William Percy Cuno Zuniga, Jordan Jonathan Quispe Navarro, Juan Daniel Pocohuanca Diaz, Jesus Talavera S., Andres Montoya A.

Universidad Nacional de San Agustin de Arequipa, Peru

doi:10.15199/48.2023.05.46

Design of a Terrain Mapping System for Low-cost Exploration Robots based on Stereo Vision

Abstract. *This paper proposes a low-cost system for terrestrial mapping and exploration of inaccessible or subterranean environments by using stereo vision with two cameras, image processing and a developed algorithm based on disparity maps that reconstructs a 3D map of the explored environment. The tests were performed on a robot with two stereo vision cameras mounted on a turret with 360° freedom of movement. The tests showed that this proposed system allows to visualize the depth of the objects around the robot and builds a 360° scenario of the explored place.*

Streszczenie. *W artykule zaproponowano niedrogi system do mapowania naziemnego i eksploracji niedostępnych lub podziemnych środowisk przy użyciu stereowizyjnego widzenia z dwiema kamerami, przetwarzania obrazu i opracowanego algorytmu opartego na mapach rozbieżności, który rekonstruuje trójwymiarową mapę badanego środowiska. Testy przeprowadzono na robocie z dwiema kamerami stereowizyjnymi zamontowanymi na wieżyczce ze swobodą ruchu 360°. Testy wykazały, że proponowany system pozwala na wizualizację głębokości obiektów wokół robota i buduje scenariusz 360° badanego miejsca.* (***Projekt systemu mapowania terenu do tanich eksploracji Roboty oparte na Stereo Vision***)

Keywords: Low Cost; Terrain mapping; Stereo vision; Distance estimation
Słowa kluczowe: mapowanie terenu, sterewizja, robot

Introduction

The exploration of environments where people can't access due to the dimensions or the danger it may present has always been a major issue in various fields. In rescue operations, after natural disasters, the exploration for the search of survivors is primordial the shortest time and the largest search area to perform the task more effectively. In cases of subway cave entrapment it is equally relevant. Most of the time these structures to be explored are not safe and reliable enough for a person to adventure to inspect, because of this at international level has increased the use of robots for dangerous work, in case something were to happen, only expendable and recoverable material would be lost. The use of robots has had an enormous impact on the support of dangerous tasks for people, such as explosive deactivation tasks [1], search and rescue tasks[2], among others. To get the most out of these robots, they are equipped with various systems that help them to function better and perform tasks efficiently without generating stress to the operator. Among these integrated systems we have systems for their control [3] and visual systems [4, 13]; as well as the combination of several systems [5, 1] which provide support and data for a better manipulation of them increasing the confidence with the robots [6].

The technological development in the programming of visual systems has also had a huge boom in recent years, with image processing has had a lot of applications in different fields; in agriculture has been quite advanced in the classification of products [7], estimates of sizes of different products [8] or measurements of crop sizes [9].

In the field of exploration there are sensors such as Light Detection and Ranging (LIDAR) used in SLAM technology, which provides a mapping of environments by infrared scanning, the disadvantage of these modern advances in technology is that they are not accessible to the economy of all people because they require the purchase of extra sensors in addition to those already existing in these types of systems (most scanning robots incorporate cameras for visualization). Different methods have also been developed to achieve terrestrial scans or mapping such as the use of synthetic aperture interferometric radar methods[10], grid maps[11] or the optimization of position graphs[12].

The technological development through image processing with two stereo cameras achieves results similar to these expensive systems with triangulation methods and disparity maps. These algorithms that are developed allow us to estimate data required in exploration and terrestrial mapping tasks. For that reason, a low-cost terrestrial mapping system is proposed, which is based on image processing algorithms, using stereo vision with cameras mounted on a 360° rotating turret placed on a robot explorer, with all the literature reviewed it was decided to perform disparity mapping combined with the triangulation method for the construction of a terrestrial map because it will take advantage of the cameras that a scanning robot normally has, thus avoiding the purchase of other sensors.

This article is divided as follows: In the Related Works Section, the work related to this research is presented, the methodology is developed in the Methodology Section, in the Development of the Proposed System Section, the description of the proposed stereo vision algorithm is performed, the description of the developed experimentation is explained in the Experimental Results Section with the obtained results and discussions, and the conclusions of this research are presented in the Conclusion Section with the future work we will perform.

Related Works

Distance estimation using stereo cameras has been developed in a variety of works as in [14], where a new method for distance measurement using a stereo camera based on frequency domain analysis was proposed; using frequency analysis, it was possible to obtain the distance between the camera and the object with precision using less time than other methods considered in the comparison. In addition, the experimental results showed that the main advantage of this new method is that the disparity value can be calculated through the phase in the frequency domain.
through the phase in the frequency domain.

In [15] a Python-based algorithm is developed to find parameters of each camera, rectify images, create the disparity map and finally use these maps to estimate the distance to the object. The results were obtained by performing the experimentation with images taken in real time with real distances between 60 cm and 200 cm; the results showed an accuracy of 99.86% with a processing time for the calculation of less than 0.355 seconds.

The article [16] emphasizes the necessity of using the triangulation method to calculate distances between a robot and an object, however, this technique lacks effectiveness when the objects to be measured are located to the right or left of the robot. To solve this problem, the look-up table and curve fitting methods were added to the triangulation method to obtain an average accuracy of 97.69%; in addition, Manhattan and Euclidean distances were calculated for objects that are not in the same line of the robot, obtaining 98.24% and 98.03% accuracy for Manhattan and Euclidean respectively.

In the aspect of camera distribution and setup for distance estimation using stereo cameras, [17] has concluded that changes in the angle between the two cameras have a significant impact on the measurement range and the measurement distance, changes in the distance between the two cameras have a greater impact on the measurement distance but a smaller impact on the measurement range, concluding that, cameras with a longer field of view are better for small measurement ranges. It is also [18], emphasized that the camera position is crucial to obtain good results, so after the analysis the results show that the optimal distance between the cameras to obtain accurate measurements is 10 cm where the measurement range with 1% error is between 40 cm and 200 cm.

In addition, distance estimation with stereo cameras has been applied in agricultural scenarios such as in [9], where a measurement system was developed for a crop robot, which converts stereo images to disparity maps through stereo matching, so that the disparity of each pixel is calculated to determine the distance between the camera and the crop. The distance measurement by means of stereo cameras has also been tested in underwater environments [19] where an embedded system is developed to perform underwater image measurement using stereo vision techniques also making use of an FPGA together with a display for image visualization. Also it has been implemented in [20] service robot, where a stereo camera is implemented to help measure the distance between the robot and possible collisions so that it can handle in time in indoor environments, in this way it was found that the robot has a good performance when performing various tests.

Methodology

This article presents the development of a system that performs low-cost terrain mapping using stereo vision algorithms. In Figure 1, the operating cycle of the proposed system is detailed. First, the robot arrives at the location to be mapped and stands still in a strategic place, usually in an area where the turret can rotate without any impediment. After that, the operator starts the mapping process, which consists of letting the turret rotate and capture images of the environment, this process is automatic so the operator only has to wait for the mapping system to finish the task. As the cameras capture images, the algorithm performs distance estimation using triangulation and disparity mapping methods. Once the system has finished its work, the robot will send the final images to the control station through its WiFi network.

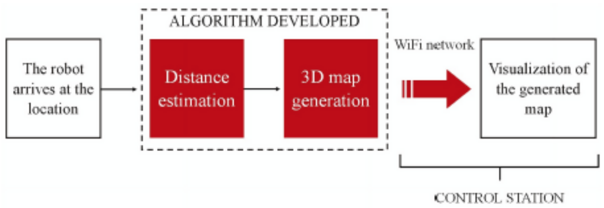

Fig. 1. Block diagram of proposed system

The algorithm starts when the turret cameras capture the first images. In Figure 2 it is observed that the algorithm when receiving the two images, left and right, starts the triangulation method to estimate the distance. First, having the cameras previously calibrated, the intrinsic and extrinsic parameters are used to correct the images and get better captures. Then, the depth map is made, to separate the objects in the images according to the distance at which they are located. After that, a filtering is performed to correct possible blurs in the depth map. Once the depth map is obtained, the distances of the objects are estimated and these are sent to the 3D map generation algorithm, which is in charge of placing the estimated distances on the original image and adding the following images obtained. At the end of the turret rotation, the 3D map generation algorithm will have generated the complete map of the environment explored by the robot.

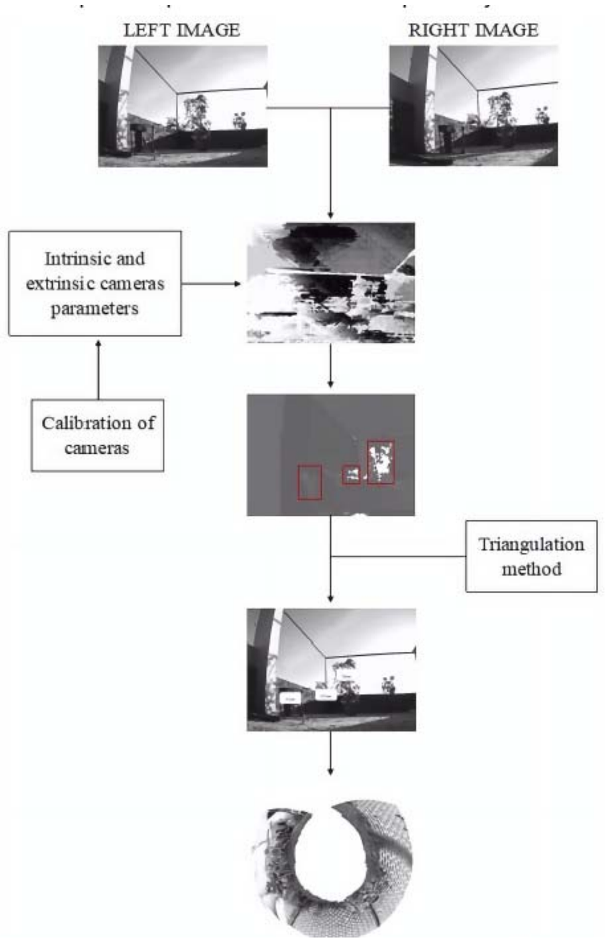

Fig. 2. Flowchart of the proposed algorithm

Development of the Proposed System

When we use a camera to capture a 3D object and it is converted into a 2D image, the depth information is lost, making it impossible to estimate distances to the captured objects. In nature this phenomenon is corrected by using both eyes to capture images and reconstruct 3D objects in the mind. Stereo vision attempts to copy this concept by extracting 3D information from two 2D images, making it possible to recover the depth information of the captured object if images of the object are captured from different perspectives. For that reason, stereo vision allows the operator to have reference to the depth at which the object is located, while if a single camera were used, it would not be possible to obtain this information. The two cameras can be on a line common base with angles of inclination, this is known as con- verging cameras, or they can be parallel, which is known as non-converging cameras, see Figure 3.

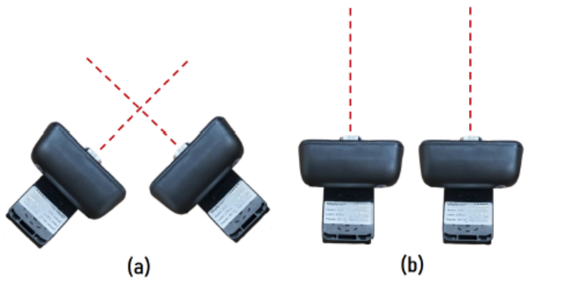

Fig. 3. (a) Configuration of convergent cameras. (b) Configuration of non-convergent cameras

0.1 Stereo image analysis

The stereo cameras used in this system to acquire images have the non-convergent configuration, this configuration was chosen for its greater mathematical simplicity. In Figure 4, we observe the two parallel cameras located on a common linebase (T represents the distance between the centers of the cameras O_L and O_R, respectively). The point P represents the object in the real world; f is the focal length of the lens; Z is the distance between P and the center point between the two cameras [21].

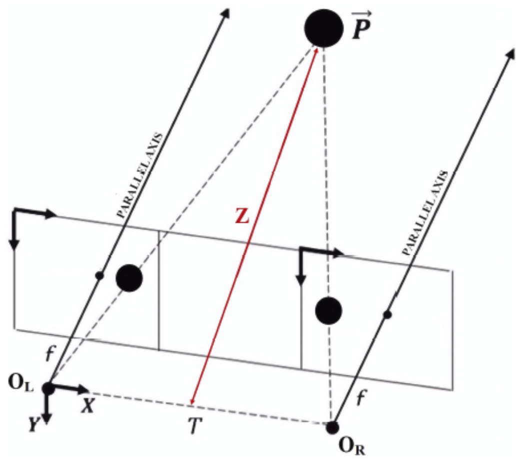

Fig. 4. Schematic of the stereo vision model

Therefore, the distance (Z) between the camera and the object can be calculated by Equation (1), the height (Y) and width (X) values will be used to calculate the area of the object captured in the generated map. The relative equations were obtained from the relation of similar triangles in Figure 5. [22]. The difference of the coordinates in both images is given by Equation (2).

$$Z = f\frac{T}{d} \quad : \quad Height$$
$$X = x\frac{Z}{f} \quad : \quad Width$$
$$Y = y\frac{Z}{f} \quad : \quad Distance \tag{1}$$

$$d = |x_L - x_R| \tag{2}$$

where the distance of the target point (P) projected on the left (x_L) and right (x_R) image sensors is calculated with depth information acquired from the disparities between the two images at the horizontal pixel (x-axis) distance from the target point.

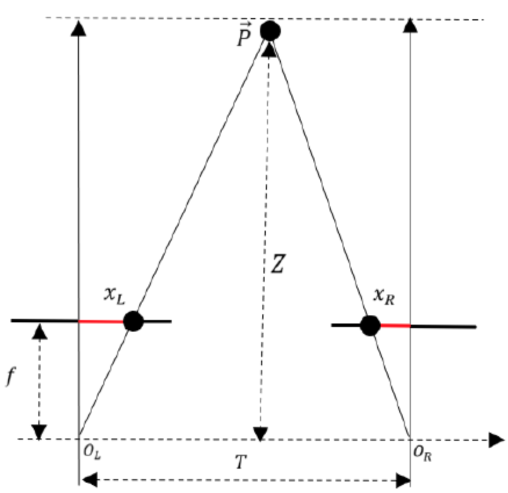

Fig. 5. Triangulation scheme of stereo vision

In order to correctly calculate disparity, stereo matching is required to match 2D objects corresponding to the same region in the two camera views. The stereo matching process finds dense correspondences in image pairs to obtain distance information [23]. Figure 6 shows the calculation performed to determine how closely pixels match by calculating the matching cost while moving across the patch horizontally. The square patch was taken as the reference image and the homologous patch was found along a corresponding horizontal line in the target image [24].

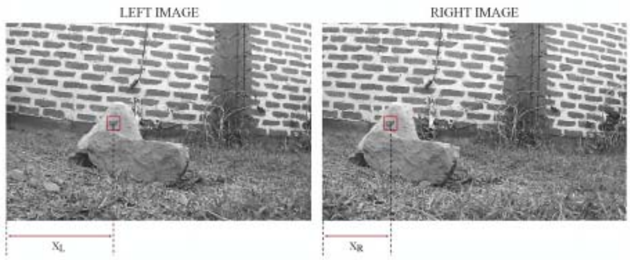

Fig. 6. Disparity model between the two images

0.2 3D map generation

In this study we have developed an algorithm that generates a 3D map of the scanned terrain. With the triangulation done correctly, we proceed to make the disparity map, see Figure 7 (a). To this disparity map generated, the distances estimated with the previous algorithm are assigned, so that it is easier for the operator to observe the distances at which the objects are located in the images, so that the disparity map will be the same as shown in Figure 7 (b). Finally the estimated distances are positioned to the original image captured by one of the cameras, by convention we work with the image of the left camera, leaving finally the image as shown in Figure 7 (c).

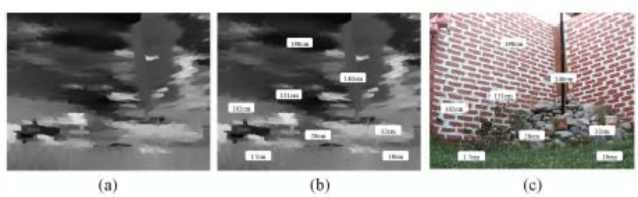

Fig. 7. Generation sequence of the 3D map images

The terrain mapping algorithm will repeat this entire process of estimating object distances and generating the disparity map for each pair of images obtained for each turn of the turret until it completes a 360° turn. Once it completes the turn, the algorithm will send the generated image to the control station via the local network server.

Experimental results

To test the proposed system, two Xiaomi CMSXJ22A webcams mounted on a stand in parallel configuration were used, as shown in Figure 8. These cameras have an ON Semiconductor 1/3-inch CMOS digital image sensor, 90° field of view, Full HD (1020 x 1080 pixels) resolution at a speed of 60 fps and separated at a base distance (T) of 125 mm. The acquired images were transmitted to a rapsberry Pi 4 model B+ with 8GB RAM memory, and the algorithm was developed in the Python programming language (Visual Studio 2017 software, Microsoft, USA).

Fig. 8. Support for both webcams

The intrinsic and extrinsic parameters of the cameras were obtained by calibrating them by Zhang's method [25]

and making use of the StereoCalibrate function (Opencv 2019). The cameras were calibrated using an 8 x 8 checkerboard with a grid size of 30 x 30 mm, as shown in Figure 9. For a better calibration, the left and right cameras were left immobile while the chessboard was moved, so that both cameras always captured all the squares of the chessboard and as the images were captured, the chessboard was rotated to capture different angles. A total of 22 images for each camera were captured to perform the correct calibration.

0.3 Results of distance estimation

The results shown in Figure 10 and 11 refer to the results obtained in [26] and [21]. In Figure 10, distances between 50 cm and 150 cm were taken into account to compare with the results when applying the estimation using the Extraction Method; the maximum error when comparing is 13.3% while the minimum error is 4.2%. In addition, the precision decreases as the object moves away from the camera. Figure 11 shows the results of [21], measurements were performed in the range of 50 cm to 150 cm using the Other Method for the calculation of the estimated distance, the minimum error when comparing is 0.2% while the maximum error is 6.5% which occurs as in the first case the farther away the object
is.

Figure 12 shows the results of testing our proposed method, the tests were performed with distances between 20 cm and 230 cm for the estimation. It also shows a maximum error of 3.7% and a minimum error of 0.0%, being the maximum error found in the minimum distance of 20 cm. Figure 13 shows a general comparison between the 3 proposed methods, where a much lower general error is evidenced by the proposed method, in addition, the proposed method shows a better performance even in distances that were not considered by the two previous methods.

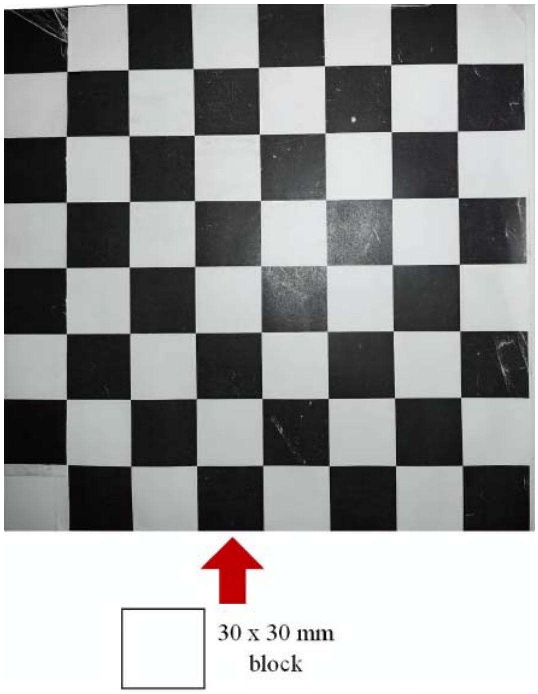

Fig. 9. Checkerboard pattern used to calibrate cameras

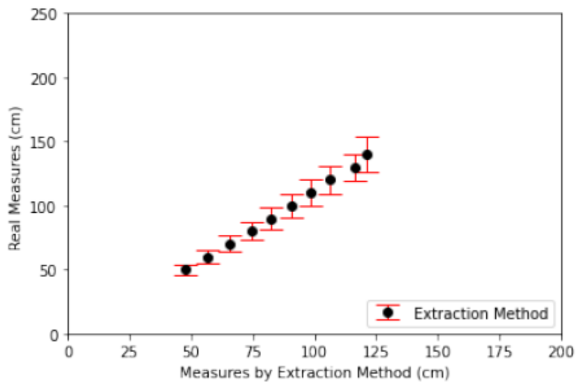

Fig. 10. Graph of extraction method measurements

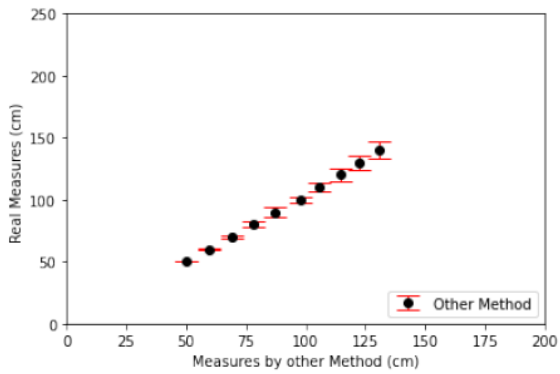

Fig. 11. Graph of the measurements of another estimation method [21]

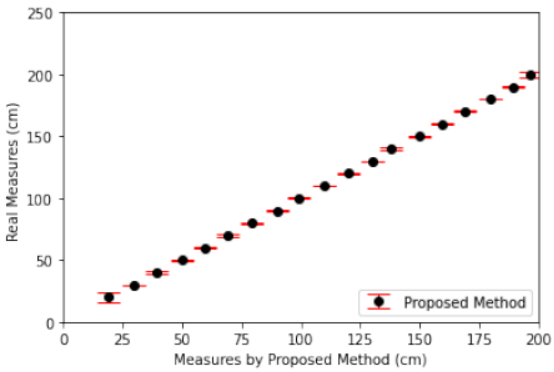

Fig. 12. Graph of the measurements of the proposed method

0.4 Results of 3D map generation

Figure 14 shows some of the most representative disparity maps obtained in this study. The disparity maps are best shown in grayscale, where the elements that are closer to the cameras are shown in white; those farther away are shown in dark gray; and those in black are the objects that fail to coincide in both cameras. For example, Figure 14 (a) shows the disparity map of some rocks of different sizes and located at different distances, generating a disparity map quite dispersed. Figure 14 (b) represents the disparity map of the remains of a house, and where only rubble remains. Figure 14 (c) shows the disparity map of an environment with a rock in the middle to test the effectiveness of the algorithm in detecting objects. Figure 14 (d) shows the map of a rooftop in a more urban environment. And finally, Figure 14 (e) shows the disparity map of a place with grass and vegetation, the intention was to test the system in various environments and conditions. The results of the disparity maps show that it can be used to differentiate the elements present in an image.

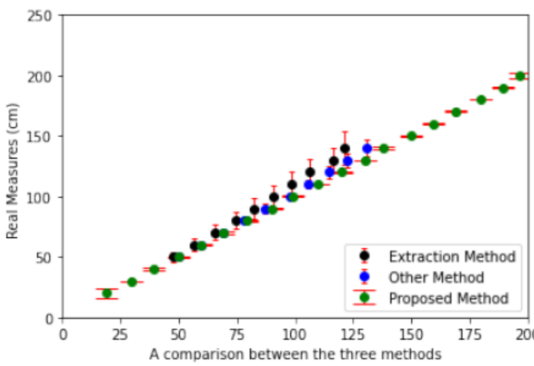

Fig. 13. Graph of the measurements of the three methods together

Figure 15 shows the 3D reconstruction of a scanned terrain, this would be the image sent to the server so that the operator can visualize the terrain with the estimated distances of the objects around the scanning robot. The robot is represented in Figure 15 as the central blue dot, while the other dots represent the objects detected by the system, each at its respective distance; The color of each point depends on how close or far it is from the robot, with red being the closest (distance < 50cm), followed by orange (50cm <= distance < 80cm), then yellow (80cm <= distance < 120cm), then lemon green (120cm <= distance < 160cm) and finally green (distance >= 160cm).

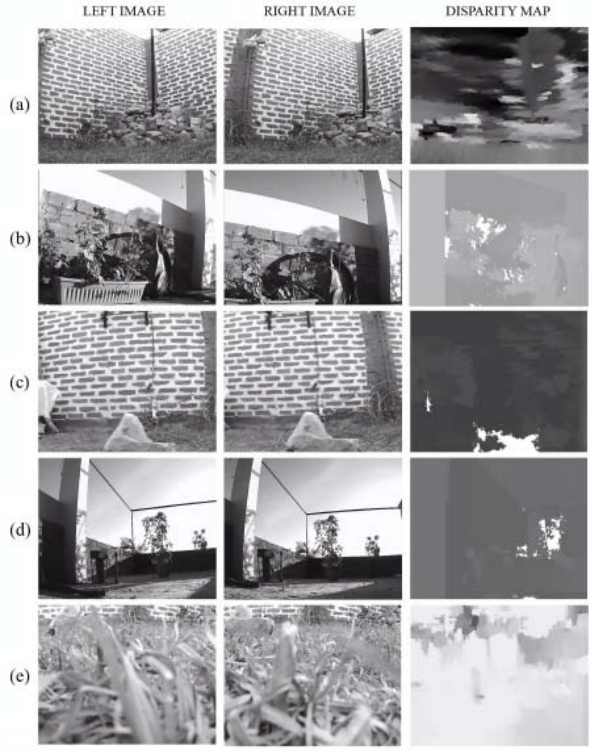

Fig. 14. Disparity map

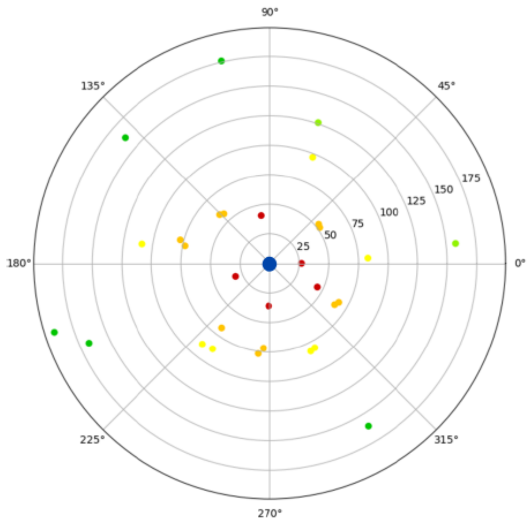

Fig. 15. 3D map

Conclusion

This paper presents the development of a map generation system that uses stereo vision to estimate the distances of the environment and project them on the images. The algorithm is in charge of generating a disparity map from the images captured by both cameras, thus obtaining a map with distances. The results shown verify the effectiveness of this method to be applied in real time with high accuracy, the accuracy obtained was around 98.9%, which is quite high and within the acceptable percentage. In addition, the map reconstruction was achieved correctly, recreating the environment with distance estimation of objects with high accuracy. This proposed system makes it possible to achieve similar results to the complex and high-cost systems mentioned above, which would have a great application in exploration tasks of places inaccessible to people.

In future work will seek to create routes by analyzing the maps built to generate an autonomous driving for the exploration robot, in addition, we will seek to improve the accuracy of the estimated distances.

Acknowledgment

The authors would like to thank Universidad Nacional de San Agustin de Arequipa.

***Authors**: William Percy Cuno Zuniga, Universidad Nacional de San Agustin de Arequipa, Peru, E-mail: wcunoz@unsa.edu.pe; Jordan Jonathan Quispe Navarro, Universidad Nacional de San Agustin de Arequipa, Peru, E-mail: jquispenav@unsa.edu.pe; Juan Daniel Pocohuanca Diaz, Universidad Nacional de San Agustin de Arequipa, Peru, E-mail: jpocohuancad@unsa.edu.pe; Dr. Jesus Talavera Suarez, Univ rsidad Nacional de San Agustin de Arequipa, Peru, E-mail: jtalaveras@unsa.edu.pe; Eng. Andres Montoya Angulo, Universidad Nacional de San Agustin de Arequipa, Peru, E-mail: jmontoyaan@unsa.edu.pe*

REFERENCES

[1] T. Yamamoto, Y. Takagi; A. Ochiai, K. Iwamoto, Y. Itozawa, Y. Asahara, K. Ikeda. Human Support Robot as Research Platform of Domestic Mobile Manipulator. Springer: Cham, Switzerland, 2019; Volume 11531.

[2] Rafael Verano M, Jose Caceres S, Abel Arenas H, Andres Montoya A, Joseph Guevara M, Jarelh Galdos B and Jesus Talavera S, Development of a Low-Cost Teleoperated Explorer Robot (TXRob) International Journal of Advanced Computer Science and Applications(IJACSA), 13(7), 2022.

[3] Vilcapaza Goyzueta, D., Guevara Mamani, J., Sulla Espinoza, E., Supo Colquehuanca, E., Silva Vidal, Y., Pinto, P.P. Evaluation of a NUI Interface for an Explosives Deactivator Robotic Arm to Improve the User Experience. HCI International 2021 Late Breaking Posters. HCII 2021. Communications in Computer and Information Science, vol 1498. Springer, Cham.

[4] M. A. Andres, L. Pari and S. C. Elvis. Design of a User Interface to Estimate Distance of Moving Explosive Devices with Stereo Cameras. 2021 6th International Conference on Image, Vision and Computing (ICIVC), 2021, pp. 362-366.

[5] E. Triantafyllidis, C. Mcgreavy, J. Gu and Z. Li, Study of Multimodal Interfaces and the Improvements on Teleoperation. IEEE Access, vol. 8, pp. 78213-78227, 2020.

[6] Deborah R. Billings, Kristin E. Schaefer, Jessie Y.C. Chen, and Peter A. Hancock. Human-robot interaction: developing trust in robots. Proceedings of the seventh annual ACM/IEEE international
conference on Human-Robot Interaction (HRI 2012). Association for Computing Machinery, New York, NY, USA, 109–110.

[7] Malabag, Beverly & Santiago Jr, Cereneo & Cahapin, Erwin & Reyes, Jocelyn & Legaspi, Gemma. (2022). Fuzzy Logic-Based Size and Ripeness Classification of Banana using Image Processing Technique. International Journal of Emerging Technology and Advanced Engineering. 12. 11-18, 2022.

[8] Nirunsin, Surasi & Shinonawanik, Praween & Thintawornkul, Tawan &Wongratanaphisan, Theeraphong. Size Estimation of Mango Using Mask-RCNN Object Detection and Stereo Camera for Agricultural Robotics. International Journal of Emerging Technology and Advanced Engineering. 12. 161-168, 2022.

[9] Wan-Soo Kim, Dae-Hyun Lee, Yong-Joo Kim, Taehyeong Kim, Won-Suk Lee, Chang-Hyun Choi. Stereo-vision-based crop height estimation for agricultural robots. Computers and Electronics in Agriculture. Volume 181,2021.

[10] M. Pieraccini, G. Luzi and C. Atzeni. Terrain mapping by ground-based interferometric radar IEEE Transactions on Geoscience and Remote Sensing. vol. 39, no. 10, pp. 2176-2181, Oct. 2001.

[11] Z. Li, Y. Li, X. Rong and H. Zhang. Grid Map Construction and Terrain Prediction for Quadruped Robot Based on C-Terrain
Path. IEEE Access. vol. 8, pp. 56572-56580, 2020.

[12] B. Zhou et al., A Graph Optimization-Based Indoor Map Construction Method via Crowdsourcing. IEEE Access., vol. 6, pp. 33692-33701, 2018.

[13] Montoya Angulo, A.; Pari Pinto, L.; Sulla Espinoza, E.; Silva Vidal, Y.; Supo Colquehuanca, E. Assisted Operation of a Robotic Arm Based on Stereo Vision for Positioning near an Explosive Device. MDPI Robotics 2022. 11, 100.

[14] Yoo, I.-S., & Seo, S.-W. Accurate object distance estimation based on frequency-domain analysis with a stereo camera. IET Intelligent Transport Systems. 11(4), 248–254, 201, 2017

[15] Adil, E., Mikou, M., Mouhsen, A. novel algorithm for distance measurement using stereo camera. CAAI Trans. Intell. Technol. 7(2), 177– 186 (2022).

[16] SOLAK, S., & BOLAT, E. D. A new hybrid stereovisionbased distance-estimation approach for mobile robot platforms. Computers Electrical Engineering., 67, 672–689, 2018.

[17] A. J. Golkowski, M. Handte, P. Roch and P. J. Marron, Quantifying the Impact of the Physical Setup of Stereo Camera Systems on Distance Estimations. 2020 Fourth IEEE International Conference on Robotic Computing (IRC). 2020, pp. 210-217.

[18] Setyawan, R. A., Soenoko, R., Mudjirahardjo, P., & Choiron, M. A. Measurement Accuracy Analysis of Distance Between Cameras in Stereo Vision. 2018 Electrical Power, Electronics, Communications, Controls and Informatics Seminar (EECCIS).

[19] Sanchez-Ferreira, C., Mori, J. Y., Farias, M. C. Q., & Llanos, C. H. A real-time stereo vision system for distance measurement and underwater image restoration. Journal of the Brazilian Society of Mechanical Sciences and Engineering- 38(7), 2039–2049, 2016.

[20] Budiharto, W., Santoso, A., Purwanto, D., & Jazidie, A. Multiple moving obstacles avoidance of service robot using stereo vision. Telkomnika, 9(3), 433-444, 2012.

[21] Du, YC., Taryudi, T., Tsai, CT. et al. Eye-to-hand robotic tracking and grabbing based on binocular vision. Microsyst Technol
27., 1699–1710 (2021).

[22] Z. Liang, W.J. Tam, Stereoscopic image generation based on depth images for 3D TV, IEEE Transactions on Broadcasting, vol. 51, issue: 2, 2005, pp. 191–199.

[23] A. A. Ahmed, M. K. Elbashir and A. A. Osman, Distance alert system using Stereo vision and feature extraction. International Conference on Communication, Control, Computing and Electronics Engineering., Khartoum, Sudan, 2017, pp. 1-5.

[24] V. G. Posugade and R. P. Patil, FPGA based design and implementation of disparity estimation for stereo vision system. 2016 International Conference on Computing Communication Control and automation., Pune, India, 2016, pp. 1-5.

[25] Z. Zhang. A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence., Nov. 2000, vol. 22, no. 11, pp. 1330-1334.

[26] J. Zhang, J. Chen, Q. Lin and L. Cheng. Moving Object Distance Estimation Method Based on Target Extraction with a Stereo Camera. IEEE 4th International Conference on Image, Vision and Computing (ICIVC)., Xiamen, China, 2019, pp. 572-577.