Research on Visual Perception of Intelligent Robots Based on ADMS

Changzhong Wu1,*, Fan Ge2, Guangchao Shang1, Hengshuai Guo1, Liang Wu1, Mingpeng Zhao1, Guitao Wang1
1School of Mechanical Engineering, University of Jinan, 250022, China
2Qilu Normal University, Jinan, Shandong 250013, China
* Corresponding author’s e-mail: me_wucz@ujn.edu.cn

Abstract: In application environments such as logistics, home services, and venue welcome, reliable and efficient autonomous environment modeling has become an urgent research need and technical challenge in the field of service robots. This paper introduces the three-dimensional simultaneous positioning and map creation (SLAM) of today's highly potential (ADMS) technology research service robots, and uses cross-disciplinary research methods to carry out innovative theoretical research, principle derivation and experimental verification. The application of ADMS intelligent robots provides strong technical support.

1. Introduction
Building a harmonious and aging social background is driving a huge social demand, including intelligent service robots serving welfare homes and rehabilitation centers, such as stairs assistance and health care, which will definitely form a broad market in the future[1]. The essence of service robots different from industrial robots is that the working environment they face is unknown, dynamic, and uncertain, and requires an intelligent level of interaction with people. These are the challenging problems facing service robot technology. Among them, environmental modeling and accurate positioning under unknown environments are prerequisites for the robot to complete other tasks[2].

With the advent of Kinect, a number of industries at home and abroad, including distance education, surgery, industrial control, human tracking, etc., have carried out application research on ADMS technology with a high degree of enthusiasm.

2. ADMS Technology
ADMS includes two sets of cameras: RGB and infrared. It is more important to obtain the respective internal parameters. Calibrating the installation attitude of the sensor relative to the robot is a prerequisite for the body sensor to work. Although ADMS technology has excellent performance and low product prices, the technical insider of the ADMS system has not been cracked in China, and the measurement principle of such structured light sensors has not been revealed[3]. Therefore, a principle prototype based on this technology principle for a wider range of applications cannot be developed. Failed to propose a suitable sensor calibration method.
In the field of mobile robot navigation, although environmental modeling methods such as topological maps and emerging semantic maps have demonstrated incomparable superiority at a higher level, they still rely on the basic environment description of metric maps at the bottom[4]. For relatively complicated tasks, the robot must inevitably complete the modeling of the three-dimensional environment. Lidar-based simultaneous robot positioning and map creation technology provides an effective solution for creating a globally consistent environmental map, but it is limited to two-dimensional maps and is difficult to handle more complex robot tasks. In order to build an accurate three-dimensional metric map, a feasible method is to use two-degree-of-freedom lidar to collect 3D environmental data, and realize the full estimation of the robot's 6DOF pose through the registration between real-time laser point cloud data and the established map. This method can generate 3D environment map in SLAM framework. The significant advantage of the laser-based method is its high accuracy, but the lack of necessary textures makes the subsequent recognition process based on gray textures limited. At the same time, three-dimensional lidar sensors require higher development costs. In recent years, the rapid development of machine vision technology has created sufficient technical conditions for this, but the measurement accuracy and noise level of the vision system have largely restricted the improvement of 3D map performance[5].

Microsoft's 3D somatosensory Kinect has multiple functions such as 3D environment reconstruction, dynamic human skeleton recognition, and voice recognition. At the end of 2010, commercial 3D body-sensing camera products came out, which greatly attracted research interest in related fields. In view of its outstanding technical advantages such as high accuracy, high reliability, and friendly interface, it has received great attention from the academic community and has been applied to scientific research. It is an active infrared 3D measurement sensor[6]. Because it uses active light sources for reference, the implementation of vision algorithms is relatively simple and the operation is more efficient[7]. In terms of performance level, the reconstruction effect provided by Kinect is unmatched by the stereo vision method. It is foreseeable that Kinect technology provides a new opportunity for development of robot environment modeling research for quite some time to come. Inspired by this, this article further expands on the indoor SLAM research based on lidar, and proposes to use Active Depth Measurement Sensor (ADMS) technology to obtain a dense 3D point cloud in space, breaking the technical bottleneck of robot 3D environment perception. Making it possible to realize practical 3D SLAM in indoor environment; in order to facilitate the realization of complex tasks such as robot path planning, and based on the consideration of storage efficiency, further complete the robot topology environment modeling[8]. The development of ADMS technology will greatly improve the accuracy and efficiency of three-dimensional simultaneous positioning and map creation of service robots, giving service robots a higher ability to model autonomous environments, and will definitely provide innovative ideas and practical ideas for service robot navigation and control solution.

3. Simultaneous Positioning and Map Creation for Robots based on Lidar and Stereo Vision
So far, lidar is the most widely used robotic environment modeling sensor. Since the measurement range of the two-dimensional lidar is limited to the plane, an additional motion servo mechanism is needed to complete the laser three-dimensional scanning in order to realize the creation of the three-dimensional environment of the robot. Many scientific research institutions at home and abroad, such as Carnegie Mellon University(CMU), German Institute for Independent Intelligent Systems(AIS), and Zhejiang
University, have adopted the addition of one-dimensional scanning devices to realize the three-dimension transformation of lidar, and applied it to robots[9]. Environment map creation and navigation positioning. In principle, lidar uses a time-of-flight (TOF) -based ranging method. On the one hand, it has a significant advantage in measurement accuracy; on the other hand, its high manufacturing and maintenance costs limit its large-scale application.

In recent years, with the emergence of a number of high-performance vision algorithms and excellent open source software, the development of computer vision technology has been strongly promoted. However, long-term practice has shown that visual methods are very sensitive to changes in the physical characteristics of the surface of the object and changes in ambient lighting, and the requirements for texture are extremely harsh. In a non-cooperative environment, a large number of outliers in the disparity map, Missing points, or even block noise, make the reconstruction of point cloud and surface level unavoidable, and become a prominent technical obstacle[10]. At present, the more successful visual SLAM systems mainly use SIFT/HARRIS corners or edge straight lines as environmental features to construct sparse environmental maps. Due to the limited information provided by sparse maps, their practicability is not strong, and there is still a considerable gap with practical applications. Osaka University of Japan has studied the method of creating indoor maps based on omnidirectional vision, and has made certain research results. However, the measurement accuracy and noise level of the vision system have restricted the improvement of 3D map performance to a considerable extent.

3.1 Active Structured Light Depth Measurement

The structured light vision method based on the principle of triangulation is a high-precision, active measurement technology that shows unique advantages in industrial precision measurement, robot obstacle avoidance, etc., thanks to the principle Opened technical obstacles to stereo matching[11]. For the measurement of free surface in space, the light band mode needs to be expanded to a multi-line structured light mode (grating structure mode). On the one hand, multiple light bands can be processed in one image at the same time, which improves the utilization of image information. On the other hand, multiple light strips on the surface of the object are covered, which increases the amount of measured information and helps to obtain Larger range of depth information on the surface of the object. The measurement efficiency and range of the multi-line structured light mode have been greatly expanded, but at the same time, the increase in the complexity of calibration and the recognition of light bars have been introduced. At present, the more mature structured light technology applied to robots is still limited to the light bar mode. Due to the limited detection range, it is mainly limited to applications in robot welding and obstacle avoidance.

The Kinect three-dimensional body-sensing camera launched by Microsoft in 2010, with its outstanding cost-effectiveness, allows people to fully appreciate the unique charm of computer vision technology in people's lives. It was officially launched in the United States on November 4, 2010, and the response was overwhelming, with sales exceeding one million in just ten days! The value of its scientific research is far higher than that of game entertainment. Because of its functions such as 3D reconstruction, dynamic image recognition, and speech recognition, researchers have used it for scientific research in tracking, recognition, and control in a short period of time[12]. In principle, it is an active 3D measurement sensor. The applicant summarizes its technical essence as Active Depth Sensor (ADMS) technology. Since the active light source is used for reference, the implementation of the vision algorithm is relatively simple and the operation is more simple and effective.

It should be pointed out that, for applications such as visual environment perception, ADMS must be accurately calibrated in order to provide acceptable environmental measurements. Most countries in the world, including China, have not yet mastered the core technology of ADMS. The difficulty is that it uses a new Light Coding method instead of the traditional light bar coding technology[13]. It is a "volume coding" with three-dimensional depth. To a certain extent, the large-scale application of this technology in scientific research has been limited[14].
3.2 ADMS Calibration and Data Processing Methods
ADMS calibration includes calibration of internal and external parameters of the depth camera, relative pose calibration between the depth camera and the RGB color camera, and the installation attitude of the ADMS system (between the sensor coordinate system os-xsyszs and the robot coordinate system or-xyrzr in Figure 2). Rotation matrix R and translation vector T) calibration. Precise calibration is the prerequisite for accurate measurement. The difficulty of calibration is that infrared cameras collect infrared speckle images, so accurate feature extraction is a very challenging technical obstacle[15].

Second, the pattern projected by the active infrared light source and traditional structured light The difference is the densely arranged speckles, which puts forward new research requirements and technical challenges for the calibration of the depth sensor ADMS. The significance of studying this calibration problem lies not only in the measurement, but more importantly, to clear the technical obstacles for the self-developed ADMS principle sensor. In addition, the calibration of external parameters, that is, determining the installation attitude of the sensor relative to the robot, has a significant impact on the accuracy of the mapping.

![Figure 2. Kinect vision system installed on robot platform](image1)

![Figure 3. Kinect vision system installed on robot platform](image2)

Although ADMS vision sensors can easily obtain three-dimensional information in space without parallax matching, the original depth image acquired by ADMS often has holes, noise, etc. due to camera overexposure, black indicates absorption of light, or specular emission from the surface of the object A large number of invalid areas, as shown in Fig. 3, have the consequence of the lack of depth data, which in turn leads to a decrease in map accuracy and incompleteness.

3.3 ADMS Calibration Method and Depth Map Parallel Preprocessing Method
Through researching the principle of structured light visual measurement based on active laser speckle, it provides a reliable basis for the calibration of sensors; a feature extraction method such as corner points on infrared speckle images is proposed to solve the problem of calibration of such sensors[16].

Optically, the technology of measuring the surface vibration of an object using laser speckle has been more mature. However, if you use laser speckle to measure the surface vibration of an object, the average area of the spot is very small, generally only a few microns, so it is very sensitive to small displacements, and it is difficult to capture with a vision camera. Although the size of the speckle can be adjusted by
changing the diameter of the laser beam, if it wants to make the imaging area in the camera larger than one pixel, the price is to greatly reduce its sensitivity to displacement in depth.

In addition, infrared cameras need to collect very "pure" speckle images without being affected by ambient light and the reflection efficiency of the object itself; this project combines optical theory to further explore these specific issues.

The joint median filter is based on the traditional median filter and uses color information to assist in filling the depth value of the failed pixels. Similar to other image operations, its time complexity and space complexity are high. This project uses the CUDA program to accelerate in parallel. The technology realizes the real-time repair of depth images. The parallel tracking and mapping ideas are used to separate the camera tracking module from the map generation module. Two independent threads are established in the computer operation process, which can achieve parallel visual tracking and map creation, and improve the computational efficiency. CUDA is a potential parallel computing architecture. Based on the powerful computing capabilities of the GPU, a more efficient dense data computing solution is established. Based on the advanced CUDA technology in computer science in the process of solving the depth map, the algorithm runs faster on the TX1 or TX2 hardware platform, and the accuracy of the environment map is improved.

4. Conclusion
This paper systematically reveals the principle of the three-dimensional depth of ADMS, and then proposes a method of sensor calibration in the context of service robot application.

Acknowledgments
This work was supported by the National Science Foundation for Young Scientists of China (Grant No.51705200).

About the author
Changzong Wu(1977-), male, Weifang, Shandong Province, E-mail: me_wucz@ujn.edu.cn; Tel: 13698637583.

References
[1] M. Ning, M. X. Ren, Q. L. Fan, et al. Mechanism design of a robotic chair / bed system for bedridden aged [J]. Advances in Mechanical Engineering, 2017, 9 (3), pp.1-8.
[2] S. Ushio, K. Okada, Y. Kido, etc. A Home Security Service Robot System Using the Network Service Platform and Its Implementation [C]. 11th IEEE / IPSJ International Symposium on Applications and the Internet (SAINT), 2011, Munich, Bavaria, Germany, pp. 402-407.
[3] W. Budiharto, A. Santos, D. Purwanto, etc. A Navigation System for Service Robot Using Stereo Vision and Kalman Filtering [C]. 2011 11th International Conference on Control, Automation and Systems (ICCAS), 2011, Gyeonggi -do, South Korea, pp. 1771-1776.
[4] W. Budiharto, D. Purwanto, A. Jazidie, etc. A Novel Static and Moving Obstacles Avoidance Method Using Bayesian Filtering for Service Robot [C]. 2010 Second International Conference on Advances in Computing, Control and Telecommunication Technologies (ACT ), 2010, Jakarta, Indonesia, pp. 156-160.
[5] X. Zhong, Y. Zhou, H. Liu. Design and recognition of artificial landmarks for reliable indoor self-localization of mobile robots [J]. International Journal of Advanced Robotic Systems, 2017, 14 (1), pp. 1-13.
[6] A. Nuchter, K. Lingemann, J. Hertzberg, H. Surmann. 6D SLAM-3D Mapping Outdoor Environments [J]. Journal of Field Robotics, 2007, 24 (8/9), pp.699-722
[7] J. Welle, D. Schulz, T. Bachran. Optimization Techniques for Laser-Based 3D Particle Filter SLAM [C]. IEEE International Conference on Robotics and Automation. Alaska, USA, May 2010, pp. 3525-3530
[8] J. Tuley, N. Vandapel, M. Hebert. Analysis and Removal Artifacts in 3-D LADAR Data [R]. CMU Technical Report CMU-RI-TR-04-44, January, 2000.

[9] H. Surmannn, K. Lingermann, A. Nuchter. A 3D Laser Range Finder for Autonomous Mobile Robots [C]. Proceedings on 32nd International Symposium on Robotics, 2001, pp.153-158.

[10] Xiang Z Y, Li B. Design and Implementation of Fast 3D Ranging System Based on 2D Lidar [J]. Journal of Transduction Technology, 2006, 19 (6), pp. 2638-2642.

[11] X Z Li, S M Jia, W Cui. Consistent Map Building by a Mobile Robot Equipped with Stereo Sensor and LRF [C]. IEEE International Conference on Computer Science and Automation Engineering, Shanghai, 2011.

[12] G. Klein and D. W. Murray. Parallel Tracking and Mapping for Small AR Workspaces [C]. Proceedings of the Inter-national Symposium on Mixed and Augmented Reality (IS-MAR), Nara, 2007, pp.225-234.

[13] AJ Davison, ND Molton, I. Reid, and O. Stasse. MonoSLAM: Real-time Single Camera SLAM [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29 (6), pp.1052-1067.

[14] P. Smith. I. Reid. A.J. Davison. Real-Time Monocular SLAM with Straight Lines [J]. Proceeding of British Machine Vision Association, 2006, 2, pp. 17-26.

[15] J. Miura, S. Ikeda. A simple modeling Systems Technologies environments for mobile robots [J]. Intelligent Systems Technologies and Applications, 2009, pp.166-177.

[16] S. Ikeda, J. Miura. 3D Indoor Environment Modeling by a Mobile Robot with Omnidirectional Stereo and Laser Range Finder [C]. IEEE / RSJ International Conference on Intelligent Robots and Systems (IROS), Beijing, China, 2006, pp. 3435-3440.