Fast Infrared Small Target Detection by Using Hadamard Product for Spatial-Temporal Matrices

PING-HSIU WU1, YU-PIN LAN1, CHAO-LUNG CHOU2, AND CHUN-TING LIN3

1Institute of Lighting and Energy Photonic, National Yang Ming Chiao Tung University, Tainan City, Guiren, Tainan 711, Taiwan
2Department of Computer Science and Information Engineering, Chung Cheng Institute of Technology, National Defense University, Taoyuan 33551, Taiwan
3Institute of Photonics System, National Yang Ming Chiao Tung University, Tainan City, Guiren, Tainan 711, Taiwan

Corresponding authors: Chao-Lung Chou (chaolung.chou@gmail.com) and Chun-Ting Lin (jinting@nycu.edu.tw)

This work was supported by the Ministry of Science and Technology (MOST), R.O.C., under Grant 109-2221-E-009-154-MY3, Grant 110-2221-E-A49-085-MY2, and Grant 111-2221-E-606-013.

ABSTRACT

Infrared (IR) small target detection has recently been widely used in civil and military applications. In IR small target detection, IR images probed at long distances are easily disturbed by complex backgrounds, light changes, and other noises, which makes it pretty challenging. Most existing literature focuses on low false alarms and high detection accuracy rather than computational efficiency. We propose a fast and reliable IR small target detection method that deals with spatial and temporal domains based on the Hadamard product for spatial-temporal matrices (HPSTM). In the spatial domain, a weighted tri-layer window is used to convolute the IR image and obtain the gradient matrices. In the temporal domain, three interval frames were used instead of conventional adjacent frames and contrast matrices were extracted. Finally, the two spatial-temporal matrices use the Hadamard product to multiply and a simple threshold to transfer into a binary image for target detection. We also proposed the optimal mask size selection (OMSS) method, which can adjust the optimal tri-layer window size in the spatial domain to obtain the best detection result. The experimental results show that the proposed HPSTM method has a high detection accuracy and fast running time compared with other methods, indicating that the proposed HPSTM method is reliable and suitable for real-time IR small target detection.

INDEX TERMS

Infrared small target detection, spatial and temporal domains, Hadamard product, weighted tri-layer window, real-time.

I. INTRODUCTION

Infrared (IR) small target detection is a critical technology in the infrared search and track (IRST) system and has received much attention in recent years. It is widely used in civil and military applications, such as weather forecasts, remote sensing, forest fires detection, meteorological analysis, sea surface search and rescue, night navigation, surveillance, and long-distance discovery [1], [2], [3].

Because IR images are usually acquired by a long distance in an IRST system, the size of the small target in the image is usually small, approximately $2 \times 2 - 9 \times 9$ [4], which means that their detailed shape and texture features are difficult to obtain. Also, IR images frequently suffer from different types of interference, such as atmospheric absorption, complex background, edges, and brightness changes, making IR small target detection quite challenging [5].

In the past decade, many IR small target detection methods have been developed to overwhelm the challenging task under such low signal-to-clutter ratio (SCR) and complex conditions by using different approaches. Common IR small target detection algorithms can be mainly divided into three categories: spatial-domain based, temporal-domain based, and deep-learning based.

Spatial-domain based methods are also called single-frame based methods. Spatial-domain based methods highlight the target, separate the background, and obtain the actual target location. Deshpande et al. [6] proposed max-mean and max-median filters to estimate the background. Cao et al. [7] suggested using the two-dimensional least mean square (TDLMS) filter based on neighborhood information to achieve more accurate background estimation.

The associate editor coordinating the review of this manuscript and approving it for publication was Mehul S. Raval 116830.
and isolation. Lv et al. [8] combine the TDLMS algorithm and neighborhood gray-scale difference to improve the performance of target enhancement and background suppression in a natural space-based environment. Chen et al. [9] proposed the local contrast measure (LCM) algorithm. Han et al. [10] developed a multiscale tri-layer window local contrast measure (TLLCM) to estimate the background. Wu et al. [11] proposed the double neighborhood gradient method (DNGM), which is characterized by detecting targets of different sizes under the fixed size of the tri-layer window while avoiding the expansion effect and reducing the amount of computation. Han et al. [12] developed a weighted strengthened local contrast measure (WSLCM) which includes a strengthened local contrast measure (SLCM) and weighting function to enhance the target and estimate the background.

Temporal-domain based methods focus on detecting moving targets in a sequence of infrared images, also called multi-frame based methods. Gao et al. [13] proposed the temporal variance and spatial contrast filter (TVPCF) based on the fact that the background between adjacent images can be regarded as stationary, and the position of small targets changes significantly between adjacent images to improve the contrast of the target and suppress the background. Deng et al. [14] proposed a spatio-temporal local contrast filter (STLCF) by calculating the difference between the target and the surrounding background grayscale value in each frame to enhance the target. Both TVPCF and STLCF extract the local contrast of moving objects by calculating the intensity difference between the average intensity of one frame and its adjacent frames, so they require much computing time. Subsequently, Zhao et al. [15] developed the spatial-temporal local contrast map (STLCM) using a mean filter to predict the predicted values of eight directions, which operates in different directions and extracts the target from the background to obtain the grayscale difference of the spatial domain. STLCM first proposed a method using interval frames instead of the traditional adjacent frame method, effectively improving computing performance in the temporal domain. Du et al. [16] proposed the spatial-temporal local difference measure (STLDM) based on the HVS model to identify IR small targets using the local difference between the center of the moving target and the background in the spatial domain. Pang et al. [17] proposed a novel spatiotemporal saliency method (NSTSM) based on a tri-layer window and through the internal layer and external layers of the tri-layer window to model the target pixels and the background pixels’ changes and the variances. NSTSM is one of the few algorithms that detect low-altitude small target. The experimental results also prove that NSTSM has excellent detection performance for detecting low-altitude small targets with tri-layer windows.

Deep-learning methods based on learning models and large amounts of data also has received particular attention recently. Hou et al. [18] proposed a robust infrared small target detection network (RISTDNet). Dai et al. [19] proposed the attentional local contrast networks (ALCNet) to modularize traditional LCM into an end-to-end network of depth-wise, less nonlinear feature refinement layers. However, the final detection performance of existing methods is sensitive to the setting of the module’s hyper-parameters. Yu et al. [20] proposed a novel multiscale local contrast learning network (MLCL-Net). Which is an end-to-end fully convolutional infrared small target detection network. Note that deep learning-based methods require many samples for training and testing. A significant issue for applying deep learning-based techniques is that actual IR small target scenes are still scarce to the public. Therefore, most IR small target detection methods are still mainly based on spatial-domain and temporal-domain.

In general, the spatial-domain-based algorithms have lower computation complexity and are easy to implement in hardware, so it is applicable for real-time applications. Their drawback is limitations for overall performance. The temporal-domain-based methods are usually an extension of spatial-domain-based algorithms and are good at dealing with IR images with low SCR and complex backgrounds. Their drawback is that higher computation costs make them unsuitable for real-time applications. Therefore, this paper proposes a fast IR small target detection method that combines the advantages of both spatial and temporal domains, which maintains good detection performance and needs less computing costs. In spatial domain we used a weighted tri-layer window and convolute with the original image in the spatial domain to obtain gradient matrices. The matrices of contrast between the maximum and minimum grayscale values in the forward and backward frames are calculated in the temporal domain which is directly follow the temporal domain calculation method of [17]. The target is extracted according to the Hadamard product for the gradient and contrast matrices.

The main contributions of this paper are as follows:

1) We propose the optimal mask size selection (OMSS) and discuss how to adjust the optimal weighted tri-layer window size to achieve the best detection results.

2) The proposed HPSTM method significantly reduces the computational complexity by using the Hadamard product of the spatial-temporal matrices, which has advantages for real-time applications.

3) Extensive experiments and comparisons on five state-of-the-art spatial-temporal benchmarks demonstrate that the proposed HPSTM method achieves the best performance in terms of accuracy, computational cost, target signal enhancement, and background noise suppression.

The rest of this paper is organized as follows. The second section introduces the related background. The third section explains the detail of the proposed method. Experimental results and analysis are demonstrated in section four. Section five gives the conclusion.

II. BACKGROUND

An IR small target image can be modeled in three parts, namely the target, the background, and the noise, as shown in (1) [21], where \((x, y)\) are the coordinates of each pixel.
in the image $I$, $I_B$ is the background, $I_T$ is the target, and $I_N$ is the noise. IR small target detection aims to separate the target from the background by enhancing the target as much as possible and suppressing background noise.

$$I(x, y) = I_B(x, y) + I_T(x, y) + I_N(x, y)$$

The small target is usually located within a complex environment facing different light changes or climate variations, making it hard to separate the target and backgrounds [22]. We conclude three typical cases may encounter for discussion and illustrated by real scenes, as shown in Fig. 1.

1) The target stays within a smooth background, as Fig. 1(a) shows. The image intensity difference between the target and background is noticeable and separating them using contrast is the most straightforward way of target detection. Fig. 1(d) shows the 3D distributions of Fig. 1(a); it can be seen that the peak of the target is more evident than the surrounding.

2) The target stays within a complex background, as Fig. 1(b) shows. The image intensity in the complex background changes significantly, and the target signal hiding in the background may be misjudgment as noise and suppressed. Fig. 1(e) shows the 3D distributions of Fig. 1(b), the contrast between the target and surrounding background are not obvious, the target resulting in a slightly higher peak than the high-brightness background. This situation makes detection difficulties for many algorithms.

3) The target stays at the edge of the image, as Fig. 1(c) shows. Edge is discontinuities in image intensity and is commonly used to help detect objects’ boundaries. While the target stays at the edge, its intensity is similar to the background and hard to distinguish based on its intensity difference.

Fig. 1(f) shows the 3D distributions of Fig. 1(c), the changes in the central and surrounding areas are not noticeable, and the distribution is relatively flat.

According to the above observation, detecting small targets at the edge or within a complex background is more challenging than a smooth background. The proposed method aims to improve the efficiency of IR small target detection in such difficult situations.

A. CONSTRUCTION OF PROPOSED TRI-LAYER WINDOW

Many recent works have proposed the tri-layer window filter to enhance the target and suppress the background noise around the target [10], [11], [12]. The three layers consist of the inner layer, the guard layer, and the outer layer, as shown in Fig. 2 (a).

The inner layer is mainly used to detect and enhance the target, and the window size is primarily to match the small target size as much as possible. The small target is indeed close to the Gaussian distribution [17], so as long as we can detect the strongest energy of the target by the inner layer.
The outer layer is used to distinguish the real target and the surrounding background, so the size cannot be smaller than the size of the actual target; otherwise, the outer part of the target pixels will be mistaken for background pixels and decrease the detection accuracy. The guard layer is used to separate the inner and outer layers and is usually spaced by one pixel.

To reduce the computational cost, [23] proposed a gradient kernel model using convolution with a weighted tri-layer window. It is necessary to compute derivatives in the horizontal and vertical directions for detecting image intensity changes. The gradient of an image $\nabla f(x, y)$ can be calculated by its directional derivatives given in (2).

$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f(x, y)}{\partial x} \\ \frac{\partial f(x, y)}{\partial y} \end{bmatrix}$$  \hspace{1cm} (2)

where $\frac{\partial f(x, y)}{\partial x}$ and $\frac{\partial f(x, y)}{\partial y}$ are the derivatives in the x-axis and y-axis respectively. The sum of the second-order derivative in the x and y directions for detecting image intensity changes is approximated as a Gaussian distribution [17]. The target’s grayscale distribution will differ from backgrounds because the target position will change over time while the background will not. It can be found that when the target is moving, the gray value of the surrounding area will also vary significantly. Fig. 3 shows an example of moving target position changing frame by frame.

The conventional temporal domain method assumes that the background can be regarded as still in consecutive frames [24]. The moving target, which usually changes significantly, can be extracted by subtracting the background in consecutive frames [13], [14]. Recently, [17] suggested a temporal gray saliency map (TGSM) of three frames extracted from the image sequence for target detection. The TGSM(i, j) is defined as (5).

$$\text{TGSM}(i, j) = (I_{\text{max}} - I_{\text{min}})^2$$  \hspace{1cm} (5)

where $I_{\text{max}}$ and $I_{\text{min}}$ are the largest and smallest grayscale pixel values in each of the three frames derived from (6).

$$\begin{cases} I_{\text{max}} = \max \left\{ I_{T}^{k-n}(i, j), I_{T}^{k}(i, j), I_{T}^{k+n}(i, j) \right\} \\ I_{\text{min}} = \min \left\{ I_{T}^{k-n}(i, j), I_{T}^{k}(i, j), I_{T}^{k+n}(i, j) \right\} \end{cases}$$  \hspace{1cm} (6)

where $I_{T}^{k-n}, I_{T}^{k},$ and $I_{T}^{k+n}$ are the grayscale values of the $(k-n)\text{th}, (n)\text{th},$ and $(k+n)\text{th}$ frames respectively.

The advantage of TGSM is that calculating consecutive frames in the temporal domain will reduce the computation cost to only three frames. The proposed method utilizes this approach for frames extraction in the temporal domain.

C. THE HADAMARD PRODUCT

General matrix multiplication requires the number of columns in the first matrix to equal the number of rows in the second matrix. The resulting matrix is known as the matrix product. The Hadamard product (also known as the Schur product) [25] is distinguished from the standard matrix product, which takes two matrices of the exact dimensions and produces another matrix of the same size. If two matrices, A and B, have the same dimension $r \times c$, then their Hadamard product $A \circ B$ is a matrix with the same size, and its element values are given in (7).

$$[A \circ B]_{ij} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \circ \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & a_{13}b_{13} \\ a_{21}b_{21} & a_{22}b_{22} & a_{23}b_{23} \\ a_{31}b_{31} & a_{32}b_{32} & a_{33}b_{33} \end{bmatrix}$$  \hspace{1cm} (7)

In the proposed method, the matrices processed in the spatial and temporal domains have the same size. Therefore,
we take advantage of the properties of the Hadamard product to merge the spatial gradient and temporal contrast matrices without additional normalization or calibration. If two matrices have the same dimension $N \times N$, in general matrix multiplication would require $O(N^2)$ floating point operations. For Hadamard matrix, this computational cost can be reduced to $O(N \log_2 N)$ which significantly reduce the computational complexity [26].

III. THE PROPOSED METHOD
We proposed a fast IR small target detection method using the Hadamard product for spatial-temporal matrices. The input is a continuous IR image processed separately in the spatial and temporal domains and then fused into a binary image for target detection. In the spatial domain, a weighted tri-layer window is used to enhance the small target, suppress the background clutter, and obtain the gradient matrices using convolution. In the temporal domain, three interval frames are used instead of conventional adjacent frames and extract the contrast matrices. To reduce the computational cost, the map dimensions in the spatial and temporal domains are the same. A binary image is obtained according to the Hadamard product of spatial-temporal matrices (HPSTM). After that, small targets are extracted through an adaptive threshold, and the decomposed possible target positions are aggregated into a final position of the detected target through the target aggregation process. Fig. 4 shows the flowchart of the proposed method, and the detailed steps will be described below.

A. SPATIAL GRADIENT MATRIX ($M_s$)
We defined the gradient mask $M_G(s, t)$ based on the weighted tri-layer windows described in Section II.A. Here $s$ and $t$ represent the row and column of the gradient mask $M_G(s, t)$, respectively, in which $-m \leq s \leq m$, and $-n \leq t \leq n$ are each element within the gradient mask, $m$ and $n$ are positive integers.

The spatial gradient matrix $M_s(r, c)$ can be obtained by convolute the original infrared image $F(r, c)$ and gradient mask $M_G(s, t)$, as shown in (8).

$$M_s(r, c) = \sum_{s=-m}^{m} \sum_{t=-m}^{m} M_G(s, t) * F(r - s, c - t)$$

where $(r, c)$ and $(s, t)$ are the sizes of the original image and gradient mask, * denotes convolution operation.

Schematic diagrams of convolution are shown in Fig. 5. The gradient mask starts with the pixels in the top left corner of the image, then slides from left to right and from top to bottom. If the mask is out of the image range, the edges of the image will fill by 0. The calculated value is either positive or negative, which indicates the degree of change in the grayscale value. To reduce the amount of computation, we take the absolute value of $M_s$ to facilitate subsequent calculations. The larger the value of the gradient matrix, the larger the sum of the original image’s gradients in that pixel and the surrounding pixels, which means there is an excellent chance that the target will be the location. The range and weight of the surrounding pixels are determined by the size and coefficient of the gradient mask, respectively.

B. TEMPORAL CONTRAST MATRIX ($M_T$)
As mentioned before, to pursue fast computational time and high detection accuracy, we use the same temporal calculation method as the literature [17]. Sometimes because small targets move slowly, the calculation of adjacent frames results in a great chance that the target will be considered as background and suppressed, such as the problem encountered by STLCM [15]. Therefore, we choose the number of interval frames $n = 5$; even if the target moves slowly, we can still clearly get the comparison of the forward and backward frames. We use three frames with a fixed number of frames $n$ spaced apart from each other to process at a time, and we use (6) to get the largest and smallest grayscale pixel values $I_{max}$ and $I_{min}$. Subtracting $I_{min}$ from $I_{max}$ will get the maximum contrast of the grayscale values. We square the result to obtain the temporal contrast matrix $M_T(r, c)$ to make the contrast constant positive and maximize the contrast to highlight the target’s position, as (9) shows.

$$M_T(r, c) = (I_{max} - I_{min})^2$$

where $(r, c)$ are the size of the original image and the temporal contrast matrix $M_T$.

C. HADAMARD PRODUCT FOR SPATIAL-TEMPORAL MATRICES($M_s$) AND TARGET AGGREGATION
To achieve the highest detection rate and the best detection result, we combine the calculation results of the spatial domain and the temporal domain. We use the Hadamard product [25] to multiply the spatial gradient matrix $M_s(r, c)$ by the temporal contrast matrix $M_T(r, c)$ and obtain the matrix of final result $M_F(r, c)$ as shown in (10).

$$M_{Fr,c} = [M_s \circ M_T]_{rc}$$

where $\circ$ is the Hadamard product operator and $1 \leq i \leq r$, $1 \leq j \leq c$, in which $i, j$ denotes the position of the matrices and $r, c$ represents the row and column of all three matrices with the same size $r \times c$.

Below is an example of the Hadamard product of the spatial gradient matrix $M_s(r, c)$ and the temporal contrast matrix $M_T(r, c)$.

$$M_{Fr,c} = [M_s \circ M_T]_{rc} = \begin{bmatrix} 61 & \cdots & 27 \\ \vdots & \ddots & \vdots \\ 196 & \cdots & 2 \\ \vdots & \ddots & \vdots \\ 2352 & \cdots & 78 \end{bmatrix} = \begin{bmatrix} 28 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 12 & \cdots & 39 \end{bmatrix}$$

Note that in $M_s(r, c)$, the more significant value represents a higher gradient, implying the location is more likely to be a small target. The values in $M_T(r, c)$ and $M_F(r, c)$ are always positive.
After obtaining the $M_F$, the target detection is conducted using a threshold $T_h$, as shown in (11), and converting the original IR image into a binary image.

$$T_h = \mu_{M_F} + k \times \sigma_{M_F}$$  \hspace{1cm} (11)

where $k$ is an adjustable constant, $\mu_{M_F}$ represents the mean and $\sigma_{M_F}$ represents the standard deviation of $M_F$, as shown in (12) and (13).

$$\sigma_{M_F} = \sqrt{\frac{1}{r \times c} \sum_{i=1}^{r \times c} (x_i - \mu_{M_F})^2}$$  \hspace{1cm} (12)

$$\mu_{M_F} = \frac{1}{r \times c} \sum_{i=1}^{r \times c} x_i$$  \hspace{1cm} (13)

where $x_i$ denotes the value of each gradient mask.

The advantage of the adaptive threshold is that under the same $k$ value, different threshold values can be obtained according to different scenarios. If $k$ is too large so that $T_h$ also becomes more significant, pixels smaller than the threshold will be suppressed, resulting in a low detection rate; $k$ too small will make $T_h$ also become smaller, more significant than the threshold of too many pixels, increasing in false alarm rate. Based on our experiments, the parameter $k$ is usually set between 12 and 13.

The original IR image is converted into the binary image $I_B$ based on whether $M_F$ is greater than $T_h$ or not, and the non-zero pixel is the small target, as (14) shows.

$$I_B(r, c) = \begin{cases} 1, & \text{if } M_F(r, c) \geq T_h \\ 0, & \text{otherwise} \end{cases}$$  \hspace{1cm} (14)

In the proposed method, if the target size is greater than that of the inner layer of the tri-layer window, only edges will be detected, resulting target’s shape being decomposed. We used the 8-connected components labeling to solve this problem, as Fig. 6 shows. The connected area can be expanded, and the surrounding decomposed parts can be aggregated into a complete one. Algorithm 1 shows the procedure of target aggregation, in which input is the binary image $I_B$; output is the final image $I_F$; $(r, c)$ is the coordinates of the detected target; $R$ and $C$ are the dimensions of the rows and columns of the original image.

Fig. 7 shows the illustration of the proposed method. Fig. 7(a) is the original IR image. Fig. 7(b) is the binary image after using the Hadamard product for the spatial- temporal matrices. Fig. 7(c) is the final detection result after target aggregation using Algorithm 1.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. EXPERIMENTAL SETUP

In order to objectively evaluate the proposed method, the experiments use six IR small target sequence images with
Algorithm 1 Aggregate Decomposed Target

Input: Binary image $I_b(r,c)$.  
Output: Final Image $I_F(r,c)$. 

1: for $r = 1 : R$ do  
2:     for $c = 1 : C$ do  
3:         if $I_b(r,c) == 1$ Then  
4:             $I_F(r-1, c-1) = 1$  
5:         end if  
6:         $I_F(r-1, c) = 1$  
7:         $I_F(r, c-1) = 1$  
8:         $I_F(r, c+1) = 1$  
9:         $I_F(r+1, c-1) = 1$  
10:        $I_F(r+1, c) = 1$  
11:       $I_F(r+1, c+1) = 1$  
12:    end if  
13: end for  
14: end for  

FIGURE 8. The datasets used in the experiment.

The image resolution of Dataset 2 is 320 × 240 and has a total of 66 frames, the target size is an aircraft with point shape moving fast from the left side of the sky to the right. The image resolution of Dataset 3 has an image resolution of 320 × 240, and has a total of 100 frames, the target size is about 4 × 9, the target is an aircraft moving slowly from left to right. The target is an aircraft with point shape moving fast from the left side of the sky to the right. The image resolution of Dataset 4 is 256 × 200 and has a total of 40 frames, the target size is about 4 × 15, and the target is an aircraft with strip shape moving fast from the right side of the sky to the left. The target is an aircraft with point shape moving fast from the right side of the sky to the left. The target is an aircraft with point shape moving fast from the right side of the sky to the left. The target size is about 4 × 2 and it moves slowly from right to left, sometimes approaching a circulating flight. The target is a personal computer with a 3.7 GHz AMD CPU. For performance evaluation, we use the following three indicators: signal to clutter ratio gain (SCRG), background suppression factor (BSF), and receiver operator characteristic (ROC) curve.

1)The signal-to-clutter ratio (SCR) describes the intensity of the target relative to the background and is defined in (15).

$$SCR = \frac{\mid \mu_t - \mu_b \mid}{\sigma_b}$$

(15)

where $\mu_t$ is the average grayscale value of the target, $\mu_b$ is the average grayscale value of the surrounding area, and $\sigma_b$ is the standard deviation of the grayscale value of the surrounding region.

The SCRG is defined as (16), in which $SCR_{in}$ and $SCR_{out}$ respectively the signal-to-noise ratio of the original image and the signal-to-noise ratio of the resulting image processed by the algorithm.

$$SCRG = \frac{SCR_{out}}{SCR_{in}}$$

(16)

The larger the SCR, the target is easier to be detected [29]. In the same way, the larger the SCRG, the better the detection performance. Table 1 shows the SCR for the four datasets in the experiment. According to the SCR, Dataset 1 is the most challenging, and Dataset 2 is the easiest for target detection.

2)Background suppression factor (BSF) is mainly used to indicate the ability of the detection method to suppress the background, the larger the value of BSF, the stronger the suppression ability of the method to the background, and the better the performance, and the equation is shown as (17).

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

(17)

where $C_{in}$ and $C_{out}$ represent the standard deviation of the entire image before and after processing by the algorithm, respectively.

3)ROC curve is a commonly used indicator to measure the detection results of infrared small targets, which can reflect

| Dataset | Image Resolution | Target Size (average) | Frame Number | Background description | SCR |
|---------|-----------------|----------------------|--------------|-----------------------|-----|
| 1       | 256 × 200       | 4 × 6                | 30           | Edge                  | 1.74|
| 2       | 320 × 240       | 8 × 9                | 100          | Smooth                | 8.89|
| 3       | 320 × 240       | 11 × 11              | 66           | Complex               | 2.64|
| 4       | 256 × 200       | 4 × 15               | 40           | Smooth                | 5.48|
| 5       | 320 × 240       | 4 × 2                | 100          | Complex               | 1.17|
| 6       | 320 × 240       | 5 × 3                | 99           | Complex               | 2.76|

Differences in backgrounds for testing from [27], [28], as Fig. 8 shown. Dataset 1 has an image resolution of 256 × 200 and a total of 30 frames, the target size is about 4 × 6, clouds mainly cover the target, and several image targets happen to be at the edge of clouds. The target is an aircraft with point shape moving fast from the left side of the sky to the right. The image resolution of Dataset 2 is 320 × 240 and has a total of 66 frames, the target size is about 8 × 9, the target is an aircraft moving slowly at the left corner of the image. And thick clouds cover the right side of the image. Dataset 3 has an image resolution of 320 × 240, a total of 100 frames, the target size is about 11 × 11, and the target is an aircraft flies from a clear sky to a thick cloud layer in the upper right, and the target is mainly in the high brightness clouds. The image resolution of Dataset 4 is 256 × 200, has a total of 40 frames, the target size is about 4 × 15, and the target is an aircraft with strip shape moving fast from the right side of the sky to the left. The target is submerged in a stationary cloud background, but interference by several pixel-sized noises with high brightness(PNHB).

Dataset 6 has an image resolution of 320 × 240, has a total of 99 frames, the target size is an aircraft with ellipse shape of about 5 × 3 and moves slowly from left to right. The background clutter is complex and cover almost half image. For details of the test dataset, see Table 1.

All experiments are implemented by Matlab R2020A on a personal computer with a 3.7 GHz AMD CPU. For performance evaluation, we use the following three indicators: signal to clutter ratio gain (SCRG), background suppression factor (BSF), and receiver operator characteristic (ROC) curve.
FIGURE 9. Detection results and 3-D gray distribution map obtained by six detection methods on Dataset 1. (a) Original image. (b) to (g) denote the corresponding detection results obtained by six detection methods. (h) 3-D gray distribution map of original image. (i) to (n) denote the 3-D gray distribution map obtained by six detection methods. The real target is displayed as red rectangle. STLDM detect two false targets at the upper right corner and TVPCF detect one false at the left corner.

FIGURE 10. Detection results and 3-D gray distribution map obtained by six detection methods on Dataset 2. (a) Original image. (b) to (g) denote the corresponding detection results obtained by six detection methods. (h) 3-D gray distribution map of original image. (i) to (n) denote the 3-D gray distribution map obtained by six detection methods. The real target is displayed as red rectangle. Each algorithm can successfully detect the target.

the relationship between the false alarm rate ($P_f$) and the detection probability ($P_d$), and can be used to compare the detection performance. The false alarm rate $P_f$ is defined as the number of pixels detected by the false alarm divided by the number of all pixels, as shown in (18); the detection rate $P_d$ is defined as the number of pixels detected to the actual target divided by the number of pixels of the real target, as shown in (19).

$$P_f = \frac{The\ number\ of\ false\ alarm\ detected\ pixels}{The\ number\ of\ all\ pixels}$$  \hspace{1cm} (18)

$$P_d = \frac{The\ number\ of\ true\ target\ detected\ pixels}{The\ number\ of\ real\ target\ pixels}$$  \hspace{1cm} (19)

B. EXPERIMENTAL RESULTS

The proposed method compared with five representative spatial-temporal based methods: TVPCF [13], STLCF [14], STLCM [15], STLDM [16], and NSTSM [17]. Table 2 shows the detailed parameters setup in this experiment.

Fig. 9 to Fig.14 demonstrate the detection results and 3D plots of six algorithms in the experiments. Fig. 9 shows that STLCM is the only algorithm that misses the target, and others can successfully detect the target. It is worth noting that while all methods except STLCM can detect the small target, STLCF, STLDM, and TVPCF show additional false targets. The main reason is that the background of Dataset 1 is more complex, readily resulting in other misjudgment errors. For Dataset 1, only the proposed method and NSTSM perform better in suppressing complex backgrounds. In Fig.10 to Fig.12, each algorithm can successfully detect the target. Fig.13 is a high-brightness background, and the contrast between the target and the background is not obvious, and there is also PNHB interference, resulting in STLCF, STLCM miss the target, NSTSM and STLDM produced many false targets. Only the proposed method and TVPCF perform better in suppressing complex backgrounds. Fig.14 shows that STLCF, STLCM still miss the target, it can clearly found that the most of the noise can be suppressed by the proposed method while other methods still produced many false targets. From Fig. 9 to Fig.14, it can be observed that the proposed method performs very well in the detection of the small target and can effectively suppress the complex background and PNHB in the experimental datasets.

Table 3 compares the six algorithms’ false alarm rate and detection rate, and the proposed method is the best and can detect all targets even in complex scenarios. Almost all algorithms have high detection accuracy in smooth scenes, such as Dataset 2 to Dataset 4. Only the STLCM is somewhat...
FIGURE 11. Detection results and 3-D gray distribution map obtained by six detection methods on Dataset 3. (a) Original image. (b) to (g) denote the corresponding detection results obtained by six detection methods. (h) 3-D gray distribution map of original image. (i) to (n) denote the 3-D gray distribution map obtained by six detection methods. The real target is displayed as red rectangle. Each algorithm can successfully detect the target.

FIGURE 12. Detection results and 3-D gray distribution map obtained by six detection methods on Dataset 4. (a) Original image. (b) to (g) denote the corresponding detection results obtained by six detection methods. (h) 3-D gray distribution map of original image. (i) to (n) denote the 3-D gray distribution map obtained by six detection methods. The real target is displayed as red rectangle. Each algorithm can successfully detect the target.

FIGURE 13. Detection results and 3-D gray distribution map obtained by six detection methods on Dataset 5. (a) Original image. (b) to (g) denote the corresponding detection results obtained by six detection methods. (h) 3-D gray distribution map of original image. (i) to (n) denote the 3-D gray distribution map obtained by six detection methods. The real target is displayed as red rectangle. STLCF and STLCM missed detecting the target, NSTSM and STLDM detect several false targets.

worse because the movement of small targets is too small, and they may be mistaken as background and suppressed. In more complex scenes, such as Dataset 1, the target stays at the edge, and the grayscale value of the background and the target is quite close. Most algorithms will suppress the potential target, resulting in a low detection rate, in which TVPCF is less than average. When the background become more complicated such as Dataset 5 and Dataset 6, the target is even submerged in the background, the detection performances of STLCF and STLCM deteriorate dramatically, and the target is not even detected. Only the TVPCF and the proposed method can suppress most background noise and PNHB.

Table 4 shows the SCRG and BSF of the six algorithms. The proposed method significantly outperforms other algorithms, with SCRG ranging from 6.115 to 75.54 and BSF
FIGURE 14. Detection results and 3-D gray distribution map obtained by six detection methods on Dataset 6. (a) Original image. (b) to (g) denote the corresponding detection results obtained by six detection methods. (h) 3-D gray distribution map of original image. (i) to (n) denote the 3-D gray distribution map obtained by six detection methods. The real target is displayed as red rectangle. STLCF and STLCM missed detecting the target, NSTSM and STLDM and TVPCF detect many false targets.

TABLE 2. Detailed parameter settings of the baseline methods.

| No. | Methods | Abbreviations | Parameter settings |
|-----|---------|---------------|--------------------|
| 1   | Temporal variance and spatial patch contrast filter[13] | TVPCF | Mean filter: 3×3;5×5;7×7;Threshold:25 |
| 2   | Spatial-temporal local contrast filter[14] | STLCF | Patch size: 3×3; Threshold: 25 |
| 3   | Spatial-temporal local contrast map[15] | STLCM | Inner window:3×3;Outer window:7×7;Threshold:18 |
| 4   | Spatial-temporal local difference measure[16] | STLDM | Frame interval:5; Threshold:15 |
| 5   | Novel spatiotemporal saliency method[17] | NSTSM | Frame interval:5; Threshold:20 |
| 6   | The proposed | HPSTM | Inner layer:2;Guard layer:4;Outer layer:6;Threshold: 12~13; Frame interval: 5 |

TABLE 3. Comparison of the detection rate $P_d$ of six methods. The bold means the best detection rate.

| Dataset | STLCF | STLCM | STLDM | TVPCF | NSTSM | Proposed |
|---------|-------|-------|-------|-------|-------|----------|
| 1       | 70%   | 60%   | 56.6% | 36.6% | 76.6% | 100 %    |
| 2       | 100%  | 97%   | 100%  | 100%  | 100%  | 100%     |
| 3       | 100%  | 90.9% | 100%  | 100%  | 100%  | 100%     |
| 4       | 100%  | 95%   | 100%  | 100%  | 100%  | 100%     |
| 5       | 0 %   | 0 %   | 22%   | 26%   | 77%   | 85%      |
| 6       | 0 %   | 0 %   | 23.2% | 21.2% | 20.2% | 53.5%    |

ranging from 0.346 to 0.931. For dataset 1, the BSF of STLCF is the highest at 0.994, and the BSF of the proposed method is 0.931. However, from Table 3, the overall detection performance of the proposed method is still significantly better than STLCF.

Fig.15 shows the ROC curve of the six algorithms, and it shows the proposed method is closest in the upper left corner of the figure, indicating that the detection performance of our method is the best. Fig. 15. (a) shows the ROC curve of Dataset 1, in which the complex background and the target stay at the edge. Most methods except the proposed method resulted in an increased false-positive rate. When the false alarm rate is greater than $0.2\times10^{-3}$, TVPCF is poorer in performance, and the detection rate is only near 0.4. Fig. 15. (c) shows the ROC curve of Dataset 3 with a complex background. Among six algorithms, the proposed method, STLDM, and NSTSM, have better performance, while STLCF, STLCM, and TVPCF are relatively poor.

Fig. 15. (b) and Fig. 15. (d) show the ROC curve of Dataset 2 and Dataset 4, both of which are smooth backgrounds. Most methods can reach a good detection performance. STLDM performs relatively poorly in Dataset 4 when the false alarm rate exceeds $0.2\times10^{-3}$.

Fig. 15. (e) and Fig. 15. (f) show that because STLCF and STLCM did not detect the target, the curve is at the bottom. Although our method does not achieve the most perfect detection rate, it is the best performance among these methods. In summary, whether the scenes are smooth or complex backgrounds, the proposed method achieves the best detection performance.

IR small target detection usually requires real-time processing in military applications, such as early warning systems, air defense systems, coastal defense systems, missile
Table 4. Comparison of the evaluation metrics SCRG and BSF of six methods. The bold means the best result.

| Methods   | Dataset 1 | Dataset 2 | Dataset 3 | Dataset 4 | Dataset 5 | Dataset 6 |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
|           | SCRG      | BSF       | SCRG      | BSF       | SCRG      | BSF       |
| STLCF     | 11.20     | **0.994** | 19.01     | 0.257     | 27.79     | 0.215     |
| STLCM     | 15.44     | 0.373     | 27.73     | 0.215     | 15.45     | 0.121     |
| STLDM     | 13.28     | 0.449     | 25.99     | 0.264     | 23.58     | 0.318     |
| TVPCF     | 9.53      | 0.851     | 24.24     | 0.350     | 15.98     | 0.230     |
| NSTSM     | 15.93     | 0.474     | 23.99     | 0.220     | 31.11     | 0.313     |
| Proposed  | **21.76** | 0.931     | **31.50** | **0.427** | **44.79** | **0.381** |

Table 5. Runtime comparisons of the of six methods (in second). The bold means the best result.

| Dataset 1 | STLCF | STLCM | STLDM | TVPCF | NSTSM | Proposed |
|-----------|-------|-------|-------|-------|-------|----------|
|           | 0.435 | 0.339 | 1.201 | 0.072 | 0.289 | **0.016** |
| Dataset 2 | 0.688 | 0.416 | 1.831 | 0.075 | 0.433 | **0.023** |
| Dataset 3 | 1.012 | 0.413 | 1.829 | 0.095 | 0.400 | **0.021** |
| Dataset 4 | 0.431 | 0.545 | 1.206 | 0.084 | 0.273 | **0.016** |
| Dataset 5 | 0.387 | 0.401 | 2.087 | 0.094 | 0.297 | **0.036** |
| Dataset 6 | 0.406 | 0.402 | 2.337 | 0.092 | 0.344 | **0.032** |

We can find that the proposed method is the best of all methods. On average, the proposed method is faster than at least 0.05 seconds to over ten times, proving that our method is quite fast in computation while maintaining a good detection ability.

C. OPTIMAL MASK SIZE SELECTION (OMSS)

The size of the spatial domain tri-layer window gradient mask $M_G$ is highly critical, which must match the target size; too large or too small will seriously affect the detection results [31]. However, in practical applications the target size is usually unknown. We observed the relationship between the small target and the tri-layer window gradient mask $M_G$ as follows.

1) When the inner layer is smaller than the target size, the inner layer will detect the edge of the target, causing the detection result to be inconsistent with the original shape, and also making the potential target covered by the background, resulting in a decrease in the detection rate and an increase in the false alarm rate.

2) When the size of the inner layer is larger than the target size, the inner layer can completely cover the entire target. In this case, the target is easily divided and because the window is too large and has a smoothing-like effect, causes the contrast is too small, which closes the grayscale value of the target and the background, making the target difficult to be detected.

To ensure that the tri-layer window gradient mask $M_G$ achieves the best detection performance and the slightest noise interference without significantly increasing the computational cost simultaneously. We proposed the optimal mask size selection (OMSS) method for the tri-layer window gradient mask $M_G$, denoted as (20).

$$M_G(s_{lower}, t_{lower}) \leq M_G(s, t) \leq M_G(s_{upper}, t_{upper})$$  

where $M_G(s, t)$ represents the tri-layer window gradient mask, $s$ and $t$ represent the row and column of the gradient mask $M_G(s, t)$, respectively. $M_G(s_{lower}, t_{lower})$ and $M_G(s_{upper}, t_{upper})$ are the lower bound and upper bound for $s$ and $t$.

The range of the $s$ and $t$ for the optimal gradient mask size is calculated by (21) to (24).

$$s_{lower} = \arg \min \left\{ \frac{M}{2} + (2 \times 2) , M + (2 \times 2) \right\}$$  

$$s_{upper} = \arg \max \left\{ \frac{M}{2} + (2 \times 2) , M + (2 \times 2) \right\}$$

$$t_{lower} = \arg \min \left\{ \frac{N}{2} + (2 \times 2) , N + (2 \times 2) \right\}$$

$$t_{upper} = \arg \max \left\{ \frac{N}{2} + (2 \times 2) , N + (2 \times 2) \right\}$$

where $M$ and $N$ denote the small target size.

To successfully detect the small target, the inner layer size is at least half of the small target size, and both the guard layer and the outer layer are annular structures, occupying two rows and two columns respectively, so the size of the lower bound and the upper bound needs to be increased by four.

Table 6 shows the optimal gradient mask size range based on the OMSS for each experiment dataset. Except for the
6 × 6 weighted tri-layer window, three more different gradient masks were conducted in the experiments to evaluate the performance with different sizes. Fig. 16 shows the three gradient masks. Fig. 16. (a). shows the window size is 8 × 8 and the inner layer is 4 × 4. Fig. 16. (b). shows the window size is 10 × 10 and the inner layer is 6 × 6. Fig. 16. (c). shows the window size is 12 × 12 and the inner layer is 8 × 8. Note that the weight values of the inner layers are all set to positive integers, the guard layers are all set to 0, and the outer layers are all set to negative integers, so according to formula (4), the sum of the overall weight values is zero.
TABLE 7. Performance with gradient mask $M_{s}(s, t)$ of different sizes. The bold means the best detection rate.

| Dataset (Target size) | Dataset 1 (4x6) | Dataset 2 (8x8) | Dataset 3 (11x11) | Dataset 4 (4x15) | Dataset 5 (4x2) | Dataset 6 (5x3) |
|-----------------------|-----------------|-----------------|-------------------|-----------------|----------------|----------------|
| $M_{s}(6, 6)$, inner layer = 2x2 | 100% | 100% | 100% | 100% | 85% | 53.5% |
| $M_{s}(8, 8)$, inner layer = 4x4 | 26.6% | 100% | 100% | 100% | 0% | 0% |
| $M_{s}(10, 10)$, inner layer = 6x6 | 0% | 100% | 100% | 100% | 0% | 0% |
| $M_{s}(12, 12)$, inner layer = 8x8 | 0% | 100% | 100% | 100% | 0% | 0% |

FIGURE 16. The weighted setting of different gradient masks. (a) 8 × 8 gradient mask. (b) 10 × 10 gradient mask. (c) 12 × 12 gradient mask.

The experimental results are shown in Table 7. Because the target of Dataset 1 is 4 × 6, only the 6 × 6 and 8 × 8 gradient masks can detect the target. Although the 8 × 8 window can detect the target, the result is not very good. The reason may be that Dataset 1 is an image with a low SCR, which is not easy to detect. InDatasets 2 to 4, because the sizes of the targets are relatively larger than that of the gradient masks, the targets can all be successfully detected. Dataset 5 and Dataset 6 are also low SCR images, and the target’s size is smaller than 6 × 6. Thus only the 6 × 6 gradient mask can detect the target. The experimental results show that better detection results can be obtained as long as the appropriate gradient mask size is selected.

V. CONCLUSION

This paper proposes a fast IR small infrared target detection method that combines spatial-temporal matrices. In the spatial domain, the weighted tri-layer window is used to convolute with the original image and obtain the spatial gradient matrix $M_{s}$. In the temporal domain, the difference between the maximum and minimum grayscale values in the three forward and backward frames are calculated to extract the temporal contrast matrix $M_{t}$. The $M_{s}$ and $M_{t}$ matrices are computed via the Hadamard product, which significantly reduces computation cost and further uses a simple threshold and aggregation algorithm to locate the target.

We also proposed the optimal mask size selection (OMSS) method to obtain the optimal gradient mask size in the spatial domain. The experimental results show that the proposed HPSTM method has excellent performance, high detection accuracy and low computational cost, and is suitable for real-time IR small target detection.

REFERENCES

[1] R. Eysa and A. Hamdulla, “Issues on infrared dim small target detection and tracking,” in Proc. Int. Conf. Smart Grid Electr. Autom. (ICSGEA), Xiangtan, China, Aug. 2019, pp. 452–456.
[2] 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.
[3] M. Zhao, W. Li, L. Li, J. Hu, P. Ma, and R. Tao, “Single-frame infrared small-target detection: A survey,” IEEE Geosci. Remote Sens. Mag., vol. 10, no. 2, pp. 87–119, Jun. 2022.
[4] Y. Qin and B. Li, “Effective infrared small target detection utilizing a novel local contrast method,” IEEE Geosci. Remote Sens. Lett., vol. 13, no. 12, pp. 1890–1894, Dec. 2016.
[5] 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. Remote Sens. Lett., vol. 15, no. 4, pp. 612–616, Apr. 2018.
[6] S. D. Deshpande, M. H. Er, V. Ronda, and P. Chan, “Max-mean and max-median filters for detection of small-targets,” Proc. SPIE, vol. 3809, pp. 74–83, Oct. 1999.
[7] Y. Cao, R. Liu, and J. Yang, “Small target detection using two-dimensional least mean square (TDLMS) filter based on neighborhood analysis,” J. Infr., Millim., THz Waves, vol. 29, no. 2, pp. 188–200, Feb. 2008.
[8] P.-Y. Lv, S.-L. Sun, C.-Q. Lin, and G.-R. Liu, “Space moving target detection and tracking method in complex background,” Infra. Phys. Technol., vol. 91, pp. 107–118, Jun. 2018.
[9] C. L. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, “A local contrast method for infrared small-target detection,” IEEE Geosci. Remote Sens. Lett., vol. 52, no. 1, pp. 574–581, Jan. 2014.
[10] J. Han, S. Moradi, I. Faramarzi, C. Liu, H. Zhang, and Q. Zhao, “A local contrast method for infrared small-target detection utilizing a tri-layer window,” IEEE Geosci. Remote Sens. Lett., vol. 17, no. 10, pp. 1822–1826, Oct. 2020.
[11] L. Wu, Y. Ma, F. Fan, M. Wu, and J. Huang, “A double-neighborhood gradient method for infrared small target detection,” IEEE Geosci. Remote Sens. Lett., vol. 18, no. 8, pp. 1476–1480, Aug. 2021.
[12] J. Han, S. Moradi, I. Faramarzi, H. Zhang, Q. Zhao, X. Zhang, and N. Li, “Infrared small target detection based on the weighted strengthened local contrast measure,” IEEE Geosci. Remote Sens. Lett., vol. 18, no. 9, pp. 1670–1674, Sep. 2021.
[13] J. Gao, Z. Lin, and W. An, “Infrared small target detection using a temporal variance and spatial patch contrast filter,” IEEE Access, vol. 7, pp. 32217–32226, 2019.
[14] L. Z. Deng, H. Zhu, C. Tao, and Y. T. Wei, “Infrared moving point target detection based on spatial-temporal local contrast filter,” Infr. Phys. Technol., vol. 76, pp. 168–173, May 2016.
[15] B. Zhao, S. Xiao, H. Lu, and D. Wu, “Spatial-temporal local contrast for moving point target detection in space-based infrared imaging system,” Infr. Phys. Technol., vol. 95, pp. 53–60, Dec. 2018.
[16] P. Du and A. Hamdulla, “Infrared moving small-target detection using spatial-temporal local difference measure,” IEEE Geosci. Remote Sens. Lett., vol. 17, no. 10, pp. 1817–1821, Oct. 2020.
[17] D. Pang, T. Shan, P. Ma, W. Li, S. Liu, and R. Tao, “A novel spatiotemporal saliency method for low-altitude slow small infrared target detection,” IEEE Geosci. Remote Sens. Lett., vol. 19, pp. 1–5, 2022.
[18] Q. Hou, Z. Wang, F. Tan, Y. Zhao, H. Zheng, and W. Zhang, “RISTDNNet: Robust infrared small target detection network,” IEEE Geosci. Remote Sens. Lett., vol. 19, pp. 1–5, 2022.
[19] Y. Dai, Y. Wu, F. Zhou, and K. Barnard, “Attentional local contrast networks for infrared small target detection,” IEEE Trans. Geosci. Remote Sens., vol. 59, no. 11, pp. 9813–9824, Nov. 2021.
[20] C. Yu, Y. Liu, S. Wu, Z. Hu, X. Xia, D. Lan, and X. Liu, “Infrared small target detection based on multiscale local contrast learning networks,” Infr. Phys. Technol., vol. 123, Jun. 2022, Art. no. 104107.
[21] Y. Gu, C. Wang, B. Liu, and Y. Zhang, “A kernel-based nonparametric regression method for clutter removal in infrared small-target detection applications,” IEEE Geosci. Remote Sens. Lett., vol. 7, no. 3, pp. 469–473, Jul. 2010.
[22] J. Han, C. Liu, Y. Liu, Z. Luo, X. Zhang, and Q. Niu, “Infrared small target detection utilizing the enhanced closest-mean background estimation,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 14, pp. 645–662, 2021.
[23] T.-H. Hsieh, C.-L. Chou, Y.-P. Lan, P.-H. Ting, and C.-T. Lin, “Fast and robust infrared image small target detection based on the convolution of layered gradient kernel,” IEEE Access, vol. 9, pp. 94889–94900, 2021.
[24] T.-W. Bae, “Spatial and temporal bilateral filter for infrared small target enhancement,” Infr. Phys. Technol., vol. 63, pp. 42–53, Mar. 2014.
[25] C. R. Johnson, “The Hadamard product,” in Matrices: Theory and Applications, vol. 40, Phoenix, AZ, USA: American Mathematical Society, 1990, pp. 87–169.
[26] V. Antun, “Coherence estimates between Hadamard matrices and Daubechies wavelets,” M.S. thesis, Dept. Math., Oslo Univ., Oslo, Norway, 2016.
[27] P. Zhang, X. Wang, X. Wang, C. Fei, and Z. Guo, “Infrared small target detection based on spatial-temporal enhancement using quaternion discrete cosine transform,” IEEE Access, vol. 7, pp. 54712–54723, 2019.
[28] Y. Li, (May 2022). An-Infrared-Small-Target-Dataset-1. [Online]. Available: https://github.com/liforeversong/An-infrared-small-target-dataset-1
[29] C. Gao, L. Wang, Y. Xiao, Q. Zhao, and D. Meng, “Infrared small-dim target detection based on Markov random field guided noise modeling,” Pattern Recognit., vol. 76, pp. 463–475, Apr. 2018.
[30] H. Liu, L. Zhang, and H. Huang, “Small target detection in infrared videos based on spatio-temporal tensor model,” IEEE Trans. Geosci. Remote Sens., vol. 58, no. 12, pp. 8689–8700, Dec. 2020.
[31] X. Yi, B. Wang, H. Zhou, and H. Qin, “Dim and small infrared target fast detection guided by visual saliency,” Infr. Phys. Technol., vol. 97, pp. 6–14, Mar. 2019.

PING-HSIU WU was born in Kaohsiung, Taiwan, in 1998. He received the B.S. degree in electrical engineering from the National Chianghua University of Education, Chianghua, Taiwan, in 2020. He is currently pursuing the M.S. degree with the Institute of Lighting and Energy Photonics, National Yang Ming Chiao Tung University, Tainan, Taiwan. His current research interests include image processing and infrared target detection.

YU-PIN LAN received the M.S. and Ph.D. degrees from the Institute of Electro-Optical Engineering, National Chiao Tung University, Hsinchu, Taiwan. She has an extensive professional career both in research and industry. Since 2016, she has been an Assistant Professor with the College of Photonics, National Yang Ming Chiao Tung University (NYCU), Taiwan. Her research interests include the laser engineering and physics.

CHAO-LUNG CHOU received the Ph.D. degree in electrical and electronics engineering from the Chung Cheng Institute of Technology, National Defense University, Taiwan, in 2012. Since 2015, he has been an Associate Professor with the Computer Science and Information Engineering Department, Chung Cheng Institute of Technology, National Defense University. His research interests include image processing, and biometrics.

CHUN-TING LIN received the B.S. and M.S. degrees in material science and engineering from the National Tsing Hua University, Hsinchu, Taiwan, in 1997 and 2001, respectively, and the Ph.D. degree in electro-optical engineering from the National Chiao-Tung University, Hsinchu, in 2007. Since 2015, he has been a Professor with the College of Photonics, National Yang Ming Chiao Tung University (NYCU), Taiwan. His research interests include radio-over-fiber systems, optical data processing.