An Improved Harris-SIFT Algorithm Based on Rotation-invariant LBP Operator

Lei Yang[1], Yanyun Ren[1], Jiyuan Cai[1] and Huosheng Hu[2]

[1]*Shanghai Key Laboratory of Power Station Automation Technology,
School of Mechatronic Engineering and Automation, Shanghai University,
Shanghai, 200072, China*
[2]*School of Computer Science and Electrical Engineering, University of Essex,
CO4 3SQ, United Kingdom*
*yangyoungya@sina.com

Abstract

Feature-points matching is an important concept in binocular stereo vision. The procession of multi-scale feature-points matching in classical Harris-SIFT algorithm is time-consuming and has high complexity when describing the feature-points. This paper proposed a new improved Harris-SIFT algorithm based on rotation-invariant LBP (Local binary patterns) operator. Firstly, the Harris operator is used to extract feature points from DOG (Difference of Gaussian) scale space. Then, the dominant direction of feature point is calculated and 81-dimensional rotation-invariant LBP descriptors are extracted when the rotation matching window is coordinated to this direction. At last, Best-Bin-First (BBF) algorithm is used to search the matching points between the two sets of feature points. Experimental results show that the proposed algorithm is lower time-consuming than classical Harris-SIFT algorithm and remains the similar matching correct rate.

Keywords: Harris-SIFT; DOG (Difference of Gaussian) scale space; rotation-invariant LBP operator; Best-Bin-First (BBF) algorithm

1. Introduction

Feature-points matching is an important concept in binocular stereo vision. The accuracy of matching is closely related to the reliability of the acquired depth information about scene. How to extract feature points accurately, choose the reliable feature descriptors and reduce the time-consuming of matching operation are the huge bottlenecks that feature-point matching algorithms encounter. And lots of improved algorithms have been proposed to solve these problems. According to the difference of feature points extraction operator, matching algorithms can be divided into Moravec operator, Harris operator, SUSAN(Smallest Univalue Segment Assimilating Nucleus) operator and so on. Moravec operator has been used successfully to match 3D point sets which captured from stereo vision platform [1]. The operator is vulnerable to oblique edge because it only considers the variation of per 45 degrees, which does not consider all other directions. And its window function is a two-valued function which endows same weight to every element. The distance to the center point is not considered. Therefore, Moravec operator is easy to be affected by noise and its location is not much accurate. Harris operator is proposed in 1988 by C. Harris and M. Stephens. Since it can maintain invariant to rotation and affine transformation, it has been widely used in corner detection and image matching [2]. The operator has low computational complexity because only the

---

*Corresponding Author

gradients in x and y axis needs to be calculated. SUSAN operator shows well performance in noise resistance. Grey value of Feature point and its neighbors are compared and gradients calculations are not involved in the procession. It is also has isotropic characteristic as Harris operator. As it can detect edge and extract corner points accurately, the operator has been used in image feature detection and matching [3-4].

The SIFT algorithm is a kind of feature matching algorithm which is depend on DOG scale space. It has excellent performance in solving rotation, scaling or radiation transformation problems of images matching [5]. SIFT algorithm has good characteristics in partial feature extraction of feature-points [6] and is widely used in remote sensing image matching [7], biological feature recognition [8], and so on. But it also has some drawbacks such as too many feature points, too large dimension of the descriptor and too high time consuming of computation in matching process. These made the SIFT algorithm can't meet the real-time requirement in the practical application. Some new methods have been proposed to improve execution speed of SIFT algorithm such as by using GPU on SIFT [9], combining PCA-based local descriptors with SIFT algorithm [10], SURF (Speeded Up Robust Features) algorithm [11] and so on. A new Harris-SIFT algorithm [12] has been proposed which use Harris corner detection to replace extreme point detection of scale space. It has well effect on reducing the amount of calculation in feature-points extraction and makes the feature-points have physical characteristics and visual characteristics. In order to make the Harris-SIFT algorithm satisfy the real-time requirement, some algorithms have been proposed to reduce the dimension of feature-points descriptor [13-14]. When calculating the same number of feature-points, improved algorithm can outperform SIFI from the perspective of the consuming time, in which time-consuming nearly reduces 30%. It is achieved by translating the feature-points descriptor with 128-dimension into 32-dimension. Although the running speed of improved algorithm has been improved, the accuracy of feature-points description decreases at the same time. Reduction of descriptor dimension leads to the difficulty of matching.

The LBP (Local binary patterns) is a descriptor to describe local texture feature [15]. It has low amount of computation and high grey monotone, rotation invariance. It is widely used in facial image analysis, texture retrieval and facial recognition [16-18]. CS-LBP (center- symmetric local binary pattern) is used as the local feature descriptor in the SIFT algorithm and receives great performance in [19]. In order to improve the computation speed of SIFT, the rotation-invariant LBP has been used as the feature-descriptor [20-21]. The rotation-invariant LBP can be received by doing bitwise right circulation for LBP. The speed of rotation-invariant LBP can be improved 2.14 times compared with SIFT algorithm. Only simple arithmetical operations are needed in rotation-invariant LBP and it has the characteristics of rotational invariance. We use Harris operator to extract feature-points in this paper. In order to speed up operation of the algorithm, rotation-invariant LBP is used as the feature operator in our proposed algorithm. Experimental results show that the proposed algorithm has almost same accuracy compared with Harris-SIFT and slightly lower than SIFT. And the proposed algorithm is faster than SIFT and it reduces the running time as 28.07% when describing same numbers of feature-points.

The remainders of this paper are organized as follows. The process of Harris feature points detection is presented in Section 2. The LBP methods which include classical LBP and rotation-invariant LBP are analyzed in Section 3. The proposed algorithm and matching process are discussed in Section 4. Experimental results and discussion are presented in Section 5. At last, Section 6 reports the conclusion.

2. Harris Feature-Points Detection

In this section, the DOG scale space is described firstly. Then, Harris operator is discussed. The feature point orientation assignment is introduced in the end of this section.

2.1. DOG Scale Space

The original image collected by camera is inevitable has different scale even for the same object. We have to eliminate the influence of different scale before using descriptor to describe feature points. One method is image registration. However, it will change the original image. Another method is building several images based on different scale space from original image which is proposed by Iijima firstly [22]. The SIFT algorithm chooses the latter method. In [23], it is proved that the only possible scale space kernel is Gaussian function. And the scale space of image can be defined as $L(x, y, \sigma)$, which is obtained by the convolution of a Gaussian function $G(x, y, \sigma)$ and image $I(x, y)$ such as

$$L(x, y, \sigma) = G(x, y, \sigma) \otimes I(x, y) \tag{1}$$

and

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp(-\frac{x^2 + y^2}{2\sigma^2}) \tag{2}$$

where $\otimes$ is convolution operation, and σ is the scale of Gaussian function. The scale decides the smoothness of the image. Large scale is corresponding to general view of image. Oppositely, small scale is corresponding to detail view of image. In order to detect feature points locations, the DOG (difference of Gaussian) scale space had been proposed by Lowe in 1999. The DOG function can be expressed by $D(x, y, \sigma)$, which is computed from subtracting the nearby scales space such as

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) \otimes I(x, y) \tag{3}$$

where k is a constant multiplicative factor and its value has almost no impact. In order to contain different scale, a DOG images pyramid has been build as shown in Figure. 1. The DOG images pyramid contains several octave and each octave also contains several DOG images. It is clear that the more DOG images each octave has, the more accurate this pyramid will be. At the same time, the algorithm will get too much computation time.
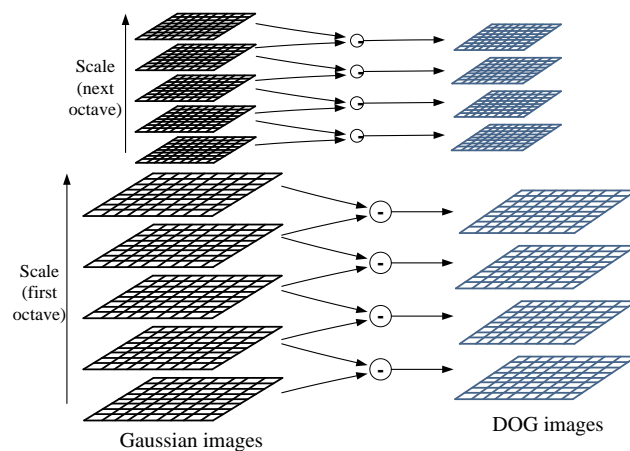

Figure 1. DOG Images Pyramid

2.2. Harris Operator

Harris operator has a lot of advantages in feature-points detection. It takes small amount of calculation. The extracted feature-points are evenly distributed. And the operator has visual characteristics. The improved version of Harris obtains the best results when several interest point detectors are evaluated by two evaluation criteria such as repeatability rate and information content [24]. Even though there has low scale change, points are geometrically stable under all tested image variations. It is also more representative than extreme detection which used in SIFT algorithm. In DOG scale space, we use Harris operator to extract feature- points. Matrix M must be constructed before Harris operator is used to extract feature- points. Matrix M is calculated as

$$M = G(x, y, \sigma) \otimes \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{4}$$

where I_x and I_y are the GOD scale space gradient in horizontal and vertical direction respectively. And we extract feature points by the equation as

$$R = (AB - CD)^2 - a(A + B)^2 \tag{5}$$

where a is an experienced value between 0.04 and 0.06. We compare R with some threshold value. If R is great than threshold value, we define the monitoring point is feature point. Otherwise, we define the monitoring point is not feature point. Harris operator is sensitive to Gaussian scale window. Large scale window is corresponding to few feature-points, and small scale window is corresponding to superior feature-points. Harris operator is translation- invariant and rotation-invariant because it is just related with gradient operation.

2.3. Orientation Assignment

In order to reduce matching difficulty of rotation image, orientation assignment is a necessary step in feature-points description. Gradient magnitude $m(x, y)$ and orientation $\theta(x, y)$ of every element in neighborhood of feature-points are calculated as

$$m(x, y) = \sqrt{(L(x + 1, y) - L(x - 1, y))^2 + (L(x, y + 1) - L(x, y - 1))^2} \tag{6}$$

$$\theta(x, y) = \tan^{-1}((L(x, y + 1) - L(x, y - 1))/(L(x + 1, y) - L(x - 1, y))) \tag{7}$$

Orientation histogram is formulated and it has 36 bins, which established ten degree for each interval. Each gradient magnitude and orientation values are added to the histogram after being weighted by Gaussian filter. The highest peak is considered to be the dominant orientation. And other peaks which are within 80% of highest peak are considered to be the auxiliary orientation of feature-points for the sake of improving the robustness. In order to get better accuracy, parabolic curve is used for every orientation which has detected. Before describing feature-points, the coordination should be rotated to its dominant orientation and auxiliary orientation.

3. LBP Feature Operator

Classical LBP operator, Rotation-invariant LBP operator and rotation-invariant LBP operator are discussed separately in this section.

3.1. Classical LBP Operator

LBP is a kind of texture description operator. It is used in supporting local contrast of image and measuring and extracting the image local texture information. It is invariant to illumination because threshold processing is used. For a pixel at (x_o, y_o), its pixel value will be compared with its 8 neighbors in a $3*3$ neighborhood by subtraction. If the result is negative, it will be assigned with value of 0, the others with 1. Then all these binary values can be encoded to a binary number in a clockwise direction and starting point is the pixel $(x_o - 1, y_o - 1)$. LBP can be obtained by converting the binary number to its decimal form as

$$LBP\ (x_o, y_o) = \sum_{i=0}^{7} s(l_i - l_0)2^i \tag{8}$$

where l_i and l_o are the values of surrounding pixels and central pixel respectively. The function $s(x)$ is defined as

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{9}$$
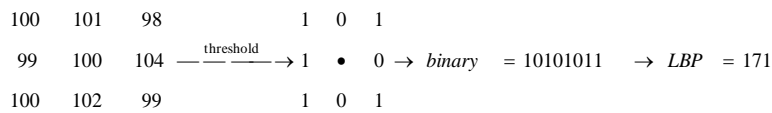
Figure 2 shows an example.

100	101	98			1	0	1					
99	100	104	$\xrightarrow{threshold}$		1	•	0	$\to$ *binary*	= 10101011	$\to$	*LBP*	= 171
100	102	99			1	0	1					

Figure 2. Example of Classical LBP Operator

It is clear in Figure 2 that the values of 8 neighbor pixels are compared with the value of centre pixel and the compared results are assigned with binary value by Eq.(9). Then all binary values are encoded as a binary number. At last we use Eq.(8) to get final LBP value, which is equal to 171.

3.2. Rotation-Invariant LBP Operator

Since the basic LBP operator cannot meet the need of feature-points matching in a pair of rotate images, rotation-invariant LBP operator with radius and several numbers of sampling points will be used. Rotation-invariant LBP operator can be defined as $LBP_{P,R}^{ri}(x_o, y_o)$, where P is the number of sampling points, R is the radius of sampling circle. As the values of P and R can be adjusted to satisfy practical application, several classical rotation-invariant LBP operator are shown in Figure 3.
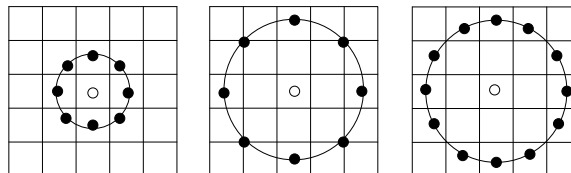

Figure 3. Several Classical Rotation-Invariant LBP Operator

The binary number of LBP is required to bitwise right circulation and each binary number is convert to its decimal form. Here we can select the minimum value as the rotation-invariant LBP operation as

$$LBP_{P,R}^{ri}(x_0, y_0) = \min\{ ROR \ (LBP_{P,R}, i), i = 0, 1 \ldots, P-1 \} \tag{10}$$

where ROR is performed as circular bitwise right circulation, i performs how many times bitwise right circulation is done, and $LBP_{P,R}^{ri}(x_o, y_o)$ is the minimum value of all the P different binary numbers. So it is rotation-invariant. For the example shown in Figure 2, all its 8 binary numbers are described in Figure 4.

As shown in Figure 4, 87 is the minimum value. So all the 8 rotation-invariant LBP operator is equal to 87 such as $LBP_{8,1}^{ri} = 87$. It is well known that 8-bit binary number can represent 256 decimal numbers. All $LBP_{8,R}^{ri}$ have 36 values, which are shown as in [21].
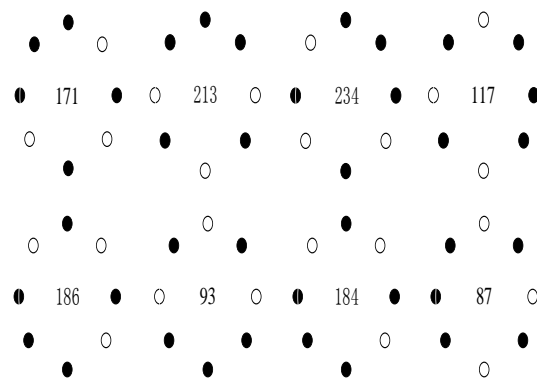

Figure 4. Eight Binary Numbers of Example in Figure 2

3.3. Rotation-Invariant LBP Feature Operator

Because feature-points marching need to describe enough information, all the pixels in $9*9$ neighborhood of feature-points must extract their rotation-invariant LBP operators. The 81 result operators can be arranged as vector $H = (h_1, h_2 \ldots \ldots h_{81})$, where h_i denotes operator of ith pixel. To eliminate the light influence, we normalize the vector H as

$$H_1 = \frac{H}{\|H\|} \tag{11}$$

The vector H_1 also contains the influence of camera saturation change. We set a threshold to cut off the beyond elements and then do the nest normalization processing until it is convergent. The final vector will be used in matching process as feature descriptor.

4. The Proposed Algorithm

By using the classical Harris-SIFT, we proposed an improved algorithm based on rotation-invariant LBP operator. Firstly, the DOG scale spaces of match image are established. The more Gaussian scales are used, the DOG scale space will show more detailed. Because extreme point detection does not have visual and physical characteristics, Harris operator is used to extract feature-points in Harris-SIFT algorithm. In our algorithm, 81-dimensional vector is used to construct the rotation-invariant LBP operator for every feature-point. This will make the algorithm has low computational complexity. And the matching algorithm of Best Bin First (BBF) is used. After getting the descriptor vectors of two images, we use

the Euclidean distance of vectors to measure the similarity of different feature-points. Firstly, we select a description vector (corresponding to the feature point A) in image 1 and calculate Euclidean distances between it and all description vectors in image 2, then sort all the distance values and choose the minimum value (corresponding to the feature point B), second-minimum value(corresponding to feature point C). At last, we compute the ratio of minimum value and second-minimum value. If the ratio is less than a threshold value, we determine that feature point A has matching feature point B successfully, which is shown as

$$t > \frac{\| H_A - H_B \|}{\| H_A - H_C \|} \qquad (12)$$

The procedure of the proposed algorithm can be summarized as following six steps:

Step 1: Get the images to be matched.
Step 2: Structure the DOG scale space of images with different Gaussian scale.
Step 3: Use Harris operator to extract feature points on DOG scale space.
Step 4: Compute dominant orientation of feature-points and rotate axis to this dominant orientation.
Step 5: Use rotation-invariance LBP operator to describe the pixels in neighborhood of feature points and arrange to 81-dimensional vector.
Step 6: Matching feature points by comparing the ratio of minimum and second-minimum Euclidean distance with a threshold value.

5. Experimental Results and Discussion

In this section, several groups of matching experiments for standard images Cones and Tsukuba and our captured images are performed by using SIFT algorithm, classical Harris-SIFT algorithm and the proposed algorithm in MATLAB 7.0. Experimental results are presented and discussed.

5.1. Experiment on the Benchmark Images

Experimental results of image Cones are shown as Figure. 5. Figure 5(a) and (b) are the test images cones, Figure (c-e) illustrate the matching results of SIFT algorithm, classical Harris-SIFT algorithm and the proposed algorithm respectively. The cyan lines denote the correct matching points and red lines are the error matching points. The marching results of numerical comparing are shown in Table 1.
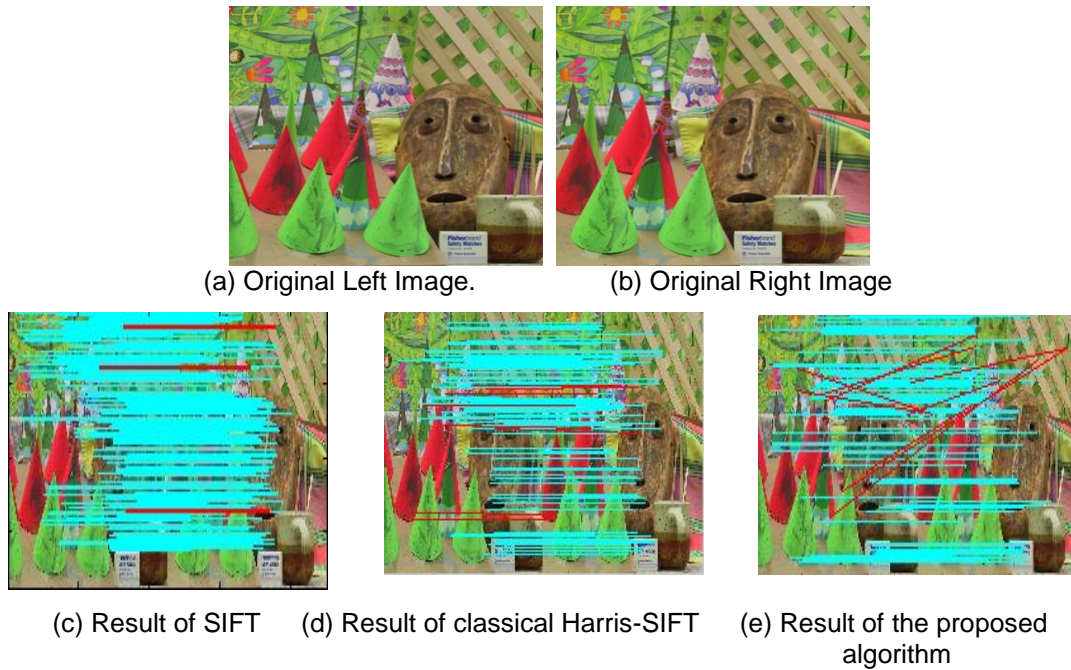
(a) Original Left Image.   (b) Original Right Image

(c) Result of SIFT   (d) Result of classical Harris-SIFT   (e) Result of the proposed algorithm

Figure 5. Experimental Results of Cones

Experimental results of image Tsukuba are shown as Figure 6. Figure 6.(a) and (b) are the test images, figure (c-e) are the matching results of SIFT algorithm, classical Harris-SIFT algorithm and our proposed algorithm respectively.
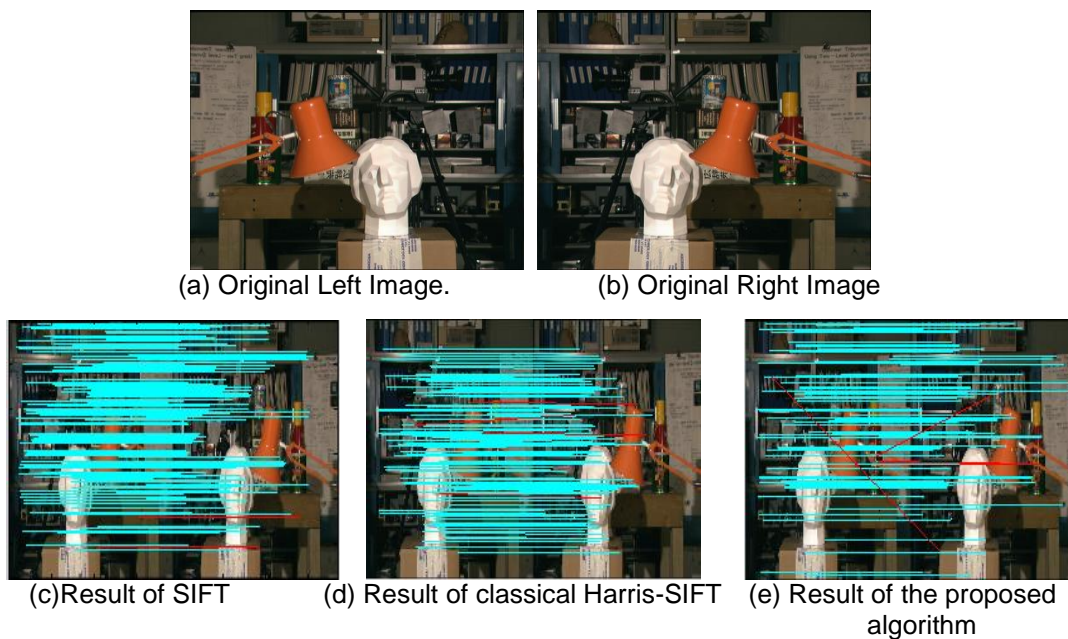

(a) Original Left Image.   (b) Original Right Image

(c)Result of SIFT   (d) Result of classical Harris-SIFT   (e) Result of the proposed algorithm

Figure 6. Experimental Results of Tsukuba

Table 1 gives the numerical comparing results of SIFT algorithm, classical Harris-SIFT algorithm and our proposed algorithm.

Table 1. Matching Results of Benchmark Images

Algorithm	Matching number		Error matching number		Matching rate(%)	
	Cones	Tsukuba	Cones	Tsukuba	Cones	Tsukuba
SIFT	400	344	3	2	99.25	99.42
Harris-SIFT	255	204	10	4	96.08	98.04
The proposed method	256	262	9	5	96.48	98.09

5.2. Experiment on Our Captured Images

In this subsection, experimental results of our captured images by us are reported. Figure. 7 shows the experimental results. Figure 7.(a) and (b) are a pair of captured images, Figure (c-e) are the matching results of SIFT algorithm, classical Harris-SIFT algorithm and our proposed algorithm respectively. In Figure 7. (c-e), the cyan lines denote the correct matching points and red lines are the error matching points. The numerical comparing results are shown in Table 2. From Table 2, we see that the matching rate of our proposed algorithm is very close to Harris-SIFT and slightly lower than SIFT. This means that our algorithm has relative high precision.

(a)Original Left Image.　　　　(b)Original Right Image

(c)Result of SIFT　　　(d)Result of classical Harris-SIFT　　　(e) Result of the proposed algorithm

Figure 7. Experimental Results of Our Captured Images.

Table 2. Matching Results of our Captured Images

Algorithm	Matching number	Error matching number	Matching rate(%)
SIFT	191	2	98.95
Harris-SIFT	158	2	98.73
The proposed method	149	3	97.99

As we know, the most time-consuming part in marching algorithm is in the process of describing feature-points. In order to compare time-consuming of feature-points description in classical Harris-SIFT algorithm and the proposed algorithm, we use these two algorithms to describe 100 feature points respectively. The results of time-consuming are shown in Table 3. We see that the classical Harris-SIFT algorithm used 114ms to describe 100 feature points and our proposed algorithm consumed 82ms. The time-consuming is reduced by 28.07%. We also notice that two factors speed up the

algorithm. The first one is descriptor vector. In classical Harris-SIFT algorithm, the descriptor vector is 128-dimension. However, our proposed algorithm just needs 81-dimension. Short vector is corresponding to low time complexity. The second factor is that the gradient magnitude and orientation of pixels are used to determine descriptor vector in classical Harris-SIFT algorithm. Several square root and anti-tangent operation are needed in the procession of computation. All these are more complex than our proposed algorithm which just use comparing and bitwise circulate operation.

Table 3. Time-Consuming Comparing Results

Algorithm	Matching number	Time-consuming(ms)
Harris-SIFT	100	114
The proposed method	100	82

6. Conclusion

In this paper, we have proposed an improved Harris-SIFT algorithm based on rotation- invariant LBP operator. The feature-points in classical Harris-SIFT algorithm are visual corners. The rotation-invariant LBP operator has advantages of multi-scale, low computing time-complexity. In our method, the Harris operator is used to extract feature-points from DOG (Difference of Gaussian) scale space, and 81-dimensional rotation-invariant LBP descriptors are used. Best-Bin-First (BBF) algorithm is also used to search the matching points between the two sets of feature-points. In order to evaluate the performance of the proposed algorithm, experiments on Cones and Tsukuba are performed. Experimental results show that the proposed algorithm is lower time- consuming than that of classical Harris-SIFT algorithm. And the proposed algorithm has similar matching rate compared with classical Harris-SIFT algorithm.

Acknowledgements

References

[1] L. Y. Jiang, J. M. Liu, D. C. Li and Z. L. Zhu, "3D Point Sets Matching Method Based o Moravec Vertical Interest Operator", Proceedings of the 2011 2nd International Congress on Computer Applications And Computational Science, vol. 144, (**2012)**, pp. 53-59.

[2] M. Y. Ding, L. L. Li, C. P. Zhou and C. Cai, "A multi-sensor image registration method based on Harris corner matching", Interactive Technologies And Sociotechnical Systems, vol. 4270, **(2006)**, pp. 174-183.

[3] M. Y. Weng and M. Y. He, "Image feature detection and matching based on SUSAN method", First International Conference on Innovative Computing, Information and Control, Vol 1, Proceedings, **(2006)**.

[4] E. Vincent and R. Laganière, "Detecting and matching feature points", Journal of Visual Communication and Image Representation, vol. 16, no. 1, **(2005)**, pp. 38-54.

[5] D. G. Lowe, "Distinctive image features from scale-invariant keypoints", International Journal of Computer Vision, vol. 60, no. 2, **(2004)**, pp. 91-110.

[6] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors, IEEE Trans Pattern Anal Mach Intell, vol. 27, no. 10, **(2005)**, pp. 1615-30.

[7] B. Kupfer, N. S. Netanyahu and I. Shimshoni, "An Efficient SIFT-Based Mode-Seeking Algorithm for Sub-Pixel Registration of Remotely Sensed Images", IEEE Geoscience And Remote Sensing Letters vol. 12, no. 2, **(2015)**, pp. 379-383.

[8] S. Malathi and C. Meena, "Improved partial fingerprint matching based on score level fusion using pore and sift features, Process Automation, Control and Computing", (PACC), 2011 International Conference on, **(2011)**, pp. 1-4.

[9]   Acharya and R. V. Babu, "Speeding up SIFT using GPU, Computer Vision, Pattern Recognition", Image Processing and Graphics (NCVPRIPG), 2013 Fourth National Conference on, **(2013)**, pp. 1-4.

[10] Y. Ke and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors", Computer Vision and Pattern Recognition, 2004. Proceedings of the 2004 IEEE Computer Society Conference on, vol. 2 , **(2004)**, pp. 506 -513.

[11] H. Bay, T. Tuytelaars and L. V. Gool "SURF: Speeded up robust features", Computer Vision - Eccv 2006 , Part 1, Proceedings, vol. 3951, **(2006)**, pp. 404-417.

[12] Q. Zhao and H. Wei, "Harris-SIFT Algorithm and Its Application in Binocular Stereo Vision, Journal of University of Electronic Science and Technology of China, vol. 39, no. 4, **(2010)**, pp. 546-550.

[13] M. Wang and W. Liu. Advanced algorithm based on SIFT and its application in binocular stereo vision, Jisuanji Gongcheng yu Yingyong(Computer Engineering and Applications) vol. 49, no.2, **(2013)**, pp. 203-206.

[14] Y. Liu, J. Wu, Y. Fan and G. Huang, "Application of improved Harris-SIFT algorithm in binocular stereo vision", Transducer and Microsystem Technologies, vol. 33, no. 6, **(2014)**, pp. 151-160.

[15] T. Ojala, M. Pietikainen and D. Harwood, "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions", Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on, vol. 1, **(1994)**, pp. 582-585.

[16] D. Huang, C. F. Shan, M. Ardabilian, Y. H. Wang and L. M. Chen, "Local Binary Patterns and Its Application to Facial Image Analysis: A Survey", IEEE Transactions on Systems Man And Cybernetics Part C-Applications And Reviews, vol. 41, no. 6, **(2011)**, pp. 765-781.

[17] L. Wang, R.-F. Li, K. Wang and J. Chen, "Feature representation for facial expression recognition based on FACS and LBP", International Journal of Automation and Computing, vol. 11, no. 5, **(2014)**, pp. 459-468.

[18] X. H. Yang, X. Y. Yao, L. J. Cai and D. F. Li, "A Generalized Local Binary Pattern Descriptor and Its Applications to Texture Retrieval", 2012 Ieee Fifth International Conference on Advanced Computational Intelligence (Icaci), **(2012)**, pp. 590-596.

[19] M. Heikkilä, M. Pietikäinen and C. Schmid, "Description of interest regions with local binary patterns", Pattern recognition, vol. 42, no. 3, **(2009)**, pp. 425-436.

[20] T. Ojala, M. Pietikainen and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", IEEE Trans. on Pattern Analysis And Machine Intelligence vol. 24, no. 7, **(2002)**, pp. 971-987.

[21] Y.-b. Zheng, X.-s. Huang and S.-j. Feng, "An image matching algorithm based on combination of SIFT and the rotation invariant LBP, Journal of computer-aided design & computer graphics, vol. 22, no. 2, **(2010)**, pp. 286-292.

[22] T. Iijima, "Basic theory of pattern normalization (for the case of a typical one dimensional pattern)", Bulletin of the Electro technical Laboratory, vol. 26, **(1962)**, pp. 368-388.

[23] J. J. Koenderink, "The structure of images", Biol Cybern, vol. 50, no. 5, **(1984)**, pp. 363-70.

[24] C. Schmid, R. Mohr and C. Bauckhage, "Evaluation of Interest Point Detectors", International Journal of Computer Vision, vol. 37, no. 2, **(2000)**, pp. 151-172.

Authors

Lei Yang was born in 1976. He received his Ph.D degrees in control science and engineering from the School of Electronics and Information Engineering, Xi'an Jiaotong University in 2008. He worked as post-doctoral researcher in Tongji University from July 2008 to June 2010. Now he is an associate researcher and postgraduate advisor in School of Mechatronics Engineering and Automation, Shanghai University. His research interests are in artificial intelligence and pattern recognition, computer vision and image processing. E-mail: yangyoungya@sina.com.
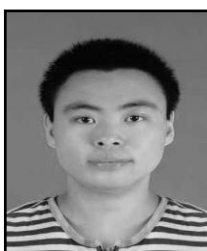
Yanyun Ren was born in 1991. He received his B.S. degree in architecture electric and intelligent from Jilin Architecture University in 2013. He is currently working towards his M.S degree in pattern recognition and intelligent system in Shanghai University. His research interests are obstacle and moving target detection. E-mail:973911878@qq.com.

Jiyuan Cai was born in 1991. He received his B.S. degree Electrical and Information Engineering from Nangjing University of Posts and Telecommunications in 2012. He is currently working towards his M.S degree in pattern recognitionand intelligent system in Shanghai University. His research interests are object detection and tracking. E-mail:caijiyuan95@126.com.

Huosheng Hu is a professor in School of Computer Science and Electronic Engineering, University of Essex, UK, leading the Robotics Research Group. His research interests include autonomous mobile robots, human-robot interaction, evolutionary robotics, embedded systems, pervasive computing, data fusion, multi-robot collaboration, and networked robotics. He has published over 380 papers in journals, books and conferences in these areas. He received the MSc degree from the Central South University, China in 1982, and the PhD degree from the University of Oxford, U.K. in 1993. He is Fellow of IET and InstMC and senior member of IEEE and ACM. E-mail: hhu@essex.ac.uk