Graph Laplacian regularization based edge-preserving background estimation for single frame small target detection

Kun Bai\textsuperscript{a,b}, Yuehuang Wang*\textsuperscript{a,b}, Qiong Song\textsuperscript{b}, Mingna Liu\textsuperscript{c}, Jiandong Wu\textsuperscript{c}

\textsuperscript{a}School of Automation, Huazhong University of Science and Technology, Luoyu Road 1037, Wuhan, China, 430074; \textsuperscript{b}National Key Laboratory of Science and Technology on Multi-spectral Information Processing, Luoyu Road 1037, Wuhan, China, 430074; \textsuperscript{c}Institute of Shanghai Aerospace Control Technology, Zhongchun Road 1555, Shanghai, China, 201109

ABSTRACT

Small target detection in the clutter infrared image is a tough but significant work. In this paper, we will propose a novel small target detection method. First, Graph Laplacian regularization is utilized to model similarity feature of graph structure in the image. And Graph Laplacian regularization is incorporated in the background estimation model to preserve edges of background in single frame infrared image. At last, the edge-preserving estimated background is eliminated from original image to get foreground image which is used to detect the small target. Experimental results show that our proposed method can achieve edge-preserving estimation of background, suppress clutter efficiently, and get better detection results.

Keywords: Graph Laplacian regularization, small target detection, edge-preserving, background estimation

1. INTRODUCTION

Small target detection is widely used in many military and civil applications. It has been a hot area of research in recent decades. Many researchers have developed several methods for small target detection in the infrared image, and these detection methods can be classified into two categories: detect-before-track (DBT) and track-before-detect (TBD). DBT methods are detecting targets in single frame first, and then track the detected target; the sample pre-detection methods in single frame are first used in TBD methods and consecutive frames are utilized to accumulate target energy and suppress noise. However, sensor ego-motion in practice limits TBD methods applications, and a better single image small detection method can have improve the result of TBD methods too. So, we keep our focus on DBT methods in this paper.

The most important task of small target detection in single frame is suppressing the clutter background in infrared image. Lots of methods have been proposed to segregate the background and foreground and suppress the clutters in the background [1, 2, 3]. These methods utilize the local contrast between small target and background, the grayvalue of target is always different with its surrounding pixels and is not spatially correlated with local neighborhood [1], to estimate background image. So these methods share similar shortcoming. In these methods, the edges in infrared image cannot be suppressed efficiently. It is because that the edges in the infrared image can also give rise to the local contrast. If small target detection methods enhance the possible targets, the edges are usually enhanced by the way. Addressing this problem, lots of edge-preserving background estimation methods are proposed. The edge direction is detected by analyzing the surrounding blocks around current background estimation window [4, 5]. Then two-dimensional least mean square (TDLMS)[5] is adjusted to the direction of edge to preserve edges. But if edges with various directions exist in the estimation window, such as at corner, they will get bad edge-preserving results. Edge-preserving ability of bilateral filter is also used to estimate the background [6, 7] and these two methods get better results using multi-frames infrared images. So the robust edge-preserving background estimation method for single frame infrared small target detection is still needed.

In this paper, the Graph Laplacian regularization is introduced to estimate clutter background in small target detection. And the edge-preserving ability is specially taken into account when we design clutter background estimation method. This approach is motivated by literature [8], in which Graph Laplacian regularization is used to describe grayvalue similarity of same graph structure. When we want to estimate background grayvalue of a pixel on a graph structure in the image, the grayvalue similarity feature can be used to estimate background grayvalue. It is capable of preserving the graph structure. So the Graph Laplacian regularization is cooperated in background estimation model.

*yuehwang@hust.edu.cn; phone 86 135-0718-1806; fax 8687558913
In the reminder of this paper, we will introduce the proposed Graph Laplacian regularization based background estimation model in section 2. The details of our infrared small target detection method will be described in section 3. Experimental results are presented in section 4. We will reach the conclusion in section 5.

2. GRAPH LAPLACIAN REGULARIZATION BASED BACKGROUND ESTIMATION

2.1 Infrared image model for small target detection

For lots of small target detection algorithms, the infrared image is always modeled using three component images: foreground image which is also called as target image, background image and noise image. And this infrared image model can be illustrated using following equation.

\[ G(p) = T(p) + B(p) + N(p) \] (1)

Here, \( G(p) \) is grayvalue of pixel \( p \) in the original infrared image. \( T(p) \) is grayvalue of target image. \( B(p) \) is grayvalue of background image at pixel \( p \). \( N(p) \) is grayvalue of noise image, and the noise is usually white Gaussian noise. The clutter in the background image causes more false alarms than noise image for the small target detection result. Therefore, we should preserve the clutter such as edge in the background image to separate the clutters and target precisely.

2.2 Grayvalue Similarity of graph structure

Natural images generally have enough redundancy in the gray value domain. In the flat region of image, there are lots of pixels which share similar grayvalues. Even straight or curve edges have a complete line of pixels which share similar grayvalues. In the Fig. 1, four patch images are selected to calculate the grayvalue similarity in the different region of image. Patch (a) contain one edge which is caused by sea and air, the edge in patch (b) is caused by clouds, edge in the patch (c) is caused by sea and island, and patch (d) is a flat region in the sea.

![Figure 1. Four patch images are selected from original image to testify the similarity of edges in the infrared image.](image)

Fig. 2 illustrates the grayvalue similarity of the pixels that are on a structure with other pixels in the patch image, and the grayvalue similarity \( S(p, p_i) \) is calculated using Equ. (2). We can find that a pixel on a structure can get lots of pixels which share similar grayvalue. Thus, when we want to estimate the background grayvalue of one pixel that is on one structure, we can first find the similar pixels, and then we can these similar pixels can be utilized to estimate the grayvalue of estimated pixel. Since that these pixels own similar grayvalue we can preserve the edges in background estimation process.
Here, $p$ is the pixel in the patch image, $p_c$ is the center pixel which is labeled using red point in Fig. 2. $G(p)$ is the grayvalue of pixel $p$, and $G(p_c)$ is grayvalue of pixel $p_c$. Parameter $\delta$ is used to control the similarity distribution.

$$S(p, p_c) = \exp\left(-\frac{(G(p) - G(p_c))^2}{2\delta^2}\right)$$  \hspace{2cm} (2)$$

Figure 2. Display the grayvalue similarity of the pixels that are on a structure or in the flat region, the similarity goes from 1 (white) to zero (black).

2.3 Graph Laplacian regularization based background estimation model

In the above subsection, we have detailed the grayvalue similarity of the graph structure such as the edges in the infrared image. Here, the background estimation model will be introduced. The grayvalue similarity feature of graph structures will be formulated using Graph Laplacian, and then the Graph Laplacian regularization will be cooperated in background estimation model.

Background image estimation methods are usually processing in patch image. Some methods use all the pixels in patch image to estimate background image, and L2 norm are also used in these models to formula background estimation model. However, L2 norm are sensitive to outliers, then existence of target region will cause a bias of background estimation. Some other methods use a portion of pixels to estimation the background to prevent the effect the target region, as shown in Fig.3. In these methods, pixels between two boxes are used to predict the background image gray value. In our semi-supervised background estimation model, we do not only consider the effect of target region, but also use information of all pixels in patch. Pixels between two boxes are used as labeled samples set to avoid the effect of target region, and other pixels in patch image are used as unlabeled samples.

Figure 3. The structure used in background estimation model, the pixels between the outer box and inner box are selected as labeled samples.
Set patch image has n pixels, all of these pixels can be regarded as samples in our semi-supervise learning model. Among all these samples, the labeled samples $X_l = \{(x_1, y_1), ..., (x_l, y_l)\}$ is used as training set, $x_i$ is the feature vector of $i$th labeled pixel $p_i$ and $y_i$ is gray value of $i$th labeled pixel $p_i$, $l$ is the number of labeled samples. Set $f(x)$ is the estimated background image, we formulated the estimation process as minimization of an energy function as follows:

$$\arg \min_{f \in \mathcal{H}_\kappa} J(f) = \arg \min_{f \in \mathcal{H}_\kappa} \sum_{i=1}^{l} \|y_i - f(x_i)\|^2 + \lambda \|f\|^2$$

(3)

Here, $\mathcal{H}_\kappa$ is Reproducing Kernel Hilbert Space which is associated with kernel $\kappa$. $\mathcal{H}_\kappa$ is completion of the linear span given by $\kappa(x, \cdot)$ for all $x_i$ in patch image.

In above, we have introduced the background estimation model. This background estimation model only utilized the grayvalues of pixels between outer box and inner box to estimation the grayvalue of test pixel. But if the test pixel on a graph structure, its similar pixels on the same structure are more suitable to estimate the grayvalue of test pixel. So, we will utilized Graph Laplacian regularization to represent this feature and give these similar pixels more weight in Equ.(3).

The patch image is modeled as undirected graph, the vertices and edges in graph represent image pixel gray value and the relationships between pixels. These intrinsic geometrical structures of patch image pixels can be described by Graph Laplacian. Through Graph Laplacian regularization, geometrical structures preserving term can be incorporated in optimal Equ. (3). The Graph Laplacian regularization can be mathematically implemented by minimizing the following term:

$$R(f) = \frac{1}{2} \sum_{i,j} W_{ij} (f(x_i) - f(x_j))^2$$

(4)

Here, $W_{ij}$ reflects the similarity of pixel $p_i$ and pixel $p_j$, pixel $p_i$ and pixel $p_j$ are in one patch image. The $W_{ij}$ is designed as follow:

$$W_{ij} = \frac{1}{C} \exp \left\{ -\frac{\|u_i - u_j\|^2}{\sigma^2} \right\} \exp \left\{ -\frac{\|y_i - y_j\|^2}{\epsilon^2} \right\}$$

(5)

Where, $u_i = [u_x, u_y]$ is the pixel $p_i$’s coordinates in patch image, and $y_i$ is the grayvalue of pixel $p_j$, $C$ is the normalization factor.

From Eq. (2,3), we can see that if pixel $p_i$ and pixel $p_j$ are in same structure, they are more likely to share similar gray value. So $W_{ij}$ will become bigger, and when $x_j$ is labeled sample, estimated gray value $f(x_j)$ will approximate to $f(x_j)$ to minimize $R(f)$. Then edges can be preserved in background estimation.

The Equ (4) can be rewritten as following equation:

$$R(f) = \mathbf{f}^T \mathbf{L} \mathbf{f}$$

Here, $\mathbf{f} = [f(x_1), ..., f(x_n)]^T$. $\mathbf{D}$ is a diagonal matrix, whose diagonal elements $\mathbf{D}_{ij} = \sum_{j=1}^{n} W_{ij}$. And $\mathbf{L} = \mathbf{D} - \mathbf{W}$.

Graph Laplacian regularization now can be incorporated in the semi-supervised background estimation model:
\[
\arg\min_{f \in \mathbb{H}} \left\{ J(f) = \|Y_f - f_r\|^2 + \lambda \|f\|^2 + \gamma f^T L f \right\}
\]

Here, \(Y_f = [y_1, \cdots, y_T]^T\), \(f = [f(x_1), \cdots, f(x_n)]\), and \(f = [y_1, \cdots, y_T]^T\).

2.4 The optimization scheme for background estimation

We can define the solution of Equ. (6) as following equation by using the kernel regression model[9]:

\[
f = \left[ f(x_1), \ldots, f(x_n) \right]^T = \sum_{i=1}^{n} \alpha_i \kappa(x_i, x_n) = Ka
\]

Here, \(K\) is the kernel gram matrix, and \(K_{ij} = \kappa(x_i, x_j)\) for \(i, j = 1, \ldots, n\). The kernel function \(K_{ij} = \kappa(x_i, x_j)\) should be able to adapt locally to the edges for preserving the edges in the background image. So we will define an edge adapted kernel in following.

In Fig. 1, 2, it is easily to find that the grayvalues of one pixel who is on same edge correlates with other edge pixels in the patch image, the target grayvalue is totally biased from its surrounding neighborhood pixels. So the edge adapted kernel should utilize grayvalue relationship difference between target and edges. Meanwhile, if the distance of two pixels is small, and the two pixels have more probability to share similar grayvalue. Thus, the edge adapted kernel can be defined as:

\[
\kappa(x, x_j) = \exp \left( -\frac{\|u_i - u_j\|^2}{\sigma_x^2} \right) \exp \left( -\frac{\|y_i - y_j\|^2}{\delta_x^2} \right)
\]

Here, \(u_i\) and \(u_j\) are the coordinates of sample \(x_i\) and \(x_j\). And \(\sigma_x\) is the parameter to control Gaussian model, and \(\sigma_x\) can be related with the patch image size, if patch image size is \((2p+1)(2p+1)\), we can set the \(\sigma_x = p\) to ensure each sample in the structure do not get too small spatial weight. \(y_i\) and \(y_j\) are grayvalues of sample \(x_i\) and \(x_j\) in the patch image. \(\delta_x\) is the parameter to control Gaussian distribution.

Then the closed-formed solution of Equ.(6) is shown in Equ. (9), and \(\hat{f} = Ka\). And \(K_L\) is the submatrix consisting of rows of \(K\) which are corresponding to the labeled samples in set \(X_L\), we also have \(f_L = K_L a\).

\[
\hat{a} = \left( K_L^T K_L + \lambda K + \gamma K L K \right)^{-1} K_L^T y_L
\]

3. EXPERIMENTS AND RESULTS

In this section, the experiments will be performed in real infrared images with sea-sky, urban building and land background. The original infrared images are shown in Fig. 4. In image 1, a boat is in the harbor and this image has sea-sky background. The land background exists in image 2. The background in image 3 is urban building. The building and
land background both exist in last image. In the experiments, the parameters $\lambda$ and $\gamma$ in Equ.(6) are set as 0.1 and 0.7 from experience. The Graph Laplacian parameters $\sigma = 5, \epsilon \in [5,12]$ and the edge adapted kernel $\sigma_\epsilon = 5, \delta_\epsilon \in [5,12]$.

![Image 1](image1.jpg) ![Image 2](image2.jpg) ![Image 3](image3.jpg) ![Image 4](image4.jpg)

(a) image 1  (b) image 2  (c) image 3  (d) image 4

Figure 4. The original images used in the experiments.

The estimated background images results are shown in Fig. 5, we can easily find that our method preserves the edge details in original images exactly in the background estimation results. We can get foreground image by eliminating the estimated background image from the original image. If foreground contains fewer edges, the better edge-preserving background estimation results we get. The foreground images of our method and other baseline methods are shown in Fig. 6. The foreground results of bilateral filter based method are displayed in the first row, images in second row are results of edge directional TDLMS method, and our results are shown in last row. And we can find that the edge directional TDLMS method get worse edge-preserving result in the corner. From this visual comparison, we can find that foregrounds of our methods contain least edges and most close to the ground truths. So our proposed method is capable of preserving best edge in background.

![Image 1](image5.jpg) ![Image 2](image6.jpg) ![Image 3](image7.jpg) ![Image 4](image8.jpg)

(a) image 1  (b) image 2  (c) image 3  (d) image 4

Figure 5. The estimated background images results.

Additionally, the signal-to-clutter ratio gain (SCR Gain) is calculated to compare the performance quantitatively. And the SCR gain is defined as follows:

$$SCR \ Gain = \frac{(S/C)_{\text{out}}}{(S/C)_{\text{in}}}$$  \hspace{1cm} (10)

Here, $S$ is the signal amplitude and $C$ is the clutter standard deviance. The indicators $\text{in}$ and $\text{out}$ denote the original image and foreground image. The SCR gains comparison is shown in Fig. 7. From the figure, we can easily conclude that our edge-preserving background method get highest SCR gain in the different scenarios.
Figure 6. The foreground images results. The first row is results of bilateral filter, the second row is results of TDLMS, and last row is results of our proposed method.

Figure 7. The SCR gain comparison, our method achieves highest SCR gains than other two baseline methods.
4. CONCLUSION

In this paper, we proposed the Graph Laplacian regularization based background estimation method. The Graph Laplacian regularization describes the grayvalue similarity of pixels on the same structure to preserve the edges in the single frame infrared image. Experimental results show that our method can achieve best edge-preserving results than the state-of-the-art baseline methods.

ACKNOWLEDGEMENTS

This work has been supported by Natural Science Foundation of China No.61227007, research foundation No.9140A01060113JW05016, research project No.51301040201, and Shanghai Aerospace Innovation Foundation (Contract No. SAST201348).

REFERENCES