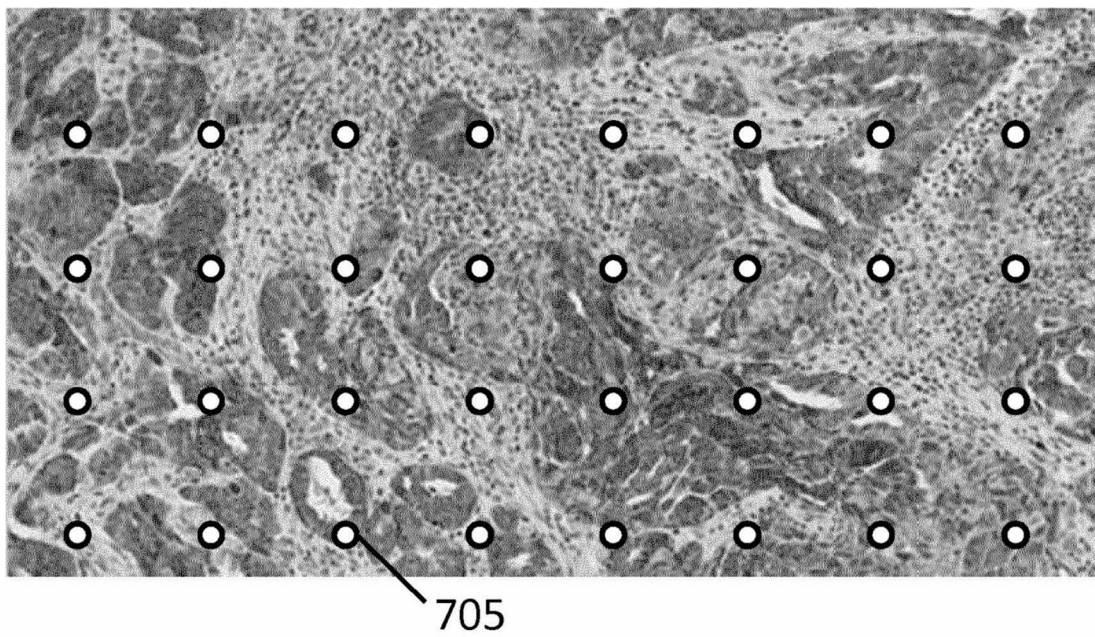


Figura 9a

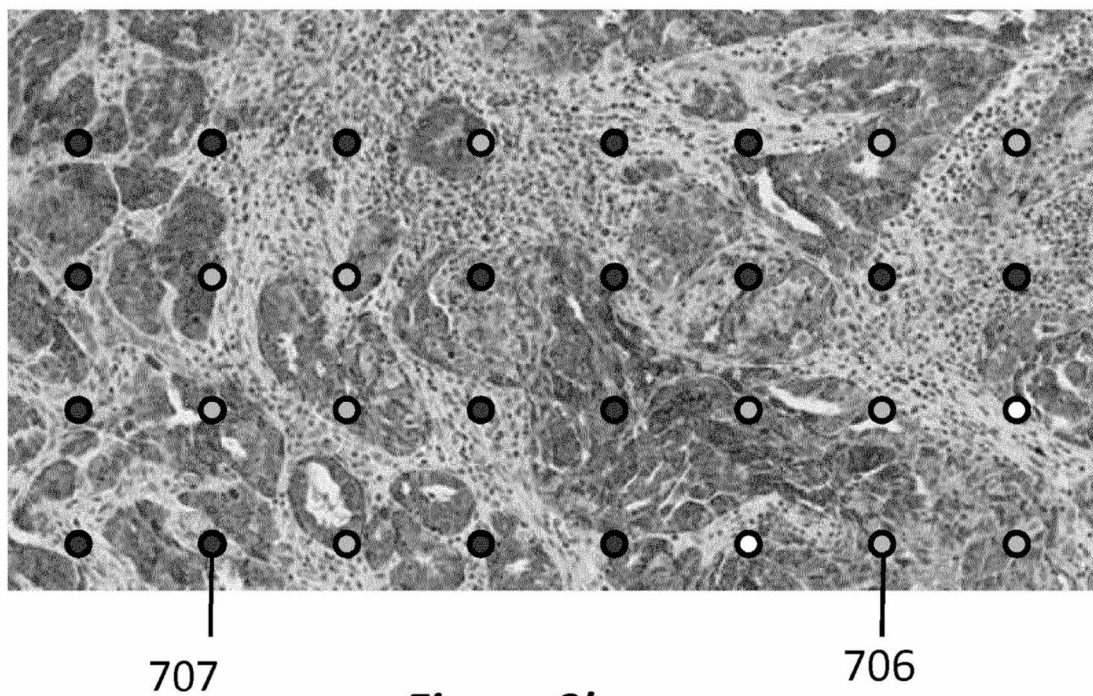


Figura 9b

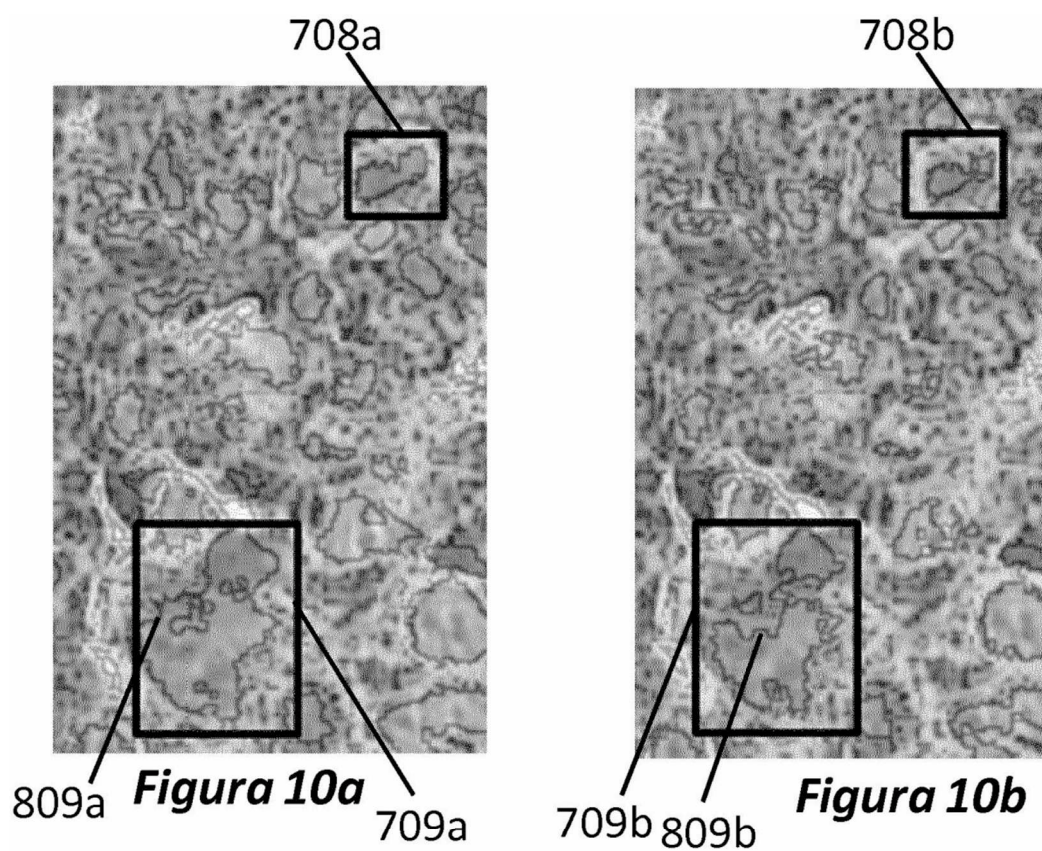
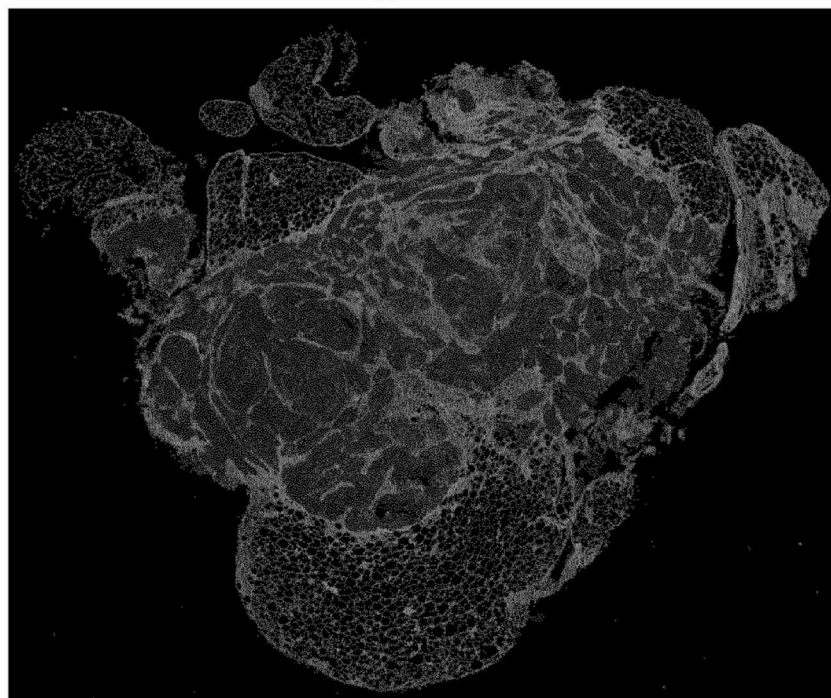


Figura 11



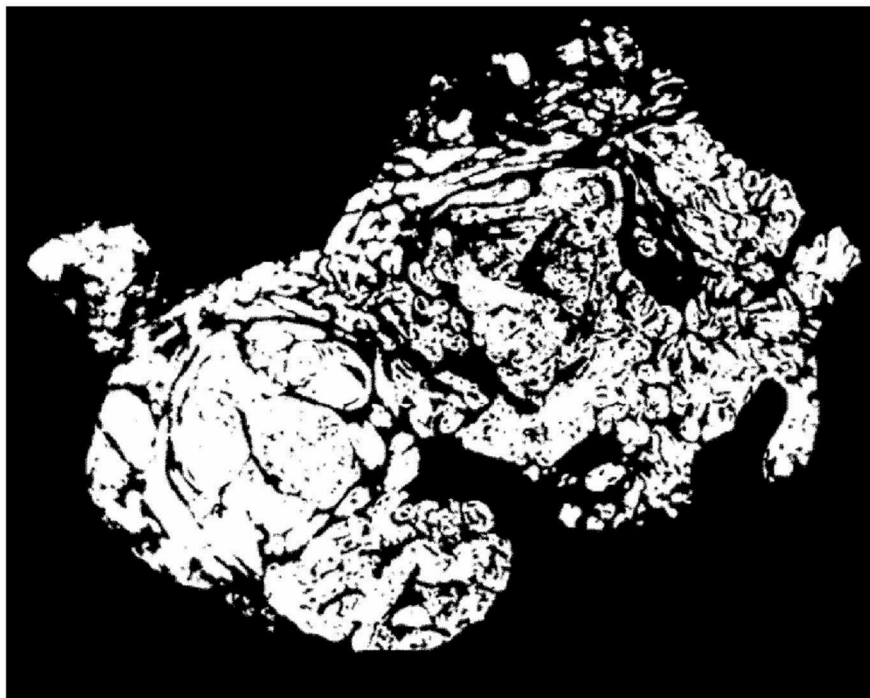


Figura 12

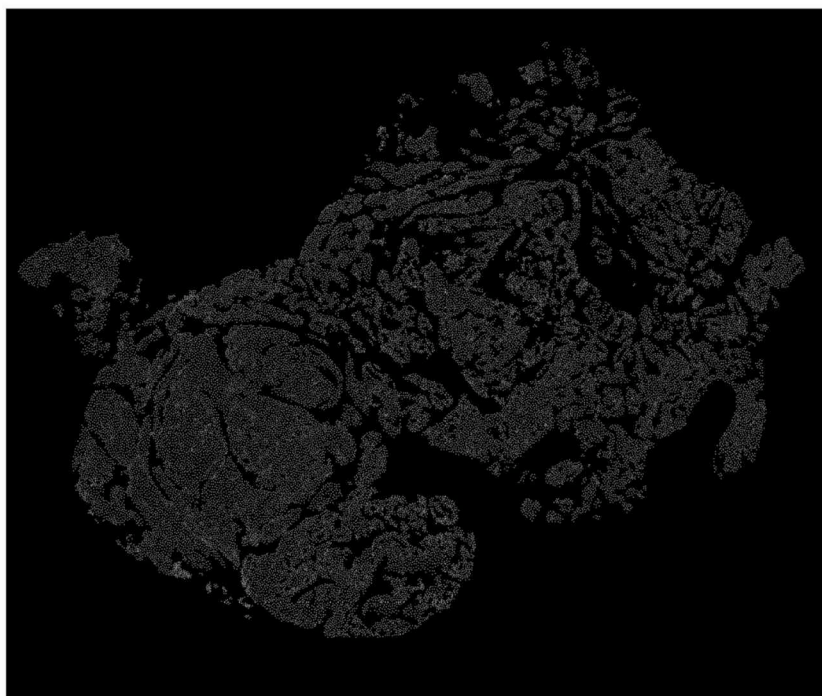


Figura 13

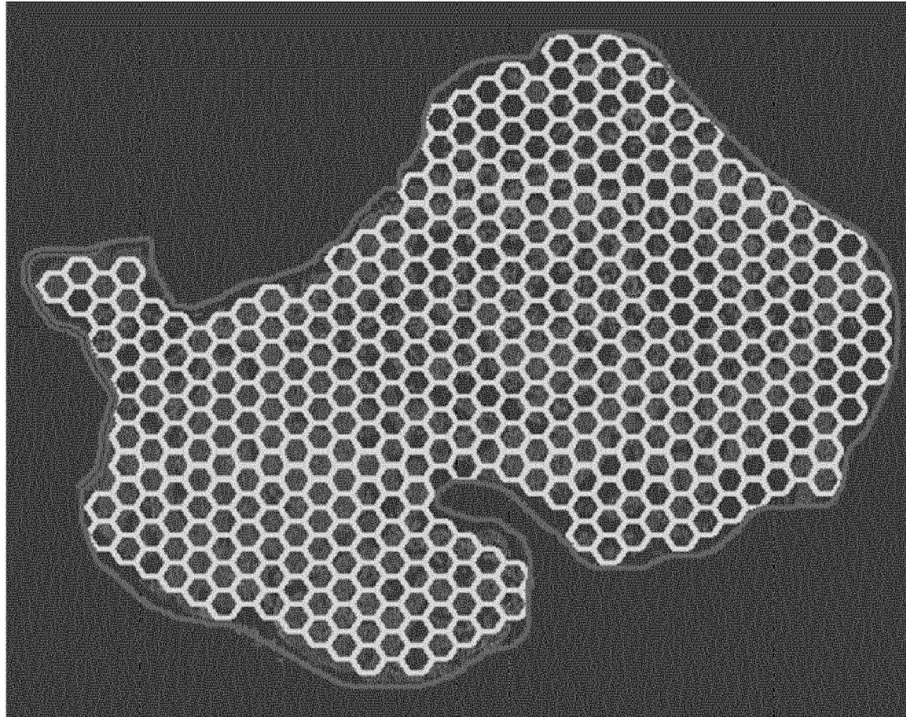


Figura 14

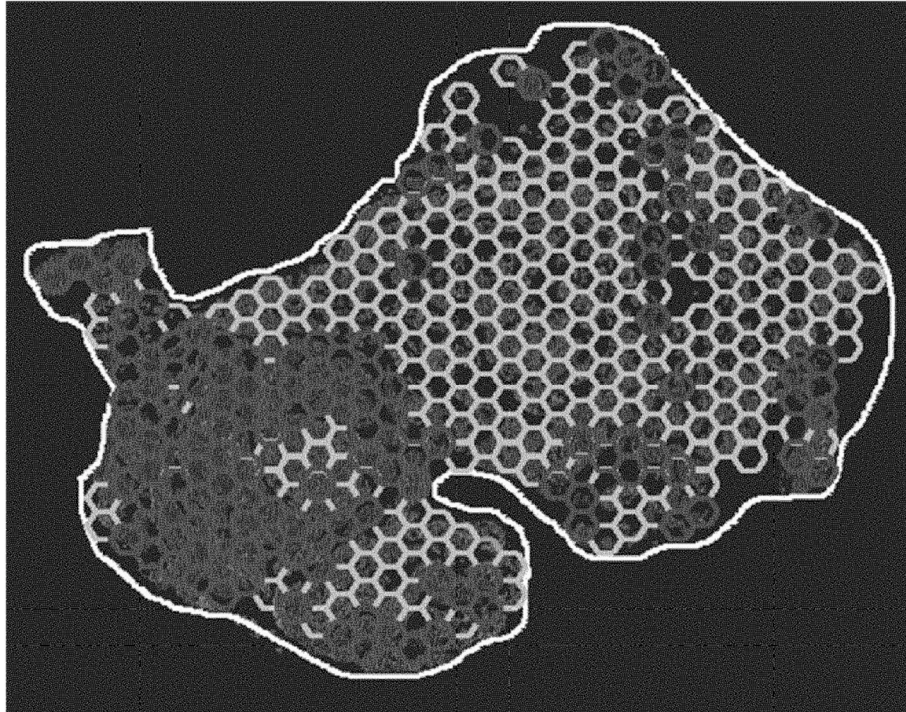


Figura 15

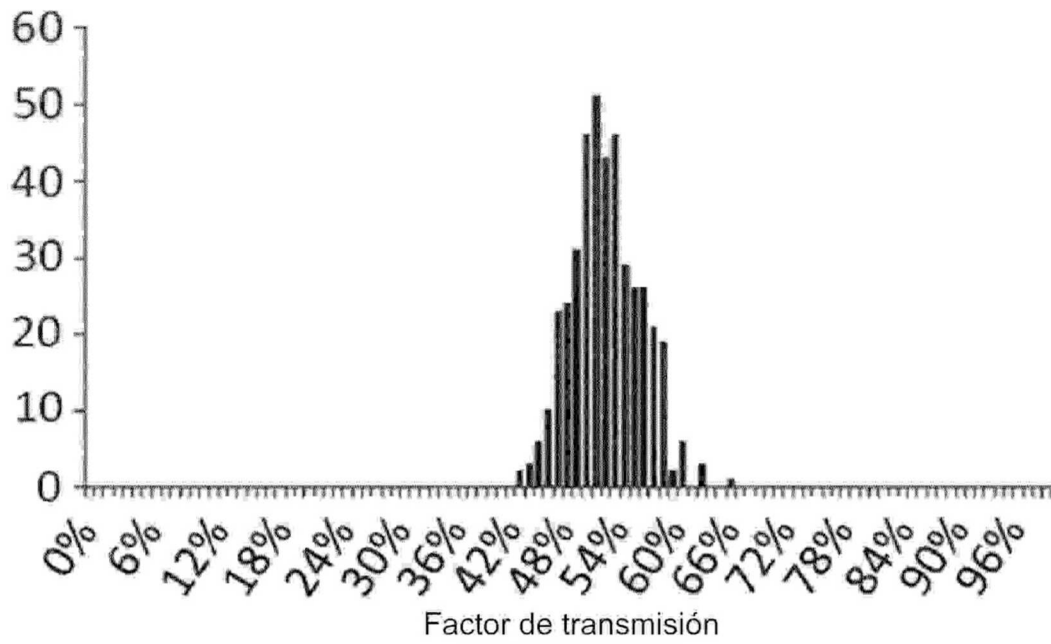


Figura 16

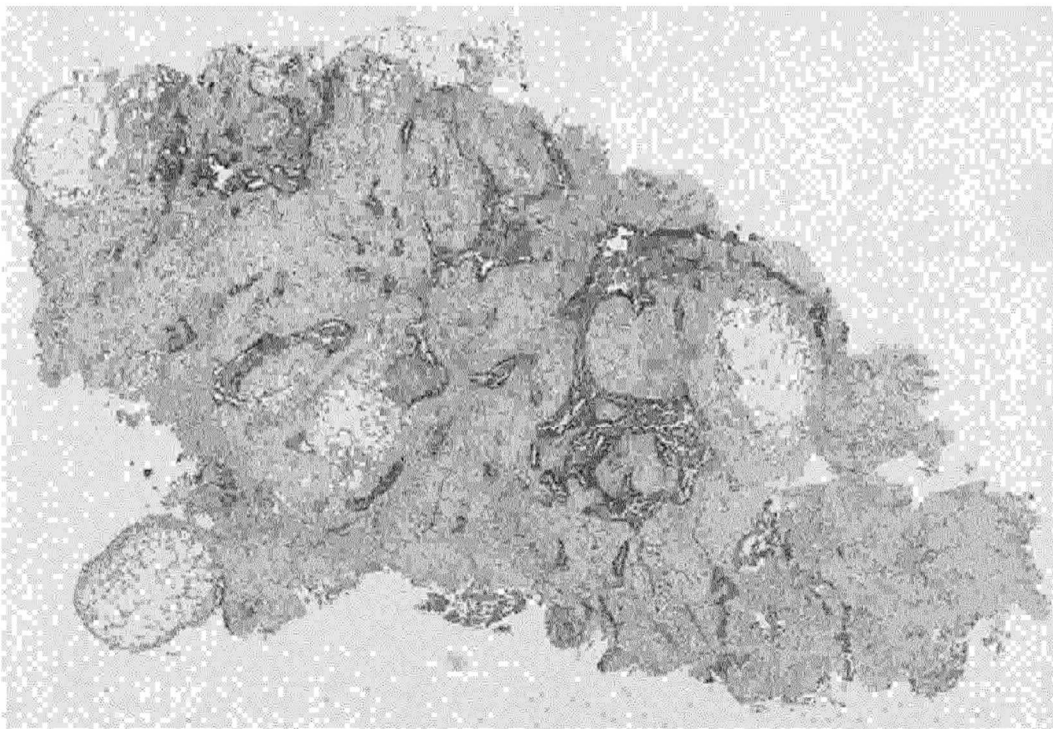


Figura 17

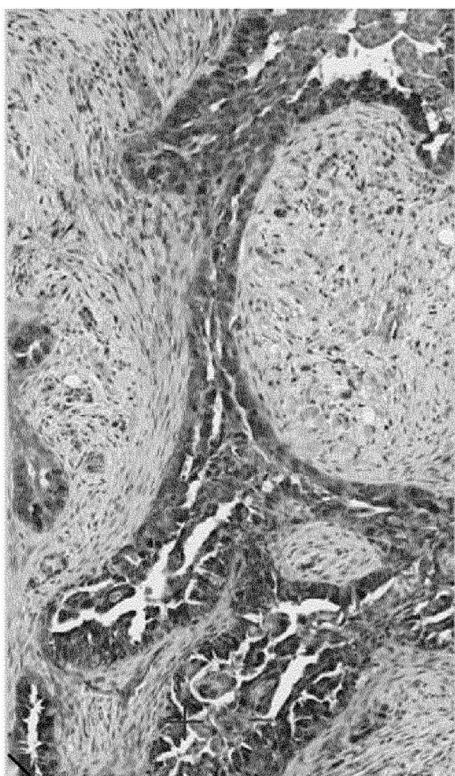


Figura 18a

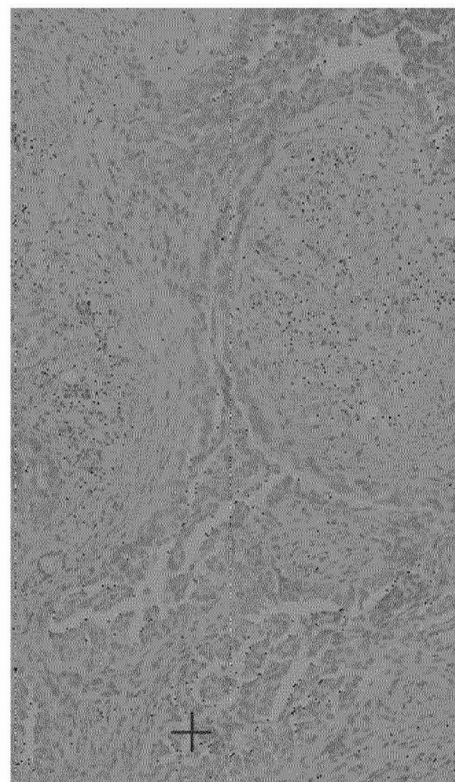


Figura 18b

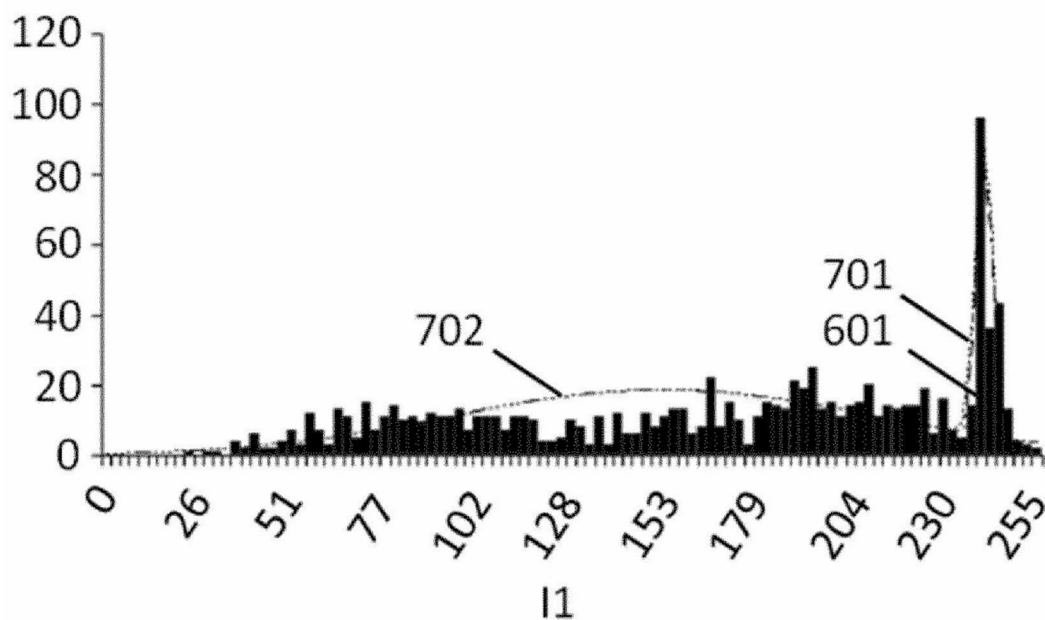


Figura 19

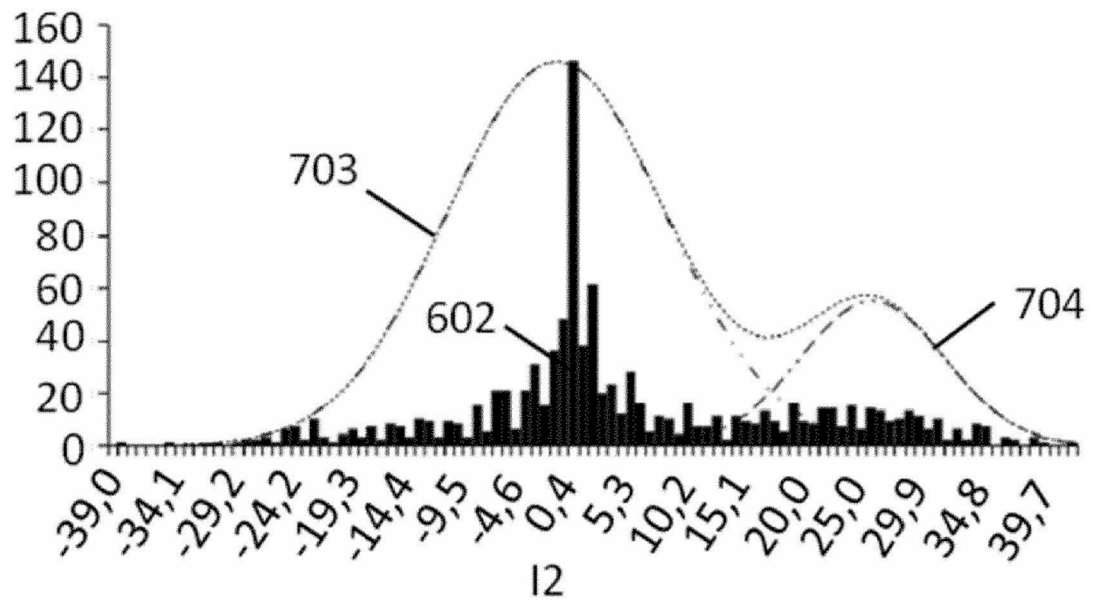


Figura 20

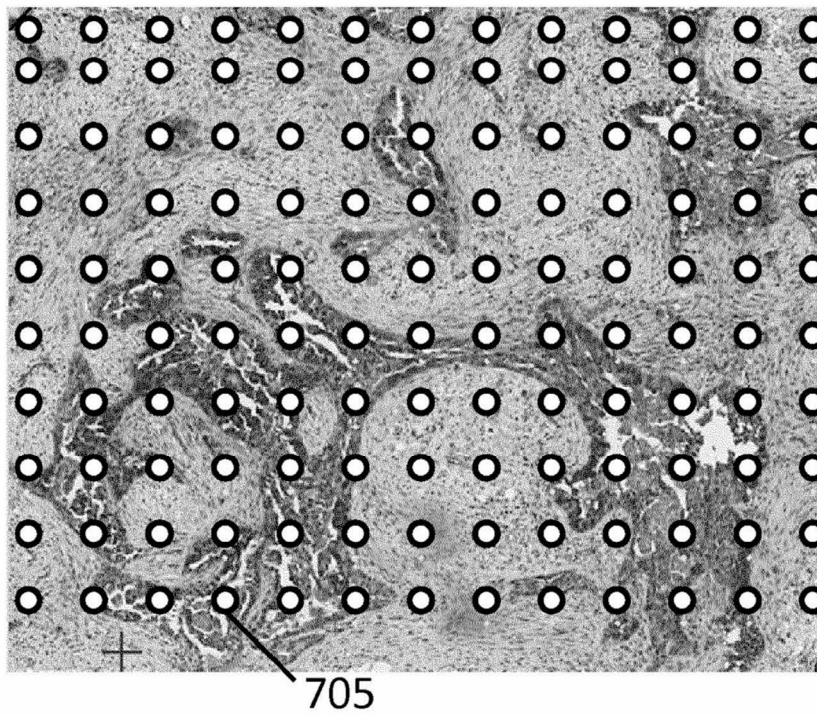


Figura 21a

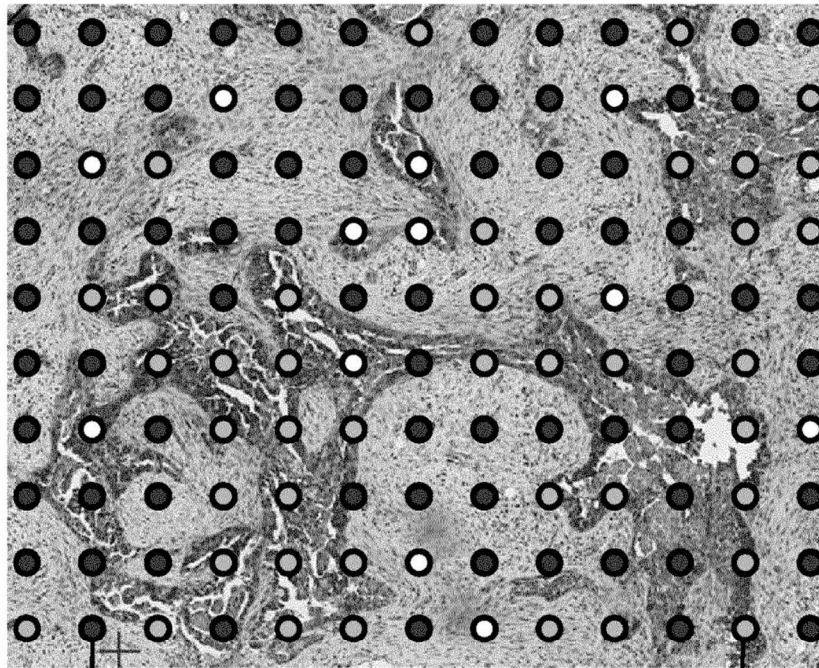
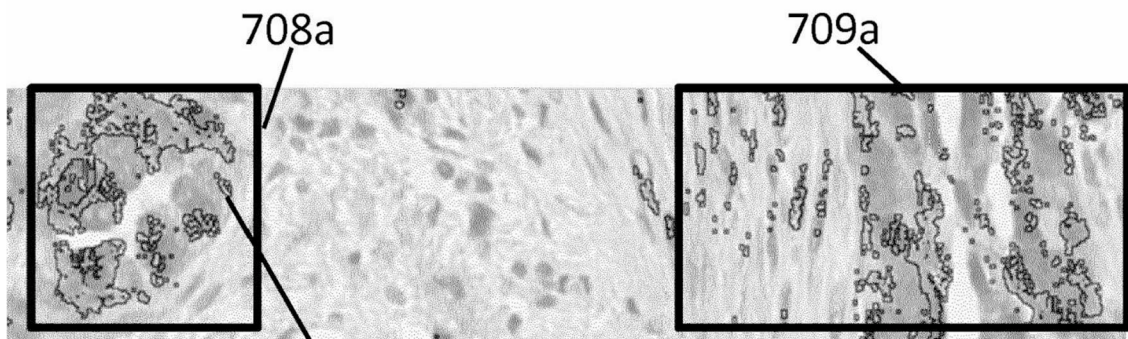


Figura 21b

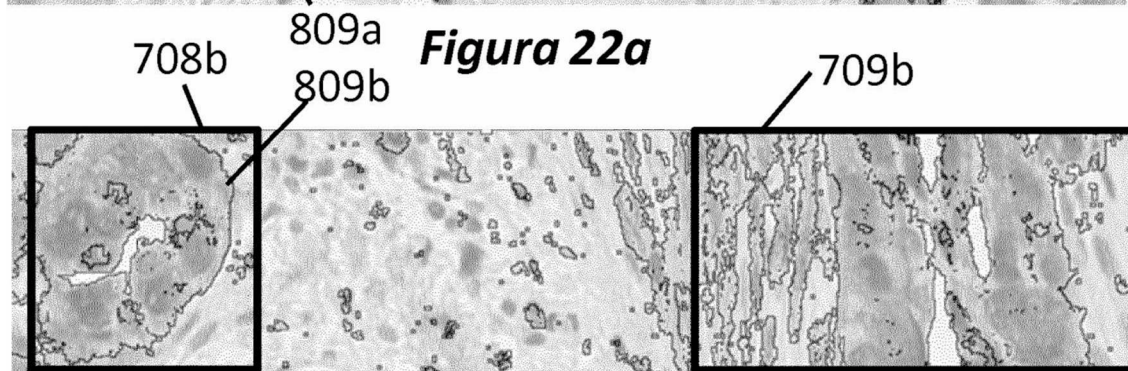
707

706



708a

709a



708b

809a

809b

709b

Figura 22a

Figura 22b

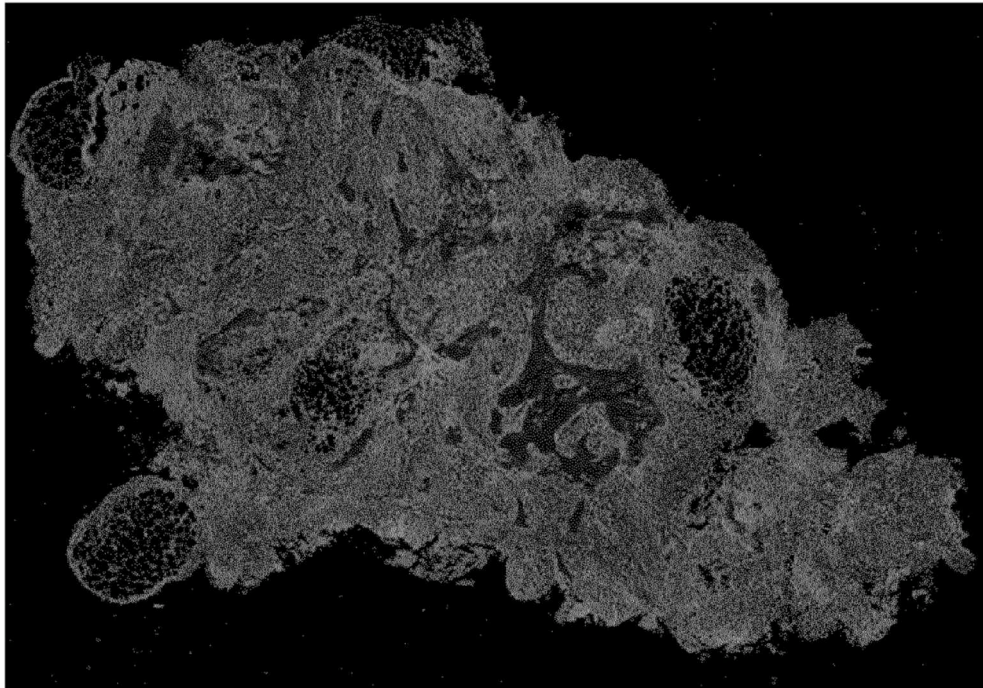


Figura 23

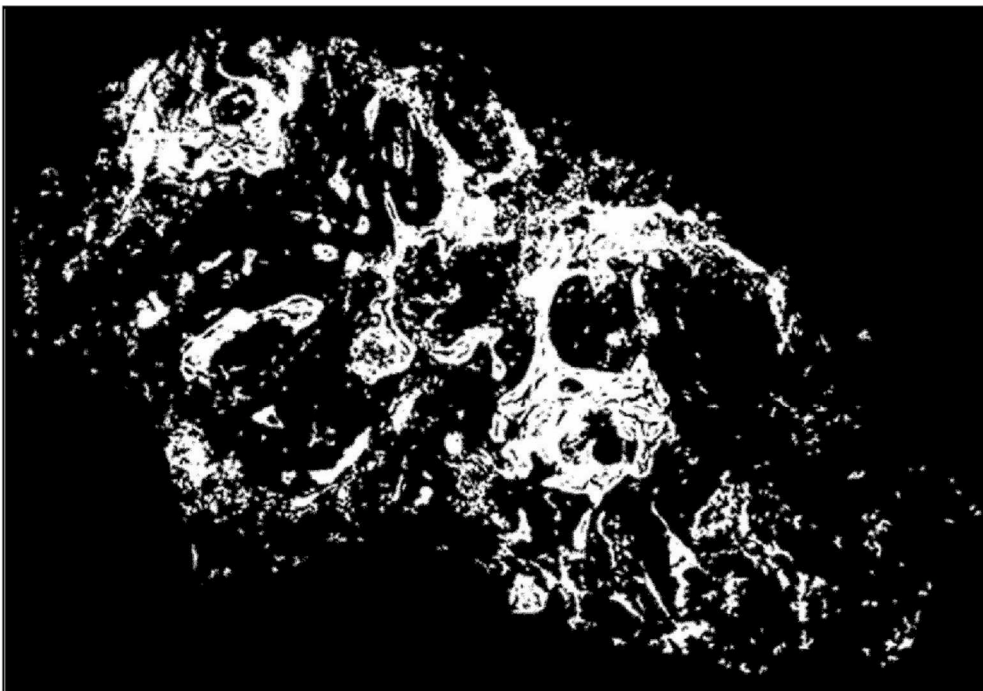


Figura 24

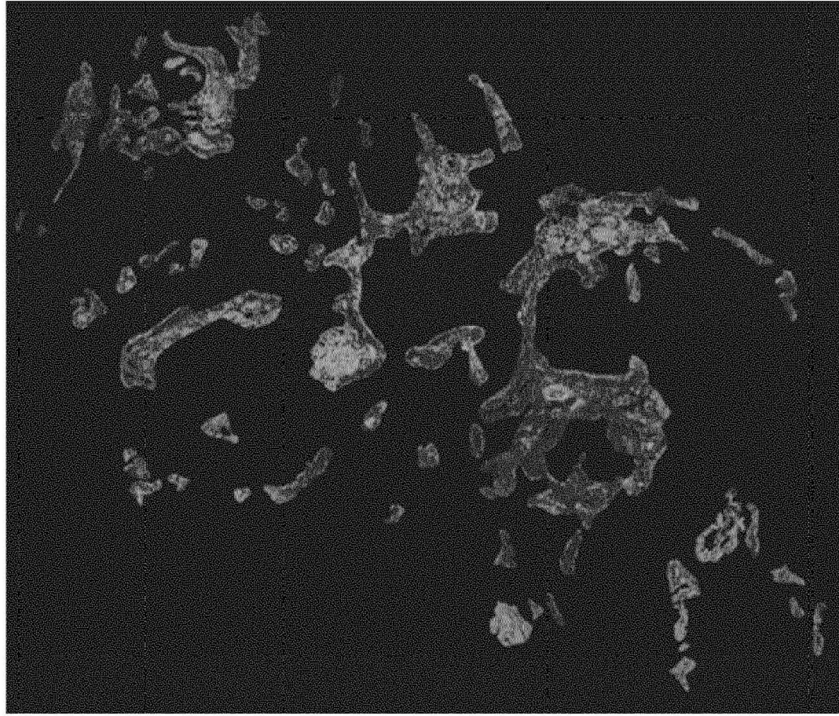


Figura 25

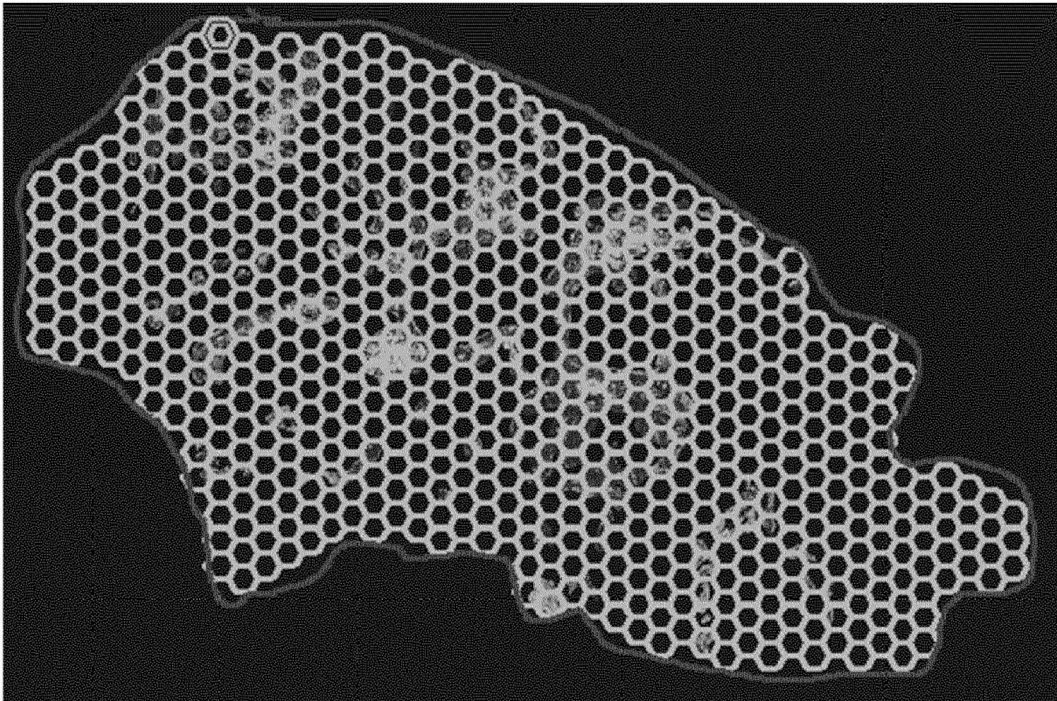


Figura 26

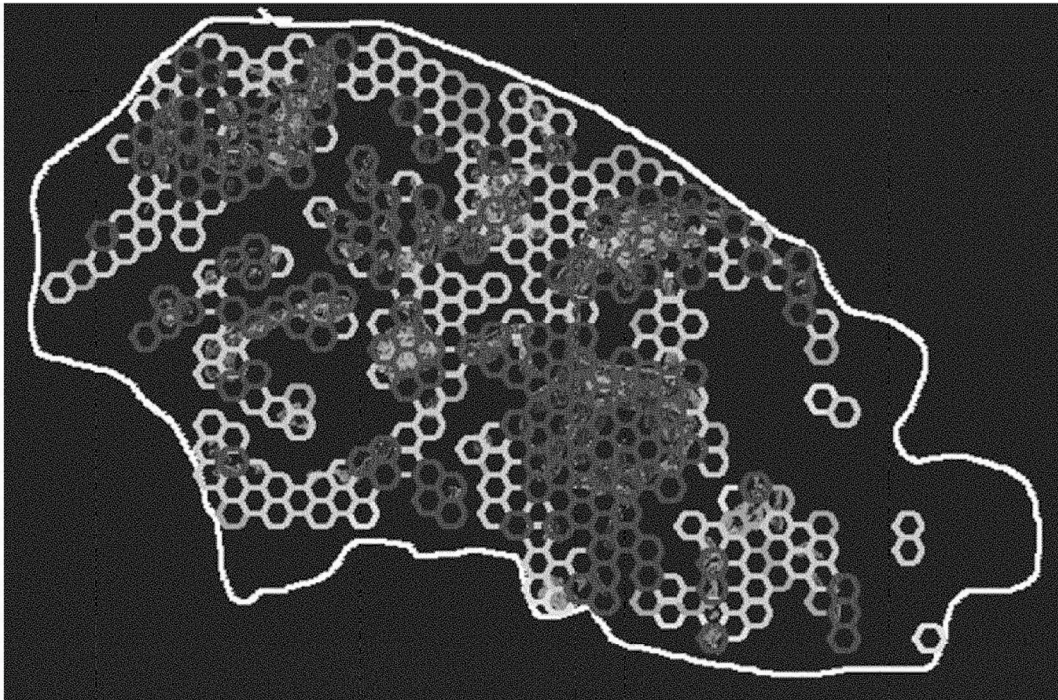


Figura 27

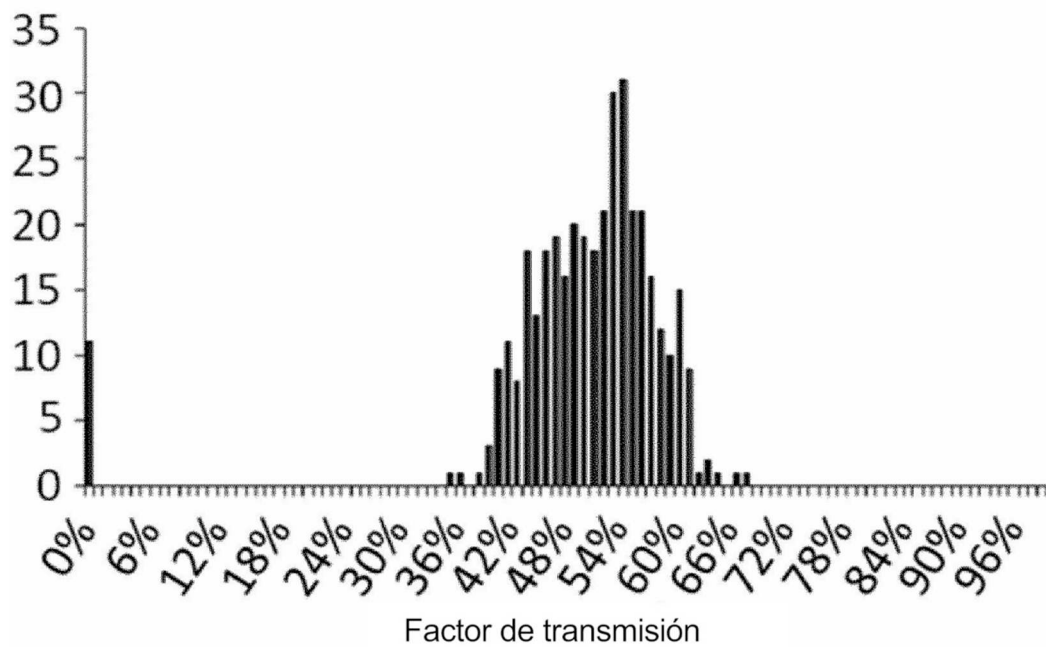


Figura 28



Figura 29

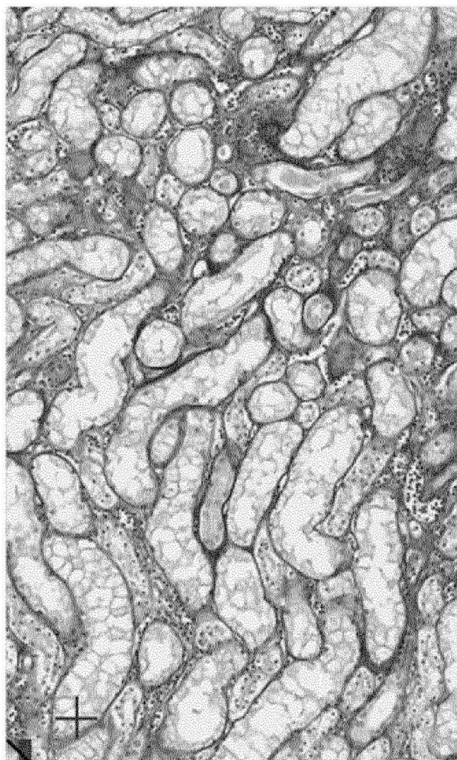


Figura 30a

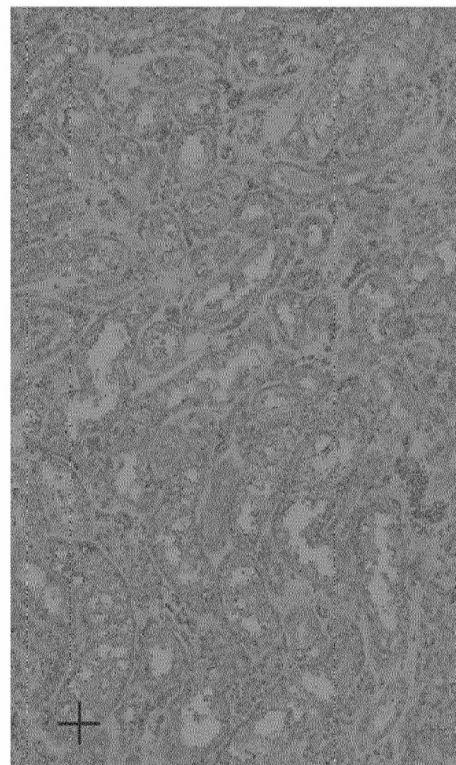


Figura 30b

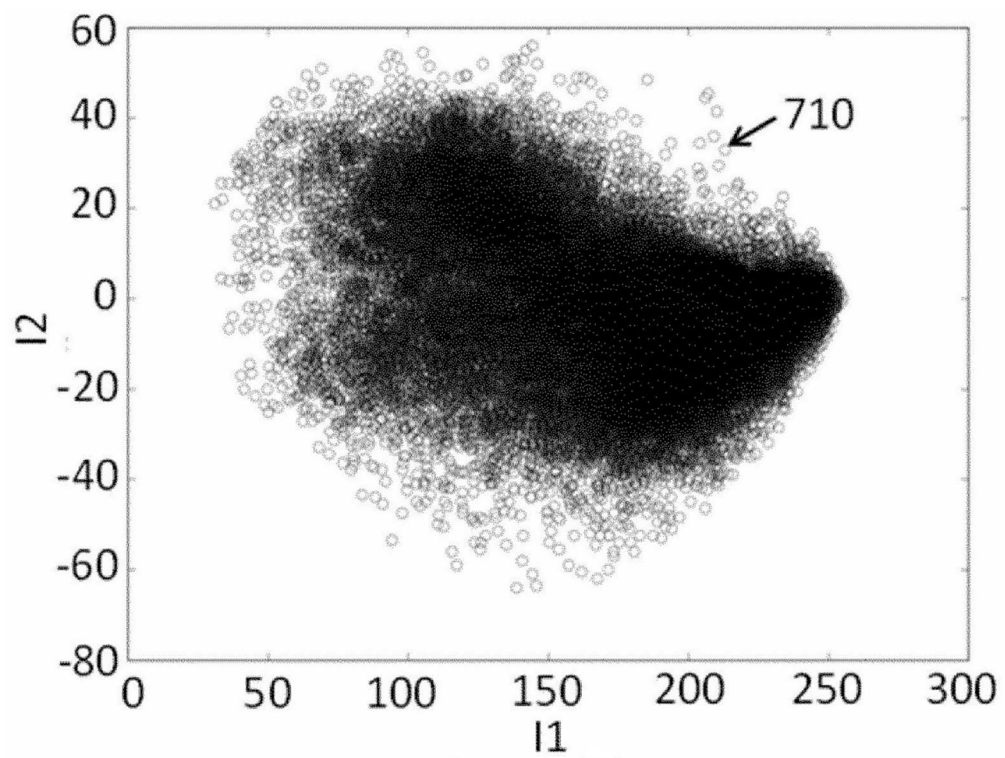


Figura 31a

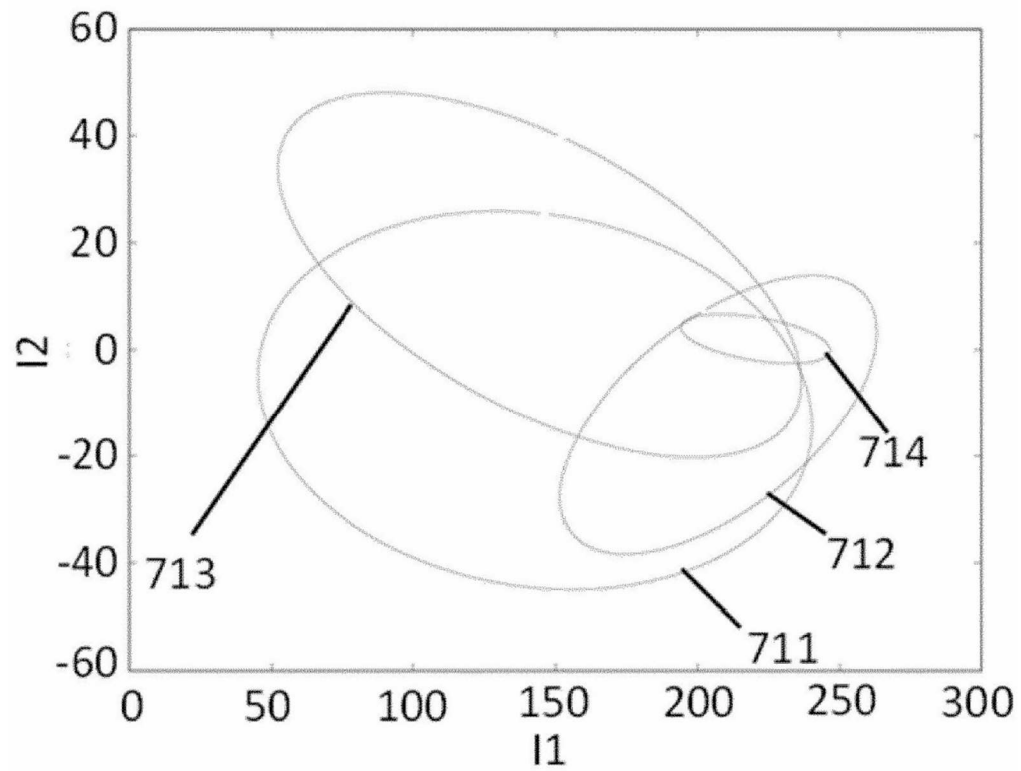


Figura 31b

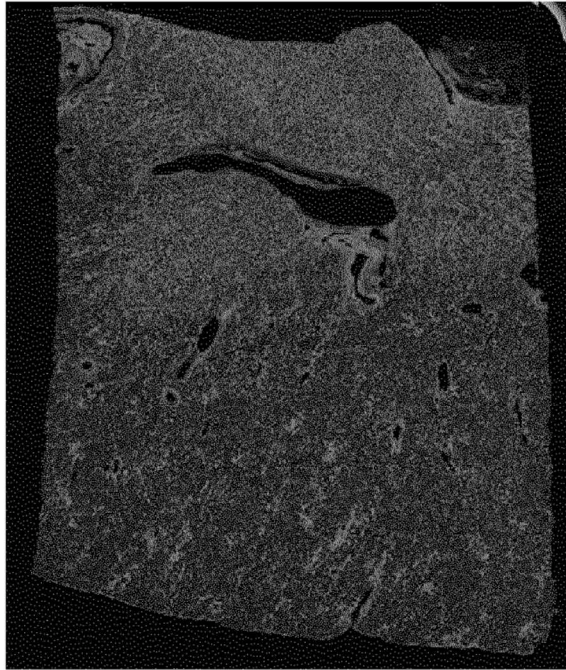


Figura 32

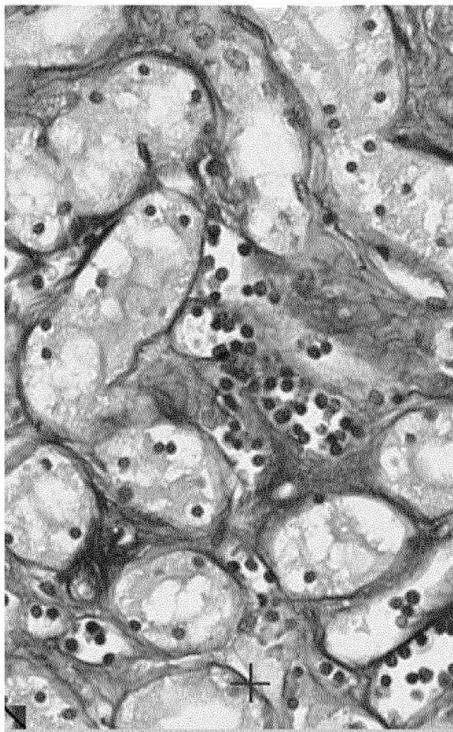


Figura 33a

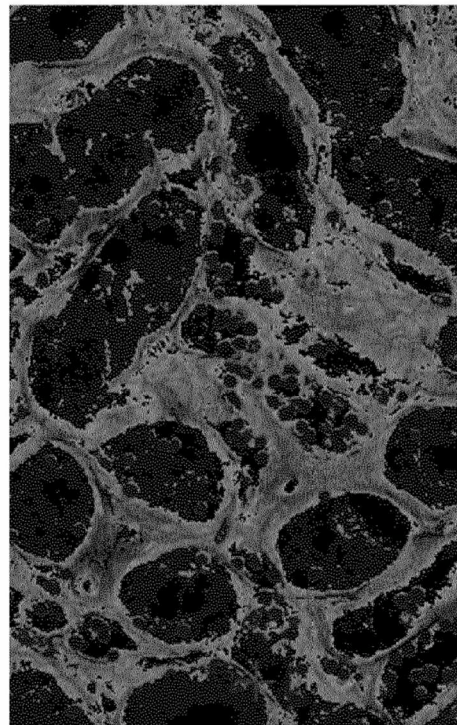


Figura 33b



Figura 34



Figura 35

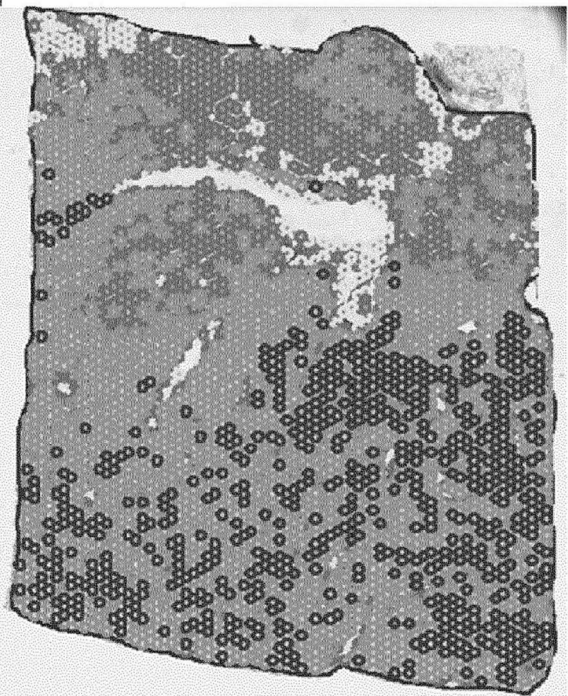


Figura 36

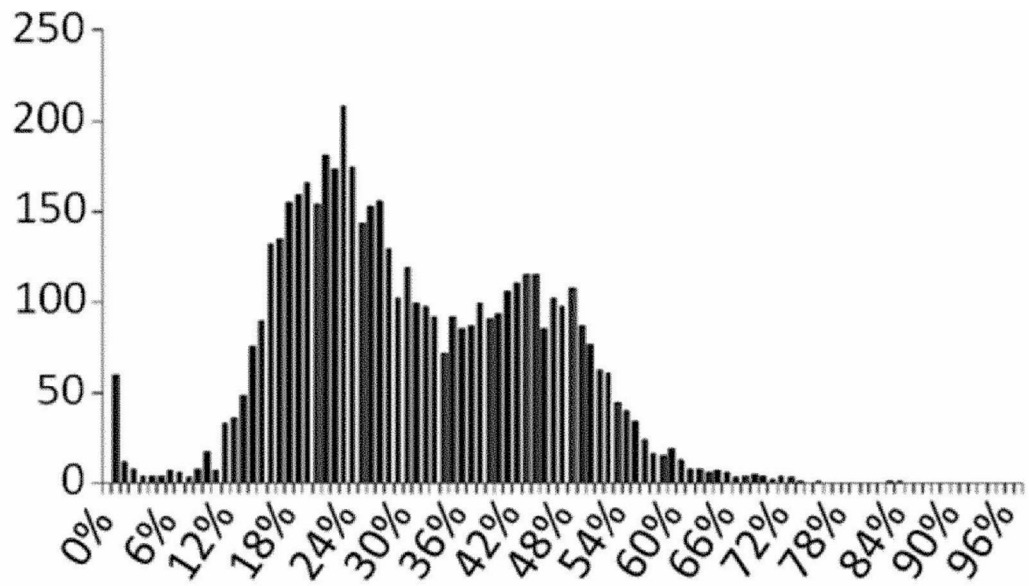


Figura 37

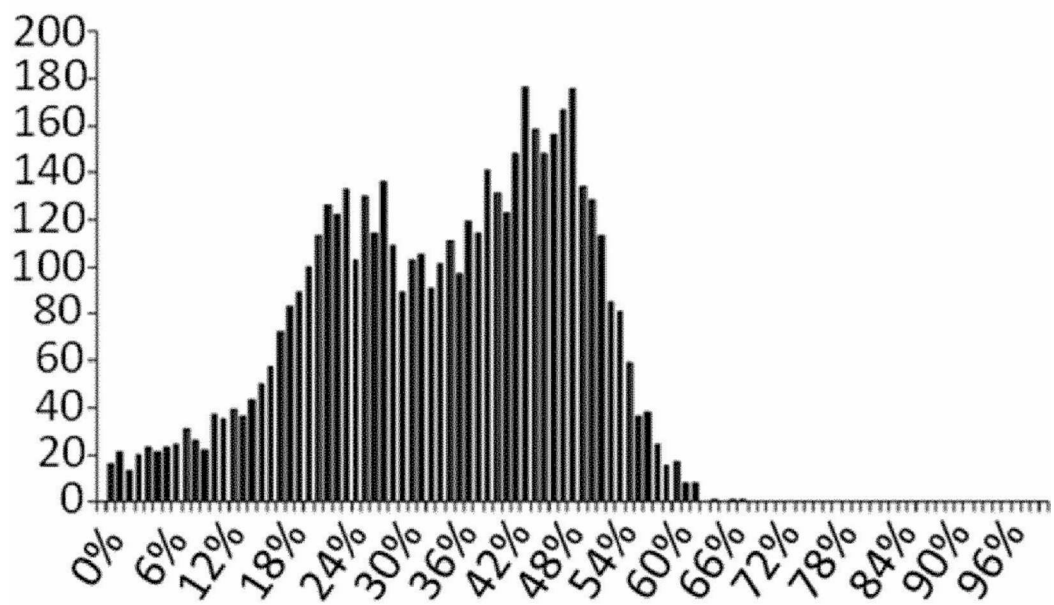


Figura 38

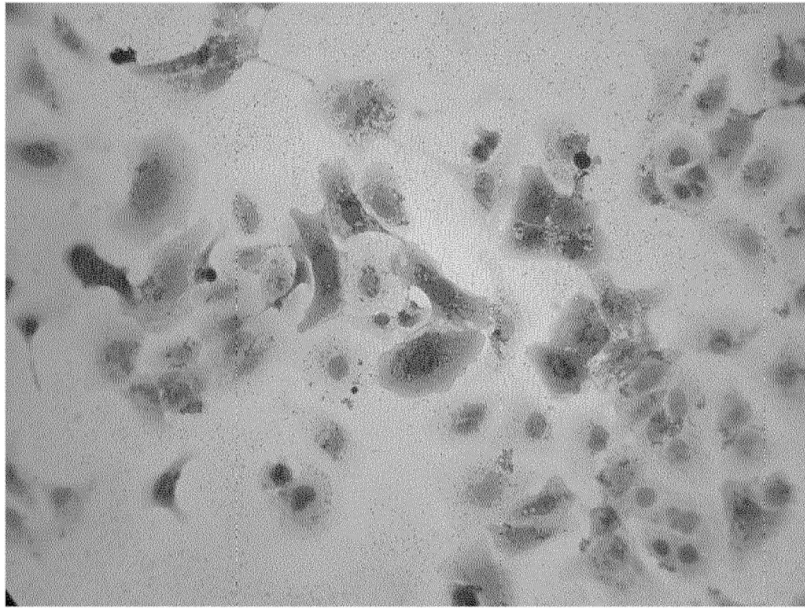


Figura 39

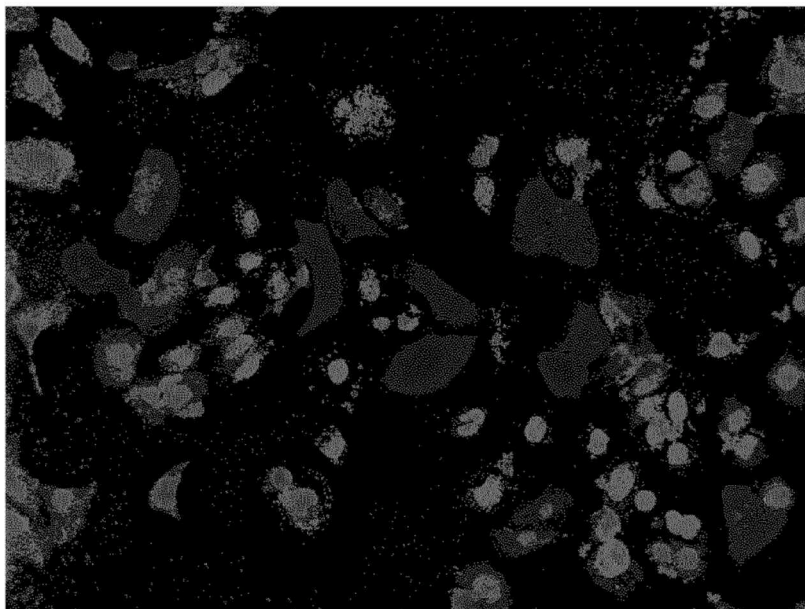


Figura 40

REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Esta lista de referencias citada por el solicitante es únicamente para mayor comodidad del lector. No forman parte del documento de la Patente Europea. Incluso teniendo en cuenta que la compilación de las referencias se ha efectuado con gran cuidado, los errores u omisiones no pueden descartarse; la EPO se exime de toda responsabilidad al respecto.

Documentos de patentes citados en la descripción

- WO 2004084135 A

Literatura no patente citada en la descripción

- Automatic color space selection for biological image segmentation. **MEAS-YEDID V et al.** PATTERN RECOGNITION, 2004. ICPR 2004. PROCEEDINGS OF THE 17TH INTERNATIONAL CONFERENCE ON CAMBRIDGE, UK AUG. 23-26, 2004, PISCATAWAY, NJ, USA. IEEE, 23 August 2004, vol. 3, 514-517