Article

Parallel Three-Branch Correlation Filters for Complex Marine Environmental Object Tracking Based on a Confidence Mechanism

Yihong Zhang *, Shuai Li, Demin Li, Wuneng Zhou, Yijin Yang, Xiaodong Lin and Shigao Jiang

College of Information Science and Technology, Engineering Research Center of Digitized Textile & Fashion Technology, Ministry of Education, DongHua University, Shanghai 201620, China; 2191514@mail.dhu.edu.cn (S.L.); deminli@dhu.edu.cn (D.L.); wnzhou@dhu.edu.cn (W.Z.); 2171318@mail.dhu.edu.cn (Y.Y.); 2191511@mail.dhu.edu.cn (X.L.); 2191460@mail.dhu.edu.cn (S.J.)

* Correspondence: zhangyh@dhu.edu.cn; Tel.: +86-138-1782-6259

Received: 15 August 2020; Accepted: 10 September 2020; Published: 12 September 2020

Abstract: Marine object tracking is critical for search and rescue activities in the complex marine environment. However, the complex marine environment poses a huge challenge to the effect of tracking, such as the variability of light, the impact of sea waves, the occlusion of other ships, etc. Under these complex marine environmental factors, how to design an efficient dynamic visual tracker to make the results accurate, real time and robust is particularly important. The parallel three-branch correlation filters for complex marine environmental object tracking based on a confidence mechanism is proposed by us. The proposed tracker first detects the appearance change and position change of the object by constructing parallel three-branch correlation filters, which enhances the robustness of the correlation filter model. Through the weighted fusion of response maps, the center position of the object is accurately located. Secondly, the Gaussian-triangle joint distribution is used to replace the original Gaussian distribution in the training phase. Finally, a verification mechanism of confidence metric is embedded in the filter update section to analyze the tracking effect of the current frame, and to update the filter sample from verification result. Thus, a more accurate correlation filter is trained to prevent model drift and achieve a good tracking effect. We found that the effect of various interferences on the filter is effectively reduced by comparing with other trackers. The experiments prove that the proposed tracker can play an outstanding role in the complex marine environment.

Keywords: parallel three-branch correlation filters; confidence mechanism; complex marine environment; object tracking

1. Introduction

Due to the complexity of the marine environment, it is meaningful to study the tracking of complex marine environmental objects [1]. Owing to changes in appearance, changes in lighting, fast motion, blur and partial occlusion, it is difficult for traditional visual object trackers to achieve accurate and robust tracking of complex marine environmental objects [2]. Therefore, based on the in-depth analysis of the visual object tracking technology, this passage proposes a series of strategies to improve the tracking performance for complex marine environmental objects.

Correlation filtering is a key point of object tracking research, which originated from signal processing [3]. Given the object in the first frame, perform feature extraction on the object and analyze the region of interest. Then, find similar features and regions of interest in subsequent images and make predictions for the position of the next frame. We take the coordinates corresponding to maximum peak as the location of the object. Following this, the earliest tracker, MOSSE [4], was proposed, which exploited the most simple tracking method of related filtering ideas. Subsequently, there were many
related improvements based on MOSSE, such as CSK [5], KCF [6], etc. According to introducing nuclear method, they have achieved good results, especially the KCF using the cyclic matrix calculation makes the tracking speed amazing. Subsequently, the DSST [7] tracker innovatively handles scale changes. Two independent correlation filters, the translation filter and scale filter, are used for object positioning and scale evaluation. Due to the problem that all the above-mentioned methods are influenced by the boundary effect. The SRDCF [8] tracker came into being. The tracker takes the image signal as large as possible, retains more real information of the object, and then punishes samples that are farther from the center of the object through a spatial weight coefficient.

Then the ECO_HC [9] tracker was proposed at the top computer vision conference CVPR2017. This method uses efficient convolution operations to extract the feature matrix and performs multi-feature fusion to obtain the feature map. Correlate the response spectrum and analyze it to get the position of the object. The ECO_HC tracker effectively improves the problems of low filter efficiency and the overfitting of related filters.

However, the ECO_HC tracker also has its shortcomings. When the object occurring deformation, light changes, or self-rotation, it is no reason to renew the relevant model with the same learning rate. In addition, the use of a smooth Gaussian distribution as the expected output will make the positioning inaccurate [10]. Subsequently, the PCF [11] tracker came into being. Compared with the ECO_HC tracker on the video sequences, precision accuracy and overlap accuracy of the PCF tracker are better than the former. When the object rotates significantly, the light changes and deforms and the PCF tracker can successfully track the object. However, the above trackers do not effectively track the object in the complex marine environment specifically. So as to promote the precision of tracking performance in the complex marine environment, we put forward parallel three-branch correlation filters for tracking the object, which can efficiently improve the robustness of object tracking [12]. Aiming at the problem that parallel three-branch correlation filters lack effective supervision mechanism, this paper proposes a confidence mechanism to analyze the distribution of related responses to verify whether it is reliable [13], and supervise the update of the sample model with inspection results.

The contributions of this paper are as follows:

First, target samples with three different weights are respectively trained with three parallel correlation filters and, where the values of the weights are determined by the learning rate. Three different learning rates are used to renew the model and perform weighted fusion to effectively improve the robustness of the tracker to overcome bad conditions in the marine environment.

Second, the confidence metric proposed in this paper judges whether to update the filter, that is, to supervise the update of the sample model. Confidence Response (CR) aims to evaluate the numerical difference of the largest n responses on the relevant response graph. When the CR value is above the threshold, the result is valid and the sample model can be updated, otherwise it will be not updated.

Third, we have done a lot of experiments in the ocean environment, which fully proves that the proposed tracker can still show excellent results when the background is blurred and the waves interfere. In addition, we also compared nine other representative trackers on the OTB-2015 dataset, showing that the proposed tracker can deal with object tracking efficiently in complex and changing scenarios.

2. Related Work

The PCF tracker is the baseline tracker of this article, and it has a good effect on object tracking in complex scenes. In the initial frame, the PCF tracker uses the shared generated samples and improved expected output to train two parallel correlation filters. Then, in the present frame, according to balancing the response maps of two filters, the PCF tracker is able to detect the location of the object with the Newton method [14]. Subsequently, the PCF tracker uses Gaussian Mixture Model to add a new sample or merge the two most similar samples to generate a new sample. Then, we use the new sample to renew the PCF1 and PCF2 at different learning rates.

PCF tracker uses object samples $f_{\text{tr}}^{1}$, $f_{\text{tr}}^{2}$, ..., $f_{\text{tr}}^{m}$ with different weights $w_{\text{tr}}^{1}$ to train two parallel correlation filters, $h1_{\text{tr}}^{1}$ and $h2_{\text{tr}}^{1}$, respectively. The value of the weights $w_{\text{tr}}^{1}$ and $w_{\text{tr}}^{2}$ are changed.
where $\alpha$ is the fusion factor, and $H_{t-1}^l$ and $H_{t-1}^l$ are two filters in the frequency domain.

After object positioning is completed, the image features of the object are collected from the positioning area as new samples, and then the similarity between the new samples and the existing training set samples is calculated using the Gaussian mixture model. If new samples and the training samples have a particularly big feature difference, the new sample is added and the old sample added previous is removed to make sure that the number of samples remains unchanged. If the similarity between the new sample and the training sample is high, the two most similar samples are merged.

The weight of each training sample is updated by Equation (4):

$$w_k^l = (1 - \eta)w_k^{l-1}$$

where $\eta$ is the learning rate and $t$ represents the $t$-th frame. It can be drawn from the above formula that different learning rates determine different sample weights.

To judge the scale of the object, the DCF tracker is used to extract the scale features under different scale factors [15], and then the relevant output response score of each scale factor is calculated using Equation (5). Finally, the scale corresponding of the max output response score is the scale of the object.

$$y_{t, \text{scale}} = \mathcal{F}^{-1} \left\{ \frac{\sum_{l=1}^{d} \hat{A}_{l-1, \text{scale}} Z_{l, \text{scale}}}{B_{l-1, \text{scale}} + \lambda} \right\}$$

where $\hat{A}_{l-1, \text{scale}}$ and $B_{l-1, \text{scale}}$ are the scale factors and $\lambda$ is the regularization parameter used to prevent the overfitting of the correlation filtering model. The greater the $\lambda$, the greater the punishment. The common methods such as the L-curve and the GCV for solving ill-conditioned equations are compared with the analysis of examples. The calculation example shows that the Tikhonov regularization parameter optimization method is a feasible and effective method to get the optimal regularization parameter $\lambda$.

The fast Fourier transform can effectively transform Equations (1) and (2) into the frequency domain for efficient calculation. Since the regularization parameters $\omega$ break the closed solution of the function and the solution cannot be directly obtained for the objective equation, the conjugate gradient method can obtain the solution sum of the above two objective functions. Then, the output response score of the relevant filter is calculated by Equation (3), and finally the location of the object is obtained by the Newton iteration method.

$$y_{t, \text{pos}} = \mathcal{F}^{-1} \left\{ \sum_{l=1}^{d} (\alpha H_{t-1}^l Z_{t, \text{pos}} + (1 - \alpha) H_{t-1}^l Z_{t, \text{pos}}) \right\}$$

where $\alpha$ is the fusion factor, and $H_{t-1}^l$ and $H_{t-1}^l$ are two filters in the frequency domain.

The objective functions of PCF1 and PCF2 are shown in Equations (1) and (2), respectively:

$$\epsilon(h_1) = \frac{n}{\sum_{k=1}^{n} \sum_{l=1}^{d} (w_1 f_1^l - g_{sk})^2 + \lambda \sum_{l=1}^{d} \| \omega h_1^l \|^2}$$

$$\epsilon(h_2) = \frac{n}{\sum_{k=1}^{n} \sum_{l=1}^{d} (w_2 f_2^l - g_{sk})^2 + \lambda \sum_{l=1}^{d} \| \omega h_2^l \|^2}$$

where $p,k$ represents the amount of samples generated by training data and cyclic shift method, respectively, and $g_{sk}$ represents the improved joint-expected output distribution, which is a linear weighted fusion of Gaussian distribution and triangular distribution. The second term of the two equations are regularization terms, and $\omega$ represents spatial regularization parameters, which are used to prevent the model from overfitting and used to solve boundary effects. The $\lambda$ is a regularization parameter used to prevent the overfitting of the correlation filtering model. The greater the $\lambda$, the greater the punishment.
where $A_{l-1, scale}$ and $B_{l-1, scale}$ represent the scale correlation filter of the previous frame, and $\lambda (\lambda \geq 0)$ is the regularization parameter.

The expected output of the traditional DCF method generally follows a two-dimensional Gaussian distribution [16], and the peak value is relatively smooth. However, because the DCF model uses a cyclic shift window to generate lots of synthetic training samples, the only real sample is in the center position, so the sharp expected output peak is more reasonable and the drift of the relevant filtering model can be avoided. In the expected output response graph, the areas farther away from the center represent negative sample labels, so the values of these locations should be set to tend to zero, while the triangular distribution graph has a smaller decline rate at the valley. Therefore, Gaussian distribution (as shown in Equation (6)) and triangular distribution (as shown in Equation (7)) are not suitable for DCF tracker. The joint expected output response distribution is used to replace the original Gaussian distribution graph. As shown in Equation (8), the Gaussian distribution and the triangular distribution are organically fused together by a multiplication operator [17], which can promote the robustness of the tracker.

$$g(x, y) = \frac{1}{2\pi\sigma^2}e^{-\frac{(x^2+y^2)}{2\sigma^2}}, \left(-\frac{a}{2} \leq x \leq \frac{a}{2}, -\frac{b}{2} \leq y \leq \frac{b}{2}\right)$$ (6)

$$s(x, y) = \left(1 - \frac{2}{a}|x|\right)\left(1 - \frac{2}{b}|y|\right), \left(-\frac{a}{2} \leq x \leq \frac{a}{2}, -\frac{b}{2} \leq y \leq \frac{b}{2}\right)$$ (7)

$$m(x, y) = \frac{1}{2\pi\sigma^2}\left(1 - \frac{2}{a}|x|\right)\left(1 - \frac{2}{b}|y|\right)e^{-\frac{(x^2+y^2)}{2\sigma^2}}, \left(-\frac{a}{2} \leq x \leq \frac{a}{2}, -\frac{b}{2} \leq y \leq \frac{b}{2}\right)$$ (8)

where $\sigma$ represents the standard covariance of the distribution map, $x, y$ represent the axis of the distribution map coordinates, $a, b$, respectively, the width and height of the distribution map, and the origin of the coordinates of the distribution map is $(\frac{a}{2}, \frac{b}{2})$.

3. Proposed Approach

3.1. Parallel Three-Branch Correlation Filters

The DCF tracking algorithm can effectively train the discriminant related filter model in the frequency domain. The difference between the object and the background is effectively used. When the object is in a more complex scene, the algorithm has better tracking and discrimination capabilities. After the object is located, the target image feature is extracted, and the related filter model is updated online. The formula is as follows:

$$\min_{HF} = \|\sum_{l=1}^{d} h_{l} f_{l}^{1} - f_{2}\|^2 + \lambda \sum_{l=1}^{d} \|h_{l}\|^2$$ (9)

where $l$ is the dimension, and $h$ is an improved correlation filter. $\lambda$ is the adjustment parameter to control the influence of regularization [18]. Calculate the partial derivative after FFT and simplify to get:

$$H_{l} = \frac{\mathcal{G}f_{l}}{\sum_{k=1}^{d} \mathcal{F}_{l}^{k} + \lambda}$$ (10)

where $G, F$ are after Fourier transform. Calculate the numerator and denominator, respectively, in the process of resolving:

$$y = \mathbb{F}^{-1}\left\{\frac{\sum A_{l} Z_{l}}{B + \lambda}\right\}$$ (11)
where $F^{-1}$ is the inverse Fourier transform, $A, B$ represent the numerator and denominator, respectively, and $z$ is the calculated feature map. The value of $y$ is changed by changing the value of $z$. When $y$ is the largest, it is the time when the correlation filter is most suitable.

$$
A_l^t = (1 - \eta)A_{l-1}^t + \eta F_{t+1}F_l^t
$$

(12)

$$
B_t = (1 - \eta)B_{t-1} + \eta \sum_{k=1}^d F_k^tF_l^t
$$

(13)

where $\eta$ is the learning rate, and $t$ represents the number of frames. From the first frame, neither the numerator nor the denominator has the previous item, and the initial numerator and denominator are well obtained.

Then, the tracker with parallel three-branch correlation filters is proposed by us [19]. The tracker uses object samples $f_1^t, f_2^t, \ldots, f_m^t$ with different weights $w_{kp}^t$ to train three parallel correlation filters, $h_1^t, h_2^t$ and $h_3^t$, respectively. The value of the weights $w_{1k}^t, w_{2k}^t$ and $w_{3k}^t$ are decided by the learning rate. The functions of PCF1, PCF2 and PCF3 can be shown as Equations (14), (15) and (16), respectively:

$$
e(h_1^t) = \sum_{k=1}^n ||\varepsilon(h_1^t) - g_{k}||^2 + \lambda \sum_{l=1}^d ||\omega h_1^t||^2$$

(14)

$$
e(h_2^t) = \sum_{k=1}^n ||\varepsilon(h_2^t) - g_{k}||^2 + \lambda \sum_{l=1}^d ||\omega h_2^t||^2$$

(15)

$$
e(h_3^t) = \sum_{k=1}^n ||\varepsilon(h_3^t) - g_{k}||^2 + \lambda \sum_{l=1}^d ||\omega h_3^t||^2$$

(16)

The conjugate gradient method is used to obtain the sum of the solutions of the above three objective functions [20]. The output response score of the relevant filter is calculated by Equation (17), and finally the position of the object in the present frame is obtained by the Newton iteration method. Three correlation filters use three different learning rates to determine the location of the object.

$$
y_{t,pos} = F^{-1} \left\{ \sum_{l=1}^d \left( \alpha H_1^{l-1}Z_{t,pos} + \beta H_2^{l-1}Z_{t,pos} + (1 - \alpha - \beta)H_3^{l-1}Z_{t,pos} \right) \right\}
$$

(17)

where $\alpha, \beta$ are fusion factors, and $H_1^{l-1}, H_2^{l-1}$ and $H_3^{l-1}$ represent three parallel correlation filters. Three different learning rates are used to renew the model and perform weighted fusion [21] to effectively improve the performance of the tracker for complex marine environmental objects.

The weight of each training sample is updated by Equation (18):

$$
t_{kp} = (1 - \eta)w_{kp}^{-1}
$$

(18)

Then, the relevant output response score of each scale factor is calculated using Equation (19).

$$
y_{t, scale} = F^{-1} \left\{ \sum_{l=1}^d \frac{A_{l-1, scale}Z_{t, scale}}{B_{l-1, scale} + \lambda} \right\}
$$

(19)
3.2. Confidence Mechanism

Without an effective supervision mechanism during the tracking of the parallel three-branch correlation filters method, errors are easily generated when the object is interfered with by external environmental factors such as occlusion and illumination changes, and then the errors are brought into the update of the sample model. The accumulated error will cause the tracking frame to drift, which will eventually cause the object to be lost. Aiming at the problem that the method of parallel three-branch correlation filters lacks effective supervision mechanism, this paper proposes a verification mechanism [22] to analyze the reliability of current results by evaluating the distribution of related responses, and supervise the update of the sample model based on the inspection results. Before inputting the tracking object area into the relevant filter, a verification mechanism of confidence measurement is added to analyze and verify the distribution of the relevant response and tracking results, and judge whether to update the sample model according to the verification mechanism [23]. When the tracking result meets the set conditions, it is used to update the sample model, otherwise it is not adopted. The confidence metric in the update stage of the filter acts a pivotal part in tracking the quality of the tracking effect. The traditional tracking method compares the value of the related responses and takes the object area with the highest related response as the correct tracking result. In practical applications, this evaluation method has a large error. When there is an interference area with a high similarity, the object may be lost, but the value of its related response is still high, as shown in Figure 1. When the form of object changes, the relevant response value may be low, which will affect the judgment of the tracking effect.

![Comparison of peak distribution](image1)

**Figure 1.** Comparison of the peak distribution of the PCF tracker and the proposed tracker when occurring occlusion. The red one is the proposed tracker and the green one is the baseline tracker PCF. There are three status: (a) The object has not affected by interference; (b) The object has affected by interference; (c) The object gradually breaks away from the interference.

In most cases, we usually take the largest point of the peak as the object location, which is not reliable [24]. This paper introduces a new confidence metric \( CR \) (Confidence Response) and embeds it in the filter update module to analyze the peak distribution of the relevant response and evaluate the tracking results. The solution process is as follows: select \( n \) related responses from the current frame image and arrange them in order from largest to smallest. Use \( R(i) \) to indicate the response value in...
the i-th position, \( p(R(i)) \) is the probability of the response value appearing. Find the expected \( \mu \) of \( n \) response values:

\[
\mu = \sum_{i=1}^{n} R(i) \times p(R(i))
\] (20)

Then, introduce the concept of standard deviation to measure the response value \( R(i) \). The available CR formula is as follows:

\[
CR = \sqrt{\sum_{i=1}^{n} (R(i) - \mu)^2}
\] (21)

CR aims to evaluate the numerical difference of the largest \( n \) responses on the relevant response graph [25]. The larger the CR value, the more dispersed the distribution between the interference area and the true area in the frame, and the more reliable the tracking effect; otherwise, the more concentrated the distribution between the interference area and the true area in the frame, the more unreliable the tracking effect.

As shown in Figure 1, the peak response is generally regarded as the location of the object [26]. However, in reality, when occurring occlusion, deformation, etc., the peak response does not represent the location of the actual object. As shown in Figure 1a, when there is no interference, the response peak is single, which can track the object well. As shown in Figure 1b, when there is interference, the response peak gradually increases. At this time, the highest peak is not the actual object position. As shown in Figure 1c, taking the highest peak position as the object position will cause tracking failure such as object drift [27]. Therefore, the confidence metric proposed in this paper judges whether to update the filter, that is, to supervise the update of the sample model. When the CR value is above the set threshold, the result is valid and the sample model can be updated, otherwise it is not updated. As shown in Figure 1c, when the CR value is lower than the threshold we set, the model update is stopped [28].

When the object is occluded, the baseline method PCF continuously updates the sample model, resulting in the accumulation of errors gradually and the occurrence of tracking drift [29]. The proposed method judges whether to update the sample model with the comparison result of the confidence metric CR value and the threshold value, and the update is no longer allowed when the object is occluded. Therefore, the improved sample update strategy can effectively solve the problem that the baseline PCF method easily causes tracking frame drift and object loss. According to the confidence mechanism, we can effectively improve the performance of the tracker.

### 3.3. Algorithm of the Proposed Tracker

As shown in Algorithm 1, we use three parallel correlation filters with different learning rates to make the tracking more robust. Firstly, if we use more branches, the speed will drop drastically, which makes tracking not time sensitive. Secondly, too many branches will cause overfitting. We combine the confidence mechanism and three-branch correlation filtering to be almost the same as the baseline tracker in terms of rate, but greatly improve the accuracy. The specific algorithm is shown below, which specifically explains the working principle step by step. In the initial video frame, the algorithm uses the prior information of the object to initialize three parallel correlation filters. Firstly, update the position by fusing the values of the three-branch parallel correlation filter, then use the scale filter to calculate the scale of the object, and finally extract the features of the object based on the new position and scale, and use the GMM to generate a new sample set to train the relevant filter model online. Finally, use the confidence mechanism to determine whether the model is updated. In subsequent frames, repeat the above steps.
Algorithm 1: Parallel Three-Branch Correlation Filters Tracking Algorithm Based on a Confidence Mechanism

Input:
1: Image $I_t$.
2: Detected target position $P_{t-1}$ and scale $S_{t-1}$.

Output:
1: Detected target position $P_t$ and scale $S_t$.

Loop:
1: Initialize $H_1^1$, $H_2^1$, $H_3^1$ and $A_1$, $B_1$, scale in the first frame by Equations (14)-(16) and (12)-(13).
2: for $t \in [2, t_f]$ do.
3: Position detection:
4: Extract position features $Z_{t, pos}$ from $I_t$ at $P_{t-1}$ and $S_{t-1}$ by a search region.
5: Compute three parallel correlation scores $y_{t, pos}^1$, $y_{t, pos}^2$, $y_{t, pos}^3$.
6: Merge the three correlation scores to $y_{t, pos}$ by Equation (17).
7: Set $P_t$ to the target position by Newton iterative method.
8: Scale detection:
9: Extract scale feature $Z_{t, scale}$ from $I_t$ at $P_{t-1}$ and $S_{t-1}$ by a search region.
10: Compute correlation scores $y_{t, scale}$ by Equation (19).
11: Set $S_t$ to the target scale that maximizes $y_{t, scale}$.
12: Model update:
13: Extract new sample features $F_{t, pos}$ and $F_{t, scale}$ from $I_t$ and $S_t$.
14: Generate new training set by the GMM method.
15: Compute response scores $CR$ by Equation (21).
16: If $CR >$ threshold, return step 2; else continue.
17: Update the model $H_1^t$, $H_2^t$, $H_3^t$ by the learning rate $\eta_1$, $\eta_2$, $\eta_3$.
18: Update the scale model $A_{t, scale}$, $B_{t, scale}$ by the learning rate $\eta_0$.
19: Return $P_t$, $S_t$.
20: end for.

4. Results

4.1. Experimental Results and Analysis in the Complex Marine Environment

To prove that our tracker is superior to the baseline tracker, Figure 2 shows the comparison between the baseline tracker PCF and the tracker proposed in this paper in different complex marine environments. The experiments clearly show that the proposed tracker can effectively suppress the drift caused by tracking failure, which makes the tracking effect more robust.

For the purpose of verifying that the proposed tracker can track the complex marine environmental objects with better performance, we conduct a comparative experiment on the complex marine environmental objects to compare the proposed tracker with the current representative trackers, including the PCF tracker [11], ECO-HC tracker [9], Staple tracker [30] and CSK tracker [5].

Figure 3 shows that the proposed tracker behaves significantly better than that of several other trackers when tracking complex marine environmental objects with motion blur and rotation deformation; the proposed tracker can effectively learn changes in appearance and accurately locate the object. The following three video paragraphs are representative, highlighting the target tracking in the three most common marine environments, especially the tracking of small targets in the ocean, the interference of ocean waves, and the interference of other ships. We select a few representative frames of each data, so the selection is meaningful. We can see from several sets of pictures that when other trackers have already drifted (once they drift, they will continue to drift), our trackers can still perform well and robustly, and this is an intuitive experience. We have a lot of data later to verify the excellent performance of our tracker.
18: Update the scale model, \( \text{scale}_A \), \( \text{scale}_B \) by the learning rate \( \eta \).

19: Return \( t_P \), \( t_S \).

20: end for.

4. Results
4.1. Experimental Results and Analysis in the Complex Marine Environment

To prove that our tracker is superior to the baseline tracker, Figure 2 shows the comparison between the baseline tracker PCF and the tracker proposed in this paper in different complex marine environments. The experiments clearly show that the proposed tracker can effectively suppress the drift caused by tracking failure, which makes the tracking effect more robust.

![Figure 2. Comparison results of the proposed tracker and the baseline tracker PCF in the complex marine environment.](image)

4.2. Comparison of the Proposed Tracker and the Baseline Tracker on OTB-2013

We have done experiments on the OTB-2013 dataset [31] to compare the benchmark tracker with the proposed tracker. Tables 1 and 2 show the center position accuracy and overlap accuracy of the tracker under 11 different attributes, the attributes include scale variation of the object (SV), illumination variation of the environment (IV), rotation of the object out of plane (OPR), the object is occlusion (OCC), the background of the object appears cluttered (BC), the object deforms (DEF), the object produces motion blur (MB), the fast motion of the object (FM), the rotation of the object in the plane (IPR), the movement of the object out of view (OV), and the low resolution of the camera (LR). Compared with the baseline tracker, the tracking accuracy of the proposed tracker has been improved in nine attributes. The results are shown in Figure 4, which further verified that the tracker in this paper can performs better than baseline tracker in complex environment, especially for the following two attributes: the illumination variation of the environment and the background of the object appears cluttered. These attributes determine whether the proposed tracker can perform excellently in the complex marine environment.

![Table 1. Central position accuracy and average accuracy of the proposed tracker and baseline tracker in 11 different attributes. Red represent the better performance.](image)

| Trackers | SV  | IV  | OPR | OCC | BC  | DEF | MB  | FM  | IPR | OV  | LR  | AP  |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Baseline | 83.5| 82.3| 87.0| 90.1| 84.5| 91.7| 76.8| 77.7| 81.2| 81.0| 53.3| 88.0|
| Ours     | 85.9| 84.9| 88.6| 92.2| 87.8| 91.6| 79.6| 81.4| 83.3| 80.1| 53.9| 89.3|

![Table 2. Overlap accuracy and average accuracy of the proposed tracker and baseline tracker in 11 different attributes. Red represent the better performance.](image)

| Trackers | SV  | IV  | OPR | OCC | BC  | DEF | MB  | FM  | IPR | OV  | LR  | AUC |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Baseline | 77.1| 78.5| 79.6| 84.7| 80.5| 91.0| 76.8| 75.3| 72.8| 82.8| 53.9| 82.5|
| Ours     | 77.7| 79.4| 79.9| 85.3| 81.6| 90.7| 76.8| 75.6| 73.2| 81.9| 55.3| 82.9|
Our trackers can still perform well and robustly, and this is an intuitive experience. We have a lot of data later to verify the excellent performance of our tracker.

Figure 3. Qualitative experimental results from marine environmental videos. Five different colors represent representative five trackers, respectively.
which can deal with changes in the complex marine environment and make the tracking e...

we can see that, among these 10 trackers, the proposed tracker performs best with the AP of 87.8%.

The experiments further verify that the proposed tracker can solve the problem of blurred background
performance than other nine trackers in the perspective of position estimation and scale estimation,
proposed tracker achieved the highest AUC among all 10 attributes. In summary, the proposed tracker
video sequences), MB (contains 29 video sequences) and OV (contains 14 videos). According to the
Sensors
2020

outstanding for marine objects.

are particularly important in the marine environment. The proposed tracker has more excellent
object tracking for changes in form, interference from obstructions, changes in light. These factors
are particularly important in the marine environment. The proposed tracker has more excellent
performance than other nine trackers in the perspective of position estimation and scale estimation,

Figure 4. Comparison experimental results of the proposed tracker and the baseline tracker PCF.

4.3. Results and Analysis on the Dataset OTB-2015

The OTB-2015 dataset [32] is a benchmark dataset commonly used for object tracking experiments. Compared with the OTB-2013 dataset, the OTB-2015 dataset is a more challenging dataset that includes 100 videos, and also has eleven various attributes. On this very challenging dataset, we do the experiments with other nine trackers: ECO-HC [9], BACF [33], LMCF [34], DSST [7], LCT [35], SAMF [36], Staple [30], SRDCF [8] and KCF [6].

Figure 5 shows the tracking accuracy of each tracker under the three attributes of BC (contains 31 video sequences), MB (contains 29 video sequences) and OV (contains 14 videos). According to the figure, the proposed tracker achieved the highest tracking accuracy among the three attributes, the AP and the AUC under these three attributes were 88.6% and 83.7%, 85.3% and 81.0%, 85.0% and 73.4%. The experiments further verify that the proposed tracker can solve the problem of blurred background and rapid change.

Figure 6 shows the average accuracy of ten trackers on the OTB-2015 dataset. From the figure, we can see that, among these 10 trackers, the proposed tracker performs best with the AP of 87.8% and the AUC of 80.9% respectively. Besides, this section also evaluates the 10 trackers in 11 attributes. From the Table 3, the proposed tracker has the highest AP among 10 attributes. From Table 4, the proposed tracker achieved the highest AUC among all 10 attributes. In summary, the proposed tracker has significantly improved tracking accuracy, robustness and real-time performance. Besides, we show a qualitative experiment results on the OTB-2015 dataset with four representative video selected sequences. The results are shown in Figure 7. We select a few representative frames of each data so the selection is meaningful. The following video paragraphs are representative, highlighting the object tracking for changes in form, interference from obstructions, changes in light. These factors are particularly important in the marine environment. The proposed tracker has more excellent performance than other nine trackers in the perspective of position estimation and scale estimation, which can deal with changes in the complex marine environment and make the tracking effect outstanding for marine objects.
The proposed tracker has more excellent performance than other nine trackers in the perspective of the AP and the AUC under these three attributes were 88.6% and 83.7%, 85.3% and 81.0%, 85.0% are representative, highlighting the object tracking for changes in form, interference from obstructions, changes in light. These factors are particularly important in the marine environment.

Table 3.

| Trackers | SV | IV | OPR | OCC | BC | DEF | MB | FM | IPR | OV | LR | AP  | AUC  |
|----------|----|----|-----|-----|----|-----|----|----|-----|----|----|-----|------|
| Ours     | 80.5 | 79.2 | 81.1 | 80.6 | 82.4 | 81.8 | 78.0 | 79.2 | 78.3 | 73.7 | 79.8 | 84.5  |
| BACF     | 76.7 | 80.3 | 77.9 | 73.0 | 80.1 | 76.4 | 76.4 | 78.7 | 79.2 | 74.8 | 74.1 | 81.6  |
| DSST     | 63.3 | 71.5 | 64.4 | 58.9 | 70.4 | 53.3 | 56.7 | 55.2 | 69.1 | 48.1 | 56.7 | 67.9  |
| KCF      | 63.5 | 72.4 | 67.7 | 63.2 | 71.3 | 61.9 | 60.1 | 62.1 | 70.1 | 50.1 | 56.0 | 69.6  |
| SAMF     | 70.1 | 70.8 | 73.9 | 72.2 | 68.9 | 68.0 | 65.5 | 65.4 | 72.1 | 62.8 | 68.5 | 75.1  |
| LCT      | 67.8 | 74.3 | 74.6 | 67.9 | 73.4 | 68.5 | 66.9 | 68.1 | 78.2 | 59.2 | 53.7 | 76.2  |
| Staple   | 71.5 | 78.7 | 73.0 | 72.1 | 76.6 | 74.3 | 70.7 | 69.7 | 77.0 | 66.1 | 63.1 | 77.9  |
| SRDCF    | 74.1 | 78.6 | 74.2 | 73.0 | 77.5 | 72.8 | 76.7 | 76.9 | 74.5 | 59.7 | 65.5 | 79.0  |
| LMCF     | 72.3 | 79.5 | 76.0 | 73.6 | 82.2 | 72.9 | 73.0 | 73.0 | 75.5 | 69.3 | 67.9 | 78.4  |
| EOC-HC   | 80.5 | 79.2 | 81.1 | 80.6 | 82.4 | 81.8 | 78.0 | 79.2 | 78.3 | 73.7 | 79.8 | 84.5  |
| Ours     | 84.3 | 84.0 | 86.5 | 84.8 | 88.6 | 84.3 | 85.3 | 82.5 | 82.3 | 85.0 | 79.0 | 87.8  |

Figure 5. Experimental results show that the proposed tracker performs excellently, particularly in three attributes: (a) The PP of the OPE on attribute BC; (b) The SP of the OPE on attribute BC; (c) The PP of the OPE on attribute MB; (d) The SP of the OPE on attribute MB; (e) The PP of the OPE on attribute OV; (f) The SP of the OPE on attribute OV.

Figure 6. Experimental results for the representative ten trackers represented by ten different color curves: (a) The PP of the OPE on all sequences; (b) The SP of the OPE on all sequence.
we will add a deep-learning neural network, such as the siamese network, to improve the tracker. The siamese neural network breaks the limitation that the tracker based on the deep neural network proposed in this paper, have lost the target, resulting in continuous drift. As part of the next step, see from the experimental results in Figure 8 that, in this case, many trackers, including the tracker, cannot be real time. At the same time, the tracker based on the siamese neural network also has high accuracy.

**4.4. Failure Case**

Although the proposed tracker can achieve outstanding results in most marine environments, it is easy to lose an object in cases of low resolution and large movements at the same time. We can see from the experimental results in Figure 8 that, in this case, many trackers, including the tracker proposed in this paper, have lost the target, resulting in continuous drift. As part of the next step, we will add a deep-learning neural network, such as the siamese network, to improve the tracker, and effectively improve the effect when adapting to scenarios in various complex environments. The siamese neural network breaks the limitation that the tracker based on the deep neural network cannot be real time. At the same time, the tracker based on the siamese neural network also has high robustness. By calculating the degree of correlation between the candidate domain to be detected and the target area, it is determined that the position with the highest similarity value is the predicted position of the object to be tracked. So, we will add the siamese network to our tracker to make the tracking effect more robust.
parallel correlation filters with three different learning rates are applied to improve the robustness of the tracking effect. In addition, a new tracking performance evaluation index is proposed, and the measurement result is used as a reference index for the filter update. The proposed tracker effectively reduces the interference caused by marine environmental factors, which leads to excellent performance of object tracking. Through the experimental results, we can conclude that the tracker proposed in this paper has outstanding attributes in light changes, interference occlusion, fast motion and scale changes. It can effectively solve the target tracking loss caused by interference from waves and

![Figure 7](image_url) Qualitative experimental results from four representative videos. Ten different colors represent representative ten trackers, respectively.

![Figure 8](image_url) Failure case of proposed tracker on complex marine environmental objects. The results for representative five trackers are expressed in different colors.

5. Conclusions

The baseline tracker PCF is easily interfered with by occlusion, strong light and other factors, which leads to tracking failure. In order to track the complex marine environmental objects efficiently, the parallel three-branch correlation filters based on a confidence mechanism is proposed. Three different parallel correlation filters with three different learning rates are applied to improve the robustness of the tracking effect. In addition, a new tracking performance evaluation index is proposed, and the measurement result is used as a reference index for the filter update. The proposed tracker effectively reduces the interference caused by marine environmental factors, which leads to excellent performance of object tracking. Through the experimental results, we can conclude that the tracker proposed in this paper has outstanding attributes in light changes, interference occlusion, fast motion and scale changes. It can effectively solve the target tracking loss caused by interference from waves and
ships, light changes, and fast movement in the complex marine environment. Compared with other trackers, the proposed tracker shows excellent performance with accuracy and success rate. It shows good robustness under the conditions of background changes and its own non-rigid transformation, verifying that the proposed tracker is outstanding in the complex marine environment. However, when the object is blocked and there is a large displacement at the same time, it cannot perform well. Based on the end-to-end feature fusion framework of the siamese network, the framework can effectively fuse CNN features and hand-designed features, solve the problem of parameter learning in feature fusion, and improve the versatility of the target tracker. So, we will try to improve the tracker with the siamese network in the future.

Author Contributions: Y.Z. and S.L. conceived the main idea, designed the main algorithm and wrote the manuscript. S.L. designed the main experiments under the supervision of Y.Z., D.L., W.Z. and Y.Y. The experimental results were analyzed by Y.Z., S.L., W.Z., Y.Y., X.L., S.J. and D.L. provided suggestions for the proposed algorithm. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by [the Fundamental Research Funding for the Central Universities of Ministry of Education of China] grant number [18D110408], [the Special Project Funding for the Shanghai Municipal Commission of Economy and Information Civil-Military Inosculation Project “Big Data Management System of UAVs”] grant number [JMRH-2018-1042], and [the National Natural Science Foundation of China (NSFC)] grant number [18K10454]. The APC was funded by [the Fundamental Research Funding for the Central Universities of Ministry of Education of China].

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Chen, X.; Xu, X.; Yang, Y.; Wu, H.; Tang, J.; Zhao, J. Augmented Ship Tracking Under Occlusion Conditions From Maritime Surveillance Videos. IEEE Access 2020, 8, 42884–42897. [CrossRef]
2. Singh, M.; Khare, S.; Kaushik, B.K. Performance Improvement of Electro-Optic Search and Track System for Maritime Surveillance. Def. Sci. J. 2020, 70, 66–71. [CrossRef]
3. Kalal, Z.; Mikolajczyk, K.; Matas, J. Tracking-learning-detection. IEEE Trans. Pattern Anal. Mach. Intell. 2012, 34, 1409–1422. [CrossRef] [PubMed]
4. Bolme, D.S.; Beveridge, J.R.; Draper, B.A.; Lui, Y.M. Visual Object Tracking Using Adaptive Correlation Filters. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, CA, USA, 13–18 June 2010; pp. 2544–2550.
5. Henriques, J.F.; Caseiro, R.; Martins, P.; Batista, J. Exploiting the Circulant Structure of Tracking-by-Detection with Kernels. Lect Notes Comput Sc 2012, 7575, 70715.
6. Henriques, J.F.; Caseiro, R.; Martins, P.; Batista, J. High-Speed Tracking with Kernelized Correlation Filters. IEEE Trans. Pattern Anal. Mach. Intell. 2014, 37, 583–596. [CrossRef]
7. Danelljan, M.; Hager, G.; Khan, F.S.; Felsberg, M. Discriminative Scale Space Tracking. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 39, 1561–1575. [CrossRef]
8. Danelljan, M.; Hager, G.; Khan, F.S.; Felsberg, M. Learning Spatially Regularized Correlation Filters for Visual Tracking. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7-13 December 2015; pp. 4310–4318.
9. Danelljan, M.; Bhat, G.; Khan, F.S.; Felsberg, M. ECO: Efficient Convolution Operators for Tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 6931–6939.
10. Aftab, W.; Hostettler, R.; De Freitas, A.; Arvaneh, M.; Mihaylova, L. Spatio-Temporal Gaussian Process Models for Extended and Group Object Tracking With Irregular Shapes. IEEE Trans. Veh. Technol. 2019, 68, 2137–2151. [CrossRef]
11. Yang, Y.; Zhang, Y.; Li, D.; Wang, Z. Parallel Correlation Filters for Real-Time Visual Tracking. Sensors 2019, 19, 2362. [CrossRef]
12. Lu, X.; Li, J.; He, Z.; Liu, W.; You, L. Visual object tracking via collaborative correlation filters. Signal Image Video Process. 2020, 14, 177–185. [CrossRef]
13. Fang, Y.; Ko, S.; Jo, G.S. Robust visual tracking based on global-and-local search with confidence reliability estimation. *Neurocomputing* **2019**, *367*, 273–286. [CrossRef]  
14. Tjaden, H.; Schwaneeck, U.; Schömer, E.; Cremers, D. A Region-Based Gauss-Newton Approach to Real-Time Monocular Multiple Object Tracking. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *41*, 1797–1812. [CrossRef] [PubMed]  
15. Yuan, Y.; Chu, J.; Leng, L.; Miao, J.; Kim, B.-G. A scale-adaptive object-tracker with occlusion detection. *Eurasip J. Image & Video Process.* **2020**, *1*, 1–15.  
16. Guo, Y.; Li, Y.; Xue, A.; Tharmarasa, R.; Kirubarajan, T. Simultaneous tracking of a maneuvering ship and its wake using Gaussian processes. *Signal Process.* **2020**, *172*, 107547. [CrossRef]  
17. Zheng, H.; Tang, Y. A novel failure mode and effects analysis model using triangular distribution-based basic probability assignment in the evidence theory. *IEEE Access* **2020**, *8*, 66813–66827. [CrossRef]  
18. Zhang, Q. Using Wavelet Network in Nonparametric Estimation. *IEEE Trans. Neural Netw.* **1997**, *8*, 227–236. [CrossRef] [PubMed]  
19. Mercorelli, P. Biorthogonal wavelet trees in the classification of embedded signal classes for intelligent sensors using machine learning applications. *J. Frankl. Inst.* **2007**, *344*, 813–829. [CrossRef]  
20. Mahdavi-Amiri, N.; Shaeiri, M. A conjugate gradient sampling method for nonsmooth optimization. *A Q. J. Oper. Res.* **2020**, *18*, 73–90. [CrossRef]  
21. Gao, J.; Wang, Q.; Xing, J.L.; Ling, H.B.; Hu, W.M.; Maybank, S. Tracking-by-Fusion via Gaussian Process Regression Extended to Transfer Learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 939–955. [CrossRef]  
22. Gao, Y.; Zhao, J.; Luo, J.; Zhou, H. Adaptive feature fusion with the confidence region of a response map as a correlation filter tracker. *Opt. Precis. Eng.* **2019**, *27*, 1178–1187.  
23. Vojir, T.; Noskova, J.; Matas, J. Robust Scale-Adaptive Mean-Shift for Tracking. *Pattern Recognit. Lett.* **2014**, *49*, 250–258. [CrossRef]  
24. Sun, H.; Chen, X.; Xiao, H. A deep object tracker with outline response map. *CAAI Trans. Intell. Syst.* **2019**, *14*, 725–732.  
25. Li, X.; Zhou, J.; Hou, J.; Zhao, L.; Tian, N. Research on improved moving object tracking method based on ECO-HC. *J. Nanjing Univ. Nat. Sci.* **2020**, *56*, 216–226.  
26. Oron, S.; Bar-Hillel, A.; Levi, D.; Avidan, S. Locally Orderless Tracking. *Int. J. Comput. Vis.* **2014**, *111*, 213–228. [CrossRef]  
27. Han, Y.; Deng, C.; Zhao, B.; Tao, D. State-Aware Anti-Drift Object Tracking. *IEEE Trans Image Process.* **2019**, *28*, 4075–4086. [CrossRef] [PubMed]  
28. Li, B.; Xie, W.; Zeng, W.; Liu, W. Learning to Update for Object Tracking With Recurrent Meta-Learner. *IEEE Trans. Image Process.* **2019**, *28*, 3624–3635. [CrossRef] [PubMed]  
29. Lin, Y.; Chen, J.; Lin, K. Integration of Texture and Depth Information for Robust Object Tracking. In Proceedings of the 2014 IEEE International Conference on Granular Computing (GrC), Noboribetsu, Japan, 22–24 October 2014; pp. 170–174.  
30. Bertinetto, L.; Valmadre, J.; Golodetz, S.; Miksik, O.; Torr, P.H. Staple: Complementary Learners for Real-Time Tracking. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 1401–1409.  
31. Wu, Y.; Lim, J.; Yang, M.H. Online Object Tracking: A Benchmark. In Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013; pp. 2411–2418.  
32. Wu, Y.; Lim, J.; Yang, M.-H. Object Tracking Benchmark. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1834–1848. [CrossRef] [PubMed]  
33. Galoogahi, H.; Fagg, A.; Lucey, S. Learning background-aware correlation filters for visual tracking. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 1144–1152.  
34. Wang, M.; Liu, Y.; Huang, Z. Large Margin Object Tracking with Circulant Feature Maps. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 4800–4808.
35. Ma, C.; Yang, X.; Zhang, C.; Yang, M.H. Long-Term Correlation Tracking. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 5388–5396.
36. Li, Y.; Zhu, J. A Scale Adaptive Kernel Correlation Filter Tracker with Feature Integration. In Proceedings of the IEEE European Conference on Computer Vision Workshops (ECCVW), Zurich, Switzerland, 6–12 September 2014; pp. 254–265.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).