HYBRID METHOD FOR RETINAL IMAGE REGISTRATION

Saeed Bamatraf, Buu Phan

University of Waterloo
Department of Electrical and Computer Engineering

ABSTRACT

In the field of medical imaging, retinal image registration is an important process for detecting and diagnosing eye diseases. This task, however, is complicated due to the nature of retinal image pairs, such as difference in intensity, texture, high resolution and small overlapping area. In this research, we propose a hybrid registration method, which utilizes the advantages of intensity and feature based registration technique to overcome the difficulties in registering retinal images. Unlike traditional feature based method, which use known feature points for registering, we consider the small area contains the optic disc as features which is detected and extracted by employing existing segmentation technique. The intensity based method is used for improving the solution from the feature-based one. We evaluate our method on the public available retina dataset and compare the results between different similarity measure used in this hybrid method. Our analysis shows that there is still a room for development in terms of similarity measure and optic disc segmentation.

1. INTRODUCTION

In the field of computer vision, image registration is a procedure which aims to align two or more images of the same view but different in time, viewpoints or capturing techniques, so that they are spatially matched to each other on pixel-by-pixel basis. Particularly, this means finding a transformation model such that after being applied, all images are registered to a reference image. This process is widely used in medical imaging [2], remote sensing [3], fingerprint or face recognition [4], image compression [5], video enhancement [6], and so forth.

In medical imaging, a common application of image registration is analyzing various eye diseases and conditions such as myopia and diabetic retinopathy using retinal images [7]. Depending on the analysis requirements, registration is practiced to aid task such as super-resolution, image mosaicking, and longitudinal study [8]. In super-resolution, information in the same area from several images are combined together to sharpen blurring features such as edges on retinal vessels. Images in this task often share a large overlapping area and thus combining their information can produce a high quality image. Image mosaicking, on the other hand, involves aligning different images within a small overlapping area. The purpose of this task is to enlarge the view of the retina, since ophthalmic camera's field of view (FOV) is limited with angle lesser than $50°$ [8]. Generating a wider view will enable better treatment for cases such as diabetic retinopathy and neonatal [9,10,11]. Longitudinal study is used for analyzing disease over time such as glaucoma. Images in longitudinal study often share large overlapping area, but different in terms of feature visual anatomical [12,22].

There are two main approaches for registering retinal image: feature-based and intensity-based registration. In feature-based registration, corresponding features, such as points and descriptors, between two images are located by using some domain knowledge or extraction algorithm, from which transformation which aligns two images are derived. This approach, however, requires features to be stable and detectable [16]. Intensity-based registration aims to optimize a predefined similarity measure (i.e loss function), such as mutual information (MI), cross correlation, sum of absolute value differences (SAV) [13,14], to find best set of parameters in the transformation model. The optimization process, however, is computationally expensive for images with high resolution, especially medical images, and is susceptible to local optima. Intensity patterns, non-overlapping area also affect the result of intensity-based registration [13, 15].

Retinal image registration is considered as a challenging task because intensities of retinal images have the difficulty of modalities and sensitivity with background. These difficulties motivate the utilization of robust features, such as optic disc and vasculature, instead of intensity in retinal image registration [17]. Feature-based methods can be classified into region and point-matching categories. In the region matching approaches, a feature images are extracted from original ones and the transformation parameters are identified by maximizing/minimizing the similarity measures. For example, the error between two binary vessel images is defined the cost function in [18] with consideration of various transformation models. Another similarity measure is formulated in [19] as the entropy-based MI. However, there are some drawbacks of region-matching methods because they depend on their huge searching space and in case of high-order transformation models, they encounter local convergence problem.

Furthermore, other factors such as presenting features in non-overlapping area, feature discontinuities, can critically affect the final result of this method. In contrary, point-matching methods rely on the matched features in both images. Since bifurcation point is a prominent indicator of vasculature, most of the point matching methods use it as landmark [20]. In [21] the retinal image registration for the optic disc segmentation in the fundus images is based on nonlinear filtering, Canny edge detector and a modified Hough transform. Mutual information using simulated annealing is used to define the registration and help finding maxima.

In this paper, hybrid technique is used in which both feature-based and intensity based methods are combined in a way to exploit the best of both techniques. For feature-based, optic disc is extracted first as a region of interest (ROI) and then local registration and translation of retinal is done based on that ROI. After that, intensity-based methods using metrics such as MI, SAV are used to fine tune the registration.

The structure of this paper is organized as follows. Section 2 describes the methodology that we adopted in the registration. Section 3 discusses the results. Section 4 gives conclusion and future work from this research.

2. METHODOLOGY

2.1. Dataset

In this paper, the Fundus Image Registration dataset (FIRE) is used for experimenting the hybrid method. Each image in the dataset is captured at 2912×2912 resolutions and a FOV of $45°$. The dataset provides a set of control points as a ground truth to evaluate the registration quality [22].

The dataset contains 134 retinal image pairs, which are divided into 3 categories: S, P and A. For the scope of this research, we focus on category P (or P-type), which consists of 49 pairs and is characterized by small overlapping area between the images in each pair. Fig. 1 shows a sample P-type pair in the dataset. Other information about the dataset can be found at [22].

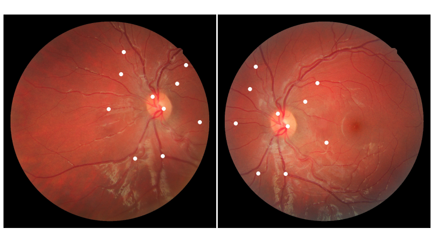

Fig. 1: Reference image (left) and moving image (right) with the control points shown as white dots

2.2. Hybrid Method for Image Registration

In this subsection, we briefly discuss about the advantages of our proposed hybrid method for this problem. The detailed procedure is described in section 2.4.

As it has been discussed in many literature, intensity-based methods are effective since, unlike feature-based method,they do not require prior knowledge and special features, about the images [1]. However, for high resolution images with small overlapping area such as category P in the dataset, intensity-based methods are likely to produce a faulty solution, since there are potentially many local optimal solutions along the non-overlapping region. Furthermore, it cannot be guaranteed that global optima is the true solution of the problem, because aligning correctly two images at the overlapping region does not ensure that the value of the similarity measure over the full area of two images is a global optima value [13,15]. The similarity measure,in addition, are only likely to achieve global optima over the overlapping region at the true solution as the effectiveness of the measure also depends on the nature of the images. For example, MI is more effective than SAV as similarity measure when working with images from different modalities [1].

In the proposed hybrid method aims to overcome those issues by combining feature and intensity based registration. The method first detects the ROI in each image, which is the small area around the optic disc,as features and then use feature based method to avoid potential local optima. Unlike point features, surface feature in this case are relatively heuristic to find, which is based on the position of the optic disc. Since the ROI also reflects the overlapping region between two retinal images, the intensity-based is applied in the overlapping region to produce the final solution to the problem.

2.3. Optic Disc Segmentation

As discussed previously, one heuristic for detecting mutual region between two retinal images is the optic disc. In this paper, we rely on existing optic disc segmentation method to extract the ROI.

The field of medical imaging, optic disc segmentation is an active field of research. Determining this region is currently performed by professionals, which is expensive and time-consuming [23]. Compare to segmentation in natural images, which contains popular objects such as vehicle, pedestrians, animals, etc., one notable challenge for medical image segmentation relates to the scarcity of data. Medical images, unlike natural images, are expensive and difficult to capture; furthermore, tasks such as segmentation, classification in medical images requires specialized knowledge and training [24]. In [24], authors introduces a deep convolutional neural network architecture, which is designed to work in such situation where data is limited. In the case of optic disc segmentation, research from [25] introduces a modified neural network version from [24] and shows competitive results with several previous existing methods such as entropy sampling [26] and deep neural network in [27]. Fig. 2 shows one segmentation result from a P-type image in FIRE using method from [25].

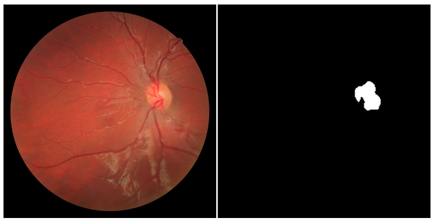

Fig. 2: Input image (left) and segmentation result (right)

In this paper, the method presented in [25] is employed as the first step to extract the ROI. For the P-type images in FIRE dataset, this method is able to detect and segment regions in 66 images out of 98. For some pairs, only one of them is successfully segmented while another is not, resulting 21 successful segmented pairs out of 49 pairs. This also shows that there is a room for development in this task.

2.4. Registration Procedure

In the procedure of registering two P-type retinal images I_M (reference image) and I_N (moving image),it assumes that transformation model $T(\theta)$ is rigid, and uses a predefined similarity measure $d(m, n)$ for intensity registration. The rigid transformation parameter θ consists of a 2D translation vector and a rotation angle. Image undergone rigid transformation with parameter θ is defined to be rotated and then translated accordingly. Also, for any image g, we denote $[T(\theta)g]$ as image g after undergoing transformation T with parameters θ and $g(A)$ as a sub-image or region obtained by extracting set of locations A , which is defined by 4 pixel locations forming a rectangular region, from image g.

This method first converts these two RGB images I_M and I_N to grayscale ones M and N respectively. After conversion, the procedure for registration retinal image pair M and N is proceed as follow:

- For each images in the pair, the deep neural network method in [25] is used to segment the optic disc. The set of points (SOP) defining the ROI for each image is obtained by first extracting the maximum and minimum locations along the x and y axis from segmented region of the image. Since the optic disc area is small and the color intensity is similar for each location (white pixels) in the segmented regions, the extracted pixel locations are extended vertically and horizontally, resulting the final SOP of ROI in each image, which is denoted as R_M and R_N (Fig. 3(a) and (b)). By extending the ROIs to neighbor area, the intensity-based method will be able to avoid local optima solutions which is caused by similar intensity over large region in both images.

- Perform local intensity registration on ROIs with the similarity measure $d(m, n)$ and transform the moving image N to N_α according to the optimization result (Fig. 3(c)). Formally, this step can be expressed as:

$$\theta^\alpha = \underset{\theta}{\operatorname{argmax}}\, d(M(R_M), [T(\theta)N(R_N)]) \tag{1}$$

$$N_\alpha = [T(\theta^\alpha)N] \tag{2}$$

- After obtaining N_α, pixel locations which defines R_N is also reallocated to a new positions. Due to the rotation from previous transform, edges of the rectangle formed by reallocated pixels will not parallel to the edges of the images (or perpendicular). This is resolved by extracting maximum and minimum values from the reallocated positions, by which a SOP R_{N_α} is obtained.

- Calculate two center locations c_M and c_{N_α} of the rectangles formed by R_M and R_{N_α}, which are used for determining a vector $\vec{\eta}$ pointing from c_{N_α} to c_M. Moving image N_α is translated with vector $\vec{\eta}$, resulting a new image N_β (Fig. 3(d)). The translation process can also be interpreted as rigid transformation with zero degree rotation. By this way, the overlapping region between the reference and moving image is heuristically adjusted and we are able to jump over many potential undesired optima which has been discussed previously.

- Consider two images M and N_β, we perform a fine tune intensity registration over ROI $M(R_M)$ of image M. By the end, we move the image N_β into N_{final} (Fig. 3(e)) by the optimization result.

$$\theta^* = \underset{\theta}{\operatorname{argmax}}\, d(M(R_M), [T(\theta)N_\beta(R_M)]) \tag{3}$$

$$N_{final} = [T(\theta^*)N_\beta] \tag{4}$$

In Fig. 3(f), difference between two images after registration is shown. Details such as retinal vessels can be seen to be overlapping each other, indicates by the black lines near the optical disc.

3. RESULTS AND DISCUSSION

We evaluate our method on the P-type and compare the result with different similarity measure $d(m, n)$. These similarity measure includes: SAV, sum of squared error (SSE), MI, Normalized MI (NMI), structural similarity (SSIM) and cross correlation (XCORR). As it has been mentioned previously, only 21 pairs out of 49 are successfully segmented are therefore, we can only assess our hybrid method on 21 pairs.

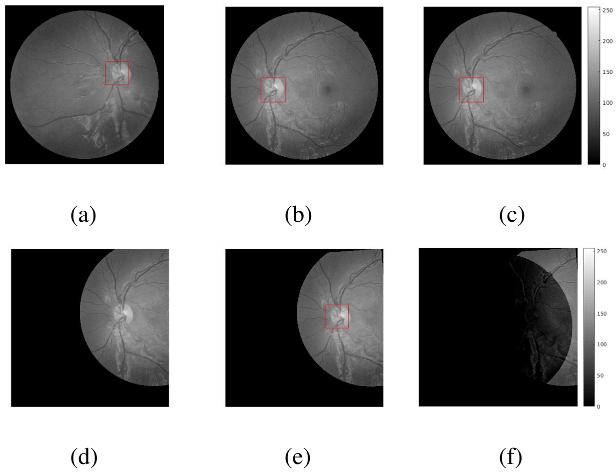

Fig. 3: Registration procedure. Fig. 3(a) and 3(b) are reference and moving image with their ROI. Fig. 3(c) is 3(b) after local registration. Fig. 3(d) is 3(c) after translation. Resulting image is 3(e) after final registration and 3(f) is difference between Fig. 3(e) and 3(a)

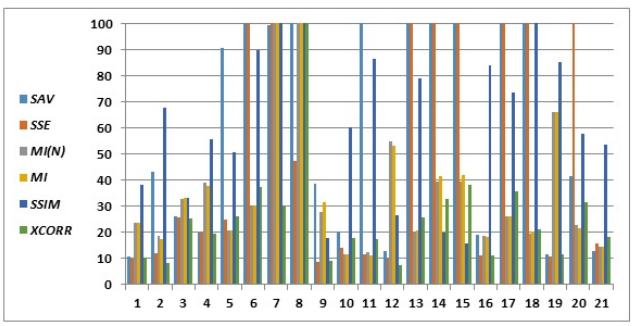

Fig. 4: TRE of 21 image pairs of Type P using 6 similarity measures

Target Registration Error (TRE) in Eq. (5) is used for quantifying the registration accuracy. A high registration accuracy is represented by a small TRE value and vice versa. TRE is an average distance measured in pixel from 10 corresponding landmarks (n) or ground truth control points between fixed (x_1, y_1) and moving (x_2, y_2) images after registration expressed in Eq.1. The landmarks identified by experts are provided by FIRE dataset.

$$TRE = \sum_{n=1}^{n=10} \sqrt{(x_2 - x_1)^2 - (y_2 - y_1)^2} \tag{5}$$

The results shows that some similarity measures work well with some pairs and others do not. Our analysis for that is because of the nature of the pairs i.e. in some pairs, one image is darker than the other and furthermore different in color structure. For visualization purposes, we just limit the maximum value to 100, however, some measures exceeds that limit and we consider that any error exceeds 100 as unsuccessful result.

4. CONCLUSION AND FUTURE WORK

In this paper, we proposed a hybrid methodology for retinal image registration, using both feature-based and intensity-based. For feature-based, optic disc is extracted as ROI and used as a background and starting point for the transformation of moving images to reference images. After that, similarity measures like MI, SAV are used to fine-tune the registration accuracy. Some similarity measures show good results for some pairs while others do not. There are two sources of error which are responsible for incorrect registration. The first one is the limitation of the segmentation technique, which may be due to the different between FIRE retinal images and their training images. Another source is the similarity measure, whose effectiveness is depends on the nature of the images. Although this paper assumes that the transformation model is rigid, it is not likely that this was a source of error, since in all 21 pairs, there is at least one similarity measurement that performs well.

The overall results are acceptable and as a future work we will continue improving the accuracy and try to include other categories in FIRE dataset. One possible direction for enhance the quality is to look for a better and more accurate way to detect optic discs. To scale up this method to other situations, one possible research direction is to automatically identify the overlapping area between two images. Developing a better similarity measure is another direction to improve the effectiveness of this method.

5. REFERENCES

[1]- B. Zitova and J. Flusser, Image Registration Methods: A Survey, Image and Vision Computing, vol. 21, pp. 977-1000, 2003.

[2]- A. Klein et al., Evaluation of 14 Nonlinear Deformation Algorithms Applied to Human Brain MRI Registration, Neuro-Image, vol. 46, pp. 786-802, 2009.

[3]- H. Li, B.S. Manjunath, and S.K. Mitra, A Contour-Based Approach to Multisensor Image Registration, IEEE Trans. Image Processing, vol. 4, no. 3, pp. 320-334, Mar. 1995.

[4]- L. Liu, T. Jiang, J. Yang, and C. Zhu, Fingerprint Registration by Maximization of Mutual Information, IEEE Trans. Image Processing, vol. 15, no. 5, pp. 1100-1110, May 2006.

[5]- F. Dufaux and J. Konrad, Efficient, Robust, and Fast Global Motion Estimation for Video Coding, IEEE Trans. Image Processing, vol. 9, no. 3, pp. 497-501, Mar. 2000.

[6]- M. Irani and S. Peleg, Motion Analysis for Image Enhancement: Resolution, Occlusion and Transparency, J. Vi-

sual Comm. And Image Representation, vol. 4, no. 4, pp. 324-335, Dec. 1993.

[7]- P. S. Reel, L. S. Dooley, K. C. P. Wong, and A. Borner, ”Multimodal retinal image registration using a fast principal component analysis hybrid-based similarity measure,” in Image Processing (ICIP), 2013 20th IEEE International Conference on, pp. 1428-1432, 2013.

[8]- R. Ramli, M. Y. I. Idris, K. Hasikin, A. Karim, N. Khairiah, A. W. Abdul Wahab, I. Ahmedy, F. Ahmedy, N. A. Kadri, and H. Arof, ”Feature-Based Retinal Image Registration Using D-Saddle Feature,” Journal of Healthcare Engineering, vol. 2017, 2017.

[9]- D. M. Brown and A. Ciardella, Mosaic fundus imaging in the diagnosis of retinal diseases, Investigative Ophthalmology Visual Science, vol. 46, pp. 25812581, 2005.

[10]- A. Bontala, J. Sivaswamy, and R. R. Pappuru, Image mosaicking of low quality neonatal retinal images, in 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), Barcelona, Spain, 2012.

[11]- B. H. Lee, G. Xu, K. Gopalakrishnan et al., AEGIS-augmentedeye laser treatment with region guidance for intelligent surgery, in presented at the 11th Asian Conference on Computer Aided Surgery (ACCAS 2015), 2015.

[12]- K. M. Adal, P. G. van Etten, J. P. Martinez, L. J. van Vliet, and K. A. Vermeer, Accuracy assessment of intra- and intervisit fundus image registration for diabetic retinopathy screening accuracy assessment of fundus image registration, Investigative Ophthalmology Visual Science, vol. 56, pp. 18051812, 2015.

[13]- P. A. Legg, P. L. Rosin, D. Marshall, and J. E. Morgan, Improving accuracy and efficiency of mutual information for multi-modal retinal image registration using adaptive probability density estimation, Computerized Medical Imaging and Graphics, vol. 37, pp. 597606, 2013.

[14]- R. Kolar, V. Sikula, and M. Base, Retinal image registration using phase correlation, Analysis of Biomedical Signals and Images, vol. 20, pp. 244252, 2010.

[15]- P. S. Reel, L. S. Dooley, K. C. P. Wong, and A. Brner, Enhanced retinal image registration accuracy using expectation maximisation and variable bin-sized mutual information, in 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 2014.

[16]- C. Y. Liu, J. Y. Ma, Y. Ma, and J. Huang, Retinal image registration via feature-guided Gaussian mixture model, Journal of the Optical Society of America A, vol. 33, pp. 12671276, 2016.

[17]- P. A. van den Elsen, E. D. Pol, and M. A. Viergever, Medical image matching - A review with classification, IEEE Engineering in Medicine and Biology Magazine, vol. 12, no. 1, pp. 26-39, Mar. 1993.

[18]- G. K. Matsopoulos, N. A. Mouravliansky, K. K. Delibasis, and K. S. Nikita, Automatic retinal image registration scheme using global optimization techniques, IEEE Trans. Information Technology in Biomedicine, vol. 3, no. 1, pp. 47-60, Mar. 1999.

[19]- N. Ritter, R. Owens, J. Cooper, R. H. Eikelboom, and P. P. van Saarloos, Registration of stereo and temporal images of the retina, IEEE Trans. Medical Imaging, vol. 18, no. 5, pp.404-418, May 1999.

[20]- L. Chen, Y. Xiang, Y. Chen, and X. Zhang, ”Retinal image registration using bifurcation structures,” in Image Processing (ICIP), 2011 18th IEEE International Conference on, 2011, pp. 2169-2172.

[21]- R. Chrstek, M. Skokan, L. Kubecka, M. Wolf, K. Donath, J. Jan, G. Michelson, H. Niemann, and R. Chrstek, ”Multimodal retinal image registration for optic disk segmentation,” 2004.

[22]-C. Hernandez-Matas, X. Zabulis, A. Triantafyllou, P. Anyfanti, S. Douma, A.A. Argyros, FIRE: Fundus Image Registration Dataset, Journal for Modeling in Ophthalmology, vol. 1, no. 4, pp. 16-28, Jul. 2017.

[23]- A. Almazroa, R. Burman,K. Raahemifar, V. Lakshminarayanan (2015). Optic disc and optic cup segmentation methodologies for glaucoma image detection: a survey. Journal of ophthalmology, 2015.

[24]- O. Ronneberger, P. Fischer, T. Brox. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234-241. Springer, Cham, 2015.

[25]- A. Sevastopolsky. Optic Disc and Cup Segmentation Methods for Glaucoma Detection with Modification of U-Net Convolutional Neural Network. arXiv preprint arXiv:1704.00979, 2017.

[26]- Zilly, J. M. Buhmann, and D. Mahapatra, Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation, Computerized Medical Imaging and Graphics, vol. 55, pp. 2841, 2017.

[27]- K. Maninis, J. Pont-Tuset, P. Arbel aez, and L. Van Gool, Deep retinal image understandingin International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 140148, Springer, 2016.