Virtual Experiments on ArUco and AprilTag Systems Comparison for Fiducial Marker Rotation Resistance under Noisy Sensory Data

1st Aufar Zakiev
Intelligent Robotics Department
Kazan Federal University
Kazan, Russian Federation
zaufar@it.kfu.ru

2nd Tatyana Tsoy
Intelligent Robotics Department
Kazan Federal University
Kazan, Russian Federation
tt@it.kfu.ru

3rd Ksenia Shabalina
Intelligent Robotics Department
Kazan Federal University
Kazan, Russian Federation
ks.shabalina@it.kfu.ru

4th Evgeni Magid
Intelligent Robotics Department
Kazan Federal University
Kazan, Russian Federation
magid@it.kfu.ru

5th Subir Kumar Saha
Department of Mechanical Engineering
Indian Institute of Technology Delhi
Delhi, India
saha@mech.iitd.ac.in

***Abstract*—Modern robotic researches propose various machine vision methods for accomplishing robotic tasks. The recognition quality in these tasks is very important for successful performance. A large number of them use fiducial marker systems as a main element of algorithms. However, only a few researches are comparing standard marker systems. This paper is dedicated to the comparison of AprilTag and ArUco markers resistance to rotations in the presence of synthetic noise. Experiments were conducted in ROS/Gazebo virtual environment in order to provide a fair comparison of marker detection and recognition algorithms while eliminating external environment conditions that influence the algorithms' performance. The presented virtual environments allow collecting a significant amount of data by experiment process automation. Different levels of additive white Gaussian noise were applied to input sensory data in order to simulate the imperfection of real digital cameras. The main contribution of the paper is the systematic comparison of AprilTag and ArUco markers for rotation resistance in the presence of optical sensor noise.**

***Index Terms*—robotics, fiducial marker system, recognition algorithms, experimental comparison, ROS, Gazebo**

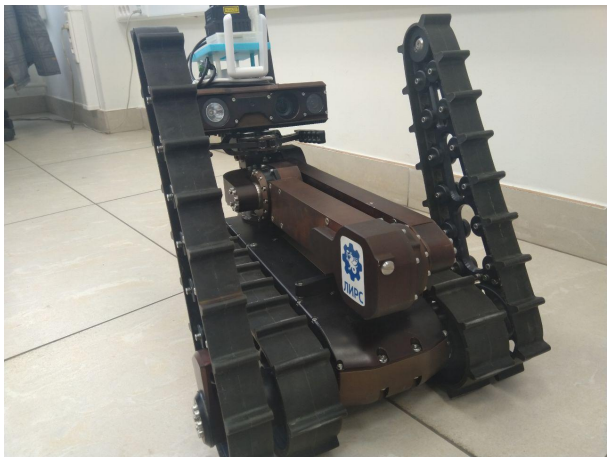

Fig. 1. Servosila Engineer robot.

I. Introduction

Fiducial marker system (FMS) is a set of graphical tokens developed for recognition by machine vision systems. These graphical signs are mostly designed to be printed and attached to rigid surfaces. Each distinct marker in FMS has its unique ID and a unique graphical representation to be recognized and decoded by a corresponding computer vision algorithm. FMSs are widely used in a broad range of tasks in physics, metrology, medicine, augmented reality, etc. In robotics, they act as a main element in navigation [1]–[3], localization [4]–[7], camera calibration [8]–[10] and mapping [11]–[13] tasks. Wide FMS usage implies multiple FMS types existence and development; each of them has its strengths and weaknesses depending on features of a particular task to be accomplished.

Our long-term target is to perform a camera calibration on AR-601M humanoid robot [14] (Fig. 2) and mobile robot Servosila Engineer [15] (Fig. 1) in an automated manner. Therefore, we need to choose the most appropriate FMS that will be further placed on the robot body surface and allow the robot to calibrate its cameras without human assistance, which is typical for standard calibration procedures [16].

Considering our long-term goal, we selected a set of criteria, which are the most important for our task. In a robot self-calibration scenario, a marker could be placed on the robot's manipulator, which will allow capturing multiple frames with the marker in its different projections. During manipulator

This work was supported by the Russian Foundation for Basic Research (RFBR), project ID 18-58-45017.

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

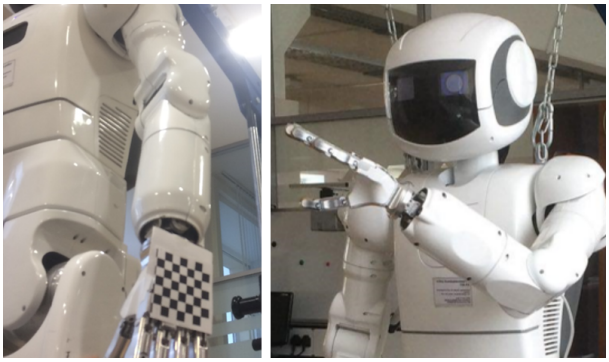

Fig. 2. AR-601M manipulator with a checker-board marker attached to its left hand palm (left); the robot performs marker observation (right).

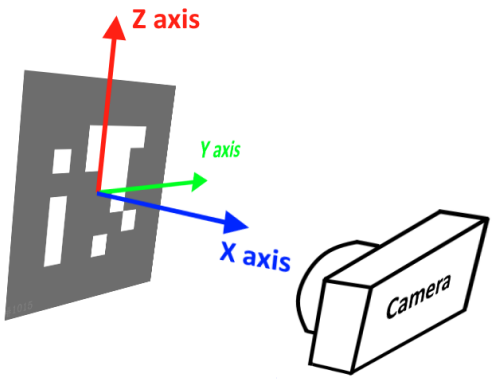

Fig. 3. Marker rotation axes alignment relatively to the camera.

movements, the marker is rotated and may be partially overlapped by other elements of the robot. Thus, chosen markers should be resistant to marker rotations, partial occlusions, and sensor noise. However, previous researches lack systematic approaches and fair comparisons, for example, occlusion resistance tests were performed in a random manner [17]. Another case is a comparison between two versions of the same FMS [18].

In our previous researches [19], [20], we compared AprilTag [13], [21], ArTag [22], [23] and CALTag [24] FMSs using manual experiments, including rotations about axes, systematic and arbitrary occlusions. Results analysis showed that all three FMSs are practically resistant to rotations around X axis; however, CALTag demonstrated significantly better results in experiments for occlusions. However, manually conducted experiments have strong disadvantages:

- Very high costs in terms of time and human resources. Collecting a statistically significant amount of data requires conducting multiple iterations of each experiment.
- Experiments' fairness control problem. Environment conditions are typically unstable and during real-world experiments a precise inclination angle, a marker position with regard to a camera, lighting conditions, etc. are difficult to track and control.
- Limited hardware choice. Real sensors have unavoidable noises and constrain comparison possibilities since their physical properties, e.g., camera resolution, lens distortion level, optical sensor sensitivity are static and could not be modified arbitrarily.

These disadvantages limit the reproducibility of experiments and therefore decrease the scientific significance of obtained results. In order to overcome these disadvantages, in our work the experiments were transferred to a virtual environment with strictly controlled external conditions [25]. This paper introduces the developed virtual experiment environment in the ROS/Gazebo simulator and the results of experiments for rotation resistance under different noise levels of input sensory data. Experiments are focused on different marker families' resistance to the noise. This noise is unavoidable in practice and should have minimum influence on the marker detection results. Therefore, this paper provides AprilTag [26] and ArUco marker families comparison.

Section 2 briefly introduces the Robot Operating System framework and the Gazebo simulator that were used in research. Section 3 describes virtual environment properties and experimental setup. The results of the virtual experiments are presented in section 4. The last section concludes the presented research and identifies further researches direction.

II. Tools: ROS and Gazebo

Robot Operating System (ROS) is a framework that unites multiple executables into a single structure with unified data streams between them. Executables in ROS system are called *nodes* and data streams for their communication are called *topics*. Nodes get data from input topics, process them, and send them to output topics. Such distributed nature and flexible design of ROS allow creating various data processing schemes, including complex navigation systems or computer vision systems [27]. ROS has a software distribution system, which is based on the so-called *packages*. Packages are software units responsible for a particular robot functionality, e.g., for robot visualization, locomotion control, path planning, sensory data gathering, etc. A package could be easily replaced due to standardized data exchange interfaces.

Gazebo is a 3D-simulator, which is fully integrated into ROS as its node. The Gazebo is responsible for simulation of real-world features, including physics, lighting and shadow rendering, collision detection, etc. A wide range of Gazebo tools for creating and manipulating with complex environments allows employing it for virtual environment visualization. Gazebo uses OGRE rendering engine [28], which is a well-developed open-source rendering engine. Since ROS-Gazebo combination is probably the most popular standard tool among robot developers today [29], all virtual experiments, which are run within this framework, could be easily

and independently reproduced by other researchers. An open-source code of projects gives an opportunity for anyone to peer review internals of any experiments. This standardizes experiments' architecture and makes it independent on a particular fiducial marker system or marker detection and recognition algorithms. Such approach allows adding new marker systems and, moreover, provides equal experimental conditions for all existing and newly created markers.

For our virtual experiments we utilized official ROS packages for the ArUco and AprilTag markers detection and recognition, *aruco_ros* package [30] and *apriltag_ros* package [31] accordingly. These packages use canonical recognition algorithms implementations. Using the Gazebo we created required external environment conditions and controlled them during the virtual experiments. The Gazebo allowed simulating real hardware properties, including lens distortion, sensor resolution, etc., and adding various levels of sensor noise.

III. Experiment setup

Multiple external environment conditions may influence a marker detection process. Virtual experiments are designed to eliminate their influence since an external environment is always kept ideal and constant. Experiment conditions are presented in Table I. This paper is dedicated to fiducial markers resistance for rotations with ideal sensor (no noise) and in presence of Gaussian noise, therefore, the sensory noise level is defined at the beginning of each experiment; a marker rotation angle is changing during the experiment. Rotations are performed about X and Z axes (shown on Fig. 3). Rotations about X-axis do not affect a visible segment of a marker; Y-axis and Z-axis rotations gradually decrease the visible percent of the marker in a very similar manner, therefore, only Z-axis rotation experiments were conducted.

Noise added to the rendered images was zero-mean additive white Gaussian noise (AWGN) [32]. This noise had Gaussian probability distribution and was independent of a rendered pixel color. Since the mean value was static, we were changing a standard deviation value to determine noise influence on detection results. This noise type approximates noises of a digital camera optical sensor under difficult conditions [33], e.g., poor lighting or low optical sensor quality. Noise levels used in experiments are shown in Fig. 4. Other constant parameters (light properties, distances, etc.) were chosen in such a manner that they could allow constructing an ideal environment and did not affect recognition results.

Virtual environment parameters were controlled by a user and were used to conduct experiments in various conditions. All user-defined parameters' values were set using a developed graphical user interface (GUI) that allows easy and fast manual setup of a virtual experiment. User-defined parameters, their possible values and particular values that were chosen for the experiments are listed in Table II.

The virtual environment includes two robots: a robot-performer and a robot-observer (shown on Fig. 5). The robot-performer was designed to hold a marker and rotate it within

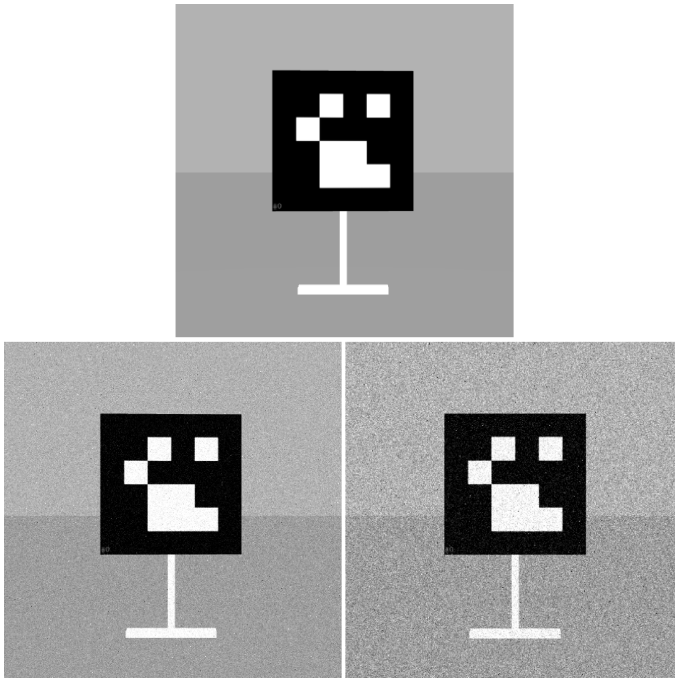

Fig. 4. Gaussian noise levels: ideal sensor (0%, top); moderate noise level (5%, bottom left); high noise level (10%, bottom right).

TABLE I
Constant parameters of experimental setup

Parameter	Value
Camera distortion level	0 (ideal lens)
Lens mapping function	Gnomonic (perspective)
Marker size	1 m. x 1 m.
Light angle of incidence	45°
Light spectrum	White light
Light conditions	Uniform in entire marker area

a particular interval between the detection series. The robot-observer was static and acted as a stand for a virtual camera. The robots were spawned on a user-defined distance between them and an experiment began. The experiment flow for each level of Gaussian noise was as follows (sequence diagram is shown on Fig. 6):

1) The robot-observer captures the user-defined number of frames with its virtual camera. In our experiments, 100 frames were captured in each detection series.
2) Each captured frame is sent as input for corresponding fiducial marker detection and recognition algorithm.
3) A logging system records the detection result of each captured result into a file and duplicates it into the user console. The result could be a success or a failure. Detection series at a particular angle (which is a set of 100 camera shots in the same pose of the marker under inspection and the camera) with a particular noise level gives a number of successful detections that lies in the

TABLE II
USER-CONTROLLED EXPERIMENTAL SETUP PARAMETERS

Parameter	Possible values range	Selected values
Tag family and type	AprilTag (4 types) Alvar [34] ArTag ArToolKitPlus [35] ArUco (2 types) ChiliTag [36]	AprilTag (type 25h7) & ArUco (type 25h7)
Camera resolution	0.3 MP (640 x 480 px) 1 MP (1280 x 960 px) 2 MP (1600 x 1200 px)	0.3 MP
Distance between a camera and a marker (m.)	[0; +∞]	2
Border presence	With or without white border	Without white border
Gaussian noise standard deviation (%)	[0;100]	0 & 5 & 10
Number of frames analyzed for each marker state	[10;100]	100
Simulation threads	Up to hardware limit	4
GUI mode	GUI or non-GUI	non-GUI

range of 0 to 100.

4) If a rotation limit is not reached, the robot-performer rotates the marker one degree further and the current algorithm returns to step 1; otherwise, the current algorithm selects a next marker type and returns to step 1.

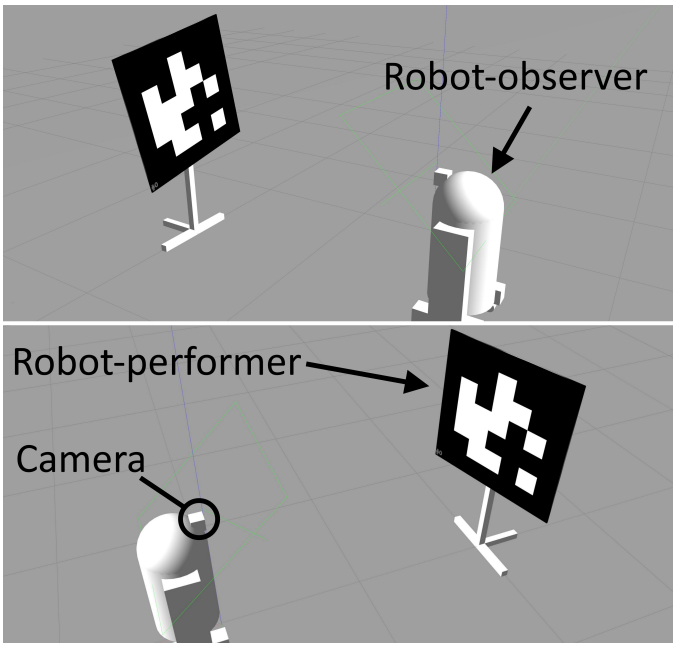

Fig. 5. Robot-performer rotated a marker by 10 degrees (top) around X-axis; robot-performer rotated marker by 15 degrees around Z-axis (bottom).

IV. RESULTS

Sets of markers were chosen to provide a fair comparison between detection and recognition algorithms of AprilTag and ArUco families, therefore, markers of similar type (i.e. 25h7 type) were selected. Similar types of FMSs eliminated the difference between them in encoding schemes: both types

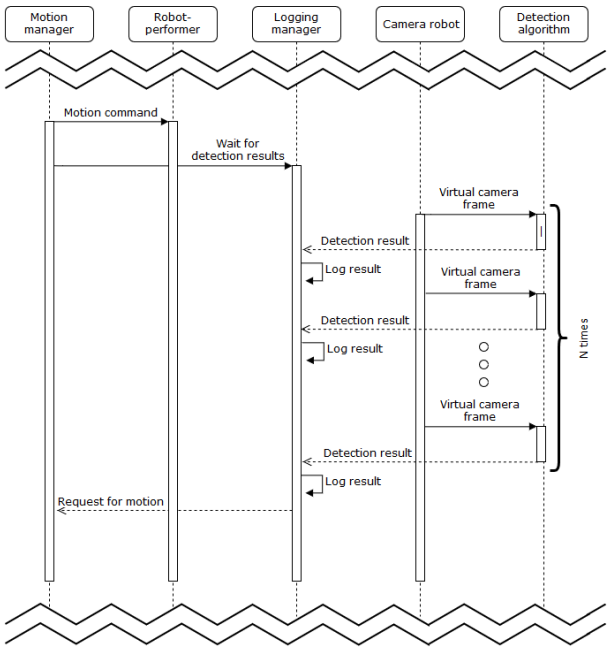

Fig. 6. Experiments sequence diagram.

have 25 encoding pixels and a minimum Hamming distance between any two markers in a set is equal or more than 7. Hamming distance equality means that both types have equal chances to detect a marker properly in cases when a part of the marker is impossible to recognize or visual data is corrupted (e.g., noisy). For instance, let us assume there are two markers to be detected: marker A encodes "10100011" and marker B encodes "11001101". In the case of corrupted input data, a decoded sequence may be "10001011". The Hamming distance value is computed using a digit-by-digit comparison between an input sequence and markers' sequences. In the case of marker A, its value equals to 2, in the case of marker B, the value equals to 3. Thus, a decoded marker is more likely to be marker A, allowing the detection algorithm to restore

TABLE III
RELIABLE DETECTION RATES IN VIRTUAL EXPERIMENTS

	Gaussian noise level (%)					
	0 (ideal)		5 (moderate)		10 (high)	
Rotation axis	X	Z	X	Z	X	Z
AprilTag - 25h7	99.94	69.96	30.87	5.42	0.0	0.0
ArUco - 25h7	99.97	86.06	99.79	48.96	0.0	0.0

corrupted data.

In order to estimate practical applicability, results of virtual experiments are provided in a form of reliable detection series rates. The detection series is called reliable if a marker is detected in at least 95% of the captured frames. Rates that are presented in Table III show the percentage of reliable detection series among all series for a particular marker family, noise level and rotation axis.

Results analysis showed predictable behavior of the markers in a presence of Gaussian noise. Noisy sensory data decreased successful detection probability and reliable detection became more difficult for detection algorithms. Reliable detection series number decreased significantly with Gaussian noise level increase and tended to zero at a high noise level. Yet, a moderate noise level did not stop the detection completely and both tested marker types were still reliably detectable with a significant ArUco advantage (Fig. 7 and 8).

A high level of noise prevented reliable detection for both tested marker systems, however, AprilTag markers still provided at least 25% detection rate (Fig. 9 and Fig. 10).

It is worth to note that detection rate distributions on X-axis rotations are clearly periodic with periods being equal to approximately 90 degrees. In our opinion, this pattern could be caused by two reasons:

- Virtual camera frame render features;
- The detection algorithms features.

Assuming that a pattern is caused by the virtual frame render features, this should lead to the same pattern for different plots. Moreover, in such case the detection behavior must be the same for different marker systems under the same conditions. However, as it is shown in Fig. 7 and Fig. 9, it differs for AprilTag and ArUco systems. Therefore, the assumption about render engine is wrong and the actual reason is that the detection algorithms perform better at particular angles.

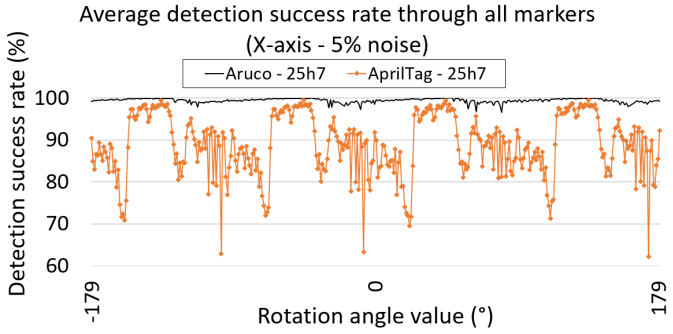

Fig. 7. Detection rates distributions in X-rotation experiments with 5% Gaussian noise.

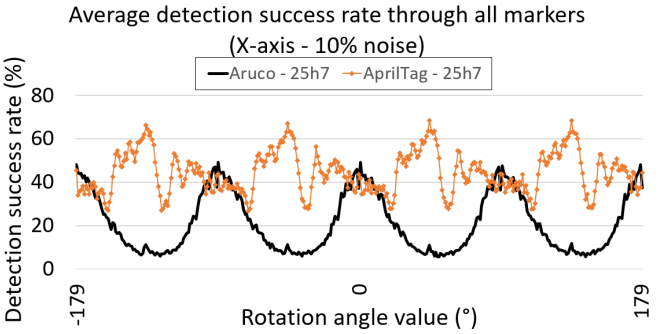

Fig. 9. Detection rates distributions in X-rotation experiments with 10% Gaussian noise.

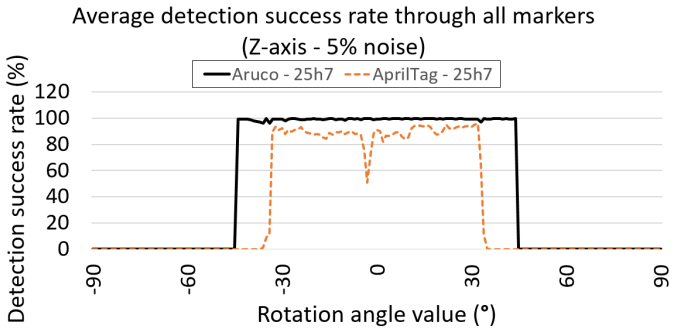

Fig. 8. Detection rates distributions in Z-rotation experiments with 5% Gaussian noise.

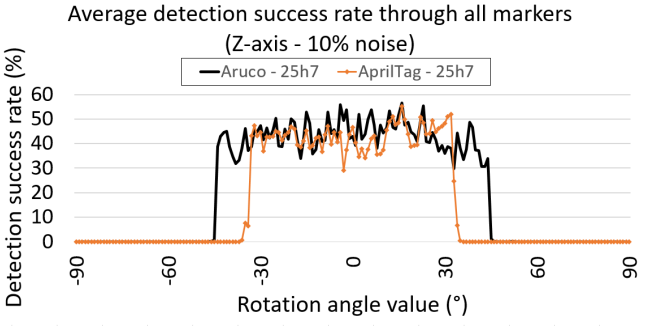

Fig. 10. Detection rates distributions in Z-rotation experiments with 10% Gaussian noise.

V. Conclusions and future work

This presented the comparison of detection and recognition algorithms for AprilTag and ArUco fiducial marker families concentrating on their resistance to rotations in the presence of synthetic noise. The constructed virtual environment allowed to eliminate external environment conditions influence and accurately control them. Synthetic Gaussian noise was added into experiments to estimate detection algorithms' resistance to noisy sensory data presence. The virtual experiments results demonstrated that ArUco (type 25h7) and AprilTag (type 25h7) were equally insensitive to X-axis rotations when no Gaussian noise is added to pure sensory data; however, ArUco markers were significantly better in experiments for Z-axis rotations. The same tendency was seen in conditions of moderate Gaussian noise in sensory data, and the reliable detection rate of ArUco markers was significantly higher than AprilTag markers detection rate. However, both marker systems decreased their successful detection rates and failed to perform reliable detection at 10% of the Gaussian noise level.

Proper fiducial marker choice increases machine vision methods efficiency. In our future research, we will consider more fiducial marker families and types, which will be used for comparison. Marker systems will be tested for their resistance to occlusions, both systematic and arbitrary. In addition, the angular size influence on the detection rate will be studied further.

References

[1] R. Kuriya, T. Tsujimura, and K. Izumi, "Augmented reality robot navigation using infrared marker," in *24th International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2015, pp. 450–455.

[2] A. Mutka, D. Miklic, I. Draganjac, and S. Bogdan, "A low cost vision based localization system using fiducial markers," *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 9528–9533, 2008.

[3] H. Lim and Y. S. Lee, "Real-time single camera slam using fiducial markers," in *2009 ICCAS-SICE*. IEEE, 2009, pp. 177–182.

[4] V. Dhiman, J. Ryde, and J. J. Corso, "Mutual localization: Two camera relative 6-dof pose estimation from reciprocal fiducial observation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 1347–1354.

[5] F. J. R. Ramírez, R. M. Salinas, and R. M. Carnicer, "Fractal markers," *Creando Redes Doctorales*, p. 693, 2019.

[6] R. Chang, Y. Li, and C. Wu, "Robust identification of visual markers under boundary occlusion condition," in *International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2019, pp. 2908–2913.

[7] M. Neunert, M. Bloesch, and J. Buchli, "An open source, fiducial based, visual-inertial motion capture system," in *19th International Conference on Information Fusion (FUSION)*. IEEE, 2016, pp. 1523–1530.

[8] M. Fiala, "Comparing ARTag and ARToolkit Plus fiducial marker systems," in *International Workshop on Haptic Audio Visual Environments and their Applications*. IEEE, 2005, p. 6.

[9] S. Daftry, M. Maurer, A. Wendel, and H. Bischof, "Flexible and user-centric camera calibration using planar fiducial markers." in *BMVC*. Citeseer, 2013.

[10] D. Claus and A. W. Fitzgibbon, "Reliable automatic calibration of a marker-based position tracking system," in *7th Workshops on Applications of Computer Vision*, vol. 1. IEEE, 2005, pp. 300–305.

[11] A. Bokovoy and K. Yakovlev, "Original loop-closure detection algorithm for monocular vslam," in *International Conference on Analysis of Images, Social Networks and Texts*. Springer, 2017, pp. 210–220.

[12] R. Lavrenov, "Smart spline-based robot navigation on several homotopies: Guaranteed avoidance of potential function local minima," in *Int. Conf. on Artificial Life and Robotics (ICAROB)*, 2018, pp. 407–410.

[13] E. Olson, "AprilTag: A robust and flexible visual fiducial system," in *International Conference on Robotics and Automation*. IEEE, 2011, pp. 3400–3407.

[14] E. Magid and A. Sagitov, "Towards robot fall detection and management for russian humanoid ar-601," in *KES International Symposium on Agent and Multi-Agent Systems: Technologies and Applications*. Springer, 2017, pp. 200–209.

[15] I. Moskvin and R. Lavrenov, "Modeling tracks and controller for servosila engineer robot," in *Proceedings of 14th International Conference on Electromechanics and Robotics "Zavalishin's Readings"*. Springer, 2020, pp. 411–422.

[16] E. A. Martínez-García, *Numerical Modelling in Robotics*. Omnia-Science, 2015.

[17] M. Fiala, "Comparing ARTag and ARToolkit Plus fiducial marker systems," in *International Workshop on Haptic Audio Visual Environments and their Applications*. IEEE, 2005, pp. 6–pp.

[18] J. Wang and E. Olson, "AprilTag 2: Efficient and robust fiducial detection," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016, pp. 4193–4198.

[19] K. Shabalina, A. Sagitov, L. Sabirova, H. Li, and E. Magid, "ARTag, AprilTag and CALTag Fiducial Systems Comparison in a Presence of Partial Rotation: Manual and Automated Approaches," in *International Conference on Informatics in Control, Automation and Robotics*. Springer, 2017, pp. 536–558.

[20] A. Sagitov, K. Shabalina, R. Lavrenov, and E. Magid, "Comparing fiducial marker systems in the presence of occlusion," in *International Conference on Mechanical, System and Control Engineering (ICMSC)*. IEEE, 2017, pp. 377–382.

[21] S. M. Abbas, S. Aslam, K. Berns, and A. Muhammad, "Analysis and improvements in apriltag based state estimation," *Sensors*, vol. 19, no. 24, p. 5480, 2019.

[22] M. Fiala, "ARTag, a fiducial marker system using digital techniques," in *Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2. IEEE, 2005, pp. 590–596.

[23] M. Hirzer, "Marker detection for augmented reality applications," in *Seminar/Project Image Analysis Graz*, 2008, pp. 1–2.

[24] B. Atcheson, F. Heide, and W. Heidrich, "Caltag: High precision fiducial markers for camera calibration." in *VMV*, vol. 10, 2010, pp. 41–48.

[25] A. Zakiev, K. Shabalina, T. Tsoy, and E. Magid, "Pilot virtual experiments on aruco and artag systems comparison for fiducial marker rotation resistance," in *Proceedings of 14th International Conference on Electromechanics and Robotics "Zavalishin's Readings"*. Springer, 2020, pp. 455–464.

[26] F. J. Romero-Ramirez, R. Muñoz-Salinas, and R. Medina-Carnicer, "Speeded up detection of squared fiducial markers," *Image and vision Computing*, vol. 76, pp. 38–47, 2018.

[27] I. Z. Ibragimov and I. M. Afanasyev, "Comparison of ros-based visual slam methods in homogeneous indoor environment," in *Workshop on Positioning, Navigation and Communications (WPNC)*, 2017, pp. 1–6.

[28] O. Team, *OGRE - Open Source 3D Graphics Engine*, 2020. [Online]. Available: https://www.ogre3d.org/

[29] R. Safin and R. Lavrenov, "Implementation of ros package for simultaneous video streaming from several different cameras," in *International Conference on Artificial Life and Robotics (ICAROB)*, 2018, pp. 220–223.

[30] B. Magyar, *aruco - ROS Wiki*, 2020. [Online]. Available: http://wiki.ros.org/aruco

[31] J. Stechschulte, *apriltag_ros - ROS Wiki*, 2020. [Online]. Available: http://wiki.ros.org/apriltag_ros

[32] R. A. Boie and I. J. Cox, "An analysis of camera noise," *Transactions on Pattern Analysis & Machine Intelligence*, no. 6, pp. 671–674, 1992.

[33] W. Liu and W. Lin, "Additive white gaussian noise level estimation in svd domain for images," *Transactions on Image processing*, vol. 22, no. 3, pp. 872–883, 2012.

[34] C. Woodward, S. Valli, P. Honkamaa, and M. Hakkarainen, "Wireless 3d cad viewing on a pda device," in *Proceedings of the 2nd Asian International Mobile Computing Conference*, vol. 14, 2002, p. 17.

[35] D. Wagner and D. Schmalstieg, "Artoolkitplus for pose tracking on mobile devices," *Computer Vision Winter Workshop*, 2007.

[36] Q. Bonnard, S. Lemaignan, G. Zufferey, A. Mazzei, S. Cuendet, N. Li, A. Ozgur, and P. Dillenbourg, "Chilitags 2: Robust fiducial markers for augmented reality and robotics," *CHILI, EPFL, Switzerland*, 2013.