Infrared Physics & Technology 73 (2015) 141–148

Contents lists available at ScienceDirect

Infrared Physics & Technology

journal homepage: www.elsevier.com/locate/infrared

Small and dim target detection by background estimation

Jing Hu, Yi Yu *, Fan Liu

Science & Technology on Multi-Spectral Information Processing Laboratory, School of Automation, Huazhong University of Science and Technology, Wuhan 430074, PR China

ARTICLE INFO

Article history:
Received 17 January 2015
Available online 30 September 2015

Keywords:
Small and dim target detection
Complicated background
Background modeling
Non-local means

ABSTRACT

An effective method for small and dim moving target detection in complicated background is proposed. The proposed approach takes advantage of the Non-local means filter, and applies a novel weight calculation model based on circular mask to the original background estimation pattern. By associating similarity of grayscale distribution of the images with temporal information, the extended method estimates the complicated background precisely and extracts point target successfully. To compare existing target detection methods and the proposed one, signal-to-clutter ratio gain (SCRG) and background suppression factor (BSF) are employed for spatial performance comparison and receiver operating characteristics (ROC) is used for detection-performance comparison of the target trajectory. Experimental results demonstrate good performance of the proposed method for infrared images in complicated scene, especially for images with low signal-to-noise ratio.

1. Introduction

For infrared (IR) images with complicated background and low signal-to-noise ratio, it's a challenge to detect small and dim moving point target. There is no significant shape and construction feature of target for extraction and detection, and the intense clutter interference caused by complicated background makes it more difficult. Early work involved spatial processing of single images followed by temporal association and tracking is computationally simple but performs poorly for small and dim moving targets in such condition [1]. Large amounts of research in this field have been carried out for decades [2–4], solving the problem from perspectives of time, space, frequency domain, etc.

Background estimation provides new thinking for small and dim moving target detection, it's based on self-similarity of images. The grayscale, gradient, and other features of any point in images can be estimated by the points around it, if the estimation falls down, then we can take this point as the target. Background estimation is one of the most important approaches for target detection. The main difficulty of background estimation is to precisely estimate the background starting from the observed image, that is, to obtain an accurate background model, while being robust to the presence of targets [5]. The classical algorithms, for instance, max-mean and max-median algorithms [6], and recent proposed methods such as Kalman filter [7], bilateral filter [8], morphological filter [8,9], are all effective background estimation algorithms. The edge directional 2D least means square filter [10] proposed by Tae-Wuk Bae takes full advantage of edge information of prediction pixels. It optimizes the predicted value by directional gradient and improves the convergence speed by new model of step size. It's superior to the original LMS algorithm in both of detection performance and calculation speed. C. L. Philip Chen proposed a new kind of algorithm named as Local Contrast Method [11]. It's inspired by the contrast mechanism of human vision system and derived kernel model. The algorithm takes advantage of the neighborhood of the current location, and measures the similarity between the current location and its neighborhoods to obtain a local contrast map of the input image. The algorithm works well even in low SNR images.

2. Problem proposed

The NL-means filter proposed by Buades et al. [12] is one of background estimation methods as well, it defines the difference between pixels with structural similarity, and makes an overall comparison in whole area around the central pixel, which decides the weight according to the similarity of grayscale distribution of the images. The method takes advantage of a weighted average in expressing self-similarity information between central pixel and its neighborhood, and suppresses the background with non-iteration and local operation.

The soul of NL-Means is estimating all pixels by neighborhood based on the principles of local regularity and self-similarity. It's exactly the theory foundation of background estimation. The target

* Corresponding author. Tel.: +86 15927372585.
E-mail address: hust_h@163.com (Y. Yu).

http://dx.doi.org/10.1016/j.infrared.2015.09.009
1350-4495/

can be taken as noise, it appears in the scene and changes the grayscale of the original pixels. If the image background can be precisely estimated by NL-means algorithm, then the target can be found in the residual between original image and the recovered one. According to NL-means, the pixels can be estimated by the equations below,

$$\tilde{v}(i) = \sum_{j \in I} \omega(i,j) v(j) \tag{1}$$

$$\omega(i,j) = \frac{1}{Z(i)} e^{-\frac{\|v(N_i)-v(N_j)\|_{2,a}^2}{h^2}} \tag{2}$$

$$Z(i) = \sum_{j \in I} e^{-\frac{\|v(N_i)-v(N_j)\|_{2,a}^2}{h^2}} \tag{3}$$

where $\tilde{v}(i)$ refers to the predicted grayscale of central pixel i of similarity window [12], $v(j)$ refers to the grayscale of each central pixel j of similarity window in squared window [12], $v(N_i)$ and $v(N_j)$ denote separately the intensity grayscale vector of similarity window N_i and N_j, and $\|v(N_i) - v(N_j)\|_{2,a}^2$ refers to the Gaussian weighted Euclidean distance [12] between N_i and N_j. Estimation procedure is shown schematically in Fig.1.

The NL-means is superior to most kinds of typical algorithms in image denoising [12]. For target detection, it works well with large similarity window, but the computational complexity force us to reduce the size of window, which results in increasing of missing alarm rate.

The detection result of NL-Means algorithm is shown in Fig. 2, Fig. 2(a) is the original image with Gaussian noise and Fig. 2(b) is the estimation, it indicates that the pixels with target on it is still bright. And we can't find the target in the corresponding place of the residual between them shown in Fig. 2(c).

Laure Genin et al. proposed an improvement based on the original method [13], nevertheless, the new method aims at being robust to high intensity point object, its performance on complicated background is still limited. Our method modifies the similarity criterion to make it less sensitive to point target even with small size of similarity window, and results in a better effect in complicated background estimation. In our method, a circular mask and the corresponding weight model are proposed, and the NL-means filter is used to temporally filter the image sequences, we named it R-NLM for short.

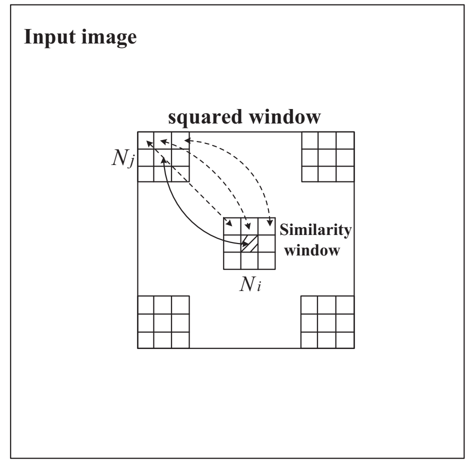

Fig. 1. The original NLM algorithm filtering sketch.

3. Overview of R-NLM algorithm

In order to precisely estimate the background, while being robust to the presence of target, the similarity window with target should be separated with windows with background only, however, there is no prior knowledge about which pixel belonging to the target and which belonging to the background, furthermore background information is also needed when processing pixel with target on it. Actually, every pixel should be estimated with grayscale of background rather than target, then the target can be extracted successfully from the residual between the original image and the estimated one.

Consequently, we separate the similarity window to two parts, as shown schematically in Fig. 3, in which the circular area peripheral is marked as area A and the smaller square area in center as area B. Moreover, we apply the match principle between two continuous frames, setting the squared window on the next frame. It's possible that the target may appear in area A, the probable situations are shown in Fig. 4. The target extends to several pixels owing to the point spread function (PSF) [14]. Generally speaking, in our mask there are more pixels belonging to area A than to area B, and the target changes the grayscale of only several pixels, as the Euclidean distance accumulated from each pair of pixels in similarity window need to be averaged at last, the effect made by the target is insignificant when it locates in area A, therefore this situation can be ignored, only the cases when the target locates in area B are taken into account in the paper, just as the 5 kinds shown in Fig. 5.

The similarity level between two match windows is measured by $\|v(N_i) - v(N_j)\|_2^2$, where $v(N_i)$ and $v(N_j)$ denote separately the intensity grayscale vector of similarity window N_i and N_j, and Euclidean distance is used instead of the Gaussian weighted Euclidean distance in NL-means. The relative distance between similarity windows in all the situations can be estimated by a simple principle, which is, obviously, if the distance equals to zero approximately, such as between target and target or same background, we think it's "small", on the contrary, for the distance between different backgrounds or target and background, we think it's "large". Table 1 explains the details, in which the target, background and different background are respectively marked by TA, BK and NB.

Before analyzing the expected weight, let's take a look at the situations of two contiguous frames. As the target is moving between two frames, it means the target is impossible located in the pixels with same background. It declares two kinds of situations as following: ① if there is a target in the area B, the backgrounds in its neighborhood in two frames must be different, just as shown in Fig. 5(a); ② if there are two match windows with same background, one of them has the target in center, then the center of the other must be background only, just as shown in Fig. 5(b).For the first condition, similarity of area A is low, while that of area B is high; for the second situation, on the contrary, the similarity of area B is low, while that of area A is high.

Except the situation mentioned before, if the similarity of area A and area B are both low, that must be two match windows with different backgrounds, or one with target as its center and the other is a totally different patch of background, which are shown separately in Fig. 5(c) and (d); and two match windows of similar background if the similarity level of area A and area B are both large, just as shown in Fig. 5(e).

After making clear of all the possible situations involved in the estimation procedure, we work backwards based on the conclusion above. The derivation procedure is shown schematically in Fig. 6, and the corresponding relationship of similarity, Euclidean distance and expected weight is given in the Table 2.

The Euclidean distance of area A and area B are defined as $DIS_A$and DIS_B respectively in Eqs. (4) and (5).