Proceeding of IC-ITS 2017 e-ISBN:978-967-2122-04-3
3th International Conference on Information Technology & Society
July 31st & August 1st, 2017, PENANG, MALAYSIA

Histogram-based Threshold Selection of Retinal Feature for Image Registration

Roziana Ramli[1], Mohd Yamani Idna Idris[1]*, Khairunnisa Hasikin[2] & Noor Khairiah A. Karim[3]

[1]*Department of Computer System & Technology, Faculty of Computer Science & Information Technology, University of Malaya*
[2]*Department of Biomedical Engineering, Faculty of Engineering, University of Malaya*
[3]*Regenerative Medicine Cluster/Imaging Unit, Advanced Medical & Dental Institute, Universiti Sains Malaysia*

*yamani@um.edu.my**

Abstract. Retinal image registration is performed to align two or more retinal images for super-resolution, image mosaicking and longitudinal study applications to assist diagnosis and monitoring retinal diseases. In fundus retina image, local feature such as Scale Invariant Feature Transform (SIFT) with illumination invariant Difference of Gaussian (*ii*DoG) operator is stable and capable of detecting keypoints from non-uniform image. However, *ii*DoG-SIFT detects keypoints along the vessels and on the background. Estimating geometrical transformation based on background keypoints can lead to inaccurate registration. Therefore, we introduce a histogram-based threshold selection computed from grayscale pixels of 9x9 patch to select keypoints along the vessels. The threshold selection is determined based on standard deviation of the histogram to exclude keypoints on the background while preserving keypoints along the vessels. The proposed method is tested in registering 134 retinal image pairs from Fundus Image Registration Dataset (FIRE) that consists of super-resolution, image mosaicking and longitudinal study applications. There are two main aspects evaluated during the experiment namely, success rate (%) and target registration error (TRE) of the successful registration. The experimental results show that the proposed method successfully registered 100% of retinal images in super-resolution application with mean TRE of 1.796 pixels, 86% in longitudinal study application with mean TRE of 5.161 pixels and 45% in image mosaicking application with mean TRE of 12.565 pixels. Furthermore, the comparative experiment is performed with GDB-ICP, Harris-PIIFD and H-M. The success rate and percentage of accurate registration of the proposed method outperformed others in super-resolution and longitudinal applications.

Keywords: Feature-based registration, fundus retinal image, local feature.

INTRODUCTION

In ophthalmology, retinal image registration is performed to assist diagnosis and monitoring the progression of retinal disease [1]. The image registration process includes the alignment of moving image to the orientation of fixed image through geometrical transformation. The geometrical transformation can be estimated

according to intensity patterns or feature information between images. Intensity-based registration searches for the similarity of intensity patterns between images whereas feature-based registration establishes the corresponding features based on anatomical or local features to estimate the geometrical transformation.

The intensity-based registration utilizes the similarity metrics such as mutual information [2] and phase correlation [3] to measure the similarity of intensity patterns between fixed and moving images. This process is repeated during the optimization process to refine the geometrical transformation until the maximum number of iteration or optimum registration is achieved. The intensity-based registration considers all the intensity information in the image during the optimization process. Consequently, the intensity-based registration is susceptible in registering retinal image pairs with significant difference of anatomical appearance such as variation of cotton-wool and microaneurysms between images. These variations commonly observed in longitudinal study application in which, retinal image pairs acquired from different period are registered to monitor the progression of retinal disease such as glaucoma and age-related macular degeneration [1].

The feature-based registration using local feature detects keypoints based on maxima and minima points such as Harris corner [4], Scale Invariant Feature Transform (SIFT) [5] and Speeded-Up Robust Features (SURF) [6, 7] on fixed and moving images. Then, the corresponding keypoints between the image pairs are established during matching process. These corresponding keypoints or matches can be further refined to obtain inliers using Random Sample Consensus (RANSAC) [8] or M-estimator SAmple Consensus (MSAC) [9] algorithms. Finally, the inliers are used to estimate the geometrical transformation between image pairs.

In feature-based retinal image registration, SIFT has been used in Generalized Dual-bootstrap Iterative Closest Point (GDB-ICP) [10], Uniform Robust Scale Invariant Feature Transform with Partial Intensity Invariant Feature descriptor (URSIFT-PIIFD) [11] and URSIFT with Frangi's Vesselness Measure selection criterion [12]. SIFT with illumination invariant Difference of Gaussian (*ii*DoG) operator detects keypoints on non-uniform illumination image by targeting the underexposed region. This characteristic is beneficial in detecting keypoints on fundus retinal images as fundus retinal image typically suffers from non-uniform illumination. Acquiring fundus retinal image with a perfect illumination is a challenging task that requires experience of the ophthalmologist and cooperation from the patient. However, *ii*DoG-SIFT detects keypoints on both background and along the vessels in retinal image. The estimation of geometrical transformation according to the keypoints on the background can results in inaccurate registration. Therefore, we propose a histogram-based threshold selection to exclude the keypoints detected on the background and preserved the keypoints detected along the vessels.

METHODOLOGY

Feature-based registration can be divided into four main modules namely, feature detection, feature descriptor, matching and estimating geometrical transformation. Further details of these modules are described as follows.

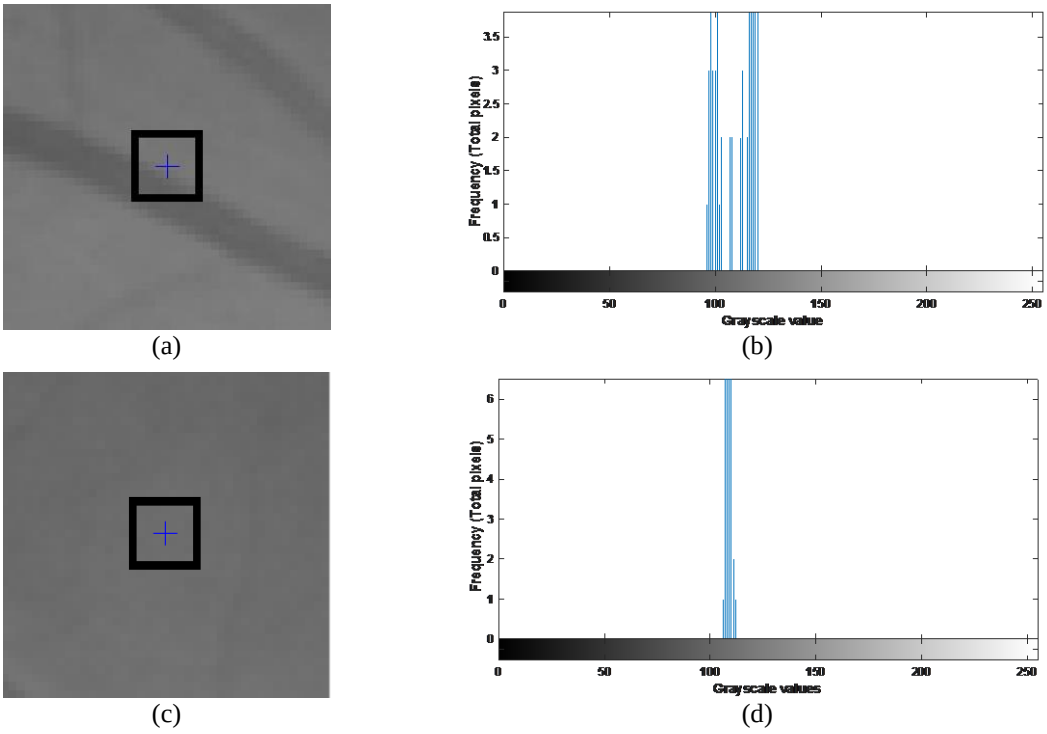
Feature Detection

In this study, the local features are detected based on *ii*DoG-SIFT through two main sub-modules; detection and selection. These sub-modules are important to extract the keypoints and distinguished it from the noise.

Detection sub-module: This sub-module begins with the representation of image $I(x, y)$ in *ii*DoG scale space that consists of normalized difference of Gaussian (*n*DoG) and DoG operators. Vonikakis et al. combined these two operators based on the piecewise function to improve the feature detection in image with non-uniform illumination.

This operation will result in hierarchical *ii*DoG scale space that holds five levels and four octaves. The local extrema in *ii*DoG scale space is detected by comparing each point to 26 points of its three-dimensional neighbors. The sub-position for an extremum is estimated by fitting a 3D quadratic function based on second order Taylor expansion to its neighborhood.

Selection sub-module: After the position of the extremum is refined, the keypoints on retinal image are extracted by rejecting the unstable extremum that is low in contrast with *ii*DoG magnitude less than $T_C = 0.03$ and along the edge. However, *ii*DoG-SIFT detects keypoints along the vessels and on the background. Estimating the geometrical transformation based on the background keypoints can be inaccurate, thus, we propose a histogram-based threshold selection to exclude the background keypoints described as follows.



(a) (b) (c) (d)

FIGURE 1. Examples of (a) vessel and (c) background features detected on fundus retinal image. Their respective histogram in 9x9 patch shown by black rectangle are given in (b) for vessel and (d) for background features.

The histogram for each keypoints is computed from a patch of grayscale pixels with 9x9 dimensions representing the distribution of grayscale values in 256 bins. The histogram of keypoint along the vessel will yield a wider distribution of grayscale value whereas the histogram of keypoint on the background will yield a narrow distribution of grayscale value as shown in Fig. 1(b) and Fig. 1(d) respectively. Therefore, we utilize the distribution of grayscale value in histogram as threshold selection to exclude keypoints on the background.

In this study, we compute standard deviation of all 256 bins in the histogram to describe the distribution of grayscale pixels in a patch. A smaller standard deviation of the histogram indicating the keypoints is along the vessel while a larger standard deviation indicating the keypoints is on the background. Through observations, we find that the keypoints on the background yields standard deviation value larger than 2, thus, we initially exclude all these keypoints. Then, we determine the suitable threshold value between 0.1 to 2.0 with an increasing step of 0.1 by performing registration on 10 randomly selected retinal image pairs. The average target registration error of 10 retinal image pairs relative to the threshold based on standard deviation of histogram is depicted in Fig. 2. According to Fig. 2, the lowest average target registration error is obtained when the standard deviation is 1.8, therefore, we choose this value as threshold selection to exclude keypoints on the background.

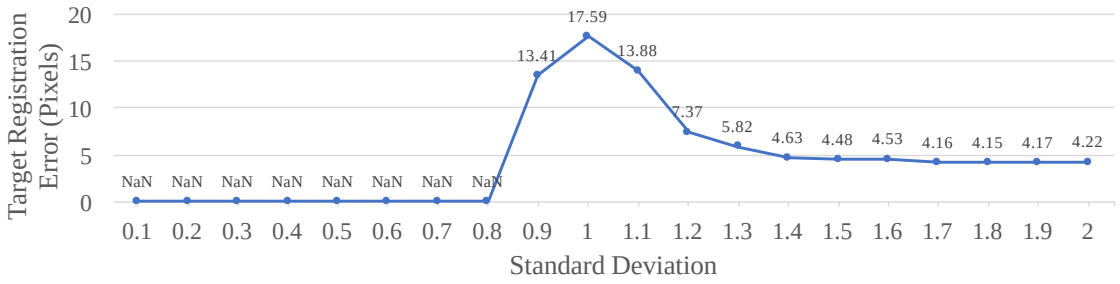


FIGURE 2. Average of target registration error relative to threshold based on standard deviation of histogram.

Feature Descriptor, Matching and Geometrical Transformation

Each keypoint along the vessel is characterized by Histogram of Oriented Gradients (HOG) descriptor [13] with cell size of 8-by-8 pixels and block size of 2-by-2 cells. Then, these descriptors are matched between image pair using nearest neighbour search with ratio threshold of 1.0. The outliers from this process are further excluded using MSAC algorithm [9] to obtain the inliers. The maximum distance between corresponding points in fixed and moving for MSAC is set to 13 values; 2, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110 and 120. Each maximum distance is repeated 50 times to ensure convergence due to randomized nature of MSAC algorithm. The similarity transformation that consists of rotation, translation and scaling is estimated based on these inliers wherein the best registration performance from this repetition is chosen as the final registration.

RESULTS & DISCUSSION

Design and experimental of the proposed method are implemented in Matlab running on Intel(R) Core(TM) i7-4770 CPU@3.40GHz 8.00GB RAM and tested on FIRE: Fundus Image Registration Dataset [14]. We compare the registration performance of the proposed method with GDB-ICP [10], Harris-PIIFD [15] and H-M [16] obtained from [14]. To reduce the processing cost, the proposed method processed the retinal image pairs in 583×583 resolution but evaluated in original resolution of 2912×2912 as GDB-ICP [10], Harris-PIIFD [15] and H-M [16] are evaluated according to original resolution.

FIRE dataset consists of 134 retinal image pairs with ground truth annotation and classified into super-resolution (71 image pairs), image mosaicking (49 image pairs) and longitudinal study (14 image pairs) applications. The purpose of super-resolution application is to combine multiple retinal images to improve pathological information particularly in the case of motion artefact whereas image mosaicking application combines multiple retinal image to create a wider view of retinal. The longitudinal study application usually performed to monitor the progression of retinal disease over time.

Super-resolution and longitudinal study applications consist of retinal image pairs with large overlap area while retinal image pairs in image mosaicking application has smaller overlap area. Pathological cases are observed in the dataset but anatomical appearance is unchanged in super-resolution and image mosaicking applications while the changes of anatomical appearance can be seen in longitudinal study application in terms of vessel tortuosity, microaneurysms and cotton-wool.

Evaluation Metrics

Registration performance of the proposed method and state-of-the-arts in FIRE dataset are evaluated according to target registration error (TRE) wherein a small TRE value representing a more accurate registration and vice versa. TRE measured in pixel is an average distance between 10 corresponding landmarks identify by experts between fixed and moving images. These landmarks are provided by FIRE dataset.

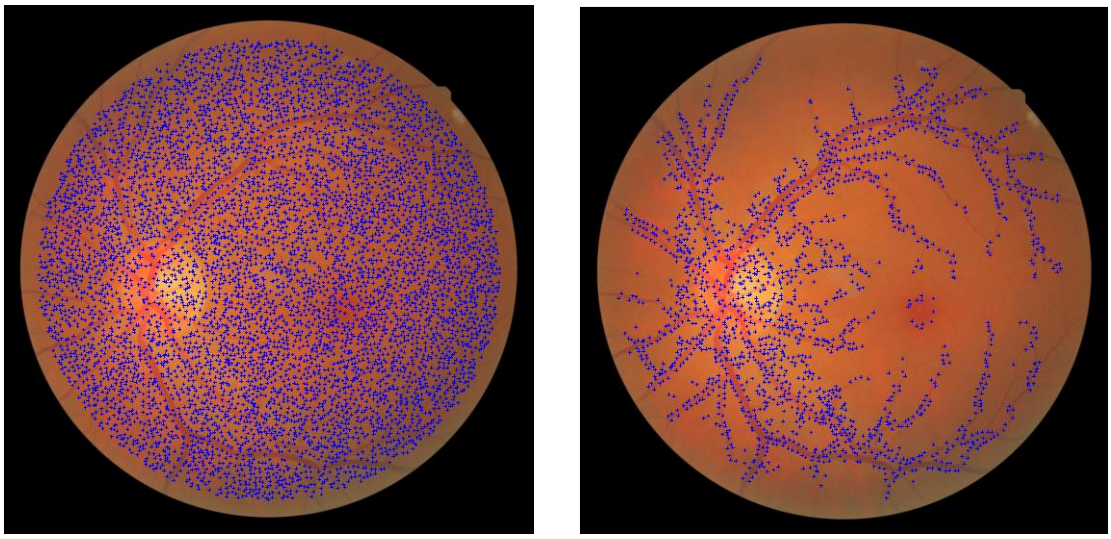
The registration that results in TRE value equal or less than 5 pixels is classified as an accurate registration while TRE value larger than 5 pixels but equal or less than 15 pixels is classified as acceptable registration. The registration with acceptable performance may requires an intervention from ophthalmologist by performing manual registration to recover the remaining distance and enable its utilization for clinical applications. The registration with TRE value larger than 15 pixels is too large and difficult for manual registration, thus, classified as failed.

The overview of successful registration in the dataset is represented by success rate (%). Success rate is ratio of total pairs with successful registration (TRE $\leq$ 15 pixels) to the total of image pairs in the dataset or application. Another aspect measured in this study is the registration consistency of the method between image pairs. The consistency is represented by the standard deviation of all TRE including the successful and failed registrations. A method with a more consistent performance between image pairs will yields a smaller value of standard deviation and vice versa.

Registration Performance

Examples of keypoints detected on fundus retinal image with and without the proposed histogram-based threshold selection are depicted in Fig. 3. Implementation of the proposed histogram-based threshold excludes the keypoints detected on the background while preserving the keypoints along the vessels. This enable the proposed method to accurately register 55% of the image pairs in FIRE dataset and 23% with acceptable registration as summarized in Table 1. Moreover, the highest registration performance of the proposed method is observed in super-resolution followed by longitudinal study and image mosaicking applications. Examples of keypoints detected on image pairs and their registered image for each application are shown in Fig. 4. Overall, the percentage of accurate registration of the proposed method is ranked in second after GDB-ICP but our proposed method obtained smaller percentage of failed registration with 22% compared to GDB-ICP with 39%.

In super-resolution application, the proposed method, H-M and Harris-PIIFD successfully registered all the image pairs whereas GBD-ICP failed to register 39% of the image pairs. The proposed method accurately registered 97% of the retinal image pairs in super-resolution application and the highest among all compared to H-M (96%), Harris-PIIFD (86%) and GDB-ICP (85%). Furthermore, the proposed method attained 1.796 pixels of the mean TRE for successful registration while H-M (1.949 pixels), Harris-PIIFD (2.981 pixels) and GDB-ICP (1.426 pixels).



(a) (b)

FIGURE 3. Example of iiDoG-SIFT (a) without and (b) with the proposed histogram-based threshold selection.

The proposed method successfully registered 86% of the image pairs in longitudinal study application and outperformed H-M (71%), Harris-PIIFD (50%) and GDB-ICP (64%). Specifically, 36% of the successful registration of the proposed method is an accurate registration and 50% of the successful registration is within the acceptable range of registration performance. The proposed method failed to register two of the image

pairs in longitudinal study application that exhibits the variation of vessel tortuosity and vessel thickness between image pairs with TRE of 23.952 pixels and 17.346 pixels respectively. The failed registration in these images are mainly due to lack of distribution and total inliers between image pairs as HOG descriptor employed in the proposed method is sensitive to the changes of structure information. Similarly, H-M and Harris-PIIFD failed to register these image pairs with higher TRE values while GDB-ICP unable to estimate any geometrical transformation.

Retinal image pairs in image mosaicking application are more challenging due to small overlapping area and degradation of image quality such as dark spot artifact covering the underlying tissues. Consequently, all successful registration of the proposed method (45%) in image mosaicking application is only within the acceptable range of registration. A similar performance in image mosaicking application is observed in the successful registration of H-M (61%) and Harris-PIIFD (10%). Conversely, GDB-ICP accurately registered 33% of the image pairs in image mosaicking application and 2% with acceptable performance.

TABLE (1). Details of registration performance in FIRE dataset for the proposed method, H-M, Harris-PIIFD and GDB-ICP.

	Proposed	**H-M (2016)**	**Harris-PIIFD (2010)**	**GDB-ICP (2007)**
Overall				
Accurate	55%	54%	48%	60%
Acceptable	23%	29%	14%	1%
Success Rate	78%	83%	62%	61%
Failed	22%	17%	38%	39%
Mean TRE (Success)	4.437	4.728	3.854	1.983
Standard Deviation TRE	7.236	39.918	365.754	1.479
Super-resolution				
Accurate	97%	96%	86%	85%
Acceptable	3%	4%	14%	0%
Success Rate	100%	100%	100%	85%
Failed	0%	0%	0%	15%
Mean TRE (Success)	1.796	1.949	2.981	1.426
Standard Deviation TRE	1.264	1.886	1.969	0.778
Image Mosaicking				
Accurate	0%	0%	0%	33%
Acceptable	45%	61%	10%	2%
Success Rate	45%	61%	10%	35%
Failed	55%	39%	90%	65%
Mean TRE (Success)	12.565	11.054	13.411	3.260
Standard Deviation TRE	3.971	18.980	580.486	1.134
Longitudinal Study				
Accurate	36%	29%	21%	29%
Acceptable	50%	43%	29%	7%
Success Rate	86%	71%	50%	36%
Failed	14%	29%	50%	64%
Mean TRE (Success)	5.161	5.482	5.884	4.317
Standard Deviation TRE	6.000	110.853	396.753	3.444

Overall consistency of registration performance between image pairs represented by standard deviation of all TRE in FIRE dataset shows that GDB-ICP is more consistent

than others with 1.479 pixels followed by the proposed method (7.236 pixels), H-M (39.918 pixels) and Harris-PIIFD (365.754 pixels). Among the applications, performance of the proposed method is more consistent in super-resolution application with standard deviation of 1.264 pixels compared to its performance in image mosaicking (3.971 pixels) and longitudinal study (6.000 pixels) applications.

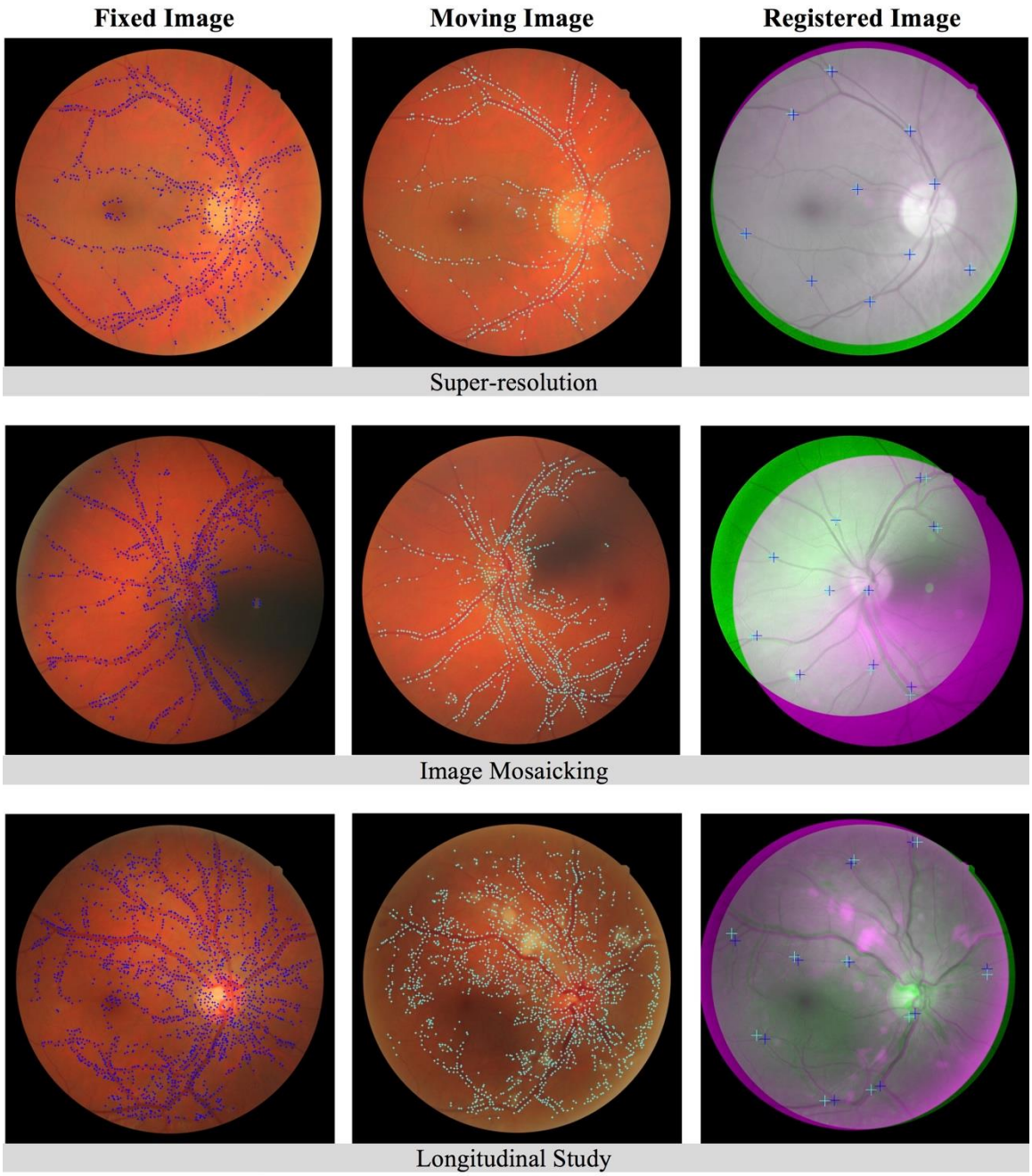


FIGURE 4. Examples of keypoints detected by the proposed method in fixed (column 1) and moving (column 2) images for super-resolution, image mosaicking and longitudinal study applications. Column 3 shows the registered images with 10 corresponding landmarks between fixed and moving images.

CONCLUSION

In this paper, we propose histogram-based threshold selection to exclude *ii*DoG-SIFT keypoints that detected on the background while preserving features detected along the vessels for retinal image registration. For each keypoint, the histogram of grayscale pixels is computed from a patch with 9x9 dimensions in which, a wider distribution of grayscale value representing keypoints along the vessels and a narrow distribution representing keypoints on the background. The standard deviation of the histogram is used to describe the distribution of the grayscale value in 256 bins and as threshold section to exclude keypoints on the background. Then, keypoints along the vessels are characterized by HOG descriptors and matched between image pairs using nearest neighbour search. The outliers from the initial matches are removed using MSAC algorithm to obtain inliers for estimation of geometrical transformation.

The proposed method is tested on FIRE dataset consists of 134 fundus retinal image pairs that divided into super-resolution, image mosaicking and longitudinal study applications. In FIRE dataset, the proposed method registered 55% of the retinal image pairs with accurate performance (TRE ≤ 5 pixels) and 23% with acceptable performance (5 pixels < TRE ≤ 15 pixels) which constitute for 78% of the retinal image pairs in the dataset. In super-resolution application, the proposed method successfully registered all retinal image pairs and attained the highest percentage of accurate registration with 97% while H-M, Harris-PIIFD and GDB-ICP attained 96%, 86% and 85% respectively. This demonstrates that the proposed method is more accurate in registering retinal image pairs in super-resolution application consisting large overlapping area and minimal degradation of image quality.

In longitudinal study application, the proposed method attained the highest success rate (86%) compared to H-M (71%), Harris-PIIFD (50%) and GDB-ICP (36%). Furthermore, the proposed method accurately registered 36% of the image pairs in longitudinal study application and 50% of the successful registrations are within the acceptable range of registration performance. However, the proposed method is susceptible in the presence of vessels variation in terms of tortuosity and thickness between retinal image pairs.

Among the applications in FIRE dataset, the proposed method recorded the lowest registration success rate (45%) in image mosaicking application consisting smaller overlap area between image pairs and black spot artifact. Therefore, in future we will focus on improving the proposed method in registering retinal image pairs with smaller overlap.

ACKNOWLEDGMENTS

This work was supported by University Malaya Postgraduate Research Grant (PG039-2015B).

REFERENCES

1. Adal, K. M., van Etten, P. G., Martinez, J. P., van Vliet, L. J. & Vermeer, K. A. (2015) Accuracy Assessment of Intra- and Intervisit Fundus Image Registration for

Diabetic Retinopathy ScreeningAccuracy Assessment of Fundus Image Registration, *Invest Ophth Vis Sci.* **56**, 1805-1812.

2. Legg, P. A., Rosin, P. L., Marshall, D. & Morgan, J. E. Improving accuracy and efficiency of mutual information for multi-modal retinal image registration using adaptive probability density estimation, *Comput Med Imag Grap.* **37**, 597-606.
3. Kolar, R., Sikula, V. & Base, M. (2010) Retinal Image Registration using Phase Correlation, *Biosig Brno*, 244-252.
4. Harris, C. & Stephens, M. (1988). A combined corner and edge detector. Paper presented at the *Alvey Vision Conference*.
5. Lowe, D. G. (2004) Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision.* **60**, 91-110.
6. Bay, H., Tuytelaars, T. & Van Gool, L. (2006). SURF: Speeded Up Robust Features. Paper presented at the *Computer Vision – ECCV 2006: 9th European Conference on Computer Vision*, Graz, Austria.
7. Bay, H., Ess, A., Tuytelaars, T. & Van Gool, L. (2008) Speeded-Up Robust Features (SURF), *Computer Vision and Image Understanding.* **110**, 346-359.
8. Fischler, M. A. & Bolles, R. C. (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Communications of the ACM.* **24**, 381-395.
9. Torr, P. H. & Zisserman, A. (2000) MLESAC: A new robust estimator with application to estimating image geometry, *Computer Vision and Image Understanding.* **78**, 138-156.
10. Yang, G., Stewart, C. V., Sofka, M. & Tsai, C.-L. (2007) Registration of challenging image pairs: Initialization, estimation, and decision, *IEEE Transactions on Pattern Analysis and Machine Intelligence.* **29**.
11. Ghassabi, Z., Shanbehzadeh, J., Sedaghat, A. & Fatemizadeh, E. (2013) An efficient approach for robust multimodal retinal image registration based on UR-SIFT features and PIIFD descriptors, *Eurasip J Image Vide*.
12. Ghassabi, Z., Shanbehzadeh, J., Mohammadzadeh, A. & Ostadzadeh, S. S. (2015) Colour retinal fundus image registration by selecting stable extremum points in the scale-invariant feature transform detector, *IET Image Processing.* **9**, 889-900.
13. Dalal, N. & Triggs, B. (2005). Histograms of oriented gradients for human detection. Paper presented at the *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*
14. Hernandez-Matas, C., Zabulis, X., Triantafyllou, A., Anyfanti, P., Douma, S. & Argyros, A. A. (2016) FIRE: Fundus Image Registration Dataset.
15. Chen, J., Tian, J., Lee, N., Zheng, J., Smith, R. T. & Laine, A. F. (2010) A Partial Intensity Invariant Feature Descriptor for Multimodal Retinal Image Registration, *IEEE T Bio-Med Eng.* **57**, 1707-1718.
16. Hernandez-Matas, C., Zabulis, X. & Argyros, A. A. (2016). Retinal image registration through simultaneous camera pose and eye shape estimation. Paper presented at the *IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC)*.