Retinal OCT Image Registration: Methods and Applications

Lingjiao Pan and Xinjian Chen*, *IEEE Senior Member*

Abstract: Retinal image registration is a critical task in the diagnosis and treatment of various eye diseases. And as a relatively new imaging method, optical coherence tomography (OCT) has been widely used in the diagnosis of retinal diseases. This paper is devoted to retinal OCT image registration methods and their clinical applications. Registration methods including volumetric transformation-based registration methods and image features-based registration methods are systematically reviewed. Furthermore, to better understanding these methods, their applications in correcting scanning artifacts, reducing speckle noise, fusing and splicing images and evaluating longitudinal disease progression are studied as well. At the end of this paper, registration of retina with serious pathology and registration with deep learning technique are also discussed.
[1]

Index Terms-medical image registration, optical coherence tomography, retina, deep learning.

1 Introduction

Optical coherence tomography (OCT) imaging technology is a non-invasive, non-contact biological tissue imaging technology. Compared with magnetic resonance imaging (MRI), computed tomography (CT) and other imaging techniques, OCT is relatively new and has developed rapidly in recent years. At present, its main application is for the study of retinopathy, such as macular hole, glaucoma, retinal detachment and age-related macular degeneration, all of which can lead to blindness [1-5]. The principle of OCT is the estimation of the depth at which a specific backscatter originated by measuring its time of flight [6]. As shown in Fig. 1, the light is split into two beams by a beam splitter. One beam reflected from the retinal tissue is called the sample arm and the other beam reflected from the reference mirror is called the reference arm. The interferogram energy between the sample arm and the reference arm is converted into image intensity by CMOS sensor, CCD or photo sensor. A depth scan generated by the interferogram intensities is called A-scan. By acquiring a series of A-scans in a raster scanning pattern, the cross sectional slicer called B-scans are generated. Composing successive B-scans yields a 3D OCT imaging of retina as shown in Fig. 2.

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 particular, medical image registration can be used for studying longitudinal and cross-sectional data, quantitatively monitoring disease progression and guiding computer assisted diagnosis and treatments [7]. In the past decades, many registration techniques have been developed for various types of data and applications. Since OCT imaging is relatively new compared with other medical imaging modalities such as MRI and CT, the requirement for processing OCT images has a shorter history. With the fast development of OCT technique, the demand for advanced image analysis techniques is rapidly growing. As one of the main tasks in retinal OCT image processing, retinal OCT image registration can be used for studying longitudinal and cross-sectional data, quantitatively monitoring disease progression and guiding computer assisted diagnosis and treatments [7]. Nevertheless, the development of registration technique which enables more precise and quantitative comparison can be challenging for 3D retinal OCT imaging. The main reasons are as follows: 1) OCT image is inherently noisy. OCT image is inevitably disturbed by noises, especially speckle noise [8-10]. A raw OCT image usually has very poor image quality due to speckle noise, which often obscures the retinal structures. Intensity-based registration methods are usually sensitive to speckle noise and have poor registration performance for low quality images. 2) Retina is a non-rigid moving organ inside a moving body. Only rigid transform is not enough to describe the deformation of retina. Furthermore, the non-rigid mixed eye motion makes the registration more complex. 3) Compared with the brain, the retina exhibits fewer stable and accurate anatomical landmarks, especially when the lesions occur, the retinal structure can change drastically. Therefore, it brings great challenges to feature detection and matching for registration. 4) Compared with other traditional imaging methods such as MRI, OCT has higher image resolution, which greatly increases the computational complexity of non-rigid registration. These characteristics lead to relatively poor performance when directly applying existing medical image registration approaches to retinal OCT images. The experiments show that the highly-rated deformable registration algorithms such as SyN, DRAMMS and HAMMER are unreliable when applied to the whole OCT image and take much longer time to finish registration process than for a standard MR brain registration [11, 12]. The successful development of accurate and fast OCT image registration technology can greatly enhance our understanding of retinal morphology and function and help the computer-aided disease analysis and diagnosis.

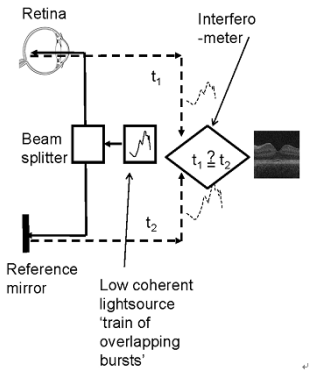
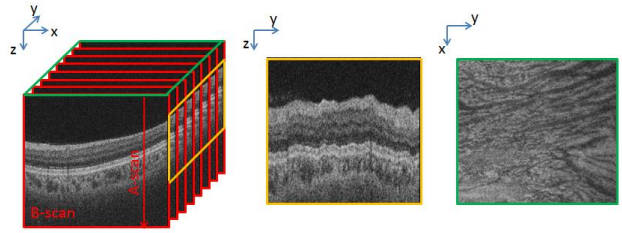


Fig. 1 Schematic diagram of OCT image formation. (Cited from Ref. [6].)

This work was supported by the National key research and development program under Grant 2018YFA0701700 and the National Natural Science Foundation of China under Grant 62001196. Lingjiao Pan is with the School of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou, China (E-mail: jsjshedy@jsut.edu.cn) and Xinjian Chen is with the School of Electronic and Information Engineering, Soochow University, Suzhou, China (E-mail: xjchen@suda.edu.cn). *Asterisk indicates corresponding author*.



(a) 3D view (b) y-z view (c) x-y view

Fig. 2 Example of a 3D OCT scan of retina.

Despite a lot of research works have been done on medical image registration methods [13-22], very few review papers focusing on retinal image registration have been published. In the two review papers on medical image registration, retinal image registration was briefly mentioned [7, 23]. Saha *et al.* reviewed the state-of-the art registration methods of 2D fundus imaging [24]. Baghaie *et al.* [25] briefly reviewed the use of image registration techniques in OCT image analysis. Sonka *et al.* provided a brief overview of registration approaches for both 2D fundus imaging and 3D OCT imaging in a separate section in their paper [6]. Our review is devoted to 3D retinal OCT imaging registration methods and their application issues. To the best of our knowledge, this is the first paper to systematically summarize the research work in the field of OCT retinal registration. Compared with other documents which cover the theoretical aspects of image registration, our paper has some novel points. 1) This paper systematically reviews the registration methods of retinal OCT imaging, including volumetric transformation-based registration methods and image features-based registration methods. This paper not only simply introduces the registration methods, but also discusses their advantages and disadvantages. 2) This paper summarizes the main clinical applications of retinal OCT registration technology and provides ideas for how to use registration technology to assist clinical computer diagnosis and analysis. 3) This paper discusses the technical difficulties of retinal OCT image registration, especially the image registration of retinopathy. How to combine deep learning technique with retinal OCT registration method and how to embody the robustness of deep learning while retaining the advantage of traditional registration methods can be studied from these reviewed works.

2 Registration Methods

In this paper, retinal OCT image registration methods are divided into two major classes, namely volumetric transformation-based registration methods and image feature-based registration methods. Fig.3 provides a brief overview of these methods. More details are given in the following sections.

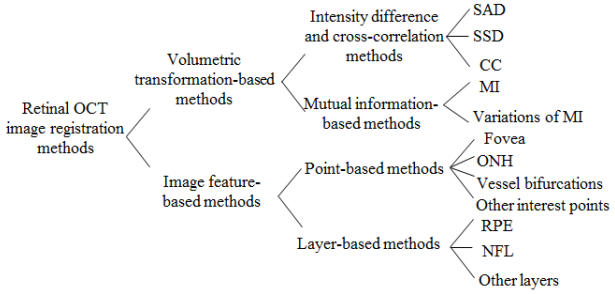


Fig. 3 Overview of the methods for retinal OCT image registration

2.1 Volumetric Transformation-based Registration Methods

Volumetric transformation-based registration methods seek to maximize the voxel similarity between the template image and subject image. These methods do not need to construct a specific anatomical model, but generally requires that the template image and the subject image are obtained by the same imaging protocol.

2.1.1 Intensity Difference and Cross-correlation Methods

SAD (sum of absolute difference) and SSD (sum of squared intensity difference) are two commonly used intensity similarity criteria. In SAD method, the template image and the deformed image are subtracted pixel-by-pixel and the sum of absolute intensity differences of all the pixels in the subtracted image is calculated. The best deformation is determined by minimize the SAD value. SSD measure is similar to SAD measure, but the squared intensity difference is computed instead of the absolute intensity difference. Cheng *et al.* [26] used SAD measure to align a group of A-scans in the neighborhood of the target A-scan and reconstructed the target A-scan using low rank matrix completion to reduce speckle noises in OCT images. Wang *et al.* [27] adopted the Matlab built-in Powell's optimizer to minimize SSD between different OCT images in the image registration process to alleviate the influences of motion artifacts. Chen *et al.* [11] checked the SSD error produced by the current deformation relative to the target image in their energy function to solve the deformable transformation for each A-scan.

Since these registration algorithms are based on intensity differences, the similar intensity ranges and consistent intensity values in a particular tissue type between the subject and target images are required. However, OCT data often shows considerable intensity variability. One potential improvement to address this is to use a similarity metric that is more robust to intensity variability. It is suggested that the cross-correlation measure may be a good replacement for SAD and SSD [11]. When there is a linear relationship between the intensity values of the images to be registered, cross-correlation is the best measure for registration. The optimum alignment is determined by the maximum value of the cross-correlation. Zhang *et al.* [28] proposed a two-step image registration scheme that combined global and local registrations for speckle reduction. The method began with a global registration to compensate for overall motion. The cross-correlation coefficient was selected as the cost function that determined the similarity between the template image and subject image. Then each A-scan was aligned by cross-correlation using a graph-based algorithm in local registration. This method did not rely on any information about the retinal layer boundaries and was able to correct translation, rotation, and local deformation in the axial direction. Zheng *et al.* [29] used registration to assist OCT retinal layer segmentation. In their paper, the deformable registration was conducted using the ANTS Symmetric Normalization algorithm [30], with the cross-correlation similarity metric and a Gaussian regularizer. After registration, reference segmentations from each of the training images were warped into the test image space.

2.1.2Mutual Information- based Methods

Mutual information (MI) is derived from information theory. It quantifies the amount of statistical information that one image depends on the other or the statistical dependence between two random variables. The mutual information is maximized for the optimal alignment. Since the nature of the relationship between the image intensities in the registered image is not assumed [22],

mutual information-based methods can be applied to both mono-modal and multi-modal registration. Although mutual information-based registration is widely used for both mono-modal and multi-modal medical image registration of various parts of the human body, some recent studies have confirmed that the mutual information may not be a universal method to solve all registration problems, especially for thin structure images, such as retinal images [22, 24]. In retinal image registration, one solution to improve the registration accuracy is combining some other information into the calculation of mutual information. An example is to embed spatial information into the calculation of mutual information such as regional mutual information (RMI) [31], localized mutual information (LMI) [32], conditional mutual information (CMI) [33] and spatially encoded mutual information (SEMI) [34]. In the recent work of Gong *et al.* [35], the spatially region-weighted correlation ratio (SRWCR) which incorporates the spatial information into the functional mapping relationship was used as an alternative intensity-based metric for image registration. Their experiment results show that SRWCR is more suitable for the registration of retinal OCT image compared with MI and SEMI. And another solution to improve the registration accuracy is to combine the mutual information with some other retinal features, also known as hybrid method. Wei *et al.* [36] applied a non-rigid registration method which combined normalized mutual information based registration with the feature landmark points based coherent point drift (CPD) registration to align retinal OCT volumes.

2.2 Image Features-based Registration Methods

Different from volumetric transformation-based methods, feature-based methods utilize a number of distinct anatomical features to determine transformation parameters. Such features are normally distinct landmark points, curves, surfaces or a combination of them. Feature-based registration methods are not as general as volumetric transformation-based methods. For different registration problem, different features have to be designed. However, feature-based registration methods still have several advantages. First, 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. Second, well selected features have less ambiguity than intensity similarity in defining correspondences especially for aligning images with intensity distortions. Feature-based registration methods have been widely used in retina registration [37-42]. These methods can be generally divided into point-based registration methods and layer-based registration methods.

2.2.1Point- based Methods

Point-based registration methods are mainly used for rigid transformation and affine transformation. If the landmark points used for computing the registration transformation are rich enough, they can also be applied to more complex deformations in theory. Since SD-OCT image is inherently noisy and the structure of retina can drastically change for diseased retina, few stable and distinct landmarks can be extracted for retina. The normally used landmarks obtained from OCT images include the fovea position as the functional center of vision, the optic nerve head connecting the eye to the brain and the retinal vasculature. The fovea position is usually used for initial registration. For example, Chen proposed to use the position of fovea which was approximated by the superior point of the thinnest portion of the retina as the landmark point in their global translation. As a result, the fovea in the subject was aligned with the fovea in the target image [11]. However, this method does not work for the data with severe lesions such as choroid neovascularization (CNV) where the fovea is no longer the thinnest portion of the retina. The optic disc location and size are also used in initial registration for multimodality images [43]. Since images from different modalities differ in size and resolution, the optic discs are registered to bring all images to a similar scale and resolution [44]. For fundus-fundus registration and OCT-fundus registration, most existing registration approaches utilize features derived from retinal vasculature [43, 45-49]. After the vessel extraction and skeletonization steps, the branching point can be easily used as a stable marker to determine the image-to-image correspondence. Since the retinal vasculature's structure is stable between modalities, bifurcations are reasonable landmark point candidates. However, these methods require the blood vessel segmentation which could be challenging for poor quality images. To overcome this limitation, interest point detection schemes including scale-invariant feature transform (SIFT) [40, 50, 51], salient feature region (SFR) [41], speeded up robust features (SURF) [52-54], Low-dimensional Step Pattern Analysis (LoSPA) [69] and so on are adopted. Instead of bifurcations, interest points are detected and used as landmark point candidates. For example, the image corner points are always used as the interest points since corner points are sufficient and uniformly distributed across the image domain. In Harris-PIIFD framework, partial intensity invariant feature descriptor (PIIFD) for all corner points was calculated and a bilateral matching technique was applied to identify corresponding PIIFDs matches between image pairs [55]. PIIFD is invariant to image rotation, partially invariant to image intensity, affine transformation, and viewpoint/perspective change. In another work, image corners detected by the features from accelerated segment test (FAST) were used as the landmark points. The histograms of oriented gradient (HOG) features were calculated form the neighborhood of each landmark point. Finally, the best transformation parameter was computed using the random sample consensus (RANSAC) method [44].

2.2.2Layer- based Methods

Retina has multi-layer structure. From top to bottom, the retina can be divided into ten layers including nerve fiber layer (NFL), ganglion cell layer (GCL), inner plexiform layer (IPL), inner nuclear layer (INL), outer plexiform layer (OPL), outer nuclear layer and inner segment layer (ONL+ISL), connecting cilia (CL), outer segment layer (OSL), Verhoeff's membrane (VM), and retinal pigment epithelium (RPE). Fig. 4 shows macular centered OCT B-scan of a normal eye with 11 surfaces that define 10 retinal layers. Layer-based registration methods can be applied to both rigid registration and non-rigid registration. The retina layers need to be segmented in advance before registration and it is important to choose which surface to use for registration.

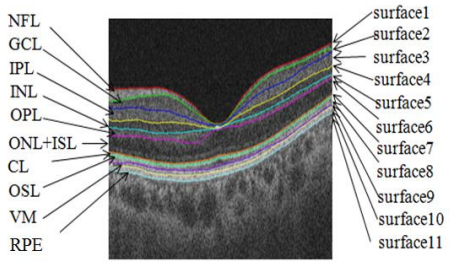


Fig. 4 OCT image with 10 retinal layers.

The top and bottom surfaces of the retina are relatively easy to segment compared with other surfaces, and they are the two most commonly used layer surfaces in layer-based registration. Usually these two surfaces are used for initial registration before the deformable transformation [11]. Layer flattening is a step to flatten the reference surface by A-scan registration and is used to correct eye movement artifacts and provide a more consistent retinal shape for visualization [56-60]. The reference surface for flattening should be robust and easy to detect. In Garvin's approach [59], by segmenting the RPE surface in a lower resolution, fitting a thin-plate spline to RPE and vertically realigning the OCT volume to make RPE completely flat, the misaligned B-scans caused by motion distortion can be corrected. However, for 3D OCT data with serous retinal disease such as pigment epithelial detachments (PED), flattening with RPE may ruin the natural curvature of the surfaces, e.g., the dome like shape of the elevated RPE and the smooth surface that forms the bottom of the retina. To keep the natural curvature of retinal layers, alignment using surface 1 is applied [61-63] instead of image flattening. In Shi's method, after segmentation of surface 1, the average z position of the left most and right most 20% points in surface 1 in each B-scan was used to estimate the displacement of each B-scan. Each B-scan was thus shifted up or down to make sure that the average z positions of peripheral surface 1 become the same. Adding more layer information can be used to deal with more complex registration problems. In OCTRexpert algorithm [12], 7 out of 11 surfaces are used for registration. Different importance coefficient was assigned to different surfaces to hierarchically deform the image. The voxels in the top and bottom surfaces were assigned with 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. After the subject image and the template image have been approximately aligned, voxels on other surfaces were gradually added into the registration by relaxing the importance coefficient selection criterion to finish the coarse to fine registration.

Theoretically, layer-based registration method can be applied to many pathological cases such as glaucoma, diabetic macular edema and central serous chorioretinopathy. However, the registration performance highly depends on the accuracy of layer segmentation. As one of the main tasks of OCT image processing, retinal layer segmentation technology is well developed and work quite well on retinal OCT data sets when no dramatic change in the layer structure happens. Representative segmentation algorithm is graph-based surface segmentation algorithm [64]. When additional structures, such as intraretinal cysts, subretinal, or sub-RPE fluid exist, accurate layer segmentation becomes challenging.

3 Applications

Medical image registration is mainly used to match multiple images taken from different viewpoints, different sensors or different time points. It is an important tool to achieve more accurate and quantitative comparison of disease. As a relatively new imaging technology compared with MRI and CT, OCT imaging is mainly used in the diagnosis of retinal diseases. The main applications of OCT image registration in retinal diagnosis and treatment are as follows.

3.1 Application in Correcting Scanning Artifacts

In ophthalmic application of OCT, involuntary eye movement is one of the biggest problems, which can cause artifacts, distortions and missing regions. Involuntary eye movement caused by heart beat and respiration during OCT acquisition process is known as axial motion. It can be observed by viewing the data orthogonal to the plane of acquired B-scans. Another artifact which is caused by eye movement in the *en face* view of the OCT data is called transverse motion. These motion artifacts cause difficulty in 3D image analysis. One way to deal with motion artifacts is to use eye tracking equipment to compensate the eye movement during image acquisition by hardware solution. And another way is the software approach which is more general and applicable without the need for additional imaging equipment. Recently, several commercial OCT scanners have already included additional eye motion correction hardware. For example, Heidelberg Engineering HRA-OCT Spectralis uses a Scanning Laser Ophthalmoscopy (SLO) device with the OCT to track the eye movements during imaging. The tracking system is very efficient for correcting transverse motion. Since SLO and OCT data are captured simultaneously, the motion estimates from SLO image registration can be applied to correct transverse motion artifacts in the simultaneously acquired OCT volumes [65]. With the help of additional hardware transverse motion can be minimized, while axial motion correction still requires a software solution [66]. Common software approaches for reducing these scanning artifacts often use cross-correlation maximization technique. The axial motion is corrected by maximizing cross-correlation of either A-scans or B-scans [66-72]. For retinal layer segmentation, flattening [57, 63, 73] or B-scan alignment [61, 62] is often used before segmentation to correct eye movement artifacts and provide a more consistent retinal shape for visualization. In flattening process, each A-scan is aligned to a reference surface to obtain flattened images. For B-scan alignment, the displacement of each B-scan is estimated and each B-scan is thus shifted up or down to make sure that the average z positions of the reference surface become the same. The upper or lower surface of the retina is usually used as the reference surface. Fig.5 shows the axial motion correction result after B-scan alignment.

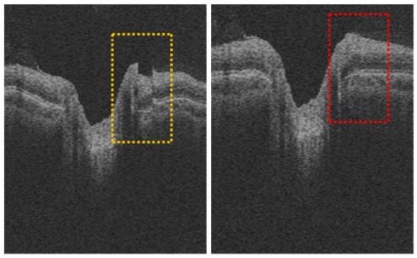


Fig. 5 The y-z image before and after B-scan alignment.

As an extension of OCT technique, optical coherence tomography angiography (OCTA) has been widely used in ophthalmology in recent years. OCTA is a non-invasive approach which provides both functional and structural information. It can also be affected by motion artifacts. Previous works used only information from OCTA image for motion estimation and correction [74-77]. However, when patients show low or no blood perfusion due to circulation abnormalities, the features that can be used for motion estimation will be reduced. Recently, more and more works focus on using both OCT and OCTA image for motion correction [72, 78, 79]. A multi-image cross-correlation method that employs spatial features in both OCTA and OCT images to improve transverse motion estimation performance was introduced in [72]. By combing OCT and OCTA image, the performance was improved. Zang *et al.* reported a 3D registration algorithm using both angiographic and structural OCT information [79]. The vasculature on *en face* retinal OCTA was used to correct

transverse motion and the inner limiting membrane surface on OCT was used as the reference to correct axial registration.

3.2 Application in Reducing Speckle Noise

Speckle noise inherently exists in OCT image and degrades the image quality. Speckle noise not only affects visual diagnosis, but also increases the challenges in automatic image analysis. The common approach to reduce speckle noise and improve the OCT image quality is to calculate the average image through overlapping scans. The data to be averaged should be acquired from the same position. However, in practice, slight position changes always happen during scanning due to the movement of the beam and the sample. Therefore, registration is important to overcome the misalignment. Various registration-based algorithms are proposed to reduce speckle noise in OCT image [26-28, 80-86]. Cheng *et al*. proposed to formulate the speckle noise reduction problem as a low rank matrix decomposition problem in 3D OCT. The proposed method aligned a group of A-scans in the neighborhood of the target A-scan and reconstructed the underlying clean A-scan using low rank matrix completion through iterative bilateral random projections [26]. The A-scan registration was done in two steps including a global B-scan alignment and a local A-scan alignment step. Liu *et al.* applied an axial and horizontal registration method to suppress the speckle noise in retinal OCT adapted to single-line HD mode hardware implementation [86]. In their method, regularized dynamic programming algorithm and hill-climbing algorithm were used to estimate the axial and lateral shifts and correct motions efficiently. Thus the image quality of averaged OCT images can be improved. In deep learning-based denoising methods, the input pairs of original noisy image and clean ground truth image are needed for deep network training. However, there's no ground truth image readily available. Therefore, registration and averaging of B-scans is also a necessary step to obtain high quality training image set [87-89].

3.3 Application in Image Fusing and Splicing

Registration is the first and important step in image fusing and splicing. Retinal image fusion and splicing can be used to enlarge retinal coverage by registering imaging data from different viewpoints or to analyze multimodal data by registering images from different sensors.

3.3.1 Multi-view analysis

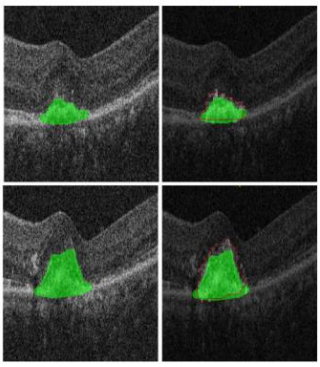
Clinically, it is expected that a larger field of retinal vision can be observed, and the retinal features across larger regions can be investigated. The main factor limiting the field of vision is the scanning speed of OCT scanner, because the image acquisition time is constrained by the patient's ability to gaze without eye movements (usually less than two seconds). Image fusing is an alternative strategy for obtaining wide-field OCT images. However, most of the works are focused on the stitching of 2D fundus photos, while the work on the registration and stitching of 3D OCT data is relatively less. In the early work, only pairs of OCT image are involved, because the multiple images mosaic is more challenging [40]. [90] is the first work for mosaicing multiple OCT images. In the work, eight partially overlapping OCT data sets were merged together to form a new wide-field 3D OCT data set. First, the OCT fundus images were registered using blood vessel ridges as the feature of interest. Then the OCT data sets were merged to form a full 3D montage using cross correlation. Inspired by this work, a novel approach was proposed for layer segmentation of multi-field retinal OCT images [42]. Instead of segmenting each field independently and then stitching the results together, co-segmentation of all the fields was involved to ensure consistent segmentation in the overlapped areas. All the fields were segmented simultaneously after 2D en-face alignment using SURF descriptors. For 2D en-face alignment, simultaneously acquired SLO images were used instead of projected OCT images with low resolutions. They showed that the NFL thickness map obtained by the independent segmentation method of each field had obvious distortion in the overlapping area, while the thickness map obtained by the proposed co-segmentation method was distortion-free. Clinical studies have also validated the benefit of wide-field OCTA, since it allows the observation of overall vascular abnormalities for clinicians. Features-based automated registration and montage for wide-field OCTA can be found in [91].

3.3.2 Multimodality analysis

In clinical practice, doctors usually use a variety of imaging information for better diagnosis. Compared with OCT images, fundus imaging can provide different information about the retina. Fundus imaging acquires a 2D representation of the 3D retina by means of reflected light. The typically used fundus imaging methods in clinic include color fundus photography (CFP), SLO, fluorescein angiography (FA) and so on. CFP uses the imaging light of red, green and blue wavebands. It can be applied for detecting retinal abnormalities such as macular hemorrhages and geographic atrophy. SLO uses single wavelength laser light to produce high contrast 2D en-face retinal images. The combination of SLO and OCT not only provides a more comprehensive imaging system but also corrects motion distortion of OCT images. FA is an invasive imaging technique that clearly illustrates vasculature through injection of fluorescein into the subject's blood. It is a gold standard imaging modality to depict neurovascular structure of retina and is used for diagnosing neurovascular related diseases. These fundus images are different from OCT image in imaging dimension, resolution, contrast and luminosity, which makes multimodality registration difficult. The common method of registration between 3D OCT image and 2D fundus image is to register the fundus image with the z-axis projection image of OCT [44, 92], or use SLO image which is collected simultaneously with OCT image as intermediate image for registration [42, 43, 93-95]. Most existing registration approaches utilize detection and extraction of features derived from retinal vasculature segmented separately from the multimodal images [48, 49]. Vascular features are the most commonly used structural features in multimodal image fusing since retinal vasculature is the most dominant structure of retina. After vascular skeletonization, vascular branching points can be easily detected and used as stable landmarks for correspondence detection. However, vessel segmentation errors may cause corresponding errors in the following registration process. Instead, other features such as the HOG feature [44], SURF and PIIFD descriptor [96] are used to capture the structural information in the images thus the segmentation of blood vessels is not required.

In an earlier study, retinal vasculature was extracted from color fundus and OCT projection image respectively. And then a registration was performed between vessel images using similarity function for the identification of neovascularization at the early stage of proliferative diabetic retinopathy [92]. This work could also be extended for microaneurysm detection at the earlier nonproliferative stage of diabetic retinopathy by registration of fundus and OCT images followed by classification algorithms. In another application, in order to analyze the asymmetry of retinal layers in right and left eyes, the OCT images form left and right eyes were registered by

alignment of retinal raphe. However, due to limited region size, the retinal raphe of macular OCTs was not calculable. Therefore, the alignment was achieved by aligning corresponding retinal raphe of fundus images and then registration of the OCTs to aligned fundus images [97]. Recent research reported accurate registration of FA and OCT images via SLO photographs, which can exactly detect the location and morphological appearance of diabetic retinopathy symptoms in OCT B-scans [94]. In their method, retinal vasculature segmentation was applied to both FA and SLO images. And then, FA vessel map was globally aligned with SLO vessel map photograph using SURF and FAST features together. Finally, a local non-rigid transformation was exploited to register two images perfectly. Combining OCT and OCTA image can not only improve the motion correction performance as introduced in section 3.1, but also greatly help the assessment of foveal avascular zone enlargement in different stages of diabetic retinopathy as reported in [98]. Moreover, several studies have shown that image fusing of multimodal data is beneficial for automated segmentation of optic disc and cup boundaries [99], retinal blood vessels [100] and intra-retinal surfaces [101].



(a) (b)

Fig. 6 Two examples of CNV growth prediction results. (a) OCT images with the CNV segmented in green; (b) Comparison between the clinical segmentation and the prediction.

3.4 Evaluation of longitudinal disease progression

Assessment of progression or regression of retinal disease can be obtained from temporal changes of longitudinal OCT images. The change of retinal thickness with time in OCT image dataset is an indicator of some retinal diseases. For example, retinal nerve fiber layer thickness measurement from OCT images has been widely used as an important clinical index to indicate glaucoma progression. Thinning of retinal layers is associated with some neurodegenerative disorders such as multiple sclerosis and Alzheimer's disease. In order to compare the changes of retinal morphology or function over time, it is necessary to register longitudinal OCT image data sets. Registration of OCT images for longitudinal analysis of retinal layer thickness measurement can be found in [102-108]. Another typical application is using the registration method to help the growth prediction of choroid neovascularization (CNV). As shown in Fig.6, CNV is a choroid lesion caused by new blood vessels growing in choroid. CNV related diseases are one of the important causes of visual disability. As so far, the pathogenesis is still not clear. A standard and effective treatment for CNV is to inject anti-vascular endothelial growth factor (anti-VEGF) agents into the eye to suppress further blood vessel growth. Such treatment requires frequent re-treatments. Therefore, in order to ensure the curative effect and reduce the risk, it is very important to predict the growth of CNV according to 3D longitudinal OCT images, so as to formulate a reasonable treatment plan. However, since longitudinal OCT images are collected at different time points, the retina displacements which are caused by different positions of eye during the scanning severely affect the accuracy of prediction. Therefore, it is necessary to register the longitudinal OCT images before prediction to guarantee the prediction accuracy and observe the change of lesion area in the same region [109].

4 Registration Performance

Due to the lack of open datasets and gold standard of OCT images, it is difficult to evaluate and directly compare the performance of OCT image registration. The definition of registration accuracy is non-unique and the methods of performance measurement reported in literatures are different. 1) For intra-modality registration, the commonly used quantitative performance metrics are dice similarity coefficient (DSC), mutual information, root mean squared error (RMSE), retinal surface distance and mean absolute retinal thickness difference obtained from layer segmentation; 2) For inter-modality registration, measuring the similarity of vascular images such as DSC of vessel binary maps, RMSE of manually selected vascular landmarks and the success rate of registration are commonly used to evaluate the registration performance. The success rate is the ratio of the image pairs with successful registration to the total number of image pairs. The successful registration is determined by the RMSE value of the matched points, and for clinical purposes, the RMSE of less than 5 is acceptable [110]. Table 1 shows the performances for some typical OCT registration algorithms. It should be noted that, in retinal OCT image registration, the registration accuracy of lesion region is more significant than other regions.

5 Discussion and Conclusion

Retinal OCT image registration is crucial for the diagnoses and treatments of various eye diseases. However, fast and accurate registration of retinal OCT images is still a challenging problem, in particular because of the low content contrast, large intensity variance as well as deterioration of unhealthy retina caused by various pathologies.

Representative retinal OCT image registration methods, which are coarsely divided into volumetric transformation-based registration methods and feature-based registration methods, have been systematically reviewed in this paper. Volumetric transformation-based registration methods use intensity information form the overlapped area of two images and seek to maximize the voxel similarity by measuring SAD, SSD, cross correlation or mutual information to obtain the best alignment. Feature-based registration methods calculate the transformation based on a number of anatomical correspondences established manually or automatically on a number of distinct anatomical features. Commonly used descriptors include the position of retinal landmarks such as the fovea center, the optic disc center and vascular bifurcation points. Moreover, high-level feature descriptors such as SIFT, SURF, Harris, HOG and so on are also used for feature-based registration. According to the corresponding relationship between the extracted feature descriptors, an objective function is used to find the best transformation parameters. Volumetric transformation-based registration methods are more simple and general than feature-based registration methods because preliminary extraction of specific anatomical features is not needed each time when applied to a new problem. However, the computation complexity is high for non-rigid registration of 3D OCT images since they should consider the similarity of each voxel in the 3D

Table 1 Overview of the performances for typical registration methods

Modalities	Reference	Method	Performance	Application
Intra-modality	Cheng *et al*. [26]	SAD	MSE：0.0415(normal eye), 0.0558(diseased eye) PSNR:23.48(normal eye),21.35(diseased eye) MSSI:0.769(normal eye),0.794(diseased eye)	Speckle Reduction
	Wang *et al.* [27]	SSD	PSNR:33dB; SSIM:0.92	Speckle Reduction
	Chen *et al.* [11]	SSD	DSC:0.77; Average layer boundary surface errors:8.0um	Average atlas and normalized space; Statistical atlas; RAVENS analysis of multiple sclerosis
	Zhang *et al.* [28]	CC	Improvement in SNR: $\sqrt{N}$; Improvement in CNR:11	speckle reduction
	Zheng *et al.* [29]	CC	Average unsigned positioning error:15.2 millimicron	Retinal Layer Segmentation
	Gong *et al*. [35]	SRWCR	HD:256.11um MHD:34.98um	Longitudinal study
	Wei *et al.* [36]	NMI	DSC:0.73	Longitudinal study
	Bogunović *et al.*[42]	SURF	Average unsigned positioning error:4.58±1.46um	Image Fusing and Splicing
	Pan *et al.*[12]	OCTRexpert	DSC: 0.92(normal eye), 0.78(diseased eye) AUSE:3.3um(normal eye),10.3um(diseased eye)	Longitudinal study
Inter-modality	Miri[44]	HOG	RMSE:0.055mm; Success rate:97.72%; Time:2.34s	Image Fusing and Splicing
	Cheng *et al*. [69]	LoSPA	RMSE:4.3; MSE:10.3; Success rate:14/18	Image Fusing and Splicing
	Almasi *et al*. [94]	SURF+FAST	RMSE:0.23 pixel; Success rate:97.23%; Time:26.1s	Image Fusing and Splicing
	Golkar *et al.* [95]	Global CPD + Intensity + Deformable	Target registration error: 0.09um; Success rate:92% (equal FOVs) and 89% (different FOVs)	Image Fusing and Splicing

MSE: mean square error
PSNR: peak signal to noise ratio
MSSI: mean structure similarity index
SSIM: structure similarity
CC: cross-correlation
SNR: signal-to-noise ratio
CNR: contrast-to noise ratio
NMI: normalized mutual information
SRWCR: spatially region-weighted correlation ratio
HD: Hausdorff distance
MHD: the maximum-likelihood Hausdorff distance
RANSAC : random sample consensus
DSC: Dice similarity coefficient
AUSE: average unsigned surface error

image. Volumetric transformation-based registration methods are sensitive to intensity change. Therefore, these methods may have poor registration performance for low quality retinal images or multimodal retinal images. Furthermore, considering each layer of retina has practically uniform image intensity, optimality criteria based on image similarity can suffer from local minima, which are caused by the ambiguity in defining correspondence. Feature-based registration methods rely on a relatively small number of significant features, and evaluate a matching criterion on features instead of on every single voxel in an image thus the computation complexity is much lower. The well selected features have less ambiguous than intensities in determining the correspondences. However, detecting stable features in pathological retina images and low quality images is difficult. Table 2 summarized the advantages and disadvantages of retinal OCT image registration methods. At present, there are no general automatic registration methods to deal with various clinical situations (different acquisition protocols, different modalities, different viewpoints, etc.). The choice of proper retinal OCT image registration method needs to consider both the registration computation complexity and accuracy requirement according to different clinical applications.

Table 2 Advantages and disadvantages of the methods for retinal OCT image registration

Volumetric transformation-based methods (rely on intensity similarity)	Pros	· simple and general
	Cons	· high computation complexity · sensitive to intensity change · local minima problem
Image feature-based methods (rely on distinct features)	Pros	· less ambiguous · low computation complexity
	Cons	· require extra feature extraction step · difficult to detect stable features in pathological retina images and low quality images

Compared with ordinary retinal image registration, the registration of pathological retina is more significant in clinic. However, for some lesions, such as choroidal neovascularization, registration is facing greater challenges. 1) The internal structures of diseased retinas are complex and

difficult to be recognized. In addition to the normal multiple retinal layers, there may be other additional structures such as neovascularization, fluid and so on; 2) Abnormalities lead to low contrast and blurred boundaries in OCT images between retinal layers. The degradation of image quality caused by abnormalities may affect the registration performance; 3) The layer structures of retina change dramatically in serious pathological regions which makes it difficult to extract stable features for registration. Therefore, new registration methods that can deal with retinas with abnormalities are needed for quantitative analysis of these diseases.

In recent years, deep learning method has achieved great success [111]. Many deep neural network models have been adopted successfully in various fields. Recent works also extended deep learning technique to solve complex medical image registration problem [112-114]. Krebs *et al.* proposed a deformable registration algorithm based on unsupervised learning of a low-dimensional probabilistic parameterization of deformations for cardiac MR sequences [115]. Islam *et al.* introduced a deep learning based registration framework for 3D multi-modal medical images of the head [116]. Yang *et al.* introduced a fast deformable image registration method called Quicksilver for 3D brain images [117]. They emphasized that their framework can be directly applied to many other registration techniques. For retinal registration, Abanovie *et al.* proposed DNN-based feature descriptor for fundus image registration and the learned feature descriptor was compared to other well-known descriptors using Fundus Image Registration dataset (FIRE) [118]. Silva *et al.* reported a deep learning based registration algorithm that aligns multi-modal retinal images (color fundus photography, fundus autofluorescence and infrared reflectance image modalities) to achieve accuracy and robustness required for large-scale longitudinal data structural changes analysis [119]. The training was performed end-to-end using the architecture shown in Fig. 7, where the combination of Siamese feature detectors, correlation matrix, and the regression layer directly output the 6-parameter affine transformation vector for a source and target image pair. Tian *et al.* presented an unsupervised learning method for deformable registration between color fundus image and OCT image [120]. At present, the research of deep learning in retinal registration is limited to 2D images. There is no publicly reported deep learning registration method for three dimensional OCT images. However, deep learning has great advantages in feature extraction. The deeply learned features can directly be combined with the traditional registration method to achieve better registration performance. It can be expected that in the near future, with the help of deep learning technology, more accurate 3D feature information can be obtained and the complex registration problem of abnormal 3D retinal OCT images can be solved.

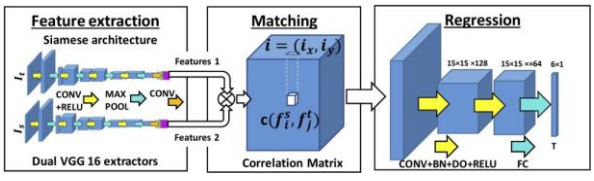


Fig.7 Feature based registration using deep-learning (DLAFFINE). CNN-based architectures for feature extraction, matching, and computing the registration transform. The network outputs 6-parameter affine transformation. CONV: convolution layer, BN: batch normalization, DO: drop off, RELU: rectified linear unit, FC: fully-connected layer. (Cited from [102].)

In the future, compared with rigid, mono-modality and 2D registration, retinal OCT image registration technology will pay more attention to non-rigid, multi-modality and 3D registration. For the eye movement artifact and the complex deformation caused by retinal lesions in OCT images, it is difficult to compensate these deformations with rigid registration. Therefore, non-rigid retinal OCT image registration which can describe more complex spatial relationship between images is preferred. Normally, 2D registration can only describe the spatial relationship between 2D planes. As a 3D image, 3D registration for OCT scans can provide more useful information than 2D registration. Although most of the previous studies focused on single-modal image registration, more and more attention has been paid to multi-modal image registration in recent years. From the clinical prospective, the combination of complementary information from different imaging modalities is not only helpful for doctors to diagnose and monitor ophthalmic diseases, but also advantageous for computer-aided image analysis and diagnosis technology. The non-rigid and poorly understood mechanics of the retina along with differences in imaging strategies, image geometries, and inter-patient variations will continue to challenge registration algorithms. It can be expected that there will be a growing demand for solutions as OCT imaging becomes more clinically relevant.

References

[1] J. Fujimoto, W. Drexler, J. Schuman, and C. Hitzenberger, "Optical Coherence Tomography (OCT) in Ophthalmology: Introduction," *Optics express*, vol. 17, pp. 3978-9, 2009.

[2] D. Kiernan, W. Mieler, and S. Hariprasad, "Spectral-Domain Optical Coherence Tomography: A Comparison of Modern High-Resolution Retinal Imaging Systems," *American journal of ophthalmology*, vol. 149, pp. 18-31, 2010.

[3] Y. Jia, S. Bailey, D. Wilson, O. Tan, M. Klein, C. Flaxel, *et al.*, "Quantitative Optical Coherence Tomography Angiography of Choroidal Neovascularization in Age-Related Macular Degeneration," *Ophthalmology*, vol. 121, 2014.

[4] A. Khanifar, A. Koreishi, J. Izatt, and C. Toth, "Drusen Ultrastructure Imaging with Spectral Domain Optical Coherence Tomography in Age-related Macular Degeneration," *Ophthalmology*, vol. 115, pp. 1883-90, 2008.

[5] E. Gao, B. Chen, J. Yang, F. Shi, W. Zhu, D. Xiang, *et al.*, "Comparison of Retinal Thickness Measurements between the Topcon Algorithm and a Graph-Based Algorithm in Normal and Glaucoma Eyes," *PLOS ONE*, vol. 10, p. e0128925, 2015.

[6] M. D. Abràmoff, M. K. Garvin, and M. Sonka, "Retinal Imaging and Image Analysis," *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 169-208, 2010.

[7] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable Medical Image Registration: A Survey," *IEEE Transactions on Medical Imaging*, vol. 32, pp. 1153-1190, 2013.

[8] Y. Ma, X. Chen, W. Zhu, X. Cheng, D. Xiang and F. Shi, "Speckle noise reduction in optical coherence tomography images based on edge-sensitive cGAN," *Biomedical Optics Express*, vol. 9, pp. 5129-5146, 2018.

[9] M. Szkulmowski, I. Gorczynska, D. Szlag, M. Sylwestrzak, A. Kowalczyk, and M. Wojtkowski, "Efficient reduction of speckle noise in Optical Coherence Tomography," *Optics express*, vol. 20, pp. 1337-59, 2012.

[10] A. Wong, A. Mishra, K. Bizheva, and D. Clausi, "General Bayesian estimation for speckle noise reduction in optical coherence tomography retinal imagery," *Optics express*, vol. 18, pp. 8338-52, 2010.

[11] 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, pp. 2196-2214, 2014.

[12] L. Pan, F. Shi, D. Xiang, K. Yu, L. Duan, J. Zheng, *et al.*, "OCTRexpert: A Feature-Based 3D Registration Method for Retinal OCT Images," *IEEE Transactions on Image Processing*, vol. 29, pp. 3885-3897, 2020.

[13] T. Makela, P. Clarysse, O. Sipila, N. Pauna, P. Quoc Cuong, T. Katila, *et al.*, "A review of cardiac image registration methods," *IEEE Transactions on Medical Imaging*, vol. 21, pp. 1011-1021, 2002.

[14] A. Khalil, S.-C. Ng, Y. Liew, and k. w. lai, "An Overview on Image Registration Techniques for Cardiac Diagnosis and Treatment," *Cardiology Research and Practice*, vol. 2018, pp. 1-15, 2018.

[15] M. A. Viergever, J. B. A. Maintz, S. Klein, K. Murphy, M. Staring, and J. P. W. Pluim, "A survey of medical image registration - under review," *Medical Image Analysis*, vol. 33, pp. 140-144, 2016.

[16] Y. Guo, R. Sivaramakrishna, C.-C. Lu, J. Suri, and S. Laxminarayan, "Breast image registration techniques: A survey," *Medical & biological engineering & computing*, vol. 44, pp. 15-26, 2006.

[17] G. Haskins, U. Kruger, and P. Yan, "Deep Learning in Medical Image Registration: A Survey," *Machine Vision and Applications*, vol. 31, 2020.

[18] R. Sivaramakrishna, "3D Breast Image Registration — A Review," *Technology in cancer research & treatment*, vol. 4, pp. 39-48, 2005.

[19] H. R. Boveiri, R. Khayami, R. Javidan, and A. Mehdizadeh, "Medical Image Registration Using Deep Neural Networks: A Comprehensive Review," *Computers & Electrical Engineering*, vol. 87, pp. 106767, 2020.

[20] A. Gholipour, N. Kehtarnavaz, R. Briggs, M. Devous, and K. Gopinath, "Brain Functional Localization: A Survey of Image Registration Techniques," *IEEE Transactions on Medical Imaging*, vol. 26, pp. 427-451, 2007.

[21] P. Pearlman, A. Adams, S. Elias, W. Mali, M. Viergever, and J. Pluim, "Mono- and multimodal registration of optical breast images," *Journal of biomedical optics*, vol. 17, pp. 80901-1, 2012.

[22] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transactions on Medical Imaging*, vol. 22, pp. 986-1004, 2003.

[23] V. R. s. Mani and A. Selvaraj, "Survey of medical image registration," *Journal of Biomedical Engineering and Technology*, vol. 1, pp. 8-25, 2013.

[24] S. Saha, D. Xiao, A. Bhuiyan, T. Y. Wong, and Y. Kanagasingam, "Color fundus image registration techniques and applications for automated analysis of diabetic retinopathy progression: A review," *Biomedical Signal Processing and Control*, vol. 47, pp. 288-302, 2019.

[25] A. Baghaie, Z. Yu, and R. M. D'Souza, "State-of-the-art in retinal optical coherence tomography image analysis," *Quantitative imaging in medicine and surgery*, vol. 5, pp. 603-617, 2015.

[26] J. Cheng, D. Tao, Y. Quan, D. Wong, C. M. G. Cheung, M. Akiba, *et al.*, "Speckle Reduction in 3D Optical Coherence Tomography of Retina by A-Scan Reconstruction," *IEEE Transactions on Medical Imaging*, vol. 35, pp. 1-1, 2016.

[27] X. Wang, X. Yu, X. Liu, S. Chen, S. Chen, N. Wang, *et al.*, "A two-step iteration mechanism for speckle reduction in optical coherence tomography," *Biomedical Signal Processing and Control*, vol. 43, pp. 86-95, 2018.

[28] 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, pp. 36013, 2015.

[29] Y. Zheng, R. Xiao, Y. Wang, and J. Gee, "A Generative Model for OCT Retinal Layer Segmentation by Integrating Graph-Based Multi-Surface Searching and Image Registration," presented at the International Conference on Medical Image Computing & Computer-assisted Intervention-MICCAI, Nagoya, Japan, 2013.

[30] B. B. Avants, C. Epstein, M. Grossman, and J. Gee, "Symmetric Diffeomorphic Image Registration with Cross-Correlation: Evaluating Automated Labeling of Elderly and Neurodegenerative Brain," *Medical image analysis*, vol. 12, pp. 26-41, 2008.

[31] S. Chen, X. Li, L. Zhao, and H. Yang, "Medium-low resolution multisource remote sensing image registration based on SIFT and robust regional mutual information," *International Journal of Remote Sensing*, vol. 39, pp. 3215-3242, 2018.

[32] S. Klein, U. Heide, I. Lips, M. Vulpen, M. Staring, and J. Pluim, "Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information," *Medical physics*, vol. 35, pp. 1407-17, 2008.

[33] D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, and P. Suetens, "Nonrigid Image Registration Using Conditional Mutual Information," *IEEE Transactions on Medical Imaging*, vol. 29, pp. 19-29, 2010.

[34] X. Zhuang, S. Arridge, D. J. Hawkes, and S. Ourselin, "A Nonrigid Registration Framework Using Spatially Encoded Mutual Information and Free-Form Deformations," *IEEE Transactions on Medical Imaging*, vol. 30, pp. 1819-1828, 2011.

[35] L. Gong, C. Zhang, L. Duan, X. Du, H. Liu, X. Chen, *et al.*, "Nonrigid Image Registration Using Spatially Region-Weighted Correlation Ratio and GPU-Acceleration," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, pp. 766-778, 2019.

[36] Q. Wei, F. Shi, W. Zhu, D. Xiang, H. Chen, and X. Chen, "Nonrigid registration of 3D longitudinal optical coherence tomography volumes with choroidal neovascularization," in *SPIE2*, USA, 2017, pp.101330X.1-101330x.8..

[37] Z. Nougrara and B. Nasr-eddine, "Retina registration for biometrics based on retinal feature points," *Modelling, Measurement and Control C*, vol. 79, pp. 242-246, 2018.

[38] H.M.Taha, E.-B. Nashwa. A.E.Hassanien,Y.Badr, V. Snasel, "Retinal Feature-Based Registration Schema," International Conference on Informatics Engineering and Information Science, ICIEIS 2011,vol. 252, pp.26-36, Berlin, Heidelberg: Springer, 2011.

[39] Z. Li, F. Huang, J. Zhang, B. Dashtbozorg, S. Abbasi-Sureshjani, Y. Sun, *et al.*, "Multi-modal and multi-vendor retina image registration," *Biomedical Optics Express*, vol. 9, 2018.

[40] M.Niemeijer, M.K.Garvin, K.Lee, *et al.*, "Registration of 3D spectral OCT volumes using 3D SIFT feature point matching," in *Proc. SPIE*, Orlando, USA, 2009, pp. 72591I- 72591I -8.

[41] J. Zheng, J. Tian, K. Deng, X. Dai, X. Zhang, and M. Xu, "Salient Feature Region: A New Method for Retinal Image Registration," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, pp. 221-232, 2011.

[42] H. Bogunović, M. Sonka, Y. H. Kwon, P. Kemp, M. D. Abràmoff, and X. Wu, "Multi-Surface and Multi-Field Co-Segmentation of 3-D Retinal Optical Coherence Tomography," *IEEE Transactions on Medical Imaging*, vol. 33, pp. 2242-2253, 2014.

[43] R. Kolar and P. Tasevsky, "Registration of 3D Retinal Optical Coherence Tomography Data and 2D Fundus Images, " *Lecture Notes in Computer Science*, vol.6204, pp. 72-82, 2010.

[44] M. S. Miri, M. Abramoff, Y. Kwon, and M. Garvin, "Multimodal registration of SD-OCT volumes and fundus photographs using histograms of oriented gradients," *Biomedical Optics Express*, vol. 7, pp. 5252-5267, 2016.

[45] M. Mokhtari, H. Rabbani, and A. Dehnavi, "Alignment of optic nerve head optical coherence tomography B-scans in right and left eyes," *2017 IEEE International Conference on Image Processing (ICIP)*, pp. 2279-2283, 2017.

[46] M. Esmaeili, M. Golabbakhsh. H. Rabbani, "Detection and registration of vessels of fundus and OCT images using curevelet analysis," the 13th IEEE International Conference on BioInformatics and BioEngineering, Los Alamitos, CA, USA, 2012.

[47] M. Golabbakhsh and H. Rabbani, "Vessel-based registration of fundus and optical coherence tomography projection images of retina using a quadratic registration model," *IET Image Processing*, vol. 7, pp. 768-776, 2013.

[48] Y. Li, G. Gregori, R. Knighton, B. Lujan, and P. Rosenfeld, "Registration of OCT fundus images with color fundus photographs based on blood vessel ridges," *Optics express*, vol. 19, pp. 7-16, 2011.

[49] S. Niu, Q. Chen, H. Shen, L. Sisternes, and D. Rubin, "Registration of SD-OCT en-face images with color fundus photographs based on local patch matching, " OMIA in MICCAI, 2014.

[50] S. Saleem and R. Sablatnig, "A Robust SIFT Descriptor for Multispectral Images," *IEEE Signal Processing Letters*, vol. 21, pp. 400-403, 2014.

[51] B. Béouche-Hédias, D. Helbert, C. Malézieu, N. Leveziel, and C. Fernandez-Maloigne, "Neovascularization Detection in Diabetic Retinopathy from Fluorescein Angiograms," *Journal of Medical Imaging*, vol. 4, pp.044503, 2017.

[52] S. H. Cheon, I. K. Eom, and Y. H. Moon, "Fast descriptor extraction method for a SURF-based interest point," *Electronics Letters*, vol. 52, pp. 274-275, 2016.

[53] Z. Yang, D. Shen, and P.-T. Yap, "Image mosaicking using SURF features of line segments," *PLOS ONE*, vol. 12, pp. e0173627, 2017.

[54] J. Liu and F. Bu, "Improved RANSAC features image-matching method based on SURF," *The Journal of Engineering*, vol. 2019, pp. 9118-9122, 2019.

[55] J. Chen, J. Tian, Lee, Noah, J. Zheng, Smith, *et al.*, "A Partial Intensity Invariant Feature Descriptor for Multimodal Retinal Image Registration," *IEEE Transactions on Biomedical Engineering*, vol. 57, pp. 1707-1718, 2010.

[56] A. Lang, A. Carass, O. Al-Louzi, P. Bhargava, H. Ying, P. Calabresi, *et al.*, "Longitudinal graph-based segmentation of macular OCT using fundus alignment," in *Proc. SPIE*, vol. 9413: 2015.

[57] 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, pp. 376-386, 2013.

[58] 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, pp. 1337-1345, 2012.

[59] M. K. Garvin, M. D. Abramoff, 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, pp. 1436-1447, 2009.

[60] J. Tian, B. Varga, G. M. Somfai, W.-H. Lee, W. E. Smiddy, and D. C. DeBuc, "Real-Time Automatic Segmentation of Optical Coherence Tomography Volume Data of the Macular Region," *PloS one*, vol. 10, pp. e0133908-e0133908, 2015.

[61] F. Shi, X. Chen, H. Zhao, W. Zhu, D. Xiang, E. Gao, *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, pp. 441-452, 2015.

[62] K. Yu, F. Shi, E. Gao, W. Zhu, H. Chen, and X. Chen, "Shared-hole graph search with adaptive constraints for 3D optic nerve head optical coherence tomography image segmentation," *Biomedical Optics Express*, vol. 9, pp. 962, 2018.

[63] D. Xiang, H. Tian, X. Yang, F. Shi, W. Zhu, H. Chen, *et al.*, "Automatic Segmentation of Retinal Layer in OCT Images With Choroidal Neovascularization," *IEEE Transactions on Image Processing*, vol. 27, pp. 5880-5891, 2018.

[64] X. Chen and L. Pan, "A Survey of Graph Cuts/Graph Search Based Medical Image Segmentation," *IEEE Reviews in Biomedical Engineering*, vol. 11, pp. 112-124, 2018.

[65] F. LaRocca, A.-H. Dhalla, M. Kelly, S. Farsiu, and J. Izatt, "Optimization of confocal scanning laser ophthalmoscope design," *Journal of Biomedical Optics*, vol. 18, pp. 076015, 2013.

[66] A. Baghaie, Z. Yu, and R. D'Souza, "Involuntary Eye Motion Correction in Retinal Optical Coherence Tomography: Hardware or Software Solution?," *Medical image analysis*, vol. 37, 2017.

[67] B. Antony, M. D. Abràmoff, L. Tang, W. D. Ramdas, J. R. Vingerling, N. M. Jansonius, *et al.*, "Automated 3-D method for the correction of axial artifacts in spectral-domain optical coherence tomography images," *Biomedical Optics Express*, vol. 2, pp. 2403-2416, 2011.

[68] A. Montuoro, J. Wu, S. Waldstein, B. Gerendas, G. Langs, C. Simader, *et al.*, "Motion Artefact Correction in Retinal Optical Coherence Tomography Using Local Symmetry," MICCA 2014,Cham, vol.8674, pp. 130-137, 2014,

[69] J. Cheng, J. Lee, G. Xu, Y. Quan, E. Ong, and D. Wong, "Motion Correction in Optical Coherence Tomography for Multi-modality Retinal Image Registration," in *International Conference on Medical Image Computing and Computer Assisted Interventions (MICCAI)*, Greece, 2016.

[70] M. F. Kraus, J. J. Liu, J. Schottenhamml, C.-L. Chen, A. Budai, L. Branchini, *et al.*, "Quantitative 3D-OCT motion correction with tilt and illumination correction, robust similarity measure and regularization," *Biomedical Optics Express*, vol. 5, pp. 2591-2613, 2014.

[71] Z. Li, V. P. Pandiyan, A. Maloney-Bertelli, X. Jiang, X. Li, and R. Sabesan, "Correcting intra-volume distortion for AO-OCT using 3D correlation based registration," *Optics Express*, vol. 28, pp. 38390-38409, 2020.

[72] S. Makita, M. Miura, S. Azuma, T. Mino, T. Yamaguchi, and Y. Yasuno, "Accurately motion-corrected Lissajous OCT with multi-type image registration," *Biomedical Optics Express*, vol. 12, pp. 637-653, 2021.

[73] D. Xiang, G. Chen, F. Shi, W. Zhu, Q. Liu, S. Yuan, *et al.*, "Automatic Retinal Layer Segmentation of OCT Images With Central Serous Retinopathy," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, pp. 283-295, 2019.

[74] P. Zang, G. Liu, M. Zhang, C. Dongye, J. Wang, A. D. Pechauer, *et al.*, "Automated motion correction using parallel-strip registration for wide-field en face OCT angiogram," *Biomedical Optics Express*, vol. 7, pp. 2823-2836, 2016.

[75] M. Heisler, S. Lee, Z. Mammo, Y. Jian, M. J. Ju, A. Merkur, *et al.*, "Strip-based registration of serially acquired optical coherence tomography angiography," *Journal of Biomedical Optics*, vol. 22, pp. 036007, 2017.

[76] Y. Chen, Y.-J. Hong, S. Makita, and Y. Yasuno, "Eye-motion-corrected optical coherence tomography angiography using Lissajous scanning," *Biomedical Optics Express*, vol. 9, pp. 1111-1129, 2018.

[77] A. Athwal, C. Balaratnasingam, D.-Y. Yu, M. Heisler, M. V. Sarunic, and M. J. Ju, "Optimizing 3D retinal vasculature imaging in diabetic retinopathy using registration and averaging of OCT-A," *Biomedical Optics Express*, vol. 12, pp. 553-570, 2021.

[78] S. B. Ploner, M. F. Kraus, E. M. Moult, L. Husvogt, J. Schottenhamml, A. Yasin Alibhai, *et al.*, "Efficient and high accuracy 3-D OCT angiography motion correction in pathology," *Biomedical Optics Express*, vol. 12, pp. 125-146, 2021.

[79] P. Zang, G. Liu, M. Zhang, J. Wang, T. Hwang, D. Wilson, *et al.*, "Automated three-dimensional registration and volume rebuilding for wide-field angiographic and structural optical coherence tomography," *Journal of Biomedical Optics*, vol. 22, pp. 026001, 2017.

[80] J. Zhao, Y. Winetraub, E. Yuan, W. H. Chan, S. Z. Aasi, K. Y. Sarin, *et al.*, "Angular compounding for speckle reduction in optical coherence tomography using geometric image registration algorithm and digital focusing," *Scientific Reports*, vol. 10, pp. 1893, 2020.

[81] D. Alonso-Caneiro, S. Read, and M. Collins, "Speckle reduction in optical coherence tomography imaging by affine-motion image registration," *Journal of biomedical optics*, vol. 16, pp. 116027, 2011.

[82] A. Baghaie, R. M. D'Souza, and Z. Yu, "Application of Independent Component Analysis techniques in speckle noise reduction of retinal OCT images," *Optik*, vol. 127, pp. 5783-5791, 2016.

[83] A. Baghaie, R. M. D'Souza, and Z. Yu, "Sparse And Low Rank Decomposition Based Batch Image Alignment for Speckle Reduction of retinal OCT Images," *arXiv:1411.4033*, 2015.

[84] H. Cai, X. Han. S. Lou, Y. Wang and X. Chen, "Speckle noise reduction in swept-source optical coherence tomography by retinal image registration," *Chinese Optics*, vol. 13, pp. 1-10, 2020.

[85] H. Zhang, Z. Li, X. Wang, N. Nan, and X. Wang "A novel method for speckle reduction in optical coherence tomography by image registration," Proc. SPIE 9230, Twelfth International Conference on Photonics and Imaging in Biology and Medicine (PIBM 2014), https://doi.org/10.1117/12.2068846.

[86] M. Liu, X. Chen, and B. Wang, "Axial and horizontal registration guided speckle suppression in single-line HD mode for retinal optical coherence tomography images," *Optics Communications*,vol.487, pp.126807, 2021.

[87] Y. Ma, X. Chen, W. Zhu, *et al.*, "Speckle noise reduction in optical coherence tomography images based on edge-sensitive cGAN," *Biomedical Optics Express*, vol. 9, pp. 5129-5146, 2018.

[88] F. Shi, N. Cai, Y. Gu, D. Hu, Y. Ma, Y. Chen, *et al.*, "DeSpecNet: a CNN-based method for speckle reduction in retinal optical coherence tomography images," *Physics in Medicine and Biology*, vol. 64, pp. 175010, 2019.

[89] M. Wang, W. Zhu, K. Yu, Z. Chen, and X. Chen, "Semi-Supervised Capsule cGAN for Speckle Noise Reduction in Retinal OCT Images," *IEEE Transactions on Medical Imaging*, vol. 99, pp. 1-1, 2021.

[90] Y. Li, G. Gregori, B. Lam, and P. Rosenfeld, "Automatic montage of SD-OCT data sets," *Optics express*, vol. 19, pp. 26239-48, 2011.

[91] J. Wang, A. Camino, X. Hua, L. Liu, D. Huang, T. S. Hwang, *et al.*, "Invariant features-based automated registration and montage for wide-field OCT angiography," *Biomedical Optics Express*, vol. 10, pp. 120-136, 2019.

[92] N. Padmasini and R. Umamaheswari, "Detection of neovascularisation using K-means clustering through registration of peripapillary OCT and fundus retinal images," in *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, 2016, pp. 1-4.

[93] Z. G. Kamasi, M. Mokhtari , and H. Rabbani . "Non-rigid registration of Fluorescein Angiography and Optical Coherence Tomography via scanning laser ophthalmoscope imaging. " IEEE Engineering in Medicine and Biology Conference (EMBC2017) IEEE, 2017.

[94] R. Almasi, A. Vafaei, Z. Ghasemi, M. R. Ommani, A. R. Dehghani, and H. Rabbani, "Registration of fluorescein angiography and optical coherence tomography images of curved retina via scanning laser ophthalmoscopy photographs," *Biomedical Optics Express*, vol. 11, pp. 3455-3476, 2020.

[95] E. Golkar, H. Rabbani, and A. Dehghani, "Hybrid Registration of Retinal Fluorescein Angiography and Optical Coherence Tomography Images of Patients

with Diabetic Retinopathy," *Biomedical Optics Express*, vol. 12, 2021.

[96] H. Tang, A. Pan, K. Yang, *et al.*, "Retinal Image Registration Based on Robust Non-Rigid Point Matching Method," *Journal of Medical Imaging & Health Informatics*, vol. 8, pp. 240-249,2018.

[97] T. Mahmoudi, R. Kafieh, H. Rabbani, A. Dehnavi, and M.-R. Akhlaghi, "Evaluation of Asymmetry in Right and Left Eyes of Normal Individuals Using Extracted Features from Optical Coherence Tomography and Fundus Images," *Journal of Medical Signals & Sensors*, vol. 11, pp. 12, 2021.

[98] G. Lynch, J. S. A. Romo, R. Linderman, B. D. Krawitz, S. Mo, A. Zakik, *et al.*, "Within-subject assessment of foveal avascular zone enlargement in different stages of diabetic retinopathy using en face OCT reflectance and OCT angiography," *Biomedical Optics Express*, vol. 9, pp. 5982-5996, 2018.

[99] M. S. Miri, M. D. Abràmoff, K. Lee, M. Niemeijer, J. Wang, Y. H. Kwon, *et al.*, "Multimodal Segmentation of Optic Disc and Cup From SD-OCT and Color Fundus Photographs Using a Machine-Learning Graph-Based Approach," *IEEE Transactions on Medical Imaging*, vol. 34, pp. 1854-1866, 2015.

[100] Z. Hu, M. Niemeijer, M. D. Abramoff, and M. K. Garvin, "Multimodal Retinal Vessel Segmentation From Spectral-Domain Optical Coherence Tomography and Fundus Photography," *IEEE Transactions on Medical Imaging*, vol. 31, pp. 1900-1911, 2012.

[101] M. S. Miri, V. A. Robles, M. D. Abràmoff, Y. H. Kwon, and M. K. Garvin, "Incorporation of gradient vector flow field in a multimodal graph-theoretic approach for segmenting the internal limiting membrane from glaucomatous optic nerve head-centered SD-OCT volumes," *Computerized Medical Imaging and Graphics*, vol. 55, pp. 87-94, 2017.

[102] S. Lee, M. Heisler, P. J. Mackenzie, M. V. Sarunic, and M. F. Beg, "Quantifying Variability in Longitudinal Peripapillary RNFL and Choroidal Layer Thickness Using Surface Based Registration of OCT Images," *Translational vision science & technology*, vol. 6, pp. 11-11, 2017.

[103] L. Balk, M. Mayer, B. M. J. Uitdehaag, and A. Petzold, "Physiological variation of segmented OCT retinal layer thicknesses is short-lasting," *Journal of Neurology*, vol. 260, pp. 3109-3114, 2013.

[104] K. Togashi, *et al.*, "Registration Algorithm of Fundus Oculi for 3-D Optical Coherence Tomography Images," *Ieice Technical Report*, vol. 113, pp. 133-137, 2014.

[105] A. Lang, A. Carass, O. Al-Louzi, P. Bhargava, S. Solomon, P. Calabresi, *et al.*, "Combined registration and motion correction of longitudinal retinal OCT data ," in *Proc. SPIE*, vol. 9784, 2016.

[106] K. Lee, M. Abràmoff, M. Sonka, and M. Garvin, "Automated segmentation of intraretinal layers from spectral-domain macular OCT: reproducibility of layer thickness measurements," in *Proc. SPIE*, vol. 7965, 2011.

[107] S. Lee, M. L. Heisler, K. Popuri, N. Charon, B. Charlier, A. Trouvé, *et al.*, "Age and Glaucoma-Related Characteristics in Retinal Nerve Fiber Layer and Choroid: Localized Morphometrics and Visualization Using Functional Shapes Registration," *Frontiers in neuroscience*, vol. 11, pp. 381-381, 2017.

[108] S. Lee, M. Heisler, P. J. Mackenzie, M. V. Sarunic, and M. F. Beg, "Quantifying Variability in Longitudinal Peripapillary RNFL and Choroidal Layer Thickness Using Surface Based Registration of OCT Images," *Translational Vision Science & Technology*, vol. 6, 2017.

[109] S. Zhu, F. Shi, D. Xiang, W. Zhu, H. Chen, and X. Chen, "Choroid Neovascularization Growth Prediction With Treatment Based on Reaction-Diffusion Model in 3-D OCT Images," *Biomedical and Health Informatics, IEEE Journal of*, vol. 21, pp. 1667-1674, 2017.

[110] Z. Ghassabi, J. Shanbehzadeh, A. Mohammadzadeh, and S. S. Ostadzadeh, "Colour retinal fundus image registration by selecting stable extremum points in the scale-invariant feature transform detector," *IET Image Processing*, vol. 9, pp.889-900,2015.

[111] X. Hao, G. Zhang, and S. Ma, "Deep Learning," *International Journal of Semantic Computing*, vol. 10, pp. 417-439, 2016.

[112] G. Haskins, U. Kruger, and P. Yan, "Deep learning in medical image registration: a survey," *Machine Vision and Applications*, vol. 31, pp. 8, 2020.

[113] V. Villena Martinez, S. Oprea, M. Saval-Calvo, J. Azorin-Lopez, A. Guillóo, and R. Fisher, "When Deep Learning Meets Data Alignment: A Review on Deep Registration Networks (DRNs)," *Applied Sciences*, vol. 10, pp. 7524, 2020.

[114] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, and C. I. Sánchez, "A Survey on Deep Learning in Medical Image Analysis," *Medical Image Analysis*, vol. 42, pp. 60-88, 2017.

[115] J. Krebs, T. Mansi, B. Mailhé, N. Ayache, and H. Delingette, "Unsupervised Probabilistic Deformation Modeling for Robust Diffeomorphic Registration," arXiv:1804.07172, pp. 101-109, 2018.

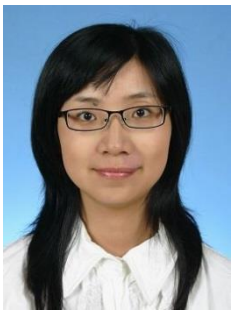
[116] K. T. Islam, S. Wijewickrema, and S. O'Leary, "A deep learning based framework for the registration of three dimensional multi-modal medical images of the head," *Scientific Reports*, vol. 11, pp. 1860, 2021.

[117] X. Yang, R. Kwitt, M. Styner, and M. Niethammer, "Quicksilver: Fast Predictive Image Registration - a Deep Learning Approach," *NeuroImage*, vol. 158, pp.378-396, 2017.

[118] E. abanovi è, G. Stankevi èius, and D. Matuzevi èius, "Deep Neural Network-based Feature Descriptor for Retinal Image Registration," in *2018 IEEE 6th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE)*, 2018, pp. 1-4.

[119] T. D. Silva, E. Y. Chew, N. Hotaling, and C. A. Cukras, "Deep-Learning based Multi-Modal Retinal Image Registration for Longitudinal Analysis of Patients with Age-related Macular Degeneration," *Biomedical Optics Express*, vol. 12, pp.619-636,2021.

[120] Y. Tian, Y. Hu, Y. Ma, H. Hao, L. Mou, J. Yang, *et al.*, "Multi-scale U-net with Edge Guidance for Multimodal Retinal Image Deformable Registration ," *42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada, 2020, pp. 1360-1363, doi: 10.1109/EMBC44109.2020.9175613.



Lingjiao Pan, received her Ph.D. degree from the School of Electrical and Information Engineering, Soochow University in 2020. She worked as a R&D engineer at LG Electronics (2007-2008) and Electronics and Telecommunications Research Institute (2008-2011) in Korea. Since 2012, she has worked as Assistant Professor at the School of Electrical and Information Engineering, Jiangsu University of Technology. Her research interests include medical image processing and analysis and data compression.

Xinjian Chen, received his Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences in 2006. He is a Distinguished Professor of the School of Electrical and Information Engineering, Soochow University and Director of Medical Image Processing, Analysis and Visualization Laboratory. His research interests include medical image processing and analysis, pattern recognition and machine learning. He has published more than 150 papers in prestigious journals and conferences and has applied for more than 70 Chinese invention patents, and more than 30 have been granted. He has served as Associate Editor for IEEE Transactions on Medical Imaging, and IEEE Journal of Translational Engineering in Health and Medicine.