Available online at www.sciencedirect.com

SCIENCE DIRECT®

Mechatronics 15 (2005) 251–269

MECHATRONICS

A vision based row-following system for agricultural field machinery

Björn Åstrand *, Albert-Jan Baerveldt

School of Information Science, Computer and Electrical Engineering, Halmstad University, P.O. Box 823, S-301 18 Halmstad, Sweden

Accepted 4 May 2004

Abstract

In the future, mobile robots will most probably navigate through the fields autonomously to perform different kind of agricultural operations. As most crops are cultivated in rows, an important step towards this long-term goal is the development of a row-recognition system, which will allow a robot to accurately follow a row of plants. In this paper we describe a new method for robust recognition of plant rows based on the Hough transform. Our method adapts to the size of plants, is able to fuse information coming from two rows or more and is very robust against the presence of many weeds. The accuracy of the position estimation relative to the row proved to be good with a standard deviation between 0.6 and 1.2 cm depending on the plant size. The system has been tested on both an inter-row cultivator and a mobile robot. Extensive field tests have showed that the system is sufficiently accurate and fast to control the cultivator and the mobile robot in a closed-loop fashion with a standard deviation of the position of 2.7 and 2.3 cm, respectively. The vision system is also able to detect exceptional situations by itself, for example the occurrence of the end of a row.

Keywords: Agricultural robots; Row-following; Vision-guidance; Crop-row location; Hough transform

* Corresponding author. Fax: +46 35 12 03 48.
E-mail address: bjorn.astrand@ide.hh.se (B. Åstrand).

0957-4158/$ - see front matter
doi:10.1016/j.mechatronics.2004.05.005

1. Introduction

Increasing cost of chemicals and the soil pollution caused by the herbicide residues ask for alternative methods of crop protection. A potential way of reduction of chemicals is the use of precision techniques for various types of agricultural operations, so that the chemicals can be placed where they have an optimal effect with 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. There is political interest in the European Union to increase the amount of ecologically grown products. The goal is that about 5–10% of the total field area should be processed by organic farming methods by the year 2005. Organic farming is not only a political goal; there is also a need from the market. More and more customers are asking for products that are organically grown. This has led to a problem for companies that need to increase their supplies of organically grown products to meet customer demands. For example, it is difficult to extend the amount of organically grown sugar beets at present, because weed control in the seedline of sugar beet is done by human labour, which implies high costs and difficulties in recruiting workers. The motivation for the work reported here is to reduce the amount of herbicides used for crop protection in agriculture by replacing chemical weed control by mechanical weed control. Our vision is that in the future, autonomous mobile robots will navigate through the fields to perform these operations. As most crops are cultivated in rows, an important step towards this long-term goal is the development of a row-recognition system, which will allow a robot to accurately follow a row of plants.

The goal of this paper is to present a new method for robust recognition of plant rows based on the Hough transform that is able to guide agricultural machines, such as inter-row cultivators or mobile robots. The main contribution of the paper is on the perception part of a row-following system. To demonstrate the capabilities of our row recognition method, the method was implemented as part of a row-following system both on an inter-row cultivator and on a mobile robot. In Section 2 we describe our approach and briefly describe previous reported work and motivate the reason for development of a new algorithm for row-recognition. In Section 3 we report the result from three different experiments with the system. And finally, in Section 4 we preset a method for exception handling, e.g. detecting the end of a row.

2. Approach

During the last two decades progress in vision-based navigation has been made on two separate fronts: indoor navigation and outdoor navigation [1]. In outdoor vehicle navigation there has been a large interest in road following using the road lane as guidance signal. In the agricultural area vision-based navigation has been used in guidance of, for example, a harvester by tracking the crop boundary [2]. As most crops are cultivated in rows there have been a number of publications on deriving guidance signals from plant structures [3–7].

In row guidance the problem is to find a guidance signal that is accurate and stable under different environmental conditions. Using a guidance signal derived from living materials has its difficulties. The rows can be incomplete, i.e. there can be missing plants and plants can be at different stages of growth (size) along the field. Another problem is the fact that there might be a lot of weed which disturbs the appearance of a nice clear row structure of green plant material coming from the crops. These factors make the recognition task of the row structure much more difficult.

Marchant [3] and Marchant and Brivot [4] performed several steps of image processing based on near-infrared images to discriminate crops from weeds. They calculated the center of each crop in the image and used Hough transform to calculate the offset and orientation relative to the row structures. They combined information coming from three row segments to increase performance. Billingsley and Schoenfish [5] used the chrominance signal from a color camera to acquire an image representing the "greenness" in the image. To threshold the image they used an adaptive method based on the expected plant material in the image. To find line structures they used linear regression based on information from three row segments to find the best fit. Slaughter et al. [6] used a Bayesian classifier based on information from a color camera to discriminate weed from plants. The camera was centered above the row and they sum up all pixels in every column thus forming a histogram of the row structure. They used the median value of the histogram to estimate the position of the row. Olsen [7] used near infrared images taken from almost above the rows to minimize the effect of perspective geometry. To find rows he sums up all gray values column-vise to obtain a histogram of the crop rows and then applies a low-pass filter on the histogram. The peaks in the filtered signal correspond to the row positions. Using gray values directly minimizes the effects of shadows from the implement but he reported occurrences of misalignment of the row structure under different direction of illumination.

Common with previous reported methods is that they can deal with larger plants (>5 cm in diameter) or smaller plants, but then only with low weed density (typically 3 weeds/plant). An important aspect of our work is that the organically grown cultivation we aim at, should be robust against relatively high weed pressure, typically up-to twelve weed-plants per crop (in our case the crop is sugar beet plants). It is also important that the system is robust in all stages of growth, i.e. from cotyledon stage to full-grown plants. Moreover, sugar beet plants are sown and not transplanted which means that the crop and weed have about the same size at first cultivation. This means that in most cases segmentation of weed and crops based on plant features, such as size or shape, is difficult because of high similarity. The relative high weed pressure also makes occlusion more likely, which makes the segmentation task even worse. This means that methods that rely on a segmentation step between crops and weeds more likely will fail, especially for small sized plants.

Existing algorithms for row-recognition were not found to be sufficient for our application [8]. Therefore we decided to develop a new row-recognition algorithm which fit the requirements imposed by organically grown farming of sown crops.

Instead of relying of a segmentation step between crops and weeds, we use the context information of living material of both crop and weed in the scene to derive the position of the row structure. We consider weed as noise and employ the theoretical fact that if weeds are uniform distributed a row structure always will be found. To get a good signal to noise ratio, which means the ability to handle large weed pressure, we look at a row segment of about five-meter long. To be able to deal with missing crops in the rows and to increase robustness and accuracy we look at two rows at the same time.

2.1. Near-infrared images

We use a gray-scale camera with a near-infrared filter to detect plants. Using the filter gives a high-contrast image, where living plant material is bright and the soil is dark [9]. However due to disturbances, in form of different light conditions and due to the fact that all white pixels do not correspond to crops, some additional image processing has to be done.

To decrease the effect of different light conditions we perform an opening operation on the image [10] and subtract the result from the original image. This results in an intensity-independent gray-level image from which a binary image can be derived with a fixed threshold. In the resulting binary image, plant material from both weeds and crops is white and the rest, coming from soil, stones and residues is black.

2.2. Geometric model

Based on the binary image the next step is to detect the row of crops in the image. To control for example a mobile robot or other tool, we must know the offset and the heading angle of the camera relative to the row structure. To be able to extract this information from the image we must have a geometric model of the row structures and know the transformations between the world coordinate system and the image coordinate system. To start with we assume that the row of crops can be approximated with a straight line. The world coordinates of the row can first be transformed to camera coordinates and finally to image coordinates through a perspective transformation. Fig. 1 shows the geometry of the world and camera coordinate system. The heading angle, α, and the offset, s, are the unknown parameters which need to be found. It is assumed that the field can be approximated by a plane and the world coordinate system is chosen in such a way that this plane can be described as $z_w = 0$ and that the rows are lying in parallel with the y_w-axis.

In Appendix A we show, by using perspective transformation, that for every pixel (x_i, y_i) in the image corresponding to a plant in the row, there is a linear relation between s and α: for small values of the heading angle, α.

$$A\alpha + Bs + C = 0 \tag{1}$$

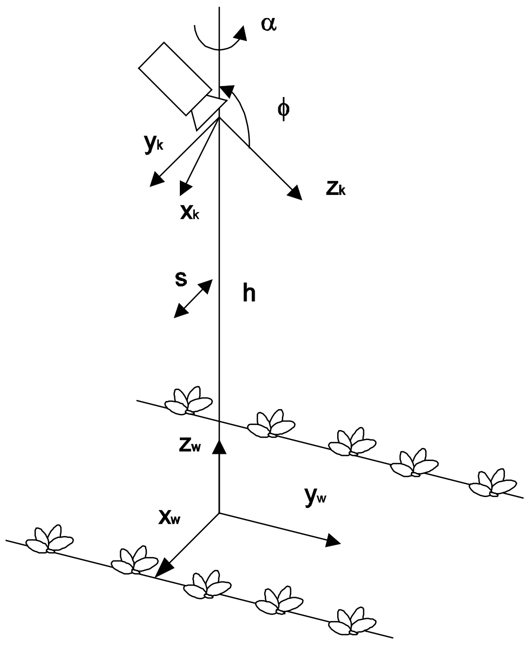

Fig. 1. World-Camera coordinate system.

where

$$A = (-hy_i \cos(\phi) - fh \sin(\phi))$$
$$B = (f \cos(\phi) - y_i \sin(\phi))$$
$$C = -x_w B + hx_i$$

x_w is known as it is the position of the row in world coordinates. The camera angle, ϕ, and its height above the ground, h, are constant as well as the focal length of the lens, f. Eq. (1) plays a central role to find the heading angle, α, and the offset, s, which is described in the next section.

2.3. Finding row-structures

The Hough transform is a well-known and robust method to find lines, especially if the lines cover the whole image as in our case [10]. For example, in road lane navigation a number of approaches uses Hough transform for line detection [11,12]. Normally the lines are found with their equation in the image space, e.g. $y_i = ax_i + b$, where the coefficients a and b are found with the Hough transform. This could also be done based on the binary image of plant material. All pixels coming from the crops contribute to the line and all pixels from the weeds are just noise. However, as shown in Eq. (1), there is a linear relation between s and α for a given pixel (x_i, y_i). (See Eq. (1), where the constants A, B and C are functions of (x_i, y_i).) This means that we directly can perform the Hough transform for s and α. The Hough-space, $H(s, \alpha)$, is then an accumulator where each pixel coordinate (x_i, y_i) in the binary image, which is on, contributes even if it is belonging to a weed plant. Every such pixel (x_i, y_i) forms a straight line in the Hough-space. The recognition of the row of crops among the weeds is achieved due the fact that weeds are uniformly distributed in the field, whereas all the crops grow exactly in a row, thus leading to a peak in the Hough-space, as shown in Fig. 2.

The novelty of this algorithm is that we model a plant row with a rectangular box instead of a line. The width of the box is equal to the average width of the plants and the length of the box is "unlimited" as it fills the whole image. The rectangular box can be described by a set of parallel adjacent lines, which appear in the image as a set of lines which intersect in one virtual point outside the image, as shown Fig. 3, due to

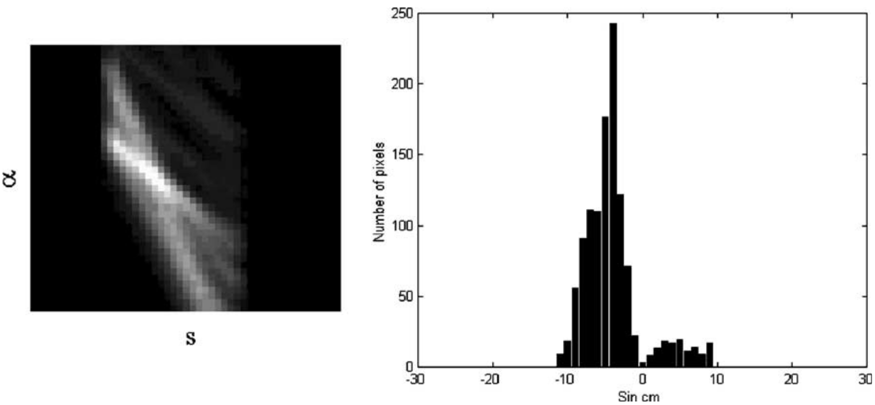

Fig. 2. Hough space accumulator and distribution of s for a specific α.

Fig. 3. Rectangular box corresponds to a number of adjacent s-cells.

perspective geometry. The number of lines is the width of the box divided by the line thickness, which is determined by the pixel size of the image. In our Hough-space this means that for one value of α the rectangular box corresponds to a number of adjacent s-cells. By summing up the contributions of the adjacent s-cells we obtain the support for a rectangular box, i.e. for the row of plants. The best estimate of s and α is found by searching for the maximum of the sum of adjacent s-cells in our Hough-space. Adaptation to different size of the plants can easily be done, by reducing or increasing the number of adjacent s-cells.

Fig. 2 shows an example of the Hough-space and the corresponding distribution of s for the correct value of α. In this example we look at the sum of three adjacent s-cells to find the most likely values of s and α, i.e. the middle line of the box and the corresponding orientation.

If more than one row is used each row has then its own corresponding Hough-space. Information from the different rows can be fused together by calculating the average of s and α derived from the individual Hough-spaces. Another possibility is to sum up the contributions from all Hough spaces for each cell (s, α), thus forming a common Hough-space and extract the most likely value of s and α from this one. The resolution of the Hough-space, i.e. the size of the cell (s, α) must be chosen carefully, where the resolution of the camera plays a major role. The size of s and α are chosen so that this corresponds to at least one pixel difference. In our case the resolution of s was set to 1 cm and of the heading angle, α, to 0.2°.

3. Experiments and result

3.1. Measurement performance

To evaluate the row-recognition system a number of real images were used to verify the system. To capture images we use an industrial PC, equipped with a Pentium II 333 MHz processor and a Matrox Pulsar frame-grabber. The vision system is a COHU CCD gray-scale camera with a near-infrared filter (780 nm) and a lens with 8.5-mm focal length. The size of the images was 384×288 pixels, which represent about 6 m of the row-structure. Three sets of images from sugar-beet plants at three different stage of growth were included. Also one set of images from a rape field was used (set 1 in Fig. 4). From each set of images a subset of 70 spatially distributed images was chosen. The most relevant properties of each set are given in Table 1 and one example image of each set is shown in Fig. 4. For all images, the real position of the camera relative to the rows was estimated twice by a human observer. The average value from the two observations was taken as the final estimate of the real values and used to evaluate the performance of the row recognition system.

As we look at two rows simultaneously, a left and right one, each row has its own corresponding Hough-space, leading to two different observations for s and α. The Hough-spaces of the left and right row are also fused together in a common Hough-space by summing up the entry of each corresponding cell (s, α). In Table 2 the error of the measured offset s based on the left and right row separately, and

Fig. 4. Four sets of real images.

Table 1
Four sets of real-images

Set	Stages	∅ cm	$\frac{\text{Weed}}{\text{Plant}}$
1	First true leaf (rape)	7–10	1
2	First-second true leaf	7	3.5
3	Cotyledon	5	<1
4	Second true leaf	10–13	<1

Table 2
Standard deviation of S in cm

Set	S_{right}	S_{left}	S_{common}	$\frac{S_{\text{right}}+S_{\text{left}}}{2}$
1	1.6 (1.8)	2.3 (2.2)	1.2 (1.3)	1.4 (1.4)
2	1.0 (1.2)	0.8 (0.8)	0.6 (0.8)	0.7 (0.8)
3	1.1 (2.2)	1.4 (2.6)	0.9 (1.2)	0.9 (1.5)
4	1.0 (1.3)	1.4 (1.5)	0.8 (1.0)	0.9 (0.9)

based on the common Hough-space is shown. The corresponding values for the heading, α, is shown in Table 3. Another way of combining the results from the two rows is to use the average result from the left and right row. The result of this is also shown in Tables 2 and 3. Based on the result reported in Table 2 we may conclude that the row-recognition system has a good performance ranging from 0.6 cm standard deviation of error for set 2 to 1.2 cm for set 1. The mean of the error is 0 cm.

Table 3
Standard deviation of α in degree

Set	α_{right}	α_{left}	α_{common}	$\frac{\alpha_{right}+\alpha_{left}}{2}$
1	0.27 (0.33)	0.63 (0.63)	0.21 (0.28)	0.35 (0.37)
2	0.27 (0.31)	0.14 (0.13)	0.11 (0.16)	0.16 (0.17)
3	0.19 (0.51)	0.41 (0.80)	0.17 (0.23)	0.21 (0.48)
4	0.17 (0.22)	0.21 (0.23)	0.13 (0.17)	0.14 (0.15)

Moreover, it is shown that by using two rows instead of one the accuracy is improved significantly. It seems that fusion of the information from the two rows should be done in the Hough-space directly rather than by averaging the final results, as this leads to slightly less standard deviation of the error. The reason for this is demonstrated in Fig. 5. Suppose that plant Nos. 2, 3 and 4 in the right row are weeds and there are three missing plants in the row. In the right row the dashed line will be selected as the most probable row structure. Taking the average value from the left and right row will then give an incorrect result. Instead when summing up both Hough-spaces into one, as suggested in Section 2.3, plant 1 and 5 contribute both to the correct value, giving the correct position of the row structure. Please note that calibration of the necessary system parameters, i.e. row distance and tilt, ϕ, of the camera as well as camera height above ground, h, is done manually.

The Hough transform calculation time is in our case data-dependent; i.e. only white pixels are processed. That means that for large crops there will be a lot of pixels to process. A possibility is then to sub-sample the image in order to reduce the amount of data. The above described test was repeated for all images which were reduced in size by a factor two (192 × 144 pixels). The result, shown also in Table 2 between parenthesis ( ), is an increase of the error with 0.2 cm for large crops and 0.3 cm for small crops. This means that for large crops, when processing time might be critical, the image can be sub-sampled without a significant loss of accuracy.

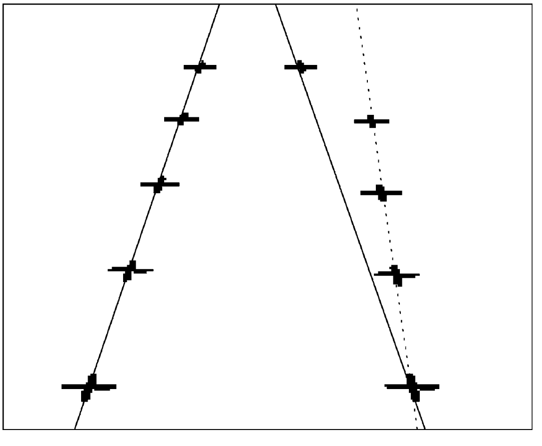

Fig. 5. Rows found in image.

3.2. Control of an inter-row cultivator

The row recognition system was implemented on an inter-row cultivator [13]. The system consists of a tractor that the farmer drives along the rows with the cultivator mounted at the rear of the tractor. A steering-unit based on a very thin steering wheel, which cuts through the soil, is used to control the position of the cultivator based on the input of the row recognition system (see Fig. 6). The controller was designed as a PID-controller. A number of field tests with a 6-row cultivator have been done to investigate the performance of the row-recognition system, including tests under different light conditions, weed density and size of plants. The system was tested on two types of crops: sugar beet plants and rape. The row distance was typically 48 cm and the distance between the plants in the row was about 12 cm for the sugar beet plants and 15 cm for the rape. The weed density varied from field to field and could reach a level up to 12 weed/plants (200 weeds/m^2), see Fig. 7. The result of one of these field tests is shown in Table 4 (6-row cultivator).

As a final evaluation of the system, field tests have been performed at two commercial farmers. Farmer one has a 9-row cultivator (4.5 m width) and the field was almost planar. The second farmer has a 18-row-cultivator (9 m width) with two steering wheels and the field was rather hilly and has a tilt of up to 4° perpendicular to the rows at some part at the fields. The controller was able to effectively deal with this disturbance due to its integral part. At farmer one the system was tested at a total distance of 3.5 km and at farmer two at a total distance of 2 km. Table 4 shows the accuracy of the row-tracking system for specific runs during these tests.

The results in column "Error σ_{observed}" in Table 4 are the position errors from the row-tracking system as observed by the row-recognition system. However, as the row-recognition system itself contains a measurement error (see Table 2) and the row-recognition system is part of the row-tracking system (feed-back to the controller),

Fig. 6. Tractor with row-cultivator.

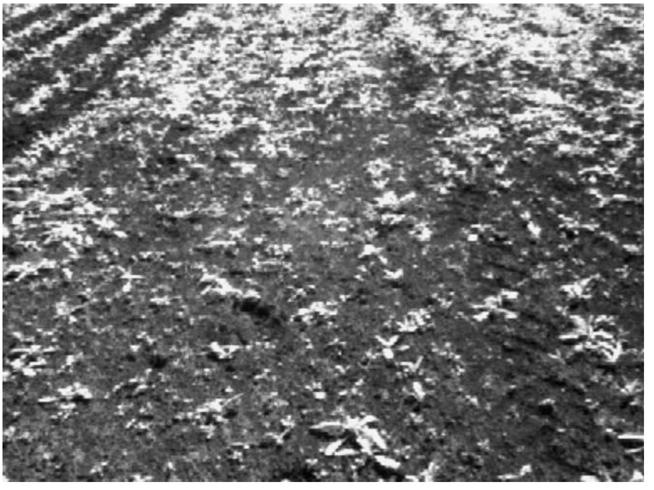

Fig. 7. Example of high weed pressure, 200 weeds/m^2.

Table 4
Performance of row-recognition system

Type	Error mean$_{observed}$ (cm)	Error $\sigma_{observed}$ (cm)	Error σ_{wcf} (cm)	Error σ_{wc} (cm)	Speed (km/h)	Weed density (m^2)	Plant size (cm)	Distance (m)
6 rows	0.2	2.0	2.7	3.0	3.0	≈0	10–15	200
9 rows	0.2	2.1	2.6	3.1	2.5	<50	15–20	220
18 rows	0.0	2.4	2.7	3.4	3.5–4.5	<50	15–30	390

we need to analyze the observed error. The observed position error is the true position error added with the measurement error. The relation can be described as the sum of two stochastic variables:

$$X_{\text{observed}} = X_{\text{position}} + X_{\text{me}} \tag{2}$$

where the measurement error and the observed error assumed to have zero mean, see Table 4. This gives the following equation of the position variance:

$$\sigma^2_{\text{position}} = \sigma^2_{\text{observed}} - \sigma^2_{\text{me}} - 2^* r^* \sigma^*_{\text{position}} \sigma_{\text{me}} \tag{3}$$

where r $[-1, 1]$ is the correlation coefficient between the position and the measurement error. An estimation of the measurement error can be derived from Table 2 and is here assumed to be $\sigma_{\text{me}} = 1.0\,\text{cm}$. From Eq. (3) a worst case scenario can be derived in case $r = -1$, which gives:

$$\sigma_{\text{wc}} = \sigma_{\text{me}} + \sigma_{\text{observed}} \tag{4}$$

So we have to correct the observed error according to (4) to obtain a worst-case estimation of the error of the true position (see Column Error σ_{wc} in Table 4). Due to the fact that the measurement is used as error signal into the control loop, a negative correlation is likely to occur. However, $r = -1$ is a too pessimistic assumption. Taking into account the fact that the row cultivator system acts as a low-pass filter, the high frequency part of the observed signal is only due to uncorrelated measurement noise. Filtering the observed signal with a low-pass filter will thus give a more realistic observation of the position. By first using a low-pass filter with a cut off frequency of 1 Hz, and then repeating the worst-case calculations, as given in Eq. (4), a more likely estimation of the true position error is obtained. This is shown in Table 4 as σ_{wcf}, with values varying from 2.6 to 2.7 cm.

3.3. Control of an autonomous mobile robot

The row recognition system was also implemented on a mobile robot [14]. The robot should be able to follow a row of plants guided by the row-following vision system, see Fig. 8. Another vision-system will then be utilized to recognize the individual sugar beet plants among weed within the row, thus enabling the robot to perform intra-row cultivation (in the seedline of the row). To evaluate the row recognition system a number of tests where performed both indoor and outdoor.

For the indoor tests a number of artificial plants were placed in a corridor. The camera for the row recognition system was mounted at the front of the robot, see Fig. 8. This is also the control point of the robot. To perform intra-row cultivation, the robot has an intra-row-weeding tool mounted at the rear of the robot. At the weeding position, a second camera was mounted to measure the actual position of the robot. The robot uses Ackerman steering [15] and was driving at 0.2 m/s. The error of the lateral offset measured by the vision-system was ±1 cm and the error measured with the down looking camera was ±0.5 cm, as illustrated in Fig. 9.

The system was also tested outdoors on a sugar beet field. The robot was tested for two runs. On run one the robot was tested for 78 m and the weed-pressure was

Fig. 8. Mobile robot.

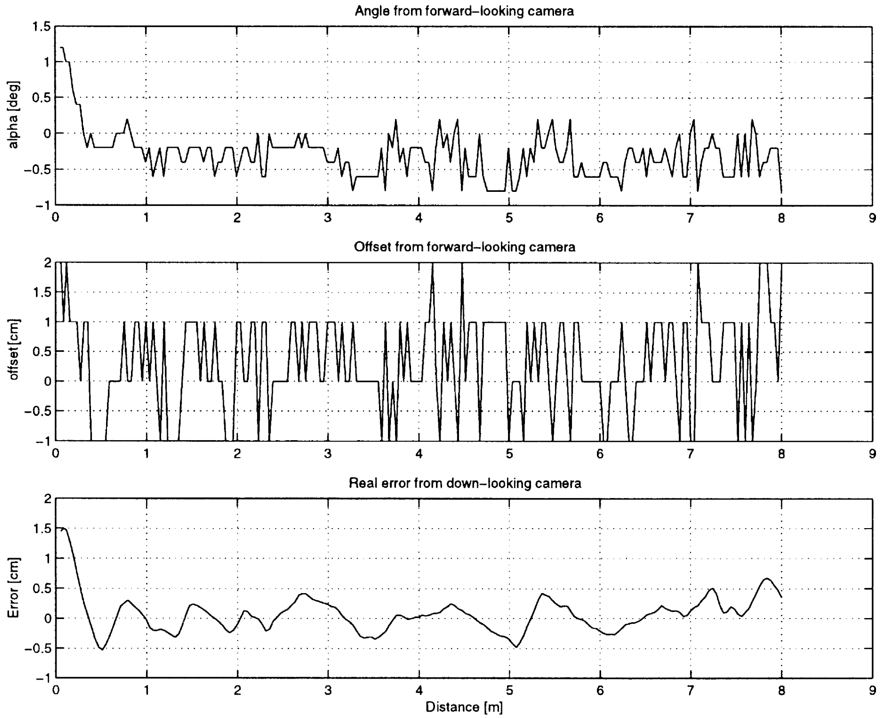

Fig. 9. Result at 0.2m/s with an error of 1.5cm at the beginning.

Table 5
Performance of autonomous mobile robot

Run	Error mean$_{observed}$ (cm)	Error $\sigma_{observed}$ (cm)	Error σ_{wcf} (cm)	Distance (m)
1	−0.5	1.3	1.9	78
2	−0.5	1.8	2.3	34

25 weed/m^2. On run two the robot was tested for 34 m and the weed-pressure was in average 100 weeds/m^2 but had parts with a weed-pressure of 200 weeds/m^2, see Fig. 7. The robot was driving at a speed of 0.1 m/s and Table 5 shows the accuracy during these tests. The estimated worst case error was calculated the same way as for the row cultivator described in Section 3.2. At the tool position, the rear of the robot, half of the values obtained in Table 5 can be expected according to the results obtained indoors.

3.4. Algorithm speedup

To increase the computational speed, the vision algorithms are implemented with full use of the MMX-utilities of the processor. Both the opening algorithm and the

Hough transform are well suited for parallel implementation. The optimized Hough transform is between 2 and 3 times faster compared to using the non-optimized code of the algorithm [16]. A typical processing time of the whole algorithm, opening and Hough transform, takes about 27 ms for midsize plants on a Pentium II 333 MHz compared to 112 ms without optimization.

4. Exception handling

It is important for the vision algorithm to detect situations were the system does not produce reliable results. Examples are the end of a row, when the rows are not straight enough or when there are too many missing plants. In these cases the control system should be notified by the vision system so it can take the appropriate actions. For example the steering wheel of the inter-row cultivator can be put in the neutral position and the tractor driver can be notified in case of an exception. Here we present an approach for the detection of exceptions and preliminary results for the end of row situation. The approach is to calculate the offset and heading angle of the row-structures based on the upper third of the image and compare this to the offset and heading angles calculated for the whole image. This can be done at very low computational cost as the Hough-space can be saved separately when the upper third of the image is processed before the rest of the image is processed. As we look at two rows at the same time six comparisons are possible, i.e. heading angles and offsets for each row separately and for the common calculated heading angle and offset. A training-set of 500 images was used to estimate the standard deviation of the difference between the parameters of the upper third of the images and the whole image. If the difference of a parameter falls outside a three σ confidence interval we counted this as a fault ($\sigma_s = 1.6$ cm and $\sigma_\alpha = 0.17°$). If there are more than two faults of the six parameters an exception flag is raised. The algorithm was tested on the test sets described in Section 2.3 to see what would happen in a "normal" situation. The result is presented in Table 6. Set Nos. 2 and 4, see Fig. 4, give very nice results as no or almost none exceptions are detected. Set Nos. 1 and 3 however indicate a few exceptions. This is naturally however as set 3 contains very small plants which are the limit for our algorithm and set 1 was in reality not a very straight row.

The end of row situation was tested in three different situations. One situation is illustrated in Fig. 10 where the end of row appears in the image from image No. 65. At image No. 80 one sixth of the image contains the end of row information. At image No. 90, one third of the image is covered with the end of row information. We can see in Fig. 10 that the first exception is raised at image No. 74 and then frequently appears. Finally from No. 92 and onwards the exception flag is raised continuously. Based on this one can easily detect an end of row situation and distinguish

Table 6
Exceptions

Set	1	2	3	4
No. of exceptions	5	0	5	2

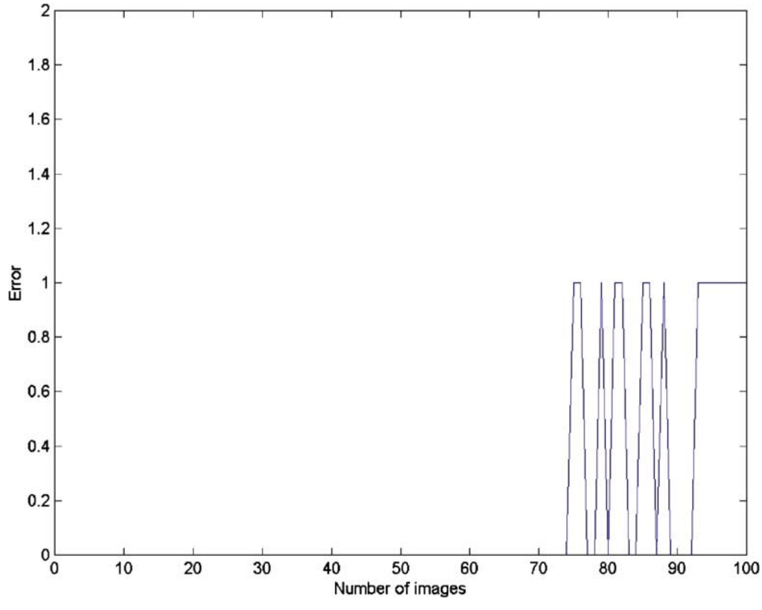

Fig. 10. End of row detection.

it from occasionally raised exceptions due to non-straight rows etc. as shown in Table 6.

5. Conclusion and future work

In this paper we presented a new method for robust recognition of plant rows based on the Hough transform that is able to guide agricultural machines, such as inter-row cultivators or mobile robots. The system is not restricted to a specific crop and has been tested on both sugar beets and rape. It uses the context information of living material in the scene to derive the row position. We presented a new way of modeling the plant rows. Instead of using a line as model, we use a box as model, where the width of the box corresponds to the size of the plants. The Hough space representation proved to be very suitable to match this model in a computationally efficient way with the real scene in order to locate the row structure. The representation makes it possible to merge information coming from several rows and even from more than one camera. The representation also makes it possible to adapt to different plant sizes. Finally, the representation makes it possible to detect exceptional situations by itself, for example the occurrence of the end of a row.

The accuracy of the position estimation relative to the row proved to be good, with a standard error between 0.6 and 1.2 cm depending on plant size, which varied between 5 and 30 cm. The system was implemented on two types of agricultural machinery, a row cultivator and an agricultural mobile robot. Extensive field tests showed that the system is sufficiently accurate and fast to control the cultivator and the mobile robot in a closed-loop fashion with a standard deviation of the position of 2.7 and 2.3 cm, respectively. The tests also showed that the system is able to deal with incomplete row structures due to missing plants combined with high weed pressure (up to 200 weeds/m^2). It can be summarized that the field tests proved that the new method of row guidance fulfils our requirements imposed by ecological farming.

Future work may follow two paths. The first is to proceed the development of the row recognition system. The system is able to work in a relatively hilly environment and is able to compensate for tilt that is perpendicular to the row structure. The restriction of today's system is that it has to be locally flat along the row structure for about 10 m. To be able to capture larger variations, an on-line automatic calibration of the forward-looking angle of the camera must be implemented. As crop size varies along the rows, it would also be interesting to implement an automatic adaptation to crop size. This value is currently set by the user based on the expected average plant size in the field. Another interesting area is to investigate the possibilities for tracking rows that are not straight for example with the method suggested by [11]. Today's system can only handle rows that are straight or slightly curved.

The second path is to continue the development of our robot platform. Our long-term goal is the development of an autonomous mobile robot able to perform mechanical weed control both between and within the rows. To make this possible, a system for distinguishing between crops and weeds within a row must be developed. The employment of autonomous robots instead of human-labour for mechanical weed control will make it possible to increase the amount of ecologically grown products significantly.

Acknowledgments

This work was carried out in the Mech-Weed project, which has been a co-operative effort between Halmstad University and Danisco Sugar AB. The project was sponsored by the Swedish KK-foundation "Stiftelsen för Kunskap och Kompetensutveckling" and by "Skogs-och Jordbrukets Forskningsråd", SJFR. The project is currently sponsored by FORMAS: The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning.

Appendix A

The transformation from a position expressed in the camera coordinate system k to a position expressed in the world coordinate system w is in general [17]:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}^w = [R]_k^w \begin{bmatrix} x \\ y \\ z \end{bmatrix}^k + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}_{k_{\mathrm{ORG}}}^w \tag{A.1}$$

where R is the rotation-matrix between world coordinate and the camera coordinate system. In our case R is a function of ϕ and α:

$$[R]_k^w = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha)\cos(\phi) & \sin(\alpha)\sin(\phi) \\ \sin(\alpha) & \cos(\alpha)\cos(\phi) & -\cos(\alpha)\sin(\phi) \\ 0 & \sin(\phi) & \cos(\phi) \end{bmatrix} \tag{A.2}$$

In our case the translation-vector is a function of α, s and h:

$$\begin{aligned} t_z &= h \\ t_x &= s \cdot \cos(\alpha) \\ t_y &= s \cdot \sin(\alpha) \end{aligned} \tag{A.3}$$

The inverse transformation is:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}^k = [R]_w^k \begin{bmatrix} x \\ y \\ z \end{bmatrix}^w + [t]_{w_{\mathrm{ORG}}}^k \tag{A.4}$$

where

$$[R]_w^k = [R]_k^{w^{\mathrm{T}}} \quad \text{and} \quad [t]_{w_{\mathrm{ORG}}}^k = -[R]_w^k [t]_{k_{\mathrm{ORG}}}^w \tag{A.5}$$

which gives:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}^k = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) & 0 \\ -\cos(\phi)\sin(\alpha) & \cos(\alpha)\cos(\phi) & \sin(\phi) \\ \sin(\alpha)\sin(\phi) & -\cos(\alpha)\sin(\phi) & \cos(\phi) \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}^w + \begin{bmatrix} -s \\ -h \cdot \sin(\phi) \\ -h \cdot \cos(\phi) \end{bmatrix} \tag{A.6}$$

As said before, the row structure can be regarded as a line lying in the plane $z_w = 0$, which gives us the three following equations from (A.6):

$$\begin{aligned} x_k &= \cos(\alpha)x_w + \sin(\alpha)y_w - s \\ y_k &= -\cos(\phi)\sin(\alpha)x_w + \cos(\alpha)\cos(\phi)y_w - h\sin(\phi) \\ z_k &= \sin(\alpha)\sin(\phi)x_w - \cos(\alpha)\sin(\phi)y_w - h\cos(\phi) \end{aligned} \tag{A.7}$$

The camera can be simply modeled as a pinhole camera where f is the so-called focal length [18]. The x_i, y_i coordinate system is the image-sensor coordinate system of which the axis are parallel to x_k and y_k, respectively.

$$x_i = -f \cdot \frac{x_k}{z_k} \quad y_i = -f \cdot \frac{y_k}{z_k} \tag{A.8}$$

Eqs. (A.7) and (A.8) together give:

$$x_i = -f \frac{\cos(\alpha)x_w + \sin(\alpha)y_w - s}{\sin(\alpha)\sin(\phi)x_w - \cos(\alpha)\sin(\phi)y_w - h\cos(\phi)} \tag{A.9}$$

$$y_i = -f\frac{-\cos(\phi)\sin(\alpha)x_w + \cos(\alpha)\cos(\phi)y_w - h\sin(\phi)}{\sin(\alpha)\sin(\phi)x_w - \cos(\alpha)\sin(\phi)y_w - h\cos(\phi)} \tag{A.10}$$

Given a known position in the field described as x_w, y_w, Eqs. (A.9) and (A.10) give the position in the image described as x_i, y_i assuming we know the position and orientation of the camera described by α, s, h, ϕ. In our case we know that every pixel belonging to a plant row has a certain fixed position x_w, which depends on the row distance but generally we do not know the value y_w. Therefore we eliminate y_w from Eqs. (A.9) and (A.10) which gives:

$$\begin{aligned} s\cdot\cos(\alpha)\cdot(-f\cos(\phi) + y_i\sin(\phi)) = {} & -fx_w\cos(\phi) + hx_i\cos(\alpha) \\ & - hy_i\cos(\phi)\sin(\alpha) + x_w y_i\sin(\phi) \\ & - fh\sin(\alpha)\sin(\phi) \end{aligned} \tag{A.11}$$

For small values of the heading angle, α, we can simplify the equation by assuming that:

$$\sin(\alpha) = \alpha \quad \text{and} \quad \cos(\alpha) = 1. \tag{A.12}$$

which gives:

$$\begin{aligned} & a(-hy_i\cos(\phi) - fh\sin(\phi)) + s(f\cos(\phi) - y_i\sin(\phi)) \\ & \quad - x_w(f\cos(\phi) - y_i\sin(\phi)) + hx_i = 0 \end{aligned} \tag{A.13}$$

This can also be written as

$$A\alpha + Bs + C = 0 \tag{A.14}$$

where

$$\begin{aligned} A &= (-hy_i\cos(\phi) - fh\sin(\phi)) \\ B &= (f\cos(\phi) - y_i\sin(\phi)) \\ C &= -x_w B + hx_i \end{aligned}$$

Finally, a straightforward transformation from image-sensor coordinates to image-pixel coordinates is needed:

$$x_i = -(x_\mathrm{p} - x_\mathrm{c})dpixelx, \quad y_i = -(y_\mathrm{p} - y_\mathrm{c})dpixely \tag{A.15}$$

where *dpixelx* and *dpixely* is the size of the sensor elements in the x- and y- direction of the CCD-array. x_c and y_c are the coordinates of the optical center of the camera in pixels and set to half the size of the image.

This gives the following equation:

$$x_\mathrm{p} = \frac{A}{h\cdot dpixelx}\cdot\alpha + \frac{B}{h\cdot dpixelx}\cdot s - \frac{B}{h\cdot dpixelx}\cdot x_w + x_\mathrm{c} \tag{A.16}$$

This equation plays a central role to find the heading angle, α, and the offset, s. Please note that x_w is known as it is the position of the row in world coordinates.

References

[1] Guilherme N, Kak AC. Vision for mobile robot navigation: survey. IEEE Trans Pattern Anal Mach Intell 2002;42(2):237–67.

[2] Pilarski T, Happold M, Pangels H, Ollis M, Fitzpatrick K, Stentz A. The demeter system for automated harvesting. Autonom Rob 2002;13:9–20.

[3] Marchant JA. Tracking of row structure in three crops using image analysis. Comput Electron Agric 1996:161–79.

[4] Marchant JA, Brivot R. Real-time tracking of plant rows using a hough transform. Real-Time Imag 1995:363–71.

[5] Billingsley J, Schoenfisch M. Vision-guidance of agricultural vehicles. Autonom Rob 1995(2):65–76.

[6] Slaughter DC, Chen P, Curley RG. Computer vision guidance system for precision cultivation. ASAE Meeting Presentation, Minneapolis, Minnesota, August 10–14, 1997.

[7] Olsen HJ. Determination of row position in small-grain crops by analysis of video images. Comput Electron Agricult 1995;12:147–62.

[8] Åstrand B, 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.

[9] Baerveldt AJ. Vision-guided mechanical weed-control. The 2nd International Workshop on Mechatronical Computer Systems for Perception and Action (MCPA'97), Pisa, Italy, 10–12 February 1997, p. 135–8.

[10] Shapiro LG, Haralick RM. Computer and robot vision volume I and II. Addison-Wesley Publishing Company Inc.; 1992–1993.

[11] Yu B, Jain A.K. Lane boundary detection using a multiresolution hough transform. proceedings. International Conference on Image Processing, 26–29 Oct. 1997, vol. 2, p. 748–51.

[12] Jin-Woo Lee, Sung-Uk Choi, Young-Jin Lee, Lee KS. A study on recognition of lane and movement of vehicles for port AGV vision system, SICE 2001. In: Proceedings of the 40th SICE Annual Conference. International Session Papers, 25–27 July 2001, p. 38–41.

[13] Åstrand B, Baerveldt AJ. Robust tracking of plant rows using machine vision. In: The Sixth International Conference on Mechatronics and Machine Vision in Practice, M^2 VIP'99, September 1–3, 1999.

[14] Åstrand B, Baerveldt AJ. Vision-based perception for an agricultural mobile robot for mechanical weed control. Autonom Rob 2002;13:21–35.

[15] Gillespie TD. Fundamentals of vehicle dynamics. Society of Automotive Engineers Inc.; 1992.

[16] González Menéndez RO. Image processing and speeding-up calculations using MMX-technology. Degree Project Report, Halmstad University, March, 1999.

[17] Craig JJ. Introduction to robotics mechanics and control. Addison-Wesley Publishing Company Inc.; 1989.

[18] Jähne B. Digital image processing. Springer-Verlag; 1997.