An Analysis on Camera Pose Estimations based on Varying Physical Variables of ArUco Markers

Nicanor R. Roxas Jr
Evelyn D. And - Institute of Biomedical Engineering and Health Technologies (EDA - IBEHT)

Aron Jake P. Vidad
Evelyn D. And - Institute of Biomedical Engineering and Health Technologies (EDA - IBEHT)

Rudwin Paul B. Abastillas
Evelyn D. And - Institute of Biomedical Engineering and Health Technologies (EDA - IBEHT)

Renann G. Baldovino
Evelyn D. And - Institute of Biomedical Engineering and Health Technologies (EDA - IBEHT)

他

https://doi.org/10.5109/7395509

出版情報：Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES). 11, pp.137-143, 2025-10-30. International Exchange and Innovation Conference on Engineering & Sciences
バージョン：

An Analysis on Camera Pose Estimations based on Varying Physical Variables of ArUco Markers

Nicanor R. Roxas Jr[1,2], Aron Jake P. Vidad[1], Rudwin Paul B. Abastillas[1], Renann G. Baldovino[1,2], Nilo T. Bugtai[1,2], Elmer P. Dadios[2], Ryan Rhay P. Vicerra[2], Argel A. Bandala[3], Aaron Raymond See[4]
[1]Evelyn D. And - Institute of Biomedical Engineering and Health Technologies (EDA - IBEHT), [2]Department of Manufacturing Engineering and Management (DMEM), [3]Department of Electronics and Computer Engineering (DECE), De La Salle University (DLSU), 2401 Taft Avenue, 0922 Manila, Philippines
[4]Department of Electronic Engineering, National Chin-Yi University of Technology, No. 57, Section 2, Zhongshan Rd, Taiping District, Taichung City, 411030, Taiwan
Corresponding author email: nicanor.roxas@dlsu.edu.ph

Abstract: *Alignment of point clouds and camera pose estimations are key concepts when discussing modeling of real-world objects and space. Proper alignment techniques enable accurate digital representation of objects, which has uses across industries such as extended reality, unmanned vehicles, and medicine. One method of alignment is the use of stationary markers, such as ArUco markers, which is the focus of the study. The research demonstrates varying physical characteristics of the ArUco markers such as amount and size, and comparison of the point cloud alignment gathered from the RealSense D455 camera. The alignment procedure was done via ArUco marker image processing, finding transformation matrixes, and using these matrixes for alignment. The data resulting from the alignment showed a possible optimal number of markers to be used with the D455 camera, along with the marker size and number of planes to be used.*

Keywords: camera pose estimation, realsense, ArUco

1. INTRODUCTION

Creating a virtual representation, i.e. a model, of existing physical objects is a key aspect in the latest technologies of the 21st century in fields such as extended (virtual or augmented) reality, unmanned aerial vehicles [1], and even medicine, such as prosthesis [2], or creation of medical training tools [3]. A popular technique for creating a digital representation of a real object is via 3D scanning, where devices emit signals and calculate the distance of objects around the area of interest based on a signals return. However, a single scan can only provide one side of a 3D object, and therefore multiple scans must be performed to view as many sides of the object as necessary. The scans are then aligned to one another, or to a suitable reference frame, to obtain an accurate digital model of the object of interest.

Point cloud alignment and camera pose estimation are two concepts that are strongly connected. Point cloud alignment involves matching or "stitching" points in 3D-space from two or more different angles. To gain a more comprehensive render of a 3D object, it is necessary to acquire point clouds from different angles which can come from one source that is moved in space or from an array of cameras. However, movement of the source also entails movement of the coordinate axis' origin and therefore this necessitates aligning the point clouds appropriately using methods such as iterative-closest-point (ICP) algorithm [4]. On the other hand, camera pose estimation is the approximation of the location and angle of a point cloud source, usually a camera, relative to a given universal coordinate frame. The two concepts, point cloud alignment and camera pose estimation, are closely linked since alignment of point clouds can be calculated based on the location and angle of the camera source, and the camera pose estimation can vice versa be calculated based on the transformation matrix used for point cloud alignment.

A common method of estimating the camera pose is by using fiducial markers, such as ArUco [5] markers, of which has improved over the previous decade in terms of robustness [6], and speed [7]. In the paper by Garrido-Jurado et al [6] they describe a methodology of creating fiducial markers that prevent confusion for marker detection, error checking, and a reduction of error due to occlusion. Detection rate of fiducial markers has also been increased in the paper by Romero-Ramirez et al [7], allowing for faster camera pose estimation calculations. Fiducial markers have also been used in tandem with cameras for applications in physics such as in the paper by Cepon et al [8], where they explored the response of a structure to an impact excitation via impact-pose estimation using ArUco markers. With the advent of RGB-D cameras, camera pose estimation techniques were further researched such as by Aalerud et al [9], where they used ICP to calibrate an array of RGB-D cameras using fiducial markers. In Madeira et al [10], they also utilized RGB-D cameras for alignment and further added the removal of the fiducial markers after having aligned the point clouds to one another.

With 3D depth perception technology becoming increasingly commonplace, it is imperative to perform a thorough analysis on all the aspects of the said technology, including further examination on properties of fiducial markers for alignment. In this study, the focus is on finding the optimal number and size of ArUco markers that allow for the closest digital representation of physical space. This will allow for optimization of camera pose estimation techniques which can be useful for technologies such as live point cloud streaming. Moreover, this paper is a part of a larger research project that focuses on accurate distance estimation based on live camera streams. Therefore, the paper is limited to the analysis of the effect of varying the physical aspects of the ArUco markers.

2. DESIGN METHODOLOGY

The framework to be used includes the creation of the ArUco markers and their detection, variables to be explored, set-up of the testing areas and camera settings, alignment method, and the metrics for evaluation of the alignment. The detailed flow of the framework can be seen in Figure 1. This section will explore each part of the framework in more detail, along with other limitations.

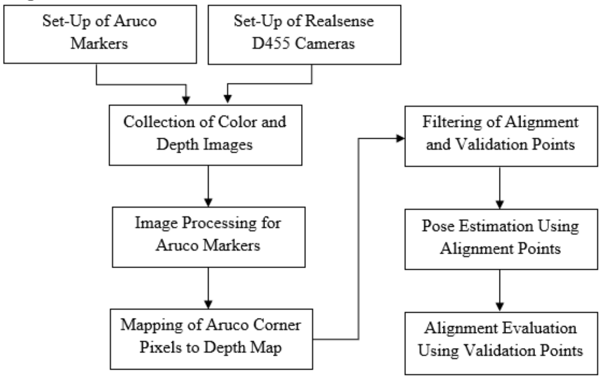

Fig. 1. Conceptual Framework

It should be noted that the camera used for this study is the Intel RealSense D455 RGB-D camera [11]. This camera was selected as it has an extensive python library (*pyrealsense2*) [12] which supports alignment of the color and depth images produced by the camera.

2.1 Set-Up

A sample of the RealSense camera can be seen in Figure 2, and the corresponding configurations for the color and depth images can be seen in Table 1.

Due to the depth limitation RealSense cameras, which is set at 6 meters from the technical specifications, the area of where the study will be conducted should not exceed 4.6 meters in either length or width. This allows the images of the cameras to overlap at least one-third of the length of the area's side. Figure 3 illustrates the visualization and calculations performed to acquire the maximum side length to be used for the study. It should also be noted that the height of the camera was set to 2.1 meters and that the distance equation uses 3 dimensions for the calculations.

Fig. 2. RGB-D Camera Hardware (RealSense D455)

Table 1. Software configuration of camera frames

Frame Type	Width (pixels)	Height (pixels)	Number Per Second (FPS)	Frames Captured
Color	1280	720	15	20
Depth	1280	720	15	20

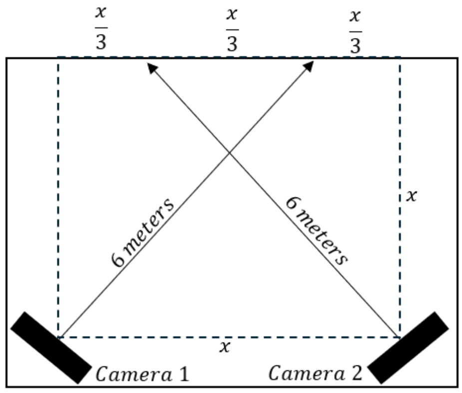

Fig 3. Maximum room dimension calculation

Given the aforementioned height of 2.1 meters, the distance formula can then be calculated using the following equation,

$$x^2 + (x/3 + x/3)^2 + 2.1^2 = 6^2$$

Which gives the length (x) a value of 4.6765 meters. This value is the maximum working dimension that will be used for both the length and width in the study.

2.2 ArUco Markers and Image Processing

ArUco Markers are a part of the OpenCV library that generate QR-like images that can be used for pose-estimation applications. As the markers would be printed with varying sizes, it was best to allocate the first digit of the marker's index as an indicator of its size for ease of naming and image processing. Therefore, the 7x7 dictionary was used as this had 1000 generated markers of where 50 markers were selected based on their index numbers. Of the 136 markers, 16 were printed individually while the other 120 were grouped into 4, and were printed to create a checkerboard-like pattern with varying sizes. The design of the printed markers can be seen in Figure 4, and the index ranges with their corresponding sizes can be seen in Table 2.

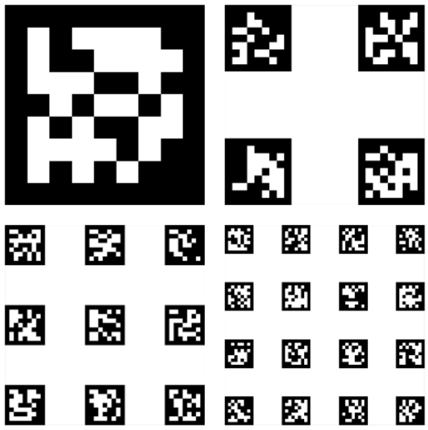

Fig. 4. Checkerboard creation of ArUco markers

Table 2. ArUco marker indexes and sizes

Index	Number of Markers	Marker Size
1-16	16	30.00 cm
104	4	50.00 cm
301 – 316	16	16.67 cm
501 – 536	36	10.00 cm
901 – 964	64	7.14 cm

For image processing, 20 images are firstly collected from each camera. Each frame undergoes alignment of its color image and depth maps using the RealSense python library. Each frame is then identified for ArUco markers utilizing the OpenCV library. The corners of the ArUco markers are then given a pixel coordinate which is cross-reference with the aligned depth map. By following this process, it is possible to find the index of identified ArUco markers, along with identifying the location of each marker's four corners in space, relative to the camera of interest.

2.3 Variables to Study

The study will focus on varying three physical qualities of ArUco markers, namely amount, positioning, and size. For the amount and positioning, this has been merged into one case study as varying the amount in two orthogonal planes tackles both physical qualities.

2.3.1 Amount and positioning

To test the effect of the amount and positioning of the ArUco markers, the study will limit the number of ArUco markers for alignment to six markers only and limit the study to only one pair of orthogonal planes. The number of markers per corresponding plane for each iteration of the study can be seen in Table 3. It should be noted that a symmetry is assumed between using the planes with their intersection acting as the line of symmetry. Therefore, no repetition is needed wherein the two planes would be swapped as plane 1 and plane 2 can be thought of as interchangeable.

For testing the alignment error for varying the amount and positioning, the ArUco markers to be used for validation are those with varying sizes (i.e. the markers with sizes 7.14 cm, 10.00 cm, 16.67 cm, and 50.00 cm).

2.3.2 Size

To study the effect of varying the size of the ArUco markers, four different sizes of the markers were created and placed in a checkerboard like pattern against a plane. To test the effect, the ArUco markers with similar sizes are filtered and used for the alignment procedure, with at most two markers from each size used for each case. With four different sizes of the markers to be used in this study, there will be a maximum number of eight iterations for this portion of the study. It should be noted that there is also a possibility that the set conditions of point cloud alignment might prevent ArUco markers of small sizes from being detected and therefore not used for alignment.

Table 3. Number of markers for each iteration

Identifier	Plane 1 Markers	Plane 2 Markers
1P10P2	1	0
1P11P2	1	1
2P10P2	2	0
2P11P2	2	1
2P12P2	2	2
3P10P2	3	0
3P11P2	3	1
3P12P2	3	2
3P13P2	3	3

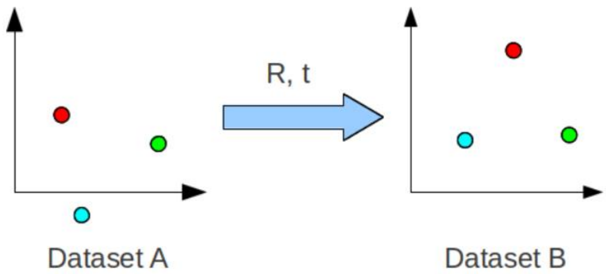

Fig. 5. Rotation and translation of datasets [13]

For testing the alignment error of varying the sizes of the ArUco markers, the markers of other sizes are grouped together and the average error within each group is found and treated as the errors for varying the size of the markers.

2.4 Alignment Procedure

The alignment procedure is the sequence of steps necessary to determine the rotation and translation matrix for aligning the two point clouds. This procedure is described accurately by Ho [13] and comprises a series of steps, including matching of identified ArUco markers, mapping of ArUco corners to depth map coordinate, determining the centroid of the points, finding the rotation matrix via singular value decomposition, and lastly finding the translation matrix by vector subtraction. An example of alignment can be seen in figure 5.

Firstly, the ArUco markers from the frames of interest are detected and their respective corners mapped in the color image via pixel coordinates. The index list of the markers is then cross-checked between the two frames of interest to determine which markers appear in both frames. The markers for the alignment procedure are then filtered based on the criteria of the user, and these markers are saved along with their respective pixel coordinate. The pixel coordinates are then mapped to the respective aligned depth maps and the 3D coordinates are found.

With the filtered points, the next step is to determine the centroid of each coordinate points for each of the two frames. This is done by finding the numerical average of the abscissas, ordinates, and applicates. By subtracting each coordinate point from the coordinate of the centroid, it is possible to find the coordinates of each point in the set in reference to the centroid, allowing the centroid to be the origin reference for that frame. This is done to the second frame as well.

For finding the rotation matrix, the technique of singular value decomposition is to be used. For this technique, the two new sets of coordinates are first multiplied together, and then the numpy method of singular value decomposition [14] is used on the resulting matrix. This results in two rotation matrixes, and one scalar matrix, however the scalar matrix can be ignored as we are assuming rigidity of the points of interest. The two rotation matrixes are then multiplied to each other, resulting in the overall rotation matrix of interest.

Lastly, the centroid of the non-reference set of points is then transformed using the found rotation matrix and then subtracted from the centroid of the reference set of points to find the translation matrix. In this method, the necessary rotation and translation matrixes are found and can be applied to all ArUco corner points found, including those for evaluating the alignment.

2.5 Alignment Evaluation

For evaluation of the alignment, only corner points of ArUco markers not used during the alignment procedure are to be used. The markers of interest are then grouped according to size and the error measured per group. To determine the error, the root mean square of the aligned values vs the reference values were taken, as seen in the following equation,

$$Error = \sqrt{\frac{1}{N}\sum(x_{aligned} - x_{reference})^2}$$

This is because the difference of aligned points and reference points could have negative values, which in turn would offset positive values. However, finding the square of the value will negate the negative value and provide a more accurate estimate of the average error of the alignment evaluation. Also, finding the error value through the root mean square method will place a greater emphasis on larger differences. These are necessary to spot as they are unfavorable aspects for alignment.

3. RESULTS AND DISCUSSION

3.1 Set-up

The set-up of the experiment can be seen in figure 6, with the cameras away from the midpoints of the ArUco marker axis of symmetry by 2.3 meters. Six ArUco markers of length 30 centimeters were grouped into three and each group was placed co-planar to each other. A checkerboard ArUco design was placed on each respective plane, and had markers of varying sizes (50 centimeters, 16.67 centimeters, 10 centimeters and 7.14 centimeters).

Fig. 6. ArUco marker set-up

Table 4. Errors for ArUco markers with size 50 cm

Index	No.	Smallest	Largest	Average
1P10P2	1	40.97 mm	401.37 mm	167.96 mm
1P11P2	2	63.50 mm	230.57 mm	124.04 mm
2P10P2	2	93.00 mm	260.77 mm	151.02 mm
2P11P2	3	93.23 mm	188.88 mm	137.36 mm
2P12P2	3	106.45 mm	257.56 mm	150.37 mm
3P10P2	4	81.16 mm	198.18 mm	159.49 mm
3P11P2	4	89.39 mm	186.25 mm	135.45 mm
3P12P2	5	101.58 mm	192.19 mm	155.44 mm
3P13P2	6	107.66 mm	196.36 mm	149.06 mm

Table 5. Errors for ArUco markers with size 16.67 cm

Index	No.	Smallest	Largest	Average
1P10P2	1	64.68 mm	385.81 mm	195.92 mm
1P11P2	2	57.25 mm	206.71 mm	117.34 mm
2P10P2	2	99.10 mm	251.30 mm	151.78 mm
2P11P2	3	81.04 mm	136.26 mm	185.04 mm
2P12P2	3	81.70mm	263.46 mm	159.27 mm
3P10P2	4	110.34 mm	197.51 mm	155.98 mm
3P11P2	4	80.66 mm	181.20 mm	132.47 mm
3P12P2	5	110.64 mm	194.47 mm	152.18mm
3P13P2	6	100.34 mm	185.58 mm	144.55 mm

Table 6. Errors for ArUco markers with size 10.00 cm

Index	No.	Smallest	Largest	Average
1P10P2	1	31.74 mm	293.50 mm	118.91 mm
1P11P2	2	42.02 mm	174.21 mm	91.08 mm
2P10P2	2	52.49 mm	100.26 mm	189.56 mm
2P11P2	3	53.25 mm	142.38 mm	94.80 mm
2P12P2	3	47.12 mm	173.03 mm	100.00 mm
3P10P2	4	63.17 mm	109.22 mm	157.52 mm
3P11P2	4	55.20mm	140.16 mm	93.45 mm
3P12P2	5	64.27mm	151.56 mm	106.11 mm
3P13P2	6	68.59 mm	141.41 mm	102.71 mm

3.2 ArUco amount and position variation

Tables 4 to 6 show the smallest, highest, and average errors for different sizes of the ArUco markers. It should be noted that the ArUco marker size of 7.14 centimeter was not included as the markers were not detected at the set range (which was close to the maximum range).

The graph of varying the number of ArUco markers can be seen in figures 7 to 9, where each plot shows the smallest, largest, and average error for each ArUco marker size versus how many ArUco markers were used for alignment.

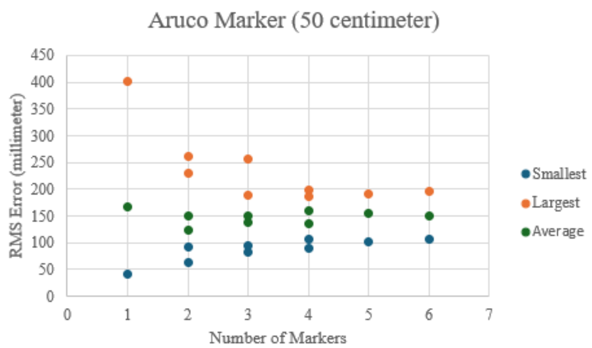

Fig. 7. Scatter plot for RMS error for 50 cm marker

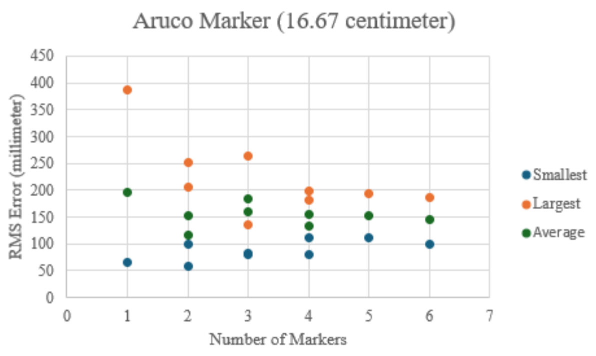

Fig. 8. Scatter plot for RMS error for 16.67 cm marker

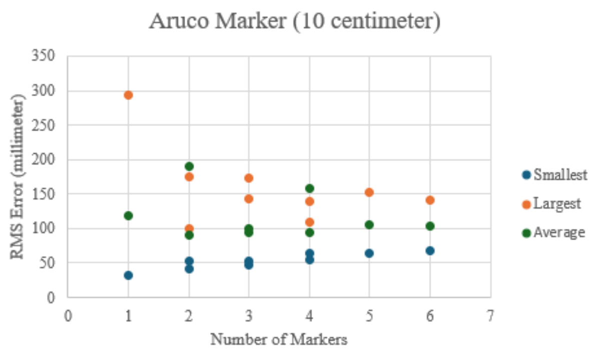

Fig. 9. Scatter plot for RMS error for 10 cm marker

Based on the scatter plots solely on the number of markers, the smallest and largest errors have a greater range for less than 3 markers, and this can be clearly seen when only one marker is used. When using 4 markers or more, there does not appear to be a greater change in errors when using more markers. This shows that the possible optimal number of markers for the study is 4 markers as more markers will not lead to a decrease in alignment error.
For checking the effect of the position of the ArUco markers, the errors were grouped according to whether one or two planes were used for the ArUco alignment. Figures 10 to 12 show the scatter plot results of grouping the errors together.
Based on the scatter plots for the number of planes, the grouping of the errors can be seen closer when two planes are used for alignment as opposed to only using one plane. The plots may show that the error can be smaller for only one plane, however the average error and largest errors are more stable for when more than one plane is used.

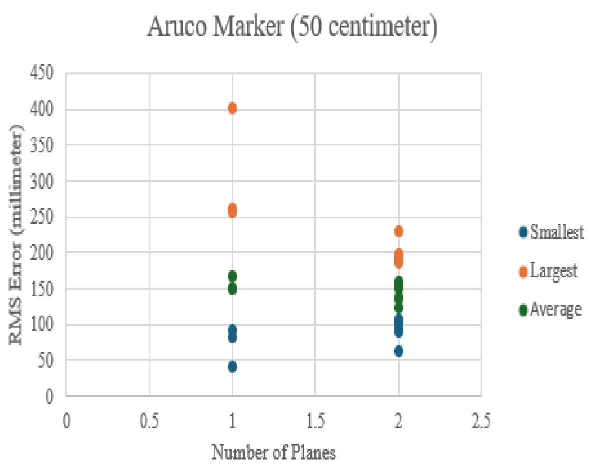

Fig. 10. Scatter plot for planes for 50 cm marker

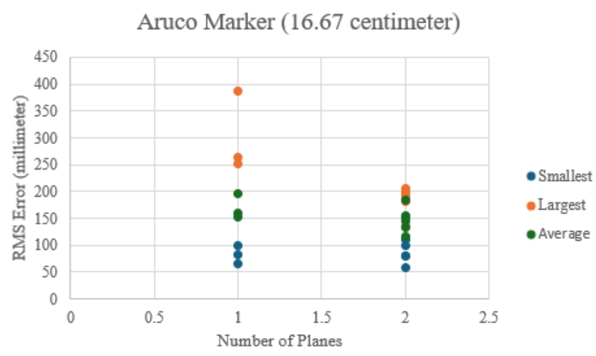

Fig. 10. Scatter plot for planes for 16.67 cm marker

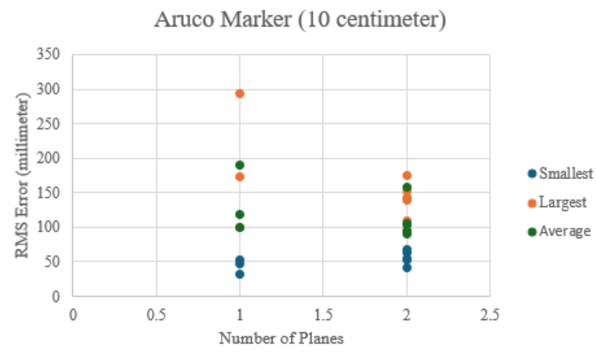

Fig. 11. Scatter plot for planes for 10 cm marker

3.3 ArUco marker size variation

The ArUco marker sizes for alignment were varied, and the RMS error for was calculated for the other ArUco markers not used for alignment. The data was grouped together according to the RMS error of an ArUco marker in questions. As can be seen in figures 12 to 14, the x-axis of the plot was labelled as the size of the ArUco marker used for alignment, and how many planes were used in the alignment. It should also be noted that the 7.14 cm markers were once again not used as they were not detected by the algorithm, and therefore could not be used for alignment nor for validation. The data for the RMS errors can also be seen in tables 7 to 9.

Table 7. Errors for markers aligned with 50.00 cm

Size	Pl	Smallest	Largest	Average
16.67 cm	1	69.06 mm	250.67 mm	180.28 mm
16.67 cm	2	58.42 mm	89.96 mm	70.43 mm
10.00 cm	1	59.97 mm	338.19 mm	167.44 mm
10.00 cm	2	16.86 mm	85.52 mm	60.14 mm

Table 8. Errors for markers aligned with 16.67 cm

Size	Pl	Smallest	Largest	Average
50.00 cm	1	76.51mm	636.17 mm	356.09 mm
50.00 cm	2	59.66 mm	223.95 mm	110.77 mm
10.00 cm	1	70.32mm	715.21 mm	362.81 mm
10.00 cm	2	37.80mm	260.59 mm	104.80mm

Table 9. Errors for markers aligned with 10.00 cm

Size	Pl	Smallest	Largest	Average
50.00 cm	1	78.23 mm	1631.69 mm	534.45 mm
50.00 cm	2	44.78 mm	1483.61 mm	365.50 mm
16.67 cm	1	103.78 mm	1462.75 mm	538.64 mm

16.67 cm	2	58.91 mm	1328.39 mm	379.28 mm

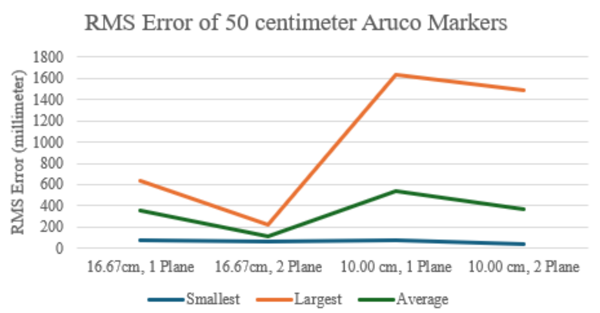

Fig. 12. Plot for RMS error of 50 cm ArUco markers

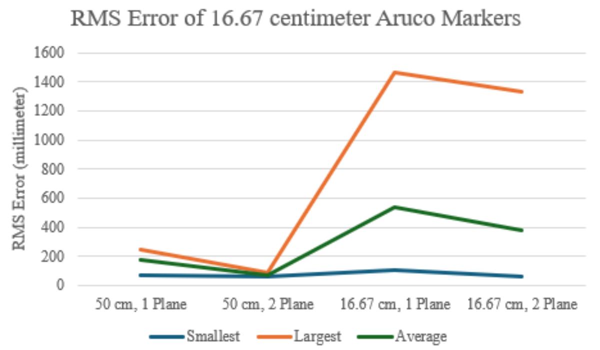

Fig. 13. Plot for RMS error of 16.67 cm ArUco markers

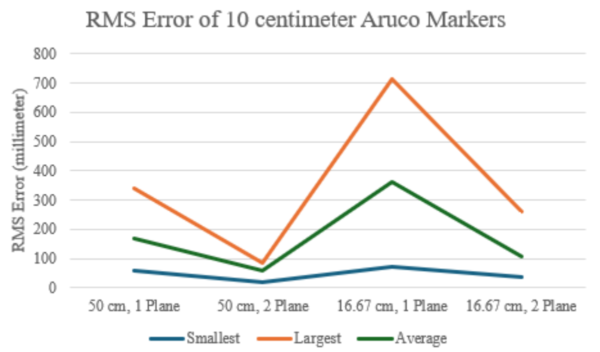

Fig. 14. Plot for RMS error of 10 cm ArUco markers

As can be seen from the data points and plots, the use of smaller ArUco markers for alignment reduces the accuracy significantly. Therefore, it is recommended to use larger ArUco markers for alignment. It can also be seen that using two planes instead of one will output more consistent results for alignment, which was also demonstrated in the previous section.

4. CONCLUSION AND RECOMMENDATIONS

The data from the study suggests that the optimal number of ArUco markers is 4 markers for when using the Intel RealSense D455 camera. Using fewer markers could lead to a wider range of errors, while using more markers would not provide a significant increase in the stability of the errors. The use of bigger markers is also recommended as shown in the latter half of the results as big markers will have a more stable error range. Both sections of the results also showed that using more than one plane will result in a narrower error range which would be ideal for real-time uses of alignment.

For future studies, it is recommended to study more than one set-up of the ArUco markers, and with variables conditions as well such as low-light or bright light along with a lux-meter. Another physical quality that can be altered is the material of the ArUco marker as this could potentially alter the depth map and potentially dampen or, conversely, increase the noise. One more potential avenue worth exploring is using three planes instead of two such as a corner and determining the error of the planes relative to one another. Lastly, a different camera can also be used for further studies to cover more recent advancements in RGB-D camera technology.

5. REFERENCES

[1] T. Patterson, S. McClean, P. Morrow, G. Parr, C. Luo, Timely autonomous identification of UAV safe landing zones, Image and Vision Computing 32 (2014) 568–578.

[2] W. Fdzil, M. Mazlan, F. Hanapiah, A. Abdullah, Development of 3D Printed Socket for Transtibial Prosthetic Leg, Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES). 5 (2019) 44-46.

[3] A. Abdullah, F. Alimarizan, N. Aziz, F. Hazwani, Development of Ophthalmology Eye Examination Device Using 3D Printing Technology, Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES). 10 (2024) 457-462.

[4] J. Park, Q. Zhou, V. Koltun, Colored Point Cloud Registration Revisited, 2017 IEEE International Conference on Computer Vision (ICCV) (2017) 143-152

[5] ArUco Markers: https://docs.opencv.org/4.x/d5/dae/tutorial_ArUco_detection.html (accessed 20.05.05)

[6] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas, M.J. Marín-Jiménez, Automatic generation and detection of highly reliable fiducial markers under occlusion, Pattern Recognition 47 (2014) 2280-2292

[7] F. Romero-Ramirez, R. Munoz-Salinas, R. Medina-Carnicer, Speeded Up Detection of Squared Fiducial Markers, Image and Vision Computing 76 (2018) 38-47

[8] G. Cepon, D. Ocepek, M. Kodric, M. Demsar, T. Bregar, M. Boltezar, Impact-Pose Estimation Using ArUco Markers in Structural Dynamics, Experimental Techniques 48 (2024) 369-380

[9] A. Aalerud, J. Dybedal, G. Hovland, Automatic Calibration of an Industrial RGB-D Camera Network Using Retroreflective Fiducial Markers, Sensors 7 (2019) 1561

[10] T. Madeira, M. Oliveira, P. Dias, Enhancement of RGB-D Image Alignment Using Fiducial Markers, Sensors 5 (2020) 1497

[11] Intel Realsense D455 Datasheet: https://www.intelrealsense.com/download/21345/?tmstv=1697035582

[12] Realsense Python Library: https://pypi.org/project/pyrealsense2/

[13] Finding optimal rotation and translation between corresponding 3D points: https://nghiaho.com/?page_id=671

[14] Numpy numpy.linalg.svd: https://numpy.org/doc/2.2/reference/generated/numpy.linalg.svd.html