Research Article

Target Recognition Algorithm Based on Optical Sensor Data Fusion

Chunlei Lv [ID] and Lihua Cao

Changchun Institute of Optics, Fine Mechanic and Physics, Chines Academy of Science, Changchun Jilin 130033, China

Correspondence should be addressed to Chunlei Lv; lvchunlei@ciomp.ac.cn

Optical sensor data fusion technology is a research hotspot in the field of information science in recent years, which is widely used in military and civilian fields because of its advantages of high accuracy and low cost, and target recognition is one of the important research directions. Based on the characteristics of small target optical imaging, this paper fully utilizes the frontier theoretical methods in the field of image processing and proposes a small target recognition algorithm process framework based on visible and infrared image data fusion and improves the accuracy as well as stability of target recognition by improving the multisensor information fusion algorithm in the photoelectric meridian tracking system. A practical guide is provided for the solution of the small target recognition problem. To facilitate and quickly verify the multisensor fusion algorithm, a simulation platform for the intelligent vehicle and the experimental environment is built based on Gazebo software, which can realize the sensor data acquisition and the control decision function of the intelligent vehicle. The kinematic model of the intelligent vehicle is firstly described according to the design requirements, and the camera coordinate system, LiDAR coordinate system, and vehicle body coordinate system of the sensors are established. Then, the imaging models of the depth camera and LiDAR, the data acquisition principles of GPS and IMU, and the time synchronization relationship of each sensor are analyzed, and the error calibration and data acquisition experiments of each sensor are completed.

1. Introduction

With the rapid development of modern optoelectronic reconnaissance technology, the image acquisition, transmission efficiency, and imaging accuracy of visible and infrared reconnaissance systems have been greatly improved, and the simultaneous carrying of these two optical reconnaissance systems on a single platform (on water or in the air) has also become a mainstream practice to further improve the effectiveness of reconnaissance platforms in single sortie conditions [1–3]. These optical sensing platforms obtain a large number of digital images and transform them into useful intelligence for the target situation on the battlefield but also need to rely on subsequent image processing methods to detect, segmentation, and tracking of the target [4, 5]. Therefore, the out-of-situ rate of the optical reconnaissance system directly depends on the effectiveness of the image processing methods. In recent years, image processing, as a popular

technology for both military and civilian use, has been developed significantly, and a large number of mature methods for various types of image enhancement, target detection, target segmentation, and other application problems have emerged, greatly promoting the intelligent development process of computer vision. These advanced image processing methods, applied to the field of photoelectric reconnaissance, are sufficient to maximize the efficiency of a single sensor out of the situation [6]. However, in the field of intelligence reconnaissance, including optoelectronic reconnaissance, the problem of multisensor data fusion has always been a major bottleneck, which restricts the further improvement of intelligence reconnaissance effectiveness. Recently, some scholars have made some breakthroughs in the research of data fusion of similar sensors, but the heterogeneous data generated by different types of sensors still cannot be effectively fused. Specifically, in the field of optoelectronic reconnaissance, data fusion of visible reconnaissance images with infrared images

has not yet emerged as a mainstream breakthrough solution. Modern imaging systems mainly include radar (synthetic aperture, phased array, and millimeter wave), visible TV, infrared, and laser imaging means, of which optical sensors, as an important part, rely on the target's thermal radiation work, which is a passive means of detection [7, 8]. Compared with radar systems, optical imaging systems have the advantages of strong anti-interference capability, simple structure, small size, lightweight, and good concealment, but there are also shortcomings such as close detection distance and inability to the range. Initially, the optical imaging system in the field of military reconnaissance is a complementary means of radar systems used to overcome the blind spot of the radar system and platform load capacity and other limitations, in close range target detection, tracking, identification, and other aspects to play a role. With the development of new technologies, new weapons and equipment increasingly focus on the development of radar stealth performance, radar as the main means of early warning detection of reconnaissance intelligence system gradually cannot meet the requirements of combat use, and the photoelectric system gradually becomes an indispensable and important means of reconnaissance, which began in all-weather, high-precision, long-range direction. The expansion of the role of optical reconnaissance equipment distance naturally gives rise to this paper to focus on the problem of small target identification. Small target refers to the imaging system detection moderate distance (about a few hundred to a few thousand meter range) through the sensor acquisition of the image element area of the small target imaging. Small targets are usually only a few tens to hundreds of pixels in visible images and a bright spot or a bright spot in infrared images. If the detection distance is close (e.g., within 100 meters), the target pixel size is large, the outline is clear, and the common means of image processing is easy to achieve detection and identification; if the detection distance is too far (e.g., greater than 10 km), the target pixel size is too small, the outline is not clear, and it is easy to drown in the background clutter and difficult to find. Therefore, the detection and identification of small targets directly affect the scope of the role of optoelectronic reconnaissance equipment; the effectiveness of the intelligence reconnaissance surveillance system is of great significance. In the problem of small target identification, optical sensors have a strong climate adaptation and smoke and dust permeability, can work around the clock, and have other unique advantages; visible TV has a high resolution and access to color information [9, 10].

This paper focuses on the target recognition algorithm of optical sensor data fusion, given full consideration to the advantages and features of the two imaging means and, after an in-depth study, designs a target recognition method framework based on optical sensor data fusion; focuses on the characteristics of the images obtained by the two imaging means; analyzes their advantages in solving the problem of small target recognition; clarifies the general idea of data fusion; and introduces the target detection method using infrared images. The method means of target detection using infrared images, the proposed cyclic clustering method based on visible image target segmentation, and the method framework of fusion processing based on optical sensor data fusion and visible image target segmentation results to achieve comprehensive target recognition are given, which can provide a clear idea for the solution of this bottleneck problem.

2. Related Work

Sensor information fusion technology, also known as sensor data fusion, first appeared at the end of World War II when both optical sensors and radar in an antiaircraft artillery fire control system are utilized. In this system, optical sensors were fully utilized to detect the presence of targets, and radar was used to measure the distance to the targets, which overcame the effects of the harsh battlefield environment and improved the hit rate of the artillery system. However, information fusion at that time was done by manual calculation by technicians, and the processing speed of information was low and the quality of processing was poor, so the information fusion technology was not accepted by people at that time. To end this issue, it was first formally introduced in research institutions and was reflected in sonar processing systems. After extensive experiments, the researchers fused the optical signals that did not interfere with each other to calculate and pinpoint the location. In this incidental use, information fusion technology showed its excellent comprehensive performance, which made information fusion technology gain widespread attention in military applications and rapidly develop into the field of people's livelihood. An example of this application is the Command, Control, Communication, and Intelligence (C3I) system, which pioneered the use of multiple sensors to collect battlefield information and demonstrated the power of information fusion technology. The C3I system is the first to use multiple sensors to collect information on the battlefield, demonstrating the power of information fusion technology, and has received wide attention from countries around the world. The C3I Technical Committee established the Data Fusion Subpanel (DFS) to improve the performance of information fusion and other metrics to overcome the technical challenges in the field of data fusion. Since then, multisensor information fusion technology has been introduced.

In recent years, researches in the field of related technologies have also been ongoing [11, 12]. Muzammal et al. [13] proposed a mathematical model based on a multisensor data fusion algorithm. Bakalos et al. [14] used multimodal data fusion and adaptive deep learning to monitor critical systems. Zhang et al. [15] proposed a method based on multisensor data fusion for UAV safety distance diagnosis. In order to achieve more accurate bearing fault diagnosis, Wang et al. [16] propose a new method to fuse multimode sensor signals collected by an accelerometer and a microphone. Researches on information fusion technology still have some problems, and the development speed is relatively slow. In order to improve the efficiency of target search in large-scale high-resolution remote sensing images, Yin et al. [17] propose an optimized multiscale fusion method for airport detection of large-scale optical remote sensing images.

3. Target Recognition Algorithm Based on Optical Sensor Data Fusion

3.1. Structure of Optical Sensor Data Fusion. Depending on the environment in which the fusion system works, Heistand proposes that there are three fusion processing architectures, namely, centralized fusion, distributed fusion, and hybrid fusion. The centralized fusion structure sends the target information acquired by each sensor directly to the fusion center for processing, and its structure is shown in Figure 1. Although this structure has the advantages of high real-time performance and low loss, it is not easy to implement in practical engineering because of the high communication requirements and large computational effort of the system.

In the distributed fusion structure, the most important feature is that the corresponding local trajectory is firstly obtained based on the separate processing and estimation of each sensor's tracking target state and enters the fusion center, where the data are correlated and filtered according to the local trajectory of each sensor, and finally the fusion estimation of the whole trajectory is completed, also known as sequential fusion, whose structure is shown in Figure 2. Compared with the centralized one, this structure reduces the communication requirements and computational complexity of the system [18]. In addition, it improves the reliability of the multisensor data fusion target recognition system. However, the recognition accuracy is reduced due to the large loss of information.

The hybrid fusion architecture mainly consists of distributed and centralized fusion architecture, whose structure is shown in Figure 3. It inherits the advantages of these two architectures but also retains their shortcomings. In addition, compared with the first two, the hybrid fusion architecture is relatively complex and has increased communication burden and computational complexity, which is not easy to implement in engineering. In practical engineering, the distributed fusion architecture is the highly popular multisensor fusion architecture [18]. Meanwhile, continuous improvement of multisensor fusion methods and algorithms can improve the fusion tracking performance under the distributed fusion architecture.

At present, the data fusion algorithm techniques commonly used for target tracking are mainly divided into four categories based on the model, statistical theory, information theory, and artificial intelligence, while this paper focuses on the study of model-based data fusion algorithms, which mainly establish a motion model for a moving target and use estimation algorithms to fuse the target states obtained from multiple sensors through certain criteria. The commonly used methods are Kalman filter, weighted average method, particle filter, etc. The research method explored in this paper is the particle filter fusion tracking algorithm.

The nonlinear, non-Gaussian state and measurement model of the system can be expressed as

$$R_k = \sum_{i=1}^{k} f_k(w_k), \tag{1}$$

$$L_k = \sum_{i=1}^{k} h_k(v_k), \tag{2}$$

where $w_k \in R^{n_v}$ and $v_k \in R^{n_n}$ are the state noise and the measurement noise, respectively, and both are non-Gaussian noise. Bayes' theorem assumes that the estimated state is a random variable and establishes its prior distribution. Let $z_1, z_2, z_3, \ldots$ be a set of random variables that are uncorrelated with each other but have the same distribution and can be measured. Each variable in this set of random variables maps an unknown parameter x, whose conditional probability density is p_x; then, the posterior probability density of the unknown parameter x is expressed as

$$P_{xz} = \frac{P_{zx}(x) \cdot P_x(z)}{\int P_{zx}(x) \cdot P_x(z) dx}, \tag{3}$$

where p_x denotes the likelihood function of the parameter x data uncorrelated with each other. p_x denotes the probability density of x, also known as the prior probability or prior distribution, which is usually determined based on prior experience before the measurement value is obtained. P_{zx} denotes the density of the posterior distribution of x, also known as the posterior probability or posterior distribution, which is determined after the measurement value is obtained. From equation (2), the Bayesian estimation theory is to pass an unknown parameter x as a random variable and, at the same time, introduce a prior probability p_x. Simply put, Bayes' theorem is to obtain the posterior distribution by updating the prior distribution of the parameter x. Then, Bayesian filtering is divided into two main parts: prediction and update [19–22].

For prediction, the state model of the system for predicting the posterior distribution function from the current moment of measurement to the next moment of measurement is used, i.e.,

$$P(x \mid z) = \int P_{zx}(x) \cdot P_{xz}(z) dx + z \int P_{x_k}(x) \cdot P_x(z) dx. \tag{4}$$

For update, the posterior distribution function is corrected using the most recent quantiles at the current moment, i.e.,

$$P(x_k \mid z_k) = \frac{P_{zx}(x) \cdot P_{xz}(z) + P_{kx}(x) \cdot P_{zx}(x) + P_{kx}(x) \cdot P_{xz}(z)}{\int P_{kx}(x) \cdot P_x(z) dx}. \tag{5}$$

It can be seen that Bayesian filtering saves storage space by not having to save and reprocess past measurement data. However, the method of calculating the posterior probability from equation (3) and equation (4) is only theoretical because the actual equation (5) is difficult to calculate to get the exact value. In a linear system, the optimal solution can be obtained by Kalman filtering. In the case of nonlinear models, the EKF, UKF, and CKF can be used to solve the posterior probabilities.
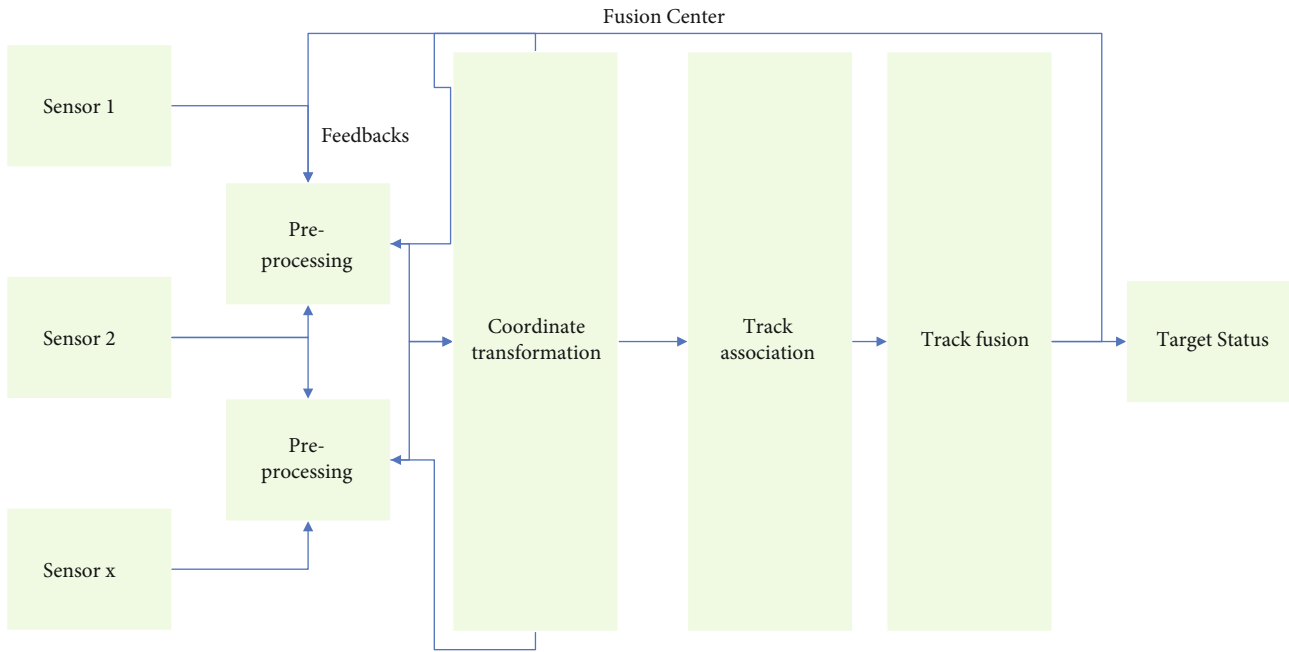
Fusion Center

Sensor 1

Feedbacks

Pre-
processing

Sensor 2

Coordinate
transformation

Track
association

Track fusion

Target Status

Pre-
processing

Sensor x

FIGURE 1: Centralized convergence structure diagram.

Association

Local track
management

Sensor 1

Gating

Filtering&pre
diction

Spatial &
temporal
alignment

Association

Local track
management

Sensor 2

Gating

Filtering&pre
diction

Spatial &
temporal
alignment

Fusion of
local
tracks

Association

Local track
management

Sensor x

Gating

Filtering&pre
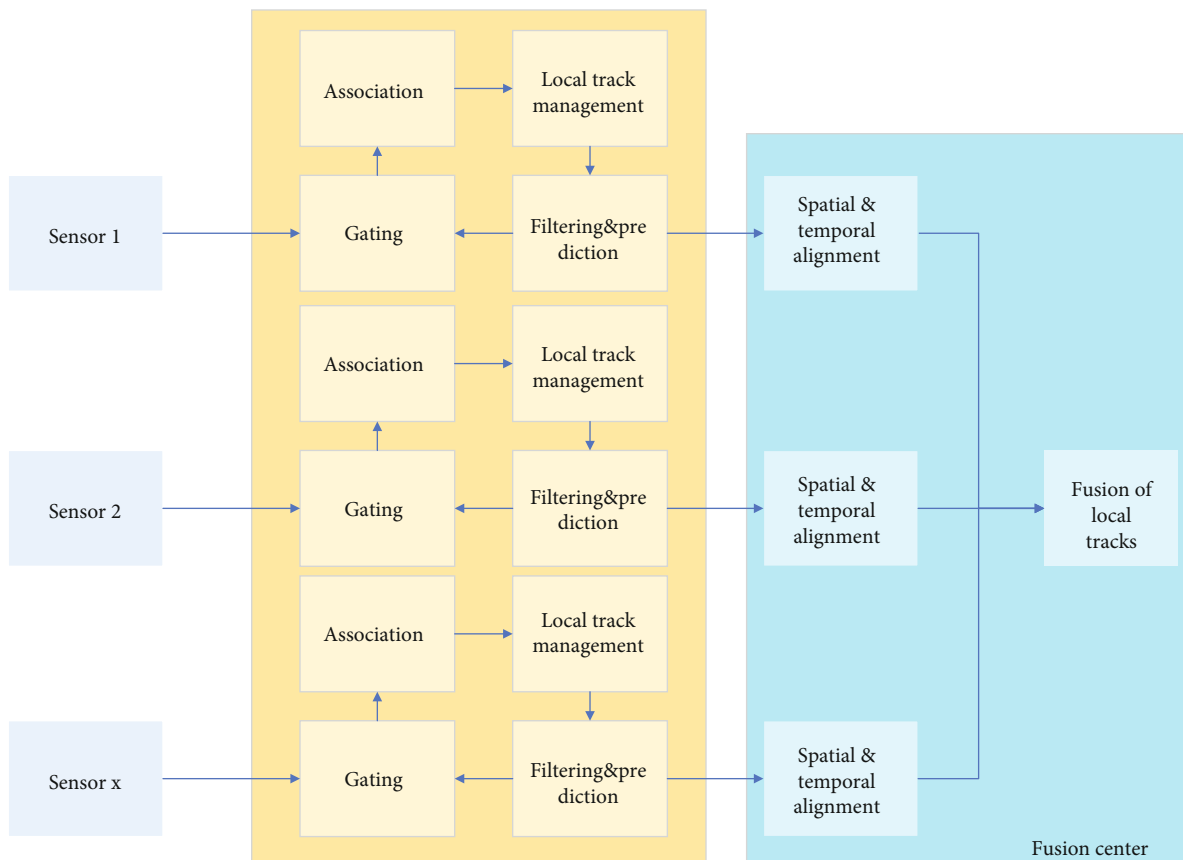diction

Spatial &
temporal
alignment

Fusion center

FIGURE 2: Distributed fusion structure diagram.

The proposal of Deep Convolution Generative Adversarial Networks (DCGAN) has given a great impetus to the development of GAN by combining the convolutional neural network model (CNN) and GAN, which enables the quality and diversity of the generated images. Compared with the traditional GAN, DCGAN has been improved in
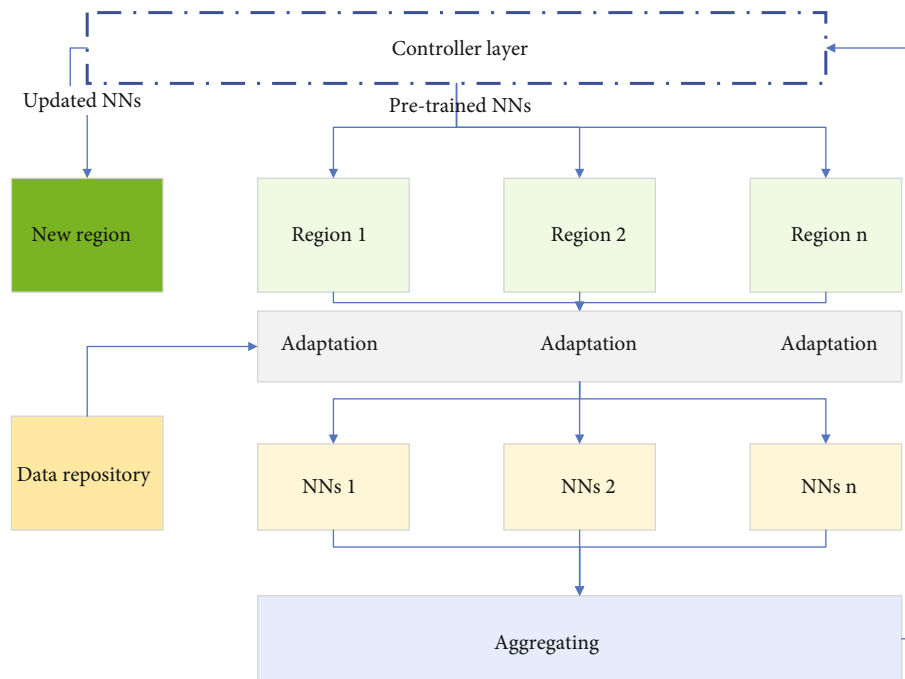
FIGURE 3: Hybrid fusion structure diagram.

several aspects. Mainly, the pooling layer is eliminated, the fully connected layer is removed, and a series of training techniques are used, such as using batch normalization (BN) to stabilize the training and using the REL activation function to reduce the risk of gradient disappearance [23]. First of all, the DCGAN model replaces all pooling layers with convolutional layers. The discriminator uses stepwise convolution instead of pooling layers, and the generator uses fractional stepwise convolution instead of pooling. The second point is to remove the fully connected layer from the model and use the global pooling layer instead of the fully connected layer, which effectively reduces the parameters of the model on one hand and improves the operation speed of the network on the other hand. The third point is that the use of batch normalization can alleviate the problem of "gradient dispersion" in deep neural networks and accelerate the convergence of the model. The discriminator uses the LeakyReLU activation function for all layers, and the generator uses the ReLU activation function except for the output layer, which uses the hyperbolic tangent function Tanh.

The input size of the discriminator model based on the SAR dataset expansion of the generative adversarial network is 8888, and the model mainly consists of 4 convolutional layers with a convolutional kernel size of 33 and convolutional depths of 32, 64, 128, and 256 and finally a flattening layer to obtain the prediction results. In addition, a set of ReLU activation function layers and deactivation layers are connected after each convolutional layer. The ReLU activation function layer is mainly to increase the representation capability of the discriminative model, and the deactivation layer is mainly to reduce the overfitting problem of the model during training and to improve the generalization capability of the discriminative model. The batch normaliza-

tion layer is added after the latter convolutional layers of the discriminant model. Since the batch normalization process will normalize the features, it is beneficial to speed up the convergence of the discriminant model. The discriminator outputs the class token of its input SAR image according to the source of the input SAR image and outputs 1 if the input is from a real sample or 0 if the input is from a sample generated by the generator. The specific discriminator model is shown in Figure 4.

3.2. Improved Data Fusion Structure.
In this paper, to achieve real-time detection of infrared targets in complex environments, the YOLOv3 algorithm with speed advantage is selected as the base network for infrared target detection and improved on it. The improved network structure is shown in Figure 5. Since YOLOv3 uses three sizes of feature maps for target detection and fuses shallow features with deep features to improve the detection capability of small targets, but the feature maps for detecting large targets do not have a large enough sensory field, so the SPP module is added after the feature extraction network to fuse local features with global features and enhance the feature expression capability, to solve the detection problem caused by the change of target scale. The SPP module is added after the feature extraction network to fuse the local features with the global features and enhance the feature representation, to solve the problem of the decrease of detection accuracy caused by the change of target scale [24]. The regression loss function in the original YOLOv3 network is replaced by the GIoU loss function for the regression of the prediction frame, and the prediction frame is considered as a whole to calculate the loss with the true value frame to improve the accuracy of the whole network localization.
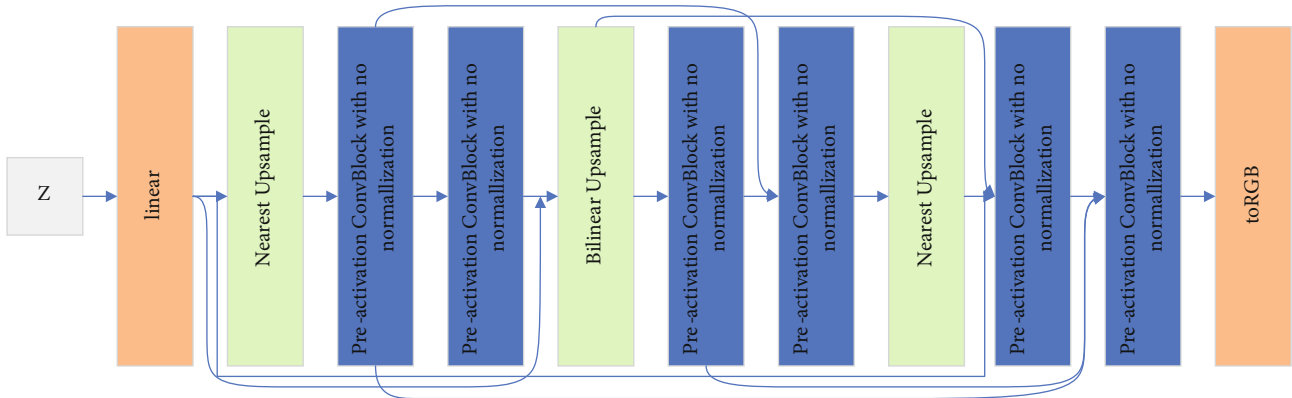
FIGURE 4: Discriminative network for generative adversarial network model.
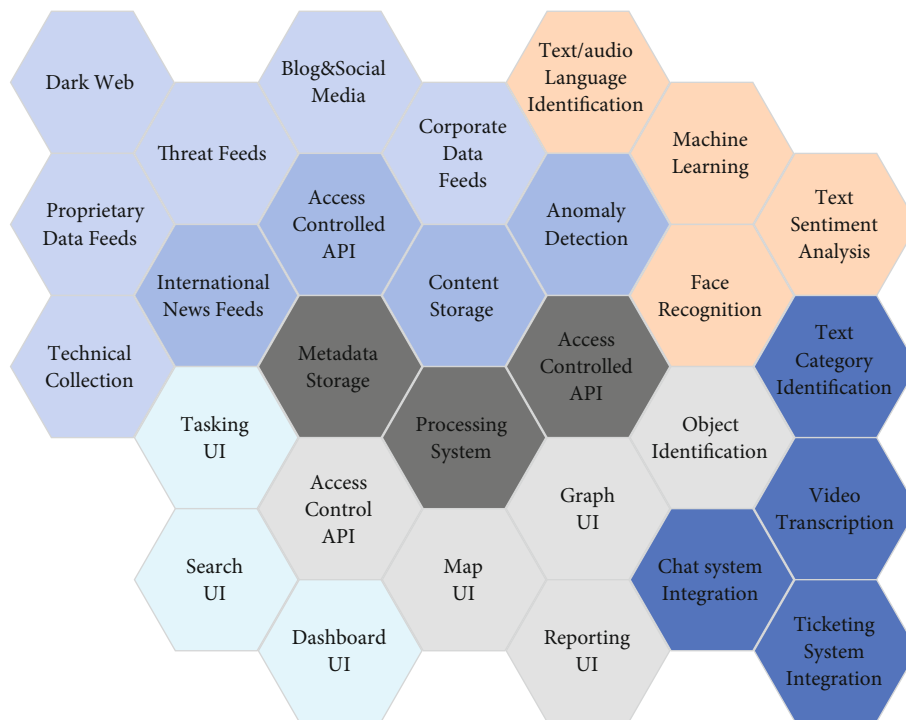

FIGURE 5: Improved data fusion structure.

In YOLOv3, the possible reason for the degradation of large-scale target detection accuracy is that the deepest feature map perceptual field is not large enough. Therefore, an SPP module is added after the feature extraction network. The design of the whole SPP module is based on the idea of a spatial pyramid, which uses multiple channels to process the input feature map in parallel, and the four branches use different sizes of pooling kernels. Firstly, the input feature map is downscaled by a $1*1$ convolution kernel to fuse the features of different channels. Then, it passes through a $1*1$ size pooling layer, i.e., to obtain global features, and then passes through $5*5, 9*9$, and $13*13$ to obtain feature maps of different sizes, to obtain different feature information from the input feature maps through different channels and finally fusing the obtained features.

In general, to address the problem of the perceptual field in the deep detection layer, a multichannel pooling kernel is used to fuse local and global features of different sizes to enrich the feature expression capability of the network and expand the perceptual field of the feature map, while avoiding the reduction of network training speed due to the use of convolution, which is helpful to improve the situation of accuracy loss due to the relatively large span of target scales in the image to be detected.

The MSRCR algorithm, multiscale retinal enhancement with color recovery, was developed based on the single-scale Retinex algorithm and the multiscale weighted average MSR algorithm. Retinex is an image enhancement algorithm as a word consisting of the words retina (retina) and cortex (cortex) [25]. Retinex theory is based on the idea that the
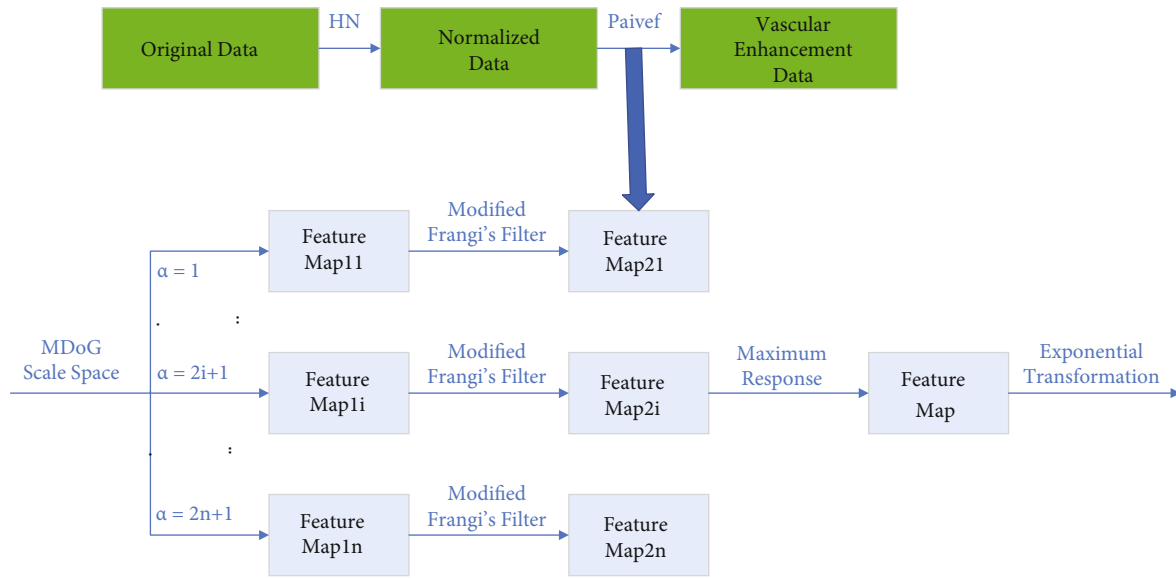
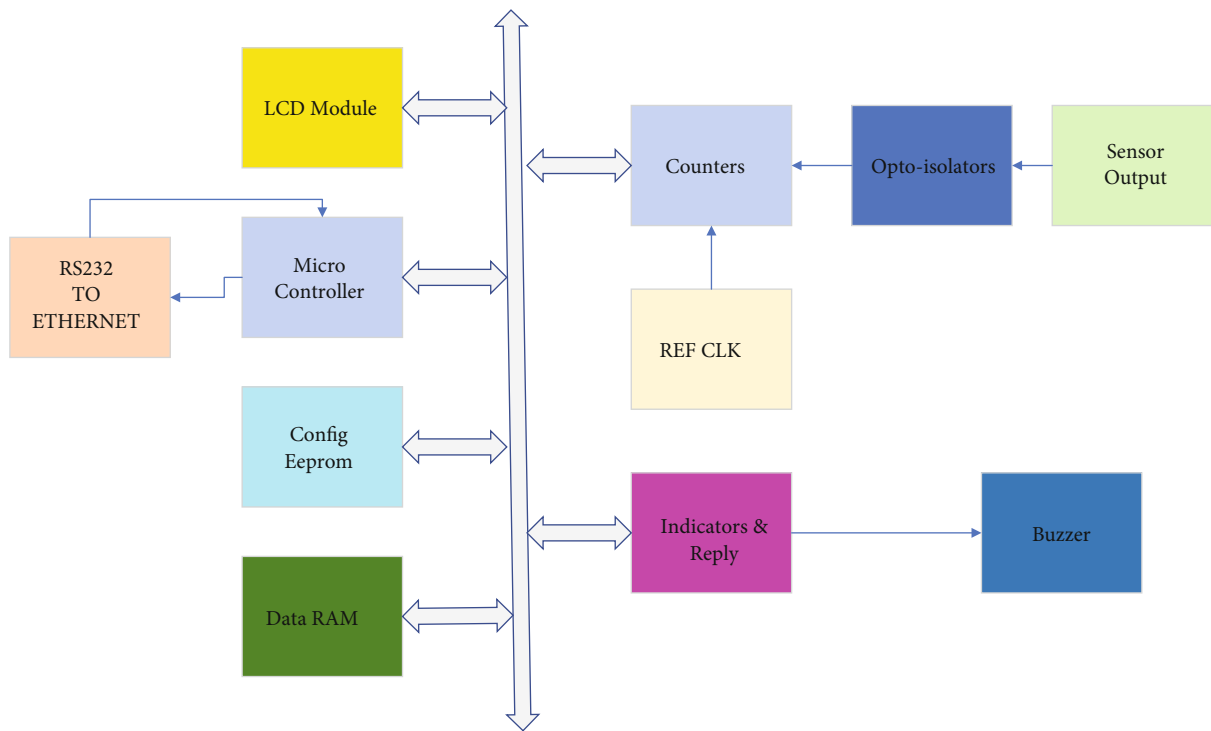FIGURE 6: Flow chart of MSRCR image enhancement algorithm.

FIGURE 7: Flow chart of optical sensor ranging program.

color of an object is consistent regardless of lighting nonuniformity, and unlike traditional linear and nonlinear image enhancement methods, it can achieve a balance in edge enhancement, color constancy, and dynamic range compression, so that adaptive enhancement can be performed for many types of images. The MSRCR algorithm is developed on this basis, which can maintain the high fidelity of the image and compress the dynamic range of the image, as well

as perform the color enhancement of the image and perform the local and global dynamic range compression. However, the abovementioned image enhancement process may distort the color of local details and deteriorate the overall visual effect due to the increase of noise. Therefore, the MSRCR algorithm is proposed, and a color recovery factor C is added to the MSR algorithm to solve the problem of color distortion due to the contrast enhancement of local

Table 1

(a) Optical sensor distance measurement experimental data

Serial number	Actual value	Measurements	Error
1	40	40.866	0.866
2	80	81.256	1.256
3	120	119.342	-0.658
4	160	161.864	1.864
5	200	199.066	-0.934
6	240	239.874	-0.126
7	280	281.798	1.798
8	320	320.826	0.826
9	360	361.356	1.356
10	400	399.574	-0.426

(b) Optical sensor test data after calibration

Serial number	Actual value	Measurements	Error
1	40	39.866	0.866
2	80	79.887	-0.133
3	120	120.502	0.502
4	160	160.302	0.302
5	200	200.904	0.904
6	240	240.291	0.291
7	280	280.757	0.757
8	320	319.125	-0.875
9	360	359.725	-0.276
10	400	400.333	0.333

areas of the image [26–28]. The flow of the MSRCR image enhancement algorithm designed in this paper is shown in Figure 6.

4. Experiment and Analysis

4.1. Experimental Design. The optical sensor target recognition experiment is based on the distance between the target identifier and the unmanned cart measured by the sensor, and the sensor performance directly affects the accuracy of the measured distance [29–33]. To test the performance of the sensor, experimental verification of the performance of individual sensors is required. The STM32F407, the PC host computer, and the optical sensor are used for the distance measurement experiments. The program flow chart is shown in Figure 7.

The HC-SR04 optical ranging module has four pins: Vcc, Trig, Echo, and GND. In this design, the STM32F407 development board (hereinafter referred to as STM32) is directly connected to the optical ranging module, Vcc is connected to the 5 V voltage output port on STM32, Gnd is connected to the ground, Trig is connected to PF6, and Echo is connected to PF5. The HC-SR04 optical range module is a trigger-type range measurement; in each measurement, PF6 transmits a high level of $10\,\mu s$-$20\,\mu s$ to Trig to make the transmitter emit the optical; Echo detects the level of the receiver; and the timer on STM32 calculates the duration of Echo's high level, and stores it in the register to calculate the distance. In the experiment, the experimental distance is varied in the range of 20 cm-280 cm, and the measured distance information is displayed on the PC host computer utilizing serial printing. To reduce the influence of sensor jitter on the data during the measurement, the same distance was measured five times and the average value was taken as the output. The test results are shown in Table 1(a). From Table 1(a), it can be seen that the partial measurement error of the sensor is greater than ±2 cm, and there is a large distortion in the measured value, which obviously cannot be used directly. In this regard, MATLAB can be used to correct the experimental data by performing a linear fit to the experimental data. Table 1(b) shows the data obtained after correction. From the data in the table, the error of the sensor measurement data after correction is reduced to within ±1 cm, which satisfies the experimental requirements of the target recognition experiment.

Through an in-depth study of the fuzzy control target recognition algorithm, the fuzzy controller collects not only the distance between the unmanned cart and the target identifier measured by the light sensor but also the direction of the unmanned cart relative to the target point. In the experimental design of the target recognition algorithm of the unmanned vehicle, the idea of this paper is to set the initial direction of motion of the unmanned vehicle to a constant 90, that is, to the front, and the steering angle obtained by the unmanned vehicle in the subsequent target recognition calculation should be added or subtracted from the constant to obtain the new angle value and stored in the register of the control core. After the first turn of the unmanned vehicle, the angle value is used as an input to the fuzzy controller for each target recognition calculation as the relative direction to the target point. The purpose of this design is to minimize the influence of the target identification process of the unmanned trolley on the original driving direction as much as possible and to maintain the original direction of driving. Target recognition experiments are conducted according to the analyzed target identifier situations, with a total of six cases and 32 possible target identifier arrangements. The experiment verifies that the unmanned trolley can detect the target identifier well in all possible environments and can make the corresponding form of movement away from the target identifier; because of the distribution of the target identifier and the possible causes of the shape of the target identifier, this paper does not enumerate them one by one.

In the design of this paper, the deflection angle of the next move of the unmanned trolley output by the target recognition algorithm is converted into the rotational speed of the left and right wheels of the unmanned trolley by the angle-velocity relationship equation, and the PWM signal with corresponding duty cycle is generated by the internal timer and register of the STM32 main control chip to control the rotational speed of the DC motor. To achieve the smooth motion of the unmanned trolley, the speed of the two motors in the trolley should be consistent, so the output PWM signal needs to be regulated by the PID algorithm. In
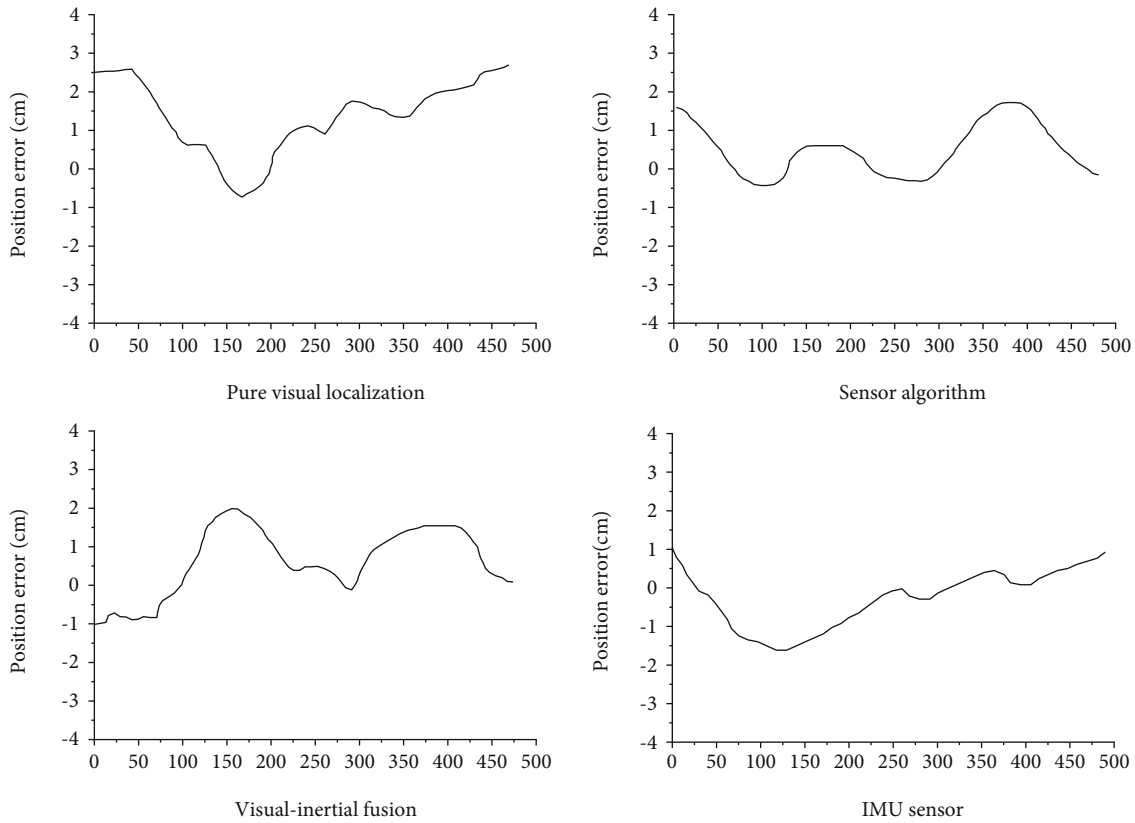
FIGURE 8: Position error trajectory curve analysis.

this design, the left motor speed is used as the reference, and the PID algorithm controls the right motor speed to be the same as the left motor speed. The speed and steering of the unmanned vehicle are controlled by the STM32 and the L298N, which receive the PWM signal from the STM32 and control the motor speed and direction of forward and reverse rotation. The PWM signal is generated using the TIM3 timer channel 1 of STM32 to control the motor speed, and the PWM signal is output to ENA and ENB of L298N through GPIO port PA7; the STM32 connects to IN1 and IN2 of L298N through PA4 and PA5, respectively, to control the forward and reverse rotation of the left motor through PA13 and PA14. The STM32 controls the forward and reverse rotation of the right motor by connecting the IN3 and IN4 interfaces of the L298N through PA13 and PA14, respectively. In other words, IN1, IN2, IN3, and IN4 form an H-bridge circuit to control the motor steering, respectively.

4.2. Experimental Results. To improve the localization accuracy and robustness of the pure vision SLAM system, a tightly coupled vision, IMU sensor algorithm is used in this thesis. The localization system consists of sensor data preprocessing, system initialization, sliding window pose solver module, loopback detection, and global pose graph optimization module. The initialization correction and minimum error objective function are constructed by the preintegration model of the IMU sensor and visual image frames. Con-
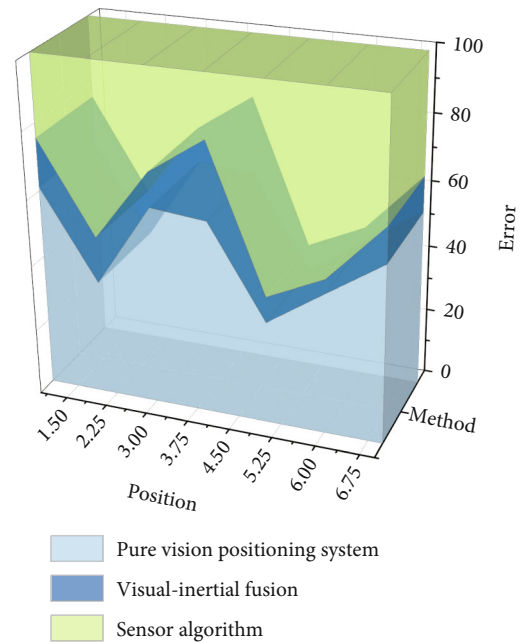

FIGURE 9: Attitude error trajectory curve analysis.

sidering the real-time of the system, the sliding window algorithm and keyframe extraction model are used to save computation, and finally, the accurate positional trajectory
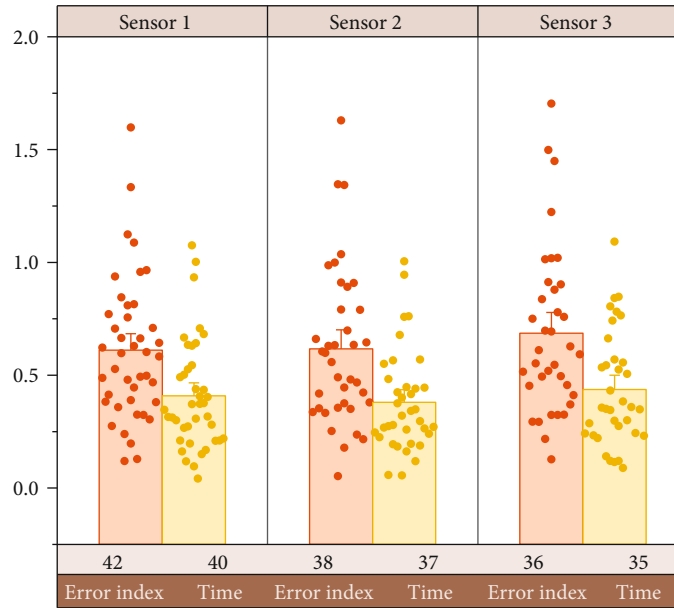
FIGURE 10: Accelerometer error analysis diagram.

and sparse point cloud information are obtained by combining the graph optimization library and loopback detection correction. The experimental validation of the localization algorithm is carried out according to the above fusion theoretical methods and steps, using the mechanical experimental building of MH-02 in the EuRoc dataset, which contains ground-truth values based on motion capture device acquisition. Then, the error comparison analysis of pure visual localization and visual-inertial fusion localization algorithms was performed, as shown in Figure 8 below. Because the adopted dataset is light stable and rich in feature points, the position errors in both pure visual and visual-inertial fusion localization systems in space are not very large and are very close to the ground-truth values, and relatively, the visual-inertial fusion localization system is more accurate in the z-axis.

However, the attitude error comparison shown in Figure 9 shows that the pure vision positioning system has a large error, with a maximum angular error of 60 degrees, but the vision-inertial guidance fusion positioning system incorporates the IMU sensor, and the attitude information is almost close to the ground-truth value.

The error and bias of the accelerometer and gyroscope in the IMU were also corrected and calibrated and analyzed as shown in Figure 10 below, which can be used as an a priori error to influence the weight determination problem of multisensor data fusion, and the determined error and bias can be used to compensate the IMU sensors for more accurate measurements.

5. Conclusion

With the development of optical sensors, target recognition algorithms based on optical sensor data fusion are widely valued and have a better prospect in the field of robotics in the future. At present, multisensor fusion localization and navigation technology is the foundation and key function in aerospace, military defense, logistics and transportation, smart factory, and biomedical fields. In this paper, target recognition with multisensor data fusion is investigated mainly for the combination of multiple optical sensors equipped with a depth camera, LIDAR, and IMU sensors in outdoor working scenarios as well as the indoor environment with light influence, and the following results are achieved:

(1) In this thesis, we designed an algorithm based on adaptive extended Kalman filtering to fuse GPS and IMU sensor acquisition data for the problem of signal occlusion in an outdoor environment. At the same time, a predictive tracking model based on multisensor target recognition is designed, and an environment sensing algorithm program is designed based on the point cloud imaging model of LiDAR. The fusion improves the robustness of navigation trajectory tracking and positioning accuracy, reduces the maximum error by 1.5 m, and achieves centimeter-level positioning accuracy by combining RTK technology

(2) In this thesis, a visual-inertial guidance tight coupling algorithm based on nonlinear optimization is designed for the indoor interferer dense problem. Firstly, an image feature point extraction and IMU preintegration model based on the improved feature point method is designed to solve the positional pose in combination with PNP positional estimation algorithm and back-end map optimization algorithm. The least-squares error objective optimization function and sliding window model are also constructed to realize the real-time positional problem solution. The analysis results show that the average error is

improved by nearly 50% after the algorithm fusion, and the minimum error is only 0.02 m, and the attitude trajectory information is closer to the real value on the ground

(3) To address the visual-inertial guidance fusion algorithm's degradation in localization accuracy and robustness under low-light conditions, which affects the target recognition task of optical sensors. In this thesis, a procedure is designed to switch to LIDAR localization mode when the number of feature extraction matches is lower than a set threshold, and the laser localization and map building algorithms are verified by building a Gazebo smart car simulation experimental platform. Finally, a laser vision inertial guidance fusion localization system based on ROS architecture is designed for low-light conditions, and the constructed raster maps are used for navigation tasks. The trajectory of the fused LIDAR sensor in the low-light environment is compared to be smoother and with reduced error, with a maximum drop of about 0.53 m

Although this paper verifies the correctness of adaptive extended Kalman filtering and nonlinear optimal fusion in multioptical sensor target recognition tasks, improvements are still needed in practical engineering, mainly from the following two aspects: first, to improve the autonomy and environmental adaptivity of optical sensors and to complete the autonomous switching of different target recognition modes indoors and outdoors by judging the number of received light source signals. Second is to improve the perception ability of optical sensors, combined with deep learning technology and laser vision fusion to build three-dimensional semantic information to better adapt to dynamic and unstructured scenes.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

We declare that there is no conflict of interest.

Acknowledgments

References

[1] D. Gozdowski, M. Stępień, E. Panek et al., "Comparison of winter wheat NDVI data derived from Landsat 8 and active optical sensor at field scale," *Remote Sensing Applications: Society and Environment*, vol. 20, article 100409, 2020.

[2] A. N. Bishop, B. Fidan, B. Anderson, K. Doğançay, and P. N. Pathirana, "Optimality analysis of sensor-target localization geometries," *Automatica*, vol. 46, no. 3, pp. 479–492, 2010.

[3] M. Hu and Q. Hu, "Design of basketball game image acquisition and processing system based on machine vision and image processor," *Microprocessors and Microsystems*, vol. 82, no. 1, article 103904, 2021.

[4] T. Meng, X. Jing, Z. Yan, and W. Pedrycz, "A survey on machine learning for data fusion," *Information Fusion*, vol. 57, pp. 115–129, 2020.

[5] X. Wang, S. Wang, and J. J. Ma, "An improved particle filter for target tracking in sensor systems," *Sensors*, vol. 7, no. 1, pp. 144–156, 2007.

[6] P. Ghamisi, R. Gloaguen, P. M. Atkinson et al., "Multisource and multitemporal data fusion in remote sensing: a comprehensive review of the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 7, no. 1, pp. 6–39, 2019.

[7] R. T. Wu and M. R. Jahanshahi, "Data fusion approaches for structural health monitoring and system identification: past, present, and future," *Structural Health Monitoring*, vol. 19, no. 2, pp. 552–586, 2020.

[8] K. Zhang, J. Wei, T. Wang, S. Li, and X. Yang, "Air target recognition algorithm based on mixed depth features in the interference environment," *Optik*, vol. 245, article 167535, 2021.

[9] L. Kong, X. Peng, Y. Chen, P. Wang, and M. Xu, "Multi-sensor measurement and data fusion technology for manufacturing process monitoring: a literature review," *International Journal of Extreme Manufacturing*, vol. 2, no. 2, article 022001, 2020.

[10] Y. Xu, B. du, L. Zhang et al., "Advanced multi-sensor optical remote sensing for urban land use and land cover classification: outcome of the 2018 IEEE GRSS data fusion contest," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 6, pp. 1709–1724, 2019.

[11] W. Ding, X. Jing, Z. Yan, and L. T. Yang, "A survey on data fusion in internet of things: towards secure and privacy- preserving fusion," *Information Fusion*, vol. 51, pp. 129–144, 2019.

[12] D. Nikolic, N. Stojkovic, Z. Popovic et al., "Maritime over the horizon sensor integration: HFSWR data fusion algorithm," *Remote Sensing*, vol. 11, no. 7, p. 852, 2019.

[13] M. Muzammal, R. Talat, A. H. Sodhro, and S. Pirbhulal, "A multi-sensor data fusion enabled ensemble approach for medical data from body sensor networks," *Information Fusion*, vol. 53, pp. 155–164, 2020.

[14] N. Bakalos, A. Voulodimos, N. Doulamis et al., "Protecting water infrastructure from cyber and physical threats: using multimodal data fusion and adaptive deep learning to monitor critical systems," *IEEE Signal Processing Magazine*, vol. 36, no. 2, pp. 36–48, 2019.

[15] W. Zhang, Y. Ning, and C. Suo, "A method based on multi-sensor data fusion for UAV safety distance diagnosis," *Electronics*, vol. 8, no. 12, p. 1467, 2019.

[16] X. Wang, D. Mao, and X. Li, "Bearing fault diagnosis based on vibro-acoustic data fusion and 1D-CNN network," *Measurement*, vol. 173, article 108518, 2021.

[17] S. Yin, H. Li, L. Teng, M. Jiang, and S. Karim, "An optimised multi-scale fusion method for airport detection in large-scale optical remote sensing images," *International Journal of Image and Data Fusion*, vol. 11, no. 2, pp. 201–214, 2020.

[18] K. Heckel, M. Urban, P. Schratz, M. D. Mahecha, and C. Schmullius, "Predicting forest cover in distinct ecosystems: the potential of multi-source Sentinel-1 and -2 data fusion," *Remote Sensing*, vol. 12, no. 2, p. 302, 2020.

[19] A. Diez-Olivan, J. Del Ser, D. Galar, and B. Sierra, "Data fusion and machine learning for industrial prognosis: trends and perspectives towards Industry 4.0," *Information Fusion*, vol. 50, pp. 92–111, 2019.

[20] N. Long, K. Wang, R. Cheng, W. Hu, and K. Yang, "Unifying obstacle detection, recognition, and fusion based on millimeter wave radar and RGB-depth sensors for the visually impaired," *Review of Scientific Instruments*, vol. 90, no. 4, article 044102, 2019.

[21] M. Belgiu and A. Stein, "Spatiotemporal image fusion in remote sensing," *Remote Sensing*, vol. 11, no. 7, p. 818, 2019.

[22] J. Gao, P. Gu, Q. Ren, J. Zhang, and X. Song, "Abnormal gait recognition algorithm based on LSTM-CNN fusion network," *IEEE Access*, vol. 7, pp. 163180–163190, 2019.

[23] P. Wang, L. Wang, H. Leung, and G. Zhang, "Super-resolution mapping based on spatial–spectral correlation for spectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 3, pp. 2256–2268, 2021.

[24] C. Gong, Y. Hu, J. Gao, Y. Wang, and L. Yan, "An improved delay-suppressed sliding-mode observer for sensorless vector-controlled PMSM," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 7, pp. 5913–5923, 2020.

[25] X. Song, J. Huang, J. Cao, and D. Song, "Multi-scale joint network based on Retinex theory for low-light enhancement," *Signal, Image and Video Processing*, vol. 15, no. 6, pp. 1257–1264, 2021.

[26] M. Jahanbakht, W. Xiang, L. Hanzo, and M. Rahimi Azghadi, "Internet of Underwater Things and big marine data analytics—a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 904–956, 2021.

[27] S. Yang, X. Wei, B. Deng, C. Liu, H. Li, and J. Wang, "Efficient digital implementation of a conductance-based globus pallidus neuron and the dynamics analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 494, pp. 484–502, 2018.

[28] F. Orujov, R. Maskeliūnas, R. Damaševičius, and W. Wei, "Fuzzy based image edge detection algorithm for blood vessel detection in retinal images," *Applied Soft Computing*, vol. 94, article 106452, 2020.

[29] Y. Gong, Z. Ma, M. Wang, X. Deng, and W. Jiang, "A new multi-sensor fusion target recognition method based on complementarity analysis and neutrosophic set," *Symmetry*, vol. 12, no. 9, p. 1435, 2020.

[30] J. Yang, M. Xi, B. Jiang, J. Man, Q. Meng, and B. Li, "FADN: fully connected attitude detection network based on industrial video," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2011–2020, 2021.

[31] X. X. du, Y. Mu, Z. W. Ye, and Y. J. Zhu, "A passive target recognition method based on LED lighting for industrial internet of things," *IEEE Photonics Journal*, vol. 13, no. 4, pp. 1–8, 2021.

[32] L. Tao, X. Jiang, X. Liu, Z. Li, and Z. Zhou, "Multiscale supervised kernel dictionary learning for SAR target recognition," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 9, pp. 6281–6297, 2020.

[33] D. Neupane and J. Seok, "A review on deep learning-based approaches for automatic sonar target recognition," *Electronics*, vol. 9, no. 11, p. 1972, 2020.