International Journal on Advanced Science Engineering Information Technology

Vol.8 (2018) No. 4
ISSN: 2088-5334

Counting and Separating Damaged Seeds of Soybean Seeds using Image Processing

Monlica Wattana[#1], Buris Siriluk[*2], Suwanan Khotwit[#3]

Department of Computer Science, Faculty of Science, Khon Kaen University
Khon Kaen, 40002, Thailand
*E-mail: [1]monlwa@kku.ac.th;[*2]burissiriluk@gmail.com; [3]skyline13rd@gmail.com*

***Abstract*— This research proposed an image processing technique for counting soybean seeds and separating damaged seeds. The technique used was to adjust the image to black and white so that the soybean seeds differ from the background color. The second part, of the research, was the separation of soybean seeds using the distance transform method and the region growing method, to count soybean seeds. In the third part, it focused on the separation of damaged seeds by the size, the circle shape and HSV of soybean seeds. From 30 soybean seed images, the percentage of accuracy of counting and separating damaged seeds by naked eyes was 100 percent, and the average time spent was 13.70 seconds. The percentage of accuracy of counting the seeds by the developed program was 100 percent, and the accuracy of the separation of damaged seeds was 99.80 percent. The average time spent was 6.49 seconds. The experimental results showed that the developed program took 2 times less to count the soybean seeds than the naked eyes. Therefore, the proposed algorithm of the program can help save the time for counting the soybean seeds and separate the damaged seeds.**

***Keywords*— image processing; seed count; damaged seeds; soybean seed quality inspection**

I. INTRODUCTION

Soybeans are an important economic crop of the world and Thailand. It is food for consumption and processing while also being used in the oil industry and extraction industry. Soybean cultivation also helps to nourish soil [1]. However, good soybean cultivation depends on the process of selecting quality seeds for planting. The process of counting and selecting soybean seeds is done using only the naked eye. This means that to count and separate bad soybeans is a prolonged process, which can create problems. Also, if there are high numbers of soybeans, it takes a long time to count, mistakes may occur. Therefore, this article presents the method of counting and separating soybean seeds using morphological image processing technique.

Numerous studies have been conducted on the process of counting the numbers of objects. For example, Aniket A. et al. proposed a method for counting objects by calculating cutoff values, which were calculated from the average size of the objects, derived from connected component labeling. It was defined that if the mean value of the object is the range of the mean of the object, it will be counted [3]. Also, Y. H. Toh et al. suggested a method of counting fish using image processing. However, there was a problem with this method because of overlapping fish [4]. Thus, Nattapon Tidmerng et al. proposed a method for solving the problem of overlapping objects, by region growing technique with distance transform. This research studied the counting of birds with the images [5].

In addition, for the research in the medical field, Omid Sarrafzadeh et al. proposed the counting of White Blood Cells for blood smears by converting RGB image to L* a* b* and applying color values a* and b* to use K-means Clustering (KMC) and to find out the beginning of white blood cells. The experiment results showed that the mean segmentation error was 6.46 and the Jaccard Similarity Function which was used to compare the similarity of the data group was 93.71% [6]. A study proposing a technique to find out the perfect area of blood cells for determining cell composition by using image segmentation. Region Growing was used to find the area of the complete blood cell [7]. The study of Sanaha S. Pathan and Avinash D. Harale proposed the use of image processing for counting silkworm eggs. The image segmentation technique was used to find out eggs, and the thresholding technique was used to remove noise. Following this, the object count function in the MATLAB program was used [8].

For the research on counting the seeds in agriculture, Chanchai Nampol studied the verifying of the red rice seed mixed in the white rice seed by digital image processing. He proposed a method for examining the ratio of red rice seeds mixed in the white rice seeds, by digital image processing.

During this research, 98 percent of white rice and 2 percent of red rice were shot. The images were then converted to gray color, and noise signals were eliminated. After that, the red and white rice images were then separated by the level of the gray color scale, and then both images were converted to binary images in order to determine the number of pixels of red rice and white rice. The final step was to calculate the ratio of red rice mixed with white rice [9].

Additionally, Anol Paisal and Teerasit Kasetkasem studied the separation of mingling varieties of mung bean seeds by image processing. It was the separation of 100 mixed Mung bean seeds. The first step was to shoot each Mung bean, by placing the white area of Mung bean to the camera. The background of each image was separated from the seed by converting the RGB image into HSI image. The next step was to select the distinctive characteristics by the data in the order of the moment of HU. Then it was machine learning and classification. The seeds were divided into two groups equally. For the first group and the first group, 100 patterns were used for machine learning, and Support Vector was investigated.

The second group consisted of 100 patterns which were mixed, and they were separated from the moment of HU by using SVM [10]. Li Pengfei et al. studied the estimation of the number of crop seeds using edge detection and segmentation techniques, and the watershed algorithm technique was improved. However, some seeds cannot be counted due to the overlapping. Also, some seeds had a striped pattern inside so that they cannot be separated [11]. Jayme Garcia Arnal Barbedo also studied counting clustered soybean seeds by using morphological operations to find out the area of soybean seeds. The localization of the plateaus was used to show the most prominent of each seed for counting the soybean seeds [12]. However, this study failed to identify damaged soybean seeds. There was the study on the damaged soybean seeds conducted by Dejun Liu et al. which studied discriminating and elimination of damaged soybean seeds caused by insects, downy mildew, and other defects.

In the study, the researchers developed a camera that was used to obtain the details of each soybean seed. A total of eight features, including Perimeter, Area, Circularity, Elongation, Compactness Eccentricity, Elliptic axle ratio, Equivalent diameter were considered. The three values of colors were obtained from the mean, standard deviation and Mean Color Difference from the Mean (MCDM) of L*, a* and b* components. The values of texture features were Uniformity, Contrast, and Entropy. These three types of data were the inputs of BP neural network in separating damaged soybean seeds [13]. Additionally, the study of counting and sorting out soybean seeds using the size of soybean seeds which was presented by the author previously showed that the accuracy of the sorting out of the damaged soybean seeds was 99.21% [14].

Therefore, in this article, the shape, the size, as well as the color of soybean seeds were considered to increase the accuracy of the counting of seeds and the separation of damaged seeds from intact seeds. Segmentation along with distance transformation and the region growing methods were used to separate damaged seeds from intact seeds, by considering the size, shape, and color of soybean seeds.

II. MATERIAL AND METHOD

The process of counting soybean seeds and separating damaged seeds was presented in Fig. 1 and the details of each step are described as follows.

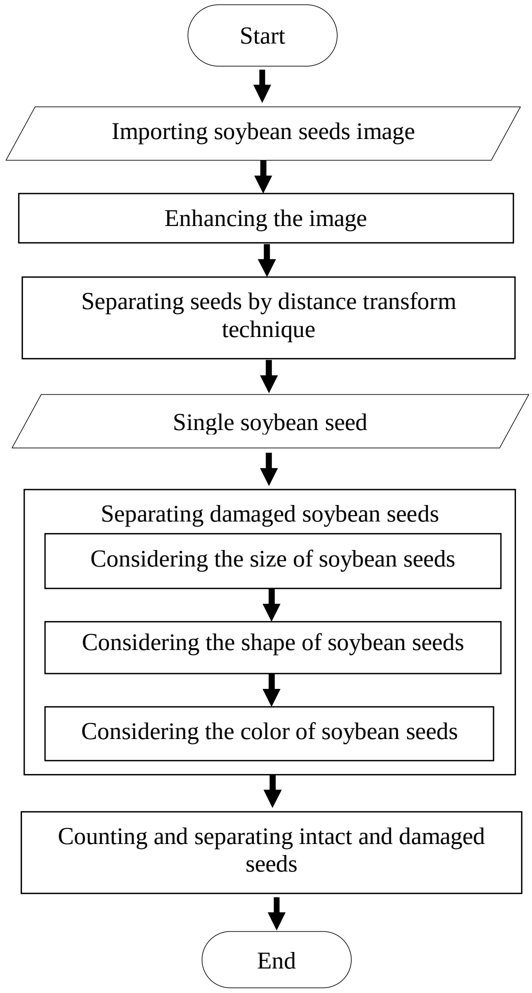

Fig. 1. The process of Counting and Separating Damaged Seeds of Soybean Seeds

A. RGB image of soybean seed

The images used in this research were RGB images of soybean seeds that were 2560 x 1920 pixels. The images were taken with the distance of 15 cm between the camera and seed. The background of the picture was black, as shown in Fig. 2.

Fig. 2. Soybean seeds image

B. Enhancing the soybean seed image

The image was adjusted from RGB color image to a gray color. Then it was adjusted to a black and white image so that the color of the soybean seeds is different from the color of the background. Then the noise was reduced using erosion and dilation to improve the image as shown in Fig. 3. The erosion is used to remove the small white pixel, and the dilation is used to add the white pixel of soybean seeds that were removed by the erosion operation. The erosion and dilation operation used a 5 X 5 structure element.

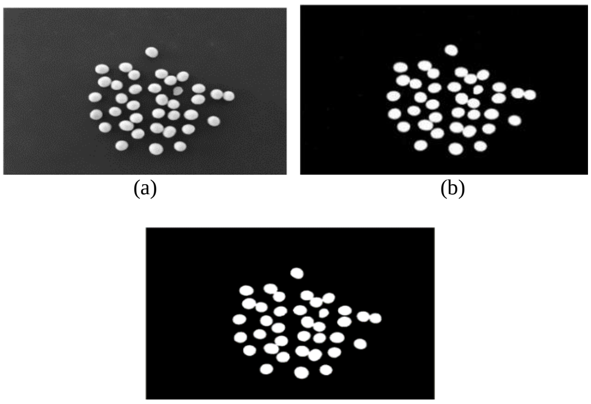

Fig. 3. Image enhancement (a) Grey color image
(b) Black-white color image (c) Reduced noise image

C. Separating intact soybean seeds

The distance transformation method was used to separate soybean seeds from the nearby seeds. Binary images were improved to find out the distance between the edges of the pixels. The white area of the soybean seeds was set to 1, and the black background was set to 0 as shown in Fig 4.

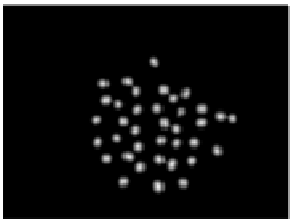

Fig. 4. The image processed through Distance Transform

Then threshold was employed to get only the white area which was calculated from Equation 1.

$$f_{thr}(x,y) = \begin{cases} 1, & f_f(x,y) < Threshold \\ 0, & f_f(x,y) \geq Threshold \end{cases} \quad (1)$$

The result from the threshold is shown in Fig. 5. Then the areas of the white pixel groups were counted, which the number of soybean seeds was.

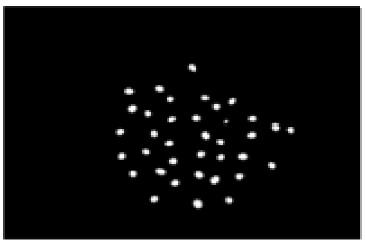

Fig. 5. The resulting image from the threshold

D. Separating damaged soybean seeds

After counting the soybean seeds, the separation of damaged seeds from intact seeds was done by the following procedure.

1) Considering the pixel size of each soybean seeds: The image used in this step was the image proceeded through the process of Region Growing. The black and white image was considered, and the number of white pixels was counted. If the area of the white pixel was less than the value set, it was considered a damaged seed. In the experiment, the average pixel of an entire soybean seed image was greater than 9,000 pixels. Therefore, if the number of pixels was less than 9,000 pixels, the seeds were considered damaged. The set pixel value was derived from the experiment with all soybean seed images. Fig 6 showed the separation of damaged soybean seeds.

Fig.6 . The separation of damaged soybean seeds (a) Damaged seed
(b) Intact seed

2) Considering the pixel size of each soybean seeds: After the procedure of separating soybean seeds by the size, a circle was created from the longest side of the soybean seed to consider the shape of soybean seed. After that, the circle created image was used to find out the overlapping area of the soybean seeds. When the overlapping area value was obtained, it was calculated into a percentage as in Equation 2.

$$\%\ \text{of overlapping area} = \frac{\text{overlapping area}}{\text{the longest side of the seed}} \quad (2)$$

In the experiment, the average percentage of the overlapping area of 50 single intact soybean seed images and the circle shape is 70%. Thus, if the overlapping area is greater than 70%, it is considered an intact seed. This value was derived from the experiment with all soybean seed images as shown in Fig 7.

Fig. 7. The images of created circle and soybean seed (a) Created circle
(b) Soybean seed

3) Considering the color of the soybean seeds: The HSV color was used to convert the RGB soybean color image into an HSV color image. Then, only the V color value was used to obtain the histogram. The average of the range from 232 -

256 (from 0 - 256) was calculated. If the average was less than 180, it was considered damaged seed. Otherwise, it is an intact seed. The HSV color of soybean seed is shown in Fig. 8.

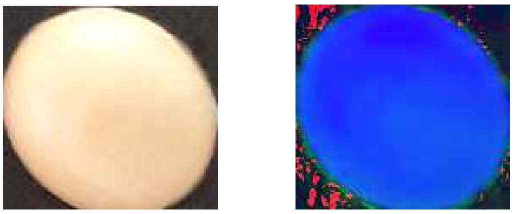

Fig.8. RGB and HSV soybean color images (a) RGB soybean color image (b) HSV soybean color image

E. Counting and separating intact and damaged seeds

During the process of separating soybean seeds by size, by the comparison with the circle and with the HSV color value, if any of the soybean seeds are presented in at least two-thirds of these conditions, they are considered intact seeds.

III. RESULTS AND DISCUSSION

During the experiment focused on counting and separating damaged soybean seeds using image processing, 30 images of soybean seeds were used. Each image contained 21-45 soybean seeds. The images used in the experiment were RGB images of soybean seeds taken with 2560 x 1920 pixels. The images were taken from above the soybean seeds with the distance of 15 cm between the camera and seed. The background of the photo was black, and the soybeans could not be overlapped. Moreover, the program was implemented in the Matlab program.

The experiment was divided into 2 parts. The first part was the counting and separating of the damaged seeds, by 3 persons with the naked eyes, each of which counted soybean seeds three times per one image. Then the average of the time spent counting was calculated. Another experiment component was the counting and separating of damaged seeds by the developed program. The time spent counting and separating the damaged soybean seeds 3 times per 1 image was considered. After that, the average time spent and the accuracy of counting and separating damaged soybean seeds was calculated. The average time spent and the accuracy of counting and separating damaged soybean seeds from the two parts of the experiment was then compared.

The experiments showed that the accuracy of counting and separating damaged soybean seeds was 100%. For the developed program, the accuracy of counting soybean seeds was 100%. However, for the separation of damaged soybean seeds, the results can be presented in three ways, as follows.

For the separation of damaged soybean seeds using the size of soybean seeds of 1,019 seeds, 8 seeds were incorrectly separated by the program. It was found that 3 damaged seeds were selected as intact seeds and 5 whole seeds were selected as damaged seeds. Therefore, the accuracy of the program in separating the damaged seeds was 99.21%. The results of the experiments on the separation of damaged soybean seeds with naked eyes and the developed program using the size of the soybean seeds were shown in Table 1. Confusion Matrix presented the results as in Table 1 below.

TABLE I
CONFUSION MATRIX TABLE OF THE ACCURACY OF SEPARATING DAMAGED SOYBEAN SEEDS WITH NAKED EYES AND THE DEVELOPED PROGRAM BY USING THE SIZE OF SOYBEAN SEEDS

	The program separating seeds as intact seeds	**The program separating seeds as damaged seeds**
Naked eyes separating seeds as intact seeds	(TP) 992	(FN) 5
Naked eyes separating seeds as damaged seeds	(FP) 3	(TN) 19

From Table 1, it can be concluded as follows.

Accuracy: (TP+TN)/(TP+TN+FP+FN) = 99.21%
Recall: TP/(TP+FN) = 99.50%
Specificity: TN/(TN+FP) = 86.36%
Precision: TP/(TP+FP) = 99.70%

For the separation of damaged soybean seeds using the comparison with the circle of 1,019 seeds, 13 seeds were incorrectly separated by the program. It was found that 9 damaged seeds were selected as intact seeds and 4 intact seeds were selected as damaged seeds. Therefore, the accuracy of the program in separating the damaged seeds was 98.72% as shown in Table 2. Confusion Matrix presented the results. Confusion Matrix table of the accuracy of separating damaged soybean seeds with naked eyes and the developed program by the comparison with the circle

TABLE II
CONFUSION MATRIX TABLE OF THE ACCURACY OF SEPARATING DAMAGED SOYBEAN SEEDS WITH NAKED EYES AND THE DEVELOPED PROGRAM BY THE COMPARISON WITH THE CIRCLE

	The program separating seeds as intact seeds	**The program separating seeds as damaged seeds**
Naked eyes separating seeds as intact seeds	(TP) 993	(FN) 4
Naked eyes separating seeds as damaged seeds	(FP) 9	(TN) 13

From Table 2, it can be concluded as follows.

Accuracy: (TP+TN)/(TP+TN+FP+FN) = 98.72%
Recall: TP/(TP+FN) = 99.60%
Specificity: TN/(TN+FP) = 59.09%
Precision: TP/(TP+FP) = 99.10%

For the separation of damaged soybean seeds using HSV color system of 1,019 seeds, 3 seeds were incorrectly separated by the program. Therefore, the accuracy of the program in separating the damaged seeds was 99.71% as shown in Table 3. Confusion Matrix presented the results as in the Table 3 below..

TABLE III
CONFUSION MATRIX TABLE OF THE ACCURACY OF SEPARATING DAMAGED SOYBEAN SEEDS WITH NAKED EYES AND THE DEVELOPED PROGRAM BY HSV COLOR SYSTEM

	The program separating seeds as intact seeds	**The program separating seeds as damaged seeds**
Naked eyes separating seeds as intact seeds	(TP) 997	(FN) 0
Naked eyes separating seeds as damaged seeds	(FP) 3	(TN) 19

From Table 3, it can be concluded as follows.

Accuracy: (TP+TN)/(TP+TN+FP+FN) = 99.71%

Recall: TP/(TP+FN) = 100%

Specificity: TN/(TN+FP) = 86.36%

Precision: TP/(TP+FP) = 99.70%

From the separation of soybean seeds by three methods and the scoring each method from 1,019 soybean seeds, 2 seeds were incorrectly separated by the program. It was found that 2 damaged seeds were selected as intact seeds. Therefore, the accuracy of the program in separating the damaged seeds was 99.80% as shown in Table 4. Confusion Matrix presented the results

Also, for the comparison of counting and separating damaged seeds by naked eyes and the developed programs, 3 test takers counted 30 soybean seeds. Each person counted 3 separate times and the average time was calculated. Then the average time from all three test takers was calculated again for obtaining the average time spent counting of 30 soybean seeds. The average time of counting was 13.70 seconds, and the accuracy was 100%. The results of the experiment were shown in Table 5. For the experiment of counting and separating soybean seeds using the developed program, the time spent counting and separating the damaged soybean seeds 3 times per 1 image was considered. After that, the average time spent was calculated, which revealed that the average time in counting and separating soybean seeds per image was 6.49 seconds.

TABLE IV
CONFUSION MATRIX TABLE OF THE ACCURACY OF SEPARATING DAMAGED SOYBEAN SEEDS WITH NAKED EYES AND THE DEVELOPED PROGRAM BY THREE METHODS

	The program separating seeds as intact seeds	**The program separating seeds as damaged seeds**
Naked eyes separating seeds as intact seeds	(TP) 997	(FN) 0
Naked eyes separating seeds as damaged seeds	(FP) 2	(TN) 20

The results of the accuracy and time spent in counting and separating damaged soybean seeds by naked eyes and the developed program were summarized in Table 5.

TABLE V
THE COMPARISON OF THE ACCURACY AND TIME SPENT IN COUNTING AND SEPARATING DAMAGED SOYBEAN SEEDS

Methods used to count soybean seeds	**Percentage of the accuracy of counting**	**Percentage of the accuracy of separation**	**Time spent in counting (second)**
Naked eyes	100	100	13.70
Developed program	100	99.80	6.49

From all experiment results, the accuracy of counting and separating soybean seeds of the program and naked eyes was similar in that the accuracy was 100%. Moreover, the average processing time of the program was less than the naked eyes. However, the environment was controlled for taking the soybean pictures. Therefore, if the environment is changed, such as the distance between the camera and seed is closer, the size of the soybean image becomes bigger. The image brightness changes, the accuracy of counting and separating soybean seeds would decrease. Therefore, if the environment is changing, the new parameters have to be set.

IV. CONCLUSION

According to the experiment of 30 soybean seed images where each image contained 21-45 soybean seeds, it can be concluded that the accuracy of counting and separating soybean seeds with naked eyes was 100%, with the average processing time being 13.70 seconds. Also, the accuracy of counting soybean seeds by the developed program was 100%. For the separation of soybean seeds by three methods, the accuracy of separation was 99.80 %. Moreover, the average processing time was 6.49 seconds. Therefore, it is clear that the program used less time than the naked eyes, a difference of 7.21 seconds. The results also showed that the developed program could count and separate damaged soybean seeds faster than the naked eyes. The accuracy of the counting and separating of damaged soybean seeds was also similar.

REFERENCES

[1] T. Eksomtramage, “Soybean: Glycine max,” 2011. [Online]. Available: http://www.natres.psu.ac.th/Department/PlantScience/510-211/pages/soybean.htm.

[2] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. U.S.A: Prentice-Hall, 2002.

[3] A. A. Khule, M. S. Nagmode, and R. D. Komati, “Automated Object Counting for Visual Inspection Applications,” in *2015 International Conference on Information Processing (ICIP)*, 2015.

[4] B. K. Toh, Y. H. Ng, T.M. Liew, “Automated fish counting using image processing,” in *International Conference on Computational Intelligence and Software Engineering, 2009(CiSE2009)*, 2009.

[5] N. Tidmerng, W. Songpan, and M. Wattana, “Solving bird image overlapping for automatic population counts of birds using image processing,” in *2016 Management and Innovation Technology International Conference (MITicon)*, 2016.

[6] O. Sarrafzadeh, A. M. Dehnavi, H. Rabbani, and A. Talebi, “A simple and accurate method for white blood cells segmentation using K-means algorithm,” in *IEEE Workshop on Signal Processing Systems (SiPS)*, 2015.

[7] M. Wattana and T. Boonsri, “Improvement of Complete Malaria Cell Image Segmentation,” in *2017 Twelfth International Conference on Digital Information Management (ICDIM)*, 2017.

[8] S. S. Pathan and A. D. Harale, “Automated silkworm eggs count,” in *2016 IEEE International Conference on Advances in Electronics, Communication and Computer Technology (ICAECCT)*, 2016.

[9] C. Nampol, “The Verifying of the Red Rice Seed Mixed in the White Rice Seed by Digital Image Processing,” *J. Ind. Technol. Ubon Ratchathani Rajabhat Univ., 1 (January - June 2015)*, pp. 60–70, 2015.

[10] A. Paisal and T. Kasetkasem, “Separation the mingling varieties of the mungbean seeds by image processing,” *Khon Kaen Agric. J. 39, 3*, pp. 240–247, 2011.

[11] L. Pengfei, M. Qi, and Z. Weihong, “Estimation of the number of crop seeds based on image identi.”

[12] J. G. A. Barbedo, “Counting Clustered Soybean Seeds,” in *12th International Conference on Computational Science and Its Applications (ICCSA)*, 2012.

[13] D. Liu, X. Ning, Z. Li, D. Yang, H. Li, and L. Gao, “Discriminating and elimination of damaged soybean seeds based on image characteristics,” *J. stored Prod. Res. v.60*, pp. 67–74, 2015.

[14] B. Siriluk, S. Khotwit, and M. Wattana, “A Soybean Seeds Counting and Sort out Using Size of Soybean Seeds,” in *The 5th ASEAN Undergraduate conference in Computing (AUC2)*, 2017.