

Study of Computer Vision for Manipulators

Perizat Rakhmetova
Dept. of Robotics and technical means
of automation
Satbayev University
Almaty, Kazakhstan
p.rakhmetova@gmail.com

Aldabergen Bektilevov
Dept. of Transport engineering and
logistics named after M. Tynyshpayev
Satbayev University
Almaty, Kazakhstan
a.bektilevov@satbayev.universsity

Bakhtiyar Kopenov
Dept. of Motor vehicles and life-safety
Academy Logistics and Transport
Almaty, Kazakhstan
madiiskakov69@mail.ru

Makabbat Zhamuratova
Dept. of Robotics and technical means
of automation
Satbayev University
Almaty, Kazakhstan
m.zhamuratova@satbayev.university

Saltanat Yussupova
Dept. of Electronic Engineering
Almaty University of Power
Engineering and Telecommunications
Almaty, Kazakhstan
s.yusupova@aes.kz

Tolkyn Toktamyssova
Dept. of Motor vehicles and life-safety
Academy Logistics and Transport
Almaty, Kazakhstan
tolkyn_aktobe@mail.ru

Abstract—In recent years, computer vision technology has made significant progress, expanding its application from simple image recognition tasks to complex real-world problems. One such area where computer vision promises to be is the control of manipulators in various fields such as manufacturing, healthcare and robotics. Manipulators, mechanical arms designed to perform tasks, are traditionally controlled using manual input or a preprogrammed sequence. The purpose of this manuscript is to research and analyze computer vision applicable to improve the control and efficiency of manipulators.

Keywords—computer vision, manipulator, object recognition, control.

I. INTRODUCTION

Computer vision for manipulator control is an advanced field of research combining the capabilities of artificial intelligence, robotics and image processing. This multidisciplinary approach aims to improve the performance and capabilities of robotic weapons, allowing them to interact more responsibly and efficiently with the environment. At its core, technology relies on computer vision algorithms that can interpret and analyze visual data in real time. These algorithms serve as the «eyes» of the robotic hand, providing it with important information about the environment. The main purpose of this article is to provide a comprehensive research of computer vision for robotic manipulators, which will become the basis of theoretical and practical significance. To achieve this goal, it is necessary to solve the following tasks:

- Study and analysis of scientific articles and works devoted to the problems of computer vision in robotics;
- Assessment and comparison of existing methods and algorithms for object recognition, position estimation, motion and other aspects of computer vision;
- Building a manipulator model and conducting experiments to evaluate the effectiveness of existing and new methods and algorithms;
- Data processing and analysis;

- Analysis of possible areas of application of the developed methods and systems, including their economic efficiency.

II. LITERATURE REVIEW

In recent years, computer vision has evolved into a transformative domain with diverse applications in different fields. This rapid evolution is evident in studies presented in numerous scientific publications. Significant work in the realm of healthcare computer vision has been conducted by Q. T. Do and J. Chaudri [1], particularly focusing on identifying and locating respiratory devices on a patient's face. Their research underscores the possibilities for enhancing medical diagnostic procedures. Another intriguing direction in computer vision research is the development of biotechnology, as demonstrated by M. J. Domínguez-Morales et al. [2]. The research is deepened into the field of dynamic vision sensors and their perspectives for the development of computer vision capabilities. Developed by scientists a biotechnological stereovision system in which a new method of calibration of neuromorphic stereovision system, implemented on specialized equipment with the interpretation of programmable logic integrated circuits (FPGA) Allows for reduced hardware implementation delays for stand-alone systems and real-time operation. The same applications with FPGA were used in the work of D. Castells-Rufas et al. [3], which analyses the maximum performance, energy efficiency, and strengths and weaknesses of these platforms in autonomous cars.

The integration of computer vision technology into manipulator control systems presents a transformative approach to automation and robotics. This multidisciplinary field converges computer science, electrical engineering, and mechanical engineering to foster enhanced perception, decision-making, and execution capabilities in robotic manipulators. The utility of computer vision spans various domains including indoor navigation [4], data compression [5], and even medical applications such as vein recognition [6]. With the increasing complexity of tasks that manipulators are expected to perform, the role of computer vision becomes even more pivotal. Recent advancements in algorithms, particularly deep learning techniques, have significantly improved the accuracy and reliability of computer vision-based perception systems [7]. However, the

real challenge lies in efficiently translating this high-dimensional data into meaningful control commands for manipulators. This review aims to provide a comprehensive overview of the methodologies, algorithms, and frameworks that leverage computer vision for manipulator control. We cover a wide array of applications including but not limited to object recognition, path planning, and human-robot collaboration. Special attention is given to vision-based control in adverse environmental conditions, such as in post-flood scenarios [8] and specific domains like agriculture [9] and healthcare.

In particular, this review also addresses the computational limitations and challenges of implementing computer vision systems in manipulators. Issues such as lossy compression to reduce data overhead [5] and the trade-off between computational efficiency and perceptual accuracy are discussed. In addition, the latest trends in algorithms used in computer vision applications are studied. Finally, future perspectives for this growing field are explored, including the role of advanced styles of machine learning [10], adversarial attacks and defenses, and how new technologies such as Vision Transformers can be adapted to control manipulators [6]. This review serves as a comprehensive guide for researchers, engineers, and practitioners seeking to explore the untapped potential of computer vision in the field of robotic manipulators.

III. METHODS

Traditionally, manipulators were controlled by a set of predefined instructions or by manual control. However, these techniques lack the flexibility and adaptability required for complex tasks such as sorting objects, precise assembly, or human-robot collaboration [11]. Computer vision fills this gap, allowing robotic hands to “see” and “understand” the world around. This is achieved through various methods such as object detection, image segmentation and depth assessment [12].

The integration of computer vision into manipulator control systems ensures more dynamic and operational behavior [13]. Robots can now adjust their actions depending on what they see, thus optimizing their performance for a wide range of tasks. For example, a robotic hand equipped with computer vision can determine the shape and orientation of an object before attempting to grab it, thereby increasing the chances of successful capture.

This technology has not only industrial applications, but is also increasingly used in health care, agriculture and even space exploration [14]. However, there have been challenges. Problems such as computational complexity, lighting conditions and sensor noise may affect the reliability of computer vision algorithms.

In the study of computer vision analysis for manipulator control, the methodology described below includes a multi-stage process that provides accurate and adaptable robotic manipulation. The study begins with a thorough review of the literature to understand the current situation, including existing algorithms, equipment configuration, and problems in computer vision and robotic manipulators [10, 15].

The first stage involves the selection of suitable equipment. For example, a manipulator with six degrees of freedom (6-DOF) in the Fig. 1 and a high-resolution camera. The camera serves as a video sensor and is strategically

positioned to capture the working space of the manipulator [4, 14]. It is essential that the camera and the manipulator be well calibrated to provide a single frame of reference.

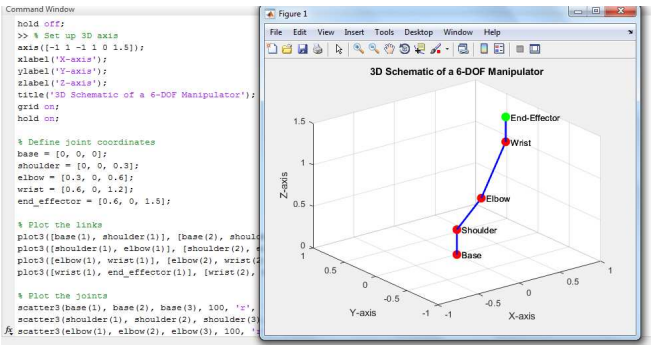


Fig. 1. 3D model of manipulator with 6-DOF.

It then focuses on data collection. Objects of various shapes, sizes [16] and materials are placed in the workspace. The camera captures the images [17, 18], which are then annotated to create a dataset for training computer vision models. Special attention is paid to lighting conditions and changes in the viewing angle to ensure model reliability. Computer Vision, a branch of AI, is concerned with teaching computers to grasp and interpret visual data [16, 17]. It entails image processing, object identification, and picture segmentation (see Fig. 2). Image processing is the application of algorithms [18] to improve photographs or extract meaningful information from them. This signal processing method takes an image as input and outputs an upgraded picture or features connected with that image. A block diagram of computer vision, in particular the stages of object recognition in interpretation with a manipulator, is shown in the Fig. 3.

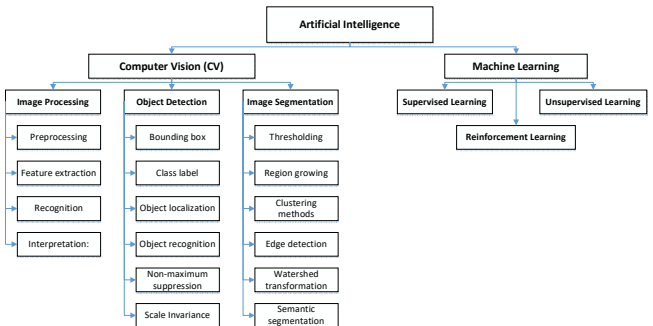


Fig. 2. Structure diagram of Artificial Intelligence.

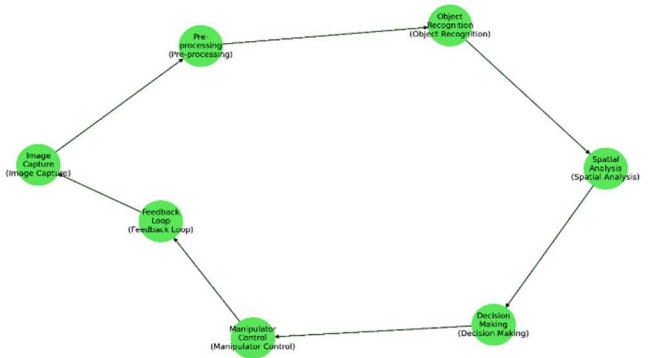


Fig. 3. Structure diagram of Computer Vision for Manipulators.

The third stage uses deep learning techniques [19] to develop computer vision algorithms. Algorithms are designed to perform tasks such as object recognition [20], segmentation, and depth estimation. The Circadian Neural Network (CNN) [15, 19, 21] is used for object recognition and the U-Net architecture for image segmentation [22]. Depth assessment is achieved using techniques [10, 22] such as stereovision or depth cameras (Fig. 4).

After training, the models are integrated into the robot manipulator control system. This involves developing an intermediate level of software that provides real-time communication between the computer vision system [23] and the manipulator control unit.

To assess performance, a series of experiments are conducted, including tasks such as selecting, sorting and assembling objects. Key performance indicators (KPI) such as accuracy, speed and energy efficiency are measured. Variables such as overwhelmed environment and moving objects are introduced to assess the adaptability of the system.

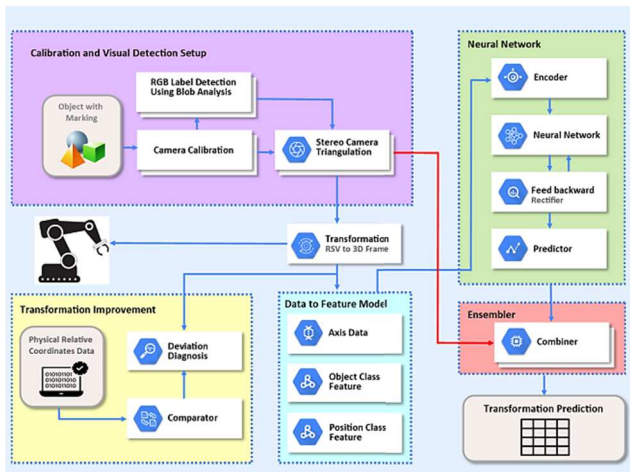


Fig. 4. Robotic manipulator platform with stereovision for tracking object and detection.

After data collection, statistical analysis is carried out to evaluate the effectiveness of our approach [24]. These results are then compared with established methods to identify areas that could benefit from improvement.

As a final step, the methodology includes an iterative fine-tuning step. Depending on the assessment, both models and control algorithms are adjusted to optimize performance.

In general, the technique outlined above is intended to be thorough and iterative, with the goal of broadening the scope of what is achievable in computer vision analysis for manipulator control. It blends cutting-edge machine learning techniques, robotics, and software development to create a system that is both automated and intelligent and flexible.

IV. RESULTS

The story emerging from these references paints a vibrant and ever-changing tableau of advancements in computer vision, especially as they pertain to control systems for robotic manipulators. As we probe further into next-generation algorithms, machine learning techniques, and practical applications, the possibilities for a seamless integration of computer vision with manipulator control systems become increasingly clear. This merger signifies a

new phase in automation, where machines are not merely task executors but also have the capabilities to sense, analyze, and adapt to their intricate and ever-changing surroundings.

In results, the application of computer vision in manipulator control represents a groundbreaking innovation with the potential to transform various industries. It melds state-of-the-art computational methods with robotic components, resulting in systems that are not just automated but also intellectually capable. As the field continues to evolve, we anticipate the development of even more advanced and resilient systems that will redefine the scope of what is achievable in robotic manipulation.

The incorporation of computer vision enables manipulators to perceive and interpret their environment in a manner akin to human vision. This technological leap has wide-ranging implications, elevating the control of manipulators to unparalleled levels of precision, adaptability, and effectiveness.

V. DISCUSSION

Understanding the capabilities and limitations of computer vision in manipulator management is critical to developing more autonomous and efficient systems. These studies can pave the way for innovations in industrial automation, surgical robotics and other applications

Although the potential of computer vision in manipulator control is enormous, it is not without problems. Ensuring reliability, scalability and security in these systems remains a priority. In addition, ethical privacy considerations and decision algorithms need to be taken into account as technology develops.

Looking ahead, we may anticipate even more dramatic advances in computer vision for manipulator control. Machine learning and artificial intelligence will be critical in allowing manipulators to adapt to changing contexts, make sophisticated judgments, and interact freely with humans.

VI. CONCLUSION

The synergy between computer vision and manipulator management changes the face of industries in all directions. These systems are not just automation tools; they are the vanguard of innovation, opening up new possibilities and redefining the boundaries of what robots and manipulators can achieve. As the world continues to implement this transformative technology, we can foresee a future in which computer-assisted manipulator management will be an indispensable cornerstone of progress and efficiency.

In this research article, the integration of a computer vision system with a manipulator was carried out, the possibilities of using the developed methods in various fields were assessed, and recommendations for further research were formulated. Thus, the implementation of the assigned tasks made it possible not only to study in detail current problems in the field of computer vision for robotic manipulators, but also to propose specific methods and solutions that have both theoretical and practical significance. This study serves as a starting point for further developments and adaptations, as well as for the application of the acquired knowledge in various industrial and scientific fields.

REFERENCES

- [1] Q. T. Do and J. Chaudri, "Creating Computer Vision Models for Respiratory Status Detection," *IEEE Engineering in Medicine & Biology Society, Glasgow, Scotland, United Kingdom*, 2022, pp. 1350-1353, doi: 10.1109/EMBC48229.2022.9871978.
- [2] M. J. Domínguez-Morales, Á. Jiménez-Fernández, G. Jiménez-Moreno, C. Conde, E. Cabello and A. Linares-Barranco, "Bio-Inspired Stereo Vision Calibration for Dynamic Vision Sensors," in *IEEE Access*, vol. 7, pp. 138415-138425, 2019, doi: 10.1109/ACCESS.2019.943160.
- [3] D. Castells-Rufas et al., "A Survey of FPGA-Based Vision Systems for Autonomous Cars," in *IEEE Access*, vol. 10, pp. 132525-132563, 2022, doi: 10.1109/ACCESS.2022.3230282.
- [4] L. Ruotsalainen, A. Morrison, M. Mäkelä, J. Rantanen and N. Sokolova, "Improving Computer Vision-Based Perception for Collaborative Indoor Navigation," in *IEEE Sensors Journal*, vol. 22, no. 6, pp. 4816-4826, 15 March 15, 2022, doi: 10.1109/JSEN.2021.3106257.
- [5] L. Guo, D. Zhou, J. Zhou, S. Kimura and S. Goto, "Lossy Compression for Embedded Computer Vision Systems," in *IEEE Access*, vol. 6, pp. 39385-39397, 2018, doi: 10.1109/ACCESS.2018.2852809.
- [6] X. Li and B. -B. Zhang, "FV-ViT: Vision Transformer for Finger Vein Recognition," in *IEEE Access*, vol. 11, pp. 75451-75461, 2023, doi: 10.1109/ACCESS.2023.3297212.
- [7] M. Faisal, F. Albogamy, H. Elgibreen, M. Algabri and F. A. Alqershi, "Deep Learning and Computer Vision for Estimating Date Fruits Type, Maturity Level, and Weight," in *IEEE Access*, vol. 8, pp. 206770-206782, 2020, doi: 10.1109/ACCESS.2020.3037948.
- [8] M. Rahneemounfar, T. Chowdhury, A. Sarkar, D. Varshney, M. Yari and R. R. Murphy, "FloodNet: A High Resolution Aerial Imagery Dataset for Post Flood Scene Understanding," in *IEEE Access*, vol. 9, pp. 89644-89654, 2021, doi: 10.1109/ACCESS.2021.3090981.
- [9] Y. Tian, G. Pan and M. -S. Alouini, "Applying Deep-Learning-Based Computer Vision to Wireless Communications: Methodologies, Opportunities, and Challenges," in *IEEE Open Journal of the Communications Society*, vol. 2, pp. 132-143, 2021, doi: 10.1109/OJCOMS.2020.3042630.
- [10] A. Shahzad, X. Gao, A. Yasin, K. Javed and S. M. Anwar, "A Vision-Based Path Planning and Object Tracking Framework for 6-DOF Robotic Manipulator," in *IEEE Access*, vol. 8, pp. 203158-203167, 2020, doi: 10.1109/ACCESS.2020.3037540.
- [11] I. Domingues, I. L. Sampaio, H. Duarte, J. A. M. Santos and P. H. Abreu, "Computer Vision in Esophageal Cancer: A Literature Review," in *IEEE Access*, vol. 7, pp. 103080-103094, 2019, doi: 10.1109/ACCESS.2019.2930891.
- [12] H. Chen and Z. Deng, "Bibliometric Analysis of the Application of Convolutional Neural Network in Computer Vision," in *IEEE Access*, vol. 8, pp. 155417-155428, 2020, doi: 10.1109/ACCESS.2020.3019336.
- [13] J. M. Ponce, A. Aquino, B. Millan and J. M. Andújar, "Automatic Counting and Individual Size and Mass Estimation of Olive-Fruits Through Computer Vision Techniques," in *IEEE Access*, vol. 7, pp. 59451-59465, 2019, doi: 10.1109/ACCESS.2019.2915169.
- [14] M. R. Gubbi, E. A. Gonzalez and M. A. L. Bell, "Theoretical Framework to Predict Generalized Contrast-to-Noise Ratios of Photoacoustic Images With Applications to Computer Vision," in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 69, no. 6, pp. 2098-2114, June 2022, doi: 10.1109/TUFFC.2022.3169082.
- [15] G. Balbayev, A. Mussina, A. Zhauyt, B. Shingissov and M. Kalekeyeva, "Neural Control for Image Stabilisation Using a Reference Model," vol. 10, no. 1, 2021, pp. 17-21, doi: 10.18178/IJMERR.10.1.17-21.
- [16] D. Mery, D. Saavedra and M. Prasad, "X-Ray Baggage Inspection With Computer Vision: A Survey," in *IEEE Access*, vol. 8, pp. 145620-145633, 2020, doi: 10.1109/ACCESS.2020.3015014.
- [17] F. Alrowais et al., "Modified Earthworm Optimization With Deep Learning Assisted Emotion Recognition for Human Computer Interface," in *IEEE Access*, vol. 11, pp. 35089-35096, 2023, doi: 10.1109/ACCESS.2023.3264260.
- [18] C. Chambers et al., "Computer Vision to Automatically Assess Infant Neuromotor Risk," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 11, pp. 2431-2442, Nov. 2020, doi: 10.1109/TNSRE.2020.3029121.
- [19] S. Dey, A. K. Singh, D. K. Prasad and K. D. McDonald-Maier, "SoCodeCNN: Program Source Code for Visual CNN Classification Using Computer Vision Methodology," in *IEEE Access*, vol. 7, pp. 157158-157172, 2019, doi: 10.1109/ACCESS.2019.2949483.
- [20] W. Zhang, B. Cheng and Y. Lin, "Driver drowsiness recognition based on computer vision technology," in *Tsinghua Science and Technology*, vol. 17, no. 3, pp. 354-362, June 2012, doi: 10.1109/TST.2012.6216768.
- [21] A. Yousefzadeh, G. Orchard, T. Serrano-Gotarredona and B. Linares-Barranco, "Active Perception With Dynamic Vision Sensors. Minimum Saccades With Optimum Recognition," in *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, no. 4, pp. 927-939, Aug. 2018, doi: 10.1109/TBCAS.2018.2834428.
- [22] B. Tultayev, G. Balbayev, A. Zhauyt, A. Sultan, and A. Mussina, "Kinematic Synthesis of Mechanism for System with a Technical Vision," *MATEC Web of Conferences*, vol. 237, Nov. 2018, doi: 10.1051/mateconf/201823703009.
- [23] F. Zhao, X. Xie and M. Roach, "Computer Vision Techniques for Transcatheter Intervention," in *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 3, pp. 1-31, 2015, Art no. xxxxxx, doi: 10.1109/JTEHM.2015.2446988.
- [24] A. Osipov et al., "Identification and Classification of Mechanical Damage During Continuous Harvesting of Root Crops Using Computer Vision Methods," in *IEEE Access*, vol. 10, pp. 28885-28894, 2022, doi: 10.1109/ACCESS.2022.3157619.