Weed and Crop Classification by Automated Digital Image Processing

A Master Thesis by
Mattias Andersson, 731127-6633, CTH

February 15, 1998

Examiner
Mats Viberg, Department of Applied Electronics

Supervisors
Mats Rudemo, Department of Mathematical Statistics
and
Tomas Gustavsson, Department of Applied Electronics

Chalmers University of Technology

ABSTRACT

Classification of carrot and seven weed species in digital color images are performed. The images are segmented to separate plant from soil. Three methods for finding thresholds for segmentation are investigated. A method to remove unwanted noise from segmented images is developed. Features of both plant shape and plant color are calculated. Discriminant analysis with the Bayes-classifier is used to perform classification. The images were taken in a carrot field and the image scenes were protected against direct sunshine. Of every species 27 images were taken yielding a total of 216 images. The best feature combination found consisted of seven features including both shape and color features. With this combination a classification rate of nearly 85% was accomplished using cross-validation. Two of the weed species were specially included in the study because they are by laymen often mistaken as carrots. The automatic method developed here showed to have problem with these weeds too, but also a third weed, not at all similar to carrot, was mistaken for being carrot. Classification results based on separate feature groups showed that the color features and the moment invariants among the shape features were good classifiers. With four color features a classification rate of 70.83% was achieved and 62.96% with two moment invariants.

SAMMANFATTNING

Klassificering av morot och sju ogräsarter i digitala färgbilder genomförs. Bilderna segmenteras för att separera planta från jord. Tre metoder för att finna tröskelvärde för segmentering utreds. En metod för att ta bort oönskat skräp i segmenterade bilder tas fram. Egenskaper, ofta kallade features, på både plantors form och färg beräknas. Diskriminantanalys med Bayes-klassificerare används för att utföra klassificeringen. Bilderna togs i ett morotsfält och bildscenen skyddades från direkt solljus. För varje art togs 27 bilder vilket medför totalt 216 bilder. Den bästa kombinationen av features bestod av sju features innefattande både form- och färgegenskaper. Med denna kombination erhölls en klassificeringsgrad på nästan 85% då kryssvalidering användes. Två av ogräsarterna var medtagna speciellt, därför att de ofta mistas för att vara morötter av lekmän. Den automatiska metod som utvecklats här visades sig också ha problem med dessa två ogräs, men den misstog även en tredje ogräsart, som inte alls liknar morot, för att vara just morot. Resultat från klassificeringar baserade på separata feature-grupper visade att färgegenskaper och invarianta moment bland formegenskaperna var goda klassificerare. Med fyra färgegenskaper erhölls en klassificeringsgrad på 70,83% och 62,92% med två invarianta moment.

PREFACE

This rapport describes a master thesis project performed at the Department of Applied Electronics at Chalmers University of Technology in co-operation with the Department of Mathematical Statistics at Chalmers University of Technology. This project is an expression for my interest in agriculture and digital image processing. I would like to thank my supervisors Mats Rudemo at Mathematical Statistics and Tomas Gustavsson at Applied Electronics.

CONTENTS

1 INTRODUCTION

The ability to locate and identify crop and weed automatically in digital images could lead to many useful inventions. In conventional agriculture there is a desire to be able to control the amount of herbicide applied when spraying to kill weeds. There are at lest two reasons why this is desired, first the increasing use of herbicides in agriculture is believed to be a big contribution to the pollution of the environment and secondly herbicides are a major cost for farmers. In 1991, the world market for herbicides was estimated to $26 800 000 000 (Paice *et al*., 1995). Today, when spraying, the same dose of herbicide is applied through out the entire field while, in fact, weeds are frequently distributed non-uniformly within a field. The idea is to apply a normal dose of herbicide where there grow a lot of weeds and apply a smaller dose where there are little or no weeds and by this reduce the amount of herbicide used. To do this a video camera would be mounted on the tractor or sprayer and a computer would continuously estimate the weed density and from that estimation control the sprayer.

In organic farming the situation is different. In organic farming herbicides are not allowed, instead weeds are removed mechanically and to a minor extent thermally. For instance when growing vegetables such as carrots, onions and cabbage the weeds in between the rows can be removed by special harrows, but there is always a band about 15 cm wide around the row that can not be weeded by machinery. The weeds in this band have to be removed by humans. Weeding is a very hard work. At Skrekarhyttans Gård in Sweden, where the images for this master thesis were taken, about 100 km of carrot row length has to be weeded by humans every year. The work pays per meter and the price has during the time period 1993 - 1997 varied between 0,30 SKr to 0,90 SKr depending on how difficult the weeding is. It is not easy to find all people needed to do all the weeding, therefor there is a desire to mechanize this work. The market for any kind of machinery in the organic farming business is today very small, in Sweden which is considered to be at the forefront of organic farming only about 3% of all vegetables consumed are grown organically. But this figure is increasing as people become more aware of the potential side effects of conventional agriculture. What is desired is a machine that by the use of video cameras, image processing and robotics could remove the weeds. This might sound rather futuristic, but could become reality if the technology is found and it is cheap enough.

A system, either to control a sprayer or to do weeding, would have to analyze images where crop and weeds are mixed together as they are in fields. In the sprayer application the positions of crops and weeds do not have to be determined, only the weed density, but in applications for organic farming the exact positions of weeds and crop would be crucial. In this work it was decided that a carrot field was to be analyzed. Before the work begun it was considered that the real case of having images where carrot plants be mixed with weeds, partly covering and crossing each other, was too difficult to be solved within the time limit of a master thesis work. Instead, it was supposed that images of that type could be segmented into individual plants, then the task would be to find those features that would distinguish carrots and the weeds from each other. This is the main aim of this master thesis.

Previous works: Images taken in barley fields during the early stages of growth were examined by Andreasen *et al*. (1997) to estimated the weed density. The automatic crop area estimator was satisfactory, but the automatic weed number estimator was not, so the resulting weed density estimate was unsatisfactory. The use of only to features were insufficient and problems with weeds partly covered by crop were considered to have led to the poor result. Guyer et al. (1986) were able to distinguish between eight species with an error probability of 8.5 - 9.8 % using eight features. By using up to nine features of plant shape Petry & Kühlbauch (1989) could distinguish six weed species in different growth stages with an average classification rate of 82.3%. Gerhards *et al*. (1993) used Fourier descriptors and shape parameters to distinguish between 10 weed species with an average classification rate of 81.9%.

2 IMAGE ACQUISITION

The images to be used for this project were taken in one of the carrot fields at Skrekarhyttans Gård in Sweden, see Figure 1. At Skrekarhyttans Gård vegetables are grown organically using no pesticides, herbicides or fertilizer. The particular carrot field was sown on May 27 and 28. Approximately one week later when the carrot plants were just breaking through the surface the field was flamed to burn weeds. After another three weeks the images were taken between June 23 and 25.

An area was selected in one of the two carrot fields. This area was selected because it contained many different weed species and the weeds were sufficiently uniformly distributed. The other parts of the carrot fields had a more one-sided weed population. In the area several weed species were more or less common, of these seven were chosen to be in this study. Table 1 lists the Latin, English and Swedish names of carrot and the weeds. In Table 1 the ”English name” is not necessarily the name used in England and for some of the species several other names exists. The names listed are though the ones that will be used through out this paper. Figure 3 and Figure 4 show examples of the eight species.

Table 1. Selected species

Latin name	English name	Swedish name
Daucus carota L.	*Carrot*	*Morot*
Chenopodium album L.	Lamb’s quarters	Svinmålla
Cirsium arvense (L.) Scop.	Canada thistle	Åkertistel
Fumaria officinalis L.	Fumitory	Jordrök
Matricaria inodora L.	Mayweed	Baldersbrå
Polygonum persicaria L.	Ladythumb Smartweed	Åkerpilört
Spergula arvensis L.	Corn Spurry	Åkerspergel
Stellaria media (L.) Vill.	Chickweed	Våtarv

The selected area was 162 rows wide and row length varied linearly from row 1 being 50 meters up to row 75 being 136 meters and then length decreased to 64 meters at row 162, refer to Figure 1. To locate spots for image acquisition every sixth rows were chosen yielding a total of 27 rows. For every row a random number of meters were measured and from that point the row was searched backwards to find carrot and weeds. A calculator was used to generate random numbers between 0 and 1, that number was multiplied with the current row length to get the position. In a few cases the end of the row was struck before all weed species had been found, when this happened the row was searched in the opposite direction starting from the random position. When a plant was found the surrounding plants were removed with in a 10 cm radii.

Carrot and five of the weed species were easily found, in general less then two meters row length had to be searched to find them. Mayweed and fumitory were however more difficult to find. There were though a strong argument why these two weeds still should be in the study and that was that they are by amateur human weeders often mistaken as carrots.

The images were taken with a Canon EOS500N still camera equipped with a 35 - 80 mm zoom lens. The camera was mounted on a tri-pod pointing directly towards the ground with the back of the camera 45 cm above ground level. The lens was always zoomed to 80 mm and the automatic program for close-up pictures was used. With these settings a picture would cover a 15 cm by 10 cm ground area. The image scene was protected against direct sunshine with a white umbrella. A few times when the sky was covered with dark clouds the camera automatically engaged the flash. When this happened no picture was taken. The films were developed and transferred to Kodak Photo CD by Kodak. Since MATLAB 5 was used in this work the images had to be converted from Kodak Photo CD-format to JPG-format, this was done with Adobe Photoshop.

On Kodak Photo CD all images come in five different resolutions; 3072*2048, 1536*1024, 768*512, 384*256 and 192*128. The images are so called ‘true color’ with 8 bits per color component. In the computer analysis part of this work the 768*512 resolution was used to reduce calculation time. With this resolution one pixel width corresponded to about 0,195 mm at ground level.

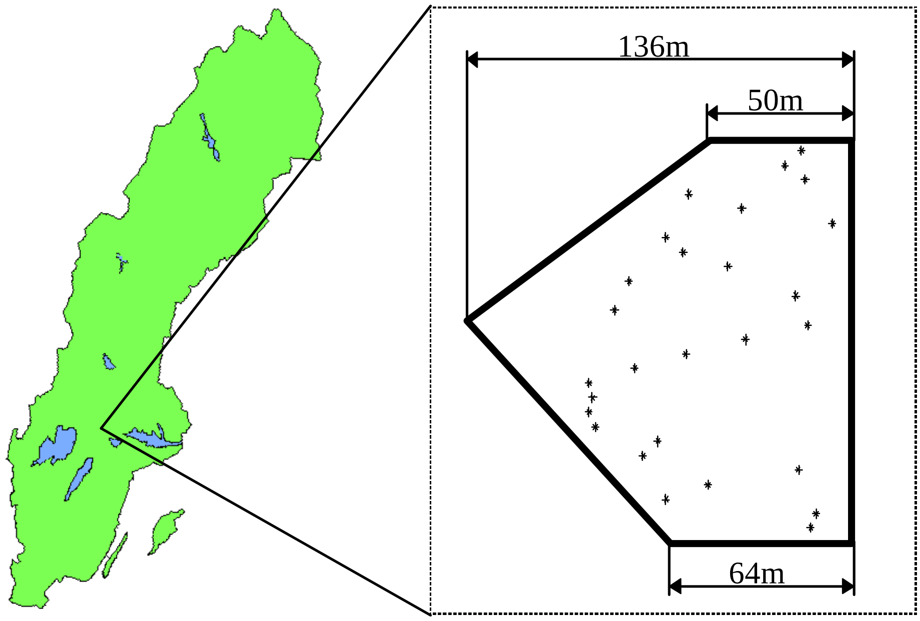


Figure 1. Area where images where taken. Stars mark photograph spots.

The photos were enumerated in the order they were taken. This number was included in the image scene on a small piece of paper. In Figure 2, to the right, one can see the number paper mounted on a stand made of rigid paper and steel wire. The stand actually has three different regions made of red, green and blue paper respectively. This was originally intended to be used as a color normal to give the computer possibility to examine light condition. This were however not used.

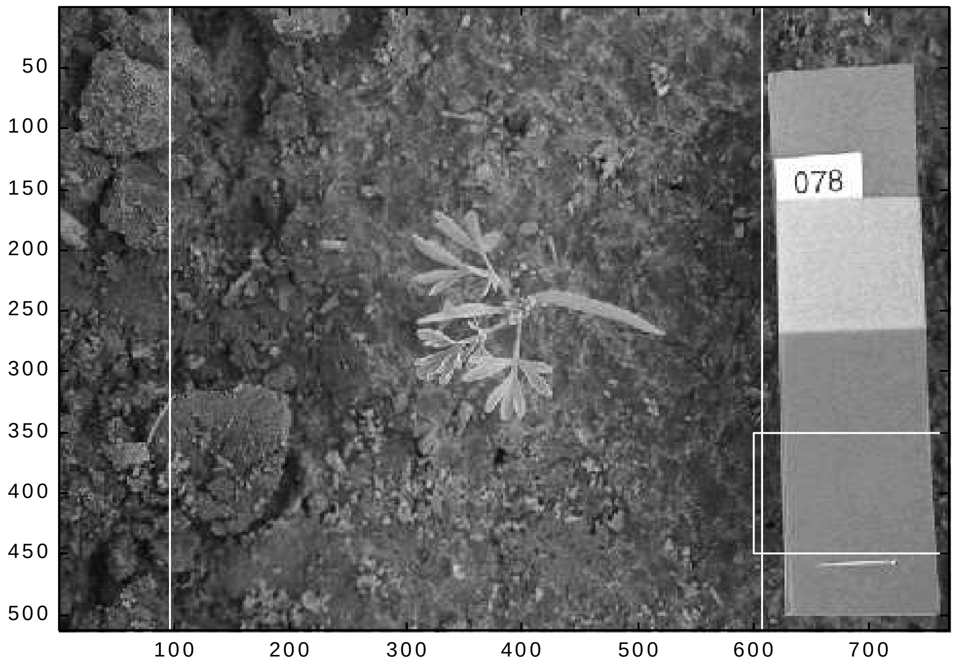


Figure 2. Sample image. The two vertical lines mark the square cut out positions and the rectangle in the lower right corner was used when calculating the cut out positions.

When the images had been gather the remaining main steps of the image analysis that follows were; *segmentation* i.e. separating plant from soil, calculating *features* that gives quantitative values of what the image contains and finally *classification*.

Figure 3. Sample images of plants; (a) Carrot; (b) Ladythumb smartweed; (c) Canada thistle; (d) Lamb's quarters.

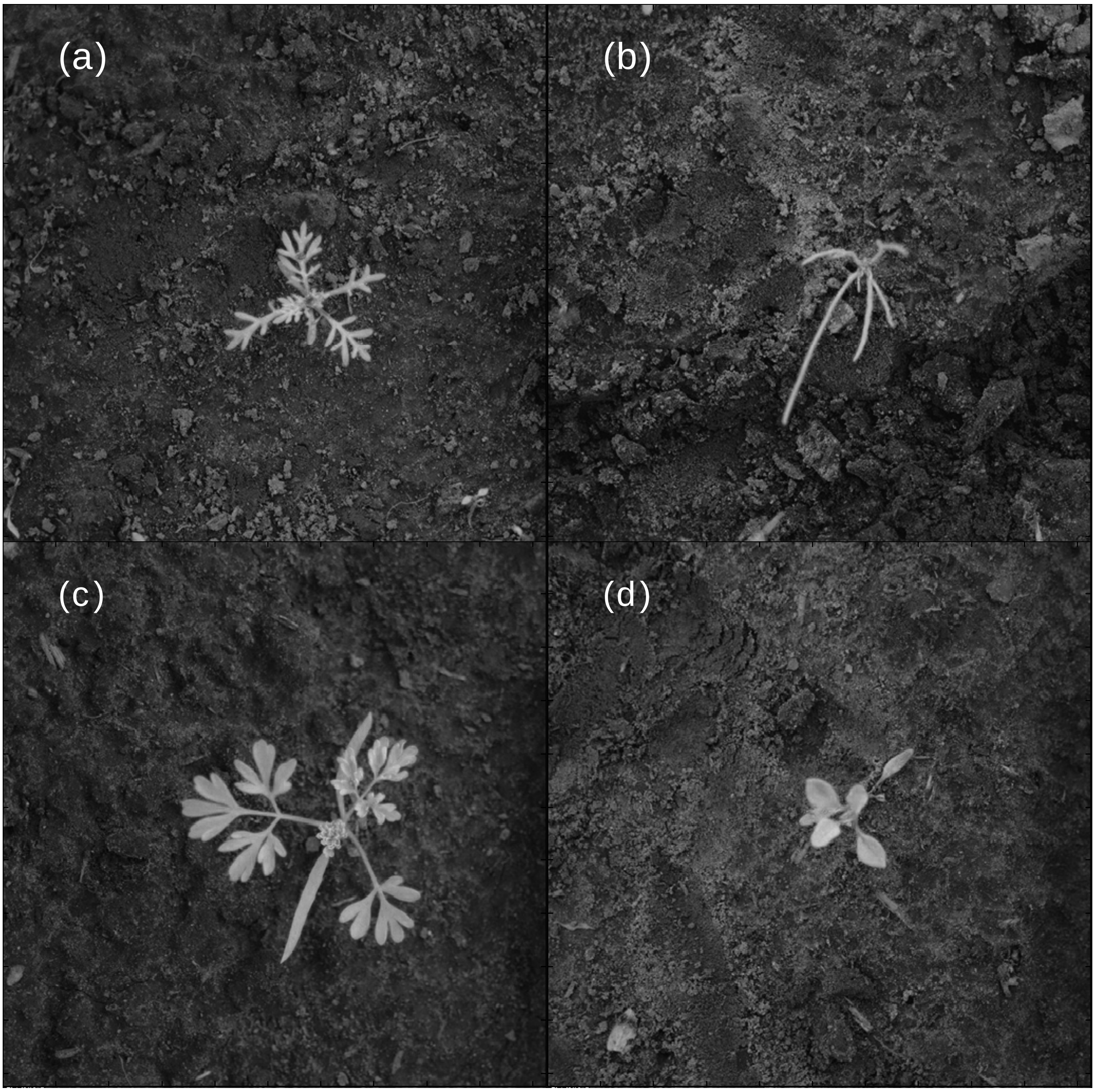


Figure 4. Sample images of plants; (a) Mayweed; (b) Corn Spurry; (c) Fumitory; (d) Chickweed.

3 SEGMENTATION

To be able to do anything useful, one must first find out what is plant and what is soil in the images, this is the task of segmentation. However to begin with every image was cut to a square from 768*512 pixels down to 512*512 pixels. This was done first of all to cut away the number holder and secondly to decrease computing time by removing some of area that only contained soil. When this was to be done a tricky problem immediately was encountered. The square could not always be centered in the image, for in some images, when the plant was small the number holder had been placed more left, toward the middle of the image. To be sure to cut away the number holder the square had to be moved left. On the other hand in images with big plants the number holder was always placed rightmost and the plant was nearly touching it. So having the square always moved left did not work either. The square had to be positioned individually for every image, in Figure 2 the square has been marked with to vertical lines.

To position the square automatically the position of the number holder had to be found. This was done by examining red part of the holder. In the rectangle, marked in the lower right corner of Figure 2, the centroid of the normalized red color component was calculated. Let *R*, *G* and *B* denote red, green and blue color component respectively then the normalized red color component is

$$r = \frac{R}{R+G+B} \qquad [\ 1\ ]$$

The value of *r* ranges from 0 to 1. In the rectangle, *r* was thresholded with the threshold equal to 0.5 so that pixels over 0.5 were set to ‘1’ and pixels lower were set to ‘0’. Then the centroid was calculated. An offset was then subtracted from the x-value of the centroid to get the right cut away position and another offset of 511 was subtracted from this position to get the left cut away position. This would yield a image that contained no part of the number holder and with a resolution of 512*512 pixels.

3.1 THRESHOLDING

To do segmentation, the fact that plants are more green than soil were used. Let, as before, *R*, *G* and *B* denote the color components, then for each image the following was calculated

$$g = 255 * \frac{G}{R+G+B} + 1 \qquad [\ 2\ ]$$

The *g*-value ranged from 1 to 256. Figure 8 (a) shows the green color component of a Lamb’s quarter and (b) shows the image of the *g*-value calculated from it. For every image a histogram of the *g*-value was created with 256 bins. A threshold value was then to be found so that pixels with *g*-value greater than this threshold were treated as plant pixels and lower were soil pixels. To find the threshold from the histogram several methods exist, three methods were explored in this project. They will now be described.

Method 1: This method is called the midpoint method. The threshold, T_0 in Figure 5, is initiated with an arbitrary value (e.g. the center of the histogram). Then from T_0, with n=0

$$\mu_0(T_n) = \frac{\sum_{z=1}^{T_n} z * h(z)}{\sum_{z=1}^{T_n} h(z)}, \quad \mu_1(T_n) = \frac{\sum_{z=T_n}^{256} z * h(z)}{\sum_{z=T_n}^{256} h(z)} \qquad [3]$$

are calculated, marked in Figure 5. These are the means of the histogram to the left and right respectively of T_0. The new threshold T_1 is then constructed, with again n=0, as

$$T_{n+1} = \frac{\mu_0(T_n) + \mu_1(T_n)}{2} \qquad [4]$$

From T_1, $\mu_1(T_1)$ and $\mu_0(T_1)$ are calculated and so on. Expressions [3] and [4] are repeated until T does not change.

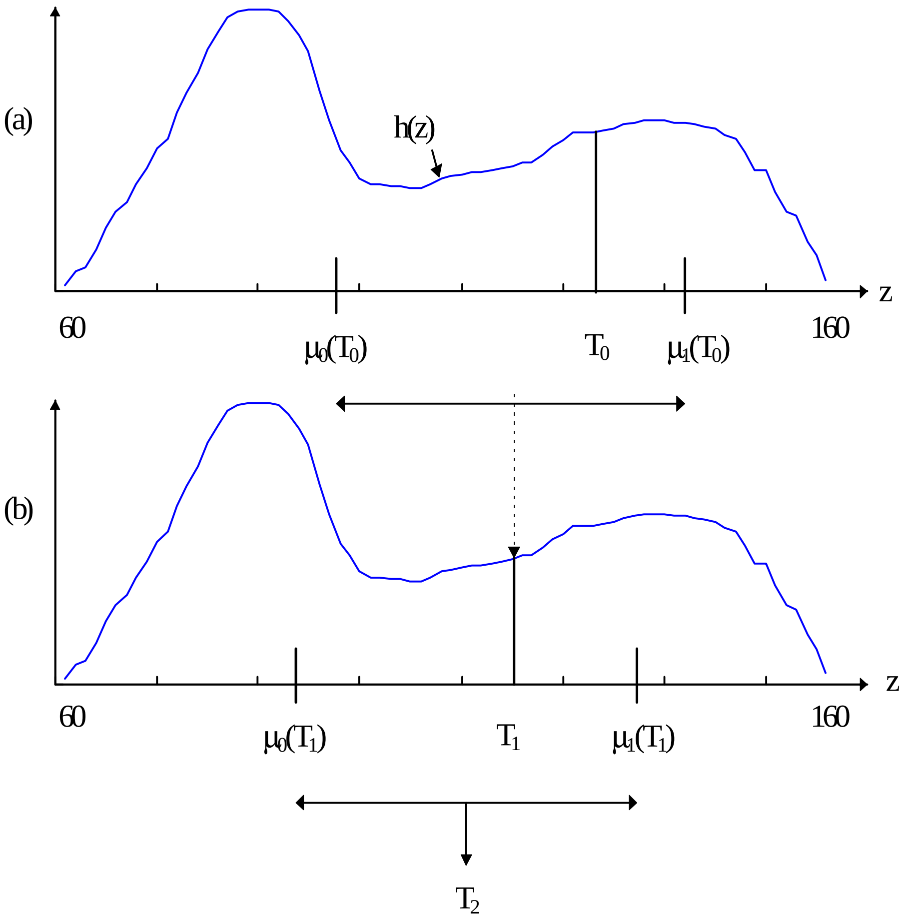


Figure 5. Midpoint method. The curve, h(z), is the histogram of an image to be segmented. The threshold is calculated iteratively, starting at an arbitrary value.

When the midpoint method was tested on the images in this project it showed to suffer from a problem. Since the soil mode in the histograms was generally much greater than the plant mode, the midpoint method tended to put the threshold in the middle of the soil mode. For splitting the soil mode in two halves would put the threshold right between μ_0 and μ_1 because the plant mode would contribute almost nothing to pull μ_1 more to the right. To solve this problem the natural logarithm of the histogram was calculated to reduce the difference between the soil mode and plant mode. The new histogram was also median filtered, which had none or little effect on the midpoint method, but was very important in method described next.

Method 2: In this method the minimum point between the two modes is used as the threshold value. To find the local minima, the global maximum is first located, marked as *Start* in Figure 6. The threshold T is initiated with the value of *z* at *Start*. Then as long as h(T+1)≤h(T) one is added to T. This means that the threshold is moved right until h(z) increase.

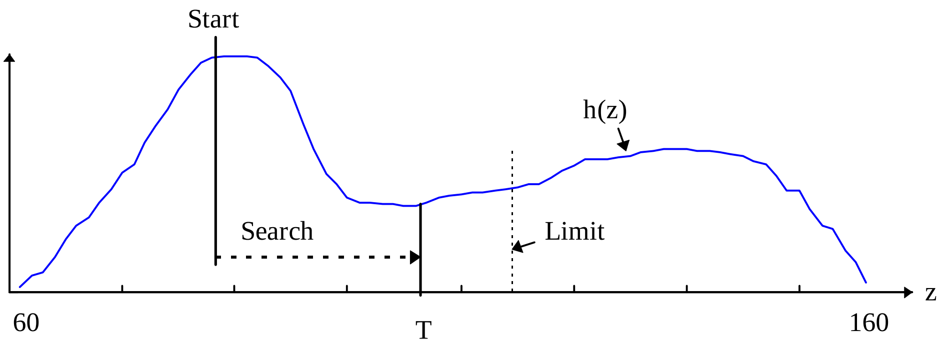


Figure 6. Using local minimum as point of threshold.

The histograms generated from the images in this work had to be median filtered before they could be searched for local minimums. This because the search often got caught on spikes in the histograms. Median filtering was done with the width w=5 and as in the midpoint method the natural logarithm were calculated from the histograms, this has no effect on finding the minimum but was done for the sake of symmetry with method 1 and also with method 3. Median filtering solved the problem with spikes, but another problem was that in some histograms there were no local minimum. A limit was therefore introduced, marked as *Limit* in Figure 6. When the search past the limit it was considered that there existed no local minimum and that the search had failed. The limit had a constant value which was set manually by observing histograms from several images. In all images with no exceptions was the soil mode the tallest mode and it was always located left of the plant mode.

Method 3: This method is sometimes called the minimum error method. It works by fitting the sum of two normal distributions,

$$p(z) = \frac{P_1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(z-\mu_1)^2}{2\sigma_1^2}\right] + \frac{P_2}{\sqrt{2\pi}\sigma_2}\exp\left[-\frac{(z-\mu_2)^2}{2\sigma_2^2}\right] \quad [5]$$

to the histogram of an image. In [5] μ_1 and μ_2 are the mean values of the two modes, σ_1 and σ_2 are the standard deviation about the means and P_1 and P_2 are the a priori probabilities of the two modes. The a priori probabilities are constrained by

$$P_1 + P_2 = 1 \quad [6]$$

which must be satisfied, so there are five unknown parameters. A minimum square error approach is used to fit *p(z)* to the actual histogram. To speed up calculations and to waste no effort on fitting *p(z)* to the left most and right most parts of the histograms these parts were left out. This was done by only calculating [5] between $0.95*z_{max}$ and 150 in the histograms, where z_{max} is the location of the soil mode maximum. The error function to be minimized is then

$$e_{ms} = \sum_{z=0.95z_{max}}^{150} \left[p(z) - h(z)\right]^2 \qquad [\,7\,]$$

where an 256-point histogram is assumed. Minimizing [7] is done iteratively. In every step the gradient of [7] is calculated. The gradient is a R^5-vector that points in the direction in which [7] increase the steepest. So adjusting the five parameters in the opposite direction as the gradient is pointing will decrease the value of [7]. The iteration is continued until the parameters settle.

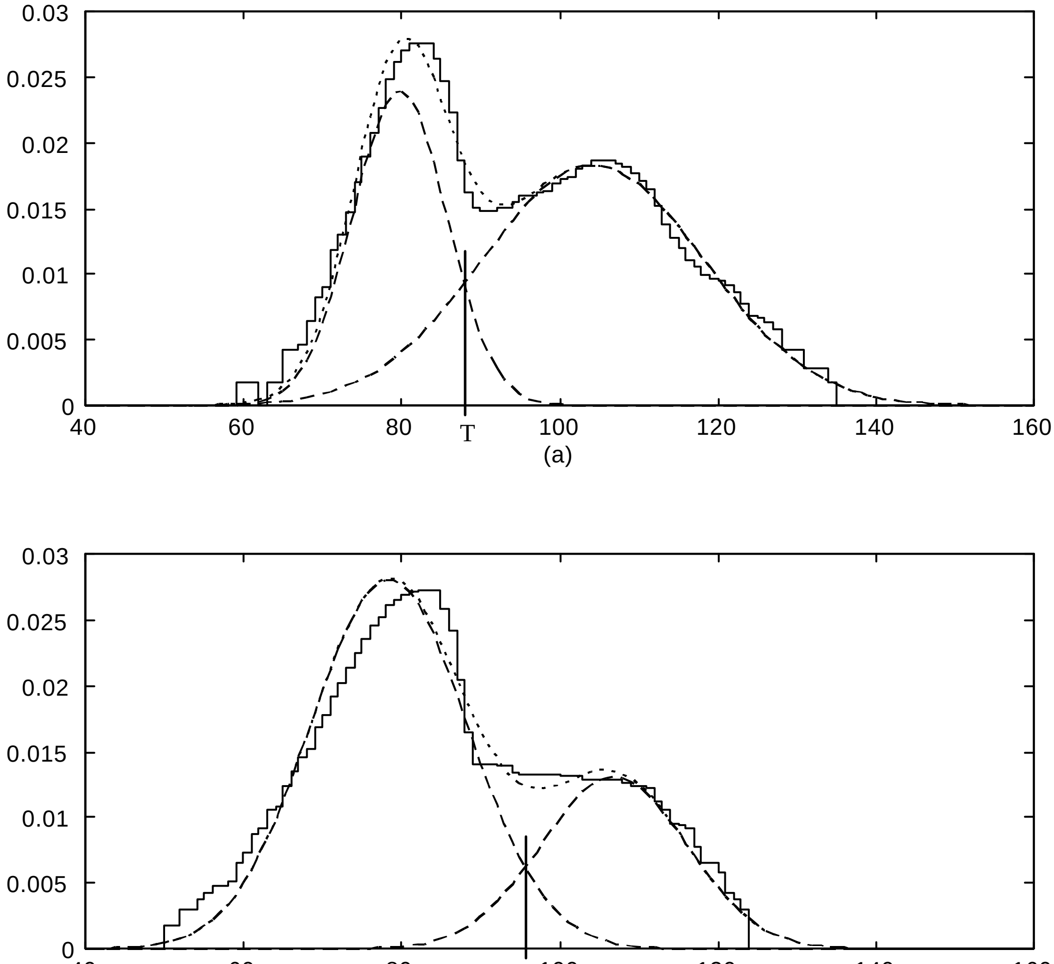


Figure 7. The stair shaped curves in (a) and (b) shows histograms that are bimodal and unimodal respectively. In each plot dashed curves show the two normal distributions who's sum, dotted curve, is fitted to the actual histogram in the least square error sense.

At first when the minimum error method was tested on the original histograms the thresholds obtained were not particularly good. Since the soil mode was in general much bigger then the plant mode the contribution to the error function was much greater from the soil mode then from the plant mode. This had the effect that the normal distribution was poorly fitted to the plant mode and thereby yielding a poor threshold. Instead as in method 1 and 2 the natural logarithm was calculated to reduce the difference between the modes. When this was done the normal distributions were much better fitted to the histograms and the thresholds were better.

Conclusions: The three threshold methods were evaluated visually by observing the resulting binary images from segmentation performed with each of the three thresholds for several images. The thresholds obtained with the midpoint method varied a lot, for some images the threshold was placed too much to the left, segmenting bright soil regions as plant. On the other hand, sometimes the threshold was placed too much to the right, loosing parts of the plant. Using the local minimum between the two modes as threshold gave good segmentation, except for those images where no local minimum was found. The minimum error method tended to put the threshold a little bit left of the optimal threshold.

When a local minimum was found this method was used, but on those occasion, about 5% of the images, when no minimum was found the threshold obtained with the minimum error method was used. Figure 8 (c) shows the resulting binary image of a lamb's quarter.

3.2 NOISE REMOVAL

As seen in Figure 8 (c) the binary image contains some noise. Some of the noise regions consist of just a few pixels, these regions can be removed with a standard *open*-operation. To the left in Figure 8 (c) a larger noise region is seen. This region will not be removed by an *open*-operation, instead a special method was developed to deal with larger noise regions. First some fundamental image processing concepts will be described.

Neighbors: The concept of neighboring pixels is important in image processing. The two most common definitions are

4-connectivety 8-connectivety

where void squares are neighbors to black squares.

Erode & dilate: These operations are defined for binary images where '0' represent background and '1' represent object. Erode is the operation of changing all pixels with value '1' that neighbors at least one pixel with value '0' to '0', neighbors in the sense of 8-connectivety. Dilate does the opposite thing. It changes a background pixel to '1' if it has at least one neighbor pixel that is '1', again using 8-connectivety.

Open & Close: Erode and dilate can be combined to form open and close operations. In open, an erode operation is followed by a dilate operation. It has the effect of removing small pixel regions. If erode and dilate are combined in the opposite order (i.e. first dilate followed by erode) they form a close-operation. The close-operation will fill small holes in an object.

To the binary images obtained in this study an *open*-operation was first asserted to remove noise regions followed by a close-operation to fill holes in the plants. This were however not enough, because large noise regions did not disappear. Instead a method was developed that first of all calculated the centroid of the entire binary image. The centroid was generally place some where near the middle of the plant. When this was done the binary image was searched for pixel regions and there size was determined as the number of pixels they consisted of. The centroid was also calculated for every individual pixel region. Finally, the size of a region was divided by the distance between the total centroid and the centroid for that particular region and when this quotient was below a certain constant the pixel region was removed. This method had the effect of removing larger regions further from the plant and leaving the pixel region or regions that made up the plant it self. As seen in Figure 8 (d) the noise region to the left of the plant and the noise regions to the right of the plant in (c) are removed.

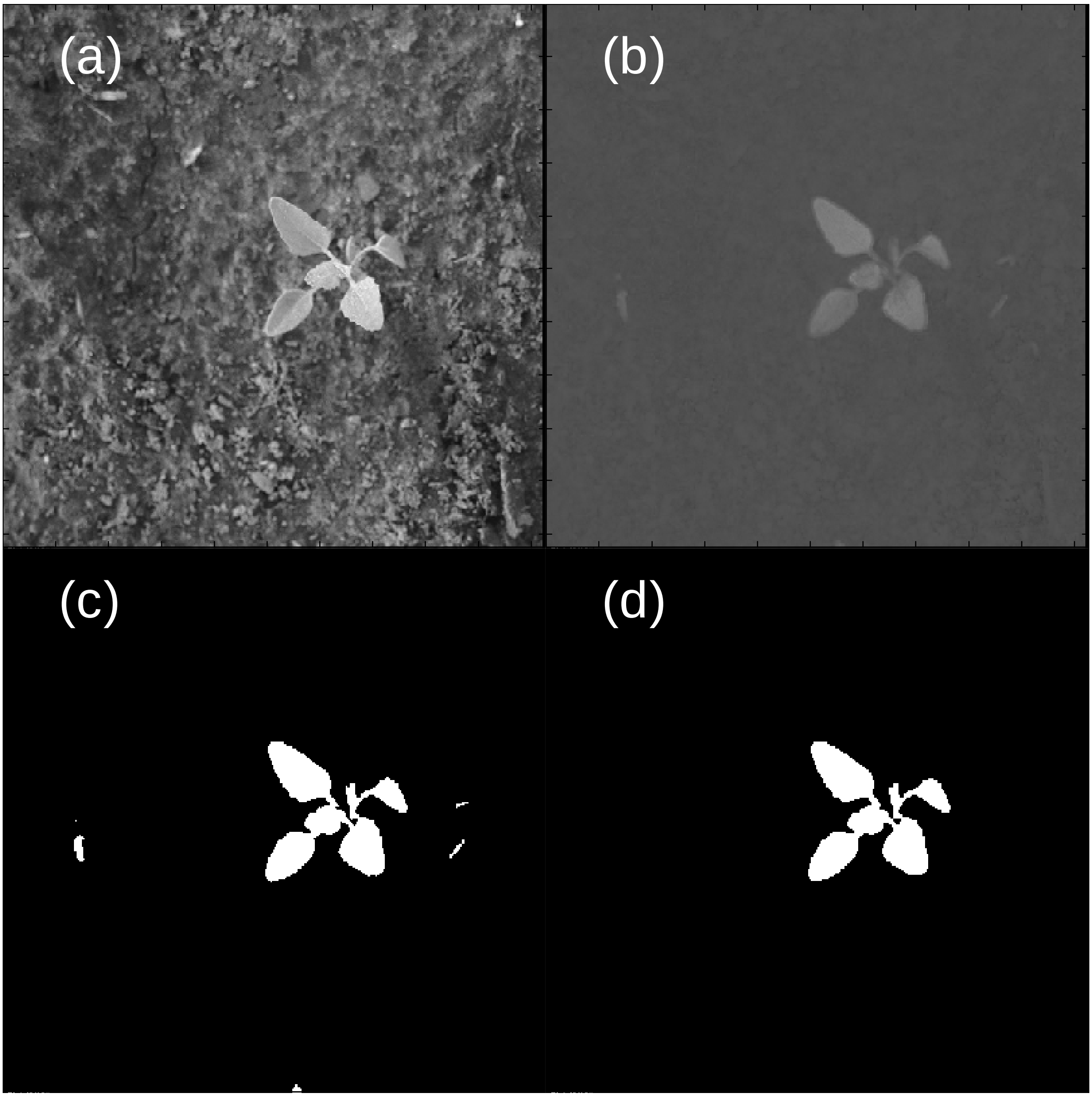


Figure 8. Images of a Lamb's quarter (a) Green color component; (b) g-value; (c) segmented binary image; (d) binary image after noise removal.

4 DESCRIPTION OF FEATURES

4.1 SIZE DEPENDENT OBJECT DESCRIPTORS

When the pictures have been segmented into binary images some descriptors of plant shape can be calculated. Descriptors calculated are:

1) *Area*, defined as the number of pixels with value ‘1’ in a binary image.

2) *Perimeter* is defined as the number of pixels with value ‘1’ for which at least one of the eight neighboring pixels is a soil pixel, e.g., the number of pixels removed by one erode-operation applied to a binary image. Figure 9 (b) shows the perimeter for a fumitory.

3) *Thickness* is twice the number of shrinking steps (elimination of border pixels one layer per step) to make an object within an image disappear. This is the same definition as used by (Guyer *et al.*, 1986).

4) *Convex area* is the area of the smallest convex hull that contains all objects in an image. The convex hull can also be described as the area formed if one would tighten a rubber band around all objects. Figure 9 (a) shows the convex hull for a sample fumitory.

5) *Convex perimeter* is the perimeter of the convex hull described in (4).

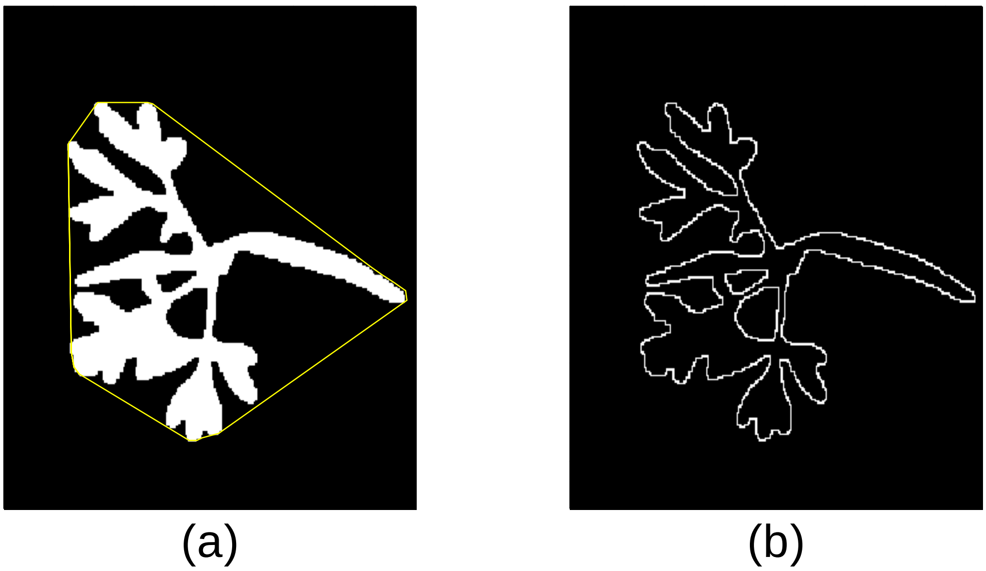


Figure 9. (a) Segmented fumitory with convex hull; (b) Perimeter of the same fumitory.

4.2 SIZE INDEPENDENT SHAPE DESCRIPTORS

The size dependent object descriptors can be combined to produce size independent shape descriptors. The following features were calculated:

1) $$\text{Formfactor} = 4\pi \frac{\text{Area}}{\text{Perimeter}^2}$$

2) $$\text{Elongatedness} = \frac{\text{Area}}{\text{Thickness}^2}$$ (Guyer *et al.*, 1986)

3) $$\text{Convexity} = \frac{\text{Convex\_perimeter}}{\text{Perimeter}}$$

4) $$\text{Solidity} = \frac{\text{Area}}{\text{Convex\_area}}$$

All these features are dimensionless and independent of size. Formfactor equals one for a circle and less than one for all other shapes. Elongation is generally high for a long narrow object and low for a short wide object. This feature, however, has a problem, if the object is short and wide but has holes in it, perhaps due to errors in segmentation, the object will shrink away faster than it should and yield a high value instead of a low value. For an already convex object convexity will be close to one and decrease when the object becomes more straggly.

For all 216 plants these four features were calculated and the results are presented in two plots in Figure 10. These two plots show that each species is more or less tightly group and that the groups overlap each other.

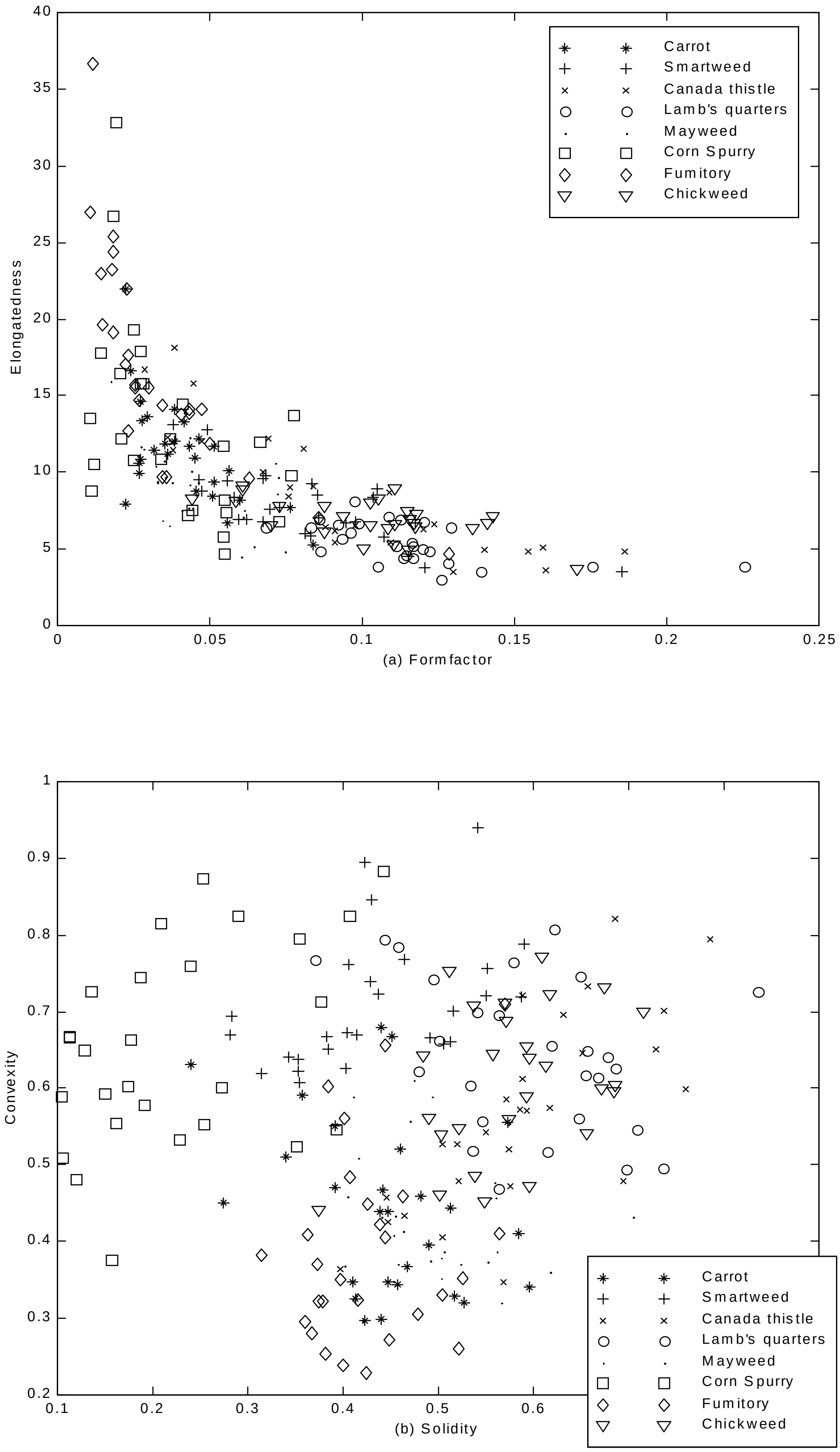


Figure 10. (a) Elongatedness vs. formfactor; (b) Convexity vs. solidity. The plots show that there exist groupings but that they overlap each other.

4.3 MOMENT INVARIANTS

The moments are measures of how spread an object's area is. The moment invariants will only be briefly described, for more details refer to (Jain *et al.*,1989). By always considering central moments

$$\mu_{p,q} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x,y) \qquad [\,8\,]$$

the moments become invariant to translation. In [8] *f(x,y)* is '1' for those pairs of *(x,y)* that correspond to plant pixels and '0' for soil pixels. The moments are made invariant to scaling by calculating the normalized moments, defined as

$$\eta_{p,q} = \frac{\mu_{p,q}}{(\mu_{0,0})^\gamma} \qquad \gamma = (p + q + 2)/2 \qquad [\,9\,]$$

which makes the them independent of size. Finally the moments are made invariant to rotation and reflection. Below the moment invariants used in this study are listed.

$$\begin{aligned}
\Phi_1 &= \eta_{2,0} + \eta_{0,2} \\
\Phi_2 &= (\eta_{2,0} - \eta_{0,2})^2 + 4\eta_{1,1}^2
\end{aligned} \qquad [\,10\,]$$

$$\begin{aligned}
\Phi_3 &= (\eta_{3,0} - 3\eta_{1,2})^2 + (\eta_{0,3} - 3\eta_{2,1})^2 \\
\Phi_4 &= (\eta_{3,0} + \eta_{1,2})^2 + (\eta_{0,3} + \eta_{2,1})^2 \\
\Phi_5 &= (\eta_{3,0} - 3\eta_{1,2})(\eta_{3,0} + \eta_{1,2})[(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2] \\
&\quad + (\eta_{0,3} - 3\eta_{2,1})(\eta_{0,3} + \eta_{2,1})[(\eta_{0,3} + \eta_{2,1})^2 - 3(\eta_{1,2} + \eta_{3,0})^2] \\
\Phi_6 &= (\eta_{2,0} - \eta_{0,2})[(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2] + 4\eta_{1,1}(\eta_{3,0} + \eta_{1,2})(\eta_{0,3} + \eta_{2,1})
\end{aligned} \qquad [\,11\,]$$

The moment invariants were calculated on object area and also on object perimeter. When plotting the calculated moments they showed to be exponential, hence taking the natural logarithm would make them more linear. One problem occurred when calculating the logarithm, because for some plants Φ_5 and Φ_6 had negative values and that makes the logarithm complex. It was though early noticed, by observing plots, that Φ_2, Φ_3, Φ_4, Φ_5, and Φ_6 contained, roughly, the same information as Φ_1 for both area and perimeter, so for simplicity Φ_5, and Φ_6 were disregarded during classification. The moment invariants were given the following names:

1) area_moment_1, ln(Φ_1) of area.
2) area_moment_2, ln(Φ_2) of area.
3) area_moment_3, ln(Φ_3) of area.
4) area_moment_4, ln(Φ_4) of area.
5) perimeter_moment_1, ln(Φ_1) of perimeter.
6) perimeter_moment_1, ln(Φ_2) of perimeter.
7) perimeter_moment_1, ln(Φ_3) of perimeter.
8) perimeter_moment_1, ln(Φ_4) of perimeter.

4.4 COLOR FEATURES

Six features regarding plant color were calculated. The digital color images are represented by three color components; red, green and blue. To try to make the color features independent of different light condition every component was divided by the sum of all three components, e.g. the gray level, before calculating mean value and standard deviation. Let R, G and B denote the red, green and blue components respectively, then the new components are:

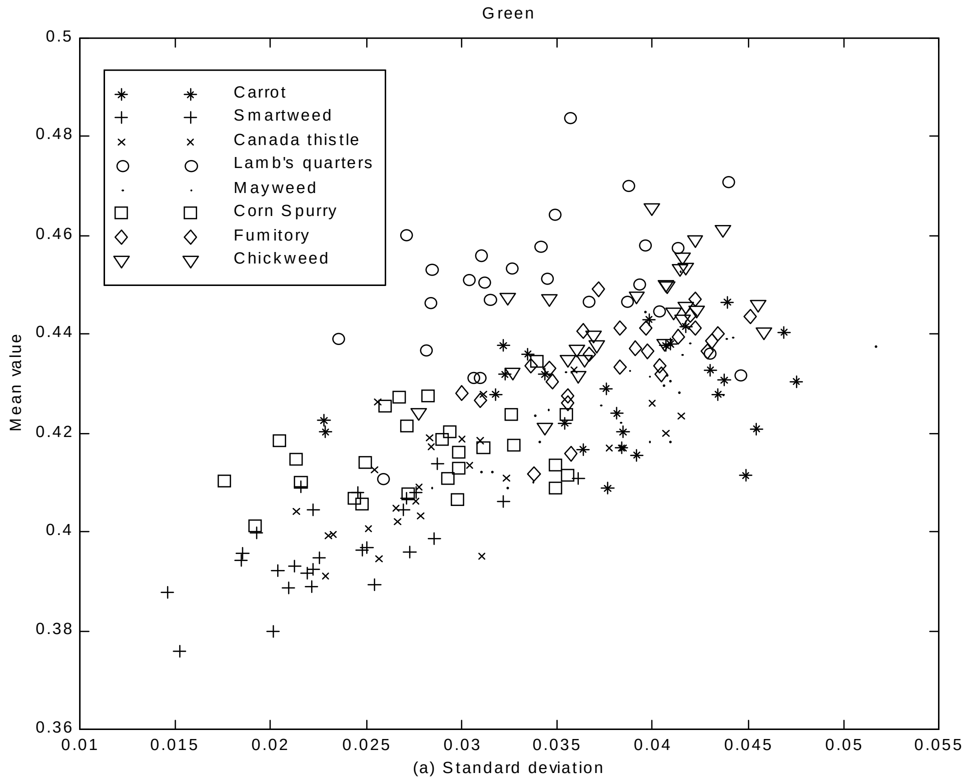
$$r = \frac{R}{R + G + B}, \quad g = \frac{G}{R + G + B}, \quad b = \frac{B}{R + G + B} \qquad [12]$$

The new variables r, g and b ranges from 0 to 1. For every component mean value and standard deviation was calculated. Only plant pixels, e.g. pixels with value '1' in the binary image, were used when calculating the color features, so the features are only based on plant color not soil color. The six color features were given the following names:

1) *Red_mean*, mean value of red.
2) *Red_std*, standard deviation of red.
3) *Green_mean*, mean value of green.
4) *Green_std*, standard deviation of green.
5) *Blue_mean*, mean value of blue.
6) *Blue_std*, standard deviation of blue.

These names will be used in the classification section.

It is easy to realize that the sum of the mean values for all three components is always equal to one, hence one mean value can always be derived from the two others. Figure 11 shows how the 216 plants were distributed for green and blue color features.



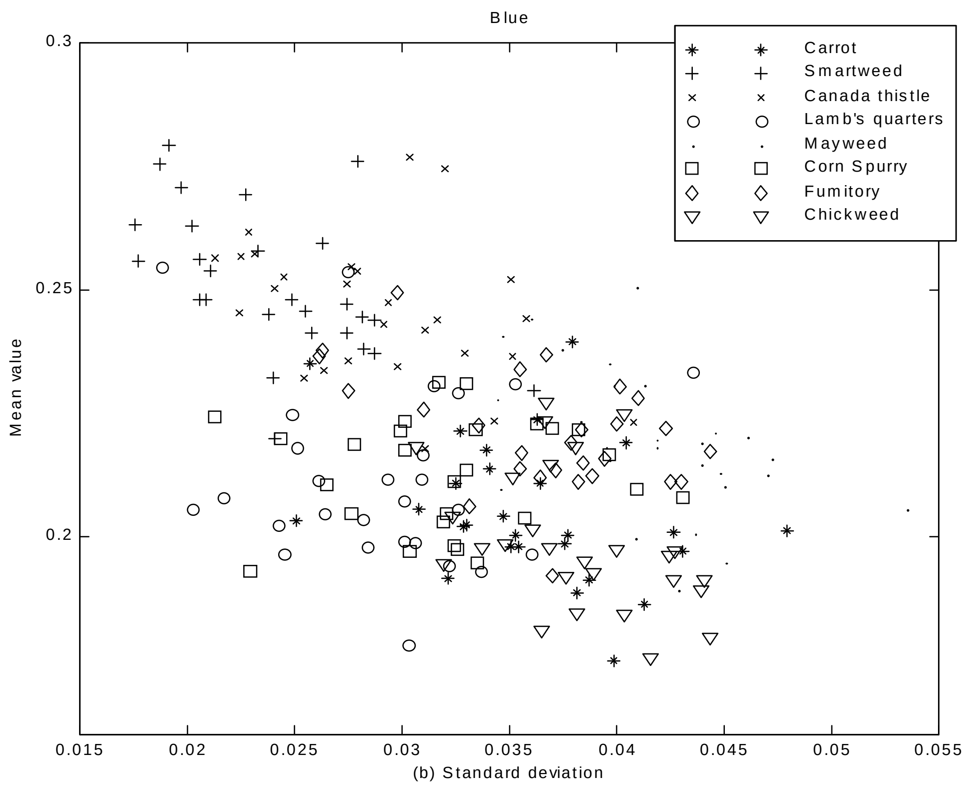


Figure 11. Mean value vs. standard deviation of (a) green and (b) blue. Species are ordered into overlapping groups.

5 PLANT CLASSIFICATION

To do image classification a probabilistic approach will be used that assign objects into defined groups. The statistical method for this classification is called discriminant analysis and is a common method in image recognition. To each species a density function is assigned called the Bayes decision function. The density functions have parameters that are calibrated by a training set. For more detailed information about the theories and expressions used in the discriminant analysis refer to (Gonzalez *et al*., 1993).

To describe this method let us consider a case with only two species or pattern classes and one feature. With means m_1 and m_2 and standard deviations σ_1 and σ_1 the Bayes decision function have the form:

$$d_j(x) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left[-\frac{(x-m_j)^2}{2\sigma_j^2}\right] P(\omega_j) \qquad j = 1,2 \qquad [\,13\,]$$

where x denote the pattern to be classified and ω_j denote a particular pattern group or species. In this study all species are considered equally likely to occur, hence $P(\omega_j)$ is equal for every density function and is dropped. In the 1-D case the border between the two groups can easily be found. Border is marked as x_0 in Figure 12. This means that when classifying, a pattern or plant is said to belong to group 1 when its x-value is below x_0 and group 2 when it is above.

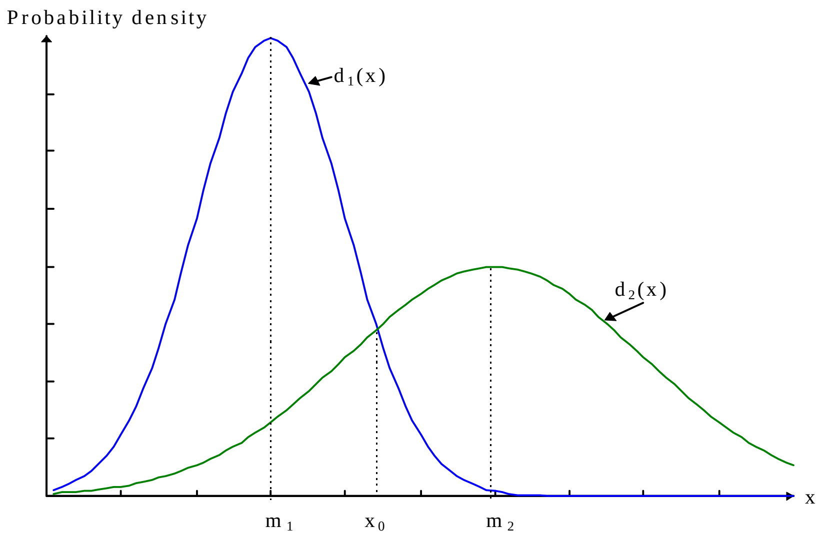


Figure 12. Probability density functions for two 1-D pattern classes. The x_0 line marks the boundary between the two classes.

In the case of multiple features and multiple species the borders, which are hyper spheres, are not as easily found. Instead, when classifying a plant, probability densities are calculated for all species and the one with the highest density is the winner and the plant is classified as belonging to that species.

In the *n*-dimensional case, the Gaussian density of the vectors in the *j*th pattern class has the form:

$$p(\mathbf{x}/\omega_j) = \frac{1}{(2\pi)^{n/2}\left|\mathbf{C}_j\right|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\mathbf{m}_j)^T \mathbf{C}_j^{-1}(\mathbf{x}-\mathbf{m}_j)\right] \qquad [\ 14\ ]$$

where each density function is completely specified by its mean vector $\mathbf{m}_j$ and covariance matrix $\mathbf{C}_j$, which are estimated by

$$\hat{\mathbf{m}}_j = \frac{1}{N_j}\sum_{\mathbf{x}\in\omega_j}\mathbf{x} \qquad [\ 15\ ]$$

and

$$\hat{\mathbf{C}}_j = \frac{1}{N_j}\sum_{\mathbf{x}\in\omega_j}\mathbf{x}\mathbf{x}^T - \mathbf{m}_j\mathbf{m}_j^T \qquad [\ 16\ ]$$

where N_j is the number of pattern vectors of class ω_j, and the summation is taken over these vectors.

For discrimination at least three different ways of using the image data exists. First there is cross-validation or jackknifing. In this method all images but the image to be classified are used to estimate the parameters for the decision function. The image to be tested is removed from the training set. Cross-validation is the method that will be used in this study. Secondly there is resubstitution where the parameters of the discriminant functions are calculated from the same set of images which is classified in to groups (i.e. training set and test set are the same). Resubstitution will be performed on the best feature combination found for comparison with cross-validation. Last there is a method where the training set and test set are to different sets of images. This method uses a separate group to construct the decision function and another group for testing.

Several methods exist to select those features that give the best classification, a few of them will now be described. There is forward-selection, where the selection process starts with one feature and all other features are added, one at a time, to the first feature to see which of them increases the classification rate the most. The feature that was best is then selected, there are now two features and a third is added the same way as the second. This process continues until no further improvement of classification rate is noticed. It should be said that forward-selection does not guarantee to find the optimal feature combination. The opposite way is backward-elimination. In backward-elimination the examination starts with all features included. The task is then to remove that feature which contributes the least to the classification rate. This process is repeated so that in every step the least significant feature is eliminated. The elimination process is ended when a certain stopping rule is met. Forward-selection and backward-elimination can be combined to form stepwise-selection. Stepwise-selection starts with forward-selection and for each feature selected backward-elimination is applied to the selected subset. Thus it is possible to eliminate a feature previously selected. Other, even more refined methods exist, but in this work forward-selection and backward-elimination will be used for simplicity and no other method is therefore examined further.

6 RESULTS

In the following sections, classification will be performed on separate feature groups as well as on all features together. The feature groups are; size dependent features, size independent shape features, moment invariants and color features. Three classifications where all features are used, excluding the size dependent features, are performed. Finally a classification result based on a feature combination where the size dependent features are included is presented.

In 9 of 27 images of carrots the image actually contains two carrots standing very close to each other. The second carrot plant had accidentally been left in the scene, this due to sometimes rather stressful conditions in the field, caused by sudden showers. It will be noted how these images were classified. In confusion matrices and brief result tables presented in the following sections values in brackets show the number of images containing two carrots that were classified in a particular way.

In the following sections a classification rate will be presented as a percent value. This classification rate is the number of correctly classified plants divided by the total number of images (i.e. the sum of the diagonal elements in a confusion matrix divided by 216).

6.1 CLASSIFICATION RESULTS OF SEPARATE FEATURE GROUPS

Classification using separate feature groups will be examined. For each group forward-selection was used and the search was repeated a few times with different starting feature to try to find the best combination within that group.

6.1.1 SIZE DEPENDENT SHAPE FEATURES

This is the feature group described in 4.1, it contains a total of five features. The following combination was the best found: **{area, perimeter, convex_perimeter, thickness}.**

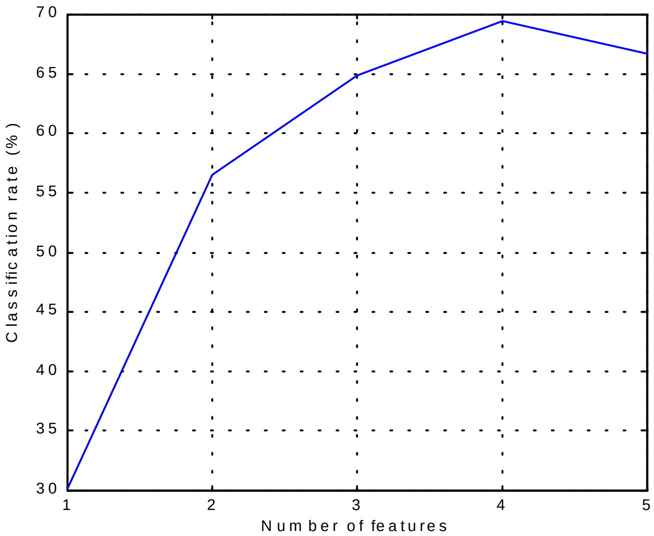


Figure 13. Classification rates for each step, selected in order as above. When the fifth and also last features was added the classification rate decreased.

Table 2. Confusion matrix. Rows are true species and columns are classification.

		1	2	3	4	5	6	7	8
1	Carrot	15(4)	1	1(1)		3(1)	4(2)	1	1(1)
2	Ladythumb Smartweed	2	20		1			1	3
3	Canada thistle			17	4				6
4	Lamb’s Quarters		3	2	18				4
5	Mayweed	9		1		15		1	1
6	Corn Spurry		1	1			24		1
7	Fumitory	1	1	1		1		21	2
8	Chickweed		2	2		3			20

Table 3. Brief results. **Size dependent shape features**.

Total number of misclassifications	Misclassified carrots	Weeds classified as carrot	Classification rate
66	**11(5)**	**12**	**69.44%**

6.1.2 SIZE INDEPENDENT SHAPE FEATURES.

There are four size independent shape features and they are described in 4.2 . Best combination found include three features: **{formfactor, convexity, elongatedness}**.

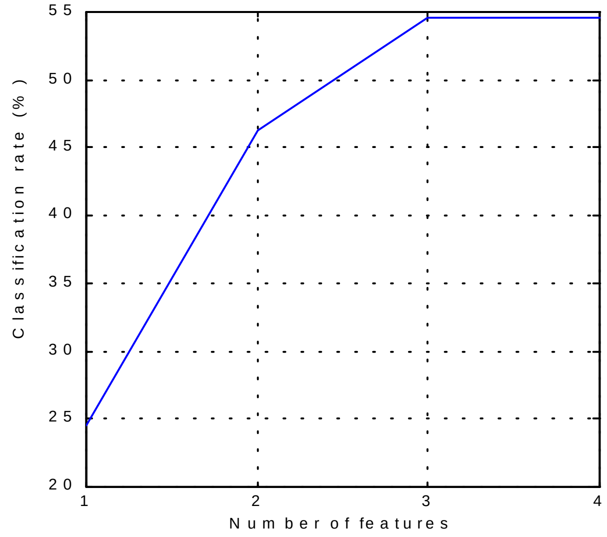


Figure 14. Classification rates during forward selection of size independent shape features. Best classification with three features.

Table 4. Confusion matrix.

		1	2	3	4	5	6	7	8
1	Carrot	12(3)	3		1	7(4)	1(1)	2	1(1)
2	Ladythumb Smartweed		20		3		3		1
3	Canada thistle	4		6	8			3	6
4	Lamb's Quarters		4	3	12				7
5	Mayweed	4		1	1	19		1	1
6	Corn Spurry	2	1	1			23		
7	Fumitory	12	1		1	2		11	
8	Chickweed		1	2	4	5			15

Table 5. Brief results. **Size independent shape features**.

Total number of misclassifications	Misclassified carrots	Weeds classified as carrot	Classification rate
98	**15(6)**	**20**	**54.63%**

6.1.3 MOMENT INVARIANTS

The moment invariants include eight features. The best combination found consisted of only two features they were **{area_moment_1, perimeter_moment_1}**. No improvement was done by adding another moment feature. Figure 15 shows how the plants were distributed for these two features and the lines show the borders between the species groups if resubstitution would be used to perform classification. The borders appear to be second degree curves as supposed. All feature combinations of length two were tested and the best combination was the same as the one presented here.

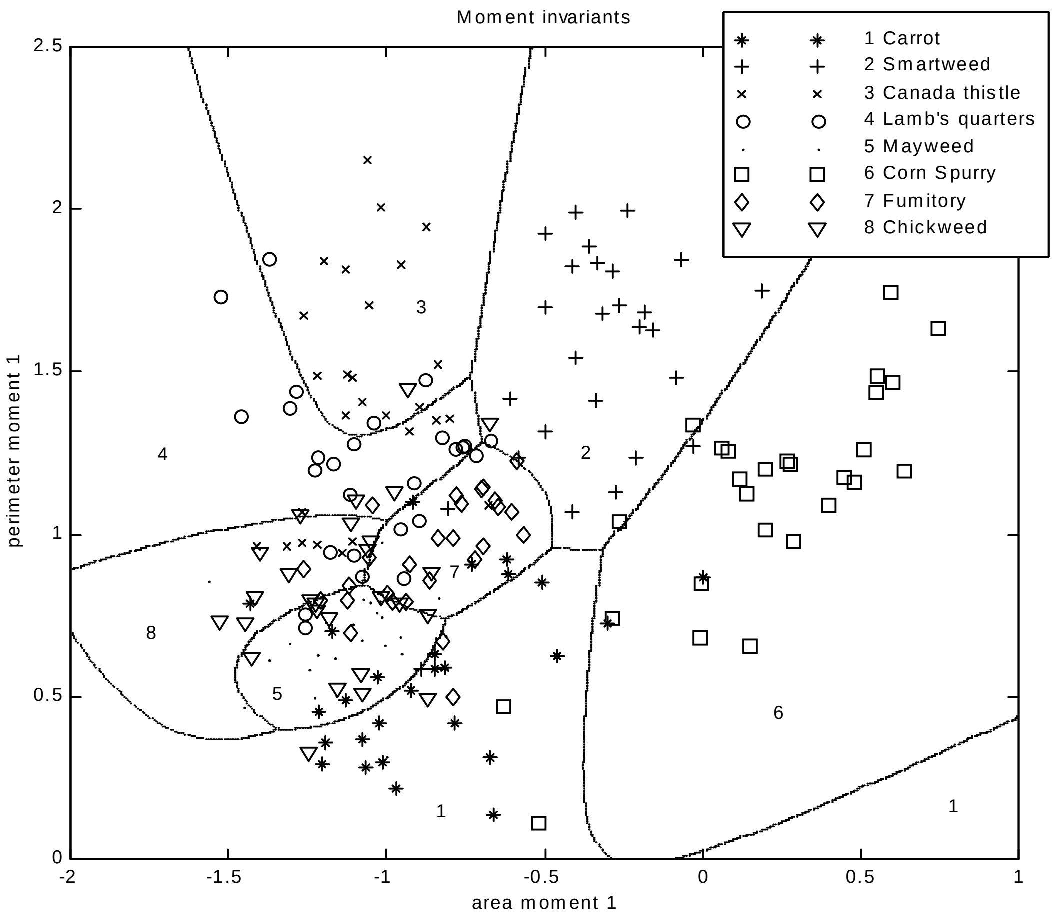


Figure 15. Perimeter_moment_1 vs. area_moment_1 with group borders drawn.

Table 6. Confusion matrix.

		1	2	3	4	5	6	7	8
1	Carrot	16(4)				4(2)	2(1)	4(1)	1(1)
2	Ladythumb Smartweed	1	24				1	1	
3	Canada thistle			16	4			1	6
4	Lamb's Quarters		1	3	14	2		4	3
5	Mayweed	3				19		3	2
6	Corn Spurry	2	2				23		
7	Fumitory	2	1		1	4		16	3
8	Chickweed	2	1	1	3	9		3	8

Table 7. Brief results. **Moment invariants features**.

Total number of misclassifications	Misclassified carrots	Weeds classified as carrot	Classification rate
80	**11(5)**	**10**	**62.96%**

6.1.4 COLOR FEATURES.

All six color features described in 4.4 were used. All three mean value features were included during forward-selection even though one can always be derived from the two other. Only two of them were selected by the forward-selection method for adding the third would be redundant. The best combination found consisted of four features and they were **{read_mean, green_mean, blue_std, green_std}**.

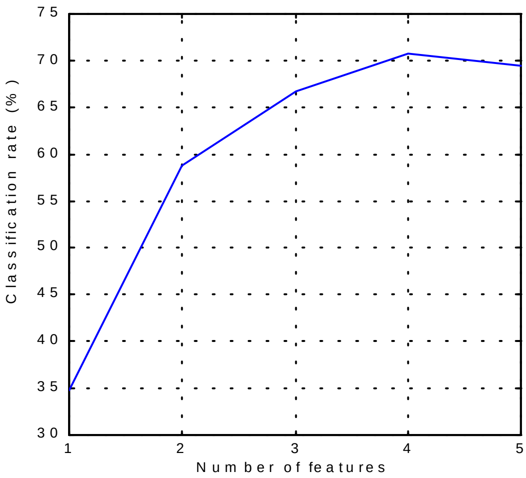


Figure 16. Classification rates for each step under forward-selection of color features. Best result is obtained with four features.

Table 8. Confusion matrix.

		1	2	3	4	5	6	7	8
1	Carrot	18(6)		2		1	3(1)		3(2)
2	Ladythumb Smartweed	1	20	5			1		
3	Canada thistle	1	8	15	1	1		1	
4	Lamb's Quarters	2	1		22			1	1
5	Mayweed	1		2		20	1	1	2
6	Corn Spurry	1	1	2		2	21		
7	Fumitory			1	3	1		19	3
8	Chickweed	2			2	2	1	2	18

Table 9. Brief results. **Color features**.

Total number of misclassifications	Misclassified carrots	Weeds classified as carrot	Classification rate
63	**9(3)**	**8**	**70.83%**

The best pair of features were area_moment_1 and perimeter_moment_1, the second best pair were green_mean and blue_mean. This example is included to illustrate the borders between species groups formed by the Bayes classifiers.

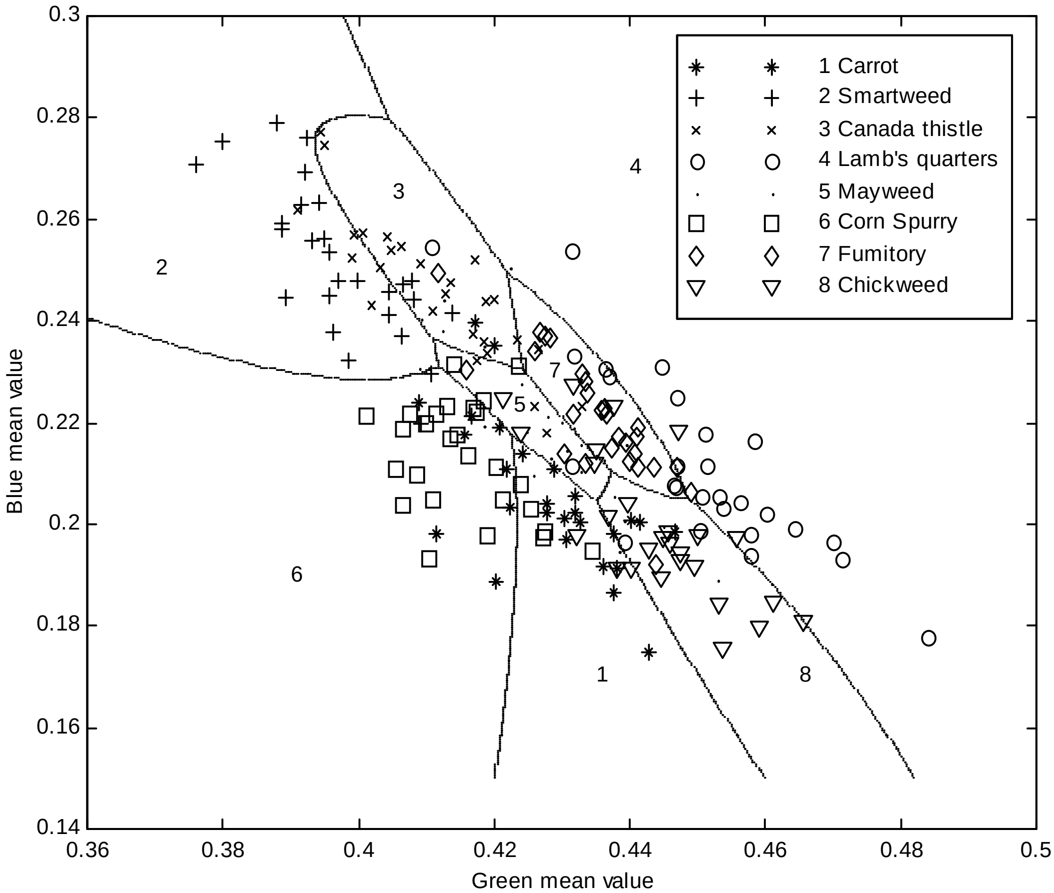


Figure 17. Blue_mean vs. green_mean with group borders drawn.

Table 10. Brief results. **Green and blue mean value**.

Total number of misclassifications	Misclassified carrots	Weeds classified as carrot	Classification rate
89	**15(3)**	**10**	**58.80%**

6.2 CLASSIFICATION RESULTS BASED ON ALL FEATURES (SIZE DEP. EXCL.)

To try to find the best feature combination forward-selection and backward-elimination were used, but the best combination was found by interactively adding and removing features. In 6.2.1 - 6.2.3 the size dependent shape features are excluded to make the classification less specific of the exact time when the images were taken. In 6.3 all features are included.

6.2.1 BEST WITH FORWARD-SELECTION

The forward-selection search was repeated a few times starting from different features to try to get different combinations. Size dependent shape features excluded. The best combination found were

{ red_mean, green_mean, convexity, formfactor, perimeter_moment_1, red_std, elongatedness}

where features are listed in order selected.

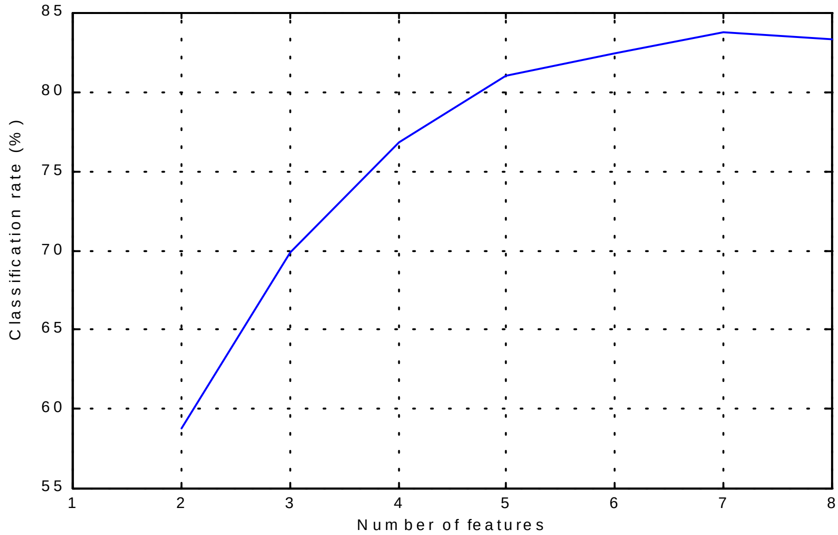


Figure 18. For the first seven features the classification rate increase, but when the eighth feature is added it decrease and forward-selection is stopped.

Table 11. Brief results. Forward-selection.

Total number of misclassifications	Misclassified carrots	Weeds classified as carrot	Classification rate
35	**7(2)**	**7**	**83.80%**

6.2.2 BEST WITH BACKWARD-ELIMINATION

When backward-elimination were to be done, red_mean, area_moment_3, area_moment_4, perimeter_moment_3 and perimeter_moment_4 had to be excluded because they were redundant and inaccuracies may occur from computing the inverse of almost non-singular covariance matrices. The following combination was found **{red_std, green_mean, blue_mean, blue_std, formfactor, elongatedness, solidity, perimeter_moment_1}**.

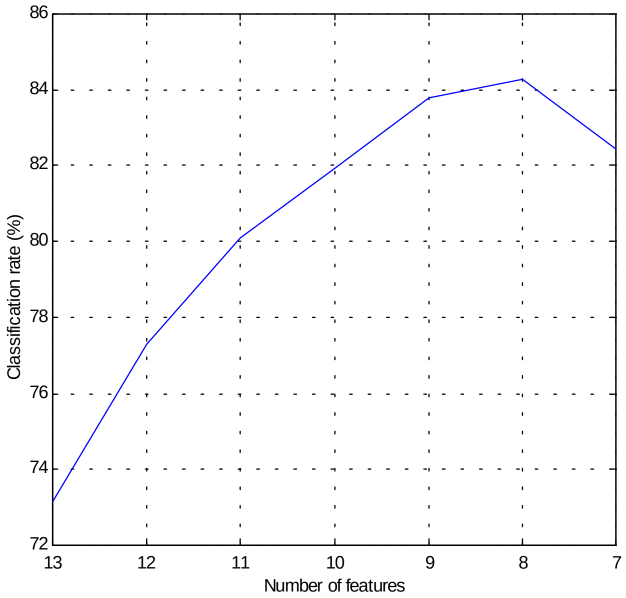


Figure 19. Classification rate increase as features are removed during backward-elimination. Best combination with eight features.

Table 12. Brief results. Backward-elimination.

Total number of misclassifications	Misclassified carrots	Weeds classified as carrot	Classification rate
34	**6(2)**	**8**	**84.26%**

6.2.3 BEST COMBINATION

The best combination found with the size dependent shape features excluded was: **{ g_mean, b_mean, area_moment_1, formfactor, b_std, perimeter_moment_1, elongatedness}**

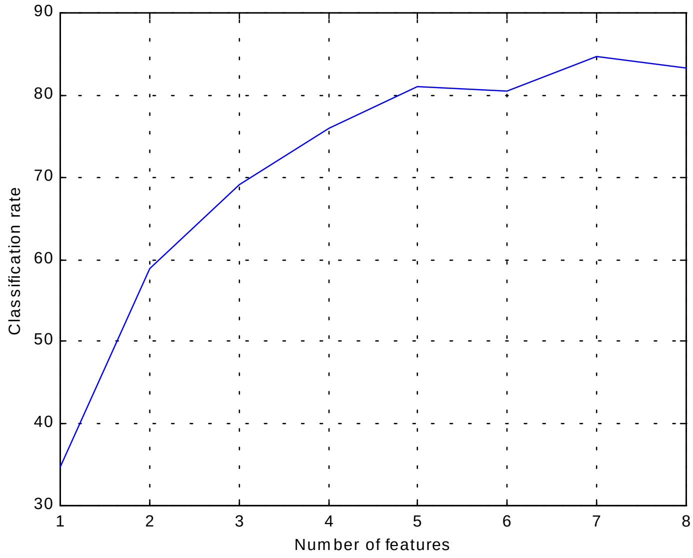


Figure 20. Plot shows how classification rate increase when features are added.

Table 13. Confusion matrix.

		1	2	3	4	5	6	7	8
1	Carrot	19(6)				5(2)	2(1)		1
2	Ladythumb Smartweed		25	1	1				
3	Canada thistle		1	22	4				
4	Lamb's Quarters		1	2	24				
5	Mayweed	3				22		2	
6	Corn Spurry	3					24		
7	Fumitory	2		1		1		22	1
8	Chickweed				1	1			25

Table 14. Brief results. **7 - best features**.

Total number of misclassifications	Misclassified carrots	Weeds classified as carrot	Classification rate
33	**8(3)**	**8**	**84.72%**

Figure 21 shows those carrots that were correctly classified, Figure 22 shows misclassified carrots and Figure 23 shows those weeds that were classified as carrot.

The 7 - best features were also tested with **resubstitution** instead of cross-validation. The following results were obtained.

Table 15. Confusion matrix.

		1	2	3	4	5	6	7	8
1	Carrot	20(6)				5(2)	2(1)		
2	Ladythumb Smartweed		26	1					
3	Canada thistle			26	1				
4	Lamb's Quarters				27				
5	Mayweed					25		2	
6	Corn Spurry	1					26		
7	Fumitory	1				1		24	1
8	Chickweed				1				26

Table 16. Brief results. **7 - best features** with resubstitution.

Total number of misclassifications	Misclassified carrots	Weeds classified as carrot	Classification rate
16	**7(3)**	**2**	**92.59%**

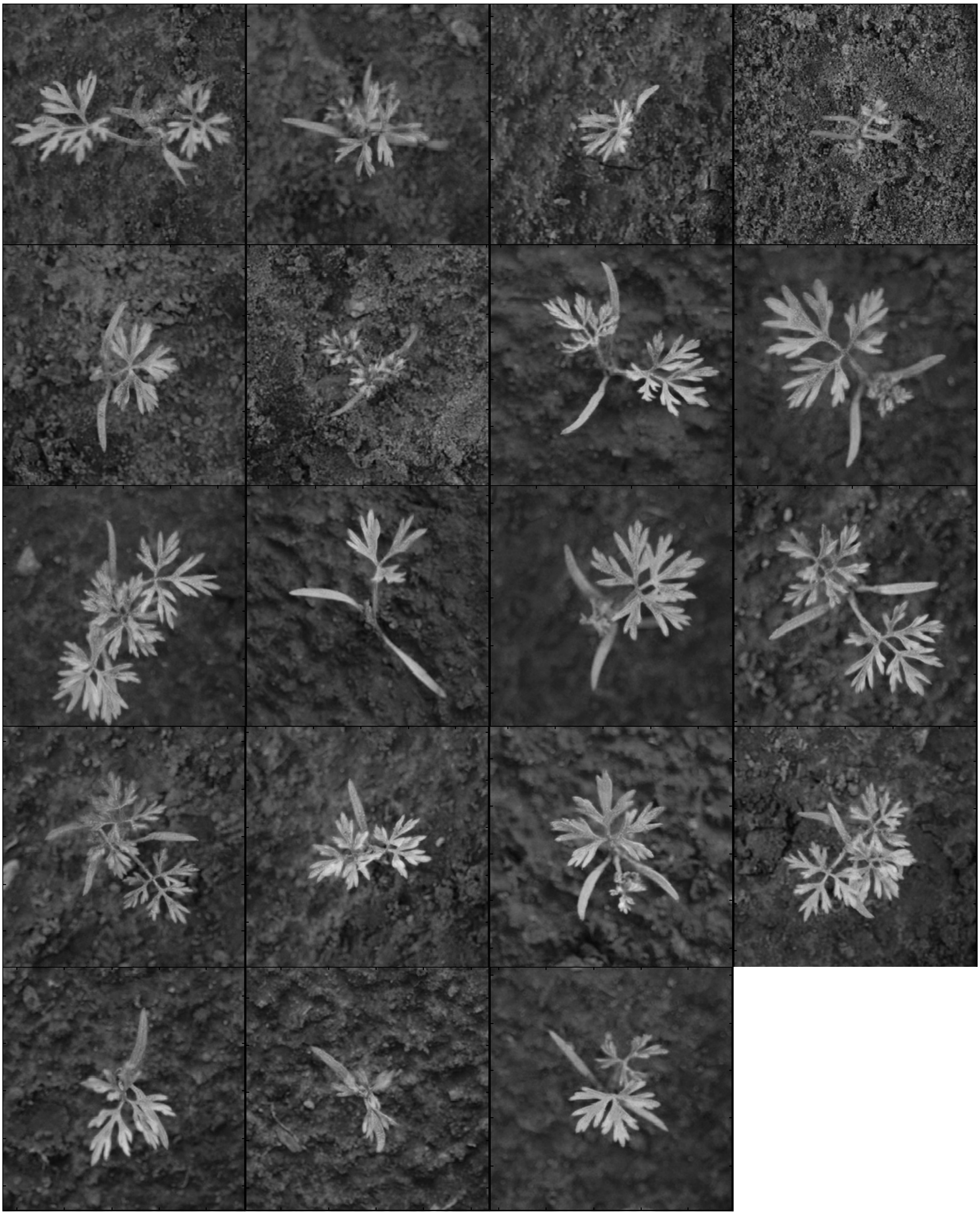


Figure 21. Correctly classified carrots.

Figure 22. Misclassified carrots.

Figure 23. Weeds classified as carrots.

a	b	c	d
e	f	g	h

(a) - (c) Mayweed; (d) - (f) Corn Spurry; (g) - (h) Fumitory.

6.3 CLASSIFICATION RESULTS BASED ON ALL FEATURES (SIZE DEP. INCL.)

This combination was found by manually adding and removing features. It includes eight features:

{ red_std, green_mean, blue_mean, blue_std, formfactor, convexity, convex_perimeter, thickness}

Table 17. Confusion matrix.

		1	2	3	4	5	6	7	8
1	Carrot	21(8)				1	4		1(1)
2	Ladythumb Smartweed		25	1					1
3	Canada thistle		1	24	1				1
4	Lamb's Quarters		1	1	24				1
5	Mayweed	2				21		4	
6	Corn Spurry	2					25		
7	Fumitory	1			1	3		20	2
8	Chickweed	1			2				24

Table 18. Brief results.

Total number of misclassifications	Misclassified carrots	Weeds classified as carrot	Classification rate
32	**6(1)**	**6**	**85.19%**

7 DISCUSSION

Critical assessment of classification rates

The most important classification is that between carrots and weeds. Crucial figures are then how many carrots are misclassified as weeds and how many weeds are misclassified as carrots. Less important are those weeds that are misclassified as another weed species. The feature combination that give the overall best classification rate does not necessarily give the best classification between carrots and weeds. The best feature combination found (size dependent features excluded) presented in 6.2.3 misclassified eight carrots and classified eight weeds as carrots, whereas, for instance, the combination found with forward-selection (6.2.1) misclassified only seven carrots and classified seven weeds as carrots but had a lower overall classification rate. A classification with only two groups, carrots and weeds, could give a different result.

If the figures in Table 11 would be interpreted as that a potential weeding machine would remove 7 out of 27 carrots and replace them with 7 weed plants then this would by far be much worse than what a human can do. Still this interpretation is very optimistic. This would suggest that better features of both shape and color have to be developed to increase classification rate.

Comparison with human weeders

Two weed species in this study are of special interest, because they are by human weeders sometimes mistaken as carrots. These two weed species are fumitory and mayweed, where fumitory is the one that cause most problems. Carrot and fumitory have similar shapes which can be seen by comparing Figure 3 (a) with Figure 4 (c). Classification based on only size independent shape features, presented in Table 4, show that as many as 12 fumitory are misclassified as carrots. To get the amateur weeders to learn to separate carrot and fumitory they are told to look at the difference in color. This we can do too by looking at the color features. If Table 8, confusion matrix of four color features, is studied it shows that there is no misclassification in either direction between carrot and fumitory. This shows that the automatic method developed here has the same problem and solution to the problem as humans when it comes to separating carrot and fumitory.

The second weed, mayweed, is not at all mistaken as carrot as often as fumitory, but it happens and by comparing Figure 3 (a) with Figure 4 (a) one can see that they, in a way, are quite similar in shape. Again Table 4, confusion matrix of size independent shape features, shows that 7 carrots are classified as mayweed and that 4 mayweed are classified as carrot. Table 6, confusion matrix of moment invariants, also show a misclassification in both direction. In this case there is no easy tip to give to the weeders and no feature group does entirely separate carrot and mayweed.

Finally when all features are combined, Table 13 shows, as could be expected, that a few carrots are classified as mayweed and fumitory and that some of these two weed species are classified as carrots. Beside this, Table 13 also shows that one carrot is classified as chickweed and three corn spurry are classified as carrot. This is rather surprising because chickweed and corn spurry are very different from carrot and human weeders never mistake them as carrots.

Two features of the convex hull were introduced, solidity and convexity, as an attempt to separate carrot and corn spurry. The hypothesis was that the long and narrow leaves of a corn spurry would span a great convex area whereas the actual area would be small. These two new features did however not solve the problem.

Images with two carrots

In the results, special notes have been made about those carrot images that contained two carrot plants. It could be expected that for the shape features these images would be overrepresented among the misclassified carrots while there should be no overrepresentation of them among the misclassifieds for color features. From Table 5 we have that 15 of 27 carrots are misclassified, 6 images of those 15 contain two carrots. This means that overall 56% (15/27) of the carrots are misclassified, 67% (6/9) of the two carrot images are misclassified and 50% (15-6)/(27-9) of the single carrot images are misclassified. These figures show that the two carrot images are slightly overrepresented among the misclassified carrots. The moment invariants also show an overrepresentation of two carrot images among the misclassifieds. While as expected the color features shows no uneven distribution.

This indicates that the shape features are sensitive to whether a plant can be entirely separated from all other plants even if the neighboring plants belong to the same species.

Plants covered with sand

Some plants were partly covered with sand. This was considered to have been caused by recurring showers that made the sandy soil bounce up on the leaves. The problem was most common with carrots and Canada thistles even though some other plants also had minor regions covered with sand. When the images were thresholded the plant parts covered with sand were lost since they were no longer green enough. No attempt was made to deal with this problem. For all feature combinations where all features are used, carrot is always the species that has the individually worst classification rate, this could be caused by bad segmentation.

The problem with plants covered with sand will most likely occur in any application and ways to deal with it must be found. It should be pointed out that a human observing the photographs of plants have no problem to tell if a sand region is actual soil background or plant covered with sand.

Choice of size dependent and size independent features

In section 6.2 the size dependent shape features, such as area and perimeter, have been excluded to make the classification less specific of the exact time when the images were taken. The carrots are sown at one time and germinate in a narrow time period while weeds germinate all the time, the size of the carrots could therefore be more compactly distributed than the size of the weeds. Including size dependent features could improve classification, but section 6.3 shows only a minor increase.

Sampling plan for choice of plants

The purpose of the image acquisition plan was to get pictures of plants where no subjective choices had been made concerning which plant to be selected. The shape of the selected area, Figure 1, makes the rows in the middle of the area longer than the rows at the edges. Since images only were acquired at one spot per every sixth row there are actually fewer spots per area in the middle of the field than at the edges. This means that, statistically the acquired images do not represent the field correctly, but this is not important since the aim was only to get images of plants with varying size and shape.

Applicability of methods

The methods developed in this study assumes that images where crops and weeds are mixed together can be segmented into individual plants. This assumption is highly questionable, for in the real images plants will partly cover each other making it impossible to see all plants entirely. This study suggests that simple features of shape and color can be used to classify between weeds and carrots quite well, but probably not good enough for any useful application. The features used here can hopefully be applied to do classification in the real type of images, but they can not be applied as simply and straight forward as in this study.

8 REFERENCES

Andreasen C, Rudemo M and Sevestre S (1997). Assessment of weed density at an early stage by use of image processing. *Weed Research* 37, 5-18.

Gonzalez RC and Woods RE (1993). *Digital Image Processing*. Addison-Wesley Publishing Company, Reading (Massachusetts), Menlo Park (California), New York, Don Mills (Ontario), Wokingham (England), Amsterdam, Bonn, Sidney, Singapore, Tokyo, Madrid, San Juan, Milan, Paris. ISBN 0-201-50803-6.

Jain AK (1989). *Fundametals of Digital Image Processing*. Prentice-Hall International, Inc. ISBN 0-13-332578-4.

Guyer DE, Miles GE, Schreiber MM, Mithell OR and Vanderbilt VC (1986). Machine Vision and Image Processing for Plant Identification. *American Transaction of the American Society of Agricultural Engineering* 29, 1500-07.

Paice MER, Miller PCH and Bodle JD (1995). An experimental sprayer for the spatially selective application of herbicides. *Journal of Agricultural Engineering Research* 60, 107-16.

Gerhards B, Nabout A, Sökefeld M, Mümbauch W and Nour Eldin HA (1993). Automatische Erkennung von Zehn Unkrautarten mit Hilfe digitaler Bildverarbeitung und Fourier-transformen. *Journal of Agronomy and Crop Science* 171, 321-8.

Petry W and Kühbauch W (1989). Automatisierte Unterscheidung von Unkrautarten nach Formparametern mit Hilfe der quantitativen Bildanalyse. *Journal of Agronomy and Crop Science* 163, 345-51.