Adaptive Segmentation of Plant Images, an Integration of Color Space Features and Self-Organizing Maps

Mahmood Golzarian

1- Phenomics and Bioinformatics Research Centre (PBRC), School of Mathematics and Statistics
University of South Australia, Australia
2- Department of Mechanical and Agricultural Engineering,
Ferdowsi University of Mashhad, Iran.
mahmood.golzarian@unisa.edu.au

Abstract—We developed an adaptive learning for segmentation of plant images into plant and non-plant regions. In this study, we used Kohonen’s self organizing map (SOM) algorithm for segmentation of plant images using image series of two complexity levels; images taken in a controlled environment of plant facility and also images taken in the field. Nine color features of three color models of normalized Red Green Blue (RGB), Hue Saturation and Intensity (HSI) and L*a*b* made up the feature map. The results showed good performance for the images with less complexity. However, for images with higher complexity where there are more regions with similar color pattern, the method produces some noise.

Keywords—Image processing, Self Organizing Map (SOM), unsupervised learning, image segmentation, plant images, cereal plants

I. INTRODUCTION

Locating and segmenting plants from the background in an automated way is a common challenge in the analysis of plant images. Segmentation is an essential preliminary to an effective identification or other post-processing step. Any post-processing such as textural analysis for crudely segmented images will yield unreliable results due to background features mixed with those of plants [1, 2]. Direct techniques such as pixel based segmentation methods rely on a single value to represent the region of interest (ROI). These methods may not segment real world images with good results because intensity inhomogeneities, which occur quite often in real-world images, cause these methods to segment a ROI into several regions [3]. Also, most of these segmentation methods rely intensively on some prior knowledge about the background and lighting conditions against which the images of plants were taken. This information is not always easily available to be used for tailoring a segmentation method. Therefore, creating a robust algorithm that can be used for all situations is desirable. The objective of this study was to develop a robust image segmentation algorithm that separates plant region from background of the images regardless of background and lighting conditions. The method we developed is based on the concept of self learning unsupervised pattern recognition and Self Organizing Map (SOM) algorithm. For evaluation of the method, we tested the algorithm for two sets of plant images of two complexity levels.

II. PATTERN RECOGNITION

For most object recognition algorithms, we look for a particular pattern a ROI has in an image. This pattern is typically meant to make the ROI distinct in some way from the remaining area of background. A pattern is a set of features that quantify the perceptual information of ROI and the background. The features are in fact descriptors, which can be either quantitative such as area, length and texture [4-6], or structural descriptors such as Point symmetry groups and Frieze symmetry groups [7-11]. The pattern recognition algorithms can be therefore fall into two broad classes of decision theoretic and structural pattern recognition, respectively. An important step in a pattern recognition process is to determine the similarity between different regions or patterns. The similarity can be quantified by measuring distance between the feature vectors of two patterns in theoretic pattern matching applications. The best match occurs where there is smallest distance between the feature clusters [12, 13].

The objective of this study is to extract a set of quantitative descriptors and use in a SOM algorithm for clustering a plant image into partitions among which plant regions form a particular cluster.

III. PLANT IMAGE SEGMENTATION

Identification or extraction of any characteristics from the plants requires a good segmentation process. In this study, we developed a SOM segmentation method and tested this method on two sets of images of narrow leaf cereal plants. The first set was the images of wheat seedlings imaged every second day by cameras fixed in stations at the Australian Plant Phenomics Facility (APPF) [14]. The APPF provides state-of-
the-art plant growth environment and the latest technology in high throughput plant imaging for the repeated measurements of the phenotype of plants automatically and non-destructively. It can produce thousands of plant images by different cameras from different angles (two side views and a top view) in one day, so it is highly desirable to be able to process these images automatically. The side view images provide more information, including the height of the plant. In these views, narrow leaf wheat plants usually appear with some sharp leaf tips and deep axils on the images either at their early growth stage or when they are fully mature, as shown in Fig. 1.

The second set was the images taken in the field like situations. These images were taken from the same plot on different days and at different time of the day, and therefore contain several complex patterns, such as crop plants, soil, stones, dried crop residue and stalks ant weeds with large variation in illumination and therefore background color among images (Fig. 2).

The second set was the images taken in the field like situations. These images were taken from the same plot on different days and at different time of the day, and therefore contain several complex patterns, such as crop plants, soil, stones, dried crop residue and stalks ant weeds with large variation in illumination and therefore background color among images (Fig. 2).

Fig. 1. Two sets of images (side views and top views) of narrow leaf wheat seedlings taken at a controlled condition of greenhouse with homogenous background

Fig. 2. Sample images of narrow leaf wheat seedlings taken from the field situation on two different days. The plants are seen as the green regions. You will see the variation in illumination and background colour for two images taken. The blue pegs are used as markers for the purpose of locating the plants.

IV. UNSUPERVISED SELF ORGANIZING MAP

Self Organizing Map (SOM) is a special architecture of neural network with unsupervised learning mode. Learning is an essential mechanism in any neural networks for feature discovery and knowledge representation, or classification [15, 16]. Learning is the process in which the behavior of the network changes through adjusting the neurons’ weights and yields based from the desired output response to a specific input. There are three learning modes: supervised, unsupervised and batch learning. In unsupervised learning, we do not know the desired response of the network [17, 18]. In unsupervised learning, some input data, e.g. x, is given and the cost function to be minimized, that can be any function of the input data, x, and the network’s output, f(x). The cost function is dependent on the task we intend to model and the implicit properties of our model, its parameters and the observed variables. As an example, consider the model f(x) = k where k is a constant and the cost C = E [x - f(x)]^2. Minimizing this cost will give us a value of k that is equal to the mean of the data. Depending on the application, the cost function can be more complicated. To extract the knowledge from the input data, the unsupervised network responds and adjusts the learning parameters, such as network constants and weights, this is called self organizing [19]. A schematic diagram of a SOM learning architecture is shown in Fig 3. Regardless of the classification theme, the objective of classification is to reduce complexity and facilitate the interpretation of the real world by grouping similar elements together to provide a convenient abstraction from the original observations. Classification of data is critical to promoting clear and effective decision making [20].

V. FEATURE SELECTION AND FEATURE MAP

Human vision is able to recognize and classify the objects in a scene even if no priori knowledge about the objects available. This is done usually according the objects color or shape. This is the foundation for feature selection for SOM. Color can be quantifies in term of the values of components of a specific color space. In this study, nine components from three color models used to form the feature map in SOM. The color features are normalized red, green, blue, L*a*b* color components, and Hue, Saturation and Intensity.

The normalized red, green, blue are:

\[ r = \frac{R}{R+G+B} \]  \hspace{1cm} (1)

\[ g = \frac{G}{R+G+B} \]  \hspace{1cm} (2)

\[ b = \frac{B}{R+G+B} \]  \hspace{1cm} (3)

Fig. 3. Unsupervised neural network - Self-organizing map

The normalized red, green, blue are:

\[ r = \frac{R}{R+G+B} \]  \hspace{1cm} (1)

\[ g = \frac{G}{R+G+B} \]  \hspace{1cm} (2)

\[ b = \frac{B}{R+G+B} \]  \hspace{1cm} (3)
In L*a*b* system, "L" represents luminance and "a" and "b" represent the Red/Green" and "Yellow/Blue" color-opponent dimensions, respectively. Also, Hue, Saturation and Intensity. The hue value is obtained from the HSI color model which is represented by a cylindrical coordinate system as given in Fig. 5 [21]. Hue (H) implies the angle measured in the counterclockwise direction from the horizontal saturation axis where the pure red color is located. Green and blue colors are located at 120° and 240° apart from the red color, respectively. Saturation represents the degree to which a pure color is diluted to grey depending on the intensity as it moves from the out-most contour line. Note that when saturation is zero, the color varies in grey scale levels from black when intensity is the lowest value, to white when intensity is the highest value [22]. These nine features used to transform the high dimensional image space to low dimensional feature space.

The Euclidean distance is defined according to:

\[ D = \| X_i - X_j \| = \sqrt{(X_i - X_j)^T (X_i - X_j)} \]

where \( X_i \) and \( X_j \) are two input feature vectors and \((X_i - X_j)^T\) is the transpose of the difference vector.

The similarity can also be mathematically measured using the cosine of the angle between two vectors:

\[ \cos \theta = \frac{X_i \cdot X_j}{\| X_i \| \| X_j \|} \]

where in the nominator we have the inner product of the two feature vectors. In the denominator we have the multiplication of the magnitude of the vectors.

In this study, we used the Euclidean distance for measuring similarities between clusters.

VI. THE SELF-ORGANIZING PROCESS

A fixed size neighborhood of pixels in a color image pixels make up the neurons constructing the network grid of SOM. Each neuron is described by a 3-dimentional weight vector of the introduced color features (Fig. 5).

For the process of a SOM, the weights at each node are initialized. Every node is examined to find the Best Matching Unit (BMU). In each training step, one sample feature vector \( X \) from the input image is chosen randomly and a similarity measure is calculated between it and all the weight vectors of the map. For a given vector of weights, the Best Matching Unit (BMU) is the node in the Self-Organizing Map that is most similar (minimum distance) to the input vector (Fig. 6).

In this study, the following exponential decay function was used to shrink the neighborhood on each iteration until eventually the neighborhood is just the BMU itself:

\[ \sigma (t) = \sigma_0 \exp \left( -\frac{t}{k} \right) \]

where \( \sigma_0 \) is the width of lattice at the time zero, \( t \) is the current time step and \( k \) is constant and depending on \( \sigma_0 \) and the number of iterations.

The iterations are repeated a convergence occurs.

VII. EXPERIMENTAL RESULTS AND DISCUSSION

The results showed that for the first set of images which
were taken in a controlled glasshouse condition, the color features extracted and the SOM network can clearly segment the plant region from the background (Fig. 7).

Fig. 7. Results of SOM image segmentation for plant images taken in the glasshouse a) original image b)g) selected ROI and background clusters h) final output: combined ROI clusters to form the plant region

However, applying this method on the images taken from the field with many small and medium size regions of complex, non-linear and non-Gaussian distributed colors missed out some plant regions and misclassified some background regions as plant. The method leaves out some noise in the final output image. This is mainly due to illumination and reflection of several different plant and non-plant objects in the image (Fig. 8). If the noise clusters are isolated and smaller than plant clusters they can be removed easily in a post-processing step using a spatial filtering or morphological noise removal.

VIII. METHOD LIMITATION

While the procedures developed in this research appear promising for color image segmentation, we should consider the limitations of the method. First, the number of neurons selected for the SOM is a very important factor and has impact on the accuracy and performance of the method. Unfortunately, there are no set logic rules for determining the correct number of neurons to use, and finding the ideal balance between the number of neurons is a trial-and-error process. The number should be large enough to represent the image data effectively, sufficiently and properly [23]. Also, as a small neighborhood of a pixel is considered as a node in the network, the whole process is very slow depending on the size of the image.

IX. SUMMARY AND CONCLUSIONS

In this study, we developed an unsupervised SOM for the segmentation of the color images of narrow leaf plants. We used the color features of normalized r, g, b, L*a*b* and Hue, Saturation and Intensity to create the feature map. The images were of two distinct complexities: the first set of images was taken in plant phenomics facility with relatively homogenous background and the second set of images was taken from the field condition with larger variability in illumination and reflection and higher scene complexity. The features used for this study were color features.

The results showed that the performance of the method is unquestionable in the case of having a linear mixture of two
relatively homogenous regions and the algorithm is robust with respect to deviations from linearity. However, it fails for complex images with several small and medium size regions dispersed non-linearly over the image. In these cases, it is possible to obtain similar results using less-computational-demanding algorithms such as K-means or direct thresholding.

ACKNOWLEDGMENT

We would like to thank the Australian Plant Phenomics Centre for providing us with the plant images for this research.

REFERENCES


