Military Technical College
Kobry El-Kobbah,
Cairo, Egypt

ICEENG

6th International Conference
on Electrical Engineering
ICEENG 2008

A NEW TARGET-TRACKING APPROACH IN FLIR IMAGERY USING MONOCULAR THERMAL IMAGE SEQUENCE

By

Gouda Ismail Salama *   Osama Fathy Hegazy **   Wael Mohamed Yousf ***

Abstract:
In this paper, we present a new approach for tracking a moving target in forward-looking infrared (FLIR) imagery taken from an anti-tank moving platform. First, the target image is captured in a frame with resolution 128×128 pixels by using IR camera, and then the image is enhanced by using a non uniformity correction algorithm to reduce the main dominant noises associated with infrared images. Second, the existing moving target is detected with time (τ) within the image by using 2-D Gabor filter kernels. Next, Gabor filter with 4 orientations applying on the window created with resolution 32×32 pixels around the predicted target center using Kalman filter estimator. Subtraction process is performed between the predicted center using Kalman filter estimator and the center of the detected target using Gabor filter, which is used to predict the next center by using Kalman filter, and to guide the detection of the target location in the tracking window of the next frame respectively. The proposed technique fasts and reduces the detection time into 1/16 of τ with high performance of tracking. The experiments performed on several real image data of an existing infrared imaging system. Results show the robustness of the proposed method, which combines high speed of detection and good performance of tracking even with strong ego-motion.

_____

*   Egyptian Armed Forces

1. Introduction:

Detection and tracking of moving or stationary targets in FLIR imagery are challenging research topics in computer vision. Though a great deal of effort has been expended on detecting and tracking targets in visual images, there has been only limited amount of work on thermal images in the computer vision community [13]. Thermal images are obtained by sensing the radiation in the infrared (IR) spectrum, which is either emitted or reflected by the object in the scene. Due to this property, IR images can provide information which is not available in visual images. However, in contrast to visual images, the images obtained from an IR sensor have extremely low signal to noise ratio (SNR), which results in limited information for performing detection or tracking tasks. In addition, in anti-tank missile forward looking infrared (FLIR) images, non-repeatability of the target signature, competing background clutter, lack of a priori information and the artifacts due to the weather conditions make detecting or tracking targets even harder. To overcome the shortcomings of the nature of the FLIR imagery, different approaches impose different constraints to provide solutions for a limited number of situations. For instance, many target detection methods require that the targets are hot spots which appear as bright regions in the images [4]. Similarly, several target tracking algorithms require one or both of the following assumptions to be satisfied: (1) no sensor ego-motion and (2) target features do not drastically change over the course of tracking [13]. However, in realistic tracking scenarios, neither of these assumptions is applicable, and a target tracking algorithm must successfully deal with these problems. Combination kernel tracking with global motion compensation for forward-looking infrared (FLIR) imagery was presented in [12]. Night vision for pedestrian detection and tracking, where the detection is performed by a support vector machine and the tracking is kernel-based was addressed by [15].

In this paper, we present a new approach to detect and track targets in FLIR imagery in the presence of changes in target features, i.e. shape and intensity. Moreover, the targets are not required to have constant velocity or acceleration. Our proposed algorithm is composed of four major modules: Target Acquisition Module (TAM), which is based on acquiring an initial target image with resolution 128×128 pixels (first frame) from the sequences of moving target recording frames using infrared camera. The second module is Target Enhancement Module (TEM) that is based on Non-uniformity correction algorithms, which is used to suppress the main dominant noises associated with thermal images. The third module is Target Detection Module (TDM), which is based on hierarchical Gabor filter approach to select the target manually or automatically and determine the initial center of target for the first frame. The last module is Target Prediction Module (TPM) which is based on Kalman filter that is used to predict the target location in the next frame relaying on the location of the target in the first frame.

The rest of this paper is organized as follows: In Section 2, we review related works and compare the proposed method with existing ones. In Section 3, we describe the Target Acquisition Module (TAM). In Section 4, we introduce the Target Enhancement Module (TEM). In Section 5, we discuss the Target Detection Module (TDM). The Target Prediction Module (TPM) is outlined in Section 6. The experimental results for the proposed algorithm are presented in Section 7 and conclusions are drawn in Section 8.

2.Related Work:

In this section, we examine some of the representative works reported in the literature on detecting and tracking targets in FLIR imagery. In general, existing methods on IR images work for a limited number of situations due to the constraints imposed on the solution. For detection of FLIR targets, many methods rely on the ithot spot technique, which assumes that the target IR radiation is much stronger than the radiation of the background and the noise. The goal of the target detectors is then to detect the center of the region with the highest intensity in the image, which is called the ithot spot. The hot spot seekers use various spatial filters to detect targets in the scene.

Takken et al. [9] developed a spatial filter based on least mean square (LMS) to maximize the signal to clutter ratio for a known and fixed clutter environment. Chen et al. [4] modeled the underlying clutter and noise after local demeaning as a whitened Gaussian random process, and developed a constant false alarm rate detector using the generalized maximum likelihood ratio. Temporal filters like Triple Temporal Filter (TTF), Infinite Impulse Response (IIR) and Continuous Wavelet Transform (CWT) have been widely used. Lim et al. [8] have presented a multistage IIR filter for detecting dim point targets. Tzannes et al. [14] presented a Generalized Likelihood Ratio Test (GLRT) solution to detect small (point) targets in a cluttered background when both the target and clutter are moving through the image scene.

Similar to the target detection methods, target tracking approaches also impose constraints on the solution, such as no sensor ego-motion or no target modal change. However, even with these assumptions, the tracking performance of most methods is not convincing. Below, we will briefly summarize commonly cited methods which have attempted to deal with these problems. To compensate the global motion, Strehl et al. [11] have used a multi-resolution scheme based on the affine motion model. The affine model has its limitations and for FLIR imagery, which is obtained from an airborne sensor; it is unable to capture the skew, pan and tilt of the planar scene.

Similarly, Shekarforoush et al. [10] first compensate for the sensor ego-motion to stabilize the FLIR sequence, then detect very hot or very cold targets. The stabilization and tracking is based solely on the goodness of the detection and the number of targets, i.e. if the number of targets is not adequate, or there is significant background texture, the system is not able to detect sufficient number of targets. Therefore stabilization fails to correctly register the image. Braga-Neto et al. [3] have presented a method based on morphological operators for target detection and tracking in FLIR imagery. Their tracker is based on the assumptions that the targets do not vary in size, they are both very hot or very cold spots, and sensor ego-motion is small. However, these assumptions contradict the nature of airborne FLIR imagery. Davies et al. [6] proposed a multiple target tracker system based on Kalman filters for small targets, which uses the output of the Daubechies wavelet family for FLIR imagery. The method assumes constant acceleration of the target, which is not valid for maneuvering targets. In addition, the method works only for sequences with no global motion.

Yilmaz et al. [13] have presented an approach for real-time target tracking in FLIR imagery in the presence of high global motion and changes in target features, i.e. shape and intensity. The method uses the positions and the sizes of targets determined by the target detection scheme. For target detection, they apply steerable filters and compute texture energies of the targets, which are located using a segmentation-based approach. Once the targets are detected, the tracking method employs three modules to perform tracking. The first module is based on finding the translation vector in the image space that minimizes the distance between the distributions of the model and the candidate. The distributions are obtained from the intensity and local standard deviation measure of the frames. The local standard deviation measure is obtained in the neighborhood of each pixel in the frame and provides a very good representation of frequency content of the local image structure. Based on the distance measure computed from the target feature distributions, the other two modules compensate for the sensor ego-motion and update the target model. The global motion estimation module uses the multi-resolution scheme that assuming a planar scene under perspective projection. It uses Gabor filter responses of two consecutive frames to obtain the pseudo perspective motion parameters.

3. TARGET ACQUISITION MODULE (TAM) :

In the image acquisition phase, infrared camera capture an image (reference image) from the sequence of moving target recording has been prepared to display with resolution 128×128 pixels. The system tracked a missile movement towards anti-tank vehicle over a monocular thermal image of different sequences frames. The difficulties

of this type missile tracking are small and low contrast, ego-motion and short firing distance, which need high performance of detection and high speed of tracking.

4. TARGET ENHANCEMENT MODULE (TEM):

Focal-plane-array (FPA) sensors typically exhibit spatial non-uniformity among the array elements. This non-uniformity results in the so-called fixed pattern noise (FPN). Using Fourier transform of the image, were used to reduce fixed pattern noise measured in the absence of illumination. By taking the logarithm of the image, this noise will be turned into additive noise, which then can be reduced by applying a high pass filter. This is sometimes referred to as homomorph filtering as shown in equation:

$$H(u,v) = (\gamma_H - \gamma_L)[1 - e^{-c(D^2(u,v)/D_o^2}] + \gamma_L \qquad (1)$$

Where:

       c is constant to control the sharpness of the filter function,

       γ_H and γ_L are chosen parameter so that $\gamma_L < 1$ and $\gamma_H > 1$.

5. TARGET DETECTION MODULE (TDM):

FLIR detector arrays generally produce spatially under sampled images, which results in poor, noisy data for further processing, such as target detection, tracking and recognition. Besides the limited capabilities of FLIR sensors, the images obtained using them often contain artifacts such as dirt on the sensor and fading brightness along with limited viewing resulting form cloudy weather. Moreover, FLIR sequences for anti-tank missile, suffer from abrupt discontinuities in motion. Targets in FLIR sequences can appear as either cold or hot spots due to thermal electro-magnetic spectrum. However, there is no a priori information about the targets temperature or shape. Due to target's low contrast with the background, our system uses both enhanced images and images filtered by 2D Gabor filter kernels, which are oriented sine-wave gratings that are spatially attenuated by a Gaussian window. Two-dimensional Gabor filters, which have also been used for object detection in FLIR images [2], have the form

$$Gi(x, y) = \exp\{-[1/(4\pi\sigma^2)][x'^2 + \alpha^{-2}y'^2]\} * \exp\{j\omega i[\cos\Phi i + y\sin\Phi i]\}. \qquad (2)$$

Where
      (x , y)……..are the horizontal and vertical image coordinates, respectively
      x' , y' ……..the axes that are rotated by ϕ_G

ωi , Φi …….modulation frequency and orientation, respectively
σ , α …….scale and eccentricity of the elliptical Gaussian window

We used 4 directions, 0, 45, 90, 135 and a fixed scale, 0.5 for our experiments. The responses in four directions are summed for obtaining the input image to the tracker. In Figure 1 (a), a frame from one of the FLIR sequences and in (b) sum of Gabor responses are shown. In Figure 1(b), the target regions are clearly emphasized.

Detection of targets in the FLIR sequences is a hard problem because of the variability of the appearance of targets due to atmospheric conditions, background, and thermodynamic state of the targets. The challenges in this regard are extremely low SNR, non-repeatability of the target signature, competing clutter, which forms similar shapes as those of the actual targets, obscured targets and lack of a priori information. In this paper, we assume that the location of target in the initial frame is given. In the next section, we will give details of our model for tracking the target in two consecutive frames, which will be generalized to track the target in the whole sequence.

(a)   (b)
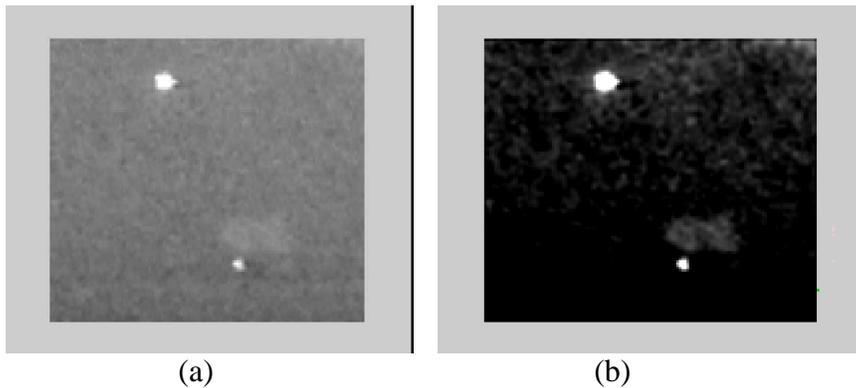
Figure(1) *(a) Sample frame from one of the FLIR sequences, (b) summation of four Gabor response of the frame in (a).*

Center of gravity
From the image processing point of view, the center of gravity of an image-also known as the "centroid" can be defined as being a pixel inside the object border located as far as possible from all border pixels, and it can be easily calculated by averaging the x and y coordinates of the object pixels as shown in the following equations (3) using MATLAB software we create a program to calculate the center of target in 2D image (x_c , y_c)

$$x_c = \sum {x_i}/{n}$$
$$y_c = \sum {y_i}/{n}$$

(3)

Where:

n ………………… is numbers of object pixels
(x_c , y_c) ……,,……is center point of the target

6. TARGET PREDICTION MODULE (TPM):

The Kalman filter estimates a process by using a form of feedback control: The filter estimates the process state at some time and then obtained feedback in the form of noisy measurement [8]. The Kalman filter addresses the general problem of trying to estimate the state $x \in R^n$ of a discrete-time controlled process that is governed by the linear stochastic difference equation

$$x_{k+1} = A_k x_k + Bu_k + w_k \tag{4}$$

With a measurement $z \in R^m$ that is

$$z_k = H_k x_k + v_k \tag{5}$$

The random variable w_k and v_k represent the process and measurement noise, respectively. $\hat{x}_k^- \in R^n$ (note the "super minus") is defined to be a priori state estimate at step k given knowledge of the process prior to step k, and $\hat{x}_k \in R^n$ to be the posteriori state estimate at step k given measurement z_k. Then a priori and posteriori estimate errors can be defined as:

$$e_k^- \equiv x_k - \hat{x}_k^-, and$$
$$e_k \equiv x_k - \hat{x}_k$$

The a priori estimate error covariance is then,

$$P_k^- = E[e_k^- e_k^{-T}] \tag{6}$$

And the a posteriori estimate error covariance is:

$$P_k = E[e_k e_k^T] \tag{7}$$

The goal is to find an equation that computes an a posteriori state estimate $\hat{x}_k$ as a linear combination of an a priori estimate $\hat{x}_k^-$ and a weighted difference between an actual measurement z_k and measurement prediction $H_k \hat{x}_k^-$ as shown below in equation (8). Some justification for equation (8) is given in "the probabilistic origins of the filter" found below

$$\hat{x}_k = \hat{x}_k^- + k(z_k - H_k \hat{x}_k^-) \tag{8}$$

One form of the resulting K that minimizes equation (7) is given by:

$$K_k = p_k^- H_k^T \left(H_k p_k^- H_k^T + R_k \right)^{-1} = \frac{p_k^- H_k^T}{H_k p_k^- H_k^T + R_k} \tag{9}$$

The typical tracking paradigm comprises a repeating cycle of measurement process and prediction as shown in Figure 2. The measurement process, in turn, invokes both target detection using Gabor filter and calculation of target centroid. In the prediction phase, a feedback loop is established by predicting the new locations with Kalman filters and using them to guide and fast the detection of the target location in the next frame.

Proposed Algorithm:
1- Start: Z = Empty, TW (Tracking Window) = Empty
2- Acquire the first frame
3- Z = first frame
4- Apply enhancement filter on Z
5- Apply 2-D Gabor filter to detect target within Z
6- Get center of the target within Z
7- Apply Kalman filter to predict the next center within Z
8- Acquire New frame
9- Z = New frame
10- Create a TW with 32X32 resolution at the predicted center within Z
11- Z = TW
12- Is end of track, If YES goto STOP, otherwise Step 4

Fig. 2 The proposed algorithm

7. EXPERIMENTAL RESULTS:

We have applied the proposed method to FLIR dataset for tracking targets. The data set was made available to us in grayscale format and was composed of 30 sequences where each frame was 128x128. The proposed method was developed using MATLAB running on Windows XP on a Pentium 4 platform and current implementation of the algorithm is capable of tracking one target at a time at 15 frames per second. The tracking results of the algorithm were visually confirmed.

Figure 3 shows a set of selected frames from a sequence of anti tank missile where there is high but smooth global motion, i.e. the global motion compensation module is not executed.

(a)   (b)   (c)   (d)

(e)   (f)   (g)   (h)
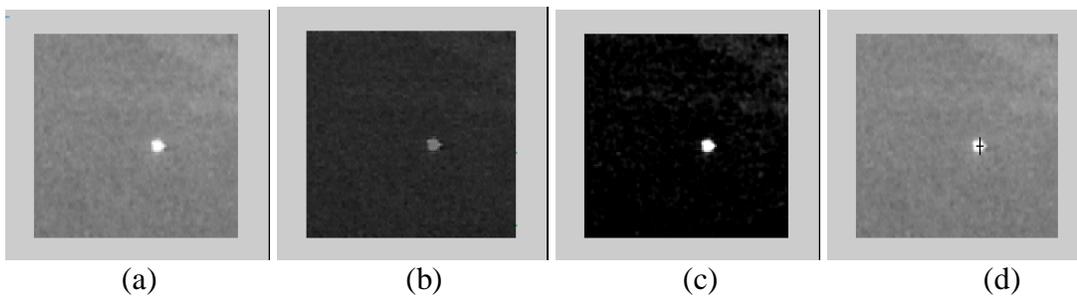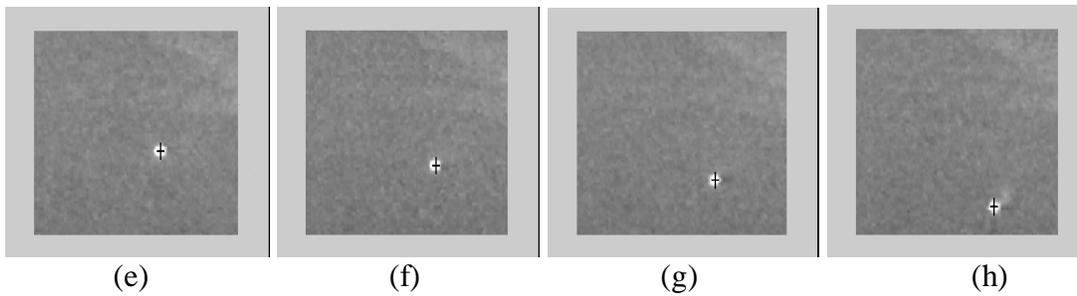
Figure(3) *Target tracking result for a sequence; (a) original image, (b) enhanced image, (c) applied Gabor filter for target detection, automatic tracked target shown by '+' in (d) frame 4, (e) frame 8, (f) frame 12, (g) frame 16, (h) frame 20.*

The original image is treated with non-uniformity correction algorithm then Gabor filter kernel used for target detection by contrast the brightness of the target, target center were calculated using Matlab program then, these calculation were the initial target position (a prior state) to Kalman filter estimator to predict the target position in the next frame (a posterior state). Target, is detected and the black cross in the other frames show the tracked target center. The target tracker correctly tracked the target.

Figure 4, showing the motion of the target center in (X-Y) directions using Kalman filter estimator, it is clear that both of measured values of target center approx. the same to the Kalman filter predicted values for the target center.

Figure 5, showing a plot of Root Mean Square Error (RMSE) values resultant from applying Kalman filter predictor for target tracking.

(a)   (b)
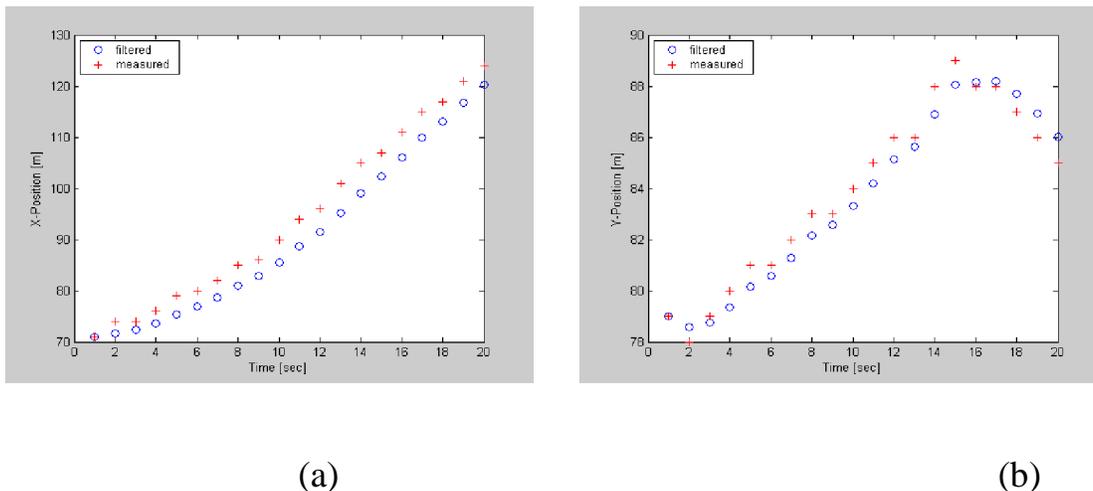
Figure(4) *Example plot of predicted center using Kalman filter with measured value of target center. (a) In y-direction. (b) In x-direction.*
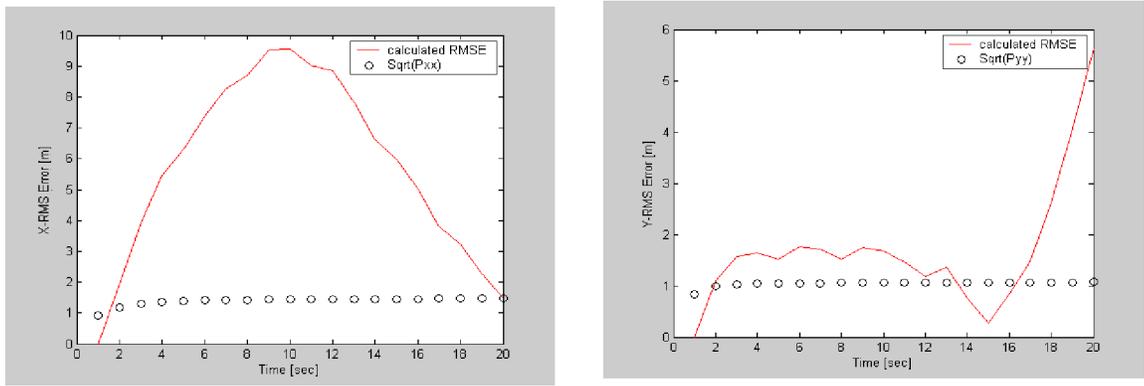
(a)   (b)

Figure(5) *Example plot of RMSE resultant from applying Kalman filter predictor. (a) In X-direction. (b) In Y-direction.*

Figure 6 shows a set of selected frames from a sequence of missile moving towards anti-tank. The original image were treated by NUC algorithm then the target center (missile) were detected manually to initiate Kalman filter estimator as a prior information to detect the motion of the target in the next frame and for the whole sequences. Then Gabor filter were applied on the window size of the target to ensure the efficiency of Kaman prediction result, and compensate the sensor ego- motion (global) compensation. The window size around the target was selected automatically with 32 x 32. Target, is detected and the black cross in the other frames show the tracked target center. The target tracker correctly tracked the target. It is shown from the result of applying Gabor filter on the window size 32×32 to evaluate the center of gravity of the moving target that as follows:

$$XC = 16 \pm 3, \ YC = 16 \pm 3$$

That means that the center of gravity of moving target is approximately coincide with the predicted center using Kalman filter predictor.
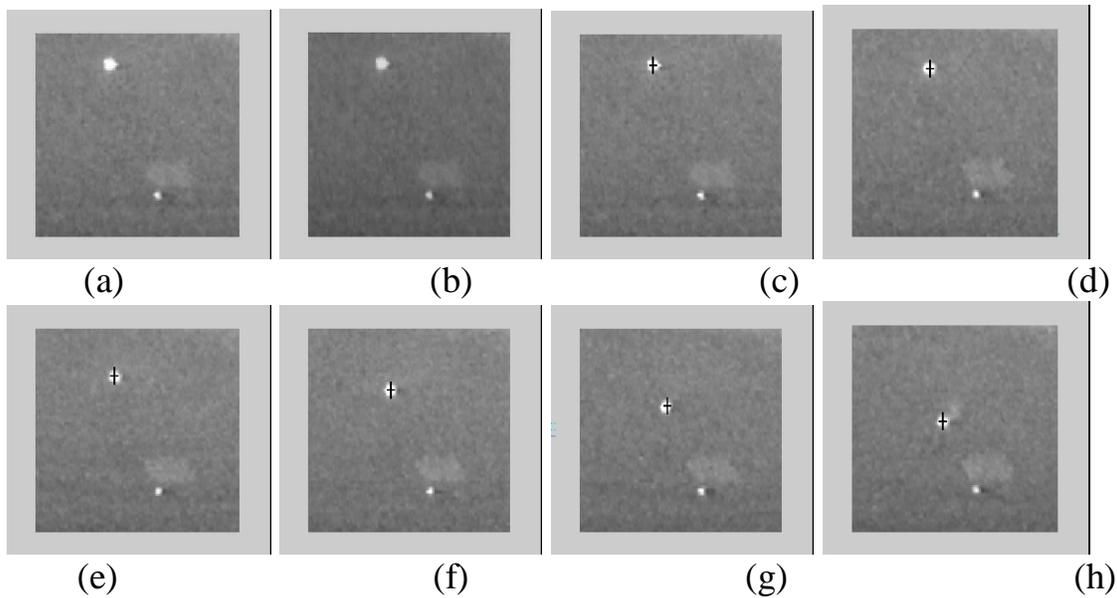
Figure(6) *Target tracking result for a sequence; (a) manually initialized target for frame 0, (b) enhanced image, automatic tracked target shown by '+' in (c) frame 4, (d) frame 8, (e) frame 12, (f) frame 16, (g) frame 20, (h) frame 25*
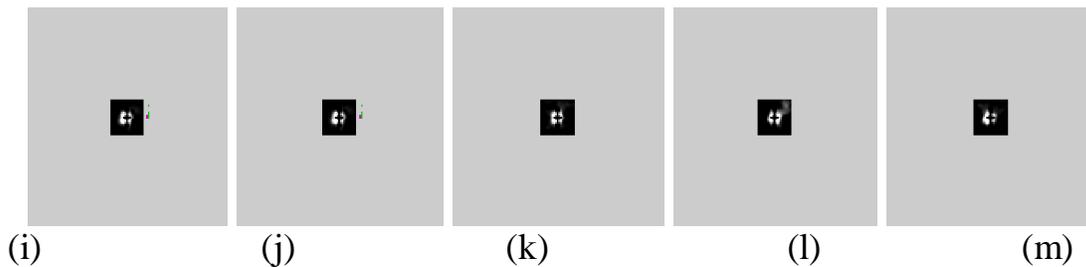
Figure(7) *Continue, Applying Gabor filter on the window size to evaluate the center of gravity of target center of first frame. (i), (j), (k), (l), (m) Applying Gabor filter on the window size to evaluate the center of gravity of target center of frame no. (4, 8, 16, 20, 25) respectively.*

Figure 7, showing the motion of the target center in (X-Y) directions using Kalman filter estimator, it is clear that both of measured values of target center approx. the same to the Kalman filter predicted values for the target center.

Figure 8, showing a plot of root mean square error (RMSE) values resultant from applying Kalman filter predictor for target tracking.
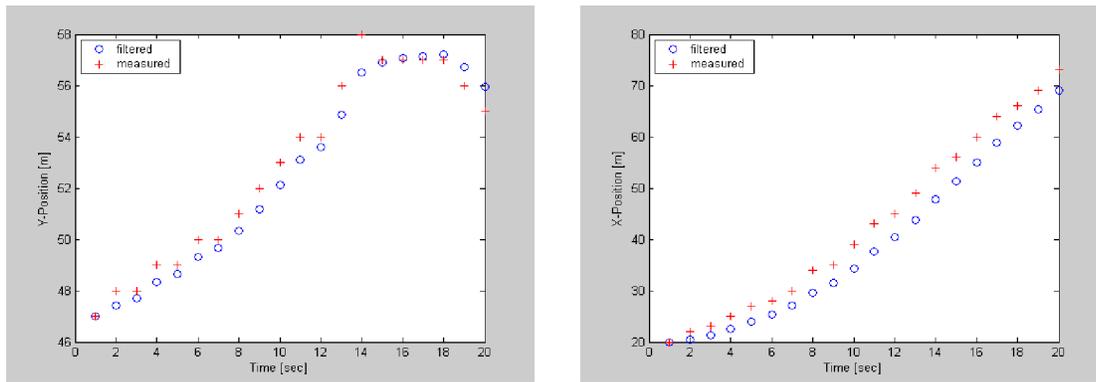
<center>(a)   (b)</center>

Figure(8) *Example plot of predicted center using Kalman filter with measured value of target center. (a) In y-direction. (b) In x-direction.*

Figure 10 shows a set of selected frames from a sequence of anti tank missile where there is high but smooth global motion, i.e. the global motion compensation module is not executed. The original image is treated with non-uniformity correction algorithm then Gabor filter kernel used for target detection by contrast the brightness of the target, target center were calculated using Matlab program then, these calculation were the initial target position (a prior state) to Kalman filter estimator to predict the target position in the next frame (a posterior state). Target, is detected and the black cross in the other frames show the tracked target center. The target tracker correctly tracked the target.
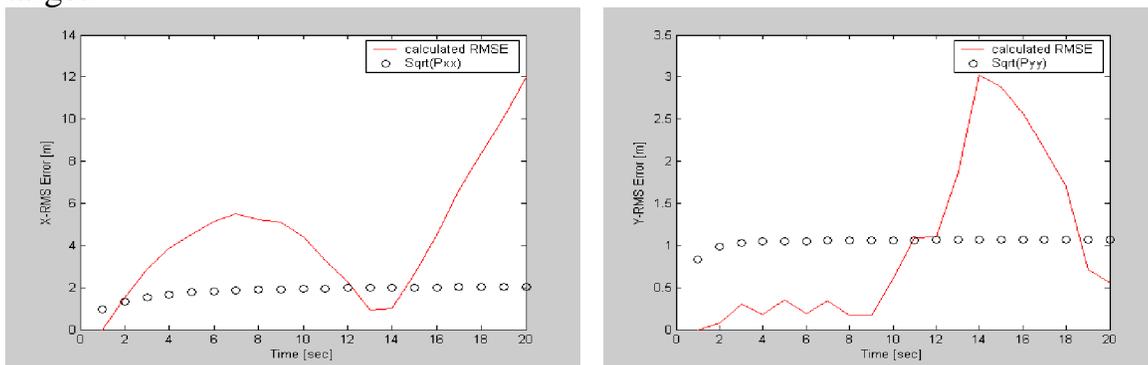
<center>(a)   (b)</center>

Figure (9) *Example plot of RMSE resultant from applying Kalman filter predictor. (a) In X-direction. (b) In Y-direction.*

Figure 11, showing the motion of the target center in (X-Y) directions using Kalman

filter estimator, it is clear that both of measured values of target center approx. the same to the Kalman filter predicted values for the target center. Figure 11, showing a plot of root mean square error (RMSE) values resultant from applying Kalman filter predictor for target tracking.
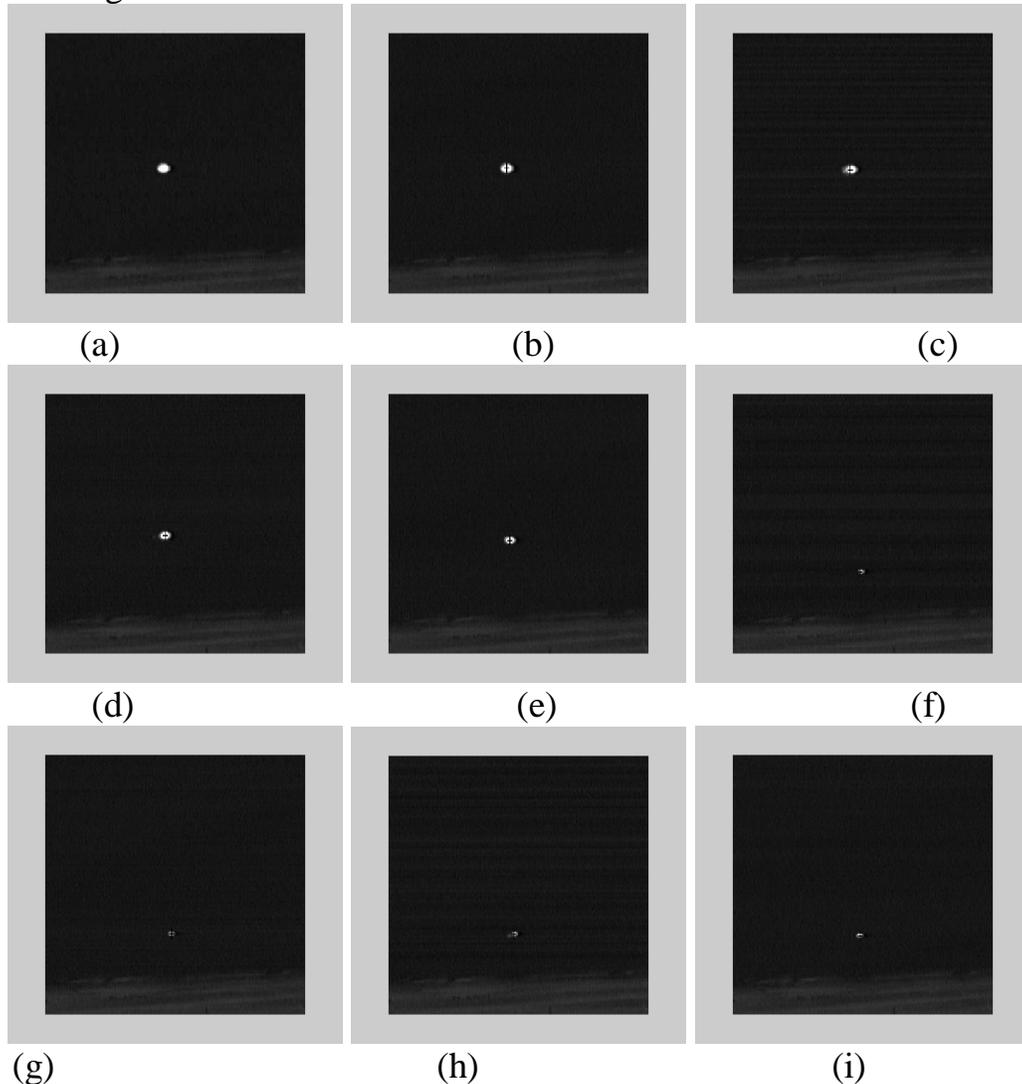
Figure(10) *Target tracking result for a sequence; (a) first frame (original), automatic tracked target shown by '+' in (b) frame 2, (c) frame 5, (d) frame 8, (e) frame 11, (f) frame 60, (g) frame 65, (h) frame 70, (i) frame 100.*
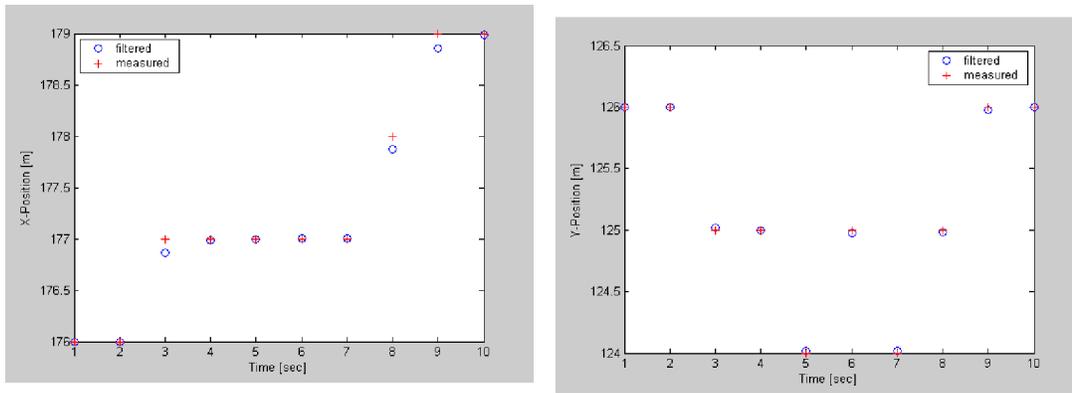
(a)   (b)

Figure(11) *Example plot of predicted center using Kalman filter with measured value of target center. (a) In x-direction. (b) In y-direction.*
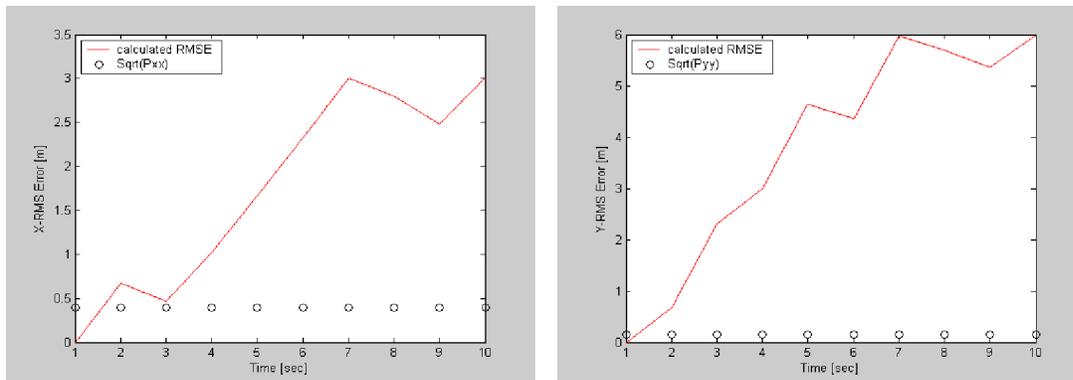
(a)   (b)

Figure(12) *Example plot of RMSE resultant from applying Kalman filter predictor.*
(a) In X-direction. (b) In Y-direction.

8. Conclusions:

The proposed algorithm combines the discriminating power of the wavelet transform with a Kalman filter to detect and track the motion of small, low contrast and high speed objects in image sequences. We have demonstrated the proposed algorithm using a

sequence of FLIR images which incorporate slow and fast motion of an approaching missile. Results show the robustness of the proposed method, which combines high speed accurate tracking with good performance of detection even with strong ego

References:

[1] *Bergen J.R., Anandan P., Hanna K.J., Hingorani R., "Hierarchical model-based motion estimation", In: Europen Conference on Computer Vision, pp. 237– 252, (1992).*

[2] *Braithwaite R.N., Bhanu, B.,"Hierarchical Gabor filters for object detection in infrared images," Proceedings of IEEE Conf. on Computer Vision Pattern Recognition, pp. 628- 631, (1994).*

[3] *Braga-Neto U., Goutsias J., "Automatic target detection and tracking in forward-looking infrared image sequences using morphological connected operators", In: 33rd Conference of Information Sciences and Systems, March, (1999).*

[4] *Chen J.Y. , Reed I.S., "A detection algorithm for optical targets in clutter", IEEE Transactions on Aerospace and Electronic Systems 23 (1) pp. 46–59, (1987).*

[5] *Comaniciu D., Ramesh V., Meer P., "Real-time tracking of non-rigid objects using mean shift", In: IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 142–149, (2000).*

[6] *Davies, D., Palmer, P., Mirmehdi ,"Detection and Tracking of Very Small Low contrast objects", Proceedings of the 9th British Machine Vision Conference, Sept. (1998).*

[7] *Gouda I. Salama, Mohamed E. Ghoneimy, and Wael Mohamed Yousef, "Performance improvement of thermal images," 5th International conference on Electrical Engineering, ICEENG 2006, PP. RE-3-1: RE-3-13, 16-18 Nov. (2006), Military Technical College , Cairo , Egypt.*

[8] *Lim E.T., Deshpande S.D., C.W. Chan, R. Venkateswarlu, "Adaptive spatial filtering techniques for the detection of targets in infrared imaging seekers", In: Proceedings of SPIE, vol. 4025, pp. 194–202, (2000).*

[9] *Longmire M.S., Takken E.H., "LMS and matched digital filters for optical clutter suppression", Applied Optics, 27 (6), 1141–1159, (1988).*

[10] *Shekarforoush H., Chellappa R., "A multi-fractal formalism for stabilization, object detection and tracking in FLIR sequences", In: IEEE International Conference on Image Processing, vol. 3, (2000).*

[11] *Strehl A., Aggarwal J. K., "Detecting moving objects in airborne forward looking infrared sequences," Machine Vision Applications, Vol. 11, pp. 267-276, (2000).*

[12] *Yilmaz A., Shafique K., Olson T., Lobo N. and Shah M., "Target Tracking in FLIR Imagery Using Mean-Shift and Global Motion Compensation," proceedings of IEEE Workshop on Computer Vision Beyond Visible Spectrum (CVBVS), Hawaii, (2001).*

[13] *Yilmaz A., Shafique K. and Shah M., "Target tracking in airborne Forward looking Infrared imagery", Image and Vision Computing 21 pp. 623–635, (2003).*

[14] *Tzannes A.P. , Brooks D.H., "Detection of point targets in image sequences by hypothesis testing: a temporal test first approach", In: Proceedings of ICASSP, 1999.*

[15] *Xu F., and Fujimura K., "Pedestrian Detection and Tracking with Night Vision," Proc. IEEE Intelligent Vehicle Symp. (2002).*