Survey of Mobile Robot Vision Self-localization

Wei A. Shang
ACM Research (Shanghai), Inc., Shanghai, China
Email: 1169623366@qq.com

***Abstract*—Visual self-localization of mobile robots is a fundamental problem in robot navigation. It is the premise of trajectory planning or vSLAM. In this paper, the research results of mobile robot vision self-localization in recent years are reviewed. The principle, advantages and disadvantages of various methods are introduced.**

***Index Terms*—mobile robot, machine vision, self-localization**

I. Introduction

Location problem is to find the current position of a mobile robot by knowing the location of landmarks in its driving environment. If we know the actual position of the robot to solve the location of the landmarks, it is the problem of map creation. When the position of robot and environment map are unknown, it is a vSLAM problem [1] to build environment navigation map incrementally by using visual sensor and solve the problem of map creation and self-localization which are complementary and inseparable simultaneously. The self-localization problem of mobile robots is an important research field at present. It is the premise and foundation of mobile robots' vSLAM.

In order to achieve vision self-localization, mobile robots also need a variety of auxiliary sensors to work together, such as ultrasonic sensors, laser rangefinders, radar sensors [2], lidar [3], odometer, magnetometer [4], radio frequency identification [5], inertial navigation system. In this paper, the research results of visual self-localization for mobile robots in recent years are summarized, and their advantages and disadvantages are analyzed. The basic flow of visual self-localization is shown in Fig. 1.

II. Robot Driving Environment

Mobile robot driving environment is mainly divided into indoor environment, outdoor environment, structured environment, unstructured environment. Indoor robots are mostly home-based, including home care robots [6], sweeping robots, etc. These robots need to know their location in static physical space when they work. In the early research, the robot environment map and path are usually given, and the edge of the image is extracted and then compared with the expected map to determine the location of the robot. Akihisa Ohya's method [7] allows the robot to identify the safe channel in the foreground field of vision. The camera's field of view distance is 1 m and the range is 60 degrees. This method only focuses on the front area and eliminates a lot of irrelevant data processing. But there are two drawbacks: 1. The accuracy of obstacle recognition depends on the size and color of obstacles. 2. When there are many obstacles filling the camera's view, the robot cannot find a safe way to move forward. Liangchen Pan [8] proposed a method of visual odometer based on feature tracking. This method is applied in indoor environment with unknown map. It has high real-time performance and strong robustness to illumination changes. The speed of the robot is 0.3m/s, and the ratio of cumulative error to distance is less than 3%.

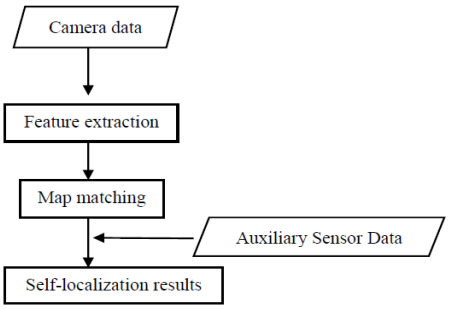

Figure 1. Visual positioning flow chart for robots

In the structured competitive environment, RoboCup [9] and FIRA Cup[10] are the frontiers of the application of visual self-localization technology, and many visual robots [11] and strategic simulation software [12] are produced for soccer matches. C. Marques [13] uses the visual system to reset the odometer periodically, which can eliminate the accumulated error of the odometer intermittently. So the soccer robot has a faster reaction speed and smooth trajectory to avoid obstacles when encountering opponents. The positioning error is 0-10cm related to the position of the robot in the field. Xiaohan Zhang[14] designed Monte Carlo self-positioning system for walking soccer robot, which probability distribution was used to represent the optimal estimation of the current robot position. A shorter side of the rectangular competition field has a greater impact on the positioning accuracy. In 20 experiments, the average error between the average positioning point and the actual position point of the robot at rest was 3.30% of the wide side of the field, 2.20% of the long side, and the floating range of the positioning point was 4.4 cm. The average error of walking is 3.97% of the wide side and 2.65% of the long side. The floating range of the error is 6.5 cm.

Manuscript received April 10, 2019; revised November 6, 2019.

doi: 10.18178/joace.7.2.98-101

III. Different Kinds of Vision Systems

The vision system of mobile robot is mainly divided into monocular vision, binocular vision, trinocular vision, panoramic vision and infrared vision. Monocular vision, that is, a single camera takes pictures in a single position. Binocular vision can take pictures of scenes either by two cameras or by image information obtained by one camera in different positions, such as installing a camera that can change position on a robot.

Panoramic vision has a panoramic perspective of nearly 360 degrees, which can obtain more complete environmental information and enhance the self-localization ability of mobile robots. Huimin Lu [15] proposed a robust omnidirectional vision soccer robot self-localization method based on particle filter and matching optimization localization algorithm. This method can achieve global localization efficiently and accurately. At the same time, the camera parameters (exposure time and gain) can be adjusted by feedback according to the change of image entropy, so that the output of omnidirectional vision can adapt to the change of ambient light. When the robot is occluded by 1/8, 1/4 and 1/2 shown in the Fig. 2, the average positioning error is less than 8 cm, the average direction error is less than 0.064 rad, and the particle filter algorithm takes 15 ms to 25 ms. In most cases, the matching optimization algorithm takes 1 ms to 3 ms, and the camera parameter adjustment can be completed in hundreds of milliseconds. Jingchuan Wang [16] designed a panoramic near infrared vision self-localization system based on coded landmarks. The auxiliary light source is used to overcome the influence of illumination, shadows, occlusion and other factors. Using the global absolute positioning advantage of vision and the local relative positioning advantage of odometer, it has high robustness and positioning accuracy in outdoor environment.

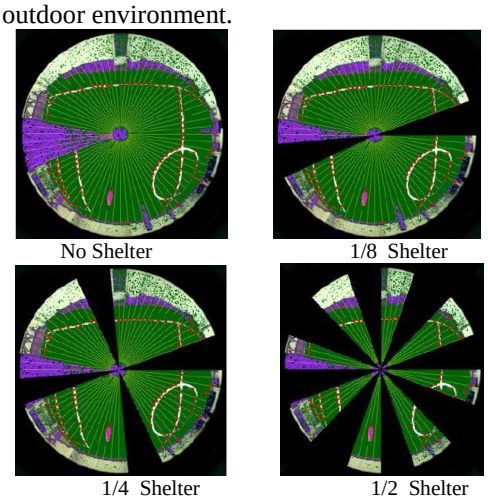

Figure 2. The results of image processing and the detection of white line points under different occlusions

Multi-sensor cooperative positioning makes up for the shortcomings of single sensor and has good robustness. However, in most cases, data fusion of different sensors is carried out at the lower level, and errors will propagate upward to affect the internal map representation at the higher level. Hyukseong Kwon [17] used interval-based logic to limit the uncertainty of sensor data. Once the robot went beyond the specified interval, it checked the consistency of different sensor data elements at all levels. The overall error was only 0.91% of the loop path that needed to be constructed. Inspired by pigeon orientation, Zhen Luo [18] introduced a positioning method based on decision-level multi-sensor information fusion. The self-positioning of mobile robot was divided into two modes: panoramic vision positioning and odometer positioning. The robot switched the positioning mode according to certain criteria. This method can reduce the amount of data calculation and balance the accuracy and efficiency of positioning. In the experiment, the soccer robot travelled a curve at 0.3m/s linear speed, and carried out 416 positioning, 181 visual positioning and 235 odometer positioning. The positioning accuracy is related to the position of the robot in the field and is also affected by the size of the site, the number and shape of obstacles, and illumination conditions.

IV. Visual Positioning Features

The vision self-localization of mobile robots needs a reference standard, i.e. road signs, which can be divided into artificial landmarks and natural landmarks. Robots are easier to recognize artificial landmarks, so they have faster positioning speed and accuracy. Maximilian Beinhofer [19] optimize the location of landmarks, reduce the number of landmarks while ensuring the reliability of location, and reduce the number of location calibrations, which is suitable for long-term self-location operation of mobile robots.

Natural landmarks are self-localization by directly utilizing the color, texture and edge density of environmental scenery. The disadvantage is that it is susceptible to the change of illumination, and it consumes resources to store a large amount of environmental information. Environmental features can be summarized as shape features and color features. Shape feature can be divided into line feature [20] and point feature which can also be divided into edge point and corner point. Fairul Azni Jafar [21] used the learning algorithm of neural network to match the color and shape features to realize the self-localization of mobile robot. Under different experimental conditions, the recognition success rate is more than 91%. Mattia Castelnovi [22] use clustering algorithm to match the perception model's color with the pre-stored environment model's color to achieve the indoor self-localization of mobile robots. However, it is susceptible to the influence of illumination and the storage environment has limited pixels, so the matched environment map is smaller. In addition, visual features can also be divided into global appearance features and local appearance features. Global appearance features include WGII features and WGOH features; local appearance features include Harris corners, SUSAN corners, SIFT features, SURF features, KLT features, etc.

V. Self-Localization Algorithm

Absolute location algorithm mainly includes: perspective method, line-of-sight method, map matching algorithm. El Mustapha Mouaddib [23] describes an absolute localization algorithm for mobile robots, which

divides the evolutionary region of the robot into rectangles limited by angle, matches maps with interpretation tree search method, and locates in indoor environment with pedestrian interference with a precision of 100 mm. Clark F. Olson [24] uses the variant of Markov localization algorithm to fit the likelihood function of Hausdorff distance of image parameterized surface, and then matches the local map with the global map by maximum likelihood estimation. He uses the branch and bound method to find the best relative position, realizes the location accuracy of 0.1M in unstructured terrain, and has strong fault-tolerant ability.

Relative localization algorithm, also known as trajectory inference algorithm, will accumulate errors over time. Liwei Han [25] proposed a visual inference positioning method based on straight line and single feature point, which takes two intersecting lines and intersections of indoor ceiling as features to carry out self-localization of mobile robots. After transforming the captured image into gray image, the global search is carried out. The average time of each localization is about 31.8 Ms. The average time of each localization is about 10.8 ms after introducing local search. The maximum error of 945 locations is 27.7 mm, and the average error is 11.72 mm. If the visual self-localization of mobile robots is regarded as an optimization problem, the particle swarm optimization algorithm can also solve it well [26].

Jingchuan Wang [27] used Kalman filter algorithm to fuse multi-sensor data to observe landmarks for global self-localization. The error between the self-localization curve and the expected path was less than 200 mm, and the error changes were not cumulative. Ofir Cohen [28] proposed an online sensor fusion framework to select the most reliable logical sensor and the most suitable algorithm for data fusion.

VI. Map Representation Method

The self-localization map representation of mobile robots can be divided into raster map, geometric map, topological map, semantic map which adapts to Visual Self-localization of Mobile Robots in Semi-static or Dynamic Environment [29], Hybrid Map and so on. In recent years, there have been many studies on semantic mapping of spatial information [30, 31, 32]. Self-localization of mobile robots requires not only spatial information, but also a deeper understanding of the environment to improve autonomy and intelligence. With the development of artificial intelligence, a high level of human-computer interaction mode has emerged, which also promotes the robot navigation based on semantic map [33, 34]. The representation of the robot's knowledge is shown in the Fig. 3. Traditional map representation is static. If a table or chair in space randomly changes its position permanently, it will affect the accuracy of map representation and generate more storage data. A two-layer navigation scheme [35] in indoor environment has good recognition and classification ability. Hybrid map is a combination of several single maps to represent the spatial location of mobile robots. Wu Hao [36] constructed a hybrid map combining the semantic map based on QR code label and topological map. The location accuracy and search time are better than the traditional single map representation. In addition, there are some MAP-free navigation methods, such as optical flow-based navigation [37], appearance-based navigation [38], behavior-based navigation and so on.

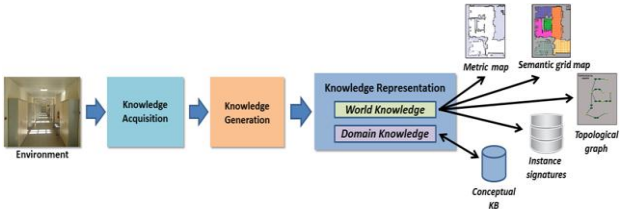

Figure 3. Representation of the robot's knowledge

VII. Summary and Development Direction

This paper summarizes the research achievements of visual self-localization for mobile robots in recent years. Visual self-localization technology still has many problems to be solved:

1. Real-time localization problem, a large number of image data need to be processed and analyzed by computer, which is a great challenge to the computer's computing and storage capacity. Most of the research is carried out in limited space, because when the space expands, a large number of image data will be generated. Logarithmic polar coordinate imaging technology [39] can reduce data processing time, save resources, and enable mobile robots to achieve real-time positioning in a wide range.

2. Occlusion problem, realizing real-time and robust localization under occlusion has been puzzling us.

3. Multi-sensor data fusion problem, the selection of effective data among sensors, the efficiency of the fusion algorithm and the reduction of error transmission also need to be solved.

In many cases, formation of robots is needed to complete tasks. Cooperative localization[40] or map creation[41] among robots is also a technical problem. Of course, we also hope that robots can perform visual self-localization in dynamic environment for a long time. Bladimir Bacca Cortes [42] proposed a histogram method with stable features to solve this kind of problem, and a spherical view method [43] can also update the topological representation of dynamic environment.

The visual self-localization system of mobile robot has many modular software architectures [44], shared libraries and related test platforms [45], which provide convenience for developers.

References

[1] J. Y. Xu, *Research on Panoramic VSLAM of Mobile Robot*, Shanghai: Shanghai Jiaotong University, 2008.

[2] G. Reina, A. Milella, and J. Underwood, "Self-learning classification of radar features for scene understanding," *Robotics and Autonomous Systems*, vol. 60, no. 11, pp. 1377-1388, 2012.

[3] M. Yang, H. Wang, et al. "Survey of pose estimation methods for mobile robots based on lidar," *Robot*, vol. 24, no. 2, pp. 177-183, 2002.

[4] J. Haverinen and A. Kemppainen, "Global indoor self-localization based on the ambient magnetic field," *Robotics and Autonomous Systems*, vol. 57, no. 10, pp. 1028-1035, 2009.

[5] Y. Takahashi, et al. "Mobile robot self localization based on multi-antenna-RFID reader and IC tag textile," in *Proc. 2013 IEEE Workshop on Advanced Robotics and Its Social Impacts*, ARSO 2013, November 7, 2013 - November 9, 2013. 2013. Tokyo, Japan: IEEE Computer Society.

[6] D. Fischinger, et al., "Hobbit, a care robot supporting independent living at home: First prototype and lessons learned," *Robotics and Autonomous Systems*, vol. 75, Part A: p. 60-78, 2016.

[7] A. Ohya, A. Kosaka, and A. Kak, "Vision-based navigation by a mobile robot with obstacle avoidance using single-camera vision and ultrasonic sensing," *IEEE Transactions on Robotics and Automation*, vol. 14, no. 6, pp. 969-978, 1998.

[8] L. C. Pan, W. D. Chen, "Research on visual location method of indoor mobile robot," *Robot*, vol. 28, no. 5, pp. 504–509, 2006.

[9] Robocup. [Online]. Available: http://www.robocup.org/

[10] Fira. [Online]. Available: http://www.fira.net/

[11] S. Y. Chiang, X. Guo, and H. W. Hu, "Real time self-localization of omni-vision robot by pattern match system," in *Proc. 2014 International Conference on Advanced Robotics and Intelligent Systems*, ARIS 2014, June 6, 2014 - June 8, 2014. 2014. Taipei, Taiwan: IEEE Computer Society.

[12] M. L. Wang, et al. "Development of a vision system and a strategy simulator for middle size soccer robot," in *Proc. 2013 International Conference on Advanced Robotics and Intelligent Systems*, ARIS 2013, May 3, 2013 - June 2, 2013. 2013. Tainan, Taiwan: IEEE Computer Society.

[13] C. Marques and P. Lima, "Avoiding obstacles - multisensor navigation for nonholonomic robots in cluttered environments," *IEEE Robotics & Automation Magazine*, vol. 11, no. 3, pp. 70-82, 2004.

[14] X. H. Zhang, "A Monte Carlo self-localization system based on visual walking robot," *Robot*, 2006.

[15] H. Lu, et al., "A robust self-localization method based on omnidirectional vision for soccer robots," *Jiqiren/Robot*, vol. 32, no. 4, pp. 553-559+567, 2010.

[16] J. C. Wang, "Robot outdoor positioning system based on near infrared vision," *Robot*, 2010.

[17] H. Kwon, K. M. A. Yousef, and A. C. Kak, "Building 3D visual maps of interior space with a new hierarchical sensor fusion architecture," *Robotics and Autonomous Systems*, vol. 61, no. 8, pp. 749-767, 2013.

[18] Z. Luo, "Self-localization of mobile robot based on information fusion of vision and odometer," *Robot*, 2006.

[19] M. Beinhofer, J. Müller, and W. Burgard, "Effective landmark placement for accurate and reliable mobile robot navigation," *Robotics and Autonomous Systems*, vol. 61, no. 10, pp. 1060-1069, 2013.

[20] T. He and S. Hirose, "A global localization approach based on line-segment relation matching technique," *Robotics and Autonomous Systems*, vol. 60, no. 1, pp. 95-112, 2012.

[21] F. A. Jafar, et al. "Autonomous mobile robot self-localization based on environmental visual features," in *Proc. 4th International Conference on Autonomous Robots and Agents*, 2009.

[22] M. Castelnovi, A. Sgorbissa, and R. Zaccaria, "Self-localization through color features detection," in *Proc. 12th International Conference on Advanced Robotics*, 2005. ICAR '05, July 18, 2005 - July 20, 2005. 2005. Seattle, WA, United States: Institute of Electrical and Electronics Engineers Computer Society.

[23] E. M. Mouaddib and B. Marhic, "Geometrical matching for mobile robot localization," *IEEE Transactions on Robotics and Automation*, vol. 16, no. 5, pp. 542-552, 2000.

[24] C. F. Olson, "Probabilistic self-localization for mobile robots," *IEEE Transactions on Robotics and Automation*, vol. 16, no. 1, pp. 55-66, 2000.

[25] L. W. Han, D. Xu, "Visual estimation and location of mobile robots based on linear and single feature points," vol. 30, no. 1, pp. 79–84, *Robot*, 2008.

[26] M. Kronfeld, C. Weiss, and A. Zell, "Swarm-supported outdoor localization with sparse visual data," *Robotics and Autonomous Systems*, vol. 58, no. 2, pp. 166-173, 2010.

[27] J. C. Wang, "Research on self-localization method of mobile robot based on panoramic vision and odometer," *Robot*, 2005.

[28] O. Cohen and Y. Edan, "A sensor fusion framework for online sensor and algorithm selection," *Robotics and Autonomous Systems*, vol. 56, no. 9, pp. 762-776, 2008.

[29] E. Einhorn and H. M. Gross, "Generic NDT mapping in dynamic environments and its application for lifelong SLAM," *Robotics and Autonomous Systems*, vol. 69, pp. 28-39, 2015.

[30] C. Galindo, et al., "Robot task planning using semantic maps," *Robotics and Autonomous Systems*, vol. 56, no. 11, pp. 955-966, 2008.

[31] A. Bouguerra, L. Karlsson, and A. Saffiotti, "Monitoring the execution of robot plans using semantic knowledge," *Robotics and Autonomous Systems*, vol. 56, no. 11, pp. 942-954, 2008.

[32] H. Zender, et al., "Conceptual spatial representations for indoor mobile robots," *Robotics and Autonomous Systems*, vol. 56, no. 6, pp. 493-502, 2008.

[33] G. Gemignani, et al., "Living with robots: Interactive environmental knowledge acquisition," *Robotics and Autonomous Systems*, vol. 78, pp. 1-16, 2016.

[34] W. Sheng, et al., "Robot semantic mapping through human activity recognition: A wearable sensing and computing approach," *Robotics and Autonomous Systems*, vol. 68, pp. 47-58, 2015.

[35] I. Kostavelis and A. Gasteratos, "Learning spatially semantic representations for cognitive robot navigation," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1460-1475, 2013.

[36] H. Wu, et al., "Spatial semantic hybrid map building and application of mobile service robot," *Robotics and Autonomous Systems*, vol. 62, no. 6, pp. 923-941, 2014.

[37] F. Kendoul, I. Fantoni, and K. Nonami, "Optic flow-based vision system for autonomous 3D localization and control of small aerial vehicles," *Robotics and Autonomous Systems*, vol. 57, no. (6–7), pp. 591-602, 2009.

[38] F. M. Campos, L. Correia, and J. M. F. Calado, "Global localization with non-quantized local image features," *Robotics and Autonomous Systems*, vol. 60, no. 8, pp. 1011-1020, 2012.

[39] V. Javier Traver, and A. Bernardino, "A review of log-polar imaging for visual perception in robotics," *Robotics and Autonomous Systems*, vol. 58, no. 4, pp. 378-398, 2010.

[40] S. Bhuvanagiri and K. M. Krishna, "Motion in ambiguity: Coordinated active global localization for multiple robots," *Robotics and Autonomous Systems*, vol. 58, no. 4, pp. 399-424, 2010.

[41] D. Benedettelli, A. Garulli, and A. Giannitrapani, "Cooperative SLAM using -Space representation of linear features," *Robotics and Autonomous Systems*, vol. 60, no. 10, pp. 1267-1278, 2012.

[42] B. Bacca, J. Salvi, and X. Cufí, "Long-term mapping and localization using feature stability histograms," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1539-1558, 2013.

[43] F. Dayoub, G. Cielniak, and T. Duckett, "Long-term experiments with an adaptive spherical view representation for navigation in changing environments," *Robotics and Autonomous Systems*, 2011. vol. 59, no. 5, pp. 285-295.

[44] C. M. Costa, et al., "Robust 3/6 DoF self-localization system with selective map update for mobile robot platforms," *Robotics and Autonomous Systems*, vol. 76, pp. 113-140, 2016.

[45] A. Jiménez-González, J. R. Martinez-de Dios, and A. Ollero, "Testbeds for ubiquitous robotics: A survey," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1487-1501, 2013.

Wei A. Shang was born in Heilongjiang Province on June 20, 1989, got Undergraduate degree in Northeast Petroleum University, then got Master's Degree in Mechanical and Electronic Engineering, Shanghai University of Engineering Science and got a first-class scholarship. The main research area is robot control. He is a MECHANICAL ENGINEER in ACM Research (Shanghai), Inc. which is located in Building 4, No. 1690 Cailun Rd. Zhangjiang, Hi-Tech Park Shanghai, China. The main work is the research and development of semiconductor manufacturing equipment, involving automated robots to carry wafers. He published "Research on Control Algorithm of Differential Drive Robot Based on Lyapunov Direct Method" in the Journal of Computer Measurement and Control, obtained a patent for utility model. Research interests are robotic control and artificial intelligence.