This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

A Local Contrast Method for Small Infrared Target Detection

C. L. Philip Chen, *Fellow, IEEE*, Hong Li, *Member, IEEE*, Yantao Wei, Tian Xia, and Yuan Yan Tang, *Fellow, IEEE*

***Abstract*—Robust small target detection of low signal-to-noise ratio (SNR) is very important in infrared search and track applications for self-defense or attacks. Consequently, an effective small target detection algorithm inspired by the contrast mechanism of human vision system and derived kernel model is presented in this paper. At the first stage, the local contrast map of the input image is obtained using the proposed local contrast measure which measures the dissimilarity between the current location and its neighborhoods. In this way, target signal enhancement and background clutter suppression are achieved simultaneously. At the second stage, an adaptive threshold is adopted to segment the target. The experiments on two sequences have validated the detection capability of the proposed target detection method. Experimental evaluation results show that our method is simple and effective with respect to detection accuracy. In particular, the proposed method can improve the SNR of the image significantly.**

***Index Terms*—Derived kernel (DK), infrared (IR) image, local contrast, signal-to-noise ratio (SNR), target detection.**

I. Introduction

IT IS well known that target detection has found wide applications in such areas as remote sensing, surveillance, aerospace, and so on [1]–[4]. As an important technique in target detection, infrared (IR) imaging has received a lot of attentions [3], [5]–[7]. Detecting a small IR target of unknown position and velocity at low signal-to-noise ratio (SNR) is an important issue in IR search and track system, which is necessary for military applications to warn from incoming small targets from a distance, such as enemy aircraft and helicopters [3], [6]–[11]. The target immersed in heavy noise and clutter background presents as spotlike feature. The gray value of a target is usually higher than that of its immediate background in IR image and is not spatially correlated with that of the local neighborhood. Due to the effects of inherent sensor noise and natural factors, there exist some high gray regions in the IR image (e.g., irregular sunlit spot). These effects make the target detection more difficult. In order to detect a small target effectively, various algorithms have been developed in the past few decades [12]–[22].

Conventional small target detection methods such as the median subtraction filter [12], top-hat filter [13], and max-mean/max-median filter [14] are widely used to reduce the background clutters. Genetic algorithm and morphological filters are also combined to detect the small targets [21]. Image information generated by wavelet in different scales supplies the feature information that could distinguish the target and background. Based on this idea, many methods of small target detection in clutter background using wavelet are proposed [23]. Fractal approaches have been widely used for IR target detection [24]. However, they are sometimes inaccurate and time consuming. Sun *et al.* propose target detection methods based on center-surround difference [19], [20]. In addition, some methods based on statistical regression have been proposed [17], where the complex background clutter is predicted and eliminated by a regression model. Recently, a small target detection method based on sparse ring representation (SRR) has been proposed [25]. SRR is an effective graphical structure which can describe the difference between the background and targets. There are still many other algorithms for target detection, such as methods based on manifold learning, empirical mode decomposition, and neural network [8], [15], [26]. However, small target detection under complex background is still a challenge.

In order to design a small target detection method, it is inspiring to imitate the computational architecture of biological vision [27]–[29]. Recently, bioinspired hierarchical algorithms such as the hierarchical model and X [28], and derived kernel (DK) [30], [31] have attracted much attention. Among these methods, the DK model studies the theoretical properties of hierarchical algorithms from a mathematical perspective. Consequently, the DK model can guide the design of a new algorithm. Noticed that the small target has a signature of discontinuity with its neighboring regions and concentrates in a relatively small region, which can be considered as a homogeneous compact region, and the background is consistent with its neighboring regions [32]. Consequently, we conceive that, after some target enhancement operation, the local region whose contrast is larger than the given threshold in some scale may be a position where the target appears. With these considerations

Manuscript received September 7, 2012; revised December 12, 2012 and January 1, 2013; accepted January 2, 2013. This work was supported in part by the National Natural Science Foundation of China under Grant 61075116, by the Natural Science Foundation of Hubei Province under Grant 2009CDB387, by the Chinese National Basic Research 973 Program under Grant 2011CB302800, by the Macau Science and Technology Development Fund under Grant 008/2010/A1, by the Multi-Year Research of University of Macau under Grants MYRG205(Y1-L4)-FST11-TYY and MYRG187(Y1-L3)-FST11-TYY, and by the Start-up Research of University of Macau under Grant SRG010-FST11-TYY.

C. L. P. Chen, T. Xia, and Y. Y. Tang are with the Faculty of Science and Technology, University of Macau, Macau, China (e-mail: Philip.Chen@ieee.org; yb17404@umac.mo; yytang@umac.mo).

H. Li is with the School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: hongli@mail.hust.edu.cn).

Y. Wei is with the College of Information Technology, Journalism and Communications, Central China Normal University, Wuhan 430079, China. He was with the Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: yantaowei@mail.ccnu.edu.cn).

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

Digital Object Identifier 10.1109/TGRS.2013.2242477

0196-2892/$31.00 © 2013 IEEE

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

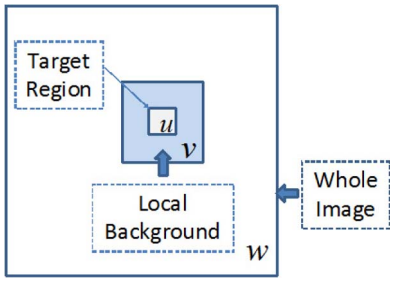


Fig. 1. Multiple windows.

in mind, a small target detection method inspired by the vision mechanism and the DK model has been designed in this paper. The contributions of this paper can be summarized as follows.

1) An effective contrast measure inspired by the biological visual mechanism is proposed. It can enhance the target and suppress background clutter simultaneously.
2) A small IR target detection algorithm is designed. Experimental results show that it is simple and effective with respect to detection accuracy. In particular, it can improve the SNR of the image significantly.
3) A mathematical framework has been given. This framework can provide a useful foundation for understanding the proposed method and guiding future developments.

This paper is organized as follows. In Section II, we introduce the target detection algorithm based on local contrast. Section III presents the theoretical analysis of the proposed method. Section IV contains experiments on some IR image sequences to demonstrate the effectiveness of the proposed method. Finally, we conclude this paper in Section V.

II. Target Detection Based on LCM

In this paper, we introduce the basic concept of the local contrast measure (LCM) inspired by the robust properties of the human vision system (HVS) and DK model. Vision scientists conclude that contrast is the most important quantity encoded in the streams of visual system. This fact is true in the whole target detection process [3], [33]. In addition, a multiscale representation is supported by vision research since the effect of pattern adaptation cannot be explained by a single resolution theory. After the multiscale representation, the visual system can tune into an appropriate size sensitive to the spatial extent.

A. LCM

LCM describes a pixel (position) by generating a signal value. It can be applied directly to grayscale images. As mentioned before, the region where the target is located is different from its neighborhood. This means that the target is conspicuous in a local region. Taking this fact into account, we use a hierarchical architecture inspired by the DK model to measure the local contrast [30]. As shown in Fig. 1, u denotes the target region, and the area between u and v is the local background region. Moreover, the whole image (frame) is denoted by w. In this case, the windows v can move on the w.

Based on the aforementioned preparation, we can find that different image patches of size v can be obtained by moving the window v on the whole image. Furthermore, u can move on

Fig. 2. Image patch obtained by sliding window can be divided into nine cells.

v. Consequently, the obtained image patches can be given by nine cells (see Fig. 2). Note that the central cell denoted by "0" is a region where the target could appear. The gray mean of the ith cell is denoted by m_i $(i = 1, 2, \ldots, 8)$, i.e.,

$$m_i = \frac{1}{N_u} \sum_{j=1}^{N_u} I_j^i \tag{1}$$

where N_u is the number of the pixels in the ith cell and I_j^i is the gray level of the jth pixel in the ith cell. In this paper, the width and the height of the window v are both three times those of window u. Consequently, the contrast between the central cell and the ith surrounding cell is defined by

$$c_i^n = \frac{L_n}{m_i} \tag{2}$$

where L_n represents the maximum of the gray value of the central cell in the nth image patch of size v.

In addition, only the visual brightness of the small target in IR image is usually greater than that of their neighborhoods, despite that the discrimination between the target and the area of the neighborhood is usually small. In order to enhance the target, the LCM is defined as follows:

$$C_n = \min_i L_n \times c_i^n = \min_i L_n \times \frac{L_n}{m_i} = \min_i \frac{L_n^2}{m_i}. \tag{3}$$

This definition means that the larger the C_n is, the more likely a target appears. If $(L_n/m_i) = (L_{n'}/m_{i'})$, then the larger of L_n and $L_{n'}$ is more likely to correspond to a target.

Algorithm 1 LCM computation

Input: Examined image patch.
Output: C_n.
1: Divide the examined image patch into 3×3 cells.
2: Compute the maximum of the central cell

$$L_n = \max_{j=1,2,\ldots,N_0} I_j^0.$$

3: **for** $i = 1, 2, \ldots, 8$ **do**

$$m_i = \frac{1}{N_u} \sum_{j=1}^{N_u} I_j^i.$$

4: **end for**
5: $C_n = \min_i(L_n^2/m_i)$.
6: Replace the value of the central pixel with the C_n.

In summary, the method to compute the LCM is shown in the Algorithm 1, where I_j^0 is the gray level of the jth pixel in

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

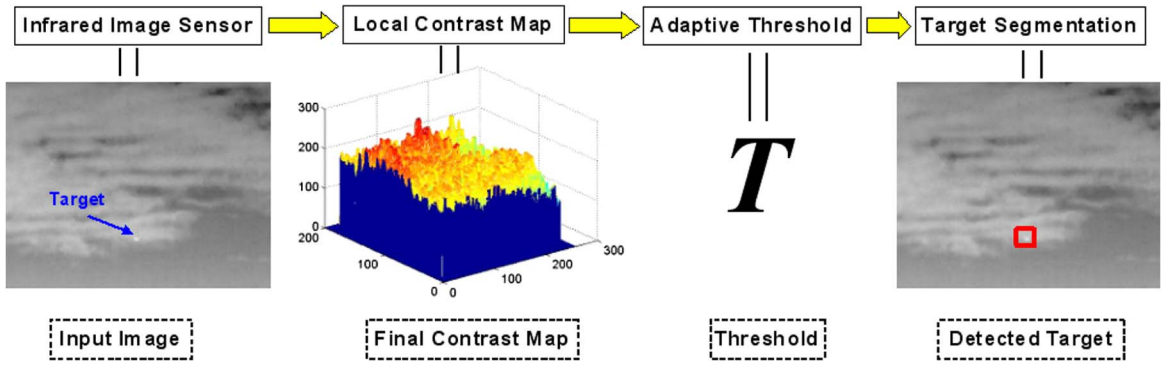


Fig. 3. Proposed target detection system.

the central cell, and N_0 is the number of the pixels of it. In this way, we can obtain a contrast map, $\mathcal{C}$, for each image on a given scale.

B. Multiscale LCM

As a matter of fact, the sizes of small targets are changing. In the ideal case, the size of window u should be the same as the target size. In order to deal with this problem, an effective way can be given in Algorithm 2, where $\hat{\mathcal{C}}$ is the final contrast map, l_{max} is the largest scale, and p_1 and q_1 are the number of rows and columns of the contrast map, respectively. In this algorithm, maximum pooling operation on multiple scales has been used. In this way, the scale invariance can be achieved.

Algorithm 2 Multiscale LCM

Input: Given frame.
Output: $\hat{\mathcal{C}}$.
1: Enough scales of the window are given.
2: **for** $l = 1 : l_{max}$ **do**
3: Compute $\mathcal{C}^l$ according to Algorithm 1.
4: **end for**
5: **for** $p = 1 : p_1$ **do**
6: **for** $q = 1 : q_1$ **do**

$$\hat{\mathcal{C}}_{p,q} = \max_{l=1,2,\ldots,l_{max}} \mathcal{C}^l_{p,q}$$

7: **end for**
8: **end for**

In addition, a problem that we must consider is how to decide the size of u. If the size of u is too small, the potential target may be masked by the background. On the other hand, if the size of u is too large, the operation will consume more computation. The small target in IR image often occupies several pixels. Society of Photo-Optical Instrumentation Engineers defines the small target from the perspective of imaging. A small target is defined to have a total spatial extent of less than 80 pixels. This classification includes: point source targets, small extended targets, and clusters of point source targets and small extended targets [34]. If the size of the input image is 256×256, the small target will occupy less than 0.15% of it. This standard is usually independent of the size of the input image. Consequently, the appropriate size of u is not larger than $9 \times 9 = 81$ pixels in the experiments.

C. LCM-Based Small Target Detection Method

It is easy to find that the target region has a signature of discontinuity with its neighboring regions and concentrates in a relatively small region, which can be considered as a homogeneous compact region, and the background is consistent with its neighboring regions. Consequently, we can conceive that, if the final contrast map is obtained, it is likely that the most salient point in the scene is a target. Based on this fact, the LCM-based target detection method can be described in Algorithm 3, where $\overline{I}^c$ and σ_{I^c} are the mean and the standard deviation of the final contrast map, respectively. In practice, k usually ranges from 3 to 5. In order to intuitively show the proposed method, a target detection system is given in Fig. 3.

Algorithm 3 Target detection method

Input: One frame.
Output: Target position.
1: Obtain $\hat{\mathcal{C}}$ according to Algorithm 2.
2: Compute the threshold according to

$$T = \overline{I}^c + k \times \sigma_{I^c}. \tag{4}$$

3: Segment targets from the background according to T.

III. Algorithm Analysis

A. Detection Ability Analysis

From the definition, we can find that the LCM can enhance the target. If the current location is the region where the target appears (marked by a solid arrow in Fig. 4), then the local maximum of the current region (central cell) is larger than the means of the surrounding cells. Therefore, the obtained local contrast is larger than the local maximum of the current region. In this way, the target can be enhanced. On the contrary, if the current location is the background (marked by a hollow arrow in Fig. 4), then the local maximum of the current region (central cell) may be smaller than the means of the surrounding cells. Therefore, the obtained local contrast may be smaller

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

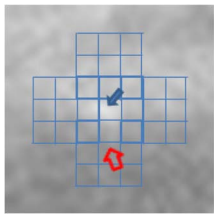


Fig. 4. Illustration of the target enhancement and background suppression.

Fig. 5. Detection result of one frame in sequence P.

than the local maximum of the current region. In this way, the background can be suppressed.

Let (x_0, y_0) be the center pixel of the target. Therefore, there is an image patch of size v centered on (x_0, y_0), which can be found in Fig. 2. The contrast of the pixel belonging to the target is given by

$$C_j^t = \min_i \frac{\left(L_j^t\right)^2}{m_i} \tag{5}$$

where L_j^t is the local maximum of the target region. This means that a larger L_j^t will lead to larger C_j^t. It is easy to find that

$$\min_i \frac{L_j^t}{m_i} > 1 \tag{6}$$

since the target is brighter than the background. The contrast of the pixel belonging to the background is obtained by

$$C_j^b = \min_i \frac{\left(L_j^b\right)^2}{m_i} \tag{7}$$

where L_j^b is the local maximum in the background region. In most cases, we have

$$\min_i \frac{L_j^t}{m_i} \leq 1. \tag{8}$$

The contrast in the target region is higher than that in the background region. Consequently, we can use the contrast map to detect the target. This reveals that the proposed method considers the problems of target enhancement and background suppression simultaneously.

B. Mathematical Framework

The proposed method is inspired by the DK model [30], [31], [35], [36]. For further reading, we refer to the detailed description provided by Smale *et al.* in [30]. We will give the mathematical framework of the proposed method hereinafter.

1) Notation and Preliminaries: Let patches u, v, and w in $\mathbb{R}^2$ $(u \subset v \subset w)$ be pieces of the domain on which the images (or image patches) are defined. Given patch w, an image $\boldsymbol{f}$ is defined on w. In the same way, the image patches $\boldsymbol{f}^u$ and $\boldsymbol{f}^v$ can be defined respectively on u and v.

We further assume that we are given an image set on w, denoted by $\mathcal{I}_w$ (that contains image $\boldsymbol{f}_i$, $i = 1, 2, \ldots$), and define $\mathcal{I}_u$ and $\mathcal{I}_v$ as all the collections of image patches $\boldsymbol{f}_i^u$ defined on u $(i = 1, 2, \ldots)$ and $\boldsymbol{f}_j^v$ defined on v $(j = 1, 2, \ldots)$, respectively. Let $\mathcal{H}_u$ be a set of transformations that consists of finite elements

$$\mathcal{H}_u = \{h_0^u, h_1^u, \ldots, h_8^u\} \tag{9}$$

where $h^u : u \to v$, and similarly, $\mathcal{H}_v$ is defined by $h^v : v \to w$. Similarly

$$\mathcal{H}_v = \left\{h_1^v, h_2^v, \ldots, h_{|\mathcal{H}_v|}^v\right\} \tag{10}$$

where $|\mathcal{H}_v|$ is the cardinality of the $\mathcal{H}_v$. Without loss of generality, we only consider translations here. The translation can be thought of as selecting a receptive field that "looks" at different regions of the image.

In this paper, one key property is given as follows [30].

Property 1: Let $\boldsymbol{f} \in \mathcal{I}_w$ and $h^v \in \mathcal{H}_v$, where each h^v corresponds to a position of a sliding window of size v. Then, we can obtain an image patch $\boldsymbol{f}^v \in \mathcal{I}_v$ by moving the sliding window on $\boldsymbol{f}$, i.e., $\boldsymbol{f} \circ h^v = \boldsymbol{f}^v \in \mathcal{I}_v$, where $\circ$ is the translation action of h^v on $\boldsymbol{f}$.

This property indicates that $\boldsymbol{f} \circ h^v$ is an image patch obtained by restricting $\boldsymbol{f}$ to a given patched domain. $\boldsymbol{f}^v$ can be also considered as an image patch obtained by scanning $\boldsymbol{f}$ with a sliding window of size v. Similarly, we have

$$\boldsymbol{f}^u = \boldsymbol{f}^v \circ h^u. \tag{11}$$

Note that, in the definition given above, the sliding window of size v moves pixel by pixel. Consequently, the obtained image patches are overlapping.

2) Multiscale Local Contrast Map: We give the computational procedure of the multiscale local contrast map as follows.

1) The computation begins with a dissimilarity measure on scale l denoted by $C_i^l(x_a, y_b)$, where (x_a, y_b) is the position to be considered. Let D be a function to measure the dissimilarity between the central cell and the ith surrounding cell. Then, we have

$$C_i^l(x_a, y_b) = D\left(\boldsymbol{f}^v \circ h_0^u, \boldsymbol{f}^v \circ h_i^u\right) \tag{12}$$

where $\boldsymbol{f}^v$ is centered on (x_a, x_b). Note that D can be any dissimilarity measure. In this paper, D takes the following form for the purpose of target detection:

$$D\left(\boldsymbol{f}^v \circ h_0^u, \boldsymbol{f}^v \circ h_i^u\right) = \frac{\left(\max\left(\boldsymbol{f}^v \circ h_0^u\right)\right)^2}{\operatorname{mean}\left(\boldsymbol{f}^v \circ h_i^u\right)}. \tag{13}$$

TABLE I
COMPARISON OF SEVERAL TARGET DETECTION METHODS

Method	Top-hat	AGADMM	Proposed method
DR	76.67	83.33	**86.67**
FA	0.7333(44/60)	0.3500(21/60)	**0.2833(17/60)**

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE II
SNR OF THE ORIGINAL IMAGES AND THE CORRESPONDING CONTRAST MAPS

Image	1	2	3	4	5	6	7	8
Original image	1.7952	0.9885	3.2887	1.7181	0.7368	1.0784	1.6503	1.2599
Contrast map	8.8165	6.1022	8.2400	5.5070	5.8453	5.9988	4.9717	4.3849

Therefore, we have

$$C_i^l(x_a, y_b) = \frac{\left(\max\left(\boldsymbol{f} \circ h_n^v \circ h_0^u\right)\right)^2}{\text{mean}\left(\boldsymbol{f} \circ h_n^v \circ h_i^u\right)} = \frac{\left(\max\left(\boldsymbol{f}_n^v \circ h_0^u\right)\right)^2}{\text{mean}\left(\boldsymbol{f}_n^v \circ h_i^u\right)} = \frac{L_{a,b}^2}{m_i} \tag{14}$$

where h_n^v is the nth translation centered on (x_a, x_b), h_i^u is the ith translation on v, $i = 1, 2, \ldots, 8$, and C_i^l is the dissimilarity between the central cell and the ith surrounding cell on scale l. The LCM at (x_a, y_b) can be obtained by

$$C^l(x_a, y_b) = \min_{i=1,2,\ldots,8} C_i^l(x_a, y_b). \tag{15}$$

2) The final LCM of the position (x_a, y_b) is defined as

$$\mathcal{C}(x_a, y_b) = \max_l C^l(x_a, y_b). \tag{16}$$

It is easy to find that each position in the input image $\boldsymbol{f} \in \boldsymbol{I}_w$ corresponds to a final LCM. Consequently, the input image can be represented by a final LCM map. In addition, the final LCM can also be called the neural response of $\boldsymbol{f}$ at a given position.

3) General Framework of the Proposed Method: As mentioned earlier, the proposed method has multilayer architecture. Similarly, it consists of two basic operations: comparing and pooling. A general definition can be given as follows.

1) Comparing

The comparing operation measures the dissimilarity (or similarity) between the target region and background regions. This operation is carried out at different scales. The dissimilarity is given by a function defined as follows:

$$S_l = \Phi_l(t) \tag{17}$$

where t is the target region, Φ_l is the comparing operation at scale l, and S_l is the dissimilarity measure. In the proposed method, the Φ_l is defined by (3). The Φ_l defined by us is more robust to complex background than that defined by the method proposed in [32] since the surrounding region is analyzed more finely.

2) Pooling

The pooling operation can be described as a function that summarizes the input with a single value. The inputs to the pooling operation will, in general, have different lengths. We can define a pooling function to be a function

$$\Psi : R^* \to R \tag{18}$$

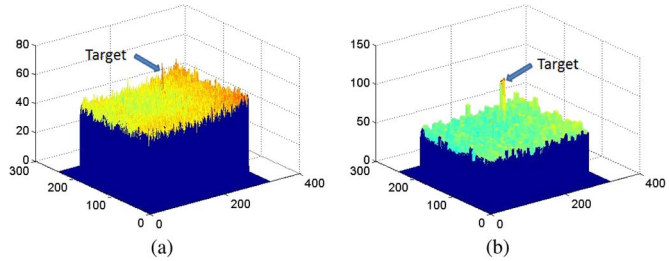


Fig. 6. Three-dimensional projection of (a) the original image in sequence P and (b) the corresponding contrast map.

TABLE III
COMPARISON OF SEVERAL TARGET DETECTION METHODS (GAUSSIAN WHITE NOISE OF VARIANCE 0.00001)

Method	Top-hat	AGADMM	Proposed method
DR	73.33	81.67	**86.67**
FA	0.75(45/60)	0.5000(30/60)	**0.2833(17/60)**

TABLE IV
COMPARISON OF SEVERAL TARGET DETECTION METHODS (GAUSSIAN WHITE NOISE OF VARIANCE 0.00005)

Method	Top-hat	AGADMM	Proposed method
DR	68.33	76.67	**83.33**
FA	0.9000(54/60)	0.5833(35/60)	**0.3333(20/60)**

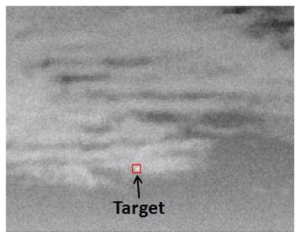


Fig. 7. Detection result of one frame in sequence S.

where R^* is the input data set. For example, the average pooling function can be described by

$$R = \frac{1}{n}\sum_{l=1}^{n} S_l. \tag{19}$$

In the literatures, the pooling operation is usually the maximum pooling operation. The maximum pooling will lead to scale invariance [28]. Consequently, the response at the candidate target is defined as

$$R = \Psi\left(\Phi(t)\right) = \Psi(S_l, S_2, \cdots) \tag{20}$$

where $\Phi(t) = (\Phi_1, \Phi_2, \cdots)$.

The definition given above allows us to integrate some of the existing target detection algorithms in practice into our framework [32], [37]. We believe that this framework will provide a useful foundation for guiding future developments.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE V
SNR OF THE ORIGINAL IMAGES AND THE CORRESPONDING CONTRAST MAPS

Image	1	2	3	4	5	6	7	8
Original image	2.4818	2.3411	2.8254	2.0161	2.5278	2.3738	2.2820	1.5488
Contrast map	3.3229	2.8277	3.0702	2.3413	3.0004	2.4744	3.1086	2.1371

C. Relationship Between Proposed Method and DK

The proposed method and DK model both imitate the visual information processing in primates. However, we can find some differences between them. The differences which are determined by the purpose of target detection can be summarized as follows.

1) It is easy to find that we only compare the central cell with each surrounding cell. Moreover, we do not compute the dissimilarity among surrounding cells. This is determined by the assumption that the target appears at the center of the receptive field while the surroundings are backgrounds.
2) The DK is constructed recursively using an alternating process of two basic operations. However, in consideration of the computation time, we do not define the LCM in this manner.
3) Unlike the DK model, we only carry out the pooling operations over different scales. Since translation invariance is not necessary, this benefits the accurate positioning.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In the experiments, we consider two noisy IR image sequences, denoted by P and S, to demonstrate the effectiveness and practicality of the proposed method. The experiments were conducted on a computer with 4-GB random access memory and 2.50-GHz Intel i5 processor, and the code was implemented in MATLAB.

In this section, we will show that the proposed method can improve the SNR of the image and then increase detection accuracy. In the experiments, the detection rate (DR) and false alarms (FAs) per image are used as standards for comparison. We have $DR = (NC/NT) \times 100\%$ and $FA = NIC/N$, where NC is the number of correctly detected targets, NT is the number of true targets, NIC is the number of incorrectly detected targets, and N is the length of the sequence.

A. Test on Sequence P

For the sequence P, the imaging sensor which was located at a distance of about 15–25 km from the target was on a ground-looking aircraft in the sky. The operating wavelength of the IR camera is 8.3 μm, and the resolution of it is 320 $\times$ 240 pixels. This sequence was captured on a sunny day.

The detection result of one frame of this sequence is given in Fig. 5. The position of the target in the sequence presented here is visually verified. It is easy to see that the proposed method gives impressive result. The proposed method is compared with the top-hat method since it is well studied and shows good performance in terms of the small target detection problem [3], [13], [38]. In addition, the average gray absolute difference maximum map (AGADMM) is an efficient method for multiscale small target detection using template matching [11]. It can be included in the proposed mathematical framework. Consequently, it can be taken as a baseline method. The DR and FA are shown in Table I. Obviously, our method performs the best. In addition, the SNR of the original images and that of the final contrast maps are given in Table II. We can find that the SNR of the original image is low and the proposed method can improve the SNR of the image significantly.

TABLE VI
COMPARISON OF SEVERAL TARGET DETECTION METHODS

Method	Top-hat	AGADMM	Proposed method
DR	88.57	92.86	**100**
FA	0.1857(13/70)	0	**0**

In order to illustrate the effectiveness of the proposed method more intuitively, the 3-D surf view of one frame and the corresponding contrast map (the image shown in Fig. 5) are given in Fig. 6. It is easy to observed that the proposed method can enhance the target.

Furthermore, in order to demonstrate the robustness of the proposed method, it was tested on the image sequence corrupted with known noise. In this experiment, Gaussian white noise of variance 0.00001 and that of variance 0.00005 are added to each frame, respectively. The experimental results are given in Tables III and IV. These results show that the proposed method achieves low FA and high DR. We can find that the proposed method is robust to noise.

B. Test on Sequence S

In this section, we evaluate the performance of the proposed method on the sequence S. In this sequence, the target is a fighter, and the the resolution is 256 $\times$ 200 pixels. The detection result of one frame is shown in Fig. 7, where the detected target position is marked by rectangles. As can be seen from Fig. 7, the target is correctly detected. In addition, the SNR of this sequence is low (see Table V). The SNRs of the final contrast maps corresponding to eight consecutive frames are given in Table V. Consequently, we can conclude that the proposed method can improve the SNR of the image. The comparison with other methods is given in Table VI. The superiority of our method is very clear. The 3-D surf view of the original image and the corresponding contrast map of one frame (the image shown in Fig. 7) of this sequence are also given in Fig. 8. It is easy to observe that the target can be enhanced by our method.

Finally, the trajectory and error curve of eight consecutive frames of this sequence are given in Fig. 9(a) and (b), respectively. In Fig. 9(a), the tracking trace does almost match that of the target movement. It can be observed from Fig. 9(b) that the vertical error is less than 2 pixels and the horizontal error is less than 4 pixels. Usually, if the distance between a ground truth and a detected position is within a threshold (5 pixels), then the detection is declared as being correct [3]. Consequently, the proposed method is effective.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

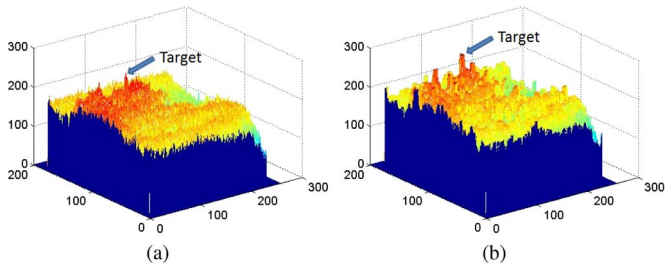


Fig. 8. Three-dimensional projection of (a) the original image in sequence S and (b) the corresponding contrast map.

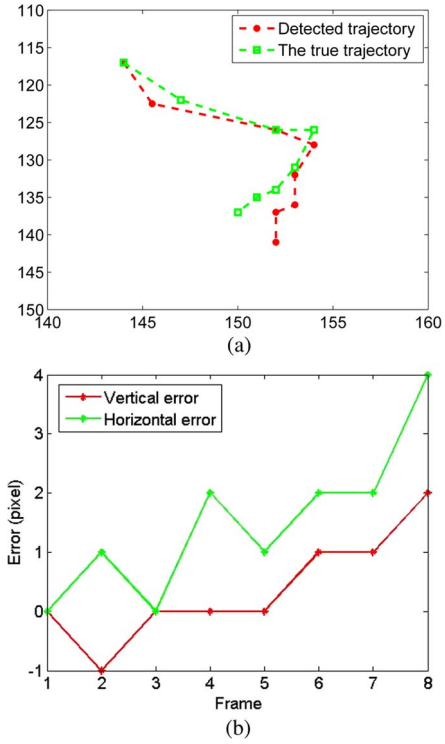


Fig. 9. (a) Obtained trajectory of the first eight frames. (b) Error curves of the first eight frames.

V. Conclusion

In this paper, an effective small target detection algorithm motivated by the robust properties of HVS and the hierarchical architecture of the DK model is presented. The key idea of the proposed method is to use the LCM to enhance the targets. The effectiveness of the proposed method has been demonstrated on two different detection scenarios. The experimental results show that local contrast can improve the SNR of the image and the LCM-based detection method significantly outperforms the conventional methods top-hat and AGADMM. The proposed method not only is simple but also has a good performance. Consequently, we can conclude that our method is suitable for IR small target detection.

Although the experiments confirm that the proposed method is robust and provide empirical evidence supporting the claim that it is fit for small target detection, we can improve it from different directions in the future work. For example, we can extend the definition of the LCM by making use of different comparing and pooling operations.

References

[1] M. Malanowski and K. Kulpa, "Detection of moving targets with continuous-wave noise radar: Theory and measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 9, pp. 3502–3509, Sep. 2012.

[2] S. Bourennane, C. Fossati, and A. Cailly, "Improvement of target-detection algorithms based on adaptive three-dimensional filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 4, pp. 1383–1395, Apr. 2011.

[3] S. Kim and J. Lee, "Scale invariant small target detection by optimizing signal-to-clutter ratio in heterogeneous background for infrared search and track," *Pattern Recognit.*, vol. 45, no. 1, pp. 393–406, Jan. 2012.

[4] L. F. Zhang, L. P. Zhang, D. Tao, and X. Huang, "A multifeature tensor for remote-sensing target recognition," *IEEE Geosci. Remote Sensing Lett.*, vol. 8, no. 2, pp. 374–378, Mar. 2011.

[5] N. Thành, H. Sahli, and D. Hao, "Infrared thermography for buried land-mine detection: Inverse problem setting," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 12, pp. 3987–4004, Dec. 2008.

[6] J. Zhao, Z. Tang, J. Yang, and E. Liu, "Infrared small target detection using sparse representation," *J. Syst. Eng. Electron.*, vol. 22, no. 6, pp. 897–904, Dec. 2011.

[7] R. Liu, Y. Lu, C. Gong, and Y. Liu, "Infrared point target detection with improved template matching," *Infrared Phys. Technol.*, vol. 55, no. 4, pp. 380–387, Jul. 2012.

[8] H. Li, Y. Wei, L. Li, and Y. Y. Tang, "Infrared moving target detection and tracking based on tensor locality preserving projection," *Infrared Phys. Technol.*, vol. 53, no. 2, pp. 77–83, Mar. 2010.

[9] T. Bae, "Small target detection using bilateral filter and temporal cross product in infrared images," *Infrared Phys. Technol.*, vol. 54, no. 5, pp. 403–411, Sep. 2011.

[10] E. Ashton, "Detection of subpixel anomalies in multispectral infrared imagery using an adaptive Bayesian classifier," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 2, pp. 506–517, Mar. 1998.

[11] G. Wang, T. Zhang, L. Wei, and N. Sang, "Efficient small-target detection algorithm," in *Proc. SPIE*, Jul. 1995, vol. 2484, pp. 321–329.

[12] J. Barnett, "Statistical analysis of median subtraction filtering with application to point target detection in infrared backgrounds," in *Proc. SPIE*, Jun. 1989, vol. 1050, pp. 10–15.

[13] V. Tom, T. Peli, M. Leung, and J. Bondaryk, "Morphology-based algorithm for point target detection in infrared backgrounds," in *Proc. SPIE*, Oct. 1993, vol. 1954, pp. 2–11.

[14] S. Deshpande, M. Er, and R. Venkateswarlu, "Max-mean and max-median filters for detection of small-targets," in *Proc. SPIE*, Oct. 1999, vol. 3809, pp. 74–83.

[15] H. Li, S. Xu, and L. Li, "Dim target detection and tracking based on empirical mode decomposition," *Signal Process. Image*, vol. 23, no. 10, pp. 788–797, Nov. 2008.

[16] J. J. Soraghan, "Small-target detection in sea clutter," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 7, pp. 1355–1361, Jul. 2004.

[17] 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 Trans. Geosci. Remote Sens. Lett.*, vol. 7, no. 3, pp. 469–473, Jul. 2010.

[18] W. Diao, X. Mao, and V. Gui, "Metrics for performance evaluation of preprocessing algorithms in infrared small target images," *Progr. Electromagn. Res.*, vol. 115, pp. 35–53, Jun. 2011.

[19] S. G. Sun, D. M. Kwak, W. B. Jang, and D. J. Kim, "Small target detection using center-surround difference with locally adaptive threshold," in *Proc. IEEE Int. Symp. Image Signal Process. Anal.*, Sep. 2005, pp. 402–407.

[20] S. G. Sun and D. M. Kwak, "Automatic detection of targets using center-surround difference and local thresholding," *J. Multimedia*, vol. 1, no. 1, pp. 16–23, Apr. 2006.

[21] Z. Shao, X. Zhu, and J. Liu, "Morphology infrared image target detection algorithm optimized by genetic theory," in *Proc. Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, 2008, vol. XXXVII, pp. 1299–1303.

[22] X. Bai, F. Zhou, and B. Xue, "Infrared dim small target enhancement using toggle contrast operator," *Infrared Phys. Technol.*, vol. 55, no. 2/3, pp. 177–182, Mar. 2012.

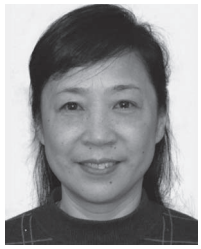
[23] G. Boccignone, A. Chianese, and A. Picariello, "Small target detection using wavelets," in *Proc. Int. Conf. Pattern Recognit.*, Aug. 1998, vol. 2, pp. 1776–1778.

[24] H. Zhang, X. Liu, J. Li, and Z. Zhu, "The study of detecting for IR weak and small targets based on fractal features," in *Proc. Adv. Multimedia Model.*, Jan. 2007, vol. 4352, pp. 296–303.

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

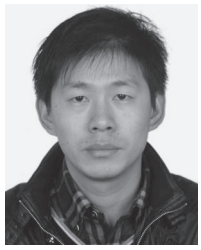
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[26] M. Shirvaikar and M. Trivedi, "A neural network filter to detect small targets in high clutter backgrounds," *IEEE Trans. Neural Netw.*, vol. 6, no. 1, pp. 252–257, Jan. 1995.
[27] K. Huang, D. Tao, Y. Yuan, X. Li, and T. Tan, "Biologically inspired features for scene classification in video surveillance," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 41, no. 1, pp. 307–313, Feb. 2011.
[28] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Robust object recognition with cortex-like mechanisms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 3, pp. 411–426, Mar. 2007.
[29] D. Song and D. Tao, "Biologically inspired feature manifold for scene classification," *IEEE Trans. Image Process.*, vol. 19, no. 1, pp. 174–184, Jan. 2010.
[30] S. Smale, L. Rosasco, J. Bouvrie, A. Caponnetto, and T. Poggio, "Mathematics of the neural response," *Found. Comput. Math.*, vol. 10, no. 1, pp. 67–91, Feb. 2010.
[31] H. Li, Y. Wei, L. Li, and Y. Yuan, "Similarity learning for object recognition based on derived kernel," *Neurocomputing*, vol. 83, pp. 110–120, Apr. 2012.
[32] G. Wang, T. Zhang, L. Wei, and N. Sang, "Efficient method for multiscale small target detection from a natural scene," *Opt. Eng.*, vol. 35, no. 3, pp. 761–768, Mar. 1996.
[33] R. VanRullen, "Visual saliency and spike timing in the ventral visual pathway," *J. Physiol. Paris*, vol. 97, pp. 365–377, Mar.–May 2003.
[34] W. Zhang, M. Cong, and L. Wang, "Algorithm for optical weak small targets detection and tracking: Review," in *Proc. IEEE Int. Conf. Neural Netw. Signal Process.*, Dec. 2003, vol. 12, pp. 643–647.
[35] J. Bouvrie, T. Poggio, L. Rosasco, S. Smale, and A. Wibisono, "Generalization and properties of the neural response," Massachusetts Institute of Technology, Cambridge, MA, USA, MIT-CSAIL-TR-2010-051/CBCL-292, Nov. 19, 2010.
[36] H. Li, Y. Wei, L. Li, and C. L. P. Chen, "Hierarchical feature extraction with local neural response for image recognition," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, 2013, doi: 10.1109/TSMCB.2012.2208743, in press.
[37] S. Kim, Y. Yang, J. Lee, and Y. Park, "Small target detection utilizing robust methods of the human visual system for IRST," *J. Infrared Millim. Terahz Waves*, vol. 30, no. 9, pp. 994–1011, Sep. 2009.
[38] J. Luo, H. Hi, and J. Liu, "An algorithm based on spatial filter for infrared small target detection and its application to an all directional IRST system," in *Proc. SPIE*, Jan. 2007, vol. 6279, pp. 62 793E1–62 793E6.



Hong Li (M'07) received the M.Sc. degree in mathematics and the Ph.D. degree in pattern recognition and intelligence control from the Huazhong University of Science and Technology, Wuhan, China, in 1986 and 1999, respectively.

She is currently a Professor with the School of Mathematics and Statistics, Huazhong University of Science and Technology. Her current research interests include approximation theory, wavelet analysis, machine learning, neural networks, signal processing, and pattern recognition.

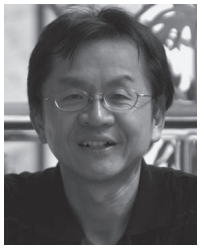


Yantao Wei received the Ph.D. degree in control science and engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2012.

He is currently with the College of Information Technology, Journalism and Communications, Central China Normal University, Wuhan. His current research interests include computer vision and pattern recognition.

Tian Xia received the Honours degree in computer science and technology from the University of Adelaide, Adelaide, Australia, in 2009. He is currently working toward the Ph.D. degree in the Faculty of Science and Technology, University of Macau, Macau, China.

His current research interests include computer vision and pattern recognition.



C. L. Philip Chen (S'88–M'88–SM'94–F'07) received the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1988.

He is currently a Chair Professor of the Department of Computer and Information Science and the Dean of the Faculty of Science and Technology at the University of Macau, Macau, China.

Dr. Chen is an American Association for the Advancement of Science Fellow. He is currently the President of the IEEE Systems, Man, and Cybernetics Society.

Yuan Yan Tang (S'88–M'88–SM'96–F'04) received the Ph.D. degree in computer science from Concordia University, Montreal, Canada.

He is currently a Chair Professor with the Faculty of Science and Technology, University of Macau, Macau, China. He is the Founder and Editor-in-Chief of the International Journal on Wavelets, Multiresolution, and Information Processing and Associate Editor-in-Chief of the *International Journal on Frontiers of Computer Science.*

Dr. Tang is an International Association of Pattern Recognition Fellow.