The Applications of Fiducial Markers in Robotic Grasping: A Comparison Between AprilTag and ArUco Markers

Junlin Wang[1], Dong Li[2]

***Abstract*—This paper investigates the possible applications of two fiducial markers (AprilTag and ArUco) on hand eye calibration and localization procedures in a robotic grasping task and compares their performance on each procedure under different lighting conditions. Firstly, an experimental platform is established based on JAKA MiniCobo manipulator, DH-Robotics PGE-50-26 gripper and Intel RealSense D455 camera. Next, a hand eye calibration experiment and a grasping experiment are designed using AprilTag and ArUco markers for pose estimation and localization, respectively. Then, the two experiments are conducted under normal light condition and dim condition successively. Finally, the precision of hand eye calibration and the success rate of grasping are evaluated. The experimental result indicates that ArUco marker possesses higher precision in hand eye calibration but a lower success rate in grasping than AprilTag marker. Besides, ArUco marker shows low resistance to the interference of lighting condition in both procedures, while AprilTag marker possesses high robustness in grasping process but low resistance in hand eye calibration process.**

***Index Terms*—Fiducial marker, Robotic grasping, Hand eye calibration, Localization, AprilTag, ArUco**

I. INTRODUCTION

HAND eye calibration and localization are two important procedures in robotic grasping. The former acts as a builder establishing the bond between the visual sensor and the manipulator. The latter serves as a guide leading the gripper to grab the targets. In both procedures, it is essential to obtain the pose information of various items. Generally, the acquisition of pose information is through the depth image of the item. Nevertheless, processing depth images is often troublesome and time-consuming and often fails in poor light conditions. In such cases, fiducial markers become another option for obtaining pose data.

A fiducial marker is a label with a highly distinguishable pattern that can be detected by an external visual sensor even under a poor light condition. The pose information of the marker is stored in the pattern and can be extracted from the marker's RGB image using specific algorithms. Given the fact that the pose information is easily accessed under different light conditions, the fiducial marker is a possible substitute for some conventional pose estimation methods in robotic grasping.

There are many different fiducial markers available now. Each of them has some unique features that distinguish itself from others. The markers used in this work are AprilTag [1], [2] and ArUco [3] markers. They are both mature enough and widely utilized in manifold robotic missions for pose estimation and localization. This work will focus on the applications of AprilTag and ArUco markers in the hand eye calibration process as well as the localization process and compare the performance of each marker under normal light condition and dim condition.

This paper is organized as follows: Section II offers an overview of the fiducial marker and explores its applications in robotic grasping. Section III describes the experimental setup in detail and outlines the procedures of different experiments. Section IV presents and analyzes the experimental results. Section V summarizes the work and draws a conclusion.

II. OVERVIEW

A. Fiducial Markers

Fiducial markers are the labels carrying special patterns used for identification and pose estimation. Most fiducial markers are monochromatic, with some binary message encoded in the patterns that can be interpreted by particular algorithms. In a single marker family, every marker has its own ID which determines the outline of the pattern. Markers with different ID numbers can be distinguished by an external visual sensor due to the difference in the stored binary message. Besides identification, the pattern of a marker also provides the marker's pose information, which can then be used in navigation, pose estimation and localization.

Among all the fiducial markers, ARTag [4], [5], AprilTag [1], [2] and ArUco [3] are commonly used in robotics. ARTag is based on ARToolkit, but uses a digital coding system to generate patterns. AprilTag inherits the framework from ARTag and makes some improvements to the image processing algorithm. ArUco is created on the basis of ARToolkit and ARTag, and allows users to generate their personally customized libraries according to their needs [6]. All these three markers have their custom ROS packages that help users to process obtained data. However, the ROS package of ARTag has not been updated for many years. Given this drawback, ARTag is not used in this work.

B. Applications

Fiducial markers are widely used in robotics, especially in hand eye calibration and localization.

1) Hand Eye Calibration: Hand eye calibration is an essential process in robotics. The purpose of hand eye calibration is to build up the relationship between the camera coordinate system and the manipulator coordinate system. Generally, the relationship is expressed by a transformation matrix mapping from the camera frame to either the manipulator's end effector frame or the manipulator's base link frame, depending on whether it is eye-in-hand or eye-to-hand. In an eye-in-hand situation, the camera is mounted on the manipulator's end effector, while in an eye-to-hand situation, the camera is mounted on an external fixed stand. No matter what situation it is, the calibration process can be simplified into solving the equation AX=XB [7], where X is either the transformation matrix from end effector to camera (1) or the transformation matrix from base link to camera (2). To solve this equation, it is necessary to acquire A and B. A is the product of two transformation matrices between base link and end effector, while B is the product of two transformation matrices between camera and an external target. A classic method to obtain the transformation matrix from target to camera is using a checkerboard target. In recent years, a new option of using a fiducial marker target became available in many robotic grasping tasks. [8] developed an online calibration method based on ChArUco board that has higher flexibility, robustness and accuracy than the conventional works. [9] adopted the hand eye calibration method using ArUco markers in a support vector machine (SVM) based robotic grasping assignment and achieved a relatively low position error. [10] performed the hand eye calibration based on AprilTag markers in a robot-assisted surgery project and fulfilled a good outcome. Compared with the traditional checkerboard method, calibration with the use of fiducial markers achieves higher flexibility, robustness and accuracy. Moreover, when the occlusion occurs, a fiducial marker shows higher resistance to the interference than the checkerboard [8]. Due to these strengths, fiducial markers are now widely used for hand eye calibration in robotic grasping.

$$ {}^{base}_{end}T_j^{-1} * {}^{base}_{end}T_i * {}^{end}_{cam}T = {}^{end}_{cam}T * {}^{cam}_{target}T_j * {}^{cam}_{target}T_i^{-1} \quad (1) $$

$$ {}^{end}_{base}T_j^{-1} * {}^{end}_{base}T_i * {}^{base}_{cam}T = {}^{base}_{cam}T * {}^{cam}_{target}T_j * {}^{cam}_{target}T_i^{-1} \quad (2) $$

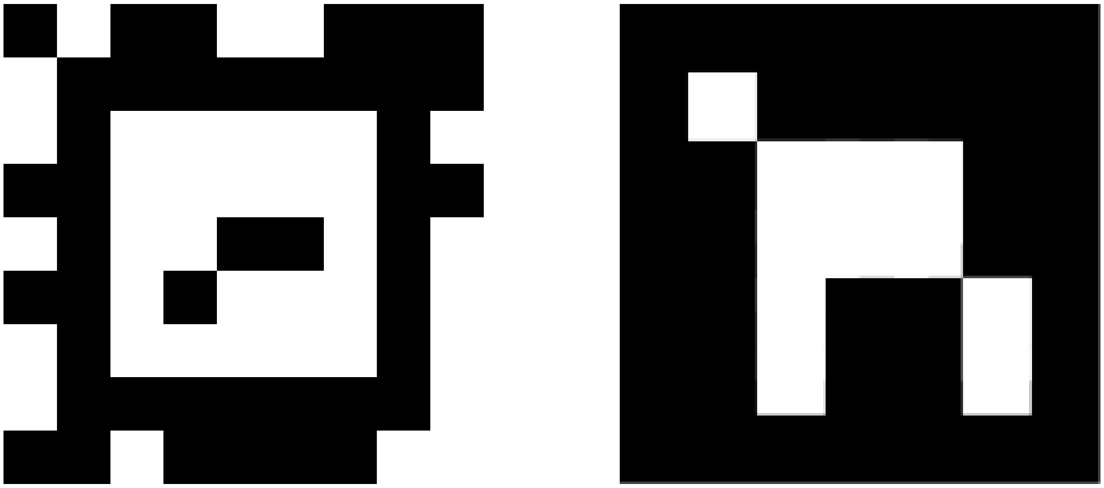
2) Localization: Before the appearance of fiducial markers, the location of the target object in a robotic grasping task was usually obtained by processing the depth images. Although this approach has a high localization accuracy in surroundings with strong visual features, it often fails in environments where there are few visual features or little light. Fortunately, the application of fiducial markers partly solves this problem. Due to their highly discernible patterns encoded with binary information, fiducial markers can be identified by visual sensors under different light circumstances. This provides great robustness and flexibility and allows the fiducial markers to be widely used in multifarious robotic missions. In the past few years, many researches of this approach have been done. In [11], the researchers make a comparison among the four freely available libraries for AprilTag detection and propose two novel techniques for improving the localization accuracy of AprilTag. In [12], ArUco markers are used for locating the BROOK camera in a robotic pipe-cutting assignment. In this work, fiducial markers are pasted on the target object in order to help the visual sensor localize its position, which will be further discussed in the next section.

III. EXPERIMENTAL SETUP

As is mentioned above, fiducial markers are widely used in robotic grasping projects. Nevertheless, not all the fiducial markers are suitable for performing every single task. Each fiducial marker system has a set of criteria that influence its performance in certain application areas [13]. A marker system may perform well in one particular area but inversely react terribly in other areas. Apart from this, the performance also involves with multiplex external factors such as distance, orientation, lighting condition and motion blur [6]. Two markers are likely to perform distinctly even if they have the same systems due to the difference in the surroundings. Hence, it is essential to investigate the performance of different fiducial markers under different circumstances.

In this work, we choose two commonly used markers, AprilTag [1], [2] and ArUco [3], as our research targets (Fig. 1). The two markers are both monochromatic, encoded with some useful information in the patterns. There are two main differences between them: (i) All the markers in ArUco library are square while AprilTag library contains both square markers and circular markers, (ii) ArUco markers are all bounded by black while AprilTag markers are sometimes black and white interlaced in the margin (Fig. 1). Due to the differences in the patterns, the valid regions for encoding information can differentiate in size for AprilTag and ArUco markers. To reduce this differentiation, markers used in this work are all square and of the same outer length.

To investigate the performance of ArUco and AprilTag markers on hand eye calibration and localization under different lighting conditions, an experimental platform for performing hand eye calibration and grasping is constructed (Fig. 2). A camera is mounted on a stand with adjustable height and orientation. A manipulator is fixed on the table, with a gripper installed at the end as the end effector. The camera used in this experiment is the Intel RealSense D455. It can capture both RGB images and depth images with certain pixels at different FPS (Frames Per Second). In a sense, using normal RGB camera is also feasible in this experiment since we only need the RGB images of fiducial markers. The strength of using Intel RealSense D455 is that the camera's intrinsics can be directly obtained from the published ROS topic. In this case, the procedure of camera calibration can be omitted in this work. As for the manipulator and the gripper, we use Chinese JAKA MiniCobo and DH-Robotics PGE-50-26. JAKA MiniCobo has 6 DOF (Degree OF Freedom) and is controlled by a customized IPC (Industrial Personal Computer). DH-Robotics PGE-50-26 is a parallel electric gripper with a maximum finger width of 2.5 cm and is powered by the IPC of JAKA MiniCobo



(a) (b)

Fig. 1. The fiducial markers used in this experiment (a) AprilTag (ID: 250). (b) ArUco (ID: 250).

in this experiment. Apart from the IPC of JAKA MiniCobo, we use another IPC, AAEON EPIC-KBS9, which is configured with Ubuntu 20.04. During the experiment, the AAEON EPIC-KBS9 works as the upper computer and directly controls the gripper and the IPC of JAKA MiniCobo. The IPC of JAKA MiniCobo, which serves as the lower computer, then controls the manipulator to move about.

A. *Hand Eye Calibration Experiments*

To investigate the precision of hand eye calibraion, we first print an AprilTag marker and an ArUco marker on the paper, respectively. Each marker has a size of 4 × 4 cm and an ID number of 250. The marker is then grabbed by the gripper and placed in the field of view of the camera (Fig. 3). Every time, the manipulator is controlled to move around while the camera is capturing RGB images with 1280 × 800 pixels and 30 FPS. One way to acquire the image data is subscribing to the ROS topic that is enabled by Intel RealSense D455. In this way, we can easily obtain the real-time image data published by Intel RealSense D455. To figure out the relationship between the camera and the manipulator, we need several samples containing the two transformation matrices: marker-to-camera matrix and end-to-base matrix. To obtain the marker-to-camera matrix, we use the ROS package of AprilTag and ArUco to process image data. Nodes in each package can subscribe to the camera image topic, decode the obtained image data, and publish the marker's pose w.r.t the camera frame that can then be used to calculate the marker-to-camera matrix. As for the end-to-base matrix, it can be directly calculated from the end's pose data published by JAKA MiniCobo. In every single experiment, a total of 17 different pairs of transformation matrix samples are recorded and then used to calculate the camera-to-base matrix by OpenCV Python. For each marker, this procedure repeats 10 times under normal light condition and another 10 times under dim condition (Fig. 4). After that, the mean and variance of the calibration results for every 10 repetitions are evaluated.

Fig. 2. The experimental platform made up of Intel RealSense D455 camera, JAKA MiniCobo manipulator and DH-Robotics PGE-50-26 gripper.

B. *Grasping Experiments*

In this experiment, we use a small object with a marker pasting on it (Fig. 5) as the target object. Both the AprilTag marker and the ArUco marker have a size of 4 × 4 cm and an ID number of 250. The object is then placed on the table

Fig. 3. The marker grabbed by the gripper.

where the manipulator can reach. When everything is set up, the manipulator moves to grasp the object according to the location obtained from the fiducial marker (Fig. 6). To fulfill this grasping task, A ROS package is developed by ourselves to determine the actual position of the object based on the marker's pose data and the camera-to-base transformation matrix. For each marker, the grasping practice repeats 100 times under normal light condition and another 100 times under dim condition, during which both the successful grasp and the failed grasp are counted. When the practice is finished, the success rate for each marker under each lighting condition is calculated, respectively.

C. Package Configurations

In the previous subsections, some ROS packages are mentioned. All of them are indispensable in this experiment since they are used to process a diversity of data. Besides those packages, two other ROS packages must be included in this experiment. One is the package that controls the motions of the manipulator. The other is the package controlling the gripper to open or close. All the required packages are listed below, where you can directly download them from source.

1) https://github.com/IntelRealSense/realsense-ros
2) https://github.com/AprilRobotics/apriltag_ros
3) https://github.com/pal-robotics/aruco_ros
4) https://github.com/JAKARobotics/JAKA_ROS_Driver
5) https://github.com/DH-Robotics/dh_gripper_ros
6) https://github.com/HenryWJL/jaka_grasping

IV. RESULT

In this section, we present and analyze the results of the hand eye calibration experiment and the grasping experiment. All the experimental results are listed in Tables and marked in different colors. In Table I and II, the greatest value in a column is marked in magenta while the least is marked in cyan. In Table III and IV, marker with the highest success rate is labelled in red.

A. Hand Eye Calibration Experiment

In this experiment, precision is evaluated by the magnitude of variance. Lower variance means higher precision and vice versa. According to Table II, ArUco marker shows lower variance than AprilTag marker in all seven components under the same lighting condition, indicating that ArUco possesses higher precision in hand eye calibration than AprilTag. Besides this observation, both ArUco and AprilTag face a variance growth when the lighting condition changes from normal to dim. This reveals the negative influence of poor lighting condition on the precision of hand eye calibration. It is also noticeable that ArUco marker under normal light condition acquires the least variance in nearly all the components, which represents the highest precision in this experiment.

B. Grasping Experiment

Table III and IV display the results of the robotic grasping experiment. Under the normal light condition, AprilTag marker possesses a success rate of up to 93%, which is much higher than the success rate of ArUco marker. This indicates that AprilTag marker has higher localization accuracy than ArUco marker under the same lighting condition. When the lighting condition alters from normal to dim, the success rate of AprilTag marker remains unchanged, indicating the marker's strong resistance to the interference of lighting condition. However, the success rate of ArUco marker drops from 67% to 55%, which shows that ArUco marker has not only low localization accuracy but also poor resistance to the interference of lighting condition.

TABLE I
THE MEAN OF CALIBRATION RESULTS

		Camera-to-Base Matrix denoted as Quaternions						
		Translation (m)			Rotation (m)			
Light Condition	Marker	x	y	z	w	x	y	z
Normal	AprilTag	-0.571	-0.288	-0.083	0.541	-0.813	0.003	-0.197
	ArUco	-0.278	-0.146	0.0228	0.548	-0.771	0.0437	-0.138
Dim	AprilTag	-0.400	-0.045	-0.040	0.573	0.753	0.057	0.169
	ArUco	-0.512	0.049	0.080	0.419	0.892	-0.075	0.099

TABLE II
THE VARIANCE OF CALIBRATION RESULTS

		Camera-to-Base Matrix Denoted as Quaternions						
		Translation (m)			Rotation (m)			
Metric	Marker	x	y	z	w	x	y	z
Normal	AprilTag	0.033	0.005	0.004	0.028	0.044	0.008	0.005
	ArUco	0.010	0.002	0.004	0.003	0.001	0.002	0.001
Dim	AprilTag	0.043	0.009	0.010	0.026	0.029	0.013	0.004
	ArUco	0.016	0.005	0.002	0.006	0.002	0.002	0.003

V. CONCLUSION

In this paper, we first offer an overview of fiducial markers, in which we summarize the common features of fiducial markers and outline several presently available packages. Based on these features, we discuss the possible applications of fiducial

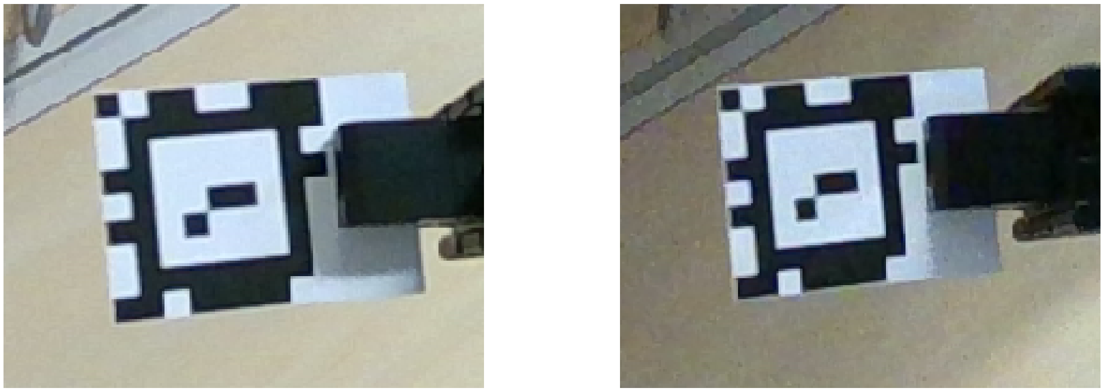

(a)
(b)

Fig. 4. The marker used (a) under normal light condition. (b) under dim condition.

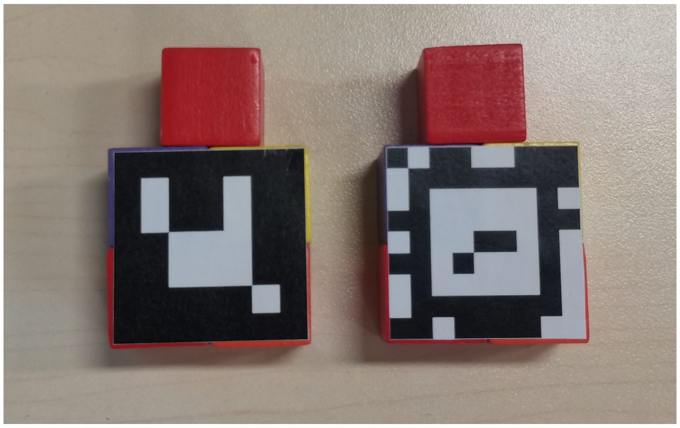


Fig. 5. The objects labelled with fiducial markers.

TABLE III
GRASPING RESULTS UNDER NORMAL LIGHT CONDITION

Marker	Success	Failure	Total	Success Rate
AprilTag	93	7	100	93%
ArUco	67	33	100	67%

markers in different areas of robotic grasping. After that, we establish an experimental platform for conducting hand eye calibration and grasping tasks under two different lighting conditions to compare the performance of AprilTag and ArUco markers. Finally, the precision of hand eye calibration and the success rate of grasping are presented and analyzed.

TABLE IV
GRASPING RESULTS UNDER DIM CONDITION

Marker	Success	Failure	Total	Success Rate
AprilTag	93	7	100	93%
ArUco	55	45	100	55%

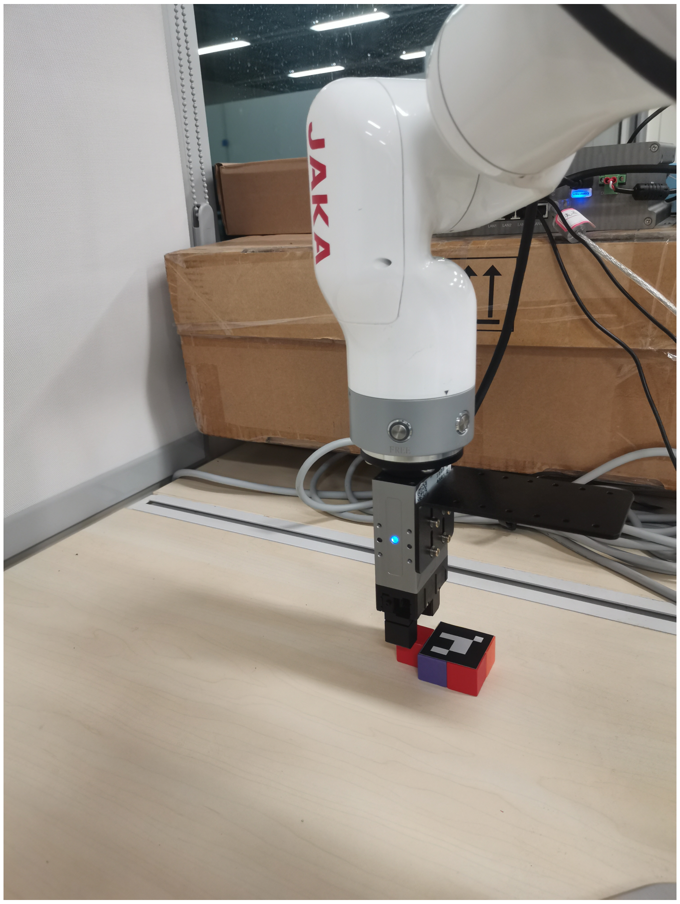


Fig. 6. The scene where the manipulator moves to grasp the target object.

Through this experiment, we obtain two observations. Firstly, ArUco marker has high precision but low resistance to poor lighting condition in hand eye calibration process. Secondly, AprilTag marker has a success rate of 93% in grasping process, indicating the marker's high accuracy of localization. Additionally, it also shows high accuracy and robustness under poor lighting condition. To summarize, ArUco marker shows a better performance in hand eye calibration process while AprilTag marker does much better in grasping process.

REFERENCES

[1] E. Olson, "AprilTag: A robust and flexible visual fiducial system," 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, 2011, pp. 3400-3407, doi: 10.1109/ICRA.2011.5979561.

[2] J. Wang and E. Olson, "AprilTag 2: Efficient and robust fiducial detection," 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Korea (South), 2016, pp. 4193-4198, doi: 10.1109/IROS.2016.7759617.

[3] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas and M.J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," Pattern Recognition, vol. 47, no. 6, pp. 2280-2292, Jun. 2014.

[4] M. Fiala, "ARTag, a fiducial marker system using digital techniques," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 590-596 vol. 2, doi: 10.1109/CVPR.2005.74.

[5] M. Fiala, "Designing Highly Reliable Fiducial Markers," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 7, pp. 1317-1324, July 2010, doi: 10.1109/TPAMI.2009.146.

[6] M. Kalaitzakis, B. Cain, S. Carroll and et al, "Fiducial Markers for Pose Estimation," *J Intell Robot Syst 101*, vol. 71, Mar. 2021.

[7] Y. C. Shiu and S. Ahmad, "Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form AX=XB," in IEEE Transactions on Robotics and Automation, vol. 5, no. 1, pp. 16-29, Feb. 1989, doi: 10.1109/70.88014.

[8] W. Lin, P. Liang, G. Luo, Z. Zhao, and C. Zhang, "Research of Online Hand–Eye Calibration Method Based on ChArUco Board," Sensors, vol. 22, no. 10, p. 3805, May 2022, doi: 10.3390/s22103805.

[9] Y. Huang, J. Zhang and X. Zhang, "Robotic grasping based on machine vision and SVM," 2019 IEEE 9th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), Suzhou, China, 2019, pp. 1027-1030, doi: 10.1109/CYBER46603.2019.9066460.

[10] T. -H. Ho and K. -T. Song, "Supervised Control for Robot-Assisted Surgery Using Augmented Reality," 2020 20th International Conference on Control, Automation and Systems (ICCAS), Busan, Korea (South), 2020, pp. 329-334, doi: 10.23919/ICCAS50221.2020.9268278.

[11] J. Kallwies, B. Forkel and H. -J. Wuensche, "Determining and Improving the Localization Accuracy of AprilTag Detection," 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 2020, pp. 8288-8294, doi: 10.1109/ICRA40945.2020.9197427.

[12] T. Burrell, C. West, S. D. Monk, A. Montezeri and C. J. Taylor, "Towards a Cooperative Robotic System for Autonomous Pipe Cutting in Nuclear Decommissioning," 2018 UKACC 12th International Conference on Control (CONTROL), Sheffield, UK, 2018, pp. 283-288, doi: 10.1109/CONTROL.2018.8516841.

[13] K. Shabalina, A. Sagitov, M. Svinin and E. Magid, "Comparing Fiducial Markers Performance for a Task of a Humanoid Robot Self-calibration of Manipulators: A Pilot Experimental Study," 2018 Interactive Collaborative Robotics (ICR), Leipzig, Germany, 2018, vol. 11097, pp. 259-268.