

Patch Image Based LSMR Method for Moving Point Target Detection

Weina Zhou^{1,2}(✉), Xinwei Lin¹, Zhijing Xu¹, and Xiangyang Xue²

¹ Fudan University, Shanghai 201203, China
{wnzhou, xwlin, zjxu}@shmtu.edu.cn

² Shanghai Maritime University, Shanghai 201306, China
xyxue@fudan.edu.cn

Abstract. The quick and high accuracy detection of point targets is a difficult but important technique in infrared surveillance. In this work, we proposed a patch image based low-rank and sparse matrices recovery (LSMR) method to detect moving point target in infrared marine surveillance. After analyzing the relationship between detection speed and image size, we wisely use frame difference and local threshold to efficiently exclude majority non-target patches before LSMR, which greatly accelerated the detection speed. Then integrated with the powerful classification ability of LSMR, the method proposed in this paper finally gains a high accuracy in point target detection in a marine environment. The experiment results show that the proposed method could effectively enhance the system's signal-to-clutter ratio gain and background suppression factor while keeping a high detection speed compared to other similar algorithms.

Keywords: Point target detection · Patch image · Low rank and sparse matrix recovery · Frame difference · Local threshold

1 Introduction

Moving point target detection is the most difficult and significant task in infrared target surveillance. Although a lot of methods have been proposed to detect point target during the past few years, their performance is still unsatisfactory due to many reasons. The main reasons resulting in the problems lie in two aspects: infrared point targets always have less texture, color and shape information, and the detection environment is commonly very harsh.

Recently, low-rank and sparse matrices recovery (LSMR) theory was proposed and proved to be much more effective than many traditional methods in a lot of fields [1–11]. And it is also applied in target detection and tracking to further improve the performance. They also can be classified into two kinds. The first kind prefers to detect target in a single image [5–8]. It can solve the problems when imaging backgrounds change quickly due to rapid relative motion between the imaging sensor and the target. And the other kind of methods makes use of information of multi-frames to predict the target position more quickly and accurate [9–11]. However, both of their performance could degrade rapidly when the signal-to-clutter ratio (SCR) is low, or the computation is too time-consuming to be realized in real-time.

In this paper, we fully exploited the advantage of LSMR algorithm, and proposed a patch-image based LSMR method to detect moving point targets in infrared marine surveillance. Different from previously proposed LSMR based algorithms, our method takes full advantage of the information from intra and inter frames, and greatly reduces computation time while obtaining a good performance on signal-to-clutter gain (SCRG) and background suppression factor (BSF).

This paper is organized as follows. In Sect. 2, preliminary work including the LSMR algorithm and its application in small target detection will be introduced briefly. In Sect. 3, the relationship between the computation amount and image size is analyzed, thus a patch image based LSMR method is proposed and discussed in detail. In Sect. 4, experiments are executed to validate the performance of the proposed approach. And in final section, a conclusion and our future work is presented for further research.

2 LRMR Theory and Small Target Detection

2.1 Low-Rank and Sparse Matrices Recovery (LRMR) Algorithm

As a theory stretching from compressive sensing (CS) and sparse represent algorithm, the LRMR supposes an ideal image or video as a low-rank matrix. Take image signal as an instance, its rank is usually much smaller than its size.

$$r \ll \min(m, n) \quad (1)$$

In (1), r is the rank of an ideal image, m, n are image's width and height respectively. In fact, the entries of the matrix are often corrupted by errors or noises, making the image or video not a low-rank matrix any more. Consequently, every image/video can then be regarded as the combination of a low-rank matrix and a noise matrix in LRMR theory, which could be represented as

$$I = L + S \quad (2)$$

I is an usual image/video, L is an ideal image/video which is low-rank, S represents random noise and errors, or other unregular signal caused by various of factors, which can be regarded as sparse matrix. It has been proved that, under surprisingly broad conditions, L and S can be exactly recovered from I via Robust Principal Component Analysis (RPCA) by solving the following convex optimization problem.

$$\min_{L, S} \|L\|_* + \lambda \|S\|_1, \quad \text{subject to } I = L + S \quad (3)$$

Here, $\|\bullet\|_*$ represents the nuclear norm of a matrix, $\|\bullet\|_1$ denotes the norm-1, and λ is a positive weighting parameter, which could be utilized to fine tune L and S for optimum results.

2.2 LSMR in Point Target Detection

Generally speaking, there are three components in an infrared image (\mathbf{I}): the background (B), the point targets (T), and various kinds of noises (N). And the infrared image can be expressed as follow:

$$\mathbf{I} = \mathbf{B} + \mathbf{T} + \mathbf{N} \quad (4)$$

When detecting point target in a single image, we found out that small targets have no concrete shape and texture, and are commonly not larger than 10×10 pixels because of the long imaging distance. They can be considered as “sparse” with respect to the extensive background. However, background patches are always “low-rank” and approximately context correlated with each other, even though the pixel distance between two patches may be large. As for detecting targets in multi-frames, background in each frame is also ‘low-rank’ due to the consistency between adjacent frames. And target is “sparse” because of its moving characteristics.

Thus the point target detection could be regarded as a typical problem of recovering a sparse component from the infrared image in LSMR theory. This assumption has been proved valid in [8], under the assumption that the random noise is i.i.d. and its Frobenius norm is smaller than some σ ($\sigma > 0$). Positive weighting parameter λ plays a promoting role in enhancing the detection stability and suppressing random noise.

$$\|\mathbf{I} - \mathbf{B} - \mathbf{T}\|_F \leq \sigma \quad (5)$$

Nevertheless, the recovery of low-rank and sparse matrix usually leads to high computational cost due to demanding convex optimization of RPCA especially with multi-frames. And the method of detecting in a single frame always cannot ensure the detection accuracy.

3 Patch Based LSMR Algorithm

How to improve the detection speed while keeping a high accuracy? A figure was drawn up to show the factor that may increase the detection time. In Fig. 1, we can see the relationship between the processing time of five randomly selected infrared marine images and their downscaled versions. Accelerated proximal gradient (APG) algorithm is used to solve the convex optimization problem in detection. From Fig. 1, we can find out that it is hard to meet the real-time requirements of most applications for detecting targets even in a single frame. However, at the same time, we also find out that the processing time is closely related to the image size. As the image shrinks, the processing time falls sharply. In that case, the processing time could be decreased when the image size is downscaled.

In view of the hugeness of the computation amount in detection with all frames in video, we use image difference to extract information between frames and divide difference image into patches to further smaller the candidate region to be processed by LSMR. These methods restrict the processing time to a reasonable range. Figure 2 is the schematic diagram of the proposed method, which can be described in detail as the following steps.

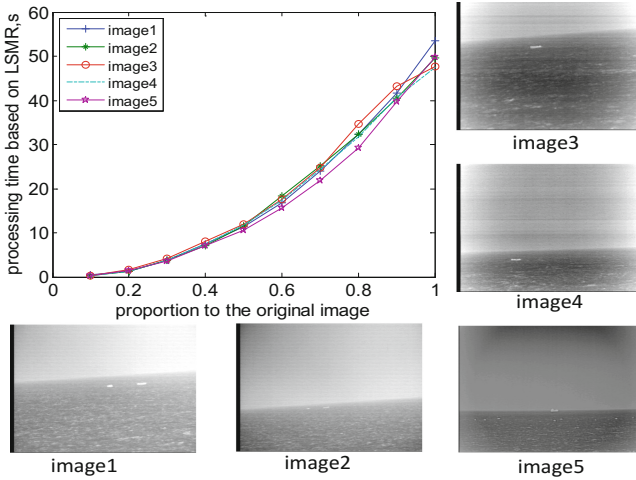


Fig. 1. Processing time of different sized images

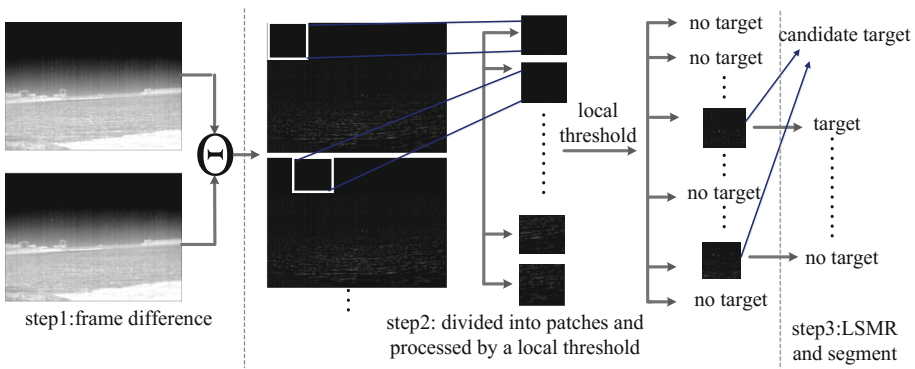


Fig. 2. Schematic diagram of the proposed method framework

Step 1: Frame difference

The frame difference is used to narrow the candidate region with little computation cost by utilizing information between adjacent frames. It could not only extract target position information from adjacent frames but also exclude the stationary disturbances for obtaining the candidate moving target regions which we are most interested in. Suppose C as the candidate area including moving point target, it can be obtained by (6).

$$C = |\Pi(t) - \Pi(t - 1)| \tag{6}$$

Here, $\Pi(t)$ is the t th frame, $\Pi(t - 1)$ is the frame before the $\Pi(t)$. $|\cdot|$ is the absolute operation.

Step 2: Divided into patches and processed by a local threshold

After obtaining the frame difference image C , we divide the image into partially overlapped patches, which further decreases the data to be processed at a time.

Obviously, the size of the patch and the overlapping area will influence the detection capability. Since point target is usually smaller than 10×10 pixels, the vertical and horizontal steps are set as P-10 to ensure no loss of targets with minimum redundant calculation. P is the width of patches, which will be discussed in the following section.

The patch is then processed by a local self-adaptive threshold to further narrow the detection region. The local self-adaptive threshold is described in (7).

$$C(x, y) = \begin{cases} 1 & \text{if } (C(x, y) \geq \alpha M_{\max}) \& (M_{\max} - M_{\min} > \beta) \\ 0 & \text{others} \end{cases} \quad (7)$$

where M_{\max} and M_{\min} are respectively the maximum and minimum pixel values of the local region. β is the threshold coefficient and set as 0.6, α is the threshold for judging the existing of the target. For the targets brightness is usually relative high in a local region, no target is expected to exist in the patches whose difference result of M_{\max} and M_{\min} is smaller than β . In the paper, we choose $\beta = 40$.

Step 3: LSMR and segment

Last but not least, we recover the target region T by LSMR in the left candidate region C and then use a simple segment method to extract targets. APG algorithm is used to solve the convex optimization problem of (3). It can achieve dramatically better performance among all the RPCA algorithms [6]. The algorithm is explained as follows:

Algorithm: RPCA via APG

Input: infrared difference image with target $C \in \mathbb{R}^{m \times n}$, λ

1. $B_0 = B_{-1} = 0; T_0 = T_{-1} = 0, t_0 = t_{-1} = 1; \mu_0 > 0; \bar{\mu} = \delta \mu_0; 0 < \eta < 1$.

While not converged do

2. $Y_k^B = B_k + \frac{t_{k-1} - 1}{t_k} (B_k - B_{k-1}), Y_k^T = T_k + \frac{t_{k-1} - 1}{t_k} (T_k - T_{k-1})$.

3. $G_k^B = Y_k^B - \frac{1}{2} (Y_k^B + Y_k^T - C)$.

4. $(U, S, V) = \text{svd}(G_k^B), B_{k+1} = US \frac{\mu_k}{2} [S]V'$.

5. $G_k^T = Y_k^T - \frac{1}{2} (Y_k^B + Y_k^T - C)$.

6. $T_{k+1} = S \frac{\lambda \mu_k}{2} [G_k^T]$.

7. $t_{k+1} = \frac{1 + \sqrt{4t_k^2 + 1}}{2}; \mu_{k+1} = \max(\eta \mu_k, \bar{\mu})$.

8. $k=k+1$.

End while

Output: $B = B_k; T = T_k$

4 Experimental Results and Analysis

4.1 The Performance Comparison with Different Weight of Sparse Error Term (λ) and Patch Size

To verify the capability of the proposed method, a real and representative cluttered marine surveillance video sequence with 48 frames is used to execute experiments. The resolution of the frames is 498×696 , and the moving point targets in the frames occupy about 5×5 pixels. All experiments were implemented by Matlab software on a PC with 2-GB RAM and 2.60-GHz Intel-i5 processor. Average detection time (ADT), SCRG, BSF are employed for objective evaluation. SCRG, BSF are same to the metrics presented in [12]. They are defined by (8) and (9). Where, S is the target amplitude and C is the clutter standard deviation within the original frame or the processed frame.

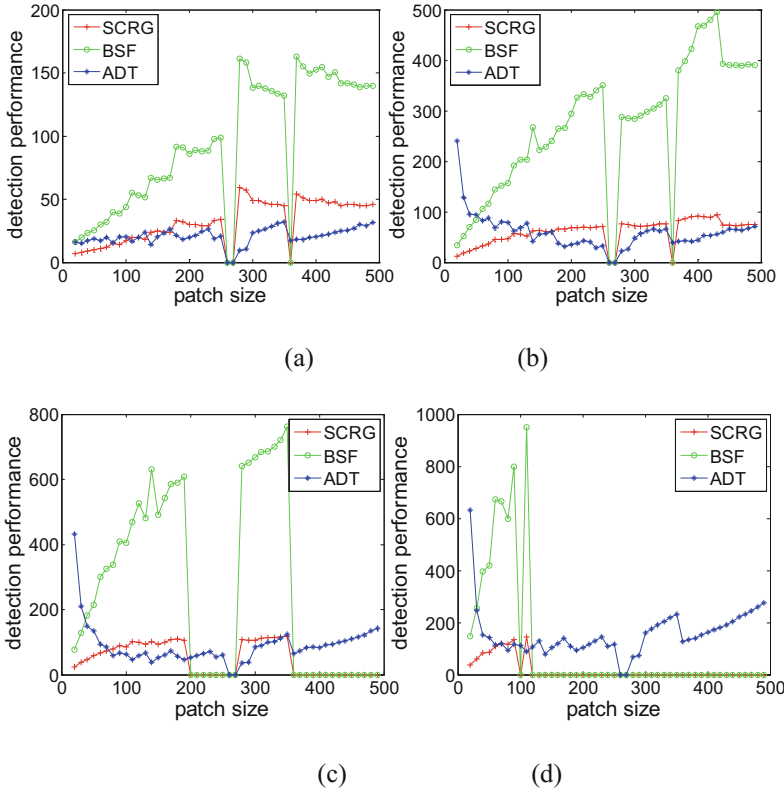


Fig. 3. Proposed performance comparison with different weight of sparse error term (λ). (a) Performance with different patch sizes ($\lambda = 0.1$); (b) Performance with different patch sizes ($\lambda = 0.2$); (c) Performance with different patch sizes ($\lambda = 0.3$); (d) Performance with different patch sizes ($\lambda = 0.4$).

$$\text{SCRG} = \frac{(S/C)_{out}}{(S/C)_{in}} \quad (8)$$

$$\text{BSF} = \frac{C_{in}}{C_{out}} \quad (9)$$

The patch sizes range from 20×20 pixels to 490×490 pixels with an increasing of 10 pixels. And the average results of all frames in the above mentioned video sequence are used for comparison.

Figure 3 shows the performance comparison with different weight of sparse error term (λ), which is an important parameter in APG algorithm. The ADT has been enlarged by 5, 10, 20 times in Fig. 3(b), (c) and (d) respectively for convenient observation. And it means the loss of targets when the SCRG and BSF are zero. In addition, it can be seen from Fig. 3 that, optimum detection time always exists with each determined λ , which is directly related to the patch size. Patches with too small or too large sizes always consume more time to be processed. The least detection time is obtained with patch size of 280×280 pixels in our experiment. The SCRG and BSF roughly increase with the patch size under determined λ , and increase with λ under determined patch size. Taking every aspect into consideration, we finally choose the optimum patch size as 280×280 pixels with $\lambda = 0.3$ in the proposed method. Its ADT is 3.53 s. And SCRG, BSF reach 107.0 and 641.1 respectively.

4.2 Evaluation Comparison

The detection capability of the proposed method is also compared with the result of the same algorithm with non-partitioned image and IPI model [6].

From Table 1, we can see that, the proposed method has better performance than other two methods under the same circumstance. Compared to IPI, our method reduces the detection time by 97.4% while improving SCRG, BSF by 180.1% and 389.4% respectively. What's more, the detection speed could be further accelerated by parallelism due to independence of patches in our method.

Table 1. Performance comparison of point target detection methods

Detection methods	IPI	Non-partitioned image	Proposed
SCRG	38.2	85.6	107.0
BSF	131.0	141.5	641.1
ADT(s)	136.67	21.29	3.53

By analyzing the results in above figure and table, we can see that: the proposed method could really improve the detection capability. And the reasons are as follows. First, the difference between adjacent frames and local threshold excluded a large amount of non-target patches before LSMR algorithm, and greatly decreased the computation load of LSMR. Since the frame difference and threshold process need

much less computation than LSMR, the total computation time could be reduced tremendously. Second, the LSMR is still effective in detection as long as the point targets are “sparse” in each patch. Third, the proposed method integrates methods of frame difference, local threshold, image segment and LSMR, which takes full advantage of multi-methods to exactly distinguish targets from complex background.

5 Conclusion and Future Work

In this paper, we proposed a patch-image based LSMR method for moving point target detection in infrared coastal surveillance. It is the first time the patches of frames are independently processed by LSMR algorithm, and the optimum patch size is also analyzed with metrics of ADT, SCRG and BSF. Experiment results show that the propose method prompts the detection speed while keeping a high detection capability.

However, detection results of the proposed method would be affected by dynamic noises which are also irregular. Thus, more research work should continually focus on suppressing noise and improving detection speed and accuracy in future. A comprehensive method which can both handle the detection velocity and accuracy parameters is essential for actual industrial applications.

Acknowledgments. This research was financially supported in part by the National Natural Science Foundation of China under Grant Nos. 61404083 and 51579143, the China Postdoctoral Science Foundation under Grant No. 2015M581527, the Innovation Program of Shanghai Municipality Education Commission under Grant No. 15ZZ080.

References

1. Filho, K.N., Romano, J.M.T.: Low-rank decomposition based on disjoint component analysis with applications in seismic imaging. *IEEE Trans. Comput. Imaging* **3**(2), 275–281 (2017). IEEE Press
2. Guo, K., Liu, L., Xu, X., Xu, D., Tao, D.: GoDec+ : fast and robust low-rank matrix decomposition based on maximum correntropy. *IEEE Trans. Neural Netw. Learn. Syst.* **18**(3), 636–653 (2016)
3. Hao, S., Ma, X., Fu, Z., Wang, Q., Li, H.: Landing cooperative target robust detection via low rank and sparse matrix decomposition. In: 2016 International Symposium on Computer, Consumer and Control, Xi’an, pp. 18–29 (2016)
4. Yang, X., Gao, X., Tao, D., Li, X., Han, B., Li, J.: Shape-constrained sparse and low-rank decomposition for auroral substorm detection. *IEEE Trans. Neural Netw. Learn. Syst.* **27**(1), 32–46 (2016)
5. Dai, Y., Wu, Y.: Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 499–508 (2017)
6. Gao, C.Q., Meng, D.Y., Yang, Y., Wang, Y.T., Zhou, X.F., Hauptmann, A.G.: Infrared patch-image model for small target detection in a single image. *IEEE Trans. Image Process.* **22**(12), 4996–5009 (2013)
7. Li, L., Li, H., Li, T., Gao, F.: Infrared small target detection in compressive domain. *Electron. Lett.* **50**(7), 510–512 (2014)

8. Zheng, C., Li, H.: Small infrared target detection based on harmonic and sparse matrix decomposition. *Opt. Eng.* **52**(6), 066401 (2013)
9. Wen, J.J., Xu, Y., Tang, J.H., Zhan, Y.W., Lai, Z.H., Guo, X.T.: Joint video frame set division and low-rank decomposition for background subtraction. *IEEE Trans. Circ. Syst. Video Technol.* **24**(12), 2034–2048 (2014)
10. Yang, C.W., Liu, H.P., Liao, S.Y., Wang, S.C.: Structured sparse coding method for infrared small target detection in video sequence. In: *International Joint Conference on Neural Networks (IJCNN)*, Beijing, pp. 1179–1184, July 2014
11. Kim, H., Paik, J.: Low-rank representation-based object tracking using multitask feature learning with joint sparsity. *Abstr. Appl. Anal.* 1–12 (2014)
12. Hilliard, C.I.: Selection of a clutter rejection algorithm for real-time target detection from an airborne platform. In: *Proceedings of the SPIE*, pp. 74–84, Orlando, USA (2000)