CONEM 2022
CONGRESSO NACIONAL DE ENGENHARIA MECÂNICA
TERESINA - PI

ABCM

XI CONGRESSO NACIONAL DE ENGENHARIA MECÂNICA
DE 07 A 11 DE AGOSTO DE 2022, TERESINA-PI, BRASIL

POSE MEASUREMENT OF A PARALLEL ROBOT MANIPULATOR USING AUGMENTED REALITY FIDUCIAL MARKERS

Leonardo Afonso Ferreira Bortoni, bortonileonardo@gmail.com[1]
Mauricio de Campos Porath, mauricio.porath@ufsc.br[1]
Lucas de Camargo Souza, lucas_camargo@hotmail.com.br[1]

[1]Universidade Federal de Santa Catarina, Centro Tecnológico de Joinville, Laboratório de Geodésia Industrial, R. Dona Francisca, 8300 - Bloco U - Zona Industrial Norte, Joinville - SC, 89219-600.

Abstract: *The control of robotic manipulators is a fundamental task to ensure accuracy and reliability of industrial processes. The real-time acquisition of high accurate pose- data of a manipulator, such as to enable online error correction, usually demands expansive equipment. In this work we explore the employment of augmented reality fiducial markers and a stereo digital camera as a low-cost alternative to track the pose of a Stewart platform. Three approaches were implemented and assessed. The first consists of obtaining the pose of the robot using just one fiducial marker. The second states the pose of a point amidst a plane generated by the position information from three markers. And the third repeats the last, adding a photogrammetry algorithm. Data fusion is implemented in the third scenario. Fiducial markers detection and computation is performed by the ArUco library, using the ROS environment coded in C++. The photogrammetry algorithm is implemented by means of the OpenCV library, also coded in C++. The stereo camera employed was a ZED2 model. Its intrinsic and extrinsic camera parameters are provided by the manufacturer. A robotic total station was employed in the frame calibration. The best strategy resulted in average positioning errors smaller than 1 mm, and orientation errors smaller than 0.2 °, with respect to the nominal pose stated by the previously calibrated Stewart platform.*

Keywords: *fiducial marker, photogrammetry, pose measurement, stewart platform.*

1. INTRODUCTION

The measurement of a robot's pose in space is required for both online and offline error correction schemes. These strategies are used to enhance pose accuracy and thus improve process quality and reliability. The applicability of sophisticated real-time six degrees of freedom (6-DoF) measurement system, such as *indoor*-GPS (Porath *et al.*, 2019) and *laser trackers* (Muralikrishnan *et al.*, 2016), has been demonstrated. However, the high cost of these equipment is still a limiting factor. As an alternative, low cost solutions that employ digital cameras and image processing can be explored.

A set of image processing techniques with fiducial markers was originally developed for augmented reality applications and is now widely employed in dynamic robotics. In this context, an augmented reality fiducial marker (AR-FM) is an encoded plane marker that is positioned in the environment or on the object to be tracked and observed by a camera (Muñoz-Salinas *et al.*, 2018). After the calibration of the system, it is possible to estimate via AR-FM the pose of the marker, or the camera pose in a given coordinated system, in real time, by evaluating the positions of the marker vertices on the image. Besides, when one has the coordinate in pixels of the same point in space with several cameras placed on different spots, it is possible to integrate online close-range photogrammetry algorithms (O-CRP) to estimate the position of this point in space. Since managing and synchronising frame images from several cameras is not easily practicable, an interesting approach would rely on the deployment of stereo cameras to empower O-CRP.

In this work we explore both AR-FM and O-CRP, employing digital image processing to track in real time markers attached to a calibrated parallel robot manipulator, namely a Stewart platform, and through the pose of the markers estimate the pose of the robot. A stereo camera is employed. Three measurement strategies were implemented and evaluated. The first estimates the pose of the robot from calculations of a single marker using just AR-FM. The second explores the performance of pose estimation from three markers, also using AR-FM, but only the position information of each marker. The third scenario repeats the last, but employing also the O-CRP strategy, so that a data fusion is deployed between AR-FM and O-CRP in order to evaluate the manipulator's pose.

2. Materials and Methods

In this section we describe the experimental setup and methodology.

2.1 Augmented reality fiducial marker and photogrammetry

In this work we employed the AR-FM technology from Garrido-Jurado *et al.* (2015), named as ArUco. The algorithm consists of applying image processing techniques to detect the black outline of a marker (Figure 1) and consequently its corners. Known the pixel coordinates of the corners, a series of computations and estimations are performed to finally determine the pose of the centre of the detected marker. The estimation of the marker's pose considers the marker's resolution, in pixels, in the image frame processed. Therefore, the more pixels represent the marker, more accurate is the estimation of the pose. Thus, markers that are closer to the camera, or larger in size, will be more accurately characterised. The resolution of the image frame is also relevant in this context.

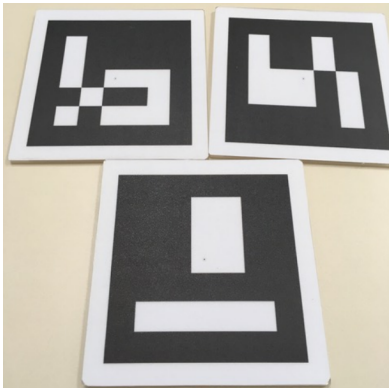

Figure 1: **ArUco's fiducial markers.**

Each marker has a unique code that allows the detection of multiple markers in the same image frame. In this work the maximum amount of markers deployed was three. We used an open source library, coded in C++, that the developers of ArUco made available (Romero-Ramirez *et al.*, 2018). The library has been assessed by Kalaitzakis *et al.* (2021) and the authors concluded that ArUco has good position and orientation results, great detection rate and low computational cost, when contrasted against other AR-FM technologies. The authors also alert that ArUco is sensitive to smaller marker's size and larger distances, and that the computational cost scales with multiple markers.

There are several photogrammetry image processing techniques available. Since the goal of this work is to estimate the pose of a manipulator from data obtained form a set of markers, we adopted the linear triangulation technique (Hartley and Zisserman, 2003). We used the implementation in C++ of the algorithm provided by the OpenCV open source library (Bradski, 2000).

The execution of the triangulation requires the image pixel frame, the coordinate in pixels of the same point seen by two cameras and the projection matrix of the two cameras. All the intrinsic and extrinsic camera parameters employed in this work were adopted as stated by the manufacturer's datasheet. The triangulation also requires rectification of the image frame, thus the very same rectified frame of pixels is deployed in ArUco and in the triangulation's algorithm. The pixel coordinate of the centre of the marker is yielded by ArUco's application programming interface (API). The Figure 2 pictures the procedure's sequence. In order to render an application that manages the camera operation and the ArUco and triangulation computations, we employed the Robot Operation System (ROS) environment (Quigley *et al.*, 2009), which contributes to ease the exchange of data among modules through a node based publisher-subscriber paradigm.

2.2 Stewart platform

This work has been carried out in the UFSC[1] laboratory LGI[2], which has its own Stewart platform. The robot was originally designed and constructed to investigate the feasibility of automated alignment and assembly processes in the shipbuilding industry. The nominal pose of the robot regards the position and orientation at the centre point of its movable platform w.r.t. a stationary frame attached to the robot's base. It is adopted the ZYX (roll, pitch, yaw) angle convention, assuming α, β and γ being the rotations of the platform frame in respect to the base frame about axes x, y and z.

The Stewart platform was previously calibrated (determination of the kinematic parameters) using an indoor-GPS and a total station. The average pose errors after calibration are presented in Table 1. The calibration and pose accuracy evaluation is presented by Porath *et al.* (2020).

[1]Universidade Federal de Santa Catarina
[2]Laboratório de Geodésia Industrial

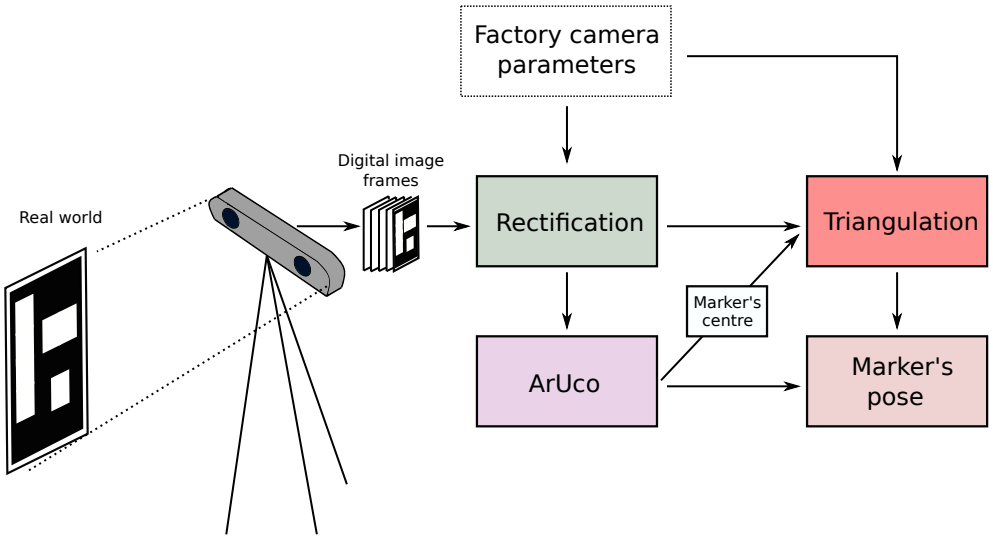

Figure 2: **Sequence of dependences and operations performed to achieve the marker's pose.**

	$\overline{E}_x$ [mm]	$\overline{E}_y$ [mm]	$\overline{E}_z$ [mm]	$\overline{E}_\alpha$ [°]	$\overline{E}_\beta$ [°]	$\overline{E}_\gamma$ [°]
Average errors	-0.25	0.27	0.09	0.052	-0.025	0.023

Table 1: **Stewart platform average pose errors.**

2.3 Robotic total station

A robotic total station was employed in this work to enable a mathematical correlation among the coordinate frames of the camera, the movable platform and the fixed base. Total stations are measurement equipment evolved from theodolites. The LGI owns a model Leica TS 12, produced in 2015. The precision (repeatability), according to ISO 17123-4 (International Organization for Standardization, 2012), is 1.0 mm + 1.5 ppm for length measurement and 7 " for angle measurement.

2.4 Camera ZED2

A digital stereo camera of model ZED2, manufactured by Stereolabs, was employed in this work. The maximum camera's resolution, which was deployed, is 2208x1242. For the application hereby presented, a frame rate of 15 fps was configured. The camera was powered through USB 3.0 connected to a computer, which is also the communication protocol of the device. All the methods and libraries employed to communicate with the camera were programmed in C++, according to the API provided by the manufacturer. Figure 3 portraits the camera device used in this work.

Figure 3: **Stereo ZED2 camera.**

2.5 Pose measurement strategies

The final pose measurement at the centre point of the moving platform w.r.t. the fixed base, deploying the camera, is deducted from a calibration procedure that correlates the system coordinate frame B,P, and F (Base, Platform and Fiducial

marker, respectively) measured by the robotic total station in the frame TS, with the frame F measured by the camera in the frame C. The Figure 4 illustrates the concept.

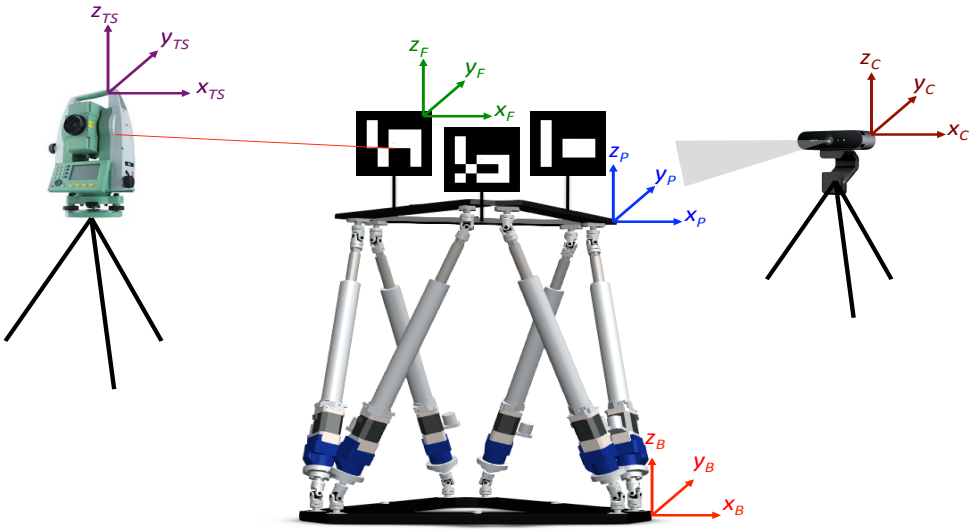

Figure 4: **Calibration approach representation.**

This calibration is performed with the moving platform at a fixed initial pose, termed as position *home*. Three fiducial markers yield the homogeneous transformation matrices (HTM) $\mathbf{T_C^B}$ and $\mathbf{T_P^F}$. Afterwards, when performing movements with the platform and measuring the markers' pose with the camera, it is possible to deduct the pose at the platform's centre point from Equation 1 and find the translation and orientation values from Equation 2, 3, 4 and 5, respectively.

$$\mathbf{T_P^B} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ t_{31} & t_{32} & t_{33} & t_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{T_C^B T_F^C T_P^F} \tag{1}$$

$$\mathbf{o_P^B} = [x_P, y_P, z_P]^T \tag{2}$$

$$R_\alpha = atan2(t_{32}, t_{33}) \tag{3}$$

$$R_\beta = atan2(-t_{31}, \sqrt{t_{32}^2 + t_{33}^3}) \tag{4}$$

$$R_\gamma = atan2(t_{21}, t_{11}) \tag{5}$$

It is possible to derive the very same calibration and approach for just one single marker. In this case, the marker's corners are used to simulate the other two markers. Since the markers' size is know, namely a square of size 100 mm, and since the HTM from the marker's frame w.r.t. the camera frame is also provided via ArUco, two other markers centre can be mimicked and the same procedure so far presented might be followed.

As mentioned, this work aims to evaluate three strategies of pose measurement using digital video frames. The strategies are as follows.

2.5.1 Strategy 1

Strategy 1 employs just one marker, and the pose is estimated with AR-FM. The stereo camera yields two image frames, for the left and right camera. Thus, the ArUco's AR-FM algorithm is deployed for both the image frames. The average of the pose reported by both cameras is taken as final pose measurement at the platform's centre point.

2.5.2 Strategy 2

Strategy 2 employs three markers. Only position data, provided by AR-FM, is used for each marker. The pose is calculated from these three point positions. As for strategy one, here the average value of both cameras is taken as final measurement.

2.5.3 Strategy 3

Strategy 3 employs three markers. The pose is estimated from AR-FM and O-CRP. As in Strategy 2, AR-FM is used to determine the position of the center point of each marker with each camera and the pose is calculated from these points. Additionally, the position of each marker is determined by O-CRP using both cameras simultaneously and the pose

calculated from these points. Data fusion is performed attributing weights to each of both pose measurements reported. Let PL, PR and PT be the pose reported by the left camera, right camera and O-CRP, respectively, and let the indices T and R stand for the translation and rotation parts of the pose. The following equations are proposed to state the final pose measurement:

$$Pose_T = c_1 \cdot PL_T + c_2 \cdot PR_T + c_3 \cdot PT_T \tag{6}$$

$$Pose_R = c_4 \cdot PL_R + c_5 \cdot PR_R + c_6 \cdot PT_R \tag{7}$$

Where $[c_1, c_2, c_3]$ and $[c_4, c_5, c_6]$ are weight coefficients of real values, so that $c_1 + c_2 + c_3 = 1$ and $c_4 + c_5 + c_6 = 1$. The values of theses constants are determined empirically. The goal of this strategy is to explore the redundancy of data in order to reduce biases in the final measurement, contributing therefore to a higher accuracy.

2.6 Evaluation experiments

A set of verification measurements has been performed to evaluate the accuracy of the pose measurement after implementing each one of the strategies. A total of 13 different poses has been selected, shown in Table 2. Poses 1 to 5 refer to positions placed on a diagonal plane of the workspace of the Stewart Platform as shown in Figure 5. This pattern follows the recommendation of ISO 9283:1998 (International Organization for Standardization, 1998).

Pose nº	x [mm]	y [mm]	z [mm]	α [°]	β [°]	γ [°]
1	-120	120	1235	0	0	0
2	-120	-120	1235	0	0	0
3	0	0	1175	0	0	0
4	120	120	1115	0	0	0
5	120	-120	1115	0	0	0
6	0	0	1175	-5	0	0
7	0	0	1175	5	0	0
8	0	0	1175	0	-5	0
9	0	0	1175	0	5	0
10	0	0	1175	0	0	-5
11	0	0	1175	0	0	5
12	0	0	1175	-5	-5	-5
13	0	0	1175	5	5	5

Table 2: **Poses for performance evaluation of the three strategies.**

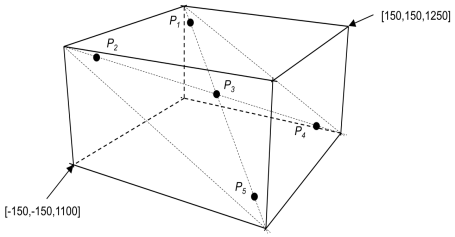

Figure 5: **Measurement volume.**

The 13 poses were evaluated three times, resulting in 39 pose measurements. For each measurement, a new frame calibration was carried out. The three strategies were processed with the same set of image frames, for each pose. The measurement of each pose was carried out 100 times, discarding the first 30. So that the final measurement of a pose, for each strategy, is the average value of 70 samples.

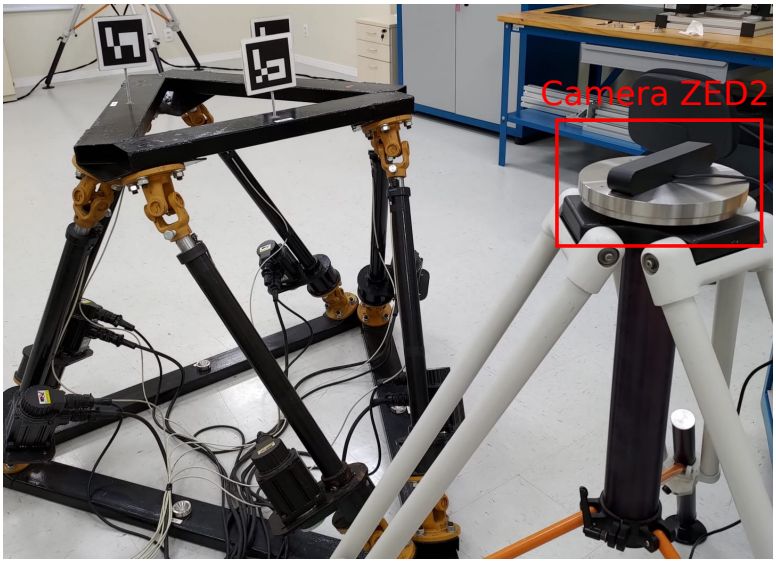

Figure 6: **Camera and robot experiment's setup.**

The final measured pose reported by each strategy is compared to the nominal robot pose.

The Figure 6 portraits the placement of the markers, the Stewart platform and the camera. The distance between the camera and the closer marker is approximately 120 mm.

3. RESULTS

Pose errors of each camera using ArUco for pose estimation through one single marker (strategy 1), are presented in Figure 7. The pose error is computed as the difference of the measured pose from the nominal pose. The figure presents the performance of just one of the three pose measurement runs. The repeatability of strategy 1, measured as the greatest absolute value of the difference between the maximum and the minimum error value among the three runs of a pose, among all the poses' numbers, is as presented in Table 3.

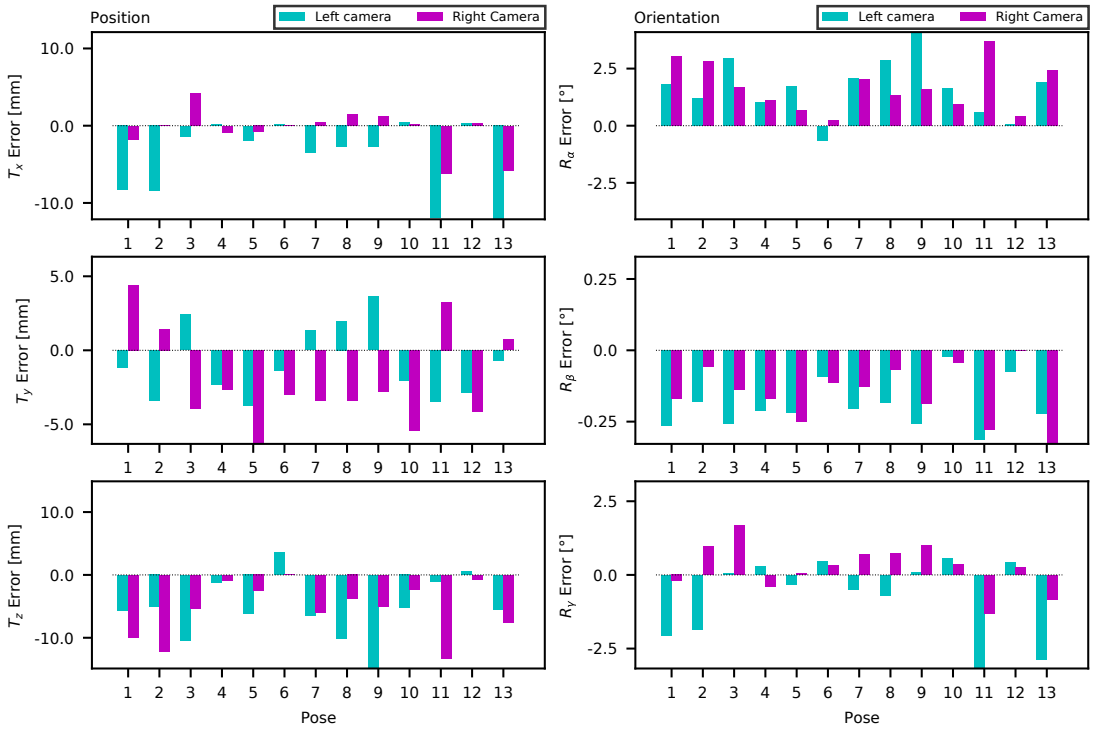

Figure 7: **Pose error performance of data for strategy 1.**

The pose errors for strategies 2 and 3, where the pose estimation is deducted from three markers, are presented in Figure 8. The figure presents the outcomes of strategy 2, namely from each of the cameras, and for strategy 3, where both cameras are used simultaneously in a O-CRP scheme. Again, the figure presents the performance of just one of the three runs. The repeatability of strategy 2 and 3, measured as the greatest absolute value of the difference between the maximum and the minimum error value among the three runs of a pose, among all the poses' numbers, is as presented in Table 3.

	Repeatibility		
	Strategy 1	Strategy 2	Strategy 3
T_x [mm]	4.1	2.9	0.8
T_y [mm]	3.7	3.5	1.0
T_z [mm]	14.2	0.2	0.1
R_α [°]	3.5	0.1	0.0
R_β [°]	0.1	0.0	0.0
R_γ [°]	1.1	1.0	0.2

Table 3: **Repeatability of each strategy of pose estimation.**

For strategy 3, the coefficients of Equations 6 and 7 were empirically optimised, aiming at a general improvement of performance. The resulting coefficients are:

$$[c_1, c_2, c_3] = [0.4, 0.4, 0.2]$$

$$[c_4, c_5, c_6] = [0.25, 0.25, 0.5]$$

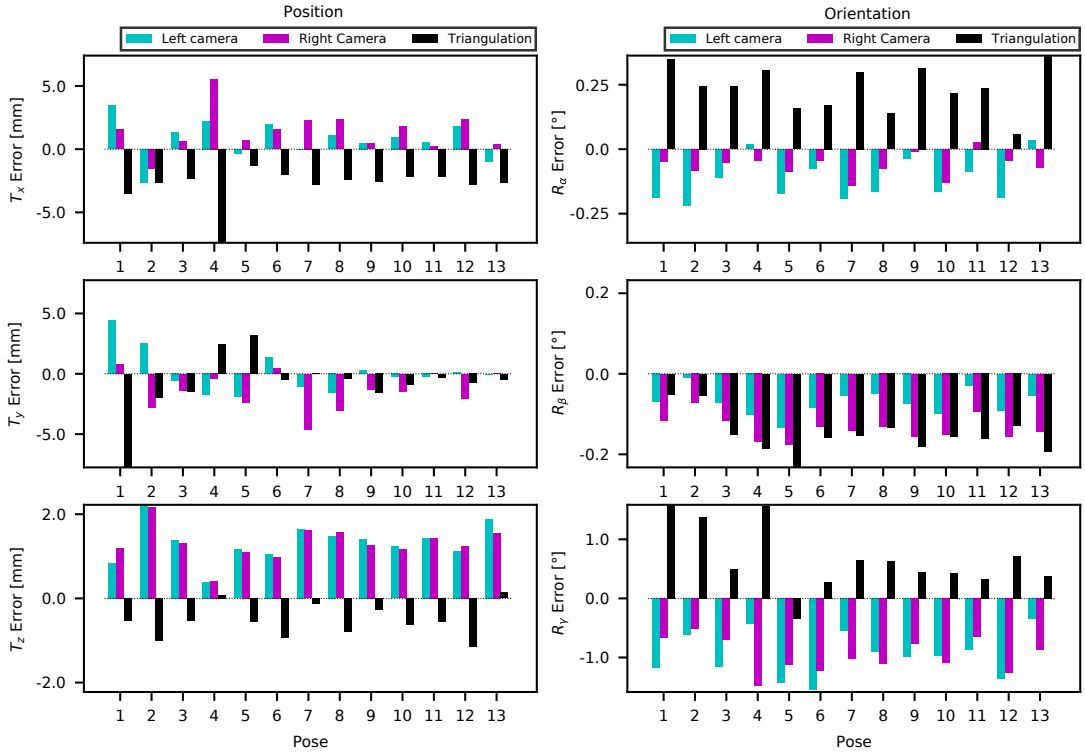

Figure 8: **Pose error performance of data for strategy 2 and 3.**

Figure 9 presents the average of the absolute pose errors, considering all poses and all measurement runs, for each strategy. The error bars represent the standard deviation of all measurements for each pose. The nomenclature S_i stands for strategy 1, 2 or 3. Table 4 presents the data from the figure.

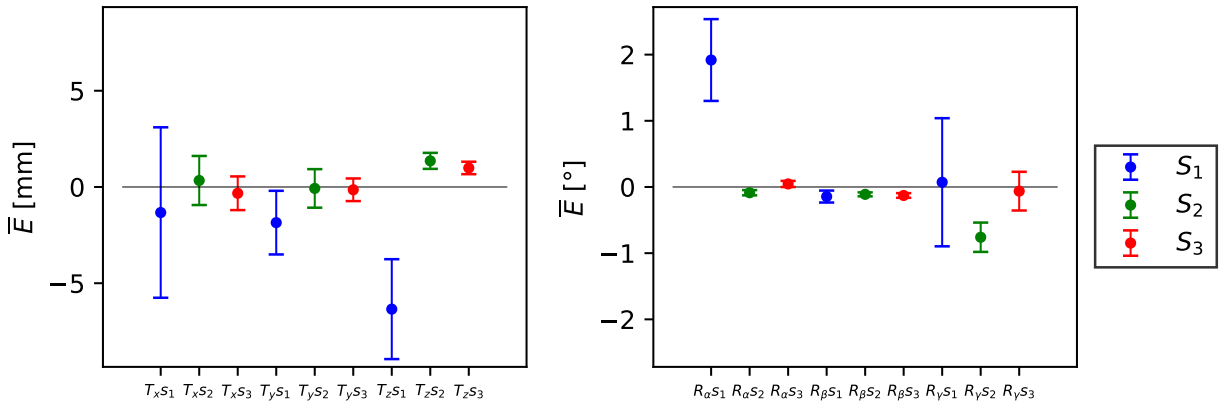

Figure 9: **Average pose's errors, respectively for each strategy.**

	Strategy 1		Strategy 2		Strategy 3	
	$\overline{E}$	s	$\overline{E}$	s	$\overline{E}$	s
T_x [mm]	-1.3	4.4	0.34	1.3	-0.33	0.88
T_y [mm]	-1.9	1.7	-0.075	1.0	-0.15	0.59
T_z [mm]	-6.3	2.6	1.4	0.42	0.99	0.32
R_α [°]	1.9	0.62	-0.087	0.040	0.46	0.047
R_β [°]	-0.14	0.090	-0.11	0.030	-0.13	0.035
R_γ [°]	0.071	0.97	-0.76	0.22	-0.063	0.29

Table 4: **Average and standard deviation of the pose's errors of each strategy.**

4. DISCUSSION AND CONCLUSION

The performance of pose measurement using just one marker has achieved results with the highest bias and variability.

The pose estimation using three markers has been proved more accurate than the estimation from one single marker, and the strategies' results presented a good repeatability, especially strategy 3. Overall, translation errors were below 4 mm, and orientation errors 1°, in average.

The integration of O-CRP triangulation algorithm and data fusion could contribute to enhance measurement accuracy concerning the translation along the z-axis and orientation around the same axis. But apart from this benefit, we could not identify other advantages that justify its employment.

The employment of the strategy 1 might be justified in some applications with line-of-sight limitations, and just the identification of less than 3 markers is feasible. Yet, the pose measurement from the three markers approach, deployed in strategy 2 and 3, achieved a competitive accuracy. The fact that the robot's positioning errors are in the same order of magnitude of the outcomes of our experiments might suggest that part of the pose's errors can actually come from the positioning uncertainty of the manipulator. However, this hypothesis has not yet been further investigated.

In this work we did not investigate the distance of the camera to the markers. All the procedures reported in this work were carried out with the camera distancing 120 mm from the closest marker. However, since it is reported that the ArUco's accuracy is sensible to the distance (Kalaitzakis *et al.*, 2021), it is expected that the achievable performance with the three strategies here employed would be affected.

From the measurements performed in this work using strategy 3, one may conclude that it is quite competitive when assessing the translation along the x and y-axis, and the orientation around the z-axis. However, performance in z-direcion and rotations around the x an y-axis are rather poor. This is expected, since the later measurements strongly depend on measurements taken in the direction of the optical axis of the camera.

The authors conclude that using a high accuracy equipment is still preferable for applications that require high accuracy and precision, however, the low cost strategy might be an alternative, depending on the application's requirements.

5. ACKNOWLEDGEMENT

This study was partially supported by CNPq (National Council for Scientific and Technological Development).

6. REFERENCES

Bradski, G., 2000. "The OpenCV Library". *Dr. Dobb's Journal of Software Tools*.

Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. and Medina-Carnicer, R., 2015. "Generation of fiducial marker dictionaries using mixed integer linear programming". *Pattern Recognition*, Vol. 51. doi:10.1016/j.patcog.2015.09.023.

Hartley, R. and Zisserman, A., 2003. *Multiple view geometry in computer vision*. Cambridge university press.

International Organization for Standardization, 1998. "ISO 9283:1998 Manipulating industrial robots - Performance criteria and related test methods".

International Organization for Standardization, 2012. "ISO 17123-4:2012 Optics and optical instruments – Field procedures for testing geodetic and surveying instruments – Part 4: Electro-optical distance meters (EDM measurements to reflectors)".

Kalaitzakis, M., Cain, B., Carroll, S., Ambrosi, A., Whitehead, C. and Vitzilaios, N., 2021. "Fiducial markers for pose estimation". *Journal of Intelligent & Robotic Systems*, Vol. 101, No. 4, pp. 1–26.

Muñoz-Salinas, R., Marín-Jimenez, M.J., Yeguas-Bolivar, E. and Medina-Carnicer, R., 2018. "Mapping and localization from planar markers". *Pattern Recognition*, Vol. 73, pp. 158–171. ISSN 00313203. doi:10.1016/j.patcog.2017.08.010.

Muralikrishnan, B., Phillips, S. and Sawyer, D., 2016. "Laser trackers for large-scale dimensional metrology: A review". *Precision Engineering*, Vol. 44, pp. 13–28. ISSN 01416359. doi:10.1016/j.precisioneng.2015.12.001. URL `http://dx.doi.org/10.1016/j.precisioneng.2015.12.001`.

Porath, M.D.C., Bortoni, L.A.F., Simoni, R. and Eger, J.S., 2020. "Offline and online strategies to improve pose accuracy of a Stewart Platform using indoor-GPS". *Precision Engineering*, Vol. 63, No. May, pp. 83–93. ISSN 0141-6359. doi:10.1016/j.precisioneng.2020.01.003. URL `https://doi.org/10.1016/j.precisioneng.2020.01.003`.

Porath, M.d.C., Simoni, R., Nunes, R.d.A. and Bertoldi, A., 2019. "Feasibility of measurement-assisted assembly of ship hull blocks". *Marine Systems & Ocean Technology*, Vol. 14, No. 1, pp. 23–33. ISSN 2199-4749. doi:10.1007/s40868-018-00053-w. URL `http://dx.doi.org/10.1007/s40868-018-00053-w`.

Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Wheeler, R., Ng, A.Y. *et al.*, 2009. "Ros: an open-source robot operating system". In *ICRA workshop on open source software*. Kobe, Japan, Vol. 3, p. 5.

Romero-Ramirez, F., Muñoz-Salinas, R. and Medina-Carnicer, R., 2018. "Speeded up detection of squared fiducial markers". *Image and Vision Computing*, Vol. 76. doi:10.1016/j.imavis.2018.05.004.

7. COPYRIGHT LIABILITY

The authors are solely responsible for the content of this paper.