
Segmentación automática de imágenes para el análisis de materiales bidimensionales sintetizados mediante deposición química de vapor

Grado en Ingeniería Informática
Ingeniería de Computadores

Proyecto de Fin de Grado
2015

Autor:
Daniel Franco Barranco

Supervisores:
Ibai Gurrutxaga Goikoetxea
Javier Francisco Muguerza Rivero



Euskal Herriko
Unibertsitatea



Agradecimientos

Estos son los agradecimientos.

Índice general

1. Introducción	1
1.1. Motivación	1
1.2. Estructura de la memoria	2
2. Conceptos básicos	5
2.1. Definición de segmentación	5
2.2. Evolución de la segmentación	5
2.3. Aplicaciones	7
3. Clasificación de técnicas de segmentación	9
3.1. Introducción	9
3.2. Segmentación de imagen basada en el tratamiento de los <i>pixels</i>	10
3.2.1. Thresholding	10
3.2.2. Clustering	13
3.2.3. Morphology	14
3.3. Segmentación de imagen basada en la detección de bordes . . .	15
3.3.1. Detección de borde usando gradientes	16
3.3.2. Active Contours	18
3.3.3. Level set	20
3.4. Segmentación de imagen basada en regiones crecientes	22
3.4.1. Region growing	22
3.4.2. Split/Merge	27
3.4.3. Graph cut	27
3.5. Conclusiones de los tipos de segmentación	28
4. Necesidades del cliente	29
4.1. Requisitos del proyecto del DIPC	29
4.2. Necesidades extra	31
5. Optimizaciones y mejoras del <i>level set</i> original	33
5.1. Aproximación a la técnica de <i>level set</i>	34

5.2. Evolución del contorno	40
6. Implementación de la aproximación y Ofeli	41
6.1. Ofeli	41
6.1.1. Estructura	41
6.1.2. Esquema general	43
7. Paralelización de la aproximación del <i>level set</i>	45
7.1. Planteamiento de la paralelización	45
8. Comparación de resultados temporales y rendimiento	47
9. Funciones extras añadidas en base a la segmentación	49
10. Gestión del proyecto	51
10.1. Gestión del alcance	51
10.2. Gestión del tiempo	51
10.3. Gestión de los riesgos	51
10.4. Gestión del costes	51
11. Conclusiones	53
Bibliografía	55

Índice de figuras

2.1. Resultados de búsqueda en [1] con las palabras «image segmentation»	6
3.1. Clasificación de las técnicas de segmentación	10
3.2. Histograma que muestra tres aparentes segmentos de la imagen con dos umbrales	11
3.3. Segmentación con <i>thresholding</i>	12
3.4. Segmentación con <i>thresholding</i> con varios umbrales	12
3.5. Segmentación con <i>thresholding</i> en imagen con iluminación gradiante	12
3.6. Segmentación utilizando K-means con K=16	14
3.7. Segmentación utilizando K-means con K=6	14
3.8. Ejemplo de dilatación de la técnica de morfología	15
3.9. Detección de bordes con el uso de los operadores	18
3.10. Ejemplo de un contorno activo	19
3.11. Ejemplo de segmentación de plantas presentando en [2] utilizando los contornos activos	20
3.12. Ilustración de la superficie de <i>level set</i>	21
3.13. Ilustración del método de <i>level set</i>	22
3.14. Ejemplo del funcionamiento de la técnica de <i>region growing</i> . .	23
3.15. Ejemplo de la técnica de <i>region growing</i>	24
3.16. Ejemplo del funcionamiento de la técnica de <i>watershed</i>	26
3.17. Funcionamiento de la técnica de <i>watershed</i>	26
3.18. Funcionamiento de la técnica de <i>watershed</i>	27
3.19. Funcionamiento de la técnica de <i>graph cut</i> en una imagen 3x3	28
4.1. Ejemplo de proyecciones en láminas de materiales sintetizados mediante la deposición química de vapor	30
5.1. Representación de la curva C	34

5.2. Ilustración de las operaciones realizadas en el paso 3 del algoritmo	39
5.3. Posibles inicializaciones del contorno con el método de Chan-Vese	40
6.1. Esquema de las funciones utilizadas en el trabajo de Ofeli para realizar el algoritmo de level set aproximado	43

Índice de tablas

3.1.	Máscaras utilizadas por el operador de Roberts de tamaño 2x2	17
3.2.	Máscaras utilizadas por el operador de Sobel de tamaño 3x3 .	17
3.3.	Máscaras utilizadas por el operador de Prewitt de tamaño 3x3	17
5.1.	Algoritmo completo de la aproximación del <i>level set</i>	38

Resumen

Puedes poner un resumen de lo que trata el PFC.

Capítulo 1

Introducción

Este proyecto se ha desarrollado para dar respuesta a un proyecto emergente dentro de la fundación Donostia International Physics Center (DIPC). Por ello, ciertas de las actividades a desarrollar han sido propuestas por dicho proyecto ya que son necesarias para su avance y desarrollo. Así pues, se ha mantenido contacto con el director del proyecto interno del DIPC durante el ciclo de vida de este proyecto.

El objetivo de este proyecto es el análisis computacional y experimental de la segmentación de imágenes, en este caso, producidas mediante la deposición química de vapor de materiales sintetizados. En el tratamiento de estas imágenes, aparte del objetivo principal de poder encontrar los diferentes elementos que se encuentran en ella, también se presentan varias tareas complementarias: porcentaje de recubrimiento de los elementos encontrados, aislamiento y densidad espectral de cada elemento.

Una vez realizadas estas tareas, se requiere paralelizar el tipo de técnica de segmentación escogida para poder satisfacer necesidades como el tratamiento de cada *frame* de una secuencia de vídeo de manera que se pueda dar una respuesta en tiempo real.

1.1. Motivación

Como se explicará más adelante en el apartado 2.2 la segmentación es algo que ha avanzado mucho durante las últimas décadas. Esto se debe a que muchas de las tecnologías que utilizamos diariamente hacen uso en cierta manera de la segmentación de imagen para realizar ciertas funciones. La segmentación de imagen está presente en la detección de movimiento y posicionamiento humano en las consolas, en aplicaciones fotográficas en las que se detectan los rostros humanos, en imágenes de satélites espaciales para

la predicción temporal del telediario, en la localización de tumores u otros síntomas patológicos en radiografías médicas etc.

A parte de ayudar a mejorar ciertas funciones de la vida cotidiana, la segmentación se ha vuelto en muchos campos como la medicina indispensable para hacer ciertos trabajos. Por no decir, que es la pieza clave en trabajos como la cirugía robotizada, ya que es necesario que al estar realizando la operación se detecten con precisión todas las partes físicas del cuerpo humano para no poner en riesgo la vida del paciente. Otra utilidad también sería la reconstrucción craneal y/o cerebral en 3D de pacientes.

Las técnicas de segmentación más avanzadas consiguen información detallada sobre la imagen tratada logrando así un resultado mejor, sin embargo cuanta más precisión se necesita más costoso es el algoritmo. Por lo tanto, muchas veces se suele buscar un equilibrio entre la precisión necesaria y la respuesta temporal mínima exigida a la hora de elegir un algoritmo adecuado. Por esta razón, hay muchos trabajos que tratan de mejorar esta respuesta temporal de los algoritmos de segmentación en base a optimizaciones y avances. El resultado de esto son miles de artículos relacionados con implementaciones paralelas e implementaciones en GPU aparte de optimizaciones matemáticas y algorítmicas.

En conclusión, la paralelización de métodos de segmentación está a la «orden del día» puesto que es necesaria una respuesta rápida en determinados ámbitos de uso. Este trabajo pretende crear una versión paralela robusta sobre una aproximación al método de *level set* para proporcionar una respuesta rápida sobre este tipo de segmentación.

1.2. Estructura de la memoria

Este documento está estructurado en varios apartados.

En la primera parte de la documentación, formada por el primer capítulo, se presentan varios conceptos básicos: la definición de la segmentación y la evolución histórica de esta. Realizada esta introducción, se comentan las aplicaciones que hoy en día tiene este ámbito que, como se podrá observar, serán muchas. Con esta primera parte se espera poder ubicar al lector sobre este análisis de imágenes, pieza angular de este proyecto, para que pueda entender correctamente el resto de la documentación.

En la segunda parte, formada por el segundo y tercer capítulo, se presenta una clasificación de los diferentes tipos de segmentación de imágenes, explicando individualmente sus características y trabajos que se hayan realizado de estas técnicas.

La cuarta parte, que lo forma el cuarto capítulo, se centra en la gestión

con el cliente del DIPC. Se nombran sus necesidades, requisitos mínimos y extras necesarios para el desarrollo de su proyecto.

La quinta parte de la documentación, formada por el quinto y sexto capítulo, se centra en una técnica de segmentación concreta: *Level Set*. Se comentarán variaciones de la técnica y mejoras hasta llegar a una aproximación la cuál mejorará el tiempo de ejecución significativamente y con la que se ha trabajado en el proyecto. Se trabajará en base a un trabajo encontrado que tiene las características idoneas para abordar el proyecto [3]. La explicación y de este trabajo se encuentra en el sexto capítulo.

En la sexta parte del documento, formada por el séptimo y octavo capítulo, se encuentra todo lo relacionado con la paralelización de la técnica utilizada. La primera toma de contacto o la primera propuesta de paralelización creada, las posteriores fases desarrolladas, las conclusiones y decisiones tomadas con cada fase y las respuestas temporales y rendimientos.

En la séptima parte, formada por el octavo capítulo, se presentan las funcionalidades extra añadidas, a posteriori de la segmentación, que ha necesitado el cliente para el desarrollo de su proyecto, intentando satisfacer así todas sus necesidades.

En la última parte, noveno capítulo, se presentan las conclusiones del trabajo realizado y algunos aspectos de la gestión del mismo. Al final del documento puede encontrarse la bibliografía y los anexos.

Capítulo 2

Conceptos básicos

2.1. Definición de segmentación

La segmentación de imagen, también nombrada a veces como *labelling*, es el proceso de dividir la imagen en grupos o regiones contiguas cuyos elementos(p. e. pixels o voxels) tienen propiedades o características comunes. Estas regiones servirán para identificar los objetos de la imagen que posteriormente podrán ser clasificados y etiquetados en base a sus propiedades [4, pag. 121].

El resultado final de la segmentación de la imagen será un conjunto de regiones o segmentos que formarán la imagen original. Cada uno de los pixels de una región tendrá una característica común con los pixels de su región y una diferencia significativa respecto a pixels de otra región(por ello se habrán agrupado en distinto segmento), ya sea por ejemplo, en el color, la textura o la intensidad. Por lo tanto, se podrán extraer los segmentos de interes de la imagen, es decir, los objetos que esta contiene.

2.2. Evolución de la segmentación

Los primeros desarrollos en el ámbito de la segmentación de imagen se remontan a hace 50 años. En 1965 se desarrolló un operador para detectar bordes entre diferentes partes de una imagen, conocido como *Roberts operator* o *Roberts Edge Detector*. Este detector fue el primer paso hacia la descomposición de imágenes en diferentes segmentos o regiones. En esa misma década también se propusieron varios detectores de bordes como *Sobel* y *Prewitt edge detectors*. A partir de ahí, comenzaron a surgir diferentes algoritmos y técnicas de segmentación. Junto con esto, también se aumentó el alcance de estas técnicas: de imagen 2D a 3D, de imágenes fijas a imágenes en «movimiento» o secuencias de imagen, de escalas de gris a imágenes a color etc [5, Capítulo

1].

A pesar de los años de investigación dedicados a estas técnicas y el gran número de ellas existentes, la segmentación de imagen sigue siendo un tema de investigación desafiante y no existe aún un estandar de segmentación que funcione bien para cualquier tipo de imagen. Estas técnicas están en continua evolución y aún estan lejos de su madurez. Prueba de ello está en que muchas conferencias de técnicas de imagen tienen apartados de segmentación de imagen, el número de artículos de este ámbito aumenta cada año y muchos libros de procesamiento de imagen tienen capítulos referidos a la segmentación [5, Capítulo 1].

Para que el lector se haga una idea, se ha realizado una búsqueda en [1] con las palabras «image segmentation» en varios años. Este buscador encuentra artículos, conferencias, estandares, libros, revistas y cursos de aprendizaje relacionados con las palabras introducidas. La figura 2.1 muestra el número de resultados de esa búsqueda desde 1960 hasta 2015. Como se puede observar, el número de resultados aumenta notablemente cada lustro.

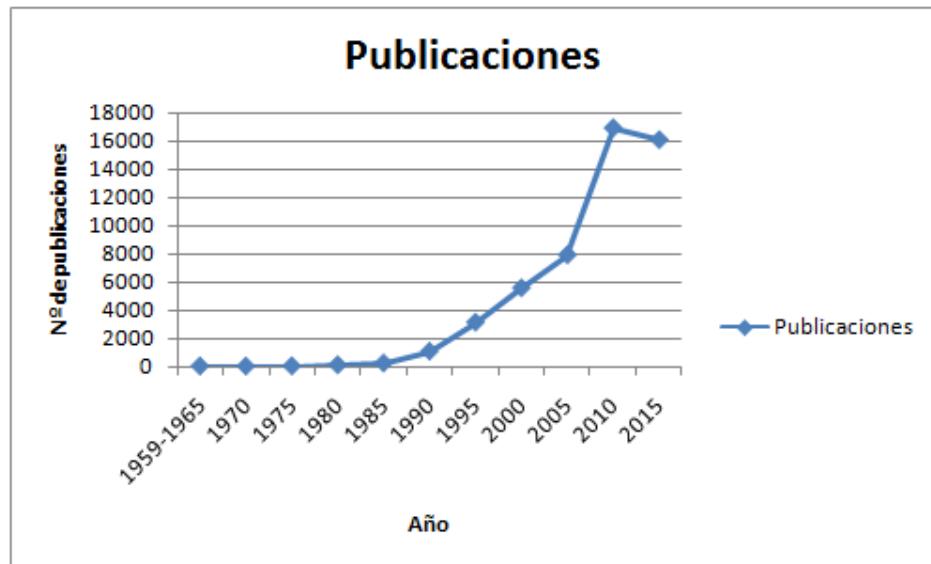


Figura 2.1: Resultados de búsqueda en [1] con las palabras «image segmentation»

Fuente: [1]

2.3. Aplicaciones

Las aplicaciones de segmentación de imágenes son muchas y muy diversas. Cualquier proceso que requiera la extracción de información de una imagen utilizará, en cierta medida, una técnica de segmentación. A continuación nombraremos algunas de las aplicaciones que se han ido recopilando en la realización de esta documentación, por lo que el número de aplicaciones totales será mucho mayor:

- Localización de moléculas en imágenes microscópicas.
- Aplicaciones médicas.
 - Localización de tumores y otras patologías.
- Detección de cuerpos para aplicaciones de seguimiento de movimientos como Kinect.
 - Operaciones guiadas por ordenador.
- Localización de objetos en imágenes de satélite.
- Visión por computador.
- Reconocimientos faciales.
- Reconocimiento de plantas

Capítulo 3

Clasificación de técnicas de segmentación

3.1. Introducción

Hay bastante controversia en cuanto a la clasificación de las diferentes técnicas de segmentación, por el gran número de estas técnicas existentes, por las diferentes maneras en las que cada una tiene representada la imagen, las diferentes características que utilizan de la imagen etc. Hay trabajos que realizan una clasificación de los algoritmos desde dos puntos de vista diferentes: en función de como puede ser utilizado el algoritmo, es decir, las aplicaciones que pueda tener y otra en base al algoritmo en sí, fijándose en como realiza la segmentación [6]. Por otro lado, también hay otros que realizan esta clasificación para ámbitos muy concretos [6, pag. 11].

Zhang propone una clasificación más clara [5], creando la división en cuatro grupos: algoritmos basados en detectar la discontinuidad de las diferentes regiones de la imagen, los llamados *edge-based algorithms*, o basados en detectar la continuidad o la similitud de las regiones, los llamados *region-based algorithms*. Posteriormente hace una subdivisión de estos dos grupos en función de la estrategia de procesamiento: los que realizan un procesamiento secuencial, donde el procesamiento de pasos previos se tienen en cuenta en pasos posteriores, y los que realizan un procesamiento paralelo, es decir, decisiones independientes y simultáneas [5, Sección 1].

La clasificación elegida tiene varios aspectos en común con la última presentada y descrita [7]. Se ha preferido esta clasificación al ser clara en cuanto a la división de las técnicas y las características que estas tienen frente a otras clasificaciones. En la figura 3.1 se muestra el esquema de la clasificación elegida.

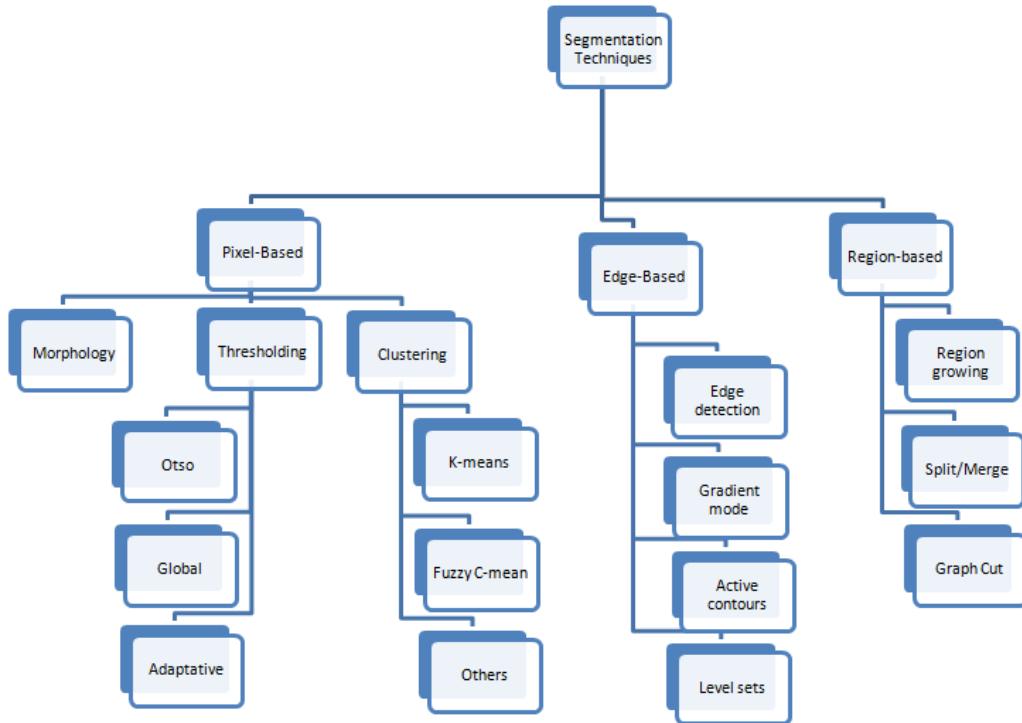


Figura 3.1: Clasificación de las técnicas de segmentación

Fuente: [7]

3.2. Segmentación de imagen basada en el tratamiento de los *pixels*

Este tipo de segmentación consiste en dividir la imagen en segmentos o conjuntos de *pixels* (conocidos como *superpixels*). Cada pixel de la imagen será tratado y agrupado en función de sus características. Existen varios subgrupos dentro de esta clasificación: *Thresholding*, o «Método del valor umbral» en castellano, y *Clustering*, o algoritmos de agrupamiento en castellano.

3.2.1. Thresholding

Este tipo de segmentación es la más simple de todas y se basa en clasificar los *pixels* en dos grupos en función de la intensidad de estos: los que superan la intensidad umbral definida y los que no la superan. El resultado de esta segmentación sería una imagen binaria.

Esta técnica puede ser definida como:

Para una imagen NxM :

for $i = 1, 2, \dots, N$ and $j = 1, 2, \dots, M$

$$f(n) = \begin{cases} 1 & \text{si } I(i, j) \geq T \\ 0 & \text{si } I(i, j) < T \end{cases} \quad (3.1)$$

También existe la posibilidad de definir varias intensidades umbrales (como en la figura 3.4b), con el fin de particionar la imagen en más segmentos. En sí, se podrán definir tantos umbrales como niveles de gris contiene la imagen, aunque habrá que buscar un buen equilibrio.

La ventaja de este tipo de segmentación es que es relativamente sencilla comparada con otros tipos de segmentación más avanzada como *watershed* o *level set*. Aun así, esta segmentación funciona bien cuando el fondo y los objetos siguen una distribución bimodal, es decir, que hay una diferencia «notable» entre las dos partes. Comunmente esta característica no se da en todas las imágenes, por lo que no tendrá buenos resultados con imágenes en las que el fondo no se distinga bien de los objetos [7]. Además, esta segmentación tampoco se comporta bien con imágenes que tienen una iluminación gradiente grande como 3.5b. Aunque es verdad que para ello también existen varias mejoras que hacen que la segmentación se «adapte» para conseguir mejores resultados.

En la figura 3.1 se muestran varias subclases de *thresholding*. Aunque no se explicarán en profundidad conviene saber que hay varias maneras de realizar esta segmentación. Primeramente se nombra el método de Otsu, *Otsu's method* en inglés, que adapta el umbral en función de la dispersión de los niveles de gris. El segundo método, el método *Global*, es el más simple y el que hemos estado explicando hasta ahora. Y por último, el método *Adaptative*, mencionado anteriormente, que se adapta a la intensidad de la imagen.

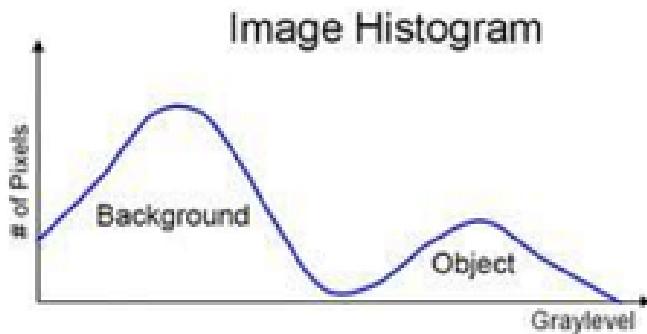


Figura 3.2: Histograma que muestra tres aparentes segmentos de la imagen con dos umbrales

Fuente: [7]

Las figuras 3.3, 3.4 y 3.5 muestran varios ejemplos de segmentación utilizando *thresholding*.

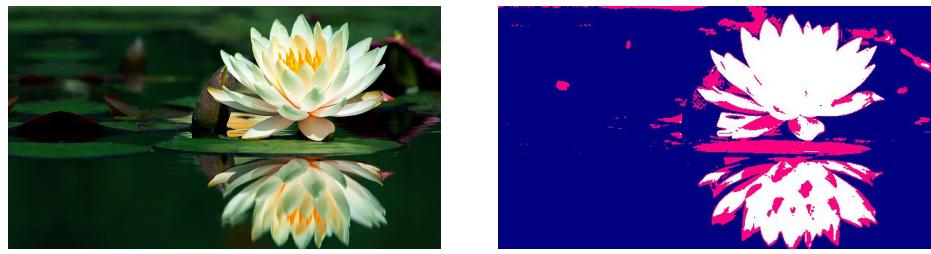


(a) Imagen original

(b) Imagen segmentada

Figura 3.3: Segmentación con *thresholding*

Fuente: commons.wikimedia.org



(a) Imagen original

(b) Imagen segmentada

Figura 3.4: Segmentación con *thresholding* con varios umbrales

Fuente: rosavallsformacio.tv y photo-kako.com para la realización de la segmentación



(a) Imagen original

(b) Imagen segmentada

Figura 3.5: Segmentación con *thresholding* en imagen con iluminación gradiente

Fuente: homepages.inf.ed.ac.uk/rbf/HIPR2/

3.2.2. Clustering

Junto con la técnica de *thresholding* el *clustering* son las técnicas de segmentación más utilizadas. En general esta técnica divide los puntos en varios *clusters* o grupos en función de la distancia entre ellos. En este caso concreto, los puntos serán los pixels de la imagen y la distancia entre ellos estará relacionada con la intensidad, color y textura, pudiendo combinar varios de estos factores.

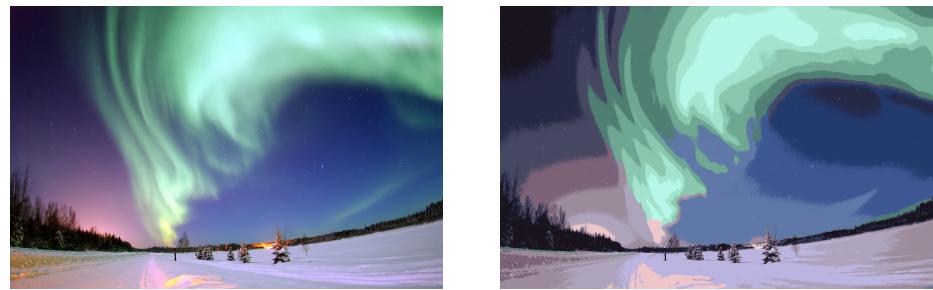
Los algoritmos de *clustering* se pueden dividir entre jerárquicos y particionales, donde la principal diferencia entre los dos está en que el jerárquico produce una serie de particiones anidadas de forma jerárquica y el partitacional genera particiones disjuntas. Los algoritmos jerárquicos suelen ser más precisos, sin embargo, no valen para una cantidad de datos grande como en una imagen ya que el coste computacional es muy elevado. Por lo tanto, la opción escogida suele ser el *clustering* partitacional. No obstante, esta técnica tiene varias desventajas [8]:

1. Generalmente es necesario saber previamente el número de *clusters* que hay en la imagen.
2. No utilizan información espacial inherente a la imagen.
3. En algunos algoritmos de clustering, como el K-means que se explicará a continuación, no se asegura un resultado óptimo, ya que distintas inicializaciones dan como diferentes resultados.

En la clasificación presentada en la figura 3.1 se muestran varias subclases de la técnica de *clustering* que corresponden a diferentes algoritmos de creación de los *clusters*. El algoritmo *Fuzzy C-Means* agrupa los pixels utilizando la lógica difusa, donde cada pixel tendrá un grado de pertenencia a su *cluster*. En general, hay muchos tipos de técnicas de clustering por lo que las técnicas restantes se agruparán en la sección de *Others* de la figura 3.1. El algoritmo K-means es el más utilizado y por ello se explicará con más detalle su funcionamiento:

1. Se asignan K primeros *pixels* como centroides.
2. Se agrupan los *pixels* restantes con los centroides definidos en función de la distancia con estos.
3. Se calculan los nuevos K centroides como los baricentros de los K conglomerados obtenidos.

4. Se alternan los pasos 2 y 3 hasta que se alcance un determinado criterio de convergencia.

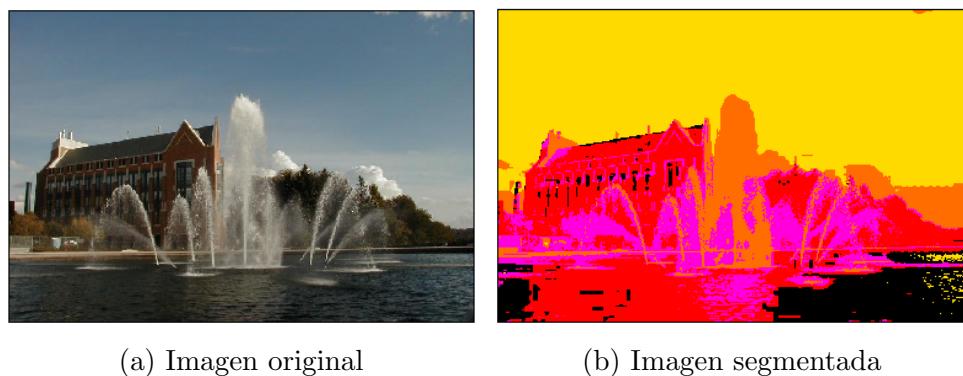


(a) Imagen original

(b) Imagen segmentada

Figura 3.6: Segmentación utilizando K-means con K=16

Fuente: commons.wikipedia.org



(a) Imagen original

(b) Imagen segmentada

Figura 3.7: Segmentación utilizando K-means con K=6

Fuente: imagedatabase.cs.washington.edu/demo/kmcluster/

3.2.3. Morphology

La técnica de la morfología, o *morphology* en inglés, se clasificaría dentro de las técnicas basadas en el tratamiento de los *pixels*. Hay varios métodos de morfología y se basan en una máscara llamada *structuring element* para investigar cada *pixel*. El valor de cada *pixel* está determinado por el de sus vecinos que pertenecen a esa máscara. Los métodos más simples de esta técnica son la dilatación y la erosión. Para una imagen binaria, la dilatación convierte en uno todos los *pixels* de la máscara si los pixels «debajo» del pixel central son ceros como se muestra en la imagen 3.8. La erosión es el

caso contrario, es decir, convierte en ceros todos los elementos de la máscara si esta contiene algún elemento que sea cero. La combinación de estas simples operaciones junto con otras como el complemento, la unión y la intersección, se pueden llegar a realizar operaciones más avanzadas y complejas.

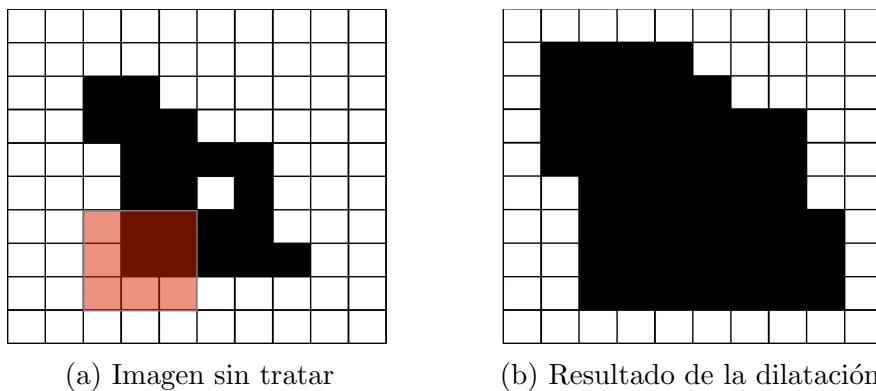


Figura 3.8: Ejemplo de dilatación de la técnica de morfología

Fuente: [9]

3.3. Segmentación de imagen basada en la detección de bordes

Este tipo de segmentación consiste en encontrar los bordes de los objetos contenidos en la imagen con el fin de poder dividir la imagen en función de los bordes encontrados. Los detectores de bordes tradicionales suelen utilizar los operadores diferenciales de detección de bordes comentados en 2.2, es decir, los operadores *Sobel*, *Roberts* y *Prewitt edge detectors* que están basados en el gradiente de la función de intensidad de una imagen. Normalmente los bordes suelen detectarse en las intersecciones de dos regiones de la imagen que tienen diferentes intensidades.

La ventaja de este tipo de segmentación frente a la basada en el tratamiento de pixels es que aparte de hacer la división de las diferentes regiones sabremos exactamente donde se encuentran los bordes de estas, siendo útil para poder extraerlas y poder tratarlas individualmente. Por lo tanto, esta técnica funcionará mejor cuando la diferencia entre las regiones tenga buena calidad. Una de las posibles «desventajas» puede ser que la detección de muchos bordes dificulte la extracción de las regiones de interés.

Existen varios subgrupos dentro de esta clasificación: *Edge detection* o «Detección de bordes» en castellano, técnicas que tienen que ver con el gra-

diente de la imagen *Gradient mode*, *Active contours* o «Contornos activos» y *Level sets* o técnicas del «Conjunto de nivel».

3.3.1. Detección de borde usando gradientes

En la figura 3.1 se diferencian las dos primeras clasificaciones *Edge detection* y *Gradient mode* pero al estar directamente relacionadas se ha decidido explicarlas en conjunto.

Las técnicas clásicas de detección de bordes se basan en encontrar la derivada respecto a los ejes que forman la imagen, o dicho de otro modo, el gradiente. El gradiente de un punto de una función escalar, representado con ∇ , se representa en forma vectorial. Este vector indica la dirección en la cual la función varía más rápidamente y su módulo representa el ritmo de variación de la función en la dirección de dicho vector. Este módulo se utilizará para determinar si un punto es borde o no si supera un valor umbral dado. Para encontrar la máxima variación en ese punto se deben de hacer las derivadas parciales respecto a cada eje y coger el máximo valor de estas. En general el gradiente se suele aproximar con la fórmula¹ $|G| \approx |G_x| + |G_y|$ que es mucho más simple de implementar en la práctica. Valiéndonos de esto, se desarrollaron los primeros operadores diferenciales ya comentados *Sobel*, *Roberts* y *Prewitt edge detectors*. Estos operadores no son más que máscaras aplicadas al pixel a tratar y a cierta vecindad de este para calcular una aproximación a dichas derivadas G_x y G_y . De ahí el nombrarlos como «operadores».

■ Roberts operator

Este operador es el más simple de los tres mencionados y approxima las derivadas tomando la diferencia de dos valores contiguos. La gran desventaja de este operador es que es muy sensible al ruido al tratar pocos vecinos y sólo permite marcar los puntos del borde pero no su orientación. A pesar de todo ello, es un operador que computacionalmente es poco costoso debido a su simpleza y que trabaja bien con imágenes binarias.

¹Válida para una imagen de dos dimensiones. En caso de tener tres dimensiones la fórmula sería $|G| \approx |G_x| + |G_y| + |G_z|$

$$\begin{array}{|c|c|} \hline +1 & 0 \\ \hline 0 & -1 \\ \hline \end{array}
 \qquad
 \begin{array}{|c|c|} \hline 0 & +1 \\ \hline -1 & 0 \\ \hline \end{array}$$

G_x G_y

Tabla 3.1: Máscaras utilizadas por el operador de Roberts de tamaño 2x2

■ Sobel operator

Este operador utiliza una máscara más grande que el *Roberts operator*, 3x3, por lo que implicará a más vecinos. Enfatiza más los pixels de alrededor del centro. La ventaja de este operador es que es menos sensible al ruido, detecta muy bien los bordes horizontales y verticales y además proporciona un suavizado. Las desventaja de este operador es que computacionalmente es más costoso, no tiene buena detección de bordes diagonales y no da información sobre la orientación del borde.

$$\begin{array}{|c|c|c|} \hline -1 & 0 & +1 \\ \hline -2 & 0 & +2 \\ \hline -1 & 0 & +1 \\ \hline \end{array}
 \qquad
 \begin{array}{|c|c|c|} \hline +1 & +2 & +1 \\ \hline 0 & 0 & 0 \\ \hline -1 & -2 & -1 \\ \hline \end{array}$$

G_x G_y

Tabla 3.2: Máscaras utilizadas por el operador de Sobel de tamaño 3x3

■ Prewitt operator

Este operador es parecido al operador de Sobel pero este no enfatiza los pixels cercanos al centro y los coeficientes son diferentes. Las ventajas son que aumenta la respuesta a los bordes diagonales poniéndole peso a pixels vecinos que antes no tenían, tiene poca sensibilidad al ruido y proporciona la magnitud y orientación del borde (hasta 8 direcciones).

$$\begin{array}{|c|c|c|} \hline -1 & +1 & +1 \\ \hline -1 & -2 & +2 \\ \hline -1 & +1 & +1 \\ \hline \end{array}
 \qquad
 \begin{array}{|c|c|c|} \hline +1 & +1 & +1 \\ \hline -1 & -2 & +1 \\ \hline -1 & +1 & +1 \\ \hline \end{array}$$

0 45

Tabla 3.3: Máscaras utilizadas por el operador de Prewitt de tamaño 3x3

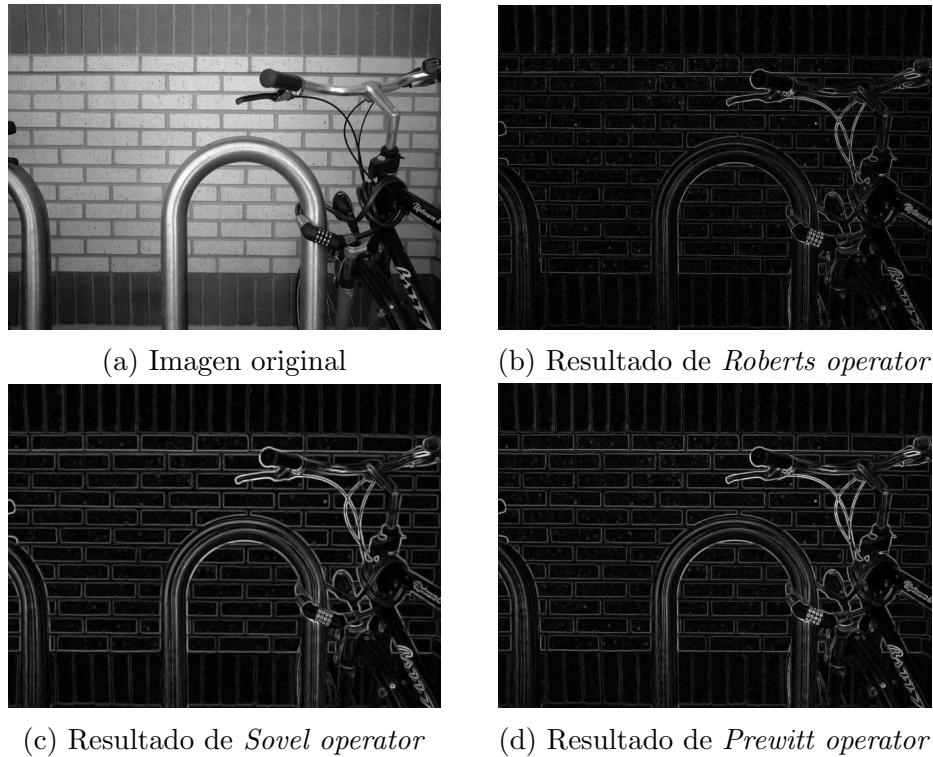


Figura 3.9: Detección de bordes con el uso de los operadores

Fuente: commons.wikipedia.org

3.3.2. Active Contours

Desde que fueron introducidos por Kass y colaboradores en 1988 [10], los contornos activos o más comúnmente nombrados como *Snakes*, han ganado popularidad desde entonces gracias a los buenos resultados que se pueden llegar a obtener en la segmentación de imágenes.

Un contorno activo o *Snake* es una curva elástica que comienza a moverse dada una posición inicial de manera que llega a delimitar las regiones de interés de la imagen. La curva se irá moviendo de manera que se minimice su energía hasta llegar a un punto de convergencia. El contorno puede ser definido paramétricamente como $V(s) = [x(s), y(s)]$ donde $x(s)$ e $y(s)$ son las coordenadas de la parte s del contorno. La energía del contorno está compuesta por una energía interna y otra externa, E_{int} y E_{ext} respectivamente. La definición formal sería:

$$E = \int_0^1 E_{int}(v, s) + E_{ext}(v(s))ds \quad (3.2)$$

E_{int} da las características de deformación del contorno elástico, por lo tanto depende de la forma que este tenga. La E_{int} puede ser definida como:

$$E_{int} = \frac{1}{2}(\alpha|\frac{\delta v}{\delta s}|^2 + \beta|\frac{\delta^2 v}{\delta s^2}|^2) \quad (3.3)$$

y los valores α y β determinan el grado en el que el contorno se puede estirar o curvar. Un aumento en la magnitud α incrementaría la tensión de la curva y un aumento de β incrementaría la rigidez de la curva, haciendo que sea menos flexible.

En cuanto a la energía externa hay varias maneras de definirla. Una elección popular sería la magnitud negativa del gradiente de la imagen que se definiría como:

$$E_{ext} = |\nabla[G_\alpha I(x)]| \quad (3.4)$$

donde G_α es una convolución con un filtro pasabajo gaussiano. Esta propuesta de energía hace que el contorno se expanda hasta los bordes que haya en la imagen como se puede ver en la figura 3.10.

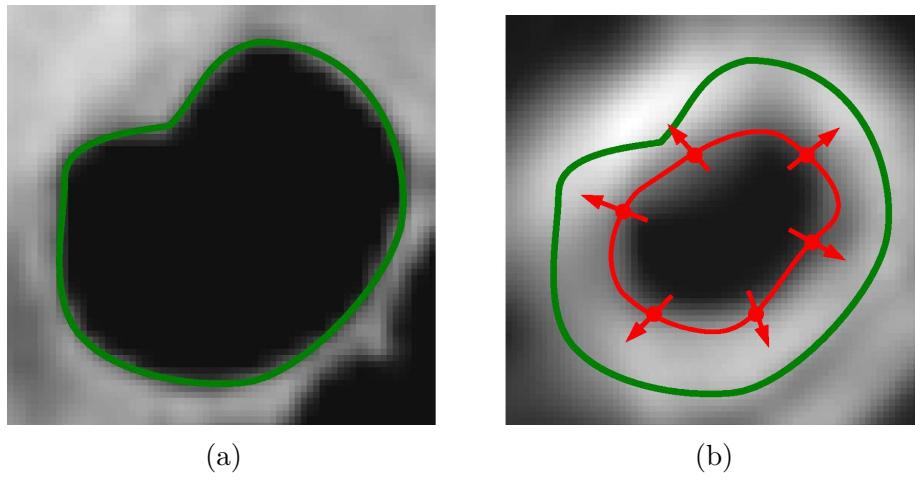


Figura 3.10: Ejemplo de un contorno activo

Fuente: [9]

La imagen 3.10a corresponde a la imagen de entrada, mientras que la imagen 3.10b corresponde a la convolución de la magnitud gradiente de la imagen de entrada con un filtro pasabajo gaussiano. La línea roja en la imagen 3.10b es el contorno activo que se moverá hasta la línea verde que corresponde con el borde de la imagen original.



Figura 3.11: Ejemplo de segmentación de plantas presentando en [2] utilizando los contornos activos

Fuente: [2]

3.3.3. Level set

Conjuntos de nivel, o *Level set* en inglés, es un tipo de segmentación muy parecida a la que hemos explicado en la sección 3.3.2 ya que también se trata de expandir un contorno dado previamente para encontrar los bordes de la imagen. La ventaja de este método frente al referido es que permite juntar y dividir contornos sin ningún cálculo extra necesario.

El contorno está representado por la función de level set, que está definida en una dimensión más que las dimensiones de la imagen a segmentar, es decir, para una imagen 2D tendríamos una superficie de level set de tres dimensiones mientras que para una imagen de 3D tendríamos una superficie de 4D. En el caso de una superficie de 3D esta tiene una forma cónica como se puede ver en 3.12. Suponiendo entonces que la imagen a tratar tiene dos dimensiones, podríamos definir la función de level set como $z = \phi(x, y, t)$ que devuelve la altura de la superficie de level set en el punto (x, y) del plano de la imagen en el tiempo t . El contorno es definido implícitamente como «zero level set», donde la altura del plano respecto a la superficie es cero ($\phi(x, y, t) = 0$). Esto es justo la intersección entre el plano de la imagen y la superficie.

Para propagar el contorno se mueve la superficie de level set de manera que esta se expande, sube y baja para encontrar el frente. Suponiendo que cada punto del contorno se mueve en una dirección normal frente al contorno

con una velocidad F , el contorno evoluciona usando la siguiente PDE (*partial differential equation* o ecuación en derivadas parciales en castellano):

$$\frac{\delta\phi(x, y, t)}{\delta t} = F(x, y, I)|\nabla\phi(x, y, t)| \quad (3.5)$$

Esta función de velocidad varía dependiendo del punto de la imagen I a tratar y hace que el contorno se expanda a ciertas áreas de la imagen y no lo haga en otras zonas de esta. Normalmente la función de velocidad se define por la intensidad o el gradiente de los pixels y por la curva de la función de level set.

De la idea de que modificaciones de *pixels* lejanos al contorno no afectan a este surgen varias mejoras de esta técnica que tienen en cuenta los *pixels* con los que se trabajará en cada iteración: *narrow band* y *sparse field methods*. El método de *narrow band* actualiza los *pixels* en una línea estrecha alrededor del contorno. El método *sparse field* actualiza los *pixels* vecinos del contorno únicamente.

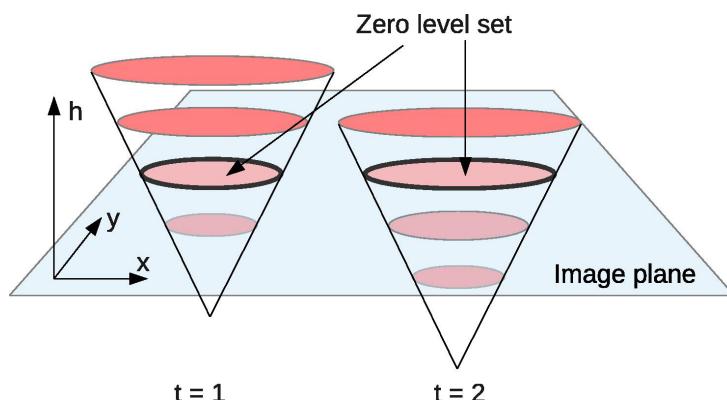


Figura 3.12: Ilustración de la superficie de *level set*

Fuente: [9]

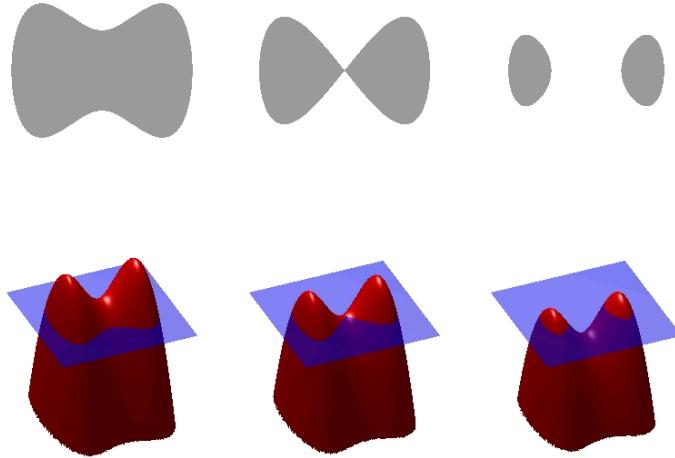


Figura 3.13: Ilustración del método de *level set*

Fuente: commons.wikipedia.org

3.4. Segmentación de imagen basada en regiones crecientes

Este tipo de segmentación, más conocida como *region growing*, se basa en la idea de que los *pixels* de una región tienen características comunes, como puede ser la intensidad de gris etc. Por ello, al tener que tratar un nuevo *pixel* si este tiene una intensidad de gris parecida a la intensidad de grises que contiene la región significará que ese punto pertenece a ella.

Existen varios subgrupos dentro de esta clasificación: técnicas basadas en la idea principal de regiones crecientes, o *region growing* en inglés, técnicas de *Split/Merge* y técnicas basadas en grafos *Graphs cuts*.

3.4.1. Region growing

Este subgrupo agrupa todos las técnicas relacionadas con la técnica «habitual» explicada en la introducción de esta sección (3.4). Hay dos técnicas conocidas que siguen esta metodología, por lo que en esta ocasión explicaremos más de una técnica es este subgrupo de segmentación.

■ ***Region growing o Seeded-based region growing segmentation***

Este tipo de segmentación comienza con la selección de un *pixel*, nombrado a menudo como semilla o *seed* en inglés, que está dentro del objeto de interés. Normalmente la semilla se elige manualmente. A partir de ese *pixel* semilla (primer punto de la región) se comenzará a extender la región procesando sus vecinos y añadiéndolos en base a un criterio predefinido. Este criterio de inserción será en base a la intensidad, color o textura de la semilla y los puntos que pertenezcan a la región. Cada vez que se inserta un nuevo punto a la región la característica que se este utilizando para realizar la inserción se volverá a calcular, por ejemplo, si se utiliza el nivel de gris, se volverá a calcular el valor medio de los niveles de gris que hay en la región. De esta manera, la región se irá expandiendo añadiendo vecinos hasta que encuentre alguno que no cumplan con la condición de inserción impuesta por el criterio. Si un punto no ha sido añadido a ninguna región se podrá añadir a una región cercana suya si la diferencia entre el nivel de gris de este punto y el nivel de gris medio de la región no supera un valor umbral T dado.

Esta técnica es útil cuando la intensidad del fondo y del objeto son muy parecidas pero están separadas por un borde «notable» o por otra región.

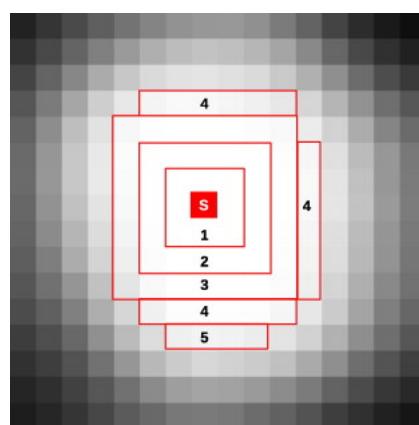
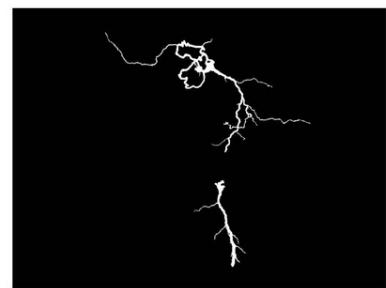
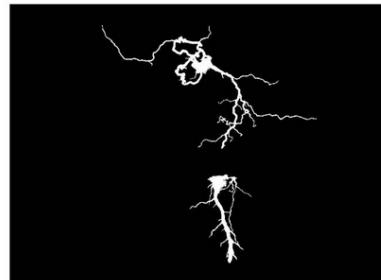
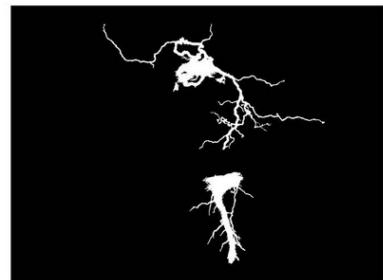


Figura 3.14: Ejemplo del funcionamiento de la técnica de *region growing*

Fuente: [9]



(a) Imagen original

(b) *Region growing* con un valor umbral $T=255$ (c) *Region growing* con un valor umbral $T=225$ (d) *Region growing* con un valor umbral $T=190$ (e) *Region growing* con un valor umbral $T=155$ Figura 3.15: Ejemplo de la técnica de *region growing*

Fuente: commons.wikipedia.org

Ejemplo de la técnica de *region growing* donde se quiere encontrar la parte del rayo más fuerte en la imagen. Para ello se eligen como semilla los puntos con mayor valor de gris posible(255). El criterio de inserción de los *pixels* es tener el "mismo" nivel de gris. En este caso se ha decidido que los puntos que no han sido insertados en una región se inserten en una cercana si superan el

valor umbral T dado. Por ello en 3.15b solo aparecen los puntos semilla ya que no se habrán unido puntos con este criterio y el valor T es demasiado alto. En las imágenes 3.15c, 3.15d y 3.15e ese valor T se va disminuyendo y cada vez se insertan más puntos en la región.

■ Watershed

La idea principal de esta técnica se basa en ver la imagen como una imagen tridimensional donde la tercera dimensión es la altura del *pixel*. La altura del *pixel* está determinada por su nivel de gris. Topológicamente quedará algo como se puede ver en 3.16. En este «terreno» creado se podrán diferenciar hasta tres puntos. Estos puntos son determinados con la analogía de como una gota de agua caería y se movería si se precipitase en ese punto. Hay tres tipos de puntos:

1. Puntos con un nivel de gris mínimo local donde la gota se estancaría.
2. Puntos en los que la gota caería o se deslizaría hacia otros puntos más bajos.
3. Puntos en los que la gota podría caer en más de un punto mínimo local.

El segundo tipo de puntos son nombrados como *watersheds*, o cuencas en castellano, y los del tercer tipo son nombrados como *watersheds lines*, bordes o líneas de las cuencas en castellano.

El objetivo final de esta técnica será encontrar los bordes de las cuencas, que representarán los bordes de la imagen original. Para ello existen varias maneras de hacerlo, la más común es la técnica de *flooding*. La idea es simple, imaginemos que se empieza a hechar agua en los puntos de tipo uno, es decir, los que representan un mínimo local y son «cuencas». El nivel del agua empezará a subir hasta que llegue a un punto en el que la cuenca se empiece a desbordar y vaya a juntarse con otra cuenca. En ese momento, se construye una presa o un muro de manera que el agua no se desborde. Esas presas o muros construidos serán los bordes de las cuencas. Con todas estas presas se habrá conseguido la segmentación de la imagen.

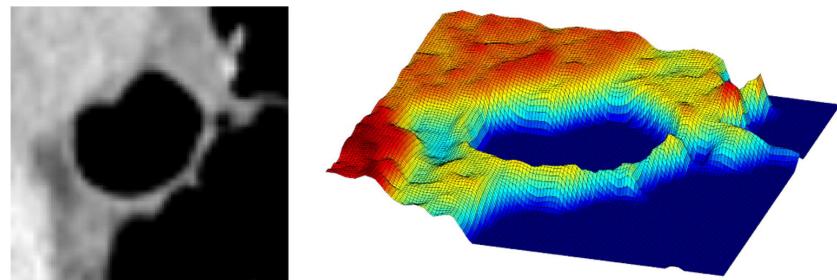


Figura 3.16: Ejemplo del funcionamiento de la técnica de *watershed*

Fuente: [9]

Ejemplo del funcionamiento de la técnica de *watershed* donde la intensidad de los pixels de la imagen de la izquierda serán la altura de ese mismo punto en la imagen derecha, creando así ese terreno.



Figura 3.17: Funcionamiento de la técnica de *watershed*

Fuente:

<http://www.di.ubi.pt/~agomes/cvm/teoricas/07-regionsegmentation.pdf>

Funcionamiento de la técnica de *watershed* donde se puede ver como va creciendo el nivel de agua hasta encontrar ese «punto» de desbordamiento en el que se marcan los bordes.

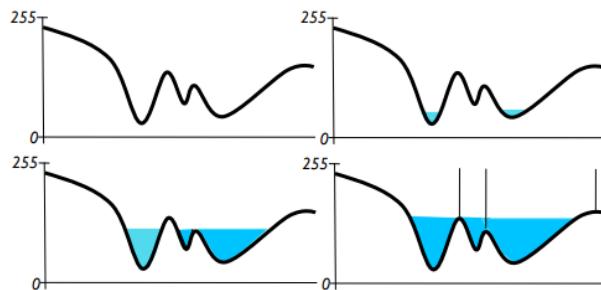


Figura 3.18: Funcionamiento de la técnica de *watershed*

Fuente:

<http://www.di.ubi.pt/~agomes/cvm/teoricas/07-regionsegmentation.pdf>

3.4.2. Split/Merge

La técnica de *split*, ha diferencia de la técnica de *region growing* que empieza con una serie de puntos «semilla», empieza con la imagen entera como una única región y va subdividiendo la imagen recursivamente en regiones más pequeñas en base a un criterio de homogeneidad.

En cuanto a la técnica de *merge*, es lo contrario que la de *split*, ya que esta empieza con pequeñas regiones de 2x2 o 4x4 y las va juntando entre sí en base a si tienen o no características comunes entre ellas como el nivel de gris, el color, la textura etc.

3.4.3. Graph cut

La idea principal de esta segmentación es representar la imagen a tratar como un grafo, donde normalmente cada *pixel* es un nodo y tienen aristas con los nodos vecinos. La ventaja de estos algoritmos esta en que pueden llegar a trabajar bien incluso si la separación entre dos regiones esta «rota» o es dudosa. Hay varios algoritmos distintos dentro de esta clasificación, en este caso se ha elegido explicar el método de segmentación de *Markov*, o *Markov random field (MRF) segmentation* y concretamente una variante de este llamada *graph cut*.

MRF considera a cada *pixel* como un nodo del grafo y tienen aristas con cada *pixel* vecino. Sin embargo, cada nodo tiene dos conexiones más a un par de nodos especiales llamados *source* (S) y *sink* (T) como se muestra en la figura 3.19. Se les añade un peso a las aristas entre los nodos de manera que los pixels que pertenecen al fondo tienen un peso pequeño en la arista que los une con uno de esos dos nodos anteriores y un peso grande con el otro nodo. De forma inversa, los *pixels* que pertenezcan al primer plano tendrán

un peso pequeño con uno de ellos y un peso grande con el otro. Por otro lado, los pesos de las aristas entre los *pixels* son grandes cuando los *pixels* tienen características comunes y un peso pequeño en caso contrario.

La segmentación se realiza aplicando un algoritmo de corte de grafos. El objetivo es minimizar la suma de las aristas por las que se va a cortar el grafo para ello hay varios algoritmos para buscar el mínimo corte ha realizar.

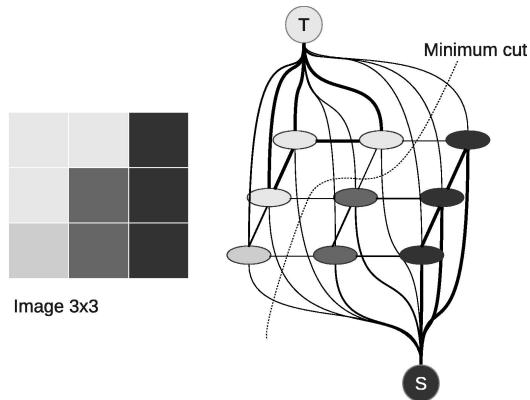


Figura 3.19: Funcionamiento de la técnica de *graph cut* en una imagen 3x3

Fuente: [9]

3.5. Conclusiones de los tipos de segmentación

Como se ha podido ver a lo largo de este capítulo, hay muchos tipos de técnicas de segmentación de imágenes. Además, existen más técnicas de segmentación que no se han explicado en este trabajo ya que se han añadido las más conocidas y populares con el fin de establecer una base para el lector.

A parte de las técnicas que no se han explicado hay que decir que existen muchas mejoras, optimizaciones y variantes de todas las técnicas de segmentación. Por ello en este trabajo se han explicado las ideas principales de los algoritmos sin tener en cuenta esto ya que el trabajo se habría extendido demasiado.

Capítulo 4

Necesidades del cliente

La primera toma de contacto con el director del proyecto interno de la fundación del DIPC fue en el mes de febrero. El proyecto que se está desarrollando es confidencial por lo que no se puede explicar en esta memoria ningún dato sobre el mismo exceptuando la parte que se ha desarrollado en este proyecto. Por lo que, lo primero que nos presentó el cliente fueron las necesidades que el tenía sobre la segmentación de imágenes.

4.1. Requisitos del proyecto del DIPC

El en proyecto que el cliente está desarrollando hay una parte que necesita realizar un tratamiento de una imagen para sacar de esta varias conclusiones. El cliente conoce la teoría sobre un algoritmo llamado *level set* que realiza segmentación de imágenes con bastante precisión y que está bastante extendido. Sin embargo, no conoce ninguna implementación de este algoritmo ni el tiempo que puede costar realizar dicha segmentación. Recordando la introducción del proyecto, las imágenes se han sacado de una proyección en cierta lámina de materiales sintetizados mediante la deposición química de vapor, por lo que tendrán una apariencia similar a las presentadas en la figura 4.1. Las características que el cliente quiere obtener de la imagen son las siguientes:

1. Encontrar las islas¹ contenidas en la imagen
2. Determinar el número de islas existentes
3. Porcentaje de recubrimiento de las islas respecto al tamaño de la imagen
4. Aislamiento de cada isla

5. Densidad espectral de cada isla

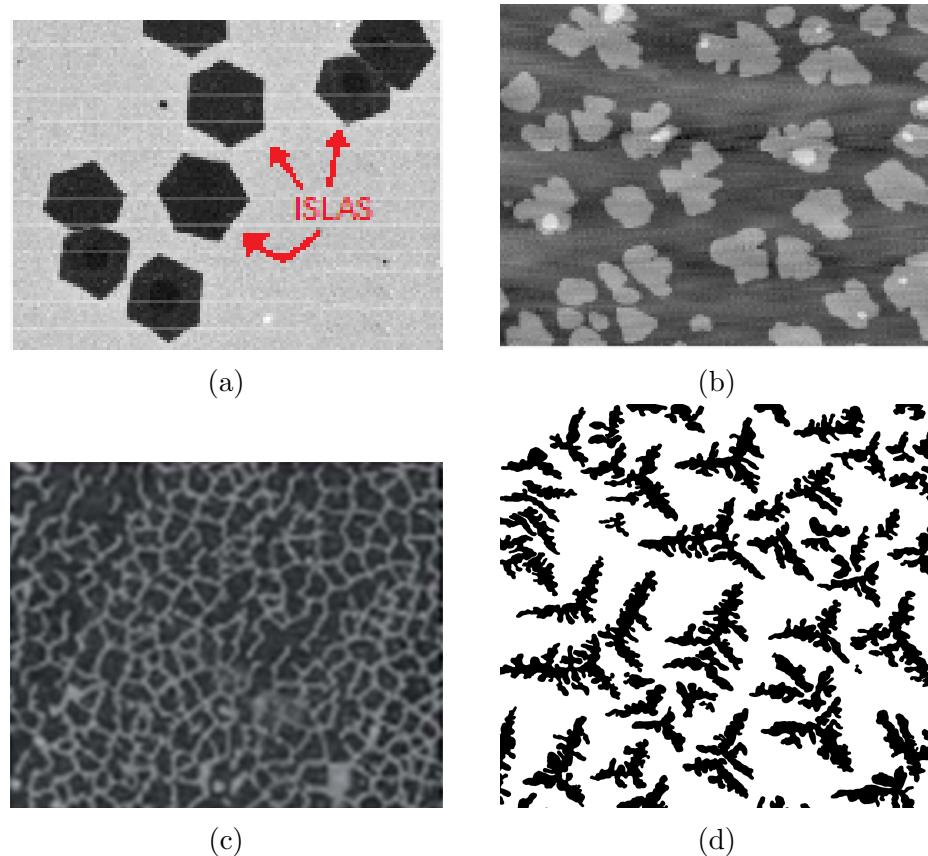


Figura 4.1: Ejemplo de proyecciones en láminas de materiales sintetizados mediante la deposición química de vapor

Las islas pueden estar muy separadas, como se puede ver en la imagen 4.1a o muy juntas como en la imagen 4.1c. También se puede observar que la forma de estas islas no es siempre la misma. Estas características dependen de ciertas variables físicas que hacen que el material se «pegue» de una determinada manera en la lámina, dando lugar a estas imágenes abstractas.

¹se llamará «isla» a cada superficie que se distinga sobre el fondo de la imagen. Ver ejemplo en 4.1a

4.2. Necesidades extra

A parte de los principales requisitos también se propusieron necesidades extras que el proyecto del cliente necesitaría. Estas son las necesidades extra:

1. Realizar la segmentación en muy poco tiempo, en fracciones de segundo o en muy pocos segundos.
2. Ser capaz de analizar un vídeo de la evolución de la deposición de estos materiales *frame a frame*. Para ello es necesario cumplir con la anterior idea, ya que un vídeo esta compuesto en cada segundo por multiples *frames*.

Capítulo 5

Optimizaciones y mejoras del *level set* original

Como se ha dado a conocer en el capítulo 3.5 hay muchos tipos de mejoras y optimizaciones de cada técnica de segmentación existentes. A la hora de afrontar el problema que el cliente proponía se tenía que elegir un tipo de segmentación de las técnicas estudiadas. Entre todas ellas se escogió el algoritmo de *level set*, por ser un algoritmo de segmentación famoso, lo que facilitaría la búsqueda de información sobre él, y además, el cliente había asegurado que se iban a tener buenos resultados con esta técnica según su propia experiencia.

Repasando este algoritmo, las ventajas que tiene son que la segmentación conseguida es precisa y que no añade sobrecoste al hacer la división o unión del contorno. Sin embargo, el algoritmo se basa en la resolución de ecuaciones diferenciales, lo que supone que el algoritmo sea costoso y lento. Se han desarrollado varias implementaciones en GPU [9], también sobre imágenes 3D [11] y hasta en proyectos de fin de carrera también se utilizan GPUs [12] etc. De la misma manera, también se han desarrollado paralelizaciones en arquitecturas SMP(*Symmetric Multi-Processing*) [13] o en arquitecturas de memoria distribuida con MPI(*Message Passing Interface*) [14].

Hay muchos artículos sobre este algoritmo y alguna implementación paralela pero sobre el algoritmo original, por lo tanto, a pesar de la paralelización consiguen tiempos elevados por encima de las necesidades de este proyecto. Estas paralelizaciones y otras técnicas de optimización que presentan varios trabajos se centran siempre en resolver las PDEs asociadas a la evolución del *level set*. Sin embargo, para muchos problemas de imagen, como la segmentación, no es necesario tanta precisión ya que el objetivo final es encontrar los bordes de los objetos. En este caso, el proceso evolutivo no tiene tanto interés como el resultado final. Siguiendo esta idea Shi y Karl presentaron un

artículo muy interesante que será la pieza clave de la segmentación realizada en este proyecto [15].

5.1. Aproximación a la técnica de *level set*

Como se ha mencionado anteriormente, no se tienen por qué resolver las PDEs en el algoritmo de *level set* ya que, en la segmentación, importa más el resultado final que la evolución propia del algoritmo. Además, la resolución de PDEs conlleva el que se tengan que realizar reinicializaciones de la función de *level set* (representada con la letra ϕ) lo que implica aún más cálculo. El trabajo presentado por Shi y Karl [15] elimina la reinicialización al tener una colección de enteros (los cuales representan la función ϕ) que cambian dinámicamente según va propagándose el contorno y no calculan PDEs. Estas dos mejoras elementales hacen que su algoritmo sea mucho más rápido que el *level set* original.

En este trabajo se presenta una nueva estrategia de implementación del método de *level set*. En el ejemplo de un espacio Euclídeo de dos dimensiones, una curva C es representada implícitamente como el *zero level set* de una función ϕ definida en una cuadrícula como se muestra en la figura 5.1. La función ϕ tendrá un valor negativo dentro de la curva C y un valor positivo fuera de ella, por eso se dice que representa implícitamente a la curva C .

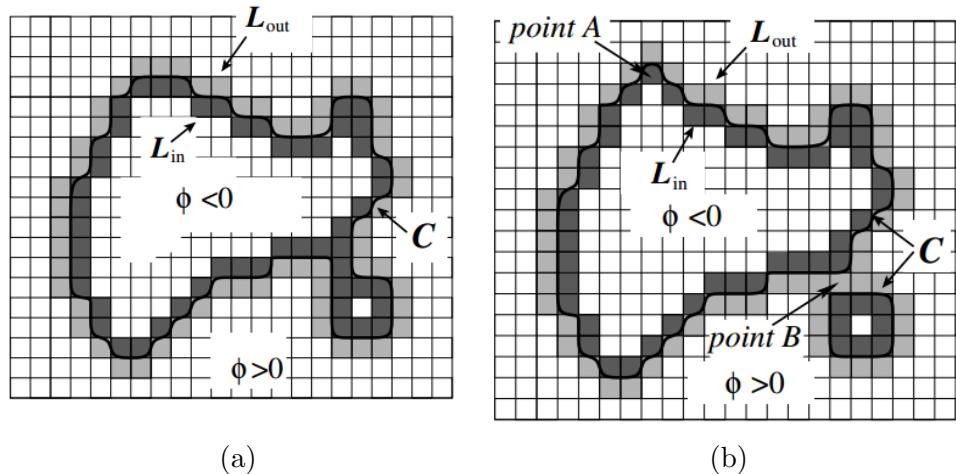


Figura 5.1: Representación de la curva C

Fuente: [15]

Se definen dos listas de vecindad en esta cuadrícula, L_{in} y L_{out} . En la

Imagen se puede ver que el movimiento de la curva C se puede conseguir moviendo un punto de una lista a otra.

Estando entonces en una representación de dos dimensiones, y siguiendo el ejemplo de la rejilla mostrada anteriormente se pueden definir dos listas de vecindad del contorno C: L_{in} y L_{out} .

$$L_{in} = \{x | \phi(x) < 0 \text{ y } \exists y \in N(x) \text{ tal que } \phi(y) > 0\}$$

$$L_{out} = \{x | \phi(x) > 0 \text{ y } \exists y \in N(x) \text{ tal que } \phi(y) < 0\}$$

Siendo x una coordenada de la rejilla denotada como $x = x_1, x_2, \dots, x_k$ y $N(x)$ como un vecino de x con valor discreto definido como:

$$N(x) = y \in D | \sum_{k=1}^K |y_k - x_k| = 1 \forall x \in D \quad (5.1)$$

Como podemos observar en la figura 5.1 la lista L_{in} está formada por los puntos de la rejilla que están dentro de la curva C y L_{out} está formada por los puntos de la rejilla que están fuera de C. Por lo tanto, como se puede ver en las definiciones formales de las listas, cada punto de ellas tiene que tener un punto vecino de la otra lista, de manera que las dos estén "pegadas".

Recordando lo visto en la sección de explicación del level set (3.3.3) en el clásico *level set* la siguiente PDE es resuelta para evolucionar el contorno C bajo una función de velocidad F:

$$\frac{d\phi}{dt} + F|\nabla\phi| = 0 \quad (5.2)$$

La figura 5.1b muestra el proceso evolutivo de la curva C de la figura 5.1a. En el punto marcado como A la curva se ha movido hacia fuera lo que ha modificado el valor de la función ϕ de positivo a negativo. En el punto B la curva se ha movido hacia dentro, partiendo la curva en dos y cambiando el valor de la función ϕ de negativo a positivo. Todo esto también ocurre en el *level set* original, con la diferencia de que la resolución de la PDE mostrada en la ecuación 5.2 tiene un coste elevado. Se consiguen los mismos resultados finales fácilmente si se usa la relación entre C, L_{in} y L_{out} . Para mover la curva hacia fuera en el punto A de la rejilla tendremos que pasar el punto de la lista L_{out} a L_{in} . De manera similar, para mover la curva hacia dentro en el punto B tendremos que cambiar el punto B de L_{in} a L_{out} . En general, con aplicar dichas operaciones se va moviendo la curva hacia cualquier punto con el mínimo coste operacional.

Algoritmo

Para la realización dle algoritmo son necesarias estas estructuras:

- Un array para la función de level set ϕ ;
- Un array para la velocidad (F) con la que se propagará la curva;
- Dos listas de vecindad de la curva: L_{in} y L_{out} .

Nombraremos a los puntos que están dentro de C pero no en la lista L_{in} como «puntos interiores» y a los puntos que están fuera de C pero que no pertenecen a L_{out} como «puntos exteriores». Para agilizar aún más el cálculo, los valores que puede tomar la función ϕ son cuatro enteros: -2, -1, 1, 3.

$$\phi(x) = \begin{cases} 3 & \text{si } x \text{ es un punto exterior} \\ 1 & \text{si } x \in L_{out} \\ -1 & \text{si } x \in L_{in} \\ -3 & \text{si } x \text{ es un punto interior} \end{cases} \quad (5.3)$$

Para la función F de velocidad sólo se usa el signo, por lo que también es un *array* de enteros con los valores: 1, 0, -1. En cuanto a las listas, son listas ligadas de manera que la inserción y el borrado de puntos se pueden hacer de manera rápida.

Antes de presentar el algoritmo se aclararán ciertas cuestiones. Para empezar, hay dos operaciones básicas que se utilizan en el algoritmo:

- La operación $switch_in()$ para un punto $x \in L_{out}$ se define como: $switch_in(x)$:
 - Paso 1: Se quita el punto de L_{out} y se pasa a L_{in} . Se cambia el valor de ϕ en ese punto: $\phi(x) = -1$.
 - Paso 2: Se añaden los puntos vecinos de x a L_{out} y se cambian sus respectivos valores en ϕ . Más formalmente: $\forall y \in N(x)$ que satisfaga $\phi(y) = 3$ se añade y a L_{out} y se pone $\phi(y) = 1$

Con esta operación se mueve el contorno un punto de la rejilla hacia fuera.

- Similarmente se define la operación $switch_out()$ para un punto $x \in L_{in}$: $switch_out(x)$:
 - Paso 1: Se quita el punto de L_{in} y se pasa a L_{out} . Se cambia el valor de ϕ en ese punto: $\phi(x) = 1$.

- Paso 2: Se añaden los puntos vecinos de x a L_{in} y se cambia sus respectivos valores en ϕ . Más formalmente: $\forall y \in N(x)$ que satisfaga $\phi(y) = -3$ se añade y a L_{out} y se pone $\phi(y) = -1$

Con esta operación se mueve el contorno un punto de la rejilla hacia dentro.

Por otro lado, la función F de velocidad presentada antes, normalmente se suele separar en dos velocidades: F_{ext} que depende de los datos y F_{in} para la realización de un suavizado del contorno. La velocidad F_{int} es normalmente la curvatura de la curva [16]. Sin embargo, esta evaluación de la curva usando la función ϕ suele ser computacionalmente costosa. Tras varios pasos esta función de velocidad F_{int} puede llegar a ser determinada por un filtro Gaussiano que puede ser aproximado con operaciones con enteros, por lo que se reduce el cómputo. Además, al separar las dos velocidades, no tiene porque hacerse el suavizado después de cada iteración de evolución del contorno como en otros trabajos [15], si no que se realizará cuando se satisfaga cierta condición, lo que reduce aún más el coste. Con esto entonces, el algoritmo tendrá dos ciclos principales: el primer ciclo en el que se expandirá o evolucionará el contorno y el segundo ciclo(que se realizará de vez en cuando) en el que se suavizará el contorno para que se siga expandiendo con normalidad.

- Paso 1: Inicializar el «array» ϕ , F_{ext} y las dos listas L_{out} y L_{in} .
- Paso 2: Primer ciclo donde se escanean las dos listas para actualizar ϕ , L_{out} y L_{in} .
 - Se calcula la velocidad para cada punto de L_{out} y L_{in} .
 - Evolución hacia fuera. Recorremos la lista L_{out} y se hace la operación $switch_in(x) \forall x \in L_{out}$ si $F(x) > 0$.
 - Se eliminan los puntos redundantes en L_{in} (véase la figura 5.2). Para ello se tendrá que recorrer la lista y para cada punto $x \in L_{in}$, si $\forall y \in N(x); \phi(y) < 0$, se borra x de L_{in} y se cambia $\phi(x) = -3$.
 - Evolución hacia dentro. Recorremos la lista L_{in} y se hace la operación $switch_out(x) \forall x \in L_{in}$ si $F(x) < 0$.
 - Se eliminan los puntos redundantes en L_{out} . Para ello se tendrá que recorrer la lista y para cada punto $x \in L_{out}$, si $\forall y \in N(x); \phi(y) > 0$, se borra x de L_{out} y se cambia $\phi(x) = 3$.
 - Se comprueba la condición de parada y si se satisface se continua al paso 3, si no, seguiremos en el paso 2.
- Paso 3: Segundo ciclo donde se realiza un suavizado del contorno con un filtro Gaussiano.
 - Evolución hacia fuera. Recorremos la lista L_{out} y se calcula $G \oplus \phi(X)$. Si $G \oplus \phi(X) < 0$ se realiza la operación $switch_in(x)$
 - Se eliminan los puntos redundantes en L_{in} (véase la figura 5.2). Para ello se tendrá que recorrer la lista y para cada punto $x \in L_{in}$, si $\forall y \in N(x); \phi(y) < 0$, se borra x de L_{in} y se cambia $\phi(x) = -3$.
 - Evolución hacia dentro. Recorremos la lista L_{in} y se calcula $G \oplus \phi(X)$. Si $G \oplus \phi(X) > 0$ se realiza la operación $switch_out(x)$
 - Se eliminan los puntos redundantes en L_{out} . Para ello se tendrá que recorrer la lista y para cada punto $x \in L_{out}$, si $\forall y \in N(x); \phi(y) > 0$, se borra x de L_{out} y se cambia $\phi(x) = 3$.
- Paso 4: Si se satisface la condición de parada del ciclo uno, se termina el algoritmo, si no, se vuelve al paso 2.

Tabla 5.1: Algoritmo completo de la aproximación del *level set*

Fuente: [15]

Aclaradas las cuestiones anteriores, el rápido algoritmo de *level set* [15] se muestra en la tabla 5.1. Como se puede observar el segundo ciclo es prácticamente igual al primer ciclo, a excepción de que la condición de cambiar un punto de una lista a otra, es decir, de hacer una de las operaciones *switch()* que se han explicado anteriormente, depende de un filtro Gaussiano en el segundo ciclo (velocidad F_{int}) y de los datos en el primer ciclo (velocidad F_{out}). Las operaciones de *switch-in()* se realizarán cuando el valor de F_{int} sea positivo, que querrá decir que el contorno deberá expandirse hacia fuera

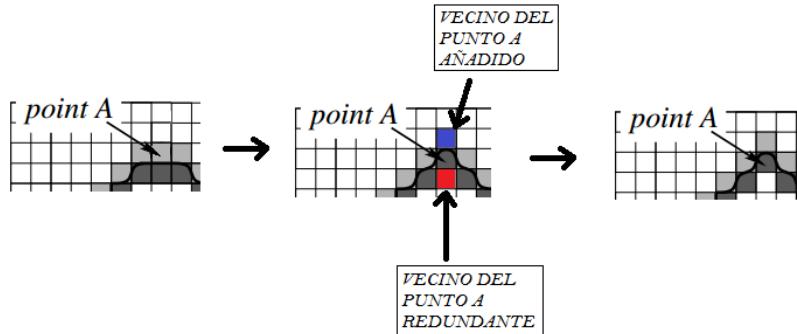


Figura 5.2: Ilustración de las operaciones realizadas en el paso 3 del algoritmo

Fuente: modificaciones de imágenes en [15]

y las operaciones de *switch_out()* se realizarán cuando el valor de F_{int} sea negativo, contrayendo el contorno en ese punto.

Repasando un poco más el algoritmo 5.1 queda por aclarar las condiciones de parada del algoritmo. Así pues, el algoritmo se parará en el caso de que se cumpla una de estas dos condiciones:

1. Que la velocidad F_{ext} de toda la vecindad del contorno cumpla con:
 - a) Formalmente: $F(x) \leq 0 \forall x \in L_{out}$, es decir, que no se tenga que expandir el contorno por ninguno de sus puntos.
 - b) Formalmente: $F(x) \geq 0 \forall x \in L_{in}$, es decir, que no se tenga que contraer el contorno en ninguno de sus puntos.
2. Que se alcance un determinado número de iteraciones establecido.

Para finalizar, el coste del algoritmo es de orden $O(2A(P_1 + P_2))$, donde A es el número de puntos entre el primer contorno creado y el último ya evolucionado, P_1 y P_2 el coste de las operaciones de *switch()* y el coste que tiene el borrado y la inserción de los puntos en las listas respectivamente. El escalar 2 es debido a que los puntos pueden llegar a ser pertenecer a una lista primero, y luego a otra, realizando así las operaciones dos veces. Nótese que ese A será mucho menor que la $O(mn)$ que sería la anchura (m) por la altura(n) de la imagen, ya que si en la imagen se detectan varios objetos, el tamaño en puntos o *pixels* de cada objeto se restará a ese $O(mn)$.

5.2. Evolución del contorno

Se ha querido añadir esta sección para destacar y hablar un poco más en detalle sobre la evolución de la curva que se ha elegido en esta aproximación: la presentada en [16]. Este popular trabajo es conocido como *Chan-Vese model* y a diferencia de como se realizaba la segmentación, utilizando el uso de detectores de bordes basados en el gradiente, este trabajo realiza la detección de objetos cuyos bordes no tienen porque estar definidos por el gradiente. Este método es flexible y potente, capaz de realizar la segmentación de muchos tipos de imágenes que con métodos tradicionales como *thresholding* o métodos basados en el gradiente no podían realizarse. Este método también permite que el contorno se inicialice dentro del objeto o incluso entre el objeto y el fondo ya que el resultado final será siempre el mismo. La figura 5.3 muestra las posibles inicializaciones. Esta característica será de utilidad como se comentará en posteriores capítulos.

Para tener una mínima idea del avance que esto supuso en la segmentación de imágenes, si volvemos a utilizar el buscador [1] y encontramos dicho artículo, el número de citaciones de otros artículos registrados en ese explorador alcanza casi la cifra de 2.000, lo que tampoco quiere decir que no haya más trabajos no registrados en [1] que lo citen.

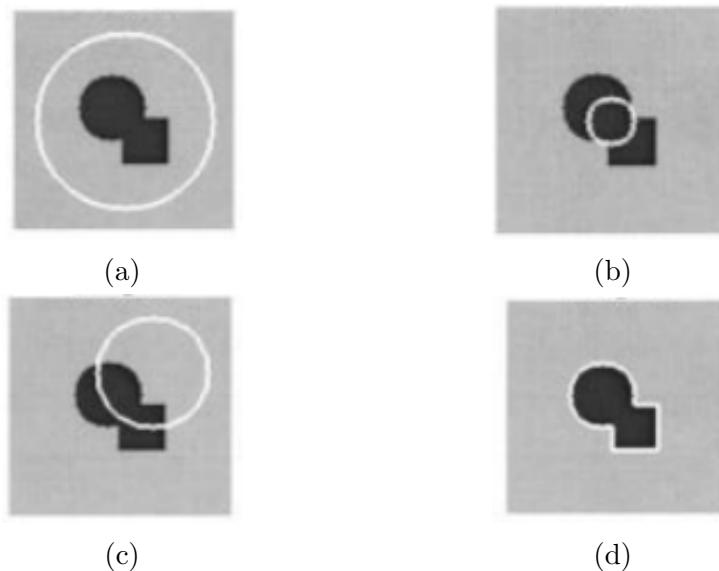


Figura 5.3: Posibles inicializaciones del contorno con el método de Chan-Vese

Fuente: [16]

Capítulo 6

Implementación de la aproximación y Ofeli

Una vez encontrada una técnica que satisficiera los requisitos del proyecto, se empezó a buscar alguna implementación sobre esta técnica. En esa búsqueda se encontró un trabajo llamado Ofeli (*Open, Fast and Efficient Level set Implementation*) [3]. Este trabajo está desarrollado con Qt que es una de librerías multiplataforma para la realización de aplicaciones con GUI (*graphical user interface*) en C++. Esta aplicación es muy completa ya que aparte del algoritmo de interés implementa varias funcionalidades extra como varios tipos de filtrado, preprocesamiento, varios tipos de evolución del contorno etc.

6.1. Ofeli

En esta sección se explicará la estructura de la implementación y las clases que se utilizan en este trabajo. Se realizará un esquema general para que el lector pueda comprender cuales son los procesos o pautas que se dan en esta implementación para poder segmentar una imagen.

6.1.1. Estructura

La implementación está formada por varias clases que se han dividido en dos tipos: las clases o ficheros que pertenecen a la GUI (etiquetadas con la palabra «GUI-» por delante de sus nombres) y las clases que pertenecen a la implementación del algoritmo (etiquetadas con la palabra «Impl-»).

1. **Impl-ActiveContour:** clase padre de todos los tipos de contornos.
Formada por los ficheros:

- a) activecontour.cpp
 - b) activecontour.hpp
2. **Impl-ACwithoutEdges**: clase que implementa el contorno que evolucionará en una imagen a escala de grises. Formada por los ficheros:
 - a) ac_withouthedges.cpp
 - b) ac_withouthedges.hpp
 3. **Impl-ACwithoutEdgesYUV**: clase que implementa el contorno que evolucionará en una imagen a color. Formada por los ficheros:
 - a) ac_withouthedges_yuv.cpp
 - b) ac_withouthedges_yuv.hpp
 4. **Impl-list**: implementación de una lista ligada genérica. Formada por los ficheros:
 - a) linked_list.tpp
 - b) linked_list.hpp
 5. **Impl-Filters**: clase que implementa los filtros que se le pueden aplicar a la imagen antes de realizar la segmentación. Formada por los ficheros:
 - a) filters.cpp
 - b) filters.hpp
 6. **Impl-GeodesicAC**: clase que implementa un contorno geodésico. Formada por los ficheros:
 - a) geodesic_ac.cpp
 - b) geodesic_ac.hpp
 7. **Impl-HausdorffDistance**: clase que implementa la distancia de Hausdorff. Formada por los ficheros:
 - a) hausdorff_distance.cpp
 - b) hausdorff_distance.hpp
 8. **GUI-ImageViewer**: formada por los ficheros:
 - a) imageviewer.cpp
 - b) imageviewer.hpp

9. GUI-PixmapWidget: formada por los ficheros:

- a) pixmapwidget.cpp
- b) pixmapwidget.hpp

Se han resaltado las tres clases que interesarán: *ActiveContour*, *ACwithoutEdges* y *list*. Las demás clases añaden funcionalidades extra que no son necesarias en este proyecto.

6.1.2. Esquema general

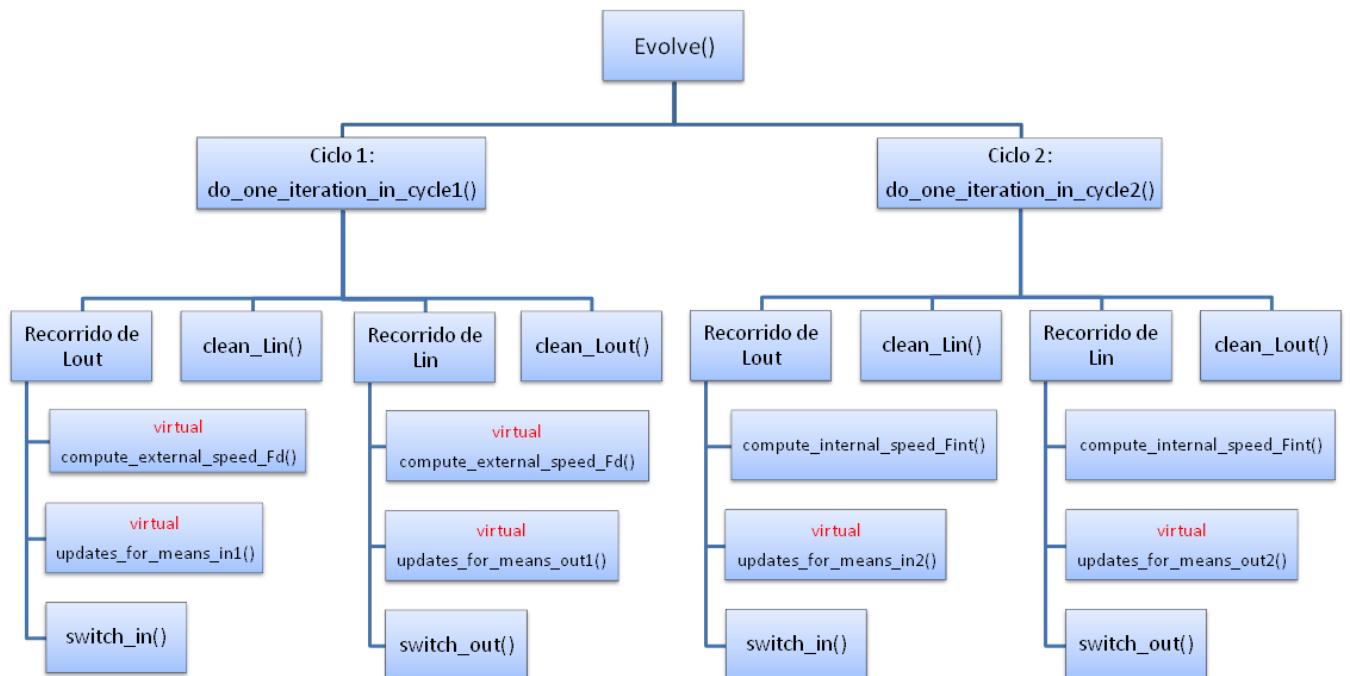


Figura 6.1: Esquema de las funciones utilizadas en el trabajo de Ofeli para realizar el algoritmo de level set aproximado

Como se puede observar la estructura del código es prácticamente igual al algoritmo de aproximación presentado en 5.1 exceptuando que en esta implementación la velocidad de cada punto se calcula a la hora de tratar este, no se calculan todas las velocidades de todos los puntos como se sugiere en el algoritmo 5.1. El cálculo de estas velocidades son las funciones `compute_(internal/external)_speed_(Fd/Fint)` que se pueden ver en el esquema 6.1. Las funciones `update_for_means_(in/out)(1/2)` que se salen del algoritmo

de aproximación presentado en 5.1 se realizan para poder «adaptar» el contorno a la imagen de manera que se pueda evolucionar este independiente-mente de los niveles de gris que se estén utilizando en la imagen. Esto se refiere a que se va haciendo una media de las intensidades de los puntos per-tenecientes a las listas que representan implicitamente el contorno, es decir, L_{out} y L_{in} , de manera que el umbral con el que se va decidiendo la velocidad de cada punto va cambiando en función de los puntos que representan el contorno.

Otra cuestión a comentar es la etiqueta «virtual» que tienen varias fun-
ciones en el esquema 6.1 que significa exactamente, que la función es virtual.
Una función virtual en el lenguaje de programación C++ se utiliza en cuestio-
nes de herencia y polimorfismo, de manera que clases hijas puedan redefinir
funciones definidas como virtuales en la clase padre. Así pués, las funciones
compute_external_speed_Fd y *update_for_means_(in/out)(1/2)* están definidas
con esta cláusula ya que dependen de los datos, en este caso de las imágenes,
y hay varias clases para trabajar con ellas en este trabajo. Si son imágenes a
color se utilizará la clase *ACwithoutEdgesYUV*, mientras que si es una ima-
gen a escala de grises se trabajará con la clase *ACwithoutEdges* como se ha
explicado en el anterior apartado.

Comentadas estas cuestiones se deduce que todas las funciones presenta-
das en el esquema 6.1 están implementadas en *ActiveContour*, la clase padre,
exceptuando aquellas que tienen la etiqueta «virtual» que las implementarán
las clases hijas *ACwithoutEdges* y *ACwithoutEdgesYUV* dependiendo de las
características de la imagen.

Para finalizar, la implementación de la lista ligada se utiliza para repre-
sentar las listas L_{out} y L_{in} con las que estaremos trabajando continuamente
en los dos ciclos.

Capítulo 7

Paralelización de la aproximación del *level set*

Las clases a modificar sobre el trabajo de Ofeli [3] para poder realizar una implementación paralela son: *ActiveContour*, *ACwithoutEdges* y *list*. De manera que se paralelizará el código para realizar la segmentación de imágenes a escala de grises, ya que en un principio el cliente sólo está interesado en este tipo de imágenes. Aún así, una vez realizada la segmentación para este tipo de imágenes sería muy sencillo poder realizarlas para imágenes a color, ya que como se ha visto en el capítulo anterior, la mayoría de las funciones son comunes a la clase *ActiveContour*.

La parallelización se realizará mediante OpenMP[17] por lo que las pruebas se realizarán en máquinas con una arquitectura SMP. Para aprovechar las mejores opciones que esta API(*Application Programming Interface*) nos ofrece para poder sacarle el máximo rendimiento posible al algoritmo se ha consultado un tutorial de *Livermore Computing Center*, uno de los centros computacionales de primera clase del mundo [18].

7.1. Planteamiento de la paralelización



Capítulo 8

Comparación de resultados temporales y rendimiento

Capítulo 9

Funciones extras añadidas en base a la segmentación

Capítulo 10

Gestión del proyecto

10.1. Gestión del alcance

10.2. Gestión del tiempo

10.3. Gestión de los riesgos

El único riesgo más aparente que se puede producir en este proyecto es la pérdida de la memoria escrita o de la implementación paralela realizada a lo largo del proyecto. Para evitar que pueda producirse tal situación se requiere el uso de dos herramientas online: el sistema de almacenamiento de Google, *Drive*, y un sistema de gestión de versiones, en este caso *Github*. Con estas dos herramientas se ha podido realizar un sistema de *backup* sencillo como gestión de este riesgo. La memoria con todos sus documentos se ha guardado diariamente en el *Drive* de manera que en caso de pérdida se pueda recuperar los avances de un día anterior. La implementación se ha ido creando en manera de versiones en las que se le han ido añadiendo funcionalidades extras en cada versión. Cada una de estas versiones se ha ido subiendo a la plataforma *Github* del autor de esta memoria en [19].

10.4. Gestión del costes

Los costes de este proyecto ha sido únicamente la dedicación de horas humanas con las que se ha realizado. Esto se puede ver en la tabla ??.

Capítulo 11

Conclusiones

CONCLUSIONES

Bibliografía

- [1] IEEE Xplore Digital Library. <http://ieeexplore.ieee.org/>.
- [2] Suta L., Bessy F., Veja C., and Vaida M.-F. Active contours: Application to plant recognition. *Intelligent Computer Communication and Processing (ICCP), 2012 IEEE International Conference on*, pages 181 – 187, 2012.
- [3] Fabien Bessy. Open, fast and efficient level set implementation. 2012-2013. <https://code.google.com/p/ofeli/>.
- [4] Terry S. Yoo. *Insight into Images: Principles and Practice for Segmentation, Registration, and Image Analysis*. A K Peters, Ltd, 2004.
- [5] Yu-Jin Zhang. *Advances in Image and Video Segmentation*. 2006.
- [6] Kevin McGuinness. *Image Segmentation, Evaluation, and Applications*. PhD thesis, Dublin City University, 2010. Pages 11-17.
- [7] Basavaprasad B. and Ravi M. A comparative study on classification of image segmentation methods with a focus on graph based techniques. *International Journal of Research in Engineering and Technology*, vol. 3, 2014.
- [8] Arnau Oliver Malagelada. *Automatic Mass Segmentation In Mammographic Images*. PhD thesis, University of Girona, <http://eia.udg.edu/~aoliver/publications/tesi/>, 2008. Chapter 2.3.3.
- [9] Erik Smistad, Thomas L. Falch, Mohammadmehti Bozorgi, Anne C. Elster, and Frank Lindseth. Medical image segmentation on GPUs – A comprehensive review. 2015.
- [10] Kass M., Witkin A., and Terzopoulos D. Snakes: Active contour models. *International Journal of Computer Vision*, pages 321–331, 1988.

- [11] Aaron Hagan and Ye Zhao. Parallel 3d image segmentation of large data sets on a gpu cluster. vol. 5876:960 – 969, 2009.
- [12] Hormuz Mostofi and Keble College. *Fast Level Set Segmentation of Biomedical Images using Graphics Processing Units*. PhD thesis, University of Oxford, 2009. <https://code.google.com/p/cudaseg/>.
- [13] M. Jeon, M. Alexander, and N. Pizzi. Parallel image segmentation with level set methods. 2015.
- [14] Mohammad Reza Hajihashemi and Magda El-Shenawee. Mpi parallelization of the level-set reconstruction algorithm. pages 1 – 4, 2009.
- [15] Yonggang Shi and William Clem Karl. A fast implementation of the level set method without solving partial differential equations. 2005. <http://iss.bu.edu/yshi>.
- [16] T. Chan and L. Vese. Active contours without edges. vol. 10:266–277, Feb 2001.
- [17] OpenMP® API specification for parallel programming. <http://openmp.org/>.
- [18] Livermore Computing Center (LC). <https://computing.llnl.gov>.
- [19] Franco D. Implementacion paralela. 2015. <https://github.com/dfranco007/parallelOfeli>.