Maritime Target Detection and Tracking

Liu Jian*, 1, 2, a, Gao Wen2, b

1Changchun University of Science and Technology, China
2Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, China

aaky1_liujian@sina.com, bbcwenzi@126.com

Keywords: object tracking, object detection, line detection

Abstract: Method of object detection and tracking is one of hot topics in image processing and computer vision field, which is significant not only in military such as imaging guidance and military target tracking, but also for civil use such as security and monitoring and the intelligent man-machine interaction. For marine targets, how to eliminate interference effectively and achieve accurate detection of moving targets is a key problem to be solved. In this paper, a method of image segmentation and line detection is proposed to detect the horizon, which can help to determine the area to be detected. Prior knowledge is used to detect the ship target in this area. To identify the suspected ship, we detect it in multiple frames. After identifying the ship, the mean shift algorithm is applied to lock tracked target. The experiment shows that the algorithm we proposed has good effect of detection and tracking for the ship target in the infrared image, and this algorithm has low computational complexity and small computation. It has been realized in the embedded hardware system and has high application value.

1. Introduction

China is a major maritime country with more than 18,000 kilometers of coastline, more than 6,000 islands and 3 million square kilometers of sea area under its jurisdiction. The sea has provided a steady flow of resources for China's economic development. At the same time, the pollution and destruction of marine environment, the over-exploitation of marine are becoming increasingly prominent and severe. The environmental protection and management problems of marine China facing have become more complex and arduous, especially with the increasing frequency of marine economic activities and the increasing competition among countries for marine resources in recent years.

Thus, with all kinds of maritime activities, for the marine sustainable development, and to protect and manage marine environment effectively, as well as to make full use of Marine resources, it’s important to carry out the research of maritime target detection and tracking [1-3].

For different data acquisition mode, the current detection methods for maritime moving targets are generally based on the following four types: satellite remote sensing images, radar images, infrared images and video images [4][5][16]. Most of the traditional detection methods focus on satellite remote sensing images and radar images. The former has good performance on shape detection and ground penetration. The latter has strong detection capability and accurate distance calculation. So both of them has been widely studied and applied. However, they all take high operation cost[6-8].
Meanwhile, they all can’t provide the real-time, all-weather or visual monitoring. Infrared camera can work continuously at night or under insufficient light, but it’s image resolution is low and the power consumption is too high [9][10].

What is different from the above three is that the video detection technology has good adaptability, good real-time, high spatial resolution and low cost. It can provide dynamical monitoring of the waters around promptly so you can know site conditions at any time[11]. As the decrease of camera equipment cost and the matures of video processing technology, maritime moving targets detection based on video technology is gradually applied in ocean environmental protection, marine monitoring, marine rights and protection, and so on. It has effectively made up for the inadequacy of traditional detection methods.

In the traditional video monitoring system, the maritime target recognition and behavior analysis is accomplished by monitoring personnel. But it is neither practical nor reliable for a person to perform this kind of boring routine monitoring for a long time. So there is an imperative requirement of intelligent target detection and identification technology. At present, the intelligent video monitoring system is mainly used in the land environment. It is a challenging task to realize the detection of moving objects in the complex sea surface environment. Due to the light, the wind and waves, the sea surface is always changing with a lot of relevant noise, including light of ripples and irregularity waves. At the same time, the mobile cloud, broken foam of waves inshore and swaying branches also make a lot of interference to target detection [12]. Therefore, how to effectively eliminate interference and realize accurate detection of moving objects is a key problem to be solved.

In this paper, a method of image segmentation and line detection is proposed to detect the horizon, which can help to determine the area to be detected. We use prior knowledge to detect the ship target in this area. To identify the suspected ship, we detect it in multiple frames. After identifying the ship, the mean shift algorithm is used to lock tracked target.

2. Target Detection

This section covers image segmentation and line detection to detect the horizon as well as the target detection after determining the detection area.

2.1 Image Segmentation based on the OTSU Method

Suppose the original brightness level is L. The number of pixels with grayscale i is ni. The total number of pixels in the image is N. So the probability of grayscale i is shown in Equation (1).

\[
P_i = \frac{n_i}{N}
\]  

(1)

According to the grayscale threshold T, the image grayscale is divided into \(C_0\) and \(C_1\), where \(C_0\) is composed of grayscale less than T, and \(C_1\) is composed of grayscale greater than T. So the probability of \(C_0\) and \(C_1\) occurring can be expressed as Equation (2).

\[
w_0 = \sum_{i=0}^{T} P_i, \quad w_1 = \sum_{i=T+1}^{L-1} P_i
\]  

(2)

Where \(w_1 = 1 - w_0\). The average gray level of \(C_0\) and \(C_1\) is shown in Equation (3) and (4).
\[ u_0 = \frac{\sum_{i=0}^{T} iP_i}{w_0} = u(T) \]  

\[ u_1 = \frac{\sum_{i=T+1}^{T+L} iP_i}{w_1} = \frac{\sum_{i=T}^{T+L} iP_i - \sum_{i=0}^{T} iP_i}{1-w_0} = \frac{u - u(T)}{1-w_0} \]  

where \( u = \sum_{i=0}^{T} iP_i \), \( u(T) = \sum_{i=0}^{T} iP_{Ti} \). The variance of \( C_0 \) and \( C_1 \) are shown in Equation (5) and (6).

\[ \sigma_0^2 = \sum_{i=0}^{T} \frac{(i-u_0)^2}{w_0} \]  

\[ \sigma_1^2 = \sum_{i=T+1}^{T+L} \frac{(i-u_1)^2}{w_1} \]  

Define intra-class \( \sigma_w^2 \) and inter-class variance \( \sigma_B^2 \), shown in Equation (7) and (8).

\[ \sigma_w^2 = w_0 * \sigma_0^2 + w_1 * \sigma_1^2 \]  

\[ \sigma_B^2 = w_0 * (u_0 - u)^2 + w_1 * (u_1 - u)^2 \]  

When the judgment criterion is the maximum, the corresponding threshold is the optimal threshold.

The method of maximum inter-class variance is widely used because it does not require preset parameters and has strong adaptability.

### 2.2 Line Detection based on Hoff Transform

Hoff transform is a method to detect simple geometric shapes such as straight lines and circles in image processing. It was originally used for line detection in binary images. For a line in the image, the cartesian coordinate system can be expressed as \( y = kx + b \). So any point on this line \((x,y)\) becomes a point when it goes to the k-b parameter space.

\[ y = x \cos \theta + y \sin \theta \]
In other words, when all non-zero pixels in the image space are converted to the k-b parameter space, they will be focused on one point. In this way, a local peak point in the parameter space is likely to correspond to a line in the original image space. However, since the slope of the line may be infinite or infinitesimal, it is not easy to describe the line in the k-b parameter space.

Therefore, the researchers proposed the use of polar coordinate parameter space for line detection. In polar coordinates, a line can be expressed in the following form: \( \rho = x \cos \theta + y \sin \theta \), shown in Figure 1.

For angle \( \theta \in [0, 360^\circ] \), \( \rho \) can be set in a certain range by the resolution of the image, and the line can be obtained.

For classification of binary image, start from the zero point, set the angle of horizon between 150° to 210°, traverse angle resolution for 2°, and obtain the line whose length, which is greater than the eighth width. Connect similar lines by line position and angle. Finally, the location of the horizon is determined. Then, the target detection was carried out in the area 100 pixels high above the sea level and 50 pixels high below.

### 2.3 Ship Target Detection

After the target area is determined, different types of ships are found by infrared images. Although the stealth performance to infrared is different, there's one area which is the brightest. According to this feature, highlight area detection is carried out in the area to be inspected. Whether the highlighted area is a ship target can be determined by limiting the size of the highlighted area and judge whether there is regular movement within the neighborhood of 5 consecutive frames. When the ship target is determined, the target tracking is carried out.
3. Target Tracking

Target tracking problem is actually a problem of state estimation when it is abstracted to a mathematical problem. So the tracking problem can be divided into several processes, such as the establishment of state space, the evaluation and correction of observation model, and so on. So these processes are conducted to describe real-time tracking algorithm of variable resolution kernel histogram of particle swarm acceleration.

3.1 Establishment of State Space

In this section, the target state is discussed. The best and most robust tracking systems are the visual processing parts of the human eye and brain, so designing human-like visual processing systems is a worthwhile undertaking.

3.1.1 Eye Attention Mechanism

From careful study to the human eye perception of things, it can be found that the images seen by the human eye are different from those in the real world. The human eye makes a clearer image of the object it cares about but a blur around. This is determined by the distribution of retinal vertebral cells in the human eye. Therefore, we propose to establish a target model similar to retina that is centered on the target.
3.1.2 Kernel-histogram

In view of the multi-scale nature of human eyes, the target template is established by using the variable-scale kernel histogram. The introduction of kernel function into gray histogram results in the introduction of spatial position factor into gray statistic feature. The closer from benchmark \( \delta(x) = \begin{cases} 0 & x \neq 0 \\ 1 & x = 0 \end{cases} \), the more important sample point \( x_i \) to estimate the statistical properties of \( x \) are.

Through statistical histogram, the probability of each grayscale in the target area is obtained. With the use of nuclear weighted histogram, the probability of grey level of the central region is increased by giving higher weight to the central region. In this way, the edge weight is small. When the edge changes and is shielded, it has little impact on the whole. So it improves the robustness of the occlusion deformation. The kernel weighted histogram can be expressed in Equation (9).

\[
q_u = C \sum_{i=1}^{n} K(X_i^*, H) \delta(b(X_i^*) - u)
\]

(9)

Where \( q_u \) represents the nuclear histogram, \( K(X_i^*, H) \) is the kernel function. \( u = 1, 2, \ldots, m \) denotes grayscale, the maximum value for \( m \). \( X_i = (x_i, y_i) \) is the coordinate of \( i \), the central pixel coordinates is \( X_c = (x_c, y_c) \), the distance from \( X_i \) to \( X_c \) is \( X_i^* = (x_i - x_c, y_i - y_c) \). \( b(X_i^*) \) represents the gray value of pixel \( i \) at \( X_i^* \). \( H \) is the bandwidth of the kernel function, called the scaling factor.

\[
C = \frac{1}{\sum_{j=1}^{n} K(X_j^*, H)}
\]

is the normalization factor, making \( \sum_{u=1}^{m} q_u = 1 \) correct. \( \delta(x) = \begin{cases} 0 & x \neq 0 \\ 1 & x = 0 \end{cases} \) is the function of Kronecker delta.

Several common kernel functions [4] are unit gaussian kernel function and Epanechnikov [5] kernel function, shown in Equation (2) and (3).

\[
K \left( \frac{X_i^*}{H} \right) = \exp \left( -\frac{x_i^*}{h_x} - \frac{y_i^*}{h_y} \right)
\]

(10)

\[
K \left( \frac{X_i^*}{H} \right) = 1 - \left( \frac{x_i^*}{h_x} \right)^2 - \left( \frac{y_i^*}{h_y} \right)^2
\]

(11)

Although all of the above kernel functions are applicable to target tracking, the unit gaussian function drops too fast, while the Epanechnikov kernel function needs to calculate a highly complex exponent calculation.

(a) Distribution of vertebral body membrane cells in the optic network  
(b) Epanechnikov kernel function
Therefore, this paper uses the pyramid surface equation instead of the type of elliptic paraboloid equation, using planar approximate surface, without changing the nature of kernel that function is monotone decreasing at the same time. This make dense information more outstanding.

\[ K(X, H) = h_x - |x| + h_y - |y| \] (12)

It can be seen that the kernel function used in this paper is of low computational complexity and requires simple addition. At the same time, the size of kernel function can be controlled by adjustment, so that the weighted histogram of kernel function includes not only the grayscale feature of image, but also the position and scale feature of image.

### 3.1.3 Adaptive Variable Scale Kernel

In each frame, three scales (They are 1, 1.2 and 0.8 times the optimal size of the previous frame, respectively) are used to calculate the optimal size of the current frame and for subsequent calculations.

### 3.1.4 State Space

The adaptive scale-varying kernel histogram is used as the state space, first the target kernel function is calculated, and then the kernel function of the candidate target in each frame is calculated.

### 3.2 Observation Model and Evaluation Standard

The purpose of establishing the observation model and evaluation standard is to evaluate the matching degree between the target and the original target after the state transition of the target, so as to obtain the optimal target position.

The evaluation criterion for target tracking is the target similarity measurement, which describes the similarity between the target model and the target candidate model. Similarity measurement is calculated by similarity function. Common similarity functions include Fisher linear criterion [13], histogram intersection [14] and Kullback divergence [15]. According to the description of Comaniciu in Reference [15], Bhattacharyya coefficient, as a similarity measurement function of the target model and observation model, has better resolution than other methods.

Bhattacharyya coefficient are defined as Equation (13).

\[ \hat{\rho}(y) \equiv \rho[\hat{\rho}(y), \hat{q}] = \sum_{u=1}^{m} \sqrt{\hat{p}_u(y)\hat{q}_u} \] (13)

The range of Bhattacharyya coefficient is \([0,1]\). The higher the value, the more similar the two models are. Ideally, the probability distribution is exactly the same, with a value of 1.
The angle cosine between two \( m \) dimensional vectors \( \vec{q} = (\sqrt{q_1}, \ldots, \sqrt{q_m}) \) and \( \vec{p}(y) = (\sqrt{p_1(y)}, \ldots, \sqrt{p_m(y)}) \) is shown in Equation (14).

\[
f(y) = \cos \theta_y = \frac{\vec{p}(y) \cdot \vec{q}'}{||\vec{p}(y)|| \cdot ||\vec{q}'||} = \sum_{s=1}^{c} \sqrt{p_s(y)q_s}
\]

From Equation (14), it can be seen that the smaller the angle between vectors, the bigger the Bhattacharyya coefficient is, the more similar the two vectors are. Conversely, the larger the angle between the two vectors, the smaller the Bhattacharyya coefficient and the smaller the similarity between the two vectors. Although the application of Bhattacharyya coefficient is good, open computing is complicated in hardware. Therefore, it is considered to simplify the Bhattacharyya coefficient and reduce the calculation amount without affecting the performance of similarity measurement.

The evaluation function shown in Equation (15) is adopted in this paper.

\[
\rho(Y) = \sum_{s=1}^{c} \min(q_s, p_s)
\] (15)

Equation (15) can avoid the open-square operation, and the value range of this coefficient is still within \((0,1)\), and the effect is satisfactory.

### 3.3 Tracking Algorithm Flow

1) Initialization, determine the target size and location,
2) In the current frame, \( N \) uniform particles distribution are produced within the range of \( W*H \) at the target position of the center of the above frame. Calculate the kernel histogram in the template size area centered on the particle position
3) Calculate the similarity measurement between the histogram of each particle and the template histogram, and obtain the local optimal value and global optimal value
4) Update the position and velocity of each particle
5) Judge whether the termination condition is satisfied. If so, output the target position; otherwise, jump to 3) continue processing.
6) Whether it is the last frame, if not, it will go to the next frame and jump to 2), otherwise it will end.

(a) 21 frame  (b) 35 frame
4. Summary

In this paper, a method of image segmentation and line detection is proposed to detect the horizon. To identify the suspected ship, we detect it in multiple frames. After identifying the ship, the mean shift algorithm is applied to lock tracked target. The experiment shows that the algorithm we proposed has good effect of detection and tracking for the ship target in the infrared image, and this algorithm has low computational complexity and small computation. It has been realized in the embedded hardware system, and has high application value.

References