Real-time image processing for the guidance of a small agricultural field inspection vehicle

Richard Gottschalk, Xavier Paolo Burgos-Artizzu* and Ángela Ribeiro

Artificial Perception Group (GPA),
Instituto de Automática Industrial (IAI),
Consejo Superior de Investigaciones Científicas (CSIC),
Carretera de Campo Real km 0,200,
28500 La Poveda, Arganda del Rey,
Madrid, Spain
E-mail: richardgottschalk@gmx.de
E-mail: xpburgos@iai.csic.es
E-mail: angela@iai.csic.es
*Corresponding author

Gonzalo Pajares

Dpto. de Ingeniería del Software e Inteligencia Artificial,
Universidad Complutense de Madrid (UCM),
Madrid, Spain
E-mail: pajares@fdi.ucm.es

Álvaro Sánchez Miralles

Escuela Técnica Superior de Ingeniería – ICAI,
Instituto de Investigación Tecnológica,
Universidad Pontificia Comillas,
Madrid, Spain
E-mail: alvaro.sanchez@iit.upcomillas.es

Abstract: This paper describes the image processing for an autonomous field inspection vehicle that uses a webcam for the navigation between two rows of agricultural crop. The relative vehicle position is calculated by segmentation and classification of the images and then by extracting geometrical lines corresponding to the crop rows. An autonomous vehicle was built and tested successfully in an agricultural environment.

Keywords: autonomous guidance; real-time image processing; precision agriculture; weed mapping.

Reference to this paper should be made as follows: Gottschalk, R., Burgos-Artizzu, X.P., Ribeiro, A., Pajares, G. and Miralles, Á.S. (2010) 'Real-time image processing for the guidance of a small agricultural field inspection vehicle', *Int. J. Intelligent Systems Technologies and Applications*, Vol. 8, Nos. 1-4, pp.434–443.

Biographical notes: Richard Gottschalk studied Electronic Engineering in Karlsruhe, Germany. He took part in a special European programme called 'Tripartite' in which he studied two semesters at the 'Ecole Supérieure des Ingénieurs en Electronique et Electrotecnique' (ESIEE) in Paris and finished his studies in the following two semesters at the Industrial Automation Institute (Spanish Council for Scientific Research – CSIC) in Madrid, Spain. This paper is a result of his final thesis which consisted of developing and testing the hardware and software for an autonomous field inspection vehicle based on computer-vision algorithms using a simple webcam.

Xavier Paolo Burgos-Artizzu is a PhD Student working at the Industrial Automation Institute (Spanish Council for Scientific Research – CSIC) in Madrid, since 2006. He received his BSc and MSc in Computer Engineering from the Complutense University of Madrid in 2005 and 2007, respectively. His main research interests are machine visual perception, robotics and machine learning. His work has been published in journals and conference proceedings in computer vision, computational intelligence and robotics.

Ángela Ribeiro is a Staff Scientist at the Industrial Automation Institute (Spanish Council for Scientific Research – CSIC) in Madrid, since 1990. She received her MSc and PhD in Physics from the Complutense University of Madrid in 1984 and 1989, respectively. Her current research interests include machine visual perception, pattern recognition, spatial-temporal reasoning and genetic algorithms applied to the design and development of perception-planning-action systems. She has been participant and leader of several European and Spanish Research Projects related to all the previous topics. Her work has been published in journals, books and conference proceedings in computer science, computer vision, engineering in agriculture and robotics.

Gonzalo Pajares received MSc and PhD in Physics from UNED (distance University from Spain) (1987, 1995) discussing a thesis on the application of pattern recognition techniques to stereovision. He was working in Indra Space and INTA developing remote sensing applications. He joined the Complutense University in 1995 as an Associated Professor and from 2004 as a Professor at Full Time on the Faculty of Informatics in the Department of Software Engineering and Artificial Intelligence. The areas covered are computer vision and artificial intelligence. His current research interests include machine visual perception, pattern recognition and neural networks.

Álvaro Sánchez Miralles is a Professor of Mobile Robotics at the Engineering School of Pontificia Comillas University (UPCo) in Madrid. He obtained his MSc and PhD in Industrial and Electrical Engineering from Pontificia Comillas University in 1998 and 2002, respectively. He joined Instituto de Investigación Tecnológica at UPCo (IIT-UPCo) in 1998. He has a large experience in industry joint research projects in the field of artificial intelligence. His areas of interest are artificial intelligence applied to diagnosis systems, electricity networks and mobile robotics, simulation, wireless communication and electronic and control systems.

1 Introduction

The economic demands of reducing tariff protection for first world farmers, shortages of skilled farm labour in agricultural regions, food and fibre requirements of a growing world population, and increasingly stringent standards for agricultural production, will continue to drive the commercial need for fully autonomous robotic devices within agricultural environments and ensure this becomes an increasingly exciting field of endeavour (Stafford, 2000; Thorp and Tian, 2004).

Precision agriculture (PA) is a concept that addresses the field in variability of factors that influence crop growth. It seeks to avoid applying same management practices to a crop regardless of site conditions (Kropff et al., 1997). The most significant benefits of site-specific management are cost reduction in the production of the crop and reduction in environmental pollution (Earl et al., 1996). In order to accomplish a PA practice, it is essential to generate a treatment map showing the degree of weed coverage present in each area of the field.

Recently, the GPA (Artificial Perception Group) of the IAI-CSIC in Madrid (Spain) has developed an image processing platform that enables the percentages of weed, crop and soil present in a image to be computed (Burgos-Artizzu et al., 2007, 2008). This platform, in conjunction with an autonomous agricultural field inspection vehicle would provide the necessary technology to automatically construct a weed risk map, considerably facilitating the task that such map construction usually involves. The vehicle should be able to move autonomously across agricultural fields, taking high-resolution photos in sampling points defined by the user, for the generation of a weed distribution map. The autonomous guidance in agriculture and crop row detection has been an active field of research for some years now (Billingsley and Schoenfisch, 1997; Hague and Tillet, 2001; Sogaard and Olsen, 2003).

This paper presents the vehicle developed at the IAI-CSIC for inspection of agricultural fields and describes the vision algorithms implemented that allow the detection of crop rows in real time, adjusting the vehicle movement for the autonomous navigation between rows.

The rest of this paper is structured as follows. Section 2 describes the vehicle and the system mounted on-board. Section 3 describes the real-time image processing. And finally, Section 4 outlines the conclusions and future work.

2 System description

The used vehicle was the Traxxas E-MAXX electric-driven model truck, Figure 1(a). Approximately, it is 52 cm long, 42 cm wide and has a ground clearance of 10 cm. The vehicle is powered by two 7.2 V battery packs and has two main driving motors and two steering motors. It is a four-wheel-driven which enables the vehicle to move through rough terrain. All the important steering and drive chain parts which need to withstand aggressive movement are made up of metal http://www.traxxas.com/products/electric/emaxx3905/trx_emaxx3905.htm.

The only industrial component used for the entire vehicle was the on-board PC/104 computer system including its power supply. The variety of connections of the PC/104 was a big advantage for connecting peripheral devices such as the camera, the motor controller, an external hard drive and a wireless access point used for external vehicle supervision and remote control. Figure 1(b) and (c) shows the final vehicle and its hardware set-up.

3 Image processing

A simple 8-bit colour webcam could be used for orientation purposes since no high resolution image is required. Typically, the only problem is that webcams have a viewing angle of less than 50^o which would make it necessary to mount the camera at an approximate height of 150 cm. Therefore, the ‘Live! Cam Notebook Ultra’ developed by Creative was chosen. It has a viewing angle of approximately 80^o, which allowed to mounting the camera at an height of only 70 cm.

3.1 Image segmentation

In order to create a reliable segmentation of vegetation against non-vegetation algorithm, the original RGB image was converted into an HSI image which made it easier to analyse hue and saturation values separately. The first step was to filter all those parts of the image which had the same colour (i.e. hue value) as plants of maize, taking into account that the colour perception of the camera could vary greatly depending on the lighting. To distinguish between plants and background a reference histogram was created from an image showing only a plant.

Figure 2(a) and (b) shows an original image of plants of maize and the colour histogram of the entire image. The histogram shows that most of the pixels have an orange hue value due to the soil which covers most of the image. The second largest colour occurrence is green deriving from the plants of maize.

Figure 2(c) shows the colour histogram which can be obtained by analysing a small region showing only plants (black square region in Figure 2(a)). The plants included in the region mainly contain similar colour values but varying in saturation. In order to successfully classify all plants as vegetation, this histogram can be applied to the original image. This is done following Equation (1). All pixels with values not contained in the histogram (hist = 0) will be set to zero (black), while all other pixels will be set to different shades of grey accordingly to the value of the histogram for that value, ranging from 1 to 255 (8 bit image). The result is shown in Figure 2(d). Only those parts that appeared green in the original image have values other than black.

$$\begin{gathered}\forall i, j \in \text{rows, columns} \\ \text{value}(i, j) = (\text{hist}(\text{value}(i, j)) \times 255) / \text{MAX}(\text{hist})\end{gathered} \tag{1}$$

Figure 3 shows how flexible this method is with respect to lighting conditions and changing camera perception. The image was taken with the mounted vehicle camera during one of the vehicle tests. Although the grass appears to be green in the original image the histogram of the region of interest shows that it mainly contains shades of yellow. Filtering the entire image with this histogram makes the grass region appear in light greyscale values. Although the rest should appear black there is still a significant amount of bright pixels in the non-grass region. These pixels in fact have one of the colours represented in the histogram but what makes them appear so different to the human eye is their brightness and saturation value. Therefore, false detection can be further limited by applying a simple threshold filter on the saturation and brightness channels.

After applying the histogram method on the image, plants are represented by different shades of grey and everything else should be black. However, the aim of segmentation is to have a binary image showing regions of interest as white areas on a black background. The conversion to binary values can be easily obtained by applying a threshold function. Although a simple threshold function converts the greyscale image into a binary image, the edges of the green regions will not be clearly defined. Single and isolated white pixels should be erased, while the parts belonging to the crop (i.e. big agglomerations of white pixels) should be amplified. This was done by using morphological imaging functions such as *erosion* and *dilation*.

Figure 1 (a) The electrically driven Traxxas E-MAXX (b) the vehicle in its final state in the testing field and (c) the hardware set-up (see online version for colours)

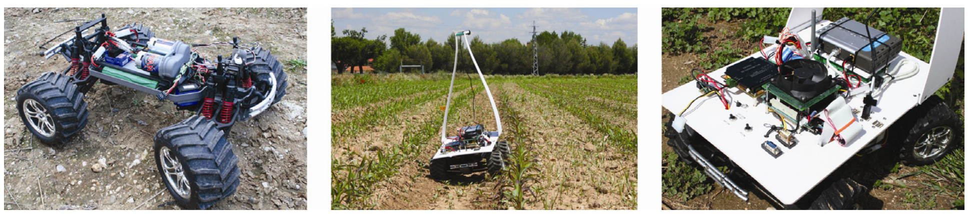


Figure 2 The use of a colour histogram for classifying objects can be easily adapted to varying lighting conditions: (a) original image with the region for the classification histogram; (b) the normalised colour histogram of the entire image; (c) the normalised colour histogram of the classification region and (d) applying the classification histogram (see online version for colours)

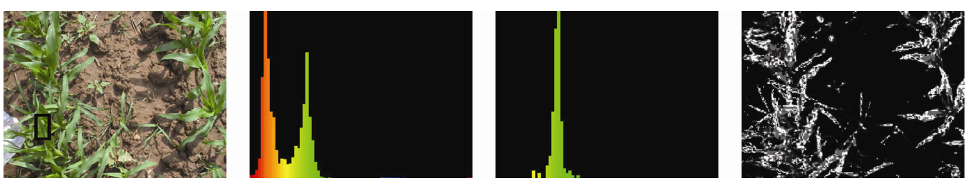
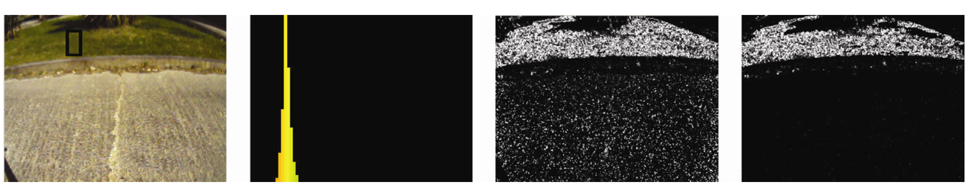


Figure 3 Depending on the lighting, the colour histogram may appear very different from what one would expect: (a) the original image with the region of interest; (b) the normalised colour histogram of the entire image; (c) the normalised colour histogram of the classification region and (d) applying the classification histogram (see online version for colours)



Note: In the original image, the grass appears to be green but the histogram shows that the camera detects it as shades of yellow. If a saturation and brightness filter is not applied in addition to the histogram, many pixels of the concrete pavement will be classified as plants.

3.2 Finding non-parallel lines using horizontal segments

After having performed the segmentation, the next step was to extrapolate crop rows. A very common method used for line detection in an image is the Hough transform, explained detail in Gonzalez and Woods (2003). However, a major disadvantage is the complexity of the algorithm which makes it too time-consuming for real-time applications.

Considering the images are taken as overhead views, crop rows after segmentation will appear like two almost vertical white columns. Therefore, crop rows can be approximated by computing the average vertical values of the white pixels. To reduce processing time, instead of computing the average for each image row, the image was divided in horizontal slices. Although best results were achieved using a fairly large number of slices this increased the processing time remarkably and tests showed that four or five slices were sufficient to achieve a good approximation of the lines. Figure 4(a)–(c) shows an example.

Having defined the points of highest probability of plant occurrence the set of points for each crop row need to be used for calculating an approximation for the equation of a straight line $y = m \times x + b$. Two methods were tested in order to compare their processing time and results.

3.2.1 Top and bottom mean values

The simplest and quickest way of computing a line equation is to use only twopoints. The 'top and bottom mean values' method computes the average x and y values of the higher and the lower set of points and uses the two resulting points to compute the line equation. The number of points defining each set can be chosen arbitrarily from one to $n/2$ using the parameter depth, where n is the total number of extracted points (number of slices in which the image was divided in the previous step). In the example image, Figure 4(d), a *depth* of $n/2$ was used. This means, that for the left crop row, the first point will be computed as the average x and y values for the three higher black points, and the second point as the average of the lowest three black points. For the right crop row, the same operations are performed using the yellow points. If the variable *depth* is chosen to be one only, only the highest and lowest point would be used. Reducing the *depth* makes this method slightly faster, but more exposed to false line recognition if false positives appear due to weed occurrences (see Section 3.3). Once both points computed ((x_1, y_1) and (x_2, y_2)), the line parameters are calculated as follows:

$$m = \frac{y_2 - y_1}{x_2 - x_1} \quad b = y_1 - mx_1$$

3.2.2 The least squares method

This method is a classical mathematical approach but it is also slightly more time-consuming. Its aim is to find a straight line in such a way that the sum of the distances between each point and the line is minimum. In order to avoid difficulties due to the sign of the distances each of them is squared before adding it to the sum. The line parameters are defined by:

$$m = \frac{\sum_i (y_i - \overline{y})(x_i - \overline{x})}{\sum_i (y_i - \overline{y})^2} \quad b = \overline{x} - m\overline{y}$$

For these calculations two loops with n steps need to be programmed: one for calculating the average values $\overline{x}$ and $\overline{y}$ and one for the sums in the fraction. The extra processing effort is negligible and therefore this algorithm can also be considered suitable. However, tests showed that the results were not necessarily much better than the one obtained using the method first described. Figure 4(d) shows a comparison.

3.3 *Statistic elimination of false recognition*

Using the previous methods, weed occurrences can lead in some cases to a false recognition of the crop rows, due to apparition of false crop points. These false positives can be easily identified because they normally tend to be located far from the other detected crop points. The single dot shown in Figure 4(a) represents an example. If this type of weed is located in line with a crop row, its location cannot be used to define it as a false positive. Although the dot in the example is obviously a false positive if you look at the segmented image, in image processing, however, this is not a trivial task.

Having performed the previous steps, such a false positive could be identified by measuring its distance to the computed line. If the distance is greater than a certain threshold it could be classified as an erroneous point and eliminated from the set. This would have to be done for each point, and as soon as a false positive was detected the line would have to be recalculated and the check would need to be restarted. Since this procedure was too time-consuming a similar but less accurate method was chosen.

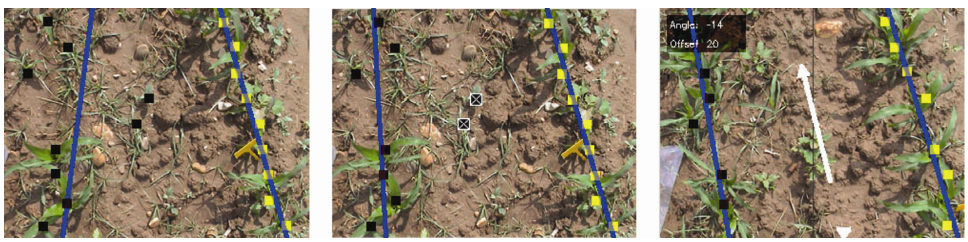
The extracted lines were assumed to be more or less vertical. Therefore, the shortest connection to a false positive (such us defined by its distance to the line) would have a much larger x component than the y component. This reduces the problem to one dimension since only the x component of the false positive is important. Its x value can be then compared to the average x value of the extracted line. If the difference is greater than a certain threshold it can be classified as a false positive and eliminated from the set.

Obviously, this method works best if the line is parallel to the y-axis since the average x value of a line is the exact x value for every point on the line and the distance to a false positive is exactly the difference between its x value and the average x value of the line. Figure 5 shows an example where the threshold value was chosen to be 60 pixels. Although, the left line of a crop ends at about half of the height of the image some weed occurrence results in a total of five false positives in the image. By using the previously explained method the two points on the right side are eliminated from the set and are marked with a white cross in the resulting image. The lines are recalculated and a better approximation is obtained. The three false positives at the top cannot be eliminated using this method. Fortunately, they do not affect the angle of the approximated line. But if it was necessary to recognise the end of the line of plants a better algorithm would need to be developed.

Figure 4 Approximation of non-parallel lines: (a) binary image after segmentation; (b) the vertical average pixel values for five horizontal segments; (c) applying a threshold filter; (d) the centres of the widest blocks on the left and right side are marked with black and yellow squares, respectively in the original image (see online version for colours)

Note: Those points were used to approximate the linear equations, using the 'least squares' (white) and the 'top and bottom average' (blue) methods.

Figure 5 Eliminating false positives and calculation of yaw angle and offset: (a) five false positives on the left side; (b) filtering with a threshold value of 60 pixels eliminated two false positives due to heir position (they are marked with a white cross) and (c) visualisation of the vehicle displacement (white triangle) and pathway angle (white arrow) (see online version for colours)



3.4 Calculation of yaw angle and offset

Having optimised the segmentation and extracted two lines represented by linear equations all that were left to do was to calculate the relative position of the vehicle between the lines. This last step of the image processing function need to reduce all acquired information to a minimum necessary for controlling the vehicle. The most important information for a controller in general is the current state of the system, which in this case is its position, since the main task was navigating the vehicle autonomously.

The controller needed to manage two types of localisation.

1 The global position on the field, which will be a future task once the GPS system installed.

2 The relative position of the vehicle between the crop rows.

Latter it can be computed using the on-board camera. The relative position is defined by two values: the displacement (offset) between its centre and the middle point between the crop rows and the angle of its direction of movement. Any other values the controller might need in the future such as vehicle speed or steering angle could not be deduced from image processing information.

Before calculating the offset and angle values the question was how these values could be defined concerning the information that had been extracted from the images. Since the camera was mounted centred at the front of the vehicle, at a height of about 70 cm and in such a way that the perspective was almost vertical to the ground, it was assumed that the angle of the moving direction was the negative mean value of the angles between the extracted lines and the y-axis of the image, taking as reference angle (zero degrees) that of a vertical line. Moreover, considering that the camera is mounted centred on the front of the vehicle a vertical centred line in the image would represent the extended axis of the vehicle in the direction of movement. Therefore, the offset could be approximated by the distance between this vertical middle axis of the image and the mean x value of the extracted lines at the bottom of the image. Figure 5(c) shows a visualisation of these values.

4 Conclusions

PA is a very important concept to reduce environmental pollution in farming and at the same time increase the efficiency. In this paper, the development of the image processing for an autonomous field inspection vehicle used for generating a global map of weed occurrences in an agricultural field is presented.

To perform the segmentation of vegetation against non-vegetation in uncontrolled outdoor conditions, a process based on a colour histogram was used. This procedure analyses the colour spectrum of a target area – a section showing only plants can be selected for calibration – and extracts those sections of the original image where the colour spectrum concurs. This approach takes into account the hue, saturation and lighting value and is very flexible because processing parameters can be altered very simply.

Having used segmentation for converting the image into binary data (plants/no plants) the next step was to extract geometrical equations for lines representing the rows of plants shown in the image. Assuming that the camera was pointed almost vertically to the ground from a height of 70 cm in front of the vehicle and that the vehicle was placed in the centre between the lines initially, two rows would appear as more or less vertical agglomerations of white areas on each side of the image. Their statistical centres were found by first calculating an average pixel value for each column of the pixel grid and applying a threshold filter to preserve binary data. This resulted in various vertical lines of different width. The widest line on each side of the image was chosen to be the approximate centre of the rows of plants. Taking into account that, when in action, an angular displacement was probable, this method was not applied to the entire image at once. Instead, it was divided into several horizontal segments each of which was then processed one by one resulting in two vertical lines representing the centres of the crop rows for each segment. The lines of each segment were reduced to their middle point and the resulting set of all points was a statistical approximation of two non-parallel lines. In order to find the geometrical equations two methods were compared: the ‘least squares’ method and using the ‘top and bottom mean values’. The results were very good and equally satisfying in respect to processing time and approximation error. The two equations were then used to calculate the vehicle’s offset and angle between the rows of crop.

Initial tests in the field have proven that the vehicle is successfully able to move autonomously among the crops, and therefore, main future work tasks are the inclusion of a GPS system, and the mounting of the high resolution photo camera in order to take pictures at the sample points and being able to automatically generate a weed coverage map.

Acknowledgements

The Spanish Ministry of Education and Science (MEC) and the CAM provides full and continuing support for this research work through projects AGL2005-06180-C03-03 and PRICIT-CAM-P-DPI-000176-0505(ROBOCITY2030). Currently, the second author holds a PhD scholarship given by the Spanish National Research Council (CSIC).

References

Billingsley, J. and Schoenfisch, M. (1997) 'The successful development of a vision guidance system for agriculture', *Computers and Electronics in Agriculture*, Vol. 16, pp.147–163.

Burgos-Artizzu, X.P., Ribeiro, Á. and de Santos, M. (2007) 'Controlador borroso multivariable para el ajuste de tratamientos en agricultura de precisin', *Revista Iberoamericana de Automática e Informática Industrial*, Vol. 4, No. 2, pp.64–71.

Burgos-Artizzu, X.P., Ribeiro, Á., Tellaeche, A. and Pajares, G. (2008) 'Optimisation of natural images processing using different evolutionary algorithms', *Proceedings of 2008 IEEE Congress on Evolutionary Computation*, ISBN: 978-1-4244-1823-7, Hong Kong, June.

Earl, R., Wheeler, P.N., Blackmore, B.S. and Godwin, R.J. (1996) 'Precision farming: the management of variability', *Landwards*, Vol. 51, No. 4, pp.18–23.

Gonzalez, R.C. and Woods, R.E. (2003) *Digital Image Processing*, 2nd edition, Prentice-Hall.

Hague, T. and Tillet, N. (2001) 'A bandpass filter-based approach to crop row location and tracking', *Mechatronics*, Vol. 11, pp.1–12.

Kropff, M.J., Wallinga, J. and Lotz, L.A.P. (1997) 'Modelling for precision weed management', *Precision Agriculture: Spatial and Temporal Variability of Environmental Quality*, Chester: Wiley, pp.182–204.

Sogaard, H. and Olsen, H. (2003) 'Determination of crop rows by image analysis without segmentation', *Computers and Electronics in Agriculture*, Vol. 38, pp.141–158.

Stafford, J.V. (2000) 'Implementing precision agriculture in the 21st century', *Journal of Agricultural Engineering Research*, Vol. 76, pp.267–275.

Thorp, K. and Tian, L. (2004) 'A review on remote sensing of weeds in agriculture', *Precision Agriculture*, Vol. 5, pp.477–508.