Iranian Journal of Information Science and Technology

# COLOR SEGMENTATION SCHEME FOR CLASSIFYING WEEDS FROM SUGAR BEET USING MACHINE VISION

**A. JAFARI, Ph.D. student**
Agricultural Machinery Department
College of Agriculture
Tehran University, I. R. of Iran
email: ajafari@shirazu.ac.ir

**H. EGHBALI JAHROMI, Ph.D.**
Computer Engineering Department
College of Engineering
Shiraz University, I. R. of Iran
email: eghbali@cse.shirazu.ac.ir

**S.S. MOHTASEBI, Ph.D.**
Agricultural Machinery Department
College of Agriculture
Tehran University, I. R. of Iran
email: mohtaseb@ut.ac.ir

**M. OMID, Ph.D.**
Agricultural Engineering Department
College of Agriculture
Tehran University, I. R. of Iran
email: omid@ut.ac.ir

**Abstract -** In recent years, machine vision and optical sensors, which can be used in an autonomous weed killing equipment, are being used extensively to detect weeds from crops. In this study, seven types of weeds that grow in most of the sugar beet fields in Iran, especially in Fars province, were considered in real outdoor conditions. Several color feature extraction algorithms have also been investigated to separate soil from the plants as well as weeds from the sugar beets. The performance of the proposed algorithm was evaluated by determining correct classification rates (CCR) and misclassification rates (MCR) of the results. The findings revealed that the proposed method could successfully detect five of the seven types of the weeds, including Chenopudium album L., *Amaranthus retrofelexus* L., *Physalis alkekengi* L., *Convolvulus arvensis* L., *Setaria vertidis* L. Beauv and *Echinochloa crus-gali* (L.) Beauv.

**Keywords** - Segmentation, Color, Machine Vision, Feature Extraction, Sugar Beet, Weed Detection.

## INTRODUCTION

Weed suppression constitutes an important part of sugar beet production costs. Many approaches have been tried to reduce this cost. Detection and localization of weeds in the field, for special crop plants, is one of the most challenging tasks for automatic weeding. Machine vision is a smart method that can be used to classify weeds from sugar beet.

Early studies of machine vision systems were done by Parrish and Goksel [12]. In their work, they studied the use of machine vision for fruit harvesting. The main idea in this work is to segment weeds from sugar beet based on the color features. Weed infestation is one of the basic problems in crop production. Farmers are interested in reducing herbicide usage, while maintaining adequate weed control. Brown et al. [1] stated that, weed control could be maintained with a 25% reduction in herbicide use, if herbicides applied properly.

Their work involved applying herbicides only to the infested areas rather than the entire field. Selective herbicides selectively kill only weeds, not crop plants, thus they are so useful for weed control. But it is difficult for the operator to locate the weeds and control the sprayer at the same time. Therefore spot spraying with intermittent sprayers has been proposed by some researchers to reduce the herbicide consumption [9].

Weed detection is investigated in two ways: machine vision and optical sensors. Reflected light wavelengths from the weeds, main plant, soil and residues are different. This difference constitutes the main basis for weed detection by means of optical sensors [5,9,13,14].

Felton et al. [3] introduced a system that sprays herbicide only on areas detected by optical sensors which detect green object on the soil background. The problem with this work is that in sugar beet field both plants and weeds are green objects. Thus a robust color segmentation scheme is necessary.

Weed color is different from that of main plant and soil. Machine vision can be used to classify weeds from the soil and crop based on their colors. Color analysis, using reflectivity signature, has been studied by Brown *et al.* [1]. They used remote sensing techniques for possible weed detection. Shape features are used in many fruit sorting systems such as pistachio [6] and cereal grain classification [10]. Moreover, each plant has its special leaf shape and features that are almost unique for each plant. Some researchers have used these features to distinguish different plants [4,8,16]. Although these characteristics are different among different plants, they can't be used to segment a special plant in a scene that includes more than one plant, since both weeds and crops appear in a single frame of an image and such features cannot be extracted separately. This case is not a problem in color segmentation. Thus, it seems that working on color features is an alternative for segmentation of weeds from sugar beet.

Woebbecke *et al.* examined some color indices in order to distinguish weeds from the soil and residues [15]. In their work no special crop was considered to be segmented from the weed; therefore, it was a segmentation of plants from the soil and residues.

The objective of this study is to explore various color space and different composition of color components to find a suitable condition for constructing a region in which weeds can be distinguished from the sugar beets in real outdoor conditions.

## METHODOLOGY

Since the results of this study are expected to be applied in an automatic weed suppression device, natural field conditions have been considered. It includes different daylights, from sunny to cloudy sky, from morning to afternoon. Different lighting may appear by device shading on plants and ground.

A total of 300 digital images were acquired from different agricultural fields using digital camera (Fotoclip 2.1 Mega pixel) in order to collect whole common weeds in sugar beet fields around Shiraz, Fasa, Zarghan and Marvdasht. Weeds included Wild Spinach (Chenopudium album L.), Redroot Pigweed (Amaranthus retroflexus L.), Chinese Lantern Plant (Physalis alkekengi L.), Little Hogweed (Portulaca oleracea L.), Field Bindweed (Convolvulus arvensis L.), Green Foxtail (Setaria vertidis L. Beauv) and Barnyardgrass (Echinochloa crus-gali (L.) Beauv). Images were acquired with a resolution of 1600×1200 pixel, pertaining to a field of view about 70×60 cms on the ground at a distance about 1.2 meters from the soil surface, having 24-bit data field and JPEG format. Figure1 shows a schematic diagram of the procedure.

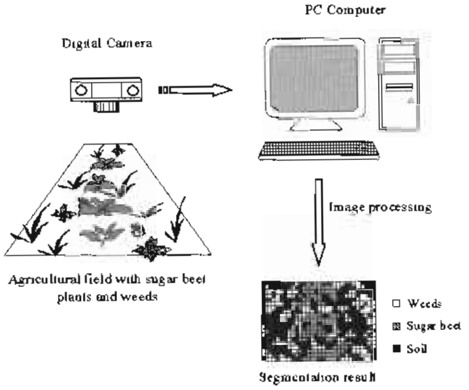

Figure 1: Weed segmentation by means of machine vision.

One of the most important problems in a vision system is to identify subimages that represent objects. The partitioning of an image into region is called segmentation. A region represents an object or part of an object. In this study, the object is weed. Generally segmentation can be defined by the following definitions:

1. A region is a subset of an image.
2. Segmentation is grouping pixels into regions, such that:

$$\bigcup_{i=1}^{k} P_i = \text{Entire image} \qquad P_i \text{ is a partition} \tag{1}$$

$$P_i \bigcap P_j = \phi \quad , \ i \neq j \tag{2}$$

Each region $P_i$ satisfies a predicate; it means that all parts of the partition have the same common property.

Pixels belonging to adjacent regions, when they are taken jointly, do not satisfy the predicate, i.e. a simple predicate may have uniform intensity.

Weed segmentation includes two stages. The first one is separation of whole plants from the background soil and residues and the second is separation of weeds from the main plants. Images were preprocessed with Matlab Image Processing Toolbox version 6.5 [11]. Image Tools (IT) version 3.0 [15] was used to test different positions of image histogram and choose the best point for threshold.

In as much as plants are green, it may be thought that we can segment plants from the soil, using threshold on green component of the RGB images.

A distinct histogram valley may exist in image histogram, but it doesn't completely conform to the separation region of plants and soil. Such a wrong result can be seen in Figure 2. This problem occurs as the change in lighting intensity will change all the three components: red (R), green (G) and blue (B). Thus, some parts of the soil that have received more lights may be classified as leaf due to their high value of G component, and some leaves in shadows may be segmented as soil because of their low value of G.

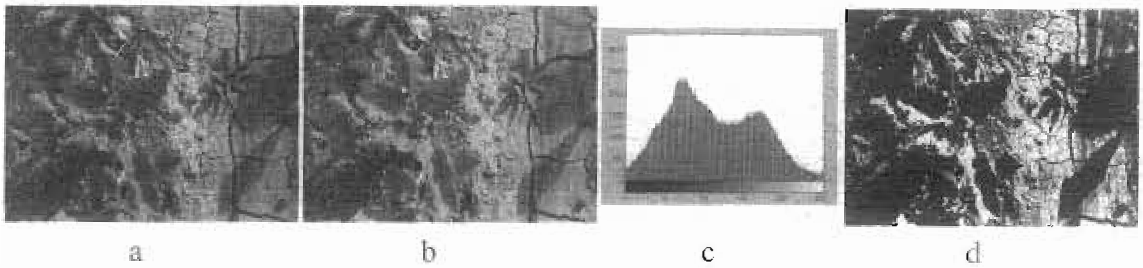

Figure 2: Segmentation based on RGB components:
a) Main image b) Green component, c) image histogram and d) Binary image after threshold.

Different lighting intensity is a problem in outdoor circumstances and may be appeared by changes in leaf angle with respect to sunlight, shadow of leaf on soil and devices shading on both soil and plants. On the other hand, sunlight varies during the day and growth season. It also causes a constant threshold value not to be able to be applied. Therefore, light intensity must be omitted in defining the index. Methods that can exclude light intensity from the image are examined as follows.

## - USING CHROMATICITY

Chromaticity just defines color components of the images without considering the light intensity. The components which define chromaticity coordinates are: [7]

$$r = \frac{R}{R+G+B} \tag{3}$$

$$g = \frac{R}{R+G+B} \tag{4}$$

$$b = \frac{R}{R+G+B} \tag{5}$$

Where "R", "G" and "B" are red, green and blue components "r", "g" and " b" are red, green and blue chromaticities respectively. Chromaticity coordinates are less sensitive to illumination. Now, plants can be segmented from the soil using green chromatic component "g" (Figure 3).

a b c

Figure 3: Using chromaticity components for segmentation:
a) original image, b) chromaticity" "g" and c) binary result.

This method, however, can separate plants from the soil, but as it can be seen from Figure (3), it produces a lot of noises. In as much as some dark regions that exist in the images have R, G and B values equal to zero, the denominator of equations (3), (4) and (5) becomes zero. To avoid dividing by zero, a value of 0.0001 was added to images matrices.

### - HSI COLOR SPACE

Another approach for omitting illumination intensity effect is using HSI (Hue, Saturation, and Intensity) space. Naturally, we can distinguish plants and soil based on their colors. RGB space are converted to HSI space, using the following equations: [7]

$$H = \cos^{-1}\left\{\frac{\tfrac{1}{2}[(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)^{1/2}]}\right\} \tag{6}$$

$$S = 1 \frac{3}{(R + G + B)} [\min(R,G,B)] \tag{7}$$

$$I = \frac{1}{3}(R + G + B) \tag{8}$$

Where $H$ is hue; $S$ is saturation and $I$ is intensity.

Defining hue and saturation space is closer to human visual system [6], so its possibility for segmentation was examined. Figure (4) shows the result of segmentation only based on Hue difference between soil and plants.

Figure 4: Hue difference: a) original image, b) hue image and c) saturation image.

Since plants are more colorish than soil, they may have significant differences between their saturation components, which produce almost the same result (Figure 5).

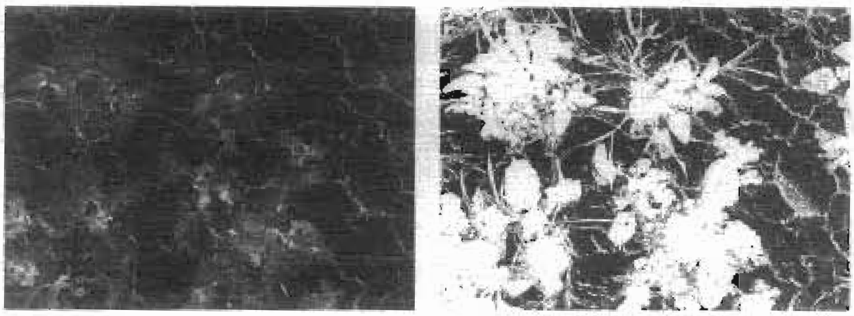

Figure 5: Saturation difference: a) Saturation and b) segmentation result.

**- GREENNESS**

In as much as plants are green, those pixels of image that relate to plants have greater G components than R and B. It means that:

$$G > \frac{(R + B)}{2} \quad \Rightarrow \quad 2G - R - B > 0 \tag{9}$$

An important preference of this formula is eliminating light intensity. If we suppose that white light constituted from equal values of red, green and blue components, thus when light intensity increases by a magnitude of "i" (Figure 6) then we have:

$$I_2 = I_1 + i \tag{10}$$

$$R_2 = R_1 + i\cos\alpha = R_1 + i\,(\frac{R_1}{I_1}) \tag{11}$$

$$R_2 = R_1\,(1 + \frac{i}{I_1}) \tag{12}$$

$$G_2 = G_1\,(1 + \frac{i}{I_1}) \tag{13}$$

$$B_2 = B_1\,(1 + \frac{i}{I_1}) \tag{14}$$

Where $I_1$ and $I_2$ are the light intensities at the first and second conditions; $\alpha$ is the angle between I axis and R coordinate. $R_1$, $G_1$ and $B_1$ are red, green and blue components at first condition of illumination and $R_2$, $G_2$ and $B_2$ are red, green and blue components at second condition of illumination.

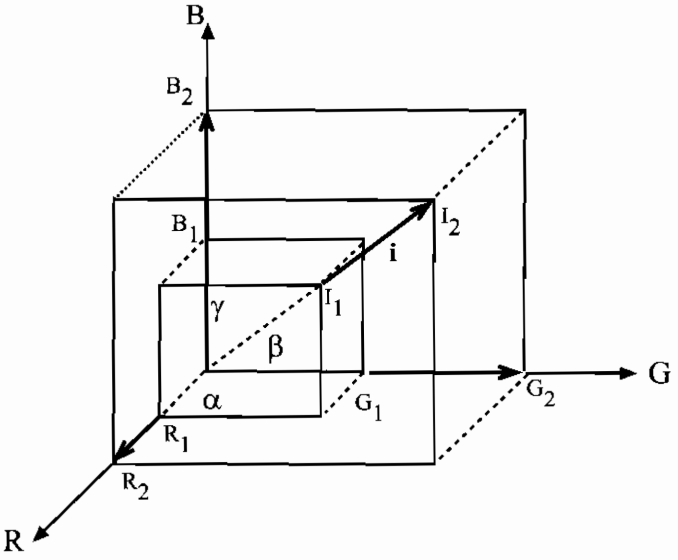

Figure 6: Light intensity changes in RGB space.

It means that when illumination changes, each one of the three components R, G and B multiplies by a factor k = (1+i/I1). So the equation (7) in new condition will become:

$$2G_2 - R_2 - B_2 = \left(1+\frac{i}{I_1}\right)(2G_1 - R_1 - B_2) \tag{15}$$

Since the point of thresholding in this relation is set to zero, the multiplication of factor k doesn't affect the segmentation result. It means that equation (9) is independent with respect to light and shadows.

It must be pointed out that these relations are related to illuminating light components and not the reflected light from the object. It is important how much of the light the surface absorbs and how much it reflects. The light reflectance is also different for different surface colors, as El-faki et al. [2] have shown, for standard red, green and blue color plates the changes in reflected light components are almost linear. Therefore, equations that have an equal degree of these components in their numerator and denominator will remain constant when light intensity changes. Thus, they will be suitable for omitting the effect of light and shadow in the images.

## WEED SEGMENTATION

After the separation of soil from plants, it is the time to distinguish weeds and sugar beets. The last formulas and equations couldn't separate weeds from sugar beet because these two classes have considerable overlapping based on those relations. However, in some special conditions, they may be successful but the algorithm should include all aforementioned circumstances.

Similar to the procedure used for soil, several parts of sugar beets and weeds from different images and conditions were collected. In order to eliminate different light intensity and shadows on leaves, chromaticity components of all images were determined. Pixels related to soil, sugar beet and weeds were manually segmented. Mean value and standard deviation of pixel values related to each one of these three classes were determined by the following equations:

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{16}$$

$$\sigma = \left[ \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \mu)^2 \right]^{1/2} \tag{17}$$

Where $\mu$ is the mean value; $x_i$ is the sampling value; $\sigma$ is the standard deviation; $n$ is the sum of the numbers of sampling pixels.

Investigating these components showed that sugar beet plants have greater "r" with respect to weeds. On the other hand, "b" component of weeds and soil is greater than that of sugar beet. There is not a considerable difference between "g" component of sugar beet and weeds because both of them are green plants. These relations are summarized in Table 1.

Table 1: Mean value and standard deviation of red, green and blue components for soil, sugar beet, and weeds.

| | Soil<br>$\mu \pm \sigma$ | Sugar beet<br>$\mu \pm \sigma$ | Weeds<br>$\mu \pm \sigma$ |
|---|---|---|---|
| **r** | 89± 15.8 ↑ | 83± 6.3 ↑ | 69±3.9 ↓ |
| **g** | 82± 10 | 121±14.5 | 93±4.7 |
| **b** | 77± 12.8 ↑ | 45±12.8 ↓ | 87±4.8 ↑ |

The arrows in Table 1 lead us to define a relation that increases the differences between sugar beet and weeds that is the ratio r/b. This ratio is equal to R/B because factor 1/(R+G+B) can be omitted from denominator and numerator. This ratio is also independent with respect to light intensity. Equations (12) and (13) yield:

$$\frac{R_2}{B_2} = \frac{R_1}{B_1} \Rightarrow \frac{R}{B} = Const. \tag{18}$$

It means that the ratio R/B is constant while light intensity changes. So the parts in shadow and light have almost the same magnitude of R/B. Using this technique, we were successful to separate weeds from sugar beet. Figure 6 shows schematic diagram of the proposed algorithm.

As it is seen in Figure 7, some parts of soil that had small values of "r" remained in the image Figure 7d due to low luminance in cracks and leaf shadows.

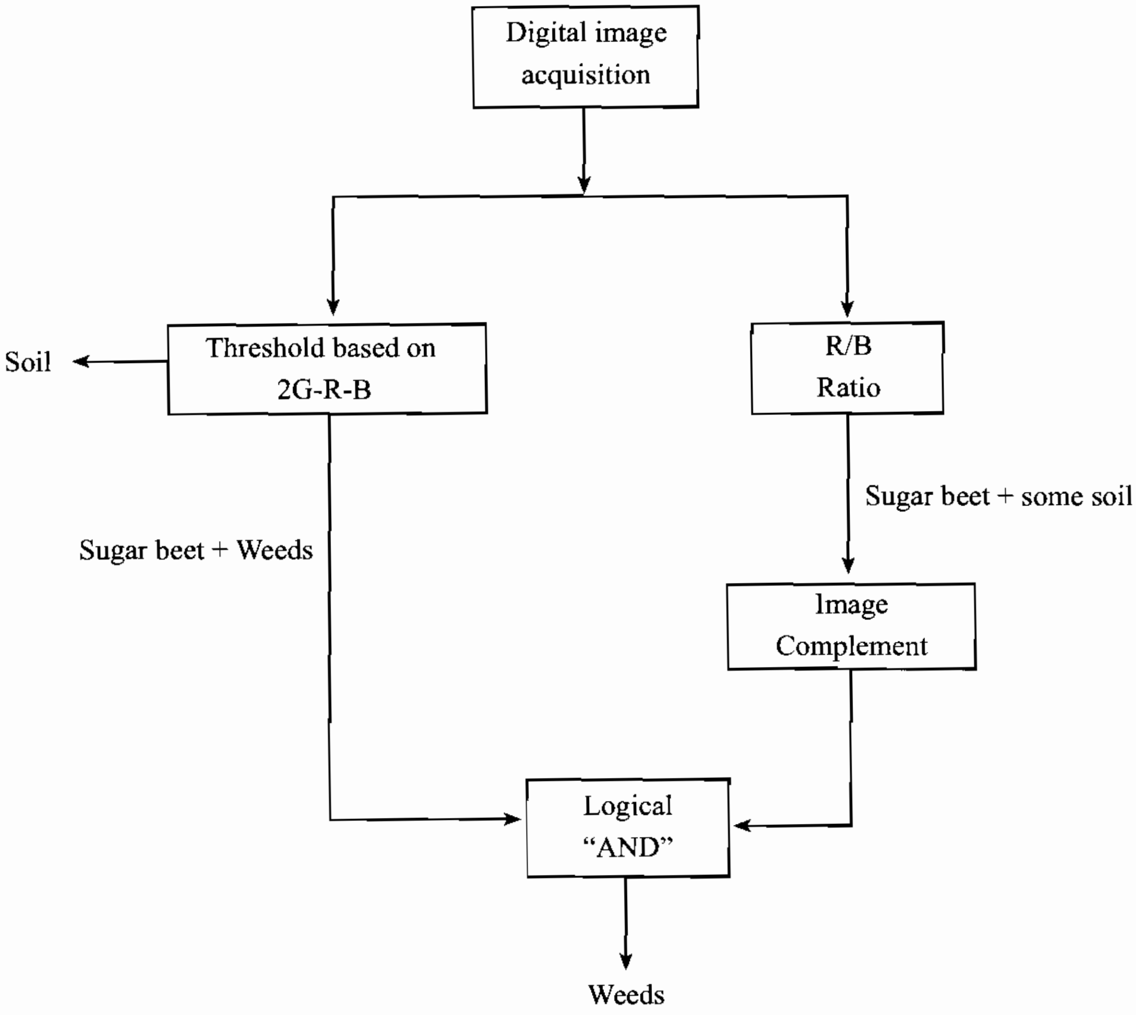

Figure 7: Block diagram of weed segmentation algorithm.

This problem was solved by constituting the complement of binary result of R/B and exerting "logical AND" between this image, Figure 8e, and the result of soil segmentation, Figure 8c. This action removed most of the remained soil and cracks. Final result is shown in Figure 8f.

Other compositions of RGB components, i.e. G/B, B/G, R/G and G/R were also examined. To avoid rounding problem that occurs during calculating such relations, it is better to consider these relations so that the smaller value will be on the denominator. On the other hand, in as much as G component of plants are greater than R and B, relations G/B and G/R were not successful in segmentation because of the same truncation problem that occurs when large values appear in the image matrices.

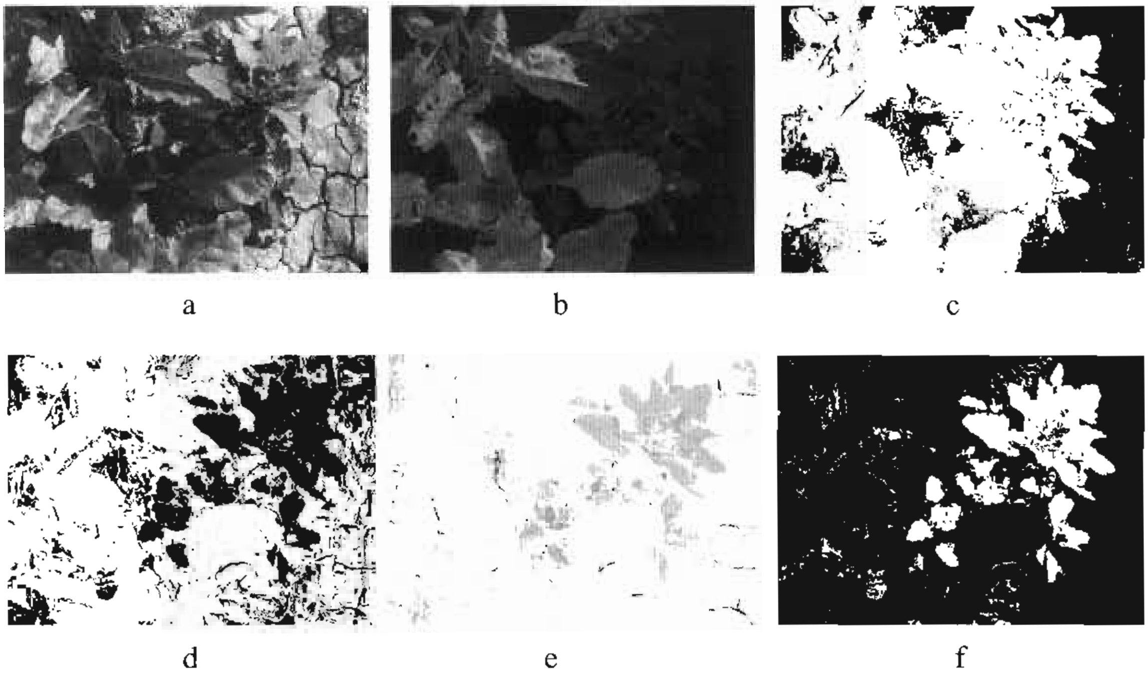

Figure 8: a) main image, b) 2G-R-B, c) binary result of soil segmentation, d) R/B ratio, e) complement of image d, f) result of weed segmentation.

It must be pointed out that chromaticity values can not completely omit light intensity; because, RGB composition in sunlight is different during the day, sunny or cloudy weather and shadow. On the other hand its R, G and B reflection from different color surfaces are not equal. It means that we compromised to accept that these three components change equally when they are subjected to different light intensities.

## RESULTS AND DISCUSSION

In order to evaluate algorithms, performance, correct classification rates (CCR) and misclassification rates (MCR) were determined. CCR is defined as the ratio between the number of weed pixels correctly classified and the total number of weed pixels in an image [2].

It is important for sugar beet plants not to be segmented as weeds. So MCR was defined as the ratio between those pixels of sugar beets that were classified as weeds and total sugar beet pixels in an image. Binary results of segmentation were compared to their initial images in order to determine CCR and MCR for 150 different images. Mean values of results are shown in Table 2. A value of about 14% was acquired for MCR.

Table 2: CCR values for different weed species.

| Weed variety | CCR % |
|---|---|
| Chenopudium Album L. | 81.5 |
| Amaranthus retroflexus L. | 76.1 |
| Physalis alkekenji L. | 81.3 |
| Portulaca oleracea L. | 41.7 |
| Convolvulus arvensis L. | 80.3 |
| Setaria virdis (L.) Beauv. | 78.2 |
| Echinochloa crus-gali (L.) Beauv | 79.9 |

Comparing the results of segmentation methods for separating soil from plants, we concluded that the best way to achieve this aim is the relation “2G-R-B” because it can eliminate light intensity effect better than other described methods.

In this experiment, however, saturation had almost good result in separating the soil but it was due to more deep color of plants with respect to that of soil. It means that we may have more difficulty to separate soils that are not grayish.

In order to validate an algorithm, both CCR and MCR are important. MCR shows what percentage of non-weeds incorrectly were classified as weeds. If we had defined the MCR as the ratio of incorrectly classified sugar beets as weed, then this magnitude would become lesser than those given in Table 2.

Results in Table (2) show that this method can be used for segmentation of seven weed species but it is weak for recognition of Portulaca, because its color is very close to sugar beet. Further research is to be done to find the features that can separate this weed species from sugar beet. It is probable that shape features may be helpful.

As young leaves had more blue color, CCR of young Chenopodium weeds was less than the old ones.

## REFERENCES

[1] Brown, R. B., Steckler, J. P. and Anderson, G. W., “Remote Sensing for Identification of Weeds in No-Till Corn.” *ASAE Paper*, St. Joseph, Mich.: ASAE, pp. 91-105, 1991.

[2] El-Faki, M. S., Zhang, N. and Peterson, D. E., “Factors Affecting Color-Based Weed Detection.” *Transaction of the ASAE* , Vol. 43, No. 4, pp. 1001-1009, 2000.

[3] Felton, W. L., Doss, A. F., Nash, P. G. and McCloy, K. R., “A Microprocessor Controlled Technology to Selectively Spot Spray Weeds.” *In Proc. Automated Agriculture for the 21st Century*, St. Joseph, Mich.: ASAE, pp. 427-432, 1991.

[4] Franz, E., Gebhardt, M. R. and Unklesbay. K. B., “Shape Description of Completely

Visible and Partially Occluded Leaves for Identifying Plants in Digital Images." *Transaction of the ASAE*, Vol. 34, N0. 2, pp. 673-681, 1991.

[5] Franz, E., Gebhardt, M. R. and Unklesbay, K. B., "The Use of Local Spectral Properties of Leaves as an Aid for Identifying Weed Seedlings in Digital Images." *Transaction of the ASAE,* Vol. 34, No. 2, pp. 682-687, 1991.

[6] Ghazanfari, A., Irudayaraj, J., Kusalik, A. and Romaniuk, M., "Machine Vision Grading of Pistachio Nuts Using Fourier Descriptors." *Journal of Agricultural Engineering Research*, Vol. 68, pp. 247-252, 1997.

[7] Gonzalez, R. C. and Woods, R. E., *Digital Image Processing.* Reading, Mass.: Addison-Wesley Publishing Co, 1992.

[8] Guyer, D. E., Miles, G. E. Gaultney, L. D. and Schreiber, M. M., "Application of Machine Vision to Shape Analysis in Leaf and Plant Identification." *Transaction of the ASAE*, Vol. 36, No. 1, pp. 163-171, 1993.

[9] Haggar, R. J., Stent, C. J. and Isaac, S. A., "Prototype Hand-Held Patch Sprayer for Killing Weeds Activated by Spectral Differences in Crop/Weed Canopies." *Journal of Agricultural Engineering Research*, Vol. 28, pp. 349-358, 1983.

[10] Majumdar, S. and Jayas, D. S., "Classification of Cereal Grains Using Machine Vision: I. Morphology Models." *Transaction of the ASAE*, Vol. 43, No. 6, pp. 1669-1675, 2000.

[11] Mathworks. *MATLAB Image Processing Toolbox.* User's Guide Version 3.00. Natick, MA: The Mathworks, Inc., 2002.

[12] Parrish, Jr., E. A. and Goksel. A. K., "Pictorial Pattern Recognition Applied to Harvesting." *Transactions of the ASAE*, Vol. 20, No. 5, pp. 822-827, 1997.

[13] Shearer, S. A. and Jones, P. T., "Selective Application of Post Emergence Herbicides Using Photoelectric." *Transaction of the ASAE*, Vol. 34, No. 4, pp. 1661-1666, 1991.

[14] Shropshire, G. J. and Glas, C., "Spectral Band Selection for Color Machine Vision Used for Plant Identification." *In Proc. SPIE 1836, Optics in Agriculture and Forestry.* Eds. J. A. Deshazer and G. E. Meyer, Bellingham, Wash.: SPIE, pp. 220-230, 1992.

[15] "Image Tool for Windows Version 3.00." UTHSCSA, The University of Texas Health Science Center in San Antonio, 2002.

[16] Woebbeck, D. M., Meyer, G. E., Von Bargen, K. and Mortensen, D.A., "Color Indices for Weed Identification under Various Soil, Residue and Lighting Conditions." *Transaction of the ASAE*, Vol. 38, No. 1, pp. 259-269, 1995.

[17] Woebbeck, D. M., Meyer, G. E. , Von Bargen, K. and Mortensen, D. A., "Shape Features for Identifying Young Weeds Using Image Analysis." *Transaction of the ASAE*, Vol. 38, No. 1, pp. 271-281, 1995.