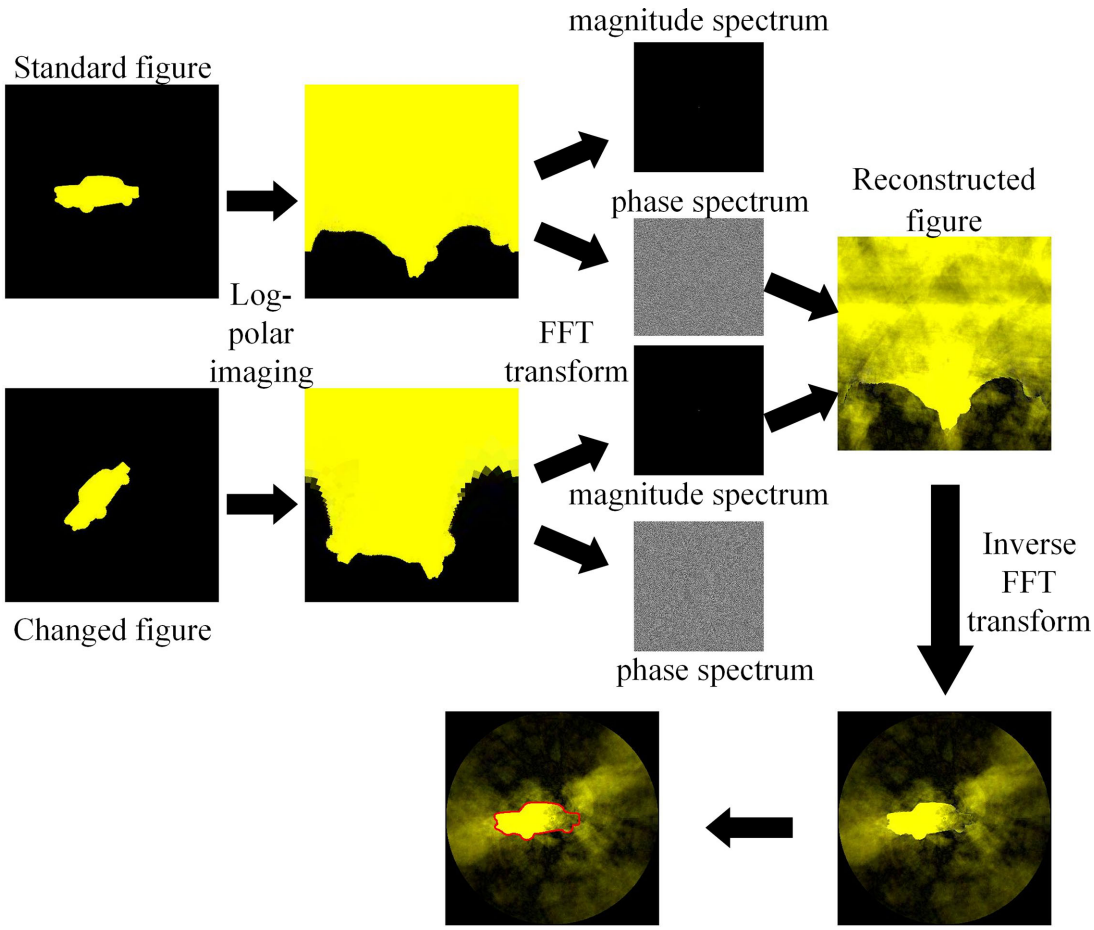
IEEE Photonics Journal
An IEEE Photonics Society Publication

Open Access

Study on Rotation and Scaling Invariance of Retina-Like Imaging

Volume 12, Number 5, October 2020

Qun Hao
Mingyuan Tang
Jie Cao
Saad Rizvi
Huan Cui
Zhimin Wu



DOI: 10.1109/JPHOT.2020.3025872

Study on Rotation and Scaling Invariance of Retina-Like Imaging

Qun Hao,[1] **Mingyuan Tang**,[1] **Jie Cao**,[1] **Saad Rizvi**,[1] **Huan Cui**,[1] **and Zhimin Wu**[2]

[1]Key Laboratory of Biomimetic Robots and Systems, Ministry of Education, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China
[2]School of Mechanical and Electrical Engineering, Shenzhen Polytechnic, Shenzhen 518055, China

DOI:10.1109/JPHOT.2020.3025872

Manuscript received August 17, 2020; revised September 18, 2020; accepted September 19, 2020. Date of publication September 22, 2020; date of current version October 2, 2020. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grants 61871031 and 61875012, in part by Beijing Natural Science Foundation under Grant 4182058, in part the funding of foundation enhancement Program 2019-JCJQ-JJ-273, and in part by the Science and Technology Program of Shenzhen of China under Grant JCYJ20170818114408837. Corresponding authors: Jie Cao; Zhimin Wu (e-mail: ajieanyyn@163.com; zhimin_wu@szpt.edu.cn.)

Abstract: Inspired by human vision, retina-like imaging has the feature of rotation and scaling invariance, which is beneficial for target tracking and recognition. However, there is a limitation that the center of field of view (FOV) should match with the target centroid. Currently, there is no effective method to quantitatively estimate the effects from this mismatch. To solve the issue, a novel evaluation method is proposed by applying the log-polar transform and Fourier transform. Then reconstruct the target image by exchanging magnitude and phase information between the target image and its rotated or scaled counterpart. Using two current methods of digital image processing, scale-invariant feature transform (SIFT) and biaxial projection similarity analysis (BPS), as evaluation criteria, the proposed method performs efficiently for all eccentric angles. Simulation results show that for a specific target, the invariance of rotation and scaling works effectively for a value above 0.69.

Index Terms: Image analysis, retina-like imaging, rotation and scaling invariance.

1. Introduction

Machine vision is widely used in many applications, including robotics, navigation, and biology [1]. To capture target information efficiently, it is important to have a large field of view (FOV) and high resolution of the captured target [2]. Recently, different techniques like sparse reconstruction [3], image fusion [4] and image stitching [5] have been proposed for image processing. However, among the above-mentioned schemes, achieving large FOV, high resolution and imaging efficiency is still very challenging. In other words, to improve imaging performance, it is important to strike a balance between the FOV and image resolution. For example, to maintain high resolution density in a large FOV, more complex hardware setup is required such as multi-camera array, curved image sensor or freeform optics [6], [7].

Inspired by the space-variant sampling structure of human eye, retina-like imaging provides a novel approach to solve the issues above [8]–[10]. The retina-like method reconstructs an

image with high-resolution at the center and low resolution in its periphery. Therefore, unlike conventional methods, the retina-like imaging achieves a large FOV with high resolution while suppressing redundant information in the background. In addition, this imaging method, due to its characteristic of rotation and scaling invariance, is used for different image processing applications such as target and text recognition, baseline matching, and image stitching [11]–[13].

However, if there is an eccentricity between the center of FOV and target centroid, retina-like imaging may be distorted because of the distorted target [14], [15]. Although the problem of eccentricity can be solved by adjusting the fixation point to target centroid, an additional device is needed, e.g. Omni-directional rotating platform. However, it is still very difficult to match these two points precisely, which effect a lot to target recognition and localization tasks. Therefore, the question that arises here is: how precise the match between the center of FOV and target centroid should be to allow normal operation of retina-like imaging? The answer to this question is not found in the existing literature. To the best of our knowledge, current studies focus on the application of retina-like imaging and most metrics used in retina-like imaging are concentrated around the imaging quality effected by noisy [16], [17]. However, analysis for the boundary conditions of eccentricity is neglected, let alone an effective method to estimate the applicability of rotation and scaling invariance.

Therefore, in this paper we study and propose a novel method (which can act as an evaluation metric) to evaluate the rotation and scaling invariance. For the target with rotation, scaling and possible eccentricity, Fourier transform (FT) is used on the log-polar transformed figure. To quantify the invariance in retina-like imaging, we reconstruct the target image by exchanging the magnitude spectrum and phase spectrum (obtained with Fourier transform) between the original target and its rotated counterpart, which we call it LPT-MP. We imply this method on arbitrary eccentricity, and get the trend of the ability of rotation and scaling invariance. The estimation for this process is used with scale-invariant feature transform (SIFT) and biaxial projection similarity analysis (BPS) to determine the reference LPT-MP value and further divide the tolerable range. The retina-like imaging can keep the ability of rotation and scaling invariance when the eccentricity is limited in this specific range. This tolerance for misalignment bring great convenience to practical application of retina-like imaging.

2. Method

2.1 Related Work for Rotation and Scaling Invariance

In order to solve the problems of target's rotation and scale changes, convolutional neural networks (CNNs), directional feature extraction and coordinate transform are widely used. Deep learning methods based on CNNs like rotation-invariant CNN [18] and rotation-invariant and Fisher discriminative CNN [19] is a novel technology. By implying these methods under popular object detection frameworks, these algorithms can get good recognition capabilities for rotated targets. Directional feature extraction is the main solution to the rotation change of the widely used local feature descriptor like SIFT [26] and speeded up robust features [20]. According to the local gradient direction of the image, each key point is given a reference direction. All the subsequent operations are based on the directional feature when it comes to rotation and scaling, thereby providing the invariance of rotation and scaling. Radon transform [21] and log-polar transformation (LPT) [22] are two typical rotating and scaling invariance transforms. Radon transformation is an integral transformation. It takes the image along any line with a line integral. By rotating the line with different angles, the target's rotation is transform into shift which makes it invariance to rotation. On the other hand, scaling change of the target effects only the absolute value of the integration, while the relative proportion remains the same. LPT is one of the theoretical basis of retina-like imaging which also have the ability of rotation and scaling invariance, and will be detailed introduced in the Section 2.2.

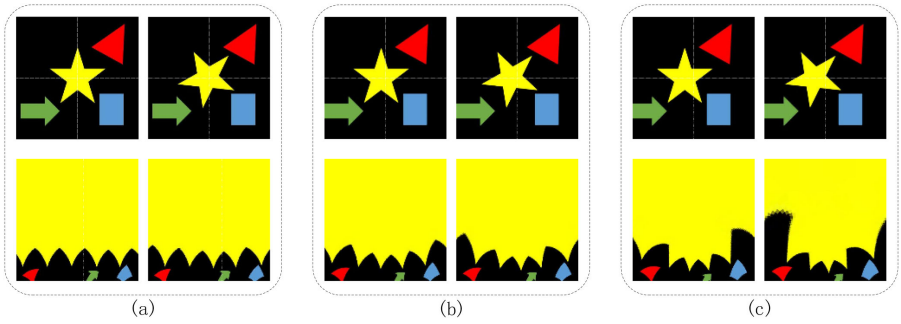


Fig. 1. Rotation invariance with different eccentricity. (a) No eccentricity. (b) 15 pixels. (c) 30 pixels.

2.2 Rotation and Scaling Invariance in Retina-Like Imaging

The kernel of retina-like imaging includes two aspects, namely, space-variant sampling, and LPT, written as

$$\begin{cases} \theta = \arctan \frac{y}{x} \\ \xi = \log_z(\sqrt{x^2 + y^2}) \end{cases}, \tag{1}$$

where z is the increasing coefficient, (x, y) are the Cartesian coordinates, and (ξ, θ) are the log-polar coordinates. Because of the existence of the parameter z, the base number of the logarithms, the resolution density becomes sparse in the periphery. Thus, the retina-like structure allows efficient suppression of redundant data.

In contrast, the rotation and scaling of the target in the Cartesian coordinate system only affect the axes θ and ξ, respectively, as represented by Eq. 2.

$$\begin{cases} \theta' = \theta + \alpha \\ \xi' = \log(\sqrt{(qx)^2 + (qy)^2}) = \log(qr) \end{cases}$$

$$= \log(q) + \log(r) = \xi + \log(q), \tag{2}$$

where α stands for the rotation, and q stands for the scale change. θ and ξ stands for the two axis of LPT, respectively. In this way, the relationship of rotation and scaling is rewritten by the shift in θ and ξ directions, which is called the rotation and scale change invariance.

Fig. 1 shows the rotation invariance of retina-like imaging. When the target (yellow star) is matched with the FOV, as shown in Fig. 1(a), the rotation invariance performs as desired. In this case, the rotation of the target is transformed into a horizontal shift. The next two parts show the distortion of the invariance. The target shape will distort in an unusual manner when some eccentricity existed, making it difficult to apply retina-like imaging for further applications. To make matters worse, the distortion become more severe with the eccentricity increases. In addition, the same problem also occurs in scaling invariance of retina-like imaging, as shown in Fig. 2.

This distortion of rotation and scaling invariance caused by eccentricity in retina-like imaging is a feature of LPT. Although this one-to-one mapping in the spatial domain can bring us the advantage of rotation and scale invariance directly in the imaging process, the existence of the central blind hole (polar coordinate origin) also limits the target's position. Therefore, the study of the tolerable range of the match between the center of field of view and the target centroid, as studied in this paper, is significant for the development of retina-like imaging.

2.3 Intensity Feature Based on the Rotation and Scaling Invariance

As shown in Eq. 2, by mapping the target to log-polar coordinates, the rotation and scaling of the target translates into lateral shifts without any change in its shape. Here, the Fourier transform

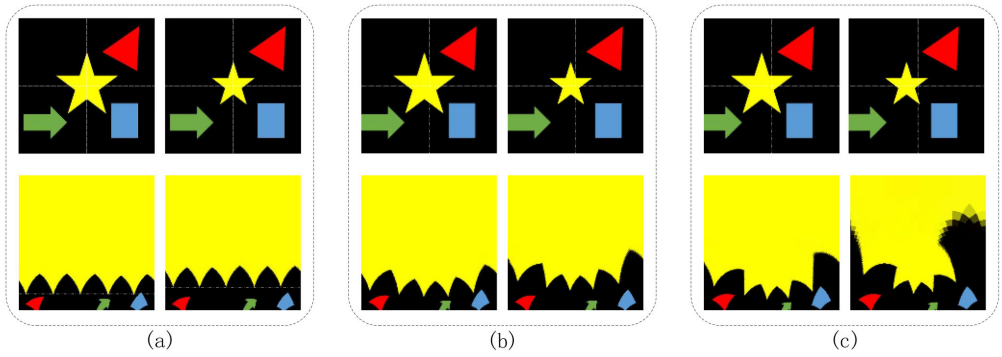


Fig. 2. Scaling invariance with different eccentricity. (a) No eccentricity. (b) 15 pixels. (c) 30 pixels.

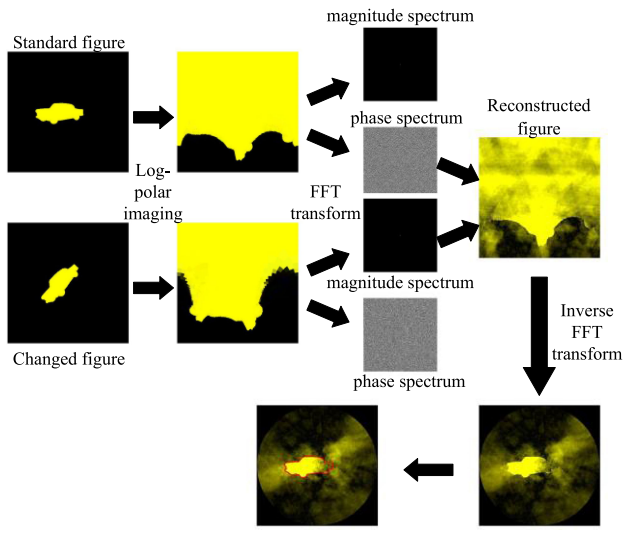


Fig. 3. The process of the LPT-MP.

can be applied to the LPT image, resulting in two further representations of magnitude spectrum for energy information, and phase spectrum for position information [23]–[25]. It is observed that the rotation and scaling change only affects the phase spectrum of the LPT, while the magnitude spectrum stays the same. Therefore, by utilizing only the magnitude spectrum of the rotated or scaled target image, and combining it with the phase spectrum of the original target image, the effects of shift changes in the log-polar mapping can be corrected. However, if the target is eccentric, the rotation or scaling will bring some changes to the shape in LPT. Therefore, not only the phase spectrum, but also the magnitude spectrum for an eccentric target image will be different. The difference makes the pixel intensity distribution of the reconstructed figure beyond the limited area. In this way, by calculating the intensity distribution, the evaluation of rotation and scaling invariance can be obtained.

Fig. 3 shows the process of the LPT-MP. The method employs retina-like imaging to get corresponding images for both the standard target and the changed (e.g., rotated) target. Next, Fourier transform is applied to the two transformed images to get their corresponding magnitude and phase spectrum information. This information is further used to reconstruct an image with the help of inverse Fourier transform (IFT) by using the phase spectrum of the standard target image and the magnitude spectrum of the rotated target image. Finally, the reconstructed log-polar image is transformed back to Cartesian coordinate representation via the inverse log-polar transform. The

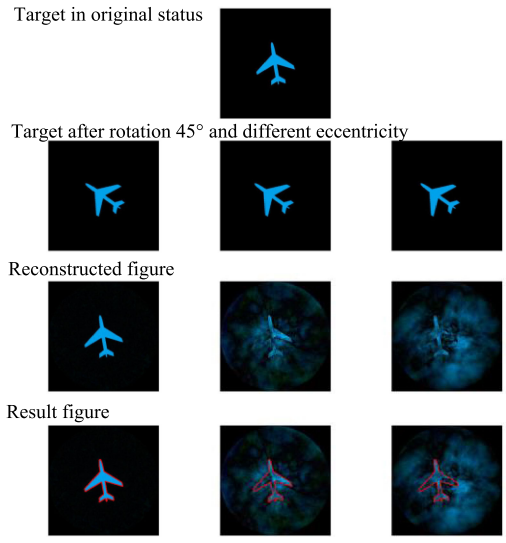


Fig. 4. Examples of LPT-MP.

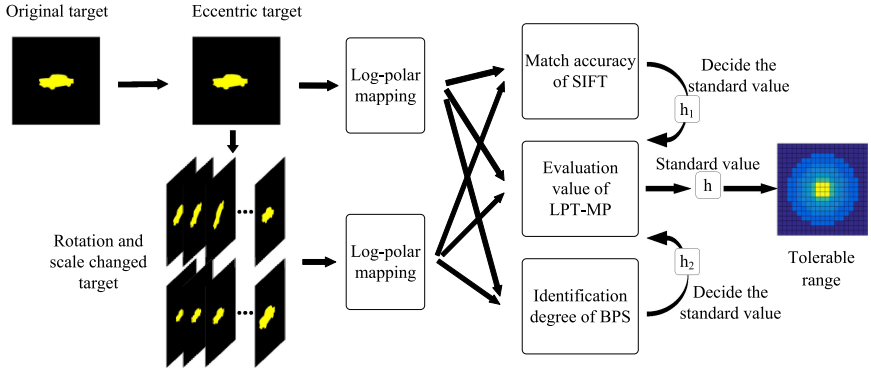


Fig. 5. Process of the simulation.

intensity distribution of the resulting image is then calculated with Eq. 3 to get the evaluation value *h*.

$$\begin{cases} I_j(x, y) = \begin{cases} 1, I(x, y) \in \text{interested target} \\ 0, esle \end{cases} \\ h = \frac{\sum\sum I_{re}(x,y) \otimes I_j(x,y)}{\sum\sum I_{re}(x,y)} \end{cases}, \tag{3}$$

where $I_j(x, y)$ is a judgement of the original figure $I(x, y)$ about whether this point (x, y) belongs to the area of interest, and $I_{re}(x, y)$ is the resulting image reconstructed with the inverse log-polar transform in final step. The value h is the ratio of the energy distributed in the interested area and the total energy distributed of the whole figure. So this value can represent the ability of rotation and scaling invariance evaluated by LPT-MP.

Fig. 4 shows some examples of the LPT-MP. The first row of the figure is the target in original status. The three changed targets in the second row are rotated 45° with the eccentricity (shifted left) of 0 pixel, 10 pixels and 30 pixels, respectively. The third row shows the reconstructed figures obtained as a result of combining the phase spectrum of the standard target and the magnitude spectrum of the changed target (via subsequent FT, IFT, and Inverse LPT operations). The last row of the figure shows the result of the LPT-MP. From the results of Fig. 5, it can be seen clearly that

without any eccentricity the processed figure gathered most intensity at the initial position. But with the increase of the eccentricity, the intensity distribution becomes more chaotic.

3. Simulation and Result

3.1 Simulation Step

We propose the evaluation method LPT-MP is aimed at expanding the scope of application of retina-like imaging in machine vision. So the threshold of the proposed method should be restricted by the algorithms of target recognition or image stitching. In this way, to verify the performance of the LPT-MP, we establish the relationship between the proposed method two existing methods, scale-invariant feature transform (SIFT) [26], and biaxial projection similarity (BPS) analysis algorithm [27].

SIFT is a common methods used in local feature point matching. By implying the directional feature, it is invariance to rotation and is one of the basis method of image stitching. So SIFT can be used to provide a reference value for feature matching results. By considering both the precision and recall of the matching result, as represented by Eq. 4, we can evaluate whether the reduction of rotation and scale invariance is significant in feature point matching or not, with the value h_1.

$$\begin{cases} \text{precision} = \frac{\text{number of correct matches}}{\text{number of matches}} \\ recall = \frac{\text{number of correct matches}}{\text{number of feature points}} \\ h_1 = 1.25 \times \frac{\text{precision} \times \text{recall}}{0.25 \times \text{precision} + \text{recall}} \end{cases}. \tag{4}$$

BPS is a fast target identification algorithm which gives good performance in terms of recognition speed and accuracy. Since BPS is based on the log-polar translation, it is also invariant to rotation and scaling. By comparing the differences between the target shapes, the similarity between two targets can be quantified with a number between 0 to 1. In this paper, we use correlation-based similarity comparison to get the value of h_2.

$$h_2 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \cdot \sqrt{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2}}, \tag{5}$$

where x_i stands for the edge of standard figure and y_i stands for the rectified edge of changed figure.

As shown in Fig. 5, we select a 400 $\times$ 400 image with a yellow car as the target, which covers about 10% (160 $\times$ 80) of the whole image. The car is moved 2 pixels each step within 50 pixels. For each step, we rotate the car from 10° to 360° with a step of 10°, and change its size from 0.9 to 1.1 times with a step of 0.02 times. By applying the rotation and scaling transformation, we can get 46 (36 from rotation and 10 from scaling) changed targets from the standard target. Next, these transformed images are mapped to log-polar coordinates to evaluate the match accuracy of SIFT, the identification degree of BPS, and the evaluation value of LPT-MP. For the particular eccentricity, the minimum value among 46 values is chosen as the final result in each group.

3.2 Result and Tolerable Range

The comparison of SIFT, BPS and LPT-MP with a left shift (for eccentricity) is shown in Fig. 6. For all three methods, the values are represented within a range of 0 to 1. In the plot, the gray line with circular markings shows the impact of eccentricity on LPT-MP, the red line with rectangles represents the change in the matching result of SIFT, and the yellow line with triangles represents the accuracy of BPS. All three index values decrease with an increase in eccentricity (moving steps). In case of slight eccentricity, the fovea captures most of the target area, and the effect of eccentricity on the characteristic of invariance to rotation and scaling is insignificant. As the eccentricity increases, the corresponding quantitative values of all three methods decrease rapidly, and becomes low for a high number of step size.

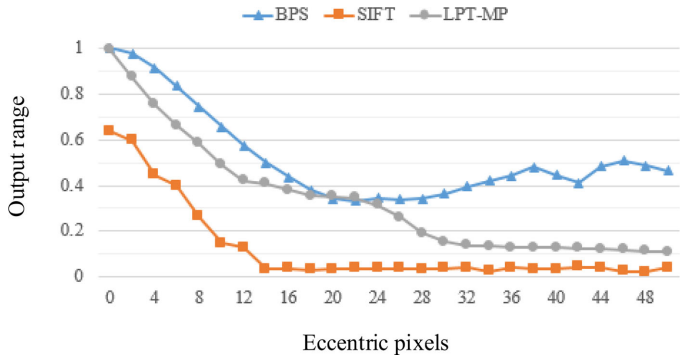


Fig. 6. Comparison of SIFT, BPS, and LPT-MP with left shift.

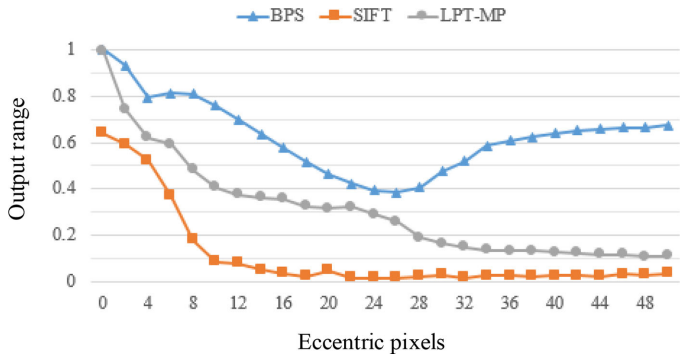


Fig. 7. Comparison of SIFT, BPS, and LPT-MP with up shift.

In practice, the matching result of SIFT should larger than 0.45 for acceptable performance, and the degree of BPS should large than 0.85. In order to meet the rate of feature point matching and the degree of target recognition at the same time, the eccentric pixel should choose the minimum one. Therefore, the value of LPT-MP should large than 0.70, which represents the eccentricity of 5 pixels.

Fig. 7 shows the comparison by up-shifting the target. To meet the standard such that the matching result of SIFT should large than 0.45 and the degree of BPS should large than 0.85, the eccentricity could be 5 pixels and 3 pixels, respectively. By selecting the minimum one, the eccentricity could be 3 pixels, which gives the value of LPT-MP to be above 0.69.

The same evaluation can be performed for the other two directions, i.e., in the right and downward directions. The results show that the value of LPT-MP should large than 0.67 and 0.71 for the down and right directions, respectively. In this way, we calculate the value of LPT-MP on the entire imaging plane with the eccentricity of 30 pixels for each direction, and simulate the tendency of the rotation and scaling invariance, as shown in Fig. 8. Similar to the previous results, the result of LPT-MP shows a high value when the eccentricity is low, and decreases rapidly with an increase in the eccentricity distance. It can be concluded from the results that the decrease in the value of LPT-MP is mostly influenced by the shape of the target, and less effected by the eccentric direction. Thus, the response of LPT-MP can be characterized around rings. That is to say the characteristic of invariance to rotation and scaling in retina-like imaging is only affected by the eccentric offset distance. The larger the eccentricity, the poorer the ability of invariance to rotation and scale change in the retina-like imaging. For optimal performance, the LPT-MP should maintain a value above 0.69, which is an area marked by the red ellipse shape in Fig. 8. Thus, the circular area around the target centroid is the area which represents the tolerable range of eccentricity, at which the characteristic of rotation and scaling invariance are maintained in retina-like imaging.

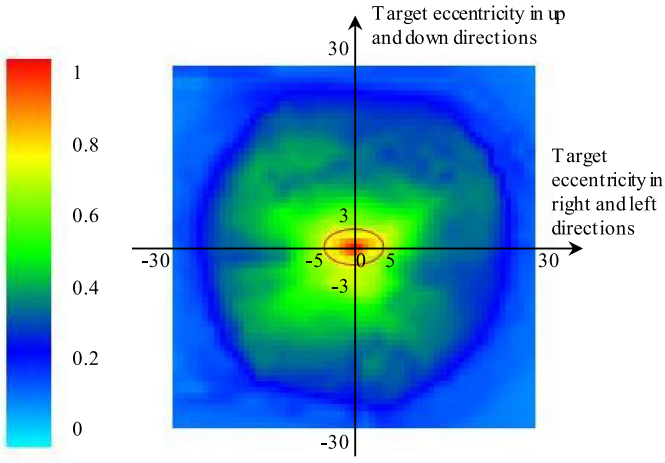


Fig. 8. Tolerable range of the target of car.

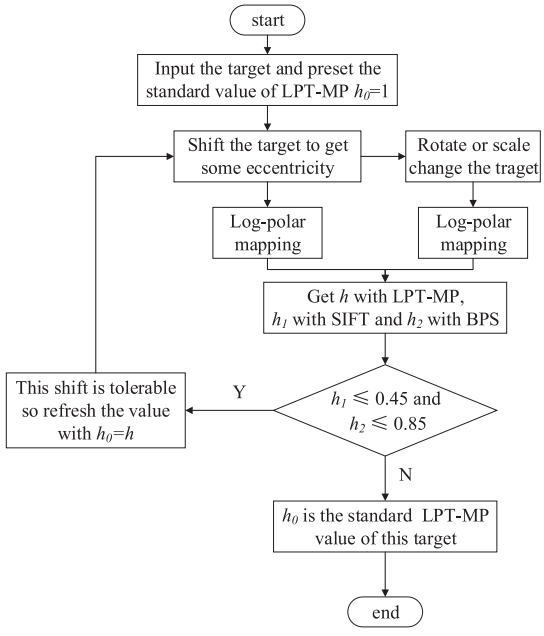


Fig. 9. The flowchart of the determining of h.

4. Discussion

4.1 Applicability Evaluation of LPT-MP

To evaluate the performance and applicability of LPT-MP, it is applied on different targets with varying shapes. We choose another two different targets, human and airplane, with the size of 400 $\times$ 400 and shift them in four directions.

The targets used in this experiment are the most common occurring targets in practical scenarios with different features. Through the method shown in Fig. 9, we can get the results as presented in Table 1. We move each target in different directions and determine the shifted pixel with the desired threshold values of SIFT and BPS. The pixel numbers are further used to estimate the standard value for LPT-MP. For each target, the four directional values of LPT-MP are averaged together to get the final standard value, which is used to defines the tolerable range. It can be observed from Table 1 that for the same offset direction of the same target, the pixel number of “0.45 of SIFT”

TABLE 1
Quantitative Comparison of LPT-MP for Different Targets

Target	Moving direction	Pixel number (shifted) for different standards		Corresponding value of LPT-MP	Average value of LPT-MP for each target
		0.45 of SIFT	0.85 of BPS		
car	up	5	**3**	0.69	0.69
	down	4	**3**	0.67	
	left	6	**5**	0.70	
	right	**5**	**5**	0.71	
human	up	**3**	4	0.54	0.51
	down	**4**	**4**	0.51	
	left	**3**	**3**	0.48	
	right	**3**	**3**	0.49	
airplane	up	6	**5**	0.48	0.46
	down	**5**	**5**	0.45	
	left	**5**	**5**	0.42	
	right	**4**	**4**	0.50	

and “0.85 of BPS” have very small difference. Similarly, it can be seen that for a particular target, the values of LPT-MP in different directions are almost the same. However, the standard value for LPT-MP varies significantly across different targets. From the simulation experiments, we find these standard values of LPT-MP for car, human and airplane to be 0.69, 0.51 and 0.46 respectively. By carefully analyzing the data in Table 1., we can see a clear relationship between the target shape and the standard value of LPT-MP. Through simulations, we find that the concavity or convexity of the target shape effects the standard value. For example, in case of ‘car’, the concave area around the target is very small, and the LPT-MP values is high. In contrast, for the ‘human’ target, there exist a lot of concavity in the area between the arms and legs, and around the target; and, therefore, the standard value is low in this case. Similarly, for the ‘airplane’, due to target’s large wings and slender body shape, the concave area is large, and the standard values of LPT-MP is low. Thus, the LPT-MP values can be used to distinguish different targets.

4.2 Limitation of LPT-MP

Although the proposed LPT-MP can be efficiently applied on a variety of targets, there still exists some limitations. Firstly, we find that the value of LPT-MP is the same for all the targets. This standard value changes with target shape. The higher the target concavity, the lower the standard value of LPT-MP is. Secondly, even though the ability to evaluate the rotation and scale changes with LPT-MP is efficient, the tolerable range can only be estimated when the scale change is close to 1. A large scale change has a high impact on the log-polar mapping, especially in the presence of eccentricity.

5. Conclusion and Future Work

A novel method is proposed to evaluate the invariance of rotation and scaling in the retain-like imaging. By incorporating the proposed evaluation method with the conventional methods for target matching and target recognition task, we can find the tolerance of the mismatch between the center

of FOV and the target centroid. Simulations are carried out to test the feasibility of the proposed method. From the results, we find that the proposed method applies to both rotation and scale change, and works efficiently in different directions of eccentricity. The standard value of LPT-MP changes with the change in target shape. If a relationship between the values of LPT-MP and target shape can be found, or a database of these values can be maintained for comparison, the LPT-MP can also be used as target matching algorithm. We leave this for future work. In this paper, we evaluate the performance of the proposed LPT-MP metric on various commonly occurring targets and find their standard values. The evaluation by the proposed metric will motivate the use of retina-like imaging for various applications.

References

[1] Z. Zhang, A. Beck, and N. Magnenat-Thalmann, "Human-like behavior generation based on head-arms model for robot tracking external targets and body parts," *IEEE Trans. Cybern.*, vol. 45, no. 8, pp. 1390–1400, 2015.

[2] J. Cao *et al.*, "Design and realization of retina-like three-dimensional imaging based on a MOEMS mirror," *Opt. Lasers Eng.*, vol. 82, pp. 1–13, Jul. 2016.

[3] S. Chao *et al.*, "High-resolution ISAR imaging of maneuvering targets based on sparse reconstruction," *Signal Proc.*, vol. 108, pp. 535–548, Mar. 2015.

[4] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2864–2875, 2013.

[5] W. Jiang and J. Gu, "Video stitching with spatial-temporal content-preserving warping," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2015, pp. 42–48.

[6] Y. Cheng *et al.*, "Review of state-of-the-art artificial compound eye imaging systems," *Bioinspiration Biomimetics*, vol. 14, no. 3, 2019.

[7] K. Akşit *et al.*, "Manufacturing application-driven foveated near-eye displays," *IEEE Trans. Visual. Comput. Graph.*, vol. 25, no. 5, pp. 1928–1939, 2019.

[8] J. J. Kim, H. Liu, A. Ousati Ashtiani, and H. Jiang, "Biologically inspired artificial eyes and photonics," *Rep. Prog. Phys.*, vol. 83, no. 4, 2020, Art. no. 047101.

[9] A. Wannig, L. Stanisor, and P. R. Roelfsema, "Automatic spread of attentional response modulation along Gestalt criteria in primary visual cortex," *Nature Neuroscience*, vol. 14, no. 10, pp. 1243–1244, 2011.

[10] F. Wang *et al.*, "Optimization of retina-like sensor parameters based on visual task requirements," *Opt. Eng.*, vol. 52, no. 4, 2013, Art. no. 043206.

[11] S. Gauglitz, T Höllerer, and M. Turk, "Evaluation of interest point detectors and feature descriptors for visual tracking," *Int. J. Comput. Vis.*, vol. 94, no. 3, pp. 335–360, 2011.

[12] A. Mikulik, O Chum, and J Matas, "Image retrieval for online browsing in large image collections," *Similarity Search Appl.*, vol. 8199, pp. 3–15, 2013.

[13] D. Mishkin, J. Matas, and M. Perdoch, "MODS: Fast and robust method for two-view matching," in *Proc. Comput. Vis. Image Understanding*, 2015, pp. 81–93.

[14] R. Narayanan *et al.*, "Considerations and framework for foveated imaging systems," *Photonics*, vol. 5, no. 3, 2018.

[15] E. Akbas and M. P. Eckstein, "Target detection through search with a foveated visual system," *PLOS Comput. Biol.*, vol. 13, no. 10, 2017, Art. no. e1005743.

[16] S Lee, M S Pattichis, and A C Bovik, "Foveated video quality assessment," *IEEE Trans. Multimedia*, vol. 4, no. 1, pp. 129–132, 2002.

[17] R.-D. Snježana, M. Vranješ, and D. Žagar, "Foveated mean squared error—A novel video quality metric," *Multimedia Tools Appl.*, vol. 49, no. 3, pp. 425–445, 2010.

[18] G. Cheng, P. Zhou, and J. Han, "Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7405–7415, 2016.

[19] C. Gong *et al.*, "Learning rotation-invariant and fisher discriminative convolutional neural networks for object detection," *IEEE Trans. Image Process.*, vol. 28, pp. 265–278, 2018.

[20] B Herbert *et al.*, "Speeded-up robust features," *Comput. Vis. Image Understanding*, vol. 110, no. 3, pp. 404–417, 2008.

[21] Y. Wan and N. Wei, "A fast algorithm for recognizing translated, rotated, reflected, and scaled objects from only their projections," *IEEE Signal Process. Lett.*, vol. 17, no. 1, pp. 71–74, 2009.

[22] S. Jiao, C. Zhou, W. Zou, and X. Li, "Shift, rotation and scale invariant optical information authentication with binary digital holography," *Opt. Commun.*, vol. 405, pp. 53–59. 2017.

[23] T. Fujisawa and M Ikehara, "High-accuracy image rotation and scale estimation using radon transform and sub-pixel shift estimation," *IEEE Access*, vol. 7, pp. 22719–22728, 2019.

[24] T. Dobashi and H. Kiya, "A parallel implementation method of FFT-based full-search block matching algorithms," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, 2013, pp. 2644–2648.

[25] X. Zhang *et al.*, "A faster 1-D phase-only correlation-based method for estimations of translations, rotation and scaling in images," *IEEE Trans. Fundamentals Electron. Commun. Comput. Sci.* vol. E97.A, no. 3, pp. 809–819, 2014.

[26] D G Lowe, "Distinctive image features from scale-invariant Keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[27] J. Yan, M. Ding, and P. Zhou, "A fast target recognition algorithm based on LPT," *CAAI Trans. Intell. Syst.*, vol. 3, no. 4, pp. 370–376, 2008.