Infrared Moving Small-Target Detection Using Spatial–Temporal Local Difference Measure

Peng Du and Askar Hamdulla

***Abstract*— Effectiveness and false alarm (FA) suppression are key issues in infrared (IR) moving small-target detection. In this letter, we propose a novel spatial–temporal local difference measure (STLDM) algorithm to detect a moving IR small target. The method we propose involves three steps. First, we block off three frames in a certain range of time (temporal) domain. Next, in a 3-D spatial–temporal domain, we analyze the local grayscale intensity difference of the small targets moving between the three frames and calculate the difference between the grayscale intensity in the center area and the grayscale intensity in the eight directions in the area surrounding the target. Finally, we segment the small target in a detection result map. The results of our experiment demonstrate that our proposed STLDM method has a higher rate of detection of IR moving small targets, as well as fewer FAs than other existing methods.**

***Index Terms*— Infrared (IR) moving small target, spatial–temporal domain, spatial–temporal local difference measure (STLDM).**

I. Introduction

INFRARED search and track (IRST) systems have broad use in applications such as precision guidance, prewarning, remote sensing, and aerospace [1], [2] to detect and track objects that give off IR radiation. The ability to detect a small IR target of unknown position and velocity at a low signal-to-noise ratio (SNR) is a necessity in IRST systems so that military applications can warn of incoming small targets, such as enemy aircraft and helicopters [3], [4], from a distance detecting an IR small target can be difficult because the target is only a dim and small spot, without a concrete shape, texture, or structure information [3], [5], and the image can easily be drowned out by complex backgrounds. Therefore, the detection of a moving IR small target under low SNR and complex background conditions is a challenge.

Small-target detection methods are generally based either on single-frame detection using spatial filtering or on multiframe detection using spatial–temporal filtering.

Most human visual system (HVS)-based models are based on single-frame-based spatial filtering. For example, Wang *et al.* [6] proposed a filter template called difference of Gaussians (DoG) that can enhance a target image, but cannot suppress the background clutter very well. Qin and Li [7] proposed the novel local contrast measure (NLCM), which uses a nested structure in which the surrounding area is divided into eight directions. This method's detection performance is better than that of traditional LCM [2]. Wei *et al.* [8] proposed the multiscale patch-based contrast measure (MPCM), which merged two corresponding directions and used difference operations to detect the target. The difference operations do a good job of suppressing the background clutter, but the enhancement of the target image is poor. Han *et al.* [5] proposed a multiscale relative LCM (RLCM) that uses both ratio and difference operations to simultaneously enhance the target and suppress the background clutter. However, these single-frame algorithms based on spatial filters are often limited in their detection capability, and it is difficult for them to remove false alarms (FAs) completely because they neglect the spatial and temporal correlation of moving targets in multiframe sequences.

In recent years, some researchers have had success in detecting small targets in a multiframe-based spatial–temporal domain. For example, Sengar and Mukhopadhyay [9] proposed a moving object detection method based on frame difference and W4. Wan *et al.* [10] proposed a method based on a saliency histogram and geometrical invariability. Deng *et al.* [11] proposed a spatial–temporal local contrast filter (STLCF) to enhance the target's contrast. Zhao *et al.* [12] proposed a spatial–temporal LCM (STLCM) that uses mean filter prediction of the differences in a spatial domain and an enhanced time domain to detect small moving targets. Li *et al.* [13] proposed a novel spatio-temporal saliency model and a prediction method (STP). These multiframe-based methods have achieved strong detection results; however, due to the simple time-domain differential operation in some methods, the local grayscale intensity difference information of small moving targets between frames may be missed. As a result, when the background is complex or cluttered, the FA rate is still high, and the detection rate is not significantly better than with methods using the single-frame algorithm.

In this letter, unlike the above mentioned multiframe-based methods, which view the space and time domains as two aspects to calculate, our proposed method directly detects targets in a 3-D space–time domain.

We analyze the characteristics of the local grayscale intensity difference of small moving targets in the process of interframe motion and propose a simple and effective spatial–temporal detection method called spatial–temporal local difference measure (STLDM). As shown in Fig. 1, it includes two processes, the DMAP process and IMAP process, which are described in detail in the following. The results of our experiments demonstrate that our proposed

Manuscript received June 28, 2019; revised October 22, 2019; accepted November 15, 2019. This work was supported by the National Natural Science Foundation of China under Grant 61563049. *(Corresponding author: Askar Hamdulla.)*

The authors are with the Institute of Information Science and Engineering, Xinjiang University, Ürümqi 830046, China (e-mail: hysddp@163.com; askar@xju.edu.cn).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2019.2954715

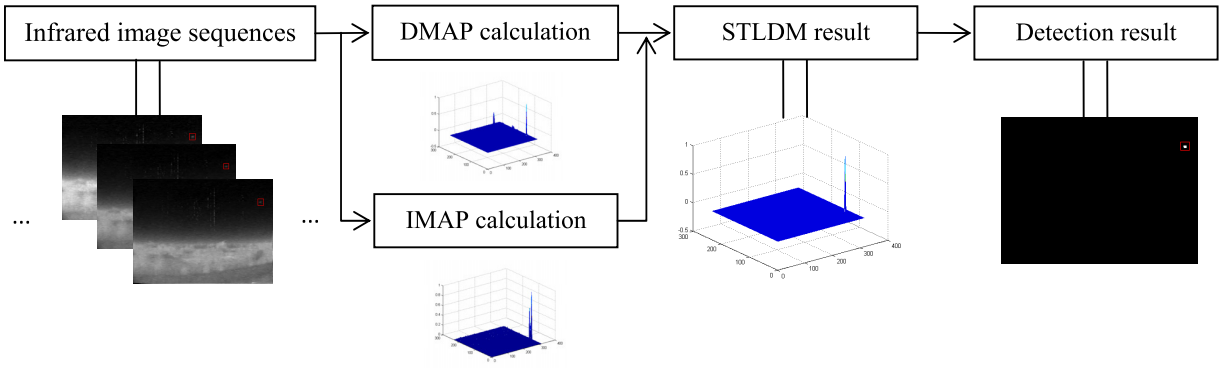


Fig. 1. Proposed IR moving small-target detection system.

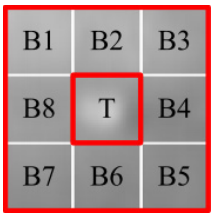


Fig. 2. Nested structure of the sliding window in one frame.

method has a higher rate of IR small-target detection than other methods.

II. Proposed Algorithm

The visual saliency of the HVS model identifies a local difference between the target center and the surrounding backgrounds of a moving IR small target in the spatial domain. In addition, due to the change in the target position in the temporal domain, the grayscale intensity of the target center and the surrounding area will correspondingly change. In the case of the 3-D spatial–temporal domain, for moving small targets, we use the feature in which the local gray intensity changes during the moving process to detect it.

A. STLDM Calculation

Let $\ldots f_{n-l}, \ldots, f_{n-1}, f_n, f_{n+1}, \ldots, f_{n+l} \ldots$be an IR image sequence, in which f_n denotes the current calculated frame and l is the number of forward or backward frames to be considered. Here, we use the nested structure of the sliding window to block the three frames of f_{n-l}, f_n, and f_{n+l} in order to calculate the local differences between frames in a 3-D spatial–temporal domain, as shown in Fig. 2.

Nine subblocks make up an image patch. The central subblock, labeled as T, B_i $(i = 1, 2, \ldots, 8)$, denotes the eight directions of the surrounding subblocks. Each subblock measures 3×3 pixels, which can better capture the details of grayscale intensity changes of a moving small target. In order to detect moving targets, we first define three equations

$$D_{i\_}f_{n-l} = M_T - I_{Bi} \quad (i = 1, 2, \ldots, 8) \tag{1}$$

$$D_{i\_}f_n = M_T - I_{Bi} \quad (i = 1, 2, \ldots, 8) \tag{2}$$

$$D_{i\_}f_{n+l} = M_T - I_{Bi} \quad (i = 1, 2, \ldots, 8) \tag{3}$$

where M_T and I_{Bi} denote the maximum grayscale value in areas T and the average grayscale value in areas B_i, respectively. Therefore, $\{D_{i\_}f_{n-l}, D_{i\_}f_n, D_{i\_}f_{n+l}\}$ represent

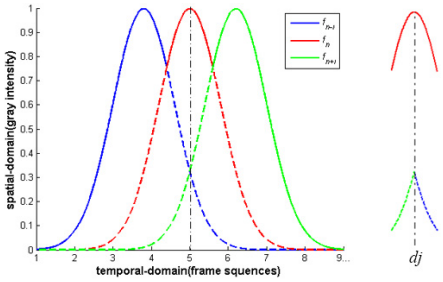


Fig. 3. Schematic of the proposed STLDM method.

the local grayscale difference in different directions, between T in the center area and Bi in the surrounding area of the three frames $\{f_{n-l}, f_n, f_{n+l}\}$, respectively.

Next, we calculate four groups of eight directional differences of local grayscale intensity in a spatial–temporal domain

$$d_j = (D_{j\_}f_n - D_{j\_}f_{n+l})^3 + (D_{j+4\_}f_n - D_{j+4\_}f_{n-l})^3 \quad (j = 1, 2, 3, 4). \tag{4}$$

At this point, we can obtain a map based on the difference between the grayscale intensity between the center area and the surrounding area of the moving target. This map is called DMAP

$$\mathrm{DMAP}(i, j) = \max d_j \quad (j = 1, 2, 3, 4) \tag{5}$$

$$\mathrm{DMAP}(i, j) = \frac{\mathrm{DMAP}(i, j)}{\max_{i,j}\{\mathrm{DMAP}(i, j)\}}. \tag{6}$$

In addition, we use the grayscale intensity change of the T center region in a spatial–temporal domain to further increase the moving target detection rate and suppress FAs. This we call IMAP

$$\mathrm{Imax} = \max\{M_{T\_}f_{n-l}, M_{T\_}f_n, M_{T\_}f_{n+l}\} \tag{7}$$

$$\mathrm{Imin} = \min\{M_{T\_}f_{n-l}, M_{T\_}f_n, M_{T\_}f_{n+l}\} \tag{8}$$

$$\mathrm{IMAP}(i, j) = (\mathrm{Imax} - \mathrm{Imin})^2 \tag{9}$$

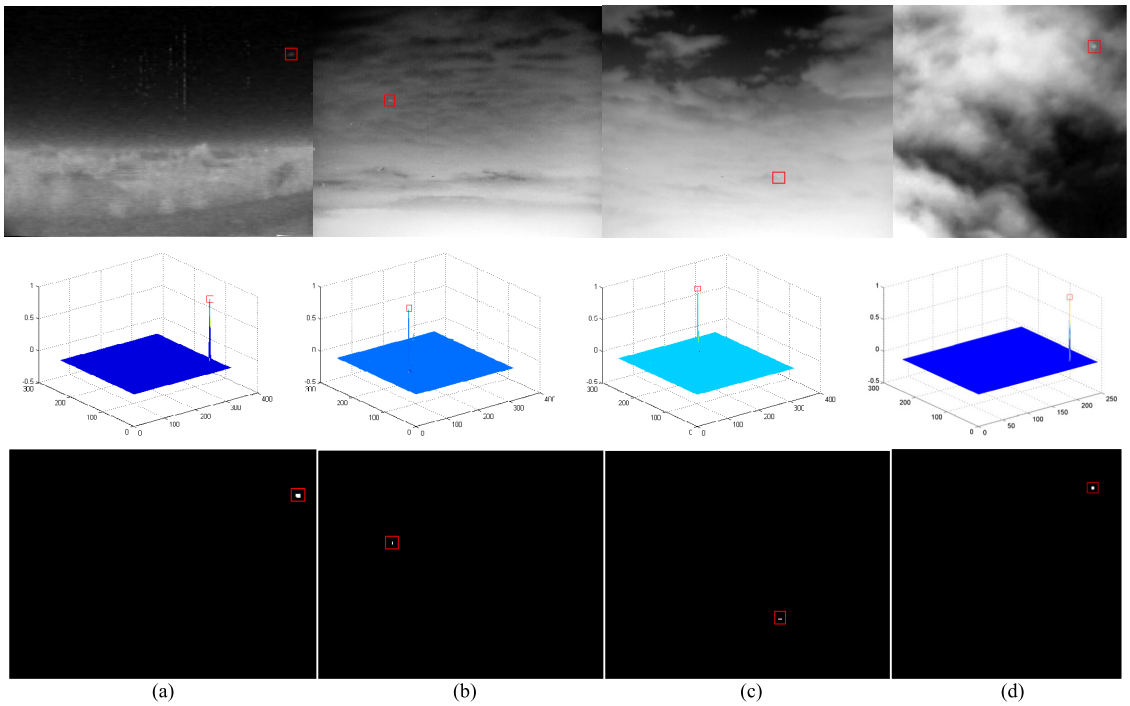
$$\mathrm{IMAP}(i, j) = \frac{\mathrm{IMAP}(i, j)}{\max_{i,j}\{\mathrm{IMAP}(i, j)\}}. \tag{10}$$

The detection result map based on the STLDM is defined as

$$\mathrm{STLDM}(i, j) = \mathrm{DMAP}(i, j) \times \mathrm{IMAP}(i, j). \tag{11}$$

The proposed method is illustrated by the schematic shown in Fig. 3.

The blue, red, and green curves, respectively, represent the local grayscale intensity distribution of the three frames



(a) (b) (c) (d)

Fig. 4. Detection results of four sequences using the proposed STLDM algorithm. From top to bottom: the raw IR image, the calculation result of STLDM ($l = 5$), and the final detection result for (a) Seq. 1, (b) Seq. 2, (c) Seq. 3, and (d) Seq. 4.

near the moving small target, with the red curve representing the frame currently being calculated. Combining Fig. 3 with (1)–(11), we can see that the proposed STLDM method, by taking advantage of the local differences in grayscale intensity of small moving targets in a 3-D spatial–temporal domain, can effectively enhance the target and reduce FAs.

B. Target Segmentation

After obtaining the detection result map through the use of STLDM, the true small target will be the most salient. In this article, a simple adaptive threshold operation [2] is used to extract the target, and the threshold Th is defined as

$$\mathrm{Th} = \mu_{\mathrm{STLDM}} + k \times \sigma_{\mathrm{STLDM}} \tag{12}$$

where μ_{STLDM} and σ_{STLDM} are the mean and standard deviations, respectively, of the final salient STLDM map, and k is the constant adjustment coefficient. Our experiments show that the optimal range of k is from 5 to 15.

III. Experimental Results

To demonstrate the effectiveness of our algorithm, we used four groups of true IR image sequences (Seq. 1 [14], Seq. 2, Seq. 3, and Seq. 4). The details of the four sequences are listed in Table I. In the experiment, STLCF, STLCM, and our proposed method all performed calculations based on the use of five forward or backward frames, that is, $l = 5$.

We conducted all our experiments on a computer with 8 GB of RAM and an Intel Core i7-920 2.66 GHz processor, running MathWorks' MATLAB R2014b software environment.

Fig. 4 shows the detection results of our proposed method. We can see that the small targets on the four image sequences are very dim, particularly in the third sequence, and that all four backgrounds are very complex. After the STLDM calculation, the target is salient and the background suppression is very good, so that the target can be easily segmented without any FAs.

TABLE I
Details of the Four Sequences

	Frames	Size	Target number	Target size	Target type
Seq. 1	400	320×240	1	5×7	Plane target
Seq. 2	200	320×256	1	3×5	Plane target
Seq. 3	200	320×256	1	3×3 to 5×3	Plane target
Seq. 4	300	250×250	1	5×5 to 7×7	Plane target

Fig. 5 displays a comparison of the detection results from the use of seven similar state-of-the-art algorithms—DoG [6], NLCM [7], MPCM [8], RLCM [5], STLCF [11], STLCM [12], STP [13], and the same samples, as shown in Fig. 4. In the sample images of Sequences 1, 2, and 4, our method and some of the other methods can detect the target, but the target detected by our method is brighter and more salient, and there are no FAs. In Sequence 3 in particular, the small target is extremely weak, and the contrast with the surrounding background is very low. In comparative experiments, the seven methods employed in our experiment could not detect the target, while our proposed STLDM algorithm could detect the real target without any FAs. Therefore, it can be stated that the STLDM algorithm was able to detect targets better than the other seven algorithms.

To further verify the effectiveness and robustness of our method, we used two numerical metrics—SCR gain (SCRG) and background suppression factor (BSF)—to evaluate the

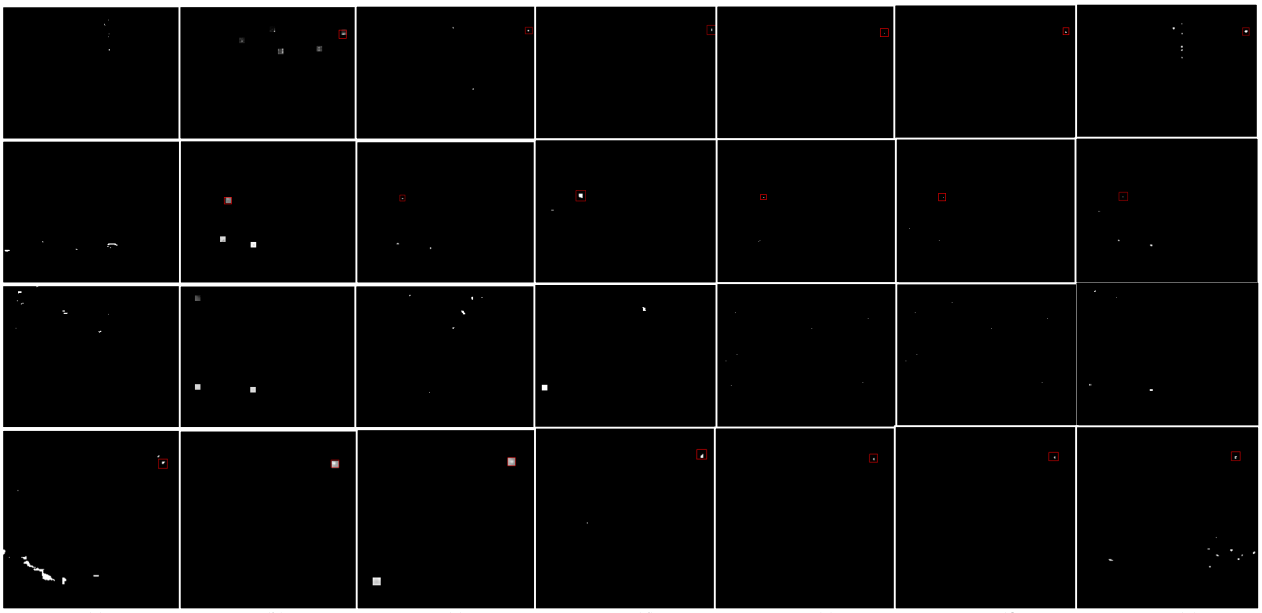


Fig. 5. From top to bottom, the detection results for Seq. 1–Seq. 4 using (a) DoG, (b) NLCM, (c) MPCM, (d) RLCM, (e) STLCF, (f) STLCM, and (g) STP.

TABLE II
SCRG AND BSF VALUES FOR THE DIFFERENT ALGORITHMS

	Seq.	DoG	NLCM	MPCM	RLCM	STLCF	STLCM	STP	Proposed
SCRG	1	4.0830	13.7804	5.5608	19.4122	95.3506	1.9532E2	1.0028E2	**2.0447E2**
	2	5.6458	21.4261	8.0379	12.5257	1.4699E2	2.4960E2	8.0249	**2.7848E2**
	3	2.1488	16.8341	3.4409	1.3223	95.2506	87.7870	8.7265	**1.1144E3**
	4	4.3236	29.8577	4.2024	2.1810	1.8272E2	2.0355E2	12.5418	**2.2702E2**
BSF	1	2.0768	1.2955	0.5762	5.9242	15.3668	38.1836	22.3307	**39.3031**
	2	3.2626	0.0286	0.0273	11.1019	31.1893	67.5444	39.6089	**78.3595**
	3	4.1657	10.5651	4.5133	12.1450	15.0912	39.0749	37.5210	**75.8990**
	4	3.4304	6.6738	4.9563	1.6624	40.4801	48.7341	2.7426	**49.9550**

degree of difficulty of small-target detection [3]. SCRG and BSF are defined as

$$\text{SCRG} = \frac{\text{SCR}_{\text{out}}}{\text{SCR}_{\text{in}}} \quad \text{and BSF} = \frac{\sigma_{\text{in}}}{\sigma_{\text{out}}} \tag{13}$$

where SCR_{in} and SCR_{out} are the SCR values of the raw image and STLDM map, respectively, and σ_{in} and σ_{out} represent the standard deviation of the raw image and STLDM map, respectively.

The SCRG and BSF evaluations are given in Table II. The highest values among the different methods are marked in bold. The method we proposed achieved the highest SCRG and BSF values in all sequences.

Next, in order to further discuss the accuracy of the detection results, we quote three metrics in [10], which are called precision rate P, recall rate R, and an evaluation index η, respectively. They can be defined as

$$P = \frac{N_C}{N_D} \times 100\%, \quad R = \frac{N_C}{N_T} \times 100\% \tag{14}$$

$$\eta = \frac{(1+\lambda^2) \cdot P \cdot R}{\lambda^2 \cdot P + R} \times 100\% \tag{15}$$

where N_D is the pixel number of targets detected by the tested method, N_T is the pixel number of true targets existing in the current frame, $N_C = N_D \cap N_T$ is the pixel number of targets detected correctly, η is a precision evaluation index, and λ is a harmonic coefficient. Here, we set $\lambda = 1$. In this letter, $\eta = 1$ means that all the real targets are discovered without any FAs and $\eta = 0$ means that no real targets are found. In Table III, we can see that our proposed method achieves high detection accuracy.

In the end, we used a receiver-operating characteristic (ROC) curve [15] to assess the detection performance. The probability of detection (PD) and the FA rate (FAR) are given in (16) and (17). The ROC curves of the eight algorithms on the four image sequences are shown in Fig. 6(a)–(d)

$$\text{PD} = \frac{\text{number of detected true targets}}{\text{total number of real targets}} \tag{16}$$

$$\text{FAR} = \frac{\text{number of detected false targets}}{\text{total number of pixels in the whole image}}. \tag{17}$$

Fig. 6 shows the ROC curves of our proposed method as being located in the upper left corner in all four sequences, which means that our method has the highest rate of detection. There is much background clutter in Seq. 1, and the detection performed by the single-frame algorithm is usually low. The STLCF and STLCM algorithms perform well in terms of detection, but not as well as our proposed algorithm. In Seq. 2, RLCM offers satisfactory detection performance, and the performance of STLCF and STLCM is slightly worse than that

TABLE III

STATISTICAL RESULTS OF $\eta(\%)$ FOR THE DIFFERENT DETECTION ALGORITHMS

	DoG	NLCM	MPCM	RLCM	STLCF	STLCM	STP	Proposed
Seq. 1	8.92	30.93	35.25	38.29	58.02	56.23	28.31	**85.23**
Seq. 2	15.45	65.83	68.04	80.25	62.23	58.55	19.60	**97.15**
Seq. 3	5.34	78.65	15.28	18.59	51.76	37.82	77.52	**95.93**
Seq. 4	10.53	83.82	16.38	52.35	92.98	94.23	21.53	**98.05**

TABLE IV

COMPUTATIONAL COST COMPARISON AMONG THE PROPOSED ALGORITHM AND OTHER SEVEN ALGORITHMS FOR A SINGLE FRAME (IN SECONDS)

	DoG	NLCM	MPCM	RLCM	STLCF	STLCM	STP	Proposed
Time consuming(s)	0.0085	0.0253	0.8342	0.0297	0.0558	0.0609	0.2458	0.0439

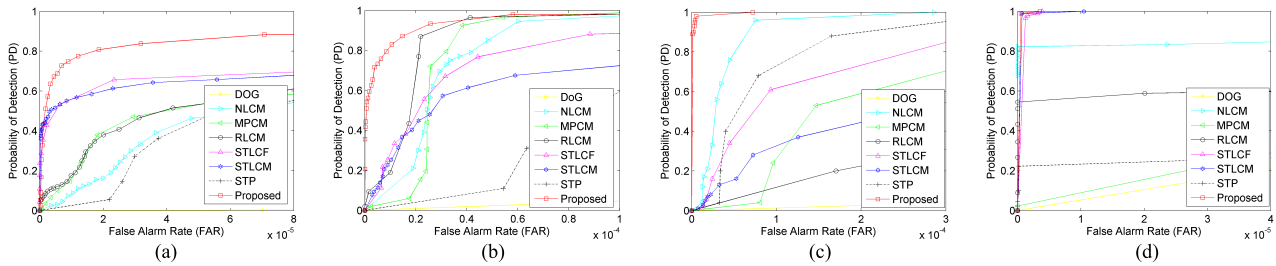


Fig. 6. ROC curves of different algorithms for (a) Seq. 1, (b) Seq. 2, (c) Seq. 3, and (d) Seq. 4.

of a single-frame algorithm, but again, our algorithm offers the highest performance. In Seq. 3, the performance of the STP and NLCM algorithms is good, but our method still obtains almost ideal PD, with few or even no FAs. In Seq. 4, our algorithm is still better than others, especially the single-frame algorithm.

Finally, we compare the average computational cost of the eight algorithms for a single image (see Table IV). Although the computational time required by our algorithm is not the lowest in all comparison methods, it is the lowest in the multiframe method.

IV. CONCLUSION

In this letter, we propose an IR moving small-target detection algorithm, the STLDM. The proposed method utilizes two features of the moving small target—the center grayscale intensity change and the surrounding grayscale intensity change—and detects the moving small target directly in the 3-D spatial–temporal domain. The results of our experiment demonstrate that our method is significantly more effective than of the other methods that it was compared to in various scenarios.

REFERENCES

[1] H. Deng, X. Sun, M. Liu, C. Ye, and X. Zhou, "Small infrared target detection based on weighted local difference measure," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 4204–4214, Jul. 2016.

[2] 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.

[3] C. Q. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, and A. G. Hauptmann, "Infrared patch-image model for small target detection in a single image," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4996–5009, Dec. 2013.

[4] J. Nie, S. Qu, Y. Wei, L. Zhang, and L. Deng, "An infrared small target detection method based on multiscale local homogeneity measure," *Infr. Phys. Technol.*, vol. 90, pp. 186–194, May 2018.

[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] X. Wang, G. Lv, and L. Xu, "Infrared dim target detection based on visual attention," *Infr. Phys. Technol.*, vol. 55, no. 6, pp. 513–521, Nov. 2012.

[7] 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.

[8] Y. Wei, X. You, and H. Li, "Multiscale patch-based contrast measure for small infrared target detection," *Pattern Recognit.*, vol. 58, pp. 216–226, Oct. 2016.

[9] S. S. Sengar and S. Mukhopadhyay, "Moving object detection based on frame difference and W_4," *Signal, Image Video Process.*, vol. 11, pp. 1357–1364, Oct. 2017.

[10] M. Wan *et al.*, "Infrared small moving target detection via saliency histogram and geometrical invariability," *Appl. Sci.*, vol. 7, no. 6, p. 569, Jun. 2017.

[11] L. Deng, H. Zhu, C. Tao, and Y. Wei, "Infrared moving point target detection based on spatial–temporal local contrast filter," *Infr. Phys. Technol.*, vol. 76, pp. 168–173, May 2016.

[12] 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.

[13] Y. Li, Y. Zhang, J.-G. Yu, Y. Tan, J. Tian, and J. Ma, "A novel spatio-temporal saliency approach for robust dim moving target detection from airborne infrared image sequences," *Inf. Sci.*, vol. 369, pp. 548–563, 2016.

[14] *IEEE OTCBVS WS Series Bench, Roland Miezianko, Terravic Research Infrared Database*. Accessed: Oct. 1, 2018. [Online]. Available: http://vcipl-okstate.org/pbvs/bench/index.html

[15] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," in *Proc. 23rd Int. Conf. Mach. Learn.*, 2006, pp. 233–240.