

Capítulo 2

LA SEGMENTACIÓN DE IMÁGENES

2.1. ¿QUÉ ES LA SEGMENTACIÓN DE IMÁGENES?

La **segmentación** es la **división o separación de una imagen en regiones para su posterior procesamiento**. Por ejemplo, se podría separar la imagen en diferentes características para clasificarlas según nos interese. A cada una de las regiones en las que se ha dividido (segmentado) la imagen se le denomina *modo*.

En nuestro caso dividiremos la imagen en dos modos: **objeto** y **fondo**, considerando el primero como la región principal que nos interesa y fondo al resto de regiones de la imagen.

2.1.1. TIPOS DE SEGMENTACIÓN

Se pueden clasificar las segmentaciones en tres grandes grupos:

A) Segmentación basada en propiedades de similitud.

Se trata de unir píxeles de una imagen que tengan características comunes. Suelen dar buenos resultados y por ello son las técnicas de segmentación más usadas. Se pueden distinguir las siguientes:

- **Umbralización**. Es la forma más sencilla de segmentar una imagen cuando los distintos modos se caracterizan por niveles de gris muy diferentes. Para ello se emplea el **histograma**, el cual indica el número de puntos de la imagen que poseen un determinado nivel de gris, para separar fondo y objeto por medio de un *umbral*.

En la **Figura 2.1** se muestra un ejemplo de umbralización. La *figura a* es la imagen a umbralizar, en la que se puede distinguir el fondo de color negro y el objeto de diferentes grises. En la *figura b* se muestra el histograma de la imagen anterior, donde se escogerá un umbral intermedio entre los dos lóbulos principales que se han creado. Poniendo un umbral de 0.18 la imagen quedará dividida en dos colores, de manera que como se observa en la imagen umbralizada (*figura c*), todos los niveles de gris se mapean en el blanco y se obtiene una imagen bimodal (objeto y fondo) por haber aplicado un sólo umbral.

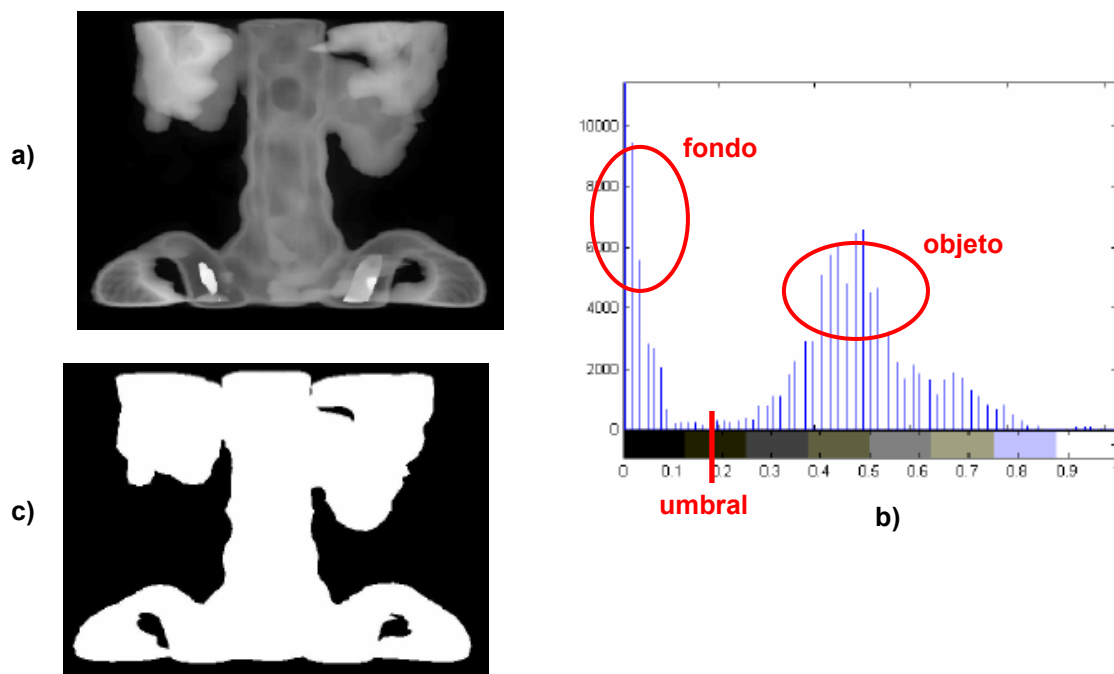


Figura 2.1. a) Imagen original. b) Histograma de la imagen. c) Imagen segmentada con un umbral de 0.18.

Como se ha visto, la umbralización se realiza respecto a niveles de grises que tiene la imagen; para **imágenes a color** existen tres métodos principalmente [w1]:

- Umbralización de tres histogramas 1-D. Se realiza el histograma en niveles de intensidad (gris) para cada *plano de color RGB*.
- Umbralización de histogramas 2-D. Transformación del plano RGB al *plano HSI* (tono - saturación - intensidad), separando en *luminancia* y *crominancia*¹.
- Umbralización del histograma 3-D. Se buscan agrupaciones de píxeles con el mismo valor RGB en un espacio 3-D.

- **Crecimiento de regiones.** En esta técnica se escoge un píxel semilla, o grupo de ellos, y una *condición de similitud* (de color, textura, etc.) para una región, de forma que un píxel será añadido a dicha región sólo si cumple el criterio. La región irá creciendo hasta que no existan píxeles que cumplan el criterio y, por tanto, no se incluyan en la región.

¹ crominancia es la información de color y luminancia es la intensidad de dicho color

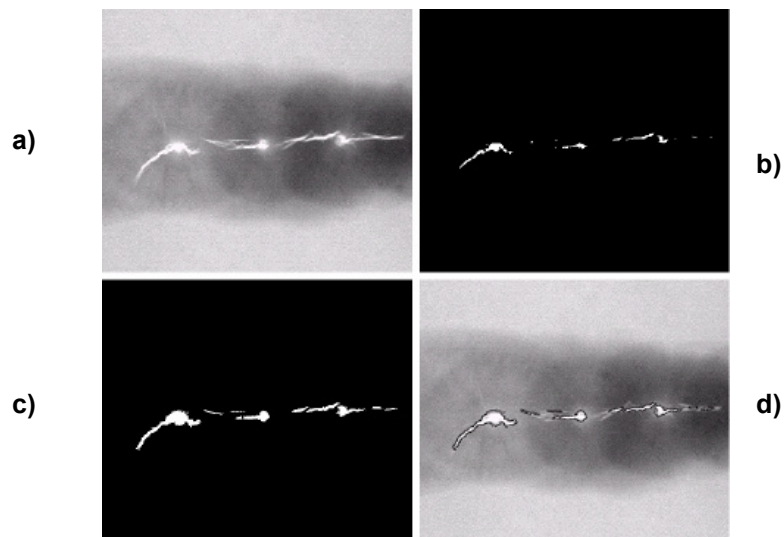


Figura 2.2. Ejemplo de segmentación por crecimiento de regiones. **a)** Imagen que muestra soldaduras defectuosas. **b)** Píxeles semilla. **c)** Resultado del crecimiento de región. **d)** Fronteras de soldaduras defectuosas ya segmentadas (en negro).

El problema de este método de segmentación reside en la dificultad de encontrar el píxel semilla y la condición de similitud adecuados.

Como alternativa para solucionar el problema de la elección de semillas, se encuentra la **División y fusión de regiones** [1], que consiste en dividir la imagen en cuatro regiones que no cumplan el criterio de similitud y unir regiones adyacentes similares entre sí. Cuando no haya más regiones que unir se subdividen otra vez las regiones en cuatro cuadrantes no similares y se mira si es posible unir regiones similares. Este proceso se repite iterativamente hasta que sea imposible subdividir en más cuadrantes no similares o no se puedan unir regiones similares, en cuyo caso la segmentación habrá finalizado.

La limitación de esta alternativa es que las imágenes a segmentar deben ser cuadradas y tener un tamaño potencia de 2 [w1].

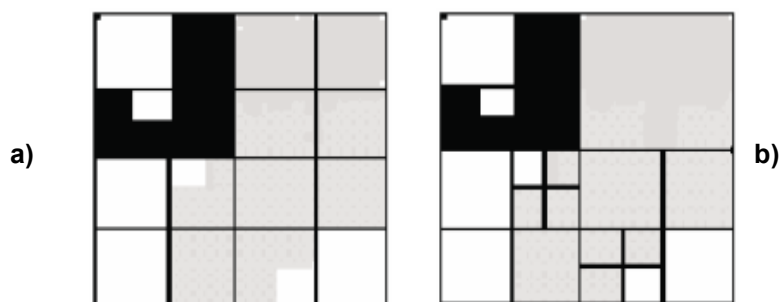


Figura 2.3. Proceso de formación de regiones de una imagen usando el método de división y fusión. **a)** División inicial de la imagen. **b)** Regiones similares resultantes después de aplicar el método en su primera etapa.

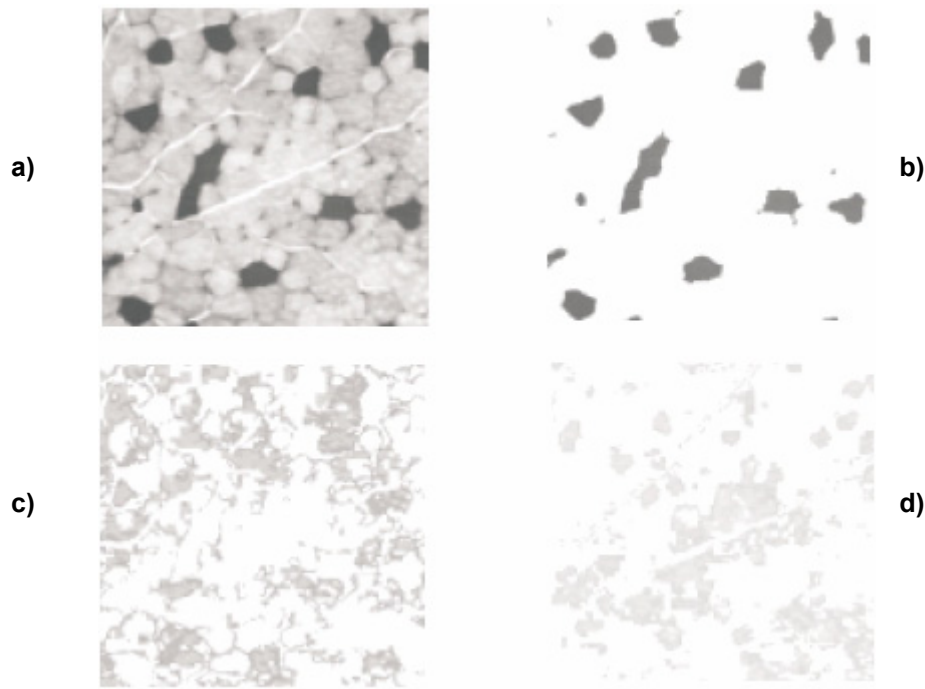


Figura 2.4. Segmentación de imágenes de secciones transversales del músculo EDL de rata tratado con la reacción de la ATPasa miofibrilar preincubada a pH 4.33. **a)** Imagen original. **b)** Imagen que agrupa las fibras más oscuras. **c)** Imagen que agrupa las fibras de tinción intermedia. **d)** Imagen que agrupa las fibras más claras.

B) Segmentación basada en propiedades de discontinuidad.

- **Detección de contornos.** En este caso lo que se busca es localizar los puntos de la imagen donde se produzcan discontinuidades o bordes. Los *bordes* son fronteras entre dos regiones con propiedades de colores diferentes. Existen varias técnicas para la detección de bordes de una imagen:

1. - Métodos basados en las primeras derivadas [w2]. En estos métodos primero se miden los cambios de intensidad en todos los puntos de la imagen, haciendo uso de la primera derivada, y después se seleccionan como puntos de bordes aquellos puntos en los que el cambio de intensidad rebasa algún umbral preestablecido.

Una imagen puede representarse matemáticamente por una función f cuyo valor en el punto (x,y) representa la iluminación existente en ese punto. El cambio de iluminación de un punto en la imagen está representado gráficamente por la pendiente que allí tiene la superficie que representa a f y, normalmente, resulta diferente según sea la dirección considerada. La dirección en la que el cambio es más intenso está dada por el vector *gradiente* en ese punto

$$\nabla f(x,y) = \frac{\partial f(x,y)}{\partial y} \vec{i} + \frac{\partial f(x,y)}{\partial x} \vec{j} \quad (2.1)$$

y la magnitud del campo vendrá dada por el módulo del *gradiente*, por lo que la medición de los cambios de iluminación se puede realizar por medio del módulo del gradiente. Para evitar que estos procesos consuman mucho tiempo de procesamiento, normalmente sólo se hace una aproximación del gradiente $|\nabla f(x,y)| \approx |\nabla i| + |\nabla j|$, donde $|\nabla i|$ y $|\nabla j|$ son las aproximaciones discretas de las derivadas parciales. A continuación se describen cuatro de los principales métodos basados en la primera derivada.

2. - Método basado en la segunda derivada [w2]. Consiste en determinar directamente en qué puntos de la imagen aparecen los picos de los cambios de iluminación utilizando para ello las segundas derivadas. En los puntos en que los cambios de intensidad tienen un máximo se anula la derivada segunda de la intensidad tomada en la dirección en que tiene lugar la transición de claro a oscuro, para esto se puede cambiar el cálculo de las derivadas por la *laplaciana*:

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial y^2} + \frac{\partial^2 f(x,y)}{\partial x^2} \quad (2.2)$$

Este operador tiene la ventaja de que sólo exige calcular dos derivadas segundas. Las siguientes máscaras implementan la laplaciana para usar en imágenes digitales.

0	1	0
1	-4	1
0	1	0

-1	-1	-1
-1	8	-1
-1	-1	-1

Este método identifica como puntos de contorno aquéllos de coordenada (i,j) en los que la laplaciana se anula; aunque puede resultar preferible tomar aquéllos en los que cambie de signo, llamados *cruces por cero*, que localizan la posición exacta de los bordes. Otras veces se suelen tomar los puntos en los que la laplaciana se anula y además el módulo del gradiente pasa cierto umbral.

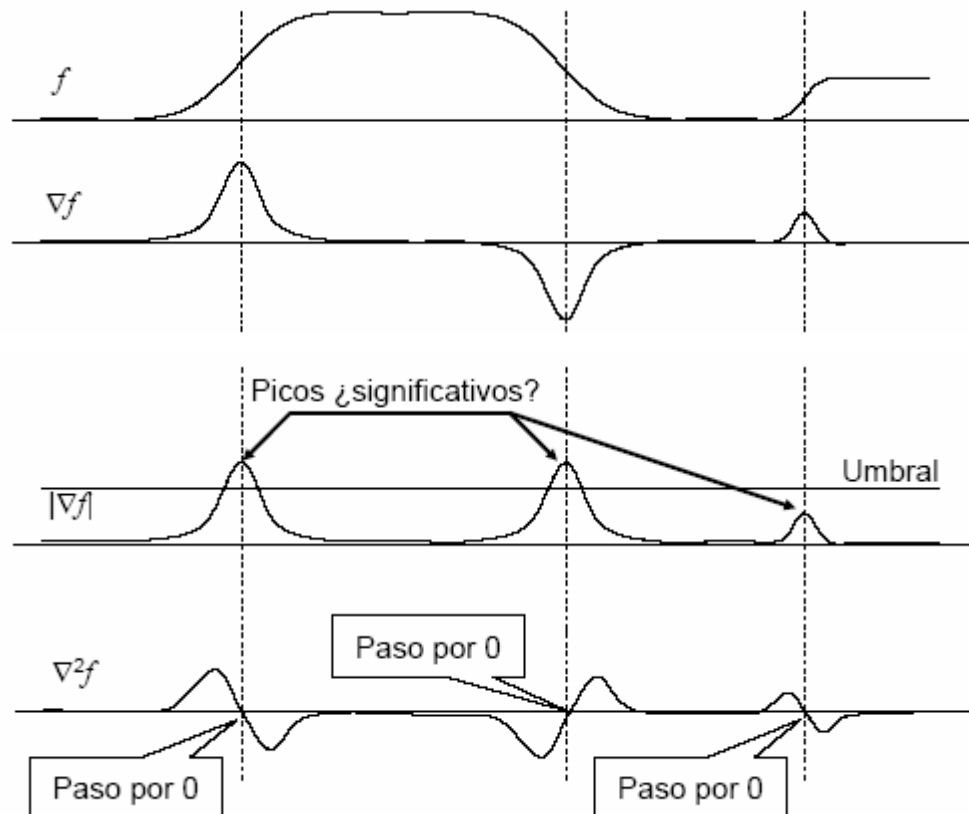


Figura 2.5. Representación gráfica del gradiente, su módulo y la laplaciana de una función.

La ventaja de utilizar la laplaciana como detector de contornos es que los puntos producidos al ser agrupados forman trazos finos, por lo que no se requiere de un posterior suavizado. Entre los inconvenientes se tiene que no ofrece información sobre la dirección y el sentido en que tiene lugar la transición del punto del contorno, ni la magnitud de la misma; además es muy sensible al ruido y por cada borde detecta un doble borde.

3. - Método de Canny. Existe un tercer método, denominado *método de Canny*, que minimiza la tasa de error (no se omiten bordes ni se crean falsos bordes) que también da una muy buena localización de los bordes y cuyo estudio se muestra ampliamente en el Capítulo 3 ya que es el utilizado en este trabajo [w1].

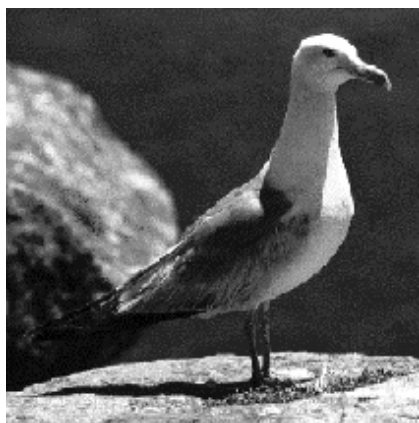
C) Segmentación *region matching*.

Este tipo de segmentación supone un conocimiento previo global de la imagen o de alguna de sus regiones. Si se tiene conocimiento de la estructura a segmentar y esta estructura no cambia mucho normalmente, como ocurre en algunas aplicaciones industriales, se puede buscar directamente la estructura dentro de la imagen. Para ello se crea un patrón de la estructura a partir de una serie de imágenes de ésta, se compara el patrón con la imagen, y las zonas que superen el *umbral de similitud* de la estructura serán las regiones a segmentar.

La forma más sencilla para ver la similitud entre dos imágenes es la *correlación*. La correlación consiste en convolucionar el patrón con la imagen, obteniéndose valores elevados allí donde el patrón sea similar a una región, por tanto, la región buscada aparecerá con el mayor nivel de intensidad de la imagen resultante.

$$\underbrace{g(x)}_{\text{resultado}} = \underbrace{f(x)}_{\text{imagen}} * \underbrace{h(x)}_{\text{patrón}} = \sum_k f(k) h(x+k) \quad (2.3)$$

Por ejemplo, para detectar el ojo de la gaviota de la imagen siguiente, se convoluciona el patrón (ojo de la gaviota) con la imagen completa. El resultado se observa que no llega a ser correcto ya que el proceso tiende a emborronar la imagen (filtrado paso-baja). La solución es realizar una correlación normalizada, dividiendo el resultado de la correlación por la intensidad total de los píxeles de la ventana. El punto más brillante de la imagen corresponde al ojo, el patrón que buscábamos.



a)



b)



c)



d)

Figura 2.6. a) Imagen original. b) Patrón del ojo que buscamos. c) Imagen correlada con efecto negativo paso-bajo. d) Imagen con correlación normalizada, donde se observa que el patrón es la zona más brillante.

En este trabajo la segmentación se realiza por **detección de contornos**, concretamente con el *método de Canny*, ya que lo que se pretende es separar el contorno de la quemadura de una imagen médica del resto de ésta para facilitar su posterior análisis y tratamiento.

2.2. DIFICULTADES EN LA SEGMENTACIÓN DE IMÁGENES

Aunque a simple vista la detección de regiones pueda parecer una tarea sencilla, realmente existen una serie de dificultades a la hora de realizar la segmentación de una imagen digitalmente.

- **Artefactos.** Por problemas en la etapa de adquisición de la imagen se puede encontrar ruido en la misma que distorsione las características de las distintas regiones.
- **Volumen parcial.** Al ser las imágenes representaciones 2D de objetos 3D se puede encontrar que la imagen información no es propiamente de ese plano y limite la definición de las regiones.
- **Inhomogeneidad.** Los objetos no poseen siempre una iluminación uniforme, sino que muchas veces es sólo una textura característica lo que los diferencia de otros.
- **Forma.** Hay que tener en cuenta la forma del objeto que se pretende segmentar, ya que a veces esta tarea puede ser muy complicada si tiene formas delgadas y/o muy ramificadas, de manera que se dificulte la detección de sus bordes o la diferenciación respecto a los bordes de objetos adyacentes.
- **Ruido.** La imagen puede ser de baja calidad debido a la presencia de ruido o a una mala adquisición de la misma.
- **Conocimiento previo.** Muchas veces se emplea en la segmentación visual el conocimiento previo que se tiene sobre una determinada región. Este conocimiento previo no es sencillo de obtener ni de implementar en estructuras aplicables a algoritmos.

2.3. FORMAS DE SEGMENTACIÓN

Las formas en que se puede segmentar una imagen vienen determinadas por los diferentes grados de intervención que pueda tener el usuario. Es un aspecto muy importante ya que la calidad de la segmentación de la imagen varía de una a otra forma.

2.3.1. SEGMENTACIÓN MANUAL

El usuario realiza la segmentación él mismo con la ayuda de una herramienta informática. Consiste en seleccionar manualmente las fronteras de las regiones que se desea segmentar, bien mediante el marcado de puntos de las mismas bien usando algunas herramientas de apoyo más complejas. Es un método muy preciso pero es muy lento y se hace impracticable cuando necesitamos segmentar un número alto de imágenes. El usuario puede marcar las regiones libremente o elegir formas predefinidas (rectángulos, círculos, etc.) que sitúa sobre la región que desea segmentar. El programa presentará automáticamente los parámetros extraídos de la región segmentada.

Es obvio que su realización consume mucho tiempo y cuando se trata un número importante de imágenes (como en una *reconstrucción*) es inaceptable por parte del usuario, de manera que se hace necesaria la automatización en la segmentación.

2.3.2. SEGMENTACIÓN AUTOMÁTICA

El ordenador realiza todo el proceso de forma automática, por tanto, no existe intervención alguna por parte del usuario. Al no consumirse *tiempo de meditación*² la segmentación tarda menos en realizarse que en el caso anterior. El inconveniente es que esta forma, completamente automática, es todavía un problema sin resolver en la mayoría de las aplicaciones de imágenes médicas debido a la amplia variedad de modelos de imágenes, propiedades de objetos y variabilidad biológica.

2.3.3. SEGMENTACIÓN SEMI-AUTOMÁTICA

Viendo las características de las formas de segmentación anteriores parece lógico buscar una alternativa que fusione sus mejores cualidades. Esta es la **segmentación semiautomática o interactiva [w3]**, en la que el ordenador realiza el proceso automáticamente, pero el usuario es capaz de seguir dicho proceso en todo momento e intervenir en determinados instantes sobre el mismo para definir parámetros o corregir resultados. Es la forma más empleada generalmente y existen varios métodos basados en ella, como el **método de los modelos de contorno activo de dos dimensiones**, también conocido como '**snakes**', introducido por M. Kass, A. Witkin y D. Terzopoulos, que ha adquirido una considerable atención durante los últimos años.

Estos algoritmos tratan de minimizar una funcional de energía basada en la deformación del contorno y en las fuerzas de imagen externas

$$E(v) = \int_{\Omega} \underbrace{\left(w_1 |v'(s)|^2 + w_2 |v''(s)|^2 \right)}_{\text{Energía interna}} - \underbrace{|\nabla I(v)|^2}_{\text{Energía externa}} ds \quad (2.4)$$

Sin embargo, para su inicialización se requiere una aproximación poligonal del contorno deseado del objeto, normalmente dibujada a mano por el usuario, o generada a partir de conocimientos a priori. Otra desventaja es que los contornos activos son sensibles a la configuración de sus numerosos parámetros y a la calidad del contorno inicial, lo cual limita su aplicación en imágenes médicas. Además el control por parte del usuario no es aplicable durante la extracción y, a menudo, el método queda atrapado en un mínimo local de energía.

² se denomina tiempo de meditación al que se toma el usuario en su intervención en la segmentación

Como consecuencia se ha seguido otra línea en la detección de fronteras de imágenes interactivas: el **método Live Wire**, también conocido como '**intelligent scissors**'. Con este método, basado en el *algoritmo de Dijkstra*, el proceso de segmentación es directamente guiado por el usuario, quien tiene control inmediato sobre los contornos de objetos. Estos contornos son hallados como caminos minimizados con respecto a una función de coste similar a la función de energía externa de los contornos activos descritos con anterioridad.

2.4. APLICACIONES

La segmentación es un paso imprescindible en diversos procesos de tratamiento de imagen. Entre otros, es necesaria para:

- La determinación de superficies y volúmenes para tomar medidas sobre una región
- Realizar reconstrucciones tridimensionales de una zona de la imagen
- Clasificación o diagnóstico automático, por ejemplo, detección de defectos en materiales, tejidos anómalos, etc.
- Reducir la información de las imágenes. Si de una serie de imágenes para un determinado estudio sólo interesa una región concreta, se puede segmentar y almacenar exclusivamente ésta para su análisis posterior.

2.5. RESUMEN

- La **segmentación** es la división o separación de una imagen en regiones para su posterior procesado.
- El tipo de segmentación utilizada está **basada en propiedades de discontinuidad** dentro de la imagen.
- Dentro de este tipo de segmentación, el detector de contorno usado en el trabajo es el diseñado por el **método de Canny**.
- La segmentación es **semi-automática**, es decir, el usuario interviene en el proceso en determinados instantes para guiar al programa que será el que realice toda la parte computacional.
- El método semi-automático seleccionado es el **Live Wire** o **intelligent scissors**.
- La aplicación principal para la que se ha implementado el algoritmo es la **segmentación de imágenes médicas**.