Multiple sensor based weed segmentation

Sarath Kodagoda and Zhengzhi Zhang

ARC Centre of Excellence for Autonomous System (CAS),

University of Technology, Sydney,

Broadway, Ultimo, NSW 2007

Tel.: +61 2 9514 7629, Fax: +61 2 9514 2655

Email: s.kodagoda@cas.edu.au

Abstract: Bidens pilosa L (commonly known as cobbler's peg) is an annual broad leaf weed in tropical and sub-tropical regions and reported to be a weed that needs to be identified and eliminated when farming thirty one different crop varieties. This paper presents a multi-modal sensing approach for detecting Bidens leaves within wheat plants. Visual cue based automatic discrimination of Bidens and wheat leaves is nontrivial due to the curled up nature of the wheat leaves. Therefore, spectral responses of Bidens and wheat leaves are first analysed to understand the discriminative spectral bands. Then a multi-modal sensory system consisting of a near infra red and a visual camera set up is proposed. Information retrieved from the sensory setup is then processed to generate a series of cues that are fed into a classification algorithm. Classification results are validated through experimentation. The proposed technique is able to achieve an accuracy of 88% - 95% even when there is substantial overlapping between Bidens and wheat leaves. Further, it is also shown that the algorithm is robust enough to discriminate some other commonly available plant species.

Keywords Sensing, precision agriculture, weed detection, classification
1. Introduction

There are a number of field operations in agriculture that can be executed by autonomous vehicles [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 the environmental impact of farming [2]. For example, 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].

The authors are collaborating with Eaton et al [4] to realize an autonomous weed control system based on CASPA weeder, which is 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 global positioning system (GPS) localization system and weighs 89kg with approximate size of 800L x 550W x 400H (mm). 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 caster wheel as the ground electrode. The authors' main contribution to this project is to synthesize a low cost, real time crop-weed classification algorithm.
In the literature, crop-weed classification is most commonly carried out by spectral based or computer vision based algorithms. Spectral based technology relies on the difference between the spectral responses of each plant species. Jurado et al [5] used a NIRS monochromator to measure plants' spectral reflectance in a lab environment. The spectral reflectance data was then analysed statistically. They had shown that the spectral difference in the band of 750nm to 950nm was suitable for discriminating sunflower and wheat stubbles. However, it was also reported that the results were affected by larger variance in the response of a particular type of plant species. Borregarrd et al [6] 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 provided additional spatial and textural information to that of normal spectrometers with a spot footprint. Classification methods such as linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), principal component analysis (PCA), and partial least-squares regression (PLS) were used to classify the spectral data. However, the accuracy of Borregarrd's algorithm was as low as 60% in some cases.

Eddy et al [7] used a hyper spectral camera with a resolution of 640x480 pixels. Each pixel of the camera has 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 could capture more information than that of a line scan spectrometer, however, it comes with a higher cost which could not be justified in most practical weed detection systems.

Computer vision based techniques have also been extensively exploited for crop-weed classification. A colour image is rich of information providing cues such as colour, texture, shape, etc. Perez et al [8] chose colour and shape as appropriate visual cues for crop-weed classification. The cues were then analysed and classified using three algorithms; heuristical approach, Bayes classification and k-nearest Neighbour classification. Detection rate of 71% to 89% was achieved. Aitkenhead et al [9] and Hemming et al [10] proposed to use shape features and achieved detection rate ranging from 50% to 90%. Large variation of the detection rates could be due to poor estimation of shape parameters. Generally, estimation of shape parameters is highly dependent on complexity of occlusions.

Focus of this paper is to segment and classify weed species among crops. As a case study, Bidens (weed) and wheat (crop) were chosen. Bidens 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 [11]. Wheat and Bidens species may have similar colour and near infra-red (NIR) properties. This causes a single modality based classification susceptible to errors. Therefore, in this paper a systematic way of choosing appropriate sensors and sensor cues for improved weed detection results is proposed. In particular, the contributions of this paper are, 1. Synthesis of a novel, low cost multi-modal sensory approach for classification of different plant species; a case study of wheat (crop) and Bidens (weed) is provided 2. Design of a highly
accurate spectrum based Bidens/wheat classifier, 3. Extensive analysis of the classification using spectrometer, colour camera and near infra-red camera data.

2. Spectrometer based classification

2.1 Spectral analysis

It is a well-known fact that optical properties of plants can be used to monitor plant growth, disease, crop status, water and nutrient content, and plant type discrimination. Spectral responses of leaves depend on pigments in the UV and visible wavelengths, and on chemical composition in the NIR range. Numerous leaf tissue structures have significant influence on the spectral properties. Leaf pigments, which contains Chlorophyll absorbs 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). The leaf pigments do not have a good NIR light absorption property, hence the leaves specularly reflect or diffuse large amount of incident radiation. The surface conditions of leaves, such as presence of hair, can affect the specular reflections. The diffuse reflection is due to part of the transmitted radiation, which is affected by the leaf internal structure and air cavities. Specular and diffuse reflections can be effectively exploited for discrimination of plant species.

Fig. 2 (a) shows spectral responses of Bidens and wheat leaves. The plants were grown in trays and spectral responses were measured with artificial lighting using an Ocean Optic™ spectrometer [16]. The steep slope in the response seen between 700nm to 750nm is called the "red edge". It shows the transition from high absorption by chlorophyll in the red to low absorption in the NIR, which generally are used to calculate the "biomass" or "vegetation indices" in remote sensing.
After analysing large number of data sets, it is noted that the spectral responses within a given plant type could significantly vary causing a straightforward segmentation based on spectral analysis erroneous. This could be explained using Fig. 2 (a) and (b). Although in Fig. 2 (a), the spectral responses of Bidens and wheat leaves are seen to be different in green (550nm) and NIR bands (750nm to 950nm) of the spectrum, Fig. 2 (b) shows the spectral variations of Bidens and wheat leaves leading to obvious
overlapping regions. It is to be noted that the term "Max" in the legend refers to the maximum spectral response received for a particular leaf and "Min" refers to the minimum spectral response received for another leaf of the same family of plant species. This variation of spectral response can be due to age of leaves, their orientation, spatial position, plant health, etc.

In order to understand the variation of the spectral responses due to orientation of the leaves with respect to the light source and spectrometer, following experiment was conducted. The measurements were taken for different viewing azimuth angles (0 to 90 degrees at 15 degrees intervals) while keeping the lighting source zenith of 60 degrees and viewing zenith angle of 0 degrees. The angles are defined in the Fig. 3. As shown in Fig. 4, it could be noted that the spectral responses are greatly affected by the viewing azimuth angles. This is due to the variations in specular reflections received by the sensing head.

![Fig.3 Definitions of angles: the ellipsoid refers to the measuring surface](image)
2.2 Classification based on leaf spectral reflectance

In this section, spectral reflectance based classification of Bidens and wheat is presented. Hundred spectral responses of wheat leaves and 100 spectral responses of Bidens leaves were captured. The light source zenith was kept approximately 60 degrees and the view azimuth was kept approximately at 0 degrees to the leaves throughout the experimentation. Regularly spaced 80 data points of the full spectral range (270nm to 1400nm) were considered in the analysis. The classification was implemented using the Weka [12], a popular machine learning software. NaiveBayes, BayesNet, support vector machine (SVM), and sequential minimal optimization (SMO) were chosen as classifiers with K-fold cross validation [17] with the number of samples equal to K (=200). The results are listed in Table 1. SVM and SMO provided the best classification results, although SVM took a longer time to converge. On the other hand, Naive Bayes classifier provided faster response, however with slightly poor performance.
Table 1 Classification results based on the entire spectrum. *Bidens (%) – Detection accuracy of Bidens, wheat (%) – Detection accuracy of wheat*

<table>
<thead>
<tr>
<th>Classification method</th>
<th>Bidens (%)</th>
<th>wheat (%)</th>
<th>Time consumed (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaiveBayes</td>
<td>94</td>
<td>93</td>
<td>1.08</td>
</tr>
<tr>
<td>BayesNet</td>
<td>95</td>
<td>95</td>
<td>3.27</td>
</tr>
<tr>
<td>SVM</td>
<td>100</td>
<td>99</td>
<td>4.36</td>
</tr>
<tr>
<td>SMO</td>
<td>100</td>
<td>100</td>
<td>2.81</td>
</tr>
</tbody>
</table>

Table 2 Classification results based on RGB, NIR and vegetation indices. *Bidens (%) – Detection accuracy of Bidens, wheat (%) – Detection accuracy of wheat*

<table>
<thead>
<tr>
<th>Classification method</th>
<th>Bidens (%)</th>
<th>wheat (%)</th>
<th>Time consumed (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaiveBayes</td>
<td>99</td>
<td>98</td>
<td>0.42</td>
</tr>
<tr>
<td>BayesNet</td>
<td>99</td>
<td>98</td>
<td>1.04</td>
</tr>
<tr>
<td>SVM</td>
<td>99</td>
<td>100</td>
<td>2.29</td>
</tr>
<tr>
<td>SMO</td>
<td>100</td>
<td>99</td>
<td>1.82</td>
</tr>
</tbody>
</table>

Although classification based on the entire spectrum has performed well, utilization of the most discriminative features could lead to improved computational efficiency without degrading the accuracy. Therefore, the visible spectral band (red, green, blue), near infra-red band and 4 vegetation indices [13] derived from data, namely, Vegetation Index (VI), Ratio Vegetation Index (RVI), Transformed Vegetation index (TVI) and red/green ratio were chosen as features. The classification results based on those features are presented in Table 2. When comparing with the Table 1, the feature based classification lead to higher computational savings and
better accuracy due to the low data redundancy and high discriminative nature of the data.

Despite the high accuracy of spectrometer based methods, their feasibility of application to on-line crop/weed classification is constrained by the higher costs or practical constraints. One such practical limitation is the requirement of a reference spectral response at the measuring point for determining the relative reflectance. Another practical limitation is the inherent small footprint, which requires a sweeping mechanism for obtaining the special information. The spectrometer also needs to be closely located to the measuring surface for accurate readings. However, based on the outcomes presented in Table 2, it is clear that the spectral responses in the RGB and NIR bands are adequate for discriminating Bidens from wheat. Therefore, it is proposed to use information gathered from a sensor package incorporating a colour camera (400nm - 650nm) and a NIR camera (720nm - 1000nm) for crop and weed discrimination.

3. Multi-modal sensor based classification

In this section, multi-modal sensor based classification using a colour and NIR camera setup is presented. Firstly, solution to the correspondence problem in the sensory suite is addressed.

3.1 Correspondence problem

Multi-modal sensor based classification in this article requires capturing both the intensity of the light in the RGB and NIR spectrum reflected from a particular point on an object at the same time. There are commercially available cameras, which simultaneously capture visible and NIR through the same optical path achieving the
pixel to pixel correspondence. However such cameras are costly. Therefore, two separate low cost cameras are used in the sensory suite. The two cameras are of multi-modal nature prohibiting the use of conventional stereo matching algorithms. Therefore, in this paper, a unique algorithm based on template matching is proposed to solve the correspondence problem.

A block diagram of the template matching based image registration algorithm is shown in Fig. 5. Firstly, region of interests (ROI) of the colour and NIR images are chosen considering the interested features in the images (Fig. 6 (a) and Fig. 6 (b)). Then the colour image is segmented based on a colour threshold to identify the foreground and background pixels. The foreground pixels are then edge detected to obtain a template image, \( T \). Edges in the NIR image are detected and distance transform (\( I \)) is calculated. Then a Chamfer distance based template matching [15] is carried out for establishing the correspondence.

\[
D_{	ext{chamfer}} (T, I) = \frac{1}{|T|} \sum_{t \in T} d_i (t)
\]  

(1)

where \(|T|\) denotes the number of features in \( T \) and \( d_i (t) \) denotes the chamfer distance between feature \( t \) in \( T \) and the closest feature in \( I \). The minimum value obtained for \( D_{	ext{chamfer}} (T, I) \) is related to a match. Fig. 6 (c) shows the results of the Chamfer distance based template matching algorithm. This algorithm could be applied to many practical scenarios with various shapes of features in multi-modal image registration. However, if the features are at different depths (resulting different disparities), a two stage (global matching and local matching) classification process needs to be implemented for fine tuning the local alignments.
**Global matching:** As detailed above, in the global matching, the whole cropped, edge detected colour image is matched with the NIR distance transformed image. As the Bidens leaves are not at the same height (and hence different disparity values), the global template matching resulted in poor correspondences of some Bidens leaves (see Fig. 7 (a)).

**Local matching:** The global correspondence results are fine tuned using the local matching procedure. A smaller processing area is defined by going through each connected group of foreground pixels in the thresholded colour image for local refinement. Chamfer distance based template matching is performed on this local region for fine alignment leading to a better correspondence of Bidens leaves even with slight height differences (see Fig. 7 (b)).
3.2 Classification algorithm

In this section, the classification algorithm is initially discussed based on colour images only and then generalized to include NIR images.

First, as an offline activity a small window of a plant species (e.g. Bidens leaf in Fig. 8 (a)) in an incoming image is manually selected and used to generate the Bidens colour model, \( m(\mu, \sigma) \); \( \mu \) and \( \sigma \) are the mean and the standard deviation of the colour model distribution. Once the model is established, it is possible to synthesize a classification algorithm based on a probabilistic measure [14], which is shown in Fig. 8 (b). Firstly, the incoming image (Fig. 9 (a)) is processed to eliminate the soil by a colour based classifier (Fig. 9 (b)). Then by going through each foreground pixel, \( I'(i, j) \), the Mahalanobis distance (\( D_{mahal} \)) to the established model is calculated. For all foreground pixels \((i, j)\),

![Fig.7](Image)

**Fig.7** Correspondence results of experimental images: The leaf contours of RGB image is plotted on the NIR image
The resulting "image", $D_{mahal}(i,j)$, describes the confidence that a pixel belongs to the established model (Fig. 9 (c)). The confidence image is then thresholded and morphological operations are performed to improve the connectivity of highly probable pixels while rejecting outliers (Fig. 9 (d)). Fig. 9 (e) shows the final results of the classification (Biden leaves in this case).

The same algorithm can be used for the multi-modal sensory data with little modifications. As discussed in the previous section, once the correspondence procedure is completed, each colour pixel is assigned with a corresponding NIR value. Therefore each colour pixel and hence the leaf model now have an extra dimension corresponding to the NIR value. The rest of the above discussed classification algorithm remains the same with an added NIR pixel dimension.

\[
D_{mahal}(i,j) = f\left(I'(i,j), m(\mu, \sigma)\right)
\]

\[
D_{mahal}(i,j) = \begin{cases} 
1 & \text{if } D_{mahal}(i,j) \leq 1 \\
\frac{1}{D_{mahal}(i,j)} & \text{if } D_{mahal}(i,j) > 1 
\end{cases}
\]  

(a) Colour model generation  
(b) Classification algorithm

Fig.8 Colour model generation and classification algorithm
4 Experimental results

4.1 Experimental Setup

The data logging setup is equipped with an NIR and a colour camera, a laptop computer, and data logging software. The colour camera that is operated in visual band (400nm to 650nm) of the spectrum and the NIR camera that is operated in 720nm to 1000nm band of the spectrum are used to capture images with a resolution of 1024x768. Both cameras have identical lenses, with vertical and horizontal field of views of 43 degrees and 56 degrees respectively. They are rigidly mounted with a baseline of 80mm. NIR camera was fitted with a visual light block filter while the colour camera was fitted with a NIR block filter to minimize noise. The colour and NIR images were captured with natural lighting. Fig. 10 shows a snapshot of the data logging setup.
There are three cases considered. In Case 1 and 2, tray planted Bidens and wheat species were imaged using colour and NIR cameras from the top. The trays were prepared to achieve two levels of complexity. In the first case (case 1), the plants were sparsely located causing minor effects of occlusions (Fig. 11 (a)); wheat leaves do not occlude Bidens leaves.

Fig. 11 Three cases of experimentation
In the second case (case 2), the species were planted close enough to each other leading to significant inter-plant occlusions (some Bidens leaves were occluded by wheat leaves) as shown in Fig. 11 (b). Case 3 consists of other plant species including *Hardenbergia* (*An Australian native plant*), *Thickhead* (*Crassocephalum crepidioides*), and *Eucalyptus kruseana*. Purpose of Case 3 is to generalize the performance of the algorithm to other commonly available plant species.

In this section, firstly the classification results based on visual band (RGB) of the spectrum is presented. Then it will be expanded to utilize both the visual and NIR bands. For the experimentation, 45 number of images with an average of 30 number of plant species in each image (total of 1350 plant species) were used. A model is learned from a Bidens leaf and classification is performed on all images. It is to be noted that all the classification results presented in this section are based on the detected plant leaves and do not reflect pixel level accuracy. The true number of leaves in an image is determined by manual counting.

### 4.2 Classification results based on visual images

Fig. 12 and Table 3 show the colour based classification results. As can be seen from the Table 3, the colour based classification provides high Bidens detection rates in both simple (Case 1) and complex (Case 2) scenarios. However, the false alarm rates are unacceptably high. This is due to the similarity in colour of some parts of the two plant species. In the next section we describe a way to reduce the false alarm rates.
Fig. 12 Colour based Bidens classification results: Left column - raw images, right column – results

Table 3 Classification results based on colour.

<table>
<thead>
<tr>
<th></th>
<th>Case 1 (%)</th>
<th>Case 2 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detection rate</td>
<td>91</td>
<td>90</td>
</tr>
<tr>
<td>False alarm rate</td>
<td>27</td>
<td>43</td>
</tr>
</tbody>
</table>
4.3 Classification results based on multi-modal sensing

Fig. 13 and Fig. 14 show the multi-modal (colour and NIR) sensor based Bidens/wheat classification results for Case 1 and Case 2 data sets respectively. Table 4 clearly shows a significant reduction of false alarm rates when compared with Table 3. In Case 1, the false alarm rate is reduced by 44%, whereas in Case 2, it is reduced by 74%. Slight reduction in the detection rate in the complex case (Case 2) is a result of twisted wheat leaves, which had similar optical properties as Bidens and poor correspondence of some parts of the wheat leaves.

The correspondence problem of wheat leaves is nontrivial to solve due to the taller curled up nature, which has resulted in shape differences of the same leaf imaged by two sensors. Therefore, poor correspondence in some parts of wheat leaves is unavoidable with the current camera set up. However, as the aim of the paper is to detect a particular leaf type with minimal false detections, it is only crucial to have accurate correspondence of the same type of leaves in two images. Sometimes, poor correspondence of other type of leaves in two images can give rise to more discriminative NIR values (generally corresponding to soil) contributing to lower false detections.
Fig. 13 Classification results, Case 1: left most column - colour images, middle column- NIR images, right most column – results
Fig. 14 Classification results, Case 2: left most column - colour images, middle column- NIR images, right most column – results

Table 4 Classification results based on multi-modal sensors.

<table>
<thead>
<tr>
<th></th>
<th>Case 1 (%)</th>
<th>Case 2 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detection rate</td>
<td>95</td>
<td>88</td>
</tr>
<tr>
<td>False alarm rate</td>
<td>15</td>
<td>11</td>
</tr>
</tbody>
</table>
4.4 Classification of other types of plant species

The same algorithm was tested on other plant species without any parametric tuning. Model of Eucalyptus kruseana was learnt and classified based on colour and NIR images, which has resulted in accurate leaf segmentation (Fig. 15). This shows the generality of the algorithm to classify other common plant species.

![Image](image1.png)  ![Image](image2.png)  ![Image](image3.png)

**Fig. 15** Classification results, Case 3: left most column - colour images, middle column- NIR images, right most column – results

Main challenges in real field scenarios are associated with variable sun light as reported by Nieuwenhuizen et al [18]. It is also reported that overcast leads to better classification results as it would alleviate varying intensity values and shadows due to sunlight. Therefore, the problems due to sunlight could be minimized by either fixing a cover on the robotic platform overcastting direct sunlight of the area to be imaged, or producing a reasonable size box (open bottom) containing artificial light sources and the camera set up.
5. Conclusions

In this paper, three types of sensing technologies for classification of leaves of plant species were discussed, namely spectrometer, colour camera and NIR camera. Spectrometer based method provided satisfactory classification results. However, it has practical implementation difficulties. Nevertheless, it allowed us to understand the optical properties to design a practically feasible sensory suite for plant type discrimination. The spectral responses clearly showed the visual and NIR bands are most discriminative for Bidens and wheat leaves. Therefore, the colour camera based classification was first synthesized, which has lead to high detection and false alarm rates. NIR cues were incorporated for improving the false alarm rates while keeping a high detection rate. Multi-modal sensing has lead to a correspondence problem, which was solved using a two step template matching algorithm. The same algorithm was tested on other types of leaves, which has resulted in accurate classifications showing easier adaptability to other plant species. In the future work, it is intended to implement the algorithm on real hardware for online testing.

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.
Reference


