Weed Technology. 2006. Volume 20:853–861

A Mobile Field Robot with Vision-Based Detection of Volunteer Potato Plants in a Corn Crop[1]

FRITS K. VAN EVERT, GERIE W.A.M. VAN DER HEIJDEN, LAMBERTUS A. P. LOTZ, GERRIT POLDER, ARJAN LAMAKER, ARJAN DE JONG, MARJOLIJN C. KUYPER, ELTJE J. K. GROENDIJK, JACQUES J. NEETESON, and TON VAN DER ZALM[2]

Abstract: Volunteer potato is a perennial weed that is difficult to control in crop rotations. It was our objective to build a small, low-cost robot capable of detecting volunteer potato plants in a cornfield and thus demonstrate the potential for automatic control of this weed. We used an electric toy truck as the basis for our robot. We developed a fast row-recognition algorithm based on the Hough transform and implemented it using a webcam. We developed an algorithm that detects the presence of a potato plant based on a combination of size, shape, and color of the green elements in an image and implemented it using a second webcam. The robot was able to detect potatoes while navigating autonomously through experimental and commercial cornfields. In a first experiment, 319 out of 324 images were correctly classified (98.5%) as showing, or not showing, a potato plant. In a second experiment, 126 out of 141 images were correctly classified (89.4%). Detection of a potato plant resulted in an acoustic signal, but future robots may be fitted with weed control equipment, or they may use a global positioning system to map the presence of weed plants so that regular equipment can be used for control.

Nomenclature: Corn, *Zea mays* L.; Potato, *Solanum tuberosum* L.

Additional index words: Autonomous navigation, autonomous weeding, glyphosate, machine-vision, site-specific weed control.

Abbreviations: DIP*lib*, Delft image-processing library; DSP, digital signal processor; GPS, global positioning system; JPEG, Joint Photographic Experts Group; NiMh, nickle metal hydride; PC, personal computer.

INTRODUCTION

Volunteer potato is a perennial weed that is difficult to control in crop rotations (Boydston 2001). In addition to reducing crop yield, volunteer plants allow the survival of harmful diseases (e.g., *Phytophthora infestans*), nematodes (e.g., the potato cyst nematode *Globodera pallida*), and insects during years when no potatoes are grown, possibly causing infections to nearby potato fields. Because of these diseases, farmers in The Netherlands are under a statutory obligation to control volunteer potato plants. The most common and widely used method for this is a selective application of glyphosate to the potato volunteers by hand. The cost of this control is high because of the labor involved. For example, these costs range from $63 to $375/ha in sugarbeet (*Beta vulgaris* L.) crops in the Netherlands (K. Westerdijk, personal communication). Potato volunteers occur on ca. 50% of the sugarbeet hectarage. The density of volunteers varies mostly between 80 and 10,000 plants/ha, but higher densities are also observed.

Small robots have been proposed as a way to reduce the cost of agricultural operations (Anonymous 2004; Blackmore et al. 2005). Robotic control of volunteer potatoes involves at least three major steps: navigation through the crop, detection of potato plants, and control of potato plants.

The sensors that can be used by an agricultural robot to support navigation include mechanical feelers, a global positioning system (GPS) receiver, geomagnetic direction sensors, and machine-vision

[1] Received for publication September 15, 2005, and in revised form January 9, 2006.

[2] First through fourth, and ninth authors: Researcher, and eight and tenth authors, Research Technician, Plant Research International, P.O. Box 16, 6700 AA Wageningen, The Netherlands; fifth author: System Administrator, Wageningen UR, Wageningen, The Netherlands; sixth and seventh authors: Software Engineer, Center for Geo-Information, P.O. Box 47, 6700 AA Wageningen, The Netherlands. Corresponding author's E-mail: frits.vanevert@wur.nl.

(Reid et al. 2000). Machine-vision has been used in a cauliflower (*Brassica oleracea* L.) weeding robot (Tillett et al. 1998) and more recently in sugarbeets (Åstrand and Baerveldt 2002; Åstrand and Baerveldt 2003). These robots target whole-field control of weeds and, thus, need significant power either to control weeds by mechanical means or to carry the necessary amount of herbicide. But if the amount of herbicide to be carried is small, as will be the case when volunteer potato plants occurring at low density are to be controlled, it becomes feasible to consider a robot that is small enough to fit between two adjacent crop rows. Small robots, however, carry their sensors close to the ground and to the crop, which makes them sensitive to dirt and malfunctioning (Van Straten 2004).

Navigation through a row crop requires detection of the exact location of the rows of plants. Any plant growing between rows can then be safely controlled. The problem of detection of volunteer potato plants in a row crop can be described more precisely as the problem of detecting the potato plants growing in the rows of crop plants. Detection of weeds against a soil background has been addressed by many authors and is quite feasible, but detection of weeds against a background of other plants is much more difficult. We are not aware of reports in the literature about detection of potato plants against a background of other plants.

The objective of the work reported here was to build a small, low-cost, autonomous vehicle able to navigate up and down the successive rows in a commercial cornfield, making turns at the headland. While moving, this vehicle will use a low-cost detection system to recognize volunteer potato plants growing in the rows and generate a signal that in future work will be used to actuate weed-control equipment. The characteristics of the generated signal and the method of weed control are not considered here.

In this pilot study a corn crop was used, even though potato in corn is easily controlled using a selective herbicide that allows monocots to escape. This study is a continuation of the work done for the 2005 Wageningen Field Robot Event,[3] in which corn was used.

[3] 2005 Wageningen Field Robot Event information on Web page: http://www.fieldrobot.com.

MATERIALS AND METHODS

Autonomous Vehicle and Navigation. We based our robot on an electric toy truck[4] (Figure 1). The radio-control parts were removed and a microcontroller board[5] was installed to control speed and steering. Sensors for navigation included an ordinary webcam,[6] mounted atop a mast at a height of 1.5 m; a home-built revolution counter, affixed to the main axle and a solid-state gyroscope.[7] Images captured by the webcam were processed by an onboard mini-ITX personal computer (PC) board,[8] equipped with a 1.6 GHz Pentium-M processor.[9] Power to the motors, the microcontroller board, and the PC is provided by 4 7.2 V nickel metal hydride (NiMh) batteries.

A typical image captured by the webcam is shown in Figure 2a. This image has a resolution of 320 by 240 pixels. Plants are separated from the background by taking twice the value of the green channel and subtracting the values of the red and blue channels ($2G - R - B$). Pixels with a value higher than a certain threshold are considered plant, and all other pixels are considered background. The white pixels in Figure 2b represent the plant material detected with this method. In a second processing step, plant rows are recognized with an algorithm inspired by the Hough transform (Hough 1962). Imaginary lines are drawn over the segmented image and scored by the number of white (plant) pixels that they cover. These scores become pixel values in a new image, where the vertical coordinate of each pixel corresponds to the slope of the imaginary line, and where the horizontal coordinate corresponds to the intercept of that line with the vertical axis.

Because a plant row typically contributes pixels to several imaginary lines, plant rows show up in the new image as areas with bright pixels (Figure 2c). Thresholding (Figure 2d) and dilation on the resulting binary image are used to merge areas of bright pixels into contiguous areas (Figure 2e); after which, the

[4] Electric toy truck, TXT-1, Tamiya, 3-7 Ondawara, Shizuoka City, Japan.

[5] Microcontroller board, LPC2129, Olimex, 89 Slavjanska Street, P.O.Box 237, Plovdiv 4000, Bulgaria.

[6] Webcam, NX Ultra, Creative Resource, 31 International Business Park, Singapore 609921.

[7] Solid-state gyroscope, ADXRS150, Analog Devices, 3 Technology Way, Norwood, MA 02062.

[8] Mini-ITX PC, G5M100-N, DFI Inc., 100, Huan-Ho Street, Hsi-Chih City Taipei Hsien, Taiwan, R.O.C.

[9] 1.6 GHz Pentium-M processor, Intel, 2200 Mission College Blvd, Santa Clara, CA 95054.

center of gravity of each contiguous area is taken to represent a plant row. Figure 2b shows recognized rows superimposed on the thresholded image. Depending on the size of crop plants, the level of weed density, lighting conditions, and algorithm parameters, three or more plant rows may be detected in this way. In that case, rules are used to filter out unlikely rows. Two examples of such rules are "plant rows must be on either side of the robot" and "plant rows must be reasonably parallel." After filtering, at most two rows are left, and the robot's heading is obtained as the average of the headings with respect to those two rows.

Finally, the end of the plant row needs to be detected. First, the total number of plant pixels in each vertical line of the image is calculated. If the coefficient of variation of the number of plant pixels per vertical line is higher than a predetermined threshold value, it is assumed that the image contains an end of the plant row. The location of the end of the rows is then determined as the point where the moving average of the number of plant pixels per vertical line changes most. The complete analysis yields the angle and the offset of the robot relative to the plant rows, as well as the distance from the robot to the end of the plant row.

When the robot reaches the end of a row, it executes a three-part turn by driving forward and making a 90° turn, driving backward a prescribed distance, and driving forward making a second 90° turn. At the end of this maneuver, the robot has executed a 180° turn and has moved sideways by exactly one interrow distance. The turning is measured either by blind reckoning or by using the gyroscope. With blind reckoning, use is made of the fact that there is almost no slip between the large tires and the ground surface, and a prescribed distance is traveled using maximum steering. With the gyroscope, the robot drives until the gyroscope indicates that a 90° turn has been made.

Low-level control is performed by the microcontroller. The microcontroller program is a state machine written in C, using the PK-ARM development environment,[10] and runs under the FreeRTOS real-time operating system.[11] Among the signals it processes are heading, lateral displacement, and distance to end-of-row from the image-processing program running on the PC. The PC program runs under Windows XP[12] and is written in Delphi[13] and C++, using the DSPack[14] wrapper for accessing DirectX and the VXL image-processing library[15] for general image processing.

Detection of Volunteer Potato. The high position of the navigation camera means that the images taken with it do not have enough resolution to detect potato plants. Therefore, volunteer potatoes are detected using a second webcam,[16] which is mounted on the mast at a height of 0.5 m from the ground, aimed perpendicular to the direction of travel (i.e., toward the crop row) and downward at an angle of 45° to the horizontal. This camera also captures images with a resolution of 320×240 pixels.

Initially, we calculated the excessive green image and performed a Canny edge detection (Canny 1986) on it, using the number of pixels on edges as an indicator of the presence of potato plants. This method was used during the 2005 edition of the Field Robot Event, where three potato plants were successfully detected in a row of corn plants of about 6 m (unpublished data). Here, we report on an extended method of analysis. We used MATLAB,[17] extended with the DIPImage Scientific Image Processing Toolbox[18] (DIP*lib*, Delft image-processing library) for image processing and PRTools[19] (Van der Heijden et al. 2004) for statistical analysis.

Image Processing. It was determined that the approximately 0.3 m of row length covered by the images is large enough that a potato plant often fills only a small part of the image. Therefore, images were split in two halves of 160 by 240 pixels, and each half was processed separately. Six features were extracted from each half-image: (1) average red for plant-pixels, (2) average green for plant-pixels, (3) average blue for plant-pixels, (4) total number of plant-pixels in the binary image, (5) total number of contour pixels in the

[10] PK-ARM development software, Keil Software, Bretonischer Ring 15, D-85630 Grasbrunn, Germany.

[11] FreeRTOS real-time operating system. Web page: http://www.freertos.org.

[12] Windows XP, Microsoft, 1 Microsoft Way, #8, Redmond, WA 98052.

[13] Delphi, Borland Software, 20488 Stevens Creek Boulevard, Cupertino, CA 95014.

[14] DSPack free software, Progdigy Web page: http://www.progdigy.com.

[15] VXL image-processing library Web page: http://vxl.sourceforge.net.

[16] Webcam, PCV740K, Royal Philips Electronics, 1096 BC, Amsterdam, The Netherlands.

[17] MATLAB, The Mathworks Inc., 3 Apple Hill Drive, Building 3, Natick, MA 01760.

[18] DIPImage Scientific Image Processing Toolbox, Technical University, Julianalaan 134, 2628 BL Delft, The Netherlands.

[19] PRTools, Technical University, Julianalaan 134, 2628 BL Delft, The Netherlands.

Figure 1. The potato-detecting robot shown in a field of corn.

Figure 2. Steps in image processing for detection of crop rows. (a) Original image. (b) Thresholded excessive green image, in which white pixels correspond to green material and black pixels to background, with detected crop rows superimposed. (c) Hough-space image calculated from (b). Each pixel represents a line drawn from left to right in (b); the brightness of the pixel corresponds to the number of plant pixels under that line. The horizontal coordinate of the pixel corresponds to the intercept with the vertical axis of the line, and the vertical coordinate corresponds to the angle between the line and the horizontal. (d) Only the brightest pixels of (c) are retained. (e) Dilation of (d) results in groups of contiguous pixels, each representing a crop row in the original picture.

binary image, and (6) total number of edge pixels in the binary Canny image.

The features relate to color (feature 1, 2, and 3), size (feature 4), shape (combination of 4 and 5), and texture (feature 6). They are extracted in a number of processing steps, which are shown schematically in a flowchart (Figure 3), and results for two representative images are shown in Figure 4. Parameters for processing were chosen as follows. Parameter α determines the weighing of the red and blue channel in the excessive green image: $I = \alpha(G - R) + (1 - \alpha)(G - B)$. Depending on the camera and lighting conditions, α can vary between zero and one. For this camera, the blue channel was too variable to take into account, and hence, $\alpha = 1$ was chosen.

Parameter T is the threshold to distinguish between plant and background in the excessive-green image. This threshold was determined by looking at the segmentation result of several pictures; $T = 10$ seemed appropriate.

The number of edges in the green image is expected to be higher if it contains a potato plant because potato plants have a more clearly defined nerve structure than corn plants and because potato plants have a large number of small, overlapping leaves, whereas young corn plants have a few smooth leaves. The Canny edge detector (Canny 1986) was used to determine the edges in the green image. The background of the image is set to a fixed value (zero) before applying the edge detector to remove spurious edges in the background. The Canny edge detector has three parameters: σ, the width of the Gaussian blurring kernel; and the high and low hysteresis thresholds, which are used to remove weak edges while retaining connectivity of stronger edges. An edge must contain at least one pixel with a value above the high threshold, and all pixels of the edge are at least as large as the low threshold. The parameters used were $\sigma = 2$, lower threshold fraction = 0.5, and the higher threshold fraction = 0.9. The threshold settings were the default settings of the DIP*lib*-implementation of the Canny edge detector. The smoothing was slightly higher to account for the rather noisy images.

Image Classification. Each image was classified manually as either containing a potato plant (one) or not (zero). A Fisher linear-discriminant classifier (Hastie et al. 2001) was trained to optimally

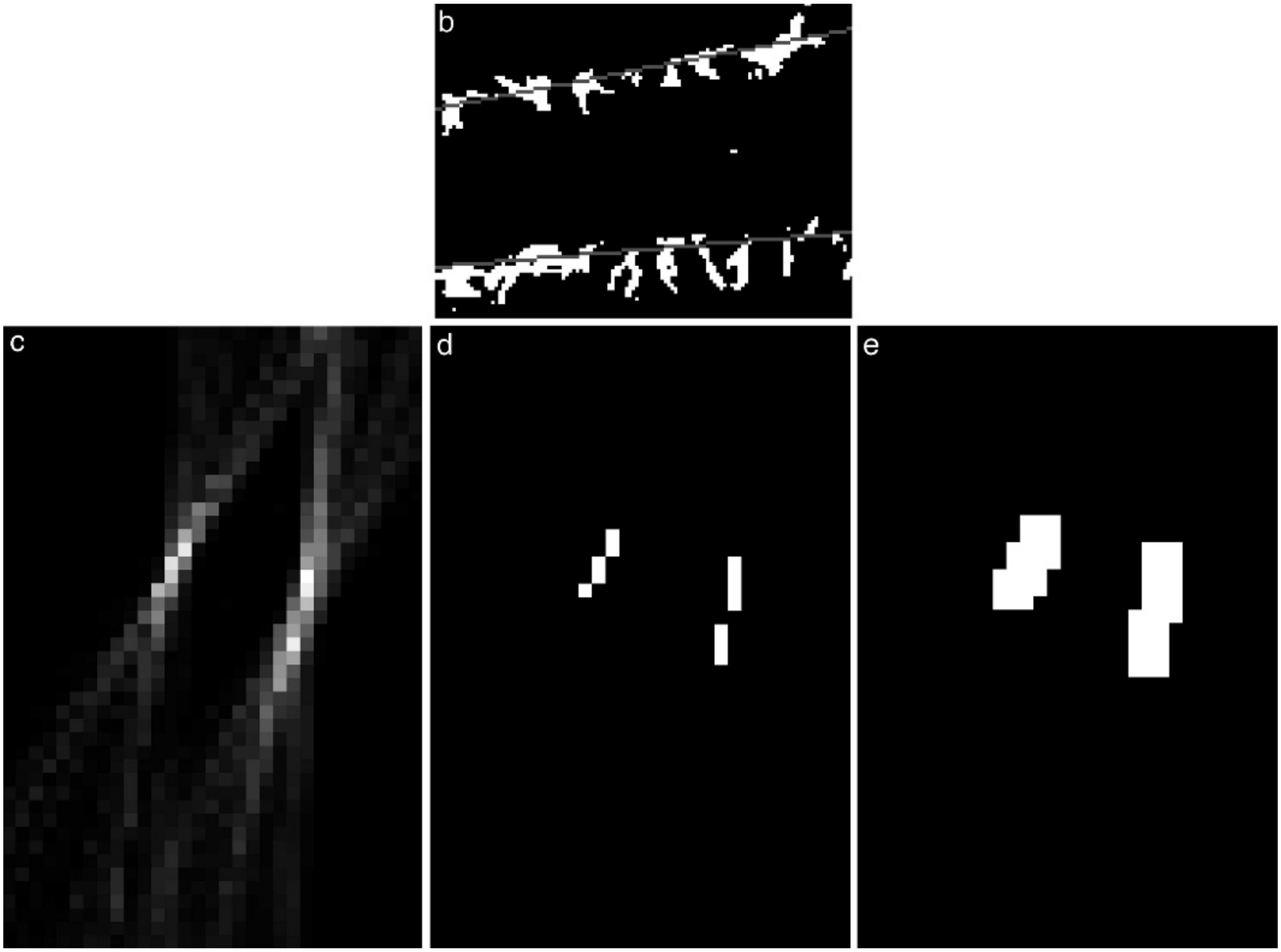

Figure 2. Continued.

distinguish between half-images containing potato plants and half-images without these plants. The error rate was determined using leave-one-out cross validation, i.e., all but one half-images were used to train the classifier, and the trained classifier was applied to this left out half-image. The predicted classification for this half-image was compared with the true (visually assessed) classification. This was repeated for all half-images. In this way an unbiased estimate of the error-rate can be obtained.

Data. Two sets of images were available to us:

1. 162 images taken on June 17, 2005, at the site of the 2005 Field Robot Event in Wageningen. Corn was sown on May 4. Images were taken at a rate of 2 images/s while the robot was moving autonomously at a speed of approximately 0.2 m/s. These were visually scored as 324 half-images.
2. 93 images taken on June 24, 2005, at an experimental field near Wageningen. Corn was sown on May 18. Images were taken at a rate of 2 images/s while the robot was moving autonomously at a speed of approximately 0.2 m/s. Both the corn and the volunteer potato plants were quite large. Some half-images had to be discarded when the robot navigated too close to the row and a single, large leaf covered most of the image. In total 141 half-images were used.

RESULTS AND DISCUSSION

Autonomous Navigation. Images were processed at or near the rate at which they are captured by the camera (30 frames/s). At this frame rate, use of the PC processor (as measured by the Windows Task Manager) is about 90%. The high frame rate makes

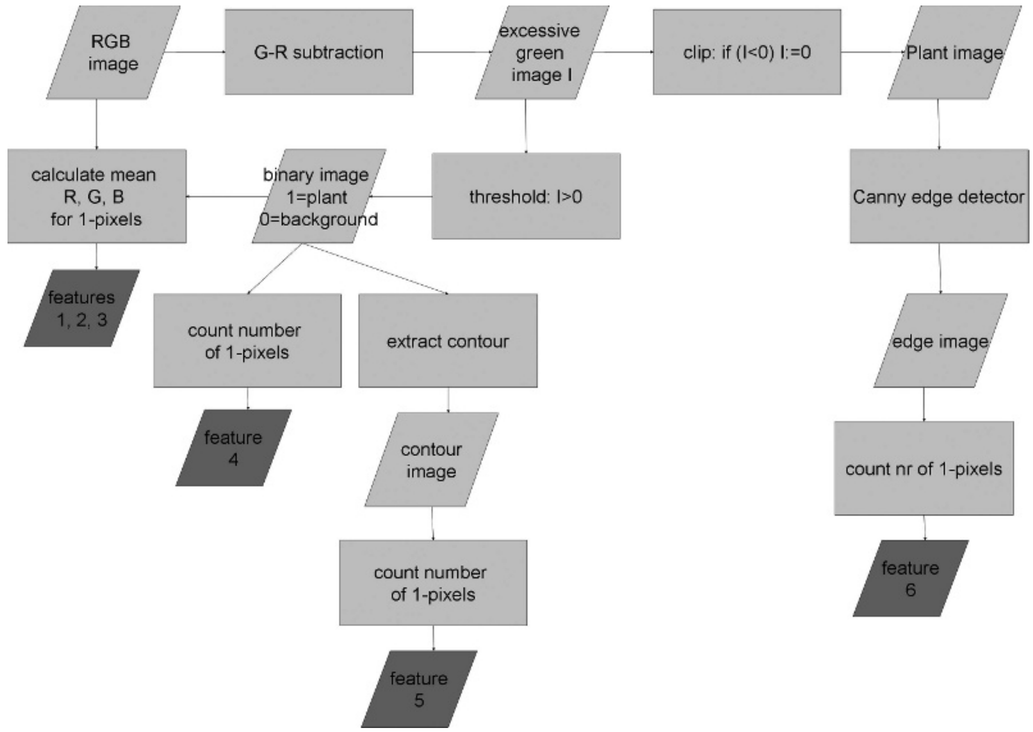

Figure 3. A flowchart outlining the steps in image processing for detection of potato plants.

it possible for the robot to travel fast and accurately through experimental and commercial cornfields. At the 2005 Field Robot Event, our robot covered slightly more than 120 m of sharply curving rows in 3 min, including three headland turns, for an overall average speed of 0.67 m/s.

The robot is small compared with the size of clods and other elements of roughness that are common even in well-prepared cornfields. As a result, the mast sways from left to right, often by 30° or more, but we observed little effect on the accuracy of navigation.

The high position of the navigation camera enabled the robot to operate through many hours of tests without encountering problems with dirt, even when dust was blowing. Likewise, for the potato detection camera, a distance of 0.5 m from the ground appeared to be sufficient to prevent problems with dirt.

Detection of Volunteer Potato. Results of the potato plant detection algorithm on the two data sets are given in Table 1. The overall error rate of the potato detection classifier was 1.5% for the first data set and 10.6% for the second data set. In some images in the first data set, the potato was not detected, but all corn images were correctly classified. However, some images classified as "no potato" contained only a small portion of the plant; in subsequent images with a larger portion of the same plant visible, potato was detected. In the second data set, misses and false detections occurred at about the same rate.

The data sets available were recorded on different fields and on different days. Corn and potato plants in the second data set were larger than the plants in the first data set. These differences between the data sets meant that a separate classifier had to be used for each data set, using the same six features. This means that, at present, calibration is needed. On the other hand, the plants in the second data set were larger than they normally are at the time of control.

It seems likely that our algorithm can be improved by including more information from the images. For example, Åstrand and Baerveldt (2002) used 19 features to identify sugarbeet plants among weeds. Alternatively, it may be useful to include one or more features extracted with other methods of image analysis. For example, the wavelet entropy used by Schut (2003) to determine the fraction of ground covered by clover in a grass–clover mixture might provide information about the presence of potato plants, too.

We used coarse measures of shape and color to detect potato plants; both have been expressed with more sophistication by other authors. Gerhards et al. (1993) and Åstrand and Baerveldt (2002) used shape

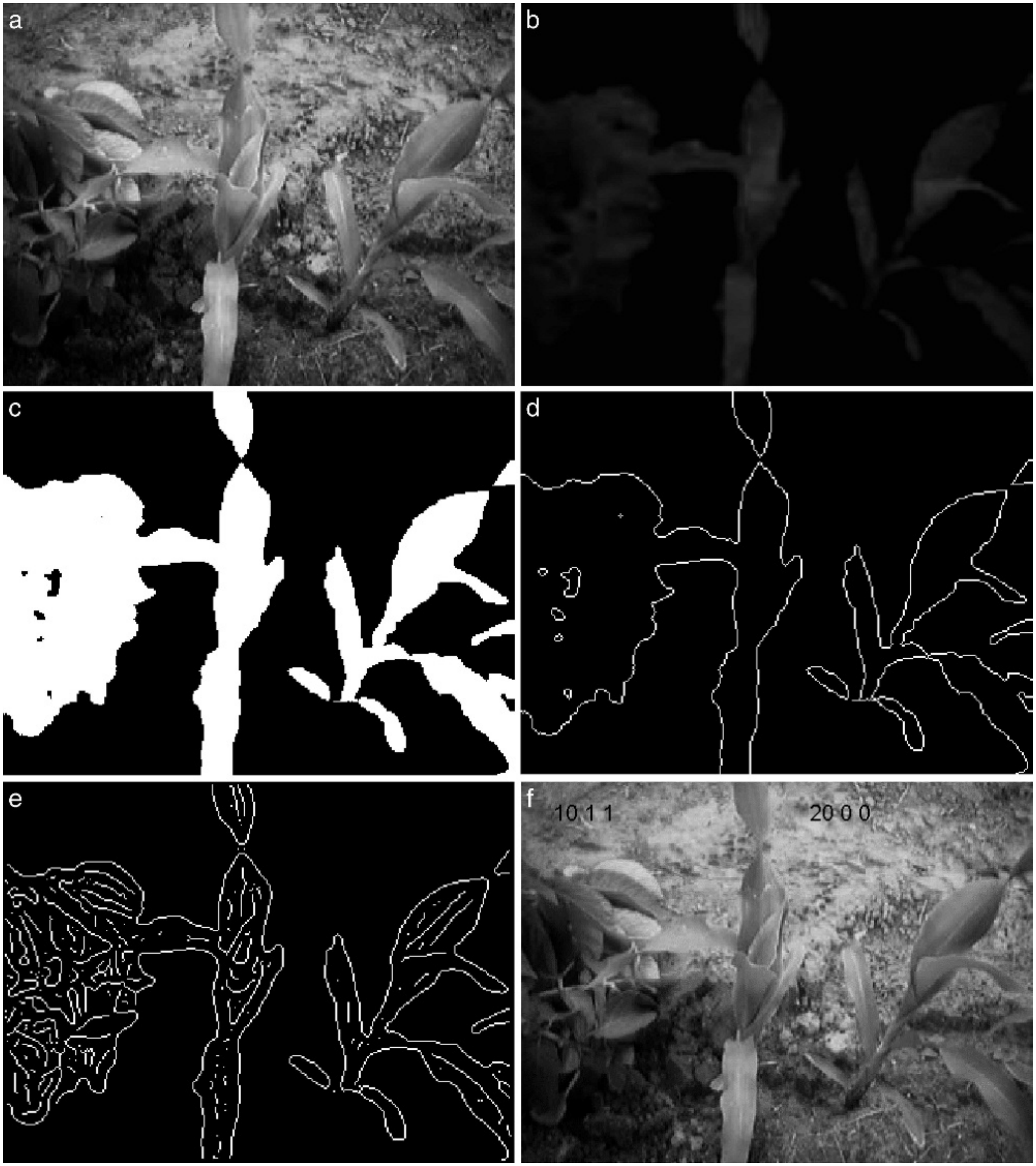

Figure 4. Steps in image processing for detection of potato plants. (a) Green image; (b) excessive green ($G - R$); (c) thresholded binary image, showing plants; (d) contour image; (e) result of the Canny edge detector; and (f) classification result: left subimage (nr 10) contains a potato (first one, visual score, second one, classification result), right subimage (nr 20) contains no potato (first zero, visual score, second zero, classification result).

to identify single weed plants against a soil background. We are not aware of reports where shape has been used to identify single potato plants. Shape is not useful in detecting weed plants growing in a row of crop plants when overlapping between plants occurs (Gerhards et al. 1995).

Color has received much attention as a means of identifying plants belonging to different species, but

Table 1. Classification results for the two data sets used in the study. Values shown are the number of images and the percentage of images classified correctly.

	Data set 1		Data set 2	
	Potato (109 images)	Corn (215 images)	Potato (56 images)	Corn (85 images)
Classified as potato	104	0	47	6
Classified as corn	5	215	9	79
Classified correctly (%)	95	100	84	93

in reviews by Zwiggelaar (1998) and Scotford and Miller (2005), it is concluded that it is next to impossible to distinguish plant species in field conditions based on differences in spectral reflectance. In work preliminary to this study, we were able to measure in the lab small, but significant, differences in reflection at red and infrared bands between leaves of young potato and corn plants; these differences were too small to have been picked up in the field by the low-cost camera we wanted to use.

We used a low-cost camera, in contrast to other authors. The quality of this camera was not determined, but it can be assumed that its optics produce significant distortions (projection, color shifts). Similarly, it can be assumed that its diffraction and sensing elements yield measurements of wide and poorly defined spectral bands. Furthermore, our images were Joint Photographic Experts Group (JPEG)-compressed, which caused loss in quality of color and structure. The accuracy and precision of the measurements, however, was limited more by the conditions encountered in the field than by the instruments themselves. Conditions encountered in the field include variable lighting, dust, and constant movement of both robot and plants. Many of the images showed motion blur, and some were incorrectly exposed. In view of the above, we were pleased that the potato-detection algorithm turned out to be robust enough to overcome these problems of image quality.

The potato-detection algorithm was fairly demanding computationally. We ran the detection program and the navigation program simultaneously on the same PC. The processing time of the detection program was such that it dropped the frame rate of the navigation program from 30 to about 8 frames/s. This low frame rate forced us to lower the speed of the robot from 0.67 m/s to approximately 0.2 m/s. By implementing these algorithms on separate processors (possibly, digital signal processors, [DSPs]), the impact of the detection algorithm on robot movement could be eliminated.

We have demonstrated the simultaneous implementation of two technologies needed to create a subcanopy robot to control volunteer potato plants in row crops: autonomous navigation and detection of potato plants. Significantly, the robot described uses low-cost cameras. Together with the availability of off-the-shelf robot bases offered by a number of companies, this brings the next step in precision weed management tantalizingly close. Although the current robot just uses its speakers to sound "potato!" whenever a potato plant is detected, future robots may be fitted with weed-control equipment, or they may use GPS to map the presence of weed plants so that regular equipment can be used for control.

ACKNOWLEDGMENTS

We thank Joop van Westeneng for access to the field in which the second data set was collected. We thank Jacques Davies, Ben Rutgers, and Gradus Leenders for technical assistance. The microcontroller development environment was provided free of charge jointly by Keil Software Inc., Grasbrunn, Germany, and Tritec, Sliedrecht, The Netherlands. The robot base and supplies were provided at a reduced price by Reyne and Sons, Krommenie, The Netherlands.

LITERATURE CITED

Anonymous. 2004. Proceedings of the 2nd Field Robot Event. Farm Technology Group. Wageningen, The Netherlands: Wageningen University.

Åstrand, B. and A. J. Baerveldt. 2002. An agricultural mobile robot with vision-based perception for mechanical weed control. Auton. Robot. 13:21–35.

Åstrand, B. and A. J. Baerveldt. 2003. A mobile robot for mechanical weed control. Int. Sugar J. 105:89–95.

Blackmore, B. S., W. Stout, M. Wang, and B. Runov. 2005. Robotic agriculture—the future of agricultural mechanisation? 5th European Conference on Precision Agriculture. Wageningen, The Netherlands: Wageningen Academic.

Boydston, R. A. 2001. Volunteer potato (*Solanum tuberosum*) control with herbicides and cultivation in field corn (*Zea mays*). Weed Technol. 15:461–466.

Canny, A. 1986. A computational approach to edge detection. IEEE Trans. Pattern Anal. 8:769–698.

Gerhards, R., A. Nabout, M. Sokefeld, W. Kuhbauch, and H. A. N. Eldin. 1993. Automatic identification of 10 weed species in digital images using Fourier descriptors and shape-parameters. J. Agron. Crop Sci. 171:321–328.

Gerhards, R., M. Sökefeld, W. Kühbauch, and A. Nabout. 1995. Using an expert system and digital image processing for site-specific weed control. 9th International Symposium on Challenges for Weed Science in a Changing Europe. Doorwerth, The Netherlands: European Weed Research Society. [In German.]

Hastie, T., R. Tibshirani, and J. Friedman. 2001. The Elements of Statistical Learning: Data Mining, Inference and Prediction. New York: Springer.

Hough, P. V. C. 1962. Method and means for recognizing complex patterns. USA Patent 3069654.

Reid, J. F., Q. Zhang, N. Noguchi, and M. Dickson. 2000. Agricultural automatic guidance research in North America. Comput. Electron. Agric. 25:155–167.

Schut, A. G. T. 2003. Imaging spectroscopy for characterisation of grass swards. Ph.D. thesis, Wageningen University, Wageningen, The Netherlands.

Scotford, I. M. and P. C. H. Miller. 2005. Applications of spectral reflectance techniques in Northern European cereal production: a review. Biosys. Eng. 90:235–250.

Tillett, N. D., T. Hague, and J. A. Marchant. 1998. A robotic system for plant-scale husbandry. J. Agric. Eng. Res. 69:169–178.

Van der Heijden, F., R. P. W. Duin, D. De Ridder, and D. M. J. Tax. 2004. Classification, parameter estimation and state estimation—an engineering approach using MATLAB. New York: J. Wiley.

Van Straten, G. 2004. Field Robot Event, Wageningen, 5–6 June 2003. Comput. Electron. Agric. 42:51–58.

Zwiggelaar, R. 1998. A review of spectral properties of plants and their potential use for crop/weed discrimination in row-crops. Crop Prot. 17:189–206.