

DetECCIÓN DE BANDAS EN COLOR EN UNA IMAGEN DE ELECTROFORESIS EN GEL DE UNA DIMENSIÓN USANDO UN ALGORITMO DE LOCALIZACIÓN BASADO EN MÁXIMOS Y MÍNIMOS

COLOR BANDS DETECTION ON A GEL ELECTROPHORESIS IMAGE IN ONE DIMENSION APPLYING A LOCATION ALGORITHM BASED ON MAXIMUMS AND MINIMUMS

Cesar Augusto Salazar-Centeno
Ingeniero Electrónico
Universidad Industrial de Santander
Bucaramanga, Colombia
cesar.salazar@correo.uis.edu.co

Carlos Andrés Niño-Niño
M.Sc. Ingeniería Eléctrica
Universidad Manuela Beltrán
Bucaramanga, Colombia.
carlos.nino@docentes.umb.edu.co

Ricardo Andrés Díaz-Suárez
M.Sc. Ingeniería Eléctrica
Universidad Santo Tomás
Bucaramanga, Colombia
ricardo.diaz@ustabuca.edu.co

Resumen– En este artículo se escribe una metodología desarrollada en MATLAB para el procesamiento de imágenes de electroforesis en gel de una dimensión las cuales contienen información de antígenos de insectos (*Triatoma Dimidiata* o *Pito*) y de parásitos (*Fasciola*). La metodología propuesta permite complementar la interpretación visual de las imágenes basándose en técnicas de contraste, realce de cada una de las bandas y la implementación de filtros para eliminar discontinuidades; adicionalmente se propone un algoritmo de máximos y mínimos para la identificación automática de cada una de las bandas. Como resultado final se obtuvieron 5 imágenes a color, ubicando e identificando con una eficiencia del 100% las bandas características de la electroforesis en gel en comparación con sus patrones originales, complementado con un análisis cuasi-cuantitativo de la cantidad de muestra de cada una de ellas en relación con la banda con mayor concentración.

Palabras claves– Algoritmo, contraste, electroforesis en gel, filtrado, máximos y mínimos.

Abstract– This article describes a developed methodology on MATLAB for the images processing of the gel electrophoresis image in one dimension which contain information of insect antigens (*Triatoma Dimidiata* or *Pito*) and parasites (*Fasciola*). The proposed methodology permits to complement the visual interpretation

of the images basing on contrast technics, heighten each of the bands and the implementation of filters to eliminate discontinuities; additionally, it's proposed an algorithm of maximums and minimums for the automatic identification of each band. As a final result it's obtained 5 color images, locating with a 100% efficiency the characteristic bands of the gel electrophoresis in comparison with their original patterns, complementing with a quasi-quantitative analysis of the sample amount of each of them in relation with the highest concentration band.

Keywords– Algorithm, contrast, filtered, gel electrophoresis, maximums and minimums.

Palabras claves– Algoritmo, contraste, electroforesis en gel, filtrado, máximos y mínimos.

1. INTRODUCCIÓN

La electroforesis en gel es el método más sencillo para la obtención de información sobre macro-moléculas como proteínas, enzimas y ácidos nucleicos (ADN y ARN), en el que las moléculas se redistribuyen en función de su forma y tamaño, generando conjuntos de bandas características. Estas bandas son resaltadas y finalmente captu-



radas como una imagen para su estudio [1]. Las macro-moléculas se separan utilizando una diferencia de potencial eléctrico, provocando un desplazamiento basado en su masa y carga eléctrica que adicionalmente por fenómenos de fricción y difusión producidos por cambios de temperaturas, no es homogéneo [2], [3].

Los problemas más importantes en la interpretación de las imágenes de electroforesis en gel de una dimensión se presentan por el bajo contraste, altos niveles de ruido y variaciones de intensidad en los niveles de grises contenidas en las mismas [4]. Una solución para este problema fue dado por Esquivel y Randall (2012) [5] quienes abarcaron el problema centrándose en las bandas características ubicadas en los carriles de la electroforesis, basándose en los parámetros de su función objetivo tanto en la estimación inicial de la desviación estándar como en la posición inicial de las bandas de un carril, a través del método de optimización *Downhill Simplex* el cual minimiza la función objetivo en cuestión.

Por otra parte, Taher et al. (2013) [6] corrigieron las imágenes utilizando dos tipos de filtros, los cuales fueron implementados en el dominio espacial y frecuencial. Para la comparación de estos se utilizó la evaluación de calidad *The Pick Signal-to Noise Ratio* (PSNR), dando como resultado una mejora en la detección de las bandas características con el filtro espacial. Otra solución al problema fue dada por Intarapanich et al. (2015) [7], donde se presentó una segmentación en las líneas de las bandas características para resolver la distorsión, deformidad, niveles de ruido de fondo y las bandas que se encuentran muy cercanas (dobletes), llegando a su detección de manera automática. Hlavatý (2014) [8] segmentó las líneas de las imágenes en escala de grises, obteniendo un algoritmo de detección automática constituido por segmentaciones horizontales que proyectaban un perfil vertical en el cual, tanto el mínimo local como el borde de cada uno de ellos son localizados, interpolando los puntos de las particiones en los bordes vecinos. De igual manera Ismail et al. (2014) [9] presentaron un sistema de detección y segmentación automática en las líneas de la electroforesis en gel, en las cuales se alcanzó

una precisión del 99.5% en imágenes de buena calidad.

E. Leal y N. Leal (2010) [10] combinaron técnicas de visión artificial y redes neuronales de pulso acoplado (PCNN) en las que lograron establecer un sistema de lectura automática, obteniendo mayor precisión en la clasificación fenotípica de un individuo. La Serna Palomino y Pró Concepción (2010) [11] utilizaron el algoritmo *Watershed* como método de segmentación de las imágenes, basándose en sus características morfológicas como texturas homogéneas y baja intensidad. Skutkova et al. (2013) [12] solucionaron la distorsión y la poca uniformidad en las imágenes, haciendo una compensación en las direcciones de los ejes X y Y, generando un algoritmo de detección en las tendencias de las muestras en escala de grises.

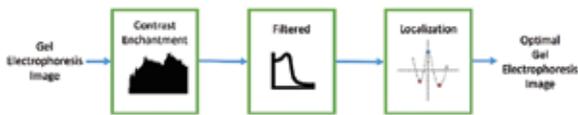
A continuación se presenta una metodología para el procesamiento de imágenes de electroforesis en gel de una dimensión, la cual consiste en detectar cada una de las bandas a partir de la fase de pre-procesamiento, compuesta por una mejora de contraste centrándose tanto en la luminosidad como en el filtrado para suavizar las bandas en la imagen. Seguido por una fase de desarrollo y la ejecución del algoritmo basado en máximos y mínimos para lograr ubicar cada banda característica. Continuando con el resultado final de las diferentes imágenes de electroforesis al procesamiento y, por último, las conclusiones más importantes de esta investigación, las cuales hablan sobre la importancia de la etapa de contraste al generar un incremento en un 200% de las bandas características con menor intensidad, la eficiencia del 100% del algoritmo de localización basado en máximos y mínimos en comparación con los patrones originales al ubicar las bandas, y la obtención del valor aproximado de la cantidad de muestra en cada banda característica.

2. DESARROLLO DEL ARTÍCULO

La metodología presentada como procesamiento de la imagen considera tres aspectos, los cuales se muestran en la Fig. 1. Un pre-procesamiento que comprende las etapas de mejora de contraste y filtrado en la imagen, seguido por la apli-

cación del algoritmo de localización basado en máximos y mínimos.

Fig. 1. DIAGRAMA DE BLOQUES DEL PROCESAMIENTO DE LA IMAGEN

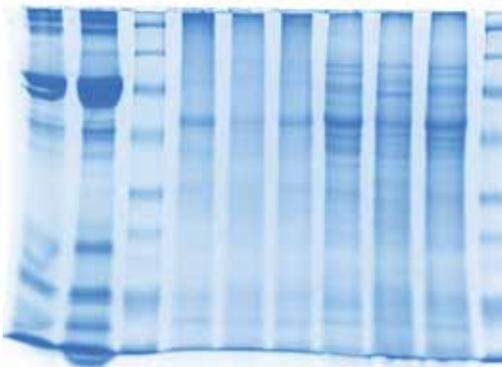


Fuente: Los autores.

El algoritmo propuesto surge frente a la necesidad de poder localizar con facilidad y de manera precisa las bandas características en las imágenes de electroforesis en gel de una dimensión, centrándose en las bandas con problemas de bajo contraste, altos niveles de ruido y variaciones de intensidad en sus niveles de grises, con el objetivo de conseguir una mejor interpretación de las mismas. Este algoritmo utiliza la información suministrada por las bandas características de la imagen de electroforesis para generar una señal de bandas, en la cual se ubican los máximos y mínimos de cada una de ellas, creando grupos de 2 mínimos y 1 máximos para la localización de cada banda.

Para la implementación, tanto del procesamiento de imágenes como del algoritmo basado en máximos y mínimos, se utilizaron las imágenes de electroforesis en gel en una dimensión, suministradas por el Grupo de Inmunología y Epidemiología Molecular (GIEM) adscrito a la Universidad Industrial de Santander (UIS). Estas imágenes se obtuvieron a través del método de electroforesis en gel de agarosa, conteniendo información de antígenos de insectos (*Triatoma Dimidiata* o *Pito*) y de parásitos (*Fasciola*) como se observa en la Fig. 2.

Fig. 2. IMAGEN DE ELECTROFORESIS EN GEL DE UNA DIMENSIÓN DE LA PROTEÍNA DE HEMOLINFA DE UN *RHODNIUX PROLIXUS*



Fuente: Grupo de Inmunología y Epidemiología Molecular (GIEM).

2.1 Pre-procesamiento de la imagen de electroforesis

La imagen de electroforesis en gel de una dimensión viene representada en el comúnmente espacio de color RGB, este se personifica en la Fig. 3 por un cubo en coordenadas rectangulares donde cada eje simboliza la intensidad de los colores rojo (R), verde (G) y azul (B); el color negro se interpreta como la ausencia de estos colores y el blanco como el máximo valor de cada uno de ellos [13].

La etapa de contraste se realizó en el espacio de color $L^*a^*b^*$ por su facilidad de separar el brillo del resto de la imagen. En la Fig. 4 se presenta este espacio de color el cual consiste en una esfera donde sus componentes se definen como luminosidad (L^*) en el eje Z, las posiciones tanto de magenta ($-a^*$) y verde ($+a^*$) en eje X como de amarillo ($+b^*$) y azul ($-b^*$) en el eje Y [14]. Por esta razón se ejecutó un cambio de RGB a $L^*a^*b^*$ en el cual se utilizó *Contrast Limited Adaptive Histogram Equalization* (CLAHE) [15] en la componente L, generando una mejora en la luminosidad de la imagen de electroforesis, retornando nuevamente al espacio RGB para la ejecución de la siguiente etapa.

El objetivo principal del filtrado es suavizar y eliminar el ruido presente en las imágenes de electroforesis en el momento de su captura, especialmente la información contenida en las bandas características, por esta razón el filtro más cercano en características frecuenciales al ideal en banda de paso, banda de rechazo y pendiente sería el *Butterworth* en comparación al *Chebyshev*, *Gaussiano* y *Cauer* [16]; este se identifica por tener una respuesta plana en la banda de paso, hasta la frecuencia de corte $2\pi f_c$; una pendiente pronunciada a partir de esta frecuencia con pendiente de $20ndB$ por década, siendo n el orden del filtro, y una respuesta en su banda de rechazo cercana a cero [17].

El filtro *Butterworth* bidimensional, para aplicaciones en procesamiento de imágenes, se define como (1)

$$H(x,y) = \left(1 + \left[\frac{D(x,y)}{D_0}\right]^{2n}\right)^{-1} \quad (1)$$

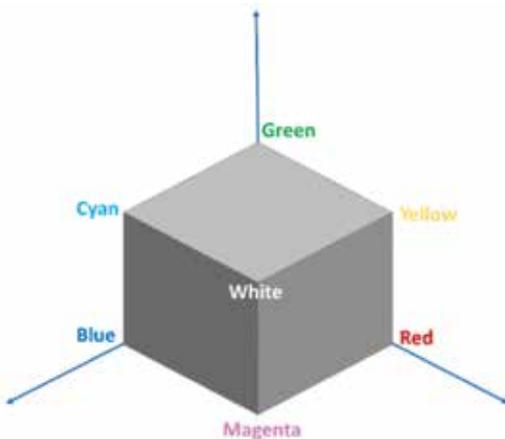
donde $H(x,y)$ se relaciona con la función distancia ($D(x,y)$), la distancia de corte (D_0), y el orden del filtro (n) [18].

El filtrado de la imagen se implementó en el dominio de la frecuencia, realizando el producto de las transformadas rápidas de Fourier bidimensionales tanto de la imagen como del filtro. En esta etapa y luego de diferentes pruebas con variaciones de D_0 y n se compararon los resultados visuales y se seleccionó el filtro pasa bajas *Butterworth* bidimensional con las siguientes especificaciones: una distancia de corte de $D_0 = 60$ unidades espaciales, y un orden de cinco ($n = 5$); cumpliendo con los requerimientos para generar homogeneidad en las bandas características.

2.2 Localización de las bandas características

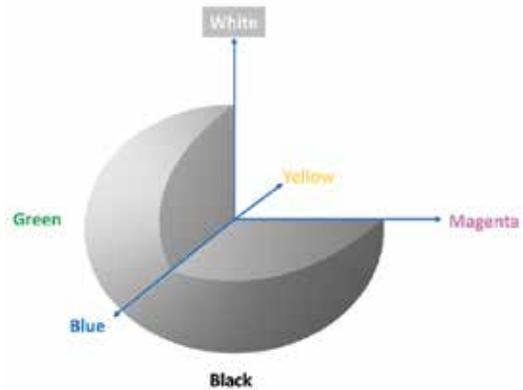
Para la localización de las bandas se desarrolló el siguiente procedimiento: En la primera fase se realizó la ubicación del centro de las bandas características separando la imagen en sus canales RGB, tomando como base el promedio de los canales para generar un vector vertical centrado en la columna de las bandas. En la segunda fase se implementó una transformación de la imagen al espacio de color HSV, ya que este tiene una representación geométrica en la Fig. 5 que toma la forma de un cono de base cuasi-hexagonal, donde cada color se relaciona con los componentes, tales como el matiz (o la clase de color), la saturación (o la pureza) y la intensidad (o brillo) [13].

Fig. 3. ESPACIO DE COLOR RGB



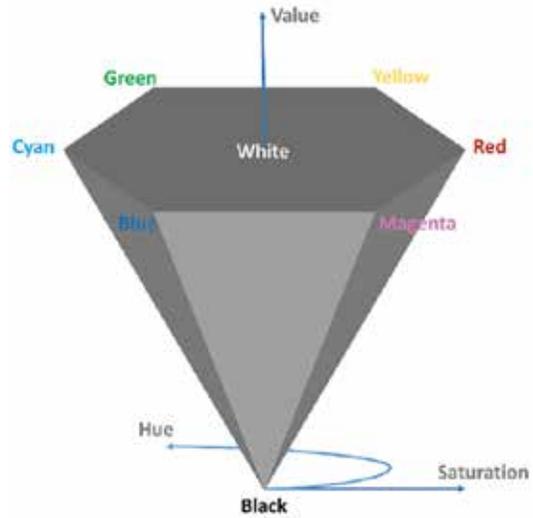
Fuente: Los autores.

Fig. 4. ESPACIO DE COLOR L*a*b*



Fuente: Los autores.

Fig. 5. ESPACIO DE COLOR HSV

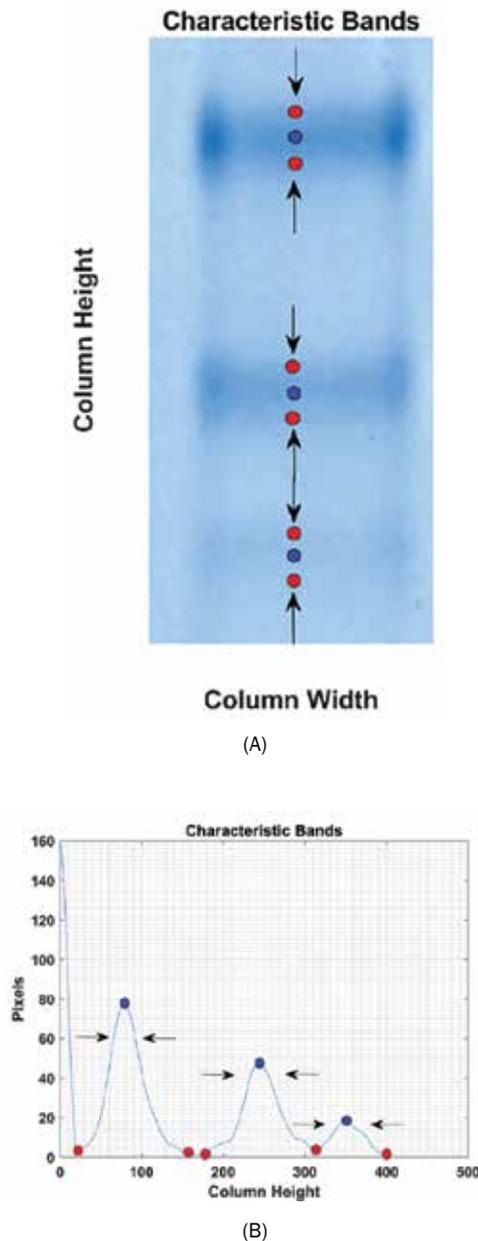


Fuente: Los autores.

En esta parte se utilizó la componente H para distinguir los colores de las diferentes imágenes de electroforesis y se retorna a RGB. En la tercera fase se proyectaron las bandas en función de la altura de la columna y el valor de cada pixel (Señal de Bandas), ubicando sus máximos (Azul) y mínimos (Rojo) con la función *findpeaks* de MATLAB e invirtiendo la señal de bandas en el caso de los mínimos, y así localizando cada banda en la señal, Fig. 6. En la cuarta fase se eliminó el ruido en la señal de bandas con un filtro pasa bajas; adicionalmente se utilizaron tanto los máximos como los mínimos para generar una aproximación triangular de cada banda, de esta forma se eliminan las bandas más pequeñas considerando los perímetros. En la última fase se localizaron las bandas características en la imagen, usando los

máximos como centros de cada banda, generando un recorte rectangular alrededor de estas, y con ello resaltándolas.

Fig. 6. (A) IMAGEN DE ELECTROFORESIS EN GEL DE 3 BANDAS Y (B) SEÑAL DE BANDAS (BANDAS CARACTERÍSTICAS)



Fuente: Los autores.

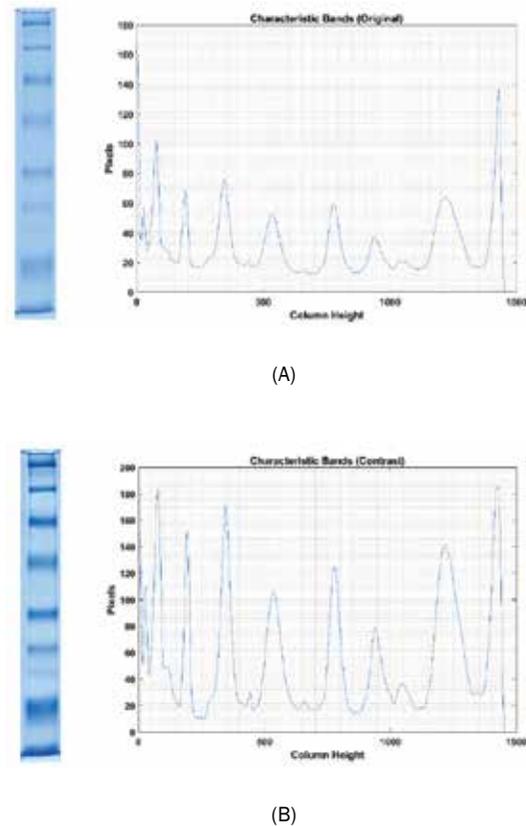
3. RESULTADOS

3.1 Pre-procesamiento de la imagen de electroforesis

Como resultado al pre-procesamiento se obtiene una mejora en contraste donde se resalta las ban-

das características de la imagen de electroforesis en gel, i.e. una variación de pixeles en la banda de 87, 165, 250 a 0, 111, 190 en sus canales R, G y B respectivamente, mejorando la percepción para el observador como se puede apreciar en la Fig. 7. Adicionalmente, comparando las señales de bandas se aprecia el incremento de cada banda en 42.16% para la de mayor intensidad y 219.05% para la de menor intensidad.

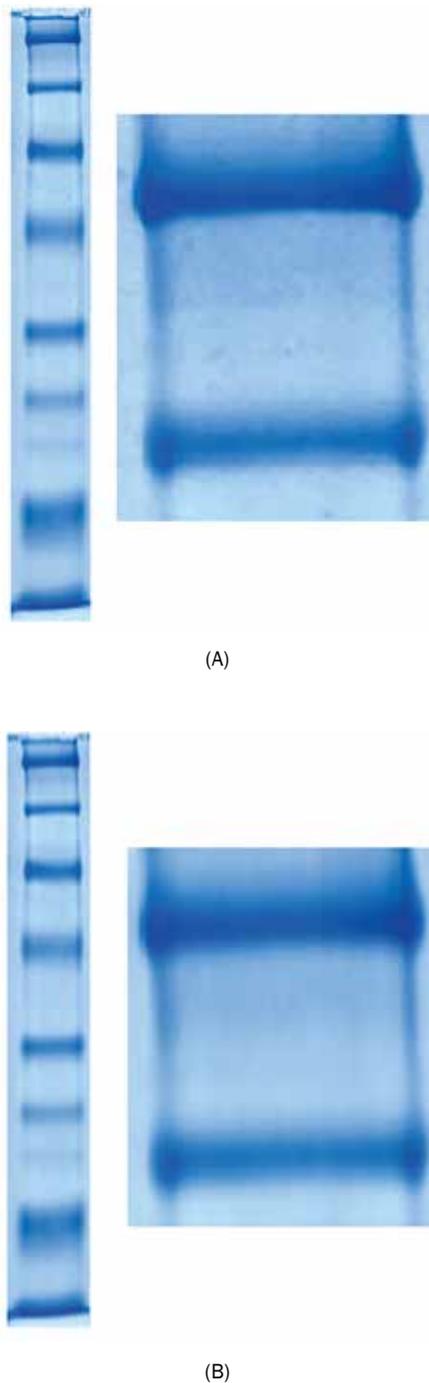
Fig. 7. (A) IMAGEN DE ELECTROFORESIS EN GEL ORIGINAL E (B) IMAGEN DE ELECTROFORESIS EN GEL CON MEJORA DE CONTRASTE



Fuente: Los autores.

En la Fig. 8 se observa el resultado final de la etapa del pre-procesamiento de la imagen; el filtrado, obteniendo una imagen de electroforesis en gel totalmente suavizada, eliminando su ruido visual. Adicionalmente, en esta se perciben las bandas características más compactas y homogéneas, las cuales ayudan a la proyección de la señal de bandas promedio; con un valor de los pixeles en la banda de 0, 111, 190 en sus canales R, G y B, respectivamente.

Fig. 8. IMAGEN DE ELECTROFORESIS EN GEL CON MEJORA DE CONTRASTE E (B) IMAGEN DE ELECTROFORESIS EN GEL FILTRADA



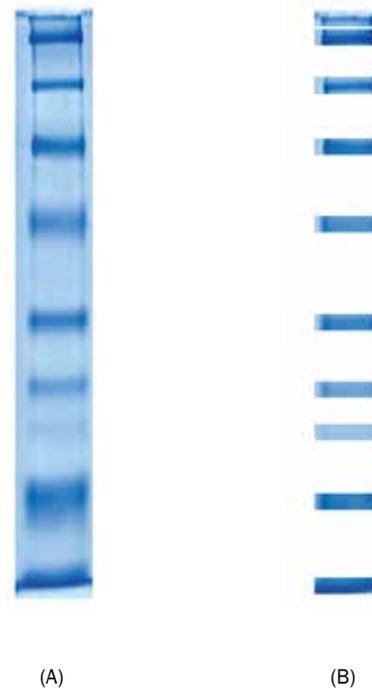
Fuente: Los autores.

3.2 La localización de las bandas características

El resultado final del procesamiento de imágenes se representa en la Fig. 9, donde se ubicaron cada una de las bandas y se elimina la información no

necesaria, obteniendo únicamente las bandas características en la imagen de electroforesis en gel.

Fig. 9. (A) IMAGEN DE ELECTROFORESIS EN GEL FILTRADA Y (B) LOCALIZACIÓN DE BANDAS EN LA IMAGEN DE ELECTROFORESIS EN GEL



Fuente: Los autores.

3.3 Resultados obtenidos para imágenes de electroforesis en gel

Como complemento se presentan los resultados obtenidos en cada una de las etapas llevadas a cabo durante la investigación Tabla I. Para esto, se emplearon las imágenes de electroforesis en gel de una dimensión de las proteínas de los antígenos *Rhodnius Prolixus* y *Fasciola*. Realizando una comparación entre la imagen original con la señal de bandas del proceso final, se extrae la información de las bandas características relacionando cada una con la banda de mayor intensidad (concentraciones normalizadas), ya que en esta se encuentra la mayor parte de la concentración de la muestra corrida en la electroforesis en gel, obteniendo un análisis cuasi-cuantitativo en relación a su máximo, con ello se logra una comparación referente ya sea al estudio efectuado o a las razones por las cuales se ejecutó la electroforesis en gel.

TABLA I
 RESULTADOS DE LAS ETAPAS DEL PROCESAMIENTO DE LAS IMÁGENES DE ELECTROFORESIS EN GEL

	ORIGINAL	CONTRASTE	FILTRADO	LOCALIZACIÓN	CONCENTRACIONES NORMALIZADAS
PROTEÍNAS DE HEMOLINFA DEL ANTÍGENO RHODNIUX PROLIXUS					87.36% 78.59% 99.77% 94.32% 52.02% 25.97% 24.39% 35.13% 58.98% 42.52% 100%
PROTEÍNAS DE HEMOLINFA DEL ANTÍGENO RHODNIUX PROLIXUS					95.80% 82.12% 33.48% 100% 68.66% 25.52% 14.76% 63.46% 70.87% 60.60% 99.25%
PROTEÍNAS DE HEMOLINFA DEL ANTÍGENO RHODNIUX PROLIXUS					62.22% 84.32% 67.95% 71.10% 60.72% 63.79% 53.34% 45.67% 65.76% 100%
PROTEÍNAS DE HEMOLINFA DEL ANTÍGENO RHODNIUX PROLIXUS					80.28% 55.34% 61.89% 44.73% 48.73% 32.24% 16.93% 51.87% 100%
PROTEÍNAS DEL ANTÍGENO FASCIOLLA					46.55% 74.02% 59.81% 70.03% 53.24% 61.98% 44.93% 29.26% 59.74% 100%

Fuente: Los autores.

4. CONCLUSIONES

Los resultados finales obtenidos al procesamiento son unas imágenes de electroforesis en gel de una dimensión con únicamente sus bandas características, facilitando su análisis comparativo con los patrones utilizados para evaluar la cantidad de ADN o ARN contenidos en la muestra.

Una de las etapas más importantes en el pre-procesado de la imagen fue el contraste, ya que en este se logró resaltar e incrementar las bandas características en alrededor de un 200% para las bandas con menor intensidad, ayudando a su identificación con el algoritmo basado en máximos y mínimos.

Con el algoritmo de localización basado en máximos y mínimos se identificaron con una eficiencia del 100% en comparación a sus patrones originales, las posiciones de cada una de las bandas características, tanto en la señal de bandas como en la imagen de electroforesis en gel de una dimensión, logrando identificar el número de bandas y la banda con mayor intensidad en las cinco imágenes trabajadas.

Se obtuvo un valor aproximado de la cantidad de muestra existente en cada banda característica, realizando un análisis cuasi-cuantitativo relacionando la imagen original con la señal de bandas del proceso final, permitiendo la comparación referente al estudio efectuado o a las razones por las cuales se ejecutó la electroforesis en gel.

AGRADECIMIENTOS

Se agradece la colaboración de la doctora Marta Lucía Díaz por la gestión de las imágenes de electroforesis en gel de una dimensión y a la microbióloga Laura Elise Corzo por su guía y explicación de los conceptos referentes a la electroforesis en gel.

REFERENCIAS

[1] P. Y. Lee, J. Costumbrado, C.-Y. Hsu, and Y. H. Kim, "Agarose Gel Electrophoresis for the Separation of DNA Fragments," *J. Vis. Exp.*, no. 62, 2012.

[2] N. C. Stellwagen, "Electrophoresis of DNA in agarose gels, polyacrylamide gels and in free solution," *Electrophoresis*, vol. 30, no. SUPPL. 1, 2009.

[3] Á. G. Hernández, "Metabolismo lipídicos, dislipemias," in *Principios de bioquímica clínica y patología molecular*, 2014, pp. 170-178.

[4] K. Jonak, K. Jędrasiak, A. Polański, and K. Puszyński, "Application of Image Processing Algorithms in Proteomics: Automatic Analysis of 2-D Gel Electrophoresis Images from Western Blot Assay," in *Computer Vision and Graphics*, 2012, pp. 433-440.

[5] A. Esquivel and J. Randall, "Optimización de la detección de cantidad, ancho y posición de bandas en imágenes de geles de electroforesis," *Hdl.handle.net*, 2012.

[6] R. S. Taher, N. Jamil, S. Nordin, F. H. Yusof, and U. M. Bahari, "Poor DNA gel electrophoresis image enhancement: Spatial vs. frequency domain filters," in *Proceedings - 2013 IEEE Conference on Systems, Process and Control, ICSPC 2013*, 2013, pp. 175-180.

[7] A. Intarapanich, S. Kaewkamnerd, P. J. Shaw, K. Ukosakit, S. Tragoonrung, and S. Tongshima, "Automatic DNA Diagnosis for 1D Gel Electrophoresis Images using Bio-image Processing Technique," *BMC Genomics*, vol. 16 Suppl 1, p. S15, 2015.

[8] M. Hlavatý, "Preprocessing of 1D Gel Electrophoresis Image," *Proc. 20th Conf. STUDENT EEICT 2014*, vol. 1, no. 1, pp. 61-63, 2014.

[9] I. Ismail, G. S. Eltaweel, and H. Nassar, "Bands detection and Lanes segmentation in DNA Fingerprint images," *J. Inf. Comput. Sci.*, vol. 9, no. 4, pp. 243-251, 2014.

[10] E. Leal and N. Leal, "Automatización de la Prueba HLA Mediante Análisis de Imágenes de Gel de Electroforesis Empleando Visión Artificial," *Prospect*, vol. 1, no. 1, pp. 13-20, 2010.

[11] N. La Serna Palomino and L. E. Pró Concepción, "Watershed: un algoritmo eficiente y flexible para segmentación de imágenes de geles 2-DE," *Rev. Investig. Sist. e Informática*, vol. 7, no. 2, pp. 35-41, 2010.

[12] H. Skutkova, M. Vitek, S. Krizkova, R. Kizek, and I. Provaznik, "Preprocessing and classification of electrophoresis gel images using dynamic time warping," *Int. J. Electrochem. Sci.*, vol. 8, no. 2, pp. 1609-1622, 2013.

[13] T. Hassanzadeh, H. Vojodi, and A. M. E. Moghadam, "An image segmentation approach based on maximum variance Intra-cluster method and Firefly algorithm," in *Proceedings - 2011 7th International Confe-*

- rence on Natural Computation, *ICNC 2011*, 2011, vol. 3, pp. 1817-1821.
- [14] M. Wirth and D. Nikitenko, "The effect of colour space on image sharpening algorithms," *CRV 2010 - 7th Can. Conf. Comput. Robot Vis.*, pp. 79-85, 2010.
- [15] K. Zuiderveld, "Contrast Limited Adaptive Histogram Equalization," in *Graphics Gems*, 1994, pp. 474-485.
- [16] J. Kuriakose, S. Ristic, and G. de Cremoux, "An automated toolbox for the design and analysis of lattice wave digital filters using MATLAB," in *IEE Colloquium on DSP enabled Radio*, 2003, pp. 1-8.
- [17] J. W. Nilsson, "Filtros Activos," in *Circuitos Eléctricos*, 2005, pp. 754-757.
- [18] P. Y. Tan and H. Ibrahim, "Implementation of Band Pass Filter for Homomorphic Filtering Technique," *Int. J. Comput. Sci. Mob. Appl.*, vol. 1, no. 5, pp. 1-6, 2013.