Robust Occlusion Recognition Based on Ring Projection Transforms

Yun-su Choi[1], Jin-hoon Park[1], Ghulam Hussain[1], Jong-hak Kim[2], Jin-won Park[1] and Jun-dong Cho[13]
[1] Department of Electrical and Computer Engineering, Sungkyunkwan University
Cheoncheon-Dong, Jangan-Gu, Suwon 300, Korea
[2] Department of IT convergence, Sungkyunkwan University
Cheoncheon-Dong, Jangan-Gu, Suwon 300, Korea
[3] Distinguished Visiting Professor, North University of China, Taiyuan, Shanxi, China
ys.choi9343@gmail.com, jhakkim@vada.skku.ac.kr, jdcho@skku.edu

Abstract: In this paper, we propose a method for locating occluded object and measuring its angle using Ring Projection Transform (RPT) template matching. The current template matching method does not accurately recognize an object that is occluded in a scene image because correlations between the reference template and scene image decrease with occlusion. However, the proposed method divides the template image into sectors to allow partial recognition for locating an occluded object. The angle of the object can also be calculated using the center point of the partially recognized area.

1. Introduction

Object recognition is an important topic in computer vision. Typical techniques for locating an object include feature point extraction and the template matching method which measure the correlation between an object template and a scene image. The representative method for feature point extraction is known as the Scale Invariant Feature Transform (SIFT) [2]. The SIFT is effective in scale rotation. There are various methods of template matching. The Sum of Absolute Differences (SAD) [3] and Normalized Cross Correlation (NCC) [3] are representative methods of template matching. These methods can perform matching without extracting feature points, but they are vulnerable to rotation or scale change.

The manufacturing industry continues to increase the use of automation systems in the product inspection, assembly, and manufacturing process. In such automation environments, a vision algorithm[4][5][6] is essential for object location and recognition. A recognition algorithm that is strong in rotation and occlusion rather than scale and brightness is needed in automation environments at factories because camera views and lighting are fixed and partially limited in these circumstances. Although there has been research on methods for recognizing objects that are not rotatable, like Ring Projection Transform (RPT) [1], such algorithms are weak at locating occluded objects. Therefore, in this paper, we propose a method which is strong in locating occluded objects, regardless of rotation or occlusion, by using template image matching based on RPT and dividing the template image for recognition.

2. Related work

In vision algorithms, template matching is typically used to locate the template image in the scene image. RPT is a type of template matching which is strong in rotation processing. RPT uses projective transformation with a ring, and can be defined as shown in Figure 1 and Equation (1).

$$\vec{T}(r) = \frac{1}{2\pi r}\int_0^{2\pi} I\left(\frac{w}{2} + r\cos\theta, \frac{h}{2} + r\sin\theta\right) d\theta \quad (0 < r \le R) \quad (1)$$

In Equation (1), $\vec{T}(r)$ is the sum of all pixel values located in the circumference of each circle which is gradually increasing the radius r from the center of the template image. Parameter $I(x, y)$ is the pixel value that corresponds with (x, y) in the image. Parameters w and h are the width and height of the template image, respectively, and R is the smaller value of radius $w/2$ and $h/2$.

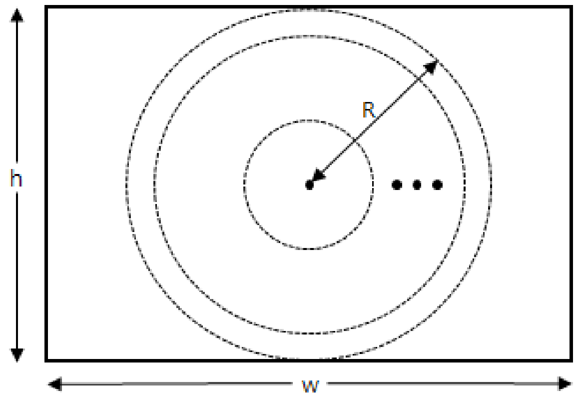

Figure 1. Template image projected to ring with size ($\boldsymbol{w}$ x $\boldsymbol{h}$).

After calculating the T vector, it needs to be compared to all points of the scene image. Figure 2 shows the matching process. The template image is moved, by pixel unit, in the scene image, and at the location of each template image point, the similarity with the scene image is calculated according to Equations (2) and (3).

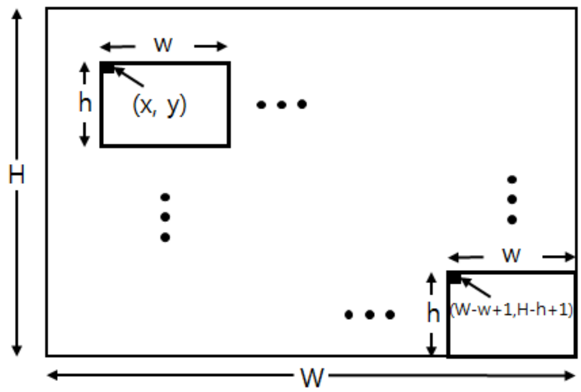

Figure 2. Matching process on the scene image with size ($\boldsymbol{W}$ x $\boldsymbol{H}$) by template with size ($\boldsymbol{w}$ x $\boldsymbol{h}$).

$$O(x,y) = NCC(\vec{T},\vec{S}_{(x,y)}) = \frac{Co-variance(\vec{T},\vec{S}_{(x,y)})}{(R-1)\sigma(\vec{T})\sigma(\vec{S}_{(x,y)})} \quad (2)$$

$$\vec{S}_{(x,y)}(r) = \frac{1}{2\pi r}\int_0^{2\pi} I'\left(x+\frac{w}{2}+r\cos\theta, y+\frac{h}{2}+r\sin\theta\right)d\theta$$
$$\begin{pmatrix} 0 \le x \le W-w+1 \\ 0 \le y \le H-h+1 \\ 0 < r \le R \end{pmatrix} \quad (3)$$

where $O(x,y)$ is the value calculated from scene image (x,y), and NCC refers to the normalized correlation coefficient. Parameter $\vec{S}_{(x,y)}$ is the sum of all pixels in each of the circle circumferences created by gradually increasing the radius r to maximum R. Parameter σ(x) refers to the standard deviation of X. After analyzing all points, the location with the largest NCC value refers to the location of the object. Since template matching calculates the correlation rate between the scene image and template image, accurate recognition is difficult when an object is partially occluded due to the reduction in correlation.

3. Proposed Method

In our proposed method, the template image is divided into four sectors: A, B, C, and D. In the center of each divided image, Equation (1) is applied to calculate the vector value. Here, the maximum radius R is the smaller value of w/4 and h/4. Using Equation (4), a total of six distances are calculated from the center points of the four sectors of the divided template image. In Equation (4), (x_1, y_1) and(x_2, y_2) is the center location of the divided template.

$$\bar{D} = \sqrt{(x_1-x_2)^2 + (y_1-y_2)^2} \quad (4)$$

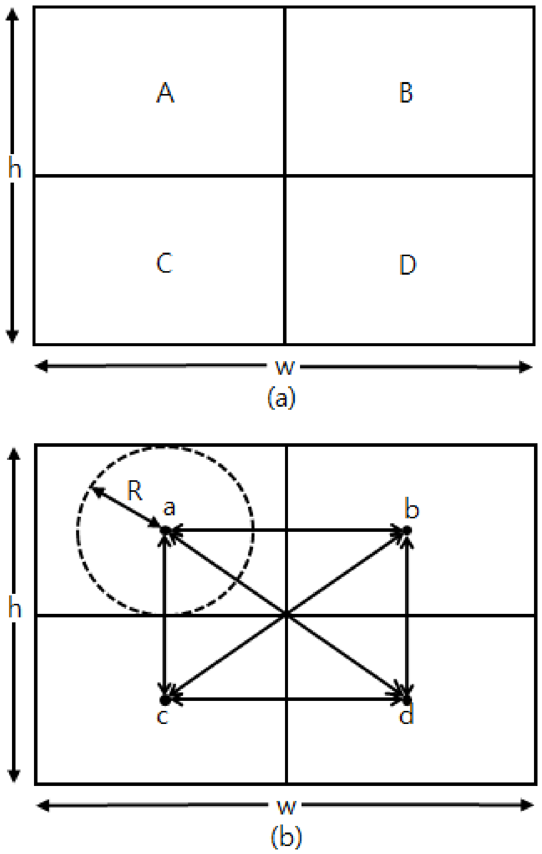

Figure 3. (a) Divided template image (b) Distance from each center.

For each of the sectors in the divided template image, Equations (2) and (3) are applied to calculate correlation coefficients with the scene image and determine the largest coefficient. In the scene image, the number of divided sectors differ according to how much the object image is occluded. When at least two images are found, Equation (4) is used in the scene image to find the distance. The accuracy of the object location can be increased by comparing this value with the value calculated from the template image.

When at least two divided images are found, the rotation angle of the entire template image can be calculated. The segment of the line calculated from the scene image is moved parallel to the point of origin, and the segment of the line that matches the template image is also moved to the point of origin. The angle between the two segments is then calculated.

4. Experimental Result

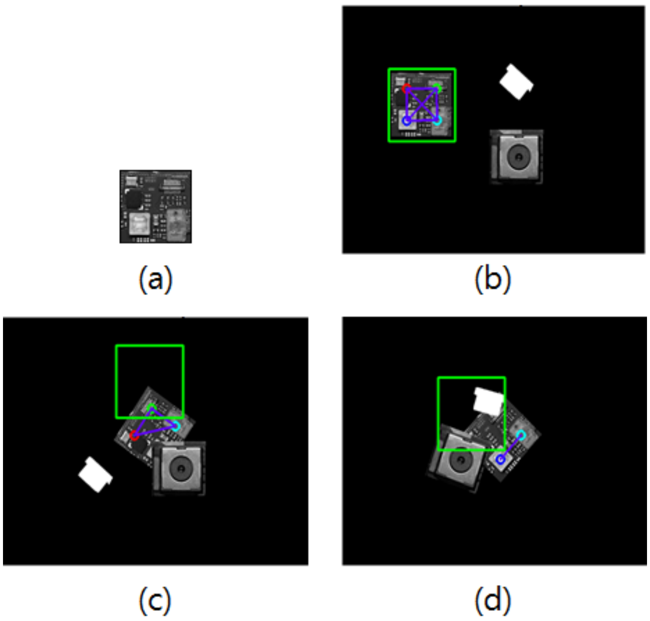

Figure 4. (a) Template image (b) Unoccluded Object (c) Occluded Object #1 (d) Occluded Object #2.

Table 1. Processing Time Comparison

Image	Existing method	Proposed method
B	2884.297[ms]	3605.621[ms]
C	2875.767[ms]	3586.238[ms]
D	2846.792[ms]	3568.233[ms]

The experiments are conducted on a PC with running Windows7 64bit, 16GM RAM, Intel Core i5-4670 3.4GHz processor. The test image set contains scene images with size 640x480 and template image with size 128x123.

Figure 4 shows comparison results between the existing and proposed method. The existing method identifies the area shown in the green square. The proposed method divided image and identified the area shown in the circle. In Figure 4(b), with an object that is not occluded, both the existing method and proposed method locate the object. In Figure 4(c) and (d), where the object is occluded, the existing method does not correctly locate the object while the proposed method locates the divided part of the object. Table 1 shows the processing times for the existing method and proposed method in the different scenarios. The processing time duration for the proposed method is an average of 20% higher than the existing method due to the increased time required for image matching.

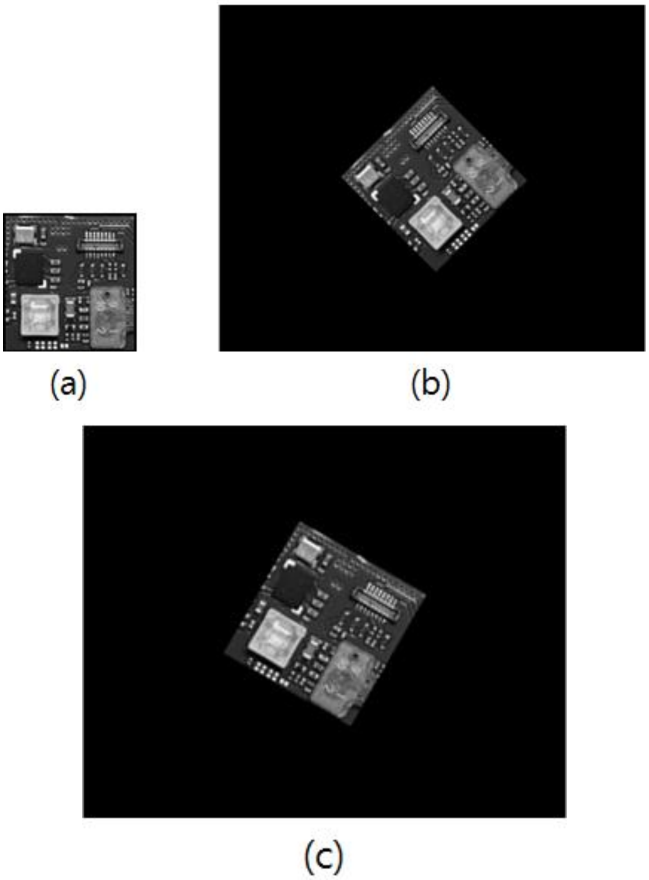

Figure 5 (a) Template image (b) Rotation image #1 (c) Rotation image #2.

Table 2. Measurement of Angle Values

Image	Real Rotation Angle	Measurement Value
B	45.0°	45.33 °
C	330.0 °	330.77 °

Figure 5 and Table 2 show the results of calculating the angle of the object using the proposed algorithm. The error rate was less than 1% of the actual measured angle using the proposed algorithm.

5. Conclusion

In this paper, we proposed a method for enhancing object location accuracy using the existing RPT method with divided images that partially locate the object using the distances from the partially located center point. Experimental results indicated a strong location rate when the object was occluded compared to the existing RPT method. The rotated angle of the object was also calculated, with an error rate of less than 1%, using the partially located center point. The proposed algorithm can be applied to automation systems in the manufacturing process where objects can be occluded

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2013R1A1A2058942)

This work was supported by the Technological Innovation R&D Program funded by the Small and Medium Business Administration(SMBA, Korea)" [S2343584 , Development of automated manufacturing robot system technology integrating with the 6 DOF robot mechanism and the S/W platform for assembling mobile IT products]

Reference

[1] H. Y. Kim, and S. A. De Araujo, “Grayscale template matching invariant to rotation, scale, translation, brightness and contrast,” Advances in Image and Video Technology. Springer Berlin Heidelberg, pp. 100-113, 2007.

[2]D. G. Lowe, "Distinctive image features from scale-invariant keypoints," IJCV, 2004.

[3] R. Gonzales and R. Woods, Digital Image Processing. Pearson Prentice Hall, third ed., 2008.

[4] Sang-hyeob Song, Seong-muk Kang, Hyeong-jun Cho and Jun-dong Cho “Ring Projection Transforms by Using CUDA Implementation for Recognizing Objects in Automation System” *International SoC Design Conference*, 2014

[5] Seung-ho Kim, Sang-hyeob Song, Jong-hak kim, Zhongyun Yuan and Jun-dong Cho “Fast Rotation-Invariant Template Matching with Candidate Reduction Using CUDA” *International Symposium on Consumer Electronics*, 2015

[6] SangKwon Sim, Jong-hak Kim, Zhongyun Yuan, Seong-muk kang and Jun-dong Cho “Robust-Rotation Recognition Based on Contour Matching Using CUDA in Automation System” *International Symposium on Consumer Electronics*, 2015

Yun-su Choi received the B.S degree Electronic Engineering from Suwon University, Suwon, Korea, in 2014. He is a M.S candidate in the Electrical and Computer Engineering at Sungkyunkwan University, Suwon, Korea. His current research interests include Object Recognition, GPGPU Computing

Jin-hoon Park received the B.S degree Electronic Engineering from Suwon University, Suwon, Korea, in 2015. He is a M.S candidate in the Electrical and Computer Engineering at Sungkyunkwan University, Suwon, Korea. His current research interests include 3D Registration, 3D visual fatigue reduction and parallel architecture.

Chulam Hussain is combined (M.S & PhD) student in electronic & electrical engineering department, Sungkyunkwan University South Korea. He has been researching in VLSI Algorithm Design Automation Lab of electronic & electrical department since April 2015. He received the Bachelor degree in Electronic Engineering from the Mehran University of Engineering and Technology, Jamshoro, Parkistan (2012). His current research interests include the Artificial Neural Networks, Biomedical applications and 3D technology, with emphasis on System.

Jong-hak Kim received the B.S. degree in radio communication engineering from the Kyunghee University, Yongin, Korea, in 2009, the M.S. degree from the Department of Electrical and Computer Engineering, Sungkyunkwan University, in 2012, and he is studying for a Ph. D degree at Sungkyunkwan University.
He is interested in the efficient low power hardware implementation for a real-time image processing system in mobile equipment. He currently studies a visual fatigue reduction scheme for a stereo vision system.

Jin-won Park received the B.S degree Electronic Engineering from Wonkwang University, Iksan, Korea, in 2016. He is a M.S candidate in the Electrical and Computer Engineering at Sungkyunkwan University, Suwon, Korea. He is current research interests include low power hardware implementation for a real-time image processing system.

Jun-Dong Cho was born in Seoul, Korea on 1957. He received B.S degree in Electronic Engineering, Sungkyunkwan Univ. in Seoul, Korea, 1980, M.S degree from Polytechnic University, Brooklyn, NY, 1989, and Ph.D. degree from Northwestern University, Evanston, IL, 1993, both in computer science. He was a Senior CAD Engineer at Samsung Electronics, Co., Ltd. He is now Professor of Dept. of Electronic Engineering, Sungkyunkwan University, Korea. He received the Best paper award at the 1993 Design Automation Conference. He has been an IEEE Senior Member since April 1996. He established an "H-Lab." and department of Human ICT Convergence at Sungkyunkwan University, Sep. 2013. He has been published more than 260 papers, 30 patents, and 5 books. He received the Best Convergence Research Award, Ministry of Science, ICT, and Future Planning, Oct. 2014, and Semiconductor Research Contribution Award from Ministry of Knowledge and Economics, Oct. 2012, respectively. His research topics include Wearable /Healthcare Device UI/UX with Activity-driven Cognition.