Research on the Mechanism and Path of Augmented Reality Empowering Robot 2D Image Perception

Benbo Cao
School of Waterford, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China

Abstract: This article focuses on the mechanisms and paths by which augmented reality empowers 2D image perception in robots. Firstly, the significant meaning of 2D image perception for robots and the current challenges they face are expounded, and then the necessity of empowering it with augmented reality technology is introduced. An in-depth analysis was conducted on the internal mechanisms by which augmented reality empowers 2D image perception in robots, including information fusion mechanisms, spatial mapping mechanisms, etc. On this basis, specific paths to achieve this empowerment were explored, covering paths such as technology integration and application scenario expansion. It aims to provide theoretical support and direction guidance for promoting the improvement of 2D image perception capabilities of robots and the wide application of augmented reality technology in the field of robotics.

Keywords: Augmented Reality; Robot; 2D Image Perception; Mechanism; Path.

1. Introduction

In today's era of exponential technological development, robotics is like a powerful wave of change and has become a key force driving profound transformation in various industries. From the precise and efficient automated production lines in the industrial manufacturing field to the precision robots assisting doctors in performing complex surgeries in the medical field; From intelligent handling robots that can quickly sort goods in logistics and warehousing scenarios to companion robots capable of completing various household tasks in the home service field, robots are everywhere, profoundly changing our production and lifestyle[1].

The 2D image perception ability of robots, as an important cornerstone for their effective interaction with the environment, plays a decisive role in core functions such as autonomous navigation, target recognition, and operation control of robots [2]. Precise 2D image perception is like the "eyes" of a robot, enabling it to quickly and accurately understand the information of its surrounding environment. During the process of autonomous navigation, robots can identify road signs, obstacles, etc. through 2D image perception, thereby planning a safe and reasonable driving path. In the task of target recognition, it can accurately determine the category, position and other information of the target object, providing a basis for subsequent operation control. Only by possessing efficient and accurate 2D image perception capabilities can robots make reasonable decisions and take actions in complex and ever-changing environments, and achieve various preset functions.

But now, robots still have many troublesome challenges in 2D image perception. In a complex environment, the problem of image interference is particularly obvious. In the real scene, the combination of light change, occlusion and complex background will greatly reduce the image quality [3]. For example, outdoors, strong direct light may lead to serious overexposure of the image, completely covering the details of the object; In the dark corner of the warehouse, the image may be underexposed, making the edge of the object blurred. Mutual occlusion between objects is also a common problem. When key information is blocked, robots can easily admit their mistakes. In addition, the robot's own perception ability is also limited. Traditional 2D image sensing methods usually only rely on a single image data, lacking multi-dimensional understanding of the environment [4]. This makes it difficult for robots to accurately distinguish those things that look alike but have different functions when facing some complex and changeable scenes, such as identifying objects with similar appearance but different uses only by 2D image features, which seriously limits the further improvement of robot performance and the expansion of application scenarios.

Augmented reality technology, as a cutting-edge technology that ingeniously integrates virtual information with the real world, has unique advantages and huge potential [5]. It can provide robots with rich virtual information supplements. By superimposing virtual signs, guiding information, etc. on the images of real scenes, it enhances the robots' understanding and cognitive abilities of the real environment. For instance, during the operation and control process of robots, augmented reality technology can display operation steps and key prompts in real time, helping robots complete operation tasks more accurately. Empowering 2D image perception of robots with augmented reality technology is expected to break through the existing bottlenecks and bring new opportunities and breakthroughs to the development of robot technology. Therefore, in-depth research on the mechanism and path of augmented reality empowering robots' 2D image perception holds significant theoretical value and practical significance. It not only promotes the further development of robot technology but also provides strong support for application innovation in various industries.

2. The Current Situation and Challenges of 2D Image Perception in Robots

2.1 The Importance of 2D Image Perception for Robots

2D image perception of robots is one of the important ways for robots to obtain environmental information. Through image acquisition devices such as cameras, robots can capture two-dimensional images of their surrounding environment, and then analyze and process these images to extract key information, such as the shape, color, and position of objects [6]. This information is crucial for the robot's tasks such as navigation and positioning, target recognition and grasping, and scene understanding. For instance, on industrial production lines, robots need to accurately identify the positions and postures of different components in order to perform precise assembly operations. In the field of logistics and warehousing, robots need to be capable of quickly locating target goods and planning the optimal route for handling. All of these are inseparable from efficient and accurate 2D image perception capabilities.

2.2 Current Challenges Faced

Although the robot's 2D image sensing technology has made some progress, there are still many problems in practical use. On the one hand, image interference in complex environment is very serious [7]. In real scenes, there are often various situations, such as light changes, things are blocked, and the background is messy. These factors will make the image quality worse, increase the difficulty of processing, and make it difficult for robots to find useful information accurately. For example, in direct strong light or very dark environment, the image may be too bright or too dark, which will affect the edges and details of the recognized object. Occlusion between objects will also cause some key information to be lost, resulting in recognition errors. On the other hand, the limitation of the robot's own perception ability also affects its application range [8]. Traditional 2D image recognition methods usually only rely on a single picture data, lacking a stereoscopic understanding of the environment, so it is difficult to cope with some complex and changeable scenes. For example, it is difficult for robots to accurately distinguish objects that look alike but have different functions only by 2D image features.

3. Overview of Augmented Reality Technology and Its Necessity in Empowering 2D Image Perception for Robots

3.1 Overview of Augmented Reality Technology

Augmented reality technology is a technology that can combine virtual information with real-world scenes in real time. It superimposes computer-generated virtual elements, such as pictures, characters and 3D models, on objects in the real environment, so that when we look at the real world, we can also see additional virtual information. Augmented reality technology is characterized by the combination of reality and reality, real-time interaction and three-dimensional registration. The combination of reality and reality is the core feature of augmented reality, which can seamlessly combine virtual information with real scenes and bring us a rich and immersive visual experience. Real-time interaction allows us to interact with these virtual information naturally and smoothly, and enhances our ability to perceive and control virtual information. Three-dimensional registration technology can ensure that virtual information can be accurately located in a specific location in the real scene, so that the effect of combining reality with reality will be accurate and stable.

3.2 The Necessity of Empowering Robots with 2D Image Perception through Augmented Reality

Given the challenges faced by robots in 2D image perception, the introduction of augmented reality technology is of great necessity. First of all, augmented reality technology can provide robots with rich virtual information supplements. By fusing pre-stored virtual models, annotation information, etc. with real-time collected 2D images, robots can obtain more detailed information about objects, such as the three-dimensional structure and functional attributes of the objects, thereby making up for the insufficiency of single 2D image information and improving the ability to recognize and understand objects. Secondly, augmented reality technology helps solve the problem of image interference in complex environments. Enhancing real images by using virtual information, such as increasing the brightness and contrast of images through virtual lighting models under low-light conditions, or virtually repairing occluded parts, can improve image quality and enhance the perceptual stability of robots in complex environments. In addition, the real-time interactivity of augmented reality technology can also provide a more convenient and efficient communication method between robots and human operators, enabling operators to convey instructions and information to robots in real time through virtual interfaces, further expanding the application scenarios and flexibility of robots.

4. The Intrinsic Mechanism by which Augmented Reality Empowers 2D Image Perception in Robots

4.1 Information Fusion Mechanism

Information fusion is one of the core mechanisms by which augmented reality empowers 2D image perception in robots. It acquires a more comprehensive and accurate environmental understanding by integrating and processing information from various sources. In the 2D image perception of robots, information fusion mainly involves the fusion of real image information and virtual information. The real image information is collected by devices such as cameras carried by the robot, including real-time visual features of the environment. Virtual information originates from pre-constructed model libraries, knowledge bases, etc., and is rich in semantic and structural information. By integrating these two types of information, robots can simultaneously obtain the appearance features and internal attributes of objects, thereby achieving more accurate recognition and understanding of the objects. For instance, when identifying a mechanical part, a robot can not only obtain the shape, color and other appearance information of the part through real images, but also learn detailed information such as the model, specification and assembly requirements of the part by fusing virtual information, which greatly improves the accuracy and practicality of recognition.

4.2 Spatial Mapping Mechanism

Spatial mapping mechanism plays a key role in allowing robots to see 2D images through AR. Its task is to accurately "paste" virtual information to a specific location in the real scene, and to ensure that the virtual and real things blend accurately and stably. In robot application, spatial mapping needs to solve two key problems: one is how to determine the coordinate position of virtual information in the real world, and the other is how to ensure that the spatial relationship between virtual information and real objects is correct. In order to achieve accurate spatial mapping, it is usually necessary to rely on various sensor technologies, such as visual sensor and inertial measurement unit (IMU), to obtain the position and attitude information of the robot itself and the 3D structural information of the environment. Using this information, together with computer vision algorithm and 3D registration technology, we can accurately superimpose the virtual model on the real object. In this way, the robot can perceive and operate those objects with virtual information in the correct spatial position. For example, when the robot is carrying out the assembly task, it can accurately display the virtual assembly guidance information in the corresponding position of the actual parts through a spatial mapping mechanism, and guide itself to complete the accurate assembly operation.

4.3 Interactive Feedback Mechanism

Interactive feedback mechanism is an important part of augmented reality to help robots understand 2D images. It enables two-way information exchange between robots and virtual information, as well as human operators. Through this mechanism, robots can not only perceive virtual information, but also adjust their behaviors and decisions according to the interaction results. On the one hand, robots can get more task-related content by interacting with virtual information, such as clicking virtual buttons to get operation instructions, or touching virtual objects to get item attribute information. On the other hand, the robot can also feed back its own state information and recognition results to the operator, so that the operator can know the working condition of the robot in time and make intervention and adjustment. For example, when the robot is remotely controlled to perform tasks, the operator can view the 2D images collected by the robot and the fused virtual-real scene in real time through the augmented reality interface. At the same time, the robot can also feed back the information of operation execution status and problems encountered to the operator to realize efficient man-machine cooperation.

5. The Implementation Path of Empowering Robots with 2D Image Perception through Augmented Reality

5.1 Technology Integration Path

Technology fusion is one of the key paths for robots to perceive 2D images through augmented reality technology. First of all, we should strengthen the deep integration of computer vision technology and augmented reality technology. Computer vision technology provides basic data processing and analysis capabilities for robot 2D image perception. By combining with augmented reality technology, the results of computer vision processing can be seamlessly integrated with virtual information, further improving the image perception effect. For example, after computer vision algorithm is used to detect and recognize the target, the recognition result is superimposed on the real image in the form of virtual annotation, so that the robot can understand the image content more intuitively. Secondly, promote the integration of sensor technology and augmented reality technology. The cooperation of various sensors can provide more comprehensive environmental information for the robot, for example, the depth sensor can obtain the distance information of the object, and IMU can provide the attitude information of the robot. Combining these sensor data with AR technology can make robots know the surrounding environment more clearly, and better integrate virtual things with the real world, so that they can see more clearly in complex places. In addition, we should also pay attention to the good cooperation between communication technology and AR technology, so as to ensure that the robot can transmit data with the cloud server and other devices in real time and quickly, and realize remote work and resource sharing.

5.2 Path for Expanding Application Scenarios

Expanding application scenarios is an important way to promote the development of two-dimensional image perception of augmented reality-empowered robots. With the continuous progress of technology, augmented reality-enabled 2D image perception technology for robots will be applied in more fields. In the industrial field, apart from traditional tasks such as assembly and inspection, it can also be applied to intelligent warehouse management, flexible production line scheduling and other aspects. For instance, by using augmented reality technology to provide real-time goods location information and navigation guidance for warehouse robots, efficient storage and handling of goods can be achieved. In flexible production lines, augmented reality technology is utilized to provide robots with dynamic task instructions and process parameter adjustments, enhancing the flexibility and adaptability of the production line. In the medical field, augmented reality-enabled 2D image perception technology for robots can be applied in scenarios such as surgical navigation and rehabilitation training. For instance, during surgery, robots can provide doctors with more accurate surgical navigation information by integrating patients' medical image data with real-time collected 2D images, assisting doctors in completing complex surgical operations. In rehabilitation training, robots can provide personalized training guidance and feedback to patients based on their rehabilitation conditions through an augmented reality interface. In addition, in fields such as service robots and agricultural robots, augmented reality technology also has broad application prospects. It can endow robots with more powerful environmental perception and interaction capabilities, and improve service quality and efficiency.

5.3 Path for Standard and Specification Formulation

Formulating unified standards and norms is an important path to ensure the healthy development of 2D image perception technology for robots empowered by augmented reality. At present, due to the lack of unified standards, there are compatibility issues between augmented reality devices developed by different manufacturers and robot systems. Inconsistent data formats and communication protocols have seriously restricted the promotion and application of the technology. Therefore, it is necessary for relevant institutions and enterprises within the industry to jointly participate in formulating standards and norms covering aspects such as hardware interfaces, data transmission, and software algorithms. For instance, stipulate the communication interface standards between augmented reality devices and robots to ensure seamless connection and data interaction among different devices; Establish a unified image data format and annotation standard to facilitate the processing and understanding of image information by robots. The formulation of standards and norms can promote collaborative cooperation among enterprises in the upstream and downstream of the industrial chain, and facilitate the large-scale application and industrial development of 2D image perception technology for robots empowered by augmented reality.

6. Conclusion

In conclusion, augmented reality has great potential and value in empowering 2D image perception for robots. By deeply analyzing its internal mechanisms, including information fusion mechanism, spatial mapping mechanism and interactive feedback mechanism, etc., we have clarified how augmented reality technology enhances the 2D image perception ability of robots in multiple aspects. At the same time, specific paths to achieve this empowerment were explored, covering aspects such as technology integration, expansion of application scenarios, and formulation of standards and norms. In the future, with the continuous breakthroughs and innovations in related technologies, augmented reality-enabled 2D image perception technology for robots will be widely applied in more fields, playing a significant role in promoting the intelligent and autonomous development of robot technology, further changing people's lifestyles and working methods, and creating greater social and economic benefits. We should continuously pay attention to the development trends in this field, increase research investment, and actively explore innovation to achieve the deep integration and coordinated development of augmented reality technology and 2D image perception technology for robots.

References

[1] Dzedzickis, A., Subačiūtė-Žemaitienė, J., Šutinys, E., Samukaitė-Bubnienė, U., & Bučinskas, V. (2021). Advanced applications of industrial robotics: New trends and possibilities. Applied Sciences, 12(1), 135.

[2] Ji, T., Sivakumar, A. N., Chowdhary, G., & Driggs-Campbell, K. (2022). Proactive anomaly detection for robot navigation with multi-sensor fusion. IEEE Robotics and Automation Letters, 7(2), 4975-4982.

[3] Martin-Martin, R., Patel, M., Rezatofighi, H., Shenoi, A., Gwak, J., Frankel, E., ... & Savarese, S. (2021). Jrdb: A dataset and benchmark of egocentric robot visual perception of humans in built environments. IEEE transactions on pattern analysis and machine intelligence, 45(6), 6748-6765.

[4] Hua, C. (2022, October). A Review of Visual Perception for Mobile Robot Navigation: Methods, Limitation, and Applications. In 2022 2nd International Conference on Algorithms, High Performance Computing and Artificial Intelligence (AHPCAI) (pp. 729-737). IEEE.

[5] Makhataeva, Z., & Varol, H. A. (2020). Augmented reality for robotics: A review. Robotics, 9(2), 21.

[6] Bonci, A., Cen Cheng, P. D., Indri, M., Nabissi, G., & Sibona, F. (2021). Human-robot perception in industrial environments: A survey. Sensors, 21(5), 1571.

[7] Wu, J., Gao, J., Yi, J., Liu, P., & Xu, C. (2022, November). Environment perception technology for intelligent robots in complex environments: A Review. In 2022 7th

International Conference on Communication, Image and Signal Processing (CCISP) (pp. 479-485). IEEE.
[8] Wang, S., Nikolić, M. N., Lam, T. L., Gao, Q., Ding, R., & Zhang, T. (2025). Robot Manipulation Based on Embodied Visual Perception: A Survey. CAAI Transactions on Intelligence Technology.