Rice Seed Germination Analysis

Benjamaporn Lurstwut
Faculty of Information and Communication Technology, Mahidol University, Thailand

Chomtip Pornpanomchai
Faculty of Information and Communication Technology, Mahidol University, Thailand

Abstract: This research aimed to develop the computer software called “Rice Seed Germination Analysis (RiSGA)” which could predict rice seed image for rice germination by using an image processing technique. The RiSGA consisted of five main process modules: 1) image acquisition, 2) image pre-processing, 3) feature extraction, 4) quality control analysis and 5) quality results. Six variations of Thai rice seed species (CP111, RD41, Chiang Phattalung, Sang Yod Phattalung, Phitsanulok 2 and Chai Nat 1) were used for the experiment. The RiSGA extracted three main features: 1) color, 2) morphological and 3) texture feature. The RiSGA applied four well-known techniques: 1) Euclidean Distance (ED), 2) Rule Based System (RBS), 3) Fuzzy Logic (FL) and 4) Artificial Neural Network (ANN). The RiSGA precision of ED, RBS, FL, and ANN was 87.50%, 100%, 100%, and 100%, respectively. The average access time was 4.35 seconds per image, 5.29 seconds per image, 7.04 seconds per image, and 159.65 seconds per image, respectively.

Keywords: Rice seed, seed germination, rice seed features, image processing, computer vision

1. INTRODUCTION
Nowadays, the agricultural industry is more widespread in the world. Oryza Sativa (Rice) is a vital worldwide agricultural produce which is very popular [47]. Thailand is one of the agricultural countries which produces a large number of food products e.g. cereals, flowers, vegetables, fruits, rubber, and especially rice. In Thailand, there are more than 114 well-known Thai rice species [33, 39, 46]. Paddy rice based on the good quality of products is mostly offered [40]. Therefore, it is essential to grade the quality of these commodities in order to command the better price in the market competition. The quality of the rice is mostly based on the quality of the seeds. However, it is very difficult to identify the quality of rice seeds by using only human vision. Thus, this research aims to apply the standard seed germination test from the International Seed Testing Association (ISTA 1996) with the top of paper method for rice germination [15, 16, 22, 23, 31]. Moreover, images are collected to predict rice seed images by using image processing techniques [18] which can identify the quality of products. The objective of this research is to develop the computer software which can predict rice seed microscopic image for rice germination by using image processing techniques. This research focuses on six variations of Thai rice seed species including CP111, RD41, Chiang Phathalung, Sang Yod Phathalung, Phitsanulok 2 and Chai Nat 1.

2. RELATED WORKS
Many researchers [1, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 17, 19, 20, 21, 24, 26, 27, 29, 30, 32, 34, 35, 36, 37, 38, 41, 42, 43, 44, 45, 48, 49, 50, 51] have applied image processing for rice germination in their studies. This is because using image processing is rapid, economic, consistent, and objective. However, there is no research conducted by using machine vision and image processing in the microscopic level which is more efficient for analyzing the information. Therefore, this research aims to implement the automatic system in the microscopic level. The Rice Seed Germination Analysis (RiSGA) system is developed and used as the inspection tool for measuring the rice seed quality in the rice germination which is essential to grade the quality of products in order to command the better price in the market competition.

3. MATERIALS AND METHODS
The experiment was conducted by using the following computer hardware specifications: 1) CPU Intel(R) Core(TM) i5-2400 CPU @ 3.10GHz 2) Memory DDR3 4 GB and 3) Hard disk 500 GB. For the computer software, Microsoft Windows 7 (Microsoft Corp.; Redmond, WA, USA) was used as the operating system. For the development tool, MATLAB R2013a (The Math Works Inc.; Natick, MA, USA) [28] was used.

Analysis and design were described by using the system conceptual diagram and system structure chart. The details of each element are described below.

3.1 System conceptual diagram
The RiSGA system is a system to imitate the abilities of the classification of the quality control of the seeds. Fig. 1 shows the overview of the system conceptual diagram. The operation of the RiSGA system is divided into two phases: 1) the preparation of the system data set for training data set and testing data set, and 2) the preparation of the system data set for testing the system on another data set of unseen or unknown images. The two phases has to be operated in order to observe its performance and accuracy. However, the second phase has to be used to validate and evaluate the RiSGA system efficiently.
3.2 System Structure Chart

The RiSGA structure chart elaborates on how each model works is shown in Fig. 2. The RiSGA consists of five main process modules: 1) image acquisition, 2) image pre-processing, 3) feature extraction, 4) quality control analysis and 5) quality results. Each process module has the following details.

3.2.1 Image Acquisition

In this module, it consists of three main parts which are capturing image, cropping microscopic image, and storing image. In capturing the image, the rice seed is placed on the plate and then the rice seed is taken in a bird-eyes-view angle from a digital microscope camera on a light mode or a simple camera. The rice seed image will be used as the input to the RiSGA system. In cropping the image, the size of each image captured from the microscope camera and the simple camera is very large and there is a lot of noise. Therefore, the raw microscopic image must be cropped in the same scale of 1900 x 1900 pixels at the same position. In storing image, the rice seed images from the microscope camera and the digital camera are collected and stored in the system database.

3.2.2 Image Pre-processing

In the second module, the image pre-processing module prepares an image before processing in the feature extraction process. This module consists of two sub-modules which are seed image segmentation and seed image binarization.

3.2.2.1 Seed Image Segmentation

The sample of rice seed microscopic image segmentation is shown in Figure. 3. Figure. 3 (a) shows a rice seed microscopic image. Figure. 3 (b) shows a bounding box of rice seed microscopic image. Figure. 3 (c) shows an enhanced bounding box of rice seed microscopic image. Figure. 3 (d) shows a segmented rice seed microscopic image.

First, the RiSGA system finds the bounding box of the rice seed microscopic input image by changing the RGB rice input microscopic image to the binarization image (Figure. 3 (a) - Figure. 3 (b)). Next, the RiSGA system will perform morphological closing to close any opening area. The RiSGA system fills holes and removes noise in order to get the enhanced binarization image. Then, the RiSGA system labels the eight connected components of the enhanced binarization image to build the rectangle which can cover and fit the size of the rice seed object in order to be the label bounding box of the rice seed microscopic image (Figure. 3 (c)). Finally, the RiSGA system uses the label bounding box of the rice seed microscopic image to segment only the rice seed object in the input image (Figure. 3 (d)).

3.2.2.2 Seed Image Binarization

The sample of rice seed microscopic image binarization is shown in Figure. 4. Figure. 4 (a) shows a segmented rice seed microscopic image. Figure. 4 (b) shows a rice seed microscopic gray-scale image. Figure. 4 (c) shows a rice seed microscopic binarization image. Figure. 4 (d) shows an enhanced rice seed microscopic binarization image.

First, the RiSGA system changes the RGB color image to the gray-scale image (Figure. 4 (a) - Figure. 4 (b)). Next, the RiSGA system transforms the gray-scale image to the binary image (Figure. 4 (c)). Then, the RiSGA system fills holes and removes noise in order to get the enhanced binarization image (Figure. 4 (d)).
3.2.3 Features extraction

In the third module, the feature extraction module is the most important module. This is because this module contains rice images which are cropped from Module 1. Moreover, this module is used to describe the accuracy of the large set of the data. In the classification process, there are three main sub-modules or seven combination optimal features as follows:

- Color feature consists of (1) average red color, (2) average green color, (3) average blue color
- Morphological feature consists of (4) aspect ratio, (5) edge,
- Texture feature consists of (6) entropy, (7) energy

To define the optimal features, the feature selection process is used to observe the most relevant features of the seed characteristics. In this research, the Wrapper Feature Selection Approach [25] is selected to search the optimal feature subset in order to achieve the best performance. After the Wrapper Feature Selection Approach had been applied, the three main sub-modules or the seven combination optimal features occurred which were (1) average red color, (2) average green color, (3) average blue color, (4) aspect ratio, (5) edge, (6) entropy, and (7) energy.

3.2.3.1 Color Feature

There are many color spaces in color features. In this research, the RGB color space is considered as the main color feature. The RGB feature consists of three features: (1) average read color, (2) average green color, and (3) average blue color. The three features are considered as the optimal selected color features.

3.2.3.2 Morphological Feature

The morphological feature measures the features based on the seed morphology, especially for the shape and the size of the seed. The morphological feature consists of two optimal features which are (1) aspect ratio and (2) edge. The aspect ratio is calculated by the major axis length divided by minor axis length. The edge is calculated by applying the Sobel edge detection with threshold values 0.03 to measure the remaining pixels in the seed area.

3.2.3.3 Texture Feature

The RiSGA system applies gray level co-occurrence matrices (GLCM) [2] for measuring the seed surface texture. The GLCM is extracted from each of the gray-tone spatial-dependence matrices. There are many features in the GLCM. However, the two optimal features: (1) energy and (2) entropy are applied in this research. Each texture feature is calculated based on equation (1) - (2):

Where

\[ P_{ij} = \text{entry in a normalized gray-tone spatial-dependence matrix,} \]

\[ N = \text{number of distinct gray levels in the quantized image.} \]

(1) Energy Texture Feature

The energy texture feature known as uniformity is the sum of squared elements in the GLCM. The energy texture is calculated in equation (1):

\[ \mu_i = \sum_{i,j=0}^{N-1} i P_{i,j}, \mu_j = \sum_{i,j=0}^{N-1} j P_{i,j} \]  

(2) Entropy Texture Feature

The entropy texture feature is a statistical measure of randomness which is used to characterize the texture of the input image. The entropy texture is calculated in equation (2):

\[ \sum_{i,j=0}^{N-1} P_{i,j} \left( - \ln P_{i,j} \right) \]  

3.2.4 Quality Control Process

In the fourth module, the RiSGA system uses the quality control analysis to apply four techniques to predict rice seed image for rice germination: (1) Euclidean Distance (ED), (2) Rule Based System (RBS), (3) Fuzzy Logic (FL) and (4) Artificial Neural Network (ANN). Next, the RiSGA system compares the features of the testing rice seed image data set with the training rice seed image data set in the system database.

3.2.4.1 Euclidean Distance (ED)

The Euclidean Distance (ED) is used to measure the similarity of the distance between every feature of a sample data set and every feature of each training data set in the RiSGA system. The RiSGA system applies the Euclidean Distance based on the minimum distance. The minimum Euclidean Distance determines that the sample data set and the training data set are very similar. The Euclidean Distance is calculated in equation (3):

\[ ED = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \]  

Where \( ED = \) the Euclidean distance value between two objects which are \( x \) and \( y \), \( n = \) number of features, \( x_i = \) the value of feature \( i \) in the system database and \( y_i = \) the value of feature \( i \) in a sample image.

3.2.4.2 Rule Based System (RBS)

The rule based system is used to store and manipulate knowledge. The knowledge is stored as the rule based representation in IF-THEN structure. The fact is represented in the IF part (antecedent). The action is represented in the THEN part (consequent). The RiSGA system applies the rule based system from the area under the normal curve in the normal distribution of data set.

3.2.4.3 Fuzzy Logic (FL)

The fuzzy logic is determined as a set of mathematical principles for knowledge representation based on degree of membership function. To define the appropriate degree of
membership function, the RiSGA system applies the mean and the quantity of three times the standard deviation to define the degree of membership function. The feature values in the data set are normalized to the scale of 0 and 1 which is easier to determine the degree of the membership function. The RiSGA system applies the trapezoidal function for representing the degree of the membership function. The input range were between 0 and 1 while the output range were between -0.3 and 1.5.

3.2.4.4 Artificial Neural Network (ANN)

ANN is a mathematical model or computational model that is inspired by the structure and functional aspects of biological neural networks. ANN is composed of many artificial neurons that are linked together according to specific network architecture. The patterns feed the input into the network and then the network will return the output. The objective of the neural network is to transform the input into the meaningful output. In this research, the artificial neural network (ANN) classifies the rice images by using the neural network structure of 7-6-2. The seven input nodes are equal to seven features of each seed image and the two output nodes are equal to two kinds of germinated seed and non-germinated seed in the training data set. The hidden nodes are 2/3 of average between input nodes and output nodes which is the rule of thumb [5].

![Diagram of neural network](image)

Figure 5: The design of the neural network.

The neural network consists of input layer, hidden layer, and output layer. The input layer consists of 7 nodes. The hidden layer consists of 6 nodes. The output layer consists of 2 nodes.

3.2.5 Quality Result

The quality result module shows rice seed germinating prediction results displaying as the graphic user interface. This module is categorized into two parts which are Display Quality Result Image and Display Details.

3.2.5.1 Display Quality Result Image

Once the user load the unknown input image to the system, the system will extract the features of the unknown input image. Next, the system compared the features of unknown image with the system database. Then, the RiSGA system will analyze the result by using one of five techniques. Finally, the quality result image will display in the graphic user interface as shown in Figure 6.

![GUI of selecting the classification technique](image)

Figure 6: The GUI of selecting the classification technique.

3.2.5.2 Display Details

The analyzing image result was displayed in the germination analysis result image box. The details of the analyzing image were displayed in the analyze result text box. Figure 7 shows the GUI of displaying the quality control analysis result.

![GUI of displaying the quality control analysis result](image)

Figure 7: The GUI of displaying the quality control analysis result.

4. RESULTS AND DISCUSSIONS

The RiSGA process employs six Thai rice seed species which are 1) CP111, 2) RD41, 3) Chiang Phatthalung, 4) Sang Yod Phattalung, 5) Phisanulok 2 and 6) Chai Nat 1 from Bureau of Rice Seed, Ministry of Agriculture and Cooperatives, Rice Department of Thailand. The sample of rice seed germination testing based on the top of paper method. For the evaluation of the five learning algorithms, each kind of rice species contains 600 samples in the microscope training data set and 120 samples in the microscope unknown data set.
4.1 Experiment Result based on training data set
The experiment result based on training data set is shown in Table 1.

<table>
<thead>
<tr>
<th>TQ</th>
<th>Amt</th>
<th>G (GP)</th>
<th>N-G (NG)</th>
<th>A (%)</th>
<th>T (s/img)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ED</td>
<td>600</td>
<td>360</td>
<td>0</td>
<td>240</td>
<td>100.00</td>
</tr>
<tr>
<td>RBS</td>
<td>600</td>
<td>348</td>
<td>12</td>
<td>11</td>
<td>229</td>
</tr>
<tr>
<td>FL</td>
<td>600</td>
<td>360</td>
<td>0</td>
<td>240</td>
<td>100.00</td>
</tr>
<tr>
<td>ANN</td>
<td>600</td>
<td>360</td>
<td>0</td>
<td>240</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Where TQ = Technique, Amt = Amount, G = Germination, NG = Non-Germination, GP = Germinated Prediction, N-GP = Non-Germinated Prediction, A = Accuracy, T = Time, and s/img = seconds per image.

4.2 Experiment Result on unknown data set
The experiment result based on unknown data set is shown in Table 2.

<table>
<thead>
<tr>
<th>TQ</th>
<th>Amt</th>
<th>G (GP)</th>
<th>N-G (NG)</th>
<th>A (%)</th>
<th>T (s/img)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ED</td>
<td>120</td>
<td>68</td>
<td>0</td>
<td>15</td>
<td>37</td>
</tr>
<tr>
<td>RBS</td>
<td>120</td>
<td>68</td>
<td>0</td>
<td>0</td>
<td>52</td>
</tr>
<tr>
<td>FL</td>
<td>120</td>
<td>68</td>
<td>0</td>
<td>0</td>
<td>52</td>
</tr>
<tr>
<td>ANN</td>
<td>120</td>
<td>68</td>
<td>0</td>
<td>0</td>
<td>52</td>
</tr>
</tbody>
</table>

Where TQ = Technique, Amt = Amount, G = Germination, NG = Non-Germination, GP = Germinated Prediction, N-GP = Non-Germinated Prediction, A = Accuracy, T = Time, and s/img = seconds per image.

5. CONCLUSIONS
In this research, the Rice Seed Germination Analysis (RiSGA) system developed from MATLAB R2013a running on Windows 7 Ultimate and the Wrapper Feature Selection Approach were selected for performing feature selection. Later, the system was developed to predict rice germination in the seed quality control process. Six Thai rice seed species which were CP111, RD41, Chiang Phathalung, Sang Yod Phattalung, Phisanulok 2, and Chai Nat collected from Bureau of Rice Seed, the Ministry of Agriculture and Cooperatives, and the Rice Department of Thailand were selected for the experiment. The RiSGA data set which consisted of 720 microscopic rice seed images were used for training data set and unknown data set. The germination experiment was conducted under the standard of the International Seed Test Association (ISTA) 1996. The environment of germination test was in the temperature of 20-30°C and the relative humidity was 72% only. The time of the experiment was from September 2013 to October 2014. Each germination tray was tested within 2 weeks. The system was applied four techniques for quality control process. The accuracy based on training data set of ED, RBS, FL, and ANN was 100%, 96.17%, 100% and 100% respectively. The average access time based on ED, RBS, FL, and ANN was 4.36, 5.39, 6.08, and 162.05 seconds per image, respectively. The accuracy based on unknown data set of ED, RBS, FL, and ANN was 87.50%, 100%, 100% and 100% respectively. The average access time based on ED, RBS, FL, and ANN was 4.35, 5.29, 7.04, and 159.65 seconds per image, respectively.

6. RECOMMENDATIONS
According to the experiment results, it was proved that the RiSGA system was efficient, effective, accurate, upgradable, and objective. However, the RiSGA system was used for the experiment for the only six Thai rice seed species collected in Thailand. Therefore, the RiSGA system should be used for the experiment for more data set in order to cover all rice seed species not only for rice seeds in Thailand but also for rice seeds in all parts of the world. To increase the efficiency of the system, the RiSGA system can be applied other feature selection to observe the key important features which are suitable for the rice seed characteristics in the rice germination, can be enhanced the Data Set to support not only Thai rice seed species but also all kinds of the rice seed species by increasing the variation of rice seed species in system database, can be applied other substratum in ISTA to enhance the ability of processing rice seed quality control process, can be improved image acquisition for supporting mobile phone devices which can be further developed for the online system application.

7. ACKNOWLEDGMENTS
This research is supported by Bureau of Rice Seed, Ministry of Agriculture and Cooperatives; Rice Department of Thailand [6] and Faculty of Science, Mahidol University for supplying rice seed species and facilities to the research.

8. REFERENCES


[40] Ricepedia - The online authority on rice [homepage on the Internet]. Retrieved November 17, 2013 available from: http://ricepedia.org


9. AUTHOR BIOGRAPHIES

Benjamaporn Lurstwut received her B.S. in ICT and M.S. in computer science from Mahidol University, Bangkok, Thailand. She is currently a Ph.D. candidate in computer science in the faculty of Information and Communication Technology, Mahidol University, Bangkok, Thailand. Her research interests include image processing, image segmentation, machine learning, and pattern recognition.

Chomtip Pornpanomchai received his B.S. in general science from Kasetsart University, M.S. in computer science from Chulalongkorn University and Ph.D. in computer science from Asian Institute of Technology. He is currently an assistant professor in the faculty of Information and Communication Technology, Mahidol University, Bangkok, Thailand. His research interests include artificial intelligence, pattern recognition and object-oriented systems.