Research and analysis of camera-based robot ranging technology

Bohan Huang

One Direction Academy, Toronto, Ontario, M3B 2T5, Canada

alan99092023@163.com

Abstract. In today's society, with the progress of science and technology, robots are increasingly being applied in all walks of life. As a result, there is an increasing need for robots to accurately measure distances in different aspects of life and work. This paper explores the principles, applications, and methods of camera-based robot ranging technology through reference and research. The research shows that different ranging principles and various methods have their advantages and disadvantages, and their applicable scenarios are also very different. This article can promote the optimization of robot ranging technology and provide valuable insights for its application in more diverse scenarios.

Keywords: Robot ranging, ranging principle, method research, application scenarios.

1. Introduction

Under the premise of the rapid development of science and technology in today's society, robots have been deeply integrated into various industries. In order to enable the robot to complete the task more efficiently and accurately and help the progress of human society, the precise measurement of distances and the consideration of environmental factors are crucial. The ranging technology of the camera has also become a research hotspot because of its high accuracy.

Some researchers designed a set of underwater vehicle target ranging system based on Deeplabv3+ semantic segmentation and binocular vision to meet the needs of target detection and ranging in underwater vehicle operations. The system adopts relevant calibration and algorithm, and the experiment shows that the measurement error within 1 meter is less than 5%. However, underwater light scattering refraction will cause more imaging interference and large errors when the distance increases, while semantic segmentation can increase the measurement speed by nearly 30% without changing the accuracy [1].

This paper summarizes the application of visual recognition and picking positioning technology in fruit picking robots, summarizes the platform technology of different vision systems of robots, and analyzes the development of target recognition and positioning technology from two perspectives of algorithm design and sensor hardware. The difficulties, development and prospects of robots in fruit picking operations are discussed [2]. By reviewing and summarizing the literature, this paper deeply discusses the common principles and methods of visual ranging and its application in the camera-based robot ranging technology. The research provides a theoretical reference for the innovation and

improvement of robot ranging technology, and provides technical and theoretical support for its technical development and practical application.

2. Literature review

2.1. Development stage

In the early stage, some modern physicists used simple geometric principles and basic models, such as the pinhole imaging principle and the Euclidean formula, to estimate the distance. However, these methods had significant limitations in terms of measurement accuracy and practical application. The 1950s to 1970s was the golden period of development in the field of photogrammetry. At this stage, lens aberration expression and model were proposed and further improved, which laid a solid foundation for the later development.

In the mid-1980s, Tsai proposed a calibration algorithm based on RAC, which used radial consistent constraints to solve other off-camera parameters other than the translation of the camera optical axis, and then solved other parameters, making the calibration process faster and more accurate [3]. In 2010, Janschek et al. designed a monocular camera ranging scheme, which used monocular vision to correct the robot's course, and this technology could guide and position the robot while sensing only one landmark at a time [4]. Some scholars put forward a new scheme relying on monocular camera ranging. After calibrating the camera parameters, they integrated the data through formula and exponential function. Researchers in the Laboratory of Vehicle Safety and Energy Conservation of Tsinghua University have developed a real-time ranging scheme based on monocular vision, and calculated the pitch Angle in real time through the parallel constraints of the road to improve the accuracy of the ranging [5].

Nowadays, people use the computer vision library OpenCV, based on geometric principles such as the similar triangle principle to achieve distance measurement. For example, companies like Google and Baidu train models for autonomous vehicles using neural networks to estimate distances, with Baidu's Apollo being a notable example.

2.2. The principle of classification

Monocular camera ranging primarily utilizes geometric optics and image processing technology. By analyzing characteristics such as size, shape, and texture of objects in an image, it can roughly estimate distances. This method is relatively inexpensive and widely available. Monocular cameras have certain limitations. Due to their reliance on fixed scenes and lack of advanced intelligence, they are not suitable for high-precision measurements in complex environments [6]. The principle of binocular camera ranging is to imitate the visual structure of human eyes. Two cameras capture the same scene from different angles simultaneously, and through a series of calculations and matching processes, the visual disparity between the object and the cameras is determined. The advantage of this method lies in its high measurement accuracy and the ability to quickly and directly obtain distance information; however, it performs poorly in complex scenes and under varying light conditions, where large errors can easily occur [7]. Depth camera ranging can calculate the distance between objects by analyzing their deformation from a particular pattern projected on them, or it can use the pulse flight time of light to determine their depth and distance. The advantage of this camera is its fast measurement speed and good practicality, but it is also susceptible to changes in the light environment, and the resolution is low to a certain extent, resulting in less clear images [8].

2.3. Key algorithms and technologies

In camera ranging, it is very important to accurately extract image features and establishing the corresponding matching relationship. Common extraction algorithms include feature extraction, accelerated robustness, invariant feature scale, etc. The common features of these images are edge texture, which tend to remain stable even with image rotation, lighting changes, and scaling. The feature

matching algorithm involves identifying corresponding feature points in different images, often using methods based on the grayness of the image.

Stereo matching is the core step of binocular camera ranging, it is to calculate the visual difference by using the same pixel corresponding to the left and right images of two cameras. Common methods are region - and feature-based matching. These algorithms have their own advantages and disadvantages in computing efficiency, matching accuracy, ranging time and adaptability to different scenes.

Camera calibration is to determine some parameters of the camera, such as aperture size, focal length, main point distortion coefficient, etc., but also to determine some external parameters of the camera, such as the position and change of attitude in the real coordinate system. Accurate camera calibration is vital for ensuring the accuracy of the range measurement. The traditional calibration method is Zhang Zhengyou calibration method, which uses a specific calibration plate to calibrate the object.

Deep learning techniques, particularly convolutional neural networks (CNNS), show particular promise in camera-based ranging. Trained on a large amount of image data, CNNS are able to learn complex relationships between image features and distances, allowing for more accurate ranging. For example, some studies use end-to-end deep learning models to predict depth maps directly from images [9].

2.4. Challenge

The variation of light intensity in different environments will have a great impact on the quality of the image, which will make it difficult to extract image features and match corresponding positions, thus reducing the accuracy of ranging. The changing complexity of the environment can also be a major challenge, such as reflective glass surfaces, noisy dance parties, etc. In these complex scenes, it is difficult to ensure the acquisition of effective and significant features, which reduces the accuracy, reliability, and speed of the ranging. In many specific cases, ranging results need to be reached quickly, which requires advanced machines and intelligent operations to make fast decisions, so the calculation rate and accuracy of the algorithm are also crucial challenges. The fusion with other sensors also creates a certain complexity. The application of rangefinder camera with other software and hardware, such as ultrasonic sensor or laser radar, will improve the rangefinder performance to a certain extent. However, the data between different sensors may produce some errors and errors in unit ranging rate and comprehensive calculation, which requires calibration to improve accuracy.

3. Case study

3.1. Accurate grasp of industrial robots

Industrial robots often use a variety of camera-based ranging technologies when performing accurate grasp. The binocular vision ranging technology can obtain the stereo image of the object from two cameras, and determine the position and attitude through three-dimensional calculation, so as to realize the high-precision capture. In some mechanical parts assembly, the technology can be used to grab small, complex parts and place them in the desired location. However, the challenge of this technology lies in the complex installation and alignment of binocular cameras, and the accuracy of the position parameters of the same object that both cameras are locked on needs to be improved. There will also be some matching errors for reflective and transparent objects in complex scenes, and the hardware and technical support for them are also high. The application of structured light ranging technology is to project specific light to an object, obtain information through the deformation image of the object, and accurately calculate the position and attitude when grasping the regular object. However, this technology faces challenges in complex environments, as it is highly sensitive to changes in lighting, leading to significant errors in strong or weak light conditions. At the same time, its high equipment cost limits the application of some enterprises and individuals [10].

3.2. Obstacle Avoidance Navigation of service robots

The technologies used for obstacle avoidance navigation in robots include monocular visual ranging and depth camera ranging technology. Monocular visual ranging is cost-effective and simple in terms of calculation, making it suitable for basic obstacle avoidance scenarios, such as the car obstacle avoidance radar used in junior high school physics laboratories or indoor corridor electric doors to detect positions. However, ts measurement accuracy is limited, with significant errors in remote and outdoor scenes and less reliable performance in three-dimensional measurements. Depth camera ranging technology, on the other hand, can obtain the depth information of the environment in real time and provide a complex 3D scene perception system. It can quickly identify the shape and distance of obstacles in complex environments such as outdoor areas, shopping malls, and karaoke halls, enabling fast and accurate obstacle avoidance navigation. However, in the strong light and low light environment, the performance will be reduced to a certain extent, and the resolution of objects at very long distances is low, making the imaging results unclear, and there may be a certain visual field blind area.

To sum up, different ranging technologies have their corresponding advantages and difficulties and challenges for the accurate grasp of industrial robots and obstacle avoidance navigation in service robots. Practical applications must consider factors such as cost, application scenarios, and efficiency to select the most appropriate ranging technology for optimal integration and performance.

4. Conclusion

In this paper, the visual principles and methods of camera-based robot ranging technology are deeply and comprehensively studied. Besides, the study summarizes the advantages and disadvantages of these technologies in practical applications and explores how to integrate various ranging methods to enhance measurement accuracy and reliability, reduce costs, and assess future development prospects. The research yields foundational theories and practical insights.

However, there are still some shortcomings in this study, such as the lack of large-scale and large-scale testing of some outdoor actual scenes. Future research could benefit from expanding the experimental scope to further promote the widespread application of robot ranging technology. It is believed that in the future, robot camera ranging technology will be more mature, the cost will be cheaper, and it can be effectively applied in a complex environment, providing a strong guarantee for both travel and various practical applications.

References

[1] Hu, Q., Wang, K., Ren, F. et al. Research on underwater robot ranging technology based on semantic segmentation and binocular vision. Sci Rep 14, 12309 (2024). https://doi.org/10.1038/s41598-024-63017-8

[2] 2. Yingyan Yang, Yuxiao Han, Shuai Li, Yuanda Yang, Man Zhang, Han Li,Vision based fruit recognition and positioning technology for harvesting robots,Computers and Electronics in Agriculture,Volume 213,2023,108258,ISSN 0168-1699,

[3] https://doi.org/10.1016/j.compag.2023.108258.

[4] 3. Tsai, Roger Y. (1986) "An Efficient and Accurate Camera Calibration Techniquefor 3D Machine Vision," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, 1986, pp. 364–374.

[5] B. Lu, B. Li, Q. Dou and Y. Liu, "A Unified Monocular Camera-Based and Pattern-Free Hand-to-Eye Calibration Algorithm for Surgical Robots With RCM Constraints," in IEEE/ASME Transactions on Mechatronics, vol. 27, no. 6, pp. 5124-5135, Dec. 2022, doi: 10.1109/TMECH.2022.3166522.

[6] GU Dong-hua, WANG Hong and SUN Dong. An improved camera calibration method of robot distance measuring[J]. Journal of Light Industry, 2015, 30(5-6): 121-123. doi: 10.3969/j.issn.2095-476X.2015.5/6.025

[7]   J. Zeng, X. Pei and W. Xi, "The monocular ranging based on stable target detection and tracking for intelligent vehicle," 2022 6th CAA International Conference on Vehicular Control and Intelligence (CVCI), Nanjing, China, 2022, pp. 1-5, doi: 10.1109/CVCI56766.2022.9964692.

[8]   B. Wei et al., "Remote Distance Binocular Vision Ranging Method Based on Improved YOLOv5," in IEEE Sensors Journal, vol. 24, no. 7, pp. 11328-11341, 1 April1, 2024, doi: 10.1109/JSEN.2024.3359671.

[9]   S. Lee, "Depth camera image processing and applications," 2012 19th IEEE International Conference on Image Processing, Orlando, FL, USA, 2012, pp. 545-548, doi: 10.1109/ICIP.2012.6466917.

[10] Ersavas, T., Smith, M.A. & Mattick, J.S. Novel applications of Convolutional Neural Networks in the age of Transformers. Sci Rep 14, 10000 (2024). https://doi.org/10.1038/s41598-024-60709-z

[11] R. Li and H. Qiao, "A Survey of Methods and Strategies for High-Precision Robotic Grasping and Assembly Tasks—Some New Trends," in IEEE/ASME Transactions on Mechatronics, vol. 24, no. 6, pp. 2718-2732, Dec. 2019, doi: 10.1109/TMECH.2019.2945135.