

Sensors-Based Stereo Image System for Precision Control of Weed in the Agricultural Industry

Bruno M. Moreno^{1,2}, Paulo E. Cruvinel^{1,2,3}

¹ Embrapa Instrumentation (CNPDIA), P.O. Box 741, 13560-970, São Carlos, SP, Brazil

² Electrical and Computer Engineering Program, University of São Paulo, São Carlos, SP, Brazil

³ Computer Science Program, Federal University of São Carlos, São Carlos, SP, Brazil

Emails: bruno.moraes.moreno@usp.br; paulo.cruvinel@embrapa.br

Abstract—This contribution proposes a solution for agricultural weed control based on stereo digital image sensors domain-oriented devices, technologies, and applications. In the world, agriculture has been developed by combining the production, value aggregation, environmental and social responsibility. The sector is primarily responsible to supply food for people, as well as fibers and energy. To keep such results, farmers have faced the need to seek, increasing the rational use of inputs, as is the use of pesticides, plant regulators, and liquid fertilizers. This paper presents a discussion related to the design and development of stereo image sensors for precision spraying to control weed species in agricultural crops, i.e., based on advances in sensor's instrumentation and image processing. The yield of a crop can vary depending on the species of invasive plants involved, its percentage of occupation by area, competition period, stage of development of the crop and soil, as well as weather conditions. In this context, the selection of an adequate image sensor can result in better recognition process, i.e., improvement in sprayer quality at variable rates based on management zones.

Keywords—Camera sensors; Embedded platform; Stereo vision; Image sensor; Decision making; Weed control; Agricultural industry.

I. INTRODUCTION

The current agriculture must face the challenge of increasing production in response to the demand of the growing population. Based on this directive and under the precision-based management, an increasing use of remote sensing technologies have been used to attend the rural areas. Additionally, embedded instruments and analog and digital cameras are found frequently, as well as series of computer modelling to aid decision-making based on the site-specific management, which has brought improvements related to the management systems that can promote the rationalization in terms of the application of inputs, decreased production costs, and impacts on the environment [1] [2]. Among the steps aimed at rationalizing the use of agricultural inputs, one may find the application of pesticides, which has required major efforts since its processes are related not only to the treatment of the pests into a cultivated area, but also related to the care about its possible impacts in both the cultivated area, as well as in adjacent areas [3]. The idea related to the technification of the agricultural processes to improve production methods has been gaining ground in recent decades. Real-time process monitoring and control, along with advances in positioning systems, can provide more details of the fields of production and therefore can improve the decision-making based on variability of the processes in agricultural production. The

agricultural production chain starts from the adaptation to the land to the sale of the product in the supermarket. Within this chain, there are several involved processes that have important relevance in order to obtain the minimization of environmental impact, better efficiency in the amount of production and quality of the final product.

The clearest example of these processes is the application of agricultural pesticides in pulverized form, in which is necessary to improve safety and efficiency of application, for agricultural pest control. The spraying of agricultural pesticides is used in most large-scale production crops and requires precision and effectiveness to avoid the impact it may have on the soil, the crop and the environment, if not applied in a responsible way and with the highest quality standards. To obtain an efficient agricultural spraying the factors, such as efficiency of the applied chemical, factors which come from the weather conditions, biological factors and the quality of the application must be considered. Thus, several studies are based on the search for quality and efficiency of the chemical products [4]. Also, there is a large number of scientists working on biological and climatic factors which can affect the quality and efficiency of the production. On the other hand, several works show the importance of the application, from the point of view of the used methods, the machinery and the automation of the application processes. In addition, technological advances and the evolution of embedded systems to aid decision-making processes have provided to agriculture a new way of seeing the rural property, where the consideration of spatial and temporal variability of soil and plants have improved the concepts related to the risk management in agriculture, i.e., the way to find efficiency and cost minimization in the use of agricultural inputs. Furthermore, in this context, and taking into account advanced trends, sensors-based technologies can provide appropriate tools to achieve solutions and to implement the strategies above mentioned. In precision agriculture area, sensor-based technologies play an important role.

In this context, the precision agriculture area has been seeking sensor-based technologies to improve agricultural processes, and such field of knowledge is playing today an important role. One technology with the use of camera sensors that adds more information to decision making in the agricultural field is the one based on stereo vision methods. These methods can be used for various applications, and there are studies that analyze the vision at night for the development of an advanced driver assistance system, monitor the conditions of long pipelines with autonomous robots, and assist in the

reconstruction of virtual plant leaf model [5]–[7]. This paper presents a method based on the use of a customized stereo image sensor for the precision application of herbicides to control weed in an agricultural field. The rest of the paper is structured as follows. Section II presents the mathematical theory behind the selection of the image sensors, based on the transfer functions. Section III discusses the stereo vision methods and the materials used. Section IV presents the results and discussions. Finally, we conclude our work in Section V.

II. THEORY BEHIND THE SELECTION OF IMAGE SENSORS FOR AGRICULTURAL WEED CONTROL

The camera function has been widespread applied for various circumstances [8] [9]. Lens and cameras designers encounter challenges for developing systems with high image quality. The major concern is how to optimize the lens parameters such as curvatures and thicknesses to get high image quality, for example, a high image resolution. Several optimizations have been proposed to improve the aberrations of lens systems [10]–[12]. The Modulation Transfer Function (MTF) is the amplitude term of Optical Transfer Function (OTF) that is similar to the transfer function of linear system. The transfer function is regarded as a major characteristic in the linear system. There are some proposals that present their methods to approach the transfer function [13]–[15]. A simple one is to receive the impulse response at output as input being an impulse signal. This impulse response is related to the transfer function. Using the same procedure, a point source is respected as the impulse signal to help estimate the image response in a lens system. The image of point source shown in the image plane is called the point spread function (PSF), which is the inverse Fourier transform of the OTF. Therefore, the MTF applied to determine the image resolution could be derived from the amplitude term of the Fourier transform of the PSF. MTF is in principle classified into three categories: scanning method, autocorrelation, and crossrelation methods. Of these, the scanning method seems to be most commonly used in many fields, and the Line Spread Functions (LSF) measurement is equally preferable because it can be obtained with a simple slit, and it is measure of fundamental definition.

The MTF can be evaluated from either the geometrical optics or diffraction calculation. Ray tracing methods are widely applied to simulate the image response of object point source in the image system. Using the irradiance model, Lin and Liu [16] presented a MTF computation method without counting the number of rays traveling to each grid. Tseng et al. [17] proposed skew ray tracing method to simulate the geometrical PSF, i.e., using a homogeneous coordinate transformation matrix.

In this paper, the use of a camera in agricultural environment can be defined taking into account the calculation of the LSF of the camera lens and MTF, which represents the magnitude response of the optical system to sinusoids of different spatial frequencies, i.e., retrieved by Fourier transform of the LSF. Several key aspects of optical instrumentation are related with the implementation of a linear source for a given optical system, the impact of the finite size of the source on the measurement, and the choice of the optical elements to image the response of specific patterns of plants and its relation with the lens used in the camera. By taking a linear source the solution to measure the MTF is in 1D, orthogonally to the

direction of the line. This can be proven considering a given source $S(x, y) = \delta(x) \cdot C$, and a lens of diameter equal to (a), that means:

$$R(k_x, k_y) = \int \int_{-a/2}^{a/2} \delta(x) C e^{j(k_x x + k_y y)} dx dy \quad (1)$$

The response of the objective can be expressed as the square of the Fourier transform of the product of the source with the aperture of the lens $R^2(k_x, k_y)$, with (k_x, k_y) the spatial frequencies associated with the spatial coordinator (x, y) . Besides, looking for a solution of (1) and solving the integral by parts is possible to reach:

$$R^2(k_x, k_y) \propto \frac{\sin^2(ak_y)}{(ak_y)^2} \quad (2)$$

Equation (2) corresponds to the LSF. The Fourier Transform of the LSF then gives the 1D MTF in the yy-direction.

III. MATERIALS AND METHODS

A. Stereo Vision

The stereo vision system used in this work is based on the use of two cameras with the aim of simulating the human vision system and obtain depth of objects, with the camera plane as a reference. The depth is resulted through the comparison of the objects position between each captured image.

The simplest way of comparing the images is guaranteed when the cameras are coplanar and aligned, as shown in Figure 1. The variables defined by the camera are the baseline b and the focal distance f . The $P(X, Y, Z)$ represents a point that would be recorded by the two cameras and $u_L = (X_L, Y_L)$ and $u_R = (X_R, Y_R)$ are the projections of this point in each image. From the concepts of geometry and similarity of triangles, it is possible to obtain:

$$\frac{X_L}{f} = \frac{X}{Z} \Rightarrow X_L = \frac{Xf}{Z} \quad (3)$$

$$\frac{-X_R}{f} = \frac{b - X}{Z} \Rightarrow X_R = \frac{(X - b)f}{Z} \quad (4)$$

$$Y_L = Y_R = \frac{fY}{Z} \quad (5)$$

Considering that the difference between the x coordinates in each image is equal to $d = X_L - X_R$, where d is the disparity, (3) and (4) can be rewritten as:

$$X = \frac{bX_L}{d} \quad (6)$$

$$Y = \frac{bY_L}{d} \quad (7)$$

$$Z = \frac{bf}{d} \quad (8)$$

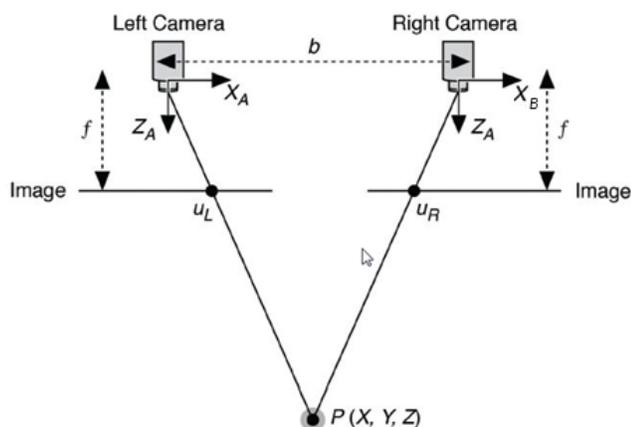


Figure 1. Stereo vision model (taken from [18]).

It is then realized that the depth of the point in relation to the plane of the cameras can be found from three parameters. The focal distance f and the baseline b could be defined previously and are held constant. Thus, to find the depth of objects in the image, it is enough to find the disparity of all the pixels, what we call disparity map.

It is also important to note the distortion that variations in the disparity map can cause in the depth estimation, i.e., verify the measurement obtained accuracy. So, for a variation in depth, it is possible to find [18]:

$$\Delta Z = Z - \frac{bf}{d + \Delta d} = \frac{Z^2 \Delta d}{bf + Z \Delta d} \approx \frac{Z^2 \Delta d}{bf} \quad (9)$$

According to (9), to decrease the distortions, the variable Z cannot be very large while b and f cannot be very small. Since f is an intrinsic feature of the camera, the system is then planned so that b is a distance that allow the desired Z , in order to have less distortion.

So, the problem becomes to find the desired map of disparity. Therefore, to find the map, one must take into account the corresponding pixels of one of the images in the other, of course, if it exists. As the system configuration ensures that the Y coordinate between the images is the same (guaranteed by the synchronization between the cameras to ensure the validity of (5)), given a pixel (x_i, y_i) in image 1, one must look for the pair (x_j, y_i) in image 2 to obtain the local disparity, which is given by $|x_j - x_i|$.

The difficulties in solving the correspondence problem could be presented in the form of ambiguities (more than one matching pair) generated from regions with similar characteristics as texture and color intensity. Furthermore, the method considers that the luminous intensity received by both cameras of a given object will be the same, despite the variation in angle observed (i.e., the surfaces analyzed follow Lambert's law for emission), and the cameras have the same properties on the receiver, such as the gain and the bias. Another problem is the presence of hidden points between the cameras, which means that one point in one image does not correspond to another because it is visible only by one the camera. Such effect is called as occlusion.

Today, the stereo methods are well used and researched, especially where the region of interest has enough information to avoid ambiguities, with wider disparity maps. The basic procedures to implement such methods are as follows [19]:

- 1) Preprocess the images (optional).
- 2) For each disparity under consideration, compute all pixels for matching.
- 3) Consider to aggregate support spatially (only in non-global methods).
- 4) Across all disparities, compute and find the best correspondent match.
- 5) Compute a sub-pixel disparity estimate (optional).

The first step consist of preparing both images for analysis and can be done via hardware or software. The captured image can have high noise or even low noise, the last one possibly caused by the camera bias. One way to remove these image noises is to decrease the resolution (hardware filtering) or apply bandpass filters in the algorithm (filtering by software). In this work, it was used an averaging filtering as procedure to eliminate errors caused during the images acquisition and recording, sometimes treated for the captured images having differentiated illumination. This operation is useful to deal with random noise [20].

For the second step, it is possible to use global methods, where the image as a whole is analyzed (increases the accuracy and also the processing time); local methods, where the area used for calculation a pixel disparity is restricted; and hybrid methods, derived from local methods. Regardless of which is chosen, the first step of any dense stereo matching algorithm is a similarity measure that compares pixel values to determinate its correspondence. The most common pixel-based matching costs include sums of squared intensity differences (SSD) and absolute intensity differences (SAD), but there are also methods which are invariant in relation to gain and camera bias, to scale and to illumination, such as the Census Transform [21] [22].

In global stereo matching methods, the goal is to find a solution d that minimizes the global energy $E(d)$ in (11), where $E_d(d)$ measure how the disparity function d agrees with the input image pair, with L being the initial matching cost, and $E_s(d)$ encodes the smoothness assumptions made by the algorithm.

$$E(d) = E_d(d) + \lambda * E_s(d) \quad (10)$$

$$E_d(d) = \sum_{(x,y)} L(x, y, d(x, y)) \quad (11)$$

Meanwhile, the local methods utilize a support region in the disparity space image $L(x, y, d)$, described in (12), to aggregate the matching cost. It is important that the selected support region is large enough to contain texture variability and small enough to not contain many depth discontinuities.

$$L(x, y, d) = w(x, y, d) * L_0(x, y, d) \quad (12)$$

To ensure that the best matching pair of pixels are selected in the local methods, the easiest way is to choose the lowest cost associated with each pair of pixels in a given disparity,

a strategy known as "winner-take-all" (WTA). Although it may happen that only one image has single correspondence (reference image) while the other has points paired to multiple pixels, this can be done correctly via dynamic programming, as cooperative algorithms with symmetric constraints. The partial occlusion can also be treated by explicit pairing a pixel in the reference image to a group of pixels in the other one.

Finally, the estimation of sub-pixel disparity works to estimate the disparity at the points where it cannot be found, as occlusion points. With methods such as iterative gradient and interpolation using discrete disparity values, these points also earn a specific disparity and then increase the resolution of the stereo vision system, especially when the disparity variation is smooth. Another post-process tool applied to the disparity map is the association of confidence levels at each pixel depth estimation.

The stereo vision algorithm used to validate the captured images is a Fast SAD local method, which performs pixelwise matching based on a maximum disparity and a windows size entered by the user [23].

B. Raspberry Pi Architecture

The embedded platform Raspberry Pi has been considered fit for use due to its compact format, good processing and low cost, even if the use in computer vision area is new. The Raspberry Pi used was the Raspberry Pi 1 model B, which features 26 GPIO pins, Ethernet connect, two USB connectors, HDMI output and power supply via micro USB. Others features are shown in Table I. To ensure sufficient storage space for digital images to be recorded, it is necessary to use a SD card of at least 8 GB. The kernel used was Raspbian, offered by the manufacturer. The power system provides a voltage of 5 V and a maximum current of 2 A, which is sufficient because in tests it has been analyzed that the basic current for use of the Raspberry is between 300 mA and 550 mA, while each camera Pi requires approximately 250 mA when activated.

The kernel Raspbian is a version of the Debian Hard Float, which supports different program languages, such as Python. The command for capturing the images inserted via terminal enables the software code to record in each Raspberry the image captured by the cameras, via serial communication.

C. Acquisition of the Digital Images

According to official information, the properties of the camera Pi template 1.3 follow the information presented in Table II.

The implemented stereo system consists in the two Camera Pi in the same plane and with equal height, connect via CSI on the Raspberry Pi and correctly powered. It is chosen to use a fixed resolution to capture the images and convert them to grayscale before the stereo vision algorithm, although it will be select only the region of interest. Anyway, the RGB images will also be held in memory because the color is being used as an attribute for the weed's segmentation. The first necessary step to allow the use of the stereo vision system is to calibrate the cameras, which consist of images taken from a frame obtained from a chessboard to extract their intrinsic characteristic. The stereo matching is part of a high-level system, as shown in Figure 2.

To acquire the digital images in the agricultural field, we used the developed stereo digital camera, for validation, which allowed images having 640 x 480 pixels. Also, there was used a wooden frame measuring 0.5 x 0.5 m, which served as a scale factor for characterizing the size of the plants. Figure 3 shows the details of the arrangement for the experimental plots in the agricultural area.

TABLE I. RASPBERRY PI 1 MODEL B FEATURES

Processor	BCM2835 ARM1176JZFS		
Clock	700 MHz	GPIO	26 pins
Memory	512 MB RAM	Ethernet	1 conector
USB Ports	2 USB 2.0 ports	HDMI	1 conector
Camera serial interface (CSI)	Display serial interface (DSI)		
3.5mm jack for audio out	SD card slot		

TABLE II. CAMERA HARDWARE SPECIFICATIONS

Size	25 x 24 x 9 mm
Still resolution	5 MP
Video modes	1080p30, 720p60, 640x480p60/90
Sensor	OmniVision OV5647
Sensor resolution	2592 x 1944 pixels
Sensor image area ($W_s \times H_s$)	3.76 x 2.74 mm
Pixel size	1.4 μm x 1.4 μm
Optical size	1/4"
Full-frame SLR lens equivalent	35 mm
S/N ratio	36 dB
Dynamic range	67 dB @ (times of gain equal to 8)
Fixed focus	1 m - ∞
Focal length	3.60 0.01 mm
Horizontal field of view (HFOV)	53.50 \pm 0.13
Vertical field of view (VFOV)	41.41 \pm 0.11
Focal ratio (F-stop)	2.9

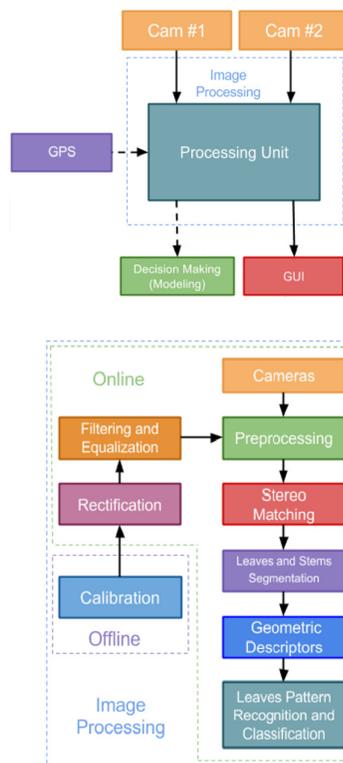


Figure 2. High-level system architecture diagram (taken from [18]).

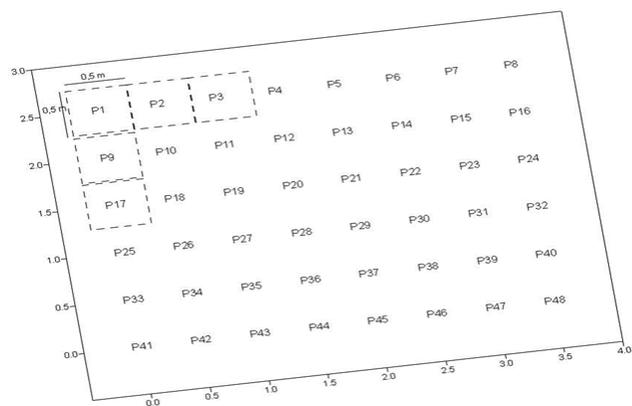


Figure 3. Details of the arrangement for validation based on the use of plots (from P1 to P48) in an agricultural area.

For weed recognition, the moment invariants method has been used, i.e., based on its application for each stereo image obtained from the plots. In fact, any geometrical pattern can be represented by a density distribution function, with respect to a pair of axes fixed in the visual field [24]. Based on such concept, the patterns can also be represented by their two-dimensional moments, with respect to the pair of fixed axes. Such moments of any order can be obtained by a number of methods [25]–[27]. Using the relations between central moments and ordinary moments, the central moments can also be obtained. In addition, by normalizing the central moments in size by using the similitude moment invariants a set of moment invariants can still be used to characterize patterns. Obviously, these are independent of the pattern position in the visual field and also independent of the pattern size. The algorithm for weed recognition based on both the stereo digital images and moment invariants is presented elsewhere [28].

Additionally, after the weed recognition process in each plot, the percentages of occupancy in such areas were considered to estimate a map of weed distribution based on the use of geostatistics. In such way it was possible to calculate the semivariogram, as well as constructing the map related to the management zones for the precision application at variable rate of the herbicide. The semivariogram depicts the spatial autocorrelation of the measured sample points. Once each pair of location is plotted, a model is fit through them. There are certain characteristics that are commonly used to describe these models [29]–[31].

IV. RESULTS AND DISCUSSIONS

With the assembled stereo vision system, as shown in Figure 4, the image quality analysis of the system modulation was evaluated by the Modulation Transfer Function (MTF) of one Camera Pi. To obtain the MTF, it was taken a image containing a linear white-to-black edge, converted the RGB image to grayscale and analyzed the normalized intensity values in a perpendicular line to the edge. The resulting profile is shown as the raw Edge Spread Function (ESF) in Figure 5a, compared with a gaussian smoothed ESF. After this, the LSF was calculated by the derivate of the ESF as shown in Figure 5b and the Fast Fourier Transform (FFT) was used to obtain the MTF, as shown in Figure 5c. Thus, the quantitative

evaluation of the system is demonstrably good because of the MTF.

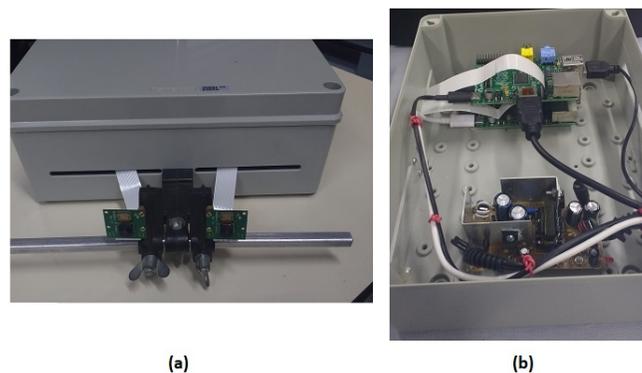


Figure 4. Implementation of the stereo system, where (a) is the stereo rig with the cameras and (b) the two Raspberry Pi on the top and a rectifier on the bottom.

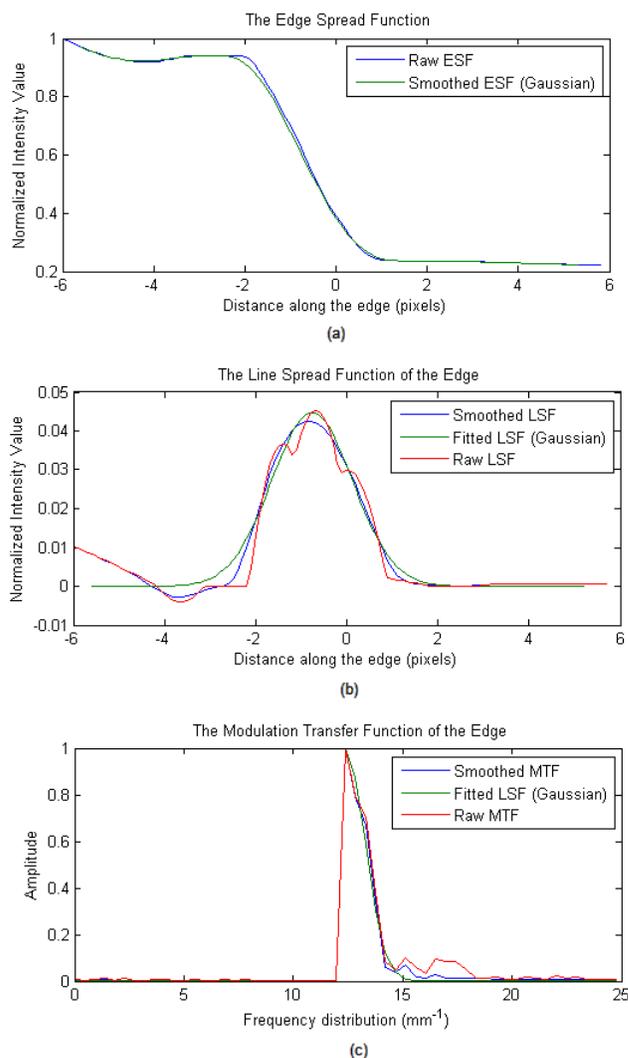


Figure 5. (a) The density profile of an edge; (b) The line spread function (LSF) of this edge; (c) The resultant modulation transfer function (MTF).

It was also desirable to verify whether the values of the field of view provided by the manufacturer are in agreement, as shown in Table II. For this, it was considered the pixel size of $1.4 \mu\text{m} \times 1.4 \mu\text{m}$, the focal length 3.6 mm, and the sensor resolution of 2592×1944 pixels were correct to evaluate the validity of the system, such as:

$$W_s = (s_x)(W) = (1.4\mu)(2592) = 3.63 \text{ m}$$

$$H_s = (s_y)(H) = (1.4\mu)(1944) = 2.72 \text{ m}$$

$$D_s = \sqrt{W_s^2 + H_s^2} = 4.54 \text{ mm}$$

$$HFOV = 2 \arctan(W_s/2f) = 53.5^\circ$$

$$VFOV = 2 \arctan(H_s/2f) = 41.4^\circ$$

$$DFOV = 2 \arctan(D_s/2f) = 64.4^\circ$$

From Table II, it is also possible to obtain $D_s = 4.65 \text{ mm}$, $DFOV = 65.7^\circ$, and so far such calculated values are in agreement with those provided by the cameras manufacturer.

The next step was to set the stereo system parameters. Considering that the distance between the camera plane and the captured image is equal to 50 cm, a good baseline to avoid distortions and to ensure a good stereo representation can be considered as equal to 6 cm, as inferred in Figure 6, obtained by the variation of b and Z in (8) and (9), respectively. Such result also has shown that the maximum disparity was around 300 pixels.

Besides, the fast SAD stereo method was chosen because offered a reasonable accurate stereo matching associated with a low processing time. The sub-pixels accuracy could be used to refine the results, although they significantly increased the processing time. Additionally, the pseudo-code of the algorithm can be observed in Figure 7.

With the parameters configured, the left and right images were obtained by each Raspberry Pi and its camera, as shown in Figure 8, with Figure 9 as the composite of the images with a red-cyan effect, which allows one to be able to view the image as a 3D shape with the aid of special glasses that has different filters in each lens.

Next, the images were entered as inputs of the chosen stereo vision method. The settings used were the maximum disparity equal to 300 and a window size equal to 5, adequate to catch the small details of the image.

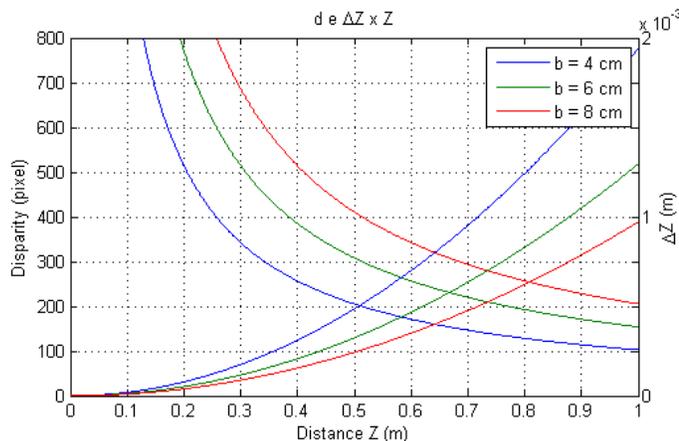


Figure 6. Graph of disparity in pixel and depth resolution against distance.

In addition, we obtained the map of disparity as a result (Figure 10), with association of the values of the baseline and the focal length. Therefore, the depth could be calculated through Equation (8). The entire process can be summarized according to the flowchart in Figure 11. Such flowchart presents aspects related to: the preparation of the hardware for images acquisition; the data transfer to the memory, which can be not only in the Raspberry architectures but also in other external computer; as well as the image processing algorithm for stereo vision and weed recognition.

Begin

- Set parameters (window size w and number of disparities n);
- Read image sizes; they must have the same height and width;
- Initiate the disparity map, matching cost, pixel cost and window cost with zero values;
- Calculate pixel cost for the n disparities using the absolute difference of two images;
- Calculate integral cost of the pixel cost with cumulative sums;
- Calculate window cost using the integral cost convoluted with a matrix h (h is a square matrix of size w with $h(1, 1) = h(w, w) = 1$, $h(1, w) = h(w, 1) = -1$ and otherwise equal to zero);
- Search disparity value by finding the minimum window cost;
- If specified, do the subpixel interpolation using polynomial fit;

end.

Figure 7. Stereo matching pseudo-code algorithm.



Figure 8. The images captured by the cameras from the agricultural area.



Figure 9. Red-cyan composite view of the stereo image, in which the 3D shape can be observed with the aid of special glasses.

Figure 12 and 13 show, respectively, the resulting semivariogram based on a Gaussian fitting, and the resulting map for the occupancy of the weed species *Bidens pilosa* in the whole agricultural area used as a pilot for validation.

In general, the herbicide isopropylamine salt of glyphosate is used to control invasive plants and the dose used is of the order of 3 L/hectare. This in practice occurs when the occupation rate of these invasive plants is in the order of 100% in relation to the crop area. Therefore, for the validation pilot that contains 48 plots with a total area of 12 m², only 3.6 mL of this agrochemical would be necessary. However, from the acquisition of the images and the digital recognition of the invasive plant patterns, as well as based on the prescription map obtained by the interpolation procedure, it is possible to observe the real occupancy rate in each plot, and the necessary volumes of this agrochemical for an effective localized control, as presented in Table III.



Figure 10. Disparity map.

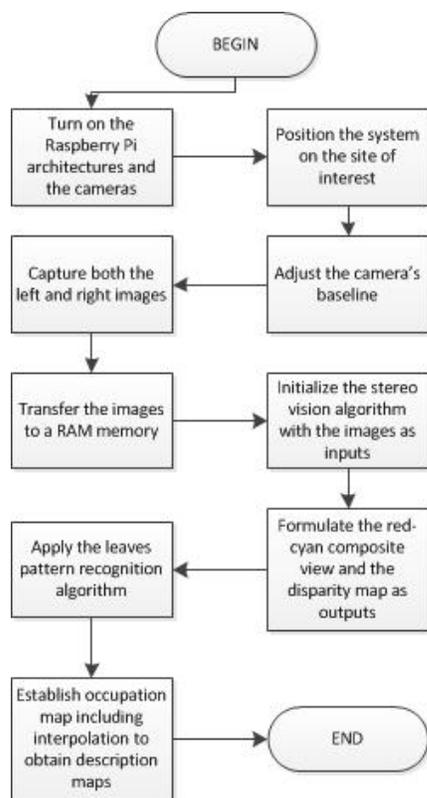


Figure 11. Image capture and processing flow.

In the traditional agricultural management systems, it is still necessary for someone to actually walk through the crops for analyzing the presence or absence of the weeds to support decision making in the crop area.

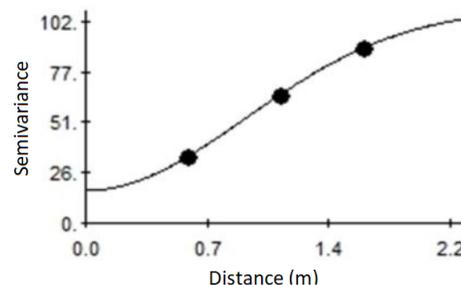


Figure 12. Semivariogram for the occupancy rate of the weed species *Bidens pilosa* (Model: Gaussian, $C_0 = 16.5$; $C_1 = 92.3$; $A = 1.3$). The model parameters are respectively represented by C_0 = nugget effect, C_1 = structural variance, and A = the range in meters.

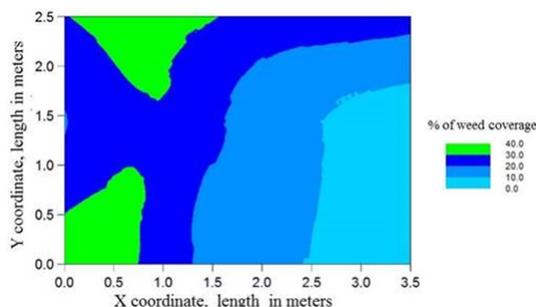


Figure 13. Map for the occupancy of the species *Bidens pilosa* in the whole plot area, i.e., based on the weed pattern recognition and the interpolation process.

TABLE III. THE CALCULATED VOLUME OF THE PRODUCT (HERBICIDE) TO BE APPLIED TO VARIABLE RATE DEPENDING ON THE TOTAL COVERAGE PROVIDED BY THE OCCURRENCE OF THE WEED IN THE PLOTS.

Experimental Plot From P1 to P24	Volume (ml) at precision application	Experimental Plot From P25 to P48	Volume (ml) at precision application
P1	0.040	P25	0.033
P2	0.045	P26	0.035
P3	0.045	P27	0.037
P4	0.100	P28	0.027
P5	0.040	P29	0.035
P6	0.043	P30	0.025
P7	0.041	P31	0.031
P8	0.040	P32	0.017
P9	0.037	P33	0.047
P10	0.032	P34	0.057
P11	0.036	P35	0.046
P12	0.033	P36	0.026
P13	0.021	P37	0.039
P14	0.021	P38	0.100
P15	0.023	P39	0.029
P16	0.024	P40	0.100
P17	0.025	P41	0.047
P18	0.035	P42	0.049
P19	0.044	P43	0.048
P20	0.032	P44	0.034
P21	0.029	P45	0.100
P22	0.027	P46	0.035
P23	0.020	P47	0.029
P24	0.019	P48	0.030

However, based on the stereo vision system, one is able to collect images to be analyzed autonomously, in order to obtain information for the prescription of maps for weed control based on herbicides application at variable rate.

V. CONCLUSIONS

The results have shown the usefulness of the developed stereo vision system customized for agricultural application, which is able to provide quality images, as a well to allow a depth model of plants, allied with a low cost portable embedded platform. Therefore, the identification of the weed species, integrated with the occupancy rate, its spatial distribution and depth defined a new approach to weed control, based on the precision application of herbicides into management zones. Besides, the case studied presented the opportunities to minimize costs and environmental impacts, as well as the promotion of the gains in competitiveness.

Future works can be found based on the developed stereo vision system customized for agricultural application taken into account its integration with a machine learning platform in order to aggregate intelligence in the decision making for automatic weed control in agriculture.

ACKNOWLEDGMENT

The authors would like to thank the Embrapa Instrumentation for providing the resources for conducting this research (grants project 01.14.09.001.05.05).

REFERENCES

[1] J. Behmann, A. Mahlein, T. Rumpf, C. Rmer, and L. Plmer, "A review of advanced machine learning methods for the detection of biotic stress in precision crop protection," *Precision Agriculture*, vol. 16, 06 2015, pp. 239–260.

[2] B. S. Blackmore and C. P. Blackmore, "People, robots and systemic decision making," J. V. Stafford, Ed. Netherlands: Wageningen Academic Publishers, 2007, pp. 433–439, <http://oro.open.ac.uk/10032/> [accessed in: July 17th, 2018].

[3] J. C. Christofoletti, "Considerations on the technology for pesticides application." Technical bulletin, vol. 5, 1999, p. 14.

[4] S. Vieira, J. Hatfield, D. R. Nielsen, and J. W. Biggar, "Geostatistical theory and application to variability of some agronomical properties," *Hilgardia*, vol. 51, 06 1983, pp. 1–75.

[5] Y. Xu, Q. Long, S. Mita, H. Tehrani, K. Ishimaru, and N. Shirai, "Real-time stereo vision system at nighttime with noise reduction using simplified non-local matching cost," *IEEE Intelligent Vehicles Symposium (IV)*, June 2016, pp. 998–1003.

[6] K. Rumyantsev, S. Kravtsov, I. Korovin, and G. Schaefer, "A statistical method for estimating the accuracy of scene reconstruction using a collinear digital stereo vision system," 6th International Conference on Informatics, Electronics and Vision and 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMHT), Sept 2017, pp. 1–4.

[7] Z. C. Liu, L. H. Xu, and C. F. Lin, "An improved stereo matching algorithm applied to 3d visualization of plant leaf," 8th International Symposium on Computational Intelligence and Design (ISCID), vol. 2, Dec 2015, pp. 354–358.

[8] M. Liang, X. Huang, C. Chung-Hao, G. Zheng, and A. Tokuta, "Robust calibration of cameras with telephoto lens using regularized least squares," *Mathematical Problems in Engineering*, 2014.

[9] L. Wang, F. Duan, K. Lv, and S. Chen, "Fisheye-lens-based visual sun compass for perception of spatial orientation," *Mathematical Problems in Engineering*, 2014.

[10] Y. C. Fang, C. M. Tsai, J. MacDonald, and Y. C. Pai, "Eliminating chromatic aberration in gauss-type lens design using a novel genetic algorithm," *Applied Optics*, vol. 46, no. 13, May 2007, pp. 2401–2410, <http://ao.osa.org/abstract.cfm?URI=ao-46-13-2401> [accessed in: July 17th, 2018].

[11] Y. C. Fang and C. M. Tsai, "Miniature lens design and optimization with liquid lens element via genetic algorithm," *Journal of Optics A: Pure and Applied Optics*, vol. 10, no. 7, 2008, <http://stacks.iop.org/1464-4258/10/i=7/a=075304> [accessed in: July 17th, 2018].

[12] C. C. Chen, C. M. Tsai, and Y. C. Fang, "Optical design of lcos optical engine and optimization with genetic algorithm," *Journal of Display Technology*, vol. 5, no. 8, Aug 2009, pp. 293–305.

[13] N. Baddour, "Multidimensional wave field signal theory: Mathematical foundations," *AIP Advances*, vol. 1, no. 2, 2011, <https://aip.scitation.org/doi/abs/10.1063/1.3596359> [accessed in: July 17th, 2018].

[14] K. F. Chen and F. L. Yan, "On the integration schemes of retrieving impulse response functions from transfer functions," *Mathematical Problems in Engineering*, 2010.

[15] M. Li, S. C. Lim, and S. Chen, "Exact solution of impulse response to a class of fractional oscillators and its stability," *Mathematical Problems in Engineering*, 2011.

[16] P. Lin and C. Liu, "Geometrical mtf computation method based on the irradiance model," *Applied Physics B*, vol. 102, no. 1, Jan 2011, pp. 243–249, <https://doi.org/10.1007/s00340-010-4349-3> [accessed in: July 17th, 2018].

[17] K. Tseng, C. Kung, T. Liao, and H. Chang, "Calculation of modulation transfer function of an optical system by using skew ray tracing," *Transactions of the Canadian Society for Mechanical Engineering*, vol. 33, 09 2009, pp. 429–442.

[18] N. S. Rosa, P. E. Cruvinel, and J. M. Naime, "Stereo vision embedded system proposal for plant phenotyping," *International Journal of Semantic Computing*, vol. 11, no. 03, 2017, pp. 293–309, <https://www.worldscientific.com/doi/abs/10.1142/S1793351X17400116> [accessed in: July 17th, 2018].

[19] D. Scharstein, *View Synthesis Using Stereo Vision*. Berlin, Heidelberg: Springer-Verlag, 1999.

[20] A. Konieczka, J. Balcerek, and A. Dbrowski, "Iterative average filtering for image denoising," *Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, Sept 2013, pp. 302–305.

[21] R. Szeliski, *Computer Vision: Algorithms and Applications*, 1st ed. Berlin, Heidelberg: Springer-Verlag, 2010.

[22] J. Lee, D. Jun, C. Eem, and H. Hong, "Improved census transform for noise robust stereo matching," *Optical Engineering*, vol. 55, 2016, <https://doi.org/10.1117/1.OE.55.6.063107> [accessed in: July 17th, 2018].

[23] W. Abbeloos, "Real-time stereo vision," Master's thesis, Karel de Grote-Hogeschool University College, Belgium, May 2010, unpublished.

[24] Y. S. Abu-Mostafa and D. Psaltis, "Recognitive aspects of moment invariants," *Pattern Analysis and Machine Intelligence IEEE Transactions*, vol. PAMI-6, 1984, pp. 698–706.

[25] H. Ming-Kuei, "Visual pattern recognition by moment invariants," *Information Theory, IRE Transactions*, vol. 8, 1962, pp. 179–187.

[26] M. R. Teaque, "Image analysis via the general theory of moments," *Journal of the Optical Society of America*, vol. 70, 1980, pp. 920–930.

[27] J. F. Boyce and W. J. Hossack, "Moment invariants for pattern recognition," *Pattern Recognition Letters*, vol. 1, 1983, pp. 451–456.

[28] P. E. Cruvinel and D. Karam, "Development of variable rate herbicide application maps for maize crops (zea mays l.) based on computational vision and occupation of broad leaf weed," In: *Brazilian Congress of Precision Agriculture (ConBAP)*, 2010, p. 4.

[29] P. Diamond and M. Armstrong, "Robustness of variograms and conditioning of kriging matrices," *Journal of the International Association of Mathematical Geology*, vol. 16, 1984, pp. 809–822.

[30] A. McBratney, R. Webster, and T. Burgess, "The design of optimal sampling schemes for local estimation and mapping of regionalized variables," *I. Theory and method. Computers and Geosciences*, vol. 7, 1981, pp. 331–334.

[31] J. Warnes and B. Ripley, "Problems with likelihood estimation of covariance functions of spatial gaussian processes," *Biometrika*, vol. 74, 1987, pp. 640–642.