

Eyes for Machines: A Computer Vision Approach to Enhance Robotic Arm Dexterity and Autonomy

Balkrishna Rasiklal Yadav^{1,*}

Abstract

By combining sophisticated robotics and visual awareness, computer vision operated robotic arms have revolutionized technology. These devices are having a profound effect on several industries, from manufacturing processes to healthcare. Computer vision–controlled robotic arms are altering the game with their ability to see, comprehend, and interact with their surroundings. In this study we have tried to develop and implement software and hardware to improve the freedom of movement in a robotic arm which will have a similar look to hand gestures. In the modern world, robotic arms are becoming essential, with uses in a wide range of sectors, including military, defense, healthcare, and industrial automation. These amazing devices can replicate the movements and hand gestures of a human, which makes them incredibly useful in a variety of scenarios. The most innovative automation technology available now is robotic systems. Initially, robots were employed on manufacturing floors in the 1960s and early 1970s. Evaluating the precision of system was done by methods such as sorting, kinematic modeling, and center recognition which elevate the movements of robotic arms. More innovation to elevate the free movement of robotic arms can enhance its applications at defense and industrial level.

Keywords: Robotics, defense, kinematic modeling, hand gestures, point to point trajectory

INTRODUCTION

Robotic arms have become indispensable in today's world, finding applications in many different industries like military, defense, medical surgeries, and industrial automation. These incredible machines possess the ability to mimic the gestures and actions of human hands, making them invaluable assets in various tasks and situations. Robotic systems are the most unique automation technology used today. In the 1960s and early 1970s, robots were initially used on factory floors. As robotic systems spread throughout the industrial sector, particularly in the automobile sector, they gained notoriety in the 1980s. Robotic arms are widely used in the industry, but most of them are used in a point to point (PTP) trajectory, the moves are learned previously by the robotic arm. Very few robots in the industry are programmed to be smart, or to make decisions. In the future to completely replace humans with robots, we need robotic arms which can make decisions [1–4].

*Author for Correspondence

Balkrishna Rasiklal Yadav
E-mail: bkrishnay@gmail.com

¹Researcher, Institute of Electrical and Electronics Engineers,
New Jersey, United States

Received Date: October 21, 2024

Accepted Date: October 24, 2024

Published Date: October 29, 2024

Citation: Balkrishna Rasiklal Yadav. Eyes for Machines: A Computer Vision Approach to Enhance Robotic Arm Dexterity and Autonomy. International Journal of Advanced Control and System Engineering. 2024; 2(2): 28–33p.

What Is a Computer Vision–Controlled Robotic Arm?

Before we jump into the intricacies of a computer vision–controlled robotic arm, let us first grasp the concept of computer vision. In simple terms, computer vision is a branch of study that aims to teach computers how to interpret and comprehend visual data, much like humans do with their eyes and brains.

Using advanced algorithms and cutting-edge machine learning techniques, computer vision

enables machines to process and analyze visual information from various sources such as images, videos, or even live camera feeds. All sorts of new opportunities for robots to "see" their environments and make smart judgements emerge from this.

LITERATURE SURVEY

Nowadays, robotics is becoming one of the most advanced fields of technology. A gesture-controlled robot is a unique type of robot that can be operated using human hands instead of traditional buttons and switches [5]. The primary objective of this project is to utilize hand gestures and computer vision to manipulate a robotic arm. The purpose of this device is to make it simple and cheap so that it can be produced and used for many purposes [6, 7]. An efficient, responsive tracking algorithm allows the hand to accurately track and respond to gestures. Gesture recognition and imitation is an efficient method used in situations that pose a threat to humans, such as radioactive environments, bomb disposal missions, cuts, epidemics, etc. [8] The robotic arm has been used in many applications in various industries. Applications have been received. These are designed to ensure maximum flexibility and utility for its planned application [9].

The hand gesture sensor in this paper utilizes the data from the accelerometer and gyroscope. A gyroscope is a sensor that captures the position of the operator's hand while working in a driven vehicle. It is typically attached to an arm. Experienced users can utilize the power of visual analysis to efficiently operate the system, while it may pose a slight challenge for those who are new to it. This system consists of two primary components: the ground station and the floor station. The ground station utilizes a sensor for recognizing hand gestures, while the floor station allows the user to control the robot's hands. Accelerators and gyroscopes are installed in the joints of the hands [10].

Many elements of building robots are described in John Iovin's book. It deals with various types of weapons design, control techniques, vehicle design etc. [11]. The *Uno Handbook* explores the complexities and methods of interconnecting many parts such as servomotors, electric motors, radio frequency receivers and radio frequency transmitters. Typically, individuals frequently use gestures that align with their verbal communication. Typically, these movements are known as conversational hand gestures that communicate semantic information alongside speech [12]. Sign language is a significant illustration of communicative gestures. Sign language follows a well-organized structure, enabling individuals who are deaf to effectively communicate with both hearing individuals and computer systems. Controlled gestures are commonly employed in various applications, including remote control systems for consumer electronics and robot control. Gesture-based interactions are used to engage with Digital things. Some great uses of manipulating gestures are digital fabrication and remote operation [1]. Rai et al. [5] provide an autonomous robotic framework for training, education, and employment. The suggested platform has the ability to execute and evaluate sophisticated algorithms for object handling and grasping, trajectory development, path planning, and other tasks, as well as to teach technical courses in subjects like robotics, control, electronics, image processing, and computer vision.

A robotic system can make use of several different types of sensors, such as infrared, temperature, color, mass, etc. However, there has been a recent shift towards using a single camera and advanced technology to replace these various sensors. With the help of advanced technology, computers are now able to understand and interpret the world around them. They can identify and analyse objects using various methods like machine learning and contouring. Control systems that utilize computer vision allow robotic arm systems to perform sophisticated operations including filtering, pick-and-place and sorting without requiring person involvement. Several systems, like the ones mentioned earlier, are included in the systems explained by Rai et al. [5]. Currently, robots have the capability to execute intricate tasks, but they still face limitations when it comes to quickly adjusting to familiar objects or environments without explicit programming. Thus, the concept of developing systems to enhance robots' perceptual abilities and object grasping recognition is a significant area of research. Prior to interacting with objects in the environment, a robot usually needs to physically grasp and elevate them.

Robots employ advanced technology such as laser distance sensors, stereoscopic vision, and a 3D camera to address this challenge.

Given the nature of the system, it involves the transmission of data between multiple points without the need for physical connections. The prevailing wireless technology employed in this scenario is radio. Wireless transmission needs to be dependable and have quick response time. There are several wireless transmission options available, including Bluetooth, infrared, Zigbee, and Wi-Fi technologies. Zigbee technology is favored for its extensive control capabilities and cost-effectiveness. Hsu et al. [6] introduced an innovative method for JACO, an assistive machine, to help with its limbs. This algorithm utilizes an affordable 3D depth camera and an enhanced reverse kinematic method to facilitate autonomous or semi-autonomous JACO functioning [13].

For use in classrooms, workplaces, and other training environments, Rai et al. [5] suggested an autonomous robotic framework. While sitting a robotic hand in the middle of the goal, they have two cameras that offered top as well as side viewpoints to study items of varying heights.

METHODOLOGY

A robotic arm controlled by computer vision utilizes cameras to capture visual data, which is then processed through advanced image recognition algorithms. The block diagram as shown in Figure 1 consists of three primary components: image acquisition (cameras), image processing (algorithms that extract relevant information), and actuation (robotic arm movement based on processed data). The cameras capture the surroundings, the algorithms analyse the visual data, and the robotic arm carries out actions accordingly. This integration improves the robot's capability to complete tasks in changing environments with immediate responsiveness.

Motor Configuration

For precise movement and control, a robotic arm configuration incorporates four servo motors that are strategically positioned. The base servo motor facilitates rotational motion, enabling the limb to perform a horizontal pivot. The shoulder servo, which is linked to the base, regulates vertical motion by raising and lowering the arm. The elbow servo, situated at the midline of the arm, improves flexibility by enabling flexion and extension. A gripper actuator is located at the end effector and controls the grasping mechanism, allowing the arm to precisely grasp and release objects. The utilization of a four-servo configuration guarantees a dynamic range of motion, which enables the robotic arm to execute tasks with versatility, efficiency, and precision.

KINEMATIC MODELING

To ensure proper functioning of the system, it is necessary to translate the position of the chosen items in the workspace of the robotic limb into angular positions for the limb. Firstly, the algorithm determines the exact position of the object being targeted via a specific formula.

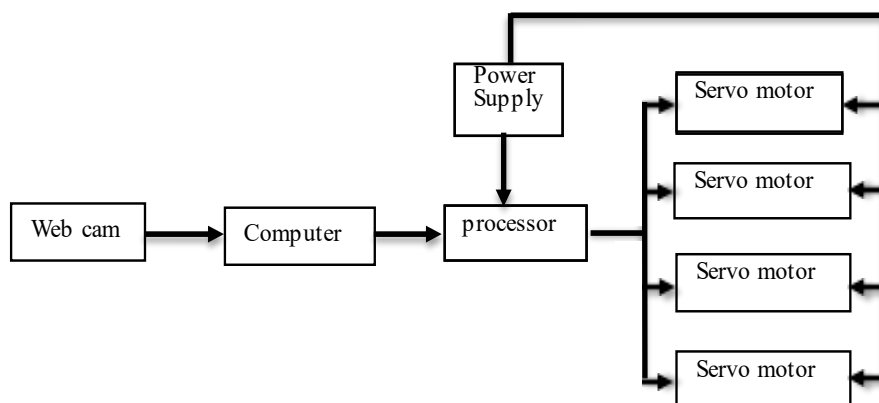


Figure 1. Block diagram of proposed concept.

Assume that there are no points in the contour. To calculate the center of a point, we use the formula $(C_x, C_y) = (x_i, y_i)$.

$$C_x = \frac{1}{n} \sum_{i=1}^n x_i$$

$$C_y = \frac{1}{n} \sum_{i=1}^n y_i$$

At this time, the positions of a points are specified in position of a points in a pixel; the unitary technique is employed to convert them to positions of a point that are Cartesian. Subsequently, joint position of a points for the robotic arm are determined using inverse kinematic modeling of the entirety of the system. These positions of a point are utilized to position the last effector's point at the center of the target object. The system's kinematic model is illustrated in Figure 2.

S1, S2, and S3 represent the joint positions of the robotic limb's limb servos, base, and elbow, respectively. The concatenated joint positions are transmitted to the processor that regulates the arm of the robot to execute collection and deliver operations and generate the intended motion.

The subsequent segment elucidates the outcomes and results acquired through the system's implementation. This consists of both the constructed physical device and the user interface.

In addition, we conducted a trial on the device to determine its effectiveness in performing the task for which it was built. This entails evaluating the precision of the system in relation to the three methods of sorting, kinematic modeling, and center recognition. We concluded the following part by comparing the outcomes and layout to those of prior endeavors.

Flowchart of the Code

In Figure 3a and 3b images, you can see the ways that we are utilizing to program the whole system:

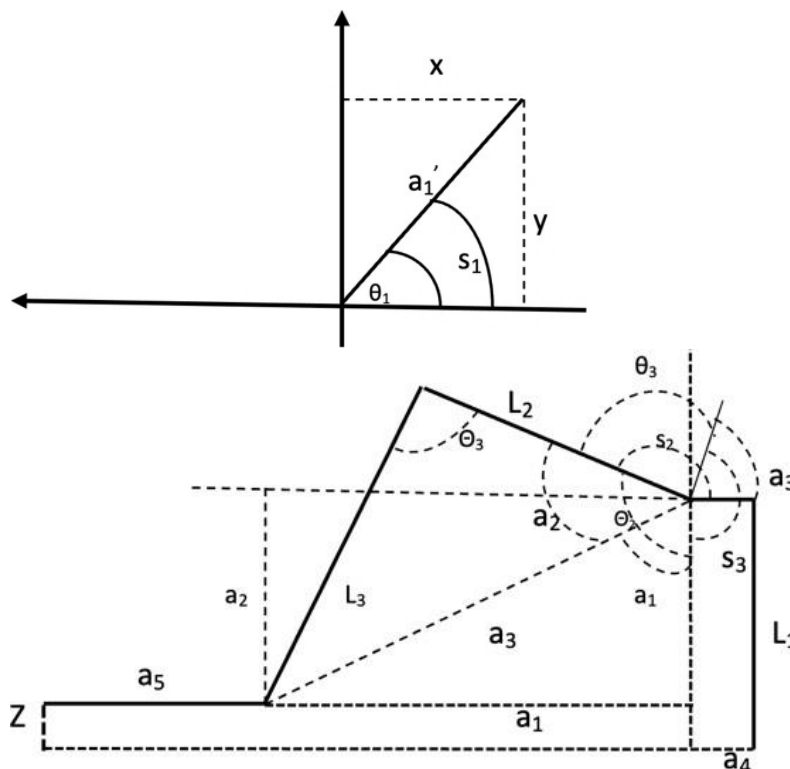


Figure 2. Side view and top view of the kinematic diagram.

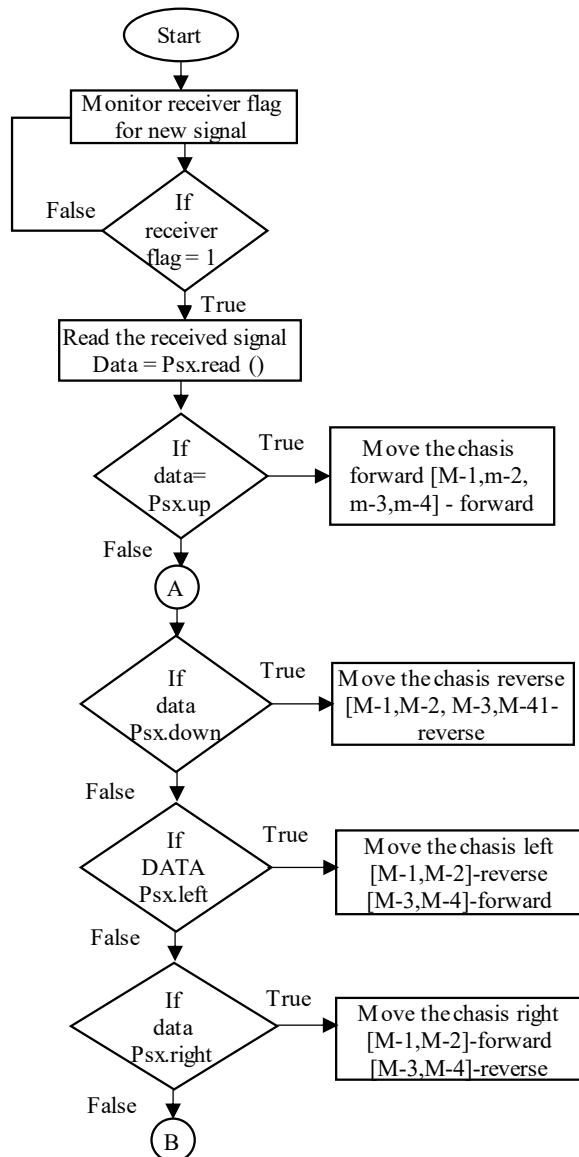


Figure 3. Block forms of utilizing the program to the whole system in two ways.

CONCLUSION AND FUTURE SCOPE

Computer vision–controlled robotic arms have taken technology to new heights, blending the power of visual perception with advanced robotics. From transforming manufacturing processes to revolutionizing healthcare, these machines are making a significant impact in various industries. With their ability to see, understand, and interact with the world around them, computer vision controlled robotic arms are truly changing the game. So, get ready to be amazed by the endless possibilities that lie ahead in this captivating field of robotics.

The ever-improving field of engineering and robotics will continue to grow because of our desire and the interesting possibilities it brings to its inventors and creators. Every person dreams of building a machine that can do the same job, or even more. Our project is a robotic arm that mimics hand gestures. It is a simple project so what we have improved is freedom of movement. We mainly focus on military and medical fields.

Developing the necessary software and the necessary hardware for a robotic limb that relies on gestures was the primary objective of the project. Features and capabilities of this branch will include

sensor capabilities, intelligence, mobility and navigation, universal gripper, and systems integration and networking.

REFERNNCES

1. Cabré TP, Cairol MT, Calafell DF, Ribes MT, Roca JP. Project-based learning example: controlling an educational robotic arm with computer vision. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*. 2013; 8 (3): 135–142. doi: 10.1109/RITA.2013.2273114.
2. Gomathy CK, Niteesh G, Sai Krishna K. The gesture controlled the robot. *Int Res J Eng Technol*. 2021; 8 (4): 1721–1725.
3. Dheeban SS, Harish DV, Harivignesh A, Prasanna M, Senthil Kumar N. Gesture controlled robotic arm. *Int J Sci Technol*. 2016; 4 (3): 101–112.
4. Patil C, Sharma S, Singh S. Design and implementation of gesture controlled robot with a robotic arm. *Int Res J Eng Technol*. 2019; 6 (9): 1351–1356.
5. Rai N, Rai B, Rai P. Computer vision approach for controlling educational robotic arm based on object properties. In: *Proceedings of the 2014 2nd International Conference on Emerging Technology Trends in Electronics, Communication and Networking*, Surat, India. December 26–27, 2014. pp. 1–9.
6. Hsu YH, Hsu H-Y, Lin J-S. Control design and implementation of intelligent vehicles with robot arm and computer vision. In: *Proceedings of the 2015 International Conference on Advanced Robotics and Intelligent Systems (ARIS)*, Taipei, Taiwan, May 29–31, 2015. pp. 1–6.
7. Chen X, Huang X, Wang Y, Gao X. Combination of augmented reality-based brain-computer interface and computer vision for high-level control of a robotic arm. *IEEE Trans Neural Syst Rehabil Eng*. 2020; 28: 3140–3147. doi: 10.1109/TNSRE.2020.3038209.
8. Yu F, Xian W, Chen Y, Liu F, Liao M, Madhavan V, Darrell T. BDD100K: a diverse driving video database with scalable annotation tooling. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, June 13–19, 2020. pp. 2633–2642.
9. Pinto L, Gupta A. Supersizing self-supervision: Learning to grasp from 50K tries and 700 robot hours. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*, Stockholm, Sweden, May 16–21, 2016. pp. 3406–3413. doi: 10.1109/ICRA.2016.7487517.
10. Luo W, Sun P, Zhong F, Liu W, Zhang T, Wang Y. End-to-end active object tracking via reinforcement learning. *Proc Mach Learn Res*. 2018; 80: 3286–3295.
11. Iovin J. *PIC Microcontroller Project*. New York: McGraw-Hill; 2000.
12. Levine S, Pastor P, Krizhevsky A, Ibarz J, Quillen D. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *Int J Robotics Res*. 2018; 37 (4–5): 421–436.
13. Geiger A, Lenz P, Stiller C, Urtasun R. Vision meets robotics: The kitti dataset. *Int J Robotics Res*. 2013; 32 (11): 1231–1237.