

Pairwise retinal image registration by radial distortion correction

ABSTRACT

We describe a new method for pairwise retinal image registration. The proposed method is unique in that the radial distortion due to image acquisition is corrected prior to the geometric transformation. Since the overlapping region is typically small in a retinal image pair, only a few correspondences are available, thus limiting the applicable model to an affine transform at best. To recover the distortion due to curved-surface of retina and lens optics, a combined approach of an affine model with a radial distortion correction is proposed. Experimental results confirmed that the affine model with distortion correction could register retinal image pairs to within 1.88 ± 0.35 pixels accuracy assessed by vessel line error, which is 17% better than the affine-only approach. Because the proposed method needs only two correspondences, it can be applied to obtain good registration accuracy even in the case of small overlap between retinal image pairs.

Keywords: Registration, retinal imaging, radial distortion, Hessian, vessel tracing

1. INTRODUCTION

The human retina is the only part of the central nervous system that can be imaged directly, and retinal imaging has become an important tool for diagnosis and screening of ophthalmologic diseases such as diabetic retinopathy, age-related macular degeneration (AMD), glaucoma, etc.¹ Similar to looking through a small keyhole into a large room, imaging the fundus with an ophthalmologic camera allows only a limited region of interest (ROI) to be viewed at a time.² This limited view of the fundus camera makes it difficult to perform a comprehensive examination of the retina. Thus retinal image mosaicing, the process of stitching together multiple images to obtain a larger ROI, allows the ophthalmologist a panoramic view of the human retina.

Fundus imaging generally begins with taking a series of overlapping images showing different parts of the retina. The adjacency information found in the overlapping regions can be used to align pairs of images. Mosaicing is then accomplished by extending the pairwise registration to multiple cases. Considering global mosaicing as the graphical representation of pairwise connectivity weighted with registration error,^{3,4} the importance of pairwise registration is paramount.

Retinal image registration is the problem of seamlessly stitching together partially overlapped images that exhibit a radial distortion due to the imaging geometry. The partial overlap weakens the reliability of intensity measures, but also decreases the feasible number of matching features. The radial distortion due to camera optics and the sphere-to-plane projection during image acquisition should be corrected prior to registration. We propose a new method for registering retina image pairs with limited overlap and radial distortion. Our method is a feature-based strategy that uses the retinal vessel network to identify features, uses the affine registration model estimated using only two feature correspondences, and corrects the radial distortion to minimize the overall registration error. We demonstrate the method using a set of images acquired using standard 30° and 60° field of view (FOV) fundus cameras.

2. METHOD

We first detect feature correspondence points (CPs) within the images to be registered. We use measurements derived from vessel lines and their branchpoints as feature vectors, since the vessel network is easy to distinguish from the background and is distributed over most of the retina. The two CPs with the highest spatial and vector space correlation are used to initialize the affine model. The registration is refined by searching for the model parameters and radial distortion terms that give the minimum registration error as measured by the vessel centerline error measure (CEM) between reference and target images.

2.1. Feature Detection

The vasculature in the retinal images provides the most prominent features for image registration. Our method extracts vessel centerlines and identifies vessel branch points as CP candidates.

An interesting characteristic of retinal image is the appearance of *vessel reflex* caused by the specular reflection from the tube-like vessel geometry. Vessel reflex results in a bright line along the centerline of the normally dark vessel segment. We use a fast morphological reconstruction to suppress vessel reflex and image background noise. The morphological reconstruction using *h-max* operator⁵ is a fast version of an iterative reconstruction with three steps: raster scanning, anti-raster scanning, and propagation. Since the retinal vessel has an elongated tube-like shape, only the first two steps are needed in our method and as a result, the reconstruction can be completed with just two passes through the image data. Fig. 1 shows an example illustrating retinal vessel reflex and results of the preprocessing step with *h-max* set to 10% of intensity range. As is evident from Fig. 1(c), the vessel reflexes (see circles) and background noise are eliminated while keeping the edge location. This preprocessing step might inadvertently remove small vessel lines, but overall will contribute to more consistent outcomes under intensity variations between images.

Explicit vessel segmentation (such as described in Ref. 6) followed by line thinning can produce vessel centerlines. Since we will use detected vessel lines to estimate the registration error during optimization, the vessel lines need to be detected with a high degree of consistency between images. Extraneous line segments without true correspondences may cause erroneous vessel line matches which are almost impossible to exclude during the optimization step. To address this problem, we use a conservative vessel line extraction that tracks candidate vessel segments by tracing the intensity valleys in vessel profiles. Tube-like vascular structures are enhanced using the Hessian⁶ operator. For an image $I(x, y)$, the Hessian operation $H(x, y)$ is given by

$$\begin{aligned}
 H(x, y) &= I(x, y) \otimes G(x, y; \sigma) \\
 &= I(x, y) \otimes \begin{bmatrix} G_{xx}(x, y; \sigma) & G_{xy}(x, y; \sigma) \\ G_{yx}(x, y; \sigma) & G_{yy}(x, y; \sigma) \end{bmatrix},
 \end{aligned}$$

where $G(x, y; \sigma)$ is a 2D Gaussian probability distribution function with zero mean and standard deviation (scale) σ , and \otimes is the convolution operator. The Hessian operator produces a 2×2

matrix for each pixel, and the matrix eigenvalues and eigenvectors provide information on the ridgeness of the intensity profile. The largest eigenvalue is associated with the eigenvector that is orthogonal to vessel line direction. Consequently, our tracing strategy successively probes points that correspond to the perpendicular direction of eigenvector with the largest eigenvalue. The tracing starts from seed points found using a method similar to that given in Ref.7. For each row and column on virtual grid lines overlaid on the image, center points of the local gradient pair found in an average vessel width range are pushed into the seed list. Vessel tracing is done independently for each seed, and terminates when a trace meets another trace, when the Hessian eigenvector changes its direction abruptly, or when two eigenvalues are similar in magnitude.

When the tracing is completed, end points of the traced segments are collected to make an endpoint list. The list is utilized to delete short line segments due to false seed detection and to connect local breaks or lateral branching in line segments. A more important use of the endpoint list is the generation of line branches. In case the trace reaches vessel branches or bifurcations, tracing stops and will give disconnected vessel segments since the Hessian operator only responds strongly for elongated structures. Our method traces back from the endpoints to extrapolate vessels. The crossing of two or more extrapolated vessel lines is considered a branch center. As shown in Fig. 2(b), incomplete line segments are either merged or eliminated during endpoint processing. In the following steps, the vessel centerlines of Fig. 2(c) and the branch vectors of Fig. 2(d) are used for registration error calculation and for correspondence selection respectively.

2.2. Correspondence Points Selection

Branch centers are characterized by their location and the orientation of the branches leaving the branch center. For a branch center located at $\vec{C} = (C_x, C_y)$, the branch vector \vec{B} is given by

$$\vec{B} = \left\{ \vec{C}, \vec{\theta}_{\text{near}}, \vec{\theta}_{\text{far}} \right\} = \left\{ C_x, C_y, \theta_{\text{near}}^0, \theta_{\text{near}}^1, \theta_{\text{near}}^2, \theta_{\text{near}}^3, \theta_{\text{far}}^0, \theta_{\text{far}}^1, \theta_{\text{far}}^2, \theta_{\text{far}}^3 \right\},$$

where θ_{near}^i and θ_{far}^i are the three branch angles measured at 3σ and 6σ from \vec{C} for i^{th} branch, $i = 0, 1, 2, 3$ (see Fig. 3(a)).

Feature correspondence is first evaluated using branch angles, so a branch in the reference image is paired with the branch in the target image that gives the smallest sum of angle difference. Then a graph-based spatial correspondence is considered for further refinement of branch pairs. Fig. 3(b) is an example where three branch pairs are involved and (3) and (3)' are incorrectly paired by angle criterion. In the graph, the node represents branch center and the edge weight is determined by Euclidean distance between nodes. By measuring the difference in edge angle and weight, the edge pair A and A' turns out to have the highest similarity, so (1)-(1)' and (2)-(2)' are chosen for the CPs. In case more than two CPs have high similarity within tolerance, we pick a pair of CPs with the highest edge weight. After the two CPs are selected, the CPs are used to estimate the initial parameters of the affine model.

2.3. Radial Distortion Model

During the retinal imaging process, light rays reflect from the retina after traveling the two-way path through the cornea and a series of optical lenses, creating a nonlinear distortion. The retinal image is acquired by projecting the 3D spherical retinal surface, through the lenses, onto the 2D imaging plane. If we model the optical effects using a single lens equivalent to the optical cascade, the distortion in retinal imaging can be described as a combination of two coordinate deformations: the optical distortion and the spherical-to-planar projection. The former, despite using a fundus camera that is well calibrated to compensate for path deformation at the cornea, still causes a pincushion distortion as images are acquired away from the optical axis.⁸ A more significant distortion is caused by the projection of the curved retinal surface to a planar image space, resulting in a barrel distortion. The amount of distortion of both types is dominated by the distance from the optical center, commonly known as radial distortion.⁸⁻¹⁰

The combined radial distortion can be modeled as a coordinate transformation where a point u on the spherical surface of the eye is transformed as a function of position off of the optical axis to give the distorted coordinate v . The geometric distortion can be corrected by estimating

\hat{v} from u using

$$\hat{v} = u \cdot \left(1 + \sum_{n=1}^{\infty} k_n r^{2n} \right),$$

where the k_n terms are the radial distortion coefficients and r is the distance from optical axis center to u . In this paper, we consider only the $n = 1$ term in the series, i.e., $\hat{v} = u \cdot (1 + kr^2)$, since this accounts for 95% of radial distortion.^{8,10} Consequently $k = k_1$ becomes the only parameter in our distortion correction model. The sign of k determines the type of distortion; $k < 0$ indicates barrel distortion and $k > 0$ indicates pincushion distortion.

2.4. Estimating the Registration and Distortion Correction Model Parameters

The radial distortion parameter k and the parameters of the affine transformation are estimated during an optimization step. The radial distortion parameter k is initially set to zero (no radial distortion). The affine model matrix M is initialized by the relative scaling, rotation, and translation calculated from the two CPs selected during correspondence points selection. Since k and M are independent of each other, we update them separately. k is determined using bisection search and M is refined using Powell’s method¹¹ to reduce the centerline error measure, CEM,

$$\text{CEM} = \frac{1}{N} \sum_{n=0}^{N-1} |v_n^{(r)} - \hat{v}_n| = \frac{1}{N} \sum_{n=0}^{N-1} |v_n^{(r)} - v_n^{(t)} M|,$$

where N is the vessel line length and $v^{(r)}$ is the closest point in the reference image after transforming a point $v^{(t)}$ in the target image.

After completing the optimization step, intensity blending is used to combine two registered images to form a composite image. In this paper we use pixel averaging to combine the two images together in order not to artificially suppress ghost vessels that serve as a visual indicator of registration performance. More sophisticated image blending techniques, such as pyramidal blending¹² and gradient blending,¹³ could also be used.

3. RESULTS

We demonstrate the method using 20 pairs of images taken of three subjects using a Canon 60 color fundus camera (Canon, Tokyo, Japan). Images for one subject were acquired

using a 30° FOV and with a 60° FOV for the other two subjects. The camera produces a three-channel RGB color image, but we use the only green channel for these experiments since the green channel has the highest information content and best contrast for the retinal vessels. The 8-bit green channel images were resized from 2392×2048 to 598×512 pixels. The percentage overlap between image pairs varied from 12.8% to 59.5%.

The grid line spacing for seed detection was set to 10 pixels, and the scale factor σ for the Hessian operator was set to 2.0 pixels. The vessel tracing continued as long as the eigenvalue difference was less than the scale parameter and eigenvector angle change was no more than 90°. The search range for the optimal radial distortion coefficient k was [-0.2, +0.2].

Fig. 4 is the registration result of the image pair with 28.9% overlap. Visual inspection reveals that the neither the similarity transform (Fig. 4(a)) nor the affine transformation (Fig. 4(b)) is able to register the image pairs without obvious vessel mismatches. If the non-linear barrel distortion ($k < 0$) is corrected and then registration is performed with the affine model (Fig. 4(c)), most of the ghost vessels are eliminated.

The overall registration accuracy as measured by CEM as given by Table 1. Based on the CEM, our proposed method can register retinal image pairs with 17% higher accuracy than the affine-only approach. Execution time required to match a single image pair using the affine model and distortion correction is about two minutes on a Pentium IV 2.0 GHz PC, but the actual runtime depends on the vessel line lengths and the percent overlap between images. Our method is implemented using Java programming language as a plugin for ImageJ.¹⁴

4. DISCUSSION

We believe this is the first use of a radial distortion correction model in the retinal image registration literature. The results show that the affine model plus the radial distortion correction can produce better image registrations (as measured by the CEM) than with registrations using the similarity transform or affine model alone. Visual inspection also indicates that the radial distortion correction plus affine model produces fewer ghost vessels. The cost of the improved

registration is that one extra parameter, the distortion parameter k , must be estimated during the optimization step.

The hybrid registration method¹⁵ uses the affine model and cannot adjust for the distortion due to the spherical shape of the retina. Others have proposed a quadratic model to account for the curved-surface of human retina.¹⁶ The quadratic model produces good registrations in the overlapping regions in many cases, however, the model requires at least six CPs, which is not common in real fundus imaging, and a small error in feature detection may cause significant coordinate deformation in the non-overlapping region as reported in Ref. 15.

Our method can correct for distortion sources such as the optical lenses and the retinal surface convexity. Lens optics can produce pincushion distortion. The shape of the eye and retinal surface causes the barrel distortion during image acquisition. Depending on the FOV, the distortion type is automatically determined, and the amount of distortion is expressed as a single parameter k in radial distortion model. It is interesting to note that the distortion correction is applied to both images, which violates the general image registration rule that the reference image is never modified. We correct the distortion in the reference image because each image is acquired from the same eye and has spherical radius. Extending this idea from pairwise registration to general image montaging, after the first pairwise registration in the montage is performed, we can dramatically narrow the search range for k for the remainder of the montage because we expect that k should be almost constant for a given eye.

5. CONCLUSION

Retinal image pair registration is a fundamental processing operation for mosaicing together separate fundus images. In this paper, we propose a new pairwise retinal image registration method that incorporates radial distortion correction. Our method is unique in that the optical distortion due to the lenses and the spherical-to-planar projection are corrected during registration, but we retain the ability to use a low-order transformation model. The results show that the new method is superior to the affine-only approach, yet does not require the 6 CPs

needed by the quadratic model. Future work will include extending these preliminary results by estimating a set of global distortion parameters and applying these methods to automatic retinal image mosaicing and the 3D reconstruction of human fundus image.

REFERENCES

1. J. Kanski, *Diseases of the ocular fundus*, Elsevier/Mosby, New York, 2005.
2. R. E. Hackel and P. J. Saine, "Creating retinal fundus maps," *J. Ophthalmic Photography* **27**(1), pp. 10–18, 2005.
3. A. Can, C. Stewart, and B. Roysam, "A feature-based technique for joint, linear estimation of high-order image-to-mosaic transformations: Mosaicing the curved human retina," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **24**, pp. 412–419, 2002.
4. T. Choe, I. Cohen, M. Lee, and G. Medioni, "Optimal global mosaic generation from retinal images," *Int. Conf. on Pattern Recognition* **3**, pp. 681–684, Aug. 2006.
5. L. Vincent, "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms," *IEEE Trans. on Image Processing* **2**, pp. 176–201, April 1993.
6. J. Staal, M. Abràmoff, M. Niemeijer, M. Viergever, and B. Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. on Medical Imaging* **23**, pp. 501–509, April 2004.
7. A. Can, H. Shen, J. Turner, H. Tanenbaum, and B. Roysam, "Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms," *IEEE Trans. on Information Technology in Biomedicine* **3**, pp. 125–138, June 1999.
8. G. Toderici, A. Esquivel, H. Li, and I. Kakadiaris, "Evaluation of variability and significance of fundus camera lens distortion," in *Proc. 26th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, pp. 1–5, (San Francisco, CA), Sep. 2004.
9. J. Weng, P. Cohen, and M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **14**(10), pp. 965–980, 1992.

10. J. Xu and O. Chutatape, “Comparative study of two calibration methods on fundus camera,” in *Proc. 26th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, pp. 576–579, (Cancun, Mexico), Sep. 2003.
11. W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C; The Art of Scientific Computing*, Cambridge University Press, 2002.
12. E. Adelson, C. Anderson, J. Bergen, P. Burt, and J. Ogden, “Pyramid methods in image processing,” *RCA Engineer* **29**(6), pp. 33–41, 1984.
13. A. Zomet, A. Levin, S. Peleg, and Y. Weiss, “Seamless image stitching by minimizing false edges,” *IEEE Trans. on Image Processing* **15**, pp. 969–977, April 2006.
14. M. Abràmoff, P. Magelhaes, and S. Ram, “Image processing with imagej,” *Biophotonics International* **11**(7), pp. 36–42, <http://rsb.info.nih.gov/ij>, 2004.
15. T. Chanwimaluang, G. Fan, and S. Fransen, “Hybrid retinal image registration,” *IEEE Trans. on Information Technology in Biomedicine* **10**, pp. 129–142, Jan. 2006.
16. A. Can, C. Stewart, B. Roysam, and H. Tanenbaum, “A feature-based robust, hierarchical algorithm for registering pairs of images of the curved human retina,” *IEEE Trans. on Pattern Analysis and Machine Intelligence* **24**, pp. 347–364, 2002.

Table 1: Mean \pm standard deviation of vessel line error in pixels across 20 image pairs.

	Similarity transform	Affine transform	Proposed method
Subject I	3.03 \pm 0.63	2.39 \pm 0.38	1.58 \pm 0.13
Subject II	2.71 \pm 0.66	1.97 \pm 0.32	1.87 \pm 0.37
Subject III	2.63 \pm 0.21	2.28 \pm 0.22	2.20 \pm 0.18
Average	2.78 \pm 0.54	2.20 \pm 0.35	1.88 \pm 0.35

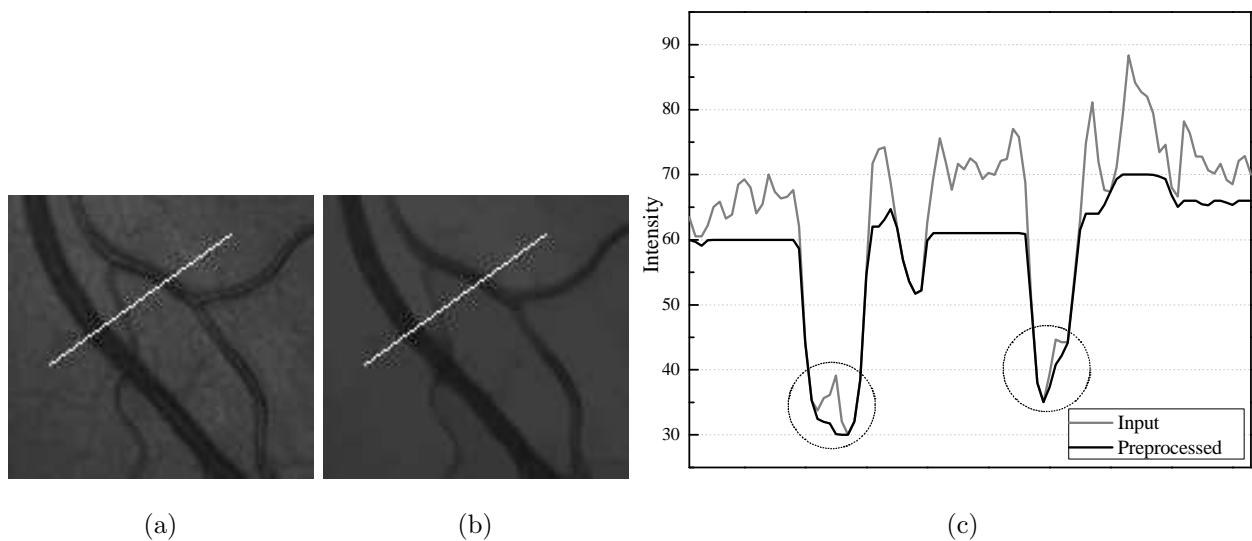


Figure 1: Preprocessing; (a) image with vessel reflex, (b) image after morphological reconstruction, and (c) intensity profile before and after preprocessing.

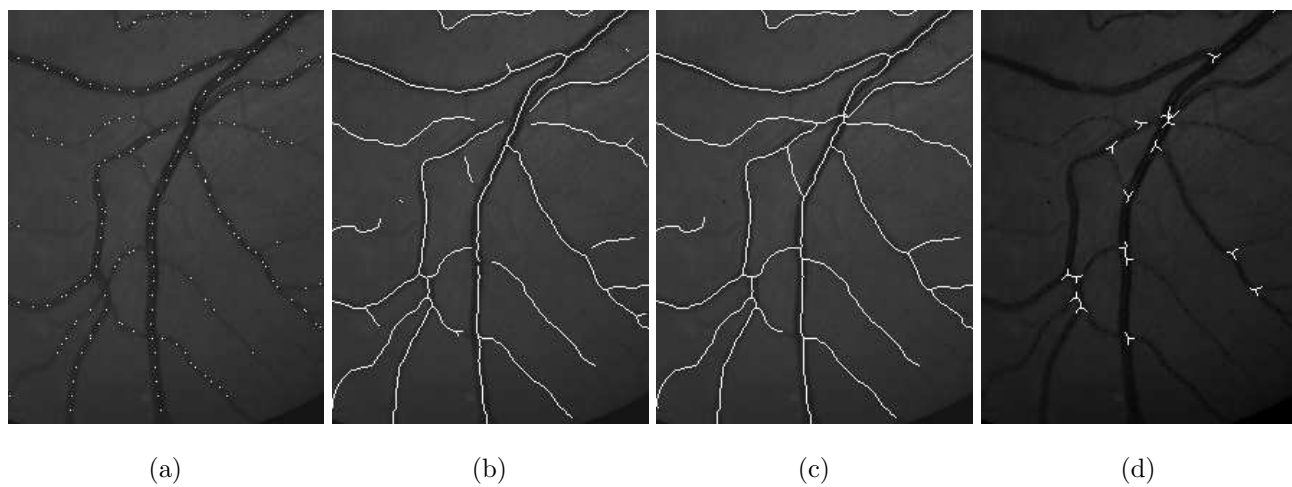


Figure 2: Feature detection: (a) seed detection, (b) vessel tracing, (c) endpoint extension, and (d) branch estimation. In all cases the lines and points are superimposed over the preprocessed input image.

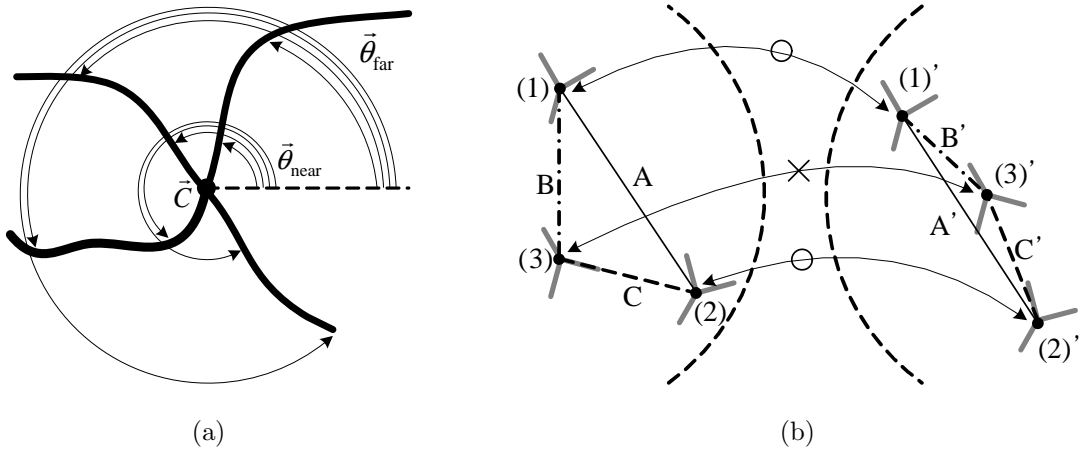


Figure 3: CP selection: (a) branch vector elements and (b) CP selection using branch vectors and spatial locations. In this example branch centers (1) and (1)', (2) and (2)', and (3) and (3)' matched as CPs. CPs (1)-(1)' and (2)-(2)' are selected to initialize the affine model.

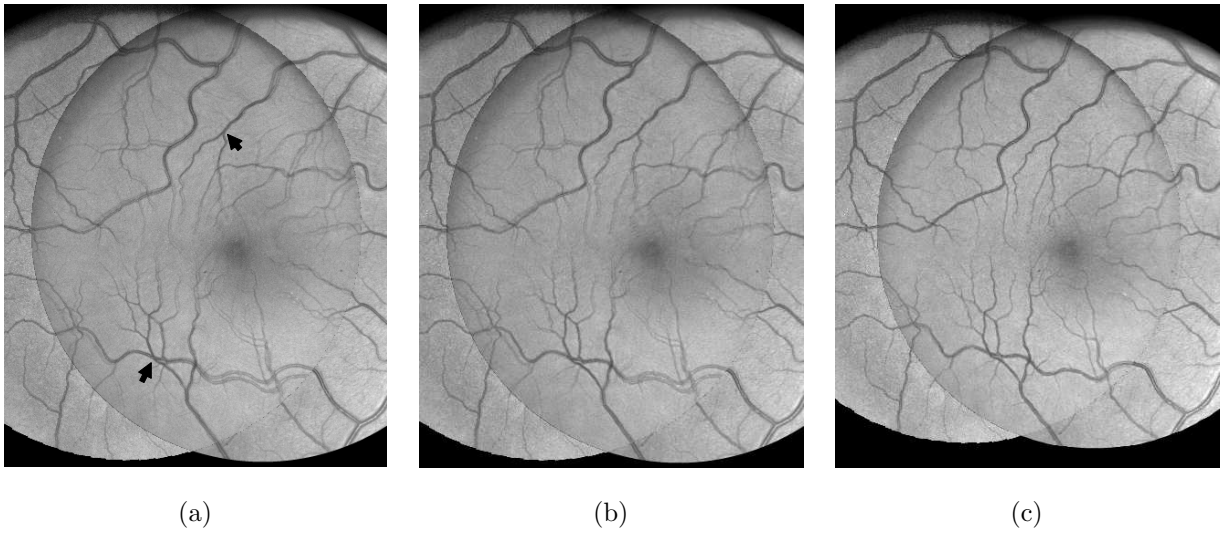


Figure 4: Registration results of retinal image using (a) similarity transform (CEM=2.85), (b) affine transform without distortion correction (CEM=2.56), and (c) proposed method (CEM=1.58, $k=-0.18$). CPs are marked by black arrows.