Proceedings of 2021 IEEE
International Conference on Mechatronics and Automation
August 8 - 11, Takamatsu, Japan

A Novel Target Tracking System for the Amphibious Robot based on Improved Camshift Algorithm

Shuxiang Guo[1,2], Handong Cheng[1] | Jian Guo[1*] | and Jigang Xu[3*]

[1]Tianjin Key Laboratory for Control Theory&Applications
In Complicated systems and Intelligent Robot Laboratory
Tianjin University of Technology
Binshui Xidao Extension
391,Tianjin,300384,China
guo@eng.kagawa-u.ac.jp; 497210593@qq.com

[2] Department of Intelligent Mechanical Systems Engineering,
Faculty of Engineering
Kagawa University
Takamatsu,Kagawa,Japan
*corresponding author : jianguo@tjut.edu.cn

[3]Unit68709
Qinghai Haidong,810700,China
*corresponding author : xujigang216@163.com

***Abstract* –In the past, there are some problems in target tracking, such as poor effect and easy loss. This paper propose a target recognition system based on improved Camshift algorithm, which is mainly used in marine rescue, marine garbage search, aquaculture surface target tracking and other fields. To the flexible application of amphibious robot in sea, land and other fields, RGB-D camera was equipped on the amphibious robot for target recognition and tracking. We have done the research of target recognition in the early stage. In this paper, CAMSHIFT and Kalman filter were used together. Kalman filter predicted the moving position of the target in the next frame and fed it back to the previous frame.Because the judgment of the edge node can greatly shorten the communication time with the terminal, so as to improve the efficiency of target tracking. Experiments show that the amphibious robot can accurately identify and track the target, and the delay is controlled within one second**

Index Terms – Camshift, Kalman filter, target tracking.

I. INTRODUCTION

With the development of tourism and aquaculture. With the development of coastal tourism, there is a risk of maritime distress. China's maritime safety awareness is relatively weak, and the number of ships and people in distress is on the rise. The maritime rescue ship has become one of the effective means for coastal countries to ensure the safety of human maritime activities, so it is urgent to develop an autonomous target recognition and tracking system to assist the safety personnel in rescue.

Zhang's team uses templates to track moving objects, which is to track objects by calculating the appearance similarity between consecutive frames.In the literature, Kalman filter is used to estimate the image distribution, and target trajectory is used to predict.The literature uses the characteristics of particle filter to track the target in the image. Combined with the features of mean square histogram and its color model, the literature uses Meanshift algorithm to track, and then searches for potential targets in the next image [1]-[4].

Voles et al first collected the video, smoothed the image after the acquisition, then extracted the image using two frame difference method, and finally detected the moving object. But the problem of this method is that the detection is not accurate, which is easy to cause the phenomenon of missing detection, and it is not suitable for the complex scene of sponge. J.G.Sanderson In this paper, a template matching algorithm is used to deal with the waveform. The frequency domain characteristics of the wave are used as the template to match the sea surface motion state. Then the matching results are analyzed. The moving targets are distinguished by the analysis results, and then the moving targets are tracked by the target estimation method. The operation speed of this method is high, and the tracking accuracy can basically meet the requirements, However, the database samples of this method are limited, so it can not track a wide range of common targets, and can not detect, track and recognize different types of targets. Nasim et al. Proposed the method of using edge detection combined with morphological processing to detect and track moving objects. Firstly, the image collected from the video is processed by dilation and erosion, followed by morphological processing such as open close operation, and then the background difference method and Sobel algorithm are used Edge detection method detects moving objects, but the time and delay of object detection is high, which can not reflect the real-time performance [5]-[8].

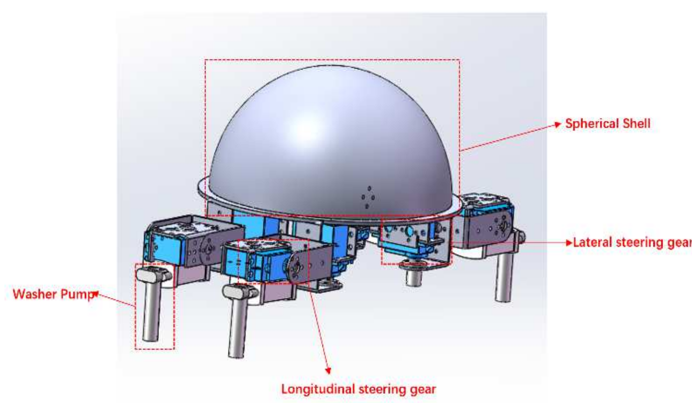


Fig.1 The platform of Amphibious Robot.

II. The Overview of Amphibious Robot Platform

The overall structure of the amphibious machine is shown in Fig.1. The experimental platform is a spherical amphibious robot, which is a kind of bionic robot. It can not only walk on land, but also complete horizontal movement and water rotation in the air.

Its areas of work include land and sea. The upper half of the platform adopts hemispherical design, in which the control element of the robot is placed, the lower half of the platform adopts the movement mode of bionic quadruped, the waterproof steering gear and water sprayer are placed in the lower half of the platform, the main controller controls the drive controller to control the steering of the waterproof steering gear, and the experimental platform adopts 16 way steering gear control board to control the steering gear[9]-[10].

In the complex land, the inner part of the spherical robot will call the action group of each steering gear in the lower part of the spherical body to make the horizontal steering gear and the longitudinal steering gear move to complete the whole action. When underwater, the controller controls four water spraying motors to complete the water spraying action by connecting the control board, so that it can move back and forth on the water surface and underwater, making the spherical underwater robot platform to complete amphibious switching and movement. In addition, the floating level switch is designed to switch the land and underwater action groups, so as to complete the target recognition and tracking of complex terrain [11]-[18].

The bottom diagram of star amphibious robot is shown in Fig.2, This experimental platform adopts the design of sealed connector to ensure the waterproof. Raspberry pi4 is used as the control board of edge target recognition in the spherical amphibious robot shell, and STM32F407 is used to control the steering gear and water spraying motor After the target recognition, After the target is locked, the target will be tracked, and then the signal will be continuously transmitted to the control board,the IIC communication interface will be used to transmit the signal to STM32F407 to complete the control of the steering gear and water jet motor [19]-[25].

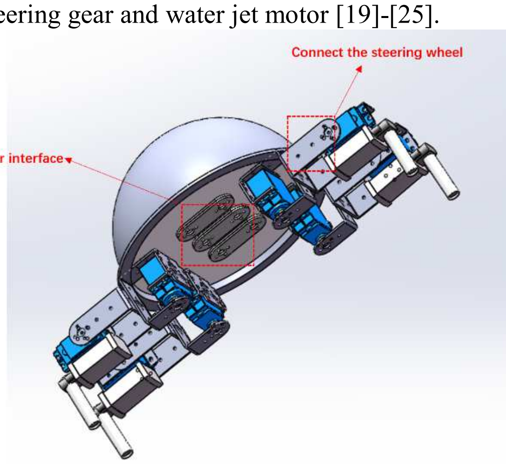


Fig.2 The Schematic diagram of amphibious robot platform bottom.

III. Study of Camshit Algorithm and Construction of Kalman Filter

A. Study of Camshit Algorithm

In this paper, Camshift algorithm is based on Meanshift, which is a continuous adaptive mean shift algorithm, called continuous adaptive mean shift algorithm. Because Meanshift algorithm has very obvious shortcomings, it is because of Meanshift The adaptive window size of the algorithm can't change with the moving target. If the moving target has the change of distance and shape, it will lead to the failure of target tracking. But Camshift algorithm can change this shortcoming very well. The algorithm can change the adaptive window according to the change of target size and distance, and can accurately locate and determine the center of the target position.

Camshift algorithm running demonstration is shown in Fig.3, The image shows the movement of the two search boxes. Once the initial position is determined, the initial color distribution will be recorded, and the center particle will be determined to complete the movement [26].

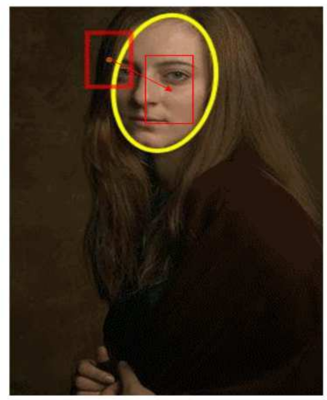


Fig.3 Camshift algorithm running demonstration.

Camshift algorithm is a dynamic feedback process. Firstly, the RGB color space of the image is transformed into the HSV color space distribution. After the transformation, the H component of the HSV color space is used to establish the color histogram, and then the back projection is carried out. The purpose is to obtain the two-dimensional discrete probability color distribution map of the moving target. In the calculation window, the zero order matrix and the pair X are used, First order matrix of Y, get the first-order matrix of X, y, use the first-order matrix of X,Y to compare with the zero order matrix of the above window according to the data, then get the ratio, finally get the centroid position of the target, when the centroid position moves, repeat the above algorithm to search the target[27].

The Zero order matrix is:

$$M_{00}^{i} = \sum_{x_i} \sum_{y_i} I(x_i, y_i) \tag{1}$$

The first order matrix of X:

$$M_{10}^{i} = \sum_{x_i} \sum_{y_i} I(x_i, y_i) \tag{2}$$

The first order matrix of Y:

$$M_{01}^{i} = \sum_{x_i} \sum_{y_i} I(x_i, y_i) \tag{3}$$

It is necessary to solve the target centroid formula:

$$x_{iC} = \frac{M_{10}^{i}}{M_{00}^{i}}, y_{iC} = \frac{M_{01}^{i}}{M_{00}^{i}} \tag{4}$$

Where $I(x_i, y_i)$ represents the pixel value with (x_i, y_i) coordinate in the image, and the change range of (x_i, y_i) is the size of the search window S_i.

Then the window will be updated. First, a threshold value ∈ will be set, and then the target centroid position and the centroid position in the color histogram model will be set. If the distance between them is greater than ∈, we will determine that the centroid position is not correct, and we will make a new comparison and judgment. Then the new search window s is solved according to formula 5. The size of S_i, we take the range of pixel value is 0-255, The flow chart of target tracking is shown in Fig.4 [28].

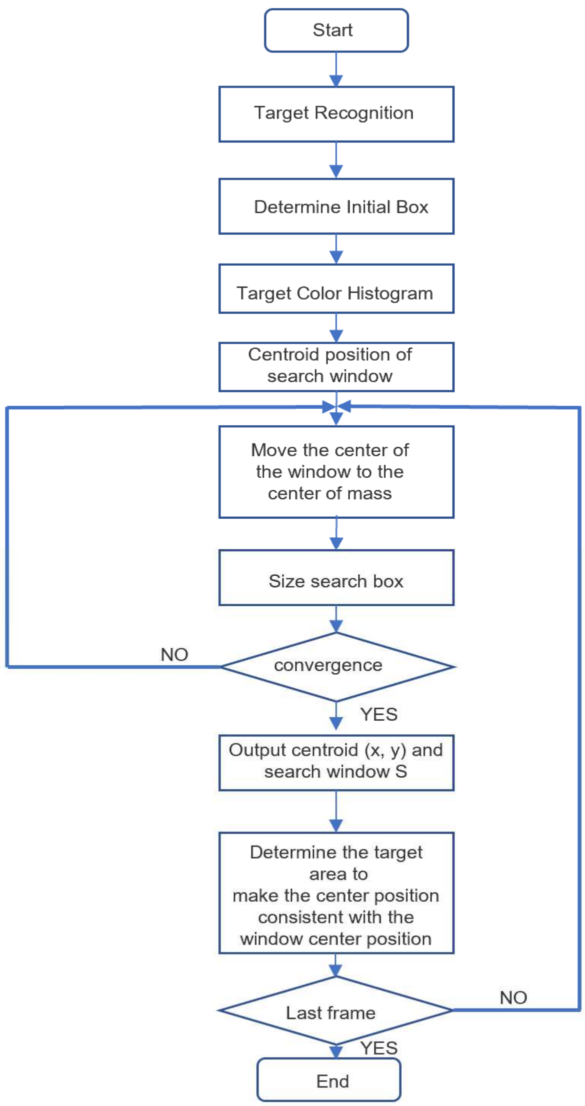


Fig.4 Flow chart of target tracking.

The Camshift algorithm is based on the color information of the target in the video to track the target. In addition, Camshift can continuously adjust the size of the search window according to the scale change of the specific target, so each frame of the video adopts the Meanshift algorithm to find the iterative result. We add the Kalman filter to the Camshift algorithm to predict the position The position of the previous frame is predicted by the movement of one frame.

B. Construction of Kalman Filter

The Kalman filter is based on the minimum mean square error as the best estimation criterion, assuming that the process equation of linear dynamic discrete system is:

$$X(k) = AX(k-1) + BU(k) + W(k) \tag{5}$$

In addition, the measured value of the system is:

$$Z(k) = HX(k) + V(k) \tag{6}$$

The process model of the system is used to predict the current state of the system :

$$X(k \mid k-1) = AX(k-1 \mid k-1) + BU(k) \tag{7}$$

Where the covariance of $X(k \mid k-1)$ is:

$$P(k \mid k-1) = AP(k-1 \mid k-1)A' + Q \tag{8}$$

Then we update the current time estimate, which is :

$$X(k \mid k) = X(k \mid k-1) + K_g(k)(Z(k) - HX(k \mid k-1)) \tag{9}$$

Then we get the Kalman gain :

$$K_g(k) = P(k \mid k-1)H'/(HP(k \mid k-1)H' + R) \tag{10}$$

Where the Kalman gain of $K_g(k)$.Finally, in order to continuously track a specific target, we update the covariance of $X(k \mid k-1)$ in K state :

$$P(k \mid k) = \big(\boldsymbol{I} - K_g(k)H\big)P(k \mid k - \boldsymbol{I}) \tag{11}$$

In order to reduce the computational complexity of the algorithm, we can fuse the information of a specific target, and then use Kalman filter to predict the location area of a specific target in the next image, so as to achieve the purpose of strengthening the target tracking.

We add the Kalman filter, so that Camshift can effectively solve the problem of target deformation and occlusion, and track the target more accurately. The system resource requirements are not high, and it is more suitable for our embedded platform. When we do not add the Kalman filter, due to the low time complexity, we can achieve good tracking effect in a simple background, but when we are in a simple environment In the case of complex background, or the interference of many pixels with similar color to the target, the tracking will fail. Because it only considers the color histogram and ignores the spatial distribution characteristics of the target, we need to add the prediction algorithm of the tracking target, that is, our Kalman filter.

First, the neural network will recognize the target, then determine the target, and then read the video sequence, select the specific target, determine the search window and initialize the relevant parameters of Kalman filter; use the Kalman filter to predict the position of the specific target in the next frame; according to the color characteristics of the specific target, use Camshift The algorithm counts the histogram of the search window and obtains the color probability distribution map by back projection; each pixel in the search window of the current frame is compared with the pixel values of its surrounding points as the input image of Camshift algorithm, so as to determine the candidate target information of the current frame; in the next frame, the information of the target in the previous step is updated by Kalman filter Filter, determine the search window, return to the previous row, continuous tracking target.

We add a computing node to the amphibious robot for neural network recognition. An edge computing board is added inside the amphibious robot, which is mainly used for image recognition. The result of target recognition by raspberry is shown in Fig.5. The information collected by our RGB-D camera is transmitted to the edge computing board. The edge computing board uses raspberry to perform neural network operation to complete the edge calculation, and then transmits the information of the first confirmation box to stm32f4 through USB communication. Then STM32F4 will carry out motion decision processing on the collected target information.

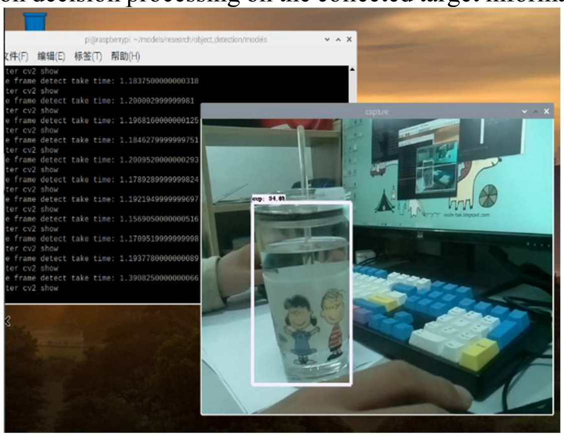


Fig.5 The results of target recognition by raspberry.

In this experiment, raspberry uses raspberry for gateway conFig.uration, and then raspberry uses WiFi for remote connection with the computer to transmit the target recognition signal to the display in real time, which is convenient for viewing.

IV. EXPERIMENTS AND RESULTS ANALYSIS

Before target tracking, we need to recognize the target. First, we train the neural network. In this paper, we use the framework of Yolo-tiny to train. After 3000 falls, loss tends to be stable. Fig.6 shows the training process.

This paper uses Python language and open source code opencv joint programming. The video captured by the camera installed on the amphibious spherical robot is used as the test sequence. The neural network is used to identify the target, and then the target in the video image is tracked. The results show that the improved algorithm has high accuracy and robustness in target tracking. In this paper, several groups of video images are used in OT100 test set to verify the performance of the proposed algorithm under the condition of fast motion and similar color background.

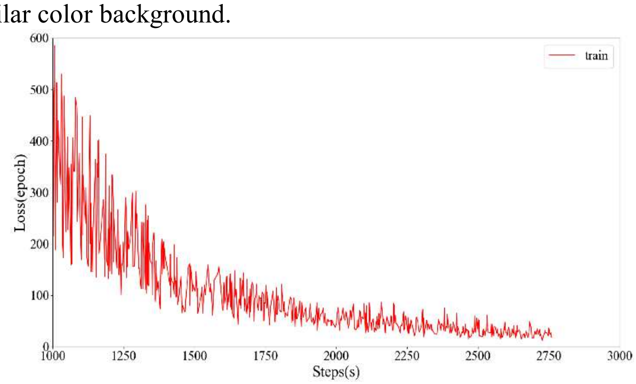


Fig.6 The the training process

The color of the target in the image is similar to that of the surrounding environment, but the tracking is not affected by the interference. The target is well tracked, and there is no loss of tracking target.

The Fig.7 shows result of target recognition by Separate Camshift algorithm, The Fig.8 shows result of target recognition by Camshift algorithm with Kalman filter.

(a) (b) (c)

Fig.7 The result of target recognition by Separate Camshift algorithm.

(a) (b) (c)

Fig.8 The result of target recognition by Camshift algorithm with Kalman filter.

The traditional Camshift algorithm is used to track the target, and the surrounding occlusion interferes with the target tracking. It can be seen that the rectangular frame of the tracking target gradually becomes larger, and the target is lost. This is because the traditional Camshift algorithm takes the color feature as the main basis of target tracking, so the tracking effect is poor when there is color interference. Using the improved tracking algorithm, the color of the target in the video image is similar to the surrounding environment, but the tracking is not affected by the interference, the target is well tracked, and there is no loss of tracking target.

The experimental data processing is mainly to discuss the coordinates, regardless of the size of the tracking box. Generally speaking, the ordinate of its curve is the precision, and the abscissa is the threshold. Setting different thresholds, we can get different precision, and then according to Matlab, we can draw a curve. The higher the accuracy is, the better it is. Generally, the threshold is set at 20 pixels. Here we mainly focus on 0-20 pixels, that is, a small threshold means better tracking when the distance is closer, The Fig.9 shows difference of abscissa of center point. It means the pixel difference between the abscissa of the center point of the original tracking object and the center point of the tracking frame.

The Fig.10 shows difference of ordinate of center point. It means the pixel difference of the ordinate between the center point of the original tracking object and the center point of the tracking frame.

Then, this paper carries on the entity experiment, when controlling the amphibious spherical robot movement, the movement range of the robot should not be too large, otherwise it is easy to make the target out of the screen range, causing false detection, because Camshift algorithm also has a return value when there is no real target. Therefore, the movement interval of the robot should not be too large or too small. After repeated experiments and debugging, the robot can carry out tracking experiments.

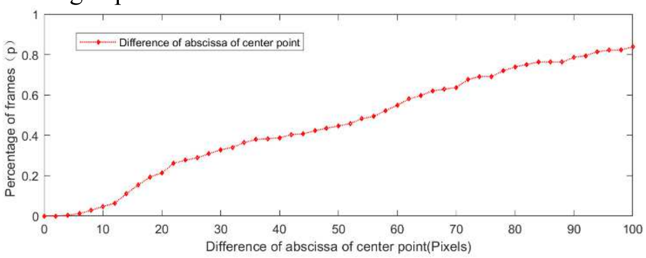


Fig.9 The difference of abscissa of center point.

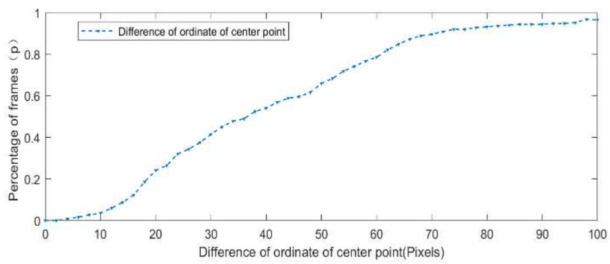
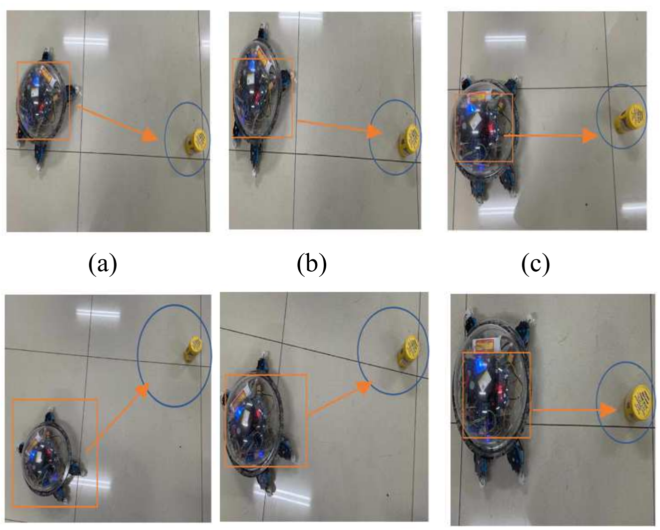


Fig.10 The difference of ordinate of center point .



(a) (b) (c)

(c) (d) (e)

Fig.11 The result of target recognition by raspberry.

This paper uses raspberry to install opencv environment to program the Camshift algorithm, uses Python environment language, and builds the tensorflow environment, so that the target recognition before target tracking is accurate, and then target tracking is carried out. When raspberry detects the target, the target distance is determined by the size of the target box in the whole screen, and then the target box is used to track the target When the target frame is on the left or right side of the screen, raspberry transmits a signal to STM32 to control the steering gear to rotate and adjust the position, so that the target frame is in the center of the screen to complete the drive.

In order to view the effect more conveniently, as shown in Fig.10, we use the rectangular box to label the two ball spherical robot. At the same time, we use the circular box to label the target object. After the verification of experiments, we carry out the comparison of multiple groups of experiments, which proves the feasibility of the target recognition algorithm.

Fig.11 shows the experimental results. Where a, b and c represent the robot's tracking experiment in the right front of the target. Where d, e and f represent the robot's tracking experiment in the left front of the target.

According to the experiment of the above pictures, we can see the credibility of the target recognition algorithm, and the Kalman filter plays a key role. The driver has been repeatedly debugged, so that the robot can run effectively.

V. Conclusions and Future Work

This paper presents a target tracking system based on Camshift algorithm. To improve the accuracy and robustness of traditional Camshift algorithm in target tracking, an improved Camshift algorithm combined with Kalman filter was fused. The algorithm tracks the target features and solved the influence of the surrounding color.The Kalman filter was used to estimate the position of a specific target, which solves the problem of target tracking in marine autonomous rescue and target loss in aquaculture water target tracking. The experimental results showed that the improved algorithm achieves the desired effect in target tracking, had high accuracy and robustness, and could be widely used. In the future, we will continue to improve the accuracy of the algorithm and reduce the recognition delay, and carry out repeated verification and comparative experiments.

Acknowledgments

This research is supported by National Natural Science Foundation of China (61703305), Key Research Program of the Natural Science Foundation of Tianjin (18JCZDJC38500) and Innovative Cooperation Project of Tianjin Scientific and Technological (18PTZWHZ00090).

Reference

[1] Y.Zhang, Z Deng, et al. “Sea Surface Target Detection based on Complex ARMA-GARCH Processes,” *Digital Signal Processing*, vol. 22, No. 2,pp.1–13,2017.

[2] S.Fefilatyev, D.Goldgof et al, “Detection and Tracking of Ships in Open Sea with Rapidly Moving Buoy-mounted Camera, ” *Ocean Engineering* , vol. 20, No.3,pp.1-12,2012.

[3] G.Vivone, P.Braca, et al. “Knowledge-Based Multi-Target Ship Tracking for HF Surface Wave Radar Systems, ” *IEEE Transactions on Geoscience andRemote Sensing* , Vol.53,No.7,pp.3931-3949, 2015.

[4] Li Xiao bo, Sun Wen fang, Li Li. “Ocean Moving Ship Detection Method for Remote Sensing Satellite in Geostationary Orbit, ” *Journal of Electronics &Information Technology*, Vol.37,No.8,pp.1862-1867,2015.

[5] P.Voles et al. “Nautical Scene Segmentation Using VariableSize Image Windows and Feature Space Reclustering, ” *Computer Vision -ECCV*, Vol.39,No.7,pp. 324-335,2000.

[6] M.Murshed et al. “Moving Edge Segment Matching for the Detection of Moving Object, ” *Image Analysis and Recognition-8th International Conference*, Vol.402,No.8,pp.22-24,2011.

[7] J.Meltzer, R.Gupta, et al. "Multiple View Feature Descriptors from Image Sequences ViaKernel Principal Component Analysis, " *Computer Vision ECCV* , Vol.40,No.12,pp. 215-227,2008.

[8] N.Arshad, et al. "An Adaptive Moving Ship Detection and Tracking Based on Edge Information and Morphological Operations, " *International Conference on Graphic and Image Processing* , vol. 20, No.3,pp.8285-8290,2011

[9] Jian Guo, Shuxiang Guo, et al, "Design and CharacteristicEvaluation of a Novel Amphibious Spherical Robot, " *Microsystem Technologies*, Vol.23, No.6, pp.1-14,2017.

[10] Huiming Xing, Shuxiang Guo et al , "Hybrid Locomotion Evaluation for a Novel Amphibious Spherical Robot", *Applied Science*s, Vol.8, No.2, pp.1-8,2016.

[11] S.Guo,Y.He, et al. "Modeling and experimental evaluation of an improved amphibious robot with compact structure, " *Robotics and Computer-Integrated Manufacturing*, vol. 20, No.3,pp,51:37-52,2018.

[12] Shuxiang Guo, Yanlin He et al, "Modal and fatigue analysis of critical components of an amphibious spherical robot, " *Microsystem Technologies*, Vol.23, No.6, pp.1-15, 2016.

[13] Xihuan Hou, Shuxiang Guo, et al, "Hydrodynamic Analysis-Based Modeling and Experimental Verification of a NewWater-Jet Thruster for an Amphibious Spherical Robot, " *Sensors*, Vol.19, No.259, pp.1-15,2016.

[14] Huiming Xing, Shuxiang Guo, et al, "Hybrid Locomotion Evaluation for a Novel Amphibious Spherical Robot, " *Applied Sciences*, Vol.8, No.2, pp.1-15, 2018.

[15] Shuxiang Guo, Shaowu Pan et al, "A system on chip-based real-time tracking system for amphibious spherical robots, " *International Journal of Advanced Robotic Systems*, Vol.14, No.4, pp.1-9,2018.

[16] X.Yao,et al, "A Novel Independent RNN Approach to Classification of Seizures against Non-seizures, " *arXiv* , vol. 20, No.3,pp.1903,2015.

[17] Wildes R P, Asmuth J C, Green G L, et al, "A machine-vision system for iris recognition , " *Machine Vision & Applications* , vol.50, No.3,1996,pp.1-8.

[18] W.Liu,et al, "DSSD : Deconvolutional Single Shot Detector, " *IEEE Conference on Computer Vision and PatternRecognition (CVPR)*. vol. 20, No.3 ,pp.1-8,2017.

[19] , S.Li, et al, "Visual model-predictive localization for computationally efficient autonomous racing of a 72-gram drone, " *arXiv* , vol. 22, No.4,pp.1–13,2019.

[20] Namiki; Matsushita, S.; Ozeki, T.; Nonam`i, K , et al, "Hierarchical processing architecture for an air-hockey robot system, " *Robotics and Automation, Karlsruhe*, Vol.24, No.6, pp. 1187–1192,2013.

[21] Shuxiang Guo, Yanlin He, et al "Modal and fatigue analysis of critical components of an amphibious spherical robot, " *Microsystem Technologies* , Vol.23,No.6, pp.1-15,2016.

[22] Maoxun Li, Shuxiang Guo, et al, "Designand performance evaluation of amphibious spherical robot, " *Roboticsand Autonomous Systems*, Vol.64, No.6, pp.21-34, 2018.

[23] Liang Zheng, Shuxiang Guo, et al, "The communication and stability evaluation of amphibious spherical robots, " *Microsystem Technologies*, Vol.24, No.6, pp.1-12,2016.

[24] Jian Guo, Shuxiang Guo et al, Li L. "Design and characteristic evaluation of a novel amphibious spherical robot, " *Microsystem Technologies*, Vol.23, No.6, pp.1999-2012, 2017.

[25] Yanlin He , Shuxiang Guo et al, "Preliminary Mechanical Analysis of an Improved Amphibious Spherical Father Robot, " *Microsystem Technologies*, Vol.22, No.8, pp.2051-2066,2016.

[26] D.Pathak，P.Krahenbuhl,et al, "Context encoders:Feature learning by inpainting, "*2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*，Vol.25, No.8, pp.2536-2544,2016.

[27] B.Dolhansky et al, "Eye in-painting with exemplar generative adversarial networks, "*Conference on Computer Vision and Pattern Recognition*, vol. 20, No.3,pp.7902 － 7911,2018.

[28] S.Iizuka et al, "Globally and locally consistent image completion," *ACM Transactions on Graphics*, vol. 36, No.4, pp. 1-14, 2017.