

Research Article

Moving Object Tracking in Satellite Videos by Kernelized Correlation Filter Based on Color-Name Features and Kalman Prediction

Wenjing Pei  and Xuhui Lu 

The Seventh Research Division and The Center for Information and Control, School of Automation Science and Electrical Engineering, Beihang University (BUAA), Beijing 100191, China

Correspondence should be addressed to Xuhui Lu; luxuhui2020@buaa.edu.cn

Received 21 April 2022; Accepted 21 May 2022; Published 9 June 2022

Academic Editor: Maode Ma

Copyright © 2022 Wenjing Pei and Xuhui Lu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper studies moving object tracking in satellite videos. For the satellite videos, the object size in the images may be small, the object may be partly occluded, and the image may contain an area resembling dense objects. To handle the above problems, this paper puts forward a kernelized correlation filter based on the color-name feature and Kalman prediction. The original image is mapped to the color-name feature space so that the tracker can process the image with multichannel color features. The Kalman filter is used to predict the moving object position in the tracking process, and the detection area is determined according to the predicted position. The Kalman filter is updated with the detection results to improve the tracking accuracy. The proposed algorithm is tested on Jilin-1 datasets. Compared with the other seven tracking algorithms, the experiment results show that the proposed algorithm has stronger robustness for several complex situations such as rapid target motion and similar object interference. Besides, it is also shown that the proposed algorithm can prevent the problem of tracking failure when the moving object is partially occluded.

1. Introduction

With the development of remote sensing technologies, the earth observation satellites are extensively applied to several fields [1–7]. However, traditional earth observation satellites can only take a single image of a certain area. The valuable and interesting dynamic data in the object area is hard to obtain only based on the static medium and high resolution optical remote sensing data, which may limit the reconnaissance capability of the earth observation satellites in emergencies [8].

On the other hand, video satellites can overcome the limitation of traditional earth observation satellites and can obtain high time-resolution images [9–11]. Video satellites are used to obtain a series of images of fixed object areas on the ground. Therefore, the videos can be formed to obtain the dynamic object information directly. Currently, the continuous monitoring capability of high-resolution sat-

ellites in a certain time range has been realized [12]. Due to the above advantages, video satellites are used to observe and track the states of moving objects on the ground [13] and have wide application potential in the fields of vehicle real-time monitoring [8, 14, 15], rapid response to natural disaster emergency [16], major engineering monitoring [17], and so on. In recent years, some representative video satellites are Skysat-1 and Skysat-2 in the United States [18, 19], TUB-SAT series satellites by the Technical University of Berlin [20] and Jilin-1 [7, 21] and Tiantuo-2 (TT-2) [22] in China.

Besides, moving object tracking (MOT) in traditional videos has been a research hotspot in computer vision. MOT is widely used in automatic monitoring, automatic driving, human-computer interaction, and so on [23, 24]. The task of MOT is to predict the size and position of the object in subsequent frames based on the size and position of the object in the initial frame. Much research has been done in this aspect, and numerous algorithms have been

investigated for accurate tracking in ordinary videos [25–30]. Notice that the commonly used visual object tracking methods can be divided into generation methods [25–27] and discrimination methods [28–30].

On the one hand, the generative methods establish the object model and describe the real-world object, so as to search for the position with the highest similarity with the object template in the next frame [31]. Many breakthroughs have been made and several generative methods have been utilized in MOT, including mean shift, sparse coding, dictionary learning, particle filter, and sliding window [32–37]. On the other hand, compared with generative methods, the discriminative methods regard object tracking as a detection problem, also known as tracking by detection (see [38, 39] and the related references). The discriminative methods generally train the classifier in the first frame to separate the object from its surrounding background. In particular, considering the complex background such as background changes, the discriminative methods establish an online discrimination classifier model to distinguish objects from cluttered backgrounds to provide more effective features and avoid unwanted model drift. Recently, some representative machine learning techniques are adopted into the discriminative methods, such as Boosting, Support Vector Machines (SVM), Multiple Instance Learning (MIL), Random Forests, Semisupervised Learning, and Structured Output Support Vector Machines (SOSVM) [40–42]. A theoretical framework of dense sampling in tracking-by-detection is presented [43]. In [44], a tracking learning detection (TLD) algorithm is also proposed, where learning and detection are introduced into the long-term tracking of the objects in the videos to enhance the tracking accuracy. Hare et al. [45] propose an adaptive tracking-by-detection method called STRUCK, which exploits Gaussian kernels and SVM as a structured output to accurately locate the objects. A cooperative model tracking algorithm based on sparse representation is presented, which is suitable for the situation of object occlusion and blur [46].

Apart from the machine-learning-based discriminative methods, the kernelized correlation filter (KCF) can also accelerate the calculation speed and improve tracking accuracy simultaneously [43, 47, 48]. In particular, the KCF-based method can efficiently handle the object in changing environment [47]. A scale adaptive tracker is proposed based on the separate discriminative correlation filters, where the computational cost is reduced [49]. Galoogahi et al. [50] demonstrate a background-aware correlation filter (BACF) for real-time visual tracking, where the handcrafted features are introduced to effectively describe the changing background. In [48], a KCF and the Normalized Cross-Correlation (NCC) template matching is proposed for long-term target tracking of UAVs to improve the tracking performance.

In addition, note that the moving objects in satellite videos include vehicles, airplanes, rockets, and ships. Compared with moving target tracking in ordinary videos, real-time object tracking in satellite videos should overcome three main challenges as follows [51, 52].

- (1) Compared with the ordinary videos on the ground, the object size is small in satellite videos, which may lose their effective features (see Figure 1). For instance, the length of a car is about 4–5 m in real life, while it is about four pixels in satellite videos. The size of these objects in the satellite videos is so small that it may be difficult to track these objects reliably
- (2) Since remote sensing images have a relatively large field of vision, there is low contrast between the background and the objects (see Figure 1(a))
- (3) The background of the image sequence of the satellite videos may be fuzzy and chaotic (see Figure 1(b)). Besides, there may be more than one moving target in the image sequence, which have the characteristics of high similarity, serious mutual interference, and low resolution. This will result in partial or complete occlusion between moving targets (see Figure 1(c))

Due to these above challenges and the increasing demand for MOT in the field of remote sensing, up until now, several moving object detection/tracking algorithms have been carefully designed for MOT in satellite videos [51–55]. Lei and Guo [53] propose a road masking and Gaussian mixture method to achieve multiple object detection and tracking of the remote sensing video satellite. Meanwhile, the method can improve the reconnaissance capability of the remote sensing satellite for the dynamic mobile small target. In [51], a fusion tracker is introduced where the kernel correlation filter and the three-frame-difference method are synthesized for satellite videos to improve the performance of the tracker. Li and Man [52] put forward an optical-flow-based detection algorithm with video attention saliency for the moving ships in the satellite videos, where the Gabor filter is utilized to extract the texture feature, and the registration of the sea-scene images can be avoided. Liu et al. [54] present a kernelized correlation filter where multifeature fusion and motion trajectory compensation are employed for satellite videos to mitigate the tracking drifts. In [55], an object tracking algorithm is also proposed for the high-resolution multispectral satellite images with multiangular observation capability, where a novel regional operator is constructed and the tracking capability is verified in the WorldView-2 satellite images.

However, the above trackers [51, 52, 54, 55] generally rely on the original pixel information (such as HSV or HOG), which ignores the color information. In fact, several satellite videos contain some color-name (CN) features. Compared with other color features, CN features show better discriminative capability in MOT [56, 57]. In addition, due to undesirable environmental factors, such as shadows, similar backgrounds, and other interferences, it will become more complicated to realize the MOT.

On this foundation, aiming at MOT in satellite videos, this paper proposes the kernelized correlation filters based on color-name features and Kalman prediction (CNK-KCF). The images are mapped to CN feature space so that the tracker can process the image with multichannel color

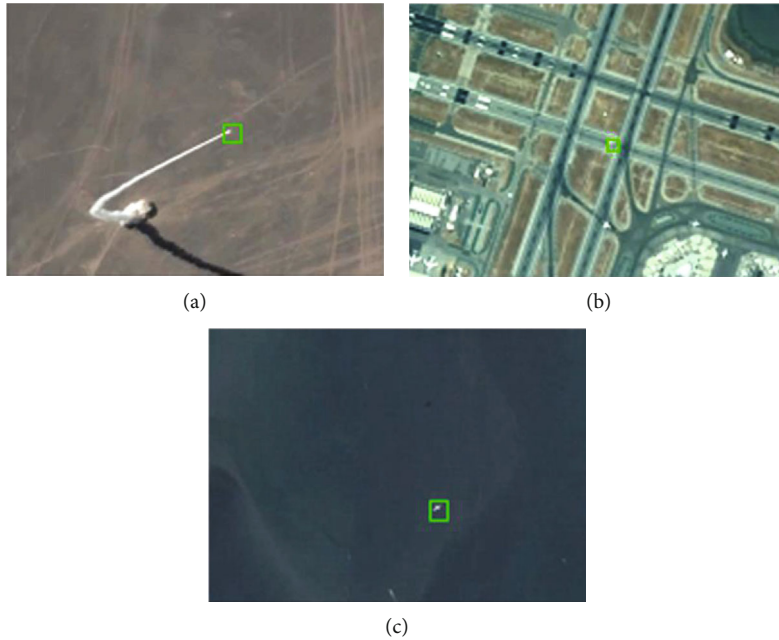


FIGURE 1: The moving objects in satellite videos. (a) There is low contrast between the background and the object. (b) The background is fuzzy and chaotic. (c) The target is partially occluded.

features. Moreover, the KCF can improve the calculation speed in the Fourier translation based on the cyclic matrix and kernel trick. Meanwhile, the Kalman filter can help correct and update the predicted position of the moving object and, together with the kernelized correlation filters based on the color-name features (CN-KCF), can improve the tracking accuracy. The proposed algorithm is tested on Jilin-1 datasets. Experimental results are analyzed and show that the proposed method is robust to some environmental factors such as partial occlusion, background similarity, and rapid motion.

In summary, the contributions of this paper are twofold.

- (1) For MOT in satellite videos, a framework named CNK-KCF is carefully designed based on the CN feature and the Kalman filter. Besides, the proposed CNK-KCF algorithm is stable in complex situations such as rapid target motion, occlusion, and similar object interference, which can solve the problem of tracking failure when a moving object is partially occluded
- (2) In the experiment section, the CNK-KCF algorithm is compared with other algorithms, and it is shown that the CNK-KCF algorithm possesses better tracking accuracy and success rate for the airplane, rocket, and ship in Jilin-1 satellite videos. The performance of each type of tracker in satellite videos is analyzed in detail

The rest of this paper is organized as follows. Section 2 presents the design of the CNK-KCF for satellites videos MOT. Section 3 introduces the experiment results and some analysis. Finally, Section 4 concludes this article.

2. Materials and Methods

In this section, to solve the problem of MOT in satellite videos, the CNK-KCF is developed carefully. As shown in Figure 2, the CNK-KCF mainly consists of 3 parts: (1) KCF [58, 59], (2) CN, and (3) KF prediction. In the rest of this section, each part is described in one subsection each.

2.1. Kernelized Correlation Filter (KCF). The KCF has a relatively low computational cost and relatively high tracking accuracy, especially for rapid deformation. KCF is also robust to illumination changes. Hence, KCF is suitable for MOT in satellite videos. The procedures of the KCF tracking algorithm are shown as follows. Firstly, in order to construct the tracking area, the tracking object is selected from the initial frame in the satellite videos. Then, according to the cyclic matrix theory, the tracking area is cyclically shifted. The kernel function is applied to calculate the similarity between the possible region of the target location and the tracking objects. Finally, the area with the largest output response is selected as the new target, and the classifier is trained based on the Fourier transform to reduce the calculation time.

In the KCF, the following regression function

$$f(x) = \omega^T z \quad (1)$$

is trained to obtain weight coefficients $\omega = [\omega^1, \omega^2, \dots, \omega^n]^T$, where $z = [z^1, z^2, \dots, z^n]^T$ is an n -dimensional vector. Correspondingly, the cost function can be minimized as

$$f(x) = \min_{\omega} (X\omega - y)^T (X\omega - y) + \lambda \|\omega\|^2, \quad (2)$$

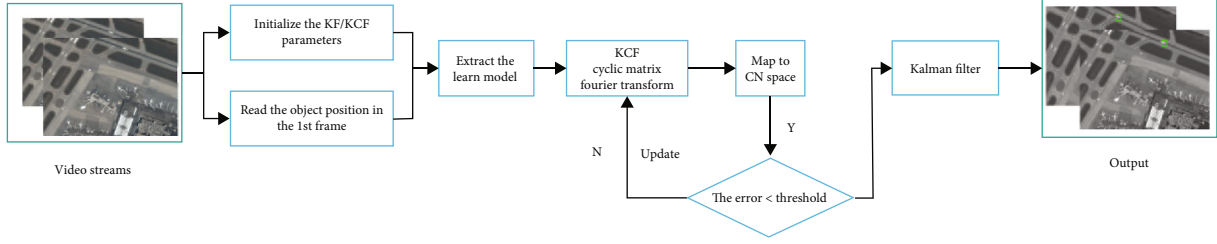


FIGURE 2: The flow charts of our proposed tracker (CNK-KCF).

where X is the data matrix, y is the desired output, and λ is the regularization parameter to prevent overfitting. Based on [64, 65], the extreme value of Equation (2) can be obtained as

$$\omega = (X^T X + \lambda I)^{-1} X^T y, \quad (3)$$

where I is an identity matrix. In Equation (3), calculating the inverse matrix $((X^H X + \lambda I)^{-1})$ is very time-consuming. Therefore, the calculation is performed in the Fourier domain, and Equation (3) can be rewritten in a complex field as

$$\omega = (X^H X + \lambda I)^{-1} X^H y, \quad (4)$$

where X^H represents the Hermitian transpose of X . Obviously, if X is a real matrix, Equation (4) can be considered as equivalent to Equation (3).

Besides, in order to accelerate the calculation speed of MOT, the cyclic shift is also introduced. The cyclic shift operator is a permutation matrix, which can be used to simulate the one-dimensional translation of this vector. The permutation matrix can be shown as

$$P = \begin{bmatrix} 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad (5)$$

so the cyclic shift of x can be presented as

$$Px = [x^n, x^1, \dots, x^{n-1}]^T. \quad (6)$$

Since the product Px shifts x by one element, u shifts are used to chain with the matrix power $P^u x$ and achieve more translations. The same signal x can be obtained periodically every n -time translation based on the cyclic property of the cyclic matrix, which means that all shifted signals can be represented as

$$\{P^u x \mid u = 0, 1, 2, \dots, n-1\}. \quad (7)$$

Correspondingly, the data matrix X can be the circulant

matrix and is denoted as

$$X = C(x) = \begin{bmatrix} x^1 & x^2 & x^3 & \dots & x^n \\ x^n & x^1 & x^2 & \dots & x^{n-1} \\ x^{n-1} & x^n & x^1 & \dots & x^{n-2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ x^2 & x^3 & x^4 & \dots & x^1 \end{bmatrix}. \quad (8)$$

One-dimensional vector cyclic displacement is given in Figure 3. All cyclic matrices can be diagonalized through the discrete Fourier transform (DFT), which is independent of the generated vector X . Therefore, X can be diagonalized as

$$X = C(x) = F \text{diag}(\hat{x}) F^H, \quad (9)$$

where the constant matrix F is known as the discrete DFT that does not depend on x . The Hermitian transpose of F is represented as F^H . The matrix $\text{diag}(\bullet)$ is the diagonal matrix. Accordingly, the vector \hat{x} is the DFT of x and is defined as

$$\hat{x} = F(x) = \sqrt{n} F x. \quad (10)$$

In the following section, the DFT of the vector will be represented by the hat ($\hat{\bullet}$). Due to the diagonalization property of the matrix $\text{diag}(\bullet)$, the matrix X^H can be obtained as

$$\hat{x} = X^H = (X^*)^T = (F^* \text{diag}(\hat{x}^*) F^{H*})^T = F \text{diag}(\hat{x}^*) F^H, \quad (11)$$

where X^* is the complex-conjugate of X and $(\bullet)^*$ can be defined as a conjugate symbol. In addition, $X^H X$ is represented to be a noncentral covariance matrix. From (9) and (11), it follows that

$$X^H X = F \text{diag}(\hat{x}^*) F^H F \text{diag}(\hat{x}) F^H. \quad (12)$$

Because the diagonal matrices are symmetric, the Hermitian transpose is used only after complex-conjugation \hat{x} . In this way, the factor $F^H F = I$ can be eliminated. Moreover, the operations on diagonal matrices are done by elements,

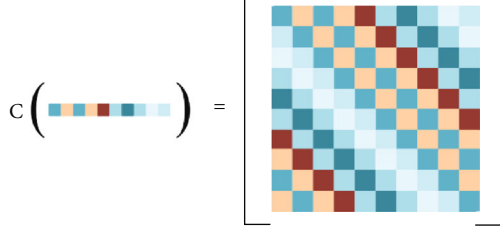


FIGURE 3: One-dimensional vector cyclic displacement.

so Equation (12) can be rewritten as

$$X^H X = F \text{diag}(\hat{x}^*) \text{diag}(\hat{x}) F^H = X^H X = F \text{diag}(\hat{x}^* \odot \hat{x}) F^H, \quad (13)$$

where \odot is the dot product of two vectors. It can be seen in (13) that the original complex matrix operation is transformed into a simple vector and dot product operation based on the diagonalization property of the cyclic matrix. Based on (4) and (13), the DFT of ω is obtained as

$$\hat{\omega} = \text{diag} \left(\frac{\hat{x}^*}{\hat{x}^* \odot \hat{x} + \lambda} \right) \hat{y} = \frac{\hat{x}^* \odot \hat{y}}{\hat{x}^* \odot \hat{x} + \lambda}, \quad (14)$$

where fractions represent the division element. ω can be recovered in the spatial domain on the basis of the inverse DFT.

In addition, a nonlinear mapping function $\varphi(x)$ is employed to make the mapped samples linearized in the new space. Therefore, $f(x)$ can be transformed into

$$f(x) = \omega^T \varphi(x). \quad (15)$$

Furthermore, the kernel technique can be used to map the input of a linear problem to a nonlinear feature space $\varphi(x)$. Correspondingly, the solution ω is rewritten as

$$\omega = \sum_i \alpha_i \varphi(x^i), \quad (16)$$

where x^i is a column vector. Therefore, the parameters of the solution are changed from ω to α . Meanwhile, Equation (2) can be rewritten as

$$\min_{\omega} \sum_i (\varphi(X)\omega - y)^T (\varphi(X)\omega - y) + \lambda \|\omega\|^2. \quad (17)$$

Besides, the dot products can be defined as $\varphi^T(x) \odot \varphi(x')$ = $\kappa(x, x')$, where $\kappa(x, x')$ is the Gaussian kernel expressed as

$$\kappa(x, x') = \exp \left(-\frac{1}{\sigma^2} \left(\|x\|^2 + \|x'\|^2 - 2\mathcal{F}^{-1}(\hat{x} \odot \hat{x}') \right) \right), \quad (18)$$

where \mathcal{F}^{-1} is a mapping function in the high dimensional space. Furthermore, the dot products $\varphi^T(x) \odot \varphi(x') = \kappa(x,$

$x')$ between all pairs of samples are stored in a $n \times n$ kernel matrix $K \triangleq [K_{ij}]_{n \times n}$, which can be defined as

$$K_{ij} = \kappa(x_i, x_j). \quad (19)$$

The complexity of the regression function increases with the increase of the sample size, so that $f(X)$ can be rewritten based on (16) as

$$f(X) = K^T \alpha. \quad (20)$$

Meanwhile, Equation (17) can be rewritten as

$$\min_{\omega} (K^T \alpha - y)^T (K^T \alpha - y) + \lambda \alpha^T K \alpha. \quad (21)$$

The optimal solution of (21) is given by

$$\alpha = (K + \lambda I)^{-1} y. \quad (22)$$

If the kernel function satisfies Equation (22), K is circulant for any permutation matrix M as

$$\kappa(x, x') = \kappa(Mx, Mx'). \quad (23)$$

In addition, if the kernels can make K circulant, Equation (22) can be diagonalized as

$$\hat{\alpha} = \frac{\hat{y}}{\hat{k}^{xx} + \lambda}, \quad (24)$$

where $k^{xx'}$ is the first row of the kernel matrix K defined as

$$k^{xx'} = [k(x_1, x_1), k(x_1, x_2), \dots, k(x_1, x_n)]. \quad (25)$$

Furthermore, K is a circulant matrix and can be expressed as

$$K = C(k^{xx}), \quad (26)$$

so it further follows that

$$f(X) = C(k^{xx})^T \alpha = \left(F \text{diag}(\hat{k}^{xx}) F^H \right)^T \alpha, \quad (27a)$$

$$\hat{f}(X) = \text{diag}(\hat{k}^{xx}) \odot \alpha. \quad (27b)$$

2.2. KCF Tracking Algorithm Based on Color-Name (CN) Feature. In the traditional KCF algorithm, the original pixel and the directional gradient histogram (HOG) feature are required, and the original pixel features are utilized to convert the images into gray images. The pixel gray value is regarded as the image feature. However, the generation process of the HOG feature descriptor is lengthy, which results in slow speed and poor real-time performance for MOT in satellite videos. In addition, the HOG feature descriptor is relatively sensitive to noise, so it is difficult to deal with occlusion based on the gradient characteristics. Compared



FIGURE 4: Overview of the MOT in our experiments. There are six scenes of the satellite videos. In every scene, there is only one object. (a) The moving airplane-1 is selected as the object; (b) the moving airplane-2 is selected as the object; (c) the moving airplane-3 is selected as the object; (d) the moving airplane-4 is selected as the object; (e) the moving rocket is selected as the object; (f) the moving ship is selected as the object.

TABLE 1: The size of the objects in satellite videos.

The object	Airplane-1	Airplane-2	Airplane-3	Airplane-4	Rocket	Ship
Size (pixels)	44×44	28×28	28×28	22×22	46×46	30×30

TABLE 2: The precision and success rate in comparison with other seven trackers for airplane-1.

Methods	KCF	K-KCF	CN-KCF	CSK	STRUCK	MeanShift	CamShift	CNK-KCF (ours)
Precision	0.929	0.888	0.907	0.943	0.889	0.621	0	0.957
Success rate	0.873	0.816	0.841	0.895	0.827	0.495	0.020	0.917

TABLE 3: The precision and success rate in comparison with other seven trackers for airplane-2.

Methods	KCF	K-KCF	CN-KCF	CSK	STRUCK	MeanShift	CamShift	CNK-KCF (ours)
Precision	0.732	0.284	0.939	0.955	0.272	0.932	0	0.967
Success rate	0.632	0.260	0.856	0.881	0.224	0.824	0.0003	0.902

TABLE 4: The precision and success rate in comparison with other seven trackers for airplane-3.

Methods	KCF	K-KCF	CN-KCF	CSK	STRUCK	MeanShift	CamShift	CNK-KCF (ours)
Precision	0.963	0.952	0.959	0.958	0.977	0.220	0	0.981
Success rate	0.889	0.866	0.878	0.873	0.925	0.186	0.010	0.935

TABLE 5: The precision and success rate in comparison with other seven trackers for airplane-4.

Methods	KCF	K-KCF	CN-KCF	CSK	STRUCK	MeanShift	CamShift	CNK-KCF (ours)
Precision	0.691	0.848	0.973	0.973	0.970	0.936	0	0.974
Success rate	0.506	0.632	0.917	0.916	0.890	0.803	0.0003	0.918

TABLE 6: The precision and success rate in comparison with other seven trackers for the rocket.

Methods	KCF	K-KCF	CN-KCF	CSK	STRUCK	MeanShift	CamShift	CNK-KCF (ours)
Precision	0.962	0.950	0.980	0.889	0.960	0.050	0	0.980
Success rate	0.903	0.879	0.941	0.822	0.898	0.075	0.048	0.941

TABLE 7: The precision and success rate in comparison with other seven trackers for the ship.

Methods	KCF	K-KCF	CN-KCF	CSK	STRUCK	MeanShift	CamShift	CNK-KCF (ours)
Precision	0.640	0.698	0.975	0.978	0.962	0.458	0.279	0.979
Success rate	0.563	0.625	0.948	0.954	0.931	0.426	0.251	0.954

TABLE 8: The basic parameters of the algorithms.

Methods	σ	λ	η
KCF	0.5	0.0001	0.02
K-KCF	0.5	0.0001	0.02
CN-KCF	0.2	0.01	0.075
CSK	0.2	0.01	0.075
STRUCK	0.2	0.01	0.075
MeanShift	0.2	0.01	0.075
CamShift	0.2	0.01	0.075
CNK-KCF (ours)	0.2	0.01	0.075

with the original pixel feature and HOG feature, the CN feature has better stability properties. As a result, the proposed algorithm applies a CN statistical feature to extract the image feature.

CN feature space is a special color space based on a potential probability model. CN space contains 11 color channels: yellow, red, black, blue, gray, pink, white, brown, green, orange, and purple. In the form of mathematical description, a color image $M(x)$ represents the color pixel value at position x , and the image is mapped to CN space so that $M(x)$ can be converted into an 11-dimensional (11D) probability feature vector $f(x)$. Specifically, the RGB value is represented by an 11D color with a total probability sum of 1, so as to realize the low-dimensional extraction of

color information. The model can be expressed as follows:

$$M(x_0) = \arg \max_i \left\{ i \mid \sum_{x \in \Omega_c(x_0)} \phi_i(x) \cdot N(x_0, \sigma), i = 1, 2, \dots, 11 \right\}, \quad (28)$$

where Ω_c presents a region with x_0 as the center and c as the radius. $N(\bullet)$ is a Gaussian function. σ is the standard deviation. However, in the process of tracking, since not all of useful object information can be provided by the 11D color attributes, the 11D color attributes are firstly reduced to 10D. Subsequently, the PCA is employed to reduce 10D to 2D, which can reduce calculation and accelerate the calculation speed of the algorithm.

We supposed that x_{CN}^k is the CN feature extracted from the target region in the k th frame and \tilde{x}_{CN}^k is the Fourier transform of x_{CN}^k . To reduce dimensions, the dimension reduction matrix is given as

$$\tilde{x}_{CN}^k = B_k \tilde{x}_{CN}^k, \quad (29)$$

where \tilde{x}_{CN}^k is the CN feature through the dimension-reduction operation in the k th frame and B_k is a dimension-reduction matrix. The reconstructed minimum cost function as the decision function is given to obtain the

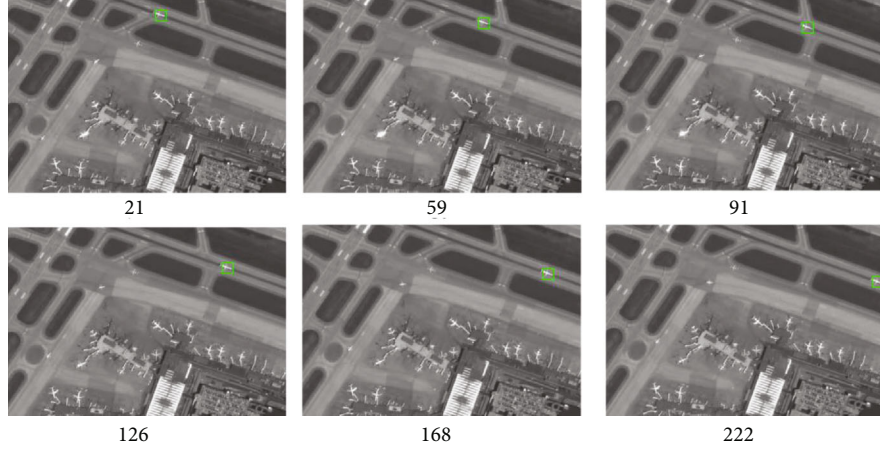


FIGURE 5: The results of moving airplane-1 tracking by CNK-KCF.

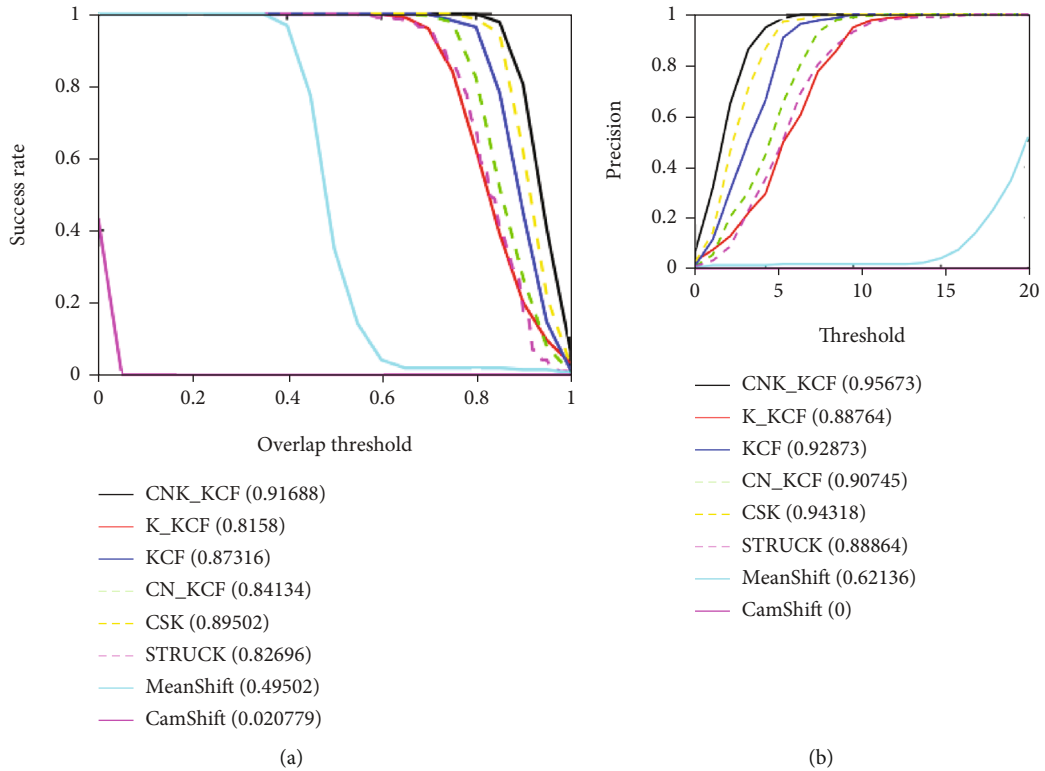


FIGURE 6: The success and precision plots in comparison with the other seven trackers: (a) the success plot for airplane-1; (b) the precision plot for airplane-1.

dimension-reduction matrix B_k as

$$\tilde{x}_{CN}^k = \min_{B_k} \left(\left(\alpha_k \eta_{\text{DATA}} + \sum_{i=1}^{k-1} \alpha_k \eta_{\text{SMOOTH}}^i \right) \right), \quad (30)$$

where α_k and α_i are weight coefficients and η_{DATA} can be used for the solution of the dimension reduction matrix B_k in Equation (29). Meanwhile, due to the poor discriminant performance in this process, η_{SMOOTH}^i is introduced to increase the robustness of B_k in (30), where the first i th

frames are added for training. The forms of η_{DATA} and η_{SMOOTH}^i are

$$\eta_{\text{DATA}} = \frac{1}{MN} \sum_{m,n} \left\| x_{CN}^k(m,n) - B_k B_k^T \hat{x}_k(m,n) \right\|^2, \quad (31a)$$

$$\eta_{\text{SMOOTH}}^i = \sum_{p=1}^s \lambda_i^p \left\| b_i^p - B_k B_k^T b_i^p \right\|^2. \quad (31b)$$

In Equations (31a) and (31b), x_{CN}^k represents the CN



FIGURE 7: The results of moving airplane-2 tracking by CNK-KCF.

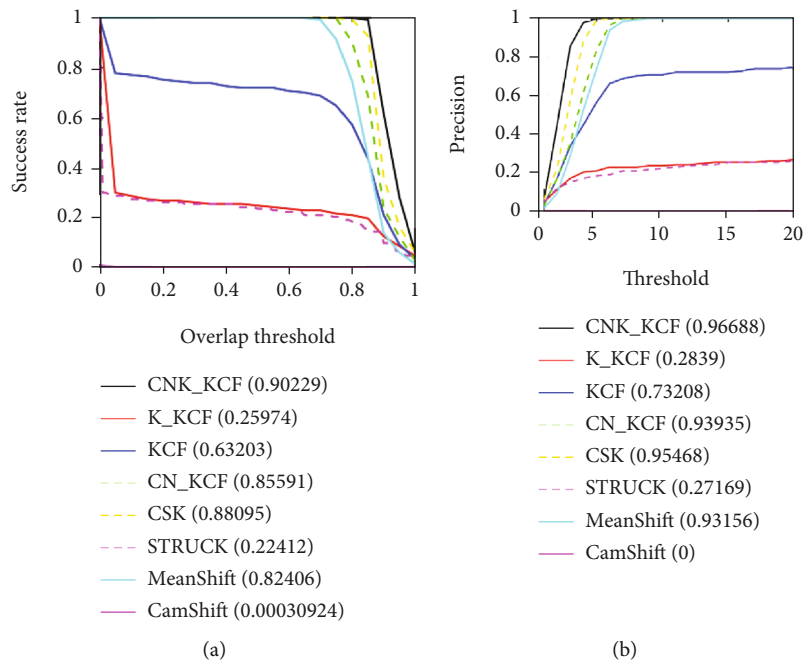


FIGURE 8: The success and precision plots in comparison with the other seven trackers: (a) the success plot for airplane-2; (b) the precision plot for airplane-2.

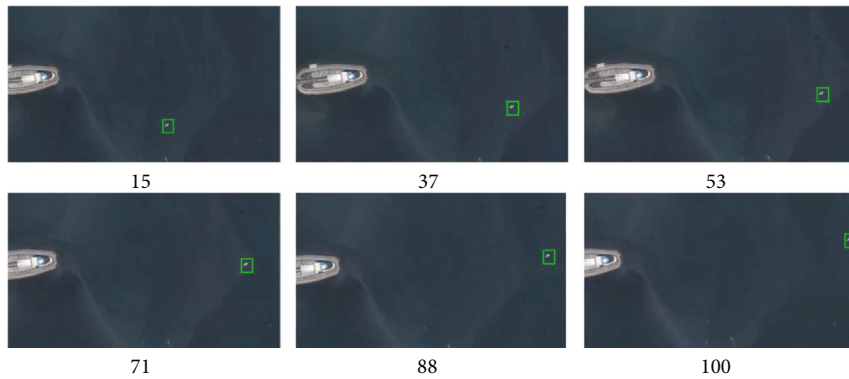


FIGURE 9: The results of moving airplane-3 tracking by CNK-KCF.

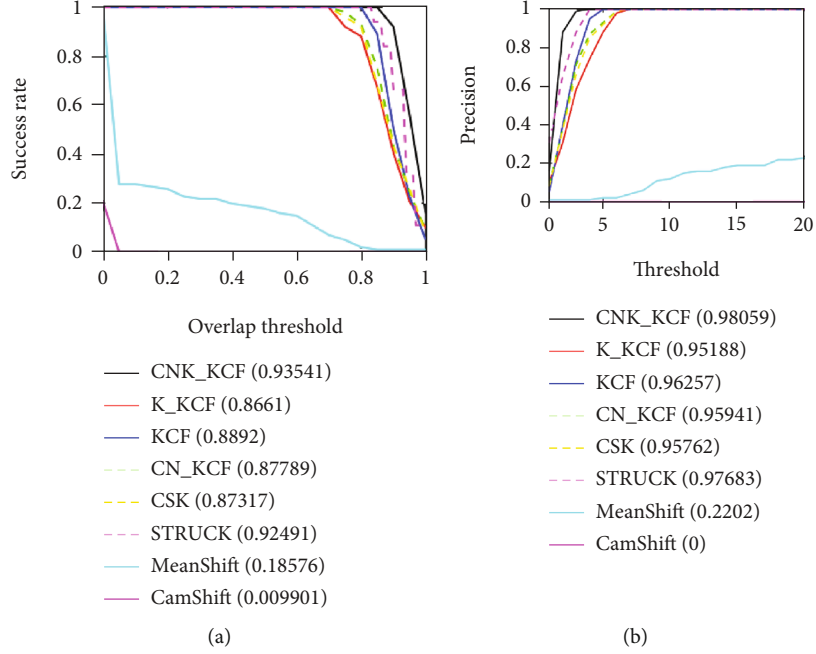


FIGURE 10: The success and precision plots in comparison with the other seven trackers: (a) the success plot for airplane-3; (b) the precision plot for airplane-3.

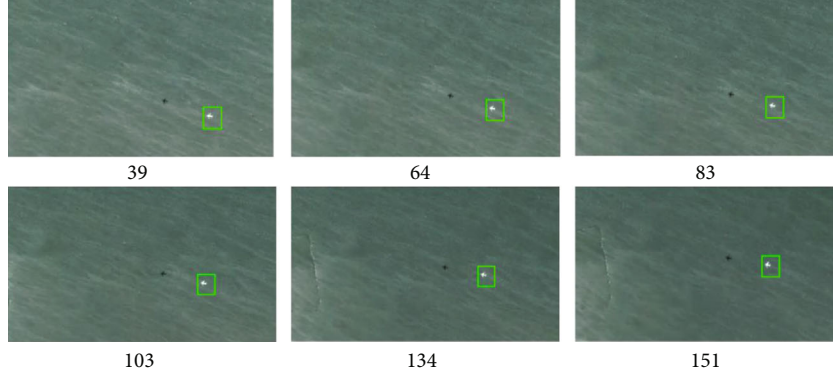


FIGURE 11: The results of moving airplane-4 tracking by CNK-KCF.

features in the k th frame, s is the total number of basis vectors, and λ_i^p is a nonnegative weight. Accordingly, based on Equations (31a) and (31b), Equation (30) can be rewritten as

$$\tilde{x}_{CN}^k = \min_{B_k} \left(\left(\alpha_k \frac{1}{MN} \sum_{m,n} \left\| x_{CN}^k(m,n) - B_k B_k^T \tilde{x}_k(m,n) \right\|^2 + \sum_{i=1}^{k-1} \alpha_k \sum_{p=1}^s \lambda_i^p \| b_i^p - B_k B_k^T b_i^p \|^2 \right) \right). \quad (32)$$

Subsequently, based on (30), B_k obtained from (32) and the values \tilde{x}_{CN}^k and \tilde{x}_{CN}^k can be updated as

$$\tilde{x}_{CN}^k = B_k \tilde{x}_{CN}^k = B_k \left[(1 - \gamma) \tilde{x}_{CN}^{k-1} + \gamma \tilde{x}_{CN}^k \right], \quad (33b)$$

$$\tilde{x}_{CN}^k = (1 - \gamma) \tilde{x}_{CN}^{k-1} + \gamma \tilde{x}_{CN}^k, \quad (33a)$$

where γ is a learning rate parameter.

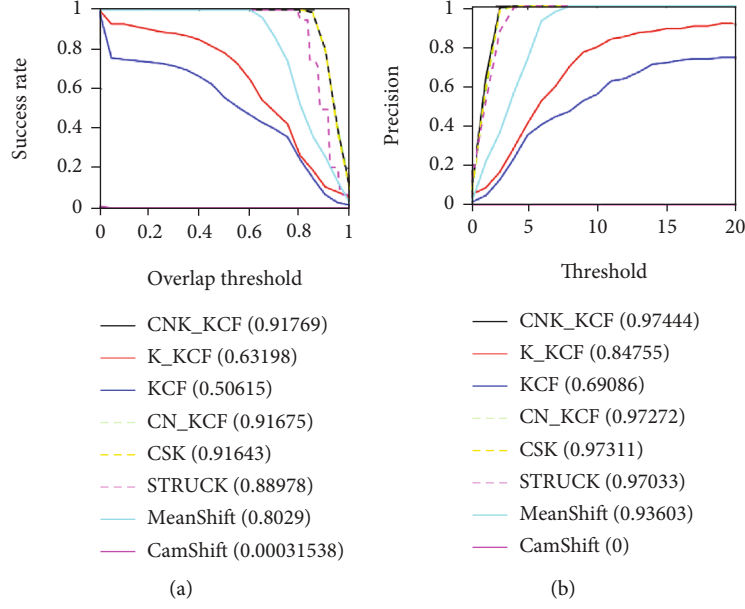


FIGURE 12: The success and precision plots in comparison with the other seven trackers: (a) the success plot for airplane-4; (b) the precision plot for airplane-4.

2.3. Motion Estimation by Kalman Filter. The Kalman filter is initialized before tracking, and the initial state vector containing the manually labeled real coordinate value of the target center and the velocity component on the coordinate axis is obtained. The state equation and observation equation of the Kalman filtering algorithm are, respectively, shown as follows:

$$U_t = A_{t,t-1}U_{t-1} + W_{t-1}, \quad (34a)$$

$$Z_t = H_t U_t + V_t, \quad (34b)$$

where U_t is the state vector of the system at time t . Z_t is the observation vector at time t . $A_{t,t-1}$ is a state transition matrix defined as

$$A_{t,t-1} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (35)$$

and H_t is an observation matrix. Let the initial state $U_0 = [x_0, y_0, x'_0, y'_0]$, where (x_0, y_0) is the initial coordinate of the target center point, and (x'_0, y'_0) is the velocity component at x -axis and y -axis. Both process noise W_{t-1} and observation noise V_t are white noise sequences with the mean value of 0, which are uncorrelated.

The current state matrix and covariance are used to predict the speed and position of the target in the next frame according to the recursive estimation principle. Finally, the prediction equation is obtained:

$$U_{t,t-1} = A_{t,t-1}U_{t-1} + W_{t-1}, \quad (36a)$$

$$P_{t,t-1} = A_{t,t-1}P_{t-1}A_{t,t-1}^T + Q_t, \quad (36b)$$

where $U_{t,t-1}$ is the state prediction vector, $P_{t,t-1}$ is the covariance matrix, Q_t is the covariance matrix of the process noise W_t , and Δt is time interval that is usually taken as 1.

After the predicted coordinates are obtained, the target sampling area is expanded to make the sampling area 3.5 times that of the target image, and the CN features of the model are extracted. Samples are constructed using a cyclic matrix. At the same time, the Fourier transform is performed. The Gaussian kernel is also obtained based on the properties of the cyclic matrix and (18).

Combined with the new actual observations and the prior estimates obtained in the previous step, an a posteriori estimate is obtained using the feedback method. The correction equation is given as

$$P_t = (I - K_t H_t)P_{t,t-1}, \quad (37a)$$

$$K_t = P_{t,t-1}H_t^T (H_t P_{t,t-1}H_t^T + R_t)^{-1}, \quad (37b)$$

$$U_t = U_{t,t-1} + K_t(Z_t - H_t U_{t,t-1}), \quad (37c)$$

$$H_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad (37d)$$

where R_t is the covariance matrix of the noise vector V_t . K_t is the Kalman gain matrix.

In theory, the trajectory of a moving target can be regarded as a smooth curve in a short time. However, on the one hand, there is a certain jitter in the target trajectory curve obtained by KCF. Besides, when there are shadows, similar backgrounds, or other interference, it can easily lead to the failure of MOT. The above two problems can be improved by using KF to correct the tracking results. On

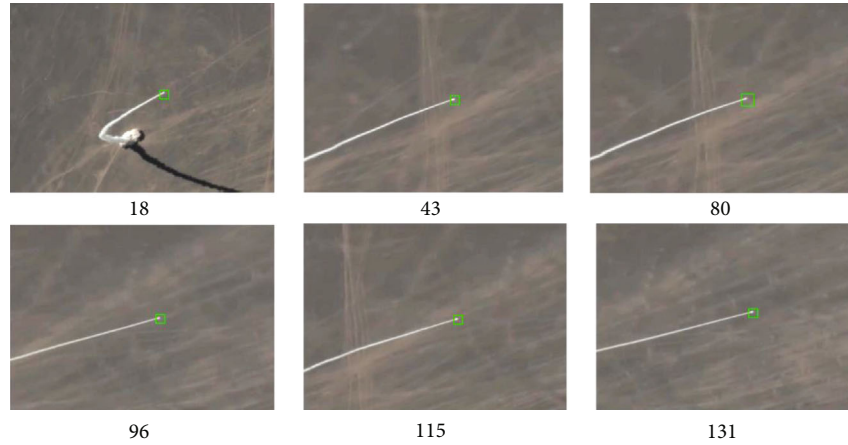


FIGURE 13: The results of moving rocket tracking by CNK-KCF.

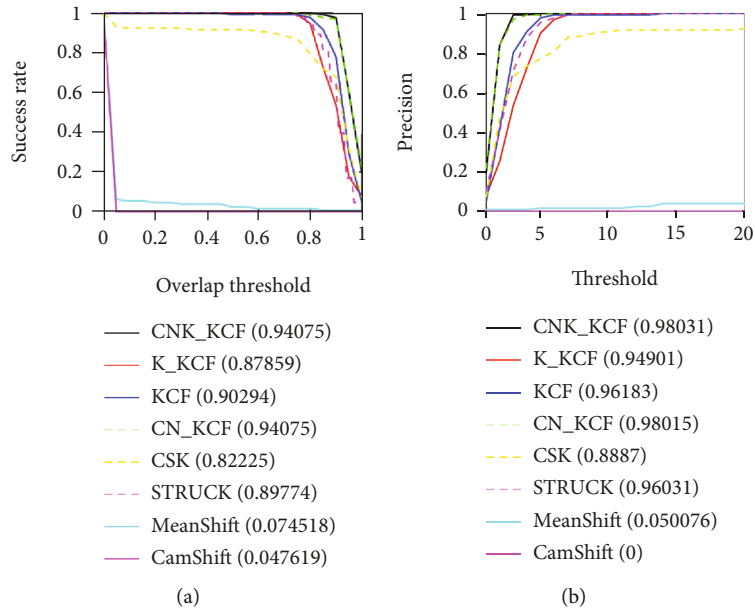


FIGURE 14: The success and precision plots in comparison with the other seven trackers: (a) the success plot for the rocket; (b) the precision plot for the rocket.

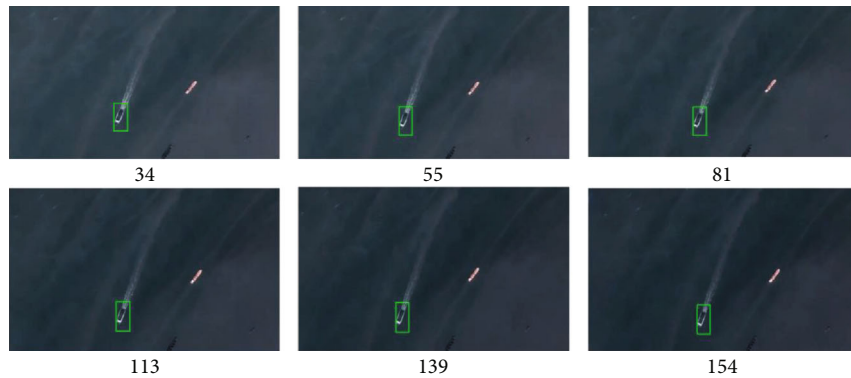


FIGURE 15: The results of moving ship tracking by CNK-KCF.

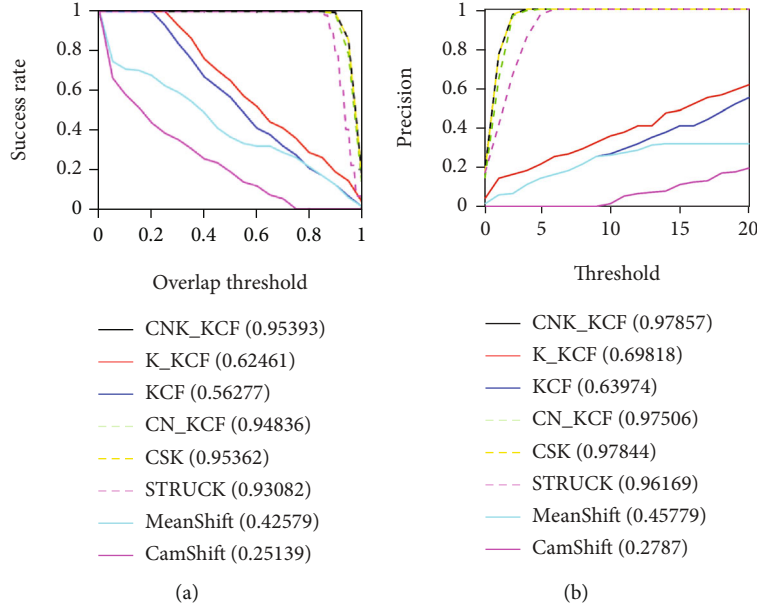


FIGURE 16: The success and precision plots in comparison with the other seven trackers: (a) the success plot for the ship; (b) the precision plot for the ship.

the other hand, the reason is that KF can predict the possible position in the next frame based on the current frame. Hence, the search area can be relatively reduced. Correspondingly, the tracking accuracy can be relatively improved and is robust to some environmental factors such as partial occlusion, background similarity, and rapid motion.

3. Results and Discussion

3.1. Datasets and Compared Algorithms. In this paper, the experimental datasets are from the Jilin No. 1 satellite constellation developed by the China Changchun Satellite Technology Co., Ltd. There are six videos used in our experiments. Among the videos, six scenes are selected from the satellite videos. Besides, the moving objects are the airplanes in the first four videos, and in the other two videos, the moving objects are the rocket and the ship. The number of the objects is only one in each video. The videos show the part of the process of aircraft flying, ship driving, and rocket launching. Each object is manually labeled through a directional bounding box to describe the position. Figure 4 illustrates the overview of the moving objects in the datasets. In addition, the sizes of the targets in the satellite videos are shown in Table 1. Notice that the sizes of moving objects are small, which may result in tracking failure.

Besides, there are many problems for MOT in satellite videos, which will affect the tracking performance. As a result, to analyze the MOT performance of the proposed method, the authors choose seven trackers, KCF, the kernelized correlation filter-based Kalman filter (K-KCF), CN-KCF, the circulant structure of tracking-by-detection with kernels (CSK), STRUCK, MeanShift, and CamShift, to compare with the proposed CNK-KCF algorithm. In Tables 2–7, the comparison among the above 8 trackers is made in detail.

3.2. Details on the Setting of Parameters. The KCF, K-KCF, CN-KCF, CSK, STRUCK, MeanShift, CamShift, and CNK-KCF are implemented in MATLAB R2018b and NVIDIA GeForce GTX 2080Ti GPU. Therefore, for all tested video datasets, the basic parameters of the algorithms are shown in Table 8, where σ is the standard deviation of the Gaussian kernel, λ is the regularization coefficient, and η is the learning factor. Besides, the parameters of each algorithm are consistent in all video sequences.

3.3. Evaluation Metrics. In this paper, two common evaluation criteria are used, that is, precision plot and success plot [60]. The horizontal axis of the accuracy chart is the center location error (CLE). In a frame image, CLE is described as

$$\text{CLE} = \sqrt{(x_{tr} - x_{gr})^2 + (y_{tr} - y_{gr})^2}, \quad (38)$$

where CLE represents the average Euclidean distance, (x_{tr}, y_{tr}) is the tracked target center position coordinates, and (x_{gr}, y_{gr}) is the manually marked real coordinates. When CLE is less than this prescribed threshold, it indicates that the tracking target is correct. That means the smaller the prescribed CLE value is, the more accurate the MOT is. The vertical axis of the accuracy map is the percentage of frames in which tracking accuracy is greater than the threshold.

On the other hand, the horizontal axis of the success rate graph is the overlap threshold of the bounding box. The mathematical expression of overlap rate S is as follows:

$$S = \frac{R_t \cap R_a}{R_t \cup R_a}, \quad (39)$$

where R_t is the predicted tracking box obtained from the algorithm, R_a is the real target box marked manually, and S is the ratio of the overlapping area of R_t and R_a to the total area of R_t and R_a . $|\bullet|$ represents the number of pixels in the region. The vertical axis of the success rate graph is the proportion of successful frames to all image frames. In this experiment, the area under the success rate curve (AUC) is used as the performance evaluation criterion of the algorithm. The larger the value is, the better the tracking performance is.

3.4. Experimental Analysis on Moving Airplane Tracking. The moving objects are the airplanes in the first four real satellite videos of this experiment, recorded as airplane-1, airplane-2, airplane-3, and airplane-4. Note that the objects are small, but the scales are large. First, the results of the moving airplane-1 tracking by CNK-KCF are shown in Figure 5. Note that the airplanes at the airport are not completely occluded. It is shown that for the airplane-1 whose shape is clear, CNK-KCF performs best with a precision of 0.957 and a success rate of 0.917. However, the other seven trackers do not possess similar performance as the CNK-KCF in the aspects of success rate and precision (in fact, in Tables 2–7, we can see the CamShift tracking failure in the satellite videos). Figure 6 shows the success rate and precision plots containing the eight trackers for airplane-1.

For the airplane-2 video, Figure 7 shows the results of the moving airplane-2 tracking. Because the environment around the moving target is relatively complex, the K-KCF and STRUCK have a weaker performance than the proposed CNK-KCF, which have a precision of 0.284 and 0.260 and success rate of 0.272 and 0.224, respectively, followed by the KCF. However, the proposed CNK-KCF has better performance with a precision of 0.902 and a success rate of 0.967. Figure 8 shows the success rate and precision plots containing the eight trackers for airplane-2.

For the airplane-3 video, Figure 9 shows the results of the moving airplane-3 tracking. In Table 4, the KCF, CN-KCF, K-KCF, CSK, STRUCK, and CNK-KCF have similar precision performance. However, the success rate of CN-KCF, KCF, K-KCF, and CSK is less than 0.9, compared with that of the proposed CNK-KCF. Figure 10 shows the success rate and precision plots containing the eight trackers for airplane-3.

For the airplane-4 video, Figure 11 shows the results of the moving airplane-4 tracking. In Table 5, the proposed CNK-KCF has better results than the other 7 trackers. The CN-KCF and CSK have slightly weaker results, followed by STRUCK. However, the KCF and K-KCF have relatively weaker results, only with a precision of 0.691 and 0.848 and a success rate of 0.506 and 0.632, respectively. Figure 10 shows the success rate and precision plots containing the eight trackers for airplane-4. In all, in the experimental results, on the MOT of airplanes, it can be seen that the designed CNK-KCF has better tracking success rate and precision, compared with the other seven algorithms. Figure 12 shows the success rate and precision plots containing the eight trackers for airplane-4.

3.5. Experimental Analysis on Moving Rocket Tracking. The results of moving rocket tracking are shown in Figure 13. Because a rocket over the sea will be partly occluded in the satellite videos, it is necessary to consider the occlusion detection in the MOT. Besides, the rocket is small in satellite videos. Hence, the texture features and shape are not clear. In Table 6, CNK-KCF performs best with a precision of 0.980 and a success rate of 0.941, as well as CN-KCF. KCF, K-KCF, STRUCT, and CSK showing weaker performance in the aspect of precision and success rate. Besides, MeanShift fails to track the rocket. Figure 14 shows the success rate and precision plots containing the 8 trackers for the rocket.

3.6. Experimental Analysis on Moving Ship Tracking. The results of moving ship tracking are shown in Figure 15. CNK-KCF performs best with a precision of 0.979 and a success rate of 0.954. KCF and K-KCF show weaker performance in the aspect of precision and success rate. The performance of MeanShift and CamShift is also weaker with a precision of 0.458 and 0.279 and a success rate of 0.426 and 0.251, respectively. Figure 16 shows the success rate and precision plots containing the 8 trackers for the ship.

In all, compared with the other 7 algorithms, the proposed CNK-KCF possesses better tracking performance.

4. Conclusions

In this paper, the authors propose an effective tracker called CNK-KCF based on the framework of the correlation filter for the MOT in satellite videos. Based on the CN feature, the proposed tracker can process the videos in the multi-channel color feature. The Kalman filter is also utilized to improve the tracking success rate and accuracy.

The improved algorithm is tested on Jilin-1 datasets many times. Meanwhile, we compared with other seven tracking algorithms. The experimental results are analyzed and it has been determined that the proposed method is robust in several complex situations such as rapid target motion, occlusion, and similar object interference. The proposed algorithm solves the problem of tracking failure when a moving object is partially occluded. In the future work, the above tracker will be combined with the controller rather than working separately.

Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (62133001 and 61520106010) and the National Basic Research Program of China (973 Program) (2012CB821200 and 2012CB821201).

References

- [1] P. d'Angelo, G. Mátyus, and P. Reinartz, "Skybox image and video product evaluation," *International Journal of Image and Data Fusion*, vol. 7, no. 1, pp. 3–18, 2016.
- [2] A. Hawbani, X. Wang, S. Karmoshi et al., "GLT: grouping based location tracking for object tracking sensor networks," *Wireless Communications and Mobile Computing*, vol. 2017, 19 pages, 2017.
- [3] J. Zhang, X. Jia, and J. Hu, "Local region proposing for frame-based vehicle detection in satellite videos," *Remote Sensing*, vol. 11, no. 20, p. 2372, 2019.
- [4] Z. Tang, "Intelligent target detection and tracking algorithm for martial arts applications," *Wireless Communications and Mobile Computing*, vol. 2022, 10 pages, 2022.
- [5] X. Li, L. Zhang, and J. You, "Domain transfer learning for hyperspectral image super-resolution," *Remote Sensing*, vol. 11, no. 6, p. 694, 2019.
- [6] F. Nunziata, X. Li, A. Marino et al., "Microwave satellite measurements for coastal area and extreme weather monitoring," *Remote Sensing*, vol. 13, no. 16, p. 3126, 2021.
- [7] A. Xiao, Z. Wang, L. Wang, and Y. Ren, "Super-resolution for "Jilin-1" satellite video imagery via a convolutional network," *Sensors*, vol. 18, no. 4, p. 1194, 2018.
- [8] G. Kopsiaftis and K. Karantzas, "Vehicle detection and traffic density monitoring from very high resolution satellite video data," in *Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1881–1884, Milan, Italy, July 2015.
- [9] Y. Guo, D. Yang, and Z. Chen, "Object tracking on satellite videos: a correlation filter-based tracking method with trajectory correction by Kalman filter," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 9, pp. 3538–3551, 2019.
- [10] T. Yang, X. Wang, B. Yao et al., "Small moving vehicle detection in a satellite video of an urban area," *Sensors*, vol. 16, no. 9, p. 1528, 2016.
- [11] Z. Shi, X. Yu, Z. Jiang, and B. Li, "Ship detection in high-resolution optical imagery based on anomaly detector and local shape feature," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 4511–4523, 2014.
- [12] W. Ao, Y. Fu, and X. Hou, "Needles in a haystack: tracking city-scale moving vehicles from continuously moving satellite," *IEEE Transactions on Image Processing*, vol. 29, pp. 1944–1957, 2020.
- [13] F. Kocadag and A. Demirkol, "Real time tracking of TV satellites on moving vehicles using Kalman filter," in *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 1–5, Noida, India, February 2015.
- [14] W. L. Wang, Q. Q. Li, and L. L. Tang, "Algorithm of vehicle detection in low altitude aerial video," *Journal of Wuhan University of Technology*, vol. 32, pp. 155–158, 2010.
- [15] Y. Luo, Y. Liang, and Y. Wang, "Traffic flow parameter estimation from satellite video data based on optical flow," *Computer Engineering and Applications*, vol. 54, pp. 204–207, 2018.
- [16] L. Gueguen and R. Hamid, "Large-scale damage detection using satellite imagery," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1321–1328, Boston, MA, USA, June 2015.
- [17] L. H. Hu, "Evaluation research on the application of GF-1 satellite for monitoring major engineering land," *Journal of North China Institute of Science and Technology*, vol. 12, pp. 110–115, 2015.
- [18] A. Hawkins, J. Carrico, S. Motiwala, and C. Maclachlan, "Flight dynamics operations and collision avoidance for the Skysat imaging constellation," in *9th International Workshop on Satellite Constellations and Formation Flying (IWSCFF)*, University of Colorado, Boulder, CO, USA, July 2017.
- [19] A. K. Murthy, M. Shearn, and B. D. Smiley, "Sky Sat-1: very high-resolution imagery from a small satellite," in *Proceedings of the Sensors, Systems, and Next-Generation Satellites XVIII*, Amsterdam, Netherlands, September 2014.
- [20] M. Steckling, U. Renner, and H. P. Röser, "DLR-TUBSAT, qualification of high precision attitude control in orbit," *Acta Astronautica*, vol. 39, no. 9–12, pp. 951–960, 1996.
- [21] N. Chen, *Jilin-1: China's first commercial remote sensing satellites aim to fill the void*, Chinese Academy of Sciences, 2016.
- [22] X. Zhang, J. Xiang, and Y. Zhang, "Space object detection in video satellite images using motion information," *International Journal of Aerospace Engineering*, vol. 2017, 9 pages, 2017.
- [23] M. Yoo, Y. Na, H. Song et al., "Motion estimation and hand gesture recognition-based human-UAV interaction approach in real time," *Sensors*, vol. 22, no. 7, p. 2513, 2022.
- [24] A. Yilmaz, O. Javed, and M. Shah, "Object tracking," *ACM Computing Surveys*, vol. 38, no. 4, pp. 13–45, 2006.
- [25] X. Mei and H. Ling, "Robust visual tracking using L1 minimization," in *2009 IEEE 12th International Conference on Computer Vision*, pp. 1436–1443, Kyoto, Japan, 2009.
- [26] C. Bao, Y. Wu, H. Ling, and H. Ji, "Real time robust L1 tracker using accelerated proximal gradient approach," in *2012 IEEE conference on computer vision and pattern recognition*, pp. 1830–1837, Providence, RI, USA, July 2012.
- [27] X. Mei, H. Ling, Y. Wu, E. P. Blasch, and L. Bai, "Minimum error bounded efficient l1 tracker with occlusion detection," in *2011 IEEE conference on computer vision and pattern recognition*, pp. 1257–1264, Colorado Springs, CO, USA, June 2011.
- [28] Z. Kalal, K. Mikolajczyk, and J. Matas, "Tracking-learning-detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 7, pp. 1409–1422, 2012.
- [29] Y. Bai and M. Tang, "Robust tracking via weakly supervised ranking SVM," in *2012 IEEE conference on computer vision and pattern recognition*, pp. 1854–1861, Providence, RI, USA, June 2012.
- [30] H. Grabner, C. Leistner, and H. Bischof, "Semi-supervised online boosting for robust tracking," in *10th European Conference on Computer Vision (ECCV 2008)*, pp. 234–247, Marseille, France, October 2008.
- [31] K. Zhang, L. Zhang, Q. Liu, D. Zhang, and M. H. Yang, "Fast visual tracking via dense spatio-temporal context learning," in *13th European Conference on Computer Vision (ECCV 2014)*, pp. 2661–2675, Zurich, Switzerland, September 2014.
- [32] Q. He, R. Yang, J. Lau, and M. H. Yang, "Visual tracking via locality sensitive histograms," in *2013 IEEE conference on computer vision and pattern recognition*, pp. 2427–2434, Portland, OR, USA, June 2013.
- [33] X. Jia, H. Lu, and M. H. Yang, "Visual tracking via adaptive structural local sparse appearance model," in *2012 IEEE conference on computer vision and pattern recognition*, pp. 1822–1829, Providence, RI, USA, June 2012.
- [34] T. Vojir, J. Noskova, and J. Matas, "Robust scale-adaptive mean-shift for tracking," *Pattern Recognition Letters*, vol. 49, pp. 250–258, 2014.

- [35] Q. Wang, J. Fang, and Y. Yuan, "Multi-cue based tracking," *Neurocomputing*, vol. 131, pp. 227–236, 2014.
- [36] T. Zhang, B. Ghanem, S. Liu, and N. Ahuja, "Robust visual tracking via multi-task sparse learning," in *2012 IEEE conference on computer vision and pattern recognition*, pp. 2042–2049, Providence, RI, USA, June 2012.
- [37] Y. Tian, T. Guan, and C. Wang, "Real-time occlusion handling in augmented reality based on an object tracking approach," *Sensors*, vol. 10, no. 4, pp. 2885–2900, 2010.
- [38] H. Shi, Z. Lin, W. Tang, B. Liao, J. Wang, and L. Zheng, "A robust hand tracking approach based on modified tracking-learning-detection algorithm," in *8th International Conference on Multimedia and Ubiquitous Engineering (MUE 2014)*, pp. 9–15, Zhangjiajie, China, May 2014.
- [39] C. Jia, Z. Wang, X. Wu et al., "Tracking-learning-detection (TLD) method with local binary pattern improved," in *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 1625–1630, Zhuhai, China, December 2015.
- [40] S. Avidan, "Support vector tracking," in *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition (CVPR 2001)*, pp. 1064–1072, Kauai, HI, USA, December 2001.
- [41] H. Grabner, M. Grabner, and H. Bischof, "Real-time tracking via online boosting," in *2006 17th British Machine Vision Conference (BMVC 2006)*, pp. 47–56, Edinburgh, United Kingdom, September 2006.
- [42] B. Babenko, M. H. Yang, and S. Belongie, "Robust object tracking with online multiple instance learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1619–1632, 2011.
- [43] M. Danelljan, A. Robinson, and F. Shahbaz Khan, "Beyond correlation filters: learning continuous convolution operators for visual tracking," in *21st ACM Conference on Computer and Communications Security (CCS 2014)*, pp. 472–488, Scottsdale, AZ, United states, November 2014.
- [44] R. S. Chavan and M. Patil, "Object tracking based on tracking-learning-detection," *IEEE Transactions on Software Engineering*, vol. 7, 2011.
- [45] S. Hare, A. Saffari, and P. Torr, "Struck: structured output tracking with kernels," in *2011 IEEE International Conference on Computer Vision (ICCV 2011)*, pp. 263–270, Barcelona, Spain, November 2011.
- [46] W. Zhong, H. Lu, and H. Yang, "Robust object tracking via sparse collaborative appearance model," *IEEE Transactions on Image Processing*, vol. 23, no. 5, pp. 2356–2368, 2014.
- [47] J. Zhang, Y. Liu, H. Liu, and J. Wang, "Learning local-global multiple correlation filters for robust visual tracking with Kalman filter redetection," *Sensors*, vol. 21, no. 4, p. 1129, 2021.
- [48] J. Yang, W. Tang, and Z. Ding, "Long-term target tracking of UAVs based on kernelized correlation filter," *Mathematics*, vol. 9, no. 23, p. 3006, 2021.
- [49] M. Danelljan, G. Häger, and F. S. Khan, "Discriminative scale space tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 8, pp. 1561–1575, 2017.
- [50] H. K. Galoogahi, A. Fagg, and S. Lucey, "Learning background-aware correlation filters for visual tracking," in *16th IEEE International Conference on Computer Vision (ICCV 2017)*, pp. 1144–1152, Venice, Italy, October 2017.
- [51] B. Du, Y. Sun, S. Cai, C. Wu, and Q. Du, "Object tracking in satellite videos by fusing the kernel correlation filter and the three-frame-difference algorithm," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 168–172, 2018.
- [52] H. Li and Y. Man, "Moving ship detection based on visual saliency for video satellite," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2016)*, pp. 1248–1250, Beijing, China, July 2016.
- [53] L. Lei and D. Guo, "Multitarget detection and tracking method in remote sensing satellite video," *Computational Intelligence and Neuroscience*, vol. 2021, 7 pages, 2021.
- [54] Y. Liu, Y. Liao, C. Lin, Y. Jia, Z. Li, and X. Yang, "Object tracking in satellite videos based on correlation filter with multi-feature fusion and motion trajectory compensation," *Remote Sensing*, vol. 14, no. 3, p. 777, 2022.
- [55] L. Meng and J. P. Kerekes, "Object tracking using high resolution satellite imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 1, pp. 146–152, 2012.
- [56] F. S. Khan, J. Van De Weijer, and M. Vanrell, "Modulating shape features by color attention for object recognition," *International Journal of Computer Vision*, vol. 98, no. 1, pp. 49–64, 2012.
- [57] M. Danelljan, G. Bhat, and F. S. Khan, "ECO: efficient convolution operators for tracking," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017)*, pp. 6931–6939, Honolulu, HI, USA, July 2017.
- [58] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 583–596, 2015.
- [59] R. Rifkin, G. Ye, and T. Poggio, "Regularized least-squares classification," *Nato Science Series Sub Series III Computer and Systems Sciences*, vol. 190, pp. 131–154, 2003.
- [60] M. Kristan, J. Matas, A. Leonardis et al., "The seventh visual object tracking VOT2019 challenge results," in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pp. 2206–2241, Seoul, Korea (South), October 2019.