Weed detection and classification for autonomous farming

S. Kodagoda[a], Z. Zhang[a], D. Ruiz[b], G. Dissanayake[a]

[a] *ARC Centre of Excellence for Autonomous Systems (CAS), University of Technology, Sydney, AUSTRALIA*
[b] *Australian Centre for Precision Agriculture (ACPA), University of Sydney, Sydney, AUSTRALIA*

Abstract

Autonomous weed control concepts have recently being extensively researched due to the advantages that they possess. One of the critical modules of such systems is the sensing and classification of weeds within crops. In this paper, we systematically chose the sensing setup and cues to be used for classification of two common weed species (Bidens pilosa L. and Lolium rigidum L.) in a wheat crop. An automatic cue selection followed by classification procedure is proposed. Some classification results are presented while discussing problems leading to future direction of research.

Keywords: Weed detection, spectral analysis, classification

1. Introduction

There are a number of field operations that can be executed by autonomous vehicles, giving more benefits than conventional man-guided machines [1]. These vehicles would be able to work unattended over long periods of time, carrying out useful tasks such as cultivation and seeding, weed control, soil scouting, application of fertilizers, irrigation, and harvesting. Automatic guidance may avoid over-application of chemicals and fertilizers, reducing environmental impact [2]. As weed populations have been found to be distributed heterogeneously in time and space within agricultural fields, weed control systems based on vision have been developed to spray specifically the weed infested areas in real-time, reducing treatment costs as well as herbicide loading to the environment [3-4].

The authors are collaborating with [5] to realize an autonomous weed control system based on CASPA weeder shown in Fig. 1. The CASPA weeder has the capabilities of remote control, joystick based control, sensing and data logging. It can be programmed for autonomous navigation. It has accurate GPS localisation system and weighs 89kg with approximate size of 800L x 550W x 400H (mm).

Fig. 1 CASPA weeder

It is planned to implement a fully insulated and isolated electrocution cradle extending out at the back of the robot to be used to destroy the weeds. It will have five independent electrodes at 20kV covering 250 mm width and a spiked castor wheel as the ground electrode. The authors' main contribution to this project is to synthesise a low cost, real time crop-weed classification algorithm.

In the literature, crop-weed classification falls into two basic categories: spectral data based classification and computer vision based classification. Spectral data based technology relies on the difference of spectral response of each plant species. Jurado et al [6] use a NIRS monochromator to measure plants' spectral reflectance in a lab environment. The spectral reflectance data was then analyzed statistically. They have shown that the spectral difference in the band of 750nm to 950nm is suitable for discriminating sunflower and wheat stubble. However, the experimental results also contained overlapping spectral responses from dissimilar plants. They proposed to use an airborne hyper spectral camera for better classification results albeit higher costs of the equipment. Borregaard et al [7] adopted a line scan spectral device to implement on site weed detection. In their work, a line scan spectrometer was mounted pointing down observing plants. The line scanning spectrometers provide additional spatial and textural information to that of normal spectrometers with a spot foot print. Classification methods such as LDA, QDA, PCA, and PLS were used to classify the spectral data. However, the accuracy of Borregaard's algorithm was between 60% and 90%. Eddy et al [8] used a hyper spectral camera with a resolution of 640x480. Each pixel from this camera can have a spectral resolution of 10nm within the spectral band of 400nm to 1000nm. A feed forward neural network was trained to classify the spectral data, and a detection rate of 88%-95% was achieved. The hyper spectral camera can capture more information than that of line scan spectrometers. However, the high cost of the hyper spectral camera prevent its feasibility of using in low cost weed detection systems.

Computer vision based techniques have also been extensively exploited for crop-weed discrimination. A color image is rich in information providing cues such as color, texture, shape,…. etc. Perez et al [9] have chosen color and shape as appropriate visual cues. The cues were then analyzed by a heuristical approach, Bayes classification and k-nearest Neighbor classification. They managed to achieve a detection rate of 71% to 89%. Aitkenhead et al [10] proposed to use shape features in a neural network to achieve a detection rate of 50%-90% for crop/weed classification. Hemming et al [11] could only achieve a detection rate of 50% to 85% using shape features. Large variation of the detection performance indicates that the accuracy of the shape parameter calculation is not robust. This can be due to occlusions introduced by proximity leaves.

Lolium rigidum L (commonly known as ryegrass) is a widespread grass weed in cereal fields, causing important yield reductions and having evolved resistance to five major groups of herbicides. In wheat fields in southern Australia yield reduction can be up to 50% at weed densities of 200-500 plants/m2 [12]. *Bidens pilosa* L (commonly known as cobbler's peg) is an annual broad leaf weed widely distributed in tropical and subtropical regions of the world and is reported to be a weed of 31 crops [13]. Therefore, in this work our particular attention is paid to classification of wheat, Bidens and Lolium species.

The rest of the paper is arranged as follows. Section 2 discusses the spectral analysis of the concerned plant species. Section 3 discribes the experimental system. Crop weed classification algorithm is given in Section 4. In Section 5, experimental results are given. Section 6 concludes the paper.

2. Spectral analysis

It is a well-known fact that various plant species have different optical and light absorption properties. That is mainly due to pigments in the UV and visible wavelengths, chemical composition in the NIR range, and numerous leaf tissue structures. Leaf pigments, which contain Chlorophyll absorb large amount of light in the UV band, blue (450nm) and red (680nm) part of the spectrum, whilst slightly lower absorption in the green (550nm) band (see Fig. 2). The leaf pigments do not have a good NIR light absorption property, hence the leaves reflect or diffuse large amount of radiation. The transmitted radiation is further affected by the leaf internal structure, which can be used for classification of diverse plant species.

Fig.2 shows the spectral response of three types of plant species under laboratory conditions. The plants were grown in trays and spectral responses were measured with artificial lighting. Oceanoptic spectrometer was used and measurements were taken on leaves of the species. Fig. 2 shows a large variation in the responses of three species within the spectral band, 750nm to 950nm (NIR). It also shows up another band around 550nm (green). The steep slope 700nm to 750nm is called the “red edge”.

After analyzing several number of data sets taken from the same type of species, it was noted that the spectral responses within a particular plant species can significantly vary causing a straightforward segmentation based on spectral analysis is erroneous. Fig.3 shows the spectral variations of each type of species and the obvious overlapping regions. For example, considering the spectral band 750nm to 900nm, it can be seen that the spectral variation of wheat is large enough to contain some Bidens and rye grass responses.

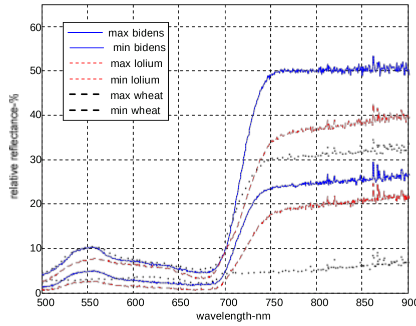


Fig. 2 Variance in spectral reflectance of different plant species

This variation in spectral response can be due to age of leaves, their orientation, spatial position of the leaves and plant health. The spectral responses of a wheat leave obtained at different orientations from the light source are shown in Fig. 3. It could be noted that the spectral responses become weak with larger angles to the light source. In some cases, relatively larger footprint of the spectrometer caused the measurements to be affected by the background, especially for thin leaves such as rye grass.

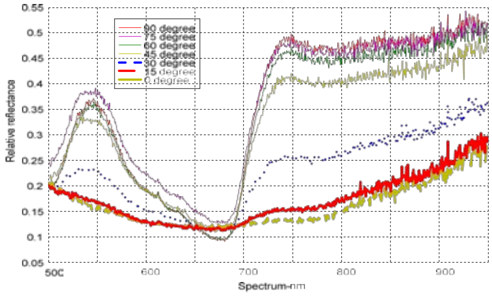


Fig. 3 Variance in spectral reflectance of different measuring angles. Angles are measured from the wheat surface. The light source is positioned at 90 degrees to wheat surface.

The spectrometers may not be a viable solution for online weed/crop classification due to either their higher costs or due to practical constraints. One such practical limitation is the requirement of a reference spectral response at the measurement point for calculating the ratio. Further, it needs a sweeping mechanism for obtaining the special information due to its limited circular footprint. The spectrometer probe also needs to be closely located to the surface before taking measurements. Therefore, it is proposed to use a color camera and a NIR (Near Infra Red) camera to capture the discriminative bands, which are around 550nm and 750nm – 950nm.

3. Experimental setup

Experiments were carried out in a laboratory consisting of a camera set up, data logging software and trays of three plant species namely wheat, Bidens and Lolium. As in Fig. 4, the experimental setup is consisted of a color camera, a NIR camera, a LED lighting source, a laptop computer and data logging software. NIR camera is a monochrome camera fitted with a visual light block filter. Lenses of both cameras were chosen to be identical. The lighting source is specially designed consisting of 5x4 LED arrays with spectral band of 780nm. The colour images were captured with fluorescent lighting and the NIR images were captured with the LED lighting. The image resolution is 640x480.

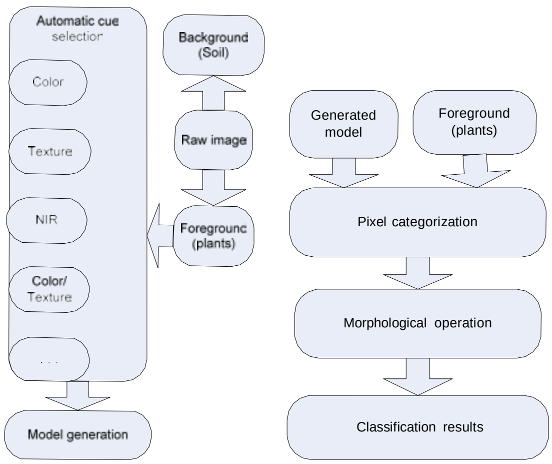
Fig. 4 Experimental setup

4. Crop/Weed classification algorithm

Vision cues have been extensively used in the literature for crop/weed classification. In this section, an unsupervised crop/weed classification algorithm based on vision information alone is presented. Fig. 5 shows a block diagram of the proposed crop/weed classification algorithm. It mainly consisted of two parts, namely model generation and classification.

4.1. Automatic model generation

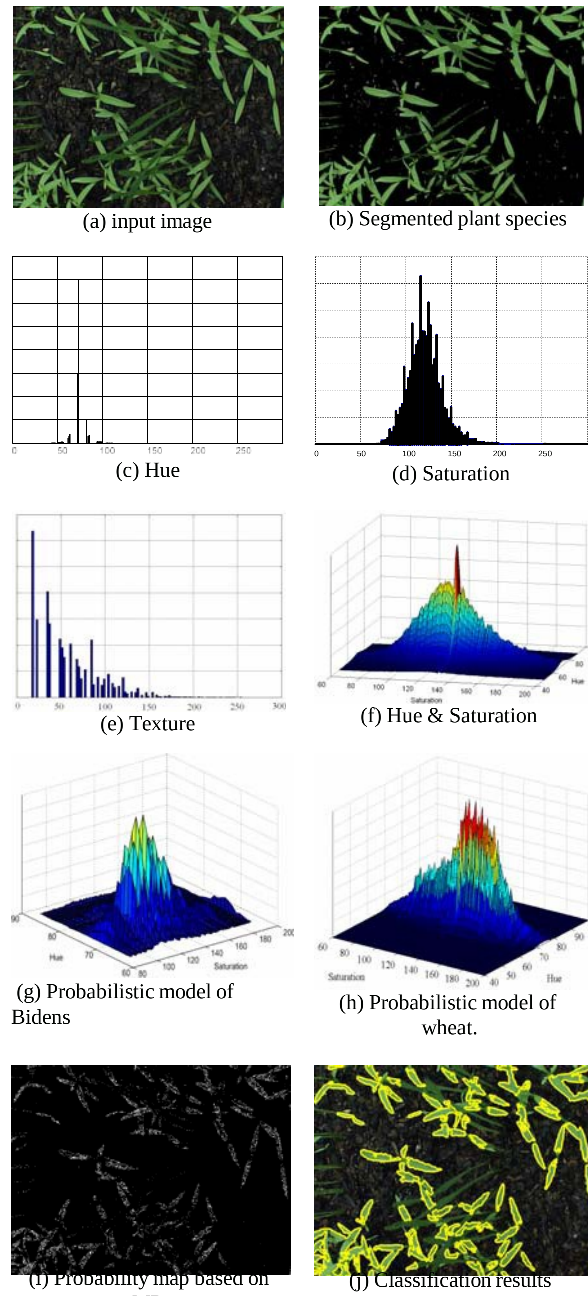
In order to classify plant species, we need to choose appropriate vision cues and models defining the species. We adopt a strategy, where such a selection is carried out automatically requiring minimal input from the operator. This leads the algorithm to be used in diverse classification applications enhancing the generality. Further, it will help an unskilled worker to carry out the operations without having to consult vision experts in selecting appropriate parameters.



(a) Model generation (b) classification

Fig. 5 Classification algorithm

The automatic model generation algorithm is given in Fig. 5(a). Once a raw image is received, a simple color based classification is used for detection of plant/soil (see Fig. 6 (b)). Foreground consisting of pixels belonging to plant species are further processed for calculating various visual cues such as, color, texture, NIR, etc. The cues and their combinations are grouped into several number of clusters based on k-means algorithm. The estimated clusters are further analyzed by Mahalanobis distance (MD), which is a good measure for discrimination of the clusters. The cues or their combinations corresponding to larger MDs are chosen as the appropriate cues or their combinations for the classification problem. The MD is a better choice than Euclidian distance (ED) between cluster centers as latter does not consider the distribution of the clusters. For example, high ED between cluster centers does not necessarily leading to good classification results due to possible larger overlap between the cluster distributions. Given the problem, once the cues or their combinations are chosen as most appropriate, data surrounding the cluster centers are used to determine the model of a particular plant species.



(a) input image (b) Segmented plant species (c) Hue (d) Saturation (e) Texture (f) Hue & Saturation (g) Probabilistic model of Bidens (h) Probabilistic model of wheat. (i) Probability map based on MD (j) Classification results

Fig. 6 Automatic model generation and classification: wheat/ Bidens

4.2.Classification algorithm

Once the model is established, it is possible to synthesize a classification algorithm based on [14], which is shown in Fig. 5(b). By going through each foreground pixel, the MD is calculated with the established model. Pixels that have MD values less than 1 are given a value of 1 and the other pixels are given a value of 1/MD. This is resulted in an “image” describing how confidence a pixel to be belonging to the established model. The confidence image is then thresholded and morphological operations are carried out to improve the connectivity of highly probable pixels while rejecting outliers. The final image shows classified pixels based on the established model.

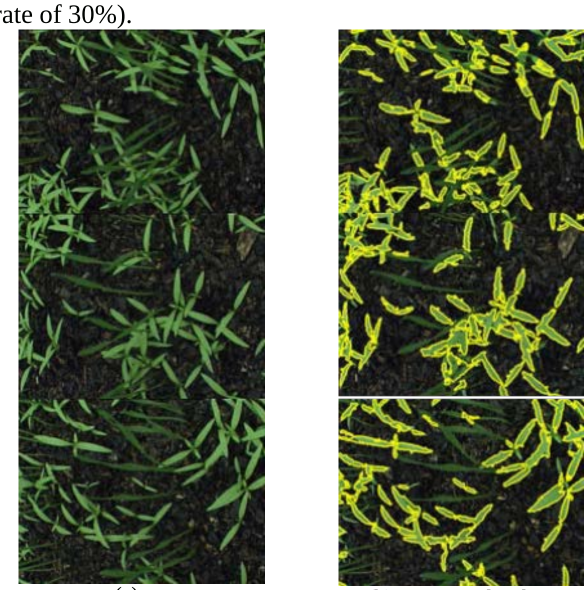
5. Experimental results

In this section, some classification results are presented with an analytical discussion. The plant species considered here are wheat, Bidens and Lolium. Cues considered are, Hue, Saturation, Texture and their combinations. For all scenarios, we assumed the number of clusters to be 2 for the k-means algorithm.

5.1.Classification of Wheat and Bidens

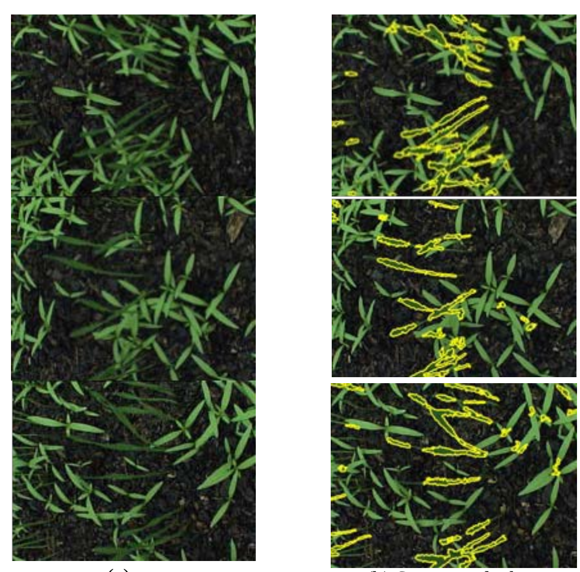
Fig.6 (a) shows a color image of wheat and Bidens plants captured in a laboratory condition with artificial lighting. We used a classification algorithm given in Section 4. Fig.6 (b) shows the classified plant species from soil based on color. The distribution of *hue*, *saturation*, *texture*, and *hue and saturation* of the image Fig. 6 (b) are shown in Fig. 6(c,d,e,f) respectively. Texture reveals the spatial distribution of an image, typically repeated patterns. The common methods of determining texture are Gabor filter, run-length statistics and co-occurrence matrices. In this work, we have used the output of Gabor filter as a measure of texture. It could be noted that the hue is tightly distributed and the saturation has a broad distribution. Although hue and saturation as individual components do not show separable clusters, they form discriminative clusters once combined as in Fig. 6 (f). Therefore, the model generation algorithm has chosen hue and saturation combination to be the most appropriate for Bidens/wheat classification. The resulting models of Bidens and wheat are shown in Fig.6 (g) and (h) respectively. Fig. 6 (i) shows the probability map image generated using the model of the Bidens (Fig. 6 (g)) followed by classification results in Fig. 6 (j).

Further, classification results based on Bidens model is shown in Fig.7. The results are appealing with 97% detection rate and just 2% of false alarms (see Table 1). Fig. 8 shows classification results based on the derived wheat model. The probabilistic model of wheat is ambiguous as shown in Fig. 6 (h), when comparing with that of Bidens (Fig. 6 (g)). Therefore, the classification results (Fig. 8) shows degraded performances than that of using Bidens model (as in Table 1, it has a detection rate of 77% and false alarm rate of 30%).



(a) raw (b) Segmented Bidens

Fig. 7 Classification based on Bidens model



(a) raw (b) Segmented wheat

Fig. 8 Classification based on wheat model

Table 1: Classification results of wheat and Bidens

	detection rate %	false alarm rate %
wheat	77	30
Bidens	97	2

5.2. Classification of wheat and Lolium

We used the same algorithm for wheat/Lolium classification without much success (Fig. 9). As given in Table 2, while using the wheat model, it has a reasonably good detection rate of 85% accompanied by bad false alarm rate of 82% bringing overall classification results to be unacceptable. While using the Lolium model, the detection rate is too low (26%) and the false alarm rate is reasonable high (27%) causing the overall classification to be unacceptable. That is mainly due to the similarities of the two species in the distributions of *hue*, *saturation,* and *texture* cues. Therefore, we are currently working on incorporating NIR images and their cues for improved classification accuracies in such challenging situations.

(a) Raw image (b) Segmented Lolium (c) Segmented wheat

Fig. 9 Classification results: wheat/Lolium

Table 2: Classification results of wheat and Lolium

	detection rate %	false alarm rate %
wheat	85	82
Lolium	26	27

6. Conclusions and future work

Automatic crop - weed classification is a crucial module in autonomous weed control. In this paper, we have investigated a method of classifying wheat from Bidens. We have also shown that visual cues such as color and texture alone do not provide enough information to obtain high classification accuracies especially for wheat and Lolium. This made us to choose NIR image cues along with color camera cues. It leads to other challenges, related to the behavior of the lighting source and multimodal correspondence problem, which we are currently working on.

ACKNOWLEDGEMENTS

This work is supported by the ARC Centre of Excellence program, funded by the Australian Research Council (ARC) and the New South Wales State Government. The authors like to thank Dr Jay Katupitiya and Mr. Kim Son Dang for providing a picture of CASPA weeder and it's description.

References

[1] Pedersen S. M., Fountas, S., Have H. and Blackmore B. S., Agricultural robots-systems analysis and economic feasibility. Precision Agriculture (2006), 7:295-308.

[2] Blackmore S., Fountas S.,Tang L. and Have H., Design specifications for a small autonomous tractor with behavioural control, J. Int. Comm. Agric. Eng. (CIGR). Manuscript PM 04 001. vol 6. July 2004.

[3] Gerhards R. and Oebel H., Practical experiences with a system for site-specific weed control in arable crops using real-time image analysis and GPS-controlled patch system. Weed Research (2006), 46:185-193.

[4] Gee C. H., Bossu J., Jones G. and Truchetet F., Crop/weed discrimination in perspective agronomic images. Com. and Electr. in Agri. (2008), 60:49-59.

[5] Eaton R., Katupitiya J., Cole A. and Meyer C., Architecture of an automated agricultural tractor: Hardware, software and control systems, Proceedings of the 2005 IFAC World Congress. Prague: Elsevier, July 2005.

[6] Jurado-Exposito M., Lopez-Granados F., Stenciano S., Garcia-Torres L., and Gonzalez-Sndujar J. L., Discrimination of Weed Seedlings, Wheat (Tritium aestivum).

[7] Borregarrd T., Nielsen H., Norgarrd L. and Have H., Crop-weed Discrimination by Line Imaging Spectroscopy, J. agric. Eng. Res. vol. 75, pp 389-400, 2000.

[8] Eddy, P.; Smith, A.; Hill, B.; Peddle, D.; Coburn, C.; Blackshaw, R.; Proceedings of Geoscience and Remote Sensing Symposium, 2006. IEEE International Conference; p116 – 119; 2006.

[9] Pérez A. J., López F., Benlloch J. V. and Christensen S., Colour and shape analysis techniques for weed detection in cereal fields, Computers and Electronics in Agriculture, 25(3):197-212, 2000.

[10] Aitkenhead M. J., Dalgetty I. A., Mullins C. E., McDonald A. J. S. and Strachan J. J. C., Weed and crop discrimination using image analysis and artificial intelligence methods, Computers and Electronics in Agriculture 39:157-171, 2003.

[11] Hemming J. and Rath T., Computer-Vision-based Weed Identification under Field Conditions using Controlled Lighting, J. agric. Engng Res., 78(3): 233-243, 2001.

[12] Lemerle D., Verbeek B. and Orchard B., Ranking the ability of wheat varieties to compete with *Lolium rigidum.* Weed Research 41 (3): 197–209, 2001.

[13] Holm G. L., Plucknett D. L., Pancho J. V. and Herberger J. P. , The World's Worst Weeds. Distribution and Biology, FL, USA, 1991.

[14] Bradski G., Kaehler A. and Pisarevsky V., Learning-Based computer vision with Intel's open source computer vision library, Intel Technology Journal, 9(2):119-130,2005.