OCTRexpert：A Feature-based 3D Registration Method for Retinal OCT Images

Lingjiao Pan, Fei Shi, Dehui Xiang , Kai Yu, Luwen Duan, Jian Zheng and Xinjian Chen

***Abstract*—Medical image registration can be used for studying longitudinal and cross-sectional data, quantitatively monitoring disease progression and guiding computer assisted diagnosis and treatments. However, deformable registration which enables more precise and quantitative comparison has not been well developed for retinal optical coherence tomography (OCT) images. This paper proposes a new 3D registration approach for retinal OCT data called OCTRexpert. To the best of our knowledge, the proposed algorithm is the first full 3D registration approach for retinal OCT images which can be applied to longitudinal OCT images for both normal and serious pathological subjects. In this approach, a pre-processing method is first performed to remove eye motion artifact and then a novel design-detection-deformation strategy is applied for the registration. In the design step, a couple of features are designed for each voxel in the image. In the detection step, active voxels are selected and the point-to-point correspondences between the subject and template images are established. In the deformation step, the image is hierarchically deformed according to the detected correspondences in multi-resolution. The proposed method is evaluated on a dataset with longitudinal OCT images from 20 healthy subjects and 4 subjects diagnosed with serious Choroidal Neovascularization (CNV). Experimental results show that the proposed registration algorithm consistently yields statistically significant improvements in both Dice similarity coefficient and the average unsigned surface error compared with the other registration methods.**

***Index Terms*—Image registration, Optical coherence tomography (OCT), retinal image.**

This work was supported by the National key R & D plan "key scientific issues of transformative technology" under 2018YFA0701700， National Nature Science Foundation of China for Excellent Young Scholars under 61622114，in part by the National Natural Science Foundation of China (NSFC) under Grant 61601317 and 61601207, in part by the Natural Science Foundation for colleges and universities of the Jiangsu Province under Grant 17KJB510015 , in part by the Changzhou Social Development Foundation under Grant CE20165021, in part by Collaborative Innovation Center of IoT Industrialization and Intelligent Production, Minjiang University under Grant IIC1702. (Corresponding *author: Xinjian Chen.)*

L. Pan is with the School of Electronic and Information Engineering, Soochow University, China. She is also with the School of Electrical and Information Engineering, Jiangsu University of Technology (e-mail: 20154028001@stu.suda.edu.cn). F. Shi and Kai Yu are with the School of Electronic and Information Engineering, Soochow University, China. They are also with the Collaborative Innovation Center of IoT Industrialization and Intelligent Production, Minjiang University. L.Duan and J. Zheng is with Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, China. X. Chen and D. Xiang are with the School of Electronic and Information Engineering, Soochow University, China, and also with the State Key Laboratory of Radiation Medicine and Protection, School of Radiation Medicine and Protection, Soochow University, Suzhou China. (e-mail: xjchen@suda.edu.cn,).

I. Introduction

Medical image registration has attracted more and more attention due to its valuable applications in clinical studies [1]. In particular, registration can be used for studying longitudinal and cross-sectional data, quantitatively monitoring disease progression and guiding computer assisted diagnosis and treatments. Registration is a fundamental task in medical image processing used to match multiple images taken from different viewpoints, different sensors or different time points. In the past decades, many registration techniques have been developed for various types of data and applications. However, spectral-domain optical coherence tomography (SD-OCT) image registration which enables more precise and quantitative comparison of retinal disease has not been well developed.

3D SD-OCT imaging technique is a noninvasive and non-contact scan of retina, which has been widely used in investigation of retinal pathology, such as Choroidal Neovascularization (CNV), age-related macular degeneration, glaucoma and so on [2]-[6]. SD-OCT acquires information through one dimension profile scans called A-scans. The A-scan is constructed by evaluating frequency spectrum of the interference signal between the reflected light from the tissue and a reference mirror. SD-OCT device acquires a single A-scan at a time. By acquiring a series of A-scans in a raster scanning pattern, the cross sectional slicer B-scan is generated. Composing successive B-scans yields a 3D SD-OCT image of retina. Fig.1 shows a 3D SD-OCT scan of retina which consists of B-scan slices. Each B-scan slice consists of A-scans. Since SD-OCT imaging is relatively new compared with other medical imaging modalities such as magnetic resonance imaging and computed tomography, the requirement for processing SD-OCT images has a shorter history. With the fast development of SD-OCT technique, the demand for advanced image analysis techniques is rapidly growing. Nevertheless, the development of such techniques can be challenging as SD-OCT image is inherently noisy and the structure of retina can change drastically when pathology occurs. Therefore, although many image registration techniques have been well developed, they cannot be directly applied to SD-OCT images. The main reasons are as follows:

1) In SD-OCT, voxels on an A-scan are highly correlated. For example, the presence of blood vessel in an A-scan will totally change the gray value of voxels below it. Furthermore, parallel A-scans in the OCT image actually should be fanning outwards from the center in physical space which depends on

the optics of individual's eye [7]. The geometry of retinal OCT image makes the deformation across different A-scans challenging.

2) Eye movements during the scan process lead to significant spatial distortions (large gaps) between B-scans. These distortions make the application of existing registration algorithms difficult especially for interpolation and regularization.

3) Serious speckle noise of OCT results in poor signal to noise ratio (SNR) compared with other medical imaging modalities. The normal intensity based registration methods tend to be more sensitive to the noises during the registration process.

4) The layer structures of retina change dramatically in serious pathological regions. For example, fovea is often deformed by neovascularization. This makes it difficult to extract stable features for registration.

5) High resolution of 3D OCT data causes dramatic increase of the computational complexity for registration and makes the clinical application difficult.

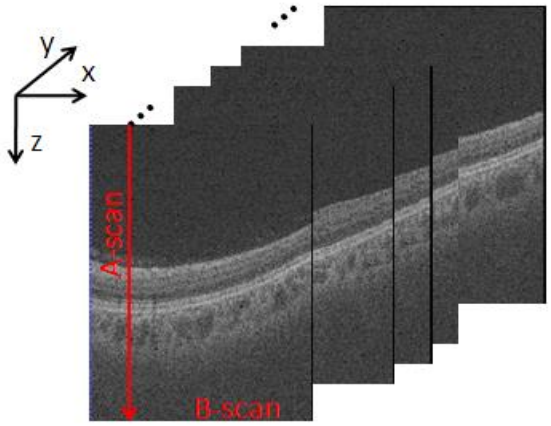


Fig. 1. Example of a 3D SD-OCT scan of retina.

This paper proposes a new 3D registration algorithm which can be applied to the longitudinal retinal OCT images for both normal data and pathological data. Compared with other existing algorithms, the proposed method presents several contributions.

1) To the best of our knowledge, this is the first longitudinal retinal OCT image deformable registration algorithm which considered both normal data and serious pathological data.
2) It is a feature-based registration method. In this method, in order to depress the serious speckle noise in OCT images, intensity-based region feature, surface-based structure feature and vessel-like feature are designed to distinguish different structures of retina.
3) Since the computation complexity of non-rigid registration is usually very high, a hierarchical deformation mechanism is designed to reduce local minima as well as to speed up the computation speed.

II. Related Work

To develop a special registration method for retinal OCT, many works have been done. Most of the works are restricted to rigid registration [8]-[12]. However, it is not enough to describe the deformation of retina using such a low dimensional model. Therefore, a deformable registration method is quite necessary. To the best of our knowledge, limited deformable registration approaches have been reported for OCT images recently. The highly-rated deformable registration algorithms such as SyN and DRAMMS are showed to be unreliable when applied to the whole OCT image and they took long time (around 8-10 hours) to finish registration process [13]. Zheng et al. used the SyN registration algorithm to aid the OCT retinal layer segmentation [14]. Zhang et al. proposed a two-step image registration method to reduce speckle noise in OCT images [15]. However, these registration methods were applied to individual layers or successive B-scans instead of the whole OCT image. They are not a full 3D registration. Chen et al. proposed a 3D deformable registration method using A-scan similarity [13]. They considered the physical geometry of retina and assumed that the deformable transformation only occurred along A-scan direction. In their method, the position of the foveae which was approximated as the superior point of the thinnest portion of the retina was used for the initial registration. However, this method does not work for the data with severe lesions such as CNV where the fovea is no longer the thinnest portion of the retina. In our previous work, a two-stage registration method for OCT volumes with CNV was reported [16]. We used the projection of 2D vessel points as landmark points and applied the coherent point drift method followed by the B-spline-based registration method to find the optimal deformation. This work was further extended by Du and Gong et al. by using conditional correlation ratio instead of mutual information as the similarity measurement for the B-spline-based registration to improve the registration accuracy [17], [18]. However, these two methods still have some obstacles. First, their deformable transformation depends on the intensity similarity and is very sensitive to the speckle noise especially for the CNV case when the intensity contrast is low due to the occurrence of neovascularization. Second, the computation complexity is usually very high for such intensity-based methods.

In this paper, we propose a feature-based 3D registration method which can be used for both normal retinal OCT data and serious pathological OCT data. Although intensity-based registration methods are more general, feature-based registration methods have several advantages compared to them: 1) Feature-based methods have low computational complexity because they evaluate a matching criterion on a relatively small number of feature points instead of on every single voxel in an image. 2) They can achieve higher registration accuracy because well selected features have less ambiguity than intensity similarity in defining correspondences. 3) They show better robustness since the selected features can avoid the negative impact of image noise effectively. One of the typical feature-based medical image registration methods is HAMMER, which is successfully applied to the registration of brain MRI images and various other organ images [19], [20]. Although HAMMER achieved big success in brain image registration, it tends to be unreliable when directly applied to

3D retinal SD-OCT images. This is due to two main reasons. First, the structure of the retina is different from that of the brain. Geometric moment invariants (GMIs) can distinguish different geometric structures of brain, but it fails to distinguish the plate-like structure of retinal layers. Second, retinal SD-OCT images have much higher resolution than brain MRI images. In HAMMER, the driving voxels are selected by using a fuzzy clustering method. Since the fuzzy clustering method needs to calculate the distance between voxels, it is time consuming especially when the data is large and may have local minima. Our work is inspired by HAMMER and tries to extend HAMMER to OCT registration. Compared with other retinal OCT image registration methods, we adopt HAMMER's hierarchical attribute matching mechanism to improve the registration accuracy while reducing the computation complexity. Furthermore, our work presents several novel elements compared with HAMMER: 1) We propose to use intensity-based region feature, surface-based structure feature and vessel-like feature instead of GMI feature to distinguish different structures of retina. 2) We propose an efficient driving voxel selection method to further reduce the computation complexity. In our method, rather than randomly selecting the active voxels, the active voxels are hierarchically selected following a strategy for importance coefficient assignment. 3) A preprocessing step is designed to remove the motion distortion in retinal OCT data before registration. To the best of our knowledge, the proposed method is the first longitudinal retinal OCT image registration algorithm which considered both normal data and serious pathological data.

III. Methods

A. Overview of the Approach

The deformable transformation is a free form mapping at each voxel x. It can be solved by finding a transformation of each voxel such that the energy function is minimized. Considering the high resolution of OCT data, the energy function would be a very high dimension function which makes it extremely difficult to find the global optimal solution. The main difficulties are the computation complexity and the local minima problem. To speed up registration process and reduce local minima, a novel design-detection-deformation mechanism is designed. The proposed method consists of four steps: preprocessing, feature design, correspondence detection and hierarchical deformation. The overall flowchart is shown in Fig. 2. In preprocessing step, OCT data are first segmented by detecting 7 surfaces using graph search-based method and then B-scans are flattened to correct eye movement. In the design step, a couple of features are designed for each voxel in the image. In the detection and deformation step, active voxels are hierarchically selected and point-to-point correspondences between the subject and the template images are established. The image is then hierarchically deformed according to the detected correspondences in multi-resolution. The detail of each step is discussed in the following parts.

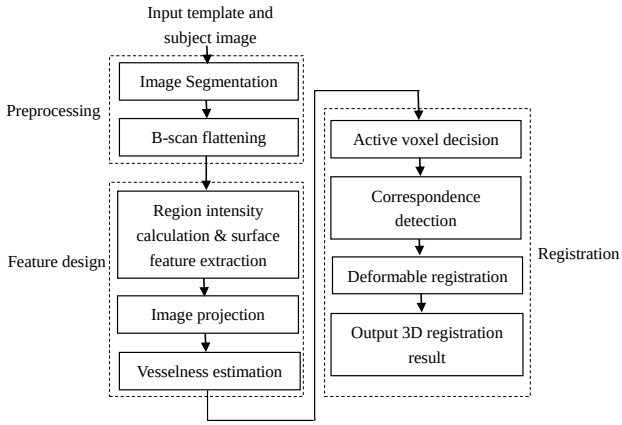


Fig. 2. Flowchart of the proposed algorithm.

B. Preprocessing

1) Multi-resolution graph search: 3D graph-based optimal surface segmentation method can detect multiple interacting surfaces simultaneously [21]. The basic idea is to transform the optimal surface detection problem into computing a minimum cut in an arc-weighted directed graph. This method and its variations are successfully applied to retinal layer segmentation of macular optical coherence tomography images [22]-[31]. The surface segmentation methods designed for normal retinas can also be used to segment the retinas with glaucoma and multiple sclerosis or other diseases without dramatic change in the layer structure in the early stage. However, it is difficult to segment the retinas with additional layer structures such as sub-RPE fluid and intra-retinal cysts by using the same methods. In our algorithm, by considering both normal and serious pathological data such as CNV, the surface segmentation is conducted based on our previously proposed constrained graph search method [32]. In preprocessing, all B-scans of the template image and the subject image are segmented by seven retinal surfaces as shown in Fig.3. These seven surfaces partition an OCT dataset into six layers. They are 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 and inner/outer segment layer (ONL+IS/OSL) and retinal pigment epithelium (RPE).

2) B-scans flattening: Eye movement artifact occurring during 3D OCT scanning is a problem for retinal OCT imaging and makes image registration difficult. During OCT acquisition process, since the volume is acquired in a few seconds, eye movement caused by heart beat and respiration in the scan process results in motion artifacts. In motion distorted data, the positions of layers varies greatly in consecutive B-scans which make interpolation and regularization difficult. The position shifts of B-scans can be viewed in the y-z slices, as shown in Fig. 4 (a), where each column corresponds to a B-scan. Flattening the 3D OCT volumes is often used to correct eye

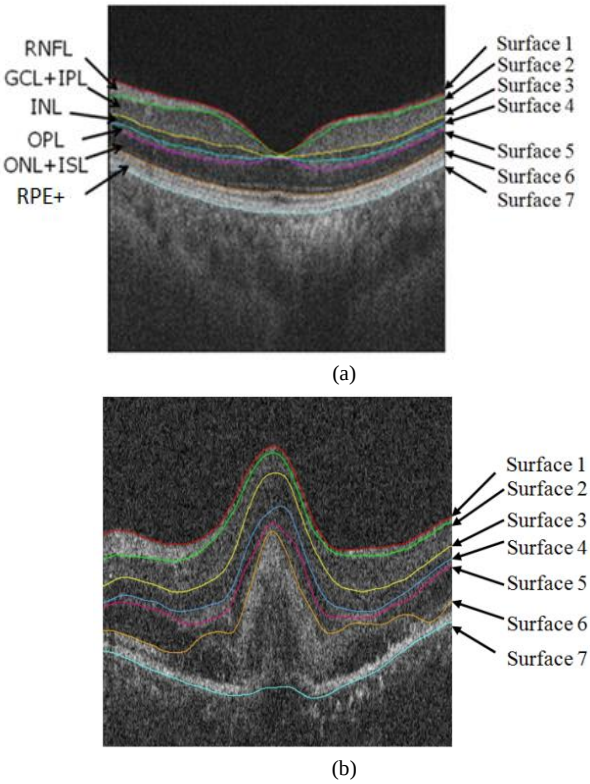


Fig. 3. Retina surface segmentation result of normal eye (a) and diseased eye with severe lesion (b).

movement artifacts and provide a more consistent retinal shape for visualization. In this paper, the surface segmentation results are used for flattening. The lowest position of surface 7 is set as the base position l_{base}. For each A-scan $A_{i,j}$, the displacement is estimated by calculating the vertical distance $\Delta z_{i,j}$

$$\Delta z_{i,j} = l_{\text{base}} - l_{7i,j} \tag{1}$$

where $l_{7i,j}$ is the position of surface 7 in each $A_{i,j}$. And then, each A-scan is shifted down in the z-direction according to $\Delta z_{i,j}$, such that the z positions of surface 7 become the same for all A-scans. The flattening process results in a smoothed appearance of all the layers in the y-z slice, as shown in Fig.4 (b). The 3D rendering of surface 7 shows that the motion artifacts are greatly reduced after flattening. It should be noted that the retina has its own curvature. Since it is difficult to distinguish axial artifacts from retinal curvature, B-scans flattening method eliminates axial eye movement and the curvature of the retina as well. Although this is not always a bad thing, the curvature of retina is useful in some applications. In our paper, after registration, the flattened surfaces can be converted back to their positions to recover the original shape. However, we can only recover the curvature in B-scan. To obtain the true curvature of retina in 3D space, additional orthogonal volume scans are required.

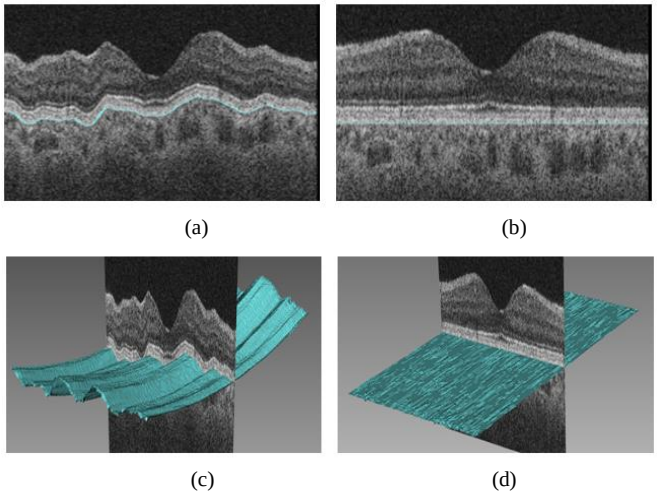


Fig. 4. The y-z slice before (a) and after alignment (b) with surface7 overlaid; 3D rendering of surface 7 before (c) and after alignment (d).

C. Feature Design

As mentioned before, feature extraction is very important for registration. Feature vector of each voxel in a 3D image can reflect the underlying local structure. A good feature set used for image registration does not have to be very detailed, but it must be robust to structural variations and invariant to image rotations. If the feature set is rich enough and robust enough, it can distinguish different parts of the anatomy. However, we should note that more features will not always produce better results if some features are ambiguous. Since SD-OCT image is inherently noisy and the structure of retina can drastically change for diseased retina, few stable and distinct features can be extracted for retina. In the proposed method, the following sets of features are designed.

1) Surface-based structure feature: Between two regions, there is a surface. The voxels on the surface are more distinct than other voxels. According to the surface segmentation results in the preprocessing, value 0 is assigned to the voxels which are not belonging to any surface and seven real values between 0 and 1.0 are assigned to seven surfaces respectively to represent the structure information. From the top surface to the bottom surface, the assigned values are increased by Δs in turn. Suppose N is the total number of the segmented surfaces (here N=7), the surface-based structure feature can be denoted as $n\Delta s$, where n=0, 1 … N correspond to non-surface voxels and voxels on N surfaces, respectively. For example, let Δs equals to 0.1, then value 0.1 is assigned to the voxels on surface 1, value 0.2 is assigned to the voxels on surface 2, and so on. From the top surface to the bottom surface, the structure feature value is increased by 0.1in turn, and finally a value of 0.7 is assigned to voxels on surface 7.

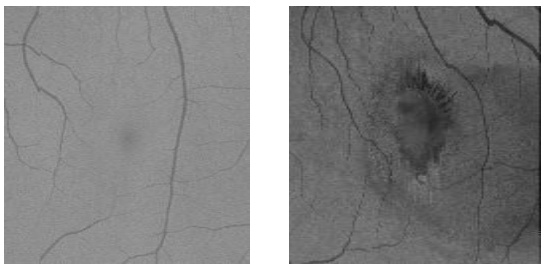
2) Intensity- based region feature: Region features based on mean intensity value of each layer are designed. Since different intensity ranges represent different layers in retina OCT images [33], intensity-based feature can describe different layer regions. For a retinal OCT scan, the segmented surfaces partition the OCT volume into eight regions including six segmented layer regions, choroid and sclera region below

surface 7 and vitreous region above surface 1. The region-specific mean intensities are calculated for each region. Since close to the fovea, RNFL region is too thin to distinguish from inner limiting membrane (ILM). And RPE region can have fluid and neovascularization in CNV case. Therefore, only the leftmost and rightmost 20% pixels are used in the calculation. Then for each region, voxel intensities are modified according to the following intensity transformation function

$$I\ (r) = I_R \times \frac{I_M(r) - I_{Mvit}}{I_{MRPE} - I_{Mvit}} \tag{2}$$

where $I_M(r)$ is the mean intensity value of region r. r =0,1,...7 denotes eight regions. I_{Mvit} and I_{MRPE} denote the mean intensity of vitreous and RPE, respectively. The vitreous and RPE are specifically chosen because they are consistently the darkest and brightest layers in OCT scan. I_R is the normalized intensity range which is set to 1 and $I(r)$ denotes the modified intensity of region r. As a voxel located on the surface between two regions, it is difficult to say whether it belongs to the region above or below the surface. But considering the consistency and stability, the surface is regarded as the lower surface of the upper region, and its region feature refers to the upper region of the surface.

3) Vessel-like feature: Since the retinal layers in OCT image are plate-like structure, structure feature and region feature can differentiate the voxels in vertical direction. However, they cannot differentiate the voxels in horizontal direction. Therefore, vessel-like feature based on vessel shape responses is designed for differentiating the voxels in horizontal direction. According to the fact that the blood vessels produce a hyper intense area in the inner retina and their shadows produce a hypo intense area in the outer retina, vessel information can be extracted from the projection image of 3D OCT data. To obtain 2D projection image, the intensity values of voxels along each A-scan between the upper surface of RPE layer and Bruch's membrane are averaged. Segmentation of vessels in OCT images is much more challenging than in fundus images due to the speckle noise and lower resolution. Especially, in CNV case, sub-retinal fluid will cause non-vessel shadows which have the similar intensity with the blood vessels as shown in Fig.5.



(a) (b)

Fig. 5. Illustration of the projection image. (a) Projection image of the normal eye. (b) Projection image of the diseased eye.

However, they have different shapes. The blood vessels are generally long and thin. Therefore, the shape characteristics can be used to distinguish two kinds of shadows. In our method, a multi-scale vessel enhancement filter based on Hessian matrix [34] is applied to the projection image to detect tubular structure of vessels. A vesselness measure function is defined as

$$v_o(\sigma) = \begin{cases} 0 & \lambda_2 > 0 \\ \exp\left(-\frac{R_B(\sigma)^2}{2\varepsilon^2}\right)\left(1-\exp\left(-\frac{S(\sigma)^2}{2c^2}\right)\right) & otherwise \end{cases} \tag{3}$$

$$R_B(\sigma) = \frac{|\lambda_1(\sigma)|}{|\lambda_2(\sigma)|} \tag{4}$$

$$S(\sigma) = \|H\|_F = \sqrt{\lambda_1(\sigma)^2 + \lambda_2(\sigma)^2} \tag{5}$$

where $\lambda_1(\sigma)$ and $\lambda_2(\sigma)$ denote the eigenvalues of the Hessian, which are computed at scale σ. For an ideal tubular structure in 2D image, the eigenvalues should satisfy $|\lambda_1| \approx 0$ and $|\lambda_1| << |\lambda_2|$. R_B is the blobness measure and accounts for the eccentricity of the second order ellipse. S is the second order structureness measure. The ε and c are the thresholds which control the sensitivity of the line filter to the measures R_B and S. The vesselness measure is analyzed between $\sigma_{\min}$ and $\sigma_{\max}$ at different scales. The response of the line filter will be maximized at a scale that approximately matches the shape of the vessel. The final estimation of vesselness is obtained as

$$v = \max_{\sigma_{\min} < \sigma < \sigma_{\max}} v_o(\sigma) \tag{6}$$

Since each pixel in the projection image corresponds to an A-scan, the estimation of vesselness is used as the vessel-like feature for the voxels in each A-scan.

D. Correspondence Detection

The problem of finding out the deformable transformation between the template image and the subject image is usually formulated as optimization of an energy function which evaluates the similarity of two images. The solution of the energy function directly determines the results of the registration. The following is the energy function used in the deformable registration.

$$E = \sum_{x \in V_S} w_T(x) \left(\frac{\sum_{z \in n(x)} w(z)(1 - m(f_S(z), f_T(d(z))))}{\sum_{z \in n(x)} w(z)} \right) + \beta \sum_{x \in V_S} \left\| \nabla^2 (d(x) - x) \right\| \tag{7}$$

There are two terms in this energy function. The first term measures the summation of differences between template feature vector and subject feature vector deformed by the current estimation of deformation. For each voxel x in subject volume V_S, $n(x)$ denotes a neighborhood of x for computing

the similarity. $d(x)$ defines the deformable transformation of voxel x. z is a neighboring voxel in the neighborhood. $f_S(z)$ presents the feature vector of subject voxel z, while $f_T(d(z))$ denotes the feature vector of the template voxel at position $d(z)$. The function $m(f_S(z), f_T(d(z)))$ measures the similarity between $f_S(z)$ and $f_T(d(z))$. $w(z)$ and w_T are weighting parameters which assign large value to the voxels with distinct features such as voxels on the surfaces or active voxels. The second term is smoothness constraint to make the result deformation field smooth. ∇^2 is Laplacian operator and the parameter β controls the smoothness of the deformation field.

In this part, we define the criterion to choose the best correspondence. The similarity of two voxels x and y is defined as

$$m(f(x), f(y)) = \begin{cases} 0, & f_1(x) \neq f_1(y) \\ \prod_{i=2}^{3} (1 - |f_i(x) - f_i(y)|), & otherwise \end{cases} \quad (8)$$

where $f_i(x)$ denotes the *i-th* feature value of voxel x and i=1, 2, 3 corresponds to the surface-based structure feature, intensity-based region feature and vessel-like feature, respectively. To find the best correspondence of voxel x, the similarity is calculated in a neighborhood $n(x)$ centered at x rather than a single voxel to improve the robustness. For a neighboring subject voxel z, its feature vector $f_S(z)$ is compared with the feature vector $f_T(d(z))$ of the corresponding deformed voxel $d(z)$ in the template. The similarity is defined as $m(f_S(z), f_T(d(z)))$. Thereby the difference is $1\text{-}m(f_S(z), f_T(d(z)))$. Given the weighting function $w(z)$ which is defined as

$$w(z) = \begin{cases} 0, & if \ \ f_1(z) = 0 \\ 1, & otherwise \end{cases} \quad (9)$$

the energy function in Eq.(7) can be simplified and only the surface voxels in the neighborhood are used to dominate the correspondence detection. By using this similarity criterion, the subject's boundaries can match well with the corresponding template boundaries. In the implementation, the size of the neighborhood and the search range are both defined in a hierarchical way. In the initial registration stage, they are large to deform the OCT image globally. In the later stage, they are hierarchically reduced to improve the accuracy of the registration.

E. Hierarchical deformation

The deformable transformation $d(x)$ in Eq. (7) is a free form mapping at each voxel x. It can be solved by finding a transformation of each voxel such that the energy function is minimized. Considering the high resolution of OCT data, the energy function would be a very high dimension function which makes it extremely difficult to find the global optimal solution. The main difficulties are the computation complexity and the local minima problem. In our method, to speed up registration process and reduce local minima, hierarchical deformation mechanism is applied. Hierarchical deformation mechanism approximates the multivariate energy function by a sequence of significantly lower dimensional energy function of only the selected active voxels. It allows the registration process focus on different sets of active voxels at different stage of image deformation. In the beginning of the image registration, only a small number of active voxels are selected to guide the entire deformation. With the registration performance, the number of active voxels is progressively increased. The active voxels should satisfy the following conditions: 1) They should have distinct geometric characteristics compared with other ambiguous voxels such that they can provide more reliable information to the registration. 2) They should cover the whole structure of the object to guide the registration. Using the active voxel in the registration can alleviate the ambiguity in correspondence detection and thus reduce the local minima. Since the structure of retina is relatively simple, few distinct features can be extracted from 3D OCT image. However, 3D graph-based optimal surface segmentation provides useful surface information which is distinct and robust. First, voxels on the surface have strong boundary characteristic which are very different from other retinal voxels. Second, the number of voxels belonging to the retinal surfaces is much smaller than other non-retinal voxels and they can cover the whole retinal structure. Therefore, the retinal surface voxels are used as active voxels. By defining the weighting function w_T as

$$w_T(x) = \begin{cases} 1, & if \ \text{active} \ voxel \\ 0, & otherwise \end{cases} \quad (10)$$

the correspondence detection can be performed only on active voxels and the active voxels are used to steer the registration. Initially, the number of active voxels is small. Then, the number of active voxels is progressively increased. In our method, active pixels are not selected randomly, but are selected hierarchically by assigning importance coefficients. The importance coefficient is assigned to each voxel following the criterion that the active voxels should be uniformly distributed in the retinal space. Otherwise, the deformation will be probably dominated by only one part of retina over the other parts. Since from the top surface to the bottom surface, the surface-based structure feature $f_1(x)$ is increased by Δs in turn, given N segmented surfaces (here N=7), the importance coefficients for each voxel can be determined as follows:

- If $f_1(x)=\Delta s$ *or* $f_1(x)=N\Delta s$, the importance coefficient $M(x)=(\frac{N-1}{2}+1)\Delta s$;
- If $f_1(x)=\frac{N+1}{2}\Delta s$, the importance coefficient $M(x)=\frac{N-1}{2}\Delta s$;
- If $f_1(x)=(N-i)\Delta s$ *or* $f_1(x)=(1+i)\Delta s$, $(1\le i\le\frac{N-1}{2}-1)$, the importance coefficient $M(x)=(\frac{N-1}{2}-i)\Delta s$;
- If $f_1(x)=0$, the importance coefficient $M(x)=0$;

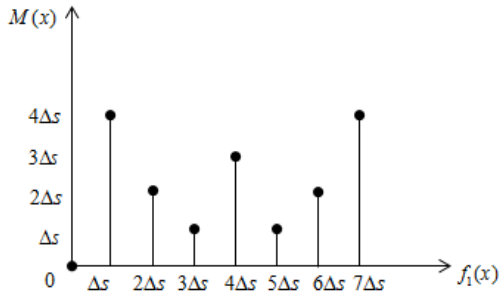


Fig. 6. Demonstration of M(x).

The voxels with larger importance coefficient will join in the process first. As shown in Fig. 6, since surface 1 and surface 7 are the first and last surface of the retina structure, the voxels in these two surfaces are assigned the largest importance coefficient to ensure the major structure of the subject retina deform to the corresponding position of the template retina in the initial stage. Then, the middle surface, surface 4 is added into the process to steer the deformation of the middle structure of retina. After the subject image and the template image have been approximately aligned, voxels on other surfaces are gradually added in to the registration by relaxing the importance coefficient selection criterion to finish the coarse to fine registration. Therefore, the active voxels can be gradually added into the deformation process by four rounds. Table I reports the active voxels used in each round for the image which consisted of 256 B-scans with 512 A-scans and 992 pixels per A-scan. S1-S7 denotes surface1-surface7. Although more advanced method can be used for active voxel selection, we use the simple and elastic importance coefficient strategy to reduce the computation complexity.

TABLE I
ACTIVE VOXELS SELECTION

Round	Selected voxel position	Total voxel no.
1	S1, S7	262144
2	S1, S4, S7	393216
3	S1, S2, S4, S6, S7	655360
4	S1, S2, S3,S4, S5,S6, S7	917504

Given the determined correspondence for all the active voxels, the transformations of all the other voxels in the neighborhood are calculated from the displacement Δ of the active voxels according to a Gaussian function *G*(z),

$$G(z)=\Delta\cdot e^{(-\frac{\|x-z\|^2}{\gamma})} \tag{11}$$

where γ is a control parameter to make the Guassian function close to 0 for the boundary voxels in the neighborhood. The Gaussian function naturally describes the influence of the active voxels on the displacement of other neighborhood voxels, fading away with distance.

In our method, to reduce the discontinuities in the total deformation, two smoothing techniques are used. The first one is the use of smooth constraint in Eq. (7). By requiring the total Laplacian value of displacement field to be as small as possible, the displacement field can be smoothed. The second smoothing technique uses neighboring deformation information. For voxel x, the final deformation at this iteration is calculated by averaging the deformations of its neighborhood as

$$\left(\frac{\sum_j \mathrm{d}(x_j)}{j}-\mathrm{d}(x)\right)\times\alpha+\mathrm{d}(x) \tag{12}$$

where $d(x_j)$ is the deformation of j-*th* neighbor and α is a weighting parameter. In our implementation, we use the average deformation of eight neighbors around voxel x in the B-scan to calculate the final deformation of voxel x.

To further speed up registration process, the multi-resolution framework is also applied. The original image (high-level) is subsampled by a factor of two and a factor of four to form mid-level and low-level image, respectively. The deformation between template and subject is first calculated in the lower level images and then up sampled and linearly interpolated to generate the initialization for the higher level.

F. Summary

The input parameters of this algorithm are the template image T, the subject image S and the general search range δ_k. Given T and S, different levels of the template image and subject image T^k and S^k can be obtained by defining the level parameter $k\in\{0,1,2\}$ corresponding to the low, middle and high levels, respectively. With the general search range δ_k for level k, the algorithm can decide the best matching point in a certain search range by calculating the similarity of the feature values F^k, and determine the complete deformations $d^k(x^S)$ at level k. Therefore, the final output is $d^2(x^S)$, which corresponds to the deformation of the subject image at high level. The proposed OCTRxpert registration algorithm can be summarized as follows:

OCTRexpert Algorithm:

1: Input: T, S and δ_k;
2: Segment retinal layers and correct motion distortions;
3: Compute feature vectors F (c.f. Section III. C);
4: **for** k∈{0,1,2} **do**
5: Get T^k, S^k and F^k;
6: **for** $x \in S^k$ **do**
7: Update position of x according to the deformation $d^{k-1}(x^S)$ if $d^{k-1}(x^S)$ exists;
8: Assign importance coefficient for each x (c.f. Section III. E);
9: Determine active voxels according to importance coefficient (c.f. Section III. E);
10: Correspondence detection for active voxels (c.f. Section III.D);
11: Calculate deformation for each x (c.f. Section III.E);
12: Smooth and save the current deformation $d^k(x^S)$;
13: **end for**
14: **end for**
15. Output: Save $d^2(x^S)$ as the final output;

IV. Experiment and Result

A. Experiment Data

To evaluate the performance of the proposed method, both data from normal subject eyes and data with severe lesions were tested in the experiments. Dataset I and dataset II are retinal OCT data from health people and people with serious CNV, respectively. OCT scans in dataset I are macula-centered SD-OCT scans acquired by using Topcon DRI OCT-1 scanner with image dimension of 512×992×256 and resolution of 11.72×2.62×23.44 um^3. The dataset includes 60 retinal OCT scans from 20 subjects, with each subject having 3 scans acquired in an interval of half a month. OCT scans in dataset II are macula-centered SD-OCT scans acquired using Zeiss 4000 scanner with image dimension of 512×1024×128 and resolution of 11.74×1.96×47.24 um^3. The dataset includes 40 retinal OCT scans from 4 subjects with each subject having 10 scans acquired once a month. The OCT scan from time point 1 was selected as the template. Scans from other time points were registered to the selected template. This study was carried out following the principles of the Declaration of Helsinki and approved by the volunteers and patients for publication. Since the datasets are related to other ongoing research projects, they are not disclosed for the time being, but we have plans to disclose them after completion of the relevant research projects.

B. Evaluation Metrics

To quantitatively assess the accuracy of the registration algorithm, the Dice similarity coefficient (DSC) and the average unsigned surface error (AUSE) which are typically used as the quantitative performance metrics in OCT registration were calculated [35]-[37]. The first metric DSC is a positive performance measure, in which higher value indicates more accurate registration, while another metric is negative performance measure, in which lower value reflects more accurate registration. DSC for each layer between template and registered OCT images is calculated as follows:

$$\mathrm{DSC}_l = \frac{2\left|\mathrm{T}_l \cap \mathrm{S}_l\right|}{\left|\mathrm{T}_l\right| + \left|\mathrm{S}_l\right|} \tag{13}$$

where DSC_l is the voxel-wise overlap ratio for retinal layer l. T_l and S_l are the set of voxels labeled as layer l in the template and the warped subjects. The overlap ratio takes the values between 0 and 1.0. DSC_l of 1.0 corresponds to full overlap between T_l and S_l, while DSC_l of 0 corresponds to none overlap. AUSE is computed for each surface between the template and the registered OCT images by measuring the average absolute Euclidean distance in z-axis. The layer segmentation and labeling were finished manually by retinal specialist as the gold standard.

C. Registration Parameter Selection

Most parameters used in our algorithm are designed to evolve as the algorithm progresses. Therefore, the algorithm is ensured to converge in the end. For example, the search range r_x is designed to reduce with the increasing of iteration number.

$$r_x = 0.5\delta_k \mathrm{e}^{\left(-\frac{I_{ratio}^{\ 2}}{0.16}\right)} + 1 \tag{14}$$

r_x is related to two parameters: δ_k (the fixed general search range) and I_{ratio} (the ratio of the current iteration number and the maximum iteration number). Since r_x is defined as the radius of search size, it should be half of the general search range δ_k. Therefore δ_k is multiplied by 0.5. The constant 0.16 is a control parameter which is designed to make the search range evolve as the algorithm progresses to ensure convergence. With the increasing iteration number, the search range finally reduced to one voxel. The general search range δ_k is large for low-level image to find the coarse correspondence position and gradually reduce for mid-level and high-level to refine the registration. In our implementation, δ_k is estimated according to the retinal thickness to cover enough search area. Known the general retinal thickness T_r, the general search range for different level δ_k can be estimated by

$$\delta_k = \frac{T_r}{4A_{res}} - 4k \tag{15}$$

where A_{res} is the axial resolution and k=0,1,2 corresponds to low-level, mid-level and high-level image, respectively. Considering the general retinal thickness and the axial resolution in the experimental datasets, δ_0 = 24, δ_1=20, and δ_2 = 16 are used in our experiments.

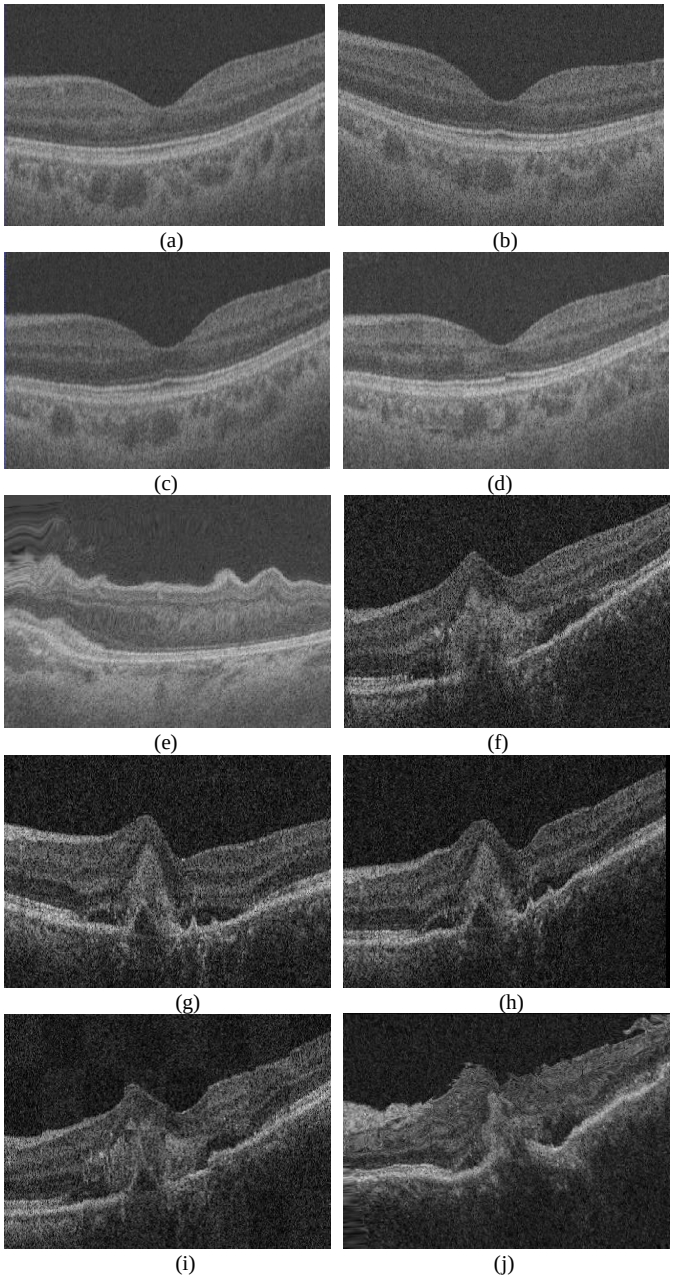


Fig. 7. Registration result of the same slice by using HAMMER and the proposed method. (a) Template image from dataset I. (b) Subject image from dataset I. (c) Registration result of the subject image from dataset I by using the proposed method. (d) Checkerboard comparisons of the proposed method for dataset I. (e) Registration result of the subject image from dataset I by using HAMMER. (f) Template image from dataset II. (g) Subject image from dataset II. (h) Registration result of the subject image from dataset II by using the proposed method. (i) Checkerboard comparisons of the proposed method for dataset II. (j) Registration result of the subject image from dataset II by using HAMMER.

Since the low-level image and the mid-level image were subsampled by the high-level image, the general search range 24 at low-level and 20 at mid-level actually correspond to a search range 96 and 40 at the original image size (high-level), respectively. The neighborhood size of the active voxel is defined in the same way as r_x.

D. Registration Performance Evaluations

1) Registration performance of the proposed method: The registration results of dataset I and dataset II by using the proposed registration method are demonstrated in Fig. 7. In Fig.7, the checkerboard images of dataset I and dataset II are smooth which demonstrate that the subject images are well aligned to the template images by using the proposed method for both datasets.

2) Comparison with HAMMER: Since this work is inspired by HAMMER, experiments were carried out to compare the performance of HAMMER and the proposed method. In the experiments, the identical set of parameters and preprocessing method were used for testing HAMMER and the proposed algorithm. Fig. 7 shows the registration results from dataset I and dataset II by using HAMMER and the proposed method, respectively. The experiment results demonstrate that although HAMMER algorithm works very well on various organ image registrations, the registration results are poor if it is directly used in retinal OCT images. The main reason is that although the use of geometric moment invariants as feature vectors can distinguish different geometric structures of brain, they fail to distinguish the horizontal structure of retinal layers. Furthermore, motion distortions of B-scans in SD-OCT image make the regular interpolation and regularization difficult. In the proposed algorithm, B-scans flattening method is used to correct eye movement artifacts and intensity-based region feature, surface-based structure feature and vessel-like feature are designed instead of GMI feature to distinguish different structures of retina. Table II and Table IV report the Dice similarity coefficient of each retinal layer and the average unsigned surface error by using HAMMER and the proposed registration method, respectively. The p-values of the Dice similarity coefficient and the average unsigned surface error from paired t-tests are shown in Table III and Table V, respectively. The p-values less than 0.05 indicate that the proposed method has statistically significantly better performance. The experiment results on dataset I show that compared with HAMMER, the DSC of the GCL+IPL layer is statistically indistinguishable and the DSCs of the other layers are statistically significantly better. The AUSEs of surface 2 and 3 are statistically indistinguishable and the AUSEs of the other surfaces are statistically significantly smaller. The experiment results on dataset II show that compared with HAMMER, the DSC of the INL layer is statistically indistinguishable and the DSCs of the other layers are statistically significantly better. The AUSE of surface 3 is statistically indistinguishable and the AUSEs of the other surfaces are statistically significantly smaller.

3) Comparison with other OCT registration method: To demonstrate the excellent performance of our algorithm, the proposed method was compared with the state-of-the-art OCT registration method, called conditional correlation ratio (CCR) method [18]. Fig. 8 shows the registration results from dataset I and dataset II by using CCR method and the proposed method respectively. The experiment results reveal that for dataset I, all

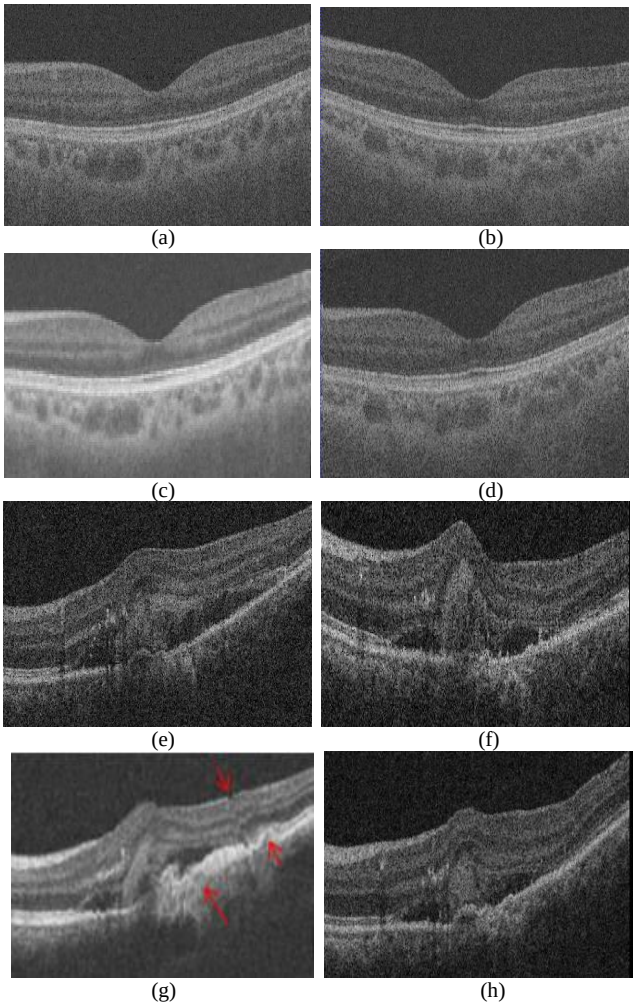


Fig. 8. Registration result of the same slice by using CCR and the proposed method. (a) Template image from dataset I. (b) Subject image from dataset I. (c) Registration result of the subject image from dataset I by using CCR. (d) Registration result of the subject image from dataset I by using the proposed method. (e) Template image from dataset II. (f) Subject image from dataset II. (g) Registration result of the subject image from dataset II by using CCR. (h) Registration result of the subject image from dataset II by using the proposed method. The red arrows mark the distortions after registration.

the two methods achieve satisfying registration results. However, for dataset II, CCR method shows different distortions in the registration results. The main reason is that CCR method is intensity-based non-rigid registration method. Its deformable transformation depends on the intensity similarity and is very sensitive to the speckle noise especially for the CNV case when the intensity contrast is low due to the occurrence of neovascularization. Therefore, for the CNV scans in dataset II, CCR method leads to the distortions of layer structure. Table II and Table IV report the Dice similarity coefficient of each retinal layer and the average unsigned surface error by using CCR method and the proposed registration method, respectively. The p-values of the Dice similarity coefficient and the average unsigned surface error from paired t-tests are shown in Table III and Table V, respectively. The experiment results on dataset I show that compared with CCR, the DSCs of the ONL+ISL layer and RPE are statistically indistinguishable and the DSCs of the other layers are statistically significantly better. The AUSEs of surface 2 and 5 are statistically indistinguishable and the AUSEs of the other surfaces are statistically significantly smaller. The experiment results on dataset II show that compared with CCR, the DSCs of the RNFL layer and GCL+IPL layer are statistically indistinguishable and the DSCs of the other layers are statistically significantly better. The AUSE of surface 1 is statistically indistinguishable and the AUSEs of the other surfaces are statistically significantly smaller. The experiment results of both dataset I and dataset II show that the proposed registration method consistently yields statistically significant improvements in both DSCs and AUSEs compared with the state-of-the-art OCT registration method consistently.

TABLE II
DICE OVERLAP RATIOS OF 6 RETINAL LAYERS

Dataset	Layer	HAMMER	CCR	OCTRexpert
Dataset I	RNFL	0.66 ±0.23	0.87 ±0.02	0.91 ±0.03
	GCL+IPL	0.89 ±0.07	0.83 ±0.08	0.94 ±0.02
	INL	0.8 ±0.03	0.69 ±0.06	0.89 ±0.03
	OPL	0.49 ±0.1	0.74 ±0.08	0.85 ±0.04
	ONL+ISL	0.76 ±0.1	0.86 ±0.09	0.94 ±0.02
	RPE	0.55 ±0.2	0.85 ±0.22	0.97 ±0.01
	Mean	0.69 ±0.14	0.81 ±0.07	0.92 ±0.04
Dataset II	RNFL	0.52 ±0.26	0.81 ±0.11	0.78 ±0.13
	GCL+IPL	0.67 ±0.18	0.73 ±0.11	0.78 ±0.13
	INL	0.59 ±0.19	0.48 ±0.19	0.67 ±0.19
	OPL	0.39 ±0.18	0.46 ±0.2	0.69 ±0.16
	ONL+ISL	0.56 ±0.23	0.78 ±0.15	0.86 ±0.09
	RPE	0.4 ±0.3	0.83 ±0.14	0.89 ±0.08
	Mean	0.52 ±0.11	0.68 ±0.16	0.78 ±0.09

TABLE III
P-VALUES OF OVERLAP RATIOS

Dataset	Layer	OCTRexpert vs -HAMMER	OCTRexpert vs CCR
Dataset I	RNFL	0.0038	0.0367
	GCL+IPL	0.0632	0.0183
	INL	0.0003	0.0078
	OPL	0.0001	0.0333
	ONL+ISL	0.0001	0.081
	RPE	0.0003	0.1787
	Mean	0.0113	0.0593
Dataset II	RNFL	0.0002	0.6305
	GCL+IPL	0.0396	0.2257
	INL	0.1422	0.0011
	OPL	<<0.001	0.0016
	ONL+ISL	<<0.001	0.0261
	RPE	<<0.001	0.0282
	Mean	0.03	0.1522

TABLE IV
AVERAGE UNSIGNED SURFACE ERROR (UM)

Dataset	Surface	HAMMER	CCR	OCTRexpert
Dataset I	1	10.9 ±7.1	1.6 ±3.3	3.3 ±3.7
	2	7.2 ±5.8	4.2 ±4.1	4.5 ±3
	3	3.9 ±5.6	10.2 ±7.1	4.5 ±4.2
	4	9.4 ±4.9	12.3 ±13.1	2.1 ±2.2
	5	17 ±6.9	13.9 ±15.7	9.2 ±7.8
	6	19.7 ±13.9	6.3 ±7.6	1.5 ±2.4
	7	25.3 ±20.2	6 ±9.1	0 ±0
	Mean	13.3 ±7.6	7.8 ±8.6	3.3 ±2.9
Dataset II	1	41.9 ±46.5	15.4 ±22.9	17.2 ±14.7
	2	31.4 ±37.8	17.6 ±22.5	10.8 ±8
	3	8.6 ±8.8	25.7 ±19.6	10.4 ±7.7
	4	20.3 ±22.7	26.6 ±19.5	10.7 ±7.5
	5	29.7 ±40.2	27.4 ±19.7	10.5 ±7.7
	6	34 ±29.9	18.7 ±22.3	6.7 ±5.3
	7	47.6 ±39.5	18.9 ±21.7	5.8 ±5.1
	Mean	30.5 ±32.2	21.5 ±21.2	10.3 ±8

TABLE V
P-VALUES OF AVERAGE UNSIGNED SURFACE ERRORS

Dataset	Surface	OCTRexpert vs HAMMER	OCTRexpert vs CCR
Dataset I	1	0.0101	0.0612
	2	0.1087	0.4979
	3	0.1315	0.0023
	4	0.0002	0.0017
	5	0.0141	0.1326
	6	0.0008	0.0007
	7	0.0012	0.0066
	Mean	0.038	0.1004
Dataset II	1	0.0008	0.1726
	2	0.0044	0.0027
	3	0.3054	<<0.001
	4	0.007	<<0.001
	5	0.01	<<0.001
	6	<<0.001	0.0002
	7	<<0.001	0.0001
	Mean	0.047	0.0251

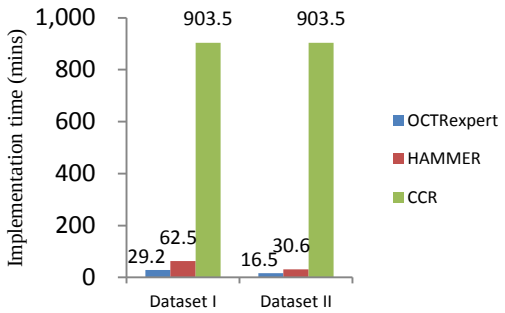


Fig. 9. Implementation time for different methods.

4) Computation complexity: Besides the accuracy evaluation, the computation complexity for different methods is evaluated as well. The proposed method was implemented in C++ and tested on a PC with Intel(R) Core(TM) i7-4790 3.6GHz CPU and 8GB RAM. Fig.9 shows the average implementation time required for each method. The retinal layer segmentation time is excluded. Since in CCR method, the original images are down sampled to the dimension of 512×512×128 during the registration, the average running time is the same for both datasets. Without the hierarchical strategy, CCR method shows very high computation complexity for the two datasets. By using HAMMER method or OCTRexpert method with the hierarchical strategy, the registration computation complexity is greatly reduced. For the OCT image which consisted of 256 B-scans with 512 A-scans and 992 pixels per A-scan, without the hierarchical strategy, more than 130 million voxels have to be matched during the registration process. The computation complexity is extremely high and there exist many ambiguous voxels which may mislead the deformation. By using the hierarchical strategy, the number of the voxels is reduced to around 260000 in the beginning of iteration, and then, gradually increased to approximately 900000 in the end. Between HAMMER method and OCTRexpert method, OCTRexpert achieves lower computation complexity. The main reason is that in HAMMER, the driving voxels are selected by using a fuzzy clustering method. Since the fuzzy clustering method needs to calculate the distance between voxels, it is time consuming especially when the data is large. In our method, we use the efficient driving voxel selection method instead of fuzzy clustering method. Therefore, the computation complexity is further reduced and the computation time of our method is very competitive.

V. CONCLUSION AND DISCUSSION

This paper proposed a retinal OCT image registration approach called OCTRexpert which can be applied to the longitudinal SD-OCT images for both normal and severe lesions subjects. In the proposed method, a novel design-detection-deformation mechanism is designed. According to the segmented 7 retinal surfaces, intensity-based region feature, surface-based structure feature and vessel-like feature are designed for the registration. To speed up registration process and reduce local minima, a hierarchical strategy is also applied. Active voxels are hierarchically selected and the point-to-point correspondences between the subject and the template images are established. The image is then hierarchically deformed according to the detected correspondences in multi-resolution. The experiment results on both healthy OCT dataset and CNV OCT dataset demonstrate that the OCTRexpert registration algorithm overcomes the shortcomings of existing registration methods and consistently yields high registration accuracy. The computation complexity experiments also show the efficiency of the proposed method. Furthermore, the proposed algorithm can also be extended to other pathological cases such as glaucoma, diabetic macular edema and central serous chorioretinopathy. It should be noted that the accuracy of the proposed registration algorithm is affected by the segmentation accuracy. For segmentation errors, we also designed strategies to mitigate this issue. 1) To find the best correspondence of a voxel, the similarity is calculated in a neighborhood rather than a single voxel to improve the robustness. Therefore, even if segmentation fails for one layer, pixels on other layer surfaces in the neighborhood can correct this error. 2) The correspondence detection is not only decided by the segmentation result but also by the intensity and the vessel information. Segmentation error will change the intensity-based feature and reduce the correspondence

similarity during the registration, thus reduce the sensitivity to segmentation errors. By using OCTRexpert, we can achieve a powerful registration for retinal OCT images where the superior performances were demonstrated.

The success of the proposed algorithm is achieved by overcoming the following challenges during the registration of retinal OCT images：1) The large gap between B-scans caused by eye movement during the scan process makes it difficult for interpolation and regularization. To correct the eye movement artifacts, B-scan flattening is used in our pre-processing step. 2) Normal intensity-based registration methods tend to be sensitive to the serious speckle noises in OCT images. To overcome this problem, we apply a feature-based registration method. Three kinds of features are designed to drive the registration. 3) High resolution of OCT data compared with other medical imaging modalities causes dramatic increase of computation complexity for registration. We propose two solutions to speed up the registration process. One is the use of active voxels. The deformation mechanism focuses on different sets of active voxels at different stage of deformation. The number of active voxels is initially small and progressively increases with the registration process. The other solution is the use of multi-resolution framework. The deformation between template and subject is first calculated in the lower resolution and then up-sampled and linearly interpolated to generate the initialization for the next resolution.

Although OCTRexpert overcomes these challenges and shows significant improvements, it still has some limitations: 1) Inconsistent correspondence between the subject and template images is a common limitation for most registration algorithms. In retinal OCT, it always happens in pathological regions. The disrupted or missing layers maybe appear due to pathology (detachment, edema, etc.). In such case, OCTRexpert is designed to relax the matching forces when no matches are found. However, further development will be necessary to estimate the transformation in these regions. 2) OCTRexpert is a feature-based registration algorithm. The registration result highly depends on the accuracy and robustness of the features. For example, the deformable registration relies on the surface information obtained from the surface segmentation. The segmentation accuracy will directly affect the accuracy of the registration. Although, the multi-resolution graph search segmentation results are generally found accurate and robust in our experiment, automatic accurate surface segmentation is not easy for pathological region with dramatic change in the layer structure. Recent works extend deep learning technique to solve complex medical image segmentation problem [38]-[40]. For example, retinal blood segmentation problem [41], pathologic OCT image layer boundary segmentation problem [42], choroidal neovascularization segmentation problem [43], segmentation of optic disk in fundus images, fluid in retinal optical coherence tomography images and fetal head in ultrasound images [44], etc. Deep neural networks also achieved great success in image recognition and analysis. For example, unsupervised image features learning for lung nodule [45], skin disease recognition [46] and lung mass density analysis [47], etc. In the near future, we will focus on the registration techniques of pathological data and consider using deep learning to learn features and improve the segmentation results of retinal vessels and layers. With the improved segmentation results of retinal vessels and layers, the registration accuracy can be further improved. Furthermore, we are going to optimize the codes to obtain additional speeding up to make the reported approach more suitable for clinical practice in the future.

REFERENCES

[1] A. Sotiras, C. Davatzikos, and N. Paragios. "Deformable medical image registration: a survey, " *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1153-1190, June 2013.

[2] J. G. Fujimoto, W. Drexler, J. S. Schuman, and C. K. Hitzenberger, "Optical Coherence Tomography (OCT) in ophthalmology: introduction," *Optics Express*, vol. 17, no. 5, pp. 3978-3979, Mar. 2009.

[3] D. F. Kiernan, W. F. Mieler, and S. M. Hariprasad, "Spectral-domain optical coherence tomography: a comparison of modern high-resolution retinal imaging systems," *American Journal of Ophthalmology*, vol. 149, no.1, pp. 18-31.e2, Jan. 2010.

[4] Y. Jia et al., "Quantitative Optical Coherence Tomography Angiography of Choroidal Neovascularization in Age-Related Macular Degeneration," *Ophthalmology*, vol. 121, no. 7, pp. 1435-1444, July 2014.

[5] A. A. Khanifar, A. F. Koreishi, J. A. lzatt, and C. A. Toth, "Drusen ultrastructure imaging with spectral domain optical coherence tomography in age-related macular degeneration," *Ophthalmology*, vol. 115, no. 11, pp. 1883-1890, Nov.2008.

[6] E. Gao, B. Chen, J. Yang, F. Shi, W. Zhu, D. Xiang, H.Chen, M. Zhang and X. Chen, "Comparison of retinal thickness measurements between the Topcon algorithm and a graph-based algorithm in normal glaucoma eyes," *PloS ONE*, vol. 10, no. 6, pp.e0128925, June 2015.

[7] A. N. Kuo, R. P. McNabb, S. J. Chiu, M. A. El-Dairi, S. Farsiu, C. A. Toth, and J. A. Izatt, "Correction of ocular shape in retinal optical coherence tomography and effect on current clinical measures," *American Journal of Ophthalmology*, vol. 156, no.2, pp. 304–311, Aug. 2013.

[8] T. M. Jørgensen, J. Thomadsen, U. Christensen, W. Soliman, and B. Sander, "Enhancing the signal-to-noise ratio in ophthalmic optical coherence tomography by image registration--method and clinical examples," *Journal of Biomedical Optics*, vol. 12, no. 4, pp. 041208, July 2007.

[9] M. Niemeijer, M. K. Garvin, K. Lee, B. V. Ginneken, M. D. Abràmoff, M. Sonka, "Registration of 3D spectral OCT volumes using 3D SIFT feature point matching," in *Proc. SPIE*, Orlando, USA, 2009, pp. 72591I- 72591I -8.

[10] Y. M. Liew, R. A. Mclaughlin, F. M. Wood, and D. D. Sampson, "Motion correction of in vivo three-dimensional optical coherence tomography of human skin using a fiducial marker," *Biomedical Optics Express*, vol. 3, no. 8, pp. 1774-1786, June 2012.

[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*, vol. 31, no, 7, pp. 1337-1345, Jan. 2012.

[12] S. Ricco, M. Chen, H. Ishikawa, G. Wollstein, and J. Schuman, "Correcting Motion Artifacts in Retinal Spectral Domain Optical Coherence Tomography via Image Registration," in *International Conference on Medical Image Computing & Computer-assisted Intervention- MICCAI*, London, UK,: Springer-Verlag, 2009, pp. 100–107.

[13] M. Chen , A. Lang, H. S. Ying, P. A. Calabresi, J. L. Prince, and A. Carass, "Analysis of macular OCT images using deformable registration," *Biomedical Optics Express* , vol. 5, no. 7, pp. 2196-2214, July 2014.

[14] Y. Zheng, R. Xiao, Y. Wang, and J. C. Gee, "A generative model for OCT retinal layer segmentation by integrating graph-based multi-surface searching and image registration, " in *International Conference on Medical Image Computing & Computer-assisted Intervention- MICCAI*, Nagoya, Japan, 2013, pp. 428-35.

[15] H. Zhang, Z. Li, X. Wang, and X. Zhang, "Speckle reduction in optical coherence tomography by two-step image registration," *Journal of Biomedical Optics*, vol. 20, no. 3, pp. 36013, Mar. 2015.

[16] Q. Wei, W. Zhu, D. Xiang and X. Chen, "Nonrigid registration of 3D longitudinal optical coherence tomography volumes with choroidal

neovascularization, " in *Proc. SPIE*, Orlando, USA,2017, pp. 101330X .
[17] X. Du, L. Gong, F. Shi, X. Chen, X. Yan and J. Zheng, "Non-rigid Registration of Retinal OCT Images Using Conditional Correlation Ratio," Fetal, Infant and Ophthalmic Medical Image Analysis, pp.159-167, 2017.
[18] L. Gong, C. Zhang, L. Duan, X. Du, H. Liu, X. Chen and J. Zheng, "Non-rigid image registration using spatially region-weighted correlation ratio and GPU-acceleration," *IEEE Journal of Biomedical and Health Informatics*,vol.23, no.2, pp.766-778,2019.
[19] D. Shen and C. Davatzikos, "HAMMER: Hierarchical Attribute Matching Mechanism for Elastic Registration," *IEEE Transactions on Medical Imaging*, vol. 21, no. 11, pp. 1421-1439, Nov. 2002.
[20] D. Shen, "Image registration by local histogram matching," *Pattern Recognition*, vol. 40, no. 4, pp. 1161-1172, Apr. 2007.
[21] K. Li, X. Wu, D. Z. Chen, and M. Sonka, "Optimal Surface Segmentation in Volumetric Images—A Graph-Theoretic Approach," I*EEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 28, no. 1, pp. 119-134, Nov. 2005.
[22] M. K. Garvin, M. D. Abramoff, R. Kardon, S. R. Russell, X. Wu, and M. Sonka, "Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search, " *IEEE Transactions on Medical Imaging*, vol. 27, no.10, pp. 1495-1505, Apr. 2008.
[23] M. K. Garvin, M. D. Abràmoff, X. Wu, S. R. Russell, T. L. Burns, and M. Sonka, "Automated 3-D Intraretinal Layer Segmentation of Macular Spectral-Domain Optical Coherence Tomography Images," *IEEE Transactions on Medical Imaging*, vol. 28, no, 9, pp. 1436-47, Sept. 2009.
[24] Q. Song, J. Bai, M. K. Garvin, M. Sonka, J. M. Buatti, and X. Wu, "Optimal multiple surface segmentation with shape and context priors," *IEEE Transactions on Medical Imaging*, vol. 32, no, 2, pp. 376-86, Feb. 2013.
[25] P. A. Dufour, L. Ceklic, H. Abdillahi, and S. Schroder, "Graph-Based Multi-Surface Segmentation of OCT Data Using Trained Hard and Soft Constraints," *IEEE Transactions on Medical Imaging*, vol. 32, no. 3, pp. 531-543, Oct. 2013.
[26] X. Chen, J. K. Udupa, U. Bağcı, Y. Zhuge, and J. Yao, "Medical image segmentation by combining graph cut and oriented active appearance models, " *IEEE Trans. Image Processing*, vol. 21, no. 4, pp. 2035 – 2046, Jan. 2012.
[27] D. Xiang, U. Bagci, C. Jin, F. Shi, W. Zhu, J. Yao, M. Sonka, X. Chen, "CorteXpert: A Model-based Method for Automatic Renal Cortex Segmentation," *Medical Image Analysis*, vol. 42, pp.257-273, Aug. 2017.
[28] X. Xu, M. Niemeijer, Q. Song, M. Sonka, M. K. Garvin, J. M. Reinhardt, M. D. Abramoff, "Vessel Boundary Delineation on Fundus Images Using Graph-Based Approach, " *IEEE Transactions on Medical Imaging*, vol. 30, no.6, pp. 1184-1191, Jan. 2011.
[29] X. Chen, M. Niemeijer, L. Zhang, K. Lee, M. D. Abramoff, and M. Sonka, "Three-dimensional segmentation of fluid-associated abnormalities in retinal OCT: probability constrained graph-search-graph-cut," *IEEE Transactions on Medical Imaging*, vol. 31, no, 8, pp. 1521-1531, Mar. 2012.
[30] K. Lee, M. Niemeijer, M. K. Garvin, Y. H. Kwon, M. Sonka, and M. D. Abramoff, "Segmentation of the optic disc in 3-D OCT scans of the optic nerve head, " *IEEE Transactions on Medical Imaging*, vol. 29, no.1, pp. 159-168, Sep. 2010.
[31] F. Shi, X. Chen, H. Zhao, et al. "Automated 3-D Retinal Layer Segmentation of Macular Optical Coherence Tomography Images with Serous Pigment Epithelial Detachments," *IEEE Transactions on Medical* Imaging, vol. 34, no.2, pp. 441-452, Sep. 2015.
[32] D. Xiang, H. Tian, X. Yang, F. Sh, W. Zhu, H. Chen and X. Chen, "Automatic Segmentation of Retinal Layer in OCT Images with Choroidal Neovascularization," *IEEE Transactions on Image Processing*, vol. 27, no.12, pp. 5880-5891, Dec. 2018.
[33] X. Chen, P. Hou, C. Jin, W. Zhu, X. Luo, F. Shi, et al., "Quantitative Analysis of Retinal Layer Optical Intensities on Three-Dimensional Optical Coherence Tomography," *Investigative ophthalmology & visual science*, vol. 54, no, 10, pp. 6846–51, 2013.
[34] A. Frangi, W. J. Niessen, K. L.Vincken, M. A.Viergever, "Multiscale vessel enhancement filtering," in *International Conference on Medical Image Computing & Computer-assisted Intervention- MICCAI*, London, UK,: Springer-Verlag, 1998, pp. 130-137.
[35] Z. Li, F. Huang, J. Zhang, T. Tan, "Multi-modal and multi-vendor retina image registration,",*Biomedical Optics Express*, vol. 9, no, 2, pp. 410-422, 2018.
[36] M. P. Heinrich, J. Schnabel, and F. V. Gleeson, "Non-rigid multimodal medical image registration using optical flow and gradient orientation," in *Proc. Medical Image Understanding and Analysis*, Warwickshire, UK, 2010, pp. 1–5.
[37] B. Antony, M. D. Abràmoff, L. Tang, W. D. Ramdas, J. R. Vingerling, N. M. Jansonius, K. Lee, Y.H. Kwon, M. Sonka and M. K. Garvin, "Automated 3-D method for the correction of axial artifacts in spectral-domain optical coherence tomography images," *Biomedical Optics* Express, vol. 2, no..8, pp. 2403-2416, Aug. 2011.
[38] G. Litjens, T. Kooi, B. E. Bejnordi, et al, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, no. 9, pp. 60-88, 2017.
[39] L. Jin, P. Yi, L. Min, et al, "Applications of Deep Learning to MRI Images:A Survey," *Big Data Mining & Analytics*, vol. 1, no. 1, pp. 3-20, 2018.
[40] H. Greenspan, B. V. Ginneken, and R. M. Summers, "Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1153-1159, Apr. 2016.
[41] P. Liskowski and K. Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE Transactions on Medical Imaging*, vol. 35, no. 11, pp. 2369–2380, Nov. 2016.
[42] L. Fang, D. Cunefare, C. Wang, R. H. Guymer, S. Li and S, Farsiu, "Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search," *Biomedical Optics Express*, vol. 8, no. 5, pp. 2732-2744, Apr. 2017.
[43] X. Xi, X. Meng, L. Yang, X. Nie, G. Yang, H. Chen, X. Fan, Y. Yin and X. Chen, "Automated segmentation of choroidal neovascularization in optical coherence tomography images using multi-scale convolutional neural networks with structure prior," Multimedia Systems, vol. 5801, pp. 1-8, Jan. 2018.
[44] Y. Rong, D. Xiang, W. Zhu, et al., "Deriving external forces via convolutional neural networks for biomedical image segmentation," *Biomedical Optics Express*, vol.10, no.8, pp. 3800-3814, 2019.
[45] M. Chen, X. Shi, Y. Zhang, D. Wu and M. Guizani, "Deep Features Learning for Medical Image Analysis with Convolutional Autoencoder Neural Network," *IEEE Transactions on Big Data*, vol. 3, pp. 1-1, 2017.
[46] M. Chen, P. Zhou , D. Wu , et al, " AI-Skin : Skin Disease Recognition based on Self-learning and Wide Data Collection through a Closed Loop Framework," *Information Fusion*, vol. 54, pp. 1-9, Feb. 2020.
[47] B. Zhou and X. Zhang, "Lung Mass Density Analysis using Deep Neural Network and Lung Ultrasound Surface Wave Elastography," *Ultrasonics*, vol.89, pp. 173-177, 2018.