Stable Registration of Pathological 3D SD-OCT Scans using Retinal Vessels

Jing Wu[1], Bianca S. Gerendas[1], Sebastian M. Waldstein[1], Georg Langs[2], Christian Simader[3], and Ursula Schmidt-Erfurth[1]

[1] Christian Doppler Laboratory for Ophthalmic Image Analysis, Department of Ophthalmology and Optometry, Medical University of Vienna, Austria, ⋆
jing.a.wu@meduniwien.ac.at,
[2] Computational Imaging Research Lab, Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna, Austria

Abstract. We propose a multiple vendor registration method for pathological retinal 3D spectral domain optical coherence tomography volumes based on Myronenko's Coherent Point Drift and our automated vessel shadow segmentation. Coherent point drift is applied to the point sets comprising the landmarks generated by the vessel shadow segmentation method to generate the registration parameters required. In contrast to other registration methods, our solution incorporates a landmark detection and extraction method that specifically limits the extraction of false positives and is capable of handling any such noise in the registration point sets. Our experiments show better performance than current methods, with low relative distances between target and registered vessel point sets as well as agreeable correlation with ground truth.

1 Introduction

Spectral-domain Optical Coherence Tomography (SD-OCT) is a non-invasive modality for acquiring high resolution, 3D cross sectional volumetric images of the retina and the sub-retinal layers, in addition to retinal pathology. Today SD-OCT is the most important ancillary test for the diagnosis of sight degrading diseases such as age-related macular degeneration (AMD) and glaucoma [1]. Disease diagnosis, assessment, and treatment requires a patient to undergo multiple OCT scans, possibly using different scanning devices, to gauge disease activity, progression and treatment success. However, using OCT devices from different vendors, combined with patient movement may result in poor scan spatial correlation, potentially leading to incorrect patient diagnosis or treatment analysis. Image registration has been applied to solve this problem using iterative closest point (ICP) in [2] to compare disease states by registering different volumes. Accurate and reproducible landmarks are required to register 3D scans from

⋆ The Financial support of the Austrian Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development is gratefully acknowledged.

different timepoints and vendors, the most suitable being the retinal vasculature. The authors in [2] apply a landmark detection and point based registration method that is untested on wet-AMD (used here, caused by blood vessel leakage, opposed to the more common dry-AMD) and is validated on only a single scan vendor. As far as we are aware, other works in this area focus on OCT to fundus registration such as in [3, 4] where in addition, healthy cases were used.

In Sec.2 we present an automated retinal vessel segmentation method for landmark acquisition, less susceptible to pathology, and in Sec.3 perform registration using coherent point drift (CPD), suitable for point sets with limited overlap, similarity and the presence of noise. Our approach is evaluated in Sec.4 using scans from 3 major OCT vendors (Zeiss Cirrus, Nidek RS3000 & Topcon 3D 2000). We show that in addition to being suitable for multiple OCT vendors and performing reproducible registration (Fig.3), our method is accurate based on expert analysis of resulting registered retinal vessels and shows improved performance over current methods.

2 Automated Landmark Extraction

We define an OCT volume as $\mathbf{V}(Z, X, Y)$ were Z is the axial, X the primary and Y the secondary scan directions and is comprised of B-scans slices $\mathbf{B}_s(Z, X)$ (2D image perpendicular to fundus where X is the horizontal axis and Z is the vertical axis). We denote the retinal pigment epithelium (RPE) layer, the bright layer in the lower part of the retina, its lowest outer surface point $RPES_{min}$, the difference ($RPES_{diff}$) between $RPES_{min}$ and the original RPE surface position $RPES_{orig}$ for a given column point. The inner limiting membrane (ILM) is the upper most surface of the retina and we denote the transformations τ for affine and non-rigid registration as τ_A and τ_N respectively.

2.1 OCT Scan Pre-processing

Retinal vessels are not clearly visible directly within SD-OCT scans due to high attenuation of blood. However, shadows are generated as a result of the absorption of light transmitted from the scanner, visible most prominently in the RPE layer and can be used to locate the retinal vessels [5]. Two major steps are required to prepare the scans for vessel segmentation, firstly each $\mathbf{B}_s$ is "flattened" to accurately and reproducibly locate the RPE layer and secondly a composite "projection" image $\mathcal{P}$ is generated from the RPE layers.

Flattening adjusts a given $\mathbf{B}_s$ in the Z axis such that the lower RPE surface has the same Z value across the X axis, compensating for varying retinal curvatures across different $\mathbf{B}_s$. Using the RPE surface segmentation by the vendor system, each column in $\mathbf{B}_s$ is shifted in the positive Z direction by $RPES_{diff}$, flattening the RPE surface as seen in Fig.1(a).

A cross-section (RPE_{CS}) is taken from the flattened $\mathbf{B}_s$ covering 10 pixels from the outer RPE surface towards the ILM, based on examination of RPE

thickness from 21 training scans across 3 vendors ensuring the clinically significant RPE is included. Averaging the column intensities within RPE_{CS} improves vessel shadow visibility [5] resulting in a single line profile (RPE_{LP}) representing the RPE layer for a given $\mathbf{B}_s$. This process is repeated for each $\mathbf{B}_s$ creating a composite image $\mathcal{P}(X, Y)$ (where Y is the vertical axis as in the fundus) from all RPE_{LP} ranging in dimensions from 512×128 to 256×256 pixels.

2.2 Projection Denoising

Each RPE_{LP} discussed in Sec.2.1 features a degree of speckle noise [6] which hinders the detection system. A block matching based sparse transform domain collaborative filtering approach [7], applied to $\mathcal{P}$, has shown to be effective. In their two step method, an initial basic estimate is calculated by first taking block-wise estimates where, for each block within the image, grouping of similar blocks to the current block generates a 3D array. Collaborative hard-thresholding is applied to the 3D transformed array coefficients to attenuate the noise. Estimates of the grouped blocks are obtained by inverting the 3D transform which are then returned to their original locations within the image. The basic estimate is computed by weighted averaging ($w_{x_R}^{ht}$) of all the block-wise estimates that overlap. In step two, a final estimate is calculated from the basic estimate. For each block, two 3D arrays are created using block-matching, one from the noisy image and one from the basic estimate. A 3D transform is applied on both arrays in addition to a Wiener filter applied to the noisy array using the energy spectrum of the basic estimate. Estimates of all grouped blocks are generated from the inverse 3D transform on the filtered coefficients before returning the block estimates to their original positions within the image. Finally a weighted average ($w_{x_R}^{wie}$) is used to aggregate all the local estimates to compute the final estimate ($\hat{y}^{final}$) as described by Eqn.1 where $\chi_{x_m} : X' \rightarrow 0{,}1$ is the characteristic function of the square support of a block located at $x_m \in X'$ and $\hat{Y'}_{x_m}^{wie,x_R}$ are the block-wise estimates.

$$\hat{y}^{final}(x) = \frac{\sum\limits_{x_R \in X'} \sum\limits_{x_m \in S_{x_R}^{wie}} w_{x_R}^{wie} \hat{Y'}_{x_m}^{wie,x_R}(x)}{\sum\limits_{x_R \in X'} \sum\limits_{x_m \in S_{x_R}^{wie}} w_{x_R}^{wie} \chi_{x_m}(x)}, \forall x \in X', \tag{1}$$

A α parameter is specified depending on the level of noise to define the block size used. Noise filtering with $\alpha = 20$ (specifying a relatively small block size) was found to be most suitable for removing speckle noise while maintaining the integrity of vessel shadows in cross vendor experiments (Fig.1(b)).

2.3 Pathology Removal & Tiling

In addition to noise, the presence of non-vessel shadows caused by pathology adversely affects the detection system and their removal is required to prevent their usage as landmarks. Such shadows appear in $\mathcal{P}$ with similar intensity to

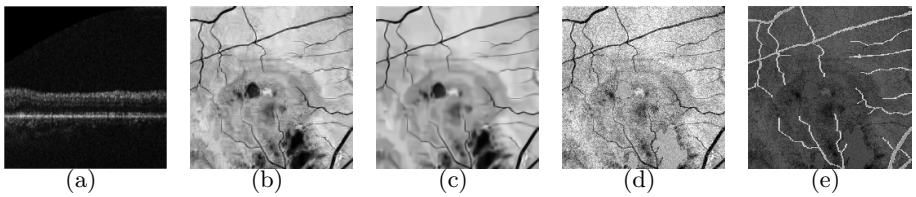

Fig. 1. (a) Exemplar flattened $\mathbf{B}_s$. Denoised $\mathcal{P}$ using (b) $\alpha = 20$ & (c) $\alpha = 40$. (d) $\mathcal{P}$ from (b) & (c) with pathology shadow removed. (e) Segmented vessel shadows overlayed onto (d).

vessel shadows but generally possess different size and shape characteristics. Vessels are generally long, thin connected regions, thus the opposite characteristics can be used to describe pathology shadows. However, due to imaging artefacts at acquisition and patient motion, the shadows can become distorted and may feature similar characteristics. Thus to distinguish between these two sets of shadows, shape was used as a feature, given that pathology shadows are amorphous and not thin and narrow. In addition, they seldom reach the boundaries of $\mathcal{P}$ and are are not simultaneously long, thin and with a low area. For each detected pathology shadow region, randomly selected boundary intensities are used to fill the pathology region, masking it from $\mathcal{P}$ (Fig.1(d)). This is repeated for all pathology regions at which point the masked $\mathcal{P}$ is denoised again giving $\mathcal{P}_M$.

Due to the presence of pathology, motion artefacts and noise, intra-projection intensity and appearance variation is present. Issues related to this problem can be limited by reducing the window size for vessel detection using tiling denoted as $\mathcal{P}_{MT}$. The search space is reduced to equally sized tiles of the form 4×4, 3×3, 2×2, 2×1 and 1×2.

2.4 Vessel Shadow Detection

The vessel enhancement filter proposed in [8] is applied on the tiled projection sub-images to detect tubular geometric structures and suppress remaining noise and background. Applied to $\mathcal{P}_{MT}$, the Taylor expansion approximates the image structure to the second order resulting in the gradient vector ($\nabla_{o,s}$) and Hessian matrix ($\mathcal{H}_{o,s}$) computed in x_o at scale s. Using eigenvalue analysis of the Hessian, the principal directions in which the local second order image structure is decomposed are extracted. This results in three orthonormal directions giving a spherical neighbourhood centred at x_o, $\mathcal{N}_{x_o}$, mapped by $\mathcal{H}_o$ to an ellipsoid with directions given by the Hessian eigenvectors and axis semi-lengths as the magnitudes of eigenvalues. This ellipsoid describes the second order structure where eigenvalues are ordered $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$. Two geometric ratios based on the second order ellipsoid are used as a dissimilarity measure. The first defines the deviation from blob-like structures whereas the second distinguishes between plate-like and line-like structures. Thus a measure defining second order structuredness of an image is used to distinguish between background and vessel. The authors explain this measure will be low for the background and larger in regions

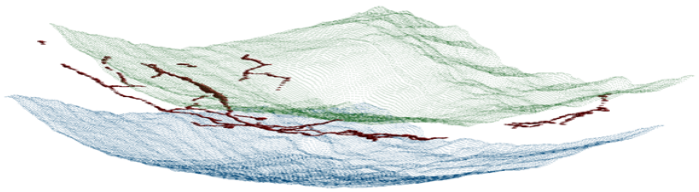

Fig. 2. Exemplar retinal vessel segmentation showing modelled in red the vessels between the green ILM surface points and blue RPE surface points.

of vessels, however the opposite would be true for OCT projection images. This is due to the vessel shadows appearing dark against a lighter background whereas the method in [8] was originally proposed for MRA and CTA images which are commonly interpreted using maximum intensity projection and feature brighter vessels against darker background. However the dark-to-light, light-to-dark distinction is still present and prominent in $\mathcal{P}_{MT}$. In addition, other similarities include the vessel structures occupying a small volume of the image as a whole.

Region growing using the highest intensity pixel from the result of the vessel enhancement filter segments the vessel shadow(s) within each tile of $\mathcal{P}_{MT}$. This process is repeated for each combination of tile patterns, with the intersection of all candidate segmentations taken as the final result. This can be seen in Fig.1(e) where the segmented vessel shadows are overlayed onto the pathology shadow removed projection image.

2.5 Segmented Vessel Point Post Processing

The vessel shadows are acquired from $\mathcal{P}$ which is a 2D representation of vessel positions. However, the vessel structure is 3D and the third dimensional spatial information is extremely important for volume registration. Thus the segmented vessel shadows are skeletonized by removal of the boundary pixels until only a single pixel remains along the structure while preventing the structure from breaking apart. The X and Y coordinate values for each vessel shadow skeleton point are compared with the corresponding coordinates from the vendor segmented ILM to identify the relative depth (Z) position within the retinal OCT volume for each skeleton point. This can be visualized as shown in Fig.2.

3 Point Based Vessel Tree Registration

In this section, we show the use of CPD registration to align retinal OCT volumes using the landmarks identified in Sec.2.

3.1 Coherent Point Drift vs. Iterative Closest Point

Registration of macula centred OCT volumes aims to improve the ability for clinicians to assess and analyse disease progression and treatment success by aligning source and target scans into the same coordinate system. This must

be accurate and reproducible to ensure that pathology and treatment progress analysis is not affected by artefacts. Registration of retinal OCT volumes has primarily been performed rigidly using ICP registration to obtain the primary transformation parameters for enface rotation, translation and scaling as well as depth translation [2]. However, a major limitation of ICP is the assumption that every source point corresponds with the closest point to it in the target point set. In addition, ICP assumes the source point set is sufficiently close to the target point set which cannot be assumed in all cases of retinal OCT. CPD [9] is suitable in this case as it is capable of performing rigid, non-rigid and affine transformations in the presence of noise, outliers and most importantly for temporally acquired retinal OCT volumes, missing points.

3.2 Point Set Registration with CPD

The alignment of two point sets, defined here as the source and target 3D retinal vessel centrelines obtained in Sec.2, is considered in CPD as a probability density estimation problem by [9]. One point set represents the Gaussian mixture model (GMM) centroids and the other represents the data points. The two point sets become aligned at the optimum and the correspondence is defined by the maximum of the GMM posterior probability for a given data point.

Given the GMM centroid points and the data points generated by the GMM, the GMM probability density function (PDF) is defined. In addition, the authors in [9] also add a weighted uniform distribution to the mixture model allowing it to account for noise and outliers. For all GMM components, equal isotropic covariances σ^2 is used as well as equal membership properties. The GMM centroid locations are re-parametrized using the parameter set θ, estimated using maximum likelihood allowing the i.i.d data assumption to be made. Thus the correspondence probability between two points can be defined as the posterior probability of the GMM centroid. Expectation Maximization (EM) is used to find θ and σ^2 where as the expectation (E) step, the parameters θ and σ^2 are initially guessed and then the Bayes' theorem is used to compute a posteriori probability distributions of mixture components. The maximisation (M) step is then used to find the new parameter values by minimizing the expectation of the complete negative log-likelihood function with respect to the new parameters. The E and M steps are processed alternately until convergence, generating the posterior probabilities of the GMM components. Unlike ICP, in CPD, the GMM centroids are forced to move coherently as a group to preserve the topological structure of the point sets.

τ_N and τ_A are examined in this paper. For τ_A, optimization is unconstrained and defined as $\mathcal{T}(y_m; \mathbf{R}, \mathbf{t}, s) = \mathbf{B}\mathbf{y}_m + \mathbf{t}$, where $\mathbf{B}_{D\times D}$ is an affine transformation matrix and $\mathbf{t}_{D\times 1}$ is the translation vector. For τ_N, transformation is defined as the initial position plus a displacement function, $\mathcal{T}(\mathbf{Y}, v) = \mathbf{Y} + v(\mathbf{Y})$, where v is the displacement function. In the method for τ_N, the authors in [9] regularize the norm of the displacement function. This approach is based on the motion coherence theory (MCT)[10] stating that points close to one another have a tendency to move coherently, thus the displacement function between points sets should

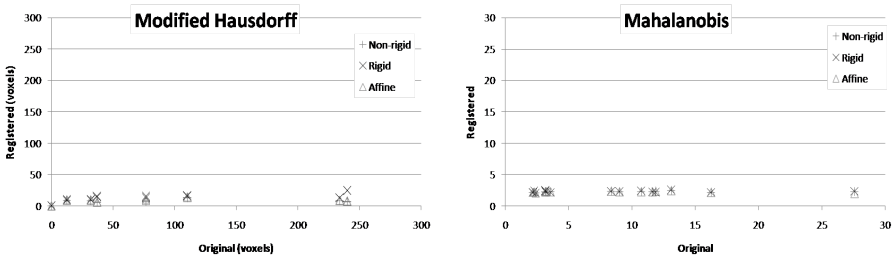

Fig. 3. (a) MHD & (b) MD between pairs of registered OCT scans comparing distances prior to registration (horizontal axis) against distances post registration (vertical axis).

be smooth. The regularization term in [9] is equivalent to the MCT, thus implies that point based motion coherence is imposed. This regularization is integral to the CPD algorithm and its ability to register point sets coherently, important for retinal vessel registration. Thus we use a high regularization parameter for τ_N to maintain vessel structure coherence.

A transformation matrix is generated from the registered vessel points to perform rigid registration of the retinal OCT volume. We currently only perform rigid volume transformation as the clinical suitability of non-rigid volume registration of retinal volumes is uncertain.

4 Registration Results

To evaluate registration performance, two sets of results are presented. In our experiments, intra-vendor registration was carried out between pairs of OCT scans acquired from 8 patients at different time points. 7 pairs feature unique scans and the eighth pair was used as a control and featured one unique scan twice. Thus a total of 15 unique 3D SD-OCT fovea-centred diseased volumes were used for testing. Of the 15 scans, levels of pathology ranged from low to high primarily due to the presence of cysts, exudates and drusen. In addition, motion artefacts in the X direction are seen in $\mathcal{P}$ due to patient motion, resulting in misaligned $\mathbf{B}_s$.

Firstly, we used CPD to register between scan pairs in 3D using τ_N and τ_A. Performance for these methods was quantified by calculating the modified Hausdorff [11] (MHD, Fig.4) and Mahalanobis distances (MD, Fig.4) between point sets, where point set distance is expected to decrease post registration as the retinal vessels become aligned. A mean MHD of 92.89 $\pm$ 92.41 voxels was calculated for the original vessel pairs which was reduced to 8.319 $\pm$ 3.91 voxels and 7.579 $\pm$ 3.569 for τ_A and τ_N respectively. A mean MD of 8.192 $\pm$ 6.999 was calculated for the original vessel pairs which was reduced to 2.259 $\pm$ 0.078 and 2.249 $\pm$ 0.109 for τ_N and τ_A respectively. Thus the mean MHD across all scan pairs was reduced by 92% & 91% for τ_N and τ_A respectively and mean MD was reduced by 72% & 73% for τ_N and τ_A respectively. Comparison with the method proposed in [2] when applied to the same test scans resulted in mean MHD of 94.39 $\pm$ 93.33 voxels and mean MD of 7.804 $\pm$ 5.373 between registered and target scans, showing our method to perform better.

Our second experiment qualitatively assesses system performance to account for the disjoint segments of the landmark retinal vessels between scan pairs. The registration results obtained previously were given to 3 expert graders who were asked to identify, in 2D, the major bifurcations from both the target and registered vessel points that were well and poorly aligned. Across the graders, a mean of 136 ± 7 and 4 ± 1 bifurcations using τ_A were well and poorly aligned respectively. With τ_N, a mean of 139 ± 10 and 2 ± 1 bifurcations were well and poorly aligned respectively. Thus 94% and 96% of bifurcations correctly overlapped between target and registered vessel sets using τ_A and τ_N respectively.

5 Conclusion

We have proposed a method of retinal OCT volume registration using CPD registration of segmented retinal vessel shadows as landmarks. We first showed that the retinal vessels could be used to accurately and reproducibly represent the retinal OCT volume for both intra-patient acquisition and multi-vendor acquisition. By combining these segmented vessel shadows and modelling them as the retinal vessels, based on their geometric parameters obtained from the respective volume and layer information, CPD registration was applied to transform the temporally acquired vessel point sets. Our experiments quantifying the MHD and MD between target and registered point sets show the minimisation of the distances between original and target point sets after transformation, by up to 92% using non-rigid transformation and 91% using affine transformation. Qualitative analysis of the transformed point set bifurcations with their respective targets by experienced graders shows a mean of 96% correctly overlapping bifurcations when using non-rigid transformation. In addition, comparison against current state of the art methods for retinal OCT volume registration show significantly improved performance for pathological, multiple vendor scans.

References

1. Geitzenauer, W., Hitzenberger, C.K., Schmidt-Erfurth, U.M.: Retinal optical coherence tomography: past, present and future perspectives. Br J Ophthalmol **95** (July 2010) 171–7
2. Niemeijer, M., Lee, K., Garvin, M.K., Abràmoff, M.D., Sonka, M.: Registration of 3d spectral oct volumes combining icp with a graph based approach. In: SPIE Medical Imaging. Volume 8314. (February 2012)
3. Kolar, R., Tasevsky, P.: Registration of 3d retinal optical coherence tomography data and 2d fundus images. BIR **6204** (2010) 72–82
4. Golabbakhsh, M., Rabbani, H.: Vessel-based registration of fundus and optical coherence tomography projection images of retina using a quadratic registration model. IET Image Processing **7** (2013) 768–76
5. Niemeijer, M., Garvin, M.K., van Ginneken, B., Abràmoff, M.D.: Vessel segmentation in 3d spectral oct scans of the retina. In: SPIE Medical Imaging. Volume 6914. (March 2008)

6. Schmitt, J.M., Xiang, S.H., Yung, K.M.: Speckle in optical coherence tomography. J. Biomed. Opt. **4** (1999) 95–105
7. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3d transform-domain collaborative filtering. IEEE Trans on Image Proc **16**(8) (August 2007)
8. Frangi, A., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. MICCAI **1496** (1998) 130–37
9. Myronenko, A., Song, X.: Point set registration: Coherent point drift. PAMI **32** (2010) 2262–2275
10. Yuille, A.L., Grzymacz, N.M.: The motion coherence theory. ICCV **3** (1988) 344–353
11. Dubuisson, M., Jain, A.K.: A modified hausdorff distance for object matching. ICPR **1** (1994) 566–568