

Transformaciones morfológicas y descripción de objetos

Extracto de “Visión por Computador. Fundamentos y métodos”. Arturo de la Escalera. Prentice Hall, 2001

Copia para el alumno con fines didácticos

8. Transformaciones morfológicas y descripción de objetos.

Una vez terminados los algoritmos de segmentación, cada objeto presente en la imagen puede ser analizado de forma individualizada. Primero habrá que resolver algún problema heredado de la segmentación como píxeles mal clasificados (como objetos cuando debían ser fondo o viceversa). Para la labor de clasificación o interpretación de la imagen interesará tanto obtener descripciones numéricas de los objetos como de la forma externa que tienen.

8.1. Transformaciones Morfológicas en imágenes binarias.

Las transformaciones morfológicas son aquellas que modifican la estructura o forma de los objetos que hay presentes en la imagen. Inicialmente estas técnicas se desarrollaron para imágenes binarias (dos niveles de gris) aunque después se extendieron los mismos conceptos a imágenes con varios niveles de gris. Estas herramientas, además de ser útiles para la extracción de características, permiten la eliminación de ruido que se produce en todo proceso de segmentación.

La morfología matemática se basa en la teoría de conjuntos y en la topología. Algunas consideraciones previas son:

- *Inclusión*. Un conjunto Y estará incluido en otro X , $Y \subset X$, si todo elemento de Y pertenece a X , $\forall y \in Y \Rightarrow y \in X$
- *Complemento*. El complemento de X , X^c , son todos los elementos que no pertenecen a X
- *Unión*. La unión de dos conjuntos X e Y , $X \cup Y$, son todos los elementos que están incluidos en uno de los dos, y se define como:

$$X \cup Y = \{x | x \in X \vee x \in Y\}$$

- *Intersección*. La intersección de dos conjuntos son los elementos que tienen comunes y se define utilizando los conceptos de unión y complemento:

$$X \cap Y = (X^c \cup Y^c)^c$$

- *Traslación*. Un conjunto X es trasladado por un vector v , cuando cada uno de los elementos de ese conjunto sufre esa translación. Se denominará al nuevo conjunto Xv .
- A cada transformación, $\Psi(X)$, se le puede asociar una *transformación dual*, $\Psi^*(X)$, que cumple:

$$\Psi^*(X) \rightarrow (\Psi(X^c))^c$$

Como ejemplo, la intersección es dual de la unión.

8.1.1. Elemento estructural, transformaciones “acierta o falla” (*hit or miss*).

Los elementos estructurales son un conjunto de puntos que servirán para determinar la estructura de un conjunto X . Uno de ellos constituirá el centro del

elemento. Como ejemplo se tienen dos elementos estructurales uno con forma de rombo y otro de circunferencia con el origen de ambos en el pixel central.

		X		
	X	X	X	
X	X	X	X	X
	X	X	X	
		X		

	X	X	X	
X	X	X	X	X
X	X	X	X	X
X	X	X	X	X
	X	X	X	

Una transformación acierta o falla, (*hit or miss* en la literatura anglosajona), se aplicará a cada punto de un conjunto X de la siguiente manera:

- Del elemento estructural B se formará un conjunto B_x que será ese elemento estructural desplazado para todo elemento x perteneciente al conjunto X. Además, B_x estará formado por dos subconjuntos B_x^1 y B_x^2 , correspondiendo el primero a los elementos del primer plano y el segundo a los del fondo. Obviamente su intersección será el conjunto vacío.
- Un punto x pertenece a la transformación acierta o falla, $X \otimes B$, si y solo si B_x^1 está incluido en X y B_x^2 está incluido en X^c :

$$X \otimes B = \{x \mid B_x^1 \subset X^1; B_x^2 \subset X^c\}$$

En otras palabras esta transformación indica donde coincide exactamente el elemento estructural B dentro del conjunto X y de ahí provienen el nombre “acierta o falla”. Se tiene por ejemplo la siguiente imagen y elemento estructural:

	1	
0	I	0

0	0	0	0	0	0
0	0	0	0	0	0
0	0	1	1	0	0
0	0	1	1	0	0
0	0	1	0	0	0
0	0	0	0	0	0

B

Imagen X

Se obtendrán los elementos estructurales para la parte del objeto y del fondo, así como el complemento del conjunto X.

1	
I	

1		1
---	--	---

1	1	1	1	1	1
1	1	1	1	1	1
1	1	0	0	1	1
1	1	0	0	1	1
1	1	0	1	1	1
1	1	1	1	1	1

B_x^1

B_x^2

Imagen X^c

Por último se realizarían las operaciones que definen a la transformación acierta o falla:

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

0	1	1	1	1	0
0	1	1	1	1	0
0	0	0	0	0	0

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

0	0	1	1	0	0
0	0	1	0	0	0
0	0	0	0	0	0

 $B_x^1 \subset X^1$

0	0	0	0	0	0
0	0	1	0	1	0
0	1	1	1	1	0

 $B_x^2 \subset X^2$

0	0	0	0	0	0
0	0	1	0	0	0
0	0	0	0	0	0

 $X \otimes B$

Comprobando que efectivamente solo aquel punto que coincide con el elemento estructural se mantiene a nivel alto. Con esto se puede pasar a la definición de las dos operaciones morfológicas principales, la erosión y la dilatación.

8.1.2. Erosiones y dilataciones.

Erosión.

Es la degradación progresiva de uno de los campos (0 ó 1). Un elemento del campo a degradar seguirá perteneciendo al mismo si está rodeado de elementos iguales a él. En caso contrario, pasará al otro campo. En un proceso iterativo terminaría por destruir la imagen. Matemáticamente se expresará de la siguiente forma: si se toma el elemento estructural simétrico respecto al origen de B , \tilde{B} , la erosión de un conjunto X respecto al elemento B es:

$$X \ominus \tilde{B} = \{x | B_x \subset X\}$$

Lo que es igual a una transformación acierta o falla donde B_x^2 es el conjunto vacío. El símbolo \ominus representa la resta de Minkowski.

Dependiendo del tamaño del elemento estructural, la erosión será más pronunciada o no. Por otro lado a veces es necesario realizar varias erosiones seguidas. Por ello se ha encontrado la relación entre el tamaño del elemento y el número de veces que se aplica:

$$d = 1 + (n - 1)N^\circ \text{ de iteraciones}$$

siendo n la dimensión del elemento inicial, y d la del elemento equivalente. Así aplicar dos erosiones de un elemento de 3×3 es igual a una de 5×5 , y tres a otra de 7×7 .

Una interpretación gráfica de la erosión puede verse en la figura 8.1. De nuevo se tienen el objeto y el elemento estructural. La erosión consistiría en pasar el elemento estructural por el borde del objeto, pasando a valer cero todos aquellos puntos por los que pase.

Dilatación.

Es el crecimiento progresivo de uno de los campos (0 ó 1). Un elemento del campo contrario a crecer será convertido si posee algún vecino perteneciente al campo que se expansiona. En caso contrario, permanecerá igual. Los elementos pertenecientes al campo a expansionar evidentemente no se modifican. En un proceso iterativo terminaría por destruir la imagen. Nótese que la erosión no es conmutativa con la dilatación, lo que más adelante se verá como una gran utilidad.

Una interpretación gráfica de la dilatación puede verse en la figura 8.1. De nuevo se tienen el objeto y el elemento estructural. La dilatación consistiría en pasar el elemento estructural por el borde del objeto, pasando a valer la unidad todos aquellos puntos por los que pase. También se presenta la dilatación mediante otro elemento estructural, que aunque obtenga un resultado parecido, presentará resultados distintos.

Matemáticamente la dilatación es la operación dual de la erosión:

$$X \oplus \tilde{B} = (X^c \ominus \tilde{B})^c$$

Algunas propiedades de la erosión y dilatación son:

$$(X \cup Y) \oplus \tilde{B} = (X \oplus \tilde{B}) \cup (Y \oplus \tilde{B})$$

$$(X \cap Y) \ominus \tilde{B} = (X \ominus \tilde{B}) \cap (Y \ominus \tilde{B})$$

$$(X \oplus \tilde{B}_1) \oplus \tilde{B}_2 = X \oplus (\tilde{B}_1 \oplus \tilde{B}_2)$$

$$(X \ominus \tilde{B}_1) \ominus \tilde{B}_2 = X \ominus (\tilde{B}_1 \oplus \tilde{B}_2)$$

8.1.3 Apertura y Cierre.

Apertura.

Las dos transformaciones antes vistas no son conmutativas. Después de una erosión no se puede recuperar la imagen original con una dilatación ya que se habrán perdido pequeños detalles. Sin embargo esto mismo tiene una gran utilidad ya que la información que se tendrá serán las partes más importantes y además se habrá eliminado pequeños elementos debidos al ruido (como se hizo referencia en el capítulo 7). La apertura (*opening* en la literatura anglosajona) se definirá entonces como una combinación de erosiones y dilataciones en las que lo primero es un erosión y la última una dilatación, siempre con el mismo elemento estructural:

$$X \circ B = (X \ominus \tilde{B}) \oplus B$$

Se denomina apertura ya que al empezar por una erosión se tiende a romper las piezas en sus partes constitutivas.

Cierre.

El cierre (*closing* en la literatura anglosajona) es la operación dual del opening por lo que su definición es:

$$X \bullet B = (X \oplus \tilde{B}) \ominus B$$

Como puede verse en la figura 8.2, la apertura suaviza los contornos del objeto, rompe enlaces delgados, y elimina pequeñas protusiones. El cierre también suaviza los contornos pero rellena fisuras, pequeños agujeros.

8.1.4 Esqueletización.

La esqueletización busca representar una región u objeto por su grafo. Es útil para la representación y es una de las características que definen un objeto. Hay que notar que pequeñas variaciones en la región afectan en gran medida a su esqueleto. Esto puede convertirse en una ventaja si lo que se propone es un control de calidad ya que será fácil detectar las piezas defectuosas.

Una manera de obtenerlo se basa en la denominada transformación del eje medial (*medial axis transformation* o MAT en la literatura anglosajona). Para cada elemento de la región se busca el punto de borde más cercano. Si más de un punto del borde está a esa distancia el elemento pertenece al eje medial que se equipara al esqueleto (figura 8.3). Gráficamente se explica mediante el siguiente ejemplo: si el objeto fuera un campo al que se pegase fuego a la vez en todos los puntos del borde, y

las llamas avanzasen a la misma velocidad, el esqueleto serían los puntos a los que llegan las llamas de dos frentes distintos al mismo tiempo.

Otro enfoque similar es sustituir cada pixel por la distancia más pequeña al borde (figura 8.4). El esqueleto vendría dado por los máximos locales.

Como puede deducirse de la definición, el coste computacional es prohibitivo. Por ello se han desarrollado otros algoritmos que consigan prácticamente el mismo fin, pero de forma más rápida. Uno de ellos es el que se basa en la sucesiva aplicación de una serie de filtros morfológicos. Para ello se aplican una serie de transformaciones acierta o falla que van disminuyendo el tamaño de las regiones. Para no llegar a destruir la imagen si se quiere encontrar el esqueleto tendrán que cumplir:

- No deben destruir los pixeles extremos.
- No deben romper la conectividad.

Así si se quiere conseguir un esqueleto con conectividad tipo 4 se aplicarán de manera sucesiva varias transformaciones acierta o falla con los siguientes elementos estructurales. Se eliminan aquellos puntos para los que la transformación de un nivel alto:

X	0	X		X	0	0		X	X	1		0	0	X
X	1	X		1	1	0		0	1	1		0	1	1
1	1	1		1	1	X		X	X	1		X	1	1
1	1	1		X	1	1		1	X	X		1	1	X
X	1	X		0	1	1		1	1	0		1	1	0
X	0	X		0	0	X		1	X	X		X	0	0

Cuando después de aplicar alguna transformación ningún pixel sea eliminado, será señal de que se ha obtenido el esqueleto. Si lo que se quiere es encontrar el esqueleto siguiendo conectividad de tipo 8 se aplicarán los elementos estructurales:

0	0	0		X	0	0		0	X	X		0	0	X
X	1	X		1	1	0		0	1	1		0	1	1
1	1	X		X	1	X		0	X	1		X	1	X
X	1	1		X	1	X		1	X	0		X	1	X
X	1	X		0	1	1		1	1	0		1	1	0
0	0	0		0	0	X		X	X	0		X	0	0
0	0	0		0	X	1		1	1	X		X	X	0
X	1	X		0	1	1		X	1	X		1	1	0
X	1	1		0	X	X		0	0	0		1	X	0

En todos ellos X indica que da igual que el pixel valga 0 ó 1. Los resultados pueden verse en la figura 8.5.

Otro algoritmo es el de Zhang-Suen. Se definen los ocho vecinos de un pixel como:

p8	p1	p2
p7		p3
p6	p5	p4

y de forma iterativa se aplican dos conjuntos de condiciones. Aquellos pixeles que cumplan todas las de una etapa son eliminados:

Etapa A.

- Si se cumple que el número de vecinos distintos de 0 es mayor o igual que dos y menor o igual que 6. Esta condición asegura que los puntos finales se preservan.
- Que solamente una vez se pasa de valor cero a uno si se recorre el borde. Se preservan los puntos que se encuentran entre los extremos.
- Que alguno de p1,p3 y p5 es cero.
- Y que alguno de p3 p5, o p7 son cero.

Etapa B.

Una vez eliminados los pixeles que cumplían estas condiciones se eliminan aquello que cumplen:

- Si se cumple que el número de vecinos distintos de 0 es mayor o igual que dos y menor o igual que 6.
- Que solamente una vez se pasa de valor cero a uno si se recorre el borde.
- Que alguno de p1,p3 y p7 es cero.
- Y que alguno de p1 p5, o p7 son cero.

Como puede verse las dos etapas coinciden en las dos primera condiciones; las dos últimas indican un giro en la búsqueda de los pixeles a eliminar.

Este algoritmo se realiza, de nuevo, de forma iterativa hasta que ningún pixel cambia de valor. Una manera de realizar este algoritmo de forma rápida es mediante una LUT. Si se convoluciona la imagen con el filtro:

128	1	2
64		4
32	16	8

los ocho vecinos han quedado codificados en un número entre cero y 255. Puede calcularse de antemano los valores numéricos que darían las condiciones expresadas antes. Por ello se deben eliminar los puntos que cumplen:

S1= {3,6,7,12,14,15,24,28,30,48,56,60,62,96,
112,120,129,131,135,143,192,193,195,199
207,224,225,227,231,240,241,243,248,249}

S2= {3,6,7,12,14,15,24,28,30,31,48,56,60,62,
63,96, 112,120,124,126,129,131,135,143,
159,192,193,195,224,225,227,240,248,252}

representado S1 los posibles casos que cumplen el primer conjunto de condiciones y S2 el segundo. Así por ejemplo el valor 207 corresponde a la convolución con una imagen:

1	1	1
1		1
0	0	1

Donde el número de vecinos distintos de 0 es dos, solamente una vez se pasa de valor cero a uno si se recorre el borde y p5 es cero. Los resultados se tienen en la figura 8.6.

El esqueleto obtenido es distinto en los tres casos. Dependerá de la implementación concreta el que interese más uno u otro dependiendo de la rapidez o de lo aproximado

que sea al caso concreto. De todas maneras los tres representan la forma del objeto que es lo que se busca con la esqueletización.

8.1.5. Adelgazamiento.

El adelgazamiento es similar a la esqueletización pero sin llegar hasta el extremo. Suele aplicarse cuando después de una detección de bordes se quiere adelgazar la respuesta al operador gradiente (figura 8.7). Debe cumplir las mismas condiciones de la obtención del esqueleto para no destruir la imagen.

8.1.6. Extracción del perímetro.

Una de las características que definen los objetos es su perímetro (figura 8.8). El método exacto para la extracción del mismo consiste en buscar un pixel perteneciente al borde del objeto e ir siguiendo la forma de este encontrando los sucesivos vecinos mientras se cuenta. El gran inconveniente es su coste computacional. La morfología presenta una forma más rápida pero inexacta para realizar obtener este parámetro. Así basta hacer restar a la imagen original el resultado de erosionarla:

$$P(A) = A - (A \ominus B)$$

y contar los pixeles que están a nivel alto, con lo que se obtiene el denominado perímetro interior; o restar a la imagen dilatada la original:

$$P(A) = (A \oplus B) - A$$

para obtener el perímetro exterior.

8.1.7. Eliminación de ruido.

Después de una segmentación se tomaran pixeles como pertenecientes al objeto que no pertenecen a él y viceversa. Para rellenar esos huecos o eliminar esos pixeles se realizan combinaciones de erosiones y dilataciones. Así se tendrá:

- Eliminación de pixeles que no pertenecen al objeto (figura 8.9-b): $A_1 = A \ominus B$
- Para recuperar la forma inicial (figura 8.9-c): $A_2 = A_1 \oplus B$
- Para rellenar huecos en el objeto (figura 8.9-d): $A_3 = A_2 \ominus B$
- Para recuperar la forma inicial (figura 8.9-e): $A_4 = A_3 \oplus B$

Luego todo junto queda:

$$A_4 = (((A \ominus B) \oplus B) \oplus B) \ominus B$$

o lo que es lo mismo un opening seguido de un closing:

$$A_4 = (A \circ B) \bullet B$$

Un ejemplo puede verse en la figura 8.10 donde se tiene el resultado de la umbralización recursiva vista en el capítulo anterior, en donde no se había podido realizar bien la segmentación de la parte más clara. Con las transformaciones morfológicas es ahora posible.

8.1.8. Cerco convexo.

Si rodeásemos un objeto con una goma elástica obtendríamos el cerco convexo (*convex-hull* en la literatura anglosajona). Es la región más pequeña que contiene al objeto de tal forma que dos puntos cualesquiera pertenecientes a él pueden conectarse a través de una línea recta, perteneciendo a la región todos los puntos de ella. Para obtenerlo se utiliza un método parecido a la esqueletización, solo que ahora se añaden nuevos puntos a la imagen original. Así se realizarían las siguientes transformaciones acierta o falla de forma sucesiva e iterativa:

1	1	1	1	1	X	1	X	X	X	X	X
1	0	X	1	0	X	1	0	X	1	0	X
X	X	X	1	X	X	1	1	X	1	1	1
X	X	X	X	X	1	X	1	1	1	1	1
X	0	1	X	0	1	X	0	1	X	0	1
1	1	1	X	1	1	X	X	1	X	X	X

En la figura 8.11 pueden verse los resultados. Las diferencia con la imagen inicial pueden ser de dos tipos: si se trataba de un agujero interno se denominan lagos y al resto bahías. Su número y distribución pueden ser características que luego ayuden a clasificar o diferenciar los objetos presentes en la imagen.

8.1.9 Segmentación mediante el operador distancia.

Supóngase que se pidiese contar el número de objetos presentes en la figura 8.12-a. Con lo visto hasta ahora se obtendría el histograma y se realizaría una umbralización (figura 8.12-b) seguida de un etiquetado (figura 8.12-c). Sin embargo, debido a que los objetos están en contacto, se le asigna la misma etiqueta a varios de ellos. Una posible solución es aplicar el operador distancia (figura 8.12-d), para determinar más tarde la distancia mínima al borde (o diámetro mínimo del objeto). Como puede verse ahora en la figura 8.12-e los objetos está ahora separados por lo que si se aplica un etiquetado (figura 8.12-f) cada objeto recibe una etiqueta distinta.

8.1.10 Eliminación de ramas.

Si se considera a la imagen resultante de una esqueletización como si fuera un árbol, se podría considerar que se tiene un tronco principal del que salen unas ramas. Algunas de estas ramas dificultan el que pueda reconocerse el objeto analizando el esqueleto ya que no aportan información sustancial acerca de él. Puede ser interesante por tanto “podar” (*pruning* en ingles) el esqueleto del objeto. Esto puede observarse en la figura 8.13 donde se tienen un par de letras con su esqueleto. Para eliminar las ramas se pasarían los ocho elementos estructurales:

0	0	0	0	0	0	0	0	X	0	0	0
0	1	0	0	1	0	0	1	1	0	1	0
X	1	X	1	0	0	0	0	X	0	0	1

$$\begin{array}{cccccc}
X & 1 & X & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{array}
\quad
\begin{array}{cccccc}
X & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 \\
X & 0 & 0 & 0 & 0 & 0
\end{array}$$

con los que se conseguiría reducir en un pixel el tamaño de ellas. Depende el caso concreto el que esta operación se realice un número fijo de veces.

8.1.11 Localización de elementos.

La morfología también sirve para localizar la posición de un elemento determinado (figura 8.14). Se tiene la imagen I, donde está el elemento A. A partir de ellas se obtendrá la imagen I^c , que es la inversa de la anterior y el elemento B que es X-A, siendo X una imagen toda ella a valor alto de dimensiones un poco más grandes que A. A partir de esto los pasos son:

- Erosionar I con A: $I \ominus B$
- Erosionar I^c con X: $I^c \ominus X$
- Obtener la intersección de ambas imágenes: $(I \ominus B) \cap (I^c \ominus X)$

con lo que se detecta la presencia del objeto. En el fondo es realizar la transformación hit o miss a la imagen I.

8.1.12 Transformaciones morfológicas en imágenes no binarias.

La idea es extender lo visto anteriormente a imágenes que tengan más de dos niveles de gris. Se define ahora la dilatación de una imagen $f(x,y)$ con un elemento estructural $b(x,y)$ a la operación:

$$f(x, y) \oplus b(x, y) = \max\{f(x-i, y-j) + b(i, j) \mid (x-i, y-j) \in Df; (i, j) \in Db\}$$

donde Df es el dominio de la imagen $f(x,y)$ y Db el del elemento estructural $b(x,y)$. Con los dominios lo que se asegura es que las dos imágenes coincidan en algún pixel. Como la dilatación elige el máximo de la suma entorno a un punto, si todos los valores del elemento estructural son positivos la imagen resultante es más clara que la original. (figura 8.15).

De forma análoga, la fórmula de la erosión es (figura 8.16):

$$f(x, y) \ominus b(x, y) = \min\{f(x-i, y-j) - b(i, j) \mid (x-i, y-j) \in Df; (i, j) \in Db\}$$

Nótese que mientras la dilatación es una operación conmutativa la erosión no lo es.

Las fórmulas de la apertura y cierre son las mismas que para imágenes binarias (figura 8.17).

$$\begin{aligned}
X \circ B &= (X \ominus \tilde{B}) \oplus B \\
X \bullet B &= (X \oplus \tilde{B}) \ominus B
\end{aligned}$$

Si se toma la imagen como un mapa tridimensional donde el nivel de gris indica la altura el efecto de realizar una apertura es desplazar el elemento estructural por debajo del terreno siendo el valor del opening el máximo entre los puntos que toca el elemento estructural. El cierre es desplazar el elemento por encima de la superficie y tomar su mínimo valor.

Aplicaciones de la morfología para imágenes con varios niveles de gris son (figura 8.18):

Suavizado de la imagen.

Ello se consigue realizando un opening seguido de un closing. Se eliminan así los píxeles o muy claros o muy oscuros.

Gradiente morfológico.

Para detectar los bordes se resta a la imagen obtenida con una dilatación la obtenida por una erosión.

$$G = (X \oplus B) - (X \ominus B)$$

Depende menos de la dirección de los bordes que otros métodos como Sobel

Transformación Top-Hat

La transformación Top-Hat se define:

$$TH = X - (X \circ B)$$

consiguiéndose un filtro paso alto de la imagen.

8.2 Características de regiones.

Una vez que los objetos están separados y se han corregido los errores cometidos en la segmentación, se está en disposición de obtener una serie de valores numéricos que describan los objetos. Este segundo grupo de características, junto con el que se utilizó en la segmentación ayudará a la última etapa para asignar al objeto detectado a uno de los posibles casos que se conocen.

8.2.1. Tamaño, perímetro, compactidad.

El tamaño de un objeto vendrá dado por la suma de sus píxeles:

$$A = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} g(i, j)$$

siendo $g(x,y)$ la imagen binaria.

El perímetro lo constituiría el número de píxeles exteriores al objeto y la compactidad viene definida por la fórmula matemática:

$$\frac{P^2}{A}$$

Fórmula cuyo valor mínimo corresponde al círculo con un valor 4π .

Una variación es la rectangularidad que se define como el cociente entre el área del objeto y la del rectángulo más pequeño que lo engloba. El valor máximo sería la unidad para un objeto rectangular y $\frac{\pi}{4}$ para un círculo.

8.2.2. Posición.

El conocimiento de la posición de un objeto en una imagen es importante ya que en la mayoría de las aplicaciones industriales los objetos se encuentran situados sobre una superficie plana (una mesa, una cinta transportadora) y se conocen los parámetros de la cámara: tamaño del CCD, distancia focal, y la distancia y orientación de la cámara.

respecto a las superficie sobre la que se encuentran los objetos, por lo que a partir de la posición en la imagen puede obtenerse las coordenadas del objeto que pueden ser utilizadas por otros elementos del sistema automático como un robot (tal como se explicó en el segundo capítulo). Usualmente se obtiene el centro de gravedad del objeto, definido por:

$$\bar{x} = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} i g(i, j)}{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} g(i, j)}$$

$$\bar{y} = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} j g(i, j)}{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} g(i, j)}$$

por lo que:

$$\bar{x} = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} i g(i, j)}{A}$$

$$\bar{y} = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} j g(i, j)}{A}$$

8.2.3. Orientación.

Para que un objeto tenga una orientación debe ser alargado (un círculo no tiene orientación). Un aspecto adicional es definir qué se entiende por orientación de un objeto. Aquí se entenderá que aquel eje que minimice la distancia a él de todos los puntos de la imagen será el que determine la orientación del objeto. En el caso de objetos simétricos (figura 8.19) no existirá ningún problema, pero si la pieza no lo es este eje no coincidirá con la dirección a la que apunta el objeto o la parte por la que el robot agarrará la pieza. Sin embargo al ser la diferencia un ángulo constante se podrán realizar las correcciones que sean necesarias.

Puede demostrarse que dicha recta es:

$$\rho = x \cos \theta + y \sin \theta$$

Siendo los valores de ρ y θ :

$$\rho = \bar{x} \cos \theta + \bar{y} \sin \theta$$

$$\tan 2\theta = \frac{b}{a - c}$$

y los valores de a, b, c son los momentos de segundo orden:

$$a = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (x_{ij} - \bar{x})^2 g(i, j)$$

$$b = 2 \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (x_{ij} - \bar{x})(y_{ij} - \bar{y}) g(i, j)$$

$$c = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (y_{ij} - \bar{y})^2 g(i, j)$$

8.2.4. Proyecciones.

Ya se habló en su momento de los histogramas laterales (figura 8.20). Estos son el caso concreto de la proyección de una imagen sobre una línea vertical u horizontal. En general la proyección de una imagen sobre una recta se obtiene viendo para cada punto de ella el número de píxeles que pertenecen a la perpendicular que pasa por ese punto. Conociendo la proyección horizontal y vertical puede conocerse el área y la posición del objeto:

$$A = \sum_{i=0}^{N-1} H(i) = \sum_{j=0}^{N-1} V(i)$$

Para obtener la posición se utilizan las mismas fórmulas vistas anteriormente; la recta vendrá definida por los valores de ρ y θ :

$$\rho = \bar{x} \cos \theta + \bar{y} \sin \theta$$

$$\tan 2\theta = \frac{b}{a - c}$$

solo que ahora se calculan de otra forma:

$$\bar{x} = \frac{\sum_{i=0}^{N-1} y_i H(i)}{A}$$

$$\bar{y} = \frac{\sum_{j=0}^{N-1} x_j V(i)}{A}$$

$$a = \sum_{i=0}^{N-1} x^2 H(i)$$

$$c = \sum_{j=0}^{N-1} y^2 V(j)$$

Para obtener el parámetro b se utiliza la proyección diagonal y puede demostrarse que:

$$b = 2 \sum_{t=0}^k t^2 P(t) - \frac{1}{2} \sum_{i=0}^{N-1} x^2 H(i) - \frac{1}{2} \sum_{j=0}^{N-1} y^2 V(j)$$

siendo p(t) la proyección oblicua.

8.2.5. Momentos invariantes a la rotación la escala y la traslación.

Con estas características se pretende poder reconocer los objetos aunque no se encuentren siempre en la misma posición, estén girados o su tamaño sea distinto. Para ello se parte de los momentos ordinarios:

$$m_{uv} = \sum \sum g(x, y) x^u y^v$$

Para conseguir la invarianza a la posición se les resta a las coordenadas de cada punto el del centro de gravedad:

$$\mu_{uv} = \sum \sum g(x, y) (x - \bar{x})^u (y - \bar{y})^v$$

Siendo los valores, hasta los momentos de tercer orden:

$$\mu_{00} = m_{00}$$

$$\mu_{10} = 0$$

$$\begin{aligned}
\mu_{01} &= 0 \\
\mu_{20} &= m_{20} - \bar{x}m_{10} \\
\mu_{02} &= m_{02} - \bar{y}m_{01} \\
\mu_{11} &= m_{11} - \bar{y}m_{10} \\
\mu_{30} &= m_{30} - 3\bar{x}m_{20} + 2\bar{x}^2m_{10} \\
\mu_{03} &= m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2m_{01} \\
\mu_{12} &= m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}^2m_{10} \\
\mu_{21} &= m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}^2m_{01}
\end{aligned}$$

El siguiente paso sería obtener momentos invariantes a la rotación y a la escala. Para ello se definen primero los momentos centrales normalizados:

$$\eta_{ij} = \frac{\mu_{ij}}{(\mu_{00})^k}$$

siendo $k = \frac{i+j}{2} + 1$ y $i+j \geq 2$. Al dividir cada momento por el área (momento de orden cero) se consigue que sean valores independientes de la escala. Para obtener valores independientes de la rotación las operaciones matemáticas son más complicadas. En el artículo de la bibliografía realizado por Hu puede encontrarse que los valores:

$$\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})\{(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\} \\
&\quad + (2\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})\{3(\eta_{03} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\} \\
\phi_6 &= (\eta_{20} - \eta_{02})\{(\eta_{30} + \eta_{12}) - (\eta_{21} + \eta_{03})^2\} + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})\{(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\} \\
&\quad - (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})\{3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\}
\end{aligned}$$

son invariantes a la rotación, la escala y la traslación. Debido a su utilidad son conocidos como *los siete magníficos de Hu*. A pesar de todas las operaciones necesarias para obtenerlos merece la pena recordar que los parámetros iniciales se obtiene mediante las proyecciones por lo que pueden obtenerse con gran rapidez.

Los valores para los objetos representados en la figura 8.21 (ver el libro de Vernon para más detalles) serían:

	R1	R2	R3	R4	R5	R6
ϕ_1	1.67 E-1	1.94 E-1	2.08 E-1	1.67 E-1	1.94 E-1	1.94 E-1
ϕ_2	0.00 E+0	6.53 E-3	1.56 E+0	0.00 E+0	6.53 E-3	6.53 E-3
ϕ_3	0.00 E+0	1.02 E-3	0.00 E+0	0.00 E+0	1.02 E-3	1.02 E-3
ϕ_4	0.00 E+0	4.56 E-5	0.00 E+0	0.00 E+0	4.56 E-5	4.56 E-5
ϕ_5	0.00 E+0	4.25 E-9	0.00 E+0	0.00 E+0	4.25 E-9	4.25 E-9
ϕ_6	0.00 E+0	1.70 E-6	0.00 E+0	0.00 E+0	1.70 E-6	1.70 E-6
ϕ_7	0.00 E+0	-8.85 E-9	0.00 E+0	0.00 E+0	-8.85 E-9	-8.85 E-9

8.2.6. Descriptores topológicos.

Los descriptores topológicos no tratan de dar un número exacto sino indicar alguna idea sobre la forma del objeto.

- Número de agujeros en el objeto.
- Número de componentes conectados. Son aquellos elementos separados que forman un objeto.
- Número de Euler: es la diferencia entre los dos anteriores. Así el número de Euler de:

A

E=0

B

E=-1

i

E=2

8.2.7. Códigos encadenados.

Consiste en asociar un código al camino recorrido cuando se va siguiendo los píxeles del borde que están conectados. Permite obtener el perímetro, área y momentos del objeto. A partir de un píxel determinado se asocia un número dependiendo de la dirección del vecino:

3	2	1
4		0
5	6	7

El código encadenado lo formará una sucesión de números entre 0 y 7. Los problemas que lleva son por un lado que los códigos suelen ser largos y que errores debidos al ruido perturban en gran manera al código. Una solución es disminuir la resolución. Otro posible problema es que depende del punto inicial.

La función de correlación usada es:

$$\rho(j) = \frac{1}{m} \sum_{i=1}^m \cos\left(\frac{\pi}{4} (c(i) - (c'(i+j) \bmod(n)))\right)$$

donde:

- $c(i)$ es el código del objeto buscado.
- m es el número de elementos de código del objeto buscado.
- $c'(i)$ es el código del objeto presente en la imagen.
- n es el número de elementos de código del objeto presente en la imagen.
- j es la desviación respecto al código inicial.

El valor de j para cuyo $\rho(j)$ sea más próximo a 1 dará la mayor correlación.

8.2.8. Signatura, energía.

La signatura consiste en obtener la curva que represente la distancia al centroide del objeto de todos los puntos de su borde (figura 8.22). Las signaturas son invariantes a la traslación pero no a la rotación o a la escala. Para eliminar la influencia de la rotación puede elegirse siempre para empezar, o el punto más alejado del centroide o el eje principal que se estudió antes.

Otra representación de la forma del objeto viene dada por su energía. Para cada punto del borde se calcula cual es el círculo tangente a él y se obtiene su radio. Tomando la inversa se calcula la curvatura de cada punto del borde:

$$k(i) = \frac{1}{r(i)}$$

Se define entonces como energía de la forma del objeto:

$$E = \frac{1}{n} \sum_{i=1}^n |k(i)|^2$$

siendo n el número de puntos. El valor mínimo el correspondiente a los círculos.

8.2.9. Descriptores de Fourier.

Como en el caso anterior se trata de pasar de tener una información bidimensional a otra unidimensional. Para ello se siguen los siguientes pasos:

- Se elige un origen de coordenadas arbitrario.
- Se toma un punto al azar como primer elemento de una sucesión de puntos tomados en sentido antihorario.
- Estas parejas de puntos se transforman a números complejos:

$$h(k) = x_k + jy_k$$

- Se obtiene la transformada de Fourier:

$$H(\omega) = \frac{1}{N} \sum_{k=0}^{N-1} h(k) e^{-j2\pi\omega k/N}$$

El módulo de las partes reales e imaginarias es invariante a la traslación y a la rotación de los objetos. Si se toma la relación de ellos con el primer coeficiente se tiene invariancia a la escala. Entre 10 y 15 descriptores son suficientes para definir cualquier forma (figura 8.23).

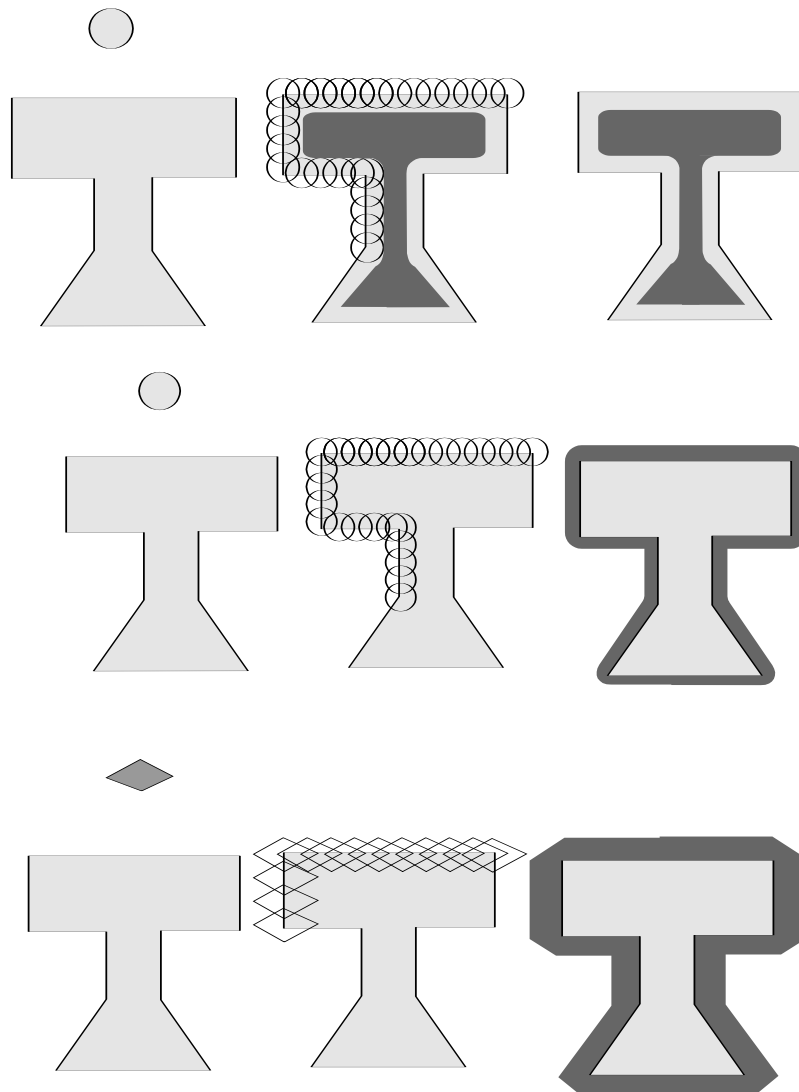


Figura 8.1. Erosión y dilatación.

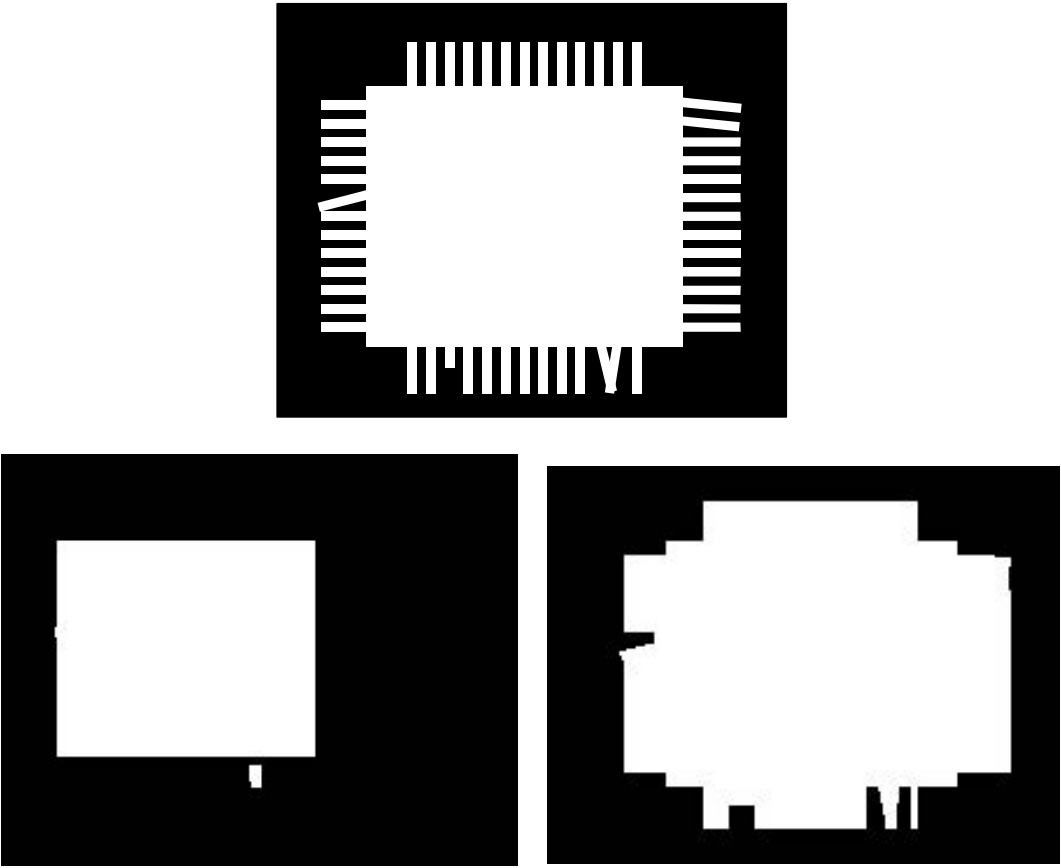


Figura 8.2. Apertura y cierre.

										0		0					
											1						
											1						
											1						
											1						
											1					0	
											1	1	1	1	1	1	0
										0						0	

Figura 8.3 Medial Axis Transform

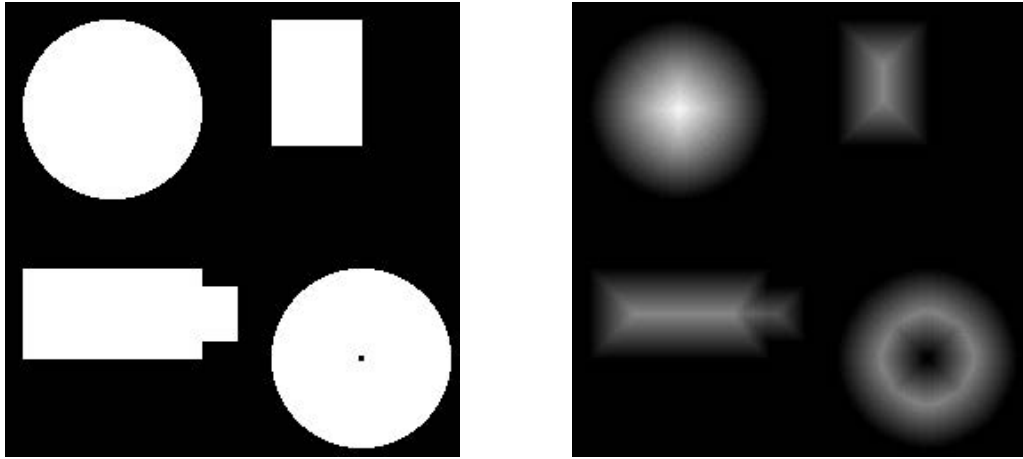
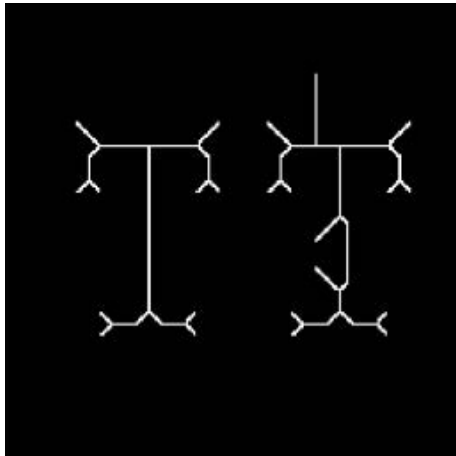


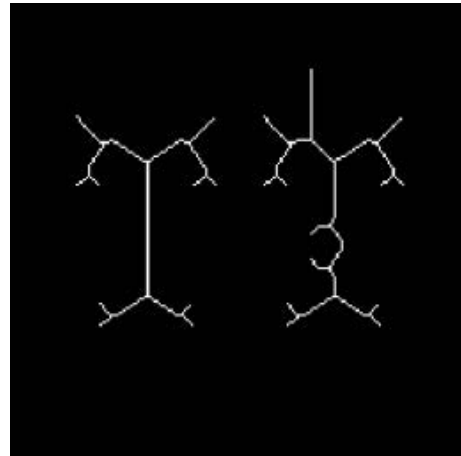
Figura 8.4 Distancias



(a)



(b)



(c)

Figura 8.5 Esqueletización. (a) Imagen original (b) Esqueleto con conectividad 4 (c) Esqueleto con conectividad 8

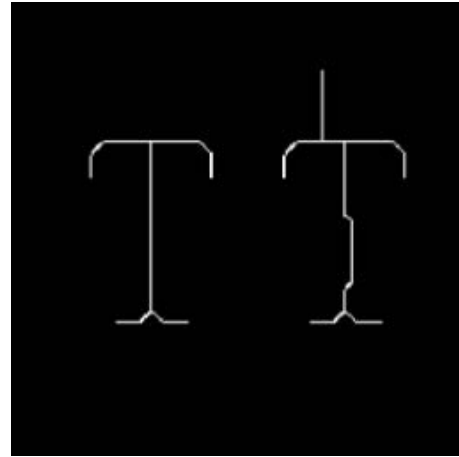
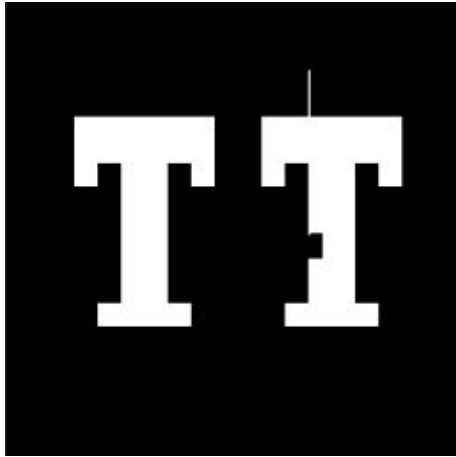


Figura 8.6. Método de Zhang Suen para obtener el esqueleto



Figura 8.7. Adelgazamiento.

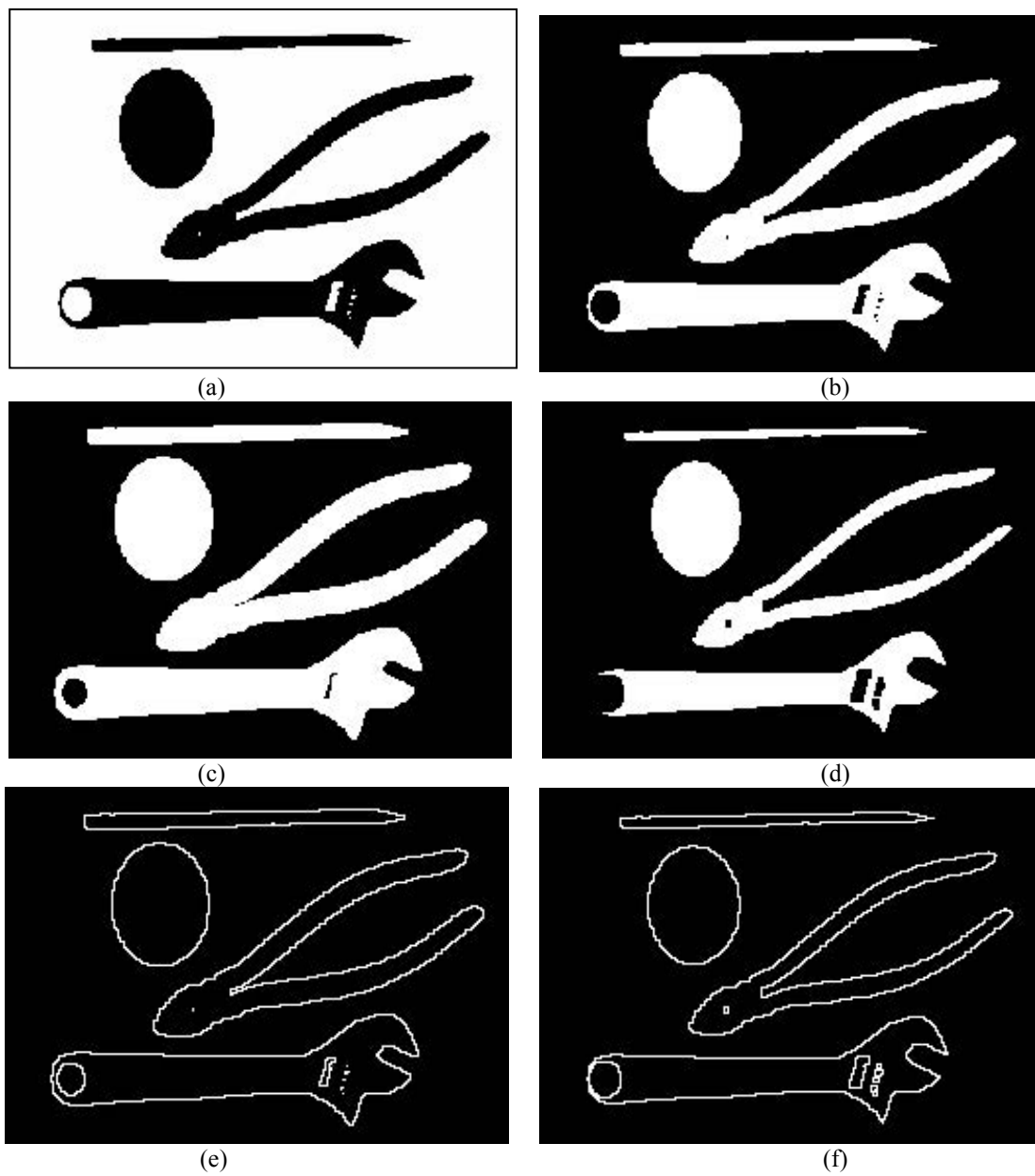


Figura 8.8. Extracción del perímetro. (a) Imagen a contraluz (b) Imagen binarizada (c) Imagen dilatada (d) Imagen erosionada (e) Perímetro exterior (f) Perímetro interior

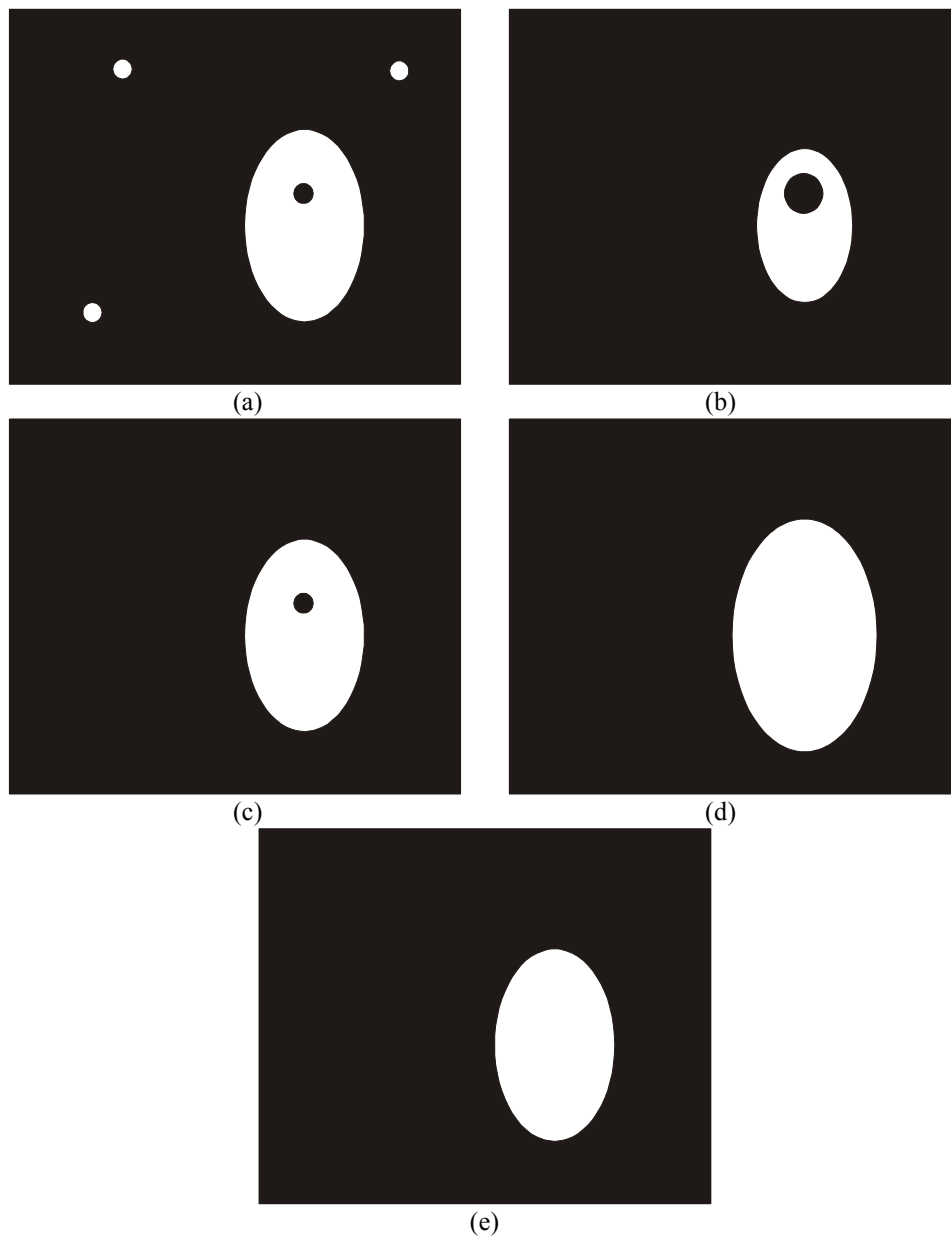


Figura 8.9. Eliminación de ruido.

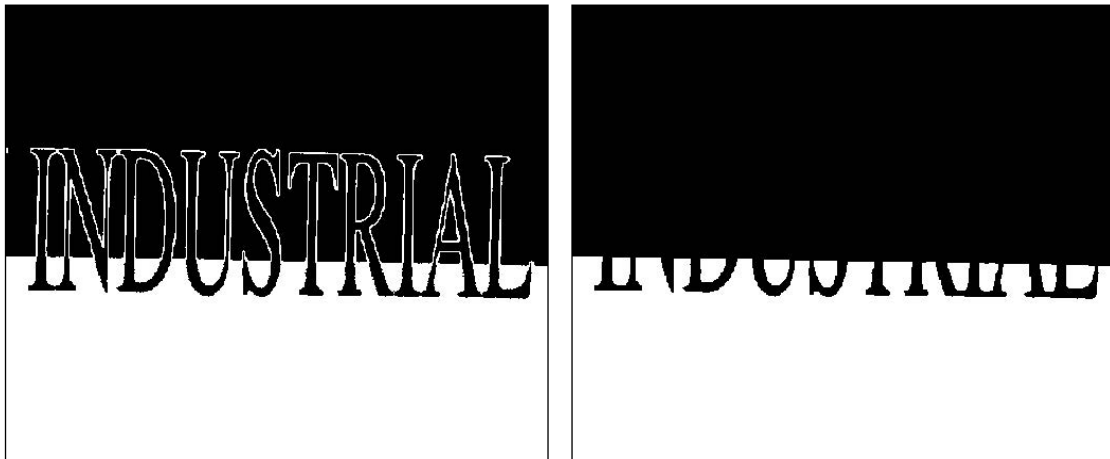


Figura 8.10. Eliminación de ruido.

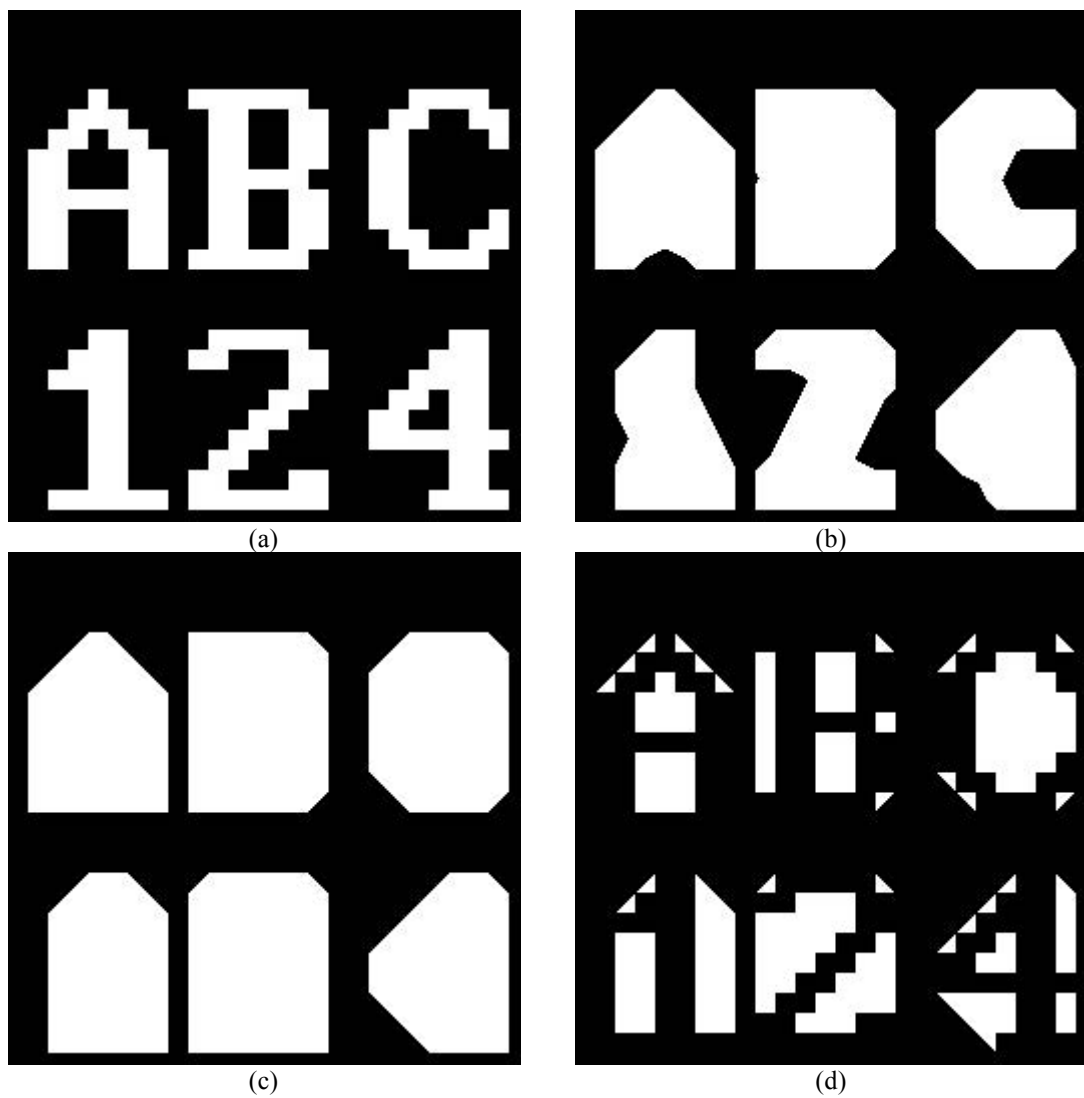


Figura 8.11. Convex-hull. (a) Imagen original (b) Crecimiento de los elementos (c) Convex-hull (d) Lagos y bahías

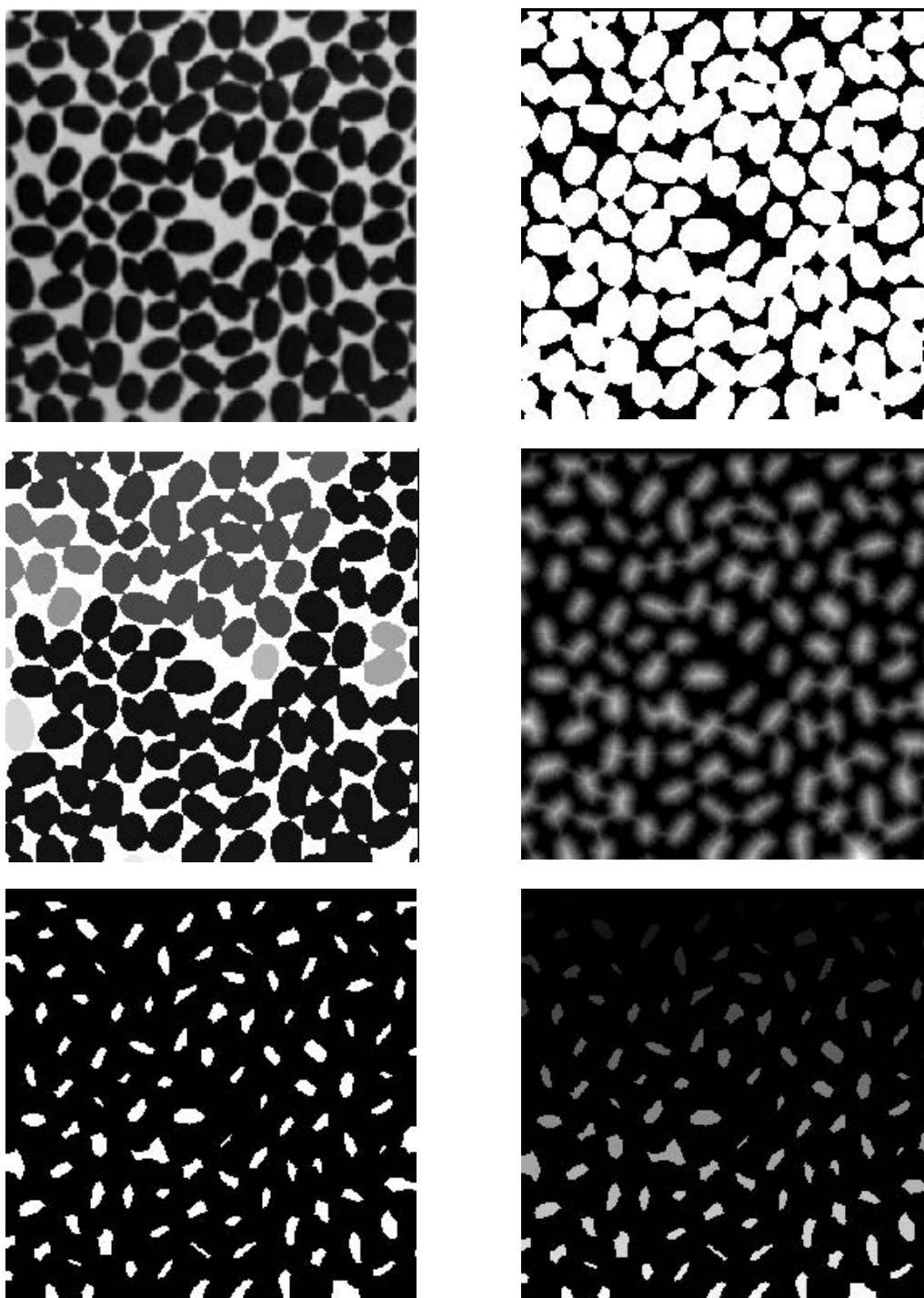


Figura 8.12. Segmentación de objetos mediante el operador distancia. (a) Imagen original (b) Imagen binarizada (c) Imagen etiquetada (d) Operador distancia (e) Umbralización de la distancia (f) Etiquetado de la distancia

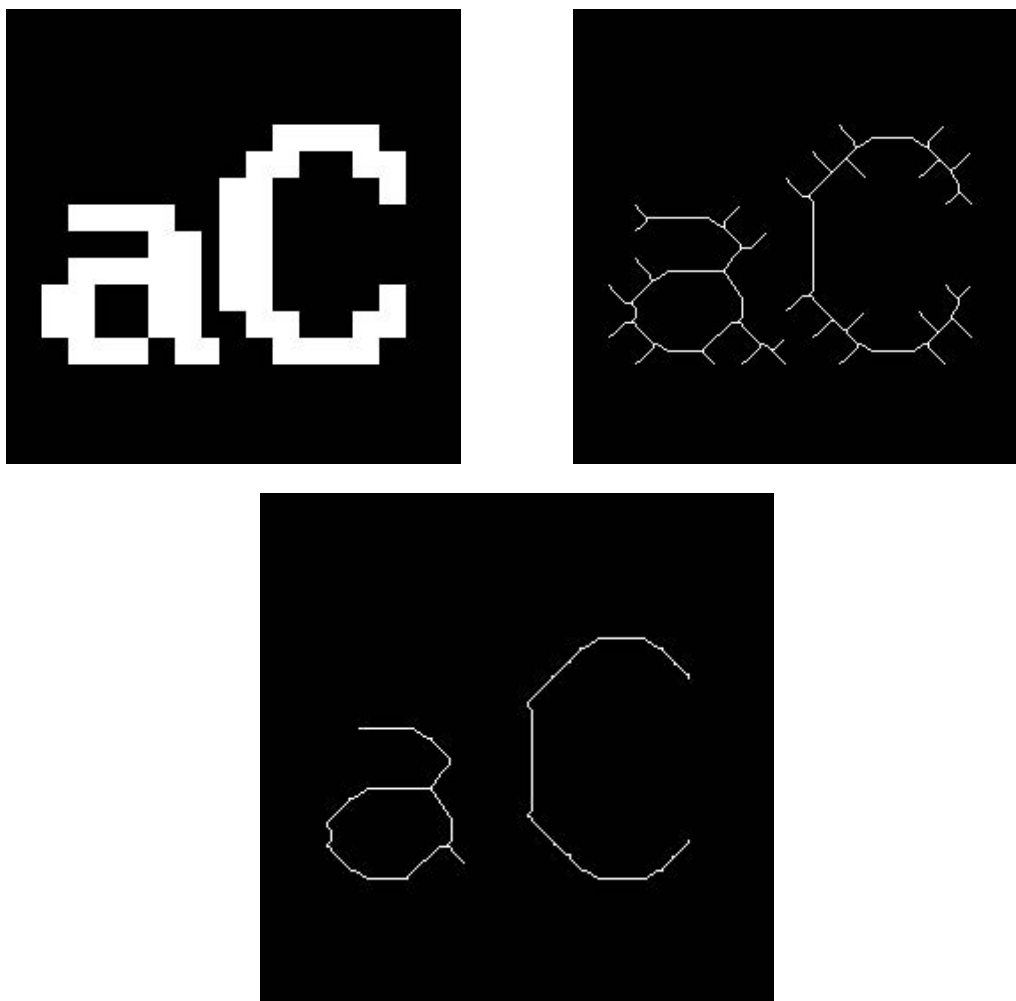


Figura 8.13. Pruning.

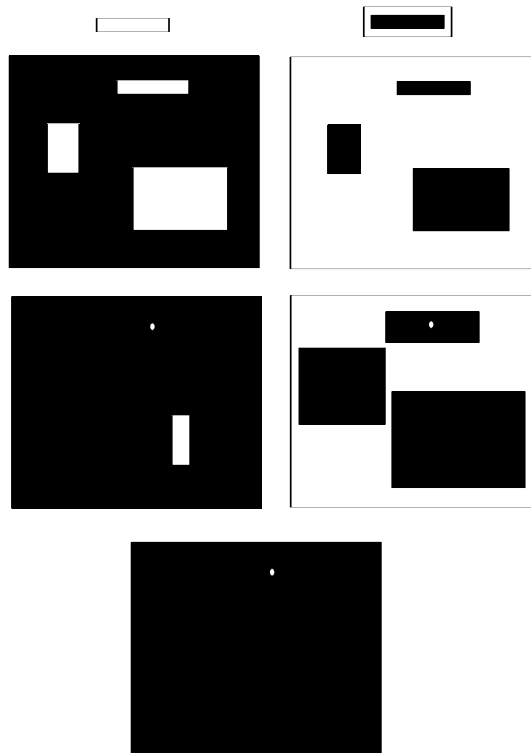


Figura 8.14. Determinación de la posición de un elemento.

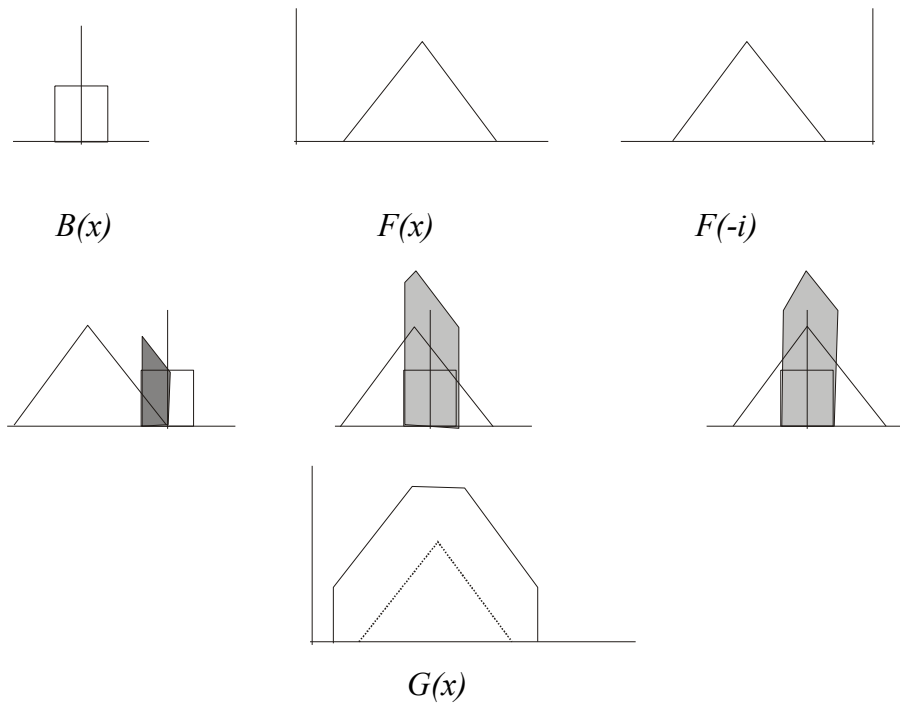


Figura 8.15. Dilatación de una imagen.

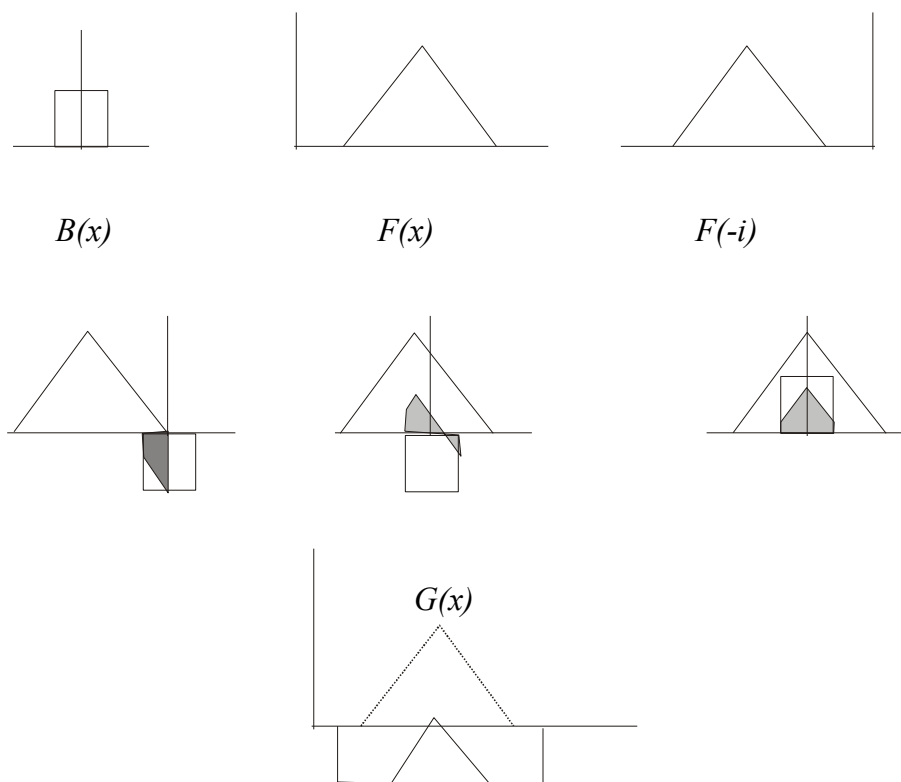


Figura 8.16. Erosión de una imagen.

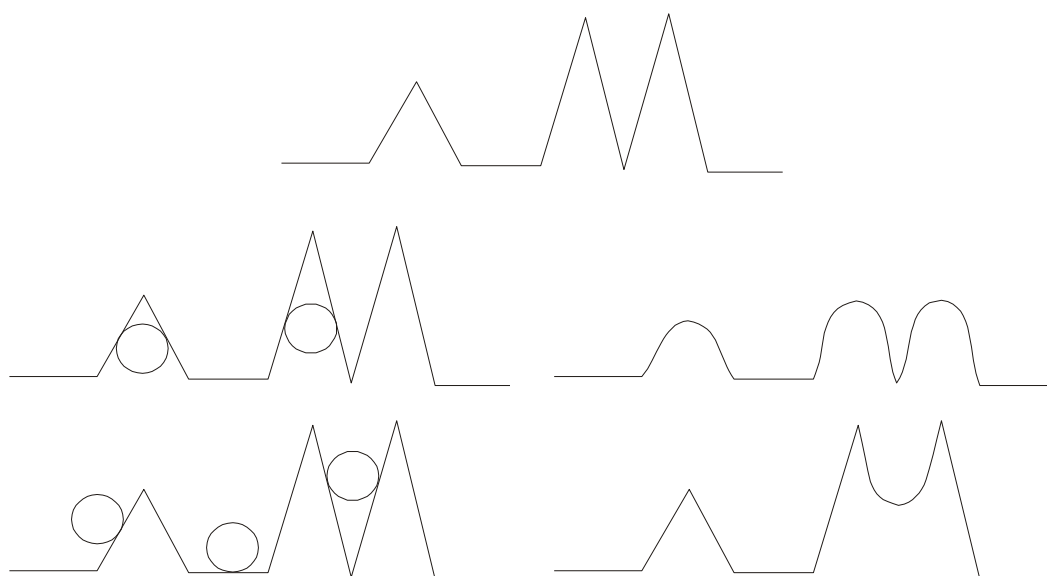
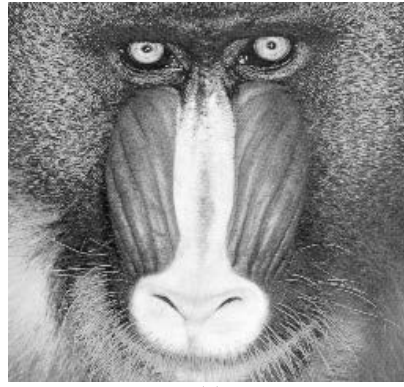


Figura 8.17. Openig y closing



(a)



(b)



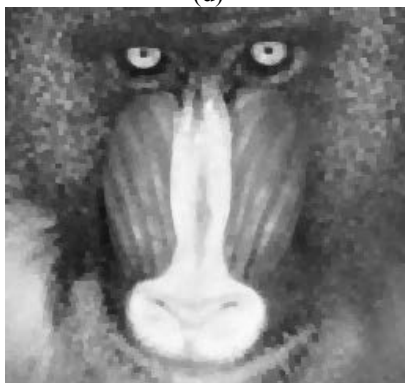
(c)



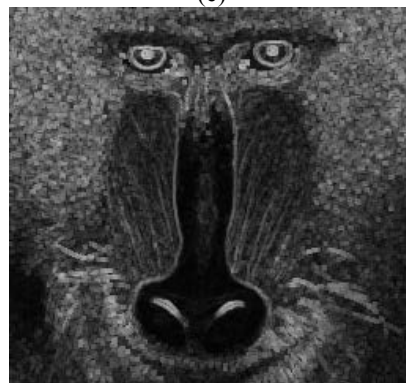
(d)



(e)



(f)



(g)

Figura 8.18: (a) Imagen original (b) dilatación (c) erosión (d) opening (e) closing (f) suavizado (g) detección de bordes

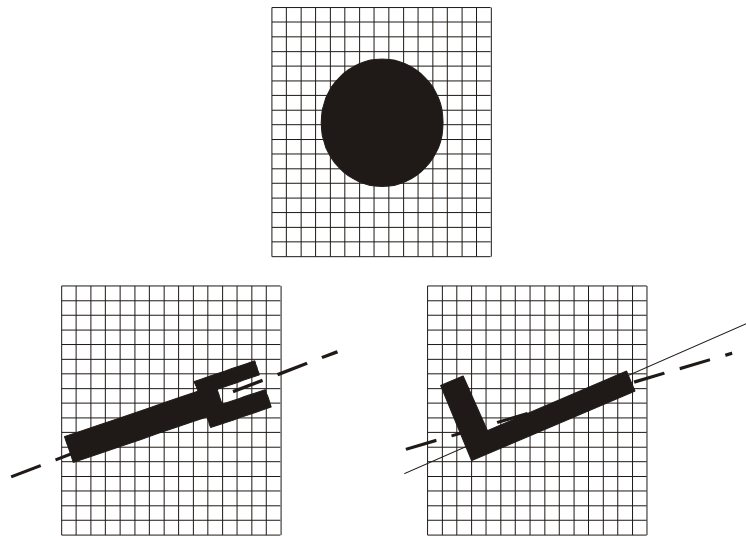


Figura 8.19. Orientación de imágenes binarias.

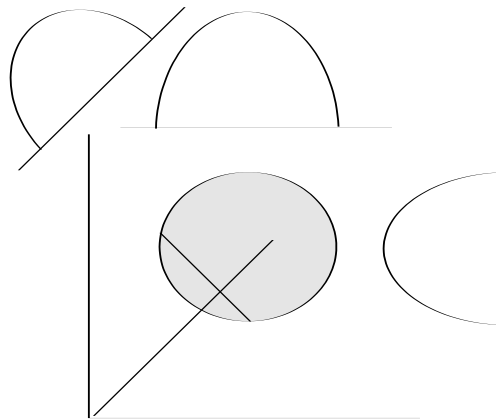


Figura 8.20. Proyecciones.

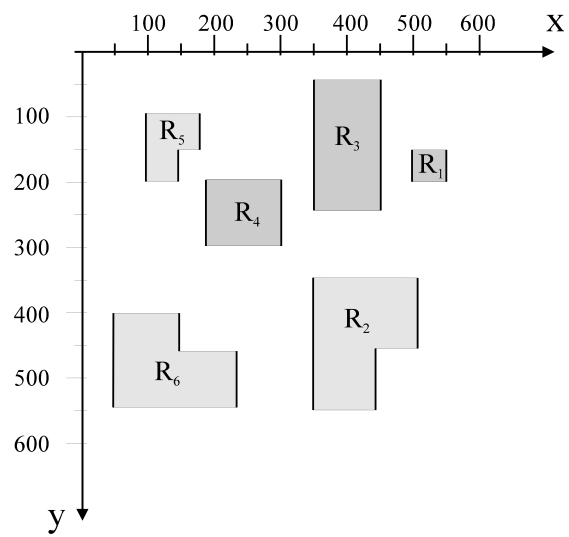


Figura 8.21. Momentos invariantes

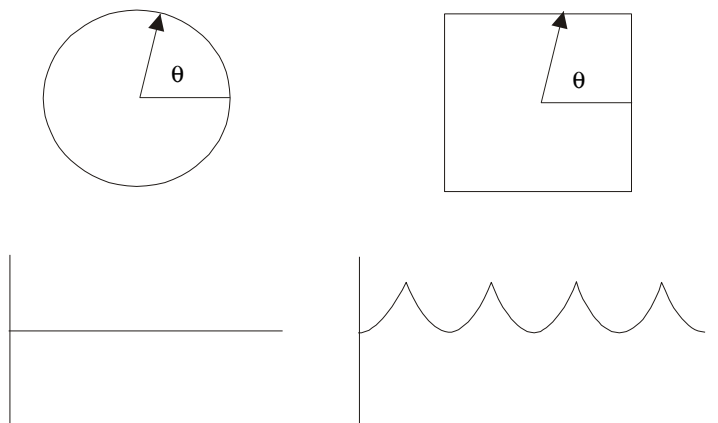


Figura 8.22. Signatura

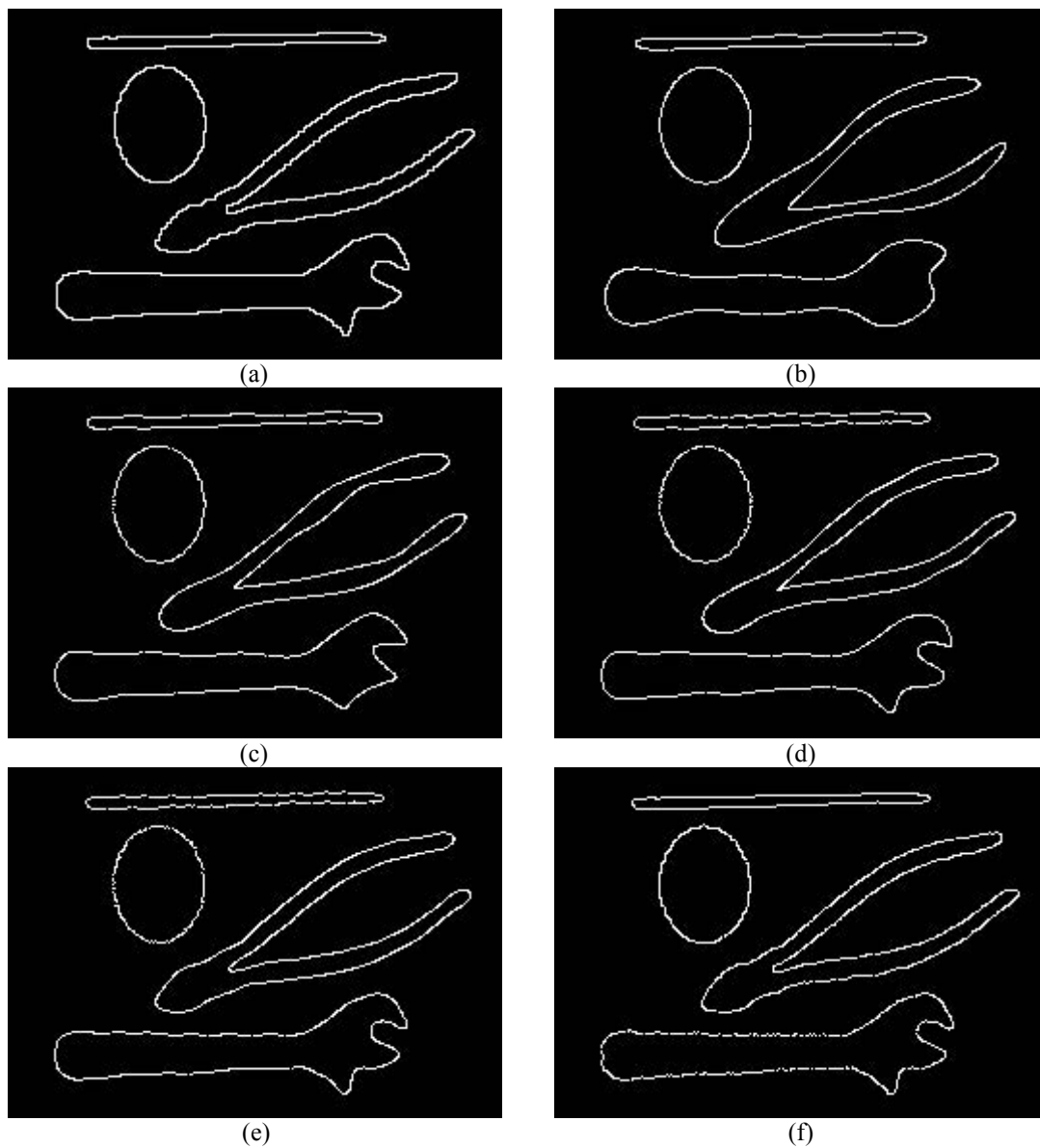


Figura 8.23. Descriptores de Fourier. (a) imagen original (b) resultado de usar 10 descriptores de Fourier (c) 20 descriptores (d) 30 (e) 50 (f) 100