Motion Correction in Optical Coherence Tomography for Multi-modality Retinal Image Registration

Jun Cheng, Jimmy Addison Lee, Guozhen Xu, Ying Quan, Ee Ping Ong, and Damon Wing Kee Wong

Institute for Infocomm Research, A*STAR, Singapore

Abstract. Optical coherence tomography (OCT) is a recently developed non-invasive imaging modality, which is often used in ophthalmology. Because of the sequential scanning in form of A-scans, OCT suffers from the inevitable eye movement. This often leads to mis-alignment especially among consecutive B-scans, which affects the analysis and processing of the data such as the registration of the OCT *en face* image to color fundus image. In this paper, we propose a novel method to correct the mis-alignment among consecutive B-scans to improve the accuracy in multi-modality retinal image registration. In the method, we propose to compute decorrelation from overlapping B-scans and to detect the eye movement. Then, the B-scans with eye movement will be re-aligned to its precedent scans while the rest of B-scans without eye movement are untouched. Our experiments results show that the proposed method improves the accuracy and success rate in the registration to color fundus images.

Keywords: optical coherence tomography, registration, motion correction

1 Introduction

Optical coherence tomography (OCT) is a micrometer-scale, cross-sectional imaging modality for biological tissue[1]. With the advancement of the technology, current commercial machines such as ATLANTIS (Topcon Corporation, Japan) are able to provide a 3D volume scan in seconds. However, one challenge to get a good scan is the eye-movement which often leads to mis-alignment between consecutive B-scans. Fig. 1 illustrates the motion artifact, where Fig. 1(a) shows typical B-scans of a 3D scan in $y - z$ plane, Fig. 1(b) shows the first frame from $x - z$ plane and Fig. 1(c) shows the *en face* image in $x - y$ plane. Fig. 1(d) shows the color fundus of the same subject eye. Benefited from the fast scanning speed of the machine, the motion artifact of a B-scan in $y - z$ plane is usually small and can be ignored. However, the large time delay among consecutive B-scans makes the motion between consecutive B-scans non-ignorable. For example, the $x - z$

view of the data in Fig. 1(b) is significantly distorted because of vertical movement in z-axis. This will affect the use of information from consecutive B-scans of the data and mislead the diagnosis. Besides that, the movement in horizontal direction in y-axis can also be identified from the mis-matched blood vessels and disc boundaries in the *en face* image in Fig. 1(c). This poses some challenges in
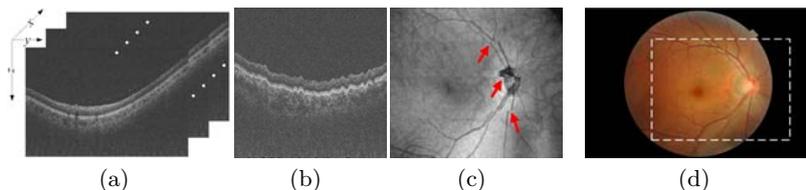

(a)	(b)	(c)	(d)

Fig. 1. Illustration of motion artifact: (a)B-scans view from $y - z$ plane of a 3D OCT (b) view from $x - z$ plane (c) *en face* image in $x - y$ plane (d) color fundus image

some applications, e.g., the multi-modality image registration between an OCT *en face* image and a color fundus image, which is important to ophthalmologists [2].

Motion correction in OCT starts to receive some attention in the last decade. Although additional hardware (e.g. eye tracker) can be used to acquire data to correct the motion artifact, it is not available in some machines including AT-LANTIS. Therefore, some software algorithms are proposed for motion correction. Zawadzki *et al.* [3] use autocorrelation method to align stack of B-scan slides. In the algorithm, each B-scan is used as the reference with respect to which the next B-scan is aligned. However, register every frame to its precedent scan leads to excessive registration especially for B-scans without eye movement. Ricco *et al.* [4] proposed to register the OCT *en face* image to an instantaneous artifact-free reference image. Xu *et al.* [5] proposed a particle filtering approach which is updated frame by frame to detect A-scan movement. Montuoro *et al.* [6] propose to correct the motion artifact using local symmetry. By using the symmetry, the method can effectively reduce the artifact in $x - z$ plane.

The main objective of this paper is to re-align the B-scans such that the OCT *en face* image has less motion artifact for easier registration to another modality such as color fundus. Therefore, we mainly focus on the motion artifact due to horizontal movement while the vertical movement in z-axis is not an issue after computing the *en face* image. The horizontal movement in y-axis introduces the mis-matched blood vessels and disc boundaries in the *en face* image and makes the registration between OCT *en face* image and image from other modality a difficult problem. A straightforward way is to register each B-scan to its precedent scan as in [3]. Taking B_i as a reference, each B-scan B_i is aligned to its precedent B-scan B_{i-1} such that their difference is minimized or the correlation is maximized. However, this leads to excessive registration for B-scans without eye movement. Unnecessary registration shall be avoided

as 1) the B-scan slice itself contains high amount of speckle noise which poses challenge to register them. 2) the consecutive B-scans are mostly similar except for the location of the vessels. The registration tends to align them such that the vessels overlaps exactly. However, this is not necessarily the case in reality. Therefore, excessive registration may lead to distortion of vessels. To overcome this, we need to detect the eye movement so that we can keep those B-scans without eye movement untouched.

The rest of paper is organized as follows. In Section 2, we introduce the method including motion detection using decorrelation from overlapping scans, the motion correction using block matching, and the application in multi-modality retinal image registration. Section 3 shows the experimental results and is followed by conclusions in the last section.

2   Method

2.1   Motion Detection

In order to conduct motion correction between and only between B-scans with eye movement, this paper proposes a method to detect the presence of motion and apply the motion correction when a motion is detected. Motivated by the fact that pixels from static tissue yield small decorrelation values while pixels from the faster moving tissue yield high decorrelation values [7][8], we propose to compute decorrelations from overlapping scans and use the *en face* decorrelation image. This implies that the proposed method is not suitable to data without any overlapping scan. Fortunately, this is not an issue in practice as overlapping scan has been used commonly to improve the image quality.

Decorrelation   Given a 3D scan V with N B-scan slices where each B-scan has M overlapping scans. For convenience, we denote the amplitude of these B-scans as $A_{i,j}, i = 1, \cdots N, j = 1, \cdots, M$. The decorrelation is computed as

$$D(x,y,z) = \frac{1}{M} \sum_{m=1}^{M-1} 1 - \frac{2A_{x,m}(y,z)A_{x,m+1}(y,z)}{A_{x,m}(y,z)^2 + A_{x,m+1}(y,z)^2} \tag{1}$$

After computing the deccorelation, we compute the *en face* image E of the 3D deccorelation volume D:

$$E(x,y) = \frac{1}{Z} \sum_{z=1}^{Z} D(x,y,z), \tag{2}$$

where Z is the number of pixels in each A-scan. Fig. 2 shows an example of decorrelation in volume and its corresponding 2D *en face* image.
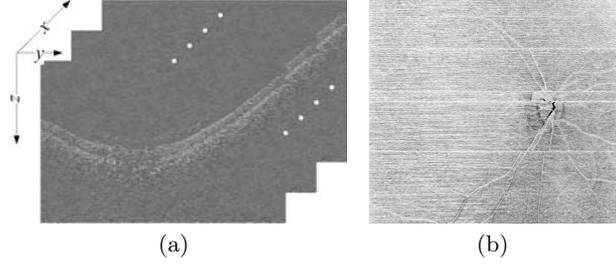
Fig. 2. Decorrelation Image: (a) 3D (b) *en face* image of (a)

White line detection Because of the eye movement, the deccorelation values become high, therefore, we observe a line or a few consecutive lines in white in the *en face* images as shown in Fig. 2(b). It can be seen that the locations of these white lines match well with those from the artifact in Fig. 1(c). Thus, we detect these white lines and use them as indicators of eye movement.

In order to do that, we compute the horizontal profile from the decorrelation *en face* image E as follows:

$$P(x) = \frac{1}{Y} \sum_{y=1}^{Y} E(y, x), \tag{3}$$

where Y is the number of columns in the *en face* image. Its gradient is then computed as

$$P_g(x) = P(x) - P(x-1) \tag{4}$$

We detect the local peaks from $P_g(x)$ as the position of the white line. In this paper, we use a simple thresholding approach. The xth line is a white line if $P_g(x) > T$. In our implementation, T is determined as

$$T = \mu_{P_g} + 1.5\sigma_{P_g}, \tag{5}$$

where μ_{P_g} and σ_{P_g} denote the mean and standard deviation of P_g.

2.2   Motion Correction

After detecting a B-scan B_i that has been affected by eye movement, we align it to its precedent scan B_{i-1}. We search for the alignment $(\Delta x_i, \Delta y_i)$ between B_i and B_{i-1} such that their difference is minimized, i.e., we compute $(\Delta x, \Delta y)$ which minimizes the following objective function:

$$\min_{\Delta x, \Delta y} \sum_{x,y} |B_i(x,y) - B_{i-1}(x + \Delta x, y + \Delta y)| \tag{6}$$

A full search can be used. However, it is time consuming. In our implementation, we use the iterative diamond search strategy [9] because of its efficiency. In the

searching procedure, the algorithm iteratively searches among the neighboring points for the alignment between B_i and B_{i-1} until the error in between is no longer decreasing. The diamond search algorithm employs two search patterns as illustrated in Fig. 3. The first pattern is called large diamond search pattern (LDSP). It comprises nine checking points to compose a diamond shape ($\diamond$). The second pattern is called small diamond search pattern(SDSP), consisting of five checking points forms in a diamond shape as well. In the searching procedure, LDSP is iteratively conducted until the step in which the minimum difference determined by (6) occurs at the center checking point. The search pattern is then switched from LDSP to SDSP. Among the five checking points in SDSP, the position yielding minimum difference determines the best matching between B_i and B_{i-1}. In the proposed method, we apply the above motion correction when we detect a white line in the *en face* decorrelation image.
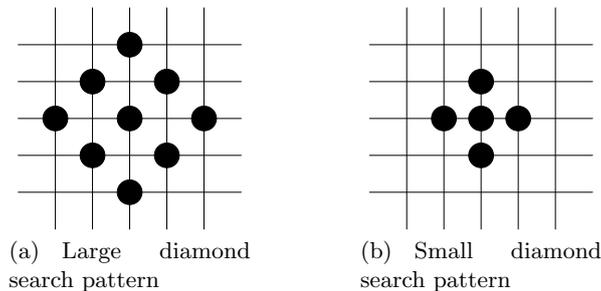

(a) Large diamond search pattern

(b) Small diamond search pattern

Fig. 3. The search patterns.

2.3   Multi-modality Registration

In order to quantitatively justify the benefit of the proposed decorrelation based motion correction method, we need to compare the *en face* OCT image with a ground truth. Although a straightforward way to obtain such a ground truth is to conduct a scan without eye movement, however, we find it very challenging in practice. Since our objective for motion correction is to improve the multi-modality registration, we use color fundus image where the image is captured instantaneously instead of a few seconds (or longer) in OCT scan as a reference and register the OCT *en face* image to the color fundus image. By comparing some key points between the OCT *en face* image with the color fundus, we are able to quantitatively evaluate how the proposed method reduce the eye movement. The registration results also reflects the benefit of the proposed motion correction for the registration. Registering OCT *en face* image with color fundus image is an important application. Therefore, such an evaluation has practical meaning.

In this paper, we register the OCT *en face* image to the color fundus using the matching method based on the Low-dimensional Step Pattern Analysis (LoSPA) [2]. LoSPA is the state-of-the-art key points based image registration algorithm. It has been proven to be superior than traditional scale invariant feature transform based algorithms [10]. A very brief introduction of LoSPA method is given here while details of the algorithm can be find in [2]. In the LoSPA method, geometric corners are first extracted from edges that form a corner in both the OCT *en face* and color fundus images. Then, the intensity change patterns based on 28 pre-defined step patterns are described as features. In the matching, the geometric corners are matched by Euclidean distance. Finally, random sample consensus (RANSAC) [11] is used with an affine transformation to produce the final registration.

3   Experimental Results

We conduct comparative experiments on a dataset comprising of 18 3D OCT scans and the corresponding color fundus images from different subject eyes. The 3D OCT data were obtained from a Topcon DRI OCT-1 machine using a customized mode: each 3D scan has 900 B-scans composed of 225 (N=225) slices with four overlapping scan (M=4) captured from a $12mm \times 9mm \times 2mm$ cube of the retina. Each B-scan slice has 992×256 pixels. Therefore, the OCT *en face* image has a dimensionality of 256×225. The color fundus photographs were acquired with a TRC-NW8 non-mydriatic fundus camera. The color fundus images are resized from 3216×2316 pixels to 536×366 as the OCT *en face* images have a much smaller size. In order to evaluate the accuracy of motion correction and also how it improves the registration, we register the OCT *en face* images to the reference color fundus images. We manually select 8 pairs of corresponding points in each OCT *en face* image and color fundus image pair to generate the ground truth. We tried our best to select the points uniformly distributed in the wider area of the images. After conducting the motion correction and the registration, we compute the root-mean-square error (RMSE) between the corresponding points in each registered image pair. For successful registration, we consider the RMSE < 5 pixels in proportion to the image resolution in [10]. In addition, a significant error such as the maximum absolute error (MAE) >10 pixels [12] also results in a registration failure.

Table 1. Performance comparison in registration

	RMSE	MAE	Median Error	Success Rate
No Correction	8.4	21.8	6.1	8/18
Zawadzki's Method [3]	5.1	11.8	4.2	12/18
Proposed	4.3	10.3	3.4	14/18

To justify the effectiveness of the proposed method, we compare the registration results with and without motion correction. In addition, we have also compared our method with the case where every B-scan is aligned with precedent scan in [3]. Table 1 summarizes the RMSE, MAE, median error and the success rate in registration. Results show that the proposed method achieves lowest RMSE, MAE and median error. As a result, the successful registration increased from 8 pairs without motion, 12 pairs with Zawazki's method to 14 pairs by the proposed method out of all 18 pairs of images. Fig. 4 show three examples where the image registration fails without motion correction while succeeding after motion correction by the proposed method.
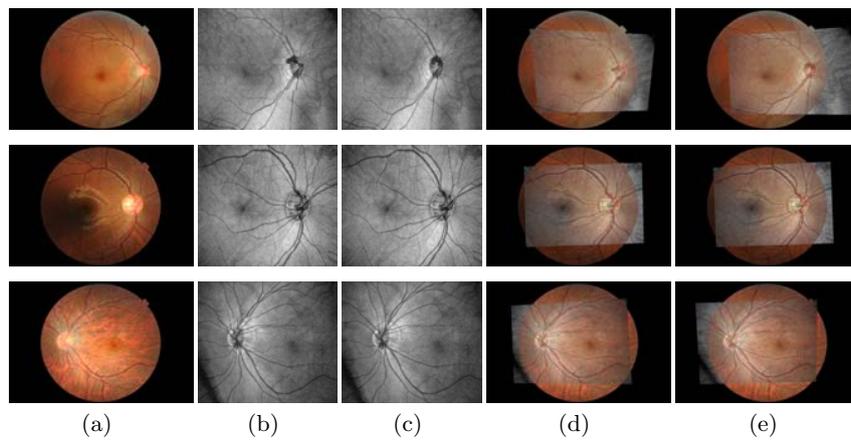

(a)   (b)   (c)   (d)   (e)

Fig. 4. Sample Results: (a) Color fundus; (b) *en face* image of original 3D OCT; (c) *en face* image of 3D OCT after proposed motion correction; (d) (a) registered with (b); (e) (a) registered with (c)

4   Conclusions

Motion artifact in 3D OCT is an important issue which might affect the interpretation of the data. In this paper, we propose to compute decorrelation from overlapping scans to detect the motion. A fast block matching search algorithm is applied to register the B-scan slides with motion to their precedent scans. The *en face* images are then computed from the motion corrected volume and the LoSPA method is used for the registration to color fundus image. Our results show that the proposed method is able to reduce the motion artifact and improve the registration between the OCT *en face* images and the color fundus images. A limitation of the method is that it requires at least two overlapping scans at each position to compute the decorrelation, therefore, it is not applicable to data

without overlapping scan. Fortunately, this limitation is not an issue in practice as overlapping is commonly used to improve the image quality.

References

1. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, and J. G. Fujimoto, "Optical coherence tomography," *Science*, vol. 254, no. 5035, pp. 1178–1181, 1991.
2. J. A. Lee, J. Cheng, G. Xu, E. P. Ong, B. H. Lee, D. W. K. Wong, and J. Liu, "Registration of color and oct fundus images using low-dimensional step pattern analysis," *N. Navab et al. (Eds.): MICCAI 2015, Part II, LNCS 9350*, pp. 214–221, 2015.
3. R. J. Zawadzki, A. R. Fuller, S. S. Choi, D. F. Wiley, B. Hamann, and J. S. Werner, "Correction of motion artifacts and scanning beam distortions in 3d ophthalmic optical coherence tomography imaging," *Proc. SPIE*, vol. 6426, pp. 642607–1– 642607–11, 2007.
4. S. Ricco, M. Chen, H. Ishikawa, G. Wollstein, and J. Schuman, "Correcting motion artifacts in retinal spectral domain optical coherence tomography via image registration," *G.-Z. Yang et al. (Eds.): MICCAI 2009, Part I. LNCS 5761*, pp. 100–107, 2009.
5. J. Xu, H. Ishikawa, G. Wollstein, L. Kagemann, and J. S. Schuman, "Alignment of 3-d optical coherence tomography scans to correct eye movement using a particle filtering," *IEEE Trans Med Imaging*, vol. 31, no. 7, pp. 1337–1345, 2012.
6. A. Montuoro, J. Wu, S. Waldstein, B. Gerendas, G. Langs, C. Simader, and U. Schmidt-Erfurth, "Motion artefact correction in retinal optical coherence tomography using local symmetry," *P. Golland et al. (Eds.): MICCAI 2014, Part II, LNCS 8674*, pp. 130–137, 2014.
7. Y. Jia, J. C. Morrison, J. Tokayer, O. Tan, L. Lombardi, B. Baumann, C. D. Lu, W. Choi, J. G. Fujimoto, and D. Huang, "Quantitative oct angiography of optic nerve head blood flow," *Biomed. Opt. Express*, vol. 3, no. 12, pp. 3127–3137, 2012.
8. Y. Jia, O. Tan, J. Tokayer, B. Potsaid, Y. Wang, J. J. Liu, M. F. Kraus, H. Subhash, J. G. Fujimoto, J. Hornegger, and D. Huang, "Split-spectrum amplitude-decorrelation angiography with optical coherence tomography," *Opt. Express*, vol. 20, no. 4, pp. 4710–4725, 2012.
9. S. Zhu and K. Ma, "A new diamond search algorithm for fast block-matching motion estimation," *IEEE Trans. Image Processing*, vol. 9, pp. 287–290, 2000.
10. Z. Ghassabi, A. Sedaghat, J. Shanbehzadeh, and E. Fatemizadeh, "An efficient approach for robust multimodal retinal image registration based on ur-sift features and piifd descriptors," *EURASIP Journal on Image and Video Processing*, , no. 1, pp. 1–16, 2013.
11. M.A. Fischler and R.C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Comm. ACM*, vol. 24(6), pp. 381–395, 1981.
12. J. Chen, J. Tian, N. Lee, J. Zheng, R. T. Smith, and A. F. Laine, "A partial intensity invariant feature descriptor for multimodal retinal image registration," *IEEE Trans. Biomedical Engineering*, vol. 57(7), pp. 1707–1718, 2010.