A Novel Method for Real-time Multiple Moving Targets Detection from Moving IR Camera

Fenghui Yao[1], Ali Sekmen[1], and Mohan Malkani[2]
[1]*Department of Computer Science,* [2]*Department of Electric and Computer Engineering*
Tennessee State University, Nashville, TN, USA
{fyao, asekmen, mmalkani}@tnstate.edu

Abstract

This paper presents a novel method for detecting multiple moving targets in real-time from infrared (IR) image sequences collected by an airborne IR camera. This novel method is based on dynamic Gabor filter and dynamic Gaussian detector. First, the ego-motion induced by the airborne platform is modeled by parametric affine transformation based on feature point matching, and the IR video is stabilized by eliminating the background motion. Then, a dynamic Gabor filter is employed to enhance the image changes for accurate detection and localization of moving targets. The orientation of Gabor filter is dynamically changed according to the orientation of optical flows. Next, the specular highlights generated by the dynamic Gabor filter are detected. The outliers and specular highlights are fused to indentify the moving targets. The experimental results show that the proposed detection algorithm is effective and efficient. And the detection speed is approximate 2 frames per second.

1. Introduction

Detection of moving targets in infrared (IR) imagery is a challenging research topic in computer vision. Detecting and localizing a moving target accurately is important for automatic tracking system initialization and recovery from tracking failure. Although many methods have been developed on detecting and tracking targets in visual images (generated by daytime cameras), there exits limited amount of work on target detection and tracking from IR imagery in computer vision community [1]. In comparison to the visual images, the images obtained from an IR camera have extremely low signal to noise ratio, which results in limited information for performing detection and tracking tasks. In addition, in airborne IR images, non-repeatability of the target signature, competing background clutter, lack of a priori information, high ego-motion of the sensor, and the artifacts due to weather conditions make detection or tracking of targets even harder. To overcome the shortcomings of the nature of IR imagery, different approaches impose different constrains to provide solutions for a limited number of situations. For instance, several detection methods require that the targets are hot spots which appear as bright regions in the IR images [2] [3] [4]. However, in realistic target detection scenarios, none of these assumptions are applicable, and a robust detection method must successfully deal with these problems.

This paper presents an approach for robust real-time target detection in airborne IR imagery. This approach has the following characteristics: (1) it is robust in presence of high global motion and significant texture in background, (2) it does not require that targets have constant velocity or acceleration, (3) it does not assume that target features do not drastically change over the course of tracking. The main contribution of this paper is the complete algorithm presented. There are two foci in this algorithm. The first one is the *dynamic Gabor filter*, where the orientation of Gabor filter is controlled by the orientation of the optic flows. The second one is *dynamic Gaussian detector*, which is used to identify the target location. The following shows the algorithm in detail.

2. Algorithm description

This algorithm can be formulated in four steps: (i) motion compensation, (ii) dynamic Gabor filtering, (iii) specular highlights detection, and (iv) target

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

localization. The following will describe these processing steps in detail.

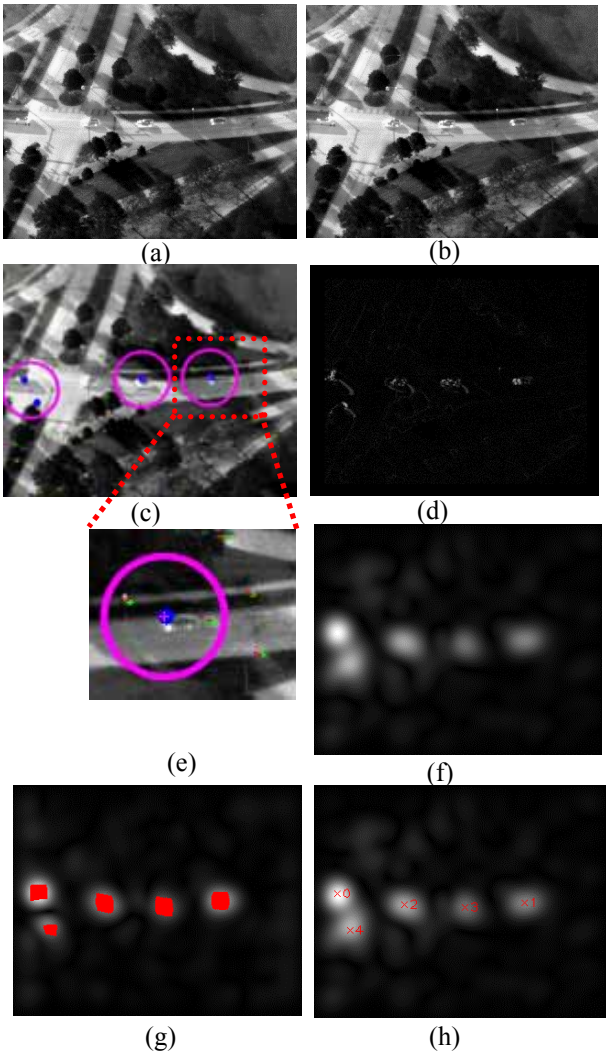

Fig. 1 (a) and (b) Two input images; (c) Detected optical flows; (d) Image changes; (e) Partially enlarge of (c) to show the outliers and inliers; (f) Dynamic Gabor filter response; (g) Specular highlights; (h) Clusters of specular highlights.

2.1. Motion compensation

The motion compensation contains the feature point extraction, optical flow detection, global parametric motion model estimation, and motion detection.

A. Feature point detection. The feature point extraction is used as the first step of this algorithm. Harris corner detector, Shi-Tomasi's corner detector, SUSAN, SIFT, SURF, and FAST are some representative feature point detection algorithms developed over past two decades. We evaluated these algorithms according to two criteria, processing time and detection accuracy. Our experiment results show that Shi-Tomasi's method is more reliable than others, and is pretty fast. Therefore, this work employs Shi-Tomasi's method for feature point detection. For two input images, $I^{t'}$ and I^{t}, let $P^{t'} = \{ p_1^{t'}, \ldots, p_M^{t'} \}$, and $P^{t} = \{ p_1^{t}, \ldots, p_N^{t} \}$ denote the feature points detected from $I^{t'}$ and I^{t}, respectively, where $t' = t - \Delta$, $p_i^{t'} = (x_i^{t'}, y_i^{t'})$, $p_j^{t} = (x_j^{t}, y_j^{t})$, $i = 1, 2, \ldots, M$ and $j = 1, 2, \ldots, N$. In the following, $I^{t'}$ is called previous image, I^{t} is called current image or reference image.

B. Optical flow detection. There are many optical flow detection algorithms. Recently there are several new developments on this topic. The evaluation results of these algorithms in [5] show that Bouguet's method [6] is the best for the interpolation task. In our algorithm, we employed Bouguet's method for optical flow detection. Let $F^{t't} = \{ \overrightarrow{F_1^{t't}}, \overrightarrow{F_2^{t't}}, \ldots, \overrightarrow{F_K^{t't}} \}$ denote the detected optical flows. For the feature points in set $P^{t'}$ and P^{t}, from which no optical flow is detected, they are filtered out. Therefore, after this filtering operation, the number of feature points in two sets, $P^{t'}$ and P^{t}, becomes the same with the number of optical flows in optical flow set $F^{t't}$, that is, K. Fig. 1 (a) and (b) show two input images, $I^{t'}$ and I^{t}, where Δ is set at 3. Fig. 1 (c) shows the optical flows detected from the feature points, where the optical flow are marked by red line segments and the endpoints of the optical flows are marked by green dots (refer to (e) for the partially enlarged picture).

C. Motion model estimation. Generally, the approach to find the coordinate transformation relies on assuming that it will take one of the following six models, (i) translation, (ii) affine, (ii) bilinear, (iv) projective, (v) pseudo perspective, and (vi) biquadratic, and then estimating the two to twelve parameters in the chosen models. The translation model is not applicable to the problem that contains rotation. The complicated models such as projective and biquadratic are computationally heavy, and parameter estimation are difficult. Here we tested affine, bilinear, and pseudo perspective model by adding some error to a parameter and checking how the image is distorted. The experiment results show that affine model is robust to parameter estimation error. Therefore, our method use affine model. Six parameters in affine model are estimated by using feature points.

(1) Feature points are separated into two categories, inliers set $P_{in}^{t'}$ and P_{in}^{t}, and outliers set $P_{out}^{t'}$ and P_{out}^{t}. The feature points associated with the moving targets are called *outliers*. Those associated with the background are called *inliers*.

(2) Outliers are clustered by distance-based clustering algorithm, which will be used for target identification. Inliers are used to estimate the affine model by employing RANSAC-like algorithm. Let $\boldsymbol{A}_b$ denote the estimated affine model.

D. Motion image generation. Here, in airborne imagery, the motion image means changes caused by the moving targets. The previous image is transformed by the affine model $\boldsymbol{A}_b$, and subtract from the current image. Fig. 1 (d) shows the motion image generated by Eq. (1) from two input images in Fig. 1 (a) and (b).

2.2. Dynamic Gabor filter

A Gabor wavelet is defined as,

$$\psi_{\mu,\nu}(z) = \frac{\left\|k_{\mu,\nu}\right\|^2}{\sigma^2} e^{-\frac{\left\|k_{\mu,\nu}\right\|^2 \times \|z\|^2}{2\sigma^2}} \left[e^{ik_{\mu,\nu}z} - e^{-\frac{\sigma^2}{2}} \right] \quad (1)$$

where $z = (x, y)$ is the point with the horizontal coordinate x and the vertical coordinate y. The parameters μ and ν define the orientation and scale of the Gabor kernel, $\|\cdot\|$ denotes the norm operator, and σ is related to the standard derivation of the Gaussian window in the kernel and determines the ratio of the Gaussian window width to the wavelength. The wave vector $k_{\mu,\nu}$ is defined as follows

$$k_{\mu,\nu} = k_\nu e^{i\phi_\mu} \quad (2)$$

where $k_\nu = k_{\max} / f^\nu$ and $\phi_\mu = \pi\mu/8$, $k_{\max}$ is the maximum frequency, and f^ν is the spatial frequency between kernels in frequency domain.

In our algorithm, we fix the following parameters, $k_{\max} = \pi/2$, $\sigma = 2\pi$, $f = \sqrt{2}$, and ν =3. The orientation μ is dynamically changed according to optical flows from inliers. We call it *dynamic Gabor filter*. The orientation μ is defined as,

$$\mu = \frac{1}{K_{in}} \sum_{i=1}^{K_{in}} \theta(\vec{F}_i^{t't}), \quad (3)$$

where $\theta(\vec{F}_i^{t't})$ is the orientation of the optical flow $\vec{F}_i^{t't} \in F_{in}^{t't}$, and is given by,

$$\theta(\vec{F}_i^{t't}) = \arctan \frac{y_i^{t'} - y_i^{t}}{x_i^{t'} - x_i^{t}}. \quad (4)$$

Fig. 1 (f) shows the Gabor filter response by performing convolution for the frame difference image in Fig. 1 (d) and the dynamic Gabor kernel. Dynamic Gabor filter enhanced the frame difference.

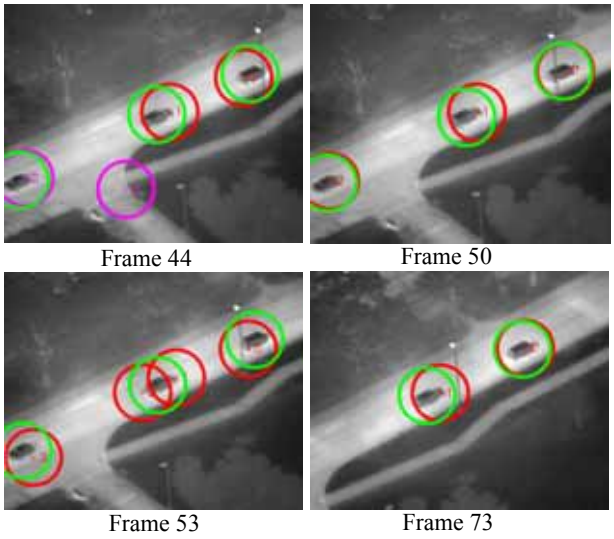

Fig. 2 Target detection results in frame 44, 50, 53, and 73. Green circles mark the ground truth target positions, labeled manually. Red circles means targets detected based on outliers clustering and specular highlights. Purple circles mark the output of the dynamic Gaussian detector.

2.3. Specular highlights detection

As can be seen in Fig. 1 (f), the image changes appear as high intensity in the dynamic Gabor filter response. We call these high intensity *specular highlights*. Therefore, the target detection problem becomes the specular highlights detection problem. Because the intensity of highlights changes for the moving targets (some specular highlights are dimmer than others), the thresholding algorithms cannot detect all specular highlights successfully. Here, we employ the pixel intensity on the circular circles C_1, C_2, and C_3, centered at C_0, where C_0 is the pixel under examination. The detector compares the intensity at C_0 and the intensity of pixels on three circular circles C_1, C_2, and C_3, with radius R_1, R_2, and R_3, respectively. C_1, C_2, and C_3 are sampled at $\pi/6$ interval, hence the detector will only compare the intensity at C_0 and 12 sample points, $C_{j,1}$, $C_{j,2}$, …, $C_{j,12}$, from each circular circle. Let $G(z)$ denote the dynamic Gabor filter response at z, the discrimination of specular highlights is as follows. If $G(C_0) \geq G(C_{j,i})$ and $G(C_{j,i}) \geq G(C_{j+1,i})$, C_0 is a specular

highlight, otherwise not, where j = 1, 2, and i = 1, 2, …, 12.

The specular highlight points detected from the dynamic Gabor filter response in Fig. 1 (f) are shown in Fig. 1 (g) by red dots. The specular highlight clustering results are sown in Fig. 1 (h).

2.4. Moving target localization

Outliers and specular highlights are used to localize the moving targets. The identification process is as follows.

(i) For a specular highlight, if its center lies in the terrain of a outlier cluster, it is considered as a target. If its center does not lie in any cluster, the dynamic Gaussian detector is employed.

(ii) For the general 2-D Gaussian function, its orientation is controlled by the orientation of the specular highlight. Here we call it *dynamic Gaussian detector*. LACC is used as to similarity measure. If LACC is larger than threshold T_G, this specular highlight is considered as a moving target.

3. Experiment results

The entire algorithm described in Section 3 is implemented by using C++ and OpenCV on windows platform. Δ is set at 2, the similarity threshold T_G at 0.93, and A, σ_x, and σ_y are set at 1, 25.0, and 15.0, respectively. The radii, R_1, R_2, and R_3, of three circles for specular highlights detection is 7, 10, and 13, respectively. The IR video data from the Vivid datasets, provided by the Air Force Research Laboratory, is used. Fig. 2 shows target detection results at frame 44, 50, 53 and 73 for an input image sequence. Green circles mark the ground truth target positions, labeled manually, red circles means targets detected based on outlier clustering and specular highlights, and purple circles marks the output of the dynamic Gaussian detector.

4. Performance analysis

To evaluate the performance of this algorithm, we selected four image sequences with the significant background as the test data. Each sequence contains 100 frames, and each frame contains 2 to four moving targets. The ground truth targets are labeled manually. The total number of targets in these 4 datasets is 1231. We examined the correct detection rate, hit rate, and processing time. The hit rate is defined as the ratio for the intersected area of detected target and ground truth target and the area of the ground truth target. The experiments are conducted on a Windows Vista machine mounted with a 2.33 GHz Intel Core 2 CPU and 2GB main memory. The total average correct detection rate is 86.6%, and hit rate is 78.6%, respectively. The detail detection results are shown in Table 1. The average processing time is 581 ms/frame.

Table 1. Target detection results.

	Data 1	Data 2	Data 3	Data 4
Total targets	381	266	287	297
Detected targets	326	221	249	270
Missed targets	55	45	38	27
Correct detection	85.6%	83.1%	86.8%	90.9%
Miss detection	14.4%	16.9%	13.2%	9.1%
Hit rate	85.9%	81.3%	70.7%	76.6%

5. Conclusions and future work

This paper described a method for multiple moving target detection from airborne IR imagery. We tested the algorithm by using the airborne IR videos from AFRL Vivid datasets. The correct detection rate is 86.6%, and the hit rate for the correct detection is 78.6%. The processing rate is 581 ms/frame, that is, approximate 2 frames per second. This speed meets the requirement for many real-time target detection and tracking systems. The future work is to apply this algorithm to the target tracking systems.

Acknowledgement

This work was partially supported by a grant from AFRL under Minority Leaders Program, contract No. TENN 06-S567-07-C2.

References

[1] A. Yilmaz, K. Shafique, and M. Shah: “Target tracking in airborne forward looking infrared imagery,” Image and Vision Computing, 21 (2003) pp.623-635.

[2] H. Shekarforoush, R. Chellappa: “A multi-fractal formalism for stabilization, object detection and tracking in flir sequences,” IEEE International Conference on Image Processing, vol. 3, 2000.

[3] D. Davies, P. Palmer, Mirmehdi: “Detection and tracking of very small low contrast objects,” Ninth British Machine Vision Conference, September, 1998.

[4] A. Strehl, J.K. Aggarwal: “Detecting moving objects in airborne forward looking infra-red sequences,” Machine Vision Applications Journal 11 (2000) 267–276.

[5] S. Baker, D. Scharstein, J.P. Lewis, S. Roth, M, J. Black, and R. Szeliski: “A Database and Evaluation Methodology for Optical Flow,” Proceedings of the IEEE International Conference on Computer Vision, 2007.

[6] J. Y. Bouguet: “Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the algorithm,” Intel Corporation, 2003.