MBR-SIFT: A Novel Descriptor for Image Matching

Mingzhe Su, Yan Ma, Xiangfen Zhang, Shunbao Li, Yuping Zhang

***Abstract*—The traditional SIFT method is capable of extracting distinctive feature for image matching. However, it is extremely time consuming in the SIFT matching due to applying the Euclidean distance measure. Recently, many binary SIFT (BSIFT) methods have been developed to improve matching efficiency, while none of them is invariant to mirror reflection. To address these issues, this paper presents a mirror reflection invariant binary descriptor, named MBR-SIFT. In addition, the Hamming distance has been employed as similarity measure for MBR-SIFT descriptor. The method is compared with other BSIFT methods, and is shown that MBR-SIFT is better both in accuracy and efficiency.**

***Index Terms*—mirror reflection, local descriptor, image matching, binarization**

I. INTRODUCTION

Local feature point has been successfully used in pattern recognition and computer vision applications such as image retrieval [1], object recognition [2], gesture recognition [3], texture recognition [4], and wide baseline matching [5]. A good descriptor should have robustness on photometric transformations such as brightness and highlight while being invariant to geometrical transformations including: rotation, scaling, viewpoint, and reflection [6].

Until recently, there are numerous feature descriptors have been proposed, in which scale invariant feature transform (SIFT) descriptor proposed by Lowe [7] is one of the most successful and popular local image feature descriptor. However, image matching in the SIFT method to large-scale image database would be extremely time-consuming. To solve this problem, several binary SIFT (BSIFT) methods have been proposed. Ni [8] first proposed a binary string approach for SIFT keypoints. His method exploits the Hamming distance to measure the similarity of two BSIFT vectors. Chen et al. [9] proposed to compare the absolute difference of two adjacent values in a descriptor with the threshold, which generated 128-bit BSIFT descriptor string. Zhou et al. [10] compared the 128 values of a SIFT descriptor individually with two threshold values. And correspondingly, 256-bit BSIFT descriptor string has been obtained.

The above mentioned BSIFT methods and their improved algorithms mostly ignore the problem of mirror reflection, which will result in significant increase in mismatch rate in face of mirror image pair. Guo [11] presented a mirror reflection invariant descriptor (MIFT) which is inspired from SIFT. Nevertheless, the matching time of MIFT is comparable to that of SIFT.

To address these issues, this paper presents a horizontal or vertical mirror reflection invariant binary descriptor, named MBR-SIFT. MBR-SIFT not only binarizes SIFT descriptor, but also takes into consideration the problem of mirror reflection. The Hamming distance is employed to measure the similarity of SIFT features in the matching. Experimental results on UKBench dataset show that the proposed method not only ensures the matching accuracy and speed, but also solves the issue of mirror reflection.

II. BSIFT

The SIFT algorithm extracts image features by searching the keypoints in the image, and then calculates the descriptors from the local region around the keypoints. As shown in Fig. 1(a), the local region is first divided into 16 cells with 8 directions in each cell, and each direction is

Manuscript received Nov. 18, 2016; revised Nov. 30, 2016. This work was supported in part by the National Natural Science Foundation of China (No.61373004) and Shanghai Normal University Innovation Team Project (No. A700115001005).

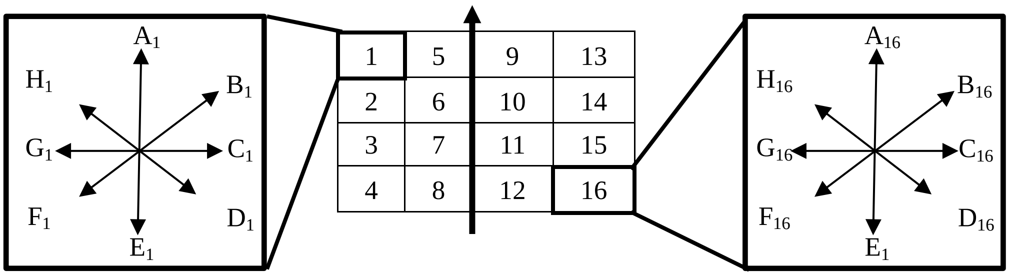
Mingzhe Su is with the College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai, China (e-mail: 1365702732@qq.com).

Yan Ma is with the College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai, China (e-mail: ma-yan@shnu.edu.cn).

Xiangfen Zhang is with the College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai, China (e-mail: xiangfen@shnu.edu.cn).

Shunbao Li is with the College of Mathematics & Science, Shanghai Normal University, Shanghai, China (e-mail: lsb@shnu.edu.cn).

Yuping Zhang is with the College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai, China (e-mail: yp_zhang@shnu.edu.cn).



(a)

A_1 B_1 C_1 D_1 E_1 F_1 G_1 H_1	2	3	4	5	6	7	8	…	15	A_{16} B_{16} C_{16} D_{16} E_{16} F_{16} G_{16} H_{16}

(b)

Fig. 1. Illustration of the descriptor organizations of SIFT. (a) The descriptor organizations in the original image; (b) SIFT descriptor for the original image.

given a value. Finally, 128-D SIFT descriptor, as shown in Fig. 1 (b), is obtained.

The SIFT binarization approach is to transform 128-D descriptor ($d_0, d_1, ..., d_{127}$) into a set of binary numeric string. The commonly used binarization approaches can be classified into two categories. The first category proposed to compare the differential value Ad_i of the two adjacent values in a descriptor with the predefined threshold *M*. The comparison result b_i is 0 or 1, only denoted by 1-bit [9].

The second category directly compares each d_i of 128-D descriptor ($d_0, d_1, ..., d_{127}$) with two thresholds, M_1 and M_2. The comparison result is 11, 10, and 00, denoted by 2-bit [10].

However, the two categories of BSIFT methods do not have a horizontal or vertical mirror robustness.

III. MBR-SIFT

Intuitively, an idea to make a BSIFT descriptor mirror reflection invariant is by artificially reflecting one of the matching image pair and performing once again image matching. This approach is simple but the time for matching is increased due to the repetitive execution of SIFT algorithm and binarization operation. BSIFT descriptor of mirror reflection image can be achieved if conducting a simple operation on the original BSIFT descriptor, which will lead to savings of computation time. The proposed binarization method (referred to as MBR-SIFT) is based on this idea.

A. SIFT Descriptor Reconstruction

By analyzing the structure of SIFT descriptor, we found that the connection between the BSIFT descriptors before and after mirror reflection can be built by reconstructing SIFT descriptor.

As shown in Fig. 2(a), the second and fourth columns of 16 cells are reorganized in reverse order respectively comparing to their original order in Fig. 1(a), and correspondingly, in Fig. 2(b), SIFT descriptor is reconstructed, in which the order of 16 cells is “1,2,3,4,8,7,6,5,9,10,11,12,16,15,14,13”. Similarly, in Fig. 2(c) and Fig. 2(d), the order of 16 cells in its horizontal or vertical mirror image is reconstructed as “13,14,15,16,12,11,10,9,5,6,7,8,4,3,2,1”. It can be seen that the order of 16 cells for the image and its mirror image just meets the reversal relation. In addition, each cell consists of 8 oriented gradients, *i.e.*, “$A_1B_1C_1D_1E_1F_1G_1H_1$” for the first cell in Fig. 1(a) and “$A_1H_1G_1F_1E_1D_1C_1B_1$” for the same cell after mirror reflection. It can be seen that there is no reversal relation. To ensure the reversal relation, 8 oriented gradients in 16 cells in Fig. 2(a) are reorganized by their respective direction and the reorganized SIFT descriptor (hereinafter referred to as R-SIFT descriptors) is obtained, as shown in Fig. 2(e) and (f). The 128 elements of original SIFT descriptor are organized by 16 cells and then 8 directions. In contrast, the 128 elements of R-SIFT descriptor are organized by 8 directions and then 16 cells, which ensures the reversal relation between 16 elements in same direction for the original image and its mirror image.

1	**8**	9	**16**
2	**7**	10	**15**
3	**6**	11	**14**
4	**5**	12	**13**

(a)

1	2	3	4	8	7	6	5	9	10	11	12	16	15	14	13

(b)

13	**12**	5	**4**
14	**11**	6	**3**
15	**10**	7	**2**
16	**9**	8	**1**

(c)

13	14	15	16	12	11	10	9	5	6	7	8	4	3	2	1

(d)

A_1	A_2	A_3	A_4	A_8	A_7	A_6	A_5	A_9	A_{10}	A_{11}	A_{12}	A_{16}	A_{15}	A_{14}	A_{13}	$B_1 \ldots B_{13}$	……	$G_1 \ldots G_{13}$	$H_1 \ldots H_{13}$

(e)

A_{13}	A_{14}	A_{15}	A_{16}	A_{12}	A_{11}	A_{10}	A_9	A_5	A_6	A_7	A_8	A_4	A_3	A_2	A_1	$H_{13} \ldots H_1$	……	$C_{13} \ldots C_1$	$B_{13} \ldots B_1$

(f)

Fig. 2. Illustration of the descriptor organizations of R-SIFT in the circumstance with and without mirror reflection. (a) The descriptor reorganizations in the original image; (b) SIFT descriptor for the original image; (c) The descriptor reorganizations in the horizontally reflected image; (d) SIFT descriptor for the mirror reflected image; (e) R-SIFT descriptor for the original image; (f) R-SIFT descriptor for the mirror reflected image.

Here, we denote R-SIFT descriptor with 128-D vector ($D_0, D_1, ..., D_{127}$) and the differential value with $AD_i (i = 0,1,...,127)$ which is calculated according to (1). Modulo operation in (1) ensures that AD_i is the difference of adjacent values in the same direction.

$$AD_i = \begin{cases} D_{i+1} - D_i & if \ \mathrm{mod}(i+1,16) \neq 0 \\ D_{i-15} - D_i & otherwise \end{cases} \quad (1)$$

The SIFT binarization method for AD_i is to compare AD_i with 0 and the comparison result is 0 or 1, only denoted by 1-bit. As a consequence, 128-D R-SIFT descriptor is transformed into 128-bit binary string (denoted as BR-SIFT) $\{b_0^1, b_1^1, \ldots, b_{127}^1\}$. This procedure can be illustrated by (2).

$$b_i^1 = \begin{cases} 1 & if \ \ AD_i \geq 0 \\ 0 & otherwise \end{cases} \quad (2)$$

B. Reverse Coding

After constructing BR-SIFT descriptor for the original image, we will illustrate how to construct their corresponding descriptors for its horizontal or vertical mirror image (denoted as MBR-SIFT).

From TABLE I, it can be seen that the R-SIFT descriptors of the original image and its mirror image are reversal relation. As a consequence, after differential operation, except that b_{15}^1 is changed into $\sim b_{15}^1$ (~ represents NOT operator), the rest bits in direction A of BR-SIFT and MBR-SIFT descriptors are mirror of each other. As for the first 15 bits in direction A, MBR-SIFT descriptor can be recovered by scanning BR-SIFT descriptor in reverse order and then conducting bitwise NOT operation. In addition, the order of 8 directions in BR-SIFT descriptor for the original image is “ABCDEFGH”, nevertheless, the order of 8 directions in MBR-SIFT descriptor for its mirror image is “AHGFEDCB”, therefore, it is required to exchange 16 binary values of direction B and H, direction C and G and direction D and F in BR-SIFT descriptor, respectively. Under this scheme, MBR-SIFT descriptor is constructed from BR-SIFT descriptor.

ISBN: 978-988-14047-3-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE I

THE COMPARISON OF DESCRIPTORS WITH OR WITHOUT REFLECTION IN A DIRECTION

original	R-SIFT descriptor	A_1	A_2	A_3	A_4	A_8	A_7	A_6	A_5	A_9	A_{10}	A_{11}	A_{12}	A_{16}	A_{15}	A_{14}	A_{13}
original	Differential value	A_2-A_1	A_3-A_2	A_4-A_3	A_8-A_4	A_7-A_8	A_6-A_7	A_5-A_6	A_9-A_5	$A_{10}-A_9$	$A_{11}-A_{10}$	$A_{12}-A_{11}$	$A_{16}-A_{12}$	$A_{15}-A_{16}$	$A_{14}-A_{15}$	$A_{13}-A_{14}$	$\mathbf{A_1-A_{13}}$
original	BR-SIFT descriptor	b_0^1	b_1^1	b_2^1	b_3^1	b_4^1	b_5^1	b_6^1	b_7^1	b_8^1	b_9^1	b_{10}^1	b_{11}^1	b_{12}^1	b_{13}^1	b_{14}^1	$\boldsymbol{b_{15}^1}$
mirror	R-SIFT descriptor	A_{13}	A_{14}	A_{15}	A_{16}	A_{12}	A_{11}	A_{10}	A_9	A_5	A_6	A_7	A_8	A_4	A_3	A_2	A_1
mirror	Differential value	$A_{14}-A_{13}$	$A_{15}-A_{14}$	$A_{16}-A_{15}$	$A_{12}-A_{16}$	$A_{11}-A_{12}$	$A_{10}-A_{11}$	A_9-A_{10}	A_9-A_9	A_6-A_5	A_7-A_6	A_8-A_7	A_4-A_8	A_3-A_4	A_2-A_3	A_1-A_2	$\mathbf{A_{13}-A_1}$
mirror	MBR-SIFT descriptor	$\sim b_{14}^1$	$\sim b_{13}^1$	$\sim b_{12}^1$	$\sim b_{11}^1$	$\sim b_{10}^1$	$\sim b_9^1$	$\sim b_8^1$	$\sim b_7^1$	$\sim b_6^1$	$\sim b_5^1$	$\sim b_4^1$	$\sim b_3^1$	$\sim b_2^1$	$\sim b_1^1$	$\sim b_0^1$	$\boldsymbol{\sim b_{15}^1}$

IV. MATCHING

Suppose I_1 and I_2 are image pair to be matched. For example, denote BR-SIFT descriptor of keypoint a_1 in I_1 with B_1, and denote BR-SIFT and MBR-SIFT descriptors of keypoint a_2 in I_2 with B_2 and MB_2, respectively. Then calculate the Hamming distance between B_1 and B_2, B_1 and MB_2, respectively and take a smaller value as the distance between keypoints a_1 and a_2. BR-SIFT descriptor of each keypoint of I_1 is compared to those (including BR-SIFT and MBR-SIFT descriptors) of all keypoints of I_2. And then determine the matching pairs according to (3).

$$match?\begin{cases} yes & if\ \ vals(1) < vals(2) * distratio \\ no & otherwise \end{cases} \tag{3}$$

where $vals(1)$ and $vals(2)$ represent the smallest distance and the secondary smallest distance, respectively. In addition, the predefined threshold *distratio* is empirically set to 0.65 in the experiment.

V. EXPERIMENT

We evaluate the proposed approach on the public dataset, *i.e.*, UKBench dataset [12]. The UKBench dataset contains 10,200 images from 2,550 object/scene groups. Each group consists of four images taken in different views or imaging conditions.

Fig. 3 illustrates the matching performance of SIFT, Chen', Zhou', and MBR-SIFT for images undergone reflection transformations, in which white lines and black thick lines represent correct matches and false matches, respectively. It can be seen from TABLE II that the performance of MBR-SIFT is more superior to the rest in terms of accuracy and recall. And the computational time is MBR-SIFT is slightly higher than Chen' method while lower than Zhou' method, and is almost half of that of SIFT method.

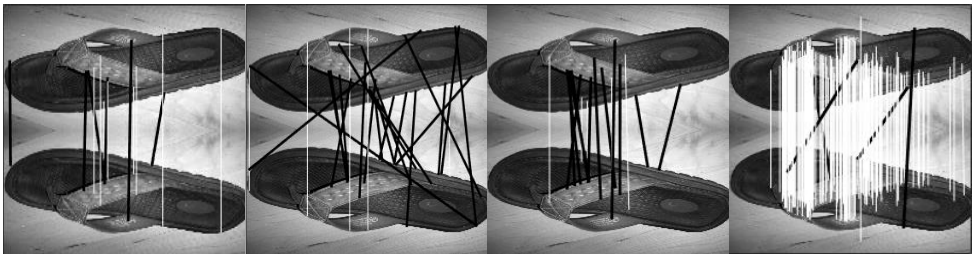
In addition, in order to verify the rotation, scale, viewpoint, lighting, and blur invariance to the proposed method, we randomly select 200 image pairs from the UKBench dataset for matching experiments. Some examples are shown in Fig. 4. And the corresponding number of correct matches and false matches are (190, 53, 163, 25, 25) and (3, 0, 0, 4, 4), respectively.

VI. CONCLUSION

This paper presents a binary SIFT descriptor (MBR-SIFT) which is achieved by reconstructing the SIFT descriptor. MBR-SIFT descriptor is invariant to mirror reflection while being robust on rotation, scaling, viewpoint, lighting, and blur changes. The experimental results show that the proposed method can achieve higher matching accuracy and recall for mirror reflection matching. How to further improve accuracy and recall will be the subject of future research.

ISBN: 978-988-14047-3-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

(a)



(b)

Fig. 3. Comparing the matching performance of SIFT, Chen', Zhou', and MBR-SIFT (from left to right) under mirror reflected transformation. (a) The original image and its horizontally reflected image; (b) The original image and its vertically reflected image.

TABLE II
MATCHING RESULTS OVER MIRROR REFLECTION

Image pair	Method	# of feature points (left/right) or (up/down)	# of matches	# of false matches	Accuracy (%)	Recall (%)	Computational Time(s)
(a)	SIFT	1051/1072	18	2	88.89	1.52	31.22
	Chen'	1051/1072	36	22	38.89	1.33	12.51
	Zhou'	1051/1072	15	9	40.00	0.57	20.81
	MBR-SIFT	1051/1072	76	0	100.0	7.23	12.79
(b)	SIFT	926/900	15	6	60.00	1.00	23.51
	Chen'	926/900	27	21	22.22	0.67	8.480
	Zhou'	926/900	17	12	29.41	0.56	15.67
	MBR-SIFT	926/900	205	4	98.05	22.33	8.580

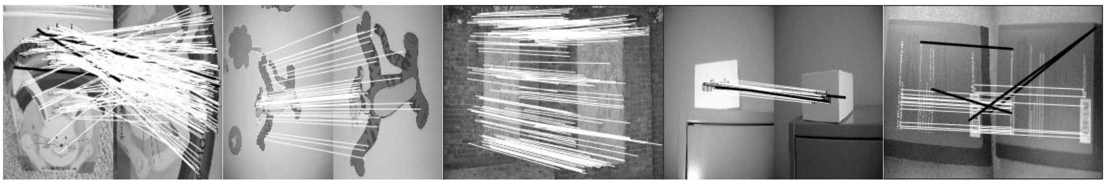


Fig. 4. Examples of matching.

REFERENCES

[1] K. Mikolajczyk, C. Schmid, "Indexing based on scale invariant interest points," In: *Proc. International Conf. on Computer Vision*, vol. 1, pp. 525-531, 2001.

[2] V. Ferrari, T. Tuytelaars, L. Van Gool, "Simultaneous object recognition and segmentation by image exploration," In: *Proc. European Conf. on Computer Vision*, pp.40-54, 2004.

[3] P. Scovanner, S. Ali, M. Shah, "A 3-dimensional SIFT descriptor and its application to action recognition," In: *ACM International Conference on Multimedia*, pp. 357-360, 2007.

[4] S. Lazebnik, C. Schmid, J. Ponce, "A sparse texture representation using affine-invariant regions," In: *IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 2, pp.319-324, 2003.

[5] J. Matas, O. Chum, M. Urban, T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," *Image Vision Computing*, vol. 22, no. 10, pp. 761-767, 2004.

[6] Li, C., Ma, L., "A new framework for feature descriptor based on SIFT," *Pattern Recognition Letters*, vol. 30, no. 5, pp. 544-557, 2009.

[7] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[8] Z.S. Ni, "B-SIFT: a binary SIFT based local image feature descriptor," In: *Proceedings of the Fourth International Conference on Digital Home*, pp. 117–121, November 2012.

[9] Chun-Che Chen, Shang-Lin Hsieh, "Using binarization and hashing for efficient SIFT matching," *Journal of Visual Communication and Image Representation*, vol. 30, pp. 86–93, 2015.

[10] Wengang Zhou, Houqiang Li, "BSIFT: Toward data-independent codebook for large scale image search," *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 967–979, 2015.

[11] Xiaojie Guo, Xiaochun Cao, Jiawan Zhang, Xuewei Li, "MIFT: a mirror reflection invariant feature descriptor," *Proceedings of the Ninth Asian Conference on Computer Vision*, vol. 5995, pp. 536-545, September 2009.

[12] D. Nister, and H. Stewenius, "Scalable recognition with a vocabulary tree," In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recogniton*, vol. 2, pp. 2161–2168, Sep. 2006.

ISBN: 978-988-14047-3-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)