

Research Article

Research on Active Intelligent Perception Technology of Vessel Situation Based on Multisensor Fusion

Ruixin Ma ^{1,2}, Yong Yin ¹, Zilong Li ², Jing Chen,² and Kexin Bao²

¹Key Laboratory of Marine Simulation and Control, Dalian Maritime University, Dalian 116026, China

²Tianjin Research Institute for Water Transport Engineering, M.O.T, Tianjin 300456, China

Correspondence should be addressed to Yong Yin; bushyin@163.com

Received 31 May 2020; Accepted 3 August 2020; Published 26 August 2020

Academic Editor: Jian G. Zhou

Copyright © 2020 Ruixin Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we focus on the safety supervision of inland vessels. This paper especially aims at studying the vessel target detection and dynamic tracking algorithm based on computer vision and the target fusion algorithm based on multisensor. For the vessel video target detection and tracking, this paper analyzes the current widely used methods and theories. Additionally, facing the application scenarios and characteristics of inland vessels, a comprehensive vessel video target detection algorithm is proposed in this paper. It is combined with a three-frame difference method based on Canny edge detection and a background subtraction method based on mixed Gaussian background modeling. Besides, for the multisensor target fusion, the processing method of laser point cloud data and automatic identification system (AIS) data is analyzed in this paper. Based on the idea of fuzzy mathematics, this paper proposes a method for calculating the fuzzy correlation matrix with normal membership function, which realizes the fusion of vessel track features of laser point cloud data and AIS data under dynamic video correction. Finally, through this method, a set of vessel situation active intelligent perception systems based on multisensor fusion was developed. Experiments show that this method has better environmental applicability and detection accuracy than traditional manual detection and any single monitoring method.

1. Introduction

With the rapid development of inland shipping, the number of water transportation vessels is increasing, and the pressure on water transportation is rising. Relying solely on manual labor for on-site or video inspections has low efficiency, and intelligent supervision methods are still insufficient. In the field of intelligent vessel supervision, the United States applies differential GPS navigation technology and radar systems to the Mississippi River Basin; the European Union launched the “Rhine Navigation Plan” to dynamically supervise ship information in the Rhine Basin; China has constructed the “Yangtze River Digital Waterway” to realize the informatization of the Yangtze River waterway and the intelligent management of vessels with the help of BeiDou Navigation Satellite System, GIS, AIS, and so on. However, there is no perfect supervision system to effectively supervise and warn against illegal acts such as vessel

overload, overrun, obscure vessel name, and failure to open AIS according to regulations currently.

For the intelligent supervision of vessels, video, AIS, RFID, radar, and other technologies are mainly used currently. At present, the traditional video surveillance in water transport is already common, but how to extract the information of interest from the video is still cumbersome. Due to the nonstandard vessel name identification and unclear characters, there are few intelligent analysis applications based on computer vision. AIS is also widely used, but it must rely on the installation and correct use of shipborne terminals. Besides, radar is widely used in coastal ports, but in inland rivers there are fewer applications due to mountain forest blocking and shoreline blocking.

In addition, any single supervision method has corresponding drawbacks, and at the same time, it is impossible to effectively carry out data collaboration. There is no unified and efficient fusion method to discover accurate and

valuable information from multisource disordered supervision data. How to accurately and effectively extract ship targets from existing videos to achieve automatic object detection and tracking of vessel video targets and cooperate with multisensor multisource data fusion analysis to improve the level of supervision precision and realize the intuitive expression of vessel supervision is currently a problem that needs to be solved urgently.

2. Research Overview

In the field of vessel video target detection and tracking, Liu of Zhejiang University used image analysis, deep learning, and other computer vision technologies to detect and recognize vessel name identification characters and proposed a reference data set for vessel name identification characters [1]; Fefilatyev et al. from University of South Florida presented an algorithm that detects and tracks marine vessels in video taken by a nonstationary camera installed on an untethered buoy [2]; Cui of Dalian Maritime University in China used multisources remote sensing images as data sources and proposed rapid and effective detection methods towards different typical marine search targets and then used high resolution images to detect oil slick targets [3]. In addition, some scholars used deep learning methods such as YOLOv3 and CNN to achieve vessel target detection in surveillance video [4, 5].

In the field of multisource sensor data fusion, Lei studied the application of cameras and laser rangefinders in ship locks and studied the motion detection system of vessel in ship locks based on the combination of multiple sensors [6]; Xue used two types of laser sensors to measure the length, height, width, speed, and other information of vessel and then extracted the vessel target through data association and data fusion methods [7]; Achiri et al. presented a novel approach which using a constant false alarm rate (CFAR) algorithm to fuse synthetic aperture radar (SAR) images and AIS data for maritime surveillance [8].

According to the current research progress, there are two main types of vessel target automatic detection methods based on computer vision, which are traditional algorithms and deep learning algorithms [2, 9, 10]. The vessel target detection algorithm based on deep learning requires a large number of samples which are difficult to obtain in the mountain river environment, so it is limited to some application scenarios, while traditional target detection methods mainly include interframe difference method, optical flow method and background subtraction method, and so on. In terms of multisensor data fusion, the current inland river vessel track fusion is mainly the fusion of AIS and radar data, mostly through the Kalman filter denoising method, the modified K-nearest neighbor method, and so on [11]. The core of these algorithms is to determine whether two trajectories from multiple systems represent the same vessel target. Although the above methods have achieved certain results, they still have the following limitations, and it is difficult for a single sensor to respond to complex environmental perception information [12]. The laser sensor has the advantages of high detection accuracy and is not affected

by weather conditions, but it cannot obtain information such as color and texture and can only reflect the characteristics of the target's outline and other characteristics [13–15]. Video makes up for the disadvantages of laser in identifying objects and can capture a lot of details of the target and the environment, but it is greatly affected by the weather and the recognition is not accurate enough [16]. AIS contains structured and standardized data, but its acquisition method depends on the use of shipboard terminals and cannot be actively supervised [17].

In view of the previously mentioned problems in vessel video target detection, this paper proposes a vessel video target detection algorithm combining a three-frame difference method based on Canny edge detection and a background subtraction method based on mixed Gaussian background modeling, which effectively improve the success rate of video target detection. For multisensor target fusion, this paper analyzes the processing method of laser point cloud data and AIS data, and based on the idea of fuzzy mathematics, this paper proposes a method for calculating the fuzzy correlation matrix with normal membership function, which realizes the fusion of vessel track features of laser point cloud data and AIS data under dynamic video correction.

3. Vessel Video Target Detection and Dynamic Tracking

A vessel video target detection algorithm combined with a three-frame difference method based on Canny edge detection and a background subtraction method based on mixed Gaussian background modeling are proposed. Then, the CV model is used to realize the motion tracking of the vessel target through the Kalman filter algorithm.

3.1. Realization Process of Vessel Video Target Detection.

First, this method preprocesses the video image sequence through histogram equalization and median filtering and processes the uneven brightness or contrast of the image to improve the accuracy of vessel target extraction. Then, calculate separately the difference image between two adjacent frames between three consecutive frames and use the OTSU method to obtain the threshold of the difference image. After that, binarize the two difference images separately and finally perform a logical AND operation on the two binary images to obtain the common target part of the two binary images. Thus, the outline information of the vessel moving target is achieved.

Second, the edge information of the N th frame image is obtained by performing Canny edge detection on the image, The use of Canny edge detection is to be able to identify the more actual and complete edge of the vessel target in the image as accurately as possible. After obtaining the edge of the image, perform logical AND operation with the contour information obtained by the three-frame difference method to obtain the foreground image.

Finally, perform background subtraction operation on the N th frame image to obtain the foreground image of the

current frame and perform the binary process to obtain the binary image and then perform the logical OR operation on the foreground image calculated by the above logical “AND” and the binary image to obtain complete image information of the edge information and the vessel target information. The specific implementation process is shown in Figure 1.

3.2. Mixed Gaussian Background Modeling. The key of the background subtraction method is to obtain accurate background images. In order to overcome the complex scenes with large changes in the external environment, this paper uses a mixed Gaussian background modeling method to obtain background images. The mixed Gaussian background modeling method uses multiple single Gaussian probability density functions, and the density distribution function of a certain pixel value is approximately represented by the weighted average of all density functions. Let $I(x, y, t)$ denote the pixel value of the pixel (x, y, t) at time t :

$$P(I(x, y, t)) = \sum_{i=1}^K \omega_i^t \eta_i(x, \mu_i^t, \sigma_i^t). \quad (1)$$

In the formula, ω_i^t is the weighting coefficient of the Gaussian component i at time t , and K Gaussian components of a certain pixel are arranged according to the value of (ω/σ) from large to small. The image background model is the first B Gaussian distributions that satisfy

$$B = \arg \min_b \left\{ \sum_{k=1}^b \omega_k > T \right\}. \quad (2)$$

The value of T in the formula is set to 0.7 according to the experimental results, which is the minimum proportion of the Gaussian distribution occupied by the background model. According to the background, the larger the value of T is, the more complex dynamic background the model can be described. For the current pixel (x, y, t) , if its value $I(x, y, t)$ matches the k th ($k \leq B$) Gaussian distribution in its background model, then the pixel is determined as the image background pixel; otherwise, it is determined as the target foreground pixel. If the *ImageOutput* is the function of output image, it can be determined by

$$\text{ImageOutput}(x, y, t) = \begin{cases} 0, & \mu_k^t - \lambda * \sigma_k^t \leq I(x, y, t) \leq \mu_k^t + \lambda * \sigma_k^t \cap k \leq B, \\ 1, & \text{otherwise.} \end{cases} \quad (3)$$

After the image automatically recognizes the target foreground, if the pixel is regarded as the target foreground, a new Gaussian distribution needs to be updated to replace the Gaussian distribution which possesses the smallest weight. The expected value of this new Gaussian distribution is represented by the current pixel value. If the pixel is regarded as the image background, the weight needs to be updated, and this weight value is the weight value of each Gaussian distribution of the background pixels of the image. The Gaussian distribution, which is based on target pixel matching, needs to update both the expected value and the deviation value in the Gaussian model simultaneously.

3.3. Vessel Target Dynamic Tracking. After identifying the moving vessel target through the image, we need to track it, and determining the target’s motion model is the key to target tracking. Most of the vessels keep moving at a constant speed through the supervision area, and the movement state will not change in a short time. Therefore, this paper establishes a CV model for the movement of vessels. Suppose the state of the vessel at time t is X_t , which can be expressed as

$$X_t = \begin{bmatrix} P_t \\ V_t \end{bmatrix}. \quad (4)$$

In the formula, V_t is the speed of the vessel at time t and P_t is the position of the vessel at time t . According to the CV model, the conversion relationship between the vessel’s position and speed is

$$\begin{bmatrix} P_t \\ V_t \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} P_{t-1} \\ V_{t-1} \end{bmatrix} + \begin{bmatrix} \frac{\Delta t^2}{2} \\ \Delta t \end{bmatrix} w_{t-1}. \quad (5)$$

Equation (5) can be written as follows:

$$\hat{x}_t = F_t \hat{x}_{t-1} + w_{t-1}. \quad (6)$$

In the formula, F represents the transition matrix of the ship from the previous state to the current state, w_{t-1} is the added process noise, and \hat{x} indicates that the current state is the estimated motion state.

Kalman filter algorithm is used to track the ship’s running track based on the CV model, and the real-time motion state of the vessel is obtained. This paper uses the above formula to predict the current vessel’s motion state. Because the prediction of the current vessel motion state contains noise, the covariance matrix P is needed to represent the noise in the prediction.

$$P_t = F \text{cov}(X_{t-1}, X_{t-1}) F^T + Q_{t-1} = F P_{t-1} F^T + Q_{t-1}. \quad (7)$$

In the formula, Q is the covariance of process noise. The specific process is shown in Figure 2.

$$\begin{aligned} K_t &= P_t^- H^T (H P_t^- + R)^{-1}, \\ P_t &= (I - K_t H) P_t^-, \\ \hat{x}_t &= \hat{x}_t^- + K_t (Z_t - H \hat{x}_t^-), \\ Z_t &= H X_t + v_t, \\ H &= (1, 0), \\ \hat{x}_t &= F_t \hat{x}_{t-1} + w_{t-1}, \\ P_t &= F P_{t-1} F^T + Q_{t-1}, \end{aligned} \quad (8)$$

In Figure 2, Z_t is the measured value of the vessel’s movement state at the time t , and v_t is the measured noise. K_t is used to represent the relationship between the measured value and the predicted value of the vessel state, and R is the variance of the measured noise; at this time, it is still necessary to update the measured value \hat{x}_t

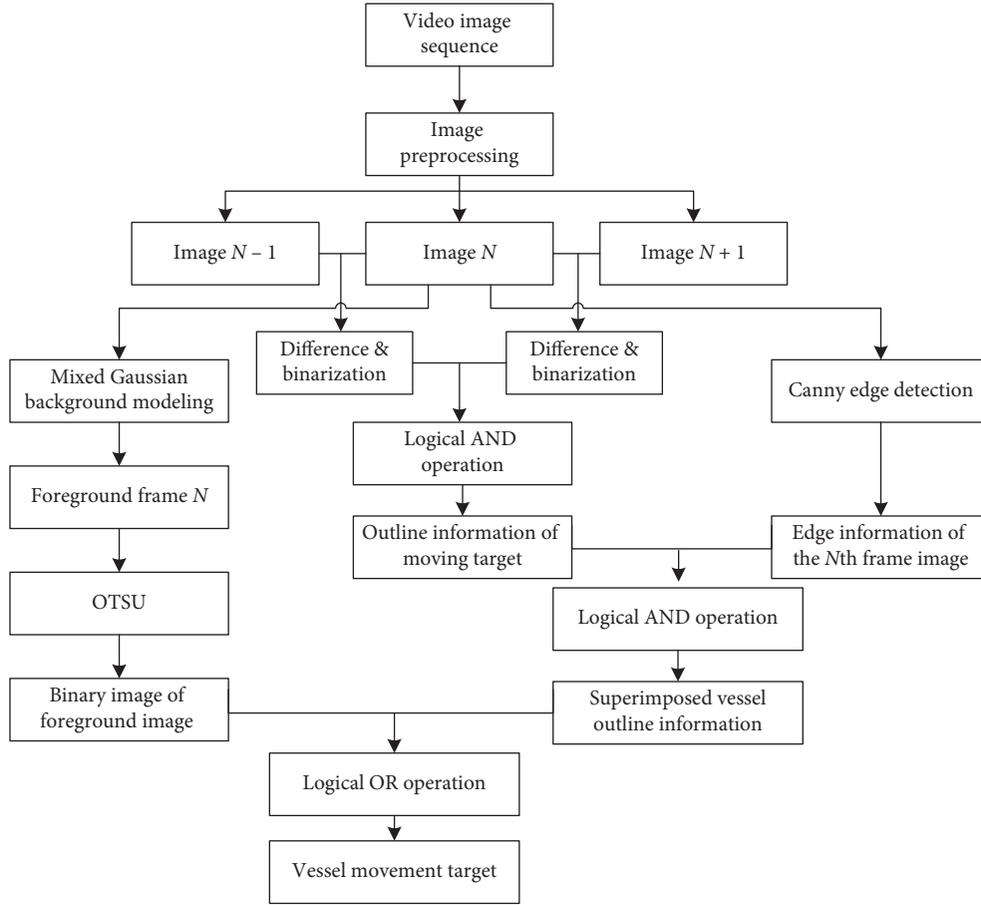


FIGURE 1: Video image processing.

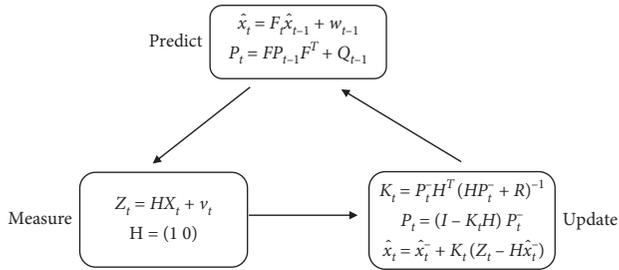


FIGURE 2: Kalman filter tracking.

and the state covariance P of the vessel's motion state. This algorithm uses the Kalman filter method to complete the prediction and correction of the vessel's moving track in the process of vessel state prediction, measurement, and update and finally reaches the state of convergence.

4. Multisensor Target Fusion Algorithm

4.1. Fusion Process Analysis. The individual sensors (laser, video, and AIS) used in this method can independently acquire and form their own vessel target motion models. After comparing and matching multiple independent vessel motion models, this paper determines the unified

vessel motion model through multisensor fusion calculation. In this method, laser point cloud data and AIS data are first used for fusion calculation of vessel's track, which is the key to fusion calculation. After the vessel's moving target is detected by the video, the vessel's track after the fusion of the laser point cloud data and AIS data can be corrected and displayed comprehensively on the dynamic video, so that the vessel's target has rich real-time environmental characteristics and the better-ranging accuracy and higher recognition rate overall. The specific process is shown in Figure 3. Multisensor integration and information fusion effectively solve the fuzzy point of a single sensor, making the fused data more advantageous in terms of redundancy, complementarity, and accuracy.

4.2. Laser Point Cloud Data of Vessel Processing. In this paper, the laser scanner is used to scan the vessel to achieve noncontact measurement of the vessel's outline and freeboard based on its advantages of high measurement accuracy, strong smoke penetration ability, no influence of light, and so on.

The laser scanner obtains the distance of the vessel and the water surface through the time difference between the continuous emission and reflection of the laser. And then, a plurality of laser curtains in the direction of the cut surface

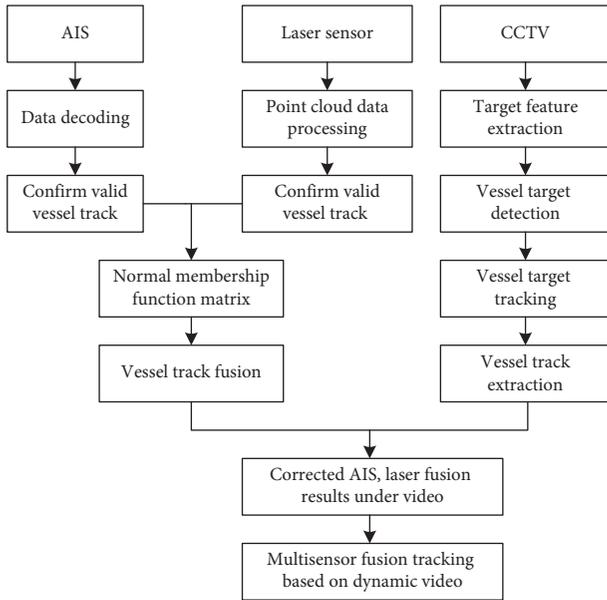


FIGURE 3: Process of multisensor target fusion.

of the waterway is formed by the rotation of the laser beam. When the vessel goes through the laser curtain, the laser scanner continuously scans each section of the vessel to collect the laser reflection data of each section of the vessel. The laser scanner will obtain the tangential multi-section profile data of the vessel, and each section is composed of countless laser reflection data. As the vessel sails through the supervision area, the laser scanner can collect the whole laser reflection data of several sections. Therefore, the complete outline data of the vessel can be obtained. The process of the vessel passing the laser curtain is shown in Figure 4.

The amount of point cloud data obtained by direct scanning is very large. In this paper, a simplified method of point cloud based on octree is used to reduce the amount of point cloud data while maintaining the geometric characteristics of vessel point cloud data. At the same time, the scene point cloud automatic cropping algorithm based on cluster analysis is used to realize the three-dimensional reconstruction of the water surface and the vessel [18]. Moreover, through the adaptive filtering algorithm to filter out the shake of the water surface and hull, the vessel's three-dimensional size and other parameters, including the ship's driving direction, speed, freeboard, and other characteristics, can be obtained [19]. The actual scene of the vessel and the simplified point cloud diagram are shown in Figure 5.

In order to reduce the blind zone interference, this method requires two laser scanners to scan together; that is, a fixed-angle laser scanner is installed on each side of the channel to achieve blind zone complementarity, as shown in Figure 6. For the point cloud data obtained by different scanners, this paper uses the iterative closet point (ICP) algorithm to match the point cloud data to a unified coordinate system through continuous rotation and translation misalignment.

4.3. AIS Data of Vessel Processing. The AIS system is a vessel automatic identification system that can send and receive vessel's dynamic and static information within the coverage area for vessel target identification and information exchange. The AIS base station realizes data collection of vessel through the receiver, serial port, and network switching device, and its specific connection diagram is shown in Figure 7.

The AIS base station can obtain relatively complete and standardized structured data, mainly including the following types:

- (1) Static information: IMO number, call sign and vessel name, length and moulded breadth, and vessel type
- (2) Dynamic information: position of the vessel, UTC time, course over ground, speed over ground, heading, navigation state, and steering rate
- (3) Voyage-related information: vessel draft, type of dangerous goods, destination port and estimated time of arrival, and route plan
- (4) Security-related information: broadcast and notification information

The information of AIS is divided into plain code and secret code. The plain code starts with the "\$" symbol, which can directly parse the meaning it represents. Although the plain code is easy to read, it takes up more bandwidth resources. Therefore, the IEC has clear character restrictions on the plain code, and at the same time introduced the secret code for the data encapsulation. The secret code is an encapsulated information packet, which starts with "!", and its format is

$$!aacc, x, y, z, u, c-c, v * hh<CR><LF>$$

where aacc is an identifier, indicating the background information encapsulated by this sentence; x represents the number of expression sentences required to transmit the information (up to 9); y represents the number of sentences in this sentence (1–9); z represents unified identification of the same sequence (0–9 cycle); u indicates the corresponding channel (A/B) when receiving the information; $c-c$ is the encapsulated information, which needs to be mapped with 6 bit ASCII code; v indicates the filled characters; and hh indicates the check characters.

For example, the following two sentences represent a piece of static information related to vessel navigation:

!AIVDM, 2, 1, 1, A, 544RLM01oOMEDA5L001<4A8T@@Tr05TpT0000016<0N=32no0=i0C2@C, 0 * 0E!

!AIVDM, 2, 2, 1, A, P00000000000000, 2 * 45.

$x = 2$ means that two sentences convey a message together, and the $c-c$ package information is

"544RLM01oOMEDA5L001<4A8T@@Tr05TpT0000016<0N=32no0=i0C2@C"+"P00000000000000".

At first, the encapsulated information is converted into 6 bit codes, and then all codes are integrated into corresponding information according to the message format of the corresponding ID. The following uses a received sentence as an example to explain the decoding process and decoding method of AIS information. The content of the received information is as follows:

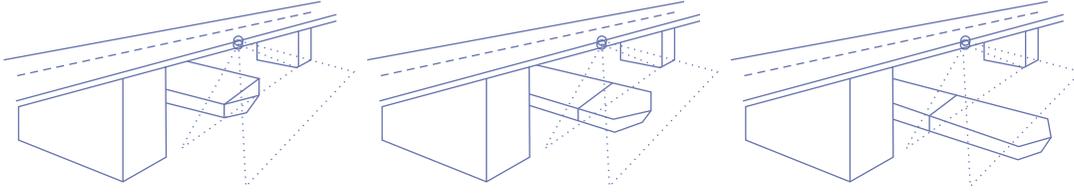


FIGURE 4: Vessel goes through the laser curtain.

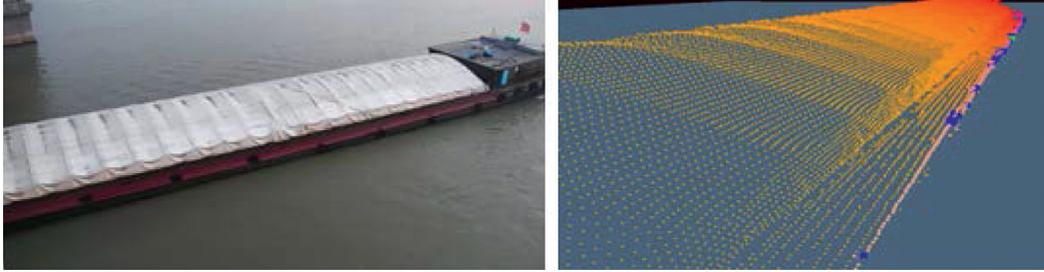


FIGURE 5: Schematic diagram of the vessel's physical object and laser point cloud.

```
!AIVDM,1,1,A,15      Cgah00008      LOnt>1
Cfs6NT00SU,0 * 3D
```

_____Decoding information start

Decoding process: at first, take the ASCII code of the received frame of characters and then check whether the received frame of characters is correct through the check characters. If the received characters are correct, the ASCII codes are converted to 6-bit ASCII codes, and then the various contents are decoded. In this frame information, the information starts with the character "1," each character is 6-bit ASCII code, the total length of the entire frame information is 168 bit, and the information content and its corresponding bit position refer to ITU-RM.1371-1 technical standard. AIS information decoding is a process of data decompression and information extraction.

4.4. The Method of Vessel Track Fusion. The method of multisensor fusion proposed in this paper mainly considers the vessel track fusion of AIS and laser point cloud. AIS data and laser point cloud data must be unified into the same time system. First, the dynamic data of the supervised ship is extracted independently, then the position coordinate transformation and time correction are performed, and the track correlation calculation is performed on the unified fusion time node. Finally, the track fusion is achieved based on the weight of the associated target.

For the unified time system, the scanning period of the laser sensor is fixed and linear interpolation can be performed. The time interval of AIS sending dynamic information varies with the vessel's navigation state, and the cubic spline interpolation can be used to calculate the time calibration of AIS. The target vector $X(t)$ of AIS at time t can be calculated as follows:

$$X(t) = A_0 + A_1(t - t_0) + A_2(t - t_0)^2 + A_3(t - t_0)^3. \quad (9)$$

In the formula, A_0, A_1, A_2, A_3 are transformation coefficients, which can be calculated by substituting four target vectors by changing AIS time to GMT time.

The key of vessel track fusion is to judge whether the two track lines from AIS and laser point cloud data represent the same vessel. This method uses fuzzy mathematics membership function to represent the similarity of its track.

$$\xi(\zeta_p) = \exp\left[-\eta_p \left(\frac{\zeta_p^2}{\sigma_p^2}\right)\right]. \quad (10)$$

In the formula, $\xi(\zeta_p)$ is the membership function of the p th factor in the fuzzy factor set. $\eta_p, \zeta_p,$ and σ_p are the weight, Euclidean distance, and spread of the p th factor in the fuzzy factor set. In this paper, the normal membership function is used to calculate the target track fusion. As shown in (10), the larger the membership function value, the greater the correlation between the two tracks.

The specific calculation process of this method is as follows:

- (1) Determine the fuzzy factor set. According to actual experience, the factors affecting the track mainly include vessel position, speed, and course. Therefore, the fuzzy factor set is established as $F(\zeta_1, \zeta_2, \zeta_3)$, where ζ_1 is the position factor, ζ_2 is the speed factor, and ζ_3 is the course factor.
- (2) Determine the weight of the fuzzy factor set. In this paper, the weight of the three fuzzy factors of η is taken as $\eta_1 = 0.55, \eta_2 = 0.35, \eta_3 = 0.10$ through the Delphi method. It can be seen that the weight of the ship position is the largest, the speed is the second, and the course is the least.
- (3) Calculate the Euclidean distance of fuzzy factors.

The Euclidean distance of the ship position is

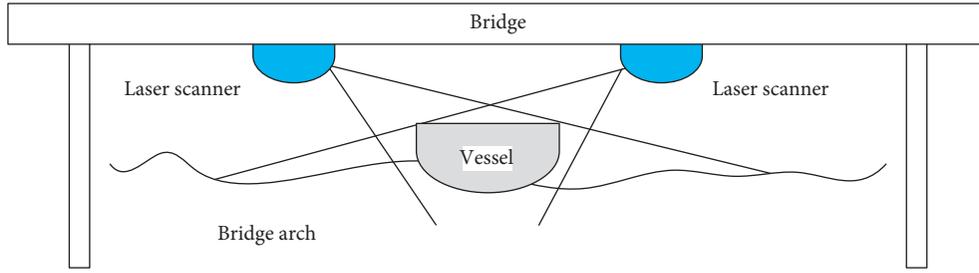


FIGURE 6: Scanning diagram of laser scanner.

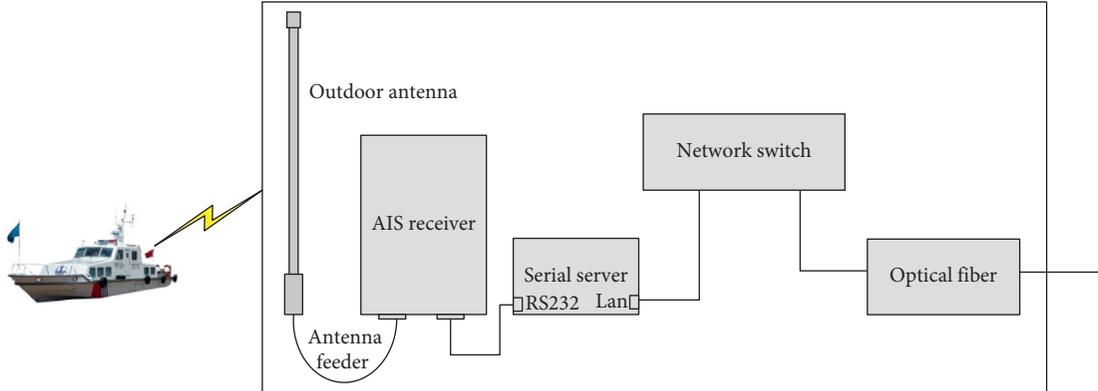


FIGURE 7: Schematic diagram of AIS base station installation.

$$\zeta_1(t) = \sqrt{[x_a(t) - x_l(t)]^2 + [y_a(t) - y_l(t)]^2}. \quad (11)$$

In the formula, $\zeta_1(t)$ is the Euclidean distance of the vessel position; $x_a(t)$ and $y_a(t)$ are the vessel position coordinates obtained by AIS at time t , and $x_l(t)$ and $y_l(t)$ are the vessel position coordinates obtained by laser.

The Euclidean distance of the speed is

$$\zeta_2(t) = |v_a(t) - v_l(t)|. \quad (12)$$

In the formula, $\zeta_2(t)$ is the Euclidean distance of the speed; $v_a(t)$ and $v_l(t)$ are the target speeds obtained by AIS and laser at time t .

The Euclidean distance of course is

$$\zeta_3(t) = |\varphi_a(t) - \varphi_l(t)|. \quad (13)$$

In the formula, $\zeta_3(t)$ is the Euclidean distance of course; $\varphi_a(t)$ and $\varphi_l(t)$ are the target course obtained by AIS and laser at time t .

(4) Calculate the comprehensive similarity. At first, the $\zeta_1(t)$, $\zeta_2(t)$, and $\zeta_3(t)$ are brought into (10) to calculate the membership function value of each fuzzy factor and then perform function weighted calculation according to the weight value of each factor.

$$\lambda_{al}(t) = \sum_{p=1}^3 \eta_p \xi[\zeta_p(t)], \quad a \in m, r \in n. \quad (14)$$

(5) Establish a fuzzy correlation matrix at time t for m tracks from AIS data and n tracks from laser data:

$$\lambda(t) = \begin{bmatrix} \lambda_{11}(t) & \lambda_{12}(t) & \dots & \lambda_{1n}(t) \\ \lambda_{21}(t) & \lambda_{22}(t) & \dots & \lambda_{2n}(t) \\ \dots & \dots & \dots & \dots \\ \lambda_{m1}(t) & \lambda_{m2}(t) & \dots & \lambda_{mn}(t) \end{bmatrix}. \quad (15)$$

In the formula, $\lambda_{mn}(t)$ represents the overall similarity between the m th track in the AIS and the n th track in the laser at time t .

(6) Track similarity check. The specific steps are as follows:

- (i) Determine the size of the threshold ε , usually $\varepsilon \geq 0.5$.
- (ii) The largest element $\lambda_{mn}(t)$ in the matrix $\lambda(t)$ is picked, and if $\lambda_{mn}(t) \geq \varepsilon$, it is determined that the AIS track m is related to the laser track n . Then, remove the row and column elements of $\lambda_{mn}(t)$ from the matrix $\lambda(t)$ to obtain a new $(m-1) \times (n-1)$ -dimensional reduced-order fuzzy matrix $\lambda_1(t)$.
- (iii) Repeat the above process for $\lambda_1(t)$ to obtain $\lambda_2(t) \dots \lambda_k(t)$ until all the elements in $\lambda_k(t)$ are less than ε , and then the track represented by the row and column number of the remaining elements is not relevant at time t .



FIGURE 8: Schematic diagram of the on-site comprehensive collection terminals.

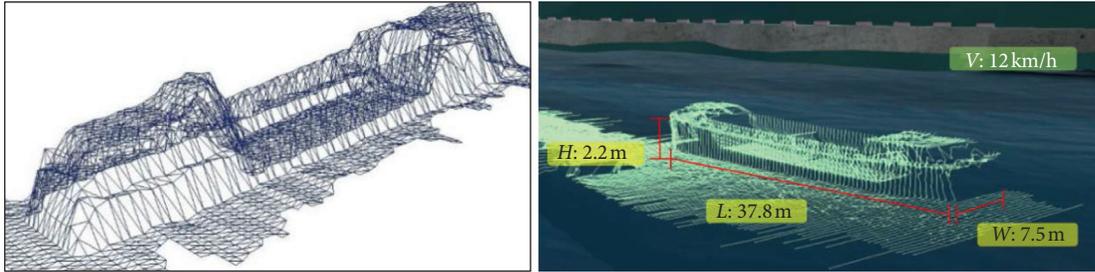


FIGURE 9: Vessel outline scanning and freeboard measurement.

TABLE 1: Test results.

Functions	Total number of vessels	Test results	Correct results	R	P
Vessel target detection	658	623	608	0.92	0.97
Vessel course detection	658	623	618	0.94	0.99
Vessel freeboard detection	658	623	562	0.85	0.90

Note. Because the vessel is affected by wind and waves, the height between the vessel deck and the water surface changes at any time, so in this test, the error between the measured value and the true value of the vessel's freeboard is allowed to be within 2 cm.

- (iv) For the tracks related at time t , if they are related at any time, it can be determined that their information comes from the same target vessel, and then the two track information can be related.

5. Test Verification Analysis

5.1. Vessel Situation Active Intelligent Perception System. For the safety supervision of inland vessels, the method proposed in this paper was tested in an inland waterway in China, and a set of active intelligent perception system for vessel situation was developed based on this method. The system can realize the functions of vessel outline recognition, vessel freeboard measurement, vessel position and speed supervision, vessel traffic flow statistics, and vessel automatic illegal evidence collection for vessels entering and leaving the jurisdiction. It can realize high-precision, all-weather, uninterrupted automatic supervision, and statistical analysis of vessels in inland waterways.

The on-site sensor of the system adopts laser scanner, CCTV, AIS, and other equipment comprehensively. The schematic diagram of on-site comprehensive collection environment is shown in Figure 8.

5.2. Effect Analysis. The system can realize vessel outline scanning and freeboard measurement through point cloud data extracting, as shown in Figure 9.

The system uses dynamic video as the basis for display, combines AR technology, and performs fusion calculations with AIS information and laser point cloud data to provide ship information integration and intelligent information services based on dynamic video.

The system has been deployed and tested for one month, and a total of 623 results have been collected. After on-site video verification, it was found that a total of 658 vessels passed through the supervision area. The recall rate (R) and precision rate (P) are used to evaluate the recognition results in this paper:

$$R = \frac{TP}{TP + FN}, \quad (16)$$

$$P = \frac{TP}{TP + FP}.$$

In the formula, TP represents the number of correct detection results, FN represents the number of missed detections, and FP represents the number of false detections. The test results are shown in Table 1, and key indicators such

as vessel target detection, vessel course, and vessel freeboard can be obtained.

According to the analysis of the existing collected data, the R of vessel target detection is 92%, and the P is 97%; the R of vessel course detection is 94%, and the P is 99%; the R of vessel freeboard detection is 85%, and the P is 90%. The system is more effective than the single use of video, AIS, and laser sensors, which can effectively reduce the labor intensity of frontline workers and improve the safety supervision level of maritime vessels.

6. Conclusions

To sum up, to make up for the deficiencies of existing methods in vessel safety supervision, this paper proposes a method for automatic detection and dynamic tracking of ship video targets and analyzes the processing of laser point cloud data and AIS data. The normalized membership function matrix is used to realize the fusion of vessel track features under dynamic video correction. A set of multi-sensor fusion vessel situation active intelligent perception system was developed based on the method proposed in this paper, which can achieve active all-weather high-precision autonomous supervision without relying on visible signs such as vessel name and number. The system is of great significance for improving the level of maritime management and innovating vessel supervision methods.

Data Availability

The data used to support the findings of this paper are included within the article. Any reader or researcher who wishes to obtain the other related data of this article can contact the author by e-mail.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The paper was funded by the Transportation Science and Technology Fund of Tianjin (no. 22118061) and the Basic Research Fund of Central-Level Nonprofit Scientific Research Institutes (no. TKS20200308).

References

- [1] B. Liu, "Research on ship identification characters detection and recognition based on image analysis and deep learning," Master thesis, Zhejiang University, Hangzhou, China, 2018.
- [2] S. Fefilat'ev, D. Goldgof, and C. Lembke, "Tracking ships from fast moving camera through image registration," in *Proceedings of the 20th International Conference on Pattern Recognition*, pp. 3500–3503, Istanbul, Turkey, August 2010.
- [3] C. Cui, "The application of multi-source remote sensing in maritime target search," Master thesis, Dalian Maritime University, Dalian, China, 2018.
- [4] X. Nie, M. Yang, and R. W. Liu, "Deep neural network-based robust ship detection under different weather conditions," in *Proceedings of the 22nd IEEE International Conference on Intelligent Transportation Systems*, pp. 47–52, IEEE, Auckland, New Zealand, October 2019.
- [5] Z. Shao, Z. L. Wang, Z. Wang, W. Du, and W. Wu, "Saliency-aware convolution neural network for ship detection in surveillance video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 3, pp. 781–794, 2020.
- [6] Y. Lei, "Research on state detection system of ship passing through lock based on multi-sensor technology," Master thesis, Wuhan University of Technology, Wuhan, China, 2012.
- [7] Y. Xue, "Application of multi-sensor data fusion in ship feature extraction," *Ship Science and Technology*, vol. 38, no. 18, pp. 73–75, 2016.
- [8] L. Achiri, R. Guida, and P. Iervolino, "SAR and AIS fusion for maritime surveillance," in *Proceedings of the IEEE 4th International Forum on Research and Technology for Society and Industry (RTSI)*, pp. 1–4, Palermo, Italy, September 2018.
- [9] S. Tan, J. Armstrong Piepmeier, and D. L. Kriebel, "A computer vision system for monitoring vessel motion in conjunction with vessel wake measurements," in *Proceedings of the 2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pp. 1830–1834, Pacific Grove, CA, USA, November 2012.
- [10] H. Li, P. Wang, M. You, and C. Shen, "Reading car license plates using deep neural networks," *Image and Vision Computing*, vol. 72, pp. 14–23, 2018.
- [11] Li Tao, "Multi-source sensor data fusion and its applications in the target detection," Master thesis, University of Electronic Science and Technology of China, Chengdu, China, 2015.
- [12] A. Weintrit, "The need to develop new geodesic-based computational algorithms for marine navigation electronic devices and systems," in *Proceedings of the International Association of Institutes of Navigation World Congress (IAIN)*, pp. 1–12, Prague, Czech Republic, October 2015.
- [13] B. Zhu and L. Xie, "A point cloud registration algorithm using depth image," *Navigation of China*, vol. 42, no. 2, pp. 8–11, 2019.
- [14] J. Zhang, X. Cheng, and X. Cheng, "Application of three-dimensional laser scanning technology in measurement of ship displacement," *Chinese Journal of Lasers*, vol. 43, no. 12, pp. 162–168, 2016.
- [15] Q. Zhu, S. Li, H. Hu et al., "Multiple point clouds data fusion method for 3D city modeling," *Geomatics and Information Science of Wuhan University*, vol. 43, no. 12, pp. 1962–1971, 2018, in Chinese.
- [16] M.-b. Li and S.-p. Zhang, "A new video surveillance method fused with AIS message for ship targets detection," *Electronic Design Engineering*, vol. 20, no. 7, pp. 157–159, 2012.
- [17] A. Milios, K. Bereta, K. Chatzikokolakis, D. Zissis, and S. Matwin, "Automatic fusion of satellite imagery and AIS data for vessel detection," in *Proceedings of the 22th International Conference on Information Fusion (FUSION) 2019*, pp. 1–5, Ottawa, ON, Canada, July 2019.
- [18] X. Zhang and S. Zhang, "Automatic registration of point cloud and image data based on point-to-plane ICP," *Control Engineering of China*, vol. 45, no. 12, pp. 2510–2514, 2017.
- [19] R.-h. Zhang, L. Fu-liang, X. Zhou et al., "A pedestrian detection method under time-space data fusion based on laser and video information," *Journal of Transportation Systems Engineering and Information Technology*, vol. 15, no. 3, pp. 49–55, 2015.