Obstacle detection for autonomous systems using stereoscopic images and bacterial behaviour

Fredy Martínez, Edwar Jacinto, Fernando Martínez
Facultad Tecnológica, Universsdad Distrital Francisco José de Caldas, Colombia

Article Info	ABSTRACT
	This paper presents a low cost strategy for real-time estimation of the position of obstacles in an unknown environment for autonomous robots. The strategy was intended for use in autonomous service robots, which navigate in unknown and dynamic indoor environments. In addition to human interaction, these environments are characterized by a design created for the human being, which is why our developments seek morphological and functional similarity equivalent to the human model. We use a pair of cameras on our robot to achieve a stereoscopic vision of the environment, and we analyze this information to determine the distance to obstacles using an algorithm that mimics bacterial behavior. The algorithm was evaluated on our robotic platform demonstrating high performance in the location of obstacles and real-time operation.

Corresponding Author:

Fredy Martínez,
Universidad Distrital Francisco José de Caldas,
Carrera 77B No.64C-74 Villaluz, Bogotá D.C., Colombia.
Tel: (+57) 3005585481
Email: fhmartinezs@udistrital.edu.co

1. INTRODUCTION

Active robotic sensors have today become a high-performance tool with great acceptance at commercial and military level [1, 2]. These are embedded systems equipped with sensors that provide specific primary data, from which a real-time processor produces information relevant to the tasks of the robot [3]. This kind of sensors has promoted research in information-driven strategies for the development of tasks with robots, as well as the implementation of algorithms for digital signal processing and control schemes oriented to these sensors [4].

When faced with the design of motion strategies for autonomous robotic systems, these sensors prove to be very convenient, and even fundamental [5, 6]. When environments are dynamic (a typical problem for service robots) it is necessary for the robot to be able to identify nearby obstacles in real time [7, 8]. Unstructured environments are more complex due to their dynamics and lack of knowledge of identifiable characteristics. In addition, not all obstacles are the same, this means that the behavior of the robot in front of each of them must be different and appropriate in each case.

Between the minimum capacities that a robot must have is its capacity to define its relative size and dimensions in the environment. In other cases, it is also necessary to know its height to define interaction strategies (pick up a bottle from a table, for example). Depending on the application it is possible to use different kinds of sensors, but those capable of providing visual information are the ones that provide more relevant information [9]. In this sense, systems with two cameras turn out to be more advantageous than systems with a single camera [8], since they provide information on the depth and orientation of the obstacle [4,10-12].

Digital cameras as fundamental elements of optical sensors have been used extensively for the robotic arm motion control solution. The camera provides the required feedback information in relation to the position of the objects to be manipulated. This strategy is known as Visual Servoing or Vision-Based Robot Control (VS) and is characterized by having as feedback information the image of a camera [13]. The aim is to support the robot's decision making with eyes that take optical information from its own perspective and in parallel (separated by a certain distance) [11]. The distance between the robot and the obstacle can be determined depending on the distance between the obstacle positions in both images, and the focal distance of the cameras [14]. The field of vision can be increased considerably by adding a hyperboloid mirror or a conic mirror in front of the camera lenses, which provides an omnidirectional view to the cameras [15].

The reconstruction of 3D models from 2D perspectives (stereoscopic vision) is a strategy inspired by animal biology that allows the collection of three-dimensional information from the navigation environment. However, the process of generating 3D models is computationally expensive [16], and requires good camera calibration, making it very difficult to implement in real time on embedded systems [17]. In addition, the generation of 3D models is highly dependent on the quality of two-dimensional images, which are strongly affected by lighting conditions [18]. The computation of the distance to the obstacle takes into account the angular distance, the distance between cameras and the pixels of the images [7, 11]. However, in many applications, it is not necessary to rebuild the entire environment, which considerably reduces the computational requirement [19]. In fact, the human brain does something similar by processing information from the eyes, only focusing on a portion of the entire image that the eye detects. This information can then be processed to find specific shapes [20, 21].

There are two strategies for estimating the distance to the obstacle in stereoscopic vision: active method and passive method [10, 22]. In the first case, the sensor system sends signals to the obstacle such as visible light or laser signals, which are then detected and analyzed [11]. The ability of these sensors to establish distances is superior to human vision, but they are also costly and complex to implement, and they have unresolved problems. For example, the laser delivers the distance of a single point. In fact, these methods do not determine the exact 3D positions of all points of the obstacle. Another negative aspect is their speed, they are very slow for real-time operation [23]. On the other hand, the passive methods estimate the location of the obstacle from the images of the environment captured by cameras [19]. They use digital processing on the images to estimate the distance. This passive strategy has the additional advantage of working with different setups (cameras, light conditions, and embedded hardware). It should be clarified, however, that there are two problems that cannot be solved with this strategy: occlusions and overlapping of objects [24].

In order for the solutions to be real, it must be possible to massify them, and for this a low cost and high performance is essential [18, 23]. In this sense, processing algorithms must have very low computational cost in order to reduce processing time and hardware cost, while demonstrating to solve the problem. This paper attempts to address some of the critical problems of the strategy by maintaining a low computational cost, in particular reducing the impact of lighting on image quality, and improving the coincidence between 2D image points.

The main idea of our strategy is to identify points of obstacles by means of a movement in the images based on bacterial interaction, these points are mapped in the planes of projection of the environment in order to establish the distance to the obstacle, all this without the need to make modifications to the environment [12]. The firmware used to control the hardware setup, as well as data acquisition and processing, is written in Python. We detail the methods and algorithms used for image processing and estimation of the distance to obstacles. The results presented are the product of real laboratory tests carried out on our robot. Our proposed bio-inspired algorithm for three-dimensional obstacle reconstruction and the resulting motion control scheme have a number of advantages over other methods that directly control the entire nonlinear system or rely on dynamic programming for planning [25].

2. PROBLEM FORMULATION

We want an autonomous robot with low resource consumption to be able to identify obstacles in an unknown environment. In this sense, we define our robot in a W workspace. Let $W \subset \mathbb{R}^3$ be the clousure of a contractible open set in space that has an open interior connected with obstacles that represent inaccessible volumes. Let $\mathcal{O}$ be a set of obstacles finite in number in which each $O \subset \mathcal{O}$ is closed and pairwise-disjoint. Let $E \subset W$ be the free space in the environment, which is the open subset of W with the obstacles removed.

The robot has two cameras that form an optical system of stereoscopic vision. This system is located in $\boldsymbol{r}\,(t) \in \mathbb{R}^3$ and has $R\,(t) \in SO\,(3)$ orientation, where $SO\,(3)$ denotes the special orthogonal group of dimension three with respect to a global frame of reference for every instant $t \geq 0$.

To determine the position of the obstacles with respect to the robot, we define a relative frame of reference with respect to the axis of the two cameras as shown in Figure 1. We denote the two cameras by Left camera (L_c) and Right camera (R_c). The L_c and R_c centers are located at $(-0.14, 0, 0)$ and $(0.14, 0, 0)$ in the relative reference framework. The distance between the cameras is $b = 0.14 + 0.14 = 0.28\,m$.
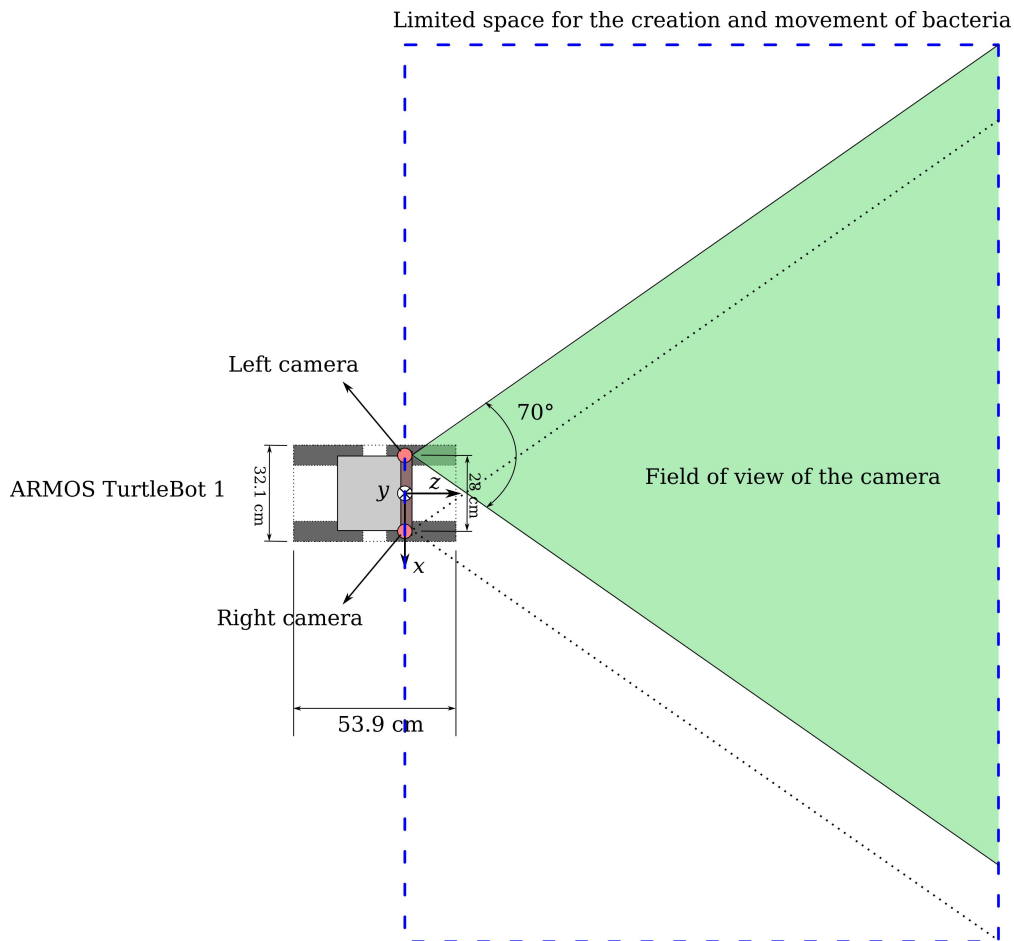

Figure 1. Dimensions of the prototype with detailed location of the cameras, three-dimensional axes for the location of bacteria, and their limited space in the navigation environment (top view)

The obstacles, indexed by $i \in \mathcal{H} = \{1, 2, 3, \cdots, n\}$ have unknown position $\boldsymbol{x}_i\,(t)$, and can be moved in E over time. The position for the obstacle O_i with respect to the global frame of reference can be expressed as (1):

$$\boldsymbol{x}_i\,(t) = R\,(t)\,\boldsymbol{p}_i\,(t) + \boldsymbol{r}\,(t) \tag{1}$$

where $\boldsymbol{p}_i\,(t)$ corresponds to the position of the obstacle with respect to the frame of reference relative to the cameras.

The cameras produce two parallel images at instant t with the location information $\boldsymbol{p}_i\,(t)$. However, obstacles are not points, they are volumes whose surface is made up of a large number of points. We do not want to determine the position of all points of obstacles. Instead, we want to identify the position of a small group of points that will ideally move to the surface of the obstacles.

We define a population of m bacteria in the space in which the robot may encounter obstacles when moving forward as shown in Figure 1. The initial position of each bacterium is random but known. From the images of the two cameras, we can establish trigonometric relationships for the three-dimensional position of each bacterium. If the bacteria are on the surface of the obstacle, then we can determine the distance to these points of the obstacle as depicted in Figure 2. We propose a search algorithm (obstacle search) in which bacteria move three-dimensionally according to local information detected in their 2D projections. In addition, the algorithm is accelerated according to the bacterial Quorum Sensing (QS), i.e. large populations of bacteria in a space make the space more attractive to other bacteria.
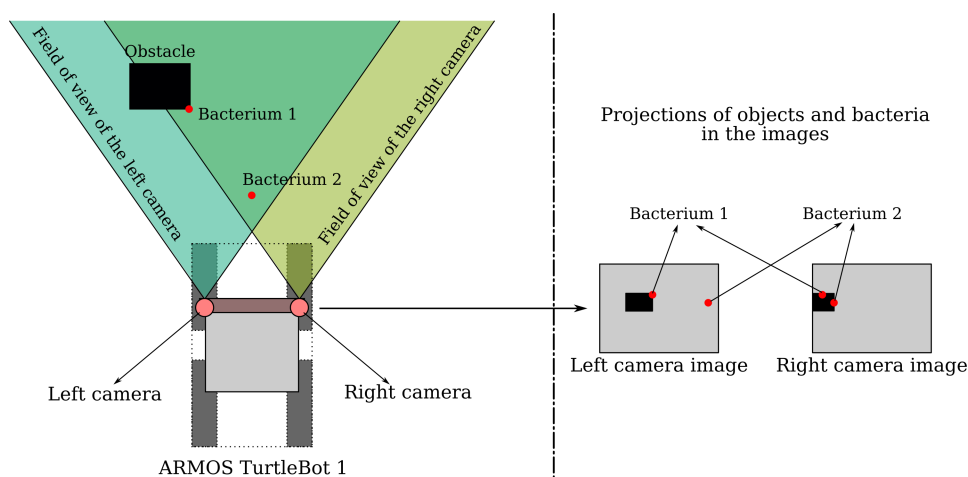

Figure 2. Layout of elements in the test hardware and images resulting from the two cameras with details of an obstacle and two bacteria (top view)

The m bacteria (or agents), all identical to each other, move in W searching for areas of great interest to them (for example, in search of food). The value of a given position is determined from local readings (local interaction with the medium) evaluated from its projection on 2D images. Each bacterium is defined by its position in the environment (2):

$$V = (p) \tag{2}$$

where p is a point in *3*-dimensional space $\left(p \in \mathbb{R}^3\right)$. The population density is evaluated using the distance between bacteria (3):

$$d_{ij} = d\left(V_i, V_j\right) \tag{3}$$

as the distance between bacteria V_i and V_j, which is calculated by an appropriate norm.

The function used to evaluate the value of the region where the bacterium is found in the left and right projections considers the similarity of the neighboring pixels to the bacterium in the two projections is depicted in Figure 2. The mathematical expression is (4):

$$F = \frac{|\nabla\left(M_L\right)| \cdot |\nabla\left(M_R\right)|}{\sum\limits_{colors}\sum\limits_{(i,j)\in N}\left[L_{(x_L+i,y_L+j)} - R_{(x_R+i,y_R+j)}\right]^2} + f\left(QS\right) \tag{4}$$

where (x_L, y_L) and (x_R, y_R) are the coordinates of the left and right projections of the current bacterium, $L_{(x_L+i,y_L+j)}$ is the grey value at the left image at pixel $(x_L + i, y_L + j)$ (in a similar way for the right image), N is the neighborhood around the projection of each bacterium, and $|\nabla\left(M\right)|$ is Sobel gradient norm on left and right projections (to penalize uniform regions).

Bacterial QS is activated if the population density within a space is greater than a threshold value T called the quorum threshold. It is the parameter defining whether or not it has reached the quorum. The behaviors of bacteria (search in the environment) are coordinated by the following rule:

- If the bacterium $V_k \subset W$ is located near to the bacterium $V_i \subset W$, i.e. (5):

$$d_{ik} < h \tag{5}$$

and the number of bacteria within the sphere with radius $\frac{h}{2}$ and origin in V_k is greater than T, then the value of the region increases for V_i.

3. RESEARCH METHOD

We initialize the bacterial population randomly within the field of action of the robot (red dotted line in the top view of Figure 1, 3 m along the x-axis, 2 m depth on the z-axis, and 2 m height above ground). The coordinates of each bacterium are defined with respect to the frame of reference relative to the cameras. The size of the population was taken as a performance variable parameter with values between 10 and 1000.

The cameras are located on the robot at a height of 0.5 m from the ground. The origin of the frame of reference relative to these cameras is at this height, in the middle of the two cameras. The positive x-axis corresponds to the right side of the robot, the positive z-axis corresponds to the direction of advance of the robot, and the positive y-axis grows above the robot.

The images of L_c and R_c are scaled to 800×600 pixels. The projection of each bacterium i on the images is determined with the following equations (the position (0,0) of the image is in the upper left side):

$$\text{Left image}: \begin{cases} x_p = 400 + \frac{(x_i + 0.14)800}{2z_i tan(35°)} \\ y_p = 300 - \frac{(y_i)600}{2z_i tan(30°)} \end{cases} \tag{6}$$

$$\text{Right image}: \begin{cases} x_p = 400 + \frac{(x_i - 0.14)800}{2z_i tan(35°)} \\ y_p = 300 - \frac{(y_i)600}{2z_i tan(30°)} \end{cases} \tag{7}$$

where (x_i, y_i, z_i) is the three-dimensional coordinate of the bacterium i, and (x_p, y_p) is the two-dimensional coordinate of the bacterium projected in the image.

The performance of the area adjacent to the bacteria at each projection is determined by (4). The bacteria move in the limited space according to this function. If the bacterium is on the obstacle surface, then it will have similar neighboring pixels in both projections as shown in Figure 2, the illumination affects both cameras equally), and the function will assign a high value to the position of the bacterium. The more the neighboring pixels differ, the less value the function assigns. The position of the bacteria is updated with the gradient looking for the high values (movement of the bacteria). The QS forces the bacteria that are slow to find the obstacle surface to move towards the large groups of bacteria. A bacterium that does not appear in any of the projections obtains the lowest position value (it is outside the robot's range of vision).

4. RESULT AND ANALYSIS

We evaluate the performance of the strategy with different configurations varying the bacteria population, the QS threshold and the correlation window used in the denominator of the evaluation function. A larger number of bacteria allows for reconstructing larger portions of the obstacles without significantly influencing the computational cost of the algorithm. The QS threshold reduces the convergence time when it does not exceed the range of 100, above this value, does not have a significant effect. The most important effect was observed in the size of the correlation window of the function, which greatly affects the bacteria's ability to locate the obstacle. Large values improve the behavior but considerably increase the computational cost. Figures 3 and 4 show the result of one of the laboratory tests.

Figure 3. Left and right images captured by parallel cameras, scaled to 800×600 and converted to grayscales

Figure 4. Image of the left camera converted to grayscale, scaled to 800×600 and with the bacteria overlapped in its final position, most of them on the obstacle

We perform more than 50 laboratory tests with different obstacles and more or less constant lighting conditions for a human indoor environment (the day with natural lighting and night with LED type lighting). The distances from the objects to the robot were established in a straight line between 0.3 and 2 m. The accuracy of the distance values determined by the optical sensor was established by comparison with the actual value, measured in the setup with a tape measure. These results were related to the distance of the obstacle. Figure 5 shows these percentages of accuracy with respect to the estimated distance.
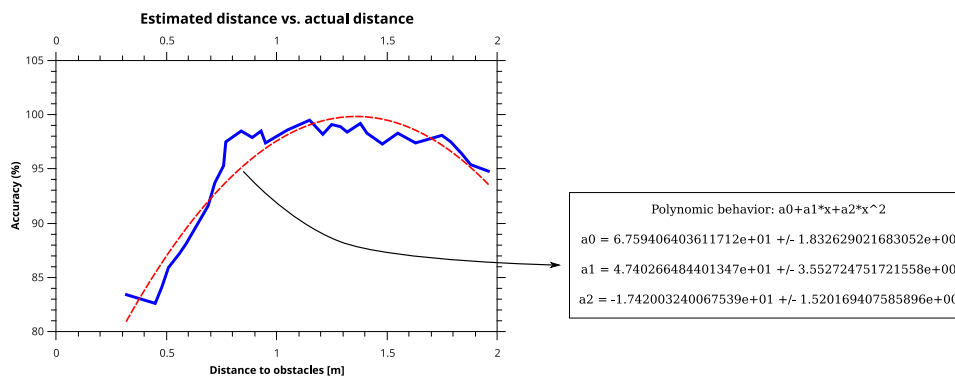

Figure 5. Percentages of accuracy with respect to the estimated distance

Our intention is to use the strategy to identify obstacles in the environment, and with this information coordinate the movement of the robot. The proposed motion planning strategy based on the detection and stereoscopic identification of obstacles considers three elements: capture and pre-processing of images, determination of obstacles and application of motion policies according to the information feedback as shown in Figure 6.
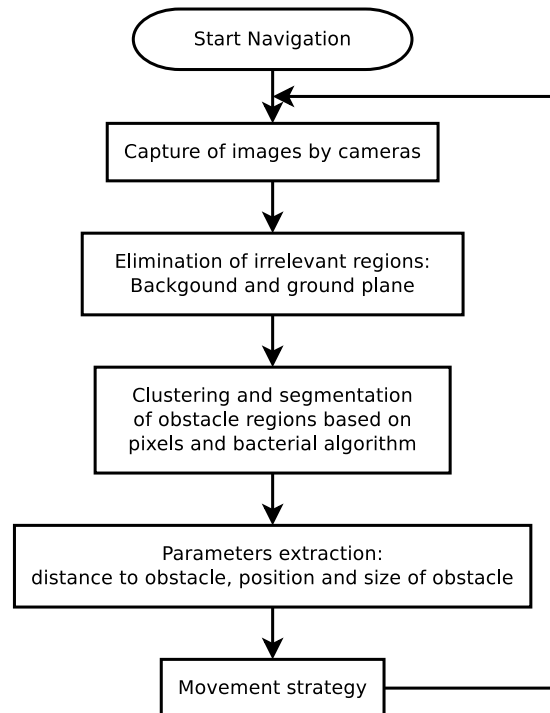
Figure 6. General scheme of the proposed motion planning strategy based on the stereoscopic detection of obstacles

5. CONCLUSION

Considering the problem of motion planning of small autonomous robots in unknown environments, particularly for service robots with direct and continuous interaction with the human being, we propose a low-cost computational stereoscopic vision strategy that allows autonomous navigation in dynamic environments. Service robots perform their tasks in indoor environments, unknown, with a high probability of constant change in the location of obstacles and people. The stereoscopic vision systems allow to establish with precision the three-dimensional location of obstacles and therefore provide complete information for the design of navigation strategies. However, their computational cost is high, making it impossible to use them in real-time on moderate performance platforms. Our strategy proposes a local reconstruction of a finite set of points of obstacles in the environment, which guarantees a low cost and a high performance. We performed the calculation of about 100 points corresponding to the surface of the obstacles. These points are identified using an uninformed search algorithm inspired by bacterial interaction. The bacteria defined in the 2D projections of the cameras move in the three-dimensional space looking for similar neighboring regions in their projections. The algorithm converges with most bacteria on the obstacles. In the experiments carried out, it was possible to verify percentages of accuracy to the obstacle distance higher than 95% and low computational consumption, making it useful for embedded implementations. The future development of the scheme includes improvements in the determination of obstacle surfaces using larger bacterial populations, and reduction in convergence times through the use of the Quorum Sensing (QS) model.

REFERENCES

[1] H. Himanshu, D. Deepanshu, K. Amit, and G. Aashish, "Autonomous robots for military," *National Journal of Multidisciplinary Research and Development*, vol. 3, no. 1, pp. 994–997, 2018.

[2] M. Ghute, K. Kamble, and M. Korde, "Design of military surveillance robot," in *First International Conference on Secure Cyber Computing and Communication (ICSCCC 2018)*, pp. 270–272, 2018.

[3] B. Schlotfeldt, V. Tzoumas, D. Thakur, and G. Pappas, "Resilient Active Information Gathering with Mobile Robots," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2018)*, pp. 4309–4316, 2018.

[4] C. Freundlich, Y. Zhang, A. Zhu, P. Mordohai, and M. Zavlanos, "Controlling a robotic stereo camera under image quantization noise," *The International Journal of Robotics Research*, vol. 36, no. 12, pp. 1268–1285, 2017.

[5] N. Aklil, B. Girard, L. Denoyer, and M. Khamassi, "Sequential action selection and active sensing for budgeted localization in robot navigation," *International Journal of Semantic Computing*, vol. 12, no. 1, pp. 109–127, 2018.

[6] B. Calli, W. Caarls, M. Wisse, and P. Jonker, "Active Vision via Extremum Seeking for Robots in Unstructured Environments: Applications in Object Recognition and Manipulation," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 4, pp. 1810–1822,2018.

[7] S. Solak and E. Bolat, "Distance estimation using stereo vision for indoor mobile robot applications," in *9th International Conference on Electrical and Electronics Engineering (ELECO 2015)*, pp. 685–688, 2015.

[8] Y. Hongshan, Z. Jiang, W. Yaonan, J. Wenyan, S. Mingui, and T. Yandong, "Obstacle Classificationand 3D Measurementin Unstructured EnvironmentsBased on ToF Cameras," *Sensors*, vol. 14, no. 1, pp. 10753–10782, 2014.

[9] Z. Yuanshen, G. Liang, H. Yixiang, and L. Chengliang, "A review of key techniques of vision-based control for harvesting robot," *Computers and Electronics in Agriculture*, vol. 127, no. 1, pp. 311–323, 2016.

[10] Y. Dawood, K. Ruhana, and E. Kamioka, "Distance measurement for self-driving cars using stereo camera," in *6th International Conference on Computing and Informatics (ICOCI 2017)*, pp. 235–242, 2017.

[11] A. Mohamed, Y. Chenguang, and A. Cangelosi, "Stereo Vision based Object TrackingControl for a Movable Robot Head," in *4th IFAC International Conference onIntelligent Control and Automation Sciences*, pp. 161–168, 2016.

[12] M. Ferreira, P. Costa, L. Rocha, and A. Moreira, "Stereo-based real-time 6-DoF work tool tracking for robot programing by demonstration," *The International Journal of Advanced Manufacturing Technology*, vol.85, no.1, pp. 57–69, 2016.

[13] C. Mao-Hsiung, L. Hao-Ting, and H. Chien-Lun, "Development of a Stereo Vision Measurement System for a 3D Three-Axial Pneumatic Parallel Mechanism Robot Arm," *Sensors*, vol. 11, no. 2, pp. 2257–2281,2011.

[14] M. Mahammed, A. Melhum, and F. Kochery, "Object distance measurement by stereo vision," *International Journal of Science and Applied Information Technology*, vol. 2, no. 2, pp. 5–8, 2013.

[15] J. Yamaguchi, *Three Dimensional Measurement Using Fisheye Stereo Vision*, Intech, 2011.

[16] H. Martins, I. Oakley, and R. Ventura, "Design and evaluation of a head-mounted display for immersive 3D teleoperation of field robots," *Robotica*, vol. 33, no. 10, pp. 2166–2185,2015.

[17] S. Dreier, M. Savran, L. Konge, and F. Bjerrum, "Three-dimensional versus two-dimensional vision in laparoscopy: a systematic review," *Surgical Endoscopy*, vol. 30, no. 1, pp. 11–23, 2015.

[18] K. Panjvani, A. Dinh, and K. Wahid, "LiDARPheno - A Low-Cost LiDAR-based 3D Scanning System for Leaf Morphological Trait Extraction," *Frontiers in Plant Science*, vol. 10, no. 147, pp. 1–17, 2019.

[19] S. Boonkwang and S. Saiyod, "Distance measurement using 3D stereoscopic technique for robot eyes," *7th International Conference on Information Technology and Electrical Engineering (ICITEE 2015)*, pp. 232– 236, 2015.

[20] O. Bertel, C. Moreno, and E. Toro, "Aplicación de la transformada Wavelet para el reconocimiento de formas en visión artificial," *Tekhnê*, vol. 6, no. 1, pp. 3–8, 2009.

[21] S. Mehta, "Vision-based localization of a wheeled mobile robot for greenhouse applications: A daisy-chaining approach," *Computers and Electronics in Agriculture*, vol. 63, no. 1, pp. 28–37, 2008.

[22] D. Patel, P. Bachani, and N. Shah, "Distance measurement system using binocular stereo vision approach,"

International Journal of Engineering Research & Technology, vol. 2, no. 12, pp. 2461–2464,2013.

[23] Y. Si, G. Liu, and J. Feng, "Location of apples in trees using stereoscopic vision," *Computers and Electronics in Agriculture*, vol. 112, no. 1, pp. 68–74, 2015.

[24] M. Mehrabi, E. Peek, B. Wuensche, and C. Lutteroth, "Making 3D Work: A Classification of Visual Depth Cues, 3D Display Technologies and Their Applications," in *Proceedings of the Fourteenth Australasian User Interface Conference (AUIC 13)*, vol. 139, 2013, pp. 91–100.

[25] R. Lins, S. Givigi, and P. Gardel, "Vision-Based Measurement for Localization of Objects in 3-D for Robotic Applications," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 11, pp. 2950–2958,2015.

BIOGRAPHIES OF AUTHORS

Fredy Martínez is a professor at the Facultad Tecnológica, Universidad Distrital Francisco José de Caldas, Bogotá D.C.-Colombia. He obtained his Bachelor's Degree in Electrical Engineering and his Ph.D in Engineering - Systems and Computing from the National University of Colombia (Colombia) in 1997 and 2018 respectively. Since 2000 he leads the ARMOS research group at the Universidad Distrital Francisco José de Caldas (Colombia). His research focuses on electronics, control systems, hybrid architectures, autonomous robotics and intelligent systems. The application of robotic systems in the provision of services to people has recently been addressed.

Edwar Jacinto is a professor at the Facultad Tecnológica, Universidad Distrital Francisco José de Caldas, Bogotá D.C.-Colombia. He obtained his Bachelor's Degree in Control Engineering and his Master Degree in Sciences of the Information and Communications from the Universidad Distrital Francisco José de Caldas (Colombia) in 2004 and 2015 respectively. His research focuses on the fields of electronics, control systems, embedded systems, communication solutions and custom encryption. The application of hardware-based encryption for decentralized communication of mobile nodes has recently been addressed. He is affiliated with IEEE as professional member.

Fernando Martínez is a professor at the Facultad Tecnológica, Universidad Distrital Francisco José de Caldas, Bogotá D.C.-Colombia. He obtained his Bachelor's Degree in Control Engineering and his Master Degree in Electronic and Computer Engineering from the Universidad Distrital Francisco José de Caldas (Colombia) in 2004 and 2012 respectively. His research focuses on the fields of electronics, instrumentation systems, real-time image and video processing, embedded signal processing solutions. Recently, the development of autonomous navigation strategies based on images has been tackled.