IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 31, 2019, accepted April 14, 2019, date of publication April 24, 2019, date of current version May 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2912976

Infrared Small Target Detection Based on Spatial-Temporal Enhancement Using Quaternion Discrete Cosine Transform

PING ZHANG[1], XIAOWEI WANG[1], XIAOYANG WANG[2], CHUN FEI[3], AND ZHENGKUI GUO[1]

[1]School of Photoelectric Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
[2]Department of Electrical and Electronic Engineering, University of Bristol, Bristol BS8 1TH, U.K.
[3]School of Computer Science & Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

Corresponding author: Ping Zhang (pingzh@uestc.edu.cn)

This work was supported by the Science and Technology Planning Project of Sichuan Province, China, under Grant 2018GZ0166 and Grant 2019YFG0307.

ABSTRACT Infrared small target detection plays an important role in the infrared search and track system. However, infrared small target images often suffer from low contrast. In this paper, we propose an infrared small target detection method that improves the target contrast and suppresses background clutters based on spatial-temporal enhancement using the quaternion discrete cosine transform (QDCT). The proposed method is twofold: 1) we propose to detect the infrared small target by constructing the quaternion feature map for infrared images. The quaternion integrates four feature maps, including the kurtosis feature, two directional feature maps extracted by steerable filtering in spatial domain and motion feature in the temporal domain. 2) Then the quaternion is input into QDCT, and the saliency maps of each feature channel can be obtained by sign function processing. The final detection result is obtained by inverse QDCT to the quaternion. Compared to several state-of-the-art algorithms, the proposed method has a lower false alarm rate when the same positive detection rate is achieved.

INDEX TERMS Infrared (IR) small target detection, kurtosis feature, steerable filter, quaternion discrete cosine transform (QDCT).

I. INTRODUCTION

Detecting small targets from infrared (IR) surveillance videos is a highly demanded application in both military and business fields. Accurate information of the interested targets is necessary for these applications. However, the infrared small targets, which size ranges from 2×2 to 9×9 pixels, may suffer from low contrast, low signal-to-clutter ratio, and lack of some critical information like color, shape, and texture because of long imaging distance and complex backgrounds. This will lead to a considerable false alarm rate [1]–[6]. Therefore, it is of great practical significance to detect small targets in IR images with complex backgrounds.

Several common infrared small target images are listed in Fig. 1. There are several specific problems in detecting IR small targets, which cannot be seen in other targets' detection. Firstly, the background and the target are both complex and variable, such as cirrus, regiment, banded clouds and so on. Secondly, compared with the visible images, IR images seem to be much simpler, which lack texture information, color information, and poor imaging quality as shown in Fig. 1.

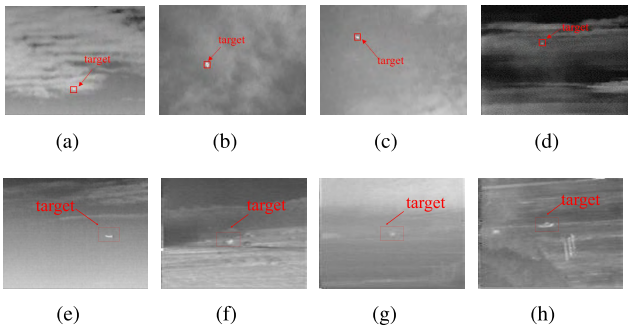


FIGURE 1. Representative IR common background and target types. (a) Cirrus cloud; (b) Flocculent target; (c) Regiment cloud; (d) Banded cloud; (e) Strip target; (f) Land dot target; (g) Land weak target; (h) Land boat target.

The associate editor coordinating the review of this manuscript and approving it for publication was Sudipta Roy.

The lack of texture and color information means there are fewer features available for distinguishing targets from the backgrounds and the unclear boundary between different parts in images. Moreover, little interference caused by the small size of targets might lead to faulty detection. These characteristics of IR images pose a big challenge when performing small target detection. Thus, regular detection approaches in visible images are no longer valid in IR images, especially in complex scenes [7], [8] and it is necessary to develop specific IR small target detection methods.

Existing IR small target detection methods can be broadly classified as spatial domain or transform domain methods. Spatial domain methods mainly extract the difference between the foregrounds and the backgrounds by measuring image contrast, including local contrast measure based methods [9]–[11], morphological filtering based methods [12], [13], statistical regression-based methods [14] and human visual system (HVS) [15], [16] based methods. The local contrast measure based methods calculate the contrast value according to the intensity difference between the small target area and its surrounding background area and then separates the small target by setting a threshold. [9]. The morphological filtering based methods use morphological operators to design a filter to suppress the backgrounds clutters [12]. Statistical regression-based methods perform detection by estimating the distribution of the backgrounds or the regression function with the input data [14]. HVS based methods achieve detection by improving the local contrast between the targets and the backgrounds [15]. Transform domain methods often use a special transform such as wavelet transform [17], Fourier transform [18], phase spectrum of quaternion Fourier transform (PQFT) [19] and so on. Detection in complex backgrounds performs better in transform domain than in the spatial domain in spite of higher computation complexity. These transform domain methods quantitatively define a frequency difference between the target and the background in the transform domain, and achieve detection by magnifying the measured difference.

Since more details in an image imply higher phase information values in the transform domain, good detection results can be achieved by extracting the phase information. According to this, Qi *et al.* [18] used Phase spectrum of Fourier transform (PFT) to enhance targets. After that, Qi *et al.* [19] put forward a new method using a phase spectrum of quaternion Fourier transform (PQFT) to detect infrared small targets. In this model, they used the facet model [20] to get four directional features which are applied to perform PQFT. Then the amplitude spectrum values are set to a constant value for each PQFT feature so that the saliency map is constituted by the inverse transform of the phase spectrum. However, these two methods only use the phase spectrum to recover signals, which will cause some information loss. It may also lead to missed detection in the test results. Yang *et al.* proposed a small target detection method based on another direction feature in [21]. However, the filter is realized by Gauss kernel function, which can blur the original image while extracting the edge direction features. Besides, comparing with Fourier transform, cosine transform only performs on real even function and contains only real cosine terms. It is a real number of domain transformation. The actual signals are mostly in the real number domain, which makes the Fourier transform very redundant. Therefore, cosine transform greatly reduces computational complexity.

Considering the advantages of cosine transform, a method of infrared small target detection based on spatial-temporal enhancement using quaternion Discrete Cosine Transform (QDCT) is proposed in this paper. QDCT could support four channels, providing more information for DCT, whereas traditional DCT can only use gray information. The motivation is mainly based on the imaging characteristics of small targets, which might obtain more information from the transform domain by combining different weak features. Besides, considering the fact that a small target often makes HVS have a higher response, the infrared small target detection can be formulated as a salient region detection problem [2]. Knowing this, we apply QDCT in infrared sequences small target detection. Our method takes advantage of motion feature in the temporal domain, kurtosis feature, and two directional features in the spatial domain: 1). The correlation between the frames is modeled by the motion feature map, which is used to improve the validity of sequence target detection. 2). Kurtosis reflects the difference between the distribution of image patch and the distribution of target, so there are different kurtosis values at the target and the background respectively, which is conducive to distinguishing them. 3). Combining isotropic characteristics of small infrared targets with Gauss-like distribution and the local direction characteristics of background, a steerable filter is used as a directional filter to filter each frame of the input video, then two directional characteristic maps are obtained. Four characteristic maps constitute a quaternion. Through the above three steps, the quaternion of input is obtained. Then we process the quaternion by QDCT, a sign function, and inverse QDCT sequentially to get the spatial representation (also a quaternion). Finally, this amplitude spectrum is filtered by a Gauss low-pass filter to obtain the final detection result of infrared small targets. Among them, the sign function plays a role in suppressing the background and detecting the significant small infrared targets with a Gaussian-like distribution.

The main contributions of this paper are as follows: (1). QDCT is applied to infrared small target detection field by designing specific features for QDCT for the first time, we have successfully detected the infrared small targets with a high detection rate and low false alarm rate. (2). For the first time, the relationship between kurtosis and contrast is analyzed and discussed and then the kurtosis is used as the indicator of contrast to reflect the difference between the target area and the surrounding background, thus promoting the detection of infrared small targets. (3). Motion feature is specially designed for high-speed infrared small targets and uses motion accumulation to extract the correlation between adjacent frames while retaining the motion trajectory.

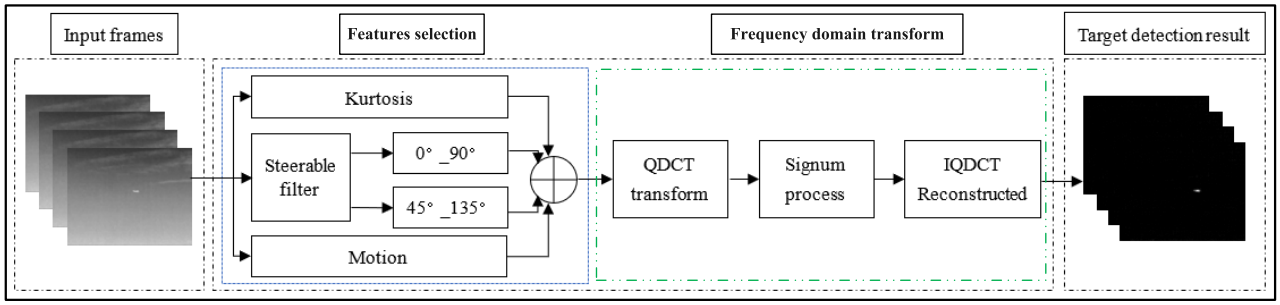


FIGURE 2. The framework of the proposed method, composed of input frames, the specific algorithm and target detection results. (The blue rectangle represents building the quaternion.)

The remainder of this paper is organized as follows. In Section II, the brief introduction of QDCT is presented. Section III introduces elaborate steps of the proposed method, including the reasons for choosing four features, construction of the quaternion, the process of QDCT, and the final small target detection result. In Section IV, several experiments on public data sets and relative discussions are presented to demonstrate the efficiency of the proposed method. An overall conclusion is drawn for a summary of this paper in Section V.

II. RELATED WORK

QDCT was first proposed by Feng and Hu [23]. They extended the discrete cosine transform (DCT) to the quaternion field and presented a simple but efficient algorithm to calculate the QDCT of a quaternion matrix using its Cayley-Dickson form. Recently, QDCT has been applied to detection, such as face detection [24] and flame detection [25]. All these two papers achieve target detection by extracting features into QDCT. Therefore, we endeavor to explore the features of small targets and input them into QDCT, and then salient small target detection is realized by sign function processing and IQDCT. Detailed QDCT algorithm process is described below.

A quaternion can be represented in a hypercomplex form as [25]:

$$I_q(m,n) = x_1 + x_2 \cdot i + x_3 \cdot j + x_4 \cdot k \tag{1}$$

where $x_1, x_2, x_3, x_4 \in R$, and $x_1, x_2, x_3, x_4 \in R$ satisfy the following multiplication rules: $i^2 = j^2 = k^2 = -1, i \cdot j = -j \cdot i = k, j \cdot k = -k \cdot j = i, k \cdot i = -i \cdot k = j$. This demonstrates that the multiplication of quaternions is not commutative. Thus, the form of Forward QDCT (FQDCT) and Inverse QDCT (IQDCT) have two categories, left-handed form and right-handed form [23]. In this paper, we will use the left-handed $FQDCT^L$ and $IQDCT^L$ as default QDCT and IQDCT, respectively to indicate the process of saliency detection.

$$FQDCT^L(u,v) = \alpha_u^M \alpha_v^N \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mu_q I_q(m,n) N(u,v,m,n) \tag{2}$$

where $I_q(m,n)$ is a two-dimensional $M \times N$ quaternion matrix. μ_q is a unit pure virtual quart which meets the constraint that ${\mu_q}^2 = -1$. The values of α_u^M, α_v^N and $N(u,v,m,n)$ are taken as in [25].

Accordingly, the $IQDCT^L$ is defined as follows:

$$IQDCT^L(m,n) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \alpha_u^M \alpha_v^N \mu_q C_q(u,v) N(u,v,m,n) \tag{3}$$

where $C_q(u,v)$ is a two-dimensional $M \times N$ quaternion matrix, which represents the result of $FQDCT^L$ of $I_q(m,n)$.

III. OUR METHOD

According to the principle of infrared small target imaging, it is known that small targets can be used with fewer features, which is also a difficult problem in detecting problems. We should use both the features of the target and the local background for detection as much as possible. In a sequence, the target moves at a high speed, and the background exhibits local directional. These two types of features are obviously beneficial for the detection of small targets. In addition, the small target brightness is large, and the difference with the surrounding background is obvious, so the use of contrast characteristics is also the mainstream direction of infrared small target detection in recent years.

Fig. 2 shows the framework of the proposed method of infrared small target detection. We first build the quaternion for infrared images by constructing the four feature maps: the kurtosis, two directional maps, and the motion feature map. Then, we enhance the salient region in the transform domain by QDCT to the quaternion. Finally, the target is recovered by applying the inverse QDCT (IQDCT) to the result of the saliency detection.

A. CONSTRUCT THE DIRECTIONAL FEATURE MAP BY STEERABLE FILTER

Steerable filter [26] is an orientation-adjustable convolution kernel used for feature extraction and image enhancement that can be expressed via a linear combination of a small set of rotated versions of itself. The steerable filter can get the edge of an image by generating templates in different directions and then deconvoluting the image with templates in different

directions. In infrared small target detection, we use steerable filter to extract direction features, but also to maintain the edges of the small target, which is an important reason why it is superior to other directional filters [27]. The directional feature map $D_\theta(x, y, t)$ is given by:

$$D_\theta(x, y, t) = f(x, y, t) * G_\theta(x, y, t). \tag{4}$$

where $G_\theta(x, y, t)$ denotes the steerable filter with directional angle $\theta \in [0, 2\pi]$. $f(x, y, t)$ represents the input image in which t is frame index satisfying that $t = 1, 2, \ldots, N$. And N is the number of video frames. $*$ signifies the convolution operator.

In order to make a balance between computational complexity and detection performance, we extract four direction feature maps, i.e., $D_0, D_{45}, D_{90}, D_{135}$ are extracted. The combination of mutually orthogonal directions has a certain effect on background suppression (donated by 0°_90° and 45°_135°), which can be given as follows, where f_1 and f_2 represent the two channels of QDCT respectively.

$$f_1(x, y, t) = D_0(x, y, t) \cdot D_{90}(x, y, t) \tag{5}$$
$$f_2(x, y, t) = D_{45}(x, y, t) \cdot D_{135}(x, y, t) \tag{6}$$

Here, $\cdot$ means multiplication of corresponding elements.

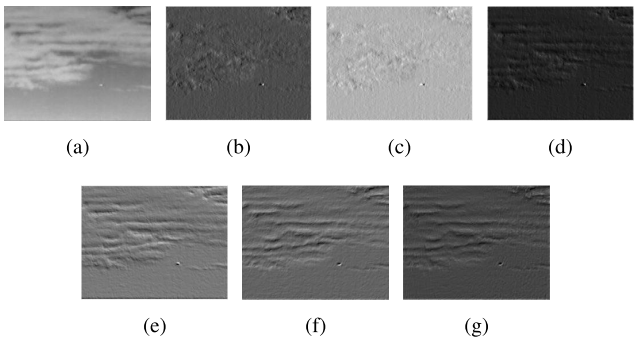


FIGURE 3. Directional feature maps and edge preserving result are obtained using steerable filter. (a) Original image; (b) 0°; (c) 90°; (d) 0°_90°; (e) 45°; (f) 135°; (g) 45°_135°.

In order to highlight the advantages of the steerable filter, the cirrus image with a complex background is selected as an example to illustrate the effect of using the steerable filter to extract direction features and preserve edges. Fig. 3 shows the directional feature maps extracted using a steerable filter. In order to prove that the combination of directions can suppress the background, we use the mean value and Signal-noise ratio (SNR) to evaluate it in Table. 1. SNR reflects the ratio of target signal to the background signal. The larger the value, the more prominent the small target is. Mean value reflects the overall brightness of the image, the smaller the value, the better the background suppression is. By describing SNR and mean value of the image, it can be seen that Fig. 3(d) and Fig. 3(g) have a darker background and less clutter, which makes the small target even more prominent. Therefore, the combination of orthogonal directions is beneficial to target detection.

TABLE 1. Compare the mean and SNR of different directional feature maps and after combination.

Indices	0°	90°	0°_90°	45°	135°	45°_135°
Mean (μ)	82.1575	122.2798	**41.1614**	179.3436	102.1427	**72.0388**
SNR	63.6714	57.9094	**76.5298**	53.3519	59.7876	**64.5569**

B. KURTOSIS FEATURE MAP

Considering that the intensity of infrared radiation changes at the boundary between the small target and the background, the small target can be smoothly distributed in a small range, and the local contrast of the small target is high, Therefore, the use of contrast features facilitates the detection of small targets. In addition, due to the small target has a Gaussian-like distribution characteristic, the statistical kurtosis reflects the difference between the sample distribution and the normal distribution. Therefore, in image processing, the kurtosis reflects the difference between the distribution of selected image patch and the Gaussian-like distribution of the small target.

According to the statistical principle of kurtosis, the statistical kurtosis of local image patch at location (x, y) is used to represent the kurtosis value at (x, y). The specific kurtosis feature channel can be given by:

$$\begin{aligned} kur(f(x, y, t)) &= \kappa - C = \frac{E(X-\mu)^4}{\sigma^4} - C \\ &= \frac{1}{mn} \frac{\sum_{i=1}^{mn} (f(x, y, t) - \mu)^4}{\sigma^4} - C \end{aligned} \tag{7}$$

where κ is the four-order center distance of image patch distribution to be processed, $m \times n$ is the size of the selected image patch. $f(x, y, t)$ is the gray value of the t frame image at (x, y). μ and σ denote the gray mean and standard deviation of image patch respectively. C represents the kurtosis value of the target distribution, which is a fixed value for a specific target. For example, the C of the normal distribution is 3. In the algorithm, we divide the infrared image into small patches of size 5×5 and calculate kurtosis of the patch center pixel instead of local kurtosis of each patch.

Therefore, the third kurtosis feature channel of an image for QDCT is given by f_3:

$$f_3(x, y, t) = kur(f(x, y, t)) \tag{8}$$

Fig. 4 shows the maps related to the kurtosis feature. Fig. 4(a) is the schematic diagram of a local image patch with a central position (x, y). The kurtosis takes full account of all the sample points in the image patch. Fig. 4(c) is the kurtosis feature map of Fig. 4(b). As can be seen from Fig. 4(c), the kurtosis around the target is larger than that around the background. Since the target in the image is strip-shaped, even if the image patch is selected as 5×5, the kurtosis feature map is rectangular distribution. That is to say, the kurtosis feature map is similar to the shape of the target, and there is a clear boundary between the target and the background. This feature is helpful to the detection of the target.

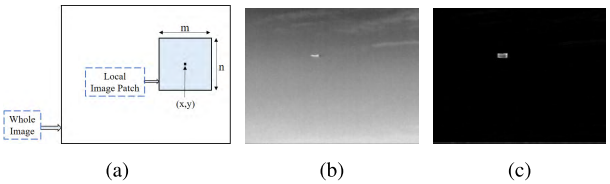


FIGURE 4. Kurtosis feature map. (a) The schematic diagram of the selected local image patch; (b) Original infrared image (the 3th frame of Seq6); (c) Kurtosis feature map.

C. MOTION FEATURE MAP

Because in a group of infrared sequences, the motion of the target is fast and continuous, while the change of background and noise is random. At present, most of the methods of small target detection in infrared sequences are based on a single frame, without considering the motion information, which often results in false detection in the whole video detection. This is a fatal problem in practice. In order to solve the above problem, the introduction of the motion feature is conducive to small target detection. Therefore, we introduce a motion charge (MC) method [28] to extract motion feature, which can not only eliminate singularities and strong background changes in isolated background, but also reduce the impact of background changes on target detection while highlighting the target, and can also characterize the cumulative effect of the target in time and reflect the motion trajectory of the target. Assuming that $f(x, y, t)$ is the t frame of the current infrared sequence, its motion information can be expressed as:

$$Mov(x, y, t) = \begin{cases} 0 & if \ f(x, y, t) = f(x, y, t-1) \\ 1 & if \ f(x, y, t) \neq f(x, y, t-1) \end{cases} \quad (9)$$

Mov represents whether there is motion between adjacent frames. The cumulative motion feature is defined as:

$$Ch_{Mov}(x, y, t) = \begin{cases} Ch_{min} & if \ \ Mov(x, y, t) = 1 \\ Ch & if \ \ Mov(x, y, t) = 0 \end{cases} \quad (10)$$

where $Ch = min(Ch_{Mov}(x, y, t-1) + A, Ch_{max})$. When $t = 0$, the cumulative motion is initialized to $Ch_{Mov}(x, y, 0) = Ch_{\max}$:

In this paper, we inverse the cumulative result image as the motion feature, expressed as:

$$m(x, y, t) = 255 - Ch_{Mov}(x, y, t) \quad (11)$$

Here A is the incremental parameter of motion cumulative, which is set to $C = 1$. $Ch_{\max}$ and $Ch_{\min}$ are set to 255 and 0 respectively. It can be seen from Eq. (10), if $Ch_{Mov} = Ch_{\min}$, the motion is detected at (x, y, t); If $Ch_{Mov} = Ch_{\max}$, no motion is detected and the time of the last motion is unknown; If $Ch_{Mov}(x, y, t) = Ch_{\min} + k \cdot A$, it means that there is no motion detected at (x, y, t), but the last motion occurred at $t - k \cdot \Delta t$. Here, $k = [Ch_{Mov}(x, y, t) - Ch_{\min}]/A$. The process of motion charge method is shown in Fig. 5.

In Table. 2, the contribution of each feature component is evaluated for Seq1 in order to construct the most suitable quaternion. We can see that the motion feature is the most

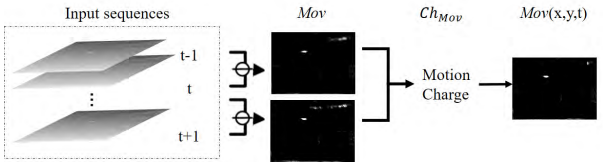


FIGURE 5. Flow chart of motion charge (MC) method.

TABLE 2. Compare the SCR and BSF of different feature.

Feature of input to QDCT	SCR	BSF
Motion Feature	3.1577	5.2714
Kurtosis Feature	1.9086	4.1195
Direction Feature	2.9106	4.7532
Space domain: Direction+ Kurtosis	6.4150	7.6924
Proposed: Motion + Kurtosis + two Directions	8.9514	9.0693

important feature. Therefore, we regard the motion feature as the real part of the quaternion. The final constructed quaternion can be expressed as:

$$q(x, y, t) = m(x, y, t) + f_1(x, y, t)i + f_2(x, y, t)j + f_3(x, y, t)k \quad (12)$$

where i, j, k satisfy $i^2 = j^2 = k^2 = -1, i \perp j, j \perp k, i \perp k, \perp$ is a vertical symbol.

In order to intuitively represent the situation of the quaternion constructed by four channels, the real or imaginary parts of the quaternion occupied by four components are shown in Fig. 6.

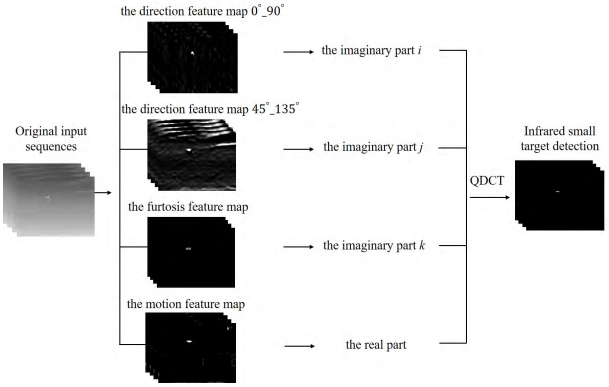


FIGURE 6. The four component display of the constructed quaternion.

D. INFRARED TARGET DETECTION BASED ON QDCT

We apply QDCT to the $q(x, y, t)$ to obtain the frequency feature map $Q(u, v, t)$. Given that resolution of the quaternion $q(x, y, t)$ is $M \times N$. Where M, N represents the length and width of the image, respectively. QDCT of the $q(x, y, t)$ is given by [24], expressed as:

$$\begin{aligned} QDCT(q(x, y, t)) &= Q(u, v, t) \\ &= \alpha_u^M \alpha_v^N \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mu_Q q(x, y, t) N(u, v, m, n) \end{aligned} \quad (13)$$

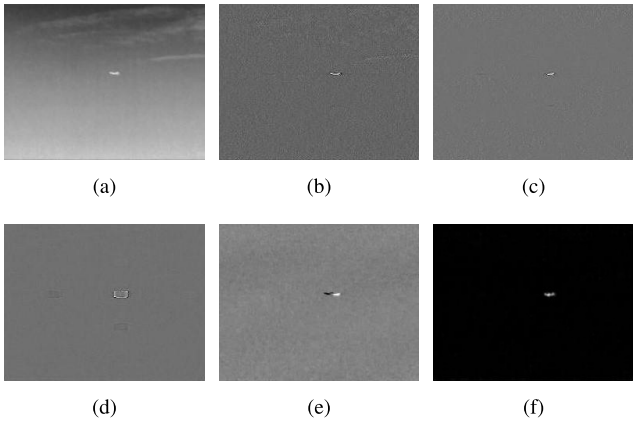


FIGURE 7. QDCT result of four channels and the final target detection result. (a) Original image (the 35th frame of Seq6); (b) 0^o_90^o; (c) 45^o_135^o; (d) Kurtosis feature map; (e) Motion feature map; (f) Detection result.

where $N(u, v, m, n) = \cos\left[\frac{\pi}{M}\left(m+\frac{1}{2}\right)u\right]\cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}\right)v\right]$ and α_u^M and α_v^N are the coefficients of QDCT, which can be taken as in [25]. μ_Q is the unit pure virtual quart, which is expressed as follows.

$$\mu_Q = -\sqrt{\frac{1}{3}}i - \sqrt{\frac{1}{3}}j - \sqrt{\frac{1}{3}}k \tag{14}$$

As long as the QDCT is obtained, we use the sign function to extract the salient area in transform domain since the sign function for quaternion corresponds with the visual response that how the visual system perceives the salient object in an image [24], [29] and it can remove the backgrounds clutters as well. The sign function $sgn(Q)$ of quaternion $Q = Q(u, v, t)$ is defined as:

$$Q' = sgn(Q) = \begin{cases} \frac{x_0}{|Q|} + \frac{x_1}{|Q|}i + \frac{x_2}{|Q|}j + \frac{x_3}{|Q|}k, & |Q| \neq 0 \\ 0, & |Q| = 0 \end{cases} \tag{15}$$

where x_0, x_1, x_2, x_3 are four components of Q, respectively. $|Q|$ is magnitude of the quaternion, $|Q| = \sqrt{Q \cdot \bar{Q}}$. $\bar{Q}$ is the complex conjugate of Q.

Finally, IQDCT is applied to transform Q' back to the spatial domain. In order to get a smoother result, we use (17) to obtain the final reconstructed target detection result.

$$\begin{aligned} q'(x, y, t) &= IQDCT(sgn(QDCT(q(x, y, t)))) = IQDCT(Q') \\ &= \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} \alpha_u^M \alpha_v^N \mu_q Q'(u, v, t) N(u, v, m, n) \end{aligned} \tag{16}$$

$$S(x, y, t) = g(x, y) * [q'(x, y, t) \odot \bar{q}'(x, y, t)] \tag{17}$$

where $g(x, y)$ is a Gauss smoothing filter with $\sigma = 1.5$. $\odot$ represents element wise product.

The final target result and the corresponding four components of Q' are shown in Fig. 7. Fig. 7(b) to Fig. 7(e) are four component feature maps of Q'. Because a single feature map is weak, it's necessary to fuse four feature maps together to promote the detection result, as it's shown in Fig. 7(f).

IV. EXPERIMENTS AND RESULTS

In this section, the performance and robustness of the proposed method are demonstrated by comparing it with eight state-of-the-art methods on nine relevant test sequences. All the experiments are implemented by Matlab 2016b on a PC with a 2.50 GHz Intel i5 CPU. The datasets and the code demo can be found on Github. [1]

A. EVALUATION INDICES

We select the signal-to-clutter ratio (SCR) and the background suppression factor (BSF) as evaluation indices. The SCR_{Gain} and the BSF are defined as:

$$SCR_{Gain} = \frac{(S/C)_{out}}{(S/C)_{in}} \tag{18}$$

$$BSF = \frac{C_{out}}{C_{in}} \tag{19}$$

where S and C are the average target intensity and clutter standard deviation, respectively. $(\cdot)_{in}$ represents the original image while $(\cdot)_{out}$ is the detection result. Additionally, C_{in} and C_{out} are the standard deviation of the backgrounds for original infrared image and detection result map, respectively. The SCR index measures the magnification of target relative to the backgrounds before and after processing. The BSF represents the suppression effect of backgrounds without any information about the target.

The receiver operating characteristic (ROC) curve quantitatively describes the varying relationship of the false positive rate (FPR) and true positive detection rate (TPR), which are defined as follows:

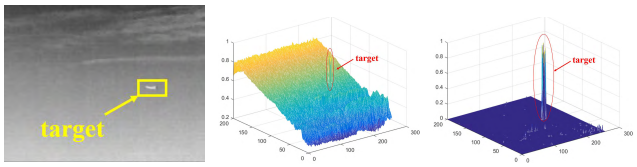
$$FPR = \frac{number\ of\ detected\ false\ targets}{total\ number\ of\ pixels\ in\ the\ whole\ image} \tag{20}$$

$$TPR = \frac{number\ of\ detected\ true\ targets}{total\ number\ of\ real\ targets} \tag{21}$$

B. RESULTS AND ANALYSIS IN SKY BACKGROUND

The detection of infrared small targets in complex backgrounds can be treated as a saliency detection problem, which needs to suppress complex backgrounds and enhance the target at the same time, and then separate the moving and prominent targets from the backgrounds to achieve the purpose of detection. Besides, the 3D surface graph can reflect the position of the target. Therefore, in order to verify the effectiveness of the proposed method, Fig. 8 shows the 3D surface graph before and after processing using this method. Comparing Fig. 8(b) with Fig. 8(c), we can see that the background of the original image is very complex, and the target is almost submerged in the serious clutters from Fig. 8(b). In Fig. 8(c), the background is obviously suppressed, and small targets are more prominent, which is conducive to locating the target accurately. It also proves that the proposed method is effective for small target detection in complex backgrounds.

[1] https://github.com/wxw211311/small-target.



(a) (b) (c)

FIGURE 8. Comparison of the 3D meshes before and after treatment. (a) Original image (the 39th frame of Seq6); (b) 3D mesh of original image; (c) 3D mesh of the detection result.

TABLE 3. Details of the test data set.

Sequences	Number	Size	Target Details	Background Details
Seq1	70	256 × 200	An aircraft with slender shape. Moving fast in the sequence.	Changing sky backgrounds with light noise.
Seq2	100	320 × 240	Extremely small moving target.	Backgrounds with huge and complex clutter.
Seq3	32	320 × 240	Weak target with irregular movement.	Blurred backgrounds with extremely low signal to clutter ratio.
Seq4	52	128 × 128	Extremely tiny moving target.	Sky scene with banded cloud.
Seq5	32	320 × 240	An aircraft with changing intensity.	Changing backgrounds with heavy cloud clutter.
Seq6	40	256 × 200	An aircraft with slender shape. Moving fast in the sequence.	Backgrounds with fixed clutter.

TABLE 4. Detailed parameter settings of the eight test methods.

No.	Methods	Abbreviations	parameter settings
1	Adaptive Butterworth high-pass filter	A-BHPF	Template size: 5 × 5.
2	Local contrast method	LCM	$interval = 2, Threshold : k = 3, 4, 5.$
3	Non-negativity-constrained variational mode	NVMD	Gaussian Filter: $\sigma_1 = 0.5, \sigma_2 = 2$; $\tau = 5$; Iterations: $N = 100$.
4	Derivative entropy contrast measure	DECM	Filter Template: 5 × 5. Adjustable coefficient $\lambda = 0.1$.
5	Stable multi-subspace learning	SMLS	Regular Factor: $\lambda = (1/(\min(m,n))^{0.5}) * 3$, Patch size: $\sqrt{p} \times \sqrt{q}$.
6	Frequency-Tuned	FT	Factor: $\beta = 0.3$; Kernel size: 5 × 5.
7	Spectral Residual	SR	Gaussian Filter: $\sigma = 8$.
8	Phase spectrum of quaternion fourier transform	PQFT	Gaussian Filter: $\sigma = 1.5$, Neighborhood window size: 5 × 5.
9	Proposed quaternion discrete cosine transform	Proposed	Gaussian Filter: $\sigma = 1.5$, $\tau = 1$, Patch size: 7 × 7.

In order to fairly evaluate the performance of infrared detection methods, the data sets consisted of six public infrared sequences with different complex backgrounds are used and the detailed features are listed in Table. 3. These databases are used in most relevant papers. These sequences have a long imaging distance, which leads to small targets, and they are all cloud or sky backgrounds with low signal-to-noise ratio. On the whole, the data sets contain various situations in airborne infrared target detection. Therefore, using the given databases could fairly show the performance of infrared detection methods.

The compared nine methods are the adaptive Butterworth high-pass filter (A-BHPF) [30], the LCM [9], the NVMD [32], the Derivative entropy contrast measure (DECM) [8], the stable multi-subspace learning (SMSL) [7], the Frequency-Tuned (FT) [33], the Spectral Residual (SR) [31], the PQFT [19] and the proposed method, respectively. The first five methods are in the spatial domain, while the last four are in the frequency domain. The detailed parameter settings used in the experiments are described in Table. 4 for reproduction. Experimental results of these methods with the index are shown in Table. 5. The highest value of each evaluation index in each column is marked red, and the second-highest one is marked blue.

From Table. 5, it can be seen that the proposed method achieves the highest values of SCR_{Gain} and BSF in test Seq1, Seq2, Seq4, and Seq6. The proposed method can improve SCR_{Gain} to some extent, and backgrounds edges are also enhanced. It means the detected small target is more prominent than the backgrounds while achieving a better

TABLE 5. Evaluation indices comparison of SCR_{Gain}, BSF.

Methods	Evaluation indices	Seq1	Seq2	Seq3	Seq4	Seq5	Seq6
A-BHPF	SCR_{Gain}	5.4441	6.3202	4.3424	8.2873	3.1847	4.9993
	BSF	5.7940	6.3302	10.0769	5.7776	5.4137	6.6330
LCM	SCR_{Gain}	6.8327	7.9143	6.4393	3.1285	7.5129	12.0083
	BSF	8.3709	6.5534	4.5440	3.4407	5.6598	8.2904
NVMD	SCR_{Gain}	8.7927	11.9417	28.8934	14.0527	4.8195	11.2148
	BSF	1.9432	5.3073	8.9912	9.9832	5.6990	4.8252
DECM	SCR_{Gain}	2.4107	3.6982	11.1416	4.1627	5.5311	1.9486
	BSF	8.9276	6.0803	12.7720	9.0029	8.5631	7.9391
SMSL	SCR_{Gain}	7.6500	12.7613	17.7629	13.9977	6.1374	9.7163
	BSF	6.6333	5.2354	8.0814	9.3976	5.4073	8.4853
FT	SCR_{Gain}	0.9417	2.6620	1.3629	2.1013	3.6705	1.7811
	BSF	1.1701	1.2675	1.4807	2.8051	1.4275	0.9381
SR	SCR_{Gain}	7.8808	9.3162	20.2515	7.4176	8.3057	11.1710
	BSF	2.4324	4.5654	9.6430	5.9103	5.1150	7.1150
PQFT	SCR_{Gain}	6.2069	7.7745	12.8211	9.7753	5.2869	9.6814
	BSF	6.6411	5.3644	11.0308	7.1316	6.0371	8.6902
Proposed	SCR_{Gain}	8.9514	14.5348	27.3039	14.4161	7.5972	15.4176
	BSF	9.0693	7.0875	13.6718	10.5222	10.7471	9.4569

suppression effect on the background. That is to say, the proposed method outperforms the compared methods in both target enhancement and background suppression from the angle of numerical indicator values for these test sequences. For test Seq3 and Seq5, the proposed method with the highest BSF and second-highest SCR_{Gain}, which is only slightly lower than the NVMD or SR method. Although the NVMD method has the highest SCR_{Gain} in test Seq3, its BSF is obviously lower than our method, which makes the detection result still noisy. For Seq5, SCR_{Gain} of the proposed method is slightly weaker than that of SR mainly because of its thick, heavy cloud, and even occlusion. In general, the method proposed in this paper can achieve the best detection effect, which is also true in the follow-up discussion of ROC. In particular, the effect of background suppression is the most prominent.

Furthermore, to give a more intuitionistic presentation of the processing results and to show the target/background enhancement ability visually, three-dimensional (3D) meshes of the original images and detection results are given in Fig. 9. Obviously, almost all of the components of the backgrounds are suppressed. Besides, there always exist backgrounds regions whose gray level is near or higher than targets as we can see from the original infrared images. This leads to the difficulty in detecting small targets from infrared sequences. By contrast, detection results of our method have relatively smooth backgrounds in 3D meshes. It is easier to determine the location and size of the small target from the detection results of our method, with fewer false alarms. In Fig. 9(d), the proposed method shows good robustness and effectiveness for different scenes and target pixels that approximately obey the highlighted varietal two-dimensional Gaussian model.

Note that in Fig. 10 the red circles denote the detected real targets, and the green circles denote some target-like false alarms. The first column refers to the original sequence frames. The #1 frame, #15 frame, #35 frame, #19 frame, #35 frame and #30 frame of Seq1 to Seq6 are selected to show the contrastive results, respectively. From the experimental results of Fig. 10, the proposed method has excellent performance on target enhancement and background suppression. The characteristics of each method are described according to the experimental results. According to Fig. 10(g), the FT method has poor performance on suppressing the complex

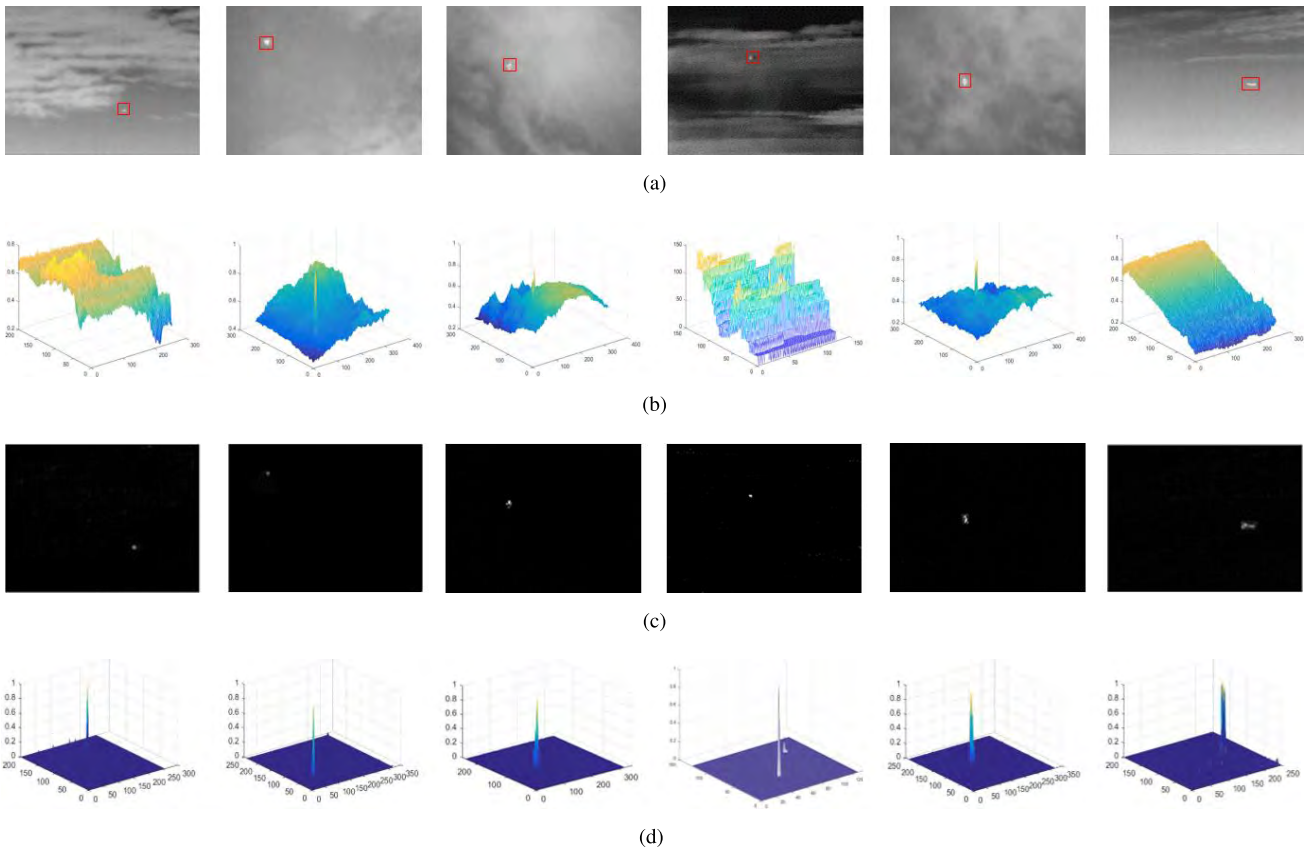


FIGURE 9. (a) Original images; (b) 3D meshes of original images; (c) Detection results; (d) 3D meshes of detection results.

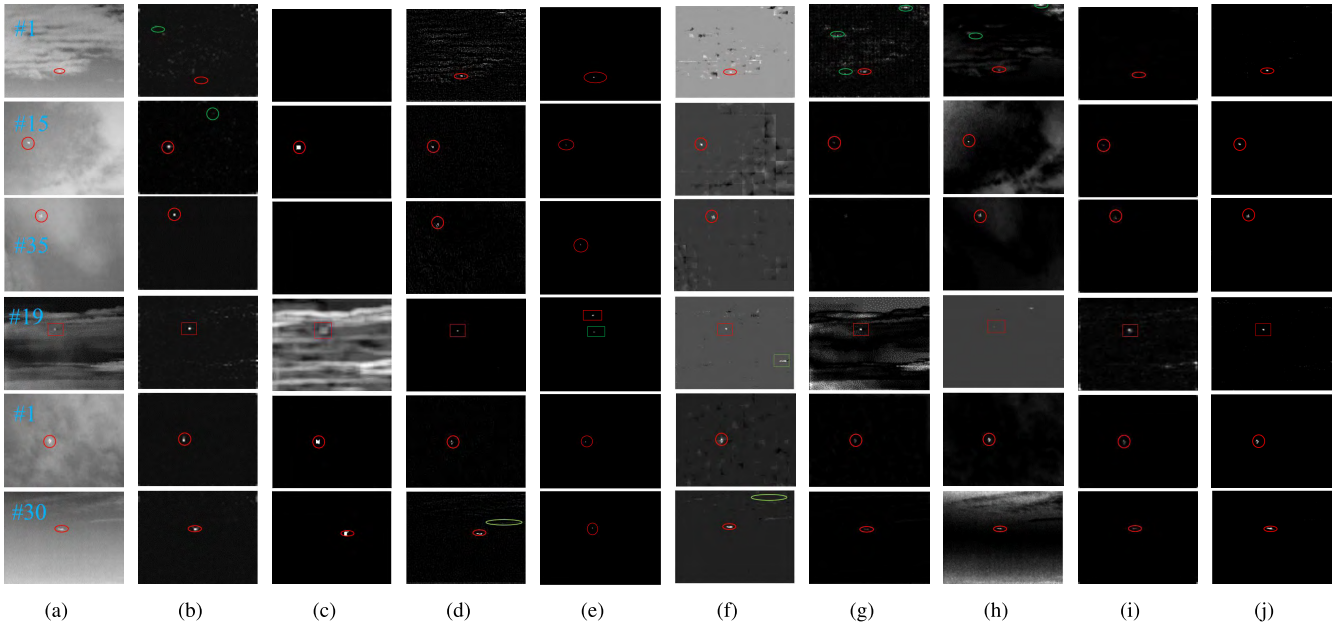


FIGURE 10. Subjective visual contrast, original images and small target detection results processed by different methods on Seq1 to Seq6. (b) to (j) denote the correspond detection results using A-BHPF, SR, LCM, NVMD, DECM, SMSL, FT, PQFT and Proposed method, respectively.

background, especially in Seq1 and there are still some background clutters. It used frequency difference to classify high-frequency clutter into small targets. Besides, A-BHPF and PQFT also can restrain the intensity of a small target to a certain extent, which means weakening the targets. From the subjective vision, SR method can enhance the targets

in general cases, but its effect of background suppression is poor. For example, there is still too much cloud interference and wrong targets. From Fig. 10(d) we can see that the result of LCM has good target enhancement ability in three cases. However, because the IR target detection results depend heavily on the appropriate threshold in LCM method, the adaptive threshold selected in this paper does not well detect the targets in Seq1, Seq3 and Seq4. Especially in Seq4 with a banded cloud, the blocking effect is too obvious and the detection effect is not good. According to Fig. 10(e), the NVMD method has a good performance on extracting small targets from a complex background in general cases. However, when the original sequences have many noises, a lot of noises still remains in Seq1 and Seq6. For SMSL, when the target was so dim that it was difficult to distinguish the target from the background clearly, the learning mechanism would assign the target to a subspace of the background, thus the target would disappear. And as can be seen from Fig. 10(g), the block effect of SMSL is obvious in some results due to the use of sliding window patching technology. Although the target is enhanced, the background is not well suppressed, and there are still some clutter and obvious block defects, which makes its *BSF* lower. According to Fig. 10(j), the dim small targets are extracted accurately from a complex background using QDCT to suppress the background and enhance the target because the proposed method uses four features to preserve more information. This indicates that the proposed method has excellent performance on enhancing target and suppressing background.

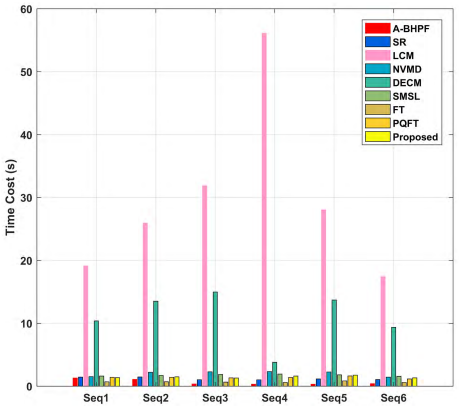


FIGURE 11. Time-consumption comparison of seven detection methods on six test image sequences.

We also make a comparison with respect to the processing time of all methods. The average time consumption comparison of nine methods is given in Fig. 11. Bars with the same color represent the average time required to process an image with a specific method. Although the LCM method has a relatively high detection accuracy, which is still lower than the proposed method for Seq1, Seq2, and Seq6, it is too expensive in time. The average running time of each frame in these test sequences is more than 15 seconds or even up to 50 seconds, which is obviously not suitable for the infrared small target detection for video sequences. We can also see that the proposed method is a little slower than the FT and A-BHPF methods, and faster than other methods. In short, the proposed method is better, whose detection effect makes up for subtle deficiency than the FT and A-BHPF methods.

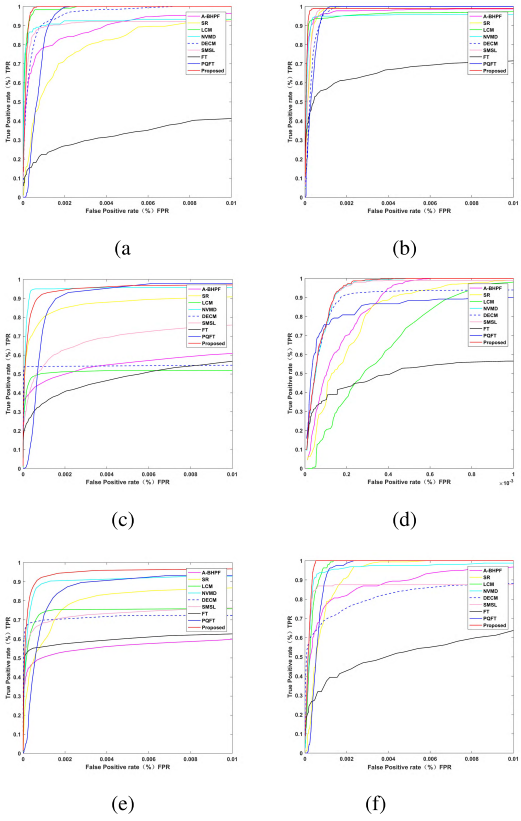


FIGURE 12. ROC curves of different methods. (a) to (f) correspond to the results of test sequences Seq1 to Seq6. The closer a curve is to the top-left corner, the better the corresponding method is.

To further demonstrate the advantages of the developed method, we provide the receiver operating characteristic (ROC) curves of the test sequences used in Table. 5, whose vertical axis was the false positive rate of the sequence and whose horizontal axis was the true positive rate of the sequence. The ROC curves are illustrated in Fig. 12. FT is almost always performed the worst since the simple frequency difference could not handle the complex background. The performance of SMSL fluctuated greatly. For Seq3, SMSL worked poorly but succeed in the other sequences, which is mainly because there is the wrong detection. LCM and DECM achieved similar results, which are unstable. This shows that they have low robustness in complex background. The biggest difference was in Seq3, which showed that the proposed is slightly lower than NVMD, probably caused by the weak target which can be detected mistakenly when the target is blocked. But the background suppression effect of the proposed method is better than NVMD. In general, the proposed method almost always reached the highest TPR with respect to the same FPR, implying that

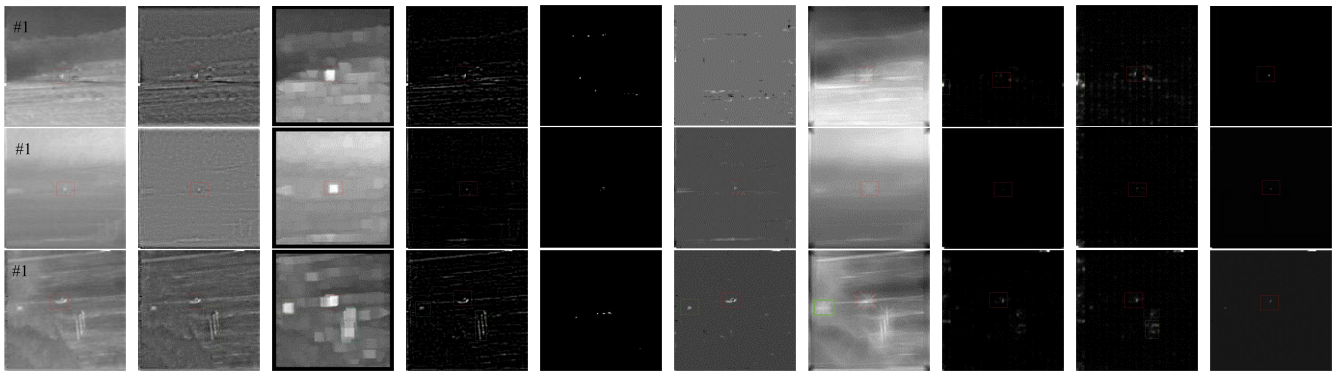


FIGURE 13. Subjective visual contrast, original images and small target detection results processed by different methods on L15NSS, L1415S, L19NSS. (b) to (j) denote the correspond detection results using A-BHPF, LCM, NVMD, DECM, SMSL, FT, SR, PQFT and Proposed method, respectively.

TABLE 6. Area under curve (AUC) values of the nine methods.

Methods	Seq1	Seq2	Seq3	Seq4	Seq5	Seq6
A-BHPF	0.9969	0.9943	0.9961	0.9978	0.9958	0.9976
LCM	0.9976	0.9952	0.9957	0.9971	0.9976	0.9984
NVMD	0.9973	0.9976	0.9984	0.9983	0.9979	0.9983
DECM	0.9974	0.9955	0.9968	0.9979	0.9973	0.9973
SMSL	0.9972	0.9981	0.9978	0.9982	0.9974	0.9974
FT	0.9935	0.9949	0.9959	0.9953	0.9961	0.9961
SR	0.9967	0.9980	0.9969	0.9974	0.9973	0.9976
PQFT	0.9970	0.9972	0.9978	0.9977	0.9976	0.9980
Proposed	0.9982	0.9984	0.9980	0.9983	0.9979	0.9986

the proposed method outperformed the other state-of-the-art methods. Besides, the AUC values are also listed in Table. 6. The AUC is the area enclosed by the ROC curve, which can also directly reflect the detection performance.

C. RESULTS AND ANALYSIS IN LAND BACKGROUND

In order to further prove the performance of the proposed algorithm, this section compares three small infrared target sequences in the land background. The eight-comparison method and the setting of their parameters are the same as the sky background in the previous section. The data sets can be downloaded in Github, which is provided by Harbin University of Technology [34] and the detailed features are listed in Table. 7. These sequences have a long imaging distance, which leads to small infrared targets and low signal-to-noise ratio. Therefore, using the given databases are appropriate to prove the performance of infrared detection methods.

Subjective analysis is made by comparing the results of nine methods, which are shown in Fig. 13. Three rows represent the detection results of three land background sequences (select representative frames). Each column represents a detection method, from the second to the tenth column represent the results obtained using A-BHPF, LCM, NVMD, DECM, SMSL, FT, SR, PQFT, and proposed method. As can be seen from Fig. 13, the LCM method produces an obvious block effect, and the target area detected is the same size as the selected image patch. A-BHPF, NVMD,DECM, and SMSL can detect targets, but there are still some clutters and background filtering is not clean. Although the results

TABLE 7. Details of the test data set.

Sequences	Number	Size	Target Details	Background Details
L15NSS	100	128 × 128	Small and dim dot infrared target. Moving fast in the sequence.	Changing land backgrounds with a lot of clutter.
L1415S	100	128 × 128	Extremely small moving target because of the far imaging distance.	Backgrounds with huge and complex clutter.
L19NSS	100	128 × 128	Weak target with regular movement.	Blurred backgrounds with extremely low signal to clutter ratio.

TABLE 8. Evaluation indices comparison of SCR_{Gain}, BSF and $Time(s)$.

Sequences	Evaluation indices	A-BHPF	LCM	NVMD	DECM	SMSL	FT	SR	PQFT	Proposed
L15NSS	SCR_{Gain}	3.4169	2.9143	5.4634	0.8615	3.4169	1.9860	2.7621	5.5525	6.1826
	BSF	3.8154	1.7836	3.9824	5.3300	3.2025	1.3740	4.2747	4.6934	5.4716
	$Time$	0.0680	6.3255	0.9486	2.9457	0.2211	0.0308	0.2398	0.1014	0.1265
L1415S	SCR_{Gain}	4.9587	4.2027	6.1791	3.5991	5.4654	1.5989	3.9094	5.9529	6.3144
	BSF	4.7538	2.7215	5.5189	5.1600	4.7538	1.5337	3.5266	5.5239	5.6277
	$Time$	0.0703	6.1719	0.8975	3.3356	0.1276	0.0255	0.2378	0.0741	0.1005
L19NSS	SCR_{Gain}	1.8996	2.8330	2.9461	1.3186	3.8330	1.0810	2.7293	2.9684	3.3472
	BSF	2.2934	1.4513	3.6271	3.3177	2.5708	1.5576	3.1367	3.2888	4.2532
	$Time$	0.0723	4.9555	0.9044	3.1127	0.2148	0.0202	0.2393	0.1025	0.0885

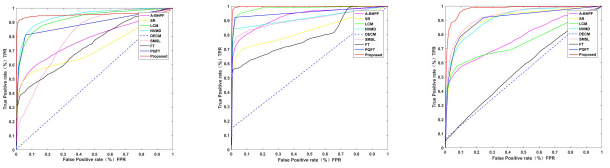
TABLE 9. AUC values of the nine methods for land background.

Sequences	A-BHPF	LCM	NVMD	DECM	SMSL	FT	SR	PQFT	Proposed
L15NSS	0.7460	0.9108	0.9177	0.5120	0.7696	0.7143	0.7108	0.8875	0.9308
L1415S	0.9331	0.9918	0.9216	0.6020	0.9076	0.6884	0.8508	0.9467	0.9995
L19NSS	0.8532	0.8308	0.9636	0.5810	0.9732	0.7943	0.9240	0.9344	0.9825

of PQFT method have better background suppression, there are many noise points. The proposed method may have the problem that the enhancement of the target is not obvious, the effect of background suppression is obviously the best. In addition, the other eight contrast methods are almost based on single frame detection because they do not select the motion feature, so they may also produce false detection and detect small false targets.

In order to further compare the detection performance of different methods, we still use SCR_{Gain} and BSF for quantitative and objective comparison in Table. 8. It can be seen that the proposed method achieves the highest values on both indices except for sequence L19NSS, which is slightly lower than SMSL. For this sequence, the infrared target is the larger, not a small target in the full sense. Besides, although the background suppression of SMSL method is not obvious, retains the contour of the small target, and is closer to the ground-truth of the target, thus making it has a larger SCR_{Gain}.

Similar to the experimental analysis in the sky background, the ROC curves of the nine methods for three real infrared small target sequences in Fig. 14 are given. As can be seen from Fig. 14, the method proposed in this paper achieves the



(a) (b) (c)

FIGURE 14. ROC curves of nine methods for three real infrared image sequences with land background. (a) L15NSS, (b) L1415S, (c) L19NSS.

best detection results. When the false alarm rate is the same, the proposed algorithm has the highest detection accuracy and the largest area under the curve. Besides, the AUC values are also listed in Table. 9.

V. CONCLUSION

Addressing the problems of target detection in a complex background, this paper introduces a novel QDCT-based method to realize the infrared small target detection. In the proposed method, we presented a novel approach to construct the quaternion, which considers two direction features, a motion feature, and a kurtosis feature. The sign function in QDCT serves as a saliency signature, which contributes to the solution of the small target detection problem. To demonstrate the performance of the proposed method in complex scenes, we evaluated the proposed approach by comparing with the most widely applied target detection algorithms on relevant data sets. Experimental results show that the proposed method can extract the small target in a complex background, and it is superior in the SCR_{Gain}, BSF, and ROC curve compared to the methods. In summary, experiments verify the effectiveness and efficiency of this approach. We were able to achieve higher results in highly complex backgrounds. At present, we are trying to accelerate the proposed method with GPU, hoping to be more widely used in practical applications, such as infrared and visible image fusion, saliency detection and so on.

ACKNOWLEDGMENT

The authors would like to thank our co-researchers for providing the image datasets. They would also grateful to our teachers for their great help in checking the grammar, and to the anonymous reviewers and the editor for their valuable comments and suggestions which are very helpful in improving the quality of the paper.

REFERENCES

[1] S. Kim and J. Lee, "Scale invariant small target detection by optimizing signal-to-clutter ratio in heterogeneous background for infrared search and track," *Pattern Recognit.*, vol. 45, no. 1, pp. 393–406, Jan. 2012.
[2] Y. Chen and Y. Xin, "An efficient infrared small target detection method based on visual contrast mechanism," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 7, pp. 962–966, Jul. 2016.
[3] S. Qi, D. Ming, J. Ma, X. Sun, and J. Tian, "Robust method for infrared small-target detection based on Boolean map visual theory," *Appl. Opt.*, vol. 53, no. 18, pp. 3929–3940, Jun. 2014.
[4] S. Qi, G. Xu, Z. Mou, D. Huang, and X. Zheng, "A fast-saliency method for real-time infrared small target detection," *Infr. Phys. Technol.*, vol. 77, pp. 440–450, Jul. 2016.
[5] G. Gu, K. Ren, M. Wan, Q. Chen, and W. Qian, "Robust infrared small target detection via non-negativity constraint-based sparse representation," *Appl. Opt.*, vol. 55, no. 27, pp. 7604–7612, Sep. 2016.
[6] S. Zhao, Y. Song, Y. Li, L. Li, and Q. Hao, "Infrared target detection method based on the receptive field and lateral inhibition of human visual system," *Appl. Opt.*, vol. 56, no. 30, pp. 8555–8563, Oct. 2017.
[7] X. Wang, Z. Peng, D. Kong, and Y. He, "Infrared dim and small target detection based on stable multisubspace learning in heterogeneous scene," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5481–5493, Oct. 2017.
[8] X. Bai and Y. Bi, "Derivative entropy-based contrast measure for infrared small-target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2452–2466, Apr. 2018.
[9] C. L. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, "A local contrast method for small infrared target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 574–581, Jan. 2014.
[10] Y. Shi, Y. Wei, H. Yao, D. Pan, and G. Xiao, "High-boost-based multiscale local contrast measure for infrared small target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 1, pp. 33–37, Jan. 2018.
[11] J. Han, K. Liang, B. Zhou, X. Zhu, J. Zhao, and L. Zhao, "Infrared small target detection utilizing the multiscale relative local contrast measure," *IEEE Geosci. Rempte Sens. Lett.*, vol. 15, no. 4, pp. 612–616, Apr. 2018.
[12] X. Bai, F. Zhou, and T. Jin, "Enhancement of dim small target through modified top-hat transformation under the condition of heavy clutter," *Signal Process.*, vol. 90, no. 5, pp. 1643–1654, May 2010.
[13] W. Wang, N. Yang, Y. Zhang, F. Wang, T. Cao, and P. Eklund, "A review of road extraction from remote sensing images," *J. Traffic Transp. Eng.*, vol. 3, no. 3, pp. 271–282, 2016.
[14] K. E. Matthews and N. M. Namazi, "A Bayes decision test for detecting uncovered-background and moving pixels in image sequences," *IEEE Trans. Image Process.*, vol. 7, no. 5, pp. 720–728, May 1998.
[15] J. Han, Y. Ma, B. Zhou, F. Fan, K. Liang, and Y. Fang, "A robust infrared small target detection algorithm based on human visual system," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 12, pp. 2168–2172, Dec. 2014.
[16] J. Han, Y. Ma, J. Huang, X. Mei, and J. Ma, "An infrared small target detecting algorithm based on human visual system," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 452–456, Mar. 2016.
[17] N. Imamoglu, W. Lin, and Y. Fang, "A saliency detection model using low-level features based on wavelet transform," *IEEE Trans. Multimedia*, vol. 15, no. 1, pp. 96–105, Jan. 2013.
[18] S. Qi, J. Ma, C. Tao, C. Yang, and J. Tian, "A robust directional saliency-based method for infrared small-target detection under various complex backgrounds," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 495–499, May 2013.
[19] S. Qi, J. Ma, H. Li, S. Zhang, and J. Tian, "Infrared small target enhancement via phase spectrum of quaternion Fourier transform," *Infr. Phys. Technol.*, vol. 62, pp. 50–58, Jan. 2013.
[20] C. Yang, J. Ma, M. Zhang, S. Zheng, and X. Tian, "Multiscale facet model for infrared small target detection," *Infr. Phys. Technol.*, vol. 67, pp. 202–209, Nov. 2014.
[21] C. Yang, J. Ma, S. Qi, J. Tian, S. Zheng, and X. Tian, "Directional support value of Gaussian transformation for infrared small target detection," *Appl. Opt.*, vol. 54, no. 9, pp. 2255–2265, Mar. 2015.
[22] H. Liu, J. Gu, M. Meng, and W. Lu, "Fast weighted total variation regularization algorithm for blur identification and image restoration," *IEEE Access.*, vol. 4, pp. 6792–6801, 2016.
[23] W. Feng and B. Hu, "Quaternion discrete cosine transform and its application in color template matching," in *Proc. Congr. Image Signal Process.*, May 2008, pp. 252–256.
[24] B. Schauerte and R. Stiefelhagen, "Predicting human gaze using quaternion DCT image signature saliency and face detection," in *Proc. IEEE Workshop Appl. Comput. Vis. (WACV)*, Jan. 2012, pp. 137–144.
[25] H. Zhang, S. Gao, and J. Xue, "Flame detection via quaternion discrete cosine transform based spectral saliency," *Int. J. Control Automat.*, vol. 9, no. 9, pp. 343–352, 2016.
[26] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 9, pp. 891–906, Sep. 1991.
[27] Y. Liu, L. Mejias, and Z. Li, "Fast power line detection and localization using steerable filter for active Uav guidance," in *Proc. 12th Int. Soc. Photogramm. Remote Sens.*, 2012, pp. 491–496.
[28] T. López *et al.*, "Motion features to enhance scene segmentation in active visual attention," *Pattern Recognit. Lett.*, vol. 27, no. 5, pp. 469–478, 2006.

[29] H. Guo, X. Wu, S. Cai, N. Li, J. Cheng, and Y.-L. Chen, "Quaternion discrete cosine transformation signature analysis in crowd scenes for abnormal event detection," *Neurocomputing*, vol. 204, pp. 106–115, Sep. 2016.
[30] L. Yang, J. Yang, and K. Yang, "Adaptive detection for infrared small target under sea-sky complex background," *Electron. Lett.*, vol. 40, no. 17, pp. 1083–1085, Aug. 2004.
[31] X. Hou and L. Zhang, "A spectral residual approach," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Aug. 2007, pp. 1–8.
[32] X. Wang, Z. Peng, P. Zhang, and Y. He, "Infrared small target detection via nonnegativity-constrained variational mode decomposition," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1700–1704, Oct. 2017.
[33] R. Achanta, S. Hemami, F. Estrada, and S. Süsstrunk, "Frequency-tuned salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*, Jun. 2009, pp. 1597–1604.
[34] Z. Shi, C. Wei, and J. Li, "Parallel search strategy in kernel feature space to track FLIR target," *Neurocomputing*, vol. 214, pp. 671–683, Nov. 2016.

PING ZHANG was born in Sichuan, in 1976. She received the bachelor's degree in electronics and information processing from Lanzhou University, in 1999, and the master's degree in signal and information processing from the University of Electronic Science and Technology of China (UESTC), in 2004, where she graduated in signal and information processing, in 2011.

In 2004, she stayed as a Teacher. From 2014 to 2015, she was a Visiting Scholar with the University of California at San Diego, San Diego. She is currently a Professor with the School of Optoelectronic Science and Engineering, UESTC. In recent years, she has hosted and participated in more than ten projects of the National Natural Science Foundation of China, the Doctoral Fund of the Ministry of Education, the Postdoctoral Fund, the Sichuan Provincial Project, the Guangdong Science and Technology Project, and related enterprises and institutions. She has published more than 20 academic papers and has been searched for more than ten articles by SCI/EI. Her current research interests include image and video signal processing, target detection and tracking, computer vision, and intelligent information processing.

XIAOWEI WANG was born in Shanxi, in 1995. She received the bachelor's degree in information display from the University of Electronic Science and Technology of China, in 2017. She is currently pursuing the Degree in signal and information processing with the School of Optoelectronic Science and Engineering, University of Electronic Science and Technology of China.

As a Principal Researcher, she has participated in two scientific and technological support projects in Sichuan, Infrared and Visible Light Image Fusion Based on Visual Saliency and Key Technology Research on Low-Altitude Small Target Detection and Tracking Based on Visual Fusion. She has applied for a patent. Her research interests include high-speed image processing algorithms in 3D additive manufacturing, based on deep learning target detection and tracking research.

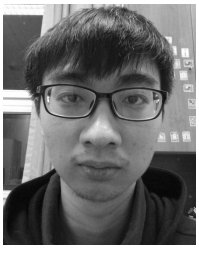
XIAOYANG WANG was born in Shanxi, in 1992. She received the bachelor's degree from the University of Electronic Science and Technology, in 2009, and the Ph.D. degree from UESTC, in 2018. After graduation, she studied at the University of Bristol as a Postdoctoral Fellow. She is currently in signal and information processing.

During her Ph.D., she has hosted and participated in more than six projects, including the National Natural Science Foundation, the Central University Research Fund, and the Chinese Academy of Sciences Beam Control Laboratory Fund. She has published more than ten academic papers, and has been searched by SCI/EI for more than seven articles and more than seven invention patents. From 2017 to 2018, she received the joint training from the University of Bristol, U.K., funded by the National Scholarship Fund. Her current research interests include trajectory prediction, target detection based on deep learning, infrared weak and small target detection, recognition and tracking, compressed sensing applications, deep learning, and machine learning.

She received school and business scholarships several times.

CHUN FEI was born in Sichuan. He received the master's degree in computer application from the University of Electronic Science and Technology of China, in 2005, and the Ph.D. degree in information and communication engineering from the School of Information and Communication Engineering, in 2015. He is currently a Professor with the School of Computer Science and Engineering, University of Electronic Science and Technology of China, where he is involved in relevant administrative management.

In recent years, he has hosted and participated in more than ten national natural science funds, postdoctoral funds, Sichuan Provincial Planning Projects, and Ph.D. funds of the Ministry of Education. He has published more than 20 academic papers and has been searched for more than ten articles by SCI/EI. His current research interests include deep learning, intelligent optimization, and image processing direction.



ZHENGKUI GUO was born in Anhui. He received the B.S. degree in optoelectronic information science and engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2018, where he is currently a Research Assistant with the School of Optoelectronic Science and Engineering. His research interests include visual object tracking and object detection.