Cultivos Tropicales, 2016, vol. 37, special edition, pp. 106-110

ISSN impreso: 0258-5936
ISSN digital: 1819-4087
DOI: 10.13140/RG.2.1.1416.6649
http://dx.doi.org/10.13140/RG.2.1.1416.6649

Ministerio de Educación Superior. Cuba
Instituto Nacional de Ciencias Agrícolas
https://ediciones.inca.edu.cu

COMPUTER SYSTEM TO COUNT SEEDS OF *Nicotiana tabacum*

Sistema informático para contar semillas de *Nicotiana tabacum*

Yunieski Martínez-Espinosa[1]✉, Juan L. Pérez-Rodríguez[2], Cosme E. Santiesteban-Toca[3] and Evelio L. Báez-Perez[3]✉

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. Seedsimagesare captured ingray levelsina conventionalscanner. 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.

Key words: quality, genetic material, digital image processing

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 Spíritus, 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.

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

INTRODUCTION

In our country, tobacco (*Nicotiana tabacum* Linneaus) has historically been one of the most prestigious marketable agricultural products worldwide. In order to conserve this genetic material in any of its reproductive formulas (seeds, cuttings, tubers, etc.), "germplasm banks" have been created. Its mission is to locate, collect, preserve and characterize the germplasm of plants

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

[2] Estación Experimental del Tabaco carretera Santa Lucía km 2. Cabaiguán, Sancti Spíritus, Cuba.

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

✉ ebaez@bioplantas.cu

that, because of their attributes, are considered of priority interest, as well as providing scientific knowledge aimed at optimizing the conservation and use of plant genetic resources. This makes it possible to conserve the plant genetic heritage and, therefore, the genetic variability of this genus, which is the essential basis of genetic improvement and genotype selection programs.

At the same time, cryopreservation techniques are developed and used to achieve conditions of low molecular kinetic energy and extremely slow diffusion, so that the chemical reactions are practically paralyzed; under these conditions, extremely long longevities are postulated. However, studies show that after a period of storage, where there are no symptoms of decreased seed quality, a rapid decline in the viability of the cryopreserved material occurs (1).

The conditions of cultivation, harvesting and postharvest, as well as humidity and storage temperature are fundamental factors that the operator must take into account in the management of seeds in a germplasm bank (2). Therefore, the influence of these factors should be evaluated before using cryopreservation as a safe strategy for storing seeds of the genus *Nicotiana* (1, 2, 3). To determine the optimal harvest day, the degree of seed maturity plays a key role. This parameter is closely related to the dry or fresh mass of the seeds. However, considering the small size of these seeds, between 0,1 and 0,5 mm (Table I), their manual quantification is one of the main obstacles in processing them, factor that introduces a high level of error in the process.

In the last years the incorporation of the techniques of Digital Image Processing (DIP) constitutes an alternative for the solution to this problem. These techniques allow the implementation of a wide range of applications with different objectives, such as inspection, classification, object recognition, measurements, quality control and process control, which makes them very attractive to the food and agricultural sector. Systems for a wide range of products, because it allows the construction of tools capable of inspection and classification in an objective, fast, reliable and non-invasive way, avoiding affecting the quality.

The DIP allows to make measurements as exact as possible of the objects and to extract the maximum information contained in the image, depending on the problem itself. These measures can be as simple as the number of objects, size or color of the object or more complicated as the shape, connectivity or appearance (texture) of objects. There are numerous examples in the literature of DIP technique application in the analysis and quantification of seeds (3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15). However, none of these applications allows the counting of seeds as small as those of *Nicotiana tabacum*. For this reason, the present work proposes a computer system based on DIP techniques for the quantification of tobacco seeds contained in an image, which allows reducing the risk of error that is introduced in the counting process of the same.

MATERIALS AND METHODS

The research was carried out in the laboratory of Applied Informatics of the Center of Bioplants, University of Ciego de Ávila, in collaboration with the Tobacco Experimental Station of Cabaiguán in Sancti Spíritus. The development, implementation and tuning of the algorithms as well as the treatment of the images, was done with the image processing platform AForge.NET and the development environment SharpDeveloment ver. 4.4. The variety *Nicotiana tabacum* Linneaus cv Sancti Spíritus 96 was used and the images were captured with a conventional scanner CANNON, Image CLASS, MF4570dw.

For the set-up of the algorithms and the system in general, 15 images of tobacco seeds containing 1492 seeds were used, each image was captured with different resolutions (150 dpi, 300 dpi, 600 dpi and 1200 dpi). One of these images and two enlarged sections are shown in Figure 1. Statistical analysis was performed using the SPSS statistical package (16).

Table I. Dry mass of 1000 seeds of five accessions stored in the germplasm bank of the Institute of Tobacco Research of Cuba

Accesion	Dry mass of 1000 seeds (g)
Nicotiana tabacum Linneaus cv Sancti Spíritus 96	0,0790
Nicotiana sylvestris Espegazzini and Comes	0,0321
Nicotiana plumbaginifolia Viviani	0,0157
Nicotiana megalosiphon Heurck and Mueller	0,1101
Nicotiana rustica Linneaus	0,2895

Figure 1. Image of tobacco seeds and two enlarged sections

RESULTS AND DISCUSSION

To carry out the counting of seeds, it is first necessary to be able to separate in the image what is seed and what is background. This process is known as segmentation and it is a fundamental step in the DIP, where each case has its particularities depending on the type of image to analyze. Later it is necessary to detect the objects of interest (seeds) and eliminate the noises present in the image, either by the process of capturing the image or by the very nature of it. In order to perform the segmentation and improvement of the image, a series of steps were followed that were shown in the algorithm (Algorithm 1) and allowed to extract each seed or group of seeds present in the image.

Algorithm 1. Process of improvement, segmentation and detection

1) Normalize the image, using as a rank the lowest and highest values of grayscales of the histogram.
2) Threshold with a cutoff point that is automatically calculated depending on the histogram.
3) Perform a morphological transformation "Opening" with the aim of eliminating possible broken fragments of seeds or other noises present in the image.
4) Detect the seeds or seed groups through the detection of their edges.

The overlap between seeds is very common in the counting process. This considerably affects the precision of the count performed since each group is considered a single seed. For this reason, a counting algorithm is applied (Algorithm 2) that includes the timely application of statisticians for the identification of the quantity of seeds.

Algorithm 2. Process of counting the seeds

1) Calculate the average area of each seed or seed group.
a) Calculate the area of each seed or seed group.
b) Calculate the average temporary area.
c) Calculate the average area of the areas smaller than the average temporary area.
2) Compare the average area with the areas of the elements detected, in this way it is possible to determine the amount of seeds present in the groups.
3) Determine the total number of seeds.

Evaluation of the proposed tool

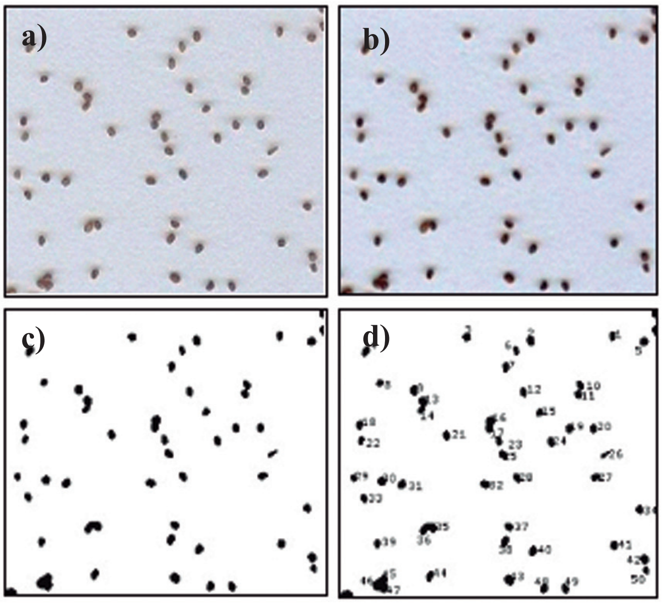
The effectiveness of the proposed method is directly dependent on the quality of the image-taking process. The appearance of undesirable elements as impurities in the samples directly influences the final result. For the evaluation of the proposed tool, a simplification of the problem was carried out, reducing it to two classes: seeds (positive) and impurities (negative). Where correct classifications would be correctly detected seeds that are quantified as true positives (VP) and undetected impurities as seeds that are quantified as true negative (NPV). While the incorrect classifications would be impurities detected as seeds and seeds classified further into groups, which are quantified as false positives (PF) and seeds identified as impurities and unidentified in the groups, which are quantified as false negatives (FN). The effectiveness (Acc) of the tool is calculated through Equation 1 (8, 14).

$$Acc = (VP + VN)/ (VP + FP + VN + FN) \qquad (1)$$

The method developed is explained through Figure 2, where the image of a small amount of tobacco seeds is shown, in order to show more detail (a); The result of transforming the original gray scale image and its normalization, a necessary process for subsequent segmentation (b); A binary image is obtained (c); the noises present in the image are eliminated as a result of the "opening" morphological operation, and the amount of seeds in the image is calculated (d).

For the evaluation of the proposed tool, a total of 1492 seeds were processed in 15 groups. In turn, four images of each group were taken with resolutions (150 dpi, 300 dpi, 600 dpi and 1200 dpi) for a total of 60 images. In all cases, the counting was done with the tool, applying the same configurations. The results are shown in Table II.

The resolution results of 150 dpi are poor, with an effectiveness of 89 %. This is because in this resolution the contours of the seeds are not well defined, which increases in the presence of seed groups. For resolutions of 300, 600 and 1200 dpi, and the results are similar with an effectiveness higher than 99 %, so it is recommended to use a resolution of 300 dpi for obvious reasons to save memory and decrease the execution time.



A) Original image. B) Gray scale and normalization. C) Thresholding. (D) Opening and counting of seeds

Figure 2. Seed counting process

Table II. Results of the evaluation made to the proposed tool

Image	150 dpi	300 dpi	600 dpi	1200 dpi	Mean
	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
Mean	0,89	0,99	0,99	0,99	0,97

CONCLUSIONS

- In the present investigation a computer system was proposed for the quantification of *Nicotiana tabacum* seeds contained in an image, applying DIP techniques.
- Experimental validation was performed on 15 images with 1492 seeds and for resolutions of: 150 dpi, 300 dpi, 600 dpi and 1200 dpi, demonstrating that the system is able to count seeds of small size (between 0,1 and 0,5 mm), with an accuracy greater than 99 % for resolutions equal to or greater than 300 dpi, without there being any statistically significant differences among them.
- The system is being applied at the Experimental Station of Tabaco de Cabaiguán, in Sancti Spíritus, with satisfactory results.

BIBLIOGRAPHY

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.crvi.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>.

Received: May 15th, 2015
Accepted: January 27th, 2016