Automatic target detection using binary template matching

Dong-San Jun
Korea Advanced Institute of Science and Technology
Department of Electrical Engineering
373-1 Guseong-dong
Yuseong-gu
Daejeon 305-701, Korea

Sun-Gu Sun
Agency for Defense Development
Daejeon, Korea

HyunWook Park, MEMBER SPIE
Korea Advanced Institute of Science and Technology
Department of Electrical Engineering
373-1 Guseong-dong
Yuseong-gu
Daejeon 305-701, Korea
E-mail: hwpark@athena.kaist.ac.kr

Abstract. This paper presents a new automatic target detection (ATD) algorithm to detect targets such as battle tanks and armored personal carriers in ground-to-ground scenarios. Whereas most ATD algorithms were developed for forward-looking infrared (FLIR) images, we have developed an ATD algorithm for charge-coupled device (CCD) images, which have superior quality to FLIR images in daylight. The proposed algorithm uses fast binary template matching with an adaptive binarization, which is robust to various light conditions in CCD images and saves computation time. Experimental results show that the proposed method has good detection performance. [DOI: 10.1117/1.1869997]

Subject terms: automatic target detection; binary template matching; clutter rejection; receiver operating characteristic (ROC).

Paper 040193 received Apr. 6, 2004; revised manuscript received Aug. 30, 2004; accepted for publication Sep. 30, 2004; published online Mar. 23, 2005.

1 Introduction

Human beings can easily recognize targets with good visibility in a peaceful environment. However, human perception capability deteriorates drastically in a restricted environment with low visibility or in hazardous circumstances such as on the battlefield. Thus, an accurate and robust automatic target recognition (ATR) technique is required to compensate for these limitations.

ATR is a military application of pattern recognition that detects and identifies types of targets. In order to model target features, an ATR system generally utilizes CCD images and/or forward-looking infrared (FLIR) images, and may also receive data from other information sources such as range information from a laser rangefinder, position information from a global positioning system (GPS), and weather data. At present, ATR is an extremely important component for various defense missions dealing with air-to-ground, ground-to-ground, and air-to-air applications. For instance, the ATR system of battle tanks significantly reduces the workloads of tank commanders by suggesting effective responses in real time.

An ATR system typically consists of several algorithmic components such as detection, segmentation, feature extraction, and classification,[1] as shown in Fig. 1. Among these components, we are interested in the detection stage, referred to as automatic target detection (ATD). Generally, ATD includes both initial detection and clutter rejection. Initial detection localizes areas in the image where a potential target is likely to be present. Initial detection is computationally simple, since it must be applied to the entire image. In general, the initial detection has a large number of false positives. The false positives are filtered out by clutter rejection, whereas real targets should not be. ATD is certainly one of the most important components, because it strongly affects the performance of the whole ATR system.

Many researchers have developed the ATD of FLIR images through approaches such as wavelet transforms,[2] Gabor filters,[3] morphology processing,[4] and a fusion of several algorithms.[5] In general, principal-component analysis (PCA) and multilayer perceptrons (MLPs) are used in the developed ATD methods.[6–11]

Installing an ATD system for both FLIR and CCD images on a battle tank would help the tank commander find the target more efficiently, because CCD images are generally superior to FLIR images during daytime. We propose an ATD algorithm to detect military vehicles in CCD images using binary template matching, which is fast and robust to various light conditions.

Section 2 briefly introduces a general template matching process and shows the necessity of binary template matching. Section 3 describes the proposed target detection method including binary template matching. In section 4, clutter rejection is performed to classify the candidate target chips into targets and false positives. In section 5, the proposed ATD method is evaluated experimentally. Finally, conclusions are presented in section 6.

2 Template Matching

The template-matching process finds the location having the best similarity between the template and the image area over which the template is positioned. The correlation (corr) between the template and the image is generally used as a measure of similarity for template matching, and is defined as

$$\mathrm{corr}(a,b)=\frac{E[(a-E[a])(b-E[b])]}{\mathrm{sd}[a]\mathrm{sd}[b]}, \tag{1}$$

where $E[x]$ and $\mathrm{sd}[x]$ are the mean and the standard deviation of a data set $\{x\}$, respectively. Since computation of

0091-3286/2005/$22.00

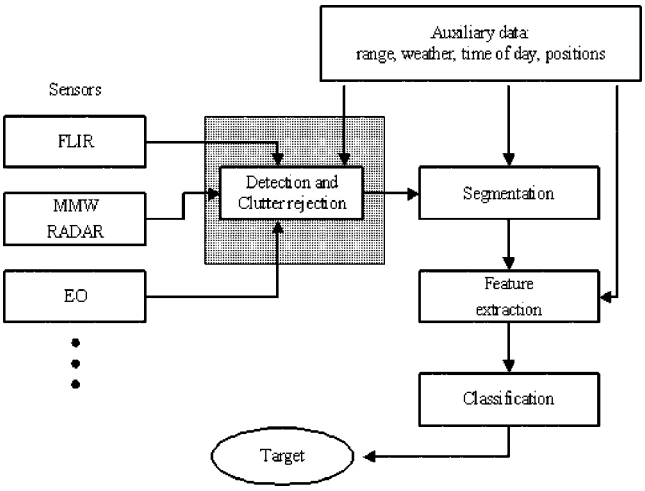

Fig. 1 Block diagram of a typical ATR system.

the correlation coefficient is expensive, the sum of absolute differences (SAD) or other similar measures can be used to reduce the computational burden. In order to further reduce the computation time, template matching is performed with a skip size, as shown in Fig. 2. We use correlation-based adaptive predictive search (CAPS) to reduce the computation time when performing template matching.[12] CAPS is based on the statistics of the template and utilizes the high correlation between neighboring pixels.

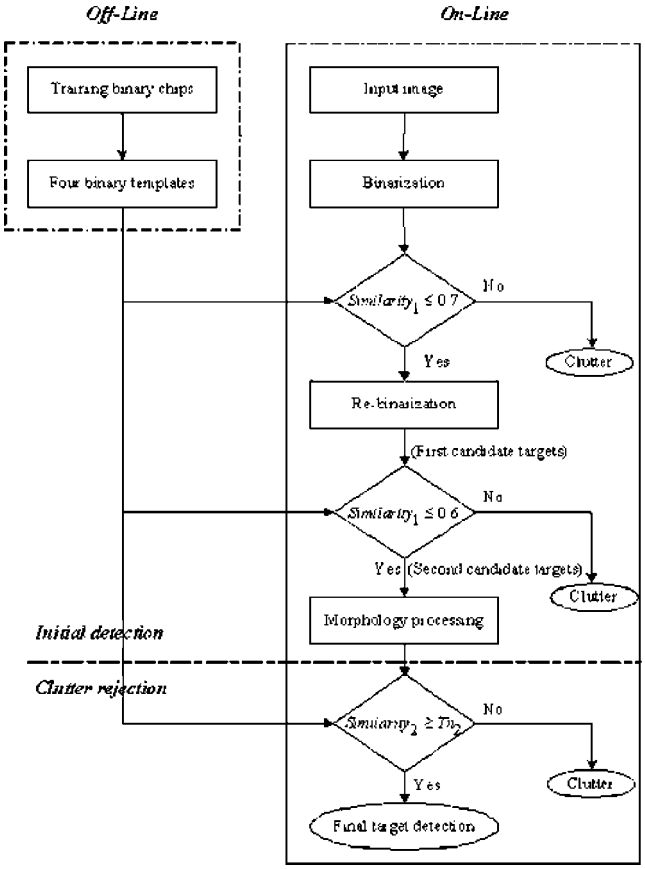

Fig. 2 Overall block diagram of the proposed ATD algorithm.

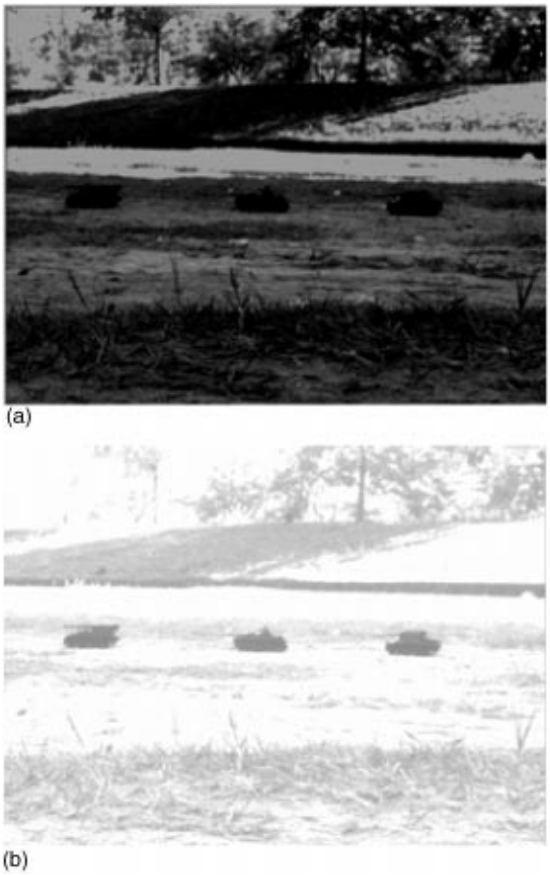

Fig. 3 CCD image with (a) dark background, (b) bright background.

3 The Proposed Target Detection Method

We propose a binary template matching based on the properties of the CCD images, which are susceptible to light conditions, as shown in Fig. 3. In our application, we assume the real targets are darker than the background. We acquired monochrome CCD images of five targets, namely, two battle tanks and three armored personnel carriers, as shown in Fig. 4, with various aspect views. A total of 72 images for each target were taken for different views with an increment of 5 deg on a clear background. We converted the acquired gray-level images into binary chips. We removed both antenna and gun barrels to avoid partially inconsistent features and filled holes in the binary chips, with which binary templates are defined as shown in Fig. 5. After obtaining the binary chips, we computed four binary templates using nonsupervised learning, which could represent all binary chips in the template-matching process. We investigated the relation between the detection rate and the number of binary templates, as shown in Fig. 6, which shows the detection rate is adequate if the number of binary templates is four. Therefore, we selected this number of binary templates in the present study.

Fig. 4 Five targets used in this article.

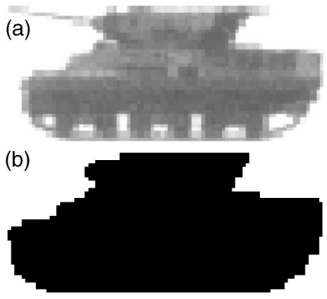

Fig. 5 An example of a target image: (a) original image, (b) binary image.

Before describing nonsupervised learning, we introduce a similarity feature defined in terms of AND and XOR logic operations on binary images, which are used as a classification criterion. The similarity between binary images A and B is defined as follows:

$$\text{Similarity}=\left[\frac{\#(A\ \text{AND}\ B)}{\#(A\ \text{XOR}\ B)+1}\right], \tag{2}$$

where $\#(X)$ denotes the number of pixels having 1 in the binary image X.

We classify the binary chips into several groups according to similar characteristics defined in Eq. (2). Based on Fig. 7, the nonsupervised learning procedure is given as follows:

(a) Randomly select four initial templates from the binary chips.

(b) Cluster the binary chips into four groups using similarity features: Four similarity values according to Eq. (2) are computed between an input chip and four initial templates, respectively. The initial template having the highest similarity among the four determines the group of the binary chip. The binary chip is then classified into one of four groups, each of which is represented by a template.

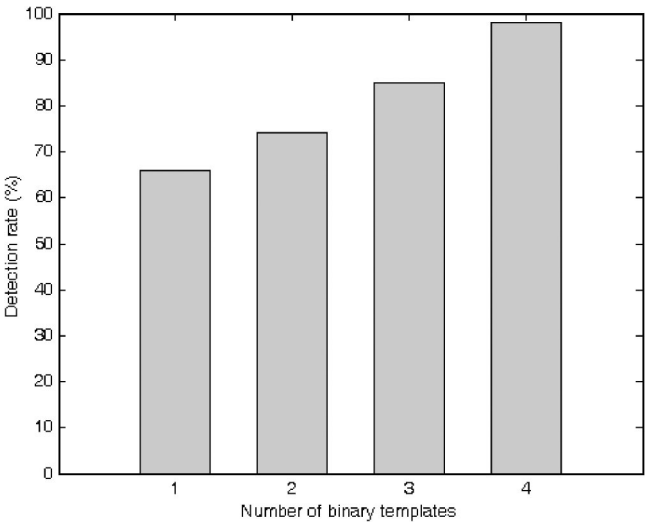

Fig. 6 Detection rate with respect to the number of binary templates.

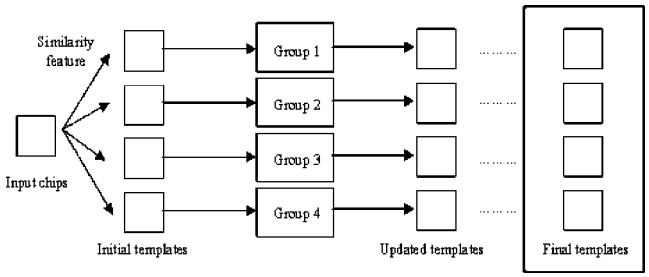

Fig. 7 Block diagram of nonsupervised learning.

(c) After clustering, the binary templates are updated for the four groups. Let the number of binary chips in a group be N_c. If the number of binary chips having target pixels at (x,y) is larger than or equal to $\frac{1}{2}N_c$, the specific coordinate (x,y) is regarded as a valid pixel and updated as a target pixel in the updated template, as shown in Fig. 8.

(d) Replace the initial templates with the updated templates in Fig. 9, and repeat the steps from (b) to (d) until there is no further change in the template update.

When we perform the initial detection, we convert the input image into a binary image using adaptive thresholding. Global thresholding for image binarization may result in distortion in a real image whose intensities vary gradually due to illumination. Therefore, we divided the input image into many blocks. The size of the blocks is determined by the largest target size, and adaptive thresholding is applied to each block. Then, thresholding values for binarization are computed by mean (mean) and standard deviation (std) of each block as follows:

$$\text{Thresholding value}=\text{mean}-\text{std}. \tag{3}$$

After image binarization, we define the skip size of template matching for each binary template. Finding the optimal skip size is based on the statistics of the template. Each binary template is expanded with a white background to calculate the parameter XOP as follows:

$$\text{XOP}(u,v)=\frac{\#(T(x,y)\,\text{XOR}\,T_E(x-u,y-v))}{\text{area of } T(x,y)}, \tag{4}$$

where $T(x,y)$ is the binary template and $T_E(x,y)$ is the expanded template. The resulting $\text{XOP}(u,v)$ has a valley at (0,0), and we define the skip width and height as one-half

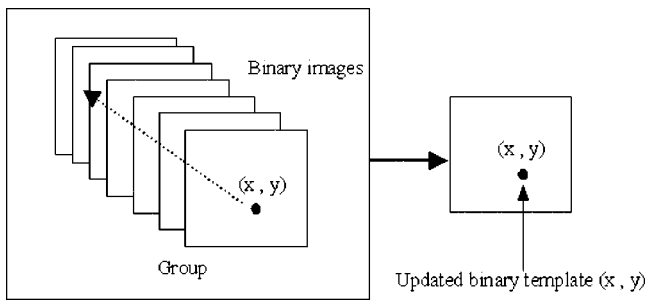

Fig. 8 Procedure to update the binary template from binary images in a group.

Fig. 9 Four binary templates.

the horizontal width W and one-half the vertical height H, respectively, at a cutoff value, as shown in Fig. 10. We set the cutoff at 0.6, and the four horizontal and vertical skip sizes are correspondingly obtained from four binary templates, as shown in Table 1.

The initial detection is performed only to localize candidate targets; the precise location of the candidate targets is not determined. The exact location is refined in the clutter rejection stage. We carry out binary template matching using four binary templates with corresponding skip sizes. In this initial detection stage, the candidate targets must include as many real targets as possible while reducing the total number of candidate targets. We define a criterion

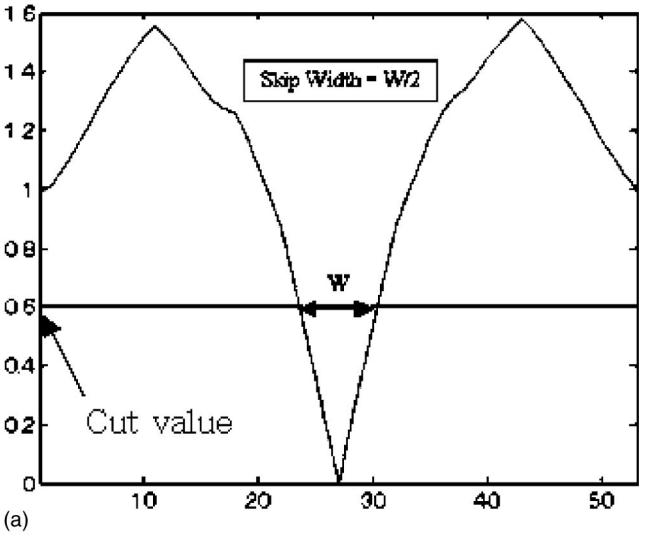

Fig. 10 Graphical representation to determine the horizontal and vertical skip sizes: (a) horizontal XOP and (b) vertical XOP.

Table 1 Skip sizes of four binary templates.

Template	Skip width	Skip height
1	3	4
2	12	4
3	11	4
4	5	4

Similarity_1 to detect the first candidate targets by template matching as follows:

$$\text{Similarity}_1 = \frac{\#(T \text{ XOR } A)}{\text{area of binary template}} \leqslant \text{Th}_1, \tag{5}$$

where T is the binary template and A is the input binary image over which the template is positioned in the test image. We set Th_1 to 0.7 in consideration of sensor noise and image distortion in the binary conversion.

A large number of extracted candidate targets increases the computation time for the clutter rejection. Thus, we reduce the number of candidate targets before carrying out the clutter rejection. We choose the expanded block for each first candidate target as shown in Fig. 11. The first candidate target block is expanded by the skip width in the horizontal direction and the skip height in the vertical direction. Because the first candidate targets give rough information on the target localization, we can assume that the expanded block area completely includes the real targets. We then reperform binarization of the original image on the expanded block with a more precise thresholding value, which can effectively discriminate the real targets from the background in the expanded block area. Thresholding values for the expanded block are computed by the difference between the mean and standard deviation of the expanded block, as described in Eq. (3). We thereafter reapply binary template matching to the expanded block with Eq. (5), where Th_1 is 0.6, which corresponds with the condition of Fig. 10. In addition, satisfying Eq. (5), we perform binary dilation on the second candidate target block, which is termed the second candidate target, with a 3×3 structuring element to reduce noise and loss of target pixels. As shown in Fig. 12, the number of second candidate targets is remarkably reduced from the 20 randomly selected test images.

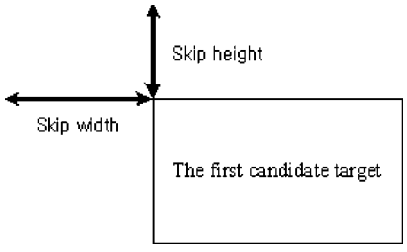

Fig. 11 Choice of the expanded block for the first candidate target.

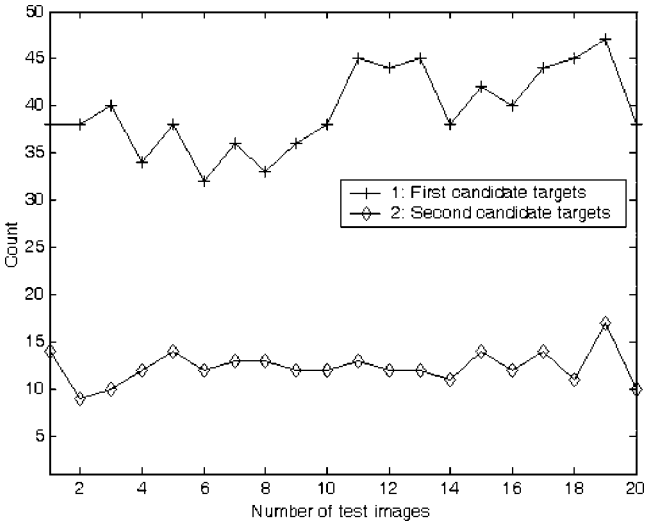

Fig. 12 Reducing the number of candidate targets.

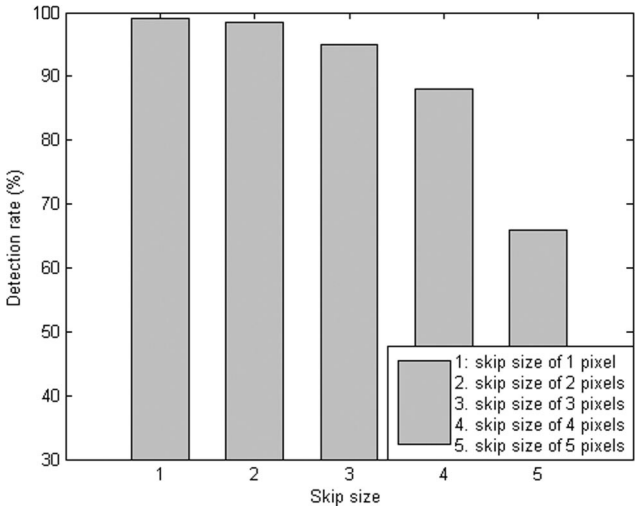

Fig. 14 Detection results for various skip sizes.

4 Clutter Rejection

We again perform template matching, with a skip size of 2 pixels on the candidate block, and with a more accurate criterion Similarity_2 as follows:

$$\text{Similarity}_2 = \frac{\frac{\#(T \text{ AND } A)}{\text{area of } T} + \frac{\text{chip size} - \#(T \text{ XOR } A)}{\text{chip size}}}{2} \times 100. \quad (6)$$

The range of Similarity_2 is from 0 to 100. If Similarity_2 is smaller than the thresholding value Th_2, the binary image is classified as clutter, as shown in Fig. 2. The thresholding value of Th_2 is 85, which is empirically determined here to yield a low false-positive rate. Figure 13 shows a receiver operating characteristic (ROC) curve measured with $\text{Th}_2 = 85$. Considering the experimental results in Figs. 14 and 15, we select an optimal skip size of 2 pixels for the detailed template matching.

5 Experimental Results

To assess the performance of the suggested approach, the proposed detection method was applied to numerous CCD test images having various light conditions and Gaussian noise. In the experiment, the size of the binary template is 35×18 and the input image is 320×240.

First, we generated test images comprising various situations such as multiple targets, nearby targets, and overlapped targets in a real background. Second, we modified the test images for various light conditions, such as bright and dark test images, as follows:

$$f_p(m,n) = \begin{cases} 255 & \text{if } f_o(m,n) + \delta > 255, \\ 0 & \text{if } f_o(m,n) + \delta < 0, \\ f_o(m,n) + \delta & \text{otherwise,} \end{cases}$$

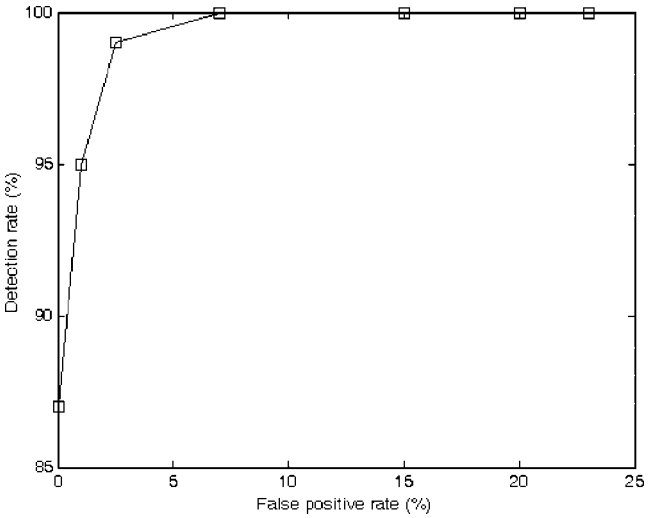

Fig. 13 Detection rate and false-positive rate.

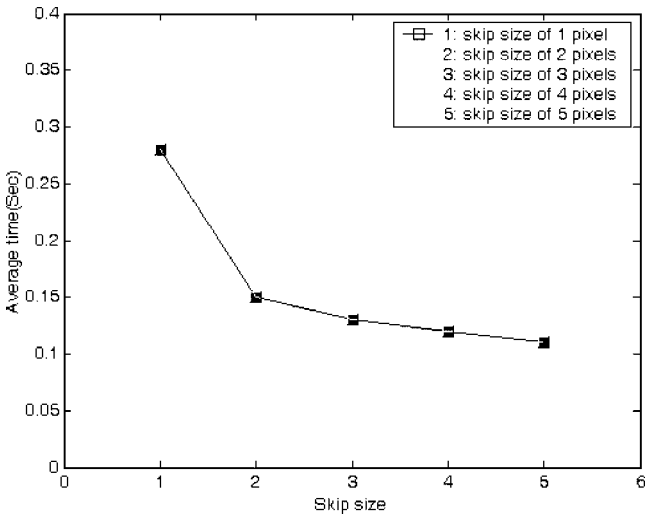

Fig. 15 Average computation time for various skip sizes.

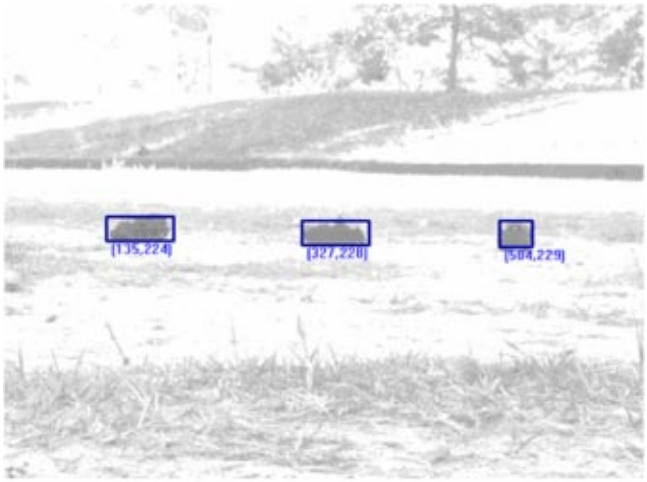

Fig. 16 Example of a bright test image (mean value: 231).

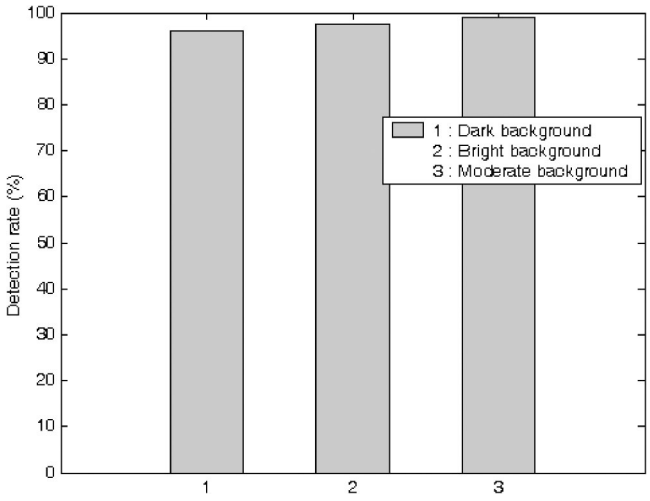

Fig. 18 Detection results for various light conditions.

where $f_o(m,n)$ is the original image and $f_p(m,n)$ is the modified image simulated for the light conditions. In the experiment, the bright and dark test images were obtained with $\delta=100$ and -100, respectively, as shown in Figs. 16 and 17. The range of detection rate was from 96% to 98% for the various light conditions, as shown in Fig. 18. Finally, we added Gaussian noise with various standard deviations (std). We then investigated the detection rate with respect to the standard deviation of the added noise, as shown in Fig. 19. The detection rate approached 98% until the standard deviation was 25, while the false-positive rate was less than 2%, as shown in Fig. 20.

We compared our method with a strict binary template matching process without the enhancement of an adaptive skip size or a second candidate target scheme. As anticipated, the computation time of the strict binary template matching was near 2.2 s, which is slower than that of our method, primarily because the strict binary template matching does not exploit the adaptive skip size. In addition, the false-positive rate is larger than that of the proposed method. Finally, we compared the proposed method with morphology processing,[4] which has been used in various

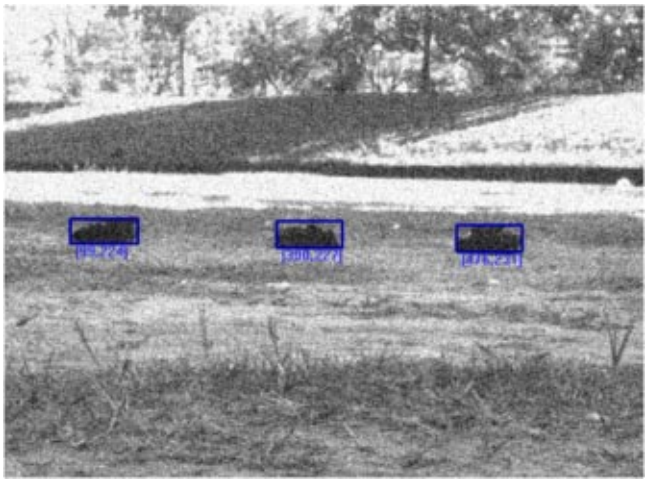

Fig. 19 Example of a test image with Gaussian noise.

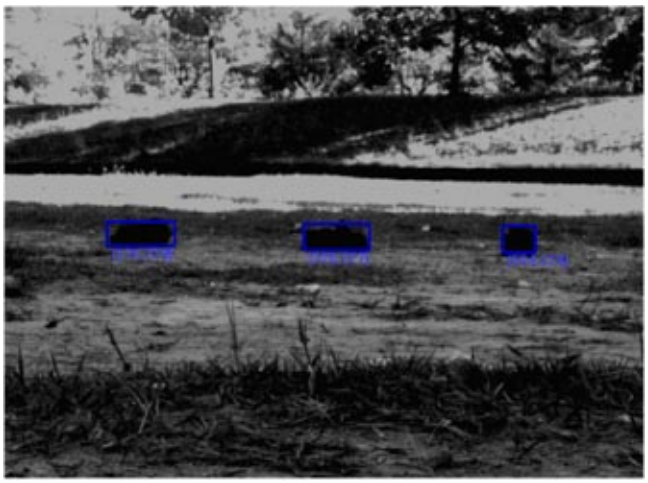

Fig. 17 Example of a dark test image (mean value: 59).

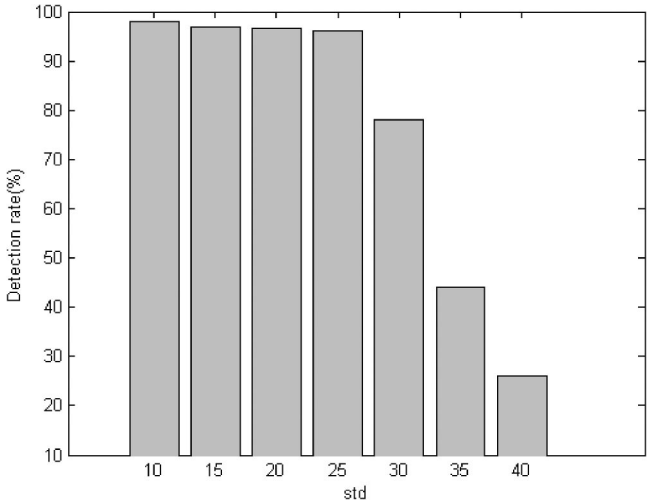

Fig. 20 Detection result when random Gaussian noise is added.

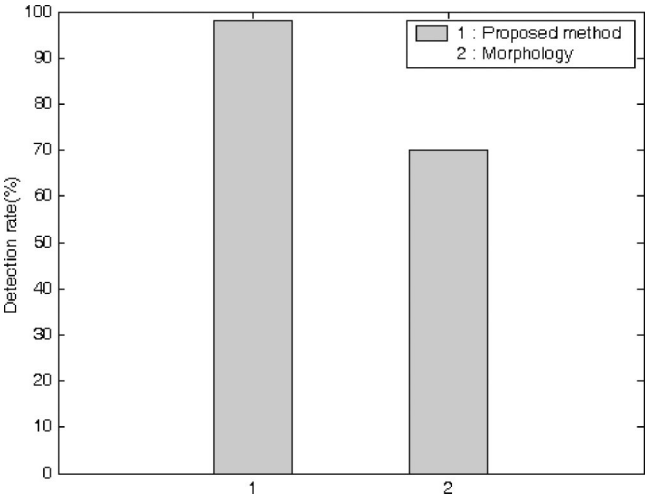

Fig. 21 Comparison of the proposed method and the morphology processing method.[4]

ATR systems for many years. As shown in Fig. 21, the detection rate of our algorithm is near 98%, while that of morphology processing is smaller than 70%.

Lastly, the average computation time of our algorithm is 0.13 s for a 320×240 image, while that of the fast (1-D) morphology processing is near 1 s in a Pentium III (2 GHz with 512 Mbyte of memory).

6 Conclusions

This paper has proposed a new automatic target detection method for CCD images in ground-to-ground scenarios. The proposed algorithm utilizes binary template matching, which is performed with an adaptive skip size obtained from the template in order to reduce the computation time. After initial detection, a detailed search is executed on the extracted candidates. Then, a predefined similarity measure is computed and the candidate target chips are classified into real targets and clutter. We applied the proposed method to numerous test images with various light and noise conditions. Our proposed method could acceptably detect military vehicles in CCD still images.

Acknowledgment

This work was supported by the Agency for Defense Development (ADD) in Korea.

References

1. S. G. Sun and H. W. Park, ''Automatic target recognition using boundary partitioning and invariant features in FLIR images,'' *Opt. Eng.* **42**(2), 524–533 (2003).
2. X. Yu, I. Reed, W. Kraske, and A. Stocker, ''A robust adaptive multispectral object detection by using wavelet transform,'' in *ICASSP-92, 1992 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Vol. 5, pp. 141–144 (1992).
3. D. Casasent and J. Smokelin, ''Real, imaginary, and clutter Gabor filter fusion for detection with reduced false alarm,'' *Opt. Eng.* **33**(7), 2255–2263 (1994).
4. Q. H. Pham, T. M. Brosnan, and M. J. T. Smith, ''Sequential digital filters for fast detection of targets in FLIR image data,'' in *Automatic Target Recognition VII, Proc. SPIE* **3069**, 62–73 (1997).
5. D. Casasent and A. Ye, ''Detection filters and algorithm fusion for ATR,'' *IEEE Trans. Image Process.* **6**(1), 114–125 (1997).
6. L. A. Chan, N. R. Nasrabadi, and D. Torrieri, ''Selectively optimized networks for automatic clutter rejection,'' in *Int. Joint Conf. on Neural Networks, 1999. IJCNN'99*, Vol. 5, pp. 3161–3164 (1999).
7. L. A. Chan, S. Z. Der, and N. M. Nasrabadi, ''Joint compression and discrimination algorithm for clutter rejection,'' *IEEE Trans. Image Process.* **1**, 525–528 (2001).
8. S. A. Rizvi, T. N. Saadawi, and N. M. Nasrabadi, ''A modular clutter rejection technique for FLIR imagery using region-based principal component analysis,'' in *Proc. 2000, Int. Conf. on Image Processing*, Vol. 2, pp. 475–478, IEEE (2000).
9. L. A. Chan, S. Z. Der, and N. M. Nasrabadi, ''Improved target detector for FLIR imagery,'' in *2003 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP'03). Proceedings*, Vol. 2, pp. 401–404 (2003).
10. L. A. Chan, S. Z. Der, and N. M. Nasrabadi, ''Automatic target detection using dualband infrared imagery,'' in *2003 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing. ICASSP'00. Proceedings*, Vol. 6, pp. 2286–2289 (2000).
11. L. A. Chan, N. M. Nasrabadi, and D. Torrieri, ''Bipolar eigenspace separation transformation for automatic clutter rejection,'' in *1999 Int. Conf. on Image Processing. ICIP 99. Proceedings*, Vol. 1, pp. 139–142 (1999).
12. S. Sun, H. W. Park, David R. Haynor, and Y. Kim, ''Fast template matching using correlation-based adaptive predictive search,'' *Int. J. Imaging Syst. Technol.* **13**(3), 169–178 (2003).

Dong-San Jun received his BS degree in electrical engineering from Pusan National University, Pusan, Korea, in 2002, and his MS degree in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2004. He has been a researcher and a member at Electronics and Telecommunications Research Institute (ETRI) in Korea since 2004. His current research interests include image processing, pattern recognition, video coding, MPEG 7, and MPEG 21.

Sun-Gu Sun received his BS and MS degrees in electrical engineering from Hanyang University, Korea in 1987 and 1989, respectively. He received his PhD degree in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Korea. He has been a researcher and a senior member at Agency for Defense Development (ADD) in Korea since 1989. His current research interests include image processing, pattern recognition, computer vision, and automatic target recognition.

HyunWook Park is a professor in the Department of Electrical Engineering at KAIST. He received his BS degree in electrical engineering from Seoul National University, Seoul, Korea, in 1981, and his MS and PhD degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Seoul, Korea, in 1983 and 1988, respectively. There he has been a professor in the Electrical Engineering Department since 1993 and an adjunct professor in the Biosystem Department since 2003. He was a research associate at the University of Washington from 1989 to 1992 and was a senior executive researcher at the Samsung Electronics Co., Ltd., from 1992 to 1993. He is a senior member of the IEEE and a member of SPIE. He has served as an associate editor for the *International Journal of Imaging Systems and Technology*. His current research interests include image computing systems, image compression, medical imaging, and multimedia systems.