Estimation of AUV Position and Attitude Based on Multi-sensor Fusion

A Dissertation Submitted for the Degree of Master

Candidate： Geng Yixuan

Supervisor：Prof. Ferrero Alessandro

Department of Electronics, Information and Bioengineering

Polytechnic University of Milan, Milan, Italy

Abstract

This paper mainly focus on the position and attitude estimation of AUV based on multi-sensor data fusion. The study includes the basic structure of multi-sensor data fusion systems, attitude and position calculating model based on inertial navigation system, single-camera vision system and USBL positioning system, an improved D-S evidence fusion method based on TOPSIS and its application on image segmentation, as well as the state estimation algorithm based on studies above. Position and attitude estimation is the key point in the docking of AUV. In view of the complexity of underwater environment, such as heavier noise interference, data received by sensors may be not accurate enough, and under special circumstances, the sensor data can severely distort. In order to improve the accuracy and reliability of attitude and position estimation, data fusion of multiple sensors is required. For the estimation and attitude estimation in short range, the inertial navigation sensors and monocular camera are chosen, and firstly the positioning model of monocular camera was modified for better accuracy. Considering the underwater light scattering, a single feature may not be accurate enough in the image feature extraction, so several features are fused to segment the region of reference lights. Common fusion methods require relatively clear prior knowledge, which is hard to obtain in actual cases, and their performances can be poor when dealing with systems with strong uncertainties. In this regard, D-S evidence theory has fewer demands and constraints on prior information, and is good at dealing with uncertain information. However, due to the fact that its classic evidence combination rule often produces results inconsistent with intuition when processing evidences with high conflict, the application of classic D-S method is limited. In order to apply D-S evidence theory to image feature fusion, a new discounting evidence method based on TOPSIS is proposed in this paper for highly conflicting cases. The feasibility of this method is verified through simulation and comparison with similar methods proposed in these years. When the lights are segmented based on fused feature, the Blob operator is used to locate the centroids of guiding lights, which are also the basis of three-dimensional positioning based on the modified

monocular camera model. After obtaining the attitude and position from inertial components and monocular camera separately, the error state model of attitude estimation and state model of position estimation are built respectively. By fusing the accelerometer and magnetometer data together with gyroscope data, and fusing the monocular camera data with accelerometer data, integrated attitude and position estimation is conducted. In the relatively long range, the inertial navigation unit and acoustic positioning system are combined to estimate the position of AUV. To compensate the accumulated error of dead reckoning, the positioning based on USBL system is introduced, and the main sources of error are analyzed. On that basis, the filtering models are built and the position of AUV is estimated through a weighting fusion algorithm. Besides, considering the fluctuation of noise, the Sage-Husa algorithm is used to achieve adaptive Kalman filter. The effectiveness of proposed algorithms is validated by simulation and experiment. In the process of research, some deficiencies were found and discussed, and some prospects for further research were proposed.

Key words: Position and Attitude Estimation, Single-camera Vision, Inertial Navigation, USBL Positioning, Multi-sensor Data Fusion.

Sommario

Questo tesi studia principalmente l'atteggiamento dell'AUV e la stima della posizione basata sulla fusione di dati multi-sensore, inclusa la struttura di base del sistema di fusione multi-sensore, modello di calcolo della posa inerziale, visione monoculare, modello di posizionamento acustico subacqueo USBL, algoritmo D-S basato su prove evidence based su TOPSIS e la sua applicazione nella segmentazione dell'immagine e algoritmo di stima dello stato AUV basato sulla ricerca di cui sopra.La stima di posa sott'acqua AUV è la tecnologia chiave del recupero a distanza ravvicinata. A causa del complesso ambiente sottomarino e dell'esistenza di rumore e altri fattori, l'accuratezza dei dati di vari sensori spesso diminuisce. In alcuni casi, si discosta anche seriamente dal valore reale, che ha una forte incertezza.La fusione di informazioni multi-sensore può migliorare l'accuratezza e l'affidabilità della stima della posa in una certa misura.A distanza ravvicinata, la navigazione inerziale e la visione monoculare vengono selezionate per completare la stima di posa combinata e il modello di posizionamento della visione monoculare viene modificato in una certa misura. In vista della grave dispersione della luce subacquea, la singola funzione potrebbe non essere sufficientemente accurata, quindi questo tesi fonde varie misure per estrarre la fonte di luce di riferimento dall'immagine visiva.Al momento, gli algoritmi di fusione comunemente utilizzati spesso richiedono forti informazioni preliminari e le loro prestazioni sono scarse quando si affrontano forti incertezze. La teoria delle prove D-S ha meno requisiti e vincoli sulle informazioni preliminari e ha i suoi vantaggi nel trattare l'incertezza.Basato sul processo decisionale dell'intervallo di TOPSIS, questo tesi propone un metodo di dimostrazione dello sconto D-S migliorato e ne verifica la fattibilità e l'efficacia attraverso la simulazione numerica e il confronto dei risultati.Dopo la segmentazione dell'immagine basata sull'algoritmo di fusione, questo tesi utilizza l'operatore Blob per individuare la sorgente luminosa guida e calcola le coordinate relative tridimensionali basate sul modello di visione monoculare modificato.Puntando sui risultati della navigazione inerziale e della stima della posa della visione monoculare, vengono costruiti il modello dello stato di errore per la stima degli assetti

e il modello di stato per la stima della posizione. Il giroscopio viene modificato dai dati dell'accelerometro e del magnetometro e l'accelerometro viene modificato dal risultato della posizione visione monoculare. Viene proposto un algoritmo di stima della posa in fusione basato sul filtro di Kalman.Mirando a lunga distanza, combinato con la navigazione inerziale e il posizionamento USBL, l'errore cumulativo della stima di assetto viene compensato dai risultati di posizionamento acustico sott'acqua. Questa parte introduce il principio del posizionamento USBL, analizza le principali fonti di errore, utilizza la fusione ponderata per filtrare e stimare e utilizza l'algoritmo Sage-Husa per realizzare un filtro di Kalman considerando l'incertezza del rumore.Infine, il metodo proposto è verificato mediante simulazione ed esperimento. Nel corso della ricerca sono state anche trovate e riassunte alcune carenze e sono state prospettate alcune prospettive per ulteriori ricerche.

Parole chiave: stima della posa,visione monoculare,navigazione inerziale, posizionamento USBL,fusione di dati multi-sensore

Contents

Figures and Tables

Chapter 1 Introduction

1.1 Background and Significance of the Research

The twenty-first century is the century of the ocean, which contains inexhaustible valuable resources and has attracted many research workers to the exploration. Considering the particularity, complexity and harshness of the undersea environment, the underwater vehicle, which can carry a complete detection system and accomplish the task independently, is bound to become the backbone of the marine reconnaissance, topographic exploration, resource exploration and even military confrontation. Due to energy constraints, AUV needs to complete the recovery independently when the task is completed or the energy is insufficient, and the accomplishment of recovery depends on precise position and attitude estimation [1]. Commonly used underwater positioning sensors on AUV include sonar sensors, inertial sensors, visual sensors and electromagnetic sensors. Among them, the sonar sensor may have large blind area in close range, which induces large measurement error and makes it not suitable for AUV docking at short range [2]. Inertial sensors do not rely on external information, therefore have high accuracy and can resist external interference [3, 4]. In consequence, they have been widely used in underwater vehicle navigation. However, the integral error of IMU increases with time significantly, as a result, when inertial navigation is used separately, the measurement deviates from the real position and attitude seriously after a period of time. The vision sensor has higher accuracy and is suitable for short distance positioning. However, considering the complex underwater environment and the serious influence of noise and light scattering, the positioning still needs further improvement. Based on the idea of multi-sensor data fusion, this paper fuses the data of inertial navigation system, magnetometer and monocular vision sensor. The position and attitude of AUV are estimated by an integrated system in order to make full use of the sensors, as well as improve the accuracy and reliability.

1.2 Research Status of the Subject

1.2.1 Research Status of UUV

According to the operation mode and autonomous ability, UUV can be divided into ROV which needs remote control and AUV which has autonomous underwater operation ability. The latter is also the main application target of algorithms proposed in this thesis. The research on unmanned underwater vehicle (UUV) in western countries can be traced back to the middle of last century [5]. However, until the 1970s, the UUVs developed successfully and put into use are still mainly the ROV operated by remote control [6]. After the 1970s and 1980s, with the progress of hardware, some UUVs had formed semi-autonomous operation capability to a certain degree, but hindered by the cost, they were not put into volume production and large-scale use [7]. Nevertheless, by 2000 or so, the maturity and commercialization of relevant technology had promoted the cost reduction of UUV. Since then, not only UUVs with special functions such as anti-submarine had begun to be formally equipped in the army, but also civil UUVs had been increasingly used in exploration, scientific investigation and other fields [8]. After entering the new century, the continuous expansion and upgrade of tasks has put forward higher requirements for the autonomous operation ability of UUV. In the first decade of the 20th century, the United States and Europe have formulated plans for the development of unmanned marine systems, intensifying relevant research, procurement and assembly of UUVs.

According to incomplete statistics, there are more than ten countries involved in UUV research all over the world. The research institutions are from private companies, military laboratories, universities and research institutes, etc. [9]. At present, the leading position of research and application of UUV are still occupied by the United States and European countries, including Norway, France, Italy, etc. In general, the products launched and put into application are pretty few, while mature products are still mainly ROV, and most of AUVs with ability of self-operation or switching to self-mode are still in the stage of development and test. According to the Interoperability Level Standard STANAG 4586 published by the United States in 2007(as shown in Figure 1) [9], most current UUVs are at Level 3 or Level 4.

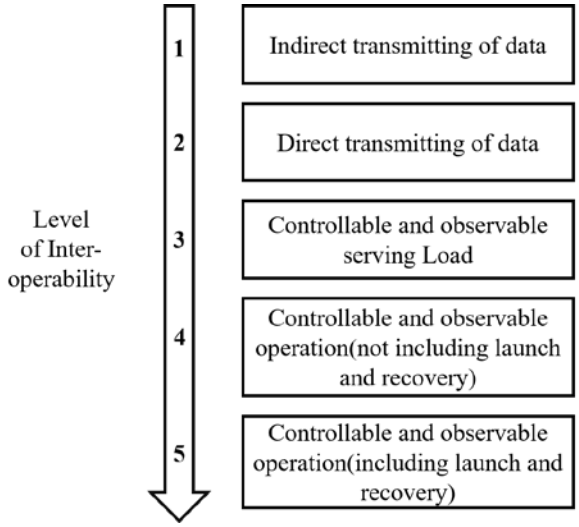


Figure 1 Grading Standard of Interoperability in STANAG 4586

At present, the AUVs that have passed the final test mainly includes: ABE AUV launched by WHOI of the United States (whose maximum submergence depth is 6 km and can cruise more than 50 hours at a time, as shown in Figure 2 [10]), REMUS 6000 produced by Hydroid of the U.S. (with submarine depth ranging from 25 to 6000 m, which basically represents the state of the art in this field, as shown in Figure 3) and AUV62-MR launched by Sweden (already assembled in submarines, as shown in Figure 4). Chinese research on UUV started late, and institutions involved are relatively fewer, but at present, some good results have been achieved [10]. CR-01 developed by Shenyang Institute of Automation, Chinese Academy of Sciences (as shown in Figure 5) can submerge to a depth comparable to REMUS 6000, and has played an important role in the investigation of seabed resources during the international scientific examination activities. Since the 1980s, Harbin Institute of Technology has also carried out a lot of researches on UUV. The flat UUV developed by Harbin Institute of Technology has approached the international advanced level in technology. However, in general, there is still much room for improving the autonomous operation ability of UUV, as well as the R&D and

application of AUV in China.

Figure 2 The ABE Developed by U.S.

Figure 3 The REMUS 6000 Developed by U.S.

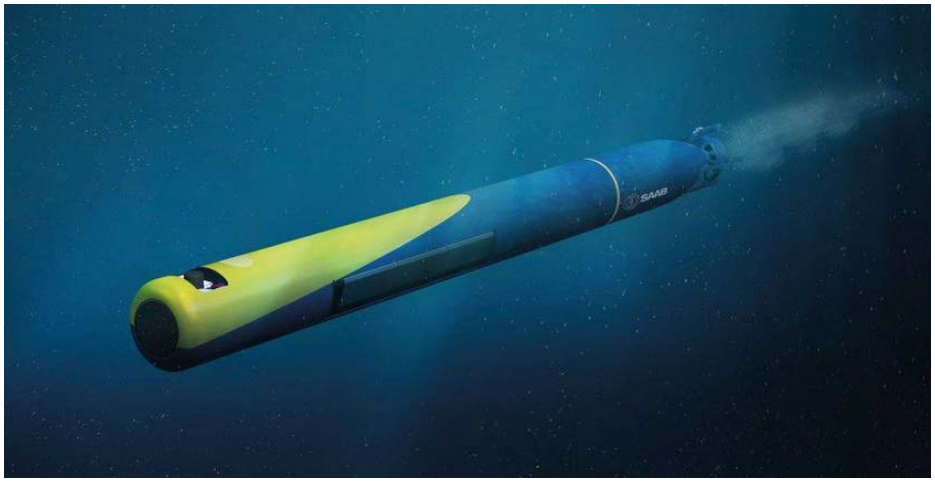


Figure 4 The AUV62-MR Developed by Sweden

Figure 5 The CR-01 Developed by Shenyang Institute of Automation

1.2.2 Research Status of Position and Attitude Measurement of AUV

Given the limitation of AUV's volume and weight together with the particularity of its working environment, it is difficult to achieve accurate estimation of AUV's position and attitude. The existing products usually use one or more of inertia and acoustic sensors, odometer, and geophysical information to navigate, and visual navigation is usually used as reference in closer docking. The commonly used navigation methods and the comparison are shown in Figure 6.

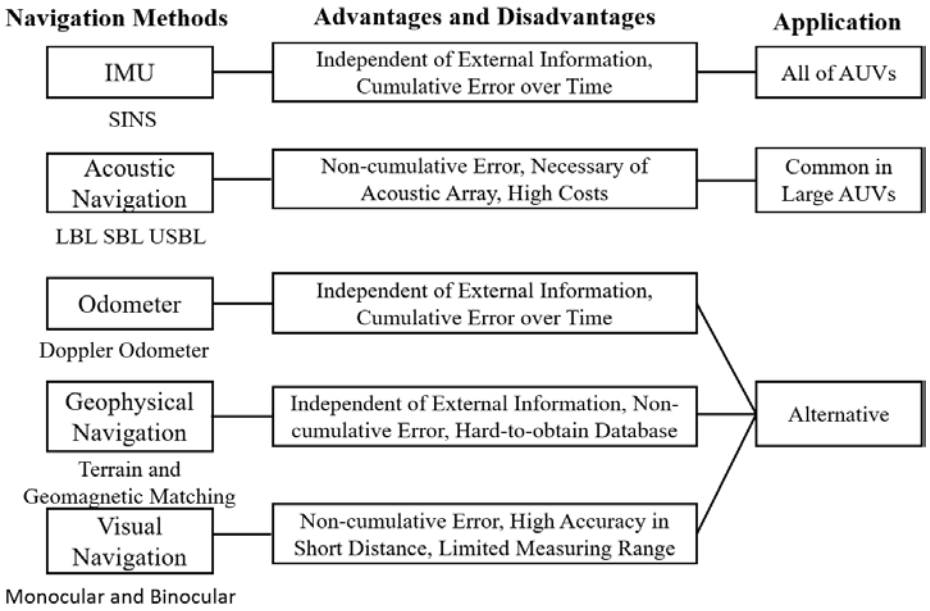


Figure 6 Common Navigation Methods and Comparison

At present, AUVs with the most accurate navigation systems are the BPAUV [11] of the United States and the "Hugin -1000" of Norway [12]. Both of them integrate the information

from inertial sensor, odometer, GPS and other signals to improve the navigation accuracy and reliability of the system. There are fewer studies on the integrated navigation system of AUV in China. Up till now, it is still a mainstream and promising research direction to fuse multi-sensor data, developing an integrated and autonomous navigation mode that rely less on the outside.

With the rapid booming of computer and image processing technology, machine vision has been widely used in the position and attitude estimation of robots. Except for some special illumination conditions, there is no great difference between the underwater vision system and those conventional vision systems. By calibrating the parameters inside and outside the camera, the spatial relationship between the image and real object can be obtained based on perspective principle. The amount of information obtained by visual guidance positioning is more, and the volume, weight and cost of visual sensors are relatively lower than that of sonar sensors, so good positioning accuracy can be guaranteed when enough visual information can be collected. Therefore, underwater visual navigation has drawn more and more attention from researchers and has been equipped on more and more AUVs. Many scholars have also carried out research on Vision-Based underwater positioning and tracking. As early as 2001, the AUV Kamraba developed in Australia was equipped with visual positioning system and it is extended to the subsequent generations. The camera deployment of Kamraba is shown in Figure 7. In 2010, American ASL Laboratory launched SAUVIM, which can achieve full-autonomous operation. The vision sensor is successfully used in target search and location and the enlarged view of its front-end is shown in Figure 8. Up to now, vision sensors have been successfully applied to underwater welding, target tracking and the docking of AUV.

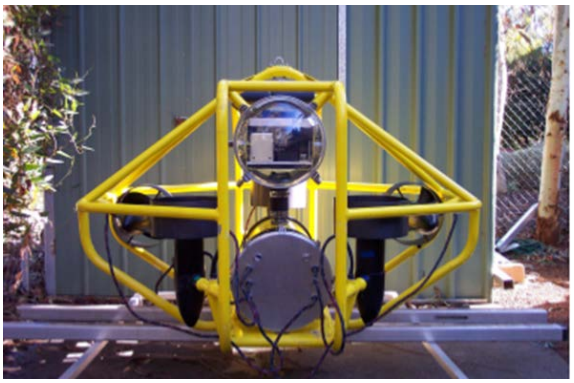


Figure 7 Layout of Camera in Kamraba

Figure 8 The Enlarged View of Front-end Camera in SAUVIM

Although visual sensors have outstanding advantages in accuracy, due to the strong underwater noise and serious light scattering, in real cases its range of application is relatively small, which makes it only used for high-precision proximal positioning of AUV [16]. Moreover, the computational burden of visual positioning is usually much larger than that of inertial and acoustic navigation, which worsens the real-time performance of system. Therefore, other sensors with lower accuracy but better real-time performance are still needed as supplement to achieve higher frequency pose estimation. At present, monocular vision positioning has been successfully applied in many fields, such as warehouse positioning, robot obstacles avoidance, medical surgery and autonomous navigation of unmanned vehicles. Its positioning model is simple and convenient, but it can't solve the depth independently without external auxiliary equipment or image correlation. Binocular vision positioning has higher accuracy in depth measurement, but considering that it needs to calculate the time difference based on feature points matching to get depth, the calculation burden is even heavier. The experimental platform

in this thesis is equipped with a monocular camera with high accuracy, and a fixed-size docking assembly can be used as reference during the docking of AUV. Therefore, monocular vision positioning method is used and studied in the follow-up contents. Considering that there may be large errors in calculating the vertical distance between the AUV and the docking assembly by monocular vision positioning, and that AUV may have certain yaw and pitch angles under the influence of external forces underwater, it is necessary to modify the general positioning model to improve the accuracy of position estimation as much as possible.

With its advantages of less dependent on external information, strong anti-jamming ability and better concealment, inertial navigation has a wide application prospect in underwater positioning. Inertial navigation system consists of gyroscope sensitive to angular velocity and accelerometer sensitive to gravity, according to the initial position and attitude information, it can determine the current attitude, velocity and position of mobile carrier by integral calculation. At present, one common application of inertial navigation system is the strap-down inertial navigation system [17], in which the inertial sensors are fixed in the motion carrier. The advantages of strap-down system are simple structure, small volume and weight, low cost, and high reliability. However, it is undeniable that SINS has its own unavoidable drawbacks. The main manifestation is that in the long term, with the increase of integration time, the integration error will accumulate gradually at each step, which eventually leads to the serious deviation of estimated position and attitude from the true value. The four stages of the development of inertial navigation elements are shown in Figure 9.

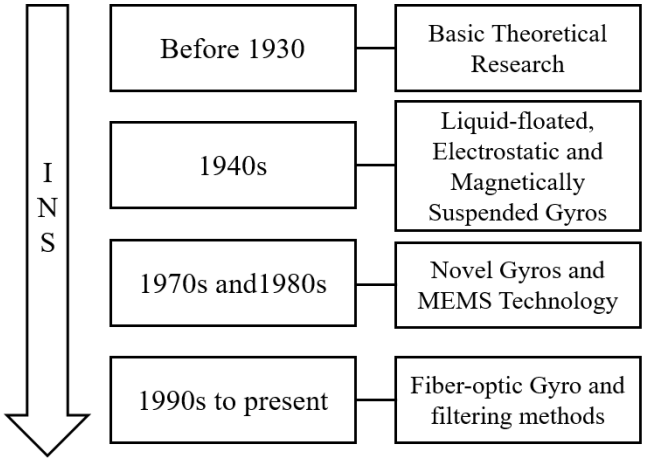


Figure 9 Development Stages of Inertial Navigation System

At present, there are two main ways to reduce the cumulative error of inertial elements: one is improving the inertial elements in principle and structure, and the other one is improving the filtering algorithm. In addition to the classical filtering algorithms, some scholars have also introduced the theory of fuzziness and self-adaptation into the field of inertial data filtering, and have achieved some good results. On the other hand, with the continuous advancement of inertial navigation algorithm and industrial technology, the accuracy of inertial navigation equipment has been increasing, with the volume and power consumption decreasing [19]. The specifications and prices of some mainstream inertial navigation devices are shown in Figure 10. Most of current AUVs under R&D or put into use are equipped with inertial navigation system. It can be predicted that inertial navigation will remain as a key link in the research towards underwater vehicles in the future, and play an important role in the autonomous position and attitude estimation of underwater robots.

Models	Range of Gyro (deg/s)	Range of G-meter (G)	Weight (g)	Price
Landmark 20 IMU	300	10	–	$3995
Spatial	2000	16	25	$3000
Nano IMU	1200	5	15	€2100
Terrella	2000	2	17	$1300
ADIS 16355	300	10	16	$600
Nav Board M3	2000	3	40	$299
CHR-6d	400	3	1.5	$125
MPU 9150	2000	16	0.1	$80

Figure 10 Specification and Price of Some Mainstream Inertial Navigation Devices

Through the above introduction, it can be concluded that both visual positioning and inertial measurement have their own advantages and disadvantages. A single navigation mode may not meet the requirements of accurate position and attitude measurement in AUV docking. In order to make up for the shortcomings and further improve the accuracy and reliability of positioning, researchers pay more and more attention to multi-sensor combination, fusing the visual and inertial data to obtain fused position and attitude estimation. Inertia-vision integrated navigation system obtains image information through visual sensors, then after certain

processing, the position of the carrier is calculated and used to correct the cumulative integral error of the inertial system. On the other hand, inertial sensors make up for the disadvantage of real-time visual navigation by their high update frequency and good adaptability to underwater environment. The idea of integration of inertial navigation and visual navigation has also attracted attention in many fields, and has achieved good fusion effects in the simulation and tests of automatic vehicles and some atmospheric aircrafts [20]. Therefore, aiming at the position and attitude estimation of AUV, the inertia-vision integrated system has great feasibility and prospect of research and application.

1.2.3 Research Status of Evidence Theory

In this thesis, after analyzing the characteristics of underwater visual imaging and image processing technology, based on previous research experience, the idea of improving evidence theory and applying it to underwater image segmentation is determined. Evidence theory has prominent advantages in the representation and integration of uncertainty as well as fusion of imperfect information. Since its inception, it has attracted wide attention of scholars at home and abroad. After decades of enrichment and development, it has been successfully applied in many fields such as target recognition, decision-making, semantic recognition and fault diagnosis. Dempster's classical evidence combination rule provides a general way for decision-making by integrating multiple evidences, but the premise of correct decision-making is to control the degree of conflict between evidences within a certain range. However, in actual application scenarios, due to structural defects of sensors or inaccurate measurement induced by noise interference in the ambient, the evidences used for fusion are very likely to have a high degree of conflict. Under this premise, the classical rule of evidence combination shows greater limitations. If the evidences are fused by the rule of combination without modification, wrong conclusions which are contrary to intuition can be got, which is also one of the reasons why D-S evidence theory is criticized by people, and limits its application to broader fields in some degree. In order to synthesize conflicting evidences reasonably, many scholars have proposed their improved methods, but in general, the mainstream amendment ideas can be divided into

two categories: one is modifying the rule of evidence combination and the other is modifying the sources of evidences. To modify the combination rule, it is necessary to redistribute and recombine the conflicting evidence information. The difficulty lies in how to build a logical and effective way to achieve it. The key to modifying the sources of evidences is to evaluate the reliability of individual evidence and the mutual support degree between different evidences, and to preprocess the evidences before fusion.

Based on the first idea, Yager [24], Dubois [25] and Lefevre [26] have put forward some influential conflict information redistribution methods, which improve the fusion effect of conflict evidences in some situations. In recent years, some scholars have carried out further exploration and supplement based on similar ideas. However, these redistribution and combination rules often have some problems, such as complex reasoning process and huge computational load, which limits the application of the improved rules from another aspects. DSmT evidence theory, Yager's and Sun Quan's algorithms are representatives of this kind. The improved method does improve the reasonableness of the fusion results of conflict evidences, but it also results in the decline of convergence speed and the expansion of computation. Up to now, none of the above algorithms has broken through their own limitations and become a recognized general algorithm in this field.

Meanwhile, based on the second idea, scholars take reasonable and effective evidence pretreatment as the direction of exploration. Under the guidance of it, many methods are proposed to measure the degree of conflict between evidences and modify the information. For example, [27] proposes that the belief function of conflict evidences should be allocated to unknown areas to reduce the degree of conflict between focal elements, which also inspires a series of subsequent algorithms. [28] defines the evidence quality measure based on the correlation and complementarity between evidences, and discounts individual evidences based on the measure. [29] introduces a measure function to form a set of evidences and measures the support degree of the set for each individual evidence. [30] uses the idea of adaptive weight to adjust the significance of evidence bodies in the fusion using the weights updated in real time. However, in general, in order to ensure the rationality of the fusion results while taking into

account the clarity of the final decision, the uncertainty of belief function can't be over-amplified when reducing evidence conflicts through pretreatment. Up to now, there hasn't been an accepted improved method which ensures the balance between low conflict and low uncertainty, so the improvement of evidence theory still has research significance and prospects.

1.3 Main Contents and Arrangement of Thesis

1.3.1 Main Contents

Based on the AUV docking platform of cooperative project, the position and attitude estimation of AUV in the process of close docking is studied. Firstly, the basic models of inertial and monocular-vision navigation are studied. The model of monocular-vision positioning is modified appropriately considering yaw and pitch of AUV to improve the positioning accuracy. As for visual image processing, an image segmentation method based on improved D-S evidence theory is proposed, which lays a foundation for precise extraction of lights centroids and position using Blob operators. With regard to integrated position and attitude estimation, a fusion algorithm based on Kalman filter is proposed, which integrates the data from inertial sensor, magnetometer and monocular vision to serve the close-range docking. Finally, the effectiveness of the algorithm is verified by simulation and experiments.

The main contents of this thesis includes:

（1） Study and modification on models of inertial and monocular-vision navigation: The basic principle and calculation models of monocular vision system and nine-axis gyroscope system (including a three-axis magnetometer) are analyzed. In order to compensate for the monocular-vision positioning error caused by AUV attitude, considering the yaw and pitch of AUV, the monocular-vision positioning model is improved to further improve the accuracy.

（2） Image segmentation based on improved D-S evidence theory: In order to achieve effective image segmentation adaptive to the particularity of underwater ambient and ensure the real-time performance, three image segmentation measures are defined, and then the classical D-S evidence theory is modified using TOPSIS interval decision-making

to propose an improved D-S evidence theory method. The basic probability assignments of evidences are fused to evaluate the quality of image segmentation and determine whether current threshold needs to be adjusted. After adaptive image segmentation, the Blob operator is used to extract the centroid of guidance lights for subsequent positioning.

（3） Design of an integrated position and attitude estimation algorithm based on multi-sensor fusion: After obtaining the inertial and monocular-vision measurements, the position and attitude of AUV are estimated by Kalman filter. Due to the inconsistency of update frequencies of different sensors, the output of inertial system is used for system positioning during monocular-vision solution. When the solution of visual sensor is completed, its result is fused with that of inertial system, and the updated result is used to initialize the inertial system at the next step. In this way, the cumulative error induced in the long-term integration of inertial system is compensated.

（4） Simulation, experiments and analysis of results: Based on MATLAB, the simulation program is compiled to verify the integrated attitude and position estimation algorithm. Then the practical experiment of short-range AUV docking is designed and carried out. The inertial navigation data and monocular-vision data are collected during the process. The integrated estimation is carried out according to proposed algorithm, and the effectiveness of the algorithm is evaluated. At last, the causes of the error are analyzed, and the focus of follow-up study is put forward.

1.3.2 Arrangement of Thesis

The thesis is divided into five chapters. Its arrangement is as follows:

The first chapter is the introduction, which mainly introduces the background and significance of study, research status at home and abroad, paper organization and research ideas, focusing on the R&D status of AUV, commonly used measurement methods and their advantages and disadvantages, the characteristics of inertial and monocular-vision navigation, as well as the development direction of inertial-visual integrated estimation.

In the second chapter, the basic principle and model modification of inertial and

monocular-vision measurement systems are introduced. Firstly, the sensor layout scheme of AUV is given. Secondly, the principle of inertial navigation calculation is illustrated, which includes the reason why the quaternion calculation method is chosen. Thirdly, the perspective positioning principle of monocular camera and the modification of positioning model considering yaw and pitch are introduced.

The third chapter presented the image segmentation and extraction of reference lights based on the improved D-S evidence theory. Firstly, given the particularity of underwater image, the commonly used image preprocessing methods are introduced and compared, and the one adopted in this thesis is determined. Next, in order to extract the region of reference lights, three measures of segmentation quality are defined and fused based on the improved DS evidence theory to achieve adaptive adjustment of segmentation threshold. The effectiveness of the algorithm is verified with practical underwater video. Then, the centroids of extracted region, which are also the feature points in positioning, are located by Blob operator, laying a foundation for subsequent integrated estimation.

Chapter 4 mainly introduces the design of integrated position and attitude estimation algorithm, including the fusion scheme design, sensor error modeling, and inertial-visual integrated algorithm design based on Karman filter. The attitude is estimated fusing the data from inertial sensor and magnetometer with a Karman filter based on error-state, while in the position estimation process, the first step is to determine whether the monocular vision sensor is in the process of updating. In the absence of the monocular-vision data, the measurement output of the inertial sensor is used for positioning. If the visual positioning data is updated, it is used to correct the three-axis output of accelerometer by a Kalman filter, to realize an integrated estimation of position.

The fifth chapter is the experiments and results analysis. Based on the actual nine-axis gyroscope chip (consisting of one gyroscope, one accelerometer and one magnetometer) and the AUV docking platform, the inertia-vision integrated estimation algorithm is tested. Then corresponding experimental results are analyzed, and the study prospect is put forward.

1.4 Chapter Summary

In this chapter, firstly the background and significance of this topic and the research status at home and abroad are introduced. Then, the characteristics of inertial navigation system, underwater visual positioning system and inertial-visual integrated positioning algorithm are illustrated. In this part, the advantages of integrated algorithm are paid more attention, and its development trend is also discussed. Finally, the main contents of this thesis are given, then the outline of following chapters are summed up.

Chapter 2 The Models of Monocular Camera and IMU

This chapter mainly focus on the models of monocular camera and inertial measurement unit (IMU). Firstly, the basic scheme of integrated position and pose estimation fusing data of IMU and monocular camera is given. Then main coordinate systems of monocular vision positioning system and their mutual transformation are introduced. On this basis, the perspective positioning principle of monocular vision is introduced, and the general model is modified considering yaw and pitch of AUV. Then, the components of IMU are introduced briefly. The outputs of sub-sensors and the method to solve position and attitude are presented.

2.1 Basic Scheme of Integrated Position and Pose Estimation

In this project, the IMU and monocular camera are used to achieve high precision positioning in AUV docking. Vision positioning is generally used for small-scale and high-precision positioning underwater, but considering that monocular camera spends about several hundred milliseconds in imaging processing, IMU is used as the main measuring system for high frequency updating, while monocular camera and magnetometer are used to correct it. The nine-axis gyroscope studied in this paper refers to an integrated chip consisting of one three-axis gyroscope, one three-axis accelerometer and one three-axis magnetometer. The gyroscope solves the current attitude of AUV by quaternion calculation. After corrected by the data from magnetometer, it is used to correct the output of accelerometer and eliminate the influence of gravity acceleration component to extract the linear acceleration of three axes. The velocity and position of AUV are obtained by once and twice differential respectively. Based on the attitude obtained by IMU, the monocular camera calculates the position according to the principle of visual perspective positioning, and uses it to obtain the position of AUV with higher accuracy. The fused results at the current moment will be fed back for the initialization of next step. The schematic of corresponding sensors on AUV is shown in Figure 11.

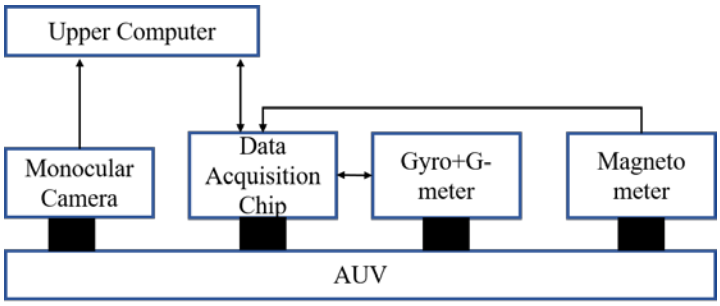


Figure 11 Configuration of Integrated Estimation System

The gyroscope, accelerometer, magnetometer and monocular camera are all installed on the AUV carrier. Video acquired by monocular camera is uploaded to the PC interface through serial port. Data of IMU are collected by data acquisition module and uploaded through serial port. Based on the collected angular velocity, acceleration, attitude angle and visual positioning results, the upper computer first synchronize the data, then update the position and attitude according to the multi-sensor data fusion algorithm, and outputs the estimation which is more accurate and reliable.

2.2 Positioning Model of Monocular Camera

2.2.1 Main Coordinate Systems and Mutual Transformation

The monocular camera positioning system includes three main coordinate systems: image coordinate system, camera coordinate system and world system. As shown in Figure 12, the image coordinate system is located on the imaging plane of monocular camera, and the positioning data need to be converted from that in the pixel coordinate system. The origin of pixel coordinate system is usually located at a corner of the image plane, while the pixel is taken as unit, any point on the image plane is represented by the row number u and column number v. In the positioning, it needs to be converted to data with real physical units.

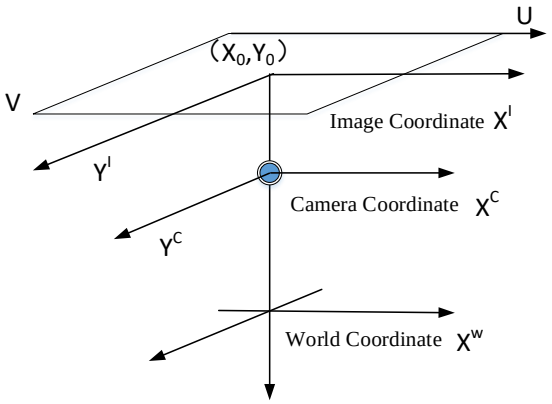


Figure 12 Coordinates of Monocular-camera Positioning System

In this view, the intersection of optical axis and imaging plane is set as the origin of image coordinate system O_I, while the axis crossing it and parallel to *u*-axis is set to be x_I-axis, and the axis crossing it and parallel to *y*-axis is set to be y_I-axis, then the image coordinate system is built。 Assume that the coordinate of O_I in the image coordinate system is (u_o, v_o), while the physical dimensions of each pixel on *u*-axis and *y*-axis are *dx* and *dy* respectively。 Then the relationship between the coordinate of one point in the image coordinate system (x_I, y_I) and that in the pixel coordinate system (u, v) is:

$$\begin{cases} x_I = (u - u_o)\,dx \\ y_I = (v - v_o)\,dy \end{cases} \tag{2.1}$$

which can be rewritten by a homogeneous matrix as:

$$\begin{bmatrix} x_I \\ y_I \\ 1 \end{bmatrix} = \begin{bmatrix} dx & 0 & -u_o dx \\ 0 & dy & -v_o dy \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{2.2}$$

The origin of the camera coordinate is at the optical center of camera O_c, while the horizontal and vertical axes, marked by x_c and y_c respectively, both cross that point and are parallel to the x_I-axis and y_I-axis。 Unlike the two-dimensional image coordinate system, it is necessary to set the z_c-axis overlapping the optical axis, and its direction is from the optical

center to the object. The distance between O_c and O_I is defined as the focal distance of camera，which is represented by f . Then according to the linear camera model, namely the pore imaging model, for any point *P* on the object plane, its coordinate in the camera coordinate is (x_c, y_c, z_c). Given that the coordinate of its image point in the image coordinate is (x_I, y_I)，then based on the similitude theorem，we can get：

$$\begin{cases} x_c = \dfrac{x_I \times z_c}{f} \\ y_c = \dfrac{y_I \times z_c}{f} \end{cases} \tag{2.3}$$

World coordinate is the benchmark for the absolute position of AUV and the camera on it, the relationship between the coordinate of one point in camera coordinate (x_c, y_c, z_c) and that in world coordinate (x_w, y_w, z_w) can be depicted with a homogeneous matrix as bellow:

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{2.4}$$

in which the rotatory matrix *R* is a 3×3 orthogonal unit matrix，representing the rotatory relation between two coordinates, while translation vector *p* is a 3×1 column vector，representing corresponding translation relation.

2.2.2 Principle of Monocular-camera Positioning under ideal conditions

Under ideal conditions, the yaw and pitch angle of AUV are both 0, thus the plane of object is vertical to the optical axis of camera. In this thesis, the docking platform are installed with guiding lights, when the arrangement of guiding lights are fixed, the relative distances between individual lights are constant and known. Therefore, the docking platform can be positioned based on each single image collected by the camera.

The straight distance between the AUV and the front plane of docking platform can be determined with the images of two guiding lights (whose relative positions are known)

collected by the camera. The schematic is given in Figure 13.

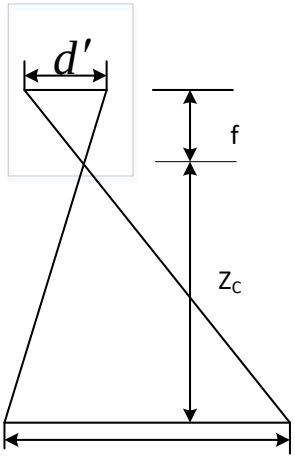


Figure 13 The Schematic of Distance Measurement by Monocular Camera

Assume that in the world coordinate, the distances between two guiding lights are d, while the corresponding distance in the image coordinate is d_I, then the straight distance z_c between the optical center of camera and the plane of object can be calculated by:

$$z_c = d \times f / d_I \tag{2.5}$$

After obtaining the straight distance, the three-dimensional relative position between the camera and docking device can be solved based on it. Assume that the coordinate of one point P on the guiding light plane in the camera coordinate is (x_c, y_c, z_c), while the coordinate of its image in the image coordinate is (x_I, y_I), then the relative position of P in the camera coordinate can be calculated by:

$$\begin{cases} x_c = \dfrac{x_I \times z_c}{f} \\ y_c = \dfrac{y_I \times z_c}{f} \end{cases} \tag{2.6}$$

in which x_c and y_c represent the horizontal and vertical coordinate of P with respect to the optical center of camera.

2.2.3 Compensated Positioning Considering Yaw and Pitch

Although the optical axis is vertical to the plane of object under ideal conditions, in

practical cases, due to the external disturbances in the ambient and the control error of AUV, most of time the optical axis of camera can't stay vertical to the object plane. Under this condition, to guarantee the accuracy of visual positioning, the model has to be compensated by the angles of yaw and pitch.

When the AUV rolls, although the camera also rolls together with the AUV, the image plane and the object plane stay parallel to each other, so the straight distance z_c can still be calculated by Formula (2.5)。On this basis, the coordinate of points in camera coordinate remains the same as that obtained by Formula (2.6). However, at this time there is some angular deviation between the camera coordinate and guidance coordinate, which can be given by

$$\begin{bmatrix} x_n \\ y_n \\ z_n \end{bmatrix} = \begin{bmatrix} \cos\gamma & 0 & \sin\gamma \\ 0 & 1 & 0 \\ -\sin\gamma & 0 & \cos\gamma \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \tag{2.7}$$

On the other hand, when the yaw and pitch angle of AUV is not zero, the image plane and object plane are not parallel to each other anymore, thus the results solved by the model must be compensated in terms of the attitude angles. To illustrate it, when the pitch angle is nonzero, the relative position between the camera and object plane, as well as the pseudo-object plane is given by Figure 14.

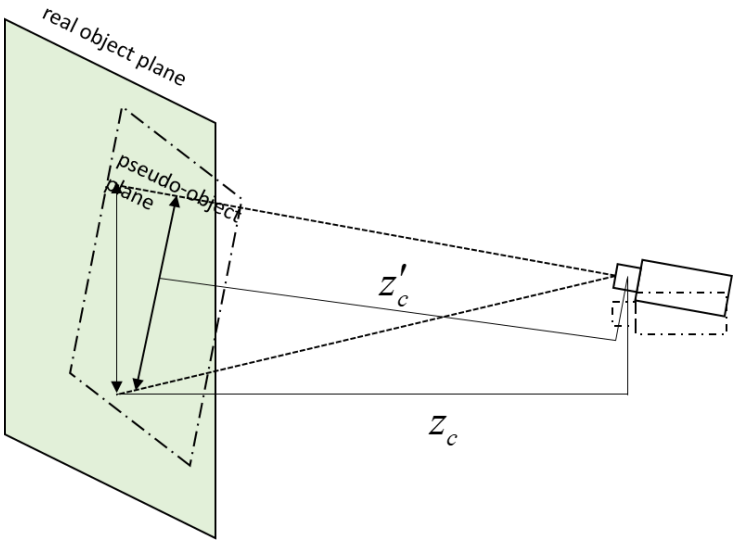


Figure 14 Relative Position of Camera, Object Plane and Pseudo-object Plane

In the virtual rotation analysis of camera, it is assumed that the camera is rotated by an

angle reverse to the pitch direction around its x_c-axis, namely to a new position where its optical axis is vertical to the object plane again. In this project, because that the camera is fixed at the front of AUV and the center of gravity does not coincide with the optical center, when there exists a nonzero pitch angle, the rotation of camera is illustrated by Figure 15.

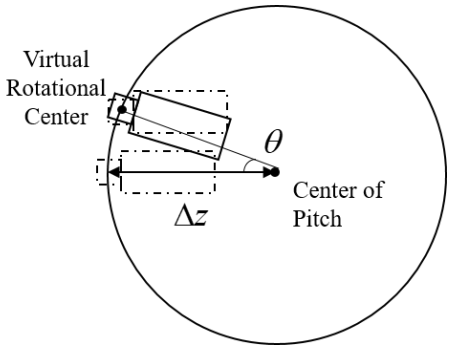


Figure 15 Position of Camera Before and After Rotation

When the pitch angle is θ（positive values for the elevated ones），the camera rotates with the AUV, so its optical center deviates to O_c'， as shown in Figure 16.

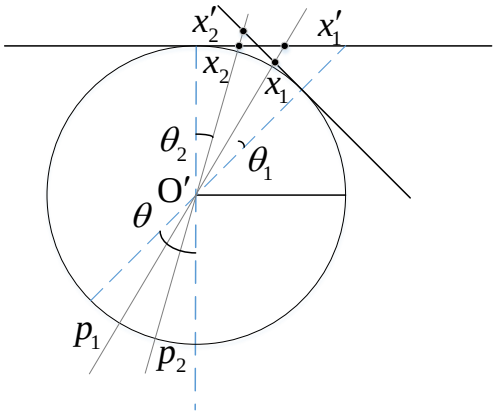


Figure 16 Image of Points in Two Coordinates

After rotation, the image coordinate is (x_I, y_I). Let the camera rotates reversely around the center O_c' and x_c-axis until its optical axis is vertical to the object plane again, then the virtual image coordinate is (x_I', y_I')。Assume that the coordinates of two reference lights p_1 和 p_2 in the coordinate (x_I, y_I) are (x_1, y_1) and (x_2, y_2) respectively，and their coordinates in the virtual image coordinate (x_I', y_I') are (x_1', y_1') and (x_2', y_2'). Given that the angle

between the optical axis and the two lights are θ_1 and θ_2, then according to the geometrical relationship, we can get:

$$\begin{cases} x_1' = -f \times \tan(\theta - \theta_1) \\ x_2' = -f \times \tan(\theta - \theta_2) \end{cases} \tag{2.8}$$

Besides,

$$\begin{cases} x_1 = f \times \tan\theta_1 \\ x_2 = f \times \tan\theta_2 \end{cases} \tag{2.9}$$

On this basis, the coordinates of reference lights on the virtual image plane are:

$$\begin{cases} x_1' = \dfrac{f(x_1 - f\tan\theta)}{f + x_1\tan\theta} \\ x_2' = \dfrac{f(x_2 - f\tan\theta)}{f + x_2\tan\theta} \end{cases} \tag{2.10}$$

When the pitch angle is nonzero, the transverse distance on the image plane is not influenced, at this time, the transverse distance between two reference lights are $\Delta y' = y_2' - y_1' = y_2 - y_1$, while the longitudinal distance is $\Delta x' = x_2' - x_1'$. Then the distance between two reference lights are $d' = \sqrt{(\Delta x')^2 + (\Delta y')^2}$, and the straight distance from the optic center to the object plane in the virtual camera coordinate is $z_c' = \dfrac{d \times f}{d'}$。According to the geometrical relationship, the straight distance between the optic center and the object plane in the real camera coordinate is $z_c = z_c' - \Delta z(1 - \cos\theta)$。

After determining z_c, the three-dimensional coordinate of reference lights can be obtained. Because the transverse positioning is not influenced, so in the virtual camera coordinate $y_c = \dfrac{z_c}{f} y' = \dfrac{z_c}{f} y$ and $x_c = \dfrac{z_c}{f} x'$. When converted to the real camera coordinate, the y_c remains the same but $x_c = \dfrac{z_c}{f} x' + \Delta z \sin\theta$。

Similarly, the three-dimensional position when the yaw angle is nonzero can be determined.

Assume that the positive rotation is counter-clockwise, and the yaw angle is ψ, then the straight distance from the optic center to the object plane is $z_c = \frac{d \times f}{d'} - \Delta z\left(1-\cos\psi\right)$, d and d' can be calculated as above, but in this case the longitudinal distances remains unchanged. Therefore, $\Delta x' = x_2' - x_1' = x_2 - x_1$ and $\Delta y' = y_2' - y_1'$, besides,

$$\begin{cases} y_1' = \dfrac{f\left(y_1 - f\tan\psi\right)}{f + y_1\tan\psi} \\ y_2' = \dfrac{f\left(y_2 - f\tan\psi\right)}{f + y_2\tan\psi} \end{cases} \tag{2.11}$$

in which the meaning of symbols are similar to that in the case of pitch. The compensated coordinates of object can be calculated by $y_c = \frac{z_c}{f}y' + \Delta z\sin\psi$ and $x_c = \frac{z_c}{f}x'$。

2.3 Model of Nine-axes Gyroscope

The AUV used in this project is equipped with a nine-axes gyroscope, namely an inertial sensor and a magnetometer, while the inertial sensor can be subdivided into a gyroscope and an accelerometer, which measure the three-dimensional angular velocity and acceleration respectively. In this thesis, the calculating process of nine-axes gyroscope is shown in Fig.17.

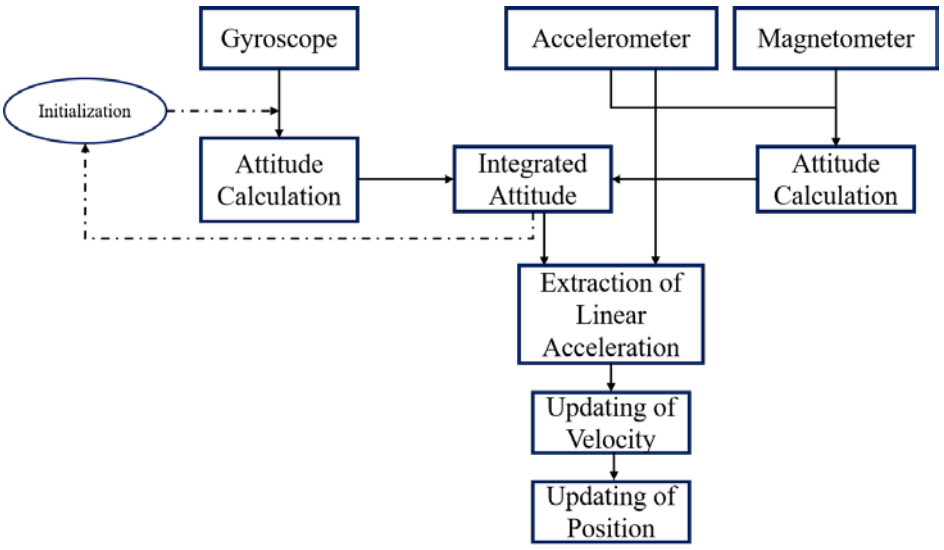


Figure 17 Calculation Process of Nine-axis Gyroscope

In the process, the output of the accelerometer in each axis is the sum of inertial force and corresponding gravitational component [33]. When the AUV is immobile or in the state of

uniform motion, the vector sum of the three outputs of accelerometer is equal to the local gravitational acceleration. On this basis, the pitch and roll angle of AUV can be calculated, but the yaw angle can't be obtained. However, the outputs of accelerometer are quite sensitive to the motion mode, in the state of unbalanced motion, the outputs will be the vector sum of gravitational acceleration and inertial acceleration. If the gravity acceleration is not eliminated, the three-dimensional velocity and displacement will have serious error.

On the premise of eliminating the component of gravity acceleration, the three-axis inertial acceleration measured by accelerometer can be used to calculate the three-axis velocity and displacement by once or twice integration. In the commonly used strap-down inertial navigation systems, the tri-axial angular velocity measured by gyroscope is used to calculate three attitude angles. Based on that, the projection of gravity acceleration on each axis is obtained and eliminated from the output of the accelerometer. That is to say, the tri-axial acceleration which can be used to integrally solve the velocity and displacement is obtained.

In Strap-down Inertial Navigation System (SINS), the data measured by the system is relative to the inertial coordinate system. To establish the transformation relationship between the carrier and the local navigation coordinate system is the task of attitude updating algorithm. Attitude updating algorithm refers to the process of calculating three attitude angles of the carrier relative to the navigation coordinate system using the three-axis angular velocity output by the gyroscope. The commonly used integration methods include Euler angle method, direction cosine method , quaternion method and so on. Although the angle solution algorithm is intuitive, considering that Euler angle and direction cosine method may encounter the problem of universal joint locking in the process of solution, the quaternion solution method is adopted in this thesis. In addition to the avoidance of joint locking, quaternion is easier to get difference and normalization than other methods, and the two rotations of the carrier can be expressed by simply multiplying two sets of quaternions. In this connection, the quaternion solution method is adopted in this subject.

Moreover, A tri-axial magnetometer is also equipped in the AUV, which can determine the deviation angle of AUV relative to the direction of the earth's magnetic field and output a more

accurate yaw angle. Compared with the gyroscope and accelerometer, the advantage of the magnetometer is that it does not need additional coordinate transformation and the output form is simpler. The magnetometer can guarantee the accuracy of its output with no disturbance or only static disturbance, but it will produce relatively large errors when exposed to dynamic electromagnetic interferences. Therefore, when installing the magnetometer, it is necessary to take account the position of the AUV motor and provide appropriate shielding.

After understanding the layout of sensors and their working principles, the solving process of attitude, velocity and position is as follows:

Representing the body coordinate of carrier with *b* and the navigation coordinate with *n*, then the transformation matrix from coordinate *b* to *n* is referred to the attitude matrix of carrier. In essence, the updating of attitude is the process of updating the attitude matrix in terms of the outputs of SINS. When the origins of these two coordinates overlap each other after translation, the relative relationship between the coordinates can be seen as the rotation around one certain point.

The structure of quaternion is as below：

$$Q\left(q_0,q_1,q_2,q_3\right)=q_0+q_1i+q_2j+q_3k \tag{2.12}$$

which can also be rewritten as

$$Q=\cos\frac{\theta}{2}+u^R\sin\frac{\theta}{2} \tag{2.13}$$

to represent the rotation of an rigid body around a fixed point more intuitively. In this way, the u^R is the instantaneous rotational axis and θ is the rotational angle. The attitude matrix from the body coordinate to the navigation coordinate can be calculated based on the quaternion as

$$C_b^n=\begin{bmatrix}1-2\left(q_2^2+q_3^2\right) & 2\left(q_1q_2-q_0q_3\right) & 2\left(q_1q_3+q_0q_2\right)\\ 2\left(q_1q_2+q_0q_3\right) & 1-2\left(q_1^2+q_3^2\right) & 2\left(q_2q_3-q_0q_1\right)\\ 2\left(q_1q_3-q_0q_2\right) & 2\left(q_2q_3+q_0q_1\right) & 1-2\left(q_1^2+q_2^2\right)\end{bmatrix} \tag{2.14}$$

When the yaw, pitch and roll angle of AUV are marked by ψ, θ and ϕ respectively， and

the ENU is chosen as the navigation coordinate, the attitude matrix can also be represented by:

$$C_b^n = \begin{bmatrix} \cos\phi\cos\psi + \sin\phi\sin\psi\sin\theta & \sin\psi\cos\theta & \sin\phi\cos\psi - \cos\phi\sin\psi\sin\theta \\ -\cos\phi\sin\psi + \sin\phi\cos\psi\sin\theta & \cos\psi\cos\theta & -\sin\phi\sin\psi - \cos\phi\cos\psi\sin\theta \\ -\sin\phi\cos\theta & \sin\theta & \cos\phi\cos\theta \end{bmatrix} \quad (2.15)$$

In this connection, after obtaining the attitude matrix base on quaternions, by comparing formulas above, the yaw, pitch and roll angle of carrier can be calculated by:

$$\begin{cases} \theta = \arcsin\left(2\left(q_2q_3 + q_0q_1\right)\right) \\ \phi = \arctan\left(-\dfrac{2\left(q_1q_3 - q_0q_2\right)}{1-2\left(q_1^2 + q_2^2\right)}\right) \\ \psi = \arctan\left(\dfrac{2\left(q_1q_2 - q_0q_3\right)}{1-2\left(q_1^2 + q_3^2\right)}\right) \end{cases} \quad (2.16)$$

To update the quaternions in terms of the outputs of IMU, the following differential equations of quaternions are needed:

$$\begin{bmatrix} \dot{q}_0 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{bmatrix}\begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} \quad (2.17)$$

Considering the low velocity of AUV underwater, the additional terms related to the carrier velocity and earth rotation are quite small, so ω in the following formula can be seen as the angular velocity output by the gyroscope. The angle increment measured by the gyroscope is

$$\Delta\Theta \approx \begin{bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{bmatrix} dt = \begin{bmatrix} 0 & -\Delta\theta_x & -\Delta\theta_y & -\Delta\theta_z \\ \Delta\theta_x & 0 & \Delta\theta_z & -\Delta\theta_y \\ \Delta\theta_y & -\Delta\theta_z & 0 & \Delta\theta_x \\ \Delta\theta_z & \Delta\theta_y & -\Delta\theta_x & 0 \end{bmatrix} \quad (2.18)$$

Assume that $\Delta\theta^2 = \Delta\theta_x^2 + \Delta\theta_y^2 + \Delta\theta_z^2$, then the first-order formula of quaternion updating is

$$\begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} = \left(I + \frac{1}{2}\Delta\Theta \right) \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} \tag{2.19}$$

In terms of the precision demand, second-,third- even fourth-order algorithms can also be used, and the updating formula of fourth-order algorithm is

$$\begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} = \left(I\left(1 - \frac{\Delta\theta^2}{8} + \frac{\Delta\theta^4}{384} \right) + \left(\frac{1}{2} - \frac{\Delta\theta^2}{48} \right)\Delta\Theta \right) \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} \tag{2.20}$$

At the same time, one set of attitude angel can be determined based on the outputs of accelerometer and magnetometer by the following formula

$$\begin{cases} \phi = \arctan\dfrac{f_x}{f_z} \\ \theta = \arctan\dfrac{f_y}{f_z} \\ \psi = \arctan\left(-\dfrac{M_y\cos\phi - M_z\sin\phi}{M_x\cos\theta + M_y\sin\theta\sin\phi + M_z\sin\theta\cos\phi} \right) \end{cases} \tag{2.21}$$

Considering the serous integral error of gyroscope and strong sensitivity of accelerometer and magnetometer to the noise，these two sets of attitude angles are fused to form the final attitude estimation by Kalman filter。Considering the practical application, after determining the attitude matrix, the three-axial velocity and displacement can be updated by

$$V_n = V_{n-1} + C_b^n \times \int_{t_{n-1}}^{t_n} f dt \tag{2.22}$$

$$P_n = P_{n-1} + \int_{t_{n-1}}^{t_n} V dt \tag{2.23}$$

To sum up, the updating process of the nine-axes gyroscope is as below. Firstly, the angular velocity data measured by gyroscope and the acceleration data measured by accelerometer in relative inertial coordinate system are transformed into the navigation coordinate system, and then the attitude angles are calculated. The calculation results of gyroscope are modified by the set of attitude angle calculated by the accelerometer and magnetometer, then the line velocity

and displacement are obtained. With the improvement of computer processing ability, the attitude estimation can be updated at high frequency by simple iteration. Therefore, when calculating the velocity and position, the relatively simple and feasible trapezoidal integration method is chosen.

2.4 Chapter Summary

This chapter introduces the scheme of vision-inertia integrated position and attitude estimation briefly. Firstly, the modified positioning model of monocular-camera under ideal and non-ideal conditions are introduced. Then, the structure, outputs and fusion idea of nine-axis gyroscope are introduced, and they also build the foundation of contents in the following chapters.

Chapter 3 The Light Positioning Algorithm Based on D-S Evidence Theory

This chapter focuses on the area extraction and centroid location of underwater light sources. Firstly, the imaging characteristics of underwater reference lights and several basic image preprocessing methods are introduced. Then several classical edge extraction methods are introduced and on the basis of comparing their extraction effects, a light-source region extraction method based on improved D-S evidence theory is proposed, and that is also the most important part of this chapter. After the extraction of light-source region, the centroid location process based on Blob analysis is introduced, and the locations are combined with the content of the previous chapter, to obtain the position of AUV relative to the docking device.

3.1 Characteristics of Underwater Image and Image Preprocessing Methods

Because of the density and optical characteristics of water itself, the imaging effect of light sources underwater is often much worse than that in air. With the increase of water turbidity, the attenuation of water to light will increase exponentially. A large amount of energy of light source will be dissipated by absorption and scattering of water, and the attenuation ratio caused by the above two terms in clear sea water is about 1:1 [35]. In order to increase illumination as much as possible and provide reference for the docking of AUV, we use incandescent lamps as the guiding lights in this thesis. By adjusting them to an appropriate brightness, we can control the scattering and ensure the image quality of monocular camera. In addition, considering that the image of underwater lights has some undesirable problems, such as blurred contour, uneven gray scale from the center to the edge, and the almost evitable noise interference in the transformation and transmittal of image, in order to ensure the quality of the image and better serve the centroid extraction and subsequent fusion, it is necessary to pretreat the original image.

In the AUV docking process, filtering can effectively remove part of the noise due to subjective and objective factors. Considering the computational burden and the applicability of actual platform, the de-noising effect of basic spatial filtering is sufficient in this thesis. The

commonly used image filtering algorithms are mean filtering, median filtering and Gaussian filtering [36]. The following is a brief introduction to them, and the effects of the three smoothing filtering algorithms are compared using the actual underwater images.

3.1.1 Mean Filtering

In the mean filtering, a template is built for each pixel (x, y) in the original image, whose center is the pixel itself and the size is $M \times N$. The gray value average of all the pixels in the template is set as the filtered gray value of the center pixel, and the formula is as below:

$$gray(x, y) = \frac{1}{M \times N} \sum_{(i,j) \in n(x,y)} gray(i, j) \tag{3.1}$$

where $n(x, y)$ is the neighboring region of pixels centered by pixel (x, y), including M rows and N columns.

Mean filter replaces the gray value of each pixel in the original image with the average gray value of all the pixels in the neighboring region. It has a good effect in restraining the abrupt change of gray value. Meanwhile, because of its simple calculation formula and low computational complexity, it has been widely used in practical engineering. In principle, the larger the size of the template is, the more obvious the effect of mean filtering will be. However, with the expansion of the template, some details in the image will be destroyed, and the edge features in the image will be weakened, resulting in blurred edges and difficulty in identifying. In order to verify the effect of mean filtering, we added Gaussian noise and salt-and-pepper noise to underwater images, and process them with mean filtering with different template sizes. The comparison between results before and after filtering is shown in Figure 18.

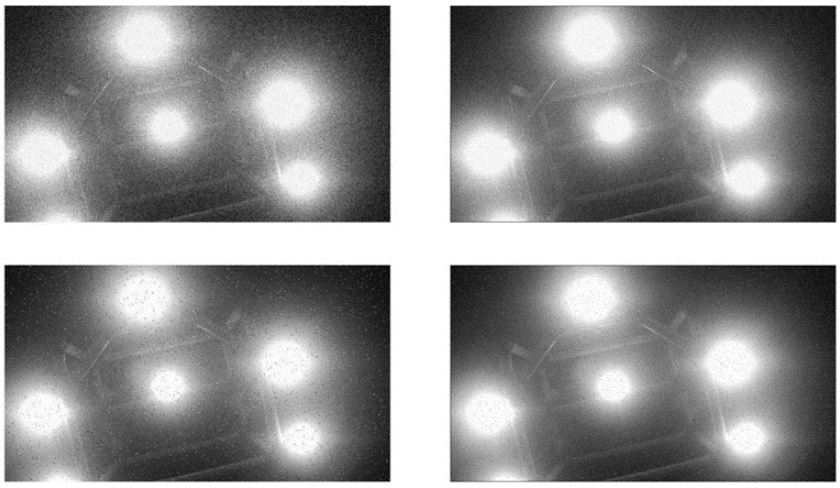


Figure 18 Effects of Mean Filtering (Top Left: Original Image Mixed by Gaussian Noise; Top Right: Filtered Image of Top Left; Bottom Left: Original Image Mixed by Pepper&Salt Noise; Bottom Right: Filtered Image of Bottom Left)

By comparing above results, we can find that the mean filter has some effect on Gaussian noise, but when salt-and-pepper noise is mixed in the image, the suppression of noise is not satisfactory. Generally speaking, no matter whether the underwater image is mixed with Gaussian noise or salt-and-pepper noise, the mean filtering can't get a relatively ideal filtering effect, which is probably because the attenuation of light underwater reduces the contrast of the image. Therefore, mean filtering it is not an ideal image de-noising and smoothing method for images underwater.

3.1.2 Median Filtering

Similar to mean filtering to some degree, the median filtering also needs a template centered by the pixel (x, y), whose common size is 3×3 or 5×5. After obtaining all the gray values within that template and ranking them , the median of gray value is assigned to the centered pixel (x, y). The results of median filtering to images mixed with Gaussian noise and salt-and-pepper noise are given by Fig.19.

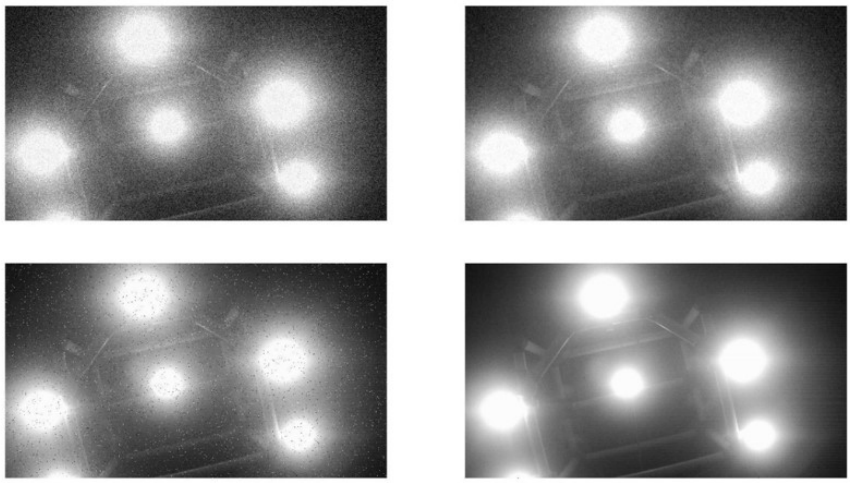


Figure 19 Effects of Median Filtering (Top Left: Original Image Mixed by Gaussian Noise; Top Right: Filtered Image of Top Left; Bottom Left: Original Image Mixed by Pepper&Salt Noise; Bottom Right: Filtered Image of Bottom Left)

Observing the results above, we can find that median filtering has a good effect on salt-and-pepper noise, but has little effect on Gaussian noise. However, by comparing the results of median filter and mean filter, it is obvious that median filter can preserve the edge information better, and the damage and fuzzification to the details of the image is relatively small.

3.1.3 Gaussian Filtering

Just as the filtering methods above, template (whose common size is 3×3 or 5×5) is necessary in Gaussian filtering. However, the main idea of Gaussian filtering is to adjust the weight of each pixel in the template, assigning higher weights to the pixels near the center, reducing the influence of pixels far from the center, and to use the weighted calculation of gray value as the filtered results. The formula of two-dimensional Gaussian function is as follows:

$$gray(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3.2}$$

The weight assignment of template used (5×5) is

$$\frac{1}{273}\times\begin{bmatrix}1 & 4 & 7 & 4 & 1\\ 4 & 16 & 26 & 16 & 4\\ 7 & 26 & 41 & 26 & 7\\ 4 & 16 & 26 & 16 & 4\\ 1 & 4 & 7 & 4 & 1\end{bmatrix} \tag{3.3}$$

The results of Gaussian filtering to images mixed with Gaussian and salt-and-pepper noise are given by Figure 20.

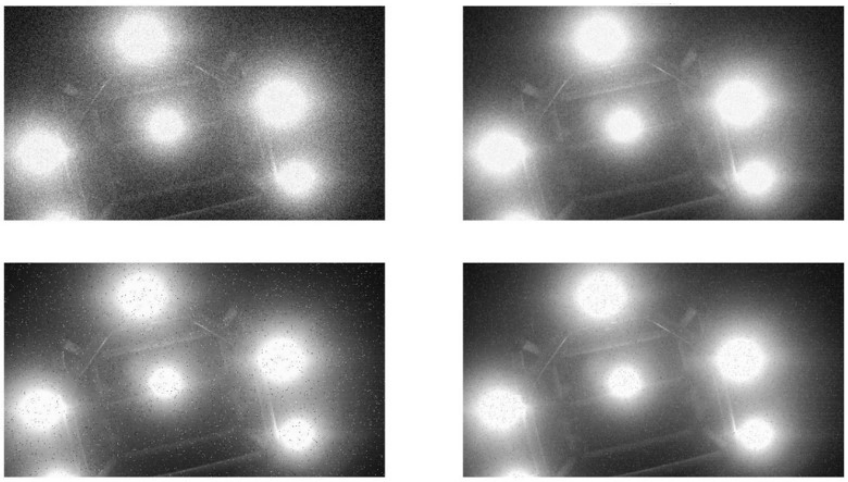


Figure 20 Effects of Mean Filtering (Top Left: Original Image Mixed by Gaussian Noise; Top Right: Filtered Image of Top Left; Bottom Left: Original Image Mixed by Pepper&Salt Noise; Bottom Right: Filtered Image of Bottom Left)

Based on the comparison between above results, it can be found that the Gaussian filter has a good suppression effect on both Gaussian noise and salt-and-pepper noise, as well as retaining the edge information and details. Besides, the intensity of the filter can also be adjusted by changing the parameter σ， so the filtering effect is more ideal. Therefore, in the follow-up contents of this thesis, the images are smoothed and de-noised using Gaussian filtering method, to better serve the segmentation of light-source region and centroid extraction.

3.2 Feature Extraction of Underwater Light-source Images

In the process of AUV docking, estimating the relative position between AUV and the docking device needs to obtain the position parameters of the guiding light-source. Considering the serious light scattering underwater, to guarantee the accuracy of positioning, firstly the idea

of extracting the guiding light contour is chosen, and the extracted contour serves as the important basis of centroid location. Commonly used contour extraction operators include first-order gradient, Sobel operator, Robert operator, Canny operator, etc. [37]. Given the particularity of underwater imaging environment, this thesis proposes an edge extraction algorithm based on improved D-S evidence theory, and its principle will be introduced below.

3.2.1 An Overview of Image Segmentation Methods

At present, there are many image segmentation methods, but in general, the idea can be divided into three kinds: threshold segmentation, edge detection and region-based segmentation [38]. The precondition for threshold segmentation is that the gray value of the pixels within the target and background regions themselves are similar and relatively even, but the gray value of the pixels between different regions has obvious difference, and there is obvious gray gradient on the boundary of the two regions. In commonly used histogram segmentation method, by analyzing the gray histogram distribution of the image, the gray value peaks corresponding to the target and background regions are found. Then the valleys between the two peaks are selected as the thresholds to segment the foreground and background. According to the specific number of thresholds, they can be divided into single-threshold and multi-threshold methods. Early threshold segmentation methods include iteration, maximum variance between classes, etc. Later, with the collision and fusion between other fields and image segmentation, some concepts such as fuzzy mathematics were introduced into this field, drawing forth a series of new segmentation methods [39].

The main idea of edge detection is to convolute the differential template with the image, find the pixels with large gradient of gray value in the image, and combine them as the image edge, so as to achieve the segmentation between the foreground and the background. However, when the image is mixed with noise, the edge points obtained by differential templates are likely to be discontinuous, so in order to separate the closed foreground area from the background, it is usually necessary to use some algorithm such as Hough transform to connect the discontinuous edge points into closed contours. Region-based image segmentation mainly

includes region growth and synthesis. In practical applications, it is often combined with edge detection to achieve image segmentation.

3.2.2 Definition of Edge Gradient Operators

After using several common edge extraction operators to extract the edge of underwater image smoothed by Gaussian filtering, it is found that the effects of common operators are not satisfactory enough. The edge extracted can be incomplete, or there are a lot of interference edges, which is difficult to meet the needs of subsequent centroid location. Therefore, we introduce three edge gradient operators, SUM measure, uniform measure and structural measure. Firstly, the horizontal and vertical original gradients of the gray image are calculated, and the geometric average value of each pixel is obtained. Then, the pixels are pre-screened according to the average value of the gradient, so as to reduce the subsequent computational burden. For each point in the image (x, y), its horizontal and vertical gray gradients are $G_x(x, y)$ and $G_y(x, y)$ respectively, then the three operators are defined as below：

SUM operator： $SUM(x, y) = |G_x(x, y)| + |G_y(x, y)|$

Before defining the uniform and structural operator, some subsidiary variables are introduced. For each pixel (x, y), we define

$$f_{11} = gray(x-1, y-1) + gray(x, y-1) + gray(x+1, y-1); \tag{3.4}$$

$$f_{21} = gray(x-1, y+1) + gray(x, y+1) + gray(x+1, y+1); \tag{3.5}$$

then the subsidiary variable of horizontal measure $r_1 = \frac{1}{3}d_1 = \frac{1}{3}|f_{11} - f_{21}|$。

Similarly, we can define

$$f_{12} = gray(x-1, y-1) + gray(x-1, y) + gray(x-1, y+1); \tag{3.6}$$

$$f_{22} = gray(x+1, y-1) + gray(x+1, y) + gray(x+1, y+1); \tag{3.7}$$

then the vertical measure $r_2 = \frac{1}{3}d_2 = \frac{1}{3}|f_{12} - f_{22}|$。

$$f_{13} = gray(x+1, y) + gray(x+1, y+1) + gray(x, y+1); \tag{3.8}$$

$$f_{23} = gray(x-1, y-1) + gray(x-1, y) + gray(x, y-1); \tag{3.9}$$

45° diagonal measure $r_3 = \frac{1}{3} d_3 = \frac{1}{3} \left| f_{13} - f_{23} \right|$。

$$f_{14} = gray(x, y-1) + gray(x+1, y-1) + gray(x+1, y); \tag{3.10}$$

$$f_{24} = gray(x-1, y) + gray(x-1, y+1) + gray(x, y+1); \tag{3.11}$$

and 135° diagonal measure $r_4 = \frac{1}{3} d_4 = \frac{1}{3} \left| f_{14} - f_{24} \right|$。

Besides these, there are four intermediate parameters as below

$$t_1 = gray(x-1, y) + gray(x, y) + gray(x+1, y)\text{；} \tag{3.12}$$

$$t_2 = gray(x, y-1) + gray(x, y) + gray(x, y+1)\text{；} \tag{3.13}$$

$$t_3 = gray(x-1, y+1) + gray(x, y) + gray(x+1, y-1)\text{；} \tag{3.14}$$

$$t_4 = gray(x-1, y-1) + gray(x, y) + gray(x+1, y+1)\text{；} \tag{3.15}$$

On this basis, we can define

Uniform operator： $R(x, y) = \max\left[r_1, r_2, r_3, r_4\right]$, in which $r_k = \frac{1}{3} d_k, k = 1, 2, 3, 4$.

Structural operator： $C(x, y) = \max\left[l_1, l_2, l_3, l_4\right]$, in which $l_k = \frac{1}{3} t_k, k = 1, 2, 3, 4$.

To synthesize the three operators defined above to judge whether a pixel belongs to the edge of image, D-S evidence theory is introduced in this thesis, and an improved D-S evidence algorithm is proposed to overcome the drawbacks of classical D-S evidence theory in dealing with conflicting evidence. The next part will explain the principle of D-S evidence theory and the improved algorithm.

3.3 Improved D-S Evidence Theory Algorithm

3.3.1 Basic Concepts of D-S Evidence Theory

D-S evidence theory was first proposed by Dempster in 1967 to solve the problem of multi-valued mapping. Because of its outstanding advantages in uncertain reasoning, such as distinguishing uncertainties and unknowns and requiring less prior knowledge, it has been widely used in target recognition, decision-making and fault diagnosis. The combination rule

of evidence is the core of D-S evidence theory, which directly determines the effect of the algorithm. However, when dealing with evidences that are conflicting with each other, the traditional combination rule may get wrong results which are contrary to intuition. Therefore, many scholars have been trying to solve this problem starting from two main ideas: modifying the combination rule and preprocessing evidences. Various improvement schemes have been put forward. Before explaining the problem of traditional evidence combination rule and the improved method in detail, some basic concepts are introduced.

In classic D-S theory, the frame of discernment is defined as an exhaustive set whose elements are mutually exclusive from each other. In this thesis, the frame of discernment is denoted by Θ hereafter，and the *n* elements within it are notified by $x_1, x_2, \cdots, x_n$. On this basis, the power set 2^{Θ} contains all the subsets of discernment framework，and can be expressed by $2^{\Theta}=\{\varnothing,\{x_1\},\{x_2\},\cdots,\{x_n\},\{x_1,x_2\},\cdots,\Theta\}$. Then, through a mapping $m(\bullet)$ named by Basic Probability Assignment(BPA)，each element of the power set can be mapped to value within the interval $[0,1]$. For the element A of the power set，mapped by $m(\bullet)$，its corresponding BPA can be denoted by $m(A)$. In the D-S evidence theory, for each evidence, the BPA of empty set is 0 and the sum of all BPAs are equal to 1. Specially, $m(\Theta)$ represents the amount of belief that can't be assigned to any specific subset of discernment framework, namely it represents the possibility that the target must be one of the $x_1, x_2, \cdots, x_n$ but it's not clear which one it is, so in essence it depicts the uncertainty of BPA. When $m(A)$ is nonzero, A is also called a focal element. Focal elements can be divided into singleton element and non-singleton element. The former includes only one element while the latter can have more than one elements. Obviously, non-singleton elements in the framework increase the uncertainty of data, but the existence of non-singleton elements is pretty common in practice.

Base on $m(\bullet)$, there are two important functions in the D-S evidence theory, which are the belief function *Bel* and plausibility function *Pl*. For each focal element A, its belief

function can be calculated by：

$$Bel(A)=\sum_{B\subseteq A} m(B) \quad \forall A\subseteq \Theta \tag{3.16}$$

which is a measure of justified support for corresponding A. The plausibility function can be expressed by:

$$Pl(A)=\sum_{A\cap B\neq\varnothing} m(B) \quad \forall A\subseteq \Theta \tag{3.17}$$

and it is a measure of potential support for corresponding A. Generally speaking, the difference between $Pl(A)$ and represents the level of uncertainty. In the improved D-S method proposed in this thesis, in the preprocessing of evidences the $Pl(A)$ and $Bel(A)$ form the upper and lower limits of the interval. The specific contents of TOPSIS theory will be introduced in the following part.

3.3.2 Classical D-S Combination Rule and The Discounting Method

Assuming that $m_1(\bullet)$ and $m_2(\bullet)$ are derived from two diverse data sources, then their combine BPAs can be attained in accordance with classical Dempter's rule as below：

$$m(A)=\frac{1}{K}\times\sum_{x\cap y=A} m_1(x)\times m_2(y) \quad \forall A\subseteq \Theta \tag{3.18}$$

where K denotes the conflicting factor whose value is in the interval $[0,1]$, reflecting the degree of incompatibility between two BPAs and can be calculated by：

$$K=\sum_{\cap A_i\neq\varnothing}\prod_{1<i<n} m_i(A_i) \tag{3.19}$$

The larger the value of K is, the severer the contradiction between two evidences is。

Although classical D-S combination rule can achieve evidence fusion, it may make mistakes in dealing with highly conflicting data [42]. To illustrate this problem, the famous counterexample proposed by Zadeh in 1986 is introduced. Assume that the framework of discernment $\Theta=(A,B,C)$, while the $m_1(\bullet)$ and $m_2(\bullet)$ are as below：

$$\begin{aligned} &m_1: m_1(A)=0.99, m_1(B)=0.01, m_1(C)=0\\ &m_2: m_2(A)=0, m_2(B)=0.01, m_2(C)=0.99 \end{aligned} \tag{3.20}$$

where the values of unlisted terms are zero. It can be seen that these two evidences are highly conflicting. The combined BPA attained by Dempster's rule is $m(A)=0, m(B)=1, m(C)=0$. Through simple analysis we can find that, evidence 1 strongly supports focal element A and evidence 2 supports focal element C. However, the combined BPA indicates that the target must be B, which is undoubtedly unreasonable.

In order to solve the problem above, Shafer proposed an evidence-discounting method [43], namely preprocessing the evidences with conflicts by discounting and then combining them still in terms of the classical combination rule. Assume that α is the discounting factor for certain evidence（whose value is in the interval $[0,1]$）, then the original BPAs can be modified in accordance with following formula:

$$m_{\alpha}(A)=\begin{cases}\alpha\cdot m(A) & A\neq\Theta\\ 1-\alpha\cdot(1-m(A)) & A=\Theta\end{cases} \tag{3.21}$$

Broadly speaking, the value of discounting factor reflects the quality of evidence, in that evidences with larger discounting factors play more crucial parts in the data fusion.

3.3.3 The Improved D-S Evidence Algorithm Based on TOPSIS

In practice, lacking of prior information, the confidence of evidence sources can't be estimated directly in most cases. In order to evaluate the quality of each evidence and allocate corresponding discount coefficients, this paper uses some ideas of TOPSIS decision-making [44]. The calculation of discounting factors can be roughly divided into four steps as below:

Step 1：Build the evaluation matrix. Based on the BPAs of evidences, all the belief and plausibility functions of focal elements are obtained and coupled into interval numbers, forming an evaluation matrix X. The columns of matrix correspond to the elements of power set while the rows correspond to evidences. Namely, for an evidence set containing M evidences and N elements in the discernment framework，the size of X is $M\times\sum_{i=1}^{N}C_N^i$。

Step 2：Standardize the evaluation matrix. In the situation of target recognition

(generalized), in order to standardize the matrix, all the values are mapped to a number in the interval $[-1,1]$. The largest value of each column is mapped to 1 while the least value of each column is assumed to be -1. Intermediate values are mapped to the interval $[-1,1]$ proportionally.

Step 3：Construct the best and the worst evaluation, calculate the correlation coefficient of each evidence with regard to these two benchmarks. The best and the worst evaluation are both row vectors with $\sum_{i=1}^{N} C_N^i$ columns. For each focal element, the upper/lower bound of best evaluation is equal to the average value of all the upper/lower bounds of this column. For instance, the j_{th} column of the best evaluation can be expressed by:

$$E_j^+ = \left[\overline{a}_{ij}, \overline{b}_{ij}\right]\left(i = 1,2,\cdots,M;\ j = 1,2,\cdots,\sum_{i=1}^{N} C_N^i\right) \tag{3.22}$$

Similarly, for each focal element, the worst evaluation is composed by the interval number deviating the most from the best evaluation. The j_{th} column of the worst evaluation can be computed by:

$$E_j^- = \left[a_{ij}, b_{ij}\right] = ArgMax\left(\left|a_{ij} - \overline{a}_{ij}\right| + \left|b_{ij} - \overline{b}_{ij}\right|\right)\left(i = 1,2,\cdots,M;\ j = 1,2,\cdots,\sum_{i=1}^{N} C_N^i\right) \tag{3.23}$$

After obtaining the benchmarks, the comprehensive distance between each evidence and the best evaluation is:

$$D_i^+ = \sum_{j=1}^{P} d\left(X_{ij}, E_j^+\right) \quad P = \sum_{i=1}^{N} C_N^i \tag{3.24}$$

in which $d\left(X_{ij}, E_j^+\right)$ denotes the Euclid distance between two zone numbers and can be computed by：

$$\begin{aligned} &d\left(z_1, z_2\right) = \sqrt{\left(a_1 - a_2\right)^2 + \left(b_1 - b_2\right)^2} \\ &z_1 = \left[a_1, b_1\right] \quad z_2 = \left[a_2, b_2\right] \end{aligned} \tag{3.25}$$

Similarly, the formula of comprehensive distance relative to the worst evaluation is:

$$D_i^- = \sum_{j=1}^{P} d\left(X_{ij}, E_j^-\right) \quad P = \sum_{i=1}^{N} C_N^i \tag{3.26}$$

Step 4: Compute and normalize the comprehensive correlation degree, and assign weights to all evidences. The correlation degree of each evidence comparing with the best and the worst evaluations can be given by the formula below:

$$C_i = \frac{D_i^-}{D_i^- + D_i^+} \quad \left(i = 1, 2, \cdots, M\right) \tag{3.27}$$

And the comprehensive correlation degree after standardization is:

$$\varepsilon_i = \frac{C_i}{\max\left(C_i\right)} \tag{3.28}$$

In this connection, these comprehensive correlation degree values are used as discounting factors for corresponding evidences. Then all the discounted evidences are fused in accordance with the classical Dempster's rule to get the final BPA.

To verity the effectiveness of proposed method, Case 1 is simulated and Table 1 shows the combined BPA obtained by classical Dempster's rule, Yager's method， method proposed in reference [10] and method proposed in this paper. From the Table I it can be concluded that although evidence 2 is highly conflict with other evidences, after fusion the proposed method can recognize target as A correctly. The probability assignment of target A is depicted in Figure 21 and it can be seen that the m(A) of proposed method increases monotonically and remains the highest among four results. It can be concluded that the method proposed here has satisfying performance when coping with conflicting evidences.

Case 1：

$$\begin{aligned}
&m_1 : m_1\left(A\right) = 0.5, m_1\left(B\right) = 0.2, m_1\left(C\right) = 0.3;\\
&m_2 : m_2\left(A\right) = 0, m_2\left(B\right) = 0.9, m_2\left(C\right) = 0.1;\\
&m_3 : m_3\left(A\right) = 0.55, m_3\left(B\right) = 0.1, m_3\left(A, C\right) = 0.35;\\
&m_4 : m_4\left(A\right) = 0.55, m_4\left(B\right) = 0.1, m_4\left(A, C\right) = 0.35;\\
&m_5 : m_5\left(A\right) = 0.6, m_5\left(B\right) = 0.1, m_5\left(C\right) = 0.3.
\end{aligned} \tag{3.29}$$

Table 1 Fusion Results of Case 1 and Comparison

Methods	Step	Combined BPAs		
		m(A)	***m(B)***	***m(C)***
Classical D-S	**1**	0	0.8571	0.1429
	2	0	0.6316	0.3684
	3	0	0.3288	0.6712
	4	0	0.1228	0.8772
Yager's Method[47]	**1**	0	0.1800	0
	2	0	0.0180	0.0105
	3	0	0.0018	0.0037
	4	0	0.0002	0.0013
Method of [48]	**1**	0	0.8571	0.1429
	2	0.6280	0.1460	0.1920
	3	0.8710	0.0140	0.0670
	4	0.9230	0.0050	0.0600
Method Proposed	**1**	0.5	0.2	0.3
	2	0.7543	0.0513	0.1944
	3	0.8911	0.0103	0.0986
	4	0.9444	0.0020	0.0536

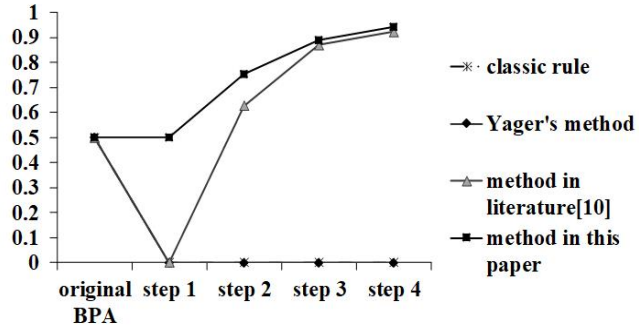


Figure 21 The Probability Assignment of $m(A)$

3.3.4 The Abstraction of BPA and Localization of Guiding Lights

To extract the contour of guiding lights using the three operators defined comprehensively, the improved D-S evidence algorithm is used. In order to abstract the BPAs from the values of operators, we define the mapping function curve as shown in Figure 22.

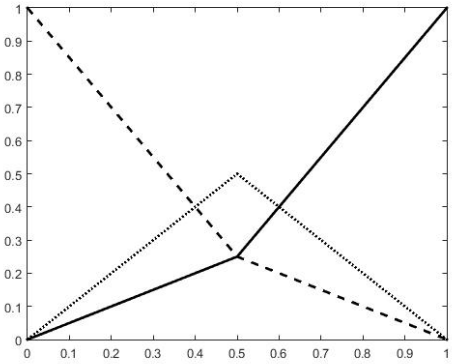


Figure 22 Mapping Curve between Values of Operators and BPA

Based on preliminary experiments and experience, the thresholds of all three operators are set to be 0.5. The closer the measure value is to 1, the more possible that it doesn't belong to the edge, and the less possible it belongs to the edge. At the same time, some BPA characterizes the uncertainty, and the sum of the BPAs is constant to 1. The closer the value of operator approaches the threshold 0.5, the higher the probability assigned to uncertainty is, so as to minimize the probability of misjudgment. On the basis of the BPA defined above, the improved D-S evidence method is used to fuse the three operators to extract the edge pixels. Considering the light scattering, an image with low contrast inside is selected, and the result of edge extraction using the improved method is shown in Figure 23.

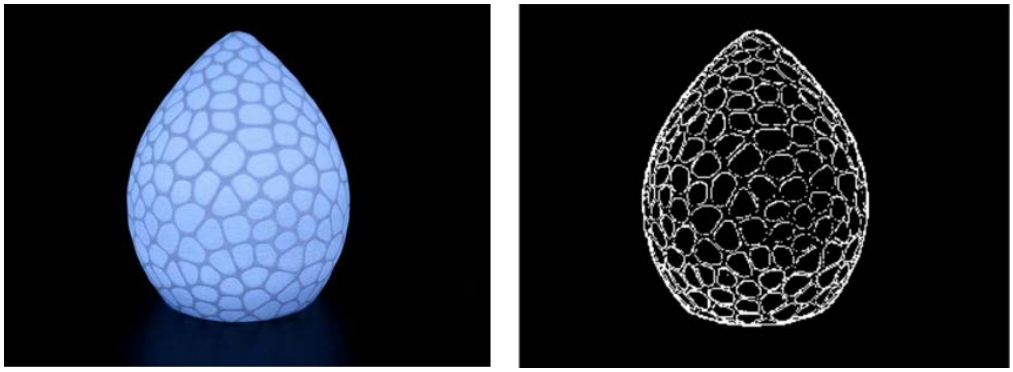


Figure 23 Example Image and Results of Edge Extraction

However, when using the improved edge extraction algorithm for underwater image segmentation and centroid localization, it is found that the computational burden of the algorithm is relatively large, given that the algorithm performs basic probability allocation and fusion for each pixel in the image. Considering the hardware configuration of the actual

verification platform, although the algorithm has good image segmentation effect, it can cause a lot of computational burden to the system, and may lead to real-time problems. In order to further improve the real-time performance of the algorithm, based on the analysis of actual experimental environment, hardware and software conditions, an image segmentation method with self-adaptive threshold is proposed based on the improved D-S evidence theory.

As mentioned above, the Otsu method is a commonly used threshold segmentation method. After many years of development, it has developed from the original one-dimensional to two-dimensional and even three-dimensional. Compared with the basic algorithm, the improved two-dimensional Ostu algorithm improves the anti-noise performance, but it needs to solve the trace of every pixel matrix in the operation, thus the computational burden expands rapidly. In fact, the time complexity of the algorithm will reach the fourth power of the image gray level, therefore it is also not suitable for the experiments. One-dimensional Ostu algorithm is commonly used in threshold segmentation, and it still shows good results in many fields, but its anti-noise performance is relatively poor. Besides, one single threshold can not adapt to the changes of light, and its accuracy may be not enough in the cases where the foreground and background contrast not too much, such as the cases with atomization, light scattering and so on. The experimental image underwater is segmented using one-dimensional Ostu algorithm and the results are given in Figure 24. For better analysis, Figure 25 shows the gray distribution of the image.

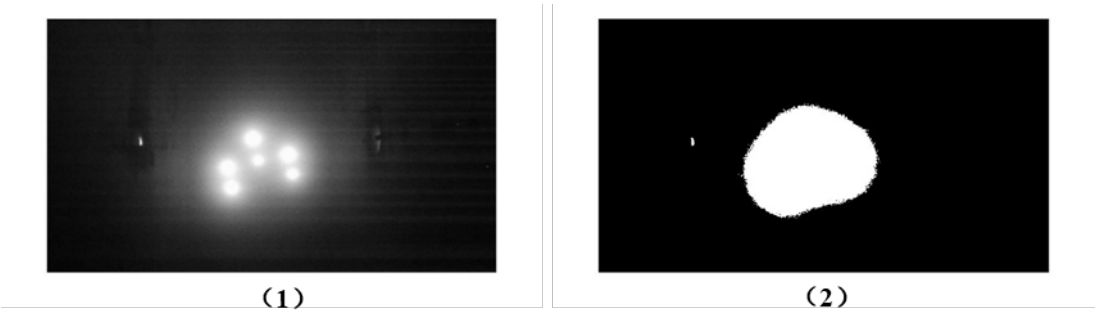


Figure 24 Results of Image Segmentation by One-dimensional Ostu Algorithm

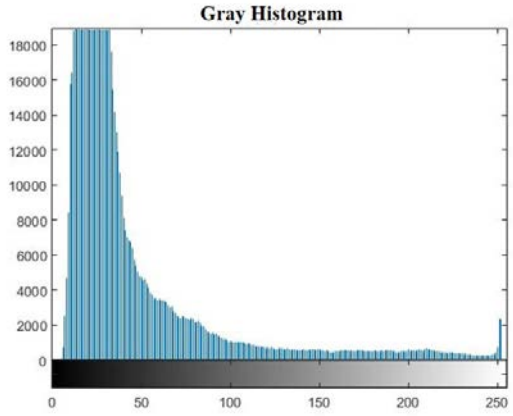


Figure 25 Gray Histogram of Image

It can be seen that due to the scattering of light, the second gray peak of gray histogram is not obvious enough. In this case, with one-dimensional Otsu segmentation, the six guiding lights can not be distinguished completely, let alone using them for subsequent three-dimensional positioning. For further comparison, fixed thresholds are set manually according to the experimental experience. The results of segmentation are shown in Figure 26.

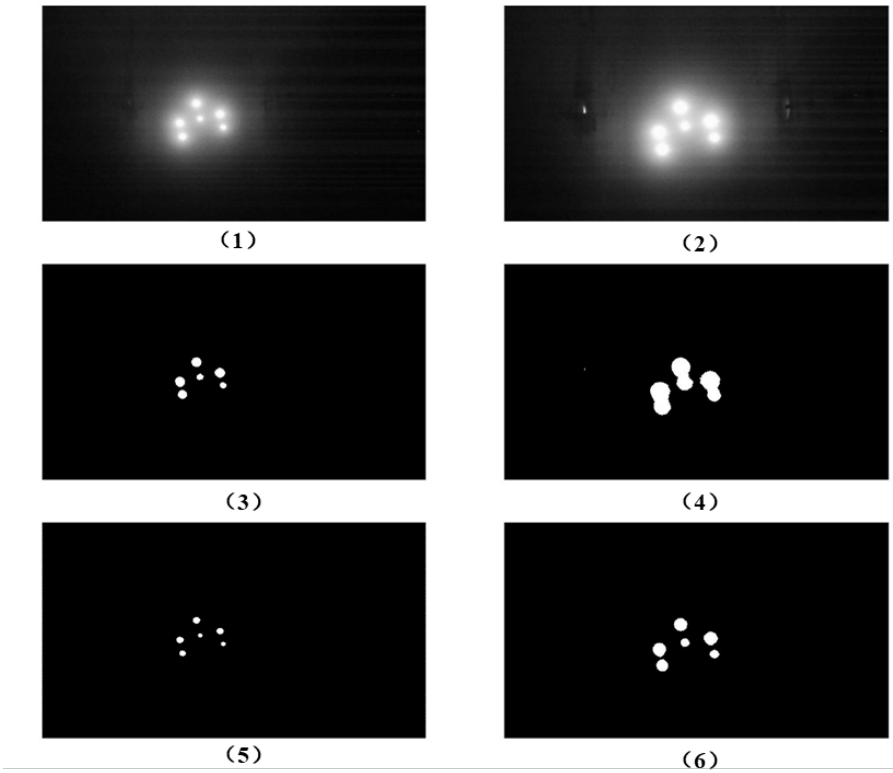


Figure 26 Results of Image Segmentation with Preset Thresholds

Among them, (1), (2) are the results of Gauss filtering, while (3), (4) are the results of

segmentation with threshold set to be 0.86, and (5) and (6) are the results of segmentation with threshold set as 0.96. Figures (1), (3), (5) correspond to the first frame of video, and Figures (2), (4), (6) correspond to the tenth frame. It can be seen that because the artificial threshold can not be adjusted in the motion, due to the change of illumination intensity, the segmentation with low threshold failed in the short distance, while in the long distance the foreground area is too small and irregular. All of these increase the risk of larger error. In order to adapt to the change of light intensity, ensure the segmentation effect and reduce the error risk as much as possible, it is necessary to find a way to adjust the threshold automatically.

Following the ideas of the previous section, three measures are defined in the process of segmentation to measure current image segmentation quality and to determine whether the threshold adjustment is to be carried out. The three measures defined are as follows:

Blob Number Measure： $M_{num}=\frac{|num-6|}{6}$， in which num is the number of Blobs in current segmentation;

Blob Area Measure： $M_{area}=1.5\times\left|\frac{mea}{0.4}\right|$， in which mea is the ratio between the smallest area and the largest area in current segmentation;

Blob Balance Measure： $M_{bal}=|mea_2-1|$， in which mea_2 is the ratio between the largest area and the sub-maximum area in current segmentation.

Considering the layout of actual guiding lights, the ideal number of segmentation is 6, while the ratio of minimum area to maximum area of Blob is about 0.4, and the ratio of maximum area to sub-maximum area is about 1. When the quality of segmentation worsens so some of those guiding lights cluster, the value of M_{num} will deviate，thus it can be seen as the symbol of inapposite threshold. However, it is found in the experiment that in some cases, although six Blobs can be segmented from the image, they still do not meet the needs of location because they are the mixture of real guiding lights and pseudo-lights (high-light noise points，etc.). Besides, there are other reasons why M_{num} is not suitable to be the only criterion of

segmentation quality, for instance, in some cases although M_{num} deviates from the ideal value, all the real guiding lights are separated successfully and can be extracted by simple screening，thus there is no need to elevate the threshold, taking the risk of increasing location error. In this connection, the Blob area measure and balance measure are defined, and these three measures are fused to evaluate the quality of segmentation. Similarly, the mapping between the measure and the BPA refers to Fig.23. The curves of possibility assigned to increasing the threshold, not increasing the threshold and uncertainty are in solid line, dash line and chain line respectively, and they are denoted by PY、PN、PU hereafter.

To test the effectiveness of algorithm, a 26-second underwater video is sampled every 0.5 seconds for segmentation and threshold evaluation. The segmentation results of auto-adjusted threshold based on measure fusion are shown in Figure 27, where subplots (1) - (6) corresponds to the segmentation results of frame 1, 10, 20, 30, 40 and 50, respectively. It can be seen that the image segmentation results meet the needs of monocular camera positioning, and the real-time performance is significantly improved, which can also meet the needs of experiments.

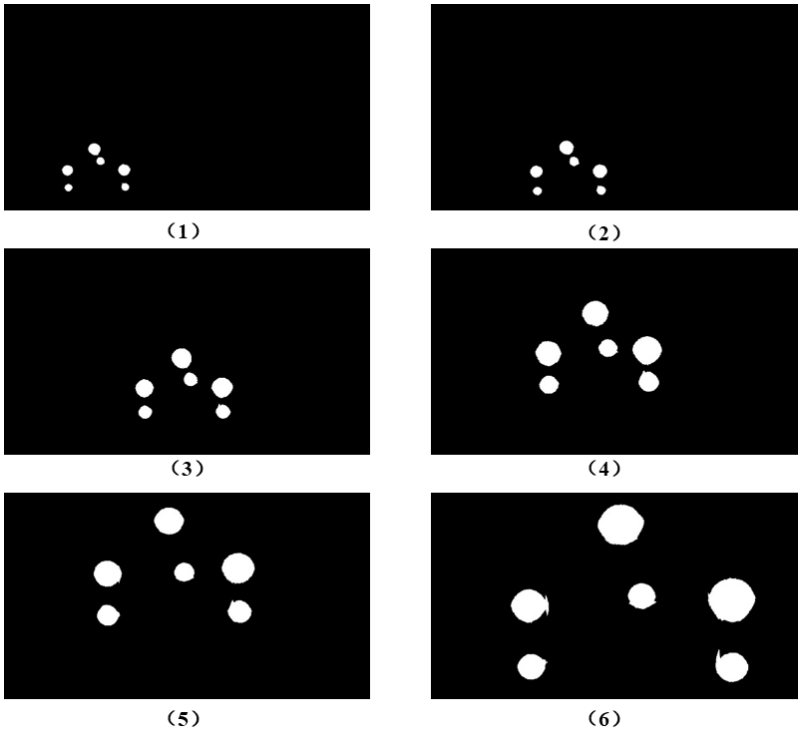


Figure 27 Results of Image Segmentation with Self-adaptive Thresholds

The frame-by-frame scatter plots of PY、PN、PU and the curve of threshold are shown

in Figure 28.

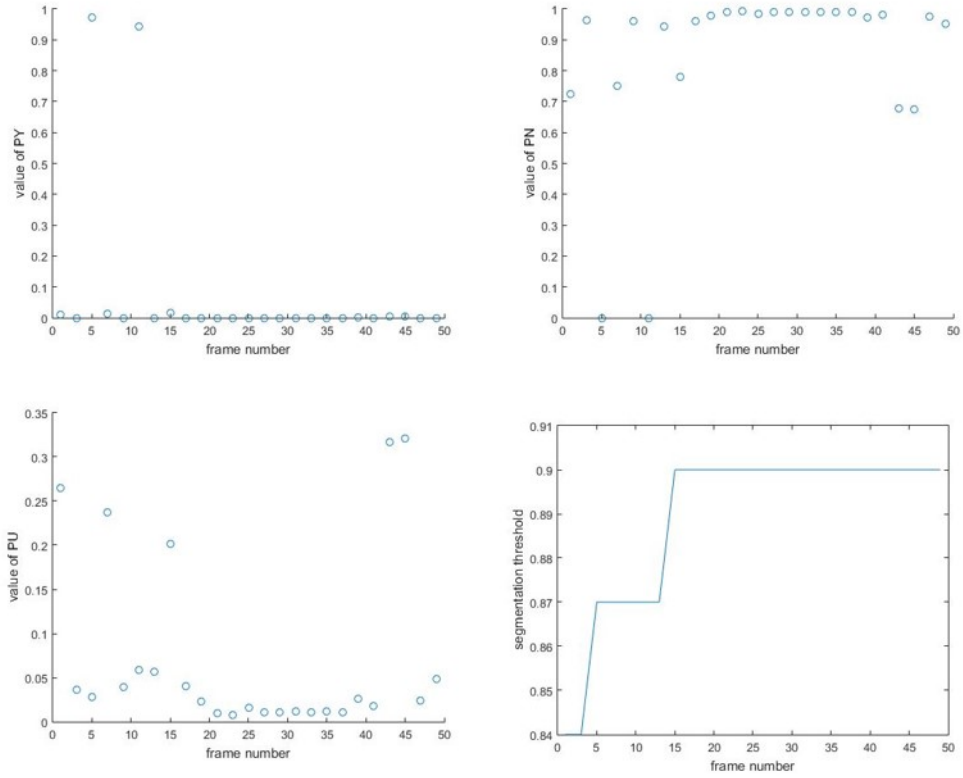


Figure 28 Curves of PY、PN、PU and the Threshold

It can be seen that in the moving process, the threshold value rises from 0.85 to 0.9 in terms of the quality of segmentation, and the uncertainty of decision is basically controlled within 0.2. The effectiveness of the improved D-S evidence theory is also verified from another side.

In order to further verify the effectiveness of algorithm, a video with more serious light scattering is selected for segmentation experiment. The video duration is 13 seconds and it is sampled every 0.5 seconds. The segmentation results of automatically adjusted threshold are shown in Figure 29, where the subplots (1) - (6) correspond to frames 1, 10, 20, 30, 40 and 50, respectively. It is obvious that the segmentation results can still meet the needs of location.

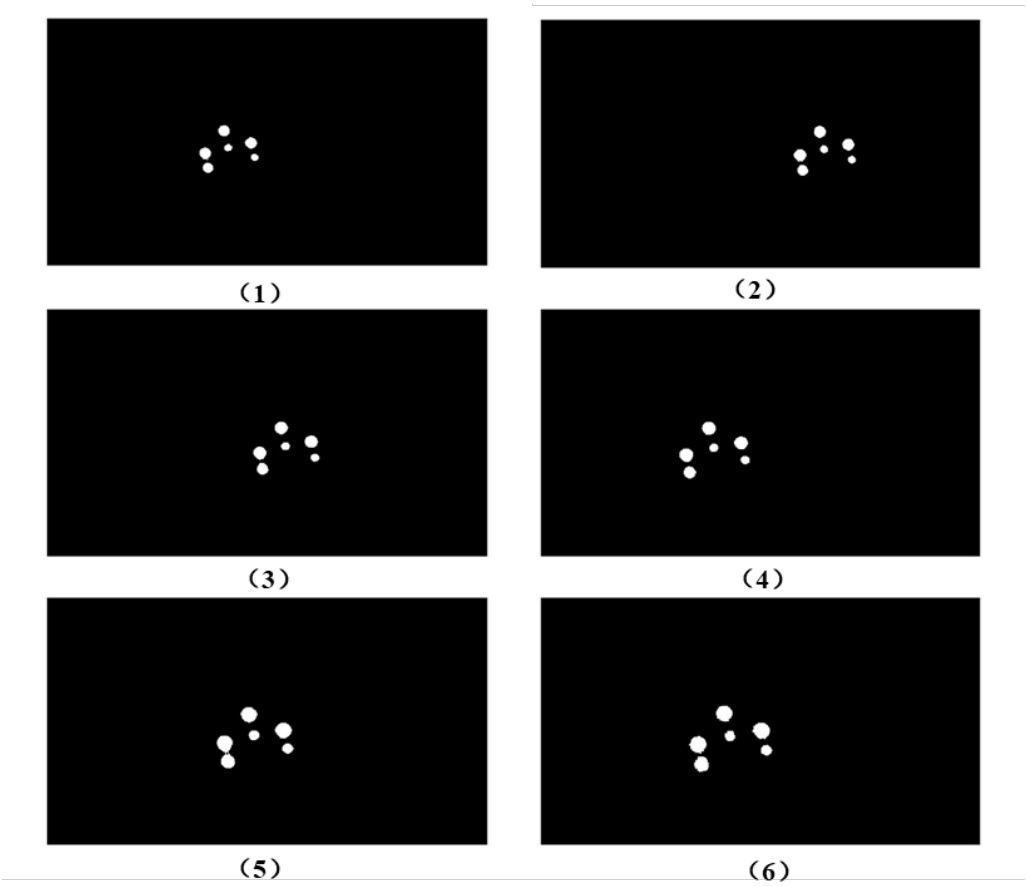


Figure 29 Results of Image Segmentation with Self-adaptive Thresholds

The frame-by-frame scatter plots of PY、PN、PU and the curve of threshold are shown in Figure 30.

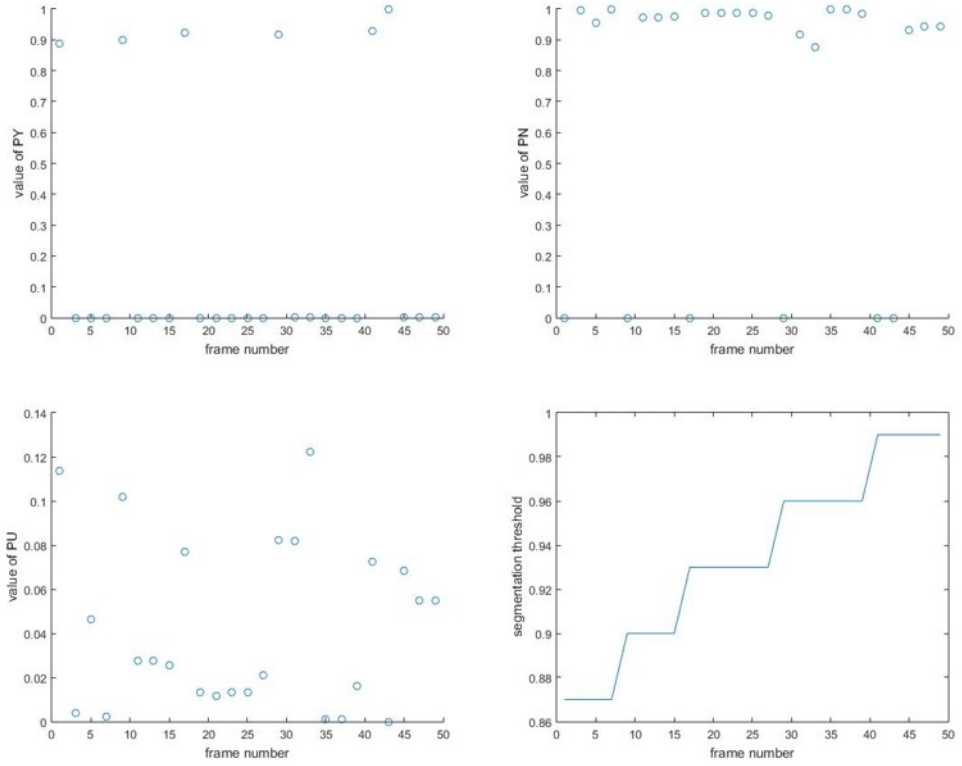


Figure 30 Curves of PY、PN、PU and the Threshold

It can be seen that in the moving process, the threshold value rises from 0.87 to 0.99 gradually, and the uncertainty of decision is kept within 0.2. The effectiveness of proposed threshold adjusting method has been further verified.

The improved image segmentation method is used to segment the video captured by AUV, and then for the Blob operator is used to locate the centroids of lights [49], which are put into the modified camera model for three-dimensional positioning. In order to verify the effectiveness of algorithm, the AUV is switched to ROV remote control mode to make it uniformly approach the docking platform. Figure 31 shows the location results of several frames, in which the centroids of guiding lights are marked. Figure 32 is the relative distance calculation result when AUV is uniformly approaching the platform. The validity of algorithm is preliminarily verified through the experiment.

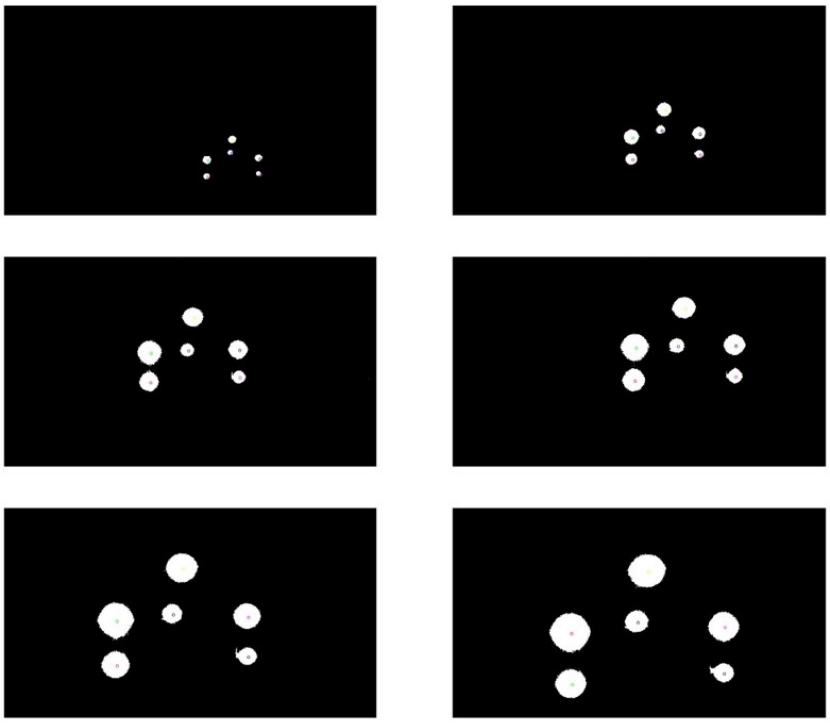


Figure 31 Results of Centroids Extraction

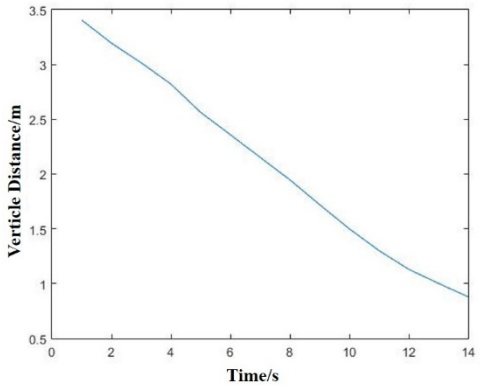


Figure 32 Results of Relative distance

3.4 Chapter Summary

In this chapter, some research has been carried out on the positioning algorithm based on monocular camera in the docking of AUV. Firstly, this chapter introduces the imaging characteristics underwater and several basic image preprocessing methods. After comparing their effects on actual underwater images, the Gaussian filter is selected and applied. Then, several edge extraction methods are introduced. After comparing their extraction effects, a light region extraction method based on D-S evidence theory is proposed, and the BPAs of three edge

extraction measures are abstracted and fused. In order to improve the effect of edge extraction, the D-S evidence theory is improved, and an improved D-S evidence algorithm based on TOPSIS is proposed. The effectiveness of improved algorithm in fusion of conflicting evidences are verified by simulation and comparison, and it is applied to image feature extraction. Considering the requirements on real-time performance, based on ideas above, a self-adaptive threshold adjusting method is proposed to achieve image segmentation. Finally, based on the improved visual positioning model in the previous chapter, the position of AUV relative to the guiding lights is calculated, and the algorithm is preliminarily tested by experiment.

Chapter 4 Position and Attitude Estimation Integrating Nine-axes Gyroscope and Monocular Camera

On the basis of attitude calculation of nine-axis gyroscope and monocular camera positioning, the task of this chapter is to obtain a more accurate and reliable estimation of position and attitude through a fusion algorithm combining the results of two sets of sensor systems. Kalman filtering theory, proposed by Kalman and Bucy, was first applied to the accurate orbit prediction of spacecraft. It has been widely used in the field of state estimation, after experiencing the test of time and different application scenarios, it still shows good estimation performance[51]. In the application of Kalman filter, the system equations and state model should be established according to kinetics, then, the optimal estimation of the system should be obtained by fusing observation quantities.

4.1 Theory of Kalman Filter

4.1.1 Characteristics of Kalman Filter

Kalman filter is a recursive linear filtering method based on least mean square error, which is suitable for state estimation of time-varying random sequences. The basic idea of Kalman filter is to analyze the statistical characteristics of state noise and observation noise according to the established state equation. The system observation values are taken as input, and the state estimation is updated and output after iteration. In the iterative process, the state estimated at previous moment are "predicted" to produce the state estimates of the current moment, then combining current observations, the final state estimates are generated by "updating", and the estimates of system noise and covariance are modified. Kalman filtering algorithm can be realized by iteration, which requires less storage space of the system, so it is simple and easy to implement, and various engineering applications have fully verified its ability of state estimation.

4.1.2 Discrete Kalman Filter

In order to explain the working principle of discrete Kalman filter, the state equations and observation equations of the system have to be established first. The object of state estimation can be described by the following equation:

$$\begin{cases} X_k = \Phi_{k,k-1} X_{k-1} + \Gamma_k W_k \\ Y_k = H_k X_k + V_k \end{cases} \tag{4.1}$$

where X_k is the state vector to be estimated by Kalman filter, while Y_k is the observation vector， $\Phi_{k,k-1}$ denotes the state-transition matrix， H_k denotes the observation matrix， Γ_k is driving matrix of process noise, W_k is the vector of process noise and V_k denotes the vector of observation noise. Assuming that：

$$\begin{cases} E(W_k) = 0 \quad Cov(X_0, W_k) = 0 \\ E(V_k) = 0 \quad Cov\left(X_0, V_k\right) = 0 \\ \text{cov}(W_k, W_j) = Q_k \delta_{kj} \\ \text{cov}(V_k, V_j) = R_k \delta_{kj} \\ \text{cov}(W_k, V_k) = 0 \end{cases} \tag{4.2}$$

Namely, the process noise and observation noise of the system are both zero-mean white noise, and their statistical characteristics are independent of each other. Besides, they are also not correlated with the statistical characteristics of the initial value of the state vector. Q_k denotes the variance matrix of process noise, while R_k is the variance matrix of observation noise, and δ_{kj} subjects to $\delta_{kj} = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases}$。

In this connection, the recursive formula of Kalman filter is as follows:

$$\begin{gathered} \hat{X}_{k,k-1} = \Phi_{k,k-1} \hat{X}_{k-1} \\ \hat{X}_k = \hat{X}_{k,k-1} + K_k \left(Y_k - H_k \hat{X}_{k,k-1} \right) \\ P_{k,k-1} = \Phi_{k,k-1} P_{k-1} \Phi_{k,k-1}^T + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T \\ K_k = P_{k,k-1} H_k^T \left(H_k P_{k,k-1} H_k^T + R_k \right)^{-1} \\ P_k = \left(I - K_k H_k \right) P_{k,k-1} \end{gathered} \tag{4.3}$$

where K_k denotes the gain matrix, which determines the weight of the difference between the observer and the model estimator and is closely related to the final filtering effect. Before starting the iteration, to update the state variables step by step, besides constructing the state transition matrix, the measurement matrix and the process noise driving matrix, the initial values of the state variable X_0 and the covariance matrix of state estimation P_0 should also be specified.

4.2 Integrated Attitude Estimation Based on Kalman Filter

The gyroscope has the advantages of low cost, small size and light weight. However, due to the limitation of its manufacturing technology, the accuracy of gyroscope is often relatively low, and the integral drift will accumulate seriously over time, severely affecting the accuracy of attitude calculation. Accurate attitude calculation is the precondition of eliminating the additional acceleration caused by the attitude angle deviation and extracting the linear acceleration of the moving carrier from the accelerometer measurements, so the integral drift of the gyroscope will seriously affect the accuracy of calculated velocity and position. The accelerometer, which also belongs to the inertial device, also has a certain zero drift, and is vulnerable to environmental noise interference as well as the magnetometer. In this section, the error sources of these three components of the nine-axis gyroscope are briefly analyzed, and the error state model is constructed. Integrating the model and measurements, the attitude is estimated based on Kalman filter.

4.2.1 Error of Components and Error Models

Due to the influence of processing and installation accuracy, gyroscope will have several kinds of errors in operation [55]. Among them, zero drift, Gaussian white noise error and first-order Markov error are the main errors which have great influence on attitude calculation. The error model of gyroscope can be expressed by:

$$\delta\omega = n_{zero} + n_{markov} + n_{gauss} \tag{4.4}$$

where n_{zero} denotes the constant zero drift and $\dot{n}_{markov} = -\frac{1}{T_g} n_{markov} + \xi_g$. Similarly, the error models of accelerometer and magnetometer can be built as below.

$$\begin{aligned} \delta a &= a_{zero} + a_{markov} + a_{gauss} \\ \dot{a}_{markov} &= -\frac{1}{T_a} a_{markov} + \xi_a \end{aligned} \tag{4.5}$$

$$\begin{aligned} \delta m &= m_{zero} + m_{markov} + m_{gauss} \\ \dot{m}_{markov} &= -\frac{1}{T_m} m_{markov} + \xi_m \end{aligned} \tag{4.6}$$

4.2.2 Error-state Model of Attitude Estimation

In this paper, because the quaternion method is used in attitude calculation, and the four elements of quaternion are all numbers less than or equal to 1, if the attitude angles are selected as the state variables to be estimated, the magnitude difference between quaternion and state variables may be large, therefore large errors may be introduced in the calculation. After studying the indirect filtering method [56], in which the state variables are chosen as the errors of state variables in direct estimation, thus are close to each other in order of magnitude, it is chosen and used considering the advantages in accuracy. Therefore, the error state model of attitude estimation is to be built.

In terms of the quaternion formula in chapter 2 we can get

$$\begin{bmatrix} \dot{q}_0 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{bmatrix} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} \tag{4.7}$$

In order to construct the error model, the upper equations are differentiated to obtain

$$\begin{bmatrix} \delta\dot{q}_0 \\ \delta\dot{q}_1 \\ \delta\dot{q}_2 \\ \delta\dot{q}_3 \end{bmatrix} = \frac{1}{2} M(q) \begin{bmatrix} 0 \\ \delta\omega_x \\ \delta\omega_y \\ \delta\omega_z \end{bmatrix} + \frac{1}{2} M^*(\omega_q) \begin{bmatrix} \delta q_0 \\ \delta q_1 \\ \delta q_2 \\ \delta q_3 \end{bmatrix} \tag{4.8}$$

in which the sub-matrixes are

$$M(q)=\begin{bmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{bmatrix} \tag{4.9}$$

$$M^*(\omega_q)=\begin{bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & -\omega_z & \omega_y \\ \omega_y & \omega_z & 0 & -\omega_x \\ \omega_z & -\omega_y & \omega_x & 0 \end{bmatrix} \tag{4.10}$$

After analyzing the outputs of system, considering that the Gaussian white noise and the term ξ_g in Markov noise are zero-mean white noises, the errors between the output of four elements of quaternion and the true value, the three-axis zero drifts and the three-axis Markov drifts of gyroscope are selected as state variables, that is, the state vector of the system is chosen as $X=\left[\delta q_0,\delta q_1,\delta q_2,\delta q_3,n_{zerox},n_{zeroy},n_{zeroz},n_{markovx},n_{markovy},n_{markovz}\right]^T$, while the process noise is $W=\left[n_{gaussx}\quad n_{gaussy}\quad n_{gaussz}\quad \xi_{gx}\quad \xi_{gx}\quad \xi_{gz}\right]$. On this basis, the state transition matrix and driving matrix of process noise can be constructed as below:

$$\Phi=\begin{bmatrix} \frac{1}{2}M^*(\omega_q) & \frac{1}{2}M^*(q) & \frac{1}{2}M^*(q) \\ 0_{3\times4} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times4} & 0_{3\times3} & T \end{bmatrix} \tag{4.11}$$

$$\Gamma=\begin{bmatrix} \frac{1}{2}M^*(q) & 0_{4\times3} \\ 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & I_{3\times3} \end{bmatrix} \tag{4.12}$$

In which the matrix $M^*(q)$ is composed by the last three columns of $M(q)$, namely

$$M^*(q)=\begin{bmatrix} -q_1 & -q_2 & -q_3 \\ q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{bmatrix} \tag{4.13}$$

$$T=\begin{pmatrix} -\frac{1}{T_g} & 0 & 0 \\ 0 & -\frac{1}{T_g} & 0 \\ 0 & 0 & -\frac{1}{T_g} \end{pmatrix} \tag{4.14}$$

Except for the attitude angles obtained by gyroscope through quaternion calculation, a set of attitude angles can also be obtained using the outputs of accelerometer and magnetometer. In order to get the measurement matrix, it is necessary to convert the attitude angles calculated by accelerometer and magnetometer into quaternion, so as to correct the estimation of state variables. The formula for converting attitude angles into quaternion is as follows:

$$\begin{aligned} q_0' &= \cos\left(\frac{\psi}{2}\right)\cos\left(\frac{\theta}{2}\right)\cos\left(\frac{\gamma}{2}\right)-\sin\left(\frac{\psi}{2}\right)\sin\left(\frac{\theta}{2}\right)\sin\left(\frac{\gamma}{2}\right) \\ q_1' &= \cos\left(\frac{\psi}{2}\right)\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\gamma}{2}\right)-\sin\left(\frac{\psi}{2}\right)\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\gamma}{2}\right) \\ q_2' &= \cos\left(\frac{\psi}{2}\right)\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\gamma}{2}\right)+\sin\left(\frac{\psi}{2}\right)\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\gamma}{2}\right) \\ q_3' &= \cos\left(\frac{\psi}{2}\right)\sin\left(\frac{\theta}{2}\right)\sin\left(\frac{\gamma}{2}\right)+\sin\left(\frac{\psi}{2}\right)\cos\left(\frac{\theta}{2}\right)\cos\left(\frac{\gamma}{2}\right) \end{aligned} \tag{4.15}$$

Because the error state filtering is used in this paper, and the state estimation model constructed is also the error model, it is necessary to differentiate the quaternion obtained by accelerometer and magnetometer from the quaternion of gyroscope. Assuming that the quaternion of gyroscope is $\begin{bmatrix} q_0 & q_1 & q_2 & q_3 \end{bmatrix}^T$, and the quaternion obtained by accelerometer and magnetometer is $\begin{bmatrix} q_0' & q_1' & q_2' & q_3' \end{bmatrix}^T$, then the observation can be defined as:

$$Y=\begin{bmatrix} q_0-q_0' \\ q_1-q_1' \\ q_2-q_2' \\ q_3-q_3' \end{bmatrix} \tag{4.16}$$

Correspondingly, the observation matrix can be determined as $H=\begin{bmatrix} I_{4\times4} & 0_{4\times6} \end{bmatrix}$. By far, besides the variance matrixes of process noise and observation noise, other matrices have been defined. In the subsequent verification, process noise and observation noise are set in terms of

the simulation and actual test conditions. It should be pointed out that the values of the state transition matrix and the driving matrix of process noise are updated according to the fused quaternion output of the previous step and the angular velocities output by gyroscope at current step.

4.2.3 Simulation of Integrated Attitude Estimation

To verify the effectiveness of proposed attitude estimation algorithm, the sensor data are simulated by MATLAB then the attitude calculation and fusion are carried out. After consulting the data sheets of related devices, the parameters of sensors are selected as shown in Table 2.

Table 2 Parameters of Nine-axis Gyroscope in Simulation

Parameters	Value
Zero drift error (gyroscope)	0.05°
First-order Markov drift (gyroscope)	0.001°/s
Related time of Markov drift (gyroscope)	1600s
SNR of white noise (gyroscope)	30dB
Zero drift error (accelerometer)	0.001g
First-order Markov drift (accelerometer)	0.001g/s
Related time of Markov drift (accelerometer)	1200s
SNR of white noise (accelerometer)	20dB
Zero drift error (magnetometer)	20nT
First-order Markov drift (magnetometer)	10nT/s
Related time of Markov drift (magnetometer)	1200s
SNR of white noise (magnetometer)	20dB

In order to compare the effects of attitude estimation before and after sensor fusion, firstly, the attitude of the carrier is calculated using the simulated sensor data. In the simulation, assuming that the yaw angle ψ, roll angle ϕ and pitch angle θ all vary with time according to sinusoidal function, and the formulas are as follows:

$$\begin{aligned}\psi &= 10\times\sin\left(\frac{2\pi}{10}t\right)\\ \phi &= 10\times\sin\left(\frac{2\pi}{10}t\right)\\ \theta &= 10\times\sin\left(\frac{2\pi}{10}t\right)\end{aligned} \tag{4.17}$$

The differences between the calculated results and real values of the yaw angle, roll angle and pitch angle are shown in Figure 33.

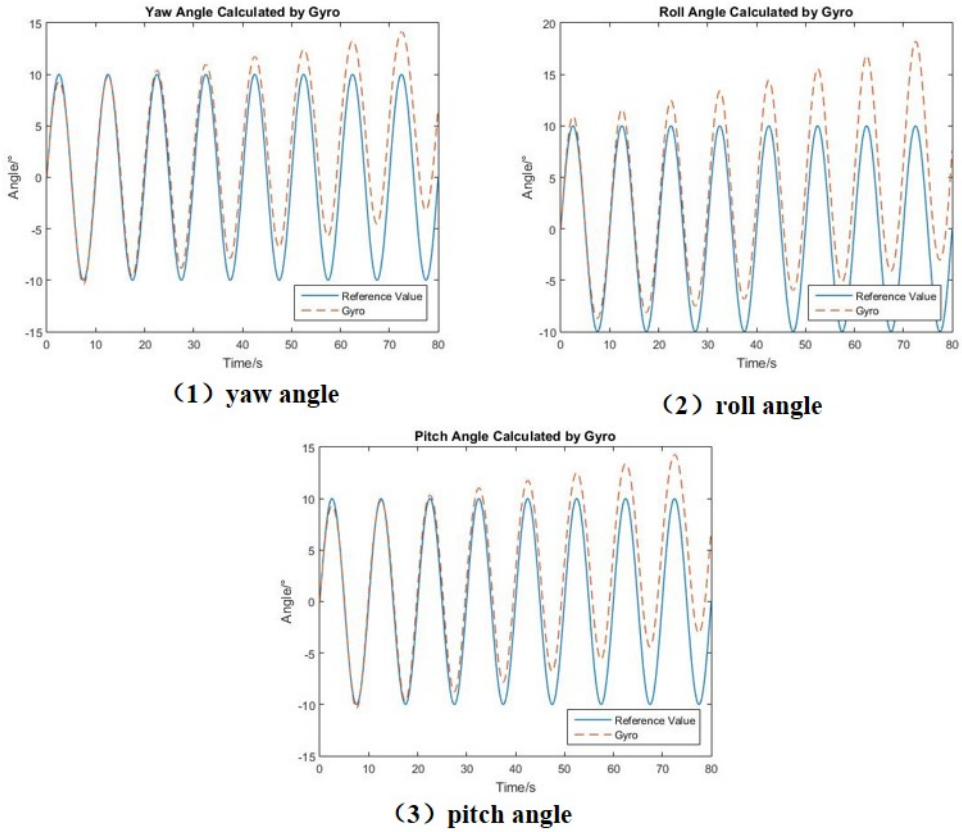


Figure 33 Attitude Calculated by Gyroscope

In the figures above, the solid lines are the true values of three attitude angles, and the dash lines are the attitude angle calculated with sensor data. It can be seen that over time, the attitude calculation results of gyroscope produce serious cumulative errors, resulting in the deviation from true attitude, which is determined by the inherent characteristics of the gyroscope sensor. Although the integral error and zero drift of gyroscope can be significantly reduced with the improvement of related technology, for most non-tactical civil equipment, the attitude angles calculated only depending on the three-axis angular velocity of gyroscope, is far from meeting

the accuracy requirements of attitude estimation. In order to compensate for this accumulated error as much as possible and improve the accuracy and reliability of results, the accelerometer and magnetometer are used to get another set of attitude angle measurements. Assuming that the yaw angle ψ, roll angle ϕ and pitch angle θ vary with time according to sinusoidal function, the comparison between the calculated results and true values o are shown in Figure 34.

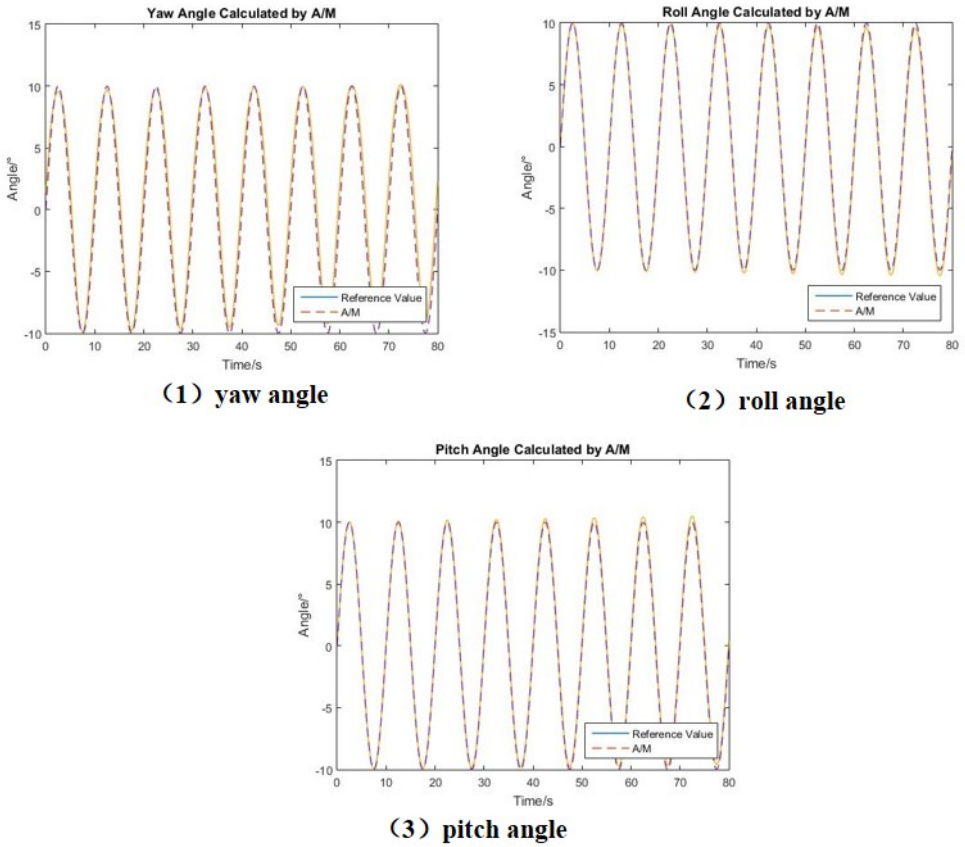


Figure 34 Attitude Calculated by Accelerometer and Magnetometer

In the figures above, the solid line are the true values of attitude angles, and the dashed lines are the results calculated with accelerometer and magnetometer data. It can be seen that the accelerometer and magnetometer will also produce some cumulative errors over time, but the errors are significantly smaller than those of gyroscopes. Since the accelerometer and magnetometer are sensitive to ambient noise while the gyroscope is more resistant to external noise interference due to their own characteristics, the attitude error model established in the previous section will be used in the next step to achieve integrated attitude estimation by

Kalman filter.

After specifying the original value of state vector x_0, covariance matrix of state vector P_0, covariance matrix of process noise Q_0 and covariance matrix of observation noise R_0, the results of integrated attitude estimation are given in Figure 35.

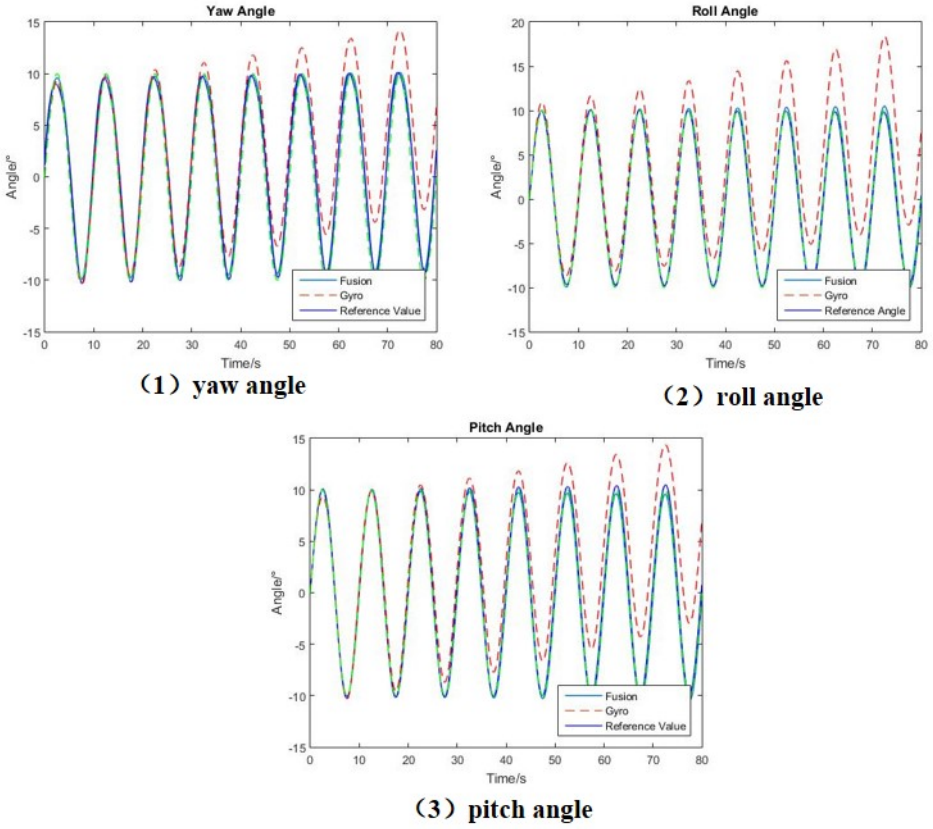


Figure 35 Attitude Calculated by Integrated Algorithm and Comparison

In the figures above, the dashed lines are the results of gyroscope data, and solid lines are the results of fusion algorithm. It can be seen intuitively that the cumulative errors decrease obviously after fusion. In order to evaluate the attitude angle errors before and after fusion more intuitively, the error curves of three attitude angles are shown in Figure 36.

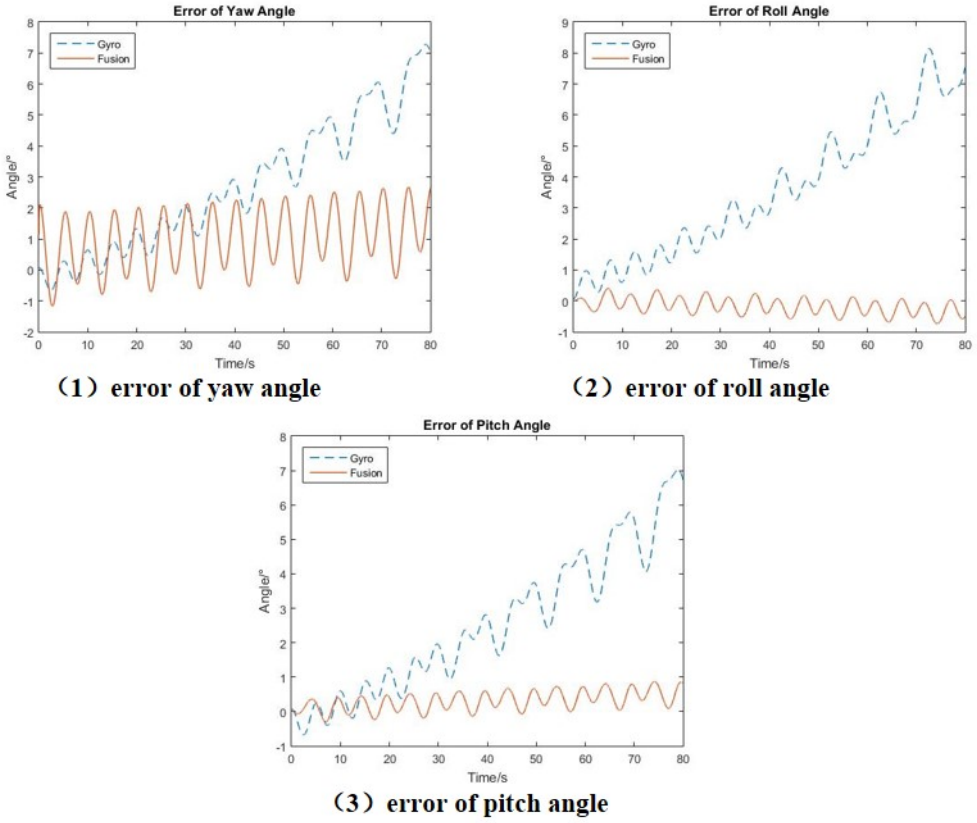


Figure 36 Errors of Attitude Before and After Fusion

Among them, the dashed lines are the errors before fusion, while the solid lines are the errors after fusion, and the root mean square values of angle errors are listed in Table 3.

Table 3 Root Mean Square Errors of Attitude Before and After Fusion

Attitude Angle	Before Fusion	After Fusion
Roll Angle	4.2053	0.3080
Pitch Angle	3.3857	0.4129
Yaw Angle	3.3378	1.3784

Comparing the results before and after fusion, it can be found that combining the accelerometer and magnetometer with gyroscope to calculate the attitude by fusion algorithm can effectively reduce the accumulated error of gyroscope and improve the accuracy of estimation in a long time.

4.3 Integrated Position Estimation Based on Kalman Filter

In the last section, the principle of integrated attitude estimation is introduced, an

d its effect is verified by simulation. In the inertial navigation unit, only the accelero meter can measure acceleration, and its value is the sum of linear acceleration and gr avity acceleration. In this section, on the basis of improved attitude algorithm, the line ar accelerations measured by accelerometer are extracted using the calculated angles, a nd the velocity and position of the carrier are obtained by twice integrals. Monocular vision has good positioning accuracy at relatively short distance, but given the large c omputational burden of image processing, its sampling speed is significantly lower tha n that of inertial navigation unit. In order to make full use of the advantages of both, a position estimation model based on Kalman filter, fusing the data of IMU and mo nocular vision unit, are established for accurate positioning of the carrier.

4.3.1 Error Analysis

As mentioned above, the output of accelerometer is the mixture of gravity acceler ation and the real linear acceleration in three axes. When the attitude of the vehicle is not clear, these two components can’t be distinguished effectively, thus the subsequen t calculation of linear velocity and position becomes unreliable. The linear acceleration of the three axes can be extracted from the output based on the attitude angles of fu sion algorithm, but in the integration, the zero drift error, cumulative error and the gr avity acceleration component which is not completely eliminated from the output will accumulate several times, leading to the serious deviation of velocity and position. Th e error increases so fast that the position determined by accelerometer and fusion attit ude alone can’t meet the requirements of our state estimation. In this paper, we use u nderwater lights to guide the AUV and the model of monocular camera positioning ha s been introduced and modified in the previous part. Considering the positioning accur acy and output frequency, in this section, the data of accelerometer and monocular ca mera are combined to estimate the position through Kalman filter.

4.3.2 Model of Position Estimation

The position algorithm and attitude algorithm are similar in principle as both of t

hem use classical Kalman algorithm. Without considering the difference of magnitude, in position calculation the general state estimation model can meet our requirements. However, considering that the position obtained by the IMU is relative to the navigati on coordinate system, but the output of monocular camera is the relative distance bet ween the carrier and the fixed docking platform, it is necessary to perform coordinate transformation before.

In the calculation, $X=\begin{bmatrix} p_x & p_y & p_z & v_x & v_y & v_z & a_x & a_y & a_z & \delta a_x & \delta a_y & \delta a_z \end{bmatrix}^T$ is c hosen as the state vector, with its terms corresponding to the three-axial positions, ve locities, accelerations and output errors. $Z=\begin{bmatrix} p'_x & p'_y & p'_z & a'_x & a'_y & a'_z \end{bmatrix}^T$ is the observ ation vector, whose terms correspond to the three-axial positions and linear accelerati ons computed by camera and accelerometer respectively. Then the state transition mat rix Φ and observation matrix H can be constructed as:

$$\Phi=\begin{bmatrix} I_{3\times3} & T\cdot I_{3\times3} & \frac{T^2}{2}\cdot I_{3\times3} & -\frac{T^2}{2}C_b^n\cdot I_{3\times3} \\ O_{3\times3} & I_{3\times3} & T\cdot I_{3\times3} & -T\cdot C_b^n\cdot I_{3\times3} \\ O_{3\times3} & O_{3\times3} & I_{3\times3} & O_{3\times3} \\ & O_{3\times9} & & I_{3\times3} \end{bmatrix} \tag{4.18}$$

$$H=\begin{bmatrix} I_{3\times3} & O_{3\times3} & O_{3\times3} & O_{3\times3} \\ O_{3\times3} & O_{3\times3} & I_{3\times3} & O_{3\times3} \end{bmatrix} \tag{4.19}$$

In the updating of Kalman filter, the discrete Kalman filter equations used are si milar to those in the previous section, so they are not repeated here. In the fusion, it is necessary to determine the state transition matrix of each step using the attitude an gles estimated by fusion attitude algorithm. The matrix C_b^n can be obtained by：

$$C_b^n=\begin{bmatrix} \cos\phi\cos\psi+\sin\phi\sin\psi\sin\theta & \sin\psi\cos\theta & \sin\phi\cos\psi-\cos\phi\sin\psi\sin\theta \\ -\cos\phi\sin\psi+\sin\phi\cos\psi\sin\theta & \cos\psi\cos\theta & -\sin\phi\sin\psi-\cos\phi\cos\psi\sin\theta \\ -\sin\phi\cos\theta & \sin\theta & \cos\phi\cos\theta \end{bmatrix} \tag{4.20}$$

where ψ, θ and ϕ denote the yaw, pitch and roll angles estimated in last step.

4.3.3 Simulation of Integrated Position Estimation

To verify the position estimation algorithm, sensor data are simulated by MATLAB and the calculation and fusion carried out. After consulting the data sheets of related devices and preliminary tests, the IMU parameters are selected as Table 2. Considering that the localization error of camera increases when the AUV is relatively far from the platform, the localization error is set to be a variable whose mean is $0.2-0.0075t \quad 0 \le t \le 20s$, subjecting to Gaussian distribution.

To compare the position estimation effects before and after fusion, firstly the position is calculated using the simulated accelerometer data. It is assumed that the carrier is moving with uniform acceleration in segments, and its acceleration is as follows:

$$\begin{cases} a_x = a_y = a_z = 0.2m/s^2 & 0 < t \le 5s \\ a_x = a_y = a_z = 0.2m/s^2 & 5 < t \le 20s \end{cases} \tag{4.21}$$

The results of three-axis acceleration, velocity and position (taking those of X-axis as examples) and the differences between them and the benchmark values are shown in Figure 37.

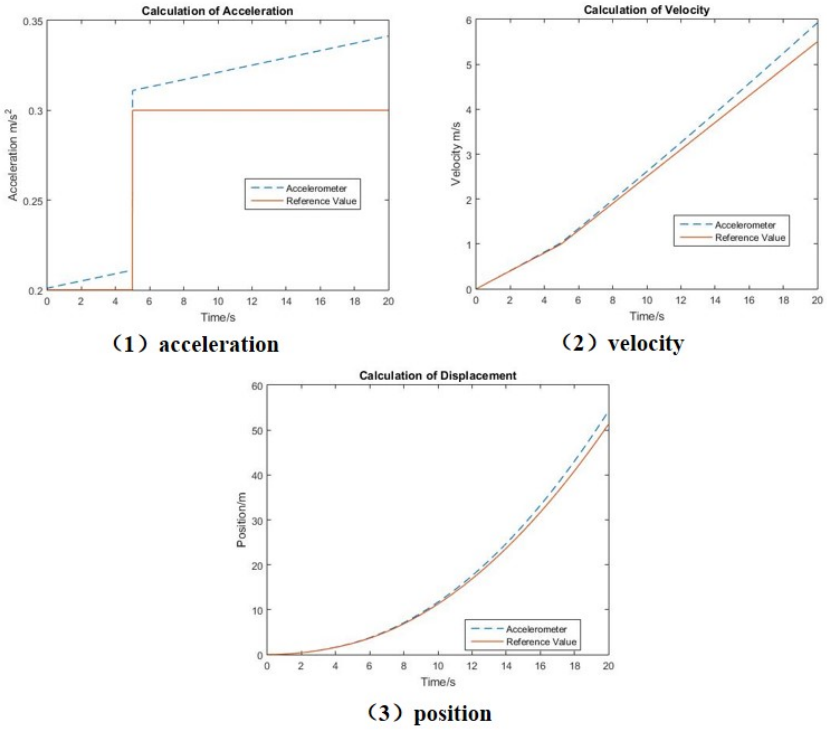


Figure 37 Acceleration, Velocity and Position Calculated by Accelerometer

In the above figures, the solid lines are the real values of linear acceleration, velocity and position, while the dashed lines are the calculated values. It can be seen that over time, the calculation results of accelerometer produce serious cumulative errors. The error of acceleration accumulates with twice integrals, resulting in the severe output deviation in a short time, which is also the important reason why the accelerometer can not be used independently for carrier positioning. Although the integral and zero drift errors of component can be reduced to some extent, most of the low-cost and widely used accelerometer still can't ensure the required positioning accuracy. In order to compensate for this cumulative error, the state model established in the previous section is used to fuse the results of IMU and monocular camera based on Kalman filter. After specifying the original value of state vector x_0, covariance matrix of state vector P_0, covariance matrix of process noise Q_0 and covariance matrix of observation noise R_0, the results of estimated velocity and position are given in Figure 38.

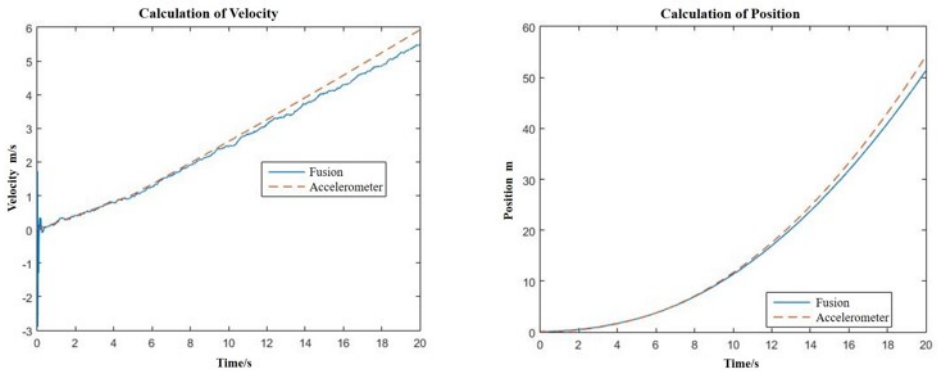
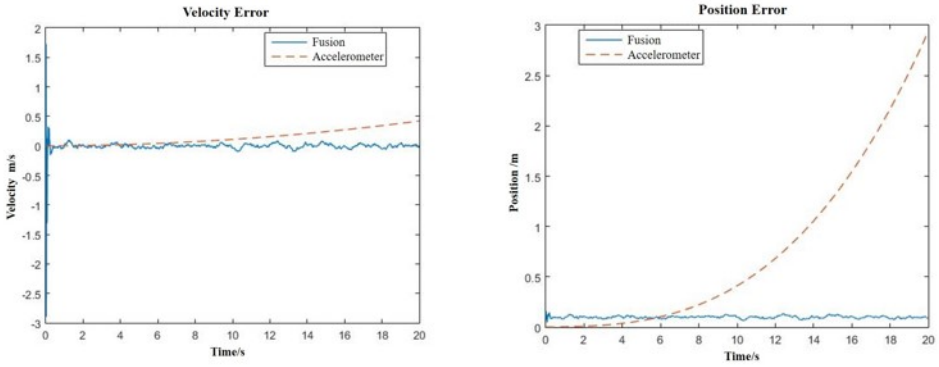


Figure 38 Velocity and Position Calculated Before and After Fusion

The dotted lines are the velocity and position computed from the accelerometer data, and the solid lines are the fused value. It can be seen intuitively that after fusion, the error of velocity and position calculation are obviously reduced. To evaluate the errors before and after fusion more intuitively and quantitatively, the error curves of velocity and position are shown in Figure 39.



（1）error of velocity　　（2）error of position

Figure 39 Errors of Velocity and Position Before and After Fusion

Among them, the dotted lines are the pre-fusion error, while the real lines are the post-fusion error, and the root mean square values of estimation error are listed in Table 4.

Table 4 Root Mean Square Errors of Velocity and Position Before and After Fusion

Variable	Before Fusion	After Fusion
RMS Error of Velocity	0.1909	0.0821
RMS Error of Position	1.1303	0.1002

Comparing the results, it can be found that although the fusion algorithm has a slightly larger initial error, it can converge fast in a short time. Besides, the slightly unsatisfactory effects of positioning in the early stage may be owing to the larger random error specified to monocular camera considering the relative distance. Generally speaking, the simulation results show that the accumulative position error of accelerometer due to two integrals can be effectively reduced and the accuracy of location is improved after fusion.

4.4 Chapter Summary

In this chapter, based on Kalman filter, the attitude is estimated combining data o f gyroscope, accelerometer and magnetometer, while the position is estimated combinin g the results of accelerometer and monocular camera. In the application, considering t he order of magnitude, the error state model and state model are established respectiv ely. Then the measurement differences or output of other sensors are taken as the obs ervation variables to compensate for the results of IMU and obtain the fusion estimati

on. Finally, the effectiveness of algorithms are verified through simulation.

Chapter 5 Experiments and Results Analysis

In order to verify the model established and the algorithm prosed, based on the actual IMU chip, AUV and docking platform in the cooperative project, the integrated attitude and position estimation algorithms are tested, in order to increase the experimental support for proposed algorithm, find out the shortcomings, and put forward the prospect.

5.1 Experimental Scenarios

5.1.1 Experimental Scenario of Integrated Attitude Estimation

In order to replicate the hardware conditions of AUV carrier and facilitate the subsequent position verification, the nine-axis IMU including three-axis gyroscope, three-axis accelerometer and three-axis magnetometer is selected in the attitude verification stage. The type of chip is MPU9250, which integrates MPU6050 (three-axis gyroscope and three-axis accelerometer) and AK8963. Arduino Leonardo chip is selected as the circuit control board to realize power supply and programing. The layout of chip is shown in Figure 40.

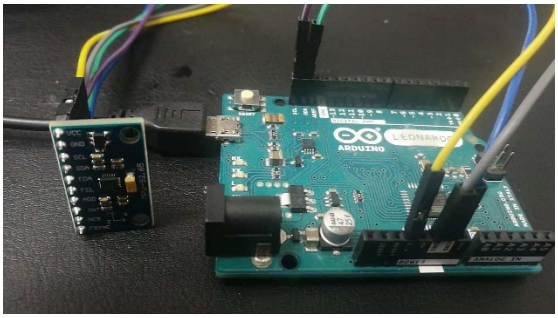


Figure 40 Nine-axis IMU Chip and the Control Board

The preliminary test is in the new main building of Beihang University, and the ambient temperature is about 18 C. The data are transmitted by I^2C, while the submodules MPU6050 and AK8963 are both capable of direct Read-write in the independent transmitting mode. Before each test, the gyroscope, accelerometer and magnetometer must be drift-corrected first, and the accelerometer and magnetometer need ellipsoidal

correction to improve data reliability. In the experiment, the sampling frequency of gyr oscope, accelerometer and magnetometer is 100 Hz while the frequency of fusion algo rithm is 50 Hz. The baud rate of data transmission is 9600 bps, and the clock is syn chronized in advance. A program is compiled to collect sensor data through the chip and complete preliminary calculation. The program interface is shown in Figure 41.

Figure 41 Interface of Data Acquisition Program

5.1.2 Experimental Scenario of Integrated Position Estimation

The position estimation test relies on the AUV carrier and the test pool of Huazhong University of Science and Technology. The width of the test section was 3.0m. The docking device was submerged and fixed near the end of the test section by cable. The underwater visibility of the pool is good. The monocular camera is installed in front of the AUV. During the movement, the camera doesn't pitch against the carrier and its relative position to the carrier is determined by calibration. The actual experimental scenario is shown in Figure 42.

Figure 42 Experimental Tank

5.2 Experimental Results and Analysis

5.2.1 Results of Integrated Attitude Estimation

To validate the proposed attitude fusion algorithm, experiments are carried out in static state and motion state respectively. During the experiment, related filter matrixes are updated according to the outputs and calculation results of gyroscope, accelerometer and magnetometer. The static test lasts for 20 seconds and the dynamic test lasts for 120 seconds. The comparison of attitude estimation results before and after fusion under static and moving states is shown in Figure 43.

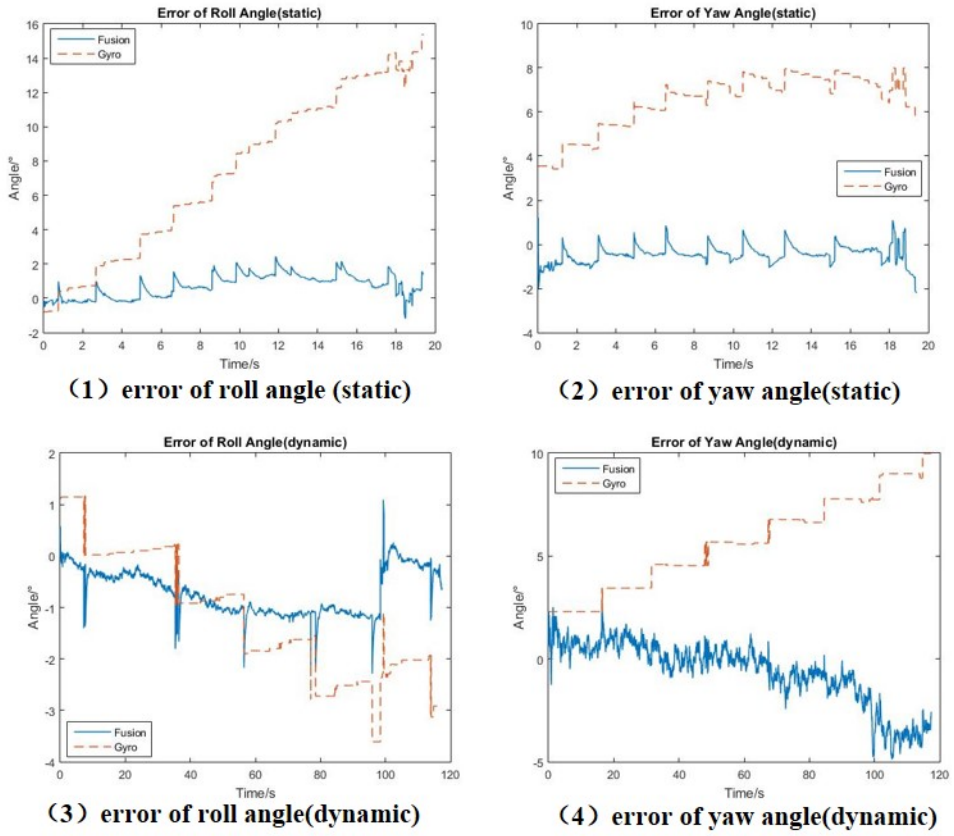


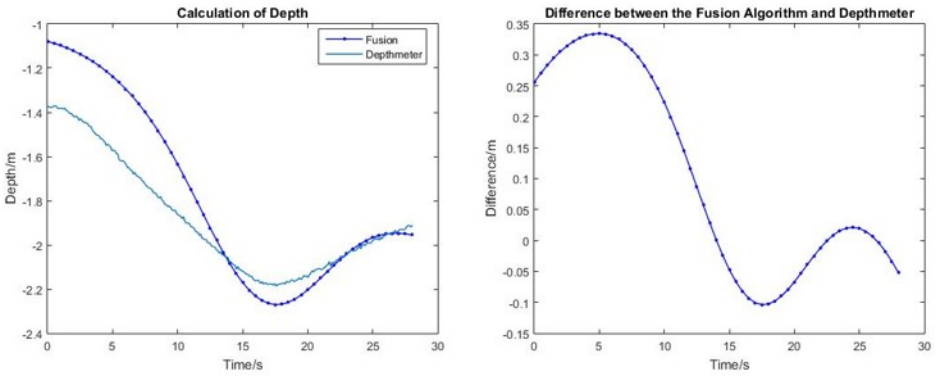
Figure 43 Errors of Attitude under Static and Dynamic Conditions

It can be seen from the figures that the fusion algorithm can achieve better attitude estimation in both cases. In static state, the performance of algorithm is improved more obviously than before because of the higher precision of accelerometer and magnetometer, while in dynamic state, the attitude accuracy of accelerometer and magnetometer decreases and the attitude estimation error increases, but in general, effective attitude estimation can still be

achieved. It is also found that the state estimation error of yaw angle is larger than that of roll angle and pitch angle. This is because the other solution of yaw angle relies on magnetometer, which is sensitive to ambient noise. However, considering that the nine-axis gyroscope used in the experiment is the basic product with limited cost and precision, in general, the attitude estimation can be achieved more accurately by proposed algorithm.

5.2.2 Results of Integrated Position Estimation

To validate the position fusion algorithm proposed in this paper, related matrixes are updated according to the results of accelerometer and monocular camera to realize iterative position estimation. Because the depth of the pool is up to 4m, it is difficult to determine the actual lateral and vertical displacement due to the limitation of test c onditions. Considering that the AUV is equipped with a high-precision depth meter, it is more convenient to determine its depth, and AUV moves relatively independently in three directions, so in position fusion test, the AUV approaches the docking platform gradually, and its depth is calculated by fusion algorithm. Comparisons between the de pths calculated by the integrated localization algorithm and depth meter are shown in Figure 44.



（1）depth of AUV calculated by integrated algorithm and depth meter

（2）error between the results of integrated algorithm and depth meter

Figure 44 Depth Calculated by Integrated Algorithm and Depth Meter

It can be seen that when AUV is relatively far from the docking device, the error of depth positioning is larger in the initial few seconds. There are two reasons for it. On the one hand,

due to the small imaging area of guiding lights, the error of centroid extraction and positioning is larger. On the other hand, as mentioned before, the error of fusion algorithm is larger at the beginning, but it converges gradually over time. With the reduction of relative distance, the positioning accuracy of monocular camera is improved, and the errors of fusion algorithms gradually declines. In the moving of AUV, although the depth data can't be measured directly, the algorithm is verified to a certain extent by the data of depth meter.

5.3 Chapter Summary

Relying on the actual IMU chip, data of actual sensors were read through the lower computer and used to calculate the attitude, verifying the effectiveness of the algorithm under static and dynamic conditions. Based on practical pool experiments, the position estimation algorithm is used to locate the AUV in the docking process and the data of depth meter basically validate the model and estimation algorithm. Generally speaking, the experimental results are consistent with the simulation results, but at the same time, some shortcomings are found, and there is still room for improvement of the algorithm.

Chapter 6 Integrated Navigation Based on IMU/USBL System

In the above sections, we focus on the docking procedure of AUV in short distance, but in practical analysis, it is the prerequisite that the AUV approach the neighbor region of docking device with an acceptable error, thus it is also meaningful to study the navigation of AUV in the relatively long distance. As discussed above, due to the inherent characteristics of dead reckoning, the positioning error of inertial system will accumulate as the time goes by and may reach a level which is inacceptable from the point of docking. Although the radio navigation, GPS navigation and celestial navigation have been widely applied for the autonomous vehicles on land, considering the sharp attenuation of electromagnetic wave, they can't fit the needs of underwater navigation to some extent. However, the sound travels underwater around a rate of 1500m/s, which is almost 4 to 5 times of that in the air, and the attenuation of sound wave is apparently less severe than that of electromagnetic wave. In this context, the acoustic positioning method has been studied and applied to a wide range of fields, including the mineral exploration, removal of mines, terrain landform analysis, and so on[58-60]. In the docking of AUV, acoustic positioning device can be used as the indicator in the middle and relatively long distance, achieving integrated navigation together with the dead reckoning process of IMU.

6.1 Acoustic Positioning and USBL Algorithm

6.1.1 Characteristics of Acoustic Positioning

As local underwater positioning systems, in terms of the array spacing, acoustic positioning systems can be divided into Long Base Line(LBL), Short Base Line(SBL) and Ultra Short Based Line(USBL) systems. The base line refers to the distance between acoustic sensors, for the LBL system, the length of base line can range from 100m to 6000m, while that length is usually less than 50m in the SBL system and even less than 1m in USBL systems. In most cases, the base lines in USBL systems are shorter than the half-wavelength of acoustic wave and are of the order of 10^{-1} m, while the acoustic responders are arranged into a certain geometrical array[58]. Different from LBL and SBL systems, the USBL positioning systems measure the relative distance based on the phase method, obtaining the azimuth angle and slope

distance in terms of the phase difference between responder signals and the answering time respectively. Due to the advantages of simple structure, convenient installation and maintenance as well as the strong flexibility, the USBL positioning technology has developed rapidly since 1990s and has been applied to a large number of underwater vehicles.

6.1.2 Positioning Principle of USBL System

For the USBL positioning system in this part, four hydrophones are installed on the docking device in one plane, and it is assumed that the two based lines are orthogonal to each other and form the x-axis and y-axis respectively. The cross point of two base lines are set as the original point of coordinate system, thus when the separation distances between hydrophones on one axis are assumed as d, then the coordinates of hydrophones can be obtained as $(\pm d/2, \pm d/2)$. Assuming that the position of AUV with respect to the center of hydrophone array is $S(x, y, z)$, then the schematic of USBL positioning is given by figure 45.

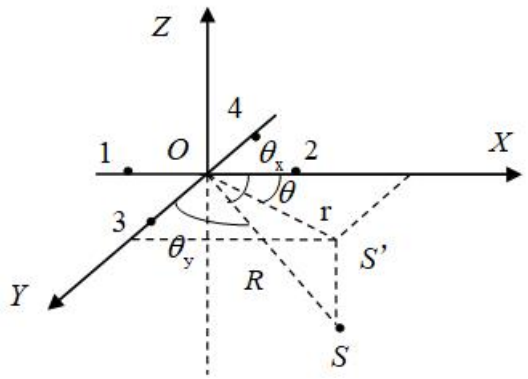


Figure 45 Schematic of USBL Positioning

As shown above, the radius vector of target is denoted by $\overline{os}$, when the included angles between it and the coordinate axes are denoted by θ_x and θ_y, the relation between the position of target and radius vector can be represented as

$$\cos\theta_x = \frac{x}{R}, \cos\theta_y = \frac{y}{R}$$
$$R = \sqrt{x^2 + y^2 + z^2} \tag{6.1}$$

In the schematic, the projection of radius vector $\overline{os}$ on OXY plane is denoted by r, based on the geometrical relation, the azimuth angle and position parameters can be calculated in terms of $\theta = \tan^{-1}\frac{y}{x} = \tan^{-1}\frac{\cos\theta_y}{\cos\theta_x}$, $r = \sqrt{x^2 + y^2}$ and $z = -\sqrt{R^2 - r^2}$. Given the small dimension of acoustic array in USBL system, when the acoustic signal is sent from a relatively far filed, according to the principle of plane wave approximation, the acoustic waves shot into the array can be regarded to be parallel with each other. Under this assumption, when the phase difference angles of hydrophones on x-axis and y-axis are denoted by φ_x and φ_y respectively, they subject to the relation that

$$\varphi_x = \frac{2\pi d\cos\theta_x}{\lambda},\ \varphi_y = \frac{2\pi d\cos\theta_y}{\lambda} \tag{6.2}$$

in which λ is the wavelength of sound wave. Substituting them into the formulas above, it can be obtained that $x = \frac{\lambda\varphi_x R}{2\pi d}, y = \frac{\lambda\varphi_y R}{2\pi d}, z = -\sqrt{R^2 - \frac{\lambda^2 R^2}{4\pi^2 d^2}\left(\varphi_x^2 + \varphi_y^2\right)}$. The slope distance from the center of hydrophone array to the target can be calculated by $R = c\cdot T$, in which c is the sound velocity underwater and $2T$ denotes the time difference between the emission and reception of acoustic signal. To sum up, the measurement values of USBL positioning system are the phase difference angles ϕ_x, ϕ_y and the slope distance R. Once these values are given, then the three-dimensional relative position of target $S(x, y, z)$ in the acoustic positioning coordinate system, the azimuth angel θ and the slope distance R can be determined.

6.2 Error Analysis of USBL Positioning System

Generally speaking, the positioning accuracy of USBL system can be measured by the ratio between the positioning error of a certain axis and the slope distance, namely the relative error. Based on the phase difference relationship, the three-dimensional position of AUV can be calculated by

$$x = \frac{\lambda\varphi_x R}{2\pi d}, y = \frac{\lambda\varphi_y R}{2\pi d}, z = -\sqrt{R^2 - \frac{\lambda^2 R^2}{4\pi^2 d^2}\left(\varphi_x^2 + \varphi_y^2\right)} \tag{6.3}$$

Taking the x-axis as an example, by calculating the full differential on both sides, it can be

obtained that

$$\Delta x = \frac{\lambda \varphi_x}{2\pi d}\Delta R + \frac{\lambda R}{2\pi d}\Delta \varphi_x + \frac{R\varphi_x}{2\pi d}\Delta\lambda - \frac{\lambda R\varphi_x}{2\pi}\frac{\Delta d}{d^2} \tag{6.4}$$

On this basis, the relative positioning error of x-axis is

$$\frac{\Delta x}{x} = \frac{\Delta R}{R} + \frac{\Delta \varphi_x}{\varphi_x} + \frac{\Delta\lambda}{\lambda} - \frac{\Delta d}{d} \tag{6.5}$$

For $R = c \cdot T$ and $c = \lambda \cdot f$,taking full differential on both sides, then we can get

$$\frac{\Delta R}{R} = \frac{\Delta T}{T} + \frac{\Delta c}{c}, \quad \Delta\lambda = \frac{\Delta c}{f} \tag{6.6}$$

Substituting them into the formulas above, the relative positioning accuracy of x-axis can be represented as

$$\frac{\Delta x}{x} = \frac{\Delta T}{T} + \frac{\Delta \varphi_x}{\varphi_x} + 2\frac{\Delta c}{c} - \frac{\Delta d}{d} \tag{6.7}$$

Similarly, the relative positioning accuracy can also be calculated in terms of the slope distance as

$$\begin{aligned} \frac{\Delta x}{R} &= \frac{\lambda \varphi_x}{2\pi d}\left(\frac{\Delta T}{T} + \frac{\Delta \varphi_x}{\varphi_x} + 2\frac{\Delta c}{c} - \frac{\Delta d}{d}\right) \\ &= \left(\frac{\Delta T}{T} + \frac{\Delta \varphi_x}{\varphi_x} + 2\frac{\Delta c}{c} - \frac{\Delta d}{d}\right)\cos\theta_x \end{aligned} \tag{6.8}$$

In the docking process of AUV, it is supposed that the acoustic sensors are installed to designated spots and firmly fixed, then the error of array spacing Δd can be neglected. Assuming that all the error terms are independent from each other, then the RMS value of relative positioning error with respect to slope distance can be given by

$$\delta_x = \frac{\lambda \varphi_x}{2\pi d}\left(\frac{\Delta T}{T} + 2\frac{\Delta c}{c}\right) + \frac{\lambda \Delta \varphi_x}{2\pi d} \tag{6.9}$$

Under normal circumstances, the relative measurement errors of time difference $\Delta T / T$ and sound velocity $\Delta c / c$ are of the order of 10^{-3}, as a result, the measurement error induced by phase difference plays the most important part in the positioning[60].

6.3 Integrated Navigation System of IMU/USBL Based on Filtering Algorithm

In this part, to synthesize the measurements of inertial navigation system and the USBL positioning system, mathematical models of integrated navigation system based on IMU/USBL are constructed. Generally, the integrated filtering algorithm includes the direct filtering and the indirect filtering. In the direct estimation method, the parameters such as velocity and position are selected as the estimated state, while in the indirect method, the state vector is composed by the estimation errors of main navigation system, which is the inertial navigation system in this work. To achieve direct estimation of navigation parameters, the kinetic model of carrier has to be built to describe the dynamic change of carrier states. However, in indirect estimation , the error state transition matrix is obtained by taking full differential of the kinetic models, and is usually linearized for simplicity. In this section, the direct filtering method is chosen to estimate the position of carrier. Simulation has been done to test the effects of different methods and the results are compared and analyzed briefly.

6.3.1 Direct Estimation of Navigation Parameters of IMU/USBL System

In this section, according to the integrated positioning model built in section 4.3.2, the navigation parameters of carrier can be estimated directly . In the direct estimation process, the measurements of IMU and USBL system are used to adjust the prediction of system state at the update step of separately, then the filtering outputs are combined to obtain the final estimation results through an adaptive fusion process. The state vector is chosen to be $X=\begin{bmatrix} p_x & p_y & p_z & v_x & v_y & v_z & a_x & a_y & a_z & \delta a_x & \delta a_y & \delta a_z \end{bmatrix}^T$, then the state transition matrix can be given by

$$\Phi=\begin{bmatrix} I_{3\times3} & T\cdot I_{3\times3} & \frac{T^2}{2}\cdot I_{3\times3} & -\frac{T^2}{2}C_b^n\cdot I_{3\times3} \\ O_{3\times3} & I_{3\times3} & T\cdot I_{3\times3} & -T\cdot C_b^n\cdot I_{3\times3} \\ O_{3\times3} & O_{3\times3} & I_{3\times3} & O_{3\times3} \\ & O_{3\times9} & & I_{3\times3} \end{bmatrix},$$

where C_b^n denotes the attitude transfer matrix from the carrier coordinate system to the navigation coordinate system. For the INS system, the measurement vector is determined as

$Z=\begin{bmatrix} p'_x & p'_y & p'_z & v'_x & v'_y & v'_z & a'_x & a'_y & a'_z \end{bmatrix}^T$, thus the measurement matrix can be built as

$$H=\begin{bmatrix} I_{3\times3} & O_{3\times3} & O_{3\times3} & O_{3\times3} \\ O_{3\times3} & I_{3\times3} & O_{3\times3} & O_{3\times3} \\ O_{3\times3} & O_{3\times3} & I_{3\times3} & O_{3\times3} \end{bmatrix} \tag{6.10}$$

It is apparent that both the transition matrix and the measurement matrix are linear, thus the discrete Kalman filter equations used here are similar to those listed in the previous sections, so they are not repeated here. For the USBL positioning system, the measurements are the phase difference angles on orthogonal axes φ_x and φ_y, as well as the slope distance R between the original point to the target. Based on analysis above, it is obvious that the state equations are linear, while the measurement equations $Z(k)=h\left[X(k)\right]+V(k)$ are nonlinear, thus the navigation state update can be boiled down to a nonlinear estimation problem and the extended Kalman filter is constructed. The Jacobian matrix of measurements can be obtained as

$$\begin{aligned} H&=\frac{\partial Z_{USBL}}{\partial X}=\frac{\partial\begin{bmatrix}\varphi_x & \varphi_y & R\end{bmatrix}^T}{\partial X} \\ &=\begin{bmatrix} \frac{2\pi d\left(R-R_x\right)}{\lambda R^2} & -\frac{2\pi dxR_y}{\lambda R^2} & -\frac{2\pi dxR_z}{\lambda R^2} & O_{1\times9} \\ -\frac{2\pi dxR_x}{\lambda R^2} & \frac{2\pi d\left(R-R_y\right)}{\lambda R^2} & -\frac{2\pi dxR_z}{\lambda R^2} & O_{1\times9} \\ R_x & R_y & R_z & O_{1\times9} \end{bmatrix} \end{aligned} \tag{6.11}$$

in which $R_x=\frac{x}{\sqrt{x^2+y^2+z^2}}, R_y=\frac{y}{\sqrt{x^2+y^2+z^2}}, R_z=\frac{z}{\sqrt{x^2+y^2+z^2}}$. The recursive equations of extended Kalman filter can be given as

$$\begin{aligned} &\hat{X}\left(k+1|k\right)=\Phi\left(k+1|k\right)\hat{X}\left(k|k\right) \\ &P\left(k+1|k\right)=\Phi\left(k+1|k\right)P\left(k|k\right)\Phi^T\left(k+1|k\right)+Q\left(k+1\right) \\ &K\left(k+1\right)=P\left(k+1|k\right)H^T\left(k+1\right)\left[H\left(k+1\right)P\left(k+1|k\right)H^T\left(k+1\right)+R\left(k+1\right)\right]^{-1} \\ &\hat{X}\left(k+1|k+1\right)=\hat{X}\left(k+1|k\right)+K\left(k+1\right)\left[Z\left(k+1\right)-h\left(\hat{X}\left(k+1|k\right)\right)\right] \\ &P\left(k+1\right)=\left[I-K\left(k+1\right)H\left(k+1\right)\right]P\left(k+1|k\right) \end{aligned} \tag{6.12}$$

6.3.2 Weighting Fusion of Sub-filters and Adaptive Kalman Filtering

Based on the model and recursive equations, the simulation program is developed in MATLAB .Considering the difference in measurement matrix, the simulated measurements of IMU and USBL positioning system are used to update the estimation of carrier position separately and independently , then the results are fused through an adaptive weighting fusion algorithm. Assuming that when the measurements of IMU and USBL systems are updated simultaneously, the weighting factors assigned to them are denoted by W_{INS} and W_{USBL}. If the local unbiased estimation results and the covariance matrices are represented by $\hat{X}_{INS}$、P_{INS} and $\hat{X}_{USBL}$、P_{USBL} respectively, then the fused estimation of AUV position $\hat{X}$ and the weighting factors satisfy the relationship that

$$\begin{cases} \hat{X} = W_{INS} \cdot \hat{X}_{INS} + W_{USBL} \cdot \hat{X}_{USBL} \\ W_{INS} + W_{USBL} = 1 \end{cases} \tag{6.13}$$

When the true state is denoted by X ,the total mean square error can be calculated by

$$P = E\left(X - \hat{X}\right)^2 = E\begin{bmatrix} W^2{}_{INS}\left(X - \hat{X}_{INS}\right)^2 + W^2{}_{USBL}\left(X - \hat{X}_{USBL}\right)^2 \\ -2W_{INS}W_{USBL}\left(X - \hat{X}_{INS}\right)\left(X - \hat{X}_{USBL}\right) \end{bmatrix} \tag{6.14}$$

Considering that $\hat{X}_{INS}$ and $\hat{X}_{USBL}$ are the unbiased estimation of carrier state X and are independent from each other, it is natural to obtain that $E\left[\left(X - \hat{X}_{INS}\right)\left(X - \hat{X}_{USBL}\right)\right] = 0$ and $P = W^2{}_{INS}P_{INS} + W^2{}_{USBL}P_{USBL}$ .To obtain the minimum total mean square error, the local estimation results are fused as

$$\begin{cases} P = \left(P^{-1}{}_{INS} + P^{-1}{}_{USBL}\right)^{-1} \\ \hat{X} = P^{-1}{}_{INS}P\hat{X}_{INS} + P^{-1}{}_{USBL}P\hat{X}_{USBL} \end{cases} \tag{6.15}$$

The simulation results are given in figure 46. Figure (a) and figure (b) show the integrated filtering output when the initial value of covariance matrices are set relatively precisely, while figure (c) and (d) are the filtering results when the initial values of covariance deviate from the truth value to some extent. It can be seen that although in both cases the integrated filtering can

reduce and restrain the positioning error to a certain degree, the satisfying effect of filtering relies on the precise evaluation of noise covariance, which is hard to achieve in practical application. In fact, in the complex underwater environment, outliers tend to occur in the measurements. To handle the uncertainty of measurement noise R, the estimation of R matrix is adjusted based on the Sage-Husa algorithm.

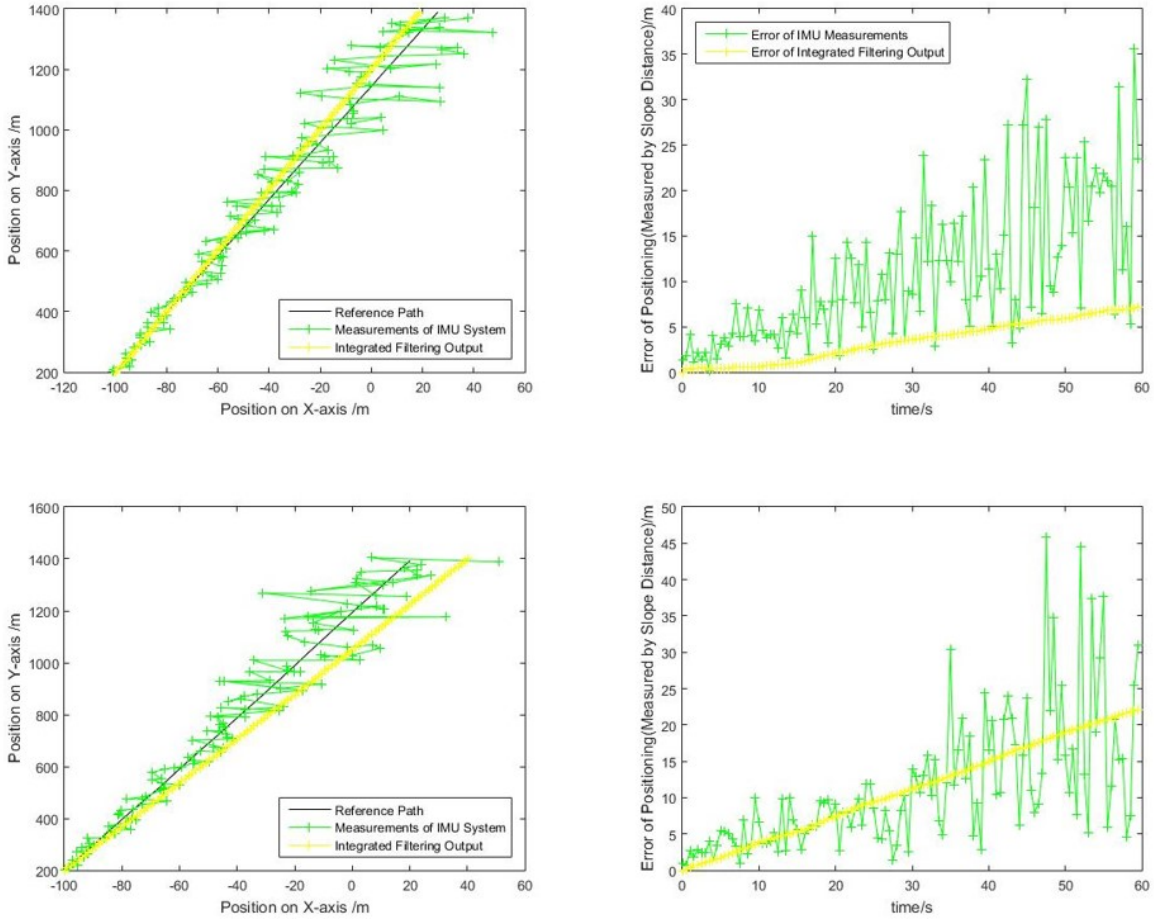


Figure 46 Results of Weighting Fusion Algorithm and the Comparison

The Sage-Husa adaptive filtering algorithm can be seen as the combination of regular Kalman filter and the time-variant noise estimator. In the filtering process, the state of system is updated based on Kalman filter, meanwhile the process noise and measurement noise are estimated by the time-variant noise estimator to fix the model of system, providing a more accurate model for the successive state estimation[61-62]. Due to the relatively simple structure and the good filtering performance, the Sage-Husa algorithm has been widely used in engineering practice.

For the linearized system, the specific formulas of Sage-Husa adaptive filtering can be given as

$$\begin{aligned}
&\hat{X}\left(k+1|k\right)=\Phi\left(k+1|k\right)\hat{X}\left(k|k\right)+\hat{q}\left(k\right)\\
&P\left(k+1|k\right)=\Phi\left(k+1|k\right)P\left(k|k\right)\Phi^{T}\left(k+1|k\right)+\hat{Q}\left(k+1\right)\\
&K\left(k+1\right)=P\left(k+1|k\right)H^{T}\left(k+1\right)\left[H\left(k+1\right)P\left(k+1|k\right)H^{T}\left(k+1\right)+\hat{R}\left(k+1\right)\right]^{-1}\\
&\hat{X}\left(k+1|k+1\right)=\hat{X}\left(k+1|k\right)+K\left(k+1\right)\left[Z\left(k+1\right)-H\left(k+1\right)\left(\hat{X}\left(k+1|k\right)\right)-\hat{r}\left(k\right)\right]\\
&P\left(k+1\right)=\left[I-K\left(k+1\right)H\left(k+1\right)\right]P\left(k+1|k\right)
\end{aligned}\tag{6.16}$$

where the $\hat{r}\left(k\right)$、$\hat{q}\left(k\right)$、$\hat{R}\left(k+1\right)$、$\hat{Q}\left(k+1\right)$ are estimated recursively by

$$\begin{aligned}
&d\left(k\right)=\left(1-b\right)/\left(1-b^{k+1}\right)\\
&\hat{r}\left(k+1\right)=\left(1-d\left(k\right)\right)\hat{r}\left(k\right)+d\left(k\right)\left(Z\left(k+1\right)-H\left(k+1\right)X\left(k+1|k\right)\right)\\
&\hat{R}\left(k+1\right)=\left(1-d\left(k\right)\right)\hat{R}\left(k\right)+d\left(k\right)\left(Z\left(k+1\right)Z\left(k+1\right)^{T}-H\left(k+1\right)P\left(k+1\right)H\left(k+1\right)^{T}\right)\\
&\hat{q}\left(k+1\right)=\left(1-d\left(k\right)\right)\hat{q}\left(k\right)+d\left(k\right)\left(\hat{X}\left(k+1\right)-\Phi\left(k+1|k\right)\hat{X}\left(k\right)\right)\\
&\hat{Q}\left(k+1\right)=\left(1-d\left(k\right)\right)\hat{Q}\left(k\right)+d\left(k\right)\begin{pmatrix}K\left(k+1\right)Z\left(k+1\right)Z\left(k+1\right)^{T}K\left(k+1\right)^{T}+P\left(k+1\right)\\-\Phi\left(k+1|k\right)P\left(k\right)\Phi\left(k+1|k\right)^{T}\end{pmatrix}
\end{aligned}\tag{6.17}$$

in which b is the forgetting factor ranging from 0 to 1, and in most cases, its value is set between 0.95 and 0.99. The simulation results are given in figure 47 and the outputs of non-adaptive and adaptive filtering are compared in the same figure. It can be seen from the figure that although the non-adaptive filtering has better performance in the early stage, in the long run, the adaptive filtering algorithm fits the fluctuation of noise covariance better and restrains the error within a smaller range.

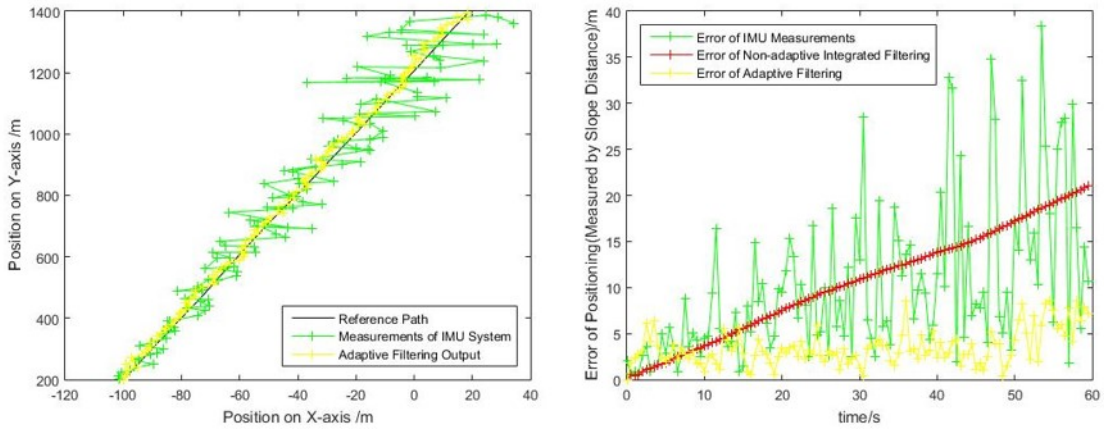


Figure 47 Results of Adaptive Filtering and Non-adaptive Filtering

Table 5 Root Mean Square Errors of Integrated Filtering Algorithm

Variable	Non-adaptive(m)	Adaptive(m)
RMS Error of X-axis	5.78	2.32
RMS Error of Slope	10.39	3.48

6.4 Chapter Summary

In this chapter, the inertial navigation unit and acoustic positioning system are combined to estimate the position of AUV in the middle and relatively long range. To compensate the accumulated error of dead reckoning, the acoustic positioning method has been studied. Considering the cost, complexity of structure and application scenario, this work only focus on the underwater positioning based on USBL system. In this section, firstly the positioning principle of USBL system is introduced, and the main sources of error are analyzed. On that basis, the filtering equations are built and the measurement matrix of USBL system is linearized to perform extended Kalman filtering given the non-linearity, and the position is estimated combining the inertial and acoustic measurements through a weighting fusion algorithm. In addition, considering the uncertainty of noise covariance, the Sage-Husa algorithm is used to achieve adaptive Kalman filter. In the programming, the covariance of measurement noise is set as time-variant, and the effectiveness of adaptive algorithm is verified through simulation.

Conclusion

This paper mainly focuses on the position and attitude estimation of AUV in the process of docking under the background of multi-sensor data fusion. The proposed algorithm is verified by simulation and experiment. Considering the average cost of AUV and the configuration of common sensors, after analyzing the principle and characteristics of common underwater positioning sensors, the IMU with high sampling rate and monocular camera are selected for close-end docking, and the USBL positioning method is studied for navigation in a relatively longer distance. When the types of sensors are fixed and the accuracy of each single sensor is limited, how to improve the accuracy of attitude and position estimation by fusing different sensor information, is the focus of this paper.

The contents and accomplishments of this paper are as follows:

(1) The basic models of inertial navigation and monocular vision are studied. Using the guiding lights whose relative position and distance are known, the three-dimensional relative coordinates are calculated based on the perspective relation. Considering the yaw and pitch of AUV, the model of monocular vision positioning is corrected to improve its accuracy. After comparing different attitude calculation algorithms, the algorithm based on quaternion is selected to lay the foundation for the subsequent error state model.

(2) Analyzing the characteristics of underwater image and several image filtering methods, using different methods to process the actual underwater image, comparing the filtering effects, and choosing the Gauss filter to preprocess the image. In order to extract the targets, several image segmentation methods are analyzed. In view of the light scatters in water, multi operators with different criteria are defined, and the idea of fusing different operators is chosen to improve the reliability of image segmentation.

(3) In order to achieve the fusion of different operators, an improved DS evidence theory is proposed. Firstly, the idea of pre-processing the evidence first and then fusing it is determined. In order to evaluate the quality of evidence, an evidence discounting algorithm based on TOPSIS is proposed, which assigns higher weights to evidences with higher quality.

By calculating the case, the improved D-S evidence theory algorithm is validated, and the results are compared with those of similar methods proposed in recent years, which proves that the algorithm has certain advantages in the final probability allocation. Based on this, a self-adaptive threshold adjusting method is proposed to achieve image segmentation. Then, the position of AUV relative to the guiding lights is calculated, and the algorithm is preliminarily tested by experiment.

(4) Considering the inherent shortcomings of monocular vision positioning and inertial sensor navigation, after analyzing the source of sensor errors, the error state model of attitude estimation and state model of position estimation are established. The gyroscope data are corrected by fusing accelerometer and magnetometer data to obtain final attitude, and the accelerometer data are corrected by visual positioning results to estimate the position. By consulting the data sheets, the simulation parameters are set. The effectiveness of proposed algorithms is verified by comparison and error analysis.

(5) Taking the acoustic positioning method into account. Considering the cost, complexity of structure and application scenario, this work only focus on the underwater positioning based on USBL system. In this section, the positioning principle of USBL system is introduced, and the main sources of error are analyzed. On that basis, the filtering equations are built and the position is estimated combining the inertial and acoustic measurements through a weighting fusion algorithm. In addition, the Sage-Husa algorithm is used to achieve adaptive Kalman filter.

(5) The effectiveness of proposed algorithm to estimate the position and attitude of AUV in close-end docking is further verified by experiments. In general, the experimental results are consistent with the simulation results, but at the same time, some shortcomings are found and further research can be carried out from the following points:

(a) Improving the accuracy of hardware system. The nine-axis gyroscope module used in the solution is a relatively low-cost basic product, which limits the accuracy of attitude measurement on AUV. As a result, the effect of attitude estimation is not ideal enough in the experiment, especially for the yaw angle.

(b) Optimizing the layout of sensors. Monocular camera is used in visual positioning. In

the experiment, it is found that the camera can ensure good positioning accuracy when the relative distance is small, but in the long distance, because the imaging areas of guiding lights are too small, the accuracy of image segmentation and centroid extraction decreases considerably, resulting in a large error at a long distance and influencing the final fusion. In the future research, maybe we can try the binocular camera positioning method to further improving the accuracy of vision positioning.

References

[1] 陈强，兰晓娟，王霜.国外UUV系统在海洋调查中的应用[J].舰船科学技术，2012,34(10)：134-140.

[2] 李壮.短基线定位关键技术研究[D].哈尔滨工程大学博士论文，2013.

[3] 孙微，王方勇.基于长基线水声定位系统误差分析以及定位精度研究[J].2016,2,14.

[4] J W O. An introduction to inertial navigation[J]. University of Cambridge, Computer Laboratory, Tech. 2007:696

[5] Monique Chyba. Autonomous underwater vehicles. Ocean Engineering. 2009(36).

[6] Kuan Meng Tan, Tommie Liddy. The advancement of an autonomous underwater vehicle(UUV) Technology. 3rd IEEE Conference on Industrial Electronics and Applications.2008,336-341.

[7] Hillenbrand Christopher. UUV science&technology trends. Sea Technology. 1997,38(12):10-13.

[8] 陈强, 张林根. 美国军用 UUV 现状及发展趋势分析[J]. 舰船科学技术, 2010, 32(7): 129-134.

[9] 钟宏伟.国外无人水下航行器装备与技术现状及展望[J].水下无人系统学报，2017, 3(25):216.

[10] 肖玉杰，邱志明，石章松.UUV国内外研究现状及若干关键问题综述[J].电光与控制.2014,21(2):46-50.

[11] 钟宏伟. 武器级 UUV 发展趋势及关键技术研究[R]. 北 京: 中国国防科学技术报告管理办公室, 2011(12).

[12] 国外舰船装备与技术发展报告编写组. 国外舰船武器装 备与技术发展报告 2013: 海上无人系统(无人潜航 器)[R]. 北京: 中国船舶重工集团公司, 2014.

[13] 孙现有，焦泽键，孙长会.无人水下航行器发展与应用[J].水雷战与舰船防护.2012,20(2):51-55.

[14] 胡必文，程彬彬.浅析无人水下作战平台现状及未来作战使命[J].水雷战与舰船防护.2013,21(4):67-72.

[15]胡玉梅.无人水下航行器的发展与展望[J].电子世界.2013,14(2):71-73.

[16]王晓娟.基于视觉的AUV水下回收导引定位技术研究[D].哈尔滨工程大学，2011.

[17]秦永元.惯性导航[M].北京：科学出版社，2006.

[18]张天光等译.捷联惯性导航技术[M].第二版.北京：国防工业出版社,2007.

[19]徐爱东，杨拥军，卢新艳. MEMS惯性器件的新进展及应用[Z].中国浙江杭州:20105.

[20]熊敏君，卢惠民，熊丹.基于单目视觉与惯导融合的无人机位姿估计[J].计算机应用.2017,37(S2),127-133.

[21]娄路.基于视觉和MEMS-IMU融合的火星车导航定向技术.航天控制.2012,4(30),32-36.

[22]Erik Murphy-Chutorian, Mohan Manubhai Trivedi. Head pose estimation and augmented reality tracking： An integrated system and evaluation for monitoring driver awareness[J]. IEEE Transactions on Intelligent Transportation Systems，2010，11(2)：300-311.

[23]钱 山. 在轨服务航天器相对测量及姿态控制研究 [D]. 长沙：国防科学技术大学航天科学与工程学 院，2010.

[24]Yager R R. Using approximate reasoning to represent default knowledge[J]. Artificial Intelligence, 1987, 31(1):99-112.

[25]Dubois D, Prade H. Represent and combination of uncertainty with belief functions and possibility measures[J]. Comput. Intell. 1988, 4(3): 244-264.

[26]Lefevre E, Colot O, et al. A generic framework for resolving the confliction the combination of belief structures[C]. The 3rd International Conference on Information Fusion, 2000, 29(12): 11-18.

[27]Yager R R. On the Dempster-Shafer framework and new combination rules[J].IEEE Trans on System. 1989,41(2):93-137.

[28]黎湘,刘永祥,付耀文等.基于D-S证据理论的修正融合目标识别模型.自然科学进展[J].2000,10(ll):1040-1043.

[29]邓勇,施文康,朱振福.一种有效处理神突证据的组合规则[J].红外与毫米波学报.2004,23(1):27-1278.

[30]陈博,万寿红,岳丽华.改进的DS证据舰船融合检测研究[J].计算机工程与应用.2010,46(28):222-224.

[31]刘和祥.面向AUV回收的水下机器视觉研究[D].哈尔滨工程大学博士论文.2009.

[32]施小成，郝丽超，张伟，吴迪.基于Blob分析和贝叶斯决策的水下目标提取方法[J].计算机应用.2012,32(11):3214-3217.

[33]Bisseling R W,Hof A L.Handling of Impact Forces in Inverse Dynamics[J].Journal of Biomechanics.2006,39(13):2438-2444.

[34]Lee K,Kim Y,Yun J,et al.Magnetic-interference-free Dual-electric Compass[J].Sensors and Actuators A:Physical.2005,120(2):441-450.

[35]Jin H,Sciammarella C,Yoshida S,et al.Three-dimensional Underwater Measuring by Structured Light Projection[M].Advancement of Optical Methods in Experimental Mechanics, Volume 3, Springer International Publishing,2014,77-83.

[36]K. Haris, S. N. Efstratiadis, N. Maglaveras. Image noise reduction based on local classification and iterated conditional modes[C],in Proc. IWISP, Manchester, U.K.1996:4-7.

[37]Yi Zhang, Kai Lu, YingHui Gao. QSobel: A novel quantum image edge extraction algorithm[J].2015,58(1),1-13.

[38]周莉莉，姜枫.图像分割方法综述研究[J].计算机应用研究,2017,34(7):1921-1924.

[39]Shilong Wang, Yuru Xu,Yongjie Pang. A fast underwater optical image segmentation algorithm based on a histogram weighted fuzzy c-means improved by PSO[J].2011,10(1):70-75.

[40]张铭钧,李煊，王玉甲.基于灰度化权值调整的水下彩色图像分割[J].哈尔滨工程大学学报,2015,36(5),707-713.

[41]Dempster, A.P.Upper and lower probabilities induced by a multivalued mapping[J],Annals of Mathematical Statistics, vol. 38, no. 2, pp. 325-339, 1967.

[42]Zadeh L.A. “A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination,” AI Mag. Vol.7, no.2, pp.85-90, 1986.

[43]Shafer G.A. A mathematical theory of evidence, Princeton: Princeton University Press, 1976.

[44]付巧峰.关于TOPSIS法的研究[J].西安科技大学学报，2008，1(1),190-193.

[45]张目，周宗放.一种基于联系度的改进TOPSIS法[J]. 系统工程，2008,8,102-107.

[46]IEEE Standard for Interval Arithmetic, IEEE Std 1788, June 2015.

[47]Yager R.R. “Hedging in the combination of evidence,” J.Inf.Optim. Sci.,vol.4, no.1, pp.73-81, 1983.

[48]Jiang, W., Zhang A., Deng Y., “Dynamic determination of sensor credibility in data fusion and its application,” Journal of Harbin Institute of Technology, vol. 42, no.7, pp. 1137-1140. 2010.

[49]石美红，王文光.基于Blob算法的织物疵点检测算法的研究[J].现代电子技术.2010,33(24):29-33.

[50]段康容，刘先勇.基于Blob算法的标记圆检测技术研究[J].传感器世界.2014,20(8):7-11.

[51]王勋，左启耀，洪诗聘.一种基于卡尔曼滤波的定位解算性能评估新方法[J].导航定位与授时,2018,5,73-75.

[52]Salmond, D., Target tracking: Introduction and Kalman tracking filters[A], in Target Tracking: Algorithms and Applications[C]. 2001:1/1-1/16.

[53]史龙，李威，雷志荣. 低成本AHRS技术综述[C]. 2016 IEEE Chinese Guidance, Navigation and Control Conference (IEEE CGNCC2016), 2016,2666-2670.

[54]顾姗姗.光纤捷联惯性组合导航系统建模及滤波技术研究[D].南京航空航天大学硕士论文，2016.

[55]伊程毅.基于地磁和微惯性器件组合的姿态测量系统研究[D]. 哈尔滨工业大学硕士论文，2013.

[56]夏家和，秦永元.摇摆状态下基于非线性误差模型的惯导对准研究[J].宇航学报，2010(2),413-415.

[57]Xue L,Yuan W-z, ChangH-l, Qin W. Application of quaternion-based ex-tended Kalman filter for MAV attitude estimation using mems sensors[J]. Nanotechnology and Precision Engineering, 2009, 7(3):163-167.

[58]Zhao J-b, Ge X-y. A Review of Underwater SINS/DVL Integrated Navigation Technology[J]. JOURNAL OF UNMANNED UNDERSEA SYSTEMS, 2018,26(1):2-4.

[59]Dinc M, Hajiyev C. Integration of Navigation Systems for Autonomous Underwater

Vehicles[J]. Journal of Marine Engineering & Technology, 2015, 14(1): 32-43.

[60]徐成龙.基于多传感器融合水下机器人导航方法的研究[D].东北大学硕士论文，2017.

[61]Xu X-s, Pan Y-f. SINS/DVL Integrated Navigation System Based on Adaptive Filtering[J].J.Huazhong Univ. of Sci.&Tech.2015, 43(3):95-98.

[62]Zhang T, Chen L-p. AUV Underwater Positioning Algorithm Based on Interactive Assistance of SINS and LBL[J]. Sensors,2016,42(16):1-6.

Paper Published During the Study

[1] Yixuan Geng, Shaoping Wang, Xingjian Wang, A Novel Conflicting Evidence Discounting Method Based on TOPSIS Decision-Making[C]. 2018 IEEE/CSAA Guidance, Navigation and Control Conference,2018: 996-1001.

Acknowledgement

The three-year experience is coming to an end, and I would like to express my gratitude to all those who helped me during the master's career.

Firstly, I gratefully acknowledge the help of my supervisors, Professor Wang Shaoping and Professor Ferrero Alessandro. They have offered me valuable suggestions and incisive comments, which contributed a lot to the completion of this thesis. I do appreciate their patience, encouragement and professional instructions. Without their help, this thesis wouldn't be what it is now.

I am greatly indebted to all the teachers participating in the double degree program of POLIMI and BUAA. They have helped me both in studies and in life, from which I benefited a lot. This program gave me a valuable experience in Milan, broadened my horizons, enriched my experience, and created a good memory for me. Thanks for you all.

Last but not the least, my gratitude also extends to my family, my boy-friend, my classmates and roommates, who have been assisting, encouraging and supporting me during the writing of this thesis, and making their comments on the draft. I owe my gratitude to those who helped me to tackle my problems, and I wish you all the best!