PAPER • OPEN ACCESS

Development of colour co-occurrence matrix (CCM) texture analysis for biosensing

View the article online for updates and enhancements.

Development of colour co-occurrence matrix (CCM) texture analysis for biosensing

Y Hendrawan[1], M R Fauzi[2], N S Khoirunnisa[1], M Andreane[1], P O Hartianti[1], T D Halim[1] and C Umam[1]

[1]Department of Agricultural Engineering, Universitas Brawijaya, Malang, Indonesia
[2]Department of Industrial Management, Graduate School of Technology Management, Institut Teknologi Sepuluh November, Surabaya, Indonesia

E-mail: yusufhendrawan@gmail.com

Abstract. In various studies on pattern recognition using image analysis, many researchers used gray-level co-occurrence matrix (GLCM) texture analysis. In this study, a colour-based texture analysis software was developed based on colour co-occurrence matrix (CCM) to be applied as bio-product sensing. CCM involves various colour space i.e. RGB, gray, HSV, HSL, Lab, Luv, LCH, XYZ, CMY and CMYK. The CCM texture analysis software has been tested successfully in four cases of bio-product sensing i.e. instant gatotqualityclassification; coffee type classification; tofu quality classification; and tempequalityclassification. The results obtained showed that CCM texture analysis software has a good performance to be implemented for biosensing.

1. Introduction

Grey Level Co-occurrence Matrix (GLCM) texture analysis method is the most commonly used method of texture analysis. However, relatively little is known about the relationship between colour information and gray-GLCM textural attributes [1]. In addition, gray-GLCM textural features tend to be globally adaptable but not locally optimized. Shearer [2] suggested use of colour texture analysis to overcome the shortcomings of traditional grey-level texture analysis. In agricultural product processing, the commonly used colour spaces such as RGB, HSL, HSV and L*a*b*, have often been used in image processing [3]. The HSL, HSV, and L*a*b* colour spaces were developed to replicate the human colour perception. Colour texture analysis using Colour Co-occurrence Matrix (CCM) method is based on the hypothesis that use of colour features in the visible spectrum provides additional image characteristics over the traditional grey-level representation. Many studies have proved the benefit of using texture analysis for biosensing techniques [4-9]. The objective of this study was developing texture analysis software based on the CCM for bio-product sensing technology.

2. Materials and Methods

First process is image acquisition (Fig. 1), in which the bio-product images were captured using digital camera (Logitec HD Webcam C270, Japan) placed at 300 mm perpendicular to the sample surface and was connected to the USB port of a computer with Intel core i7 processor. The digital camera is used for image acquisition which provides images in BMP format. Images were captured with its maximum resolution (1280 x 720). Imaging was done under controlled and well distributed light conditions.

Light was provided by two 22W lamps (EFD25N/22, National Corporation, Japan). Light intensity over the object surface was uniform at 300 lux in the *center* of the region during image acquisition.
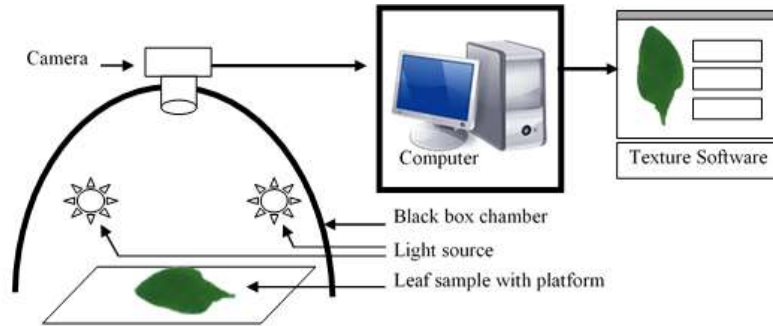

Figure 1. Design of image acquisition system.

The textural analysis can be considered as one of applicable techniques for extracting image features [10]. The CCM procedure consists of three primary mathematical processes: (1) the image is transformed from RGB colour representation to other colour representation such as gray [11], HSL and HSV [12], L*a*b* and XYZ [13], LCH [14] and Luv [15], CMY and CMYK; (2) calculation of Spatial Gray-Level Dependence Matrices (SGDMs), resulting in one CCM for each colour space, the CCM was calculated based on normalization value; and (3) determination of ten Haralicktextural features (TFs) [5].

Ten Haralick's TFsare as follows:

$$Energy = \sum_{i}^{M} \sum_{j}^{N} P^2[i,j] \tag{1}$$

$$Entropy = -\sum_{i}^{M} \sum_{j}^{N} P[i,j] \log P[i,j] \tag{2}$$

$$Contrast = \sum_{i}^{M} \sum_{j}^{N} (i-j)^2 P[i,j] \tag{3}$$

$$Homogeneity = \sum_{i}^{M} \sum_{j}^{N} \frac{P[i,j]}{1+|i-j|} \tag{4}$$

$$Inverse\,Difference\,Moment = \sum_{i}^{M} \sum_{j}^{N} \frac{P[i,j]}{|i-j|^k} \quad i \neq j \tag{5}$$

$$Correlation = \sum_{i}^{M} \sum_{j}^{N} \frac{(i-\mu)(j-\mu)P[i,j]}{\sigma^2} \tag{6}$$

$$Sum\,Mean = \frac{1}{2} \sum_{i}^{M} \sum_{j}^{N} (iP[i,j] + jP[i,j]) \tag{7}$$

$$Variance = \frac{1}{2} \sum_{i}^{M} \sum_{j}^{N} ((i-\mu)^2 P[i,j] + (j-\mu)^2 P[i,j]) \tag{8}$$

$$Cluster\,Tendency = \sum_{i}^{M} \sum_{j}^{N} (i+j-2\mu)^k P[i,j] \tag{9}$$

$$Maximum\,Probability = \underset{i,j}{Max}^{M,N} P[i,j] \tag{10}$$

Where: *P(i,j)* is the *(i,j)*^th element of a normalized co-occurrence matrix, and μ and σ are the mean and standard deviation of the pixel element given by the following relationships:

$$P[i,j] = \frac{N(i,j)}{M} \tag{11}$$

$$\mu = \sum_{i}^{M} i \sum_{j}^{N} P[i,j] \tag{12}$$

$$\sigma = \sum_{i}^{M} (i-\mu)^2 \sum_{j}^{N} P[i,j] \tag{13}$$

where: $N(i,j)$ is the number counts in the image with pixel intensity i followed by pixel intensity j at one pixel displacement to the left, and M is the total number of pixels. A total of 260 TFs were extracted i.e. 10 TFs each for R, G, B, gray, hue, saturation$_{(HSL)}$, saturation$_{(HSV)}$, lightness$_{(HSL)}$, value$_{(HSV)}$, X$_{(XYZ)}$, Y$_{(XYZ)}$, Z$_{(XYZ)}$, L*, a*, b*, C$_{(LCH)}$, H$_{(LCH)}$, u$_{(Luv)}$, v$_{(Luv)}$, C$_{(CMY)}$, M$_{(CMY)}$, Y$_{(CMY)}$, C$_{(CMYK)}$, M$_{(CMYK)}$, Y$_{(CMYK)}$, and K$_{(CMYK)}$. These are the definition of each textural features: *entropy* measures the randomness of a gray-level distribution; e*nergy* measures the number of repeated pairs; *contrast* measures the local contrast image; *homogeneity* measures the local homogeneity of a pixel pair; *summean* provides the mean of the gray levels in the image; *variance* describes the spread out the distribution of gray levels; *correlation* provides a correlation between the two pixels in the pixel pair; *maximum probability* is the results in the pixel pair that is most predominant in the image; *inverse different moment* (IDM) describes about the smoothness of the image; and *cluster tendency* measures the grouping of pixels that have similar gray level values. For an image of a biological object, statistical textural analysis is suitable.

3. Results and Discussion

The CCM texture analysis software had been developed. The performance had been tested successfully to extract CCM textural features from some bio-product images. Figure 2a shows the appearance of the CCM texture analysis software which has been developed using Visual Basic 6.0. It consisted of three form i.e. form 1 for selecting the group of colour to be extracted; form 2 for setting the number of image to be extracted and the value of distance and angle in texture analysis; and form 3 to show the process of the image during the extraction. Figure 2b shows the output data of CCM textural features which has been extracted by the software to be shown in Microsoft Excel. The next step was testing the software for bio-product sensing. Four researches of bio-product sensing were conducted i.e. (1) classifying the quality of instant gatot using textural features; (2) classifying the varieties of coffee using texture analysis; (3) classifying the quality of tofu using texture analysis combined with artificial neural network (ANN); and (4) classifying the quality of tempe using textural features combined with ANN.
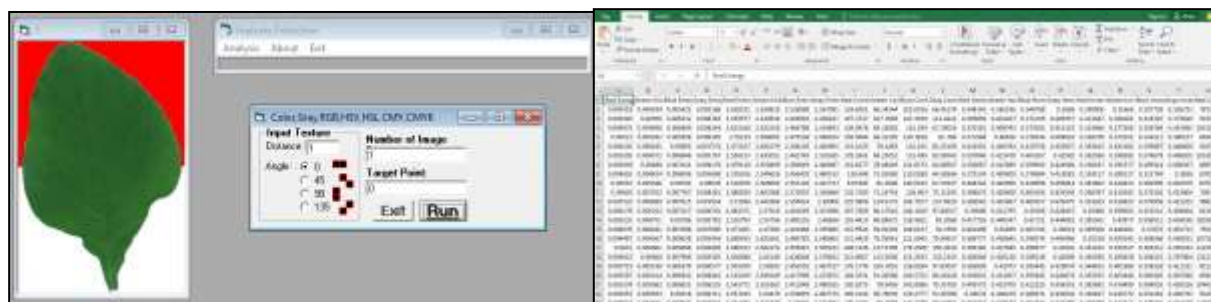

(a)   (b)

Figure2. (a) Design of CCM texture analysis software for biosensing; (b) the output data of CCM textural features shown in Microsoft Excel.

3.1. Classifying quality of instant gatot using texture analysis

The purpose of this testing case was to examine the relationship of textural features to instant gatot quality category as shown in fig. 3. This research used research object in the form of instant gatot quality classified into 3 classes. The total images of the three classes of instant gatot quality were 450 images. Of the 450 images were then extracted to obtain the value of texture. After that, the analysis was done to model the correlation of textural features to instant gatot quality category. The results of the analysis showed that from 260 textural features, there were 10 top-rank textural features which were able to classify all categories of instant gatot quality as shown in table 1, while other textures

were less suitable to be used as classification indicator. The textural features based on its ability to distinguish the category of instant gatot quality can be seen in Table 1.
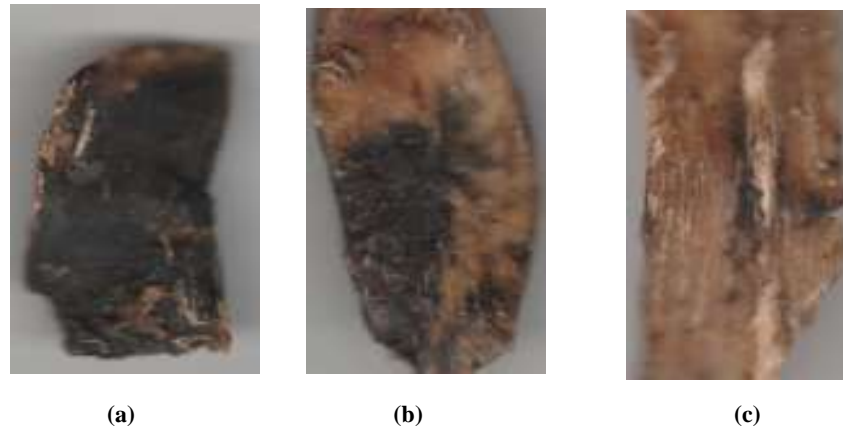

| (a) | (b) | (c) |

Figure 3. The image of instant gatot: (a) quality A, (b) quality B, and (c) quality C.

Table 1. Top ten rank of CCM textural features to classify the quality of instant gatot.

No	CCM Texture	quality A		quality B		quality C	
		Max	Min	Max	Min	Max	Min
1.	Hue Energy	0.55279	0.1317	0.04117	0.01296	0.10523	0.05755
2.	Saturation (HSL) entropy	0.49647	0.09887	0.00993	-0.0024	0.00625	0.00245
3.	Saturation (HSV) entropy	0.48646	0.09284	0.00904	-0.0027	0.00371	0.00166
4.	a (L*a*b) entropy	0.89815	0.62258	0.33579	0.21894	0.51514	0.33351
5.	Saturation (HSL) IDM	0.19134	0.07183	0.38714	0.28981	0.49231	0.41376
6.	Value IDM	0.26964	0.18307	0.36728	0.28007	0.45323	0.38512
7.	Red SumMean	57.3589	44.3369	106.696	78.7121	153.249	123.006
8.	Green SumMean	51.7867	40.416	84.9659	62.6241	124.939	96.2128
9.	GraySumMean	52.7996	41.292	88.6353	65.6779	129.182	100.583
10.	Light SumMean	52.7	41.3443	84.6655	63.7859	121.498	94.3635

3.2. Classifying type of coffee using texture analysis

The second study case aimed to classify the typical local robusta coffee bean and coffee powder using texture analysis and ANN. Image acquisition used 960 images and divided into two categories i.e. 480 images of coffee bean and 480 images of coffee powder. Two ANN models with texture inputs had been built i.e. model 1 with 12 inputs and model 2 with 10 inputs. Both models were used to produce accurate predictions.

The result of the research showed that model 1 with the smallest validation mean square error (MSE) value for the coffee bean image was in $S_{(HSV)}$TFsof 0.012 and ARE 18.26%, for the coffee powder image found on Hue TFs of 0.025 and ARE 22.45%. While on model 2, the smallest validation MSE value for coffee bean image was found in cluster tendency TFs of 0.009 and ARE 17.74%, and for coffee powder image the lowest MSE achieved by cluster tendency texture of 0.013 and ARE 22.86%. The type of robusta coffee and the ANN structure can be seen in Figure 4.
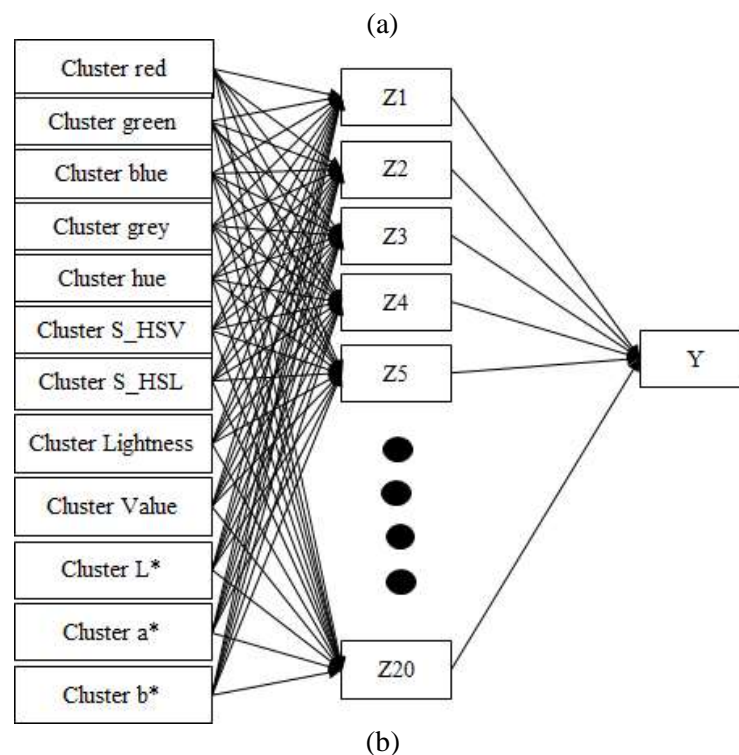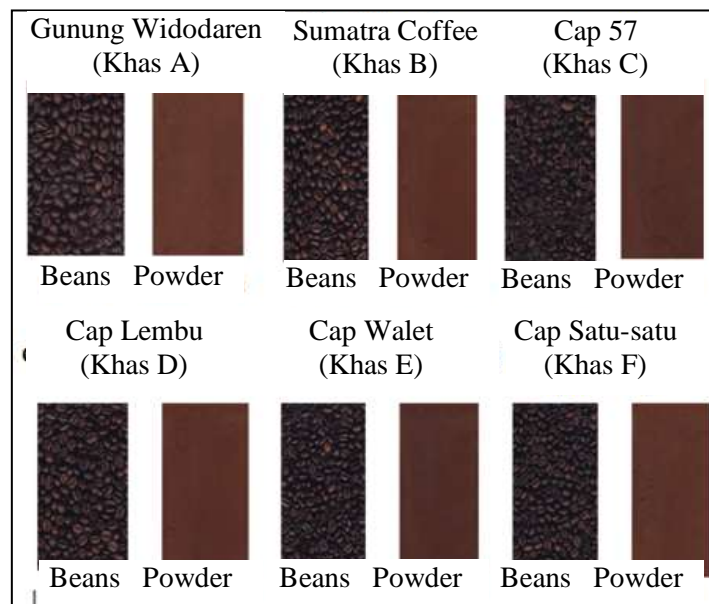
(a)

(b)

Figure 4. The varieties of Robusta coffee and the model of ANN.

3.3. Classifying the quality of tofu using texture analysis

In the third research, the quality classification on tofu based on texture analysis combined with ANN model was observed by using hidden layer neuron 20, learning rate 0.4, momentum 0.1 with 12 input of TFs. From the modeling results showed that the training data Correlation TFs had a small error value with the value of MSE 0.002 and ARE 3.34%, and for the validation data, Sum Mean TFs had the smallest error value of 0.002 (2.94%). Figure 5 shows tofu with the quality of A (fresh) averagely had a degree of high grayness and evenly distributed; tofu with the quality of B (consumption)

averagely had a loose and uneven texture with a low contrast, while tofu with quality C (reject) averagely had a low gray degree and uneven texture which was lower than quality B, and was not smooth. Table 2 shows the ANN-Prediction performance and Figure 6 shows the ANN model using 12 TFs.

Figure 5. Tofu based on its quality.

Tabel 2. Performance of ANN-Prediction model using textural features

| Type of input | Input ANN | | | | Output ANN | | | |
| | Hidden Layer | Iteration | Learning Rate | Moment um | Training | | Validation | |
					MSE	ARE (%)	MSE	ARE (%)
Energy	20	500	0.4	0.1	0.009	6.10	0.025	6.38
Entropy	20	500	0.4	0.1	0.004	5.14	0.012	7.06
Contrast	20	500	0.4	0.1	0.005	7.07	0.008	6.4
Homogeneity	20	500	0.4	0.1	0.009	5.99	0.002	5.87
Inverse Diff. Moment	20	500	0.4	0.1	0.004	5.40	0.003	6.05
Correlation	20	500	0.4	0.1	0.002	3.34	0.003	4.18
Sum mean	20	500	0.4	0.1	0.016	5.12	0.002	2.94
Variance	20	500	0.4	0.1	0.004	4.30	0.003	3.85
Cluster Tendency	20	500	0.4	0.1	0.003	3.99	0.004	6.54
Maximum Probability	20	500	0.4	0.1	0.005	7.05	0.011	6.83
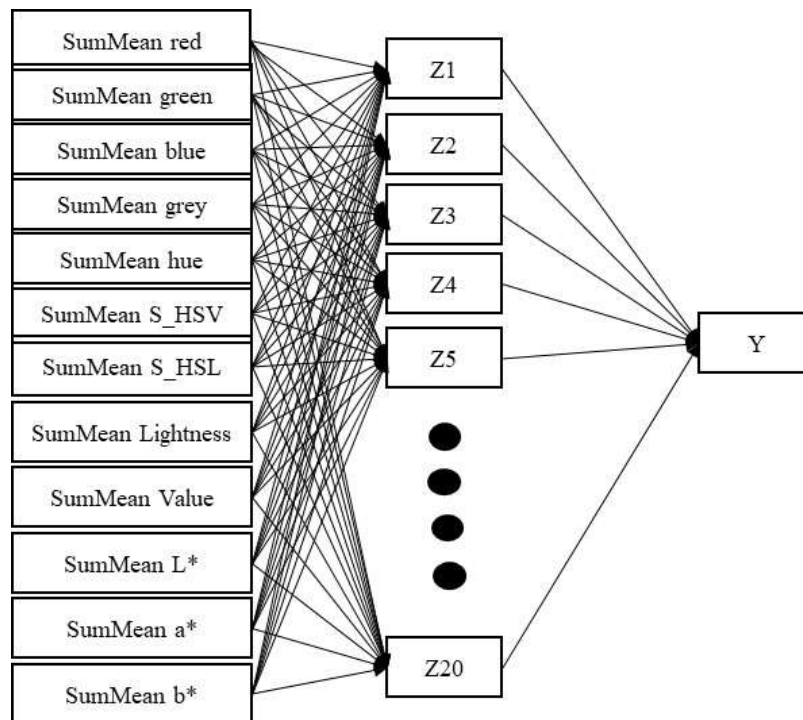
Figure 6. ANN model using texture.

3.4. Classifying the quality of tempeh using texture analysis

The fourth research was made to analyse the estimation of tempeh quality based on texture and modelled with ANN as shown in fig.7. From this research, the best modelling was using cluster tendency TFs with the training value of MSE 0.003 and ARE value 5.70%, and the validation result showed the value of MSE 0.001 and ARE 2.39%. The value of cluster tendency TFs in every colour-space for each type of (1) fresh tempeh; (2) consumable; and (3) reject, respectively i.e. red (0.0499; 0.2246; 0.4017), green (0.1067; 0.3359; 0.5481), blue (0.4174; 0.4061; 0.5135), gray (0.0891; 0.2944; 0.4963), Hue (0.6453; 0.7610; 0.4795), $S_{(HSL)}$ (0.2541; 0.2359; 0.2823), $S_{(HSV)}$ (0.4315; 0.3799; 0.4628), $L_{(HSL)}$ (0.1941; 0.3192;0.4937), $V_{(HSV)}$ (0.1759; 0.2979; 0.4114), L* (0.1022; 0.2918; 0.4664), a* (0.2019; 0.3829; 0.5384), b* (0.5434; 0.2469; 0.2038).
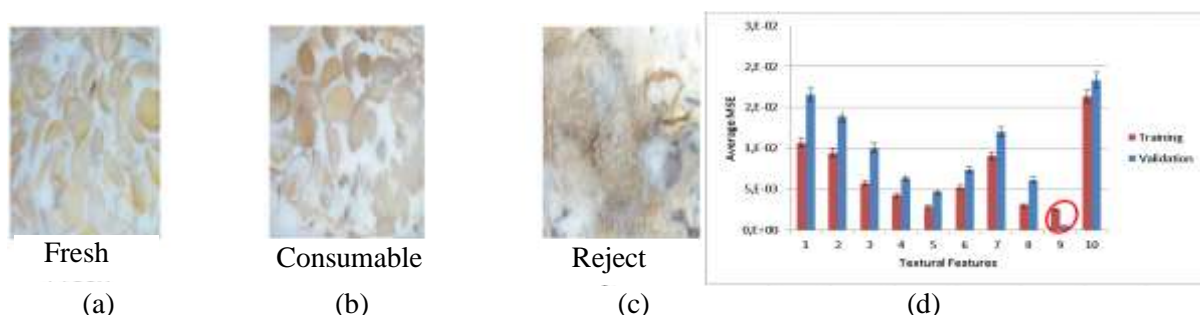

| Fresh | Consumable | Reject | |
| (a) | (b) | (c) | (d) |

Figure 7. Classification of tempeh quality: (a) fresh; (b) consumable; (c) reject; and (c) the result of ANN-prediction using texture: (1) *energy,* (2) *entropy,* (3) *contrast,* (4) *homogeneity,* (5) *IDM,* (6) *correlation,* (7) *sum mean,* (8) *variance,* (9) *cluster tendency* and (10) *maximum probability* (hidden layer 20, learning rate 0.4 and momentum 0.1).

4. Conclusion

The development of colour co-occurrence matrix (CCM) texture analysis software had been done and had been tested for its performance. The use of CCM texture analysis for bio-product had been tested very effectively in some researches i.e. instant gatot quality classification, coffee type classification, tofu quality classification, and tempeh quality classification. The result of prediction accuracy is very good, where the error value was very low. CCM texture analysis software can be an alternative testing and modelling of bio-product sensing.

References

[1]   Yin M, Panigrahi S 2004 Image processing techniques for internal texture evaluation of French fries *Appl. Eng. Agric*. **206** 803-811.

[2]   Shearer S A 1986 Plant identification using colour co-occurrence matrices derived from digitized images PhD Dissertation Department of Agricultural Engineering Ohio State University Columbus Ohio USA.

[3]   Philipp I, Rath T 2002 Improving plant discrimination in image processing by use of different colour space transformations *Comput. Electron. Agric.* 35 1-15

[4]   Hendrawan Y, Murase H 2010 Neural-Genetic Algorithm as feature selection technique for determining sunagoke moss water content *Eng. Agric. Environ. Food (EAEF) J.* **3** 1 25-31

[5]   Hendrawan Y, Murase H 2011 Non-destructive sensing for determining Sunagoke moss water content: Bio-inspired approaches *Agric. Eng. Int.: CIGR Journal* **15** 64 131.

[6]   Hendrawan Y, Murase H 2011 Bio-inspired feature selection to select informative image features for determining water content of cultured *Sunagoke moss Expert Syst. Appl.* **381** 4321-14335.

[7]   Hendrawan Y, Murase H 2011 Neural-intelligent water drops algorithm to select relevant textural features for developing precision irrigation system using machine vision *Comput. Electron. Agric.* 772 214-228.

[8]   Hendrawan Y, Murase H 2011 Neural-discrete hungry roach infestation optimization to select informative textural features for determining water content of cultured *Sunagoke moss Environ. Control Biol.* **491** 1-21.

[9]   Hendrawan Y, Al Riza D F 2016 Machine vision optimization using Nature –inspired algorithms to model *Sunagoke moss* water status *Int. J. Adv. Sci. Eng. Inform. Technol.* **6** 2088-5334.

[10]   Haralick R M, Shanmugam K, Dinstein I 1973 Textural features for image classification *IEEE T. Syst. Man Cyb*. **36** 610-621

[11]   Rotterman Y, Porat M 2006 Colour image coding using regional correlation of primary colours *Image Vision Comput.* **25** 637-651.

[12]   Angulo J, Serra J 2007 Modelling and segmentation of colour images in polar representations *Image Vision Comput*. **25** 475-495

[13]   Leon K, Mery D, Pedreschi F, Leon F 2006 Colour measurement in L*a*b* units from RGB digital images *Food Res. Int.* 39 1084-1091

[14]   Kim M C 2008 Comparative colour gamut analysis of xvYCC standard *Displays* 29 376-385

[15]   Palm C 2004 Colour texture classification by integrative Co-occurrence matrices *Pattern Recogn.* **37** 965-976.