NLEBS: automatic target detection using a unique nonlinear-enhancement-based system in IR images

Shlomo Greenberg
Raanan Yehezkel
Yaniv Gurevich
Hugo Guterman
Ben-Gurion University of the Negev
Department of Electrical and Computer
Engineering
P.O. Box 653
Beer-Sheva 84105, Israel
E-mail: hugo@ee.bgu.ac.il

Abstract. A new automatic target detection method for IR images that only requires information about the size of the targets is described. The proposed nonlinear-enhancement-based system (NLEBS) algorithm is based on a nonlinear enhancement paradigm that increases the contrast of the targets with minimal change in the clutter's contrast. The NLEBS employs several stages of processing, each with a different operational purpose. First, the nonlinear enhancement operation is performed by using an iterative procedure. After binarization, segmentation merging causes each local image region to grow by filling in holes. Then segmentation pruning is applied to remove spurious segments. Finally, a heuristic-based metric is employed to validate the possible targets. The performance of the NLEBS was tested with a large set of IR images. The results of these experiments showed a probability of detection greater than 90% and a false-alarm rate of about 1 false alarm per image.

© 2000 Society of Photo-Optical Instrumentation Engineers. [S0091-3286(00)03705-3]

Subject terms: automatic target detection; enhancement algorithms; image processing; pattern recognition.

Paper 990248 received June 21, 1999; revised manuscript received Sep. 21, 1999; accepted for publication Nov. 15, 1999.

1 Introduction

Automatic target detection (ATD) is an important task for military purposes. Its goal is to ascertain the presence of specific objects in an image automatically, without human intervention. An efficient ATD system should achieve high detection percentage for targets while maintaining a minimal false-alarm rate. This means that it must preserve an optimal balance between a high detection rate and a low error probability.

In real-world infrared (IR) images, targets of interest and the background against which they are observed are frequently obscured by clutter. In IR images, the gray level of a pixel represents the scene temperature. A high gray level (brighter) represents hotter parts in the scenery. At night, elements for which the temperature is not artificially changed have approximately the same temperature, and therefore they appear in the IR image with a similar gray level. In daytime some elements get heated by sunlight while other elements have a lower sensitivity. Targets (such as tanks, trucks, etc.) can be hot (due to a hot engine) or cold (engines shut down for a long period of time). Usually, the difference between the average temperature of a cold target and the background is minimal (low contrast), while the difference between hot targets and their local background is high (high contrast). Elements in the scenes that are not targets, but that have a high contrast and characteristics similar to the targets, are referred to as clutter. Clutter that is misinterpreted as targets is a principal source of detection error.1

An ATD system is expected to detect equally targets in scenarios with a low contrast (cold targets) and with a high contrast between natural elements and the background. However, those two situations are the main causes of false detection, and several methods have been developed to overcome these types of problems.^{3,4} In the early 1980s ATD algorithms were generally heuristic and detection was based on some sort of threshold. The threshold was determined by the contrast of an object with its local background.³ The segmentation part of such algorithms was attained by employing a standard edge operator followed by closing shapes and filling steps.^{5–8} The detection performance of those algorithms never exceeded 70%.³ Detection algorithms from the early 1990s were based on the convolution of some sort of filter with the image. 3,9,10 These algorithms were based on prior knowledge derived from the target's template. More recently, Casasent and Neiberg analyzed the performance of three wavelet-based approaches—the morphological wavelet transform (MWT), the Gabor transform (GT), and the Gabor wavelet function (GBF)—as well as their fusion.⁴

In this paper, we introduce an ATD algorithm based on adaptive nonlinear enhancement of the image. The proposed nonlinear-enhancement-based system (NLEBS) algorithm is based on a nonlinear enhancement paradigm that increases the contrast of the targets with minimal change to the clutter's contrast. This is attained by applying a nonlinear operator to the histogram. The objective is to change the histogram of a suspected area in such a way that the histogram's valley separates the target from its background and clutter (see Fig. 1). Then, every pixel that has a graylevel value above a chosen threshold is assigned as a possible target, and the others as background (or clutter). The

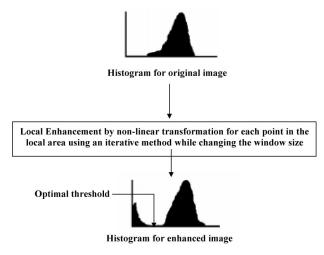


Fig. 1 Image histogram enhancement.

nonlinear enhancement operation is carried out by an iterative procedure. The NLEBS employs a segmentation method that exploits the results of the nonlinear enhancement operation and strengthens the target signature even further

The paper is organized as follows: A general description of the proposed ATD system is provided in Sec. 2, and Sec. 3 presents the enhancement procedure and its iterative. Sections 4 and 5 describe the segmentation merge and pruning stages, respectively. Section 6 deals with metrics for the detection decision. Experimental results are presented in Sec. 7, and Sec. 8 provides conclusions.

2 General Description

The proposed NLEBS system involves several different stages of processing, each stage performing a different operation. This intricate procedure reflects the complexity of the ATD problem. A general scheme of the NLEBS is shown in Fig. 2. Nonlinear enhancement is introduced to overcome the low contrast and nonuniform conditions of typical IR images. After binarization, segmentation merging causes each local image region to grow by filling in holes. This operation increases the target signature by merging bright segments that are close to each other. Segmentation pruning removes small bright clusters on a dark background and takes away bright regions that do not correspond to *a priori* knowledge of the target. Finally, a confidence measurement of the detection is obtained by applying a heuristic-based target metric.

3 A Nonlinear Enhancement Algorithm

A well-known experimental result in vision physiology and psychology is that for different background brightnesses I, a given absolute brightness difference ΔI does not excite the same variation in the human visual system, whose discriminating ability decreases with an increase in background brightness. The minimum detectable visual increment (also known as the visual incremental threshold, or the luminance difference threshold, or the brightness just-noticeable difference) is defined as the ΔI necessary to allow discrimination under a reference background field I. The human visual system presents three-well defined zones.

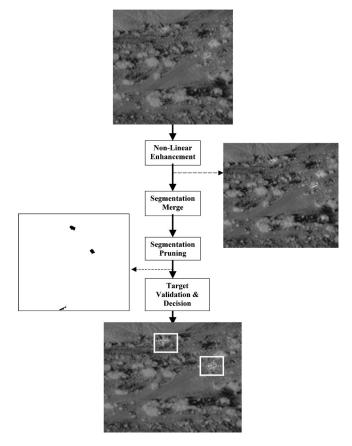


Fig. 2 General schema of the NLEBS system.

At low light intensities, near the absolute visual threshold, ΔI is constant.¹² Then, with increasing intensity, ΔI converges asymptotically to Weber behavior,

$$\Delta I = \alpha_1 I. \tag{1}$$

The intermediate zone with slope of 1/2 is known as the DeVries-Rose region:

$$\Delta I = \alpha_2 \sqrt{I},\tag{2}$$

where α_1 and α_2 are constants. If the brightness of an object is higher (or lower) than its background I by at least ΔI , the object is detectable by the human visual system.

Another capability of the human visual perception relevant to the problem at hand is its inherent spatial adaptation. The capability to perform ocular movements in order to scan an entire image and to focus on subareas that might contain objects of interest ensures global and local adaptation. For the objective of developing an enhancement method that can, as far as possible, exploit the above-described principles, we devised the following strategy for enhancement of IR images.

First, to facilitate the process, we make the following common sense assumptions:

- 1. In IR images, a target usually has brighter parts than those in its environment.
- 2. The targets are small compared to the entire scene.

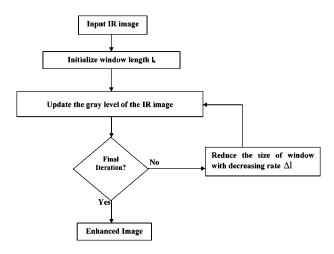


Fig. 3 Flow chart of the enhancement algorithm.

- In high-contrast images, the target usually has some brighter parts, but only relative to its local environment
- 4. The algorithm focuses on enhancing those parts in the target that appear to be brighter in the IR image.

Based on these assumptions, we developed an algorithm that divides the image into a set of subregions and then selectively enhances the image. A general flow chart of the enhancement algorithm is shown in Fig. 3, and a detailed description of the algorithm follows.

3.1 Hierarchical Image Partition

One of the most difficult problems in discriminating between a target and its background is the determination of the optimum window size. As shown in Fig. 1, the histogram of the original IR image may be too dense to allow its separation. However, by mimicking the human vision capability of spatial adaptation, we were able to observe the different shape of each local histogram. Basically, spatial adaptation can be attained by zooming into each local area with different window sizes. In the present case, the window size transformation is given by

$$l_k = l_0 - n \ \Delta l, \tag{3}$$

where l_0 is the initial length of a square window, Δl is the decreasing factor, and n is the iteration index. In the present study the following values were used: $l_0 = 100$, $\Delta l = 5$; and a total of 10 iterations were performed. As shown in Fig. 4, the window size decreases as the iteration proceeds. The nonoverlapping windows for the initial and final stages are shown in Fig. 5. The zooming operation fuses the broad and narrow neighborhood information efficiently.

3.2 Enhancement Operator

For each window $q(W_q)$, the average value (m_q) of the IR local image is calculated. Then, the following equation is employed to update each pixel $(p_{i,j})$ in the selected window:

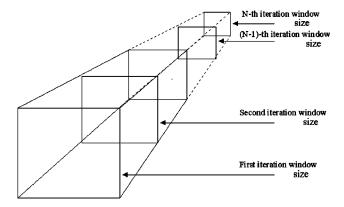


Fig. 4 Window size for each iteration.

$$\begin{aligned} p_{i,j} &= p_{i,j} + f(\Delta p_{i,j}), \\ \Delta p_{i,j} &= p_{i,j} - m_q, \end{aligned} \tag{4}$$

where $f(\Delta p_{i,j})$ is the enhancement operator.

The enhancement schema presented here, which includes image partition and the enhancement operator, is a highly complex feedback system. The positive feedback can amplify small signals $\Delta p_{i,j}$ to very high levels unless the function $f(\Delta p_{i,j})$ is nonlinear. To strengthen this point, let us apply a linear enhancement operator that has been widely used $f(\Delta p_{i,j})$ is nonlinear.

$$f_L(\Delta p_{i,j}) = d \cdot (p_{i,j} - m_a), \quad 0 < d < 1.$$
 (5)

It can be seen that when applied to IR images, this linear operator enhances the targets as well as the clutter (Fig. 6).

Nonlinearity *per se* does not eliminate background amplification, since many nonlinear operators can produce a pathological response in which all signals are amplified to equal asymptotes no matter how different and small their initial values are. ^{15,16} To avoid such situations, the operator $f(\Delta p_{i,j})$ needs to be a faster than linear function, such as

$$f_{NL}(\Delta p_{i,j}^k) = d \cdot (p_{i,j}^k - m_k)^n, \quad n > 1, \quad 0 < d < 1.$$
 (6)

The behavior of both operators (linear and nonlinear) can be seen in more detail in Fig. 7. It can be observed that the nonlinear operator does not enhance those pixels that

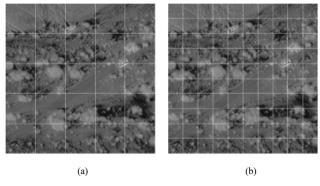


Fig. 5 Image partition: (a) Initial, (b) Final.

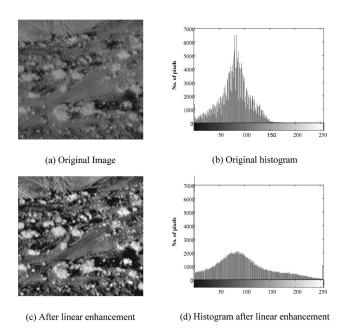


Fig. 6 Results of the enhancement using a linear operator.

are near the mean illumination level of the selected window. Only pixels that are located far away from the mean are significantly amplified.

By applying the nonlinear operator to highly cluttered IR images, we were able to detect locally defined targets despite the clutter. An example of a nonlinearly enhanced image and its corresponding histogram are shown in Fig. 8.

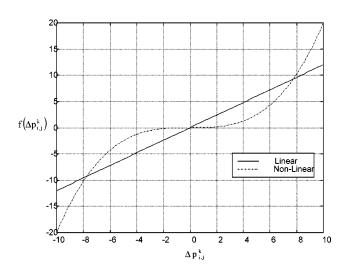


Fig. 7 General representation of the linear and nonlinear operators (n=3, d=0.2). Units are arbitrary.

It can be seen that while the targets are emphasized the clutter remains almost unchanged (Fig. 9). The histogram plots shown below support this claim.

4 Merge Segmentation

The segmentation stage increases the target signature relative to the clutter (in the binary image) by merging all the small segments belonging to the same target. The merging is done in such a way that the resultant shape of the segment is similar to the original object's shape, while having

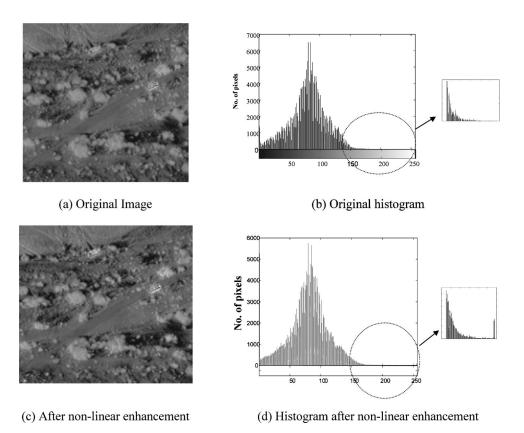


Fig. 8 Results of the enhancement using a nonlinear operator.

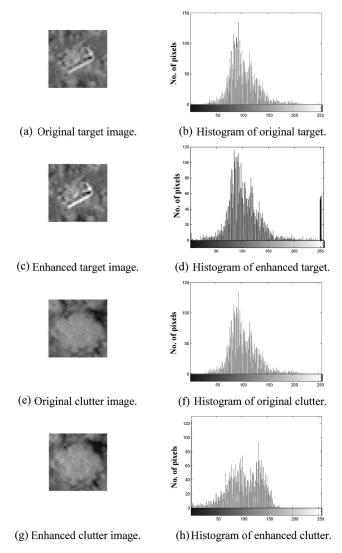


Fig. 9 Target and clutter signature after nonlinear enhancement.

a minimal effect on the clutter's shape. ^{17,18} Since a target has hot parts near relatively cold ones, it usually appears as a group of small, closely spaced segments [Fig. 10(c-1)]. Clutter usually has a more homogeneous temperature and

therefore appears as a group of larger, more compact segments [Fig. 10(c-2)].

The merge segmentation consists of four stages. In the first stage binarization of the enhanced image takes place [Fig. 10(c)]. By changing the threshold value, the similarity of the merged segment's shape to the original shape can be adjusted. ¹⁹ Selection of too low a threshold value can cause the background clutter to merge with the target segment. On the other hand, selection of too high a threshold value may not allow the formation of enough subsegments to complete a target segment. The chosen threshold must be as high as possible to ensure minimization of the clutter segments.

After binarization, three morphological operations are carried out. 20,21 First, a thickening transformation is applied [Fig. 10(d)] by which two segments having a gap between them smaller than 2m pixels are merged (m represents the size of the structuring element of the morphological operation). In order to maintain the original shape and size of the nonmerging segments, a thinning transformation was applied [Fig. 10(e)]. Finally, all the holes in the structure are filled [Fig. 10(f)].

5 Segmentation Pruning

The objective of this stage is to minimize the number of false alarms produced in the merge segmentation stage. To accomplish this, two steps of pruning are applied to the different segments. First, *a priori* geometric knowledge of the possible targets is applied. During this stage we remove those segments whose area (*A*), perimeter (*P*) or ratio *A/P* is greater than a predetermined threshold.²¹ The effect of this primary pruning on long segments can be seen in Fig. 11.

The second part of the pruning deals with the removal of small segments. Exhaustive analysis of the binary images revealed that the determination of an adaptive threshold (S_{Th}) for the minimum size of the segments strongly correlates with the total number of active pixels (P_T) in the binary image. The following empirical equation was found to provide good results:

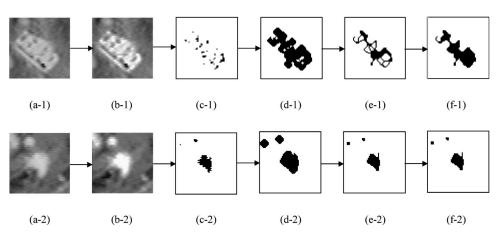
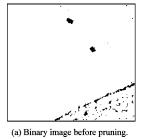


Fig. 10 Stages of merge segmentation on a target (1) and a clutter (2): (a) original image, (b) enhanced image, (c) binary image after thresholding, (d) after thickening, (e) after thinning, (f) after filling segment holes (m=3).



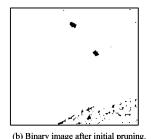


Fig. 11 Initial pruning.

$$S_{\text{Th}} = \frac{P_T}{a + bP_t}, \quad a > 1, \quad b < 1.$$
 (7)

Figure 12 demonstrates the effectiveness of the threshold on a binary image with two real targets and several segments created from enhanced clutter.

6 Target Validation

In order to assess the goodness of the detection it was necessary to apply a target-clutter discrimination metric. Several metrics have been proposed for measuring perceived target distinctness and the effect of clutter on detection systems. We found (after an exhaustive search) that a linear combination of the normalized values of the segment's intensity standard deviation (σ_j) , maximum intensity variation (ΔI_j) , maximum intensity variation (ΔI_j) , maximum intensity (ΔI_j) provides a good validation metric for the present detection system (J represents the segment index):

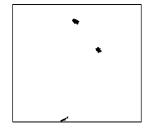
$$TV_{j} = a_{1} \frac{\sigma_{j}}{\max_{J} \sigma_{j}} + a_{2} \frac{\Delta I_{j}}{\max_{J} \Delta I_{j}} + a_{3} \frac{\max_{J} I_{j}}{\max_{J} I_{j}} + a_{4} \frac{S_{j}}{\max_{J} S_{j}},$$

$$\sum_{i=1}^{4} a_{i} = 1.$$
(8)

While the original image is employed to estimate σ_j and ΔI_j for each segment, the enhanced image is used to calculate $\max I_j$.

The selection of these features was based mainly on *a priori* knowledge of the behavior of clutter and targets in IR images. A target presents an unnatural pattern because, unlike clutter (which is natural), targets are artificial. Therefore, the following observations are generally true:

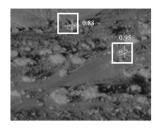


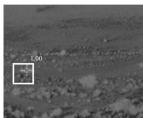


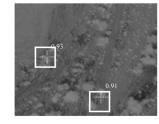
(a) Image after initial pruning

(b) After removal of small segments

Fig. 12 Final pruning.







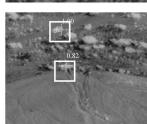


Fig. 13 Typical IR images. The crosses inside the boxes indicate the detection of a possible target. The numbers indicate the TV score [Eq. (8)].

- 1. Target segments have larger standard deviation than clutter segments.
- 2. Some parts of a target are artificially heated (e.g., the engine) while other parts remain close to air temperatures. Thus the maximum intensity variation and intensity are higher in a target than in the clutter.
- 3. Due to the segmentation and enhancement, target segments are larger than clutter segments.

7 Results

The image database used to evaluate the NLEBS consists of about 600 IR images of several areas of Israel at different hours. The field of view was $5.60 \times 5.60 \text{ deg}^2$. The size of the original images was 512×512 pixels. The database consisted of real targets (tanks, armored personnel carriers, and trucks) and was divided into three subsets that represent three levels of target contrast (high, medium, and low). About 30 images selected by human experts were provided for training and tuning, while the remainder images were employed for the test. Four typical IR images, the detected targets, and the calculated TV scores [Eq. (8)] are shown in Fig. 13. Detection performance was graded in terms of:

- *Probability of detection* (PD), defined as the probability of correct discrimination of a target in the image from background and noise
- False-alarm rate (FAR), defined as the probability of an error in detection.

The performance evaluation of the proposed algorithm is summarized in Table 1. The PD was around 90 to 92%, while the FAR was 0.5 to 1.6 FA/image, depending on the image's contrast. The optimum decision threshold, which maximizes the PD while minimizing the FAR can be found by plotting the probability of detection versus the normalized FAR (Fig. 14) and searching for the minimum gradient. The optimum threshold value was found to be $TV_{Th} \approx 0.402$. This value ensures a PD greater than 90%, with a FAR of about 1 FA/scene.

Table 1 Performance assessment of the NLEBs.

Set contrast	PD (%)	FAR (FA/image)
Low	92.12	0.56
Medium	90.6	1.64
High	90.28	0.39
Combination of all sets	90.92	1.00

Conclusions

This paper has presented a new algorithm for automatic target detection. The development of an ATD system is a complex problem, which requires the optimum integration of several image-processing stages. At the initial stages of this research very simple and well-known ideas such as thresholding methods and correlators were tested. However, it became evident that conventional approaches are not satisfactory for solving this type of problem. As human observers generally outperform ATD systems (especially in the FAR category) it was only natural to turn to the human visual system for inspiration. Consequently we derived the NLEBS algorithm, which employs a nonlinear enhancement paradigm that increases the contrast of the targets with minimal change in the clutter's contrast. This adaptive nonlinear method facilitates target segmentation from the background while target validation metrics estimate the "goodness" of the detection.

In order to compare the performance of the NLEBs with that of existing algorithms, a systematic review of the available literature was made. Only two articles describing several ATR systems appeared to be relevant for the comparison.^{3,4} In the first survey, which was performed at the U.S. Army Communications-Electronics Command Night Vision and Electronic Sensors Directorate (NVESD) at Ft. Belvoir, Virginia, eight systems were evaluated and compared with human observers.³ The algorithms cover a wide range of approaches: relational template matching, two versions of distance measuring with edge probing, template matching with triple window detection, a statistical

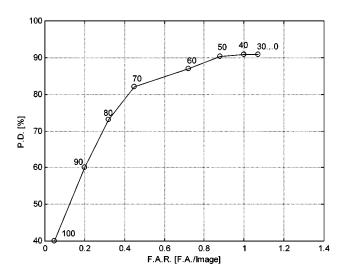


Fig. 14 Detection performance versus false-alarm rate. The numbers indicate the threshold decision level.

feature-based approach, two early model-based (computational model matching) approaches, and an analog optical correlator that located the peak in the correlation signal. 24-26 The second survey was made over four algorithms that were based on wavelets: morphological wavelet transform (MWT), the Gabor transform (GT), Gabor basis function (GBF), and a fusion of the MWT, GT, and GBF systems.

Due to the different databases and the lack of information about the evaluations performed, only a very crude comparison could be made in the present study. Most of the commonly used algorithms do not comply with the requirement of attaining a high PD and preserving a low FAR and therefore when compared with human observers fall far short.3,4

The fusion of MWT, GT, and GBTR systems, where the PD is 85% and the FAR is about 3.6 false alarms per scene, seems to give the best results, though again this is still far from the performance of human observers (PD of 83% with a FAR of 0.9).4 For our proposed algorithm, the PD was greater than 90% and the FAR was about 1 FA/scene. Although these figures also do not as yet compete with the human visual system, our results seem very promising and warrant further research.

References

- 1. B. Bhanu and R. D. Holben, "Model-based segmentation of FLIR
- images," *IEEE Trans. Aerosp. Electron. Syst.* **26**, 2–10 (1990). N. Ben-Yosef, B. Bahat, and G. Feigin, "Simulation of IR images of natural backgrounds," *Appl. Opt.* **22**, 190–193 (1983).
- 3. J. A. Ratches, C. P. Walters, R. G. Buser, and B. D. Guenther, "Aided and automatic target recognition based upon sensor inputs from image forming systems," *IEEE Trans. Pattern Anal. Mach. Intell.* 19, 1004–
- 4. D. P. Casasent and L. M. Neiberg, "Classifier and shift-invariant automatic target recognition neural networks," Neural Networks 8, 1117-1129 (1995).
- 5. B. J. Schachter, A. Lev, S. W. Zucker, and A. Rosenfeld, "An application of relaxation methods to edge reinforcement," *IEEE Trans. Syst. Man Cybern.* **7**, 813–816 (1977).

 6. D. K. W. Walters, "Computer vision model based on psychophysical
- D. K. W. Walters, Computer vision into the control of processing states of processing states
- 7. T. Poggio, "Early vision," in *Perspectives in Computing*, A. Rosenfeld, Ed., pp. 190–206, Academic Press, New York (1986).
- K. C. Marham, "Comparison of segmentation processes for object acquisition in infrared scenes," *IEE Proc. F, Commun. Radar Signal*
- Process. 136, 13–21 (1989).

 9. P. J. Burt, "Smart sensing in machine vision," in Machine Vision: Algorithms, Architectures, and Systems (Perspectives in Computing, Vol. 20), H. Freeman, Ed., Academic Press, New York (1988).
- C. R. Chatwin, R. K. Wang, and R. C. D. Young, "Wiener filter: synthetic discriminant function for target identification," Proc. SPIE **2484**, 616–635 (1994).
- 11. R. L. Harvey, Neural Network Principles, Prentice-Hall International
- Editions, Englewood Cliffs, NJ (1994).

 12. G. Buchsbaum, "An analytical derivation of visual nonlinearity,"
- IEEE Trans. Biomed. Eng. 27, 237–242 (1980).

 13. Z. Tianxu, P. Jiaxiong, and L. Zongjie, "An adaptive image segmentation method with visual nonlinearity characteristics," *IEEE Trans.* Syst. Man Cybern. 26, 619-627 (1996).
- 14. C. Joongho, H. Gunhee, J. M. Valverde, N. C. Griswold, J. F. Duque-Carrillo, and E. Sánchez-Sinencio, "Cork quality classification system," *IEEE Trans. Neural Netw.* **8**, 964–974 (1997).
- 15. S. Grossberg, "Contour enhancement, short term memory, and constancies in reverberating neural networks," Appl. Math. (Germany) **52**, 217-257 (1973).
- 16. S. Grossberg and M. Kuperstein, Neural Dynamics of Adaptive Sensory-Motor Control, Pergamon Press, New York, 1989.
- 17. B. Bhanu, S. Lee, and J. Ming, "Adaptive image segmentation using a genetic algorithm," IEEE Trans. Syst. Man Cybern. 25, 1543-1553
- 18. Y. J. Zheng, "Feature extraction and image segmentation using self-organizing networks," *Mach. Vision Appl.* 8, 262–274 (1995).
 19. S. Prasanna, C. Wilkins, and J. Yeager, "Threshold selection using

- Renyi's entropy," Pattern Recogn. 30, 71-84 (1997).
- E. K. Davies, Machine Vision: Theory, Algorithms, Practicalities, 2nd ed., Academic Press, New York, 1997.
- R. S. Gonzales, *Digital Image Processing*, Addison-Wesley, Reading, MA, 1987.
- S. R. Rotman, G. Tidhar, and M. L. Kowalczyk, "Clutter metrics for target detection systems," *IEEE Trans. Aerosp. Electron. Syst.* 30, 81–91 (1994).
- A. C. Copeland and M. M. Trivedi, "Models and metrics for signature strength evaluation of camouflaged targets," *Proc. SPIE* 3070, 194–199 (1997).
- K. Vijaya, "A tutorial survey of composite filter designs for optical correlators," Appl. Opt. 31, 4773–4798 (1992).
- D. Flannery, J. Loomis, M. Milkovich, and P. Kellin, "Applications of binary phase-only correlation to machine vision," Opt. Eng. (Bellingham) 27, 309–320 (1988).
- D. Flannery, J. Loomis, and M. Milkovich, "Transform-ratio ternary phase-amplitude filter formation for improved correlation discrimination," Appl. Opt. 27, 4079–4083 (1988).



Shlomo Greenberg received his BSc and MSc degrees with honors, and his PhD degree in electrical and computer engineering, from Ben-Gurion University, Beer-Sheva, Israel, in 1976, 1984, and 1998, respectively. He is currently a visiting researcher at Ben-Gurion University. His primary research interests are computer vision, image compression, signal and image processing, automatic target detection, pattern recognition, neural networks, and fuzzy logic.



Raanan Yehezkel received his BSc degree in 1998 from the Electrical and Computer Engineering Department of BenGurion University of the Negev, BeerSheva, Israel. His particular interests are computer vision, automatic target detection, and neural networks.



Yaniv Gurevich received his BSc degree in 1998 from the Electrical and Computer Engineering Department of Ben-Gurion University of the Negev, Beer-Sheva, Israel. His particular interests are computer vision, automatic target detection and neural networks.



Hugo Guterman received his Electronic Engineering degree from the National Technological University of Buenos Aires, Argentina, in 1978, and his MSc and PhD degrees in computer and electrical engineering from Ben-Gurion University, Beer-Sheva, Israel, in 1982 and 1988, respectively. From 1988 to 1990 he was a postdoctoral fellow at MIT. Since 1990, he has been at the Department of Computer and Electrical Engineering at Ben-Gurion

University. His research interests include control, image and signal processing, neural networks, and fuzzy logic.