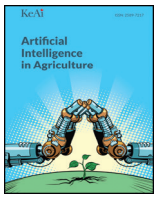
Artificial Intelligence in Agriculture 5 (2021) 72–81

Contents lists available at ScienceDirect

Artificial Intelligence in Agriculture

journal homepage: http://www.keaipublishing.com/en/journals/artificial-intelligence-in-agriculture/



An investigation into the potential of Gabor wavelet features for scene classification in wild blueberry fields

Gashaw Ayalew [a,*], Qamar Uz Zaman [b], Arnold W. Schumann [c], David C. Percival [d], Young Ki Chang [b]

[a] Independent Researcher, Thorold, ON L2V 1Z8, Canada
[b] Engineering Department, Dalhousie University Faculty of Agriculture, Truro, NS B2N 5E3, Canada
[c] Citrus Research and Education Centre, University of Florida, Lake Alfred, FL33850, USA
[d] Plant, Animal and Environmental Sciences Department, Dalhousie University Faculty of Agriculture, Truro, NS B2N 5E3, Canada

ARTICLE INFO

Article history:
Received 4 October 2020
Received in revised form 22 December 2020
Accepted 30 March 2021
Available online 1 April 2021

Keywords:
Wavelet transforms
Wild blueberry
Computer vision
Machine learning
Discriminant analysis
Precision agriculture

ABSTRACT

A Gabor wavelets based technique was investigated as a potential tool for scene classification (into one of bare patch, plant, or weed) for its ultimate utility in site-specific application of agrochemicals in wild blueberry fields. Images were gathered from five sites located in central Nova Scotia, Canada. Gabor wavelet features extracted from these images were used to classify scenes according to visually determined classes using step-wise linear discriminant analysis.

For individual fields, classification accuracy attained ranged between 87.9% and 98.3%; selected Gabor features ranged between 27 and 72; contextual accuracy for herbicide ranged between 67.5% and 96.7%, and contextual accuracy for fertilizer ranged between 63.6% and 97.1%. The pooled scenes yielded a classification accuracy of 81.4%, and contextual accuracy figures of 61.1% and 73.1% for herbicide and fertilizer, respectively, with selected Gabor features of 36.

Calibrations based on LDA coefficients from the pooled scenes could help avoid the need to re-calibrate for each field, whereas those based on individual field LDA coefficients could improve accuracy, hence enable saving on expensive agrochemicals.

© 2021 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Blueberries constitute an important crop, covering an estimated area of >76,000 ha in Canada, and distributed among >3500 farms (Statistics Canada, 2018). During the 2014 to 2018 production years, these farms produced an average of ca. 190,000 T annually (Statistics Canada, 2019). This amounted to an average of ca. CAD 253 million in farm gate price over the same duration (Statistics Canada, 2019). It is also indicated that wild blueberry makes up ca. 30% of the North American blueberry production(Yarborough, 2017) – proving the economic significance of wild blueberry production.

Due to the wild and perennial nature of wild blueberry plants, cultural practices of weed control are meagre to nonexistent. This necessitates reliance on herbicides for weed control. As a result, millions of dollars are spent by the industry on herbicides (Esau et al., 2016).

It is widely recognized that site-specific agrochemical application reduces the economic as well as environmental costs incurred in wild blueberry production (Chang et al., 2012; Tian and Reid, 1999). Various investigations on weed detection methods were made and varying degrees of success achieved using different technologies, in different cropping systems, including wild blueberry.

These technologies included digital colour image based co-occurrence matrix algorithm to classify wild blueberry plantation scenes into bare patches, weed or plant (Chang et al., 2012); ultrasonic ranging to differentiate between wild blueberry plants, weeds and bare patches (Swain et al., 2009); Gabor wavelets to differentiate between broad leaf and grassy weeds (Tang et al., 2003) and discrete wavelet transforms of digital images for the development of site-specific herbicide application (Tian and Reid, 1999).

Among these, those systems utilizing digital image classification are reported to demonstrate the ability to identify weeds and bare patches more effectively irrespective of plant or weed height (Chang et al., 2012; Tang et al., 2003; Tian and Reid, 1999). Particularly, in wild blueberry fields, image based spot application of fertilizer (Chang et al., 2017; Chang et al., 2014), herbicide (Esau et al., 2016; Esau et al., 2014a; Chang et al., 2012) and fungicide (Esau et al., 2014b) were investigated in more recent studies, and promising results obtained.

However, the features extracted in the digital image based techniques were largely dependent on colour variations, a parameter that is affected by lighting and plant conditions (Yazawa, 1977). These effects, coupled to variations in camera settings and potential similarities in colour between blueberry and weed plants, and the promising results

* Corresponding author.
E-mail addresses: gayalew@niagaracollege.ca, gayalew@addisscience.net (G. Ayalew).

https://doi.org/10.1016/j.aiia.2021.03.001
2589-7217/© 2021 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

from past texture related studies (Chang et al., 2012; Castleman, 1996), a texture based method was chosen for this study.

Nonetheless, the use of colour-based preprocessing enables to derive enhanced images that would serve as input to texture based techniques. For instance, it has been shown that chlorophyll estimation is possible from the ratio, $(Red - Blue)/(Red + Blue)$, of colour components of wheat and rye crop images under different weather conditions (Kawashima and Nakatani, 1998), nitrogen deficiency in barley (Pagola et al., 2009), and yield in wild blueberry (Swain et al., 2010). It is, therefore, held that colour based preprocessing will aid in accentuating differences between plants and weeds.

The Gabor wavelet transform enables the analysis of image scenes both in spatial and frequency domains (Daugman, 1985). Selection of this technique for the current study is inspired by previous successful investigations by Tang et al. (2003) and Tian and Reid (1999), in the identification of isolated broad-leaved and grassy weeds; Tian et al. (2000), on identification of weeds interspersed with corn plants; and Kurtulmus et al. (2014), on the detection of immature peach. It is to be noted that wavelet transform of images is an established technique of multi-resolution filtering to extract textural features. As such it can be thought of as an "orientation and scale tunable" line and edge detector (Manjunath and Ma, 1996).

A Gabor wavelet transform of an image is obtained by its convolution with a Gabor wavelet filter. Wavelet filters, as applied to images, are designed to selectively pass regions containing a certain size and orientation of a given shape structure. For instance, a wavelet transform, based on a mother wavelet designed to pass elliptic shapes, would primarily allow the passage of similarly shaped and sized objects with close orientation to the particular filter (Castleman, 1996; Daugman, 1988). In Gabor wavelet transforms such mother wavelet is generated by a Gabor function given (Manjunath and Ma, 1996) as:

$$g(x,y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jWx\right] \quad (1)$$

and its Fourier transform, given (Manjunath and Ma, 1996) by:

$$G(u,v) = \exp\left\{-\frac{1}{2}\left[\frac{(u-W)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right]\right\} \quad (2)$$

where $\sigma_u = 1/2\pi\sigma_x$, $\sigma_v = 1/2\pi\sigma_y$, and σ_x and σ_y are scale factors along the x and y dimensions, respectively. The self-similar Gabor wavelet filters are produced through scaling and rotation of the mother wavelet described in Eq. 1 according to the equations:

$$\begin{aligned} g_{mn}(x,y) &= a^{-m}g(x',y') \\ x' &= a^{-m}(x\cos\theta + y\sin\theta) \\ y' &= a^{-m}(-x\sin\theta + y\cos\theta) \end{aligned} \quad (3)$$

where $\theta = n\pi K$, $a > 1$, $m, n = 0,1,2, \ldots$, and K is the total number of orientations, and the scale factor a^{-m} is included to ensure independence of spectral energy on m – the scale for a particular filter (Manjunath and Ma, 1996). The Gaussian envelope that is responsible for the shape of the function is evident from Eq. 2. It is shown by (Manjunath and Ma, 1996; Daugman, 1988; Daugman, 1985) that redundancy of information in filtered images is reduced by ensuring that the half-peak amplitudes of filter responses (for the given set of scale and orientation parameters) do not overlap. Daugman (1985) also shows that two-dimensional Gabor filters are optimal in the sense that they closely approach the theoretically attainable limit of minimum joint uncertainty in space and frequency. Detailed treatment of the subject is available elsewhere (Manjunath and Ma, 1996; Daugman, 1993; Daugman, 1988).

Once images are convolved with the given bank of Gabor filters, the mean and standard deviations of the transformed images are computed as features of the image, for the given scale and orientation (Manjunath and Ma, 1996).

This study aimed at the assessment of the potential of Gabor wavelet features for automatic classification of digital color scenes from wild blueberry fields.

2. Materials and methods

2.1. Wild blueberry fields and imaging conditions

Digital color images were collected from five wild blueberry production fields located in central Nova Scotia, Canada. Table 1 provides data on the fields and the images collected.

Field 1 was in its fruit year, and was made up of predominantly dense plant patches – some plants being inflorescent. The lighting condition was a mid-day sunlight. Some of the blueberry leaves had hues ranging from yellow to pink. The predominant weed in this field was goldenrod (*Solidago spp.*). Dead grassy and woody plant materials were visually classified as bare patches.

Field 2 and Field 3 were in their vegetative year, and the lighting condition was cloudy in both cases. They generally consisted of sparsely populated wild blueberry plants interspersed with weeds and bare patches. Predominant weeds were lamb's-quarters (*Chenopodium album L.*), sheep sorrel (*Rumex acetosella L.*, both vegetative and inflorescence), goldenrod, poplar (*Populus spp.*) spreading dogbane (*Apocynum androsaemifolium*), mouse-eared hawkweed (*Hieracium pilosella L.*) and a few black bulrush (*Scirpus atrocinctus*).

Field 4 was in its fruit year, with scattered colony of dense wild blueberry plants. The lighting condition was a mid-afternoon sunlight. Predominant weeds were sheep sorrel, lamb's-quarters, goldenrod and wire grass (*Danthonia spicata (L.) Beauv.*). Bare patches comprised of bare soil and areas covered with dead plant remains.

Field 5 was in its fruit year, and the lighting condition was an early afternoon sunlight. Predominant weeds were goldenrod and sheep sorrel, the inflorescence of the latter being dominant. A few incidences of wire grass also existed. Some wild blueberry plant colonies had predominantly pink leaves, and others had yellowish colored leaves. The majority, however, had leaves with a mixture of these colors. In few other cases, the blueberry leaves at the shoot apex had hues ranging from yellow to pink. Some bare patches were covered with dead grass and woody weeds, others constituted completely bare soil. Anonymous (2019), LeBlanc and McCully (2005) and Sampson et al. (1990) were used as references for the identification of weed.

2.2. Acquisition and preprocessing of images

Wild blueberry canopy images were acquired using an IDS μEye 1220SE/C industrial camera (IDS Imaging Development Systems Inc., Woburn, MA 01801, USA), fitted with an LM4NCL C-mount lens (Kowa Optimed Inc., Torrance, CA 90502, USA), with 3.5 mm focal length and set at an aperture of $f/4$. The camera was kept approximately at 1.2 m above the ground surface. Resolution of the collected images was 720 × 128 pixel, and stored in the 24-bit colour bitmap format.

The 720 × 128 pixel images were first divided into two 360 × 128 pixel regions, and then each was further divided into three regions of

Table 1
Wild blueberry fields and scene characteristics.

Field	Stage of growth	Number of images	Priors		
			BARE	PLANT	WEED
Field 1	Fruit year*	108	11	67	30
Field 2	Vegetative year**	999	100	651	248
Field 3	Vegetative year	1110	117	608	385
Field 4	Fruit year	411	12	11	388
Field 5	Fruit year	653	65	387	201
Pooled	Mixed	3281	305	1724	1252

* The growth year immediately before harvest.
** The growth year following harvest and subsequent mowing of the stand.

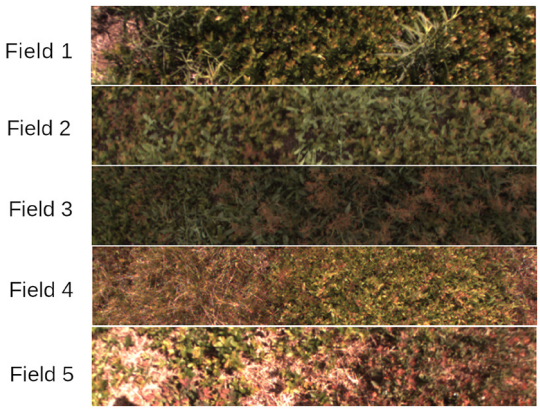


Fig. 1. Sample images from each field.

128 × 128 pixel. Overlapping patches of 12 × 128 pixel regions were allowed between consecutive 128 × 128 pixel regions to avoid the requirement for zero padding for use with the Fast Fourier Transform (Gonzalez and Wintz, 1987). Each of these 128 × 128 pixel regions corresponded approximately to a ground surface area of 27.5 cm × 27.5 cm. Fig. 1 shows representative images from the five fields.

Due to dependence on plant (Yazawa, 1977) and lighting conditions (Kawashima and Nakatani, 1998), it was decided to use differences and/or ratios of the primary color components, Red (R), Green (G) and Blue (B). 2D arrays were populated with numerical figures derived according to the relationships given in Eqs. 4a – 4j, and subsequently Gabor transformed to enable the computation of features used to carryout scene classification. Eqs. 4a to 4 are due to Kawashima and Nakatani (1998), whereas Eq. 4j represented the difference in the magnitude of the Green component on one hand and the Red and the Blue components on the other, and is applied in the differentiation of green plants and soil (Hamuda et al., 2016).

$$\text{Red ratio, } RR = R/(R+G+B) \tag{4a}$$

$$\text{Green ratio, } GR = G/(R+G+B) \tag{4b}$$

$$\text{Blue ratio, } BR = B/(R+G+B) \tag{4c}$$

$$\text{Red—Green normalized index, } RGI = (R-G)/(R+G) \tag{4d}$$

$$\text{Red—Blue normalized index, } RBI = (R-B)/(R+B) \tag{4e}$$

$$\text{Green—Blue normalized index, } GBI = (G-B)/(G+B) \tag{4f}$$

$$\text{Red—Green ratio, } RGR = (R-G)/(R+G+B) \tag{4g}$$

$$\text{Red—Blue ratio, } RBR = (R-B)/(R+G+B) \tag{4h}$$

$$\text{Green—Blue ratio, } GBR = (G-B)/(R+G+B) \tag{4i}$$

$$\text{Excess Green, } ExG = 2G-0.8R-1.2B \tag{4j}$$

2.3. Visual classification of scenes

Arguably, the most popular approach for classification of wild blueberry field scenes is along whether they are bare patch, all wild blueberry plants or weed (or weed interspersed with wild blueberry plants) (Chang et al., 2012; Swain et al., 2010; Zaman et al., 2009). To facilitate the visual classification of wild blueberry scenes along the above scheme quickly and consistently, a GUI-based tool was developed and used (Fig. 2).

The user would choose whether to start a new classification task or to resume an ongoing one. This would allow the user to open an appropriate file, containing a list of files to begin a new task, or a decision file to resume a previous task, respectively. The user would then choose the image with which to begin the classification. The six parts of the selected image would then be displayed in the respective windows. The

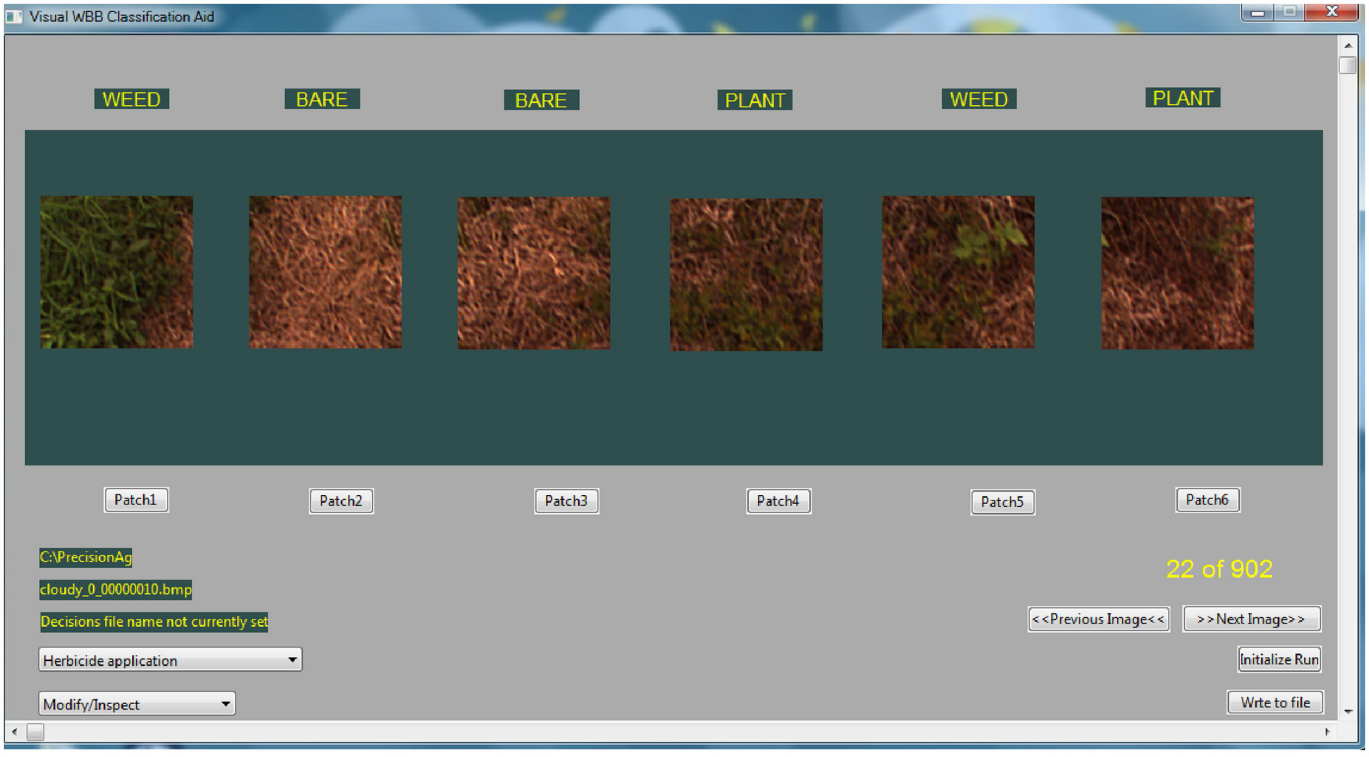


Fig. 2. Custom made visual classification tool.

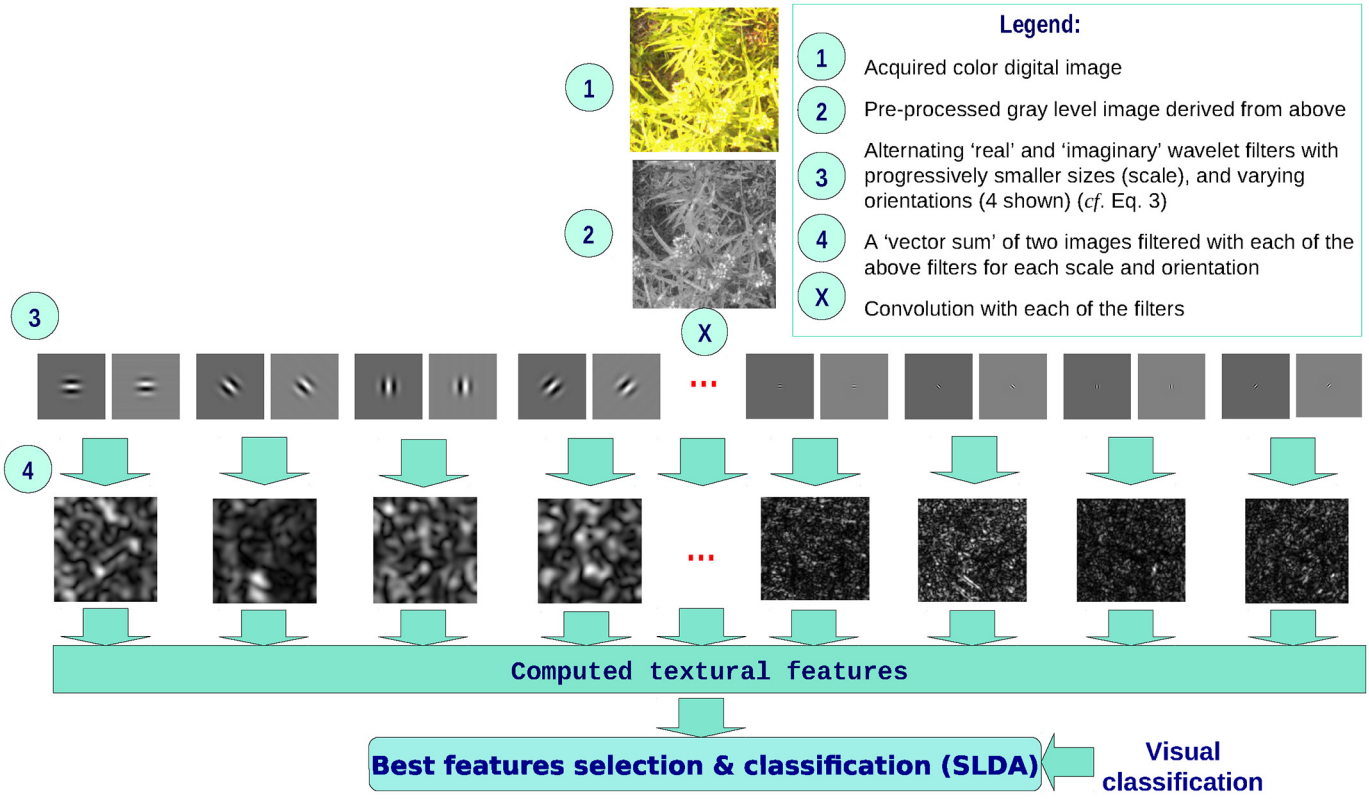


Fig. 3. Illustration of the Gabor wavelet based classification procedure.

user would then click on the button below each of the image segments to toggle the class of the concerned image segment. The classes are BARE, PLANT, WEED and DEFER, where the last class is applied when the image is considered unfit due to such effects as motion blur or partial shadow. The term "weed" refers to the usual definition in agronomy. Relative abundance of weeds was not considered in designating a scene as WEED – presence of any recognizable amount of weed sufficed.

The results of visual classification were then recorded to an ASCII decision file where each segment was identified with a unique name.

2.4. Extraction of Gabor wavelet features

Fig. 3 is a pictorial representation of the workflow for the feature extraction, and statistical classification. Each derived (preprocessed) image was filtered with a bank of Gabor wavelet filters computed with designated lower (U_l) and higher (U_h) frequencies selected to be 0.1 and 0.5, respectively (Manjunath and Ma, 1996). Four levels of orientation, and ten levels of scale were chosen. The mean (μ-features) and standard deviation (σ-features) of the magnitudes of each of the Gabor transformed derived image segments were eventually computed to represent features for the given set of parameters in a similar fashion as in Manjunath and Ma (1996), Daugman (1988), and Daugman (1985). Feature data were recorded in an ASCII file with each image segment given a unique name.

All computation required for feature extraction was carried out with custom programs coded in the **C** / **C**$^{++}$ programming languages. However, we would like to acknowledge the use of a modified version of Manjunath and Ma (1996)'s code to design the Gabor wavelet filters. One major modification was the creation of a filter dictionary at the global scope and use its portion that relates to a particular combination of scale and orientation as required – rather than re-creating filters for every image segment.

2.5. Statistical analysis

The unique names provided in the two ASCII files mentioned above (one from visual classification, the other from Gabor wavelet feature extraction) were used to match visual classes with computed features. Matched data from the two files, excluding those with a "DEFER" class, were then used for discriminant analysis.

A step-wise linear discriminant analysis (SLDA) was carried out on each of the groups of images collected from the fields. Wavelet feature selection was carried out using the stepclass() function of the klaR package (Weihs et al., 2005) for R. The criteria for selection was accuracy, with a threshold of incremental improvement of 0.1%. The direction for selecting features was set to "both" (i.e., adding features to, or removing features from the LDA model as feature selection proceeded), and cross-validation was based on a leave-one-out strategy. The same process was repeated for the pooled data set from all fields.

2.6. Overall and contextual accuracy of classification

Every individual scene that crossed over to another class according to the SLDA was automatically identified and counted. Overall accuracy, referred to as "accuracy" in short, was quoted as the percentage of correctly classified scenes out of the total number of scenes. In addition, to put the significance of accuracy of classification in the context of practical applications, two hypothetical agrochemical applications were considered: herbicide and fertilizer. The Tanimoto Similarity Criterion (Ayalew et al., 2004), as adapted to take the form shown in Eqs. 5 and 6, respectively, was used to measure the accuracy for a hypothetical site-specific application of herbicide and fertilizer. The criterion would assume values between 0.0% (complete failure), and 100% (complete success).

$$T_{\text{CHerbicide}} = \frac{\text{True}\quad \text{WEED}-\text{Missed}\quad \text{WEED}}{\text{True}\quad \text{WEED}+\text{False}\quad \text{WEED}} \times 100 \tag{5}$$

$$T_{\text{CFertilizer}} = \frac{\text{True}\quad \text{PLANT}-\text{Missed}\quad \text{PLANT}}{\text{True}\quad \text{PLANT}+\text{False}\quad \text{PLANT}} \times 100 \tag{6}$$

3. Results and discussion

3.1. Classification performance

Plots of linear discriminant analysis of visually classified field scenes, with respect to Gabor wavelet features, together with the 95% confidence limits for the clusters, are shown in Fig. 4.

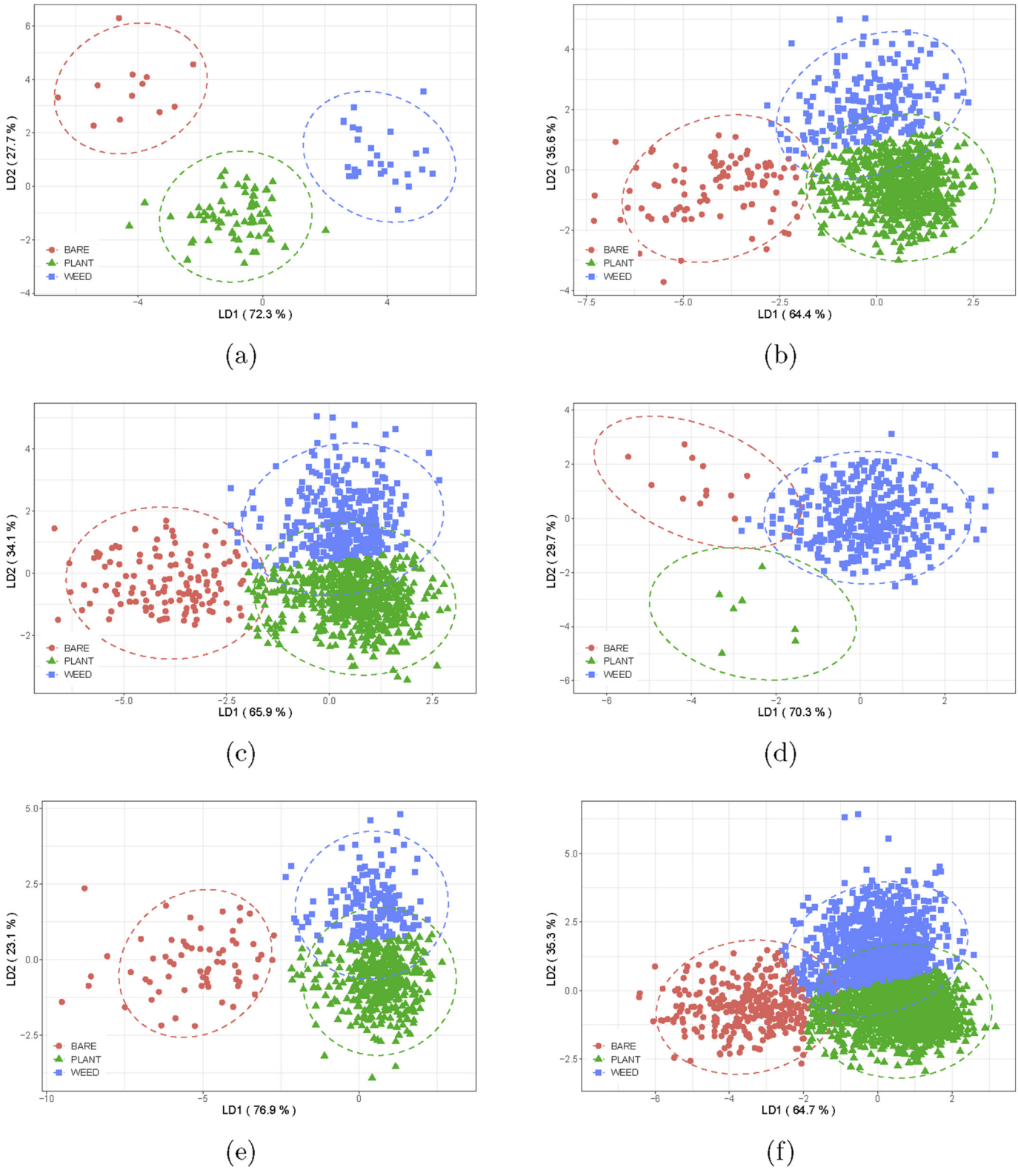


Fig. 4. Linear discriminant plots with 95% confidence limits. (a) Field 1; Field 2; (c) Field 3; (d) Field 4; (e) Field 5 and (f) Pooled field data analyzed as a single field.

Table 2
Classification performance.

Fields	Scenes	Features		Misclassifications						
			Acc.(%)	B→P	B→W	P→B	P→W	W→B	W→P	Total
Field 1	108	27	98.1	0	0	1	0	0	1	2
Field 2	999	53	89.8	7	5	5	24	3	58	102
Field 3	1110	53	87.9	4	0	7	34	9	80	134
Field 4	411	31	98.3	0	2	1	3	1	0	7
Field 5	653	72	88.5	2	1	1	21	2	48	75
Pooled *	3281	36	81.4	14	22	38	157	25	353	609

Keys: Acc.(%) = accuracy in percent; B = BARE; P = PLANT; W = WEED;
Misclassifications: X→Y ≡ visual scene X misclassified as scene Y according to SLDA;
*Pooled before SLDA and treated as a stand-alone field data.

Table 3
Classification performance in terms of the Tanimoto Similarity Criterion in the context of herbicide (T_{CH}) and fertilizer (T_{CF}).

Field	T_{CH} (%)	T_{CF} (%)
Field 1	96.7	97.1
Field 2	67.5	86.9
Field 3	70.6	81.9
Field 4	98.5	63.6
Field 5	67.7	83.5
Pooled	61.1	73.1

Table 2 shows the total number of scenes, the number of features selected, classification accuracy, inter-class misclassification, and total misclassification for the Gabor wavelet based method, for scenes collected from each field, and also for all scenes pooled together.

Number of Gabor features ranged between 27 and 72, among which 36 were selected for the pooled scenes. These 36 features, despite resulting in lower accuracy, would enable the design of a simpler system for a hypothetical sprayer or fertilizer spreader without the need to reconfigure the system for a particular field.

With regard to classification accuracy, the highest value was attained by Field 4, at 98.3%, and the least, by the set at 81.4%. Results achieved in this study were comparable with those attained with the colour co-occurrence matrix based texture analysis technique reported by Chang et al. (2012).

In addition, as it can be seen from Fig. 4 and Table 2, misclassification between plant and weed samples was the most pronounced. As a result, bare patches were more accurately classified in comparison to plant and weed scenes. This may be explained in terms of the fact that both plant and weed groups are made up of live plants that share similar hues and are generally broad leaved. On the contrary, bare patches were mostly brownish soil, or brown plant debris that may also occasionally include twigs without leaves. This is evident from Fig. 4 that greater separation existed between bare patches on one hand and weed and plant scenes on the other with respect to the first LDA coefficients (LDA1). It was with respect to the second LDA coefficients (LDA2) that separation existed between plants and weeds, albeit incompletely in some cases. As can be expected, the accuracy of SLDA classification for images was lower than that of any of the individual fields.

The contextual classification accuracy for herbicide and fertilizer applications are shown in Table 3. Due to an incidentally high weed priors, and low weed misclassification rate, Field 4 exhibited the highest contextual accuracy for herbicide application. On the contrary, this group exhibited the lowest contextual accuracy for fertilizer, owing to 4 out of 11 missed plant scenes.

With only one out of 67 plant priors misclassified as bare patch, and 1 out of 30 weed misclassified as plant, Field 1 exhibited the highest contextual accuracy with respect to fertilizer. Field 1 also exhibited one of the highest contextual accuracy for herbicide, with only 1 out of 11 weed scene misclassified as plant.

Field 5 also exhibited a large discrepancy between the two contextual accuracy figures. Low herbicide contextual accuracy resulted due to 48 out of 201 of weed misclassified to plant, 2 to bare patch, 1 from bare patch, and 21 from plant. On the other hand, contextual accuracy with respect to fertilizer was higher due to a proportionally lower misclassification of plant scenes in the form of 22 missed plants and 50 false plants. Field 2 and Field 3 performed moderately with regard to contextual accuracy.

As with accuracy of classification, the contextual accuracy figures are higher for classification based on individual scene SLDA, compared with those based on pooled scene SLDA. It can, therefore, be argued that a simple system can be designed with the LDA parameters for the pooled

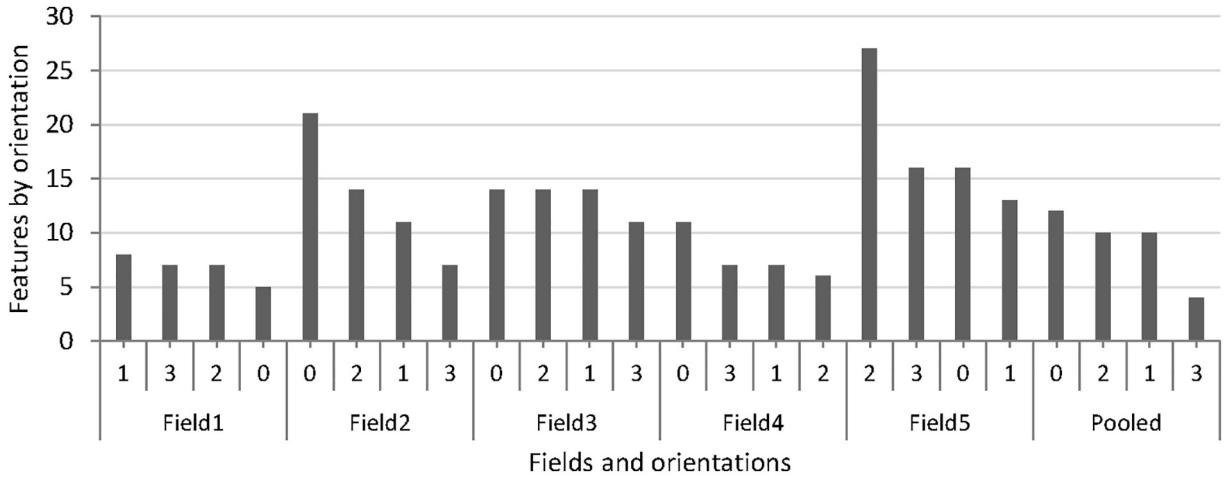


Fig. 5. Plot of count of features against orientation grouped by field, and sorted in descending order.

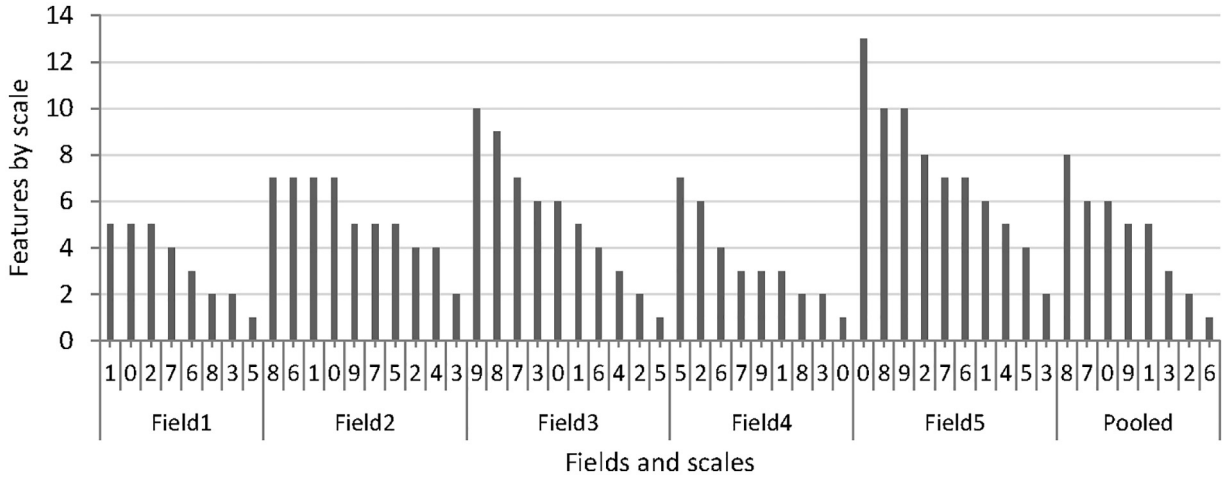


Fig. 6. Plot of count of features against scale, grouped by field, and sorted in descending order.

set for low-cost agrochemicals, and a field-specific calibration for more accurate operation to be used in the application of more expensive agrochemicals.

3.2. The effects of orientation and scale

Fig. 5 shows the number of features selected for each level of the orientation parameter, for each field, and sorted in descending order. Kendall correlation test was carried out to test if there is any order of preference, hence rank correlation between fields with regard to the numbers of features for each orientation parameter. Fig. 7(a) shows the correlogram of Kendall-τ values for fields with respect to rank of selected orientations.

Generally, the number of features grouped according to orientation shows a very weak, or nonexistent correlation between fields, as shown in the Figure. Only one of the correlations is found to be significant at the level of 0.05. This was the correlation between the set and Field 3. Even in this case, it could be argued that it is attributed to the bias caused by the highest membership of Field 3 in the set.

As the result of this weak correlation, and with the anecdotal knowledge of random orientation of leaves of wild blueberry and weed plants, it is concluded that the effect of orientation on the selection of features is of a statistically random nature.

Like the case for orientation, a rank correlation was carried out between the fields in terms of the order of number of features selected for each scale (Fig. 6), to test if there was any order of preference in the selection of scales. Fig. 7(b) depicts the associated Kendall-τ correlogram.

Comparing Figs. 7(a) and 7(b), it could be seen that there is more significant correlation between fields for scale than orientation. It is to be

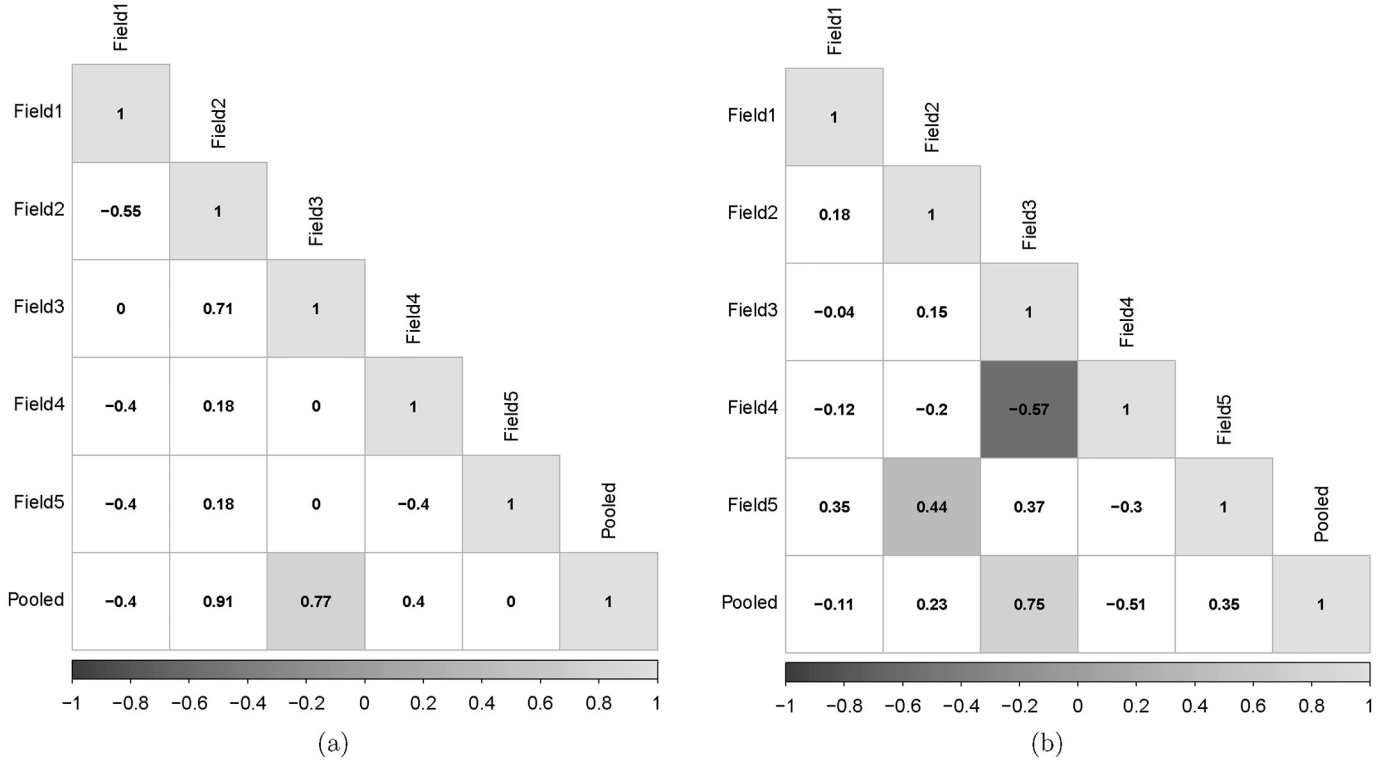


Fig. 7. Kendall-τ correlograms, with significant correlations at the level of 0.05 shaded. Cells not shaded indicate correlations that were not significant at the same significance level with respect to: (a) orientation; and (b) scale.

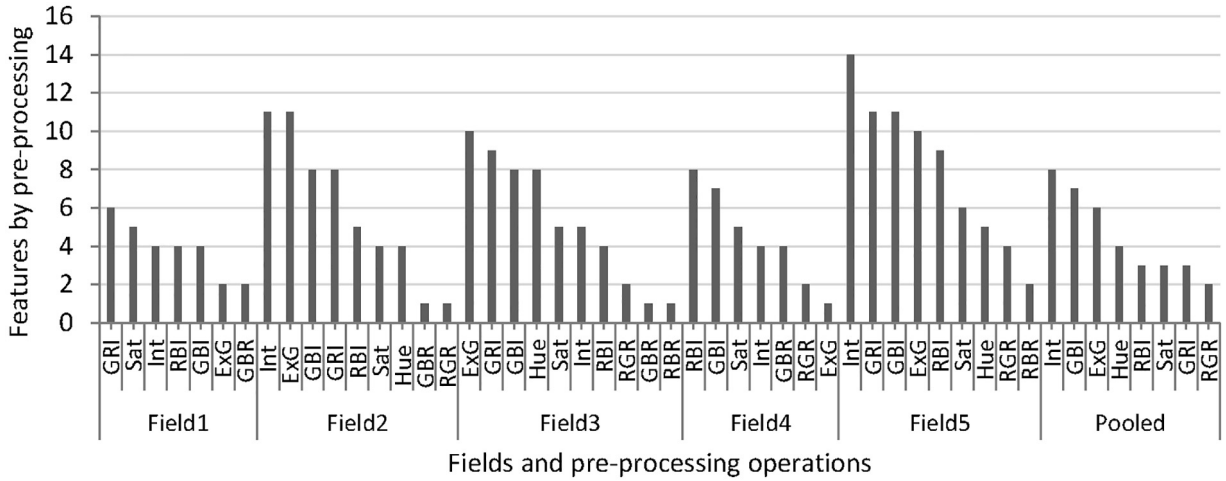


Fig. 8. Plot of count of features against preprocessing operations grouped by field, and sorted in descending order.

noted that scale is an important factor in the derivation of features, hence the fact that more correlations were significant is not surprising. However, weak correlations existed even in the significant values of Kendall-τ at the level of 0.05. The correlation between Field 3 and Field 5 and between Field 3 and Field 4 are low and incoherent. Therefore, it can be concluded that the selection of features was not determined or significantly affected by the scale parameter and depended very much on the individual scenes. The reason for this disparity may be explained in terms of uneven height of plants and weeds, whereby actual sizes of leaves appear disproportionately larger or smaller than other features in the scene – based on their location on the tall or short plant.

3.3. The effects of preprocessing operations

Fig. 8 shows the breakdown of features according to preprocessing operations of images that were used as input to the Gabor feature extraction routine. The count of features selected was plotted against image preprocessing procedure, grouped by field and sorted in descending order. As shown in the figure, the most frequently selected features were based on normalized difference indices(Eqs. 4d to 4f), ExG (Eq. 4j), Saturation, Intensity and Hue. There were also a few features based on difference ratios (Eqs. 4g to 4i).

It is also interesting to note that Field 4 and Field 1, being the fields that exhibited the top accuracy of classification, did not use Hue based features. Even in the fields for which Hue features were selected, it was not among the top three preprocessing operations. This may be attributed to the ambiguity that could result between plants and weed with respect to Hue. Another observation is that features based on simple ratios of R, G and B (Eqs. 4a – 4c) were not selected for classification indicating that indices based on color differences are more reliable than the basic colors alone.

Similar to orientation and scale, rank correlations were computed between fields to test if there was a preferred order of selection. Fig. 9 shows the correlogram for the Kendall-τ values that existed between

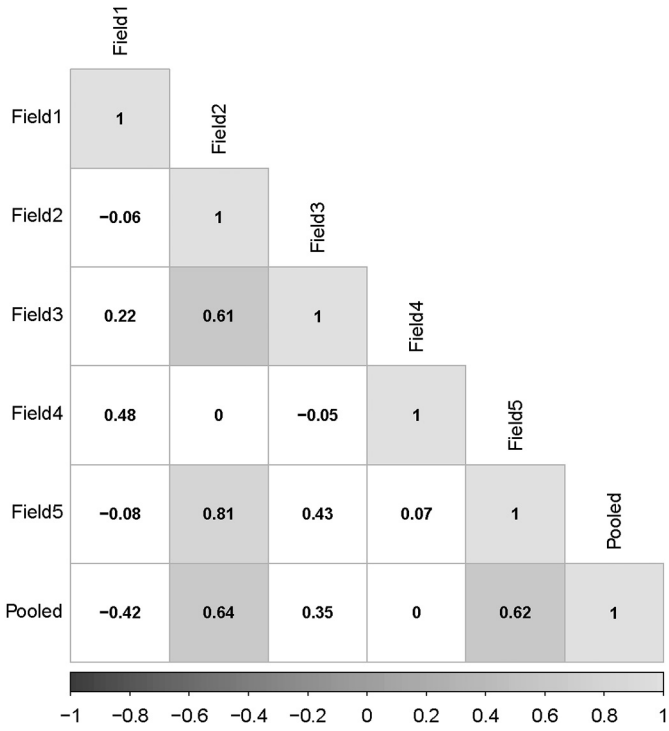


Fig. 9. Kendall-τ correlogram for count of features according to image preprocessing operations. Shaded cells in the plot signify correlations that were significant at the level of 0.05.

the fields, with respect to the ranking of feature counts according to preprocessing operations. With low correlation, and insignificance of most Kendall-τ values at the level of 0.05 (i.e., white background cells), it can be said that there is no significant correlation between fields with respect to selection of preprocessing operations, perhaps due to the diversity of colors of both plants and weeds.

3.4. General discussion

Given the diversity of hues and texture features existing in a given field depending on the stage of growth of the plant, the season, type of weed and its stage of growth, classification will be more accurate if it was based on the context of agrochemical, namely fertilizer, herbicide, or pesticide. Further improvement could be achieved if classification was based on the specific herbicide or pesticide. This would incidentally enable the decision to fall into one of two levels (i.e., "ON" or "OFF"), and hence result in a faster processing time. In addition, the use of both the first and the second linear discriminant functions simultaneously (as shown in Figs. 4) when applied to real-time control would improve accuracy, albeit with a higher computational cost.

Reduction of wavelet transform artifacts could be achieved with increased resolution of scene images. Although this could improve accuracy of classification, it reduces processing speed, for a given hardware. Therefore, with the selection of more powerful hardware, and a proper balance between the resolution and processing speed, it would be possible to apply this method for a real-time application of agrochemicals.

Consistency in classification of scenes, and hence accuracy in automatic operation of a sprayer system operating on the basis of the current method depends on the level of coherence of distances between plants and the camera. Greater uniformity between distances from the camera would enable greater accuracy in the selection of scales for the Gabor wavelet filters, which would in turn enable greater accuracy in classification. Perhaps a future study would be the development of an automatic correction for the effects of variation in plant height, whereby leaves are scaled according to their height from the ground, to offset the effect of height variations.

4. Conclusions

The potential of the Gabor wavelet based texture analysis to the identification of wild blueberry plants, weeds, and bare patches has been demonstrated. It is shown that classification accuracy levels in excess of 90% are possible for individual fields.

Results show that the scale parameter has a significant effect on the performance of the method, hence error or variations in scale of features, such as due to variations in plant height can cause deterioration of accuracy of classification. On the contrary, orientation of leaves and other image constituents has a random effect, suggesting that no need to differentiate features based on orientation as long as reasonably diverse levels of orientation (four levels in this study) have been used.

Features based on normalized difference indices, excess green, saturation, intensity, and hue were among the most frequently selected features, showing that preprocessing based on relative colors enhances the efficacy of the Gabor wavelet texture analysis.

The evaluation of a conceptual system for the application of fertilizer, herbicide or pesticide based on the results of the Gabor wavelet texture classification shows that it is feasible to build a reasonably precise "ON"/"OFF" control system. Further improvement is also possible if the training is tuned to the active ingredient suitable for a particular weed.

Further study is necessary for the assessment of the efficacy of this technique in real-time application of agrochemicals. An improvement in classification accuracy can be attained with the development of a mechanism to automatically control the distance between the plant canopy and the camera.

Credit Author Statement

Gashaw Ayalew: conceptualization, data curation, formal Analysis, investigation, methodology, software development, validation, visualization, writing, and editing.

Qamar Zaman: funding acquisition, project administration, supervision.

Arnold Schumann: funding acquisition, methodology, review, supervision.

David Percival: funding acquisition, methodology.

Young Ki Chang: data curation, resources, validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Oxford Frozen Foods Limited, Canada; Agriculture and Agri-food Canada under grant NS-Agri-Futures (ACAAF) under ACAAF-NS0153 (2007–2010) and Department of Agriculture Technology Development Program for jointly funding the project partially; and Dr. Travis Esau, Dr. Aitazaz Farooque, Mr. Kelsey Laking and Mr. Scott Read for their assistance during the collection of images. First author would also like to thank Professor Tessema Astatkie for his collegiality and his family for their support.

References

Anonymous, 2019. Integrated Pest Management Images. New Brunswick Department of Agriculture and Aquaculture URL:. https://daafmaapextweb.gnb.ca/010-002/Default.aspx?Culture=en-CA Accessed on 20 June 2019.

Ayalew, G., Holden, N.M., Grace, P.M., Ward, S.M., 2004. Detection of glass contamination in horticultural peat with dual-energy X-ray absorptiometry (DXA). Comput. Electron. Agric. 42, 1–17.

Castleman, K.R., 1996. Digital Image Processing. Prentice-Hall, Inc., Upper Saddle River, New Jersey.

Chang, Y.K., Zaman, Q., Schumann, A.W., Percival, D.C., Esau, T.J., Ayalew, G., 2012. Development of color co-occurrence matrix based machine vision algorithms for wild blueberry fields. Appl. Eng. Agric. 28, 315–323.

Chang, Y.K., Zaman, Q., Chattha, H., Reads, S., Schumann, A., 2014. Sensing system using digital cameras for spot application of fertilizer in wild blueberry fields. Paper written for presentation at the 2014 ASABE–CSBE/SCGAB Annual International Meeting. Meeting Paper Number 141913445.

Chang, Y.K., Zaman, Q.U., Farooque, A., Chattha, H., Read, S., Schumann, A., 2017. Sensing and control system for spot-application of granular fertilizer in wild blueberry field. Precis. Agric. 18, 210–223.

Daugman, J.G., 1985. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J. Opt. Soc. Am. A. 2, 1160–1169.

Daugman, J.G., 1988. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. IEEE Trans. Acoust. 36, 1169–1179.

Daugman, J., 1993. High confidence visual recognition of persons by a test of statistical independence. IEEE Trans. Pattern Anal. Mach. Intell. 15, 1148–1161.

Esau, T.J., Zaman, Q.U., Chang, Y.K., Groulx, D., Schumann, A.W., Farooque, A.A., 2014a. Prototype variable rate sprayer for spot-application of agrochemicals in wild blueberry. Appl. Eng. Agric. 30, 717–725.

Esau, T.J., Zaman, Q.U., Chang, Y.K., Schumann, A.W., Percival, D.C., Farooque, A.A., 2014b. Spot-application of fungicide for wild blueberry using an automated prototype variable rate sprayer. Precis. Agric. 15, 147–161.

Esau, T., Zaman, Q., Groulx, D., Corscadden, K., Chang, Y., Schumann, A., Havard, P., 2016. Economic analysis for smart sprayer application in wild blueberry fields. Precis. Agric. 17, 753–765.

Gonzalez, R.C., Wintz, P., 1987. Digital Image Processing. second ed. Addison-Wesley Publishing Company, Reading, Massachusetts.

Hamuda, E., Glavin, M., Jones, E., 2016. A survey of image processing techniques for plant extraction and segmentation in the field. Comput. Electron. Agric. 125, 184–199.

Kawashima, S., Nakatani, M., 1998. An algorithm for estimating chlorophyll content in leaves using a video camera. Ann. Bot. 81, 49–54.

Kurtulmus, F., Suk Lee, W., Vardar, A., 2014. Immature peach detection in colour images acquired in natural illumination conditions using statistical classifiers and neural network. Precis. Agric. 15, 57–79.

LeBlanc, L., McCully, K., 2005. Weed Identification Guide. URL. https://novascotia.ca/agri/documents/weed-identification-guide.pdf Accessed on 20 June 2019.
Manjunath, B.S., Ma, W., 1996. Texture features for browsing and retrieval of image data. IEEE Trans. Pattern Anal. Mach. Intell. 18, 837–842.
Pagola, M., Ortiz, R., Irigoyen, I., Bustince, H., Barrenechea, E., Aparicio-Tejo, P., Lamsfus, C., Lasa, B., 2009. New method to assess barley nitrogen nutrition status based on image colour analysis: comparison with SPAD-502. Comput. Electron. Agric. 65, 213–218.
Sampson, M.G., McCully, K., Sampson, D., 1990. Weeds of Eastern Canadian Blueberry Fields. Nova Scotia Agricultural College Bookstore. Truro, Nova Scotia.
Statistics Canada, 2018. Table 32-10-0417-01: Fruits, berries and nuts. URL. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3210041701 Accessed on 06 August 2018.
Statistics Canada, 2019. Table 32-10-0364-01: Estimates, production and farm gate value of fresh and processed fruits. URL. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=321003640 Accessed on 18 June 2019.
Swain, K., Zaman, Q., Schumann, A.W., Percival, D.C., 2009. Detecting weed and bare-spot in wild blueberry using ultrasonic sensor technology. Paper presented at the 2009 ASABE Annual meeting, Grand Sierra Resort and Casino, Reno, Nevada. ASABE Paper Number: 096879.
Swain, K.C., Zaman, Q.U., Schumann, A.W., Percival, D.C., Bochtis, D.D., 2010. Computer vision system for wild blueberry fruit yield mapping. Biosyst. Eng. 106, 389–394.
Tang, L., Tian, L., Steward, B.L., 2003. Classification of broadleaf and grass weeds using Gabor wavelets and an artificial neural network. Trans. ASAE 46, 1247–1254.
Tian, L., Reid, F., Hummel, J.W., 1999. Development of a precision sprayer for site-specific weed management. Trans. ASAE 42, 893–900.
Tian, L., Reid, J.F., Hummel, J.W., 2000. Development of a precision sprayer for site-specific weed management. Technical Report. University of Illinois. Agricultural Engineering Department, University of Illinois, 1304 W. Pennsylvania Ave., Urbana, IL 61801.
Weihs, C., Ligges, U., Luebke, K., Raabe, N., 2005. klaR analyzing German business cycles. In: Baier, D., Decker, R., Schmidt-Thieme, L. (Eds.), Data Analysis and Decision Support. Springer-Verlag, Berlin, pp. 335–343.
Yarborough, D., 2017. Blueberry crop trends 1996–2017. A presentation at the Wild Blueberry Producers Association of Nova Scotia, 17 November 2017.
Yazawa, F., 1977. Diagnosis of nutrition of crop plants by their leaf colors. Jpn. Agric. Res. Q. 11, 145–150.
Zaman, Q.U., Zhang, F., Schumann, A.W., Percival, D.C., 2009. Bare spots mapping in wild blueberry fields using digital photography. Paper written for presentation at the 2009 ASABE Annual International Meeting, Reno Nevada, June 21 – June 24, 2009. ASABE Paper Number 095582.