

A morphological approach of target detection on perspective plane

Soo-Chang Pei^{a,*}, Chin-Lun Lai^b

^a*Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, Republic of China*

^b*Department of Electronic Engineering, Oriental Institute of Technology, Taipei, Taiwan, Republic of China*

Received 15 June 1998; received in revised form 15 January 2001

Abstract

This paper presents a morphological approach of target detection on the perspective-distorted plane for image analysis applications. By modifying morphological operations with a dynamic-varying structuring element, distorted targets with variant contrast values, shape deformations, and different location/orientation can be detected successfully on the perspective plane even though they are in cluttered background. Two experimental simulation results are presented to show that our algorithm works successfully in detecting the targets from their blurred perspective plane image. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Mathematical morphology; Perspective transformation; Variable structuring element; Target detection

1. Introduction

Object detection is an important step in many image processing and analysis applications. After undergoing the detection process, those regions of targets in the input scene can be located for further processing such as feature extraction or scene analysis. Many applications such as robot vision need ability to recognize the 3-D world objects on their 2-D perspective image plane. Traditional recognition work uses the model-based matching technique in which features are extracted from 3-D objects and are used for model matching. However, due to the greater complexity of 3-D perspective

transformation, recognizing 3-D objects directly from their perspective transformation may be intractable, and thus some efforts are made to approximate the perspective transformation to simplify this problem [11,1].

Mathematical morphology has played an increasingly important role in object recognition applications. The main idea of using morphological operations is that solving recognition problems can be reduced to ‘set’ processes. Thus the advantage of using the mathematical morphology is its simplicity and easy implementation in parallel. Main principles of using the morphology in recognition include hit-or-miss operation, skeleton-matching, etc. Based on these methods, however, some problems still exist in designing a suitable feature model when the targets are distorted in a cluttered background. Moreover, exact selection of structuring elements for different applications is unavailable.

*Corresponding author. Tel.: +886-2-2363-5251; fax: +886-2-2367-1909.

E-mail address: pei@cc.ee.ntu.edu.tw (S.-C. Pei).

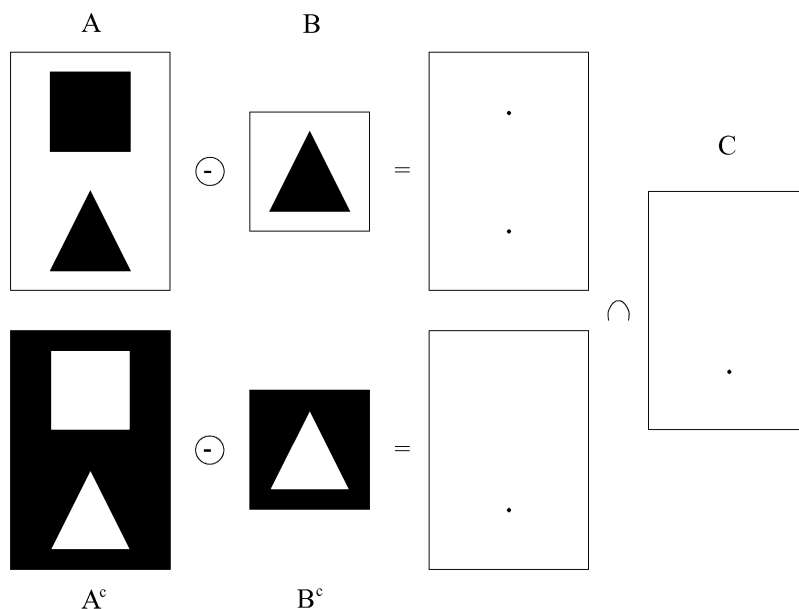


Fig. 1. The complete morphological form of target detection.

The goal of this paper is to propose a morphological approach of target detection algorithm to pick up 3-D targets from cluttered perspective plane. By adaptively varying the structuring element (size and/or shape) according to the perspective projection of the template, the proposed morphological detection algorithm performs a suitable and easy way in detecting the 3-D targets from their perspective images. Two experimental examples are presented at the end to demonstrate the performance and efficiency of the proposed algorithm. The contribution of this paper is that an effective morphological approach with parallel structure and less complexity is proposed to detect the targets in the distorted perspective plane with cluttered background.

This paper is organized as follows. In Section 2, the morphological recognition operations are explained and the principles of perspective transformation are described in Section 3. Section 4 gives the proposed algorithm which combines the perspective transformation and morphological operation to detect the interested objects in the input scene. Experimental results and discussions are

provided in Section 5. Finally, conclusions are made in Section 6.

2. Mathematical morphology in target detection

2.1. The generic target detection algorithm

Mathematical morphology has been widely used in digital image processing, and much work has been done in image analysis applications recently [14,13,5,7]. To use the morphology in target detection, it is simple to adopt a morphological erosion in recognition; however, it merely provides a way in which only the foreground of the template is considered, as shown in Fig. 1. A more precise one for target detection, however, should consider the background of the structuring element as well as its foreground. Thus, a more complete morphological operation for target detection [3] can be defined as follows. Let A and C denote binary images and B denote a binary structuring element. The complement set of A is denoted as A^c . Then a shape B , defined in a window W , is said to occur in the

image A if

$$C = (A \ominus B) \cap (A^c \ominus B^c) \quad (1)$$

is not empty, where \ominus is a morphological erosion and B^c is the complement of B and is defined as

$$B^c = W - B \quad (2)$$

and \ominus is a morphological erosion. Algorithm described in Eq. (1) is illustrated in Fig. 1. It is observed that C consists of one point if and only if A is a translate of B . Due to the use of a window W , we can detect many copies of B in the original image. However, one point in C corresponds to one location of a translate of B .

Target detection in binary morphology can be extended to gray-scale by using the umbra and top-surface concept. The gray level of the input is viewed as the height of the surface. Let $C(x, y)$ represent the target detection error function, then the most likely position of B in A is choosing the minimum of C which is represented as

$$C(x, y) = (A^c \ominus B^c)(x, y) - (A \ominus B)(x, y), \quad (3)$$

where

$$(A \ominus B)(x, y) = \min_{i,j} [A(x+i, y+j) - B(i, j)] \quad (4)$$

and

$$(A^c \ominus B^c)(x, y) = \max_{i,j} [A(x+i, y+j) - B(i, j)]. \quad (5)$$

2.2. Modified morphological recognition algorithm

The main principle of generic target recognition is to use both foreground and background as two patterns. Thus only those objects which exactly match the template (structuring element) both in the foreground and the background can be extracted. However, in many image processing applications, due to the quantization effect, it is impossible to acquire the exact form of objects. Thus, it is necessary to modify the generic morphological recognition algorithm. To fit the real image cases, a modified morphological operation, which adopts the concept of cardinality as measurement in target detection problem, is proposed as follows. Let A and B be defined as previously. A modified

erosion (ME) of A by B is defined as

$$A \tilde{\ominus} B(x, y) = \frac{\text{Card}(A_w \cap B(x, y))}{\text{Card}(A_w)}, \quad (6)$$

where $\text{Card}()$ is the cardinality value and A_w is all the elements of A within the window W . The reason to use Eq. (6) in preference to the standard erosion is that it is based on a simple Euclidean-distance like template matching and thus is much simple in maintaining the property of parallel implementation. It should be noted that the resultant values of the ME range from zero (no match) to one (match completely). Thus target detection problem involves choosing the most possible object. Using Eq. (6) as a measurement, the target recognition of B in A now becomes choosing the maximal value of the success function which is defined as

$$\tilde{C} = (A \tilde{\ominus} B)^p \diamond (A^c \tilde{\ominus} B^c)^p, \quad (7)$$

where \diamond represents an addition or multiplication operation and p represents a weighting power value. It should be noted that if the values of success function do not exceed a threshold, no object could be said to fit the template although it has a maximal success value. An example of using the modified recognition algorithm with $p = 3$ is illustrated in Fig. 2. It is observed that the most likely location of the triangle object is the position with a peak value as shown in Fig. 2(c).

To be extended to the 3-D (gray-scale) object recognition problem, the concept of umbra is also applied to both the foreground and the background. Thus, calculating the cardinality of a 3-D contour involves counting all the pixels under the top surface. If the threshold decomposition method is used, recognition of the 3-D objects can be implemented by processing each slice of the structuring elements, summing or multiplying all the success functions of each slice and picking up the maximum success function. The success function of a 3-D object of the structuring elements can be represented as

$$\tilde{C}_{3-D} = \sum_i \tilde{C}_i$$

or

$$\tilde{C}_{3-D} = \prod_i \tilde{C}_i. \quad (8)$$

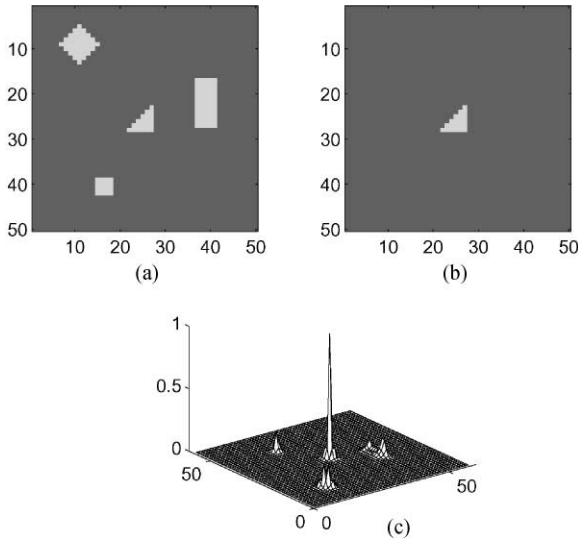


Fig. 2. Example of target detection using the modified morphological success function. (a) Input patterns. (b) Template (structuring element), and (c) results of the success function \tilde{C} .

2.3. Properties of the modified morphology

Property description of the proposed morphological operations can be divided into two parts. In the first part, since operations used in pre-processing and object brightening process (MCMO, bi-thresholding, and closing-opening) are conventional morphological operations, they satisfy all the properties of conventional morphology such as increasing, (anti)extensive, idempotent, etc. [13], and thus are not reviewed here again. Besides, one of the most important property for this part is the threshold decomposition technique [14] which can be used in implementing those functions in parallel.

In the second part, since the modified erosion (ME) described in Eq. (6) map sets into real value and are not similar to the traditional morphology, some new properties related to this modified morphology are derived as follows.

Property 1. *The ME operations increase for the structuring element, i.e. for $B_1 \subseteq B_2$, we have $A \tilde{\ominus} B_1 \leq A \tilde{\ominus} B_2$.*

Proof. This proof is trivial since $B_1 \subseteq B_2$ implies that $A \cap B_1 \subseteq A \cap B_2$, thus $\text{Card}(A \cap B_1) \leq$

$\text{Card}(A \cap B_2)$, which leads to the consequence of Property 1. \square

Property 2. *Let $B = C \cap D$, then $A \tilde{\ominus} B \leq \text{Min}(A \tilde{\ominus} C, A \tilde{\ominus} D)$, where the minimal set in $\text{Min}()$ is the set with smallest area.*

Proof. Since $B = C \cap D$, then $A \cap B = A \cap (C \cap D) = (A \cap C) \cap (A \cap D) \subseteq \text{Min}$. $\text{Set}[A \cap C, A \cap D]$, thus we have the following result that $A \tilde{\ominus} (C \cap D) \leq \text{Min}(A \tilde{\ominus} C, A \tilde{\ominus} D)$. \square

Property 3. *Let $B = C \cup D$, then $A \tilde{\ominus} B \geq \text{Max}(A \tilde{\ominus} C, A \tilde{\ominus} D)$, where the maximal set in $\text{Max}()$ is the set with largest area.*

Proof. This proof is similar to Property 2. Since $A \cap (C \cup D) = (A \cap C) \cup (A \cap D) \supseteq \text{Max}$. $\text{Set}[A \cap C, A \cap D]$, the result of this property can be easily derived. \square

Properties 2 and 3 deal with the distribution from the right, there are similar distribution properties from the left, i.e. $A = (A_1 \cap A_2)$ or $A = (A_1 \cup A_2)$, which lead to the following two properties.

Property 4. *Let $A = (A_1 \cap A_2)$, then $A \tilde{\ominus} B \leq \text{Min}(A_1 \tilde{\ominus} B, A_2 \tilde{\ominus} B)$.*

Property 5. *Let $A = (A_1 \cup A_2)$, then $A \tilde{\ominus} B \geq \text{Max}(A_1 \tilde{\ominus} B, A_2 \tilde{\ominus} B)$.*

Proof. The proof of these two properties are trivial, and are skipped. \square

Property 6. *The ME is said to be invariant under some conditions.*

Proof. First, the translation invariance is trivial since the perspective transformation is linear and ME also satisfies the ‘set’ theory. Secondly, assume that the truncation error can be ignored, the rotation invariance property can be obtained if a multi-directional structuring element model,

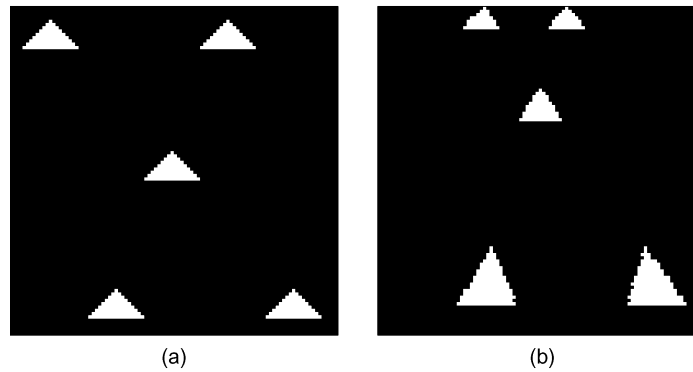


Fig. 3. (a) An image which contains five similar objects and (b) its perspective plane. It is observed in (b) that both the shape and size of the objects are changed.

16-direction or 32-direction [10], is adopted. Thus the success function is rewritten as

$$\tilde{C} = \max_k [\tilde{C}_k], \quad (9)$$

where \tilde{C}_k represent the success function value at the k th direction. Finally, the scaling invariance is also possible if normalization skills are adopted or the corner structure pattern instead of block pattern of the object is selected.

The invariance property is important especially in pattern recognition application while the distribution properties is helpful in reducing the computational complexity by decomposing the object/structuring element into the union of small subsets.

3. The 3-D perspective transformation

The perspective transformation (PT) which governs the way that the 3-D information is translated onto the 2-D image, has its validity in photogrammetric engineering and surveying applications [6]. The complete principle of the PT (including forward and backward) is not reviewed here since it is well known and thus interested readers can easily pick them out from [6].

4. The proposed target detection algorithm

A particular target detection application which needs to consider the perspective transformation is

observed in the following example. A robot, which has a single camera with some known camera parameters, is designed to pick up the pre-determined object (template) in its captured image in which there exists a lot of objects with different shapes and sizes. Fig. 3 shows a perspective plane in which some objects of the same size and shape are contained. It is noted that at different locations, the same object may appear differently in their shapes and sizes in the perspective plane. Thus, applying the modified morphological algorithm to such input scene directly by using the template (structuring element) to detect the target is impossible. In order to simplify the analysis of the proposed recognition algorithm, objects used in this paper work are assumed to be on the same plane without considering rotation and/or scaling invariance models. Moreover, aberration and out of focus conditions due to the lens are also ignored.

To solve the above problem, an algorithm which combines the modified morphological operations and perspective transformation to detect the desired objects is proposed. Fig. 4 illustrates the schematic block diagram of the proposed algorithm. Each step is stated as follows.

The input scene first undergoes a pre-processing procedure for noise reduction and object enhancement. Then a modified morphological close minus open (MCMO) operation is used to remove differences in the mean value of large background regions and to make all objects brighter than their surrounding background. However, the structuring

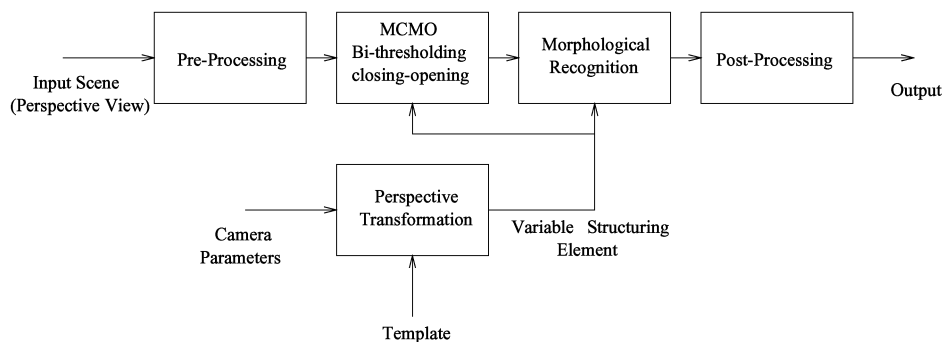


Fig. 4. Functional block diagram of the proposed target detection algorithm.

elements are replaced by the vertical and horizontal lines whose lengths vary dynamically according to their location on the perspective plane. After this process, regions corresponding to the targets can be enhanced while the other regions are depressed. Once the regions of interest are located, both a binary threshold procedure, which transforms the result to become a binary image, and an efficient alternating sequential filter (ASF) [9] are performed to make these regions clear to prevent the false alarm. After the above processes, only the candidate object regions, whose sizes are similar to that of the template, remained in the proceeded image while all the background clutter and noise are filtered out. A morphological detection algorithm based on Eq. (7) is then applied to the processed image for picking out the desired target. Finally, a post-processing procedure is helpful in further analysis processing.

It is known that the choice of the structuring element is very important in most morphological operations. The structuring element acts as a probe in extracting features of an input signal. In the proposed detection algorithm, it is impossible to use a unique structuring element to detect all the targets located at different positions in the input image if perspective distortions exist. Thus, variation of the structuring element is needed. The structuring element used both in the MCMO and the ME is determined by applying the perspective transformation on templates to obtain their exact representation of shapes and sizes at different positions. In order to obtain a suitable structuring

element for each position in the perspective plane, a 3-D template model should be constructed in advance. Then, by applying the perspective transformation to the template model at each position, a suitable structuring element can be easily obtained since the camera parameters and photographing conditions are known. That is, shape and size of the structuring element could be varied adaptively to fit the perspective view just by applying the perspective transformation on the template model. Once the foreground of the structuring element is determined as described above, a rectangular window with an area of two times the foreground region is used as the background to construct the completed structuring element.

5. Experimental results and discussions

Two experimental examples are included in this section to demonstrate the performance of the proposed recognition approach. The first example is to choose a particular brick in a cluster of bricks. The second example comes from traffic monitor system: given a road image, extract the cars in background clutter regions.

Example 1. Fig. 5 shows each step of the entire algorithm. Fig. 5(a) shows the input image, which is captured by a camera, with a lot of different shaped bricks. Our goal is to pick up the triangular cubes from the input image. Fig. 5(b) shows the result scene processed by the MCMO operation. It is

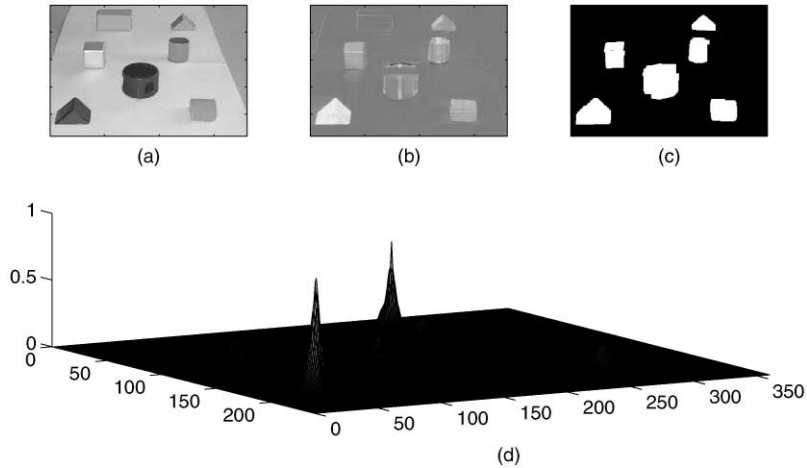


Fig. 5. (a) The input perspective image. (b) The candidate objects image after the pre-process and MCMO operation and (c) its binary image after ASF operations. (d) The final output of the target detection algorithm.

observed that only the candidate objects can be explored since the variant structuring element is used. Thus, the rectangular parallelepiped is removed from the candidate objects set since its size is not appropriate. Fig. 5(c) shows the result after taking bi-thresholding and ASF of Fig. 5(b). By applying the proposed recognition algorithm to Fig. 5(c), the triangle target can be extracted exactly as shown in Fig. 5(d) where two peaks appear at the target positions.

An important problem addressed in this example is that due to the perspective transformation, the 3-D geometric information of an object will be reduced to 2-D information. Thus, as observed in Fig. 6, a confused result is obtained for two different objects if they have the same projective form. In the proposed algorithm, the structuring element is obtained by applying the perspective transformation to all pixels of the template model. Although overlapping of points at a fixed position is possible, it can be regarded as the same structuring pixel and does not affect the recognition performance.

Example 2. Fig. 7(a) is an image captured by a handy-camera. After the pre-processing and the MCMO processes, regions of interest are enhanced and are shown in Fig. 7(b), with five car regions in

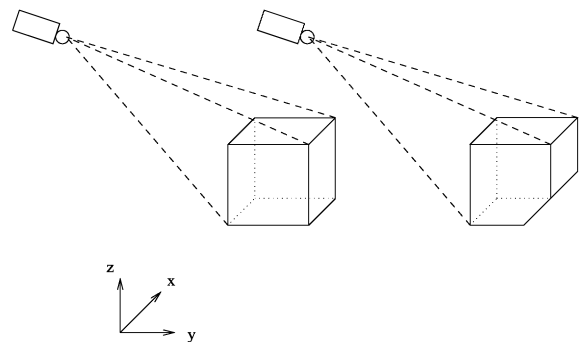


Fig. 6. An example of ambiguity in which two different objects results in the same perspective image.

it. The candidate object set image is then applied with the binary thresholding and closing-opening operations, as shown in Fig. 7(c), for further processing. Our goal is to pick up the cars in the plane. Using the same technique as in the previous example, all cars in the perspective plane can be detected successfully as shown in Fig. 7(d) where the pixels with highest grayscale represent positions of the cars. It should be noted that although there are bright regions obtained from the cluttered background, the result of the detection process is not affected significantly since they are outside the pre-determined detecting field. Moreover, by

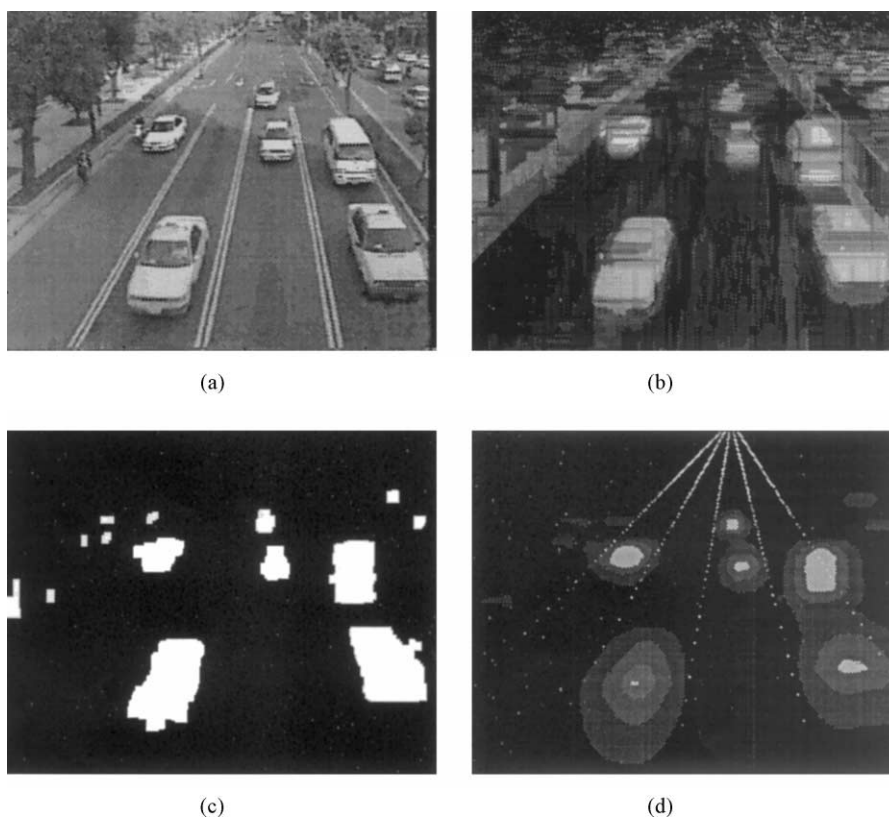


Fig. 7. (a) The captured image. (b) The resultant image after pre-processing and MCMO process. (c) The binary candidate objects after closing-opening operation, and (d) The output image after detection process. Note that five dotted lines are added in (d) to present the relative positions of the detected targets.

observing Fig. 7(d), the proposed algorithm will respond to a peak point for an exact object but a mash for a deformed object. This property leads to the possibility for further analyzing the correctness of the detection process. Consequently, since the region of car glass is also depressed due to its low contrast, an assumption should be made in this example that the distance of any two succeeding cars should be at least longer than the length of glass, thus avoiding a false alarm.

It should be noted that although there are successful systems, especially in latest results [12,8], in dealing with the traffic monitoring problems, the proposed algorithm offers an effective and simple approach in capturing the desired targets. The performance of our system is acceptable although it is

not designed to solving such problems. Due to the characteristic of the MCMO, targets in noisy and cluttered environment such as rain, fog, and moving shadow conditions, can be detected successfully although it may fail in other systems. This result also can be observed in the shadow condition of Figs. 5(a) and 7(a). Thus, it is useful in detecting targets in blurred environment such as detecting tanks on a battlefield via our detection algorithm. A comparison with the detection algorithm of [12] is made, which claims that the detection rate of 9090 images under different environment conditions (30 images for each weather condition) with 421 vehicles are used to evaluate the proposed algorithm. A successful detection is recorded if a vehicle is clearly seen by a person and a peak is marked in its corresponding location. The overall

Table 1
Detection rate of the proposed system tested under three different weather conditions

	Normal	Rainy	Shadows
Total vehicles	162	119	140
Vehicles detected	146	105	122
Vehicles missed	9	3	4
False detected	7	11	14
Detection rate	90.1%	88.2%	87.1%

result is shown in Table 1 and is found that the performance of the proposed algorithm is acceptable.

On the other hand, compared with other conventional morphological operations used in target detection such as top-hat or hit-miss [4,2], our algorithm is better in detecting the targets in noisy and blurred environment with perspective distortion, while maintaining flexible structure in parallel implementation.

Some skills can be used to enhance the performance and/or efficiency of the proposed algorithm. As stated previously, affine transformation or scaled-orthographic transformation can help in reducing the complexity of the perspective representation. Another way to reduce the complexity is to implement the morphological detection process by considering the object edge or skeleton information, whose geometry at different positions can be obtained via an algorithm similar to the proposed method, into the feature matching process. However, due to the high noise sensitivity of the skeleton/edge matching process, it is reasonable to expect that the proposed algorithm has a better performance especially in a cluttered background. By using the morphological operations, complexity of the recognition algorithm can be reduced since the ‘set’ operation is used, thus it is possible to implement the algorithm in parallel. On the other hand, since the MCMO operation is sensitive to the variation of the input scene, a more heuristic method may be needed to locate the candidate object regions exactly to increase the correctness of detection. Moreover, as long as the template model is known in advance, extension to recognizing a target from any position and any view angle

without locating at a fixed plane can be easily implemented.

6. Conclusions

A morphological target detection algorithm which can recognize the targets on the perspective plane is proposed. By varying the structuring element adaptively according to the geometry of the perspective transformed template, the proposed recognition algorithm is able to detect the targets from the cluttered perspective plane although there exist shape deformation. Two experimental examples are given to show that the proposed algorithm can detect the targets in a blurred environment successfully. Moreover, our algorithm can be easily extended to rotation and scaling invariance model for more practical applications. The proposed algorithm is a simple and effective method in adaptively detecting the 3-D target from its 2-D perspective distorted plane with cluttered background and acts as a good guide in similar applications.

References

- [1] J.Y. Aloimonos, Perspective approximations, *Image Vision Comput.* 8 (3) (August 1990) 179–192.
- [2] H. Arnarson, L.F. Pau, PDL-HM: morphological and syntactic shape classification algorithm, *Machine Vision and Applications* 7 (2) (1994) 59–68.
- [3] J. Cardillo, M.A. Sid-Ahmed, Target recognition in a cluttered scene using mathematical morphology, *Pattern Recognition* 29 (1) (January 1996) 27–49.
- [4] H. Florence, C.P. Fr, M. Juliette, Textural analysis by mathematical morphology transformations: structural opening and top-hat, *IEEE International Conference on Image Processing*, Vol. 3, September 1996, pp. 49–52.
- [5] C.R. Giardina, E.R. Dougherty, *Morphological Methods in Image and Signal Processing*, Prentice-Hall, Englewood Cliffs, New Jersey, 1988.
- [6] R.M. Haralick, Using perspective transformations in scene analysis, *Comput. Vision Graphics Image Process.* 13 (1980) 191–221.
- [7] R.M. Haralick, S.R. Sternberg, X. Zhuang, Image analysis using mathematical morphology, *IEEE Trans. Pattern Anal. Mach. Intell.* 9 (4) (July 1987) 532–550.
- [8] Y. Iwasaki, An image processing system to measure vehicular queues and an adaptive traffic signal control by using the information of the queues, *IEEE International Conference on Intelligent Transportation System, ITSC 97*, pp. 195–200.

- [9] S.C. Pei, C.L. Lai, F.Y. Shih, An efficient class of alternating sequential filters in morphology, *Graphical Models Image Process.* 59 (2) (March 1997) 109–116.
- [10] S.C. Pei, C.L. Lai, F.Y. Shih, A morphological approach to shortest path planning for rotating objects, *Pattern Recognition* 31 (8) (August 1998) 1127–1138.
- [11] S.C. Pei, L.G. Liou, Finding the motion, position and orientation of a planar patch in 3D space from scaled-orthographic projection, *Pattern Recognition* 27 (1) (January 1994) 9–25.
- [12] J.C. Rojas, J.D. Crisman, Vehicle detection in color images, *IEEE International Conference on Intelligent Transportation System, ITSC 97*, pp. 403–408.
- [13] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, New York, 1982.
- [14] S.R. Sternberg, Grayscale morphology, *Comput. Vision Graphics Image Process.* 35 (September 1986) 333–355.