Precision Agric (2006) 7:165–178
DOI 10.1007/s11119-006-9006-9

Identification of broad-leaved dock (*Rumex obtusifolius* L.) on grassland by means of digital image processing

Steffen Gebhardt · Jürgen Schellberg · Reiner Lock · Walter Kühbauch

Published online: 16 May 2006

Abstract Digital image processing has the potential to support the identification of plant species required for site-specific weed control in grassland swards. The present study focuses on the identification of one of the most invasive and persistent weed species on European grassland, the broad-leaved dock (*Rumex obtusifolius L.*, *R.o.*), in complex mixtures of perennial ryegrass with *R.o.* and other herbs.

A total of 108 digital photographs were obtained from a field experiment under constant recording geometry and illumination conditions. An object-oriented image classification was performed. Image segmentation was done by transforming the red, green, blue (RGB) colour images to greyscale intensity images. Based on that, local homogeneity images were calculated and a homogeneity threshold (0.97) was applied to derive binary images. Finally, morphological opening was performed. The remaining contiguous regions were considered to be objects. Features describing shape, colour and texture were calculated for each of these objects. A Maximum-likelihood classification was done to discriminate between the weed species. In addition, rank analysis was used to test how combinations of features influenced the classification result.

The detection rate of *R.o.* varied with the training dataset used for classification. Average *R.o.* detection rates ranged from 71 to 95% for the 108 images, which included more than 3,600 objects. Misclassifications of *R.o.* occurred mainly with *Plantago major* (*P.m.*). Between 9 and 16% *R.o.* objects were classified incorrectly as *P.m.* and 17–24% *P.m.* objects were misclassified as *R.o*. The classification result was influenced by the defined object classes (*R.o.*, *P.m.*, *T.o.*, soil, residue vs. *R.o.*, residue). For instance, classification rates were 86–91% and 65–82% for *R.o.* exclusively and *R.o.* against the remaining herb species, respectively.

Keywords Rumex obtusifolius · Weed detection · Digital image processing · Pattern classification · Precision farming

S. Gebhardt (✉) · J. Schellberg · R. Lock · W. Kühbauch
Institute of Crop Science and Resource Management—Crop Science and Plant Breeding, University of Bonn, Katzenburgweg 5, D 53115 Bonn, Germany
e-mail: s.gebhardt@uni-bonn.de

Introduction

The broad-leaved dock (*Rumex obtusifolius L.*) is one of the most harmful and invasive weeds in European grasslands because of its frequent occurrence, its large number of vital and germinable seeds, and rapid spread (Zaller, 2004). *Rumex obtusifolius* (*R.o.*) seeds persist in the rumen, as well as in slurry and farmyard manure, and remain viable throughout hard winters in the sward and soil. The availability of large concentrations of soil nutrients, especially nitrogen, promote the competitiveness of *R.o.* within the grass sward. Open spaces in damaged grazed swards improve the environmental conditions for germination and rapid establishment of *R.o.*, which favours its generative development and reproduction. Large dry matter contributions of *R.o.* reduce the average quality of the standing forage considerably because of the poor palatability of leaves and tillers.

Conventional techniques to control weeds, such as mechanical weeding or spraying, are laborious, time consuming and costly. Recently developed precision farming techniques based on weed mapping that use mainly image processing, enable site-specific spraying of weeds in arable crops (Dicke, Fries, & Gerhards, 2004; Gerhards, Sökefeld, Timmermann, Krohmann, & Kühbauch, 2000). Petry and Kühbauch (1989) and Sökefeld, Gerhards, and Kühbauch (2000) successfully discriminated between weed species using form parameters, such as Fourier descriptors and the *Shape Factor*. Furthermore, Marchant and Onyango (2003) separated cauliflower, weeds and soil using colour information and plant position. Burks, Shearer, Heath, and Donohue (2005) used the colour co-occurrence method (CCM) to discriminate weed species by textural features. On grassland however, these techniques have not been applied successfully. One of the major constraints of applying image processing to grassland is the diversity of optical plant properties within a mixed sward, such as leaf colour and shape, overlapping of leaves and tillers, shades on leaves and soil, as well as the non-uniform soil background. Compared to the identification of isolated individual plants on a moderately uniform soil background in arable crops, the processing of images from such a complex environment as grassland requires a different and more elaborate approach.

Previously there has been little research on the application of image processing to identify plant species in mixed grassland swards. Everitt et al. (1995) used colour and colour-infrared aerial photographs as well as conventional colour video imagery to detect leafy spurge (*Euphorbia esula*) at its flowering stage. The procedure was based on differences in the reflectance between flowers and vegetative organs. Pixel sizes ranged from 0.12×0.12 m to 0.5×0.5 m in the aerial images and from 1.3×1.3 m to 3.2×3.2 m in the video imagery. Glenn et al. (2005) used HyMap hyperspectral data with a 3.5 m pixel resolution to detect leafy spurge. They classified the images using the Mixture Tuned Matched Filtering algorithm. The overall accuracy of leafy spurge presence was between 84 and 94%.

The detection of individual weed plants is unlikely to be successful with far-range-remotely sensed image data because of inadequate spatial resolution. Therefore, image analysis has been introduced to near-range photographs. This technique has been applied to grassland under controlled conditions as a first step with the focus on identification of single plant species in artificially designed mixtures. For instance, Bonesmo, Kaspersen, and Bakken (2004) developed an image processing system to estimate the canopy cover of white clover in a grass–clover mixture, based on clover colour and its morphological properties. The efficiency of the system was tested in a field trial, exhibiting a regression coefficient of 0.82 with correctly identified vs. manually marked clover.

We propose that future practical applications of image processing of permanent grassland require systematic studies on plant discrimination in mixed swards. Therefore, the objective of the presented study was to develop algorithms within an image processing system to achieve a high recognition rate of *R.o.* in mixed permanent grassland swards, and to distinguish it from other accompanying herbs that are frequently associated with *R.o.* in grass swards, such as *Taraxacum officinale* (*T.o.*) and *Plantago major* (*P.m.*).

Materials and methods

Field experiment

A field experiment has been established on Grassland Research Station Rengen of the University of Bonn in the Eifel Mountains (Germany). Its location and climatic conditions are described by Schellberg, Möseler, Kühbauch, and Rademacher (1999). The experimental field enclosed a row of 36 plots; each plot was 2.7 × 2 m. On 18 of these plots sods of rye grass (*Lolium perenne* L.) were established without gaps, whereas on the other 18 plots mechanically added gaps were introduced showing the bare soil background. Together with *R.o.*, three more herb species, *Taraxacum officinale* Web. (*T.o.*), *Plantago major* L. (*P.m.*) and *Trifolium repens* L. (*T.r.*) were grown in the greenhouse and then planted in the experimental plots in May 2005. To enable the detection of non-overlapping plant individuals only, these species were planted in plots one to eight at distances far enough apart to avoid overlap. In contrast, for plots nine to 16 the above species were planted close enough to allow overlap. Each plot was planted with only one species, whereas in plots 17 and 18 a mixture of overlapping individuals of all the above-mentioned species was established. This method of planting the four species in plots one to 18 was repeated for the remaining 18 plots (19–36) including gaps.

Before the first image was recorded, all plots were cut to a height of 9 cm above the soil surface on July 11th, 2005. The images used subsequently for image processing were taken on days eight, 11 and 18 thereafter. The 36 plots were subdivided into two groups of 18 plots each with and without gaps. Hence, we analysed six datasets with 18 images each and with two datasets per recording date, giving a total of 108 images.

Image processing

To avoid any misunderstanding of the image processing procedure, a definition of the terms used is given in Table 1. We did an object oriented image classification, which is an application of pattern recognition in machine vision. The steps required for this are (1) image acquisition, (2) pre-processing and image segmentation, (3) feature extraction and (4) classification (Duda, Hart, & Stork, 2001; Gonzales & Woods, 1992). We did not classify single pixels, but instead image objects, which are contiguous regions of neighbouring pixels that were extracted in the image segmentation step. These objects were classified using their mathematical pattern, which is a vector of information comprising their geometry, colour and texture. In contrast, pixel-based image classification uses the spectral information in one or more spectral bands, and attempts to classify each pixel based only on this information. In the following paragraphs we describe the processing steps we performed.

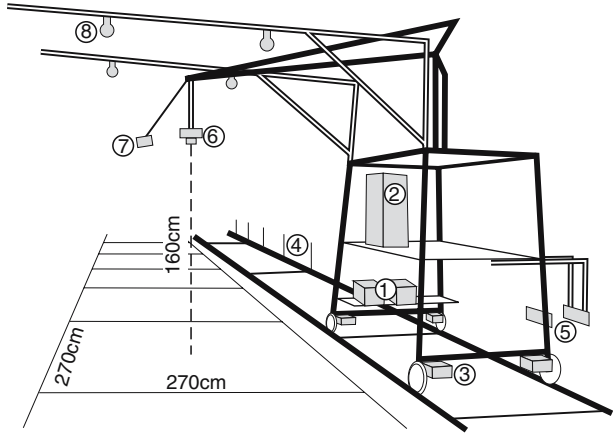
Table 1 Definition of frequently used terms in this paper

Term	Description
Data set	Record set of the images acquired per recording date
Object	Contiguous regions in the binary images derived by image segmentation
Feature vector	For every object derived by image segmentation 9 features were derived and stored into a feature vector
Feature matrix	$9 \times n$ matrix with n feature vectors, where n is the number of segmented objects per data set and 9 is the number of features
Training data set	A subset of the feature matrix for training the classifier which is the calculation of the distribution parameters for each class
Classification	MLE (maximum likelihood estimation) of the feature matrices, each matrix element is assigned to a class
Detection rate	Success of the classification measured as correctly detected objects given in percent of total object number
Total detection rate	Detection rate for $i = 5$ (*R.o.*, *P.m.*, *T.o.*, soil, residue) classification
R.o. detection rate	Detection rate of correctly classified *R.o.* objects
Misclassification	Rate of incorrectly classified objects

Image acquisition

The image recording was done in near range from about 1.60 m above ground by an autonomous vehicle (Fig. 1), driven on rails about 110 m long parallel to the row of plots. The vehicle was guided by electronic light barriers to stop automatically and to record images of every plot separately. The advantage of this system is that it is automatic and that the recording conditions are consistent, in particular the recording geometry.

We used a Canon PowerShot Pro1 digital camera with an image resolution of about 8 million pixels (3,264 × 2,448 pixels) above a ground area of 2.7 × 2 m. The pixel size was 0.6 × 0.6 mm at the nadir position. To avoid light reflectance and shadows in the plots



[Error Figure]

Fig. 1 Technical set up for image acquisition in the field. A specially constructed vehicle automatically and precisely accesses each experimental plot. The vehicle is power supplied by 12 V batteries (1). The electronic control box (2) contains the hardware and software to manage the movement, positioning and control of the vehicle. Four electronic motors (3) move the vehicle forward which stops automatically when

induced by direct sunlight, images were taken only under diffuse illumination conditions. Before each recording, a manual white balance was applied. To avoid loss of quality, the images were stored as uncompressed Canon RAW files to compact flash card and transformed to TIFF thereafter.

Pre-processing and image segmentation

All image processing steps were applied or developed within Mathworks MATLAB® using the MATLAB® Image Processing Toolbox. In a pre-processing step, the images were converted from 24-bit red, green, blue (RGB) images to 8-bit intensity images using standard formula, Eq. 1, which is the National Television System Committee (NTSC) standard for luminance,

$$I = 0.2989 \cdot R + 0.5870 \cdot G + 0.1140 \cdot B \tag{1}$$

where R, G and B is the pixel intensity in the red, green, and blue channel, respectively, and I is the intensity of the resulting grey scale pixel. The intensity depends strongly on the illumination conditions. However, constant measurement conditions in this experiment ensured nearly equal I values, thus enabling simultaneous processing of all images without illumination interference.

The aim of image segmentation is to locate certain objects of interest, which, in this case, were the leaves of the weeds. The procedure detects contiguous regions in an image by merging neighbouring pixels using a defined criterion. The texture of the leaves of *R.o.* and accompanying herbs were identified as more homogenous than those of grass. We used a measure of homogeneity to separate the herb and *R.o.* leaves from the background. The local homogeneity was calculated according to Cheng and Sun (2000). First, the standard deviation image of I was calculated by

$$S_{ij} = \sqrt{\frac{1}{n_w} \sum_{I_w \in W_d(P_{ij})} (I_w - m_{ij})^2} \tag{2}$$

S_{ij} is the standard deviation at the central pixel P_{ij}, m_{ij} is the mean of n_w intensities I within the window $W_d(P_{ij})$ of $d \times d$ pixel size with $d = 5$ in this case, n_w is the number of pixels within the 5×5 pixel window W. The gradient G_{xij} (in direction x) as well as G_{yij} (in direction y) was calculated as a measure of the change of grey values at pixel P_{ij} by applying the Sobel-operator Eqs. 3 and 4 to the intensity image I.

$$G_{x_{ij}} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \cdot W_d(P_{ij}) \tag{3}$$

$$G_{y_{ij}} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \cdot W_d(P_{ij}) \tag{4}$$

The gradient magnitude ∇f is calculated by Eq. 5 (Gonzales & Woods, 1992). The gradient is large where the maximum grey level range within the 3×3 window $W_d(P_{ij})$ is maximum and low where the grey levels are approximately the same.

$$\nabla f_{ij} = \sqrt{G_{x_{ij}}^2 + G_{y_{ij}}^2} \tag{5}$$

The local homogeneity H_{ij} is calculated using the maximum normalized gradient and standard deviation images $\nabla f_{ij}/\nabla f_{\max}$ and $S_{ij}/S_{\max}$, respectively.

$$H_{ij} = 1 - \frac{S_{ij}}{S_{\max}} \cdot \frac{\nabla f_{ij}}{\nabla f_{\max}} \tag{6}$$

By definition, the H_{ij} value ranges from zero to one. Pixels with a local homogeneity close to one were considered to be within homogeneous regions, and vice versa. A binary image B_{ij} was derived from the local homogeneous image by applying a threshold to H_{ij}

$$B_{ij} = \begin{cases} 0 & \text{if } H_{ij} < 0.97 \\ 1 & \text{otherwise} \end{cases} \tag{7}$$

A threshold value of $H_{ij} < 0.97$ was identified empirically and performed best in terms of detecting homogeneous regions. As a final image processing step, morphological opening (Gonzales & Woods, 1992) was performed on the binary image B_{ij} using Eq. 8 with a diamond structure element (S) of radius three pixels. The opening is an erosion $E(B_{ij},S)$ of image B_{ij} by S, followed by a dilation $D(E(B_{ij},S),S)$ of the result by S. The resulting binary image now consisted of several white objects (e.g. weeds) with grey level 1 and a background object with grey level 0.

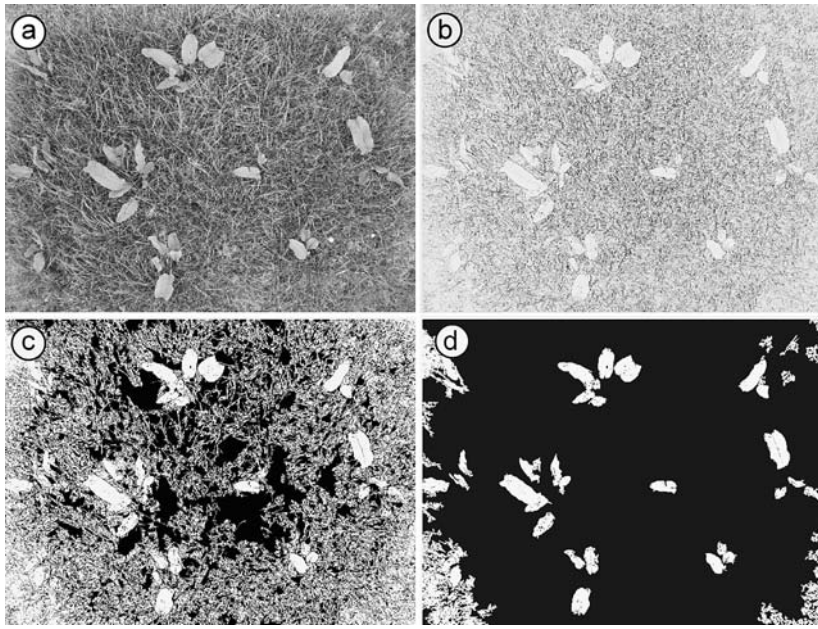
$$O(B_{ij}, S) = B_{ij} \circ S = D(E(B_{ij}, S), S) \tag{8}$$

As the leaf size of *R.o.* and the other weeds differed markedly from the background, an area threshold helped to discriminate them from smaller objects in the binary image, such as grass and clover. We applied an area threshold of 5,000 pixels which represents about 17 cm^2 on the ground. The leaves of *R.o.* covered larger areas than this threshold for all recording dates. The constancy of scale and geometry for all images allowed the area of the objects to be calculated based on the number of pixels within each object. Hence, a transformation of the image coordinate system to a metric one was unnecessary. An example of image segmentation is shown in Fig. 2.

The preceding image processing steps enabled the segmentation of leaves from the grass background. These leaves were represented as clusters (contiguous regions) of foreground pixels which are defined further on as objects (Fig. 2d).

Feature extraction

Feature extraction aims to describe the properties of objects mathematically. Specific features suitable for separating *R.o.* from the remaining object classes needed to be identified. Visual observations indicated that the leaf shape of *R.o.* differed from those of the other herbs. Consequently, we aimed to identify geometrical features describing these shapes (Du & Sun, 2004). The simplest features that could be extracted were *Area* and *Perimeter*. The *Shape Factor* gives information about the compactness of the shape of the object; the most compact shape is a circle (*Shape Factor* = 1). The *Shape Factor* decreases when the perimeter increases and the area remains constant, for example the leaves of *Taraxacum officinale* have a jagged outline. Another parameter describing compactness is



[Error Figure]

Fig. 2 Visualisation of the image segmentation procedure of grassland plots containing *Rumex obtusifolius* plants: (**a**) intensity image after conversion from RGB, (**b**) the local homogeneity image,

Circularity, which is 4π for a circle and which increases with increasing perimeter for a nearly constant area.

$$Shape\ Factor = \frac{4\pi \cdot Area}{Perimeter^2} \tag{9}$$

$$Circularity = \frac{Perimeter^2}{Area} \tag{10}$$

The *Eccentricity* was calculated from the ratio of the semi-minor and semi-major axes of the surrounding ellipse.

$$Eccentricity = \sqrt{1 - \frac{SemiMinor^2}{SemiMajor^2}} \tag{11}$$

The *Eccentricity* ranges from zero for a circle to almost one for a thin ellipse. In this approach, *Area* and *Perimeter* are not invariant in relation to scale, but we could disregard this as all images were recorded at the same scale. The geometric representation described here provided a feature matrix containing as many feature vectors as objects in the images, including the vector elements *Area*, *Perimeter*, *Shape Factor*, *Circularity* and *Eccentricity*.

Objects derived from image segmentation were not always leaves, but also homogeneous soil objects, which could not be described and discriminated from leaves by their shape alone. Therefore, colour features of the objects were calculated as well. The mean values of intensity (μ_I), as well as the mean grey values of red (μ_R) and green (μ_G) channels in the original RGB image were derived for each object using Eqs. 12–14. According to Soille (2000) the blue channel contributes little to discriminate soil from vegetation and so was disregarded.

Veins on the upper side of *R.o.* leaves and the herbs give rise to different degrees of surface roughness. Therefore, the mean value of the gradient ($\mu_{\nabla f}$) computed by Eq. 15 within an object was included in the feature extraction as a measure of the surface roughness.

$$\mu_I = \frac{1}{n}\sum_{i=1}^{n} I_i \tag{12}$$

$$\mu_G = \frac{1}{n}\sum_{i=1}^{n} G_i \tag{13}$$

$$\mu_R = \frac{1}{n}\sum_{i=1}^{n} R_i \tag{14}$$

$$\mu_{\nabla f} = \frac{1}{n}\sum_{i=1}^{n} \nabla f_i \tag{15}$$

where n is the number of pixels within an object, and I_i, G_i, R_i, ∇f_i are the values for intensity, green, red and gradient in an object at the position i, respectively. The feature matrix was extended by including the means of I, G, R and ∇f resulting in a total of nine vector elements. Six feature matrices were calculated from six available sets of data obtained for three recording dates.

Classification

A Maximum-likelihood estimation (MLE) (Stork, Yom-Tov, & Duda, 2004) was used to classify the objects into a defined number of classes ω_i, which was either $i = 5$ (*R.o.*, *T.o.*, *P.m.*, soil and residue) or $i = 2$ (*R.o.* and residue). The MLE is a parametric classification method. Each class ω_i is represented by a parameter vector θ_i. These parameters are calculated using a set of sample objects with known group membership called the training dataset. The calculation of these parameters is the training of the classifier.

There were marked morphological changes in the plant organs over time. Therefore, we wanted to determine whether the time span between establishing the training data set and obtaining the data to be classified had any influence on the result. The classifier was trained using subsets of the feature matrices obtained from data recorded on July 19th and July 22nd, and the corresponding classes for $i = 5$ and $i = 2$ were assigned manually. The first training data set had 209 objects (45 *R.o.*, 29 *T.o.*, 50 *P.m.*, 22 soil, 63 residue for $i = 5$ and 45 *R.o.* plus 164 residue for $i = 2$). We then created a new training data set using the

feature matrix of July, 22nd including 252 objects (55 *R.o.*, 37 *T.o.*, 43 *P.m.*, 39 soil, 78 residue). With these two training datasets two classifiers were created. The remaining objects, which were not included in the training data sets, were then classified separately using both classifiers. The probability densities in MLE are assumed to depend on θ_i, so that $p(X|\omega_i,\theta_i)$ is the probability of X belonging to class i. For any unknown object with feature vector X, MLE estimated the class that maximized the probability (Duda et al., 2001). To determine the error, the group membership of each object was labelled manually and compared with the classification results.

Feature evaluation

Finally we tested how single features of geometry, colour and texture influenced the classification results. Therefore, seven new training data sets were derived from the data recorded on July 22nd each comprising the following features and combination of features.

1. means of ∇f,
2. means of I, G and R,
3. eccentricity, circularity and shape factor,
4. area, perimeter, eccentricity, circularity and shape factor,
5. means of ∇f and means of I, G and R,
6. means of ∇f, means of I, G, R, eccentricity, circularity and shape factor,
7. all features.

Based on these training data sets, seven separate classifications were computed with the six feature matrices recorded on the three dates of image acquisition, plus one feature matrix that originated from merging the six entire feature matrices. The detection rates were stored in 7×7 matrices according to the seven training data sets and the seven classified feature matrices. As a result, four such matrices were derived, two matrices each for total detection rates or *R.o.* detection rates both for the classification into either $i = 5$ or $i = 2$ classes. A rank analysis was done on these four matrices to identify the training data set that performed best. As there were seven classifiers to be compared, the ranks ranged from one to seven. The classification with the highest detection rate was ranked one. The mean rank for all seven classifications was calculated for each of the four matrices, then the means of the ranks were summed. The classifier with the smallest sum of means was considered to be the best.

Results and discussion

The results of the classification based on the first training data set are given in Table 2. The total detection rate of correctly classified objects (*R.o.*, *T.o.*, *P.m.*, soil and residue) ranged from about 65 to 81% for $i = 5$, whereas for *R.o. vs.* residue it was between 71 and 87%, corresponding to between 66 and 121 objects. Detection appeared to be decreasing with time, indicating a change in phenology and morphology creating a change in geometric features.

Finally, the six feature matrices for the six data sets were merged to one feature matrix containing 3,861 objects, and then the data were classified again. Total detection rates and *R.o.* detection rates of 71 and 78% were determined, respectively. These results are confirmed by the confusion matrix of the classification given in Table 3. The largest misclassification

Table 2 Classification results for the classifiers trained with 207 objects from July 19th (*R.o.*: 45; *T.o.*: 29, *P.m.*: 50; soil: 22; residue: 63)

Data set	Total number of		Classification into					
			i = 5 groups[1]			i = 2 groups[2]		
			Detection rate (%)		Total number of *R.o.*	Detection rate (%)		Total number of *R.o.*
	Objects	*R.o.*	Total	*R.o.*		Total	*R.o.*	
July 19th—1	372	82	80.7	80.5	66	90.6	91.5	75
July 19th—2	635	128	76.4	86.7	111	90.9	92.2	118
July 22th—1	694	142	71.8	75.3	107	90.6	85.9	122
July 22th—2	780	146	70.5	82.9	121	87.4	87.7	128
July 28th—1	671	157	65.3	73.9	116	88.1	74.5	117
July 28th—2	709	127	66.4	70.9	90	89.7	77.9	99
All data sets	3,861	782	71.0	78.1	611	89.4	84.3	659

[1] *Rumex obtusifolius*, *Taraxacum officinale*, *Plantago major*, soil (gaps), residue

[2] *Rumex obtusifolius* and residue

Table 3 Confusion matrix indicating misclassifications for the classification of all 3,861 objects according to Table 2

			Predicted class					
			Residue	*R.o.*	*T.o.*	*P.m.*	Soil	Sum
Original class	Total	Residue	1,238	39	122	109	39	1,547
		R.o.	10	611	36	125	0	782
		T.o.	88	19	357	128	1	593
		P.m.	25	113	86	450	0	674
		Soil	168	8	1	2	86	265
	Percentage	Residue	80.0	2.5	7.9	7.0	2.5	100
		R.o.	1.3	78.1	4.6	16.0	0.0	100
		T.o.	14.8	3.2	60.2	21.6	0.2	100
		P.m.	3.7	16.8	12.8	66.8	0.0	100
		Soil	63.4	3.0	0.4	0.8	32.5	100

rates were for *R.o.* and *P.m.*, as 125 *R.o.* objects (16%) were assigned to *P.m.* and in turn 113 *P.m.* objects (17%) were assigned to *R.o*. Further misclassifications occurred between the classes residue and soil, with 168 out of 265 soil objects being classified incorrectly as residue. The misclassification rate between *T.o.* and *P.m.* was about 13 and 22%, respectively.

Reducing the number of classes to i = 2 increased the detection rate considerably (Table 2). The total detection rate was on average 90%, whereas the rate of *R.o.* detected correctly varied from 75 to 92%, again indicating a time dependent decrease.

The results of the same classification procedure obtained from the second and independent training set are given in Table 4. The total classification rate for the five classes of 67–82% was similar to those obtained before, whereas for *R.o.* the detection rate increased, ranging from 77 to 92%. For the classification into two groups (*R.o.* vs. residue) the total detection rate was similar to that for the July 19th training data set, whereas the *R.o.* detection rate increased to 81 and 95%, respectively. As before, all feature matrices were

Table 4 Classification results for the classifiers trained with 252 objects from July 22th (*R.o.*: 55; *T.o.*: 37, *P.m.*: 43; soil: 39; residue: 78)

Data set	Total number of		Classification into					
			i = 5 groups[1]			i = 2 groups[2]		
			Detection rate (%)		Total number of *R.o.*	Detection rate (%)		Total number of *R.o.*
	Objects	*R.o.*	Total	*R.o.*		Total	*R.o.*	
July 19th—1	581	127	81.8	92.1	117	86.4	91.3	116
July 19th—2	635	128	70.6	89.8	115	87.9	95.3	122
July 22th—1	442	87	73.3	85.1	74	88.5	90.8	79
July 22th—2	780	146	67.2	84.3	123	85.8	92.5	135
July 28th—1	671	157	70.2	77.1	121	91.2	80.9	127
July 28th—2	709	127	67.3	79.5	101	87.3	85.0	108
All data sets	3,818	727	70.6	84.0	611	88.0	89.0	647

[1] *Rumex obtusifolius*, *Taraxacum officinale*, *Plantago major*, soil (gaps), residue

[2] *Rumex obtusifolius* and residue

merged (3,818 objects) and classified. The corresponding confusion matrix is given in Table 5. There was still considerable misclassification for *R.o.* and *P.m.* (9 and 24%, respectively). Incorrect results for *P.m.* and *T.o.* were 16 and 21%, respectively, and for soil and residue it was 6 and 24%, respectively.

A comparison of Tables 2 and 4 shows that the detection rate of *R.o.* improved with the classifier trained with data from July 22nd compared to the July 19th training data set. When the classifier was trained with the data derived on July 19th (Table 2), the accuracy of the classification of the July 28th data was drastically reduced (75 and 78% *R.o.* detection rates). In contrast with the July 22nd training data set, the July 28th data showed much greater rates of detection for *R.o.* (81 and 85%). This might have been caused by visually observed changes in morphology of the plant leaves during nine days of growth and phenological development.

The classification results for the detection of *R.o.* in mixed grassland swards grown together with similar broad leaved weeds were surprisingly good. In view of the practical application of site-specific herbicide spraying or mechanical weeding, the required detection rate thresholds for *R.o.* cannot be determined yet, making an evaluation of the

Table 5 Confusion matrix indicating misclassifications for the classification of all 3,818 objects according to Table 4

			Predicted class					
			Residue	*R.o.*	*T.o.*	*P.m.*	Soil	Sum
Original class	Total	Residue	1,135	10	111	186	90	1,532
		R.o.	5	651	44	72	0	772
		T.o.	86	14	350	125	10	585
		P.m.	12	161	106	399	3	681
		Soil	59	1	2	2	184	248
	Percentage	Residue	74.1	0.7	7.2	12.1	5.9	100
		R.o.	0.6	84.3	5.7	9.3	0.0	100
		T.o.	14.7	2.4	59.8	21.4	1.7	100
		P.m.	1.8	23.6	15.6	58.6	0.4	100
		Soil	23.8	0.4	0.8	0.8	74.2	100

detection rates in Tables 2 and 4 difficult. In general, it can be argued that if we fail to identify individual *R.o.* plants during classification, it could have serious implications because of the large number of seedlings per plant and its large capacity for germination. Consequently, we aimed to obtain detection rates that were as good as possible. In this experimental approach, individual plants of *R.o.* consisted of more than one isolated object in the images. For example, with data set 2 for July 22nd, 121 objects related to *R.o.* were detected correctly out of about 60 *R.o.* plants in the field, indicating about two objects per plant on average. Consequently, the probability of plant identification must have been greater than the object identification rate of the same individual.

Three main reasons for misclassification were identified. Firstly, there were many errors in the residue class, with residue objects being assigned incorrectly. Such errors could arise as a result of the optical properties of the lens; objects cannot be projected to the image plane without radial symmetric distortions. This effect increases with the distance from the image centre. Hence, the geometry of peripheral weed leaves has been distorted and in addition, peripheral residue objects have been segmented incorrectly. Consequently, derived residue objects varied markedly in colour, texture and shape features. Some of the object features were similar to those of the herb and soil features in the respective training data set, hence causing misclassification.

Secondly, greater misclassification of *R.o.* occurred at later growth stages, indicating that differences in colour, texture and shape of *R.o.* leaves increased with plant age. Thirdly, *R.o.* and *P.m.* objects were assigned incorrectly to their respective classes because their features were unexpectedly similar. Clearly, the identification features of the two plant species used so far are not sufficiently typical for either class.

Influence of object features on the classification result

The rank analysis of the experimental data was used to identify those features of the object and their combinations that provided the best classification results. The average ranks for $i = 5$ and $i = 2$ of the total and *R.o.* specific detection rates are given in Table 6. The final column indicates the sum of these ranks. The lowest ranks indicate the best classification results for all classifications on average.

Table 6 The average ranks of the detection rates for the classifications of 7 data sets based on training data sets with different features

Training data set (feature combination)	Average ranks of detection rates of				Sum of ranks
	$i = 5$[1] classification		$i = 2$[2] classification		
	Total	*R.o.*	Total	*R.o.*	
1 (μ_f)	4.0	3.6	2.3	3.3	13.2
2 (μ_I, μ_R, μ_G)	5.7	4.9	6.4	5.3	22.3
3 (*Eccentricity*, *Circularity*, *Shape Factor*)	5.3	6.4	6.0	6.6	24.3
4 (*Eccentricity*, *Circularity*, *Shape Factor*, *Area*, *Perimeter*)	7.0	5.7	4.9	5.6	23.2
5 (μ_f, μ_I, μ_R, μ_G)	1.6	3.6	1.6	3.6	10.4
6 (μ_f, μ_I, μ_R, μ_G, *Eccentricity*, *Circularity*, *Shape Factor*)	2.7	1.9	3.0	2.1	9.7
7 (all features)	1.6	1.4	3.4	1.7	8.1

[1] *Rumex obtusifolius, Taraxacum officinale, Plantago major*, soil (gaps), residue

[2] *Rumex obtusifolius* and residue

Separately applied geometric, colour or textural features resulted in the highest ranks, and therefore the poorest classifications. For example, the classification into $i = 5$ groups using only geometry features (training data sets 2–4) resulted in total detection rates ranging from 40 to 60%. When using texture features exclusively, the detection rate was between 60 and 70%. Merging gradient and colour features (training data set 5) increased the total detection rate (65–82%). This training dataset led to the third best classification results in this analysis with a sum of average ranks of 10.4. The best results were obtained by combining all features (training dataset 7).

In general, the total detection rates of the three best ranked training data sets varied less than about 5%, but in particular cases the differences were between 10 and 20%.

Conclusions

The image processing system applied to digital images of *R.o.* and herbs within a grass sward provided surprisingly high rates of classification. The automatically driven vehicle enabled images to be recorded under constant geometric conditions in the field from which *R.o.* could be detected in mixed swards together with similar herbs and grass as the background.

We conclude that training datasets close to observation dates are required when classifying growing plants. The classification result was further determined by different feature combinations. Although object colour and geometry did not contribute strongly to the total classification rate, its integration into the classification procedure gave the best classification results.

The study shows that it should be possible to increase the total detection rates of specific plant types by eliminating known sources of error. The negative effect of distortions in the segmentation procedure can be eliminated by defining a confidence circle or ellipse around the image centre. Only those objects falling into this confidence area should then be considered for classification.

At present the image analysis is quite slow. Images were processed offline because the processing speed is too slow for online techniques in precision agriculture. Digital images require considerable computer memory. In the present study, most time elapsed in the conversion of RAW to TIFF images and in image segmentation, especially when calculating the local homogeneity image. Therefore, there is a need to study the effect of image compression and resolution on image analysis and classification results. We assume that the performance of the system could be increased by applying JPEG compression, which would reduce the size and loading time of the image considerably. A reduction in image resolution might also considerably reduce the time needed for image segmentation. For instance, by halving the image size from $2{,}448 \times 3{,}264$ pixels to $1{,}224 \times 1{,}632$ pixels would decrease the time for calculating the standard deviation image by a factor of 4.

The results of the present study support the hypothesis, that *R.o.* can be differentiated in images recorded under defined illumination conditions. The processing of GPS coded images would enable maps of weed distribution to be obtained supporting automatic site-specific spraying of herbicides. In arable crops, the economic and environmental benefit through savings of herbicides has already been demonstrated by similar applications (Dicke et al., 2004).

Acknowledgments The research was funded by the German Research Group (DFG), within the Research Training Group (Graduiertenkolleg) 722 ''Information Techniques for Precision Crop Protection'' at the University of Bonn (Germany). The authors gratefully acknowledge valuable input to the manuscript by an unknown reviewer.

References

Bonesmo, H., Kaspersen, K., & Bakken, A. K. (2004). Evaluating an image analysis system for mapping white clover pastures. *Acta Agriculturae Scandinavica, Section B–Soil and Plant Science, 54*, 76–82.

Burks, T. F., Shearer, S. A., Heath, J. R., & Donohue, K.D. (2005). Evaluation of neural-network classifiers for weed species discrimination. *Biosystems Engineering, 91*, 293–304.

Cheng, H. D., & Sun, Y. (2000). A hierarchical approach to color image segmentation using homogeneity. *IEEE Transactions on Image Processing, 9*, 2071–2082.

Dicke, D., Fries, A., & Gerhards, R. (2004). Determination of weed thresholds for site-specific weed control in malting barley. Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz—*Journal of Plant Diseases and Protection*, *Special Issue XIX*, 413–421.

Du, C. J., & Sun, D. W. (2004). Recent developments in the applications of image processing techniques for food quality evaluation. *Trends in Food Science and Technology, 15*, 230–249.

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification*. New York, USA: Wiley.

Everitt, J. H., Anderson, G. L., Escobar, D. E., Davis, M. R., Spencer, N. R., & Andrascik, R. J. (1995). Use of remote-sensing for detecting and mapping leafy spurge (Euphorbia-Esula). *Weed Technology, 9*, 599–609.

Gerhards, R., Sökefeld, M., Timmermann, C., Krohmann, P., & Kühbauch, W. 2000. Precision weed control—more than just saving herbicides. Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz—*Journal of Plant Diseases and Protection Special Issue XVII*, 179–186.

Glenn, N. F., Mundt, J. T., Weber, K. T., Prather, T. S., Lass, L. W., & Pettingill, J. (2005). Hyperspectral data processing for repeat detection of small infestations of leafy spurge. *Remote Sensing of Environment, 95*, 399–412.

Gonzales, R. C., & Woods, R. E. (1992). *Digital image processing*. Reading, MA, USA: Addison-Wesley.

Marchant, J. A., & Onyango, C. M. (2003). Comparison of a Bayesian classifier with a multilayer feed-forward neural network using the example of plant/weed/soil discrimination. *Computers and Electronics in Agriculture, 39*, 3–22.

Petry, W., & Kühbauch, W. (1989). Automatic distinction of weed species using form parameters by means of digital image processing. *Journal of Agronomy and Crop Science—Zeitschrift für Acker und Pflanzenbau, 163*, 345–351.

Schellberg, J., Möseler, B. M., Kühbauch, W., & Rademacher, I. F. (1999). Long-term effects of fertilizer on soil nutrient concentration, yield, forage quality and floristic composition of a hay meadow in the Eifel mountains, Germany. *Grass and Forage Science, 54*, 195–207.

Soille, P. (2000). Morphological image analysis applied to crop field mapping. *Image and Vision Computing, 18*, 1025–1032.

Sökefeld, M., Gerhards, R., & Kühbauch, W. (2000). Site-specific weed control—from weed recording to herbicide application. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz—Journal of Plant Diseases and Protection*, *Special Issue XVII*, 227–233.

Stork , D. G., Yom-Tov, E., & Duda, R. O. (2004). *Computer manual in MATLAB to accompany pattern classification*. Hoboken, USA: Wiley–Interscience.

Zaller, J. G. (2004). Ecology and non-chemical control of *Rumex crispus* and *R. obtusifolius* (Polygonaceae): A review. *Weed Research, 44*, 414–432.