

3D Indoor Environmental Reconstruction System that Allows to Determine the Trajectory of a Mobile Robot Through Analysis and Image Processing

Ing. Nelson Stalin Shagñay Ruiz

Investigador Independiente, stalinshs101@gmail.com

Ing. Oswaldo Geovanny Martínez Guashima

Escuela Superior Politécnica de Chimborazo, omartinez@esPOCH.edu.ec

Ing. Luis Fabian Brito Mancero

Universidad Nacional de Chimborazo, fabian.brito@unach.edu.ec

Ing. Marcela Yolanda Brito Mancero

Escuela Superior Politécnica de Chimborazo, mybrito@esPOCH.edu.ec

Ing. Joffre Satalin Monar

Escuela Superior Politécnica de Chimborazo, jmonar@esPOCH.edu.ec

Abstract

This paper presents the results of 3D reconstruction of indoor environments using as a unique technique the reconstruction technique of stereo vision and Matlab software for obtaining point clouds, subsequent stitching using the ICP algorithm and trajectory calculation using the same reconstruction making an estimate of location in the indoor area. To obtain the information of the environment to be reconstructed, a low-cost non-autonomous mobile robot commanded by bluetooth was implemented, mounted on this a stereo camera MYNT EYE S1030 was available for the capture of stereo images and having a certain trajectory in order to obtain the data of the environment in 360 grams with respect to the horizontal axis. The results, based on the statistical analysis carried out on this system, showed variations in the fidelity of the physical environment compared to the virtual environment with an average error of 3.27 meters, and is intended to be used in the topography and mapping of indoor environments.

Keywords: *Vision estéreo, Mobile robot, ICP algorithm, Matlab.*

I. INTRODUCTION

Thanks to technological progress we have managed to carry out increasingly complex processes, such as 3D reconstruction, which allows us to recreate real objects or scenarios virtually and show them through a computer while maintaining their physical characteristics. This process takes time under

study for which there are various methods to make the 3D reconstruction in order to become as faithful to the object or real scenario and then give a purpose to the virtual scenario, such as space exploration or teleoperation in areas of high human risk.

Stereo vision can be used to control the mobile robot in the teleoperation area. So the

calculation of the distance is important and it is done by 2 cameras that make up a stereo camera and has a function similar to that of the eyes of the human being. This is because you get the view of an object in different positions, the difference in angles between these positions is known as disparity and this disparity is used to calculate the distance between the camera and the object [1]–[3].

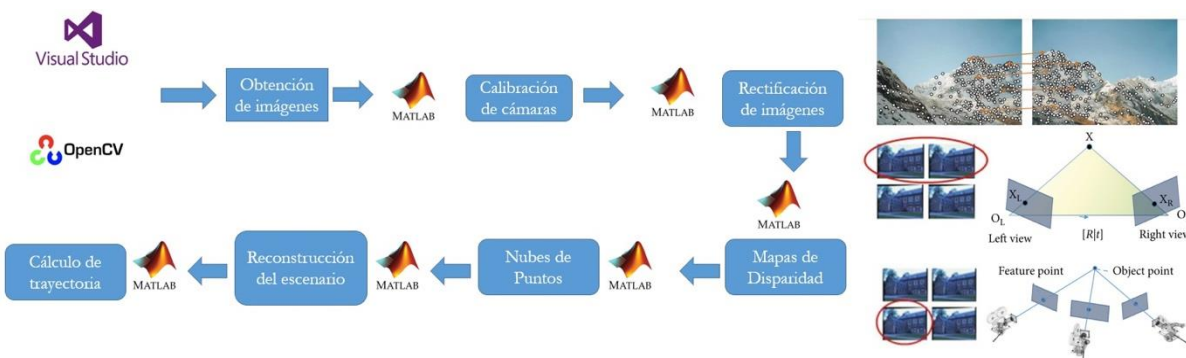
Among the advances of 3D reconstruction systems we have In case [4] a stereo catadioptric 3D reconstruction system was developed, composed of two catadioptric cameras, capable of generating semi-dense reconstructions based on DeepMatching with epipolar constraints.

They also [5] present a system called DeepMVS, which is defined as a deep convolutional neural network (ConvNet) for multi-view stereo reconstruction. Taking an arbitrary number of images as input, they first produce a set of flat-scan volumes and use the proposed DeepMVS network to predict high-quality disparity maps.

Finally, [6] they describe the reconstruction of indoor environments in 3D by using a single stereo camera mounted on a drone and the calculation of the position of the drone by using only the images captured by the drone, this is done using the SGM algorithms for the calculation of depth maps and IGP algorithm for the union of the point cloud.

The purpose of this work is the reconstruction of scenarios in 3 dimensions with fidelity as close as possible to the real scenario. This fidelity depends on the characteristics of the stereoscopic camera when capturing data in controlled environment scenarios, as well as the software to process the images. The reconstruction of 3D scenarios begins with obtaining data from the real scenario, for which, one way to obtain this data is the use of images captured by stereo cameras properly calibrated on a mobile robot, calling this process as stereo vision, it is continued by rectifying these images, creating disparity maps, Creation of point clouds, union of point clouds and trajectory calculation, to finally analyze the operation of the system. These phases are shown in Fig. 1.

Fig. 1. Stages of image processing for the reconstruction of environments and trajectory calculation.



II. 3D RECONSTRUCTION SYSTEM

For the reconstruction of indoor environments , the stages shown in Fig. 1 were followed. for

which the images are obtained using the Visual Studio tool and through the OpenCV library you can manipulate the camera to capture the data, in this case the stereo images.

The captures of the images were taken by a non-autonomous mobile robot programmed in Arduino and of low cost, which was controlled by Bluetooth technology and a mobile application. Both the design and implementation are shown in Fig. 2.

Fig. 2. Indoor 3D reconstruction system (a) Design, (b) Implementation



A. Camera calibration

After obtaining the images, it must be taken to calculate the extrinsic and intrinsic parameters of the MYNT EYE S1030 camera for the rectification of the images and thus a more accurate pixel correspondence can be made. Therefore, MATLAB and its calibration tool are used to obtain this data, where the input parameters are nothing more than captured images containing a calibration pattern, so a "chessboard" pattern was used. This process delivers calculated parameters such as the

number of patterns, parameters of each chamber individually, mean re-projection error, translation matrix, units of measurement, etc.

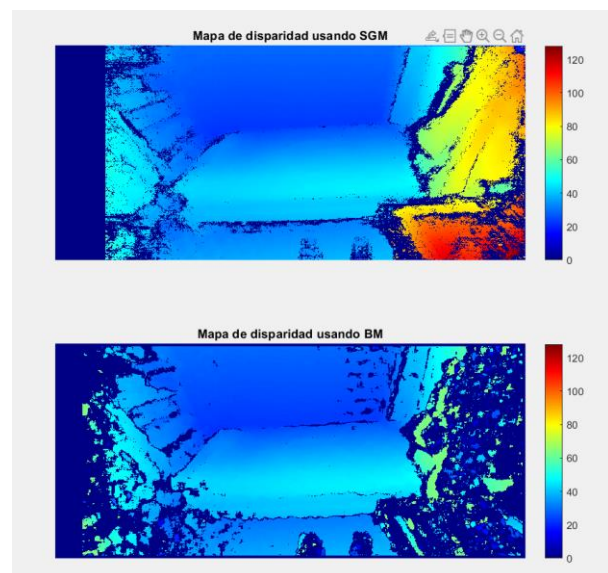
B. Image rectification

The parameters calculated in the previous stage serve for the rectification of the images obtained since, by default, the images are obtained not rectified, this means that in the images there are problems such as radial or tangential distortion that cause the images to present curvatures in the periphery. This is summarized in that there is no correct correspondence of pixels between the left image with the right image, hence the concept of epipolar plane is defined for the stereo vision technique that defines correspondence of pixels that are projected parallel to the horizontal plane that indicate similarity of values on the vertical axis. So the calculation is reduced to finding the difference of pixels on the horizontal axis and what is known as disparity.

C. Disparity calculation

Se defines as the relative difference of the position of 2 images, for which it uses the semi global matching method SGM, which is based on a computer vision algorithm for the search of dense disparity maps from a pair of rectified stereo images, where it measures the similarity of each pixel of a stereo image with each pixel within the subset of the other stereo image [7]. The algorithm consists of segmenting the image into fixed-size blocks and finding the best match of each block with the reference image. Then, the block with the highest match will be the block with the least re-projection error and that block will be selected as the match block. The result of this algorithm can be seen in Fig. 3.

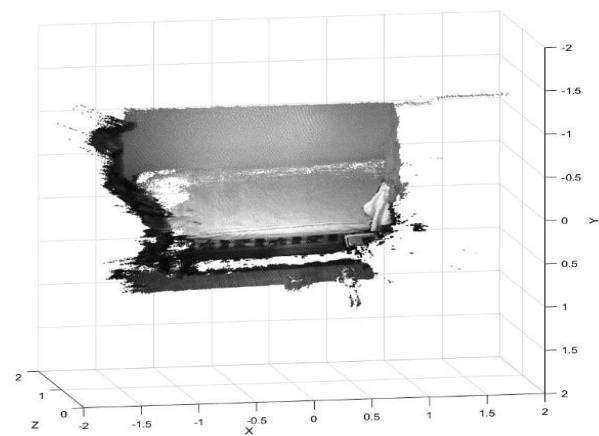
Fig. 3. Disparity map foran indoor environment



D. Point clouds

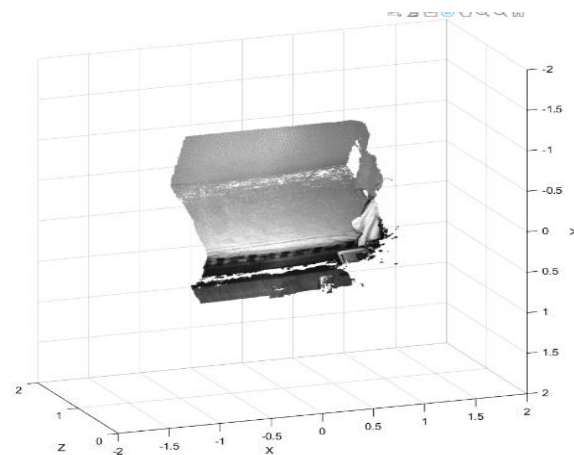
To generate the point clouds, we proceeded to find the depth which is inversely proportional to the disparity, which means that, the greater the disparity between the pair of stereo images, the lower the depth and vice versa. Therefore, the calculation of depth and its fidelity to the physical environment depends on its disparity map. The result of the point cloud obtained is shown in Fig. 4.

Fig. 4. Point cloud as a result of disparity maps



The segmentation is an important point to eliminate the noise of the point cloud obtained, this noise is due to the fact that there is no correct correspondence of pixels so the calculation of the depth is inaccurate, giving divergent points. The segmented cloud can be seen in Fig. 5.

Fig. 5. Segmented point cloud

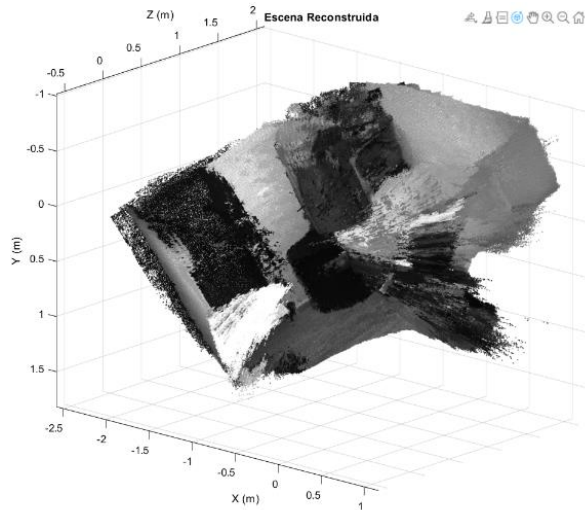


E. Rebuilding the environment

As a fundamental concept for the union / seam of point cloud has the closest iterative point algorithm ICP that has as function the sampling of a set of points of 2 point clouds as input parameters, so a previous point reduction must be made, it is necessary to find the correspondence of points between the 2 point clouds, assigns pairs of points or rejects certain pairs of points that do not have correspondence, assigns an error metric based on the pairs of points and minimizes the error obtained [8]. All this process is done to calculate both the rotation matrix, as well as the translation matrix of the second point cloud with respect to the first. This process is constantly repeated only with the slight change that must preserve a universal matrix of the environment under reconstruction and in which the following point clouds are sewn, so the correspondence between the universal

matrix and the matrix calculated by the ICP algorithm must be found to have a matrix that allows to correctly locate the point cloud that adheres to the environment under reconstruction. The result of this process is shown in Fig. 6.

Fig. 6. Scenery reconstruido

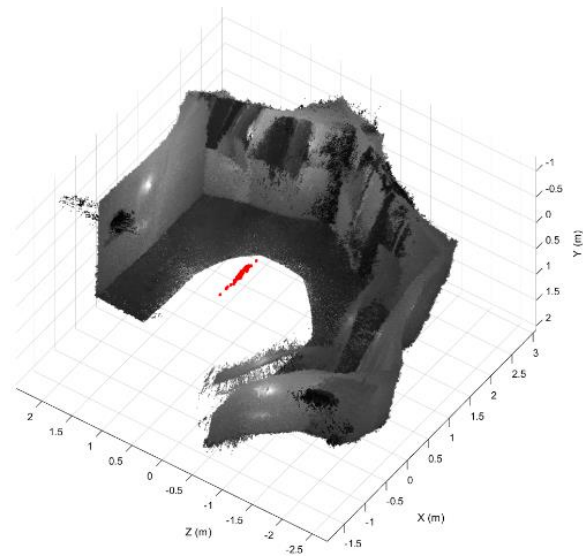


F. Trajectory

To determine the trajectory it was done by analyzing each point cloud that make up the reconstructed environment and the translation vector of each cloud with respect to the environment. To do this, it was done by creating a transformation map of normal NDT distributions from a pre-designed point cloud map of the environment. The NDT map is a compressed and memory-efficient representation suitable for localization. The object converts the point cloud map into a set of voxels (a name that refers to the same concept of pixels in a 2D image, but into a 3D object), each represented by a normal 3D distribution. To draw the trajectory it must be taken into account that this is determined by the position of the stereo camera when capturing the different images, only the translation matrices that resulted from the ICP algorithm and serve to locate the point clouds

that form the environment must be taken into account. Once the point cloud has been found within the NDT map, the location of the camera is determined when the pair of stereo images that generated the point cloud were taken. Each location is drawn in the environment and little by little it is drawing the trajectory of the mobile robot. In Fig. 7 The calculated trajectory for a closed square environment can be observed.

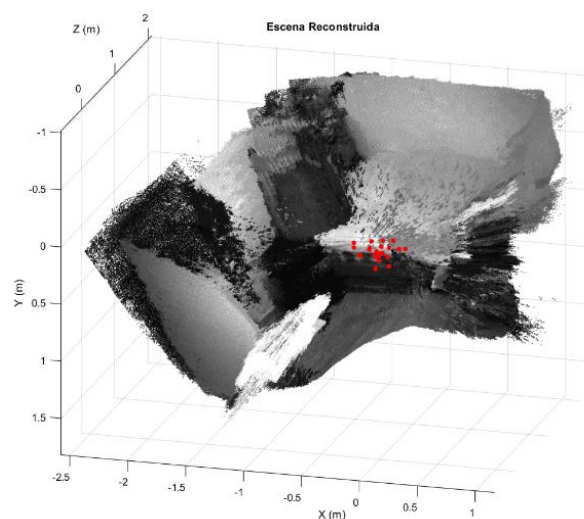
Fig. 7. Calculation of the ray



III. RESULTS

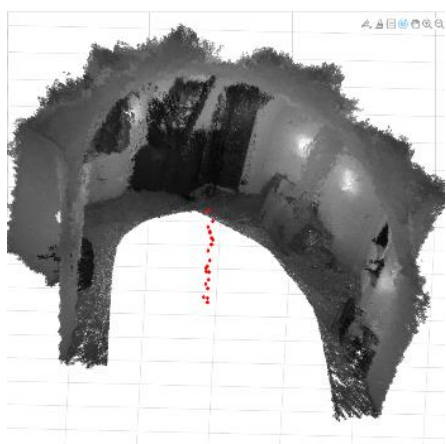
For the analysis of results, reconstruction tests were carried out using 2 environments. One without movement of the robot, only from a static place just by rotating the camera and can be seen in Fig. 7. Where the correct functioning of the trajectory calculation is determined since the points that draw the trajectory are concentrated in a single place.

Fig. 8. Trajectory calculation for Static Robot

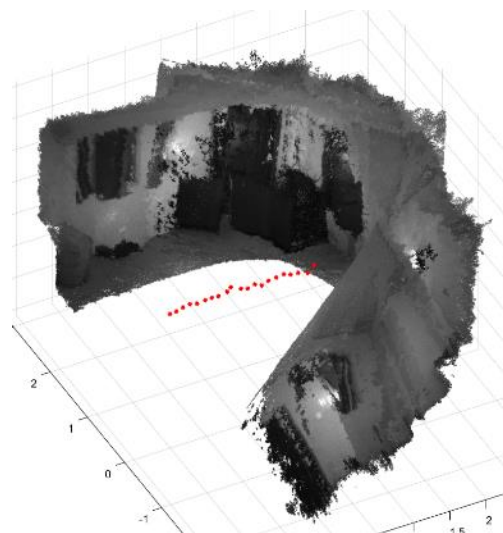


It was taken into account, already with the mobile robot, the calculation of the trajectory a larger stage and with different variants such as square, round objects and of various shapes to see the behavior of the system before the abundance of objects progressively. Giving as results those observed in Fig. 9.

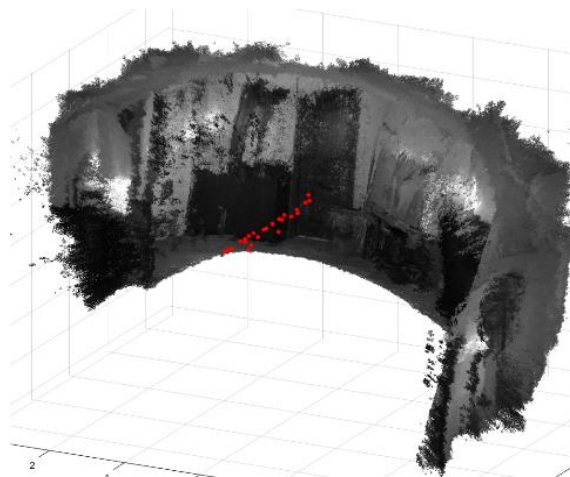
Fig. 9. Reconstructed environment (a) with box s, (b) boxes and cylinders, (c) objects of various shapes



a



b



c

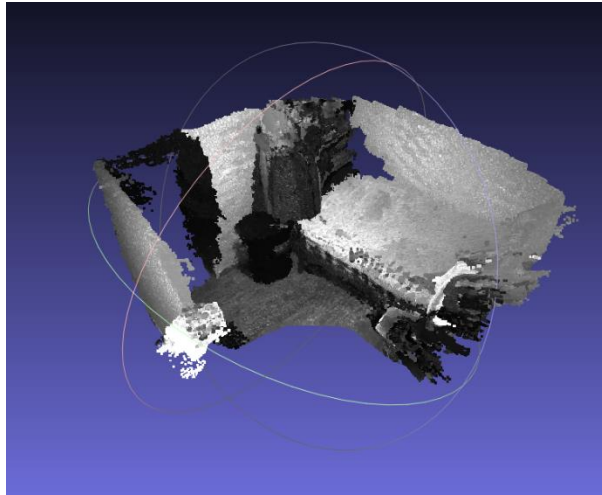
Where to check the fidelity of these virtual objects with physical objects, a calculation of the percentage error of the heights of the reconstructed objects was made. These results are summarized in Table 1.

Table 1. Square Objects

Objects	Real (cm)	Reconstructed (cm)	Error %
1	60	58.32	2.8
2	45	46.76	3.9
3	75	71.47	4.7

The reconstructed scenario presents many areas with noise so through a post-processing using the Meshlab tool the noise was removed to observe the reconstructed environment in a better way. This result can be seen in Fig. 10.

Fig.10 Post-processed environment to eliminate excess noise.



Finally, through a statistical analysis of randomness it was determined if the system has acceptable results or not, so we had parameters to measure such as the execution time of the algorithm, the precision and accuracy of the system. In Table 2. you can see the execution times of the program in Matlab, where it was divided into 4 stages that are: Data import, creation of point clouds, union of point clouds and trajectory. Giving a total of 10 min with 14 sec. The trajectory is where processing is most consumed.

Table 2. Algorithm execution times

Threads	Times (sec)
Importing data	14,768893
Creating point clouds	28,172494
Point Cloud Joining	35,406362
Trajectory	530,317418

Total	608.66
-------	--------

For accuracy, a comparison between the prototype and a commercial system called Colmap was taken into account, photogrammetry as a reconstruction technique and it was taken as a basis to take 30 samples, where the system did not exceed 95% confidence due to its high error in comparison with the commercial system. These results are shown in Table 3.

Table 3. Precision analysis

	Prototype (Stereo Vision)	Colmap (Fotogrametría)
Perimeter	9.978831267 meters	14.481146 meters
Standard deviation	$\pm 0,421681568$	$\pm 0,611988994$
Shapiro Test	0,41100158	0,943538086
F-test (F-value)	0,4747686	Equality of variances
F-Test (Critical F-Value)	0,5374	Equality of variances
T-Test	Does not exceed 95% Trust	
Absolute error	4.50 meters	

For the accuracy analysis, in the same way, a total of 30 samples were taken as a basis with which a comparison was made with the physical environment to observe the fidelity of the virtual environment with this, giving an absolute error of 3.25 meters. These results are shown in Table 4

Table 4. Accuracy analysis

Prototype	Real scenario
9.978831267 meters	13.25 meters
Absolute error of 3.27 meters	

Standard deviation of ± 0.4216 meters	
--	--

IV. CONCLUSION

It was possible to determine that the system visually reconstructs in an acceptable way, but not being faithful to the real dimensions as demonstrated in the statistical analysis of randomness, taking into account that only the stereo vision technique was used for the reconstruction without using another complementary technique that would allow to obtain the exact dimensions of the physical environment, in order to become as faithful as possible to it. So it could be concluded that, in a country where the 3D reconstruction of either environments or objects is just in its infancy, this is a satisfactory job and opens a range of possibilities to improve the project. Systems like this are of great use as far as the area of teleoperation and tele exploration is concerned.

Reference

- [1] A. Bhatti, Stereo Vision. 2008.
- [2] M. M. Martínez, «Stereoscopic vision techniques to determine the three-dimensional structure of the scene», p. 97, 2010.
- [3] G. Pajares and J. M. De la Cruz, Computer Vision. Digital Images and Applications. 2nd Edition. Paracuellos del Jarama (Madrid): RA-MA S.A. Editorial y Publicaciones, 2008.
- [4] D. Córdova, J. Terven, J. Romero, y A. Ramírez, «Three-Dimensional Reconstruction of Indoor and Outdoor Environments Using a Stereo Catadioptric System», Applied Sciences, vol. 10, n. o 24, Art. n. o 24, ene. 2020, doi: 10.3390/app10248851.
- [5] P.-H. Huang, K. Matzen, J. Kopf, N. Ahuja, y J.-B. Huang, «DeepMVS: Learning Multi-View Stereopsis», presentado en Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 2821-2830. Accedido: 15 de febrero de 2023. [En línea]. Disponible en: https://openaccess.thecvf.com/content_cvpr_2018/html/Huang_DeepMVS_Learning_Multi-View_CVPR_2018_paper.html
- [6] Y.-H. Jin, K.-W. Ko, y W.-H. Lee, «An Indoor Location-Based Positioning System Using Stereo Vision with the Drone Camera», Mobile Information Systems, vol. 2018, p. e5160543, oct. 2018, doi: 10.1155/2018/5160543.
- [7] S. I. Delgado, J. F. Troncoso, R. S. Aréchiga, A. A. Trejo, J. V. Maldonado, and J. S. Rodríguez, «Semi-Global-Matching modified for stereo vision in real-time systems», Difu100ci@, Revista de difusión científica, ingeniería y tecnologías, vol. 9, n. or 1, Art. n. o 1, Aug. 2015, Accessed April 25, 2023. [Online]. Available in: <http://difu100cia.uaz.edu.mx/index.php/difuciencia/article/view/129>
- [8] N. Correll, "The iterative closest point algorithm (ICP)", LibreTexts Español, October 30, 2022. [https://espanol.libretexts.org/Ingenieria/Ingenier%C3%ADa/Mec%C3%A1nica/Introducci%C3%B3n_a_los_Robots_Aut%C3%B3nomos_\(Correll\)/12%3A_SLAM_RGB-12.02%3A_El_algoritmo_de_punto_m%C3%A1s_cercano_iterativo_\(ICP\)](https://espanol.libretexts.org/Ingenieria/Ingenier%C3%ADa/Mec%C3%A1nica/Introducci%C3%B3n_a_los_Robots_Aut%C3%B3nomos_(Correll)/12%3A_SLAM_RGB-12.02%3A_El_algoritmo_de_punto_m%C3%A1s_cercano_iterativo_(ICP)) (accessed April 25, 2023).