

DEFORMABLE REGISTRATION OF MACULAR OCT USING A-MODE SCAN SIMILARITY

Min Chen,[†] Andrew Lang,[†] Elias Sotirchos,^{*} Howard S. Ying,[‡]
Peter A. Calabresi,^{*} Jerry L. Prince,[†] and Aaron Carass[†]

[†]Image Analysis and Communications Laboratory, The Johns Hopkins University

^{*}Department of Neurology, The Johns Hopkins School of Medicine

[‡]The Wilmer Eye Institute, The Johns Hopkins Hospital

[†]{mchen55,alang9,prince,aaron_carass}@jhu.edu ^{*}[‡]{esot,hying1,calabresi}@jhmi.edu

ABSTRACT

Optical coherence tomography (OCT) of the macular cube has become an increasingly important tool for investigating and managing retinal pathology. One important new area of investigation is the analysis of anatomic variability across a population. Such an analysis on the retina requires the construction of a normalized space, which is generally created through deformable registration of each subject into a common template. Unfortunately, state-of-the-art 3D registration tools fail to adequately spatially normalize retinal OCT images. This work proposes a new deformable registration algorithm for OCT images using the similarity between pairs of A-mode scans. First, a retinal OCT specific affine step is presented, which uses automated landmarks to perform global translations and individual rescaling of all the subject's A-mode scans. Then, a deformable registration using regularized one-dimensional radial basis functions is applied to further align the retinal layers. Results on 15 subjects show the improved accuracy of this approach in comparison to state of the art methods with respect to registration for labeling. Additional results show the ability to generate stereotaxic spaces for retinal OCT.

Index Terms— Optical coherence tomography, registration.

1. INTRODUCTION

There has been increased interest in optical coherence tomography (OCT) in recent years, in part because of the high resolution (typically 2.5-10 μ m), convenience, cost, and patient comfort of OCT imaging. Though a more important factor in its rising popularity is the direct correlations that can be seen from OCT measurements—i.e., retinal morphology and layer thickness—to the clinical status of patients, such as multiple sclerosis (MS) [2]. Atlasing of macular cube OCT images can be used to understand the deviations of a subject from the normal population. The construction of retinal stereotaxic spaces is dependent on the ability to deformably reg-

This work was supported by the NIH/NEI through grant 1R21EY022150-01 and by the National MS Society.

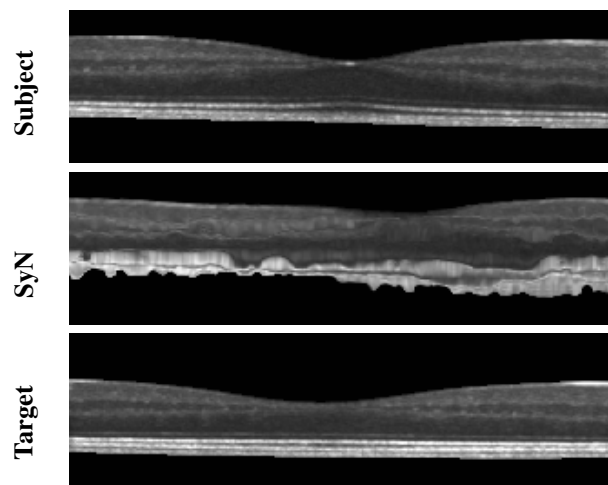


Fig. 1. Retinal OCT images used as subject and target and the result of deformably registering them using SyN [1].

ister OCT images. However, state of the art registration techniques, such as the symmetric inverse consistent diffeomorphic fluid based registration method SyN [1], perform poorly on retinal OCT data (see Fig. 1). There are several potential reasons for the failure of such registration algorithms: 1) Unlike computed tomography and magnetic resonance imaging where voxels are acquired mostly independently, OCT images are acquired as a large collection of one-dimensional reflectivity profiles in the acquisition direction (A-mode scans). Hence, voxels along these A-mode scans are highly dependent on the voxels before it, and interference at the start of a scan (such as a blood vessel) can dramatically change the entire A-mode scan; 2) the image slices (B-mode scans) are often separated by large gaps making correspondences between adjacent slices less reliable and, therefore, degrades the overall registration accuracy; 3) The field of view (FOV) is inconsistent across subjects, giving rise to boundary issues; 4) Noise in the vitreous humour and choroid can confound current registration techniques. Problem 2, can be partially addressed by upsampling the image to isotropic resolution. However, this makes the data size unmanageable, as a typ-

ical retinal OCT image would become on the order of 4GB in size. In addition, interpolation across such large gaps are typically unreliable. Problems 3 and 4 can be addressed by masking regions of interest; however this requires some *a-priori* knowledge and manual intervention.

In this work, we propose a registration algorithm that respects these limitations in OCT images. Rather than allowing the registration to deform freely across the image, our algorithm makes the assumption that, after a translation to align the fovea, accurate correspondences are best found within two A-mode scans. This assumption can be partially justified by the fact that, aside from the fovea, there are very few landmarks which can be used to find correspondences across subjects in a macular scan. This restriction is imposed by constraining the registration to find deformations in the direction of each A-mode scan, simplifying the problem into a series of registrations between every pair of A-mode scans in the subject and target images. An automated retinal boundary detection method is used to help the translation alignment of the foveae.

2. METHODS

2.1. Overview

Our proposed registration approach is divided into two steps. First, we perform a retinal OCT specific affine alignment (A-OCT) by translating the foveae between the two images and then individually scaling each subject A-mode scan to match the corresponding target A-mode scan. Then, a deformable registration (D-OCT) in the A-mode direction is used to further align the retinal layers.

Outside of the initial fovea alignment, the registration is performed by only comparing the similarity between pairs of A-mode scans. Hence, we index the images with the subscripts $\mathcal{S}_{(a,b)}$ and $\mathcal{T}_{(a,b)}$, where a and b refers to the location of an A-mode scan within a B-mode scan in the subject and target image, respectively.

2.2. Background Removal and Affine A-Mode Alignment

All images were preprocessed by first automatically removing the background above the inner limiting membrane (ILM) and below the Bruch’s membrane (BM) boundaries. This was performed by first filtering with a 1D Gaussian kernel having a standard deviation of 5 in the lateral direction, followed by a morphological closing operator with a vertical line structuring element of length 5. A vertical Sobel gradient filter was applied to extract the image gradient in the direction of each A-mode scan. The ILM is taken to be either the largest or second largest negative gradient along each A-mode scan, depending on which gradient is closest to the top of the image. The second largest gradient is also constrained to be at least 25 pixels from the largest gradient. The gradient closest to the bottom of the image is taken to be the inner segment-outer segment (ISOS) boundary. The BM is taken to be the largest positive gradient below the ISOS. Simple outlier detection and smoothing is then performed on the surfaces to

extract their final positions.

Once the background has been removed, the foveal centers are detected as the A-mode scans with the thinnest portion along the retina, denoted \mathcal{S}_f and \mathcal{T}_f for the subject and target respectively. We build a vector, c , that describes the perpendicular translation from \mathcal{S}_f to \mathcal{T}_f and apply it to all the subject A-Mode scans. This gives us the fovea aligned subject A-mode scans:

$$S_{(a,b)}^{\text{Trans}}(x) = S_{(a,b)+c}(x). \quad (1)$$

After the fovea alignment, for each individual pair of A-mode scans in the subject and target image, we perform a linear scaling of the subject scan such that the top and bottom boundaries are aligned with the target scan. This can be regarded as a landmark-based linear rescaling at each individual A-mode scan, where we use the outer edge of the ILM and BM layers as anatomical landmarks. If we denote the location of the ILM and BM boundaries in each target and subject A-mode scan as $t_{(a,b)}^I, t_{(a,b)}^B, s_{(a,b)}^I, s_{(a,b)}^B$, respectively, then the rescaled subject A-mode scans can be described as:

$$S_{(a,b)}^{(\text{A-OCT})}(x) = S_{(a,b)}^{\text{Trans}} \left((x - t_{(a,b)}^I) \frac{s_{(a,b)}^I - s_{(a,b)}^B}{t_{(a,b)}^I - t_{(a,b)}^B} + s_{(a,b)}^I \right). \quad (2)$$

We refer to this collection of preprocessing steps followed by affine A-mode alignment as A-OCT. Its purpose is to approximately align the A-mode scans between the subject and target images using linear transformations and to provide an initialization for the deformable registration.

2.3. Deformable Registration Using A-Mode Similarity

The goal of the deformable registration step (D-OCT) is to solve for a mapping $v_{a,b} : \mathcal{D}_{\mathcal{S}_{a,b}} \rightarrow \mathcal{D}_{\mathcal{T}_{a,b}}$ between each pair of subject and target A-mode scans after the A-OCT step, where \mathcal{D} is the respective domain of each A-scan. We restrict $v_{a,b}$ such that $I + v : \mathcal{D}_{\mathcal{S}_{a,b}} \rightarrow \mathcal{D}_{\mathcal{T}_{a,b}}$ is a one to one onto continuous map with a continuous inverse (i.e., a homeomorphism), where I is the identity transformation. This allows the transformation to preserve the topology of the image and prevent “folding” and “tearing” artifacts.

In our algorithm, the mapping $v_{a,b}$ is built as a summation of radial basis functions (RBFs), $\phi(x)$, a model which has been previously used in brain registration [3]. Since our registration is restricted to pairs of A-mode scans, this mapping is represented as a one-dimensional deformation field,

$$v(x) = \sum_i c_i \phi(x - x_i), \quad (3)$$

where c_i and x_i determine the size and center of each RBF, respectively. For $\phi(x)$, we chose the same RBF as presented in [3], which has properties such as smoothness, positive definiteness, and compact support. Our algorithm finds $v(x)$ by uniformly placing RBFs along the entire A-mode scan and then optimizing each c_i such that the similarity is maximized between the target and deformed subject A-mode scan. Sum of squared differences between the A-mode scans is used as

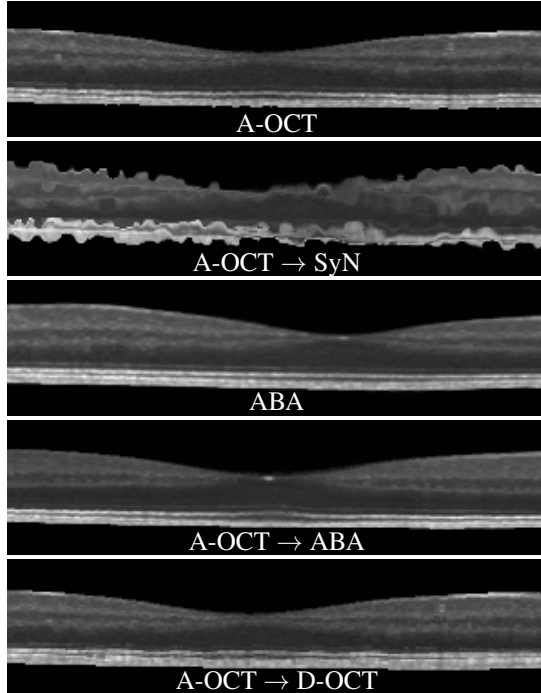


Fig. 2. Example results of various registration techniques. From top to bottom, the images are: our A-mode affine alignment A-OCT (Sec. 2.2); A-OCT followed by SyN registration [1]; ABA [3] registration without A-OCT; ABA registration after A-OCT; and A-OCT followed by our deformable registration D-OCT. See Fig. 1 for the subject and target images, and the results of using SyN without A-OCT.

the similarity measure in our algorithm, which gives us the energy function,

$$E_{(a,b)}^{\text{SSD}} = \sum_x \left(T_{(a,b)}(x) - S_{(a,b)}^{(\text{A-OCT})}(x + v_{(a,b)}(x)) \right)^2 \quad (4)$$

Optimizing the deformation field on independent pairs of A-mode scans can lead to discontinuities in the deformation, and potentially ignore useful neighboring information that can be used in the optimization. To address this, we introduce a regularization term,

$$E_{(a,b)}^{\text{Reg}} = \sum_{\substack{r=-R \\ r \neq 0}}^R \sum_x \frac{1}{r} \left(T_{(a+r,b)}(x) - S_{(a+r,b)}^{(\text{A-OCT})}(x + v_{(a,b)}(x)) \right)^2, \quad (5)$$

where R determines how many adjacent A-mode scans to use in the regularization. Note that we do not regularize across B-mode scans, which follows our premise that the large separation between B-mode scans provide poor correspondences for the registration. Naturally, for data that do not suffer from this limitation, this regularization can be easily extended to use B-mode scans as well. We combine both energy terms ($E_{a,b}^{\text{SSD}}$ and $E_{a,b}^{\text{Reg}}$), which we then use to optimize each RBF.

Name	Abbr.	
Retinal nerve fiber layer	RNFL	■
Ganglion cell layer and Inner plexiform layer	GCL+ IPL	■
Inner nuclear layer	INL	■
Outer plexiform layer	OPL	■
Outer nuclear layer	ONL	■
Inner segment	IS	■
Outer segment	OS	■
Retinal pigment epithelium	RPE	■

Fig. 3. List and partial examples of layers manually segmented in our fifteen subject cohort, and their abbreviations (Abbr.).

3. RESULTS

3.1. OCT Data

OCT volumes were acquired using a Spectralis OCT scanner (Heidelberg Engineering, Germany). Macular raster scans ($20^\circ \times 20^\circ$) consisting of 49 B-mode scans (each separated by $122 \mu\text{m}$) were acquired utilizing automatic real-time (ART) to increase image quality by averaging multiple images of the same location. Scans with $\text{ART} \geq 12$ (number of scans averaged) and with signal quality of at least 20 dB were used in this study. Each B-mode scan has 1024 A-mode scans (separated by $5.5 \mu\text{m}$), and each A-mode scan has 496 pixels (with $3.8 \mu\text{m}$ resolution). Each volume was manually segmented using an internally developed software and protocol; a list of the segmented layers is shown in Fig. 3.

3.2. Evaluation Against Manual Segmentations

Fifteen volumes with manual segmentations of eight of the retinal layers were used to evaluate our registration approach. Two of the images were randomly chosen as atlases and registered to the other thirteen images (26 registrations in total for each method). With the estimated deformation field, atlas labels were transformed to each target space and the Dice coefficient between each of the eight layers are reported. Our approach (with $R = 3$, chosen empirically) was compared to each of SyN [1] (a highly ranked algorithm for brain image registration [4]), ABA [3], and a directional constrained ABA (both available in the JIST software package [5]). In addition, all four methods were ran from both the initial background removed images and the images after our A-mode affine alignment (A-OCT), see Tab. 1 for results. The subject and target images are shown in Fig. 1, as well as the results of the SyN registration. Other registration results are shown in Fig. 2.

3.3. Average Atlas Construction

An important application of deformable registration is the ability to construct average atlases by registering a group of subjects into a common space [6]. Fig. 4 shows such an atlas created using our deformable registration approach. The registrations were performed relative to a randomly selected template. Although not ideal because of atlas bias, it demonstrates a proof of concept for the construction of retinal

Table 1. Dice overlap between segmentations from atlas registrations and the manual segmentation for eight retinal layers, averaged over 26 registrations (13 subjects \times 2 atlases). The deformable registrations were performed with either SyN [1], two variants of ABA [3], or our approach (D-OCT). We also ran each experiment with and without our A-mode affine alignment (A-OCT) to show the benefits of this preprocessing. See Fig. 3 for a list of the layer name abbreviations.

	<i>RNFL</i>	<i>GCL + IPL</i>	<i>INL</i>	<i>OPL</i>	<i>ONL</i>	<i>IS</i>	<i>OS</i>	<i>RPE</i>	<i>Mean</i>
SyN	0.21	0.22	0.11	0.11	0.17	0.06	0.09	0.14	0.14
A-OCT \rightarrow SyN	0.73	0.79	0.58	0.60	0.78	0.40	0.48	0.65	0.63
ABA	0.58	0.62	0.45	0.47	0.66	0.40	0.49	0.58	0.53
A-OCT \rightarrow ABA	0.84	0.82	0.57	0.59	0.83	0.54	0.69	0.80	0.71
Constrained ABA	0.64	0.70	0.51	0.53	0.73	0.46	0.55	0.65	0.60
A-OCT \rightarrow Constrained ABA	0.84	0.82	0.57	0.59	0.83	0.54	0.69	0.80	0.71
A-OCT	0.85	0.82	0.57	0.59	0.83	0.55	0.71	0.83	0.72
A-OCT \rightarrow D-OCT	0.86	0.84	0.65	0.66	0.86	0.67	0.75	0.84	0.77

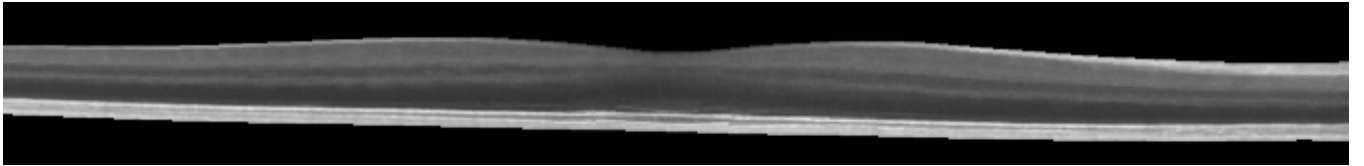


Fig. 4. Average atlas constructed using our deformable registration approach.

average atlases, which can be used for the study of population and pathology.

4. DISCUSSION AND CONCLUSION

We have proposed a new approach for performing deformable registration of retinal OCT. We show considerably higher accuracy when compared to existing 3D registration methods, when performing segmentation of the layers. In addition, we have introduced a simple and effective approach for performing affine scaling of OCT images to initialize deformable registration. Our results show this step dramatically improved the registration accuracy, regardless of which algorithm it was used with (see Tab. 1).

To extend this work, we intend to address the limitations from B-mode scan separations, such that the B-mode directions can be used in the regularization. In addition, we will construct an unbiased average atlas of the retina using an approach similar to that presented in [6]. Such an atlas will allow us to directly apply a number of existing computational anatomy tools, which will allow us to look at the variability of the retinal layers between individuals and study population differences across pathology.

5. REFERENCES

- [1] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, “Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain,” *Medical Image Analysis*, vol. 12, no. 1, pp. 26–41, 2008.
- [2] S. Saidha *et al.*, “Visual dysfunction in multiple sclerosis correlates better with optical coherence tomography derived estimates of macular ganglion cell layer thickness than peripapillary retinal nerve fiber layer thickness,” *Multiple Sclerosis Journal*, vol. 17, no. 12, pp. 1449–1463, 2011.
- [3] G. K. Rohde, A. Aldroubi and B. M. Dawant, “The Adaptive Bases Algorithm for intensity based nonrigid Image Registration,” *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1470–1479, 2003.
- [4] A. Klein *et al.*, “Evaluation of volume-based and surface-based brain image registration methods,” *NeuroImage*, vol. 51, no. 1, pp. 214–220, 2010.
- [5] B. C. Lucas, J. A. Bogovic, A. Carass, P.-L. Bazin, J. L. Prince, D. L. Pham, and B. A. Landman, “The Java Image Science Toolkit (JIST) for rapid prototyping and publishing of neuroimaging software,” *Neuroinformatics*, vol. 8, no. 1, pp. 5–17, 2010.
- [6] S. Joshi *et al.*, “Unbiased diffeomorphic atlas construction for computational anatomy,” *NeuroImage*, vol. 23, no. 1, pp. 151, 2004.