Research on Target Detection Method for Intelligent Mobile Robots based on Machine Visionr

Baitong Hou *

Southampton Ocean Engineering Union College, Harbin Engineering University, Harbin, Heilongjiang Province, 150000, China

* Corresponding Author Email: houbaitong@gmail.com

Abstract. Intelligent mobile robots use machine vision and deep learning to achieve environmental perception, path planning, and obstacle avoidance, effectively improving autonomous decision-making capabilities. In this paper, you only look once (YOLO) and Faster Region-based Convolutional Neural Network (Faster R-CNN), Region of Interest Detection (ROB), and Direct Sparse Odometry (DSO) are compared and analyzed. The advantages and disadvantages of two Simultaneous Localization and Mapping (SLAM) algorithms, single-sensor and multi-sensor fusion are compared so that different object detection methods can be used in different application scenarios. In this paper, it is concluded that algorithms such as YOLO and Faster R-CNN meet the requirements of real-time and accuracy, respectively, visual SLAM supports autonomous navigation, and multi-sensor fusion enhances adaptability and stability. Current challenges include balancing real-time performance and accuracy, improving multi-target recognition capabilities in complex scenarios, and maintaining stability in harsh environments. Future research will combine deep learning, edge computing, and 5G technologies to further promote the intelligent and efficient application of machine vision in the field of robotics.

Keywords: Machine Vision, Object Detection Methods, Visual SLAM Technology.

1. Introduction

With the rapid development of science and technology, intelligent robots are gradually replacing manual services and becoming an emerging technology in many fields such as medical care, logistics, transportation, aerospace, etc [1]. As its key technology, machine vision is mainly responsible for object recognition, path planning and obstacle avoidance, and realizes environmental perception and autonomous decision-making through image processing and deep learning.

In recent years, robotics technology has developed rapidly, and the scale of the global robotics industry has increased rapidly [2]. At present, the research direction is mainly focused on efficient navigation, the combination of deep learning and image processing, multi-sensor information fusion and hardware optimization. With the rapid development of science and technology, the application of intelligent robots in the fields of medical care, logistics and transportation has been continuously expanded, and one of the key technologies is machine vision. Machine vision realizes environmental perception and target recognition through the analysis and processing of image data, and is one of the core technologies for intelligent robot navigation and decision-making [3]. By collecting environmental images through cameras and other devices, combined with image processing and deep learning algorithms, machine vision helps robots accurately identify and understand the surrounding environment, thereby achieving path planning, obstacle avoidance, and operation in dynamic environments. Its coordinated use with other sensors further improves the robot's ability to perceive and locate the environment, enabling it to make accurate judgments under various complex conditions, such as low light or bad weather. In addition, with the combination of machine vision technology and deep learning, robots can autonomously learn and optimize their behavior, and continuously improve their intelligence level. Therefore, machine vision greatly enhances the efficiency and stability of robots in the process of navigation, perception, and decision-making.

This paper aims to study the target detection method of machine vision. It first explains the concept of machine vision, then lists two target detection and recognition algorithms, YOLO and Faster R-CNN, and then introduces the meaning and application of the SLAM algorithm. Finally, it points out the shortcomings of a single sensor and proposes the concept of multi-sensor fusion. The key technologies of machine vision in intelligent mobile robots are analyzed.

2. Machine Vision Technology

Machine vision technology is a technology that processes and analyzes images or videos through computer systems to achieve object recognition, detection, measurement, and discrimination. In the field of mechanical manufacturing, it is mainly used for quality inspection, part positioning, product identification and automatic navigation [4].This technology not only enables the robot to identify objects around it, but also helps it complete path planning and obstacle avoidance operations through complex algorithms. Therefore, machine vision directly affects the working accuracy and efficiency of intelligent robots.

2.1. Object Detection and Recognition

Object detection and recognition is one of the important links in machine vision technology. It helps robots identify and locate different objects in the environment, thereby supporting their further operations. Common object detection algorithms today include YOLO and Faster R-CNN [5].

YOLO and Faster R-CNN each have their own advantages and disadvantages: YOLO is suitable for scenarios with high real-time requirements [6]. Such as real-time monitoring and unmanned driving, while Faster R-CNN performs better in detection accuracy [7]. It is suitable for scenarios such as medical image analysis. In complex applications, the advantages of both can be combined, by first using YOLO to quickly locate the target and then using Faster R-CNN to improve the detection accuracy, thus achieving a balance between real-time performance and accuracy.

With the help of these algorithms, robots can quickly and accurately identify and classify objects in the surrounding environment, providing reliable data support for subsequent tasks. This recognition and detection capability not only improves the robot's autonomy but also lays the foundation for its multi-task execution in complex environments. With the continuous advancement of technology, the performance of target detection algorithms will be further optimized, bringing greater breakthroughs in the field of machine vision.

2.2. Visual SLAM Technology

SLAM is the core technology for intelligent mobile robots to achieve simultaneous localization and mapping in unknown environments. Visual SLAM extracts and matches features from image data captured by cameras, thereby achieving accurate environmental localization and map generation [8]. Currently, common visual SLAM algorithms include ORB-SLAM and DSO. ORB-SLAM uses ORB features for matching and localization, has good real-time robustness, and is suitable for applications in dynamic environments. DSO is a direct SLAM algorithm that directly estimates pose changes through pixel grayscale values without relying on feature point matching, and has great advantages in accuracy and flexibility. Visual SLAM provides strong support for the autonomous navigation of robots in unknown environments.

2.3. Multi-sensor fusion

A single sensor is a device that independently senses physical quantities, such as temperature sensors, pressure sensors, displacement sensors, etc., and is used to obtain specific information. However, its limitations lie in its single sensing range and single function, which makes it difficult to meet the needs of complex environments or multiple parameters. In order to improve perception accuracy and robustness, machine vision is often combined with other sensors [9], such as lidar and ultrasonic sensors. Visual sensors may have limitations in some cases (such as when the light changes dramatically or the light is insufficient). Lidar can provide high-precision distance data to make up for the lack of visual information. By fusing vision with data from other sensors, robots can obtain more comprehensive information under different environmental conditions, improving perception capabilities and decision-making efficiency. This multi-sensor fusion technology has been widely used in fields such as autonomous driving, industrial robots, and drones, showing great application potential.

2.4. Technology Application

Machine vision-based target detection, navigation, obstacle avoidance and SLAM technologies play an important role in the practical application of intelligent mobile robots. These technologies have been widely used in fields such as autonomous driving, intelligent logistics, security monitoring and industrial automation. For example, in autonomous vehicles, machine vision is combined with multiple sensors such as lidar and millimeter wave radar to support vehicles to achieve target recognition and lane-keeping functions in complex road environments [10]. In intelligent warehousing systems, visual SLAM technology helps logistics robots to navigate autonomously and operate efficiently in dynamic warehousing environments. This greatly improves the efficiency, accuracy, and safety of logistics [11]. In industrial automation, machine vision technology realizes the intelligence and automation of production processes through the guidance and control of automatic sorting systems and robotic arms [12]. With the continuous development of technologies such as deep learning, edge computing and 5G, the application of machine vision technology in intelligent mobile robots will be further expanded and improved. Machine vision technology will face the problem of insufficient development. Smaller-scale technical pilot projects should be adopted to gradually improve the technology while strengthening technology research and development to achieve technological breakthroughs.

3. Discussion

The application of machine vision in intelligent mobile robots is a reflection of the deep integration of modern technology and automation. By collecting environmental information through visual sensors such as cameras and combining them with computer vision algorithms, machine vision gives robots the ability to understand and analyze the external environment. This process involves multiple key technologies, including target detection and recognition, simultaneous localization and mapping (SLAM), path planning, and obstacle avoidance, enabling robots to autonomously navigate and perform precise operations in complex and dynamic environments.

Object detection and recognition are core functions in machine vision. With algorithms such as YOLO and Faster R-CNN, robots can quickly and accurately identify and locate specific targets. By using algorithms such as YOLO with high real-time performance or Faster R-CNN with high accuracy, robots can perform different tasks in a variety of environments, supporting the accuracy of path planning and decision-making.

SLAM technology (Simultaneous Localization and Mapping), especially visual SLAM, helps robots achieve self-localization and map building in unknown environments. Visual SLAM algorithms such as ORB-SLAM and DSO enable robots to achieve autonomous navigation and precise positioning through image feature extraction and pose estimation. ORB-SLAM has good real-time and robustness in dynamic environments, while DSO estimates pose through pixel grayscale, getting rid of the dependence on feature points and enhancing accuracy and flexibility. These technologies provide reliable support for the autonomous actions of robots in unknown environments and play an important role in applications such as autonomous driving and smart logistics.

In addition, multi-sensor fusion technology further improves the perception accuracy and robustness of machine vision. In the case of unstable or insufficient lighting, visual sensors may be limited, so they are often combined with lidar, millimeter-wave radar, ultrasonic sensors, etc. to achieve multi-layer data complementarity. Multi-sensor fusion is widely used in fields such as autonomous driving and industrial automation, providing a guarantee for robots to obtain comprehensive and reliable information under different environmental conditions.

The diversity of technology applications demonstrates the value of machine vision in various practical scenarios. For example, in autonomous driving, machine vision is integrated with multiple sensors to help vehicles achieve functions such as target detection, lane keeping, and pedestrian avoidance; in smart warehousing systems, visual SLAM technology supports autonomous navigation of logistics robots in dynamic warehousing environments, improving operational efficiency. With the development of deep learning, edge computing, and 5G, the application scenarios of machine vision in intelligent robots will become richer, and its performance will be further optimized.

4. Conclusion

This paper deeply analyzes the key application technologies of machine vision in intelligent mobile robots, including target detection, visual SLAM, multi-sensor fusion, etc., and discusses its application and effects in actual scenarios. As the core technology to enhance the robot's autonomous navigation, environmental perception and decision-making capabilities, machine vision enables it to accurately plan paths and avoid obstacles in dynamic and complex environments. Target detection algorithms such as YOLO and Faster R-CNN have their own characteristics, which meet the requirements of real-time performance and recognition accuracy in different scenarios, while visual SLAM technology provides effective autonomous navigation support for robots through environmental positioning and map construction. In addition, the multi-sensor fusion strategy effectively enhances the accuracy and stability of machine vision and broadens its adaptability in various environments.

Although machine vision technology has made remarkable achievements in the field of intelligent robots, it still faces some challenges. For example, how to balance real-time and accuracy, improve the ability to recognize multiple targets in complex scenes, and ensure stability in low-light or harsh environments are all issues to be solved. Future research can further optimize the combination of deep learning and visual algorithms, improve the efficiency of multimodal information fusion, and use edge computing and 5G technologies to enhance real-time processing performance.

In short, machine vision technology has injected a strong impetus into the development of intelligent mobile robots. With continuous breakthroughs in technology, machine vision will drive robots to achieve more intelligent and efficient automation tasks in more application fields in the future.

References

[1] WANG Shuping. Design of intelligent robot path planning system based on machine vision [J/OL]. Automation Technology and Application, 1-7 [2024 - 12 - 03]. http://kns.cnki.net/kcms/detail/23.1474.TP.20241203.1422.030.html.

[2] SUN Liangyan. Robotics Technology and Application, 2024 (05): 2 - 5.

[3] FENG Heping, ZHAO Boxin. Research on the application of machine vision technology in the field of intelligent manufacturing [J]. Communications World, 2024, 31 (10): 193 - 195.

[4] Alzigul Umar. Research on the innovative application of intelligent control technology in modern machinery manufacturing [J]. Agricultural Machinery Use and Maintenance, 2024 (11): 92 - 95.

[5] Lu Quan, He Jiasheng, Yin Linfei. Transmission tower state detection algorithm based on improved YOLOv7 [J/OL]. Computer Engineering: 1 - 21 [2024 - 12 - 05]. https://doi.org/10.19678/j.issn.1000 - 3428.0069372.

[6] Cao Jiwei, Luo Fei, Ding Weichao. BS-YOLO: Small object detection algorithm based on BSAM attention mechanism and SCConv [J/OL]. Computer Engineering: 1 - 10 [2024 - 12 - 05].

[7] Zhang Qihang, Zeng Qi. Unmanned target detection based on improved Faster-RCNN[J]. Computer and Digital Engineering, 2024, 52 (08): 2411 - 2415+2419.

[8] Zhao Boting, Liu Jun, Chai Hongxu, et al. A real-time visual SLAM algorithm for complex dynamic scenes [J/OL]. Journal of Beijing University of Aeronautics and Astronautics: 1 - 18 [2024 - 12 - 05].

[9] Chen Xiaofeng, Li Yufeng, Wang Chuansong, et al. Research on Unmanned Vehicle Target Detection System Based on Multi-sensor Fusion [J/OL]. Journal of Laser: 1 - 8 [2024 – 12 - 05]. http://kns.cnki.net/kcms/detail/50.1085.TN.20241106.1102.006.html.

[10] Wu Songbo, Yu Xinglin, Yang Guilin, et al. Application and challenges of machine vision in automotive safety system [J]. Automotive Knowledge, 2024, 24 (10):38 - 41.

[11] Ren Fang. Logistics Technology and Application, 2024, 29 (06): 82 - 83.

[12] Liu Zhenxing, Kou Shuang. Application of machine vision inspection technology in the field of industrial automation [J]. Papermaking Technology and Application, 2024, 52 (03): 46 - 48.