Plant recognition and localization using context information and individual plant features

Björn Åstrand
Halmstad University
Box 823
S-301 18 Halmstad
SWEDEN
Bjorn.Astrand@ide.hh.se

Albert-Jan Baerveldt
Halmstad University
Box 823
S-301 18 Halmstad
SWEDEN
Albert-Jan.Baerveldt@ide.hh.se

Abstract – **In this paper we present a method that addresses the problem of recognition and localisation of plants in row-cultivated crops. As most crops are cultivated in rows and sown in a defined pattern, taking advantage of the geometrical properties of the scene may improve the results of crop recognition and localisation. The method proposed is to combine geometrical features of the scene (context) with individual plant features, e.g. colour and shape features. The method has been evaluated on two datasets of sugar beets at different stages of growth. Using an individual plant classifier (IPC), the best classification rate for datasets 1 and 2 was 89.0% and 94.9% respectively. Even though the classification rate is high, the drawback of using IPC is that it must be properly trained for each dataset separately and is thus sensitive to variations in plant appearance and weed species. The context method is more robust to these variations, and a classification rate of 78.4% was achieved for dataset 1. For dataset 2 the context method fails due to high weed pressure, up to 400 weeds/m^2. However, when it works properly, as in dataset 1, it is very robust to variations in crop appearance, although with one drawback. If the emergence of crops is low it will leave many weeds standing at crop positions. The idea of sequential combination of the classifiers is that the IPC removes as many weeds as possible, i.e. it works as a weed filter, while leaving possible crop candidates to the context method. We showed that sequential combination of the classifiers increases the overall classification rate, depending on which IPC used, by 3 to 8% for dataset 1 and 3 to 4% for dataset 2 compared to using IPC only. The classification rate was 91.9% for dataset 1 and 98.4% for dataset 2.**

I. INTRODUCTION

The increasing cost of chemicals and soil pollution caused by herbicide residues call for alternative methods of crop protection. A potential way to reduce chemicals is to use precision techniques for various types of agricultural operations, so that the chemicals can be placed where they have an optimal effect with a minimum quantity. For some operations it will even be possible to abandon the use of chemicals and apply other methods, e.g. mechanical weed control.

Our goal is to carry out intra-row (the area within the crop row, see Fig. 1 for definition) weed control. In the past decade, a great deal effort has been put into automatic guidance systems for inter-row guidance [1]. This has resulted in an improvement in the inter-row (the area between the rows) treatment, i.e. the untreated band could be reduced to a few centimetres. However, the challenging task is still intra-row cultivation. The problem of weed control in the seed line is to determine the crop positions. When the crop is located, a mechanical tool or a precision sprayer could be used to remove the weed.

As most crops are cultivated in rows and sown in a defined pattern, taking

Fig. 1. Dataset 1. Area between the lines defines the intra-row area. Crops are marked with a square and the ellipses are the size of the Gaussian bells (explained in section II.D).

advantage of the geometrical properties of the scene may improve the results of crop recognition and localisation. Knowing that the crops are sown in rows with a certain, constant distance between them, we showed in [2] that it is possible to recognize and locate the crops on the basis of this information instead of looking only at individual features of a plant, e.g. morphological or spectral properties.

If the emergence of crops is high and the weed pressure is low, the context method [2] is sufficient to recognize and locate the sowing pattern and is able to classify over 90 % of the crops. However, since crops do not have 100% emergence, some weeds will be classified as crops, when they grow at positions where the crop should have grown up, see Fig. 2, third ellipse. This problem grows with increasing weed pressure and lower emergence of crops. Increasing weed pressure also has the effect of making the crop pattern “disappear” among the weeds. This may lead to a collapse of the context method as the grid can not be found correctly. To address these two problems, the context method is here combined with classification methods based on individual plant features.

The aim of this paper is to present a vision-based plant recognition and localization algorithm that combines information about the scene structure, geometrical properties, with individual plant features, e.g. colour, shape and moments, to improve the recognition and localisation of plants.

Fig. 2. Dataset 2. Area between the lines defines the intra-row area. Crops are marked with a square and the ellipses are the size of the Gaussian bells (explained in section II.D). An example of a missing crop in the third ellipse, where the context method classifies the weed in the middle as crop.

II. METHODS AND MATERIALS

A. Data collection

Two sets of colour images were collected from different fields of sugar beets at two different stages of growth. Crops in dataset 1 are in the first true leaf stage, meaning that they have two pairs of leaves, see Fig. 1. For dataset 2, the crops are in the cotyledon stage, meaning that they have one pair of leaves, see Fig. 2. The images were collected using the weeding robot platform described in [3], see Fig. 3.

Changes in illumination cause colour shifts, and the colour camera is therefore mounted inside an enclosure to avoid natural light using halogen lamps only as the light source. The most important properties of each data set are given in Table 1.

Different camera systems with different image resolution were used for the two datasets. In dataset 1 the size of each image corresponds to the resolution of the camera. In dataset 2, high resolution images were generated by merging three images into one high resolution image. Each image covers about 80 and 60 cm in datasets 1 and 2 respectively.

There was no overlap between the images and the position (defined as where the stem meets the soil) of the crops was recorded manually in all images.

TABLE 1
PROPERTIES OF EACH DATASET

Property	*Dataset 1*	*Dataset 2*
No. images	143	54
Stage of growth	First true leaf	cotyledon
Image size (pixels)	768x576	1500x480
Image resolution (pixels/mm)	1.04	2.56
Weed pressure (weeds/m^2)	50	400
No. crops	574	196
No. weeds	414	828
Emergence (%)	71	73
Width of the intra-row area (mm)	82	65

Fig. 3. Autonomous robot for intra-row weed control.

B. Image Binarization

A linear discriminant in the normalized RGB colour space [4] was used to segment plant material from soil, see (1). Pixels that were too dark (intensity below 50) were classified as soil. The discriminant function was generated by randomly selecting pixels belonging to foliage and soil from a subset of images in the data sets. The parameters of the discriminate function for each data set are given in Table 2.

$$\left.\begin{array}{l} \mathrm{g_n} - \left(r_n * k_{discrim} + m_{discrim}\right) > 0 \\ AND \\ \dfrac{(R+G+B)}{3} > I_{Threshold} \end{array}\right\} \Rightarrow \text{Plant pixel} \tag{1}$$

where g_n = G/(R+G+B) and r_n = R/(R+G+B).

TABLE 2
DISCRIMINANT PARAMETERS

	$k_{discrim}$	$m_{discrim}$	$I_{Threshold}$
Data set 1	0.86	0.06	50
Data set 2	0.54	0.18	50

C. Segmentation of plants

After binarization, the next step is to segment pixels into regions with similar characteristics. The method used for grouping pixels into objects is the connected component labelling (CCL) algorithm [5]. Not all plants are grouped into one single object after CCL. One plant sometimes corresponds to more than one object. These objects are treated by an additional merging step. Some objects consist of both crop and weed. These are further referred to as CCL error. In this paper these objects are not further analysed and are treated as being a crop. As the two datasets have different characteristics, two different merging algorithms were applied. A very simple merging algorithm was used for dataset 1, which only merges objects that lay in a certain distance from each other. The reason that this simple algorithm is sufficient is that the weed pressure is relative low and the crops often correspond to one object after the CCL. The merging distance, the distance between object centres, was found experimentally and set to 27.6 mm. The segmentation results for dataset 1 are shown in Table 3, where the three crop classes, i.e. Correctly segmented, CCL error and Merging error, are shown separately.

TABLE 3.
MERGING RESULTS. TOTAL 457 CROPS AND 474 CROP OBJECTS AFTER MERGING.

Type	1. Correct segmented	2. CCL error	3. Merging error
N	378	23	56
%	83%	5%	12%

Dataset 2 required a more complex merging algorithm owing to higher weed pressure and the fact that only a few crops correspond to one object after segmentation. The merging algorithm used following merging criteria:

- The distance between object centres.

- The distance between objects' boundaries.
- A similarity measurement between the objects.
- Only two objects could be merged (most of the plants were in the cotyledon stage).
- The algorithm was tuned to merge crops, not weeds (error in weed merging is assumed not to influence the classification result).

The algorithm is based on three thresholds: distance between object centres, distance between object boundary and the number of objects to be merged with each other. The object and boundary distance was calculated from a subset of dataset 2, see Table 4.

TABLE 4.
MERGING DISTANCES AQUIRED FROM A SUBSET OF DATASET 2.

Merging distances	Average[mm]	Std	Min [mm]	Max [mm]
Boundary	2.8	1.12	0.87	4.96
Object centre	17.9	2.89	12.48	23.4

The similarity measure is the sum of the quotient between features. The features were experimentally chosen to be: *area*, *form factor*, *compactness* and *moment1* (for a definition of features, see appendix). The distance threshold was set to 23.6 mm and the boundary threshold to one third of the distance threshold. A similarity measure was then calculated for all objects that lay close to each other. The two most similar objects were then merged. The results are given in Table 5.

TABLE 5.
MERGING RESULTS. TOTAL 196 CROPS AND 209 CROP OBJECTS AFTER MERGING.

Type	1. Correct segmented	2. CCL error	3. Merging error
N	139	47	10
%	70.9%	24.0%	5.1%

For dataset 2, the CCL error also led to a merging error in only three cases. This means that even if a weed, or part of a weed, was segmented together with a crop, the whole crop was still correctly merged.

A final remark is that neither of these algorithms is claimed to be optimal and require further investigations. The classification results for these three crop categories, i.e. Correctly segmented (Crop 1), CCL error (Crop 2) and Merging error (Crop 3), will be presented separately in this paper.

D. Context method

Knowing that the crops are sown in rows with a certain, constant distance between them, we showed in [2] that it is possible to recognize and locate the crops on the basis of this information instead of looking only at individual features of a plant, e.g. morphological or spectral properties. The intra-row distance between crops follows a Gaussian distribution [6], see Fig. 4. The distribution of the intra-crop distance was tested and proven to be normally distributed with D'Agostino's normality test with a 95% confidence level. Under the assumption of normality, the crop row can be modelled as a set of Gaussian bells [2], see Fig. 5, where the distance between the bells is the mean inter-crop distance, μ_d, with the standard deviation, s_d, and the standard deviation of the alignment error of the crops relative the crop row, s_e. It is important to note that the variance of the intra-crop distance, s_d^2, does not increase

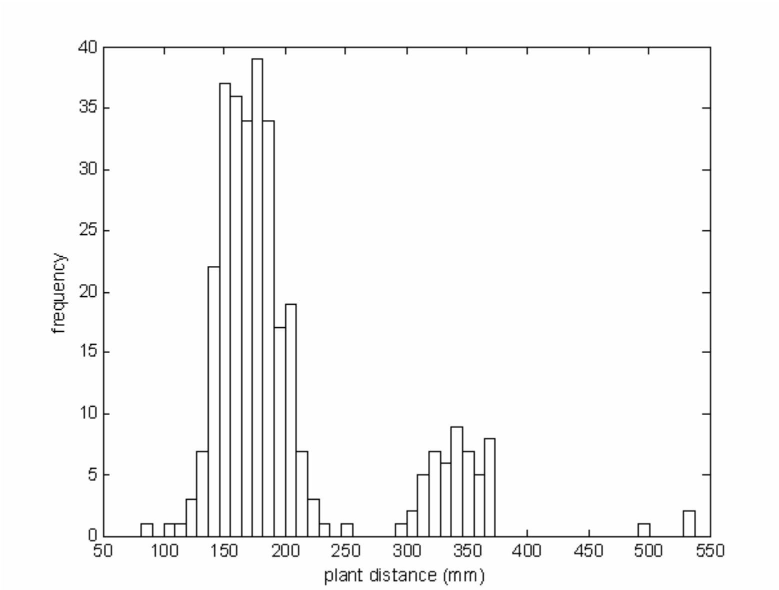

Fig. 4. Crop distance follows a Gaussian distribution. The second peek is due to missing crops.

with an increasing number of bells, i.e. the crop distances are assumed not to be a random walk (an effect caused by a new seed being sown when the sowing machine has travelled a certain distance). This is true for small numbers of N, where the variance of the intra-crop distance dominates by imprecise positioning of the crop seed when it is dropped into the seeding furrow and field conditions influence where crops emerge relative to their seed position [7].

The conditional probability density function for the crop positions (x, y), given crop grid position g_i, and the assumed sown position of this crop on the grid, denoted by the bell index n, are defined as:

$$p(x, y \mid g_i, n) = \frac{1}{2\pi\sqrt{\sigma_d^2 \sigma_e^2}} \exp\left\{ -\frac{1}{2}\left[\frac{(x - n\mu_d + g_i)^2}{\sigma_d^2} + \frac{(y - \mu_e)^2}{\sigma_e^2} \right] \right\} \qquad (2)$$

where n is the bell index [1..N] and N is the number of Gaussian bells. x is the position along the row and y the position perpendicular to the row. μ_e is the row position and is assumed to be known. The grid position, g_i, describes the placement of the crop row model along the row. It ranges from zero to plant distance μ_d and is discrete according to the image pixels.

To find the most probably grid position given the positions of the crops we use Bayes' theorem [8]. This gives:

$$P(g_i \mid (x_1, y_1), \ldots, (x_N, y_N)) = \frac{\left(\sum_{n=1}^{N} p_n(x, y \mid g_i, n) \right) P(g_i)}{\sum_{i=0}^{\mu_d} \left(\sum_{n=1}^{N} p_n(x, y \mid g_i, n) \right) P(g_i)} \qquad (3)$$

The value of function (3) can be calculated by placing the model at every possible grid position g_i for the actual image. In the ideal case, with one crop around each sow position and no weed, this is straightforward. In the presence of a weed, however we

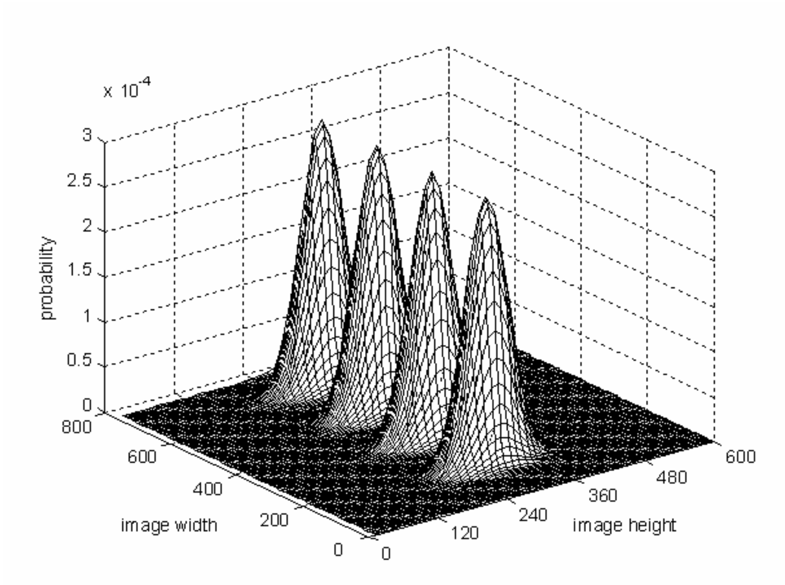

Fig.5. Gaussian bells that correspond to average plant distance and alignment error of the plant row. N = 4.

have to choose which plant contributes in each bell, i.e. $p_n(x,y|g_i,n)$, as only one crop is assumed in each bell. With no other apriori information about the plant than position, the plant that is most likely to be a crop based on the position is chosen, i.e. the highest $p_n(x,y|g_i,n)$ within each bell n. If no plant is found within the 99.7 % area (3 sigma) of a bell n, see the ellipses drawn in Figs. 1 and 2, a missing crop is assumed and $p_n(x,y|g_i,n)$ is set to zero. With no additional information about the apriori probability of the crop grid position, $P(g_i)$, it is assumed to be a uniform distribution. However, if the distance travelled between consecutive samples of the row structures (images) is known, the apriori probability for the grid position can be estimated. By maximizing (3) we obtain the most likely position of the crop. Since the denominator is the same for all possible positions of the crop row, it need not to be calculated and (3) can be reduced to:

$$GridPosition = \max_{g_i}\left(\left(\sum_{n=1}^{N} p_n(x, y \mid g_i, n)\right)P(g_i)\right) \tag{4}$$

i.e. the grid is estimated to be at the grid position with highest posteriori probability. Once the grid is found, the plants that contribute in each bell at this position are classified as crops.

As mentioned, in each Gaussian bell, it is assumed that only one plant can be a crop candidate and a crop candidate is chosen by the highest position probability. However, if further classification is applied to the plant, i.e. if colour and shape features are included, the probability for a plant to be a crop or weed on the basis of these features can be combined with the position probability before calculating the most likely position of the crop grid. This way of combining classifiers is further discussed in section II.G.

In (2) it is assumed that the bells can be placed exactly over the crop row. However, in a real situation, i.e. when driving along the row, the exact position of the row structure is not known because of the movement of the camera, see Fig. 1. This has the greatest impact on the lateral position of the row, while the error caused by heading is negligible. To solve this, (2) can either be extended by including the standard deviation of the movement of the camera (5) or get the row offset and heading from a guidance system. In this paper we use the former approach.

A second extension to (2) is to add the difference between the true crop positions and the estimated crop position. The crop position is defined as the place at which the stem meets the soil. However, this position is often difficult to estimate by computer vision, and the centre of the boundary-box around the plant foliage is therefore used to estimate the plant position. The difference between the estimated position and the true position was included in (2). The extensions for (2) are defined as:

$$\begin{aligned} s_{d'}^2 &= s_d^2 + s_{pc}^2 \\ s_{e'}^2 &= s_e^2 + s_{pc}^2 + s_{lm}^2 \end{aligned} \tag{5}$$

where s_{pc} is the standard deviation between the estimated crop position and the true crop position and s_{lm} is the standard deviation of the lateral movement of the camera. It is assumed that these errors are uncorrelated. The values used in this paper for (2) are listed in Table 6.

TABLE 6.
CONTEXT METHOD PARAMETERS FOR EACH DATA SET

Dataset	μ_d (mm)	$\sigma_{d'}$ (mm)	$\sigma_{e'}$ (mm)	N
1	170.8	24.0	13.6	5
2	118.8	14.8	10.8	5

E. Individual plant features

Using different individual plant features such as colour, shape and moments to distinguish between plants species is a well known approach and has been employed by a number of authors [8-14]. The advantage is that they are relative easy to calculate and good classification results are often achieved, often over 90%. In [3] we presented a study of classification of crops and weeds using 19 different features, six colour features, seven "shape" features and six moment-based features, all defined in the appendix. We showed that colour is an important feature for classification and, on the database used in [3], we achieved a classification rate of 96% using three features only. In this paper we apply the same features to our two new datasets.

F. Classification methods and feature selection

Three different types of classifiers were evaluated: Gaussian quadratic, k-Nearest Neighbour (kNN) and an Artificial Neural Network (Multi Layer Perceptron, MLP) classifier. For the MPL we use one hidden layer and the number of hidden nodes found by cross-validation. We use a single output node with a logistic sigmoid activation function, which allows us to interpret the output (classification result) as posterior probabilities [15]. The kNN classifier was evaluated with different numbers of neighbours, i.e. 1-, 3-, 5-, 7-, 9-NN, using a Euclidean distance measure [16].

The five-fold cross-validation scheme was used to train and evaluate the classifiers. To find the feature combination that gives the highest classification rate we used forward inclusion and backward elimination. For the MLP, each feature combination was trained five times with random initialisations of weights, and the best "net" was then used. This is done to avoid local minima. The MLP was trained using the Levenberg-Marquardt algorithm [17].

We used the *t*-test to see whether the architectures were significantly different with a 95% confidence interval according to (6):

$$\left|m_i - m_j\right| > 2.306\frac{s_{ij}}{\sqrt{K}} \qquad (6)$$
$$s_{ij}^2 = s_i^2 + s_j^2$$

where K is the number of folds in the cross-validation and μ_i, s_i and μ_j, s_j are the average and standard deviation of the *i* and *j* classifiers.

We do normalisation by "standardization" where all variables are standardized to have zero mean and unit variance, see (7)

$$z_k(n) = \frac{x_k(n) - m_k}{s_k} \qquad (7)$$

Only correctly segmented crops were used to train the classifiers (class 1 crops). The number of weeds used was reduced so that the number of weeds and crops was about the same.

G. Combining context and plant features

If different classifiers offer complementary information about the classes, combining these classifiers may increase the correct classification rate [18-19]. The goal is to find a scheme of combining these classifiers so that the complementary information can be harnessed.

In this paper we use two ways of combining the information from the context method and the individual plant feature classifiers (IPC): one parallel and one sequential. Depending on the output from the classifiers there are different ways of combining them. In our case, the output from the classifiers, the estimation of the posterior probability of being a crop $P(crop|IPC_{features})$, is made binary by a threshold: 0 for weed and 1 for crop.

The sequential combination is an extension of the context classifier described in section II.D. The problem with the context method is that it fails to locate the grid at higher weed pressures. Therefore, the idea in the sequential combination is that the individual plant classifier (IPC) removes as many weeds as possible, i.e. it works as a weed filter while leaving possible crop candidates to the context method. As mentioned above, the output of the classifier is made binary by a threshold for being a crop candidate. The reason for doing this is that we want to remove as many weeds as possible, because each weed that is not removed interferes with the grid matching of the context method. Removing as many weeds as possible, with a minimum reduction of crops, boosts the performance of the context classifier.

A second reason for using a threshold for the plant being a crop candidate is the following. If a weed is standing at the crop position because of low emergence, as in Fig. 2, where the third Gaussian bell (marked as an ellipse) contains a single weed at the crop position, the context method will classify it as a crop if a threshold for being a crop candidate is not applied. However, applying a threshold for being a crop candidate after the context classification will most likely also remove some crops, see example Fig. 6, where weed No. 1 is better positioned than the crop and thus has a higher total probability of being a crop, $P_{total}(crop) = P(crop|IPC_{features})*P(crop|x,y)$.

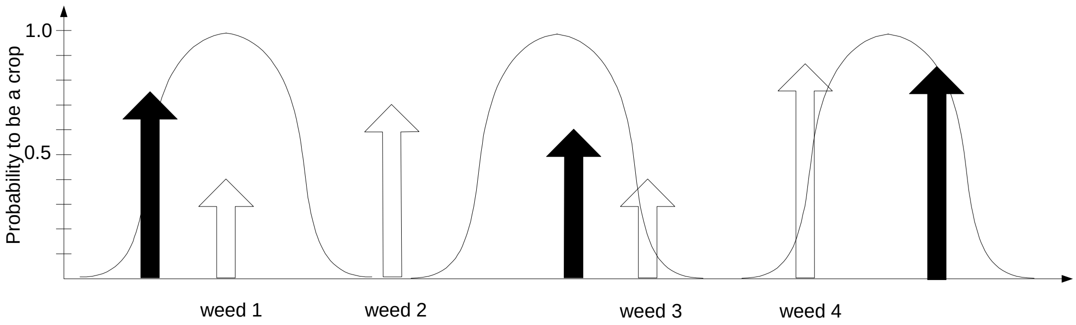

Fig. 6. In the first bell a weed (white) is better placed than a crop (black) and selected to be a crop by the context method even if it has a lower prior probability to be a crop estimated by IPC, $P(crop|IPC_{features})<0.5$. This is avoided by removing all plants with a low probability of being a crop.

There, $P(crop|x,y)$ is the position probability estimated by (3) when the grid is found, i.e. $p(x,y|g_{GridPosition})$. This example corresponds to the situation in the fourth Gaussian bell in Fig. 1 where a weed is better placed than a crop. For these two reasons, boosting the performance of the context classifier and reducing the risk of removing crops by applying a threshold for being a crop candidate after the context method, we use a threshold for being a crop candidate on the IPC output.

Using a threshold on the IPC output has the consequence that a crop classified as a weed by IPC can thus not be correctly classified by the context method. This kind of error made by IPC can not be corrected by the context method. However, a weed wrongly classified as a crop by IPC can be corrected by the context method if the weed is not located on the crop grid, like weed No. 2 in Fig. 6, or if a real crop is better placed than the weed, like weed No. 4 in Fig. 6.

For reasons of comparison, we also investigate a parallel combination. In the parallel case each classifier uses its own representation of the input pattern, i.e. the context method uses plant position and the plant feature classifier uses the individual plant feature of each plant. We use a majority vote schema to combine the classifiers and implement them as an AND-operator.

III. RESULTS

A. Context method

The classification results for the context method are shown in Tables 7 and 8. The results of crop classification are shown for each crop class, defined in section II.C. When the context method is used all plants have the same probability of being a crop. This means that the method is sensitive to high weed pressure and the emergence of crops. The emergence is about 70% in both datasets, but the weed pressure is much higher in dataset 2. In dataset 1 94% of the crops were found, as compared to 61% in dataset 2. The reason for this is that the high weed pressure makes the crop structure disappear among the weeds, i.e. the grid is not found. Table 9 lists the number of correct grid matches. The definition of a grid match is when a majority of the actual crops in the grid is found. If the image contains one or two crops, all crops must be found for a definition of a grid match and, if the image contain three or more crops, only one crop can be missed for a definition of a grid match, see Table 9. In dataset 2 only half of the matches occur as compared to 92% for dataset 1. It is evident from the results in Table 8 that the context method does not work for higher weed pressures. However, for lower weed pressures, as in dataset 1, it works with one major drawback; many weeds are still left, about 40%. The remaining weeds are mostly located at crop positions, see Table 10, missing crop, where the crop did not grow (weed at missing crop). That only 10% of the weeds are classified as crops is caused by a weed in a better position than a crop (crop/weed error), see Fig 1. Half of these crop/weed errors are caused by a failure in grid matching. The grid match error is also the cause of one third of the crop errors (Table 10). The rest are due to a crop/weed error, as mentioned above.

To conclude, the context method is a robust method when the weed pressure is low and/or the emergence is high.

TABLE 7.
RESULT OF CONTEXT METHOD DATASET 1.

Type		Classified as Crop	Classified as Weed	Classification rate
	1	355	23	93.9%
Crop	2	19	4	82.6%
	3	50	6	89.3%
Weed		148	266	64,3%
Crop 1 + weed				*78.4%*

TABLE 8.
RESULT OF CONTEXT METHOD DATASET 2.

Type		Classified as Crop	Classified as Weed	Classification rate
	1	85	54	61.1 %
Crop	2	27	23	54.0 %
	3	8	2	80.0 %
Weed		135	693	83.7 %
Crop 1 + weed				*80.5 %*

TABLE 9.
GRID FOUND DEFINED AS THE NUMBER OF CROPS FOUND COMPARED TO THE ACTUAL NUMBER OF CROPS.

Grid Match	1/1	2/2	2/3	3/4	4/5	Total
Dataset 1	12/17	19/20	38/40	45/48	18/18	132/143
Dataset 2	1/1	1/8	8/13	12/20	6/12	28/54

TABLE 10.
WEED AND CROP ERROR ANALYSIS OF DATASET 1 AND CLASS 1 CROP.

	Total	Caused by grid error
Weed at missing crop	133	21
Crop/weed error	15	7
Crop error (class 1)	23	8

B. Individual plant features

There were no large differences between the different architectures of the MLP and kNN. In most cases there were no significant differences between the number of hidden nodes used or the number of neighbours used, even if the number of features used differs in some cases. However, the best classifier for each classifier type was chosen so that redundant features were removed, thus selecting the feature combination that performs significantly best with a minimum number of features. The results of the classifiers chosen, one classifier from each classifier type, are shown in Tables 11 and 12. For dataset 1 (Table 11) best classifiers are the MLP 4 (four hidden nodes) and the 5-NN. The Gaussian classifier let most of the crops remain while only removing 65% of the weeds, as compared to the other classifiers, which remove almost 90% of the weeds. The Gaussian classifier shows similar results as the context method in Table 7. For dataset 2 there are no significant differences between classifiers.

TABLE 11.
RESULTS USING PLANT FEATURES DATASET 1.

Classifier [Features]	Type		Classified as crop	Classified as weed	Class. rate
Gauss [7,1,5,13,8]	Crop	1	354	24	93.6%
		2	20	4	83.3%
		3	49	23	68.1%
	Weed		144	270	62.5%
Crop 1 + weed					*78.8%*
5-NN [7,3,6,11,15]	Crop	1	329	49	87.0%
		2	20	4	83.3%
		3	49	23	68.1%
	Weed		52	362	87.4%
Crop 1 + weed					*87.2%*
MLP 4 [7,3,13,11,14,8]	Crop	1	340	38	89.9%
		2	19	5	79.2%
		3	44	28	61.1%
	Weed		49	365	88.2%
Crop 1 + weed					*89.0%*

TABLE 12.
RESULTS USING PLANT FEATURES DATASET 2.

Classifier [Features]	Type		Classified as crop	Classified as weed	Class. rate
Gauss [7,4,11,14]	Crop	1	133	6	95.7%
		2	44	6	88.0%
		3	14	6	70.0%
	Weed		39	789	95.3%
Crop 1 + weed					*95.3%*
5-NN [3,6,12,11,15]	Crop	1	132	7	95.0%
		2	37	13	74.0%
		3	5	15	25.0%
	Weed		42	786	94.9%
Crop 1 + weed					94.9%
MLP 2 [7,1,11,14,19]	Crop	1	135	3	97.1%
		2	38	12	76.0%
		3	11	9	55.0%
	Weed		61	767	92.6%
Crop 1 + weed					*93.4%*

We use in total 19 features (six colour, seven shape, six moments), see Appendix. The most common features used for both datasets and all classifiers are *area* and *solidity*, both in five of six cases. The *area* can be explained by the fact that there are often many small weeds, given the average weed size to be smaller than crops. The single classification rate for *area* and *solidity* is 73.2% and 58.7% for dataset 1 and 91.2% and 65.4% for dataset 2, respectively. As we pointed out in [3] colour is an important feature and, here, at least one of the colour features is used in all the selections of classifiers. Using colour alone, a classification rate of 80% was reached for both classifiers.

The features used by at least one classifier for each datasets are *green mean* and *red* *mean*, *blue* STD, *area*, *solidity*, *moment1* and *moment2*. To conclude, using individual plant features, a high correct classification rate can be achieved, for datasets 1 and 2 89.0% and 94.9% respectively. Even if the individual plant classifier achieves a high correct classification rate, it has a drawback: it must be properly trained to achieve these results. Variations in plant appearance in and between fields could easily reduce the performance of a classifier trained offline. Thus, classifiers trained offline can be enhanced by online adaptation.

C. Combining of context and plant features

Both parallel AND and sequential combination were tested for dataset 1. The context method was unable to locate the grid for dataset 2 because of high weed pressure, and the results of the AND combining for dataset 2 are thus not shown. The results of the combinations are shown in Tables 13 and 14. Even for dataset 1 (Table 13) it is shown that the sequential combination of classifiers is a slightly better approach than the parallel AND combination, although there is an improvement of 1-2% using parallel AND over using IPC alone, see Tables 11 and 13. With sequential combination, the classifiers increase the overall classification rate, depending on which IPC is used, by 3 to 8% for dataset 1 and 3 to 4% for dataset 2, as compared to using IPC only. The best combined classifier for dataset 1 is the MLP classifier, with a classification rate of 91.9%. For dataset 2, the best combined classifier is the 5-NN classifier, with a classification rate of 98.4%.

After combining, there is less difference between the classifiers as compared to the results in the case using the IPC only. The reason for this is that the feature classifiers perform differently in the classification of weeds. When the context classifier is added, a further reduction of weeds is achieved, and the difference thus becomes smaller.

The grids found, i.e. the number of grid matches, increase if the number of crops is detected by IPC. The number of times the correct grid is found is a crucial performance measure of the context method and the sequential combination. Tables 15 and 16 show the number of grids found and the number of grids found for images containing three crops or more (in parentheses). The grids found are about 90% in most cases and increase by 2-3% if only images with three or more crops are included. Compared to the context method only, there is a small performance drop in grids found for dataset 1. This is because some crops are classified as weeds by IPC, causing the grid match to fail. The drop in performance is greater if all images are included as compared to counting images that contain three crops or more. Still, errors caused by a failure in grid match are fewer than those caused by the IPC. Higher emergence of crops and more correctly segmented crops will reduce these grid errors even further.

Tables 15 and 16 also give an error analysis of the crops and weeds that are wrongly classified after the sequential combination. In most cases, weeds are classified as crops because of a missing crop (weed at missing crop), see Fig 2. Only a few misclassifications are interchanges between crops and weeds when a weed is better positioned in the grid (crop/weed error). There is no crop/weed error for dataset 2, which shows the advantage of using context to recognize and locate crops and weeds. The context method removes 50-80% of the weeds while it leaves most of the weeds that are at a position where a crop should have grown up. Weeds that remain are thus standing at a position where they interfere least with the crops from a nutriment point of view. The crop errors in Tables 15 and 16 are crop/weed errors as mentioned above, and errors caused by a misplaced crop. The majority of the crop error is caused by the IPC classifying a crop as a weed.

TABLE 13.
RESULTS USING COMBINING CLASSIFIERS DATASET 1.

Classifier	Type		Sequential C.C[a]	Sequential C.W	Sequential R. %	AND C.C	AND C.W	AND R. %
Gauss	Crop	1	346	32	91.5	336	42	88.9
		2	17	7	70.8	16	8	66.7
		3	38	34	52.8	35	37	48.6
	Weed		68	346	83.6	59	355	85.7
Crop 1 + weed (IPC only: *78.8*)					*87.4*			*87.2*
5-NN	Crop	1	324	54	85.7	312	66	82.5
		2	18	6	75.0	17	7	70.8
		3	42	30	58.3	38	34	52.8
	Weed		24	390	94.2	19	395	95.4
Crop 1 + weed (IPC only: *87.2*)					*90.1*			*89.3*
MLP 4	Crop	1	337	41	89.1	322	56	85.2
		2	16	8	66.7	15	9	62.5
		3	34	38	47.2	32	40	44.4
	Weed		23	391	94.4	21	393	94.9
Crop 1 + weed (IPC only: *89.0*)					*91.9*			*90.3*

[a]C.C = Classified as crop, C.W = Classified as weed, R. = Class rate.

TABLE 14.
RESULTS USING COMBINING CLASSIFIERS DATASET 2.

Classifier	Type		Sequential Classified as crop	Sequential Classified as weed	Sequential Class rate
Gauss	Crop	1	131	8	94.2%
		2	41	9	82.0%
		3	9	11	45.0%
	Weed		10	818	98.8%
Crop 1 + weed (IPC only: *95.3 %*)					*98.1%*
5-NN	Crop	1	131	8	94.2%
		2	37	13	74.0%
		3	5	15	25.0%
	Weed		7	821	99.2%
Crop 1 + weed (IPC only: *94.9 %*)					*98.4%*
MLP 2	Crop	1	130	9	93.5%
		2	38	12	76.0%
		3	8	12	40.0%
	Weed		17	811	97.9%
Crop 1 + weed (IPC only: *93.4 %*)					*97.3%*

TABLE 15.
WEED AND CROP ERROR ANALYSIS DATASET 1 (CLASS 1 CROP).

	Error type	Total	Caused by grid error	Grid found
Gauss	Weed at missing crop	50	13	127/143 (98/106)[a]
	Crop/weed error	7	9	
	Crop error	32	5	
5-NN	Weed at missing crop	21	3	117/143 (87/106)[a]
	Crop/weed error	3	2	
	Crop error	51	4	
MPL4	Weed at missing crop	19	2	126/143 (99/106)[a]
	Crop/weed error	4	2	
	Crop error	37	2	

[a] Counting only images containing three crops or more.

TABLE 16.
WEED AND CROP ERROR ANALYSIS DATASET 2 (CLASS 1 CROP).

	Error type	Total	Caused by grid error	Grid found
Gauss	Weed at missing crop	10	3	44/54 (42/45)[a]
	Crop/weed error	0	0	
	Crop error	8	1	
5-NN	Weed at missing crop	7	1	49/54 (41/45)[a]
	Crop/weed error	0	0	
	Crop error	8	0	
MLP 2	Weed at missing crop	17	4	50/54 (42/45)[a]
	Crop/weed error	0	0	
	Crop error	9	2	

[a] Counting only images containing three crops or more.

IV. DISCUSSION

The three main factors that influence the performance of the recognition and localisation of crops are the emergence of the crops, weed pressure and the stage of growth of crops. The emergence of crops has an impact only on the context method, while the two others have an impact on both classifiers. High weed pressure causes more plants to grow together and makes the segmentation more difficult, i.e. getting correctly segmented plants. It is difficult or even impossible to match in the context method because the crop grid "disappears" among the weeds. If crops are larger, the segmentation becomes less complex, as in dataset 1, and with less complexity there are fewer errors, i.e. larger plants are often segmented into one object while smaller plants often consist of multiple objects that must be merged.

When the classifiers are combined in sequence, the context method is able to handle higher weed pressures in our case up to 400 weeds/m^2. However, this depends on how successful classification of the individual plant classifier is. For both datasets the k-NN and MLP show an advantage over the Gaussian classifier. The drawback of the individual plant feature method is that the classifier must be trained before use. Variations in plant size and weed species within and between fields and during the season call for some type of unsupervised or reinforcement learning to make classifiers adaptive to these variations. The context method is more robust to these variations and could therefore be used to train the feature classifier.

The performance of the context method also depends on how accurately the crop row is followed. As pointed out in paper [2], compensating for these errors can improve the classification rate by 2% for dataset 1, as an example. If the precision in sowing is increased the Gaussian bells could be smaller and thus increase the robustness against high weed pressure and reduce the probability that weeds will be accepted at missing crop positions. Further improvement can be achieved by increasing the number of bells used for grid matching [2]. The number of bells used is limited by the fact that the crop distances can vary as a result of imprecise sowing causing shifts in the crop pattern that disturb the matching. More bells also require longer initialisation, i.e. a longer distance must be travelled before a match can be achieved. An alternative approach toward enhancing the performance of the grid matching is to track the grid by estimating the distance travelled between consecutive images (samples of the row structure) and thus achieving an estimate of the expected crop position P(g_i), see (3). A tracking approach is suggested by [20], and these authors used an extended Kalman filter to track the model of the crop pattern.

Combining classifiers in sequence shows a clear advantage over the parallel AND approach. The reason for this is that the individual plant classifier boosts the performance of the context classifier. The main drawback of the suggested combining schema is that the performance depends on how well the feature classifier recognizes crops. However, when the weed pressure is low, the feature classifier could be tuned to accept more plants as crops, see Table 17, at the cost that more weeds are classified as potential crops. In the example in Tables 17 and 18, the tuning is achieved by changing the degree of confidence of a 5-NN classifier, i.e. the number of neighbours to be classified as a crop. 20% confidence means that least one of the five neighbours must be a crop to be classified as a crop. For dataset 1, the confidence should be set to 40%, which implies a 4% increase in crops found compared to use a confidence of 60%, which is used in Section III. For higher weed pressures, as in dataset 2, it may be necessary to remove as many weeds as possible to boost the context classifier. As seen in Table 18, reducing the confidence to 40% does not increase the number of crops found. For dataset 2, a 60% confidence is recommended. The confidence can not be reduced and an increase in confidence does not significantly decrease the number of remaining weeds.

TABLE 17
PERFORMANCE OF 5-NN CLASSFIER WITH DIFFERENT DEGREES OF CONFIDENCE (DATASET 1).

Degree of confidence %		0[a]	20	40	60	80	100
Plant	Crop	378	359	343	329	285	209
Feature	Weed left	414	188	105	52	25	10
weed/crop ratio		*1.10*	*0.52*	*0.31*	*0.16*	*0.09*	*0.05*
Sequence	Crop	355	348	337	324	284	209
	Weed left	148	81	48	24	12	4
weed/crop ratio		*0.42*	*0.23*	*0.14*	*0.07*	*0.04*	*0.02*

[a] 0% confidence is analogous to the context method.

TABLE 18
PERFORMANCE OF 5-NN CLASSFIER WITH DIFFERENT DEGREES OF CONFIDENCE (DATASET 2).

Degree of confidence %		0[a]	20	40	60	80	100
Plant	Crop	139	134	132	132	132	125
Feature	Weed left	828	214	105	42	22	9
weed/crop ratio		*5.96*	*1.60*	*0.80*	*0.32*	*0.16*	*0.07*
Sequence	Crop	78	115	126	131	131	124
	Weed left	151	56	25	7	5	4
weed/crop ratio		*1.94*	*0.49*	*0.20*	*0.05*	*0.04*	*0.03*

[a] 0% confidence is analogous to the context method.

V. CONCLUSIONS AND FUTURE WORK

In this paper we present a method that addresses the problem of recognition and localisation of plants in row-cultivated crops. The proposed method combines geometrical features of the scene (context) with individual plant features. The advantage of using context, i.e. the crop pattern, is that this feature is stable to within and between field variations in crop appearance and weed species. The drawback is that, if the emergence is low, there will be weeds standing at crop positions that will be treated as crop. Another drawback is that when the weed pressure is high the crop pattern will disappear among the weeds, which causes the context method to fail. To address these two problems we combine the context classifier with a classifier that uses individual plant features, e.g. colour and shape features.

The methods have been evaluated on two datasets of sugar beets at different stages of growth. We tested three different types of classifiers for individual plant features:

Gaussian quadratic, k-Nearest Neighbour (kNN) and an Artificial Neural Network (Multi Layer Perceptron, MLP) classifier. For dataset 1 the k-NN and MLP show an advantage over the Gaussian classifier, and the best classification rate for datasets 1 and 2 was 89.0% and 94.9% respectively. The most important features were *area* and *solidity*, and at least one colour feature was used in all selected classifiers.

We have shown that a sequential combination of geometrical features and individual plant features is preferable to a majority vote combining (in this case a logical AND). The idea behind the sequential combination is that the individual plant classifier removes as many weeds as possible, i.e. it works as a weed filter, while leaving possible crop candidates to the context method. Combining the classifiers increases the overall classification rate, depending on which individual plant feature classifier is used, by 3-8% for dataset 1 and 3-4% for dataset 2 as compared to using individual plant features only. The classification rate for the combining was at best 91.9% for dataset 1 and 98.4% for dataset 2. The advantage of combining the context method with individual plant features is that the context method can handle higher weed pressures, in this case up to 400 weeds/m^2. Another advantage of the proposed method is that most of the weeds not recognized are located where a crop should have grown, and are thus located at positions where they interfere least with the crops from a nutriment point of view.

A problem that increases with higher weed pressures is the increasing number of occluded plants, causing more segmentation errors. This problem is not fully addressed in this paper and will be a subject for further work. Further work will also investigate the number of Gaussian bells used by the context method to match the crop grid pattern. A preliminary investigation shows that adding more bells would give a further increase of robustness against higher weed pressures. The algorithms will also be implemented on the weeding robot described [3] and evaluated in the field.

Even if the individual plant classifier achieves high classification results, it has a drawback: it must be properly trained to achieve these results. Variations in plant appearance within and between fields could easily reduce the performance of a classifier trained offline. The long term goal is to make the method adaptive, meaning that all parameters necessary for successful discrimination between crops and weeds should be automatically identified and updated. This will make the method more robust to within field and between field variations in the weed pressure and in plant appearance. The algorithm presented in this paper is a first step toward reaching that goal in the sense that the planting geometry is less affected by those variations.

VI. ACKNOWLEDGMENT

This work was carried out in the Mech-Weed project. The project was sponsored by FORMAS: the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning. The project is currently sponsored by SJV, the Swedish Board of Agriculture, and SLF, Stiftelsen Lantbruksforskning – the Swedish Farmers' Foundation for Agricultural Research.

VII. REFERENCES

[1] B. Åstrand, *Mechatronics in Agriculture - Row Guidance System for Agricultural Field Machinery*, Licentiate thesis, School of Electrical and Computer Engineering, Chalmers University of Technology, Göteborg,

Sweden, October 2000.

[2] B. Åstrand and A.-J. Baerveldt, Plant recognition and localization using context information, *Mechatronics and Robotics 2004 – special session Autonomous Machines in Agriculture*, Aachen, Germany, September 13-15, pp. 1191-1196, (2004).

[3] B. Åstrand and A.-J. Baerveldt, An agricultural mobile robot with vision-based perception for mechanical weed control, *Autonomous Robots*, vol. 13, 2002, pp 21-35.

[4] H.D. Cheng, X.H. Jiang, Y.Sun, Jingli Wang, Color image segmentation: advances and prospects, *Pattern Recognition*, vol. 34, 2001, pp.2259-2281.

[5] Shapiro, L.G., and R.M Haralick, *Computer and Robot Vision volumes I and II*, Addison-Wesley Publishing Company Inc., 1992-1993.

[6] J. Bontsema, C.J van Asselt, P.W.J. Lempens and G. van Straten, Intra-row weed control: A mechatronic approach, *1st IFAC Workshop on Control Applications and Ergonomics in Agriculture*, Athens, Greece, June 15-17, 1998.

[7] H.-W. Griepentrog, M. Nørremark, H. Nielsen, S. Blackmore, Individual plant care in cropping systems, *4th European Conference on Precision Agriculture*, Berlin, Germany, July 2003.

[8] C.M Onyango, J.A. Marchant, Segmentation of row crop plants from weeds using colour and morphology, *Computers and Electronics in Agriculture*, vol. 39, 2003, pp. 141-155.

[9] D. M. Woebbecke, G. E. Meyer, K. von Bargen, D. A. Mortensen, Shape features for identifying young weeds using image analysis, *Transactions of ASAE*, vol. 38(1), 1995, pp271-281.

[10] W. S. Lee, D. C. Slaughter, D. K. Giles, Robotic Weed Control System for Tomatoes, *Precision Agriculture*, 1, 1999, pp. 95-113.

[11] J. Hemming, T. Rath, Computer-vision-based weed identification under field conditions using controlled lighting, *Journal of Agricultural Engineering Research*, 78(3), 2001, pp. 233-243.

[12] L. Tian, D. C. Slaughter and R. Norris. (1997) Outdoor field machine vision identification of tomato seedlings for automated weed control. *Transactions of the ASAE*, Vol. 40 (6): 1761-1768

[13] A. J. Pérez, F. López, J. V. Benlloch, S. Christensen, Colour and shape analysis techniques for weed detection in cereal fields, *Computers and Electronics in Agriculture*, 25, 2000, pp.197-212.

[14] S. I. Cho, D. S. Lee, J. Y. Jeong, Weed-plant discrimination by machine vision

and artificial neural network, *Biosystems Engineering*, 83(3), 2002, pp. 275-280.

[15] C.M. Bishop, 1997, *Neural Networks for Pattern Recognition*, Oxford University Press Inc., New York.

[16] R.O. Duda, P.E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.

[17] M. T. Hagan, M. Menhaj, Training multilayer networks with the Marquardt algorithm, *IEEE Trans Neural Networks*, 6 (5), 1994, pp. 989-993.

[18] J. Kittler, M. Hatef, R. P. W. Duin, J. Matas, On combining classifiers, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, No. 3, March 1998, pp. 226-239.

[19] L.I. Kuncheva, Combining classifiers: soft computing solutions, In: S.K. Pal and A. Pal (Eds.), *Pattern Recognition: From Classical to Modern Approaches*, World Scientific, 2001, pp 427-451.

[20] B. Southall, J. A. Marchant, T. Hague, and B. F. Buxton, Model based tracking for navigation and segmentation, In: *Proceedings of the Fifth European Conference on Computer Vision*, Freiburg, Germany, June 2-6, 1998, pp. 797-811.

APPENDIX

The definition of *form factor* is that it is a measure of how much "plant mass" there is in the centre in relation to how much "plant mass" there is in the periphery.

$$MEAN_{dist} = \frac{1}{N}\sum\sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}$$

where N is the total number of object pixels and x_c and y_c are the geometrical centre of the objects, defined below.

$$x_c = \frac{1}{N}\sum_{i=0}^{N-1} x_i \qquad y_c = \frac{1}{N}\sum_{i=0}^{N-1} y_i$$

$$VAR_{dist} = \frac{1}{N-1}\sum\left[\sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} - MEAN_{dist}\right]^2$$

$$formfactor = \frac{MEAN_{dist}}{\sqrt{VAR_{dist}}}$$

TABLE A1.
LIST OF FEATURES.

#	Name	Description
1	green mean	The mean value, over the whole plant, of the normalised green colour, g = G/(R+G+B).
2	green std	The standard deviation, over the whole plant, of the normalised green colour.
3	red mean	The mean value, over the whole plant, of the normalised red colour, r = R/(R+G+B).
4	red std	The standard deviation, over the whole plant, of the normalised red colour.
5	blue mean	The mean value, over the whole plant, of the normalised blue colour, b = B/(R+G+B).
6	blue std	The standard deviation, over the whole plant, of the normalised blue colour.
7	area	Area is defined as the number of pixels belonging to the plant
8	perimeter	Perimeter is defined as the number of pixels of the plant boundary.
9	compactness	area/perimeter^2
10	elongation	area/thickness^2, where thickness is defined as the number of shrinking steps of an object until only one pixel is left in the image.
11	solidity	area/(area of convex hull), where convex hull is described as the area formed if a rubber band would be tighten around the object.
12	formfactor	See the definition above in this appendix.
13	convexity	perimeter/(perimeter of convex hull)
14	moment1	These are functions of moments, which are invariant to geometric transformations such as translation, scaling and rotation. Defined in (Jain, 1989).
15	moment2	
16	moment3	
17	moment4	
18	moment5	
19	moment6	