A novel registration method for retinal images based on local features

Some of the authors of this publication are also working on these related projects:

Project Vitamin C View project

Project VFA production and recovery from waste activated sludge by anaerobic digestion View project

A novel registration method for retinal images based on local features

Jian Chen, R. Theodore Smith, Jie Tian, and Andrew F. Laine

Abstract—**Sometimes it is very hard to automatically detect the bifurcations of vascular network in retinal images so that the general feature based registration methods will fail to register two images. In order to solve this problem, we developed a novel local feature based retinal image registration method. We first detect the corner points instead of bifurcations since corner points are sufficient and uniformly distributed in the overlaps. Second, a novel highly distinctive local feature is extracted around each corner point. These local features are invariant to rotation and contrast, and partially invariant to scaling. Third, a bilateral matching technique is applied to identify the corresponding features between two images. Finally a second order polynomial transformation is used to register two images. Experimental results show that our method is very robust and compute efficient to register retinal images even of very low quality.**

Keywords—**retinal images, registration, local feature, corner points, polynomial transformation**

I. INTRODUCTION

The purpose of retinal image registration is to spatially align two or more retinal images taken at different times or at different fields of view. Generally, retinal image registration methods are classified as area based methods and feature based methods. Area based methods usually require optimization of their similarity metric between two images, and the registration is achieved with the transformation that maximizes similarity metric. The similarity metric is chosen so that its optimum value is achieved when the two images are properly registered. Area based methods are often used in multimodal or temporal image registration applications for the reason that in those cases the nonoverlapping areas are small. Feature based methods [2-4] typically involve detecting the landmark points in retinal vascular network and extracting features around those landmark points, and then using a match metric to identify the correspondences between two images. The registration process is performed by maximizing a similarity measure computed from the correspondences. Feature based methods can be applied to register two retinal images with larger nonoverlapping areas.

J. Chen is with Institute of Automation, Chinese Academy of Science, Beijing, China; and the Department of Biomedical Engineering of Columbia University, NY, USA. (email: jc3129@columbia.edu).

Prof. R. T. Smith is with the Department of Ophthalmology, Columbia University, NY, USA (email: rts1@columbia.edu).

Prof. J. Tian is with Institute of Automation, Chinese Academy of Science, Beijing, China (email: tian@ieee.org)

Prof. A. Laine is with the Department of Biomedical Engineering of Columbia University, NY, USA (email: laine@columbia.edu).

However, the detection of landmark points is not always successful, and failure has severe effects on subsequent registration steps. Therefore, feature based methods fail to register low quality retinal images. For instance, France Laliberté et al [2] failed to register the images in figure 1 because it is hard to detect the bifurcations of vascular network in figure 1(a). In this paper, we proposed a novel feature based retinal image registration method to solve the mentioned problems. In our scheme, we first detect the corner points instead of bifurcations since corner points are sufficient and uniformly distributed in the overlaps, and are much easier to detect than bifurcations in low quality images. Second, a novel highly distinctive local feature is extracted around each corner point. These local features are invariant to rotation and contrast, and partially invariant to scaling. Third, a bilateral matching technique is applied to identify the corresponding features between two images, and those correspondences are not only at the bifurcations of vascular network, but also at some other salient landmark points in retinal image. Finally a second order polynomial transformation is used to register two images. Experimental results show that our method is very robust and computationally efficient to register retinal images even of very low quality.
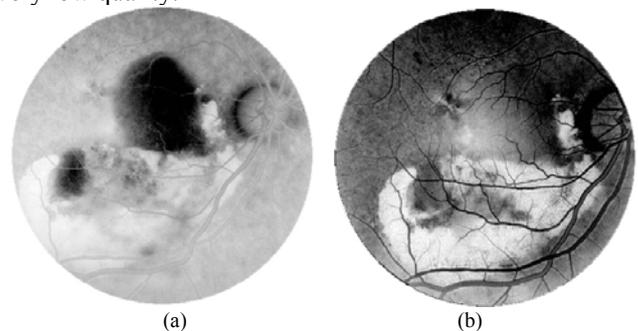

Figure 1. Retinal images taken at different stages. It is hard to detect the landmark points in (a) so that traditional feature based algorithm will fail to register these two images. This pair of image is from [2].

II. METHODS

Our suggested algorithm comprises four distinct steps:
- The Harris corner points detecting in both images.
- Extracting the local feature around each corner point.
- Bilateral matching of local features across the two images.
- Transformation based on the matched features .

A. The Harris corner detection

The Harris detector was proposed by C. Harris and M. J. Stephens [1] in 1988. This technique is computationally efficient and is easy to implement. The Harris detector is invariant to rotation so that it is applied to detect the

keypoints in retinal images in our algorithm. The basic idea of the Harris detector is to calculate the changes in all directions when convolved with a Gaussian window. Mathematically, the Harris detector is as follows:

$$M = \begin{bmatrix} G_x^2 & G_x G_y \\ G_y G_x & G_y^2 \end{bmatrix} * h \tag{1}$$

$$R = \det(M) - k \cdot tr^2(M) \tag{2}$$

Where G_x and G_y are the gradient of original image, h is the Gaussian window, k is a constant (usually $k = 0.04 \sim 0.06$). det and tr are determinant and trace of the matrix, respectively. If $R > 0$, the corresponding point in original image is a corner point. The bifurcations and the Harris corners of figure 1(a) are shown on figure 2. The bifurcations are detected by central line extraction algorithm [2]. In this case, only four bifurcations were detected and on the other hand large numbers of the Harris corners were detected and uniformly distributed on the image.
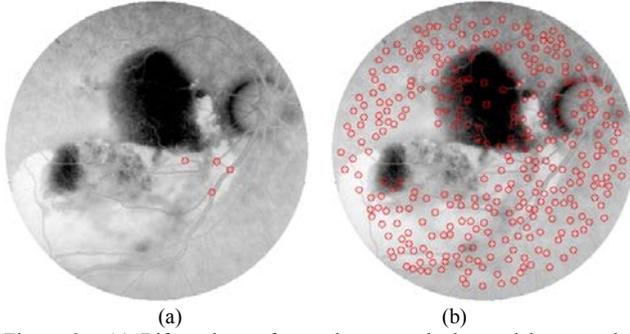

(a)   (b)

Figure 2. (a) Bifurcations of vascular network detected by central line extraction method. (b) Corner points detected by the Harris detector.

B. Local feature extraction

Before extracting the local features, a main orientation which is relative to the local gradients must be assigned to each corner point. Thus the local feature can be represented relative to this orientation and therefore achieve invariance to image rotation. In this paper, we introduced a continuous method, *averaging squared gradients* [6], to assign the orientation to each corner point (keypoint). This method uses the averaged perpendicular direction to the gradient to represent the keypoint's orientation. So the orientation has been limited from 0 to π. For each image sample, $I(x,y)$, the gradient vector $[G_x(x,y) \, G_y(x,y)]^T$ is as follows:

$$\begin{bmatrix} G_x(x,y) \\ G_y(x,y) \end{bmatrix} = sgn\left(\frac{\partial I(x,y)}{\partial y} \right) \begin{bmatrix} \partial I(x,y)/\partial x \\ \partial I(x,y)/\partial y \end{bmatrix} \tag{3}$$

The second element of the gradient vector has been chosen to always be positive. The reason for this choice is that opposite directions of gradient indicate equivalent orientations in symmetric descriptor (defined in the following context). Gradients cannot be averaged directly since opposite gradient vectors will then cancel each other, although they indicate the same orientation. A solution to this problem is proposed by squaring the gradient vector considered as a complex number before averaging. The squared gradient vector $[G_{sx}(x,y) \, G_{sy}(x,y)]^T$ is given by:

$$\begin{bmatrix} G_{sx}(x,y) \\ G_{sy}(x,y) \end{bmatrix} = \begin{bmatrix} G_x^2(x,y) - G_y^2(x,y) \\ 2G_x(x,y)G_y(x,y) \end{bmatrix} \tag{4}$$

Next, the Gaussian weighted average squared gradient $[\overline{G}_{sx}(x,y) \, \overline{G}_{sy}(x,y)]^T$ can be calculated. It is averaged in some neighborhood area, which is decided by the Gaussian weighted circular window with a σ:

$$\begin{bmatrix} \overline{G}_{sx} \\ \overline{G}_{sy} \end{bmatrix} = \begin{bmatrix} G_{sx} * h_\sigma \\ G_{sy} * h_\sigma \end{bmatrix} \tag{5}$$

Where h_σ is the Gaussian-weighted kernel, operator $*$ means convolution. Now the dominant direction of each neighborhood φ, with $0 \le \varphi < \pi$, is given by:

$$\varphi = \frac{1}{2} \begin{cases} \tan^{-1}(\overline{G}_{sy}/\overline{G}_{sx}) + \pi & \overline{G}_{sx} \ge 0 \\ \tan^{-1}(\overline{G}_{sy}/\overline{G}_{sx}) + 2\pi & for \quad \overline{G}_{sx} < 0 \cap \overline{G}_{sy} \ge 0 \\ \tan^{-1}(\overline{G}_{sy}/\overline{G}_{sx}) & \overline{G}_{sx} < 0 \cap \overline{G}_{sy} < 0 \end{cases} \tag{6}$$

For each keypoint whose coordinate is (x,y), its orientation is assigned to $\varphi(x,y)$.

The previous operations have assigned an orientation to each keypoint. Now it is possible to extract the local feature in a manner invariant to rotation and contrast [5]. We call this local feature a symmetric descriptor for the reason that it is symmetric to image contrast. First the image gradient magnitudes and orientations are sampled around the keypoint location. In order to achieve orientation invariance, the coordinates of the descriptor and the gradient orientations are rotated relative to the keypoint's main orientation. For each keypoint, the pixels that fall in a circle around the keypoint are selected to create the descriptor.

Figure 3 illustrates the computation of the symmetric keypoint descriptor. This symmetric descriptor is calculated by two sub-descriptors shown in figure 3 (c) and (e) (The figure shows a 2x2 array of histogram with 8 orientation bins, whereas our experiments are implemented with a 4x4 array with 8 orientation bins in each.). Figure 3(a) shows the magnitudes and orientations of gradients in a local neighborhood around the keypoint. In order to achieve contrast invariance, the orientations of gradient are limited from 0 to π, this is shown in figure 3(b). Figure 3(c) shows eight directions for each orientation histogram, with the length of each arrow corresponding to the magnitude of that histogram entry. It is important to avoid all boundary affects in which the descriptor abruptly changes as a sample shifts smoothly from being within one histogram to another or from one orientation to another. Therefore, trilinear interpolation is used to distribute the value of each gradient sample into adjacent histogram bins. In other words, each entry into a bin of histogram is multiplied by a weight of 1-d for each dimension, where d is the distance of the sample from the central value of the bin as measured in units of the histogram bin spacing. The sub-descriptor is formed from a vector containing the values of all the orientation histogram entries, corresponding to the lengths of the arrows in Figure 3(b).

The another sub-descriptor shown in figure 3(f) is calculated the same as figure 3(c). However, the local

neighborhood area is rotated 180 degree. This is shown in figure (d).

Assume one sub-descriptor (figure 3(c)) of keypoint is $A(4 \times 4 \times 8)$, and another sub-descriptor (figure 3 (f)) is $B(4 \times 4 \times 8)$ which is formed by rotating 180º of gradients' image. We can easily prove that:

$$B(i,j,k) = A(5-i, 5-j, k) \qquad (7)$$

where $i, j = 1, 2, \cdots, 4$, and $k = 1, 2, \cdots, 8$. So for efficiency, sub-descriptor B is not need to compute by rotating gradients image. We can get it from sub-descriptor A directly.

In order to achieve contrast invariance, the two sub-descriptor, A and B, must be combined together. Assume the symmetric descriptor of keypoint is $des(4 \times 4 \times 8)$, and then it is computed as follows:

$$des(i,j,k) = \begin{cases} c_1 \cdot |A(i,j,k) + B(i,j,k)| & i=1,2 \\ c_2 \cdot |A(i,j,k) - B(i,j,k)| & i=3,4 \end{cases} \qquad (8)$$

where c_1 and c_2 are two parameters to tune the proportion of magnitude in symmetric descriptor.
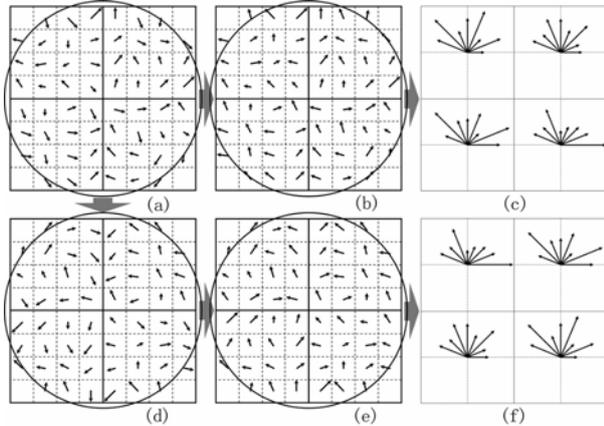

Figure 3: Symmetric descriptor. (a) the gradient magnitude and orientation at each image sample point in a region around the keypoint location. (b) all gradient orientation are restricted from 0 to π. (c) the accumulated gradient magnitude in each orientation. (d)-(f) show another accumulated gradient magnitude by rotating 180 degree of the original neighborhood around the keypoint. The symmetric descriptor is calculate from (c) and (f) by equation (8).

C. Bilateral matching method

We use the Best-Bin-First (BBF) algorithm [5]to match the correspondences between two images. It is an algorithm to identify the approximate closest neighbors of points in high dimensional spaces. This is approximate in the sense that it returns the closest neighbor with high probability. Suppose that the set of all symmetric descriptors of image I_1 is DES_1, and the set of I_2 is DES_2. For a given descriptor $des \in DES_1$, a set of distances is defined as follows:

$$Dis_{-des} = \{des \bullet des_s \mid des_s \in DES_2\} \qquad (9)$$

Where $\bullet$ means dot product of vectors. It is obvious that this set comprises all the distances between des and descriptors in I_2. The des_s correspond to the biggest element of Dis_{-des} denotes des's closest neighbor. Next we compare the distance of the closest neighbor to that of the second-closest neighbor. If the closest neighbor is significantly closer than the second-closest neighbor, then we can say it's a unilateral

match (or correspondence) from DES_1 to DES_2. Otherwise the descriptor des is discarded.

The BBF algorithm mentioned above is unilateral. It keeps the matches to be surjective, but the matches still can be injective. This means the unilateral BBF algorithm cannot exclude the following mismatch: two descriptors in I_1 are matched to the same descriptor in I_2.

The bilateral BBF algorithm is as simple as the unilateral one. The above unilateral matches are denoted as $M(I_1, I_2)$, and another unilateral matches $M(I_2, I_1)$ are also applied, then the same matches between these two set of matches are the bilateral matches.
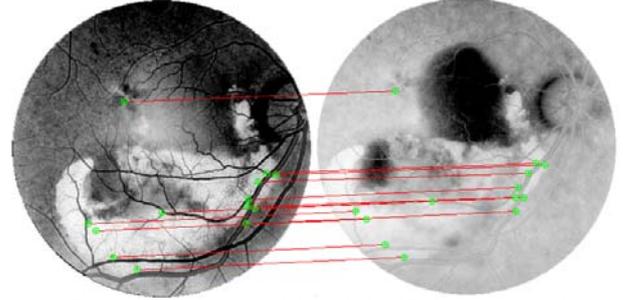

Figure 4. The correspondences which are indentified by our method between the two images in figure 1.

Even the bilateral BBF algorithm cannot guarantee all matches are correct. Fortunately it is easy to exclude the incorrect matches using the keypoints' orientations and the geometrical size of matches.

Suppose there are K matches in total, they are $m(p_{s1}, p_{t1}), m(p_{s2}, p_{t2}), \cdots, m(p_{sK}, p_{tK})$, where p_{si} is a keypoint in I_1, p_{ti} is a keypoint in I_2. It's obvious that the difference of p_{si}'s orientation and p_{ti}'s orientation is almost a constant. If the difference of orientations is much bigger or smaller than this constant, then the match is incorrect. Our experiments show that most incorrect matches are excluded by this criterion. Next we calculate the geometrical size of matches. The ratio of distances of two matches is defined as $r_{ij} = d(p_{si}, p_{sj}) / d(p_{ti}, p_{tj})$, where $d(p_{si}, p_{sj})$ means the Euclidian distance of two keypoints. If there is no large affine transformation, the ratio must be close to a constant too.

D. Second order polynomial transformation

Second order polynomial transformation was applied for general geometric distortion correction. It allows corrections of misalignment that cannot be corrected by rotation, translation and scaling only. This transformation is defined as follows:

$$X_C = A \cdot B \qquad (10)$$

$$A = \begin{bmatrix} a_{00} & a_{10} & a_{01} & a_{11} & a_{20} & a_{02} \\ b_{00} & b_{10} & b_{01} & b_{11} & b_{20} & b_{02} \end{bmatrix} \qquad (11)$$

$$B = [1 \quad x_D \quad y_D \quad x_D y_D \quad x_D^2 \quad y_D^2]^T \qquad (12)$$

Where $X_D = [x_D, y_D]^T$ are the coordinates of a point in the distorted image, $X_C = [x_C, y_C]^T$ are the coordinates of a point in the corrected image. A is the parameter matrix.

III. EXPERIMENTS AND RESULTS

We tested the proposed method on 12 retinal image pairs and compared with Dual Bootstrap-ICP algorithm [7] , and the vascular networks are difficult to extract from some of the images. Our method took about 90s to run all 12 cases on a Pentium M 1.5GHz laptop using Matlab, thus averaging about 7.5s per registration. Dual Bootstrap-ICP (download at [8]) took about 500s to run all, thus about averaging 42s per registration. In all 12 image pairs, all of them were registered very well by our method, but 5 of them failed by Dual Bootstrap-ICP algorithm. Three cases are shown in figure 5.

IV. CONCLUSIONS

A novel feature based image registration of retinal images has been presented. First, a set of uniformly distributed corner points is detected by the Harris detector. Second, a highly distinctive descriptor is extracted around each corner point. Then we apply a bilateral BBF method to identify the matches between two images, and remove the incorrect matches using the main orientation and geometric information of matches. Finally, a second order polynomial transformation is applied to register two images.

The proposed method is computationally efficient and totally automatic. It works on very low quality images that the vascular network is hard even impossible to extract. This method can deal with large initial misalignment registration like perspective distortion, arbitrary rotation and <1.5 times scaling. It also can deal with registration of field images with large non-overlapping areas.
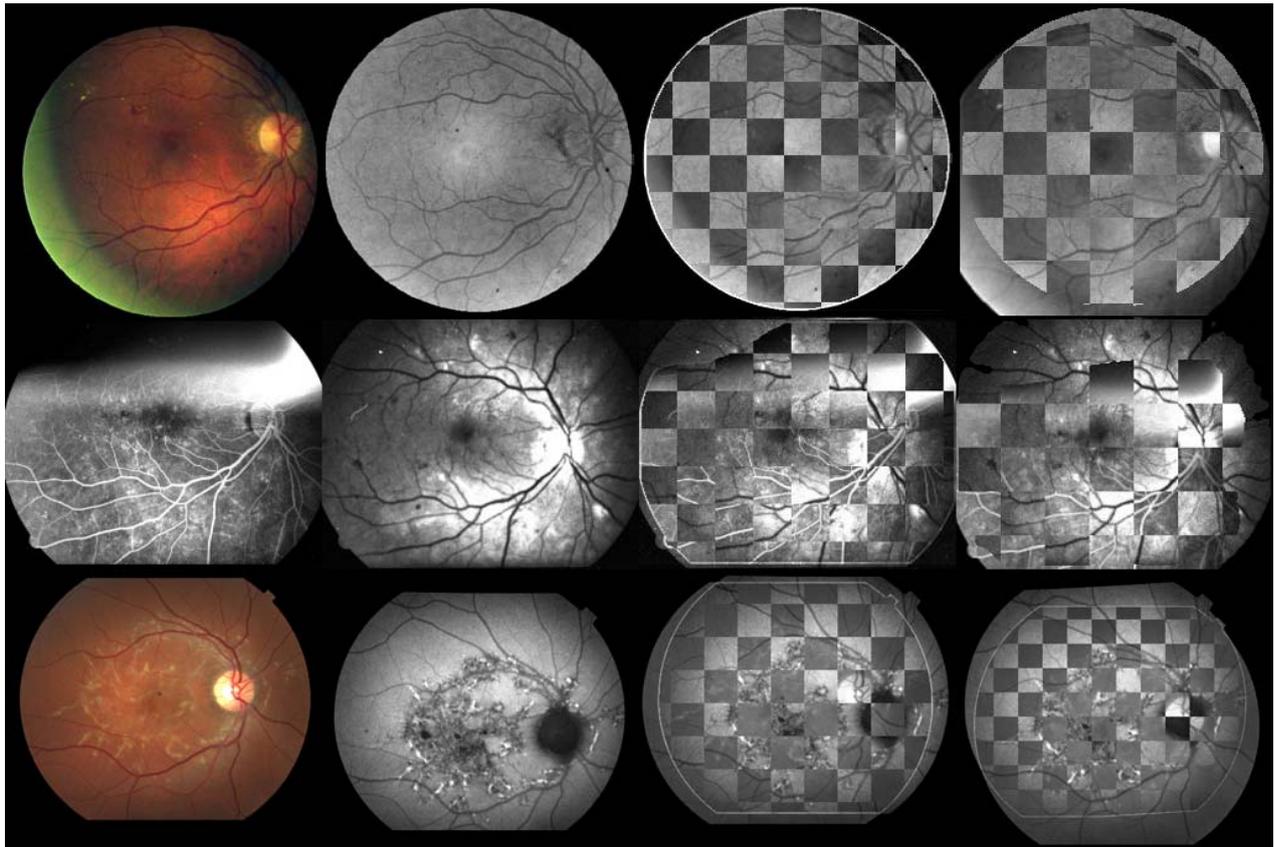

Figure 5. Comparing results between the proposed method and DB-ICP. The floating images and reference images are shown on the first and second column respectively. The results of DB-ICP are shown on the third column, and the results of the proposed method are shown on the last column.

REFERENCES

[1] C. Harris and M.J. Stephens. A combined corner and edge detector. In Alvey Vision Conference, pages 147–152, 1988.

[2] F. Laliberte, L. Gagnon and Y.L. Sheng, Registration and fusion of retinal images--an evaluation study, IEEE Trans. Med. Imag., vol.22, No. 5, pp. 661–673, 2003

[3] N. Ryan, C. Heneghan and P. de Chazal, Registration of digital retinal images using landmark correspondence by expectation maximization, Image and Vision Computing 22 (2004), pp. 883–898.

[4] T. Chanwimaluang, G. Fan, and S.R. Fransen, Hybrid retinal image registration. IEEE Trans Info. Tech. in Biomed., Vol. 10, No. 1, pp. 129-142, 2006.

[5] D.G.Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision 60(2), 91–110, 2004.

[6] Lin Hong, Yifei Wan, Anil Jain, Fingerprint image enhancement: Algorithm and performance evaluation, IEEE Trans. Pattern Analysis and Machine Intell., vol. 20, no. 8, pp. 777--789, Aug. 1998.

[7] Gehua Yang, Charles V. Stewart, Michal Sofka, and Chia-Ling Tsai, "Registration of Challenging Image Pairs: Initialization, Estimation, andDecision" IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 11, pp. 1973-1989, Nov. 2007.

[8] http://www.vision.cs.rpi.edu/download.html