Multiscale facet model for infrared small target detection

Changcai Yang<sup>a</sup>, Jiayi Ma<sup>b</sup>,⇑, Meifang Zhang<sup>c</sup>, Sheng Zheng<sup>d</sup>, Xin Tian<sup>b</sup>  

<sup>a</sup>College of Computer and Information Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China  
<sup>b</sup>Electronic Information School, Wuhan University, Wuhan 430072, China  
<sup>c</sup>Basic Subject Department of Fujian Health College, Fuzhou 350101, China  
<sup>d</sup>Institute of Intelligent Vision and Image Information, College of Science, China Three Gorges University, Yichang 443002, China

HIGHLIGHTS  
- The directional second-order directional derivative filters are deduced to enhance the targets.  
- A multiscale representation provided by the proposed method is used to reduce the false alarm rate.  
- The proposed method is effective to detect the small target in complex background with low SCR.

ARTICLE INFO  
Article history:  
Received 1 June 2014  
Available online 5 August 2014

Keywords:  
Directional second-order directional derivative filter  
Facet model  
Multiscale facet model matrix  
Small target detection

ABSTRACT  
In this paper, we proposed a new robust infrared small target detector that is more suitable for complex background with low signal-to-clutter ratio. The original image is decomposed into sub-bands in different orientations by using the directional second-order directional derivative (DSODD) filters deduced from the facet model. The multiscale facet model (MFM) analysis is developed by using a series of multiscale DSODD filters, which are obtained by filling zeros in the basic DSODD filter. Based on MFM, an MFM matrix is constructed, and the normalized determinant of this matrix is then defined as the target measure. The corresponding multiscale correlations of the target measures are computed to enhance the target signal and suppressing the background clutter. The experimental results on a set of real infrared images demonstrate that the proposed approach is effective and is superior to the traditional small target detection methods in terms of the pertained quantitative detection evaluation indexes, such as the signal-to-clutter ratio gain and the background suppression factor.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction  
The robust detection of infrared small targets in clutter is an issue of critical importance to infrared search and track (IRST) applications for self-defense or attacks [1–6]. Due to that the target is at a far distance, its projected image is usually very small and does not have available shape and texture for detection or matching [7–9]. Furthermore, because of the effects of inherent sensor noise or environment, the obtained infrared small targets are often buried in a complex background with low signal-to-clutter ratio (SCR). Therefore, it is difficult to detect the small targets in complex infrared background.

In order to detect a small target effectively, many approaches have been proposed in the past few decades. The small target regions are typically different from the surrounding clutter background. Based on this fact, some background prediction model based methods are proposed to detect targets. Gu et al. proposed a kernel-based nonparametric regression method for background prediction and clutter removal [1]. Each pixel of the observed infrared imagery is represented by using a linear mixture model. The clutter background is estimated by using kernel regression and small target is subsequently detected from the “pure” target-like region. Deng and Liu introduced an efficient method for background prediction based on the concept of self-information [2]. The self-information map (SINM) is constructed by a Parzen window function. Then the adaptive thresholding method followed by a region growing technique is adopted to detect the target from SINM. The major drawbacks of this method are the target growing effect and large calculation. To overcome these disadvantages, Deng et al. developed a weighted SINM based method and improved the region growing technique [3]. The spatial bilateral filter (BF) is integrated with temporal profile to predict background without targets [4]. An infrared patchimage (IPI) model based on
the non-local selfcorrelation property of the infrared image for small target detection was proposed by Gao et al. [6]. They assume that the target and background patch-images are a sparse and a low-rank matrix, respectively. Then the background image is reconstructed by recovering the low-rank and sparse matrices. Bae et al. [5] introduced an edge directional 2D least mean squares (LMSs) filter to predict the background excluding small targets. Small targets can be extracted by subtracting the predicted background from the original infrared image.

Traditional small target detection methods including Top-hat filtering [10], max-mean/max-median filter [11] and so on, are widely used to suppress the background clutter or enhance the small target. Based on Top-hat filtering, some new related methods have been presented. The Top-hat filtering parameters are optimized by using neural network and genetic algorithm [12]. A new Top-hat transformation method was proposed by Bai et al. [13]. They reorganized the classical Top-hat transformation by using two different but correlated structuring elements. The different information between the target and surrounding regions is also taken into account. Meng et al. proposed an adaptive method for small target detection [14]. The modified top-hat transformation using interrelated structuring elements is used to adaptively detect the darker and brighter targets and suppress the cluttered background. However, they sometimes did not give truly satisfactory results. The main difficulty is that the size of the small target may vary from 2 × 2 to 12 × 12 pixels [15].

In order to solve the scale problem, many multiscale based methods have been proposed. According to the principle of human discrimination of small targets from a natural scene, an efficient multiscale small target detection using template matching based on the average gray absolute difference maximum map was proposed by Wang et al. [16]. A small targets detection method based on support vector machines (SVM) in the wavelet domain has been presented [17]. Motivated by the robust properties of the human visual system (HVS), Kim and Lee proposed a scale invariant small target detection method [18], directional saliency based method [19], image layering based method [20], as well as local contrast method [21]. These methods can achieve good performance in typical applications but may become less effective for the small target in complex background with low SCR.

In this paper, we explore a new robust infrared small target detection method based on multiscale facet model. The experiments are provided in Section 4. We conclude this paper in Section 5.

2. Multiscale facet model

In the following, we describe the necessary algorithm on which our algorithm is based, and then a multiscale representation is developed by using a series of multiscale DSODD filters.

2.1. DSODD

Let R and C be the index sets of the neighborhood that satisfy the symmetric conditions, i.e., \( r \in R \) implies \( -r \in R \) and \( c \in C \) implies \( -c \in C \). Similar to the second-order directional derivative (SODD) based method [22,23], we deduce the DSODD filters from the facet model [24]. The set \( \{ P_1(r,c), \ldots, P_n(r,c) \} \) of discrete orthogonal polynomials over \( R \times C \) can be constructed. Then the fitted intensity surface function over the constant vector space, \( R \times C \), can be written as follows:

\[
\hat{f}(r, c) = \sum_{n=1}^{N} a_n P_n(r,c),
\]

(1)

where \( f(r, c) \) is the corresponding intensity value. For the \( 5 \times 5 \) neighboring window in which \( R = \{-2 -1 0 1 2\} \) and \( C = \{-2 \ldots -1012\} \), \( P_n(r,c) \) are shown as \( \{1, r, c, r^2 - 2, c^2 - 2, r^3 - \frac{1}{2}cr, (r^2 - 2)c, (c^2 - 2)r, c^3 - \frac{1}{2}cr^2\} \).

The fitting coefficients can be estimated by minimizing a residual error and can be written as

\[
a_n = \frac{\sum_{r \in R} \sum_{c \in C} f(r,c)P_n(r,c)}{\sum_{r \in R} \sum_{c \in C} P_n^2(r,c)}. \]

(2)

From Eq. (2), we can seen that each fitting coefficient \( a_n \) can be computed as linear combination of the data values. For each index \( (r,c) \) in the index set, the intensity value \( f(r,c) \) is multiplied by the weight

\[
W_m = \frac{P_m(r,c)}{\sum_{r \in R} \sum_{c \in C} P_m(r,c)}. \]

(3)

If we substitute \( P_m(r,c) \) into Eq. (3), the weight kernels \( W_m \) in Eq. (3) can be obtained as follows:

\[
\begin{align*}
W_4 &= \frac{1}{70} \begin{bmatrix}
2 & 2 & 2 & 2 & 2 \\
-1 & -1 & -1 & -1 & -1 \\
-2 & -2 & -2 & -2 & -2 \\
-1 & -1 & -1 & -1 & -1 \\
2 & 2 & 2 & 2 & 2 
\end{bmatrix}, \\
W_5 &= \frac{1}{100} \begin{bmatrix}
4 & 2 & 2 & -2 & -4 \\
2 & 1 & -1 & -1 & -2 \\
0 & 0 & 0 & 0 & 0 \\
-2 & -1 & -1 & 1 & 2 \\
-4 & -2 & 2 & 2 & 4 
\end{bmatrix}, \\
W_6 &= W_4^T. 
\end{align*}
\]

Assume that in each neighborhood of the image the intensity surface function \( f \) takes the parametric form of a polynomial in the row and column coordinates. Thus, in each neighborhood \( f \) can be written as:

\[
f(r,c) = k_1 + k_2r + k_3c + k_4r^2 + k_5rc + k_6c^2 + k_7r^3 + k_8r^2c + k_9rc^2 + k_{10}c^3.
\]

(4)

Evaluating the second row and column partial derivatives at the center point \((0,0)\) from Eq. (4) yields the second order directional derivatives.
The SODD of \( f \) at point \((r,c)\) in direction \( \alpha \) can be written as

\[
\frac{\partial^2 f(r,c)}{\partial r^2} \bigg|_{(0,0)} = 2K_4, \quad \frac{\partial^2 f(r,c)}{\partial r \partial c} \bigg|_{(0,0)} = K_5, \quad \frac{\partial^2 f(r,c)}{\partial c^2} \bigg|_{(0,0)} = 2K_6.
\]

(5)

The SODD of \( f \) at point \((r,c)\) in direction \( \alpha \) can be written as

\[
f_x = \frac{\partial^2 f(r,c)}{\partial x^2} \sin^2 \alpha + \frac{2\partial^2 f(r,c)}{\partial x \partial c} \sin x \cos \alpha + \frac{\partial^2 f(r,c)}{\partial c^2} \cos^2 \alpha,
\]

where \( \alpha \) is the clockwise angle from the column axis. If we substitute Eq. (5) into Eq. (6), the SODD of \( f \) at the center point \((0,0)\) in the direction \( \alpha \) can be written as

\[
f_x \big|_{(0,0)} = 2K_4 \sin^2 \alpha + 2K_5 \sin x \cos \alpha + 2K_6 \cos^2 \alpha.
\]

(7)

We set \( \alpha \) to be 90\(^\circ\) and 0\(^\circ\) for horizontal and vertical orientations, respectively. Then the horizontal and vertical SODD filters \( f^h \) and \( f^v \) can be deduced from Eq. (7).

2.2. Multiscale representation

When we represent scenes, generally the sizes of small targets are changing. It can be advantageous to represent the images at several scales. Similar to the multiscale support value filter [25–27], a series of multiscale horizontal and vertical SODD filters can be obtained by filling zeros in the DSODD filters. In this way, we can decompose the original image \( I \) into sub-bands in horizontal, vertical, and diagonal orientations according to the following rules:

\[
S^h_l = f^h_l \ast I, \quad S^v_l = I - S^h_l,
\]

\[
S^{vh}_l = f^v_l \ast I, \quad I_{a-1} = S^v_l - S^{vh}_l,
\]

\[
S^{bh}_l = f^h_l \ast I, \quad S^{bh}_l = S^h_l - S^{bh}_l, \quad l = 1, \ldots, L, I_1 = I.
\]

(8)

where \( l \) is the scale level, \( \ast \) represents convolution operation, \( S^{bh}, S^{vh} \) and \( S^h, S^v \) are the DSODD map of the original image obtained by multiscale decomposition in horizontal, vertical, and diagonal orientations, respectively.

Fig. 1 shows the DSODD decomposition at scale 1 for infrared image. Fig. 1(a) is an infrared image under sea-clutter background. Fig. 1(b) shows the 3D gray value distribution for Fig. 1(a). DSODD map with horizontal, vertical, and diagonal orientations are shown in Figs. 1(c–e), respectively. The proposed MFM gives a multiscale representation containing DSODD maps which signify every target whose size is adapted to the resolution of the filter at each scale and whose orientation is in horizontal, vertical, or diagonal direction.

3. Target detection based on MFM

The MFM can be extended to detect small targets. In the following, we construct an MFM matrix and then define the target measure using the determinant of this matrix. Each step of the proposed approach is then discussed in details.

3.1. MFM based detector

To effectively utilize the local structural information in the DSODD map, we define an MFM matrix for each location \((x,y)\) at scale \( l \) as follows:

\[
M_l(x,y) = \begin{bmatrix} S^{hh}_{l}(x,y) & S^{hv}_{l}(x,y) \\ S^{vh}_{l}(x,y) & S^{vv}_{l}(x,y) \end{bmatrix}, \quad l = 1, \ldots, L.
\]

(9)

The eigenvalues \((\lambda_1, \lambda_2)\) of the MFM matrix represent two principal signal changes in the neighborhood of a point, as shown in Fig. 1(f) and (g). For small target, the product of the eigenvalues has large values, as shown in Fig. 1(h). This property enables the extraction of small targets, where the signal change is significant in both orthogonal directions. Based on this principle, we define the small target measure as the determinant of the MFM matrix.

\[
DM_l(x,y) = \prod \left[ S^{hh}_{l}(x,y)S^{vv}_{l}(x,y) - (S^{vh}_{l}(x,y))^2 \right].
\]

(10)

Since the determinant of the matrix is equal to the product of its eigenvalues, \(DM_l(x,y)\) reaches a extremum for blob-like small targets in the image. As illustrated in Fig. 2(a–d), the new measures effectively integrate the information of the small targets at different orientations. As shown in Fig. 2, there is an inherent adaptivity of the analysis to the target size and orientation since the size of the convolution filter increases with the scale of analysis. The defined small target measures computed from the first level integrate the significant values at those locations where pixel-sized significant features are present (see Fig. 2(a)), and significant measure values in the detail images correspond more and more to significant features of increasing spatial dimension with decreasing of the resolution (see Fig. 2(b–d)). However, it is very difficult to pick up the interest features from the analysis of any single scale. To overcome the limitation of data coming from a single DSODD map, we take advantage of the multiresolution representation provided by the MFM. From results in Fig. 2(a–d), we can see that where there are small targets in an image, the local maxima in small target measure planes tend to propagate across scales, while...
others are not. We therefore design a multiscale spatial filtering scheme that the small target measure has high value in the presence of a small target, and non-significant value for the background. To this end, we compute MFM map \( C_{L}^{DM}(x, y) \) which is defined at each location \((x, y)\) by the direct spatial multiscale product of the small target measure values.

\[
C_{L}^{DM}(x, y) = \prod_{l=1}^{L} DM_l(x, y),
\]

\[
\bar{C}_{L}^{DM}(x, y) = \frac{C_{L}^{DM}(x, y) - \min(C_{L}^{DM}(x, y))}{\max(C_{L}^{DM}(x, y)) - \min(C_{L}^{DM}(x, y))},
\]

where \( L \) is the lowest level at which the correlation is computed. Fig. 2(e) illustrate the spatial multiscale products of the small target measure values, where the level \( L \) is 4. Fig. 2(f) shows the correlation image \( C_{4}^{DM} \).

We subsequently use the fact that the product of significant measure values across scales at the location \((x, y)\) results in significant values of \( C_{L}^{DM}(x, y) \) only if the local maxima propagate down to the considered scales. Obviously, if the local maxima die at some intermediate scale, this will largely decrease the values of \( C_{L}^{DM}(x, y) \).

As illustrated in Fig. 2(e), the multiscale product of the small target measure values effectively enhance the spot small targets, while degrading the background.

### 3.2. MFM based small target detection method

The MFM map can enhance the small targets and suppress the background clutter simultaneously as mentioned above. Consequently, it is likely that \( C_{L}^{DM}(x, y) \) reach maxima for target in the scene. Based on this fact, an adaptive threshold is adopted to segment the small target.

\[
t_{th} = \mu_{CI} + \tau \sigma_{CI},
\]

where \( \mu_{CI} \) and \( \sigma_{CI} \) are the mean and the standard deviation of the correlation images, respectively, \( \tau \) is a constant determined experimentally. A pixel at \((x, y)\) is segmented as a target pixel if...
4. Experiments and discussions

To evaluate the performance of the proposed method for small target detection, we test it on a set of collected real infrared images. The first column in Fig. 5 are six representative images denoted as Images A to F. The background types are sea-sky (Images A and B), sky (Image C), cloud cluttered sky (Images D and E), ground with sky (Image F). Images A to F are of sizes 276 × 224, 276 × 224, 146 × 146, 557 × 448, 146 × 146, and 216 × 136, respectively.

All the experiments are implemented in Matlab 8.1 on Windows 8 operating system and run on a computer with 4-GB random access memory and 2.60-GHz Intel i5 processor.

4.1. Evaluation metrics

In this section, we will show that the proposed method can effectively enhance the targets while suppress the background clutter. In the experiment, two common performance measures, the SCRG and BSF, are used as standards for comparison which are defined as follows [28]:

\[
\text{SCRG} = \frac{(S/C)_{\text{out}}}{(S/C)_{\text{in}}}, \quad \text{BSF} = \frac{C_{\text{in}}}{C_{\text{out}}}. \tag{13}
\]

<table>
<thead>
<tr>
<th>Image</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top-hat</td>
<td>13.1302</td>
<td>12.8424</td>
<td>4.5133</td>
<td>4.2261</td>
<td>1.4169</td>
<td>1.2851</td>
</tr>
<tr>
<td>Max-mean</td>
<td>9.8896</td>
<td>4.3303</td>
<td>2.8059</td>
<td>2.4546</td>
<td>1.6470</td>
<td>0.8047</td>
</tr>
<tr>
<td>Max-median</td>
<td>8.8410</td>
<td>1.4729</td>
<td>4.2194</td>
<td>0.9717</td>
<td>1.5262</td>
<td>0.3917</td>
</tr>
<tr>
<td>MLoG</td>
<td>6.6212</td>
<td>1.1661</td>
<td>1.8333</td>
<td>1.1035</td>
<td>3.0064</td>
<td>0.3845</td>
</tr>
<tr>
<td>MFM</td>
<td>209.6746</td>
<td>84.8699</td>
<td>25.8020</td>
<td>57.6400</td>
<td>7.5602</td>
<td>2.0841</td>
</tr>
</tbody>
</table>

Fig. 4. The signal-to-clutter ratio gain (SCRG) and background suppression factor (BSF) of the proposed MFM method as functions of the scale level.

Fig. 5. The original images and the corresponding processed results of different methods.
where \( S \) and \( C \) are the signal amplitude and clutter standard deviation, respectively; \( \text{in} \) and \( \text{out} \) represent the input original image and the output target enhanced map, respectively. SCRG reflects the amplification of target signal relative to backgrounds after and before processing, whereas BSF expresses the suppression level of backgrounds without any information about target. The larger the SCRG and BSF values, the better the filtered results are.

The ROC curve can also be used to evaluate the performance of the small target detection methods. The ROC curve reflects the varying relationship between the detection probability \( P_d \) and the false alarm rate \( P_f \) which are defined as follows [1]:

\[
P_d = \frac{\# \text{number of detected pixels}}{\# \text{number of real target pixels}}
\]

\[
P_f = \frac{\# \text{number of false alarms}}{\# \text{total number of pixels in the whole image}}.
\] (14)

### 4.2. Effect of scale level

In practice, there are “bright” and “dark” targets in the image. For “bright” targets, small target measure \( DM_l(x, y) \) reaches a maximum. For “dark” targets, small target measure \( DM_l(x, y) \) reaches a negative minimum. For convenience of segmenting target, the scale level \( L \) is set to be an even number. Correlation image \( C^l_4 \)
reaches a maximum for targets in the image. Fig. 4 shows the value of SCR for the proposed MFM method used to examine the effect of varying the scale level L. This figure is based on a small targets detection task using Images A to F. As the Fig. 4 shows, the highest SCR and BSF are obtained when sampling 4 scales each original image, and this is the number of scale level used for all other experiments throughout this paper.

4.3. Comparison to state-of-the-art methods

The proposed MFM small target detector is compared to Top-hat filter-based method (Top-hat) [10], Max-mean filter-based method (Max-mean) [11], Max-median filter-based method (Max-median) [11], Facet model (FM) [29], and Min-local-LoG filter (MLLoG) based method [30]. The columns from the second to right in Fig. 5 show the results of six methods before segmentation. It can be seen that our method has less clutter and noise residual for different clutter backgrounds compared to other methods.

Table 1 gives the values of quality of SCRG for the small target detection approaches. For Images A to D, the proposed MFM method obtains the best SCRG score and are followed by FM and Top-hat methods. For Image E, the proposed MFM method obtains the best SCRG score and are followed by FM and MLLoG methods. For Image F, the SCRG value of the MF method is 2.6660 and is a little larger than 2.0841, provided by the proposed MFM method.

Table 2 gives the values of quality of BSF for the small target detection approaches. For Images A to F, the proposed MFM method obtains the best BSF score and is followed by the FM method. From Tables 1 and 2, we can note that the proposed MFM method provides the best target enhancement and background clutter suppression performance in terms of resulting values of the quantitative evaluation indexes including SCRG and BSF.

Fig. 6 shows the ROC curves of the six methods for six real infrared images. For the Images A to E, the proposed method performs best. For Image F, however, there is no clear winner. The Max-mean and Top-hat methods have a little better performance than our method when \( P_f \leq 1.0212 \times 10^{-4} \), but our method can reach 1 (300%) faster than the Top-hat method when \( P_f > 1.0212 \times 10^{-4} \). The Top-hat method performs well on the different types of background except for cloud cluttered sky background (Images D and E), where the Top-hat method performs significantly worse than our method. The worse result for the Top-hat method is due to the gray intensity differences between the target and the background are not sufficient for the Top-hat method to enhance the target from low SCR infrared image. The Max-mean and MLLoG methods give average scores. The proposed method has better performance than other methods for all six real infrared images, which means that the proposed method is robust and effective for different clutter and noisy backgrounds and target types.

4.4. Computation efficiency analysis

The complexity and efficiency of a small target detection method is an important issue in particular when applying the method to image sequences. The computational complexity of the proposed FM method is essentially determined by the DSODD decomposition. The DSODD decomposition at each scale requires 3 convolutions and 3 subtracts. The size of the multiscale DSODD filter at scale \( l \) is \((hl – l + 1)(wL – l + 1)\), where \( h \times w \) is the original DSODD filter size. The computational complexity of each convolution is \( O((hl – l + 1)(wL – l + 1)m) \), where \( m \times n \) is the image size. The scale level \( L \) is 4. The computational complexity of each subtract is \( O(mn) \). So the total computational complexity of the DSODD decomposition is \( O(30hw – 20h – 20w + 26m) \).

In order to estimate the computation efficiency of different small target detection methods, the time costs (seconds) for different approaches are listed in Table 3. The fastest method is MLLoG since it only smooths the image with four directional local LoG filters. The time costs of the proposed MFM method and the Top-hat method are almost twice as the FM method. The main reason is that the proposed MFM method needs more time to decompose the original image into sub-bands. Therefore, its computation efficiency is lower than the FM and MLLoG methods.

5. Conclusions

In this paper, we develop an MFM analysis method, and present an MFM based small target detection method. The key ideas of the proposed method are to use the DSODD filters to enhance the targets and the multiscale representation provided by MFM to reduce the false alarm rate. We have carried out the performance comparison among the proposed MFM, Top-hat [10], Max-mean [11], Max-median [11], FM [29], and MLLoG [30]. The effectiveness of the proposed method has been demonstrated on a set of real infrared images with different background clutters. The experimental results show that the proposed MFM method outperforms the traditional small target detection methods especially for the complex background with low SCR.

Conflict of interest

There is no conflict of interest in this article.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of the Chinese Academy of Sciences astronomical union funds under Grant U1331113, by the National Natural Science Foundation of China under Grant 61102064, and by the China Postdoctoral Science Foundation under Grants 2013M540582 and 2014T70707.

References


