

Capítulo 8

Representación de formas y descripción

8.1. Introducción

Una vez segmentada la imagen y extraído el contorno de los objetos hay que analizar la forma geométrica de los mismos utilizando para su representación una estructura de datos compacta, es decir, lo que llamaremos un *esquema de representación*. Asimismo, hay que estudiar algunos *descriptores* de contornos y de regiones para el reconocimiento e identificación de los objetos de la imagen. Finalmente, se va a estudiar un conjunto de operadores morfológicos que nos van a permitir manipular la forma de los objetos, extraer el contorno de los mismo, clasificar los objetos según su tamaño, eliminar objetos pequeños, etc.

Los *esquemas de representación* de formas deben de tener ciertas propiedades deseables:

- a) Unicidad: cada objeto debe tener una única representación.
- b) Invariancia frente a transformaciones geométricas, como traslaciones, rotaciones, cambios de escala y reflexiones.
- c) Sensibilidad o capacidad para diferenciar objetos casi iguales.
- d) Abstracción del detalle o capacidad para representar los rasgos característicos básicos de los objetos y abstraer los detalles.

Vamos a distinguir entre *esquemas de representación externa*, que usan el contorno de los objetos y sus rasgos característicos, como son los códigos de cadena, los descriptores de Fourier y las aproximaciones poligonales, y los *esquemas de representación interna*, que describen la región ocupada por el objeto en la imagen binaria, como son el área, los momentos, el esqueleto, etc.

8.2. Esquemas de representación externa: Trazado y representación de contornos

Vamos a estudiar dos maneras de representar las fronteras (contornos) de las regiones (objetos) que configuran la imagen digital que son útiles para estudiar la forma del contorno de los objetos, realizar su descripción y conseguir una identificación o reconocimiento de los objetos de la imagen.

a) Códigos de cadena de Freeman

Supongamos que tenemos el contorno de un objeto, en una imagen binaria, representado por una cadena de píxeles con valor 1 en la imagen binaria. El código de cadena es un tipo de estructura de datos para representar el contorno de un objeto en una imagen binaria mediante una secuencia de segmentos, conectados consecutivamente, de longitud y orientación específica, que conectan píxeles adyacentes.

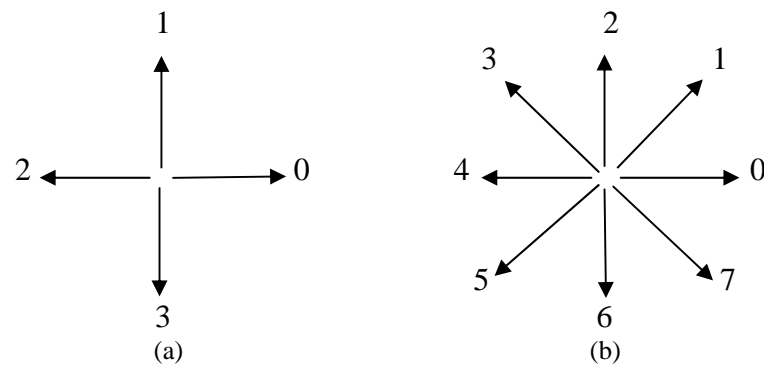


Figura 8.1. (a) Direcciones para entornos de 4 vecinos
(b) Direcciones para entornos de 8 vecinos.

La conexión de los segmentos se lleva a cabo en entornos de 4 vecinos o de 8 vecinos. Cuando se usa un entorno de 4 vecinos tendremos cuatro orientaciones para los segmentos, como se muestra en la figura 8.1(a) y utilizaremos los números 0,1,2 y 3, para especificar la orientación de los segmentos. Si se usa un entorno de 8 vecinos tendremos 8 orientaciones posibles (figura 8.1(b)).

El código de cadena genera una secuencia de números que representan las orientaciones de los segmentos conectados consecutivamente, partiendo de un punto del contorno y siguiendo el sentido de las agujas del reloj. Por ejemplo, la cadena 005577443221 representa el contorno del objeto de la figura 8.2 cuando empezamos por el píxel superior izquierdo del contorno. Aunque depende del punto de partida, ello no supone un problema pues se puede considerar el código como una lista circular en el caso de contornos cerrados. Además, como es conveniente que tenga la propiedad de unicidad, entre todas las cadenas posibles, según el punto de partida que hayamos tomado, elegimos aquella que corresponda al menor número entero (cada cadena nos da un número entero cuyos dígitos varían del cero al tres para entornos de cuatro vecinos o del cero al siete para entornos de ocho vecinos).

Cuando se trata de objetos formados por regiones no conectadas o con agujeros se necesita más de una cadena para representarlos y hay que especificar si la cadena corresponde a la parte exterior del objeto o a un agujero. En cualquier caso, es también fácil reconstruir el objeto completo a partir del código de cadena mediante una operación de rellenado.

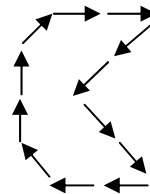
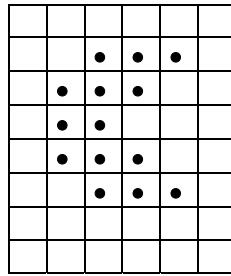


Figura 8.2. Código de cadena 005577443221.

Ventajas frente a la representación matricial de un objeto binario:

- El código de cadena es una representación invariante frente a *traslaciones*. Esta propiedad facilita la comparación de objetos.
- A partir del código de cadena se pueden obtener ciertas características del contorno, como el perímetro, el área del objeto y los descriptores de Fourier, de forma más eficiente que utilizando la representación matricial de la imagen binaria.
- El código de cadena es una representación compacta de un objeto binario. Suministra una buena compresión de la descripción del contorno ya que cada cadena se puede codificar sólo con dos bits (para entornos de 4 vecinos) en lugar de las coordenadas (x,y) de cada píxel del contorno. Por ejemplo, si un círculo de radio R lo representamos por los $(2R)^2$ píxeles del cuadrado más pequeño que lo contiene, necesitamos para su almacenamiento $(2R)^2$ bits, mientras que si utilizamos un código de cadena con entornos de cuatro vecinos necesitamos en torno a $2\pi R$ bits.

Inconvenientes:

- La cadena resultante suele ser demasiado larga
- Cualquier perturbación o ruido en el contorno produce segmentos erróneos.

Para **trazar los contornos** de objetos binarios se puede emplear un algoritmo similar al algoritmo de la **tortuga de Papert** (1973). Dicho algoritmo trabaja para entornos de 4 vecinos. La tortuga comienza en un píxel del contorno. Si el píxel en el que se encuentra la tortuga es del objeto entonces avanza al píxel vecino en la orientación actual aumentada en 90° y si es del fondo entonces avanza al píxel vecino en la dirección actual disminuida en 90° . El algoritmo

acaba cuando la tortuga alcanza el punto de partida. Dicho algoritmo puede generar bucles que se pueden eliminar en un procesamiento posterior.

b) Signaturas

Una *signatura* es una representación de un contorno mediante una función real unidimensional que sea más sencilla que la función bidimensional que define el contorno. Hay varias maneras de definir una signatura. Una de las más simples es a través de la distancia desde un punto interior, como puede ser el *centroide* del contorno, a cada uno de los puntos del contorno, como una función del ángulo, tal y como se muestra en la figura 8.3 para una circunferencia y en la figura 8.4 para un cuadrado. La signatura sólo tiene sentido usarla cuando el vector que se extiende desde el origen corta en un solo punto al contorno.

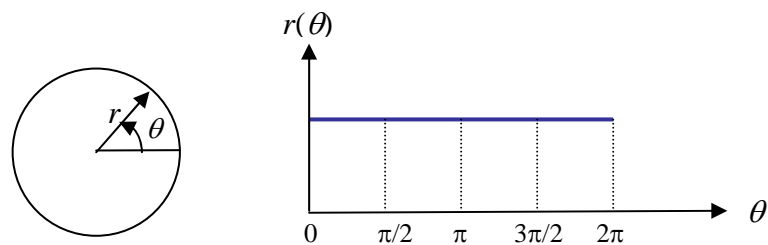


Figura 8.3. Signatura de una circunferencia.

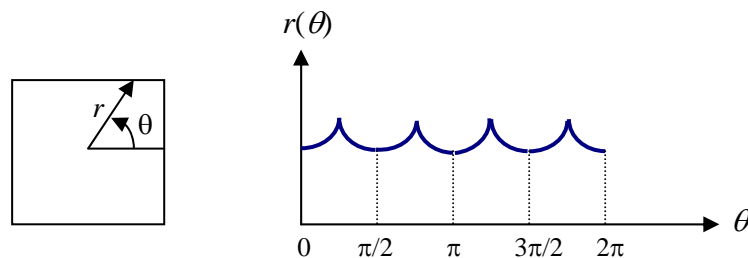


Figura 8.4. Signatura de un cuadrado.

La signatura es invariante frente a traslaciones pero no lo es frente a rotaciones o cambios de escala. Sin embargo, se puede conseguir la invariancia frente a rotaciones cuando se encuentra un punto característico del contorno a partir del cual se comienza a generar la signatura. Dicho punto puede ser, por ejemplo, el más cercano al centroide, siempre que sea único, o un punto del contorno determinado por la intersección de este con su eje mayor.

Por otra parte, como los cambios en el tamaño del contorno conllevan cambios en los valores de la amplitud de la signatura, se puede normalizar la signatura de forma que sus valores estén dentro del intervalo $[0, 1]$, sin más que dividirla por su amplitud máxima para conseguir la adimensionalidad. Sin embargo, en el caso de contornos con ruido en los que el valor máximo pueda venir afectado por dicho ruido entonces no es conveniente la normalización. Una alternativa para realizar la normalización puede ser dividir por la desviación típica de los valores de la signatura.

c) Descriptores de Fourier

Si el contorno viene dado por una curva cerrada se puede representar mediante los descriptores de Fourier. La curva del contorno va a venir dada por las coordenadas paramétricas, $(x(t), y(t))$, que son un par de funciones periódicas, donde el parámetro t describe la longitud del camino de la línea de contorno calculada desde el punto de partida.

En primer lugar seleccionamos N puntos equidistantes del contorno (muestreo),

$$(x(t), y(t)), \quad n=0,1,2,\dots,N-1,$$

y a partir de estos valores obtenemos el vector complejo

$$z = x + i y,$$

siendo $x = (x(0), x(1), \dots, x(N-1))$, $y = (y(0), y(1), \dots, y(N-1))$.

La transformada de Fourier discreta de este vector complejo viene dada por la expresión,

$$Z(u) = \frac{1}{N} \sum_{n=0}^{N-1} z(n) \exp\left(-\frac{2\pi n u i}{N}\right), \quad u=0,1,2,\dots,N-1.$$

Dichos valores se conocen con el nombre de **descriptores de Fourier Cartesianos del contorno**. Además, a partir de ellos se puede reconstruir el contorno mediante la transformación inversa,

$$z(n) = \frac{1}{N} \sum_{k=0}^{N-1} Z(k) \exp\left(\frac{2\pi n k i}{N}\right), \quad n=0,1,2,\dots,N-1.$$

El coeficiente $Z(0)$,

$$Z(0) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) + i \frac{1}{N} \sum_{n=0}^{N-1} y(n),$$

representa el “punto medio” (centro de gravedad) del contorno del objeto, llamado *centroide*.

Como se trata de la transformada de Fourier discreta, los coeficientes de Fourier $Z(k)$ representan las pequeñas variaciones en las tendencias del contorno para valores de k pequeños y las grandes variaciones para valores de k grandes. Es decir, componentes de alta frecuencia

tienen en cuenta los detalles más finos del contorno mientras que las componentes de baja frecuencia determinan la forma global del contorno. Además, se conseguirán detalles más finos del contorno del objeto conforme mayor sea N .

Supongamos ahora que en lugar de considerar todos los descriptores sólo se consideran los P primeros descriptores de Fourier. Es decir, tomamos $Z(u) = 0, u = P+1, \dots, N-1$. Entonces

$$\hat{Z}(u) = \frac{1}{N} \sum_{n=0}^{P-1} z(n) \exp\left(-\frac{2\pi n u i}{N}\right) ,$$

es una *aproximación* de $Z(u)$, $u = 0, 1, 2, \dots, N-1$ que utiliza el mismo número de puntos N pero no el mismo número de descriptores en la representación. Esta aproximación supone una pérdida de detalle en el contorno que se incrementa conforme P decrece; es una suavización del contorno.

Los descriptores de Fourier deberían ser insensibles a traslaciones o rotaciones, o a cambios de escala. Sin embargo, los descriptores de Fourier no son invariantes a tales transformaciones geométricas. Sin embargo, dichos cambios están relacionados con transformaciones simples de los descriptores.

Una traslación en las coordenadas de la curva,

$$z_t(n) = z(n) + z_o$$

afecta sólo al término $Z(0)$ de la representación, según la expresión:

$$Z_t(0) = Z(0) + z_o$$

Un cambio de escala con respecto a un sistema de coordenadas que tuviera su origen en el centro de gravedad de la curva afecta a los coeficientes de Fourier de forma similar, es decir,

$$z_e(n) = a z(n), \quad Z_e(k) = a Z(k)$$

Un cambio en el punto de partida,

$$z_p(n) = z(n_1 - n_0)$$

produce modulación en los descriptores de Fourier,

$$Z_p(k) = Z(k) \exp(-2\pi n_0 k i / N).$$

La rotación de la curva con un ángulo θ ,

$$z_r(n) = z(n) \exp(i\theta),$$

produce un desplazamiento en la fase de los coeficientes de Fourier de igual magnitud,

$$Z_r(k) = Z(k) \exp(i\theta)$$

Las magnitudes de los descriptores de Fourier, $|Z(k)|$, $k=0,1,2,\dots,N-1$, son invariantes frente a rotaciones y traslaciones.

La fase de los descriptores de Fourier, $\arg(Z(k))$, $k=0,1,2,\dots,N-1$, es invariante frente a cambios de escala.

Por lo tanto, los descriptores de Fourier nos suministran descriptores de forma invariantes, muy útiles para reconocer objetos.

Para identificar cierto tipo de objetos, según su contorno, podemos utilizar como medida el error cuadrático medio,

$$ECM = \sum_{k=0}^{N-1} (|Z_1(k)| - |Z_2(k)|)^2,$$

donde se comparan los descriptores de Fourier de la curvas $z_1(n)$ y $z_2(n)$ para ver si dichas curvas son idénticas, aunque una sea una versión girada de la otra. Dicho error debe ser prácticamente cero (teóricamente cero).

Los descriptores de Fourier sirven también para detectar *simetrías* en los objetos. Si un objeto se gira alrededor de su centro de gravedad $2\pi/n$ sin que cambie su forma se dice que tiene una simetría de n pliegues. El contorno de un objeto con este tipo de simetría contiene un patrón que se repite n veces y, por lo tanto, sólo los descriptores de Fourier con los índices $k=pn$, $p=0,1,2,\dots$, son diferentes de cero.

d) Aproximaciones poligonales

Se trata de aproximar de la mejor manera posible el contorno de un objeto mediante una curva de tramos lineales que constituye un polígono, y donde los vértices de dicho polígono son una representación del contorno del objeto. Como criterio para evaluar la calidad del ajuste se puede elegir el criterio de mínimos cuadrados.

Supongamos una curva que va del punto \mathbf{x}_1 al punto \mathbf{x}_N y sean $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N-1}$, las coordenadas de $N-2$ puntos intermedios de la curva. Si \mathbf{d}_i es el punto del segmento $[\mathbf{x}_1, \mathbf{x}_N]$ más próximo al punto \mathbf{x}_i , $i=2,3,\dots,N-1$, entonces

$$|\mathbf{x}_i - \mathbf{d}_i|$$

es el error de aproximación de la curva por el segmento lineal $[\mathbf{x}_1, \mathbf{x}_N]$ correspondiente al píxel \mathbf{x}_i , $i=2,3,\dots,N-1$ (figura 8.5).

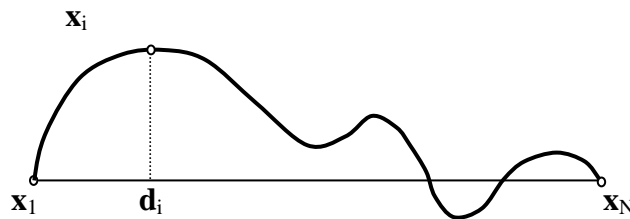


Figura 8.5. Aproximación poligonal.

Podemos elegir, como criterio para medir la bondad del ajuste, el criterio del mínimo error cuadrático medio, es decir, que el error dado por la expresión

$$\sum_{i=2}^{N-1} |\mathbf{x}_i - \mathbf{d}_i|^2$$

sea mínimo.

También se puede utilizar el criterio *minimax*, es decir, minimizar el error dado por la expresión

$$\max_{2 \leq i \leq N-1} |\mathbf{x}_i - \mathbf{d}_i|.$$

La aproximación poligonal consistirá en determinar los vértices del polígono de manera que el error total sea mínimo. Sin embargo, encontrar la solución óptima para este problema requiere un coste computacional muy alto. Por ello, hay técnicas alternativas que suministran buenas soluciones (cercanas a la óptima) y son más rápidas y eficientes. Las técnicas de partición consisten en ir dividiendo, de forma recurrente, los trozos de curva, representados por un segmento, en dos trozos, representados cada uno de ellos por un segmento, de manera que la medida local de error sea mínima. Si utilizamos el criterio *minimax* entonces para cada trozo de curva representado por un segmento se determina el punto de la curva más alejado del segmento y se toma dicho punto como nuevo vértice, como puede verse en la figura 8.6. El proceso acaba cuando el error máximo por píxel es menor que cierta cantidad prefijada. Los puntos \mathbf{x}_1 y \mathbf{x}_N pueden ser los dos puntos más alejados del contorno. Dichos puntos dividen al contorno en dos partes y el algoritmo se aplica a cada una de ellas. Obsérvese que, en cada paso, el error por píxel es igual o menor que dicho error en el paso anterior. Conforme se van introduciendo nuevos vértices se mejora la aproximación.

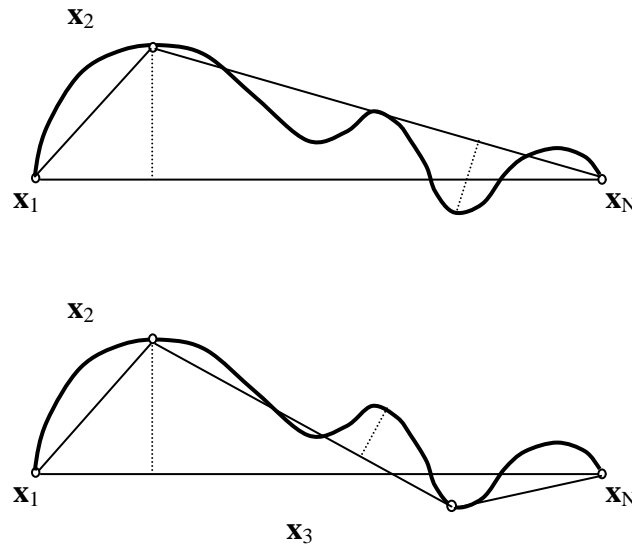


Figura 8.6. Aproximación poligonal por partición sucesiva.

Las técnicas de incorporación operan al contrario, es decir, parten de un punto \mathbf{x}_1 de la curva y van recorriendo los píxeles de la curva en el sentido de las agujas del reloj. Cuando se encuentra el primer píxel, digamos \mathbf{x}_i , para el que el error de \mathbf{x}_1 a \mathbf{x}_i supera un valor umbral T (prefijado), se declara dicho punto como vértice del polígono y se repite el proceso comenzando ahora en \mathbf{x}_i . Se genera así una secuencia de vértices que constituye la aproximación poligonal. Las técnicas de incorporación presentan el inconveniente de que los vértices no suelen ser puntos de inflexión, como ocurre con las técnicas de partición, generando vértices que no coinciden con las esquinas de la curva.

8.3 Descriptores de contornos

A continuación vamos a estudiar la forma geométrica de los contornos de las regiones (objetos) que conforman la imagen digital utilizando para ello descriptores, es decir, valoraciones numéricas que nos van a permitir identificar y reconocer los objetos de dicha imagen.

Vamos a describir la forma de los contornos de los objetos de la imagen binaria mediante parámetros de forma geométricos. El **perímetro** viene dado por la longitud de su código de cadena (con ocho direcciones) pero ponderando los pasos diagonales por $\sqrt{2}$ y los horizontales y verticales por 1, es decir, viene dado por el número total de píxel que configuran su contorno pero los píxeles de bordes diagonales se ponderan con $\sqrt{2}$.

El **diámetro** de un contorno (frontera de una región) viene dado por la distancia Euclídea entre los dos píxeles del contorno más alejados. Dichos puntos no son siempre únicos, como ocurre en una circunferencia, pero es un descriptor de interés cuando sí lo son. La recta que pasan por dichos puntos se llama *eje mayor* de la región. El rectángulo, con dos lados paralelos al eje mayor, que tiene la propiedad de que es el menor rectángulo que contiene al contorno se llama *rectángulo base* (figura 8.7). El cociente entre la longitud del lado mayor y la longitud del lado menor se llama *excentricidad* del contorno.

El centro de gravedad o *centroide* de un contorno determinado por el conjunto de píxeles $\{(x_i, y_i), i=1,2,\dots,N\}$ es el punto (\bar{x}, \bar{y}) definido por las siguiente expresiones:

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N}, \quad \bar{y} = \frac{\sum_{i=1}^N y_i}{N}$$

El *eje menor* del contorno viene definido por la recta perpendicular al eje mayor que pasa por el centro de gravedad del contorno.

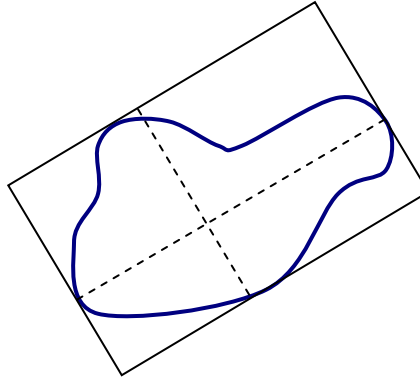


Figura 8.7. Rectángulo base de un contorno.

Todos los parámetros anteriores no dependen de la orientación de los objetos en la imagen, son invariantes frente a traslaciones pero no lo son frente a transformaciones de escala.

La **curvatura** se define como la tasa de cambio de la pendiente (tangente) del contorno, pero es difícil de obtener medidas fiables en una imagen digital porque los bordes suelen ser localmente “mellados”. Sin embargo, se pueden obtener descriptores de la curvatura bastante útiles mediante diferencia de las pendientes de segmentos adyacentes del contorno.

Otros descriptores de la forma de los contornos son los **momentos estadísticos**. La forma de una representación unidimensional de un contorno a través de función real, $g(x)$, se puede describir utilizando momentos estadísticos, como la media, la varianza o momentos de orden superior. Dicha función puede ser la signatura del contorno. En el caso de contornos abiertos, se puede utilizar la función que se obtiene de las distancias de los puntos del contorno al segmento que une los dos puntos extremos de dicho contorno (figura 8.8).

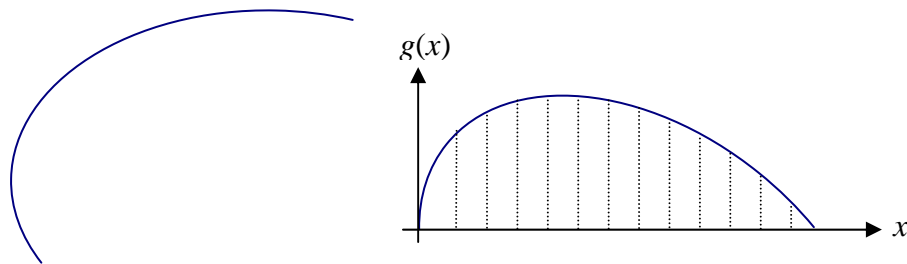


Figura 8.8. Representación unidimensional de un contorno.

Para describir la forma de $g(x)$ normalizamos dicha función para que el área que encierra con el eje de abscisas sea la unidad y se puede considerar como un histograma. Así $g(x)$ se puede interpretar como la probabilidad del valor x . Si suponemos que el contorno viene descrito por N puntos, el valor medio viene dado por la expresión:

$$m = \sum_{i=0}^{N-1} x_i g(x_i)$$

La varianza, dada por la expresión

$$\sigma^2 = \sum_{i=0}^{N-1} (x_i - m)^2 g(x_i),$$

es una medida de la dispersión (variabilidad) de los puntos con respecto a su valor medio.

El momento central de tercer orden,

$$\mu_3 = \sum_{i=0}^{N-1} (x_i - m)^3 g(x_i),$$

es una media de la simetría de la distribución con respecto a su media.

8.4 Descriptores de regiones

Aunque las regiones (objetos) pueden venir definidas por sus contornos, vamos a estudiar aquí algunas características geométricas y topológicas de las regiones que nos ayudarán a identificar y reconocer los objetos de la imagen digital.

a) Parámetros geométricos

El **área** de una región (objeto) viene dada por el número de píxeles que la componen. Se puede obtener también de forma sencilla a partir de su código de cadena.

El centro de gravedad o *centroide* de una región determinada por el conjunto de píxeles $\{(x_i, y_i), i=1,2,\dots,N\}$ es el punto (\bar{x}, \bar{y}) definido por las siguiente expresiones:

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N}, \quad \bar{y} = \frac{\sum_{i=1}^N y_i}{N}$$

Obsérvese que el centroide de una región no tiene por qué coincidir con el centroide de su contorno.

b) Parámetros topológicos

La topología es el estudio de configuraciones geométricas con propiedades específicas como la invariancia bajo ciertas transformaciones, como transformaciones de escala.

La **compacidad** (o circularidad) de una región es un parámetro que no depende del tamaño de la región (como los anteriores) y viene dado por:

$$c = \frac{A}{p^2}$$

donde A es el área y p el perímetro de la región. El perímetro se eleva al cuadrado para conseguir un parámetro adimensional. Su valor máximo corresponde a los círculos (máxima

superficie para un perímetro dado) y vale $1/(4\pi)$ (≈ 0.07957), por ello es una medida de **circularidad**; para el triángulo equilátero vale $1/(12\sqrt{3})$ (≈ 0.048) y para el cuadrado vale $1/16$ (≈ 0.0625). Los valores pequeños del parámetro indican objetos alargados. Además, es invariante frente a traslaciones, giros y cambios de escala. Conviene normalizar sus valores al intervalo $[0,1]$ dividiendo por $1/(4\pi)$ para garantizar que el valor que le asigne al círculo sea 1, y un valor menor a cualesquiera otra figura geométrica, tanto menor cuanto más afechinada sea o presente espículas. Por ello, tomaremos

$$c = 4\pi \frac{A}{p^2}$$

La **rectangularidad** de una región es otro parámetro adimensional que viene definido por el cociente entre el área de la región y el área de su rectángulo base (figura 8.9):

$$r = \frac{A}{ab}$$

siendo a la longitud del lado mayor de su rectángulo base y b la longitud de su lado menor.

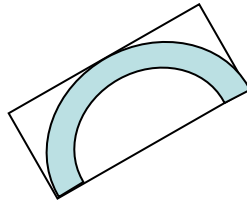


Figura 8.9. Rectángulo base de un objeto.

El **alargamiento** de una región se puede definir por el cociente entre la longitud del lado mayor a y el lado menor b de su rectángulo base. Sin embargo, para regiones curvadas no es una medida adecuada (figura 8.9). Por ello, se define el alargamiento como el cociente entre su área y el cuadrado del valor máximo de su grosura:

$$c = \frac{\text{Área}}{(2d)^2}$$

d se puede determinar como el número de veces que hay que aplicar el operador erosión 3×3 hasta que la región desaparece.

Otras características topológicas importantes son la **conectividad** y los **agujeros**. Una imagen segmentada puede estar compuesta por regiones que tienen componentes conexas que configuran los objetos, es decir, regiones tales que dos puntos cualesquiera de ellas se pueden

unir por una curva contenida en ellas. Un agujero es una región de la imagen que está completamente encerrada por una componente conexa de la imagen.

El **número de Euler** de una imagen se define como:

$$E=C-H,$$

donde C es el número de componentes conexas y H el número de agujeros de la imagen. Este número es invariante frente a traslaciones, rotaciones y cambios de escala, y nos permite de forma sencilla discriminar entre ciertas clases de objetos.

c) Medidas estadísticas para la cuantificación de la textura de una región

Una región puede ser descrita por su *textura* y una manera de cuantificar la textura es utilizando algunos momentos estadísticos del histograma de la intensidad luminosa de los píxeles de la región.

Se define el *momento central de orden r* de los valores de intensidad luminosa (tonos de gris) de los píxeles de una región mediante la expresión:

$$\mu_r = \sum_{i=0}^{L-1} (z_i - m)^r p(z_i)$$

donde $p(z_i)$ es la frecuencia relativa (probabilidad) de los píxeles de la región que presentan el tono de gris z_i ; $m = \sum_{i=0}^{L-1} z_i p(z_i)$ es el valor medio de los tonos de gris de los píxeles de la región y L es el número de tonos o niveles de gris.

Los descriptores estadísticos más utilizados para medir la textura de una región son:

- La media m que nos da el tonos de gris más representativo de la región (cuando se utiliza el criterio de mínimo error cuadrático)
- La desviación típica $\sigma = \sqrt{\mu_2}$ que es una medida de contraste medio de la imagen
- El *coeficiente de suavidad*, $R = 1 - \frac{1}{1 + \mu_2}$, que mide la suavidad relativa o uniformidad

de los tonos de gris de la región; vale cero para una región con el mismo tono de gris (intensidad constante) y se aproxima a 1 cuando hay una gran variabilidad en los tonos de gris de los píxeles de la región. En la práctica, la varianza, μ_2 , utilizada en la expresión de R , se normaliza en el rango $[0, 1]$ dividiéndola por $(L-1)^2$.

- El momento central de tercer orden, μ_3 , es una *medida de sesgo* o *asimetría* del histograma; vale 0 cuando el histograma es simétrico, es positivo cuando el histograma está sesgado a la derecha y negativo cuando está sesgado a la izquierda.
- La cantidad $U = \sum_{i=0}^{L-1} p^2(z_i)$ es una medida de la *uniformidad* de la región. Es máxima cuando todos los niveles de gris presentan la misma frecuencia relativa.

Otra medida estadística importante es la *entropía*, definida por la expresión:

$$H = - \sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$$

La entropía es una medida de la aleatoriedad de los tonos de gris de la región.

d) Medidas espectrales para la cuantificación de la textura de una región

Las medidas espectrales se basan en el espectro de Fourier que es adecuado para describir la direccionalidad de los patrones periódicos o casi periódicos de una región. La textura espectral es útil para discriminar entre patrones de textura periódicos y no periódicos. Además, permite cuantificar las diferencias entre patrones periódicos.

La interpretación de las características del espectro se simplifican expresando el espectro en coordenadas polares para dar una función $S(r, \theta)$. Para cada dirección θ , $S(r, \theta)$ se puede considerar como una función unidimensional $S_\theta(r)$. Análogamente, para cada r , vamos a tener la función unidimensional $S_r(\theta)$. Así, podremos analizar el comportamiento de $S_\theta(r)$ a lo largo de una dirección radial desde el origen y el comportamiento de $S_r(\theta)$ a lo largo de una circunferencia centrada en el origen. Se obtiene una descripción global de la textura mediante las funciones:

$$S_1(r) = \sum_{\theta=0}^{\pi} S_\theta(r) \quad \text{y} \quad S_2(r) = \sum_{r=1}^{R_0} S_\theta(r)$$

Para cada par de valores (r, θ) se obtiene un par de valores $[S_1(r), S_2(\theta)]$. Variando las coordenadas polares se obtienen dos funciones unidimensionales $S_1(r)$ y $S_2(\theta)$ que constituyen una descripción espectral de la textura de la región. Se obtienen descriptores a partir de dichas funciones, como pueden ser la localización del valor máximo, la media y la varianza de las variaciones, y la distancia entre la media y máximo de la función.

e) Medidas invariantes de momentos

Se define el **momento** de orden p y q de la imagen digital $f(m, n)$ por la expresión:

$$m_{pq} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} i^p j^q f(i, j).$$

Si lo calculamos para el objeto determinado por la región S de una imagen binaria vale:

$$m_{pq} = \sum_{(i,j) \in S} i^p j^q$$

Obsérvese que m_{00} nos da el área del objeto y que $(m_{10}/m_{00}, m_{01}/m_{00})$ es el centroide (centro de gravedad) del objeto. Los momentos de orden superior no son invariantes a traslaciones, por ello vamos a realizar una traslación del origen al centroide y obtenemos así los **momentos centrales** de orden p y q mediante la expresión:

$$\mu_{pq} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (i - \bar{i})^p (j - \bar{j})^q f(i, j)$$

donde (\bar{i}, \bar{j}) es el centroide del objeto. Los momentos centrales son así invariantes frente a traslaciones de los objetos. Sin embargo, no son invariantes frente a giros o cambios de escala.

A continuación vamos a ver la interpretación de algunos momentos centrales.

μ_{20} es una medida de la dispersión horizontal del objeto con respecto al centroide

μ_{02} es una medida de la dispersión vertical del objeto con respecto al centroide

μ_{11} es una medida de la diagonalidad indicando el cuadrante de mayor masa.

μ_{12} es una medida de la divergencia horizontal; indica la extensión de la región izquierda del objeto frente a la derecha

μ_{21} es una medida de la divergencia vertical; indica la extensión de la región inferior del objeto frente a la superior

μ_{30} es una medida del desequilibrio (o asimetría) horizontal e indica si el objeto tiene mayor extensión a la izquierda o a la derecha del centroide

μ_{30} es una medida del desequilibrio (o asimetría) vertical.

Si dividimos μ_{pq} por $\mu_{00}^{[1+(p+q)/2]}$ se obtiene el momento central η_{pq} normalizado,

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{1+\frac{p+q}{2}}}$$

que es invariante a cambios de escala, para $p+q=2,3,\dots$.

Un conjunto de siete *invariantes de momentos* que son insensibles a traslaciones, cambios de escala, rotaciones y transformaciones especulares viene dado por las siguientes expresiones:

$$\phi_1 = \eta_{20} + \eta_{02}$$

$$\begin{aligned}
\phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) + \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\
&\quad \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \\
\phi_6 &= (\eta_{20} - 3\eta_{02}) \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\phi_7 &= (3\eta_{21} - 3\eta_{02})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] + (3\eta_{12} - \eta_{03})(\eta_{21} + \eta_{03}) \\
&\quad \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right]
\end{aligned}$$

Se pueden determinar también los **momentos** de orden p y q **invariantes a giros**, que representaremos por ϕ_{pq} y que vienen dados en términos de los momentos centrales:

$$\phi_{pq} = \sum_{r=0}^p \sum_{s=0}^q (-1)^{q-s} \binom{p}{r} \binom{q}{s} (\cos \theta)^{p-r+s} (\sin \theta)^{q-s+r} \mu_{p-r+q-s, r+s}$$

donde θ es el ángulo que forma el eje de mínima inercia con el eje de abscisas, y viene dado por la relación:

$$\theta = (1/2) \arctan(2\mu_{11}/(\mu_{20} - \mu_{02})).$$

8.5 Análisis Morfológico

Una vez realizada la segmentación de la imagen sabemos ya qué píxeles pertenecen a los objetos de interés de la misma. Tendremos así una imagen binaria en la que los píxeles blancos corresponden a los objetos y los negros al fondo. El análisis morfológico va a permitir la mejora de los resultados de la segmentación. Mediante los operadores del análisis morfológico vamos a poder modificar el contorno de los objetos de la imagen binaria, suavizarlos, detectar sus contornos, eliminar objetos de pequeño tamaño, rellenar huecos e identificar objetos.

8.5.1. Operadores morfológicos binarios

Para actuar sobre la forma de los objetos vamos a utilizar plantillas, como las utilizadas para imágenes con diferentes tonos de gris. Ahora, sin embargo, la imagen es binaria por lo que debemos sustituir las operaciones aritméticas por operaciones Booleanas (lógicas); la operación

suma se sustituye por la operación *or* (disyunción) y la multiplicación se va a reemplazar por la operación *and* (conjunción).

Sea $\{f(i,j): i=0,...,M-1, j=0,...,N-1\}$ una imagen binaria de tamaño $M \times N$ y $\{g(i,j): i=-m,...,0,...,m, j=-m,...,0,...,m\}$ una plantilla (máscara) simétrica de tamaño $(2m+1)(2m+1)$ cuyos elementos son ceros o unos. La nueva imagen $f \oplus g$ que se obtiene por convolución de f y g viene dada por la expresión:

$$f \oplus g(i, j) = \bigvee_{h=-m}^m \bigvee_{k=-m}^m f(i-h, j-k) \wedge g(h, k) \quad (1)$$

donde los símbolos \vee y \wedge representan a las operaciones Booleanas *or* y *and*, respectivamente.

También se puede expresar de la siguiente manera:

$$[f \oplus g](m, n) = \begin{cases} 1 & \text{si } [f * g](m, n) \geq 1 \\ 0 & \text{si } [f * g](m, n) = 0 \end{cases}$$

Para ver como actúa este operador, es decir, la plantilla sobre la imagen original, supongamos que todos los elementos de la plantilla son unos; si en el entorno correspondiente de un píxel del objeto hay al menos un píxel con valor uno entonces el píxel toma el valor uno (se incorpora al objeto). Por lo tanto, el objeto original se agranda. Además, los agujeros pequeños o fisuras del objeto original se rellenan y las líneas del contorno se hacen más suaves. Por ello, el operador definido en (1) se conoce con el nombre de **operador de dilatación**. Este operador también viene dado como el valor máximo de los productos de cada elemento de la plantilla por el correspondiente elemento de la imagen.

Si aplicamos el operador dilatación al fondo de la imagen se erosionan los objetos de la imagen. Por ello, la erosión de los objetos de una imagen se consigue calculando el negativo de la imagen (para que el fondo sea el objeto) y dilatando el objeto (fondo), es decir, mediante el operador:

$$f \ominus g(i, j) = \overline{\bigvee_{h=-m}^m \bigvee_{k=-m}^m f(i-h, j-k) \wedge g(h, k)}$$

donde hemos vuelto a tomar el negativo para obtener la imagen con los objetos originales erosionados.

La erosión de un objeto de la imagen se puede conseguir mediante el operador que determina el valor mínimo de los productos de cada elemento de la plantilla por su correspondiente elemento del objeto, de manera que un píxel del objeto se pierde cuando en su entorno $(2m+1)(2m+1)$ hay al menos un píxel del fondo, obteniéndose de esta forma un objeto erosionado. Este operador se conoce con el nombre de **operador de erosión** y viene determinado por la siguiente expresión:

$$[f \ominus g](m, n) = \begin{cases} 1 & \text{si } f * g(m, n) = t \\ 0 & \text{si } f * g(m, n) < t \end{cases}$$

siendo t el número total de unos que tiene la plantilla utilizada. Dicho operador reduce cada objeto y elimina o filtra objetos pequeños (dependiendo del tamaño de la plantilla utilizada). La elección adecuada de la plantilla nos permite eliminar objetos con una forma determinada.

La plantilla que utilizan estos operadores se llama **elemento estructurante**, pues de su configuración y de su tamaño depende la forma y el grado de reducción o ampliación que efectúa sobre el objeto de la imagen. Además, estos operadores se llaman operadores morfológicos (nombre originado de las ciencias morfológicas para describir la forma de los objetos en biología y geología).

Los operadores de dilatación y erosión también se pueden definir como operadores sobre conjuntos de píxeles. Sea F un conjunto de píxeles con valor uno; sea G el conjunto de elementos de la plantilla con valor 1 y G_p el conjunto de píxeles del entorno del píxel p determinado por la plantilla G .

La imagen dilatada mediante el elemento estructurante G viene dada por la expresión:

$$F \oplus G = \{p: G_p \cap F \neq \emptyset\}$$

La imagen erosionada mediante el elemento estructurante G viene dada por:

$$F \ominus G = \{p: G_p \subseteq F\}$$

A continuación vamos a definir dos nuevos operadores a partir de los operadores de dilatación y erosión que son de gran interés en el análisis de la imagen binaria. Ahora pretendemos filtrar aquellos objetos pequeños de la imagen pero sin reducir el tamaño de los restantes. Hemos visto que el operador de erosión permite filtrar aquellos objetos pequeños pero reduciendo al mismo tiempo el tamaño de los restantes objetos de la imagen. Para evitar esto, basta con dilatar la imagen resultante. Así, obtenemos un nuevo operador (o) que combina las operaciones de erosión y dilatación, llamado **operador de abertura**, y que viene dado por la expresión:

$$f \circ g = (f \ominus g) \oplus g$$

Cuando se pretende rellenar los pequeños huecos o grietas en los objetos de la imagen podemos utilizar el operador de dilatación, pero éste agranda también dichos objetos por lo que a continuación se deberá aplicar el operador de erosión. Se obtiene así un nuevo operador que combina los operadores de dilatación y erosión y que se llama **operador de cierre**. Viene dado por la expresión:

$$f \bullet g = (f \oplus g) \ominus g$$

Se puede establecer la siguiente relación entre los cuatro operadores morfológicos definidos según el tamaño de la imagen que obtienen:

$$f \ominus g \subseteq f \circ g \subseteq f \subseteq f \bullet g \subseteq f \oplus g$$

Para extraer el contorno de los objetos podemos utilizar los operadores morfológicos. En primer lugar eliminamos los puntos del contorno (frontera) utilizando el operador de erosión con el elemento estructurante

$$g \equiv \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{ó} \quad g \equiv \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

El contorno del objeto se obtiene como la diferencia entre el objeto original y el objeto erosionado. Por lo tanto, el operador que nos da el *contorno de un objeto* es:

$$\text{CONTORNO}(f) = f - (f \ominus g)$$

Finalmente, vamos a definir un operador que utiliza un elemento estructurante que nos da una configuración particular de los píxeles del entorno de cada píxel de manera que cada píxel de la región es eliminado si su entorno no está completamente emparejado con el elemento estructurante. Por ejemplo, un elemento estructurante puede ser:

$$g = \begin{bmatrix} 0 & 0 & 0 \\ & 1 & \\ 1 & 1 & 1 \end{bmatrix}$$

Un píxel de la imagen es eliminado si el tono de gris de los píxeles del entorno establecido por el elemento estructurante no coinciden con los 7 valores de dicho elemento. Para que un píxel permanezca, los ceros de su entorno tienen que coincidir con los ceros del elemento estructurante y los unos con los unos. A este operador se le conoce con el nombre de *operador de acierto y fallo* (hit and miss).

Dicho operador permite identificar objetos de una forma dada. Por ejemplo, si deseamos identificar en una imagen binaria un objeto que está separado de los demás, como el definido en la figura 8.10, basta con aplicar el operador de acierto y fallo con el siguiente elemento estructurante:

$$\begin{bmatrix} 0 & 0 & 0 & & \\ 0 & 1 & 0 & & \\ 0 & 1 & 0 & & \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

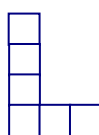


Figura 8.10. Objeto a identificar.

El *operador de acierto y fallo* también se puede expresar en términos de los operadores de erosión y dilatación. Para ello utilizamos dos elementos estructurantes disjuntos, g_1 y g_2 . El primero se utiliza con el operador erosión para determinar los píxeles cuyo entorno coincide, en los valores uno, con el elemento estructurante g_1 . El segundo se utiliza con el operador erosión aplicado al negativo de la imagen f para identificar los píxeles cuyo entorno en \bar{f} coincide, en los valores uno (cero de f), con el elemento estructurante g_2 . Por lo tanto,

$$f \otimes g = (f \ominus g_1) \cap (\bar{f} \ominus g_2) = (f \ominus g_1) \cap \overline{(f \oplus g_2)}$$

En el ejemplo anterior, utilizaríamos los dos elementos estructurantes siguientes:

$$g_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad g_2 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

La letra L se identifica teniendo en cuenta que los píxeles que la definen valen 1 o los píxeles que la rodean 0.

Por ejemplo, si consideramos los siguientes elementos estructurantes

$$g_1 \equiv \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{y} \quad g_2 \equiv \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

el operador $f \otimes g$ nos dará los píxeles del contorno derecho del objeto que forman un eje de 135° .

A partir del operador de acierto y fallo se define el operador *adelgazamiento* mediante la siguiente expresión:

$$ADELGAZAMIENTO_g(f) = f - f \otimes g$$

Nos da el conjunto de píxeles de f cuyo entorno no está emparejado con el elemento estructurante. Un elemento estructurante usado con frecuencia para el adelgazamiento es

$$\begin{bmatrix} 0 & 0 & 0 \\ & 1 & \\ 1 & 1 & 1 \end{bmatrix}$$

y los correspondientes a los giros de dicha plantilla

8.5.2 Esqueletos

Una aproximación importante para representar la forma de una región plana (objeto) es reducirla (simplificarla) a un grafo que llamaremos esqueleto de la región. Dicha reducción se puede conseguir aplicando sucesivamente un algoritmo de adelgazamiento de la región. El esqueleto de un objeto puede ser tan significativo como el propio objeto e incluso tal representación puede ser más fiable, es decir, se puede interpretar mejor y conducir a una identificación o reconocimiento del objeto óptima, sin necesidad de reconstruirlo a partir del esqueleto, como suele ocurrir en el reconocimiento de caracteres alfanuméricos, identificación de firmas o de huellas dactilares. Asimismo, cuando nuestro objetivo es el almacenamiento compacto (compresión), el esqueleto de un objeto, junto con alguna información adicional que permita reconstruir el objeto a partir de su esqueleto, suele reducir sustancialmente el número de bits necesarios para su almacenamiento, puesto que esqueleto conlleva eliminación de redundancia.

Un punto P de una región R se dice que pertenece al esqueleto de R si equidistan de, al menos, dos puntos Q_1 y Q_2 del contorno de R creando direcciones PQ_i , perpendiculares al contorno de R en cada punto Q_i , $i=1,2$. Por lo tanto, dichos puntos, Q_i , pertenecen a circunferencias con centro P . Así, se puede definir el **esqueleto** de una región de la imagen binaria como el lugar geométrico de los centros de las circunferencias maximales inscritas en dicha región. Una circunferencia maximal es aquella que no contiene circunferencias inscritas más pequeñas (las circunferencias inscritas tienen que contener dos o más puntos del contorno del objeto). Para cada punto P de la región determinamos el menor entorno (círculo) que contiene al menos un punto del contorno. Si dicho entorno contiene más de un punto del contorno entonces decimos que P forma parte de un eje *medial*, es decir, que forma parte del esqueleto del objeto (figura 8.11).

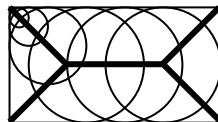


Figura 8.11. Esqueleto de una región rectangular.

Para la construcción del esqueleto vamos a utilizar un algoritmo de adelgazamiento basado en operadores morfológicos. El esqueleto viene definido por

$$S = \bigcup_{n=0}^N S_n$$

donde
$$S_n = \text{erosión}^n(f) - \text{abertura}[(\text{erosión}^n)(f)]$$

El operador $\text{erosión}^n(f)$ consiste en aplicar n veces consecutivas el operador erosión a la imagen f y N es el menor entero n tal que $\text{erosión}^{n+1}(f) = \emptyset$.

Por ejemplo, para determinar el esqueleto del rectángulo siguiente de tamaño 3×6 ,



basta con utilizar el elemento estructurante

$$\begin{bmatrix} & 1 & \\ 1 & 1 & 1 \\ & 1 & \end{bmatrix}$$

y aplicar hasta dos veces el operador erosión.

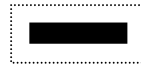
Para $n=0$ sólo hay que determinar la abertura de f ,



resultando así que S_0 viene dada por la imagen

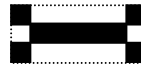


Para $n=1$ se aplica el operador erosión a f , resultando la imagen



y como su abertura es el conjunto vacío, resulta que S_2 coincide con dicha imagen.

Por lo tanto, el esqueleto del rectángulo que viene dado por $S_0 \cup S_1$ es la siguiente imagen:



8.5.3 Análisis morfológico para imágenes en tonos de gris

Los conceptos de dilatación y erosión se pueden extender para imágenes en tonos de gris. Así, se define el **operador dilatación** basado en la plantilla g , llamada *elemento estructurante*, como

$$(f \oplus g)(m, n) = \max \left\{ f(m-i, n-j) + g(i, j) \mid (i, j) \in D_g \right\}$$

donde D_g es el dominio (entorno) de g . Se supone que la imagen vale $-\infty$ fuera de su dominio. Obsérvese que esta ecuación establece un proceso similar a la convolución espacial sustituyendo cada píxel por el valor máximo de las sumas de los tonos de gris de los píxeles del entorno con los correspondientes valores de la plantilla. El dominio D_g especifica los píxeles del entorno de (m, n) a los que se les suma el valor correspondiente del elemento estructurante y sobre los que se calcula su valor máximo. Un ejemplo de dominio de g puede ser::

$$D_g = \{(i, j) : i = -r, \dots, -1, 0, 1, \dots, r; j = -s, \dots, -1, 0, 1, \dots, s\}$$

En la práctica se utilizan elementos estructurantes planos, es decir, con todos sus elementos iguales a cero, $g(i, j)=0, \forall (i, j) \in D_g$, en cuyo caso el operador se reduce al operador *máximo local* que consiste en tomar como valor del píxel el máximo de los píxeles del entorno definido por g .

Asimismo, se define el **operador erosión** basado en la plantilla g como

$$(f \ominus g)(m, n) = \min \{ f(m+i, n+j) - g(i, j) \mid (i, j) \in D_g \}$$

De manera análoga al caso binario, se definen el **operador de abertura** de f en términos de los operadores de dilatación y erosión,

$$f \circ g = (f \ominus g) \oplus g$$

Supongamos que la imagen se interpreta como una representación 3D en la que la tercera dimensión viene dada por los tonos de gris y las dos primeras por las posiciones de los píxeles sobre el plano xy . La acción del operador de abertura de f por g se puede interpretar geométricamente de la siguiente manera: aplasta un poco los pequeños picos más puntiagudos de la superficie de la imagen conforme la plantilla se traslada por todo el dominio de la misma. Por ello, permite suprimir pequeños detalles de brillo (que corresponden a zonas de tamaño inferior a la plantilla) dejando relativamente inalterados los demás tonos de gris y las zonas luminosas más grandes. Es decir, suaviza aquellas zonas pequeñas más luminosas. Análogamente, se define el **operador cierre** en términos de los operadores de dilatación y erosión como

$$f \bullet g = (f \oplus g) \ominus g$$

Dicho operador suprime los detalles oscuros que son más pequeños que la plantilla utilizada. Por lo tanto, los operadores de cierre y abertura se suelen utilizar

conjuntamente para suavizar la imagen y eliminarle ruido, aplicando primero, por ejemplo, la abertura y después el cierre. También se suele utilizar un *filtrado secuencial alternante*, es decir, combinar sucesivamente dichos operadores cambiando de elemento estructurante.

Otro operador importante es la transformación *copa del sombrero* (top-hat) de Meyer que viene dada por la expresión:

$$TopHat(f) = f - f \circ g$$

Dicho operador permite aislar objetos convexos luminosos (frente a un fondo más oscuro) del tamaño inferior a la plantilla utilizada. Es conveniente que la plantilla tenga una forma convexa.

Asimismo, la transformación *alas de sombrero* (bottom-hat) se define como el cierre menos la imagen,

$$BottomHat(f) = f \bullet g - f$$

Permite aislar objetos convexos oscuros (con respecto al fondo más luminoso) del tamaño inferior a la plantilla.

Un estimador sencillo de una superficie umbral de variación local, basado en estos operadores, es el siguiente:

$$\theta(m,n) = \frac{1}{2} [(f \oplus g) + (f \ominus g)]$$

donde g es una plantilla plana (de elementos nulos).

Los operadores morfológicos también se pueden utilizar como una técnica para aumentar el contraste local de la imagen. Una *medida del contraste local* en la ventana definida por el elemento estructurante g es la siguiente:

$$ContrasteLocal = (f \oplus g) - (f \ominus g)$$

Una transformación para conseguir aumentar el contraste local es la siguiente:

$$c(m,n) = a \frac{f(m,n) - [f \ominus g](m,n)}{[f \oplus g](m,n) - [f \ominus g](m,n)}$$

siendo a un parámetro de escala que permite ajustar la luminosidad de la imagen.

En muchas aplicaciones científicas es necesario determinar la distribución del tamaño de las partículas (objetos pequeños) que aparecen en una imagen. Ello constituye una parte importante del campo de la *granularidad*. Para ello se pueden

utilizar las técnicas morfológicas. Para partículas que son más luminosas que el fondo se aplica el operador abertura con diferentes tamaños de plantilla en orden creciente. Para cada operación de abertura se determina la suma de los valores (tonos de gris) de todos los píxeles; dicha suma se suele llamar el *área de superficie* de la imagen. Podemos representar gráficamente las área de superficies obtenidas frente al radio de los elementos estructurantes (discos). Los picos en dicha gráfica indican la presencia de muchos objetos de dicho tamaño (radio). Si la gráfica presenta muchas pequeñas variaciones se puede repetir el proceso con una imagen más suavizada.