Real-Time Target Detection against Strong Background Under of Daytime Condition

Zhenming Peng[*a,b], Qiheng Zhang[b], Xianrong Peng[b], Junping Xu[b], Peng Yan[b]
[a] College of Opto-Electronic Information,
University of Electronic Science and Technology of China (UESTC), Chengdu, China 610054;
[b]Institute of Optics and Electronics (IOE), Chinese Academy of Sciences, Chengdu, China 610209

ABSTRACT

Real-time target detection against strong (bright) background under daytime is a challenging and leading edge subject, and also is a key technique for imaging tracking system. Strong background makes CCD image sensor work in critical saturation state, and imaging target contrast is very low. It's very difficult to accurately and stably track due to the complex characteristics of imaging target, such as strong clutter background, low contrast, and low signal to noise radio (SNR). So the key techniques for detecting and tracking target are eliminating the disturbance of diffuse reflection and beacon, synchronous detection, improving the performance of real-time image processing with high frame rate and high sampling rate.

A robust strategy for detecting and tracking day-time target was proposed in this paper. A series of efficient approaches ware presented to improve performance of detection and tracking in precision and stability, including strong background and noise suppression, image enhancement, adaptive thresholding, region merging based on morphology, recognition and tracking algorithm and so on.

In the end, we summarized and built the effictive flow for detecting and tracking target against strong background under daytime. The results of combining computer simulation with practical detection experiments show that the above-mentioned approaches are feasible and significant for real-time tracking system.

Keywords: Strong background attenuation, nonlinear contrast enhancement, Adaptive thresholding, Region merging, Day-time target recognition and tracking

1. INTRODUCTION

Real-time target detection against strong (bright) background under daytime is a challenging and leading edge subject, and also is a key technique for imaging tracking system. A number of point (dim) target detection schemes have been developed over the past years [1-7]. Among these schemes are adaptive prediction filters [4], adaptive lattice algorithm [5], sequential hypothesis testing [6] and mean shift [7]. However, in practice, these methods have been only partially successful and have been shown to produce high false alarm rate. At visible-light band, the sky background luminance change range is 5~200 $W.m^{-2}$. Sr^{-1}. While the high angle of sun is 10~15° , and the angle of the line and sight (LOS) with the beam of sunlight is 90~120° , the background luminance is generally 7~15 $W.m^{-2}.Sr^{-1}$. Strong background

[*] uestc_pzm@yahoo.com.cn; phone 86 028-83208946; fax 86 028-85100966; //http: www.uestc.edu.cn

Electronic Imaging and Multimedia Technology IV, edited by Chung-Sheng Li,
Minerva M. Yeung, Proc. of SPIE Vol. 5637 (SPIE, Bellingham, WA, 2005)
0277-786X/05/$15 · doi: 10.1117/12.572424

makes CCD image sensor works in critical saturation state, and imaging target contrast is very low [8]. So there are many difficulties faced to be overcome while detecting and tracking target under the condition of daytime, such as eliminating the disturbances resulting from diffuse reflection and beacon, synchronous detection, improving the performance of real-time image processing with high frame rate and high sampling rate[9-10].

An initial image acquired from real scene at some experiment is shown in figure 1. Interesting target hidden in image is hardly observed by eye. Imaging target under day-time background has complex characteristics as follows:

(1) The S/N radio is very low as various disturbances existing in image, including platform disturbance, power disturbances, electromagnetic disturbances, system noises and various miscellaneous reflexes. It results in many difficulties for image processing. Figure 2 shows an image segmented with arbitrary threshold, and segmenting result shows shape with bright-in-center and dark around, like a cooker bottom. Equidistant stripes in image are periodic power disturbances. In addition, various random noises, clutter conglomerations caused by miscellaneous reflexes and immanent spots of optical system also appeared in image.

(2) Target in imaging plane shows complex characteristics such as misshapen and blurry edge feature, intensity change with obvious. Such imaging characteristics without shape, size and textures make it very difficult for measuring off-target value, i.e. the error between the target position and the field of view (FOV), and tracking target subsequently. Figure 3 is a result image segmented with an optimal threshold, which contour is clear but the disturbances in image couldn't be removed completely.

In this paper, we mainly describe the new methods and technique for improving performance of daytime detection and tracking precision and stability. Various effective approaches for daytime target detection were presented in this paper, including strong background attenuation, image enhancement, adaptive thresholding, region merging algorithm based on morphology, recognition and tracking algorithm and so on.

In the end, we summarized and built the effective flow for detecting and tracking target against strong background under daytime. The results of combining computer simulation with practical detection experiments show that the above-mentioned approaches are feasible and significant for real-time tracking system.

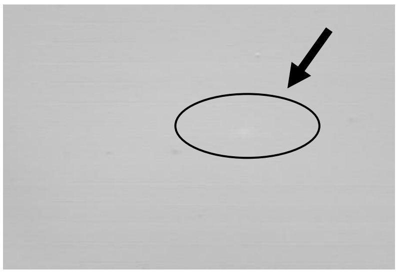


Fig.1 Initial image with strong background

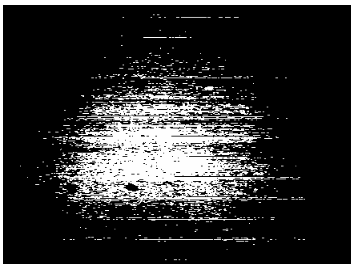


Fig.2 Various disturbances (image segmented with arbitrary threshold)

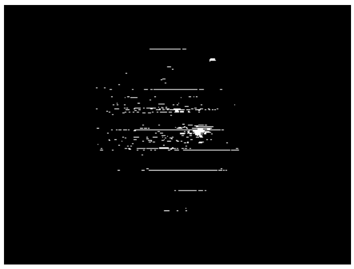


Fig.3 Segmented result with certain threshold

2. STRONG BACKGROUND ATTENUATION

2.1 Clutter elimination based on spectrum-filter

Spectrum filter is a technique which can filtrate light with expectative wave length range from continuous spectrum so that improve the modulation contrast between target and background. So we should design and select corresponding light filter according to target's spectral radiant characteristics. Light filter must satisfy certain requirement that energy loss of expected wave band should be as little as possible, and light of that unexpected wave-band should be reflected and absorbed completely. The spectrum transmissivity curve of light filter should conform to the design standard, and appear in the sensitive zone of detector. To adapt various target and background with changing luminance, we can use two modes to adjust light in optical system: piece-wise and continuous. Piece-wise adjustment uses the combination of light-balancing filter and narrow-band pass filter, so that change brightness for target and background in detector imaging plane, and eliminate the disturbance resulting from beacon and miscellaneous-light, and improve S/N radio of imaging target.

2.2 Improving imaging quality using eyelet-diaphragm

In order to remove long range miscellaneous-light outside of optical axis (reflection for non-imaging object, refraction interface and the insides of instrument crust) into FOV, eyelet diaphragm is a feasible and effective approach, which can improve imaging quality to a certain extent. In practical application, we should design the proper size for diaphragm according to FOV, target size, pupil entry, diaphragm and pupil exit based on the conjugate relationship between object and image. By calculating and experimentally analyzing, eyelet diaphragm with a diameter of 3~4mm is generally effective for daytime detection.

In real-time system, it possibly generate a so-called “cooker-bottom effect” in FOV while adopting eyelet, that is, bright in middle and dark around (see figure 2). Moreover, the smaller the eyelet is, the more obvious the effect is. Therefore, we should take account of a compromise between eliminating clutter and “cooker-bottom effect”.

3. IMAGE ENHANCEMENT

For low contrast image, enhancement must be done at preprocessing stage. Experimental analysis shows that the scale of linear gray-level transform is uniform and disadvantageous for boosting the relative difference between target and background. However, direct contrast enhancement (DCE) based on nonlinear transform（e.g. quadratic curve, logarithmic function）can solve the problem well and increase image contrast with pertinency.

First, define a contrast for each pixel within processing window（N×N window as an example）

$$C(i,j) = \frac{\left|G(i,j) - \overline{G}\right|}{\overline{G}} \qquad (1)$$

where, $G(i,j)$ is the initial intensity of pixel (i,j) before transforming within processing window, $\overline{G}$ is mean intensity summing up all pixels within processing windows, and can be written as

$$\overline{G} = \frac{1}{N^2}\sum_{j=0}^{N-1}\sum_{i=0}^{N-1} G(i,j) \qquad (2)$$

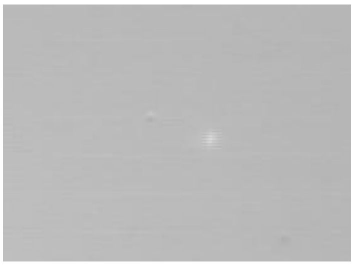
Then, define a nonlinear transform function as follows

$$C_T = \varphi(C) = 3C - 3C^2 + C^3 \qquad (3)$$

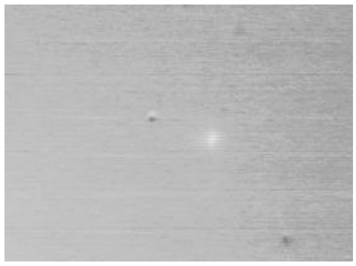
where, φ is convex transform curve, and satisfy $\varphi(0)=0, \varphi(1)=1, \varphi(C) \geq C$. This transform only enhances contrast greatly for each pixel with lower contrast, but slightly for each pixel with higher contrast. Each pixel is transformed in light of equeation（3）, then intensity of pixel (i,j) transformed can be formulated as

$$G_T(i,j)=\begin{cases} \dfrac{\overline{G}(1-C_T(i,j))}{1+C_T(i,j)} & G(i,j)\leq \overline{G} \\ R-\dfrac{(R-\overline{G})(1-C_T(i,j)}{1+C_T(i,j)} & G(i,j)>\overline{G} \end{cases} \qquad (4)$$

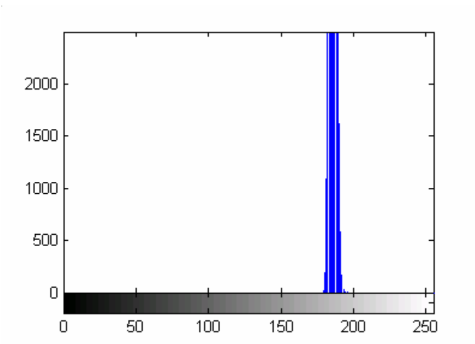
where, R is maximum intensity limited. Figure 4(a) and (b) show the comparison of an image before and after DCE. We can see, image (b) has higher contrast than image (a). Figure 4(c) and (d) show their histograms, and it is obvious for figure (d) to be extended comparing with figure (c).



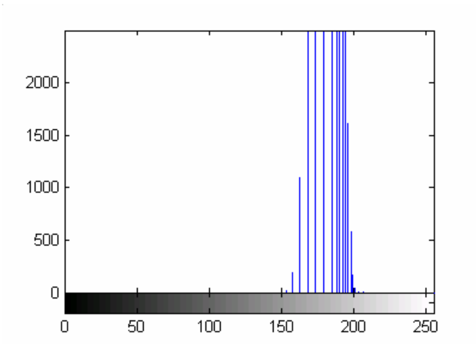
（a）



(b)



（c）



（d）

Fig.4 Nonlinear contrast extension and image enhancement. (a) Initial image; (b) Image after DCE; (c) Initial histogram for image (a); (d) Histogram for image (b).

In practical tracking, because target generally located the brightest area within tracking-gate, we can search that of bright area and give enhancement factor by a jump function, so that adjust the changing range for image data while processing with hardware. This approach also belongs to one of the nonlinear DCE for image enhancement.

3. ADAPTIVE THRESHOLDING

The method for threshold selection should combine mean with variance of image data within tracking gate, can be written as [12]

$$T = \mu + \lambda\sigma \qquad (5)$$

where μ is the mean value of intensity within processing windows, σ is variance about mean value, and λ is an adaptive coefficient changing with target and background feature. The criterion for selecting λ is that segmenting results for target point numbers are moderate. If λ were too large, segmenting results could contain more high-frequency noises, whereas may be lose the target. Tests show its value changing from 7 to 10 is generally suitable for target segmentation. In real-time detection, λ could be adjusted timely according to different background characteristics, so that determine the optimal thresholding.

In implementation, considering the bright target, equation (5) can be simplified as follows

$$T = \mu + \frac{|G_{\max} - \mu|}{2} \qquad (6)$$

where $G_{\max}$ represents the maximum gray in processed region.

Similarly, threshold for black target can be written as

$$T = \mu - \frac{|G_{\max} - \mu|}{2} \qquad (7)$$

Then, calculate threshold according to equation (6) or（7）, and implement binary image segmentation.

Target pixels segmented by the above-mentioned approach probably contain high-frequency noises. The numbers of noise points depend on threshold given. These high-frequency noise points are very similar to target points in shape at single frame image, show a point or a spot, even the gray difference between both is very small. However, target and noise have their respective characteristics. Target points show certain stability and continuity and noises show randomicity. Therefore, from pixels distribution, target and noise should locate at their respective neighborhood region. Moreover, target has fixed size and noise points are uncertain in image sequences. Selecting a proper pixel numbers with the same type as a threshold can remove noise well in future.

4. RECOGNITION AND TRACKING

4.1 Point target recognition

A simplified clustering method based on neighborhood-region analysis is presented. For one pixel segmented as probable target, if three pixels or more around it within 3×3 region are segmented as target points, this pixel must be target point, otherwise, is noise. By analogy with this, we can cluster all pixels of image within gate, and succeed in extracting target points from background clutter. This method probably increases computation time consumption, but we can bisect the gate for locating roughly target area to satisfy the requirement for real time and avoid repeating computations.

In fact, target pixel recognition is the process of counting "white" pixels around each pixel neighborhood region in binary image. As figure 5, suppose the white point in image is the probable target point (e.g. pixel（i, j）), thus the sum of target points around its 3×3 neighborhood region can be formulated as

$$D_{(i,j)} = \sum_{x=-1}^{1}\sum_{y=-1}^{1} G(i+x, j+y) \qquad (8)$$

where $G(i, j)$ is the gray value of pixel(i,j). If D (i,j)＞3, pixel (i,j)is assumed as a target point segmented, otherwise, is regarded as noise or background to be removed.

The criterion of the above-mentioned recognition is that spot with 2×2 or more size is regarded as target point. It probably results in wrong segmentation, i.e. some small clutter blocks was regarded wrong as target points, example for inherent flecks and blot in CCD target area. If these clutter blocks contribute to poid computation for target, they could influence tracking precision. Therefore, it is necessary to remove these small clutters farther.

With a rectangle (e.g. 5×5 window) covering all over the searching image (as figure 6), find the region with the white-pixel, which is regarded as the initial location for target. This approach has two advantages: avoid interlarding segmentation result with clutter block being regarded wrong as target in processing window; provide rough tracking area for computing poid of target in future.

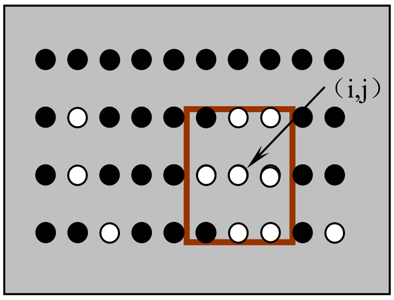


Fig.5 Clustering for binary image in neighborhood

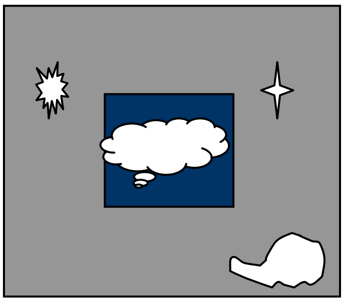


Fig.6 Roughly locating target

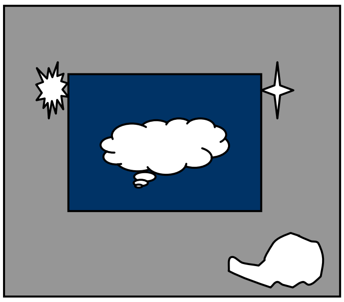


Fig.7 Accurately calculating target's poid

4.2 Poid tracking

The poid of object is also called centroid. For point target segmented in binary image, it is significant for poid tracking to locate target with high precision. Enlarge the area appropriately from initial processing window (area changing from fig.6 to fig.7), then calculate the coordinate of poid (x,y), i.e. the final result for target tracking. To avoid memory unit more than 16-bit and reduce time consumption in real-time implementation, poid computation can be divided into two processing phases.

First, we project statistically target on two directions of X and Y,

$$E(x) = \sum_{y=0}^{N-1} G(x,y)$$
$$E(y) = \sum_{x=0}^{M-1} G(x,y) \qquad (9)$$

where G(x,y) is the gray value of binary image segmented, and M×N is the size of processing window.
To obtain accuracy for poid, average E (x) and E (y), thus the total pixels of target image as follows

$$N' = \frac{E(x) + E(y)}{2} \qquad (10)$$

Secondly, calculate the coordinates (x_0,y_0) of poid according to equation (11). Note that the origin of coordinate locates at upper-left corner of FOV and convert them if necessary.

$$x_0 = \frac{1}{N'} \sum_{x=0}^{M-1} xE(x)$$
$$y_0 = \frac{1}{N'} \sum_{y=0}^{N-1} yE(y) \qquad (11)$$

4.3 Region merging

Morphology is a mathematical method that analyzes geometry and structure, which includes erosion, dilation, opening, closing and hit-or-miss transform [13].

Setting X and B are the subset of space E^N, and symbol "$\oplus$" and "Θ" represent dilation and erosion, respectively, thus B dilates X can be formulated as

$$X \oplus B = \{z \in E^N \mid z = x + b, x \in X \; and \; b \in B\} \qquad (12)$$

Dilation remains primary shape of original, but fills the concave section of unsmoothed boundary.
Similarly, B erodes X can be formulated as

$$X \Theta B = \{z \in E^N \mid z = x + b \in X, forevery \quad b \in B\} \qquad (13)$$

Erosion also remains primary shape of original, but remove the convex section of unsmoothed boundary.
Opening can be written as follows

$$Open(X, B) = (X \Theta B) \oplus B \qquad (14)$$

That is, erosion first and dilation later for one image using the same structure elements. Opening removes the isolated sub-regions and burrs.
Closing can be written as follows

$$Close(X, B) = (X \oplus B) \Theta B \qquad (15)$$

Closing is the contrary process of opening, which can fill some small cavities and merge that of close target points.

For point target, its poid has the characteristic of rotation invariability, and tracking point is steady. However, the poid coordinate calculated is probably uncertain due to imaging dispersion with target moving. In order to compensate for dispersion and deformity for imaging target, we can use opening operation to merge region. First, erosion was done for image, which can remove the isolated noise and stripe disturbances; then dilation was done to grow and merge imaging region with a local peak value starting from target point determined after being segmented roughly.

Opening operation can remove the isolated noise or noises far from target, and merge the background pixels into object pixels. If the distance between both objects is very short, two blocks segmented may be connected together. Therefore, it can compensate for deformity due to imaging dispersion, and make the segmentation result approach to real object and poid calculation is more accurate and improves detection precision, tracking stability, adaptability to various scenes and intelligentized degree for target with flickering and changing luminance. Figure 8 show the segmenting result of initial image with optimal threshold, which include lots of noises and stripe disturbances. Figure 9 is the opening operation result for figure 8 with 3×3 structure elements. It is easy to see that isolated noise and stripe disturbances are almost removed and target contour extracted is close to real shape.

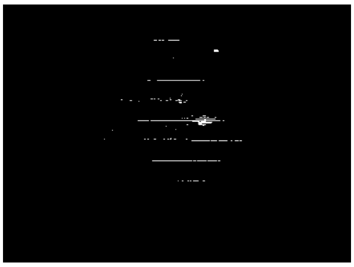
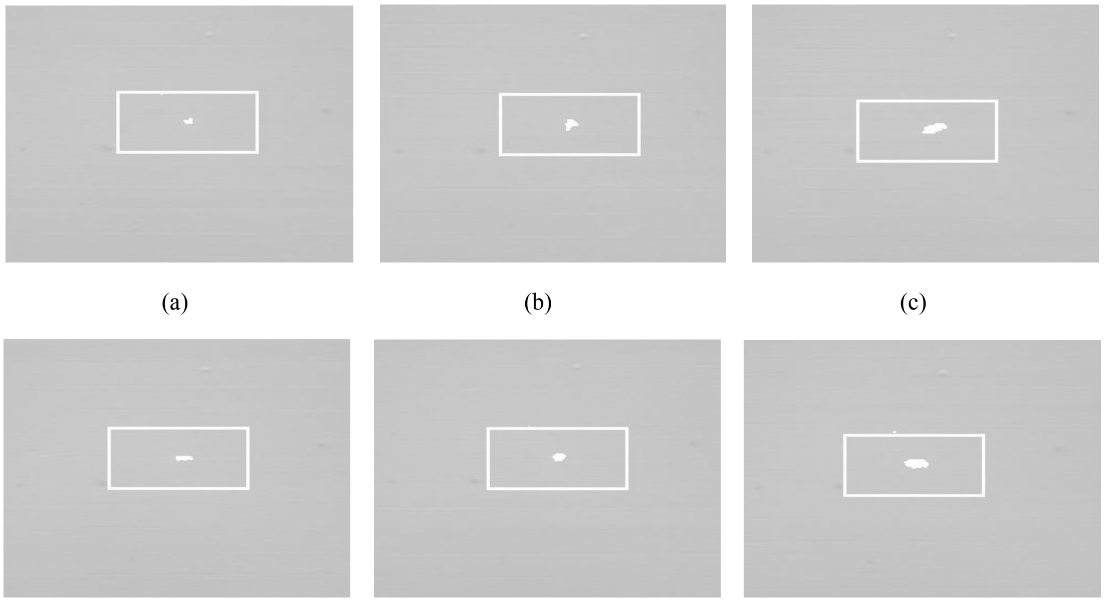


Fig.8 The result segmented with optimal threshold

Fig.9 The result of opening operation for figure 8

5. EXPERIMENTAL RESULTS

Based on the above approach and ideal depicted, we had implemented real-time tracking for image sequences acquired from real scene. The initial images size are $N\times N = 768\times576$, 200 frames in all. Figure 10 shows the tracking result(the size only for displaying images are 320×274), where (a), (b), (c), (d), (e) and (f) are frame #17, #32, #44, #83, #134 and #192 of consecutive scenes, respectively. We can see from experimental results that tracking is steady and have not lost the target. Although the target shape segmented had been changed timely during moving, they kept the real shape of target (this is because that optical system cause imaging dispersion), and could satisfy the precision requirement for real-time tracking.



(a) (b) (c)

(d) (e) (f)

Fig.10 An example set of tracking target against strong background. (a) ~ (f) shows the result for frame #17, #32, #44, #83, #134 and #192, respectively.

6. CONCLUSIONS

This paper focused studying on the key technique for real-time dim target detection against strong background under daytime, and created the effective tracking flow. The practical application and lots of experimental results for diverse scenes show that the above approach is feasible, especially in flickering and dim target detection with far distance. Suppressing miscellaneous-light, enhancing S/N radio and contrast, effective image segmentation and high-precision tracking algorithms are very significant for improving performance of detection system.

By studying and analyzing on experimental results, currently there are still some problems to be resolved in future works.

(1) System adaptability to diverse scene and the performance of ant-disturbance need to be improved due to various influences such as CCD quanta noises, current noises, and power disturbances resulted from platform vibration, and electromagnetic disturbances.

(2) System strictly depends upon weather and sky background, and the adaptability to weather is relatively weak. The adaptive strategy for background suppression, compensation and estimation need to be explored farther.

(3) Study on the adaptive correction methods for imaging dispersion due to optical system.

REFERENCES

1. C.E. Caefer, et.al, "Point target detection in consecutive framer IR imagery with evolving cloud clutter," *Proceeding of SPIE*, **2561**, 14-24, 1995

2. J.M. Mooney, et.al, "Point target detection in consecutive frame staring infrared with evolving cloud clutter," Optical Engineering, **34**, 2772-2784,1995
3. D. Davies, P. Palmer, Mirmehdi, "Detection and Tracking of Very Small Low Contrast Objects," *Proceedings of the 9th British Machine Vision Conference*, Sept. 1998.
4. A. Sarkan, et.al, "Performance evaluation of 2D adaptive prediction filters for detection of small objects in image data," *IEEE Trans. On IP*, 2, 327-340, 1997
5. P.A. Ffrench, J. R. Zeidler, and W. H. Ku. "Enhanced detectability of small objects in correlated clutter using an improved 2-d adaptive lattice algorithm," *IEEE Trans. on IP*, **6**(3):383–397, 1997.
6. S. Blostein and T. Huang. "Detecting small, moving objects in image sequences using sequential hypothesis testing," *IEEE Trans. on Signal Processing*, **39**(7):1611–1629, 1991.
7. D. Comaniciu, V. Ramesh, P. Meer, "Real-time tracking of non-rigid objects using mean shift," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, **2**, 142-149, 2000.
8. W.B. Boucher, M.R. Michnovicz, "Dynamic range selection in image processing hardware to maximize SNR while avoiding image saturation," *Proceedings of SPIE*, **3692**, 1999
9. Z. Peng, Q. Zhang, J. Wang., "Self-adaptive strategy searching for the faint target," *Acquisition, Tracking, and Pointing XVII, Proceedings of SPIE* , **5082**, 114-122,2003
10. J.N. Sanders-reed, W.M. Dimmler., "Scene generation for tracking small targets against cloud backgrounds," *Proceedings of SPIE*, 4376, 2001
11. H. Nariman, M. Alireza, B. Neil, "Automatic thresholding for change detection in digital video," *Proceeding of SPIE*, **4067**, 133-142, 2000
12. J.B. Kim, H.J. Kim, "Efficient region-based motion segmentation for a video monitoring system," *Pattern Recognition Letters*, 24, 113-128, 2003
13. U. Braga-Neto, J. Goutsias, "Automatic Target Detection and Tracking in Forward-Looking Infrared Image Sequences Using Morphological connected Operators," *33rd Conference of Information Sciences and Systems*, March 1999.