



Journal of Applied Sciences

ISSN 1812-5654

science
alert

ANSI*net*
an open access publisher
<http://ansinet.com>

Scale Invariant Feature Transform Technique for Weed Classification in Oil Palm Plantation

¹Kamarul Hawari Ghazali, ¹Mohd. Marzuki Mustafa, ¹Aini Hussain and ²Saifudin Razali
¹Faculty of Engineering, Universiti Kebangsaan Malaysia, 43600, Bangi, Selangor, Malaysia
²Faculty of Electrical and Electronic Engineering, Universiti Malaysia, Pahang, 25000 Kuantan, Pahang, Malaysia

Abstract: This study presents a new and robust technique using Scale Invariant Feature Transform (SIFT) for weed classification in oil palm plantation. The proposed SIFT classification technique was developed to overcome problem in real application of image processing such as varies of lighting densities, resolution and target range which contributed to classification accuracy. In this study, SIFT classification algorithm is used to extract a set of feature vectors to represent the input image. The set of feature vectors then can be used to classify weed. In general, the weeds can be classified as either broad or narrow. Based on this classification, a decision will be made to control the strategy of weed infestation in oil palm plantations. The effectiveness of the robust SIFT technique has been tested offline where the input images were captured under varies conditions such as different lighting effects, ambiguity resolution values, variable range of object and many sizes of weed which simulate the actual field conditions. The proposed SIFT resulted in over 95.7% accuracy of classification of weed in palm oil plantation.

Key words: SIFT, Gaussian, key descriptor, weed

INTRODUCTION

In the plantation sectors, weeding strategy plays a significant role in managing productivity of palm oil plantation as well as controlling palm oil quality. Currently, most plantation companies adopt manually sprayed herbicide as their weeding strategy which is known to be inefficient, labor intensive and also hazardous to the environment as well as the plantation workers. Furthermore, new regulations have been enforced to limit herbicides or pesticides concentration in drinking water leading to more environmentally responsible weeding strategies (Barnes *et al.*, 1996). At present, herbicides are applied uniformly on the field, even though researches have shown that the spatial distribution of weed is non-uniform. If the non-uniformity of the weed spatial distribution can be identified and detected, then it would be possible to reduce herbicide quantities through efficient weeding strategy (Lindquist *et al.*, 1998; Manh *et al.*, 2001).

In this research, we intend to study and attempt the use of SIFT in classifying different types of weed in palm oil plantation. In controlling weed population, different herbicide formulations are used for different types of weed. In order to avoid environmental pollution due to over usage of herbicides that can contaminate the drinking water supply, it is important that the use of such

formulation be controlled and the minimum amount required is being used. Consequently, an intelligent system for automated weeding strategy is greatly needed to replace manual spraying system that is able to protect the environment and at the same time, produce better and greater yields. Appropriate spraying technology and decision support systems for precision application of herbicides are available but they are unsuitable for the weed types that exist in Malaysian palm oil plantation. To date, the automatic weeding systems have demonstrated potential herbicide savings of 30 to 75% (Stafford, 1997).

Machine vision methods are based on digital images, within which, geometrical, utilized spectral reflectance or absorbance patterns to discriminate between narrow and broad weed. Machine vision methods have been used to show shape features that can be used to discriminate between corn and weeds (Meyer *et al.*, 1994). Other studies classified the scene by means of color information (Cho *et al.*, 2002). Most of the reported study did not maximize and thoroughly analyze the usage of image processing technique but rather used the general technique on standard RGB colour and shape analysis. In the present study, the proposed technique using SIFT, transforms the original image to the values called key point which is found very robust to common problem in image processing such as lighting, range of object and size of image. SIFT technique is based on scale space

fundamental theory which is known as the best technique of structural analysis for object detection. In the image analysis study, there are two common approaches of processing namely statistical and structural techniques. In (Victor, 2005), statistical approach has been used to analyze the weed presence in cotton fields. It was reported that, the used of statistical approach gave very weak detection with 15% of false detection rate. On top of that, structural approach of SIFT technique was introduced to compare the detection and classification rate of weed using statistical approach in (Ishak *et al.*, 2007).

SIFT has been proven to be the most robust technique for object detection (Lowe, 2004). The SIFT key descriptor is well localized in both the spatial and frequency domains, reducing the probability of disruption by occlusion, clutter, or noise. Large numbers of features can be extracted from typical images with efficient algorithms. In addition, the features are highly distinctive, which allows a single feature to be correctly matched with high probability against a large database of features, providing a basis for object and scene recognition. Krystian and Schmid (2001) used the scale invariant technique to detect edge base features object. SIFT technique was used by Ellekilde *et al.* (2006) to develop rescuer robot. Due to many key points generated from an image, it is difficult to analyze the keys for feature extraction. Nowak *et al.* (2006) proposed Bag-of-Feature representation to make the key point easy to analyze.

In this study, since it is impractical to detect and analyze the weed according to its exact species due to large coverage area, we have addressed the problem by using a common approach, which is, by classifying the weed population into two broad classes based on leaf size. The two classes are broad and narrow weed. Classification will be based on the weed images and SIFT-C, the abbreviation for SIFT for Classification, that will be used as the main method of image analysis.

MATERIALS AND METHODS

The overall block diagram of weed classification system is shown in Fig. 1. Firstly, coloured digital images of narrow and broad weed were captured using a C-5050 ZOOM digital camera (Olympus Digital 5.0 Mega pixel) under field lighting condition at noon. An image resolution of 320×240 pixels was used, which then translated into 2 pixels per mm. Images were downloaded using the USB cable interface and stored in JPEG file format. More than 1000 images were used in the processing and testing stage to evaluate the SIFT-C. The

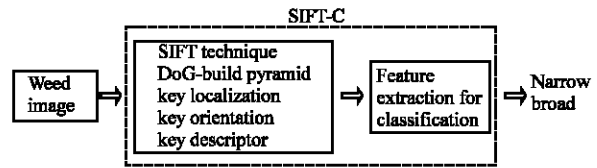


Fig. 1: Overall block diagram of the weed classification method

images were taken under varied condition of weed such as high and low distribution.

The SIFT technique started with the implementation of the Gaussian filter by applying the scale space fundamentals theory (Tony, 1993) which involved different types of Gaussian (DoG) techniques. The pyramid level of scale space consists of a very large data and the SIFT technique has minimized it by determining its peak value. The peak value is measured by implementing the key localization technique which compared the maximum and minimum values of the neighbors from all pyramid level. Practically, the maximum and minimum values of the pyramid level can be interpreted as a distinctive edge of the object in the processed image. Next, we investigate the orientation gradient value to achieve image rotation invariant and rearrange it in the form of orientation histogram. The SIFT key descriptors produce a compact feature descriptor to describe an image using its keypoints. Generally, SIFT keypoint is mainly used for object detection. To recognize and classify an object, a matching algorithm is required. Therefore, in order to utilize the SIFT algorithm for classification purpose, we have proposed a new technique that we named as SIFT-C. The new SIFT-C algorithm is used to classify targeted object, namely the narrow and broad weed. Quelhas *et al.* (2005) proposed a technique for visual scene representation using quantized colour and texture local invariants features computed over regions of point of interests. The use of quantized local invariant features has been proven to provide a robust and versatile way to model images, leading to good classification (Willamowski *et al.*, 2004), retrieval (Sivic *et al.*, 2005) and image segmentation (Dorko and Schmid, 2003). In this study, the SIFT-C is used to derive the magnitude and orientation of the local descriptors and rearrange the histogram into a single keypoint. The feature vector set of SIFT-C is then extracted and used in the design of the classifier system.

SIFT theory: Difference of Gaussian (DoG) is the first stage of SIFT computation. The main function is to search over all scales and images locations to identify potential interest points that are invariant to scale and orientation (Lowe, 2004). Firstly, the raw image with RGB colour code

is converted to gray scale and subjected to the DoG technique. The DoG is similar to the Laplace of Gaussian technique in which the image is first smoothen by convolution with the Gaussian kernel of certain width as shown in (1) and (2) (Lowe, 2004).

$$G_{\sigma_1}(x, y) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left[-\frac{x^2 + y^2}{2\sigma_1^2}\right] \quad (1)$$

$$\begin{aligned} g_1(x, y) - g_2(x, y) &= G_{\sigma_1} * f(x, y) - G_{\sigma_2} * f(x, y) \\ *f(x, y) &= (G_{\sigma_1} - G_{\sigma_2}) * f(x, y) \end{aligned} \quad (2)$$

The DoG as an operator or convolution kernel is defined as (3)

$$\begin{aligned} \text{DoG} &= G_{\sigma_1} - G_{\sigma_2} = \frac{1}{\sqrt{2\pi}} \\ &\left[\frac{1}{\sigma_1} e^{-(x^2+y^2)/2\sigma_1^2} - \frac{1}{\sigma_2} e^{-(x^2+y^2)/2\sigma_2^2} \right] \end{aligned} \quad (3)$$

The initial output of DoG produces a large pixel value at each level and each scale space. In order to reduce the size of DoG output, the local maxima and minima technique need to be implemented to search for the peak value among each level in the pyramid. This peak value contributes to rotation invariance and high level of efficiency. As such, we decide to select key locations at the maxima and minima of a difference of Gaussian function applied in scale space. Maxima and minima of scale space of DoG are determined by comparing each pixel in the pyramid to its neighbors. For instance, a pixel is compared to its 8 neighbors of the same level of the pyramid. If it is a maxima or minima at this level, then the closest pixel location is calculated at the next lowest level of the pyramid. The final step to construct SIFT keypoint is by implementing orientation gradient equation as shown in formula (4). For each DoG image sample, L_x, y , the gradient magnitude, m and orientation θ , is precompiled using pixel differences:

$$\begin{aligned} m &= \sqrt{(L_{x+1,y} - L_{x-1,y}) + (L_{x,y+1} - L_{x,y-1})^2} \\ \theta &= \tan^{-1}((L_{x,y+1} - L_{x,y-1}) / (L_{x+1,y} - L_{x-1,y})) \end{aligned} \quad (4)$$

The construction of SIFT-C feature vector from keydescriptor will involve the determination of orientation θ and magnitude m of every element in DoG smoothing output. At each element in SIFT keydescriptors, we calculate the magnitude of same angle direction and put into 1×128 keypoints. The implementation of SIFT-C formula onto keydescriptor will produce values that can be used to classify object. Here are the steps to construct SIFT-C values.

- Analyze of each k of SIFT keydescriptors by determine its magnitude at every θ (128 bin of degrees).
- SIFT technique will produce $k \times 128$ keydescriptors which is constructed by its magnitude m and orientation θ . For every angle of θ , sum its magnitude so that the keydescriptors are reduced [will reduce] to 1×128 values.
- Implement to the narrow and broad weed and plot graph to visualize the features of object.
- Finally, normalize [normalized] using [with] the maximum values of SIFT-C keypoints.

The above algorithm can be described as (5):

$$K(m, \theta) = \sum_{\theta=0} m_1 + \sum_{\theta=1} m_2 + \dots + \sum_{\theta=m} m_n \quad (5)$$

Normalize the $K(k, \theta)$

$$K = \frac{K(m, \theta)}{\max[K(m, \theta)]} \quad (6)$$

As mentioned earlier, the 1×128 SIFT-C values are distinct and can be used to represent the narrow and broad weed as unique feature vectors. This can be clearly seen in the SIFT-C graph plot of Fig. 1. In order to capture the differences, we used the statistical measurement method that determines the Geometric Mean (GM) and Harmonic Mean (HM) of the data. The geometric mean is

$$m = \left[\prod_{i=1}^n x_i \right]^{\frac{1}{n}} \quad (7)$$

and the harmonic mean is

$$m = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}} \quad (8)$$

The statistics measurements above have been used as features to the input vector for linear classifier design.

$$f = mx + b \quad (9)$$

Where:

$$m = \frac{n \sum (x_i f_i) - \sum x \sum y}{n \sum x^2 - (\sum x)^2} \quad (10)$$

$$b = \frac{\sum y - m \sum x}{n} \quad (11)$$

RESULTS AND DISCUSSION

A set of 1000 sample images consists of narrow and broad weed has been processed using SIFT-C and obtained the keydescriptors. Figure 2 shows the weed images and its SIFT-C keydescriptors plot graph. From the Fig. 2, we can see the SIFT-C values of broad and narrow that have different tabulation, with maximum peak for broad graph occurred at SIFT-C keypoints axes 40-50 and 70-80 while for narrow, its maximum peak values occurred a few time along the SIFT-C keypoints axes. Features for both types of weed have been extracted by measuring its harmonic mean (7) and geometric means (8), respectively. To prove the robustness of SIFT-C; we also test the technique onto image that has a small weed distribution. The sample image of weed with small weed distribution can be seen in Fig. 2. Normally, any changes on the content of images will impinge on the accuracies of the processing technique. In our observation (Fig. 3), using

the SIFT-C technique indicates that the pattern of graph for narrow and broad have very minimal changes as compared to the normal image in terms of their pattern and values. The maximum peak values of broad are still in the range of 40-50 and 70-80 while for narrow, its maximum peak values occurred a few time along the SIFT-C keypoints axes.

Another sample of problem image that has been tested is an image that has been enlarged to 50% of its original size of 240×320 pixels. This can be interpreted as camera movement in a real time application. Resizing from original size of image will automatically change the structure of the pixel values. This can be seen when we compare the quality of the original and resized image. Usually, any changes in size of an image will affect the final result of the classification technique. Figure 4 shows the samples of problem images with different resolution (zooming) by resize the original with 50% of its size. In Fig. 5, original image has been adjusted by increasing the

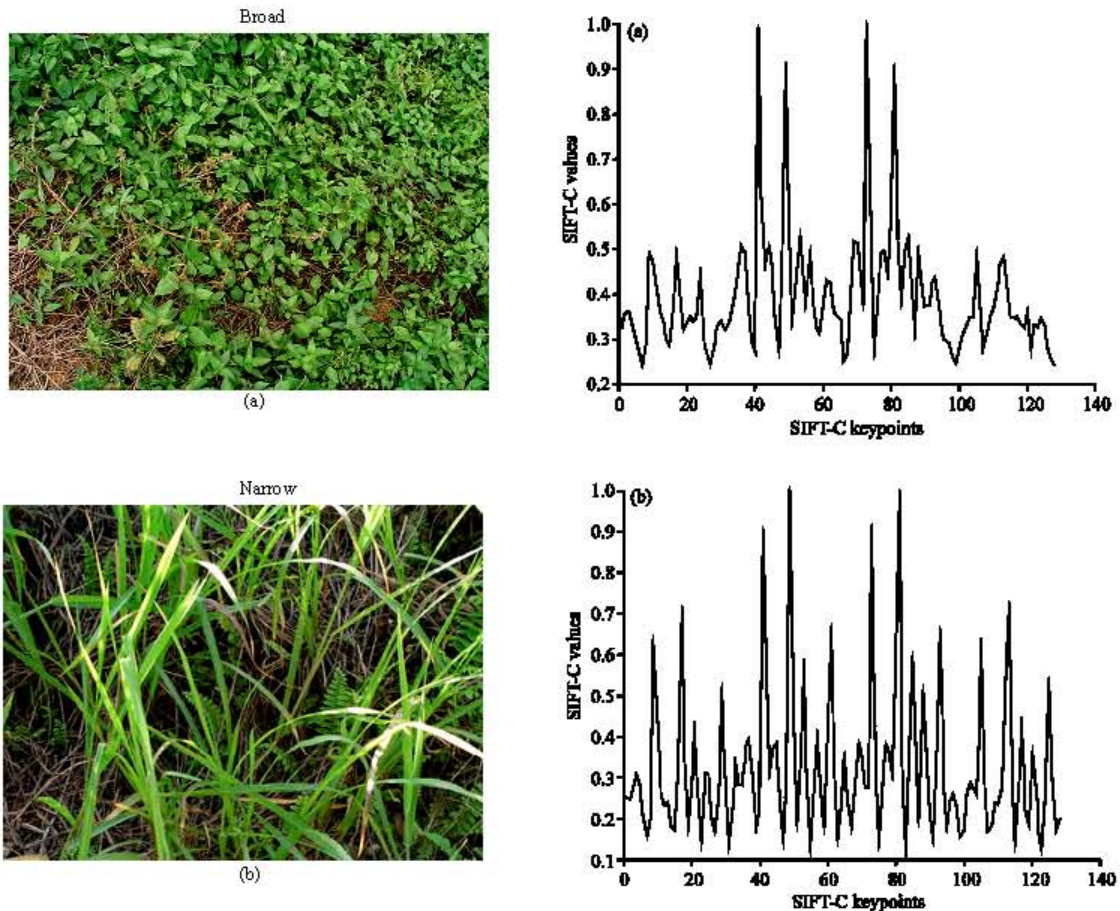


Fig. 2: Plot graph of 1×128 SIFT-C values of narrow and broad weed

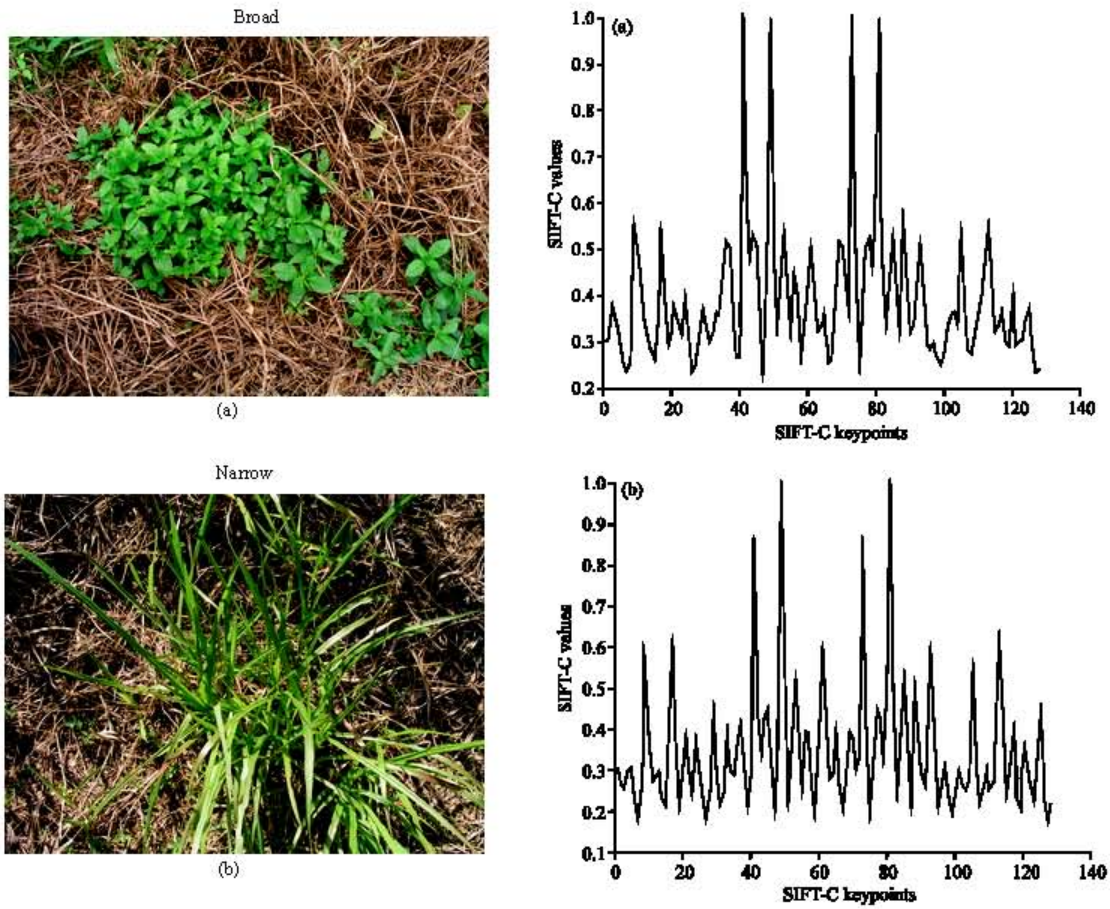


Fig. 3: SIFT-C plot for image with little weed distribution (a) Broad and (b) Narrow



Fig. 4: Sample images of 50% zooming from original images (a) Broad and (b) Narrow

brightness with different constant factor. Constant factor of 1.5 and 2.5 will increase the brightness of original images, meanwhile constant factor of 0.5 will give original images less brightness.

The sample image that used to verify the SIFT-C technique consists of all conditions as discussed above. Figure 3-5 show the types of sample images problem used for testing. These images represent the various

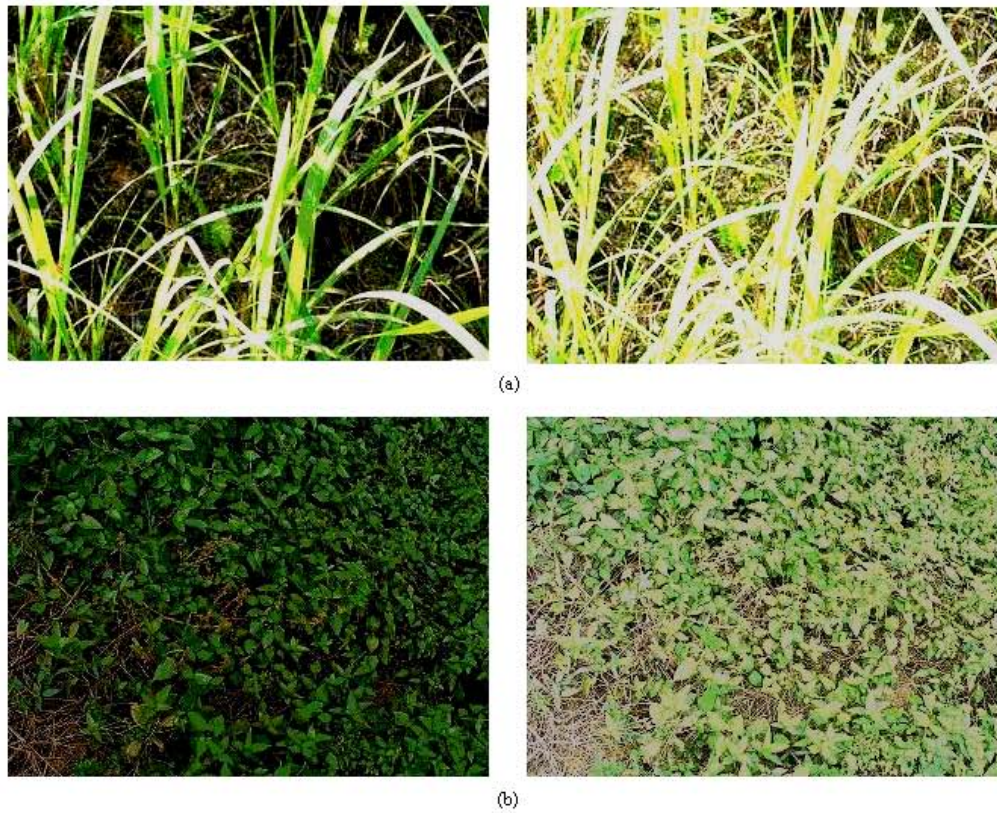


Fig. 5: Sample images of narrow and broad weed with different lighting effect. (a) Narrow with constant lighting of 1.5 and 2.5 and (b) Broad with constant lighting of 0.5 and 1.5

conditions that need to be faced and posed as a challenge to develop a robust technique. Using these types of images, the effectiveness of the SIFT-C technique can be proven. The SIFT-C data has been analyzed using statistical measurement Geometrical Mean (GM) and Harmonic Mean (HM) to extract its feature vectors. The feature vectors which are obtained from (7) and (8) can be seen in Fig. 6. From the plotted graph, the feature vector of SIFT-C has been distributed into two different clusters represent narrow and broad weed.

The feature vectors of raw images as shown in Fig. 6 measured its statistical properties. The purpose is to see the patent of the values in order to choose the best classifier tools for classification. It is clearly seen in Table 1 and 2 that the statistic values of both GM and HM are significantly justified in two different clusters. For example, the mean of GM and HM for both types of weed are 0.32, 0.23 and 0.73 and 0.69. The other values of properties such as standard deviation, max. and min. show that a linear classification can be chosen as a tool to classify narrow and broad weed types.

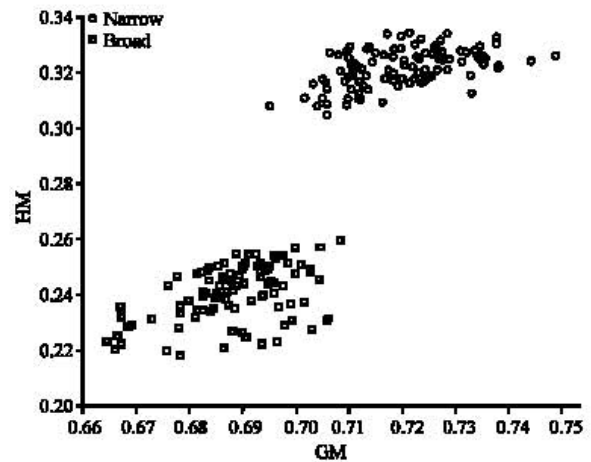


Fig. 6: Feature vector of narrow and broad weed from original sample image

The same step was taken to measure the feature vectors to the problem images which is the brightness has been increase and decrease by certain constant. Figure 6-9 show the values of feature vector of their

Table 1: Statistical properties of 1000 images of narrow and broad harmonic mean feature vector

	Narrow	Broad
Mean	0.32	0.23
Standard Deviation	0.00169	0.0119
Max	0.34	0.26
Min	0.30	0.2

Table 2: Statistical properties of 1000 images of narrow and broad geometrical mean feature vector

	Narrow	Broad
Mean	0.73	0.69
Standard Deviation	0.0168	0.0132
Max	0.75	0.70
Min	0.70	0.66

Table 3: Classification efficiency of original sample image

	Correct Classification
Narrow	99.6%
Broad	99.8%
Overall	99.7%

Table 4: Percentage of classification result for different brightness

	Constant 0.2	Constant 0.5	Constant 1	Constant 1.5	Constant 2
Narrow	90.1%	98.6%	99.6%	99.6%	94.0%
Broad	91.3%	99.8%	99.8%	99.8%	94.8%
Overall	90.7%	99.7%	99.7%	99.7%	94.4%

Table 5: Percentage of classification result for different zooming

	70%	100%
Narrow	98.6	99.6
Broad	98.2	99.8
Overall	98.4	99.7

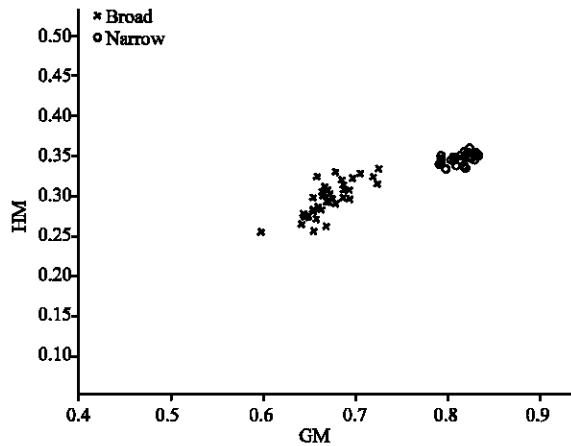


Fig. 7: Feature vector of GM and HM for brightness constant 1.5

respective brightness constant. Increasing the brightness by scaling 1.5 doesn't affect the class of feature vector. This can be seen in Fig. 6-9 where the values of GM and HM are still in two different clusters to represent broad and narrow weed.

A similar case is depicted in the zoomed image (Fig. 10, 11). We enlarge the image to 70% from its original

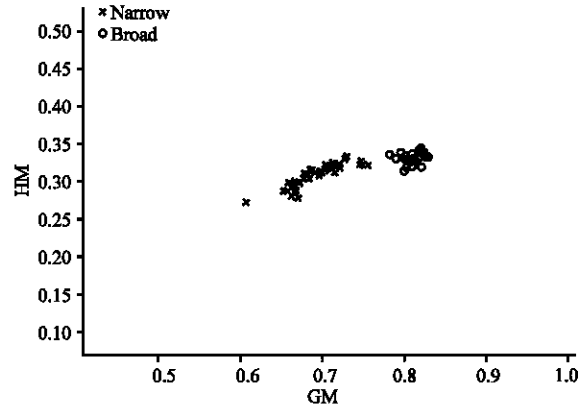


Fig. 8: Feature vector of GM and HM for brightness constant 2.0

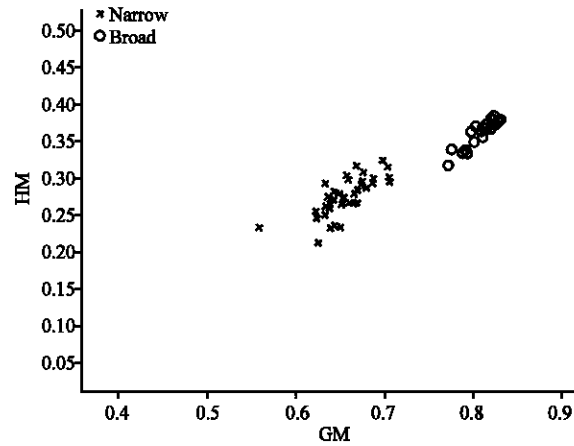


Fig. 9: Feature vector of GM and HM for brightness constant 0.5

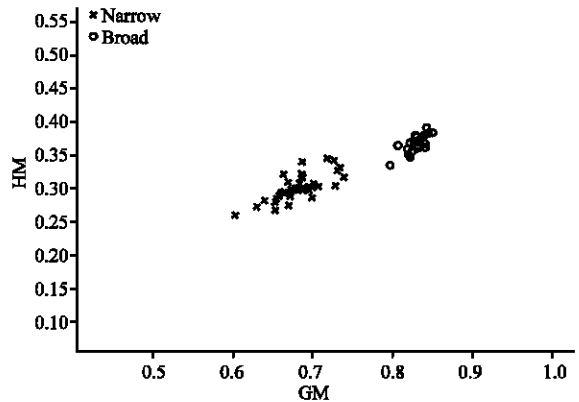


Fig. 10: Feature vector of GM and HM for zooming 70% size and the values of feature vectors are remain in the cluster of narrow and broad class. This implies that the

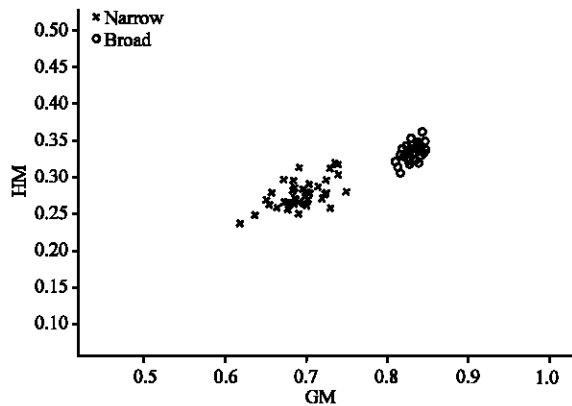


Fig. 11: Feature vector of GM and HM for zooming 50%

SIFT technique is robust towards different lightings as well as unsteady camera movement.

Finally, the feature vectors obtained through HM and GM have been used as an input to the classifier design. A simple classifier system as which was defined in (9), (10) and (11) has been used and the equation obtained was $f = -2.336559x_1 + 44.193866x_2 - 10.822155$. The linear equation f was used to test the classification efficiency of 1000 sample images. Table 3 shows the correct classification rate of original sample weed image. It is found that the narrow and broad weeds are well classified using SIFT-C with efficiency of 99.6 and 99.8%, respectively.

Table 4 and 5 show the testing results of different image conditions that were manually adjusted from the original image by varying the brightness of the images as well as resizing by 50 and 70%. The overall classification rate of brightness images with constant factor 0.5 remains unchanged. Further decrease to a factor of 0.2 causes the accuracy to drop to 90.7%. This also happened to the image with constant factor of brightness increase up to 2. The efficiency was dropped but still in the acceptable error.

The fundamental idea of image processing techniques is based on structural analysis of SIFT-C technique. Ishak *et al.* (2007) used a statistical analysis of histogram technique to classify the same weed in palm oil plantation. The result obtained from her studies showed the poor classification efficiency with the overall correct classification rate at 88.75%. Statistical approach of image processing technique is very sensitive to its pixels value as any changes of image due to noise such as lighting or bad resolution will automatically affect the final result. The use of statistical approach of SIFT-C technique was successful classified the type of weed as the efficiency is constantly above 90%.

CONCLUSION

The adoption of SIFT for object classification instead of object detection was successful in classifying narrow and broad weed. Figure 6 shows the overall results of successful classification with different image conditions. We test our image by manually adjusting the brightness to see the robustness of SIFT-C technique. After increasing the brightness of the pixel value by a factor of 1.5, the classification rate remains until the factor was doubled. At this point, the classification rate decreases to 94 and 94.8% for narrow and broad weed, respectively. We also test for different sizes of image as this can be perceived as zooming or camera moving in real application. The performance rate does not change much for the 70% decrease of the image. To conclude, the SIFT-C has a robust characteristic and shows better performance for narrow and broad weed classification.

REFERENCES

- Barnes, E.M., M.S. Moran, P.J. Pinter and T.R. Clarke, 1996. Multispectral remote sensing and site-specific agriculture: Examples of current technology and future possibilities. In: Proceedings of the 3rd International Conference on Precision Agriculture, pp: 297-302.
- Cho, S.I., D.S. Lee and J.Y. Jeong, 2002. Weed-plant discrimination by machine vision and artificial neural network. *Biosyst. Eng.*, 83 (3): 275-280.
- Dorko, G. and C. Schmid, 2003. Selection of scale invariant parts for object class recognition. In: Proceedings of IEEE International Conference on Computer Vision, Nice.
- Ellekilde, L.P., J.V. Miro and G. Dissanayake, 2006. Fusing range and intensity images for generating dense modes of three-dimensional environments. In: Proceedings of International Conference on Man-Machine Systems. Langkawi, Malaysia, 1: 200-206.
- Ishak, A.J., M.M. Mustafa, A. Hussain and K.H. Ghazali, 2007. A comparison of grey level co-occurrence matrix and histogram techniques using neural network for automatic weed detection. In: Proceedings of International Conference on Man-Machine Systems. Langkawi, Malaysia, 1: 310-314.
- Krystian, M. and C. Schmid, 2001. Indexing based on scale invariants interest points. International Conference on Computer Vision, Vancouver, Canada, pp: 525-531.
- Lindquist, J.L., A.J. Dieleman, D.A. Mortensen, A. Johnson and D.Y. Wyse-Pester, 1998. Economic importance of managing spatially heterogeneous weed populations. *Weed Technol.*, 12: 7-13.

- Lowe, D.G., 2004. Distinctive image features from scale invariants keypoints. *Int. J. Comput. Vision*, 60 (2): 91-110.
- Manh, A.G., G. Rabatel, L. Assemat and M.J. Aldon, 2001. Weed leaf image segmentation by deformable templates. *Automation and Emerging Technologies*. Silsoe Research Institute, pp: 139-146.
- Meyer, G.E., K.V. Borgen, D.M. Woebbecke and D.A. Mortensen, 1994. Shape features for identifying young weed using image analysis. *American Society of Agriculture Engineers, St. Joseph, MI USA.*, 94: 3019.
- Nowak, E., F. Jurie and B. Triggs, 2006. Sampling Strategies for Bag-of-Features Image Classification. *ECCV part IV, LNCS 3954*, pp: 490-503.
- Quelhas, P., F. Monay, L.M. Odobez, D. Gatica-Perez, T. Tuytelaars and L. Van Gool, 2005. Modeling scenes with local descriptors and latent aspect. *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV'05)*.
- Sivic, J., B.C. Russel, A.A. Efros, A. Zisserman and W.T. Freeman, 2005. Discovering object categories in image collection. *ICCV05, Beijing*.
- Stafford, J.V., 1997. Machine-Assisted Detection of Weeds and Weed Patches. In: *Precision Agriculture '97*. Stafford, J.V. (Ed.). Technology, IT and Management, BIOS Scientific Publishers Ltd., 2: 511-518.
- Tony, L., 1993. *Scale Space Theory in Computer Vision*. Royal Institute of Technology, Stockholm, Sweden, Kluwer Academic Publisher, 1993.
- Victor, A., R. Leonid, H. Amots and Y. Leonid, 2005. Weed detection in multi-spectral images of cotton fields. *Comput. Elect. Agric.*, 47 (3): 243-260.
- Willamowski, J., D. Arregui, G. Csurka, C. Dance and L. Fan, 2004. Categorizing nine visual classes using local appearance descriptors. In: *Proceedings of LAVS Workshop in ICPR 04 Cambridge*.