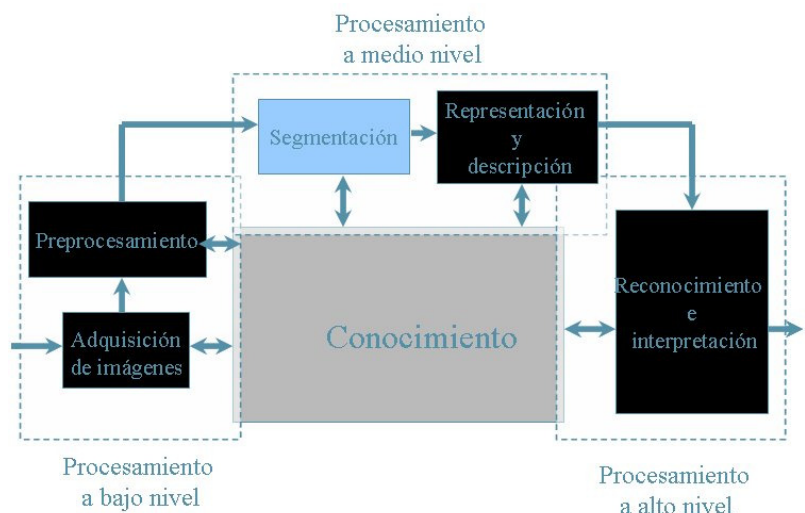


5 Segmentación

En las técnicas de procesamiento de las imágenes cada píxel era transformado de un valor a otro. Estos cambios, mayoritariamente, se realizan para facilitar la partición de la imagen en áreas de píxeles con significado. En esta nueva fase, se trata de agrupar los píxeles, por algún criterio de homogeneidad, para particionar la escena en regiones de interés. Estas áreas deben de tener algún significado físico. Por tanto, la segmentación de una imagen es un proceso de extracción de objetos de interés insertados en la escena capturada. La agrupación de los píxeles se hace a razón de que sus vecinos sean similares en criterios como de luminancia, color, bordes, texturas, movimientos,... Una vez que la imagen ha sido particionada, la unidad dejara de ser el píxel para ser la agrupación de píxeles que constituye el objeto. La imagen estará definida por un conjunto de objetos, habiendo pasado de un nivel bajo a otro más elaborado o nivel medio visual. La información estará preparada para el reconocimiento e interpretación de la imagen.



Para la segmentación de las imágenes se usan tres

conceptos básicos:

- Similitud: los píxeles agrupados del objeto deben ser similares respecto algún criterio (nivel de gris, color, borde, textura,...).
- Conectividad: los objetos corresponden a áreas de píxeles con conectividad. Las particiones corresponden con regiones continuas de píxeles.
- Discontinuidad: los objetos tienen formas geométricas que definen unos contornos. Estos bordes delimitan unos objetos de otros.

En la práctica, la imposición de estas condiciones sobre la estrategia de segmentación resulta casi imposible. Así, por ejemplo, si se empieza por los criterios de similitud y debido a la aparición del ruido, a la falta de iluminación uniforme sobre el escenario o a la creación de sombras de unos objetos sobre otros, todas estas causas producen que algo que parecía sencillo de definir como es alguna regla sobre similitud, resulte impracticable de acotar. Todas ellas fracasan sobre escenas más o menos complejas. En cuanto a la conectividad, la ocultación parcial de un objeto capturado hace fallarla. Suele ser normal en el proceso de segmentación que se produzcan varias regiones de píxeles agrupados que provienen del mismo objeto físico. No sólo debido a las propias oclusiones de éste, sino también a los propios cambios debido a la textura del objeto, a la variación de la iluminación de la escena o al reflejo de otros objetos adyacentes. Por último y tal cual se comentó en el anterior capítulo, la detección de los contornos físicos suele estar plagada de errores y de discontinuidades en los bordes. Resulta extraordinariamente difícil obtener los contornos cerrados, inmunes al ruido y sin desplazamiento entre el contorno real y el obtenido.

La conclusión es obvia, la segmentación es un proceso altamente complejo y sin solución final en muchos años. Es una labor de investigación actual y en las siguientes décadas venideras. En el paradigma actual se combinan las técnicas basadas en la localización de regiones uniformes con los contornos de los objetos.



Figura 5. 2. Estado actual de las técnicas de segmentación sobre escenarios naturales

No obstante, para la actual Visión Artificial todavía existe mucho recorrido para las aplicaciones industriales. Las citadas dificultades de la fase de segmentación son resueltas con una elección esmerada en la formación de la imagen, eligiendo una iluminación adecuada y simplificando la escena. Como se ha comentado varias veces, las técnicas de Visión Artificial resultan factibles para el análisis de imágenes simples y repetitivas.

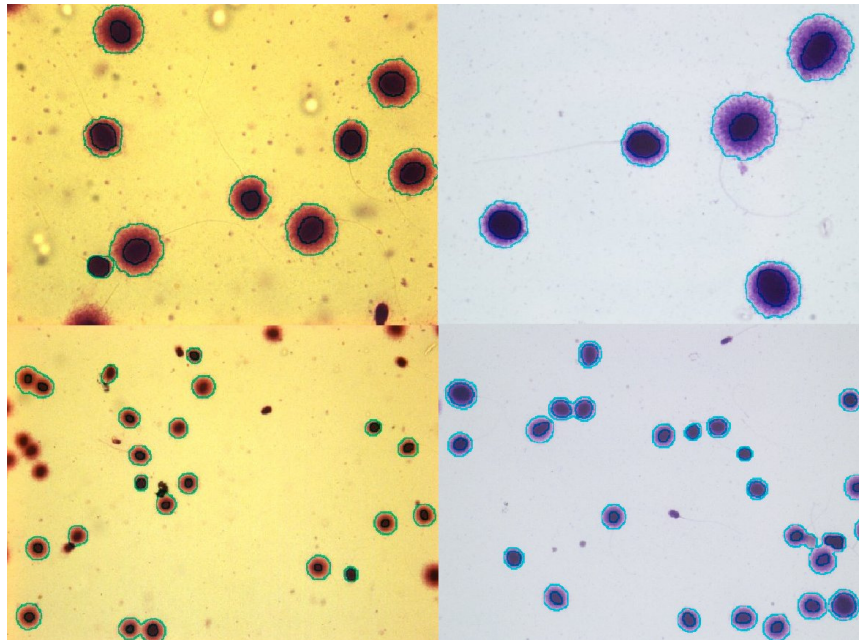


Figura 5. 3. Técnicas de segmentación sobre escenarios repetitivos y alienantes para el hombre. Aplicaciones de segmentación sobre imágenes procedentes de microscopía.

La exposición de este capítulo versa sobre las distintas técnicas de segmentación, clasificadas en:

- Contornos de los objetos: trata de localizar las fronteras de los objetos. Hace referencia a los aspectos geométricos de los objetos.
- Búsqueda de regiones homogéneas: agrupa los píxeles por que cumple algún criterio de similitud y tienen conectividad entre ellos. Está unido a las características radiométricas del objeto.

5.1 Técnicas basadas en los bordes

Para la localización de las fronteras de los objetos en la escena se emplearán las técnicas de detección de los bordes, vistas en el capítulo anterior. Sin embargo, esta etapa no es definitiva para poder segmentar los objetos presentes en la imagen. La presencia de ruido, el efecto de las sombras, la falta de iluminación uniforme y un largo etcétera de causas, produce que los contornos no sean del todo continuos y cerrados sobre los objetos. Se requiere otra etapa de post-procesamiento. Esta nueva fase emplea los resultados de la detección de bordes para elaborar las fronteras de los objetos. Se trata de agrupar los píxeles etiquetados como bordes, de la etapa anterior (de detección

de bordes), empleando la propiedad de conectividad. Para que un píxel etiquetado como borde se defina como píxel frontera de un objeto se necesitará que otros píxeles bordes tengan similar dirección y módulo del gradiente. Dos píxeles serán considerados pertenecientes a una misma frontera si presentan alguna condición de conectividad y las diferencias entre sus gradientes no superan un determinado umbral:

$$\begin{aligned} \|G_1 - G_2\| &< T_M \\ |\vartheta_1 - \vartheta_2| &< T_A \end{aligned} \quad (5.1)$$

Donde $|G_i|$ y θ_i se ha denotado el módulo y argumento del gradiente del píxel i , respectivamente. Evidentemente, en la construcción de la frontera, aparecerá ruido, píxeles que constituyen esquinas de los objetos y otras causas que producen discontinuidades en la localización de la frontera del objeto. Hay varios planteamientos para su decantación. Algunos autores emplean técnicas de crecimiento de regiones, las cuales se verán más adelante, pero que básicamente se trata de partir de un píxel borde semilla e ir agrupando píxeles con conectividad siempre y cuando cumplan algún criterio de homogeneidad, por ejemplo, el expresado en la anterior ecuación. Otros autores emplean un marco de trabajo basado en la optimización. Realizan un grafo de los posibles caminos que pueda llevar la frontera del objeto y mediante una función de costo, eligen la mejor solución. Un planteamiento alternativo es la agrupación de los píxeles que cumplen un cierto criterio en una primitiva más elaborada, como puede ser un pequeño segmento con orientación. Posteriormente, se procede a encadenar estas primitivas dando paso a la delimitación de los objetos.

Un campo muy activo en esta materia son los contornos activos. Se trata de emplear una curva cerrada que vaya adaptándose dinámicamente hasta alcanzar la frontera del objeto.

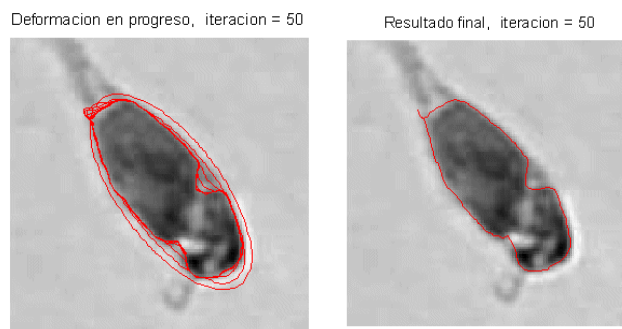


Figura 5. 4 Ejemplo de contornos activos

Muchas aplicaciones se basan en las localizaciones de líneas rectas o de otras formas básicas como círculos o elipses, ya que se consideran elementos fundamentales en la percepción de entornos estructurados o semi-estructurados. Se basa en el hecho de que las escenas artificiales, construidas por el hombre, están plagadas de contornos rectos o de objetos con formas circulares o elípticas. En este sentido, se aplican las transformadas de Hough para la localización de estas fronteras.

5.1.1 Transformada de Hough

Este algoritmo trata de detectar formas geométricas sencillas en una imagen. En su procesamiento emplea el conjunto total de la imagen, haciéndolo robusto ante la presencia del ruido o discontinuidades de las etapas previas. De hecho, para su ejecución requiere de una imagen binarizada en la que se han seleccionado previamente los bordes. Por tanto, la entrada al algoritmo de Hough es una imagen en la que se ha hecho la detección de los bordes. Hough, desde una perspectiva global, intentará extraer primitivas de más alto nivel como son líneas, circunferencias, elipses o cualquier tipo de curva parametrizada o no. El mayor inconveniente es su alto coste computacional.

5.1.1.1 Detección de líneas rectas

Considerando un píxel de coordenadas (x_i, y_i) , seleccionado como elemento que representa un borde; sobre éste pasarán infinitas rectas:

$$y_i = ax_i + b \quad (5.2)$$

El modelo de todas estas posibilidades quedan definidas por los infinitos valores de a y b . Haciendo la transformación al espacio paramétrico, los valores de a y b quedarán definidos como:

$$b = y_i - ax_i \quad (5.3)$$

Al variar el parámetro a desde $-\infty$ a $+\infty$ se obtendrá los infinitos valores de b . La representación geométrica de $b = y_i - ax_i$, en el espacio paramétrico, será una recta. La característica interesante de esta presentación consiste en que si dos píxeles que pertenezcan a una misma línea son representados en el espacio paramétrico, el modelo de la recta está definido en la intersección de las dos rectas del espacio paramétrico.

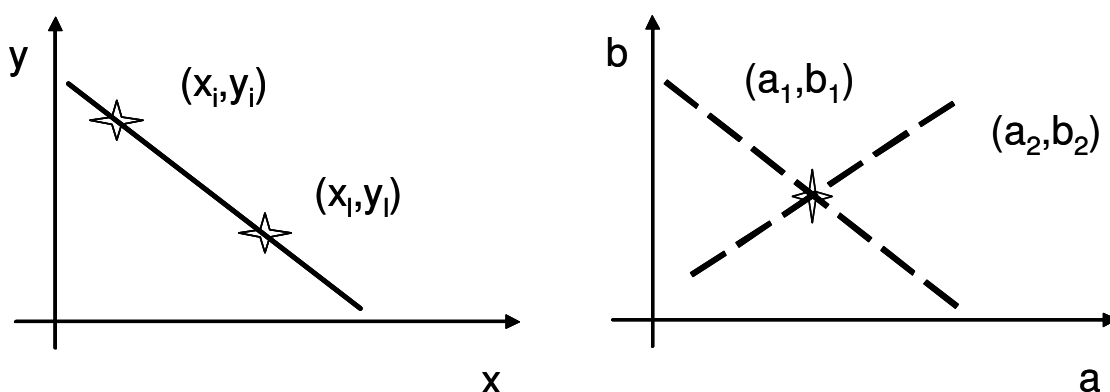


Figura 5.5 Ejemplo de conversión hacia el espacio paramétrico

La transformada de Hough (1962) aplica este concepto para la localización de líneas rectas en la imagen. Discretiza el espacio paramétrico en intervalos de $[a_{\min}, a_{\max}]$ y $[b_{\min}, b_{\max}]$, creando una rejilla de celdas de acumulación. Por cada píxel, considerado como borde, se hace recorrer el rango dinámico de a obteniendo los valores de b . Por

cada valor de a y b se le pone un voto en la celda correspondiente. Esta operación se hace con todos los píxeles etiquetados como bordes. Al finalizar, aquellas celdas con más votos indicarán la presencia de rectas en la imagen, cuyos modelos corresponderán con las coordenadas de la celda.

Sin embargo, el espacio paramétrico elegido no es el más correcto, ya que los rangos dinámicos de a y b no están limitados. En cambio, si se hace una representación en coordenadas polares, el

ángulo de la normal de la recta, θ , está limitado al rango de $[0 \pi]$:

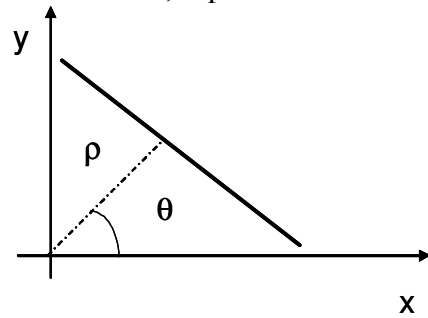


Figura 5. 6. Espacio paramétrico en coordenadas polares

$$x_i \cos \theta + y_i \sin \theta = \zeta \quad (5. 4)$$

La representación de un píxel seleccionado en este espacio paramétrico será una parte de una senoide.

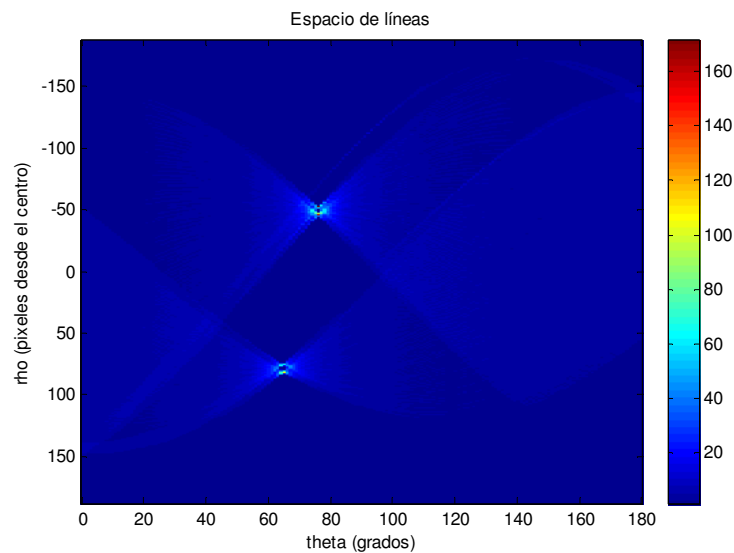


Figura 5. 7 Votaciones en el espacio paramétrico de coordenadas polares

Una forma de mejorar el coste computacional es aprovechar el ángulo del gradiente de cada píxel, pues éste, aunque es de carácter local, dará una primera aproximación del ángulo de la normal. De esta forma, se disminuye el rango dinámico de θ y por tanto del número de operaciones y de votos en la rejilla de acumulación.

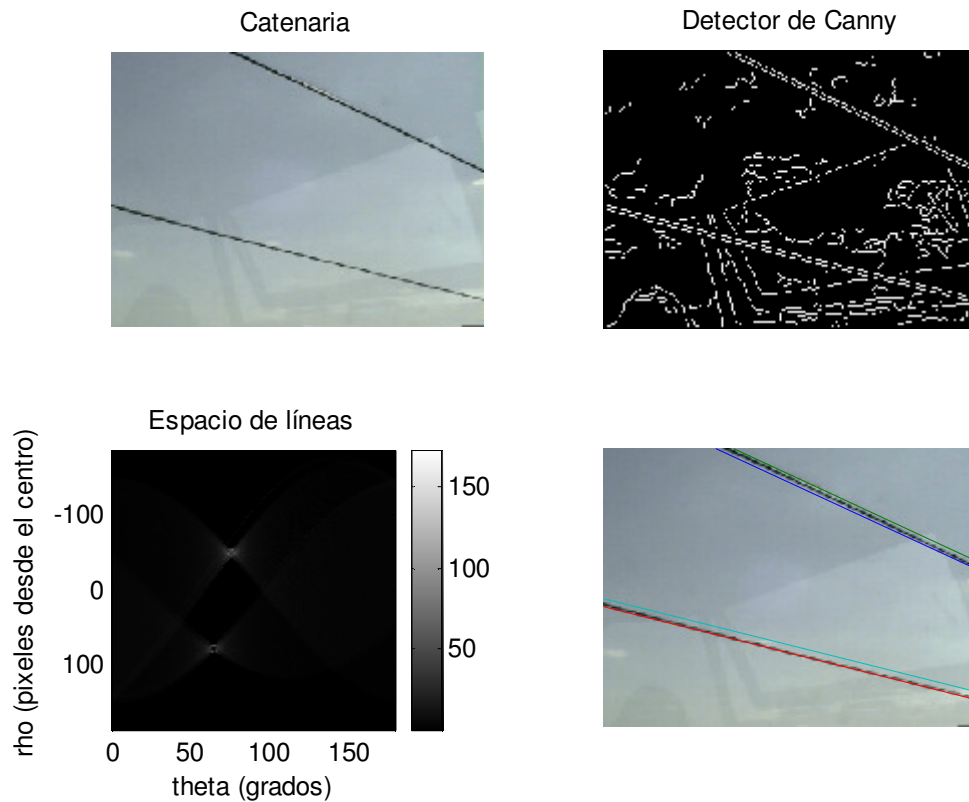


Figura 5. 8. Transformadas de Hough para líneas rectas

Resolución Matlab

```
>> imgEnt=imread('cables_gris.bmp');imshow(imgEnt);pause;
>> imgBorde=edge(imgEnt,'canny');imshow(imgBorde);
>> theta = 0:180;
>> [acum,rho] = radon(imgBorde,theta);
>> imagesc(theta,rho,acum), colorbar;
>> xlabel ('theta (grados)'), ylabel ('rho (pixeles desde el centro)')
>> title('Espacio de líneas');
```

Otro inconveniente de la transformada de Hough es que devuelve el modelo de una recta infinita y no de un segmento. No se sabe donde empieza y acaba la recta.

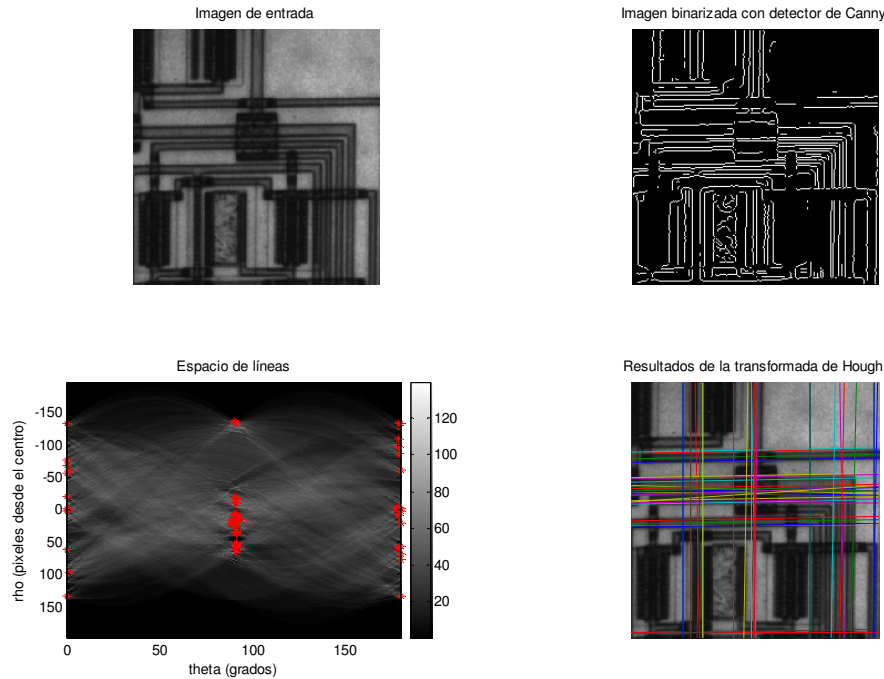


Figura 5. 9 Resultados de la transformada de Hough, si las rectas tenían al menos 80 votos

5.1.1.2 Detección de círculos

Para este caso, el espacio paramétrico será 3D, correspondiente a la posición del centro del círculo, (x_0, y_0) , y al radio, r :

$$(x_i - x_0)^2 + (y_i - y_0)^2 = r^2 \quad (5.5)$$

Las votaciones de cada celdilla del espacio paramétrico se realizará variando la posición del centro del círculo, (x_0, y_0) , para cada píxel etiquetado, (x_i, y_i) , obteniendo los distintos valores del radio, r . Por tanto, entrado en la rejilla de acumulación por posición del centro del círculo y del radio calculado se procederá a depositar los votos correspondientes. Los círculos seleccionados serán aquellos que superen un cierto umbral de votos recibidos. El mayor inconveniente es su alto coste computacional. Una forma de ahorro en el cálculo es prefijar el valor del radio del círculo. Si se conoce el radio, junto con la información de la orientación del gradiente del píxel etiquetado, (x_i, y_i) , las posibles coordenadas del centro del círculo darán una nube de puntos limitados.

Las ventajas de esta detección de círculos, provienen de las características inherentes de las transformadas de Hough, la inmunidad al ruido y que es capaz de detectar círculos, aunque exista oclusión del mismo. Por otro lado, en este tipo de curvas, al ser cerradas, no hay problemas del carácter infinito de las líneas rectas. Los círculos están definidos en la imagen, sin ninguna expansión.

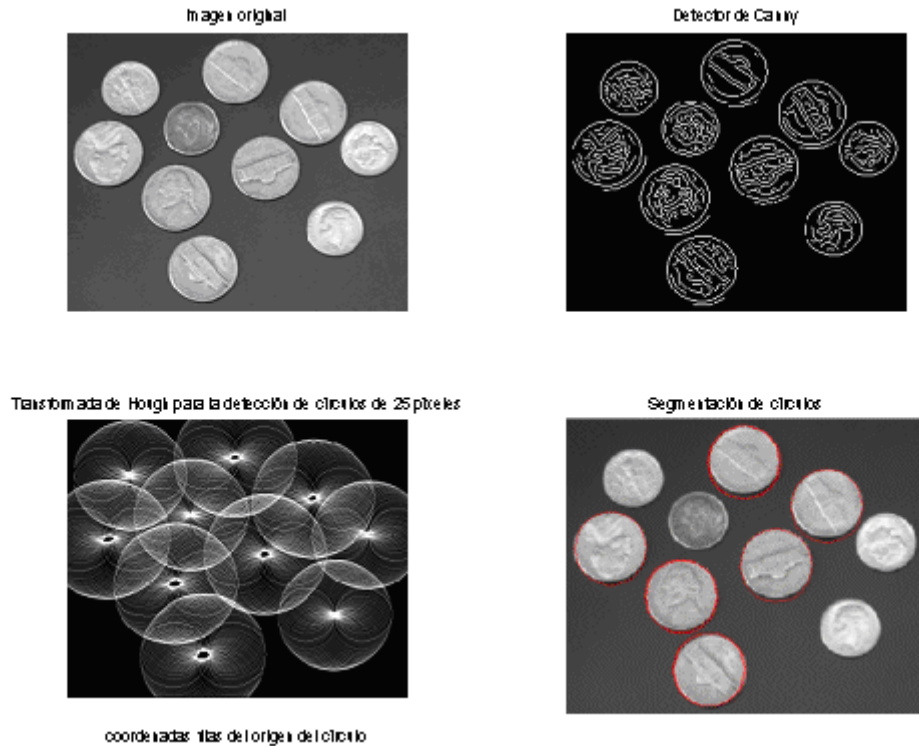


Figura 5. 10. Transformadas de Hough para círculos

Resolución Matlab

```
>>imgEnt = imread('coins.png');radioMoneda = 30;
>>[y0,x0,Accumulator]=houghcircle(edge(imgEnt,'canny'),radioMoneda,4)
>>UmbralVotaciones = 45; [x,y]=find(Accumulator>UmbralVotaciones);
>>imshow(imgEnt);hold on;
>>for i=1:size(x,1)
            dibujarCirculos(radioMoneda,y(i),x(i));
>>end;hold off
```

Las funciones *houghcircle* y *dibujarCirculos* no están en el toolbox, han sido implementadas.

5.1.1.3 Transformadas de Hough generalizadas

Cuando no hay expresión analítica en las curvas a detectar, se emplean las transformadas de Hough generalizadas. El método es esencialmente un detector de formas o curvas geométricas. Consiste en elegir un punto de referencia dentro del objeto a describir, en principio, próximo al centroide del mismo, (x_{ref}, y_{ref}) . Luego, siguiendo el borde del objeto se obtiene una tabla donde se

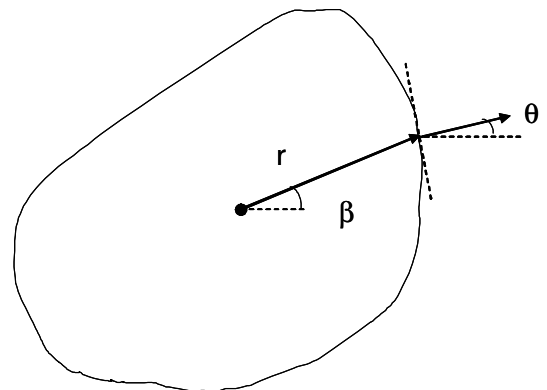


Figura 5. 11. Descripción de Hough para objetos cualesquieras

guarda la distancia al punto etiquetado, r , el ángulo de apertura, β , y la orientación del gradiente compensado, $\theta+90^\circ$. La tabla será indexada por la orientación del gradiente compensado, $\theta+90^\circ$.

Posteriormente, para cada píxel etiquetado se entrará en la tabla según la orientación del gradiente compensado, $\theta+90^\circ$, obteniéndose el radio y el ángulo de apertura, β . Con estos valores se evaluará la expresión:

$$\begin{aligned}x_{ref} &= x_i + r \cdot \cos(\beta + \alpha) \\y_{ref} &= y_i + r \cdot \sin(\beta + \alpha)\end{aligned}\quad (5.6)$$

Retornado para cada valor de α un punto de referencia distinta, (x_{ref}, y_{ref}) . El espacio de votaciones estará formado por:

- a) Las coordenadas de referencia del objeto, (x_{ref}, y_{ref}) .
- b) Y la orientación del objeto, α .

La celda con más votos indicará la posición y orientación del objeto en la imagen. Como se ha comentado, las transformadas de Hough generalizadas se emplean como descriptores de alto nivel de los bordes.

5.2 Umbralización

La umbralización es una técnica de segmentación ampliamente utilizada en las aplicaciones industriales. Se emplea cuando hay una clara diferencia entre los objetos a extraer respecto del fondo de la escena. Los principios que rigen son la similitud entre los píxeles pertenecientes a un objeto y sus diferencias respecto al resto. Por tanto, la escena debe caracterizarse por un fondo uniforme y por objetos parecidos.

Al aplicar un umbral, la imagen de niveles de grises quedará binarizada; etiquetando con '1' los píxeles correspondientes al objeto y con '0' aquellos que son del fondo. Por ejemplo, si los objetos son claros respecto del fondo, se aplicará:

$$g(x, y) = \begin{cases} 1 & f(x, y) > T \\ 0 & f(x, y) \leq T \end{cases} \quad (5.7)$$

$f(x, y)$ es la función que retorna el nivel de gris del píxel (x, y) , $g(x, y)$ será la imagen binarizada y T es el umbral. En el caso de que los objetos sean oscuros respecto del fondo, la asignación sería a la inversa:

$$g(x,y) = \begin{cases} 1 & f(x,y) < T \\ 0 & f(x,y) \geq T \end{cases} \quad (5.8)$$

El umbral puede depender de $f(x,y)$, de alguna propiedad local del píxel, $p(x,y)$ y hasta de su propia posición:

$$T = T(f(x,y), p(x,y), x, y) \quad (5.9)$$

Si el umbral sólo depende de $f(x,y)$ se dice que es un umbral global; en el caso de que además dependa de $p(x,y)$, por ejemplo, el valor medio del entorno de vecindad, el umbral es denominado local y si depende también de la posición del píxel, se denominará dinámico. El problema es encontrar el umbral; operación nada sencilla ya que las imágenes están contaminadas con el ruido. Para acotar el estudio, sólo se centrará en las técnicas globales.

5.2.1 Segmentación por histograma

La segmentación por histograma es una técnica global que se basa en considerar una definición clara de los objetos respecto del fondo, con una iluminación relativamente uniforme en toda la escena. Se supone que el histograma está constituido por dos picos y un valle. Al segmentar empleando el histograma, se denota que este proceder no considera las relaciones de vecindad de los píxeles. El umbral estará definido justamente en el valle.

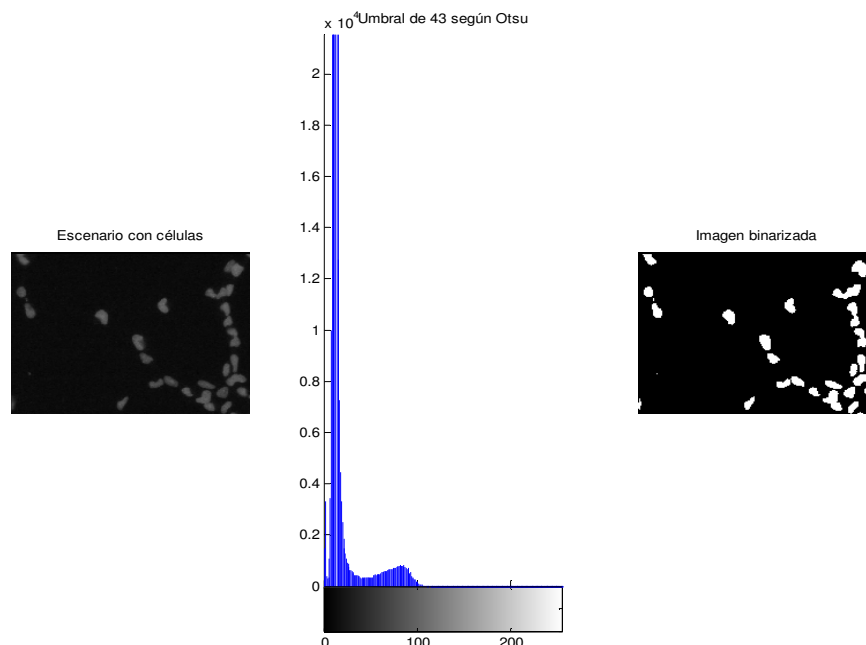


Figura 5. 12 Técnica de segmentación mediante análisis del histograma

Esta técnica es extensible a un histograma con N picos y $N-1$ valles, presentado varios umbrales. Sin embargo, a mayor número de umbrales en la imagen, menos fiables serán los resultados. La umbralización por histograma presenta ciertas limitaciones:

- Es difícil identificar correctamente los mínimos del histograma.
- Se tienen problemas cuando las regiones varían suavemente su nivel (por ejemplo, el efecto de las sombras).
- Se aplica sólo cuando hay pocas regiones.
- No se pueden distinguir regiones separadas de niveles similares de gris.
- No considera la conectividad de los píxeles. Debido a este inconveniente, se suele emplear a continuación técnicas de post-procesado (se tratarán más adelante).

5.2.1.1 Selección óptima del umbral

Hay diferentes estrategias para la elección del umbral óptimo; entre ellas destacan aquellas cuyos fundamentos se basan en:

- Ajuste gaussiano.
- Minimización de la varianza interclase.
- Entropía del histograma.
- Análisis de la concavidad.
- Métodos basados en momentos.

Un resumen detallado de los distintos algoritmos se puede encontrar en revistas especializadas¹. De entre todas las propuestas se va a destacar el umbral de Otsu, por su alta utilización en las aplicaciones de Visión Artificial.

¹ Sahoo, P.K., *A survey of Thresholding Techniques*, Computer Vision, Graphics and Image processing, 41, pp. 233-260, 1988.

5.2.1.1.1 Método de Otsu

Este método se basa en la suposición de que la función de densidad del fondo, C_f , y de los objeto, C_o , tienen modelo gaussiano, $N(\mu_f, \sigma_f^2)$ y $N(\mu_o, \sigma_o^2)$. Cada grupo estará formado por los niveles de grises fijados por el umbral, T :

$$C_f = \{0, 1, 2, \dots, T\} \quad C_o = \{T+1, T+2, \dots, I-1\} \quad (5.10)$$

El umbral debe de minimizar la suma ponderada de cada una de las varianzas de las clases, C_f y C_o , ya que se supone que conforme la suma de las dos normales se aproxime más al histograma real, las desviaciones serán menores.

Para determinar los coeficientes se toman las probabilidades de cada una de las clases. Considerando un valor fijo de umbral, T , las probabilidades de cada categoría serán:

$$P_{C_f} = \sum_{i=0}^T p_i \quad P_{C_o} = \sum_{i=T+1}^{I-1} p_i \quad (5.11)$$

Donde p_i es la probabilidad de la intensidad i en la imagen. Las medias y varianzas de cada grupo corresponderán a:

$$\begin{aligned} \mu_{C_f} &= \frac{1}{P_{C_f}} \sum_{i=0}^T i \cdot p_i & \mu_{C_o} &= \frac{1}{P_{C_o}} \sum_{i=T+1}^{I-1} i \cdot p_i \\ \sigma_{C_f}^2 &= \frac{1}{P_{C_f}} \sum_{i=0}^T (i - \mu_{C_f})^2 \cdot p_i & \sigma_{C_o}^2 &= \frac{1}{P_{C_o}} \sum_{i=T+1}^{I-1} (i - \mu_{C_o})^2 \cdot p_i \end{aligned} \quad (5.12)$$

Siendo entonces la varianza ponderada:

$$\sigma_p^2 = P_{C_f} \cdot \sigma_{C_f}^2 + P_{C_o} \cdot \sigma_{C_o}^2 \quad (5.13)$$

Se recorre todo el rango de niveles de gris, de T igual 0 a $I-1$, calculándose la varianza ponderada y se elige el umbral, T , que minimiza este valor.

5.3 Segmentación orientada a las regiones

La estrategia a emplear en este tipo de segmentación está basada en los criterios de similitud y continuidad de los píxeles que forman una región. Bajo esta perspectiva, la imagen se considera formada por n regiones disjuntas, cada una de las cuales tiene agrupada a los píxeles por alguna propiedad que los hace ser característicos de esa zona y discrepantes respecto al resto. Desde el punto de vista formal, las condiciones de la segmentación orientada a regiones serían:

$$\begin{aligned}
a) \quad I &= \bigcup_{i=1}^n R_i & b) \quad R_i \cap R_j &= \emptyset \quad i \neq j \\
c) \quad p_i \in R_j &\leftrightarrow P_{R_j}(p_i) = 1 & d) \quad p_i \notin R_j &\leftrightarrow P_{R_j}(p_i) = 0
\end{aligned} \tag{5.14}$$

Siendo I la imagen, R_i una región de ésta, p_i un píxel cualquiera de la imagen y P_{R_j} la regla de similitud de la región j . Las dos primeras condiciones describen la segmentación como un proceso de partición de la imagen en regiones disjuntas; mientras, las otras dos restantes hacen referencia a las propiedades de similitud de los píxeles agrupados y de discrepancia respecto al resto.

5.3.1 Crecimiento de regiones

A partir de píxeles semillas, se hace crecer la región considerando alguna regla de similitud y considerando la propiedad de conectividad de los píxeles. Una vez definida una semilla, se analiza los vecinos según la regla de similitud de la región, P_{R_j} , los píxeles que cumplan la propiedad se añadirán a la región de crecimiento; aquellos que no cumplan significarán que pertenecen a otra región. Con los nuevos píxeles unidos a la región de crecimiento, se volverá a procesar sus nuevos vecinos. El algoritmo de crecimiento finalizará cuando todos los vecinos hayan sido estudiados. Nótese el carácter de conectividad y similitud de la técnica de segmentación de crecimiento de regiones.

Esta estrategia tiene dos problemas: a) ¿Cuál es el criterio de similitud de las regiones?, y b) ¿Cómo se definen las semillas de las regiones?. Para la primera cuestión, las reglas de similitud de las regiones se fundamentan en alguna propiedad tal como luminancia, color, textura, etc. Por ejemplo, se presentan tres reglas basadas en intensidad:

$$a) |f(x, y) - f(x_{s,i}, y_{s,i})| < T \quad b) |f(x, y) - \mu_{R_i}| < T \quad c) |f(x, y) - \mu_{R_i}| < k\sigma_{R_i} \tag{5.15}$$

Donde $f(x, y)$ es la intensidad del píxel (x, y) , $f(x_{s,i}, y_{s,i})$ es nivel de gris de la semilla, T es un umbral, μ_{R_i} es el brillo de la región de crecimiento R_i y σ_{R_i} es su varianza o contraste. La primera regla añade píxeles vecinos que sea menores a una variación determinada del nivel de gris respecto a la semillada, fijado por el umbral T . La segunda actualiza la comparación en cada iteración con el valor medio de intensidad de la región creciente. Por último, en la tercera, se emplea como umbral k veces la desviación típica de la región.

La segunda cuestión resulta ser tan decisiva como la regla de similitud: ¿cuál es el píxel semilla?. La elección de uno u otro píxel hace que el resultado final de la segmentación pueda variar. Estas técnicas suelen ser frágiles en sus resultados por la altísima dependencia entre los píxeles semillas y las reglas de similitud. Por tanto, se suele acompañar con otras estrategias para incrementar su robustez. La elección de las semillas puede abordarse buscando el nivel de gris de mayor probabilidad en el histograma, esto es, el nivel de gris con mayor número de ocurrencias. Para su situación espacial se puede apoyarse en una detección de los bordes, los cuales delimitarán las

fronteras de las regiones. Los centroides de las áreas delimitadas por las fronteras, junto con valor del nivel de gris, dará un conjunto de semilla. En todo caso, la casuística de cada problema de análisis de imagen decidirá que algoritmia emplear en la localización de las semillas y en las reglas de similitud.

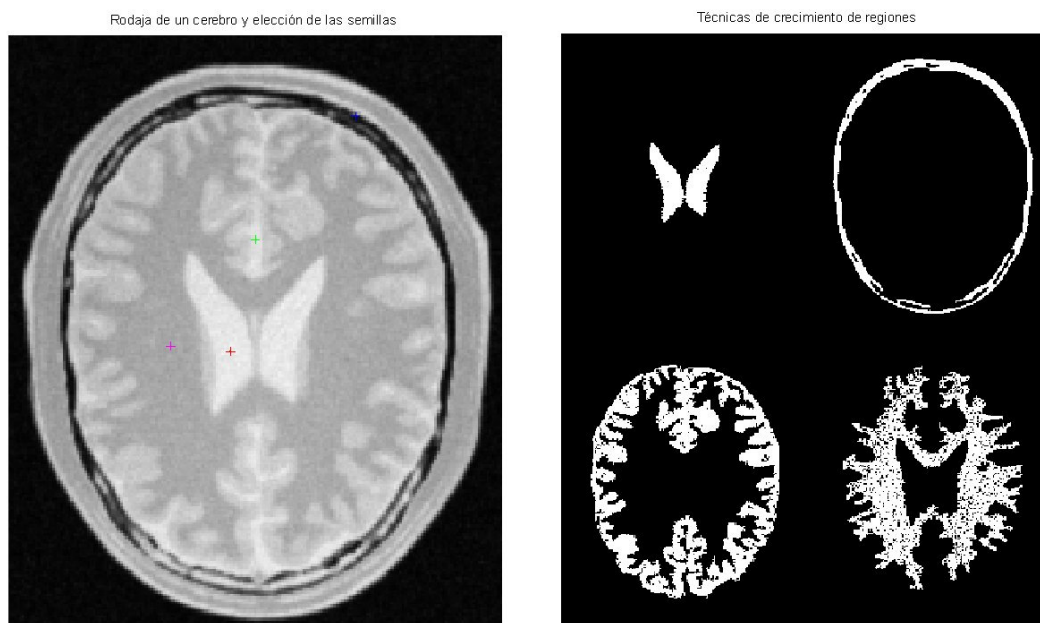


Figura 5. 13 a) imagen adquirida con etiquetado los píxeles semilla b) Resultados de la técnica de crecimiento empleando la regla del brillo de la región creciente

5.3.2 Pirámides y árboles cuaternarios

Muchas técnicas de análisis de imágenes, no sólo las técnicas de segmentación orientada a regiones, emplean en su procesamiento distintas resoluciones espaciales de la imagen. La información visual se puede organizar en forma de pirámide. En la base se encontraría la máxima resolución espacial, con M filas y N columnas y a partir de este valor, reducir el número de píxeles en cada nivel ascendente en la pirámide hasta alcanzar la cúspide con un único píxel. Normalmente, se suele emplear un promediado de los píxeles en la reducción de la resolución espacial, de un nivel a otro.

Una de las formas piramidales de resolución son los árboles cuaternarios (*quadtree*). La imagen de máxima resolución se divide en cuatro

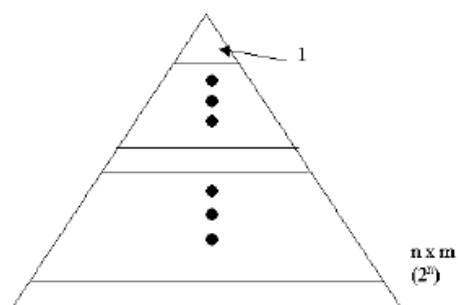


Figura 5. 14. Pirámide de resoluciones de la imagen

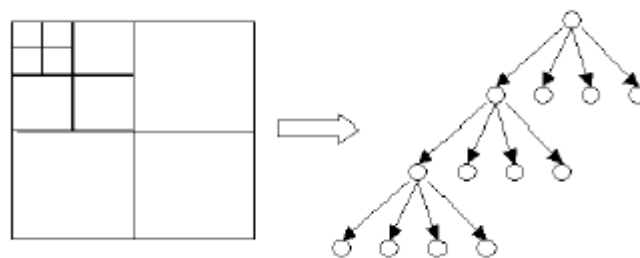


Figura 5. 15 Estructura de árboles cuaternarios

rectángulos iguales. Cada región es posteriormente dividida en otras cuatro subregiones. Estas divisiones se van haciendo recursivamente hasta alcanzar a nivel de píxel. Por tanto, por cada nodo o región padre colgarán otras cuatro subregiones. Dando motivo a denominarlo como árbol cuaternario.

Sin embargo, estas divisiones arbitraria de las regiones en otras subcuatros, en cualquier nivel de la pirámide, se hace en una división geométrica buscando cuatro subregiones de igual número de píxeles. Esta forma de dividir no tiene en cuenta el contenido de la escena, produciendo una descomposición rápida desde el punto de vista computacional, pero pobre desde el punto de vista de análisis de las imágenes. Las tendencias actuales es a una actuación mixta considerando en las resoluciones la posibilidad de solapamiento entre las regiones adyacentes.

5.3.3 División y fusión de regiones

Uno de los puntos débiles de las técnicas de crecimiento de regiones está en la localización de las semillas. Para superar esta situación se pasa a la descomposición de la imagen en regiones arbitrarias, de forma que si la región es muy discrepante con algún tipo de regla se dividirá, en caso contrario, buscará fusionarse con las regiones adyacentes.

Esta técnica suele emplear una descomposición inicial de la imagen empleando un árbol cuaternario. En esta fase, la imagen se ha particionado en un conjunto de regiones arbitrarias y disjuntas. La regla de división suele estar relacionada con el contraste o varianza de la región. Si la varianza es elevada se pasará a descomponerlo en cuatro subregiones.

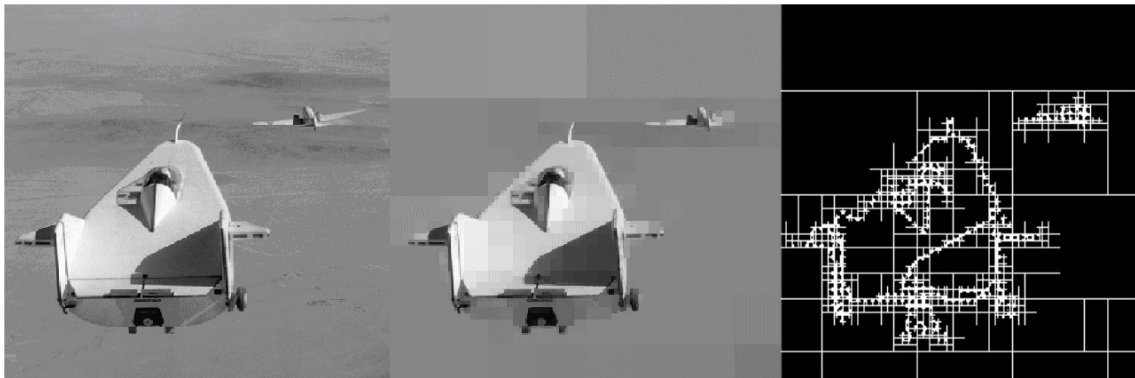


Figura 5. 16 División de la imagen de entrada empleando árboles cuaternarios

Una vez dividida la imagen, se procede a la fusión de regiones adyacentes empleando técnicas de crecimiento de regiones.

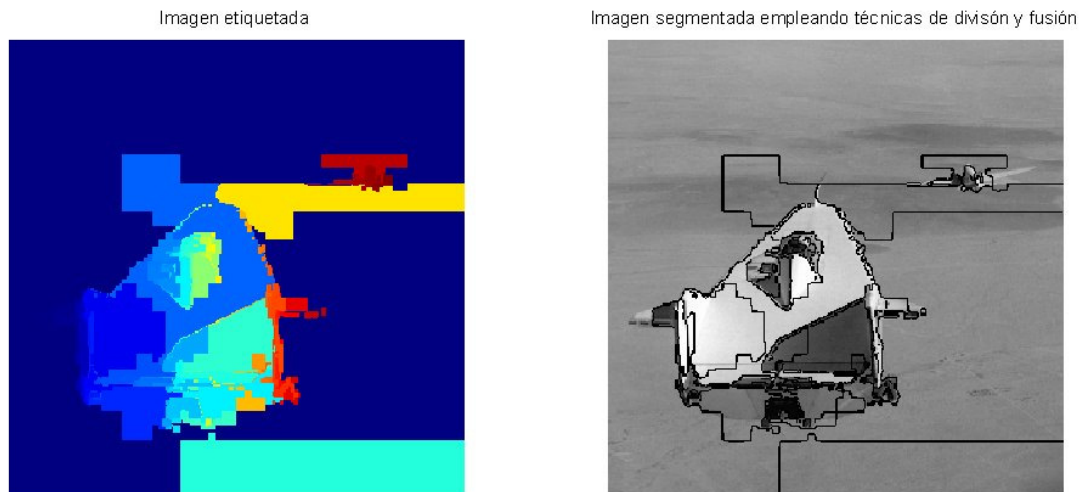


Figura 5. 17. Resultado de la fusión de las regiones empleando técnicas de crecimiento de regiones

Los resultados de esta técnica de segmentación no preservan los contornos reales. Esta consecuencia se debe al proceso de división de la imagen en cuatro áreas iguales, sin considerar el contenido existente en ellas.

Resolución Matlab

```
>>%División de la imagen
>>imgEnt = imread('liftingbody.png');
>>im1=imfilter(imgEnt,fspecial('gaussian'));
>>imgDescomp = qtdecomp(im1,.27);
>>imgDivision = imgEnt;
>>for dim = [128 64 32 16 8 4 2 1]
    [valores,filas,columnas] = qtgetblk(im1, imgDescomp, dim);
    if (~isempty(valores))
        doublesum = sum(sum(valores,1,'double'),2);
    end
    imgDivision = qtsetblk(imgDivision, imgDescomp, ...
        dim, doublesum ./ dim^2);
>>end
```

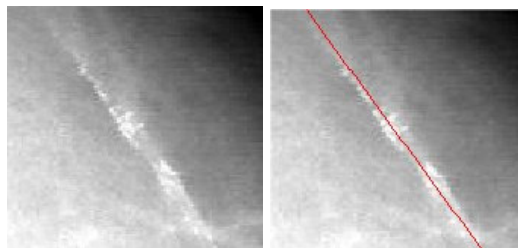
5.4 Ejercicios

1. Estado actual de las técnicas de segmentación.
2. Clasificación de las técnicas de segmentación.
3. Transformadas de Hough para líneas rectas y círculos.
4. Método de umbralización de Otsu.
5. Técnicas de crecimiento de regiones.
6. Técnicas de división y fusión de regiones.

5.5 Problemas

Ejercicio 1

Segmentar el vaso capilar de la mamografía de la figura a) como se muestra en b)



Se trata de determinar la recta que identifica el vaso capilar presente en la mamografía. Para ello se utilizará la **transformada de Hough**. Primero se aplica una detección de bordes, por ejemplo, mediante el algoritmo de Canny, donde habrá que definir la escala, a través de la varianza, y el nivel de umbral.

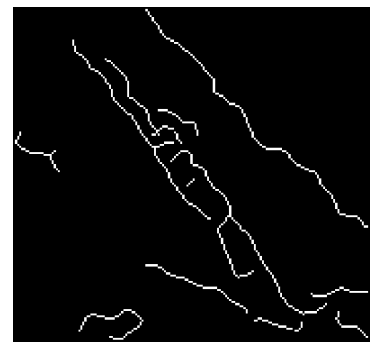
Una vez obtenida la imagen binaria de detección de bordes, $f(x,y)$, se procede al siguiente algoritmo:

- a) Inicialización de la matriz de votaciones:

$$C(\theta, r) = 0, \text{ para } \theta = \theta_1, \dots, \theta_a, r = r_1, \dots, r_b.$$

- b) Para $\theta = \theta_1, \dots, \theta_a$:

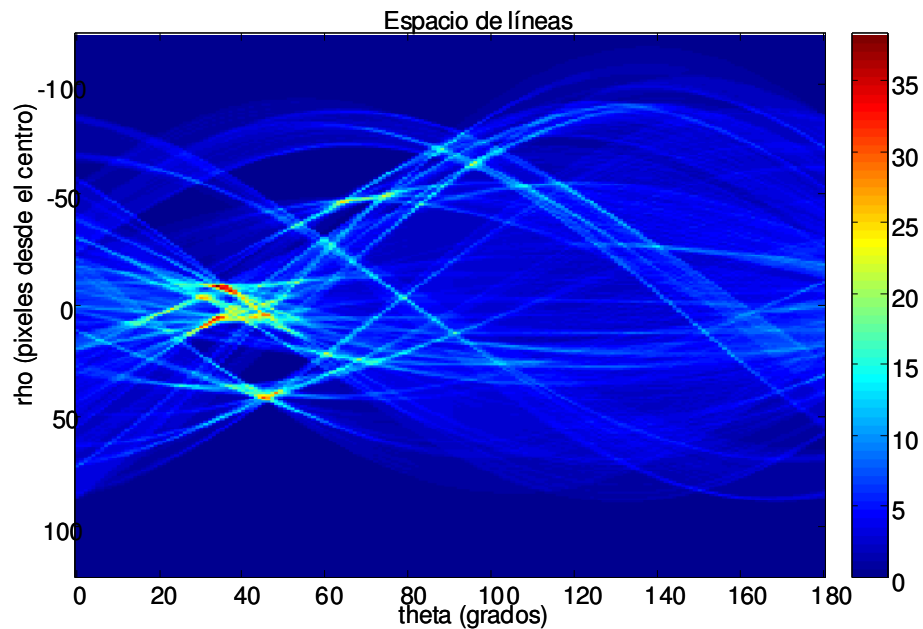
Para todos los píxeles (x_i, y_i) con $f(x_i, y_i)=1$ se determina:



$$r = x_i \cos \theta + y_i \sin \theta$$

y se añade el voto: $C(\theta, r) = C(\theta, r) + 1$.

El algoritmo da la matriz de votaciones $C(\theta, r)$. La recta se determinará con la celda de mayor votación.



Ejercicio 2

Para la siguiente figura, con 8 niveles de grises de cuantificación, obtener:

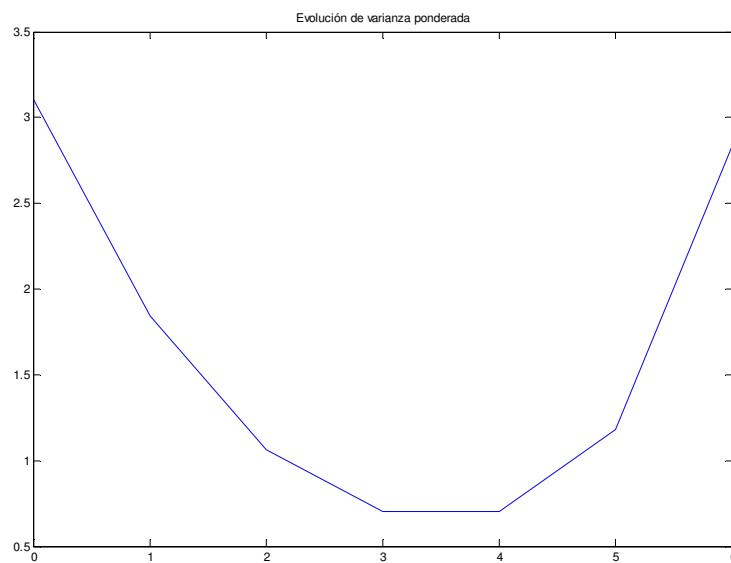
a) Umbral mediante método Otsu

1	1	2	0	2	2	1	3	1	1
1	1	3	3	3	3	2	2	1	0
1	1	2	2	5	3	2	2	1	1
3	2	3	6	6	6	2	2	1	1
2	2	6	7	7	6	6	1	1	1
2	2	6	6	7	6	6	0	3	2
2	2	2	6	5	5	3	2	1	1
1	1	2	3	5	5	3	2	1	0
1	1	1	3	3	3	3	2	2	1
0	1	1	1	1	2	1	1	1	0



nivel(i)	0	1	2	3	4	5	6	7
h(i)	6	33	26	16	0	5	11	3
p(i)	0.06	0.33	0.26	0.16	0	0.05	0.11	0.03

T	P_{cf}	P_{co}	μ_{cf}	μ_{cf}	σ_{cf}^2	σ_{cf}^2	σ_p^2
0	0.06	0.94	0	2.60	0	3.30	3.10
1	0.39	0.61	0.84	3.47	0.13	2.93	1.84
2	0.65	0.35	1.30	4.57	0.39	2.30	1.06
3	0.81	0.19	1.64	5.89	0.77	0.41	0.70
4	0.81	0.19	1.64	5.89	0.77	0.41	0.70
5	0.86	0.14	1.83	6.21	1.34	0.16	1.18
6	0.97	0.03	2.30	7.00	2.93	0	2.84



El umbral según el método de Otsu será 3 ó 4.

Ejercicio 3

Para la siguiente matriz, realizar una descomposición de árbol cuaternario empleando como regla de división de la región que la diferencia entre el nivel de gris máximo y mínimo debe de mayor a un determinado umbral:

- Obtener el árbol cuaternario si el umbral es 5.
- Obtener la matriz de salida con el umbral anterior.

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 2 & 3 & 6 & 6 \\ 1 & 1 & 2 & 1 & 4 & 5 & 6 & 8 \\ 1 & 1 & 1 & 1 & 7 & 7 & 7 & 7 \\ 1 & 1 & 1 & 1 & 6 & 6 & 5 & 5 \\ 20 & 22 & 20 & 22 & 1 & 2 & 3 & 4 \\ 20 & 22 & 22 & 20 & 5 & 4 & 7 & 8 \\ 20 & 22 & 20 & 20 & 9 & 12 & 40 & 12 \\ 20 & 22 & 20 & 20 & 13 & 14 & 15 & 16 \end{pmatrix}$$

a)

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 2 & 3 & 6 & 6 \\ 1 & 1 & 2 & 1 & 4 & 5 & 6 & 8 \\ 1 & 1 & 1 & 1 & 7 & 7 & 7 & 7 \\ 1 & 1 & 1 & 1 & 6 & 6 & 5 & 5 \\ 20 & 22 & 20 & 22 & 1 & 2 & 3 & 4 \\ 20 & 22 & 22 & 20 & 5 & 4 & 7 & 8 \\ 20 & 22 & 20 & 20 & 9 & 12 & 40 & 12 \\ 20 & 22 & 20 & 20 & 13 & 14 & 15 & 16 \end{pmatrix}$$

b)

$$\begin{pmatrix} 1.06 & 1.06 & 1.06 & 1.06 & 3.5 & 3.5 & 6.5 & 6.5 \\ 1.06 & 1.06 & 1.06 & 1.06 & 3.5 & 3.5 & 6.5 & 6.5 \\ 1.06 & 1.06 & 1.06 & 1.06 & 6.5 & 6.5 & 6 & 6 \\ 1.06 & 1.06 & 1.06 & 1.06 & 6.5 & 6.5 & 6 & 6 \\ 20.75 & 20.75 & 20.75 & 20.75 & 3 & 3 & 5.5 & 5.5 \\ 20.75 & 20.75 & 20.75 & 20.75 & 3 & 3 & 5.5 & 5.5 \\ 20.75 & 20.75 & 20.75 & 20.75 & 12 & 12 & 40 & 12 \\ 20.75 & 20.75 & 20.75 & 20.75 & 12 & 12 & 15 & 16 \end{pmatrix}$$

Derecho de Autor © 2008 Carlos Platero Dueñas.

Permiso para copiar, distribuir y/o modificar este documento bajo los términos de la Licencia de Documentación Libre GNU, Versión 1.1 o cualquier otra versión posterior publicada por la Free Software Foundation; sin secciones invariantes, sin texto de la Cubierta Frontal, así como el texto de la Cubierta Posterior. Una copia de la licencia es incluida en la sección titulada "Licencia de Documentación Libre GNU".

La Licencia de documentación libre GNU (GNU Free Documentation License) es una licencia con [*copyleft*](http://www.gnu.org/copyleft/) para [*contenidos abiertos*](http://www.gnu.org/licenses/). Todos los contenidos de estos apuntes están cubiertos por esta licencia. La version 1.1 se encuentra en <http://www.gnu.org/copyleft/fdl.html>. La traducción (no oficial) al castellano de la versión 1.1 se encuentra en <http://www.es.gnu.org/Licencias/fdles.html>