3D ORTHOGONAL SD-OCT VOLUMES REGISTRATION FOR THE ENHANCEMENT OF PORES IN LAMINA CRIBROSA

Nan Ding, Florence Rossant, Hélène Urien, Jérémie Sublime, Michel Paques

3D ORTHOGONAL SD-OCT VOLUMES REGISTRATION FOR THE ENHANCEMENT OF PORES IN LAMINA CRIBROSA

Nan Ding[⋆] *Florence Rossant*[⋆] *Hélène Urien*[⋆] *Jérémie Sublime*[⋆] *Michel Paques*[†]

[⋆] Institut Supérieur d'Electronique de Paris (ISEP), Paris, France
[†] Centre d'investigation Clinique 1423, INSERM & Direction de l'Hospitalisation et des Soins, Hôpital des Quinze-Vingts, Sorbonne Université, Paris, France

ABSTRACT

The lamina cribrosa (LC), a 3D porous structure of the eye in which optic fibers pass to reach the brain, is critically involved in the diagnosis and pathogenesis of glaucoma, a leading cause of blindness. However, segmenting the LC pores in 3D images is a task that has rarely been investigated, mainly because of the very low signal to noise ratio. To address this problem, we propose a fully automatic method to register and fuse two 3D SD-OCT volumes acquired in orthogonal directions, in order to improve the OCT image quality and thus facilitate the pore detection. Our method relies on *a priori* knowledge about *en-face* images to enhance useful pore information and initialize the process. Moreover, the optimization of the cross-correlation function is carried out in several stages in order to ensure the robustness of the process and the accuracy of the result. The quantitative evaluation shows that the proposed method is efficient to register the two 3D orthogonal OCT volumes and the annotated pores, with a distance between aligned pores around 3 pixels (below the pore size).

Index Terms— image processing in ophthalmology, optical coherence tomography, SD-OCT, lamina cribrosa, 3D image registration, image fusion, pore enhancement.

1. INTRODUCTION

The lamina cribrosa (LC) is a 3D collagenous mesh in the optic nerve head (ONH) that plays a crucial role in the mechanisms and diagnosis of glaucoma [1], one of the leading causes of blindness. The LC is composed of so-called "pores", namely axonal paths within the collagenous mesh, through which the axons pass to reach the brain [2]. In-vivo 3D observation of the pores is now possible thanks to advances in Optical Coherence Tomography (OCT) technology (Fig. 1). Our final aim is to automatically perform the 3D reconstruction of axonal paths from OCT volumes, in order to study the remodeling of the lamina cribrosa during glaucoma and better understand this disease.

Recent studies are mainly limited to the detection of the LC thickness [3] and the surface of the LC [4]. Morphological changes in pores, such as elongation, have been observed in glaucoma patients in 2D Adaptive Optics [5]. Few groups, including ourselves, have proposed methods to detect the pores of the LC in 3D and in vivo [6, 7], thanks to OCT technology. However, the limited transverse resolution of conventional OCT (∼15µm) as well as the bad signal to noise ratio (SNR) did not allow the characterization of all axonal paths with enough reliability, knowing that it is difficult even for experts to identify the pores in a single *en-face* image. Therefore, in this article, we propose to exploit the fact that OCT volumes can be acquired along two different scanning directions, horizontal or vertical (Fig. 1), to get one fused volume of better image quality, and finally a better segmentation. For that, we first need to register the two orthogonal volumes.

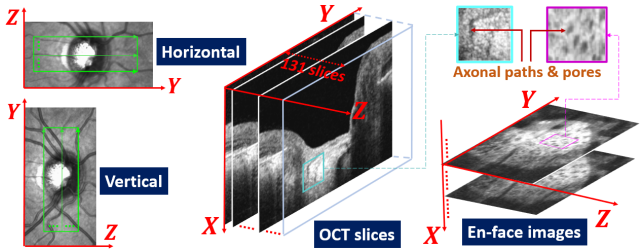


Fig. 1: 3D orthogonal OCT volumes. Each volume of 131 slices is obtained by a(n) horizontal/vertical scan, with a resolution of 7µm/pixel along the scanning direction and a sampling step of 15µm. *En-face* images (cropped in the illustration) are extracted from the OCT slices, revealing the LC pores as dark spots.The X axis corresponds to the depth in the ONH (4µm/pix).

Methods have already been proposed to register OCT images: for noise reduction [8, 9], motion estimation [10, 11], and mosaicing [12, 13]. But, to the best of our knowledge, general methods to register orthogonal OCT volumes (Fig. 1) for revealing more detailed information on the pores, have never been investigated. Aligning such OCT volumes is challenging for several reasons: the SNR is very low, local shifts may occur because of the scanning process, the structures of interest are of small size and low contrast. All of those result in poor cross-correlation. Moreover, heterogeneous qualities of the orthogonal volumes and intensity variations increase the difficulty. Finally, the volumes must be aligned in 3D, which requires robust global optimization algorithms.

Fig. 2: Top-level block diagram showing the stages of the algorithm. The notation is explained in Section 3.

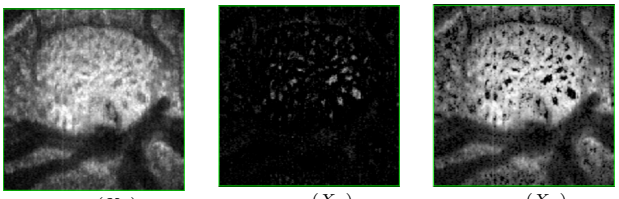
In this paper, we present a method to register orthogonal OCT volumes to obtain a better resolution of the LC pores. This method is based on pixel intensity and is divided into several steps to solve the optimization problem in a robust way and finally obtain an accurate registration. First, we rely on *a priori* knowledge on the pore appearance in *en-face* images to delimit the area of interest and enhance useful information. Then, several geometric transforms are defined and applied sequentially, each optimizing a given cross-correlation function, to automatically and accurately align the 3D OCT data. A fused volume is calculated afterwards.

2. 3D OCT DATA

Our dataset consists of 21 3D orthogonal OCT volumes, acquired from 20 eyes of 14 subjects in both transverse directions (42 volumes in total), all with the same Heidelberg Spectralis SD-OCT machine and according to the same scanning protocol. We call the two orthogonal volumes horizontal/vertical scans respectively (Fig. 1). Each acquisition contains 131 2D OCT slices of 496×768 pixels, the pixel size being approximately 7µm×4µm and the sampling step between two consecutive slices 15µm. *En-face* images of 768×131 pixels are obtained from such 2D OCT slices in both direction during the same examination. They are resized by a factor 2 along the second axis to restore almost square pixels (768×262 pixels). We denote the gray levels of intensity in horizontal and vertical OCT volumes by $I_H(x,y,z)$ and $I_V(x,y,z)$, with $x \in [1,496], y \in [1,768], z \in [1,262]$, and the corresponding *en-face* image extracted at the X coordinate from I_H or I_V by $I_H^{(X)}(y,z)$ and $I_V^{(X)}(y,z)$, respectively. Both are also denoted as $I^{(X)}(y,z)$ for simplicity.

3. METHOD

The pipeline for the registration process of the orthogonal OCT volumes is shown in Fig. 2. In the initial stage, pore features are enhanced, which is important since the volumes are of low SNR. For the same reason, we define a reference *en-face* image in each volume, namely the one offering the best contrast in pores, and we use this reliable information to initialize a registration process by translation. Finally we further register the 3D volumes by affine transformation and fuse the aligned volumes.



(a) $I^{(X_0)}$ (b) $I_{ASF}^{(X_0)}$ (c) $I_{enh}^{(X_0)}$

Fig. 3: ROI of the reference *en-face* image before (a) and after (c) enhancement based on alternate sequential filtering (b).

3.1. Pre-processing

Region of interest (ROI) in *en-face* images. The LC is located deep in the optic disc, it can be observed in the central area called the cup. On en-face scans, the cup appears as a dark connected component in top *en-face* images (small X coordinate, Fig. 1) that can be automatically detected by thresholding. The ROI is defined as the 262×262 pixels bounding this region, and the detectable part of the LC is directly below this component (Fig. 3). We denote by $ROI(I_H^{(X)})$ and $ROI(I_V^{(X)})$ the regions of interest extracted at the X coordinate from horizontal and vertical volumes (Fig. 3).

Pore feature enhancement. Pore features are enhanced by a series of alternate sequential filters (ASF) defined by closing ($\bullet$) and opening ($\circ$) operations with increasing size of structuring element D_i (up to $i = 9$):

$$I_{open}^{(X)} = R^D_{I_{asf}^{(X)(i-1)}}(I_{asf}^{(X)(i-1)} \circ D_i), \quad I_{asf}^{(X)(0)} = I^{(X)}$$

$$I_{asf}^{(X)(i)} = R^E_{I_{open}^{(X)}}(I_{open}^{(X)} \bullet D_i)$$

$$I_{ASF}^{(X)} = 1 - min(max(I_{asf}^{(X)(9)} - I^{(X)}, 0), 1)$$

$$I_{enh}^{(X)} = C(I^{(X)} - \alpha I_{ASF}^{(X)})$$

where $R_M^D(I)$ and $R_M^E(I)$ are the morphological reconstruction by dilation and by erosion of I in the mask M, and the transformation C allows to keep the mean and the standard deviation of the intensity levels in $I^{(X)}$. As shown in Fig. 3, $I_{ASF}^{(X)}$ allows keeping the main dark structures (i.e. pores) while denoising, and the enhanced image $I_{enh}^{(X)}$ helps better visualize the pores ($\alpha = 0.5$).

Reference image extraction. The reference plane in a volume is selected as the *en-face* image where the pores are the most contrasted. We rely on *a priori* knowledge of the size and shape of pores to create a binary mask of the dark spots that are likely to be pores. To this end, we first apply a sliding window along the X axis and average the images $I_{ASF}^{(x)}$ on 5 consecutive *en-face* planes to get the $\overline{I}_{ASF}^{(x)}$ images with better SNR. Then, the OTSU algorithm [14] is applied to threshold the image. Morphological filters F_{morp} remove connected components that may not be pores, those with an area smaller or larger than a disk of radius 3 to 9. The resulted image is multiplied by $I_{ASF}^{(x)}$, and finally all pixel intensity values in the current *en-face* image are summed up in $E(x)$:

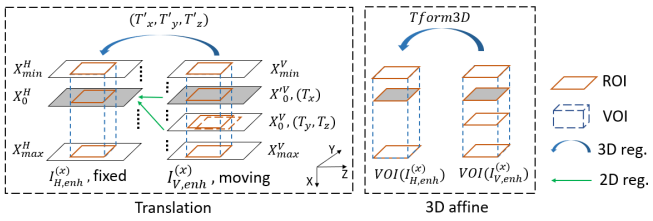


Fig. 4: Registration process.

$$E(x) = \sum_{y,z} \{ I^{(x)}_{ASF}(y,z) * F_{morp}[OTSU(\bar{I}^{(x)}_{ASF}(y,z))] \}$$

The reference plane X_0, namely X_0^H or X_0^V, is the one optimizing this cost function. Thresholding $E(X)$ at $0.25E(X_0)$ enables us to define the vertical area $[X^H_{min}, X^H_{max}]$ (resp. $[X^V_{min}, X^V_{max}]$) where pores are visible.

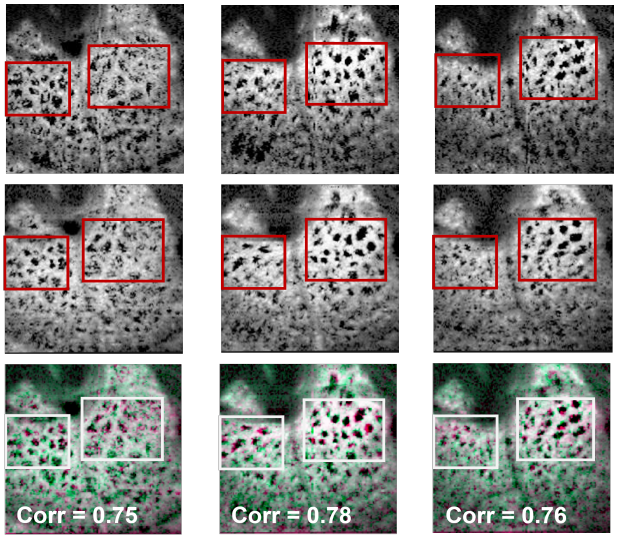
3.2. Registration

Since volumes to be registered are from the same sensor (i.e. SD-OCT), linear geometric transforms, namely affine transforms, have been chosen in this work as they are adequate and commonly used with mono-modal intensity based problems [15]. The translation is used to get a coarse but robust alignment, that is then used as initialization for refinement with an affine transform, as shown in Fig. 4. We use the cross-correlation of pixel intensity as the similarity metric for all transforms since it is adequate in mono-modal problems. Given I_H as the fixed volume, we look for transforms that align I_V to I_H.

Translation. We divide the translation alignment into 2 steps (Fig. 4, left):

- ROI of enhanced reference images, denoted by $ROI(I^{X_0^H}_{H,enh})$ and $ROI(I^{X_0^V}_{V,enh})$, are aligned by searching for the optimal translation vector (T_y, T_z) in axes Y and Z. Then, X_0^V is updated to $X_0^{'V}$ by seeking for a translation T_x in the axis X that maximizes the cross-correlation with $ROI(I^{X_0^H}_{H,enh})$. This process is iterated until the cross-correlation score cannot be improved anymore. At the end, we get a new reference plane for I_V denoted by $X_0^{'V}$.
- We define sub-volumes of same size, given the aligned ROIs in en-face images, the vertical intervals $[X^H_{min}, X^H_{max}]$, $[X^V_{min}, X^V_{max}]$ and the reference planes X_0^H and $X_0^{'V}$. The alignment of the resulting enhanced sub-volumes of interest, $VOI(I_{H,enh})$ and $VOI(I_{V,enh})$, is refined by translating again around (T_x, T_y, T_z).

This method has proven to be very robust: we first use the most reliable information, found in 2D images where the pores are numerous and contrasted, to approach the solution, and we refine the translation by considering a volume of interest. In this way, the problem of falling in a local minima of the cross-correlation function is overcome. As already highlighted, the pre-processing steps are crucial to enhance the useful information and focus on VOIs.



(a) $I^{(X_0-20)}_{enh}$ (b) $I^{(X_0)}_{enh}$ (c) $I^{(X_0+20)}_{enh}$

Fig. 5: Registration results showing the ROI of different *en-face* planes (X_0, $X_0 \pm 20$) using our method. From top to bottom: fixed horizontal planes, moving vertical planes, and the registration results. Correlation scores after registration are shown for each plane.

3D affine transform. The 3D translation transform alone is not enough since the optical axis may vary from one acquisition to another. To address this issue, we add the affine transform ($Tform3D$, Fig. 4, right), which includes translation, scaling, rotation and shearing, to further align the sub-volumes of interest. Regular step gradient descent is used for the optimization. This additional step has also proven to be very useful. Indeed, the cross-correlation is not only globally increased in the VOI, but we also verified experimentally that every 2D cross-correlation score of corresponding *en-face* planes is increased as well, which signifies a good local alignment of pores (Fig. 5). Finally, the registered volumes are fused by average, leading to *en-face* images of higher resolution (around 7µm×7µm) and better contrast (Fig. 6).

4. EXPERIMENTS

To be able to quantitatively assess the pore alignment performance, we asked a medical expert to annotate the main pore paths, in both horizontal and vertical source volumes, those that allow most axons to pass, since exhaustively annotating them is unrealistic because of the low SNR. A manually annotated pore path begins with a point in the reference plane, and continues in the neighboring planes by checking the continuity. We applied the found geometric transform ("Ours") to the annotated pores in vertical volumes and then, for each en-face image, we calculated the distance ($MinDist$) between each pore in the horizontal volume and its counterpart in the vertical registered volume. The smaller this distance, the better

Table 1: Evaluation of our method with respect to the GT. Statistics of the first line represent global registration performances on the VOI with $X \in [X_{min}, X_{max}]$, while the last 3 lines are calculated on the ROI with different X. "Trans." and "Affine" columns refer to the performances after the translation and affine transform stages.

		GT		Ours	
		Trans.	Affine	Trans.	Affine
VOI	Corr	0.59±0.09	0.70±0.06	0.69±0.11	**0.71±0.09**
	MinDist	3.59±0.65	**3.05±0.73**	4.82±0.67	3.73±0.49
X_0-20	Corr	0.56±0.09	0.68±0.10	0.69±0.07	**0.71±0.06**
	MinDist	3.99±1.03	**3.20±1.18**	5.08±1.43	4.22±0.77
X_0	Corr	0.63±0.06	0.71±0.07	0.72±0.08	**0.78±0.06**
	MinDist	3.41±0.18	**2.53±0.54**	3.95±1.09	3.16±0.70
X_0+20	Corr	0.54±0.13	0.68±0.10	0.69±0.08	**0.70±0.08**
	MinDist	4.13±0.82	**3.63±0.99**	6.42±1.51	5.12±1.23

the alignment. We also construct the ground truth registration ("GT") by aligning the annotated paths. In this case, the optimized criterion is the distance between the annotated paths in both volumes, rather than the cross-correlation ($Corr$) of the data. Then we applied the found transform to the vertical enhanced volume and calculated the cross-correlation scores. Table 1 displays the metrics obtained with Matlab on all 21 processed volumes at each stage of registration process (translation only, translation + affine transform). We present the statistics on the VOI (first line) and detail on 3 *en-face* planes.

We first analyse the results obtained with our method ("Ours"). Considering the VOI and the final registration, high cross-correlation scores between aligned sub-volumes are observed, with low standard-deviation (0.71±0.09). As expected, cross-correlation scores are the highest around the automatically selected reference images, where there is the most information with many contrasted pores. But we also obtain good matches in distant planes close to the X_{min} and X_{max} abscissas, in spite of the worse SNR. The distances between the aligned annotated pores are low (3.73±0.49) compared to the typical pore sizes (from 5 to 20 pixels, average around 10), which validates the proposed method in terms of robustness and accuracy. All results confirm that the 3D affine transform leads to better alignment compared to the translation (lower $MinDist$ and higher $Corr$). Fig. 6 illustrates these observations: the pores are well aligned in the fused image (see especially the areas in rectangles), better with the affine transform than only with the translation, and the green crosses are well inside the dark spots in the fused image (see also Fig. 7(a)).

Considering now the registration based on the annotated pores ("GT"), we also observe that the affine transform improves the registration on both criteria, cross-correlation and pore distance. The distance between the registered paths is higher with the proposed method ("Ours") since the optimization relates to the cross-correlation of the image data. However, the averaged distances are not far from those of the ground truth (3.73±0.49 against 3.05±0.73) and remain below the typical pore diameter (see Fig. 7).

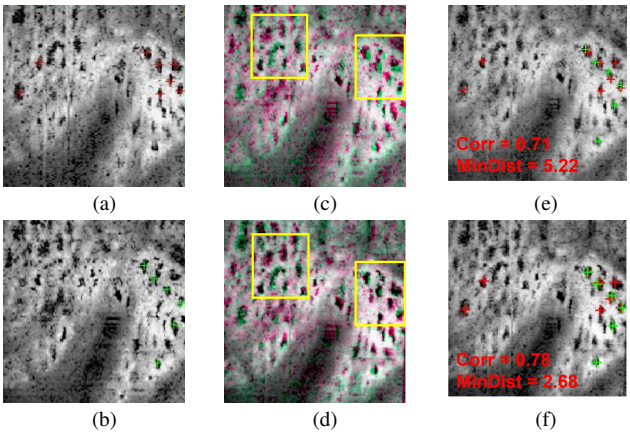


Fig. 6: Registration results of $ROI(I^{(X_0)})$ at each stage. (a) Enhanced horizontal reference image and with annotated pores (in red). (b) Enhanced vertical reference image with the annotated pores (in green). (c-d) Registration after the translation/the 3D affine transform. The registration performance is significantly improved after 3D affine transform, as can be observed in particular inside the yellow rectangle. (e-f) Fused image after the translation/the 3D affine transform, and the pore counterparts. Cross-correlation scores of *en-face* images and pore distances, are shown as well.

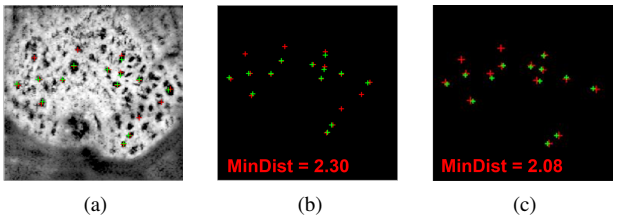


Fig. 7: Comparison between our method and the GT. (a) Fused image using our method. (b) Aligned pores with our method. (c) Aligned pores with transforms found from the GT.

5. CONCLUSION

We have proposed a fully automatic registration method to align two orthogonal OCT volumes acquired from the same OCT device and to enhance the pore features of the LC. Experiments performed on 21 annotated orthogonal volumes show that our algorithm is robust and leads to accurate alignment, with a mean distance between aligned pores around 3 pixels, below the pore size. The registration allows computing a fused image with better spatial resolution and higher pore contrast. For the first time, two orthogonal volumes can be jointly exploited, facilitating the medical interpretation. We plan now to train our context-aware attention U-net network [7] with the registered *en-face* images, in order to reach better segmentation results and, at the end, a better reconstruction of the 3D axonal for the longitudinal follow-up.

6. ACKNOWLEDGEMENT

This work was co-funded by the Institut de la Vision and ISEP (IHU M20JRAS004). The authors would like to thank the medical team of the Department of Ophthalmology III of the Quinze-Vingts hospital for the manual annotation.

7. COMPLIANCE WITH ETHICAL STANDARDS

This research was conducted retrospectively using OCT data collected in the Quinze-Vingts hospital. Written consent for reusing each data record was required from all participants involved in the study, and all OCT data were de-identified before used for the modeling. IRB approval was not required as confirmed by the doctors.

8. REFERENCES

[1] H. Quigley, E. Addicks, W. Green, and A. Maumenee, "Optic nerve damage in human glaucoma. ii. the site of injury and susceptibility to damage," *Archives of ophthalmology*, 1981.

[2] A. Dichtl, J. B. Jonas, and G. O. Naumann, "Course of the optic nerve fibers through the lamina cibrosa in human eyes," *Graefe's archive for clinical and experimental ophthalmology*, 1996.

[3] K. Omodaka, T. Horii, S. Takahashi, T. Kikawa, A. Matsumoto, Y. Shiga, K. Maruyama, T. Yuasa, M. Akiba, and T. Nakazawa, "3d evaluation of the lamina cribrosa with swept-source optical coherence tomography in normal tension glaucoma," *PloS one*, 2015.

[4] M. Tan, S. Ong, S. Thakku, C.-y. Cheng, T. Aung, and M. Girard, "Automatic feature extraction of optical coherence tomography for lamina cribrosa detection," *Journal of Image and Graphics*, 2015.

[5] A. Vilupuru, N. Rangaswamy, L. Frishman, E. Smith, R. Harwerth, and A. Roorda, "Adaptive optics scanning laser ophthalmoscopy for in vivo imaging of lamina cribrosa," *Journal of the Optical Society of America. A, Optics, image science, and vision*, 2007.

[6] F. Rossant, K. Grieve, S. Zwillinger, and M. Paques, "Detection and tracking of the pores of the lamina cribrosa in three dimensional sd-oct data," *International Conference on Advanced Concepts for Intelligent Vision Systems*, 2017.

[7] N. Ding, H. Urien, J. Sublime, F. Rossant, and M. Pâques, "Context-aware attention u-net for the segmentation of pores in lamina cribrosa using partial points annotation," *IEEE International Conference on Machine Learning and Applications*, 2022.

[8] M. R. Avanaki, R. Cernat, P. J. Tadrous, T. Tatla, A. G. Podoleanu, and S. A. Hojjatoleslami, "Spatial compounding algorithm for speckle reduction of dynamic focus oct images," *IEEE Photonics Technology Letters*, 2013.

[9] Z. Jian, L. Yu, B. Rao, B. J. Tromberg, and Z. Chen, "Three-dimensional speckle suppression in optical coherence tomography based on the curvelet transform," *Optics express*, 2010.

[10] M. F. Kraus, J. J. Liu, J. Schottenhamml, C.-L. Chen, A. Budai, L. Branchini, T. Ko, H. Ishikawa, G. Wollstein, J. Schuman *et al.*, "Quantitative 3d-oct motion correction with tilt and illumination correction, robust similarity measure and regularization," *Biomedical optics express*, 2014.

[11] 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 transactions on medical imaging*, 2012.

[12] H. C. Hendargo, R. Estrada, S. J. Chiu, C. Tomasi, S. Farsiu, and J. A. Izatt, "Automated non-rigid registration and mosaicing for robust imaging of distinct retinal capillary beds using speckle variance optical coherence tomography," *Biomedical optics express*, 2013.

[13] K. L. Lurie, R. Angst, and A. K. Ellerbee, "Automated mosaicing of feature-poor optical coherence tomography volumes with an integrated white light imaging system," *IEEE Transactions on Biomedical Engineering*, 2014.

[14] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, 1979.

[15] L. Zagorchev and A. Goshtasby, "A comparative study of transformation functions for nonrigid image registration," *IEEE transactions on image processing*, 2006.