IEICE TRANS. INF. & SYST., VOL.E99–D, NO.2 FEBRUARY 2016

LETTER

Low-Rank and Sparse Decomposition Based Frame Difference Method for Small Infrared Target Detection in Coastal Surveillance

Weina ZHOU[†,††a)], **Xiangyang XUE**[††], *Nonmembers*, *and* **Yun CHEN**[††], *Member*

SUMMARY Detecting small infrared targets is a difficult but important task in highly cluttered coastal surveillance. The paper proposed a method called low-rank and sparse decomposition based frame difference to improve the detection performance of a surveillance system. First, the frame difference is used in adjacent frames to detect the candidate object regions which we are most interested in. Then we further exclude clutters by low-rank and sparse matrix recovery. Finally, the targets are extracted from the recovered target component by a local self-adaptive threshold. The experiment results show that, the method could effectively enhance the system's signal-to-clutter ratio gain and background suppression factor, and precisely extract target in highly cluttered coastal scene.

key words: *target detection, low-rank, sparse recovery, frame difference*

1. Introduction

Small infrared target detection plays an important role in infrared search and track system, and numerous methods have been put forward in the past decades to improve the detection accuracy and robustness. However, the detection performance is still poor in situations like coastal surveillance, which always has a low signal-to-clutter ratio (SCR) due to highly cluttered and complex background or spot-like noises [1]–[4]. Recently, low-rank and sparse matrices recovery theory (LRMR) [5], [6] has been proposed and applied in target detection and tracking. And it has been proved more effective compared with conventional baseline methods in some situations [7]. These LRMR methods can be mainly classified into two kinds. The first kind detects the small infrared target in a single frame [7]–[9], but its performance could degrade rapidly when the SCR is low. The other kind [5], [6], [10]–[12] makes use of the information of all the frames in a video sequence, the moving targets could then be more easily detected in a low-SCR scene. Unfortunately, the computation amount and time of this kind of method are huge, and it makes sense on the assumption that the targets are uniformly located at the scene [10].

In this paper, we detect the small infrared target in coastal surveillance video sequence using a method called low-rank and sparse decomposition based frame difference (LRSFD). Compared to other target detection method based on LRMR, our algorithm has three main advantages. First, profiting from getting information from multi-frames, the method could be effectively used in highly cluttered situations. Second, the method detects the target based on the result of frame difference, thus it not only has high stability but also requires much less computation amount and time than the existing methods based on video sequence. Third, the LRSFD method uses a local self-adaptive threshold and frame difference to improve the result of sparse recovery, so it could greatly enhance the signal-to-clutter ratio gain (SCRG) and the background suppression factor (BSF).

Manuscript received August 17, 2015.
Manuscript revised October 18, 2015.
Manuscript publicized November 11, 2015.

[†]The author is with the College of Information Engineering, Shanghai Maritime University, Shanghai, 201306 China.
[††]The authors are with the Fudan University, Shanghai, 200433 China.
a) E-mail: wnzhou@shmtu.edu.cn

DOI: 10.1587/transinf.2015EDL8186

2. LRMR Theory and Its Application in Small Target Detection

LRMR is a theory stretching from compressive sensing and sparse represent theory. It supposes an ideal image of size m x n as a low-rank matrix, whose rank-r is much smaller than its size, i.e. $r << min(m, n)$. But, in fact, the entries of the matrix are often corrupted by errors or noises, making the image not a low-rank matrix any more. To recover the low-rank character of the image, we regard every image as the combination of a low-rank matrix and a noise matrix, which could be represented in Eq. (1).

$$D = A + E. \tag{1}$$

In which, D is a corrupted image, A is an ideal low-rank image, E represents arbitrary noise and errors caused by various factors. Recently, it has been shown that, under surprisingly broad conditions, one can exactly recover A and E from D via Robust Principal Component Analysis (RPCA) to solve the following convex optimization problem [6].

$$\min_{A,E} = \|A\|_* + \lambda\|E\|_1, \quad subject \quad to \quad D = A + E. \tag{2}$$

Here, $\|\cdot\|_*$ represents the nuclear norm of a matrix, $\|\cdot\|_1$ denote the norm-1, and λ is a positive weighting parameter, which could be utilized to enhance the detection stability and suppress random noise. It is on this premise that the problem of small target detection could also be resolved as recovery of a low-rank matrix and a noise matrix.

Generally speaking, frames of coastal surveillance video consist of three components: the background, the small target, and various kinds of noises. As found by reference [7]–[9], patches of the background are commonly approximately context correlated with each other even though the pixel distance between two patches may be large in an image. Thus, background could be assumed as a low-rank matrix. Nevertheless, the foreground target is always small with respect to the whole image, which makes the target as a sparse noise matrix. Thus the small target detection task is intrinsically a typical problem of recovering a low-rank component (A) and a sparse component (E) from a data matrix (D). This model has been improved efficient in small target detection under the assumption that the random noise is i.i.d. and its Frobenius norm is smaller than some σ ($\sigma > 0$) [7]. That is

$$\|D - A - E\|_F \leq \sigma. \tag{3}$$

Where, $\|\cdot\|_F$ is the Frobenius norm. However, not all kinds of noises can satisfy the Eq. (3) in reality, thus we proposed LRSFD method to further improve detection rate of small infrared target in highly cluttered and complex background.

3. Sparse Coding Based Frame Difference (LRSFD) Method

In coastal surveillance, targets which we focus on usually have three essential characteristics: sparse, in motion, and bright in a local region. And as can been seen from Fig. 1, they would usually been affected by noises such as clutters of sea (marked as "F"), sensor noises (marked as "N"), spot-like noises (marked as "P"), coastline (marked as "L") and buildings or other kinds of objects in the coast (marked as "B"). Since not all the disturbances satisfy the definition of noise in LRMR, the recovery of sparse matrix is not always efficient in small target detection. Therefore, we classify these disturbances into three types according to characteristics of targets themselves: the stationary disturbances, random noises of i.i.d., and dynamic and irregular sea clutter. And methods of frame difference, LRMR and a local self-adaptive threshold are integrated successively to distinguish all types of noises from targets in different manners.

The LRSFD method of small target detection proposed in this paper is shown in Fig. 2. First, image difference between adjacent frames is executed to exclude the stationary disturbances for obtaining the candidate moving target regions we are most interested in. Second, accelerated proximal gradient (APG) solution is applied to the frame difference result to estimate the sparse target matrix E. Finally, targets are extracted by a local self-adaptive threshold by excluding left clutter noises based on their brightness characteristics. The algorithm is explained in detail as follows.

Frame difference: image difference between adjacent frames is a fundamental preprocessing measure to segment out the motion objects for further analysis. It is especially efficient to exclude stationary disturbances like spot-like noises, coastline and buildings or other kinds of objects in the coast. Since these noises are commonly complex and irregular, but relative stationary in a short period. The motion objects D can then be obtained by Eq. (4), which is the absolute value of the frame difference.

$$D = |I(t) - I(t-1)|. \tag{4}$$

Here, I(t) is the t_{th} frame, I(t-1) is the frame before the I(t). $|*|$ is the absolute operation. And D is the motion objects including potential targets.

APG based LRMR: As explained in Sect. 2, targets are sparse, and much different from low-rank background and random noise of i.i.d. Thus LRMR method is then used to recover target sparse matrix, extracting more accurate target regions E from D. Algorithm 1 (APG algorithm) is applied to solve the convex optimization problem of Eq. (2). It has a relative good accuracy and velocity among all the RPCA algorithms. In this paper, we choose $\lambda = 0.1$,

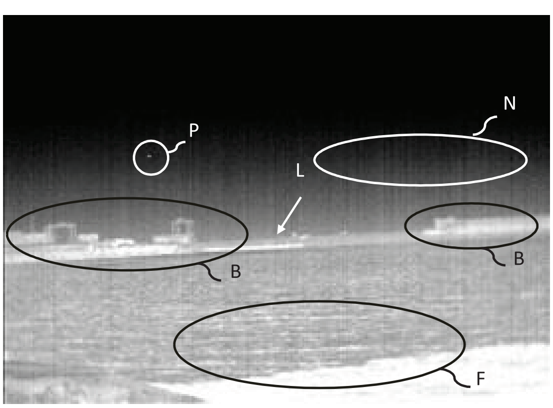

Fig. 1 A typical frame in costal surveillance video.

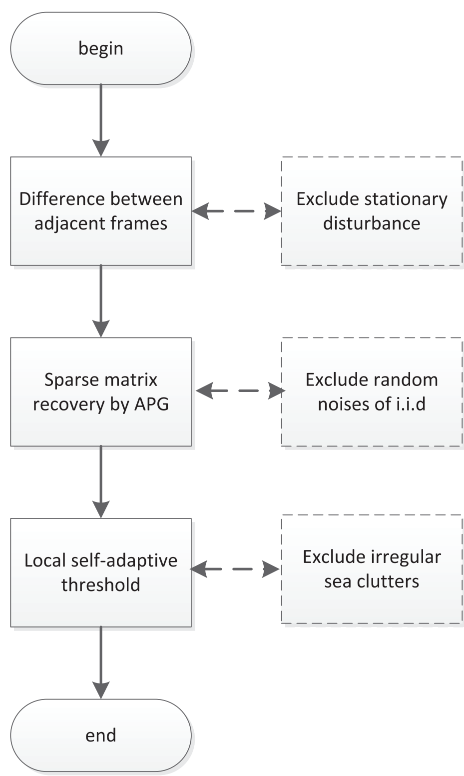

Fig. 2 The flow chart of LRSFD.

Algorithm 1: Robust PCA via Accelerated Proximal Gradient

Input: small infrared target frame $D \in R^{m \times n}$, λ

1. $A_0 = A_{-1} = 0$; $E_0 = E_{-1} = 0$, $t_0 = t_{-1} = 1$;
   $\mu_0 > 0$; $\overline{\mu} = \delta\mu_0$; $0 < \eta < 1$.
2. while not converged do
3. $Y_k^A = A_k + \frac{t_{k-1}-1}{t_k}(A_k - A_{k-1})$, $Y_k^E = E_k + \frac{t_{k-1}-1}{t_k}(E_k - E_{k-1})$.
4. $G_k^A = Y_k^A - \frac{1}{2}(Y_k^A + Y_k^E - D)$.
5. $(U, S, V) = svd(G_k^A)$, $A_{k+1} = US_{\frac{\mu_k}{2}}[S]V^T$.
6. $G_k^E = Y_k^E - \frac{1}{2}(Y_k^A + Y_k^E - D)$.
7. $E_{k+1} = S_{\frac{\lambda\mu_k}{2}}[G_k^E]$.
8. $t_{k+1} = \frac{1+\sqrt{4t_k^2+1}}{2}$; $\mu_{k+1} = max(\eta\mu_k, \overline{\mu})$.
9. $k = k + 1$.
10. end while

output: $A = A_k$; $E = E_k$

$\mu_0 = \|D\|_2$, $\delta = 10^{-5}$, $\eta = 0.99$.

Local self-adaptive threshold: To further eliminate the disturbances of sea clutter other than the random noise, we finally use a local self-adaptive threshold to extract the target by using its brightness characteristics. That is, the target whose size may vary from 2x2 to 16x16 pixels, is always the brightest in a local region. In this paper, the size of the local region is set as 50x50 pixels, and the horizontal and vertical sliding step are set as 10 pixels, which are proved efficient in detection [7]. The local self-adaptive threshold is described in Eq. (5):

$$E(x,y) = \begin{cases} 1 & if\ (E(x,y) \geq \alpha M_{max}) \\ & \quad \&\ (M_{max} - M_{min} > \beta) \\ 0 & others \end{cases} \tag{5}$$

Where M_{max} is the maximum pixel value of the local region. And M_{min} is the minimum pixel value of the local region which is bigger than zero. α is the threshold coefficient and set as 0.6, β is the threshold for judging the existing of the target. And no target is expected to exist in the frame when the difference value of the M_{max} and M_{min} is smaller than β. In the paper, we choose $\beta = 40$.

4. Experiment Work

To demonstrate the effectiveness of the LRSFD method, experiments are executed using a real and representative highly cluttered coastal surveillance video sequence with 48 frames. And all experiments were implemented by Matlab software on a PC with 2-GB RAM and 2.60-GHz Intel-i5 processor.

Figures 3 and 4 are randomly selected two adjacent frames in the video sequence, and their corresponding gray distributions. From Fig. 4, we can see that, the SCR of the frames is very low. The image difference result of the two adjacent frames is shown in Fig. 5. We can find out that those disturbances like spotlike noises, coastline and buildings are successfully eliminated. Figure 6 is the recovered target and background result by LRMR method, the vast majority of the sea clutter and random noises have been excluded as low-rank background. Finally, by applying the local threshold in recovered sparse image, the target could be obviously distinguished in Fig. 7.

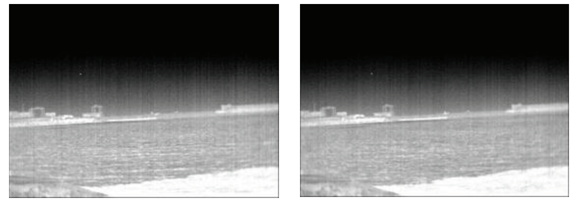

Fig. 3 Two adjacent frames in a real video sequence.

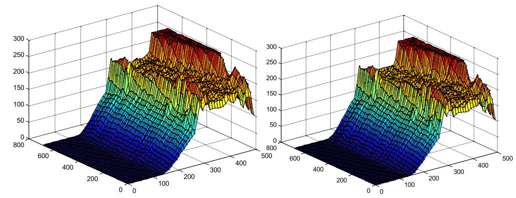

Fig. 4 The gray distributions of the frames in Fig. 3.

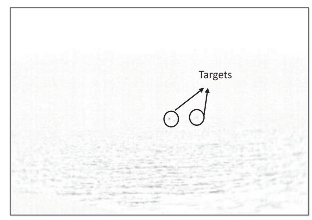

Fig. 5 The image difference result of the adjacent frames (color reversed for observation).

Fig. 6 The recovered target and background component by LRMR (color reversed for observation).

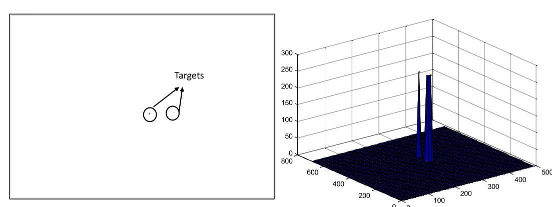

Fig. 7 The target image (color reversed for observation) and its gray distribution.

The detection capability of our proposed method is also compared to other baseline methods like TopHat [13], Maxmean, Maxmedian [14], and Infrared patch-image (IPI) model [7] which is a representative method based on LRMR. The SCRG and BSF are employed for objective evaluation.

Table 1 Comparison of small target detection methods.

Detection methods	Tophat	Maxmean (5*5)	Maxmedian (5*5)	IPI	LRSFD
SCRG	16.53	28.51	35.95	38.16	88.60
BSF	18.94	65.77	82.05	130.96	150.48

The SCRG and BSF are two common filter metrics [15]. They are respectively defined by Eqs. (6) and (7).

$$SCRG = \frac{(S/C)_{out}}{(S/C)_{in}}. \tag{6}$$

$$BSF = \frac{C_{in}}{C_{out}}. \tag{7}$$

Where, S is the target amplitude and C is the clutter standard deviation within the original frame or the processed frame. And the average results of all frames in the above mentioned video sequence are used for comparison.

From the Table 1, we can see that, the LRSFD method has better performance than other detection methods under the same circumstance.

Never mind of the influence of other process in PC, the total processing time of the proposed method is 41.35s. Although the processing time is less than that of the traditional LRMR method of video performed in the same PC environment, it is far from meeting the realtime requirement. The reason mainly lies in the non-optimized APG algorithm in LRMR, which occupies more than 99% of the total computation. Thus what we will do next is to optimize the APG algorithm and propose a FPGA based acceleration scheme.

5. Conclusion

In this paper, we present the LRSFD method for infrared small target detection in highly cluttered coastal surveillance video sequence. It is the first time that the sparse image recovery has been integrated with frame difference and local self-adaptive threshold. The proposed method takes full advantage of the integrated algorithm according to the different types of noises, and obtains better performance in detection capability compared with TopHat, Maxmean, Maxmedian and typical LRMR based method. Thus the method is expected to be effectively used in infrared surveillance system for coastal safety in future.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant No. 51279099 and 61404083, the Shu Guang Project of Shanghai Municipal Education Commission and Shanghai Education Development Foundation under Grant No. 12SG40, the Innovation Program of Shanghai Municipality Education Commission under Grant No. 15ZZ080.

References

[1] C.Q. Gao, T.Q. Zhang, and Q. Li, "Small infrared target detection using sparse ring representation," IEEE Aerosp. Elecron. Syst. Mag., vol.27, no.3, pp.21–30, March 2012.

[2] Y.F. Gu, C. Wang, B.X. 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, July 2010.

[3] S.X. Qi, J. Ma, C. Tao, C.C. Yang, and J.W. Tian, "A robust directional saliency-based method for infrared small-target detection under various complex backgrounds," IEEE Geosci. Remote Sens. Lett., vol.10, no.3, pp.495–499, May 2013.

[4] C.L.P. Chen, H. Li, Y.T. Wei, T. Xia, and Y.Y. Tang, "A local contrast method for small infrared target detection," IEEE Geosci. Remote Sens. Lett., vol.52, no.1, pp.574–581, Jan. 2014. DOI: 10.1109/TGRS.2013.2242477

[5] J. Wright, A. Ganesh, S. Rao, Y.G. Peng, and Y. Ma, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization," Advances in Neural Information Processing Systems 22, pp.2080–2088, Vancouver, Dec. 2009.

[6] E.J. Candès, X.D. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," Journal of the ACM, vol.58, no.3, pp.1–37, June 2011. DOI: 10.1145/1970392.1970395

[7] C.Q. Gao, D.Y. Meng, Y. Yang, Y.T. Wang, X.F. 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, Sept. 2013. DOI: 10.1109/TIP.2013.2281420

[8] L. Li, H. Li, T. Li, and F. Gao, "Infrared small target detection in compressive domain," Electron. Lett., vol.50, no.7, pp.510–512, March 2014. DOI: 10.1049/el.2014.0180

[9] C. Zheng and H. Li, "Small infrared target detection based on harmonic and sparse matrix decomposition," Opt. Eng., vol.52, no.6, pp.066401-1-066401-10, 2013. DOI: 10.1117/1.OE.52.6.066401

[10] J.J. Wen, Y. Xu, J.H. Tang, Y.W. Zhan, Z.H. Lai, and X.T. Guo, "Joint video frame set division and low-rank decomposition for background subtraction," IEEE Trans. Circuit Syst. Video Technol., vol.24, no.12, pp.2034–2048, Dec. 2014.

[11] C.W. Yang, H.P. Liu, S.Y. Liao, and S.C. Wang, "Structured sparse coding method for infrared small target detection in video sequence," 2014 International Joint Conference on Neural Networks (IJCNN), pp.1179–1184, Beijing, July 2014.

[12] H. Kim and J. Paik, "Low-rank representation-based object tracking using multitask feature learning with joint sparsity," Abstract and Applied Analysis, vol.2014, Article ID.147353, pp.1–12, 2014.

[13] R. Zhou, B. Liu, and J. Teng, "Morphology based particle filtering algorithm for point target tracking," 2013 25th Chinese Control and Decision Conference (CCDC 2013), pp.3760–3764, Guiyang, May 2013. DOI: 10.1109/CCDC.2013.6561603

[14] S.D. Deshpande, M.H. Er, R.Venkateswarlu, and P. Chan, "Max-mean and max-median filters for detection of small-targets," Proc. SPIE, no.3809, pp.74–83, Denvor, July 1999. DOI: 10.1117/12.364049

[15] C.Q. Gao, J.W. Tian, and P. Wang, "Generalised-structure-tensor-based infrared small target detection," Electron. Lett., vol.44, no.23, pp.1349–1351, Nov. 2008. DOI: 10.1049/el:20081781