Image Matching for Unspecified Target Objects Using Geometric Edge Features

Ryusei Shibakawa[1], Fumihiko Saitoh[1]

Z4525043@edu.gifu-u.ac.jp
[1] Gifu University Graduate School 1-1 Yanagito, Gifu-shi, Gifu-ken 501-1112 Japan
Keywords: block matching, indefinite target object, geometric edge feature, rotation and scale change

ABSTRACT

In this paper, we propose a method for detecting the same object that is mixed in two images. In this method, image matching is performed using block matching [1] and geometric edge features.

1 Introduction

In this paper, we propose image matching for the purpose that the user can match without having to prepare a template image that only a specific object is photographed.

1.1 Difference from Normal Template Matching

In normal template matching, there are methods such as normalized cross-correlation [2] and incremental code correlation [3], but in both cases, prepare a template image that shows only the object that the user wants to search, find the pixel features of the entire image, and find the target image. And match. However, the problem this time is that we do not know where in the image the object we want to search is in. Therefore, by dividing the image that is the search source into blocks, an image close to a normal template image is created. The object shown in the image created here may be rotated or have a different scale than the object shown in the target image. Therefore, matching is performed using the geometric edge feature amount that can correspond to the rotation of the object and the voting method that can correspond to the scale change.

1.2 Process of Image Processing

Here, the process of image processing will be described.

(1) Divide the image that is the search source into blocks

(2) Obtain the edge strength using the Sobel filter [4].

(3) Make the edge image thin and make it a line image.

(4) Acquire the edge features of all edge pixels of the line image.

(5) As shown in Fig. 2, the relative relationship between the edge strength and the edge angle of the surrounding pixels as seen from the pixel P is obtained.

(6) The target image is also processed from (2) to (5).

(7) Calculate the similarity between the line image of the block-divided search source image and all the pixels of the line image of the target image, and if they are similar, vote for the coordinates of the pixels of the target image.

(8) Judge that the target object exists at the coordinates with the most votes.

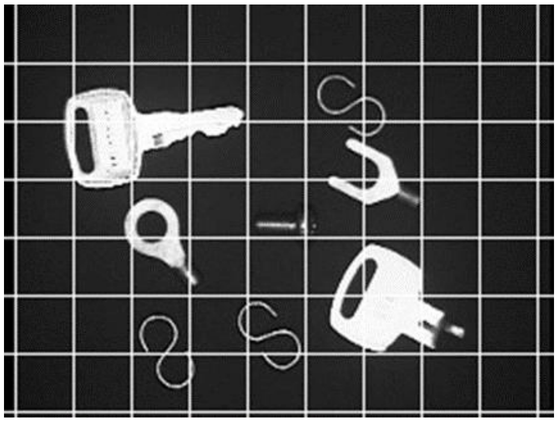

Fig. 1 Block image of original image

2 Proposed Method

Here, the proposed method of this paper will be explained.

2.1 Block Division of the Image to be searched

In this paper, regarding the block division method, blocks are divided into squares as shown in Fig. 1 regardless of the aspect ratio of the original image. This is to prevent the advantages and disadvantages of matching from occurring depending on the shape and rotation of the target object. In addition, to prevent the important features of the object from being damaged by straddling the boundary line of the division during division, the coordinates of the boundary line are shifted by 1/2 of the width of the block image in the x and y directions. Is also prepared, and matching is performed with block images for a total of two original images.

2.2 Edge Strength Measurement

Calculate the edge strength in the X and Y directions using a general Sobel filter and create an edge image.

2.3 Thinning of Edge Image

Create a line image by thinning the created edge image.

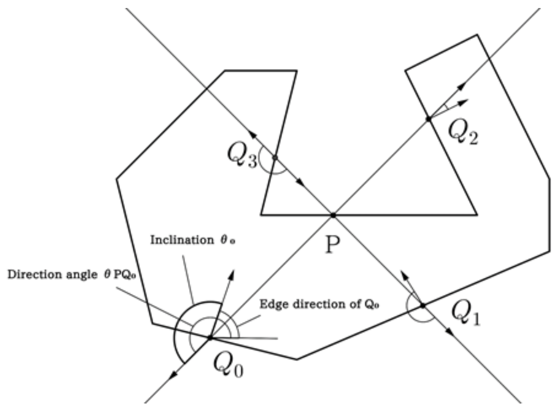

Fig. 2 Relative edge features

2.4 Get Edge Strength and Edge Angle

Get the edge strength and edge angle for every edge pixel in the line image. As shown in FIG. 2, the edge strength eP and the edge angle θ_{eP} of the pixel of interest P are obtained from the following equation using the edge strengths in the X and Y directions obtained by the Sobel filter.

$$eP = \sqrt{eP_x{}^2 + eP_y{}^2} \tag{1}$$

$$\theta_{eP} = tan^{-1}\frac{eP_x}{eP_y} \tag{2}$$

2.5 Relative Relationship

As shown in Fig. 2, 360 degrees is divided into n (4 divisions in the case of an image) based on the edge angle θ_{eP} of a certain pixel of interest P obtained in 2.4, and the destination in that direction is searched. Use the coordinates of pixel P (x_p, y_p) and the coordinates of pixel Q_n (x_{Q_n}, y_{Q_n}) for the direction angle θ_{PQ_n} with the pixel group Q_n closest to pixel P. To find it. By finding the edge direction of pixel Q_n in the same way as in Equation 2 and taking the difference from the direction angle θ_{PQ_n} seen from pixel P, the relative edge direction of the pixel seen from the pixel can be found. This is shown in the following equation.

$$\theta_n = \theta_{PQ_n} - \theta_{eQ_n} \tag{3}$$

Also, in the same way as the edge angle, the absolute value e_n of the difference is obtained to obtain the relative edge strength of the pixel Q_n as seen from the pixel P.

$$e_n = |eP - eQ_n| \tag{4}$$

2.6 Do the Same for the Target Image

The processing from 2.2 to 2.4 is also performed on the target image, and the relative relationship between the block-divided original image and the surrounding pixels as seen from all the edge pixels of the target image is obtained.

2.7 Comparison

The block-divided pixel P of the original image and the pixel P of the target image are compared in all combinations.

When comparing the attention pixel P_I of the target image and the attention pixel P_T of the block image, the relative edge angle θ_{nI} and edge strength e_{nI} of Q_0 to Q_n that the pixel P_I has as information and the relative edge angle θ_{nI} and the edge strength e_nI of the pixel P_T. The similarity is judged by the difference between the typical edge angle θ_{nT} and the edge strength e_{nT}. The formula for calculating this difference is shown below.

$$\mathrm{err}\theta_n = |\theta_{nI} - \theta_{nT}| \tag{5}$$

$$\mathrm{err}e_n = |e_{nI} - e_{nT}| \tag{6}$$

If this difference $\mathrm{err}\theta_n$ and $\mathrm{err}e_n$ is smaller than the predetermined threshold value, the number of features is counted as $m(0 \leqq m \leqq n)$. The value v obtained by dividing this number m by the number n of the compared features is voted in the voting space as the similarity of the pixels of interest P_I and P_T

$$\mathrm{v} = \mathrm{m/n} \tag{7}$$

All block images accumulate votes in one common voting space. At this time, the block image that becomes noise unrelated to the target object also votes in the same voting space, but the voting is dispersed in the voting space for the noise, whereas the block image of the target object concentrates on the similarities. I think that it will be possible to make a distinction because the votes will be targeted. The coordinates of the target object are the points where the most votes are collected after all the block images have been voted. The coordinates to vote (t_x, t_y) are calculated using the following formula

$$t_x = x + r * \cos\left(\theta_{lT} + (\theta_{eI} - \theta_{eT})\right) \tag{8}$$

$$t_y = y + r * \sin\left(\theta_{lT} + (\theta_{eI} - \theta_{eT})\right) \tag{9}$$

Currently, the coordinates of the edge pixel P_I of the target image as the pixel of interest are x and y, the distance of the edge pixel P_T of the block image from the center of the block image is r, and the direction angle from the center of the block image is θ_{lT}. In addition, $\theta_{eI} - \theta_{eT}$ represents the difference between the edge direction of P_I and the edge direction of P_T, and the rotation angle of the target object in the target image can be obtained. Therefore, the coordinates to vote (t_x, t_y) calculated by using equations (8) and (9) coincide with the center of the block image, and it is considered that the voting is concentrated on one point of the target object.

Here, the scale invariance voting is applied. Vote by changing equations (8) and (9) as follows.

$$t_x = x + s * r * \cos\left(\theta_{lT} + (\theta_{eI} - \theta_{eT})\right) \tag{10}$$

$$t_y = y + s * r * \sin\left(\theta_{lT} + (\theta_{eI} - \theta_{eT})\right) \tag{11}$$

s represents the magnification of the scale change of the target object. By voting while changing this value such as s = 0.80, 0.81, ..., 1.19, 1.20, voting corresponding to scale changes is performed. With this voting method, voting is done linearly instead of points, but by voting linearly from multiple pixels of interest, the number of votes is concentrated on the coordinates where the lines overlap

2.8 Judgement

Finally, in the voting space where all votes have been completed, it is judged that the target object is at the coordinates with the highest number of votes.

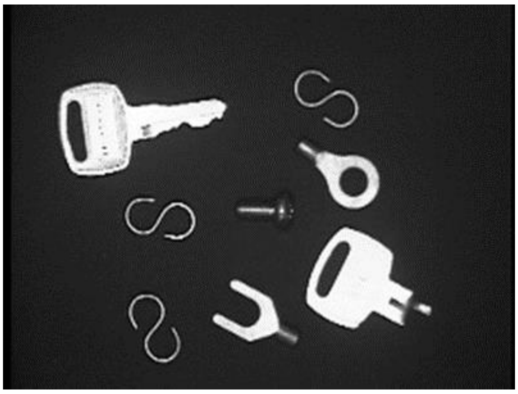

Fig. 3 (a) Original image 1

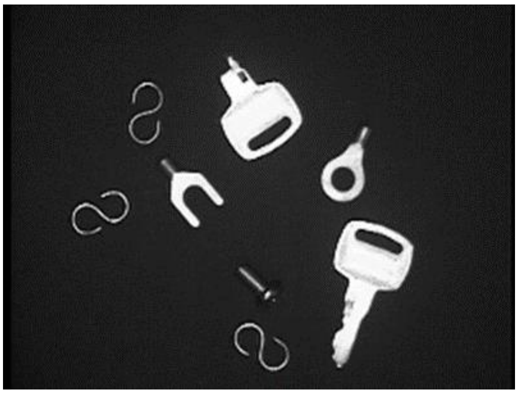

Fig. 3 (b) Original image 2

2.9 Experimental conditions

The target image and the original image had a size of 256 × 192 pixels, and a shade image of 256 gradations was used. For the original images, we used three types with a scale of 1x as shown in Fig. 3 (a) and three types with a scale of 0.85 times as shown in Fig. 3 (b). For the target images, we used three types of images as shown in Fig. 4, and a total of 108 images taken by rotating them by 10 degrees. The number of blocks in the original image was 7 × 9 as shown in Fig. 1.

The target image 1 shown in Fig. 4 (a) is an image in which an object whose characteristics are not very similar to the target object is added as noise. The target images 2 and 3 shown in Fig. 4 (b) and (c) are images in which a circular object similar to the characteristics of the target object is added as noise. Also, as shown in Fig. 4, the radius of the white circle drawn as the coordinates of the recognition result is the same as the width of the block image. This is set as the estimated size of the target object because the size of the target object is unspecified in the proposed method. The correctness of the recognition result was visually judged whether or not the coordinates of the center of the circle output as a result indicate the feature points of the target object.

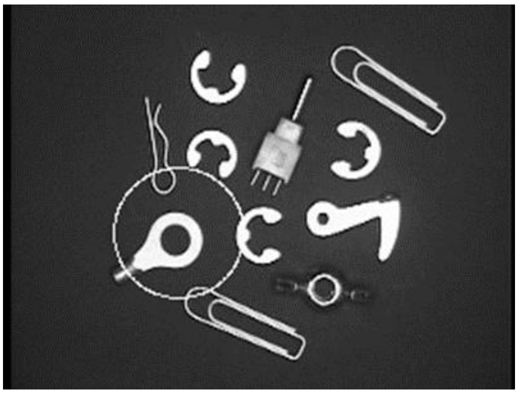

Fig. 4 (a) Target image 1

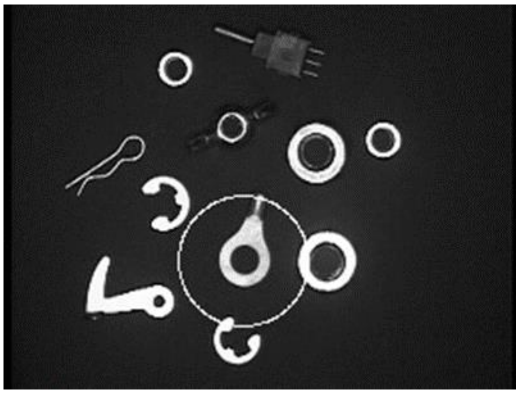

Fig. 4 (b) Target image 2

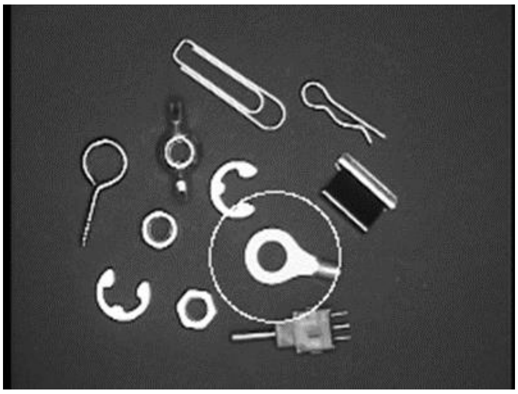

Fig. 4 (c) Target image 3

3 Results

The experimental results and their consideration in this study are shown.

3.1 Results and Consideration

Table 1 shows the recognition rate of the target object of the proposed method. The target image 1 had a recognition rate of 95% or more even when the target object of the original image was reduced, and good results were obtained. I think this is because there is no other object that has the same characteristics as the circular characteristics of the target object. In the target image 2, a circle having a size like the circle of the target object exists as noise, and the recognition rate is lower than that of the target image 1. Even so, the recognition rate is 90% or higher, which means that the matching method of the proposed method is useful. The recognition rate of target image 3 dropped significantly to 78.7% by reducing the original image. It is considered that this is because the similarity to the nut added as noise has increased by reducing the target object of the original image.

Figure 5 shows an example of recognition failure. Most of the images that failed to be recognized made a false judgment for an object with a characteristic of a circle similar to the target object.

Table 1 recognition rate(%)

	Target 1	Target 2	Target 3
Original image 1	100	95．1	91．3
Original image 2	95．4	94．4	78．7

4 Conclusion

The conclusions of this study are shown.

In this paper, we proposed a method to find an object from two images showing a common object without an ideal template image. He divided the original image into blocks to create a pseudo template image and used a voting method corresponding to geometric edge features and scale changes to respond to the rotation and scale changes of the object. As a result, the recognition rate decreased when the scale of the object changed significantly, but the overall recognition rate was good.

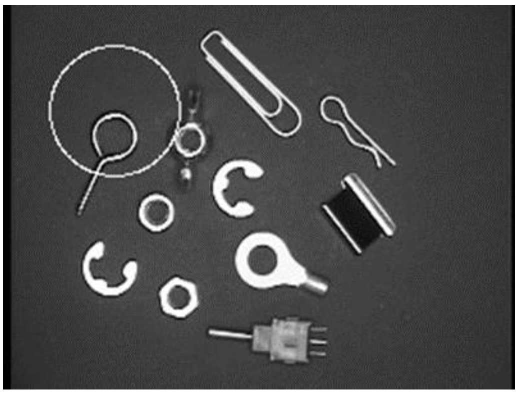

Fig. 5 (a) Failed image 1

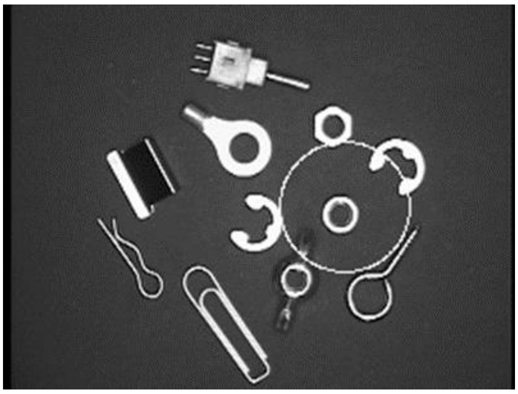

Fig. 5 (b) Failed image 2

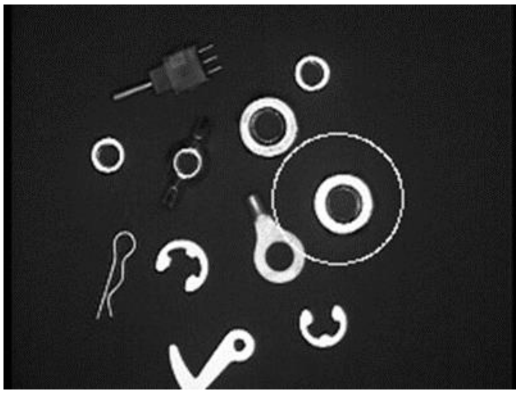

Fig. 5 (c) Failed image 3

References

[1] Fumihiko Saitoh, “Image Matching that is Strong Against Shielding Using Block Matching Voting Processing,” Journal of the Electronic Information Society, D-2, Information / System 2-Pattern processing, pp.2270-2279(2000).

[2] Koji Ikeda, Shoji Yoshida, Keisuke Nakajima, Nagaharu Hamada and Haruo Inui, “High-speed Template Matching by Monotonic Function of Normalization Correlation Operation,” Journal of the Institute of Electronics, Information and Communication Engineers, D-11, Vol.J83-D-11, No9, pp.1861-1869(2000).

[3] Shunichi Kaneko, Ichiro Murase and Satoru Igarashi, “Robust Image Registration by Increment Sign Correlation,” Transactions on Pattern Recognition (PR), Vol.35, No.10, pp2223-2234(2002).

[4] O. R. Vincent and O. Folorunso, ”A Descriptive Algorithm for Sobel Image Edge Detection,” Proceedings of Informing Science & IT Education Conference, InSITE09, pp.97-107(2009).