Research on Mobile Robot Target Recognition and Obstacle Avoidance Based on Vision

Lei Zhang, Dezhong Li

School of Electrical and Information Engineering
Beijing University of Civil Engineering and Architecture, China
leizhang@bucea.edu.cn, 528405176@qq.com

Abstract

This work investigates the method of object recognition and autonomous obstacle avoidance for mobile robot based on vision, and solves to the problem that mobile robot can move and identify objects in complex environment. Build the Mecanum wheel mobile platform and equip the Kinect sensor and laser rangefinder sensor, transmitted data to the host computer to information decision analysis, and the execution system generation mobile robot motion control commands, then upload the commands to the cloud data center and establish the SQL Server database table to store the information. Robot scan 100ms of every table, to perform the latest control information. 0bject detection uses Gauss model background difference method to detect objects. Target feature extraction adopts the SURF (Speed Up Robust Features) algorithm and uses RANSAC (Random Sample Consensus) algorithm to optimization, removing the mismatching points improves the computation speed and the detection accuracy. Autonomous obstacle avoidance module through the laser range finder to collect distance information, set the safety distance with the obstacles, can achieve the effect of self-obstacle avoidance. The experiment proves that the robot can complete the task of target recognition and autonomous obstacle avoidance, and verify the effectiveness of the system.

Keywords: Mobile robot, Machine vision, Target recognition, Autonomous obstacle avoidance, Cloud data center

1 Introduction

Mobile robot technology is one of the forefront of scientific exploration, including the intelligence, detection and recognition. With the improvement of the technology, there are many mobile robots that need improved technology, in the autonomy, independence, intelligence, especially hoping the mobile robot can independently complete the specific task. Among them, multi-sensor fusion is a very important part of the field, which refers to the use of sensor acquisition device to replace the human senses, from the surrounding environment to extract the relevant information, and then analyze and sort out the information to complete the assigned task, and better achieve harmony with human beings.

The key problem of mobile robot's target recognition when using visual sensor is how to extract the effective feature from the huge video image, that is: how to realize the target recognition task of mobile robot in complex natural environment by using a fast and effective target recognition algorithm, and at the same time to meet the robot's own real-time and flexibility requirements.

The target recognition can be completed in two steps. First, the feature description of the target should be completed, and then the feature of the target is matched with the target feature information extracted from the scene. There are three ways to characterize the target: first, based on a priori knowledge of the target's geometry. These prior knowledge is obtained by drawing tools or simple sensors, and then matches with the input image data to achieve the purpose of identifying objects. The geometric features of the object are more intuitionistic and insensitive to light, but it is difficult to meet the requirements of real-time target recognition in the process of subsequent target feature matching. Second, based on the appearance of object. Objects are photographed from different angles, and objects are represented by the appearance information of objects in the images. The appearance information includes contour, color, shape, etc. In the process of recognition, first, the complex scene is divided into several sub regions, and then the regions are matched with the target appearance features to search for targets. Such as color histogram, perceptual domain color histogram, and so on, this method is more intuitive, but because of the deformation problem introduced by multiple perspectives and the sensitivity of this method to light, it is easy to affect the accuracy of recognition. The third is a method based on local features. The common local features include point

*Corresponding Author: Lei Zhang; E-mail: leizhang@bucea.edu.cn
DOI: 10.3966/160792642018111906023

feature, line feature and regional feature, where point feature can be divided into corner feature and extreme point feature. Harris and Stephens proposed the Harris corner point detection algorithm. Smith proposed a SUSAN corner point detection algorithm. Edward Rosten proposes a fast corner point detection algorithm FAST. In 1999, David Lowe proposed a SIFT algorithm based on target local feature recognition. The SIFT extreme point features are insensitive to illumination. The SIFT descriptor is robust to rotation, scaling and perspective deformation, but the complexity of the SIFT descriptor is high, and the amount of computation is too large. In 2004, Y.Ke proposed that PCA method be used to reduce the dimension of SIFT descriptors in order to reduce computation. In 2006, the emergence of SURF algorithm greatly improved the speed of local feature extraction. In 2010, the BRIEF descriptor was proposed by Michael Calonder and so on. Compared with the corresponding matching algorithm, the matching speed of the corresponding gradient descriptor was greatly improved, and the requirement of real-time recognition could be achieved. In 2011, Ethan Rublee proposed ORB algorithm, which made BRIEF comparison descriptor be applied in feature matching. However, because the comparison descriptor is sensitive to noise and does not have affine invariance, it is not applicable in complex environment.

The feature matching method of the target can be divided into two categories: the first is the feature matching based on the region. The method is mainly used in the appearance based feature description method. By comparing the feature information of several sub regions divided by the scene, the target matching is realized. The second kind is the matching method based on feature. This method is to match the corner, edge contour and line feature of the image. It mainly solves the problem of finding the approximate target features quickly and accurately from a large number of features. Generally, the feature similarity criterion is used as the feature vector in Euclidean distance. There are several common feature matching methods. The linear search algorithm compares the target feature with the scene feature one by one. It is the most accurate matching method, but the matching time increases with the number of features linearly and is not suitable for the matching of a large number of features. Based on epipolar geometric constraint method, the feature is reduced from two-dimensional search to one-dimensional, and the search time is reduced. The Kd tree search algorithm constructs a Kd tree structure with the features of the scene to be matched, searches for the features according to the structure of the Kd tree, speeds up the search speed, but is not suitable for the matching of the high dimensional eigenvectors. BBF algorithm is an improvement of Kd tree algorithm, which solves the matching problem of high-dimensional eigenvectors, and is suitable for fast matching of large number of high dimensional features.

In this work we study the target recognition problem and integrate simple autonomous obstacle avoidance function of mobile robot based on monocular vision in indoor environment, and focus on how to transmit data efficiently.

By the study of a large number of target recognition algorithms, the object recognition algorithm based on local feature extraction and matching is studied from how to extract different image features. The vision is used to find the true targets detected by the radar and to discard those who might be false positives by the method of combining offline SVM classifier, saliency detection and online compressive tracking [1]. This paper describes algorithms for Autonomous Surface Vehicle (ASV) obstacle avoidance and target search task. In this task, ASV must avoid obstacle buoys, while it is searching for totem-shaped buoy. 2D scanning LIDAR and monocular vision sensor are also used in detecting [2]. In this paper, they propose a deep neural network for scene depth estimation that is trained on synthetic datasets, which allow inexpensive generation of ground truth data. We show how this approach is able to generalize well across different scenarios [3]. This paper presents the design of a mobile robot which could follow a track and shoots a ball to the target set at the distance in a relay race. The design of the motor control circuit, speedometer, sensors and embedded control system are detailed respectively [4]. In this work, we address the problem of human body pose recognition using RGB-D sensor, to perform user tracking by a mobile robot. User's skeleton joints orientations are used in this approach to compute torso joint orientation [5]. The system realizes target tracking by extracting the color of the target as the feature, filtering and detecting feature of the target and then displays it on the screen. It realizes target positioning according to the centroid of the tracked target, then computing the orientation which can be used for the intelligent robot vision [6]. These methods are based on HSV color segmentation and also on local features recognition with SURF (Speeded-Up Robust Features). Client-Server communication models are developed to connect the navigation control with the arm controller which is in turn also connected with the Kinect control for visual processing [7]. A wireless navigation mobile robot system is proposed in this paper for both path planning and trajectory execution. The use of Developed image processing while analyzing algorithms and orientation is based on the color markers recognition to determine the robot's position [8].

Mobile robot target recognition and autonomous obstacle avoidance technology can only use a specific target recognition technology for a task [9]. In this paper, Gaussian model background difference method is used in detection of moving objects. Gaussian modeling method is used to determine the background and prospects. The feature matching detection method based on the SURF algorithm is used to determine whether the target is detected by calculating the matching degree between the sample feature and the feature of the image to be examined. The main advantage is that it converts the rich image pixel information into the image feature, on the other hand it is robust to the target change, even if the target has a certain deformation in addition rotation can also be detected, but the template image and video capture images that exist in the mismatch point will be the target of the detection of the speed to reduced detection accuracy. In order to improve the accuracy of the retrieval, the RANSAC algorithm is used to optimize the matching accuracy of the SURF algorithm, and the mismatch points are removed [10]. The optimized algorithm does not only improves the running speed, but also enhances the matching accuracy.

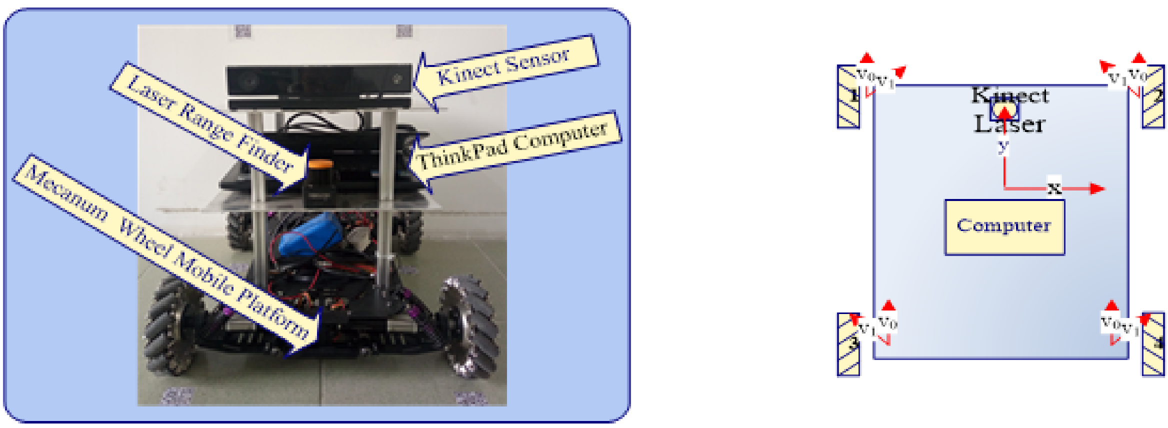
2 Mobile Robot Platform

In this work, the hardware structure of the mobile robot, according to the dual structure will be designed as two sub-modules : The upper part of PC is responsible for the sensor information data. The lower computer includes monocular vision sensor, laser range finder communication module, drive module, power module, and Mecanum wheel which makes up the mobile platform. The lower computer responsible for the implementation of the robot motion control system.

2.1 Mechanical Structure

Mecanum four-wheel mobile robot, can achieve forward, reverse, left and right movement, 45 degrees oblique movement, 360 degrees all round zero radius and other actions, in the indoor environment, to achieve free and more flexible movement. In this work, the four Mecanum wheel drive mechanism used in this paper, the wheel diameter of 15cm, thickness of 3.5cm, made of stainless steel; there are 16 of the 45 degrees driven roller, made of rubber material. The wheel weighs 0.65kg. Platform size is 60cm×45cm×60cm. Four wheels with 45 degree shear cross distribution, shaft spacing 55cm, wheel spacing 45cm. Use a DC servo brush motor drive, comes with motor reduction gear ratio 16: 1. The maximum torque output is 346.66 g·cm. Using DC motor drive, current and speed dual closed-loop control method, the drive uses USB_CAN communication mode. The robot drives the battery with a 24V 10AH DC battery.

As shown in Figure 1(a), the robot is mainly composed of two hardware modules, motion control module and information acquisition module. The motion control module is implemented using a Mecanum wheel mobile robot, which is primarily responsible for driving the mobile robot movement, according to the upper control commands. The information acquisition module is composed of Kinect sensor and laser range finder. It is mainly responsible for sensing the surrounding environment, recognizing the moving target, detecting the obstacle and completing the obstacle avoidance. Where the Kinect sensor is used to obtain the image information of the surrounding environment, and the laser range finder is used for autonomous obstacle avoidance. The main hardware contents of this paper are as follows, Mecanum wheel mobile robot body, Kinect sensor and laser range finder. The Figure 1(b), shown the robot multi-mode movement, and the movement modes of the mobile robot are shown in the Table 1.



(a) Robot prototype: the mobile robot equipped with a Kinect sensor, a laser rangefinder sensor and a host computer

(b)Multi-mode movement: Mecanum wheel four-wheel mobile robot can achieve forward, reverse, left and right movement, 45 degrees oblique movement, 360 degrees all round zero radius and other actions

Figure 1.

Table 1. Multi-mode movement table

Operation mode	Motor 1	Motor 2	Motor 3	Motor 4
forward	+1	+1	+1	+1
backward	-1	-1	-1	-1
left straight	+1	-1	-1	+1
right straight	-1	+1	+1	-1
left anterior45^0	+1	O	O	+1
right back 45^0	-1	O	O	-1
right anterior 45^0	O	+1	+1	O
left back 45^0	O	-1	-1	O
clockwise rotation	+1	-1	+1	-1
Anticlockwise rotation	-1	+1	-1	+1
stop	O	O	O	O

Note. +1: Motor forward -1: Motor reversal
O: Motor stop Motor speed 0-2000 rev/min

This work uses the Kinect sensor developed by Microsoft, which has a color camera, a dot matrix infrared module transmitter, an infrared camera and two microphones. The Kinect prototype is shown in Figure 2, and the Kinect function parameters are shown in Table 2. RGB color camera is mainly used to collect the visual information of the surrounding environment, with 640 × 480 pixel optical resolution, horizontal field of view range of 57 °, vertical field of view range of 43°. The infrared emitter and infrared cameras are used to obtain the depth of field information, infrared emitter with three-dimensional depth of the laser speckle, the speckle random high, the formation of the pattern will vary with the distance. Infrared camera to capture these speckle patterns, and compared with the pre-saved pattern, the depth of the measured object information, the effective depth of detection range of 0.8m-4.0m.

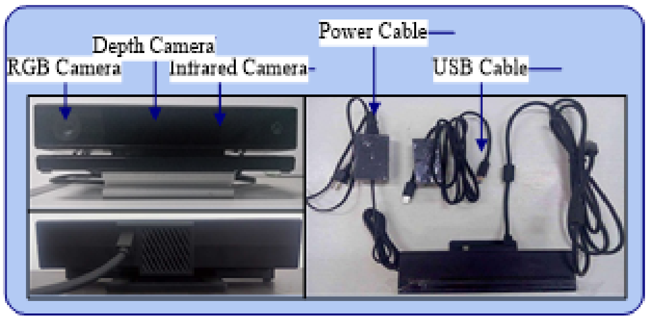


Figure 2. Kinect prototype

Table 2. Kinect function parameters

Important parameter	Data Range
Depth image	320×240
Color image	640×480
Perspective angle	vertical43°/Level57°
Frame rate	30 Frame/S
Audio	16KHz

In this work, UST-10LX laser rangefinder is produced by Hokuyo, which as a self-obstruction of the auxiliary sensor. The UST-10LX laser range finder consists of a laser, a laser detector and a measuring circuit, that measures the distance between the mobile robot and the obstacle by measuring the laser propagation time. The laser beam is emitted by the laser in a fixed time and a built-in counter starts counting. When the laser light reflected through the surface of the object is detected by the laser detector, the counter stops counting and finally calculates the robot and the obstacle by the counter value and the propagation speed of the laser. Mobile robot equipped with a laser range finder in real time to obtain polar coordinates within 270 ° range of 1080 laser point data information [11-12]. In order to obtain the distance information of the obstacle around the mobile robot, the laser data is linearly segmented, the laser data is divided into several regions and the position information of the moving robot is obtained from the obstacle. As shown in Figure 3 is the UST-10LX laser range finder of the physical figure and scanning range. At the same time, the laser range finder parameters are shown in Table 3.

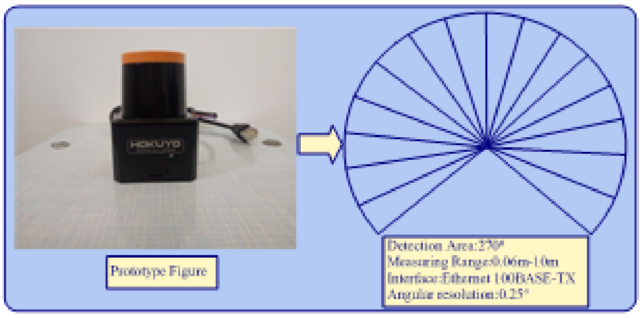


Figure 3. Laser range finder proto-type and scanning range

Table 3. Laser range finder parameters

Important parameter	Data Range
Supply Voltage	DC12V/24V
Detection Range	0.06-10m, Max: 30m, 270°(step0-1080)
Accuracy	±40mm
Scan Speed	25msec/scan
Angular Resolution	0.25°
Interface	Ethernet 100BASE-TX
Dimensions	50×50×70mm (W×D×H)
Weight	130g

2.2 Software Architecture

The main basis of the autonomous target recognition of the mobile robot is Machine Vision Based on the stereo object extraction target information, the autonomous avoidance function of the laser range finder, and the multi-sensor model is established. Finally, the weighted average method is used to complete the autonomous target recognition task. Mobile robot operating mechanism physical prototype is shown in Figure 4, and the mobile robot overall frame diagram is shown in Figure 5.

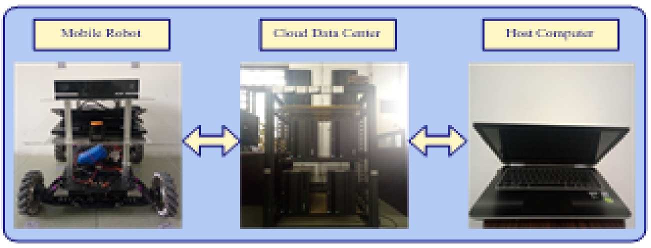


Figure 4. Mobile robot operating mechanism physical prototype

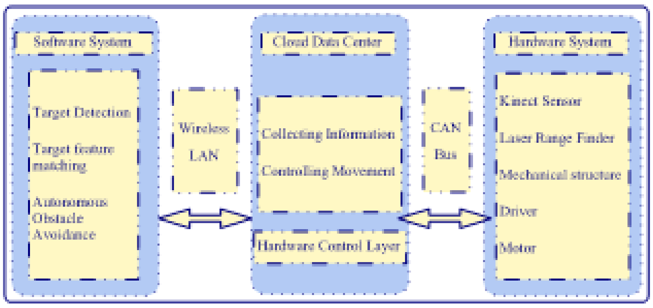


Figure 5. Mobile robot overall frame diagram

Mobile robot ontology system is mainly responsible for receiving cloud data through the center control commands, collecting sensor information and returns to the cloud data center [13]. The wireless network and cloud data center receives the control commands analyzed [14], and then the control command execution system underlying motion planning generates to control the motor according to the results of analysis [15]. In addition, it also needs to collect and package information of Kinect sensor, laser rangefinder sensor, and send them to the cloud data center through the wireless network to provide data support for the decision-making system [16]. The cloud data center system mainly provides the decision for the movement of the mobile robot body by the sensor data [17], and also provides human-computer interface, which is convenient for human-computer interaction. It consists of three modules: target detection module, target feature extraction module and autonomous obstacle avoidance module. The target detection module and feature extraction module for the surrounding environment of visual information through the visual sensor, using the Gauss background model and detection based on SURF feature matching and through RANSAC algorithm to remove the false matching points, and the specific target identification. Target information is transmitted to the target identification module frame [18]. The autonomous obstacle avoidance module is mainly responsible for ensuring that the mobile robot can walk safely in an unknown environment and avoid collisions with obstacles [19].

As shown in Figure 6 of the software system flow chart, each module is self-running and will not interfere with other parts. The upper and lower computers communicate information through the cloud data center. The host computer will be collected to determine the environmental information, analysis and processing to generate the corresponding mobile robot action instructions, and then send these instructions in real time to the cloud data center, the cloud data center will receive these instructions and stored in the form. The lower computer scans the information of the form every 100ms, and then judges that if the information is valid, the action function is called, and the mobile robot moves according to the instruction.

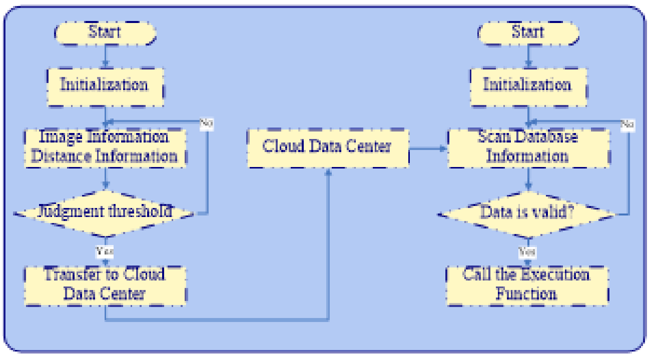
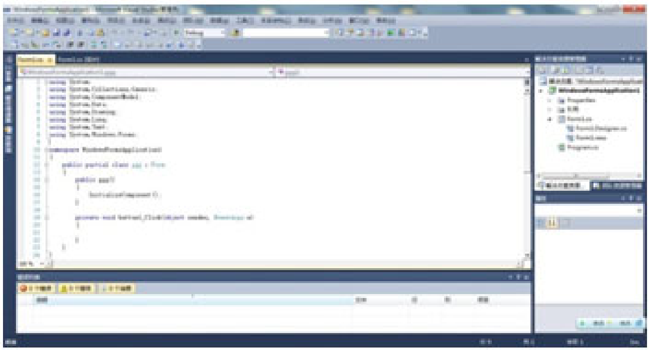
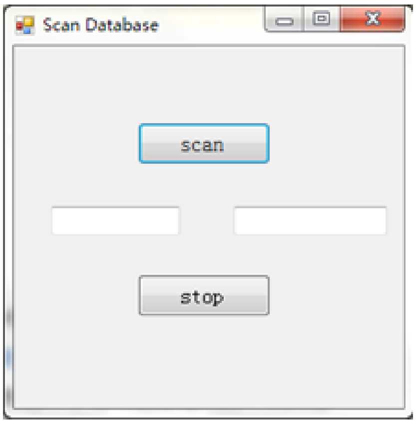


Figure 6. Software system flow chart

In this work, communication is made through a wireless transmission mode in the cloud data center to establish a database, through the wireless module to control the transmission of information stored in the cloud database [20]. It uses a SQL Server database, as shown in Figure 7 is the SQL Server development page [21]. Application of SQL Server database for data transmission is the principle of the host computer, cloud data center, the lower computer through the wireless module to connect to form a local area network. And then set up a form in SQL Server, used to store from the host computer side to send over the control information [22]. When sending data the host computer then scans the information to determine the decision-making control information sent to the cloud data center, the cloud data center will be stored after received, and then gives each data an address and number. The lower computer scans the database by address and number, thus receiving the latest control information. After receiving the information, it is necessary to judge the validity of the data. If the data is valid, the action execution function is called to perform the corresponding action, otherwise the scan will continue.



(a)SQL Server development interface



(b)The lower computer scans the database

Figure 7. Establish the SQL Server database table in the cloud data center to store the information. Robot scans the table every 100ms to perform the latest control information

3 Target Detection

Target detection methods are based on model-based detection, region-based detection, activity-based contour detection, and feature-based detection. For the recognition of moving objects, this paper extracts the foreground and background of relative motion from the image sequence containing the target motion information in the static background [23]. The background difference method can obtain the complete characteristic data, but it is sensitive to the change of illumination in the environment and the interference of the noise. The most important part is the extraction and updating of the background.

In this work, we use the Gaussian background model to achieve the background subtraction method. Gaussian model is a common method of extracting background which is suitable for the background static. The principle of background difference method of Gaussian model is to use the distribution of background model to study the degree of matching between pixel information and Gaussian background model, so as to classify the foreground and background, and use the prospect as the target. Each pixel in the image updates the Gaussian model by updating the Gaussian distribution parameters [24].

3.1 Parameter Update for the Gaussian Model

As the background of the scene is changing, in order to make the background model established by the Gaussian model adapt to the scene of the time change it is necessary to use the current frame image information of each pixel of the Gauss parameters update. It is necessary to use the current frame image information of each pixel of the Gauss parameters updated, the value of a current frame image of a pixel is represented by X_t , the mean value μ_t and the covariance $\sum_t$ of the pixel are updated by X_t , such as formula (1) and (2) shown [25]:

$$\mu_{t+1} = (1-\alpha)\mu_t + \alpha X_t . \quad \textbf{(1)}$$

$$\Sigma_{t+1} = (1-\alpha)\Sigma_t + \alpha d_t d_t^T . \quad \textbf{(2)}$$

The symbol α is the parameter update rate in the formula, generally the value of α is determined by the actual scene. If α takes a smaller value, it will cause the background model not to adapt to the changing scene. If α takes a larger value, the slow moving object will be updated to the background part of the model, which can cause the detected moving objects to appear empty or smearing.

3.2 Foreground Detection of Gaussian Model

Based on the Gaussian model, the gray value of each pixel on the background image, and the mean value of the Gaussian distribution in the background model of the scene corresponds to each other [26]. If the moving object corresponds to the gray value of the pixel it changes in the current frame image. The moving object then appears to enable the foreground detected by the equation (3):

$$\begin{cases} |X_t - \mu_t| \le T, background \\ |X_t - \mu_t| > T, foreground \end{cases} . \quad \textbf{(3)}$$

In the above equation, T is the threshold of the foreground segmentation.

Firstly, the Gauss smoothing filter is used to remove the noise. And then use the Gaussian modeling method to determine the background and foreground [27]. On the prospects of rot and incineration closed operation eliminates interference. Select the clump region using its area and threshold for screening, to identify the moving target object. The flow chart is shown in Figure 8.

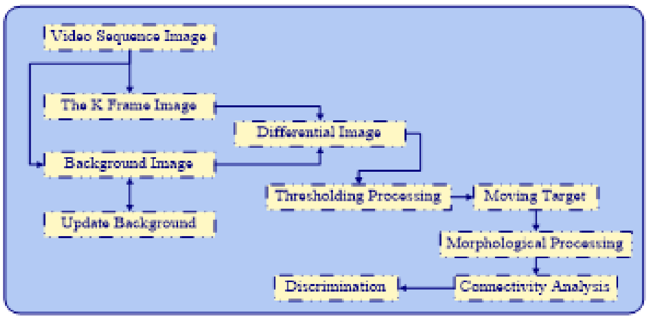
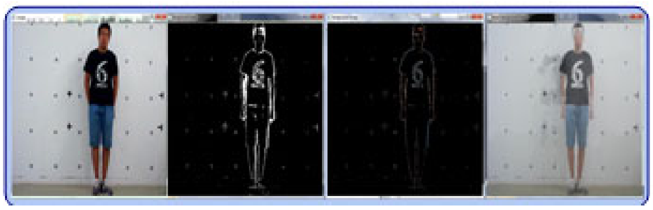


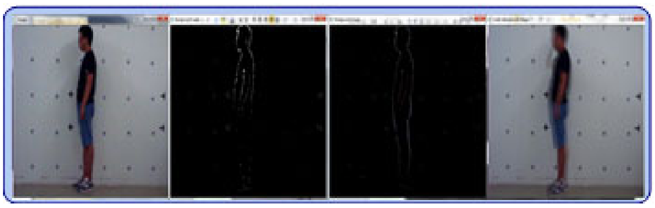
Figure 8. Gaussian model background difference method basic principle diagram

3.3 Comparison Experiments of Different Postures, Different Illumination and Different Speeds

By using the Gaussian background model subtraction method to achieve the moving target detection, compared with the three (3) groups above of the experiments respectively in different postures. Different illumination and different speeds of motion experiments were compared. Figure 9(a) and Figure 9(b) shows that the algorithm can still be well detected by moving the target from left to right. The image foreground mask, foreground image, and mean background under different illumination conditions. Figure 9(c) and Figure 9(a) compared with the different speed of movement can be drawn, when the movement speed is fast, the moving target will appear inside the noise and void phenomenon through morphological processing to reduce the cavity and improve anti-interference ability.



(a) (a) and (b) different postures contrast



(b) (b) and (c) different illumination contrast

(c) (a) and (c) different speeds contrast

Figure 9. Foreground detection of Gaussian model

4 Target Recognition Algorithm Based on Feature Matching

Target detection is to locate the object in the image by matching the specific object between the sample of the particular object, and the specific object in the input image to be detected, and identifying the specific target from the background image [28]. The feature recognition flow chart is shown in Figure 10.

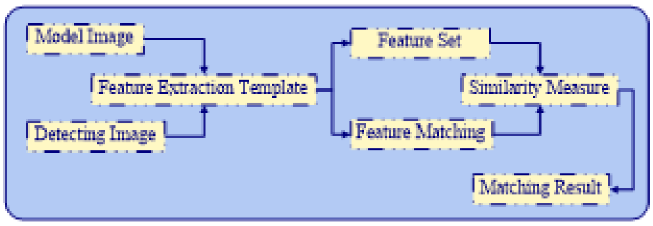


Figure 10. Feature recognition flow chart

4.1 Target Recognition Based on SURF Feature

The feature matching algorithm is used to search the target in the image based on the characteristics of the target, so as to realize the recognition of the target. The SURF target matching feature is a local feature in the image, with scale invariance and rotation invariance, which has good adaptability to scaling small viewing angle change, noise and brightness change. There are the three steps for Target recognition based on SURF matching tracking method: (1) SURF feature point detection; (2) generating feature description vector; (3) feature point matching.

The above steps are calculated as follows:

4.1.1 SURF Feature Point Detection

The full name of SURF is Speed Up Robust Features [29], which was put forward by Bay Herbert in 2006. The SURF algorithm can extract the feature of the template image and the acquired image information, and obtain the matching result in real time. The SURF algorithm has high computational speed and good robustness and is widely used in object recognition. The core of the algorithm is the Hessian matrix determinant which mainly includes the following steps: constructing the pyramid scale space, locating the feature points, selecting the main direction of the feature points, and constructing the feature point descriptors. The algorithm uses the box filter to approximate the Gaussian kernel function and filters the integral image of the original image, which greatly improves the processing speed. When constructing the feature point descriptor each feature point is a 64-dimensional vector which is much faster than the SIFT [30] (Scale-invariant feature transformation) algorithm.

Feature point detection is the first step of SURF matching. The purpose of feature point detection is to find candidate nodes with obvious features in scale space. The main process is as follows:

Constructing scale space. Using the step enlarged approximate Gauss filter to do convolution operation on the integral image of input image so as to form the scale space of input image. The construction of SURF scale space is not based on the gradual reduction of the image, but the operation of the same image with different size filters to ensure the preservation of the high frequency information of the image. After the scale space of the image is established, the fast Hessian matrix is used to detect the extreme points on each image [31].

For any point in space (x, y), the scale σ and the Hessian matrix are defined as:

$$H(X,\sigma)=\begin{bmatrix} L_{xx}(X,\sigma) & L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma) & L_{yy}(X,\sigma) \end{bmatrix}. \quad \textbf{(4)}$$

Where $L_{xx}(X,\sigma)$ is the Gaussian filter second derivative $\frac{\partial^2 g(\sigma)}{\partial x^2}$ and the input image convolution, that is, if $g(\sigma)$ is a Gaussian function, L_{yy} and L_{xy} are similar to L_{xx}. In order to facilitate the calculation, here with the box filter template and the input image convolution D_{xx}, D_{yy}, D_{xy} instead of L_{xx}, L_{yy}, L_{xy}. The 9 * 9 box filter is similar to σ = 1.2 second-order Gaussian derivative, D_{xx}, D_{xy} instead of L_{xx}, L_{xy} between the following relationship:

$$\omega=\frac{\|L_{xy}(1.2)\|_F \|D_{xx}(9)\|_F}{\|L_{xx}(1.2)\|_F \|D_{xy}(9)\|_F}\approx 0.9. \quad \textbf{(5)}$$

In the formula, $\|\bullet\|_F$ is the Frobenius norm, ω is the weight coefficient, and 0.9 is used in practice. Then the determinant of the Hessian matrix can be expressed as:

$$\det(H)=D_{xx}D_{yy}-(0.9D_{xy})^2. \quad \textbf{(6)}$$

Extreme point detection. Comparison of adjacent to each pixel and the pixel scale space in the same layer of 8 pixels, and two layers of 9 adjacent pixels in a total of 26 adjacent pixels, get the local maximum and minimum points. Then the use of 3 dimensional quadratic Taylor equation expansion of surface fitting, so as to realize the precise positioning of feature points, the coordinates of the feature points (x, y), the scale σ, will be able to determine the main direction of the feature point neighborhood information and feature vector. The results of SURF feature point detection are shown in Figure 11.

Figure 11. SURF feature point detection

4.1.2 SURF Feature Point Description Generation

The SURF feature point descriptor is calculated based on the local information of the surrounding (neighbor) of the feature point, including the main direction and the feature description vector. The calculation process is as follows:

Determination of main direction of SURF feature point [32]:

Firstly, the Haar wavelet response d_x and d_y in the x and y directions are obtained for all the pixels in the circular region with the feature point, as the center and the radius 6σ (σ is the scale of the feature point), so that each pixel has a corresponding Haar wavelet response point $Hp(dx,dy)$. Among them, Haar wavelet response is obtained by using the corresponding Haar wavelet filter and integral image convolution obtained. Then all the wavelet responses are summed by a fan-shaped sliding window of size $\pi/3$, and the longest direction is taken as the main direction of the feature points. In the process of summation, the wavelet response of each pixel is weighted by the Gaussian function centered on the feature point. In this step, the fan-shaped sliding window size is the key parameter that determines whether the correct main direction can be obtained.

Descriptor Generation Based on Haar Wavelet Response:

The feature vector extraction of the SURF feature point is carried out in a square region centered on the feature point, and parallel to the main direction. First determine a feature point as the center, the size of 20σ square area. In order to ensure that the extracted feature vector has rotational invariance, it is necessary to rotate the square region, so that it is parallel to the main direction of the feature point. Then the square region is subdivided into 4x4 sub-regions, and the sum of the Haar wavelet responses in the X and Y directions and the sum of the absolute values of the subdivision are counted in each sub-region: $\sum d_x$, $\sum d_y$, $\sum|d_x|$, $\sum|d_y|$. In the process of statistics, still with the feature point as the center of the Gaussian function for weight processing. So that each sub-region has a 4-dimensional descriptor, $v_4 = (\sum d_x, \sum d_y, \sum|d_x|, \sum|d_y|)$, the entire region has 4x4x4 = 64-dimensional eigenvector. The results of SURF feature point description generation are shown in Figure 12.

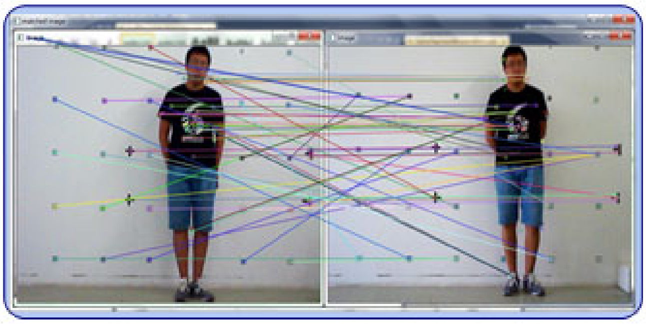


Figure 12. SURF feature point description generation

4.1.3 SURF Feature Points Match

After using the SURF method to obtain the reference image and the feature points of the image to be registered, the commonly used feature matching method is: similarity measure method. In the above-mentioned information about the feature points (position, scale, main direction and eigenvector), the eigenvector contains the information of the neighborhood of the feature point, and the nearest neighbor matching method of the vector can find the potential matching pair to calculate the amount of additional information. The nearest neighbor matching method is to calculate the distance between all the eigenvectors, and the vector in the image to be searched for a certain eigenvector, and then find the ratio of the nearest neighbor and the nearest neighbor. If the ratio is less than the preset threshold, nearest neighbor is a better match. The threshold chosen in this work is 0.65, which has a high matching accuracy rate.

In addition, the SURF algorithm adds a Laplacian identifier to each feature point when constructing a feature descriptor. It is defined by the value of the trace of the Hessian matrix. If the trace of the Hessian matrix is greater than 0, the identifier is denoted by 1, otherwise -1. The identifier is used to distinguish between two different types of feature points on bright backgrounds and dark spots on bright backgrounds. The same type of feature points can be used as matches. The target object recognition step based on SURF feature matching is as follows:

- Extract the SURF feature points of two images;
- Calculate the descriptive vector for each holding point;
- Calculate the vector distance of the feature points in the two images, and find the matching feature points for each feature point in the other graph;
- Find the two mapping points in the single mapping transformation and visualize the results.

As shown in Figure 13 is the target object identification flow chart. The results of SURF feature points match are shown in Figure 14.

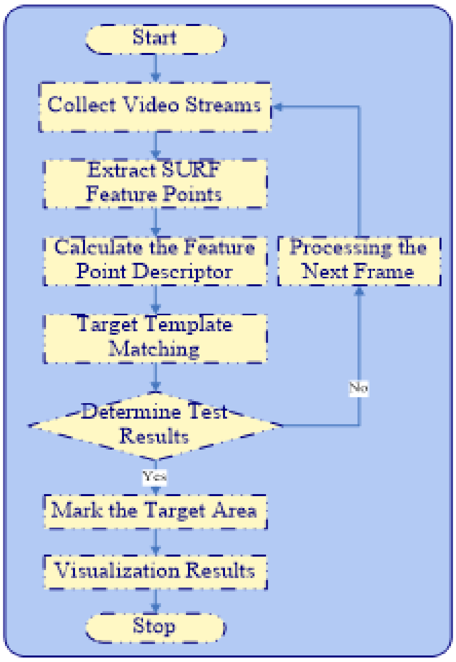


Figure 13. Target object identification flow chart

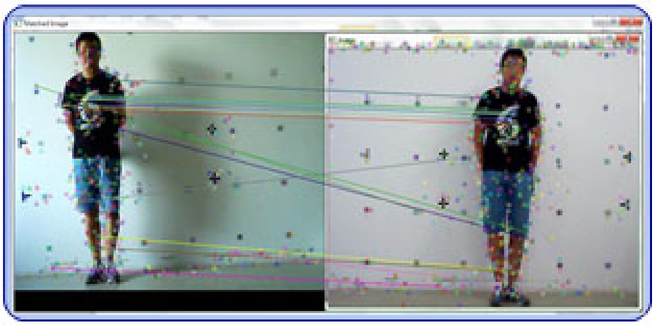


Figure 14. SURF feature points match

4.2 RANSAC Algorithm Optimization of Feature Matching

Based on the SURF feature matching, it is feasible to identify the target object, but there is a mismatch between the template image and the video acquisition image, which will reduce the detection accuracy. RANSAC algorithm is a robust algorithm for estimating the mathematical model by iterative method [33]. In this work, the RANSAC algorithm is used to improve the accuracy of retrieval. The sample data to which it is applied generally contains abnormal data, which is estimated by repeatedly selecting a set of random subsets in the data, checking with the remaining points, and removing the anomaly data. If there are enough points to match the estimated model, then the estimated model is justified. It is possible to compensate for the shortcomings of the general model estimation method and to eliminate the influence of the anomaly data on the model estimation. RANSAC algorithm optimization feature matching is essentially through the continuous cycle, seeking a model containing the most matching points, so that the rest of the mismatch can be removed. Figure 15 is a flow chart optimized for RANSAC feature matching.

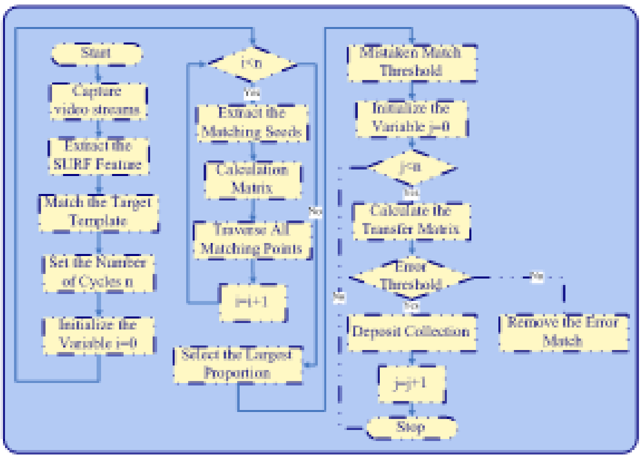
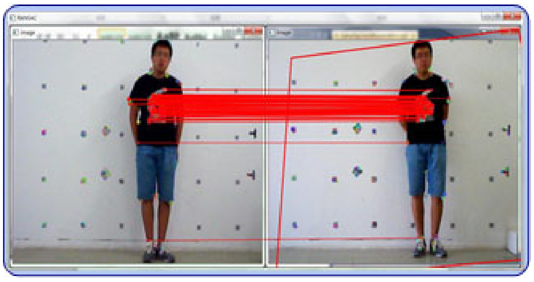


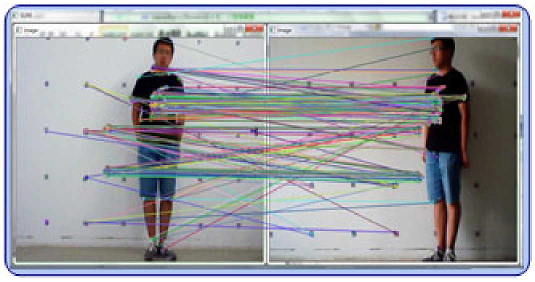
Figure 15. RANSAC algorithm optimization flow chart

According to the above ideas for the target object recognition experiment. The experiment is carried out in two steps. Figure 16(a) and Figure 16(c) shows the rough matching experiment of the SURF feature. Figure 16(b) and Figure 16(d) shows the effect of using the RANSAC algorithm to remove the mismatch. From the two experiments can be intuitive to see, into the RANSAC thinking, the matching effect has been greatly improved, good to improve the accuracy of the target object recognition. Table 4 is an optimized comparison table based on the above experiment. It can be seen from this table that there will be a large number of false matches in SURF rough matching due to the influence of factors such as light, target scale transformation and background complexity. On this basis, the RANSAC algorithm can eliminate most of the mismatched points in the rough matching. Although the method is mistakenly removed by a small part of the correct match point, but the matching effect has been greatly improved.

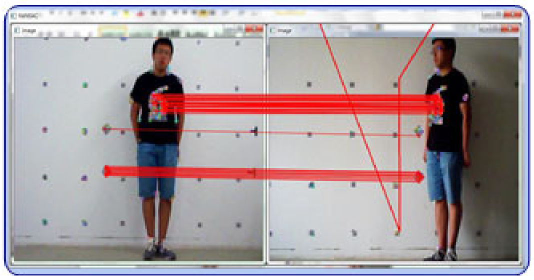
(a) (a) and (c); RANSAC matching point (b) and (d)



(b) (a) and (c); RANSAC matching point (b) and (d)



(c) (a) and (c); RANSAC matching point (b) and (d)



(d) (a) and (c); RANSAC matching point (b) and (d)

Figure 16. SURF matching point

Table 4. Comparison of RANSAC on SURF Algorithm Optimization

Experiment content	SURF matching point pairs	Error matching point pairs	Percentage of error matching points
target object feature detection	242	46	19.01%
Experiment content	RANSAC matching point pairs	Error matching point pairs	Percentage of error matching points
target object feature detection	106	4	3.78%

5 Self-obstacle Avoidance

Figure 17 shows the schematic diagram of obstacle detection. Mobile robot platform is 600mm × 450mm × 600mm, movement speed control at 0.2m/s. Considering the moving speed of the robot and the data transmission time of the system, in order to prevent the collision between the robot and the obstacle, the longitudinal safety distance and the lateral safety distance should be set for the robot. We set a longitudinal threshold of 450mm and a lateral threshold of 500mm. The laser range finder collects the distance and time stamp of the robot from the obstacle at the moment, and the full scan time (0-270^0 scanning range) is 25ms, and the collected data is uploaded to the cloud data center. It is installed in front of the second side of the mobile robot and is mounted at a height of 30cm from the ground. The robot speed is set to 0.2m/s. Considering the speed of the system, the laser data is divided into several areas without affecting the scanning accuracy. Finally, the position information of the mobile robot from the obstacle is uploaded to the cloud data center. The calculation process is as follows. The scanning period is 25ms, the minimum distance in the database, the known oblique length and the angle, and the trigonometric function are used to calculate the lateral distance d1, and the longitudinal distance d2 of the obstacle from the robot platform. The size of the mobile platform and maintain a safe distance, as shown in Figure 18, when d1 is greater than 500mm, call the forward control function, otherwise, to determine d2 less than or equal to 450mm, call the back function. When d2 is greater than 450mm is less than or equal to 1100mm, and then determine the obstacle and the robot moving platform in front of the angle θ, when θ is less than or equal to 90 ° counterclockwise turn, otherwise, turn clockwise. When d2 is greater than 1100mm, moveforward.

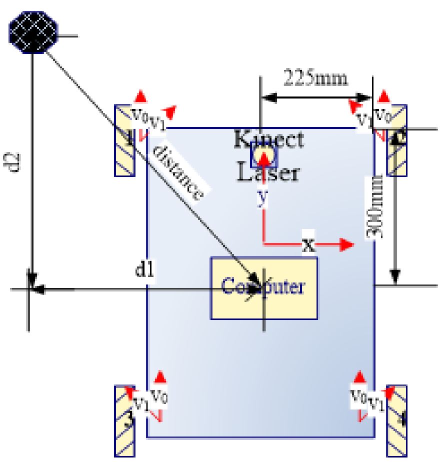


Figure 17. Schematic diagram of obstacle detection

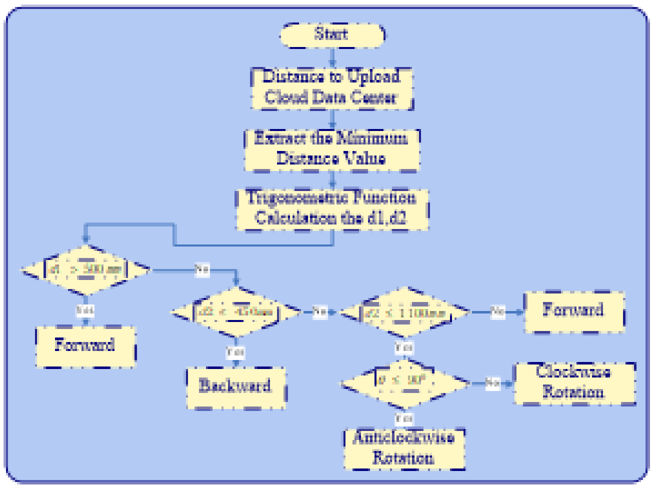


Figure 18. Obstacle detection flow chart

Experimental results: 0.5s reaction speed, with the basic ability of independent obstacle avoidance decision-making, to achieve forward and backward and other multi-mode movement. According to the above experiment can be seen based on the laser range finder independent obstacle effect is more ideal, but there is a drawback. As a result of the laser range finder scanning is a horizontal limit, when the robot encountered some low obstacles, may not be detected, resulting in cannot avoid these obstacles.

Based on the above ideas, this work designs a multi-sensor fusion system which integrates the stereo vision sensor and the laser rangefinder sensor. The comprehensive experiment proves that the mobile robot can realize the target recognition and obstacle avoidance function better.

6 Sensor Data Transfer Experiment

Data table is made up of four fields namely: step, data, timestamp, time.

With step, data, time stamp data format for the 32-bit integer, time data format is varchar (50). The total amount of data transfer is 14029488 bytes.

Experiment using the database system for the microsoft SQL server2012, using 300M wireless LAN transmission [34]. In this experiment, the time cost of the complete process from the collected data to the complete data is calculated based on the time of the computer system. The following two kinds of transmission methods are designed to compare the experiment. The experimental results show that the data transmission time is small.

Method 1. the database operation package open, data write, close into the class, in other functions call these methods. That is, each write a data should be opened separately, close the database once.

Method 2. the database to open, and close the statement written to the function, and write the data written to the function, the formation of bulk operation. All data is written only open, close the database once.

Experimental design: use two transmission methods each transmission is 1081 data, a total of ten transmission, and record the cost of each transmission time. And then calculate the mean and variance of the two methods respectively. The results are as Table 5, Table 6 and Table 7.

Table 5. Sensor data transfer time: s

number	method 1	method 2
1	1.257	1.161
2	1.171	1.157
3	1.154	1.177
4	1.160	1.127
5	1.173	1.153
6	1.181	1.173
7	1.178	1.165
8	1.176	1.164
9	1.178	1.139
10	1.192	1.186

Table 6. Sensor data transfer stability statistics

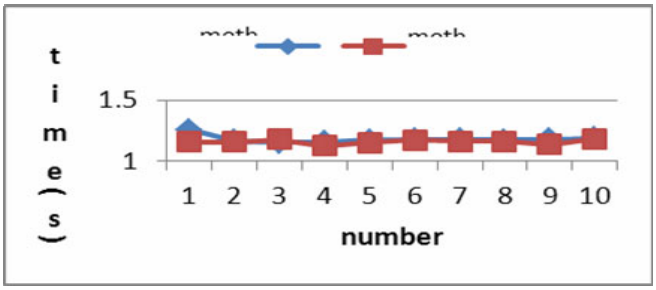


Table 7. The mean and variance of the time consumption of ten data transmission using two transmission methods are calculated as follows

	mean	variance
method1	1.182	0.007
method2	1.160	0.003

As shown in the table above, it is that the method 2 is faster and more stable. So in the next experiment we have adopted the second method for data transmission.

7 Conclusions

Through the above experiments, the target detection module can detect the target object by using the Gaussian model background difference method in the case of different lighting conditions and different moving target speed. The target feature extraction module adopts the SURF (Speed Up Robust Features) algorithm based on feature matching. By detecting the SURF feature point, the feature point description vector and feature point matching are generated to detect the feature points that describe the target object. In this paper, the RANSAC algorithm is used to optimize the SURF algorithm, and the mismatched feature points in the SURF algorithm are filtered out. The experimental results show that the percentage of error matching points of the SURF algorithm is 19.01%, the percentage of error matching points of the RANSAC algorithm is 3.78%, the RANSAC algorithm can eliminate most of the errors under the influence of light, target scale transformation and background complexity. At the same time, filtered out the mismatched feature points, a small part of the correct match point also removed, but the overall feature point matching effect has been greatly improved. The autonomous obstacle avoidance module can achieve the good obstacle avoidance effect by setting the longitudinal safety distance to 450mm and setting the crosswise safety distance to 500mm, and it can complement the shortcomings of the monocular vision sensor. The establishment of cloud data center, connected to the host computer and the lower computer between the mobile robot control commands and sensor data storage implementation, through the data transmission experiment designed a reasonable and effective transmission method.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant No. 61473027), and Beijing Key Laboratory of Robot Bionics and Function Research (GrantNo. BZ033 7), and Supported by The Fundamental Research Funds for Beijing Universities, and Beijing Advanced Innovation Center for Intelligent Robots and Systems, and Beijing Advanced Innovation Center for Future Visual Entertainment.

References

[1] X. Zhang, H. Wang, W. Cheng, Vessel Detection and Classification Fusing Radar and Vision Data, *2017 Seventh International Conference on Information Science and Technology*, Da Nang, Vietnam, 2017, pp. 474-479.

[2] J. Woo, J. Lee, N. Kim, Obstacle Avoidance and Target Search of An Autonomous Surface Vehicle for 2016 Maritime RobotX Challenge, *2017 IEEE Underwater Technology*, Busan, South Korea, 2017, pp. 1-5.

[3] M. Mancini, G. Costante, P. Valigi, T. A. Ciarfuglia, J. Delmerico, D. Scaramuzza, Toward Domain Independence for Learning-Based Monocular Depth Estimation, *IEEE Robotics and Automation Letters*, Vol. 2, No. 3, pp. 1778-1785, July, 2017.

[4] Z. Tang, F. Law, A. Thurlbeck, C. West, M. Baxter, A. Inayat, Design of a Mobile Robot Prototype for a Relay Race, *2016 31st Youth Academic Annual Conference of Chinese Association of Automation*, Wuhan, China, 2016, pp. 105-109.

[5] S. Kahlouche, N. Ouadah, M. Belhocine, M. Boukandoura, Human Pose Recognition and Tracking using RGB-D Camera, *2016 8th International Conference on Modelling, Identification and Control*, Algiers, Algeria, 2016, pp. 520-525.

[6] C. Chen, Y. Liu, C. Lyu, W. Zhou, J. Peng, X. Jiang, P. Li, Real-time Target Tracking and Positioning on FPGA, *2016 IEEE International Conference on Real-time Computing and Robotics*, Angkor Wat, Cambodia, 2016, pp. 448-453.

[7] M. M. Ali, H. Liu, N. Stoll, K. Thurow, Intelligent Arm Manipulation System in Life Science Labs Using H20 Mobile Robot and Kinect Sensor, *2016 IEEE 8th International Conference on Intelligent Systems*, Sofia, Bulgaria, 2016, pp. 382-387.

[8] S. Jose, A. Antony, Mobile Robot Remote Path Planning and Motion Control in a Maze Environment, *2016 IEEE International Conference on Engineering and Technology*, Coimbatore, India, 2016, pp. 207-209.

[9] S. Sun, N. An, X. Zhao, M. Tan, Human Recognition for Following Robots with a Kinect Sensor, *2016 IEEE International Conference on Robotics and Biomimetics*, Qingdao, China, 2016, pp. 1331-1336.

[10] X. Du, Y. He, L. Chen, S. Gao, Pose Estimation of Large Non-cooperative Spacecraft based on Extended PnP Model, *2016 IEEE International Conference on Robotics and Biomimetics*, Qingdao, China, 2016, pp. 413-418.

[11] C. Zhuang, H. Zhou, S. Sakane, Learning by Showing: An End-to-end Imitation Leaning Approach for Robot Action Recognition and Generation, *2016 IEEE International Conference on Robotics and Biomimetics*, Qingdao, China, 2016, pp. 173-178.

[12] Z. Shu, G. Liu, Q. Xie, Z. Ren, A Method of Urine Detection Based on Front Vision and Image Recognition, *2016 9th International Congress on Image and Signal Processing*, BioMedical Engineering and Informatics, Datong, China, 2016, pp. 402-405.

[13] A. Kundu, L. Luan, R. Liu, Synchronisation of Data Transfer in Cloud, *International Journal of Internet Protocol Technology*, Vol. 8, No. 1, pp. 1-24, May, 2014.

[14] C.-C. Tseng, K.-C. Ting, H.-C. Wang, F.-C. Kuo, L.-H. Chang, Construction and Analysis of a Green Clustered Architecture for RNG-based Wireless Ad Hoc Networks, The International *Journal of Ad Hoc and Ubiquitous Computing*, Vol. 19, No. 1/2, pp. 62-74, May, 2015.

[15] Y. Liu, X. Liu, Y. Shi, F. Wu, J. Zhang, A Robust Method of Mark Recognition for Robot Autonomous Navigation, *2016 International Conference on Audio, Language and Image Processing*, Shanghai, China, 2016, pp. 711-715.

[16] N. Pandeeswari, R. Karuppathal, Hypervisor Based Anomaly Detection System in Cloud Computing Using ANFIS, *Journal of Internet Technology*, Vol. 18, No. 6, pp. 1335-1344, November, 2017.

[17] S. Namasudra, P. Roy, Time Saving Protocol for Data Accessing in Cloud Computing, *IET Communications*, Vol. 11, No. 10, pp. 1558-1565, July, 2017.

[18] C. Li, X. Wang, Visual Localization and Object Tracking for the NAO Robot in Dynamic Environment, *2016 IEEE International Conference on Information and Automation*, Ningbo, China, 2016, pp. 1044-1049.

[19] C. Potthast, G. S. Sukhatme, Online Trajectory Optimization to Improve Object Recognition, *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Daejeon, South Korea, 2016, pp. 4765-4772.

[20] W. Wu, N. Xiong, C. Wu, Improved Clustering Algorithm based on Energy Consumption in Wireless Sensor Networks, *IET Networks*, Vol. 6, No. 3, pp. 47-53, May, 2017.

[21] V. Vasco, A. Glover, C. Bartolozzi, Fast Event-based Harris Corner Detection Exploiting the Advantages of Event-driven Cameras, *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Daejeon, South Korea, 2016, pp. 4144-4149.

[22] M. Panicker, T. Mitha, K. Oak, A. M. Deshpande, C. Ganguly, Multisensor Data Fusion for An Autonomous Ground Vehicle, *2016 Conference on Advances in Signal Processing*, Pune, India, 2016, pp. 507-512.

[23] S. Ricardo, D. Bein, A. Panagadan, Low-cost, Real-time Obstacle Avoidance for Mobile Robots, *2017 IEEE 7th Annual Computing and Communication Workshop and Conference*, Las Vegas, NV, 2017, pp. 1-7.

[24] L. Wang, A. D. Ames, M. Egerstedt, Safety Barrier Certificates for Collisions-Free Multirobot Systems, *IEEE Transactions on Robotics*, Vol. 33, No. 3, pp. 661-674, June, 2017.

[25] M. M. Ali, H. Liu, N. Stoll, K. Thurow, An Identification and Localization Approach of Different Labware for Mobile Robot Transportation in Life science Laboratories, *2016 IEEE 17th International Symposium on Computational Intelligence and Informatics*, Budapest, Hungary, 2016, pp. 000353-000358.

[26] T. Dewi, N. Uchiyama, S. Sano, Service Mobile Robot Control for Tracking a Moving Object with Collision Avoidance, *2015 IEEE International Workshop on Advanced Robotics and its Social Impacts*, Lyon, France, 2015, pp. 1-6.

[27] Q. Zhong, J. Zhao, C. Tong, Tracking for Humanoid Robot Based on Kinect, *2014 International Conference on Mechatronics and Control*, Jinzhou, China, 2014, pp. 1191-1194.

[28] C. Premachandra, Y. Okamoto, K. Kato, High Performance Embedding Environment for Reacting Suddenly Appeared Road Obstacles, *2014 IEEE International Conference on Robotics and Biomimetics*, Bali, Indonesia, 2014, pp. 2394-2397.

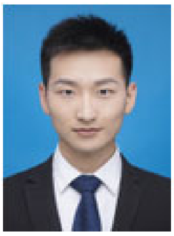
[29] G. Yasuda, Behavior-based Autonomous Cooperative Control of Intelligent Mobile Robot Systems with Embedded Petri Nets, *2014 Joint 7th International Conference on Soft Computing and Intelligent Systems and 15th International Symposium on Advanced Intelligent Systems*, Kitakyushu, Japan, 2014, pp. 1085-1090.

[30] N. Abdelkrim, K. Issam, K. Lyes, C. Khaoula, Fuzzy Logic Controllers for Mobile Robot Navigation in Unknown Environment Using Kinect Sensor, *IWSSIP 2014 Proceedings*, Dubrovnik, Croatia, 2014, pp. 75-78.

[31] R. Lin, M. Li, L. Sun, Real-time Objects Recognition and Obstacles Avoidance for Mobile Robot, *2013 IEEE International Conference on Robotics and Biomimetics*, Shenzhen, China, 2013, pp. 1157-1162.
[32] Y. Diskin, B. Nair, A. Braun, S. Duning, V. K. Asari, Vision-based Navigation System for Obstacle Avoidance in Complex Environments, *2013 IEEE Applied Imagery Pattern Recognition Workshop*, Washington, DC, 2013, pp. 1-8.
[33] G. Hoang, H. K. Kim, S. B. Kim, Control of Omini-directional Mobile Vehicle for Obstacle Avoidance using Potential Function Method, *2013 9th Asian Control Conference*, Istanbul, Turkey, 2013, pp. 1-6.
[34] L. Zhang, K. Zhang, An Interactive Control System for Mobile Robot Based on Cloud Services, *Concurrency and Computation Practice and Experiences*, Vol. 30, No. 24, e4983, October, 2018. https://doi.org/10.1002/cpe.4983

Biographies

Lei Zhang received his doctorate degree in BIT, Beijing, China in 2007. He is currently professor of School of Electrical and Information Engineering, BUCEA. And he is currently deputy director of Beijing Key Laboratory of Robot Bionics and Function Research. He has been working in the field of robot motion control and artificial intelligence for nearly 20 years. His main research interests are humanoid robot, human-machine interaction and machine vision.



Dezhong Li is a control science and engineering graduate student from BUCEA. His research interests currently lie in the human-machine interaction in human settlements, machine learning and the application of robot operating system in mobile robots.