

3. Metodología

3.1 Preprocesado de la imagen

La segmentación es la separación de imágenes digitales en múltiples regiones que son homogéneas, en relación a uno o más atributos. Por ejemplo, si estamos interesados en una única región, separaremos la imagen en objeto y fondo, considerando como fondo a todas las regiones que no nos interesan.

Aunque en principio podría parecer un proceso sencillo, a la hora de realizar la segmentación nos encontramos con una serie de dificultades, que enumeraremos a continuación:

- Debido a los artefactos de captura tenemos problemas en la adquisición de algunas imágenes de Tomografía Axial Computerizada (TAC), pues nos encontramos ruido que distorsiona sus características en distintas regiones.
- Al ser las imágenes analizadas proyecciones en dos dimensiones (2D) de imágenes en tres dimensiones (3D), podemos encontrar información que no pertenece propiamente a ese plano y que dificulta la detección de regiones.
- Los objetos pueden tener distintos tipos de forma. Puede existir objetos con una forma más o menos regular y convexa, otros objetos con bordes no definidos y ciertas ramificaciones, otros objetos finos y muy ramificados. Estos últimos pueden ser difíciles de segmentar al poder estar entrelazados con otros objetos.
- La imagen algunas veces puede ser de mala calidad debido a la presencia de ruido o a la mala adquisición de la misma.

La imagen procedente directamente del dispositivo de captura no es adecuada para el procesamiento porque puede tener problemas de iluminación, ruido o distorsión. Para solucionar estos problemas realizamos un paso previo a la segmentación consistente en técnicas de preprocesado.

En la fase de preprocesado no sólo se intenta solucionar estos problemas, sino que se prepara la imagen para que el análisis posterior sea mucho más eficiente.

Las técnicas de preprocesado que implementaremos en este proyecto serán la **normalización** y el **aumento de contraste**.

La **Normalización** de la imagen adapta el rango de variación de grises entre crestas y valles de la imagen a un rango deseado para facilitar el procesamiento de las siguientes etapas. El factor de normalización se calcula teniendo en cuenta la media y la varianza de luminancia en la imagen. La normalización se hace necesaria, para tener una cierta independencia de las propiedades de la imagen, como lo son el brillo, el contraste y así poder comparar por su índice de calidad (QI).

Esta técnica permite normalizar la imagen de tal forma que el valor de gris más pequeño es llevado a cero y el máximo valor a uno, con ello conseguimos que la imagen tratada sea independiente del instrumento de captura. Las ecuaciones que nos permiten la normalización de la imagen son las siguientes:

$$1) \quad I(x,y) = I(x,y) - \min(I(x,y))$$

$$2) \quad \max = \max(I(x,y))$$

$$3) \quad I(x,y) = \frac{I(x,y)}{\max}$$

Siendo:

- $I(x,y)$ del apartado 1): nivel de gris de la imagen de entrada.
- $\min(I(x,y)), \max(I(x,y))$: mínimo y máximo nivel de gris de la imagen respectivamente.
- $I(x,y)$ del apartado 3): corresponde con el nivel de gris de la imagen normalizada en la coordenada (x,y) .

El **contraste** se define como la tasa de cambio de la luminancia relativa de los elementos de la imagen de una reproducción. Otra definición es la que lo define como la capacidad de distinguir entre un patrón de barras y un fondo uniforme, de tal forma que un sistema más sensible necesita menor diferencia de contraste o luminancia, mientras que un sistema menos sensible necesitará de un mayor contraste para ser capaz de reconocer dicho patrón de barras. La luminancia se mantiene constante para todos los patrones. En definitiva el contraste es la mínima diferencia de luminancia que existe entre dos puntos de una imagen.

Tras realizar un aumento de contraste a las imágenes TAC mejoramos su apariencia. En ciertas ocasiones una imagen distorsionada necesita una modificación del contraste para descubrir elementos implícitos en la imagen, que a simple vista no pueden ser detectados.

Frecuentemente, los niveles de gris de una imagen proveniente de una Tomografía Axial Computerizada no ocupan todo el rango posible (generalmente 256 niveles), sino sólo una pequeña parte de los mismos. Esto produce en la imagen un efecto de falta de contraste, debido a la cercanía de las distintas intensidades. El ojo humano no puede diferenciar intensidades muy próximas en un rango de 256 niveles, ya que en general se admite que el ojo humano distingue menos de 64 niveles de intensidad en una misma imagen.

La operación de expandir el histograma de valores de intensidad, de tal manera, que sin modificar su forma, se logre que ocupe todo el rango disponible en el equipo de visualización (monitor o impresora) produce un efecto de mejora o aumento del contraste.

Las tareas de preprocesado realizadas en un principio se pensaron que realizarían la normalización de forma automática y el aumento de contraste mediante un umbral fijado de resultados experimentales.

Tras realizar un estudio de los histogramas de la imagen nos dimos cuenta que en la mayoría de los casos, el rango de 0 hasta 0.4 de intensidades se corresponden con objetos pertenecientes a fondo, mientras que el rango de intensidades de interés se encuentra entre 0.4 a 1. El rango de intensidades óptimo que podemos aplicar para realizar el ajuste era de $[0.4;1]$. Es decir que el rango de intensidades de $[0.4;1]$ de la imagen a segmentar fuera extendido a $[0;1]$. Veamos un ejemplo:

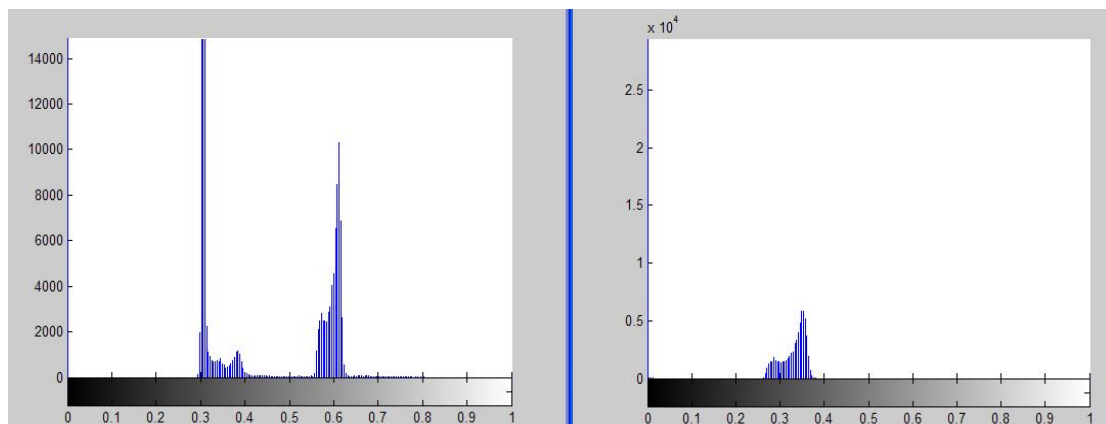


Figura 1: (a) Histograma original, (b) Histograma tras realizar el aumento de contraste comprendido en el rango $[0.4; 1]$

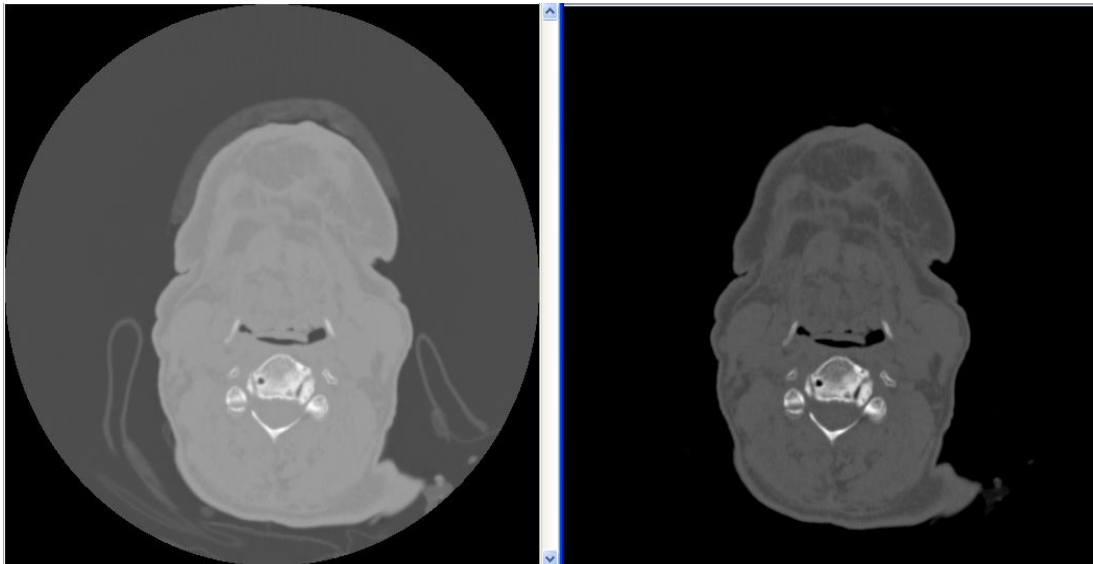


Figura 2: Aumento de contraste con rango fijo de [0.4; 1].

Como podemos observar tras eliminar los píxeles con un nivel de gris inferior a 0.4 y su posterior expansión de aquellos comprendidos entre 0.4 y 1 (figura 1(b)), observamos el resultado en la figura 2, en ella podemos apreciar que el fondo ha desaparecido y que se ha aumentado los contrastes entre los distintos tejidos de interés.

Sin embargo nos dimos cuenta que existe un porcentaje del 8% de las imágenes cuyos niveles de gris correspondientes a zonas de interés se encontraban fuera de este rango. Veamos un ejemplo de este tipo de casos:

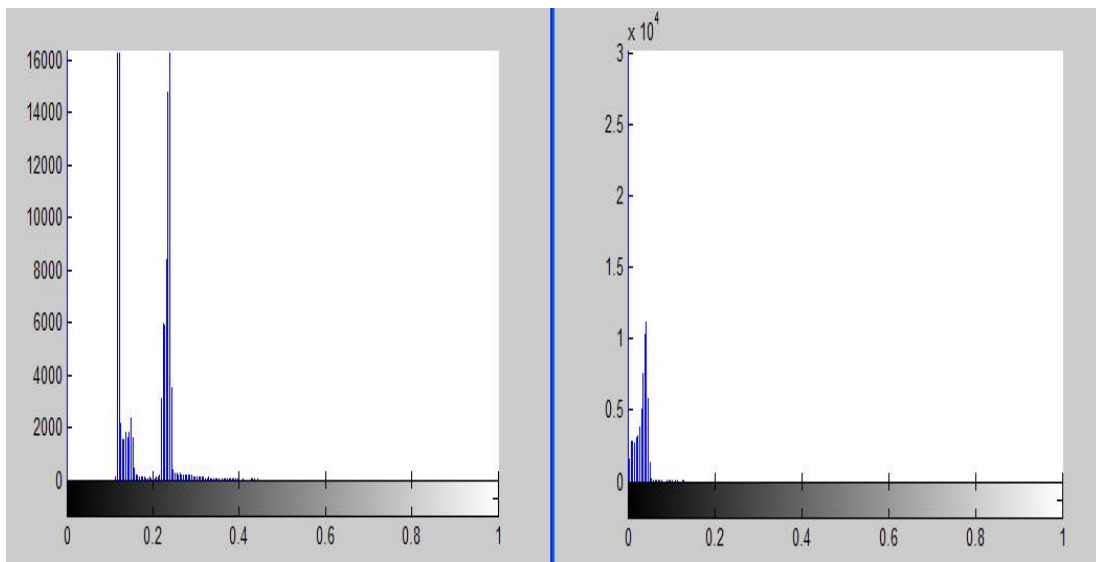


Figura 3: (a) Histograma original, (b) Histograma tras realizar el aumento de contraste comprendido en el rango [0.4; 1]

Si hacemos el aumento de contraste en el rango de intensidades fijados ([0.4;1]) la imagen se nos pierde, como podemos observar en el siguiente caso:

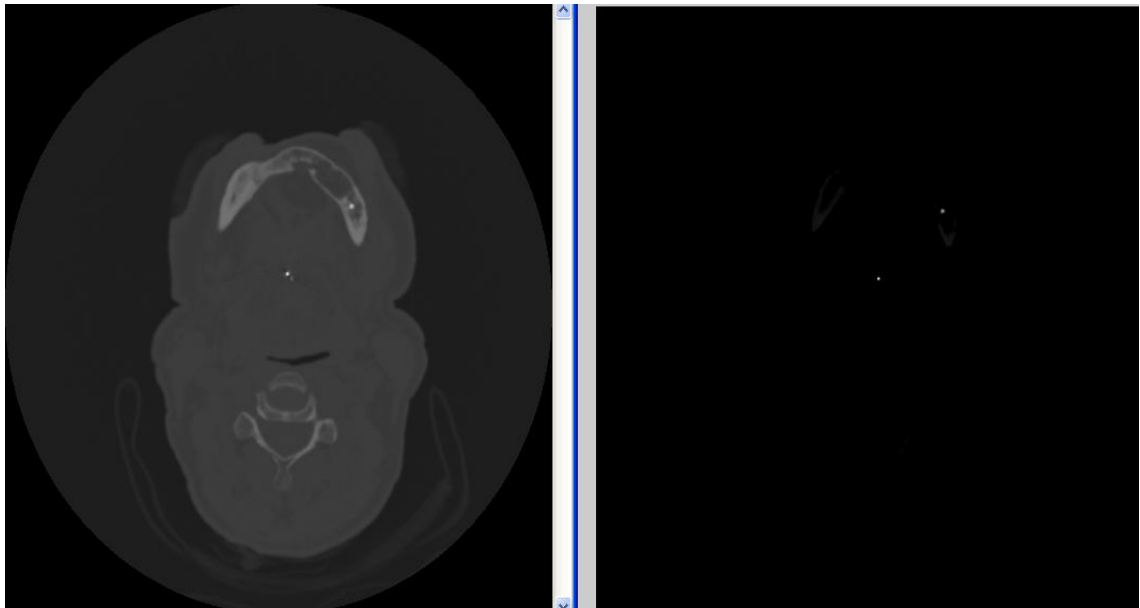


Figura 4: Resultado tras un aumento de contraste fijo inadecuado

Para solucionar este problema diseñamos un programa en Matlab que nos realizó un ajuste automático de los histogramas independiente de cada imagen y que nos proporcionó un aumento de contraste óptimo para las imágenes tomográficas.

Las claves de este aumento de contraste automático residirá en el estudio de los histogramas de la imagen, en los cuales tras analizarlos nos dimos cuenta que excluyendo el valor máximo del valor de nivel de gris existente en cero, el siguiente valor máximo que aparece se corresponde con los niveles de gris pertenecientes a los objetos de fondo, que se encuentran fuera del rango de interés.

Una vez obtenido este conocimiento aplicaremos este aumento dentro del rango de validez, a partir de dicho valor máximo. Esta mejora nos permite visualizar con más nitidez y claridad los distintos tejidos que componen la imagen. El resultado podemos contemplarlo en las siguientes figuras:

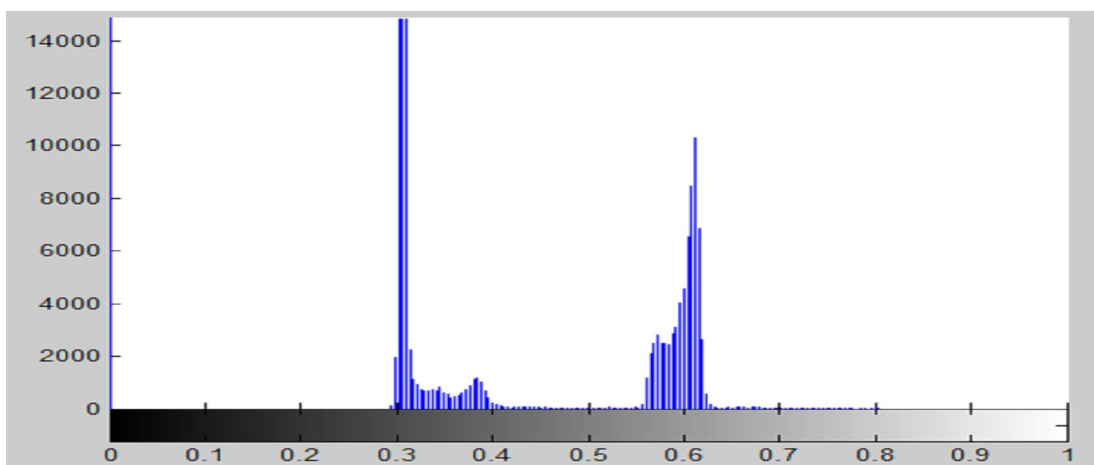


Figura 5: Histograma original de la imagen

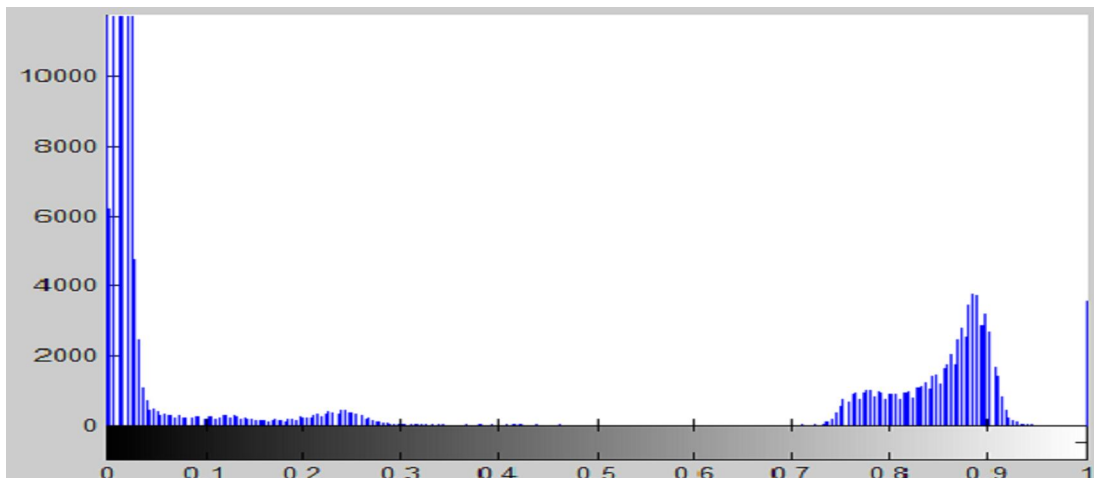


Figura 6: Expansión del histograma. Aumento del contraste

La diferencia entre el aumento del contraste fijo (figura 7) y el automático (figura 8) que hemos realizados en las imágenes procedentes de TAC se pueden visualizar en las siguientes figuras:

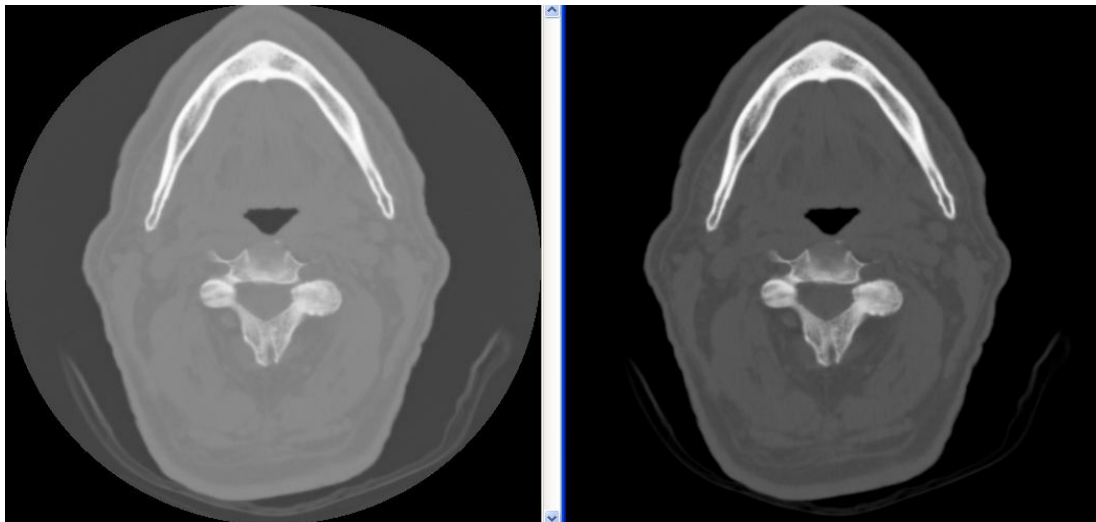


Figura 7: Imagen normalizada con la aplicación de aumento de contraste fijo

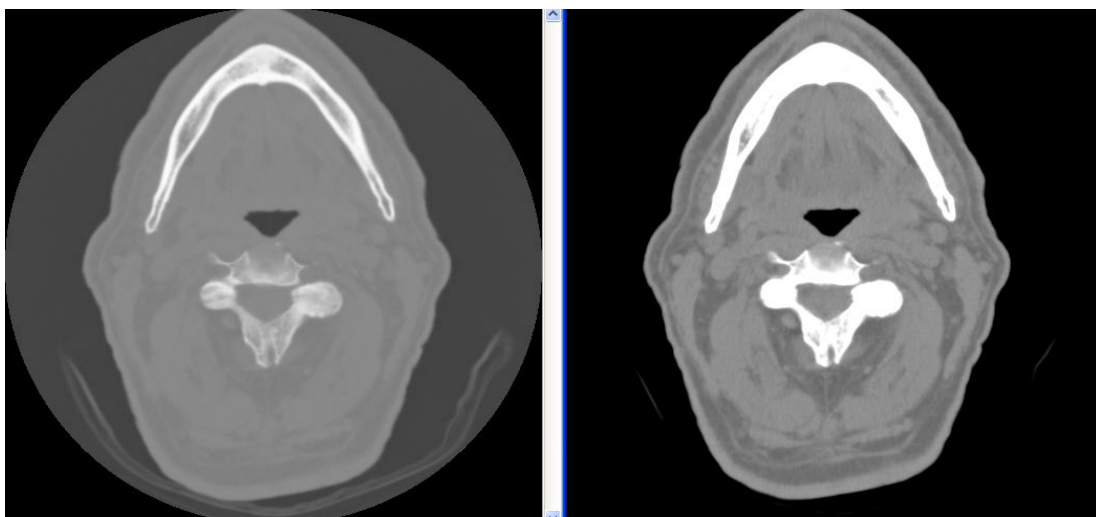


Figura 8: Imagen normalizada con la aplicación automatizada del contraste

3.2 Segmentación

Una vez realizado el preprocesado, la imagen hay que separarla en objetos compuestos por masas de píxeles conectados formando regiones homogéneas que comparten alguna característica común. Para ello se emplean técnicas de segmentación, que permiten agrupar los píxeles en torno a ciertos colores característicos presentes en la imagen.

La segmentación es un proceso común en la mayoría de aplicaciones de visión y proceso de imagen. Su importancia es tal, que muchas de las etapas posteriores de la aplicación dependen fuertemente de la calidad de la segmentación producida.

Mediante la segmentación vamos a dividir la imagen en las partes u objetos que la forman. El nivel al que se realiza esta subdivisión depende de la aplicación en particular, es decir, la segmentación terminará cuando se hayan detectado todos los objetos de interés para la aplicación. En general, la segmentación semiautomática es una de las tareas más complicadas dentro del procesado de imagen. La segmentación va a dar lugar en última instancia al éxito o fallo del proceso de análisis. En la mayor parte de los casos, una buena segmentación dará lugar a una solución correcta, por lo que, se debe poner todo el esfuerzo posible en dicha etapa.

Los algoritmos de segmentación de imágenes monocromáticas generalmente se basan en dos propiedades básicas de los niveles de gris de la imagen: discontinuidad y similitud. Dentro de la primera categoría se intenta dividir la imagen basándonos en los cambios bruscos en el nivel de gris. Las áreas de interés en esta categoría son la detección de puntos, de líneas y de bordes en la imagen. Las áreas dentro de la segunda categoría están basadas en las técnicas de umbrales, crecimiento de regiones, y técnicas de división y fusión.



Figura 9: Esquema general de Segmentación

3.2.1 Segmentación basada en características de discontinuidad

El método más común de buscar discontinuidades es la correlación de la imagen con una máscara. En la figura 10 se puede ver un caso general de máscara de 3 x 3. En este procedimiento se realiza el producto de los elementos de la máscara por el valor de gris correspondiente a los píxeles de la imagen encerrados por la máscara. La respuesta a la máscara de cualquier píxel de la imagen viene dado por:

$$R = \sum_{i=1}^9 w_i z_i$$

Donde z_i es el nivel de gris asociado al píxel de la imagen con coeficiente de la máscara w_i . Como suele ser habitual, la respuesta de la máscara viene referida a su posición central.

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

Figura 10: Mascara 3x3

Cuando la máscara esté centrada en un píxel de borde de la imagen, la respuesta se determina empleando el vecindario parcial apropiado.

Las técnicas de detección de discontinuidades idealmente, proporcionarán los píxeles correspondientes a los contornos o fronteras entre las regiones de la imagen. En la práctica este conjunto de píxeles no suele caracterizar completamente a esos contornos debido a la presencia de ruido, ruptura en los propios contornos debido a la iluminación no uniforme y otros efectos que introducen espúreos en las discontinuidades de la intensidad. En general, después de los procedimientos de detección de bordes se suele emplear técnicas de enlazado u otras técnicas de detección de contornos designadas para unir los píxeles de bordes en contornos significativos.

En este proyecto nos centraremos en el segundo bloque anteriormente comentado, en la segmentación basada en características de similitud.

3.2.2 Segmentación basada en características de similitud

Las áreas dentro de esta categoría están basadas en las técnicas de umbrales y crecimiento de regiones. En los siguientes apartados explicaremos cada uno de estos conceptos.

3.2.2.1 Segmentación por umbrales

Muchas imágenes se pueden caracterizar como un conjunto de objetos luminosos sobre un fondo de determinada oscuridad, de tal forma que los píxeles del objeto y el fondo tienen los niveles de gris agrupados en dos modos dominantes (ejemplos típicos pueden ser texto manuscritos y escritos a máquina, muestras biomédicas microscópicas, rutas de aeroplanos ...) Un modo de extraer los objetos del fondo es elegir un determinado umbral T que aisle dichos modos.

Entonces cualquier punto (x,y) para el cual $f(x,y) > T$ es llamado un punto objeto en caso contrario, el punto es llamado un punto del fondo.

La misma aproximación básica clasifica un punto (x,y) como perteneciente a un tipo de objeto si $T_1 < f(x,y) \leq T_2$, a la otra clase de objeto si $f(x,y) > T_2$, y al fondo si $f(x,y) \leq T_1$, a este tipo de umbralización se le conoce como Umbralización multinivel, y suele ser menos segura que la Umbralización binaria, a la cual también se le llama Binarización. El motivo es la dificultad de establecer umbrales múltiples que logren aislar realmente las regiones de interés, sobre todos cuando el número de modos de los histogramas es alto.

El histograma de una imagen es ampliamente utilizado como herramienta tanto cualitativa como cuantitativa. Este corresponde a un gráfico de la distribución de valores de intensidad de los píxeles de una imagen (niveles de gris) o de una porción de la misma.

Podemos denotar como $h(i)$, el número de píxeles que dentro de la región de interés tiene el valor de intensidad i , donde $i = 0, 1, 2, \dots, L-1$ es el número posible de niveles de gris para la imagen. Los valores $h(i)$, corresponderán entonces a los valores del histograma.

El gráfico del histograma es bidimensional y en él se representa $h(i)$ en función de i . Tal gráfico, puede proporcionar importante información acerca del brillo y contraste de una imagen así como

de su rango dinámico. En las siguientes figuras se muestran los dibujos de dos histogramas típicos.

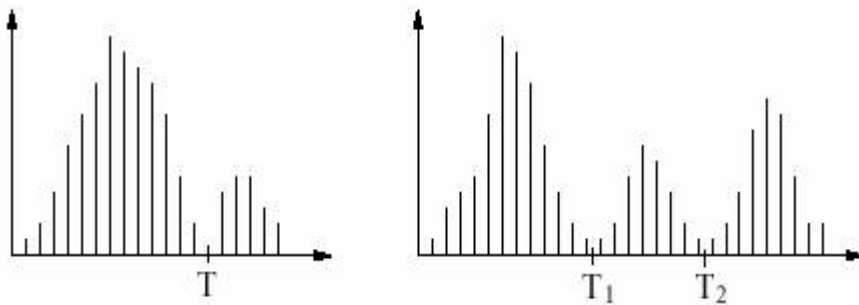


Figura 11: Histogramas de nivel de gris que se pueden dividir en: (a) un umbral sencillo y (b) umbrales múltiples

En el eje de abscisa se representan los niveles de gris, y en el eje de ordenadas se representa $h(i)$, es decir, el número de ocurrencias del nivel de gris.

Se puede decir que la iluminación de la imagen, desempeña un papel determinante en el establecimiento de la forma del histograma.

3.2.2.1.1 Umbralización binaria

Para realizar la binarización, basta con explorar todos los puntos de la imagen y determinar qué puntos corresponderían al objeto y cuáles al fondo.

Así que el éxito de éste método depende totalmente de cómo se divida el histograma.

La técnica más simple para realizar esto es dividir el histograma de la imagen usando un umbral simple T . La imagen particionada $g(x, y)$ está definida como:

$$g(x,y) = \begin{cases} 1 & \text{si } f(x,y) > T \\ 0 & \text{si } f(x,y) \leq T \end{cases}$$

Entonces los píxeles etiquetados con 1 (o cualquier otro nivel de intensidad conveniente) corresponden a los objetos, mientras que los píxeles etiquetados con 0 corresponden al fondo.

La selección del valor del umbral, se realiza generalmente a partir del histograma de la imagen. Así si una imagen está compuesta de un objeto que aparece en la escena sobre un fondo, entonces es de esperar que el histograma sea bimodal, es decir, si por ejemplo el objeto es más claro que el fondo, pues en el histograma aparecerán dos picos, el ubicado en los valores de gris más elevados correspondiente al fondo y otro pico para niveles de gris más bajos, correspondientes al objeto.

En la siguiente figura se muestra un histograma bimodal, en el cual el umbral se ubica entre los dos picos del histograma:

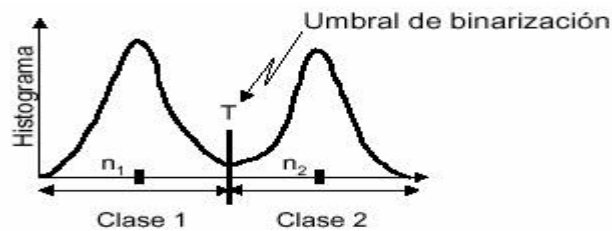


Figura 12: Histograma Bimodal

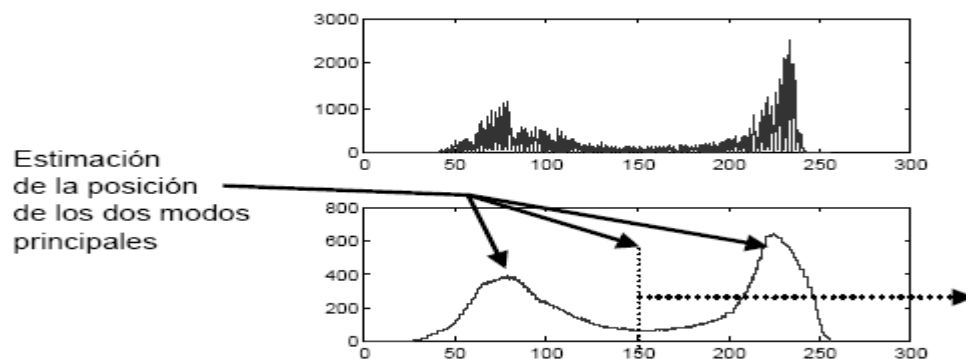


Figura 13: Estimación de posición de modos principales

La selección automática del umbral, es un problema difícil, debido a que el histograma no siempre es bimodal, en cuyo caso resulta necesario combinar la información espacial presente en la imagen, con la información referente al nivel de gris. A continuación podemos observar un ejemplo:

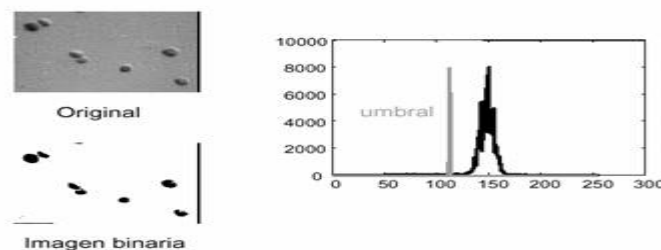


Figura 14: Histograma no modal

La determinación del umbral es clave para conseguir resultados útiles, puesto que una mala elección del umbral nos podría llevar a errores.

Supongamos que una imagen contiene sólo dos regiones de brillo principales. El histograma de una imagen de este tipo se puede considerar como una estimación de la función densidad de probabilidad de brillo $p(z)$.

Esta función de densidad total es la suma o mezcla de dos densidades unimodales, una para las regiones claras de la imagen y la otra para las oscuras. Además los parámetros de la mezcla son proporcionales a las superficies de cada región de brillo. Si se conoce o se supone la forma de las densidades, se puede determinar un umbral óptimo (en términos del error mínimo) por segmentación en las dos regiones de brillo de la imagen.

Supongamos que una imagen contiene dos valores combinados con ruido aditivo Gaussiano. La función densidad de probabilidad de la mezcla es:

$$p(z) = P_1 p_1(z) + P_2 p_2(z)$$

Para el caso Gaussiano, es:

$$p(z) = \frac{P_1}{\sqrt{2\pi} \sigma_1} \exp \left[-\frac{(z - \mu_1)^2}{2\sigma_1^2} \right] + \frac{P_2}{\sqrt{2\pi} \sigma_2} \exp \left[-\frac{(z - \mu_2)^2}{2\sigma_2^2} \right]$$

Siendo μ_1 y μ_2 los valores medios de los dos niveles de brillo, σ_1 y σ_2 son las desviaciones estándar respecto a la media y P_1 y P_2 son las probabilidades a priori de los dos niveles. Se debe satisfacer la restricción:

$$P_1 + P_2 = 1$$

Por lo que la densidad de la mezcla tiene cinco parámetros desconocidos. Si se conocen todos los parámetros, se puede determinar con facilidad el umbral óptimo.

Supongamos que las regiones oscuras corresponden al fondo y que las claras corresponden a los objetos. En este caso μ_1 menor que μ_2 y se puede definir un umbral T de forma que todos los píxeles con nivel de gris superior a T se consideren puntos objeto.

La probabilidad de clasificar (erróneamente) un punto objeto como un punto del fondo es:

$$E_1(T) = \int_{-\infty}^T p_2(z) dz$$

De forma similar, la probabilidad de clasificar un punto de fondo como un punto objeto es:

$$E_2(T) = \int_T^{\infty} p_1(z) dz$$

Así la probabilidad de error total es:

$$E(T) = P_2 E_1(T) + P_1 E_2(T)$$

Derivando $E(T)$ con respecto a T e igualando a 0, encontraremos el valor umbral para el que el error es mínimo. Procediendo así, tenemos:

$$P_1 p_1(T) = P_2 p_2(T)$$

Aplicando este resultado a la densidad Gaussiana, tomando logaritmos y simplificando obtenemos una ecuación cuadrática

$$AT^2 + BT + C = 0$$

Donde:

$$A = \sigma_1^2 - \sigma_2^2$$

$$B = 2(\mu_1 \sigma_2^2 - \mu_2 \sigma_1^2)$$

$$C = \sigma_1^2 \mu_2^2 - \sigma_2^2 \mu_1^2 + 2\sigma_1^2 \sigma_2^2 \ln(\sigma_2 P_1 / \sigma_1 P_2)$$

La posibilidad de tener dos soluciones indica que se pueden necesitar dos valores de umbral para obtener la solución óptima.

Si las varianzas son iguales $\sigma_1^2 = \sigma_2^2$, es suficiente con un umbral único

$$T = \frac{\mu_1 + \mu_2}{2} + \frac{\sigma^2}{\mu_1 - \mu_2} \ln \left(\frac{P_2}{P_1} \right)$$

Si las probabilidades anteriores son iguales, $P_1 = P_2$, como tanto si $s = 0$, el umbral óptimo es el promedio de las medias.

La determinación del umbral óptimo se puede realizar de forma similar para otras densidades de un solo modo de la forma conocida, tales como la de Raleigh y la normal.

Se puede utilizar un método de error medio por mínimos cuadrados para estimar los parámetros de una imagen a partir de un histograma. Por ejemplo, el error cuadrático medio entre la densidad de la mezcla $p(z)$ y el histograma experimental $h(z_i)$ de n puntos es:

$$e_{ms} = \frac{1}{n} \sum_{i=1}^n [p(z_i) - h(z_i)]^2$$

En general, la determinación analítica de los parámetros que minimizan este error cuadrático medio no es un tema sencillo. No obstante, para el caso gaussiano, el cálculo directo de igualar las derivadas parciales a 0 conduce a un conjunto de ecuaciones simultáneas trascendentes que normalmente sólo se pueden resolver por procedimientos numéricos. Dado que el gradiente se puede calcular fácilmente, es posible usar para minimizar e_{ms} un gradiente conjugado o un método de Newton para ecuaciones no lineales simultáneas.

Con cualquiera de estos métodos iterativos, se pueden especificar los valores de salida. Puede ser suficiente con suponer que a priori todas las probabilidades iguales. Los valores de salida para las medias y las varianzas se pueden determinar detectando modos en el histograma o simplemente dividiendo el histograma en dos partes respecto a su valor medio, calculando las medias y varianzas de las dos partes para utilizarlas como valores de salida.

La identificación fidedigna de los picos del modo de un cierto histograma, es uno de los aspectos más importantes de la selección de umbral.

Esta capacidad es particularmente importante para la selección automática de umbrales en situaciones donde las características de una imagen pueden variar en un amplio rango de distribuciones de intensidad.

De forma intuitiva se podría pensar que las posibilidades de seleccionar un “buen” umbral aumentan considerablemente si los picos del histograma son altos, estrechos, simétricos y separados por valles profundos.

Para mejorar la forma de los histogramas se puede solamente considerar aquellos píxeles que están situados en el límite entre los objetos y el fondo o cerca de él. De este modo los histogramas serán menos dependientes de los tamaños relativos de los objetos y del fondo.

Además, la utilización de píxeles que satisfacen alguna de las medidas sencillas basadas en los operadores gradiente y laplaciano tienen tendencia a hacer más profundo los valles entre los picos del histograma.

Este método tiene un problema, que es la suposición implícita de que se conoce el límite entre los objetos y el fondo. Esta información no está claramente disponible durante la segmentación, si se encuentra una división entre objetos y fondo, precisamente porque la segmentación está por todas partes.

Sin embargo, se podría obtener una indicación de si un píxel está en un borde calculando su gradiente. Además, el uso del laplaciano puede proporcionar información respecto a si un píxel dado está situado en la parte oscura (fondo) o clara (objeto) de un borde. El valor medio del laplaciano es cero en la transición de un borde, por lo que en la práctica los valles de los histogramas formados por los píxeles seleccionados por un criterio de gradiente/laplaciano, se puede esperar que estén escasamente poblados. Esta propiedad produce los deseados valles profundos de los que hablábamos.

3.2.2.1.2 Umbralización multinivel por luminancia y por segmento de color

Podemos describir dos casos en la umbralización multinivel. Uno sería umbralización multinivel de luminancia y el otro umbralización multinivel de componentes de color.

Para el caso de umbralización multinivel de luminancia, la idea es similar a la binarización, solo que con varios umbrales que definen varias clases.

La idea de lo que se haría en el histograma la podemos ver gráficamente:

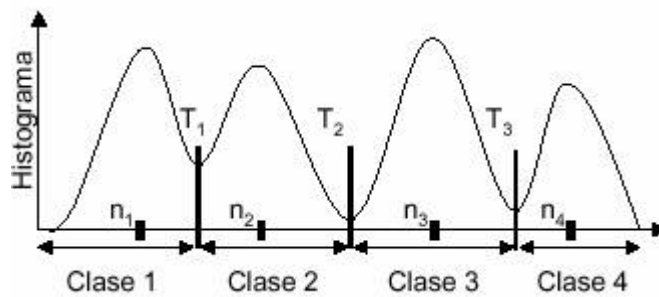


Figura 15: Histograma de umbralización multinivel de luminancia

Una efectiva segmentación multinivel se puede lograr aplicando un método iterativo. En la primera etapa del proceso la imagen es umbralizada para separar regiones con más brillo de aquellas más oscuras, para ello se localiza un mínimo entre los modos del histograma.

De cada parte segmentada, se forman otros histogramas. Si estos no son unimodales, las partes son umbralizadas otra vez. El proceso se repite hasta que los histogramas de cada parte se concierten en unimodales.

El siguiente tipo de segmentación multinivel es por segmentos de color. Un ejemplo típico es el de las imágenes en color, donde se utilizan los componentes rojo (R), verde (G) y azul (B) para formar una imagen de color compuesta. Este tipo de segmentación no la abordaremos en este proyecto por lo que simplemente haremos una breve descripción.

Cada píxel estaría caracterizado por tres valores, lo cual permitiría construir un histograma tridimensional. El concepto de umbralización se transforma ahora en encontrar agrupaciones de puntos en un espacio tridimensional.

Si suponemos que se han encontrado en el histograma K agrupaciones de puntos significativas, la imagen se puede segmentar asignando una intensidad a los píxeles cuyos componentes RGB están próximos a una agrupación y otra intensidad al resto de píxeles de la imagen.

La principal dificultad estriba en la búsqueda de agrupaciones, la cual se convierte en una tarea compleja que se incrementa conforme aumenta el número de variables.

La segmentación realizada por umbrales será válida si los objetos a segmentar tienen valores de intensidad constantes sobre

toda la imagen pero, si la iluminación no es uniforme sobre toda la escena, puede suceder que un umbral no sea suficiente para poder segmentar la imagen.

Otra de las limitaciones de estos métodos consiste en que el histograma nos da información de la distribución global de la intensidad de una imagen. Imágenes muy diferentes pueden tener diferentes distribuciones espaciales de niveles de gris, pero tener histogramas muy similares, por lo que hay que tener precaución ya que este método no realiza controles de conectividad espacial de las distintas regiones que conforman la imagen.

3.2.2.2 Segmentación por crecimiento de regiones

Los métodos de segmentación basados en regiones, parten de la idea de que la imagen a segmentar es una región homogénea desde algún punto de vista. Existe un criterio de homogeneidad P que cumple cada región individualmente, pero que no se cumple en el caso de la unión de dos regiones (texturas, niveles de gris,...).

El criterio más sencillo es la homogeneidad de nivel de gris. Si suponemos que cada región se caracteriza por un nivel de gris concreto, la segmentación consistiría en detectar esas regiones. Las técnicas orientadas a regiones funcionan bien en imágenes con ruido y distorsiones, donde los bordes son muy difíciles de detectar.

Vamos a ver el proceso de segmentación como un proceso en el cual dividimos la imagen I en n subregiones $R_1 R_2 \dots R_n$ tal que:

- 1) La segmentación debe ser completa y todo píxel debe estar en alguna región.
- 2) R_i con $i=1,2,\dots,n$ es una región conectada, es decir, todos los píxeles pertenecientes a una región han de estar conectados.
- 3) $R_i \cap R_j = 0$ para todo i y j con $j \neq i$, es decir, las regiones han de ser disjuntas y ningún píxel puede pertenecer a dos regiones.

Para algún predicado lógico P :

- 4) $P(R_i) = \text{CIERTO}$ (para $i=1,2,\dots,n$, es decir, los píxeles de una región segmentada deben satisfacer unas determinadas propiedades (por ejemplo, $P(R_i) = \text{CIERTO}$ si todos los píxeles de R_i tienen la misma intensidad).

- 5) $P(R_i \cup R_j) = \text{FALSO}$ para $j \neq i$ y R_i adyacente a R_j , es decir, las regiones adyacentes R_i y R_j son distintas según el criterio del predicado P .

Sin embargo, aunque en teoría el proceso es sencillo, decidir cuando una región satisface o no el criterio de similaridad, no es del todo sencillo.

En este método, las regiones crecen mediante agregación de píxeles similares en valor respecto a la propiedad P que se utilice para realizar la segmentación.

Este tipo de algoritmos necesita que el usuario seleccione un conjunto de *puntos semilla* en la imagen. Estos puntos semillas servirán como puntos de comienzo del proceso de crecimiento de las regiones, con lo cual, el número final de regiones ha de ser como mucho igual al número de semillas sembradas por el usuario (puede ser menor, pues en algún paso del algoritmo se puede decidir unir dos regiones para formar una sola).

Para poder realizar la agregación de píxeles similares será necesario definir el concepto de similaridad anteriormente comentado, que no tiene por qué ser el mismo para todo tipo de aplicaciones.

Posibles criterios, ya utilizados en algoritmos desarrollados, pueden ser que la diferencia entre el valor del píxel a agregar y el valor de la semilla, o el valor medio de la región ya formada, sea menor que un cierto umbral predeterminado.

Una propiedad P que se puede utilizar para incluir un píxel en una de las regiones sería que la diferencia absoluta entre el nivel de gris de un píxel N_i y el generador N_g sea menor que un umbral T . Además, será necesario utilizar un criterio para decidir en caso de empate.

Por ejemplo, en la siguiente imagen se toma como puntos semilla los puntos de coordenadas (3,2) y (3,4). Para la propiedad:

$$P = |N_i - N_g| < T.$$

Si la diferencia entre dos píxeles, T , es menor o igual (caso de empate) a 3 se tienen dos regiones, en cambio si $T=8$ se tendrá una sola región.

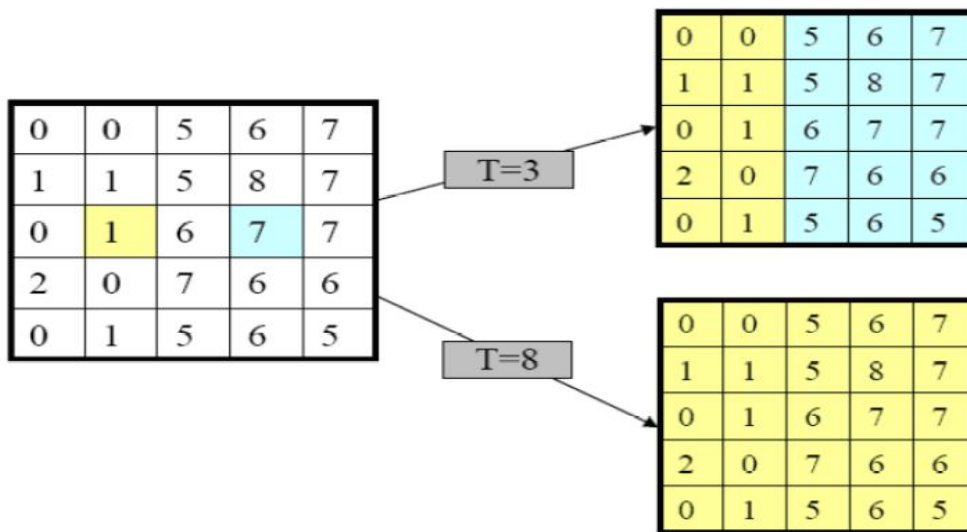


Figura 16: Crecimiento de regiones

En las siguientes imágenes podemos observar la secuencia de crecimiento, en la segmentación de un hueso en una imagen obtenida por resonancia magnética. La región va creciendo gracias a las características de similitud impuestas en el algoritmo de segmentación por crecimiento de regiones, hasta alcanzar el borde, límite impuesto que no cumple las condiciones de inclusión, donde se detiene el crecimiento de la región.

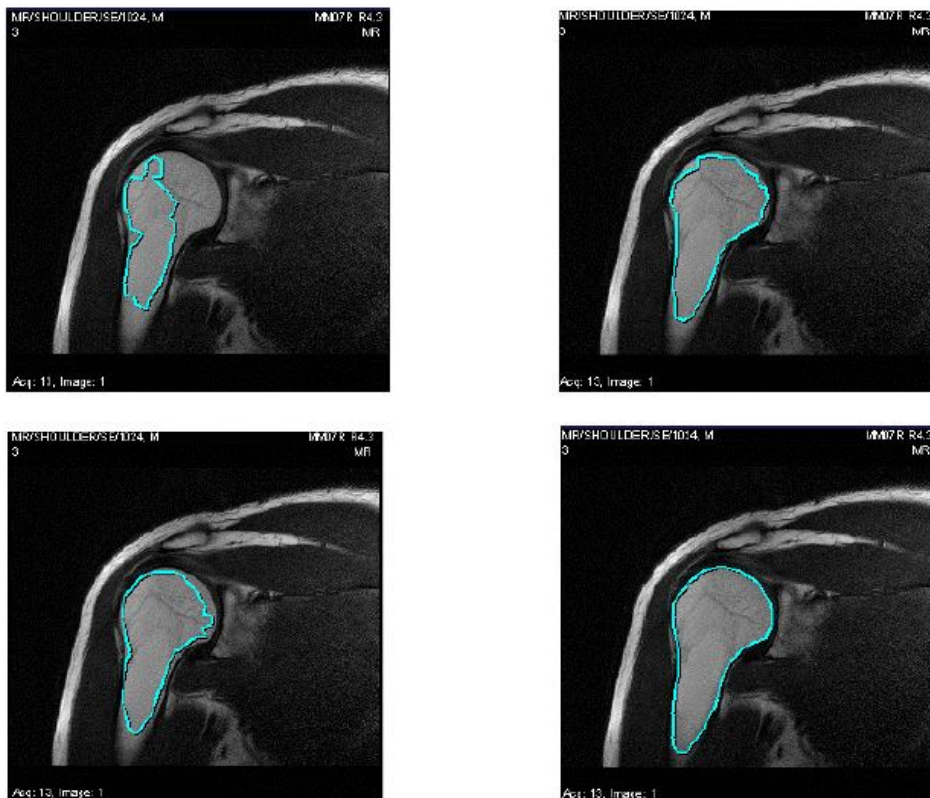


Figura 17: Secuenciación de crecimiento

Algunos problemas fundamentales en el crecimiento de regiones son:

1) La *selección de las semillas o puntos de partida* que representen adecuadamente a las regiones de interés. Para solventar este primer inconveniente, podemos actuar de dos formas:

- Basándonos en la naturaleza del problema. Por ejemplo, en imágenes de infrarrojos, los blancos de interés normalmente desprenden calor, por lo que corresponderán a píxeles claros frente a fondo oscuro. En este caso, los píxeles claros son una elección natural para las semillas.
- En ausencia de conocimiento a priori sobre el problema, un procedimiento que suele dar lugar a buenos resultados consiste en calcular para cada píxel de la imagen las mismas propiedades que luego se emplearán durante el proceso de crecimiento. Si el resultado del cálculo de dichas propiedades da lugar a agrupaciones de los valores para dichas propiedades, se pueden emplear como semillas los píxeles cercanos a los centros de dichas agrupaciones. Por ejemplo, los píxeles correspondientes a los valores máximos de cada modo del histograma de nivel de gris podrían servir como semillas cuando se emplee como propiedad la similitud en el nivel de gris.

El gran inconveniente del crecimiento de regiones es que puede dar diversos resultados dependiendo de donde caigan las semillas y del umbral que se utilice para determinar cuándo un píxel pertenece o no a una región.

Así, si tomamos la figura siguiente, se observa en el caso (a) como el nivel de gris del fondo sigue una rampa de menos a más, mientras que el objeto sigue otra en sentido contrario. En el caso (b) las semillas han caído en el centro del objeto y de cada lado del fondo mientras que el umbral T es la mitad del salto en los niveles de gris. Con esto la clasificación se realiza de forma correcta. Sin embargo, en el caso (c) las semillas del objeto han caído en los dos extremos con lo que el resultado es que la imagen puede separarse en cuatro zonas distintas. Peor resultado se observa en el caso (d) donde ahora el umbral toma la mitad del valor de los casos anteriores con lo que se obtienen siete zonas distintas. El crecimiento en regiones dará por tanto más objetos de los presentes en la imagen.

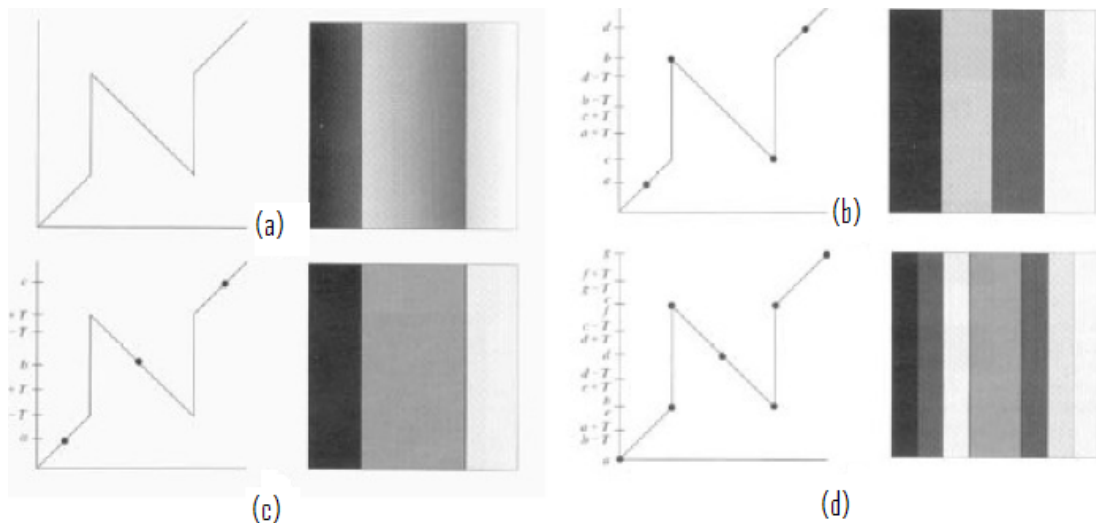


Figura 18: Crecimiento de regiones con distintos umbrales

2) La *elección de las propiedades* adecuadas que permitan ir añadiendo píxeles durante el proceso de crecimiento es otro inconveniente. Para la selección del criterio de similitud, dependeremos no sólo del problema considerado, sino también del tipo de imagen disponible. En imágenes en color o multispectrales (fotografías desde satélite, por ejemplo) se pueden emplear propiedades de similitud en los diferentes canales disponibles. Sin embargo, en general, lo habitual es trabajar con imágenes de intensidad con un único canal disponible. En este proyecto las propiedades se basarán en descriptores locales a partir de una única imagen como son la intensidad y propiedades espaciales.

En cualquier caso siempre hay que tener en cuenta la conectividad durante el proceso de crecimiento para que el resultado tenga significado dentro de su contexto.

3) Otro problema añadido en el crecimiento de regiones es la *regla de parada*. En general, una región deja de crecer cuando no existan más píxeles que satisfagan el criterio de inclusión en esa región. A veces, si se conoce la forma del objeto a segmentar, ésta es tomada en cuenta a la hora de detener el crecimiento (por ejemplo, la segmentación de objetos circulares).

4) El *crecimiento a partir de un píxel semilla* en particular, permitiendo que su región crezca completamente antes de tratar otras semillas puede tener varios efectos:

- Crecimiento de regiones dominantes: Ambigüedades en torno a los bordes de regiones adyacentes pueden no ser resueltas correctamente.
- La elección de diferentes puntos semilla, puede dar lugar a diferentes segmentaciones.

- Pueden surgir problemas si un píxel semilla (elegido arbitrariamente) pertenece a un borde.

Para tratar de resolver estos problemas se han desarrollado técnicas de crecimiento de regiones de forma simultánea en los algoritmos de segmentación diseñados en este proyecto:

- No se permite que una única región domine completamente el proceso.
- Un cierto número de regiones crecen al mismo tiempo (regiones similares, presentan un mismo comportamiento en su crecimiento).

3.3 Etiquetado

Una vez segmentada una imagen hace falta agrupar los píxeles para formar objetos. Esto se realiza mediante técnicas de etiquetado, que consiste en identificar mediante una etiqueta cada objeto existente en la imagen.

Como objeto se define aquel grupo de píxeles que presentan características similares, en concreto en los algoritmos de segmentación realizados el etiquetado se lleva a cabo teniendo en cuenta los niveles de gris de la imagen. Para etiquetar se realiza una pasada por la imagen de arriba hacia abajo y de izquierda a derecha, rellenando con la etiqueta los nuevos objetos encontrados. Gracias al etiquetado podremos reconstruir la imagen compuesta por las distintas regiones dependiendo del algoritmo y de los criterios de inclusión impuestos.

3.4 Post-Procesado de la imagen

En los algoritmos de segmentación automáticos implementados se han utilizado filtros de suavizado, en concreto filtros de mediana, que pretenden homogeneizar la imagen, eliminando pequeños huecos que quedan en las regiones segmentadas.

El filtro de mediana es un filtro que da el valor del píxel según la mediana de los 8 vecinos del píxel. Estos filtros reducen significativamente el tramado de la imagen, a costa de emborronar los detalles, especialmente los bordes de los objetos. En caso necesario se pueden aplicar sucesivamente dos o más veces. El efecto de la aplicación del filtro de mediana se muestra en el apartado 5.3 Resultado de Imágenes