A Ship Ranging and Positioning System Based on Binocular Vision Technology

Yabo Zheng, Lixia Wang, Ming Zhou and Yu Wang
Dalian Naval Academy, Dalian, China
Email: 2262109914@qq.com

Abstract. Relative ranging and positioning is the basis to ensure the navigation safety of ships, and it has a broad application prospect in various maritime activities, especially in the aspects of docking and undocking, sailing in crowded areas, towing, convoy sailing, and warship replenishment, etc. In order to improve the autonomy and automation of ship ranging and positioning and reduce the cost, a novel type of ship ranging and positioning system based on binocular vision technology is proposed. The proposed system includes seven modules, i.e. video information acquisition module, ship attitude acquisition module, camera parameter calibration module, target recognition and segmentation module, image matching module, ranging and positioning module, and information display module. After acquiring the target image sequences by using two cameras, the proposed system identifies and segments maritime targets quickly by using Canny algorithm in the artificial interactive mode, realizes reduced-space search by using epipolar constraint method, achieve fast target features matching to obtain the vision disparity of the feature points by using the SURF algorithm, and calculates the distance of the target according to the binocular vision ranging principle, so as to realize interactive visual ranging and positioning. The proposed system has the advantages of low cost, simple structure and convenient operation. It can realize relative ranging and positioning with only one computer and two cameras. Experiments demonstrate that the proposed system, with fast running speed and high accuracy, can satisfy the relative ranging and positioning needs of ships.

1. Introduction
The traditional ship ranging and positioning mainly adopts two methods, i.e. laser ranging method and GPS relative positioning method. However, both of them have obvious defects: the laser ranging method requires multi-point manual intervention and has a low degree of automation; GPS relative positioning method has a higher degree of automation, but it is highly dependent on satellite navigation system and communication system, and it will be susceptible to interference in wartime. In order to reduce the cost of equipment and improve the independence, a kind of ship relative ranging and positioning system based on binocular vision technology is proposed in this paper. The proposed system establishes the relative position model by using computer vision technology, separates and recognizes the target from the images collected by two cameras, obtains the vision disparity by image matching, calculates the relative distance and position of the target based on the imaging models of the cameras and the ranging model of binocular vision, and displays the relative distance and position information in the real time. The system can reduce the dependence on navigation system and realize high automation of ranging and positioning.
In this paper, we begin with introducing the hardware architecture and functional modules of the proposed system in the second section. Then, a detailed description of the key technologies of the proposed system is given in the third section, including ranging and position technology, camera calibration technology, image recognition and segmentation technology and image matching technology. Finally, the basic workflow of the proposed system is described in the fourth section.

2. System Architecture

2.1. Hardware Architecture
The hardware of the proposed system mainly includes two small gray image cameras, one fixed bracket, two attitude sensors and one computer. Figure 1 demonstrates the hardware architecture of the proposed system.

![Figure 1. Hardware architecture of the proposed system.](image)

The proposed system adopts gray image cameras in order to save the process of graying the original images. The two cameras are installed on a set of fixed bracket in strict accordance with the designed size. The function of the cameras is to collect the gray image sequences of the target in the sea area at the same time. Figure 2 shows the gray image cameras and the image they capture.

![Figure 2. Gray image cameras and the image they capture.](image)

The sailing attitude of the ship will affect the image acquisition and target distance calculation, and ultimately affect the accuracy of relative ranging and positioning. The pitch and rolling of the ship will lead to the pitching and rotation of the camera, and the horizontal height of the camera will be changed due to the change of the ship’s tonnage. In order to acquire the change of the attitudes and horizontal heights of the cameras in the real time, the proposed system sets two sensors: the altitude sensor and
the attitude sensor. The altitude sensor is located at the gunwale to measure the height from the gunwale to the sea surface. The attitude sensor is installed near the camera to output the ship's pitch and rolling angles.

Computer acquires the target segmentation areas of the two image frames in the artificial interactive mode, identifies the interested feature pixel points, matches these feature pixel points in the condition of epipolar constraint, and obtains the vision disparity of the corresponding feature points. Then it calculates the distance of the target based on the binocular vision ranging principle and the physical model of the cameras. Finally, the calculated result is displayed.

2.2. Function Module

The proposed system has seven modules: video information acquisition module, ship attitude acquisition module, camera parameter calibration module, target recognition and segmentation module, image matching module, ranging and positioning module, and information display module. Figure 3 shows the function module architecture of the proposed system.

![Function module architecture of the proposed system](image_url)

The video information acquisition module, with the function of acquiring the gray image sequences of the target synchronously, provides the basic data for the realization of the system.

The ship attitude acquisition module, with the function of acquiring the pitch and rolling information of the ship and the height information of the cameras in the real time, amends the image incline caused by ship shaking to ensure the calculation accuracy of the proposed system.

The camera parameter calibration module, with the function of calibrating camera parameters, obtains the interior and exterior parameters of the cameras and corrects the distortion parameters, so as to establish the physical models of the cameras.

The target recognition and segmentation module, with the function of recognizing and segmenting the target from the sea area background in the condition of artificial interaction, realizes target recognition through extracting the interested targets from the video images, reduces the image matching range and improves the matching speed.

The image matching module, with the function of matching the corresponding image frames in two groups of video sequences, realizes pixel matching rapidly in the condition of epipolar constraint, and obtains the vision disparity of the corresponding feature pixel points for distance calculation.

The ranging and positioning module, with the function of distance calculation based on the vision disparity of the corresponding feature pixel points, the physical models of the cameras and the binocular vision ranging model, determining the target distance.

The information display module has the function of displaying the video images collected by the two cameras and the distance and location information of the target in the real time.
3. Key Technologies

3.1. Ranging and Positioning Technology

The core technology of ranging and positioning is the binocular vision ranging technology which is to extract scenery information and estimate distance of the target by simulating human eyes [1, 2]. As shown in figure 4, when two cameras \( O_L \) and \( O_R \) installed in parallel take pictures of the same target (point \( P \)), the target images have vision disparity due to their different spatial positions [3-5]. The vision disparity of the two images is:

\[
d = \mu_L - \mu_R
\]

In this equation, \( \mu_L \) and \( \mu_R \) are the pixel abscissas of the points \( P_L \) and \( P_R \), which are the imaging points of the target point \( P \) in the image coordinate systems of the two cameras at the same time.

According to the geometric model, the relationship between \( Z \), i.e., the distance from the target point \( P \) to the imaging plane, and \( d \), i.e. the vision disparity, is as follows:

\[
Z = \frac{fb}{d}
\]

In this equation, \( f \) is the focal length of the camera, and \( b \) is distance between optical axis of the two cameras.

Obviously, in order to obtain the distance \( Z \) of the target, we need to obtain the focal length \( f \) of the camera, the distance \( b \) between optical axis of the two cameras, and the vision disparity \( d \) of the corresponding pixels firstly. Among them, the focal length \( f \) and the distance \( b \) between optical axis can be obtained through static camera parameter calibration, while the vision disparity \( d \) can only be obtained after the feature points in the two images are matched at the pixel level.

![Figure 4. Basic principle of binocular vision ranging.](image)

3.2. Camera Parameter Calibration Technology

The accuracy of the cameras’ internal and external parameters is related to the accuracy of the scene imaging model and the ranging and positioning model, which is an important factor that affects the accuracy of ranging and positioning. Therefore, it is necessary to calibrate the cameras’ imaging parameters and imaging distortion parameters accurately [6].

Camera parameter calibration is the premise and basic work of the proposed system. The purpose is to establish the mapping parameters from the 3D objective world to the 2D imaging plane of the camera, so as to obtain the internal and external parameters of the camera. As the proposed system adopts linear imaging cameras, its visual angle can meet the requirements of ranging and positioning, and the imaging distortion is small, so the proposed system adopts Zhang Zhengyou’s camera calibration method to calibrate camera parameters with the help of calibration checkerboard.

In the process of calibration, 30 images of the checkerboard are captured from different angles continuously by using the single frame image acquisition program firstly. Then the images are batch processed by the camera calibration program of OpenCV. Finally, the internal and external parameters
of the camera are given according to the calibration algorithm. The flow chart of camera calibration is shown in figure 5.

![Flow chart of camera calibration.](image)

The parameter matrix and distortion parameters of the camera are calculated by the least-square theory through this process. The camera matrix is composed of the optimal internal parameters of the camera, which is used for matching the image feature points. The distortion parameters are used for removing distortion of the images.

### 3.3. Target Recognition and Segmentation Technology

The purpose of target recognition and segmentation is to separate the targets accurately from the sea scene images captured by the cameras, so as to match images rapidly. It’s the premise condition of target relative position calculation. The key technologies are image edge detection and image segmentation. At present, a lot of methods for digital image segmentation are proposed. According to the actual characteristics of the sea scene, targets at sea are generally sparse, and imaging features of the target and the sea scene are quite different, so we divide the images into the sea scene texture areas and target pixel areas, realize target segmentation from the images of the sea scene by adopting a combination of texture recognition method and target edge detection method (Canny operator). Considering the complexity of the sea scene, the proposed system adopts the combination of automatic segmentation and manual assistance, as well as the combination of static segmentation and dynamic segmentation to identify the target.

Experiments demonstrate that the edge detection algorithm based on Canny operator can basically satisfy the requirements of the proposed system. As shown in figure 6, the left is the original image of the sea scene, the right is the boundary image after the operation of this algorithm. It can be seen that the image on the right has a clear boundary, and the target can be accurately separated from the sea background.

![Original image and boundary image.](image)

### 3.4. Image Matching Technology

The key to system implementation is to obtain the vision disparity \(d\) of the corresponding feature pixel points of the same image frames in two groups of video sequences. The acquisition of vision disparity requires the matching of the corresponding feature pixel points firstly. The accuracy of the image matching determines the accuracy of ranging and positioning directly.

At present, a lot of image feature matching methods are proposed, although image feature matching algorithms are different and the matching process also have certain differences, but the main steps of these image feature matching methods are much the same, which can be roughly divided into 5 steps: input image, image preprocess, extract useful information for matching, image match, and output results [7, 8]. Figure 7 is the general flow chart of image matching.
Figure 7. General flow chart of image matching

Because the original images are relatively large, the amount of calculation of image matching is large too, which affects the operating speed of the proposed system and is not conducive to practical applications. Therefore, this system adopts the method of combining epipolar constraint search method with SURF algorithm, so as to reduce the matching region and improve the matching efficiency.

As shown in figure 4, it is very time-consuming to match the corresponding feature points $P_L$ and $P_R$ on the two-dimensional image space. Points $P_L$ and $P_R$ are located at the epipolar lines where two images and the epipolar plane formed by the camera $O_L$, the camera $O_R$ and the target point $P$ intersect respectively. For feature points $P_L$ in the left image, the matching point $P_R$ in the right image will be located at the right epipolar line. Therefore, feature matching is reduced from two-dimensional search to one-dimensional search by using the epipolar constraint method. It reduces the matching search region, improves the search efficiency, reduces the error matching, and can satisfy real-time calculation requirements of video frame sequence completely.

The epipolar constraint method realizes the dimensionality reduction search and reduces the search region, but it can’t match the pixel points precisely. In order to match the specific corresponding points, the extraction and matching of feature points must be carried out. At present, the feature point extraction method includes Harris algorithm, FAST algorithm and SURF algorithm, etc. Among them, SURF algorithm has the advantages of scale invariant characteristic, strong robustness and fast computing speed. It is more suitable for the extraction of image feature points collected in complex sea environment. Therefore, the proposed system adopts SURF algorithm provided by OpenCV to conduct the feature point match. It realizes pixel-level matching of the video images captured by two cameras, and lays a foundation for vision disparity calculation and system realization.

4. System Workflow
As shown in figure 8, the entire workflow of the proposed system has the following seven steps:

- Calibrate camera parameters to obtain the interior and exterior parameters of the cameras;
- Install and debug system equipment, and set the parameters of the cameras by using flycapture software tool;
- Collect the video image information of the target, and send the image sequences and attitude information to the computer;
- Process images in the manual interaction mode, including target recognition and segmentation, image feature matching;
- Calculate the vision disparity of the corresponding pixel points according to the imaging models of the cameras.
- Calculate the distance of the target according to the binocular vision ranging model;
- Display the images and the distance information of the target.
5. Conclusion
In this paper, a ship ranging and positioning system based on binocular vision technology is proposed. It will help to solve the practical problem of ship relative positioning at sea, promote the research of engineering application of computer vision technology in ship navigation effectively and improve the autonomy and automaticity of positioning system. The proposed system has the advantages of low cost, simple structure and convenient operation. Experiments demonstrate that the proposed system has good ranging accuracy and can meet the needs of ship navigation and positioning. It will have a broad application prospect in both military and civil navigational fields.

Acknowledgments
This work is supported by the National Natural Science Foundation of China (Grant No. 61471378).

References
[1] Hayashi R, Maeda T, Shimojo S and Tachi S 2014 An integrative model of binocular vision: A stereo model utilizing interocularly unpaired points produces both depth and binocular rivalry Vision Research (44) 2367-2380.
[2] Long H M, Guo H Y, Liang F and Liu G H 2014 Distance measurement algorithm based on binocular stereo vision Applied Mechanics & Materials (3485) 948-952.
[3] Zhu Y, Xu Y C, Wang X and Xu X X 2012 Research on traffic signs distance measurement algorithm based on binocular stereo vision 2012 IEEE 2nd International Conference on Cloud Computing and Intelligence Systems pp 1719-1722.
[4] Li Q M and Li Y Y 2017 Binocular stereo distance measurement based on improved SIFT algorithm Transducer and Microsystem Technologies (36) 119-122.
[5] Vallerand S, Kanbara M and Yokoya N 2003 Binocular vision-based augmented reality system with an increased registration depth using dynamic correction of feature positions Proceedings Virtual Reality (IEEE Press) pp 271-272.
[6] Tsai R 2003 A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses IEEE Journal on Robotics & Automation 323-344.
[7] Zitova B and Flusser J 2003 Image registration methods: A survey Image Visio. Comp. (11) 977-1000.
[8] Zhang X and Liu Z 2014 A survey on stereo vision matching Algorithms Intelligent Control and Automation 2026-2031.