

Morfometría en diferentes variedades de semillas de higuera mediante procesamiento de imágenes

María Acosta, José Padilla, José Botello, Joel Quintanilla, José Barrón, Jesús Espinoza y Salvador Montes

M. Acosta, J. Padilla, J. Botello, J. Quintanilla, J. Barrón, J. Espinoza, S. Montes
Universidad Tecnológica del Suroeste de Guanajuato, Depto. de Ing. en Procesos Alimentarios, Carr. Valle-Huanimaro Km 1.2, Valle de Santiago, Guanajuato, México.
Instituto Tecnológico de Celaya, Depto. de Ing. Electrónica, Av. Tecnológico esq. Av. García Cubas/n, Celaya, Guanajuato, 38000, México.
Instituto Tecnológico de Celaya, Depto. de Bioingeniería, Av. Tecnológico esq. Av. García Cubas s/n, Celaya, Guanajuato, 38000, México.
Universidad Tecnológica del Suroeste de Guanajuato, Depto. de Ing. en Tecnologías de la Información, Carr. Valle-Huanimaro Km 1.2, Valle de Santiago, Guanajuato, México.
Instituto Nacional de Investigación Forestal y Agrícola Bajío, Carretera Celaya-San Miguel de Allende, Celaya, Guanajuato.
msacosta@utsoe.edu.mx

Abstract

This paper presents by means of image processing the morphological characterization of six collections of seed higuera (*Ricinus communis* L) originate in the state of Guanajuato, Mexico. Image processing is implemented by two image segmentation techniques.

The first segmentation techniques implemented is the k-mean algorithm and the other one is the proposed methodology in this paper. The obtained measurements in the experiments were area, perimeter, Minor Axis Length (MiAL), Major Axis Length (MaAL), Enlarged Index (EI) and Feret's diameter (DF). Data were analyzed statistically using JMP (α 0.5) program. The obtained measurement results were: area 0.81-5.7 mm², perimeter 3.54-8.52 mm, MiAL 0.060-1.96 mm, MaAL 0.70- 2.89 mm, Enlarged Index 0.08-0.18 and Feret diameter de 1.0-5.53 mm respectively. It is important to mention that proposed model in Nh3 and Nh4 varieties presents the best results.

9 Introducción

La planta de higuera (*Ricinus communis* Lineo) tiene bajos requerimientos agronómicos y gran capacidad de adaptación a diferentes entornos con amplios beneficios por su capacidad de captación de CO₂ y fitoremediación de áreas contaminadas por Plomo (Pb) y Cadmio (Cd) (Hoeneisen y Silva, 1986; De Souza Costa *et al.*, 2011); además de crear fenotipos especiales para cada lugar en donde se desarrollan (Akpan *et al.*, 2006).

El cultivo de la higuera resulta interesante por la producción de aceite como materia prima renovable en diversas aplicaciones a nivel industrial para la obtención de lubricantes, dispersantes, humectantes, resinas y monómeros para poliuretanos biodegradables (Gil *et al.*, 2007) y la generación de residuos lignocelulósicos del cultivo (Amro y Soheir., 2009).

El procesamiento de imágenes tiene diferentes aplicaciones industriales en áreas como cosecha, poscosecha y procesamiento de materia prima.

En estas áreas, es necesario conocer la dimensiones morfométricas de los objetos (como frutas, granos, cereales y oleaginosas) con geometrías definidas o atípicas. Un descriptor de una especie o genotipo es una carácter fácil de medir, y que hace referencia a la forma y estructura de una accesión (López-Santiago *et al.*, 2008).

Actualmente existen pocas investigaciones sobre las propiedades y características de la planta de higuera que puedan ser utilizadas para caracterizar los fenotipos de cada región en México; debido a esto Fundación Guanajuato Produce está apoyando la transferencia de dicho cultivo a través del proyecto "Transferencia de tecnologías para la producción sustentable de biocombustibles en Guanajuato", mediante parcelas demostrativas en terrenos productores cooperantes, y talleres de producción artesanal de biocombustible; sin embargo no se cuenta por el momento una base de datos completa por estado.

México posee gran diversidad genética, manifestada por la alta distribución, condiciones edáficas y climáticas donde prospera de manera silvestre, y no obstante es poca la información que respecto a ella existe, ha sido estudiada principalmente en el estado de Chiapas, en donde (Goytia *et al.*, 2011) muestran accesiones con diferencia en peso, tamaño, forma y color.

El objetivo de este trabajo fue realizar un estudio morfométrico, comparando dos sistemas para el procesamiento de imágenes *k-mean* y un algoritmo *propio* como herramientas de medición en accesiones de higuera del estado de Guanajuato con la finalidad de contribuir al conocimiento mediante la generación de información de la diversidad específica.

9.1 Materiales y métodos

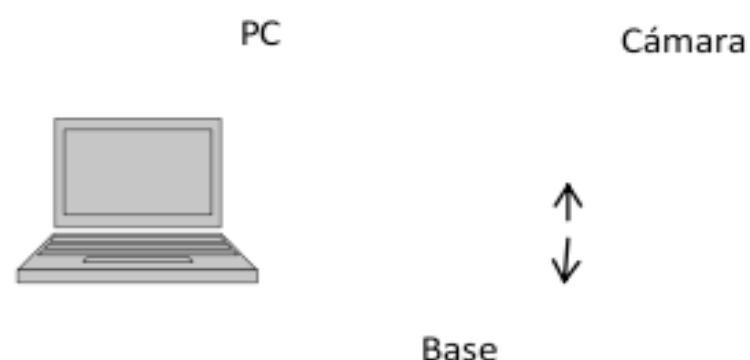
Las accesiones de semilla (Nh 1, 2, 3, 4, 10 y 12) fueron proporcionadas por el Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias (INIFAP-Bajío) de un consorcio seleccionado por alto rendimiento agronómico y contenido porcentual de aceite.

Procesamiento de imagen

Se realizó en 100 semillas para cada variedad bajo distancia focal e intensidad luminosa constantes, se utilizó una plataforma blanca a una distancia de 29 cm mediante una cámara microtouch (Dinolite digital microscope) como sistema óptico acoplada a una computadora personal, como se muestra en la figura 9.

Las imágenes capturadas fueron en RGB (rojo, verde, azul) con una resolución de 640 x 480 píxeles, se tomó como referencia un objeto de 2.1 cm, en base a lo anterior se determinó la equivalencia pixel/longitud y pixel/área que corresponden a 21/220 (0.095 mm) y 3.46/38013 ($9.102 \times 10^{-5} \text{ cm}^2$) respectivamente.

Figura 9 Esquema de toma de fotografía de las variedades de semilla



Este procedimiento permitió obtener imágenes binarizadas y delinear los bordes de las semillas, para evaluar sus dimensiones. La caracterización dimensional de las accesiones se realizó mediante la secuencia mostrada en la figura 9.1, que representa el procesamiento de imagen en el *modelo propio*.

Figura 9.1 Secuencia para el procesamiento de imagen de algoritmo propuesto



Con la finalidad de realizar una comparación del modelo propuesto, se utiliza un algoritmo de agrupamiento particional para realizar la segmentación de las imágenes, este algoritmo es conocido como *k-means*. El objetivo de los algoritmos de agrupamiento es encontrar grupos de píxeles que tienen características similares (intensidad, textura, etc.), y que en la imagen representan regiones homogéneas. El *k-means* fue presentado por (MacQueen, 1967) y de manera resumida la principal tarea de este algoritmo es hacer una partición estricta del espacio de las características. El *k-means* ha sido ampliamente utilizado en diferentes aplicaciones de segmentación de imágenes entre las cuales se pueden mencionar, segmentación de imágenes de mamografía con el objetivo de detectar microcalcificaciones (Quintanilla-Dominguez *et al.*, 2011, 2009) o masas (Ojeda-Magaña *et al.*, 2011) los cuales son signos de cáncer de mama. Otra aplicación de este algoritmo ha sido en imágenes de suelo con el propósito de encontrar regiones porosas (Cortina-Januchs *et al.*, 2011). En (Ojeda-Magaña *et al.*, 2010) se ha utilizado el *k-means* con la finalidad de obtener un grado de madurez de algunos frutos.

9.2 Caracterización morfológica de la semilla

Las variables a evaluar fueron: área (a), eje menor ($MiAL$), eje mayor ($MaAL$), Índice de alargamiento ($EI=MaAL/MiAL$) y Diámetro de Feret ($DF=\sqrt{(4-Area)/\pi}$) para cada semilla (Wilcox *et al.*, 2002); también se obtuvo promedio y desviación estándar de 100 semillas por variedad.

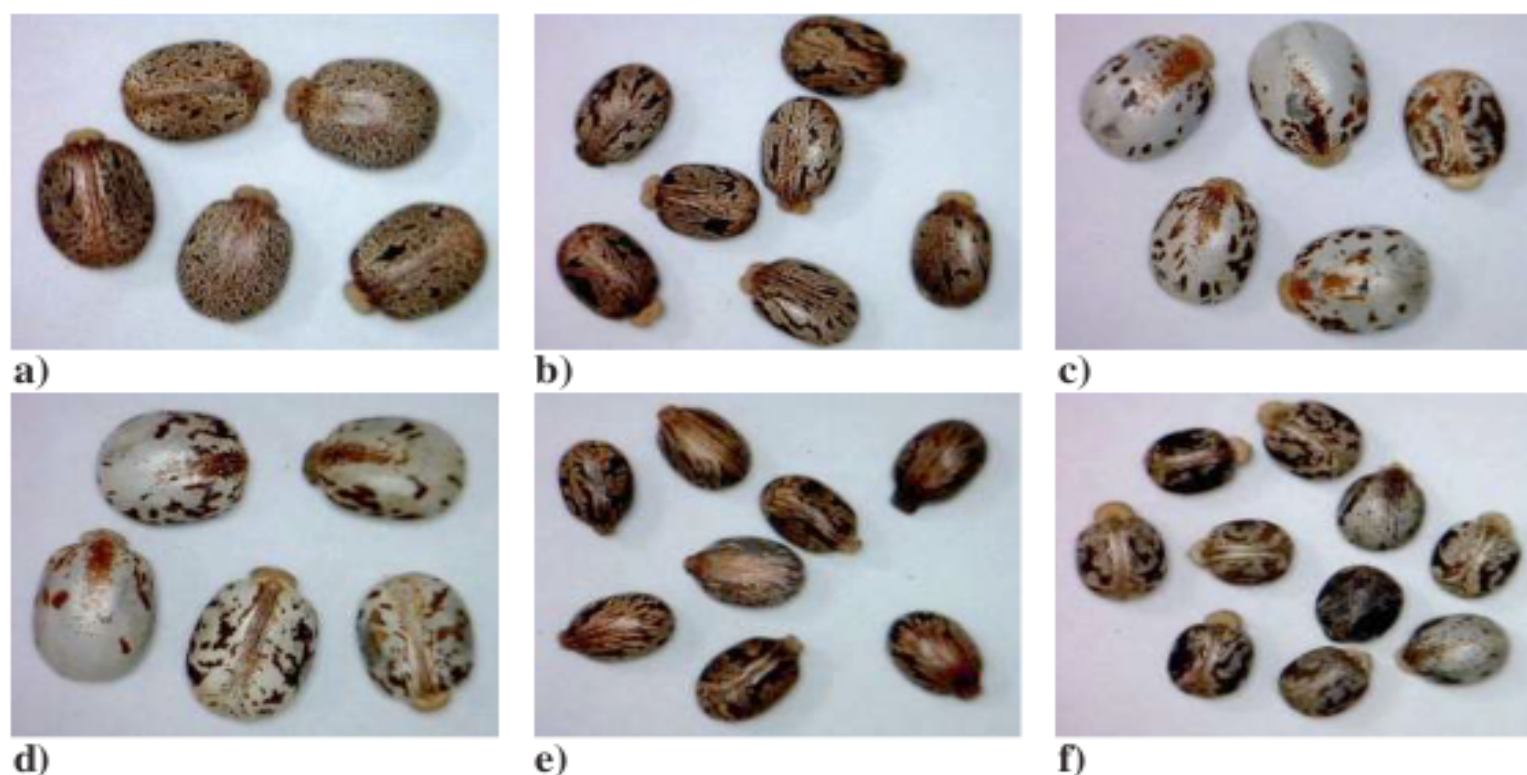
Análisis de datos

Se realizó un análisis de varianza para cada una de las variables, y de correlación para establecer la comparación entre las dos técnicas de segmentación de imágenes, mediante el programa JMP (95% de confiabilidad).

Resultados



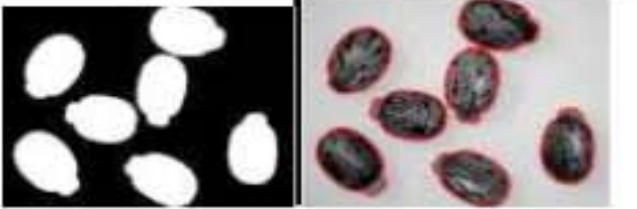


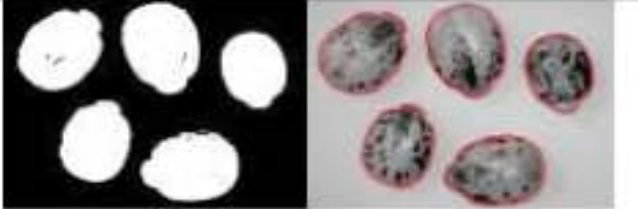
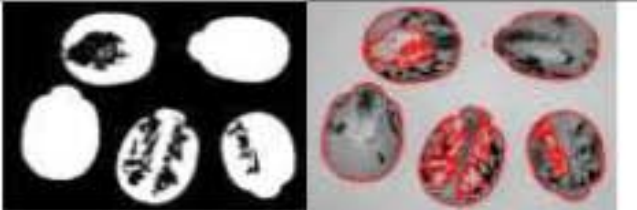
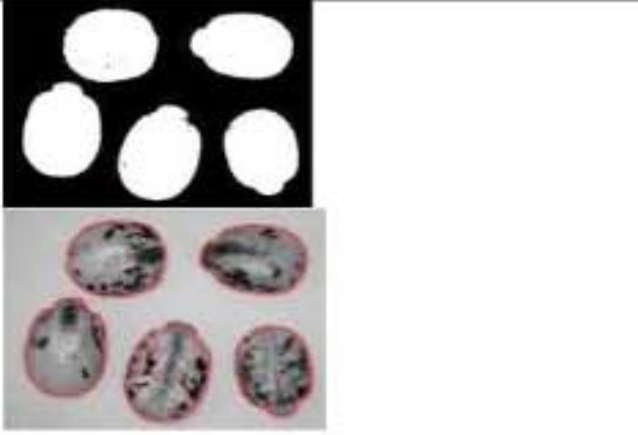

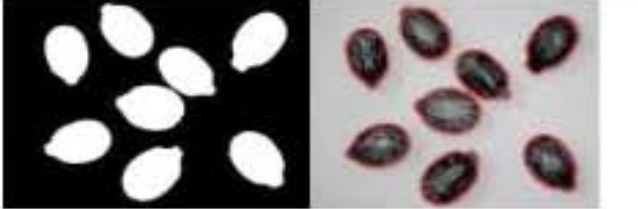
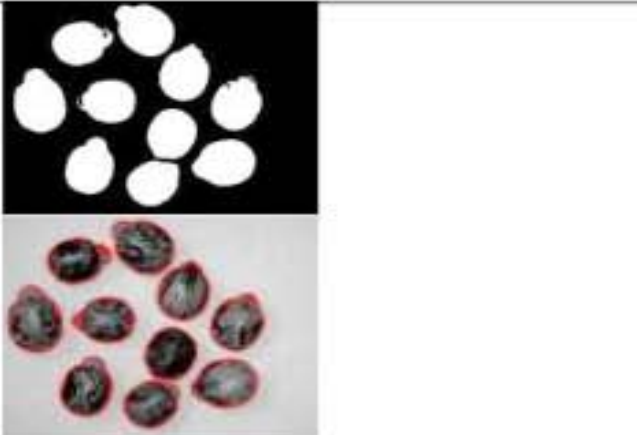
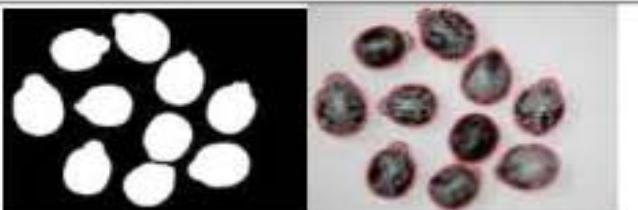
Para el inicio de la secuencia, se obtuvieron las imágenes mostradas en la figura 3, en donde existe una diferencia en color entre cada una de las accesiones, con contrastes de gris plateado (figura 3c y d), ocre (figura 3a), oscuro y negro mate (figura 3b y e); sin embargo se observó que de una misma variedad se presentan combinaciones de colores, dificultando su clasificación (figura 3f). Para esta variable, Biofuelseeds (2008) referenciado por Goytia *et al.*, (2011), propuso una clasificación basada en los siguientes colores: rojo, nordestino, blanco, negro, jaspeado y negro, clasificación que no coincide del todo con los mostrados en este documento. Por lo que se puede establecer que las accesiones crean fenotipos especiales para cada lugar.

Figura 9.2 Fotograma de semilla donde a) Nh1; b)Nh2; c)Nh3;d)Nh4; e)Nh10; f)NH12



Las imágenes fueron procesadas y segmentadas por *k-means* y un modelo propuesto; en este último se utilizó un factor de corrección, debido a las manchas que presentan principalmente las accesiones Nh 3 y 4 (tabla 9).Las imágenes mostraron ruido tanto en su binarizado como el segmentado debido principalmente al color de las accesiones antes mencionadas; en base a lo anterior se muestra que el modelo propuesto tiene la capacidad de procesar las imágenes aun cuando los colores de estos sean con contrastes de gris plateado.

Tabla 9 Muestra de imágenes

Imagen binaria y segmentada por k-mean	Modelo propuesto	Variedad
		Nh1
		Nh2
		Nh3
		Nh4
		Nh10
		Nh12

En base a las imágenes binarizadas y segmentadas de las variedades la morfometría de las accesiones fueron comparadas estadísticamente, evaluando el algoritmo y el modelo así como la comparación de cada una de las variables evaluadas por variedad como se muestra en la tabla 9.1

Tabla 9.1 Morfometría mediante dos técnicas de segmentación de imagen

Variedad	Método de procesamiento de imágenes				
	Segmentación mediante algoritmo de agrupamiento <i>k-means</i>				
	Area (mm ²)	MiAL (mm)	MaAL (mm)	IE (Φ)	DF
Nh1	2.18±0.054 D	14.58±0.410 CD	18.80±0.395 A	1.27±0.003 F	5.77±0.002 C
Nh2	1.26±0.054 F	10.88±0.410 FG	15.77±0.395 B	1.49±0.003 A	5.96±0.002 B
Nh3	2.08±0.054 D	14.05±0.410 CD	20.72±0.395 A	1.38±0.003 C	5.71±0.002 E
Nh4	2.80±0.054 C	16.77±0.410 B	19.86±1.395 A	1.28±0.003 F	5.73±0.002 D
Nh10	0.88±0.054 G	8.56±0.410 H	12.64±0.395 CD	1.47±0.003 B	6.09±0.002 A
Nh12	0.86±0.054 G	9.29±0.410 GH	12.17±0.395 DE	1.24±0.003 G	6.10±0.002 A
Modelo propuesto					
Nh1	4.40±0.044 AB	19.08±0.335 A	14.66±0.323 BC	1.30±0.002 H	5.60±0.001 G
Nh2	2.61±0.044 C	15.63±0.335 BC	10.66±0.323 EF	1.46±0.002 B	3.33±0.001 I
Nh3	4.14±0.044 B	18.82±0.335 A	13.95±0.323 BCD	1.35±0.002 D	5.27±0.001 H
Nh4	4.41±0.044 A	19.08±0.335 A	14.70±3.323 B	1.30±0.002 E	5.62±0.001 F
Nh10	1.99±0.044 D	13.02±0.335 DE	8.69±0.323 G	1.49±0.002 A	2.26±0.001 J
Nh12	1.69±0.044 E	11.82±0.335 EF	9.11±0.323 FG	1.29±0.002 EF	2.15±0.001 K

Promedio±error estándar con comparación de medias por Tuckey (95% confiabilidad)
Donde *MiAL*: eje menor; *MaAL*: eje mayor; *IE*: Índice de alargamiento; *DF*: Diámetro de Feret

Estadísticamente existe diferencia significativa ($P<0.05$) entre el método *k-means* y el *propuesto* en cada una de las variables evaluadas, en el procesamiento de imágenes se observó una serie de perturbaciones en las variedades Nh 3 y 4; lo cual no permite realizar los cálculos adecuados cuando las superficies delineadas se presentan como irregulares, las variedades con coloraciones claras presentaron esta problemática; por lo cual el algoritmo de agrupamiento *k-means* no se recomienda para variedades con estos pigmentos; no así con el algoritmo *propuesto*, el cual mostró segmentados definidos sin importar el color de la imagen.

El área superficial observada por medio de los dos modelos muestra diferencias significativas entre ellos, siendo las áreas del algoritmo *propuesto* el doble (1.69 a 4.41 mm²) que las analizadas por *k-means* (0.86 a 2.8 mm²). Al mismo tiempo se puede observar que las áreas entre las variedades fueron estadísticamente diferentes, teniendo mayor área las variedades Nh1, 3 y 4 en comparación con Nh10 y 12.

Para el eje mayor efectivamente se ve una diferencia clara en la medida dependiendo del modelo de análisis. Este comportamiento fue lo contrario a lo observado en las áreas. Siendo mayor la medida calculada por *k-means* (12.17 a 20.72 cm) en comparación con el modelo *propuesto* (8.69 a 14.7 mm); sin embargo fue el mismo comportamiento entre las variedades observándose claramente que las variedades Nh 1, 3 y 4 presentaron un eje mayor al de las variedades Nh10 y 12.

Algo similar al análisis del área sucede con el eje menor donde se observa una diferencia clara en la medida dependiendo del modelo de análisis.

El eje menor calculado por medio del algoritmo *propuesto* fue mayor (11.82 a 19.09 mm) que el analizados con *k-means* (8.56 a 16.77 mm). Entre las variedades conserva el mismo patrón que los anteriores siendo mayor el eje menor de las variedades Nh1, 3 y 4 que el de las variedades Nh10 y 12.

El comportamiento de la esfericidad muestra diferencias significativas, teniendo también un comportamiento similar al mostrado por la medición del área y el eje menor, siendo ligeramente mayor la esfericidad en el análisis del modelo *propuesto* (1.29 a 1.49) en comparación con el de *k-means* (1.24 a 1.49); sin embargo entre las variedades no se presentó el mismo comportamiento de las mediciones anteriores debido que en esta ocasión las variedades de mayor esfericidad fueron las variedades Nh 2 y 10 en comparación con la variedad Nh 12. Esto nos indicaría que la variedad Nh 2 presenta una forma más esférica y es aquella de menor tamaño en comparación con las demás variedades.

El índice de Feret presentó diferencia significativa entre los modelos de análisis, mostrando el mismo comportamiento que el de eje mayor. Siendo mayor la analizada por el *k-means* (5.71 a 6.1) en comparación con el algoritmo *propuesto* (2.15 a 5.6). Sin embargo entre las variedades se muestra una diferencia significativa en donde la variedad Nh 1 es mayor en comparación con la variedad Nh 10 y 12 (tabal 3).

Tabla 9.2 Análisis estadístico entre variedades sin factor modelo

Variedad	Comparativo
Nh 1 A	5.68 \pm 0.002
Nh 2 D	4.64 \pm 0.001
Nh 3 C	5.49 \pm 0.001
Nh 4 B	5.67 \pm 0.001
Nh10 E	4.17 \pm 0.001
Nh12 F	4.12 \pm 0.001

Media \pm desviación estándar

Letras iguales no muestran diferencia significativa (Tuckey<0.05)

Se ha dicho que las semillas mas pequeñas tienen tasas germinativas mayores que las semillas mas grandes (Edwards y Hartwig., 1971); mientras que otros autores han observado que el tamaño de la semilla en general no es tan importante cuando todos los miembros de la población tienen el mismo tamaño, pero cuando se encuentra una diferencia de este en la misma población, las plantas provenientes de semillas grandes pueden dominar y sobrevivir mejor, a expensas de los nutrientes de las semillas que se quedaron más pequeñas (Harper *et al.*, 1970).

9.3 Conclusiones

De manera general existe diferencia significativa ($P < 0.05$) desde el procesamiento de las imágenes entre el método *k-means* y el *propuesto*; siendo este último el que generó resultados de las variables más acordes comparado con el trabajo realizado por Goytia *et al.*,(2011), aun cuando existe diferenciación por los fenotipos de la región.

Las variables mostradas en la tabla muestran correlaciones directas por variedad, al área con sus ejes mayor y menor y diámetro de feret; lo anterior se debe a que son componentes de forma de las semillas, por lo que se deduce que a medida que se incrementa el eje mayor, se incrementarán en forma proporcional sus parámetros y viceversa, ya que son elementos dependientes.

En base a lo anterior es importante establecer las características morfológica en accesiones para establecer fenotipos por región, las técnicas de segmentación de imagen permiten apoyar en la medición rápida en grandes cantidades de semillas en tamaños diferentes, colores y formas.

9.4 Referencias

Amro A. A.y Soheir R. S.(2009). Degradation of Castor Oil and Lipase Production by *Pseudomonas aeruginosa*, *American-Eurasian J. Agric. & Environ. Sci.*, 5 (4): 556-563, 2009

Akpan, U. G., Jimoh, A. and Mohammed A. D. (2006). Extraction, Characterization and Modification of castor seed oil. *Leonardo Journal of Sciences*. Issue 8, January-June, p. 43-52

Cortina-Januchs M.G., Quintanilla-Dominguez J., Vega-Corona A., Tarquis Ana M. and Andina D. Detection of pore space in CT soil images using artificial neural networks. *Modeling soil system: complexity under your feet. Biogeosciences*. 8(2): 279-288, 2011.

De Souza Costa E., Guilherme L., de Melo E., Ribeiro B., Dos Santos B I., Da Costa Severiano E, Faquin V, Hale B.(2011). Assessing the Tolerance of Castor Bean to Cd and Pb for Phytoremediation Purposes. *Review Biol Trace Elem Res*. Agosto , p. 9

Edwards T., J. Jr.; Harthwig, E.E.(1971). Effects of seed size upon rate of germination in soybeans. *Agronomy Journal* vol. 63. Pp 429-430

Gil P. A., Jaramillo I. F; Cruz R. L., (2007). Poliuretanos degradables a partir de higuera. *Scientia et Technica* año XIII, No 36, septiembre 2007. Universidad Tecnológica de Pereira

Goytia J. M.A., Gallegos G. C.H., Núñez C. C.A., (2011). Relación entre variables climáticas con la morfología y contenido de aceite de semillas de higuera (*Ricinus communis L.*) de Chiapas. *Rev. Chapingo. Serie ciencias forestales y del ambiente* vol. 17. pp 41-48

Harper J.L; Lovelly P.H; Moore K.G(1970). The shapes and size of seeds. *Annual Review of Ecology and Systematics* vol.1, pp 327-357

Hoeneisen M y Silva M. (1986). Química y actividad biológica de algunas Compositae chilenas. *Rev. Latin Amer. Quim*. Vol. 17. pp 19-24

López S. J., Nieto A. R., Barrientos P. A.F., Rodríguez P. E., Colinas M.T., Borys M.W., González A. F., (2008). Selección de variables morfológicas para la caracterización del tejocote (*Crataegus spp*). *Rev. Chapingo. Ser. Hortic*. Vol.14, n.2 pp 97-111.

MacQueen J. B. Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1, 281–297, 1967.

Ojeda-Magaña B., Ruelas R., Quintanilla-Dominguez J., Corona-Nakamura M.A. and Andina D. Identification of Masses in Mammograms by Image Sub-segmentation. In *Soft Computing Models in Industrial and Environmental Applications, Advances in Intelligent and Soft Computing*. Springer Berlin / Heidelberg. Vol. 87: 589-598, 2011.

Ojeda-Magaña B., Ruelas R., Quintanilla-Dominguez J. and Andina D. Color Image Segmentation by Partitional Clustering Algorithms. In *36th Annual Conference of the IEEE Industrial Electronics Society, IECON 2010*. pp. 2828-2833, 2010.

Quintanilla-Dominguez J., Ojeda-Magaña B., Cortina-Januchs M.G., Ruelas R., Vega-Corona A. and Andina D. Image segmentation by fuzzy and possibilistic clustering algorithms for identification of microcalcifications. *Scientia Iranica Transaction D-Computer Science & Engineering and Electrical Engineering*. 18(3): 580-589. 2011.

Quintanilla-Domínguez J., Cortina-Januchs M.G., Barrón-Adame J.M., Vega-Corona A., Buendía-Buendía F.S. and Andina D. Detection of Microcalcifications Using Coordinate Logic Filters and Artificial Neural Networks. In *Bioinspired Applications in Artificial and Natural Computation, LNCS*. Vol. 5602, pp. 179-187, 2009.