



SELECCIÓN AUTOMÁTICA DEL NIVEL OPTIMO DE UMBRAL PARA SEGMENTACIÓN POR HISTOGRAMA DE IMÁGENES RGB DE CORTES HISTOLÓGICOS DE CÁNCER DE MAMA

M. Fátima De Abreu (1) Victor Barrios (2) Aldo Reigosa (3)

(1)MSc en Ingeniería Eléctrica, Sistema Digital de Imágenes Médicas. Profesor Ordinario Dedicación Exclusiva. Departamento de Computación. Facultad de Ingeniería. Universidad de Carabobo. e-mail: mdeabreu@uc.edu (2)**MSc en Ingeniería Electrónica. Profesor Ordinario Dedicación Exclusiva. Departamento de Física. Facultad de Ingeniería. Universidad de Carabobo. e-mail. vbarrios@uc.edu** (3) **Facultad de Ciencias de la Salud. Universidad de Carabobo.**

Resumen

En este trabajo se presenta un método de selección automática del óptimo valor de umbral para segmentación de núcleos celulares por histograma en imágenes de tejido canceroso. El método consiste en determinar el nivel de gris correspondiente a la mínima probabilidad existente entre las medias de dos distribuciones normales correspondientes a los histogramas de objetos y fondo. Inicialmente se obtienen tres imágenes correspondientes a los planos de tinte, saturación y brillo. Posteriormente se aplicó el algoritmo que implementa el método de selección de umbral a los planos de brillo y tinte, y finalmente se realizó la segmentación de los núcleos. Otro ensayo consistió en obtener los bordes nucleares mediante segmentación por nivel de gradientes del plano de brillo y aplicar un algoritmo de adelgazamiento de líneas a los bordes obtenidos. Para el cálculo de gradientes se utilizó el operador Prewitt. Se obtuvieron mejores bordes aplicando el procedimiento anterior sobre la imagen segmentada por tinte. Los resultados obtenidos demuestran la validez del método en este tipo de imágenes.

Palabras claves: segmentación, segmentación automática, detección de bordes, imágenes histológicas, cáncer de mama, automatic segmentation.

Abstract

In this paper a method of automatic selection of the best threshold value is presented for segmentation of cellular nuclei by histogram in cancerous tissue images. The method consists on determining the level of gray corresponding to the minimum existent probability between the means of two normal distributions corresponding to the histograms of objects and background. Initially three images were obtained corresponding to the hue, saturation and brightness planes. Later the algorithm that implements the method of threshold selection was applied to the planes of brightness and hue, and finally the segmentation of the nuclei was carried out. Another attempt consisted on obtaining the nuclear borders by segmentation of the level of gradients of the brightness plane and on applying an algorithm of line thinning to the borders obtained. The operator Prewitt was used for the calculation of gradients. It was obtained better borders

applying the previous procedure on the image segmented by hue. The results demonstrate the validity of the method in this type of images.

INTRODUCCIÓN

En el diagnóstico de cáncer, existe la necesidad entre patólogos de reproducir resultados y eliminar discrepancias que conllevan a diferencias en las terapias sugeridas y pronósticos. Esto se lograría, utilizando sistemas de análisis de imágenes computarizados que trabajen en forma automática sin la intervención de los especialistas. La morfometría de los núcleos celulares guarda relación con la patología y es de gran utilidad diagnóstica [1]. El estudio de la forma, tamaño y cantidad de los núcleos se realiza sobre la imagen de contornos nucleares, para obtener los contornos es necesario separar los núcleos dentro de la imagen. El proceso mediante el cual se separan en una imagen los objetos de interés se denomina segmentación. La segmentación por histograma es el método más simple que existe y generalmente se utiliza en imágenes con histograma bimodal. La selección del valor de umbral es crucial para lograr una exitosa segmentación por histograma. La selección automática de este valor permite realizar una segmentación objetiva sin la participación del especialista.

La preparación previa del tejido bajo estudio, involucra un proceso de coloración por hematoxilina-eosina que produce un tinte azul violeta en el núcleo y rosado en el citoplasma, logrando así un contraste objeto-fondo que permite diferenciación por color a simple vista. El histograma de tinte que se obtiene en estas imágenes es bimodal, donde una cresta del histograma corresponde al tinte violeta, el objeto y la otra al tinte rosado, el fondo. La segmentación por color, de núcleos celulares ha sido investigada en trabajos publicados anteriormente: utilizando redes neuronales [2], segmentación por histograma utilizando el método de la moda para calcular el valor de umbral [3], ISAP [4] que utiliza un método de selección automática de nivel de umbral para segmentación en imágenes de gris, basado en cálculos estadísticos. En este trabajo se aplica un algoritmo iterativo [5] de selección automática del nivel óptimo de umbral, que se basa en determinar el valor de gris, correspondiente a la mínima probabilidad entre las medias de dos distribuciones normales (histogramas de objetos y fondo). Se muestran los resultados obtenidos con segmentación por color, segmentación por nivel de gris y obtención de bordes a partir de segmentación por gradientes.

Medición de color en el sistema HSB

Existen diferentes sistemas para especificar el color de un píxel. Uno de ellos es el sistema RGB (Red, Green, Blue) utilizado por equipos electrónicos. En este trabajo se utiliza otro sistema denominado HSB (Hue, Saturation, Brightness) porque además de facilitar el procesamiento de imágenes a color, es más adecuado para la interpretación y entendimiento de las imágenes en estudio por parte de los especialistas. En este caso las imágenes se obtienen de tejidos sometidos a un proceso de tinción a objeto de utilizar el color como referencia en el diagnóstico. Es más fácil manejar la información de color en este sistema porque "el color" que el especialista observa en la lámina está representado por el valor de tinte (H). En HSB, el Tinte se refiere al color en sí. Saturación se refiere a la cantidad de mezcla del color con el blanco y Brillo se refiere a la cantidad de intensidad de luz que es percibida por el ojo humano [6]. La información de brillo o luminancia (L) de cada píxel de la imagen RGB, se determina como una mezcla de los valores R, G, B en las siguientes proporciones [3]:

$$Y = 0.30 * R + 0.59 * G + 0.11 * B \quad (1)$$

La figura 2, muestra la imagen de brillo obtenida de aplicar esta transformación a la imagen de la figura 1.

La medición de color H y S, se obtiene a partir de las siguientes expresiones:

$$R_Y = 0.70 * R - 0.59 * G - 0.11 * B \quad (2)$$

$$B_Y = -0.30 * R - 0.59 * G + 0.89 * B \quad (3)$$

$$H = \arctan\left(\frac{R_Y}{B_Y}\right) \quad (4)$$

$$S = \sqrt{B_Y^2 + R_Y^2} \quad (5)$$

Las imágenes de cada plano HSB o RGB se pueden visualizar como imágenes en tono de gris. Cada tono de gris corresponde a un valor comprendido entre 0 y 255, asignando "0" al negro y "255" al blanco y se denomina a cualquier valor en este rango "nivel de gris". Las figuras 2, 3 y 4 muestran la visualización como imágenes de gris de los planos obtenidos al transformar de RGB a HSB.

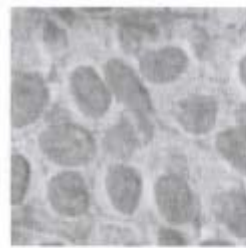


Figura 1.- Imagen RGB

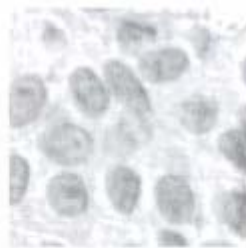


Figura 2.- Plano de Brillo



Figura 3.- Plano de Tinte



Figura 4.- Plano de Saturación

Segmentación

Por histograma de niveles de gris [7]

En una imagen de niveles de gris, los objetos se caracterizan por tener un conjunto de niveles de gris característicos, que lo diferencian del fondo. Para separar objetos del fondo de la imagen se define una transformación de una imagen de entrada F en una imagen binaria de salida como sigue:

$$\begin{aligned} g(y,x) &= 0 && \text{para } f(y,x) \geq T \\ g(y,x) &= 1 && \text{para } f(y,x) < T \end{aligned} \quad (6)$$

donde T es el nivel de umbral, para los píxeles correspondientes a los objetos y para los píxeles correspondientes al fondo o viceversa. El par representa la posición del píxel en la imagen.

Por histograma de tinte

Una imagen que posee dos tintes contrastantes (figura 1) permitiendo diferenciar a simple vista los objetos del fondo, presenta un histograma de tinte bimodal, como el de la figura 5. La segmentación por tinte en estas imágenes consiste en determinar el nivel de umbral T y aplicar la transformación definida en (6).

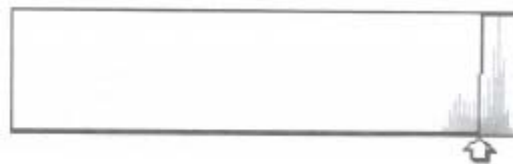


Figura 5. Histograma de tinte de la imagen RGB y nivel de umbral

Bordes

Un borde es una propiedad de un píxel y se calcula a partir del comportamiento de la función imagen en la vecindad del píxel. Es un vector variable que tiene dos componentes, magnitud y dirección. La magnitud del borde es la magnitud del gradiente y la dirección del borde está rotada -90° con respecto a la dirección del gradiente. La dirección del gradiente da la dirección del máximo crecimiento de la función, por ejemplo de negro ($f(Y,x) = 0$) a blanco ($f(y, x) = 255$). En la figura 6, las líneas cerradas son líneas de igual brillo. Los bordes se utilizan en análisis de imágenes para encontrar fronteras de regiones. Si una región tiene un brillo homogéneo su frontera se encuentra en aquellos píxeles que poseen borde con mayor magnitud, considerando el caso ideal de que no existe ruido [6].

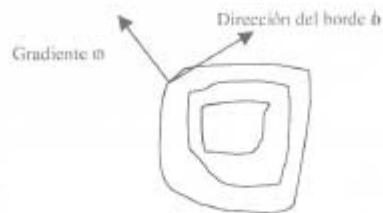


Figura 6.- Dirección de gradiente y dirección de borde

Operador Prewitt. Cálculo de gradiente

Este operador aproxima el cálculo de la primera derivada en la función imagen. El gradiente se calcula en ocho posibles direcciones utilizando una máscara de convolución de 3×3 , la dirección con el valor más alto determina la dirección del gradiente [7]. Las máscaras utilizadas son las siguientes:

$$h1 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad h2 = \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$$

$$h3 = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad h4 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix}$$

Selección automática del nivel de umbral

En un histograma de niveles de gris, las mayores frecuencias provienen de los píxeles pertenecientes a los objetos y fondo, siendo mínima la contribución de los niveles de gris de los píxeles de borde (píxeles con alto gradiente). Un método de detección automática de umbral denominado optimal thresholding consiste en determinar el nivel de gris correspondiente a la mínima probabilidad ubicada entre los máximos de dos distribuciones normales. Un algoritmo iterativo [4] que aproxima este método maximizando la varianza del nivel de gris entre objetos y fondo, es el que sigue:

1. Se asume un desconocimiento de la ubicación de los objetos dentro de la imagen y se considera como una primera aproximación, que las cuatro esquinas de la imagen contienen píxeles de fondo y el resto de la imagen contiene píxeles de objetos. Se determina un nivel de umbral utilizando las ecuaciones (7) y (8)
2. En el tiempo T, se calcula μ_B^1 y μ_0^1 , como la media de los niveles de gris de fondo y objetos respectivamente luego de haber realizado una segmentación utilizando el nivel de umbral T^t determinado en el paso anterior.

$$\mu_B^1 = \frac{\sum_{(y,x) \in \text{fondo}} f(y,x)}{\# \text{píxeles_fondo}} \quad (7)$$

$$\mu_0^1 = \frac{\sum_{(y,x) \in \text{objetos}} f(y,x)}{\# \text{píxeles_objetos}}$$

Se determina un nuevo valor de umbral

$$T^{(n+1)} = \frac{\mu_B^1 + \mu_0^1}{2} \quad (8)$$

- 3) Si $T^{(T-t)n} = T^t$, detener; de lo contrario retornar al paso 2.

Segmentación por histograma de gradiente para detectar bordes

En el histograma de valor absoluto de gradientes, existen dos principales contribuciones, una que agrupa los máximos valores de gradiente, correspondientes a píxeles de borde y la otra al resto de los píxeles de la imagen. Al aplicar segmentación, se obtiene una imagen con bordes gruesos. Adelgazamiento de bordes

Un método denominado non-maximal supresión [6], utiliza como datos de entrada, la magnitud y dirección del píxel de borde, obtenidos con algún operador de gradiente, antes de aplicar segmentación. Luego examina la magnitud de borde de los dos píxeles vecinos, perpendiculares a la dirección del borde, para cada píxel de la imagen ya segmentada por gradiente. Si al menos uno de estos píxeles tiene una magnitud de borde más alta que la del píxel bajo estudio, se le asigna valor cero a la magnitud del píxel examinado.

METODOLOGÍA

Instrumentos utilizados

- √ Lenguaje de computación técnica MATLAB 5.3.0,10183 (R11).1998. Toolbox de Imágenes versión 2.2. (R11) 1998.
- √ Imágenes digitalizadas de tejido histológico (true color 640 x 480)

Segmentación

La figura 7 representa los tres pasos aplicados a la imagen RGB original para lograr la segmentación de los núcleos.



Detección de Bordes

Se aplicó segmentación por histograma de gradientes del plano de brillo y de la imagen obtenida con segmentación por tinte, para obtener los primeros bordes. Los pasos seguidos se muestran en la figura 8.



DISCUSIÓN DE LOS RESULTADOS

Al comparar las segmentaciones obtenidas en las figuras 9 y 10, se observa que con segmentación por tinte desaparecen los restos de núcleos, que aparecen como pequeñas manchas en la segmentación por nivel de gris y los objetos presentan menos huecos. Con segmentación por nivel de gris, existe pérdida de data en los tres núcleos superpuestos del centro de la imagen, esto posteriormente se traduce en bordes errados. En ambas figuras se obtienen una extracción de los objetos bastante aceptable.

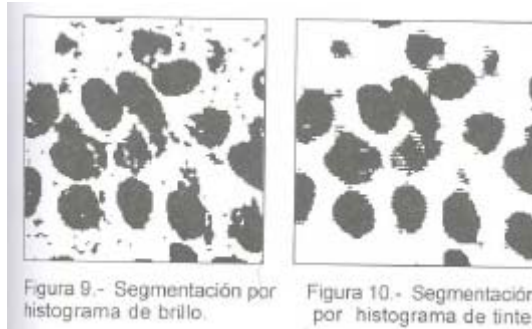


Figura 9.- Segmentación por histograma de brillo.

Figura 10.- Segmentación por histograma de tinte.

Los bordes obtenidos sobre la imagen segmentada por tinte, figuras 11 y 12, son una buena aproximación a los bordes esperados.

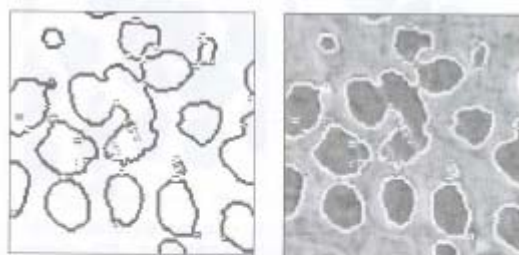


Figura 11.- Segmentación por gradiente de la imagen de segmentación por tinte. No se aplicó adelgazamiento de línea.

Figura 12.- Bordes trazados sobre la imagen original

Los bordes pertenecientes a núcleos muy próximos, están mejor definidos sobre la imagen segmentada por gradiente de brillo, figura 13. En cuanto a los núcleos superpuestos, como se mencionó en la discusión anterior, aparecen bordes que no se corresponden con los núcleos originales. El algoritmo de adelgazamiento de líneas, produce bordes muy bien definidos sin pérdida de información, figura 14.

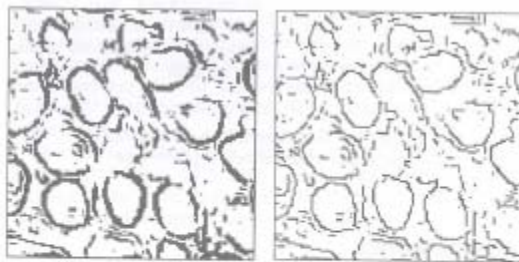


Figura 13.- Segmentación por gradiente de la imagen de brillo.

Figura 14.- Algoritmo de adelgazamiento de línea "non-maximal supresión" aplicado a la imagen de la figura 11.

En la figura 15, se puede observar la segmentación obtenida utilizando otros métodos [2j, [3]. La segmentación por histograma con selección automática de umbral, produce una mejor definición

de los contornos de los núcleos, esta característica es de gran importancia, puesto que el siguiente procesamiento sobre la imagen segmentada será trazar los bordes.

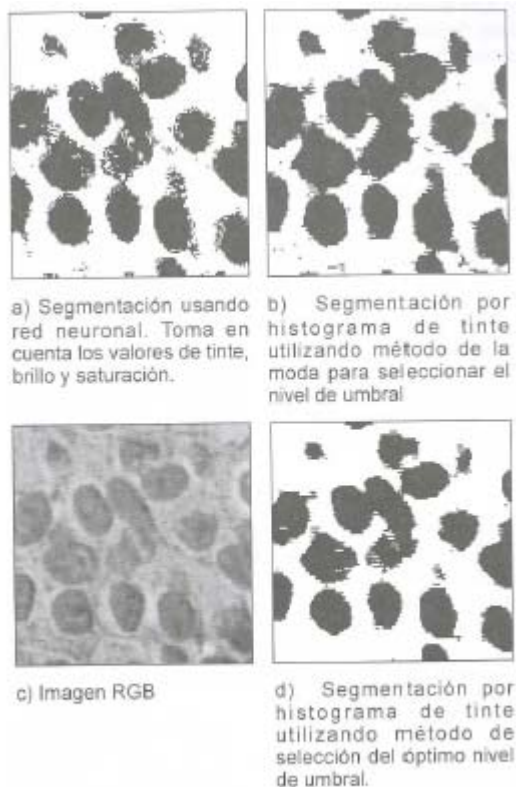


Figura 15.- Segmentación de núcleos aplicando tres métodos diferentes

CONCLUSIONES

El método de selección automática, del nivel óptimo de umbral para segmentación por histograma, facilita y produce muy buenos resultados en segmentación para este tipo de imágenes. Se comprobó que funciona perfectamente en los diferentes casos: tinte, brillo, gradientes. Es un método bastante sencillo que requiere muy poco tiempo de procesamiento.

RECONOCIMIENTOS

Este trabajo se ha desarrollado en el marco del proyecto ¡CM (Imagenología del cáncer de mama) con el financiamiento del Consejo de Desarrollo Científico Humanístico y Tecnológico de la Universidad de Carabobo y con el apoyo del Hospital Pérez Carreño.

REFERENCIAS

[1] BARRIOS, Victor, Costalat Steve, Reigosa Aldo, Montilla Guillermo (1997). Segmentación de imágenes celulares utilizando redes neuronales, Centro de Procesamiento de Imágenes. Facultad de Ingeniería. Facultad de Ciencias de la Salud. Universidad de Carabobo. Venezuela.

- [2] DE ABREU, María F., Barrios V., Reigosa A., (1997). Segmentación de imágenes RGB de cáncer de mama utilizando redes neuronales, Centro de Procesamiento de Imágenes. Facultad de Ingeniería. Universidad de Carabobo. Venezuela.
- [3] DE ABREU, María F. (1999). Detección, estandarización y medición de colores en imágenes de cortes histológicos procesados con la técnica de inmunohistoquímica. Tesis de maestría. Centro de Procesamiento de Imágenes. Facultad de Ingeniería. Universidad de Carabobo. Venezuela.
- [4] MOORE GW, Berman JJ, -Brown LA. (1994). Software for Image Segmentation and Analysis in Pathology (ISAP). U.S.Government Work, Presented at the 1994 meeting of the American Society of Clinical Pathology, Washington, DC.
- [5] Milan Sonka, Vaclav Hlavac, Roger Boyle. Image Processing, Analysis and Maching. Editorial Chapman and hall Computing, Chapter 5,1993.
- [6] Wyszecki G. and Stiles W.S.. (1982). Color Science, Concepts and Methods, Quantitative Data and Formulae. New York, Wiley.
- [7] Milan Sonka, Vaclav Hlavac, Roger Boyle. Image Processing, Analysis and Maching. Editorial Chapman and hall Computing, Chapter 4,1993.