

Visual ship tracking via a hybrid kernelized correlation filter and anomaly cleansing framework

Xinqiang Chen^{a,b}, Xueqian Xu^b, Yongsheng Yang^b, Yanguo Huang^{c,*}, Jing Chen^d, Ying Yan^{e,*}

^a Institute of Atmospheric Sciences, Fudan University, Shanghai 200438, China

^b Institute of Logistics Science and Engineering, Shanghai Maritime University, Shanghai, 201306 China

^c School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou 341000, China

^d Merchant Marine College, Shanghai Maritime University, Shanghai 201306, China

^e College of Transportation, Chang'an University, Xi'an 710064, China

ARTICLE INFO

Keywords:

Visual ship tracking
Remote maritime traffic controlling
Kernelized correlation filter
Anomaly cleansing
Smart ship

ABSTRACT

Ship tracking from maritime visual sensing data (namely maritime surveillance videos) provides various kinematic maritime traffic information, which significantly benefits remote maritime traffic controlling and management, off-site law enforcement, etc. But, it is difficult to extract distinct ship visual features when the target ship is sheltered by the neighboring ship in the maritime images (or the images are shot in low visibility condition). To address the difficulty, we proposed a hybrid ship tracking framework via the help of kernelized correlation filter (KCF) and anomaly cleaning models (including curve fitting method and Kalman filter). First, we employed the KCF model to obtain raw ship trajectories in consecutive maritime images. Second, our ship tracker is accurately initialized (to be specific, ground truth ship position in the first frame is employed to initialize the ship tracker). Third, the Kalman filter is introduced to suppress the trivial ship position oscillations in the raw ship trajectories. We verified the proposed framework performance on the four typical maritime scenarios. The experimental results indicate that the proposed ship tracker showed more accurate ship tracking results compared to other four popular ship trackers in terms of average root mean square error (RMSE), mean absolute deviation (MAD) and mean square error (MSE). The research findings can help maritime traffic participants obtain visual on-spot maritime kinematic information, and thus further enhance maritime traffic safety.

1. Introduction

Smart ship contains many efficacies (e.g., high efficiency, low carbon footprint, cost-effective, etc.), which significantly outperforms those in traditional navigation era (Shi et al., 2018). Visual ship tracking technique can tackle many remote unmanned tasks, which is considered as one of the fundamental tasks for visual perception for a smart ship. To fulfill accurate maritime situational awareness for a smart ship, both visual and non-visual sensing techniques are employed to obtain informative kinematic and static ship information from available data sources (e.g., maritime surveillance videos, long-range identification and tracking (LRIT), automatic identification systems (AIS), synthetic aperture radar (SAR)) (Henriques et al., 2015; Zhang et al., 2017; ZHANG et al., 2016; Frodella et al., 2020). Note that maritime video clips provide straightforward yet informative spatial-temporal maritime data, and thus obtaining accurate ship tracking results have become a hot

topic in maritime traffic research community (Chen et al., 2020a).

Though AIS data support a broad range of maritime applications (e.g., ship positioning, short and long-term path planning), which have enjoyed huge successfulness in tackling the ship tracking task. The AIS relevant techniques are considered as popular yet very efficient model to tackle the remote maritime traffic controlling task (Liu et al., 2018a). The main weakness is that not all ships are equipped with the AIS receiver and transmitter facilities (e.g., small fishing boats). Moreover, ship (e.g., smuggling ships, warships) officials may deliberately deactivate the AIS equipment to avoid being perceived by maritime traffic regulation department and neighbouring ships. In that manner, additional techniques are in need to help the AIS-relevant techniques obtain ship positions. The other popular ship tracking techniques (e.g., LRIT, SAR) transmits ship locations at a large time interval, which cannot satisfy ship tracking requirement (especially in coastal ship tracking tasks) (Fang et al., 2018a; Chen et al., 2018; Abebe et al., 2020; Zhou

* Corresponding authors.

E-mail addresses: yghuang@jxust.edu.cn (Y. Huang), yanying2199@chd.edu.cn (Y. Yan).

et al., 2019; Lang et al., 2018; Fang et al., 2018b; Chen et al., 2020b; Liang et al., 2020).

Visual ship tracking from videos provides various maritime traffic information (e.g., speed, course, trajectory), which helps traffic participants aware traffic situation in a timely manner. Previous ship tracking studies were conducted to solve various visual ship tracking challenges, such as ship imaging size variation, ship occlusion, etc. Xu et al. proposed a box regression based self-selective correlation filtering model to solve the coupled challenge of ship size variation and sea background interference during the ship tracking procedure (Kang et al., 2019). Chen et al. introduced a deep learning model to robustly exploit distinct ship features for the purpose of extracting ship trajectories from maritime video clips (Chen et al., 2020c). Shao et al. employed a convolutional neural network to predict ship location in image sequences, which were further corrected by the coastline imaging information (Shao et al., 2019). Similar studies can be found in Kim et al. (2018), Liu et al. (2018b), Smolenski et al. (2018), Zong et al. (2019), Tang et al. (2020), Chen et al., 2020d).

Previous ship tracking models fulfill the visual ship tracking task by exploiting meaningful ship features from maritime images. The ship tracking model performance is easily interfered by ship occlusion challenge, which is common in the maritime surveillance videos shot at coastal navigation area. Moreover, adverse weather condition (such as low visibility) can further degrade ship tracking model accuracy due to the ambiguous visual ship features. The difficulty is considered as one of the bottlenecks for fulfilling the visual perception task in the smart ship era. We found that the typical ship occlusion challenge is triggered when the target ship is small and the obstacle ship is small. In that way, the target ship visual features are sheltered by the obstacle ship counterparts, and thus mislead the ship tracking model learning from the biased ship features.

In this study, we aim to accurately track ship positions from maritime images under interference of varied ship occlusion and low visibility

condition. Note that ship tracking under severe extreme weather conditions (such as storm, snow) beyond our scope due to a few distinct visual features can be identified. Our main contributions can be summarized as follows: (1) we proposed a hybrid framework via the kernelized correlation filter and curve fitting model to robustly track ships under ship occlusion interference; (2) we employed the Kalman filter to smooth the trivial ship position oscillations in both x and y axis; (3) we verified our proposed ship tracking model performance under four typical maritime traffic situations. More specifically, the first three cases focused on performance evaluation under ship occlusion challenge, and the fourth with adverse weather condition. The remainder of the paper is organized as follows: Section II describes the ship tracking model in detail. Section III introduces the ship dataset, evaluation metrics, and experimental results, and Section IV briefly concludes the research.

2. Methodology

The proposed ship tracking framework is shown in Fig. 1, which includes extracting raw ship positions, cleansing position outliers, minimizing trivial oscillations in the overall ship trajectory. More specifically, the raw ship positions in maritime images are firstly extracted frame-by-frame with the help of KCF tracker. Then, we estimate ship positions with curve fitting model when the target ship is occluded in image sequences. Finally, a general but efficient denoising algorithm of Kalman filter is performed to suppress abnormal trivial position oscillations in the raw ship trajectories. Details for each step are provided as follows.

2.1. Obtain raw ship positions from maritime images with KCF model

The KCF model finds target ship from maritime images with the help of cross-correlation criterion in the manner of determining maximum stimulus response between the ground truth training ship sample and

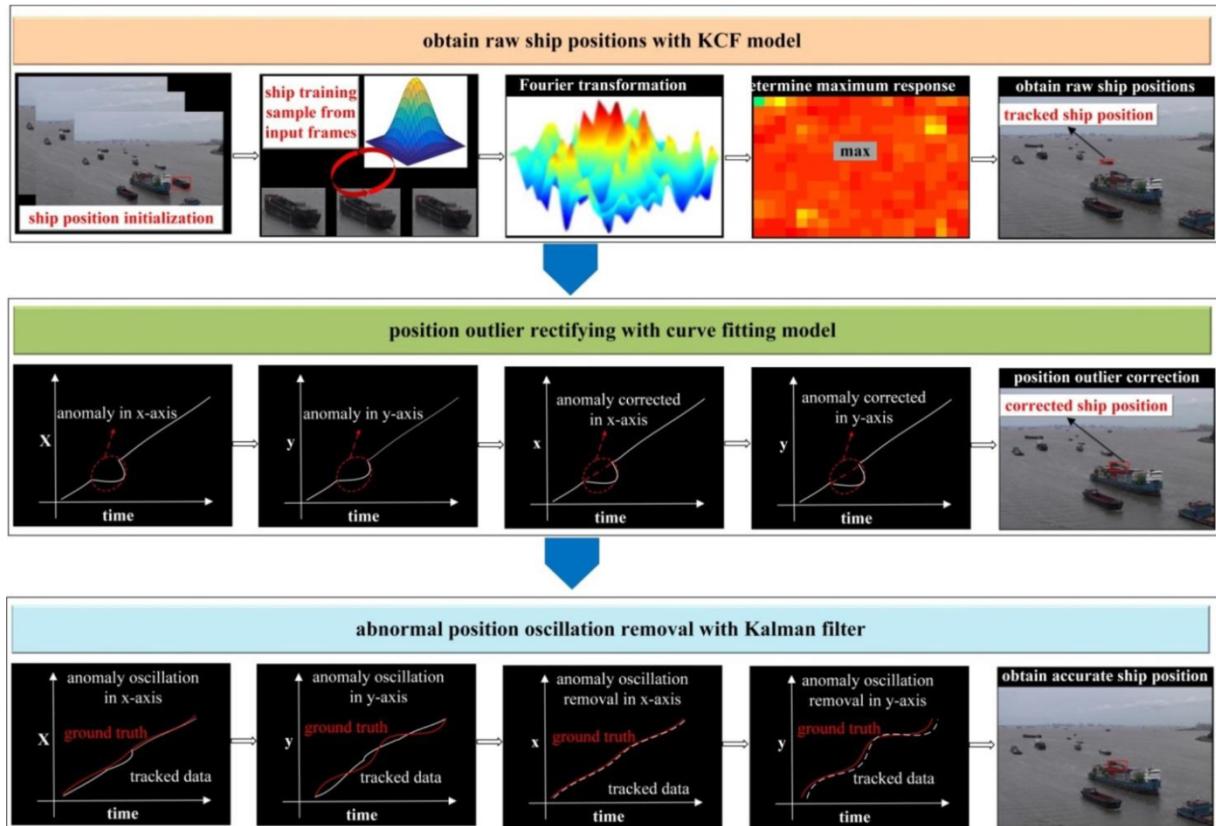


Fig. 1. Overview for the proposed ship tracking model.

ship candidates (Yang et al., 2019). The proposed KCF ship tracker is detailed explained as follows. The ship images and labels are manually labeled for the purpose of training the KCF tracker, which are denoted as (o_n, τ_n) , respectively. Note that the parameter n is the sample index. The KCF ship tracker $f(o)$ is trained in the manner of finding optimal solution for the linear regression problem (see (1)).

Previous study suggested that the optimal solution for (1) is obtained by searching the minimal distance the training ship sample (i.e., ground truth in previous frame) and ship candidates in the current frame (see (2)) (Henriques et al., 2015). The closed form solution is considered as one of the optimal solutions for the (2), which is formulated as (3). More specifically, the input ship training samples are mapped into feature space $\psi(o_n)$, which is demonstrated as kernel e . We obtain the linear combination of the training data, which is considered as the optimal solution for (2). In that way, the target ship in the current ship frame (i.e., maximum response for the KCF ship tracker) is obtained with (4). For the purpose of generalization, we reformulate the (4) in the form of (5) which can represent both linear and nonlinear ship tracking results for the KCF tracking model. Motivated by previous studies, we obtain an efficient yet simple closed form solution for the KCF model (see (6)) (Smola and Schölkopf, 2004; Ong et al., 2005).

$$f(o) = \langle w, o \rangle + b \quad (1)$$

$$\min \sum_n L(f(o_n), \tau_n)^2 + \| \sigma w \|^2 \quad (2)$$

$$w = (o^T o + \sigma I)^{-1} o^T \tau \quad (3)$$

$$w = \sum_n \gamma_n \psi(o_n) \quad (4)$$

$$f(o) = \sum_n \gamma_n e(o, o_n) \quad (5)$$

$$\gamma = (E + \sigma I)^{-1} \tau \quad (6)$$

where symbol $\langle \cdot, \cdot \rangle$ is the dot production. The parameter w is the linear combination of the input training ship samples and images, and σ determines the regularization level for the ship tracker. The parameter $L(f(o_n), \tau_n)^2$ is the loss function when training the KCF ship tracker. The symbol E is a kernel matrix comprising of a group of kernels e , and I is the identity matrix. The parameters o and τ consisting of o_n and τ_n , respectively, and γ is the obtained solution for the KCF ship tracker.

We use a cyclic-shift mechanism for the purpose of producing additional positive training ship samples, which are further employed to enhance the KCF model tracking performance. More specifically, the cyclic-shift mechanism samples on the ship base sample (i.e., ground truth ship sample) in an iterative manner. We obtain cyclic-shift training ship samples by sampling on the input ship with cosine windows in the manner of one-dimensional and two-dimensional cyclic shift (see Fig. 2). The permutation matrix used for collecting cyclic-shift candidate

samples from maritime images is shown as (7), which is denoted as $d = [d_1, d_2, d_3, \dots, d_s]$ (i.e., $1 \times s$ dimension vector).

Furthermore, we obtain the circulant kernel matrix $L(d)$ in the manner of cyclically shifting on the permutation matrix (i.e., the vector d), which are used to generate derivative ship training samples for the KCF tracking model (see (8)). The circulant kernel matrix $L(d)$ is transformed into diagonal matrix after the discrete Fourier transformation, which can significantly reduce computational complexity for the KCF ship tracker. Note that the discrete Fourier transformation applied in our study is shown in (9), where parameter F is discrete Fourier transform matrix. The target ship positions in the input maritime images (as shown in (3) and (4), respectively) are obtained by (10). Note that the operators $*$ and \odot are both the element-wise product, and the symbol \wedge is the discrete Fourier transformation generated vector.

It is found that obtaining ship tracking results in temporal-spatial domain requires large time costs for the KCF tracker. To mitigate such weakness, researchers employ the diagonalized Discrete Fourier Transform method to transform the ship data and corresponding operators (i.e., candidate ships, circulant matrices) into frequency domain (Henriques et al., 2015; Gray, 2006). In such way, the KCF operators (e.g., inversion, transposition, multiplication) are implemented on each maritime frame (used for both training and testing purpose) during the ship tracking procedure. Note that the time computation is halved compared to that in the temporal-spatial domain due to the information redundancy advantage of the diagonalized Discrete Fourier Transform model.

By defining a compact vector ρ (as shown in (11)), we re-formulate the (6) into (12) which can be solved in a more efficient manner. The target ship area in the maritime frames is determined by obtaining the maximum response between the input ship template and the to-be-tracked image. In that manner, the KCF model obtains minimal value for the (2) when the target ship is found. The KCF tracking results (i.e., ship position) in each ship frame is shown as (13).

$$v_p = \begin{bmatrix} 00\dots01 \\ 10\dots00 \\ 01\dots00 \\ \vdots \ddots \vdots \\ 00\dots10 \end{bmatrix} \quad (7)$$

$$L(d) = \begin{bmatrix} d_1 d_2 d_3 \dots d_s \\ d_s d_1 d_2 \dots d_{s-1} \\ d_{s-1} d_s d_1 \dots d_{s-2} \\ \vdots \vdots \ddots \vdots \\ d_1 d_3 d_4 \dots d_1 \end{bmatrix} \quad (8)$$

$$L(d) = F \text{diag}(\widehat{d}) F^H \quad (9)$$

$$\widehat{w} = \frac{\widehat{o}^* \odot \widehat{\tau}}{\widehat{o}^* \odot \widehat{o} + \sigma} \quad (10)$$

$$\rho_n = \rho(o, L(d)^n) \quad (11)$$

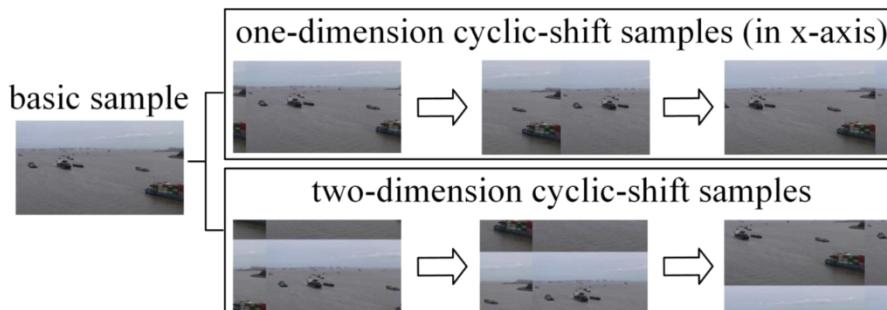


Fig. 2. Cyclic-shift samples generated by basic ship samples.

$$\gamma = F^{-1} \left(\frac{F}{F(L) + \sigma} \right) \quad (12)$$

$$\tau' = \sum_n \gamma_n \rho(o, L(d)^n) \quad (13)$$

where parameter ρ_n is a compact form for the kernel matrix $L(d)$, and τ' is the target ship position output by the KCF model.

2.2. Removing ship tracking outliers in the ship-occlusion frames

The KCF tracking performance may fall short of our expectations when the target ship is occluded by the obstacles (i.e., partial ship pixels are sheltered by obstacle pixels in maritime images), which misleads the KCF model into tracking obstacles (due to that ship visual features are temporally invisible to the KCF tracker in the ship-occlusion images). To remove the ship tracking outliers, we introduce curve fitting model to correct the abnormal data with the help of neighboring trustable ship positions. More specifically, we introduce data outlier identification rule to determine anomaly ship data (i.e., ship positions in occlusion frames). It is noted that ship movements in neighboring frames (i.e., under very short time) are very similar and consistent, indicating that obvious ship displacements in neighbouring frames are irrational data. To follow the premise, a group of constraints (see (14)) are set to detect ship tracking outliers in the raw data (i.e., the above step outputs).

$$\begin{cases} \Delta l(x) \geq \alpha \\ \Delta l(y) \geq \beta \\ \Delta \theta(x) \geq \delta \\ \Delta \theta(y) \geq \nu \end{cases} \quad (14)$$

where $\Delta l(x)$ and $\Delta l(y)$ are the ship displacements in neighbouring frames in x and y-axis, respectively. The symbol $\Delta \theta(x)$ and $\Delta \theta(y)$ are ship azimuth in the maritime images on the x-axis and y-axis, respectively. The parameters α , β , δ and ν are thresholds setting for identifying ship data outliers.

After obtaining the ship position outliers, we implement the curve fitting algorithm to rectify the anomaly data by reconstructing ship positions. More specifically, the ship tracking positions obtained from non-occlusion frames are employed to train the curve fitting model, which is then used to estimate anomaly ship positions. Following the rule in previous studies (Wan et al., 2016), the curve fitting model is trained in the manner of finding minimum loss function between the training and predicted data (see (15)). We trained the curve fitting in the x and y-axis, respectively, and thus the ship positions are rectified in x and y direction as well.

$$\min \sum_{t=1}^C (y_b(t) - \hat{y}_b(t))^2 + (x_b(t) - \hat{x}_b(t))^2 \quad (15)$$

where $x_b(t)$ and $y_b(t)$ are the tracked ship positions in x and y-axis at the t th frame. The $\hat{x}_b(t)$ and $\hat{y}_b(t)$ are the counterparts of the rectified data.

2.3. Suppressing anomaly ship position oscillations with Kalman filter

Though obvious ship tracking outliers are suppressed in the previous step, trivial ship position oscillations exist due to the tracking rectangle vibration, low video quality, insignificant background imaging interference, etc. Note that previous cleansing models are implemented without exploiting spatial-temporal features, and thus cannot be directly used to denoise the ship trivial position outliers in our study. To address the issue, we introduce the Kalman filter to re-track target ship in the maritime images for the purpose of removing trivial ship position outliers. The Kalman filter obtains ship position in each frame by combining both the observed and predicted data by assigning them with different weights (Zhou and Zhang, 2019), and the detailed explanations

are introduced in the section. The ship target position at the $(j+1)$ th frame is denoted $A(j+1)$, and the ship position measurement model is presented as $Y(j+1)$. The Kalman filter is initialized with (16) and (17), respectively as follows.

$$A(j+1) = BA(j) + \Phi W(j) + V(j) \quad (16)$$

$$Y(j+1) = CA(j+1) + U(j+1) \quad (17)$$

where B and Φ are the state transition matrix and the input transition matrix, respectively. The parameter B is the relationship matrix between two neighboring states, which are the ship positions obtained by the Kalman filter. The parameter Φ is the relationship matrix between different input positions. In addition, the parameter C is the observation transition matrix, which indicates the relationship between the tracked and ground truth ship positions. The symbol $W(j)$ is the input controlling vector. The parameter $V(j)$ is the measured noise at the j th frame, and $U(j+1)$ is the observed noise at the $(j+1)$ th frame. Note that the $V(j)$ and $U(j+1)$ are considered as independent of each other.

The Kalman filter implements the ship tracking task with prediction and update steps. More specifically, Kalman filter employs the previous estimated ship position to estimate ship position in current frame. After that, the ship position is further refined with measured data (see (18) to (22)).

$$\bar{A}(j+1) = B\hat{A}(j) + \Phi W(j) + V(j) \quad (18)$$

$$Q^-(j+1) = BQ(j)B^T(j+1) + P(j+1) \quad (19)$$

$$\Gamma(j+1) = Q^-(j+1)C^T(j+1)$$

$$[C(j+1)Q^-(j+1)C^T(j+1) + H(j+1)]^- \quad (20)$$

$$\hat{A}(j+1) = \bar{A}(j+1) + \Gamma(j+1)\zeta(j+1) \quad (21)$$

$$Q(j+1) = [I - \Gamma(j+1)C(j+1)]Q^-(j+1) \quad (22)$$

where $\hat{A}(j)$ and $\hat{A}(j+1)$ are the estimated ship positions at j th and $(j+1)$ th frames, respectively.

The parameter $\bar{A}(j+1)$ is the estimated ship position at $(j+1)$ th frame. The $B^T(j+1)$ is the transposition of state transition matrix at the $(j+1)$ th frame. The $Q(j)$, $Q^-(j+1)$, $Q(j+1)$ and $P(j+1)$ are the covariance matrices under the specific ship frame (i.e., j th and $(j+1)$ th). The $C^T(j+1)$ is transposition of state transition matrix under $(j+1)$ th frame. The $\Gamma(j+1)$ is the gain matrix, and $\zeta(j+1)$ is the measurement transition matrix at the $(j+1)$ th frame. The parameter $H(j+1)$ is the identity matrix.

2.4. Evaluation metrics

For the purpose of tracker performance measurement, we compare the ship tracking data against with the ground truth positions (which are manually labeled frame-by-frame). Following the rule in previous studies (Chen et al., 2019), we employ the root mean square error (RMSE), mean absolute deviation (MAD) and mean square error (MSE) to quantify the tracker performance. Note that we employ the ship position center for both the tracked and ground truth ship positions during procedure of calculating the statistical indicators (i.e., RMSE, MAD and MSE). More specifically, we employ the intersection point (IP) to demonstrate the ship positions (for both tracked and ground truth data), which are further employed to calculate the RMSE, MAD and MSE. The difference between the tracked and ground truth center points are evaluated with Euclidean distance. For a given ship video with M-frames, the ship ground truth position is denoted as $T(x, y)$, and symbol $Z(x, y)$ is the ship tracking results. We measure the Euclidean distance between two positions through (23) to (25). After that, the

RMSE, MAD and MSE values are further calculated by (26) to (29). It is noted that smaller RMSE indicates that the ship tracker obtains better ship tracking accuracy (i.e., the tracked ship position is closer to that of the ground truth), and vice versa. Such rule is applicable to the MAD and MSE as well.

$$N_i(x) = (T_i(x) - Z_i(x))^2 \quad (23)$$

$$N_i(y) = (T_i(y) - Z_i(y))^2 \quad (24)$$

$$N_i(T(x, y), Z(x, y)) = \sqrt{N_i(x) + N_i(y)} \quad (25)$$

$$\bar{N} = \frac{\sum_{i=1}^M N_i(T(x, y), Z(x, y))}{m} \quad (26)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^M |N_i(T(x, y), Z(x, y)) - \bar{N}|^2}{M}} \quad (27)$$

$$MAD = \frac{\sum_{i=1}^M |N_i(T(x, y), Z(x, y)) - \bar{N}|}{M} \quad (28)$$

$$MSE = \frac{1}{M} \sum_{i=1}^M |N_i(T(x, y), Z(x, y)) - \bar{N}|^2 \quad (29)$$

where parameter $N_i(x)$ measures the offset between $T(x, y)$ and $Z(x, y)$ at the x-axis direction for the frame i . The $T_i(x)$ and $Z_i(x)$ are x-coordinates for both $T(x, y)$ and $Z(x, y)$. The parameters $N_i(y)$, $T_i(y)$ and $Z_i(y)$ follow the same rule as that of the $N_i(x)$. The $N_i(T(x, y), Z(x, y))$ is the Euclidean distance between $T(x, y)$ and $Z(x, y)$, and \bar{N} is the average Euclidean distance.

3. Experiments

To evaluate ship tracking performance, we have collected four video clips shot by cameras installed on the coastal building and ships (i.e., two videos were gathered from ship-borne cameras and the rest from coastal building camera). More details are provided in the following data section. Following the rule in our previous study (Chen et al., 2020e), we implemented both typical and newly-developed models to track ships from maritime videos. More specifically, we compared our proposed framework to the Mean-shift (Chen et al., 2019) model, KCF algorithm (Henriques et al., 2015), scale adaptive with multiple features (denoted as SAMF) (Li and Zhu, 2014), and our previous proposed ship tracker shorted as KCFC (Chen et al., 2020e). The above mentioned ship trackers were implemented on Windows 10 operation system. The RAM and CPU are 8 G and Intel Core I7-4710 HQ CPU @ 3.50 GHz processor, respectively. Moreover, the GPU is NVIDIA GeForce GTX 850 M and the memory is 2 G. The experimental platform is Matlab 2016 version.

3.1. Data

It noted that public computer vision benchmarks (such as OTB100) do not contain typical maritime traffic situations and challenges, and thus cannot be used to evaluate ship tracking model performance. To fill the gap, we collected four maritime video clips which contain typical ship tracking challenges. We are willing to make our collected dataset public accessible in future. By installing maritime surveillance camera on coastal buildings at Shanghai Port, China, we captured two video clips, which are denoted as video #1 and #2, respectively. The video #1 contains ship tracking scenario when target ship size is small and the obstacle ship size is significantly larger than the target ship in the video. The video length is 35 s which can fully unveil the ship occlusion procedure in the video, and the image resolution is 1280×720. The frame rate in video #1 is 30 fps (i.e., the camera captures 30 frames per second). The ship tracking challenge in video #2 is similar to that of the

video #1, where both the target and obstacle ships are imaged in large sizes. Moreover, the shooting environment and camera setting (e.g., image resolution, weather, visibility, shooting angle) are very similar for both video #1 and #2. The time duration for video #2 is 37 s. The video #2 frame rate is 30 fps and resolution is 1280×720, which is same to those of video #1. More specific definitions of large and small ships are suggested to refer to Park et al. (2015), Cui et al. (2019).

We have collected another two video clips with the help of ship-borne cameras, which are labelled as video #3 and #4, respectively. More specifically, the video #3 was collected by when the camera is installed on the ship stem, and the video #4 was collected by a ship port-side installed camera. The lengths for video #3 and #4 are 33 s and 21 s, respectively, and frame rates for the two video clips are 30 fps. The image resolutions for video #3 and #4 are 1980×1080 and 1280×720, respectively. The ship tracking challenge in video #3 is that the small-size target ship is partially occluded by the neighbouring small-size ship. Video #4 aims to testify ship tracker performance under mist condition (a type of typical maritime traffic adverse weather). Typical frames for each of collected video clips were shown in Fig. 3. More detailed information about the collected video clips are shown in Table 1.

3.2. Experimental results and analysis

1) Ship tracking performance analysis on video #1

Ship tracking results with various models were shown in Fig. 4(a) and (b), which are the ship movements in x and y axis, respectively. More specifically, the ship tracking positions for the ground truth (denoted as TR), KCF, Mean-shift, SAMF, KCK are marked with blue, green, purple, yellow, red and black curves in the two subplots in Fig. 4. For the purpose of simplicity, we only explain the ship position distribution for the x-axis due to that the counterparts for the y-axis show similar variation tendency. The target ship was obviously occluded by neighbouring obstacle from frame #380 to #850, which was manually checked in maritime images. Both the KCF and Mean-shift models showed obvious outlier in the x axis as shown in Fig. 4(a), indicating that the two trackers tracked the obstacle ship in the image sequences. The main reason is that the KCF model is very sensitive to ship visual feature variation during the tracking procedure. In that way, the visual features for the small target ship cannot be successfully extracted by the KCF model in the ship-occlusion images.

Moreover, the KCF tracking model showed satisfied performance when the target ship was not occluded (see the green curve variation in Fig. 4(a) from frame #850 to the end). This is due to the fact that the KCF tracking model preserves initial ship training sample (i.e., the raw ship tracking template), which helps the KCF model suppress the tracking outlier. The purple curve position distribution (i.e., ship tracking results for the Mean-shift model) in Fig. 4(a) showed that the Mean-shift tracking model was misled by the obstacle ship since frame #380, and the ship tracking error significantly increased since then until the end of the video. The SAMF tracking model showed very similar performance as that of the Mean-shift model (see the red curve in Fig. 4(a)). More specifically, both the SAMF and Mean-shift models tracked the obstacle ship when the target ship was occluded (i.e., the small target ship was sheltered by the obstacle), which cannot be further corrected after the ship-occlusion procedure completed.

It is noted that the KCF model exploits ship features from both the input ship training template and previous ship tracking result (i.e., ship from previous frame). In that manner, the KCF model obtains the ship position in current frame in terms of determining the maximum response between the input maritime frame and training sample. The Mean-shift model tracked ship in the manner of iteratively learning visual features from previous tracking result, and thus tracking error in a single frame will permanently lead to tracking outlier. The SAMF model aggregated various handcrafted features (e.g., contour, color, etc.) into high-level



Fig. 3. Ship samples for each video clip (the target and obstacle ships are marked with green and red rectangles, respectively).

Table 1
Ship video clips information details.

| No. | frame rate (fps) | resolution | length | camera position | ship tracking challenge |
|----------|------------------|------------|--------|------------------|--|
| video #1 | 30 fps | 1280×720 | 35 s | coastal building | small ship completely occluded by a large ship |
| video #2 | 30 fps | 1280×720 | 37 s | coastal building | large ship completely occluded by a large ship |
| video #3 | 30 fps | 1980×1080 | 33 s | ship-stem | small ship partially sheltered by small ship |
| video #4 | 30 fps | 1280×720 | 21 s | ship port-side | mist weather condition |

but distinguished feature. Note that the pre-defined raw handcrafted features may be collected from obstacles (in the ship-occlusion relevant images), which may degrade SAMF tracker performance.

The KCFC model outperformed the KCF, Mean-shift and SAMF models due to the ship position in x-axis is closer to the ground truth (see Fig. 4(a)). The main reason is that KCFC model introduced both visual features and temporal-spatial information (with the help of curve fitting logic) to enhance ship tracking performance. More specifically, the KCFC model alleviated ship tracking interference (caused by ship-occlusion) by inferring to non-occlusion ship trajectory variation tendency. In that manner, the ship positions obtained by the KCFC model were closer to the ground truth data in both x and y axis compared to the KCF, Mean-shift and SAMF models. But, trivial position oscillations were

still found in the KCFC-obtained data due to intrinsic weakness of visual feature supported trackers (i.e., visual feature loss in the ship occlusion images cannot be 100% recovered).

Compared to the KCFC model, our proposed ship tracking framework (i.e., the KCK model) integrated the Kalman filter to suppress the unexpected tracking anomalous data when implementing the ship tracking task. The tracking positions obtained by the proposed KCK model are quite close to the ground truth position, which can be confirmed in the black curve distributions in Figs. 4(a) and 5(b). Indeed, the KCK model introduced the Kalman filter to implement the data quality control, which removed various trivial abnormal oscillations in the tracked ship positions for the video clip. Note that the abnormal tracked positions will be verified and adjusted by the Kalman filter with the help of neighbouring ship positions. Based on the above analysis, it is found that the KCF, Mean-shift, and SAMF models extract distinct ship visual features for implementing the ship tracking task for video #1, and the obstacle ship features were learned when the target ship was occluded (i.e., small target ship was completely sheltered by neighbouring large ship). The ship tracking models involving data quality control (i.e., KCFC and KCK) obtained better tracking accuracy compared those of the KCF, Mean-shift, and SAMF models. Moreover, the KCK model obtained higher tracking accuracy than the KCFC model due to the fine-tuned abnormal oscillation removal by the Kalman filter.

To further quantify the tracking performance of various models, we calculated the position difference between tracked and ground truth data (PDTG) frame-to-frame. It is noted that the KCF, Mean-shift and SAMF showed obvious tracking outliers from between #400 and #850 (see the green, purple and yellow curves in Fig. 5). The KCFC model obtained PDTG showed less oscillations compared to the KCF, Mean-shift and SAMF models. The KCK obtained PDTG (see the black curve

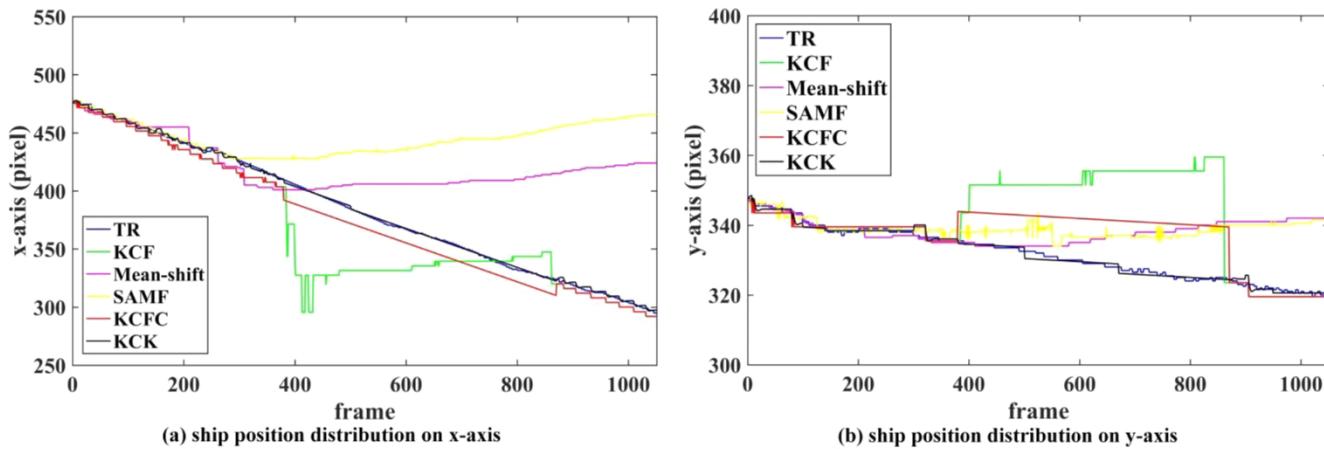


Fig. 4. Ship position distributions on both x and y axis for video #1.

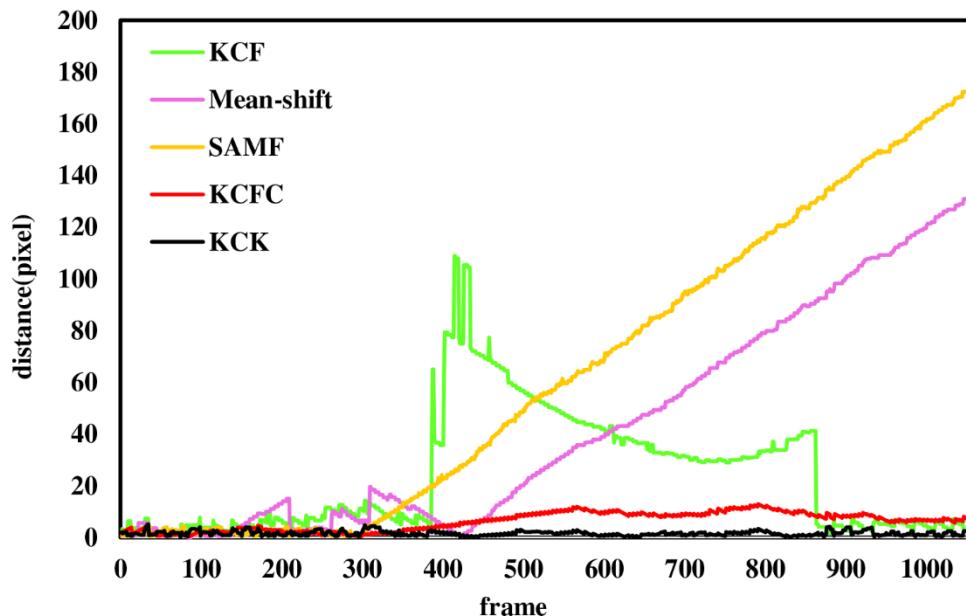


Fig. 5. PDTG distributions for various ship tracking models on video #1.

in Fig. 5) are quite close to zero, which indicates better tracking accuracy than the above-mentioned ship tracking models. By carefully checking the PDTG variation details, we found that the KCF relevant PDTG curve showed different tendency in details compared to the counterparts. More specifically, the KCF model obtained satisfied tracking performance when the target ship was not occluded by obstacle (i.e., before and after ship occlusion procedure). The PDTG variations confirmed that our proposed ship tracker outperformed the counterparts in tackling the ship tracking challenge in video #1, which is consisted with previous analysis.

The ship tracking accuracy was further measured with the RMSE, MAD, and MSE indicators, which are shown in Table 2. It is noted that

Table 2
Ship tracking accuracy comparison for various models on video #1.

| | RMSE | Mad | MSE |
|------------|-------|-------|---------|
| Mean-shift | 41.26 | 36.02 | 1702.39 |
| SAMF | 57.01 | 50.56 | 2556.31 |
| KCF | 23.12 | 19.79 | 534.53 |
| KCFC | 3.47 | 3.08 | 12.06 |
| KCK | 0.90 | 0.81 | 0.71 |

the RMSE for the Mean-shift model is 41.26, which is 45.8 times larger than that of the KCK. Moreover, the Mean-shift MAD is 44.47 times larger than that of the KCK model, and the MSE for the Mean-shift model (i.e., 1702.39) is significantly larger than the KCK counterpart. The RMSE, MAD and MSE indicators obtained by the SAMF tracker (which are 57.01, 50.56 and 2556.31) are significantly larger than those of counterparts. The RMSE and MAD values for the KCF model are approximately 25 times higher than those of the proposed KCK model considering that the two indicators for the KCK model are 0.9 and 0.81, respectively. Meanwhile, MSE indicator of the KCF model is about 750 times larger than that of the KCK (which is 0.71). The RMSE and MAD indicators for the KCFC model are approximately four-folds higher than that of the KCK model, and the KCFC obtained MSE is 15 times higher than that of the KCK model. Based on the analysis about RMSE, MAD and MSE, we concluded that the proposed KCK model obtained higher tracking accuracy compared to the other four tracking models dealing with the ship tracking challenge in video #1.

For the purpose of better visualizing tracking accuracy, we presented ship tracking results on typical frames (see Fig. 6). We tailed the region of interest in each frame for the purpose of helping readers better understand various trackers' tracking performance. The five ship trackers

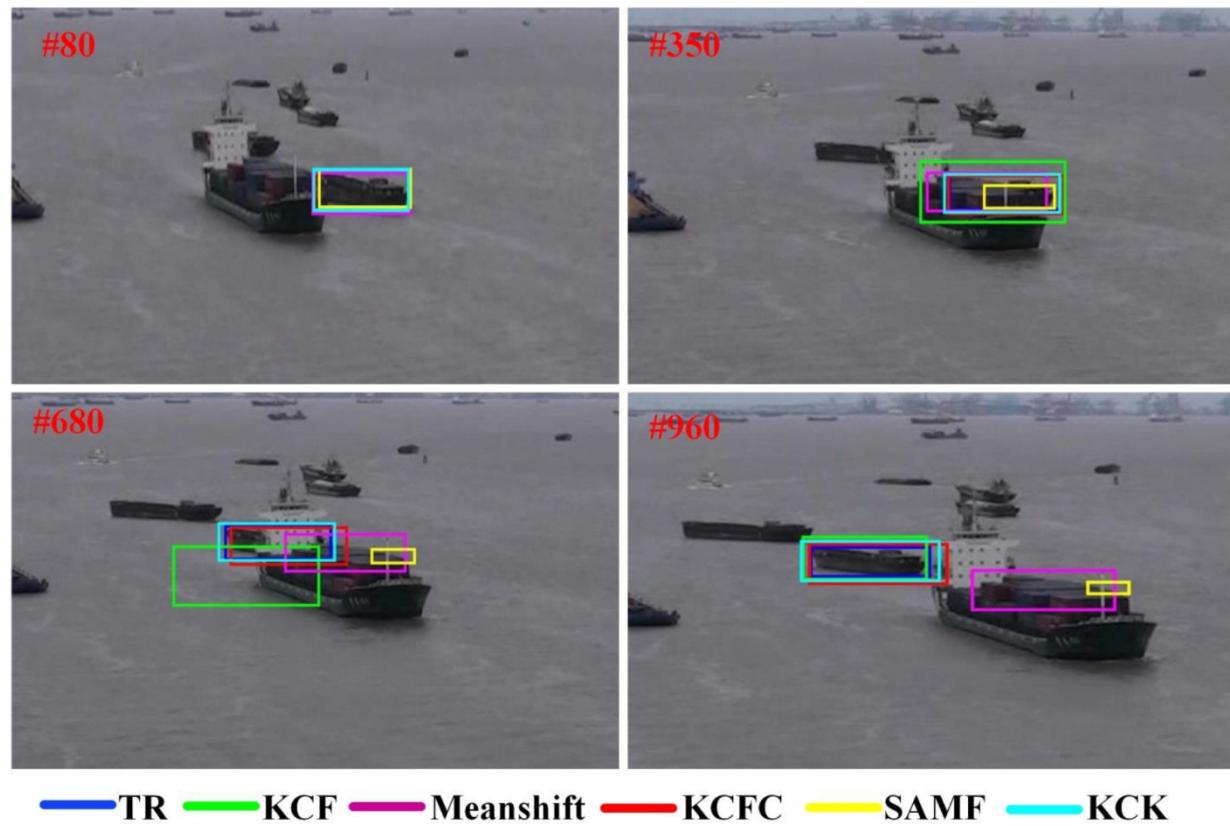


Fig. 6. Various ship trackers' performance on typical frames for video #1.

showed accurate ship tracking results before the ship occlusion happened (see the tracking results on the frame #80 in Fig. 6). The KCF, Mean-shift and SAMF trackers failed to track the target ship in the ship occlusion frames (see the tracking results on frame #350, #680 and #960 in Fig. 6). Note that the SAMF tracked ship (see the yellow rectangle in frame #960 in Fig. 6) size was obvious smaller than those of the counterparts. KCFC model obtained more accurate tracking results compared to KCF, Mean-shift and SAMF trackers. Moreover, the KCK tracked ship positions were very close to the ground truth (as shown in the dark green curves in Fig. 6). The above qualitative and quantitative analysis indicated that the proposed KCK ship tracker is more robust to the ship tracking challenge (i.e., the small ship target is occluded by the large ship).

2) Ship tracking performance analysis from video #2 to #4

For the purpose of verifying ship tracker robustness, we implemented various ship trackers on video #2, #3 and #4 with different ship tracking challenges (more details were illustrated in the previous section). For simplicity, we analyze ship tracking results from perspective of ship position variation and tracking accuracy (in terms of RMSE, MAD and MSE). The ship position variation for video #2, #3 and #4 were shown in Figs. 7, 8 and 9. We observed that the KCF, Mean-shift and SAMF models lost the target ships since the target ship was occluded by obstacle, which is similar to that of video #1. Though the KCFC model did not lose the target ship, abnormal oscillations can be found in the x and y axis distributions (see the red curves in the Figs. 7, 8 and 9). It is found that ship positions obtained by the proposed KCK tracking model were smoother and closer to the ground truth ship as shown in each subplot in Figs. 7, 8 and 9. After carefully checking the ship position distribution on x axis for video #2 (see Fig. 7(a)), we found that the KCK model successfully keeps the ship trajectory details, and the abnormal

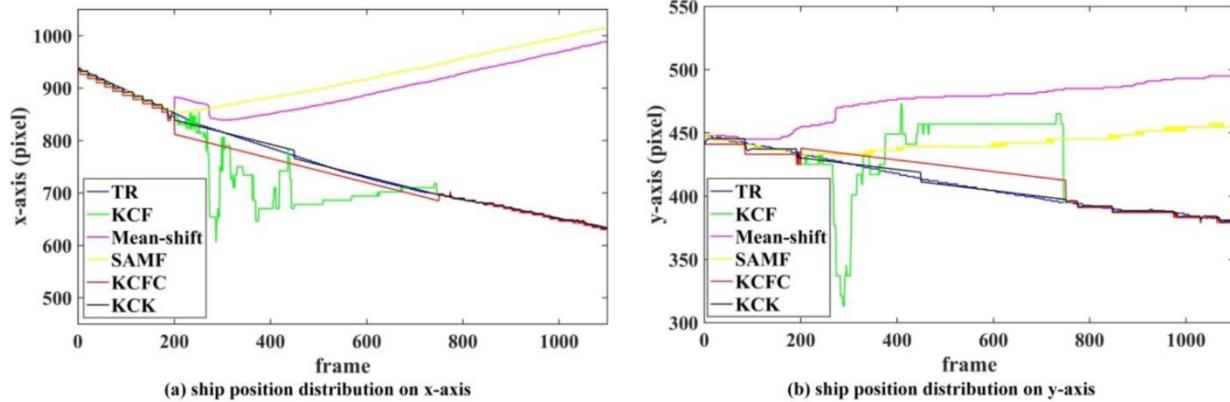


Fig. 7. Ship position distributions on both x and y axis for video #2.

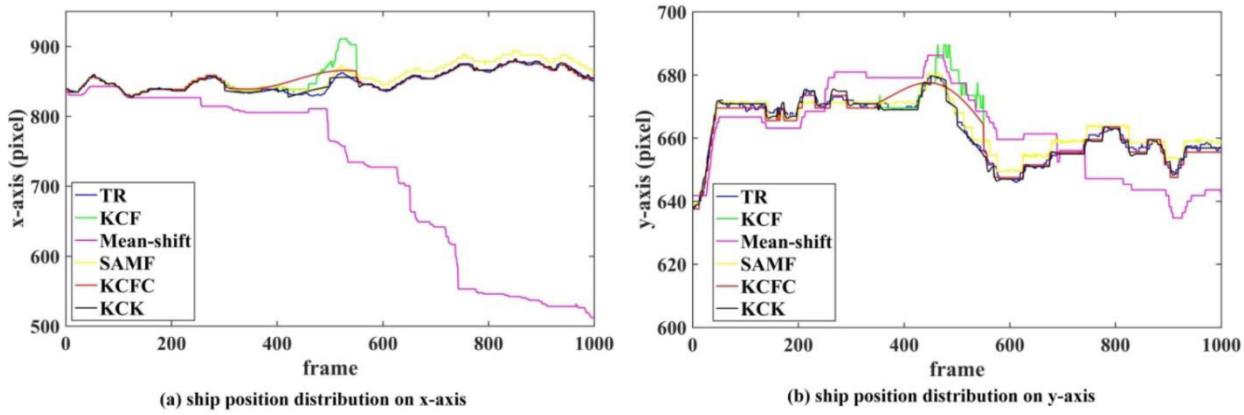


Fig. 8. Ship position distributions on both x and y axis for video #3.

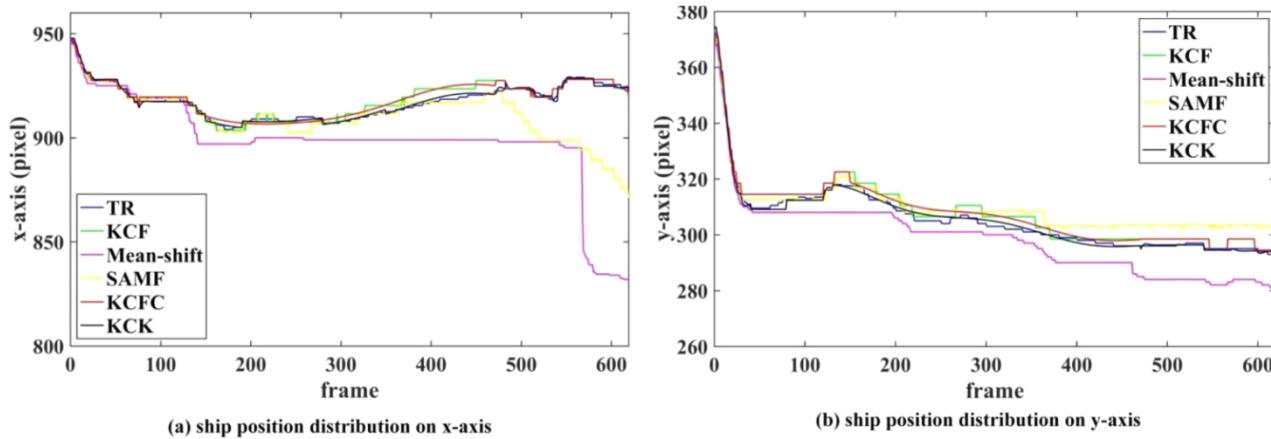


Fig. 9. Ship position distributions on both x and y axis for video #4.

trajectory outliers were suppressed. The ship position distributions on the y-axis for video #2 (see Fig. 7(b)) showed similar variation tendency.

Ship tracking accuracy for the video #2, #3 and #4 were further quantified with the RMSE, MAD and MSE (see Tables 3, 4 and 5). The RMSE indicators for the proposed KCK model for video #2, #3 and #4 are 2.47, 1.77 and 0.81, which are approximately one-fifth, one-quarter, one-half to the counterparts of the KCFC model. The Mean-shift model obtained RMSE were higher than those of SAMF, KCF, KCFC and KCK, which are 116.23, 113.19 and 22.74, respectively. Both the MAD and MSE statistics showed similar variation as that of RMSE in tackling the ship tracking challenges in video #2, #3 and #4. More specifically, the minimal MAD indicators are 1.77, 4.50 and 0.71, while the MSE are 6.09, 3.15 and 0.65, which are obtained by proposed KCK tracker. Based on the above quantitative analysis, we concluded that the KCK tracking model can successfully handle typical ship tracking challenges with accurate tracking results.

Table 3
Ship tracking accuracy comparison for various models on video #2.

| | RMSE | Mad | MSE |
|------------|-------------|-------------|-------------|
| Mean-shift | 116.23 | 103.20 | 13,509.41 |
| SAMF | 52.51 | 113.36 | 2757.30 |
| KCF | 41.31 | 33.47 | 1706.52 |
| KCFC | 13.17 | 9.55 | 173.51 |
| KCK | 2.47 | 1.77 | 6.09 |

Table 4
Ship tracking accuracy comparison for various models on video #3.

| | RMSE | Mad | MSE |
|------------|-------------|-------------|-------------|
| Mean-shift | 113.19 | 117.34 | 12,811.98 |
| SAMF | 15.13 | 5.19 | 228.92 |
| KCF | 17.33 | 7.60 | 300.33 |
| KCFC | 6.87 | 4.50 | 46.89 |
| KCK | 1.77 | 1.24 | 3.15 |

Table 5
Ship tracking accuracy comparison for various models on video #4.

| | RMSE | Mad | MSE |
|------------|-------------|-------------|-------------|
| Mean-shift | 22.74 | 14.43 | 517.11 |
| SAMF | 11.85 | 8.57 | 140.42 |
| KCF | 1.75 | 1.46 | 3.06 |
| KCFC | 1.49 | 1.26 | 2.21 |
| KCK | 0.81 | 0.71 | 0.65 |

4. Conclusion

Accurate ship tracking provide important on-site traffic information, which can significantly benefit maritime traffic efficacy and management in remote controlling manner. Ship tracking under occlusion scenario is a type of common bottlenecks, which impose big challenge to obtain high-resolution ship trajectories for remote maritime traffic awareness. We proposed a hybrid kernelized correlation filter and anomaly cleaning framework to tackle typical ship occlusion tracking

challenges. Firstly, the proposed KCK ship tracking model obtained ship trajectories in maritime images with kernelized correlation filter. Then, the KCK model removed the obvious ship tracking outliers with curve fitting mechanism. After that, the data quality of the ship trajectories was further improved by the Kalman filter in the manner of suppressing anomaly oscillations. We have tested the proposed ship tracker performance on four typical ship occlusion videos. The experimental results indicated the efficacy of proposed tracker due to that the KCK-obtained average RMSE, MAD and MSE indicators are 1.49, 1.13 and 2.65, which are significantly smaller than those of the counterparts.

In future, we can expand our work by exploring the following directions. First, we have tested the proposed model performance under single ship tracking task. Testing the model performance on the multiple ships tracking task can be an interesting expansion. Second, our ship tracker is accurately initialized (to be specific, ground truth ship position in the first frame is employed to initialize the ship tracker). We can further exploit ship tracker performance with non-perfect initialization. Third, additional general-purpose ship trackers (such as deep learning models) can be implemented to provide more holistic performance results. Last but not least, we can further verify our model performance under various adverse weather conditions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was jointly supported by the National Natural Science Foundation of China (5207120072061016, 51978069, 51708094, 51867009, 61663027), Shanghai Committee of Science and Technology, China (18DZ1206300, 18040501700, 18295801100, 17595810300).

References

- Abebe, M., Shin, Y., Noh, Y., Lee, S., Lee, I., 2020. Machine learning approaches for ship speed prediction towards energy efficient shipping. *Appl. Sci.* 10, 2325.
- Chen, X., Li, Z., Yang, Y., Qi, L., Ke, R., 2020b. High-resolution vehicle trajectory extraction and denoising from aerial videos. *IEEE Trans. Intell. Transp. Syst.* 1–13.
- Chen, X., Qi, L., Yang, Y., Luo, Q., Postolache, O., Tang, J., et al., 2020c. Video-based detection infrastructure enhancement for automated ship recognition and behavior analysis. *J. Adv. Transp.* 2020.
- Chen, X., Wang, S., Shi, C., Wu, H., Zhao, J., Fu, J., 2019. Robust ship tracking via multi-view learning and sparse representation. *J. Navig.* 72, 176–192.
- Chen, X., Wu, S., Shi, C., Huang, Y., Yang, Y., Ke, R., et al., 2020d. Sensing data supported traffic flow prediction via denoising schemes and ANN: a comparison. *IEEE Sens. J.* 1.
- Chen, X., Xu, X., Yang, Y., Wu, H., Tang, J., Zhao, J., 2020e. Augmented Ship tracking under occlusion conditions from maritime surveillance videos. *IEEE Access* 8, 42884–42897.
- Chen, X., Yang, Y., Wang, S., Wu, H., Tang, J., Zhao, J., et al., 2020a. Ship type recognition via a coarse-to-fine cascaded convolution neural network. *J. Navig.* 73, 813–832.
- Chen, Z., Xue, J., Wu, C., Qin, L., Liu, L., Cheng, X., 2018. Classification of vessel motion pattern in inland waterways based on Automatic Identification System. *Ocean Eng.* 161, 69–76.
- Cui, Z., Li, Q., Cao, Z., Liu, N., 2019. Dense attention pyramid networks for multi-scale ship detection in SAR images. *IEEE Trans. Geosci. Remote Sens.* 57, 8983–8997.
- Fang, Z., Jian-yu, L., Jin-jun, T., Xiao, W., Fei, G., 2018b. Identifying activities and trips with GPS data. *IET Intel. Transport Syst.* 12, 884–890.
- Fang, Z., Yu, H., Ke, R., Shaw, S.-L., Peng, G., 2018a. Automatic identification system-based approach for assessing the near-miss collision risk dynamics of ships in ports. *IEEE Trans. Intell. Transp. Syst.* 20, 534–543.
- Frodella, W., Elashvili, M., Spizzichino, D., Gigli, G., Adikashvili, L., Vacheishvili, N., et al., 2020. Combining InfraRed thermography and UAV digital photogrammetry for the protection and conservation of Rupelstrian Cultural Heritage Sites in Georgia: a methodological application. *Remote Sens. (Basel)* 12, 892.
- Gray, R.M., 2006. Toeplitz and circulant matrices: a review. *Found. Trends® Commun. Inf. Theory* 2, 155–239.
- Henriques, J.F., Caseiro, R., Martins, P., Batista, J., 2015. High-speed tracking with kernelized correlation filters. *IEEE Trans. Pattern Anal. Mach. Intell.* 37, 583–596.
- Kang, X., Song, B., Guo, J., Du, X.J., Guizani, M., 2019. A self-selective correlation ship tracking method for smart ocean systems. *Sensors* 19. Feb 2.
- Kim, K., Hong, S., Choi, B., Kim, E., 2018. Probabilistic ship detection and classification using deep learning. *Appl. Sci.* 8, 936.
- Lang, H., Wu, S., Xu, Y., 2018. Ship classification in SAR images improved by AIS knowledge transfer. *IEEE Geosci. Remote Sens. Lett.* 15, 439–443.
- Li, Y., Zhu, J., 2014. A scale adaptive kernel correlation filter tracker with feature integration. In: European Conference on Computer Vision, pp. 254–265.
- Liang, C., Zhang, X., Han, X., 2020. Route planning and track keeping control for ships based on the leader-vertex ant colony and nonlinear feedback algorithms. *Appl. Ocean Res.* 101, 102239.
- Liu, W., Ma, L., Chen, H., 2018b. Arbitrary-oriented ship detection framework in optical remote-sensing images. *IEEE Geosci. Remote Sens. Lett.* 15, 937–941.
- Liu, Y., Yao, L., Xiong, W., Zhou, Z., 2018a. GF-4 satellite and automatic identification system data fusion for ship tracking. *IEEE Geosci. Remote Sens. Lett.* 16, 281–285.
- Ong, C.S., Smola, A.J., Williamson, R.C., 2005. Learning the kernel with hyperkernels. *J. Mach. Learn. Res.* 6, 1043–1071.
- Park, J., Kim, J., Son, N., 2015. Passive target tracking of marine traffic ships using onboard monocular camera for unmanned surface vessel. *Electron. Lett.* 51, 987–989.
- Shao, Z., Wang, L., Wang, Z., Du, W., Wu, W., 2019. Saliency-aware convolution neural network for ship detection in surveillance video. *IEEE Trans. Circuits Syst. Video Technol.*
- Shi, W., Xiao, Y., Chen, Z., McLaughlin, H., Li, K.X., 2018. Evolution of green shipping research: themes and methods. *Maritime Policy Manage.* 45, 863–876.
- Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. *Stat. Comput.* 14, 199–222.
- Smolenski, R., Benysek, G., Malinowski, M., Sedlak, M., Styński, S., Jasinski, M., 2018. Ship-to-shore versus shore-to-ship synchronization strategy. *IEEE Trans. Energy Convers.* 33, 1787–1796.
- Tang, J., Gao, F., Liu, F., Chen, X., 2020. A denoising scheme-based traffic flow prediction model: combination of ensemble empirical mode decomposition and fuzzy C-means neural network. *IEEE Access* 8, 11546–11559.
- Wan, X., Li, Y., Xia, C., Wu, M., Liang, J., Wang, N., 2016. A T-wave alternans assessment method based on least squares curve fitting technique. *Measurement* 86, 93–100.
- Yang, K., Wu, X., Zhu, Z., Xu, J., Wan, Z., Chang, Y., et al., 2019. A high-confidence model updating correlation filtering tracker with scale adaptation for visual target tracking. *Optik (Stuttgart)*, 163730.
- Zhang, C., Xia, X., Bo, S., Yao, K., 2017. Target track recording system based on kernelized correlation filters tracking algorithm. *Comput. Syst. Appl.* 26, 113–118.
- Zhang, L., Wang, Y., Sun, H., Yao, Z., Wu, P., 2016. Adaptive scale object tracking with kernelized correlation filters. *Opt. Precis. Eng.* 24, 448–459.
- Zhou, F., Pan, S., Jiang, J., 2019. Verification of AIS data by using video images taken by a UAV. *J. Navig.* 72, 1345–1358.
- Zhou, L., Zhang, J., 2019. Combined Kalman filter and multifeature fusion Siamese network for real-time visual tracking. *Sensors* 19, 2201.
- Zong, F., Wu, T., Jia, H., 2019. Taxi drivers' cruising patterns—insights from taxi GPS traces. *IEEE Trans. Intell. Transp. Syst.* 20, 571–582.