Multimodal Retinal Image Registration method Based on Speed-up Robust Feature

Zhitao Xiao[1], Wan Zhu[2], Fang Zhang[3], Jun Wu[4], Lei Geng[5], Wen Wang[6]

[1]School of Electronics and Information Engineering, Tianjin Polytechnic University, Tianjin 300387, China.

[1]

xiaozhitao@tjpu.edu.cn, [2]15022669595@163.com, [3]hhzhangfang@126.com, [4]zhenkongwujun@163.com, [5]genglei@tjpu.edu.cn, [6]wangwen19900129@163.com

Abstract: Retinal fundus image plays an important role in the diagnosis of eye diseases. In order to have a comprehensive understanding of retinal image, doctors need to take pictures from different angles of the patient's eyes. Because of the movement of the eye and head of the patient, there are some differences between the corresponding physiological structures. So it is necessary to carry out registration before information fusion. In this paper, a registration method based on speed-up robust feature is proposed. Firstly, preprocessing is performed on the fundus images, Secondly, surf feature points are extracted from the reference image and the floating image, and the descriptors are generated later. Finally, the spatial transformation of the floating image is carried out. In this paper, 120 fundus images of 20 patients were tested, the results showed that for different resolutions, angles, and translations, our method can be achieved automatic registration, average matching rate can reach 93.3%, and precision can reach sub pixel level.

1. Introduction

With the improvement of standard of living and the changement of lifestyle, diabetes, glaucoma and hypertension have been a serious threaten to people's healthy. The occurrence of these diseases is closely related to the characteristics of the retinal images. Therefore, the study of fundus images will provide the important basis for the medical workers in the diagnosis of fundus diseases.

Color fundus images and Fundus fluorescein angiography images are acquired in different devices. Image information obtained by different devices is not the same. The optic disc in color fundus images with high brightness, by observing the optic disc doctors can diagnosis fundus diseases which is associated with the changes of optic disc. In the fundus fluorescein angiography images, the blood vessels have a high contrast. Single model images can't provide enough information for doctors. Therefore, it is necessary to fuse information from different imaging equipments. Due to the movement of eye and head could lead to scale, rotation and translation differences between the fundus images, these differences will cause physiological structure of the fundus images exist spatial position deviation, so before information fusion we need to register images first.

Recently, many related registration approaches have been proposed for retinal image registration, which can be classified into two classes: area-based, feature-based, typically. The main idea of the area-based approaches is to use the gray level information between the images to get the change parameters. In [1], the similarity measure is obtained by calculating the mutual information of the images and then the transform parameters are obtained by genetic algorithm. The feature-based approaches are another retinal image registration method. This method mainly uses the characteristic points of the images to obtain the transform parameters. Vessel is a physiological tissue that is stable and covers a wide range in the fundus images, at present feature-based method is almost based on the characteristics of blood vessel. Zhang et al. [2] use vascular crossing points as the feature points and use the method of iterative closest point (ICP) and singular value decomposition (SVD) to match the fundus images. Zana et al. [3] proposed morphological methods

to extract blood vessels, then registration is realized through the angle invariance of the feature points. Choe et al. [4] proposed to use the Y feature of the branch points as the feature points for registration. Wei et al. [5] proposed a scale invariant feature (SIFT) as the matching feature points to achieve registration.

Retinal image registration can achieved through above methods, but the area-based registration method is sensitive to the change of the gray scale and the calculation is large. On the other hand, the registration method based on blood vessel is influenced by the accuracy of vascular extraction. Fundus images background is quite complex, when a large area of lesions appeared in the fundus, the difficulty of the extraction of blood vessels increased, as result, the speed and accuracy of registration is affected.

In view of the above problems, in this paper we propose a registration method based on speed-up robust feature, which is based on the original fundus image rather than vessels. The algorithm mainly includes three parts: preprocessing, feature matching and registration.

2. Preprocessing

The main object of this paper is shown as Fig. 1, called color fundus(CF) image and fundus fluorescein angiography(FFA) image. Due to the impact of the surrounding environment and other factors, the fundus image will be uneven brightness and low contrast, therefore, preprocessing is necessary before extracting the feature points. Three channels of color fundus images are shown as Fig. 2, it is obvious that the contrast of the green channel is higher and the information is more abundant, therefore, the green channel of these two kinds of fundus images is selected for subsequent processing.
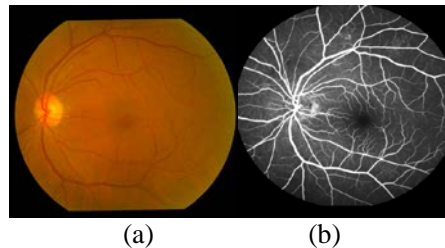

(a)   (b)

Fig. 1 (a) color fundus image (b) fundus fluorescein angiography image

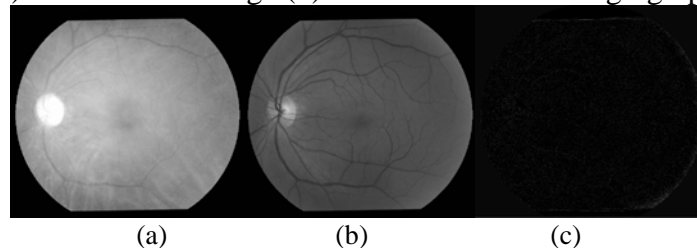(a)   (b)   (c)

Fig. 2 (a) red channel (b) green channel (c) blue channel

In this paper, contrast enhanced adaptive histogram equalization (CLAHE) [6] method is used to enhance the contrast of the green channel of the CF image, the result is shown as Fig. 3. FFA has a high contrast, so the contrast enhancement processing is no longer on the FFA image.
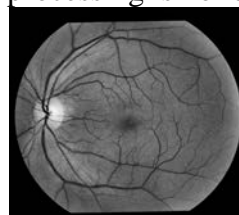

Fig. 3 CLAHE image

3. Features matching

3.1 Extract feature points

In order to get more information about the fundus, doctors need to take pictures from different

angles. The movement of the patient's head and eye will lead to the changement of scale, rotation and translation between the fundus images. In this paper, aimed at the scale invariance of feature detection problem, this paper analyzes and introduces a robust scale invariant feature, which is called speed-up robust feature. Interest point detector of SURF is based on integral images [7], Hessian matrix [8], and scale space theory. The core idea of SURF is that the interest point candidates can be detected by the maximum determinant of the Hessian matrix. Given an arbitrary point $p = (x,y)$ in an image I, we can give a 2×2 Hessian matrix $H_L(p,\sigma)$ in p at scale σ as follows:

$$H_L(p,\sigma) = \begin{bmatrix} L_{xx}(p,\sigma) & L_{xy}(p,\sigma) \\ L_{xy}(p,\sigma) & L_{yy}(p,\sigma) \end{bmatrix} \tag{1}$$

Where $L(p,\sigma) = G(p,\sigma)*I(p)$, $*$ denotes the convolution operation between image I and the Gaussian kernel, and the Gaussian function $G(p,\sigma) = 1/2\pi\sigma^2 \exp\left(-\left(x^2 + y^2\right)/2\sigma^2\right)$. L_{xx}, L_{xy}, and L_{yy} denote the convolution of the Gaussian second order derivatives with the image I in point p.

From the scale space, a new scale space is constructed using the difference of Gaussian (DOG) function as used in [9, 10, 11]:

$$D(p,\sigma) = L(p,k\sigma) - L(p,\sigma) \tag{2}$$

Where k denotes the linear scale difference between each image in each octave.

However, SURF approximates the determinant of the Hessian matrix $H_L(x,\sigma)$ by using box filters, and SURF constructs the approximation $D(p,\sigma)$ to $\det(H_L(p,\sigma))$ directly. More precisely,

$$\det\left(H_{approx}\right) = D_{xx}D_{yy} - \left(wD_{xy}\right)^2 \tag{3}$$

Where the relative weight w of the filter responses is used to balance the expression of the determinant of the Hessian matrix. Bay et al. [12] suggest $w = 0.9$.

Using the integral image, SURF approximates the different levels of scale space $D(p,\sigma)$ by adjusting the size of the box filters instead of the original image. Finally, interest points can be found by non-maximum suppression in a $3\times3\times3$ neighborhood around each sample point. The response value of the characteristic point is larger than the response value of 26 adjacent points, then the location and scale information of the feature points can be obtained. The SURF feature points of CF and FFA images are shown as Fig. 4.
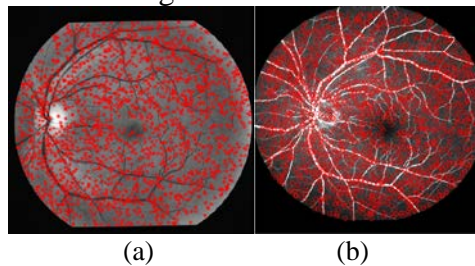

(a)   (b)
Fig. 4 (a) SURF feature of CF (b) SURF feature of FFA

3.2 Extract feature descriptors

In order to unify coordinate system, it is necessary to described the feature points, including two steps: determine main direction and build descriptors. In order to be invariant to image rotation, the Haar wavelet responses are calculated in x and y direction within a circular neighborhood of radius 6s around the feature point, s is the scale at which the feature point was detected. The Haar wavelet responses are represented as vectors. Then sum all the vector of x and y direction of the Haar wavelet responses within a sliding orientation window covering an angle of size $\pi/3$ around the feature point. The two summed response yield a new vector, called (m_w, θ_w),

$$m_w = \sum_w dx + \sum_w dy \tag{4}$$

$$\theta_w = \arctan\left(\sum_w dx \Big/ \sum_w dy\right) \tag{5}$$

The longest vector is the dominant orientation of the feature point, expression is as follows:

$$\theta = \theta_w \big| \max\{m_w\} \tag{6}$$

For extraction of the descriptor, construct a square region with a size of $20s$ and split the interest region up into 4×4 square sub-regions with 5×5 regularly spaced sample points inside. Compute the Haar wavelet response x-direction dx and the Haar wavelet response y-direction dy. Weight the response with a Gaussian kernel centered at the interest point. Sum the response over each sub-region for dx and dy separately. In order to bring in information about the polarity of the intensity changes, extract the sum of absolute value of the responses. Therefore, each sub-region is formed a 4-dimensional vector,

$$v = \left(\sum dx, \sum dy, \sum |dx|, \sum |dy|\right) \tag{7}$$

Concatenate the descriptor vector for all 4×4 square sub-regions, the length of a feature vector is 64. And compute the sum dx and $|dx|$ separately for $dy<0$ and $|dy|\geq0$, similarly for the sum of dy and $|dy|$, the length is doubled.

3.3 Feature point matching

After the establishment of feature descriptor, using Euclidean distance to realize the initial matching of feature points, then purified matching points using the random sample consensus algorithm.

Calculating the Euclidean distance of the feature points on the floating image to the reference image, get a distance sets, if the ratio of the minimum Euclidean distance and the second smallest Euclidean distance is less than the threshold 0.7, we think that the feature point and the minimum Euclidean distance of feature points are matched. Given two images I_1 and I_2, vector sets of image I_1 and I_2 are $V_1 = \{v_{11}, v_{12}, \cdots, v_{1m}\}$ and $V_2 = \{v_{21}, v_{22}, \cdots, v_{2n}\}$, where m and n respectively represent the number of SURF feature points of image I_1 and I_2. The expression of the threshold R is as follows:

$$R = \frac{dis(j_0)}{dis(j_1)} \tag{8}$$

Where $dis(j_0) = \min_j\left(dis\left(v_{1i}, v_{2j}\right)\right), dis(j_1) = \min_{j\neq j_0}\left(dis\left(v_{1i}, v_{2j}\right)\right), dis\left(v_{1i}, v_{2j}\right)$ denotes Euclidean distance of two feature vectors, R is the ratio of the minimum distance and the second smallest distance. The expression of Euclidean distance is as follows:

$$dis_{ij} = \sqrt{\sum_{k=1}^{N}\left(v_{ik} - v_{jk}\right)^2} \tag{9}$$

where N is the dimension of feature vector, the initial matching results are shown as Fig. 5.
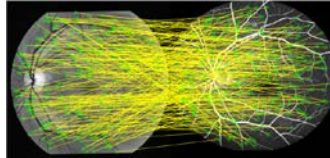

Fig. 5 initial matching results

In view of the problem of false matching, the method of random sample consensus [13] algorithm is used to remove the points with poor quality. The refined matching results are shown as Fig. 6.
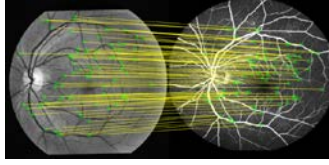
Fig. 6 refined matching results

4. Transformation model and parameter estimation

There are three kinds of models of fundus image: similar, affine and quadratic curve model, the degrees of freedom are 4, 6, and 12, respectively. Under the same registration conditions, the average performance of the three models are evaluated by the [14], respectively: 61%, 66%, and 65%. The registration accuracy results of the three models are tested by [15], and the results are as follows: 5.05 pixels, 4.58 pixels, and 0.64 pixels. Compared with the quadratic curve model, affine transformation model won't lose too much information, so we choose this model to carry on the spatial transformation. The expression of the affine transformation model is as follows:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{10}$$

where $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} s\cos\theta & -s\sin\theta \\ s\sin\theta & s\cos\theta \end{pmatrix}$, s, θ, t_x, and t_y respectively represent the scale, angle, the

translation of horizontal and vertical direction. According to the correct matching points, the least

square method is used to estimate the parameters. Registration results are shown as Fig. 7.
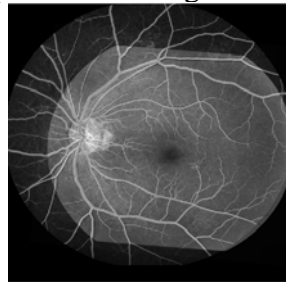

Fig. 7 registration results

5. Results

In this paper, 120 multimodal fundus images were tested, which were provided by Ophthalmological Hospital Affiliated to Tianjin Medical University, the resolution of the color fundus(CF) and fundus fluorescein angiography(FFA) image is 640×640 pixels. Partial fundus image registration results are shown as Fig. 8.
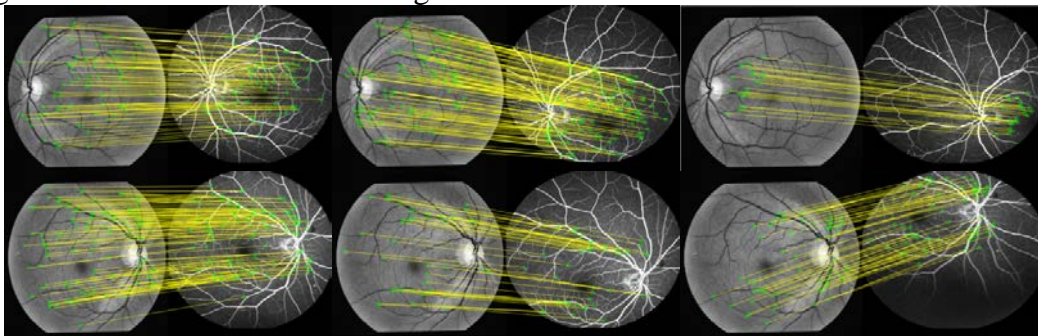

Fig. 8 Partial fundus image registration results

To evaluate the results of registration, we need a reliable and fair criteria to measure the performance of registration approaches, because of lacking public multimodal retinal image

registration database, in this paper, we tried two evaluation criterias to evaluate registration approaches: Matching rate and Root mean square error, the results were shown as Table1 and Table 2.

Table 1 Number of feature points and matching rate

number	feature points in image 1	feature points in image 2	matching points	error matching points	matching rate
1	1844	2088	78	5	93.6
2	1844	1899	121	8	93.4
3	1844	1916	24	2	91.7
4	1801	1943	122	9	92.6
5	1801	1941	46	4	91.3
6	1801	1643	58	5	91.4

Table 2 root mean square error

number	RMSE
1	0.6029
2	0.5900
3	0.5441
4	0.7451
5	0.6140
6	0.6241

In this paper, 20 patients with 120 multimodal fundus images were statistics, the average matching rate is 93.3%, and the average root mean square error is 0.8764. Compared with [16] and [2], the results are shown as Table 3.

Table 3 Comparison with other methods

methods	Average matching rate	Mean RMSE
[16]	85	——
[2]	——	1.4491
Our method	93.3	0.8764

6. Conclusion

In this paper, we focus on multimodal retinal image registration, and we found that the problem of previous registration approaches, then our proposed method not only extracts sufficient information from image pairs, but also removes incorrect matches. Multiple experiments on our 120 multimodal retinal images demonstrate that our proposed algorithm could achieve registration of multimodal fundus images in different angles, which has strong robustness, and the average matching rate can reach 93.3%, the accuracy can reach sub pixel level.

References

[1] Zhang Erhu, Yu Jiang, Retinal fundus image registration method based on mutual information, J. Journal of Xi'an University of Technology. Vol. 18 (2002) No. 1, pp. 14-17.

[2] Zhang Bo: Fundus image registration based on ICP and SVD (D, Jilin University, China 2009). pp .1-60.

[3] Zana F, Klein J C. A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform. IEEE Transactions on Medical Imaging. Vol. 18 (1999) No. 5, pp. 419-428.

[4] Choe T E, Medioni G, Cohen I, et al. 2-D registration and 3-D shape inference of the retinal fundus from fluorescein images. Medical Image Analysis. Vol. 12(2008) No. 2, pp. 174-190.

[5] Wei Lifang, Pan Lin, Yu Lun, et al. A registration method for fundus images. Chinese Journal of Biomedical Engineering. Vol. 30(2011) No. 4, pp. 549-554.

[6] Setiawan A W, Mengko T R, Santoso O S, et al. Color retinal image enhancement using CLAHE. International Conference on ICT for Smart-Society (ICISS). Jakarta, 2013, pp. 215-217.

[7] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Kauai, HI, United states, 2001, pp. 1511-1518.

[8] Lindeberg T. Feature detection with automatic scale selection. International journal of computer vision. Vol. 30(1998) No. 2, pp. 79-116.

[9] D.G. Lowe. Object recognition from local scale-invariant features. The Proceedings of the Seventh IEEE International Conference on Computer vision. Kerkyra, 1999, pp. 1150–1157.

[10] D.G. Lowe. Local feature view clustering for 3d object recognition. Proceedings 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Kauai, HI, United states, 2001, pp. 682–688.

[11] D.G. Lowe. Distinctive image features from scale-invariant keypoints. International journal of computer vision. Vol. 60(2004) No. 2, pp. 91-110.

[12] Bay H, Tuytelaars T, Van Gool L, et al. Surf: Speeded up robust features. 9th European Conference on Computer Vision, ECCV 2006. Graz, Austria, 2006, pp. 404-417.

[13] Lv Y, Feng J, Li Z, et al. A new robust 2D camera calibration method using RANSAC. Optik-International Journal for Light and Electron Optics. Vol. 126(2015) No. 24, pp. 4910-4915.

[14] Laliberté F, Gagnon L, Sheng Y, et al. Registration and fusion of retinal images: a comparative study. 16th International Conference on Pattern Recognition, ICPR 2002. Quebec City, QC, Canada, 2002, pp. 715-718.

[15] Chen J, Tian J, Lee N, et al. A partial intensity invariant feature descriptor for multimodal retinal image registration. IEEE Transactions on Biomedical Engineering. Vol. 57(2010) No. 7, pp. 1707-1718.

[16] Liu Shangping: Quality improvement, feature detection and registration in the automatic analysis of retinal images (D, Chongqing University, China 2011).pp. 1-87.