

Iterative Template Matching with Rotation Invariant Best-Buddies Pairs

Zhuo Chen, Yang Yang^(✉), Weile Chen, Qian Kou, and Dexing Zhong

The School of Electronic and Information Engineering, Xi'an Jiaotong University,
Xi'an, China
yyang@mail.xjtu.edu.cn

Abstract. In this paper, we propose a new method for template matching method with rotation invariance. Our template matching can not only find the location of the object, but also annotate its rotation angle. The key idea is to firstly rectify the local rotation patches according to their intensity centroids, and then to find the corresponding patch-features between template and target images under an iterative matching framework. We adopt the coarse-to-fine search ways, so the patch size should be updated accordingly, which is time-consuming. To tackle this problem, we use the integral image to update the intensity centroid to accelerate the computing speed. The corresponding feature matching is based on the Best-Buddies Pairs (BBPs), which is robust to the non-rigid transform of local range and outliers. Experimental results demonstrate the effectiveness and robustness of the proposed algorithm.

Keywords: Template matching · Rotation invariance
Intensity centroid · Integral image

1 Introduction

Template matching is an essential work in the computer vision field, which has many applications in the object tracking, 3D reconstruction and image registration. Corresponding techniques has rapidly developed in recent years, which can be generally divided into two types: the similarity measure algorithm and local feature matching algorithm.

The former type relies mostly on the global search matching. Some good similarity measure functions such as the Sum of Squared Differences (SSD) [1], Sum of Absolute Differences (SAD) [2] and Histogram has been used to judge the difference between the template and target images. Mattoccia et al. [3] proposed to use the bounded partial correlation (BPC) with the similarity measure, which reduced the computation cost by comparing the correlation between template and target images. Furthermore, to solve the matching problem when rotation occurs, some methods [4, 5] combined the circular projection transformation (CPT) or radial projections [6, 7] to the conventional methods. However, when the region of interest has high-level outliers like object's non-rigid transformation, background noise and partial occlusion. These global similarity methods

can not automatically distinguish the outliers. Recently, a new similarity measure called Best-Buddies Similarity (BBS) [8] was proposed. The key idea is to count the number of a subset of nearest neighbor patch-pairs – the Best-Buddies Pairs (BBPs) between two point sets. BBPs are robust to outliers. But, since the BBS is under the slide-window searching, it cannot annotate the angle of the object of interested.

The other type of algorithm is to extract local invariant features between two images. The Scale-Invariant Feature Transform (SIFT) [9] is a popular method in this type. The SIFT is robust to the light change in the scene, various object deformation and other complex conditions. SURF [10] is another famous feature matching method. It is 10 times faster than SIFT with little decline of the performance. Feature line based method [11] is also a local feature instead feature points in SIFT or SURF. The line structure provides more information to consists more accurate descriptors and achieve better matching results. However, these methods are dependent much on the assumption that enough edges or corner points in the object could be distinguished from the background. This limits the application in the wild.

This paper concentrates on how to achieve rotation invariance during the matching process, as well as to be robust to the outliers, partial occlusion, local rigid or non-rigid deformation of the object in the wild. Our purpose can not only find the location of the object, but also annotate its rotation angle. We proposed to matching rotation invariant BBPs under an iterative matching framework, which combines the advantages of the above two types of methods. The key idea is to firstly rectify the local rotation patches according to their intensity centroids, and use the integral image to accelerate the computing speed. And then we find the corresponding BBP patch-features and update the rotation and translation parameter to update the location. Compared with the similar work proposed by Luo [12], our method has a great superiority on the template matching with large rotation angles. Experimental results demonstrate the proposed method is effectiveness and robustness.

2 Proposed Method

In template matching, rotation of the object in target images differs in various scenes. How to reduce the error brought by the rotation is our first concern in template matching. Intensity centroid is chosen in our method to overcome the problem and achieve matching by improved Best-Buddies Similarity algorithm.

Figure 1 shows the process of our proposed method. The algorithm mainly includes following steps: 1. divide template image and target image into $k \times k$ patches and utilize intensity centroid to realize rotation invariance; 2. find corresponding patch-feature in processed template and target image; 3. calculate the matching parameters of position and locate the target object in this iteration.

In each iteration, we need to calculate matching parameter by acquired corresponding patch-features in template and target image to locate the target object and move the template image for next iteration. The corresponding patch-features are two point sets $P = \{p_i\}_{i=1}^N$ and $Q = \{q_j\}_{j=1}^M$, where P is the template

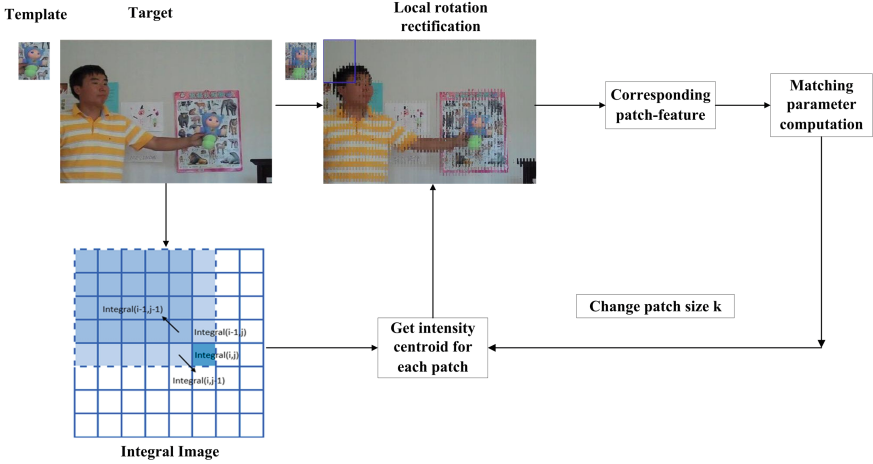


Fig. 1. Matching process.

set and Q is the target set, and p_i, q_j represents the coordinate. The matching parameters are rotation matrix R and translation vector t are computed between P and Q by minimizing the following distance function:

$$(R, t) = \underset{i=1}{\operatorname{argmin}} \sum_{i=1}^M \|R * p_i + t - q_i\|. \quad (1)$$

The template is set at the upper left corner in target image as its first iteration position. In each iteration, the position of the template changes according to the parameter. Meanwhile the patch size k also changes for matching from coarse to precise. However, different k means different intensity centroid for each patch in two images which will cost more time in computing. We choose integral image to help improve computing speed. The following subsections will discuss the details of our method.

2.1 Local Rotation Rectification

To obtain rotation invariance for our method, we use an effective measure of orientation—the intensity centroid.

Intensity centroid. The intensity centroid defines that a patch’s intensity is offset from its geometrical center, thus the vector from the geometrical center to the intensity centroid can be used to get the orientation. Rosin [13] defines the moments of a patch as:

$$m_{pq} = \sum_{x,y} x^p y^q I(x, y). \quad (2)$$

where x, y is the coordinate of each pixel and $I(x, y)$ is the grey value.

Then we can find the centroid:

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right). \tag{3}$$

We can construct a vector from a patch’s geometrical center, O, to the intensity centroid, \vec{OC} . The orientation of the patch is:

$$\theta = \text{atan2}(m_{01}, m_{10}). \tag{4}$$

where atan2 is the quadrant-aware version of arc tan.

In template matching, template image and target image are divided into $k * k$ distinct patches. For each path, we need to find the centroid to get an orientation. We choose integral image to complete it simply and quickly.

Integral image. The integral image is also called summed area table (SAT). It is a fast and efficient data structure and algorithm for calculating the sum of a rectangular sub-region of a grid. The value of each point (x, y) in the integral image, I, is the sum of all the gray-scale values, i, in the upper left corner of the corresponding position in the image:

$$I(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'). \tag{5}$$

Integral image has a good property that it can be calculated by traversing the whole image only once. Equation can be rewritten as:

$$I(x, y) = i(x, y) + I(x - 1, y) + I(x, y - 1) - I(x - 1, y - 1). \tag{6}$$

In our experiment, we construct three integral images for each image. The first one is the ordinary integral image corresponding to m_{00} in Eq. (2). The second one is calculated after multiplying gray-scale values of each position at (x, y) by x-coordinate. It corresponds to m_{10} . Analogously, the last one corresponds to m_{01} obtained by multiplying the y-coordinate.

2.2 Corresponding Patch-Features

Finding correlation points is an essential step in template matching. Feature based approach such as scale-invariant feature transform (SIFT) extract feature points by local information like edges and corners from images, but it will fail to find enough correlation points in images with less edges or corners.

Our method chooses a new proposed similarity measure called Best-Buddies Similarity (BBS) to get correlation points. BBS is robust to outliers from the background and complex deformations. These advantages enable it to obtain plenty of feature points. The key of BBS is to compute corresponding features between two sets of points. The corresponding feature is relied on the Best-buddies Pairs (BBPs) extracted from the two sets of points.

BBPs. There are two sets of points $P = \{p_i\}_{i=1}^N$ and $Q = \{q_j\}_{j=1}^M$ from two images. Each point has two kind of information. RGB shows its pixel appearance and (x, y) shows its pixel location. A pair of points from P and Q respectively can be regarded as BBP when they satisfy the following equation:

$$bb(p_i, q_j, P, Q) = \begin{cases} 1 & NN(p_i, Q) = q_j \wedge NN(P, q_j) = p_i \\ 0 & otherwise \end{cases} \quad (7)$$

The equation means that the nearest neighbor of p_i in set Q is q_j , and vice versa. $NN(p_i, Q)$ and $NN(P, q_j)$ is calculated by:

$$NN(p_i, Q) = argmin_{q \in Q} d(p_i, q). \quad (8)$$

$$NN(P, q_j) = argmin_{p \in P} d(p, q_j). \quad (9)$$

$$d(p_i, q_j) = (x_i - x_j)^2 + (y_i - y_j)^2 + \lambda * [(r_i - r_j)^2 + (g_i - g_j)^2 + (b_i - b_j)^2]. \quad (10)$$

r, g, b is the RGB value of the point. λ is a weight number and we choose $\lambda = 2$ in our experiments.

2.3 Implementation Detail

Our method aims to achieve template matching with rotation invariance. The algorithm procedure is concluded as follows:

1. Given template image, T and target image, I , calculate the integral images as mentioned in Sect. 2.1;
2. Break the image regions into $k * k$ distinct patches and update the pixel RGB value in each patch by rotating a degree which is got from intensity centroid;
3. Calculate $d(p_i, q_j)$ to obtain BBPs sets P and Q in I, T ;
4. Get matching parameters R, t to obtain the computing position for next iteration;
5. Iterate steps 2 to 4 until the variation of parameters R and t reach the threshold or iteration ends.

The innovation of our algorithm is that we use intensity centroid combined with integral image to achieve rotation invariance for each patch and change the patch size k as iteration time increases to accelerate the computing process.

Specifically, each $k * k$ patch in the image contains $k * k$ pixels. Using the three integral images, we can simply get m_{01}, m_{10} and m_{00} to compute the intensity centroid for a patch when we know the coordinates of the lower right corner coordinates of the patch in the original image. Then the rotation angle for each patch is got from the vector \overrightarrow{OC} as mentioned before and all pixels in the patch rotate to get new RGB value. When computing $d(p_i, q_j)$ the patch is regarded as a ‘point’– the center coordinates of the patch denote the location and all pixels’ new RGB value in the patch denote appearance. For different k ,

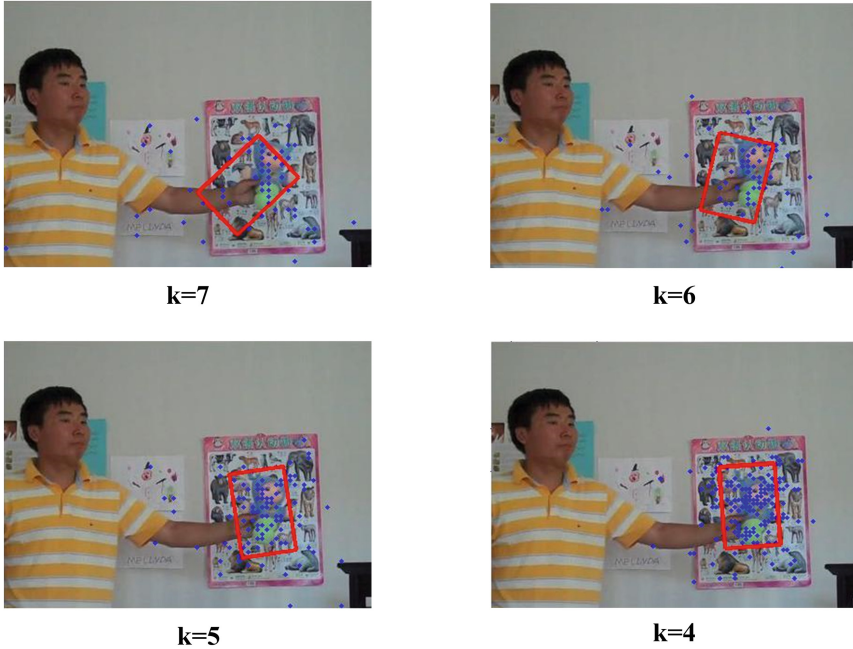


Fig. 2. Results under different patch size k from 7 to 4. (Color figure online)

we can just look up in the integral image to calculate the intensity centroid for the patch, which is much less time-consuming.

In practice, we change the patch size k in the iteration. The k is getting smaller with the iteration time increases. According to Eq. (10), in the first iteration, the BBPs of target image I mostly locate in upper left corner in I . With iteration time increases, they will gradually converge to the correct matching region. Initially, the size k is set larger. This will accelerate the convergence speed to make the BBPs in I approach an approximate region. Then the k is getting smaller ($k > 3$) for accurate matching. Smaller k will help find more BBPs and better matching result with more time consumed. Figure 2 shows the different results under different patch size k .

In Fig. 2 blue dots are the BBPs in the target image and red rectangular is the matching result. When size k is too big, the matching speed is faster but less BBPs will be got. While k is small, the speed will decrease but more BBPs are got. We combine the two properties of different size to adjust size k during the iteration process, so we can get enough BBPs to help matching and spend less time.

3 Experimental Results

Experiments are conducted on a computer with 3.10 GHz CPU and 8G memories. The images are chosen randomly. To verify the better performance of our

algorithm, testing images in our experiments are taken from real-world scenarios and some are from the BBS work [9]. To demonstrate effectiveness of the matching method, we compare the results with original BBS algorithm and an improved algorithm (B+S) by Luo [12], which can also deal with rotation in template matching.

We set up two groups of experiments. First one uses images which are rotated by 0° , 90° , 180° , 270° respectively. The second one uses images which contain rotated template in the target image. As shown in Fig. 4, our results outperform other two algorithms. For each template with rotation of 90° or 270° , the BBS fails in most cases because the sliding window search can only find locations of the objects. Since the BBS doesn't have rotation invariant features, it loses the object in Fig. 4(4). The B+S method also doesn't use rotation invariant features. So it cannot get accurate results when large rotation happens.

We adopt the overlap ratio to quantitatively compare the results, which is the percentage of the overlapping region with the ground truth. The overlap results are shown in Fig. 3. We can see that our method get the highest overlap under every large rotation condition.

Template ID ^o	Overlap ^o											
	0 ^o			90 ^o			180 ^o			270 ^o		
	Ours ^o	B+S ^o	BBS ^o	Ours ^o	B+S ^o	BBS ^o	Ours ^o	B+S ^o	BBS ^o	Ours ^o	B+S ^o	BBS ^o
(1) ^o	0.95 ^o	0.94 ^o	0.98 ^o	0.89 ^o	0.74 ^o	0 ^o	0.93 ^o	0.81 ^o	0.06 ^o	0.87 ^o	0.88 ^o	0 ^o
(2) ^o	0.93 ^o	0.92 ^o	0.89 ^o	0.89 ^o	0.73 ^o	X ^o	0.88 ^o	0.91 ^o	0.54 ^o	0.90 ^o	0.93 ^o	X ^o
(3) ^o	0.95 ^o	0.95 ^o	0.95 ^o	0.93 ^o	0.93 ^o	0.71 ^o	0.92 ^o	0.83 ^o	0.90 ^o	0.96 ^o	0.87 ^o	0.68 ^o
(4) ^o	0.96 ^o	0.94 ^o	0 ^o	0.96 ^o	0.53 ^o	0 ^o	0.93 ^o	0.82 ^o	0 ^o	0.95 ^o	0.47 ^o	0.41 ^o
Mean ^o	0.95 ^o	0.94 ^o	0.71 ^o	0.92 ^o	0.73 ^o	0.24 ^o	0.92 ^o	0.84 ^o	0.38 ^o	0.92 ^o	0.79 ^o	0.36 ^o

Fig. 3. Overlap of experiment 1

To further testify the rotation invariance of our method, we repeat the experiment after rotating the images by every 30° . Figure 5 shows the detailed results.

Under every rotation condition, overlap-rate represents the area ratio of the matching result covering the ground truth.

$$OverlapRate = \frac{area(Result \cap GT)}{area(GT)}. \quad (11)$$

The rotation accuracy is the results' rotating angle compared to the current image rotation angle:

$$Rotationaccuracy = \frac{\theta_{Result}}{\theta_{GT}}. \quad (12)$$

where the θ_{Result} is the rotation angle which is got from matching result and θ_{GT} is the ground truth angle we annotated manually.

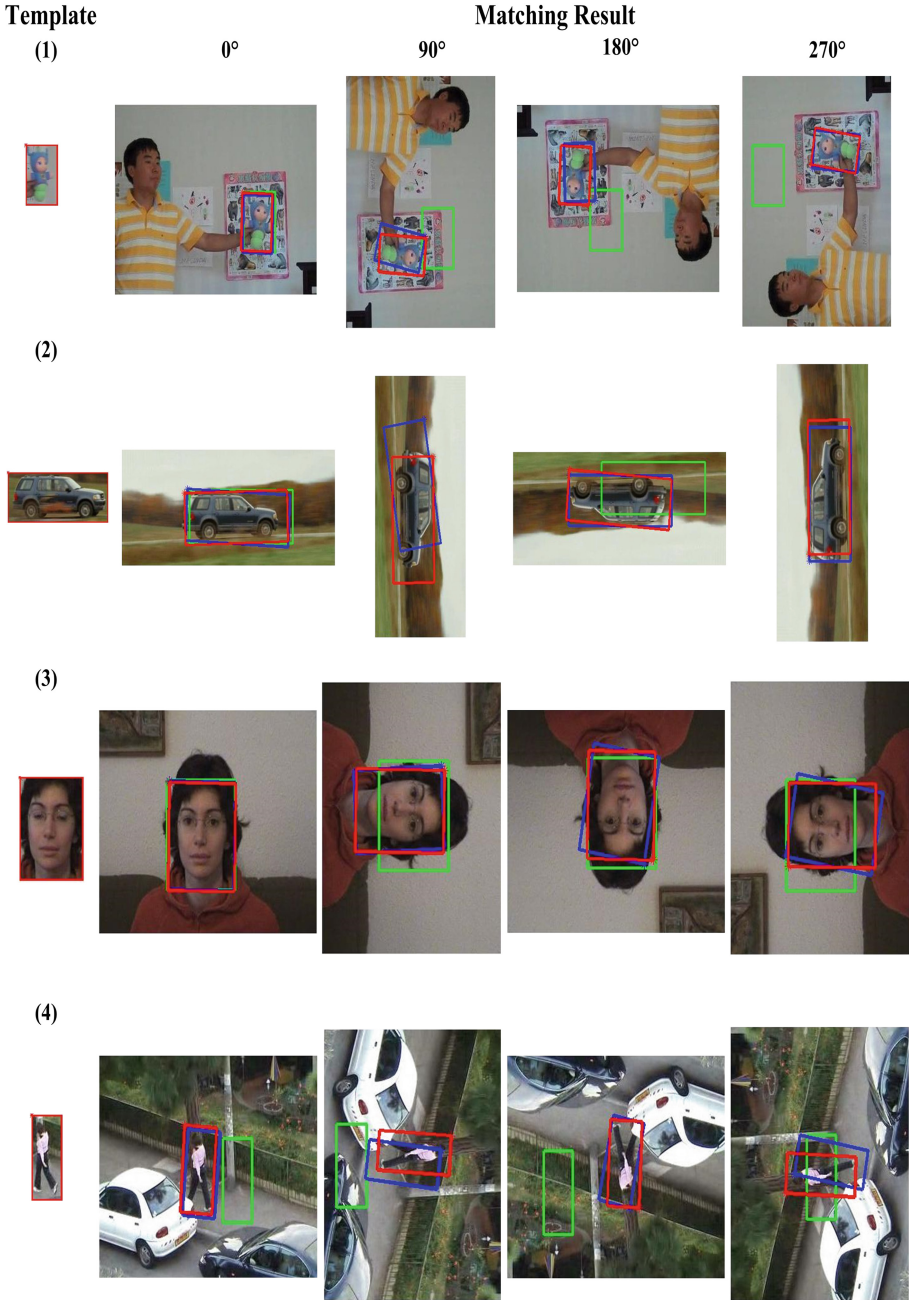


Fig. 4. Template matching results of different rotated images in experiment 1. The green rectangular is BBS's [9] result; blue one is B+S [12] result; red one is ours. (Color figure online)

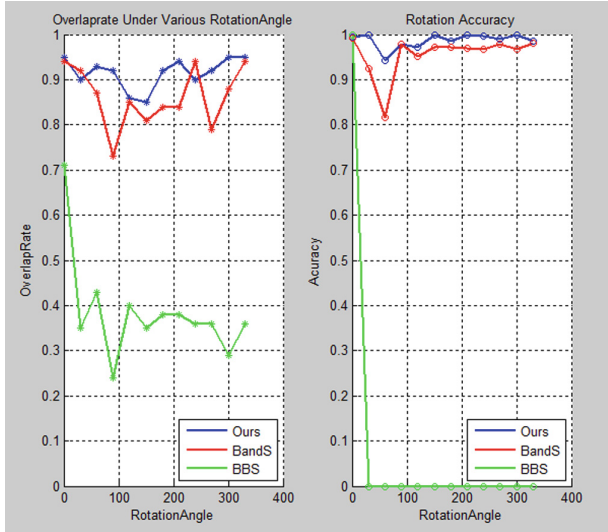


Fig. 5. The overlap-rate of compared three methods is shown on the left and the rotation accuracy is on the right.

The two parameters are between thresholds from 0 to 1. Bigger value means better result. As shown in Fig. 5, the proposed scheme has the highest overlap rate and rotation accuracy which demonstrate its good performance and rotation invariance under different angles.

In the experiment 2, we compute the rotation error (RE) as Eq. (13) by comparing the matching angle with the ground truth. The smaller RE is, the more accurate results are.

$$RE = \frac{\theta_{Result} - \theta_{GT}}{\theta_{GT}}. \quad (13)$$



Fig. 6. Template matching results of different rotated template in target image in experiment 2. The green rectangular is BBS's [9] result; blue one is B+S [12] result; red one is ours. (Color figure online)

Template ID ^o	Overlap ^o			Rotation angle ^o				Rotation error ^o		
	Ours ^o	B+S ^o	BBS ^o	Ours ^o	B+S ^o	BBS ^o	GT ^o	Ours ^o	B+S ^o	BBS ^o
(1) ^o	0.91 ^o	0.89 ^o	0 ^o	45 ^o	36 ^o	0 ^o	47 ^o	-4.3% ^o	-23.4% ^o	-1 ^o
(2) ^o	0.86 ^o	0.93 ^o	0.73 ^o	-20 ^o	-20 ^o	0 ^o	-22 ^o	-9.1% ^o	-9.1% ^o	-1 ^o
(3) ^o	0.97 ^o	0.62 ^o	0.70 ^o	-40 ^o	-5 ^o	0 ^o	-44 ^o	-9.1% ^o	-88.6% ^o	-1 ^o

Fig. 7. Result of experiment 2

The compare RE results and the corresponding images are shown in Figs. 5 and 6, respectively. Among the three algorithms, our method gets the best results in the overlap and most accurate rotation angle as shown in Fig. 7. All the above experimental results demonstrate that the proposed method has the best rotation-invariance among the three methods.

4 Conclusion

This paper proposed a template matching method with robustness and rotation in-variance. The method employed the intensity centroid to rectify the local rotation patches and used integral image to reduce the computing time. Then BBPs is used to get corresponding points and the rigid transform parameters are updated iteratively. Our method is demonstrated to have better performance under large rotation angles compared with other methods in the experimental results. Future work will introduce the scaling invariance and more robust features in our methods.

Acknowledgement. This work was in part supported by the National Natural Science Foundation of China (61627811, 61573274), Postdoctoral Science Foundation of China (2015M582661), Natural Science Foundation of Shaanxi, China (No. 2015JQ6257) and the Fundamental Research Funds for the Central Universities (xjj2017005, xjj2017067)

References

1. Stefano, L.D., Mattocchia, S.: Fast template matching using bounded partial correlation. *Mach. Vis. Appl.* **13**(4), 213–221 (2003)
2. Li, W., Salari, E.: Successive elimination algorithm for motion estimation. *IEEE Trans. Image Process.* **4**, 105–107 (1995)
3. Mattocchia, S., Tombari, F., Stefano, L.D.: Fast full-search equivalent template matching by enhanced bounded correlation. *IEEE Trans. Image Process.* **17**(4), 528–538 (2008)
4. Lee, W.C., Chen, C.H.: A fast template matching method with rotation invariance by combining the circular projection transform process and bounded partial correlation. *IEEE Signal Process. Lett.* **19**(11), 737–740 (2012)
5. Lin, Y.H., Chen, C.H., Wei, C.C.: New method for sub-pixel image matching with rotation invariance by combining the PT method and the RPT process. *Opt. Eng.* **45**(6), 067202(1)–067202(9) (2006)

6. Choi, M.S., Kim, W.Y.: A novel two stage template matching method for rotation and illumination invariance. *Pattern Recogn.* **35**, 119–129 (2002)
7. Kim, Y.H.: Rotation-discriminating template matching based on Fourier coefficients of radial projections with robustness to scaling and partial occlusion. *Pattern Recogn.* **43**(3), 859–872 (2010)
8. Dekel, T., Oron, S., Rubinstein, M., Avidan, S., Freeman, W.T.: Best-buddies similarity for robust template matching. In: *Computer Vision and Pattern Recognition*, pp. 2021–2029 (2015)
9. Lowe, D.: Distinctive image features from scale-invariant key points. *Int. J. Comput. Vis.* **60**(2), 91–110 (2004)
10. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: speeded up robust features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006*. LNCS, vol. 3951, pp. 404–417. Springer, Heidelberg (2006). https://doi.org/10.1007/11744023_32
11. Zhang, Y., Qu, H.S.: Rotation invariant feature lines transform for image matching. *J. Electron. Imaging* **23**(5), 053002 (2014)
12. Luo, S., Yang, Y., Du, S.Y., Kou, Q.: Robust image registration with rotation, scale, translation using best-buddies pairs. In: *International Joint Conference on Neural-Networks*, pp. 1728–1732 (2016)
13. Rosin, P.L.: Measuring corner properties. *Comput. Vis. Image Underst.* **73**(2), 291–307 (1999)