Joint Template Matching Algorithm for Associated Multi-object Detection

Jianbin Xie[1], Tong Liu[1], Zhangyong Chen[2] and Zhaowen Zhuang[1]
[1]Department of Electronic Science and Engineering, National University of Defense Technology
Changsha, China
[2]ZHONGCHAO Enterprise CO, LTD.
Beijing, China
[e-mail: jbxie@126.com, liutong1129@126.com]
*Corresponding author: Tong Liu

Abstract

A joint template matching algorithm is proposed in this paper to reduce the high rate of miss-detection and false-alarm caused by the traditional template matching algorithm during the process of multi-object detection. The proposed algorithm can reduce the influence on each object by matching all objects together according to the correlation information among different objects. Moreover, the rate of miss-detection and false-alarm in the process of single-template matching is also reduced based on the algorithm. In this paper, firstly, joint template is created from the information of relative positions among different objects. Then, matching criterion according to normalized cross correlation is generated for multi-object matching. Finally, the proposed algorithm is applied to the detection of watermarks in bill. The experiments show that the proposed algorithm has lower miss-detection and false-alarm rate comparing to the traditional NCC algorithm during the process of multi-object detection.

---

1. Introduction

Object detection attracts lots of interests in the field of image analysis, recognition and tracking. Generally, the common algorithm for object detection is template matching (TM) [1][2]. The TM algorithm mainly includes two parts: one based on features [3][4][5][6][7] and the other based on normalized cross correlation (NCC) [8][9][10][11][12]. The former one requires that the features of objects are invariable to the position, scale and rotation. However extracting these features is not an easy task in many situations. The latter one computes the similarity between template and unknown objects according to the cross correlation, which is widely used because of its robustness under the brightness variations [13]. However, some drawbacks of NCC algorithm are obvious when detecting multi-objects. Firstly, only one object can be detected because we can obtain only one maximum value of NCC in an image; so the rate of miss-detection is high. Secondly, the phenomenon of false detection may occur when the images of objects are influenced by the factors such as occlusion, distortion, variation of light and similar objects [14][15].

If the relative positions among different objects are known, and the set of position distribution image is numerable, this kind of multi-object detection is called associated multi-object detection. In order to reduce the rate of miss-detection and false-alarm for associated multi-object detection, we propose a joint template matching (JTM) algorithm. Firstly, joint template (JT) is constructed based on the information of relative positions among different objects. Then, matching criterion according to NCC is constructed for multi-object matching. Finally, the proposed algorithm is applied to watermarks detection of bill, and good results are obtained.

The rest of this paper is organized as follows. In Section 2, the proposed algorithm is described. Results for watermarks detection are presented in Section 3, together with a comparison between the proposed algorithm and traditional NCC algorithm. Finally, conclusions are given in Section 4.

2. The Proposed Algorithm

2.1 Construction of JT

For associated multi-object detection, we can list all position distributions of the objects appearing in the image based on the information of relative positions among different objects. The set of position distribution is called JT, which can represent all position distributions of unknown objects which may appear in the image.

There are two basic assumptions to construct JT:

(1) The set of position distribution of multi-object is numerable;

(2) The relative position between two objects is known or measurable.

Let P be the number of members in the set of position distribution, and P is known according to condition (1), then JT = $\{JT_k | 1 \leq k \leq P\}$, where $JT_k = \{T_k^l | 1 \leq l \leq Q\}$, T_k^l is the lth template in the kth member, Q represents the total number of templates in the kth member. The relationship

among different templates is determined by condition (2). Let (x_k^1, y_k^1) be the coordinate of T_k^1 in the image, then the coordinate of T_k^l is $(x_k^1+dx_k^l, y_k^1+dy_k^l)$. According to condition (2), dx_k^l and dy_k^l is known or measurable.

2.2 Joint Template Matching

The basic idea of JTM is to find similar multi-object with JT. We execute JTM by using NCC between templates and objects. The steps are as follows:

Step 1: Calculate the value of NCC between image and the kth member of JT.

Let (m, n) be the coordinates of T_k^1 in the image, then the value of NCC between the image and the kth member of JT is:

$$R_k(m,n) = \frac{1}{Q}\sum_{l=1}^{Q} R_k^{max}(m,n,l) \tag{1}$$

Where,

$$R_k^{max}(m,n,l) =$$

$$\max_{-D_x \le x \le D_x, -D_y \le y \le D_y} \left(\frac{\sum_{j=1}^{N}\sum_{i=1}^{M} f(i+m+dx_k^l+x, j+n+dy_k^l+y) \cdot T_k^l(i,j)}{\sqrt{\sum_{j=1}^{N}\sum_{i=1}^{M} f(i+m+dx_k^l+x, j+n+dy_k^l+y)^2}\sqrt{\sum_{j=1}^{N}\sum_{i=1}^{M} T_k^l(i,j)^2}} \right) \tag{2}$$

Where, $f(i, j)$ is the image under examination, $T_k^l(i, j)$ is the lth template in the kth member, of size $M\times N$ located at pixel coordinates $(m+dx_k^l+x, n+dy_k^l+y)$. Because image distortion and calculation error always exist, there is some deviation between the information of relative position in multi-template and the one in multi-object. To find the best matching position of a single template T_k^l, we need to compute the maximum value of NCC in a neighborhood region (NR) of the template. Therefore, $R_k^{max}(m, n, l)$ represents the maximum value of NCC when point (x, y) moves from one pixel to another in NR, with the size of $(2\times D_x+1)\times(2\times D_y+1)$.

Step 2: Finding the best matching position of the kth member in the template.

Let S be the region to which Tk1 can move in the image. When (m, n) is moving in S, we find the position where $R_k(m, n)$ is maximum, which is the best matching position of the kth member in the template. The maximum value of $R_k(m, n)$ is R_k:

$$R_k = \max_{(m,n)\in S} R_k(m,n) \tag{3}$$

Step 3: Finding the best matching position of JT.

After finding best matching positions of every member in JT, we find the position where R_k is maximum, which is the best matching position of JT:

$$R = \max_{0 \leq k \leq P} R_k$$

(4)

Assuming that R achieves maximum value at member JT_k, meanwhile (m, n) moves to (m_k, n_k), we can then detect all unknown multi-object, which are corresponding to the templates in JT_k. Then their positions in the image can be obtained by adding position (m_k, n_k) to the positions of the corresponding templates in JT_k.

Overall, the basic principle of JTM is to match multi-object according to the relative position among different objects, to reduce the rate of miss-detection and false-alarm for single-template matching. JTM algorithm can be used in the field of associated multi-object detection such as watermarks detection of bill and component detection of circuit board.

3. Watermarks Detection of Bill Based on JTM

3.1 Construction of JT for watermarks

In the processing of bill production, only one kind of watermark board is used to produce watermarks in bills. Therefore, the number of watermarks' types is constant, and the relative position among different watermarks is constant. A bill image is shown in **Fig. 1**, and all eight types of watermarks are shown in **Fig. 2**. The relative position among different watermarks in the image is known, and the set of position distribution of watermarks in different bills is numerable. So we can construct JT for watermarks as described in 2.1.

Let W_0 be the watermark at top-left corner of the image, then W_0 can determine the position distributions of watermarks in different bills. When W_0 varies from W_1 to W_8, there will be 8 sorts of different position distributions, so P=8. If we take no account of the incomplete watermarks on the boundaries of the image, there are 18 watermarks in each bill, so Q=18. Therefore, JT for watermarks can be described as: JT=$\{JT_k | 1 \leq k \leq 8\}$, JT_k=$\{T_k^l | 1 \leq l \leq 18\}$.

Components of JT_k are shown in **Fig. 3**, where T_k^1 is a watermark of W_1 type. W_0=W_1, and its coordinate is (x_k^1, y_k^1), dx and dy are the distance between two neighboring watermarks at horizontal and vertical direction, respectively. Thus, the coordinate of T_k^l is $(x_k^1+l_1 \times dx, y_k^1+(l-6 \times l_1) \times dy)$, where l_1=floor($l/6$), and the floor() function returns the value of a number rounded DOWNWARDS to the nearest integer.
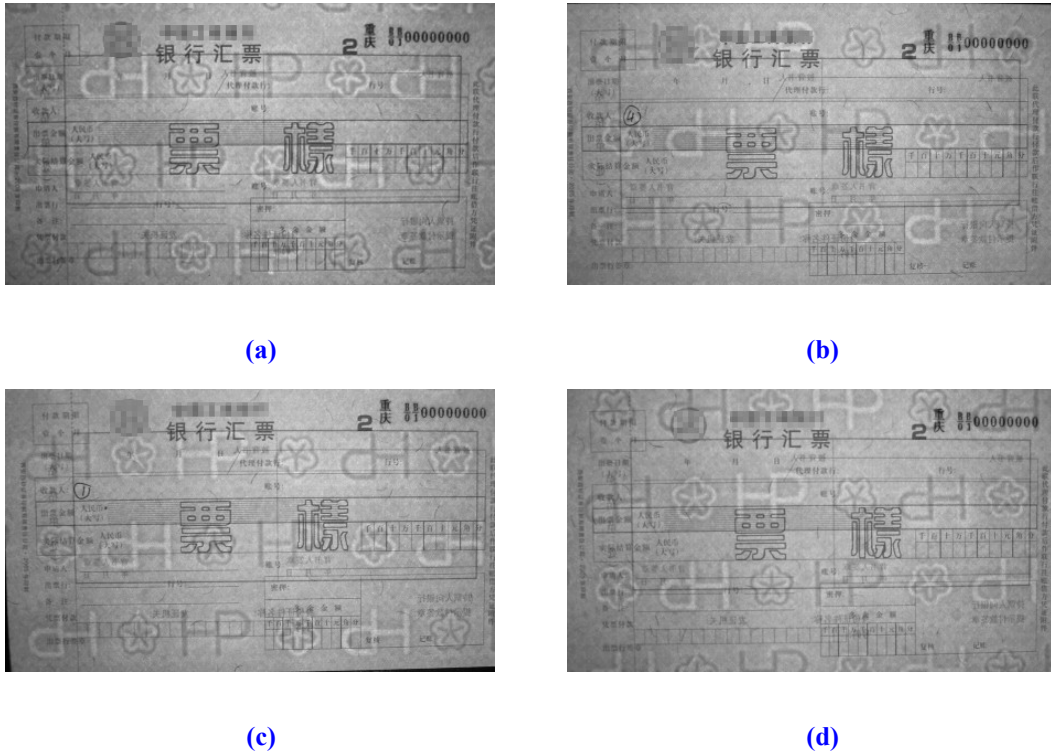
(a)   **(b)**

(c)   **(d)**

Fig. 1. Bill images

W_1 W_2 W_3 W_4

W_5 W_6 W_7 W_8

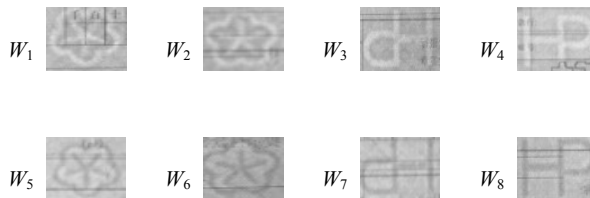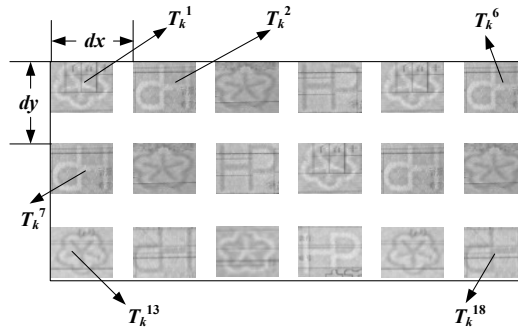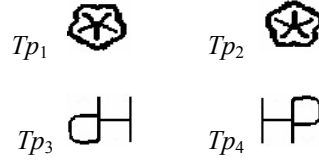**Fig. 2.** Watermarks (W_1-W_4: white watermarks, W_5-W_8: black watermarks)

Fig. 3. Components of JT_k

Fig. 4. watermark templates

3.2 Image Preprocessing

There are two tasks need to achieve in the stage of image preprocessing:

(1) Image cutting and resizing is executed to obtain normalized bill image shown in **Fig. 1**;

(2) Two-level wavelet transform is executed to obtain the low-frequency image.

The contribution of wavelet transform has two aspects: (1) to reduce computation cost of following process by reducing image size while reserve most information of the image; (2) to enhance robustness by reducing the influence of light variation and noise in the image. In this paper, Mallat wavelet is used to achieve fast wavelet transform.

3.3 Image Binarization

For different images of bills, the degrees of occlusion to watermarks from lines, texts and frames are different. Therefore, it's difficult to obtain good results by using TM algorithm based on gray features. Instead, shape features are more important. In this paper, we execute JTM algorithm based on binary image, to avoid the influence of above factors by ignoring the distribution characters of gray level.

The distinction between watermarks and background is unapparent, so we obtain binary image by using local-threshold segmentation method. Two functions F_1 and F_2 are constructed in this paper. F_1 is used to describe weighted gradient of Laplacian. The weight value is the inverse of distance, so the closer pixels from center make greater contribution to gradient; F_2 is used to describe the difference between the gray value of one pixel and the average value of its neighborhood (SR), of size $(2 \times w + 1) \times (2 \times h + 1)$ pixels. Then, F_1 and F_2 can be described as:

$$F_1(i,j) = r \cdot \sum_{y=-h, x^2+y^2 \neq 0}^{h} \sum_{x=-w}^{w} \frac{[f(i,j) - f(i+x, j+y)]}{\sqrt{x^2 + y^2}} \tag{5}$$

$$F_2(i,j) = r \cdot [f(i,j) - \frac{1}{(2w+1) \cdot (2h+1)} \sum_{y=-h}^{h} \sum_{x=-w}^{w} f(i+x, j+y)] \tag{6}$$

Where, r is used to describe watermark's attributes of black or white. If watermark is black, $r=1$, and the binary image is denoted by Img_B. Otherwise, $r=-1$, and the binary image is denoted by Img_W. Let $g(i,j)$ be gray value of pixel (i,j) in the binary image, then:

$$g(i,j) = \begin{cases} 1 & , F_1(i,j) < 0 \quad and \quad F_2(i,j) < 0 \\ 0 & , otherwise \end{cases} \tag{7}$$

There are many noises in the binary image, which may interfere with the process of object detection. Therefore, mathematical morphology method is executed to remove small noise and amalgamate neighboring pixels.

3.4 Joint Template Matching

The basic unit of JTM is to compute the value of NCC between a single template in JT and an unknown object in source image. Considering the differences of black watermarks and white ones, we need to construct the corresponding relationship among JT, binary template and image according to different values of l.

The corresponding binary templates of **Fig. 2** are shown in **Fig. 4**, of size $M \times N$ pixels. According to **Fig. 2**, **Fig. 3** and **Fig. 4**, the template-image index table of the kth member in JT is shown in **Table 1**, and the ones of the other 7 members can be constructed in the same way. Then, $R_k^{max}(m, n, l)$ is obtained by putting $T_k^l(i,j)$ and corresponding $g(i,j)$ into formula (2).

Table 1. The template-image index table

l	1	2	3	4	5	6
$T_k^l(i,j)$	Tp_1	Tp_3	Tp_2	Tp_4	Tp_1	Tp_3
$g(i,j)$	Img_W	Img_W	Img_B	Img_B	Img_W	Img_W
l	7	8	9	10	11	12
$T_k^l(i,j)$	Tp_3	Tp_2	Tp_4	Tp_1	Tp_3	Tp_2
$g(i,j)$	Img_W	Img_B	Img_B	Img_W	Img_W	Img_B
l	13	14	15	16	17	18
$T_k^l(i,j)$	Tp_1	Tp_3	Tp_2	Tp_4	Tp_1	Tp_3
$g(i,j)$	Img_B	Img_B	Img_W	Img_W	Img_B	Img_B

According to formulas (1), (3) and (4), we can compute R after obtaining $R_k^{max}(m, n, l)$. Where, the region S to which T_k^1 can move in the image can be determined based on the distribution of watermarks in bill:

$$S = \{(m,n) \mid 0 \le m \le dx, 0 \le n \le dy\} \tag{8}$$

After obtaining R with formula (4), the member JT_k corresponding to R can be regarded as the matching member of JT. Then the types and positions of unknown multi-object in the image can be determined by the templates in JT_k, as described in 2.2.

Considering that both source image and template image are binary image, and the pixels' values of background and object can be denoted by 0 and 1, respectively, so we can rewrite $R_k^{max}(m, n, l)$ as formula (9), which can reduce computational cost by avoiding a lot of multiplication operators.

$$R_k^{max}(m,n,l) = \max_{-D_x \leq x \leq D_x, -D_y \leq y \leq D_y} \left(\frac{N_{g\_T}}{\sqrt{N_{g\_g}}\sqrt{N_{T\_T}}} \right)$$

(9)

Where,

$$N_{g\_T} = \sum_{j=1}^{N}\sum_{i=1}^{M} g[i+m+l_1 \times dx + x, j+n+(l-6 \times l_1) \times dy + y] \cdot T_k^l(i,j)$$

(10)

$$= \sum_{j=1}^{N}\sum_{i=1}^{M} g[i+m+l_1 \times dx + x, j+n+(l-6 \times l_1) \times dy + y] \& T_k^l(i,j)$$

$$N_{g\_g} = \sum_{j=1}^{N}\sum_{i=1}^{M} g[i+m+l_1 \times dx + x, j+n+(l-6 \times l_1) \times dy + y]^2$$

(11)

$$= \sum_{j=1}^{N}\sum_{i=1}^{M} g[i+m+l_1 \times dx + x, j+n+(l-6 \times l_1) \times dy + y]$$

$$N_{T\_T} = \sum_{j=1}^{N}\sum_{i=1}^{M} T_k^l(i,j)^2 = \sum_{j=1}^{N}\sum_{i=1}^{M} T_k^l(i,j)$$

(12)

Where, symbol "&" represses the AND operator.

3.5 Experiment and Analysis

Experimental environment: Randomly selecting 500 pieces of bills, we capture 1500 images using 3 devices, and then sort them into 3 groups according to different device number. The size of original image, normalized image and that after wavelet transformation is 720×576, 444×256 and 111×64 (defined as *W×H*), respectively. The size of watermark template is 15×11 (defined as *M×N*). The distance parameter (*dx*, *dy*) between two neighboring watermarks is set to (17, 20).

Experimental methods: Firstly, binary templates of watermark are selected as shown in **Fig. 4**; secondly, binary images are obtained as described in 3.3; finally, two algorithms of NCC [1] and JTM are used to detect all 18 watermark objects in bills. The size of SR in the stage of image binarization is set to 5×5. And the size of NR of JTM is set to 15×11 (D_x=7, D_y=5). The results are shown in **Table 2**.

Experiment analysis:

(1) *Analysis of object detection performance*: It can be seen from **Table 2** that, the rate of miss-detection by NCC algorithm is very high, because it can only detect one object each time according to the maximum value of NCC. For JTM algorithm, all position distributions of the objects appearing in the image have been listed in JT, so the rate of miss-detection is zero. Meanwhile, because of the similarity among different types of watermarks, and the occlusion to watermarks from lines, texts and frames in image, the matching position with the maximum value of NCC may not be the best one. Therefore, the rate of false-alarm by NCC algorithm is

high. On the other hand, JTM algorithm uses the associated information among different watermarks to achieve corporate matching for all objects, and finds the best distribution from all possible position distributions in JT, which leads to great false-alarm rate reduction. In addition, the NCC algorithm is sensitive to interferences such as different occlusion of different types of watermarks, different gray distribution of watermarks due to the light's position and intensity, camera's position and other structures of different devices. But JTM algorithm reduces the influence on each object by matching all objects at one time, so it is more ~~the~~ robust than the NCC algorithm.

   (2) *Analysis of algorithm efficiency*: Most of the computation cost of JTM algorithm is to calculate the value of NCC. Let t_0 be the computation cost of NCC calculation for one time. In the process of watermarks detection, the computation cost of NCC algorithm is $t_1=(W-M)\times(H-N)\times t_0$, the one to one template matching in JTM algorithm is $t_2=(2\times D_x+1)\times(2\times D_y+1)\times t_0$, and the one to all templates matching in JTM algorithm is $t_3=P\times Q\times dx\times dy\times t_2$. In our experimental environment, $t_1=5088\times t_0$, $t_2=165\times t_0$ and $t_3=8078400\times t_0$. As a whole, the computation cost of JTM algorithm is about 1588 times higher than that of NCC algorithm. But JTM algorithm can detect 18 objects, so the average computation cost to detect one object is $(448800\times t_0)$, which is about 88 times higher than that of NCC algorithm. As the computation process of each member in JTM is independent, parallel processing can be used to improve the efficiency of JTM algorithm.

   (3) *Analysis of memory cost*: Compare to NCC algorithm, JTM algorithm needs more memory to store multi-object template and the template-image index table. The memory size is $4\times M\times N+P\times Q\times 2=948$ bytes in our experimental environment.

Table 2. The results of watermarks detection

Algorithms		NCC								JTM
Types of watermarks		W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	$W_1 \sim W_8$
The rate of false-alarm	Group 1	13.2%	13.8%	7.2%	6.8%	12.8%	14.2%	7.4%	7.8%	0.006%
	Group 2	12.8%	13.8%	6.4%	6.8%	13.4%	14.2%	7.2%	7.6%	0.004%
	Group 3	13.2%	13.2%	6.0%	7.0%	13.2%	14.2%	7.2%	7.8%	0.006%
The rate of miss-detection		94.4%	94.4%	94.4%	94.4%	94.4%	94.4%	94.4%	94.4%	0.0%

4. Conclusions

A novel JTM algorithm is proposed in this paper, which can reduce the rate of miss-detection and false-alarm for associated multi-object detection. The new algorithm can be used in the

field of associated multi-object detection such as watermarks detection of bill and component detection of circuit board. Practically, the matching criterion of NCC can be displaced by feature matching according to the specific demands. Drawback is the efficiency of new algorithm. Future work will focus on eliminating this drawback.

References

[1]   K. P. William, "Digital image processing," *of John Wiley &Sons*, 2001. Article (CrossRef Link)

[2]   B. Perret, S. Lefevre and C. Collet, "A robust hit-or-miss transform for template matching applied to very noisy astronomical images," *Journal of Pattern Recognition*, vol. 42, no. 11, pp. 2470-2480, Feb. 2009. Article (CrossRef Link)

[3]   Y. H. Lin and C. H. Chen, "Template matching using the parametric template vector with translation, rotation and scale invariance," *Journal of Pattern Recognition*, vol. 41, no. 7, pp. 2413-2421, Jan. 2008. Article (CrossRef Link)

[4]   F. Ullah and S. Kaneko, "Using orientation codes for rotation-invariant template matching," *Journal of Pattern Recognition*, vol. 37, no. 8, pp. 201-209, Jan. 2004. Article (CrossRef Link)

[5]   H. Y. Kim, "Rotation-discriminating template matching based on Fourier coefficients of radial projections with robustness to scaling and partial occlusion," *Journal of Pattern Recognition*, vol. 43, no. 3, pp. 859-872, Aug. 2010. Article (CrossRef Link)

[6]   F. Essannouni, and D. Aboutajdine, "Fast frequency template matching using higher order statistics," *IEEE Transactions on Image Processing*, vol. 19, no. 3, pp. 826-830, Mar. 2010. Article (CrossRef Link)

[7]   S. Mattoccia, F. Tombari and L. Di Stefano, "Efficient template matching for multi-channel images," *Journal of Pattern Recognition Letters*, vol. 32, no. 5, pp. 694-700, Dec. 2011. Article (CrossRef Link)

[8]   L. Di Stefano, S. Mattoccia and F. Tombari, "ZNCC-based template matching using bounded partial correlation," *Journal of Pattern Recognition Letters*, vol. 26, no. 14, pp. 2129-2134, May. 2005. Article (CrossRef Link)

[9]   S. Mattoccia, F. Tombari and L. Di Stefano, "Fast full-search equivalent template matching by enhanced bounded correlation," *IEEE Transactions on Image Processing*, vol. 17, no. 4, pp. 528-538, Apr. 2008. Article (CrossRef Link)

[10] H. Li, H. B. Duan and X. Y. Zhang, "A novel image template matching based on particle filtering optimization," *Journal of Pattern Recognition Letters*, vol. 31, no. 13, pp. 1825-1832, Dec. 2010. Article (CrossRef Link)

[11] S. D. Wei and S. H. Lai, "Fast template matching based on normalized cross correlation With adaptive multilevel winner update," *Journal of IEEE Transactions on Image Processing*, vol. 17, no. 11, pp. 2227-2235, Nov. 2008. Article (CrossRef Link)

[12] H. B. Duan, C. F. Xu, S. Q. Liu et al., "Template matching using chaotic imperialist competitive algorithm," *Journal of Pattern Recognition Letters*, vol. 31, no. 13, pp. 1868-1875, Dec. 2010. Article (CrossRef Link)

[13] J. C. Yoo and T. H. Han, "Logic operation-based template matching algorithm for one-dimensional signals," *Institution of Engineering and Technology(IET) Signal Processing*, vol. 5, no. 2, pp. 261-269, Apr. 2011. Article (CrossRef Link)

[14] C. H. Wu, D. Z. Wang, A. Ip et al., "A particle swarm optimization approach for components placement inspection on printed circuit boards," *Journal of Intelligent Manufacturing*, vol. 20, no. 5, pp. 535-549, Jun. 2009. Article (CrossRef Link)

[15] [15] A. J. Crispin and V. Rankov, "Automated inspection of PCB components using a genetic algorithm template-matching approach," *International Journal of Advanced Manufacturing Technology*, vol. 35, pp. 293-300, Oct. 2007. Article (CrossRef Link)

Jianbin Xie received his B.S., M.S. and Ph.D. degrees from the National University of Defense Technology in 1993, 1995 and 1999. He is currently working as a research fellow at National University of Defense Technology. His main research interests include biometric identification, pattern recognition and machine learning.

Tong Liu received his B.S. and M.S. degrees from the National University of Defense Technology in 2004 and 2006. He is currently working toward a Ph.D. degree at National University of Defense Technology. His current research interests include biometric identification, pattern recognition and machine learning.

Zhangyong Chen received his B.S. and M.S. degrees from the National University of Defense Technology in 1993 and 1995. His current research interests include biometric identification, pattern recognition and machine learning.

Zhaowen Zhuang received his B.S. and M.S. degrees from the National University of Defense Technology in 1981 and 1984, and received his Ph.D. degrees from Beijing Institute of Technology in 1989. He is currently working as a professor in the National University of Defense Technology. His current research interests include target recognition and satellite navigation.