Regular Paper

Small Target Detecting and Tracking Using Mean Shifter Guided Kalman Filter

Soo-Young Ye
Department of Radiological Science, Catholic University of Busan, Busan 609-757, Korea

Jae-Heum Joo
Department of Multimedia Engineering, Catholic University of Busan, Busan 609-757, Korea

Ki-Gon Nam[†]
Department of Electronics Engineering, Pusan National University, Busan 609-735, Korea

Because of the importance of small target detection in infrared images, many studies have been carried out in this area. Using a Kalman filter and mean shift algorithm, this study proposes an algorithm to track multiple small moving targets even in cases of target disappearance and appearance in serial infrared images in an environment with many noises. Difference images, which highlight the background images estimated with a background estimation filter from the original images, have a relatively very bright value, which becomes a candidate target area. Multiple target tracking consists of a Kalman filter section (target position prediction) and candidate target classification section (target selection). The system removes error detection from the detection results of candidate targets in still images and associates targets in serial images. The final target detection locations were revised with the mean shift algorithm to have comparatively low tracking location errors and allow for continuous tracking with standard model updating. In the experiment with actual marine infrared serial images, the proposed system was compared with the Kalman filter method and mean shift algorithm. As a result, the proposed system recorded the lowest tracking location errors and ensured stable tracking with no tracking location diffusion.

Keywords: Mean shifter guided kalman filter, Infrared images, Target detection

1. INTRODUCTION

Because of the importance of small target detection in infrared images, many studies have been carried out in this area. Some of the techniques used to estimate a background and obtain the target location include a linear diffusion filter, anisotropic diffusion filter [1], maximum likelihood filter [2], mean shift filter [3], and two-dimensional LMS (TDLMS) filter [4]. It is very difficult to detect a small target in infrared still images with a complex background because of the irregular noises and low SNR [5]. Any effort to detect a small target in infrared still images with a marine background inevitably faces some limitations since clouds or waves are included and appear with a strong noise. Research has been conducted on tracing a target in serial images with the target location obtained in still images in order to resolve such noises [6-9]. This research employed conventional tracing techniques to take advantage of noise irregularity in serial images including a Kalman filter [10], extended Kalman filter [11], and particle filter [12]. These algorithms demonstrate excellent performance in real-time tracing but have some limitations with the application to image processing due to the weak associations among targets in serial images with noises and background interference. Some methods can be used to resolve those problems such as the Probabilistic Multiple Hypothetic Tracking (PMHT)

[†] Author to whom all correspondence should be addressed:
E-mail: kgnam@pusan.ac.kr

188

Trans. Electr. Electron. Mater. 14(4) 187 (2013): S.-Y. Ye *et al.*

[13] algorithm and the mean shift algorithm [14]. The PMHT algorithm tracks a target by associating small targets in serial images based on the EM algorithm [15] and has a disadvantage of requiring preliminary information on the number of targets. Research has been carried out to help to overcome the disadvantage [16]. The mean shift algorithm is a powerful method used to find extreme values in the dense distribution of data sets, demonstrating great performance in moving nonlinear targets and powerful associations among targets in serial images. However, it also has some disadvantages including requiring early information on targets and research in the case of target appearance and disappearance. This study proposes an algorithm to track multiple small moving targets even in the case of target disappearance and appearance in serial infrared images in an environment with many noises by using a Kalman filter and mean shift algorithm.

Fig. 1. Overview of system block diagram.

2. PROPOSED MULTIPLE TARGET DETECTING AND TARCKING TECHNIQUE

Target detection in still images has a limitation of increasing errors detected when the image has many clutters. Such a limitation is resolved by using the non-uniformity of noise information, which includes location, shape, and brightness, in serial images as important information for target detection. Figure 1 shows the system flow chart of the proposed multiple target tracking technique. F refers to an input image, k is the time index, X^I is the location of a target detected in a still image, X^f is the location of a target being newly tracked, X^t is the location of a target that has been tracked, and H is the histogram model information. The proposed systems can be categorized into three stages: the first stage detects the locations of candidate targets in still images, including small target detection and background estimation filter blocks; the second stage classifies target candidates according to noises or targets by estimating and comparing target locations with a Kalman filter, including multi-target tracking, target selection, and target position prediction blocks; and the final stage fine-tunes each target location in serial images with NCC model matching and mean shift algorithm, including target position adjustment, local histogram production, target model update, NCC model matching, and mean shift adjustment blocks.

2.1 Background estimation filters

The study first estimated image backgrounds in order to detect small targets in still infrared images in a much cluttered environment by introducing and utilizing the existing filters. The study applied and compared several filters including a linear diffusion filter, anisotropic diffusion filter, maximum likelihood filter, mean shift filter, and 2D LMS filter to background estimation, assessing their performance. The filter with the highest performance was used to detect candidate targets in a frame for multiple target tracking.

2.1.1 Linear diffusion filter

A linear diffusion filter determines a pixel value by combining surrounding pixel values linearly. When filtered and processed in such a way, images become blurry as if a low pass filter had been used. The diffusion equation of a linear diffusion filter is shown in equation (1):

$$\frac{\partial f(x,y,t)}{\partial t} = div(f(x,y,t)) \tag{1}$$

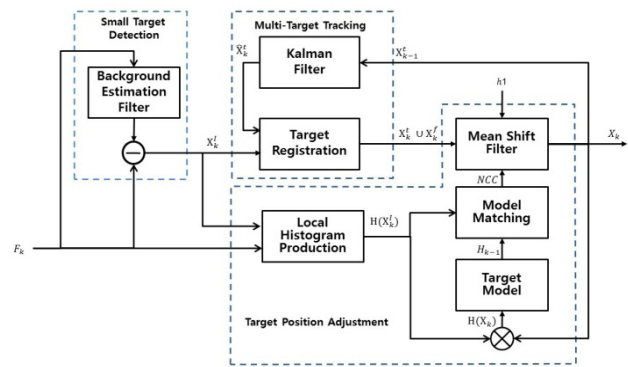$f(x,y,t)$ is the number of inputs repeated, and t is the number of diffusion repeated. The linear diffusion filter based on the discrete expression of equation (1) has the following equation (2):

$$f^{n+1}(x,y) = \sum_{(i,j)\in N} f^n(x+i, y+i) \tag{2}$$

where N is the neighbor of (x,y), $f^{n+1}(x,y)$ is a diffusion image, and $f^n(x,y)$ is an input image. The $f^{n+1}(x,y)$ of a diffusion image is obtained by combining the current pixels and their surrounding pixels.

2.1.2 Anisotropic diffusion filter

An anisotropic diffusion filter varies the filtering degrees of targets and backgrounds as a nonlinear filter to avoid the local, blurry problems of a linear diffusion filter. An anisotropic diffusion equation is expressed in equation (3):

$$\frac{\partial f(x,y,t)}{\partial t} = div\{c(|\nabla f(x,y,t)|)\nabla f(x,y,t)\} \tag{3}$$

where $f(x,y,t)$ is a parameter to determine input reckoning; t is a parameter to determine diffusion reckoning; $\nabla f(x,y,t)$ is an inclination according to directions; and $c(l)$ a weighting function. They are defined as shown in equation (4):

$$c(l) = \exp\left\{\frac{-l^2}{k^2}\right\} \tag{4}$$

This is a decreasing function with $c(l)$ converging toward 0 when l increasing to infinity. When the inclination becomes large, diffusion ceases. An anisotropic diffusion equation can be expressed in a discrete equation (5):

$$f^{n+1}(x,y) = f^n(x,y) + \sum_{(i,j)\in N}\{c(\nabla f^n(x+i, y+i))\nabla f^n(x+i, y+i)\} \tag{5}$$

Here, $f^{n+1}(x,y)$ is a diffusion image and $f^n(x,y)$ is an input image.

2.1.3 Maximum likelihood filter

A maximum likelihood filter estimates a parameter with some samples and uses it for image segmentation. The pixels of segmented images are used to determine the mean value of each

area. A likelihood function to estimate a parameter is shown in equation (6):

$$l = \sum_{k=1}^{m} \log(P(x_k | \theta_j))$$
$$S_k = \arg \max(l) \tag{6}$$

Equation (6) is used to estimate a parameter, which is, in turn, used to segment areas according to the classes of which the likelihood functions are a maximum. The means of segmented areas are obtained as in equation (7):

$$f(x,y) = \frac{1}{N_{S_k}} \sum_{f(i,j)} f(i,j) \tag{7}$$

where N_{Sk} is the number of pixels in each class and S_k represents a segmented class. The pixels of the resulting images are determined with the means of segmented areas to estimate a background.

2.1.4 Mean shift filter

A mean shift filter calculates a mean shift vector to obtain the mean locations of pixels that have a similar brightness distribution A mean shift filter evens out the brightness values while conserving the image boundaries. The mean shift vectors of pixel values are calculated as shown in equation (8):

$$Y = \begin{pmatrix} x_m \\ y_m \end{pmatrix} = \frac{\sum_{(i,j) \in N} X_{i,j} w_{i,j} f(i,j,l)}{\sum_{(i,j) \in N} f(i,j,l)} \tag{8}$$

where x_m, y_m are the mean shift vectors; w is a weight; f is a Kernel function; l is the brightness value of (i,j); and N is a window size. The mean locations shift according to such mean shift vectors. Repetition continues until the mean locations converge and are then replaced with the mean brightness value in the areas with the same convergence locations as those of the mean shift vectors.

2.1.5 2D LMS filter

A 2D LMS(TDLMS) filter is an extended version used to process first-dimensional LMS in a two-dimensional image. Such an s filter predicts the pixels of the resulting images based on the window weights. Errors are calculated by comparing the predicted pixels with the desired results as in equation (9):

$$e^k = d(x,y) - f(x,y) \tag{9}$$

where $d(x,y)$ is the desired result, and $f(x,y)$ is the predicted result. A TDLMS filter ensures that the square value of the errors obtained in equation (9) will be a minimum. The weights that should be adjusted for updating are shown in equation (10):

$$w^{(k+1)}(i,j) = w^k(i,j) +$$
$$\mu e^k f(x-i, y-i), (i,j) \in N \tag{10}$$

where μ is a step size; N is a window size; and e^k is a prediction error from equation (9). Equation (10) shows how to obtain desired resulting images by repeatedly calculating and updating weights as in equation (11) with errors between the desired results and predicted results.

$$f(x,y) = \sum_{(i,j) \in N} w_k(i,j) g(x-i, y-i) \tag{11}$$

where $g(x,y)$ is an input image; $f(x,y)$ is a predicted image; w_k is a weight matrix repeated k times; and N is a window size.

2.2 Small target detection

Difference images, which highlight the background images estimated with a background estimation filter from the original images, have a relatively very bright value, which becomes a candidate target area. Background estimation results are compared and analyzed by obtaining the means of the target areas of difference images and those of the remaining areas and comparing their performance. The filter with the highest performance is used to detect a target candidate for multiple target tracking.

2.3 Multiple target tracking technique

Multiple target tracking consists of a Kalman filter section (target position prediction) and candidate target classification section (target selection) in a system flow chart. The system removes error detection from the detection results of candidate targets in still images and associates targets in serial images. Target association refers to making a judgment to see whether a random target i detected in Image k corresponds with a random target detected in Image $k+1$. The Kalman filter used in the process can produce an error when the target path of the first order Kalman filter is nonlinear. However, the movement scope of each target is very narrow per image in the case of small infrared targets. The targets can thus be modeled linearly. Errors can also be adjusted using the accuracy enhancement technique for target tracking in the entire system.

At the stage of detecting the locations of candidate targets in still images, targets are separated from backgrounds with the critical values of the original images and the difference images from the images obtained through a background estimation filter to acquire target locations. At the stage of classifying target candidates according to noises or targets, we calculate the candidate target location $(\widehat{X_k^t})$ predicted in Frame k with the path of the candidate target location (X_{k-1}^t) tracked in Frame k-1 by using a Kalman filter. In such a case, parameters such as a Kalman coefficient are updated and preserved with a target index. For classification, the candidate targets are classified into "candidate targets first detected (X^f)" and "candidate targets that have been tracked (X^f)" with the locations (X^I) of targets detected in still images and the predicted locations $(\widehat{X^t})$ of targets that have been tracked. When the location information of targets detected in still images matches the predicted location information of targets that have been tracked, they are classified as "candidate targets that have been tracked (X^f)." When there is no match between them, they are classified as "candidate targets first detected (X^f)." The remaining are classified as "candidate targets detected with an error (X^e)." These categories are shown in equation (12) and Fig. 2 below:

$$X^I - \widehat{X^t} = X^f$$
$$X^I \cap \widehat{X^t} = X^t$$
$$\widehat{X^t} - X^I = X^e \tag{12}$$

When candidate targets are classified as "candidate targets that have been tracked X^f" according to the classification method, the reliability (C_l) of candidate targets being judged to be targets increases. When they are classified as "candidate targets detected with an error (X^e)", the reliability (C_l) decreases.

190

Trans. Electr. Electron. Mater. 14(4) 187 (2013): S.-Y. Ye *et al.*
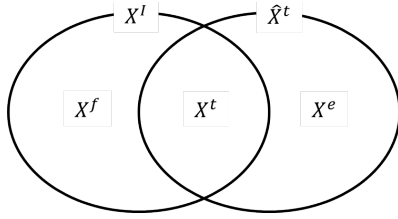

Fig. 2. Classification of candidates.

2.4 Accuracy enhancement technique of target tracking

The accuracy of target tracking can be enhanced by local histogram production, target model update, NCC model matching, and mean shift adjustment in a system flow chart. The system fine-tunes the locations of targets detected in still images and the locations of targets predicted with a Kalman filter, thus improving the stability of a target tracking system.

The stage of the fine-tuning target locations uses input images(F_k) and target locations(X_k^l) and produces local histogram models according to equation (13) below. The number of probing histogram models produced is determined by target locations(X_k^l) and M×M the certain area size in the surroundings.

$$H_l = \frac{1}{N \times N} \sum_{n=1}^{N \times N} \delta[S(X_n - l)] \tag{13}$$

where, N is the size of a local area needed to produce a histogram model; $S(X)$ is the brightness value at Location X; and l is a histogram level and becomes L at 1 that is the scope of brightness value. NCC (normalized cross-correlation) is used to estimate the similarity between the produced histogram model and the standard histogram model produced at k-1. In such a case, NCC uses the following equation (14):

$$NCC = \sum_{l=1}^{L} \sqrt{H_l \times T_l} \tag{14}$$

where H is a probing histogram, and T is a standard histogram. Since targets are associated with each other, a target location (X_k^l) corresponds to a standard histogram. The number of calculated NCC coefficients is determined using M×M, the number of probing histograms. A produced NCC map is used as a weight for the mean shift algorithm. On an NCC map, the mean shift vector of the mean shift algorithm is calculated in the direction of high NCC. In such a case, the maximum local value of an NCC map is obtained by carrying out a mean shift repeatedly until the mean shift vector becomes 0. The mean shift vector is shown in equation (15) below:

$$Y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i g\left(\left\|\frac{Y_0 - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\|\frac{Y_0 - x_i}{h}\right\|^2\right)} \tag{15}$$

where $g( )$ is the derivative function of a Kernal function. Each target produces the local extreme value of NCC found with the mean shift algorithm. As it approaches increasingly closer to 1, the reliability of being judged to be targets (C_2) increases. The location of extreme values becomes X_k, which is the result of the entire system. Of the M×M histograms $H(X_k^l)$ that correspond to each target X_k^l in the last stage of Hour k, the histogram that falls
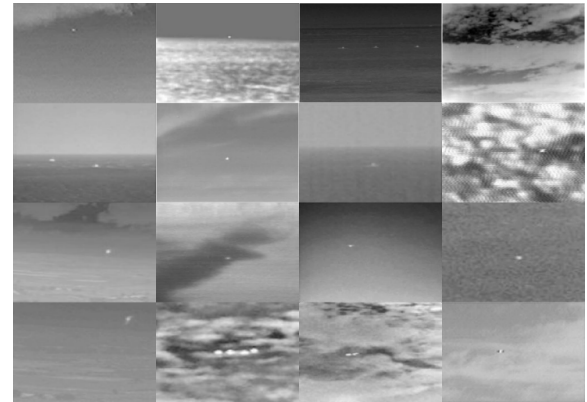

Fig. 3. Performance comparison of each background estimation filter in test images.

Table 1. Mean value of difference between the original and the filtered images in the background area and the target area.

	Linear Diffusion	Anisotropic Diffusion	Maximum Likehood	Mean Shift	TDLMS
Background	35.67	35.67	7.40	2.94	6.26
Target	39.79	39.79	19.97	25.62	27.93

into the location of the result (X_k) is updated as a standard histogram model.

3. EXPERIMENT AND REVIEW

An experiment was conducted to compare the performance of background estimation filters needed in the entire system.

The performance of background estimation filters is determined by comparing the difference operation means of the target areas and those of the remaining areas in the difference operation results of the background and original images calculated with each background estimation filter. Figure 3 shows images included in the target and background.

When the difference means of the target areas are high in the result images of the difference operation and have a greater difference from those of the background areas, the performance of the background estimation filter is considered to be excellent. Since the remaining areas other than the target areas can be seen as noises in the difference operation results, the ratio between target and background in Table 1 can be the signal-to-noise ratio (SNR). The experiment results reveal that the mean shift filter (MSF) had the highest SNR and was thus considered to be superior. It was used as a target candidate detection filter in a frame in the object tracking process.

The tracking performance of the proposed tracking system was examined by implementing and comparing the first order Kalman filter and the mean shift algorithm only. Thirty marine infrared serial images were used in the comparison, with four small targets that had a still image for one second. The proposed system was capable of processing when a new target was produced and lost, but the other systems in the comparison equation were not capable of such processing and accordingly used serial images with no target disappearance and production. The proposed system used a mean shift filter that demonstrated relatively good performance in the experiment for background estimation.

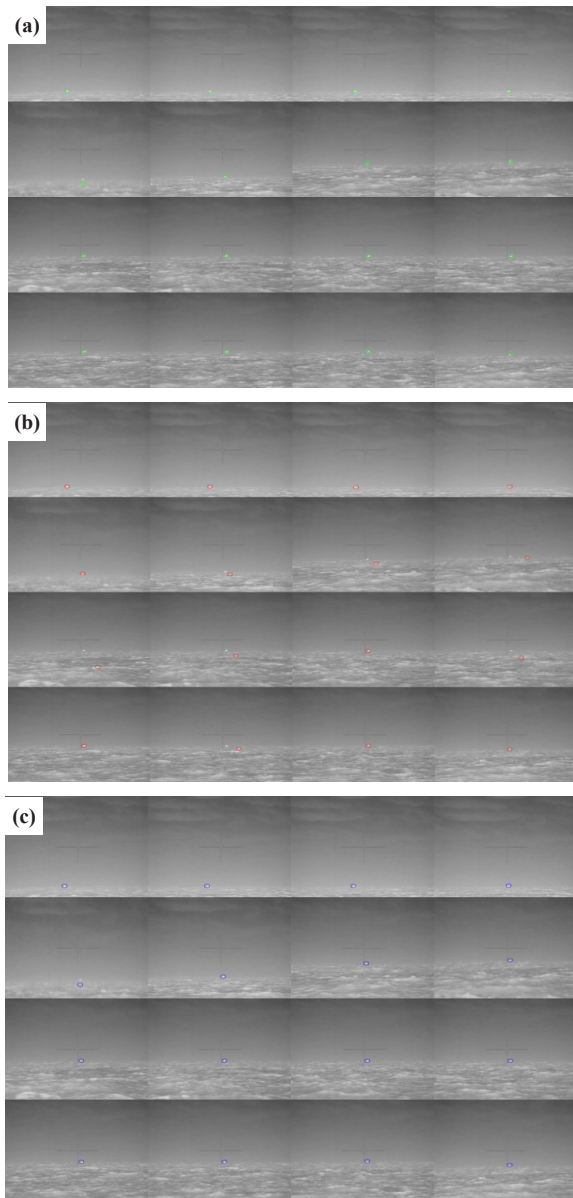Judging the error detection of candidate targets or the assess-

Trans. Electr. Electron. Mater. 14(4) 187 (2013): S.-Y. Ye *et al.*

191

(a)

(b)

(c)

Fig. 4. Image of test result (a) using Kalman filter, (b) mean shift tracker, and (c) proposed method.

Fig. 5. Comparison of the position error of target .

4. CONCLUSIONS

The paper proposed a system to track the locations of multiple small targets in serial infrared images by using a background estimation filter, Kalman filter, and mean shift algorithm. The study compared and assessed various background estimation filters for performance and used a Kalman filter and candidate target classification to allow for target tracking and ensure no location tracking diffusion even in the case of target disappearance and production. In addition, the final target detection locations were revised with the mean shift algorithm having comparatively low tracking location errors and allow for continuous tracking with standard model updating. The experiment used each background estimation filter with actual marine infrared still images, obtained SNR in the original and difference images, and found that the mean shift filter recorded the highest SNR and relatively good performance. In the experiment with actual marine infrared serial images, the proposed system was compared with the Kalman filter and mean shift algorithm. As a result, the proposed system recorded the lowest tracking location errors relatively and ensured stable tracking with no tracking location diffusion.

ACKNOWLEDGMENTS

ment of target reliability (C$_1$, C$_2$) was straightforward as follows: those candidate targets that were classified as "candidate targets that have been tracked (X^f)" two or more times were classified as targets. Those candidate targets that were classified as "candidate targets detected with an error (X^e)" of two or more were classified as error detection. Reliability C$_2$ was not used here.

Figure 4 shows part of the result image of introducing an experiment image into each tracking system. Figure 5 presents the mean location errors of four targets tracked with each tracking system. The first order Kalman filter had no diffusion in tracking locations and accordingly demonstrated relatively high stability, but its tracking location errors were high. The mean shift algorithm showed overall high tracking location errors as the target tracking location diffused and converged to a point in the sky starting with the 19th image. The proposed system recorded the lowest tracking location error except for the 18th image and showed relatively high stability with no tracking location diffusion.
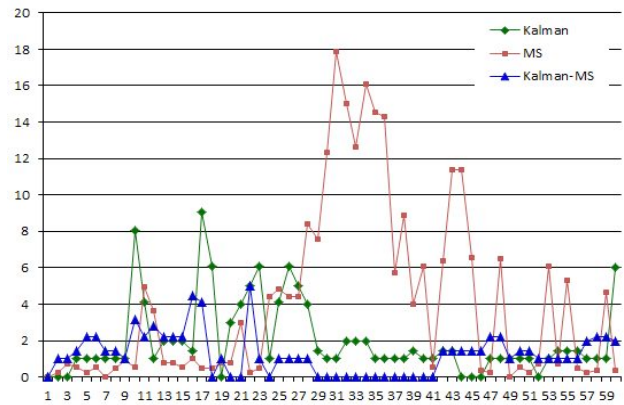
REFERENCES

[1]   B. Zhang, T . Zhang, Z. Cao, K. Zhang,: Fast new samll-target detection algorithm based on a modified partial differ ential equation in infrared clutter, Optical Engineering, 46, 10, 106401-1 -106401-6 (2007) [DOI:http://dx.doi.org/10.1117/1.2799509].

[2]   S. D. Deshpande, V. Ronda, P. Chan, :Max-Mean and Max-Median filters for detection of small targets, Proc. SPIE Conference, 3809, 74-83 (1999) [DOI: http://dx.doi.org/10.1117/12.364049].

[3]   D. Comaniciu, P. Meer, :Mean Shift: A Robust Approach Toward Feature Space Analysis, IEEE Trans. on PAMI, 24, 5, 1-18 (2002) [DOI: http://dx.doi.org/10.1109/34.1000236].

[4]   Y. Cao, R. Liu, J, Yang, : Small Target Detection Using Two-Dimensional Least Mean Square (TDLMS) Filter Based on Neighborhood Analysis, Int. J, Infrared Milli, Waves, 29, 188-200 (2008) [DOI : http://dx.doi.org/10.1007/s10762-007-9313-x].

[5]   J. Zaveri, M.A. ; Desai, U.B. ; Merchant, S.N. : PMHT based multiple point targets tracking using multiple models in infrared

192

Trans. Electr. Electron. Mater. 14(4) 187 (2013): S.-Y. Ye *et al.*

image sequence, Advanced Video and Signal Based Surveillance, Proceedings. IEEE Conference, 73-78(2003) [DOI: http://dx.doi.org/10.1109/AVSS.2003.1217904].

[6] C. Y. Li, H. B. Ji,: Marginalized Particle Filter based Track-Before-Detect Algorithm for Small Dim Infrared Target, AICI, 3, 321-325(2009) [DOI : http://dx.doi.org/10.1109/AICI.2009.232].

[7] Y. H. Liu, Q. Q. Yan, W. Liu, H. Yuan, G.Y. Zhang, :An effective target tracking algorithm in infrared images video, WICOM, 1-4(2010) [DOI : http://dx.doi.org/10.1109/WICOM.2010.5600106].

[8] O. Nichtern, S. R. Rotman, :Tracking of a Point Target in an IR Sequence using Dynamic Programming Approach, 24th Convention of IEEE, 265-269(2006) [DOI: http://dx.doi.org/10.1109/EEEI.2006.321068].

[9] L. Yang, W. Lu, J. Yang, :A New Tracking Method for Small Infrared Targets, ICIP, 3609-3612(2009) [DOI: http://dx.doi.org/10.1109/ICIP.2009.5414290].

[10] R. E. Kalman,: A New Approach to Linear Filtering and Prediction Problems, Trans. of the ASME-J.of Basic Engineering, 82(SeriesD),35-5(1960) [DOI: http://dx.doi.org/10.1115/1.3662552].

[11] R. Rosales, S. Sclaroff,: 3D trajectory recovery for tracking mul-tiple objects and trajectory guided recognition of actions. IEEE Conf. Computer Vision and Pattern Recognition, 117-123 (1999) [DOI: http://dx.doi.org/10.1109/CVPR.1999.784618].

[12] W. R. Gilks, C. Berzuini, :Following a moving target - Monte Carlo inference for dynamic Bayesian models, J. of the Royal Statistical Society B, 127-146(2001) [DOI: http://dx.doi.org/10.1111/1467-9868.00280].

[13] C. Rago, P. Willett, R. Streit, :Direct Data Fusion Using the PMHT, ACC,3,1698-1702(1995) [DOI: http://dx.doi.org/10.1109/ACC.1995. 529798].

[14] D. Comaniciu, V. Ramesh and P. Meer, :Kernel-Based Object Tracking, IEEE Trans. on PAMI.,25,5,564-577(2003) [DOI: http://dx.doi.org/10.1109/TPAMI.2003.1195991].

[15] A. P. Dempster, N. M. Laird, D. B. Rubin.,:Maximum Likelihood from In-Complete Data via the EM Algorithm,. J. of the Royal Statistical Society B, 39,1,1-38(1977) [DOI: http://dx.doi.org/10.2307/2984875].

[16] T. Liu, X. Li, :Infrared Small Targets Detection and Tracking based on Soft Morphology Top-Hat and SPRT-PMHT, Int. Congress on Image and Signal Processing, 968-972, (2010) [DOI: http://dx.doi.org/10.1109/CISP.2010.5646926].