Survey on Indoor Positioning and Navigation Technologies Based on Sensors Fusion

Yuxuan Shao *

School of Engineering, Anhui Agricultural University, Hefei, China

* Corresponding Author Email: syx21114190@ahau.edu.cn

Abstract. Due to the growth of productivity and the development of min-sensors, relevant localization and navigation algorithms are better suited today. However, indoor localization still faces many new problems, and with the popularity of usage scenarios, the variable and complex environmental information brings many new challenges to indoor localization and navigation. Currently, the widely used sensors on the market are vision sensors, ToF sensors and wireless sensors. These three types of sensors are used under different conditions and also have different accuracies and measurement ranges. Through sensors fusion algorithms can complement the strengths and weaknesses of different sensors to achieve better performance in complex using situations. The paper presents a survey on indoor positioning and navigation technologies based on sensor fusion from different sensors and relevant applications. We demonstrate the benefits of sensor fusion algorithms in practical applications by performing sensor fusion based on and single type of sensors and different types of sensors.

Keywords: Sensors; positioning; navigation; sensors fusion.

1. Introduction

As global technology has a tremendous progress in sensors, indoor positioning and indoor navigation technologies are being applied in various industrial fields. The widespread use of sensors has produced an increasing wealth of such information.

However, the applicability of indoor position and navigation technologies is still limited by a several of factors. Compared to the open assembly lines, densely populated areas like homes and hospitals with numerous obstacles and distractions pose a challenge to the popularization of indoor position and navigation technologies [1]. This is why more and more scholars are working to develop indoor position and navigation technologies for complex conditions. The information acquired by sensors in these environments is often variable, which creates significant interference and uncertainty in indoor position and navigation tasks [2]. Thus, the ability to achieve real-time information perception and accurate positioning in complex indoor environments has become an important index for indoor positioning and navigation. For traditional positioning techniques, relatively accurate mathematical models are often required and indoor obstacle avoidance is low.

Therefore, this paper proposes the use of multi-sensor fusion to establish an intelligent optimization algorithm to cope with complex indoor environments, to address the shortcomings of traditional methods. The sensors are categorized into visual sensors, ToF (Time of Flying) sensors and wireless sensors based on different implementation methods and principles. Sections 3 to 5 of this article will introduce and summarize the principles and applications of each type of sensor in detail based on vision sensor, ToF sensor, and wireless sensor respectively.

2. Related Work

Positioning systems sense spatial location information and use related algorithms to calculate the physical and logical location information of physical entities, which can be classified into outdoor positioning and indoor positioning in different environments. However, indoor positioning information is difficult to get universal due to complex environmental factors [3] As far as current positioning algorithms are concerned, they can be divided into three classes: Proximity-based method, Scene analysis method and Geometric feature method.

The proximity-based method uses a limited range of signal action, so it is possible to determine whether the point to be measured is in the vicinity of a reference point [4]. However, the proximity-based method can only provide approximate positioning information, which can be used in the requirements of certain applications. For example, using mobile phone base station positioning, it is possible to determine the place where the caller belongs, but the position can only be determined in a certain area.

The scene analysis method is used for the positioning of measurable features at a specified location. For example, measuring the strength of the received signal and comparing it with the signal strength of that location that is implemented for measurement and exists in a database. Optimally, the scene analysis method could use video recognition to locate through a particular scene, but this requires a large data base of information to support this which also means that the cost will be much higher.

The geometric feature method is the use of geometric principles of positioning algorithms, is currently the most widely used positioning algorithms, usually need more than one point and edge as a known condition for the calculation of positioning, which is divided into the following types:

(1) The trilateral localization method: In a two-dimensional plane, three sensors measure the distance to the unknown point, as shown in Figure 1, the intersection of the circles of the three sensors forms the exact position of the unknown point [5]. According to the estimated trilateral distance, combined with the position of the reference point, the position of the point to be measured is at the intersection of the three circles.

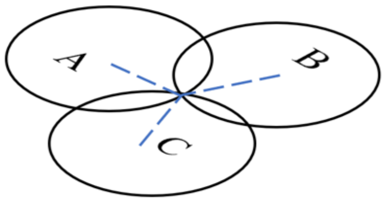

Fig 1. The trilateral positioning method.

(2) The triangulation positioning method: In a triangle shown in Figure 2, if the length of one side and the size of two angles are known, it is possible to determine the location of a third point which is the intersection of the other two sides of the triangle [6].

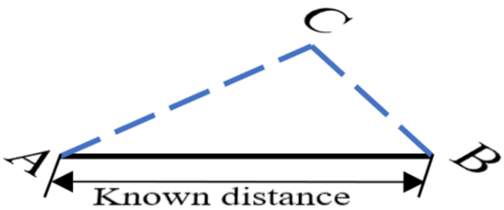

Fig 2. The triangulation positioning method.

(3) The hyperbolic positioning method: The difference between the distances of a moving point to two fixed points is constant, and its trace can form a hyperbola with these two fixed points as its focus. By adding another point it can form another hyperbola, and the intersection of these two hyperbolas is the point to be measured [7]. As shown in Figure 3.

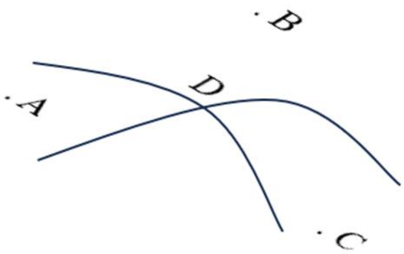

Fig 3. The hyperbolic positioning method.

Compared to outdoor position techniques like GPS, indoor positioning techniques require higher accuracy as well as cheaper costs to cope with complex indoor environments [8].

3. Vision Sensor

The basic function of the vision sensor is to acquire the raw images required by a machine vision system, the vision sensor usually has pixels of a certain resolution size to capture light and acquire images. Currently, there are two main types of methods through vision sensors: The visual SLAM method and the Deep-Learning method.

SLAM (simultaneous localization and mapping) method is the estimation of one's position by sensors in an unknown environment and the construction of a map of the environment. SLAM consists of three main components: imaging model, camera motion and map building. The imaging model of the camera represents an actual 3D-point to the image of the 2D-point correspondence, that is, the camera field of view of the 3D-point can be found in the image of the corresponding pixel point, in the absence of aberrations, the monocular camera imaging the ideal model for the pinhole model. The optical center is the center of the lens, the optical axis that passes through the optical center and the physical imaging plane perpendicular to the straight line, F on behalf the focal length is the distance from the pinhole to the physical imaging plane, P on behalf of any 3D space under the camera coordinate system, through the optical center, projected to the 2D physical imaging plane in P' points, each 3D space point corresponds to a pixel on the physical imaging plane. The pinhole model establishes a mapping relationship from the 3D world to the 2D-pixel plane through geometric relationships in the imaging process, which is mathematically described as the observation equation [9]:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{Z}\begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \frac{1}{Z}KP \tag{1}$$

Where $P'(u,v)^T$ the 2D pixel is coordinates, $P(X,Y,Z)^T$ is the 3D pixel coordinates, and K is the camera internal reference matrix, determined by the characteristics of the camera.

The input to the vision SLAM system is the camera image at different moments, and all we have to utilize is the pixel information $(u,v)^T$ of the image. In the first frame coordinate system, it is known that a certain spatial point passes through the camera model, the following two observation equations can be obtained:

$$P_1 = \frac{1}{Z_1}KP \tag{2}$$

$$P_2 = \frac{1}{Z_2}K(RP + t) \tag{3}$$

Where R and t represent the rotation and translation from the first frame coordinate system to the second frame coordinate system.

Such 8 points can match with the known n spatial point from P_1 and P_2. In the final, it is possible to derive a transformation relationship (R,t) between these two coordinate systems. The first frame coordinate system will be designated as the world coordinate system, and since the world coordinate system is fixed, the position and attitude of the camera coordinate system of the second frame can be obtained, which means that the localization has been completed, and the localization is also to obtain the position and attitude of the camera coordinate system relative to the world coordinate system at any moment.

4. ToF Sensor

The measurement principle of the ToF sensor is that the sensor emits a light (or sound wave) that immediately returns when it encounters the target object and is received again by the sensor, thereby measuring the time taken from emission to reception to calculate the distance.

$$D = \frac{vt}{2} \tag{4}$$

Where D is the distance, v is the speed of light or sound, and t is the flying of time.

ToF sensor positioning accurately determines the position of an object by measuring distance. Short-range ToF ranging has a centimeter-level accuracy in practical applications and is therefore useful for locating moving objects indoors. The principle of ToF sensor shown in Figure 4.

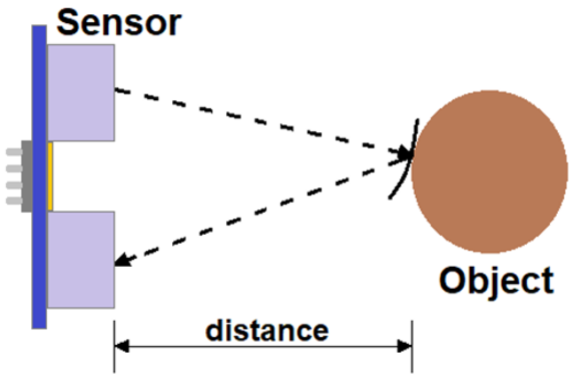

Fig 4. The principle of the ToF sensor.

Assuming that the line segment features represented by the raw data obtained by the ToF sensor are represented in polar coordinates as (ρ, α), then the equation of the line segment in the Cartesian coordinate system is [10]:

$$x cos\alpha + y sin\alpha = \rho \tag{5}$$

Where$\rho > 0, -\pi \leq \alpha \leq \pi$

Sensors based on the ToF principle are ultrasonic sensors and laser sensors. Compared to vision sensors, ToF sensors are free from light interference, have simple data processing methods and high accuracy. The piezoelectric transducer on the ultrasonic transducer is able to convert electrical signals into mechanical vibrations, which are converted into electrical signals by the receiver. Due to the severe dispersion of the ultrasonic sensors, the directional localization accuracy of the ultrasonic sensors is poor when the ultrasonic generation point is far away from the obstacle.

5. Wireless Sensor

Indoor positioning system based on wireless sensors compared to using a GPS is a more convenient and cheap method to implement. Compared with the traditional multiple sensors fixed situation for localization, wireless sensors for the indoor variable scene has better flexibility, and now more and more lightweight devices are widely used at low cost. Currently, there are two main types of wireless sensors, Wifi sensors and Bluetooth sensors.

The weakening of the Wifi signal from the access point to the receiver is converted into the distance between the two objects by using the triangulation method, which is based on the distance constraints between three or more acceptance points to the receiver The attenuation of the Wifi signal from the access point to the receiver is converted into the distance between the two using the triangulation method. We can estimate the distance from more than three access points to the receiver and locate them by the trilateral positioning method. Due to the complexity and variability of the indoor environment, it is difficult to establish a mathematical model for signal decay, so signal decay will have a greater impact on positioning.

Since the vast majority of smartphone terminals comes with Bluetooth module, it is convenient for widespread popularization and deployment of field devices. However, Bluetooth-based positioning technology is easily interfered by external noise signals, with poor signal stability and a small communication range. The most commonly used Bluetooth positioning technology is based on Bluetooth 4.0 low-power Bluetooth technology. The iBeacon technology has many advantages such as fast connection speed, high transmission rate, and stable signal transmission. However, the deployment of the iBeacon technology requires high-density deployment of Bluetooth beacons and high cost to establish the system, so it is not suitable for industrial applications.

6. Application

6.1. Sensor fusion based on the same type of sensor

When the measurement is in practice, several sensors of the same type are often used to ensure that the accuracy of each sensor is the same size, and at the same time reduce the cost of a lot. Through the inverse-variance weighting formula, we can fuse data from multiple sensors of the same type.

$$Fused\ Average\ Data(t) = \frac{1}{60}\sum_{k=0}^{59}\sum_{i=1}^{N}\overline{w_i}(t-k)\cdot x_i(t-k) \tag{6}$$

Where t is the time, $\overline{w_i}$ is the weight, x_i is the current data and Fused Average Data is the result of the inverse-variance weighting formula.

To verify the inverse-variance weighting formula, we choose to use three different ToF sensors, the laser module VL6180X, the ultrasonic modules US-100, and HCSR-04. Taking into account the measuring range and the measuring angle of the modules, we chose a piece of cardboard of A4 size as the measuring object and measured it under 5cm, 10cm and 15cm respectively. To keep the sensor emitting points in the same line, we use a protoboard to connect all the module. We draw three parallel lines spaced 5cm apart and then place the modules on the parallel lines to measure the distance. The schematic of the experiment is shown in Figure 5.

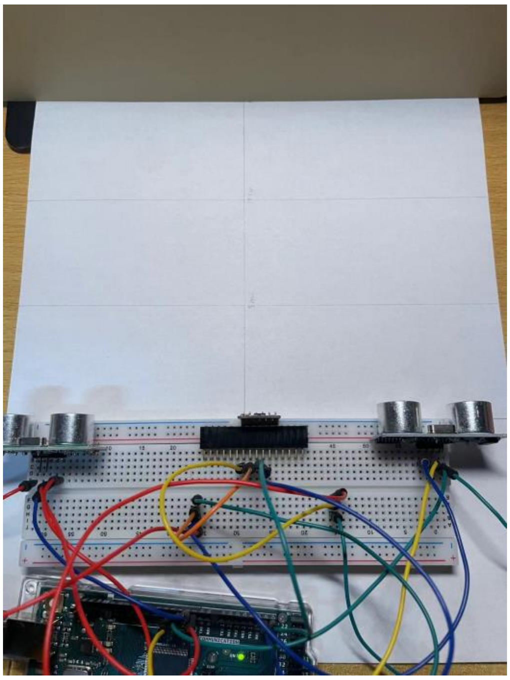

Fig 5. The schematic of the experiment to verify the inverse-variance weighting formula.

After collecting the data at 5cm (Figure 6), 10cm (Figure 7) and 15cm (Figure 8), we compared the result after the inverse-variance weighting and the original data. It can be seen that there is an excellent convergence effect for data from multiple sensors.

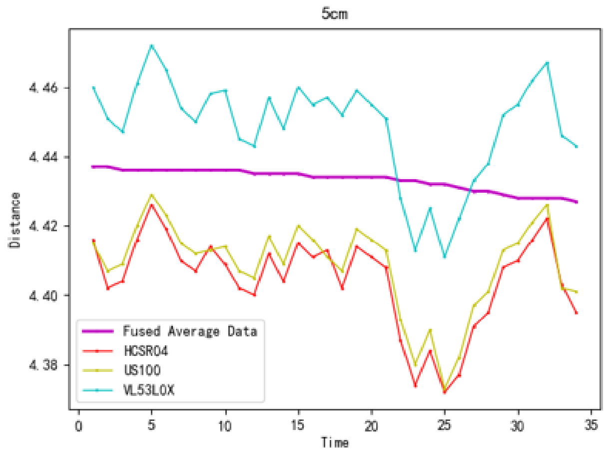

Fig 6. The result of the inverse-variance weighting data at 5cm.

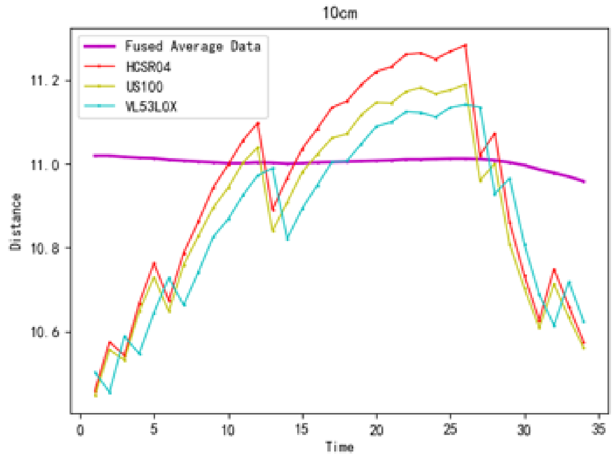

Fig 7. The result of the inverse-variance weighting data at 10cm.

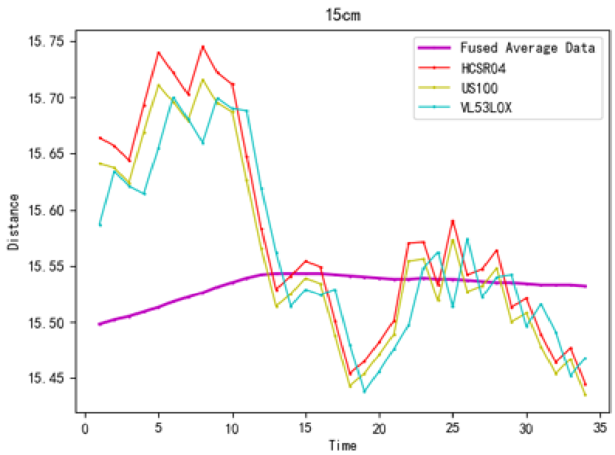

Fig 8. The result of the inverse-variance weighting data at 10cm.

6.2. Sensor fusion based on different types of sensors

ZHAO Yiheng and ZHOU Zhifeng built a testbed based on laser and vision sensors and extracted image corners based on Harris' corner points [11]. Harris corner point detection utilizes the principle that feature points have the property of local differences. By taking a custom-sized window centered on each point, if the window is moved in any direction with the feature point as the center, the surroundings produce a large change. It shows that a large gradient change is produced in this region, and this feature point is called the Harris corner point. A comparison of mapping with a single sensor and mapping with fusion of laser and visual sensors is shown in the Figure 9.

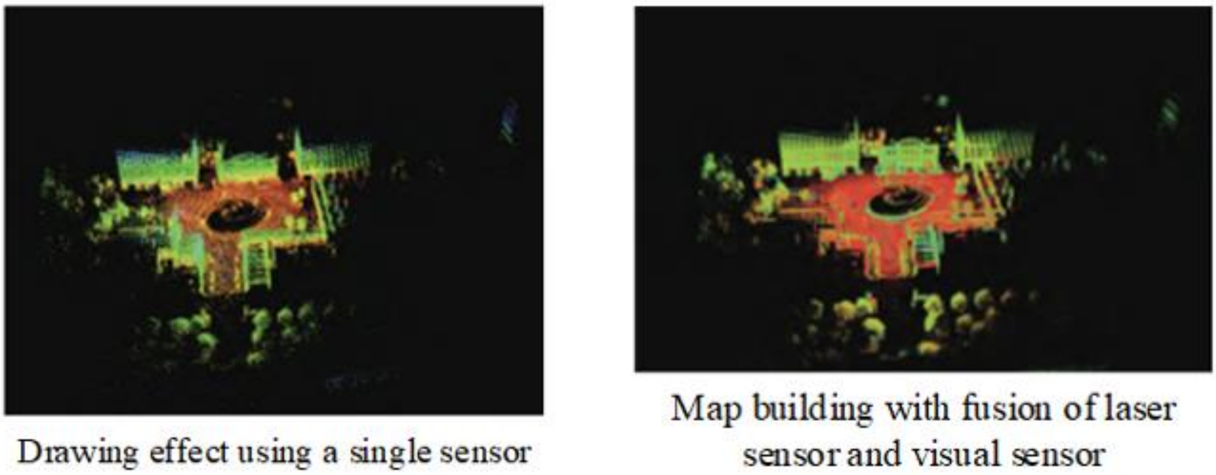

Fig 9. A comparison of mapping with a single sensor and fusion of laser and visual sensors.

It can be noticed that the map features are not distinct and less robust under a single sensor. After using sensor fusion, the class notes that the noise is reduced and the feature points are obvious.

ZHANG Wen utilized laser sensors and vision sensors to realize autonomous map building for mobile robots in indoor scenes [12]. The robot converts the laser data in real time, localizes itself through adaptive Monte Carlo in the ROS framework, uses the dynamic window method for obstacle avoidance, and generates a map based on the Gmapping method after traversing the global environment. And using a convolutional neural network to semantically classify the RGB images obtained from visual sensors. Based on the semantic map, different weights are assigned to different locations to achieve path planning under different tasks (Figure 10 and Figure 11).

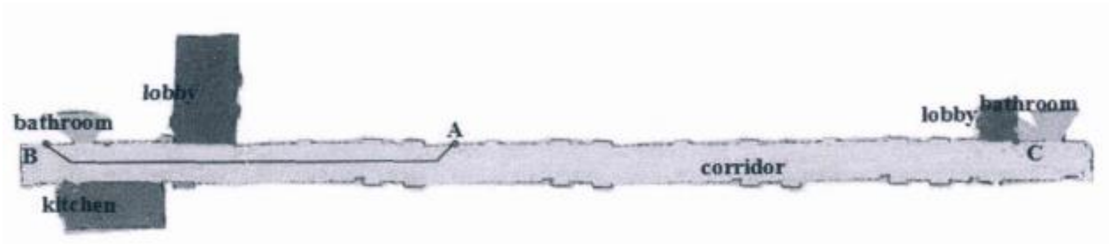

Fig 10. Selection of the shortest path during working hours.

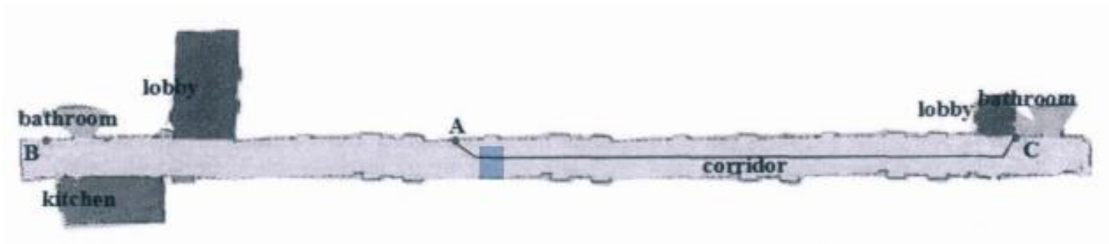

Fig 11. Avoid kitchens and halls during lunchtime.

7. Conclusion

This paper summarizes and analyzes the current indoor positioning algorithms, illustrates the performance of the current mainstream sensor positioning algorithms in indoor positioning, and conducts experiments with two sensor fusion, which illustrates that the fusion algorithms have excellent performance results in indoor positioning. However, there are still many shortcomings, due to the limitation of experimental conditions, we cannot verify the effect of the fusion algorithm in a more complex environment. However, as far as the current experimental effect is concerned, sensor fusion can realize the accuracy of indoor localization without high cost.

References

[1] Ramon F. Brena, Juan Pablo García-Vázquez, Carlos E. Galván-Tejada, David Muñoz-Rodriguez, Cesar Vargas-Rosales, and James Fangmeyer Jr. Evolution of indoor positioning technologies: a survey. Journal of Sensors, 2017, 1-21.Fangfang.

[2] M. B. Alatise, G. P. Hancke, A review on challenges of autonomous mobile robot and sensor fusion methods,in IEEE cess, vol. 8, pp. 39830-39846, 2020.

[3] Byeon TW, Jang SY. A study on the technological and environmental factors affecting the accuracy of beacon based indoor positioning system. Journal of the Korea Society for Simulation. 2016;25(2):21-9.

[4] Song J, Haas CT, Caldas CH. A proximity-based method for locating RFID tagged objects. Advanced Engineering Informatics. 2007 Oct 1;21(4):367-76.

[5] Xingqi Chen and Zhi Gao, "Indoor ultrasonic positioning system of mobile robot based on TDOA ranging and improved trilateral algorithm," 2017 2nd International Conference on Image, Vision and Computing (ICIVC), Chengdu, China, 2017, pp. 923-927

[6] Hartley RI, Sturm P. Triangulation. Computer vision and image understanding. 1997 Nov 1;68(2):146-57.

[7] Liu BC, Lin KH, Wu JC. Analysis of hyperbolic and circular positioning algorithms using stationary signal-strength-difference measurements in wireless communications. IEEE Transactions on Vehicular Technology. 2006 Mar 20;55(2):499-509.

[8] XI Rui, LI Yu-jun, HOU Meng-shu. Survey on indoor localization. Computer Science, 2016, 43(04):1-6+32. (In Chinese)

[9] LIU Yin-zhen, XU Xiang-rong, ZHANG Hui, YU Qing-song. Visual SLAM algorithm based on semantic and geometric constraints under dynamic scenes. Information and Control, 2023,11,19:1-8. (in Chinese)

[10] ZHUANG Yan, WANG Wei, WANG Ke, XU Xiao-Dong. Mobile Robot Indoor Simultaneous Localization and mapping using laser range finder and monoculer vision. Acta Automatic Sinica, 2005(06):113-121. (in Chinese)

[11] ZHAO Yiheng,ZHOU Zhifeng. Location and mapping of lidar and vision sensor fusion. Journal Of Shanghai University of Engineering Science, 2022,36(04):392-397. (in Chinese)

[12] ZHANG Wen. Research on autonomous navigation method for indoor robots based on multi-sensor fusion. (in Chinese)