

Intelligent robotic control system based on computer vision technology

Chang Che^{1,6,*†}, Haotian Zheng^{2,7,†}, Zengyi Huang^{3,8}, Wei Jiang^{4,9}, Bo Liu^{5,10}

¹Mechanical Engineering, The George Washington University, DC, USA

²Computer Engineering, New York University, New York, NY, USA

³Applied Economics, The George Washington University, DC, USA

⁴Computer Science, Xidian University, Xian, China

⁵Software Engineering, Zhejiang University, Hangzhou China

⁶cche57@gwmail.gwu.edu

⁷hz2687@nyu.edu

⁸zengyihuang@gwmail.gwu.edu

⁹jiangwei13@gmail.com

¹⁰lubyliu45@gmail.com

*corresponding author

†Chang Che and Haotian Zheng contributed equally to this work and should be considered as co-first authors.

Abstract. Computer vision is a kind of simulation of biological vision using computers and related equipment. It is an important part of the field of artificial intelligence. Its research goal is to make computers have the ability to recognize three-dimensional environmental information through two-dimensional images. Computer vision is based on image processing technology, signal processing technology, probability statistical analysis, computational geometry, neural network, machine learning theory and computer information processing technology, through computer analysis and processing of visual information. The article explores the intersection of computer vision technology and robotic control, highlighting its importance in various fields such as industrial automation, healthcare, and environmental protection. Computer vision technology, which simulates human visual observation, plays a crucial role in enabling robots to perceive and understand their surroundings, leading to advancements in tasks like autonomous navigation, object recognition, and waste management. By integrating computer vision with robot control, robots gain the ability to interact intelligently with their environment, improving efficiency, quality, and environmental sustainability. The article also discusses methodologies for developing intelligent garbage sorting robots, emphasizing the application of computer vision image recognition, feature extraction, and reinforcement learning techniques. Overall, the integration of computer vision technology with robot control holds promise for enhancing human-computer interaction, intelligent manufacturing, and environmental protection efforts.

Keywords: Computer Vision, Robotic Control, Industrial Automation, Environmental Protection.

1. Introduction

In recent years, the integration of computer vision technology with robot control and interaction has opened up new avenues of research and development. Researchers are gradually realizing that vision is very important to people, [1] and more than 90% of human information is based on the eyes. Let's look at the frontier of artificial intelligence - machine vision. Robotic vision, a rapidly growing branch of AI (artificial intelligence) that aims to give robots vision comparable to our own, has made huge strides in the past few years thanks to researchers applying specialized neural networks to help robots recognize and understand images from the real world. For example, in manufacturing environments, [2] robots equipped with computer vision can accurately identify and manipulate objects on assembly lines, increasing efficiency and productivity. In healthcare, robots equipped with computer vision can assist medical professionals with tasks such as surgery and patient care, increasing accuracy and reducing the risk of errors. In this article, we will explore the intersection of computer vision technology and robotic control, examining the latest advances, applications, and challenges in this rapidly evolving field. We will delve into how computer vision algorithms are integrated into robotic systems, the impact and advantages of such integration for specific industries, and the prospects of computer vision technology in the field of machine vision interaction and control.

2. Related work

2.1. Computer Vision

Computer vision tasks depend on image features (image information), and the quality of image features largely determines the performance of the vision system. Traditional methods usually use SIFT, HOG and other algorithms to extract image features, and then use SVM and other machine learning algorithms to further process these features to solve visual tasks. Pedestrian detection is to determine whether there are pedestrians in the image or video sequence and give accurate positioning. The earliest method is HOG feature extraction + SVM classifier, and the detection process is as follows:

- The sliding window is used to traverse the whole image, and the candidate region is obtained
- Extract the HOG feature of the candidate region
- Classification of feature maps using SVM classifiers (to determine whether it is human)
- Using the sliding window will appear duplicate areas, using NMS(non-maximum) to filter the duplicate areas

The application of machine vision in robot interaction and control is very broad, it can help robots perceive and understand the surrounding environment, so as to perform tasks and interact with humans more effectively. Below I will explain how the main ideas of machine vision can be combined with robot interaction and control.

Image processing stage: In the robot vision system, the image processing stage is very important. [3] The robot acquires image data from the environment through cameras or other sensors.

In the process of image processing, the robot uses algorithms and technologies to preprocess the image, such as removing noise, enhancing contrast, edge detection, etc., in order to better extract the features of the object.

1. Object feature screening: In the process of image processing, the robot needs to recognize and extract various physical characteristics of the object, such as shape, color, texture, etc. The screening process also involves removing disturbing factors from the image, making the final feature more realistic and effective.

2. Image recognition stage: After image processing, the robot performs the image recognition phase. This phase involves matching and analyzing the features of the screened object with data from a pre-established database. By identifying and matching the features of the object, the robot can determine the identity, location and state of the object.

3. Robot interaction and control: Once a robot recognizes objects in its surroundings, it can perform various tasks and interactive behaviors based on this information. For example, robots can navigate, grasp, control devices, and communicate with humans based on the objects they recognize.

4. Real-time feedback and adjustment: During the execution of tasks, the robot may need to constantly make real-time feedback and adjustments. This can be achieved by continuously acquiring image data and processing it.

In summary, [4]with the continuous development of intelligent robots in the future, people's requirements for human-computer interaction are becoming higher and higher, intelligent, smooth, and anthropomorphic, which deeply test our ability to apply various machine modules. Here we discuss the human-machine interaction between machine vision and artificial intelligence from three aspects: robot vision artificial intelligence and robot control.

2.2. Robot visual interaction and control

Machine vision encompasses various technologies, including image processing, lighting control, optical sensors, and computer software, aimed at enhancing production flexibility and automation. It finds applications where manual labor may be impractical, substituting it with automated processes driven by machine vision. Integrating machine vision with robotics and artificial intelligence (AI) is crucial for advancing human-computer interaction capabilities. The ultimate goal of machine vision is to mimic human recognition abilities. Technological advancements, including recognition algorithms and hardware support like light sensors and image processing units, enhance machine vision's capability to recognize objects and faces intelligently. By optimizing algorithms and increasing recognition efficiency, machine vision systems become more adept at identifying objects in real-world scenarios. Real-world applications of machine vision technology bring convenience and efficiency to people's lives. For instance, in manufacturing, automation driven by machine vision reduces production time and enhances product quality. In logistics, automated systems powered by machine vision streamline operations, ensuring timely and accurate delivery of goods. Moreover, in healthcare, robot vision aids medical professionals in delivering precise treatments and diagnoses, ultimately improving patient outcomes.

2.3. Robot visual interaction application

Human-computer interaction (HCI) [5]has become ubiquitous in various aspects of daily life, from simple radio buttons to complex control systems in critical environments like nuclear reactors. With the rapid advancement of technology, machine vision has emerged as a crucial element in HCI, offering new possibilities and dimensions to the interaction process. However, these approaches often lack sophistication, leading to limitations in HCI. The advent of intelligent robots has transformed this landscape by integrating built-in vision sensors, robotic frameworks, and rudimentary logical reasoning capabilities. These robots can now engage in meaningful communication and interaction with humans. For instance, systems like AlphaGo leverage [6]machine vision, artificial intelligence, and deep learning to facilitate HCI with human players, showcasing the potential of advanced technologies in enabling seamless human-robot interaction. Despite the fragmented integration between machine vision and robot control in traditional production processes, ongoing advancements in science and technology are driving improvements in HCI capabilities. The emergence of intelligent robots equipped with visual sensors and mechanical frameworks signifies a shift towards robots possessing rudimentary logical thinking abilities and the capacity for human-like interactions. The integration of machine vision and artificial intelligence in intelligent robots holds significant promise for enhancing HCI applications. By leveraging advanced technologies, these robots offer greater convenience and efficiency in various aspects of human life, while driving ongoing progress and innovation in HCI technology.

3. Methodology

With the acceleration of global urbanization, urban waste disposal has become an increasingly serious problem. Traditional garbage sorting methods usually rely on manual sorting, which is inefficient and error-prone. To further explore the intersection of computer vision technology and robot control, it is imperative to establish a comprehensive methodology that encompasses various aspects of research,

development, and application. Therefore, this part using robot vision technology to develop intelligent garbage sorting robots has become a new way to solve this problem

3.1. Application design

Above the garbage collection and compression bin of the garbage transfer station, an intelligent identification camera is equipped, which can recognize the license plate of the garbage truck with high precision, and carry out real-time visual detection and recognition of the dynamically dumped garbage. By using advanced algorithm deduction technology, the recognition results are compared with the graph database of garbage, so as to realize the high-precision detection and recognition of garbage. And through the early warning system to remind the relevant personnel to deal with. The robot can accurately identify various types of garbage in the garbage disposal station or household and other scenes, and place them respectively in the corresponding recycling bins, thereby improving the efficiency of garbage disposal and reducing environmental pollution.

3.2. Implementation principle

1) Computer vision image recognition, image recognition is an important research direction in the field of artificial intelligence and computer vision. Its development is inseparable from the progress of computer technology and the support of big data. [7]At present, image recognition has been widely used in face recognition, unmanned driving, intelligent security monitoring, medical image analysis, virtual reality and other fields. Computer vision algorithm is the core of image recognition, which uses mathematical and statistical methods to extract features from images, and finally realizes image classification, detection and recognition after model training and optimization.

2) Image recognition method

Feature extraction: Feature extraction is the primary task of computer vision algorithms. It extracts representative feature information, such as color, texture, shape, etc., by processing pixels in the image. Commonly used feature extraction methods include SIFT, SURF, HOG, etc.

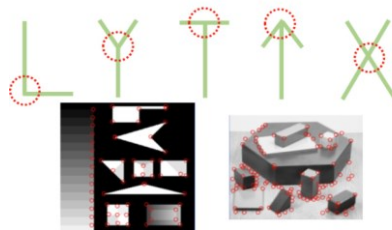


Figure 1. Principle of image feature extraction

Model training and optimization: On the basis of feature extraction, it is necessary to build a suitable model to realize image classification and recognition. Commonly used models include support vector machine (SVM), K-nearest neighbor (KNN), deep neural network (DNN), etc. The process of training the model is to learn and optimize through a large number of image samples, so that the model can accurately classify and recognize images.

Target detection and recognition: Target detection and recognition is one of the core tasks of image recognition. It is able to accurately find out the target object from the image and identify it according to the model established in advance. Target detection and recognition methods mainly include YOLO (You Only Look Once), Faster R-CNN, and SSD (Single Shot MultiBox Detector) based on deep learning.

Therefore, this application experiment is closely related to image recognition and feature extraction through the application of machine vision in intelligent garbage classification. First, image recognition technology enables the computer to identify the type of waste, such as paper, plastic, metal, etc., which helps the classification system understand the input image. [8]Second, feature extraction refers to the process of extracting useful information from images, such as color, texture, shape, etc. These features help the sorting system distinguish between different types of garbage. Therefore, the machine vision

system needs to combine image recognition technology to identify the type of garbage, and use feature extraction technology to extract the features in the image to achieve accurate garbage classification.

3.3. Robot operation and motion control

In a smart waste sorting experiment in which computer vision technology interacts with an intelligent robot, the researchers conducted experiments using a reinforcement learning system in robot classrooms. The researchers collected a large amount of trial data, both in the actual deployment environment and the simulation environment. By continuously adding data, they improved the performance of the entire system and evaluated the final system. The experimental results show that the system has achieved good results in handling garbage sorting tasks, with an average accuracy of about 84%, and the performance has steadily improved with the increase of data. Computer vision plays an important role in this system, providing critical information and feedback to the robot. Through computer vision, the robot is able to recognize the type and location of different garbage, so that it can effectively grasp and dispose of it. In addition, with RetinaGAN technology, the fidelity of the simulation images is improved, enabling the robot to better learn in the simulation environment and apply what it learns in the actual deployment environment. This method of integrating computer vision technology provides strong support for robots to perform tasks in complex environments, and also provides a good foundation for them to adapt and learn in real scenarios.

3.4. Robot motion control and real-time feedback

At present, automation is expected to drive up the total amount of waste recycling at a time when manual recycling is facing inefficiencies, and because of this, the market for garbage sorting robots is expected to reach \$12.26 billion by 2024 and continue to grow at a compound annual growth rate of 16.52%. The AMP Robotics garbage sorting robot has an unusually high sorting speed - 80 pieces per minute, much higher than manual pickup, and can automatically identify, process and monitor complex waste streams. It is modular in design, allowing managers to adapt existing workflows for a single class or different volumes of recyclable objects. AMP robots can pick out many kinds of garbage, including metal, batteries, capacitors, plastics, [9]PCBS, wires, cartons, cardboard, cups, shells, LIDS, aluminum, etc. It also includes metal blends of wood, asphalt, brick, concrete, plastics and mixtures (e.g., polyethylene terephthalate, high density polyethylene, low density polyethylene, polypropylene, polystyrene). It also classifies the film by color, clarity and transparency. AMP Robotics realizes the workflow of action control and real-time feedback for garbage classification as follows:

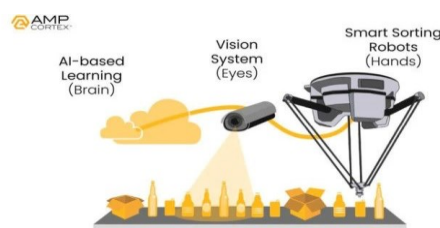


Figure 2. Robot garbage sorting principal architecture released by AMP company

As can be seen from Figure 2, in the process of garbage classification, the robot carries out the classification business through the following processes combined with computer vision technology:

1. Garbage delivery and delivery: Garbage is dropped by the customer into the garbage collection scene, and then transported to the robot's work area by conveyor belt or other means. 2. Visual recognition and localization: AMP Robotics' robots are equipped with an advanced visual recognition system to quickly and accurately identify various objects in the garbage stream. Once the target object is identified, the robot immediately locates its position. 3. Motion control and grasping: According to the visual recognition and positioning results, the robot quickly adjusts its posture and executes corresponding actions through its own motion control system to accurately grasp the target object.

The training data of AMP Robotics comes from AMP Neuron, and the built-in algorithm of the robot trains through massive images. The image database contains the various states in which the recyclables are intact, indentation or crushed. This allows robots to classify more accurately, learn new material classes more efficiently, and better adapt to packaging design and lighting changes. Sensors use computer vision to scan fast-moving objects, distinguishing them based on color, texture, shape, size, and material. [10]The robot's attachments then use suction cups to grab items from the conveyor belt and drop them into the corresponding recycling bin.

4. Conclusion

In general, this paper discusses the intersection of computer vision technology and robot control in depth, emphasising the importance and potential of this field in human-computer interaction, intelligent manufacturing and environmental protection. This paper first introduces the definition and development of computer vision technology, points out its important role in simulating human visual observation and image analysis, and emphasises its importance in human information acquisition. Then, the paper introduces the integration of computer vision technology and robot control in detail, as well as the influence and application of this integration in various fields. Especially in the fields of industrial automation, healthcare and environmental protection, this integration has brought great progress and benefits, improving production efficiency, quality and environmental protection.

References

- [1] Liu, Bo, et al. "Integration and Performance Analysis of Artificial Intelligence and Computer Vision Based on Deep Learning Algorithms." *arXiv preprint arXiv:2312.12872* (2023).
- [2] Golnabi, Hossein, and A. Asadpour. "Design and application of industrial machine vision systems." *Robotics and Computer-Integrated Manufacturing* 23.6 (2007): 630-637.
- [3] Hu, Hao, et al. "Casting product image data for quality inspection with xception and data augmentation." *Journal of Theory and Practice of Engineering Science* 3.10 (2023): 42-46.
- [4] Yang, Le, et al. "AI-Driven Anonymization: Protecting Personal Data Privacy While Leveraging Machine Learning." *arXiv preprint arXiv:2402.17191* (2024).
- [5] Yu, Liqiang, et al. "Semantic Similarity Matching for Patent Documents Using Ensemble BERT-related Model and Novel Text Processing Method." *arXiv preprint arXiv:2401.06782* (2024).
- [6] Huang, Jiaxin, et al. "Enhancing Essay Scoring with Adversarial Weights Perturbation and Metric-specific AttentionPooling." *arXiv preprint arXiv:2401.05433* (2024).
- [7] Zhang, Chenwei, et al. "Enhanced User Interaction in Operating Systems through Machine Learning Language Models." *arXiv preprint arXiv:2403.00806* (2024).
- [8] Lin, Qunwei, et al. "A Comprehensive Study on Early Alzheimer's Disease Detection through Advanced Machine Learning Techniques on MRI Data." *Academic Journal of Science and Technology* 8.1 (2023): 281-285.
- [9] Che, Chang, et al. "Advancing Cancer Document Classification with Random Forest." *Academic Journal of Science and Technology* 8.1 (2023): 278-280.
- [10] Snyder, Wesley E., and Hairong Qi. *Machine vision*. Vol. 1. Cambridge University Press, 2004.
- [11] Ang, Fitzwatler, et al. "Automated waste sorter with mobile robot delivery waste system." De La Salle University Research Congress. Manila, Philippines: De La Salle University, 2013.
- [12] Che, Chang, et al. "Enhancing Multimodal Understanding with CLIP-Based Image-to-Text Transformation." *Proceedings of the 2023 6th International Conference on Big Data Technologies*. 2023.
- [13] Li, Chen, et al. "Enhancing Multi-Hop Knowledge Graph Reasoning through Reward Shaping Techniques." *arXiv preprint arXiv:2403.05801* (2024).
- [14] Carroll, John M. "Human-computer interaction: psychology as a science of design." *Annual review of psychology* 48.1 (1997): 61-83.