Robust and fast ship tracking algorithm based on kernel correlation filter

Wenbo Ji [1, a], Chaojian Shi [1, b], Zewei Yu [2, c, *], Xiaojun Mei [1, d], Qiang Luo [3, e], Haifeng Liu [1,f], Jiansen Zhao[1,g]

[1]Merchant Marine College, Shanghai Maritime University, Shanghai 201306, China

[2]Mathematical Modelling Laboratory, Shanghai Maritime University, Shanghai 201306, China

[3]School of Civil Engineering, Guangzhou University, Guangzhou 510006, China

[a]1445945717@qq.com, [b]cjshi@shmtu.edu.cn, [c]*zwyu@shmtu.edu.cn (corresponding author), [d]576099720@qq.com, [e]luoqiang0617@yeah.net, [f] hfliu@shmtu.edu.cn, [g]jszhao@shmtu.edu.cn

Abstract

This In view of the complex and variable traffic environment at sea, this paper selects the KCF algorithm to achieve the tracking of ships. The algorithm uses the cyclic shift matrix to intensively sample the target, and utilizes the multi-channel HOG feature to enhance the description ability of the classifier. The improved discrete Fourier transform of the kernel correlation function can effectively improve the running speed of the algorithm. Through the tracking test of three sets of typical marine ship surveillance video, it shows that the algorithm has good tracking performance for target motion changes, background clutter and background blur. The excellent performance of the algorithm's calculation speed and accuracy is in line with the special environment of marine ship tracking. And there is also a good improvement in calculation speed and accuracy.

Keywords

Kernel correlation filter；ship tracking；maritime surveillance video; smart ship.

1. Introduction

In recent years, the field of computer vision has developed rapidly. Target tracking as the most important research direction of computer vision has been widely used in transportation, monitoring, medical, manufacturing and other fields. With the rapid increase in ship traffic, the maritime regulatory authorities are facing tremendous maritime regulatory pressures. Compared with onshore, the existing maritime surveillance system is still relatively backward, relying mainly on AIS radio monitoring, radar monitoring, telescope viewing and so on. These traditional methods are mostly dependent on people, monitoring is not reliable enough, and the perspective is not intuitive enough. The complex and volatile environment at sea affects the difficulty of ship monitoring. The high precision, real-time and intelligence of computer vision target tracking can greatly reduce the regulatory difficulties of supervisors, thus achieving better and safer maritime traffic management services.

2. Related Work

The existing mainstream target tracking algorithms can be mainly divided into two types: generation type and discriminant type[1,2]. The generative model is to establish a target model by learning, and then achieve target tracking by comparing the region with the smallest error of the model. The classic algorithms include mean_shift[3] and Kalman filtering[4]. However, such methods ignore the extraction of background information and are prone to tracking drift and recognition errors. The discriminant model is to extract the target and the background features at the same time through the machine learning method to train the classifier, use the trained classifier to discriminate the optimal solution in the next frame and continue to train the classifier[5,6,7]. This type of method is generally better than the generation type. For example, MOSSE(Minimum Output Sum of Squared Error) minimum squared difference and output method correlation filter proposed by Bolme et al.[8], the boundary between the target and the background is well distinguished by the discrete Fourier transform, and the calculation speed of the algorithm is enhanced. Circulant Structure of Tracking-by-detection with Kernels (CSK) was designed by Henriques et al[9]. In 2011 through a densely sampled kernel function circulant matrix plus a regularized least squares (RLS) discriminant classifier[10]. The extraction of such algorithm features is a gray value feature with a single feature. In a subsequent study, Henriques et al. improved the CSK algorithm. The original gray value feature is replaced by the HOG (histogram of oriented gradients) feature ; the core correlation filter is extended to multiple channels for processing the feature information of more complex dimensions. The improved nuclear correlation tracking algorithm (Kernelized Correlation Filter, KCF) stronger explanatory power of the target, the target illumination changes, motion blur, background clutter, occlusion have good tracking performance. And there is also a good improvement in calculation speed and accuracy.

3. KCF tracking algorithm

The KCF ship tracking algorithm is mainly a discriminant tracking algorithm for cyclic dense sampling detection. Firstly, in the process of training the ship classifier on the maritime surveillance video image, the circulatory matrix theory is used to intensively sample the ship in the target area, extract the ship features, and train all the samples in the ridge regression classifier. Then apply the kernel correlation filter algorithm to find the maximum response area and get the position of the target ship. The discrete Fourier transform of the kernel correlation function can effectively improve the running speed of the algorithm.

3.1 Cyclic matrix sampling

In view of the slow sampling and sample redundancy problems existing in the traditional tracking algorithm sparse sampling, the KCF tracking algorithm can obtain enough training samples in a short time, and uses the original sample x to continuously multiply the permutation matrix to obtain n-1 similar cyclic shifts vector, eventually forming an n_order circulant matrix (1) :

$$\mathrm{C}(x) = \begin{bmatrix} x_0 & x_1 & x_2 & & x_{\mathrm{n}-1} \\ x_{\mathrm{n}-1} & x_0 & x_1 & \cdots & x_{\mathrm{n}-2} \\ x_{\mathrm{n}-2} & x_{\mathrm{n}-1} & x_0 & & x_{\mathrm{n}-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_0 \end{bmatrix} \tag{1}$$

The obtained cyclic matrix is subjected to diagonalization of the discrete Fourier matrix in the Fourier space, and the operation of the matrix is converted into the hadmad product of the vector, that is, the dot multiplication of the elements. This method greatly reduces the computational complexity of the algorithm, shortens the computation time, and satisfies the fast requirements of tracking detection.

3.2 Classifier training

By taking a training sample set from the cyclic shift matrix（$\mathrm{x_i}, \mathrm{y_i}$), i denotes the requiredship image of each frame. We use ridge regression to train the target classifier by finding parameters that minimize the risk of regularization , in the form of (2):

$$\mathrm{f}(x) = \langle \theta, x \rangle + t \tag{2}$$

among them: ⟨·,·⟩expressed as a dot product operator. The minimization problem can be solved by the least squares method (3):

$$\min_{\theta,\mathrm{t}} \sum_i F\big(y_i, f(x_i)\big) + \lambda \|\theta\|^2 \tag{3}$$

The matrix is expressed as (4):

$$\min_{\theta,t} \sum_i \|X\theta - y\|^2 + \lambda\|\theta\|^2 \quad (4)$$

$F(y_i, f(x_i))$is loss function for training the classifier; λ Used to control the level of regularization of the classifier,which is the structural complexity of the system.X= $[x_1, x_2, x_3, \dots x_n]^T$ represents a horizontal amount of labels for each element corresponding to a sample; y is a column vector.

Ship training images are mapped to their feature space $\beta(x)$. Nonlinear mapping function $\beta(x)$ the specific form is difficult to express. Generally we let A represent the kernel matrix of the kernel space $A=\beta(X)\beta(X)^T$.We ridge regression classifier $f(x_i)=\theta^T\beta(x_i)$ in $\beta(x)$ sample new space to find a column vector, so that it can be divided into linear. Its weight coefficient (5) is:

$$\theta = \min_{\theta} \|\beta(X)\theta - y\|^2 + \lambda\|\theta\|^2 \quad (5)$$

make $\theta = \sum_i \partial_i \beta(x_i)$, Then the above minimum solution input θ linear combination. The closed form solution (6) is as follows:

$$\partial^* = (A+\lambda I)^{-1} y \quad (6)$$

among them: A= $\beta(X)\beta(X)^T$ for the kernel matrix, the process is reversible; I is the unit matrix; θ is a vector in the space formed by the row vector: $\beta(X)=[\beta(x_1), \beta(x_2) \dots \beta(x_n)]^T$.In the case where the matrix is converted to a Fourier domain, we can be in the vector Perform operations on it. We can get the formula In the case where the matrix is converted to a Fourier domain, We can operate on vector $\vec{y}$ .Therefore, we can get the optimal solution (7) of $\min_{\theta,t} \sum_i F(y_i, f(x_i)) + \lambda\|\theta\|^2$ as shown below:

$$\theta = F^{-1}\left(\frac{F(y)}{F(k)+\lambda}\right) \quad (7)$$

3.3 Quick check

After the above steps solve the problem of tracking the training detector, for the next frame, the sample set to be detected, we get the same set $t_i = P^i t$ after cyclic shift. The region of maximum response in Classifier $f(t_i) = \partial^T \beta(X)\beta(t_i)$ both (8) is considered to be the result of the movement of the target ship. In this way, each subsequent frame is detected to achieve tracking of the ship.

$$\hat{f} = (\hat{A}^{xt}) \cdot \hat{\partial} \quad (8)$$

4. Results and Discussion

In order to verify the effectiveness of the algorithm, the algorithm is written in Matlab and C language. The computer used in the experiment was configured as an Intel Core i7 4790 CPU with a frequency of 3.60GHz and 8GB of memory. In order to compare the specific performance of the marine environment, the experiment selected three typical videos from the port and onboard surveillance video. Three special scenarios of ship steering, ship from far and near driving, and ships traveling in fog were verified. The experimental results are shown in Fig. 1 to Fig. 3.

The 331-frame video of Fig. 1 depicts a process in which the ship sails from left to right and the ship is deformed during the steering. The ship's background is complex; the 299-frame video of Fig. 2 depicts the video image size changes that occur as the ship moves from near to far. The KCF algorithm adopts the multi-channel HOG feature to enhance the description ability of the classifier's detection target. The change of the tracking rectangle describes the change of the ship size well, and the tracking effect is good. The 277 frame video of Fig. 3 is derived from the tracking of another ship on a fastmoving ship in the foggy day. The scene in Fig. 3 is more complicated than the first two scenarios. The color of the tracked ship is similar to the background color. In addition to the influence of heavy fog, the ship features a lot of clutter. Due to the interference of the waves, the video taken on the ship is up and down.Two ships facing each other driving makes the relative speed very fast. In such a complicated situation, the KCF algorithm tracks stably, and there is no drift or tracking failure, and the tracking effect is good.

Fig. 1 ship steering

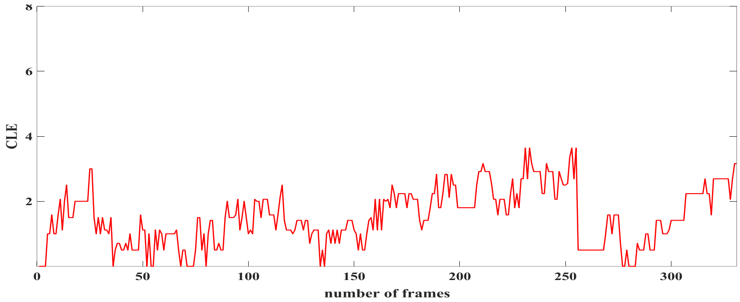
Fig. 2 ship from near and far driving

Fig. 3 ships traveling in fog

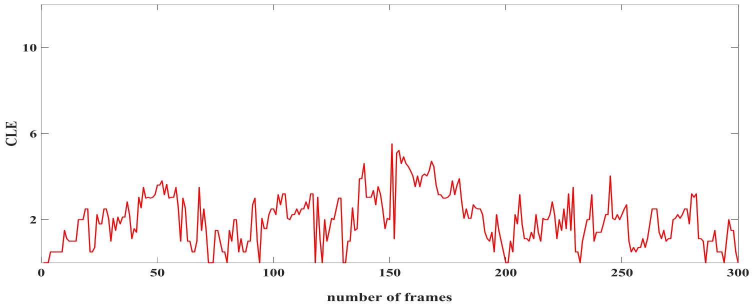
In order to quantitatively analyze the performance of the tracking method, the Center Position Error (CLE) was selected as the evaluation standard. Its calculation formula (9) is:

$$CLE = \sqrt{(x_i - x_{i\_r})^2 + (y_i - y_{i\_r})^2} \tag{9}$$

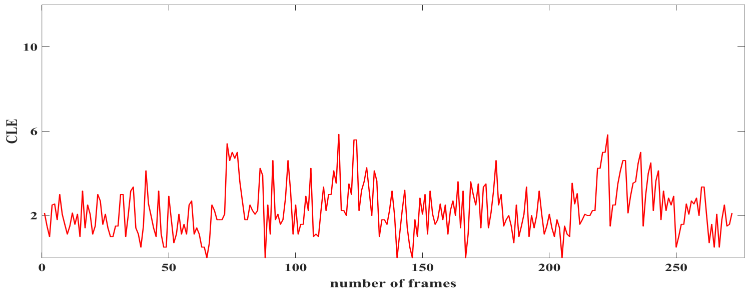
Where CLE is the central position error, (x_i, y_i) is the target center position obtained by the tracking algorithm in the ith frame video, and (x_{i_r}, y_{i_r}) is the target center position obtained by manually marking the ith frame video.The result is shown below:



ship steering



ship from near and far driving



ships traveling in fog

Fig. 4 Center Position Error curve

It can be seen from the figure that the center position error of the algorithm tracking data and the artificial marker data in the three scenes is generally between 1 and 2 pixels, and the maximum peak value is lower than 6 pixels, and there is basically no curve band with a large drop. Fig. 4 shows that the CKF ship tracking algorithm is stable and effective. Table 1 summarizes all the data from the experiment, where ACLE represents the average of CLE.

Table 1 Three Scene comparing

VIDEO	FPS	ACLE
Scene 1	5.47	1.52
Scene 2	4.15	1.99
Scene 3	1.87	2.05

5. Conclusion

The KCF tracking algorithm uses the cyclic matrix theory and the kernel correlation filtering algorithm to speed up the calculation of target sampling and data processing. The multi-channel HOG feature is used to enhance the description ability of the classifier's detection target. The algorithm changes the target's motion and background. The background is blurred and has good tracking performance. These excellent performances are in line with the special environment of marine vessel tracking and can provide a reliable tracking method for maritime video surveillance. Experimental results show that this tracking algorithm performs well in the field of ship tracking.

Acknowledgements

This work was jointly supported by the National Natural Science Foundation of China (51709167, 51579143), Shanghai Committee of Science and Technology, China (18040501700, 18295801100, 17595810300).

References

[1] Ross D A , Lim J , Lin R S , et al. Incremental Learning for Robust Visual Tracking, International Journal of Computer Vision, Vol. 77(1-3) 2008, p. 125-141.

[2] Kalal Z , Mikolajczyk K , Matas J . Tracking-Learning-Detection. Transactions on Software Engineering, Vol.34(7) 2011, p. 1409-1422.

[3] Comaniciu D , Ramesh V , Meer P . Real-time tracking of non-rigid objects using mean shift, Proceedings IEEE Conference on Computer Vision and Pattern Recognition 6(2002) No.PR00662.

[4] Kalman R E. A New Approach To Linear Filtering and Prediction Problems,Transactions of the ASME-Journal of Basic Engineering, Vol. 82(1) 1960, p.35-45

[5] Danelljan M , Khan F S , Felsberg M , et al. Adaptive Color Attributes for Real-Time Visual Tracking, Conference on Computer Vision and Pattern Recognition, 2014, p. 1090-1097.

[6] Zhang K , Zhang L , Yang M H . Fast Compressive Tracking, Transactions on Pattern Analysis & Machine Intelligence, Vol.36(10) 2014, p. 2002-2015.

[7] Ma C , Yang X , Zhang N C , et al. Long-term correlation tracking, Conference on Computer Vision and Pattern Recognition , (Boston, MA 07 Jun 2015), P. 5388-5396.

[8] Bolme D S ,Beveridge J R ,Draper B A , et al. Visual object tracking using adaptive correlation filters, Conference on Computer Vision and Pattern Recognition, June 2010, p.13-18.

[9] João F. Henriques, Caseiro R , Martins P , et al. Exploiting the Circulant Structure of Tracking-by-Detection with Kernels,Proceedings of the 12th European conference on Computer Vision, Vol. 4 (2012), p.702-715.

[10] Henriques J F , Caseiro R , Martins P , et al. High-Speed Tracking with Kernelized Correlation Filters, Transactions on Pattern Analysis and Machine Intelligence, Vol.37(3) 2015, p. 583-596.