Improving SIFT's Performance by Incorporating Appropriate Gradient Information

Guohua Lv, Md. Tanvir Hossain, Shyh Wei Teng, Guojun Lu
Gippsland School of Information Technology
Monash University
Churchill, Victoria 3842, Australia

Martin Lackmann
Department of Biochemistry & Molecular Biology
Monash University
Clayton, Victoria 3800, Australia

Email: {Guohua.Lv, Tanvir.Hossain, Shyh.Wei.Teng, Guojun.Lu, Martin.Lackmann}@monash.edu

***Abstract*—Scale Invariant Feature Transform (SIFT) has been applied in numerous applications especially in the domain of computer vision. In these applications, image information used for building the SIFT descriptor can have a significant impact on its performance. When building orientation histograms for descriptors, a critical step is how to increment the values in the orientation bins. The original scheme for this step in SIFT was improved in [6]. Two different types of gradient information are used for building orientation histograms. The limitations of the two schemes are identified in this paper and we then propose three new schemes which use both types of gradient information in the feature description and matching stages. Our experimental results show that the proposed schemes can achieve better registration performances than the schemes proposed in SIFT and [6].**

Keywrods - SIFT; Image registration; Keypoint description and matching

I. Introduction

Scale Invariant Feature Transform (SIFT) based image registration techniques [1]–[7] have been widely adopted and improved in recent years because of its effectiveness and robustness. SIFT is a technique for detecting and describing local features in images. Its main merits are summarized as follows. Firstly, SIFT features are invariant to differences in scale, rotation and illumination, and partially invariant to affine transformations, viewpoint changes and noises. Secondly, they are highly distinctive - even a single feature can be correctly matched with high probability against a large data set of features from many images [2]. Moreover, many descriptors can be built to represent a very small portion in an image. This leads to greater robustness in image registration. In addition, local feature detection is computationally efficient - thousands of local features can be extracted from a typical image in near real-time [2]. Finally, SIFT descriptor is extensible, i.e., it can be incorporated with features like color and texture [2].

SIFT descriptor is a 128 dimensional representation of a local image region around a keypoint (a SIFT keypoint is a location in the image that can be repeatedly chosen to represent the local image contents even when the image has undergone various image transformations). A SIFT descriptor is derived by concatenating 16 orientation histograms, whereby each consists of eight bins. In the process of building such a descriptor, a fundamental task is to ensure its distinctiveness. This directly affects its matching accuracy. Thus, it is critical to use appropriate image information for incrementing the values in the orientation bins, in order to appropriately represent the visual differences between images.

Initially, incrementing the values in the orientation bins using gradient magnitudes was proposed in the original SIFT [2]. Later, as identified in [6], using gradient magnitudes alone may lead to similar descriptors for images with completely different visual contents. Thus, a novel scheme which increments the values in the orientation bins based on gradient occurrences was proposed in [6]. However, based on our analysis which we will present in this paper, this scheme still has its limitations. Thus in this paper, we propose three new schemes which use both types of gradient information for building and matching the descriptors. The experimental results presented in this paper will show that our proposed schemes outperform the schemes in the original SIFT and in [6] when registering images.

The rest of the paper is structured as follows. Section II provides an overview of SIFT and the improved scheme proposed in [6]. The limitations of using only gradient magnitudes or gradient occurrences for building and matching descriptors are also illustrated in this section. We then describe our three proposed schemes in Section III. Finally, the performance evaluation of the proposed schemes and conclusions are discussed in Section IV and Section V respectively.

II. Related Work

In this section, the main stages of SIFT are described. Then, an improvement to SIFT proposed in [6] which uses gradient occurrences for building descriptors is described. Finally, the limitations of this improved scheme are discussed.

A. Overview of SIFT

SIFT is a feature detection and description technique [1], [2]. Its main stages can be briefly summarized as follows.

1) Keypoint Detection: To detect local features that are invariant to different scales, Difference-of-Gaussian (DOG) function is applied to the image and locations which are the extrema over multiple scales are selected. Specifically, a DOG image is calculated by

$$\begin{aligned} D(x,y,\sigma) =& L(x,y,k\sigma) - L(x,y,\sigma) \\ =& (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y), \end{aligned} \tag{1}$$

where I denotes the original image, G is the Gaussian kernel, * is convolution operation, σ is variance of the Gaussian filter and k is a constant factor for separating two adjacent scales. In order to detect the extrema, each pixel is compared to its 26 neighbors in 3×3 regions at the current and adjacent scales [2]. If the pixel is the maximum or minimum among the neighboring pixels, it is a keypoint candidate. Each keypoint detected has its own location and scale.

2) Orientation Assignment: By assigning dominant orientation for each keypoint, the keypoint descriptor can be represented relative to this orientation, thereby achieving invariance to image rotation. Firstly, the Gaussian smoothed image, $L(\sigma)$, is selected with the closest scale of the keypoint so that all computations are performed in a scale-invariant way. Then, a local region around the keypoint is determined by a Gaussian-weighted circular window that is derived based on the scale of the keypoint. The gradient magnitude, G_m, and orientation, G_θ, for each pixel, $L(x,y)$, within this region are calculated [2]:

$$\begin{aligned} G_m &= \sqrt{{d_x}^2 + {d_y}^2}, \\ G_\theta &= tan^{-1}(d_y/d_x), \end{aligned} \tag{2}$$

where

$$\begin{aligned} d_x &= L(x+1,y) - L(x-1,y), \\ d_y &= L(x,y+1) - L(x,y-1). \end{aligned} \tag{3}$$

Then, an orientation histogram consisting of 36 bins (covering 360° with an interval of 10°) is built based on the gradient orientations of all the pixels within the local region. The value in each orientation bin is incremented based on the gradient magnitude (weighted by the Gaussian window) of each pixel with a corresponding orientation in the neighboring region. The orientation bin with the highest value denotes the dominant orientation.

3) Keypoint Description: This stage aims to represent each keypoint based on its neighboring region that is determined by a Gaussian-weighted circular window (indicated by a circle in Fig. 1). Firstly, the gradient magnitude and orientation for each pixel located in this region are calculated. Secondly, the gradient orientations are rotated relative to dominant orientation of the keypoint to attain rotation invariance. Then, all the gradients are weighted by the Gaussian window, giving less emphasis to gradients that are further from the center of the region. After that, this region is divided into 4×4 sub-regions and an orientation histogram is built in each sub-region, with eight orientation bins evenly covering 360°(quantized with an interval of 45°). Next, each orientation histogram is built in the same way as the previous stage. Finally, the descriptor is normalized to reduce the effects of illumination changes. Fig. 1 illustrates the process of building a descriptor.

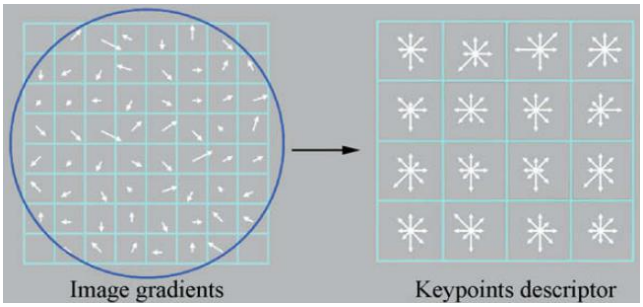


Fig. 1. The formation of SIFT descriptor

4) Keypoint Matching: The descriptors built for the reference and target images are matched using a nearest neighbor-based scheme [2], [8]. Keypoint B is considered a match to keypoint A if B's descriptor D_B is the nearest neighbor of A's descriptor D_A, with a constraint that the distance ratio between the first and the second nearest neighbor is below a threshold.

B. Improved SIFT

As stated in the previous subsection, 16 orientation histograms make up a SIFT descriptor. Each orientation histogram consists of eight orientation bins. The value in each orientation bin is incremented by the gradient magnitude of each pixel with the corresponding orientation. However, using such information for incrementing the values in the orientation bins will result in descriptors which would potentially cause ambiguity during matching under some circumstances. Fig. 2 depicts one such example.

In Fig. 2, each sub-figure denotes a sub-region around a keypoint as described in Section II A and each cell corresponds to a pixel. Each arrow indicates an occurrence of image gradient. The length and orientation of each arrow denote the gradient magnitude and orientation of a particular pixel respectively. A blank cell means that the gradient magnitude of this pixel is zero. Without loss of generality, assume that the orientations of all the arrows in the four regions are 45°. Based on the *Keypoint Description* stage of SIFT discussed in the previous subsection, the sum of gradient magnitudes at the 45° orientation bin for the four regions is 4 and the values in all the other seven orientation bins are zero. Consequently, the same orientation histogram will be built for the four regions.

To deal with this issue, incrementing the values in the orientation bins using gradient occurrences was proposed in [6]. This is done by counting the number of gradients in each sub-region that has orientations corresponding to each orientation bin and incrementing the value in that bin with this number. By applying this scheme in the example in Fig. 2, the value in the 45° orientation bin for the four regions will be 1, 4, 2 and 3 respectively. Thus, four different orientation

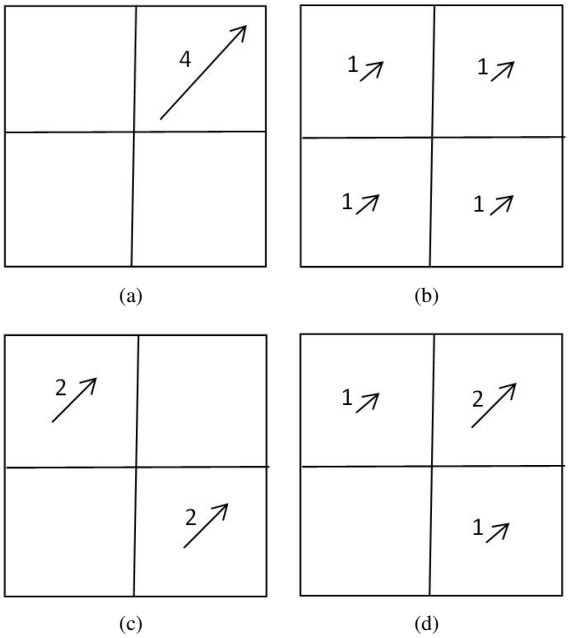


Fig. 2. Ambiguity of incrementing the values in the orientation bins based on gradient magnitudes in SIFT: the four visually different regions have the same orientation histogram.

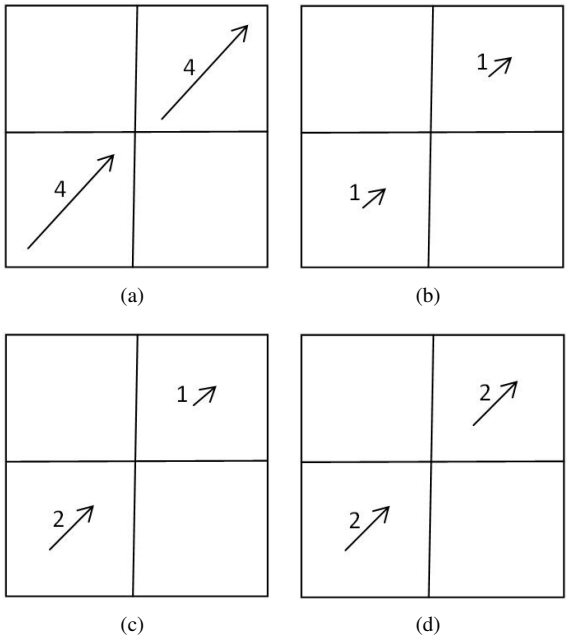


Fig. 3. Ambiguity of incrementing the values in the orientation bins based on gradient occurrences in OG-SIFT: the four visually different regions have the same orientation histogram.

histograms will be built. In [6], another scheme for incrementing the values in the orientation bins was proposed based on average squared difference of magnitudes. In this scheme, the difference between the magnitude of a particular pixel and the average magnitude is considered when incrementing the value in a particular orientation bin. Experimental results in [6] showed that, the scheme of using the number of gradient occurrences performs better than the other scheme based on average squared difference of magnitudes as well as the original scheme used in SIFT. For referencing in this paper, the improved SIFT using gradient occurrences to increment the values in the orientation bins is termed as OG-SIFT (OG: Occurrences of Gradients) in this paper.

C. Limitations of OG-SIFT Descriptors

As stated in Section II, using gradient occurrences for incrementing the values in the orientation bins can successfully distinguish those regions similar to the ones depicted in Fig. 2. However, indistinctiveness of OG-SIFT descriptors can be reflected under some circumstances. A typical example is shown in Fig. 3. All the assumptions in Fig. 3 are consistent with those in Fig. 2. Let us build the orientation histograms for the four regions in Fig. 3 using gradient occurrences to increment the value in each orientation bin. The value in the 45° orientation bin would be consistently equivalent to 2 for the four regions. Thus, the same orientation histogram will be built for them. However, the four regions represent different image contents and have completely different magnitudes. If orientation bins are incremented using gradient magnitudes, the value in the 45° orientation bin would be 8, 2, 3 and 4 respectively, leading to four different orientation histograms.

In conclusion, both SIFT and OG-SIFT fail to distinguish image local regions with different visual contents under some circumstances, like the two examples illustrated in Fig. 2 and Fig. 3. Since gradient magnitudes and gradient occurrences are both important visual properties in images, using both types of information in the description and matching stages will improve the registration performance.

III. Proposed Methods

To deal with the limitations of SIFT and OG-SIFT descriptors stated in the previous section, three schemes that use both gradient magnitudes and gradient occurrences for building and matching descriptors are proposed in this section.

A. Scheme 1

The main idea is to select matches which are visually more similar from the match set of OG-SIFT by also considering the distance of the corresponding SIFT descriptors. The algorithm can be summarized as follows.

1) Perform the *Keypoint Detection* and the *Orientation Assignment* stages of SIFT, as described in Section II A.
2) Build SIFT descriptors and OG-SIFT descriptors, i.e., increment the values in the orientation bins using gradient magnitudes and occurrences respectively.
3) Match the OG-SIFT descriptors between the reference and target images using nearest neighbor-based match-

ing scheme (described in Section II A) to obtain an initial match set, M_{OG}.

4) Reject the matches with relatively large distance of SIFT descriptors in M_{OG} based on a threshold that is predetermined. In our experiments, if the Euclidean distance of the two SIFT descriptors of a particular match in M_{OG} is greater than 0.5 (empirically determined), this match is removed from M_{OG}. The set of removed matches is named M_R. The remaining matches in M_{OG} are determined as more desirable matches, which are visually very similar in terms of both gradient magnitudes and gradient occurrences. They make up the primary match set, named M_P.
5) Match the SIFT descriptors of the reference image corresponding to M_R with all the SIFT descriptors of the target image. Then, the secondary match set, M_S, is obtained. By doing so, the OG-SIFT matches that are relatively far in terms of gradient magnitudes are matched once again using the corresponding SIFT descriptors, potentially leading to new matches.
6) Merge the primary match set, M_P, and the secondary match set, M_S, to constitute the final match set, M_F.

B. Scheme 2

Observing our experimental results, we found that the number of true matches determined by Scheme 1 is much lower as compared to SIFT. To increase the number of true matches, Scheme 2 is proposed. In this scheme, the match set determined by SIFT is taken as the initial match set and then it is refined with the distance of the corresponding OG-SIFT descriptors. The steps from 3) to 5) in Scheme 1 should be modified as follows.

1) Match SIFT descriptors between the reference and target images using nearest neighbor-based matching scheme to obtain an initial match set, M_M.
2) In M_M, a threshold is set to reject those matches with relatively large distance of OG-SIFT descriptors. Those matches with the distance of the corresponding OG-SIFT descriptors smaller than the threshold are determined as more desirable matches, that is, the primary match set, M_P. The remaining matches in the initial match set, M_M, make up the set of removed matches, M_R.
3) Match the OG-SIFT descriptors of the reference image listed in M_R with all the OG-SIFT descriptors of the target image. Then, the secondary match set, M_S, is obtained. This stage is to further match the SIFT matches that are relatively far in terms of gradient occurrences.

C. Scheme 3

The third scheme we propose is to extract the matches determined by both SIFT and OG-SIFT, i.e., the final match set, M_F, can be deduced by

$$M_F = M_M \cap M_{OG}, \tag{4}$$

where M_M and M_{OG} denote the match set obtained using gradient magnitudes and occurrences to increment the values in the orientation bins respectively. The matches obtained using this scheme hold relatively small distance for both SIFT descriptors and OG-SIFT descriptors, reducing the chances of generating false matches caused by cases that are similar to the ones depicted in Fig. 2 and Fig. 3.

In conclusion, Scheme 1 is to prune the match set of OG-SIFT with the assistance of the corresponding SIFT (OG-SIFT) descriptors. In Scheme 2, the match set of SIFT is refined with the corresponding OG-SIFT descriptors. Lastly, Schemes 3 aims at extracting a match set intersected by both SIFT and OG-SIFT. For referencing purpose, the three schemes are termed as OG-SIFT-M, M-SIFT-OG and MOG-SIFT respectively (OG: Occurrences of Gradients and M: Magnitude).

IV. Performance Evaluation

In our experiments, the test data set used consists of 40 pairs of natural images [8], [10] which are commonly used in the literature for evaluating the performance of local descriptors, and 56 pairs of medical microscopic images. Performance comparisons have been carried out on five schemes: SIFT, OG-SIFT, OG-SIFT-M (Scheme1), M-SIFT-OG (Scheme 2) and MOG-SIFT (Scheme 3).

A. Evaluation Criterion

The accuracy of an image registration technique depends to a high degree on the accuracy of the match set. The higher the percentage of true matches, the more accurate the final registration will be. Therefore, we evaluate our proposed schemes using the accuracy of the match set where

$$accuracy = \frac{\#true\ matches}{\#total\ matches} \times 100\%. \tag{5}$$

The ground truths for natural images are given in [10] and the homography for medical microscopic images is estimated using a sample of corresponding points in each image pair identified manually. A pixel error of 4, which is consistent with [9], is used when determining a true match.

B. Experimental Results

Tables I and II compare the five schemes in terms of accuracy of the match set as well as the number of true matches and false matches. For the natural images stated in Table I, each accuracy is obtained by registering each original image with its five transformed images, whereby each transformed image is derived by applying increasing amount of the transformation of the types stated in Column T of Table I. For the medical microscopic images stated in Table II, the rotation between an original image and a transformed image ranges from 0° to 90° with an interval of 15°. Thus, each accuracy in this table is an average accuracy derived by registering seven image pairs. The experimental results in these two tables show that

TABLE I
EXPERIMENTAL RESULTS FOR NATURAL IMAGES

Image	T[a]	Scheme	#True[b]	#False[b]	Accuracy(%)
bark	scale + rotation	SIFT	369	3712	9.04
		OG-SIFT	350	3552	8.97
		OG-SIFT-M	346	3386	9.27
		M-SIFT-OG	366	3550	9.35
		MOG-SIFT	340	3254	9.46
boat	scale + rotation	SIFT	6796	1203	84.96
		OG-SIFT	6553	653	90.94
		OG-SIFT-M	6484	490	92.97
		M-SIFT-OG	6718	684	90.76
		MOG-SIFT	6250	311	95.26
graffiti	viewpoint	SIFT	876	2413	26.63
		OG-SIFT	878	1863	32.03
		OG-SIFT-M	866	1706	33.67
		M-SIFT-OG	868	2065	29.59
		MOG-SIFT	832	1419	36.96
wall	viewpoint	SIFT	1650	15633	9.55
		OG-SIFT	1593	14959	9.62
		OG-SIFT-M	1587	14758	9.71
		M-SIFT-OG	1643	15473	9.60
		MOG-SIFT	1553	14329	9.78
bikes	blur	SIFT	3986	1621	71.09
		OG-SIFT	3842	1060	78.38
		OG-SIFT-M	3775	788	82.73
		M-SIFT-OG	3892	973	80.00
		MOG-SIFT	3722	583	86.46
trees	blur	SIFT	2870	3831	42.83
		OG-SIFT	2635	3468	43.18
		OG-SIFT-M	2592	3208	44.69
		M-SIFT-OG	2840	3552	44.43
		MOG-SIFT	2429	2936	45.27
leuven	illumination	SIFT	3608	3589	50.13
		OG-SIFT	3719	3408	52.18
		OG-SIFT-M	3685	3282	52.89
		M-SIFT-OG	3596	3356	51.73
		MOG-SIFT	3535	3068	53.54
ubc	JPEG compression	SIFT	9865	695	93.42
		OG-SIFT	9216	317	96.67
		OG-SIFT-M	9163	248	97.36
		M-SIFT-OG	9672	426	95.78
		MOG-SIFT	8954	161	98.23

[a] Column T denotes transformations between image pairs.
[b] Column #True and column #False indicate the number of true matches and false matches separately.

TABLE II
EXPERIMENTAL RESULTS FOR MEDICAL MICROSCOPIC IMAGES

Specimen[a]	Δ_σ^b	Scheme	#True	#False	Accuracy(%)
A	$2x$	SIFT	160	230	41.03
		OG-SIFT	117	82	58.79
		OG-SIFT-M	114	21	84.44
		M-SIFT-OG	144	79	64.57
		MOG-SIFT	110	15	88.00
B	$2x$	SIFT	3442	1015	77.23
		OG-SIFT	3351	529	86.37
		OG-SIFT-M	3333	401	89.26
		M-SIFT-OG	3414	617	84.69
		MOG-SIFT	3245	144	95.75
C	$2x$	SIFT	1324	1121	54.15
		OG-SIFT	1277	515	71.26
		OG-SIFT-M	1274	377	77.17
		M-SIFT-OG	1306	754	63.40
		MOG-SIFT	1230	133	90.24
D	$2x$	SIFT	4996	1518	76.70
		OG-SIFT	4798	764	86.26
		OG-SIFT-M	4786	636	88.27
		M-SIFT-OG	4984	1067	82.37
		MOG-SIFT	4713	411	91.98
E	$2x$	SIFT	1034	239	81.23
		OG-SIFT	929	93	90.90
		OG-SIFT-M	874	49	94.69
		M-SIFT-OG	971	104	90.33
		MOG-SIFT	860	36	95.98
F	$2x$	SIFT	898	274	76.62
		OG-SIFT	777	162	82.75
		OG-SIFT-M	760	110	87.36
		M-SIFT-OG	857	164	83.94
		MOG-SIFT	742	82	90.05
A	$4x$	SIFT	258	184	58.37
		OG-SIFT	173	72	70.61
		OG-SIFT-M	169	14	92.35
		M-SIFT-OG	221	64	77.54
		MOG-SIFT	172	13	92.97
D	$4x$	SIFT	877	1467	37.41
		OG-SIFT	841	626	57.33
		OG-SIFT-M	830	388	68.14
		M-SIFT-OG	863	667	56.41
		MOG-SIFT	804	192	80.72

[a] The first column denotes labels of the original images.
[b] The second column gives the scale difference between image pairs.

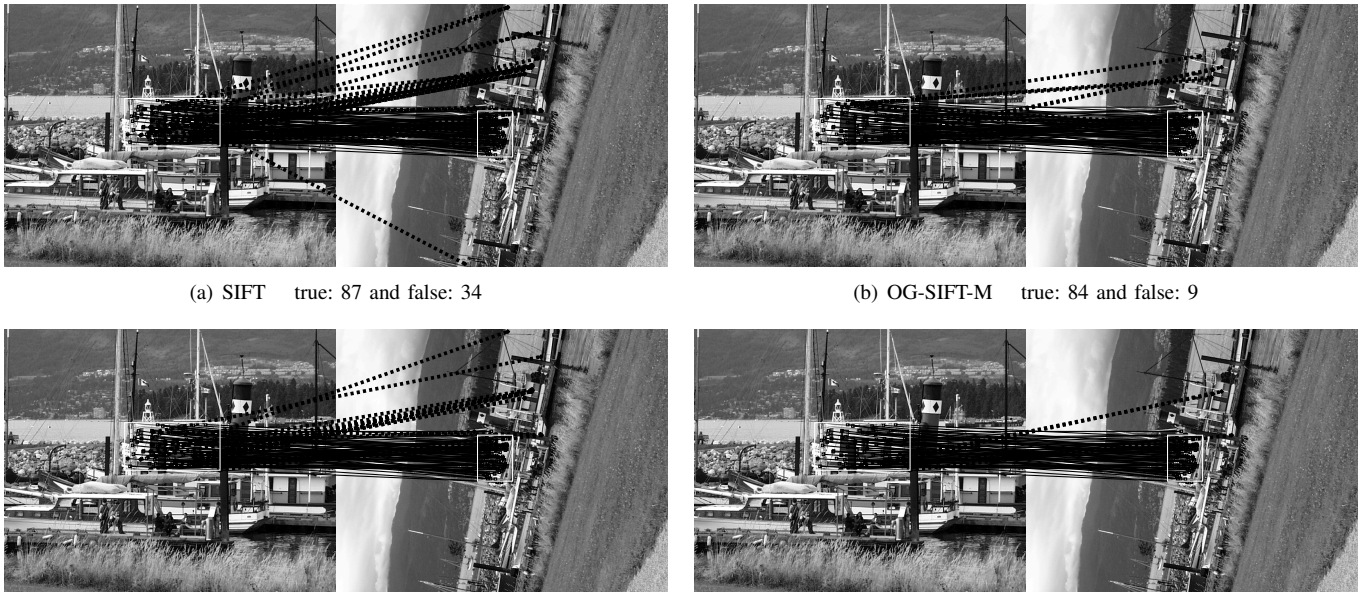
TABLE III
COMPARISON OF FIVE SCHEMES WITH THE TOTAL NUMBER OF TRUE MATCHES AND AVERAGE ACCURACY

Scheme	# Total True Matches	Average Accuracy (%)
SIFT	43009	55.65
OG-SIFT	41049	63.52
OG-SIFT-M	40638	69.06
M-SIFT-OG	42355	63.41
MOG-SIFT	39491	72.54

compared to SIFT and OG-SIFT, OG-SIFT-M and MOG-SIFT consistently achieve higher matching accuracy for all the natural and medical microscopic images. As for M-SIFT-OG, it also consistently achieves higher matching accuracy than SIFT. Compared to OG-SIFT, M-SIFT-OG's matching accuracy is higher for 37.5% of the test images and is slightly lower for the remaining test images. However, its number of true matches is higher for 87.5% of the test images.

The total number of true matches and average accuracies for all the 96 image pairs are presented in Table III. Compared to OG-SIFT, the accuracy of the match set increases by 5.54% and 9.02% on average respectively when OG-SIFT-M and MOG-SIFT are used. Compared to SIFT, the average accuracy obtained using OG-SIFT-M and MOG-SIFT improves by 13.41% and 16.89% separately. Meanwhile, the number of true matches decreases. This is because fewer matches are determined as the matching criterion is stricter by using both gradient magnitudes and gradient occurrences. In the process of refining an initial match set, some true matches are potentially lost. Besides, it should be noted that, M-SIFT-OG can achieve the accuracy approximately as high as OG-SIFT (an overall gap of 0.11%), yet can produce many more true matches than OG-SIFT. In contrast to SIFT, M-SIFT-OG increases the accuracy by 7.76%, and the number of true matches only decreases by 1.52%.

To summarize, among the proposed three schemes, OG-SIFT-M and MOG-SIFT can achieve higher accuracy of the match set than both SIFT and OG-SIFT. To further illustrate



(a) SIFT true: 87 and false: 34

(b) OG-SIFT-M true: 84 and false: 9

(c) M-SIFT-OG true: 85 and false: 18

(d) MOG-SIFT true: 81 and false: 4

Fig. 4. A matching example for a *boat* image pair. The two regions bounded by a rectangular are corresponding portions in the image pair. For a better comparison of the match set, those matches with keypoints in the reference image (left) beyond the bounded region are not plotted here. Solid and dashed lines denote true and false matches respectively.

this, Fig. 4 clearly presents a comparison between the number of true matches and the number of false matches using a specific image pair. Through observing the experimental results, we recommend OG-SIFT-M and MOG-SIFT be applied in applications which require high accuracy. If besides high accuracy, the application also requires high number of true matches, then M-SIFT-OG will be an ideal alternative as it offers a compromise between high accuracy and high number of true matches.

V. Conclusions

Two different types of gradient information are used for building and matching descriptors in the original SIFT and OG-SIFT [6]. In this paper, we have presented three new schemes which make use of both types of gradient information in the keypoint description and matching stages, which can address the limitations associated with incrementing the values in the orientation bins when building descriptors in SIFT and OG-SIFT. Our experimental results clearly show that the proposed schemes can achieve better registration performances as compared to SIFT and OG-SIFT. Our proposed new descriptors, therefore, can be used in all applications where SIFT is appropriate to improve accuracy.

References

[1] D.G. Lowe, *Object recognition from local scale-invariant features*, The Proc. of the Seventh IEEE International Conf. on Computer Vision, vol. 2, pp. 1150-1157, 1999.

[2] D.G. Lowe, *Distinctive Image Features from Scale-Invariant Keypoints*, International Journal of Computer Vision, vol. 2, no. 60, pp. 91-110, 2004.

[3] Y. Ke and R. Sukthankar, *PCA-SIFT: A More Distinctive Representation for Local Image Descriptors*, Proc. Conf. Computer Vision and Pattern Recognition, vol. 2, pp. 506-513, 2004.

[4] J. Chen and J. Tian, *Real-time multi-modal rigid registration based on a novel symmetric-SIFT descriptor*, Progress in Natural Science, vol. 19, pp. 634-651, 2009.

[5] M. Toews and W. Wells, *SIFT-Rank: Ordinal Description for Invariant Feature Correspondence*, Proc. Conf. Computer Vision and Pattern Recognition, pp. 172-177, 2009.

[6] M.T. Hossain, Shyh Wei Teng, Guojun Lu, and M. Lackmann, *An Enhancement to SIFT-Based Techniques for Image Registration*, International Conf. on Digital Image Computing: Techniques and Applications, pp. 166-171, 2010.

[7] R. Ma, J. Chen, and Z. Su, *MI-SIFT: Mirror and Inversion Invariant Generalization for SIFT Descriptor*, Proc. of the ACM International Conf. on Image and Video Retrieval, pp. 228-235, 2010.

[8] K. Mikolajczyk and C. Schmid, *A Performance Evaluation of Local Descriptors*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, pp. 1615-1630, 2005.

[9] G. Yang, C. Stewart, M. Sofka, and C. Tsai, *Registration of Challenging Image Pairs: Initialization, Estimation, and Decision*, IEEE transactions on Pattern Analysis and Machine Intelligence, pp. 1973-1989, 2007.

[10] Available from: http://www.robots.ox.ac.uk/ vgg/research/affine.