



SISTEMA INFORMÁTICO PARA CONTAR SEMILLAS DE *Nicotiana tabacum*

Computer system to count seeds of *Nicotiana tabacum*

Yunieski Martínez-Espinosa¹✉, Juan L. Pérez-Rodríguez²,
Cosme E. Santiesteban-Toca³ y Evelio L. Báez-Perez³✉

ABSTRACT. In Cuba, conservation of *Nicotiana tabacum* is vital because it is one of the main items on the export and sustainability of the economy. The experimental station of Cabaiguán in Sancti Spiritus, whose functions include the conservation of genetic material from *Nicotiana tabacum*. Due to the small size of these seeds (0,1-0,5 mm), quantification has become one of the main problems, showing inconsistency in results of the count (high error rate) of a specialist to another, labor intensive and time lost. Therefore, it is propose to develop a computer system for quantification of *Nicotiana tabacum* seeds, based on techniques of digital image processing. Seedimages are captured in gray levels in a conventional scanner. Image segmentation is performed by thresholding the histogram from the frequency of gray levels. Impurities and image noise are eliminated from size. The average seed size and number of seeds present in each cluster is calculated. Finally, the number of seeds in the image is determined. To assess the quality and effectiveness of the proposed system an experimental validation was performed on 15 images containing 1492 seeds for different resolutions (150 dpi, 300 dpi, 600 dpi and 1200 dpi). Statistical results prove the ability of the system to count seeds with an accuracy above 99 % for the same or higher than 300 dpi resolutions.

RESUMEN. En Cuba, la conservación de *Nicotiana tabacum* es de vital importancia debido a que constituye uno de los principales rubros en la exportación y sustentabilidad de la economía del país. La estación experimental de Cabaiguán en Sancti Spiritus, tiene entre sus funciones la conservación del material genético de *Nicotiana tabacum*. Debido al reducido tamaño de estas semillas (0,1-0,5 mm), su cuantificación es uno de los principales problemas, observándose inconsistencia en el resultado del conteo, alta tasa de error de un especialista a otro, labor intensiva y pérdida de tiempo. Por tal motivo, se propone desarrollar un sistema informático para la cuantificación de semillas de *Nicotiana tabacum*, basado en técnicas de procesamiento digital de imágenes. Las imágenes de las semillas se capturan en niveles de grises en un scanner convencional. La segmentación de la imagen se realiza por umbralización a partir del histograma de frecuencia de los niveles de grises. Las impurezas y ruidos en la imagen son eliminados a partir del tamaño. Se calcula el tamaño medio de las semillas y el número de semillas presente en cada clúster. Finalmente el número de semillas en la imagen es determinado. Para evaluar la calidad y efectividad del sistema propuesto se realizó una validación experimental en 15 imágenes conteniendo 1492 semillas, para diferentes resoluciones (150 dpi, 300 dpi, 600 dpi y 1200 dpi). Los resultados estadísticos demuestran la capacidad del sistema de contar las semillas con una precisión superior al 99 % para las resoluciones iguales o superiores a 300 dpi.

Key words: quality, genetic material,
digital image processing

Palabras clave: calidad, material genético,
procesamiento digital de imágenes

¹Departamento de Ciencias Informáticas, Universidad “Máximo Gómez Báez” de Ciego de Ávila, carretera a Morón km 19 ½, Ciego de Ávila, Cuba, 65400.

²Estación Experimental del Tabaco carretera Santa Lucía km 2. Cabaiguán, Sancti Spiritus, Cuba.

³Laboratorio de Informática Aplicada, Centro de Bioplantas, carretera a Morón km 19 ½, Ciego de Ávila, Cuba, 65400.

✉ ebaez@bioplantas.cu

INTRODUCCIÓN

En nuestro país, el tabaco (*Nicotiana tabacum* Linneaus), históricamente es uno de los productos agrícolas comercializables de mayor prestigio a nivel mundial. Con el fin de conservar este material genético en cualquiera de sus fórmulas reproductivas (semillas, esquejes, tubérculos, etc.) se han creado los “bancos de germoplasma”. Cuya misión consiste en ubicar, recolectar, conservar y caracterizar el plasma

germinal de las plantas que, por sus atributos, son consideradas de interés prioritario, además de aportar conocimiento científico orientado a la optimización de la conservación y uso de los recursos fitogenéticos. Esto permite conservar el patrimonio genético vegetal y por tanto, la variabilidad genética de este género, lo que constituye el fundamento esencial de los programas de mejoramiento genético y de selección de genotipos.

A la par, se desarrollan y emplean técnicas de crioconservación para lograr condiciones de baja energía cinética molecular y una difusión extremadamente lenta, de tal forma que las reacciones químicas se encuentren prácticamente paralizadas; bajo estas condiciones, se postulan longevidades extremadamente largas. Sin embargo, estudios resientes muestran que después de un periodo de almacenamiento, donde no se muestran síntomas en la disminución de la calidad de las semillas, ocurre una rápida declinación de la viabilidad del material crioconservado (1).

Las condiciones de cultivo, cosecha y postcosecha, así como la humedad y la temperatura de almacenamiento son factores fundamentales que el operador debe tomar en consideración en el manejo de semillas en un banco de germoplasma (2). Por lo tanto, la influencia de estos factores debe ser evaluada antes de utilizar la crioconservación como estrategia segura para almacenar semillas del género *Nicotiana* (1, 2, 3). Para determinar el día óptimo de cosecha, el grado de madurez de la semilla juega un papel fundamental. Este parámetro guarda estrecha relación con la masa seca o fresca de las semillas. Sin embargo, teniendo en cuenta el reducido tamaño de estas semillas, entre 0,1 y 0,5 mm (Tabla I), su cuantificación manual es uno de los escollos principales en el procesamiento de las mismas. Factor que introduce un alto nivel de error en el proceso.

En los últimos años la incorporación de las técnicas de Procesamiento Digital de Imágenes (PDI) constituye una alternativa para la solución a este problema. Estas técnicas permiten la implementación de una amplia gama de aplicaciones con diversos objetivos, tales como inspección, clasificación, reconocimiento de objetos, mediciones, control de calidad y control de procesos, lo que las hace muy atractivas para el sector alimenticio y agrícola, al desarrollar sistemas para un amplio rango de productos, debido a que

permite construir herramientas capaces de realizar una inspección y clasificación de forma objetiva, rápida, confiable y no invasiva, evitando afectar la calidad.

El PDI permite realizar mediciones lo más exactas posibles de los objetos y extraer el máximo de información contenida en la imagen, en dependencia del problema en sí. Estas medidas pueden ser tan simples como el número de objetos, tamaño o color del objeto, o más complicadas como la forma, conectividad o apariencia (textura) de los objetos. Existen numerosos ejemplos en la literatura de la aplicación de las técnicas de PDI en el análisis y cuantificación de semillas (3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15). No obstante, ninguna de estas aplicaciones permite el conteo de semillas tan pequeñas como las de *Nicotiana tabacum*.

Es por ello que en el presente trabajo se propone un sistema informático basado en técnicas de PDI para la cuantificación de semillas de tabaco contenidas en una imagen, que permita reducir el riesgo de error que se introduce en el proceso de conteo de las mismas.

MATERIALES Y MÉTODOS

Las investigaciones se realizaron en el laboratorio de Informática Aplicada del Centro de Bioplantas, Universidad de Ciego de Ávila, en colaboración con la Estación Experimental del Tabaco de Cabaiguán en Sancti Spíritus. El desarrollo, implementación y puesta a punto de los algoritmos así como el tratamiento de las imágenes, se realizó con la plataforma de procesamiento de imágenes AForge.NET y el ambiente de desarrollo SharpDevelopment ver. 4.4. Se trabajó con la variedad *Nicotiana tabacum* Linneaus cv SanctiSpíritus96 y las imágenes se capturaron con un scanner convencional CANNON, ImageCLASS, MF4570dw.

Para la puesta a punto de los algoritmos y del sistema en general se utilizaron 15 imágenes de semillas de tabaco que contienen 1492 semillas, cada una de las imágenes fue capturada con diferentes resoluciones (150 dpi, 300 dpi, 600 dpi y 1200 dpi). En la Figura 1 se muestra una de estas imágenes y dos secciones ampliadas. El análisis estadístico se realizó usando el paquete estadístico SPSS (16).

Tabla I. Masa seca de 1000 semillas de cinco accesiones almacenadas en el banco de germoplasma del Instituto de Investigaciones del Tabaco de Cuba

Accesión	Masa seca de 1000 semillas (g)
<i>Nicotiana tabacum</i> Linneaus cv Sancti Spiritus 96	0,0790
<i>Nicotiana sylvestris</i> Esgazzini y Comes	0,0321
<i>Nicotiana plumbaginifolia</i> Viviani	0,0157
<i>Nicotiana megalosiphon</i> Heurck y Mueller	0,1101
<i>Nicotiana rustica</i> Linneaus	0,2895



Figura 1. Imagen de semillas de tabaco y dos secciones ampliadas

RESULTADOS Y DISCUSIÓN

Para realizar el conteo de las semillas, primeramente es necesario poder separar en la imagen qué es semilla y qué es fondo. Este proceso es conocido como segmentación y es un paso fundamental en el PDI, donde cada caso tiene sus particularidades en dependencia del tipo de imagen a analizar. Posteriormente es necesario detectar los objetos de interés (semillas) y eliminar los ruidos presentes en la imagen, ya sean por el proceso de captura de la imagen o por la misma naturaleza de la misma. Para realizar la segmentación y mejora de la imagen se siguieron una serie de pasos que se muestran en el algoritmo (Algoritmo 1) y que permitieron extraer cada semilla o grupo de semillas presente en la imagen.

Algoritmo 1. Proceso de mejora, segmentación y detección

- 1) Normalizar la imagen, empleando como rango los valores menor y mayor de los tonos de grises del histograma.
- 2) Umbralizar con un punto de corte que se calcula automáticamente en dependencia del histograma.
- 3) Realizar una transformación morfológica "Apertura" con el objetivo de eliminar posibles fragmentos de semillas partidas u otros ruidos presentes en la imagen.
- 4) Detectar las semillas o grupos de semillas a través de la detección de sus bordes.

En el proceso de conteo resulta muy común el solapamiento entre semillas. Esto afecta considerablemente la precisión del conteo realizado ya que cada grupo es considerado una sola semilla. Por esta razón, se aplica un algoritmo de conteo (Algoritmo 2) que incluye la aplicación oportuna de estadígrafos para la identificación de la cantidad de semillas.

Algoritmo 2. Proceso de conteo de las semillas

- 1) Calcular el área promedio de cada semilla o grupo de semilla.
 - a) Calcular el área de cada semilla o grupo de semilla.
 - b) Calcular el área promedio temporal.
 - c) Calcular el área promediode las áreas menores que el área promedio temporal.
- 2) Comparar el área promedio con las áreas de los elementos detectados, de esta forma es posible determinar la cantidad de semillas presentes en los grupos.
- 3) Determinar el número total de semillas.

EVALUACIÓN DE LA HERRAMIENTA PROPUESTA

La efectividad del método propuesto está condicionada directamente a la calidad del proceso de toma de la imagen. La aparición de elementos indeseables como impurezas en las muestras influye directamente en el resultado final.

Para la evaluación de la herramienta propuesta, se realizó una simplificación del problema, reduciéndolo a dos clases: semillas (positivos) e impurezas (negativos). Donde las clasificaciones correctas serían las semillas correctamente detectadas que son cuantificadas como verdaderos positivos (VP) y las impurezas no detectadas como semillas que son cuantificadas como verdaderos negativos (VN). Mientras que las clasificaciones incorrectas serían las impurezas detectadas como semillas y las semillas clasificadas de más en los grupos, que son cuantificadas como falsos positivos (FP) y las semillas identificadas como impurezas y no identificadas en los grupos, que son cuantificadas como falsos negativos (FN). La efectividad (Acc) de la herramienta se calcula a través de la Ecuación 1 (8, 14).

$$\text{Acc} = (\text{VP} + \text{VN}) / (\text{VP} + \text{FP} + \text{VN} + \text{FN}) \quad (1)$$

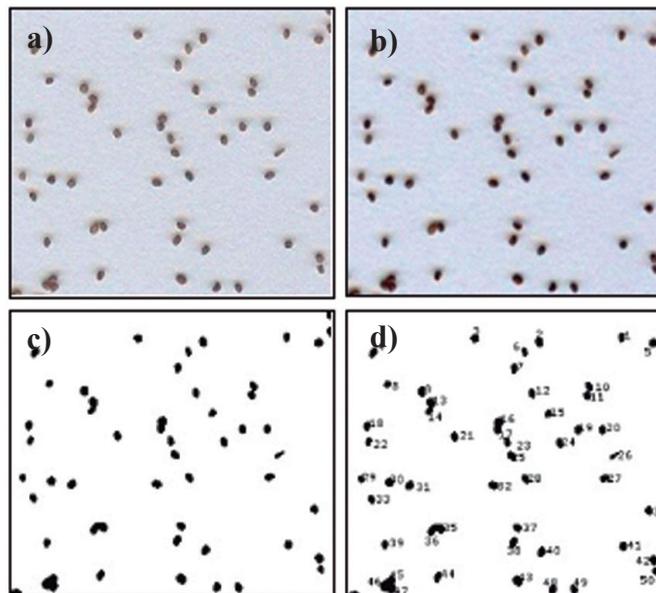
El método desarrollado se explica a través de la Figura 2, donde se muestra la imagen de una pequeña cantidad de semillas de tabaco, con el objetivo de mostrar mayores detalles (a); el resultado de transformar la imagen original a escala de gris y su normalización, proceso necesario para la posterior segmentación (b); se obtiene una imagen binaria (c); se eliminan los ruidos presentes en la imagen como resultado de la operación morfológica "apertura", y se calcula la cantidad de semillas en la imagen (d).

Para la evaluación de la herramienta propuesta, se procesaron un total de 1492 semillas separadas en 15 grupos. A su vez, se tomaron cuatro imágenes de cada grupo con las resoluciones (150 dpi, 300 dpi, 600 dpi y 1200 dpi), para un total de 60 imágenes. En

todos los casos se realizó el conteo con la herramienta, aplicando las mismas configuraciones. Los resultados se muestran en la Tabla II.

Los resultados de la resolución de 150 dpi son pobres, de una efectividad del 89 %. Esto se debe a que en esta resolución los contornos de las semillas no están bien definidos, lo que se incrementa en la presencia de grupos de semillas.

Para las resoluciones de 300, 600 y 1200 dpi los resultados son similares con una efectividad superior al 99 %, por lo que se recomienda utilizar una resolución de 300 dpi por obvias razones de ahorro de memoria y disminución del tiempo de ejecución.



a) imagen original. b) escala de gris y normalización. c) umbralización. d) apertura y conteo de las semillas

Figura 2. Proceso de conteo de las semillas

Tabla II. Resultados de la evaluación realizada a la herramienta propuesta

Imagen	150 dpi	300 dpi	600 dpi	1200 dpi	Media
	Acc	Acc	Acc	Acc	
1	0,94	0,97	1,00	0,97	0,97
2	0,94	1,00	1,00	1,00	0,99
3	0,95	0,98	1,00	0,98	0,98
4	0,94	1,00	1,00	1,00	0,98
5	0,97	1,00	0,99	0,99	0,99
6	0,98	1,00	1,00	1,00	0,99
7	0,65	1,00	0,98	0,99	0,91
8	0,91	1,00	1,00	1,00	0,98
9	0,95	0,98	0,98	0,99	0,98
10	0,76	0,99	0,98	0,99	0,93
11	0,83	0,98	1,00	0,99	0,95
12	0,90	0,99	1,00	0,99	0,97
13	0,91	1,00	0,99	1,00	0,98
14	0,97	0,99	1,00	0,99	0,99
15	0,98	1,00	1,00	0,99	0,99
Media	0,89	0,99	0,99	0,99	0,97

CONCLUSIONES

- ♦ En la presente investigación se propuso un sistema informático para la cuantificación de semillas de *Nicotiana tabacum* contenidas en una imagen, aplicando técnicas de PDI.
- ♦ Se realizó la validación experimental del mismo en 15 imágenes con 1492 semillas y para las resoluciones de: 150 dpi, 300 dpi, 600 dpi y 1200 dpi, demostrándose que el sistema es capaz de contar semillas de reducido tamaño (entre 0,1 y 0,5 mm), con una precisión superior al 99 % para las resoluciones iguales o superior a 300 dpi, sin que existan diferencias estadísticamente significativas entre éstas.
- ♦ El sistema se está aplicando en la Estación Experimental del Tabaco de Cabaiguán, en Sancti Spiritus, con resultados satisfactorios.

BIBLIOGRAFÍA

1. Walters, C.; Ballesteros, D. y Vertucci, V. A. "Structural mechanics of seed deterioration: Standing the test of time". *Plant Science*, vol. 179, no. 6, diciembre de 2010, (ser. Translational Seed Biology: From Model Systems to Crop Improvement), pp. 565-573, ISSN 0168-9452, DOI 10.1016/j.plantsci.2010.06.016.
2. Buitink, J.; Leprince, O.; Hemminga, M. A. y Hoekstra, F. A. "Molecular mobility in the cytoplasm: An approach to describe and predict lifespan of dry germplasm". *Proceedings of the National Academy of Sciences*, vol. 97, no. 5, 29 de febrero de 2000, pp. 2385-2390, ISSN 0027-8424, 1091-6490, DOI 10.1073/pnas.040554797, PMID: 10681458.
3. Bailly, C.; El-Maarouf-Bouteau, H. y Corbineau, F. "From intracellular signaling networks to cell death: the dual role of reactive oxygen species in seed physiology". *Comptes Rendus Biologies*, vol. 331, no. 10, octubre de 2008, (ser. Les graines de la vie / The seeds of life), pp. 806-814, ISSN 1631-0691, DOI 10.1016/j.cvi.2008.07.022.
4. Zapotoczny, P. "Discrimination of wheat grain varieties using image analysis and neural networks. Part I. Single kernel texture". *Journal of Cereal Science*, vol. 54, no. 1, julio de 2011, pp. 60-68, ISSN 0733-5210, DOI 10.1016/j.jcs.2011.02.012.
5. Alvarenga, R. O.; Filho, J. M. y Junior, F. G. G. "Avaliação do vigor de sementes de milho superdoce por meio da análise computadorizada de imagens de plântulas". *Journal of Seed Science*, vol. 34, no. 3, 12 de septiembre de 2012, pp. 488-494, ISSN 2317-1545.
6. Chaugule, A. "Application of image processing in seed technology: A survey". *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 4, 2012, pp. 153-159, ISSN 2250-2459.
7. Chiquito, A. A.; Junior, G.; Guilhien, F. y Marcos-Filho, J. "Assessment of physiological potential of cucumber seeds using the software Seedling Vigor Imaging System® (SVIS®)". *Revista Brasileira de Sementes*, vol. 34, no. 2, junio de 2012, pp. 255-263, ISSN 0101-3122, DOI 10.1590/S0101-31222012000200010.
8. Daniel, I. O.; Adeboye, K. A.; Oduwaye, O. O. y Pobeni, J. "Digital seed morpho-metric characterization of tropical maize inbred lines for cultivar discrimination". *International Journal of Plant Breeding and Genetics*, vol. 6, no. 4, 2012, pp. 245-251, ISSN 1819-3595.
9. Mandal, S.; Roy, S. y Tanna, H. "A low-cost image analysis technique for seed size determination". *Current Science*, vol. 103, no. 12, 2012, pp. 1401-1403, ISSN 0011-3891.
10. Pourreza, A.; Pourreza, H.; Abbaspour-Fard, M.-H. y Sadrnia, H. "Identification of nine Iranian wheat seed varieties by textural analysis with image processing". *Computers and Electronics in Agriculture*, vol. 83, abril de 2012, pp. 102-108, ISSN 0168-1699, DOI 10.1016/j.compag.2012.02.005.
11. Shahin, M. A.; Symons, S. J. y Wang, N. "Predicting dehulling efficiency of lentils based on seed size and shape characteristics measured with image analysis: Predicting lentil dehulling with image analysis". *Quality Assurance and Safety of Crops & Foods*, vol. 4, no. 1, marzo de 2012, pp. 9-16, ISSN 17578361, DOI 10.1111/j.1757-837X.2011.00119.x.
12. Belsare, P. P. y Dewasthale, M. M. "Application Of Image Processing For Seed Quality Assessment: A Survey". *International Journal of Engineering Research & Technology*, vol. 2, no. 2, 28 de febrero de 2013, ISSN 2278-0181, [Consultado: 1 de febrero de 2016], Disponible en: <<http://www.ijert.org/view-pdf/2329/application-of-image-processing-for-seed-quality-assessment-a-survey>>.
13. Pazoki, A.; Pazoki, Z. y Sorkhilalehloo, B. "Rain Fed Barley Seed Cultivars Identification Using Neural Network and Different Neurons Number". *World Applied Sciences Journal*, vol. 22, no. 5, 2013, pp. 755-762, ISSN 1818-4952.
14. Sandeep, V. V.; Kanaka, D. K. y Keshavulu, K. "Seed Image Analysis: Its Applications in Seed Science Research". *International Research Journal of Agricultural Science*, vol. 1, no. 2, 2013, pp. 30-36, ISSN 2327-3321.
15. Kaur, J. "Digital image analysis based on automated counting clustered soybean seeds". *International Journal of Computers & Technology*, vol. 12, no. 3, 3 de enero de 2014, pp. 3325-3328, ISSN 2277-3061.
16. *IBM SPSS Statistics* [en línea]. versión 20, [Windows], edit. IBM Corporation, U.S, 2011, Disponible en: <<http://www.ibm.com>>.

Recibido: 15 de mayo de 2015

Aceptado: 27 de enero de 2016

TUTORIAL

NÚMERO ESPECIAL

*Este número de la revista está dedicado
al X Congreso Internacional de Biotecnología Vegetal (BioVeg2015)*

El Centro de Bioplantitas es una institución de investigaciones científicas, adscrita a la Universidad de Ciego de Ávila “Máximo Gómez Báez” del Ministerio de Educación Superior de Cuba. El mismo surge en 1987 como un laboratorio de investigaciones y micropropagación de plantas frutales. Desde 1992, tiene como misión desarrollar, aplicar y ofrecer tecnologías, productos, asistencia técnica y servicios académicos de excelencia en el marco de la Biotecnología Vegetal.

El grupo de investigadores, técnicos de laboratorio y otro personal auxiliar altamente calificados, han sido galardonados con premios relevantes de la Academia de las Ciencias de Cuba y con reconocimientos por la labor que realizan en la transferencia de resultados científicos y tecnológicos, la producción de vitroplantitas para el comercio internacional, y la educación postgraduada. Para el trabajo científico cuenta con seis laboratorios: Cultivo de Células y Tejidos Vegetales, Agrobiología, Interacción Planta-Patógeno, Ingeniería Metabólica, Mejoramiento Genético de Plantas y Computación Aplicada. Todos con las mejores facilidades y un equipamiento de alta calidad para asegurar resultados relevantes.

El Centro de Bioplantitas desde 1997 y, como bienal, desarrolla su Congreso Internacional de Biotecnología Vegetal (BioVeg), el cual constituye un marco excepcional para el intercambio de conocimientos y experiencias entre científicos, docentes y productores. En este se debaten en forma de Conferencia Magistrales, Talleres y Mesas Redondas durante sesiones de trabajo, los resultados más relevantes y los problemas más acuciantes que enfrenta la biotecnología vegetal cubana y mundial.

Por todo lo anterior, el Comité Organizador de BioVeg2015 en su décima edición se complace en presentarles una muestra representativa de 19 trabajos científicos completos recibidos y siente profunda satisfacción en invitarlos para el próximo BioVeg2017 que se desarrollará en la fecha 22-26 del mes de mayo.

Nota:

Durante el proceso de edición no se pudo acceder al trabajo de retoque y mejoramiento de imágenes, por lo que estas han sido insertadas con la misma calidad con la que enviaron sus autores.

La Editorial