3D Reconstruction and Registration for Retinal Image Pairs

Fan Guo*, Xin Zhao, Beiji Zou
School of Information Science and Engineering
Central South University
Changsha, P. R. China
e-mail: guofancsu@163.com

Pingbo Ouyang
Ophthalmology Department of the Second Xiangya Hospital
Central South University
Changsha, P. R. China

***Abstract*—Structure analysis of the optic-nerve head (ONH) and retinal image registration are important for diagnosing eye diseases. Thus, a 3D reconstruction method and a retinal image registration method are proposed in this paper. The former reconstruction method exploits disparity map estimation as a depth cue to disambiguate the relative positions and then uses the guided filter to obtain the final depth map. Using the depth map, a red-cyan anaglyph can be computed and the 3D structure of the ONH can be reconstructed. The latter registration method exploits and establishes the vasculature relationship between retinal images as a mosaic cue to be used for feature extraction and matching, and then divides the vessel segmentation image into four quadrants and obtains the Grid-based Motion Statistics (GMS) feature correspondence technique for each region. Using the regional GMS feature correspondence, the transformation parameters can be accurately identified and a good registration result can be ensured. Experiment results demonstrate the benefit of the proposed reconstruction and registration method.**

Keywords-3D reconstruction; optic-nerve head (ONH); disparity map; image registration; point-matching

I. INTRODUCTION

For retinal image pairs, there are two important key techniques: 3D reconstruction and image registration, which are essential to improve the accuracy of computer aided diagnosis of eye diseases and help doctor to achieve full information of eye diseases. Take Glaucoma for example, it is a disease in which elevated intraocular pressure causes the optic-disc to become cupped and atrophic, which eventually causes permanent impairment of vision. The most common clinical routine for glaucoma patients is examination of the optic-nerve head (ONH) [1]. Subtle ONH properties, such as neuroretinal rim slope and curvature of the retinal nerve fiber layer (RNFL), which has significance in assessment of the nerve loss and its progression, cannot be adequately described by two-dimensional parameters. Besides, though optical coherence tomography (OCT) can be used to quantify the topography of the ONH by using a scanning laser [2, 3], it lacks long-term data and is not available in most clinics. Thus, it is necessary to reconstruct the 3D structure of ONH by using retinal image pairs.

On the other hand, retinal image registration as the process of establishing pixel-to-pixel correspondence between two retinal images of the same eye is also a useful technique. There are many circumstances in which registration algorithm of retinal images is needed, such as diagnosing or monitoring retinal abnormality, etc. The main factors that will cause the difference between retinal images are [4]: (1) change of eye positions including movements along X-, Y- and Z-axis; (2) change of camera interior parameters, such as focal length and resolution; (3) change of imaging modality; and (4) change of retinal tissue in the progression of diseases. All these factors make the issue of retinal image registration a hard task. Besides, due to the limitation of image capture angle, the imaging regions are obtained each time always relatively smaller compared to the whole fundus images. Thus, it is important to mosaic retinal image pairs together.

II. RELATED WORKS

For 3D reconstruction of ONH, it is widely accepted that estimating the 3D shape of the ONH from retinal image pairs falls under the class of problems known as shape-from-stereo [5], which has been a challenge problem in computer vision research for decades. It involves the estimation of depth differences or 3D shape using two retinal images of the same scene under slightly different geometry. By measuring the relative position differences or disparity of one or more corresponding patches or regions in the two images, the 3D shape can be estimated using the underlying structure information [6].

For retinal image registration, the commonly used registration method for retinal images is feature-based method. The technique can be classified into region- and point-matching categories [7]. The region-matching approaches consider all the features in a region as a whole and identify the transformation parameters by minimizing the similarity measures. Point-matching methods rely on the matched features in both images, and the technique mainly consists of two steps: feature matching and transformation estimation. In this paper, a novel point-matching method is proposed to mosaic retinal image pairs.

III. PROPOSED ALGORITHM

A. ONH Reconstruction Based on Disparity Mmap Estimation

In this section, we present two ways to visualize the 3D shape of the ONH by using retinal image pairs: the red-cyan anaglyph method and the 3D reconstruction method. Both methods use the disparity between the left and right stereo

pair as the depth cue to display the 3D shape. However, the former generates a 3D stereoscopic image that requires the viewer to wear red-cyan glasses, and the latter reconstructs the shape of the ONH by using a stereo matching algorithm.

The key issue for of the proposed method is the depth map estimation that used for 3D reconstruction of the ONH. By applying the stereo matching algorithm, the disparity map of retinal image pairs can be considered as a depth cue which reflects the relative positions between salient objects (e.g. optic-nerve head, blood vessels, etc.) and their neighboring regions. The disparity map is then smoothed by guided filter to capture the depth boundary and remove the redundant details of the map. The smoothed map is a pseudo depth map instead of a recovery of real depth information. Using the pseudo map, the 3D shape of the ONH can be viewed as shown schematically in Fig. 1.

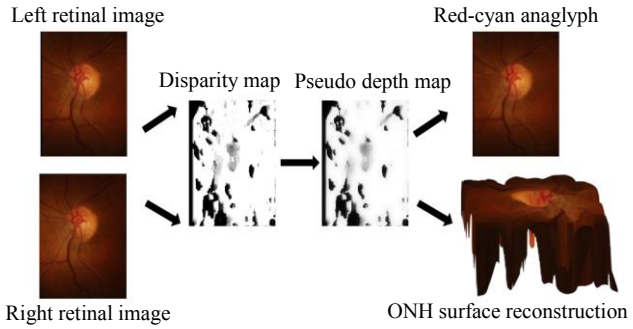

Figure 1. Flow chart for 3D reconstruction of ONH.

As can be seen in Fig. 1, the proposed algorithm is divided into the following three major steps.

1) Disparity map estimation: The aim of the stage is to estimate the dense disparity map of the input retinal image pairs. The disparity map is shown as gray level image where dark intensities corresponding to points that are far from the camera and the bright intensities correspond to points that are close to the camera. The disparity map estimation can be formulated as follows [8]:

$$SAD(x, y, d) = \sum_{x,y \in W} \left| I_l(x, y) - I_r(x, y - d) \right| \tag{1}$$

$$D(x, y) = \arg\min SAD(x, y, d) \tag{2}$$

where I_l and I_r are the intensity values in left and right image, (x, y) are the pixel's coordinates, d is the disparity value under consideration and W is the aggregated support region. The selection of the appropriate disparity value for each pixel is performed according to Eq. (2). For each pixel (x, y) and for constant value of disparity d the minimum cost is selected. Thus, the disparity map D can be obtained.

2) Pseudo depth map estimation: The aim of the stage is to estimate the pseudo depth map. Note that there are obvious noises and redundant details in the disparity map, thus the map is then smoothed by guided filter [9] to capture the depth boundary and remove the redundant details since the guided filter has excellent edge-preserving smoothing property. We observe that the depth map generated using the smoothed disparity map only reflects the relative positions between the salient objects (e.g. blood vessels, ONH, etc.) and their neighboring regions, thus the map is a pseudo depth map instead of a recovery of real depth information. The process can be written as:

$$\begin{aligned} a_k &= \frac{\frac{1}{|\omega|}\sum_{(x,y)\in\omega_k} D(x,y)^2 - u_k^2}{\sigma_k^2 + \varepsilon} \\ b_k &= u_k - a_k u_k \end{aligned} \tag{3}$$

where D is the disparity map, u_k is mean value of the input D in window ω_k. $|\omega|$ is the number of pixels in ω_k. σ_k is the pixel variance in the kth window, and the small variable ε is a regulation parameter to prevent a_k from being too large (ε = 0.04). Let $\bar{a}_k = (1/|\omega|)\sum_{i\in\omega_k} a_i$ and $\bar{b}_k = (1/|\omega|)\sum_{i\in\omega_k} b_i$, then the pseudo depth map *Depth* can be approached by:

$$Depth(x, y) = \bar{a}_k D + \bar{b}_k \tag{4}$$

3) 3D shape reconstruction: The aim of the stage is to visualize the 3D shape of the ONH. Once the pseudo depth map is estimated from a pair of retinal images, an anaglyph can be computed from the pseudo depth map and the left/right images. With a pair of red-cyan stereo glasses, one can experience the 3D effect. Furthermore, the 3D surface of the ONH can be also animated which allows the user to observe the retinal surface from different view angles. Due to the red-cyan anaglyph and the 3D reconstruction of the ONH structure, the study of patient retinal images is possible.

B. Retinal Image Registration Based on Regional Feature Point Matching

For retinal image registration, we present a regional feature point matching method to mosaic retinal images using prior knowledge and a GMS feature correspondence technique [10]. The key issue for retinal image registration is finding proper relationship between image pair. By establishing the vasculature relationship of the image pair to be matched, the blood vessels in each retinal image can be viewed as a mosaic cue to be used for GMS feature point matching. The matching process is carried out in four different local regions to ensure the correspondence established between the two feature groups in each region is reliable and the whole matched feature pairs are representative and comprehensive. Using these feature pairs, the transformation parameters can be easily and accurately identified, and the final registration result can thus be obtained as shown schematically in Fig. 2.

From Fig. 2, we can see that the proposed image registration method consists in the following three steps:

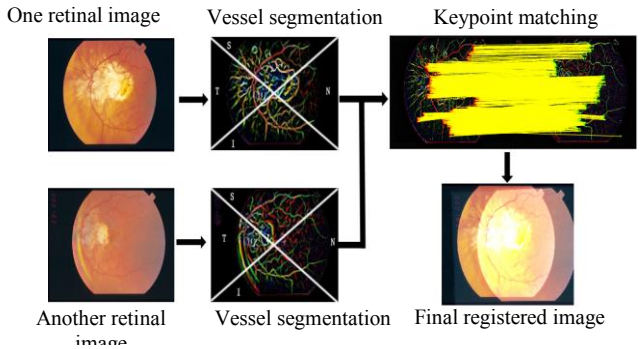

Figure 2. Flow chart for retinal image registration.

1) Vasculature relationship establishment: The aim of the stage is to establish the vasculature relationship between image pair because blood vessels are robust towards geometrical transformation and intensity changes. Whether the blood vessels can be successfully segmented determines the quality of feature points' extraction. Thus, the guided filter [9] is first used to enhance the original input images since the blood vessels in the enhanced retinal images are more obvious than those in the original images, which make it much easier for identifying the feature and structure of the vessels. Then, Jerman's 2D Hessian based vessel enhancement filter [11] is adopted here since the filter is proved to be quite effective for enhancing the retinal vasculature. The output of the Hessian-based filter is the vesselness probability map of the enhanced retinal image. The vessel segmentation results can thus be computed according to:

$$I_{vs} = I_{enh} \otimes P_v \tag{5}$$

where I_{enh} denotes an enhanced image ($0 \le I_{enh} \le 255$) obtained by the guided filter. P_v denotes the vesselness probability map ($0 \le P_v \le 1$). The vessel segmentation result I_{vs} can thus be obtained by taking a dot products operation between I_{enh} and P_v .

2) Four quadrants division: The aim of the stage is to divide the vessel segmentation image into 4 different quadrants and apply a GMS feature correspondence technique to each region. Note that the GMS method is based on the key observations that a true match may have many more supporting matches in its surrounding area than a false match, and the matching results seem quite promising for most natural scene images and robust than conventional feature descriptors (e.g. SIFT, SURF, etc.). However, from the point of view of the whole retinal images, GMS method always fails in extracting the feature points in some local regions with lower contrast and less vessels, which will cause an incorrect transformation estimation in the following steps. Therefore, the ISNT prior that widely accepted in ophthalmology field is used here to divide the vessel segmentation image into different quadrants (see Fig. 2). According to the ISNT rule [12], the neuroretinal rim can be divided into four quadrants: inferior (I), superior (S), nasal (N) and temporal (T). Using the ISNT prior, the four quadrants A_I, As, A_N and A_T are given by:

$$\begin{cases} A_I = I_{vs}(i,j) & where\ \ i > \dfrac{Mj}{N}\ \ and\ \ i \ge -\dfrac{Mj}{N} + M \\ A_S = I_{vs}(i,j) & where\ \ i \le \dfrac{Mj}{N}\ \ and\ \ i < -\dfrac{Mj}{N} + M \\ A_N = I_{vs}(i,j) & where\ \ i < \dfrac{Mj}{N}\ \ and\ \ i > -\dfrac{Mj}{N} + M \\ A_T = I_{vs}(i,j) & where\ \ i \ge \dfrac{Mj}{N}\ \ and\ \ i \le -\dfrac{Mj}{N} + M \end{cases} \tag{6}$$

In Eq. (6), $I_{vs}(i,j)$ denotes the vessel segmentation image corresponding to the pixel (i,j) , where $1 \le i \le M$, $1 \le j \le N$. M and N are the height and width of the vessel image, respectively. Once the four quadrants are obtained, the GMS technique can be adopted for each region to find more reliable feature pairs than the pairs obtained by just applying GMS to the entire image. Thus, a more robust transformation estimation can be ensured by considering some neglected local regions that have relatively few number of neighboring matches due to poor contrast and few vessels.

3) Retinal image mosaic: The aim of the stage is to identify the transformation parameters and mosaic the input retinal images. The GMS feature points from the four quadrants are all used to find the homography, which is a 3-by-3 matrix and contains the scale, rotation, and translation information of the two input images. Since the feature correspondence across two images may not be unique due to the similar angle values, the RANSAC algorithm [13] are also adopted here to eliminate false matching from the corresponding GMS points. Thus, by applying the homography to one of the input retinal images, a transformed image can be obtained to align the mosaic image.

IV. Experimental Evaluation

A. ONH Reconstruction Evaluation

To evaluate the flexibility of the proposed method, the root of mean squared (RMS) differences E_{RMS} between the estimated depth map *Depth* and the depth information $Depth^*$ obtained from OCT scans is used as a criterion and the index E_{RMS} is given by:

$$E_{RMS} = \sqrt{\frac{1}{N}\sum_x\sum_y \left(Depth(x,y) - Depth^*(x,y)\right)^2} \tag{7}$$

In Eq. (7), N is the total number of pixels. The E_{RMS} generally did not exceed 0.3, which demonstrates that the relative positions of the ONH can be well depicted in the estimated depth map just like the OCT data. Besides, our

visualization effects for both red-cyan anaglyph and 3D reconstruction of the ONH also seem very promising. Tests on retinal images were carried out on 59 image pairs. Fig. 3 shows an example obtained on two pairs of retinal images. The first row shows the stereo image pair of a confirmed glaucoma patient, and the second row is the image pair of a healthy subject. For glaucoma patients, the cupping of the optical nerve tends to get larger with more damages to optical nerve fibers compared with normal ones. This confirms our observations on Fig. 3. From the clinical point of view, the results are quite satisfactory since the depression of the optic disc is very visible for glaucoma diagnosis. From the image processing point of view, this visualization scheme is very promising for further processing to study the feasibility of automating glaucoma diagnosis using digital image processing techniques. Test of other retinal image pairs also demonstrates significant improvement of visual effects and reasonable reconstruction results. Thus, the proposed method has potential possibilities to help automatic glaucoma management with commonly used retinal image pairs instead of those more expensive OCT images.

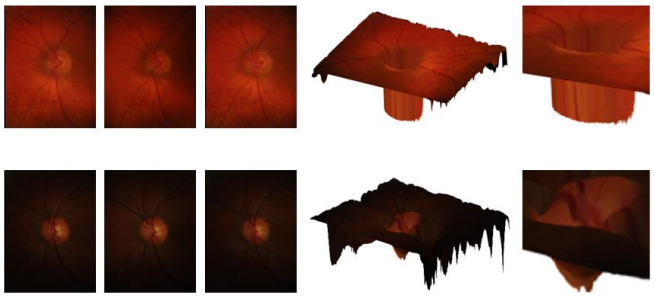

Figure 3. Example of ONH reconstruction for glaucoma patient and healthy object. (a) Left fundus images. (b) Right fundus images. (c) Red-cyan anaglyph results. (d) 3D reconstruction results. (e) Zoomed ONH region for (d).

B. Image Registration Evaluation

To evaluate the flexibility of the proposed method, the method is tested on both public database and real-captured images for evaluation purpose. There are totally 67 testing image pairs and the minimum overlapping area of the image pairs is approximately 40%. Experimental results show that the registration results seem very promising. Fig. 4 displays some registration examples obtained on two pairs of retinal images. For Figs. 4(a) and 4(b), the first row shows the image pair in the STARE database [14], where each image with a size of 700×605 pixels, and the second row is a pair of real captured images with a size of 800×791 pixels for each image. Fig. 4(c) shows the mosaic retinal images obtained by SIFT method, and Fig. 4(d) is the mosaic results obtained by the proposed method. One can clearly see that most of the vessels in the SIFT-based mosaic results are aligned well except a few local blood vessels with a little shift, while the proposed method has no such problem. Other testing image pairs also confirm this observation. After testing all image pairs, we observe that the only limitation of our method is that it requires the vascular structure to be successfully segmented. This constraint is common for point-matching registration as vessel segmentation is prerequisite. Once the vessel is successfully segmented, a good mosaic result can be obtained as long as more reliable and comprehensive feature correspondences can be found by using some relatively robust features (e.g. SIFT, GMS, etc.). Thanks to the region division strategy, the proposed regional GMS feature matching seems better than global SIFT-based matching. Thus, the proposed method has potential possibilities to help automatic retinal image analysis.

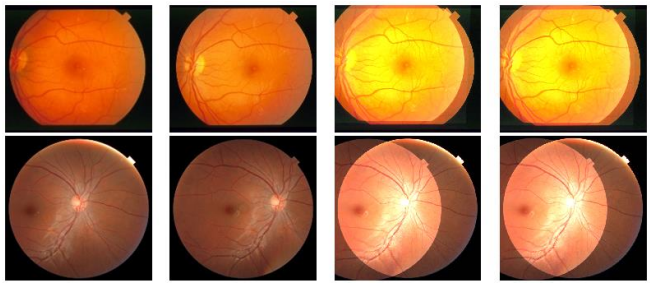

Figure 4. Image registration results. (a) One retinal image. (b) Another retinal image. (c) The mosaic retinal images obtained by SIFT method. (d) The mosaic retinal images obtained by the proposed method.

V. CONCLUSION

A 3D reconstruction method based on the estimation of the pseudo depth map for ONH and a retinal image registration method based on the regional feature point matching are proposed in this paper. Both methods proved to be efficient as an image reconstruction or registration technique and could be taken into consideration as a low-cost preliminary examination way for screening examinations and treatment monitoring of retinal pathologies, or as a suitable tool to be integrated into automated retinal image analysis system for clinical purposes. Meanwhile, other image reconstruction or registration problems in computer vision may benefit much from the proposed disparity map-based strategy or regional feature correspondence strategy.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (No. 61502537), and the Hunan Provincial Natural Science Foundation of China (No. 2018JJ3681).

REFERENCES

[1] Y. M. Zhu, “A Java program for stereo retinal image visualization,” Computer Methods and Programs in Biomedicine, 85, (3), pp. 214-219, 2007

[2] J. S. Schuman, M. R Hee., and A. V. Arya, etc., “Optical coherence tomography: a new tool for glaucoma diagnosis,” Current Opinion in Ophthalmology, 6, (2), pp. 89-95, 1995

[3] L. Tang, Y. H. Kwon, and W. L. M. Alward, etc., “3D reconstruction of the optic nerve head using stereo fundus images for computer-aided diagnosis of glaucoma,” Progress in Biomedical Optics and Imaging - Proceedings of SPIE, Vol. 7624, pp. 76243D: 1-8, 2010

[4] L. Chen, Y. Xiang, Y. J. Chen, and X. L. Zhang, “Retinal image registration using bifurcation structures,” Proceedings of IEEE International Conference on Image Processing, Brussels, Belgium, pp. 2169-2172, 2011

[5] D. Marr, and T. Poggio, “Cooperative computation of stereo disparity,” Science, 194, (4262) pp. 283-287, 1976

[6] S.T. Barnard, and W. B. Thompson, “Disparity analysis of images,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2, (4), pp. 333-340, 1980

[7] F. Guo, X. Zhao, B. J. Zou, and Y. X. Liang, “Automatic retinal image registration using blood vessel segmentation and SIFT feature,” International Journal of Pattern Recognition and Artificial Intelligence, 31, (11), pp. 1757006-1757031, 2017

[8] N. Lazaros, G. C. Sirakoulis, and A. Gasteratos, “Review of stereo vision algorithms: from software to hardware,” International Journal of Optomechatronics, 2, (4), pp. 435-462, 2008

[9] K. M. He, J. Sun, and X. O. Tang, “Guided image filtering,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 35, (6), pp. 1397-1409, 2013

[10] J. W. Bian, W. Lin, Y. Matsushita, S. K. Yeung, T. D. Nguyen, and M. M. Cheng, “GMS: Grid-based Motion Statistics for Fast, Ultra-robust Feature Correspondence,” Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, pp. 2828-2837, 2017

[11] T. Jerman, F. Pernus, B. Likar, and Z. Spiclin, “Enhancement of Vascular Structures in 3D and 2D Angiographic Images,” IEEE Transactions on Medical Imaging, 35, (9), pp. 2107-2118, 2016

[12] N. Harizman, C. Oliveira, A. Chiang, C. Tello, M. Marmor, R. Ritch, and J. M. Liebmann, “The ISNT rule and differentiation of normal from glaucomatous eyes,” Arch Ophthalmol, 124, (11), pp. 1579-83, 2006

[13] M. A. Fischler, and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” Communications of the ACM, 24, (6), pp. 381-395, 1981

[14] A. Hoover and M. Goldbaum, “Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels,” IEEE Transactions on Medical Imaging, 22, (8), pp. 951-958, 2003